М.І. Жалдак,

Н.М. Кузьміна,

Г.О. Михалін

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Підручник для студентів
фізико-математичних та інформатичних спеціальностей
педагогічних університетів
Видання четверте, доповнене

Київ. НПУ імені М.П. Драгоманова

2020

УДК 519.21 /.25
Ж24

2-ге видання затверджене Міністерством освіти і науки України (Лист від 03.11.08 №1.4/18-Г-2303)
3-тє видання схвалено рішенням Вченої Ради Національного педагогічного університету імені М.П. Драгоманова (протокол № 7 від 24 лютого 2015 року)
4-те видання схвалено рішенням Вченої Ради Національного педагогічного університету імені М.П. Драгоманова (протокол №5 від 31січня 2019)

Рецензенти:
Доктор педагогічних наук, професор Горошко Ю.В.
Доктор фіз.-мат. наук, професор Торбін Г.М.
Доктор фіз.-мат. наук, професор Самусенко П.Ф.

М.І. Жалдак, Н.М. Кузьміна, Г.О. Михалін
Ж24 Теорія ймовірностей і математична статистика: Підручник для студентів фізико-математичних та інформатичних спеціальностей педагогічних університетів. Видання четверте, доповнене / М.І. Жалдак, Н.М. Кузьміна, Г.О. Михалін. – Київ. НПУ імені М.П. Драгоманова, 2020 – 750 с.

ISBN 978-966-931-221-1

Подано основи теорії ймовірностей та математичної статистики. Матеріал перших розділів цілком може бути використаний вчителями в процесі навчання елементів стохастики в школі. Для побудови графічних зображень, обчислення значень виразів, визначених інтегралів, аналізу статистичних даних, визначення числових характеристик розподілів ймовірностей, в тому числі статистичних, передбачається використання відповідних комп'ютерних програм, зокрема Gran1.

Призначено для студентів фізико-математичних та інформатичних спеціальностей педагогічних університетів. Може стати в нагоді також студентам технічних, економічних та інших спеціальностей, вчителям математики та учням старших класів середніх навчальних закладів.

УДК 519.21 /.25
Ж24

ISBN 978-966-931-221-1

ПЕРЕДМОВА

Елементи стохастики – порівняно новий розділ в шкільному курсі математики, який в навчальних планах подано у вигляді двох частин із назвами «Елементи теорії ймовірностей» та «Елементи математичної статистики». Разом з тим існують тісні міжпредметні зв'язки основних понять стохастики з відповідними поняттями геометрії, алгебри і початків аналізу, фізики і ін., оскільки поняття ймовірності тісно пов'язане з поняттям міри множини деяких елементів (точок). Як відомо *мірою множин*, заданою на деякій їх сукупності S, називається функція $V(A)$ від цих множин $A \in S$, стосовно якої задовільняються такі основні вимоги:

1_V. $V(A) \geq 0$, $A \in S$, тобто міра множин із їх сукупності S має бути невід'ємна. Цю властивість називають властивістю невід'ємності міри V.

2_V. Якщо $A_i \in S$, $i = 1, 2, \ldots$, і $A_i A_j = \varnothing$, коли $i \neq j$, тобто множини A_i і A_j не перетинаються (не існує точок, які належать відразу до двох різних множин), і $\bigcup_i A_i \in S$, то

$$V\left(\bigcup_i A_i\right) = \sum_i V(A_i),$$

де $\bigcup_i A_i$ – об'єднання довільної скінченної або зчисленної кількості множин із сукупності S.

Цю властивість називають властивістю повної адитивності міри V.

Прикладами мір можуть бути довжини лінійних вимірних фігур, зокрема відрізків та їх об'єднань, площі плоских квадровних фігур, зокрема квадратів та їх об'єднань, об'єми просторових кубовних фігур, зокрема кубів та їх об'єднань, маси тіл та їх об'єднань, кількості успішних спроб виконати спортивну вправу, кількості влучень в окремі частини мішені та в їх об'єднання, кількості елементів в множинах та в їх об'єднаннях і т.д.

Нагадаємо, що означення площі плоскої фігури чи об'єму просторового тіла за підручниками із шкільної геометрії О.В. Погорєлова, означення маси тіла тощо за суттю співпадають з наведеним вище означенням міри $V(A)$ множин A із сукупності множин S, $A \in S$. Тут сукупність S є областю задання функції V – аргументами функції V є елементи сукупності S, які є деякими множинами. Це дає можливості використання геометричних та фізичних аналогій під час вивчення властивостей ймовірнісної міри, з'ясування внутріпредметних та міжпредметних зв'язків стохастики, геометрії, алгебри і початків аналізу, фізики та ін.

Якщо сукупність S, на якій задана міра $V(A)$, утворена із підмножин деякої множини Ω такої, що $0 < V(\Omega) < \infty$, і крім того задовільняються вимоги

1_S. $\Omega \in S$,

2_S. Якщо $A \in S$, то і $\overline{A} = \Omega \setminus A \in S$,

3_S. Якщо $A_i \in S$, то і $\bigcup_i A_i \in S$,

тоді функція $P(A) = \dfrac{V(A)}{V(\Omega)}$, $A \in S$, також буде мірою, заданою на сукупності S підмножин множини Ω. В такому разі говорять, що міра $P(A)$ утворена нормуванням міри $V(A)$, $A \in S$.

Очевидно, стосовно цієї міри $P(A)$, $A \in S$, задовільняються вимоги:

1_P. $P(A) \geq 0$, $A \in S$.

2_P. Якщо $A_i \in S$, $i = 1,2,...$, і $A_i A_j = \varnothing$, коли $i \neq j$, то $P(\bigcup_i A_i) = \sum_i P(A_i)$.

3_P. $P(\Omega) = 1$.

Якщо проводяться випробування, які полягають в тому, що із множини Ω в кожному випробуванні навмання вибирають одну точку, то можна вважати, що чим більше (ближче до 1) відношення $P(A) = \dfrac{V(A)}{V(\Omega)}$ міри множини A до міри множини Ω, тим вірогідніше, що в такому експерименті буде вибрано точку із множини A. Зокрема, якщо $A = \Omega$, тоді навмання вибрана точка обов'язково належатиме до множини A і крім того буде $P(A) = 1$.

Будь яку функцію $P(A)$, *стосовно якої задовільняються вимоги* 1_p*-*3_p *і яка задана на сукупності S підмножин множини* Ω, *стосовно якої задовільняються вимоги* 1_s*-*3_s , *називають ймовірнісною мірою (або просто ймовірністю), заданою на сукупності S підмножин множини* Ω. Самі підмножини A множини Ω, які входять до сукупності S, називають подіями (або випадковими подіями). Слід зауважити, що відбування чи невідбування події $A \in S$, $A \subset \Omega$, в результаті випробування пов'язується з тим елементом $E \in \Omega$, який вибирається навмання із множини Ω. Якщо виявляється, що $E \in A$, то говорять, що в результаті даного випробування подія A відбулася. Якщо ж $E \overline{\in} A$, тоді відбувається подія $\overline{A}$, протилежна до A (елемент E не належить до A, однак завжди $E \in \Omega$).

В теорії ймовірностей і вивчаються властивості ймовірнісних мір, заданих на сукупностях S підмножин деяких множин Ω, причому стосовно таких сукупностей повинні задовільнятися вимоги 1_s-3_s.

Вимоги 1_s-3_s, 1_p-3_p називають системою аксіом теорії ймовірностей (системою аксіом А.М. Колмогорова).

Слід зауважити, що для дослідження властивостей ймовірнісних мір не має ніякого значення, як задані чи одержані такі міри – в результаті певних припущень чи домовленостей, в результаті проведення певних досліджень чи ще якось. І для абстрактних ймовірнісних мір, і для одержаних в результаті реальних випробувань, властивості таких мір, що випливають з аксіом 1_s-3_s, 1_P-3_P, залишаються однаковими.

Наприклад, часто вважають, що ймовірність випадання герба в результаті підкидання реальної монети дорівнює 1/2, що в наступних досить довгих (чи не дуже довгих) серіях підкидань монети співвідношення кількості випадань герба до кількості проведених випробувань буде близьким до 1/2, хоч цього неможливо довести ніякими логічними міркуваннями. Те саме стосується будь яких інших випробувань з випадковими непередбачуваними наслідками – підкидання реального кубика, яким би правильним і однорідним не був кубик, вибору навмання кульки під час проведення реальних тиражів спортлото і т.д.

Разом з тим, як зауважував ще в 1693 році видатний швейцарський математик Якоб Бернуллі (1654-1705 р.р.): «*Що не дано нам вивести apriori* (тобто наперед передбачити до проведення досліду), *те принаймні можна дістати aposteriori*, тобто з численних спостережень результатів подібних дослідів. Тому можна передбачати, що деяке явище згодом може відбутися у стількох саме випадках, у скількох воно раніше було відмічене як таке, що відбулося за подібних умов …». Тобто швидше за все і в майбутньому перебіг випадкових процесів чи частоти відбування випадкових подій мало відрізнятимуться від того, як це було в минулому.

Однак, вивчаючи попередній досвід, прогнозувати результати можна лише з тією або іншою мірою точності і впевненості, не даючи разом з тим жодних запевнень і гарантій. В зв’язку з цим слід підкреслити, що ймовірність будь якої випадкової події (окрім неможливої $\varnothing$ і вірогідної Ω) гіпотетична, оскільки ніяк неможливо довести, що ймовірність події слід вважати саме такою, а не інакшою, якщо попередньо не робити ніяких припущень (гіпотез).

Зауважимо, що множина Ω не обов’язково повинна бути множиною точок в якомусь координатному просторі, а міра множини Ω (наприклад, кількість елементів) не обов’язково має бути пов’язана з якимись віддалями між точками, площами фігур, об’ємами тіл і т.п.

Оскільки статистична ймовірність (відносна частота) відбування події A в серії із n вже проведених випробувань визначається за рівністю

$$P_n^*(A) = \frac{k_n(A)}{k_n(\Omega)},\ 0 < k_n(\Omega) < \infty,\ A \in S,$$

де $k_n(A)$ – кількість відбувань події A в серії із n проведених випробувань, то $P_n^*(A)$ є нормованою мірою на S, одержаною нормуванням міри $k_n(A)$, $A \in S$, тобто $P_n^*(A)$ є *ймовірнісною мірою*. Тому в процесі вивчення статистичних ймовірностей $P_n^*(A)$, $A \in S$, заодно вивчатимуться і теоретичні основи стохастики, і елементи математичної статистики, оскільки $P_n^*(A)$, $A \in S$, одержуються на основі відповідного статистичного матеріалу за результатами серії вже проведених випробувань.

Якщо перейти від ймовірнісної міри $P_n^*(A)$ до довільної (можливо гіпотетичної) ймовірнісної міри $P(A)$, і у всіх твердженнях, що випливають з основних властивостей 1_P-3_P ймовірнісної міри $P_n^*(A)$, замінити позначення $P_n^*(A)$ на $P(A)$, а слова «статистична ймовірність» словом «ймовірність», то всі такі твердження стосовно ймовірнісної міри $P_n^*(A)$ залишаться правильними і стосовно $P(A)$.

Разом з тим в разі реальної випадкової події статистичну ймовірність $P_n^*(A)$ можна отримати в результаті реального проведення серії із n випробувань, і цей спосіб задання ймовірнісної міри є найдоступнішим для початкового ознайомлення з поняттям ймовірнісної міри серед всіх можливих способів задання такої міри. В зв'язку з цим швидше за все твердження, що стосуються статистичних ймовірностей $P_n^*(A)$, будуть доступніші для розуміння, аніж твердження, що стосуються гіпотетичної ймовірності $P(A)$ з довільним неконкретним способом задання.

Якщо задано підмножини H_i множини Ω такі, що $H_i \in S$, $H_1 + H_2 + ... + H_k = \Omega$, $H_i H_j = \varnothing$, коли $i \neq j$, і знайдено $P_n^*(H_i)$, $i \in \overline{1,k}$, такі, що $P_n^*(\bigcup_{i=1}^{k} H_i) = P_n^*(\Omega) = \sum_{i=1}^{k} P_n^*(H_i) = 1$, тоді говорять, що статистична ймовірність розподілена на множині Ω за підмножинами H_i так, що на підмножину H_i припадає статистична ймовірність $P_n^*(H_i)$, $i \in \overline{1,k}$. Тут доцільно провести порівняння (аналогію) з розподілом одиничної маси на множині Ω за підмножинами H_i таким, що на підмножину H_i припадає маса $m(H_i)$, $\sum_{i=1}^{k} m(H_i) = m(\Omega) = 1$.

Нехай окрім ймовірнісної міри $P_n^*(H_i)$, $i \in \overline{1,k}$, на S задано ще якусь міру V, таку, що $V(H_i) > 0$, $i \in \overline{1,k}$, (наприклад, кількості елементів в множинах H_i, довжини, площі, об'єми і ін.), і $\sum_{i=1}^{k} V(H_i) = V(\Omega)$. Тоді після проведення серії із n випробувань, в яких одержані числа $k_n(H_i)$, природно говорити, що спостережені елементарні події в множині H_i розміщені щільніше (відносно міри V), ніж в множині H_j, якщо $\frac{k_n(H_i)}{V(H_i)} > \frac{k_n(H_j)}{V(H_j)}$, $i \neq j$, бо у множині H_i на одиницю міри V в середньому припадає більше спостережених елементарних подій, аніж у множині H_j.

Відношення

$$\frac{k_n(H_i)/n}{V(H_i)} = \frac{P_n^*(H_i)}{V(H_i)},\ i \in 1,2,...,k,$$

природно назвати *середньою щільністю статистичної ймовірності* $P_n^*(H_i)$ *на множині* H_i *відносно міри* V.

Функцію

$$f_n^*(E) = \frac{P_n^*(H_i)}{V(H_i)},\ E \in H_i,\ i \in \overline{1,k},$$

яка на множинах H_i набуває сталих значень, називають *усередненою щільністю розподілу статистичних ймовірностей* на множині Ω за множинами H_i, $i \in \overline{1,k}$, відносно міри V.

В такому разі якщо $A = \bigcup_{i \in I} H_i$, $I \subset \{1,2,...,k\}$, то

$$P_n^*(A) = P_n^*(\bigcup_{i \in I} H_i)\ \vdots.$$

Останню суму природно назвати інтегралом від функції $f_n^*(E)$ на множині A за мірою V, який позначають $\int_A f_n^*(E) V(dE)$.

Очевидно, стосовно так введеної щільності розподілу статистичних ймовірностей задовільняються такі основні вимоги:

1. $f_n^*(E) \geq 0$, $E \in \Omega$,
2. $\int_\Omega f_n^*(E) V(dE) = 1$.

Разом з тим про функцію $F_n^*(x)$ розподілу статистичних ймовірностей на множині Ω, центр розподілу статистичних ймовірностей на множині Ω, величину розсіювання (чи скупченості) статистичних ймовірностей навколо центра розподілу (дисперсію розподілу статистичних ймовірностей) можна говорити лише тоді, коли елементи множини Ω є точками в деякому координатному просторі.

Після введення поняття випадкової величини *X* всі міркування стосовно розподілів статистичних ймовірностей на множині Ω_X значень випадкової величини *X* та їх числових характеристик без будь яких змін залишаються такими самими, як і міркування стосовно розподілів статистичних ймовірностей на множині Ω.

Матеріал посібника супроводжується достатньою кількістю прикладів, рисунків, запитань для самоконтролю, вправ для самостійного виконання.

Для обчислення значень функцій, інтегралів, побудови графіків функцій, гістограм тощо передбачається використання комп'ютера та в основному педагогічного програмного засобу Gran1.

Посібник призначений для студентів фізико-математичних та інформатичних спеціальностей педагогічних університетів. Разом з тим він може стати в нагоді вчителям математики, студентам технічних, економічних та інших спеціальностей, учням старших класів середніх навчальних закладів.

РОЗДІЛ І. ВИПАДКОВІ ПОДІЇ. ЙМОВІРНІСНІ МІРИ

1.1. Стохастичні випробування. Простір елементарних подій

Для дослідження всеможливих процесів і явищ люди часто вдаються до спеціальним чином організованих спостережень чи випробувань. Результати таких випробувань іноді є непередбачуваними. *Випробування,* точні результати яких передбачити неможливо, називають *стохастичними*. Разом з тим кожному стохастичному випробуванню відповідає певна *множина* Ω *його можливих наслідків*. Ця множина Ω називається *множиною* або *простором елементарних подій*, а її елементи називаються *елементарними подіями*. Слід підкреслити, що результатом проведення випробування є один єдиний наслідок – відбувається одна єдина елементарна подія E із множини Ω всіх елементарних подій. Іншими словами, *в результаті кожного випробування із множини* Ω *немов би вибирається один єдиний елемент E – відбувається елементарна подія E.*

Приклад 1.1.1. Грані кубика пофарбовано різними кольорами: одна грань біла, ще одна грань червона, інші чотири грані зелені. Кубик підкидають і фіксують колір грані, якою кубик впаде догори. Тоді множина $\Omega = \{$”біла”, “червона”, “зелена”$\}$ є множиною можливих наслідків випробування, тобто простором із трьох елементарних подій.

Приклад 1.1.2. Всі грані кубика пофарбовано в білий колір. Кубик підкидають і фіксують колір грані, якою кубик впаде догори. Тоді в множині $\Omega = \{$”біла”$\}$ є один єдиний елемент, а тому в даному випробуванні можливий один єдиний наслідок – верхня грань біла.

Якщо грані кубика пофарбовані в два кольори – тоді наслідків у випробуванні наведеного типу буде два, якщо в три кольори – наслідків буде три, і т.д., якщо шість кольорів – наслідків буде шість (тут мається на увазі, що на кожній грані є лише один колір).

Приклад 1.1.3. Підкидаються два кубики і фіксується пара кольорів: колір грані, якою догори впав перший кубик, і колір грані, якою догори впав другий кубик. Кожну грань першого кубика пофарбовано одним із трьох кольорів – білий, зелений, червоний, так само як і кожну грань другого кубика пофарбовано одним із тих самих трьох кольорів. Тоді простором елементарних подій є множина пар $(k_1\,;k_2)$:

$$\Omega=\{(k_1,k_2)\mid \; k_1\in\{\text{білий, зелений, червоний}\}; \\ k_2\in\{\text{білий, зелений, червоний}\}\}.$$

У цьому разі елементарні події можна інтерпретувати як всі можливі впорядковані пари кольорів (k_1,k_2), де $k_1\in\{$білий, зелений, червоний$\}$, тобто один з кольорів, якими пофарбовано грані першого кубика, $k_2\in\{$білий, зелений, червоний$\}$, тобто один з кольорів, якими пофарбовано грані другого кубика. Так, елементарна подія (білий, червоний) полягає в тому, що перший кубик падає догори білою гранню, а другий – червоною.

Тут множину Ω можна розглядати як декартовий добуток множин A=$\{$білий, зелений, червоний$\}$ і B=$\{$білий, зелений, червоний$\}$, тобто як множину всеможливих впорядкованих пар (k_1, k_2) таких, що $k_1\in A$, $k_2\in B$, де змінна k_1 набуває всіх можливих значень із множини A, і з кожним значенням k_1 змінна k_2 набуває всіх можливих значень із множини B.

Приклад 1.1.4. Підкидається кубик, на гранях якого нанесено мітки "1", "2", "3", "4", "5", "6" (на кожній грані одна цифра, на різних гранях різні цифри), і фіксується лише парна чи непарна цифра на грані, якою кубик впаде догори. Тоді множина Ω =$\{$"парна", "непарна"$\}$ є множиною двох можливих наслідків випробування, тобто простором із двох елементарних подій.

Очевидно, з підкиданням кубика можна пов'язати простори елементарних подій з одним, двома, трьома, чотирма, п'ятьма, шістьма можливими наслідками випробовування. Якщо поділити множину $\{"1","2","3","4","5","6"\}$ на деяку кількість підмножин $H_1,H_2,...,H_k$, $1\le k\le 6$, таких, що $H_iH_j=\varnothing$, коли $i\ne j$, тобто в двох різних множинах H_i та H_j немає спільних елементів, а об'єднанням всіх множин H_i, $i\in\overline{1,\,k}$, є дана множина $\{"1","2","3","4","5","6"\}$, та вважати єдино можливими наслідками випробування попадання в такі підмножини, тоді в просторі Ω елементарних подій буде k елементарних подій $E_1,E_2,...,E_k$ (k можливих наслідків випробовування, $1\le k\le 6$).

Приклад 1.1.5. На гранях кубика нанесено мітки від "1" до "6". Випробування полягає в підкиданні кубика і фіксації грані, якою кубик впаде догори. Тоді множина Ω=$\{$"1", "2", "3", "4", "5", "6"$\}$ є множиною можливих наслідків випробування, тобто простором елементарних подій. Випадання на верхній грані кубика однієї з цифр від 1 до 6 означає, що відбувається відповідна елементарна подія.

Якщо підкидаються два кубики, на гранях яких нанесені цифри від 1 до 6, і як результат випробування фіксувати сумарну кількість очок на двох кубиках, тоді простором елементарних подій буде множина $\Omega = \{2, 3, 4, \dots, 11, 12\}$.

В розглянутих прикладах простір елементарних подій був скінченним. Проте часто трапляються випадки, коли множина Ω нескінченна.

Приклад 1.1.6. В круглу мішень радіуса 1, яку можна вважати множиною точок (x, y) таких, що $x^2+y^2\le 1$, виконується один постріл. Випробування налаштовано так, що попадання кулі за межі мішені неможливе. Тоді множина $\Omega=\{(x, y) \mid x^2+y^2\le 1\}$ є множиною можливих наслідків випробування: через точки (x, y), в які можливе влучення кулі, визначаються відповідні елементарні події в даному випробуванні.

Приклад 1.1.7. Випробування полягає в тому, що з відрізка $[t_1; t_2]$ навмання вибирається два числа. Тоді множину

$$\Omega=\{(x, y) \mid x\in[t_1;t_2],\ y\in[t_1;t_2]\},$$

що складається з усіх впорядкованих пар дійсних чисел (x, y) так званого *декартового добутку* $[t_1;t_2]\times[t_1;t_2]$, можна розглядати як простір елементарних подій в даному випробуванні.

Приклад 1.1.8. Якщо результатом випробування вважати траєкторію частинки в броунівському русі, що починається в певній точці, тоді простором елементарних подій буде нескінченна множина всіх можливих траєкторій руху частинки, які починаються в заданій точці.

Можна навести багато інших прикладів стохастичних випробувань в фізиці, біології, масовому виробництві, в системах автоматичного управління, медицині тощо.

Якщо проводити велику серію спостережень за випадковим явищем, прояв якого за кожного спостереження відбувається за одних і тих самих умов, можна помітити деякі закономірності. В стохастиці саме й вивчаються такі закономірності.

В подальшому вважатимемо, що *випробування полягає в тому, що з деякої множини елементів навмання, незалежно від спостерігача, вибирається один окремий елемент. Поява того* *чи іншого елемента ототожнюється з відбуванням відповідної елементарної події, яка полягає в тому, що вибрано саме цей елемент.*

Отже, в результаті проведення випробування обов'язково відбувається якась одна елементарна подія E із множини Ω елементарних подій.

Поняття елементарної події та простору елементарних подій належать до основних понять теорії ймовірностей і не означаються через простіші поняття, аналогічно до того, як в теорії множин до основних понять належать поняття елемента множини та множини, в геометрії – поняття точки та площини, в теорії алгоритмів – поняття алгоритму і т. д.

Вправи для самостійного виконання

1.1.1. Перевірити чи правильні твердження:

1. Кожне випробування пов'язується з певним простором елементарних подій.
2. Кожна множина Ω може бути простором елементарних подій стосовно деякого випробування.
3. Елементи кожної множини Ω мають бути точками деякого координатного простору з координатами, за якими характеризується місце розташування точок в такому просторі.
4. Поняття елементарної події і простору елементарних подій є означуваними поняттями.

1.1.2. Для даного випробування вказати можливі множини Ω елементарних подій (можливих наслідків випробування):

1. Підкидання монети двічі.
2. Підкидання монети тричі.
3. Підкидання кубика двічі.
4. Підкидання кубика тричі.
5. Фіксація часу зустрічі двох осіб, які домовилися зустрітися на проміжку часу $[t_1;t_2]$.
6. Фіксація типу погоди (сонячно, хмарно, дощ) у першу суботу вересня о 9-й годині ранку.

1.1.3. Визначити, скільки різних просторів елементарних подій можна пов'язати з підкиданням кубика.

1.2. Поняття випадкової події

Нехай Ω – множина елементарних подій, що відповідна певному випробуванню.

Деяку (не будь-яку) *підмножину A множини* Ω називають *подією* (або *випадковою подією*).

Кажуть, що в результаті проведення *випробування відбулася подія A*, якщо відбулася елементарна подія *E* така (із множини Ω навмання вибрано елемент *E* такий), що $E \in A$. Множини $A = \Omega$ та $B = \varnothing$ завжди вважаються подіями (на відміну від інших підмножин множини Ω).

Оскільки Ω – множина всіх можливих наслідків випробування, то в результаті кожного випробування подія Ω обов'язково відбувається, бо завжди $E \in \Omega$. Тому подію Ω називають *вірогідною.*

В множині $B = \varnothing$ немає жодного елемента (елементарної події) з множини Ω, тому подія $B = \varnothing$ ніколи не може відбутися в результаті проведення випробування, бо завжди $E \overline{\in} \varnothing$. Тому подію $B = \varnothing$ називають *неможливою*.

Приклад 1.2.1. Випробування полягає у виборі навмання однієї кульки із 10, серед яких є білі і червоні. Як наслідок випробування фіксується колір кульки, $\Omega = \{б, ч\}$. Як події можуть розглядатися підмножини множини Ω: $A = \{б\}$, $B = \{ч\}$, $C = \{б, ч\}$, $D = \varnothing$, що означає відповідно: *A* – навмання вибрана кулька біла, *B* – навмання вибрана кулька червона, *C* – навмання вибрана кулька біла або червона, *D* – навмання вибрана кулька ні біла, ні червона. З відбуванням елементарної події *б* відбувається подія $A = \{б\}$, бо $б \in \{б\}$, і подія $C = \{б, ч\}$, бо $б \in \{б, ч\}$, але не відбувається подія $B = \{ч\}$ і подія $D = \varnothing$. Аналогічно, якщо відбувається елементарна подія *ч*, то відбувається подія $B = \{ч\}$, бо $ч \in \{ч\}$, і подія $C = \{б, ч\}$, бо $ч \in \{б, ч\}$, однак не відбувається подія $A = \{б\}$ і подія $D = \varnothing$.

Приклад 1.2.2. Випробування полягає у підкиданні кубика один раз, Ω={"1", "2", "3", "4", "5", "6"}. Якщо одна з випадкових подій *A*={"3", "6"} полягає в тому, що на верхній грані кубика випадає число, кратне 3, то коли випаде "3"$\in\Omega$, відбувається подія *A*={"3", "6"}, а коли випаде "2"$\in\Omega$, не відбувається ця подія, оскільки "2" $\overline{\in}$ *A*={"3", "6"}.

Приклад 1.2.3. Випробування полягає в тому, що монету підкидають тричі. Нехай *A* – подія, яка полягає в тому, що принаймні двічі випаде герб. Множиною елементарних подій у даному експерименті будемо вважати сукупність впорядкованих трійок результатів першого, другого та третього підкидань монети, тобто

$$\Omega = \{ГГГ, ГГЦ, ГЦГ, ГЦЦ, ЦГГ, ЦГЦ, ЦЦГ, ЦЦЦ\}.$$

Тоді

$$A = \{ГГГ,\ ГГЦ,\ ГЦГ, ЦГГ\}.$$

Приклад 1.2.4. На стіл (з бортами) навмання кидається тенісна кулька. Нехай елементарна подія E полягає в тому, що кулька опиниться в певній точці стола, A – подія, яка полягає в тому, що кулька опиниться в певній частині площини стола. Тут множина Ω нескінченна (кожній можливій точці зупинки кульки відповідає елементарна подія, яка полягає в тому, що кулька зупиниться саме в цій точці). Так само нескінченною може бути й підмножина A, через яку визначається розглядувана подія.

Приклад 1.2.5. Виконується постріл в круглу мішень одиничного радіуса, причому куля не може влучити за межі мішені. В даному разі $\Omega = \{(x, y) \mid x^2 + y^2 \le 1\}$. Нехай одна з подій $A = \Omega$ полягає в тому, що куля влучає в мішень. З вибором будь-якого елемента множини Ω відбувається подія A і не відбувається подія $B = \varnothing$. Якщо множина $C = \{(x, y) \mid x^2 + y^2 \le 0.2\}$ – подія, то коли буде вибрано точку (0, 0), відбудеться подія C, а коли буде вибрано точку $\left(\frac{1}{2}, \frac{1}{2}\right)$, ця подія не відбудеться.

Нехай підмножини A і B множини Ω є подіями. Якщо $A \subset B$, тобто кожен елемент множини A є елементом і множини B, то говорять, що через *подію A спричинюється подія B*. Отже, через подію A спричинюється подія B тоді і тільки тоді, коли з відбуванням події A відбувається і подія B (Рис. 1.2.1).

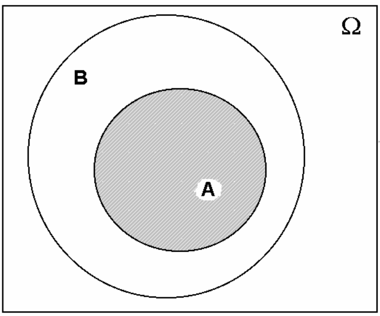


Рис.1.2.1

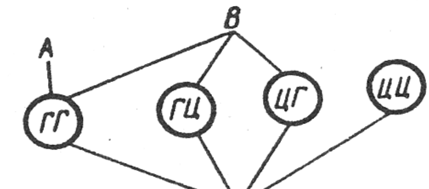


Рис. 1.2.2

Події A і B називають *рівними*, якщо $A \subset B$ і $B \subset A$, тобто через кожну з них спричинюється інша. Отже, події A і B рівні тоді і тільки тоді, коли вони одночасно або відбуваються, або не відбуваються.

Приклад 1.2.6. Випробування полягає в тому, що двічі підкидають монету. Як елементарні події розглядаються всі можливі пари випадань герба – $Г$ і цифри – $Ц$, $\Omega=\{ГГ, ГЦ, ЦГ, ЦЦ\}$. Нехай A – подія, яка полягає в тому, що двічі випаде герб. Тоді через подію A спричинюється подія B, яка полягає в тому, що принаймні один раз випаде герб. Тут

$$A=\{ГГ\},\ B=\{ГГ, ГЦ, ЦГ\},$$

і таким чином $A\subset B$, тобто через подію A спричинюється подія B (Рис. 1.2.2).

Стосовно будь-якої множини Ω та підмножин A, B, C цієї множини мають місце такі властивості.

1. $A\subset A,\ A\subset\Omega,\ \varnothing\subset A$.
2. Якщо $A\subset B$ і $B\subset C$, то $A\subset C$.
3. $A=A$.
4. Якщо $A=B$, то $B=A$.
5. Якщо $A=B$, $B=C$, то $A=C$.

Позначаючи події (множини) кругами Ейлера, дістанемо відповідні геометричні ілюстрації (Рис. 1.2.1).

Приклад 1.2.7. Якщо Ω={“1”, “2”, “3”, “4”, “5”, “6”}, то через подію A={“3”, “6”}спричинюється подія B={“1”,“3”,“6”}, через яку в свою чергу спричинюється подія C={“1”,“2”,“3”,“6”}. Подія C еквівалентна події D, яка полягає в тому, що на верхній грані кубика випадає число, відмінне від 4 і 5.

Вправи для самостійного виконання

1.2.1. Перевірити, чи правильні твердження:

1. Множина елементарних подій – це будь-яка множина.
2. Множина елементарних подій – це множина точок в деякому координатному просторі.
3. Подія – це будь-яка підмножина множини Ω елементарних подій.
4. Елементарна подія є подією.
5. З випадковим вибором навмання кожного елемента із множини елементарних подій відбувається певна подія.
6. Існує елементарна подія, з відбуванням якої не відбувається жодна подія.
7. За будь-яких подій A і B принаймні через одну з них спричинюється інша.

1.2.2. Вказати всі можливі події, що визначаються за множиною Ω елементарних подій:

1. Ω – множина наслідків випробування, яке полягає в підкиданні відразу двох монет (1 і 2 копійки). Як наслідки випробування розглядаються:

а) усі можливі пари випадань герба і цифри на двох монетах – $\Omega=\{ГГ, ГЦ, ЦГ, ЦЦ\}$, якщо першою вказується монета 1 копійка;

б) кількість випадань герба на обох монетах, тобто $\Omega=\{0,1,2\}$.

2. Ω – множина наслідків випробування – пострілу в мішень, на якій зазначені можливі кількості отриманих очків: 0, 1, 2, 3, 4, 5 в залежності

від відстані точки влучення від центра мішені. Як елементарні події розглядаються наступні наслідки – 5 чи не 5 очків отримується після пострілу.

3. Ω – множина наслідків випробування, що полягає в підкиданні відразу двох кубиків (червоного та білого кольору), на гранях кожного з яких нанесені цифри від 1 до 6. Як наслідки випробування розглядаються:

а) можливі пари цифр $(i,\ j),\ i \in \overline{1,6},\ j \in \overline{1,6}$,

б) можливі суми цифр $(i + j),\ i \in \overline{1,6},\ j \in \overline{1,6}$,

де i – цифра, що випала на верхній грані одного з кубиків (білого кольору), j – цифра, що випала на верхній грані іншого кубика (червоного кольору).

1.3. Операції над подіями

Оскільки події – це деякі підмножини множини Ω елементарних подій, то над подіями можна ввести такі самі операції, як і над множинами.

Нехай $A \subset \Omega$ і $B \subset \Omega$ – деякі події.

Означення. *Сумою подій A* і *B* називають таку подію *C*, яка відбувається тоді і тільки тоді, коли відбувається принаймні одна з подій *A* або *B*.

Це означає, що кожна елементарна подія *E*, що належить до множини *C*, належить принаймні до однієї із множин *A* або *B*, і навпаки, якщо $E \in A$ або $E \in B$, то $E \in C$. В теоретико-множинному тлумаченні сумі *C* подій *A* і *B* відповідає об'єднання відповідних множин. Суму подій *A* і *B* позначають $A \bigcup B$ або $A+B$.

Аналогічно визначається *сума довільної кількості подій* A_k, $k \in K$, – це подія, яка відбувається тоді й тільки тоді, коли відбувається принаймні одна з подій A_k, $k \in K$. *Суму скінченної кількості n* подій A_k позначають $\bigcup_{k=1}^{n} A_k$, а також $\sum_{k=1}^{n} A_k$.

Зауважимо, що для того, щоб одержати об'єднання (суму) двох множин *A* і *B*, потрібно до однієї з них приєднати з іншої ті елементи, яких немає в першій множині.

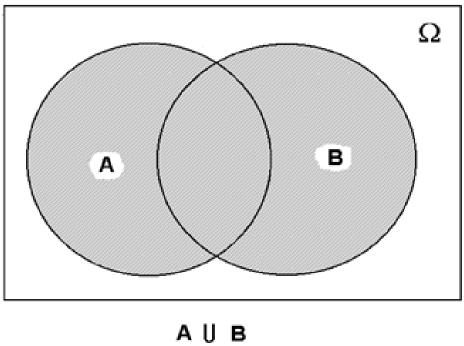


Рис 1.3.1

Геометричне тлумачення суми подій *A* і *B* подане на Рис. 1.3.1, де прямокутником зображено множину елементарних подій Ω, один з кругів – подія *A*, інший – подія *B*, заштрихована множина – подія $A+B = A \bigcup B$.

Приклад 1.3.1. Нехай Ω={"1", "2", "3", "4", "5", "6"} – множина елементарних подій, пов'язана з підкиданням кубика один раз, *A*={"3", "6"} – подія, що полягає у випаданні на верхній грані кубика числа, кратного 3, а B={“2”, “4”, “6”} – подія, що полягає у випаданні парного числа. Тоді $A+B=A\bigcup B=$ ={“2”, “3”, “4”, “6”} – подія, яка полягає в тому, що на верхній грані кубика випаде або число, кратне 2 (відбувається подія B), або число, кратне 3 (відбувається подія A) (Рис. 1.3.2).

Приклад 1.3.2. Нехай подія A полягає в тому, що навмання вибране ціле додатне однорозрядне число ділиться на 3, а подія B – навмання вибране однорозрядне число є парним. Елементарна подія E_i – навмання вибране однорозрядне число є i, де $i \in \{1,\ 2\ ...,\ 9\}$, $(E_i = "i")$.

Тоді

$$\Omega = \{"1",\ "2","3","4","5","6","7","8","9"\},$$

$$A = \{"3",\ "6",\ "9"\},$$

$$B = \{"2",\ "4",\ "6",\ "8"\}.$$

Події $C = A + B$ відповідає множина

$$C = \{"2","3","4","6","8","9"\}.$$

Кількість елементів у множині $C = A + B$ не обов’язково дорівнює сумарній кількості елементів в множинах A і B, оскільки спільні елементи враховуються лише один раз.

Зазначимо, що кожна множина є об’єднанням (сумою) її одноелементних підмножин. Якщо в наведеному прикладі E_i={i}, $i \in \overline{1,6}$, то

$$A = \{E_3\} + \{E_6\},\quad B = \{E_2\} + \{E_4\} + \{E_6\},$$

$$\Omega = \{E_1\} + \{E_2\} + \{E_3\} + \{E_4\} + \{E_5\} + \{E_6\}.$$

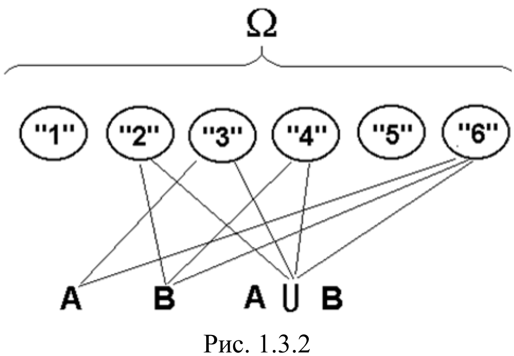


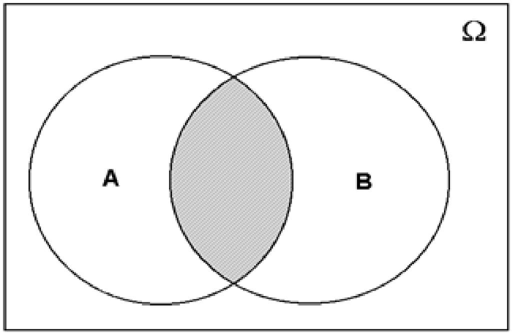
Рис. 1.3.2

Означення. *Добутком подій A* і *B* називають таку подію *C*, яка відбувається тоді і тільки тоді, коли відбуваються обидві події *A* і *B*.

Це означає, що кожна елементарна подія *E*, що належить до множини *C,* належить також до обох множин *A* і *B*, тобто $E \in A$ і

$E \in B$, і навпаки, якщо $E \in A$ і $E \in B$, то $E \in C$.

В теоретико-множинному тлумаченні добуткові подій A і B відповідає перетин відповідних множин, тобто сукупність елементів, що належать як до множини A, так і до множини B.



A ⋂ B

Рис. 1.3.3

Добуток подій A і B позначають $A \bigcap B$, або $A \cdot B$ або просто AB.

Аналогічно визначається *добуток довільної кількості подій* A_k, $k \in K$, – це подія, яка відбувається тоді й тільки тоді, коли відбуваються всі події A_k, $k \in K$, під час проведення одного і того самого випробування. *Добуток скінченної кількості n подій* позначають $\bigcap_{k=1}^{n} A_k$, а також $\prod_{k=1}^{n} A_k$.

Зауважимо, що для того, щоб одержати перетин (добуток) двох множин A і B, слід взяти всі елементи, які належать до обох цих множин.

Геометричне тлумачення добутку подій A і B подане на Рис. 1.3.3, де заштрихована множина точок – подія $A \cdot B = A \bigcap B$.

Приклад 1.3.3. Нехай подія A полягає в тому, що навмання вибране ціле додатне однорозрядне число ділиться на 3, а подія B – навмання вибране однорозрядне число є парним. Елементарна подія E_i – навмання вибране однорозрядне число є i, де $i \in \{1, 2, ..., 9\}$, $(E_i = i)$.

Тоді

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}, \ A = \{3, 6, \ 9\}, \ B = \{2, 4, 6, 8\}.$$

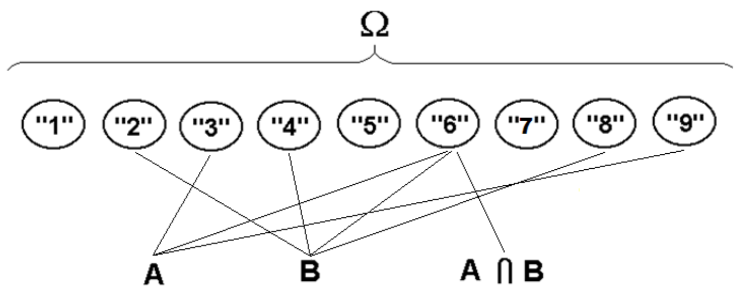


Рис. 1.3.4

Подія $C = AB$ означає, що навмання вибране число є число 6 (і ділиться на 3, і парне).

Приклад 1.3.4. Мішень складається з 10 кругів, обмежених концентричними колами з радіусами r_k, $k \in \{1, 2, ..., 10\}$, причому $r_1 < r_2 < r_3 < ... < r_{10}$. Елементарні події ототожнюються з точками мішені. Нехай A_k – подія, яка полягає у влученні в круг радіуса r_k. Тоді події $B = \bigcup_{k=1}^{6} A_k$ і $C = \bigcap_{k=1}^{10} A_k$ – множини, зображені відповідно на Рис. 1.3.5, а, б.

Означення. Події A і B називають *несумісними,* якщо $A \cdot B = \varnothing$, тобто якщо вони не можуть відбутися обидві в результаті одного і того самого випробування.

Приклад 1.3.5. Якщо Ω={”1”, “2”, “3”, “4”, “5”, “6”}, C={”1”, “3”, “5”}, D={”2”, “4”, “6”}, то C і D несумісні події, оскільки не існує елементарної події, яка належить як до множини C, так і до множини D.

Означення. *Різницею подій* A і B (A мінус B) називають таку подію C, яка відбувається тоді і тільки тоді, коли відбувається подія A і не відбувається подія B.

Це означає, що кожна елементарна подія E, яка належить до множини C, належить також до множини A і не належить до множини B, і навпаки, якщо $E \in A$ і $E \overline{\in} B$, то $E \in C$. В теоретико-множинному тлумаченні різниці подій A і B відповідає різниця відповідних множин.

Різницю подій A і B (A мінус B) позначають $A \setminus B$, а також $A - B$.

Зауважимо, що для того, щоб одержати різницю множин $A \setminus B$, слід із множини A вилучити всі елементи множини B.

Геометричне тлумачення різниці подій $A \setminus B$ подано на Рис. 1.3.6, де заштрихована множина точок – подія $A \setminus B$.

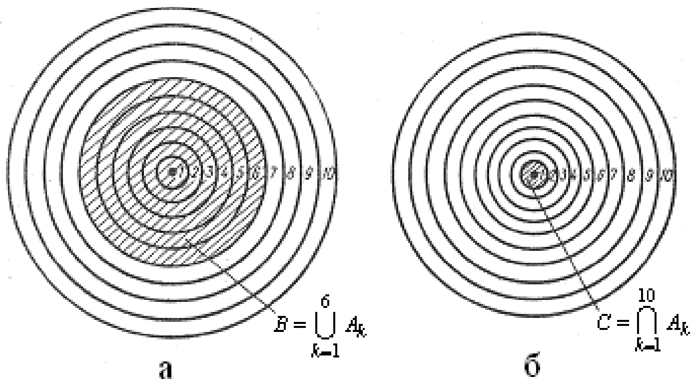


Рис. 1.3.5

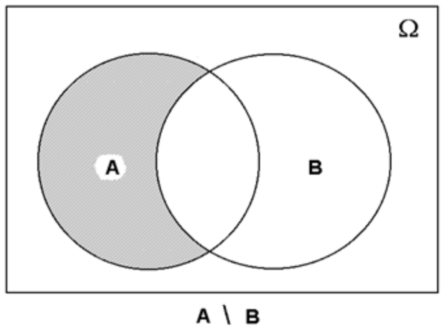


Рис. 1.3.6

Приклад 1.3.6. Серед учнів 11 класу, в якому навчаються не менше п'яти спортсменів, навмання вибирається група з п'яти учнів. Нехай *A* – подія, яка полягає в тому, що в групі не більше трьох спортсменів, *B* – подія, яка полягає в тому, що спортсменів не менше одного. Можливими наслідками даного випробування щодо кількості спортсменів у групі очевидно є: E_0 – 0 спортсменів, E_1 – 1 спортсмен, E_2 – 2 спортсмени, ..., E_5 – 5 спортсменів. Тоді

$$\Omega = \{E_0,\, E_1,\, E_2,\, E_3,\, E_4,\, E_5\}.$$

Подія $A \setminus B$ полягає в тому, що в групі з п'яти учнів спортсменів немає, а $B \setminus A$ – спортсменів не менше чотирьох (4 або 5).

Означення. *Подією, протилежною до події A,* називають різницю $\Omega \setminus A$, яку позначають $\overline{A}$.

Отже, подія $\overline{A}$, протилежна до події A, відбувається тоді і тільки тоді, коли не відбувається подія A. Зрозуміло, що подія A є протилежною до події $\overline{A}$. Тому події *A* і $\overline{A}$ називають взаємно протилежними. Очевидно, події A і $\overline{A}$ несумісні.

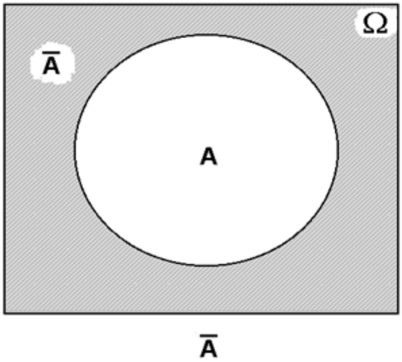


Рис. 1.3.7

Геометричне тлумачення взаємно протилежних подій подано на Рис. 1.3.7, де заштрихована множина точок – подія $\overline{A}$, протилежна до події A.

Приклад 1.3.7. Якщо Ω={“1”, ”2”, ”3”, ”4”, ”5”, ”6”} і A={“3”, ”6”} та B={“2”, ”4”, ”6”} – події, то $A \setminus B$ = {“3”} – подія, що полягає у випаданні непарного числа, кратного 3, $B \setminus A$ = ={“2”, ”4”} – подія, що полягає у випаданні парного числа, не кратного 3. $\overline{A} = \Omega \setminus A$ = {“1”, ”2”, ”4”, ”5”} – подія, що полягає у випаданні числа, не кратного 3, а $\overline{B} = \Omega \setminus B$ = {“1”,”3”,”5”} – подія, що полягає у випаданні непарного числа.

Приклад 1.3.8. Нехай A подія, яка полягає в тому, що в результаті підкидання кубика на його верхній грані випаде не менше, ніж три очка. Нехай можливими наслідками цього випробування є елементарні події $E_i = "i"$, $i \in \{1, ..., 6\}$, які полягають в тому, що на верхній грані кубика випаде i очок. Тоді

$$\Omega = \{"1", "2", "3", "4", "5", "6"\};\ A = \{"3", "4", "5", "6"\};$$

$$\overline{A} = \{"1", "2"\}$$

Таким чином, тут подія $\overline{A}$ означає, що на верхній грані кубика випаде менше, ніж три очка.

Подія $\overline{\Omega}$, протилежна до вірогідної, є неможливою. Неможливій події відповідає порожня множина $\overline{\Omega} = \Omega \setminus \Omega = \varnothing$ можливих наслідків випробування. Зрозуміло, що

$$\overline{\varnothing} = \Omega \setminus \varnothing = \Omega.$$

Якщо A_1, A_2, ... – деяка послідовність подій, то події A_1, $A_2 \setminus (A_2 \cap A_1)$, $A_3 \setminus (A_3 \cap (A_1 \cup A_2))$, ... попарно несумісні і сума їх дорівнює сумі подій A_i.

У такий спосіб можна дістати правило розкладання суми деяких подій на суму попарно несумісних подій. Слід зазначити, що $A_1, A_2 \setminus A_1,\ A_3 \setminus (A_1 \bigcup A_2), \ldots$ є також попарно несумісними, і сума їх дорівнює сумі подій A_i.

Вправи для самостійного виконання

1.3.1. Перевірити, чи правильні твердження:

1. Якщо відбувається подія A, то відбувається подія $A+B$.
2. Через подію $A+B$ спричинюється подія A.
3. Якщо відбувається подія $A \cdot B$, то відбувається і подія A.
4. Через подію A спричинюється подія $A \cdot B$.
5. Через подію $A \cdot B$ спричинюється подія A.
6. Якщо не відбувається подія $A \setminus B$, то відбувається подія B.
7. Якщо відбувається подія A, то відбувається і подія $A \setminus B$.
8. Якщо $C = A - B$, то $A = C + B$.
9. $(A + B) - B = A$; $(A - B) + B = A$.
10. $A + \overline{A} = \Omega$.
11. Неможлива подія – це добуток протилежних подій.
12. Якщо $A = \overline{B}$, то $\overline{A} = B$, тобто події $\overline{\overline{B}}$ і B рівні.
13. $A \bigcap B = \overline{\overline{A} \bigcup \overline{B}}$.
14. $A \setminus B = A \bigcap \overline{B}$.
15. $\bigcap\limits_k A_k = \overline{\bigcup\limits_k \overline{A_k}}$.

1.3.2. Мішень складається з 10 кругів, обмежених концентричними колами з радіусами $r_k : r_1 < r_2 < \ldots < r_{10}$. Подія A_k полягає у влученні в круг радіуса r_k. Що означають події:

1. $B = \sum\limits_{k=1}^{6} A_k$? 2. $C = \prod\limits_{k=1}^{10} A_k$? 3. $A_{k+1} \setminus A_k$, $k \in \overline{1,9}$?

4. $A_1 \setminus A_2$? 5. $\overline{A_k}$, $k \in \overline{1, 10}$?

1.4. Властивості операцій над подіями

Нехай Ω – простір елементарних подій, $A \subset \Omega$, $B \subset \Omega$ і $C \subset \Omega$ – довільні випадкові події, $\varnothing$ – неможлива подія. Стосовно введених *операцій над подіями задовільняються такі закони*:

1. $\overline{\overline{A}} = A$ – закон подвійного заперечення.

$$\left.\begin{array}{l} 2.\ A+B=B+A \\ 3.\ AB=BA \end{array}\right\}$$ комутативні (переставні) закони додавання і множення.

$$\left.\begin{array}{l} 4.\ (A+B)+C=A+(B+C) \\ 5.\ (AB)C=A(BC) \end{array}\right\}$$ *асоціативні (сполучні) закони додавання і множення*

6. $(A+B)C = AC+BC$ – перший дистрибутивний (розподільний) закон.

7. $AB+C=(A+C)(B+C)$ – другий дистрибутивний (розподільний) закон.

8. $A+A=A$.

9. $A\cdot A=A$.

10. $A+\overline{A}=\Omega$.

11. $A\cdot\overline{A}=\varnothing$.

12. $A+\Omega=\Omega$.

13. $A\cdot\Omega=A$.

14. $A+\varnothing=A$.

15. $A\cdot\varnothing=\varnothing$.

$$\left.\begin{array}{l} 16.\ \overline{A+B}=\overline{A}\cdot\overline{B} \\ 17.\ \overline{A\cdot B}=\overline{A}+\overline{B} \end{array}\right\}$$ – закони двоїстості (правила де Моргана).

Рівності 1–17 випливають безпосередньо з наведених вище означень.

Для прикладу розглянемо рівність 7. Геометрично цей закон можна проілюструвати так, як показано на Рис. 1.4.1, а, б.

Сумістивши ці зображення, можна впевнитися, що множини $AB+C$ і $(A+C)\ (B+C)$ співпадають.

Аналітичне доведення рівності 7 може бути таким.

Нехай E – довільна елементарна подія така, що $E \in AB+C$. Тоді справджується принаймні одне з двох – або $E \in AB$, або $E \in C$. Якщо $E \in AB$, то $E \in A$ і $E \in B$. Проте тоді $E \in A+C$ і $E \in B+C$ (за означенням суми подій), а отже $E \in (A+C)\ (B+C)$. Якщо $E \in C$, то $E \in C+A$ і $E \in C+B$, і тому $E \in (A+C)\ (B+C)$.

Таким чином, довільний елемент множини $AB+C$ є елементом множини $(A+C)(B+C)$, тобто

$$(AB+C)\subset(A+C)\,(B+C).$$

Отже, через подію $AB+C$ спричинюється подія $(A+C)\,(B+C)$.

Нехай, навпаки, $E\in(A+C)\,(B+C)$. Це означає, що $E\in A+C$ і $E\in B+C$. Якщо $E\in C$, то $E\in C+AB$. Якщо $E\overline{\in}C$, то $E\in A$ і $E\in B$, тобто $E\in AB$, а отже $E\in AB+C$.

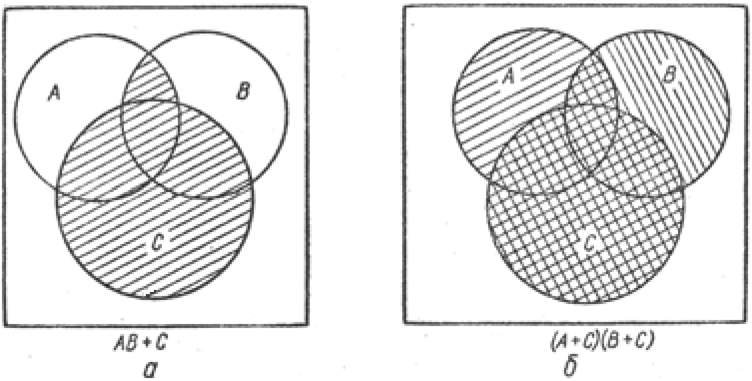


Рис. 1.4.1

Таким чином, довільний елемент множини $(A+C)\,(B+C)$ є елементом множини $AB+C$, тобто $(A+C)\,(B+C)\subset AB+C$, і через подію $(A+C)(B+C)$ спричинюється подія $AB+C$. Звідси і з того, що $(AB+C)\subset(A+C)\,(B+C)$, слідує

$$AB+C=(A+C)\,(B+C).$$

Інші рівності 1–17 доводяться аналогічно.

Приклад 1.4.1. За будь-яких подій $A\subset\Omega$ і $B\subset\Omega$ виконуються наступні співвідношення:

а) $AB\subset A,\ AB\subset B$, тому $\varnothing=\varnothing\cdot A\subset A,\ \varnothing=\varnothing\cdot B\subset B$;

б) $A\subset A+B,\ B\subset A+B$ (співвідношення а) і б) випливають безпосередньо з означення суми й добутку подій);

в) якщо $A\subset C$ і $B\subset C$, то й $A+B\subset C$ (співвідношення випливає з означення суми подій);

г) якщо $C\subset A$ і $C\subset B$, то $C\subset AB$ (співвідношення випливає з означення добутку подій);

д) якщо $A\subset B$, то $A+B=B,\ (A+B)\setminus A=B\setminus A$.

Справді, оскільки $A\subset B$ і $B\subset B$, то згідно із співвідношенням в) $A+B\subset B$, а згідно із співвідношенням б) $B\subset A+B$.

Таким чином $A+B=B,\ (A+B)\setminus A=B\setminus A$. Звідси випливає також, що коли $A+B=A$, то $B\subset A$;

е) якщо $A\subset B$, то $AB=A$. Справді, згідно із співвідношенням а) $AB\subset A$, а оскільки $A\subset A$ і $A\subset B$, то згідно із співвідношенням г) $A\subset AB$. Отже $AB=A$. Звідси, якщо $AB=A$, то $A\subset B$.

Приклад 1.4.2. Довести, що $A+\bar{A}B+\overline{A+B}=\Omega$.

На основі другого дистрибутивного закону

$$A+\bar{A}B=(A+\bar{A})\ (A+B).$$

За законами 10 і 13

$$(A+\bar{A})\ (A+B)=\Omega\ (A+B)=A+B.$$

Таким чином,

$$A+\bar{A}B+\overline{A+B}=(A+\bar{A})\ (A+B)+\overline{(A+B)}=(A+B)+\overline{(A+B)}=\Omega$$

.

За співвідношенням *а*) слідує

$$\varnothing=\varnothing\cdot A\subset A=A\cdot\Omega\subset\Omega \text{ і } A=A\cdot A\subset A,$$

за будь-якої підмножини A простору Ω.

Приклад 1.4.3. Довести, що з рівності $A+B=AB$ випливає, що $A=B$.

Справді, $A\subset A+B=AB\subset B$ і тому $A\subset B$. Аналогічно $B\subset A+B=AB\subset A$ і тому $B\subset A$. Отже, $A\subset B$ і $B\subset A$, тобто $A=B$.

Легко бачити, що події A_1 і $A_2\setminus(A_2\bigcap A_1)$ несумісні і $A_1\bigcup A_2=A_1\bigcup(A_2\setminus(A_2\bigcap A_1))$.

Означення. Якщо $A_1, A_2, \ldots$ – нескінченна послідовність множин $A_i\subset\Omega$, то множину A^*, елементи якої є елементами нескінченної кількості множин цієї послідовності, називають *верхньою границею послідовності* множин A_i і позначають $\overline{\lim_{i\to\infty}}\ A_i$.

Множину A_*, що складається з елементів, кожен з яких є елементом всіх множин послідовності, крім, можливо, скінченного числа їх, називають *нижньою границею послідовності* множин A_i і позначають $\underline{\lim_{i\to\infty}}\ A_i$.

Множину A^* можна інтерпретувати як подію, яка полягає в тому, що відбувається нескінченно багато подій з послідовності $A_1, A_2, \ldots$. Множина A_* є подією, яка полягає в тому, що відбуваються всі події, починаючи з деякого номера, або, що те саме, не відбувається тільки скінченна кількість подій з послідовності $A_1, A_2, \ldots$. Очевидно, що $A_*\subset A^*$.

Якщо $\underline{\lim}_{i\to\infty} A_i = \overline{\lim}_{i\to\infty} A_i$, то говорять, що послідовність $A_1, A_2, \ldots$ збіжна до границі, яку позначають $\lim_{i\to\infty} A_i$.

Можна довести, що коли множини $A_1, A_2, \ldots$, де $A_i \subset \Omega$, довільні, то

$$A^* = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i, \quad A_* = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i. \tag{3.1}$$

Приклад 1.4.4. Нехай задано послідовність множин $A_i = (-\infty; (-1)^i], i = 1, 2, 3, \ldots$

Надаючи n значень $1, 2, 3, \ldots$, отримаємо

$$\bigcap_{i=1}^{\infty} (-\infty;\ (-1)^i] = (-\infty; -1], \ \bigcap_{i=2}^{\infty} (-\infty; (-1)^i] = (-\infty; -1],$$

$$\bigcap_{i=3}^{\infty} (-\infty; (-1)^i] = (-\infty; -1], \ldots,$$

тому $A_* = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} (-\infty; (-1)^i] = (-\infty; -1]$.

Аналогічно

$$\bigcup_{i=1}^{\infty} (-\infty;\ (-1)^i] = (-\infty; +1]; \ \bigcup_{i=2}^{\infty} (-\infty; (-1)^i] = (-\infty; +1];$$

$$\bigcup_{i=3}^{\infty} (-\infty; (-1)^i] = (-\infty; +1]; \ldots,$$

тому $A^* = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} (-\infty; (-1)^i] = (-\infty; +1]$.

Оскільки $A_* \neq A^*$, то у розглядуваної послідовності множин $A_i = (-\infty; (-1)^i]$ границі немає.

Приклад 1.4.5. Нехай задано *зростаючу послідовність множин* A_i, $i = 1, 2, 3, \ldots$, тобто таку, що $A_1 \subset A_2 \subset A_3 \subset \ldots$.

Тоді

$$\bigcap_{i=1}^{\infty} A_i = A_1, \ \bigcap_{i=2}^{\infty} A_i = A_2, \ \bigcap_{i=3}^{\infty} A_i = A_3, \ \ldots,$$

тому

$$A_* = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i.$$

З іншого боку

$$\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=2}^{\infty} A_i \subset \bigcup_{i=3}^{\infty} A_i \subset \ldots,$$

тому

$$A^* = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i = \bigcup_{i=1}^{\infty} A_i .$$

Отже $A_* = A^* = \bigcup_{i=1}^{\infty} A_i$.

Таким чином, якщо послідовність множин A_i , $i = 1, 2, 3, ...$ зростаюча, тобто $A_i \subset A_{i+1}$, $i = 1, 2, 3, ...$, то така послідовність збіжна до границі $\lim_{i\to\infty} A_i = \bigcup_{i=1}^{\infty} A_i$.

Нехай тепер задано *спадаючу послідовність множин* A_i , $i = 1, 2, 3, ...$, тобто таку, що $A_1 \supset A_2 \supset A_3 \supset ...$.

Тоді $\bigcup_{i=1}^{\infty} A_i = A_1$, $\bigcup_{i=2}^{\infty} A_i = A_2$, $\bigcup_{i=3}^{\infty} A_i = A_3$, …,

тому

$$A^* = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i = \bigcap_{i=1}^{\infty} A_i .$$

З іншого боку

$\bigcap_{i=1}^{\infty} A_i \supset \bigcap_{i=2}^{\infty} A_i \supset \bigcap_{i=3}^{\infty} A_i \supset ...$, тому $A_* = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i = \bigcap_{i=1}^{\infty} A_i$.

Отже $A_* = A^* = \bigcap_{i=1}^{\infty} A_i$.

Таким чином, якщо послідовність множин A_i , $i = 1, 2, 3, ...$, спадаюча, тобто $A_i \supset A_{i+1}$, $i = 1, 2, 3, ...$, то така послідовність збіжна до границі $\lim_{i\to\infty} A_i = \bigcap_{i=1}^{\infty} A_i$.

Зростаючі або спадні *послідовності* множин називають *монотонними*. Таким чином, границя монотонної послідовності множин завжди існує.

Зауважимо, що коли множини $A_1, A_2, ...$, де $A_i \subset \Omega$, довільні, тоді $B_n = \bigcup_{i=n}^{\infty} A_i \supset B_{n+1} = \bigcup_{i=n+1}^{\infty} A_i$, тобто послідовність множин $B_n = \bigcup_{i=n}^{\infty} A_i$ монотонно спадаюча, а тому існує границя $\lim_{n\to\infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$. Аналогічно послідовність множин

$G_n = \bigcap\limits_{i=n}^{\infty} A_i$ монотонно зростаюча, $G_n = \bigcap\limits_{i=n}^{\infty} A_i \subset G_{n+1} = \bigcap\limits_{i=n+1}^{\infty} A_i$, а тому існує границя $\lim\limits_{n\to\infty} G_n = \bigcup\limits_{n=1}^{\infty} G_n = \bigcup\limits_{n=1}^{\infty}\bigcap\limits_{i=n}^{\infty} A_i$.

Вправи для самостійного виконання

1.4.1. Перевірити, чи правильні твердження:

1. $A \bigcap B = \overline{\overline{A} \bigcup \overline{B}}$. 2. $A \setminus B = A \bigcap \overline{B}$. 3. $\bigcap\limits_{k} A_k = \overline{\bigcup\limits_{k} \overline{A_k}}$.

1.4.2. Довести закони двоїстості.

1.4.3. Вивести правило, за яким суму довільної кількості подій можна подати у вигляді суми попарно несумісних подій.

1.5. Простір подій. Уточнення поняття випадкової події

В п.1.2 сказано, що випадковою подією називають *деяку* (не будь-яку) підмножину простору Ω елементарних подій, що відповідний певному випробуванню.

Після введення операцій над подіями можна дещо уточнити, які підмножини простору Ω можна вважати подіями, а які не можна.

Отже, нехай Ω – простір елементарних подій, що відповідний певному випробуванню, а S – *деяка* (не будь-яка, див. § 1.9) *сукупність підмножин множини* Ω, стосовно якої задовільняються умови:

1) $\Omega \in S$;

2) якщо $A \in S$, то і $\overline{A} \in S$;

3) якщо $A_i \in S$, $i \in \{1,2,3,\ldots\}$, то $\bigcup\limits_{i \in I} A_i \in S$, де $I \subset \{1,2,3,\ldots\}$ – довільна підмножина множини індексів 1, 2, 3, … .

Тоді *кожну підмножину* $A \subset \Omega$, *віднесену до сукупності S, називають випадковою подією* або просто *подією*, а сукупність S називають *простором подій*, що відповідний даному простору Ω елементарних подій.

Таким чином, поняття події вводиться лише у зв'язку з поняттям простору подій.

Зауважимо, що $\varnothing = \overline{\Omega}$ і тому за умовами 1) і 2) $\varnothing \in S$.

Оскільки $\bigcap\limits_{i \in I} A_i = \overline{\bigcup\limits_{i \in I} \overline{A_i}}$, то за умовами 2) і 3) $\bigcap\limits_{i \in I} A_i \in S$, коли $A_i \in S$, $I \subset \{1, 2, 3, \ldots\}$. Зокрема, $A \bigcap B \in S$, коли $A \in S$ і $B \in S$. Враховуючи це та рівність $A \setminus B = A \bigcap \overline{B}$, з умови 2) дістанемо, що $A \setminus B \in S$, коли $A \in S$ і $B \in S$.

Отже, *простір S випадкових подій є замкненим відносно операцій суми, різниці та добутку подій*.

Приклад 1.5.1. За будь-якої множини Ω елементарних подій покладемо $S=\{\varnothing, \Omega\}$. Тоді стосовно S задовільняються умови 1)-3), тобто S може бути простором подій (*найвужчим* з можливих), а тому можлива ситуація, коли подіями вважаються лише неможлива та вірогідна події.

Приклад 1.5.2. Якщо всі грані кубика пофарбовано в білий колір і за підкидання кубика фіксується колір грані, якою кубик падає догори, тоді в Ω={біла} є один єдиний елемент. Тому в даному випадку простір S може складатися лише з двох елементів: $S = \{\varnothing, \Omega\}$.

Приклад 1.5.3. За будь-якого простору Ω елементарних подій покладемо $S=\{A \mid A \subset \Omega\}$, тобто до сукупності S віднесемо будь-які підмножини простору Ω. Тоді стосовно S задовільняються умови 1)-3), тобто S може бути простором подій (*найширшим* з можливих), а тому можлива ситуація, коли подіями будуть будь-які підмножини простору Ω.

Приклад 1.5.4. Нехай Ω={“1”, “2”, “3”, “4”, “5”, “6”} – простір елементарних подій, відповідних випробуванню – однократному підкиданню кубика. Покладемо S={$\varnothing$, Ω, {“6”}, {“1”, “2”, “3”, “4”, “5”}}. Тоді стосовно S задовільняються умови 1)–3), тобто розглядувана сукупність S підмножин множини Ω може бути простором подій, а елементи сукупності S – подіями.

Зауважимо, що в останньому прикладі сукупність S можна визначити і інакше: наприклад, покласти S={$\varnothing$, Ω, {“1”}, {“2”, “3”, “4”, “5”, “6”}}, або S={$\varnothing$, Ω, {“1”, “5”}, {“2”, “3”, “4”, “6”}}, або S={$\varnothing$, Ω, {“1”, “2”, “3”}, {“4”, “5”, “6”}}, або визначити S як у прикладі 1.5.2: $S=\{A \mid A\subset \Omega\}$, і т.д.

Отже, одному і тому самому простору Ω *елементарних подій можуть бути поставлені у відповідність різні простори* S *випадкових подій. Множина* $A\subset\Omega$ *залежно від обраного простору* S *подій може бути подією, а може і не бути. Лише множини* $\varnothing$ *та* Ω *є подіями за будь-якого простору* S *випадкових подій.*

Приклад 1.5.5. Нехай $\Omega=[-r; r)$ – простір елементарних подій, через які характеризуються результати випробування: положення точки, в яку влучив спортивний снаряд, відносно цілі (якщо $x<0$ –недоліт на відстань $(-x)$; $x>0$ – переліт на відстань x; $x=0$ – влучення).

Поділимо проміжок $[-r; r)$ на k проміжків $[a_{i-1};\ a_i)$, $i\in\{1,2,...,k\}$, $a_0=-r$, $a_k=r$, $a_i=a_{i-1}+\dfrac{a_k-a_0}{k}$. Розглянемо сукупність S, до якої включено порожню множину $\varnothing$, проміжки $[a_{i-1};\ a_i)$, $i\in\overline{1,k}$, всеможливі об’єднання проміжків $[a_{i-1};\ a_i)$ із двох проміжків, із трьох, із чотирьох і т.д., із $(k-1)$ проміжків, всіх k проміжків $\bigcup\limits_{i=1}^{k}[a_{k-1};\ a_k)=\Omega$, тобто

$$S=\{A \mid A=\bigcup_{i\in I}[a_{i-1};\ a_j),\ I\subset\{1,2,3,\ldots,k\}\}\ .$$

Тоді сукупність S буде простором подій, оскільки стосовно такої сукупності задовільняються вимоги 1_s)-3_s). Цей *простір подій* називають *породженим за поділом множини* $\Omega=[-r; r)$ *на підмножини* $H_i=,\ [a_{i-1};a_i),\ i\in\overline{1,k}$. Підкреслимо, що до сукупності S включається тільки всеможливі об’єднання $\bigcup\limits_{i\in I}[a_{i-1};\ a_i)$, $I\subset\{1,2,...,k\}$, а також різні об’єднання, перетини та різниці таких об’єднань, які в свою чергу є об’єднаннями того самого типу (зокрема $\varnothing=\bigcup\limits_{i\in\varnothing}[a_{i-1};a_i)$).

Приклад 1.5.6. Нехай $\Omega=[-r; r]$ – простір елементарних подій, через які характеризуються результати експерименту: положення точки, в яку влучив спортивний снаряд, відносно цілі (якщо $x<0$ –недоліт на відстань ($-x$); $x>0$ – переліт на відстань x; $x=0$ – влучення).

Розглянемо сукупність $\tilde{S}$ підмножин $A\subset[-r; r]$, кожна з яких є об'єднанням скінченної кількості проміжків $\langle \alpha_k; \beta_k \rangle \subset \Omega$, де кожен проміжок $\langle \alpha_k : \beta_k \rangle$ може бути або відрізком $[\alpha_k; \beta_k]$, або інтервалом $(\alpha_k; \beta_k)$, або піввідрізком $[\alpha_k; \beta_k)$, або півінтервалом $(\alpha_k; \beta_k]$.

Тоді $\tilde{S}$ не є простором подій, оскільки не виконується умова 3). Проте, виявляється, що існує найменша сукупність $S \supset \tilde{S}$, яка вже може бути простором подій. Цей *простір подій* називають *породженим за сукупністю скінченних об'єднань числових проміжків*. Зокрема в S містяться всеможливі об'єднання $\bigcup_k \langle \alpha_k; \beta_k \rangle, \langle \alpha_k; \beta_k \rangle \subset [-r; r]$, а також перерізи та різниці таких об'єднань.

Приклад 1.5.7. Стосовно сукупності S підмножин виду $[-\infty; x)$ множини $\overline{R}^1=[-\infty; \infty)$, де $x \in \overline{R}^1=[-\infty; \infty)$, задовільняються вимоги 1_s)-3_s), а тому S може бути простором подій. Зокрема до такої сукупності відноситься множина

$$R^1=(-\infty; \infty)=[-\infty; \infty) \setminus \bigcap_{n=1}^{\infty}[-\infty; -\infty+\frac{1}{n}),$$

множини виду $(-\infty; x)=[-\infty; x) \setminus \bigcap_{n=1}^{\infty}[-\infty; -\infty+\frac{1}{n})$, $x \in (-\infty; \infty)$, множини виду $[x_1; x_2)=(-\infty; x_2) \setminus (-\infty; x_1)$, $x_1 \in (-\infty; \infty)$, $x_2 \in (-\infty; \infty)$, $x_1 \le x_2$, одноточкові множини виду $\{x_1\}=\bigcap_{n=1}^{\infty}[x_1; x_1+\frac{1}{n})$, будь які скінченні множини виду $\{x_1, x_2, ..., x_k\}$ як об'єднання скінченної кількості одноточкових множин $\{x_1\}, \{x_2\}, ..., \{x_k\}$, будь які відкриті підмножини множини R^1, як об'єднання не більше ніж зчисленної кількості відкритих околів раціональних точок, будь які замкнені множини як доповнення відкритих до R^1. Таку сукупність S підмножин множини $\overline{R}^1$ позначають через $\mathscr{B}(R^1)$ і називають сукупністю борелівських множин (за ім'ям французького математика Еміля Бореля).

Вправи для самостійного виконання

1.5.1. Перевірити, чи правильні твердження:

1. Існує простір Ω елементарних подій, за якого існує лише один простір S випадкових подій.

2. Якщо в просторі Ω елементарних подій є 2 елементи, то у відповідному йому просторі випадкових подій може бути 3 елементи.

3. В разі Ω={*Г, Ц*} у відповідному йому просторі S випадкових подій може бути лише 2 або 4 елементи.

4. Сукупність S довільних підмножин множини Ω=N натуральних чисел може бути простором подій.

5. Сукупність S будь-яких скінченних підмножин множини Ω=N може бути простором подій.

6. Поняття простору подій визначається через поняття події.

7. Поняття події визначається через поняття простору подій.

8. За кожного простору Ω елементарних подій існує єдиний простір випадкових подій.

9. Якщо Ω=[0; 1], а до S віднесено елементи ∅, Ω і довільні скінченні та зчисленні підмножини $A \subset \Omega$ і ніяких інших, то S є простором подій.

10. Сукупність S довільних проміжків множини Ω=(–∞; +∞) та будь-яких їх скінченних та зчисленних об’єднань може бути простором подій.

1.5.2. Побудувати можливі простори S випадкових подій, що відповідають просторам Ω елементарних подій:

1. Ω – множина можливих наслідків випробування, що полягає в підкиданні монети двічі. Як наслідки випробування розглядаються:
   а) однією стороною догори чи різними упала монета обидва рази?
   б) кількість випадань герба в обох підкиданнях;
   в) усі можливі пари випадань герба і цифри в обох підкиданнях.

2. Ω – множина можливих наслідків випробування, що полягає в підкиданні кубика, частина граней якого пофарбовані в білий колір, інша частина – в червоний, третя частина – в зелений. В кожній з частин є не менше однієї грані. Як наслідки випробування розглядаються:
   а) колір грані, що виявилася верхньою;
   б) зелений чи ні колір виявляється на верхній грані.

3. Ω – множина можливих наслідків випробування, що полягає в підкиданні двічі кубика, на гранях якого нанесені цифри від 1 до 6. Як наслідки випробування розглядаються:
   а) можливі суми $i+j,\ i \in \overline{1{,}6},\ j \in \overline{1{,}6}$,
   б) чи виконується нерівність $i \le j,\ i \in \overline{1{,}6},\ j \in \overline{1{,}6}$,

де i – цифра, що випадає на верхній грані кубика за першого підкидання, j – цифра, що випадає на верхній грані кубика за другого підкидання.

1.5.3. Визначити, скільки елементів може бути в просторі S випадкових подій, якщо у відповідному просторі Ω елементарних подій є:

а) один елемент,
б) два елементи,
в) три елементи.

1.6. Статистична ймовірність події

Нехай задано деякий простір Ω елементарних подій і один із можливих просторів S випадкових подій, пов'язаний з простором Ω елементарних подій. Розглянемо довільну подію $A \in S$. Нехай проведено серію із n випробувань, за результатами яких визначено кількість $k_n(A)$ відбувань події A.

1) Число $k_n(A)$, що дорівнює кількості тих випробувань, в яких подія A відбулася, називається *абсолютною частотою події* $A \in S$ в даній серії із n випробувань. Іншими словами, $k_n(A)$ – це кількість відбувань події A в серії із n проведених випробувань;

2) число $P_n^*(A) = \dfrac{k_n(A)}{k_n(\Omega)}$ називається *статистичною ймовірністю* або *відносною частотою події* $A \in S$ *в даній серії із* n *випробувань*.

Підкреслимо, що аргументами функцій $k_n(A)$ і $P_n^*(A)$ є підмножини простору Ω, які віднесені до сукупності S.

Приклад 1.6.1. Нехай простір елементарних подій $\Omega = \{Г, Ц\}$, простір подій $S = \{\varnothing, \{Г\}, \{Ц\}, \{Г, Ц\} = \Omega\}$, подія $A_0 = \{Г\}$ – випадання герба в результаті однократного підкидання монети.

Припустимо, що виконано $n = 100$ підкидань монети, в яких герб випадав 45 разів. Тоді $k_{100}(\{Г\}) = 45$, $k_{100}(\{Ц\}) = 100 – 45 = 55$, $k_{100}(\varnothing) = 0$, $k_{100}(\Omega) = 100$ – абсолютні частоти відповідних подій. Числа $P_{100}^*(\{Г\}) = \dfrac{45}{100} = \dfrac{9}{20}$, $P_{100}^*(\{Ц\}) = \dfrac{55}{100} = \dfrac{11}{20}$, $P_{100}^*(\varnothing) = 0$, $P_{100}^*(\Omega) = 1$ – це статистичні ймовірності (відносні частоти) відповідних подій в даній серії із $n = 100$ підкидань монети. Зокрема $k_{100}(A_0) = 45$, $P_{100}^*(A_0) = \dfrac{45}{100} = \dfrac{9}{20}$.

Приклад 1.6.2. Нехай простір елементарних подій $\Omega = \{Г, Ц\}$, простір подій $S = \{\varnothing, \Omega\}$, а в серії випробувань фіксувалося лише число $k_n(\Omega)$, а тому й $k_n(\varnothing) = k_n(\Omega) - k_n(\Omega) = n - n = 0$. В даному випадку множини $A = \{Г\}$, $B = \{Ц\}$ подіями не вважаються і тому не фіксувалися кількості їх відбувань.

Отже, про абсолютні і відносні частоти в даній серії із n випробувань можна говорити лише стосовно будь-якої події A, що віднесена до даного простору S випадкових подій.

Приклад 1.6.3. Нехай підкидається кубик і простір елементарних подій $\Omega = \{"1", "2", "3", "4", "5", "6"\}$, простір S подій – сукупність будь-яких підмножин простору Ω, а подія $A_1 = \{"3", "6"\}$ – випадання в результаті однократного підкидання кубика числа, кратного 3. Припустимо, що виконано n=15 таких підкидань і спостерігалися такі результати:

$E_{сп\,1}$="5", $E_{сп2}$="6", $E_{сп3}$="6", $E_{сп4}$="6", $E_{сп5}$="4",
$E_{сп\,6}$="3", $E_{сп7}$="5", $E_{сп8}$="2", $E_{сп9}$="6", $E_{сп10}$="4",
$E_{сп11}$="5", $E_{сп12}$="6", $E_{сп13}$="3", $E_{сп14}$="3", $E_{сп15}$="5".

Тоді абсолютні частоти $k_{15}(E_i)$ елементарних подій $E_i = "i"$, $(\{E_i\} \in S)$, $i \in \overline{1,6}$, дорівнюють:

$$k_{15}(\{E_1\}) = k_{15}(\{"1"\}) = 0, \quad k_{15}(\{E_2\}) = k_{15}(\{"2"\}) = 1,$$
$$k_{15}(\{E_3\}) = k_{15}(\{"3"\}) = 3, \quad k_{15}(\{E_4\}) = k_{15}(\{"4"\}) = 2,$$
$$k_{15}(\{E_5\}) = k_{15}(\{"5"\}) = 4, \quad k_{15}(\{E_6\}) = k_{15}(\{"6"\}) = 5.$$

Тепер можна обчислити абсолютну частоту $k_{15}(A)$ відбування довільної події $A \in S$.

Зокрема, $k_{15}(A_1) = k_{15}(\{"3","6"\}) = k_{15}(\{"3"\}) + k_{15}(\{"6"\}) = = 3 + 5 = 8$ – абсолютна частота випадання числа, кратного 3, в даній серії із n=15 підкидань кубика.

Статистична ймовірність або відносна частота події A_1 – це число $P_{15}^*(A_1) = \dfrac{k_{15}(A_1)}{k_{15}(\Omega)} = \dfrac{8}{15}$.

Зауважимо, що за результатами спостережень $P_n^*(A)$ визначене за довільного $A \subset \Omega$.

Зауважимо, що в разі, коли $A \subset \Omega$, $B \subset \Omega$, $A \subset B$, тоді всі спостережені елементарні події, які є елементами множини A, є також і елементами множини B, тому $P_n^*(A) \le P_n^*(B)$ за довільних $A \subset \Omega$ і $B \subset \Omega$ таких, що $A \subset B$.

Вправи для самостійного виконання

1.6.1. Перевірити, чи правильні твердження:

1. Число $k_n(A)$ може набувати будь-якого значення з проміжка $[0; n]$.

2. Статистична ймовірність (відносна частота) $P_n^*(A)$ події A може набувати будь-якого значення з проміжка $[0; 1]$.

3. Якщо $P_n^*(A) = 1$, то A – вірогідна подія.

4. Якщо $P_n^*(A) = 0$, то A – неможлива подія.

1.6.2. 1. Монету підкидали 10 разів. Якими можуть бути абсолютні і відносні частоти випадань герба і цифри за результатами цих 10 підкидань?

2. Кубик, частина граней якого пофарбовані в білий колір, друга частина – в червоний колір, третя частина – в зелений колір, підкидали 5 разів. Якими можуть бути абсолютні і відносні частоти випадань на верхній грані кожного з трьох кольорів за результатами цих 5 підкидань?

3. Грані кубика пофарбовані в білий, червоний, зелений, блакитний, жовтий, бузковий кольори, причому на кожній грані є лише один колір. Випробування полягає в підкиданні кубика і фіксації кольору грані, якою кубик падає догори. Якими можуть бути статистичні ймовірності випадання грані кожного кольору, якщо кубик підкидали: 1 раз? 2 рази? 3 рази? 4 рази? 5 разів? 6 разів? 7 разів? 8 разів? 9 разів? 10 разів?

4. Кубик, частина граней якого пофарбована в білий колір, друга частина – в червоний колір, третя частина – в зелений колір, підкидали 1000 разів. Як результати випробувань розглядали колір верхньої грані. Виявилося, що в 1000 випробуваннях білою гранню догори кубик падав 500 разів, червоною – 300, зеленою – 200. Побудувати всі можливі простори подій, відповідні даному експерименту, і визначити статистичні імовірності всіх подій із кожного з просторів, виходячи з зазначених даних.

1.7. Властивості статистичної ймовірності

Безпосередньо з означення випливають *основні властивості статистичної ймовірності:*

$1_р$. $P_n^*(A) \ge 0,\ A \in S$, тобто статистична ймовірність довільної події $A \in S$ невід'ємна;

$2_р$. $P_n^*\left(\sum_{i\in I} A_i\right) = \sum_{i\in I} P_n^*(A_i)$, $I \subset \{1,2,3,...\}$, коли $A_i \in S$ і $A_i A_j = \varnothing$, коли $i \ne j$ (події A_i попарно несумісні), тобто статистична ймовірність суми довільної кількості попарно несумісних подій дорівнює сумі статистичних ймовірностей цих подій. Це так звана *властивість повної* (або *зчисленної*) *адитивності* статистичної ймовірності.

$3_р$. $P_n^*(\Omega) = 1$, тобто статистична ймовірність вірогідної події дорівнює 1.

Властивості $1_р$-$3_р$ називають *основними або визначальними*. З них випливають інші важливі властивості статистичної ймовірності.

Наведемо деякі з цих властивостей:

4. $P_n^*(\overline{A}) = 1 - P_n^*(A)$. Справді, $A \cdot \overline{A} = \varnothing$ і $A + \overline{A} = \Omega$. Тому за основними властивостями $2_р$ і $3_р$:

$$P_n^*(A + \overline{A}) = P_n^*(A) + P_n^*(\overline{A}) = P_n^*(\Omega) = 1.$$

Звідси $P_n^*(\overline{A}) = 1 - P_n^*(A)$.

5. $P_n^*(\varnothing) = 0$, тобто статистична ймовірність неможливої події дорівнює нулю.

Справді, оскільки $\varnothing = \overline{\Omega}$, то за властивостями 4 і $3_р$

$$P_n^*(\varnothing) = P_n^*(\overline{\Omega}) = 1 - P_n^*(\overline{\Omega}) = 1 - 1 = 0.$$

6. Якщо $A \subset B$, то $P_n^*(B) = P_n^*(A) + P_n^*(B \setminus A)$, а тому $P_n^*(A) \le P_n^*(B)$. Зокрема, $P_n^*(A) \le 1$ за будь-якої події A.

Справді, якщо $A \subset B$, то $B = A + (B \setminus A)$ і $A \cdot (B \setminus A) = \varnothing$. Тому за основними властивостями $2_р$ і $1_р$:

$$P_n^*(B) = P_n^*(A) + P_n^*(B \setminus A) \ge P_n^*(A).$$

Зокрема, якщо $B = \Omega$, то $A \subset \Omega$, і тому $P_n^*(A) \le P_n^*(\Omega) = 1$ за будь-якої події A.

7. Враховуючи властивості $1_р$ та 6, дістанемо $0 \le P_n^*(A) \le 1$, тобто статистична ймовірність довільної події може набувати значення лише з відрізка $[0;1]$. (Ця властивість також випливає з означення $P_n^*(A)$).

8. Стосовно довільних подій $A \in S$, $B \in S$ правильна рівність

$$P_n^*(A+B) = P_n^*(A) + P_n^*(B) - P_n^*(AB).$$

Справді, подію $A+B$ можна подати у вигляді суми $A+(B \setminus AB)$ двох несумісних доданків (Рис. 1.7.1.). Подію B також можна подати у вигляді суми $AB+(B \setminus AB)$ двох несумісних доданків (Рис. 1.7.2.).

Тому за властивістю $2_р$:

$$P_n^*(A+B) = P_n^*(A) + P_n^*(B \setminus AB), \tag{1.7.1}$$

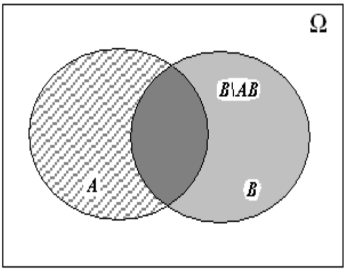
$$P_n^*(B) = P_n^*(AB) + P_n^*(B \setminus AB).$$

З останньої рівності випливає

$$P_n^*(B \setminus AB) = P_n^*(B) - P_n^*(AB),$$

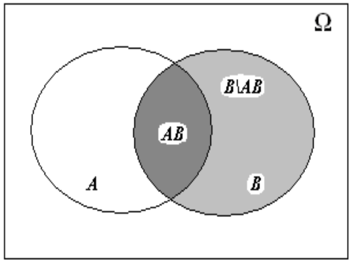
а тому, враховуючи рівність (1.7.1), дістаємо:

$$P_n^*(A+B) = P_n^*(A) + P_n^*(B) - P_n^*(AB).$$



$A+B=A+(B \setminus AB)$

Рис. 1.7.1



$B=AB+(B \setminus AB)$

Рис. 1.7.2

Вправи для самостійного виконання

1.7.1. Перевірити, чи правильні твердження:

1. Якщо $P_n^*(A+B) = P_n^*(A) + P_n^*(B)$, то A і B несумісні події.
2. Якщо $P_n^*(A) + P_n^*(B) = 1,$ то A і B протилежні події.
3. Якщо $P_n^*(A+B) \neq P_n^*(A) + P_n^*(B),$ то A і B сумісні події.
4. Якщо через подію A спричинюється подія $B,$ то $P_n^*(A) < P_n^*(B)$.
5. Якщо $P_n^*(A) < P_n^*(B)$, то через подію A спричинюється подія B.
6. За будь-яких подій $A \in S$, $B \in S$ і $C \in S$:

$$P_n^*(A+B+C) = P_n^*(A) + P_n^*(B) + P_n^*(C) -$$
$$- P_n^*(A \cdot B) - P_n^*(AC) - P_n^*(BC) + P_n^*(ABC).$$

7. Якщо $A \subset B,$ то $P_n^*(B \setminus A) = P_n^*(B) - P_n^*(A).$

1.7.2. 1. Сформулювати і довести властивості абсолютної частоти, аналогічні до властивостей статистичної ймовірності.

1.8. Ймовірнісні простори. Уточнення поняття випадкової події

Нехай Ω – деякий простір елементарних подій, і задана деяка сукупність S підмножин множини Ω, стосовно якої задовільняються вимоги:

1_s. $\Omega \in S$;

2_s. Якщо $A \in S$, то і $\overline{A} \in S$;

3_s. Якщо $A_i \in S$, $i \in \{1, 2, 3, \ldots\}$, то і $\bigcup\limits_{i \in I} A_i \in S$, $I \subset \{1, 2, 3, \ldots\}$.

Будь-яку *задану на* S *числову функцію* $V(A)$, $A \in S$, стосовно якої задовільняються вимоги:

1_v. $V(A) \geq 0$, $A \in S$;

2_v. Якщо $A_i \in S$, $i \in \{1, 2, 3, \ldots\}$, $A_i A_j = \varnothing$, коли $i \neq j$, то

$$V\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} V(A_i)\,,\; I \subset \{1, 2, 3, \ldots\}\,,$$

називають *мірою, заданою на сукупності* S *підмножин множини* Ω .

Таку сукупність S підмножини множини Ω називають *простором вимірних за мірою* V *множин*, кожну множину $A \in S$ називають *вимірною за мірою* V , а значення $V(A)$ називають *мірою множини* A . Якщо ж $A \subset \Omega$, але $A \overline{\in} S$, то *таку множину* A *називають невимірною відносно міри* V , *заданої на* S, оскільки міра визначена лише стосовно множин, що віднесені до сукупності S .

Прикладами мір можуть бути об'єми вимірних просторових фігур, площі вимірних плоских фігур, довжини вимірних лінійних фігур (зокрема об'єми многогранників та їх об'єднань, площі многокутників та їх об'єднань, довжини відрізків та їх об'єднань), кількості елементів в підмножинах даної скінченної множини, кількості попадань у певні підмножини деякої множини, з якої елементи вибираються навмання, маси фізичних тіл та їх об'єднань, тощо.

Якщо $V(\Omega) = 1$, тоді міру V називають *ймовірнісною мірою* або *ймовірністю, визначеною на сукупності* S *підмножин множини* Ω .

Таким чином, стосовно ймовірнісної міри (ймовірності) $P(A)$, визначеної на сукупності S підмножин множини Ω , повинні задовільнятися вимоги:

1_p. $P(A) \geq 0$, $A \in S$;

2_p. Якщо $A_i \in S$, $i \in \{1, 2, 3, \ldots\}$, $A_i A_j = \varnothing$, коли $i \neq j$, то

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)\,,\; I \subset \{1, 2, 3, \ldots\}\,.$$

3_p. $P(\Omega) = 1$.

Якщо задано простір елементарних подій Ω , сукупність S підмножин множини Ω , стосовно якої задовільняються вимоги

1_s–3_s, і на цій сукупності визначена числова функція $P(A)$, стосовно якої задовільняються вимоги 1_p–3_p, тоді говорять, що задано *ймовірнісний простір* (Ω, S, P).

Елементи множини Ω *із простору* (Ω, S, P) *називаються елементарними подіями, елементи сукупності* S *– подіями, а числа* $P(A)$, $A \in S$, *– ймовірностями подій* $A \in S$.

Вимоги 1_p-3_p стосовно функції $P(A)$, $A \in S$, називаються *основними* чи *визначальними* властивостями ймовірнісної міри. З них випливають і інші важливі властивості ймовірнісної міри $P(A)$.

Властивість 2_p називають *властивістю повної аддитивності функції* $P(A)$.

Будь-яку числову функцію $P(A)$, $A \in S$, *що задана на просторі подій* S *і стосовно якої задовільняються вимоги* 1_p-3_p, *називають ймовірнісною мірою (або просто ймовірністю) подій* $A \in S$.

Вимоги (властивості) 1_s-3_s, 1_p-3_p називають *системою аксіом теорії ймовірностей* (системою аксіом А.М. Колмогорова).

Зауважимо, що коли стосовно міри $V(A)$, $A \in S$, задовільняються вимоги $0 < V(\Omega) < \infty$, тоді функція $\tilde{V}(A)$ від множин $A \in S$, що визначається за рівністю

$$\tilde{V}(A) = \frac{V(A)}{V(\Omega)}, \ A \in S,$$

буде ймовірнісною мірою, оскільки стосовно неї задовільняються вимоги 1_p-3_p. Зокрема $\tilde{V}(\Omega) = 1$.

Таку функцію $\tilde{V}(A)$ від множин $A \in S$ називають *нормованою мірою, що отримується нормуванням міри* $V(A)$, *заданої на* S.

Легко бачити, що статистична ймовірність (відносна частота) $P_n^*(A)$, $A \in S$, визначена на сукупності S підмножин множини Ω, є ймовірнісною мірою (оскільки стосовно неї задовільняються вимоги 1_p-3_p).

Зауважимо, що $P_n^*(A)$, $A \in S$, є *нормованою мірою, одержаною нормуванням міри* $k_n(A)$, заданої на S:

$$P_n^*(A) = \frac{k_n(A)}{k_n(\Omega)}, \ A \in S.$$

Причому стосовно міри $k_n(A)$, $A \in S$, задовільняються вимоги $0 \le k_n(A) < \infty$ за всіх $A \in S$, зокрема $1 \le k_n(\Omega) = n < \infty$, $P_n^*(\Omega) = \dfrac{k_n(\Omega)}{k_n(\Omega)} = 1$.

Зауважимо також, що $P_n^*(A)=\frac{k_n(A)}{k_n(\Omega)}$, $A \in S$, є часткою від кількості $k_n(\Omega)=n \geq 1$ точок, вибраних із множини Ω в результаті серії із n проведених випробувань, яка припадає на множину $A \subset \Omega$, $A \in S$, тобто $k_n(\Omega)\cdot P_n^*(A)=k_n(A)$, $A \subset \Omega$, $A \in S$.

Очевидно, $P_n^*(A)=\frac{k_n(A)}{k_n(\Omega)}$ може набувати лише значень $0=\frac{0}{n},\frac{1}{n},\frac{2}{n},\ldots,\frac{n}{n}=1$ в межах від 0 до 1 – найменшого, коли серед точок, вибраних із множини Ω в результаті серії із n випробувань, $n \geq 1$, $k_n(\Omega)=n$, не виявилося жодної, яка належить до множини *A*, тобто $k_n(A)=0$, $P_n^*(A)=\frac{k_n(A)}{k_n(\Omega)}=\frac{0}{n}=0$, і найбільшого, коли серед точок, вибраних із множини Ω в серії із *n* випробувань, $n \geq 1$, $k_n(\Omega)=n$, всі вибрані точки належать до множини *A*, тобто $k_n(A)=n$, $P_n^*(A)=\frac{k_n(A)}{k_n(\Omega)}=\frac{n}{n}=1$.

Якщо, наприклад, проведено лише одне випробування, тобто $n=1$, і вибрано точку $E \in A$, $A \subset \Omega$, $A \in S$, тоді $k_n(\Omega)=1$, $k_n(A)=1$, $P_n^*(A)=\frac{k_n(A)}{k_n(\Omega)}=\frac{1}{1}=1$, $n=1$.

Якщо крім того розглядається інша подія $B \subset \Omega$, $B \in S$, і вибрана точка *E* не є елементом множини *B*, тобто $E \overline{\in} B$, а значить $E \in \overline{B}$, тоді $k_n(B)=0$, $k_n(\overline{B})=1$, $P_n^*(B)=\frac{k_n(B)}{k_n(\Omega)}=\frac{0}{1}=0$, $P_n^*(\overline{B})=\frac{k_n(\overline{B})}{k_n(\Omega)}=\frac{1}{1}=1$, $n=1$.

На практиці як правило проводять досить довгу серію випробувань для того, щоб за її результатами можна було робити більш або менш точні і надійні прогнози на майбутнє стосовно відбування деякої події *A*. Однак за теоретичних досліджень властивостей статистичної ймовірності $P_n^*(A)$ як ймовірнісної міри, заданої на деякій сукупності *S* підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, довжина серії випробувань не має ніякого значення, важливо лише, щоб задовільнялась вимога $1 \leq n < \infty$.

Приклад 1.8.1. Нехай дуже велику кількість разів проводили випробування – підкидання кубика, в результаті кожного з яких навмання вибирався один елемент із множини $\Omega=\{"1","2","3","4","5","6"\}$. Нехай, разом з тим, повідомлено лише, що за результатами цих випробувань виявилося

$P_n^*(\{"6"\}) = 0.60$, $P_n^*(\{"5"\}) = 0.30$. Тоді за наведеними даними можна побудувати один із можливих просторів подій, стосовно всіх елементів якого буде визначено статистичну ймовірність (ймовірнісну міру), слідуючим чином. Розглянемо підмножину множини Ω, яка є доповненням множини $\{"5"\} \bigcup \{"6"\}$ до множини Ω, тобто множину

$$\Omega \setminus (\{"5"\} \bigcup \{"6"\}) = \{"1","2","3","4"\}.$$

Введемо позначення: $H_1 = \{"1","2","3","4"\}$, $H_2 = \{"5"\}$, $H_3 = \{"6"\}$. Очевидно $H_i H_j = \varnothing$, коли $i \neq j$, $H_1 + H_2 + H_3 = \Omega$, тобто з підмножин H_1, H_2, H_3 утворено поділ множини Ω на підмножини, що попарно не перетинаються.

Розглянемо таку сукупність S підмножин множини $\Omega = \{"1","2","3","4","5","6"\}$:

$$S = \{\varnothing,\ H_1,\ H_2,\ H_3,\ H_1 + H_2,\ H_1 + H_3,\ H_2 + H_3,\ H_1 + H_2 + H_3 = \Omega\} =$$

$$= \{\varnothing,\ \{"1","2","3","4"\},\ \{"5"\},\ \{"6"\},\ \{"1","2","3","4","5"\},$$

$$\{"1","2","3","4","6"\},\ \{"5","6"\}, \{"1","2","3","4","5","6"\} = \Omega\}.$$

Очевидно, стосовно цієї сукупності S задовільняються вимоги 1_s-3_s. В подібних випадках говорять, що сукупність S породжена за поділом множини Ω на підмножини H_1, H_2, H_3. Очевидно також, що $P_n^*(A)$ визначена стосовно всіх елементів сукупності S:

$$P_n^*(\varnothing) = 0;\ P_n^*(\{"1","2","3","4"\}) = P_n^*(\overline{\{"5"\} \bigcup \{"6"\}}) =$$

$$= 1 - P_n^*(\{"5"\} \bigcup \{"6"\}) = 1 - (0.30 + 0.60) = 0.10;\ P_n^*(\{"5"\}) = 0.30;$$

$$P_n^*(\{"6"\}) = 0.60;\ P_n^*(\{"1","2","3","4","5"\}) = 0.40;$$

$$P_n^*(\{"1","2","3","4","6"\}) = 0.70;\ P_n^*(\{"5","6"\}) = 0.90;$$

$$P_n^*(\{"1","2","3","4","5","6"\}) = 1,$$

і стосовно неї задовольняються вимоги 1_p-3_p.

Тому в розглянутому випадку сукупність S є простором подій (вимірних за мірою $P_n^*(A)$ підмножин множини Ω), а трійка $(\Omega,\ S,\ P_n^*)$ є ймовірнісним простором.

Разом з тим підмножини, наприклад, $\{"1","2"\}$, $\{"1","3","5"\}$, $\{"2","4","6"\}$, і деякі інші підмножини множини $\Omega = \{"1","2","3","4","5","6"\}$, зокрема будь які підмножини множини $H_1 = \{1, 2, 3, 4\}$, відмінні від $\varnothing$ і від H_1, виявляються невимірними відносно розглянутої міри $P_n^*(A)$, $A \in S$, оскільки за наведеними даними неможливо визначити статистичну ймовірність (відносну частоту) попадання в такі підмножини в проведеній серії із n випробувань, і тому такі підмножини не включені до простору подій S, вони не вважаються подіями.

Приклад 1.8.2. Нехай Ω – скінченна множина точок x_i з відрізка $[0;1]$, причому $x_i = x_{i-1} + h$, $x_0 = 0$, $h = 10^{-1000000}$, $i \in \{1, 2, \ldots, 10^{1000000}\}$. Тоді (як і у випадку неперервної множини Ω типу $\Omega = [a;\, b))$, на практиці швидше за все недоцільно розглядати всі підмножини розглянутої скінченної множини Ω. В подібних випадках доцільно множину Ω поділити на деяку практично прийнятну кількість k підмножин H_i таких, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup\limits_{i=1}^{k} H_i = \Omega$, і як події разом з $\varnothing$ розглядати лише всеможливі об'єднання підмножин H_i, тобто покласти $S = \{A \mid\ A = \bigcup\limits_{i \in I} H_i,\ I \subset \{1, 2, \ldots, k\}\}$. Визначивши статистичні ймовірності $P_n^*(H_i)$, $i \in \{1, 2, \ldots, k\}$, стосовно довільної події $A = \bigcup\limits_{i \in I} H_i \in S$, $I \subset \{1,2,\ldots,k\}$, одержимо $P_n^*(A) = \sum\limits_{i \in I} P_n^*(H_i)$.

До так побудованої сукупності S подій (підмножин множини Ω) окрім неможливої події $\varnothing$, можуть бути віднесені всі події H_i, всі суми виду $H_i + H_j$ із двох подій, $i \in \{1, 2, \ldots, k\}$, $j \in \{1, 2, \ldots, k\}$, $i \neq j$, всі суми виду $H_i + H_j + H_l$ із трьох подій, $i \in \{1, 2, \ldots, k\}$, $j \in \{1, 2, \ldots, k\}$, $l \in \{1, 2, \ldots, k\}$, $i \neq j \neq l$, всі суми із чотирьох подій H_i, всі суми із п'яти подій H_i, і т.д., всі суми із $(k-1)$ подій H_i, сума всіх k подій H_i, тобто $H_1 + H_2 + \ldots + H_k = \Omega$. Очевидно стосовно так побудованої (породженої за поділом множини Ω на підмножини H_1, H_2, …, H_k) сукупності S підмножин множини Ω задовільняються вимоги 1_s-3_s, а стосовно так введеної ймовірнісної міри $P_n^*(A)$ задовільняються вимоги 1_p-3_p. В такий спосіб побудовано ймовірнісний простір $(\Omega,\ S,\ P_n^*)$.

Якщо події H_i розглядати як єдино можливі наслідки випробування, тоді фактично вводиться новий простір $\tilde{\Omega}$ елементарних подій $\tilde{\Omega} = \{H_1, H_2, \ldots, H_k\}$ і відповідно новий ймовірнісний простір $(\tilde{\Omega},\ S,\ P_n^*)$.

В розглянутому прикладі множину Ω можна поділити, наприклад, на 10 підмножин H_i в такий спосіб:

$$H_1 = \{x_i \mid 0 \le x_i < 0.1\},\quad H_2 = \{x_i \mid 0.1 \le x_i < 0.2\},$$
$$H_3 = \{x_i \mid 0.2 \le x_i < 0.3\},\quad H_4 = \{x_i \mid 0.3 \le x_i < 0.4\},$$
$$H_5 = \{x_i \mid 0.4 \le x_i < 0.5\},\quad H_6 = \{x_i \mid 0.5 \le x_i < 0.6\},$$
$$H_7 = \{x_i \mid 0.6 \le x_i < 0.7\},\quad H_8 = \{x_i \mid 0.7 \le x_i < 0.8\},$$
$$H_9 = \{x_i \mid 0.8 \le x_i < 0.9\},\quad H_{10} = \{x_i \mid 0.9 \le x_i \le 1.0\}.$$

Можливі і інші поділи множини Ω на підмножини, які попарно не перетинаються. Наприклад, $\tilde{H}_1 = H_1 + H_2$, $\tilde{H}_2 = H_3 + H_4$, $\tilde{H}_3 = H_5 + H_6$, $\tilde{H}_4 = H_7 + H_8$, $\tilde{H}_5 = H_9 + H_{10}$ і т.п.

Вправи для самостійного виконання

1.8.1. Перевірити, чи правильні твердження:

1. В понятті міри множини узагальнюється поняття кількості елементів в множині, довжини лінійної множини, площі плоскої фігури, об'єму тіла, маси тіла.
2. Міру можна задати на будь-якій сукупності множин.
3. Абсолютна частота $k_n(A)$ відбування події A, $A \in S$, є мірою, заданою на S.
4. Статистична ймовірність є ймовірнісною мірою.
5. Кожна ймовірнісна міра є статистичною ймовірністю.
6. Для того, щоб задати ймовірнісну міру $P_n^*(A),\ A \in S$, потрібно спочатку задати простір подій S.
7. Для того, щоб задати простір подій S, потрібно спочатку задати ймовірнісну міру на деяких підмножинах $A \subset \Omega$.

1.8.2. Нехай $\Omega = \{"1","2","3","4","5","6"\}$, і відомі відносні частоти $P_n^*(\{"5","6"\}) = 0.70;\ P_n^*(\{"1","2","3","4"\}) = 0.30$. Чи буде трійка $(\Omega,\ S,\ P_n^*)$ ймовірнісним простором, якщо:

а) $S = \{\varnothing, \{"1","2","3","4"\}, \{"5","6"\}, \Omega\}$?

б) $S = \{\varnothing, \Omega\}$?

в) $S = \{\varnothing, \{"5","6"\}, \Omega\}$?

г) $S = \{\varnothing, \{"1","2"\}, \{"3","4","5","6"\}, \Omega\}$?

д) $S = \{\{"1"\}, \{"2"\}\}$?

1.8.3. Нехай задано деяку множину Ω елементарних подій і ймовірнісну міру $P_n^*(A)$ стосовно деякого $A \subset \Omega$. Побудувати ймовірнісний простір (Ω, S, P_n^*), за якого $A \in S$. Скільки просторів подій можна побудувати за заданих умов?

1.8.4. Нехай задано деяку множину Ω і ймовірнісні міри $P_n^*(H_1)$, $P_n^*(H_2)$, $P_n^*(H_3)$ деяких підмножин H_1, H_2, H_3 множини Ω, таких, що $H_i H_j = \varnothing$, коли $i \neq j$. Побудувати всі можливі ймовірнісні простори (Ω, S, P_n^*) за заданих умов.

1.9. Побудова ймовірнісного простору

Будуючи ймовірнісний простір, слід мати на увазі, що трійку (Ω, S, P) можна вважати ймовірнісним простором, якщо *множина* Ω, *сукупність* S *підмножин множини* Ω *і ймовірнісна міра* $P(A)$ *узгоджені*.

Це означає, що функція $P(A)$ повинна бути визначена за кожного $A \in S$, стосовно сукупності S підмножин множини Ω повинні задовільнятися вимоги 1_s-3_s, а стосовно функції $P(A)$ – вимоги 1_p-3_p. Якщо ж хоч одна із зазначених вимог не задовільняється, тоді слід продовжити дії стосовно побудови ймовірнісного простору.

Наприклад, якщо функція $P(A)$ визначена стосовно не всіх $A \in S$, то тоді необхідно або довизначити функцію $P(A)$ стосовно всіх $A \in S$ (продовжити міру $P(A)$ на всю сукупність S), або ж вилучити із сукупності S всі її елементи $A \in S$, стосовно яких $P(A)$ не визначена, і залишивши лише такі $A \in S$, стосовно яких $P(A)$ визначена, побудувати на їх основі нову сукупність S_1 таку, стосовно якої будуть задовільнятися вимоги 1_s-3_s, і крім того $P(A)$ буде визначена стосовно всіх елементів $A \in S_1$ і будуть задовільнятися вимоги 1_p-3_p.

Множини $A \subset \Omega$, стосовно яких міра $P(A)$ не визначена, називаються *невимірними* відносно міри $P(A)$. Відповідно множини $A \subset \Omega$, стосовно яких $P(A)$ визначена, називаються *вимірними* відносно міри $P(A)$.

Невимірні (*відносно міри* $P(A)$) *підмножини* $A \subset \Omega$ *множини* Ω *не розглядаються як події* (не вважаються подіями), *а до простору подій* S *включаються лише вимірні* (*відносно міри* $P(A)$) *підмножини множини* Ω.

Приклад 1.9.1. Нехай Ω ={“1”, “2”, “3”, “4”, “5”, “6”} і нехай $H_1 = \{"1","2","3"\}$, $H_2 = \{"4"\}$, $H_3 = \{"5"\}$, $H_4 = \{"6"\}$. Тоді $H_i H_j = \varnothing$, коли $i \neq j$, $H_1 + H_2 + H_3 + H_4 = \Omega$.

Розглянемо таку сукупність S_1 підмножин множини Ω:

$$S_1 = \{\varnothing,\ H_1,\ H_2,\ H_3,\ H_4,\ H_1+H_2,\ H_1+H_3,\ H_1+H_4,\ H_2+H_3,\ H_2+H_4,\ H_3+H_4,$$
$$H_1+H_2+H_3,\ H_1+H_2+H_4,\ H_1+H_3+H_4,\ H_2+H_3+H_4,\ H_1+H_2+H_3+H_4=\Omega\} =$$
$$= \Big\{ \varnothing,\ \{"1","2","3"\},\ \{"4"\},\ \{"5"\},\ \{"6"\},\ \{"1","2","3","4"\},\ \{"1","2","3","5"\},$$
$$\{"1","2","3","6"\},\ \{"4","5"\},\ \{"4","6"\},\ \{"5","6"\},\ \{"1","2","3","4","5"\},$$
$$\{"1","2","3","4","6"\},\quad \{"1","2","3","5","6"\},\quad \{"4","5","6"\},$$
$$\{"1","2","3","4","5","6"\} \Big\}.$$

Очевидно стосовно сукупності S_1 (породженої за поділом множини Ω на підмножини H_1, H_2, H_3, H_4) задовільняються вимоги 1_s-3_s. Разом з тим, якщо задано лише $P_n^*(\{"6"\}) = 0.6$, $P_n^*(\{"5"\}) = 0.3$, тоді можна однозначно визначити $P_n^*(A)$ стосовно A із сукупності подій

$$S = \{\varnothing, \{"1","2","3","4"\}, \{"5"\}, \{"6"\}, \{"1","2","3","4","5"\},$$
$$\{"1","2","3","4","6"\}, \{"5","6"\}, \{"1","2","3","4","5","6"\}\}.$$

Таким чином P_n^* і S узгоджені. Разом з тим, ймовірнісна міра $P_n^*(A)$ виявляється неузгодженою із сукупністю S_1, оскільки неможливо однозначно визначити статистичні ймовірності попадання в підмножини $\{"1","2","3"\}, \{"4"\}, \{"1","2","3","5"\}$, $\{"4","5"\}$, $\{"1","2","3","6"\}$, $\{"4","6"\}$, $\{"1","2","3","5","6"\}$, ці підмножини виявляються невимірними відносно так заданої ймовірнісної міри. Якщо разом з тим в доповнення до попереднього повідомлено, що відносна частота попадання в множину $\{"4"\}$ дорівнює, наприклад, 0.06, а в множину $\{"1","2","3"\}$ – 0.04, то тоді одержуємо нову функцію $\tilde{P}_n^*(A)$, $A \in S_1$, визначену стосовно всіх елементів із S_1. Зокрема,

$$\tilde{P}_n^*(\{"1","2","3"\}) = 0.04 \;\; \tilde{P}_n^*(\{"4"\}) = 0.06,$$
$$\tilde{P}_n^*(\{"1","2","3","5"\}) = 0.34, \;\; \tilde{P}_n^*(\{"4","5"\}) = 0.36,$$
$$\tilde{P}_n^*(\{"1","2","3","6"\}) = 0.64, \;\; \tilde{P}_n^*(\{"4","6"\}) = 0.66,$$
$$\tilde{P}_n^*(\{"4","5","6"\}) = 0.96,$$

а стосовно елементів сукупності $S \subset S_1$

$$\tilde{P}_n^*(A) = P_n^*(A), \;\; A \in S.$$

Таким чином ймовірнісна міра $\tilde{P}_n^*(A)$ є продовженням ймовірнісної міри $P_n^*(A)$ із сукупності S на сукупність $S_1 \supset S$.

Разом з тим, якщо немає ніяких коректних міркувань щодо того, як слід продовжувати міру $P_n^*(A)$ із сукупності S на сукупність S_1, тоді слід із сукупності S_1 вилучити всі підмножини множини Ω, стосовно яких за наявними даними неможливо визначити міру $P_n^*(A)$, для того, щоб мати можливість побудувати сукупність S вимірних відносно міри $P_n^*(A)$ підмножин множини Ω, стосовно якої будуть задовільнятися вимоги 1_s-3_s, і в такий спосіб побудувати ймовірнісний простір (Ω, S, P_n^*).

Наприклад, якщо задані міри $P_n^*(H_i)$ множин H_1, H_2, …, H_k таких, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$, тоді можна побудувати сукупність S вимірних за мірою P_n^* підмножин множини Ω, включивши до S разом з $\varnothing$ всеможливі об'єднання множин H_i із одного, із двох, із трьох і т.д., із $k-1$, із всіх k доданків, тобто покласти $S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1, 2, 3, ..., k\}\}$.

Очевидно, стосовно такої сукупності S (породженої за поділом множини Ω на підмножини H_1, H_2, …, H_k) задовільнятимуться вимоги 1_s-3_s і стосовно всіх елементів такої сукупності буде визначена міра P_n^*.

Слід підкреслити, що за заданих в прикладі 1.9.1 умов існує багато різних продовжень міри P із сукупності S на сукупність S_1.

Зауважимо, що коли множина Ω елементарних подій скінченна, $\Omega = \{E_1, E_2, ..., E_k\}$, і визначені статистичні ймовірності $P_n^*(\{E_i\})$, $i \in \{1, 2, ..., k\}$, то тоді числа $P_n^*(A) = \sum_{E_i \in A} P_n^*(\{E_i\})$ виявляються визначеними за будь якого $A \subset \Omega$.

Зауважимо також, що до сукупності S не обов'язково включати всі множини H_i, всі суми $H_i + H_j$, $i \neq j$, із двох доданків, всі суми із трьох доданків і т.д. Важливо лише, щоб задовільнялися вимоги 1_s-3_s стосовно сукупності S, і щоб за кожного $A \in S$ була визначена ймовірнісна міра $P_n^*(A)$.

Приклад 1.9.2. Нехай стосовно деяких $A \subset \Omega$, $B \subset \Omega$, $AB = \varnothing$, задано $P_n^*(A)$ і $P_n^*(B)$. Тоді оскільки

$$\overline{A+B} = \overline{A}\,\overline{B},\ A + B + \overline{A}\,\overline{B} = \Omega,\ A \cdot \overline{A}\,\overline{B} = \varnothing,\ B \cdot \overline{A}\,\overline{B} = \varnothing,$$

$$P_n^*(\overline{A}\,\overline{B}) = 1 - P_n^*(A+B) = 1 - P_n^*(A) - P_n^*(B),$$

можна побудувати сукупність

$$S = \{\varnothing, A, B, \overline{A}\,\overline{B}, A+B, A+\overline{A}\,\overline{B}, B+\overline{A}\,\overline{B}, A+B+\overline{A}\,\overline{B} = \Omega\}$$

підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, і в такий спосіб буде отримано ймовірнісний простір (Ω, S, P_n^*).

Це означає, що ймовірнісний простір (Ω, S, P_n^*) можна побудувати і тоді, коли задано $P_n^*(H_i)$ стосовно деякої кількості k підмножин H_i із множини Ω таких, що $H_i H_j = \varnothing$, коли

$i \neq j$, але не обов'язково $\bigcup_{i=1}^{k} H_i = \Omega$, бо легко визначити ймовірнісну міру $P_n^*\left(\overline{\bigcup_{i=1}^{k} H_i}\right) = 1 - \sum_{i=1}^{k} P_n^*(H_i)$ доповнення множини $\bigcup_{i=1}^{k} H_i$ до множини Ω, тобто отримати поділ множини Ω на підмножини H_1, H_2, …, H_k, $\overline{\bigcup_{i=1}^{k} H_i}$, які не перетинаються, і стосовно всіх із яких визначено ймовірнісну міру P_n^*. На основі цього поділу легко побудувати деяку сукупність S підмножин множини Ω, стосовно якої задовільнятимуться вимоги 1_S-3_S і стосовно всіх елементів якої буде визначена ймовірнісна міра P_n^*.

Таким чином, остаточно сказати, які саме підмножини A множини Ω *вважаються подіями, можна лише після того, як побудовано ймовірнісний простір* $(\Omega,\ S,\ P_n^*)$.

Вправи для самостійного виконання

1.9.1. Перевірити, чи правильні твердження:

1. Якщо сукупність S підмножин $A \subset \Omega$ входить в простір подій S_1, тобто $S \subset S_1$, і на S_1 задана ймовірнісна міра $P_n^*(A)$, $A \in S_1$, то $(\Omega,\ S,\ P_n^*)$ ймовірнісний простір.

2. Якщо $(\Omega,\ S,\ P_n^*)$ – ймовірнісний простір, $S \subset S_1$, і стосовно S_1 задовільняються вимоги 1_s-3_s, то міру $P_n^*(A)$ можна продовжити із сукупності S на сукупність S_1:

а) безліччю способів; б) єдиним способом.

1.9.2. 1. Задано $\Omega = \{$“1”,“2”,“3”,“4”,“5”,“6”$\}$ і $P_n^*(\{$“1”,“3”,“5”$\})=0.5$, $P_n^*(\{$“1”,“4”,“6”$\})=0.5$. Чи можливо за заданих умов побудувати якийсь ймовірнісний простір? Якщо так, то скільки і які саме?

2. Задано простір елементарних подій Ω, підмножини $A \subset \Omega$, $B \subset \Omega$, $C \subset \Omega$, такі що $B \subset A$, $C \subset A$, $BC = \varnothing$. Крім того задано $P_n^*(A)$, $P_n^*(B)$, $P_n^*(C)$. Побудувати всі можливі ймовірнісні простори за заданих умов.

1.10. Умовна статистична ймовірність. Статистична ймовірність добутку подій.

Нехай Ω – простір елементарних подій, S – деякий простір подій, відповідний Ω.

Нехай проведено n випробувань, в яких спостерігалися події $A \in S$ і $B \in S$, і нехай $m_B \neq 0$ – число випробувань, в яких відбулася подія B. Серед цих m_B випробувань подія A відбувається тоді, коли відбуваються елементарні події, які є як в множині A, так і в множині B, тобто в множині AB. Нехай m_{AB} – число випробувань, в яких відбулася подія AB. Тоді відносна частота відбування події A серед тих випробувань, коли відбувалася подія B, дорівнює

$$\frac{m_{AB}}{m_B} = \frac{m_{AB}/n}{m_B/n} = \frac{P_n^*(AB)}{P_n^*(B)}.$$

Цю величину називають *умовною статистичною ймовірністю події A за умови, що подія B мала місце*, і позначають через $P_n^*(A/B)$. Величину $P_n^*(A) = P_n^*(A/\Omega)$ називають *безумовною статистичною ймовірністю події* A.

Зауважимо, що умовна статистична ймовірність

$$P_n^*(A/B) = \frac{P_n^*(AB)}{P_n^*(B)}$$

є нормованою мірою підмножин AB множини B, одержаною нормуванням міри $P_n^*(AB)$, $AB \subset B$, $A \in S$, $B \in S$. Стосовно міри $P_n^*(A/B)$ очевидно задовільняються вимоги

$$0 \leq P_n^*(A/B) \leq 1,$$

зокрема $P_n^*(B/B) = \dfrac{P_n^*(B)}{P_n^*(B)} = 1$.

Якщо в механічній інтерпретації $P_n^*(B)$ – це маса, що припадає на множину B, а $P_n^*(AB)$ – маса, що припадає на множину AB, то $\dfrac{P_n^*(AB)}{P_n^*(B)}$ – частка від маси, що є в множині B, яка припадає на множину A. $P_n^*(A/B) = \dfrac{P_n^*(AB)}{P_n^*(B)}$ є також часткою від кількості попадань в множину B, яка припадає на множину A, тобто на AB.

З означення умовної статистичної ймовірності:

$$P_n^*(A/B)=\frac{P_n^*(AB)}{P_n^*(B)} \tag{1.10.1}$$

випливає

$$P_n^*(AB)=P_n^*(B)\,P_n^*(A/B)\,.$$

Звідси, враховуючи, що $AB=BA$, дістанемо також рівність:

$$P_n^*(AB)=P_n^*(A)P_n^*(B/A)\,. \tag{1.10.2}$$

Приклад 1.10.1. Кубик підкидали дуже велику кількість n разів. В результаті виявилось, що відносна частота випадання на верхній грані кубика цифри “1” дорівнює 0.01, цифри “2” – 0.02, цифри “3” – 0.03, цифри “4” – 0.04, цифри “5” – 0.30, цифри “6” – 0.60.

Позначимо можливі наслідки розглядуваних випробувань (елементарні події) “i” через E_i, $i\in\{1,2,...,6\}$. Тоді простір елементарних подій $\Omega=\{E_1,E_2,E_3,\ E_4,E_5,E_6\}$. Як події розглядатимемо будь які підмножини множини Ω, тобто розглянемо простір подій, до яких віднесено події $\varnothing$ і Ω, $E_i=''i''$, $i\in\{1,...,6\}$, – всі одноелементні підмножини множини Ω, а також всі двохелементні, всі триелементні, всі чотириелементні, всі п’ятиелементні підмножини множини Ω. Нехай подія A полягає в тому, що на верхній грані кубика випадає парна цифра, B – непарна цифра, C – одна із цифр “4”, ”5”, ”6”, D – одна із цифр “1”, ”2”, ”3”, ”4” (Рис. 1.10.1).

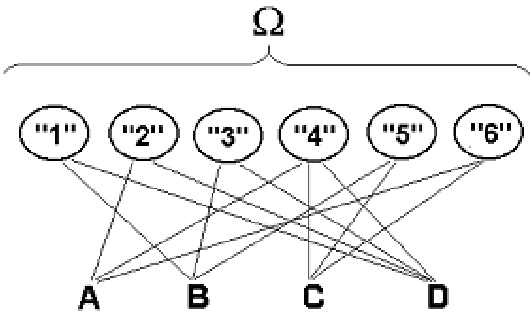


Рис. 1.10.1

Враховуючи заданий *розподіл статистичних ймовірностей на множині* Ω *елементарних подій*:

E_i	E_1	E_2	E_3	E_4	E_5	E_6
$P_n^*(\{E_i\})$	0.01	0.02	0.03	0.04	0.30	0.60

та формулу $P_n^*(A) = \sum_{E_i \in A} P_n^*(\{E_i\})$, стосовно розглядуваних подій A, B, C, D дістанемо:

$$P_n^*(A) = 0.66,\ P_n^*(B) = 0.34,\ P_n^*(C) = 0.94,\ P_n^*(D) = 0.10.$$

Оскільки $AB = \varnothing$, $AC = \{E_4, E_6\}$, $AD = \{E_2, E_4\}$, то за формулою (1.10.1) стосовно умовних статистичних ймовірностей одержуємо:

$$P_n^*(A/B) = \frac{P_n^*(AB)}{P_n^*(B)} = \frac{P_n^*(\varnothing)}{P_n^*(\{E_1, E_3, E_5\})} = \frac{0}{0.34} = 0,$$

$$P_n^*(B/A) = \frac{P_n^*(AB)}{P_n^*(A)} = \frac{P_n^*(\varnothing)}{P_n^*(\{E_2, E_4, E_6\})} = \frac{0}{0.66} = 0,$$

$$P_n^*(A/C) = \frac{P_n^*(AC)}{P_n^*(C)} = \frac{P_n^*(\{E_4, E_6\})}{P_n^*(\{E_4, E_5, E_6\})} = \frac{0.64}{0.94} = 0.68,$$

$$P_n^*(A/D) = \frac{P_n^*(AD)}{P_n^*(D)} = \frac{P_n^*(\{E_2, E_4\})}{P_n^*(\{E_1, E_2, E_3, E_4\})} = \frac{0.06}{0.10} = 0.6.$$

Стосовно добутку трьох подій A, B, C правильна рівність:

$$P_n^*(ABC) = P_n^*((AB)C) = P_n^*(AB)\, P_n^*(C/AB) =$$
$$= P_n^*(A)\, P_n^*(B/A)\, P_n^*(C/AB).$$

В механічній інтерпретації – щоб одержати масу, що припадає на множину ABC, треба спочатку знайти масу, що припадає на AB, для чого треба взяти частку, що припадає на множину B, від маси, що є в множині A, потім від маси, що є в AB, взяти частку, що припадає на множину C.

За методом математичної індукції легко вивести *формулу статистичної ймовірності добутку довільної скінченної кількості подій* A_1, $A_2, \ldots$, A_m, $A_i \in S$, $i \in \overline{1, m}$,

$$P_n^*(A_1 A_2 \ldots A_{m-1} A_m) = P_n^*((A_1 A_2 \ldots A_{m-1}) A_m) =$$
$$= P_n^*(A_1 A_2 \ldots A_{m-1}) P_n^*(A_m / A_1 A_2 \cdots A_{m-1}) = \ldots = \qquad (1.10.3)$$
$$= P_n^*(A_1) P_n^*(A_2 / A_1) P_n^*(A_3 / A_1 A_2) \cdots P_n^*(A_m / A_1 A_2 \cdots A_{m-1}).$$

Приклад 1.10.2. В скриньці є 49 кульок, які пронумеровані від 1 до 49. Дуже велику кількість n разів проводились такі випробування – навмання одна за однією без повернення в скриньку виймали шість кульок. В результаті виявилось, що відносна частота появи будь якої із 49 кульок першою дорівнює $\frac{1}{49}$. Після того, як першу кульку вийнято, відносна частота появи будь якої із решти 48 кульок дорівнює $\frac{1}{48}$ незалежно від номера першої кульки. Після того, як вийнято дві кульки, відносна частота появи будь якої із решти 47 кульок дорівнює $\frac{1}{47}$ незалежно від номерів двох вийнятих кульок. Після того, як вийнято три кульки, відносна частота появи будь якої із решти 46 кульок дорівнює $\frac{1}{46}$ незалежно від номерів трьох вийнятих кульок. Після того, як вийнято чотири кульки, відносна частота появи будь якої із решти 45 кульок дорівнює $\frac{1}{45}$ незалежно від номерів чотирьох вийнятих кульок. Після того, як вийнято 5 кульок, відносна частота появи будь якої із решти 44 кульок дорівнює $\frac{1}{44}$ незалежно від номерів п'яти вийнятих кульок.

Потрібно обчислити відносну частоту (статистичну ймовірність) відбування події A, яка полягає в тому, що на кожній із шести кульок, навмання вийнятих в одному випробуванні, був один із шести наперед задуманих номерів.

Припустимо, що задумали номери $i_1, i_2, i_3, i_4, i_5, i_6$. Позначимо через A_1 подію, яка полягає в тому, що першою вийнято кульку з номером m_1, який є серед задуманих, тобто $m_1 \in \{i_1, i_2, i_3, i_4, i_5, i_6\}$. Зрозуміло, що ця подія спричинюється через шість попарно несумісних подій, кожна з яких полягає в тому, що першою вийнято кульку, номер якої дорівнює одному із задуманих, тобто дорівнює або i_1, або i_2, або i_3, або i_4, або i_5, або i_6.

За умовою задачі статистична ймовірність кожної із цих шести подій дорівнює $\frac{1}{49}$, а тому $P_n^*(A_1) = \frac{6}{49}$.

Позначимо через A_2 – подію, яка полягає в тому, що номер m_2 кульки, яку вийнято другою, є одним із шести задуманих номерів, тобто $m_2 \in \{i_1, i_2, i_3, i_4, i_5, i_6\}$.

Якщо подія A_1 відбулася, тобто першою вийнято кульку з номером $m_1 \in \{i_1, i_2, i_3, i_4, i_5, i_6\}$, то кулька з номером m_1 не може бути вийнята вдруге. Отже за умови, що подія A_1 відбулася, подія A_2 спричинюється п'ятьма попарно несумісними подіями, кожна з яких полягає в тому, що другою вийнято кульку, номер якої m_2 є одним із задуманих, але $m_2 \neq m_1$, тобто

$$m_2 \in \{i_1, i_2, i_3, i_4, i_5, i_6\} \setminus \{m_1\}.$$

За умовою задачі умовна статистична ймовірність кожної з цих п’яти подій дорівнює $\frac{1}{48}$. Тому $P_n^*(A_2 / A_1) = \frac{5}{48}$.

Аналогічно, якщо A_3, A_4, A_5, A_6 події, які полягають в тому, що відповідно третьою, четвертою, п’ятою, шостою вийнято кульку з номерами m_3, m_4, m_5, m_6, кожний з яких є одним із задуманих номерів і крім того не співпадає з номерами раніше вийнятих кульок, які також є серед задуманих, то

$$P_n^*(A_3 / A_1 A_2) = \frac{4}{47}, \qquad P_n^*(A_4 / A_1 A_2 A_3) = \frac{3}{46},$$

$$P_n^*(A_5 / A_1 A_2 A_3 A_4) = \frac{2}{45}, \qquad P_n^*(A_6 / A_1 A_2 A_3 A_4 A_5) = \frac{1}{44}.$$

Очевидно, що подія A відбувається тоді, коли *в одному й тому самому випробуванні* (вийманні одна за однією шести кульок) відбуваються всі шість подій $A_1, A_2, A_3, A_4, A_5, A_6$, тобто

$$A = A_1 A_2 A_3 A_4 A_5 A_6.$$

Враховуючи формулу (1.10.3), одержимо

$$P_n^*(A) = P_n^*(A_1 A_2 A_3 A_4 A_5 A_6) =$$
$$= P_n^*(A_1) P_n^*(A_2 / A_1) P_n^*(A_3 / A_1 A_2) P_n^*(A_4 / A_1 A_2 A_3) \cdot$$
$$\cdot P_n^*(A_5 / A_1 A_2 A_3 A_4) P_n^*(A_6 / A_1 A_2 A_3 A_4 A_5) =$$
$$= \frac{6}{49} \cdot \frac{5}{48} \cdot \frac{4}{47} \cdot \frac{3}{46} \cdot \frac{2}{45} \cdot \frac{1}{44} \approx \frac{7}{100000000}.$$

Отже, в наведеній серії дуже великої кількості n випробувань 6 номерів із 49 можливих вдавалося вгадати в середньому 7 разів у 100 мільйонах спроб.

Вправи для самостійного виконання

1.10.1. Перевірити, чи правильні твердження:

1. Умовна статистична ймовірність $P_n^*(A/B)$ існує за будь-яких подій A та B.
2. Статистична ймовірність добутку подій дорівнює добуткові статистичних ймовірностей цих подій.
3. Умовна статистична ймовірність потрібна в разі необхідності визначення статистичної ймовірності добутку подій.
4. $P_n^*(AB) = 1 - P_n^*(\overline{A} + \overline{B})$.

1.10.2. Відомо, що подія A полягає у випаданні не менше ніж a очок за підкидання кубика, причому $P_n^*(\{"i"\}) = \frac{1}{6}$, $i \in \{1, 2, ..., 6\}$.

Знайти $P_n^*(A/B)$, де подія B полягає у випаданні 4 очок, коли 1) $a=1$; 2) $a=2$; 3) $a=3$; 4) $a=4$; 5) $a=5$; 6) $a=6$.

1.11. Залежні і незалежні відносно ймовірнісної міри P_n^* випадкові події

Нехай задано ймовірнісний простір (Ω, S, P_n^*).

Означення. *Події* $A \in S$ *і* $B \in S$ *називаються незалежними відносно ймовірнісної міри* P_n^*, якщо

$$P_n^*(AB) = P_n^*(A)P_n^*(B).$$

Оскільки $P_n^*(AB) = P_n^*(B)P_n^*(A/B)$, то події $A \in S$ і $B \in S$ незалежні тоді і тільки тоді, коли $P_n^*(A/B) = P_n^*(A)$, тобто коли умовна статистична ймовірність $P_n^*(A/B)$ події A, знайдена за умови, що подія B мала місце, співпадає з безумовною статистичною ймовірністю $P_n^*(A)$ події A.

Аналогічно, якщо $A \in S$ і $B \in S$ незалежні події, то і $P_n^*(B/A) = P_n^*(B)$. Якщо ж $P_n^*(A) = 0$ або $P_n^*(B) = 0$, то події A і B незалежні, оскільки $AB \subset A$, $AB \subset B$, і тому $0 \le P_n^*(AB) \le P_n^*(A) = 0$ або $0 \le P_n^*(AB) \le P_n^*(B) = 0$, звідки $P_n^*(AB) = 0 = P_n^*(A) \cdot P_n^*(B)$.

Слід підкреслити, що розглядувані події $A \in S$ і $B \in S$ мають бути елементами одного і того самого простору подій.

Означення. *Події* $A_1, A_2, \ldots, A_k$, $A_i \in S$, $i \in \overline{1,k}$ *називаються незалежними в сукупності відносно ймовірнісної міри* P_n^*, якщо

$$P_n^*(\prod_{i \in I} A_i) = \prod_{i \in I} P_n^*(A_i)$$

за довільного $I \subset \{1, 2, \ldots, k\}$, $I \ne \varnothing$.

З останнього означення випливає, що для того, щоб події $A_1, A_2, \ldots, A_k$ були незалежні в сукупності відносно міри P_n^*, потрібно, щоб кожна з них не залежала від будь-якої сукупності інших подій, тобто щоб мали місце рівності

$$P_n^*(A_i / \prod_{j \in I \setminus \{i\}} A_j) = P_n^*(A_i),\ I \subset \{1, 2, \ldots, k\},\ I \setminus \{i\} \ne \varnothing.$$

Зауважимо, що коли замість ймовірнісної міри $P_n^*(A)$, $A \in S$, на тому самому просторі S подій задати іншу ймовірнісну міру $\widetilde{P}_n^*(A)$, $A \in S$, то події, що були незалежні (чи незалежні в сукупності) відносно ймовірнісної міри P_n^*, не обов'язково залишатимуться такими самими відносно ймовірнісної міри $\widetilde{P}_n^*$.

Одні і ті самі події A і B, $A \in S$, $B \in S$, відносно однієї ймовірнісної міри P_n^*, заданої на S, можуть бути незалежними, а відносно іншої ймовірнісної міри $\tilde{P}_n^*$, заданої на тому самому просторі подій, можуть бути залежними. Тому говорити про незалежність чи залежність подій A і B безвідносно якоїсь ймовірнісної міри некоректно.

Приклад 1.11.1. Нехай проведено серію випробувань, в кожному з яких із скриньки, де було 40 кульок чотирьох кольорів – білі, червоні, зелені, жовті, 10 кульок кожного кольору, навмання вибирали одну кульку, фіксували її колір і повертали в скриньку (Рис. 1.11.1).

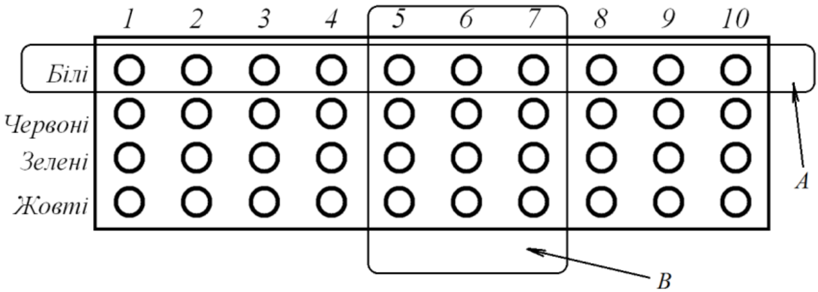


Рис. 1.11.1

Кульки кожного кольору пронумеровані від 1 до 10.

Нехай до простору подій S включаються всі підмножини заданої сорокаелементної множини Ω, подія A – кулька білого кольору, подія B – на кульці нанесено один із номерів 5, 6, 7 (Рис. 1.11.1).

а) Нехай в результаті проведеної серії випробувань отримали, що всі кульки вибирали однаково часто, тобто із статистичною ймовірністю $\frac{1}{40}$ кожна. Тоді

$$P_n^*(A) = \frac{10}{40} = \frac{1}{4},\ P_n^*(B) = \frac{12}{40} = \frac{3}{10},\ P_n^*(AB) = \frac{3}{40},$$

$$P_n^*(A/B) = \frac{3/40}{12/40} = \frac{1}{4} = P_n^*(A),\ P_n^*(B/A) = \frac{3/40}{10/40} = \frac{3}{10} = P_n^*(B),$$

$$P_n^*(AB) = P_n^*(A) \cdot P_n^*(B/A) = P_n^*(B) \cdot P_n^*(A/B) = P_n^*(A) \cdot P_n^*(B) = \frac{10}{40} \cdot \frac{12}{40} = \frac{3}{40}.$$

Таким чином відносно так заданої міри $P_n^*(A)$, $A \in S$, розглядувані події A і B незалежні.

б) Нехай в результаті іншої серії випробувань з тими самими множиною Ω, сукупністю підмножин S і подіями A, B отримали

$\widetilde{P}_n^*(AB) = 0$ (в множину AB не попадали жодного разу), $\widetilde{P}_n^*(A) = \dfrac{1}{2}$ (в половині всіх випробувань серії попадали в множину A), $\widetilde{P}_n^*(B) = \dfrac{1}{4}$ (в четвертій частині всіх випробувань серії попадали в множину B), але біла кулька з номерами 5, 6, 7 не була вибрана жодного разу. Тоді, очевидно,

$$\widetilde{P}_n^*(AB) = 0 \neq \widetilde{P}_n^*(A) \cdot \widetilde{P}_n^*(B) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

і тому відносно ймовірнісної міри $\widetilde{P}_n^*(A)$, $A \in S$, розглядувані події A і B залежні.

Приклад 1.11.2. Нехай $\Omega = \{E_1, E_2, E_3, E_4\}$ (Рис. 1.11.2) і за результатами досить великої серії випробувань отримали

$$P_n^*(\{E_1\}) = \frac{1}{4},\ P_n^*(\{E_2\}) = \frac{1}{4},\ P_n^*(\{E_3\}) = \frac{1}{4},\ P_n^*(\{E_4\}) = \frac{1}{4}.$$

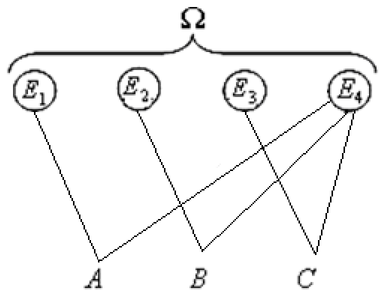


Рис. 1.11.2

Нехай $S = \{A \mid A = \bigcup\limits_{i \in I} \{E_i\},\ I \subset \{1, 2, 3, 4\}\}$, події A, B, C визначені так: $A = \{E_1, E_4\}$, $B = \{E_2, E_4\}$, $C = \{E_3, E_4\}$.

Тоді $P_n^*(A) = \dfrac{2}{4} = \dfrac{1}{2}$, $P_n^*(B) = \dfrac{2}{4} = \dfrac{1}{2}$, $P_n^*(C) = \dfrac{2}{4} = \dfrac{1}{2}$.

Оскільки $AB = \{E_4\}$, $AC = \{E_4\}$, $BC = \{E_4\}$, то

$$P_n^*(AB) = \frac{1}{4} = P_n^*(A)P_n^*(B) = \frac{1}{2} \cdot \frac{1}{2},$$

$$P_n^*(AC) = \frac{1}{4} = P_n^*(A)P_n^*(C) = \frac{1}{2} \cdot \frac{1}{2},$$

$$P_n^*(BC) = \frac{1}{4} = P_n^*(B)P_n^*(C) = \frac{1}{2} \cdot \frac{1}{2}.$$

Таким чином події A, B, C відносно розглядуваної ймовірнісної міри P_n^* попарно незалежні. Тут подія A відбувалася в половині всіх n випробувань, а також в половині тих випробувань, коли відбувалася подія B, оскільки

$$P_n^*(A) = \frac{1}{2}, \ P_n^*(A/B) = \frac{P_n^*(AB)}{P_n^*(B)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

Те саме можна сказати про безумовні статистичні ймовірності $P_n^*(B)$, $P_n^*(C)$ та умовні статистичні ймовірності $P_n^*(B/A)$, $P_n^*(A/C)$, $P_n^*(C/A)$, $P_n^*(B/C)$, $P_n^*(C/B)$.

Разом з тим $ABC = \{E_4\}$, тому

$$P_n^*(ABC) = \frac{1}{4} \neq P_n^*(A)P_n^*(B)P_n^*(C) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8}.$$

Той факт, що, наприклад, подія A залежить від добутку подій B і C, очевидний, оскільки у всіх тих випробуваннях, коли відбувалася подія BC (добуток подій B і C), відбувалася і подія A, тобто

$$P_n^*(A/BC) = 1 \left(= \frac{P_n^*(ABC)}{P_n^*(BC)} = \frac{1/4}{1/4} \right).$$

Оскільки $P_n^*(A/BC) = 1 \neq P_n^*(A) = \dfrac{1}{2}$, то подія A залежить від добутку подій B і C.

Отже, в розглядуваному випадку події A, B, C попарно незалежні, але залежні в сукупності.

Вправи для самостійного виконання

1.11.1. Перевірити, чи правильні твердження:

1. Будь-які події A і B є незалежними, якщо $P_n^*(A) = 0$ або $P_n^*(B) = 0$.
2. Події A, B і C незалежні в сукупності, якщо вони попарно незалежні.
3. Твердження, обернене до 2, правильне.
4. Якщо події незалежні відносно однієї ймовірнісної міри, то вони незалежні і відносно будь-якої іншої ймовірнісної міри.

1.11.2. Нехай $\Omega = \{ГГ, ГЦ, ЦГ, ЦЦ\} = \{E_1, E_2, E_3, E_4\}$ – простір елементарних подій, що відповідає одночасному підкиданню двох монет, причому $P_n^*(\{E_k\}) = \dfrac{1}{4}, \ k \in \overline{1, 4}$. Подія A полягає у випаданні герба на першій монеті, подія B – випадання герба на другій монеті.

1. Чи є події A і B незалежними відносно заданої міри P_n^*?
2. Чи можна визначити ймовірнісну міру P_n^* так, щоб A і B стали незалежними (чи залежними) подіями?

1.11.3. Нехай A і B події з попередньої задачі. Чи існує подія C така, що події A, B і C – незалежні в сукупності відносно заданої міри P_n^*?

1.12. Формула повної статистичної ймовірності. Формула Байєса

Нехай простір Ω елементарних подій поділено на m підмножин $H_1, H_2, \ldots, H_m$ таких, що $H_i \in S$, $H_i H_j = \varnothing$, коли $i \neq j$, $H_1 + H_2 + \ldots + H_m = \Omega$, і нехай A – подія, $A \subset \Omega$, $A \in S$ (Рис. 1.12.1). Надалі H_i називатимемо *гіпотезами*. Оскільки

$$A = A \cdot \Omega = A(H_1 + H_2 + \ldots + H_m) = AH_1 + AH_2 + \ldots + AH_m,$$

причому $AH_i \cdot AH_j = \varnothing$, коли $i \neq j$, то за властивістю аддитивності статистичної ймовірності і формулою (1.10.2):

$$P_n^*(A) = P_n^*(AH_1 + AH_2 + \ldots + AH_m) = P_n^*(AH_1) + P_n^*(AH_2) + \ldots + P_n^*(AH_m) =$$

$$= P_n^*(H_1) P_n^*(A/H_1) + P_n^*(H_2) P_n^*(A/H_2) + \ldots + P_n^*(H_m) P_n^*(A/H_m).$$

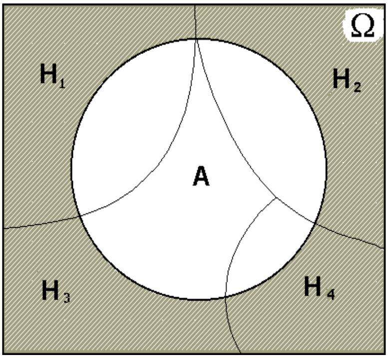


Рис. 1.12.1

Таким чином,

$$P_n^*(A) = \sum_{i=1}^{m} P_n^*(H_i) P_n^*(A/H_i) \tag{1.12.1}$$

Цю формулу називають *формулою повної статистичної ймовірності*.

Отже, коли відомі або легко знаходяться статистичні ймовірності $P_n^*(H_i)$ гіпотез H_i, а також відомі або легко знаходяться умовні статистичні ймовірності $P_n^*(A/H_i)$ події A за кожною з гіпотез H_i, $i \in \{1, 2, \ldots, m\}$, тоді статистичну ймовірність $P_n^*(A)$ події A можна обчислити за формулою (1.12.1).

Приклад 1.12.1. Є дві скриньки, в яких є білі і чорні кульки. Випробування полягає в тому, що навмання вибирається скринька, а з неї навмання вибирається кулька, яка потім повертається в скриньку. Із n випробувань m_1 разів була вибрана перша скринька, серед них в $k_1<m_1$ випадків кулька була біла, і $m_2=n-m_1$ разів була вибрана друга скринька, серед них в $k_2<m_2$ випадків кулька була біла.

Природно вважати, що простір елементарних подій $\Omega=\{(1,б),(1,ч),(2,б),(2,ч)\}$, де елементарна подія $E_{kб}=(k,б)$, $k\in\overline{1,2}$, полягає в тому, що з k-тої скриньки вийнято білу кульку. Подія $H_k=\{(k,б),(k,ч)\}$ полягає в тому, що кульку вийнято з k-тої скриньки, $k\in\overline{1,2}$. За умовою задачі $P_n^*(H_1)=\dfrac{m_1}{n}$, $P_n^*(H_2)=\dfrac{m_2}{n}$. Очевидно $H_1\cdot H_2=\varnothing$ і $H_1+H_2=\Omega$.

Нехай подія $A=\{(1,б),(2,б)\}$ – поява білої кульки. За умовою задачі $P_n^*(A/H_1)=\dfrac{k_1}{m_1}$, $P_n^*(A/H_2)=\dfrac{k_2}{m_2}$.

Тому за формулою (1.12.1) статистична ймовірність (відносна частота) появи білої кульки в розглядуваних n випробуваннях дорівнює

$$P_n^*(A)=P_n^*(H_1)P_n^*(A/H_1)+P_n^*(H_2)P_n^*(A/H_2)=\frac{m_1}{n}\cdot\frac{k_1}{m_1}+\frac{m_2}{n}\cdot\frac{k_2}{m_2}.$$

Якщо в умовах, коли відомі $P_n^*(H_i)$ і $P_n^*(A/H_i)$, потрібно визначити, чому дорівнює $P_n^*(H_i/A)$, тобто відносну частоту відбування події H_i серед тих випробувань, коли відбувалася подія A, то, виходячи з означення умовної статистичної ймовірності та формули повної статистичної ймовірності, дістанемо:

$$P_n^*(H_i/A)=\frac{P_n^*(AH_i)}{P_n^*(A)}=\frac{P_n^*(H_i)P_n^*(A/H_i)}{\sum\limits_{i-1}^{m}P_n^*(H_i)P_n^*(A/H_i)}.$$

Останню формулу називають *формулою Байєса*.

Приклад 1.12.2. За результатами багатократного передавання телеграфних повідомлень за допомогою двох знаків "•" (крапка) і "▬" (тире) було з'ясовано, що відносна частота спотворення знака "•" дорівнює $\dfrac{2}{5}$, а відносна частота спотворення знака “▬” дорівнює $\frac{2}{3}$. Крім того було помічено, що у повідомленнях, які треба передавати, відносна частота появи знака “ ●” дорівнює $\frac{5}{8}$, а відносна частота появи знака “▬” дорівнює $\frac{3}{8}$. Потрібно обчислити відносну частоту випадків, коли передавали знак “ ●”, серед тих випадків, коли було прийнято знак “ ●”.

Природно вважати, що простором елементарних подій є $\Omega = \{$“●●”, “●▬”, “▬●”, “▬ ▬”$\}$, де кожній елементарній події $E \in \Omega$ відповідає сукупність двох знаків: першим вказано переданий знак, а другим – прийнятий.

Нехай подія $A = \{$“●●”, “▬●”$\}$ полягає в тому, що прийнято знак “ ●”. Така подія може відбутися як за умови, що передавали знак “ ●”, так і за умови, що передавали знак “▬”. Отже є дві гіпотези: $H_1 = \{$“●●”, “●▬”$\}$ – передавали знак “ ●”, $H_2 = \{$“▬●”, “▬ ▬”$\}$ – передавали знак “▬”. Очевидно $H_1 H_2 = \varnothing$, $H_1 + H_2 = \Omega$. За умовами задачі $P_n^*(H_1) = \frac{5}{8}$, $P_n^*(H_2) = \frac{3}{8}$. Крім того, оскільки відносна частота спотворення знака “ ●” дорівнює $\frac{2}{5}$, а знака “▬” – $\frac{2}{3}$, то подія A за гіпотези H_1 відбувається тоді, коли знак “ ●” не спотворюється, тобто $P_n^*(A/H_1) = 1 - \frac{2}{5} = \frac{3}{5}$, а за гіпотези H_2 подія A відбувається тоді , коли спотворюється знак “▬”, тобто $P_n^*(A/H_2) = \frac{2}{3}$.

Тоді за формулою повної статистичної ймовірності

$$P_n^*(A) = P_n^*(H_1) P_n^*(A/H_1) + P_n^*(H_2) P_n^*(A/H_2) = \frac{5}{8} \cdot \frac{3}{5} + \frac{3}{8} \cdot \frac{2}{3} = \frac{5}{8}.$$

Щоб визначити статистичну ймовірність того, що передавали знак “ ●”, за умови, що прийнято знак “ ●”, тобто $P_n^*(H_1/A)$, скористаємось означенням умовної статистичної ймовірності $P_n^*(A/B) = \frac{P_n^*(AB)}{P_n^*(B)}$ або формулою Байєса.

Стосовно розглядуваного випадку одержуємо:

$$P_n^*(H_1/A) = \frac{P_n^*(AH_1)}{P_n^*(A)} = \frac{P_n^*(H_1) P_n^*(A/H_1)}{P_n^*(A)} =$$

$$=\frac{P_n^*(H_1)P_n^*(A/H_1)}{P_n^*(H_1)P_n^*(A/H_1)+P_n^*(H_2)P_n^*(A/H_2)}=\frac{\frac{5}{8}\cdot\frac{3}{5}}{\frac{5}{8}}=\frac{3}{5}.$$

Вправи для самостійного виконання

1.12.1. Перевірити чи правильні твердження:

1. Якщо $H_1+H_2+...+H_n=\Omega$, то H_i – гіпотези.

2. Якщо $H_i\cdot H_j=\varnothing$, коли $i\neq j$, то H_i – гіпотези.

3. Сума умовних статистичних ймовірностей $P_n^*(A/H_i)$, $i\in\overline{1,k}$, може бути більшою, ніж 1.

4. За кожної гіпотези H_i існує подія A така, що $P_n^*(H_i/A)=P_n^*(H_i)$.

5. Кожна гіпотеза за даного ймовірнісного простору $(\Omega,\ S,\ P_n^*)$ відбувається з єдиною статистичною ймовірністю, яка не залежить від того, відбулася чи ні якась подія A.

6. Кожна гіпотеза за даного ймовірнісного простору $(\Omega,\ S,\ P_n^*)$ може відбуватися з різними умовними статистичними ймовірностями.

1.12.2. За даними прикладу 1.12.2 обчислити:

а) відносну частоту випадків, коли передавали знак "▬", серед тих випадків, коли було прийнято знак " ●";

б) відносну частоту випадків, коли передавали знак " ●", серед тих випадків, коли було прийнято знак "▬";

в) відносну частоту випадків, коли передавали знак "▬", серед тих випадків, коли було прийнято знак "▬".

РОЗДІЛ II. РОЗПОДІЛИ СТАТИСТИЧНИХ ЙМОВІРНОСТЕЙ

2.1. Поняття розподілу статистичних ймовірностей на множині елементарних подій.

Нехай задано ймовірнісний простір (Ω, S, P) (див. §1.8), де $\Omega = \bigcup\limits_{i=1}^{k} H_i$, $H_i H_j = \varnothing$, коли $i \neq j$, $i \in \overline{1,k}$, $j \in \overline{1,k}$, $S = \{A \mid A = \bigcup\limits_{i \in I} H_i, \ I \subset \{1, 2, \dots, k\}\}$, і задані ймовірності $P(H_i)$, $i \in \overline{1,k}$, $P(H_i) \geq 0$, $\sum\limits_{i=1}^{k} P(H_i) = 1$, $k \in N$.

1. В такому разі будемо говорити, що на множині Ω задано розподіл ймовірностей за множинами H_i, $i \in \overline{1,k}$, – помножинний розподіл за множинами H_i, $i \in \overline{1,k}$.

Разом з тим може бути невідомо, як саме розподілені ймовірності $P(H_i)$ на множинах H_i, $i \in \overline{1,k}$. Це можуть бути розподіли за дрібнішими підмножинами H_{ij}, $j \in \overline{1,k_i}$, множини H_i, $i \in \overline{1,k}$, чи за скінченними кількостями точок із множин H_i, $i \in \overline{1,k}$, чи неперервно в разі, коли H_i неперервна множина рівномірно чи нерівномірно, чи якимось комбінованим чином – частина $P(H_i)$ розподілена на скінченній множині точок із H_i, а інша частина – за підмножинами H_{ij}, $j \in \overline{1,k_i}$. Крім того розподіл ймовірностей на множинах H_i, $i \in \overline{1,k}$, можуть бути різних типів.

Зрештою ймовірнісна міра $P(A)$ за довільного $A \in S$ обчислюється за формулою

$$P(A) = P\Big(\bigcup_{H_i \subset A} H_i\Big) = \sum_{H_i \subset A} P(H_i)$$

і тому в розглядуваному разі відомості про те, як саме розподілені ймовірності $P(H_i)$ на множинах H_i, не є необхідними.

2. В разі, коли всі H_i, $i \in \overline{1,k}$, є одноточковими множинами – $H_i = \{x_i\}$, $i \in \overline{1,k}$, $x_i \in R^1$, розподіл ймовірностей на множині $\Omega = \bigcup\limits_{i=1}^{k} \{x_i\}$ будемо називати *поточковими (або розподілом за точками* x_i, $i \in \overline{1,k}$ *)*.

3. В разі, коли Ω неперервна множина, яку можна ототожнювати з числовим проміжком $\langle a; b \rangle$, наприклад $\Omega = (a; b]$, а множини H_i – інтервали типу $(a_{i-1}; a_i]$, $i \in \overline{1,k}$,

$a_0 = a,\ a_k = b$, $\Omega = \bigcup_{i=1}^{k}(a_{i-1}; a_i]$, $S = \{A \mid A = \bigcup_{i\in I}(a_{i-1}; a_i],\quad I \subset \{1, 2, \ldots, k\}\}$, і задані $P((a_{i-1}; a_i]),\ i \in \overline{1,k}$, тоді розподіл ймовірностей на множині $\Omega = \bigcup_{i=1}^{k}(a_{i-1}; a_i]$ будемо називати *поінтервальним (або розподілом за інтервалами* $(a_{i-1}; a_i],\ i \in \overline{1,k}$).

Зауважимо, що невідомо, як саме розподілені ймовірності $P((a_{i-1}; a_i])$ на інтервалах $(a_{i-1}; a_i],\ i \in \overline{1,k}$.

4. В разі, коли Ω неперервна множина, яку можна ототожнювати з проміжком $\langle a; b\rangle$, а до простору подій S включено всі борелівські підмножини множини Ω, тобто $S = \mathscr{B}(\Omega)$, довжини інтервалів $(a_{i-1}; a_i]$ спрямовуються до нуля, а ймовірнісна міра на S задається через щільність розподілу ймовірностей, тоді розподіл ймовірностей на множині Ω будемо називати *неперервним*.

Приклад 2.1.1. Нехай множину Ω поділено на підмножини $H_1,\ H_2,\ \ldots\ ,\ H_k$, такі, що $H_i H_j = \varnothing$, коли $i \neq j$, тобто в підмножинах H_i і H_j немає спільних елементів, коли $i \neq j$, а також $H_1 + H_2 + \ldots + H_k = \Omega$. Нехай до сукупності S підмножин множини Ω включено порожню множину $\varnothing$, всі підмножини $H_1,\ H_2,\ \ldots,\ H_k$, а також всеможливі об’єднання цих підмножин із двох, із трьох і т.д., із k-1, із всіх k підмножин, тобто $S = \{A \mid A = \bigcup_{i\in I} H_i,\ I \subset \{1, 2, 3, \ldots, k\}\}$. В такому разі говорять, що *простір подій S породжений за сукупністю підмножин* $H_1,\ H_2,\ \ldots,\ H_k$ таких, що $\bigcup_{i=1}^{k} H_i = \Omega$, $H_i H_j = \varnothing$, коли $i \neq j$.

Якщо крім того задано $P_n^*(H_i)$, $i \in \overline{1,k}$, такі, що $0 \le P_n^*(H_i) \le 1$, $\sum_{i=1}^{k} P_n^*(H_i) = 1$, то тим самим задано $P_n^*(A)$ за довільного $A \in S$. Тоді, якщо $A = \bigcup_{i\in I} H_i$, де $I \subset \{1,2,\ldots,k\}$ – деяка підмножина множини $\{1,2,\ldots,k\}$, зокрема порожня $\varnothing$, то $P_n^*(A) = \sum_{i\in I} P_n^*(H_i)$. В такому разі говорять, що *на множині* Ω *задано розподіл статистичних ймовірностей за підмножинами* H_i такий, що статистична ймовірність попадання в множину H_i дорівнює $P_n^*(H_i)$ (або іншими словами – на множину H_i припадає статистична ймовірність $P_n^*(H_i)$ – помножинний розподіл).

Якщо H_i одноелементні підмножини множини Ω, $H_i = \{E_i\}$, $\Omega = \{E_1, E_2, ..., E_k\}$, тоді S буде найбагатшим (найширшим) простором подій і до S входитимуть всі підмножини множини Ω разом з порожньою $\varnothing$ та самою множиною Ω.

В такому разі до S входять разом з $\varnothing$ і Ω всі події виду $\{E_i\}$, $i \in \overline{1,k}$, та всеможливі підмножини множини Ω – двохелементні, триелементні і т.д., (k-1)-елементні, тобто $S = \{A \mid A = \bigcup\limits_{i \in I} \{E_i\}, I \subset \{1, 2, 3, ..., k\}\}$, а за довільного $A \in S$ буде

$$A = \bigcup_{i \in I} \{E_i\}, \quad I \subset \{1, 2, ..., k\}, \quad P_n^*(A) = P_n^*(\bigcup_{i \in I} \{E_i\}) = \sum_{i \in I} P_n^*(\{E_i\}).$$

Таким чином одержується *розподіл статистичних ймовірностей* $P_n^*(H_i) = P_n^*(\{E_i\})$ *на множині* Ω *за елементарними подіями* $E_i \in \Omega$. Такий розподіл статистичних імовірностей на скінченній множині $\Omega = \{E_1, E_2, ..., E_k\}$ називатимемо *поточковим*.

Якщо множина Ω є множиною точок деякого координатного простору (наприклад, числової прямої – одновимірного координатного простору, площини – двохвимірного координатного простору тощо), тоді елементарні події $E \in \Omega$ будуть ототожнюватися з деякими точками відповідного координатного простору.

Розглянемо *дискретний простір* Ω елементарних подій, тобто такий, до якого входить скінченна або зчисленна кількість елементів, які можна занумерувати скінченною кількістю натуральних чисел від 1 до k або всіма натуральними числами. Отже, $\Omega = \{E_1, E_2, ..., E_k\}$ або $\Omega = \{E_1, E_2, ..., E_k, ...\}$. В геометричній інтерпретації кожну елементарну подію $E_i \in \Omega$ можна ототожнювати з деякою точкою на числовій прямій, тобто з деяким дійсним числом x_i, $i \in \overline{1,k}$ або $i \in N$, причому числа x_i попарно різні: $x_i \neq x_j$, коли $i \neq j$. Тому можна вважати, що $\Omega = \{x_1, x_2, ...\}$, причому $x_1 < x_2 < x_3 < ...$. В такий спосіб здійснюється перехід від безкоординатного простору до координатного.

Приклад 2.1.2. Якщо ототожнювати випадання білої грані в результаті підкидання кубика з числом 1, випадання зеленої – з числом 2, випадання червоної – з числом 3, за умови, що кожну грань пофарбовано в один з трьох кольорів – білий, зелений, червоний, то замість простору $\Omega = \{$білий, зелений, червоний$\}$, можна розглядати простір $\Omega = \{1, 2, 3\}$.

Приклад 2.1.3. Нехай експеримент полягає в підкиданні монети до першого випадання герба. Тоді замість простору елементарних подій $\Omega=\{Г, ЦГ, ЦЦГ, ЦЦЦГ, ...\}$ можна розглядати простір $\Omega=\{1,2,3,4,...\}$.

Припустимо, що проведено n випробувань, в результаті яких спостерігалися елементарні події $E_{\text{сп}\,i}\in\Omega$, $i\in\overline{1,n}$, замість яких можна розглядати відповідні числові значення $x_{\text{сп}\,i}\in\{x_1, x_2, ...\}$, $i\in\overline{1,n}$. Ці значення $x_{\text{сп}\,i}$, $i\in\overline{1,n}$, називають *спостереженими значеннями*. Якщо їх впорядкувати за величиною, то отримається так званий *варіаційний ряд*, члени якого називаються *варіантами*.

Якщо n_i – абсолютна частота відбування події $\{E_i\}$, тобто елементарна подія E_i спостерігалася в n випробуваннях n_i разів, причому $\sum_{i=1}^{k} n_i=n$, то число n_i є абсолютною частотою появи і значення x_i, а число $P_n^*(\{E_i\})=\dfrac{n_i}{n}=P_n^*(\{x_i\})$ – статистична ймовірність або відносна частота відбування події $\{E_i\}$, а отже і події $\{x_i\}$.

Табл. 2.1.1

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

Табл. 2.1.2

x_i	x_1	x_2	...	x_k
$P_n^*(\{x_i\})$	$P_n^*(\{x_1\})$	$P_n^*(\{x_2\})$	$\cdots$	$P_n^*(\{x_k\})$

За таблицями виду 2.1.1 і 2.1.2 задають відповідно *ряд розподілу абсолютних частот* та *ряд розподілу статистичних ймовірностей (відносних частот)* на дискретній множині $\Omega=\{E_1, E_2, ..., E_k\}$, або, що те саме, на множині точок $\{x_1, x_2, ..., x_k\}$. Говорять також, що за таблицями 2.1.1 і 2.1.2 задано *дискретний (поточковий) розподіл частот* (відповідно абсолютних та відносних). В такому разі якщо $A=\bigcup_{i\in I}\{x_i\}$, $I\subset\{1, 2, ..., k\}$, то $P_n^*(A)=\sum_{i\in I}P_n^*(\{x_i\})$.

Зауважимо, що коли Ω нескінченна дискретна множина, то таблиці 2.1.1 та 2.1.2 також мають бути нескінченні, але $n_i\neq 0$ і $P_n^*(\{x_i\})\neq 0$ лише для скінченної кількості спостережених значень x_i, і на практиці до таблиць записують лише такі елементи.

Для полегшення підрахунку чисел n_i, а отже і чисел $P_n^*(\{x_i\})$, спостережені значення $x_{\text{сп}\,i}$ розташовують в порядку їх зростання, в результаті чого одержують *варіаційний ряд.*

Приклад 2.1.4. Варіаційний ряд за спостережених значень

5, 6, 6, 6, 4, 3, 5, 2, 6, 4, 5,6, 3, 3, 5

набуває вигляду

2, 3, 3, 3, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6,

а відповідні ряди розподілу абсолютних та відносних частот подані в таблицях 2.1.3 і 2.1.4.

Табл. 2.1.3

x_i	1	2	3	4	5	6
n_i	0	1	3	2	4	5

Табл. 2.1.4

x_i	1	2	3	4	5	6
$P_n^*(\{x_i\})$	0	$\frac{1}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{5}{15}$

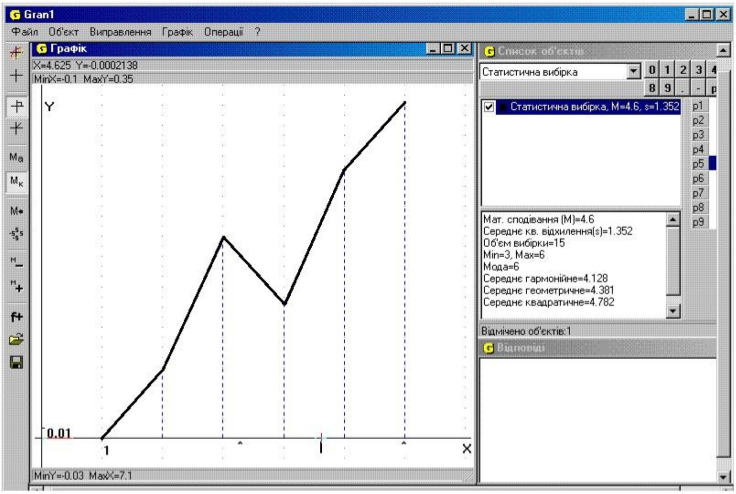


Рис. 2.1.1

В фізичній інтерпретації ряд розподілу абсолютних частот можна розглядати як цілочисельний поточковий розподіл маси n на множині точок $\{x_1, x_2, ..., x_k\}$, а ряд розподілу відносних частот – як поточковий розподіл відповідної нормованої (одиничної) маси на тій самій множині точок.

Оскільки за будь-якої події $A \subset \Omega$ $P_n^*(A) = \sum_{E_i \in A} P_n^*(\{E_i\})$, то в фізичній інтерпретації $P_n^*(A)$ – це частина одиничної маси, що припадає на всі точки із множини A, за умови, що на множині $\Omega = \{E_1, E_2, ..., E_k\}$ одинична маса розподілена так, що на точку E_i припадає маса $P_n^*(\{E_i\})$.

Якщо точки $(x_i, P_n^*(\{x_i\}))$ зобразити на координатній площині, то ламану з вершинами у цих точках називають *полігоном відносних частот* або *многокутником поточкового розподілу статистичних ймовірностей* (відносних частот).

Приклад 2.1.5. На Рис. 2.1.1 зображено полігон відносних частот, що визначається за таблицею 2.1.4.

Зауважимо, що дискретний розподіл статистичних ймовірностей можна задати і на неперервній множині Ω, вибравши дискретну підмножину множини Ω і задавши поточковий розподіл статистичних ймовірностей на цій підмножині.

Вправи для самостійного виконання

2.1.1. Перевірити, чи правильні твердження:

1. Елементарні події довільної множини Ω можна ототожнювати з натуральними числами.

2. Спостережені елементарні події E_i – це те саме, що й спостережені значення x_i.

3. В заданій серії з *n* випробувань кожному із спостережених і неспостережених значень можна поставити у відповідність абсолютні і відносні частоти їх спостережень.

4. Дискретний поточковий розподіл частот можна задати за таблицею.

5. За будь-якими таблицями виду 2.1.1 та 2.1.2 задають дискретний поточковий розподіл частот.

6. За варіаційним рядом цілком визначається дискретний поточковий розподіл частот.

7. Сукупність чисел 1, 1, 2, 2, 2, 4, 4, 4, 4, 3, 3, 3, 5, 5, 5, 5, 5, *5* можна вважати варіаційним рядом.

8. Статистичні ймовірності можна інтерпретувати як маси.

9. Кожен простір елементарних подій є дискретним.

10. Через полігон відносних частот цілком визначається дискретний поточковий розподіл частот.

11. Будь-яку ламану на координатній площині можна вважати многокутником розподілу статистичних ймовірностей.

2.1.2. 1. Побудувати варіаційний ряд спостережених значень, що відповідають спостереженим відхиленням точки падіння снаряда від цілі: -50, 20, -10, 20, 10, 20, -50, -20, -10, 40, -20,-30, -10, 10, 20, -40, 50, -10, 10, 50.

2. Використовуючи варіаційний ряд з 2.1.2.1, побудувати ряди розподілу абсолютних та відносних частот на множині можливих значень {-50, -40, -30, -20, -10, 0, 10, 20, 30, 40, 50}.

3. Побудувати полігон відносних частот стосовно розподілу з 2.1.2.2.

4. Підрахувати відносну частоту відбування події *A*, яка полягає в тому, що відхилення точки падіння снаряда від цілі не перевищує 20.

2.2. Щільність розподілу статистичних ймовірностей. Узагальнені статистичні ймовірності

Нехай задано ймовірнісний простір (Ω, S, P_n^*), де простір подій S породжений за поділом множини Ω на підмножини H_1, H_2, …, H_k, $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup\limits_{i=1}^{k} H_i = \Omega$, тобто $S = \{A \mid A = \bigcup\limits_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$, і нехай $S \subset \tilde{S}$, де $\tilde{S}$ – сукупність множин, стосовно якої задовільняються вимоги 1_s-3_s і на якій задано деяку міру $m(G)$, $G \in \tilde{S}$, за якої $0 < m(\Omega) < \infty$.

Наприклад, мірою $m(G)$ можуть бути кількості точок в підмножинах $G \in \tilde{S}$, довжини вимірних лінійних фігур $G \in \tilde{S}$, якщо G – множина точок на прямій, площі квадровних фігур, якщо G – множина точок в двохвимірному координатному просторі, об'єми кубовних фігур $G \in \tilde{S}$, маси фізичних тіл тощо.

Нехай $H_i \in S$, $i \in \overline{1, k}$, такі, що

$$H_i H_j = \varnothing \text{, коли } i \neq j \text{, } \bigcup_{i=1}^{k} H_i = \Omega \text{, } m(H_i) > 0 \text{, } i \in \overline{1, k} \text{,}$$

зокрема це можуть бути множини H_i такі, що $m(H_i) = \dfrac{m(\Omega)}{k}$, $i \in \overline{1, k}$.

Тоді відношення $\dfrac{P_n^*(H_i)}{m(H_i)}$ природно назвати *середньою щільністю* статистичної ймовірності (ймовірнісної міри) P_n^* в множині H_i відносно міри $m(A)$, $A \in S$.

Введемо до розгляду функцію

$$f_n^*(E) = \frac{P_n^*(H_i)}{m(H_i)}, \; E \in H_i, \; i \in \overline{1, k},$$

визначену за всіх $E \in \Omega$.

Таку функцію природно назвати *усередненою* або *середньою щільністю відносно міри m розподілу статистичних ймовірностей* (ймовірнісної міри P_n^*) *на множині* Ω *за підмножинами* H_i.

Із виразу $f_n^*(E) = \dfrac{P_n^*(H_i)}{m(H_i)}$, $E \in H_i$, $i \in \overline{1, k}$, слідує

$$P_n^*(H_i) = f_n^*(E) m(H_i), \; E \in H_i, \; i \in \overline{1, k}.$$

Звідси в разі $A=\bigcup_{i\in I}H_i$, $H_iH_j=\varnothing$, коли $i\neq j$, $I\subset\{1,2,...,k\}$,

$$P_n^*(A)=P_n^*(\bigcup_{i\in I}H_i)=\sum_{i\in I}P_n^*(H_i)=\sum_{i\in I,\,E\in H_i}f_n^*(E)m(H_i)\cdot$$

Підкреслимо, що ймовірнісна міра $P_n^*(A)$ визначена тільки стосовно множин A із сукупності S підмножин множини Ω.

Якщо $\tilde{\Omega}\supset\Omega$ деяка вимірна за мірою m множина, $\tilde{S}\supset S$ – сукупність вимірних за мірою m підмножин множини $\tilde{\Omega}$, відносно якої задовільняються вимоги 1_S-3_S, $G\in\tilde{S}$ – довільна вимірна за мірою m підмножина множини $\tilde{\Omega}$, тоді на сукупності $\tilde{S}\supset S$ вимірних за мірою m множин можна ввести нову ймовірнісну міру P, наприклад поклавши

$$P(G)=P(G\bigcap(\bigcup_{i=1}^{k}H_i))=P(\bigcup_{i=1}^{k}(G\bigcap H_i))=\sum_{i=1}^{k}P(G\bigcap H_i)=$$

$$=\sum_{i=1}^{k}\frac{m(G\bigcap H_i)}{m(H_i)}P_n^*(H_i)=\sum_{i=1,\,E\in G\bigcap H_i}^{k}\frac{m(G\bigcap H_i)}{m(H_i)}f_n^*(E)m(H_i)=$$

$$=\sum_{i=1,E\in G\bigcap H_i}^{k}f(E)m(G\bigcap H_i)\,. \qquad (2.2.1)$$

де $P(G)$, $G\in\tilde{S}$, вже не є статистична ймовірність $P_n^*(G)=\dfrac{k_n(G)}{k_n(\Omega)}$, $G\in\tilde{S}$, отримана за результатами серії із n проведених випробувань (за статистичними даними), а ймовірнісна міра, введена на основі припущення (гіпотези) про те, що на кожній із множин H_i статистичні ймовірності, отримані за статистичними даними (результатами серії із n проведених випробувань), розподілені рівномірно, тобто якщо є дві підмножини $Q\subset H_i$ і $G\subset H_i$ однакової міри $m(Q)=m(G)$, то і

$$P(Q)=\frac{m(Q)}{m(H_i)}\cdot P_n^*(H_i)=\frac{m(G)}{m(H_i)}\cdot P_n^*(H_i)=P(G),$$

$$f(E)=\begin{cases}\dfrac{P(G\bigcap H_i)}{m(G\bigcap H_i)}=f_n^*(E), & \text{коли } E\in G\bigcap H_i,\\[2mm] 0, & \text{коли } E\,\overline{\in}\,G\bigcap H_i.\end{cases}$$

Зауважимо, що коли $G=\Omega$, тоді $f(E)=f_n^*(E)$ за всіх $E\in\Omega$.

В такий спосіб *ймовірнісна міра* $P_n^*(A)$, $A\in S$, *задана на сукупності подій S, породженій за системою підмножин* H_i,

$i \in \overline{1,k}$ *, може бути поширена (продовжена) на сукупність* $\tilde{S} \supset S$ *множин, вимірних за мірою m*. Очевидно, стосовно $A \in S$ має місце рівність $P(A) = P_n^*(A)$.

Тут також робиться припущення (гіпотеза), що щільність розподілу статистичних ймовірностей на кожній із множин H_i стала, тобто $f_n^*(E) = c_i = \dfrac{P_n^*(H_i)}{m(H_i)}$, $E \in H_i$, $i \in \overline{1,k}$.

Насправді ж за результатами проведених випробувань (за статистичними даними) може виявитись (якщо фіксувати не тільки кількість точок, які попадають у множину H_i, а і кількості точок, які попадають у різні підмножини множини H_i, $i \in \overline{1,k}$), що $k_n(Q) \neq k_n(G)$, коли $Q \subset H_i$, $G \subset H_i$, $m(Q) = m(G)$, $Q \in \tilde{S}$, $G \in \tilde{S}$, тому ймовірнісні міри

$$P(G \cap H_i) = \frac{m(G \cap H_i)}{m(H_i)} P_n^*(H_i),\ G \in \tilde{S},$$

є гіпотетичними, які вводяться на основі припущення, що на кожній із множин H_i статистичні ймовірності розподілені рівномірно, тобто $P_n^*(G \cap H_i) = P_n^*(Q \cap H_i)$, коли $m(G) = m(Q)$, $G \subset H_i$, $Q \subset H_i$, $G \in \tilde{S}$, $Q \in \tilde{S}$, і що $f(E) = f_n^*(E)$, коли $E \in G \subset H_i$, тобто що

$$f(E) = \frac{P(G \cap H_i)}{m(G \cap H_i)} = f_n^*(E) = \frac{P_n^*(H_i)}{m(H_i)},\ E \in G \cap H_i,$$

коли $E \in G \cap H_i$, $i \in \overline{1,k}$, $G \in \tilde{S}$, що за статистичними даними може виявитись не так.

Більш загальний і практично прийнятний спосіб наближеного визначення міри $P(G)$ множини $G \subset \Omega$ за даними стосовно ймовірнісного простору (Ω, S, P_n^*) може бути таким. Нехай $G \subset \Omega$, $G \in \tilde{S}$, $\bigcup H_i$ – об'єднання деяких підмножин H_i, $G_* = \bigcup\limits_{\bigcup H_i \subset G \cap \Omega} \left(\bigcup H_i \right)$ – об'єднання всіх множин виду $\bigcup H_i$ таких, що $\bigcup H_i \subset G \cap \Omega$, тобто найширше об'єднання множин H_i таке, що $\bigcup H_i \subset G \cap \Omega$, $G^* = \bigcap\limits_{G \cap \Omega \subset \bigcup H_i} \left(\bigcup H_i \right)$ – перетин всіх множин виду $\bigcup H_i$ таких, що $G \cap \Omega \subset \bigcup H_i$, тобто найвужче об'єднання множин H_i таке, що $G \cap \Omega \subset \bigcup H_i$. Очевидно $G_* \in S$, $G^* \in S$, $G_* \subset G^*$, $G^* \setminus G_* \in S$ (Рис. 2.2.1). Покладемо

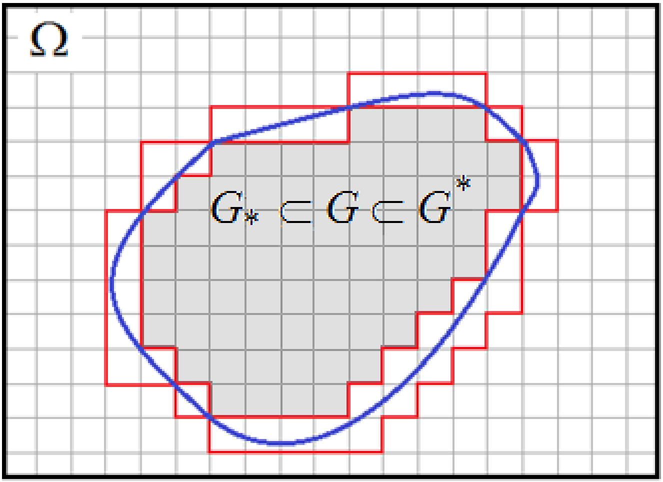


Рис. 2.2.1

$$P(G) = P_n^*(G_*) + \frac{m(G \bigcap (G^* \setminus G_*))}{m(G^* \setminus G_*)} P_n^*(G^* \setminus G_*) . \qquad (2.2.2)$$

Очевидно $P_n^*(G_*) \le P(G) \le P_n^*(G^*)$.

Справді $0 \le \dfrac{m(G \bigcap (G^* \setminus G_*))}{m(G^* \setminus G_*)} \le 1$, тому $P_n^*(G_*) \le P(G) \le$ $\le P_n^*(G_*) + P_n^*(G^* \setminus G_*) = P_n^*(G^*)$. Легко бачити, що коли $G = \bigcup\limits_{i \in I} H_i \in S$, $I \subset \{1, 2, ..., k\}$, тоді $G_* = G^* = G \in S$, $P(G) = P_n^*(G)$.

Зауважимо, що замість коефіцієнта $\dfrac{m(G \bigcap G^* \setminus G_*))}{m(G^* \setminus G_*)}$ біля $P_n^*(G^* \setminus G_*)$ в формулі (2.2.2) можна взяти довільне число $\alpha \in [0;1]$, наближено покладаючи таким чином $P(G) = P_n^*(G_*) + \alpha P_n^*(G^* \setminus G_*)$, $\alpha \in [0;1]$. Найчастіше покладають $\alpha = \dfrac{1}{2}$. В такому разі похибка наближеного значення міри $P(G)$ не перевищує $P_n^*(G^* \setminus G_*)$.

Однак в останньому випадку множина G має бути вимірною за мірою $\tilde{P}$, що вводиться наступним чином.

Поділимо множини H_i на r частин H_{ij}, $j \in \overline{1, r}$, таких, що $H_{ij} \neq \varnothing$, $H_{ij} \cdot H_{sl} = \varnothing$, коли $i \neq s$ або $j \neq l$ за $i = s$, $\bigcup_{j=1}^{r} H_{ij} = H_i$, і покладемо $P(H_{ij}) > 0$ так, щоб $\sum_{j=1}^{r} P(H_{ij}) = P_n^*(H_i)$, $\max_{1 \le j \le r} P(H_{ij}) \underset{r \to \infty}{\to} 0$ (найчастіше покладають $P(H_{ij}) = \frac{1}{r} P_n^*(H_i)$ за всіх $j \in \overline{1,\ r}$),

$S = \{A \mid A = \bigcup_{i \in I,\, j \in J} H_{ij}\ ,\ I \subset \{1, 2, ..., k\},\ \ J \subset \{1, 2, ..., r\}\}$,

$$G_r^* = \bigcap_{G \cap \Omega \subset \bigcup H_{ij}} \bigcup H_{ij}\ ,\ \ G_{*r} = \bigcup_{\bigcup H_{ij} \subset G \cap \Omega} \bigcup H_{ij}\ ,$$

$$P_r(G) = P_n^*(G_{*r}) + \alpha P_n^*(G_r^* \setminus G_{*r})\ ,\ \ \alpha \in [0;1]\ .$$

Очевидно, послідовність чисел $P_n^*(G_{*r})$ неспадна і обмежена зверху, а $P_n^*(G_r^*)$ незростаюча і обмежена знизу, тому обидві ці послідовності збіжні. В разі, коли виявляється $\lim_{r \to \infty} P_n^*(G_{*r}) = \lim_{r \to \infty} P_n^*(G_r^*) = \lim_{r \to \infty} P_r(G) = \tilde{P}(G)$, тоді множина G буде вимірною за так введеною мірою $\tilde{P}$, а $P(G) = P_n^*(G_{*r}) + \alpha P_n^*(G_r^* \setminus G_{*r})$, $\alpha \in [0;1]$, буде наближеним значенням міри $\tilde{P}(G) = \lim_{r \to \infty} P_n^*(G_{*r}) = \lim_{r \to \infty} P_n^*(G_r^*)$. Зрозуміло, що $P_n^*(G_{*r}) \le \tilde{P}(G) \le P_n^*(G_r^*)$ за довільного r. Разом з тим якщо виявиться

$$\lim_{r \to \infty} P_n^*(G_{*r}) \neq \lim_{r \to \infty} P_n^*(G_r^*)\ ,$$

то тоді множина G виявиться невимірною за так введеною мірою $\tilde{P}$, і застосування формули

$$P(G) = P_n^*(G_{*r}) + \alpha P_n^*(G_r^* \setminus G_{*r})\ ,\ \ \alpha \in [0;1]\ ,$$

для наближеного визначення значення $\tilde{P}(G)$ буде некоректним.

Зауважимо, що всі множини із сукупностей S_r, $r \in \{1, 2, 3, ...\}$, вимірні за мірою $\tilde{P}$.

Можна довести, що коли H_{ij}, $j \in \overline{1, r}$, рівні між собою рівносторонні n-вимірні паралелепіпеди в n-вимірному просторі R^n, діаметри яких спрямовуються до нуля, тоді за мірою $\tilde{P}$ будуть вимірні довільні опуклі множини та їх перетини і об'єднання в довільній кількості, множини з вимірними межами обмеженої міри тощо.

Міра P, введена на $\tilde{S}$ за формулою (2.2.2), є одним з можливих продовжень ймовірнісної міри P_n^* із сукупності S підмножин множини Ω на сукупність $\tilde{S}$, де S породжена за поділом множини Ω на підмножини H_i, $i \in \overline{1,k}$ такі, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$, $A = \bigcup_{i \in I} H_i \in S$, $I \subset \{1, 2, ..., k\}$, $\tilde{S}$ породжена за сукупністю вимірних за мірою m підмножин множини $\tilde{\Omega} \supset \Omega$, $\tilde{S} \supset S$. Крім того стосовно обох сукупностей S і $\tilde{S}$ задовільняються умови 1_S - 3_S, тобто S і $\tilde{S}$ є просторами подій, а (Ω, S, P_n^*) і $(\Omega, \tilde{S}, P)$ – ймовірнісні простори. Зауважимо, що

$$P_n^*(G_*) = \max_{\bigcup H_i \subset G \cap \Omega} P_n^*\left(\bigcup H_i\right) = \max_{\bigcup H_i \subset G \cap \Omega} \sum P_n^*(H_i),$$

$$P_n^*(G^*) = \min_{G \cap \Omega \subset \bigcup H_i} P_n^*\left(\bigcup H_i\right) = \min_{G \cap \Omega \subset \bigcup H_i} \sum P_n^*(H_i),$$

$$m(G_*) = \max_{\bigcup H_i \subset G \cap \Omega} m\left(\bigcup H_i\right) = \max_{\bigcup H_i \subset G \cap \Omega} \sum m(H_i),$$

$$m(G^*) = \min_{G \cap \Omega \subset \bigcup H_i} m\left(\bigcup H_i\right) = \min_{G \cap \Omega \subset \bigcup H_i} \sum m(H_i),$$

$$m(G^* \setminus G_*) = m(G^*) - m(G_*),$$

$$m(G \cap (G^* \setminus G_*)) \leq m(G^* \setminus G_*).$$

Зауважимо, що в разі, коли невідомо, як саме розподілені ймовірності $P_n^*(H_i)$ на множинах H_i, $i \in \overline{1,k}$, однозначно продовжити ймовірнісну міру $P_n^*(A)$, $A \in S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \overline{I,k}\}$, із простору подій S на ширший простір подій $\tilde{S}$ такий, що $S \subset \tilde{S}$, неможливо. Разом з тим за даними стосовно ймовірнісного простору (Ω, S, P_n^*) за формулою (2.2.2) можна визначити наближене значення ймовірнісної міри $P(G)$ множини $G \subset \Omega$, $G \overline{\in} S$

Надалі ймовірнісну міру події A, $A \in S$, введену на основі припущення, яке базується на результатах деякої серії із n

випробувань чи кількох таких серій, будемо позначати $P(A)$, $A \in S$, і називати *узагальненою статистичною ймовірністю*, або *гіпотетичною ймовірнісною мірою*, або *гіпотетичною ймовірністю*, або просто *ймовірністю*, якщо це не викликатиме суперечностей, невизначеності, необхідності додаткових пояснень.

Якщо ж гіпотетична ймовірнісна міра, задана на сукупності S підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, вводиться довільно, без проведення попередньо будь яких випробувань і збирання статистичних даних, лише вимагається, щоб стосовно такої ймовірнісної міри задовільнялися вимоги 1_p-3_p, тоді таку ймовірнісну міру також називатимемо *гіпотетичною ймовірнісною мірою або гіпотетичною ймовірністю* або просто ймовірністю і також позначатимемо $P(A)$, $A \in S$.

В такому разі ймовірнісний простір (Ω, S, P) не пов'язується з жодними конкретними випробуваннями і є лише теоретичною моделлю довільного ймовірнісного простору. Разом з тим всі теоретичні положення витікають із практики.

Вираз $\sum\limits_{i=1,\, E \in G \cap H_i}^{k} f(E) m(G \cap H_i)$ природно назвати інтегралом

від функції $f(E)$ на множині G за мірою m і позначити як $\int\limits_G f(E) m(dE)$.

Очевидно, стосовно функції $f(E)$ (як і $f_n^*(E)$) задовільняються такі властивості:

1. $f(E) \geq 0$.
2. $\int\limits_\Omega f(E) m(dE) = 1$.

Нехай, наприклад, на столі є чотири поля однакової міри m: H_1 – біле, H_2 – зелене, H_3 – червоне, H_4 – жовте. В ці поля навмання кидали 100 кульок. В результаті на білому полі виявилось 60 кульок, на зеленому – 30 кульок, на червоному – 6, на жовтому – 4 кульки.

Тоді природно вважати, що на білому полі кульки розташовані щільніше, ніж на зеленому, червоному, жовтому, на зеленому щільніше, ніж на червоному і жовтому, на червоному щільніше, ніж на жовтому.

Підкреслимо, що за наведених даних кількість кульок, що попали в будь які непорожні підмножини множин H_i, відмінні від самих H_i, визначити неможливо.

В розглядуваному випадку на множині Ω можна задати функцію

$$\hat{f}_{100}(E)=\begin{cases}\dfrac{60}{m}, & \text{коли } E\in H_1,\\ \dfrac{30}{m}, & \text{коли } E\in H_2,\\ \dfrac{6}{m}, & \text{коли } E\in H_3,\\ \dfrac{4}{m}, & \text{коли } E\in H_4,\end{cases}$$

за якою характеризуватиметься усереднена щільність відносно міри m розподілу кульок на множині Ω за полями H_i, $i\in\overline{1,4}$, якщо щільністю розподілу кульок на множині Ω називати кількість кульок, яка припадає на кожну одиницю міри множини Ω, де

$$m(\Omega)=m\left(\bigcup_{i=1}^{4}H_i\right)=m(H_1)+m(H_2)+m(H_3)+m(H_4),$$

а через функцію

$$f_m^*(E)=\begin{cases}\dfrac{60/100}{m}, & \text{коли } E\in H_1,\\ \dfrac{30/100}{m}, & \text{коли } E\in H_2,\\ \dfrac{6/100}{m}, & \text{коли } E\in H_3,\\ \dfrac{4/100}{m}, & \text{коли } E\in H_4,\end{cases}$$

характеризується усереднена щільність відносно міри m розподілу статистичних ймовірностей за полями H_i, $i\in\overline{1,4}$, на тій самій множині Ω.

Приклад 2.2.1. Нехай $\Omega=H_1+H_2+H_3+H_4+H_5+H_6+$ $+H_7$ і H_i є подіями із простору *S*, породженого за поділом H_1, H_2, H_3, H_4, H_5, H_6, H_7 множини Ω, $H_iH_j=\varnothing$, коли

$i \neq j$, $\bigcup_{i=1}^{7} H_i = \Omega$, $S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1, 2, 3, ..., 7\}\}$ (Рис. 2.2.2).

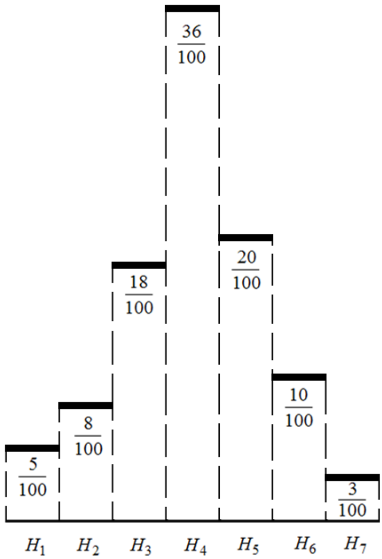


Рис. 2.2.2

Нехай проведено 100 випробувань, в кожному з яких із множини Ω навмання вибирали одну точку, і нехай одержано результати:

$$k_{100}(\Omega) = 100,\ k_{100}(H_1) = 5,\ k_{100}(H_2) = 8,\ k_{100}(H_3) = 18,$$
$$k_{100}(H_4) = 36,\ k_{100}(H_5) = 20,\ k_{100}(H_6) = 10,\ k_{100}(H_7) = 3.$$

Підкреслимо, що за наведених даних визначити кількість попадань в будь які підмножини множин H_i, відмінні від $\varnothing$ і самих H_i, неможливо.

Знайдемо статистичні ймовірності $P_{100}^*(H_i)$, $i \in \overline{1,7}$:

$$P_{100}^*(H_1) = \frac{5}{100},\ P_{100}^*(H_2) = \frac{8}{100},$$
$$P_{100}^*(H_3) = \frac{18}{100},\ P_{100}^*(H_4) = \frac{36}{100},$$
$$P_{100}^*(H_5) = \frac{20}{100},\ P_{100}^*(H_6) = \frac{10}{100},\ P_{100}^*(H_7) = \frac{3}{100}.$$

Якщо додатково задано міри $m(H_i) > 0$, $i \in \overline{1,7}$, тоді щільність відносно міри *m* розподілу статистичних ймовірностей за подіями H_i набуває вигляду

$$f_n^*(E) = \frac{P_n^*(H_i)}{m(H_i)}\,,\ E \in H_i\,,\ i \in \overline{1,k}\,.$$

Звідси

$$P_n^*(H_i) = f_n^*(E)\cdot m(H_i) = \int\limits_{H_i} f_n^*(E)\cdot m(dE)\,,\ E \in H_i\,.$$

Якщо $G \subset \widetilde{\Omega} = R^1$, $\widetilde{S} = B(R^1)$, $G \in \widetilde{S}$ – вимірна за мірою m множина, тоді за аналогією з попереднім (див. формулу 2.2.1) на $\widetilde{S}$ можна ввести ймовірнісну міру $P(G)$, $G \in \widetilde{S} \supset S$:

$$P(G) = \sum_{i=1}^{k} P(G \cap H_i) = \sum_{i=1}^{k} \int\limits_{G \cap H_i} f(E)m(dE) = \int\limits_{G} f(E)m(dE)\,.$$

Якщо Ω – відрізок числової осі довжиною 7 одиниць, H_i – інтервали довжиною 1, $m(H_i)$ – довжина інтервала H_i, тоді в геометричному тлумаченні $P_n^*(H_i) = f_n^*(E)\cdot m(H_i)$ буде площею прямокутника з основою H_i і висотою $f_n^*(E)$, тобто $P_n^*(H_i) = \int\limits_{H_i} f_n^*(x)dx$, а

$$P(G) = \sum_{i=1}^{k} P(G \cap H_i) = \sum_{i=1}^{k} f(E)m(G \cap H_i) = \int\limits_{G} f(E)m(dE) \ -$$

сума площ прямокутників з висотами $f(E)$, $E \in G \cap H_i$, над інтервалами $G \cap H_i$ (Рис. 2.2.2).

На рис. 2.2.2 подано графік функції $f_n^*(E)$, яка набуває сталих значень на проміжках H_i, $i \in \overline{1,7}$.

Приклад 2.2.2. Нехай в квадратну мішень (Рис. 2.2.3) виконано 100 пострілів, і одержано кількості $k_{100}(H_i)$ влучень в частини H_i, $i \in \overline{1,25}$, цієї мішені:

$k_{100}(H_1)=1,\quad k_{100}(H_2)=2,\quad k_{100}(H_3)=3,\quad k_{100}(H_4)=2,$
$k_{100}(H_5)=1,\quad k_{100}(H_6)=2,\quad k_{100}(H_7)=4,\quad k_{100}(H_8)=8,$
$k_{100}(H_9)=4,\quad k_{100}(H_{10})=2,\quad k_{100}(H_{11})=3,\quad k_{100}(H_{12})=8,$
$k_{100}(H_{13})=20,\quad k_{100}(H_{14})=8,\quad k_{100}(H_{15})=3,\quad k_{100}(H_{16})=2,$
$k_{100}(H_{17})=4,\quad k_{100}(H_{18})=8,\quad k_{100}(H_{19})=4,\quad k_{100}(H_{20})=2,$
$k_{100}(H_{21})=1,\quad k_{100}(H_{22})=2,\quad k_{100}(H_{23})=3,\quad k_{100}(H_{24})=2,$
$k_{100}(H_{25})=1.$

Нехай $\Omega=\bigcup_{i=1}^{25}H_i$, а частини H_i, $i\in\overline{1,25}$, мішені вважатимемо попарно несумісними подіями із простору S, породженого за поділом H_1, H_2, …, H_{25} множини Ω на підмножини H_i, $i\in\overline{1,25}$, $H_iH_j=\varnothing$, коли $i\neq j$, $S=\{A\,|\; A=\bigcup_{i\in I}H_i,\ I\subset\{1,\ 2,\ 3,\ \ldots,\ 25\}\}$. Тоді відповідні статистичні ймовірності дорівнюють:

$$P^*_{100}(H_1)=\frac{1}{100},\ P^*_{100}(H_2)=\frac{2}{100},\ P^*_{100}(H_3)=\frac{3}{100},\ P^*_{100}(H_4)=\frac{2}{100},$$
$$P^*_{100}(H_5)=\frac{1}{100},\ P^*_{100}(H_6)=\frac{2}{100},\ P^*_{100}(H_7)=\frac{4}{100},\ P^*_{100}(H_8)=\frac{8}{100},$$
$$P^*_{100}(H_9)=\frac{4}{100},\ P^*_{100}(H_{10})=\frac{2}{100},\ P^*_{100}(H_{11})=\frac{3}{100},\ P^*_{100}(H_{12})=\frac{8}{100},$$
$$P^*_{100}(H_{13})=\frac{20}{100},\ P^*_{100}(H_{14}=\frac{8}{100},\ P^*_{100}(H_{15})=\frac{3}{100},\ P^*_{100}(H_{16})=\frac{2}{100},$$
$$P^*_{100}(H_{17})=\frac{4}{100},\ P^*_{100}(H_{18})=\frac{8}{100},\ P^*_{100}(H_{19})=\frac{4}{100},\ P^*_{100}(H_{20})=\frac{2}{100},$$
$$P^*_{100}(H_{21})=\frac{1}{100},\ P^*_{100}(H_{22})=\frac{2}{100},\ P^*_{100}(H_{23})=\frac{3}{100},\ P^*_{100}(H_{24})=\frac{2}{100},$$
$$P^*_{100}(H_{25})=\frac{1}{100}.$$

Як видно з Рис. 2.2.3 а), б) точки влучення у множині H_{13} розташовані набагато щільніше, ніж в множинах H_1, H_2 і т.д. Тому природно назвати відношення статистичної ймовірності (а також і абсолютної частоти) попадання в множину H_i до міри цієї множини щільністю статистичної ймовірності (а також і точок влучення) у множині H_i відносно міри $m(H_i)$.

Таким чином, якщо

$$f_n^*(E) = \frac{P_n^*(H_i)}{m(H_i)}, \; E \in H_i, \; i \in \overline{1,25},$$

то $P_n^*(H_i) = f_n^*(E) \cdot m(H_i) = \int\limits_{H_i} f_n^*(E) \cdot m(dE)$, $E \in H_i$, а за довільного $G \subset \widetilde{\Omega}, \widetilde{\Omega} \supset \Omega$, вимірного за мірою m, $G \in \widetilde{S} \supset S$, за аналогією з попереднім (див. формули 2.2.1, 2.2.2) можна ввести нову ймовірнісну міру $P(G)$, визначену на $\widetilde{S}$:

H_{21}	H_{22}	H_{23}	H_{24}	H_{25}
H_{16}	H_{17}	H_{18}	H_{19}	H_{20}
H_{11}	H_{12}	H_{13}	H_{14}	H_{15}
H_6	H_7	H_8	H_9	H_{10}
H_1	H_2	H_3	H_4	H_5

Рис. 2.2.3 а)

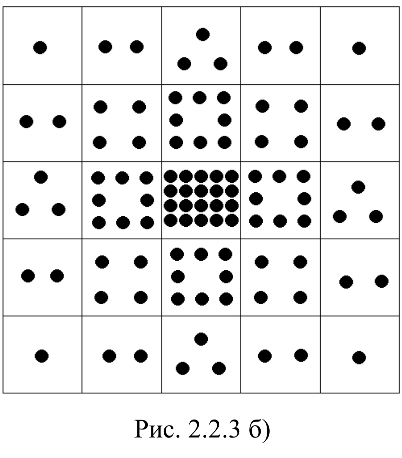


Рис. 2.2.3 б)

$$P(G)=\sum_{i=1}^{k}P(G\bigcap H_i)=\sum_{i=1}^{k}\int_{G\bigcap H_i}f(E)m(dE)=\int_G f(E)m(dE).$$

Якщо Ω квадрат із стороною довжиною 5 одиниць, H_i, $i\in\overline{1,25}$, – квадрат із стороною 1, тоді в геометричному тлумаченні $P_n^*(H_i)=f_n^*(E)\cdot m(H_i)$, $i\in\overline{1,25}$, – об'єм прямого паралелепіпеда із основою H_i і висотою $f_n^*(E)$, $E\in H_i$, тобто

$$P_n^*(H_i)=\int_{H_i}f_n^*(E)m(dE),$$

а $P(G)=\sum_{i=1}^{25}P(G\bigcap H_i)=\sum_{i=1,\,E\in G\bigcap H_i}^{25}f(E)\cdot m(G\bigcap H_i)=\int_G f(E)m(dE)$, $G\in\widetilde{S}$, – сума об'ємів прямих циліндрів з основами $G\bigcap H_i$, і висотами $f(E)$, $E\in G\bigcap H_i$ (зокрема паралелепіпедів, призм і т.п., твірні бічних поверхонь яких перпендикулярні до основ, а напрямні лінії твірних проходять вздовж меж основ) (Рис. 2.2.4). Зауважимо, що множина G не обов'язково має бути зв'язною.

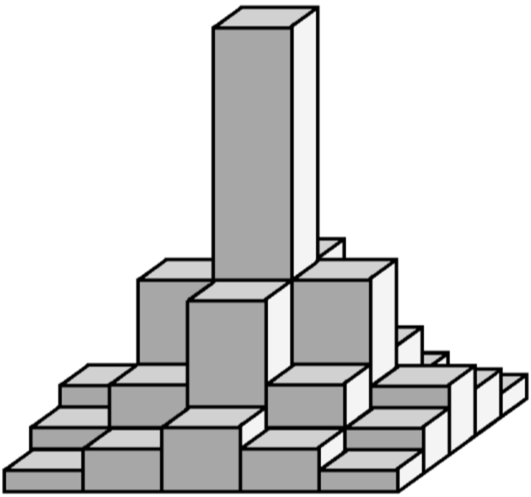


Рис. 2.2.4

Приклад 2.2.3. Нехай множина Ω скінченна, однак кількість елементів в ній дуже велика (див. приклад 1.8.2), $m(\Omega)$ – кількість елементів в множині Ω.

Поділимо Ω на k підмножин H_i додатної міри:

$$\Omega=\bigcup_{i=1}^{k}H_i,\ H_i\neq\varnothing,\ H_iH_j=\varnothing,\text{ коли } i\neq j,\ 0<m(\Omega)<\infty,$$

$$m(\Omega)=m(\bigcup_{i=1}^{k}H_i)=\sum_{i=1}^{k}m(H_i)\,.$$

Нехай простір подій S породжений за сукупністю таких підмножин H_i, $i\in\overline{1,k}$, множини Ω, тобто

$S=\{A \mid A=\bigcup_{i\in I}H_i,\ I\subset\{1,\ 2,\ 3,\ ...,\ k\}\}$, і нехай визначені $P_n^*(H_i)=\dfrac{k_n(H_i)}{k_n(\Omega)}$, $\widetilde{S}\supset S$ – сукупність вимірних за мірою m множин, стосовно якої задовільняються умови 1_s-3_s.

Нехай $G\in\widetilde{S}$ – вимірна за мірою m множина. За аналогією з попереднім (див. формули (2.2.1), (2.2.2)) на $\widetilde{S}$ можна ввести міру P, поклавши за $G\in\widetilde{S}$

$$P(G\bigcap H_i)=\frac{m(G\bigcap H_i)}{m(H_i)}P_n^*(H_i)\,,$$

$$P(G)=P\left(\bigcup_{i=1}^{k}G\bigcap H_i\right)=\sum_{i=1}^{k}P(G\bigcap H_i)=$$

$$=\sum_{i=1}^{k}\frac{m(G\bigcap H_i)}{m(H_i)}P_n^*(H_i)\,, \tag{2.2.3}$$

$f_n^*(E)=\dfrac{P_n^*(H_i)}{m(H_i)}$, $E\in H_i$, $P_n^*(H_i)=f_n^*(E)\cdot m(H_i)$, $E\in H_i$, $i\in\overline{1,k}$,

$$f(E)=\begin{cases}\dfrac{P(G\bigcap H_i)}{m(G\bigcap H_i)}=f_n^*(E), & \text{коли } E\in(G\bigcap H_i),\\ 0, & \text{коли } E\ \overline{\in}\ (G\bigcap H_i),\ i\in\overline{1,k},\end{cases}$$

$$P(G\bigcap H_i)=f(E)m(G\bigcap H_i)\,,\ E\in G\bigcap H_i,\ i\in\overline{1,k}\,,$$

$$P(A)=\sum_{i=1,\,E\in A\bigcap H_i}^{k}f(E)m(A\bigcap H_i)=\int_A f(E)m(dE)\,. \tag{2.2.4}$$

Очевидно, за $G\in S\subset\widetilde{S}$ буде $P(G)=P_n^*(G)$. Якщо $k=m(\Omega)$, а $m(H_i)=1$, тоді $f_n^*(E)$ буде статистичною ймовірністю попадання в одноелементну множину $\{E\}$ (в окрему точку $E\in\Omega$).

Приклад 2.2.4. В коробці є 100 кульок – білі, жовті, зелені, червоні, оранжеві, 20 кульок кожного кольору. В результаті деякої серії випробувань, в кожному з яких із коробки навмання діставали одну кульку, фіксували її колір, після чого її повертали в коробку, з’ясувалося, що білі кульки діставали в 5% всіх випробувань, жовті в 15% випробувань, зелені в 40%

випробувань, червоні в 25% випробувань, оранжеві – в 15% випробувань.

Нехай H_1 – множина білих кульок, H_2 – множина жовтих кульок, H_3 – множина зелених кульок, H_4 – множина червоних кульок, H_5 – множина оранжевих кульок. Нехай міра $m(H_i)$ множини H_i – кількість кульок в множині H_i.

Тоді $m(\Omega)=100$, $m(H_1)=20$, $m(H_2)=20$,

$$m(H_3)=20,\ m(H_4)=20,\ m(H_5)=20,$$

$$P_n^*(H_1)=0.05,\ P_n^*(H_2)=0.15,\ P_n^*(H_3)=0.40,$$

$$P_n^*(H_4)=0.25,\ P_n^*(H_5)=0.15.$$

$$f_n^*(E)=\begin{cases}\dfrac{0.05}{20}, & \text{коли } E\in H_1,\\ \dfrac{0.15}{20}, & \text{коли } E\in H_2,\\ \dfrac{0.40}{20}, & \text{коли } E\in H_3,\\ \dfrac{0.25}{20}, & \text{коли } E\in H_4,\\ \dfrac{0.15}{20}, & \text{коли } E\in H_5.\end{cases}$$

Таким чином кількість попадань в множину H_3 найбільша, а оскільки всі $m(H_i)$ однакові, то попадання у множину H_3 виявляються найчастішими, точки влучення у множині H_3 виявляються розташованими найщільніше в порівнянні з множинами H_1, H_2, H_4, H_5, тобто на кожну підмножину одиничної міри множини H_3 в середньому припадає більше точок влучення, ніж на кожну підмножину одиничної міри в множинах H_1, H_2, H_4, H_5.

Зауважимо, що в загальному випадку міри $m(H_i)$ множин H_i не обов'язково повинні бути однаковими.

Вправи для самостійного виконання

2.2.1. Перевірити чи правильні твердження:

1. Якщо $G\subset\tilde{\Omega}$, $G\in\tilde{S}$, то $P(G)=P_n^*(G)$.

2. $f_n^*(E)\ge f(E)$, коли $E\in G\bigcap H_i$, $i\in\overline{1,k}$, $H_iH_j=\varnothing$, коли $i\ne j$, $\bigcup\limits_{i=1}^{k}H_i=\Omega$, $S=\{A\mid A=\bigcup\limits_{i\in I}H_i,\ I\subset\{1,2,...,k\}\}$, $G\in\tilde{S}\supset S$.

3. Коли $G \in S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1, 2, ..., k\}\}$, тоді $G_* = G^*$.

4. За довільних $G \subset \Omega$ існують $m(G_*)$ і $m(G^*)$.

5. Різниця $m(G^*) - m(G_*)$ може бути як завгодно малою, коли $\max\limits_{1 \le i \le k} m(H_i) \underset{k \to \infty}{\to} 0$.

2.2.2. За заданих Ω, H_i, $i \in \overline{1, k}$, $S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1,\ 2,\ ...,\ k\}\}$, $P_n^*(H_i)$, $G \in \tilde{S}$, обчислити $m(G_*)$, $m(G^*)$, $P_n^*(G_*)$, $P_n^*(G^*)$, $P_n^*(G^* \setminus G_*)$:

1. $\Omega = [-2;\, 2]$, $H_i = [-2 + (i-1)h;\, -2 + i \cdot h)$, $h = 0.2$, $i \in \overline{1,\ 20}$,

$$P_n^*(H_i) = \frac{1}{20};\ G = [-1.3; 1.7);\ G = [-1.5; 1.5);\ G = [-0.8; 1.2).$$

2. $\Omega = \{(x, y) \mid x \in [-5; 5), y \in [-5; 5)\}$, $h = 0.5$,
$H_{i,j} = \{(x, y) \mid x \in [-5 + (j-1)h;\, -5 + j \cdot h), y \in [-5 + (i-1)h;\, -5 + i \cdot h),$
$i \in \overline{1,\ 20};\ j \in \overline{1,\ 20}\}$,

$$P_n^*(H_{ij}) = \frac{1}{400},\ G = \{(x, y) \mid x^2 + y^2 \le 16\}.$$

За заданих Ω, H_{ij}, $i \in \overline{1, k}$, $j \in \overline{1, k}$, $P_n^*(H_{ij})$, обчислити $P_n^*(G_*)$, $P_n^*(G^*)$, $P_n^*(G^* \setminus G_*)$.

2.3. Поінтервальний розподіл статистичних імовірностей на неперервній множині точок

Розглянемо нескінченну множину Ω елементарних подій, яку можна ототожнювати з числовим проміжком $\langle a;b\rangle$. Таку *множину* Ω називають *континуальною* або *неперервною.* В цьому випадку кожне число $x\in\langle a;b\rangle$, яке ототожнюється з певною елементарною подією, може бути спостереженим значенням в результаті експерименту, що пов'язаний з множиною Ω.

Приклад 2.3.1. Нехай випробування – це постріл в круглу мішень радіуса $r=1$, а кожна елементарна подія ототожнюється з відстанню точки влучення від центра мішені. Випробування налаштоване так, що попадання кулі за межі мішені неможливе. Тоді $\Omega=[0;1]$ – неперервна множина елементарних подій, кожна з яких ототожнюється з певним числом (точкою) $x\in[0;1]$.

У випадку неперервної множини $\Omega=\langle a;b\rangle$ елементарних подій подіями найчастіше вважають так звані вимірні (тобто такі, яким можна приписати довжину) підмножини множини Ω, до яких відносять, зокрема, будь які проміжки $\langle\alpha;\beta\rangle\subset\langle a;b\rangle$, а також об'єднання таких проміжків $\bigcup\limits_k\langle\alpha_k;\beta_k\rangle$ та різниці $\Omega\setminus\bigcup\limits_k\langle\alpha_k;\beta_k\rangle$. Як виявляється, в будь-якому проміжку $\langle a;b\rangle$ міститься невимірна підмножина (якій неможливо приписати довжину), а тому не завжди кожну підмножину $A\subset\Omega$ можна вважати подією.

Приклад 2.3.2. Якщо Ω з прикладу 2.3.1, то подія A може полягати у тому, що відстань точки влучення від центра мішені – число (точка), що є елементом певної вимірної підмножини $A\subset[0;1]$. Наприклад, такими подіями можуть бути підмножини

$$A=\left[\frac{1}{2};\frac{3}{4}\right],\ B=\left(0;\frac{1}{4}\right]\cup\left[\frac{3}{4};1\right).$$

Нехай проведено n випробувань (n досить велике), в результаті яких дістали спостережені значення $x_{\text{сп }1}, x_{\text{сп }2}, \ldots, x_{\text{сп }n}$, кожне з проміжка $\langle a;\ b\rangle=\Omega$.

Якщо підрахувати статистичні ймовірності $P_n^*(\{x_i\})$ за кожного x_i, що зустрічається серед спостережених значень $x_{\text{сп }i}$, $i\in\overline{1,n}$, то швидше за все стосовно більшості x_i дістанемо, що $P_n^*(\{x_i\})=0$ в межах точності обчислень, і таким чином формула $P_n^*(A)=\sum\limits_{x\in A}P_n^*(\{x\})$ є недостатньо коректною для обчислення статистичної ймовірності події $A \subset \Omega$, коли Ω – неперервна множина.

Тому у випадку неперервної множини $\Omega = \langle a;\, b\rangle$ в разі підрахунку статистичних ймовірностей (відносних частот) $P_n^*(A)$ поступають інакше, ніж у випадку, коли множина Ω скінченна і в ній є не дуже велика кількість точок.

Надалі вважатимемо, що $\Omega = [a;b)$.

Поділимо проміжок $\Omega = [a;b)$ точками $a_0 = a$, $a_i = a_{i-1} + h$, $h = \dfrac{b-a}{k}$, $i \in \overline{1,\, k}$, на k проміжків $[a_{i-1}; a_i)$ однакової довжини. Число $h = \dfrac{b-a}{k}$ називається *кроком поділу*. За кожного $i \in \overline{1,\, k}$ підрахуємо кількість спостережених значень $x_{\text{сп}\,j}$, що попадають в проміжок $[a_{i-1}; a_i)$. Дістанемо числа n_i – *абсолютні частоти попадання спостережених значень* $x_{\text{сп}\,j}$ *в проміжки* $[a_{i-1}; a_i)$, причому $\sum\limits_{i=1}^{k} n_i = n$. Число $P_n^*([a_{i-1};\, a_i)) = \dfrac{n_i}{n}$ називають *статистичною ймовірністю* або *відносною частотою* попадання спостережених значень $x_{\text{сп}\,j}$, $j \in \overline{1, n}$, в проміжок $[a_{i-1}; a_i)$, $i \in \overline{1, k}$. Як події будемо розглядати разом з $\varnothing$ та $\Omega = \bigcup\limits_{i=1}^{k} [a_{i-1};\, a_i)$ лише всеможливі об’єднання проміжків $[a_{i-1}; a_i)$, тобто $S = \{A \mid A = \bigcup\limits_{i \in I} H_i,\ I \subset \{1, 2, \ldots, k\}\}$. Тоді якщо $A = \bigcup\limits_{i \in I} [a_{i-1};\, a_i) \in S$, $I \subset \{1,\, 2, \ldots,\, k\}$, то $P_n^*(A) = \sum\limits_{i \in I} P_n^*([a_{i-1}; a_i))$.

За таблицями виду 2.3.1 задають *поінтервальний розподіл абсолютних частот*, а за таблицями виду 2.3.2 задають *поінтервальний розподіл статистичних ймовірностей* (*відносних частот*) на множині $\Omega = [a;b] = \bigcup\limits_{i=1}^{k} [a_{i-1}; a_i)$.

Табл. 2.3.1

$[a_{i-1}; a_i)$	$[a_0; a_1)$	$[a_1; a_2)$	...	$[a_{k-1}; a_k)$
n_i	n_1	n_2	...	n_k

Табл. 2.3.2

$[a_{i-1}; a_i)$	$[a_0; a_1)$	$[a_1; a_2)$	...	$[a_{k-1}; a_k)$
$P_n^*([a_{i-1}; a_i))$	$P_n^*([a_0; a_1))$	$P_n^*([a_1; a_2))$	...	$P_n^*([a_{k-1}; a_k))$

За довільного $<\alpha;\beta>$ можна покласти (див. §2.2, формулу 2.2.1)

$$P(<\alpha;\beta>) = P(<\alpha;\beta>\bigcap(\bigcup_{i=1}^{k}[a_{i-1};a_i))) =$$

$$= P\left(\bigcup_{i=1}^{k}<\alpha;\beta>\bigcap[a_{i-1};a_i)\right) = \sum_{i=1}^{k}P(<\alpha;\beta>\bigcap[a_{i-1};a_i)),$$

де $$P(<\alpha;\beta>\bigcap\ [a_{i-1};a_i)) = \frac{m(<\alpha;\beta>\bigcap[a_{a-1};a_i))}{m([a_{i-1};a_i))} P_n^*([a_{i-1};a_i)),$$

визначивши таким чином ймовірнісну міру P на просторі подій $\tilde{S}$, породженому за сукупністю всеможливих проміжків вигляду $<\alpha;\beta>$.

Зауважимо, що у випадку поінтервального розподілу статистичних ймовірностей використання формул (2.2.1) і (2.2.2) приводить до одного і того самого результату.

Тут припускається, що ймовірність $P([\alpha_i;\beta_i))$ попадання в довільний проміжок $<\alpha_i;\beta_i>\subset[a_{i-1};a_i)$ дорівнює $\frac{\beta_i-\alpha_i}{a_i-a_{i-1}}P_n^*([a_{i-1};a_i))$, тобто припускається, що на проміжках $[a_{i-1};a_i)$, $i\in\overline{1,k}$ *ймовірності* P *розподілені рівномірно*, що означає, що ймовірності P попадання в дві підмножини проміжка $[a_{i-1};a_i)$ однакової міри (довжини) однакові, $i\in\overline{1,k}$ (див. формули 2.2.1, 2.2.2).

Разом з тим в разі, коли невідомо, як саме розподілені статистичні ймовірності $P_n^*([a_{i-1},a_i))$ на кожному з інтервалів $[a_{i-1},a_i), i\in 1,k$, однозначно визначити ймовірнісну міру $P(<\alpha;\beta>)$ довільного проміжка неможливо.

Для складання таблиць поінтервального розподілу статистичних імовірностей, як і в дискретному випадку, спостережені значення записують у вигляді варіаційного ряду.

В фізичній інтерпретації поінтервальний розподіл абсолютних частот – це розподіл маси n на сукупності заданих проміжків $[a_{i-1};a_i)$, $i\in\overline{1,k}$, а поінтервальний розподіл узагальнених статистичних ймовірностей (відносних частот) – це розподіл одиничної маси на вказаній сукупності проміжків. На практиці найчастіше вважають, що розподіл узагальнених статистичних ймовірностей рівномірний на кожному із проміжків $[a_{i-1};a_i)$, $i\in\overline{1,k}$.

Приклад 2.3.3. Нехай $\Omega=[0;1]$ з прикладу 2.3.1 і виконано n=100 пострілів, результати яких наведено у таблицях 2.3.3 і

2.3.4. Ці таблиці є конкретними поінтервальними розподілами абсолютних і відносних частот.

Табл. 2.3.3

$[a_{i-1};a_i)$	$\left[0;\frac{1}{10}\right)$	$\left[\frac{1}{10};\frac{2}{10}\right)$	$\left[\frac{2}{10};\frac{3}{10}\right)$	$\left[\frac{3}{10};\frac{4}{10}\right)$	$\left[\frac{4}{10};\frac{5}{10}\right)$
n_i	50	20	10	10	4
	$\left[\frac{5}{10};\frac{6}{10}\right)$	$\left[\frac{6}{10};\frac{7}{10}\right)$	$\left[\frac{7}{10};\frac{8}{10}\right)$	$\left[\frac{8}{10};\frac{9}{10}\right)$	$\left[\frac{9}{10};1\right)$
	0	2	1	2	1

Табл. 2.3.4

$[a_{i-1};a_i)$	$\left[0;\frac{1}{10}\right)$	$\left[\frac{1}{10};\frac{2}{10}\right)$	$\left[\frac{2}{10};\frac{3}{10}\right)$	$\left[\frac{3}{10};\frac{4}{10}\right)$	$\left[\frac{4}{10};\frac{5}{10}\right)$
$P^*_{100}([a_{i-1};a_i))$	0.5	0.2	0.1	0.1	0.04
	$\left[\frac{5}{10};\frac{6}{10}\right)$	$\left[\frac{6}{10};\frac{7}{10}\right)$	$\left[\frac{7}{10};\frac{8}{10}\right)$	$\left[\frac{8}{10};\frac{9}{10}\right)$	$\left[\frac{9}{10};1\right)$
	0	0.02	0.01	0.02	0.01

Розглянемо функцію $f_n^*(x)$, що набуває нульового значення за межами проміжка $[a;b)$, тобто $f_n^*(x)=0$ за $x<a$ або $x\ge b$, і значень $\frac{1}{h}P_n^*([a_{i-1};a_i))$ на проміжках $[a_{i-1};a_i)$, $i\in\overline{1,k}$, де $h=\frac{b-a}{k}$, $a_0=a$, $a_i=a_{i-1}+h=a+i\cdot h$, $i\in\overline{1,k}$. Графік такої функції називають *гістограмою* поінтервального розподілу статистичних ймовірностей (відносних частот), який визначається за Табл. 2.3.2.

Функцію $f_n^*(x)$ називають *щільністю розподілу узагальнених статистичних ймовірностей (відносних частот) на проміжку* $[a;b)$ за інтервалами $[a_{i-1};a_i)$, $i\in\overline{1,k}$.

В фізичній інтерпретації $f_n^*(x)$ – це усереднена щільність розподілу одиничної маси на проміжку $[a;b)$, оскільки на кожному інтервалі $[a_{i-1};a_i)$ значення $f_n^*(x)$ одержується як маса $P_n^*([a_{i-1};a_i))$, що припадає на цей інтервал, поділена на довжину інтервала $h=a_i-a_{i-1}$, тобто як *середня щільність* маси $P_n^*([a_{i-1};a_i))$ на інтервалі $[a_{i-1};a_i)$.

Приклад 2.3.4. Якщо поінтервальний розподіл узагальнених статистичних ймовірностей задано за таблицею 2.3.4, то усереднена щільність цього розподілу набуває вигляду:

$$f_n^*(x)=\begin{cases} 0, & \text{коли } x<0, \\ 5, & \text{коли } 0\le x<0.1, \\ 2, & \text{коли } 0.1\le x<0.2, \\ 1, & \text{коли } 0.2\le x<0.3, \\ 1, & \text{коли } 0.3\le x<0.4, \\ 0.4, & \text{коли } 0.4\le x<0.5, \\ 0, & \text{коли } 0.5\le x<0.6, \\ 0.2, & \text{коли } 0.6\le x<0.7, \\ 0.1, & \text{коли } 0.7\le x<0.8, \\ 0.2, & \text{коли } 0.8\le x<0.9, \\ 0.1, & \text{коли } 0.9\le x<1, \\ 0, & \text{коли } x\ge 1. \end{cases}$$

На Рис. 2.3.1 зображено гістограму цього поінтервального розподілу статистичних ймовірностей на проміжку [0; 1). Оскільки за $x\in[a_{i-1};a_i)$ $P_n^*([a_{i-1};a_i))=f_n^*(x)h$ – це площа прямокутника з основою $[a_{i-1};a_i)$ і висотою $f_n^*(x)$, то

$$P_n^*([a_{i-1};\, a_i))=f_n^*(x)\cdot h=\int\limits_{[a_{i-1};\, a_i)} f_n^*(x)dx=\int\limits_{a_{i-1}}^{a_i} f_n^*(x)dx\,.$$

Отже з геометричної точки зору $P_n^*([a_{i-1};\, a_i))$ – це площа вказаного прямокутника, а статистична ймовірність $P_n^*(A)$ попадання спостережених значень в множину A, що є об’єднанням деяких проміжків $[a_{i-1};a_i)$, визначається за рівністю

$$P_n^*(A)=\sum_{[a_{i-1};a_i)\subset A} P_n^*([a_{i-1};a_i)) \;=\sum_{[a_{i-1};a_i)\subset A}\;\int\limits_{a_{i-1}}^{a_i} f_n^*(x)dx=\int\limits_A f_n^*(x)dx\,.$$

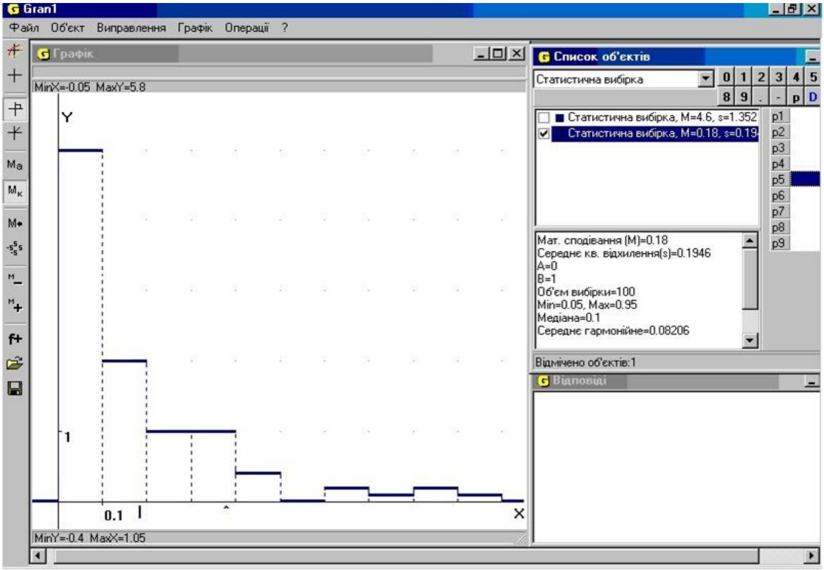


Рис. 2.3.1

Останнє число називають інтегралом від функції $f_n^*(x)$ на множині A за мірою, яку називають довжиною.

Може статися, що множина A складається з проміжків $<\alpha_j;\beta_j>$, що попарно не перетинаються і є частинами проміжків $[a_{i-1};a_i)$. Тоді вважають, що коли $<\alpha_j;\beta_j>\subset[a_{i-1};a_i)$, то узагальнена ймовірнісна міра P проміжка $<\alpha_j;\beta_j>$ дорівнює (див. формули 2.2.1, 2.2.2)

$$P(<\alpha_j;\beta_j>)=\frac{\beta_j-\alpha_j}{a_i-a_{i-1}}P_n^*([a_{i-1};a_i))=(\beta_j-\alpha_j)f(x),\ x\in\left\langle\alpha_j;\beta_j\right\rangle,$$

де $f(x)=\dfrac{P(<\alpha_j;\beta_j>)}{\beta_j-\alpha_j}=f_n^*(x)$, коли $x\in<\alpha_j;\beta_j>$, $f(x)=0$,

коли $x\overline{\in}<\alpha_j;\beta_j>$.

Це означає, що на кожному з проміжків $[a_{i-1};a_i)$ розподіл узагальнених (гіпотетичних) статистичних ймовірностей P *вважається рівномірним, тобто узагальнені (гіпотетичні) статистичні ймовірності попадання* в будь які вимірні підмножини проміжка $[a_{i-1};a_i)$ є однаковими, коли міри (довжини) таких підмножин однакові. В такому разі ймовірнісна міра попадання у множину $A = \bigcup_j < \alpha_j, \beta_j >$ дорівнює (див. формули 2.2.1, 2.2.2)

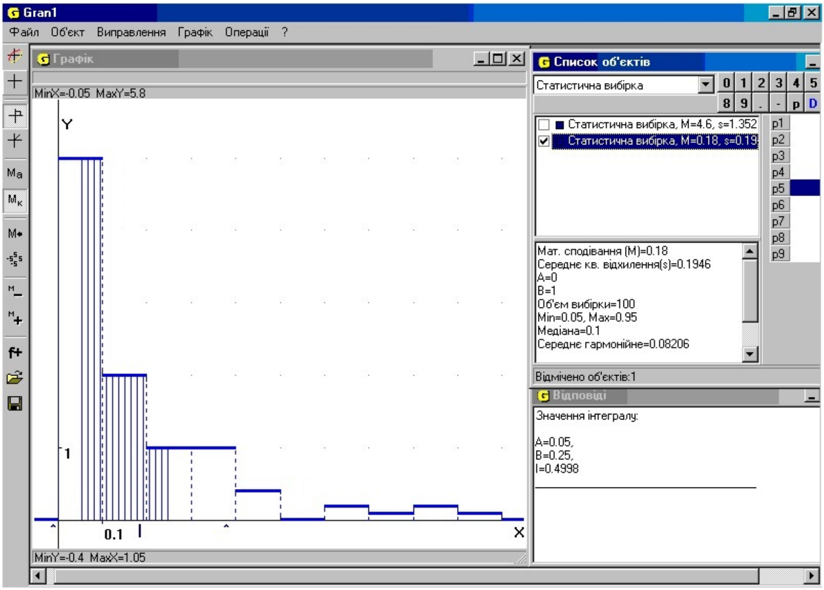


Рис. 2.3.2

$$P(A) = \sum_{<\alpha_j;\beta_j>\subset A} P(<\alpha_j;\beta_j>) = \sum_{\substack{<\alpha_j;\beta_j>\subset A \\ x\in<\alpha_j;\beta_j>}} f(x)(\beta_j - \alpha_j) = \int_A f(x)dx$$

і це число теж називають інтегралом від функції $f(x)$ на множині $A = \bigcup < \alpha_j;\beta_j >$, де $<\alpha_j;\beta_j> \subset [a_i;b_i)$ і попарно не перетинаються. Виявляється, що взагалі ймовірнісна міра (узагальнена статистична ймовірність) попадання в множину A дорівнює $P(A) = \int_A f(x)dx$ за будь-якої події (вимірної множини) $A \subset \Omega$, $A \in \tilde{S}$ (див. формули 2.2.1-2.2.4).

Приклад 2.3.5. Якщо в експерименті, описаному в прикладі 2.3.3, подія A – це попадання в множину точок, які віддалені від центра мішені на відстань, не більшу, ніж 0.25, і не меншу, ніж 0.05, то $P(A) = \int_{[0.05;0.25]} f(x)dx = 0.5 -$ інтеграл від функції $f(x)$ на проміжку $[0.05; 0.25]$, що дорівнює сумі площ заштрихованих прямокутників на Рис. 2.3.2, і значення цього інтеграла є узагальненою статистичною ймовірністю (ймовірнісною мірою) попадання в множину A, $A \in \widetilde{S}$.

Стосовно щільності $f(x)$ розподілу узагальнених статистичних ймовірностей (як і щільності $f_n^*(x)$ розподілу статистичних ймовірностей) задовільняються такі властивості:

1. $f(x) \geq 0$, $x \in (-\infty; \infty)$;
2. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$.

Вправи для самостійного виконання

2.3.1. Перевірити, чи правильні твердження:

1. Кожен простір Ω елементарних подій можна вважати неперервною множиною.
2. Неперервна множина Ω ототожнюється з певним проміжком $\langle a; b \rangle$.
3. За неперервної множини Ω подіями вважаються вимірні підмножини.
4. Кожна підмножина $A \subset [a;b)$ є вимірною, тобто подією.
5. За неперервної множини Ω статистичні ймовірності подій $A \subset \Omega$ визначаються так само, як і за дискретної множини Ω.
6. Поінтервальний розподіл статистичних ймовірностей можна задати через варіаційний ряд.
7. Щоб описати поінтервальний розподіл статистичних ймовірностей, підраховують абсолютні та відносні частоти попадання спостережених значень в проміжки, які попарно не перетинаються і об’єднання яких дорівнює Ω.
8. Гістограма – це графік щільності поінтервального розподілу статистичних ймовірностей.
9. Якщо відома щільність $f_n^*(x)$ поінтервального розподілу статистичних ймовірностей, то можна визначити статистичну ймовірність будь-якої події $A \in S$, $A \subset \Omega$.
10. Поінтервальний розподіл статистичних ймовірностей залежить від поділу проміжка $[a;b)$ на проміжки $[a_{i-1}; a_i)$.

2.3.2. 1. Побудувати варіаційний ряд спостережених значень відхилень точки падіння снаряда від цілі:

-52, -50, 47, 20, 17, 10, 14, 20,-50, -5, -20, 2, -27, -10, -15, 40, -19, 45,
34, -20, -29, -30, 15, -10, 10, 12, 14, 20, -33, -40, 50, 48, -10, 14, 10,
-32, 50, -22, 43, 2, -5, 7, 14, -8, -5, 0, 23, -40, 17, 3.

2. Використовуючи варіаційний ряд з 2.3.2.1, побудувати поінтервальні розподіли абсолютних і відносних частот (статистичних ймовірностей), взявши інтервали довжиною 10 з центрами в точках: -50, -40, -30, -20, -10, 0, 10, 20, 30, 40, 50.

3. За даними з 2.3.2.2 визначити функцію $f_n^*(x)$ і побудувати її графік.

4. Використовуючи $f_n^*(x)$ з 2.3.2.3, визначити ймовірнісну міру P_n^* влучення точки падіння снаряда в інтервал:

(-35; 35); б) (-25; 25); в) (-5; 55).

2.4. Деякі властивості узагальнених статистичних ймовірностей

Нехай $\Omega = [a;\ b) = \bigcup_{i=1}^{k} [a_{i-1};\ a_i)$, сукупність S подій породжена за поділом множини Ω на підмножини $[a_{i-1}; a_i)$, $i \in \overline{1,k}$, тобто $S = \{A \mid A = \bigcup_{i \in I} [a_{i-1};\ a_i),\ I \subset \{1, 2, \ldots, k\}\}\cdot$

Використовуючи функцію

$$f_n^*(x) = \begin{cases} \dfrac{P_n^*([a_{i-1}; a_i])}{a_i - a_{i-1}}, & \text{коли } x \in [a_{i-1}; a_i),\ i \in \overline{1,k}, \\ 0, & \text{коли } x \,\overline{\in}\, [a, b), \end{cases}$$

можна ввести нову ймовірнісну міру $\tilde{P}_n^*([\alpha; \beta))$ на просторі подій $\mathscr{B}(R^1)$ за довільних проміжків $[\alpha; \beta) \subset (-\infty; +\infty)$. За сукупністю таких проміжків породжується новий простір подій $\tilde{S} = \mathscr{B}(R^1)$, $\tilde{S} \supset S$.

Нехай $[\alpha; \beta) \subset (-\infty; \infty)$, $[\alpha; \beta) \in \tilde{S}$,

$$m_*([\alpha;\beta)) = \max_{\bigcup [a_{i-1};\ a_i) \subset [\alpha;\beta) \cap \Omega} P_n^*\left(\bigcup [a_{i-1}; a_i)\right),$$

$$m^*([\alpha;\beta)) = \min_{[\alpha;\beta) \cap \Omega \subset \bigcup [a_{i-1}; a_i)} P_n^*\left(\bigcup [a_{i-1}; a_i)\right).$$

Покладемо

$$\tilde{P}_n^*([\alpha;\beta)) = m_*([\alpha;\beta)) + \frac{1}{2}(m^*([\alpha;\beta)) - m_*([\alpha;\beta))) =$$

$$\frac{m_*([\alpha;\beta)) + m^*([\alpha;\beta))}{2}. \tag{2.4.1}$$

Очевидно

$$m^*([\alpha;\beta)) - m_*([\alpha;\beta) \le c \cdot 2h,$$

де $h = a_i - a_{i-1} > 0$, $i \in \overline{1,k}$, $c = \max_{x \in [a;b)} f_n^*(x) < \infty$,

$$P_n^*\left(\bigcup_{i \in I} [a_{i-1}; a_i)\right) = \sum_{i \in I} P_n^*([a_{i-1}; a_i)) = \sum_{i \in I} \int_{a_{i-1}}^{a_i} f_n^*(x)\,dx,\ I \subset \{1, 2, \ldots, k\}.$$

Легко бачити, що

$$m_*([\alpha;\beta)) \le \tilde{P}_n^*([\alpha;\beta)) \le m_*([\alpha;\beta)) + ch,$$

тобто

$$0 \le \widetilde{P}_n^*([\alpha\,;\beta)) - \max_{\bigcup[a_{i-1};\,a_i)\subset[\alpha;\beta)\bigcap\Omega} P_n^*\left(\bigcup[a_{i-1}\,;a_i)\right) \le ch\,.$$

Очевидно, коли $h \to 0$, то

$$m^*([\alpha\,;\beta)) - m_*([\alpha\,;\beta)) =$$

$$= \min_{[\alpha;\beta)\bigcap\Omega\subset\bigcup[a_{i-1};a_i)} P_n^*\left(\bigcup[a_{i-1}\,;a_i)\right) - \max_{\bigcup[a_{i-1};a_i)\subset[\alpha;\beta)\bigcap\Omega} P_n^*\left(\bigcup[a_{i-1}\,;a_i)\right) \to 0\,,$$

$$P_n^*([\alpha\,;\beta)) \to \max_{\bigcup[a_{i-1};a_i)\subset[\alpha;\beta)\bigcap\Omega} P_n^*\left(\bigcup[a_{i-1}\,;a_i)\right).$$

Оскільки за довільних $[\alpha;\beta)\subset(-\infty\,;\infty)$ $0 \le m_*([\alpha;\beta)) \le$ $\le m^*([\alpha;\beta)) \le 1$, то коли $m^*([\alpha;\beta)) = 1$, природно покласти $\widetilde{P}_n^*([\alpha;\beta)) = 1$. Зокрема коли $\Omega\subset[\alpha;\beta)$, тоді буде $m^*([\alpha;\beta)) = 1$. Якщо $[\alpha;\beta)\bigcap\Omega = \varnothing$, тоді природно покласти $m^*([\alpha;\beta)) = 0$, $\widetilde{P}_n^*([\alpha;\beta)) = 0$.

Використовуючи функцію $f_n^*(x)$, можна ввести ймовірнісну міру $\widetilde{P}_n^*$ на просторі подій $\widetilde{S}$ і в інший спосіб, поклавши $\widetilde{P}_n^*([\alpha;\beta)) = \int\limits_{\alpha}^{\beta} f_n^*(x)\,dx$ за довільних проміжків $[\alpha;\beta)\subset(-\infty\,;\infty)$, або

$$\widetilde{\mathrm{P}}_n^*([\alpha;\beta)) = \sum_{i=1}^{k} \frac{m([\alpha;\beta)\bigcap[a_{i-1};a_i))}{m([a_{i-1};a_i))} \mathrm{P}_n^*([a_{i-1};a_i)) =$$

$$= \sum_{i=1}^{k} \frac{m([\alpha;\beta)\bigcap[a_{i-1};a_i))}{m([a_{i-1};a_i))} f_n^*(x) m([a_{i-1};a_i)) =$$

$$= \sum_{\substack{i=1 \\ x\in[a_{i-1};a_i)}}^{k} f_n^*(x) m([\alpha;\beta)\bigcap[a_{i-1};a_i)) = \int\limits_{\alpha}^{\beta} f_n^*(x)dx$$

де $m([\alpha;\beta)) = \beta - \alpha$, $m([a_{i-1};a_i)) = a_i - a_{i-1}$.

Зауважимо, що в формулі (2.4.1) замість коефіцієнта $\frac{1}{2}$ можна взяти довільний коефіцієнт $q\in[0,\,1]$, тобто покласти $\widetilde{\mathrm{P}}_{\mathrm{n}}^*([\alpha;\beta)) = m_*([\alpha;\beta)) + q(m_*([\alpha;\beta)) - m_*([\alpha;\beta)))$. На практиці найчастіше покладають $q = \frac{1}{2}$.

Зауважимо, що із подрібненням інтервалів $[a_{i-1}\,;a_i)$ верхня оцінка $m^*([\alpha;\beta))$ міри $P_n^*([\alpha;\beta))$ не збільшується, а нижня оцінка $m_*([\alpha;\beta))$ не зменшується, тому границя значення $\lim\limits_{h\to 0} P_n^*([\alpha;\beta))$ завжди лежить в межах між $m_*([\alpha;\beta))$ і $m^*([\alpha;\beta))$, тобто $m_*([\alpha;\beta)) < \lim\limits_{h\to 0} P_n^*([\alpha;\beta)) \le m^*([\alpha;\beta))$, тому $m_*([\alpha;\beta))$ і $m^*([\alpha;\beta))$ є відповідно нижньою і верхньою оцінкою невідомої ймовірнісної міри $P_n^*([\alpha;\beta))$: $m_*([\alpha;\beta)) \le P_n^*([\alpha;\beta)) \le m^*([\alpha;\beta))$ за довільних $h \le a_i - a_{i-1}$

Ймовірнісну міру $\tilde{P}_n^*$, визначену на сукупності $\tilde{S} = B(R^1)$ вимірних за мірою $m([\alpha;\beta)) = \beta - \alpha$ підмножин $[\alpha;\beta)$ множини $\tilde{\Omega} = (-\infty;\infty)$ за мірою P_n^*, визначеною на сукупності *S*, будемо називати *узагальненою статистичною ймовірністю*, стосовно якої функція $f_n^*(x)$ також називається *щільністю розподілу узагальнених статистичних імовірностей*. Очевидно, що коли $A \in S$, тоді $\tilde{P}_n^*(A) = P_n^*(A)$.

Говорять, що міра $\tilde{P}_n^*$ є продовженням міри P_n^* із сукупності подій *S* на сукупність $\tilde{S}$.

Зауважимо, що міру P_n^* можна продовжити із сукупності *S* підмножин множини Ω на сукупність $\tilde{S}$ і за іншими способами, а не лише за вказаними (див. §2.2).

Приклад 2.4.1. Нехай задано поінтервальний розподіл статистичних імовірностей на множині $\Omega = [0;7)$ за інтервалами $[a_{i-1};a_i)$, $i \in \overline{1,7}$, $a_0 = 0$, $h = a_i - a_{i-1} = 1$:

Табл. 2.4.1

$[a_{i-1};a_i)$	[0; 1)	[1; 2)	[2; 3)	[3; 4)	[4; 5)	[5; 6)	[6; 7)
$P_n^*([a_{i-1};a_i))$	0.006	0.061	0.242	0.382	0.242	0.061	0.006

Тоді усереднена щільність $f_n^*(x)$ такого поінтервального розподілу ймовірностей набуває вигляду:

$$f_n^*(x)=\begin{cases}0.006, & \text{коли } x\in[0;1)\bigcup[6;7),\\ 0.061, & \text{коли } x\in[1;2)\bigcup[5;6),\\ 0.242, & \text{коли } x\in[2;3)\bigcup[4;5),\\ 0.382, & \text{коли } x\in[3;4),\\ 0, & \text{коли } x\overline{\in}[0;7).\end{cases}$$

Графік функції $f_n^*(x)$ подано на Рис. 2.4.1.

Якщо $[\alpha;\beta)=[1.7;5.3)$, тоді за попереднім

$$m_*=\max_{\bigcup[a_{i-1};a_i)\subset[\alpha;\beta)\bigcap\Omega} P_n^*(\bigcup[a_{i-1};a_i))=P_n^*([2;3)\bigcup[3;4)\bigcup[4;5))=$$

$$=P_n^*([2;3))+P_n^*([3;4))+P_n^*([4;5))=0.242+0.382+0.242=0.866;$$

$$m^*=\min_{[\alpha;\beta)\bigcap\Omega\subset\bigcup[a_{i-1};a_i)} P_n^*(\bigcup[a_{i-1};a_i))=P_n^*([1;2)\bigcup[2;3)\bigcup[3;4)\bigcup[4;5)\bigcup$$

$$\bigcup[5;6))=P_n^*([1;2))+P_n^*([2;3))+P_n^*([3;4))+P_n^*([4;5))+P_n^*([5;6))=$$

$$=0.061+0.242+0.382+0.242+0.061=0.988;$$

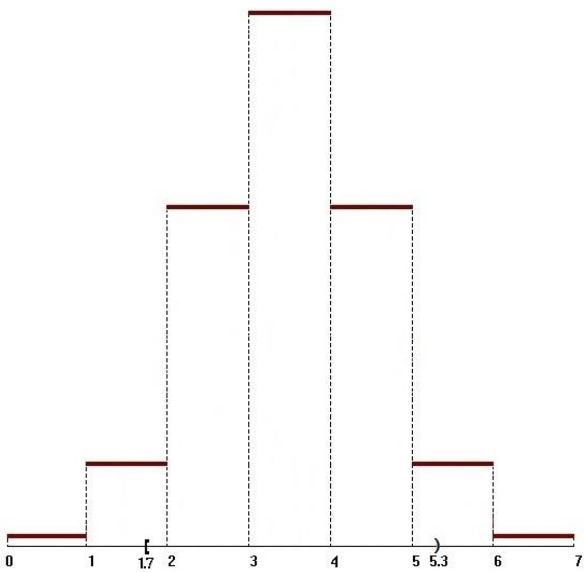


Рис. 2.4.1

$m^* - m_* = 0.988 - 0.866 = 0.122$;

$\widetilde{P}_n^*([1.7; 5.3)) = m_* + \frac{1}{2}(m^* - m_*) = 0.866 + 0.061 = 0.927$.

Зауважимо, що поінтервальний розподіл статистичних імовірностей, наведений в Табл. 2.4.1, міг бути отриманий за розподілом, поданим в Табл. 2.4.2, чи за будь яким іншим розподілом таким, що статистичні ймовірності попадання в інтервали $[a_{i-1}; a_i)$, $i \in \overline{1,7}$, $a_0 = 0$, $h = a_i - a_{i-1} = 1$, вказані в Табл. 2.4.1, набувають вказаних в цій таблиці значень.

Приклад 2.4.2. Нехай задано поінтервальний розподіл статистичних імовіностей на множині $\Omega = [0; 7)$ за інтервалами $[a_{i-1}; a_i)$, $i \in \overline{1,35}$, $a_0 = 0$, $h = a_i - a_{i-1} = 0.2$:

Табл. 2.4.2

$[a_{i-1}; a_i)$	$P_n^*[a_{i-1}; a_i))$
[0; 0.2)	0.000
[0.2; 0.4)	0.000
[0.4; 0.6)	0.001
[0.6; 0.8)	0.002
[0.8; 1.0)	0.003
[1.0; 1.2)	0.005
[1.2; 1.4)	0.007
[1.4; 1.6)	0.011
[1.6; 1.8)	0.016
[1.8; 2.0)	0.022
[2.0; 2.2)	0.030
[2.2; 2.4)	0.039
[2.4; 2.6)	0.048
[2.6; 2.8)	0.058
[2.8; 3.0)	0.067
[3.0; 3.2)	0.073
[3.2; 3.4)	0.078
[3.4; 3.6)	0.080
[3.6; 3.8)	0.078
[3.8; 4.0)	0.073
[4.0; 4.2)	0.067
[4.2; 4.4)	0.058
[4.4; 4.6)	0.048
[4.6; 4.8)	0.039
[4.8; 5.0]	0.030
[5.0; 5.2)	0.022
[5.2; 5.4)	0.016
[5.4; 5.6)	0.011
[5.6; 5.8)	0.007
[5.8; 6.0)	0.005
[6.0; 6.2)	0.003
[6.2; 6.4)	0.002
[6.4; 6.6)	0.001
[6.6; 6.8)	0.000
[6.8; 7.0)	0.000

Тоді усереднена щільність $f_n^*(x)$ такого поінтервального розподілу статистичних імовірностей набуває вигляду

$$f_n^*(x) = \begin{cases} 0.000, \text{ коли } & x \in [0; 0.2) \bigcup [0.2; 0.4) \bigcup [6.6; 6.8) \bigcup [6.8; 7.0) \\ 0.005, \text{ коли } & x \in [0.4; 0.6) \bigcup [6.4; 6.6), \\ 0.010, \text{ коли } & x \in [0.6; 0.8) \bigcup [6.2; 6.4), \\ 0.015, \text{ коли } & x \in [0.8; 1.0) \bigcup [6.0; 6.2), \\ 0.025, \text{ коли } & x \in [1.0; 1.2) \bigcup [5.8; 6.0), \\ 0.035, \text{ коли } & x \in [1.2; 1.4) \bigcup [5.6; 5.8), \\ 0.055, \text{ коли } & x \in [1.4; 1.6) \bigcup [5.4; 5.6), \\ 0.080, \text{ коли } & x \in [1.6; 1.8) \bigcup [5.2; 5.4), \\ 0.110, \text{ коли } & x \in [1.8; 2.0) \bigcup [5.0; 5.2), \\ 0.150, \text{ коли } & x \in [2.0; 2.2) \bigcup [4.8; 5.0), \\ 0.195, \text{ коли } & x \in [2.2; 2.4) \bigcup [4.6; 4.8), \\ 0.240, \text{ коли } & x \in [2.4; 2.6) \bigcup [4.4; 4.6), \\ 0.290, \text{ коли } & x \in [2.6; 2.8) \bigcup [4.2; 4.4), \\ 0.335, \text{ коли } & x \in [2.8; 3.0) \bigcup [4.0; 4.2), \\ 0.365, \text{ коли } & x \in [3.0; 3.2) \bigcup [3.8; 4.0), \\ 0.390, \text{ коли } & x \in [3.2; 3.4) \bigcup [3.6; 3.8), \\ 0.400, \text{ коли } & x \in [3.4; 3.6), \end{cases}$$

Графік цієї функції $f_n^*(x)$ подано на Рис. 2.4.2.

Якщо $[\alpha; \beta) = [1.7; 5.3)$, тоді

$$m_* = \max_{\bigcup [a_{i-1}; a_i) \subset [\alpha; \beta) \bigcap \Omega} P_n^*(\bigcup [a_{i-1}; a_i)) = P_n^*([1.8; 2.0) \bigcup [2.0; 2.2) \bigcup$$
$$\bigcup [2.2; 2.4) \bigcup [2.4; 2.6) \bigcup [2.6; 2.8) \bigcup [2.8; 3.0) \bigcup [3.0; 3.2) \bigcup [3.2; 3.4) \bigcup$$
$$\bigcup [3.4; 3.6) \bigcup [3.6; 3.8) \bigcup [3.8; 4.0) \bigcup [4.0; 4.2) \bigcup [4.2; 4.4) \bigcup [4.4; 4.6) \bigcup$$
$$\bigcup [4.6; 4.8) \bigcup [4.8; 5.0) \bigcup [5.0; 5.2)) = 0.022 + 0.030 + 0.039 + 0.048 +$$
$$+ 0.058 + 0.067 + 0.073 + 0.078 + 0.080 + 0.078 + 0.073 + 0.067 +$$
$$+ 0.058 + 0.048 + 0.039 + 0.030 + 0.022 = 0.910;$$

$$m^* = \min_{[\alpha; \beta) \bigcap \Omega \subset \bigcup [a_{i-1}; a_i)} P_n^*(\bigcup [a_{i-1}; a_i)) = P_n^*([1;6; 1.8) \bigcup [1.8; 2.0) \bigcup$$
$$\bigcup [2.0; 2.2) \bigcup [2.2; 2.4) \bigcup [2.4; 2.6) \bigcup [2.6; 2.8) \bigcup [2.8; 3.0) \bigcup [3.0; 3.2) \bigcup$$
$$\bigcup [3.2; 3.4) \bigcup [3.4; 3.6) \bigcup [3.6; 3.8) \bigcup [3.8; 4.0) \bigcup [4.0; 4.2) \bigcup [4.2; 4.4) \bigcup$$
$$\bigcup [4.4; 4.6) \bigcup [4.6; 4.8) \bigcup [4.8; 5.0) \bigcup [5.0; 5.2) \bigcup [5.2; 5.4)) =$$

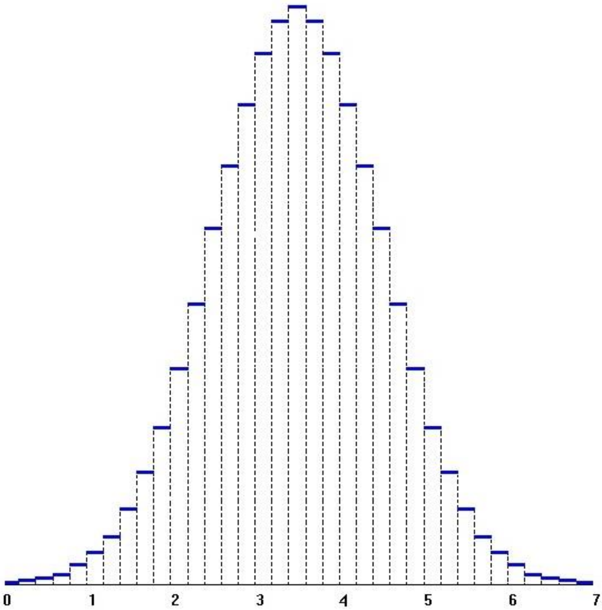


Рис. 2.4.2

$= 0.016 + 0.022 + 0.030 + 0.039 + 0.048 + 0.058 + 0.067 + 0.073 +$
$+ 0.078 + 0.080 + 0.078 + 0.073 + 0.067 + 0.058 + 0.048 + 0.039 +$
$+ 0.030 + 0.022 + 0.016 = 0.942$;

$m^* - m_* = 0.942 - 0.910 = 0.032$;

$\widetilde{P}_n^*([1.7; 5.3)) = m_* + \frac{1}{2}(m^* - m_*) = 0.910 + 0.016 = 0.926$.

Оскільки $m_* \le m^*$, бо $\bigcup_{i \in I_1} [a_{i-1}; a_i) \subset [\alpha; \beta) \bigcap \Omega \subset \bigcup_{i \in I_2} [a_{i-1}; a_i)$, де $I_1 \subset \{1, 2, ..., k\}$, $I_2 \subset \{1, 2, ..., k\}$, $I_1 \subset I_2$, то

$$m_* \le \widetilde{P}_n^*([\alpha; \beta)) \le \frac{1}{2}(m_* + m^*) \le m^* .$$

Як видно із прикладів 2.4.1, 2.4.2, із зменшенням довжини $h = a_i - a_{i-1}$ інтервала $[a_{i-1}; a_i)$, $i \in \overline{1, k}$, різниця $m^* - m_*$ також зменшується.

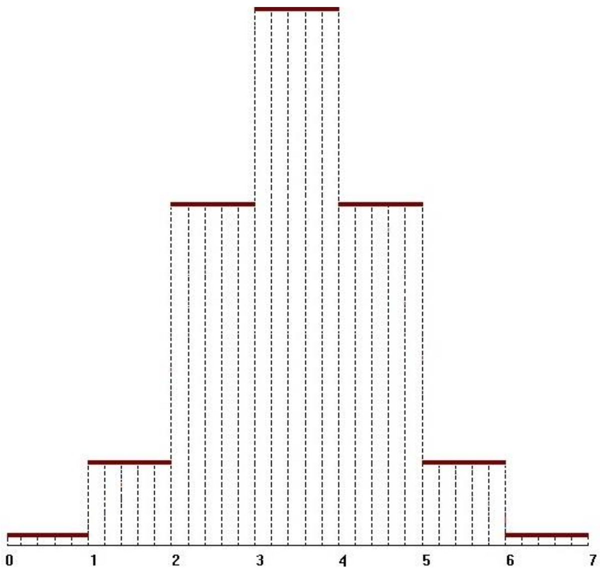


Рис. 2.4.3

Приклад 2.4.3. Нехай задано поінтервальний розподіл статистичних імовірностей на множині $\Omega=[0;\,7]$ за інтервалами $[a_{i-1};a_i)$, $i\in\overline{1,35}$, $a_0=0$, $h=a_i-a_{i-1}=0.2$, усереднена щільність $f_n^*(x)$ якого набуває вигляду (як і в прикладі 2.4.1):

$$f_n^*(x)=\begin{cases}0.006, \text{ коли } x\in[0;1)\bigcup[6;7),\\ 0.061, \text{ коли } x\in[1;2)\bigcup[5;6),\\ 0.242, \text{ коли } x\in[2;3)\bigcup[4;5),\\ 0.382, \text{ коли } x\in[3;4),\end{cases}$$

Нехай на інтервалах $[\tilde{a}_{i-1};\tilde{a}_i)$, $i\in\overline{1,7}$, $\tilde{a}_0=0$, довжиною $\tilde{h}=\tilde{a}_i-\tilde{a}_{i-1}=1$ статистичні ймовірності розподілені рівномірно за інтервалами $[a_{i-1};a_i)$ довжиною $h=a_i-a_{i-1}=0.2$, тобто якщо $[a_{i-1};a_i)\subset[\tilde{a}_{k-1};\tilde{a}_k)$, $k\in\overline{1,7}$, $[a_{j-1};a_j)\subset[\tilde{a}_{k-1};\tilde{a}_k)$, $i\neq j$, то $P_n^*([a_{i-1};a_i))=P_n^*([a_{j-1};a_j))$.

Графік розглядуваної функції подано на Рис. 2.4.3.

Якщо $[\alpha;\beta)=[1.7;5.3)$, тоді (див. Рис. 2.4.3, а також приклади 2.4.1, 2.4.2) :

$$m_*=\max_{\bigcup[a_{i-1};\,a_i)\subset[\alpha;\,\beta)\bigcap\Omega}P_n^*\Big(\bigcup[a_{i-1};a_i)\Big)=P_n^*([1.8;2.0)\bigcup[2.0;3.0)\bigcup$$

$$\bigcup[3.0;\,4.0)\bigcup[4.0;\,5.0)\bigcup[5.0;\,5.2) = P_n^*([1.8;\,2.0)) + P_n^*([2.0;\,3.0)) +$$
$$+ P_n^*([3.0;\,4.0)) + P_n^*([4.0;\,5.0)) + P_n^*([5.\,0;\,5\ ;2)) = 0.0122 + 0.2420 +$$
$$+ 0.3820 + 0.2420 + 0.0122 = 0.0244 + 0.8660 = 0.8904\ ;$$
$$m^* = \min_{[\alpha;\beta)\bigcap\Omega\subset\bigcup[a_{i-1};\,a_i)} P_n^*(\bigcup[a_{i-1};\,a_i)) = P_n^*([1;6;\,1.8)\bigcup[1.8;2.0)\bigcup$$
$$\bigcup[2.0;\,3.0)\bigcup[3.0;\,4.0)\bigcup[4.0;\,5.0)\bigcup[5.0;\,5.2)\bigcup[5.2;\,5.4)) = 0.0122 +$$
$$+ 0.0122 + 0.2420 + 0.3820 + 0.2420 + 0.0122 + 0.0122 = 0.0488 +$$
$$+ 0.8660 = 0.9148\ ;$$
$$m^* - m_* = 0.9148 - 0.8904 = 0.0244\ ;$$
$$\tilde{P}_n^*([1.7;\,5.3)) = m_* + \frac{1}{2}(m^* - m_*) = 0.8904 + 0.0122 = 0.9026.$$

Приклад 2.4.4. Нехай задано поінтервальний розподіл статистичних ймовірностей на множині $\Omega = [0;\,7]$ такий самий, як і в прикладі 2.4.3, і нехай $[\alpha;\,\beta) = [1;\,6)$.

Тоді
$$m_* = \max_{\bigcup[a_{i-1};\,a_i)\subset[\alpha;\,\beta)\bigcap\Omega} P_n^*(\bigcup[a_{i-1};\,a_i)) = P_n^*([1;\,2)\bigcup[2;\,3)\bigcup[3;\,4)\bigcup[4;\,5)\bigcup$$
$$\bigcup[5;6)) = 0.061 + 0.242 + 0.0382 + 0.242 + 0.061 = 0.988,$$
оскільки
$$[1;\,2)\bigcup[2;\,3)\bigcup[3;\,4)\bigcup[4;\,5)\bigcup[5;\,6) \subset [1;\,6),$$
$$m^* = \min_{[\alpha;\,\beta)\bigcap\Omega\subset\bigcup[a_{i-1};\,a_i)} P_n^*(\bigcup[a_{i-1};\,a_i)) = P_n^*([1;\,2)\bigcup[2;\,3)\bigcup[3;\,4)\bigcup[4;\,5)\bigcup$$
$$\bigcup[5;6)) = 0.061 + 0.242 + 0.0382 + 0.242 + 0.061 = 0.988,$$
оскільки
$$[1;\,6) \subset [1;\,2)\bigcup[2;\,3)\bigcup[3;\,4)\bigcup[4;\,5)\bigcup[5;\,6).$$
Таким чином $m_* = m^*$.

В розглядуваному випадку
$$[1;\,6) = [1;\,2)\bigcup[2;\,3)\bigcup[3;\,4)\bigcup[4;\,5)\bigcup[5;\,6) \in S\ ,$$
$$\tilde{P}_n^*([1;\,6)) = m_* + \frac{1}{2}(m^* - m_*) = m_* = m^* = P_n^*([1,6)) = 0.988\ .$$

Оскільки за довільних $[\alpha;\,\beta) \subset (-\infty;\,\infty)$ $0 \le m_* \le m^* \le 1$, то коли $m_* = 1$, природно покласти $\tilde{P}_n^*([\alpha;\,\beta)) = 1$. Зокрема коли $\Omega \subset ([\alpha;\,\beta)$, тоді буде $m_* = 1$. Якщо $[\alpha;\,\beta)\bigcap\Omega = \varnothing$, тоді природно покласти $m^* = 0$, $\tilde{P}_n^*([\alpha;\,\beta)) = 0$.

Зауважимо, що ймовірнісні міри P_n^*, розглянуті в прикладах

2.4.1, 2.4.3, є гіпотетичні, знайдені за припущенням, що щільність $f_n^*(x)$ розподілу статистичних ймовірностей на проміжках [0;1), [1;2), [2;3), [3;4), [4;5), [5;6), [6;7), набуває сталих значень відповідно 0.006, 0.061, 0.242, 0.382, 0.242, 0.061, 0.006, і значення 0 за межами проміжка [0;7), що в дійсності може виявитись не так (див. Рис. 2.4.2).

Якщо статистичні ймовірності попадання в будь-які підмножини V і W множини G однакові, коли міри цих підмножин однакові, тобто $P_n^*(V) = P_n^*(W)$, *коли* $m(V) = m(W)$, *тоді розподіл статистичних ймовірностей на множині G називається рівномірним*.

Якщо множина *G* скінченна, тоді як міру $m(V)$ підмножини $V \subset G$ розглядатимемо кількість елементів в множині *V*, якщо ж множина *G* є деякий відрізок на числовій прямій чи деяка область на площині (в двохвимірному просторі) чи в просторі (тривимірному), тоді як міру $m(V)$ множини $V \subset G$ розглядатимемо відповідно суму довжин інтервалів, з яких складено *V*, чи суму площ областей, з яких складено *V*, чи суму об'ємів просторових областей, з яких складено *V*.

Зокрема, розподіли статистичних ймовірностей на множинах Ω, розглянутих в прикладах 2.4.1-2.4.3, не є рівномірними.

Якщо розглянути інтервали $[i-1; i)$, $i \in \overline{1,7}$, із прикладу 2.4.1 і поділити кожен з них на 5 інтервалів довжиною 0.2 , як це зроблено в прикладі 2.4.3, то статистичні ймовірності попадання в інтервали $[i-1; i)$ довжиною 1 залишаються такими самими, як і задані в поінтервальному розподілі із прикладу 2.4.1, і (за припущенням) усереднена щільність $f_n^*(x)$ розподілу статистичних ймовірностей на множині $\Omega = [0; 7]$ на кожному інтервалі $[i-1; i)$, $i \in \overline{1,7}$, набуває сталих значень, відповідно 0.006, 0.061, 0.242, 0.382, 0.242, 0.061, 0.006, тобто за припущенням статистичні ймовірності на кожному інтервалі $[i-1; i)$, $i \in \overline{1,7}$ розподілені рівномірно із щільністю $f_n^*(x)$, а тому (за припущенням) статистичні імовірності попадання в дві, підмножини $[\alpha; \beta) \subset [i-1; i)$ і $[\gamma; \delta) \subset [i-1; i)$, $i \in \overline{1,7}$, складені із підінтервалів довжиною 0.2 , однакові, тобто

$P_n^*([\alpha;\beta)) = P_n^*([\gamma;\delta))$, коли міри (довжини) цих множин $[\alpha;\beta)$ і $[\gamma;\delta)$ однакові – $m([\alpha;\beta)) = m([\gamma;\delta))$, або, що те саме, однакові кількості інтервалів довжиною 0.2, з яких складені множини $[\alpha;\beta)$ і $[\gamma;\delta)$.

Однак припущення про рівномірність розподілу статистичних ймовірностей на інтервалах $[i-1;i)$, $i \in \overline{1,7}$, може виявитись таким, що не узгоджується із реальним розподілом статистичних ймовірностей на множині $\Omega = \bigcup_{i=1}^{7}[i-1;i)$ (див. приклад 2.4.2).

На Рис. 2.4.5 подано графік усередненої щільності $f_n^*(x)$ розподілу статистичних ймовірностей на множині $\Omega = [0;5)$ за інтервалами $[a_{i-1};a_i)$, $i \in \overline{1,25}$, довжиною 0.2, за припущення, що на інтервалах $[a_{i-1};a_i)$, $i \in \overline{1,25}$, статистичні ймовірності розподілені рівномірно.

Можна навести багато інших варіантів розподілів статистичних ймовірностей, за якими одержується один і той самий поінтервальний розподіл статистичних ймовірностей за інтервалами $[a_{i-1};a_i)$ з однією і тією самою усередненою щільністю $f_n^*(x)$ розподілу статистичних ймовірностей, наведеною в прикладах 2.4.1, 2.4.3, 2.4.5.

Зокрема, може виявитись, що поінтервальний розподіл статистичних ймовірностей, розглянутий в прикладі 2.4.5, одержаний за дискретним розподілом статистичних ймовірностей на множині $\{0.5; 1.5; 2.5; 3.5; 4.5\}$ із статистичними ймовірностями попадання у вказані точки відповідно 0.05, 0.20, 0.50, 0.20, 0.05, або на множині {0.20, 0.40, 0.60, 0.80, 1.00, 1.20, 1.40, 1.60, 1.80, 2.00, 2.20, 2.40, 2.60, 2.80, 3.00, 3.20, 3.40, 3.60, 3.80, 4.00, 4.20, 4.40, 4.60, 4.80, 5.00} із статистичними ймовірностями попадання у вказані точки відповідно 0.01, 0.01, 0.01, 0.01, 0.01, 0.04, 0.04, 0.04,0.04, 0.04, 0.10, 0.10, 0.10, 0.10, 0.10, 0.04, 0.04, 0.04, 0.04, 0.04, 0.01, 0.01, 0.01, 0.01, 0.01. Очевидно, можливі і інші розподіли статистичних ймовірностей на множині $\Omega = [0;5)$, за яких буде одержуватися один і той самий поінтервальний розподіл статистичних ймовірностей на множині $\Omega = [0;5)$ за інтервалами $[i-1;i)$, $i \in \overline{1,5}$, з однією і тією самою усередненою щільністю $f_n^*(x)$ розподілу статистичних ймовірностей, наведеною в прикладі 2.4.5.

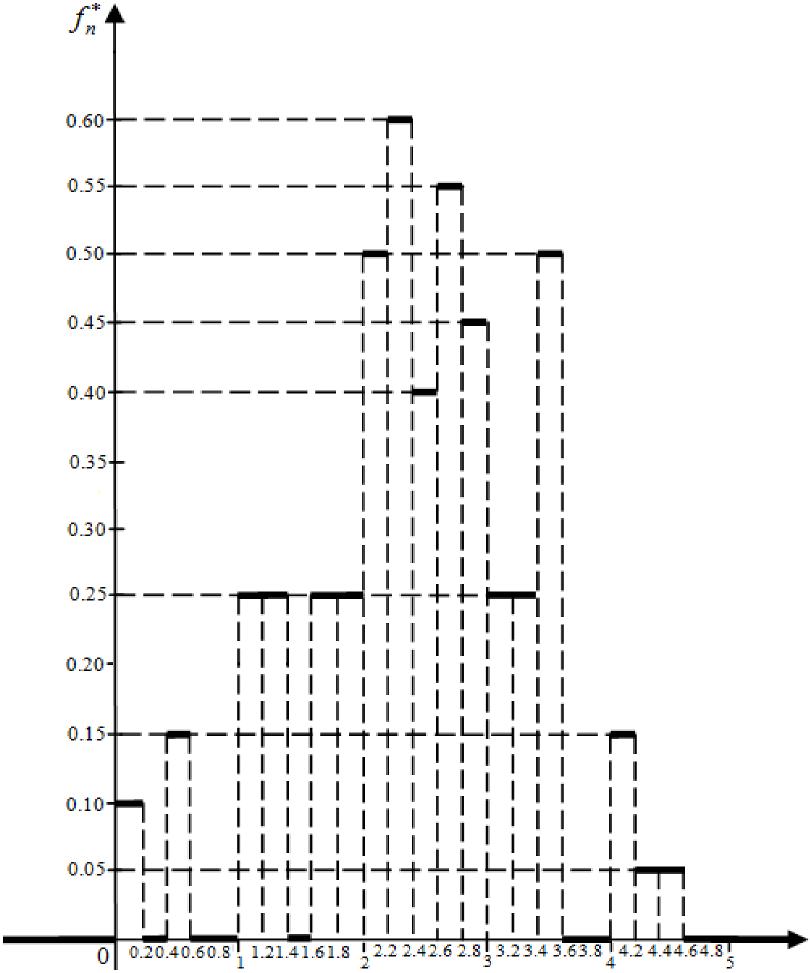


Рис. 2.4.5

Подібних прикладів можна навести як завгодно багато.

Зауважимо, що якщо не збережено варіаційний ряд, в якому зафіксовано місцеположення на числовій осі всіх спостережених точок, і не проводити додаткових спостережень для уточнення розподілу статистичних ймовірностей, то за заданим поінтервальним розподілом статистичних ймовірностей побудувати новий поінтервальний розподіл з дрібнішими інтервалами неможливо, якщо не робити припущення про те, що усереднена щільність $f_n^*(x)$ розподілу статистичних ймовірностей на заданих інтервалах стала і не змінюється з подрібненням заданих інтервалів, чи якихось інших припущень. Якщо ж таке припущення зробити, тоді новий поінтервальний розподіл з дрібнішими інтервалами, як і наявний, будуть гіпотетичними, які можуть виявитись такими, що не цілком узгоджуються із спостереженими даними, зафіксованими у варіаційному ряді.

Разом з тим за заданим поінтервальним розподілом статистичних ймовірностей можна побудувати новий поінтервальний розподіл з укрупненими інтервалами, одержаними як об'єднання кількох заданих інтервалів, коли статистичні ймовірності попадання в укрупнені інтервали одержуються як суми статистичних ймовірностей попадання в інтервали, через об'єднання яких утворено укрупнений інтервал.

Вправи для самостійного виконання

2.4.1. Перевірити, чи правильні твердження:

1. Із зменшенням довжини інтервалів $[a_{i-1}; a_i)$, $i \in \overline{1,k}$, в r разів щільність розподілу узагальнених статистичних імовірностей на кожному із дрібніших інтервалів $[\tilde{a}_{j-1}; \tilde{a}_j) \subset [a_{i-1}; a_i)$ збільшується.

2. Ймовірнісна міра $\tilde{P}_n^*([\alpha; \beta))$ довільних проміжків $[\alpha; \beta) \subset (-\infty; \infty)$ знаходиться в межах $[m_*([\alpha; \beta))\ ;\ m^*([\alpha; \beta)]$.

3. Властивості узагальнених статистичних імовірностей $\tilde{P}_n^*(A)$ такі самі, як і властивості статистичних ймовірностей $P_n^*(A)$, знайдених за результатами серії із n випробувань.

2.4.2. Задано поінтервальний розподіл статистичних імовірностей

$[a_{i-1}; a_i)$	[0; 1)	[1; 2)	[2; 3)	[3; 4)	[4; 5)
$P_n^*([a_{i-1}; a_i))$	0.05	0.20	0.50	0.20	0.05

Визначити $\tilde{P}_n^*([\alpha; \beta))$, коли: $[\alpha; \beta) = [-1; 6)$; $[\alpha; \beta) = [2.1; 2.9)$; $[\alpha; \beta) = [0.5; 1.5)$; $[\alpha; \beta) = [-5; -1)$; $[\alpha; \beta) = [7; 15)$; $[\alpha; \beta) = [-2; 3)$; $[\alpha; \beta) = [2; 9)$.

2.5. Функція поточкового розподілу узагальнених статистичних ймовірностей на скінченній множині точок в одновимірному координатному просторі

Нехай в одновимірному координатному просторі дискретний поточковий розподіл статистичних ймовірностей на множині $\Omega = \{x_1, x_2, ..., x_k\}$ визначається за таблицею 2.5.1:

Табл. 2.5.1.

x_i	x_1	x_2	...	x_k
$P_n^*(\{x_i\})$	$P_n^*(\{x_1\})$	$P_n^*(\{x_2\})$	...	$P_n^*(\{x_k\})$

а як простір подій S розглядається найширша сукупність підмножин множини Ω, тобто $S = \{A \mid A = \bigcup\limits_{i \in I}\{x_i\}, I \subset \{1, 2, ..., k\}\}$.

Тоді за припущення, що $\widetilde{\Omega} = (-\infty; \infty)$, $(-\infty; x) \in \widetilde{S}$ за будь-якого $x \in (-\infty; \infty)$, тобто множини $(-\infty; x)$ за будь-якого $x \in (-\infty; \infty)$ є подіями з простору подій $\widetilde{S}$, на якому визначена ймовірнісна міра $\widetilde{P}_n^*(G)$, $G \in \widetilde{S}$, за кожного $x \in \overline{R} = R \bigcup \{+\infty, -\infty\}$ можна підрахувати число $\widetilde{P}_n^*((-\infty; x))$ – узагальнену статистичну ймовірність попадання спостережених значень $x_i \in \Omega$ в проміжок $(-\infty; x)$, яке дорівнює сумі усіх $P_n^*(\{x_i\})$, за яких $x_i \in (-\infty; x)$ (або, що те саме, $x_i < x$), тобто

$$\widetilde{P}_n^*((-\infty; x)) = \sum_{x_i \in (-\infty; x) \bigcap \Omega} P_n^*(\{x_i\}) = \sum_{x_i < x} P_n^*(\{x_i\}). \qquad (2.5.1)$$

Стосовно цієї статистичної ймовірності $\widetilde{P}_n^*((-\infty; x))$ простором елементарних подій вважають множину $\Omega = R = (-\infty; \infty)$, а простором подій – простір, породжений за сукупністю скінченних об’єднань всеможливих проміжків. Зокрема до цього простору подій належать будь-які проміжки, а також їх скінченні і зчисленні об’єднання і перетини та доповнення цих об’єднань до R.

Зауважимо разом з тим, що $(-\infty; x) \overline{\in} S$, і тому вираз $P_n^*((-\infty; x))$ не є коректним, оскільки ймовірнісна міра P_n^* задана лише на сукупності S підмножин множини Ω. Разом з тим легко бачити, що $(-\infty; x) \bigcap \Omega \in S$.

Функцію

$$F_n^*(x) = \widetilde{P}_n^*((-\infty; x)) = \sum_{x_i \in (-\infty; x) \bigcap \Omega} P_n^*(\{x_i\}), \ i \in \overline{1, k}, \ x \in R, \qquad (2.5.2)$$

називають *функцією розподілу узагальнених статистичних ймовірностей* (*відносних частот*) *на скінченній множині* $\{x_1, x_2, ..., x_k\}$.

Зауважимо, що функцію $F_n^*(x)$ можна визначити і так:

$$F_n^*(x) = P_n^*(\bigcup_{A\in S,\, A\subset(-\infty;\, x)} A). \quad (2.5.3)$$

Приклад 2.5.1. Нехай розподіл статистичних ймовірностей задано за таблицею 2.5.2.

Табл. 2.5.2

x_i	1	2	3	4	5	6
$P_n^*(\{x_i\})$	0	$\frac{1}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{5}{15}$

Тоді

$$F_n^*(x) = \tilde{P}_n^*((-\infty;\, x)) = \begin{cases} 0, \text{ коли } x \le 2 \\ \frac{1}{15}, \text{ коли } 2 < x \le 3, \\ \frac{4}{15}, \text{ коли } 3 < x \le 4, \\ \frac{6}{15}, \text{ коли } 4 < x \le 5, \\ \frac{10}{15}, \text{ коли } 5 < x \le 6, \\ 1, \text{ коли } 6 < x \end{cases}$$

Оскільки $-\infty < x_i$ за всіх i, то $F_n^*(-\infty) = P_n^*(\varnothing) = 0$ як статистична ймовірність неможливої події, яка полягає у появі значення $x_{\text{сп}\, i}$, меншого за $-\infty$.

Так само, оскільки $x_i < +\infty$ за всіх i, то

$$F_n^*(+\infty) = \sum_{x_i\in(-\infty;\, \infty)\bigcap\Omega} \tilde{P}_n^*(\{x_i\}) = \tilde{P}_n^*((-\infty;\, +\infty)) = \tilde{P}_n^*(\Omega) = 1$$

як статистична ймовірність вірогідної події, яка полягає у попаданні точки $x_{\text{сп}\, i}$ на проміжок $(-\infty;\, +\infty)$.

На Рис. 2.5.1 зображено графік функції розподілу узагальнених статистичних ймовірностей, що визначається за таблицею 2.5.2.

Відмітимо *основні властивості функції розподілу узагальнених статистичних ймовірностей*:

1. $F_n^*(x) \ge 0$, як узагальнена статистична ймовірність події $G = (-\infty; x) \subset \Omega$.

2. $F_n^*(-\infty) = 0$, як узагальнена статистична ймовірність неможливої події $\varnothing = \{x \mid x < -\infty\}$, тобто статистична ймовірність появи значень $x_{\text{сп}\, i}$, менших за $-\infty$.

3. $F_n^*(+\infty)=1$, як узагальнена статистична ймовірність вірогідної події $\tilde{\Omega}=(-\infty;\infty)$, тобто статистична ймовірність появи значень $x_{\text{сп}\,i}$, менших за $+\infty$.

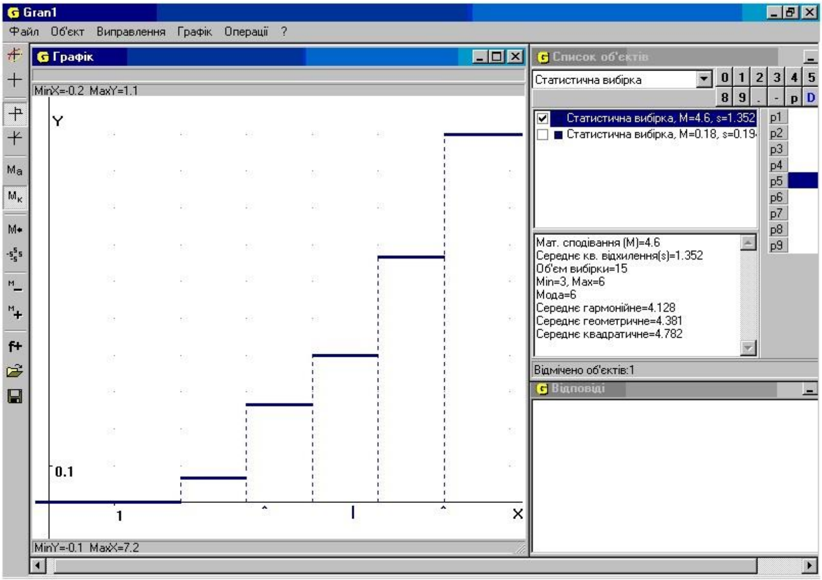


Рис. 2.5.1

4. Якщо $u<v$, то $F_n^*(u)\le F_n^*(v)$, тобто функція розподілу узагальнених статистичних ймовірностей неспадна.

Справді,

$$F_n^*(v)=\sum_{x_i<v}P_n^*(\{x_i\})=\sum_{x_i<u}P_n^*(\{x_i\})+\sum_{u\le x_i<v}P_n^*(\{x_i\})\ge\sum_{x_i<u}P_n^*(\{x_i\})=F_n^*(u),$$

оскільки $P_n^*(\{x_i\})\ge 0$ за всіх i.

З наведеного, зокрема, випливає, що

$$\tilde{P}_n^*([u;v))=\sum_{x_i\in[u,v)}P_n^*(\{x_i\})=F_n^*(v)-F_n^*(u),$$

тобто узагальнена статистична ймовірність попадання точок $x_{\text{сп}\,i}$ на проміжок $[u;v)$ дорівнює приростові функції розподілу узагальнених статистичних ймовірностей на цьому проміжку.

5. На кожному проміжку $(\alpha;x_i]$, в якому немає інших точок множини $\{x_1,x_2,\ldots\}$, окрім точки x_i, функція розподілу узагальнених статистичних ймовірностей стала, причому $F_n^*(x)=F_n^*(x_i)$, коли $x\in(\alpha;x_i]$.

Будь-яку функцію, стосовно якої задовільняються умови

1–5, називають *функцією розподілу узагальнених статистичних ймовірностей на множині* $\tilde{\Omega} = (-\infty; \infty)$, визначеною за поточковим розподілом статистичних імовірностей $P_n^*(\{x_i\})$, $i \in \overline{1,k}$, на скінченній множині $\Omega = \{x_1, x_2, ..., x_k\}$.

Зауважимо, що коли число k дуже велике, наприклад $k = 10^{1000\,000}$, $k = 10^{1000\,000\,000}$ і т.п., а всі $P_n^*(\{x_i\})$, $i \in \overline{1,k}$, рівні або близькі між собою, тоді обчислення $\tilde{P}_n^*((-\infty; x))$ безпосередньо за формулою (2.5.1) стає практично неприйнятним, оскільки всі $P_n^*(\{x_i\})$ практично дорівнюють нулеві і за довільного $x < \infty$ буде $\tilde{P}_n^*((-\infty; x)) = 0$, що призводить до суперечностей. Наприклад, якщо всі $x_i \in [0; 1]$ і віддалені одна від одної на відстань $h = 10^{-1000\,000\,000}$, а всі $P_n^*(\{x_i\})$ рівні між собою, тоді за довільного $x \le \infty$ практично буде $F_n^*(x) = \tilde{P}_n^*((-\infty;\ x)) = \sum\limits_{x_i < x} P_n^*(\{x_i\}) = 0$, що приводить до суперечностей, оскільки за довільного $x > 1$ повинно бути $F_n^*(x) = 1$.

В подібних випадках доцільно множину $\Omega = \{x_1, x_2, ..., x_k\}$ замінити на $[x_1; x_k + \varepsilon)$ і поділити цей проміжок на практично прийнятну кількість m підмножин $H_i = [a_{i-1}; a_i)$, $i \in \overline{1,m}$, де $a_0 = x_1$, $a_m = x_k + \varepsilon$, $a_i = a_{i-1} + h$, $h = \dfrac{a_m - a_0}{m}$, $\varepsilon > 0$ – досить мале число, і визначати не $P_n^*(\{x_i\})$, $i \in \overline{1,k}$, а $\tilde{P}_n^*(H_i)$, $i \in \overline{1,m}$, а $F_n^*(x)$ наближено обчислювати за формулою (див. §2.2, формули 2.2.1, 2.2.2):

$$F_n^*(x) = \sum_{a_i < x} \tilde{P}_n^*(H_i) + \frac{x - a_{i-1}}{h} \tilde{P}_n^*(H_i),\quad x \in [a_{i-1}; a_i),\ i \in \overline{1,m}.$$

Тоді за всіх x таких, що $x \le a_0$, буде $F_n^*(x) = 0$, а за всіх x таких, що $x > a_m$, буде $F_n^*(x) = 1$.

Зауважимо, що подібним чином діють і у випадку, коли множина Ω неперервна.

Приклад 2.5.2. Нехай (див. Приклад 1.8.1) задано простір елементарних подій $\Omega = \{1, 2, 3, 4, 5, 6\}$ і поділ множини Ω на підмножини $H_1 = \{1\}$, $H_2 = \{2, 3, 4\}$, $H_3 = \{5\}, H_4 = \{6\}, H_i H_j = \varnothing$, коли $i \ne j$, $\bigcup\limits_{i=1}^{4} H_i = \Omega$, $S = \{A \mid A = \bigcup\limits_{i \in I} H_i,\ I \subset \{1, 2, 3, 4\}\}$, і на S

задано ймовірнісну міру $P_n^*(H_1) = 0.04$, $P_n^*(H_2) = 0.06$, $P_n^*(H_3) = 0.30$, $P_n^*(H_4) = 0.60$.

Нехай $G = (-\infty; x)$, $\tilde{\Omega} = R^1$, $\tilde{S} = \mathscr{B}(R^1)$.

Тоді за заданого розподілу ймовірностей на просторі подій S визначити ймовірності $P_n^*(G \cap \Omega)$ за значень $x \in (2; 3]$ або $x \in (3; 4]$ неможливо, оскільки $(-\infty; x) \cap \Omega \overline{\in} S$ за вказаних значень x.

Разом з тим $(-\infty; x) \cap \Omega = \varnothing \in S$, коли $x \le 1$,

$$(-\infty; x) \cap \Omega = \{1\} \in S, \text{ коли } 1 < x \le 2,$$
$$(-\infty; x) \cap \Omega = \{1, 2, 3, 4\} \in S, \text{ коли } 4 < x \le 5,$$
$$(-\infty; x) \cap \Omega = \{1, 2, 3, 4, 5\} \in S, \text{ коли } 5 < x \le 6,$$
$$(-\infty; x) \cap \Omega = \{1, 2, 3, 4, 5, 6\} \in S, \text{ коли } 6 < x.$$

Тому $F_n^*(x) = 0,$ коли $x \le 1$,
$F_n^*(x) = 0.04,$ коли $1 < x \le 2$, $F_n^*(x) = 0.10$ коли $4 < x \le 5$,
$F_n^*(x) = 0.40$ коли $5 < x \le 6$, $F_n^*(x) = 1.00$, коли $6 < x$,
а за значень $x \in (2; 4]$ визначити $F_n^*(x)$ за заданого розподілу ймовірностей неможливо, *за таких значень* x *значення функції* $F_n^*(x)$ *виявляються невизначеними.*

В такому разі доцільно ввести новий поділ множини Ω на підмножини H_{ij} такі, що $\bigcup_{j=1}^{m_i} H_{ij} = H_i$ (див. §2.2, §2.3) і на множинах H_{ij} задати нову ймовірнісну міру $P_n^{*(1)}$ так, щоб мали місце рівності $P_n^{*(1)}(\bigcup_{j=1}^{m_i} H_{ij}) = P_n^*(H_i)$. Як простір подій $S^{(1)}$ роглядатимемо сукупність підмножин множини Ω $S = \{A \mid A = \bigcup_{i \in I} \bigcup_{j \in J} H_{ij},\ I \subset \overline{1, k},\ J \subset \overline{1, m_i}\}$ (див. §2.2, §2.3).

В розглядуваному випадку отримаємо $H_{11} = \{1\}$, $H_{21} = \{2\}$, $H_{22} = \{3\}$, $H_{23} = \{4\}$, $H_{31} = \{5\}$, $H_{41} = \{6\}$.

Розподілити статистичну ймовірність $P_n^*(H_2)$ за множинами H_{21}, H_{22}, H_{23} можна різним чином, наприклад покласти $P_n^{*(1)}(H_{21}) = 0.02$, $P_n^{*(1)}(H_{22}) = 0.02$, $P_n^{*(1)}(H_{23}) = 0.02$, або $P_n^{*(1)}(H_{21}) = 0$, $P_n^{*(1)}(H_{22}) = 0$, $P_n^{*(1)}(H_{23}) = 0.06$ і т.д., таких варіантів існує безліч.

Однак за будь якого варіанту довизначення розподілу ймовірностей на множині Ω для того, щоб функція розподілу ймовірностей $F_n^*(x)$ була визначена за всіх $x \in (-\infty, \infty)$, необхідно до простору подій S включити всі одноточкові підмножини $\{x_i\}$, $i \in \overline{1,k}$, множини $\Omega = \{x_1, x_2, \dots, x_k\}$, тобто покласти $S = \{A \mid A = \bigcup\limits_{i \in I} \{x_i\}, \quad I \subset \{1, 2, \dots, k\}\}$, і задати $P_n^*\{x_i\}$, $i \in \overline{1,k}$, тим або іншим чином так, щоб виконувались вимоги $P_n^*\{x_i\} \geq 0$, $\sum\limits_{x_i \in \Omega} P_n^*(\{x_i\}) = 1$.

Нехай в розглядуваному в прикладі випадку $P_n^{*(1)}(H_{21}) = 0$, $P_n^{*(1)}(H_{22}) = 0$, $P_n^{*(1)}(H_{23}) = 0.06$.

Тоді, визначаючи функцію $F_n^*(x)$ розподілу узагальнених статистичних ймовірностей, в розглядуваному випадку дістанемо

$$F_n^*(x) = \begin{cases} 0, \text{ коли } x \leq 1, \\ 0.04, \text{ коли } 1 < x \leq 2, \\ 0.04, \text{ коли } 2 < x \leq 3, \\ 0.04, \text{ коли } 3 < x \leq 4, \\ 0.10, \text{ коли } 4 < x \leq 5, \\ 0.40, \text{ коли } 5 < x \leq 6, \\ 1, \text{ коли } 6 < x. \end{cases}$$

Графік функції $F_n^*(x)$ за так довизначеного розподілу узагальнених ймовірностей подано на Рис. 2.5.2.

Після такого (гіпотетичного) довизначення розподілу статистичних ймовірностей на множині $\Omega = \{1, 2, 3, 4, 5, 6\}$ функція $F_n^*(x)$ розподілу статистичних ймовірностей виявляється визначеною за всіх $x \in (-\infty; \infty)$ на відміну від випадку, коли розподіл статистичних ймовірностей на множині $\Omega = \{1, 2, 3, 4, 5, 6\}$ був заданий за підмножинами $H_1 = \{1\}$, $H_2 = \{2, 3, 4\}$, $H_3 = \{5\}$, $H_4 = \{6\}$, і в такому разі значення функції $F_n^*(x)$ за значень $x \in (2; 4]$ визначити було неможливо, якщо не робити якихось припущень про розподіл статистичних ймовірностей на множині $H_2 = \{2, 3, 4\}$.

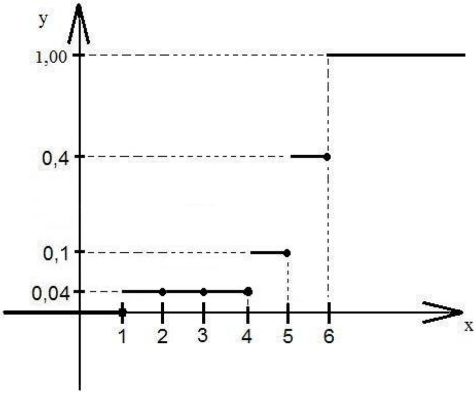


Рис. 2.5.2

Зауважимо, що функція $F_n^*(x) = \tilde{P}_n^*((-\infty; x))$, $x \in (-\infty; \infty)$, визначена за формулою (2.5.2) в разі поточкового розподілу узагальнених статистичних ймовірностей на скінченній множині $\Omega = \{x_1, x_2, \dots, x_k\}$ визначена за всіх $x \in (-\infty, \infty)$ тоді і тільки тоді, коли статистичні ймовірності $P_n^*(\{x_i\})$ визначені за кожного $x_i \in \Omega$, $i \in \overline{1, k}$.

Зауважимо також, що коли функцію $F_n^*(x)$ визначити за формулою

$$F_n^*(x) = P_n^*(\bigcup_{A \in S,\ A \subset (-\infty;\ x)} A), \qquad (2.5.4)$$

тоді $F_n^*(x)$ буде визначена за всіх $x \in R^1$ і разом з тим кусково стала, набуваючи приростів $P_n^*(H_i)$ в разі, коли із зростанням змінної x вперше виявиться $H_i \subset (-\infty; x)$, $i \in 1, 2, \dots, k$, де $\Omega = \bigcup_{i=1}^{k} H_i$, $S = \{A \mid A = \bigcup_{i \in I} H_i, \quad I \subset \{1, 2, \dots, k\}\}$.

Приклад 2.5.3. Нехай $\Omega = \{1, 2, 3, 4, 5, 6\}$, $H_1 = \{1, 3, 5\}$, $H_2 = \{2, 4, 6\}$, $H_1 \bigcap H_2 = \varnothing$, $H_1 + H_2 = \Omega$, $S = \{\varnothing, H_1, H_2, H_1 + H_2 = \varnothing\}$, $P_n^*(H_1) = 0.5$, $P_n^*(H_2) = 0.5$.

Визначаючи функцію $F_n^*(x)$ за формулою (2.5.4), дістанемо:

$$F_n^*(x) = \begin{cases} 0, \text{ коли } x \le 5, \\ 0.5, \text{ коли } 5 < x \le 6, \\ 1, \text{ коли } 6 < x. \end{cases}$$

Графік так визначеної функції $F_n^*(x)$ подано на Рис. 2.5.3

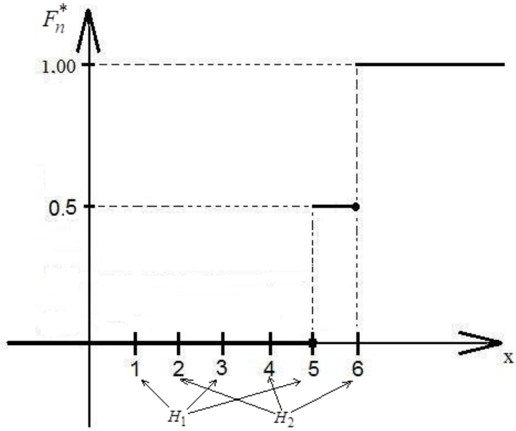


Рис. 2.5.3

Вправи для самостійного виконання

2.5.1. Перевірити, чи правильні твердження:

1. Функцію розподілу узагальнених статистичних ймовірностей неможливо визначити в безкоординатному просторі.

2. Якщо задано ряд поточкового розподілу статистичних ймовірностей, то можна побудувати відповідну функцію розподілу узагальнених статистичних ймовірностей.

3. Якщо задано функцію розподілу узагальнених статистичних ймовірностей на дискретній множині $\{x_1, x_2, ..., x_k\}$, то можна підрахувати статистичні ймовірності появи кожного значення $x_i \in \{x_1, x_2, ..., x_k\}$.

4. Якщо подія G полягає у попаданні точок в проміжок $(-\infty;\alpha]$, тобто $G = (-\infty;\alpha]$, то узагальнена статистична ймовірність цієї події дорівнює значенню відповідної функції розподілу ймовірностей в точці α.

5. Функція розподілу узагальнених статистичних ймовірностей є зростаючою.

6. Функція розподілу узагальнених статистичних ймовірностей є невід'ємною неспадною функцією.

7. Кожна невід'ємна неспадна кусково стала функція є функцією розподілу узагальнених статистичних ймовірностей стосовно деякого дискретного розподілу статистичних ймовірностей.

8. Функція розподілу узагальнених статистичних ймовірностей може набувати лише двох значень.

9. Множина значень функції розподілу узагальнених статистичних

ймовірностей скінченна.

2.5.2. Побудувати графік функції розподілу узагальнених статистичних ймовірностей за дискретним поточковим розподілом:

а)

x_i	-1	0	1
$P_n^*(\{x_i\})$	0.25	0.50	0.25

б) x_i – відхилення точки падіння снаряда від цілі, що можуть набувати значень – -50, -40, -30, -20, -10, 0, 10, 20, 30, 40, 50, а $P_n^*(\{x_i\})$ – визначаються за серією спостережених значень:

10, 0, -20, -10, 20, 0, -10, 20, -30, 10, -20, 0, 0, 10, -10, 30, 0, 10, 0, -20, 10, 10, -20, -10, 0, -10, -10, 0, -10, 20, 0, 0, 10, 10, 20, -30, 20, -20, -10, 10, 0, 10, -10, 0, 10, -10, 0, 0, 0, 20.

2.6. Функція поінтервального розподілу узагальнених статистичних ймовірностей на неперервній множині точок в одновимірному координатному просторі

Перш ніж розглянути означення функції поінтервального розподілу узагальнених статистичних ймовірностей на неперервній множині точок, нагадаємо деякі поняття з теорії множин.

Нехай Ω – деяка фіксована непорожня множина.

Означення. Система V підмножин множини Ω називається *алгеброю*, якщо:

1) $\Omega \in V$;

2) з того, що $A \in V$, випливає, що $\overline{A} \in V$;

3) з того, що $A \in V$ і $B \in V$, випливає, що $A \bigcup B \in V$.

Легко бачити, що всі розглянуті раніше операції над скінченною кількістю підмножин не виводять з алгебри V.

Приклади алгебр підмножин множини Ω:

1. Сукупність $S_* = \{\varnothing, \Omega\}$ – *тривіальна або "найбідніша" алгебра підмножин множини* Ω.

2. $\{\varnothing, A, \overline{A}, \Omega\}$ – алгебра множин, що породжується за множиною A.

3. Система S^* всіх підмножин множини Ω – "найбагатша" алгебра підмножин множини Ω.

4. Нехай $H_1, H_2, ..., H_k$ – деякі підмножини множини Ω такі, що $H_i\ H_j = \varnothing$, коли $i \neq j$, причому $H_i \neq \varnothing$, $\bigcup\limits_{i=1}^{k} H_i = \Omega$.

Якщо утворити систему множин, до якої входять порожня множина $\varnothing$, усі множини H_i, усі можливі суми множин H_i із двох, із трьох, із чотирьох і т.д. доданків, то така система множин буде алгеброю множин. Така *алгебра* множин називається *породженою за поділом* $D = \{H_1, H_2, ..., H_n\}$ *множини* Ω *на підмножини* H_i, які попарно не перетинаються, тобто $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup\limits_{i=1}^{k} H_i = \Omega$.

Множини H_i в такому разі називають *атомами поділу* D. Таким чином, за поділом D однозначно визначається алгебра $V = \alpha(D)$. Навпаки, за алгеброю множин, породженою за поділом D, можна відновити цей поділ, причому єдиний. Справді, якщо $H \in V$ таке, що за будь-якого $B \in V$ або $H \bigcap B = H$ або $H \bigcap B = \varnothing$, тоді сукупність таких H і буде шуканим поділом.

Означення. Система S підмножин із Ω називається *σ-алгеброю*, якщо:

1_s. $\Omega \in S$;

2_s. з того, що $A \in S$, випливає, що $\overline{A} \in S$;

3_s. з того, що $A_i \in S,\ i = 1, 2, \ldots,$ випливає, що $\bigcup_{i=1}^{\infty} A_i \in S$.

Легко бачити, що всі розглянуті операції над скінченною або зчисленною кількістю підмножин не виводять з σ-алгебри S. Зокрема, $\bigcap_{i=1}^{\infty} A_i \in S$, коли $A_i \in S,\ i \in N$.

Прикладом σ-алгебри є сукупність S^* всіх підмножин множини Ω.

Системи множин $S_* = \{\varnothing, \Omega\}$, $S^* = \{A \mid A \subset \Omega\}$ є й алгебрами, й σ-алгебрами, причому S_* – тривіальна, "найбідніша" σ-алгебра, а S^* – "найбагатша" σ-алгебра, до якої включаються всі підмножини множини Ω. Якщо $D = (H_1,\ H_2, \ldots)$ – деякий зчисленний поділ множини Ω на непорожні підмножини

$$\Omega = H_1 + H_2 + \ldots,\ H_i\ H_j = \varnothing, \text{ коли } i \neq j,$$

то система $S = \alpha(D)$, утворена з множин, що є сумами скінченної або зчисленної кількості елементів з D, і приєднаною до них порожньою множиною $\varnothing$, є і алгеброю, і σ –алгеброю.

Зрозуміло, що кожна σ-алгебра є й алгеброю, проте не навпаки.

Означення. Нехай W – деяка система підмножин із Ω.

σ-алгебра $\sigma(W)$ називається *найменшою, до якої входить система* W, якщо:

1) $W \subset \sigma(W)$;

2) За будь-якої σ-алгебри S, до якої також входить система $W, W \subset S$, має місце включення $\sigma(W) \subset S$.

За будь-якої системи множин W існує найменша σ-алгебра $\sigma(W)$, до якої входить система W. Справді, існує принаймні одна σ-алгебра (а саме S^*), до якої входить система W. Перетин всіх σ-алгебр, до якої входить система W, і є шуканою σ-алгеброю. Систему $\sigma(W)$ називають σ-алгеброю, породженою за системою множин W.

Нехай $\Omega = R = (-\infty; \infty)$ – дійсна пряма,

$$[a\,; b) = \{\,x \mid x \in R^1,\ a \le x < b\}$$

за всіх $-\infty \le a \le b \le \infty$. Зокрема $[a\,;a) = \varnothing$. Позначимо через $[-\infty;b)$ інтервал $[-\infty;b)$ (щоб доповнення до R^1 проміжка $[-\infty;b)$ було того самого типу, тобто відкритим праворуч і замкненим ліворуч).

Позначимо через L систему множин з R^1, що складаються із скінченних сум проміжків виду $[a\,;b)$:

$$A \in L \text{, якщо } A = \sum_{i=1}^{n} [a_i\,;b_i\,)\,,\ n < \infty\,.$$

Система L є алгеброю, але не є σ-алгеброю, оскільки якщо $A_n = \left[\frac{1}{n}\,;1\right) \in L$, то $\bigcup_{n=1}^{\infty} A_n = (0\,;1) \,\overline{\in}\, L$.

Означення. Нехай $\mathscr{B}(R^1)$ – найменша σ-алгебра $\sigma(L)$, до якої входить система L. Елементи σ-алгебри $\sigma(L)$ називають *борелівськими множинами*, а сама σ-алгебра $\sigma(L)$ називається *σ-алгеброю борелівських множин в* R^1.

Оскільки $\{a\} = \bigcap_{n=1}^{\infty} \left[a\,;a + \frac{1}{n}\right) \in \sigma(L)$, то одноточкові множини є борелівськими. Проміжки $(a\,;b) = [a\,;b) - \{a\}$, $(a\,;b] = (a\,;b) + \{b\}$, $[a\,;b] = [a\,;b) + \{b\}$ є борелівськими множинами. Кожна відкрита множина є борелівською, оскільки будь-яка відкрита множина в R^1 є сумою скінченної або зчисленної кількості інтервалів. Кожна замкнена множина є також борелівською (доповнення замкненої множини до R^1 – відкрита множина).

Якщо $\Omega = R^n$ і L – система паралелепіпедів виду

$$\prod_{i=1}^{n} [a_i\,;b_i) = \{x \mid x = (x_1, x_2, ..., x_n),\, a_i \le x_i < b_i\,,\, i = \overline{1, n}\}\,,$$

то найменша σ-алгебра $\mathscr{B}(R^n)$, до якої входить система L, називається *σ-алгеброю борелівських множин в* R^n.

Отже, якщо множину A можна дістати, виходячи із замкнених і відкритих множин, за допомогою скінченної або зчисленної кількості операцій об'єднання та перерізу, то множина A буде борелівською. *Обмежена борелівська множина називається* $\mathscr{B}$ *–вимірною* (*вимірною за Борелем*).

Нехай $\Omega = [a\,;b) = \bigcup_{i=1}^{k} [a_{i-1}\,;a_i\,)$, $a \in (-\infty;\infty)$, $b \in (-\infty;\infty)$, $0 < b - a < \infty$, $H_i = [a_{i-1}\,;a_i)$, $i \in \overline{1, k}$, $a_i - a_{i-1} = h > 0$ за всіх $i \in \overline{1, k}$, $S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \{1,\ 2,\ ...,\ k\}\}$, $\widetilde{\Omega} = R^1 = (-\infty;\infty) \supset \Omega$,

$\widetilde{S} = \mathscr{B}(R^1)$ – σ-алгебра борелівських множин із R^1, $\widetilde{S} \supset S$, $m(H_i) = a_i - a_{i-1} = h > 0$, і нехай на Ω задано поінтервальний розподіл статистичних ймовірностей за таблицею 2.6.1

Табл. 2.6.1

$[a_{i-1};a_i)$	$[a_0;a_1)$	$[a_1;a_2)$	...	$[a_{k-1};a_k)$
$P_n^*([a_{i-1};a_i))$	$P_n^*([a_0;a_1))$	$P_n^*([a_1;a_2))$	...	$P_n^*([a_{k-1};a_k))$

звідки усереднена щільність $f(x)$ поінтервального розподілу статистичних ймовірностей на множині $\Omega = [a; b)$ за інтервалами $[a_{i-1}; a_i)$, $i \in \overline{1,k}$, набуває вигляду

$$f(x) = \begin{cases} \dfrac{P_n^*(H_i)}{m(H_i)}, \text{коли } x \in H_i = [a_{i-1}; a_i), \, i \in \overline{1,k}, \\ \quad 0, \qquad \text{коли } x \,\overline{\in}\, \Omega. \end{cases}$$

Оскільки $0 \le P_n^*(H_i) \le 1$, $m(H_i) = a_i - a_{i-1} = h > 0$, $i \in \overline{1,k}$, то $0 \le f(x) \le c < \infty$ за всіх $x \in \widetilde{\Omega} \supset \Omega$, тобто значення функції $f(x)$ обмежені деяким сталим числом $c < \infty$.

Нехай $G = (-\infty; x) \subset \widetilde{\Omega}$, $G \in \widetilde{S} = \mathscr{B}(R^1)$.

Оскільки $(-\infty; x) \bigcap \Omega \,\overline{\in}\, S = \{A \mid A = \bigcup\limits_{i \in I} H_i, \; I \subset \{1, 2, \ldots, k\}\}$, то значення $P_n^*((-\infty; x) \bigcap \Omega)$ не визначені, коли $x \ne a_i$, $i = \overline{0,k}$, тому визначити $F(x)$, поклавши $F(x) = P_n^*((-\infty; x))$, неможливо.

Покладемо

$$F(x) = P_n^*(G_*) + \alpha(P_n^*(G^*) - P_n^*(G_*)) = \\ = P(G) = P((-\infty; x)), \alpha \in [0; 1], \qquad (2.6.1)$$

де $G_* = \bigcup\limits_{\cup H_i \subset G \cap \Omega} (\cup H_i)$ – об’єднання всіх об’єднань $\cup H_i$ таких, що $\bigcup H_i \subset G \bigcap \Omega = (-\infty; x) \bigcap [a; b)$, $G^* = \bigcap\limits_{G \cap \Omega \subset \cup H_i} (\cup H_i)$ – перетин всіх об’єднань $\cup H_i$ таких, що

$$G \bigcap \Omega = (-\infty; x) \bigcap [a; b) \subset \bigcup H_i = \bigcup [a_{i-1}; a_i),$$

$P(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha \in [0; 1)$, узагальнена статистична ймовірність на $(R^1, \mathscr{B}(R^1))$ – ймовірнісна міра на $(R^1, \mathscr{B}(R^1))$, одержана як одне з продовжень міри $P_n^*(A)$, $A \in S$, із простору подій S на простір $\widetilde{S} = \mathscr{B}(R^1)$ за формулою (2.6.1).

Очевидно $G_* \subset G \bigcap \Omega \subset G^*$, тому $P(G_*) \le P(G^*)$.

Легко бачити, що стосовно функції $F(x) = P((-\infty; x))$ задовільняються такі властивості:

1. $F(x) \ge 0$, оскільки $P(G_*) \ge 0$, $P(G^*) - P(G_*) \ge 0$, $0 \le \alpha \le 1$.

2. $F(x)$ неспадна функція, тобто якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$.

Справді, якщо $x_1 < x_2$, то $G_1 = (-\infty; x_1) \subset G_2 = (-\infty; x_2)$, тому $G_{1*} \subset G_{2*}$, $G_1^* \subset G_2^*$, $P(G_{1*}) \le P(G_{2*})$, $P(G_1^*) \le P(G_2^*)$, звідки $F(x_1) = (1-\alpha)P(G_{1*}) + \alpha P(G_1^*) \le (1-\alpha)P(G_{2*}) + \alpha P(G_2^*) = F(x_2)$.

3. $F(x) \to 0$, коли $x \to -\infty$.

Справді, коли $x \to -\infty$, то $(-\infty; x) \to (-\infty; -\infty) = \varnothing$, тому $G_* = \varnothing$, $G^* = \varnothing$, $P(G_*) = 0$, $P(G^*) = 0$, $(1-\alpha)P(G_*) + \alpha P(G^*) = 0$.

4. $F(x) \to 1$, коли $x \to \infty$.

Справді, коли $x \to \infty$, то $(-\infty; x) \to (-\infty; +\infty) = \tilde{\Omega} \supset \Omega$, тому за $G = \tilde{\Omega}$ буде $G \bigcap \Omega = \Omega$, $G_* = \bigcup\limits_{\bigcup H_i \subset G \bigcap \Omega} (\bigcup H_i) = \Omega$, $G^* = \bigcap\limits_{G \bigcap \Omega \subset \bigcup H_i} (\bigcup H_i) = \Omega$, звідки

$$F(x) = (1-\alpha)P_n^*(\Omega) + \alpha P_n^*(\Omega) = (1-\alpha) \cdot 1 + \alpha \cdot 1 = 1.$$

Оскільки $P_n^*(H_i) = f(x) \cdot m(H_i) = \int\limits_{H_i} f(x)dx$, $x \in H_i$, $i \in \overline{1, k}$, то

$$P(G) = (1-\alpha)P(G_*) + \alpha P(G^*) = (1-\alpha) \int\limits_{G_*} f(x)dx + \alpha \int\limits_{G^*} f(x)dx =$$

$$= \int\limits_{G_*} f(x)dx + \alpha\left(\int\limits_{G^*} f(x)dx - \int\limits_{G_*} f(x)dx\right), \ \alpha \in [0, 1] \qquad (2.6.2)$$

Оскільки $m(G^*) - m(G_*) \le h$, то

$$\int\limits_{G^*} f(x)dx - \int\limits_{G_*} f(x)dx \le \max_{x \in [a,b)} f(x) \cdot h = c \cdot h \to 0, \text{ коли } h \to 0.$$

5. Якщо $h > 0$, $x_1 \ne x_2$, $x_1 \in H_{i_0} = [a_{i_0 - 1}; a_{i_0})$, $x_2 \in H_{i_0} = [a_{i_0 - 1}; a_{i_0})$, тобто дві різні точки x_1 і x_2 лежать в одному і тому самому інтервалі, то $F(x_1) = F(x_2)$, тобто функція $F(x_1)$ набуває сталих значень c_i на кожному з інтервалів $[a_{i-1}; a_i)$.

Оскільки $F(x) = 0$, коли $x \le a$, $F(x) = 1$, коли $b < x$, то, коли $h > 0$, існують принаймні два сусідні інтервали $[a_{i_1 - 1}; a_{i_1})$ і

$[a_{i_2-1}; a_{i_2})$ такі, що $|F(x_2)-F(x_1)|>0$, коли $x_1 \in [a_{i_1-1}; a_{i_1})$, $x_2 \in [a_{i_2-1}; a_{i_2})$, навіть коли $x_1 \to x_2$.

Це означає, що коли $h>0$, то функція $F(x)$ поінтервального розподілу узагальнених статистичних ймовірностей не може бути неперервною.

Разом з тим, коли $x_1 \in [a_{i_0-1}; a_{i_0})$, $x_2 \in [a_{i_0}; a_{i_0+1})$, тобто точки x_1 і x_2 лежать в сусідніх інтервалах, тоді

$$F(x_2)-F(x_1) \le \max_{x\in[a,b)} f(x)\cdot h = c\cdot h,$$

бо при переході точки x із інтервалу $[a_{i_0-1}; a_{i_0})$ в сусідній інтервал $[a_{i_0}; a_{i_0+1})$ значення функції $F(x)$ збільшується на величину $(1-\alpha)P_n^*([a_{i_0-1}; a_{i_0})) + \alpha P_n^*([a_{i_0}; a_{i_0+1})) \le c\cdot h$, оскільки до $G_* = \bigcup\limits_{\bigcup H_1 \subset G\bigcap\Omega} (\bigcup H_i) \subset (-\infty; x)\bigcap[a;b)$ додається ще один інтервал, так само, як і до $G^* = \bigcap\limits_{G\bigcap\Omega\subset\bigcup H_i} (\bigcup H_i)$. Тому коли $h \to 0$, тоді $x_2 \to x_1$, $F(x_2) \to F(x_1)$, тобто за необмеженого зменшення $h>0$, $h \to 0+0$, функція розподілу узагальнених статистичних ймовірностей стає практично неперервною (див. Рис. 2.7.5 а) – 2.7.5 д)), тобто за як завгодно малих $\varepsilon>0$ і $\delta>0$ матиме місце нерівність $|F(x_2)-F(x_1)|<\delta$, коли $|x_2-x_1| \le 2h < \varepsilon$.

Окрім того, коли $h \to 0$, то і $G_* \to G$, $G^* \to G$, $\lim\limits_{h\to 0} G_* = \lim\limits_{h\to 0} G^* = G$, $m(G^* \setminus G_*) = m(G^*) - m(G_*) \to 0$, тому, враховуючи формулу (2.6.2), стосовно $G=(-\infty, x)$ отримаємо

$$F(x) = P((-\infty; x)) = P(G) = \int\limits_{G_*} f(x)dx + \alpha\Big(\int\limits_{G^*} f(x)dx - \int\limits_{G_*} f(x)dx\Big) \le$$

$$\le \int\limits_{G_*} f(x)dx + ch \underset{h\to 0}{\to} \int\limits_{G} f(x)dx = \int\limits_{-\infty}^{x} f(x)dx.$$

Функція $F(x)$, *визначена за формулою* (2.6.1), *називається функцією поінтервального розподілу узагальнених статистичних ймовірностей на множині*

$$\tilde{\Omega} = (-\infty; \infty) \supset \Omega = [a; b) = \bigcup_{i=1}^{k} [a_{i-1}; a_i), \ a_i - a_{i-1} = h > 0.$$

Зауважимо, що коли функцію $F(x)$ визначити за формулою

$$F(x) = P_n^*\Big(\bigcup_{A\in S;\, A\subset(-\infty;x)} A\Big), \qquad (2.6.3)$$

то вона співпадатиме з функцією $F(x)$, визначеною за формулою (2.6.1) за умови $\alpha = 0$.

Якщо щільність $f(x)$ розподілу узагальнених статистичних ймовірностей обмежена, тобто $f(x) \le c < \infty$, тоді

$$F(x) = P((-\infty;\, x)) = \int_{G_*} f(x)dx + \alpha\Big(\int_{G^*} f(x)dx - \int_{G_*} f(x)dx\Big) \xrightarrow[h\to 0]{} \int_{-\infty}^{x} f(x)dx$$

Така функція $F(x) = \int_{-\infty}^{x} f(x)dx$ *є неперервною, а відповідний граничний поінтервальний розподіл узагальнених статистичних ймовірностей, коли* $h \to 0$, *також називається неперервним.*

Зауважимо, що коли $h > 0$ досить мале додатне число, $a_i = a_{i-1} + h$, тоді

$$F(a_i) - F(a_{i-1}) = P_n^*([a_{i-1};\, a_i)) = f(x)\cdot h,\ \ x \in [a_{i-1}, a_i),$$

звідки

$$f(x) = \frac{F(a_i) - F(a_{i-1})}{a_i - a_{i-1}},\ \ x \in [a_{i-1}, a_i),$$

тобто через значення усередненої щільності $f(x)$ поінтервального розподілу статистичних ймовірностей на множині $\Omega = [a;\, b) = \bigcup_{i=1}^{k} [a_{i-1};\, a_i)$ характеризується середня швидкість зростання функції $F(x)$ під час переходу від точки $x = a_{i-1}$ до точки $x = a_i$. Якщо $h \to 0$, то $a_i \to a_{i-1}$ і

$$\frac{F(a_i) - F(a_{i-1})}{a_i - a_{i-1}} \to F'(a_i) = f(a_i),$$

тобто в разі, коли $F(x)$ – неперервна функція на $\tilde{\Omega} = R^1$, $S = B(R^1)$, то $f(x) = \frac{d}{dx}F(x)$ у всіх точках $x \in [a_{i-1};\, a_i) \subset (-\infty;\infty)$, за яких існує границя $\lim\limits_{a_i \to a_{i-1}} \frac{F(a_i) - F(a_{i-1})}{a_i - a_{i-1}}$, $x \in [a_{i-1}, a_i)$.

Приклад 2.6.1. Якщо поінтервальний розподіл статистичних ймовірностей визначається за таблицею 2.6.2:

Табл. 2.6.2

$[a_{i-1};\, a_i)$	$\left[0;\frac{1}{10}\right)$	$\left[\frac{1}{10};\frac{2}{10}\right)$	$\left[\frac{2}{10};\frac{3}{10}\right)$	$\left[\frac{3}{10};\frac{4}{10}\right)$	$\left[\frac{4}{10};\frac{5}{10}\right)$
$P_{100}^*([a_{i-1};a_i))$	0.5	0.2	0.1	0.1	0.04

$\left[\frac{5}{10};\frac{6}{10}\right)$	$\left[\frac{6}{10};\frac{7}{10}\right)$	$\left[\frac{7}{10};\frac{8}{10}\right)$	$\left[\frac{8}{10};\frac{9}{10}\right)$	$\left[\frac{9}{10};1\right)$
0	0.02	0.01	0.02	0.01

то

$$F(x)=P_n^*(G_*)=\begin{cases} 0, & \text{коли} \quad x \le 0.1, \\ 0.5, & \text{коли} \quad 0.1 < x \le 0.2, \\ 0.7, & \text{коли} \quad 0.2 < x \le 0.3, \\ 0.8, & \text{коли} \quad 0.3 < x \le 0.4, \\ 0.9, & \text{коли} \quad 0.4 < x \le 0.5, \\ 0.94, & \text{коли} \quad 0.5 < x \le 0.6, \\ 0.94, & \text{коли} \quad 0.6 < x \le 0.7, \\ 0.96, & \text{коли} \quad 0.7 < x \le 0.8, \\ 0.97, & \text{коли} \quad 0.8 < x \le 0.9, \\ 0.99, & \text{коли} \quad 0.9 < x \le 1, \\ 1, & \text{коли} \quad x > 1. \end{cases}$$

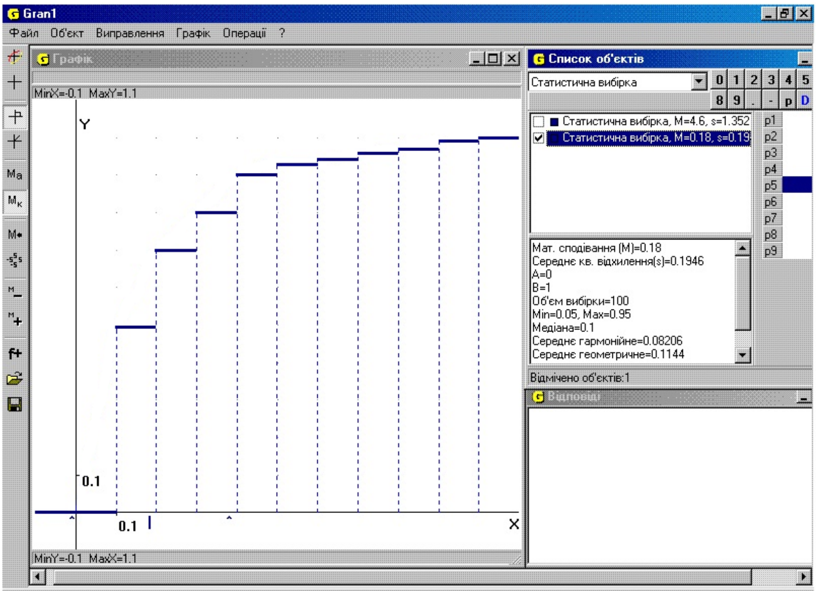


Рис. 2.6.1

На Рис. 2.6.1 подано графік функції поінтервального розподілу узагальнених статистичних (гіпотетичних) ймовірностей, визначених за поінтервальним розподілом статистичних ймовірностей, що заданий за таблицею 2.6.2.

Оскільки

$$F(a_i)=\int_{-\infty}^{a_i} f(x)dx=\int_{-\infty}^{a_{i-1}} f(x)dx+\int_{a_{i-1}}^{a_i} f(x)dx=$$
$$=F(a_{i-1})+P_n^*([a_{i-1},a_i)),$$

то

$$P_n^*([a_{i-1};a_i))=F(a_i)-F(a_{i-1}).$$

Таким чином, за заданою функцією $F(x)$ поінтервального розподілу узагальнених статистичних імовірностей на множині $\Omega=(-\infty;\infty)\supset\Omega=\bigcup_{i=1}^{k}[a_{i-1};a_i)$, $a_i-a_{i-1}=h>0$, можна визначити статистичні ймовірності $P_n^*([a_{i-1};a_i))$ (відносні частоти попадання спостережуваних значень у проміжки $[a_{i-1};a_i)$) за будь-яких проміжків $[a_{i-1};a_i)$, заданих в таблиці 2.6.1. А тому через функцію $F(x)$ розподілу узагальнених статистичних ймовірностей цілком визначається узагальнена статистична (гіпотетична) ймовірність $P(x)$ будь-якої події (вимірної множини) $G\subset\tilde{\Omega}$, як гіпотетична ймовірність попадання в об'єднання проміжків виду $G\bigcap[a_{i-1},a_i)$, $i\in\overline{1,k}$, гіпотетичні ймовірності попадання в які обчислюються за формулою (2.6.1) або за формулою

$$P(G\bigcap[a_{i-1},a_i))=\frac{m(G\bigcap[a_{i-1},a_i))}{m([a_{i-1},a_i))}P_n^*([a_{i-1},a_i))=\int_{G\bigcap[a_{i-1},a_i)} f(t)dt,$$

$i\in\overline{1,k}$, тобто як сума приростів функції $F(x)$ розподілу статистичних ймовірностей на проміжках $G\bigcap[a_{i-1},a_i)$, $i\in\overline{1,k}$.

Якщо $G=\bigcup_{i\in I}[a_{i-1},a_i)$, $I\subset\{1,\ 2,\dots,\ k\}$, тоді $P(G)=P_n^*(G)$.

Приклад 2.6.1. Нехай задано поінтервальний розподіл ймовірностей $P_n^*([a_{i-1};a_i))$, $i\in\overline{1,25}$ на множині $\Omega=[0;5)$ за інтервалами довжиною $h=a_i-a_{i-1}=0.2$, $a_0=0$, $a_i=a_{i-1}+0.2$:

$P_n^*([a_{i-1};a_i))=0.01$, коли $i\in\overline{1,5}$,

$P_n^*([a_{i-1};a_i))=0.04$, коли $i\in\overline{6,10}$,

$P_n^*([a_{i-1};a_i))=0.10$, коли $i\in\overline{11,15}$,

$P_n^*([a_{i-1};a_i))=0.04$, коли $i\in\overline{16,20}$,

$P_n^*([a_{i-1};a_i)) = 0.01$, коли $i \in \overline{21, 25}$,

$S = \{A \mid A = \bigcup_{i \in I}[a_{i-1}; a_i), \ I \in \overline{1, 25}\}$.

Тоді за $x \in \Omega^* = (-\infty; 0] \cup (\bigcup_{i=0}^{25}\{a_i\}) \cup (5; +\infty)$ одержимо :

$F(x) = 0$, коли $x \le 0$; $F(x) = 0.01$, коли $x = 0.2$, $F(x) = 0.02$, коли $x = 0.4$;

$F(x) = 0.03$, коли $x = 0.6$; $F(x) = 0.04$, коли $x = 0.8$; $F(x) = 0.05$, коли $x = 1.0$;

$F(x) = 0.09$, коли $x = 1.2$; $F(x) = 0.13$, коли $x = 1.4$; $F(x) = 0.17$, коли $x = 1.6$;

$F(x) = 0.21$, коли $x = 1.8$; $F(x) = 0.25$, коли $x = 2.0$; $F(x) = 0.35$, коли $x = 2.2$;

$F(x) = 0.45$, коли $x = 2.4$; $F(x) = 0.55$, коли $x = 2.6$; $F(x) = 0.65$, коли $x = 2.8$;

$F(x) = 0.75$, коли $x = 3.0$; $F(x) = 0.79$, коли $x = 3.2$; $F(x) = 0.83$, коли $x = 3.4$;

$F(x) = 0.87$, коли $x = 3.6$; $F(x) = 0.91$, коли $x = 3.8$; $F(x) = 0.95$, коли $x = 4.0$;

$F(x) = 0.96$, коли $x = 4.2$; $F(x) = 0.97$, коли $x = 4.4$; $F(x) = 0.98$, коли $x = 4.6$;

$F(x) = 0.99$, коли $x = 4.8$; $F(x) = 1.00$, коли $x = 5.0$; $F(x) = 1.00$, коли $x > 5$;

Нагадаємо, що значення функції $F(x) = P_n^*((-\infty; x))$ у внутрішніх точках інтервалів $[a_{i-1}; a_i)$ невизначені.

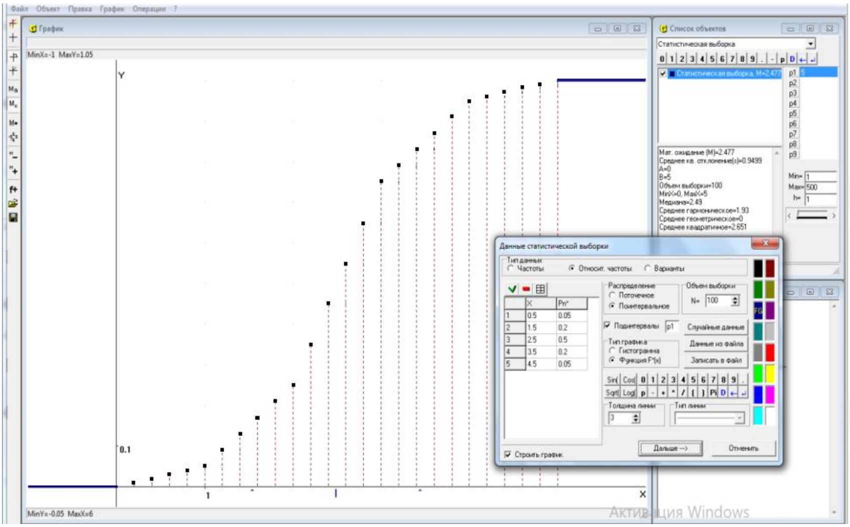


Рис. 2.6.2

Нанісши на координатну площину точки $(x_i, F(x_i))$, $x_i \in \bigcup_{i=0}^{25}\{a_i\}$, а також графік функції $F(x)$ за $x \in (-\infty; 0]$ та $x \in (5; \infty)$, одержимо наближену залежність значень $F(x)$ від довільних $x \in (-\infty;\ \infty)$, враховуючи, що в разі довизначення значень функції $F(x)$ у внутрішніх точках інтервалів $[a_{i-1};\ a_i)$ за $x \in [a_{i-1}; a_i)$ має бути $F(a_{i-1}) \le F(x) \le F(a_i)$, та $F(x_1^{(i)}) \le F(x_2^{(i)})$, коли $x_1^{(i)} < x_2^{(i)}$, $x_1^{(i)} \in [a_{i-1};\ a_i)$, $x_2^{(i)} \in [a_{i-1};\ a_i)$.

На практиці точки $(a_{i-1}, F(a_{i-1}))$ та $(a_i, F(a_i))$ з'єднують відрізками прямих, задаючи в такий спосіб значення функції $F(x)$ за всіх $x \in (-\infty; \infty)$, і крім того вважаючи, що на інтервалах $[a_{i-1};\ a_i)$ ймовірності $P_n^*[a_{i-1}; a_i)$ розподілені неперервно і рівномірно із щільністю

$$f_n^*(x) = \begin{cases} \dfrac{P_n^*([a_{i-1}; a_i))}{a_i - a_{i-1}}, & \text{коли } x \in [a_{i-1}; a_i),\ i \in \overline{1,k} \\ 0, & \text{коли } x \,\overline{\in}\, [a_0; a_k), \end{cases}$$

що насправді може виявитись не так (див. §2.2, §2.3, §2.5, а також рисунки 2.7.4.а), б), в), г), д)).

Щоб функція $F(x)$ була визначена за всіх $x \in (-\infty; \infty)$, покладемо

$$F(x) = P_n^*(G_*) + \alpha(P_n^*(G^*) - P_n^*(G_*)) = P(G) = P((-\infty; x)) \in [0;\ 1],$$

де $G = (-\infty; x) \,\overline{\in}\, S$, $G_* = \bigcup_{\cup H_i \subset G \cap \Omega} (\bigcup H_i) \in S$, $G^* = \bigcap_{G \cap \Omega \subset \cup H_i} (\bigcup H_i) \in S$,

$P(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha \in [0; 1]$, – найближене значення ймовірнісної міри $P_1(G),\ G \,\overline{\in}\, S,\ G \in \tilde{S}$, яка є продовженням ймовірнісної міри $P_n^*(A),\ A \in S,$ із простору подій $S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1, 2, \dots, k\}\}$ на простір подій $\tilde{S} = \mathscr{B}(R^1)$, яке визначається за формулою (2.6.1) (див. § 2.2).

Найчастіше (за досить малих h) покладають $\alpha = 0$. Іноді покладають $\alpha = \dfrac{m((G \bigcap \Omega) \setminus G_*)}{m(G^* \setminus G_*)}$.

Очевидно, $G_* \subset G \bigcap \Omega \subset G^*$, тому $P(G_*) \le P(G^*)$, $P_n^*(G_*) \le F(x) \le P_n^*(G^*)$.

Таким чином, за довільного $x \in [a_i;\ a_{i+1})$, $i \in \overline{0, k-1}$, $F(x) = \sum\limits_{j=1}^{i} P_n^*([a_{j-1};\ a_j)) + \alpha P_n^*([a_i;\ a_{i+1}))$, $\alpha \in [0; 1]$, $F(x) = 0$, коли $x \le a_0$, $F(x) = 1$, коли $x > a_k$.

Оскільки значення функції $F(x)$ цілком визначене лише в точках a_i, $i \in \overline{0, k}$ (див.§2.2,.§2.3,.§2.5), то в разі $x \in [a_{i-1};\ a_i)$ за досить малих $h = a_i - a_{i-1}$, $i \in \overline{1, k}$, значення $F(x)$ можна обирати довільно в межах $F(a_{i-1}) \le F(x) \le F(a_i)$ з дотриманням вимог $F(x_1) \le F(x_2)$, коли $x_1 < x_2$, $x_1 \in [a_{i-1}; a_i)$, $x_2 \in [a_{i-1}; a_i)$, причому $F(a_i) - F(a_{i-1}) = P_n^*([a_{i-1}; a_i))$. Оскільки невідомо, як саме розподілені статистичні ймовірності $P_n^*([a_{i-1}; a_i))$ на інтервалах $[a_{i-1}; a_i)$, $i \in \overline{1, k}$, значення функції $F(x) = P((-\infty;\ x))$ розподілу ймовірностей на $(R^1, \mathscr{B}(R^1)$ виявляються невизначеними у внутрішніх точках інтервалів $[a_{i-1}; a_i)$ (див. §2.2., §2.3). Розподіл ймовірностей $P_n^*([a_{i-1}; a_i))$ на інтервалах $[a_{i-1}; a_i)$ можна задати різним чином, і таких способів існує безліч (див. приклад 2.5.4). Лише після уточнення розподілів ймовірностей $P_n^*([a_{i-1}; a_i))$ на інтервалах $[a_{i-1}; a_i)$, $i \in \overline{1, k}$, значення функції $F(x) = P((-\infty; x))$ будуть визначені у всіх точках $x \in (-\infty\,;\ \infty)$, зокрема і за всіх $x \in [a_{i-1}; a_i)$, $i \in \overline{1, k}$. Одним із варіантів довизначення функції $F(x) = P((-\infty; x))$ за всіх $x \in (-\infty\,;\ \infty)$ є задання функції $F(x) = P((-\infty; x))$ за формулою (2.6.1). Часто покладають $F(x) = F(a_{i-1})$, коли $x \in [a_{i-1}; a_i)$.

Функція $F(x)$, *визначена за формулою* (2.6.1), *називається функцією поінтервального розподілу узагальнених статистичних ймовірностей на неперервній множині*

$$\tilde{\Omega} = (-\infty; \infty) \supset \Omega = [a; b) = \bigcup_{i=1}^{k} [a_{i-1}; a_i)\,,\ a_i - a_{i-1} = h > 0\,.$$

Очевидно $P_n^*(G_*) \le F(x) \le P_n^*(G^*)$. Зауважимо, що, як видно з (2.6.1), коли $x = a_i$, тоді $P_n^*(G^*) - P_n^*(G_*) = 0$ і $F(x) = P_n^*(G_*)$. Очевидно, на кожному з інтервалів $[a_{i-1};\, a_i)$ функція $F(x)$, визначена за формулою 2.6.1, набуває одного і того самого значення в межах між $F(a_{i-1})$ і $F(a_i)$.

Як видно з формули 2.6.1, функція $F(x)$ на інтервалах $[a_{i-1};\, a_i)$ набуває сталих значень, в разі $\alpha = 0$ рівних $\sum_{j=1}^{i} P_n^*([a_{j-1};\, a_j))$, і коли $h = a_i - a_{i-1} > 0$, $i \in \overline{1,k}$, функція $F(x)$ набуває не більше, ніж $k+2$ значень (найменшого – нуль, найбільшого – 1), а тому за $h > 0$ $F(x)$, визначена за формулою 2.6.1, не може бути неперервною.

Легко бачити, що $F(-\infty) = 0$, $F(+\infty) = 1$, функція $F(x)$ неспадна на $R^1 = (-\infty;\, \infty)$, оскільки на інтервалі $[a_i\,;\, a_{i+1})$ значення функції $F(x)$ (сталі на інтервалах $[a_{i-1}; a_i)$, $i \in \overline{1,\, k}$) не можуть бути менші, ніж на інтервалі $[a_{i-1}; a_i)$. Очевидно, коли $x \in [a_{i-1}; a_i)$, тоді $G_* = \bigcup_{k=1}^{i-1} [a_{k-1}\,;\, a_k)$,

$$G^* = \bigcup_{k=1}^{i} [a_{k-1};\, a_k) = \bigcup_{k=1}^{i-1} [a_{k-1};\, a_k) \bigcup [a_{i-1};\, a_i) = G_* \bigcup [a_{i-1};\, a_i),$$

$$P_n^*(G^*) = P_n^*(G_*) + P_n^*([a_{i-1}\,;\, a_i)) \ge P_n^*(G_*).$$

Зауважимо, що

$$P([a_i;\, a_{i+l})) = F(a_{i+l}) - F(a_i) = \sum_{j=0}^{l-1} P([a_{i+j}; a_{i+j+l})),\ [a_i;\, a_{i+l}) \in S,$$

$0 \le i < i + l \le k$.

Коли щільність $f(x)$ розподілу узагальнених статистичних ймовірностей обмежена, тобто $f(x) \le c < \infty$, тоді за умови $h \to 0$ $G^* \to G_* \to G \bigcap \Omega,\ S \to \mathscr{B}(\Omega)$,

$$F(x) = P_n^*(G_*) + \alpha(P_n^*(G^*) - P_n^*(G_*)) \xrightarrow[h \to 0]{} \int_{-\infty}^{x} f(x)dx.$$

Така функція $F(x)$ *є неперервною, а відповідний граничний поінтервальний розподіл статистичних ймовірностей, коли* $h \to 0$, *також називається неперервним.*

Зауважимо, що коли $F(x)$ визначена за формулою 2.6.1, тоді за довільних $k \in N$, $h = a_i - a_{i-1}$, $i \in \overline{1,k}$, $\frac{d}{dx}F(x) = 0$ майже всюди, і тому $\lim_{h\to 0}(\frac{d}{dx}F(x)) = 0 \neq \frac{d}{dx}(\lim_{h\to 0} F(x)) = f(x)$. Разом з тим $\lim_{h\to 0}\frac{1}{h}(F(a_i) - F(a_{i-1})) = \lim_{h\to 0}\frac{1}{h}P_n^*((a_{i-1};\, a_i]) = f(x)$.

Вправи для самостійного виконання

2.6.1. Перевірити, чи правильні твердження:

1. В безкоординатному просторі функція розподілу ймовірностей може бути визначена через щільність розподілу ймовірностей.
2. За кожного поінтервального розподілу статистичних ймовірностей існує відповідна функція розподілу.
3. Функція $F(x)$, що визначається за рівністю $F(x) = \int_{-\infty}^{x} f(x)dx$, є функцією розподілу гіпотетичних ймовірностей.
4. Щоб задати поінтервальний розподіл узагальнених статистичних ймовірностей, досить задати або таблицю виду 2.6.1, або щільність розподілу $f(x)$, або функцію розподілу $F(x)$.
5. Функція $F(x)$ поінтервального розподілу гіпотетичних ймовірностей є зростаючою.
6. Функція $F(x)$ поінтервального розподілу гіпотетичних ймовірностей є зростаючою на кожному проміжку $[a_{i-1};a_i)$.
7. Множина значень функції $F(x)$ поінтервального розподілу гіпотетичних ймовірностей скінченна.
8. Функція $F(x)$ розподілу гіпотетичних ймовірностей в разі, коли розподіл неперервний, є неперервною в кожній точці x_0, тобто $F(x) \approx F(x_0)$, коли $x \approx x_0$.

2.6.2. Побудувати графік функції $F(x)$ поінтервального розподілу узагальнених статистичних (гіпотетичних) ймовірностей, якщо поінтервальний розподіл статистичних ймовірностей задано за таблицею:

а)

$[a_{i-1};a_i)$	[-35; -25)	[-25; -15)	[-15; -5)	[-5; 5)	[5; 15)	[15; 25)	[25; 35)
$P_n^*([a_{i-1};a_i))$	0.04	0.06	0.20	0.40	0.20	0.06	0.04

б)

x_i	-4	-3	-2	-1	0	1	2	3	4
$P_n^*([a_{i-1};a_i))$	0.02	0.03	0.05	0.15	0.50	0.15	0.05	0.03	0.02

де x_i – центри інтервалів $[a_{i-1};a_i)$, $a_i = a_{i-1} + h$, $i = \overline{1,9}$.

2.7. Залежність властивостей функцій розподілу узагальнених статистичних ймовірностей від структури подій

Нехай задано ймовірнісний простір (Ω, S, P_n^*), $\Omega = (-\infty\ ;\infty)$, $\tilde{S} = \mathscr{B}(R^1)$ – сукупність підмножин множини $\tilde{\Omega}$, стосовно якої задовільняються вимоги 1_s-3_s, зокрема до $\tilde{S}$ відносяться підмножини $(-\infty; x) \subset \tilde{\Omega}$ за довільних $x \in (-\infty\ ;\infty)$, $S \subset \tilde{S}$.

Позначимо через $\bigcup\limits_{A \in S,\ A \subset (-\infty;\ x)} A$ суму всіх подій A із S таких, що $A \subset (-\infty; x) \in \tilde{S}$. Очевидно така сума $\bigcup\limits_{A \in S,\ A \subset (-\infty;\ x)} A$ буде належати до сукупності S, тобто буде подією.

Покладемо

$$F_n^*(x) = \tilde{P}_n^*((-\infty; x)) = \tilde{P}_n^*\left(\bigcup_{A \in S,\ A \subset (-\infty;\ x)} A \right),\ A \in S\ ,\ (-\infty; x) \in \tilde{S} \quad (2.7.1)$$

Очевидно, функція $F_n^*(x)$ визначена за всіх $x \in (-\infty\ ;\infty)$.

Легко бачити, що стосовно функції $F_n^*(x)$ задовільняються такі властивості:

1. $F_n^*(x) \geq 0$ як статистична ймовірність деякої події.
2. $F_n^*(-\infty) = 0$ як статистична ймовірність неможливої події, оскільки за довільного Ω
$$\bigcup_{A \subset (-\infty; -\infty)} A = \varnothing \text{, бо } (-\infty; -\infty) = \varnothing .$$
3. за довільного $x \in (-\infty; -\infty)$

$$m_*((-\infty; x)) = \max_{\bigcup A \subset (-\infty; x) \bigcap \Omega} P_n^*(\bigcup A) \leq F_n^*(x) \leq \min_{(-\infty; x) \bigcap \Omega \subset \bigcup A} P_n^*(\bigcup A) =$$
$$= m^*((-\infty; x)),\ A \in S .$$

4. Функція $F_n^*(x)$ неспадна, тобто якщо $u < v$, то $F_n^*(u) \leq F_n^*(v)$.

Справді, нехай $u < v$. Тоді

$$\bigcup_{A \in S,\ A \subset (-\infty; u)} A \subset \bigcup_{A \in S,\ A \subset (-\infty; v)} A ,$$

а тому за властивостями ймовірнісної міри

$$P_n^*\left(\bigcup_{A \in S, A \subset (-\infty; u)} A \right) \leq P_n^*\left(\bigcup_{A \in S, A \subset (-\infty; v)} A \right),$$

тобто $F_n^*(u) \leq F_n^*(v)$.

Очевидно

$$P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;v)} A \ \setminus \bigcup_{A\in S, A\subset(-\infty;u)} A\right)=$$

$$=P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;v)} A\right)-P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;u)} A\right)=F_n^*(v)-F_n^*(u) \quad (2.7.2)$$

Разом з тим не виключено, що $\bigcup_{A\in S, A\subset(-\infty;v)} A\setminus \bigcup_{A\in S, A\subset(-\infty;u)} A\neq\varnothing$, однак $\bigcup_{A\in S, A\subset(u;v)} A=\varnothing$. Наприклад, коли $\Omega=[a;b)$, простір подій S породжений за поділом множини Ω на підмножини $[a_{i-1};a_i)$, $i\in\overline{1,k}$, такі, що $\bigcup_{i=1}^{k}[a_{i-1};a_i)=\Omega$, $A=\bigcup_{i\in I}[a_{i-1};a_i)\in S$, $I\subset\{1,2,...,k\}$, тобто $S=\{A \mid A=\bigcup_{i\in I}[a_{i-1};a_i),\, I\subset\{1,2,...,k\}\}$, $a_{i-1}<u<a_i$, $a_i<v<a_{i+1}$, тоді буде $\bigcup_{A\in S, A\subset(-\infty;v)} A \ \setminus \bigcup_{A\in S, A\subset(-\infty;u)} A=[a_{i-1};a_i)$, однак $\bigcup_{A\in S, A\subset[u;v)} A=\varnothing$, оскільки жоден з проміжків $[a_{i-1};a_i)$ не є підмножиною проміжка $[u;v)$. Тому не виключено, що

$$P_n^*\left(\bigcup_{A\in S, A\subset(u;v)} A\right)\neq P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;v)} A\right)-P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;u)} A\right)=F_n^*(v)-F_n^*(u).$$

Оскільки $\bigcup_{A\in S, A\subset[u;v)} A \subset \bigcup_{A\in S, A\subset(-\infty;v)} A \ \setminus \bigcup_{A\in S, A\subset(-\infty;u)} A$, то

$$P_n^*\left(\bigcup_{A\in S, A\subset[u;v)} A\right)\leq P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;v)} A \ \setminus \bigcup_{A\in S, A\subset(-\infty;u)} A\right)=$$

$$=P_n^*\left(\bigcup_{A\in S, A\subset[-\infty;v)} A\right)-P_n^*\left(\bigcup_{A\in S, A\subset(-\infty;u)} A\right)=F_n^*(v)-F_n^*(u).$$

Зауважимо, що записи вигляду $P_n^*([u;v))=F_n^*(v)-F_n^*(u)$, $F_n^*(x)=P_n^*((-\infty;x))$ тощо можуть виявитись некоректними, якщо множини $[u;v]$, $(-\infty;x)$ і т.п. не є елементами простору подій *S*, тобто не є подіями, бо ймовірнісна міра P_n^* визначена лише на елементах сукупності *S* підмножин множини Ω, тобто на подіях *A* із простору подій *S*.

Із рівності (2.7.2) слідує, що і за поточкового розподілу статистичних ймовірностей на скінченній множині $\Omega=\{x_1,x_2,...,x_m\}$, коли як події $A\in S$ разом з порожньою множиною $\varnothing$ розглядаються всеможливі об'єднання $A=\bigcup_{i\in I}H_i$,

$I \subset \{1,2,...,k\}$, $k \le m$, підмножин H_i множини Ω таких, що $H_i H_j = \varnothing$, коли $i \ne j$, $\bigcup_{i=1}^{k} H_i = \Omega$ (зокрема H_i можуть бути одноелементними підмножинами множини Ω), і за поінтервального розподілу статистичних ймовірностей на нескінченній неперервній множині $\Omega = [a;b)$, коли як події $A \in S$ разом з порожньою множиною $\varnothing$ розглядаються всеможливі об'єднання $A = \bigcup_{i \in I}[a_{i-1};a_i)$, $I \subset \{1,2,...,k\}$, інтервалів $[a_{i-1};a_i)$ таких, що $[a_{i-1};a_i) \bigcap [a_{j-1};a_j) = \varnothing$, коли $i \ne j$, $\bigcup_{i=1}^{k}[a_{i-1};a_i) = \Omega$, функція $F_n^*(x)$ кусково стала, тобто набуває одного і того самого значення, рівного $F_n^*(u)$, у всіх точках x із проміжка $[u;v]$, якщо $P_n^*\left(\bigcup_{A \in S, A \subset (-\infty;v)} A \setminus \bigcup_{A \in S, A \subset (-\infty;u)} A\right) = 0$, зокрема коли $\bigcup_{A \in S, A \subset (-\infty;v)} A \setminus \bigcup_{A \in S, A \subset (-\infty;u)} A = \varnothing$.

Очевидно, коли $\Omega = \{x_1, x_2, ..., x_k\}$, $A = \bigcup_{i \in I}\{x_i\} \in S$, $I \subset \{1,2,...,k\}$, і $u \in \{x_1, x_2, ..., x_k\}$, $v \in \{x_1, x_2, ..., x_k\}$, або коли $\Omega = [a;b) = \bigcup_{i=1}^{k}[a_{i-1};a_i)$, $A = \bigcup_{i \in I}[a_{i-1};a_i) \in S$, $I \subset \{1,2,...,k\}$, і $u \in \{a_0, a_1, ..., a_k\}$, $v \in \{a_0, a_1, ..., a_k\}$, тоді буде $[u\ v) \in S$ і $P_n^*[u;v) = F_n^*(v) - F_n^*(u)$.

Приклад 2.7.1. Нехай на множині $\Omega = \{1,2,3,4,5\}$ задано розподіл статистичних ймовірностей через ряд розподілу

x_i	1	2	3	4	5
$P_n^*(\{x_i\})$	0.05	0.20	0.50	0.20	0.05

Як події разом з порожньою множиною $\varnothing$ розглядатимемо всеможливі об'єднання $A = \bigcup_{i \in I}\{x_i\}$, $I \subset \{1,2,3,4,5\}$, $x_i = i$, $i \in \overline{1,5}$, тобто всеможливі підмножини множини $\Omega = \{1,2,3,4,5\}$. Таким чином $S = \{A \mid A = \bigcup_{i \in I}\{x_i\},\ I \subset \{1,\ 2,\ ...,\ 5\}\}$.

Тоді за даного розподілу статистичних ймовірностей буде:

1) $F_n^*(x) = 0$, коли $x \le 1$, оскільки жодна із непорожніх підмножин $\bigcup_{i \in I}\{x_i\}$, $I \subset \{1,2,3,4,5\}$, $I \ne \varnothing$, множини Ω

не є підмножиною множини $(-\infty; 1)$;

2) коли буде $1 < x \le 2$, тоді до множини $(-\infty; x)$ буде входити підмножина $\{1\}$ множини Ω, і тому

$$F_n^*(x) = P_n^*(\{1\}) = 0.05, \text{ коли } 1 < x \le 2;$$

3) коли x буде змінюватись в межах від 2 до 3 включно, тобто буде $2 < x \le 3$, тоді до множини $(-\infty; x)$ будуть входити підмножини $\{1\}$, $\{2\}$, $\{1,2\}$, які є елементами сукупності S, тобто подіями, і об'єднання яких $\{1\} \bigcup \{2\} \bigcup \{1,2\}$ буде множиною, що входить до сукупності S, тобто подією, і крім того буде підмножиною множини $(-\infty; x)$, коли $x \in (2; 3]$. Тому

$$F_n^*(x) = P_n^*\Big(\bigcup_{A \in S,\ A \subset (-\infty;\, x)} A\Big) = P_n^*(\{1\} \bigcup \{2\} \bigcup \{1,2\}) = P_n^*(\{1,2\}) =$$

$$= P_n^*(\{1\}) + P_n^*(\{2\}) = 0.25,$$

коли $2 < x \le 3$, тобто $F_n^*(x) = 0.25$, коли $2 < x \le 3$.

Міркуючи аналогічно, знайдемо

4) $F_n^*(x) = 0.75$, коли $3 < x \le 4$;

5) $F_n^*(x) = 0.95$, коли $4 < x \le 5$;

6) $F_n^*(x) = 1.00$, коли $5 < x$.

Остаточно одержуємо

$$F_n^*(x) = \begin{cases} 0, & \text{коли } x \le 1, \\ 0.05, & \text{коли } 1 < x \le 2, \\ 0.25, & \text{коли } 2 < x \le 3, \\ 0.75, & \text{коли } 3 < x \le 4, \\ 0.95, & \text{коли } 4 < x \le 5, \\ 1.00, & \text{коли } 5 < x. \end{cases}$$

Графік такої функції $F_n^*(x)$ стосовно заданого дискретного поточкового розподілу статистичних ймовірностей на скінченній множині точок $\Omega = \{1, 2, 3, 4, 5\}$ подано на Рис. 2.7.1.

Підкреслимо, що в даному прикладі множини вигляду $(-\infty; x)$ не є подіями відносно ймовірнісного простору (Ω, S, P_n^*), бо таких множин немає в сукупності S підмножин розглядуваної множини Ω, тобто в просторі подій, до якого разом з порожньою множиною $\varnothing$ віднесено підмножини

$\bigcup\limits_{i \in I}\{x_i\}$, $I \subset \{1,2,3,4,5\}$, множини Ω, а тому записи вигляду $F_n^*(x) = P_n^*((-\infty;\, x))$ в розглядуваному випадку некоректні.

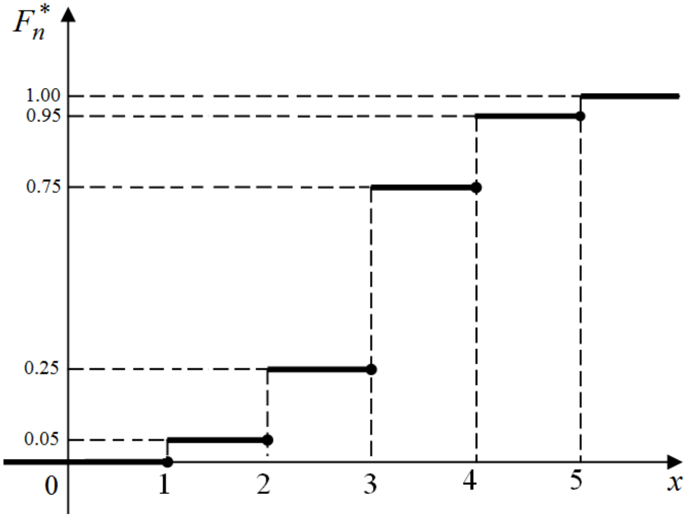


Рис. 2.7.1

Приклад 2.7.2. Нехай на множині $\Omega = (0\,;5]$ задано поінтервальний розподіл статистичних ймовірностей

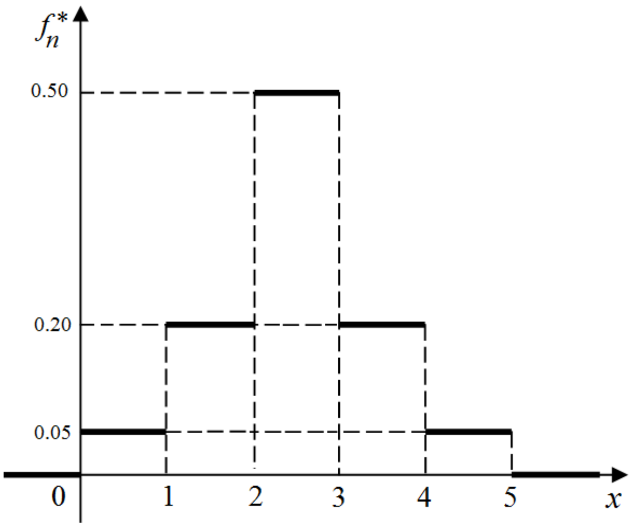


Рис. 2.7.2

$(a_{i-1};a_i]$	$(0;1]$	$(1;2]$	$(2;3]$	$(3;4]$	$(4;5]$
$P_n^*((a_{i-1};a_i])$	0.05	0.20	0.50	0.20	0.05

Графік усередненої щільності $f_n^*(x)$ цього розподілу статистичних ймовірностей за інтервалами $(i-1;i]$, $i\in\overline{1,5}$, на множині $\Omega=(0;5]=(0;1]\bigcup(1;2]\bigcup(2;3]\bigcup(3;4]\bigcup(4;5]$ подано на Рис. 2.7.2.

Як події разом з порожньою множиною $\varnothing$ розглядатимемо всеможливі об'єднання інтервалів $(a_{i-1};a_i]$, тобто

$$S=\{A \mid A=\bigcup_{i\in I}(a_{i-1};a_i],\ I\subset\{1,\ 2,\ 3,\ 4,\ 5\}\}.$$

Нехай $\tilde{\Omega}=(-\infty;\infty)\supset\Omega$, $\tilde{S}=\mathscr{B}(R^1)\supset S$. Тоді, враховуючи формулу (2.7.1), стосовно даного поінтервального розподілу статистичних ймовірностей одержуємо: $F_n^*(x)=0$, коли $x\le 1$, оскільки жодна з підмножин $A=\bigcup_{i\in I}[a_{i-1};a_i)\in S$, $I\subset\{1,2,3,4,5\}$, $I\ne\varnothing$, не є підмножиною множини $(-\infty;x)$, коли $x\le 1$. Коли x буде в межах від 1 до 2 включно, тобто $x\in(1;2]$, тоді інтервал $(0,1]$ буде підмножиною множини $(-\infty;x)$, і тому за $x\in(1;2]$ буде

$$F_n^*(x)=P_n^*((0;\ 1])=\int_0^1 f_n^*(x)dx=0.05,\ \text{коли}\ x\in(1;2].$$

Коли x буде в межах від 2 до 3 включно, тобто $x\in(2;3]$, тоді сума подій $(0;1]\in S$, $(1;2]\in S$, $(0;1]\bigcup(1;2]\in S$, тобто $(0;1]\bigcup(1;2]\bigcup((0;1]\bigcup(1;2])\in S$ буде підмножиною множини $(-\infty;x)$, $x\in(2;3]$, тому

$$F_n^*(x)=P_n^*((0;1]\bigcup(1;2]\bigcup((0;1]\bigcup(1;2]))=$$
$$=P_n^*((0;1]\bigcup(1;2])=P_n^*((0;1])+P_n^*((1;2])=$$
$$=\int_0^1 f_n^*(x)dx+\int_1^2 f_n^*(x)dx=0.05+0.20=0.25,$$

коли $x\in(2;3]$.

Міркуючи аналогічно, знайдемо

$$F_n^*(x)=0.75,\ \text{коли}\ x\in(3;4];$$
$$F_n^*(x)=0.95,\ \text{коли}\ x\in(4;5];$$
$$F_n^*(x)=1.00,\ \text{коли}\ x\in(5;+\infty).$$

Таким чином

$$F_n^*(x)=\begin{cases}0, & \text{коли } x\le 1,\\ 0.05, & \text{коли } 1<x\le 2,\\ 0.25, & \text{коли } 2<x\le 3,\\ 0.75, & \text{коли } 3<x\le 4,\\ 0.95, & \text{коли } 4<x\le 5,\\ 1.00, & \text{коли } 5<x.\end{cases}$$

Вигляд графіка так визначеної функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $\widetilde{\Omega}=(-\infty;\,\infty)\supset\Omega=(0;\,5]$, розглядуваного в прикладі 2.7.2, буде такий самий, як і вигляд графіка функції $F_n^*(x)$ дискретного поточкового розподілу статистичних ймовірностей на множині $\Omega=\{1,2,3,4,5\}$, розглянутого в прикладі 2.7.1 (див. Рис. 2.7.1, Рис. 2.7.5 а)). Однаковий вигляд мають і подання самих функцій $F_n^*(x)$ розподілу узагальнених статистичних ймовірностей в обох розглядуваних в прикладах 2.7.1 і 2.7.2 випадках.

Тому у випадках розглянутого способу побудови простору подій *S*, $S=\{A\,\big|\,A=\bigcup\limits_{i\in I}H_i,\ I\subset\{1,\,2,\,\dots,k\}\}$, *коли як події разом з порожньою множиною* $\varnothing$ *розглядаються всеможливі об'єднання* $A=\bigcup\limits_{i\in I}H_i\in S$, $I\subset\{1,2,...,k\}$, *підмножин* (*можливо одноточкових*) H_i, $i\in\overline{1,k}$, *скінченної множини* $\Omega=\{x_1,x_2,...,x_m\}$, $(k\le m)$, *чи всеможливі об'єднання підмножин* $H_i=(a_{i-1};\,a_i]$, $i\in\overline{1,k}$, *нескінченної множини* $\Omega=(a\,;\,b]$, *таких, що* $H_iH_j=\varnothing$, *коли* $i\neq j$, $\bigcup\limits_{i=1}^{k}H_i=\Omega$, *і* $(-\infty;\,x)\,\overline{\in}\,S$ *за довільних* $x\in(-\infty;\,\infty)$, *за виглядом опису функції* $F_n^*(x)$ *розподілу узагальнених статистичних ймовірностей неможливо визначити, розглядається поточковий розподіл статистичних ймовірностей на скінченній множині* $\Omega=\{x_1,x_2,...,x_m\}$, *чи поінтервальний розподіл статистичних ймовірностей на нескінченній неперервній множині*

$$\Omega=(a\,;\,b]=\bigcup_{i=1}^{k}(a_{i-1};\,a_i],\ (a_{i-1};\,a_i]\bigcap(a_{j-1};\,a_j]=\varnothing,\ \text{коли } i\neq j\,.$$

Приклад 2.7.3. Нехай на множині $\Omega=(0;\,5]$ задано поінтервальний розподіл статистичних ймовірностей такий самий, як і в прикладі 2.7.2, а простір подій *S* породжений за поділом множини Ω на 25 інтервалів довжиною 0.2, тобто як події разом з порожньою множиною $\varnothing$ розглядаються всеможливі об'єднання $\bigcup\limits_{i\in I}(a_{i-1};\, a_i]$, $I \subset \{1, 2, ..., 25\}$, інтервалів $(a_{i-1};\, a_i]$, $i \in \overline{1, 25}$, $a_0 = 0$, $a_i - a_{i-1} = 0.2$.

Це означає, що розглядається новий простір $\hat{S} = \{A \mid A = \bigcup\limits_{i\in I}(a_{i-1};a_i],\;\; I \subset \{1,\, 2,\, \dots,\, 25\}\}$, породжений за сукупністю проміжків $(a_{i-1};\, a_i]$, $i \in \overline{1, 25}$, елементами якого є множини виду $A = \bigcup\limits_{i\in I}(a_{i-1};\, a_i]$, $I \subset \{1, 2, ..., 25\}$. Очевидно, всі події із простору S, розглядуваного в прикладі 2.7.2, є також і елементами простору $\hat{S}$, тобто $S \subset \hat{S}$. Нову ймовірнісну міру $\hat{P}_n^*$ на просторі подій $\hat{S}$ визначимо за формулою

$$\hat{P}_n^*(A) = \int\limits_A f_n^*(x)dx = \sum_{x\in(a_{i-1};\, a_i]\subset A} f_n^*(x)\cdot(a_i - a_{i-1}),\;\; A \in \hat{S},$$

де $f_n^*(x)$ – щільність розподілу статистичних ймовірностей така сама, як і в прикладі 2.7.2 (див. Рис. 2.7.2).

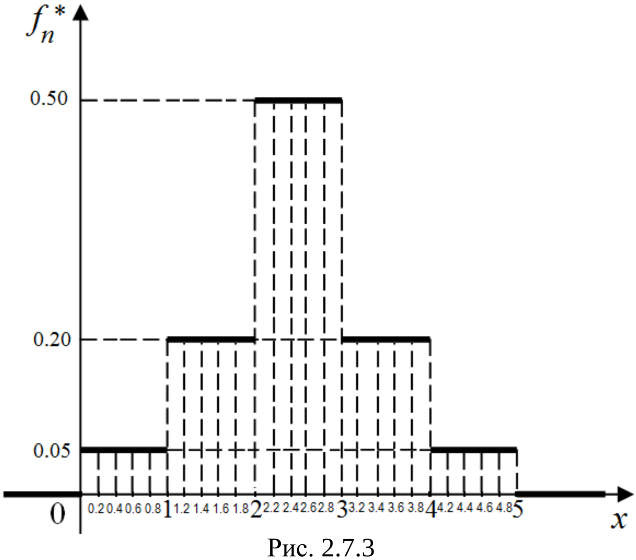


Рис. 2.7.3

Очевидно, за тієї самої усередненої щільності $f_n^*(x)$ розподілу статистичних ймовірностей, що і в прикладі 2.7.2, статистичні ймовірності попадання в проміжки, отримані подрібненням проміжків із прикладу 2.7.2 на 5 проміжків однакової довжини, в 5 разів меншої, ніж довжина вихідних проміжків, будуть в 5 разів менші, ніж статистичні ймовірності попадання у вихідні проміжки, і будуть, як і раніше, обчислюватися за формулою

$$\hat{P}_n^*((a_{i-1}, a_i]) = f_n^*(x)(a_i - a_{i-1}),\ i \in \overline{1, 25}.$$

Очевидно, $\hat{P}_n^*(A) = P_n^*(A)$, коли $A \in S$. В такому разі говорять, що міра $\hat{P}_n^*(A)$, $A \in \hat{S}$, є продовженням міри $P_n^*(A)$, $A \in S$, із простору подій S на простір подій $\hat{S}$.

Якщо проміжки $(a_{i-1}; a_i]$, з яких складаються події $A = \bigcup\limits_{i \in I}(a_{i-1}; a_i] \in \hat{S}$, $I \subset \{1, 2, ..., 25\}$, поділити кожен на якесь число ще дрібніших проміжків однакової довжини, отримаємо новий простір подій $\hat{\hat{S}}$ і нову ймовірнісну міру $\hat{\hat{P}}_n^*$ цілком аналогічно до попереднього. Таке подрібнення проміжків $(a_{i-1}; a_i]$ можна продовжувати як завгодно довго до тих пір, поки різниця $h = a_i - a_{i-1}$ стане меншою, ніж будь яке як завгодно мале наперед задане число $\varepsilon > 0$. В результаті кожного зменшення довжини $h = a_i - a_{i-1}$ проміжків $(a_{i-1}; a_i]$, однакової за всіх i, будемо отримувати все нові і нові ймовірнісні простори $(\Omega, S^{(j)}, P_n^{*(j)})$, $j \in \{0, 1, 2, 3, ...\}$, (з одними і тими самими Ω і $f_n^*(x)$).

Очевидно, функція $F_n^*(x)$, визначена за формулою (2.7.1), набуватиме сталих значень на проміжках $(a_{i-1}; a_i]$, $i \in \overline{1, 25}$, і в разі переходу через точку a_i набуватиме приріст

$$f_n^*(x) \cdot (a_i - a_{i-1}),\ x \in (a_{i-1}; a_i]. \qquad (2.7.3)$$

Міркуючи аналогічно до того, як це було зроблено в прикладі 2.7.2 під час побудови функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $\Omega = (0; 5]$ за інтервалами $(0; 1]$, $(1; 2]$, $(2; 3]$, $(3; 4]$, $(4; 5]$, в розглядуваному прикладі одержимо:

$F_n^*(x) = 0$, коли $x \le 0.2$, $F_n^*(x) = 0.01$, коли $0.2 < x \le 0.4$,

$F_n^*(x) = 0.02$, коли $0.4 < x \le 0.6$, $F_n^*(x) = 0.03$, коли $0.6 < x \le 0.8$,

$F_n^*(x) = 0.04$, коли $0.8 < x \le 1.0$, $F_n^*(x) = 0.05$, коли $1.0 < x \le 1.2$,

$F_n^*(x) = 0.09$, коли $1.2 < x \le 1.4$, $F_n^*(x) = 0.13$, коли $1.4 < x \le 1.6$,

$F_n^*(x) = 0.17$, коли $1.6 < x \le 1.8$, $F_n^*(x) = 0.21$, коли $1.8 < x \le 2.0$,

$F_n^*(x) = 0.25$, коли $2.0 < x \le 2.2$, $F_n^*(x) = 0.35$, коли $2.2 < x \le 2.4$,

$F_n^*(x) = 0.45$, коли $2.4 < x \le 2.6$, $F_n^*(x) = 0.55$, коли $2.6 < x \le 2.8$,

$F_n^*(x) = 0.65$, коли $2.8 < x \le 3.0$, $F_n^*(x) = 0.75$, коли $3.0 < x \le 3.2$,

$F_n^*(x)=0.79$, коли $3.2<x\le 3.4$, $F_n^*(x)=0.83$, коли $3.4<x\le 3.6$, $F_n^*(x)=0.87$, коли $3.6<x\le 3.8$, $F_n^*(x)=0.91$, коли $3.8<x\le 4.0$, $F_n^*(x)=0.95$, коли $4.0<x\le 4.2$, $F_n^*(x)=0.96$, коли $4.2<x\le 4.4$, $F_n^*(x)=0.97$, коли $4.4<x\le 4.6$, $F_n^*(x)=0.98$, коли $4.6<x\le 4.8$, $F_n^*(x)=0.99$, коли $4.8<x\le 5.0$, $F_n^*(x)=1.00$, коли $5.0<x$.

На Рис. 2.7.5 б) подано графік так визначеної функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $\Omega=\bigcup_{i=1}^{25}(a_{i-1};\, a_i]$, $a_0=0$, $a_i-a_{i-1}=0.2$, $i\in\overline{1,25}$, за інтервалами $(a_{i-1},a_i]$, $i\in\overline{1,25}$, із щільністю (див. Рис. 2.7.3, Рис. 2.7.4)

$$f_n^*(x)=\begin{cases}0, & \text{коли } x\,\overline{\in}\,(a_0;\, a_{25}],\\ 0.05, & \text{коли } x\in(0;1]\cup(4;5],\\ 0.20, & \text{коли } x\in(1;\,2]\cup(3;\,4],\\ 0.50, & \text{коли } x\in(2;\,3].\end{cases}$$

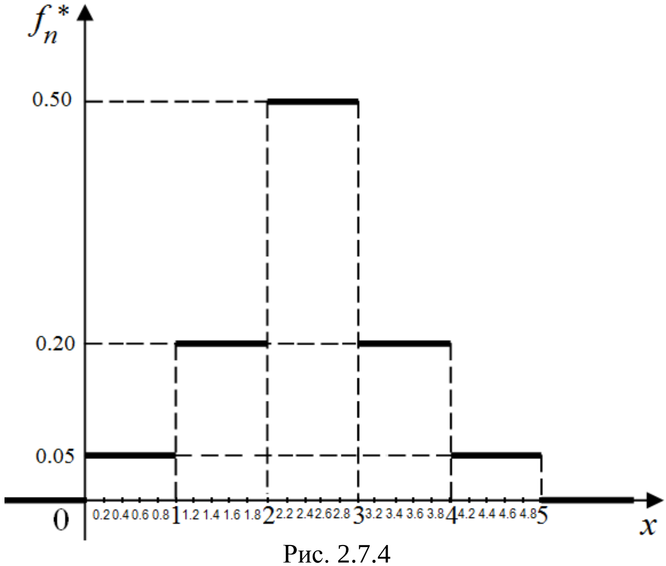


Рис. 2.7.4

Зауважимо, що в програмі Gran1 передбачено послугу, за якою із подрібненням інтервалів $(a_{i-1};\, a_i]$ (за рахунок збільшення їх кількості) автоматично перебудовується графік функції $F(x)$ відповідного поінтервального розподілу узагальнених статистичних ймовірностей (див. Рис. 2.7.5 а) -Рис. 2.7.5 д)).

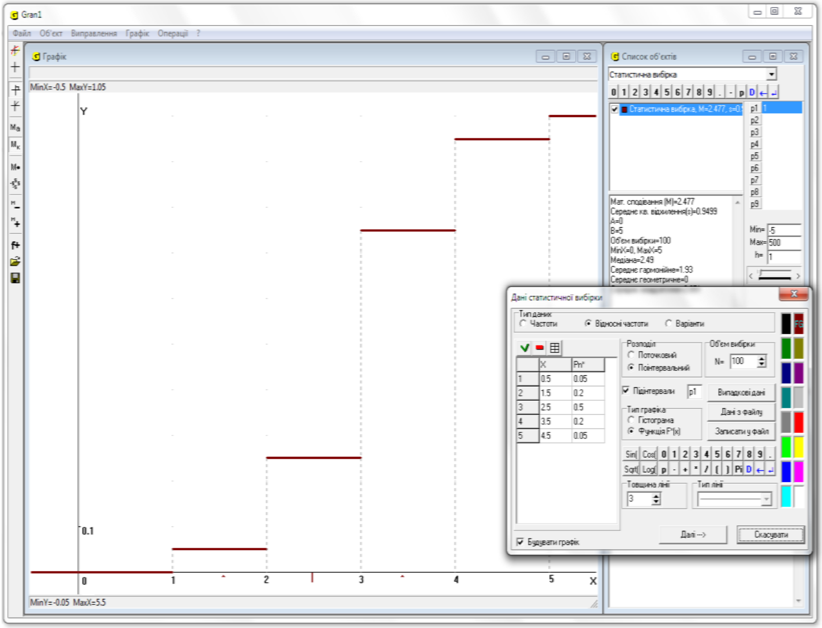


Рис. 2.7.5 а)

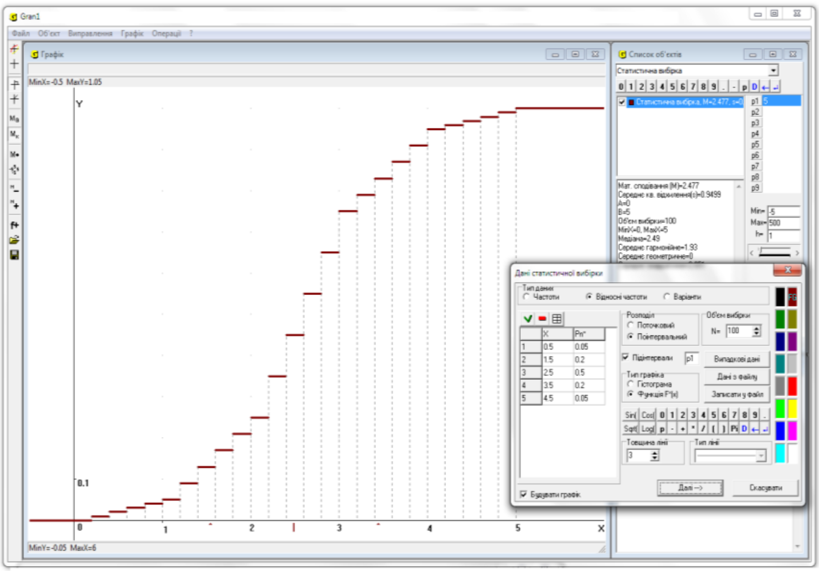


Рис. 2.7.5 б)

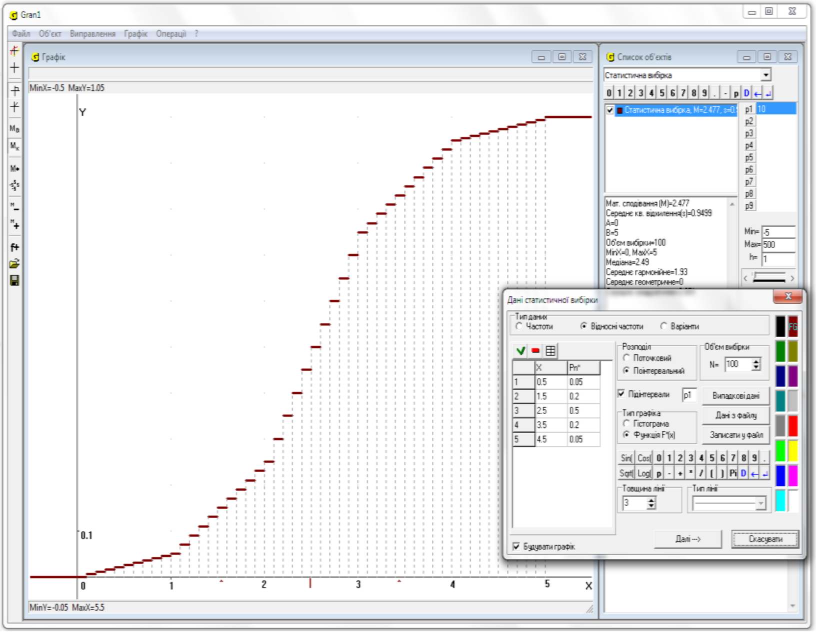


Рис. 2.7.5 в)

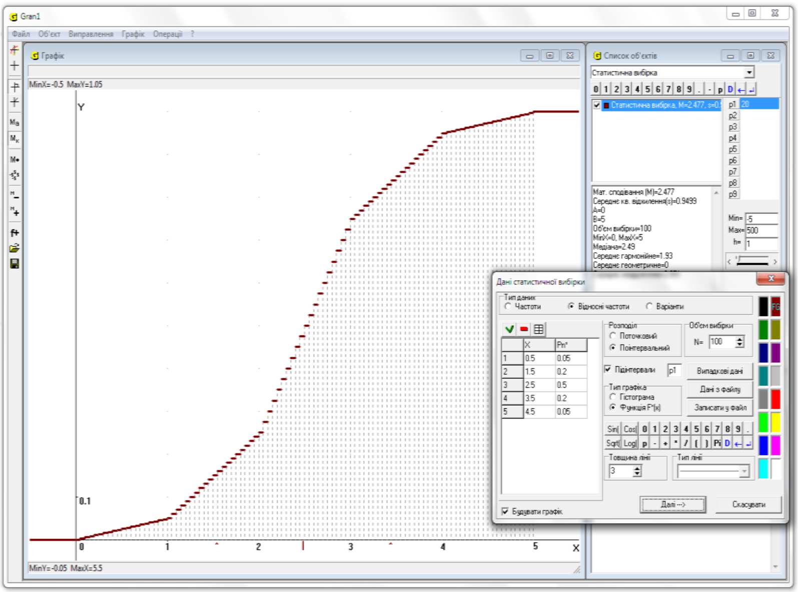


Рис. 2.7.5 г)

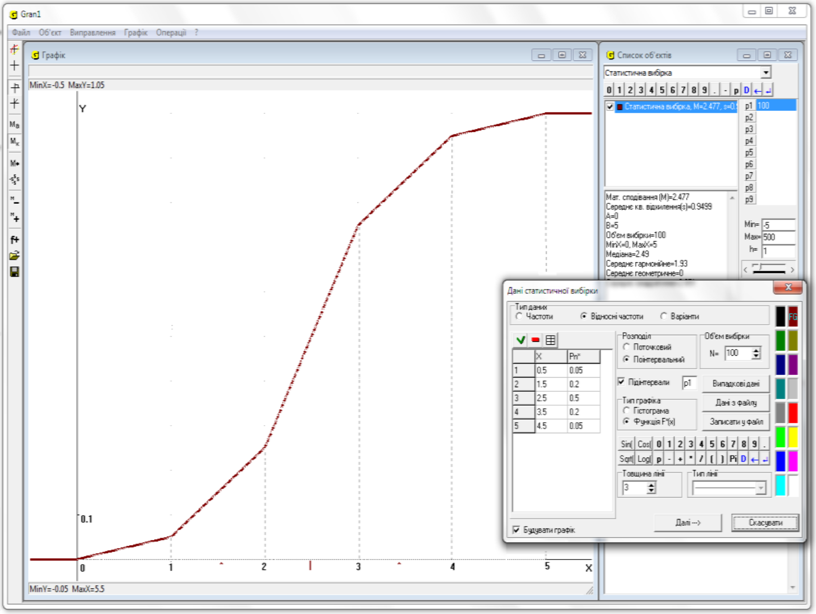


Рис. 2.7.5 д)

Для збільшення чи зменшення кількості інтервалів за рахунок їх подрібнення чи укрупнення, досить, використовуючи послуги програми Gran1, збільшувати чи зменшувати значення відповідного параметра $P \in \{P1, P2, ..., P9\}$, вибравши крок його зміни h рівним 1, щоб кількість інтервалів була цілим числом (див. Рис. 2.7.5 б)).

На Рис. 2.7.5 а) подано графік функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на проміжку $(0; 5]$, коли $h = 1$, а $f_n^*(x)$ із прикладу 2.7.2 (Рис. 2.7.4), а на Рис. 2.7.5 б) - Рис. 2.7.5 д) подано графіки функцій $F_n^*(x)$ поінтервальних розподілів узагальнених статистичних ймовірностей на тому самому проміжку $(0; 5]$, коли довжину інтервала зменшено відповідно в 5 разів (Рис. 2.7.5 б)), в 10 разів (Рис. 2.7.5 в)), в 20 разів (Рис. 2.7.5 г)), в 100 разів (Рис. 2.7.5 д)) (див. значення параметра $P1$ на рисунках), а $f_n^*(x)$ така сама, як і раніше (Рис. 2.7.4).

Зауважимо, що коли за довільних $x \in (-\infty; \infty)$ множини $(-\infty; x)$ є подіями, тобто елементами простору подій S, $(-\infty; x) \in S$, що означає, що як простір елементарних подій розглядається множина $\Omega = R^1 = (-\infty;\, \infty)$, а як простір подій – $\mathscr{B}(R^1)$, тоді $\bigcup\limits_{A\in S,\ A\subset(-\infty;\, x)} A = (-\infty;\, x)$, і формула (2.7.1) набуває вигляду $F_n^*(x) = P_n^*((-\infty; x))$, $x \in (-\infty;\, \infty)$. *В такому разі за виглядом функції $F_n^*(x)$ можна з'ясувати, така функція побудована за поточковим розподілом статистичних ймовірностей на скінченній множині точок $\{x_1, x_2, ..., x_k\} \subset \Omega = (-\infty;\, \infty)$, чи за поінтервальним розподілом статистичних ймовірностей на неперервній множині* $(a, b] \subset \Omega = (-\infty, \infty)$.

Приклад 2.7.4. В разі поточкового розподілу статистичних ймовірностей на скінченній множині $\{1, 2, 3, 4, 5\} \subset \Omega = (-\infty;\, \infty)$, розглянутого в прикладі 2.7.1, за умови $(-\infty;\, x) \in S$ за довільного $x \in (-\infty;\, \infty)$ функція розподілу статистичних ймовірностей

$$F_n^*(x) = P_n^*((-\infty; x)),\ (-\infty;\, x) \in S,\ x \in (-\infty;\, \infty),$$

буде кусково сталою і матиме той самий вигляд, що і раніше, тобто

$$F_n^*(x) = \begin{cases} 0, & \text{коли } x \le 1, \\ 0.05, & \text{коли } 1 < x \le 2, \\ 0.25, & \text{коли } 2 < x \le 3, \\ 0.75, & \text{коли } 3 < x \le 4, \\ 0.95, & \text{коли } 4 < x \le 5, \\ 1.00, & \text{коли } 5 < x. \end{cases}$$

Графік цієї функції такий самий, що і на Рис. 2.7.1.

Разом з тим, за поінтервального розподілу статистичних ймовірностей на множині $(0;\, 5] \subset \Omega = (-\infty;\, \infty)$, розглянутого в прикладі 2.7.2, за умови $(-\infty; x) \in S = \mathscr{B}(R^1)$ за довільного $x \in (-\infty;\, \infty)$ функція розподілу статистичних ймовірностей

$$F_n^*(x) = P_n^*((-\infty; x)) = \int\limits_{-\infty}^{x} f_n^*(x)dx,\ (-\infty;\, x) \in S,\ x \in (-\infty;\, \infty),$$

буде неперервною і її опис набуде вигляду

$$F_n^*(x) = \int\limits_{-\infty}^{\infty} f_n^*(x)dx = \begin{cases} 0, & \text{коли } x \le a_0, \\ f_n^*(x)\cdot(x-a_0), & \text{коли } x \in (a_0;a_1], \\ \sum\limits_{i=1}^{j-1} P_n^*([a_{i-1};a_i)) + f_n^*(x)(x-a_{j-1}), & \\ & \text{коли } x \in (a_{j-1};a_j], \quad 2 \le j \le k, \\ 1, & \text{коли } a_k < x, \end{cases}$$

тобто за наведених даних буде

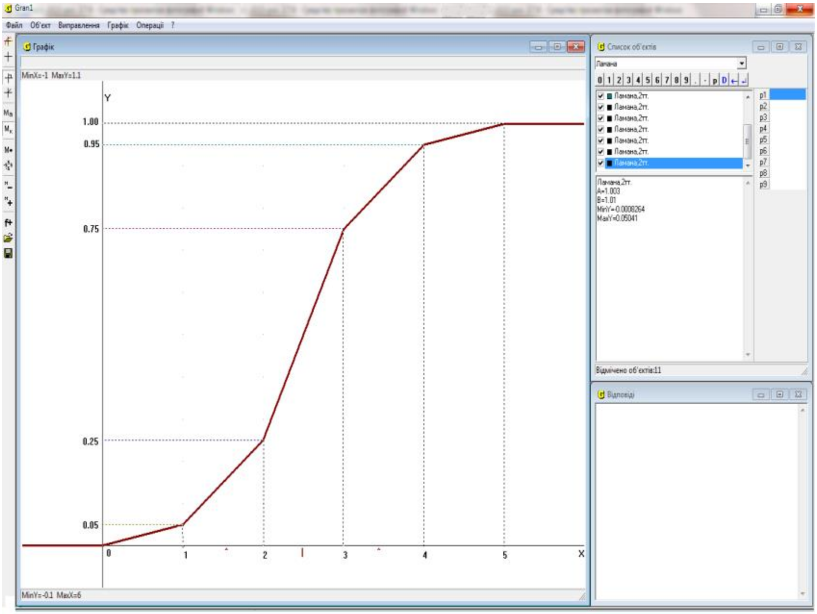


Рис. 2.7.6

$$F_n^*(x) = \begin{cases} 0, & \text{коли } x \le 0, \\ 0.05x, & \text{коли } 0 < x \le 1, \\ 0.05 + 0.20(x-1), & \text{коли } 1 < x \le 2, \\ 0.25 + 0.50(x-2), & \text{коли } 2 < x \le 3, \\ 0.75 + 0.20(x-3), & \text{коли } 3 < x \le 4, \\ 0.95 + 0.05(x-4), & \text{коли } 4 < x \le 5, \\ 1, & \text{коли } 5 < x. \end{cases}$$

Графік останньої функції $F_n^*(x)$ подано на Рис. 2.7.6.

Зауважимо, що коли усереднена щільність $f_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей задана на інтервалах $(a_{i-1}; a_i]$, $i \in \overline{1,k}$, фіксованої довжини h, таких, що $(a_{i-1}; a_i] \bigcap (a_{j-1}; a_j] = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} (a_{i-1}; a_i] = \Omega$, $a_i - a_{i-1} = h$ (наприклад, на інтервалах довжини $h = 1$, як в прикладі 2.7.2, див. Рис. 2.7.2), а простір подій S породжується за поділом множини Ω на все дрібніші і дрібніші інтервали такі, що довжина найдовшого з таких інтервалів стає все меншою і меншою (див. приклад 2.7.3, Рис. 2.7.5 а) – Рис. 2.7.5 д)), тоді функція $F_n^*(x)$ такого поінтервального розподілу узагальнених статистичних ймовірностей із вказаною щільністю $f_n^*(x)$ за все більшого і більшого подрібнення інтервалів, з яких складаються події $A \in S$, все менше і менше відрізнятиметься від неперервної функції $F_n^*(x)$ поінтервального розподілу статистичних ймовірностей, побудованої за умови, що як події розглядаються всеможливі множини $(-\infty; x)$, $x \in (-\infty; \infty)$, тобто множини $(-\infty; x)$ є елементами простору подій $S = \mathscr{B}(R^1)$ (див. приклади 2.7.3, 2.7.4).

Зауважимо, що *коли* $F_n^*(x)$ *неперервна, тоді розподіл узагальнених статистичних ймовірностей на множині* $\Omega = (-\infty; +\infty)$ *називається неперервним,* а $(-\infty; x) \in S$ за довільних $x \in (-\infty; \infty)$. У випадку, коли $(-\infty; u) \in S$, $(-\infty; v) \in S$, $u < v$, $u \in (-\infty; \infty)$, $v \in (-\infty; \infty)$, як випливає з властивостей 1_s-3_s простору подій S, $[u; v) = (-\infty; v) - (-\infty; u)$ також є елементом простору подій S, тобто $[u; v) \in S$, бо коли $A \in S$ і $B \in S$, то $B \setminus A = B \bigcap \overline{A} = \overline{\overline{B \bigcap \overline{A}}} = \overline{\overline{B} \bigcup \overline{\overline{A}}} = \overline{\overline{B} \bigcup A} \in S$ за властивостями 1_s-3_s простору подій S. Тому, як видно з формули (2.7.2), буде $P_n^*([u; v)) = F_n^*(v) - F_n^*(u)$, бо в розглядуваному випадку за $A \in S$ таких, що $A \subset (-\infty; v)$, та $A \in S$ таких, що $A \subset (-\infty; u)$, буде мати місце рівність

$$\bigcup_{A \subset (-\infty; v)} A \setminus \bigcup_{A \subset (-\infty; u)} A = (-\infty; v) \setminus (-\infty; u) = [u; v) \in S .$$

Зауважимо, що ймовірнісні міри P_n^*, розглянуті в прикладах 2.7.3 і 2.7.4, є гіпотетичні, знайдені за припущення, що щільність $f_n^*(x)$ розподілу статистичних ймовірностей набуває вигляду (див. Рис. 2.7.2, Рис. 2.7.4)

$$f_n^*(x)=\begin{cases}0, & \text{коли } x \overline{\in} (0;5],\\ 0.05, & \text{коли } x\in(0;1]\cup(4;5],\\ 0.20, & \text{коли } x\in(1;2]\cup(3;4],\\ 0.50, & \text{коли } x\in(2;3],\end{cases}$$

тобто за припущення, що щільність розподілу статистичних ймовірностей на проміжках $(0;1]$, $(1;2]$, $(2;3]$, $(3;4]$, $(4;5]$ набуває сталих значень відповідно 0.05, 0.20, 0.50, 0.20, 0.05, і значення 0 за межами проміжка $(0;5]$, що в дійсності може виявитись не так (див. §2.4, Рис. 2.4.2).

Вправи для самостійного виконання

2.7.1. Функції $F_n^*(x)$ дискретного поточкового та поінтервального розподілу узагальнених статистичних ймовірностей визначені за формулою (2.7.1). Перевірити, чи правильні твердження:

1. Функція $F_n^*(x)$ поточкового розподілу узагальнених статистичних ймовірностей на множині $\Omega=\{x_1, x_2, ..., x_k\}$ із зростанням x може спадати.

2. Значення функції $F_n^*(x)$ поточкового розподілу узагальнених статистичних ймовірностей на множині $\Omega=\{x_1, x_2, ..., x_k\}$ можна визначити в будь-якій точці проміжка $[x_1; x_k]$.

3. Значення функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $\Omega=\bigcup_{i=1}^{k}[a_{i-1}; a_i)=[a_0; a_k)$ можна визначити в будь-якій точці проміжка $[a_0; a_k)$.

4. Функція $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на інтервалі $[a_{i-1}; a_i)$ набуває тим більшого приросту, чим більша усереднена щільність $f_n^*(x)$ розподілу узагальнених статистичних імовірностей на такому інтервалі.

5. Знаючи функцію $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $\Omega=\bigcup_{i=1}^{k}[a_{i-1}; a_i)$, можна визначити усереднену щільність $f_n^*(x)$ розподілу узагальнених статистичних ймовірностей на інтервалах $[a_{i-1}; a_i)$, $i\in\overline{1,k}$.

6. Знаючи усереднену щільність $f_n^*(x)$ розподілу узагальнених статистичних ймовірностей за інтервалами $[a_{i-1}; a_i)$, $i\in\overline{1,k}$, можна визначити функцію $F_n^*(x)$ поінтервального розподілу узагальнених

статистичних імовірностей на множині $\Omega = \bigcup_{i=1}^{k} [a_{i-1}; a_i)$.

7. Знаючи усереднену щільність $f_n^*(x)$ розподілу узагальнених статистичних імовірностей на інтервалах $[a_{i-1}; a_i)$, $i \in \overline{1,k}$, можна визначити узагальнену статистичну ймовірність попадання в будь-який проміжок $(\alpha; \beta)$ такий, що $(\alpha; \beta) \subset \bigcup_{i=1}^{k} [a_{i-1}; a_i)$

8. Знаючи функцію $F_n^*(x)$ поінтервального розподілу узагальнених статистичних імовірностей на множині $\Omega = \bigcup_{i=1}^{k} [a_{i-1}; a_i)$, можна визначити узагальнену статистичну ймовірність попадання в будь-який проміжок $(\alpha; \beta)$ такий, що $(\alpha; \beta) \subset \bigcup_{i=1}^{k} [a_{i-1}; a_i)$.

2.7.2. 1. Задано простір Ω елементарних подій $\Omega = \{x_1, x_2, x_3\}$, де x_1, x_2, x_3 – точки на числовій прямій. Побудувати всі можливі простори подій, за кожного з яких задати розподіл статистичних ймовірностей та побудувати відповідні функції розподілу узагальнених статистичних ймовірностей.

2. Задано неперервну множину $\Omega = \bigcup_{i=1}^{5} [i-1; i)$. Простір S подій породжується за поділом множини Ω на підмножини $[i-1; i)$. Усереднена щільність $f_n^*(x)$ розподілу узагальнених статистичних імовірностей набуває одного і того самого значення на підмножинах $[i-1; i)$. Визначити функцію $F_n^*(x)$ поінтервального розподілу узагальнених статистичних імовірностей на множині Ω за заданих умов.

3. Використовуючи програму Gran1, з'ясувати, який вигляд матиме функція поінтервального розподілу узагальнених статистичних імовірностей на множині $[0; 5)$, заданого через щільність розподілу узагальнених статистичних ймовірностей

$$f_n^*(x) = \begin{cases} 0, & \text{коли } x \,\overline{\in}\, [0;5), \\ 0.10, & \text{коли } x \in [0;1) \bigcup [4;5), \\ 0.20, & \text{коли } x \in [1;2) \bigcup [3;4), \\ 0.40, & \text{коли } x \in [2;3), \end{cases}$$

якщо сукупність S породжена за поділом множини $[0; 5)$:

а) на 10 інтервалів $[a_{i-1}; a_i)$ довжиною 0.50, $a_0 = 0$, $a_{10} = 5$, $a_i - a_{i-1} = 0.50$, $i \in \overline{1,10}$?

б) на 20 інтервалів $[a_{i-1}; a_i)$ довжиною 0.25, $a_0 = 0$, $a_{20} = 5$, $a_i - a_{i-1} = 0.25$, $i \in \overline{1,20}$?

в) на 25 інтервалів $[a_{i-1}; a_i)$ довжиною 0.20, $a_0 = 0$, $a_{25} = 5$, $a_i - a_{i-1} = 0.20$, $i \in \overline{1,25}$?

г) на 50 інтервалів $[a_{i-1}; a_i)$ довжиною 0.1, $a_0 = 0$, $a_{50} = 5$, $a_i - a_{i-1} = 0.1$, $i \in \overline{1,50}$?

д) як зміниться вигляд функції поінтервального розподілу узагальнених статистичних ймовірностей на множині $[0;5)$, якщо вважати, що $\Omega = (-\infty; \infty) = R^1$, $S = \mathscr{B}(R^1)$, а ймовірнісна міра P на S задана через ту саму щільність розподілу ймовірностей $f_n^*(x)$.

4. Використовуючи програму Gran1, за даними із прикладів 2.7.2., 2.7.3 побудувати графік функції $F_n^*(x)$ поінтервального розподілу узагальнених статистичних ймовірностей на множині $[0;5)$, коли $h = 0.01$.

2.8. Деякі числові характеристики поточкового розподілу статистичних ймовірностей в одновимірному координатному просторі

Приклад 2.8.1. Розглянемо в одновимірному координатному просторі два розподіли статистичних ймовірностей:

a)

x_i	0	1	2
$P_n^*(\{x_i\})$	0.1	0.8	0.1

b)

$x_i^/$	100	101	102
$P_n^*(\{x_i^/\})$	0.1	0.8	0.1

Ці розподіли відрізняються тим, що кожне із значень $x_i^/$ в розподілі *b*) на 100 більше, ніж в розподілі *a*). В розподілі *a*) статистичні ймовірності розподілені на множині із трьох точок x_1=0, x_2=1, x_3=2, що розташовані не далі, ніж на відстані, рівній 1, від точки *x*=1, а в розподілі *b*) статистичні ймовірності розподілені так само на множині із трьох точок $x_1^/=100, x_2^/=101, x_3^/=102$, що віддалені не далі, ніж на відстані 1 від точки $x^/=101$. Точки $\bar{x}=1$ та $\bar{x}^/=101$ природно назвати центрами розподілу (чи розсіювання) статистичних ймовірностей відповідно стосовно розподілів *a*) та *b*). За ними характеризуються значення, навколо яких зосереджуються спостережені значення.

Нехай проведено *n* спостережень, в результаті яких дістали спостережені значення $x_{\text{сп }1}$, $x_{\text{сп }2}$, ..., $x_{\text{сп }n}$, на основі яких визначено поточкові розподіли абсолютних частот та відповідних статистичних ймовірностей, подані в таблицях 2.8.1. та 2.8.2.

Табл. 2.8.1

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

Табл. 2.8.2

x_i	x_1	x_2	...	x_k
$P_n^*(\{x_i\})$	$P_n^*(\{x_1\})$	$P_n^*(\{x_2\})$	...	$P_n^*(\{x_k\})$

Точку, абсциса якої дорівнює середньому арифметичному спостережених значень:

$$m_n^* = \frac{1}{n}\sum_{i=1}^{n} x_{\text{сп}\,i} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \sum_{i=1}^{k} x_i P_n^*(\{x_i\}),$$

називають *центром розподілу* (чи розсіювання) поточкового розподілу статистичних ймовірностей на скінченній множині $\Omega = \{x_1, x_2, ..., x_k\}$. Очевидно, центр розсіювання буде знаходитися якомога ближче до точок, на які припадає основна маса статистичних ймовірностей. Часто m_n^* позначають також через $\bar{x}$.

Важливою характеристикою розподілу статистичних ймовірностей, окрім центра розподілу, є також величина, за якою характеризують розсіювання (чи скупченість) статистичних ймовірностей навколо центра розподілу. До таких характеристик відносяться *дисперсія* розподілу статистичних ймовірностей, а також *середнє квадратичне відхилення.*

Дисперсією поточкового розподілу статистичних ймовірностей на скінченній множині $\Omega = \{x_1, x_2, ..., x_k\}$ називають число

$$D_n^* = \frac{1}{n}\sum_{i=1}^{n} (x_{\text{сп}\,i} - m_n^*)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - m_n^*)^2 =$$
$$= \sum_{i=1}^{k} (x_i - m_n^*)^2 P_n^*(\{x_i\}).$$

Середнім квадратичним відхиленням називають число $\sigma_n^* = \sqrt{D_n^*}$.

Приклад 2.8.2. Нехай є два розподіли:

c)

x_i	-1	0	1
$P_n^*(\{x_i\})$	0.1	0.8	0.1

,

d)

$x_i^/$	-10	0	10
$P_n^*(\{x_i^/\})$	0.1	0.8	0.1

.

Очевидно за кожного з цих розподілів $m_n^* = 0$, але в розподілі *d*) розсіювання статистичних ймовірностей навколо центра розподілу $m_n^* = 0$ помітно більше, ніж в розподілі *c*).

Розглянуті характеристики досить важливі для опрацювання результатів спостережень. Чим менше розсіювання (дисперсія), тим точнішим, вірогіднішим і надійнішим є усереднений результат спостережень (m_n^*) за достатньо великої кількості спостережень.

В фізичній інтерпретації центр розподілу статистичних ймовірностей є центром мас одиничної маси, розподіленої на множині точок $\{x_1, x_2, \ldots, x_k\}$ так, що на точку x_i припадає маса $P_n^*(\{x_i\})$, а дисперсія – це момент інерції цієї одиничної маси відносно центра мас.

Приклад 2.8.3. Якщо розподіл статистичних ймовірностей визначається за таблицями 2.8.3 та 2.8.4,

Табл. 2.8.3

x_i	1	2	3	4	5	6
n_i	0	1	3	2	4	5

Табл. 2.8.4

x_i	1	2	3	4	5	6
$P_n^*(\{x_i\})$	0	$\frac{1}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{5}{15}$

то
$$m_n^* = 1\cdot 0 + 2\cdot\frac{1}{15} + 3\cdot\frac{3}{15} + 4\cdot\frac{2}{15} + 5\cdot\frac{4}{15} + 6\cdot\frac{5}{15} =$$
$$= \frac{1}{15}(\,1\cdot 0 + 2\cdot 1 + 3\cdot 3 + 4\cdot 2 + 5\cdot 4 + 6\cdot 5) = \frac{69}{15} = 4\frac{3}{5},$$
$$D_n^* = \frac{1}{15}\left[\left(1-4\frac{3}{5}\right)^2\cdot 0 + \left(2-4\frac{3}{5}\right)^2\cdot 1 + \left(3-4\frac{3}{5}\right)^2\cdot 3 + \right.$$
$$\left. + \left(4-4\frac{3}{5}\right)^2\cdot 2 + \left(5-4\frac{3}{5}\right)^2\cdot 4 + \left(6-4\frac{3}{5}\right)^2\cdot 5\right] \approx 1.7,$$
$$\sigma_n^* = \sqrt{D_n^*} \approx 1.3.$$

Вправи для самостійного виконання

2.8.1. Перевірити, чи правильні твердження:

1. В безкоординатному просторі числові характеристики розподілу статистичних ймовірностей можна визначити через щільність їх розподілу.
2. За кожного дискретного розподілу статистичних ймовірностей існує центр розсіювання.
3. Всі спостережені значення можуть бути натуральними числами.
4. Координата центра розсіювання статистичних ймовірностей є натуральним числом, коли всі спостережені значення – натуральні числа.
5. В разі обчислення дисперсії треба спочатку обчислити

координату центра розсіювання статистичних ймовірностей.

6. В разі обчислення дисперсії треба спочатку обчислити середнє квадратичне відхилення.

2.8.2. 1. Знайти центр розподілу, дисперсію та середнє квадратичне відхилення поточкового розподілу статистичних ймовірностей $P_n^*(\{x_i\})$:

а)

x_i	-3	-2	-1	0	1	2	3
$P_n^*(\{x_i\})$	0.10	0.20	0.40	0.10	0.10	0.07	0.03

б)

x_i	-9	-8	-7	-6	-5	-4	-3	-2	-1
$P_n^*(\{x_i\})$	0.02	0.10	0.70	0.08	0.04	0.03	0.01	0.01	0.01

в)

x_i	-4	-3	-2	-1	0	1	2	3	4	5
$P_n^*(\{x_i\})$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

г)

x_i	-3	-2	-1	0	1	2	3
$P_n^*(\{x_i\})$	0.30	0.10	0.08	0.04	0.08	0.10	0.30

д) заданого через функцію розподілу статистичних ймовірностей:

$$F_n^*(x) = \begin{cases} 0, & \text{коли } x \le 0, \\ 0.25, & \text{коли } 0 < x \le 1, \\ 0.75, & \text{коли } 1 < x \le 2, \\ 1, & \text{коли } x > 2. \end{cases}$$

2.9. Деякі числові характеристики поінтервального розподілу статистичних ймовірностей в одновимірному координатному просторі

Центром розподілу (чи розсіювання) узагальнених статистичних ймовірностей на множині $\Omega=[a;b)=\bigcup_{i=1}^{k}[a_{i-1};a_i)$, $[a_{i-1};a_i)\bigcap[a_{j-1};a_j)=\varnothing$, коли $i\neq j$, $a_i-a_{i-1}=h=\dfrac{a_k-a_0}{k}$, $i\in\overline{1,\,k}$, $S=\{A\,|\;A=\bigcup_{i\in I}[a_{i-1};a_i),\; I\subset\{1,2,\dots,k\}\}$ в разі поінтервального розподілу статистичних ймовірностей із щільністю розподілу $f(x)$, $x\in[a;b)$, ($f(x)=0$, коли $x\overline{\in}[a;b)$), називають точку на осі *Ox,* абсциса якої дорівнює

$$\bar{x}=\int_a^b x\,f(x)\,dx=\sum_{i=1}^{k}\int_{a_{i-1}}^{a_i}x\,f(x)\,dx=\sum_{i=1}^{k}\int_{a_{i-1}}^{a_i}x\,\frac{P_n^*([a_{i-1};a_i))}{h}dx=$$

$$=\sum_{i=1}^{k}\bar{x}_iP_n^*([\,a_{i-1};a_i))=\sum_{i=1}^{k}\frac{1}{n}\bar{x}_ik_n([a_{i-1};a_i))\approx$$

$$\approx\frac{1}{n}\sum_{i=1}^{k}\sum_{x_{cn\,j}\in[a_{i-1};\,a_i)}x_{cn\,j}=\frac{1}{n}\sum_{i=1}^{n}x_{cn\,i}=m_n^*,$$

де $h=a_i-a_{i-1}$, $\bar{x}_i=\displaystyle\int_{a_{i-1}}^{a_i}x\frac{1}{h}dx=\frac{a_i^2-a_{i-1}^2}{2h}=\frac{a_{i-1}+a_i}{2}$, і це число $\bar{x}_i$ наближено дорівнює середньому арифметичному спостережених значень $x_{cn\,j}$, що потрапили в проміжок $[\,a_{i-1};a_i)$, $k_n([a_{i-1};a_i))$ – кількість спостережених значень $x_{cn\,j}$, що потрапили в проміжок $[\,a_{i-1};a_i)$

Дисперсією поінтервального розподілу узагальнених статистичних ймовірностей на множині $\Omega=[a;\,b)=\bigcup_{i=1}^{k}[a_{i-1};\,a_i)$ із щільністю $f(x)$, $x\in[\,a;b)$ ($f(x)=0$, коли $x\overline{\in}[\,a;b)$), називають число

$$D_n^*=\int_a^b(\,x-m_n^*)^2\,f(x)\,dx,$$

а число $\sigma_n^*=\sqrt{D_n^*}$ – називають *середнім квадратичним відхиленням* відповідного поінтервального розподілу статистичних ймовірностей.

В фізичній інтерпретації центр розсіювання є центром мас одиничної маси, розподіленої на проміжку $[a;b)$ із щільністю $f(x)$, а дисперсія – момент інерції цієї маси відносно центра мас.

Приклад 2.9.1. В разі розподілу узагальнених статистичних ймовірностей, що визначається за Табл. 2.9.1:

Табл. 2.9.1

$[a_{i-1};a_i)$	$\left[0;\frac{1}{10}\right)$	$\left[\frac{1}{10};\frac{2}{10}\right)$	$\left[\frac{2}{10};\frac{3}{10}\right)$	$\left[\frac{3}{10};\frac{4}{10}\right)$	$\left[\frac{4}{10};\frac{5}{10}\right)$
$P_{100}^*([a_{i-1};a_i))$	0.5	0.2	0.1	0.1	0.04

	$\left[\frac{5}{10};\frac{6}{10}\right)$	$\left[\frac{6}{10};\frac{7}{10}\right)$	$\left[\frac{7}{10};\frac{8}{10}\right)$	$\left[\frac{8}{10};\frac{9}{10}\right)$	$\left[\frac{9}{10};1\right)$
	0	0.02	0.01	0.02	0.01

дістаємо:

$$m_n^* = 0.05\cdot 0.5 + 0.15\cdot 0.2 + 0.25\cdot 0.1 + 0.35\cdot 0.1 + 0.45\cdot 0.04 +$$
$$+0.55\cdot 0 + 0.65\cdot 0.02 + 0.75\cdot 0.01 + 0.85\cdot 0.02 + 0.95\cdot 0.01 \approx 0.18,$$

Обчислюючи дисперсію D_n^*, одержимо:

$$D_n^* = \int_a^b (x-m_n^*)^2 f_n^*(x)dx = \sum_{i=1}^{k}\int_{a_{i-1}}^{a_i}(x-m_n^*)^2 f_n^*(x)dx =$$
$$= \sum_{i=1}^{k}\int_{a_{i-1}}^{a_i}(x-m_n^*)^2 \frac{P_n^*([a_{i-1},a_i))}{h}dx = \sum_{i=1}^{k}\frac{P_n^*([a_{i-1},a_i))}{h}\int_{a_{i-1}}^{a_i}(x-m_n^*)^2 dx.$$

Якщо довжини проміжків $[a_{i-1},a_i)$ досить малі, то значення інтеграла $\int_{a_{i-1}}^{a_i}(x-m_n^*)^2 dx$ наближено дорівнює значенню підінтегральної функції в середній точці проміжка $[a_{i-1};a_i)$, тобто в точці $x=\frac{a_{i-1}+a_i}{2}$, помноженому на довжину проміжка $h=a_i-a_{i-1}$, тобто $\int_{a_{i-1}}^{a_i}(x-m_n^*)^2 dx \approx \left(\frac{a_{i-1}+a_i}{2}-m_n^*\right)^2\cdot h$, звідки

$$D_n^* \approx \sum_{i=1}^{k} P_n^*([a_{i-1};a_i))\left(\frac{a_{i-1}+a_i}{2}-m_n^*\right)^2.$$

Стосовно розподілу, заданого за таблицею 2.9.1, одержимо
$D_n^* \approx (0.05-0.18)^2 \cdot 0.50 + (0.15-0.18)^2 \cdot 0.20 + (0.25-0.18)^2 \cdot 0.10 +$
$+(0.35-0.18)^2 \cdot 0.10 + (0.45-0.18)^2 \cdot 0.04 + (0.55-0.18)^2 \cdot 0.00 +$
$+(0.65-0.18)^2 \cdot 0.02 + (0.75-0.18)^2 \cdot 0.01 + (0.85-0.18)^2 \cdot 0.02 +$
$+(0.95-0.18)^2 \cdot 0.01 \approx 0.0375$, $\sigma_n^* = \sqrt{D_n^*} \approx \sqrt{0.0375} \approx 0.194$.

Вправи для самостійного виконання

2.9.1. Перевірити, чи правильні твердження:

1. В безкоординатному просторі координату центра розподілу статистичних ймовірностей можна визначити, поділивши множину Ω на підмножини H_i, $i \in \overline{1,n}$.

2. В координатному просторі за кожного поінтервального розподілу узагальнених статистичних ймовірностей існує центр розсіювання.

3. Якщо $x_{\text{сп}\,j}$, $j \in \overline{1,\ n}$, – спостережені значення, а $\bar{x}$ – координата центра розподілу статистичних ймовірностей, то

$$\min_{1\le j\le n} x_{\text{сп}\ j} \le \bar{x} \le \max_{1\le j\le n} x_{\text{сп}\ j}.$$

4. Координата центра розподілу статистичних ймовірностей завжди є раціональним числом.

5. Для обчислення дисперсії розподілу статистичних ймовірностей спочатку треба обчислити координату центра розподілу статистичних ймовірностей.

6. Для обчислення дисперсії розподілу статистичних ймовірностей спочатку треба обчислити середнє квадратичне відхилення.

2.9.2. Знайти центр розподілу та дисперсію поінтервального розподілу узагальнених статистичних ймовірностей:

а) заданого за таблицею:

$[a_{i-1}; a_i)$	[0; 1)	[1; 2)	[2; 3)	[3; 4)	[4; 5)	[5; 6)
$P_n^*([a_{i-1}; a_i))$	0.10	0.70	0.10	0.06	0.03	0.01

б) заданого через щільність розподілу:

$$f(x) = \begin{cases} 0.20, \text{ коли } x \in [0;\ 1), \\ 0.50, \text{ коли } x \in [1;\ 2), \\ 0.20, \text{ коли } x \in [2;\ 3), \\ 0.10, \text{ коли } x \in [3;\ 4), \\ \ \ 0, \ \text{ коли } x \overline{\in} [0;\ 4). \end{cases}$$

в) заданого через функцію розподілу:

$$F(x)=\begin{cases} 0, & \text{коли } x \le -1, \\ 0.2(x+1), & \text{коли } -1 \le x \le 0, \\ 0.2+0.6x, & \text{коли } 0 \le x \le 1, \\ 0.8+0.2(x-1), & \text{коли } 1 \le x \le 2, \\ 1, & \text{коли } x \ge 2. \end{cases}$$

2.10. Повторні випробування

Нехай проводиться n випробувань, в кожному з яких навмання вибирається точка $M(x_1, x_2, ..., x_m) \in R^m$ із m-вимірної множини точок $\tilde{\Omega} = \Omega^m = \Omega \times \Omega \times ... \times \Omega \subset R^m$. Нехай координати $x_1, x_2, ..., x_m$ точки M вибираються одна за однією слідуючим чином: спочатку навмання вибирається координата $x_1 \in \Omega \subset R^1$, потім вибирається навмання друга координата $x_2 \in \Omega \subset R^1$, далі навмання вибирається третя координата $x_3 \in \Omega \subset R^1$, і т.д., нарешті навмання вибирається координата $x_m \in \Omega \subset R^1$. Якщо вибір кожної координати $x_i \in \Omega$, $i \in \overline{1,m}$, розглядати як окреме випробування, тоді вибір n точок $M_k \in \Omega^m$, $k \in \overline{1,n}$, можна розглядати як n серій із m випробувань, в кожному з яких навмання вибирається координата $x_i \in \Omega$ точки $M_k \in \Omega^m$, $i \in \overline{1,m}$, $k \in \overline{1,n}$.

За результатами вказаних n випробувань можна визначити статистичні ймовірності $P_n^*(A^{(i)})$ попадання координати x_i в деяку підмножину $A^{(i)}$ множини Ω стосовно кожної координати x_i окремо, $i \in \overline{1,m}$.

Множина точок Ω^m є множиною можливих наслідків кожного із вказаних n випробувань, в кожному з яких вибирається точка $M(x_1, x_2, ..., x_m) \in \Omega^m$ (тобто Ω^m є множиною елементарних подій $\tilde{E} \in \Omega^m$). Кожна із таких елементарних подій $\tilde{E} \in \Omega^m$ характеризується впорядкованим набором $(E_1, E_2, ..., E_m)$ елементарних подій $E_i \in \Omega$, що відбуваються з вибиранням відповідних координат x_i, $i \in \overline{1,m}$. Через сукупність Ω^m таких наборів $(E_1, E_2, ..., E_m)$, де кожна змінна E_i, $i \in \overline{1,m}$, пробігає множину Ω, коли фіксовані всі можливі набори значень інших змінних $E_j \in \Omega$, $j \in \overline{1,m} \setminus \{i\}$, утворюється так званий *декартовий добуток m разів множини Ω саму на себе*, який позначається $\Omega \times \Omega \times ... \times \Omega$, або декартовий m-й степінь множини Ω, який позначається через Ω^m.

Наприклад, $\Omega \times \Omega = \{(E_1, E_2) \mid E_1 \in \Omega, E_2 \in \Omega\}$ (Рис. 2.10.1), $\Omega \times \Omega \times \Omega = \{(E_1, E_2, E_3) \mid E_1 \in \Omega, E_2 \in \Omega, E_3 \in \Omega\}$ (Рис. 2.10.2).

Декартів квадрат Ω^2 є частинним випадком декартового добутку $A \times B$ довільних непорожніх множин A і B, який визначається за рівністю $A \times B = \{(x, y) \mid x \in A,\ y \in B\}$.

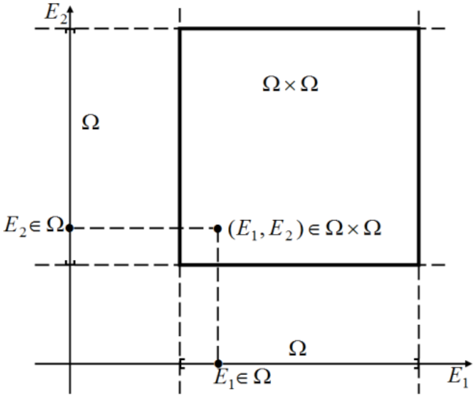


Рис. 2.10.1

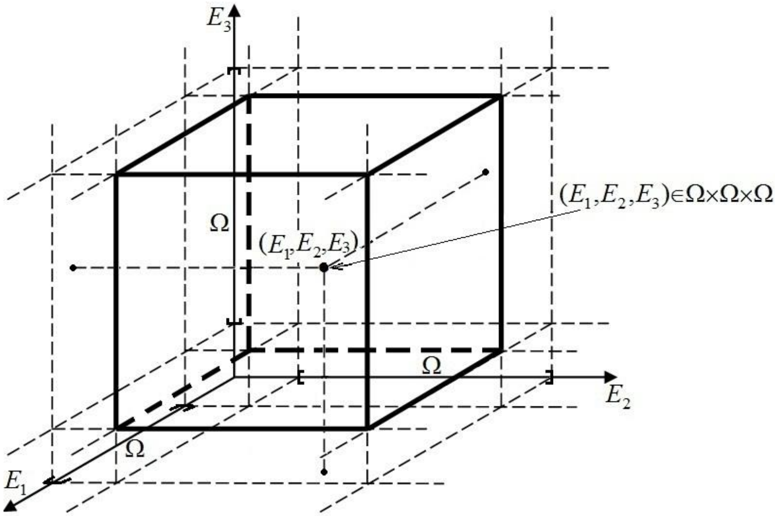


Рис. 2.10.2

Поняття декартового добутку поширюється на випадок довільної кількості множин $A_1, A_2, \ldots, A_m$: $A_1 \times A_2 \times \ldots \times A_m =$ $= \{(x_1, x_2, \ldots, x_m) \mid x_k \in A_k,\ k \in \overline{1, m}\}$.

Зауважимо, що сукупність всіх m-розрядних двійкових кодів можна розглядати як декартовий m-й степінь множини $\{0, 1\}$. Тоді, очевидно, однорозрядних двійкових кодів буде два – 0 і 1, двохрозрядних буде чотири – спочатку до кожного однорозрядного двійкового коду зліва дописуємо 0 і отримаємо:

00, 01, потім до кожного однорозрядного двійкового коду зліва дописуємо 1 і отримаємо – 10, 11, трирозрядних двійкових кодів буде вісім – спочатку до всіх чотирьох двохрозрядних двійкових кодів зліва дописуємо 0 і отримаємо – 000, 001, 010, 011, потім до всіх чотирьох двохрозрядних двійкових кодів зліва дописуємо 1 і отримаємо – 100, 101, 110, 111, аналогічно одержимо – чотирирозрядних двійкових кодів буде 16, п'ятирозрядних – 32, і т.д., m-розрядних двійкових кодів буде 2^m.

Аналогічно, якщо розглядати декартовий m-й степінь множини $\{A, \overline{A}\}$, то всіх m-розрядних наборів літер A і $\overline{A}$ буде 2^m.

Крім того декартовий добуток $\Omega\times\Omega$ можна розглядати і як перетин множин

$$\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in W\} \text{ і } \{(E_1, E_2) \mid E_1 \in W,\ E_2 \in\Omega\},$$

де W довільна множина така, що $\Omega\subset W$, тобто

$$\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in W\} \bigcap \{(E_1, E_2) \mid E_1 \in W,\ E_2 \in\Omega\},$$

а тому

$$\Omega\times\Omega=\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in\Omega\}=$$

$$=\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in W\} \bigcap \{(E_1, E_2) \mid E_1 \in W,\ E_2 \in\Omega\},$$

$$\Omega\times\Omega\times\Omega=\{(E_1, E_2, E_3) \mid E_1 \in\Omega,\ E_2 \in W,\ E_3 \in W\} \bigcap$$

$$\bigcap \{(E_1, E_2, E_3) \mid E_1 \in W,\ E_2 \in\Omega,\ E_3 \in W\} \bigcap$$

$$\bigcap \{(E_1, E_2, E_3) \mid E_1 \in W,\ E_2 \in W,\ E_3 \in\Omega\}.$$

Аналогічно

$$\Omega^m=\{(E_1, E_2, \ldots, E_m) \mid E_1 \in\Omega,\ E_2 \in W,\ \ldots,\ E_m \in W\} \bigcap$$

$$\bigcap \{(E_1, E_2, \ldots, E_m) \mid E_1 \in W,\ E_2 \in\Omega,\ E_3 \in W, \ldots, E_m \in W\} \bigcap \ldots \bigcap$$

$$\bigcap \{(E_1, E_2, \ldots, E_m) \mid E_1 \in W,\ E_2 \in W, \ldots, E_{m-1} \in W,\ E_m \in\Omega\},\ \Omega\subset W.$$

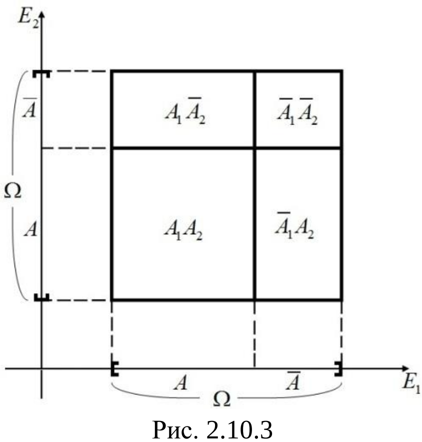


Рис. 2.10.3

Якщо, наприклад, розглядається декартовий добуток множин $\Omega\times\Omega=\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in\Omega\}=\{(E_1, E_2) \mid E_1 \in\Omega,\ E_2 \in W\} \bigcap$

$\bigcap\{(E_1, E_2) \mid E_1 \in W,\ E_2 \in \Omega\}$, $\Omega \subset W$, і $\Omega = A + \overline{A}$, та через A_1 позначити множину точок із $\Omega\times\Omega$, в яких координата $E_1 \in A$, через A_2 – множину точок із $\Omega\times\Omega$, в яких координата $E_2 \in A$, то одержимо (Рис. 2.10.3):

$$A_1 = \{(E_1, E_2) \mid E_1 \in A,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\},$$
$$A_2 = \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in A\},$$
$$\overline{A_1} = \{(E_1, E_2) \mid E_1 \in \overline{A},\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\},$$
$$\overline{A_2} = \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \overline{A}\},$$
$$A_1A_2 = \{(E_1, E_2) \mid E_1 \in A,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\}$$
$$\bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in A\} =$$
$$= \{(E_1, E_2) \mid E_1 \in A,\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in A\},$$
$$\{(E_1, E_2) \mid E_1 \in A,\ E_2 \in \Omega\} \subset \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\},$$
$$\{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in A\} \subset \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \Omega\}.$$

Аналогічно

$$A_1\overline{A_2} = \{(E_1, E_2) \mid E_1 \in A,\ E_2 \in \Omega\} \bigcap$$
$$\bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \overline{A}\},$$
$$\overline{A_1}A_2 = \{(E_1, E_2) \mid E_1 \in \overline{A},\ E_2 \in \Omega\} \bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in A\},$$
$$\overline{A_1}\,\overline{A_2} = \{(E_1, E_2) \mid E_1 \in \overline{A},\ E_2 \in \Omega\} \bigcap$$
$$\bigcap \{(E_1, E_2) \mid E_1 \in \Omega,\ E_2 \in \overline{A}\}.$$
$$A_1 = A_1A_2 + A_1\overline{A_2} \subset \Omega\times\Omega,$$
$$A_2 = A_1A_2 + \overline{A_1}A_2 \subset \Omega\times\Omega \text{ (див. Рис. 2.10.3).}$$

Нехай

$$A_i = \{(E_1, E_2, \ldots, E_m) \mid E_i \in A^{(i)},\ A^{(i)} \subset \Omega,$$
$$E_j \in \Omega,\ j \in \overline{1, m} \setminus \{i\}\},\ i \in \overline{1, m}\ \},$$

тобто координата E_i належить до множини $A^{(i)} \subset \Omega$, а інші координати E_j, $j \in \overline{1, m} \setminus \{i\}$ належать до Ω.

Позначимо через $I_1 \subset \{1, 2, ..., m\}$, $I_2 \subset \{1, 2, ..., m\}$ підмножини індексів такі, що $I_1 \bigcap I_2 = \varnothing$, $I_1 \bigcup I_2 = \{1, 2, ..., m\}$, де множина I_1, як і I_2, може бути порожньою, а також будь якою іншою підмножиною множини $\{1, 2, \ldots, m\}$. Очевидно, $I_2 = \{1, 2, ..., m\} \setminus I_1 = \{1, 2, ..., m\}$, коли $I_1 = \varnothing$, $I_2 = \varnothing$, коли $I_1 = \{1, 2, ..., m\}$.

Розглянемо сукупність $\tilde{S}$ підмножин множини $\tilde{\Omega} = \Omega^m$ (простір подій $\tilde{A} \subset \tilde{\Omega} = \Omega^m$, $\tilde{A} \in \tilde{S}$), породжену за системою паралелепіпедів виду:

$$\prod_{i\in I_1}\overline{A}_i\prod_{j\in I_2}A_j\,.$$

де $A_i \in \widetilde{S}$ означає подію, яка полягає в тому, що в i-му випробуванні відбувається подія $A^{(i)} \in S^{(i)} = \{\varnothing, A^{(i)}, \overline{A}^{(i)}, \Omega\}$. Нехай за результатами вказаної серії із n випробувань одержали

$$P_n^*(A_i),\ P_n^*\left(\prod_{i\in I_1}\overline{A}_i\prod_{j\in I_2}A_j\right),$$

і нехай стосовно добутку будь якої кількості m_1 подій $\overline{A}_i$, $i\in I_1$, і m_2 подій A_j, $j\in I_2$, де $m_1 = m(I_1)$, $m_2 = m(I_2)$ – кількості елементів в множинах $I_1 \subset \{1,2,...,m\}$, $I_2 \subset \{1,2,...,m\}$, $I_1 \bigcup I_2 \neq \varnothing$, виявилось, що практично з досить великою точністю мають місце рівності

$$P_n^*\left(\left(\bigcap_{i\in I_1}\overline{A}_i\right)\cap\left(\bigcap_{j\in I_2}A_j\right)\right) = \prod_{i\in I_1}P_n^*(\overline{A}_i)\prod_{j\in I_2}P_n^*(A_j) \qquad (2.10.1)$$

На практиці «швидше за все» так і буде. Очевидно, окремий наслідок $\widetilde{E} = (E_1, E_2, E_3, ..., E_m) \in \widetilde{\Omega}$ кожного із n випробувань, в якому вся вказана серія із m випробувань розглядається як одне випробування – вибір точки $\widetilde{E}$ із $\widetilde{\Omega}$, належить до однієї із множин (див. Рис. 2.10.3):

$$\begin{aligned}
H_0 &= \overline{A}_1\overline{A}_2\overline{A}_3\ldots\overline{A}_{m-2}\overline{A}_{m-1}\overline{A}_m,\\
H_1 &= \overline{A}_1\overline{A}_2\overline{A}_3\ldots\overline{A}_{m-2}\overline{A}_{m-1}A_m,\\
H_2 &= \overline{A}_1\overline{A}_2\overline{A}_3\ldots\overline{A}_{m-2}A_{m-1}\overline{A}_m,\\
H_3 &= \overline{A}_1\overline{A}_2\overline{A}_3\ldots\overline{A}_{m-2}A_{m-1}A_m,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
H_{k-1} &= A_1A_2A_3\ldots A_{m-2}A_{m-1}\overline{A}_m,\\
H_k &= A_1A_2A_3\ldots A_{m-2}A_{m-1}A_m.
\end{aligned} \qquad (2.10.2)$$

Легко бачити, що $H_iH_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=0}^{k}H_i = \Omega^m$. Тут $k = 2^m - 1$, а в номері $i \in \overline{0,k}$ множини H_i у його поданні у вигляді m-розрядного двійкового числа (коду) містяться нулі на тих місцях серед місць від першого до m-го, на яких в добутку H_i стоїть літера $\overline{A}$ з рискою, і одиниці на тих місцях, де стоїть літера A без риски.

Ці номери i, $i \in \overline{0,k}$, подані у вигляді m-розрядних двійкових чисел, матимуть вигляд

000…000
000…001
000…010
000…011
…………
111…110
111…111

Приклад 2.10.1. Нехай за допомогою датчика випадкових чисел навмання вибиралися точки (i, j), $i \in \overline{1,10}$, $j \in \overline{1,10}$, так, що спочатку навмання вибирали число i (одне із чисел 1, 2, …, 10), а потім навмання вибирали число j (одне із десяти чисел 1, 2, …, 10).

Як простір подій $\tilde{S}$ розглядатимемо найширшу сукупність підмножин множини $\tilde{\Omega} = \Omega^2 = \{(i, j) \mid i \in \overline{1, 10}, j \in \overline{1, 10}\}$.

Нехай $A_1 \in \tilde{S}$ – множина тих точок із $\tilde{\Omega} = \Omega^2$, перша координата яких належить до множини $\{4, 5, 6, 7\}$, $A_2 \in \tilde{S}$ – множина тих точок із $\tilde{\Omega} = \Omega^2$, друга координата яких належить до множини $\{2, 3, 4, 5, 6\}$ (Рис. 2.10.4).

Нехай кожна із 100 точок (i, j), $i \in \overline{1,10}$, $j \in \overline{1,10}$, вибиралася із статистичною ймовірністю $P_n^*(\{(i, j)\}) = \frac{1}{100}$. Тоді очевидно (див. Рис. 2.10.4)

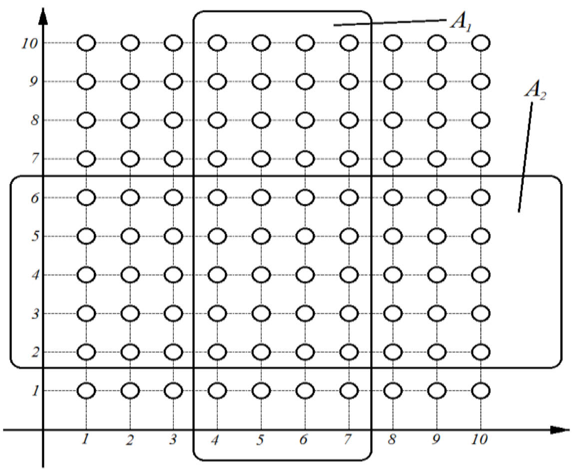


Рис. 2.10.4

$$P_n^*(A_1)=\frac{40}{100},\ P_n^*(A_2)=\frac{50}{100},\ P_n^*(\overline{A}_1)=\frac{60}{100},\ P_n^*(\overline{A}_2)=\frac{50}{100},$$

$$P_n^*(A_1A_2)=\frac{20}{100}=\frac{40}{100}\cdot\frac{50}{100}=P_n^*(A_1)P_n^*(A_2),$$

$$P_n^*(\overline{A}_1A_2)=\frac{30}{100}=\frac{60}{100}\cdot\frac{50}{100}=P_n^*(\overline{A}_1)P_n^*(A_2),$$

$$P_n^*(A_1\overline{A}_2)=\frac{20}{100}=\frac{40}{100}\cdot\frac{50}{100}=P_n^*(A_1)P_n^*(\overline{A}_2),$$

$$P_n^*(\overline{A}_1\overline{A}_2)=\frac{30}{100}=\frac{60}{100}\cdot\frac{50}{100}=P_n^*(\overline{A}_1)P_n^*(\overline{A}_2).$$

Таким чином в розглядуваному випадку події $\overline{A}_1$ та $\overline{A}_2$, $\overline{A}_1$ та A_2, A_1 та $\overline{A}_2$, A_1 та A_2 виявляються незалежними відносно ймовірнісної міри P_n^*.

Зауважимо, що події A_1, A_2 можуть виявитися незалежними і у випадку, коли статистичні ймовірності елементарних подій не обов'язково рівні між собою.

Приклад 2.10.2. Нехай, як і у попередньому прикладі, за допомогою датчика випадкових чисел навмання вибиралися пари чисел (i, j), $i\in\overline{1,3}$, $j\in\overline{1,3}$, тобто $\tilde{\Omega}=\Omega^2=\{(i, j)\,|\,i\in\overline{1,3},\ j\in\overline{1,3}\}$. Як сукупність $\tilde{S}$ подій розглядатимемо найширшу сукупність підмножин множини $\tilde{\Omega}=\Omega^2$.

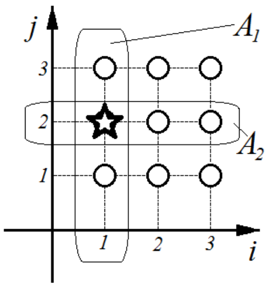


Рис. 2.10.5

Розглянемо події (Рис. 2.10.5):

$$A_1=\{(i, j)\,|\,i=1,\ j\in\overline{1,\ 3}\},\ A_2=\{(i, j)\,|\,i\in\overline{1,\ 3},\ j=2\}.$$

Нехай проведено лише одне випробування, в якому вибраною виявилась точка (1, 2) (див. Рис. 2.10.5).

Тоді $P_n^*(A_1)=1$, $P_n^*(A_2)=1$, $P_n^*(A_1A_2)=1$,

$$P_n^*(A_1A_2)=P_n^*(A_1)P_n^*(A_2),\ n=1.$$

Тобто події A_1 і A_2 виявляються незалежними відносно так визначеної міри P_n^*.

Приклад 2.10.3. Нехай за допомогою датчика випадкових чисел в 3-вимірному просторі вибираються точки (i,j,k), $i\in\overline{1,10}$, $j\in\overline{1,10}$, $k\in\overline{1,10}$. Нехай $\tilde{\Omega}=\{(i,j,k)\mid i\in\overline{1,10},\ j\in\overline{1,10},\ k\in\overline{1,10}\}$, $\tilde{S}$ – найширша сукупність підмножин множини $\tilde{\Omega}$. Нехай в результаті досить великої серії із n випробувань виявилось, що в кожну із точок (i,j,k), $i\in\overline{1,10}$, $j\in\overline{1,10}$, $k\in\overline{1,10}$, попадали практично однаково часто (статистичні ймовірності всіх елементарних подій виявились практично однаковими).

Нехай події $A_1=\{(i,j,k)\mid i\in\{3,5,7\}, j\in\overline{1,10}, k\in\overline{1,10}\}$,

$$A_2=\{(i,j,k)\mid i\in\overline{1,10}, j\in\{2,3,4,5\}, k\in\overline{1,10}\},$$

$$A_3=\{(i,j,k)\mid i\in\overline{1,10}, j\in\overline{1,10}, k\in\{1,2,3,5,7\}\}.$$

Тоді легко бачити, що відносно так визначеної ймовірнісної міри P_n^* події A_1, A_2, A_3 виявляються статистично незалежними попарно і в сукупності, тобто

$$P_n^*(A_1A_2)=P_n^*(A_1)P_n^*(A_2),\ P_n^*(A_1A_3)=P_n^*(A_1)P_n^*(A_3),$$

$$P_n^*(A_2A_3)=P_n^*(A_2)P_n^*(A_3),\ P_n^*(A_1A_2A_3)=P_n^*(A_1)P_n^*(A_2)P_n^*(A_3).$$

Приклад 2.10.4. Спортивна вправа полягає в тому, що кожен учень із команди із 10 учнів один за одним в наперед визначеній черговості кидає один раз м'яч в баскетбольну корзину. Для визначення вправності кожного учня вправу на тренуваннях виконували 100 разів (тобто вся команда 100 разів виконувала вправу – кожен із 10 членів команди один за одним у визначеній черговості виконував один кидок, і так повторювали 100 разів).

В результаті з'ясувалось, що перший учень в ста спробах влучив 80 разів, другий – також 80 разів, третій – 90 разів, четвертий – 75 разів, п'ятий – 98 разів, шостий – 95 разів, сьомий – 100 разів, восьмий – 60 разів, дев'ятий – 88 разів, десятий – 50 разів.

Якщо через A_i, $i\in\overline{1,10}$, позначити подію – влучення м'ячем в корзину в результаті i-го кидка в кожній серії із 10 кидків і довільних наслідках в інших 9-ти кидках, тоді одержимо $P_{100}^*(A_1)=0.80$, $P_{100}^*(A_2)=0.80$, $P_{100}^*(A_3)=0.90$, $P_{100}^*(A_4)=0.75$, $P_{100}^*(A_5)=0.98$, $P_{100}^*(A_6)=0.95$, $P_{100}^*(A_7)=1.00$, $P_{100}^*(A_8)=0.60$, $P_{100}^*(A_9)=0.88$, $P_{100}^*(A_{10})=0.50$.

Якщо через $\tilde{E}=(E_1,E_2,E_3,...,E_9,E_{10})$ позначити окремий можливий наслідок виконання вправи всіма 10-тьма учнями, тоді кожен такий наслідок належатиме до однієї із $2^{10} = 1024$ множин H_0, H_1, H_2, H_3, …, H_{1022}, H_{1023} (див. формули 2.10.2).

Вправність кожного учня можна визначити і в інший спосіб – спочатку всі 100 кидків виконує 1-й учень, потім 100 кидків виконує 2-й учень, і т.д., нарешті всі 100 кидків виконує 10-й учень. Або ж всі 10 учнів виконують кожен 100 кидків в різні 10 корзин і можливо в різних спортивних залах і в різний час. Важливим є лише визначення статистичних ймовірностей $P_n^*(A_i)$, $i \in \overline{1,10}$, під час виконання вправи десятьма учнями один за іншим в наперед визначеній черговості.

Зауважимо, що якщо кожного із 100 разів всі 10 кидків виконує один і той самий учень, тоді не виключено, що всі $P_{100}^*(A_i)$, $i \in \overline{1,10}$, можуть виявитися з достатньою точністю практично рівними між собою, що не виключено і тоді, коли кидки виконують різні 10 учнів.

Зрозуміло, що коли $B = \bigcup_{i \in I} H_i$, $I \subset \{0,1,2,...,k\}$, то $P_n^*(B) = \sum_{i \in I} P_n^*(H_i)$. Якщо події $A_1, A_2, \ldots, A_m$ виявляються практично незалежними в сукупності відносно ймовірнісної міри P_n^*, визначеної за результатами вказаних n випробувань (статистично незалежними), а кожна з подій H_i є добутком m подій A_i та $\overline{A}_j$, до якого множник A_i входить s разів, тоді B можна подати у вигляді

$$B = A_1 A_2 ... A_s \overline{A}_{s+1} ... \overline{A}_m + \overline{A}_1 A_2 ... A_{s+1} \overline{A}_{s+2} ... \overline{A}_m + ... + \overline{A}_1 \overline{A}_2 ... \overline{A}_{m-s} A_{m-s+1} ... A_m,$$

і таким чином з досить великою точністю практично має місце рівність

$$P_n^*(B) = P_n^*(A_1 A_2 ... A_s \overline{A}_{s+1} ... \overline{A}_m) + P_n^*(\overline{A}_1 A_2 ... A_{s+1} \overline{A}_{s+2} ... \overline{A}_m) + ... +$$
$$+ P_n^*(\overline{A}_1 \overline{A}_2 ... \overline{A}_{m-s} A_{m-s+1} ... A_m) =$$
$$= P_n^*(A_1) P_n^*(A_2) ... P_n^*(A_s) P_n^*(\overline{A}_{s+1}) ... P_n^*(\overline{A}_m) +$$
$$+ P_n^*(\overline{A}_1) P_n^*(A_2) ... P_n^*(A_{s+1}) P_n^*(\overline{A}_{s+2}) ... P_n^*(\overline{A}_m) + ... +$$
$$+ P_n^*(\overline{A}_1) P_n^*(\overline{A}_2) ... P_n^*(\overline{A}_{m-s}) P_n^*(A_{m-s+1}) ... P_n^*(A_m).$$

Очевидно, в останній сумі стільки доданків, скількома способами із m місць можна вибрати s місць для літер A_i без риски і решту m-s місць для літер $\overline{A}_j$ з рискою, або скількома способами в m-розрядному двійковому коді можна вибрати s місць для цифри 1 і решту m-s місць для цифри 0, тобто дорівнює C_m^s – кількості невпорядкованих s-елементних підмножин в m-елементній множині.

Якщо в розглядуваних випробуваннях $A_i = \{(E_1, E_2, ..., E_m) \mid E_i \in A,\ A \subset \Omega,\ E_j \in \Omega,\ i \in \overline{1,m},\ j \in \overline{1,m} \setminus \{i\}\}$, і крім того виявляється, що всі $P_n^*(A_i)$, $i \in \overline{1,m}$, практично рівні між собою – $P_n^*(A_i) = p$, $P_n^*(\overline{A}_i) = q = 1 - p$, $i \in \overline{1,m}$, (на практиці «швидше за все» так і буде), тоді, враховуючи формулу (2.10.1), одержимо, що з достатньо великою точністю практично мають місце рівності (отримані за статистичними даними):

$$P_n^*(H_0) = q \cdot q \cdot q \cdot ... \cdot q \cdot q \cdot q = q^m p^0,$$

$$P_n^*(H_1) = q \cdot q \cdot q \cdot ... \cdot q \cdot q \cdot p = q^{m-1} p^1,$$

$$P_n^*(H_2) = q \cdot q \cdot q \cdot ... \cdot q \cdot p \cdot q = q^{m-1} p^1,$$

$$P_n^*(H_3) = q \cdot q \cdot q \cdot ... \cdot q \cdot p \cdot p = q^{m-2} p^2,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$P_n^*(H_{k-1}) = p \cdot p \cdot p \cdot ... \cdot p \cdot p \cdot q = q^1 p^{m-1},$$

$$P_n^*(H_k) = p \cdot p \cdot p \cdot ... \cdot p \cdot p \cdot p = q^0 p^m.$$

Якщо попадання в множини H_i розглядати як єдино можливі наслідки випробування, яким вважається розглядувана серія із m випробувань, то тут виникає ще один ймовірнісний простір $(\breve{\Omega}, \breve{S}, \breve{P}_n^*)$ з елементарними подіями $\breve{E}_i = H_i$, найширшим простором подій $\breve{S}$ і ймовірностями елементарних подій $\breve{P}_n^*(\breve{E}_i) = p^s q^{m-s}$, $\breve{E}_i \in \breve{\Omega}$, $i \in \overline{0, 2^m - 1}$, $s \in \overline{0,m}$, де s – кількість відбувань події A в розглядуваній серії із m випробувань, $(m - s)$ – кількість невідбувань події A в розглядуваній серії із m випробувань, відповідно до кількості множників A_i та $\overline{A}_j$, $i \neq j$, у поданні $\breve{E}_i = H_i$, $i \in \overline{0, 2^m - 1}$, через добуток подій A_i та $\overline{A}_j$, $i \neq j$, $i \in \overline{1,m}$, $j \in \overline{1,m}$.

Позначимо через $B_{m,s}$ подію, яка полягає в тому, що в розглядуваній серії із m випробувань подія A відбувається s разів, $s \in \overline{0,m}$. Подія $B_{m,s}$ є об'єднанням (сумою) подій $\breve{E}_i = H_i$ таких, в яких в добутках H_i співмножник A_j без риски зустрічається s разів, $j \in \overline{1,m}$, а співмножник $\overline{A}_j$ зустрічається $(m - s)$ разів.

Тоді, враховуючи, що доданки в сумі $B_{m,s}$ попарно несумісні, а також те, що статистична ймовірність кожного доданку в сумі $B_{m,s}$ з достатньо великою точністю практично дорівнює $p^s q^{m-s}$, дістанемо, що з достатньо великою точністю практично має місце рівність

$$\breve{P}_n^*(B_{m,s}) = \sum_{\breve{E} \in B_{m,\,s}} P_n^*(\{\breve{E}\}) = C_m^s p^s q^{m-s}, \; s \in \overline{0,m}\,. \qquad (2.10.3)$$

Формулу (2.10.3) називають *формулою Бернуллі*.

Розподіл статистичних ймовірностей $\breve{P}_n^*(B_{m,s})$, що визначається за формулою (2.10.3), називають *біноміальним*, оскільки статистичні ймовірності $\breve{P}_n^*(B_{m,s})$ обчислюються так само, як члени розкладу бінома Ньютона $(p+q)^m$ за степенями p і q:

$$(p+q)^m = \sum_{s=0}^{m} C_m^s p^s q^{m-s}\,. \qquad (2.10.4)$$

Як відомо, C_m^s обчислюється за формулою

$$C_m^s = \frac{m(m-1)\cdot\ldots\cdot(m-(s-1))}{1\cdot 2\cdot\ldots\cdot s} = \frac{m!}{s!(m-s)!}, \; s \in \overline{0,m}, \; 0! = 1\,. \qquad (2.10.5)$$

Враховуючи, що події $B_{m,s}$ попарно несумісні і в m випробовуваннях принаймні одна з них обов’язково відбувається, дістаємо:

$$\sum_{s=0}^{m} \breve{P}_n^*(B_{m,s}) = \sum_{s=0}^{m} C_m^s p^s q^{m-s} = (p+q)^m = 1\,.$$

Зауважимо, що коли $B_{m,s}$, $s = 0, 1, ..., m$, розглядати як єдино можливі наслідки серії із m випробувань стосовно кількості появ події A в m випробуваннях, тоді породжується ще один простір елементарних подій (наслідків серії із m випробувань)

$$\hat{\Omega}_m = \{B_{m,0},\, B_{m,1},\, ...,\, B_{m,m}\}$$

з імовірнісною мірою $\hat{P}_n^*(B_{m,s}) = C_m^s p^s q^{m-s}$, $s = 0, 1, ..., m$, визначеною на всіх підмножинах множини $\hat{\Omega}_m$ (тобто $\hat{S}$ – найширша сукупність підмножин множини $\hat{\Omega}_m$).

В такому разі за довільного $B \subset \hat{\Omega}_m$ буде $\hat{P}_n^*(B) = \sum_{B_{m,s} \in B} \hat{P}_n^*(B_{m,s})$.

Приклад 2.10.5. Нехай велику кількість n разів (наприклад $n = 1000$) виконували 10 підкидань монети. Нехай $\Omega = \{\mathit{Г}, \mathit{Ц}\}$, подія $A = \{\mathit{Г}\}$ – випадання герба, $A \in S = \{\varnothing, \{\mathit{Г}\}, \{\mathit{Ц}\}, \{\mathit{Г}, \mathit{Ц}\}\}$, $\overline{A} = \{\mathit{Ц}\}$ – випадання цифри, $\overline{A} \in S = \{\varnothing, \{\mathit{Г}\}, \{\mathit{Ц}\}, \{\mathit{Г}, \mathit{Ц}\}\}$. Нехай подія A_i – відбування події A під час i-го підкидання монети, $\overline{A}_i$ – відбування події $\overline{A}$ під час i-го підкидання монети, $i \in \overline{1,10}$, H_0, H_1, …, H_{1023} – всеможливі добутки подій A_i і $\overline{A}_i$ із 10 співмножників (див. формули 2.10.2). Як події із простору $\tilde{S} = \{\tilde{A} \mid \tilde{A} = \bigcup_{i \in I} H_i, I \subset \{0,1,2,...,1023\}\}$ подій розглядатимемо разом з $\varnothing$ всеможливі об’єднання підмножин $H_i \subset \Omega^{10}$, $i \in \overline{0,1023}$, із одного, двох, трьох, і т.д., 1023, всіх 1024 доданків. Нехай за результатами проведення n=1000 серій із 10-ти підкидань монети одержали, що практично $P_n^*(A_i) = \frac{1}{2}$ за всіх $i \in \overline{1,10}$, тобто і під час перших підкидань, і під час других, і під час всіх інших із 10 підкидань подія $A = \{\varGamma\} \in S$ відбувалася із статистичною ймовірністю $P_n^*(A)$, з достатньою точністю досить близькою до $\frac{1}{2}$, а всі події A_i і $\overline{A}_j$, $i \in I \subset \{1,2,...,10\}$, $j \in \{1,2,...,10\} \setminus I$, виявилися практично незалежними в сукупності (статистично незалежними, див. формулу 2.10.1. На практиці «швидше за все» так і буде). Тоді, враховуючи формули (2.10.4) і (2.10.5), одержимо, що подія $B_{10,5} \in \tilde{S}$ (див. формулу 2.10.3), яка полягає в тому, що в 10 підкиданнях герб випадав 5 разів, в даній серії із n=1000 серій із 10-ти підкидань відбувалася із статистичною ймовірністю $\breve{P}_n^*(B_{10,5})$, що з досить великою точністю дорівнює

$$\breve{P}_n^*(B_{10,5}) = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^5 = \frac{10!}{5! \cdot 5!} \cdot \frac{1}{2^{10}} = \frac{252}{1024} \approx \frac{1}{4}.$$

Щоб знайти найбільшу за змінною $s \in \overline{0,m}$ статистичну ймовірність $\hat{P}_n^*(B_{m,s})$, розглянемо відношення:

$$\frac{\hat{P}_n^*(B_{m,s})}{\hat{P}_n^*(B_{m,s-1})} = \frac{C_m^s p^s q^{m-s}}{C_m^{s-1} p^{s-1} q^{m-(s-1)}} =$$

$$= \frac{\dfrac{m(m-1)\ ...(m-(s-1))}{1 \cdot 2\ ...s}}{\dfrac{m(m-1)\ ...(m-(s-2))}{1 \cdot 2\ ...\ (s-1)}} \cdot \frac{p^s q^{m-s}}{p^{s-1} q^{m-(s-1)}} = \frac{m-s+1}{s} \cdot \frac{p}{q} < 1,$$

звідки $mp + p - sp < sq$ або $mp + p < s(p+q)$, тобто $s > (m+1)p$.

Отже

$$P_n^*(B_{m,\,s}) < P_n^*(B_{m,\,s-1}) \text{ коли } \quad s > (m+1)p\ .$$

Аналогічно дістаємо, що

$$P_n^*(B_{m,\,s}) < P_n^*(B_{m,\,s-1}) \text{ коли } \quad 1 \le s < (m+1)p\ .$$

Таким чином, коли число $(m+1)p$ не є цілим і $s_0 = [(m+1)p]$ – найбільше ціле число, що не перевищує $(m+1)p$, то $\hat{P}_n^*(B_{m,s_0})$ є єдиною найбільшою статистичною ймовірністю, а число відбувань події A в серії із m незалежних випробувань, яке зустрічається найчастіше, близьке до $mP_n^*(A) = mp$.

У випадку, коли число $(m+1)p = s_0$ є цілим, дістаємо

$$\hat{P}_n^*(B_{m,s}) < \hat{P}_n^*(B_{m,s-1}), \text{ якщо } \ s \ge s_0,$$

а

$$\hat{P}_n^*(B_{m,s}) > \hat{P}_n^*(B_{m,s-1}), \text{ якщо } \ s < s_0, \text{ тобто } s \le s_0 - 1.$$

В такому разі

$$\frac{\hat{P}_n^*(B_{m,s_0})}{\hat{P}_n^*(B_{m,s_0-1})} = \frac{m - s_0 + 1}{s_0} \cdot \frac{p}{q} = \frac{(m+1) - (m+1)p}{(m+1)p} \cdot \frac{p}{q} =$$

$$= \frac{(m+1) \cdot (1-p)}{(m+1)p} \cdot \frac{p}{q} = 1,$$

тобто $\hat{P}_n^*(B_{m,s_0}) = \hat{P}_n^*(B_{m,s_0-1})$ – дві найбільші статистичні ймовірності.

Таким чином, статистичні ймовірності $\hat{P}_n^*(B_{m,s})$ *зростають за змінною s, коли* $s \in (0;\ s_0)$, *і спадають за цією змінною, коли* $s \in (\ s_0;\ m)$, *де* $s_0 = [(m+1)p]$. *Тому* $\hat{P}_n^*(B_{m,s_0})$ *є найбільшою статистичною ймовірністю, коли* $(m+1)p$ *не є цілим числом, а коли число* $(m+1)p$ *– ціле, то* $\hat{P}_n^*(B_{m,s_0}) = \hat{P}_n^*(B_{m,s_0-1})$ *– дві найбільші статистичні ймовірності.*

Крім того

$$(m+1)p - 1 < s_0 \le (m+1)p, \text{ звідки } \ p + \frac{p-1}{m} < \frac{s_0}{m} \le p + \frac{p}{m}\ .$$

Останнє означає, що коли m *досить велике, найчастіше число відбувань події A в серії із m незалежних випробувань виявляється рівним* s_0 *і таким, що із збільшенням m число* $\frac{s_0}{m}$ *стає як завгодно близьким до статистичної ймовірності* $P_n^*(A) = p$.

Приклад 2.10.6. Якщо $P_n^*(A) = p = \frac{1}{2}$ і $m = 10$, то ряд розподілу статистичних ймовірностей $\hat{P}_n^*(B_{10,s})$ набуває вигляду

s	0	1	2	3	4	5	6	7	8	9	10
$\hat{P}_n^*(B_{10,s})$	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$	$\frac{210}{1024}$	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$

Многокутник цього розподілу статистичних ймовірностей $\hat{P}_n^*(B_{10,s})$, $s \in \overline{0,10}$, зображено на Рис. 2.10.6. В розглядуваному випадку

$$s_0 = [(m+1)p] = \left[\frac{11}{2}\right] = \left[5\frac{1}{2}\right] = 5.$$

Отже, $\hat{P}_n^*(B_{10,5})$ – єдина найбільша статистична ймовірність за змінною $s \in \overline{0,\ 10}$, і крім того $\frac{s_0}{m} = \frac{5}{10} = P_n^*(A) = \frac{1}{2}$.

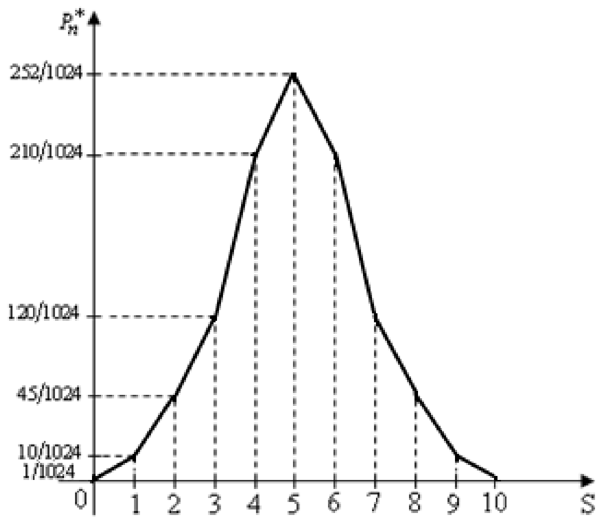


Рис. 2.10.6

Якщо статистична ймовірність відбування події A, $A \subset \Omega$, не одна й та сама в кожному з m незалежних випробувань, то така схема повторних випробувань називається *схемою Пуассона*, в якій дещо узагальнюється схема Бернуллі, за якою статистичні ймовірності відбування події A однакові в кожному із m випробувань (див. формулу 2.10.3). Формула для обчислення статистичної ймовірності того, що подія A в серії з m випробувань відбудеться рівно s разів за умови, що в i-му випробуванні вона відбувається з статистичною імовірністю p_i, набуває вигляду:

$$\breve{P}_n^*(B_{m,s}) = p_1 p_2 ... p_s q_{s+1} q_{s+2} ... \ q_m + q_1 p_2 ... p_{s+1} q_{s+2} ... \ q_m + \\ + ... + q_1 q_2 ... \ q_{m-s} p_{m-s+1} ... \ p_m. \qquad (2.10.6)$$

Якщо розглянути функцію

$$\Phi(z) = (p_1 z + q_1)(p_2 z + q_2) ... (p_m z + q_m),$$

то виявляється, що коефіцієнт біля z^s в розкладі $\Phi(z)$ за степенями z дорівнює ймовірності $\breve{P}_n^*(B_{m,s})$. *Функцію* $\Phi(z)$ називають *породжуючою*. Якщо всі p_i рівні між собою, тобто $p_i = p, \ i = 1, 2, ..., n$, то вираз породжуючої функції набуває вигляду $\Phi(z) = (pz + q)^m$.

Формули (2.10.3) і (2.10.6) можна узагальнити й на той випадок, коли в кожному випробуванні більше ніж два наслідки, причому ймовірності цих наслідків можуть залишатися однаковими в усіх випробуваннях або змінюватися із зміною номера випробування.

Якщо в кожному випробуванні є не два наслідки (A та $\overline{A}$), а k попарно несумісних наслідків $A_1, A_2, ..., A_k$, причому $A_i A_j = \varnothing$, коли $i \neq j, \quad A_1 + A_2 + ... + A_k = \Omega, \ P_n^*(A_i) = p_i, \ \sum_{i=1}^{k} p_i = 1$, то, міркуючи аналогічно до попереднього, дістанемо: статистична ймовірність того, що в серії з m незалежних випробувань події A_i відбудуться кожна m_i разів, $\left(\sum_{i=1}^{k} m_i = m \right)$, дорівнює (див. формулу 2.10.3)

$$\breve{P}_n^* \left(s_1 = m_1, \ s_2 = m_2, \ ..., \ s_k = m_k \right) = \\ = C_m^{m_1} C_{m-m_1}^{m_2} \ ... \ C_{m-m_1-m_2- \ ... -m_{k-1}}^{m_k} p_1^{m_1} p_2^{m_2} \ ... \ p_k^{m_k} = \\ = \frac{m!}{m_1! m_2! \ ... \ m_k!} \ p_1^{m_1} p_2^{m_2} \ ... \ p_k^{m_k}.$$

Надаючи змінним $m_1, m_2, ..., m_k$ значень в межах від 0 до n так, щоб завжди виконувалась умова $m_1 + m_2 + ... + m_k = m$, одержимо розподіл імовірностей на множині всіх можливих наборів значень k чисел $(m_1, m_2, ..., m_k)$. Такий *розподіл імовірностей* називається *поліноміальним*, оскільки ймовірності тут обчислюються за тими самими формулами, що й члени розкладу полінома $(p_1 + p_2 + ... + p_k)^m$ за степенями p_i.

Приклад 2.10.7. Баскетболіст робить п'ять незалежних кидків м'ячем в корзину. Потрібно знайти статистичну ймовірність того, що він влучить в корзину тричі, якщо з кожним кидком він влучає з статистичною ймовірністю 0.6.

Кожен кидок можна вважати випробуванням, в якому подія A (влучення м'яча в корзину) відбувається з статистичною ймовірністю $p = 0.6$ і не відбувається з статистичною ймовірністю $q = 1 - p = 0.4$. Таким чином, треба знайти статистичну ймовірність того, що в серії з п'яти випробувань подія A відбудеться тричі. Беручи до уваги, що $m = 5$, $s = 3$, за формулою Бернуллі (2.10.3) знаходимо

$$\breve{P}_n^*(B_{5,3}) = C_5^3 (0.6)^3 (0.4)^2 \approx 0.35 .$$

Досить часто доводиться обчислювати статистичну ймовірність події, яка полягає в тому, що величина s набуває значень в межах від m_1 до m_2, тобто статистичну ймовірність події $(m_1 \le s \le m_2)$. Оскільки подія $(m_1 \le s \le m_2)$ може бути подана у вигляді суми попарно несумісних подій

$$(m_1 \le s \le m_2) = (s = m_1) + (s = m_1 + 1) + \ ... \ + (s = m_2) ,$$

то

$$\breve{P}_n^* (m_1 \le s \le m_2) = \breve{P}_n^* (B_{m,m_1}) + \breve{P}_n^* (B_{m,m_1+1}) + \ ... \ + \breve{P}_n^* (B_{m,m_2}) =$$

$$= C_m^{m_1} p^{m_1} q^{m-m_1} + C_m^{m_1+1} p^{m_1+1} q^{m-m_1-1} + \ ... \ + C_m^{m_2} p^{m_2} q^{m-m_2} ,$$

або в більш компактній формі

$$\breve{P}_n^*(m_1 \le s \le m_2) = \sum_{k=m_1}^{m_2} C_m^k p^k q^{m-k} . \qquad (2.10.7)$$

Приклад 2.10.8. В мішень виконується 10 незалежних пострілів. Статистична ймовірність влучення під час кожного пострілу дорівнює 1/2. Знайти статистичну ймовірність того, що в мішені буде не менше восьми влучень.

Тут $m_1 = 8, \ m_2 = 10, \ m = 10, \ p = 1/2, \ q = 1 - p = 1/2$.

Тому за формулою (2.10.7) одержимо

$$\breve{P}_n^* (8 \le s \le 10) = \sum_{k=8}^{10} C_{10}^k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10-k} = \frac{56}{2^{10}} \approx 0.06 .$$

Приклад 2.10.9. Розглянемо дві серії із 5 і 6 повторних випробувань, в кожному з яких підкидається монета і тому в кожному два можливі наслідки $Г$ і $Ц$ (герб і цифра), статистичні ймовірності яких $p = \frac{1}{2}$ і $q = \frac{1}{2}$.

Тоді стосовно серії із 5-ти випробувань одержимо розподіл ймовірностей на множині можливих кількостей випадань герба:

s	0	1	2	3	4	5
$C_m^s p^s q^{m-s}$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Тут $m=5$, $s_0=(m+1)p=3$ – ціле число, і тому найбільш імовірних кількостей появ герба дві – s_0 і (s_0-1):

$$s_0=(m+1)p=3,\ s_0-1=(m+1)p-1=mp-q=5\cdot\frac{1}{2}-\frac{1}{2}=2\,.$$

Стосовно серії із 6-ти випробувань одержимо

s	0	1	2	3	4	5	6
$C_m^s p^s q^{m-s}$	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Тут

$$(m+1)p=\frac{7}{2},\quad s_0=[(m+1)p]=3\,.$$

Отже в даному випадку найбільш імовірна кількість випадань герба – 3.

Якщо m і s досить великі, то обчислення ймовірностей безпосередньо за формулами (2.10.3) і (2.10.7) стають надто громіздкими і практично нездійсненними. В таких випадках в разі наближеного обчислення ймовірності $\breve{P}_n^*(B_{m,s})$ використовують так звану *локальну асимптотичну теорему Муавра-Лапласа*, згідно з якою за досить великих m статистичну імовірність $\breve{P}_n^*(B_{m,s})$ наближено можна обчислювати за формулою

$$\breve{P}_n^*(B_{m,s})\approx\frac{1}{\sqrt{2\pi}\ \sqrt{mpq}}\,e^{-\frac{(s-mp)^2}{2mpq}}\,. \qquad (2.10.8)$$

Приклад 2.10.10. Проводиться серія з 400 випробувань, в кожному з яких подія A відбувається з статистичною імовірністю 1/2. Знайти статистичну ймовірність того, що подія A відбувається 190 разів.

Обчислення за формулою Бернуллі (2.10.3) тут досить громіздкі. Тому доцільно застосувати *локальну асимптотичну теорему Муавра-Лапласа* (див. додаток 3). Одержуємо

$$mpq=100,\ mp=200,\ s-mp=-10\,,$$

і таким чином

$$\breve{P}_n^*(B_{400,190}) \approx \frac{1}{\sqrt{2\pi}\cdot 10} e^{-\frac{1}{2}} .$$

Скориставшись для обчислення значення виразу

$$1/(SQRT(2*3.142)*\exp(-1/2)$$

послугами програми GRAN1, знайдемо

$$\widetilde{P}(\mu = 190) \approx 0.024 .$$

За досить великих m в разі обчислення ймовірності $\widetilde{P}(m_1 \le s \le m_2)$ використовують *інтегральну асимптотичну теорему Муавра-Лапласа* (див. додаток 2), згідно з якою

$$\widetilde{P}(m_1 \le s \le m_2) \approx \frac{1}{\sqrt{2\pi}\sqrt{npq}} \int_{m_1}^{m_2} e^{-\frac{(x-np)^2}{2npq}} dx ,$$

або $\left(\text{після заміни змінних } \frac{x\text{-}np}{\sqrt{npq}} = t \right)$

$$\widetilde{P}(m_1 \le s \le m_2) \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{m_1 - np}{\sqrt{npq}}}^{\frac{m_2 - np}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt .$$

За локальною асимптотичною теоремою Муавра-Лапласа отримуються досить точні результати, коли p і q близькі до 1/2. Проте виявляється, що чим ближче p або q до нуля, тим гірше наближення результату за формулою (2.10.8). В таких випадках зручно користуватися *теоремою Пуассона* (див. додаток 4):

Якщо статистична ймовірність $p > 0$ настання події A в кожному із серії з m випробувань мала, то за досить великих m

$$\breve{P}_n^*(B_{m,s}) \approx \frac{a^s}{s!} e^{-a}, \text{ де } a = mp \le c = const .$$

Справді, скориставшись формулою Бернуллі і врахувавши, що $a = mp$, звідки $p = \frac{a}{m} \le \frac{c}{m}$, одержимо

$$\breve{P}_n^*(B_{m,s}) = C_m^s p^s q^{m-s} = C_m^s \left(\frac{a}{m}\right)^s \left(1 - \frac{a}{m}\right)^{m-s} =$$

$$=\frac{m\,(m-1)\,\dots\,(m-(s-1))}{s!}\;\frac{a^s}{m^s}\left(1-\frac{a}{m}\right)^m\frac{1}{\left(1-\frac{a}{m}\right)^s}=$$

$$=\frac{a^s}{s!}\cdot 1\cdot\left(1-\frac{1}{m}\right)\dots\left(1-\frac{s-1}{m}\right)\left(1-\frac{a}{m}\right)^m\frac{1}{\left(1-\frac{a}{m}\right)^s}.$$

Оскільки за фіксованого s та досить великих m

$$1\cdot\left(1-\frac{1}{m}\right)\dots\left(1-\frac{s-1}{m}\right)\approx 1,\quad \frac{1}{\left(1-\frac{a}{m}\right)^s}\approx 1,$$

$$\left(1-\frac{a}{m}\right)^m=((1-\frac{a}{m})^{\frac{m}{a}})^a\approx e^{-a},$$

то за досить великих m

$$\breve{P}_n^*(B_{m,s})\approx\frac{a^s}{s!}e^{-a},\text{ де }\ a=mp.$$

Приклад 2.10.11. Проводиться серія з 1000 випробувань, в кожному з яких подія A відбувається з статистичною ймовірністю $p=0.0001$. Знайти статистичну ймовірність того, що в цій серії випробувань подія A відбудеться п'ять разів.

Тут m досить велике, а p близьке до нуля, тому зручно скористатися теоремою Пуассона. Одержуємо

$$a=mp=1000\cdot 0.001=1,\ \breve{P}_n^*(B_{1000,5})\approx\frac{1}{5!}e^{-1}\approx 0.003.$$

Вправи для самостійного виконання

2.10.1. Перевірити, чи правильні твердження:

1. Через серію з m випробувань задається новий стохастичний експеримент, а не той, з яким пов’язане кожне з цих m випробувань.

2. Якщо експеримент задається через серію з m випробувань, в кожному з яких фіксується відбування або невідбування заданої події A, то можливими наслідками цього експерименту є набори чисел: $(c_1,c_2,\dots,c_m)$, де $c_k=1$, коли подія A відбувається в k-му випробуванні, і $c_k=0$, коли A не відбувається в k-му випробуванні.

3. Множина $\{j \mid j\in N,\ c_j=1\}$ завжди є непорожньою.

4. Кожна елементарна подія $(c_1,c_2,\dots,c_m)$ із твердження 2 є добутком m незалежних в сукупності подій, кожна k-а з яких полягає або у відбуванні (подія A_k), або у невідбуванні (подія $\overline{A_k}$) події A в k-му випробуванні.

5. За кожної елементарної події $(c_1, c_2, ..., c_m)$ із твердження 2 існує єдине число $s \in \overline{0,\ m}$ таке, що $P_n^*(c_1, c_2, ..., c_m) = p^s q^{m-s}$.

6. Якщо подія $B_{m,s}$ полягає в тому, що в m випробуваннях подія A відбувається s разів, то ця подія складається з тих елементарних подій $(c_1, c_2, ..., c_m)$, за яких серед чисел c_k є: 1) принаймні s одиниць; 2) рівно $m-s$ нулів; 3) принаймні $m-s$ нулів; 4) рівно s одиниць.

7. Всі елементарні події $(c_1, c_2, ..., c_m)$, з яких складається подія $B_{m,s}$, є рівноможливими між собою.

8. Схема Бернуллі пов'язана з серією із m незалежних випробувань, в кожному з яких фіксована подія A відбувається з однією і тією самою ймовірністю, тобто кожне випробування пов'язане з одним і тим самим ймовірнісним простором.

9. Серії із m випробувань можуть бути пов'язані з різними просторами елементарних подій.

10. Повторні випробування пов'язані з одним і тим самим простором елементарних подій.

11. Незалежність випробувань означає незмінність умов проведення цих випробувань, тобто незмінність відповідного імовірнісного простору.

12. Сума всіх статистичних ймовірностей $\hat{P}_n^*(B_{m,s})$ не залежить від m.

2.10.2. Довести, що в формулах (2.10.2) $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup\limits_{i=0}^{k} H_i = \tilde{\Omega} = \Omega^m$.

2.10.3. Знайти статистичну ймовірність того, що в серії з m практично (статистично) незалежних випробувань подія A відбувається принаймні один раз, якщо статистична ймовірність відбування цієї події в кожному випробуванні практично дорівнює $p = P_n^*(A) > 0$.

2.10.4. Знайти статистичну ймовірність того, що в серії із 9 підкидань монети герб випадає:

а) 5 разів;

б) 4 рази,

вважаючи, що за кожного підкидання $P_n^*(\{\Gamma\}) = \dfrac{1}{2}$.

Порівняти ці статистичні ймовірності.

Побудувати многокутник розподілу статистичних ймовірностей $\hat{P}_n^*(B_{9,s})$, $s \in \overline{0,\ 9}$.

2.10.5. Розглянути приклад 2.10.6 за умов $m = 20$, $p = \dfrac{1}{2}$. Визначити найменше ціле число l, за якого матиме місце рівність $\hat{P}_n^*(s_0 - l \le m \le s_0 + l) \ge 0.95$.

РОЗДІЛ ІІІ. ЙМОВІРНОСТІ

3.1. Ймовірнісні міри. Означення ймовірності

Перш ніж розглянути означення ймовірності випадкової події, нагадаємо деякі поняття з теорії міри.

Означення. Нехай V – деяка алгебра підмножин з Ω. Дійсна функція $\mu(A)$ від множин з V, що набуває значень з $[0;\infty]$, називається *скінченно-адитивною мірою*, заданою на V, якщо за будь-яких $A \in V,\ B \in V$ таких, що $A \bigcap B = \varnothing$,

$$\mu(A+B) = \mu(A) + \mu(B).$$

Скінченно-адитивна *міра* називається *скінченною*, якщо $\mu(\Omega) < \infty$. Легко бачити, що $\mu(\varnothing) = 0$ стосовно будь-якої скінченної міри μ.

Якщо $\mu(\Omega) = 1$, тоді скінченно-адитивна міра називається *скінченно-адитивною ймовірнісною мірою*.

Означення. Простір Ω разом з σ-алгеброю S його підмножин називається *вимірним простором* і позначається (Ω, S).

Означення. Скінченно-адитивна *міра* μ, задана на алгебрі V підмножин множини Ω, називається *зчисленно-адитивною* (*σ-адитивною*), якщо за будь-яких $A_1, A_2, \ldots$ з V таких, що $\sum_{i=1}^{\infty} A_i \in V,\ A_i \bigcap A_j = \varnothing$, коли $i \neq j$, виконується рівність

$$\mu\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Означення. Зчисленно-адитивна міра P на алгебрі V, стосовно якої задовільняється умова $P(\Omega) = 1$, називається *ймовірнісною мірою*, або *ймовірністю* (визначеною на множинах алгебри V).

Отже, стосовно ймовірнісної міри (ймовірності) $P(A)$, $A \in V$, визначеної на множинах $A \in V$, задовільняються вимоги:

1_P. $P(A) \geq 0$, $A \in V$.

2_P. $P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ за будь яких $A_i \in V$ таких, що $\sum_{i=1}^{\infty} A_i \in V$, $A_i A_j = \varnothing$, коли $i \neq j$.

3_P. $P(\Omega) = 1$.

Сформулюємо теорему, за якою визначаються умови, за яких скінченно-адитивна функція є одночасно й зчисленно-адитивною.

Теорема. Нехай P – скінченно-адитивна міра, задана на алгебрі V, причому $P(\Omega)=1$. Тоді наступні умови еквівалентні:

1) P є σ-адитивною;

2) P неперервна знизу, тобто за будь-яких $A_1, A_2, \ldots$ з V таких, що $A_i \subset A_{i+1}$, $\bigcup_{i=1}^{\infty} A_i \in V$, виконується рівність

$$\lim_{i\to\infty} P(A_i) = P\left(\bigcup_{i=1}^{\infty} A_i\right);$$

3) P неперервна зверху, тобто за будь-яких $A_1, A_2, \ldots$ з V таких, що $A_i \supset A_{i+1}$, $\bigcap_{i=1}^{\infty} A_i \in V$, виконується рівність

$$\lim_{i\to\infty} P(A_i) = P\left(\bigcap_{i=1}^{\infty} A_i\right);$$

4) P неперервна в $\varnothing$, тобто за будь-яких $A_1, A_2, \ldots$ з V таких, що $A_{i+1} \subseteq A_i$, $\bigcap_{i=1}^{\infty} A_i = \varnothing$, виконується рівність

$$\lim_{i\to\infty} P(A_i) = 0.$$

5) P неперервна в Ω, тобто за будь-яких $A_1, A_2, \ldots$ з V таких, що $A_i \subset A_{i+1}$ і $\bigcup_{i=1}^{\infty} A_i = \Omega$, виконується рівність

$$\lim_{i\to\infty} P(A_i) = 1.$$

Умови 2) і 3) можна записати так:

$$P(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} P(A_i).$$

Властивість, що визначається за останньою рівністю, називається властивістю неперервності зчисленно-адитивної ймовірнісної міри.

Важливою є така теорема.

Теорема (Каратеодорі про продовження міри). Нехай P – зчисленно-адитивна ймовірнісна міра на алгебрі V. Тоді існує єдина зчисленно-адитивна ймовірнісна міра Q, визначена на найменшій σ-алгебрі $\sigma(V)$, в яку входить V, яка є продовженням міри P, тобто міра Q така, що $Q(A)=P(A)$ за кожного A з V.

Одним з основних понять теорії ймовірностей, в основу якого покладено *систему аксіом* 1_s-3_s, 1_p-3_p *А.М. Колмогорова*, є поняття ймовірнісного простору.

Означення. Набір об'єктів (Ω, S, P), де:

а) Ω – множина точок E;

б) S – σ–алгебра підмножин множини Ω;

в) P – ймовірність на S,

називається *ймовірнісним простором* (або *ймовірнісною моделлю* стохастичного випробування).

В даному разі Ω називається *простором наслідків* випробування або *простором елементарних подій*, множини A з S називаються *подіями*, сукупність S підмножин множини Ω називається *простором подій*, а числа $P(A)$, $A \in S$, – називаються *ймовірностями подій* A.

Приклад 3.1.1. Нехай випробування полягає в тому, що навмання вибирається точка з проміжка $[0;1)$, причому всі точки проміжка вважаються "рівноможливими". В такому разі неможливо приписати ймовірність кожному елементу простору $\Omega=[0;1)$ елементарних подій так, щоб за ймовірностями елементарних подій можна було визначити ймовірності попадання в будь-які підмножини множини Ω (оскільки якщо всі точки з $[0;1)$ рівноможливі, то довелося б покласти $P(E)=0,\ E\in\Omega=[0;1))$. Тоді ймовірність приписують не окремим елементам, а деяким підмножинам множини Ω. В такому разі вимагають, щоб стосовно ймовірності задовільнялися певні вимоги, наприклад, ймовірності попадання в інтервали однакової довжини були однакові (якщо точки "рівноможливі"), ймовірність попадання в підмножину деякої множини не була більшою, ніж ймовірність попадання в саму множину, і т.п.

Крім того, якщо визначено ймовірності подій A і B, то бажано мати можливість визначити і ймовірності подій

$$\overline{A},\ \overline{B},\ A\bigcup B,\ A\bigcap B,\ A\setminus B\ .$$

Нехай ймовірність попадання в проміжок $[a_i\,;b_i)\subset[0;1)$ дорівнює (b_i-a_i), а в об'єднання скінченної кількості проміжків $[a_i\,;b_i)$, які не перетинаються, $\sum\limits_{i=1}^{n}(b_i-a_i)$, $i\in\overline{1,\ n}$. Розглянемо систему L всіх проміжків виду $[a_i\,;b_i)$ і скінченних сум таких проміжків. Нехай $\mathscr{B}=\sigma(L)$ – найменша σ–алгебра, в яку входить система L, тобто σ–алгебра борелівських множин проміжка $[0;1)$. За теоремою Каратеодорі існує єдина міра $P(A)$ на $\mathscr{B}=\sigma(L)$ така, що $P([a;b))=b-a$, $[a;b)\subset[0;1)$, $[a;b)\in L$. Розглядатимемо як випадкові події лише борелівські підмножини множини $\Omega = [0;1)$. Ймовірністю події $A \in \mathscr{B}$ назвемо $P(A)$.

Таким чином, побудовано ймовірнісну модель (ймовірнісний простір) $(\Omega, \mathscr{B}, P)$ випробування, яке полягає в тому, що з проміжка $[0;1)$ навмання вибирається точка, тобто задано:

1) простір елементарних подій $\Omega = [0;1)$;
2) σ-алгебру $\mathscr{B}$ випадкових подій;
3) ймовірність $P(A)$, $A \in \mathscr{B}$.

Слід підкреслити, що оскільки в $[0;1)$ існують невимірні підмножини, то неможливо на множині всіх підмножин проміжка $[0;1)$ визначити ймовірність (ймовірнісну міру) вказаним вище способом. Тому в даному разі як випадкові події розглядаються не всі підмножини множини $\Omega = [0;1)$, а тільки частина з них, а саме σ-алгебра всіх борелівських підмножин проміжка $[0;1)$.

Сукупність $\mathscr{B}([0;1])$ борелівських підмножин відрізка $[0;1]$ можна дістати й як $\{A \bigcap [0;1]\}$, $A \in \mathscr{B}(R^1)$, де $\mathscr{B}(R^1)$ – сукупність борелівських множин на числовій прямій $R^1 = (-\infty; \infty)$.

Приклад 3.1.2. Нехай $\Omega = \{\Gamma, Ц\}$, $S = \{\varnothing, \Omega, \{\Gamma\}, \{Ц\}\}$ і $A_0 = \{\Gamma\}$. Покладемо $P(\{\Gamma\}) = p$, $P(\{Ц\}) = q = 1 - p$, $P(\varnothing) = 0$ і $P(\Omega) = 1$. Тоді стосовно функції $P(A)$, $A \in S$, задовільняються умови 1_p-3_p, а тому $P(A)$ є ймовірнісною мірою довільної події з S. Зокрема, $P(A_0) = p$ – ймовірність події $A_0 = \{\Gamma\}$.

Приклад 3.1.3. Нехай $\Omega = \{\Gamma, Ц\}$, а $S = \{\varnothing, \Omega\}$, і нехай $P(\varnothing) = 0$, а $P(\Omega) = 1$. Тоді стосовно функції $P(A)$, $A \in S$, задовільняються вимоги 1_p-3_p, отже $P(A)$ є ймовірнісною мірою довільних подій з так визначеного простору подій S.

Проте в просторі S немає множини $A_0 = \{\Gamma\} \subset \Omega$, і тому в даному випадку не можна говорити не тільки про ймовірність події $A_0 = \{\Gamma\}$, а й навіть про подію $A_0 = \{\Gamma\}$.

Таким чином, щоб говорити про подію A та її ймовірність, треба мати, по-перше – простір Ω елементарних подій, по-друге – простір S подій, кожна з яких є підмножиною простору Ω, і крім того $A \in S$, а стосовно сукупності S підмножин множини Ω задовільняються вимоги 1_s-3_s, і по-третє – функцію P, що визначена на S і стосовно якої задовільняються вимоги 1_p-3_p. Тоді підмножину $A \subset \Omega$ таку, що $A \in S$, можна називати подією, а значення $P(A)$ вказаної функції можна називати ймовірнісною мірою або ймовірністю події A.

Отже лише після того, як задано множину Ω можливих наслідків випробування, сукупність S підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, та функцію $P(A)$, яка визначена на всіх елементах $A \in S$ і стосовно якої задовільняються вимоги 1_p-3_p, тобто після того, як побудовано ймовірнісний простір (Ω, S, P), підмножини множини Ω, які входять до сукупності S, називаються *подіями*; сама сукупність S називається *простором подій*; значення функції $P(A)$ називаються ймовірностями подій $A \in S$.

Таким чином *остаточно сказати, які саме підмножини множини Ω вважаються подіями і як слід обчислювати ймовірності таких подій, можна лише після того, як побудовано ймовірнісний простір* (Ω, S, P).

Розглянемо простір Ω елементарних подій, що є скінченним або зчисленним і таким чином всі елементи цього простору можна занумерувати натуральними числами від 1 до k або всіма натуральними числами, тобто $\Omega = \{E_1, E_2, ..., E_k\}$ або $\Omega = \{E_1, E_2, \dots, E_k, \dots\}$. Такий простір елементарних подій називають *дискретним*.

Нехай простір S випадкових подій – це сукупність будь-яких підмножин дискретного простору елементарних подій Ω. Зокрема, кожна одноелементна множина $\{E_i\} \in S$. Тоді кожну подію $A \in S$ можна подати у вигляді:

$$A = \{E_{i_1}, E_{i_2}, ..., E_{i_m}\} \text{ або } A = \{E_{i_1}, E_{i_2}, ..., E_{i_m}, ...\},$$

тобто $A = \sum\limits_{E_i \in A} \{E_i\}$.

Звідси випливає, що коли простір Ω – дискретний, то для задання ймовірності $P(A)$, $A \in S$, достатньо задати $P(\{E_i\})$ за кожного можливого i, оскільки за властивістю 2_p

$$P(A) = \sum_{E_i \in A} P(\{E_i\}).$$

Крім цього слід вимагати, щоб виконувались умови $P(\{E_i\}) \ge 0$ за всіх i і $\sum\limits_i P(\{E_i\}) = 1$. Тоді

$$P(\Omega) = P(\sum_i \{E_i\}) = \sum_i P\{E_i\} = 1, \text{ а } P(\varnothing) = 1 - P(\Omega) = 0,$$

оскільки $\varnothing = \overline{\Omega}$, $\varnothing \cap \Omega = \varnothing$, $\varnothing + \Omega = \Omega$.

Приклад 3.1.4. Нехай $\Omega = \{"1", "2", "3", "4", "5", "6"\}$, S – сукупність всіх підмножин простору Ω, і $A_1 = \{"3", "6"\}$.

Покладемо $P(\{"i"\}) = \frac{1}{6}$ за кожного $i \in \overline{1,6}$. Цим самим задано ймовірність будь-якої події $A \in S$. Зокрема в даному разі буде $P(A_1) = P(\{"3","6"\}) = P(\{"3"\}) + P(\{"6"\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Приклад 3.1.5. Нехай $a < b$ і $\Omega = [a;b]$ – простір елементарних подій, а простір S випадкових подій, породжений за сукупністю скінченних об'єднань проміжків, що є частинами $[a;b]$. Тоді кожній події $A \in S$ можна приписати міру (довжину) $m(A)$ (чим узагальнюється поняття довжини відрізка). Тому існує функція $P(A)$, $A \in S$, така, що $P(A) = \frac{m(A)}{b-a}$, і стосовно якої задовільняються основні вимоги 1_p-3_p до ймовірнісної міри. Зокрема, якщо $A = \langle \alpha; \beta \rangle \subset [a;b]$, то $P(A) = \frac{\beta - \alpha}{b - a}$, а коли $A = \bigcup_i \langle \alpha_i; \beta_i \rangle$, де проміжки $\langle \alpha_i; \beta_i \rangle \subset [a;b]$ попарно не перетинаються, то

$$P(A) = \frac{1}{b-a} \sum_i (\beta_i - \alpha_i) = \sum_i \frac{\beta_i - \alpha_i}{b-a} = \sum_i P(A_i), \text{ де } A_i = \left\langle \alpha_i ; \beta_i \right\rangle.$$

Однак виявляється, що не кожна підмножина $B \subset [a;b]$ є вимірною за так заданою мірою, тобто не кожна підмножина $B \subset [a;b]$ є подією. Відомо, що в кожному проміжку $\langle \alpha; \beta \rangle$, де $\alpha < \beta$, міститься підмножина B, що не є вимірною (якій неможливо приписати довжину), а тому в розглядуваному випадку $B \overline{\in} S$, тобто B не є подією. В той же час до простору S належать події (множини) виду $\bigcup_i \langle \alpha_i; \beta_i \rangle$, а також перетини та різниці таких множин.

Зауважимо також, що в останньому прикладі кожній елементарній події $E = x \in [a;b]$ ставиться у відповідність нульова ймовірність, оскільки $P(E) = P(\{E\}) = P(\{x\}) = P([x;x]) = 0$.

Отже, якщо простір Ω не є дискретним, то за ймовірностями елементарних подій не завжди можна визначати ймовірності всіх подій $A \in S$.

Приклад 3.1.6. Нехай в результаті підкидання шестигранного кубика велику кількість n разів фіксували наслідки:

E_1 – на верхній грані випадає цифра 6, $\{E_1\} = \{"6"\}$;

E_2 – на верхній грані випадає цифра 5, $\{E_2\} = \{"5"\}$;

E_3 – на верхній грані випадає цифра не більша, ніж 4, тобто $\{E_3\} = \{"1", "2", "3", "4"\}$.

Легко бачити, що статистична ймовірність $P_n^*(A) = \dfrac{k_n(A)}{n}$, де $k_n(A)$ - кількість відбувань події A в серії із n випробувань, є ймовірнісною мірою на сукупності
$S = \{\varnothing, E_1, E_2, E_3, E_1 + E_2, E_1 + E_3, E_2 + E_3, E_1 + E_2 + E_3\}$
підмножин множини Ω, оскільки стосовно $P_n^*(A)$, $A \in S$, задовільняються вимоги $1_р$-$3_р$ до ймовірнісної міри. Нехай виявилося, що $P_n^*(\{E_1\}) = 0.6$; $P_n^*(\{E_2\}) = 0.3$; $P_n^*(\{E_3\}) = 0.1$.

Тоді стосовно множини $\{"1", "2", "3", "4", "5", "6"\}$ можна обчислити ймовірності попадання лише в підмножини

$$E_1 = \{"6"\},\ E_2 = \{"5"\},\ E_3 = \{"1", "2", "3", "4"\},$$

$$\{E_1\} + \{E_2\} = \{"5", "6"\},\ \{E_1\} + \{E_3\} = \{"1", "2", "3", "4", "6"\},$$

$$\{E_2\} + \{E_3\} = \{"1", "2", "3", "4", "5"\},\ \{E_1\} + \{E_2\} + \{E_3\} = \Omega,\ \varnothing.$$

Разом з тим за наведеного задання подій та їх ймовірнісних мір ймовірності попадання в підмножини виду $\{"1"\}$, $\{"2"\}$, $\{"2", "3"\}$, $\{"1", "5"\}$, $\{"2", "6"\}$ і т.д. обчислити неможливо – такі підмножини виявляються “*невимірними*” відносно заданої ймовірнісної міри P_n^* і не розглядаються як події.

Приклад 3.1.7. Нехай Ω – довільний простір елементарних подій, а S – деякий простір випадкових подій, відповідний Ω.

Чи завжди можна задати на просторі S функцію, стосовно якої задовільняються вимоги $1_р$-$3_р$?

Для відповіді на це питання виокремимо скінченну або зчисленну кількість елементарних подій $E_i \in \Omega$, $i = 1, 2, \ldots$, і покладемо $P(\{E_i\}) = p_i \in [0;1]$, де $\sum\limits_i p_i = 1$. Після цього вважатимемо, що $P(A) = \sum\limits_{E_i \in A} P(\{E_i\})$ за будь-якого $A \in S$, причому $P(A) = 0$, коли $E_i \overline{\in} A$ за всіх i.

Зокрема, якщо виокремити лише одну елементарну подію $E_1 \in \Omega$, то можна вважати, що

$$P(A) = \begin{cases} 1, & \text{коли } E_1 \in A, \\ 0, & \text{коли } E_1 \overline{\in} A. \end{cases}$$

Легко бачити, що стосовно так визначеної функції $P(A)$, $A \in S$, задовільняються вимоги $1_р$–$3_р$, тобто $P(A)$ є ймовірнісною мірою довільної події $A \in S$.

Отже, на будь якому просторі подій S можна задати деяку ймовірнісну міру $P(A)$, $A \in S$. Більш того, якщо в просторі S є більше двох подій, то на ньому можна задати більше однієї функції $P(A),\ A \in S$.

Підкреслимо ще раз, що поняття ймовірності події вводиться не відносно окремої події, а відносно простору подій, за ймовірностями елементарних подій не завжди можна визначити ймовірність довільної події $A \in S$, а визначеність ймовірностей подій $A \in S$ не гарантує визначеності ймовірностей попадання в будь-які підмножини множини Ω, зокрема і визначеності ймовірностей елементарних подій.

Слід зауважити, що коли на S задано довільну скінченну міру $V(A),\ A \in S$, тобто таку, що $0 < V(\Omega) < \infty$, тоді нормуванням міри $V(A),\ A \in S$, можна задати на S ймовірнісну міру $P(A) = \dfrac{V(A)}{V(\Omega)},\ A \in S$.

Підкреслимо, що на довільній непорожній сукупності S підмножин множини Ω можна задати деяку функцію множин $V(A),\ A \in S$. Якщо стосовно сукупності S задовільняються вимоги 1_S-3_S, а стосовно функції V – задовільняються вимоги 1_P-3_P, тоді *функцію V і сукупність S називають узгодженими стосовно означення ймовірнісної міри*. *В іншому разі функцію V і сукупність S називають неузгодженими* стосовно означення ймовірнісної міри.

Оскільки статистична ймовірність

$$P_n^*(A) = \frac{k_n(A)}{k_n(\Omega)},\ A \in S ,$$

є ймовірнісною мірою, визначеною за результатами деякої досить довгої чи не дуже довгої (n може набувати значень 1, 2, 3 і т.д.) серії випробувань, то всі твердження і висновки, сформульовані стосовно статистичних ймовірностей, які випливають з вимог (аксіом) 1_s-3_s, 1_p-3_p, залишаються правильними і стосовно ймовірнісної міри $P(A)$, заданої на сукупності S підмножин множини Ω будь яким чином – на основі деяких припущень, досліджень, чи ще якось.

Приклад 3.1.8. Підкидається кубик, одна грань якого пофарбована в білий колір, дві грані пофарбовані в зелений колір, три грані пофарбовані в червоний колір. Потрібно знайти ймовірність того, що кубик впаде догори зеленою гранню, якщо вважається, що ймовірності випадання всіх граней однакові.

Тут робиться припущення, чи можливо висновок за результатами якихось досліджень чи розмірковувань, що ймовірності випадання всіх шести граней однакові, тобто що ймовірність випадання кожної окремої грані дорівнює 1/6. Однак ніякими логічними міркуваннями неможливо ані довести, ані спростувати, що ймовірність випадання кожної окремої грані кубика справді дорівнює 1/6, яким би правильним і однорідним не був кубик.

Тому всі висновки в подібних випадках будуть гіпотетичними (модельними) – якщо вихідні дані саме такі, то і висновки (згідно з теорією ймовірнісної міри) будуть відповідними і правильними стосовно вказаних даних, хоча можливо вказані дані ніяк не пов’язані з реальною дійсністю і є чисто уявними (модельними).

Разом з тим розв’язування таких модельних (чисто теоретичних) задач, що розглядаються стосовно якихось гіпотетичних даних і умов, досить часто приводить до досить важливих практично значущих результатів і висновків.

Надалі будемо розглядати деяку ймовірнісну міру $P(A)$, $A \in S$, задану на тій чи іншій сукупності S підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, не з’ясовуючи питання, як ця міра отримана – в результаті якихось припущень, якихось досліджень, узагальнення результатів великої кількості випробувань чи ще якимось іншим чином.

Зауважимо, що оскільки властивості статистичних ймовірностей $P_n^*(A)$, що випливають з аксіом 1_s-3_s, 1_p-3_p, залишаються без жодних змін і за довільної ймовірнісної міри $P(A)$, досить у відповідних виразах і твердженнях стосовно статистичних ймовірностей замінити позначення $P_n^*(A)$ на $P(A)$, а слова «статистична ймовірність» замінити словом «ймовірність», щоб отримати такі самі вирази і твердження, які будуть правильні стосовно ймовірностей.

Без змін залишаються і правила визначення ймовірнісної міри суми подій, добутку подій, означення залежних і незалежних відносно певної ймовірнісної міри подій, формула повної ймовірності, формула Байєса, формула Бернуллі (стосовно серії із *m* послідовних випробувань).

Без жодних змін залишаються і твердження та висновки, що стосуються розподілів статистичних ймовірностей на множині Ω, лише потрібно замінити позначення $P_n^*(A)$ на $P(A)$, позначення щільності розподілу статистичних ймовірностей $f_n^*(E)$ замінити на позначення щільності розподілу ймовірностей $f(E)$, $f_n^*(x)$ – на $f(x)$, позначення функції розподілу статистичних ймовірностей $F_n^*(x)$ на позначення функції розподілу ймовірностей $F(x)$, позначення координати центра розсіювання статистичних ймовірностей m_n^* замінити на позначення координати центра розсіювання ймовірностей x_c, позначення дисперсії розподілу статистичних ймовірностей D_n^* на позначення дисперсії розподілу ймовірностей D і т.д.

Вправи для самостійного виконання

5.1.1. Перевірити, чи правильні твердження:

1. Система V підмножин множини Ω є алгеброю, якщо $(A+B)\in V, (A\cdot B)\in V$ за будь-яких множин $A\in V$ і $B\in V$.

2. Твердження, обернене до 1, є правильним.

3. Якщо V є алгеброю, то вона є й σ-алгеброю.

4. Твердження, обернене до 3, є правильним.

5. Сукупність V всеможливих скінченних підмножин множини N натуральних чисел є алгеброю.

6. Сукупність V всеможливих скінченних та зчислених підмножин множини N натуральних чисел є σ-алгеброю.

7. Сукупність всеможливих континуальних підмножин відрізка $[0;1]$ є σ-алгеброю.

8. Якщо V_1 і V_2 – σ-алгебри підмножин множини Ω, то $V_1 \bigcap V_2$ також σ-алгебра.

9. Система L, що складається із скінченних сум проміжків виду $[a;b)$ є: а) алгеброю, б) σ-алгеброю.

10. Кожна підмножина множини $\Omega=(-\infty;+\infty)$ є борелівською множиною.

11. Кожна числова функція, що визначена на алгебрі V і набуває невід'ємних значень, є мірою.

12. Якщо μ – міра, то $\mu(\varnothing) = 0$.

13. Існують скінченно адитивні міри, що не є зчисленно адитивними.

14. Кожна ймовірнісна міра є зчисленно адитивною і скінченною мірою .

15. Для того, щоб скінченно адитивна міра була зчисленно адитивною, необхідно і достатньо, щоб вона була неперервною зверху або знизу.

16. Кожну зчисленно адитивну міру можна продовжити з алгебри V, на якій вона визначена, на найменшу σ–алгебру $\sigma(V)$, до якої входить V.

17. Продовження із завдання 16 не є єдиним.

18. За кожною ймовірнісною мірою можна визначити певний ймовірнісний простір.

19. Кожен простір елементарних подій є дискретним.

20. За будь-якого ймовірнісного простору (Ω, S, P) ймовірність будь-якої події A цілком визначається за ймовірностями одноелементних подій $\{E\}$, де $E\in A$.

21. На будь-якому просторі подій можна задати ймовірнісну міру.

22. На будь-якому просторі подій, за якого події $A\subset(-\infty;+\infty)$, можна задати ймовірнісну міру P, стосовно якої задовільняється умова

$P(A)=P(A_\varepsilon)$, де $A_\varepsilon=\{x \mid y=x+\varepsilon,\, x\in A\}$, тобто A_ε є результат паралельного перенесення події A на величину ε.

23. За кожного простору елементарних подій Ω можна побудувати єдиний вимірний простір $(\Omega,\ S)$.

24. Якщо невід'ємна функція $P(A)$, визначена на певній сукупності S підмножин простору Ω, є цілком аддитивною і $P(\Omega)=1$, то $P(A)$, $A\in S$, є ймовірнісною мірою (ймовірністю).

25. Якщо $P(A+B)=P(A)+P(B)$, то події A і B несумісні.

5.1.2. Довести, що коли V є алгеброю, проте не є σ–алгеброю, то V – нескінченна множина.

5.1.3. Довести, що коли алгебра V є скінченною сукупністю підмножин множини Ω, то кожна міра на V є зчисленно аддитивною.

5.1.4. Нехай $\Omega=\{x_k,\, k\in N\}$ – довільна зчисленна множина, а S складається з таких підмножин A множини Ω, за яких або A – скінченна, або $\overline{A}$ – скінченна. Довести, що: 1) S є алгеброю; 2) S не є σ-алгеброю; 3) якщо $\mu(A)=0$, коли $A\in S$, і скінченна, а $\mu(A)=1$, коли $A\in S$, і нескінченна, то μ – скінченно аддитивна міра, проте не σ-аддитивна.

5.1.5. Побудувати ймовірнісний простір стосовно експерименту, який полягає у підкиданні монети до першого випадання герба, вважаючи, що ймовірність випадання герба за кожного підкидання дорівнює $p\in(0;1)$.

3.2. Функція та щільність одновимірного розподілу ймовірностей на $(R^1, \mathscr{B}(R^1))$ та їх властивості. Міра Лебега-Стільтьєса

Нехай $P(A)$ – *ймовірнісна міра, визначена на вимірному просторі* $(R^1, \mathscr{B}(R^1))$, *тобто на борелівських множинах* A *числової прямої*, $A \in S = \mathscr{B}(R^1)$. Нехай стосовно $A = (-\infty; x)$, $A \in S = \mathscr{B}(R^1)$, задана функція

$$F(x) = P((-\infty; x)), \quad x \in R^1.$$

Зауважимо, що оскільки в розглядуваному випадку $\bigcup\limits_{A \in S,\ A \subset (-\infty;\ x)} A = (-\infty;\ x)$, то $F(x)$ можна подати у вигляді

$$F(x) = P(\bigcup_{A \in S,\ A \subset (-\infty;\ x)} A).$$

Стосовно визначеної таким чином функції $F(x)$ задовільняються наступні властивості (порівняйте з § 2.7):

1°. $F(x)$ – неспадна.

2°. $F(-\infty) = 0,\ F(+\infty) = 1$, де

$$F(-\infty) = \lim_{x \to -\infty} F(x),\ F(+\infty) = \lim_{x \to +\infty} F(x).$$

3°. $F(x)$ неперервна ліворуч та існує границя праворуч в будь-якій точці $x \in R^1$.

Доведемо ці властивості.

1°. Нехай $x_1 < x_2$. Тоді $(-\infty; x_2) = (-\infty; x_1) + [x_1; x_2)$.

Звідси за властивістю 2_p аддитивності ймовірнісної міри

$$P((-\infty; x_2) = P((-\infty; x_1)) + P([x_1; x_2)),$$

тобто

$$F(x_2) = F(x_1) + P([x_1; x_2)),\ P([x_1; x_2)) = F(x_2) - F(x_1).$$

Оскільки $P([x_1; x_2)) \ge 0$, то $F(x_2) \ge F(x_1)$.

2°. Нехай $x_{n+1} < x_n$, $\lim\limits_{n \to \infty} x_n = -\infty$. Тоді $(-\infty; x_{n+1}) \subset (-\infty; x_n)$, $\bigcap\limits_{n=1}^{\infty} (-\infty; x_n) = \varnothing$.

За властивостями неперервності ймовірнісної міри

$$0 = P(\varnothing) = P\left(\bigcap_{n=1}^{\infty}(-\infty; x_n)\right) = \lim_{n\to\infty} P((-\infty; x_n)) = \lim_{n\to\infty} F(x_n).$$

Аналогічно, якщо $x_n < x_{n+1}$, $\lim\limits_{n\to\infty} x_n = \infty$, то

$$(-\infty; x_n) \subset (-\infty; x_{n+1}), \quad \bigcup_{n=1}^{\infty}(-\infty; x_n) = \Omega.$$

$$1 = P(\Omega) = P\left(\bigcup_{n=1}^{\infty}(-\infty; x_n)\right) = \lim_{n\to\infty} P((-\infty; x_n)) = \lim_{n\to\infty} F(x_n).$$

3°. Нехай $x_n < x_{n+1}$, $x_n \to x$, $x_n < x$.

Тоді $(-\infty; x_n) \subset (-\infty; x_{n+1})$ і $(-\infty; x_n) = \bigcup\limits_{n=1}^{\infty}(-\infty; x_n)$.

і за властивістю неперервності знизу ймовірнісної міри

$$F(x) = P((-\infty; x)) = P\left(\bigcup_{n=1}^{\infty}(-\infty; x_n)\right) = \lim_{n\to\infty} P((-\infty; x_n)) = \lim_{n\to\infty} F(x_n).$$

Нехай $x_{n+1} < x_n$, $\lim\limits_{n\to\infty} x_n = x$, $x_n > x$.

Тоді

$$(-\infty; x_n) \supset (-\infty; x_{n+1}), \quad \bigcap_{n=1}^{\infty}(-\infty; x_n) = (-\infty; x]$$

і за властивістю неперервності зверху ймовірнісної міри

$$P((-\infty; x]) = P\left(\bigcap_{n=1}^{\infty}(-\infty; x_n)\right) = \lim_{n\to\infty} P((-\infty; x_n)) = \lim_{n\to\infty} F(x_n) =$$

$$= F(x+0) = P((-\infty; x) + \{x\}) = P((-\infty; x)) + P(\{x\}) = F(x) + P(\{x\}).$$

Звідси випливає, що

$$P(\{x\}) = F(x+0) - F(x),$$

оскільки $\{x\} = (-\infty; x] - (-\infty; x)$, $(-\infty; x) \subset (-\infty; x]$ і

$$P(\{x\}) = P((-\infty; x]) - P((-\infty; x)) = F(x+0) - F(x).$$

Означення. Функція $F(x)$, стосовно якої задовільняються наведені властивості $1^\circ - 3^\circ$, називається *функцією розподілу ймовірностей* на числовій прямій R^1.

Таким чином, *кожній ймовірнісній мірі на* $(R^1, \mathscr{B}(R^1))$ *можна поставити у відповідність деяку функцію розподілу ймовірностей*.

Виявляється, що виконується й обернене твердження.

Якщо $F(x)$ – деяка функція розподілу ймовірностей на числовій прямій, то на $(R^1, \mathscr{B}(R^1))$ існує, і притому єдина, ймовірнісна міра така, що за будь-яких $-\infty < x_1 \le x_2 < \infty$

$$P([x_1; x_2)) = F(x_2) - F(x_1).$$

Міру P, побудовану за функцією F, називають мірою Лебега-Стільтьєса.

Якщо

$$F(x) = \begin{cases} 0, \text{ коли } x < 0; \\ x, \text{ коли } 0 \le x \le 1; \\ 1, \text{ коли } x > 1, \end{cases}$$

то відповідну ймовірнісну міру $\lambda(A) = P(A)$ називають мірою Лебега на інтервалі $[0;1]$. Очевидно, що $\lambda([a;b)) = b - a$, коли $[a;b) \subset [0;1]$. Звідси видно, що міра Лебега проміжків $[a;b), [a;b], (a;b], (a;b)$, що є частинами $[0;1]$, також дорівнює $b - a$.

Розглянемо систему $\overline{\mathscr{B}}([0;1])$ множин Λ з $[0;1]$ таких, що за кожного Λ можна знайти борелівські множини A і B такі, що $A \subseteq \Lambda \subseteq B$, $\lambda(B \setminus A) = 0$. *Система $\overline{\mathscr{B}}([0;1])$ є σ-алгеброю, яку називають системою лебегових множин відрізка* $[0;1]$. Очевидно, що $\mathscr{B}([0;1]) \subset \overline{\mathscr{B}}([0;1])$. Міру λ, визначену поки що лише на множинах з $\mathscr{B}([0;1])$, можна продовжити на $\overline{\mathscr{B}}([0;1])$, поклавши $\overline{\lambda}(\Lambda) = \lambda(A)$, якщо (порівняйте з §2.2 – §2.4)

$$\Lambda \in \overline{\mathscr{B}}([0;1]), A \subset \Lambda \subset B, A \in \overline{\mathscr{B}}([0;1]), B \in \overline{\mathscr{B}}([0;1]), \lambda(B \setminus A) = 0.$$

Визначена так функція від множин з $\overline{\mathscr{B}}([0;1])$ буде ймовірнісною мірою на $([0;1], \overline{\mathscr{B}}([0;1]))$. *Її називають також мірою Лебега* (*на системі лебегових множин відрізка* $[0;1]$).

Відповідність між імовірнісними мірами P на $(R^1, \mathscr{B}(R^1))$ і функціями розподілу ймовірностей, яка визначається за рівністю

$$P([x_1; x_2)) = F(x_2) - F(x_1), \tag{3.2.1}$$

дає змогу конструювати на $(R^1, \mathscr{B}(R^1))$ різні ймовірнісні міри, задаючи відповідні функції розподілу ймовірностей (порівняйте з §2.7, приклади 2.7.1, 2.7.2).

Якщо міра P зосереджена в точках $x_1, x_2, \dots$, тобто

$$P(\{x_k\}) = F(x_k + 0) - F(x_k) = \Delta F(x_k) \text{ і } \sum_k P(\{x_k\}) = 1,$$

тоді говорять, що задано *дискретний поточковий розподіл імовірностей* на множині $\{x_1, x_2, ...\}$, а відповідну *міру і функцію розподілу ймовірностей* $F(x)$ називають *дискретними*.

Міра, задана на $(R^1, \mathscr{B}(R^1))$, називається *абсолютно-неперервною*, якщо відповідну функцію розподілу ймовірностей $F(x)$ можна подати у вигляді

$$F(x) = \int_{-\infty}^{x} f(t)\, dt, \tag{3.2.2}$$

де $f(t)$ – деяка невід'ємна функція, $t \in R^1$ (в загальному випадку мається на увазі інтеграл Лебега). Таку *функцію* $F(x)$ *розподілу* ймовірностей також називають *абсолютно неперервною*.

Означення. Невід'ємна функція $f(x)$, $x \in R^1$, через яку визначається $F(x)$ за рівністю (3.2.2), називається *щільністю розподілу ймовірностей*.

Відмітимо основні властивості щільності розподілу ймовірностей:

1. $f(x) \geq 0$;
2. $\int_{-\infty}^{\infty} f(x) dx = 1$.

Зрозуміло, що за будь-якою невід'ємною функцією $f(x)$ такою, що $\int_{-\infty}^{\infty} f(x)\, dx = 1$, можна визначити деяку функцію розподілу ймовірностей:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt.$$

Міри на $(R^1, \mathscr{B}(R^1))$, функції $F(x)$ розподілу яких неперервні і $F'(x) = 0$ майже всюди (множина точок зростання функції $F(x)$ є множиною нульової міри Лебега), називаються *сингулярними*. Прикладом такої функції розподілу є функція, яку будують так.

Спочатку беруть

$$F_0(x) = \begin{cases} 0, & \text{коли} \quad x \leq 0, \\ x, & \text{коли} \quad 0 \leq x \leq 1, \\ 1, & \text{коли} \quad 1 \leq x. \end{cases}$$

Далі відрізок $[0; 1]$ ділять на три рівні частини, вважаючи, що

$$F_1(x)=\begin{cases}0, \text{ коли } x\le 0;\\ \dfrac{1}{2}, \text{ коли } x\in\left[\dfrac{1}{3};\,\dfrac{2}{3}\right];\\ 1, \text{ коли } x\ge 1,\end{cases}$$

а на відрізках $\left[\,0;\,\dfrac{1}{3}\,\right], \left[\,\dfrac{2}{3};\,1\,\right]$ функція лінійно зростає і неперервна на $[0;1]$ (Рис. 3.2.1).

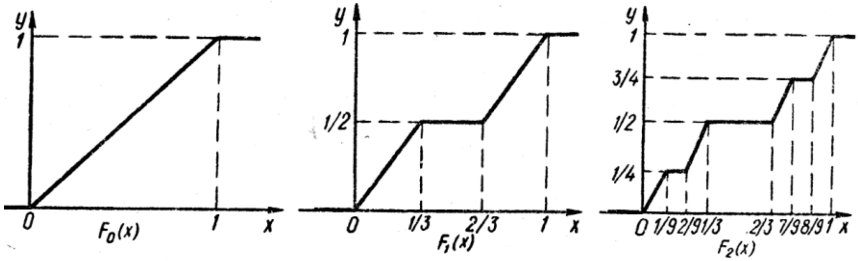


Рис. 3.2.1

Далі відрізки $\left[\,0;\,\dfrac{1}{3}\,\right], \left[\,\dfrac{2}{3};\,1\,\right]$, на яких функція лінійно зростає, також ділять на три рівних частини, вважаючи, що

$$F_2(x)=\begin{cases}0, \quad \text{коли} \quad x\le 0;\\ \dfrac{1}{4}, \quad \text{коли } x\in\left[\dfrac{1}{9};\,\dfrac{2}{9}\right];\\ \dfrac{1}{2}, \quad \text{коли } x\in\left[\dfrac{1}{3};\,\dfrac{2}{3}\right];\\ \dfrac{3}{4}, \quad \text{коли } x\in\left[\dfrac{7}{9};\,\dfrac{8}{9}\right];\\ 1, \quad \text{коли } \; x\ge 1,\end{cases}$$

і що функція лінійно зростає в інших точках і неперервна на $[0;1]$ (Рис. 3.2.1). Відрізки, на яких функція лінійно зростає, знову ділять на три рівні частини і аналогічно будують функцію $F_3(x)$ і т.д. Цей процес продовжується нескінченно. Послідовність так одержаних неперервних функцій збігається до так званої канторової функції $F(x)$, множина точок зростання якої є множиною нульової лебегової міри (x називається точкою зростання, якщо $F(x+\varepsilon)-F(x-\varepsilon)>0$ за будь-якого $\varepsilon>0$).

Справді, загальна довжина інтервалів, на яких функція $F(x)$ набуває сталих значень, дорівнює

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\ldots=\frac{\frac{1}{3}}{1-\frac{2}{3}}=1.$$

Якщо через $\mathcal{N}$ позначити множину точок зростання канторової функції $F(x)$, то з останньої рівності випливає, що міра Лебега $\lambda(\mathcal{N})=0$. Якщо μ міра, що відповідає канторовій функції $F(x)$, то $\mu(\mathcal{N})=1$. В такому разі говорять, *що міра μ сингулярна щодо міри Лебега λ*.

Множину $\mathcal{N}$ називають канторовою досконалою множиною.

Виявляється, що довільну функцію розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ можна подати у вигляді

$$p_1F_1+p_2F_2+p_3F_3,$$

де F_1 – дискретна, F_2 – абсолютно неперервна, F_3 – сингулярна функції розподілу ймовірностей; $p_1+p_2+p_3=1,\ p_i\ge 0,$ $i=1,\ 2,\ 3$.

Вправи для самостійного виконання

3.2.1. Перевірити, чи правильні твердження:

1. Якщо функція $F(x),\ x\in(-\infty;+\infty)$, є неспадною, неперервною і такою, що $\lim\limits_{x\to-\infty}F(x)=0$, а $\lim\limits_{x\to+\infty}F(x)=1$, то вона є функцією розподілу ймовірностей на числовій прямій.
2. Твердження, обернене до 1, є правильним.
3. У функції розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ існують скінченні ліва та права границі в будь-якій точці.
4. Множина точок розриву функції розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ не більш ніж зчисленна.
5. За кожною функцією розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ визначається єдина міра Лебега-Стільтьєса.
6. Якщо $\lambda=P(A)$ – міра Лебега, то вона є мірою Лебега-Стільтьєса.
7. Твердження, обернене до 6, є правильним.
8. Якщо A – лебегова множина, то вона є борелівською множиною.
9. Твердження, обернене до 8, є правильним.
10. Кожну лебегову множину можна як завгодно добре наблизити борелівськими множинами.
11. Кожна функція розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ є або дискретною, або абсолютно неперервною, або сингулярною.
12. За кожної функції розподілу ймовірностей на $(R^1,\ \mathscr{B}(R^1))$ існує щільність розподілу.

3.2.2. Побудувати функцію $F(x)$ розподілу ймовірностей на $(R^1, \mathscr{B}(R^1))$, за якої $F(0)=0,\ F(1)=\frac{1}{2}$ і ця функція : 1) розривна лише у точках 0 та 1; 2) скрізь неперервна; 3) абсолютно неперервна.

3.2.3. Побудувати функцію $F(x)$ розподілу ймовірностей на $(R^1, \mathscr{B}(R^1))$, за якої $F(-1)=0,\ F(x)=ax+b$, коли $x \in (-1;1)$ і $F(x)=1$, коли $x \geq 1$. Якими повинні бути a і b, щоб ця функція була неперервною на R^1?

3.2.4 Побудувати абсолютно неперервну функцію $F(x)$ розподілу ймовірностей на $(R^1, \mathscr{B}(R^1))$, за якої щільність розподілу набуває значень $f(0)=1,\ f(1)=2$, а в інших точках ця функція є сталою, відмінною від 1 та 2. Якою може бути ця стала? В яких точках $F'(x)=f(x)$? З'ясувати, чи існує похідна $F'(x)$ в точках $x=0$ та $x=1$?

3.3. Многовимірні розподіли ймовірнісних мір в координатному просторі R^n

Нехай в R^n задано простір елементарних подій $\Omega \subset R^n$, простір подій S, породжений за поділом множини Ω на підмножини H_i, $i \in \overline{1,k}$, такі, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$, тобто $S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$, і на S задана деяка ймовірнісна міра $P(A)$, $i \in \overline{1,k}$, така, що $P(H_i) \geq 0$, $\sum_{i=1}^{k} P(H_i) = 1$, тобто заданий розподіл ймовірностей на множині Ω за підмножинами H_i, $i \in \overline{1,k}$, $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$ (зокрема це можуть бути статистичні ймовірності $P_n^*(H_i)$).

Тоді за довільного $A \in S$ буде

$$P(A) = P(\bigcup_{i \in I} H_i) = \sum_{i \in I} P(H_i), \quad I \subset \{1, 2, ..., k\}.$$

Нехай $S \subset \tilde{S}$, де $\tilde{S}$ – сукупність множин із R^n, стосовно якої задовільняються вимоги 1_s-3_s і на якій задано деяку міру $m(G)$, $G \in \tilde{S}$, за якої $0 < m(\Omega) < \infty$ (див. §2.2).

Введемо до розгляду функцію

$$f(E) = \frac{P(H_i)}{m(H_i)}, \ E \in H_i, \ i \in \overline{1,k}, \tag{3.3.1}$$

визначену за всіх $E \in \Omega$.

Таку функцію природно назвати усередненою або середньою щільністю відносно міри m розподілу ймовірностей (ймовірнісної міри P) на множині Ω за підмножинами H_i.

Із виразу (3.3.1) отримаємо

$$P(H_i) = f(E) m(H_i), \ i \in \overline{1,k}.$$

Звідси в разі $A = \bigcup_{i \in I} H_i$, $I \subset \{1, 2, ..., k\}$, буде

$$P(A) = P(\bigcup_{i \in I} H_i) = \sum_{i \in I} P(H_i) = \sum_{i \in I, E \in H_i} f(E) m(H_i).$$

Підкреслимо, що ймовірнісна міра $P(A)$ визначена тільки на множинах A із сукупності S підмножин множини Ω.

Якщо $G \in \tilde{S} \supset S$ – довільна вимірна за мірою m множина із R^2, $G \subset R^2$, тоді природно ввести нову ймовірнісну міру $\tilde{P}$ на сукупності $\widetilde{S} \supset S$ вимірних за мірою m множин $G \subset R^2$, поклавши

$$\widetilde{P}(G) = \widetilde{P}(G \bigcap \bigcup_{i=1}^{k} H_i) = \widetilde{P}(\bigcup_{i=1}^{k} G \bigcap H_i) =$$

$$= \sum_{i=1}^{k} \widetilde{P}(G \bigcap H_i) = \sum_{i=1}^{k} \frac{m(G \bigcap H_i)}{m(H_i)} P(H_i) =$$

$$= \sum_{E \in H_i, i=1}^{k} \frac{m(G \bigcap H_i)}{m(H_i)} f(E) m(H_i) = \sum_{E \in H_i, i=1}^{k} f(E) m(G \bigcap H_i).$$

де E – точка із простору R^2.

В такий спосіб імовірнісна міра $P(A)$, $A \in S$, задана на сукупності $S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$, підмножин множини Ω, породженій за поділом множини Ω на підмножини H_i, $i \in \overline{1, k}$, такі, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$, поширюється (продовжується) на сукупність $\widetilde{S} \supset S$ множин, вимірних за мірою m. Очевидно, в разі $A \in S$ має місце рівність $\widetilde{P}(A) = P(A)$.

Вираз $\sum_{E \in H_i, i=1}^{k} f(E) m(G \bigcap H_i)$ природно називати інтегралом від функції $f(E)$ на множині G за мірою m і позначати

$$\int_G f(E) m(dE),$$

де E – точка із простору R^2 з координатами (x, y), dE – елемент площі $(dxdy)$, $m(dE)$ – міра (площа) елемента площі dE.

Очевидно: 1. $f(E) \geq 0$, $E \in \Omega$;

2. $\int_\Omega f(E) m(dE) = 1$.

Інший спосіб задання міри $\widetilde{P}$ на сукупності $\widetilde{S} \supset S$ може бути таким (див. §2.2). Нехай $G \in \widetilde{S}$, $G_* = \bigcup_{\cup H_i \subset G \cap \Omega} (\cup H_i)$ – найширше об'єднання об'єднань $\cup H_i$ множин H_i таке, що $\cup H_i \subset G \bigcap \Omega$, тобто об'єднання всіх об'єднань $\cup H_i$ таких, що $\cup H_i \subset G \bigcap \Omega$, $G^* = \bigcap_{G \cap \Omega \subset \cup H_i} (\cup H_i)$ – перетин всіх об'єднань $\cup H_i$ таких, що $G \bigcap \Omega \subset \cup H_i$, тобто найвужче об'єднання $\cup H_i$ множин H_i таких, що $G \bigcap \Omega \subset \cup H_i$. Очевидно $G_* \in S$, $G^* \in S$, $G_* \subset G^*$, $G^* \setminus G_* \in S$.

Покладемо

$$\widetilde{P}(G) = P(G_*) + \frac{m(G \bigcap (G^* \setminus G_*))}{m(G^* \setminus G_*)} P(G^* \setminus G_*) . \qquad (3.3.2)$$

Очевидно

$$P(G_*) \le \widetilde{P}(G) \le P(G^*) ,$$

оскільки

$$0 \le \frac{m(G \bigcap (G^* \setminus G_*))}{m(G^* \setminus G_*)} \le 1 , \ P(G^* \setminus G_*) \ge 0 .$$

На практиці часто покладають $\widetilde{P}(G) = P(G_*) + \alpha P(G^* \setminus G_*)$, де $\alpha \in [0;1]$ – довільне число із проміжка $[0;1]$, зокрема може бути $\alpha = 0$, $\alpha = \frac{1}{2}$, $\alpha = 1$ і ін. (див. §2.2).

Зауважимо, що коли $G = \bigcup_{i \in I} H_i \in S$, $I \subset \{1, 2, ..., k\}$, тоді $G_* = G^* = G \in S$, $\widetilde{P}(G) = P(G)$.

Так введена на $\widetilde{S}$ міра $\widetilde{P}$ є одним з можливих продовжень ймовірнісної міри P із сукупності S підмножин множини Ω на сукупність $\widetilde{S} \supset S$ множин $G \subset R^2$, вимірних за мірою m. Оскільки стосовно $\widetilde{S}$ задовільняються вимоги 1_s-3_s, а стосовно міри $\widetilde{P}$ задовільняються вимоги 1_p-3_p, то трійка $(R^2, \widetilde{S}, \widetilde{P})$ є ймовірнісним простором.

Аналогічно будується ймовірнісний простір $(R^n, \widetilde{S}, \widetilde{P})$ за заданим ймовірнісним простором (Ω, S, P), де $\Omega \subset R^n$, $S = \{A \,|\, A = \bigcup_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$, $H_i H_j = \varnothing$, коли $i \ne j$, $\bigcup_{i=1}^{k} H_i = \Omega$, $\widetilde{S} \supset S$ – сукупність підмножин із R^n, вимірних за мірою m, і така, стосовно якої задовільняються вимоги 1_s-3_s.

Приклад 3.3.1 Нехай $\Omega \subset R^2$ – квадрат розмірами 8×8 з центром в початку координат і сторонами, паралельними до координатних осей, тобто $\Omega = \{(x, y) \,|\, x \in [-4; 4), y \in [-4; 4)\}$, або, що те саме, $\Omega = [-4; 4) \times [-4; 4)$, поділений на 64 квадрати H_{ij} розмірами 1×1, $i \in \overline{1, 8}$, $j \in \overline{1, 8}$, $H_{ij} \bigcap H_{rk} = \varnothing$, коли $i \ne r$ або $j \ne k$, $\bigcup_{i=1}^{8} \bigcup_{j=1}^{8} H_{ij} = \Omega$ (див. Рис. 3.3.1).

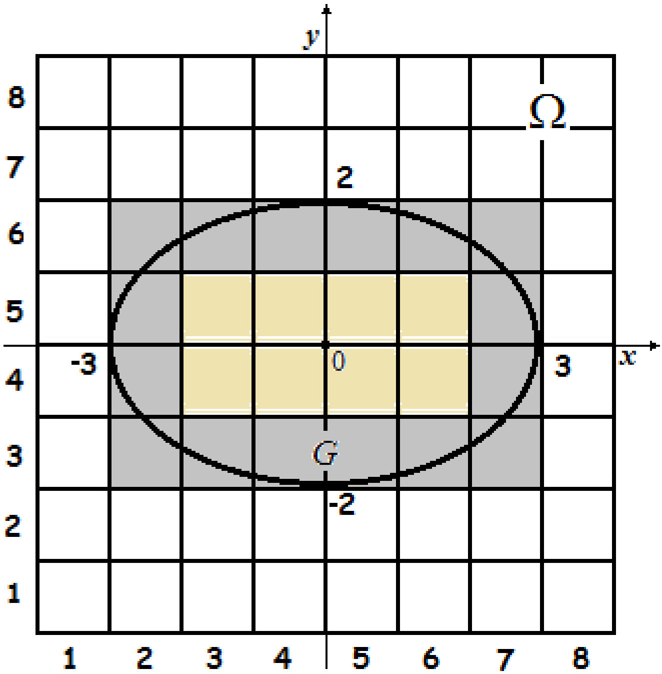


Рис. 3.3.1

Нехай G – множина внутрішніх точок еліпса з півосями $a = 3$, $b = 2$, рівняння якого $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, тобто $G = \{(x, y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1\}$.

Як відомо, площа еліпса G обчислюється за формулою $m(G) = \pi ab$, тому G – вимірна множина, $m(G) = \pi ab$, де міра множини $G \subset R^2$ – її площа. В разі $a = 3$, $b = 2$, одержимо $\pi ab \approx 3.14 \cdot 3 \cdot 2 = 3.14 \cdot 6 = 18.84 \approx 19$. За парою індексів (i, j), $i \in \overline{1, 8}$, $j \in \overline{1, 8}$, визначається місце розташування квадрата H_{ij} в квадраті Ω, причому квадрати H_{ij} нумеруються від найнижчого ліворуч, якому відповідають індекси $i = 1$, $j = 1$, до найвищого праворуч, якому відповідають індекси $i = 8$, $j = 8$. Якщо зафіксувати індекс j і збільшувати індекс i, то відбуватиметься перехід знизу вгору до квадратів, розташованих вище, але які залишаються в колонці квадратів із зафіксованим індексом j. Якщо ж зафіксувати індекс i і збільшувати індекс j, то відбуватиметься перехід зліва вправо до квадратів, розташованих правіше, але які залишаються в рядкові квадратів із зафіксованим індексом i.

Таким чином індекс i, $i \in \overline{1,\ 8}$ – це номер рядка квадратів H_{ij}, $j \in \overline{1,\ 8}$, індекс j, $j \in \overline{1,\ 8}$ – номер стовпчика квадратів H_{ij}, $i \in \overline{1,\ 8}$.

Найширшим об'єднанням $\bigcup H_{ij}$ множин H_{ij} таким, що $\bigcup H_{ij} \subset G \bigcap \Omega = G$ буде об'єднання (див. Рис. 3.3.1)

$$H_{4,3} \bigcup H_{4,4} \bigcup H_{4,5} \bigcup H_{4,6} \bigcup H_{5,3} \bigcup H_{5,4} \bigcup H_{5,5} \bigcup H_{5,6},$$

тобто $G_* = \bigcup\limits_{i=4}^{5} \bigcup\limits_{j=3}^{6} H_{ij}$. Найвужчим об'єднанням $\bigcup H_{ij}$ множин H_{ij} таким, що $G \bigcap \Omega \subset \bigcup H_{ij}$ в розглядуваному прикладі буде об'єднання (див. Рис. 3.3.1)

$$H_{3,2} \bigcup H_{3,3} \bigcup H_{3,4} \bigcup H_{3,5} \bigcup H_{3,6} \bigcup H_{3,7} \bigcup$$
$$\bigcup H_{4,2} \bigcup H_{4,3} \bigcup H_{4,4} \bigcup H_{4,5} \bigcup H_{4,6} \bigcup H_{4,7} \bigcup$$
$$\bigcup H_{5,2} \bigcup H_{5,3} \bigcup H_{5,4} \bigcup H_{5,5} \bigcup H_{5,6} \bigcup H_{5,7} \bigcup$$
$$\bigcup H_{6,2} \bigcup H_{6,3} \bigcup H_{6,4} \bigcup H_{6,5} \bigcup H_{6,6} \bigcup H_{6,7},$$

тобто

$$G^* = \bigcup_{i=3}^{6} \bigcup_{j=2}^{7} H_{ij}.$$

Якщо задані імовірності $P(H_{i,j})$, $i \in \overline{1,\ 8}$, $j \in \overline{1,\ 8}$, всі однакові між собою, тобто $P(H_{i,j}) = \frac{1}{64}$, $i \in \overline{1,\ 8}$, $j \in \overline{1,\ 8}$, тоді буде $P(G_*) = \frac{8}{64}$, $P(G^*) = \frac{24}{64}$.

Стосовно $\tilde{P}(G) \approx P(G_*) + \alpha(P(G^*) - P(G_*))$, де $\alpha \in [0;1]$, $G \in \tilde{S}$, часто покладають $\alpha = \frac{1}{2}$, якщо різниця $P(G^*) - P(G_*)$ не дуже велика (див. §2.2).

В розглянутому прикладі, поклавши $\alpha = \frac{1}{2}$, одержимо наближене значення ймовірнісної міри $\tilde{P}(G)$:

$$\tilde{P}(G) \approx P(G_*) + \frac{1}{2}(P(G^*) - P(G_*)) = \frac{8}{64} + \frac{1}{2}(\frac{24}{64} - \frac{8}{64}) = \frac{16}{64}.$$

Можна показати, що коли сторони квадратів H_{ij} зменшувати в кілька разів так, що кожний із квадратів H_{ij} буде поділено на r

рівних між собою квадратів H_{ijk}, $k \in \overline{1, r}$, крім того покласти $P(H_{ijk}) = \frac{1}{r} P(H_{ij})$ за всіх $k \in \overline{1, r}$,

$$G_{*r} = \bigcup_{\bigcup H_{ijk} \subset G \cap \Omega} \bigcup H_{ijk}, \; G_r^* = \bigcap_{G \cap \Omega \subset \bigcup H_{ijk}} \bigcup H_{ijk},$$

$$S_r = \{A \mid A = \bigcup_{i \in I,\, j \in J,\, k \in K} H_{ijk}, \; I \subset \{1, 2, ..., 8\},$$

$$J \subset \{1, 2, ..., 8\}, \; K \subset \{1, 2, ..., 8\},$$

тоді, оскільки множина G опукла і довжина її контура обмежена, буде мати місце рівність

$$\lim_{r \to \infty} P(G_{*r}) = \lim_{r \to \infty} P(G_r^*) = \tilde{P}(G),$$

тобто множина G буде вимірною за так введеною мірою $\tilde{P}$ (див. §2.2), а

$$P_r(G) = P(G_{*r}) + \alpha P(G_r^* \setminus G_{*r}), \; \alpha \in [\,0\,;1\,]$$

за довільного $r \geq 1$ буде наближеним значенням числа $\tilde{P}(G)$, і точність такого наближення буде тим більша, чим більшим буде число r. Очевидно, за довільного $r \geq 1$ має місце нерівність

$$P(G_{*r}) \leq \tilde{P}(G) \leq P(G_r^*).$$

Приклад 3.3.2 Нехай множина Ω задана в полярних координатах – круг радіуса 5 з центром в початку координат (Рис. 3.3.2)

$$\Omega = \{(\rho, \varphi) \mid \rho \in [0; 5), \varphi \in [0; 2\pi)\}, H_0 = \{(\rho, \phi) \mid \rho \in [0; 1), \phi \in [0; 2\pi)\},$$

$$H_i = \{(\rho, \varphi) \mid \rho \in [1; 2), \varphi \in [\frac{2\pi}{8}[i-1); \frac{2\pi}{8} i)\}, \; i \in \overline{1, 8},$$

$$H_i = \{(\rho, \varphi) \mid \rho \in [2; 3), \varphi \in [\frac{2\pi}{8}[i-8); \frac{2\pi}{8}(i-8+1)\}, \; i \in \overline{9, 16},$$

$$H_i = \{(\rho, \varphi) \mid \rho \in [3; 4), \varphi \in [\frac{2\pi}{8}[i-16); \frac{2\pi}{8}(i-16+1)\}, \; i \in \overline{17, 24},$$

$$H_i = \{(\rho, \varphi) \mid \rho \in [4; 5), \varphi \in [\frac{2\pi}{8}[i-24); \frac{2\pi}{8}(i-24+1)\}, \; i \in \overline{25, 32},$$

$$S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \overline{0, 32}\},$$

$$G = \{(x, y) \mid x \in [-3; 3], y \in [-3; 3]\} = [-3; 3] \times [-3; 3] \text{ -}$$

квадрат з центром в початку координат і сторонами, паралельними до координатних осей, причому сторони квадрата G дотикаються до кола радіуса $\rho = 3$.

В даному разі $G \overline{\in} S$, $G \in \tilde{S} \supset S$, де $\tilde{S}$ – сукупність підмножин із R^2, вимірних за мірою m, $m(G)$ – площа множини точок G (тобто $\tilde{S}$ сукупність квадровних множин із R^2).

Оскільки відстань від центра до вершини квадрата G дорівнює $\sqrt{3^2+3^2} = \sqrt{18} \approx 4.25 > 4$, то перетини квадрата G з підмножинами H_i, $i \in \overline{25,\ 32}$, непорожні (див. Рис. 3.3.2), тобто $G \bigcap H_i \neq \varnothing$, коли $i \in \overline{25,\ 32}$. Тому $G_* = \bigcup\limits_{i=0}^{16} H_i$, $G^* = \bigcup\limits_{i=0}^{32} H_i$.

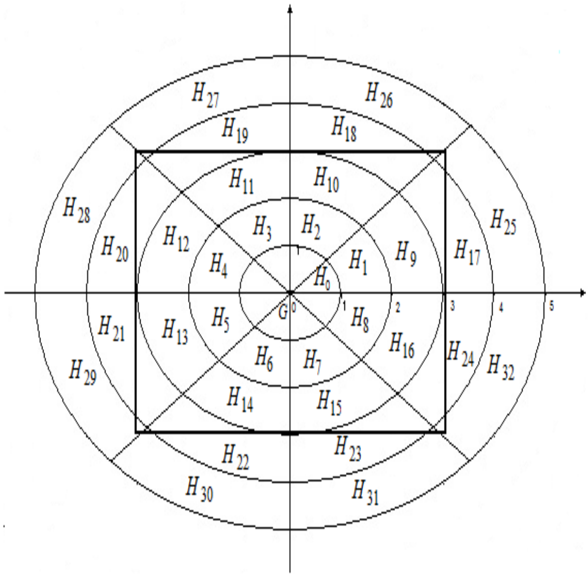


Рис. 3.3.2

Якщо ймовірнісна міра (ймовірність) P розподілена на Ω слідуючим чином:

$P(H_0) = 0.40$; $P(H_i) = \dfrac{0.32}{8}$, коли $i \in \overline{1,\ 8}$; $P(H_i) = \dfrac{0.16}{8}$, коли $i \in \overline{9,\ 16}$; $P(H_i) = \dfrac{0.08}{8}$, коли $i \in \overline{17,\ 24}$; $P(H_i) = \dfrac{0.04}{8}$, коли $i \in \overline{25,\ 32}$, тоді буде $P(G_*) = P(\bigcup\limits_{i=0}^{16} H_i) = \sum\limits_{i=0}^{16} P(H_i) = 0.88$, $P(G^*) = P(\bigcup\limits_{i=0}^{32} H_i) = \sum\limits_{i=0}^{32} P(H_i) = 1.00$. Якщо покласти

$\widetilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, коли $\alpha = \dfrac{1}{2}$, $G \in \widetilde{S}$, тоді одержимо $\widetilde{P}(G) = 0.88 + \dfrac{1}{2}(1.00 - 0.88) = 0.94$.

Зауважимо, що коли в останньому прикладі 3.3.2 подрібнювати множини H_i, $i \in \overline{0,32}$, лише зменшуючи в r разів кути між радіусами векторами, якими обмежуються H_i, і не зменшуючи віддалі між колами, якими обмежуються H_i, так, що

$$H_i = \bigcup_{j=1}^{r} H_{ij}\,,\ P(H_{ij}) = \frac{1}{r} P(H_i) \underset{r\to\infty}{\to} 0\,,$$

$$G_{*r} = \bigcup_{\bigcup H_{ij} \subset G \bigcap \Omega} \bigcup H_{ij}\,,\ G_r^* = \bigcap_{G \bigcap \Omega \subset \bigcup H_{ij}} \bigcup H_{ij}\,,$$

тоді буде

$$\lim_{r\to\infty} P(G_{*r}) \neq \lim_{r\to\infty} P(G_r^*),$$

і тому за такого способу визначення ймовірнісної міри множина G виявляється невимірною.

Якщо ж H_i подрібнювати так, щоб найбільший з діаметрів $d(H_{ij})$ множин H_{ij}, $j \in \overline{1,\, r}$, прямував до нуля, коли $r \to \infty$, і разом з тим прямувала до нуля ймовірнісна міра $\max\limits_{1 \le j \le r} P(H_{ij})$, де $d(H_{ij}) = \max\limits_{\substack{M_1 \in H_{ij} \\ M_2 \in H_{ij}}} \rho(M_1, M_2)$, $\rho(M_1, M_2)$ – віддаль між точками M_1 і M_2, $M_1 \in H_{ij}$, $M_2 \in H_{ij}$, тобто $d(H_{ij})$ – найбільша відстань між всеможливими парами M_1, M_2 точок із H_{ij}, тоді буде

$$\lim_{r\to\infty} P(G_{*r}) = \lim_{r\to\infty} P(G_r^*),$$

і за так введеною мірою $\widetilde{P}(G) = \lim\limits_{r\to\infty} P(G_{*r}) = \lim\limits_{r\to\infty} P(G_r^*)$ множина G буде вимірною. В такому разі за довільного r наближено можна покласти $\widetilde{P}(G) \approx P(G_{*r}) + \alpha P(G_r^* \setminus G_{*r})$.

Нагадаємо, що міру $\widetilde{P}$ на $\widetilde{S}$ за мірою P, заданою на S, можна визначити і в інший спосіб, поклавши

$$\widetilde{P}(G) = \widetilde{P}(G \bigcap \Omega) = \widetilde{P}(G \bigcap \bigcup_{i=1}^{k} H_i) = \widetilde{P}(\bigcup_{i=1}^{k} G \bigcap H_i) =$$

$$= \sum_{i=1}^{k} \widetilde{P}(G \bigcap H_i) = \sum_{i=1}^{k} \frac{m(G \bigcap H_i)}{m(H_i)} P(H_i) =$$

$$= \sum_{i=1, E \in H_i}^{k} f(E) m(G \cap H_i) = \int_G f(E) m(dE), \qquad (3.3.3)$$

де $f(E) = \dfrac{P(H_i)}{m(H_i)}$, $E \in H_i$, $i \in \overline{1,k}$ – усереднена відносно міри m щільність розподілу ймовірностей на множині Ω за підмножинами H_i, $i \in \overline{1,k}$.

Зауважимо, що не обов'язково повинно бути $G \subset \Omega$, тобто не обов'язково множина G повинна бути підмножиною множини Ω.

Приклад 3.3.3 Нехай

$$\Omega = \{(x, y) \mid x \in [-2; 2), y \in [-2; 2)\} = [-2; 2) \times [-2; 2),$$

квадрат з вершинами $(-2, -2)$, $(2, -2)$, $(2, 2)$, $(-2, 2)$ (Рис. 3.3.3),

$$H_{ij} = \{(x, y) \mid x \in [-2 + (i-1); -2 + i), y \in [-2 + (j-1); -2 + j)\},$$

$$i \in \overline{1, 4}, \ j \in \overline{1, 4}, \ \Omega = \bigcup_{i=1}^{4} \bigcup_{j=1}^{4} H_{ij} .$$

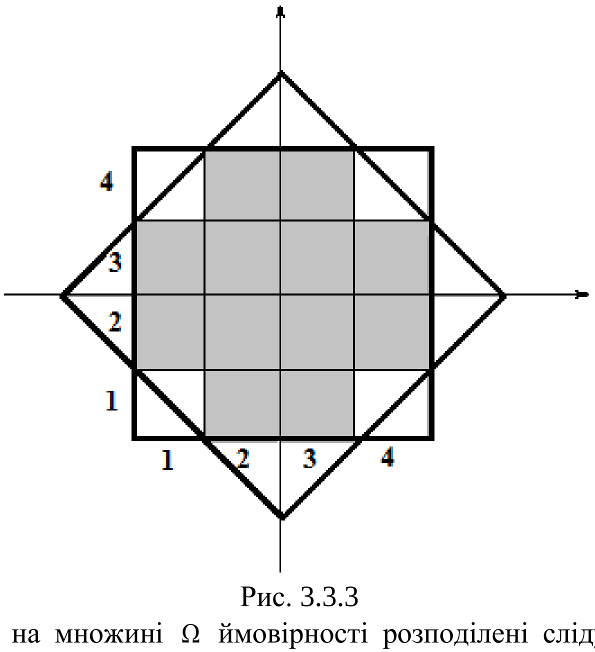


Рис. 3.3.3

Нехай на множині Ω ймовірності розподілені слідуючим чином:

$P(H_{ij}) = 0.16$, коли $i \in \overline{2, 3}$, $j \in \overline{2, 3}$;

$P(H_{ij}) = 0.03$, коли $i = 1$, $j \in \overline{1, 4}$ або $i = 4$, $j \in \overline{1, 4}$;

$P(H_{ij}) = 0.03$, коли $j = 1$, $i \in \overline{1,\ 4}$ або $j = 4$, $i \in \overline{1,\ 4}$.

Нехай $G = \{(x, y) \,|\, |x + y| < 3\} \bigcap \{(x, y) \,|\, |-x + y| < 3\}$ – квадрат із вершинами $(0,\ -3)$, $(3,\ 0)$, $(0,\ 3)$, $(-3,\ 0)$ (Рис. 3.3.3).

Тоді

$$G_* = \bigcup_{\cup H_{ij} \subset G \cap \Omega} (\cup H_{ij}) = H_{1,2} + H_{1,3} + H_{2,1} + H_{2,2} + H_{2,3} + H_{2,4} +$$

$$+ H_{3,1} + H_{3,2} + H_{3,3} + H_{3,4} + H_{4,2} + H_{4,3},$$

$$G^* = \bigcap_{G \cap \Omega \subset \cup H_{ij}} (\cup H_{ij}) = \bigcup_{i=1}^{4} \bigcup_{j=1}^{4} H_{ij} = \Omega .$$

Очевидно (див. Рис. 3.3.3)

$P(G_*) = 4 \cdot 0.16 + 8 \cdot 0.03 = 0.88$, $P(G^*) = 4 \cdot 0.16 + 12 \cdot 0.03 = 1.00$.

Якщо покласти $\tilde{P}(G) = P(G_*) + \frac{1}{2}(P(G^*) - P(G_*))$, тоді одержимо

$$\tilde{P}(G) = 0.88 + \frac{1}{2} 0.12 = 0.94 .$$

Якщо покласти (див. формулу (3.3.3))

$$\tilde{P}(G) = \sum_{i=1}^{4} \sum_{j=1}^{4} \frac{m(G \cap H_{ij})}{m(H_{ij})} P(H_{ij}) = \int_G f(E) m(dE),$$

тоді одержимо $\tilde{P}(G) = 1.00 - 4 \cdot \frac{0.03}{2} = 0.94$.

Оскільки $m(H_{ij}) = 1 \cdot 1 = 1$, $i \in \overline{1,\ 4}$, $j \in \overline{1,\ 4}$, то усереднена щільність $f(E)$ розподілу ймовірностей на множині Ω в розглядуваному прикладі набуває вигляду

$$f(x, y) = \begin{cases} 0.16, \text{ коли } (x, y) \in H_{i,j},\ i \in \overline{2,\ 3},\ j \in \overline{2,\ 3}; \\ 0.03, \text{ коли } (x, y) \in H_{i,j},\ i \in \{1,\ 4\},\ j \in \{1, 2, 3, 4\} \\ \qquad\qquad \text{або } j \in \{1, 4\}, i \in \{1, 2, 3, 4\}; \\ 0, \text{ коли } (x, y) \overline{\in} \Omega. \end{cases}$$

Вправи для самостійного виконання

3.3.1. Перевірити, чи правильні твердження:

1. За довільної множини G можна визначити G_* та G^*.

2. За довільної множини $G \subset R^2$ можна визначити ймовірнісну міру $\tilde{P}(G)$.

3. За довільної множини $G \subset R^2$ можна визначити ймовірнісну міру $\tilde{P}(G)$ за формулою 3.3.3.

4. Існують множини $G \subset R^2$ такі, що $G_* = G^*$.

5. Існують множини $G \subset R^2$, $G \in \widetilde{S}$, такі, що $\widetilde{P}(G) = P(G)$.

3.3.2. Виконати вправи:

1. В R^2 задано Ω – квадрат $[-5;5]\times[-5;5]$, тобто $\Omega = \{(x, y) \mid x \in [-5;5), y \in [-5;5)\}$. H_{ij} – квадрати розмірами 0.5×0.5, тобто

$$H_{ij} = \{(x, y) \mid x \in [-5+0.5(i-1); -5+0.5i), y \in [-5+0.5(j-1); -5+0.5j)\}$$
$$i \in \overline{1,20}, \ j \in \overline{1,20}$$

тобто квадрат Ω поділений на 400 квадратів H_{ij}, $i \in \overline{1,\ 20}$, $j \in \overline{1,\ 20}$.

$G = \{(x, y) \mid x^2 + y^2 \le 16\}$ – круг радіуса 4 з центром в початку координат. Знайти G_* та G^*, а також $P(G_*)$ та $P(G^*)$, коли

$$P(H_{ij}) = \frac{0.60}{144}, \text{ коли } i \in \overline{5,\ 16}, \ j \in \overline{5,\ 16},$$

$$P(H_{ij}) = \frac{0.40}{256}, \text{ коли } i \in \{1,2,3,4\} \cup \{17,18,19,20\}, \ j \in \overline{1,\ 20},$$
$$i \in \overline{5,\ 16}, \ j \in \{1,2,3,4\} \cup \{17,18,19,20\}.$$

Визначити $\widetilde{P}(G)$, поклавши $\widetilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha = \dfrac{1}{2}$.

Як зміняться $P(G_*)$ та $P(G^*)$, $\widetilde{P}(G)$, якщо:

а) довжини сторін квадратів H_{ij} збільшити вдвічі, поклавши

$H_{ij} = \{(x, y) \mid x \in [-5+1(i-1); -5+1\cdot i), y \in [-5+1(j-1); -5+1\cdot j)\}$, $i \in \overline{1,\ 10}$, $j \in \overline{1,\ 10}$.

б) довжини сторін квадратів H_{ij} зменшити вдвічі, поклавши

$H_{ij} = \{(x, y) \mid x \in [-5+0.25(i-1); -5+0.25i), y \in [-5+0.25(j-1); -5+0.25\cdot j)\}$, $i \in \overline{1,\ 40}$, $j \in \overline{1,\ 40}$.

в) довжини сторін квадратів H_{ij} зменшити в 5 разів, поклавши

$H_{ij} = \{(x, y) \mid x \in [-5+0.1(i-1); -5+0.1i), y \in [-5+0.1(j-1); -5+0.1j)\}$, $i \in \overline{1,\ 100}$, $j \in \overline{1,\ 100}$.

2. За даними прикладу 3.3.2 визначити міру $\widetilde{P}(G)$ за формулою (3.3.3).

3.4. Функція розподілу ймовірностей в многовимірному координатному просторі

Нехай задано ймовірнісний простір (Ω, S, P), де $\Omega \subset R^n$ – простір елементарних подій, S – простір подій, породжений за системою підмножин H_i, $i \in \overline{1, k}$, множини Ω таких, що $H_i H_j = \varnothing$, коли $i \neq j$, $\bigcup_{i=1}^{k} H_i = \Omega$, тобто

$$S = \{A \big|\ A = \bigcup_{i \in I} H_i,\ I \subset \{1,\ 2,\ ...,\ k\}\}.$$

Нехай $x = (x_1, x_2, ..., x_n) \in R^n$, $\tilde{\Omega} = R^n$,

$$G = (-\infty;\, x_1) \times (-\infty;\, x_2) \times ... \times (-\infty;\, x_n) \subset R^n,$$

або (в скороченому запису) $G = (-\infty;\, x\ ) \subset R^n$, $G \in \tilde{S} = \mathscr{B}(R^n)$, $\tilde{S} \supset S$ – сукупність підмножин із R^n, вимірних за мірою m (об'ємом n-вимірних фігур, див. §2.2).

Покладемо

$$F(x_1, x_2, ..., x_n) = P(G_*) + \alpha(P(G^*) - P(G_*)) = \tilde{P}(G), \qquad (3.4.1)$$

де $\alpha \in [0; 1]$, $G_* = \bigcup_{\bigcup H_i \subset G \bigcap \Omega} (\bigcup H_i)$ – об'єднання всіх об'єднань $\bigcup H_i$ таких, що $\bigcup H_i \subset G \bigcap \Omega$, $G^* = \bigcap_{G \bigcap \Omega \subset \bigcup H_i} (\bigcup H_i)$ – перетин всіх об'єднань $\bigcup H_i$ таких, що $G \bigcap \Omega \subset \bigcup H_i$ (див. §2.2, §3.3).

Очевидно $G_* \subset G \bigcap \Omega \subset G^*$, тому $P(G_*) \le P(G^*)$.

$\tilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha \in [0; 1]$, – це узагальнена статистична ймовірність (ймовірнісна міра) на $(R^n, \mathscr{B}(R^n))$, отримана як продовження міри $P(A)$, $A \in S$, із простору подій $S = \{A \big|\ A = \bigcup_{i \in I} H_i, I \subset \{1,\ 2,\ ...,\ k\}\}$ на простір подій $\tilde{S} = \mathscr{B}(R^n)$ за формулою (3.4.1).

Стосовно функції

$F(x_1, x_2, ..., x_n) = P(G_*) + \alpha(P(G^*) - P(G_*)) = (1 - \alpha)P(G_*) + \alpha P(G^*)$,

$\alpha \in [0; 1]$, задовільняються такі властивості

1. $F(x_1, x_2, ..., x_n) \ge 0$, оскільки $P(G_*) \ge 0$, $P(G^*) \ge 0$;

2. $F(x_1, x_2, ..., x_n)$ неспадна за будь якою сукупністю своїх аргументів, оскільки якщо $x_i \le \tilde{x}_i$, $i \in \overline{1, n}$,

$$G = (-\infty;\, x_1) \times (-\infty;\, x_2) \times ... \times (-\infty;\, x_n),$$
$$\tilde{G} = (-\infty;\, \tilde{x}_1) \times (-\infty;\, \tilde{x}_2) \times ... \times (-\infty;\, \tilde{x}_n),$$

то $G \subset \tilde{G}$, $G \bigcap \Omega \subset \tilde{G} \bigcap \Omega$, а тому $G_* \subset \tilde{G}_*$, $G^* \subset \tilde{G}^*$, $P(G_*) \le P(\tilde{G}_*)$, $P(G^*) \le P(\tilde{G}^*)$, звідки

$$F(x_1, x_2, ..., x_n) = (1-\alpha)P(G_*) + \alpha P(G^*) \le (1-\alpha)P(\tilde{G}^*) + \alpha P(\tilde{G}^*) =$$
$$= F(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n),$$
оскільки $1-\alpha \ge 0$, $\alpha \ge 0$.

3. Якщо принаймні один аргумент x_i, $i \in \overline{1, n}$, спрямувати до $-\infty$, тоді значення функції $F(x_1, x_2, ..., x_n)$ прямуватиме до нуля. Справді, якщо $x_i \to -\infty$, то
$$(-\infty; x_1) \times ... \times (-\infty; x_{i-1}) \times (-\infty; x_i) \times (-\infty; x_{i+1}) \times ... \times (-\infty; x_n)$$
прямує до
$$(-\infty; x_1) \times ... \times (-\infty; x_{i-1}) \times (-\infty; -\infty) \times (-\infty; x_{i+1}) \times ... \times (-\infty; x_n) = \varnothing,$$
бо $(-\infty; -\infty) = \varnothing$, тобто $G \to \varnothing$, $G \bigcap \Omega \to \varnothing$, а тому $G_* = \varnothing$, $G^* = \varnothing$, звідки $(1-\alpha)P(G_*) + \alpha P(G^*) = 0$.

Нагадаємо, що оскільки за довільної множини A елементів деякої природи має місце включення $A \subset A$, тому доцільно вважати, що $\varnothing \subset \varnothing$, тобто $\varnothing$ є однією із множин, до якої входить $G \bigcap \Omega = \varnothing \bigcap \Omega$, звідки $\bigcap\limits_{G\bigcap\Omega \subset \bigcup H_i} (\bigcup H_i) = \varnothing$.

4. Якщо всі x_i, $i \in \overline{1, n}$, спрямувати до $+\infty$, тоді значення функції $F(x_1, x_2, ..., x_n)$ прямуватиме до 1.

Справді, якщо $x_i \to +\infty$, за всіх $i \in \overline{1, n}$, тоді
$$G = (-\infty; x_1) \times (-\infty; x_2) \times ... \times (-\infty; x_n) \to$$
$$\to (-\infty; \infty) \times (-\infty; \infty) \times ... \times (-\infty; \infty) = R^n = \tilde{\Omega} \supset \Omega$$
а тому $G_* = \Omega$, $P(G_*) = 1$, $G^* = \Omega$, $P(G^*) = 1$, звідки $(1-\alpha)P(G_*) + \alpha P(G^*) = 1$.

Зауважимо, що коли $G = (-\infty; x) \in S$, тоді $P(G_*) = P(G^*) = P(G)$, звідки $\tilde{P}(G) = P(G)$.

Функція $F(x_1, x_2, ..., x_n)$, що визначається за формулою (3.4.1), називається функцією розподілу ймовірностей $\tilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha \in [0; 1]$, на множині $\tilde{\Omega} = R^n$, де $\tilde{P}(G)$, $G \in \tilde{S}$, є продовженням міри P, заданої на просторі подій $S = \{A \mid A = \bigcup\limits_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$, на простір $\tilde{S} \supset S$ за вказаним способом.

Приклад 3.4.1 Нехай (див. Рис. 3.4.1)
$$\Omega = \{(x_j, y_i) \mid x_j \in \overline{1, 3}, y_i \in \overline{1, 3}\} \subset R^2,$$

$$S = \{A \mid A = \bigcup_{i \in I} \bigcup_{j \in J} \{x_j, y_i\}, I \subset \{1, 2, 3\}, J\{1, 2, 3\}\} =$$
$$= \{A \mid A = \bigcup_{i \in I} \bigcup_{j \in J} H_{ij}, I \subset \{1, 2, 3\}, J\{1, 2, 3\}\},$$

де $H_{ij} = \{(x_j, y_i)\}$, $i \in \{1, 2, 3\}$, $j \in \{1, 2, 3\}$ – одноточкові підмножини розглядуваної скінченної множини Ω.

Нехай на S задана ймовірнісна міра $P(\{(x_j, y_i)\}) = \dfrac{1}{9}$ за всіх $i \in \overline{1, 3}$, $j \in \overline{1, 3}$. Нехай $G = (-\infty; x) \times (-\infty; y) \in \tilde{S} \supset S$, $G \overline{\in} S$.

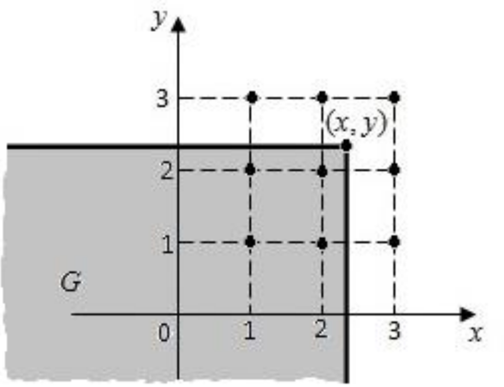


Рис. 3.4.1

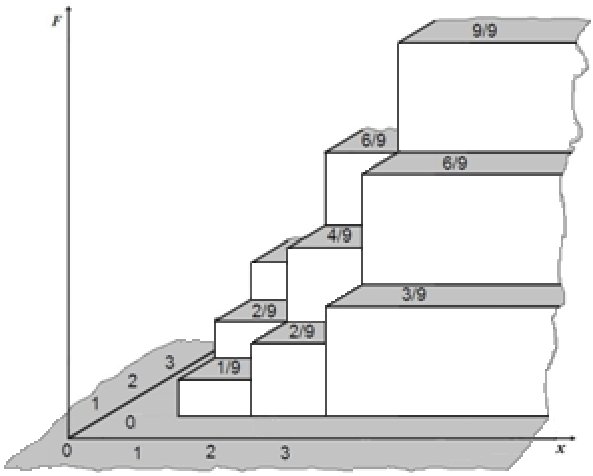


Рис. 3.4.2

Оскільки $G \cap \Omega \in S$ за довільних $(x, y) \in R^2$, то $G_* = G^*$ і за формулою (3.4.1) $F(x, y) = P(G_*) = \tilde{P}(G)$, тобто функція $F(x, y)$ в розглядуваному випадку визначена за довільних $(x, y) \in R^2$. Враховуючи заданий поточковий розподіл ймовірності $\Omega = \{(x_j; y_i) \mid x_j \in \overline{1,3},\ y_i \in \overline{1,3}\}$ одержуємо:

$$F(x, y) = \begin{cases} 0, \text{ коли } x \le 1 \text{ або } y \le 1, \\ \frac{1}{9}, \text{ коли } 1 < x \le 2 \text{ і } 1 < y \le 2, \\ \frac{2}{9}, \text{ коли } 2 < x \le 3 \text{ і } 1 < y \le 2, \\ \qquad \text{або } 1 < x \le 2 \text{ і } 2 < y \le 3, \\ \frac{3}{9}, \text{ коли } 3 < x \text{ і } 1 < y \le 2, \\ \qquad \text{або } 1 < x \le 2 \text{ і } 3 < y, \\ \frac{4}{9}, \text{ коли } 2 < x \le 3 \text{ і } 2 < y \le 3, \\ \frac{6}{9}, \text{ коли } 2 < x \le 3 \text{ і } 3 < y, \\ \qquad \text{або } 3 < x \text{ і } 2 < y \le 3, \\ \frac{9}{9}, \text{ коли } 3 < x \text{ і } 3 < y. \end{cases}$$

Графічне зображення поверхні $z = F(x, y)$ подано на Рис. 3.4.2.

Як видно з наведеного, коли множина Ω скінченна, $\Omega = \bigcup_{i=1}^{m} \bigcup_{j=1}^{s} (\{(x_j, y_i)\})$, і на ній задано поточковий розподіл імовірностей, тобто задані ймовірності $P((x_j, y_i))$, $i \in \overline{1, m}$, $j \in \overline{1, s}$, тоді функція $F(x, y)$ поточкового розподілу ймовірностей на множині Ω за точками (x_j, y_i), $i \in \overline{1, m}$, $j \in \overline{1, s}$, буде визначена за довільних $(x, y) \in R^2$, кусково стала і множина її значень буде скінченна.

Приклад 3.4.2 Нехай

$$\Omega = \{(x, y) \mid x \in (0;3], y \in (0;3]\} = (0;3] \times (0;3],$$

квадрат, довжина сторони якого дорівнює 3,

$$H_{i,j} = \{(x, y) \mid x \in (j-1;\ j], y \in (i-1; i]\} = (j-1;\ j] \times (i-1; i],$$

$i \in \overline{1,\ 3}$, $j \in \overline{1,\ 3}$, – квадрати, довжини сторін яких дорівнюють 1, $H_{i,j} \bigcap H_{m,r} = \varnothing$, коли $i \neq m$ або $j \neq r$, $\bigcup\limits_{i=1}^{3} \bigcup\limits_{j=1}^{3} H_{i,j} = \Omega$ (див. Рис. 3.4.3), $S = \{A \mid A = \bigcup\limits_{i \in I} \bigcup\limits_{j \in J} H_{i,j}, I \subset \{1,\ 2,\ 3\}, J \subset \{1,\ 2,\ 3\}\}$.

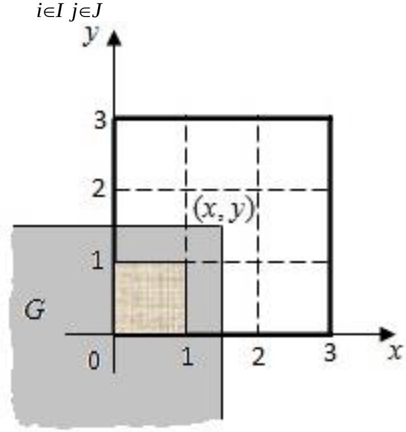


Рис. 3.4.3

Нехай на S задана ймовірнісна міра $P(H_{ij}) = \frac{1}{9}$, $i \in \overline{1,\ 3}$, $j \in \overline{1,\ 3}$, тобто задано розподіл імовірностей на нескінченній множині Ω за підмножинами $H_{i,j}$, $i \in \overline{1,\ 3}$, $j \in \overline{1,\ 3}$. Нехай $\tilde{\Omega} = R^2$, $\tilde{S} = \mathscr{B}(R^2)$, $G = (-\infty;\, x) \times (-\infty;\, y) \in \tilde{S} \supset S$, $G \,\overline{\in}\, S$.

Зауважимо, що в розглядуваному випадку $G \bigcap \Omega \,\overline{\in}\, S$, коли $(x, y) \,\overline{\in}\, \{(x, y) \mid x < 0 \text{ або } y < 0\} \bigcup \{(x, y) \mid x > 3 \text{ і } y > 3\} \bigcup$
$\bigcup \{(x, y) \mid x \in \overline{1,3} \text{ і } y \in \overline{1,3}\} \bigcup \{(x, y) \mid x \in \overline{1,3},\ y > 3\} \bigcup \{(x, y) \mid x > 3,\ y \in \overline{1,3}\}$, значення функції

$$\tilde{F}(x, y) = P(G \bigcap \Omega) = P((-\infty;\, x) \times (-\infty;\, y) \bigcap \Omega)$$

будуть невизначені, оскільки за вказаних (x, y) буде $(G \bigcap \Omega) \,\overline{\in}\, S$ (див. Рис. 3.4.4).

Разом з тим значення функції $F(x, y)$, визначені за формулою (3.4.1), будуть визначені за всіх $(x, y) \in R^2$ і набуватимуть сталих значень на множинах $\{(x, y) \mid x < 0 \text{ або } y < 0\}$; $\{(x, y) \mid x > 3 \text{ і } y > 3\}$; H_{ij}, $i \in \overline{1,3}$, $y \in \overline{1,3}$.

На Рис. 3.4.4 показана множина точок, в яких функція $F(x, y) = P(G \bigcap \Omega)$ визначена (такі точки затемнені).

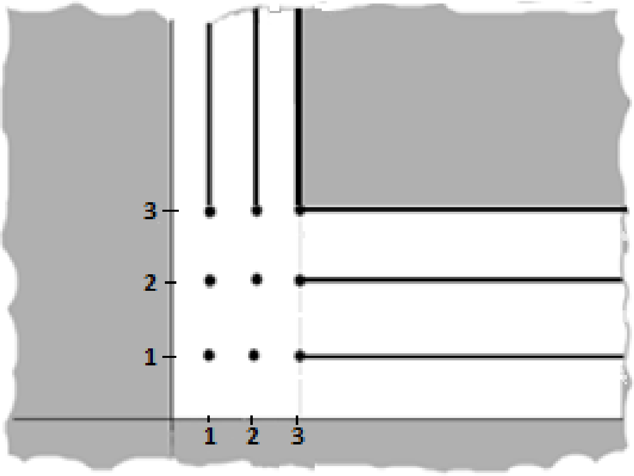


Рис. 3.4.4

Покладемо $\alpha = 0$. Тоді (див. §3.3) одержимо

$$F(x,y) = P(G_*) = \begin{cases} 0, & \text{коли } x \le 1 \text{ або } y \le 1, \\ \dfrac{1}{9}, & \text{коли } 1 < x \le 2 \text{ і } 1 < y \le 2, \\ \dfrac{2}{9}, & \text{коли } 2 < x \le 3 \text{ і } 1 < y \le 2, \\ & \quad \text{або } 1 < x \le 2 \text{ і } 2 < y \le 3, \\ \dfrac{3}{9}, & \text{коли } 3 < x \text{ і } 1 < y \le 2, \\ & \quad \text{або } 1 < x \le 2 \text{ і } 3 < y, \\ \dfrac{4}{9}, & \text{коли } 2 < x \le 3 \text{ і } 2 < y \le 3, \\ \dfrac{6}{9}, & \text{коли } 2 < x \le 3 \text{ і } 3 < y, \\ & \quad \text{або } 3 < x \text{ і } 2 < y \le 3, \\ \dfrac{9}{9}, & \text{коли } 3 < x \text{ і } 3 < y. \end{cases}$$

Очевидно в даному випадку

$$F(x,y) = P(G_*) = \iint\limits_{G_*} f(x,y)\,dxdy\,,$$

де

$$f(x,y)=\begin{cases}\dfrac{P(H_{i,j})}{m(H_{i,j})}, \text{ коли } (x,y)\in H_{i,j},\ i\in\overline{1,3},\ j\in\overline{1,3},\\ 0, \quad \text{коли } (x,y)\,\overline{\in}\,\Omega,\end{cases}$$

тобто

$$f(x,y)=\begin{cases}\dfrac{1}{9}, \text{ коли } (x,y)\in\Omega,\\ 0, \text{ коли } (x,y)\,\overline{\in}\,\Omega.\end{cases}$$

Таким чином функція $F(x,y)$ розподілу ймовірностей на нескінченній неперервній множині Ω, розглядуваній в даному прикладі, за підмножинами

$$H_{i,j}=\{(x,y)\mid\ x\in(j-1;j], y\in(i-1];i]\},\ i\in\overline{1,\ 3},\ j\in\overline{1,\ 3},$$

за умови, що $S=\{A\mid\ A=\bigcup\limits_{i\in I}\bigcup\limits_{j\in J}H_{i,j}, I\subset\{1,\ 2,\ 3\}, J\subset\{1,\ 2,\ 3\}\}$, тобто за умови, що як події розглядаються разом з порожньою множиною лише всеможливі об'єднання множин $H_{i,j}$, і тому за довільних (x,y) $G=(-\infty;x)\times(-\infty;y)\,\overline{\in}\,S$, виявляється такою самою, як і функція $F(x,y)$ розподілу ймовірностей на скінченній множині Ω, розглядуваній в прикладі 3.4.1.

Це означає, що за наведених способів задання просторів подій S і ймовірнісних мір на них не завжди за функцією $F(x,y)$ розподілу ймовірностей на множині Ω можна визначити тип розподілу ймовірностей.

Графік функції $F(x,y)$ розподілу ймовірностей на множині

$$\Omega=\{(x,y)\,|\,(x,y)\in(0;3]\times(0;3]\}$$

за підмножинами

$$H_{i,j}=\{(x,y)\mid\ (x,y)\in(j-1;j]\times(i-1;i]\},\ i\in\overline{1,\ 3},\ j\in\overline{1,\ 3},$$

коли

$$S=\{A\mid\ A=\bigcup_{i\in I}\bigcup_{j\in J}H_{i,j}, I\subset\{1,\ 2,\ 3\}, J\subset\{1,\ 2,\ 3\}\},$$

буде такий самий, як і графік функції $F(x,y)$ розподілу ймовірностей на множині $\Omega=\{(x_j,y_i)\,|\,(x_j\in\{1,\ 2,\ 3\}, y_i\in\{1,\ 2,\ 3\}\}$ за підмножинами $H_{i,j}=\{(x_j,y_i)\mid\ (i\in\{1,\ 2,\ 3\}, j\in\{1,\ 2,\ 3\}\}$, коли

$$S=\{A\mid\ A=\bigcup_{i\in I}\ \bigcup_{j\in J}\{x_j\,,\,y_i\},\ I\subset\{1,2,3\},\ J\{1,2,3\}\}=$$

$$=\{A\mid\ A=\bigcup_{i\in I}\ \bigcup_{j\in J}H_{i,j},\ I\subset\{1,2,3\},\ J\{1,2,3\}\},$$

(див. Рис. 3.4.2).

Зауважимо, що коли функцію $F(x,y)$ визначити за формулою

$$F(x,y) = P(\bigcup_{A\in S,\, A\subset(-\infty;\, x)\times(-\infty;\, y)} A),$$

то буде отримано той самий результат, що і за формулою (3.4.1) в разі $\alpha = 0$.

Лише після уточнення розподілів ймовірностей $P(H_i)$ на множинах H_i, $i \in \overline{1,k}$, (наприклад за припущення, що на множинах H_i ймовірності $P(H_i)$ розподілені рівномірно) значення $F(x_1, x_2, ..., x_n)$ будуть визначені у всіх точках $x \in R^n$, зокрема і у всіх точках $x \in H_i$, $i \in \overline{1,k}$. Одним із варіантів довизначення функції $F(x)$, $x \in R^n$, є задання функції $F(x_1, x_2, \dots, x_n)$ за формулою (3.4.1).

На практиці часто покладають $\widetilde{P}(G) \approx P(G_*) + \alpha P(G^* \setminus G_*)$, де $\alpha \in [0;1]$ – деяке фіксоване дійсне число із проміжка $[0;1]$, отримуючи в такий спосіб наближене значення $\widetilde{F}(x)$ функції $F(x)$. Іноді покладають $\alpha = \dfrac{m(G \cap \Omega \setminus G_*)}{m(G^* \setminus G_*)}$.

Зауважимо, що коли функцію $F(x)$, $x \in R^n$, визначити за формулою

$$F(x) = P(\bigcup_{A\in S,\; A\subset(-\infty;\, x]} A), \tag{3.4.2}$$

де $\Omega = \bigcup_{j_1=1}^{k_1} \bigcup_{j_2=1}^{k_2} \dots \bigcup_{j_n=1}^{k_n} (a_{j_1-1}^{(i)}; a_{j_1}^{(i)}] \times (a_{j_2-1}^{(i)};\; a_{j_2}^{(i)}] \times \dots \times (a_{j_n-1}^{(i)}; a_{j_n}^{(i)}]$,

$H_{j_1,\, j_2,\, \dots,\, j_n} = (a_{j_1-1}^{(i)};\; a_{j_1}^{(i)}] \times (a_{j_2-1}^{(i)}; a_{j_2}^{(i)}] \times \dots \times (a_{j_n-1}^{(i)}; a_{j_n}^{(i)}]$,

$j_1 \in \overline{1,\, k_1}$, $j_2 \in \overline{1,\, k_2}, \dots,$ $j_n \in \overline{1,\, k_n}$,

$S = \{A \mid A = \bigcup_{i\in I} H_i,\; I \subset \overline{1,k}\}$, тобто $P(A) = P(\bigcup_{H_i \subset A} H_i) = \sum_{H_i \in A} P(H_i)$,

тоді функція $F(x)$ буде визначена за всіх $x \in R^n$ і разом з тим кусково стала як за поточкового, так і за поінтервального розподілу ймовірностей, набуваючи сталих значень на кожному з інтервалів.

Стосовно функції $F(x_1, x_2, ..., x_n)$ задовільняються такі властивості

1. $F(x_1, x_2, ..., x_n) \geq 0$, оскільки $P(G_*) \geq 0$, $G_* \subset G^*$, $P(G_*) \leq P(G^*)$, $\alpha \in [0;1]$.

2. $F(x_1, x_2, ..., x_n)$ неспадна за будь якою сукупністю своїх аргументів, оскільки коли $x_i \leq \tilde{x}_i$, $i \in \overline{1, n}$,

$$G = (-\infty; x_1) \times (-\infty; x_2) \times (-\infty; x_n),$$
$$\tilde{G} = (-\infty; \tilde{x}_1) \times (-\infty; \tilde{x}_2) \times (-\infty; \tilde{x}_n),$$

тоді $G \subset \tilde{G}$, $G \cap \Omega \subset \tilde{G} \cap \Omega$, тому $G_* \subset \tilde{G}_*$, $G^* \subset \tilde{G}^*$, $P(G_*) \leq P(\tilde{G}_*)$, $P(G^*) \leq P(\tilde{G}^*)$, звідки

$$F(x_1, x_2, ..., x_n) =$$

$$= (1-\alpha)P(G_*) + \alpha P(G^*) \leq (1-\alpha)P(\tilde{G}^*) + \alpha P(\tilde{G}^*) =$$
$$= F(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n),$$

оскільки $1-\alpha \geq 0$, $\alpha \geq 0$.

3. Коли $x_i \to -\infty$ принаймні за одного i, тоді $F(x_1, x_2, ..., x_n) \to 0$, оскільки коли $x_i \to -\infty$, тоді $(-\infty; x_1) \times ... \times (-\infty; x_{i-1}) \times (-\infty; x_i) \times (-\infty; x_{i+1}) \times ... \times (-\infty; x_n)$ прямує до $(-\infty; x_1) \times ... \times (-\infty; x_{i-1}) \times (-\infty; -\infty) \times (-\infty; x_{i+1}) \times ... \times (-\infty; x_n) = \varnothing$, тобто $G \to \varnothing$, $G \cap \Omega \to \varnothing$, $G_* = \varnothing$, $G^* = \varnothing$, звідки $(1-\alpha)P(G_*) + \alpha P(G^*) \to 0$, $\alpha \in [0;1]$

4. Коли $x_i \to +\infty$ за всіх $i \in \overline{1, n}$, тоді $F(x_1, x_2, ..., x_n) \to 1$.

Справді, коли $x_i \to \infty$ за всіх $i \in \overline{1, n}$, тоді

$$G = (-\infty; x_1) \times (-\infty; x_2) \times ... \times (-\infty; x_n) \to$$
$$(-\infty; \infty) \times (-\infty; \infty) \times ... \times (-\infty; \infty) = R^n = \tilde{\Omega} \supset \Omega,$$

а тому $G_* = \Omega$, $G^* = \Omega$, $P(G_*) = 1$, $P(G^*) = 1$, $\tilde{P}(G) = (1-\alpha)P(G_*) + \alpha P(G^*) = 1$, $\alpha \in [0; 1]$.

Зокрема за $\alpha = 0$ $\tilde{P}(G) = P(G_*)$.

Зауважимо, що коли $G \cap \Omega \in S$, тоді $P(G_*) = P(G^*) = P(G)$, звідки $\tilde{P}(G) = P(G)$ (див. §2.2-§2.6).

Функція $F(x_1, x_2, ..., x_n)$, що визначається за формулою (3.4.1), називається функцією розподілу ймовірностей $\tilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, $\alpha \in [0; 1]$, на множині $\tilde{\Omega} = R^n$, де $\tilde{P}(G)$ є наближеним значенням невідомого значення $P_1(G)$ в разі, коли P_1, $G \in \tilde{S}$, є продовженням ймовірнісної міри P, заданої на просторі подій $S = \{A \mid A = \bigcup_{i \in I} H_i,\ I \subset \{1, 2, ..., k\}\}$, на простір $\tilde{S} \supset S$.

Зауважимо, що коли $F(x,y)$ визначити за рівністю $F(x,y)=P_n^*(\bigcup_{A\in S,\, A\subset(-\infty;\, x)\times(-\infty;\, y)} A)$, тоді значення $F(x,y)$ будуть такі самі, як і значення $F(x,y)$, визначеної за формулою (3.4.1), в разі $\alpha=0$.

Нехай задано ймовірнісний простір (Ω, S, P), де $\Omega=[a;b)\times[d;e)$, $-\infty<a<\infty$, $-\infty<b<\infty$, $-\infty<d<\infty$, $-\infty<e<\infty$, $0<b-a<\infty$, $0<e-d<\infty$, $\Omega=\bigcup_{i=1}^{m}\bigcup_{j=1}^{s} H_{ij}$, H_{ij} - квадрати зі сторонами довжиною h, тобто

$$H_{i,j}=\{(x,y)\,|\;\; x\in[x_0+(j-1)h;\, x_0+jh),\, y\in[y_0+(j-1)h;\, y_0+jh)\},$$
$$i\in\overline{1,m},\;\; j\in\overline{1,\, s},$$

або

$$H_{i,j}=\{(x,y)\,|\;\; (x,\, y)\in[x_0+(j-1)h;\, x_0+jh)\times[y_0+(j-1)h;\, y_0+jh)\}$$
$$,$$
$$i\in\overline{1,m},\;\; j\in\overline{1,\, s},\;\; x_0=a,\;\; y_0=d,$$
$$S=\{A\,|\;\; A=\bigcup_{i\in I}\bigcup_{j\in J} H_{ij},\, I\subset\{1,\, 2,\ldots,\, m\},\, J\subset\{1,\, 2,\ldots,\, s\}\},$$
$$\tilde{\Omega}=R^2,\;\; \tilde{S}=\mathscr{B}(R^2),\;\; G=(-\infty;x)\times(-\infty;y)\in\tilde{S}\supset S,\;\; G\,\overline{\in}\, S,$$
$$G_*=\bigcup_{\cup H_{ij}\subset G\cap S}(\cup H_{ij}),\;\; G^*=\bigcap_{G\cap S\subset\cup H_{ij}}(\cup H_{ij}),$$
$$F(x,y)=P(G_*)+\alpha(P(G^*)-P(G_*))=\tilde{P}(G),\;\; \alpha\in[0;1].$$

Тоді легко бачити, що коли дві різні точки (x_1,y_1) і (x_2,y_2) є точками одного і того самого квадрата H_{i_0,j_0}, тобто $(x_1,y_1)\in H_{i_0,j_0}$, $(x_2,y_2)\in H_{i_0,j_0}$, тоді $F(x_1,y_1)=F(x_2,y_2)$. Це означає, що функція $F(x,y)$ набуває одного і того самого значення у всіх точках квадрата H_{i_0,j_0}. *Таким чином, функція $F(x,y)$ є кусково сталою і набуває на множині $\tilde{\Omega}=R^2\supset\Omega$ не більше, ніж $(m\cdot s+2)$ різних значень, коли $h>0$.*

Нехай

$$f(x,y)=\begin{cases}\dfrac{P(H_{i,j})}{m(H_{i,j})}=c_{ij}, & \text{коли}\;\; (x,y)\in H_{ij},\, i\in\overline{1,\, m},\, j\in\overline{1,\, s},\\ 0, & \text{коли}\;\; (x,y)\,\overline{\in}\,\Omega.\end{cases}$$

Оскільки

$$F(x,y)=P(G_*)+\alpha(P(G^*)-P(G_*))=$$

$$= \iint\limits_{G_*} f(x,y)dxdy + \alpha(\iint\limits_{G^*} f(x,y)dxdy - \iint\limits_{G_*} f(x,y)dxdy) =$$

$$= \sum_{H_{ij}\subset G_*} P(H_{ij}) + \alpha(\sum_{H_{ij}\in G^*} P(H_{ij}) - \sum_{H_{ij}\subset G_*} P(H_{ij})) =$$

$$= \sum_{H_{ij}\subset G_*} c_{ij}\cdot m(H_{ij}) + \alpha(\sum_{H_{ij}\in G^*} c_{ij}\cdot m(H_{ij}) - \sum_{H_{ij}\subset G_*} c_{ij}\cdot m(H_{ij})) =$$

$$= (1-\alpha)\sum_{H_{ij}\subset G_*} c_{ij}\cdot h^2 + \alpha \sum_{H_{ij}\subset G^*} c_{ij}\cdot h^2 =$$

$$= h^2(\sum_{H_{ij}\subset G_*} c_{ij} + \alpha(\sum_{H_{ij}\subset G^*} c_{ij} - \sum_{H_{ij}\subset G_*} c_{ij})),$$

а $F(-\infty,-\infty)=0$, $F(+\infty,+\infty)=1$, то в множині Ω знайдеться принаймні дві сусідні множини H_{i_1,j_1} і H_{i_2,j_2} такі, що $F(x_1,y_1)\neq F(x_2,y_2)$, коли $(x_1,y_1)\in H_{i_1,j_1}$, $(x_2,y_2)\in H_{i_2,j_2}$, тобто буде $|F(x_1,y_1)-F(x_2,y_2)| = |c_{i_1,j_1} - c_{i_2,j_2}| > 0$ навіть коли $\rho((x_1,y_1),(x_2,y_2))\to 0$, $(x_1,y_1)\in H_{i_1,j_1}$, $(x_2,y_2)\in H_{i_2,j_2}$.

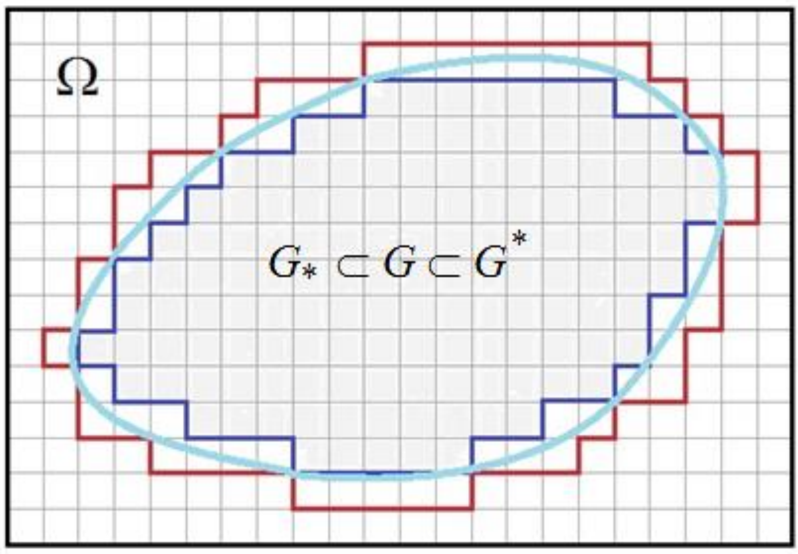


Рис. 3.4.5

Це означає, що за так заданого простору подій S, коли

$$S = \{A \mid A = \bigcup_{i\in I}\bigcup_{j\in J} H_{ij}, I\subset\{1,\ 2,\ldots,\ m\}, J\subset\{1,\ 2,\ldots,\ s\}\}$$

$$H_{i,j} = \{(x,y) \mid (x,\ y)\in[x_0+(j-1)h;\ x_0+jh)\times[y_0+(j-1)h;\ y_0+jh)\},$$

$$i \in \overline{1,\ m},\ \ j \in \overline{1,\ s}$$

функція $F(x,y)$ не може бути неперервною, якщо $h > 0$ (вона набуває не більше, ніж $m \cdot s + 2$ різних значень).

Зауважимо, що коли $h \to 0$, то $P(G^*) - P(G_*) \to 0$ (див. Рис. 3.4.5), і якщо $f(x,y) \le c < \infty$, то стосовно всіх H_{ij}, $i \in \overline{1,\ m}$, $i \in \overline{1,\ s}$ імовірність

$$P(H_{i,j}) = \iint\limits_{H_{ij}} f(x,y)dxdy \le c \cdot m(H_{ij}) = c \cdot h^2$$

стає як завгодно малою, коли $h \to 0$, і тому якщо точки (x_1, y_1) і (x_2, y_2) належать до сусідніх квадратів, різниця між значеннями $F(x_1, y_1)$ і $F(x_2, y_2)$ стає як завгодно малою, тобто $|F(x_1, y_1) - F(x_2, y_2)| \underset{h \to 0}{\to} 0$.

Отже якщо за всіх $(x,y) \in \Omega$ $f(x,y) \le c < \infty$ і $h \to 0$, тоді функція $F(x,y)$ розподілу ймовірностей на $\tilde{\Omega}$ стає неперервною, тобто якщо $0 < h < \varepsilon$, то $|F(x_1, y_1) - F(x_2, y_2)| < \delta$, де $\varepsilon > 0$ і $\delta > 0$ як завгодно малі додатні числа.

В такому разі

$$F(x,y) = P(G_*) = P(G^*) = P(G) = P((-\infty\,;\,x) \times (-\infty\,;\,y))$$

або, в скороченому записі,

$$F(x) = P((-\infty\,;\,x)),\ \ x = (x_1, x_2) \in R^2,\ \ (-\infty;\,x) = (-\infty;\,x_1) \times (-\infty;\,x_2).$$

Нехай, як і раніше,

$$H_{i,j} = \{(x,y) \mid\ (x,\ y) \in [x_0 + (j-1)h;\ x_0 + jh) \times [y_0 + (j-1)h;\ y_0 + jh)\},$$

$$i \in \overline{1,\ m},\ \ j \in \overline{1,\ s},$$

квадрати із стороною довжини h, $0 < h < \infty$, $-\infty < x_0 < \infty$, $-\infty < y_0 < \infty$, $m < \infty$, $s < \infty$,

$$\Omega = \bigcup_{i=1}^{m} \bigcup_{j=1}^{s} H_{ij} = [a\,;\,b) \times [d\,;\,e),$$

$$S = \{A \mid\ A = \bigcup_{i \in I} \bigcup_{j \in J} H_{ij}, I \subset \{1,\ 2, \ldots,\ m\}, J \subset \{1,\ 2, \ldots,\ s\}\},$$

на S задана ймовірнісна міра

$$P(A) = \sum_{i \in I} \sum_{j \in J} P(H_{ij}),\ \ I \subset \{1, 2, \ldots, m\},\ \ J \subset \{1, 2, \ldots, s\},$$

числа $0 \le P(H_{ij})$ визначені за всіх $i \in \overline{1,\ m}$, $j \in \overline{1,\ s}$, $\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{s} P(H_{ij}) = 1$, $\tilde{\Omega} = R^2$, $\tilde{S} = \mathscr{B}(R^2) \supset S$ – сукупність всіх квадровних фігур із R^2.

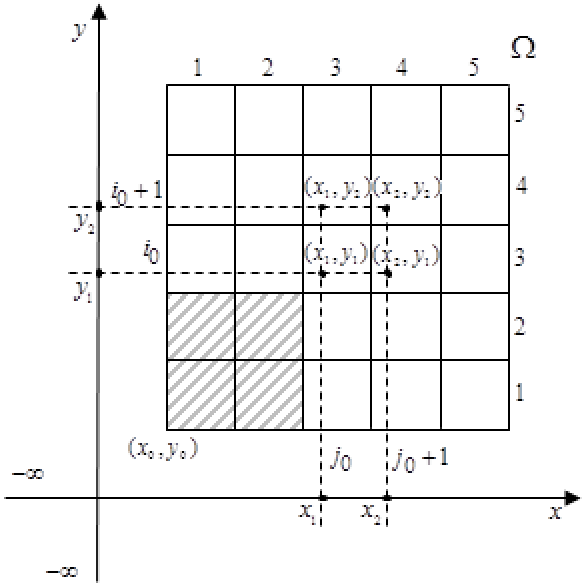


Рис. 3.4.6. а) G_{x_1,y_1}

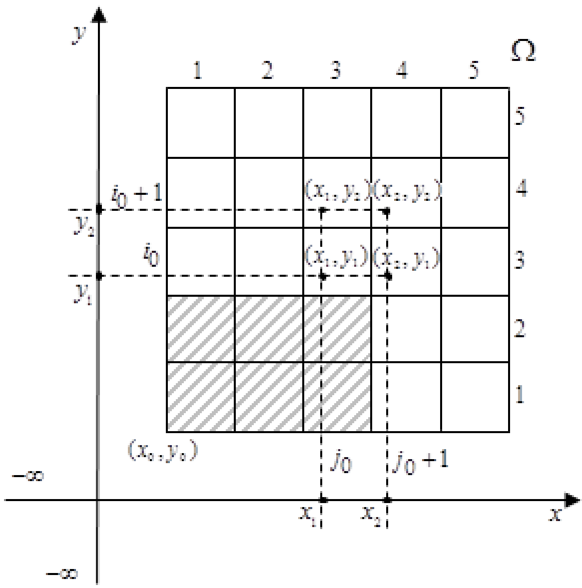


Рис. 3.4.6. б) G_{x_2,y_1}

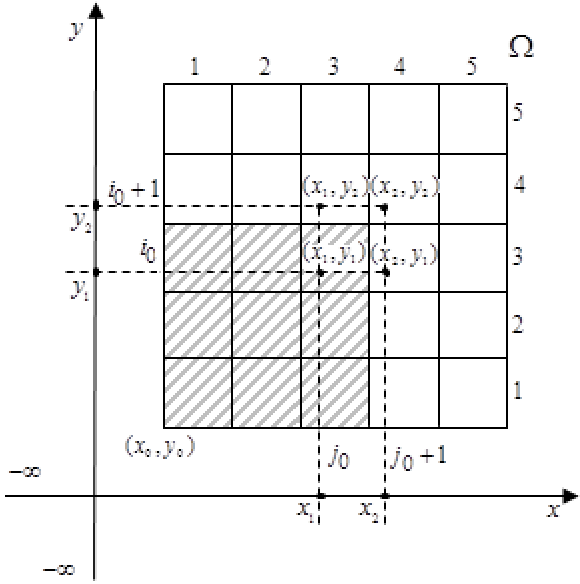


Рис. 3.4.6. в) G_{x_2,y_2}

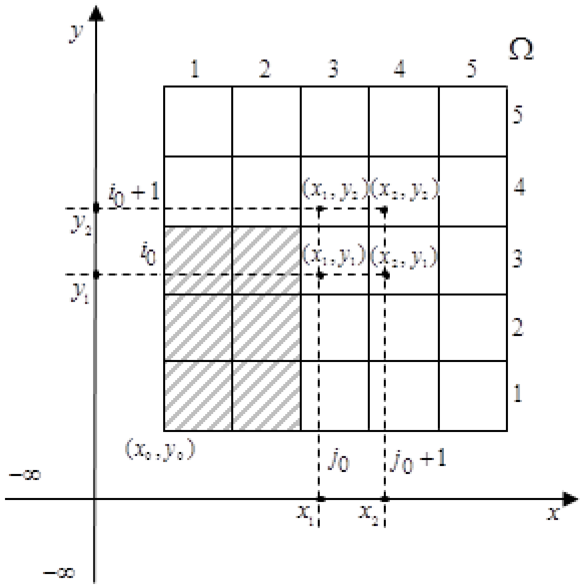


Рис. 3.4.6. г) G_{x_1,y_2}

З рис. 3.4.6. а), б), в), г) можна бачити, що $G_{x_2,y_1} \setminus G_{x_1,y_1} = H_{1,3} + H_{2,3}$,

$$G_{x_2,y_2} \setminus G_{x_1,y_1} = H_{3,1} + H_{3,2} + H_{3,3} + H_{1,3} + H_{2,3}$$

Нехай $G = (-\infty;\, x) \times (-\infty;\, y) \subset \tilde{\Omega}$, $G \in \tilde{S}$,

$$F(x, y) = \tilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*)),\ \alpha \in [0,\ 1],$$

$\tilde{P}(G)$, $G \in \tilde{S}$, – продовження міри $P(A)$, $A \in S$, із простору подій S на простір $\tilde{S}$ за формулою (3.4.1), і в такому разі $\tilde{P}(G) = P(G)$, коли $G \in S$.

Розглянемо два сусідні квадрати H_{i_0, j_0} і H_{i_0+1, j_0+1} і нехай $(x_1, y_1) \in H_{i_0, j_0}$ і $(x_2, y_2) \in H_{i_0+1, j_0+1}$.

Покладемо $\alpha = 0$,

$$G_{x_1,y_1} = (-\infty;\, x_1) \times (-\infty; y_1) \bigcap \Omega,\ G_{x_1,y_2} = (-\infty;\, x_1) \times (-\infty; y_2) \bigcap \Omega,$$

$$G_{x_2,y_1} = (-\infty;\, x_2) \times (-\infty; y_1) \bigcap \Omega,\ G_{x_2,y_2} = (-\infty;\, x_2) \times (-\infty; y_2) \bigcap \Omega,$$

і розглянемо різниці

$F(x_2, y_2) - F(x_1, y_2)$, $F(x_2, y_1) - F(x_1, y_1)$, $F(x_1, y_2) - F(x_1, y_1)$, $F(x_2, y_2) - F(x_2, y_1)$.

Легко бачити (див. Рис. 3.4.6), що

$$F(x_2, y_2) - F(x_2, y_1) = \tilde{P}(G_{x_2,y_2}) - \tilde{P}(G_{x_2,y_1}) =$$

$$= \tilde{P}(G_{x_2,y_2} \setminus G_{x_2,y_1}) = P(\bigcup_{j=1}^{j_0} H_{i_0, j}),$$

$$F(x_1, y_2) - F(x_1, y_1) = \tilde{P}(G_{x_1,y_2}) - \tilde{P}(G_{x_1,y_1}) =$$

$$= \tilde{P}(G_{x_1,y_2} \setminus G_{x_1,y_1}) = P(\bigcup_{j=1}^{j_0-1} H_{i_0, j}) .$$

Тому

$$(F(x_2, y_2) - F(x_2, y_1)) - (F(x_1, y_2) - F(x_1, y_1)) =$$

$$= P(\bigcup_{j=1}^{j_0} H_{i_0 j}) - P(\bigcup_{j=1}^{j_0-1} H_{i_0 j}) = P(\bigcup_{j=1}^{j_0} H_{i_0 j} \setminus \bigcup_{j=1}^{j_0-1} H_{i_0 j}) = P(H_{i_0 j_0}) .$$

Аналогічно

$$F(x_2, y_2) - F(x_1, y_2) = \tilde{P}(G_{x_2,y_2}) - \tilde{P}(G_{x_1,y_2}) =$$

$$= \tilde{P}(G_{x_2,y_2} \setminus G_{x_1,y_2}) = P(\bigcup_{i=1}^{i_0} H_{i, j_0}),$$

$$F(x_2, y_1) - F(x_1, y_1) = \tilde{P}(G_{x_2,y_1}) - \tilde{P}(G_{x_1,y_1}) =$$

$$= \widetilde{P}(G_{x_2,y_1} \setminus G_{x_1,y_1}) = P(\bigcup_{i=1}^{i_0-1} H_{i,j_0}) .$$

Тому

$$(F(x_2,y_2)-F(x_1,y_2))-(F(x_2,y_1)-F(x_1,y_1)) =$$

$$= P(\bigcup_{i=1}^{i_0} H_{i,j_0}) - P(\bigcup_{i=1}^{i_0-1} H_{i,j_0}) = P(\bigcup_{i=1}^{i_0} H_{i,j_0} \setminus \bigcup_{i=1}^{i_0-1} H_{i,j_0}) = P(H_{i_0 j_0}) .$$

Нехай

$$f(x,y) = \begin{cases} \dfrac{P(H_{i,j})}{m(H_{i,j})} = \dfrac{P(H_{i,j})}{h_0^2}, \text{ коли } (x,y) \in H_{i,j}, i \in \overline{1, m}, j \in \overline{1, s}, \\ 0, \qquad \text{коли } (x,y) \overline{\in} \Omega, \end{cases}$$

усереднена щільність розподілу ймовірностей на множині Ω за підмножинами $H_{i,j}$, коли h_0 – деяке фіксоване число, $0 < h_0 < \dfrac{b-a}{s_0} = \dfrac{e-d}{m_0}$, m_0 і s_0 – фіксовані. Нехай спочатку $h = h_0$, $m = m_0$, $s = s_0$. Оскільки $0 \le P(H_{ij}) \le 1$, і $h_0^2 > 0$, то $\max\limits_{(x,y)\in\Omega} f(x,y) \le c < \infty$, де c – деяке скінченне число.

Тому

$$F(x_2,y_1) - F(x_1,y_1) = P(\bigcup_{j=1}^{j_0-1} H_{i_0,j}) =$$

$$= \sum_{j=1}^{j_0-1} P(H_{i_0,j}) = \sum_{j=1,(x,y)\in H_{i_0,j}}^{j_0-1} f(x,y) m(H_{i_0,j}) \le c \cdot h \cdot (e-d) \to 0 ,$$

коли $h \to 0$, тобто $F(x_2,y_1) \to F(x_1,y_1)$, коли $h \to 0$.

Міркуючи аналогічно, одержимо:

$$F(x_2,y_2) \to F(x_1,y_2), \text{ коли } h \to 0 ,$$
$$F(x_1,y_2) \to F(x_1,y_1), \text{ коли } h \to 0 ,$$
$$F(x_2,y_2) \to F(x_2,y_1), \text{ коли } h \to 0 ,$$
$$F(x_2,y_1) \to F(x_1,y_1), \text{ коли } h \to 0 .$$

Це означає, що коли $h \to 0$, а $f(x,y) \le c < \infty$ за довільних $(x,y) \in \Omega$, тоді функція $F(x,y)$ стає неперервною, тобто $\left| F(x_2,y_2) - F(x_1,y_1) \right|$ стає як завгодно малим числом, коли віддаль $\rho((x_1,y_1),(x_2,y_2))$ між точками (x_1,y_1) і (x_2,y_2) стає досить малою. Зауважимо, що коли $h \to 0$, тоді квадрат H_{i_0,j_0} стягується в точку, а тому $\rho((x_1,y_1),(x_2,y_2)) \to 0$, коли $h \to 0$.

Отже, яким би малим не було $\varepsilon > 0$, знайдеться досить мале число δ таке, що матиме місце нерівність

$$\left|F(x_2, y_2) - F(x_1, y_1)\right| < \varepsilon \text{, коли } \rho((x_1, y_1), (x_2, y_2)) < \delta .$$

Як вже було сказано, коли $h \to 0$, то $G_* \to G, G^* \to G$, $P(G_*) \to P(G^*)$, а тому $\widetilde{P}(G) = P(G_*) + \alpha(P(G^*) - P(G_*)) \to P(G)$. Разом з тим щільність $f(x, y)$ на квадратах із стороною довжини h набуває тих самих значень, що і на квадратах із стороною довжини h_0, а зменшення довжини сторони квадрата відбувається через зменшення в кілька разів довжин сторін вже існуючих квадратів, тому координати вершин вже існуючих квадратів з подрібненням сторін не змінюються. За такого подрібнення квадратів простір $\hat{S}$ подій, утворюваних як всеможливі об’єднання подрібнених квадратів, включатиме і простір S, тобто буде $S \subset \hat{S}$, і це означатиме, що якщо $A \in S$, то $A \in \hat{S}$, але не навпаки.

Будемо вважати, що точки (x_1, y_1) і (x_2, y_2) завжди знаходяться в центрах або в лівих нижніх вершинах сусідніх квадратів H_{i_0, j_0} і H_{i_0+1, j_0+1}. Тоді буде $x_2 - x_1 = h$, $y_2 - y_1 = h$, а тому із виразу

$$(F(x_2, y_2) - F(x_1, y_2)) - (F(x_2, y_1) - F(x_1, y_1)) =$$

$$= P(H_{i_0, j_0}) = f(x, y) \cdot m(H_{i_0, j_0}) = f(x, y) \cdot h^2,\ (x, y) \in H_{i_0, j_0},$$

одержимо

$$\frac{(F(x_2, y_2) - F(x_1, y_2)) - (F(x_2, y_1) - F(x_1, y_1))}{h^2} = f(x, y),\ (x, y) \in H_{i_0, j_0},$$

або

$$\frac{\dfrac{(F(x_2, y_2) - F(x_1, y_2))}{h} - \dfrac{(F(x_2, y_1) - F(x_1, y_1))}{h}}{h} = f(x, y),\ (x, y) \in H_{i_0, j_0},$$

тобто

$$f(x, y) \approx \frac{\left.\dfrac{\partial F(x, y)}{\partial x}\right|_{\substack{x = x_1 \\ y = y_2}} - \left.\dfrac{\partial F(x, y)}{\partial x}\right|_{\substack{x = x_1 \\ y = y_1}}}{h} \approx$$

$$\approx \frac{\partial}{\partial y}\frac{\partial}{\partial x}F(x, y)\Big|_{\substack{x = x_1 \\ y = y_1}} = \left.\frac{\partial^2 F(x, y)}{\partial x \partial y}\right|_{\substack{x = x_1 \\ y = y_1}},$$

коли $(x, y) \in H_{i_0, j_0}$, h досить мале.

Отже, коли $h \to 0$, то за довільної точки $(x, y) \in \Omega$ буде $f(x, y) = \dfrac{\partial^2 F(x, y)}{\partial x \partial y}$.

Тоді $F(x,y)=\iint\limits_{G\bigcap\Omega} f(x,y)dxdy$, де $G=(-\infty,x)\times(-\infty,y)\in\widetilde{S}$, або, враховуючи, що $f(x,y)=0$, коли $(x,y)\overline{\in}\,\Omega$

$$F(x,y)=\iint\limits_{G} f(x,y)dxdy=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(x,y)dxdy\,.$$

Вправи для самостійного виконання

3.4.1. Перевірити, чи правильні твердження:

1. Якщо задано усереднену щільність розподілу ймовірностей на множині $\Omega=\bigcup\limits_{i=1}^{m}\bigcup\limits_{j=1}^{s} H_{i,j}$ за квадратами

$H_{i,j}=\{(x,y)\,|\;\; x\in[x_0+(j-1)h;\, x_0+jh),\;\; y\in[y_0+(j-1)h;\, y_0+jh)\}$,

$$f(x,y)=\begin{cases}\dfrac{P(H_{i,j})}{m(H_{i,j})}=\dfrac{c_{ij}}{h^2}, \text{ коли } (x,y)\in H_{i,j},\; i\in\overline{1,\,m},\;\; j\in\overline{1,\,s}, \\ \qquad\qquad c_{ij}\ge 0,\; \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{s}\dfrac{c_{ij}}{h^2}=1, \\ 0, \qquad \text{коли } (x,y)\overline{\in}\,\Omega, \end{cases}$$

тоді можна:

а) однозначно визначити усереднену щільність $f(x,y)$ розподілу ймовірностей на квадратах $\widetilde{H}_{i,j}$, $i\in\overline{1,\,5m}$, $j\in\overline{1,\,5s}$, сторони яких в 5 разів менші, ніж сторони квадратів $H_{i,j}$;

б) однозначно визначити усереднену щільність $f(x,y)$ розподілу ймовірностей на квадратах $\widetilde{H}_{i,j}$, сторони яких вдвічі більші, ніж сторони квадратів $H_{i,j}$, коли m і s - парні числа.

2. $\Omega=[-1;1)\times[-1;1)$, $h=0.2$,

$H_{i,j}=\{(x,y)\,|\;\; x\in[x_0+(j-1)h;\, x_0+jh)\times[y_0+(j-1)h;\, y_0+jh)\}$,

$$i\in\overline{1,\,10},\;\; j\in\overline{1,\,10};$$

$$G=\{(x,\,y)\;\;|\;\; x^2+y^2\le 0.8^2\};$$

$$f(x,y)=\begin{cases}\dfrac{1}{4}, \text{ коли } (x,y)\in\Omega, \\ 0, \text{ коли } (x,y)\,\overline{\in}\,\Omega.\end{cases}$$

Обчислити $P(G_*)$, $P(G^*)$, $P(G) = P(G_*) + \alpha(P(G^*) - P(G_*))$, поклавши $\alpha = \frac{1}{2}$.

Визначити значення функції $F(x, y) = P(G_*) + \alpha(P(G^*) - P(G_*))$, коли $G = (-\infty; x) \times (-\infty; y)$, $(x, y) \in [0; 0.2) \times [0; 0.2)$.

3.5. Функція та щільність многовимірного розподілу ймовірностей на $(R^n, \mathscr{B}(R^n))$ та їх властивості

Нехай P – деяка ймовірнісна міра на вимірному просторі $(R^n, \mathscr{B}(R^n))$. Легко бачити, що коли простір подій $S = \mathscr{B}(R^n)$, то за довільного $G \in S$ буде $P(G_*) = P(G^*) = P(G)$. Нехай $G = (-\infty; x_1) \times (-\infty; x_2) \times ... \times (-\infty; x_n)$.

Розглянемо функцію

$$F(x_1, x_2, ..., x_n) = P((-\infty; x_1) \times (-\infty; x_2) \times ... \times (-\infty; x_n)),$$

або (в скороченому запису)

$$F(x) = P((-\infty; x)),$$

де $x = (x_1, x_2, ..., x_n)$, $(-\infty; x) = (-\infty; x_1) \times ... \times (-\infty; x_n)$.

Розглянемо оператор $\Delta_{a_i b_i}$:

$$\Delta_{a_i b_i} F(x_1, x_2, ..., x_n) =$$
$$= F(x_1, ..., x_{i-1}, b_i, x_{i+1}, ..., x_n) - F(x_1, ..., x_{i-1}, a_i, x_{i+1}, ..., x_n).$$

Тоді:

$$\Delta_{a_1 b_1} \Delta_{a_2 b_2} ... \Delta_{a_n b_n} F(x_1, x_2, ..., x_n) = P([a; b)) \geq 0,$$

де $[a; b) = [a_1; b_1) \times [a_2; b_2) \times ... \times [a_n; b_n)$.

Стосовно функції $F(x_1, x_2, ..., x_n)$ задовільняються такі властивості:

1) $F(x_1, x_2, ..., x_n) \geq 0$;

2) $F(x_1, x_2, ..., x_n)$ неспадна за довільною сукупністю своїх аргументів;

3) з неперервності ймовірнісної міри P випливає, що $F(x_1, x_2, ..., x_n)$ неперервна зліва за довільною сукупністю змінних, тобто коли

$$x_i^{(k)} \to x_i, \ x_i^{(k)} \leq x_i, \ i \in I \subset \{1, 2, ..., n\}, I \neq \varnothing,$$

тоді

$$F(x^{(k)}) \to F(x), \ F(x^{(k)}) \leq F(x), \ k \to \infty;$$

4) $F(+\infty, +\infty, ..., +\infty) = 1$;

5) $\lim\limits_{x \to y} F(x_1, x_2, ..., x_n) = 0$, коли $x_i > y_i$, $i = \overline{1, n}$, $y_i = -\infty$ принаймні за одного i.

Ці властивості випливають з означення функції $F(x_1, x_2, ..., x_n)$ та властивостей ймовірності.

Означення. Будь-яка функція $F(x_1, x_2, ..., x_n)$, стосовно якої задовільняються умови 1) – 5), називається функцією розподілу ймовірностей на n-вимірному просторі $(R^n, \mathscr{B}(R^n))$.

Якщо $F(x_1, x_2, ..., x_n)$ – деяка функція розподілу ймовірностей на $(R^n, \mathscr{B}(R^n))$, то на n-вимірному просторі $(R^n, \mathscr{B}(R^n))$ існує, і притому єдина, ймовірнісна міра P така, що

$$P([a;\ b]) = \Delta_{a_1 b_1}\Delta_{a_2 b_2} \ ... \ \Delta_{a_n b_n} F(x_1, x_2, ..., x_n).$$

Нехай, наприклад, $F^{(1)}, F^{(2)}, ..., F^{(n)}$ – одновимірні функції розподілу ймовірностей в R^1. Тоді

$$F(x_1, x_2, ..., x_n) = F^{(1)}(x_1)\ F^{(2)}(x_2)\ ...\ F^{(n)}(x_n)$$

є функцією розподілу ймовірностей в R^n. Очевидно, що в такому разі

$$\Delta_{a_1 b_1}\Delta_{a_2 b_2} \ ... \ \Delta_{a_n b_n} F(x_1, x_2, ..., x_n) = \prod_{k=1}^{n} (F^{(k)}(b_k) - F^{(k)}(a_k)) \geq 0 .$$

Особливо важливим є випадок, коли

$$F^{(k)}(x_k) = \begin{cases} 0, \text{ коли } x_k \leq 0; \\ x_k, \text{ коли } 0 \leq x_k \leq 1; \\ 1 \quad \text{коли } 1 \leq x_k . \end{cases}$$

Тоді $F(x_1, x_2, ..., x_n) = x_1 \cdot x_2 \cdot ... \cdot x_n$, коли $x_i \in [0; 1)$ за всіх i.

Імовірнісну міру, відповідну такій n-вимірній функції розподілу ймовірностей, називають n-вимірною мірою Лебега на $[0; 1]^n = [0; 1] \times [0; 1] \times ... \times [0; 1]$.

Очевидно, коли $[a_i; b_i) \subset [0; 1]$, тоді

$$\Delta_{a_1 b_1}\Delta_{a_2 b_2} \ ... \ \Delta_{a_n b_n} F_n(x_1, x_2, ..., x_n) = \prod_{k=1}^{n} (b_i - a_i)$$

є об’ємом n-вимірного паралелепіпеда

$$[a_1;\ b_1) \times [a_2;\ b_2) \times ... \times [a_n;\ b_n).$$

Якщо $f(x_1, x_2, ..., x_n)$ – деяка невід'ємна борелівська функція така, що

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\ldots\int\limits_{-\infty}^{\infty} f(x_1,\, x_2,\, \ldots,\, x_n)\, d\,x_1 d\,x_2 \ldots d\,x_n = 1,$$

то

$$F(x_1,\, x_2, \ldots x_n) = \int\limits_{-\infty}^{x_1}\int\limits_{-\infty}^{x_2}\ldots\int\limits_{-\infty}^{x_n} f(t_1,\, t_2,\, \ldots,\, t_n)\, d\,t_1 d\,t_2 \ldots d\,t_n \quad (3.5.1)$$

є функцією розподілу ймовірностей на n-вимірному просторі $(R^n,\, \mathscr{B}(R^n))$.

Така функція $f(x_1,\, x_2,\, \ldots,\, x_n)$ називається щільністю n-вимірного розподілу ймовірностей.

В такому разі *розподіл ймовірностей називається абсолютно неперервним*.

Нагадаємо, що функція $y = f(x_1, x_2, \ldots, x_n)$, $(x_1, x_2, \ldots, x_n) \in R^n$, $y \in R^1$, називається борелівською, коли за відображенням $R^n \xrightarrow{f} R^1$: $f(x_1, x_2, \ldots, x_n) = y \in R^1$, прообраз $f^{-1}(G)$ довільної борелівської множини G із $\mathscr{B}(R^1)$ є борелівською множиною із $\mathscr{B}(R^n)$.

Якщо задано функцію абсолютно неперервного розподілу ймовірностей $F(x_1,\, x_2,\, \ldots,\, x_n)$, то як видно з формули (3.5.1) $f(x_1,\, x_2,\, \ldots,\, x_n) = \dfrac{\partial}{\partial x_1}\dfrac{\partial}{\partial x_2}\ldots\dfrac{\partial}{\partial x_n} F(x_1,\, x_2,\, \ldots,\, x_n)$ майже скрізь на R^n.

Вправи для самостійного виконання

3.5.1. Перевірити, чи правильні твердження:

1. Кожна функція розподілу ймовірностей на R^2 цілком визначається за двома функція розподілу ймовірностей на R^1.
2. Якщо F_x і F_y – функції розподілу ймовірностей на R^1, то за ними можна визначити єдину функцію розподілу ймовірностей на R^2.

3.5.2. Побудувати функцію $F(x, y)$ двохвимірного розподілу ймовірностей за допомогою двох функцій одновимірних розподілів ймовірностей:

$$F_1(x) = \begin{cases} 0, & \text{коли} \quad x \le 0; \\ x, & \text{коли} \quad 0 \le x \le 1; \\ 1, & \text{коли} \quad 1 \le x. \end{cases} \qquad F_2(y) = \begin{cases} 0, & \text{коли} \quad y \le 0; \\ y, & \text{коли} \quad 0 \le y \le 1; \\ 1, & \text{коли} \quad 1 \le y. \end{cases}$$

3.6. Ймовірність як нормована міра

Нехай на сукупності S підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, задано деяку міру $m(A)$, $A \in S$, таку, що $0 < m(\Omega) < \infty$. Тоді функція

$$P(A) = \frac{m(A)}{m(\Omega)}, \ A \in S, \tag{3.6.1}$$

задана на S, буде ймовірнісною мірою, оскільки стосовно такої функції задовільняються вимоги 1_p-3_p.

Міру $\dfrac{m(A)}{m(\Omega)}$ *називають нормованою мірою підмножин* $A \subset \Omega$, $A \in S$, *відносно* $m(\Omega)$.

В даному разі якщо в S існують дві множини $A \in S$, $B \in S$, міри яких однакові, тобто $m(A) = m(B)$, то і ймовірнісні міри таких множин, визначені за формулою 3.6.1 (ймовірності відповідних подій A і B), однакові.

Розглянемо простір подій S, породжений за сукупністю вимірних за мірою m підмножин H_1, H_2, …, H_k множини Ω таких, що $H_i H_j = \varnothing$, коли $i \neq j$, $H_1 + H_2 + ... + H_k = \Omega$, $m(H_i) = p > 0$ за всіх $i \in \overline{1, k}$, тобто $S = \{A \mid A = \bigcup_{i \in I} H_i, I \subset \{1, 2, ..., k\}\}$. Тоді до S входять разом з порожньою множиною $\varnothing$ всі множини H_i, всеможливі їх суми із 2-х, із 3-х, …, із $k-1$ доданків, а також сума k доданків $H_1 + H_2 + ... + H_k = \Omega$:

$$\begin{gathered} S = \{\varnothing, H_1, H_2, ..., H_k, H_1 + H_2, H_1 + H_3, ..., H_1 + H_k, \\ H_2 + H_3, ..., H_2 + H_k, ..., H_{k-1} + H_k, H_1 + H_2 + H_3, \\ H_1 + H_2 + H_4, ..., H_{k-2} + H_{k-1} + H_k, ..., \\ H_2 + H_3 + ... + H_{k-1} + H_k, H_1 + H_2 + ... + H_k = \Omega\}. \end{gathered}$$

Якщо ймовірнісну міру на такому просторі S визначено за формулою

$$P(A) = \frac{m(A)}{m(\Omega)}, \ A \in S,$$

тоді говорять, *що розподіл ймовірностей на множині* Ω (за множинами H_1, H_2, …, H_k) *рівномірний* в тому розумінні, що множинам $A \in S$ і $B \in S$ однакової міри $m(A)$ і $m(B)$ ставляться у відповідність однакові ймовірності $P(A)$ і $P(B)$.

Справді, нехай $m(A)=m(B)$, де: $A=\bigcup_{i\in I} H_i \in S$ і $B=\bigcup_{i\in J} H_i \in S$, $I\subset\{1, 2, ..., k\}$, $J\subset\{1, 2, ..., k\}$. Тоді $k(I)=k(J)$, де $k(I)$ та $k(J)$ – кількість елементів відповідно у множинах I та J – підмножинах індексів із множини $\{1, 2, ..., k\}$, а тому

$$P(A)=\frac{m(A)}{m(\Omega)}=\frac{m(\bigcup_{i\in I} H_i)}{m(\Omega)}=\frac{\sum_{i\in I} m(H_i)}{m(\Omega)}=\sum_{i\in I} P(H_i)=\frac{k(I)}{k}=$$

$$=\frac{m(B)}{m(\Omega)}=\frac{m(\bigcup_{i\in J} H_i)}{m(\Omega)}=\frac{\sum_{i\in J} m(H_i)}{m(\Omega)}=\sum_{i\in J} P(H_i)=\frac{k(J)}{k}=P(B).$$

Зокрема множини H_i можуть бути одноелементними, і тоді простір S буде найширшим з можливих.

Нехай, наприклад, $\Omega=\{"1", "2", "3", "4", "5", "6"\}$ (експеримент полягає у підкиданні шестигранного грального кубика),

$$H_1=\{"1", "2"\},\ H_2=\{"3", "4", "5", "6"\},$$

$$S=\{\varnothing, H_1, H_2, H_1+H_2\}=$$

$$=\{\varnothing,\{"1","2"\},\{"3","4","5","6"\},\{"1","2","3","4","5","6"\}\},$$

$$m(A)=k(A),\ P(A)=\frac{m(A)}{m(\Omega)}=\frac{k(A)}{k(\Omega)},\ A\in S.$$

Тоді $m(\varnothing)=0$, $m(\Omega)=6$, $m(\{"1","2"\})=2$, $m(\{"3","4","5","6"\})=4$, $P(\varnothing)=0$, $P(\Omega)=1$, $P(\{"1","2"\})=\frac{2}{6}$, $P(\{"3","4","5","6"\})=\frac{4}{6}$.

Даний розподіл ймовірностей не є рівномірним (за множинами H_1 і H_2). Разом з тим не виключено варіант, коли $m(H_1)=m(H_2)=3$. Тоді отримаємо рівномірний розподіл ймовірностей на Ω (за множинами H_1 і H_2).

Якщо із заданням ймовірнісної міри за рівністю $P(A)=\frac{k(A)}{k(\Omega)}$, $A\in S$, де $k(A)$ – кількість елементів у множині A, розглядається найширший простір подій, до якого входять всі підмножини скінченної множини Ω (порожня, одноелементні, двохелементні, триелементні, чотириелементні і т.д.), тоді виявиться, що ймовірність кожної елементарної події (яка ототожнюється з відповідною одноелементною множиною), дорівнює $\frac{1}{n}$, де $n = k(\Omega)$. Однак може трапитись, що до сукупності S входять не всі одноелементні множини і тоді питання про ймовірності всіх елементарних подій втрачають смисл, оскільки ймовірнісна міра $P(A)$ визначається лише на елементах сукупності S (сукупність S підмножин множини Ω є областю задання функції $P(A)$, $A \in S$).

Таким чином, коли ймовірнісна міра на S задається як $P(A) = \frac{k(A)}{k(\Omega)}$, $A \in S$, Ω – скінченна множина, $k(\Omega) = n$, через $P(A)$ можуть бути визначені ймовірності тільки тих елементарних подій, які є елементами S, однак не навпаки, ймовірність довільної події $A \in S$ не може бути визначена через ймовірності відповідних елементарних подій, оскільки відповідних одноелементних подій в S може не бути, а тому їх ймовірності можуть бути невизначені (невідомі).

Нагадаємо, що за означенням статистичної ймовірності

$$P_n^*(A) = \frac{k_n(A)}{k_n(\Omega)} = \frac{k_n(A)}{n}, \ A \in S,$$

тобто $P_n^*(A)$ є нормованою мірою підмножин $A \subset \Omega$, $A \in S$, відносно міри $m(\Omega) = k_n(\Omega)$.

Якщо в результаті великої серії із n випробувань, що полягали в підкиданні шестигранного грального кубика, в яких дослідника цікавило лише випадання на верхній грані кубика шести та п'яти очок, з'ясувалося, що $P_n^*(\{"6"\}) = 0.60$, $P_n^*(\{"5"\}) = 0.30$, $P_n^*(\{"1","2","3","4"\}) = 0.10$, то можна говорити про статистичні ймовірності попадання лише в множини із сукупності S:

$$S = \{\varnothing, H_1, H_2, H_3, H_1 + H_2, H_1 + H_3, H_2 + H_3, H_1 + H_2 + H_3 = \Omega\},$$

де

$$H_1 = \{"1","2","3","4"\}, \ H_2 = \{"5"\}, \ H_3 = \{"6"\}.$$

За суттю в даному експерименті лише три можливих наслідки: $H_2 = \{"5"\}$, $H_3 = \{"6"\}$, $H_1 = \Omega \setminus (\{"5"\} \cup \{"6"\}) = \{"1","2","3","4"\}$, тобто або $"5"$, або $"6"$, або і не $"5"$ і не $"6"$, хоча здається, що їх 6: $E_1 = "1"$, $E_2 = "2"$, $E_3 = "3"$, $E_4 = "4"$, $E_5 = "5"$, $E_6 = "6"$.

Очевидно $P_n^*(\varnothing) = 0$, $P_n^*(H_1) = P_n^*(\{"1", "2", "3", "4"\}) = 0.10$,

$P_n^*(H_2) = P_n^*(\{"5"\}) = 0.30$, $P_n^*(H_3) = P_n^*(\{"6"\}) = 0.60$,

$P_n^*(H_1 + H_2) = P_n^*(\{"1", "2", "3", "4", "5"\}) = 0.40$,

$$P_n^*(H_1 + H_3) = P_n^*(\{"1", "2", "3", "4", "6"\}) = 0.70\ ,$$
$$P_n^*(H_2 + H_3) = P_n^*(\{"5", "6"\}) = 0.90\ ,$$
$$P_n^*(H_1 + H_2 + H_3) = P_n^*(\{"1", "2", "3", "4", "5", "6"\}) = 1.0\ .$$

Однак питання про статистичні ймовірності попадання в будь-які інші підмножини множини Ω, окрім тих, що входять до S, наприклад в підмножини $\{"1", "3", "5"\}$, $\{"2", "4", "6"\}$, $\{"1", "2"\}$ тощо, а також питання про статистичні ймовірності елементарних подій $E_1 = "1"$, $E_2 = "2"$, $E_3 = "3"$, $E_4 = "4"$, та визначення через статистичні ймовірності елементарних подій статистичних ймовірностей всіх інших подій втрачають смисл. За даних умов знайти відповіді на такі питання неможливо.

Нехай множина Ω є лінійною, плоскою чи просторовою і вимірна, тобто задано міру $m(\Omega)$ множини Ω – довжину, площу, об'єм, масу тощо, причому $0 < m(\Omega) < +\infty$, а на сукупності S підмножин множини Ω, стосовно якої задовільняються вимоги 1_s-3_s, задана міра $m(A)$ підмножин $A \subset \Omega$, $A \in S$ (довжина, площа, об'єм, маса тощо). Тоді за формулою $P(A) = \dfrac{m(A)}{m(\Omega)}$, де $A \in S$, задається ймовірнісна міра (ймовірність) на S, причому коли в S є дві множини $A \in S$ і $B \in S$ однакової міри, тобто $m(A) = m(B)$, то і імовірності попадання в такі множини однакові, тобто $P(A) = P(B)$, $A \in S$, $B \in S$ (Рис. 3.6.1). *Таке задання ймовірнісної міри називається геометричним*.

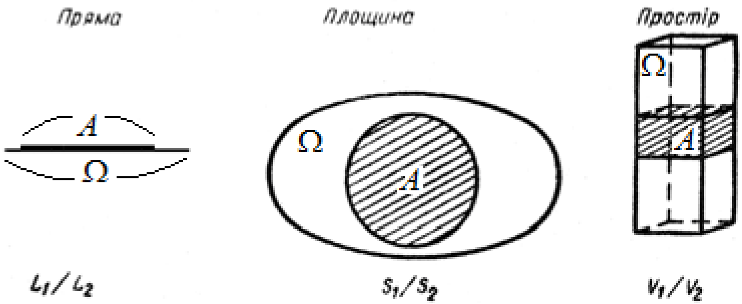


Рис. 3.6.1

Приклад 3.6.1. (Задача про зустріч) Двоє домовилися зустрітися протягом години (між 18^{00} і 19^{00}) і чекати один одного не довше, ніж 15 хв. Потрібно знайти імовірність того, що вони зустрінуться, якщо імовірності попадання моментів приходу кожного в будь які проміжки часу між 18^{00} і 19^{00} однакової довжини однакові, незалежно від того, коли приходить інший, тобто однаково імовірно, наприклад, що перший з них прийде до місця зустрічі на проміжку часу 18^{00} – 18^{01}, чи на проміжку часу 18^{20} – 18^{21}, чи на проміжку часу 18^{59} – 19^{00}, чи на будь якому іншому проміжку тієї самої довжини, незалежно від того, коли приходить інший. Те саме стосується і іншої особи. Зауважимо, що проміжки часу можна розглядати як довші за 1 хв, так і коротші (можливо коротші і за 1 сек).

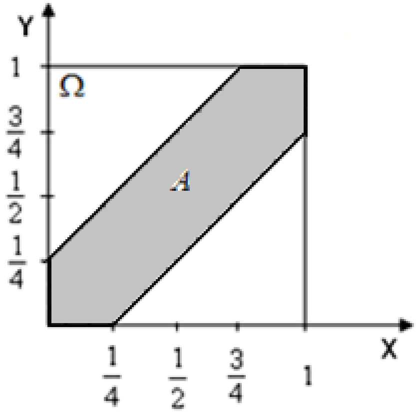


Рис. 3.6.2

Якщо через x позначити час, що пройде після 18^{00} до моменту приходу першої особи до місця зустрічі, через y – час після 18^{00} до моменту приходу другої особи до місця зустрічі, то в результаті кожного такого експерименту дістанемо пару чисел (x, y), тобто точку із 2-вимірного простору (точку на площині), причому стосовно координат (абсциси x і ординати y) цієї точки очевидно задовільняються вимоги $0 \le x \le 1$, $0 \le y \le 1$ (за одиницю вимірювання довжини проміжків часу обрано проміжок часу довжиною 1 година, початок відліку – 18^{00}). Через множину всіх пар (x, y), $x \in [0;1]$, $y \in [0;1]$ заповнюється квадрат $\Omega = [0;1] \times [0;1]$ (Рис. 3.6.2). За умовою задачі згадані особи зустрінуться, якщо різниця між моментами приходу кожного з них до місця зустрічі за абсолютною величиною не перевищуватиме $\frac{1}{4}$ години (15 хв.), тобто якщо виконуватиметься умова $|y - x| \le \frac{1}{4}$, або $x - \frac{1}{4} \le y \le x + \frac{1}{4}$. На рис. 3.6.2 заштрихована множина точок, стосовно координат яких задовільняється умова $|y-x|\le\frac{1}{4}$ (чи, що те саме, $|x-y|\le\frac{1}{4}$).

Нехай проміжок $\Delta x\subset[0;1]$, проміжок $\Delta y\subset[0;1]$. Тоді прямокутник $\Delta x\times\Delta y\subset[0;1]\times[0;1]=\Omega$. Природно покласти $\tilde{m}(\Delta x\times\Delta y)=m(\Delta x)\cdot m(\Delta y)$, тобто $\tilde{m}(\Delta x\times\Delta y)$ – це площа прямокутника $\Delta x\times\Delta y$. Якщо момент приходу першого попадає на фіксований проміжок Δx_1, то однаково імовірно, що момент приходу другого буде знаходитись на проміжку Δy_1 чи Δy_2, якщо міри (довжини) проміжків Δy_1 і Δy_2 однакові, і ймовірності попадання точки (x,y) у прямокутники $\Delta x_1\times\Delta y_1$ і $\Delta x_1\times\Delta y_2$ однакові. Те саме стосується і фіксованого проміжка Δx_2. Аналогічно, якщо міри (довжини) проміжків Δx_1 і Δx_2 однакові, а Δy_1 – фіксований проміжок, то і ймовірності попадання точки (x,y) в прямокутники $\Delta x_1\times\Delta y_1$ і $\Delta x_2\times\Delta y_1$ однакові. Таким чином, якщо міри (площі) прямокутників $\Delta x_1\times\Delta y_1$ і $\Delta x_2\times\Delta y_2$ однакові (за умови, що міри проміжків Δx_1 і Δx_2 однакові, а також однакові міри проміжків Δy_1 і Δy_2), то і ймовірності попадання точки (x,y) в такі прямокутники однакові. Легко бачити, що коли міри (площі) прямокутників $\Delta x_1\times\Delta y_1$ і $\Delta x_2\times\Delta y_2$ однакові, то і ймовірності попадання в такі прямокутники однакові і тоді, коли можливо $m(\Delta y_1)\ne m(\Delta y_2)$.

Таким чином імовірнісна міра $\tilde{P}$ на сукупності $\tilde{S}$ підмножин множини Ω, які вимірні за мірою $\tilde{m}$ (площу яких можна визначити), може бути задана за формулою

$$\tilde{P}(A)=\frac{\tilde{m}(A)}{\tilde{m}(\Omega)},\quad A\in\tilde{S}.$$

Оскільки в даному прикладі

$$\tilde{m}(\Omega)=\tilde{m}([0;1]\times[0;1])=m([0;1])\cdot m([0;1])=1\cdot 1=1,$$

а

$$\tilde{m}(A)=1-(3/4)^2=\frac{7}{16},$$

то шукана ймовірність (попадання точки (x,y) в заштриховану область на рис. 3.6.2) дорівнює

$$\tilde{P}(A)=\frac{\tilde{m}(A)}{\tilde{m}(\Omega)}=\frac{(1-(3/4)^2)}{1}=\frac{7}{16}.$$

Зауважимо, що коли $A \subset \Omega$ – довільна вимірна за мірою m підмножина множини Ω, S – сукупність всіх вимірних за мірою m підмножин множини Ω, $P(A) = \dfrac{m(A)}{m(\Omega)}$, $0 < m(\Omega) < \infty$, $m(A)$, $A \in S$, є міра Лебега, то такий розподіл ймовірностей на множині Ω називають *рівномірним*, а елементарні події (наслідки випробування) вважаються рівноімовірними (в такому разі ймовірність кожної елементарної події дорівнює нулеві).

Приклад 3.6.2. На відрізку OQ завдовжки l навмання вибирають дві точки M і N. Знайти ймовірність того, що точка M буде від точки N не далі, ніж від точки O, якщо всі точки відрізка OQ вважаються рівноімовірними.

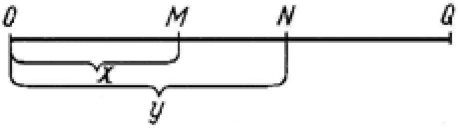


Рис. 3.6.3

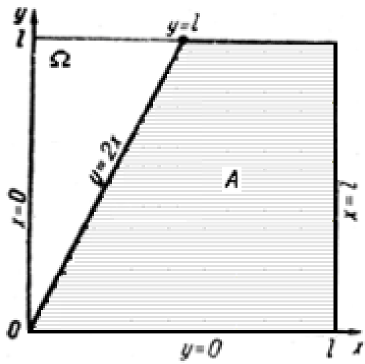


Рис. 3.6.4

Якщо позначити через x відстань від точки O до M, а через y – відстань від точки O до N (Рис. 3.6.3), то після кожного вибору M і N дістанемо пару чисел (x, y), стосовно яких задовільняються умови $0 \le x \le l$, $0 \le y \le l$. Крім того x і y набувають одного з можливих значень на відрізку $[0; l]$ незалежно одне від одного. Множина всіх можливих пар (x, y) – це множина точок квадрата Ω із стороною завдовжки l (Рис. 3.6.4), тобто $\Omega = \{(x, y) \mid 0 \le x \le l, 0 \le y \le l\}$.

В даному разі немає підстав вважати, що можливість попадання в деяку підмножину із Ω певної міри більша або менша, ніж можливість попадання в будь-яку іншу підмножину тієї самої міри, тобто розподіл ймовірностей в квадраті Ω слід вважати рівномірним (а точки квадрата Ω вважати рівноімовірними).

Оскільки кожній точці квадрата Ω відповідає елементарна подія, яка полягає в тому, що навмання вибраною виявляється саме ця точка, можна ототожнити множину точок квадрата з множиною Ω всіх елементарних подій в даному експерименті.

Відстань від точки M до N є $|x-y|$, а відстань від точки O до $M \in x$. Отже, точка M буде ближче до точки N, ніж до точки O, якщо виконується нерівність

$$|x-y| \le x,$$

або система нерівностей

$$-x \le x-y, \ \ x-y \le x,$$

тобто

$$y \ge 0, \ \ y \le 2x.$$

Перша нерівність системи тривіальна, вона виконується в кожній точці квадрата і є "зайвою", бо в ній повторюється одна з вимог $y \ge 0,$ за якими визначаються точки, що лежать в квадраті. З другої нерівності випливає, що точка M лежить ближче до N, ніж до O, якщо $y \le 2x$.

Отже, подію A, яка полягає в тому, що точка M виявиться віддаленою від точки N не більше, ніж від точки O, можна описати так:

$$A = \{(x,y) \mid 0 \le x \le l, 0 \le y \le l, \ y \le 2x\}$$

(на Рис. 3.6.4 множину A затемнено).

Прямі $y = 2x$ і $y = l$ перетинаються в точці з абсцисою $x = l/2$, тому

$$m\,(\Omega) = l^2, \ \ m\,(A) = l^2 - \frac{1}{2}l^2 \cdot \frac{1}{2} = \frac{3}{4}l^2.$$

Отже,

$$P\,(A) = \frac{m\,(A)}{m\,(\Omega)} = \frac{3}{4}.$$

Приклад 3.6.3. Між 0 і 1 навмання вибирається два числа. Знайти ймовірність того, що добуток логарифмів цих чисел не більший за 1, а добуток різниць між 1 і цими числами перевищує 0.1.

Позначимо перше число через x, а друге через y. Тоді в результаті кожного випробування дістанемо пару (x,y), причому множиною всіх таких пар є

$$\Omega = \{(x,y) \mid \ x \in [0;1], \ y \in [0;1]\}.$$

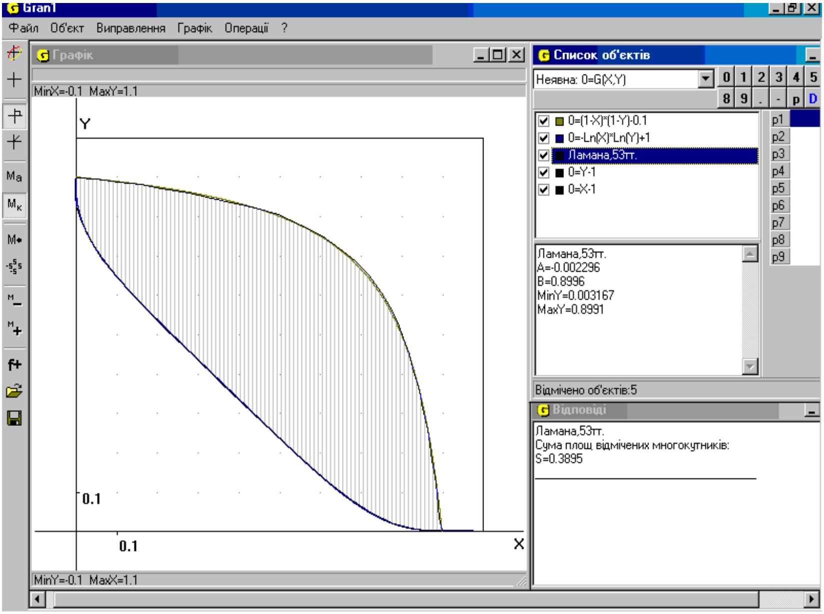


Рис. 3.6.5 а

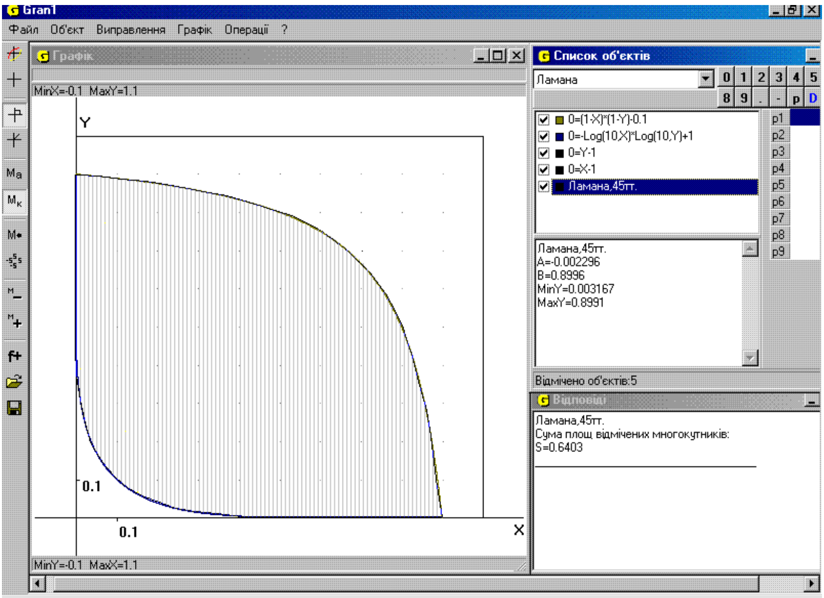


Рис. 3.6.5 б

Всі точки множини Ω вважатимемо рівноімовірними, тобто розподіл ймовірностей на множині Ω вважатимемо рівномірним.

Множиною пар (x, y), за якими визначається вказана подія, є

$$A = \{(x, y) \mid x \in [0; 1], y \in [0; 1], (1-x)(1-y) \geq 0.1, \ln x \cdot \ln y \leq 1\}$$

(на Рис. 3.6.5 а множину A заштриховано).

Використовуючи послуги програми GRAN1, знайдемо точки перетину кривих $y = 1 + \dfrac{0.1}{x-1}$ і $y = e^{\frac{1}{\ln x}}$.

Далі знайдемо площу $m(A)$ області, обмеженої зазначеними кривими (заштрихованої області на Рис. 3.6.5 а), вписавши в контур множини A замкнену ламану з 53 вершинами, розташованими приблизно на однаковій відстані одна від одної. Далі за допомогою послуги «Площа многокутника» програми GRAN1 знайдемо наближено площу між вказаними кривими $S \approx 0.39$ (див. Рис. 3.6.5а). Оскільки $m(\Omega) = 1$, то шукана ймовірність дорівнює (див. Рис. 3.6.5а)

$$P(A) = \frac{m(A)}{m(\Omega)} = 0.39.$$

Якщо замість натуральних розглядати десяткові логарифми, то відповідний результат подано на Рис. 3.6.5 б.

Приклад 3.6.4. У крузі радіуса R навмання вибирається хорда. Знайти ймовірність того, що довжина цієї хорди буде не більша від R.

Слід зауважити, що розглядувана задача не є цілком визначеною, оскільки не сказано, як саме вибирається хорда. Розв'язки цієї задачі можуть бути різними залежно від способу вибирання хорди, через що визначається множина Ω всіх можливих наслідків випробування, а також множина тих наслідків, коли довжина хорди буде не більша, ніж довжина радіуса круга.

1. Нехай хорди вибираються паралельними до заданого напрямку (Рис. 3.6.6). Тоді досить вибрати точку на діаметрі MN, перпендикулярному до напрямку хорд, щоб визначити хорду. Множина Ω всіх можливих наслідків експерименту еквівалентна множині точок відрізка MN, будь-які точки діаметра MN вважаються рівноімовірними. Згідно з Рис. 3.6.6, подія A, яка полягає в тому, що довжина хорди буде не більша, ніж R, відбудеться, якщо навмання вибрана точка лежатиме на одному з відрізків MP або NQ, де за точками P і Q визначаються хорди довжини R.

Легко бачити, що

$$m(A) = 2(R - R\cos(\pi/6)) = 2R(1 - \sqrt{3}/2).$$

Таким чином,

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{2R(1-\sqrt{3}/2)}{2R} = 1-\sqrt{3}/2\ .$$

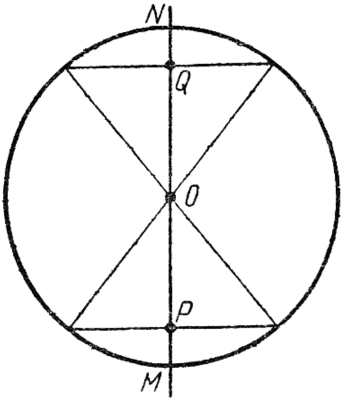


Рис. 3.6.6

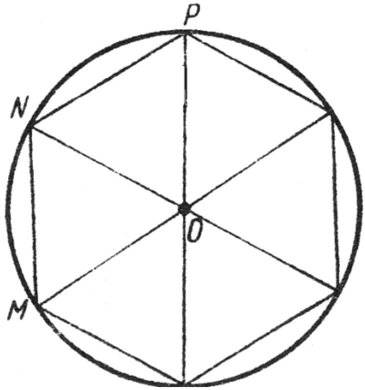


Рис. 3.6.7

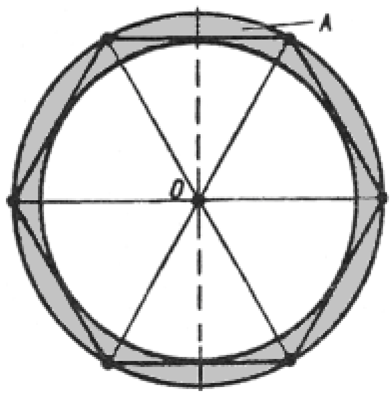


Рис. 3.6.8

2. Нехай один кінець хорди, наприклад N, закріплений, а положення другого кінця рівноможливе в будь-якій точці кола (Рис. 3.6.7).

В такому разі множиною всіх можливих наслідків експерименту, рівноімовірних між собою, вважається множина Ω всіх точок кола, яким обмежується круг, а множиною точок, відповідною розглядуваній події A, є множина точок дуги MNP, де $MN = NP = R$.

Очевидно, що

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{1/3 \cdot 2\pi R}{2\pi R} = \frac{1}{3}\ .$$

3. Хорда цілком визначається, якщо задати її середину (Рис. 3.6.8). У цьому випадку множині Ω всіх можливих наслідків експерименту відповідає множина точок круга радіуса R, а події A відповідає множина точок всередині даного круга і зовні круга, концентричного даному, з радіусом $(\sqrt{3}/2)R$ (див. випадок 1). Всі точки круга вважаються рівноімовірними. Отже дістаємо

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{\pi R^2 - (3/4)\pi R^2}{\pi R^2} = \frac{1}{4}.$$

4. Нехай обидва кінці хорди вибираються на колі навмання. Якщо на колі вибрати точку відліку і відстань вздовж кола від цієї точки вимірювати проти руху стрілки годинника (Рис. 3.6.9, а), то внаслідок кожного випробування дістанемо пару точок (x, y), причому множиною всіх таких пар є $\Omega = \{(x, y) \mid x \in [0;\ 2\pi R],\ y \in [0;\ 2\pi R]\}$ (Рис. 3.6.9, б). Всі точки множини Ω вважаються рівноімовірними.

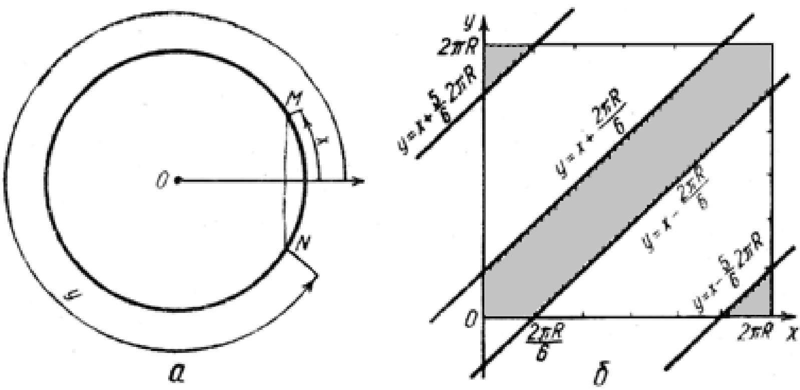


Рис. 3.6.9

В даному разі хорда буде не довша за радіус, якщо $|x - y| \le \frac{2\pi R}{6}$ або $2\pi R - |x - y| \le \frac{2\pi R}{6}$ (див. Рис. 3.6.9, а). На Рис. 3.6.9, б зображено множину точок (вона затемнена), за якими визначається подія A, яка полягає в тому, що хорда коротша, ніж радіус кола. Очевидно, що

$$P(A) = 1 - \left(\frac{5}{6}\right)^2 + \left(\frac{1}{6}\right)^2 = \frac{12}{36} = \frac{1}{3}.$$

5. Нехай в крузі навмання вибирають точку M і через неї в довільно взятому напрямі проводять хорду. Якщо напрям хорди визначати за перпендикуляром до неї, а точку, через яку проводять хорду, визначати за полярними координатами, то хорда визначається за трьома параметрами: ρ – полярний радіус точки, через яку проводять хорду; φ – полярний кут радіуса–вектора точки; α – кут, утворений перпендикуляром до хорди і радіусом-вектором точки (Рис. 3.6.10).

Таким чином,

$$\Omega = \left\{ (\rho, \varphi, \alpha) \,\middle|\, 0 \le \rho \le R,\ 0 \le \varphi \le 2\pi,\ \varphi - \frac{\pi}{2} \le \alpha \le \varphi + \frac{\pi}{2} \right\}.$$

Множиною точок Ω, за якими визначається подія A, яка полягає в тому, що хорда коротша за радіус, є множина

$$A = \left\{ \begin{array}{l} (\rho,\ \varphi,\ \alpha): R\sqrt{3/2} \le \rho \le R, \\ \qquad\qquad 0 \le \varphi \le 2\pi, \\ \qquad\qquad \varphi - \arccos \dfrac{R\sqrt{3/2}}{\rho} \le \alpha \le \varphi + \arccos \dfrac{R\sqrt{3/2}}{\rho} \end{array} \right\}.$$

Всі точки множини Ω вважаються рівноімовірними.

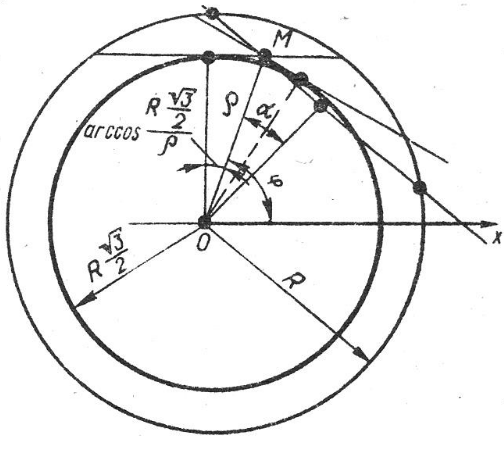


Рис. 3.6.10

Якщо вибрати тривимірну прямокутну систему координат $(\rho,\ \varphi,\ \alpha)$, то множини Ω і A можна подати так, як показано на Рис. 3.6.11 (множині Ω поставлена у відповідність фігура $EFGHH_1E_1F_1G_1$, а множині A – фігура $ABCC_1A_1B_1$). Очевидно, що фігури $EFGHE_1F_1G_1H_1$ і $EFGHE_2F_2G_2H_2$ рівновеликі, оскільки в них одна й та сама площа основи й однакові висоти. Так само рівновеликі фігури $ABCC_1A_1B_1$ і $ABCC_2A_2B_2$. Звідси об'єм фігури $ABCC_1A_1B_1$ відноситься до об'єму фігури $EFGHE_1F_1G_1H_1$ так, як площа фігури ABC до площі фігури $EFGH$ (Рис. 3.6.12).

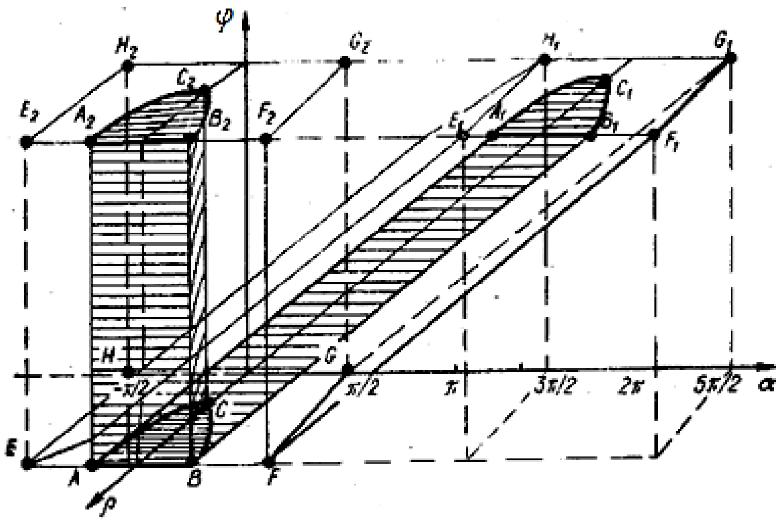


Рис. 3.6.11

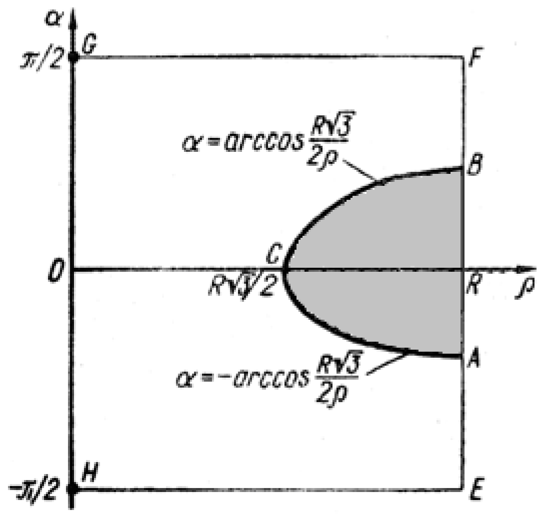


Рис. 3.6.12

Отже,

$$\frac{S_{ABC}}{S_{EFGH}} = \frac{1}{\pi R}\ 2\int\limits_{\frac{R\sqrt{3}}{2}}^{R} \arccos\left(\frac{R\sqrt{3}}{2\rho}\right) d\rho\,.$$

Після заміни змінних

$$\frac{R\sqrt{3}}{2\rho}=\frac{1}{t}, \quad \left(t=\frac{2\rho}{R\sqrt{3}}, \ \rho=\frac{R\sqrt{3}}{2}t, \ d\rho=\frac{R\sqrt{3}}{2}dt\right)$$

дістанемо

$$P(A)=\frac{2}{\pi R}\int_1^{2/\sqrt{3}}\arccos\left(\frac{1}{t}\right)\frac{R\sqrt{3}}{2}dt=\frac{\sqrt{3}}{\pi}\int_1^{2/\sqrt{3}}\operatorname{arctg}(\sqrt{t^2-1})\,dt\,.$$

Позначивши $t^2-1=s^2$, $\left(t^2=1+s^2,\ t=\sqrt{1+s^2},\ dt=\frac{sds}{\sqrt{1+s^2}}\right)$,

знайдемо

$$P(A)=\frac{\sqrt{3}}{\pi}\int_0^{1/\sqrt{3}}\operatorname{arctg} s\,\frac{sds}{\sqrt{1+s^2}}\,.$$

Після інтегрування частинами

$$\left(u=\operatorname{arctg} s,\ du=\frac{ds}{1+s^2};\ \frac{sds}{\sqrt{1+s^2}}=dv,\ v=\sqrt{1+s^2}\right)$$

одержимо

$$\frac{\sqrt{3}}{\pi}\left(\sqrt{1+s^2}\operatorname{arctg} s\Big|_0^{1/\sqrt{3}}-\int_0^{1/\sqrt{3}}\frac{\sqrt{1+s^2}}{1+s^2}\,ds\right)=$$

$$=\frac{\sqrt{3}}{\pi}\left(\sqrt{1+\frac{1}{3}}\ \operatorname{arctg}\frac{1}{\sqrt{3}}-\int_0^{1/\sqrt{3}}\frac{ds}{\sqrt{1+s^2}}\right)=\frac{\sqrt{3}}{\pi}\left(\frac{2}{\sqrt{3}}\ \frac{\pi}{6}-\right.$$

$$\left.-\ln(s+\sqrt{1+s^2})\Big|_0^{1/\sqrt{3}}\right)=\frac{1}{3}-\frac{\sqrt{3}}{\pi}\ln\left(\frac{1}{\sqrt{3}}+\sqrt{1+\frac{1}{3}}\right)=$$

$$=\frac{1}{3}-\frac{\sqrt{3}}{\pi}\ \ln\left(\frac{3}{\sqrt{3}}\right)=\frac{1}{3}-\frac{\sqrt{3}}{2\pi}\ln 3\approx 0.0305\,.$$

Обчислити

$$\frac{2}{\pi R}\int_1^{2/\sqrt{3}}\arccos\left(\frac{1}{t}\right)\frac{R\sqrt{3}}{2}\,dt \ \text{ або } \ \frac{\sqrt{3}}{\pi}\int_1^{2/\sqrt{3}}\operatorname{arctg}(\sqrt{t^2-1})\,dt$$

можна за допомогою програми GRAN1 (Рис. 3.6.13, а, б). В результаті дістанемо, як і раніше,

$$P\ (A) \approx 0.0305\,.$$

Наближений розв'язок цієї задачі можна знайти ще й так.

Вибираємо навмання (за допомогою генератора випадкових чисел з рівномірним розподілом імовірностей на проміжку $[a;b)$) щоразу три випадкових числа: перше число – в межах $[0;\,R]$ і вважаємо його значенням ρ, друге – в межах $[0;\,2\pi]$ і вважаємо його значенням φ, третє – в межах $[\varphi-\frac{\pi}{2};\varphi+\frac{\pi}{2}]$ і вважаємо його значенням α. Якщо виявиться $R\frac{\sqrt{3}}{2}\le\rho\le R$ і

$$\varphi-\arccos\frac{R\sqrt{3}}{2\rho}\le\alpha\le\varphi+\arccos\frac{R\sqrt{3}}{2\rho}\,,$$

то вибрана хорда буде коротша за радіус. Здійснюючи досить велику серію випробувань і знайшовши відношення числа випробувань, в яких подія A відбулася, до числа всіх випробувань, дістанемо наближено шукану ймовірність

$$P(A)\approx P_n^*(A)\,.$$

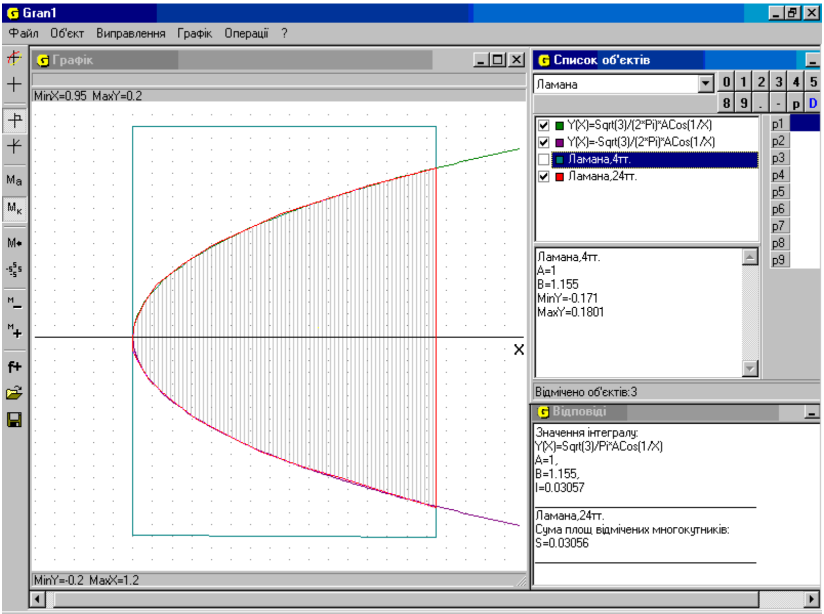


Рис. 3.6.13, а

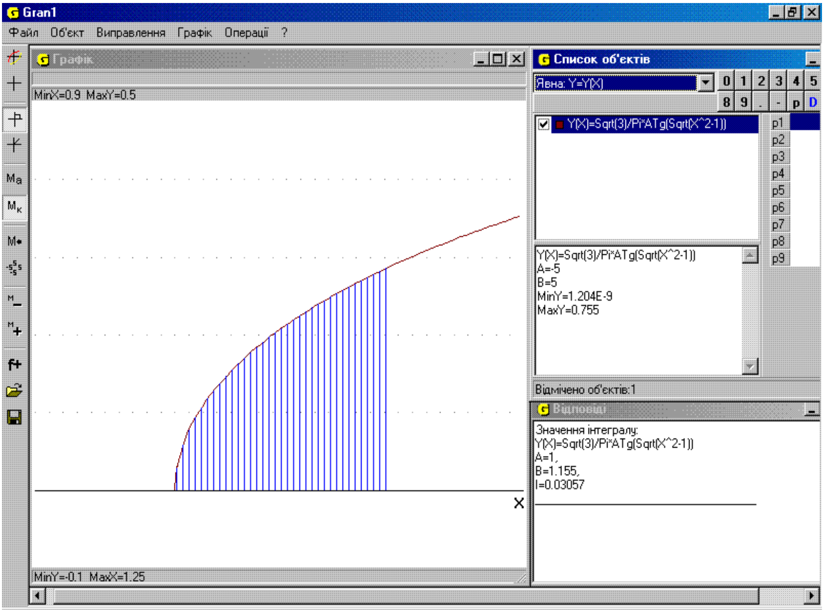


Рис. 3.6.13, б

Наведемо варіант програми для реалізації описаних випробувань, а також результати, здобуті за цією програмою коли кількість випробувань дорівнює відповідно 100000, 500000, 1000000.

```
Uses Crt
function ArcCos (X: Real): Real;
      begin
            ArcCos:=ArcTan (Sqrt (1-Sqr (X))/X);
      end;
var
          R, Alpha, Ro, Phi, W: Real;
           I, J, K: LongInt;
begin
           ClrScr;
           Write ('Input radius:');
           Readln (R);
           Write ('Count of steps:');
           Readln (J);
           Writeln;
           Randomize;
           K: =0;
             for I: =1 to J do
                    begin
                           Ro: =Random *R;
                           Phi: =Random *Pi *2;
```

```
                    Alpha: =Phi – Pi/2+Random *Pi;
            W: =ArcCos (R *Sqrt (3)/(2 *Ro));
                    if (Ro) = (R*Sqrt (3)/2)) and (Ro < R) and
                    (Alpha) =Phi – W) and (Alpha (=Phi+W)
                    then Inc (K);
                end;
    GotoXY (1, 4);
            Writeln;
            Writeln (' R=', R);
            Writeln (' Step=', J);
            Writeln (' P (A) =', K/J);
            Writeln;
             Readin
    end.
    R = 1.0000000000E+00
    Step = 100000
    P (A) = 3.1082000000E-02
    R = 1.0000000000E+00
    Step = 500000
    P (A) = 3.0820000000E-02
    R = 1.0000000000E+00
    Step = 1000000
    P (A) = 3.0521000000E-02
```

Зауваження. Як бачимо, чим більша серія випробувань, тим ближче наближений результат до здобутого раніше “точного”.

Здавалося б, задача одна й та сама, а відповіді у кожному випадку різні – так званий парадокс Бертрана за ім’ям французького математика Жозефа Бертрана, який опублікував цей “парадокс” у 1907р. Проте парадоксу тут немає. Все залежить від того, що розуміють під словами “певним способом вибирається хорда”. Залежно від того, як визначено множину всіх можливих наслідків випробування, рівноімовірних між собою, дістають різні задачі – з різними множинами Ω і A. Цілком природно, що й імовірності відповідних подій різні. Слід зазначити, що можуть бути й задачі із скінченним числом можливих наслідків експерименту, де можна неоднозначно тлумачити на перший погляд одне й те саме питання. Тому розв’язуючи задачу, треба спочатку чітко визначити, що слід вважати елементарними наслідками експерименту, що вважати подіями та їх ймовірностями, а потім вже досліджувати задачу.

Вправи для самостійного виконання

3.6.1. Перевірити, чи правильні твердження:

1. Стосовно кожної вимірної за Лебегом множини Ω можна побудувати ймовірнісний простір (Ω, S, P), в якому ймовірність подій $A \in S$ задана геометрично $P(A) = \dfrac{m(A)}{m(\Omega)}$.

2. Геометрично задати ймовірність можна лише за допомогою міри Лебега.

3. Сукупність підмножин відрізка $[0;1]$, вимірних за Лебегом, можна вважати простором подій.

4. Якщо в разі геометричного задання ймовірності мірою є міра Лебега-Стільтьєса, то розподіл ймовірностей залишається рівномірним.

5. Геометричне задання ймовірності є однаковим за формою, проте може бути різним за суттю, залежною від того, що вважати простором Ω елементарних подій і що вважати простором подій.

6. Парадокс Бертрана свідчить про неможливість застосування теорії ймовірностей до розв'язування практичних задач.

7. Ймовірність, задану геометрично за допомогою міри Лебега підмножин відрізка $[0;1]$, можна продовжити на "найширшу" σ-алгебру всеможливих підмножин відрізка $[0;1]$.

3.7. Властивості ймовірнісної міри

Нехай розглядається деякий стохастичний експеримент, Ω – відповідний простір елементарних подій, і задано систему S підмножин множини Ω, яка є σ-алгеброю, тобто виконуються умови:

1_s. $\Omega \in S$;

2_s. Якщо $A \in S$, то $\overline{A} \in S$;

3_s. Якщо $A_i \in S$, $i = 1, 2, \ldots$, то і $\bigcup\limits_i A_i \in S$.

Будь-яку числову функцію $P(A)$, визначену на σ-алгебрі S підмножин множини Ω, називають ймовірнісною мірою або *ймовірністю*, якщо виконуються такі умови (порівняйте з §1.7):

1_p. $P(A) \geq 0$ за будь-якого $A \in S$.

2_p. Якщо $A_i \in S$, $i = 1, 2, \ldots$, такі, що $A_i \bigcap A_j = \varnothing$, коли $i \neq j$, то $P\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} P(A_i)$.

3_p. $P(\Omega) = 1$.

В такому разі елементи сукупності S називаються подіями, а числа $P(A)$, $A \in S$, – ймовірностями цих подій.

Умови $1_s - 3_s$, $1_p - 3_p$ називають системою аксіом теорії ймовірностей (системою аксіом А.М. Колмогорова). Ця система аксіом несуперечлива, бо існують реальні об’єкти, стосовно яких вони задовільняються. Вона не є категоричною, бо на даному вимірному просторі можна різними способами задати ймовірність, що дає змогу охопити всі можливі ситуації.

Властивості $1_p - 3_p$ *називають основними або визначальними* властивостями ймовірності.

Встановимо деякі інші властивості ймовірності, які випливають з основних.

4. За будь-якої події $A \in S$

$$P(\overline{A}) = 1 - P(A).$$

Оскільки $A \bigcup \overline{A} = \Omega,\ A \bigcap \overline{A} = \varnothing$, то за аксіомами 2_p і 3_p

$$1 = P(\Omega) = P(A \bigcup \overline{A}) = P(A) + P(\overline{A}).$$

Тому

$$P(\overline{A}) = 1 - P(A).$$

5. Імовірність неможливої події дорівнює нулю:

$$P(\varnothing) = 0.$$

Справді, $\varnothing = \overline{\Omega}$. Тоді за властивістю 4

$$P(\varnothing) = 1 - P(\Omega) = 0.$$

6. Якщо через подію $A \in S$ спричиняється подія $B \in S$, тобто $A \subset B$, то

$$P(A) \le P(B).$$

Справді, якщо $A \subset B$, то $B = A + \overline{A}B$, і за властивостями 1_p і 2_p

$$P(B) = P(A + \overline{A}B) = P(A) + P(\overline{A}B) \ge P(A).$$

7. За будь-якої випадкової події $A \in S$

$$0 \le P(A) \le 1.$$

Справді, за будь-якої події $A \in S$ має місце входження:

$$A \subset \Omega,$$

а тому за властивостями 1_p, 6 і 3_p

$$0 \le P(A) \le P(\Omega) = 1,$$

тобто

$$0 \le P(A) \le 1.$$

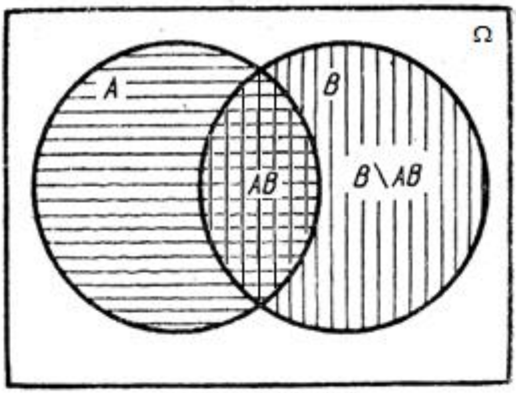
8. За будь-яких двох подій $A \in S$ і $B \in S$

$$P(A + B) = P(A) + P(B) - P(AB).$$

Справді (Рис. 3.7.1),

$$A + B = A + (B \setminus AB),$$

$$B = AB + (B \setminus AB),$$



$A + B = A + (B \setminus AB)$

Рис. 3.7.1

причому доданки в правих частинах останніх рівностей є несумісними подіями. Тому за аксіомою 2_p

$$P(A+B)=P(A)+P(B\setminus AB) \text{ і } P(B)=P(AB)+P(B\setminus AB) .$$

Звідси

$$P(B\setminus AB)=P(B)-P(AB) .$$

Тому

$$P(A+B)=P(A)+P(B)-P(AB) ,$$

що й треба було довести.

Стосовно суми трьох подій $A+B+C$ одержуємо:

$$P(A+B+C)=P((A+B)+C)=P(A+B)+P(C)-P((A+B)C)=$$

$$=P(A)+P(B)-P(AB)+P(C)-P(AC+BC)=$$

$$=P(A)+P(B)+P(C)-P(AB)-(P(AC)+P(BC)-P(AC\cdot BC))=$$

$$=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC) .$$

Використовуючи метод математичної індукції, можна довести, що за довільного скінченного числа випадкових подій $A_1, A_2, ..., A_n$

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)-\sum_{1\le i<j\le n} P(A_i A_j)+\sum_{1\le i<j<k\le n} P(A_i A_j A_k)-...+$$

$$+(-1)^{n-1} P(A_1 A_2 ... A_n) .$$

Нагадаємо, що коли події A_i, $i=1, 2, ..., n$, попарно несумісні, тобто $A_i A_j=\varnothing$, коли $i\ne j$, то ймовірність суми будь-якої кількості подій дорівнює сумі їх імовірностей (аксіома 2_p).

Приклад 3.7.1. На тролейбусній зупинці протягом 10 хв. зупиняється тільки один тролейбус маршрутів № 5, 6, 8, 9, або 17. Знайти ймовірність того, що протягом 10 хв. підійде тролейбус № 5 або № 6, якщо відомо, що тролейбуси № 5 протягом 10 хв. з’являються на зупинці з імовірністю $P(''5'')=p_1$, тролейбуси № 6 – з імовірністю $P(''6'')=p_2$, тролейбуси № 8 – з імовірністю $P(''8'')=p_3$, тролейбуси № 9 – з імовірністю $P(''9'')=p_4$, тролейбуси № 17 – з імовірністю $P(''17'')=p_5$.

В даному прикладі події $\{''5''\}$ і $\{''6''\}$ означають, що протягом 10 хв. на зупинці зупиняться тролейбуси відповідно маршрутів № 5 та № 6.

Тоді A – випадкова подія, яка полягає в тому, що протягом 10 хв. на зупинку прийде тролейбус № 5 або тролейбус № 6, тобто $A = \{"5"\} + \{"6"\} = \{"5", "6"\}$. Тому, враховуючи несумісність подій $\{"5"\}$ та $\{"6"\}$, за властивістю аддитивності імовірності одержуємо:

$$P(A) = P("5") + P("6") = p_1 + p_2.$$

Нехай $\Omega = \{E_1, E_2, ..., E_n\}$ – скінченна множина n рівноможливих наслідків експерименту, S – сукупність всіх підмножин множини Ω, ймовірнісна міра (ймовірність) на S задана так, що кожному E_i приписана ймовірність $P(E_i) = \frac{1}{n}$. Якщо A – випадкова подія, що складається з m елементарних подій, $A \subset \Omega$, $A = (E_{i_1}, E_{i_2}, ..., E_{i_m})$, тоді

$$P(A) = \sum_{E_i \in A} p_i = \sum_{E_i \in A} (1/n) = \frac{m}{n} = \frac{m(A)}{m(\Omega)}.$$

В такому разі якщо в двох підмножинах A і B множини Ω міститься однакова кількість елементів (міри $m(A)$ і $m(B)$ множин A і B однакові, де $m(A) = k(A)$ – кількість елементів у множині A, $m(B) = k(B)$ – кількість елементів у множині B), то і ймовірності подій A і B однакові – $P(A) = P(B)$. В такому разі говорять, що розподіл імовірностей на скінченній множині Ω рівномірний (порівняйте з §3.6).

Приклад 3.7.2. Монету підкидають двічі. Простір Ω елементарних подій цього експерименту складається з чотирьох елементів:

$$\Omega = \{ГГ, ГЦ, ЦГ, ЦЦ\}.$$

Припустимо, що кожна елементарна подія відбувається з ймовірністю $p_i = \frac{1}{4}$, $i = 1, 2, 3, 4$. Нехай A – подія, яка полягає в тому, що в результаті першого підкидання випав герб, подія B – в результаті другого підкидання випав герб. Тоді

$$A = \{ГГ, ГЦ\},\ B = \{ГГ, ЦГ\},\ AB = \{ГГ\},\ A + B = \{ГГ, ГЦ, ЦГ\},$$

$$P(A) = P(ГГ) + P(ГЦ) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

$$P(B) = P(ГГ) + P(ЦГ) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Розглянемо подію C: хоча б один раз випав герб. Очевидно, що

$$C = A + B = \{ГГ, ЦГ\} + \{ГГ, ГЦ\} = \{ГГ, ГЦ, ЦГ\},\ P(C) = 3/4.$$

Приклад 3.7.3. З n виробів, серед яких m бракованих, навмання вибирають k виробів. Знайти ймовірність того, що серед цих k виробів рівно r бракованих.

З умови задачі випливає, що виконуються нерівності $r \le k,\ r \le m,\ k - r \le n - m.$

Наслідки експерименту – варіанти наборів із k виробів, навмання вибраних серед даних n. Загальна кількість таких наслідків дорівнює C_n^k. Всі наслідки експерименту вважаються рівноімовірними. Нехай A – подія, яка полягає в тому, що серед навмання вибраних k виробів r виробів браковані. Наслідок експерименту, через який спричинюється подія A, відбувається, коли в наборі з k виробів є рівно r бракованих. Сформувати такі набори можна так: навмання вибрати r виробів з m бракованих (це можна зробити C_m^r різними способами) і потім додати $k - r$ виробів з решти $n - m$ якісних виробів (це можна зробити C_{n-m}^{k-r} способами). Оскільки до кожного із C_m^r наборів з r бракованих виробів можна додати будь який із C_{n-m}^{k-r} наборів з $k - r$ якісних виробів, то кількість наборів з k виробів, серед яких буде рівно r бракованих, дорівнює $C_m^r C_{n-m}^{k-r}$.

Оскільки наслідки експерименту вважаються рівноімовірними, бо немає підстав віддавати перевагу будь-якому з них, то

$$P(A) = \frac{C_m^r C_{n-m}^{k-r}}{C_n^k}.$$

Приклад 3.7.4. (задача про збіги). У ліфті знаходяться m пасажирів; ліфт зупиняється на n поверхах. Знайти ймовірність того, що принаймні два пасажири вийдуть на тому самому поверсі.

Нехай A – подія, яка полягає в тому, що принаймні два пасажири вийдуть на тому самому поверсі. Шукатимемо ймовірність протилежної події A, яка полягає в тому, що два пасажири не вийдуть на тому самому поверсі. Якщо перший пасажир вибирає будь-який один з n поверхів, і другий вибирає будь-який з n поверхів, то, очевидно, є n^2 варіантів вибору поверхів двома пасажирами. З кожним з цих n^2 варіантів третій пасажир може вибрати будь-який з n поверхів. Тому матимемо n^3 варіантів вибору поверхів трьома пасажирами. Продовжуючи міркування, дістаємо, що m пасажирів можуть вибрати поверхи n^m способами.

Таким чином, є всього n^m можливих наслідків у даному випробуванні, причому всі наслідки будемо вважати рівноімовірними, оскільки немає підстав віддавати перевагу жодному з них.

Кількість варіантів, в яких два пасажири не виходять на одному поверсі, дорівнює $n(n-1)(n-2)...(n-(m-1))$, оскільки після вибору одного з n поверхів першим пасажиром у другого залишається $n-1$ варіант. Аналогічно після вибору поверхів першим і другим пасажирами у третього залишається $n-2$ варіанти і т. д.

Отже

$$P(\overline{A}) = \frac{A_n^m}{n^m} = \frac{n(n-1)(n-2)...(n-(m-1))}{n^m} =$$

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{m-1}{n}\right),$$

$$P(A) = 1 - P(\overline{A}) = 1 - \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{m-1}{n}\right).$$

Приклад 3.7.5. Підкидають 12 гральних кубиків. Обчислити ймовірність того, що кожне з чисел $1, 2, ..., 6$ випаде двічі.

Очевидно, в даному експерименті простір Ω складається з 6^{12} елементарних подій (оскільки з кожною цифрою на першому кубикові на другому може випасти одна із 6-ти цифр, тобто є 6^2 варіантів випадання пари цифр на двох кубиках. З кожною парою цифр на перших двох кубиках на третьому може випасти одна із 6 цифр, тобто є 6^3 варіантів випадання трьох цифр на трьох кубиках і т.д.). Отже кількість всіх можливих наслідків даного експерименту дорівнює 6^{12}. Всі наслідки експерименту вважаються рівноімовірними, оскільки немає підстав надавати перевагу будь якому з них.

Щоб підрахувати кількість елементарних подій, з відбуванням яких відбувається подія A, міркуватимемо так. Пронумеруємо кубики від 1 до 12 і з 12-елементної множини місць, на які можуть бути поставлені цифри від 1 до 6, вибиратимемо можливі (невпорядковані) двохелементні підмножини місць. Далі з решти 10 місць вибиратимемо двохелементні підмножини, на які ставитимемо одну з решти 5 цифр і т.д. Тоді кількість варіантів, коли кожна з цифр 1, 2, 3, 4, 5, 6 зустрічатиметься на двох із 12-ти підкинутих кубиків, дорівнює

$$C_{12}^2 \cdot C_{10}^2 \cdot C_8^2 \cdot C_6^2 \cdot C_4^2 \cdot C_2^2 = \frac{12!}{(2!)^6} .$$

Таким чином,

$$P(A)=\frac{12!}{(2!)^6\cdot 6^{12}}\cdot$$

Приклад 3.7.6. З п'яти літер ***а, к, і, н, т*** навмання вибирають три. Знайти ймовірність події A, яка полягає в тому, що з вибраних літер можна утворити одне із слів, які можна дістати переставлянням літер у слові "тік".

В цьому прикладі істотним є і набір з трьох літер, і порядок, в якому вони розміщені. Кількість всіх можливих трилітерних слів, які можна утворити з п'яти даних літер, тобто кількість всіх можливих наслідків даного випробування, дорівнює A_5^3. Всі можливі наслідки розглядуваного випробування будемо вважати рівноімовірими, оскільки немає підстав надавати перевагу жодному з них. Число випадків, з якими відбувається подія A (число слів, які можна утворити з літер ***т***, ***і***, ***к***), дорівнює $P_3=3!$.

Отже,

$$P(A)=\frac{P_3}{A_5^3}=\frac{3!}{5\cdot 4\cdot 3}=0.1$$

Приклад 3.7.7. У записаному телефонному номері 476^{****} чотири останні цифри стерлися. Припустивши, що всі набори чотирьох невідомих цифр рівноймовірні, знайти ймовірність таких подій:

A – стерлися різні цифри, відмінні від 4, 6, 7;
B – дві цифри із стертих однакові;
C – дві пари стертих цифр однакові;
D – усі цифри, що стерлися, однакові.

В даному прикладі число всіх елементарних подій дорівнює 10^4. Розглянемо наслідки експерименту, з відбуванням яких відбувається подія A. Таких наслідків є A_7^4 - кількість впорядкованих наборів із чотирьох цифр, вибраних навмання одна за однією із семи цифр 0, 1, 2, 3, 5, 8, 9 без повернення до вихідного набору.

Тому

$$P(A)=\frac{A_7^4}{10^4}=0.084$$

Число випадків, з відбуванням яких відбувається подія B, можна підрахувати так. Спочатку вибрати два місця для двох однакових цифр. Очевидно, що таких наборів місць буде стільки, скільки існує двохелементних невпорядкованих підмножин у чотириелементній множині, тобто $C_4^2=6$. Враховуючи, що цифри можуть бути будь-які від 0 до 9, дістанемо, що дві однакові цифри з 10 цифр на два місця з чотирьох можна поставити

$6 \cdot 10 = 60$ способами. Для третьої цифри є дев'ять можливих варіантів і для четвертої – вісім.

Отже, число випадків, з відбуванням яких відбувається подія B, дорівнює $C_4^2 \cdot 10 \cdot A_9^2 = 6 \cdot 10 \cdot 9 \cdot 8 = 4320.$ Тому

$$P(B) = \frac{4320}{10^4} = 0.432\,.$$

Міркуючи аналогічно, дістанемо, що подія C спричинюється

$$C_4^2 \cdot 10 \cdot 9$$

рівноможливими наслідками експерименту (оскільки з вибором двох місць із чотирьох для першої цифри однозначно визначаються інші два місця для другої цифри).

Таким чином,

$$P(C) = \frac{6 \cdot 10 \cdot 9}{10^4} = 0.054$$

Подія D, очевидно, спричинюється 10-ма наслідками випробування, бо існує лише одна чотириелементна підмножина в чотириелементній множині і на всі місця такої чотириелементної підмножини місць ставиться одна й та сама цифра, одна з десяти. Таким чином, подія D спричинюється $C_4^4 \cdot 10$ наслідками з 10^4. Тому

$$P(D) = \frac{10}{10^4} = 10^{-3} = 0.001$$

Приклад 3.7.8. Нехай n елементів розміщено в певному порядку. Ці елементи розсипаються і потім розставляються навмання. Яка ймовірність того, що принаймні один елемент залишиться на своєму місці?

Вираз "елементи розставляються навмання" означає, що всі $n!$ перестановок елементів рівноможливі. Нехай A_i – подія, яка полягає в тому, що i-й елемент залишиться на своєму i-му місці. Тоді подією "принаймні один елемент залишиться на своєму місці" є подія $A = \sum A_i$. Одержуємо:

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}\,,$$

$$P(A_i A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}\,,$$

$$P(A_i A_j A_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}\,,\ \ldots\,,$$

$$P\left(\prod_{i=1}^{n} A_i\right) = \frac{1}{n!}.$$

Таким чином, оскільки кількість пар (A_i, A_j) дорівнює $C_n^2 = \frac{n(n-1)}{1 \cdot 2}$, кількість трійок (A_i, A_j, A_k) дорівнює $C_n^3 = \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3}$ і т.д., то

$$P(A) = P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) -$$

$$- \ \dots + (-1)^{n-1} P(A_1 A_2 \ \dots \ A_n) = 1 - \frac{1}{2!} +$$

$$+ \frac{1}{3!} - \dots + (-1)^{n-1} \frac{1}{n!} \approx e^{-1}.$$

Вправи для самостійного виконання

3.7.1 Перевірити чи правильні твердження:

1. Кожну сукупність S підмножин простору Ω елементарних подій можна вважати простором подій.

2. Якщо невід'ємна функція $P(A)$, визначена на певній сукупності S підмножин простору Ω, є зчисленно аддитивною і $P(\Omega) = 1$, то $P(A)$, $A \in S$, є ймовірнісною мірою.

3. Якщо $A = B - C$, то $P(A) = P(B) - P(C)$.

4. Якщо $P(A) = 1$, то $A = \Omega$.

5. Якщо $P(A) = 0$, то $A = \varnothing$.

6. Якщо $P(A) \le P(B)$, то через подію A спричинюється подія B.

7. Твердження, обернені до будь-якого з тверджень 4-6, правильні.

8. $P(B) = 1 - P(A) \Leftrightarrow B = \overline{A}$.

9. $P(A+B+C) = P(A) + P(B) + P(C) - P(ABC)$.

10. Якщо $P(A+B) = P(A) + P(B)$, то події A і B несумісні.

11. $P(A+B+C) = P(A+B) + P(C)$ за будь-яких подій A, B і C.

12. Якщо твердження 11 є правильним, то події A, B і C попарно несумісні.

3.7.2. Довести, що коли події $\overline{A}$ і $\overline{B}$ несумісні, то $P(AB) = 1 - P(\overline{A}) - P(\overline{B})$.

3.7.3. Довести, що $P(A+B) = 1 - P(\overline{A}\,\overline{B})$.

3.8. Умовні ймовірності. Залежні і незалежні відносно ймовірнісної міри P випадкові події

Досить часто виникають ситуації, коли необхідно враховувати певні додаткові відомості про обставини, за яких може відбутися чи не відбутися та чи інша випадкова подія.

Якщо ймовірність події A обчислюється за умови, що відбулася подія B, то таку ймовірність події A називають умовною і позначають $P(A/B)$. Якщо ж ймовірність події A обчисляється без будь-якої умови, то таку ймовірність події A називають безумовною і позначають $P(A)$ (порівняйте з §1.10).

Приклад 3.8.1. В кожній з двох урн міститься n кульок, причому в першій урні є m білих і k чорних кульок, а в другій – s білих і r чорних кульок. Із навмання вибраної урни навмання вибирають кульку. Нехай подія A полягає в тому, що вибрали білу кульку, подія B полягає в тому, що вибрано першу урну. Очевидно, що ймовірність появи білої кульки за умови, що вибрано першу урну, є

$$P(A/B)=\frac{m}{n},$$

аналогічно

$$P(A/\overline{B})=\frac{s}{n},$$

де $\overline{B}$ – подія, яка полягає в тому, що вибрано другу урну. Таким чином, імовірність події A залежить від того, відбулася подія B, чи ні.

Об'єднавши кульки обох урн, одержуємо множину Ω, в якій міститься $2n$ елементів. Вибір тієї чи іншої урни означає, що спочатку вибирається n-елементна підмножина B або $\overline{B}$ простору елементарних подій Ω, а потім з цієї підмножини навмання вибирається один елемент.

Зазначимо, що стосовно умовної ймовірності $P(A/B)$ події A множина елементарних подій, за якою визначається подія B, відіграє роль, аналогічну до тієї, яку відіграє простір елементарних подій Ω стосовно безумовної ймовірності $P(A)$ події A. В даному разі σ-алгебру S_B підмножин з B, $B\in S$, можна дістати як сукупність множин виду $A\bigcap B$, $A\in S$, якщо вибірковим простором вважати не всю множину Ω, а тільки множину B (див. §1.10). Ймовірнісна міра $P(A/B)$ (іноді позначають також $P_B(A)$) на (B,S_B) в такому випадку покладається рівною $\dfrac{P(AB)}{P(B)}$, де $A\in S$, $AB\in S_B$. Зауважимо, що умовна ймовірність $P(A/B)$ є нормованою мірою підмножин $AB\subset B$ відносно міри $P(B)$ (див. §3.6).

Якщо відомо, що подія B відбулася, то подія A відбувається з відбуванням тільки тих елементарних подій з множини Ω всіх елементарних подій, з відбуванням яких відбуваються як подія A, так і подія B, тобто подія $AB \subset \Omega$.

Означення. Нехай (Ω, S, P) – ймовірнісний простір, $B \in S$ і $P(B) \neq 0$. *Умовною ймовірністю події* A, $A \in S$, *за умови, що відбулася подія* B, *називається число*

$$P(A/B) = \frac{P(AB)}{P(B)}.$$

Таким чином, умовна ймовірність $P(A/B)$ події A за умови, що відбулася подія B, є відношенням ймовірності одночасного відбування подій A і B до ймовірності відбування події B (див. §3.6, а також §1.10).

Якщо простір елементарних подій $\Omega = \{E_1, E_2, \ldots, E_n\}$ скінченний і $\{E_i\} \in S$, $i \in \overline{1, n}$, то

$$P(A/B) = \frac{P(AB)}{P(B)} = \frac{\sum\limits_{E_i \in AB} P(E_i)}{\sum\limits_{E_i \in B} P(E_i)}.$$

Якщо елементарні події рівноймовірні, тобто $P(E_i) = \dfrac{1}{n}$, то умовна ймовірність $P(A/B)$ події A за умови, що подія B відбулася, дорівнює відношенню кількості тих елементарних подій з множини B, з відбуванням яких відбувається і подія A, тобто кількості елементарних подій у множині AB, до кількості елементарних подій у множині B.

В механічній інтерпретації умовна ймовірність $P(A/B)$ події A за умови, що подія B відбулася, є часткою відмінної від нуля маси, розподіленої на множині B, яка припадає на множину A (тобто на AB).

Якщо $P(B) = 0$, то будь-яка частка ймовірності, розподіленої на множині B, також дорівнює нулю, і $P(AB)/P(B)$ стає невизначеною (її можна визначити довільно, наприклад, покласти $P(A/B) = P(A)$).

Безумовну ймовірність $P(A)$ можна трактувати як умовну ймовірність $P(A/\Omega)$ події A за умови, що відбулася подія Ω, яка визначається за множиною всіх можливих наслідків випробування.

Приклад 3.8.2. Двічі підкидають гральний кубик. Простір елементарних подій складається з 36 рівноможливих елементів.

$$\Omega = \{(i, j) \mid i = \overline{1, 6},\ j = \overline{1, 6}\}.$$

Нехай подія A полягає в тому, що випаде принаймні одна шістка:

$$A=\{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(6,5),(6,4),(6,3),(6,2),(6,1)\}.$$

Нехай також відомо, що в результаті двох підкидань випало вісім очок (подія B):

$$B=\{(2,6)\ (3,5)\ (4,4)\ (5,3)\ (6,2)\}.$$

Із п'яти наслідків випробування, з відбуванням яких відбувається подія B, з відбуванням двох наслідків (2, 6) та (6, 2) відбувається і подія A (Рис. 3.8.1).

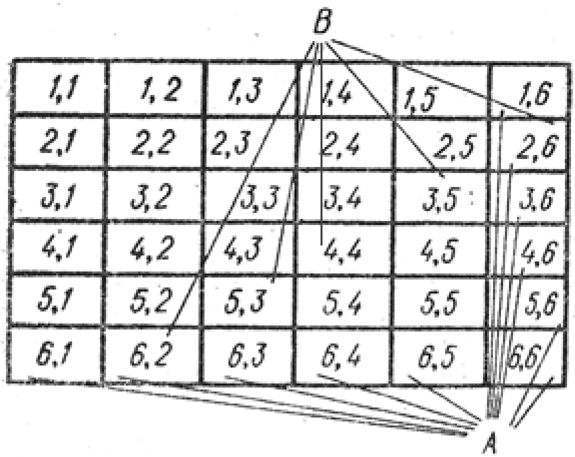


Рис. 3.8.1

Отже, в розглядуваному випадку

$$P(A/B)=2/5.$$

Із означення умовної ймовірності стосовно ймовірності добутку двох подій одержуємо:

$$P(AB)=P(A)\,P(B/A)=P(B)\,P(A/B)\,. \qquad (3.8.1)$$

Стосовно умовної ймовірності задовільняються такі властивості:

1. $P(B/B)=1$, коли $P(B)>0$.

Ця властивість випливає з означення умовної ймовірності.

2. $0\le P(A/B)\le 1$.

Зазначимо, що $AB\subset B$ і тому за властивостями безумовної ймовірності $0\le P(AB)\le P(B)$. Звідси випливає, що $0\le P(A/B)\le 1$.

3. Якщо $AB=\varnothing$, то

$$P((A+B)/C)=P(A/C)+P(B/C)\,.$$

Справді, оскільки $AB=\varnothing$, то події AC і BC несумісні, а тому $P((A+B)\ C)=P(AC)+P(BC)$, і якщо $P(C)>0$, то

$$P((A+B)/C)=\frac{P((A+B)C)}{P(C)}=\frac{P(AC)}{P(C)}+\frac{P(BC)}{P(C)}=P(A/C)+P(B/C).$$

4. Якщо $A\subset B$, то $P(A/C)\le P(B/C)$.

Справді, якщо $A\subset B$, то $B=A+\overline{A}B$, і за властивістю 3 одержуємо:

$$P(B/C)=P(A/C)+P(\overline{A}B/C).$$

Згідно з властивістю 2 $P(\overline{A}B/C)\ge 0$. Тому

$$P(B/C)\ge P(A/C).$$

5. За довільних подій A і B виконується рівність

$$P((A+B)/C)=P(A/C)+P(B/C)-P(AB/C).$$

Використовуючи формулу стосовно ймовірності суми двох довільних подій та означення умовної ймовірності, можна записати

$$P((A+B)/C)=\frac{P((A+B)C)}{P(C)}=\frac{P(AC+BC)}{P(C)}=$$

$$=\frac{P(AC)+P(BC)-P(AC\cdot BC)}{P(C)}=\frac{P(AC)}{P(C)}+\frac{P(BC)}{P(C)}-\frac{P(ABC)}{P(C)}=$$

$$=P(A/C)+P(B/C)-P(AB/C), \text{ коли } P(C)>0.$$

6. Якщо $A_iA_j=\varnothing$, коли $i\ne j$, то

$$P((A_1+A_2+...+A_n)/B)=P(A_1/B)+P(A_2/B)+...\ P(A_n/B).$$

Ця властивість випливає з властивості $2_р$ аддитивності ймовірності та означення умовної ймовірності. Так, враховуючи, що $A_iB\cdot A_jB=A_iA_jB=\varnothing$, одержуємо

$$P((A_1+A_2+...+A_n)/B)=\frac{P((A_1+A_2+...+A_n)B)}{P(B)}=$$

$$=\frac{P(A_1B+A_2B+...+A_nB)}{P(B)}=\frac{P(A_1B)}{P(B)}+\frac{P(A_2B)}{P(B)}+....+$$

$$+\frac{P(A_nB)}{P(B)}=P(A_1/B)+P(A_2/B)+...+P(A_n/B),$$

коли $P(B)>0$.

7. $P(\Omega/B)=1$.

Справді, оскільки $\Omega B = B$, то у випадку $P(B) > 0$

$$P(\Omega / B) = \frac{P(\Omega\ B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

8. $P(\varnothing / B) = 0$.

Справді, якщо $P(B) > 0$, то

$$P(\varnothing / B) = \frac{P(\varnothing\ B)}{P(B)} = \frac{P(\varnothing)}{P(B)} = \frac{0}{P(B)} = 0.$$

9. $P(\overline{A} / B) = 1 - P(A / B)$.

Оскільки $A + \overline{A} = \Omega$, то

$$P((A + \overline{A})\ B) = P(\Omega B) = P\ (B).$$

Крім того,

$$P((A + \overline{A})\ B) = P(AB) + P\ (\overline{A}B).$$

Звідси, коли $P(B) > 0$,

$$\frac{P(A\ B)}{P(B)} + \frac{P(\overline{A}B)}{P(B)} = \frac{P(B)}{P(B)} = 1,$$

тобто

$$P(A / B) + P\ (\overline{A} / B) = 1\text{, звідки } P\ (\overline{A} / B) = 1 - P(A / B).$$

Означення. Події A і B називаються незалежними відносно ймовірнісної міри P, якщо

$$P(A \cdot B) = P(A)\ P(B).$$

Легко бачити, що події A і B незалежні відносно ймовірнісної міри P тоді і тільки тоді, коли умовна ймовірність $P(A / B)$ події A за умови, що подія B відбулася, дорівнює безумовній ймовірності події A.

Справді, $P(AB) = P(B) \cdot P(A / B)$, а тому

$$P(AB) = P(A) \cdot P(B) \Leftrightarrow P(A / B) = P(A).$$

В механічному тлумаченні тут частка, що припадає на множину A від маси (ймовірності), розподіленої на множині B, така сама, як і частка, що припадає на множину A від загальної маси (ймовірності), розподіленої на множині Ω всіх елементарних подій.

Зазначимо, що коли подія A не залежить від події B відносно ймовірнісної міри P, то й подія B не залежить від події A відносно тієї самої ймовірнісної міри. Справді, нехай

$$P(B) \neq 0,\ P\ (A) \neq 0,\ P(A / B) = P(A).$$

Тоді

$$P(B / A) = \frac{P(A\ B)}{P(A)} = \frac{P(B) P(A / B)}{P(A)} = \frac{P(B)\ P(A)}{P(A)} = P(B),$$

що й треба було довести.

Зауважимо, що події A і B, незалежні відносно ймовірнісної міри P, можуть виявитись залежними відносно іншої ймовірнісної міри $\tilde{P}$ (див. §1.10, §1.11).

Приклад 3.8.3. Довести, що з умови $P(A/B) = P(A/\overline{B})$ випливає незалежність подій A і B відносно ймовірнісної міри P.

Враховуючи, що $B + \overline{B} = \Omega,\ A\Omega = A$ і $AB \cdot A\overline{B} = \varnothing$, одержуємо

$$P\ (A) = P\ (A\Omega) = P(A(B + \overline{B})) = P(AB) + P(A\overline{B}) =$$
$$= P(B)\,P(A/B) + P(\overline{B})P(A/\overline{B}).$$

Оскільки $P(A/B) = P(A/\overline{B})$, то замінивши $P(A/\overline{B})$ на $P(A/B)$, дістанемо

$$P(A) = P(A/B)\ (P(B) + P(\overline{B})) = P(A/B)\,.$$

Отже, подія A не залежить від події B відносно ймовірнісної міри P.

Приклад 3.8.4. Якщо події A і B незалежні відносно ймовірнісної міри P, то події A і $\overline{B}$ також незалежні відносно тієї самої ймовірнісної міри.

Справді, оскільки $A = AB + A\overline{B}$, причому $AB \cdot A\overline{B} = \varnothing$, то

$$P(A) = P(AB) + P(A\overline{B})\,.$$

Тоді

$$P\ (A\overline{B}) = P\ (A) - P(AB) = P(A) - P(A)\,P(B) =$$
$$= P(A)\,(1 - P(B)) = P(A) \cdot P(\overline{B})\,.$$

Отже, події A і $\overline{B}$ незалежні відносно ймовірнісної міри P. Звідси випливає, що коли події A і B незалежні відносно ймовірнісної міри P, то й події A і $\overline{B}$, B і $\overline{A}$, $\overline{A}$ і $\overline{B}$ також незалежні відносно тієї самої ймовірнісної міри.

Приклад 3.8.5. Нехай є 10 білих, 10 червоних, 10 зелених і 10 жовтих карток.

Картки кожного кольору пронумеровано цифрами від 1 до 10. Із цих 40 карток навмання вибирається одна картка. Знайти ймовірність того, що ця картка буде біла і, крім того, на ній написано одну з трьох цифр 6, 7 або 8, якщо всі картки рівноймовірні.

Позначимо через A подію, яка полягає в тому, що навмання вийнята картка біла, а через B – подію, яка полягає в тому, що картку пронумеровано однією з трьох цифр 6, 7 або 8. У цьому разі одержуємо простір Ω із 40 рівноймовірних елементарних подій (Рис. 3.8.2), 10 з яких належать до множини A і 12 – до множини B, причому в A і B є три спільні елементи б6, б7, б8.

Таким чином, $P(A)=\dfrac{10}{40}$, а $P(B/A)$ – умовна ймовірність того, що на картці буде одна з трьох цифр 6, 7 або 8 за умови, що картка біла, дорівнює 3/10. Отже,

$$P(AB)=P(A)\;P(B/A)=\frac{10}{40}\cdot\frac{3}{10}=\frac{3}{40}.$$

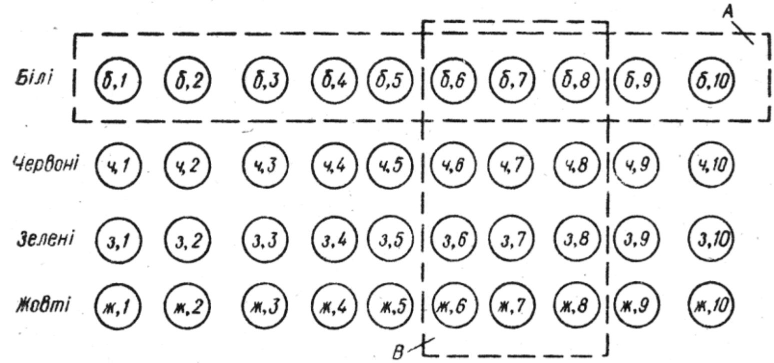


Рис. 3.8.2

Цей результат можна дістати й так. Враховуючи, що $P(B)=12/40,\ P(A/B)=3/12$, маємо

$$P(AB)=P(B)\;P(A/B)=\frac{12}{40}\cdot\frac{3}{12}=\frac{3}{40}.$$

У розглянутому прикладі події A і B незалежні відносно так заданої на Ω ймовірнісної міри, тобто (див. Рис. 3.8.2)

$$P(A/B)=P(A)=1/4,\qquad P(B/A)=P(B)=3/10.$$

Використовуючи формулу (3.8.1), стосовно добутку довільної кількості n подій $A_1, A_2, ..., A_n$ одержимо:

$$P(A_1A_2\ ...\ A_n)=$$
$$=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ ...\ P(A_n/A_1A_2\ ...\ A_{n-1}).$$

Справді

$$P(A_1A_2...A_{n-1}A_n)=P((A_1A_2...A_{n-1})A_n)=$$
$$=P(A_1A_2...A_{n-1})\;P(A_n/A_1A_2...A_{n-1})=$$
$$=P(A_1A_2\ ...\ A_{n-2})P(A_{n-1}/A_1A_2\ ...\ A_{n-2})\;P(A_n/A_1A_2\ ...\ A_{n-1})=...$$
$$...=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ ...\ P(A_{n-1}/A_1A_2\ ...\ A_{n-2})\times$$
$$\times\ P(A_n/A_1A_2\ ...\ A_{n-1}). \tag{3.8.2}$$

Означення. Події $A_1, A_2, ... A_k$ називають *незалежними в сукупності* відносно ймовірнісної міри P, якщо $P(\prod\limits_{i\in I}A_i)=\prod\limits_{i\in I}P(A_i)$ за довільного $I\subset\{1,2,...,k\}$.

Враховуючи комутативний (переставний) закон множення і формулу (3.8.2), дістаємо, що ймовірність добутку подій $A_1, A_2, \dots, A_n$ дорівнює добутку ймовірностей цих подій тоді, коли кожна з подій не залежить від будь-якої іншої події, від добутку двох будь-яких інших подій, від добутку трьох будь-яких інших подій і т.д. і, нарешті, від добутку всіх інших подій. Отже, якщо події незалежні в сукупності відносно ймовірнісної міри P, то будь-яка з них не залежить відносно цієї ймовірнісної міри від добутку будь-якої сукупності інших подій.

Зазначимо, що з попарної незалежності випадкових подій відносно ймовірнісної міри P, не випливає їх незалежність в сукупності відносно тієї самої ймовірнісної міри (див. приклад 1.11.2).

Слід зазначити також, що з того, що за деяких подій A, B, C

$$P(ABC) = P(A)\ P(B)\,P(C),$$

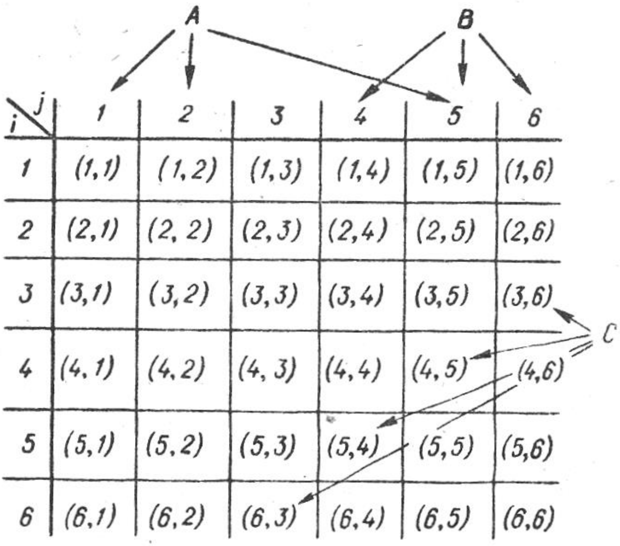
не випливає попарна незалежність цих подій відносно міри P, а отже і незалежність в сукупності.

Приклад 3.8.7. Нехай

$$\Omega = \{(i, j) \mid i = \overline{1, 6}, j = \overline{1, 6}\},\ A = \{(i, j) \mid i \in \overline{1, 6}, j \in \{1, 2, 5\}\},$$

$$B = \{(i, j) \mid i \in \overline{1, 6}, j \in \{4, 5, 6\}\},\ C = \{(i, j) \mid i + j = 9\}$$

(Рис. 3.8.3) і всі елементарні події $(i, j) \in \Omega$ рівноймовірні. Тоді



i \ j	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Рис. 3.8.3

$$AB = \{(i,\ j) \mid i \in \overline{1,\ 6}, j = 5\}\quad AC = \{(4,\ 5)\},\ BC = \{(\ 5, 4),\ (\ 4, 5), (\ 3, 6)\},$$
$$ABC = \{(\ 4,\ 5)\};$$

$$P(A)=\frac{1}{2},\ P(B)=\frac{1}{2},\ P(C)=\frac{1}{9},\ P(ABC)=\frac{1}{36}=P(A)P(B)P(C),$$
$$P(AB)=\frac{1}{6}\neq P(A)P(B)=\frac{1}{4},\ P(AC)=\frac{1}{36}\neq P(A)P(C)=\frac{1}{18},$$
$$P(BC)=\frac{3}{36}\neq P(B)P(C)=\frac{1}{18}.$$

Приклад 3.8.8. Імовірність того, що протягом доби вийде з ладу k-й блок радіопристрою, дорівнює p_k, $k=1,\ 2,\ldots,10$ Знайти ймовірність того, що протягом доби радіопристрій вийде з ладу (тобто вийде з ладу принаймні один блок), якщо кожен блок виходить з ладу незалежно від інших.

Позначимо через A_k, $k=1,\ 2,\ \ldots,10$, подію, яка полягає в тому, що виходить з ладу k-й блок, через B позначимо подію, яка полягає в тому, що виходить з ладу радіопристрій. Тоді

$$B=A_1+A_2+\ldots+A_{10},$$

$$P(B)=P(A_1+A_2+\ldots+A_{10})=1-P\ \overline{(A_1+A_2+\ldots+A_{10})}=$$

$$=1-P(\overline{A}_1\,\overline{A}_2\,\overline{A}_3\ \ldots\ \overline{A}_{10})=1-P(\overline{A}_1)P(\overline{A}_2)\ \ldots\ P(\overline{A}_{10})=$$

$$=1-(1-p_1)\ (1-p_2)\ \ldots\ (1-p_{10})\ .$$

Вправи для самостійного виконання

3.8.1. Перевірити, чи правильні твердження:

1. Якщо задано імовірнісний простір (Ω,S,P) і фіксована подія B, така, що $P(B)\neq 0$, то умовна ймовірність $P(A/B)$, $A\in S$, є ймовірністю на просторі подій S.

2. Якщо S – простір подій, а $B\in S$ – фіксована подія, то сукупність $\tilde{S}=\{\tilde{A}\mid\ \tilde{A}=AB,\ A\in S\}$ є простором подій, стосовно якого простором елементарних подій є $\tilde{\Omega}=\Omega\cdot B$.

3. Умовна ймовірність $P(A/B)=\tilde{P}(\tilde{A}),\ \tilde{A}\in\tilde{S}$, є ймовірністю на просторі подій $\tilde{S}$.

4. Якщо події A і B незалежні відносно ймовірнісної міри P, то $P(A)=0$ або $P(B)=0$.

5. Твердження, обернене до 4, є правильним.

6. Якщо події A і B незалежні відносно ймовірнісної міри P, то $P(A/\overline{B})=P(A)$.

7. Якщо події $\overline{A}$ і $\overline{B}$ незалежні, відносно ймовірнісної міри P, то і події A та B незалежні відносно тієї самої ймовірнісної міри.

8. Якщо $P(ABC)=P(A)P(B)P(C)$, то події A, B, C незалежні в сукупності відносно ймовірнісної міри P.

9. Твердження, обернене до 8, є правильним.

10. Події $A_1, A_2, ..., A_n$ незалежні в сукупності відносно ймовірнісної міри P, якщо такими є будь-які $(n-1)$ подій цієї сукупності і крім того

$$P(A_1 A_2 ... A_n)=P(A_1)P(A_2)...P(A_n).$$

3.8.2. Довести, що несумісні події A і B є незалежними відносно ймовірнісної міри P тоді і тільки тоді, коли $P(A)=0$ або $P(B)=0$.

3.8.3. Довести, що коли події $A_1, A_2, \ldots, A_n$ незалежні в сукупності відносно ймовірнісної міри P, то $P(\sum_{k=1}^{n} A_k)=1-\prod_{k=1}^{n}(1-P(A_k))$.

3.9. Формула повної ймовірності. Формула Байєса. Теорема гіпотез

Нехай простір елементарних подій Ω поділено на k підмножин $H_1, H_2, ..., H_k$ таких, що $H_i H_j = \varnothing$, коли $i \neq j$, $H_1 + H_2 + ... + H_k = \Omega$, $H_i \in S$, $i \in \overline{1, k}$, $A \in S$, і нехай проводиться випробування, в якому спостерігається подія A (Рис. 3.9.1).

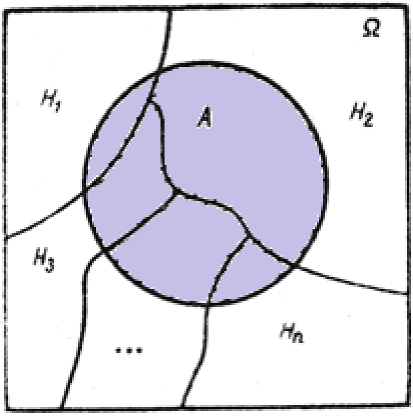


Рис. 3.9.1

Події $H_1, H_2, ..., H_k$ називатимемо *гіпотезами*, з якими може відбутися подія A (порівняйте з §1.12).

Припустимо, що ймовірності гіпотез H_i до випробування відомі (або їх можна легко знайти) і дорівнюють $P(H_i)$, $i = 1, 2, ..., k$, а також відомі (або легко знайти) умовні ймовірності $P(A/H_i)$ події A за кожною з гіпотез H_i, $i = 1, 2, ..., k$. Треба знайти ймовірність $P(A)$ події A.

Враховуючи співвідношення

$$A = A\Omega = A(H_1 + H_2 + ... + H_k) = AH_1 + AH_2 + ... + AH_k$$

і те, що доданки в останній сумі попарно несумісні, оскільки $AH_i AH_j = AAH_i H_j = A\varnothing = \varnothing$, дістаємо

$$P(A) = P(AH_1 + AH_2 + ... + AH_k) = P(AH_1) + P(AH_2) + ... + P(AH_k) =$$
$$= P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + ... + P(H_k)P(A/H_k).$$

Таким чином, ймовірність події A дорівнює сумі ймовірностей $P(H_i)$ гіпотез H_i, помножених на умовні ймовірності події A за цими гіпотезами:

$$P(A) = \sum_{i=1}^{k} P(H_i)\ P(A/H_i).$$

Цю формулу називають *формулою повної ймовірності*.

Приклад 3.9.1. Розігрують m виграшів серед n осіб $(m < n)$. Кожна особа згідно зі своїм номером в черзі $(1 \le i \le n)$ бере білет з урни (кількість білетів дорівнює n). Чи рівні шанси виграшу у кожного з учасників розіграшу, якщо кожен білет знаходиться в конверті і конверти відкриваються одночасно? Що вигідніше – бути першим чи останнім у черзі?

Якщо подія A_1 – виграш особи, яка першою бере білет, то $P(A_1) = \frac{m}{n}$, а $P(\overline{A}_1) = \frac{n-m}{n}$.

Нехай подія A_2 – виграш особи, яка бере білет другою. Стосовно події A_2 та події $\overline{A}_2$ природні гіпотези A_1 та $\overline{A}_1$. Тоді зрозуміло, що

$$P(A_2 / A_1) = \frac{m-1}{n-1},\ P(A_2 / \overline{A}_1) = \frac{m}{n-1},$$

і за формулою повної ймовірності

$$P(A_2) = P(A_1) \cdot P(A_2 / A_1) + P(\overline{A}_1) \cdot P(A_2 / \overline{A}_1) =$$

$$= \frac{m}{n} \cdot \frac{m-1}{n-1} + \frac{n-m}{n} \cdot \frac{m}{n-1} = \frac{m}{n(n-1)}(m-1+n-m) = \frac{m}{n}$$

Отже, $P(A_2) = P(A_1) = \frac{m}{n}$, а $P(\overline{A}_2) = P(\overline{A}_1) = \frac{n-m}{n}$.

Нехай подія A_3 – виграш особи, яка бере білет третьою. Стосовно подій A_3 та $\overline{A}_3$ природні гіпотези: $\overline{A}_1\overline{A}_2$, $\overline{A}_1 A_2$, $A_1\overline{A}_2$ і $A_1 A_2$. Обчислюючи ймовірності цих добутків подій за формулою (3.8.1), одержимо:

$$P(\overline{A}_1\overline{A}_2) = P(\overline{A}_1) \cdot P(\overline{A}_2 / \overline{A}_1) = \frac{n-m}{n} \cdot \frac{n-1-m}{n-1},$$

$$P(\overline{A}_1 A_2) = P(\overline{A}_1) \cdot P(A_2 / \overline{A}_1) = \frac{n-m}{n} \cdot \frac{m}{n-1},$$

$$P(A_1\overline{A}_2) = P(A_1) \cdot P(\overline{A}_2 / A_1) = \frac{m}{n} \cdot \frac{n-1-(m-1)}{n-1} = \frac{m}{n} \cdot \frac{n-m}{n-1},$$

$$P(A_1 A_2) = P(A_1) \cdot P(A_2 / A_1) = \frac{m}{n} \cdot \frac{m-1}{n-1}.$$

Враховуючи це, за формулою повної ймовірності дістаємо

$$P(A_3) = P(\overline{A}_1\overline{A}_2) \cdot P(A_3 / \overline{A}_1\overline{A}_2) + P(\overline{A}_1 A_2) \cdot P(A_3 / \overline{A}_1 A_2) +$$

$$+ P(A_1\overline{A}_2) \cdot P(A_3 / A_1\overline{A}_2) + P(A_1 A_2) \cdot P(A_3 / A_1 A_2) =$$

$$= \frac{n-m}{n} \cdot \frac{n-m-1}{n-1} \cdot \frac{m}{n-2} + \frac{n-m}{n} \cdot \frac{m}{n-1} \cdot \frac{m-1}{n-2} + \frac{m}{n} \cdot \frac{n-m}{n-1} \cdot \frac{m-1}{n-2} +$$

$$+\frac{m}{n}\cdot\frac{m-1}{n-1}\cdot\frac{m-2}{n-2}=\frac{m}{n(n-1)(n-2)}(n^2-nm-n-mn+m^2+m+nm-n-$$

$$-m^2+m+nm-n-m^2+m+m^2-2m-m+2)=\frac{m(n^2-3n+2)}{n(n-1)(n-2)}=\frac{m}{n}.$$

Отже,

$$P(A_3)=P(A_2)=P(A_1)=\frac{m}{n}, \text{ а } P(\overline{A}_3)=P(\overline{A}_2)=P(\overline{A}_1)=\frac{n-m}{n}.$$

Аналогічно переконуємося, що коли подія A_i – виграш особи, яка бере білет i-ю, то $P(A_i)=\frac{m}{n}$.

Отже, $P(A_i)=\frac{m}{n}$ за довільного i. Таким чином усі особи мають рівноможливі шанси виграти незалежно від номера в черзі, за умови, що кожен білет знаходиться в конверті і конверти відкриваються одночасно.

Якщо ж особа, яка бере білет i-ою, знає, які білети взяли раніше, то за цієї умови шанси виграти, взагалі кажучи, не дорівнюють $\frac{m}{n}$.

Наприклад, якщо друга особа знає, що перша особа взяла невиграшний білет, то її шанси виграти дорівнюють $P(A_2/\overline{A}_1)=\frac{m}{n-1}>\frac{m}{n}$.

Приклад 3.9.2. Є дві партії зошитів. В першій з них m зошитів, в другій – n зошитів, причому в кожній партії є k зошитів в клітинку, а решта зошитів в лінійку. З першої партії навмання вибирають один зошит і перекладають в другу, після чого з другої партії навмання беруть один зошит. Знайти ймовірність того, що цей зошит виявиться в клітинку (подія A), за умови, що всі зошити вибираються з однаковими ймовірностями з першої партії, так само з однаковими ймовірностями зошити вибираються і з другої партії.

Якби було відомо, який зошит, в клітинку чи в лінійку, перекладено з першої партії в другу, то ймовірність того, що після перекладання навмання взятий з другої партії зошит в клітинку, було б легко знайти. Проте оскільки невідомо, який саме зошит буде перекладено з першої партії в другу, можна лише зробити два припущення: H_1 – перекладено зошит в клітинку, H_2 – перекладено зошит в лінійку.

Очевидно, що гіпотези несумісні, ними вичерпуються усі можливі випадки, причому $P(H_1)=\frac{k}{m}$, $P(H_2)=\frac{m-k}{m}$, бо в першій партії серед m зошитів є k в клітинку і $m-k$ в лінійку.

Ймовірність того, що навмання взятий з другої партії зошит виявиться в клітинку за умови, що туди було покладено зошит в клітинку, тобто $P(A/H_1)$, дорівнює $\frac{k+1}{n+1}$, оскільки після такого перекладання в другій партії виявляється $n+1$ зошит і серед них $k+1$ в клітинку. Аналогічно

$$P(A/H_2)=\frac{k}{n+1}.$$

За формулою повної ймовірності дістаємо

$$P(A)=P\ (H_1)\ P\ (A/H_1)+P\ (H_2)\ P\ (A/H_2)=$$
$$=\frac{k}{m}\ \frac{k+1}{n+1}+\frac{m-k}{m}\ \frac{k}{n+1}=\frac{k\,(m+1)}{m\,(n+1)}.$$

Слід зазначити, що в наведених міркуваннях гіпотези роблять до випробування, тобто до того, як стало будь що відомо про наслідок випробування.

Досить часто доводиться розглядати таку задачу. Умови ті самі, що й у попередній задачі, тобто подія A може відбутися лише з однією з гіпотез $H_1, H_2, \ldots, H_k$, причому $H_iH_j \neq \varnothing$, коли $i \neq j,\ H_1+H_2+\ldots+H_k=\Omega$, ймовірності $P(H_i)$ відомі (або легко знаходяться), а також відомі (або легко знаходяться) умовні ймовірності $P(A/H_i)$ події A за кожною з гіпотез.

Нехай випробування проведено і подія A відбулася. Треба знайти ймовірність того, що подія A відбулася з гіпотезою H_i, тобто $P(H_i/A)$ (в механічному тлумаченні – частку ймовірності, розподіленої на множині A, що припадає на множину H_i, тобто на AH_i).

Згідно з означенням умовної ймовірності,

$$P(H_i/A)=\frac{P(AH_i)}{P(A)}=\frac{P(H_i)\,P(A/H_i)}{P(A)}.$$

Враховуючи, що $P(A)=\sum_{i=1}^{k}P(H_i)\,P(A/H_i)$, дістаємо

$$P(H_i/A)=\frac{P(H_i)\,P(A/H_i)}{\sum_{i=1}^{k}P(H_i)\,P(A/H_i)},\quad i=1,\ 2,\ \ldots,\ k.$$

Останню рівність називають *формулою Байєса*.

Умовну ймовірність гіпотези H_i за умови, що відбулася подія A, часто називають ймовірністю гіпотези H_i після випробування, в якому відбулася подія A.

Таким чином має місце *теорема гіпотез*: *ймовірність гіпотези після випробування, в якому відбулася подія* A*, дорівнює ймовірності цієї гіпотези до випробування, помноженій на умовну ймовірність події* A *за цією гіпотезою і поділеній на повну ймовірність події* A.

В механічному тлумаченні за формулою Байєса обчислюється частка $P(H_i / A) = \dfrac{P(AH_i)}{P(A)}$ від маси, розподіленої на множині A, яка припадає на множину H_i, якщо відомі маси $P(H_i)$, розподілені на кожній з підмножин H_i, а також частки $P(A/H_i) = \dfrac{P(AH_i)}{P(H_i)}$ від мас $P(H_i)$, що припадають на множину A (тобто на множини AH_i).

Приклад 3.9.3. Троє стрільців одночасно вистрілили в мішень, в якій виявилось лише два влучення. Знайти ймовірність того, що в мішень не влучив третій стрілець, якщо в результаті кожного пострілу перший з них влучає з ймовірністю 0.4, другий – з ймовірністю 0.5, третій – з ймовірністю 0.6.

На перший погляд може здатися, що є лише три гіпотези: H_1 – не влучив перший стрілець, H_2 – не влучив другий стрілець, H_3 – не влучив третій стрілець. Проте гіпотези роблять до того, як стало що-небудь відомо про те, які саме наслідки мали місце у випробуванні.

Позначимо через A_1 подію, яка полягає в тому, що перший стрілець влучає в мішень, $\overline{A}_1$ – перший не влучає, A_2 – другий влучає, $\overline{A}_2$ – другий не влучає, A_3 – третій влучає, $\overline{A}_3$ – третій не влучає. Тоді до випробування, до того, як стало відомо, що в мішені є два влучення, допустимі такі гіпотези:

$H_1 = A_1 A_2 A_3$ – всі троє влучають;

$H_2 = A_1 A_2 \overline{A}_3$ – перший і другий влучають, третій не влучає;

$H_3 = A_1 \overline{A}_2 A_3$ – перший і третій влучають, другий не влучає;

$H_4 = A_1 \overline{A}_2 \overline{A}_3$ – перший влучає, другий і третій не влучають;

$H_5 = \overline{A}_1 A_2 A_3$ – перший не влучає, другий і третій влучають;

$H_6 = \overline{A}_1 A_2 \overline{A}_3$ – перший і третій не влучають, другий влучає;

$H_7 = \overline{A}_1 \overline{A}_2 A_3$ – перший і другий не влучають, третій влучає;

$H_8 = \overline{A}_1 \overline{A}_2 \overline{A}_3$ – всі троє не влучають.

$\Omega = \sum_{i=1}^{8} H_i$ (порівняйте з §2.10, формули 2.10.2).

Позначимо через A подію, яка полягає в тому, що в мішені є два влучення. Очевидно, що $P(A/H_1) = 0$, оскільки за умови, що влучають всі троє, неможливо, щоб в мішені було рівно два влучення. Аналогічно $P(A/H_2)=1$, $P(A/H_3)=1$, $P(A/H_4)=0$, $P(A/H_5)=1$, $P(A/H_6)=0$, $P(A/H_7)=0$, $P(A/H_8)=0$.

Оскільки природно вважати, що події A_1, A_2, A_3 незалежні в сукупності відносно заданої на Ω ймовірнісної міри P, бо стрільці вистрілили одночасно, то

$$P(H_1)=P(A_1A_2A_3)=P(A_1)P(A_2)P(A_3)=0.4\cdot 0.5\cdot 0.6=0.12$$
$$P(H_2)=P(A_1A_2\overline{A}_3)=P(A_1)P(A_2)P(\overline{A}_3)=0.4\cdot 0.5\cdot 0.4=0.08$$
$$P(H_3)=P(A_1\overline{A}_2A_3)=P(A_1)P(\overline{A}_2)P(A_3)=0.4\cdot 0.5\cdot 0.6=0.12$$
$$P(H_4)=P(A_1\overline{A}_2\overline{A}_3)=P(A_1)P(\overline{A}_2)P(\overline{A}_3)=0.4\cdot 0.5\cdot 0.4=0.08$$
$$P(H_5)=P(\overline{A}_1A_2A_3)=P(\overline{A}_1)P(A_2)P(A_3)=0.6\cdot 0.5\cdot 0.6=0.18$$
$$P(H_6)=P(\overline{A}_1A_2\overline{A}_3)=P(\overline{A}_1)P(A_2)P(\overline{A}_3)=0.6\cdot 0.5\cdot 0.4=0.12$$
$$P(H_7)=P(\overline{A}_1\overline{A}_2A_3)=P(\overline{A}_1)P(\overline{A}_2)P(A_3)=0.6\cdot 0.5\cdot 0.6=0.18$$
$$P(H_8)=P(\overline{A}_1\overline{A}_2\overline{A}_3)=P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)=0.6\cdot 0.5\cdot 0.4=0.12$$

За формулою повної ймовірності

$$P(A)=\sum_{i=1}^{8}P(H_i)\ P(A/H_i)=$$
$$=0.12\cdot 0+0.08\cdot 1+0.12\cdot 1+0.08\cdot 0+0.18\cdot 1+$$
$$+0.12\cdot 0+0.18\cdot 0+0.12\cdot 0=0.38$$

За формулою Байєса знайдемо ймовірність того, що не влучив третій стрілець за умови, що в мішені є два влучення:

$$P(H_2/A)=\frac{P(H_2)P(A/H_2)}{\sum_{i=1}^{8}P(H_i)P(A/H_i)}=\frac{0.08\cdot 1}{0.38}\approx 0.21.$$

Вправи для самостійного виконання

3.9.1. Перевірити, чи правильні твердження:

1. Якщо події $H_i,\ i\in\overline{1,\ n}$, попарно несумісні, то вони є гіпотезами.
2. Якщо $H_1+H_2+...+H_n=\Omega$, то події $H_i,\ i\in\overline{1,\ n}$, є гіпотезами.
3. Твердження, обернене до 1 або до 2, є правильним.
4. Якщо $H_1+H_2=\Omega$, то
$$P(A)=P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2).$$
5. Ймовірність гіпотези залежить від того, відбулася подія A чи ні.
6. Умовна ймовірність гіпотези H_i за умови,що відбулася подія A, дорівнює добутку ймовірності цієї гіпотези на умовну ймовірність події A за цією гіпотезою, поділеному на повну ймовірність події A.

3.9.2. Довести, що коли через суму подій B_k, $k \in \overline{1, n}$, вичерпується простір елементарних подій Ω, тобто $\Omega = \sum_{k=1}^{n} B_k$, то події

$$H_1 = B_1,\ H_2 = B_2 - B_1,$$
$$H_3 = B_3 - (B_1 + B_2),\ \ldots,\ H_n = B_n - (B_1 + B_2 + \ldots + B_{n-1})$$

можуть бути гіпотезами.

3.9.3. Дослідити, якими можуть бути гіпотези стосовно ймовірнісного простору (Ω, S, P), що відповідає експерименту: 1) підкидання монети та фіксація її верхньої грані; 2) підкидання двох монет та фіксація їх верхніх граней.

3.10. Поточкові розподіли ймовірностей на дискретних множинах точок із R^1

Якщо множина $\Omega \subset R^1$ елементарних подій дискретна, тобто скінченна $\Omega = \{E_1, E_2, ..., E_k\}$, або зчисленна – $\Omega = \{E_1, E_2, ..., E_k, ...\}$, $S = \{A \mid A = \bigcup_{i \in I} E_i,\ I \subset \{1, 2, ..., k\}\}$ чи $S = \{A \mid A = \bigcup_{i \in I} E_i,\ I \subset \{1, 2, ...\}\}$ і на S задана ймовірнісна міра P, тобто визначені $P(\{E_i\})$ за всіх $E_i \in \Omega$, то розподіл ймовірностей на такій множині є поточковим (дискретним).

Зауважимо, що дискретний розподіл ймовірностей можна задати і на неперервній множині виду $\Omega = [a;b]$, вибравши в неперервній множині Ω дискретну підмножину і задавши розподіл ймовірностей на такій підмножині.

Поточковий дискретний розподіл ймовірностей на скінченній множині $\Omega = \{E_1, E_2, ..., E_k\}$ можна задати за таблицею виду

Табл. 3.10.1

E_i	E_1	E_2	$\dots$	E_k
$P(\{E_i\})$	$P(\{E_1\})$	$P(\{E_2\})$	$\dots$	$P(\{E_k\})$

Таку таблицю називають *рядом розподілу ймовірностей на скінченній множині* Ω.

Очевидно, що в такому разі за довільного $A \subset \Omega$, $A \in S$, буде $P(A) = \sum_{E_i \in A} P(E_i)$.

Якщо кожній елементарній події E_i поставити у взаємно однозначну відповідність деяку точку x_i на числовій осі Ox (або множина Ω з самого початку мала вигляд $\Omega = \{x_1, x_2, ..., x_k\}$), то ряд розподілу ймовірностей на множині Ω набуває вигляду

Табл. 3.10.2

x_i	x_1	x_2	$\dots$	x_k
$P(\{x_i\})$	$P(\{x_1\})$	$P(\{x_2\})$	$\dots$	$P(\{x_k\})$

Якщо точки $(x_i, P(\{x_i\}))$ зобразити на координатній площині, то ламану з вершинами в цих точках називають *многокутником розподілу ймовірностей* (Рис. 3.10.1).

Якщо покласти $\tilde{\Omega} = R^1 \supset \Omega$, $\tilde{S} = \mathscr{B}(R^1) \supset S$, тоді міру $\tilde{P}$ на $(\tilde{\Omega}, \tilde{S})$ можна задати як продовження міри P із (Ω, S) на $(\tilde{\Omega}, \tilde{S})$, тобто із (Ω, S) на $(R^1, \mathscr{B}(R^1))$, поклавши стосовно $G \subset R^1$, $G \in \mathscr{B}(R^1)$, $\tilde{P}(G) = P(G \bigcap \Omega) = P(\bigcup_{\{E_i\} \subset G} \{E_i\}) = \sum_{\{E_i\} \subset G} P(\{E_i\})$.

Тоді функція $F(x)$ розподілу ймовірностей на множині $\tilde{\Omega} \supset \Omega = \{x_1, x_2, ..., x_k\}$ набуває вигляду

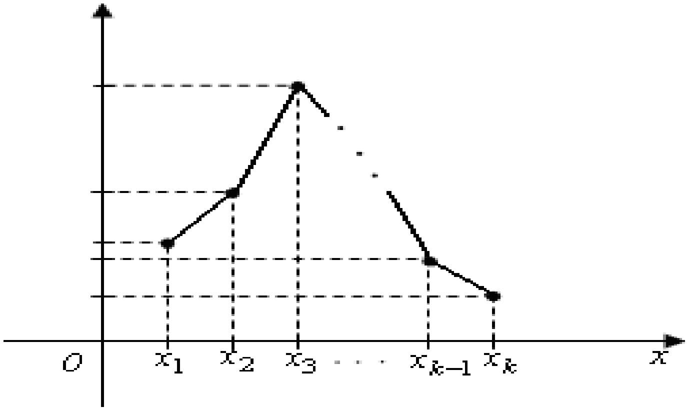


Рис. 3.10.1

$$F(x) = \tilde{P}((-\infty; x)) = \sum_{x_i \in (-\infty, x)} P(\{x_i\}), \ x \in (-\infty; \infty), \ (-\infty; x) \in \mathscr{B}(R^1).$$

Множина значень функції $F(x)$ в розглядуваному випадку скінченна, функція $F(x)$ може набувати відмінних від нуля приростів лише в точках x_i (можливо не в усіх точках x_i, якщо $P(x_i) = 0$ за деяких x_i).

Легко бачити, що виконуються умови:

1. $F(x) \geq 0$;
2. $F(-\infty) = 0$;
3. $F(+\infty) = 1$;
4. Якщо $x_1 < x_2$, то $F(x_1) \leq F(x_2)$;
5. $P([a;b)) = F(b) - F(a)$.

Приклад 3.10.1. Нехай на множині $\Omega = \{-1, 0, 1\}$ задано розподіл ймовірностей: $P(\{-1\}) = \frac{1}{4}$, $P(\{0\}) = \frac{1}{2}$, $P(\{1\}) = \frac{1}{4}$.

Тоді ряд такого розподілу ймовірностей набуває вигляду

x_i	-1	0	1
$P(\{x_i\})$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Многокутник даного розподілу ймовірностей подано на Рис. 3.10.2.

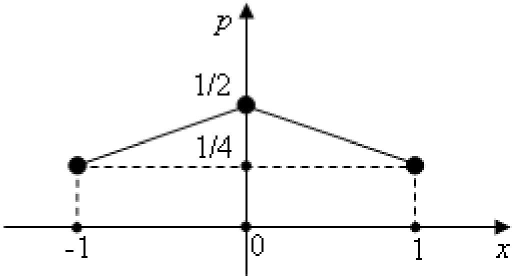


Рис. 3.10.2

Опис функції розподілу ймовірностей $F(x)$ в даному випадку набуває вигляду

$$F(x)=\begin{cases} 0, \text{ коли } & x\le -1, \\ \frac{1}{4}, \text{ коли } & -1<x\le 0, \\ \frac{3}{4}, \text{ коли } & 0<x\le 1, \\ 1, \text{ коли } & 1<x. \end{cases}$$

Графік цієї функції $F(x)$ подано на Рис. 3.10.3.

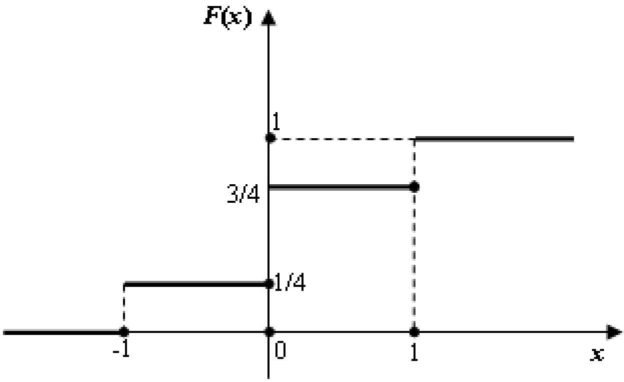


Рис. 3.10.3

Ймовірність попадання на проміжок, наприклад, $\left[-\frac{1}{2};\frac{3}{2}\right)$, (тобто в одну із точок x_i із Ω, що лежить в проміжку $\left[-\frac{1}{2};\frac{3}{2}\right)$

за заданого розподілу ймовірностей дорівнює

$$P\left(\left[-\frac{1}{2},\frac{3}{2}\right)\right)=F\left(\frac{3}{2}\right)-F\left(-\frac{1}{2}\right)=1-\frac{1}{4}=\frac{3}{4},$$

що дорівнює сумі приростів функції $F(x)$ на проміжку $\left[-\frac{1}{2},\frac{3}{2}\right)$:

$$P\left(\left[-\frac{1}{2},\frac{3}{2}\right)\right)=F\left(\frac{3}{2}\right)-F\left(-\frac{1}{2}\right)=P(\{0\})+P(\{1\})=\frac{1}{2}+\frac{1}{4}=\frac{3}{4}.$$

Вправи для самостійного виконання

3.10.1. Перевірити, чи правильні твердження:

1. Кожен ймовірнісний простір визначається за поточковим розподілом ймовірностей.
2. Якщо простір елементарних подій скінченний, то існує відповідний многокутник розподілу ймовірностей.
3. За кожного ймовірнісного простору (Ω, S, P), де Ω не більш ніж зчисленна множина, існує відповідна функція поточкового розподілу ймовірностей.
4. Кожна невід’ємна і неспадна функція може бути функцією дискретного розподілу ймовірностей.
5. Серед значень функції дискретного розподілу ймовірностей обов’язково число 0 є найменшим, а число 1 – найбільшим її значенням.
6. Твердження 5 є правильним, коли простір елементарних подій Ω є скінченним.
7. Функція поточкового розподілу ймовірностей на дискретній множині Ω є кусково-сталою, тобто область визначення цієї функції є об’єднанням зчисленної кількості проміжків, на кожному з яких дана функція є сталою.
8. Через функцію поточкового розподілу ймовірностей на дискретній множині Ω визначається ймовірність будь-якої події з відповідного ймовірнісного простору.
9. Функція поточкового розподілу ймовірностей на дискретній множині Ω розривна принаймні в одній точці.
10. Множина точок розриву функції поточкового розподілу ймовірностей на дискретній множині Ω не більш ніж зчисленна.
11. У функції поточкового розподілу ймовірностей на дискретній множині Ω можуть бути точки розриву другого роду.
12. В кожній точці розриву функція поточкового розподілу ймовірностей на дискретній множині Ω є неперервною зліва.

3.10.2. Навести приклад ймовірнісного простору з функцією поточкового розподілу ймовірностей на дискретній множині Ω:

1) з єдиною точкою розриву;
2) з двома точками розриву;
3) із зчисленною кількістю точок розриву.

3.11. Поточкові розподіли ймовірностей на дискретних множинах точок із R^2

Поточковий розподіл ймовірностей на двохвимірній скінченній множині $\Omega=\{(x_j, y_i), i=\overline{1, m}, j=\overline{1, s}\}$ (Рис. 3.11.1), $S=\{A \mid A=\bigcup\limits_{i\in I}\bigcup\limits_{j\in J}\{(x_j, y_i)\}, I\subset\{1, 2, ..., m\}, J\subset\{1, 2, ..., s\}\}$, можна задати за таблицею виду Табл. 3.11.1,

Табл. 3.11.1

$y_i \backslash x_j$	x_1	x_2	x_3	$\dots$	x_{s-1}	x_s
y_1	p_{11}	p_{12}	p_{13}	$\dots$	p_{1s-1}	p_{1s}
y_2	p_{21}	p_{22}	p_{23}	$\dots$	p_{2s-1}	p_{2s}
y_3	p_{31}	p_{32}	p_{33}	$\dots$	p_{3s-1}	p_{3s}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_{m-1}	p_{m-11}	p_{m-12}	p_{m-13}	$\dots$	p_{m-1s-1}	p_{m-1s}
y_m	p_{m1}	p_{m2}	p_{m3}	$\dots$	p_{ms-1}	p_{ms}

де всі числа p_{ij}, $i\in\overline{1, m}$, $j\in\overline{1, s}$, невід'ємні і такі, що $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{s}p_{ij}=1$.

В механічній інтерпретації за табл. 3.11.1 задано поточковий розподіл одиничної маси на множині $\Omega=\{(x_j, y_i), i\in\overline{1, m}, j\in\overline{1, s}\}$.

Згідно властивості адитивності ймовірності, якщо $A\subset\Omega$, $A\in S$, то

$$P(A)=\sum_{(x_j, y_i)\in A}p_{ij}\,.$$

Якщо покласти $\tilde{\Omega}=R^2\supset\Omega$, $\tilde{S}=\mathscr{B}(R^2)\supset S$, тоді міру $\tilde{P}$ на $(\tilde{\Omega}, \tilde{S})$ можна задати як продовження міри P із (Ω, S) на $(\tilde{\Omega}, \tilde{S})$, тобто із (Ω, S) на $(R^2, \mathscr{B}(R^2))$, поклавши стосовно $G\subset R^2$, $G\in\mathscr{B}(R^2)$,

$$\tilde{P}(G)=P(G\cap\Omega)=P\Big(\bigcup_{\{E_{i,j}\}\subset G}\{E_{i,j}\}\Big)=P\Big(\bigcup_{\{(x_j, y_i)\}\subset G}\{(x_j, y_i)\}\Big)=\sum_{\{(x_j, y_i)\}\subset G}P(\{(x_j, y_i)\})\,\cdot$$

Тоді функція двохвимірного поточкового розподілу ймовірностей на множині $\tilde{\Omega}\supset\Omega=\{(x_j, y_i), i\in\overline{1, m}, j\in\overline{1, s}\}$, $\tilde{S}=\mathscr{B}(R^2)\supset S$, набуває вигляду

$$F(x,y)=\tilde{P}((-\infty;x)\times(-\infty;y))=\sum_{y_i<y}\sum_{x_j<x}p_{ij},$$

де

$$(-\infty;x)\times(-\infty;y)\subset\tilde{\Omega},\ (-\infty;x)\times(-\infty;y)\in\tilde{S}=\mathscr{B}(R^2).$$

Міру $\tilde{P}$ на $\mathscr{B}(R^2)$ можна задати і так:

$$\tilde{P}((-\infty;x)\times(-\infty;y))=P(G_*)+\alpha(P(G^*)-P(G_*)),\ \alpha\in[0;1], \qquad (3.11.1)$$

де

$$G=(-\infty;x)\times(-\infty;y)\in\tilde{S}\supset S,\ G\,\overline{\in}\,S,$$

$$G_*=\bigcup_{\bigcup H_{ij}\subset G\cap\Omega}(\bigcup H_{ij}),\ G^*=\bigcap_{G\cap\Omega\subset\bigcup H_{ij}}(\bigcup H_{ij}),\ H_{ij}=\{(x_j,y_i)\}$$

(див. §3.4, §3.5).

Зауважимо, що у випадку поточкового розподілу ймовірностей на дискретній множині точок Ω завжди буде $G_*=G^*\in S$, $P(G_*)=P(G^*)$, і тому функція $\tilde{P}((-\infty;x)\times(-\infty;y))$, визначена за формулою (3.11.1), є мірою на $\tilde{S}=\mathscr{B}(R^2)$ (див. §2.2).

В механічній інтерпретації значення $F(x_0,y_0)$ є маса, що припадає на множину $((-\infty;x_0)\times(-\infty;y_0))=\{(x,y)\mid x<x_0, y<y_0\}$, за умови, що на множині Ω в площині xOy розподілена одинична маса так, що на точку (x_j,y_i) припадає маса p_{ij}, $i=\overline{1,\ m}$, $j=\overline{1,\ s}$.

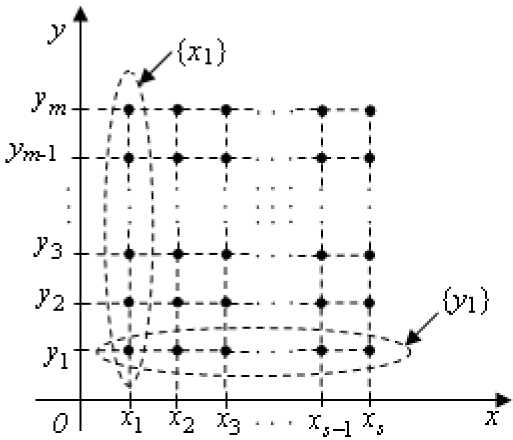


Рис. 3.11.1

Знаючи двохвимірний поточковий розподіл ймовірностей $p_{ij}\geq 0$ на множині $\Omega=\{(x_j,y_i),\ i\in\overline{1,\ m},\ j\in\overline{1,\ s}\}$, $\sum_{i=1}^{m}\sum_{j=1}^{s}p_{ij}=1$, на основі властивості адитивності ймовірності можна знайти ймовірності $P(\{x_j\})=\sum_{i=1}^{m} p_{ij}$, $j\in\overline{1,s}$, – ймовірності попадання в множину точок, абсциси яких дорівнюють x_j , а також $P(\{y_i\})=\sum_{j=1}^{s} p_{ij}$, $i\in\overline{1,\,m}$, – ймовірності попадання в множину точок, ординати яких дорівнюють y_i (Рис. 3.11.1), тобто одновимірні розподіли ймовірностей $P(\{x_j\})$, $j\in\overline{1,\,s}$, та $P(\{y_i\})$, $i\in\overline{1,\,m}$, вздовж осей Ox та Oy відповідно.

Слід зауважити, що знаючи одновимірні розподіли ймовірностей $P(\{x_j\})$, $j\in\overline{1,\,s}$, та $P(\{y_i\}),\, i\in\overline{1,\,m}$, не завжди можна знайти двохвимірний розподіл $P(\{x_j, y_i\}),\, i\in\overline{1,\,m},\; j\in\overline{1,\,s}$.

Коли виявляється, що $P(\{x_j, y_i\})=P(\{x_j\})P(\{y_i\})$ *за всіх* $i\in\overline{1,\,m},\; j\in\overline{1,\,s}$, *тоді розподіли ймовірностей* $P(\{x_j\})$, $j\in\overline{1,\,s}$, *та* $P(\{y_i\}),\, i\in\overline{1,\,m}$, *називаються незалежними відносно ймовірнісної міри* P .

Приклад 3.11.1. Нехай розподіл ймовірностей на множині $\Omega=\{-1,0,1\}\times\{-1,0,1\}\subset R^2$ (Рис. 3.11.2) такий, як поданий в Табл. 3.11.2, $S=\{A \mid A=\bigcup_{i\in I}\bigcup_{j\in J}\{(x_j, y_i)\}, I\subset\{1,2,3\}, J\subset\{1,2,3\}\}$, $x_j\in\{-1,0,1\}, y_i\in\{-1,0,1\}$.

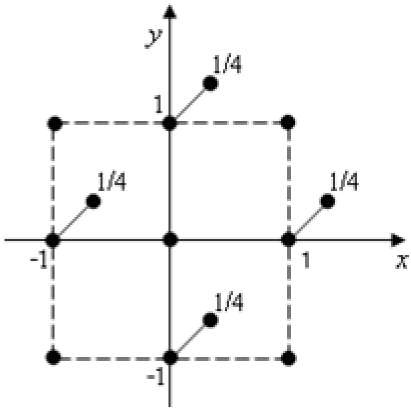


Рис. 3.11.2

Табл. 3.11.2

$y_i \backslash x_j$	-1	0	1
-1	0	1/4	0
0	1/4	0	1/4
1	0	1/4	0

Тоді розподіл $\sum_{i=1}^{3} P(\{(x_j, y_i)\}) = P(\{x_j\})$, $j \in \overline{1,3}$, набуває вигляду

x_j	-1	0	1
$P(\{x_j\})$	1/4	1/2	1/4

Аналогічно розподіл $P(\{y_i\})$, $i \in \overline{1,\,3}$, набуває вигляду

y_i	-1	0	1
$P(\{y_i\})$	1/4	1/2	1/4

Оскільки $P(\{x_j, y_i\}) \neq P(\{x_j\})P(\{y_i\})$ за всіх $i \in \overline{1,\,3}$, $j \in \overline{1,\,3}$, то в даному випадку розподіли $P(\{x_j\})$, $j \in \overline{1,\,3}$, і $P(\{y_i\})$, $i \in \overline{1,\,3}$, не є незалежними відносно заданої на

$$S = \{A \mid A = \bigcup_{i \in I} \bigcup_{j \in J} \{(x_j, y_i)\}, I \subset \{1,2,3\}, J \subset \{1,2,3\}\},$$

$$x_j \in \{-1,\ 0,\ 1\},\ y_i \in \{-1,\ 0,\ 1\},$$

ймовірнісної міри P.

Опис функції $F(x, y)$ двохвимірного розподілу ймовірностей на $(\tilde{\Omega} = R^2,\ \tilde{S} = \mathscr{B}(R^2))$, $\Omega \subset \tilde{\Omega} = R^2$, $S \subset \tilde{S} = \mathscr{B}(R^2)$, в розглядуваному випадку набуває вигляду

$$F(x,y) = \begin{cases} 0, & \text{коли } x \le -1 \text{ або } y \le -1 \text{ або } x \le 0 \text{ і } y \le 0; \\ 1/4, & \text{коли } x > 0 \text{ і } -1 < y \le 0 \text{ або } y > 0 \text{ і } -1 < x \le 0; \\ 1/2, & \text{коли } 0 < x \le 1 \text{ і } 0 < y \le 1; \\ 3/4, & \text{коли } 1 < x \text{ і } 0 < y \le 1 \text{ або } 1 < y \text{ і } 0 < x \le 1; \\ 1, & \text{коли } 1 < x \text{ і } 1 < y. \end{cases}$$

Поверхню, що описується через залежність $z = F(x, y)$, зображено на Рис. 3.11.3.

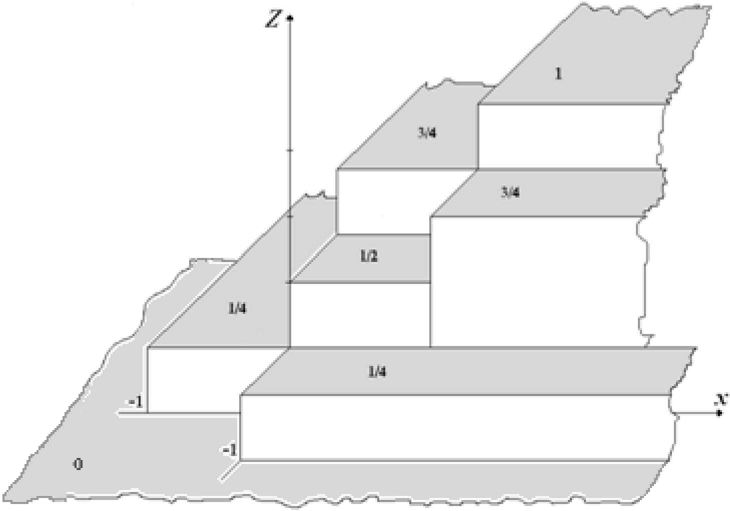


Рис. 3.11.3

Формула для обчислення ймовірності $P([a;b)\times[c;d))$ попадання в прямокутник $[a;b)\times[c;d)$ через значення функції розподілу ймовірностей $F(x,y)$ (див. § 3.4) набуває вигляду (Рис. 3.11.4):

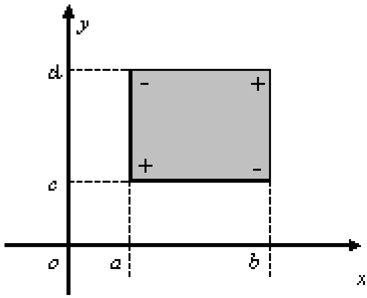


Рис. 3.11.4

$$P([a;b)\times[c;d))=\Delta_{a,b}\Delta_{c,d}F(x,y)=\Delta_{a,b}(F(x,d)-F(x,c))=(F(b,d)-$$
$$-F(b,c))-(F(a,d)-F(a,c))=F(b,d)-F(b,c)-F(a,d)+F(a,c).$$

Якщо в розглянутому прикладі 3.11.1 взяти прямокутник $\left[-\frac{1}{2};\frac{3}{2}\right)\times\left[-\frac{1}{2};\frac{3}{2}\right)$, тоді одержимо

$$P\left(\left[-\frac{1}{2};\frac{3}{2}\right)\times\left[-\frac{1}{2};\frac{3}{2}\right)\right)=$$

$$=F\left(\frac{3}{2};\frac{3}{2}\right)-F\left(-\frac{1}{2};\frac{3}{2}\right)-F\left(\frac{3}{2};-\frac{1}{2}\right)+F\left(-\frac{1}{2};\frac{1}{2}\right)=1-\frac{1}{4}-\frac{1}{4}+0=\frac{1}{2}.$$

Вправи для самостійного виконання

3.11.1. Перевірити, чи правильні твердження:

1. Кожен скінченний простір елементарних подій Ω можна вважати двохвимірним простором.
2. Кожен скінченний двохвимірний простір елементарних подій Ω можна вважати одновимірним простором.
3. За кожним двохвимірним дискретним розподілом ймовірностей визначається рівно два одновимірних дискретних розподіли ймовірностей.
4. Кожен двохвимірний поточковий розподіл ймовірностей визначається за двома незалежними одновимірними поточковими розподілами ймовірностей.
5. Стосовно значень функції двохвимірного поточкового розподілу ймовірностей на скінченній множині точок 0 є найменшим, а 1 – найбільшим значенням.
6. Функція двохвимірного поточкового розподілу ймовірностей на скінченній множині точок є кусково-сталою функцією.
7. Функція двохвимірного поточкового розподілу ймовірностей на скінченній множині точок є зростаючою за кожною змінною.
8. Через функцію двохвимірного поточкового розподілу ймовірностей на скінченній множині точок визначається ймовірність будь-якої події з відповідного ймовірнісного простору $(R^2, \mathscr{B}(R^2), \widetilde{P})$.
9. Функція двохвимірного поточкового розподілу ймовірностей на скінченній множині точок може набувати лише скінченну кількість значень.

3.11.2. Навести приклад функції двохвимірного розподілу ймовірностей, в множині значень якої міститься:
1) два елементи; 2) три елементи.

3.12. Абсолютно неперервні розподіли ймовірностей на $(R^1, \mathscr{B}(R^1))$

Одновимірний абсолютно неперервний розподіл ймовірностей на $(R^1, \mathscr{B}(R^1))$ можна задати за допомогою щільності розподілу ймовірностей $f(x)$ такої, що:

1. $0 \le f(x) \le c < \infty$,
2. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$

або за допомогою функції розподілу ймовірностей $F(x)$ такої, що:

1. $F(x) \ge 0$;
2. $F(-\infty) = 0$;
3. $F(\infty) = 1$;
4. $F(x)$ неспадна, тобто якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$,
5. $F(x)$ неперервна на $(-\infty;\infty)$.

В даному разі $F(x) = P((-\infty;x)) = \int\limits_{-\infty}^{x} f(t)dt$,

$$f(x) = \frac{d}{dx}F(x) \text{ майже всюди на } R^1.$$

Якщо $A = \langle \alpha;\beta \rangle$, то

$$P(A) = \int\limits_{\langle \alpha;\beta \rangle} f(x)dx = F(\beta) - F(\alpha);$$

якщо

$$A = \bigcup_i \langle \alpha_i;\beta_i \rangle, \ (\alpha_i;\beta_i) \bigcap (\alpha_j;\beta_j) = \varnothing, \text{ коли } i \ne j,$$

то за властивістю адитивності ймовірності

$$P(A) = P(\bigcup_i \langle \alpha_i;\beta_i \rangle) = \sum_i P(\langle \alpha_i;\beta_i \rangle).$$

Приклад 3.12.1. Нехай на множині $\Omega = (-\infty;\infty)$ задано абсолютно неперервний розподіл ймовірностей через щільність розподілу $f(x)$ (Рис. 3.12.1):

$$f(x) = \begin{cases} c(1-|x|), & \text{коли } |x| \le 1, \\ 0, & \text{коли } |x| \ge 1. \end{cases}$$

Потрібно знайти константу c, функцію розподілу ймовірностей $F(x)$, ймовірність $P\left(\left[-\frac{1}{2};\frac{1}{2}\right)\right)$ попадання на проміжок $[-\frac{1}{2};\frac{1}{2})$.

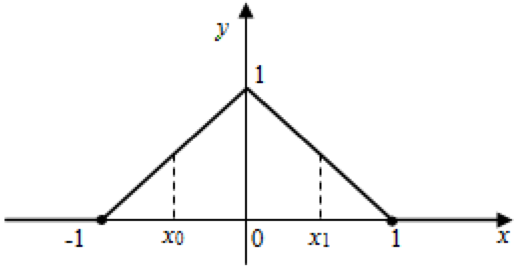


Рис. 3.12.1

Враховуючи властивості щільності розподілу ймовірностей $f(x)$, одержимо:

1) із вимоги $f(x)\ge 0$ дістаємо, що $c \ge 0$;

2) із вимоги $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1$ в даному разі випливає

$$1=\int\limits_{-\infty}^{\infty} f(x)dx=\int\limits_{-\infty}^{-1} 0\cdot dx+\int\limits_{-1}^{1} c\cdot(1-|x|)dx+\int\limits_{1}^{\infty}0\cdot dx=2c\int\limits_{0}^{1}(1-x)dx=$$

$$=2c\left(x-\frac{x^2}{2}\right)\Bigg|_0^1=2c\left(1-\frac{1}{2}\right)=c,$$

отже $c=1$.

Враховуючи, що за абсолютно неперервного розподілу ймовірностей $F(x)=P((-\infty;\,x))=\int\limits_{-\infty}^{x} f(t)\,dt$, $x\in(-\infty;\infty)$, в даному випадку одержимо:

$$F(x)=\int\limits_{-\infty}^{x} f(t)dt=\begin{cases} 0, & \text{коли} \quad x\le -1,\\ \dfrac{(x+1)^2}{2}, & \text{коли} \quad -1\le x\le 0,\\ 1-\dfrac{(1-x)^2}{2}, & \text{коли} \quad 0\le x\le 1,\\ 1, & \text{коли} \quad 1\le x.\end{cases}$$

(бо коли $x_0 \in [-1;\ 0]$, $\int\limits_{-\infty}^{x_0} f(x)dx$ дорівнює площі рівнобедреного прямокутного трикутника з основою $(-1; x_0)$ довжиною x_0+1, а $\int\limits_{-\infty}^{x_1} f(t)dt$, коли $x_1 \in [0;1]$, дорівнює площі трикутника з вершинами $(-1,0)$, $(0,1)$, $(1,0)$ без площі рівнобедреного прямокутного трикутника з основою $(x_1;1)$ довжиною $1-x_1$ (Рис. 3.12.1)).

Графік функції $F(x)$ подано на Рис. 3.12.2.

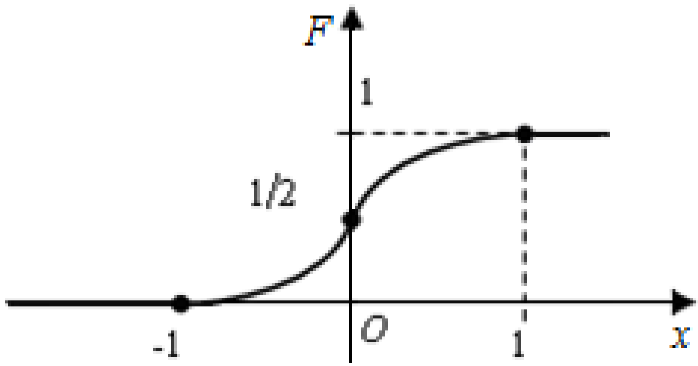


Рис. 3.12.2.

Обчислюючи ймовірність $P\left(\left[-\frac{1}{2};\frac{1}{2}\right)\right)$, знайдемо

$$P\left(\left[-\frac{1}{2};\frac{1}{2}\right)\right) = \int\limits_{-1/2}^{1/2} f(x)dx = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = \frac{7}{8} - \frac{1}{8} = \frac{6}{8} = \frac{3}{4}.$$

Оскільки в даному випадку розподіл ймовірностей абсолютно неперервний, то

$$P\left(\left[-\frac{1}{2};\frac{1}{2}\right)\right) = P\left(\left(-\frac{1}{2};\frac{1}{2}\right]\right) = P\left(\left(-\frac{1}{2};\frac{1}{2}\right)\right) = P\left(\left[-\frac{1}{2};\frac{1}{2}\right]\right),$$

бо за будь якого x_0 імовірність $P(\{x_0\})$ попадання в окрему точку x_0 за неперервного розподілу ймовірностей дорівнює нулеві, що слідує із співвідношень:

$$P(\{x_0\}) = \lim_{\Delta x \to 0} (F(x_0 + \Delta x) - F(x_0)) = 0$$

в силу неперервності функції розподілу ймовірностей $F(x)$.

Вправи для самостійного виконання

3.12.1. Перевірити, чи правильні твердження:

1. Кожен ймовірнісний простір визначається за неперервним розподілом ймовірностей.
2. Кожен неперервний розподіл ймовірностей можна задати за допомогою щільності розподілу ймовірностей.
3. Якщо розподіл ймовірностей не є неперервним, то відповідний простір елементарних подій є скінченним.
4. Твердження, обернене до 3, є правильним.
5. Кожна неперервна функція розподілу ймовірностей є абсолютно неперервною.
6. Якщо $F(x) = \int\limits_{-\infty}^{x} f(t)\,dt$, $x \in R^1$, то $F(x)$ є функцією розподілу ймовірностей.
7. Абсолютно неперервна функція розподілу ймовірностей є скрізь диференційовною.
8. Серед значень абсолютно неперервної функції розподілу ймовірностей може не бути ні найменшого, ні найбільшого значень.
9. Множиною значень будь-якої функції розподілу ймовірностей є відрізок $[0;1]$.
10. Через функцію абсолютно неперервного розподілу ймовірностей визначається ймовірність будь-якої події з відповідного ймовірнісного простору.
11. В разі функції абсолютно неперервного розподілу ймовірностей відповідна щільність розподілу:

а) завжди неперервна;
б) у неї може бути скінченна кількість точок розриву;
в) у неї може бути зчисленна кількість точок розриву;
г) така функція завжди обмежена.

12. Якщо $F(x) = \int\limits_{-\infty}^{x} f(t)\,dt$ і x_0 – точка розриву функції $f(x)$, то в цій точці і функція $F(x)$:

а) може бути розривною;
б) може бути диференційовною;
в) завжди недиференційовна.

3.12.2. Побудувати функцію абсолютно неперервного розподілу ймовірностей таку, що відповідна щільність розподілу є:

а) скрізь неперервною;
б) розривна лише в одній точці;
в) розривна в двох точках.

3.13. Двохвимірні абсолютно неперервні розподіли ймовірностей на $(R^2, \mathscr{B}(R^2))$

Двохвимірний абсолютно неперервний розподіл ймовірностей на $(R^2, \mathscr{B}(R^2))$ можна задати за допомогою щільності двохвимірного розподілу ймовірностей $f(x,y)$ такої, що:

1. $0 \le f(x,y) \le c < \infty$, $(x,y) \in R^2$;
2. $\iint\limits_{R^2} f(x,y)dxdy = 1$,

або за допомогою функції двохвимірного розподілу ймовірностей $F(x,y) = P((-\infty;\, x)\times(-\infty;\, y))$, $(-\infty;\, x)\times(-\infty;\, y) \in \mathscr{B}(R^2)$, такої, що:

1. $F(x,y) \ge 0$;
2. $F(x,y) \to 0$, коли $x \to -\infty$ або $y \to -\infty$;
3. $F(x,y)$ неспадна за кожним із своїх аргументів.
4. $F(x,y) \to 1$, коли $x \to \infty$ і $y \to \infty$;
5. $F(x,y)$ неперервна на R^2.

В даному разі

$$F(x,y) = P((-\infty;\, x)\times(-\infty;\, y)) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(x,y)dxdy,$$

$$f(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y) \text{ майже всюди на } R^2.$$

За довільної події $A \in S = \mathscr{B}(R^2)$ $P(A) = \iint\limits_{A} f(x,y)\,dxdy$.

Якщо A – прямокутник виду $A = [a;b]\times[c;d)$, тоді

$$P(A) = P([a;\, b)\times[c;\, d)) = \int\limits_{a}^{b}\left\{\int\limits_{c}^{d} f(x,y)dy\right\}dx =$$

$$= F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

В механічній інтерпретації значення $F(x_0, y_0)$ є маса, що припадає на множину $(-\infty;\, x_0)\times(-\infty;\, y_0) = \{(x,y) \mid x < x_0 \text{ і } y < y_0\}$, за умови, що на множині $\Omega = G$ в площині xOy розподілена одинична маса із щільністю $f(x,y)$.

Знаючи двохвимірний неперервний розподіл ймовірностей, можна знайти відповідні одновимірні розподіли ймовірностей вздовж осі Ox та вздовж осі Oy:

$$P((-\infty;\ x)\times(-\infty;\ +\infty)) = F(x,\ +\infty) = P_1((-\infty;\ x)) = F_1(x),$$
$$P((-\infty;\ \infty)\times(-\infty;\ y)) = F(+\infty,\ y) = P_2((-\infty;\ y)) = F_2(y)$$
або через щільність розподілу ймовірностей:
$$P((-\infty;\ x)\times(-\infty;+\infty)) = \int_{-\infty}^{x}\left\{\int_{-\infty}^{\infty} f(x,y)dy\right\}dx = \int_{-\infty}^{x} f_1(x)dx,$$
$$P((-\infty;\ \infty)\times(-\infty;\ y)) = \int_{-\infty}^{y}\left\{\int_{-\infty}^{\infty} f(x,y)dx\right\}dy = \int_{-\infty}^{y} f_2(y)dy,$$
де $f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy$, $f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx$.

Слід зауважити, що знаючи функції $F_1(x)$ та $F_2(y)$ чи щільності $f_1(x)$ та $f_2(y)$ одновимірних розподілів ймовірностей, не завжди можна знайти функцію $F(x,y)$ чи щільність $f(x,y)$ двохвимірного розподілу ймовірностей.

Коли виявляється, що
$$F(x,y) = F_1(x)F_2(y) \text{ чи } f(x,y) = f_1(x)f_2(y),$$
тоді *розподіли ймовірностей*, що визначаються за функціями $F_1(x)$ та $F_2(y)$ чи за щільностями $f_1(x)$ та $f_2(y)$, називаються *незалежними відносно ймовірнісної міри* P, заданої на $S = \mathscr{B}(R^2)$ через щільність $f(x,y)$ або через функцію $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy$ розподілу ймовірностей на множині $\tilde{\Omega} = R^2$.

Приклад 3.13.1. Нехай задано рівномірний розподіл ймовірностей в квадраті $[0;1]\times[0;1]$.

Тоді опис щільності розподілу ймовірностейна множині $\tilde{\Omega} = R^2$ набуває вигляду (Рис. 3.13.1):
$$f(x,y) = \begin{cases} 1, & \text{коли } (x,y)\in[0;1]\times[0;1], \\ 0, & \text{коли } (x,y)\,\overline{\in}\,[0;1]\times[0;1]. \end{cases}$$

За так заданого розподілу ймовірностей одержимо (Рис. 3.13.1):
$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy = \begin{cases} 0, & \text{коли } x\,\overline{\in}\,[0;1], \\ 1, & \text{коли } x\in[0;1]. \end{cases}$$

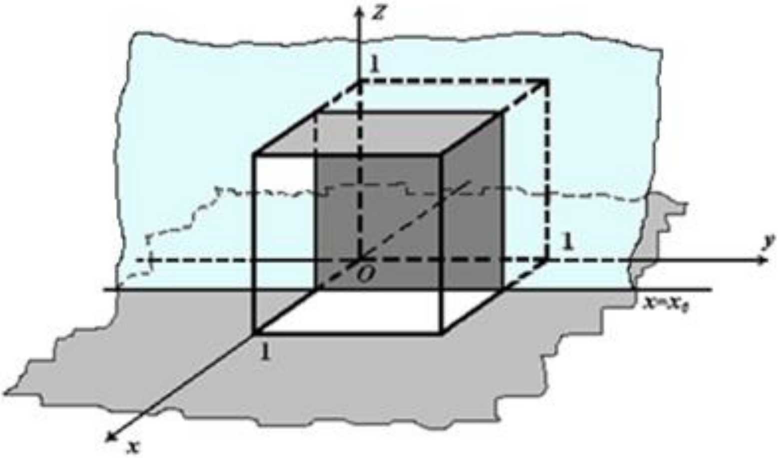


Рис. 3.13.1

Аналогічно

$$f_2(y)=\int\limits_{-\infty}^{\infty} f(x,y)dy=\begin{cases}0, \text{ коли } y\,\overline{\in}\,[0;1],\\ 1, \text{ коли } y\in[0;1].\end{cases}$$

Тоді $f(x,y)=f_1(x)f_2(y)=\begin{cases}1, & \text{коли } (x,y)\in[0;1]\times[0;1],\\ 0, & \text{коли } (x,y)\,\overline{\in}\,[0;1]\times[0;1].\end{cases}$

Отже в даному випадку розподіли ймовірностей, що визначаються за щільностями $f_1(x)$ та $f_2(y)$, виявляються незалежними відносно ймовірнісної міри P, заданої через щільність $f(x,y)$.

Стосовно $F(x,y)$ одержуємо (Рис. 3.13.2):

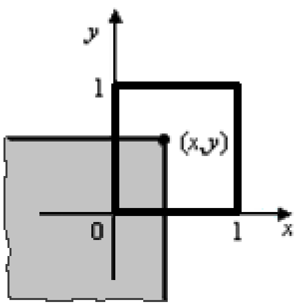


Рис. 3.13.2

$$F(x,y)=P((-\infty;\,x)\times(-\infty;\,y))=\int\limits_{-\infty}^{x}\left\{\int\limits_{-\infty}^{y} f(x,y)dy\right\}dx=$$

$$= \begin{cases} 0, \text{ коли } x \le 0 \text{ або } y \le 0, \\ xy, \text{ коли } 0 \le x \le 1 \text{ і } 0 \le y \le 1, \\ x, \text{ коли } 0 \le x \le 1 \text{ і } 1 \le y, \\ y, \text{ коли } 1 \le x \text{ і } 0 \le y \le 1, \\ 1, \text{ коли } 1 \le x \text{ і } 1 \le y. \end{cases}$$

Поверхню, що описується через залежність $z = F(x, y)$, зображено на Рис. 3.13.3.

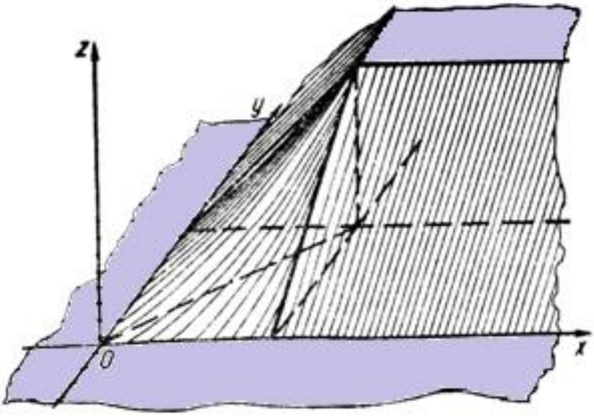


Рис. 3.13.3

Із виразу стосовно $F(x, y)$ одержуємо:

$$F(x, +\infty) = F_1(x) = \begin{cases} 0, \text{ коли } x \le 0, \\ x, \text{ коли } 0 \le x \le 1, \\ 1, \text{ коли } 1 \le x. \end{cases}$$

Аналогічно

$$F(+\infty, y) = F_2(y) = \begin{cases} 0, \text{ коли } y \le 0, \\ y, \text{ коли } 0 \le y \le 1, \\ 1, \text{ коли } 1 \le y. \end{cases}$$

Як бачимо, в розглядуваному випадку $F(x, y) = F_1(x) F_2(y)$, що видно також із співвідношень:

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dx dy = \int_{-\infty}^{x} \int_{-\infty}^{y} f_1(x) f_2(y) dx dy =$$

$$= \int_{-\infty}^{x} f_1(x) dx \int_{-\infty}^{y} f_2(y) dy = F_1(x) F_2(y).$$

Вправи для самостійного виконання

3.13.1. Перевірити, чи правильні твердження:

1. Кожен двохвимірний абсолютно неперервний розподіл ймовірностей можна задати за допомогою щільності розподілу ймовірностей.
2. Кожна функція двохвимірного неперервного розподілу ймовірностей є зростаючою за кожною змінною.
3. Існують функції двохвимірних неперервних розподілів ймовірностей, що не є абсолютно неперервними.
4. Якщо простір елементарних подій Ω є скінченним, то відповідна функція розподілу ймовірностей не є абсолютно неперервною.
5. Твердження, обернене до 4, є правильним.
6. Кожен абсолютно неперервний двохвимірний розподіл ймовірностей цілком визначається за двома незалежними одновимірними абсолютно неперервними розподілами ймовірностей.
7. Щільність розподілу ймовірностей за функції абсолютно неперервного розподілу ймовірностей: а) завжди неперервна; б) може бути розривною в одній точці; в) може бути необмеженою.

3.13.2. Побудувати функцію двохвимірного неперервного розподілу ймовірностей, за якої:

а) існує неперервна щільність розподілу;

б) існує щільність розподілу, в множині значень якої міститься лише два елементи;

в) не існує щільності розподілу.

3.14. Деякі числові характеристики поточкових розподілів ймовірностей на дискретних множинах точок

До найважливіших числових характеристик одновимірного поточкового розподілу ймовірностей $(x_i, P(\{x_i\}))$, $i \in \overline{1, k}$, (або $i \in N$), на дискретній множині $\Omega = \{x_1, x_2, \ldots, x_k\}$ (або $\Omega = \{x_1, x_2, \ldots\}$), відносяться координата (абсциса) x_c центра розподілу (розсіювання) ймовірностей та дисперсія D розподілу ймовірностей. Координата x_c центра поточкового розподілу ймовірностей $(x_i, P(\{x_i\}))$, $i \in \overline{1, k}$, (або $i \in N$), на дискретній множині $\Omega = \{x_1, x_2, \ldots x_k\}$ (або $\Omega = \{x_1, x_2, \ldots x_k, \ldots\}$), обчислюється за формулою:

$$x_c = \sum_{i=1}^{k} x_i P(\{x_i\}) \text{ або } x_c = \sum_{i=1}^{\infty} x_i P(\{x_i\})$$

за умови збіжності останнього ряду.

Очевидно центр розподілу ймовірностей буде знаходитись ближче до тих точок x_i, на які припадають більші ймовірності.

В механічній інтерпретації x_c є координатою центра мас загальної одиничної маси, розподіленої на множині $\Omega = \{x_1, x_2, \ldots\}$ так, що на точку x_i припадає маса $P(\{x_i\})$, $i \in \overline{1, k}$ або $i \in N$.

Дисперсією одновимірного поточкового розподілу ймовірностей $(x_i, P(\{x_i\}))$, $i \in \overline{1, k}$ або $i \in N$, з центром розподілу x_c називається число (скінченне або нескінченне)

$$D = \sum_{i=1}^{k} (x_i - x_c)^2 P(\{x_i\}) \text{ або } D = \sum_{i=1}^{\infty} (x_i - x_c)^2 P(\{x_i\}).$$

За допомогою дисперсії характеризують величину розсіювання (чи скупченість) ймовірностей $P(\{x_i\})$, $i \in \overline{1, k}$ або $i \in N$, навколо центра розподілу ймовірностей x_c.

Часто розсіювання характеризують за допомогою середнього квадратичного відхилення σ, яке визначається за формулою $\sigma = \sqrt{D}$.

В механічній інтерпретації дисперсія є моментом інерції системи мас $P(\{x_i\})$, $\sum_i P(\{x_i\}) = 1$, відносно центра розсіювання x_c.

Приклад 3.14.1. Нехай задано рівномірний поточковий розподіл ймовірностей на множині $\Omega=\{x_1, x_2, ..., x_k\}$, тобто $P(\{x_i\})=\dfrac{1}{k}$, $i\in\overline{1,\ k}$.

Тоді

$$x_c=\sum_{i=1}^{k}x_i\cdot\frac{1}{k}=\frac{1}{k}\sum_{i=1}^{k}x_i\ ,$$

тобто в даному випадку координата центра розподілу ймовірностей дорівнює середньому арифметичному координат x_i, $i\in\overline{1,\ k}$.

Обчислюючи дисперсію, одержимо:

$$D=\sum_{i=1}^{k}(x_i-\frac{1}{k}\sum_{i=1}^{k}x_i)^2\,\frac{1}{k}=\frac{1}{k}\sum_{i=1}^{k}(x_i-\frac{1}{k}\sum_{i=1}^{k}x_i)^2\ .$$

В разі двохвимірного поточкового розподілу ймовірностей $P(\{x_i, y_i\})$, $i\in\overline{1,\ m}$, $j\in\overline{1,\ s}$, на $\tilde{\Omega}=R^2$ координати центра розподілу ймовірностей (точки $C(x_c, y_c)$) обчислюються за формулами

$$x_c=\sum_{i=1}^{m}\sum_{j=1}^{s}x_jP(\{x_j, y_i\})=\sum_{j=1}^{s}x_j\sum_{i=1}^{m}P(\{x_j, y_i\})=\sum_{j=1}^{s}x_jP(\{x_j\}),$$

$$y_c=\sum_{i=1}^{m}\sum_{j=1}^{s}y_iP(\{x_j, y_i\})=\sum_{i=1}^{m}y_i\sum_{j=1}^{s}P(\{x_j, y_i\})=\sum_{i=1}^{m}y_iP(\{y_i\}).$$

Дисперсії розподілів $P(\{x_j\})$, $j\in\overline{1,\ s}$, та $P(\{y_i\})$, $i\in\overline{1,\ m}$, вздовж осей Ox та Oy відповідно обчислюються за формулами

$$D_1=\sum_{i=1}^{m}\sum_{j=1}^{s}(x_j-x_c)^2P(\{x_j, y_i\})=\sum_{j=1}^{s}(x_j-x_c)^2P(\{x_j\}),$$

$$D_2=\sum_{i=1}^{m}\sum_{j=1}^{s}(y_i-y_c)^2P(\{x_j, y_i\})=\sum_{i=1}^{m}(y_i-y_c)^2P(\{y_i\}).$$

Приклад 3.14.2. Нехай задано двохвимірний поточковий розподіл ймовірностей на множині $\{-1,\ 0,\ 1\}\times\{-1,\ 0,\ 1\}$ за таблицею (Рис. 3.14.1)

$y_i \backslash x_j$	-1	0	1
-1	0	1/4	0
0	1/4	0	1/4
1	0	1/4	0

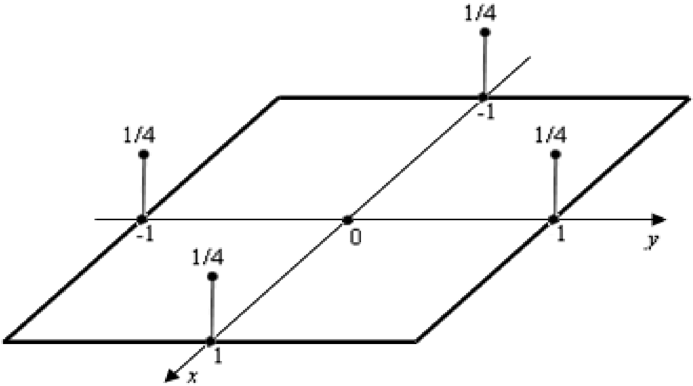


Рис. 3.14.1

Тоді

$$x_c = \sum_{i=1}^{3}\sum_{j=1}^{3} x_j P(\{x_j, y_i\}) = \sum_{j=1}^{3} x_j \sum_{i=1}^{3} P(\{x_j, y_i\}) =$$

$$= -1\cdot(0+\frac{1}{4}+0)+0\cdot(\frac{1}{4}+0+\frac{1}{4})+1\cdot(0+\frac{1}{4}+0)=0,$$

$$y_c = \sum_{i=1}^{3}\sum_{j=1}^{3} y_i P(\{x_j, y_i\}) = \sum_{i=1}^{3} y_i \sum_{j=1}^{3} P(\{x_j, y_i\}) =$$

$$= -1\cdot(0+\frac{1}{4}+0)+0\cdot(\frac{1}{4}+0+\frac{1}{4})+1\cdot(0+\frac{1}{4}+0)=0.$$

Таким чином центром даного розподілу ймовірностей є точка $(0, 0)$.

Обчислюючи дисперсії, одержимо:

$$D_1 = \sum_{i=1}^{3}\sum_{j=1}^{3} (x_j - x_c)^2 P(\{x_j, y_i\}) = \sum_{j=1}^{3} (x_j - x_c)^2 \sum_{i=1}^{3} P(\{x_j, y_i\}) =$$

$$=1\cdot(0+\frac{1}{4}+0)+0\cdot(\frac{1}{4}+0+\frac{1}{4})+1\cdot(0+\frac{1}{4}+0)=\frac{2}{4}=\frac{1}{2}.$$

Аналогічно

$$D_2 = \sum_{i=1}^{3}\sum_{j=1}^{3} (y_i - y_c)^2 P(\{x_j, y_i\}) = \sum_{i=1}^{3} (y_i - y_c)^2 \sum_{j=1}^{3} P(\{x_j, y_i\}) =$$

$$=1\cdot(0+\frac{1}{4}+0)+0\cdot(\frac{1}{4}+0+\frac{1}{4})+1\cdot(0+\frac{1}{4}+0)=\frac{2}{4}=\frac{1}{2}.$$

Вправи для самостійного виконання

3.14.1. Перевірити, чи правильні твердження:

1. За кожного одновимірного поточкового розподілу ймовірностей на дискретній множині точок існує центр розсіювання ймовірностей.

2. Якщо існує центр розсіювання одновимірного поточкового розподілу ймовірностей на дискретній множині точок, то координата цього центра є скінченним числом.

3.За кожного одновимірного поточкового розподілу ймовірностей на дискретній множині точок існує дисперсія.

4. Якщо існує центр розсіювання одновимірного поточкового розподілу ймовірностей на дискретній множині точок, то існує й дисперсія цього розподілу.

5. Твердження, обернене до 4, є правильним.

6. Дисперсія завжди є скінченним числом.

7. Якщо простір Ω елементарних подій скінченний, то центр розсіювання і дисперсія відповідного розподілу ймовірностей скінченні.

8. Твердження, обернене до 7, є правильним.

9. Дисперсія дорівнює квадрату середнього квадратичного відхилення, тобто $D = \sigma^2$, а тому $\sigma = \pm\sqrt{D}$.

3.14.2. Переформулювати твердження 1-9 із завдання 3.14.1 на випадок двохвимірного поточкового розподілу ймовірностей на дискретній множині точок та перевірити, чи правильні ці твердження.

3.14.3. Побудувати одновимірний поточковий розподіл ймовірностей на дискретній множині точок, за якого:

1) $x_c = 0,\ D = 0$;

2) $x_c = 0,\ D \neq 0$;

3) x_c не існує, та дослідити, чи може тоді існувати D;

4) x_c існує, проте $D = +\infty$.

3.15. Деякі числові характеристики абсолютно неперервних розподілів ймовірностей

До найважливіших числових характеристик абсолютно неперервного одновимірного розподілу ймовірностей на $\tilde{\Omega} = R^1$, $S = \mathscr{B}(R^1)$, що визначається за щільністю $f(x)$, відносяться координата (абсциса) центра розподілу ймовірностей та дисперсія розподілу ймовірностей.

Координата x_c центра одновимірного абсолютно неперервного розподілу ймовірностей із щільністю $f(x)$ обчисляється за формулою

$$x_c = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

за умови збіжності останнього невласного інтеграла.

В механічній інтерпретації x_c є координатою центра мас загальної одиничної маси, розподіленої вздовж осі Ox із щільністю $f(x)$.

Дисперсією одновимірного абсолютно неперервного розподілу ймовірностей називається число

$$D = \int_{-\infty}^{\infty} (x - x_c)^2 f(x)\, dx .$$

За допомогою дисперсії характеризують величину розсіювання (чи скупченості) ймовірностей навколо центра розподілу ймовірностей x_c .

Часто розсіювання характеризують за допомогою середнього квадратичного відхилення σ , яке визначається за формулою

$$\sigma = \sqrt{D} .$$

В механічній інтерпретації дисперсія є момент інерції одиничної маси, розподіленої вздовж осі Ox із щільністю $f(x)$, відносно центра мас.

Приклад 3.15.1. Нехай задано неперервний розподіл ймовірностей на множині $\Omega = (-\infty; \infty)$ через щільність $f(x)$ (Рис. 3.15.1):

$$f(x) = \begin{cases} 1 - |x|, \text{ коли } x \in [-1; 1], \\ 0, \text{ коли } x \overline{\in} [-1; 1]. \end{cases}$$

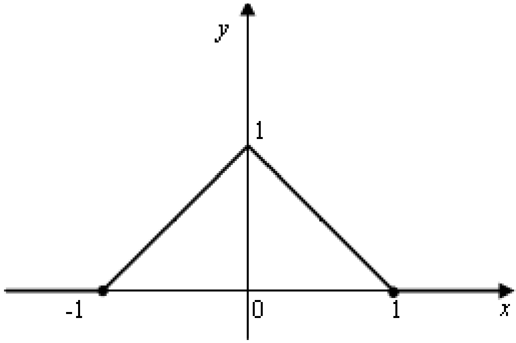


Рис. 3.15.1

Тоді

$$x_c = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{-1} x \cdot 0 \cdot dx + \int_{-1}^{1} x(1-|x|)\, dx + \int_{1}^{\infty} x \cdot 0 \cdot dx = \int_{-1}^{1} x(1-|x|)\, dx =$$

$$= \int_{-1}^{0} x(1+x)\, dx + \int_{0}^{1} x(1-x)\, dx =$$

$$= \left(\frac{x^2}{2} + \frac{x^3}{3}\right)\Bigg|_{-1}^{0} + \left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Bigg|_{0}^{1} =$$

$$= -\left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) = 0$$

$$D = \int_{-\infty}^{\infty} (x-0)^2\, f(x)\, dx =$$

$$= \int_{-\infty}^{-1} x^2 \cdot 0 \cdot dx + \int_{-1}^{1} x^2\, (1-|x|)\, dx + \int_{1}^{\infty} x^2 \cdot 0 \cdot dx =$$

$$= 2\int_{0}^{1} x^2 (1-x)\, dx = 2\left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Bigg|_{0}^{1} = 2 \cdot \left(\frac{1}{3} - \frac{1}{4}\right) = 2 \cdot \left(\frac{4}{12} - \frac{3}{12}\right) = \frac{1}{6}\ .$$

В разі двохвимірного абсолютно неперервного розподілу ймовірностей на множині $\tilde{\Omega} = R^2$, $S = \mathscr{B}(R^2)$, із щільністю $f(x, y),\ (x, y) \in R^2$, координати центра розподілу ймовірностей обчислюються за формулами:

$$x_c=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,f(x,y)\,dx\,dy=\int_{-\infty}^{\infty}x\left\{\int_{-\infty}^{\infty}f(x,y)\,dy\right\}dx=\int_{-\infty}^{\infty}x\,f_1(x)\,dx\,,$$

$$y_c=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\,f(x,y)\,dx\,dy=\int_{-\infty}^{\infty}y\left\{\int_{-\infty}^{\infty}f(x,y)\,dx\right\}dy=\int_{-\infty}^{\infty}y\,f_2(y)\,dy\,.$$

за умови збіжності невласних інтегралів.

Дисперсії двохвимірного абсолютно неперервного розподілу ймовірностей, за допомогою яких характеризують розсіювання ймовірностей навколо центра в напрямах, паралельних осям Ox та Oy відповідно, обчислюються за формулами:

$$D_1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-x_c)^2\;f(x,y)\,dx\,dy=$$

$$=\int_{-\infty}^{\infty}(x-x_c)^2\left\{\int_{-\infty}^{\infty}f(x,y)\,dy\right\}dx=\int_{-\infty}^{\infty}(x-x_c)^2\;f_1(x)\,dx,$$

$$D_2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-y_c)^2\;f(x,y)\,dx\,dy=$$

$$=\int_{-\infty}^{\infty}(y-y_c)^2\left\{\int_{-\infty}^{\infty}f(x,y)\,dx\right\}dy=\int_{-\infty}^{\infty}(y-y_c)^2\;f_2(y)\,dy.$$

Приклад 3.15.2. Нехай задано рівномірний розподіл ймовірностей на множині $G=\{(x,y)\mid\;|x|+|y|\le 1\}\subset\tilde{\Omega}=R^2$.

Опис щільності такого розподілу набуває вигляду (Рис. 3.15.2 а)):

$$f(x,y)=\begin{cases}\dfrac{1}{2},\text{ коли}\,(x,y)\in G,\\ 0,\quad\text{коли }(x,y)\,\overline{\in}\,G,\quad (x,y)\in R^2.\end{cases}$$

Тоді (Рис. 3.15.2 а), б))

$$f_1(x)=\int_{-\infty}^{\infty}f(x,y)\,dy=\begin{cases}0,\text{ коли } x\,\overline{\in}[-1;1],\\ \dfrac{1}{2}\cdot(1+x)\cdot 2,\text{ коли } x\in[-1;0],\\ \dfrac{1}{2}\cdot(1-x)\cdot 2,\text{ коли } x\in[0;1].\end{cases}$$

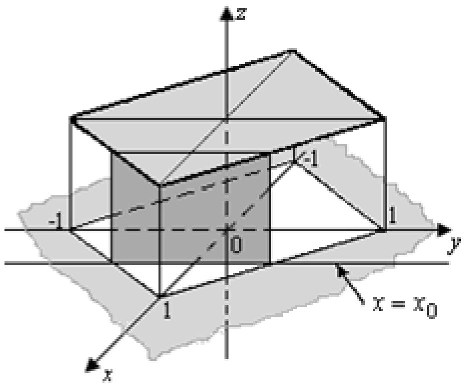


Рис. 3.15.2 а)

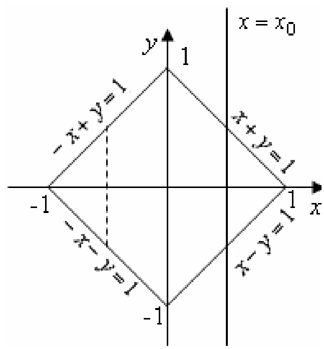


Рис. 3.15.2 б)

Аналогічно

$$f_2(y)=\int\limits_{-\infty}^{\infty} f(x,y)\,dy=\begin{cases} 0, \text{ коли } y\,\overline{\in}[-1;1], \\ (1+y), \text{ коли } y\in[-1;0], \\ (1-y), \text{ коли } y\in[0;1]. \end{cases}$$

Графіки функцій $f_1(x)$, $f_2(y)$ подано на Рис. 3.15.3.

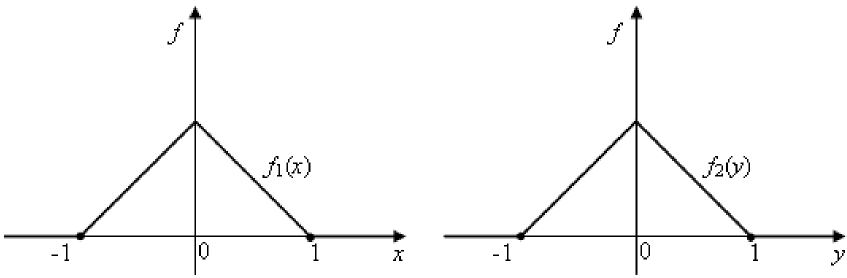


Рис. 3.15.3

Обчислюючи координати центра, одержимо:

$$x_c=\int\limits_{-\infty}^{\infty} x\, f_1(x)dx=\int\limits_{-\infty}^{-1} x\cdot 0\cdot dx+\int\limits_{-1}^{0} x(1+x)dx+\int\limits_{0}^{1} x(1-x)dx+\int\limits_{1}^{\infty} x\cdot 0\cdot dx=$$

$$=\int\limits_{-1}^{0}(x+x^2)dx+\int\limits_{0}^{1}(x-x^2)dx=\left(\frac{x^2}{2}+\frac{x^3}{3}\right)\Bigg|_{-1}^{0}+\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Bigg|_{0}^{1}=0\,;$$

аналогічно

$$y_c=\int_{-\infty}^{\infty} y\,f_2(y)dy=\int_{-\infty}^{-1} y\cdot 0\cdot dy+\int_{-1}^{0} y(1+y)dy+\int_{0}^{1} y(1-y)dy+\int_{1}^{\infty} y\cdot 0\cdot dy=0\,.$$

Обчислюючи дисперсії D_1 і D_2 одержимо:

$$D_1=\int_{-1}^{0}(x-0)^2(1+x)dx+\int_{0}^{1}(x-0)^2(1-x)dx=$$

$$=\int_{-1}^{0}(x^2+x^3)dx+\int_{0}^{1}(x^2-x^3)dx=$$

$$=\left(\frac{x^3}{3}+\frac{x^4}{4}\right)\Bigg|_{-1}^{0}+\left(\frac{x^3}{3}-\frac{x^4}{4}\right)\Bigg|_{0}^{1}=-\left(-\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)=\frac{2}{3}-\frac{2}{4}=\frac{1}{6}\,.$$

Аналогічно одержимо $D_2=\dfrac{1}{6}$.

Вправи для самостійного виконання

3.15.1. Перевірити, чи правильні твердження:

1. За кожного абсолютно неперервного одновимірного розподілу ймовірностей існує центр розсіювання, координата якого скінченна.
2. Якщо за абсолютно неперервного одновимірного розподілу ймовірностей існує центр розсіювання, то координата цього центра є скінченним числом.
3. За кожного абсолютно неперервного одновимірного розподілу ймовірностей існує дисперсія.
4. Якщо за абсолютно неперервного розподілу ймовірностей існує дисперсія, то існує й відповідний центр розсіювання.
5. Твердження, обернене до 4, є правильним.
6. Дисперсія завжди є скінченним числом.
7. Якщо за абсолютно неперервного розподілу ймовірностей відповідний простір елементарних подій $\Omega=[a;b]$, де $-\infty<a<b<+\infty$, то за такого розподілу існують скінченні центр розсіювання і дисперсія.
8. Твердження, обернене до 7, є правильним.
9. Середнє квадратичне відхилення – це довільне дійсне число, квадрат якого дорівнює дисперсії.

3.15.2. Переформулювати твердження 1-9 із завдання 3.15.1 на випадок двохвимірного абсолютно неперервного розподілу ймовірностей та перевірити, чи правильні ці твердження.

3.15.3. Побудувати одновимірний абсолютно неперервний розподіл ймовірностей, за якого:

1) $x_c=0$ і $D=1$;

2) $x_c=0$, а $D\le 2$;

3) $x_c=0$, а $D>4$;

4) $x_c=0$, а $D=+\infty$;

5) x_c існує, а D не існує.

3.16. Приклади важливих розподілів ймовірностей

Розглянемо деякі розподіли ймовірностей, які найбільш часто використовуються на практиці.

1. *Рівномірний поточковий розподіл імовірностей на скінченній множині точок*. Одновимірний поточковий розподіл імовірностей на скінченній множині $\Omega = \{x_1, x_2, ..., x_n\}$, $S = \{A \mid A = \bigcup_{i \in I} \{x_i\}, I \subset \{1, 2, ..., n\}\}$, за умови, що всі події $\{x_i\} \in S$ відбуваються з однаковими ймовірностями $P(\{x_i\}), i = \overline{1, n}$, називається *рівномірним*.

Якщо дискретний поточковий розподіл ймовірностей на множині $\Omega = \{x_1, x_2, ..., x_n\}$ рівномірний, то всі $p_i = P(\{x_i\})$, $i = 1, 2, \ldots, n$, рівні між собою і дорівнюють $\frac{1}{n}$. В такому разі якщо в двох підмножинах A і B множини $\Omega = \{x_1, x_2, ..., x_n\}$ містяться однакові кількості елементів – m (міри $m(A) = k(A)$ і $m(B) = k(B)$ підмножин A і B однакові), то ймовірності попадання в такі підмножини однакові: $P(A) = P(B) = \frac{m}{n}$ (за властивістю аддитивності ймовірнісної міри).

Ряд рівномірного поточкового розподілу ймовірностей на множині $\Omega = \{x_1, x_2, ..., x_n\}$ в такому разі набуває вигляду (Табл. 3.16.1)

Табл. 3.16.1

x_i	x_1	x_2	$\ldots$	x_n
p_i	$\frac{1}{n}$	$\frac{1}{n}$	$\ldots$	$\frac{1}{n}$

Відповідний многокутник розподілу ймовірностей зображено на Рис. 3.16.1 а), а графік функції $F(x)$ рівномірного розподілу ймовірностей на множині $\Omega = \{x_1, x_2, ..., x_n\}$ зображено на Рис. 3.16.1 б).

2. *Біноміальний розподіл імовірностей*. Розподіл імовірностей на множині $\Omega = \{x_1, x_2, ..., x_n\}$, який можна дістати за формулою Бернуллі

$$P(\{x_i\}) = C_n^{x_i} p^{x_i} q^{n - x_i},$$

коли x_i набуває значень $0, 1, 2, ..., n$, називають *біноміальним*.

Назва цього розподілу ймовірностей походить від того, що ймовірності $P(\{x_i\})$ обчислюються так само, як і члени бінома Ньютона $(p+q)^n$:

$$(p+q)^n = \sum_{x_i = 0}^{n} C_n^{x_i} p^{x_i} q^{n - x_i}.$$

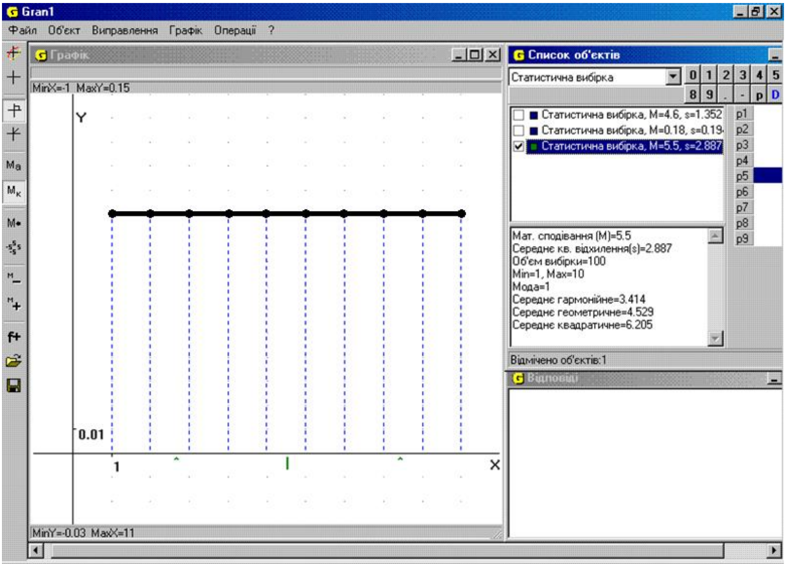


Рис. 3.16.1. а)

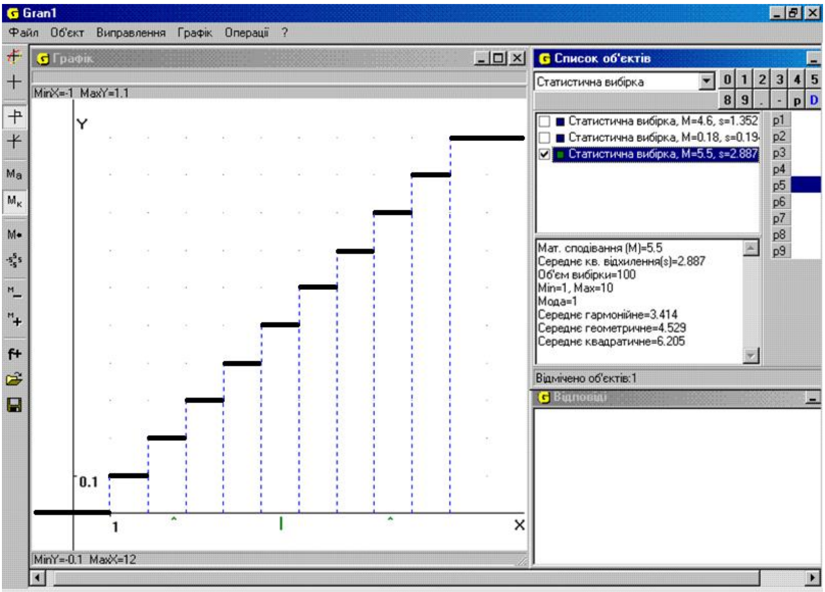


Рис. 3.16.1. б)

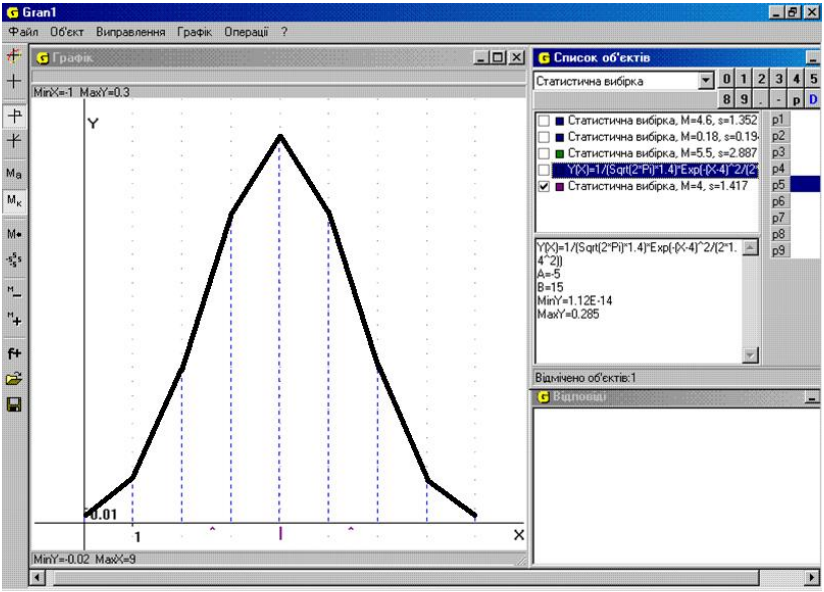


Рис. 3.16.2 а)

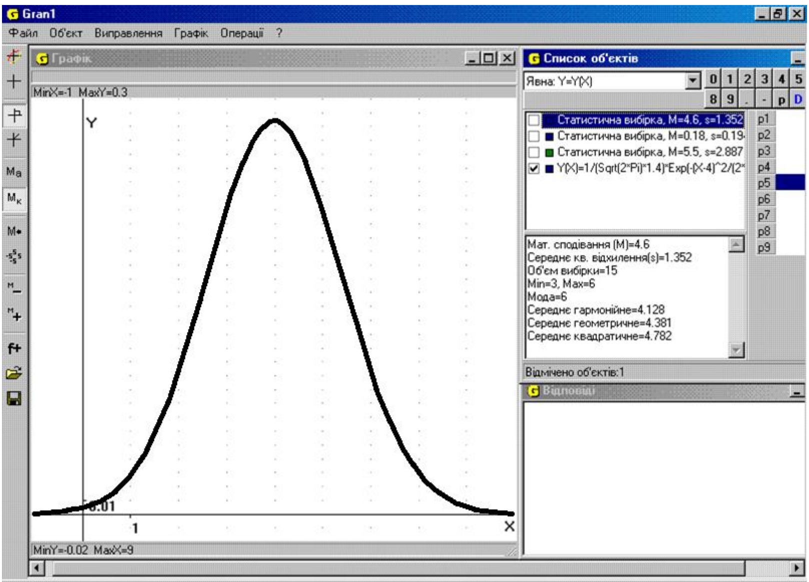


Рис. 3.16.2 б)

На Рис. 3.16.2 а) подано многокутник біноміального розподілу ймовірностей (за значень $p=\frac{1}{2}$, $q=\frac{1}{2}$, $n=8$).

Виявляється, що за досить великих n многокутник біноміального розподілу досить близький до графіка функції $\frac{1}{\sqrt{2\pi}\sqrt{npq}}e^{-\frac{(x-np)^2}{2npq}}$ (див. Рис. 3.16.2 б)). Цей факт обґрунтовується в локальній асимптотичній теоремі Муавра-Лапласа, на ньому базується й інтегральна асимптотична теорема Муавра-Лапласа (див. §2.10).

3. *Розподіл Пуассона*. Прикладом поточкового розподілу ймовірності на дискретній нескінченній множині є розподіл, який дістають за теоремою Пуассона (див. §2.10).

Надаючи x_i значень $0, 1, 2, \dots, n, \dots$, дістанемо нескінченний ряд розподілу ймовірностей (Табл. 3.16.2).

Табл. 3.16.2

x_i	0	1	2	$\dots$	n	$\dots$
p_i	$\frac{a^0}{0!}e^{-a}$	$\frac{a^1}{1!}e^{-a}$	$\frac{a^2}{2!}e^{-a}$	$\dots$	$\frac{a^n}{n!}e^{-a}$	$\dots$

Можна показати, що $\sum_{i=0}^{\infty} p_i = 1$.

Справді

$$\sum_{i=0}^{\infty} p_i = e^{-a}\left(1+\frac{a}{1!}+\frac{a^2}{2!}+\dots+\frac{a^n}{n!}+\dots\right) = e^{-a}\cdot e^{a} = 1.$$

Тут a – фіксоване додатне число. Різним значенням a відповідають різні розподіли ймовірностей. На Рис. 3.16.3 зображено многокутники розподілів імовірностей за значень $a=1$ і $a=0.5$ (Табл. 3.16.3).

Розподіл Пуассона широко застосовують на практиці, він тісно пов’язаний з біноміальним розподілом (див. §2.10).

Якщо у виразі

$$P(\{x_i\}) = C_n^{x_i} p^{x_i} q^{n-x_i}$$

для обчислення ймовірностей біноміального розподілу зафіксувати число x_i, спрямувати n до нескінченності $(n\to\infty)$, а p (або q) до нуля $(p\to 0)$ так, щоб добуток np дорівнював сталому числу a, то

$$\lim_{n\to\infty} P(\{x_i\}) = \lim_{n\to\infty} C_n^{x_i} p^{x_i} q^{n-x_i} = \frac{a^{x_i}}{x_i\,!}e^{-a}.$$

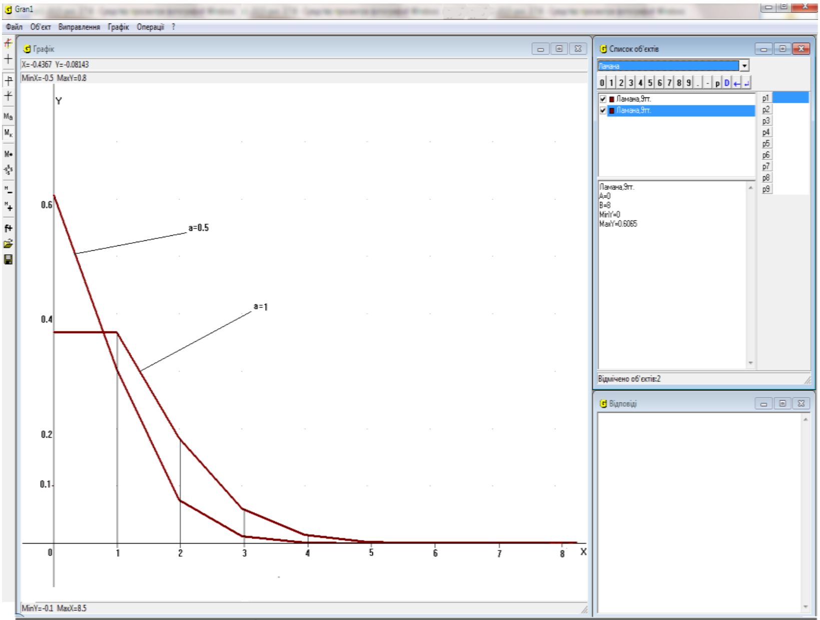


Рис. 3.16.3

Табл. 3.16.3

m	$\frac{a^m}{m!}e^{-a}$	$a=1$	$a=0.5$
0	$\frac{a^0}{0!}e^{-a}$	0.3678	0.6065
1	$\frac{a^1}{1!}e^{-a}$	0.3678	0.3032
2	$\frac{a^2}{2!}e^{-a}$	0.1839	0.0758
3	$\frac{a^3}{3!}e^{-a}$	0.0613	0.0126
4	$\frac{a^4}{4!}e^{-a}$	0.0153	0.0016
5	$\frac{a^5}{5!}e^{-a}$	0.0030	0.0002
6	$\frac{a^6}{6!}e^{-a}$	0.0005	0.00001

Таким чином, за необмеженого збільшення n за наведених умов біноміальний розподіл імовірностей асимптотично наближається до розподілу Пуассона (див. §2.10).

4. *Геометричний розподіл імовірностей*. Повторні незалежні випробування продовжуються до першого відбування події A, можливі наслідки серії випробувань – кількість проведених випробувань. Очевидно, що така кількість може набувати значень $1, 2, 3, \dots$, причому $P(\{x_i\}) = q^{x_i - 1} p$, де x_i – можливі наслідки серії випробувань, $\Omega = \{1, 2, 3, \dots, n, ...\}$. Розглядуваний розподіл імовірностей називається *геометричним*. Ряд цього розділу ймовірностей подається у вигляді (Табл. 3.16.4):

Табл. 3.16.4

x_i	1	2	3	$\dots$	n	$\dots$
p_i	p	qp	$q^2 p$	$\dots$	$q^{n-1} p$	$\dots$

Легко бачити, що $\sum\limits_{i=0}^{\infty} pq^i = \dfrac{p}{1-q} = 1$ як сума членів нескінченної спадної геометричної прогресії з першим членом p і знаменником прогресії q.

5. *Ряд гіпергеометричного розподілу ймовірностей* подається у вигляді (Табл. 3.16.5):

Табл. 3.16.5

x_i	0	1	2	$\dots$	k
p_i	$\dfrac{C_M^0 C_{N-M}^n}{C_N^n}$	$\dfrac{C_M^1 C_{N-M}^{n-1}}{C_N^n}$	$\dfrac{C_M^2 C_{N-M}^{n-2}}{C_N^n}$	$\dots$	$\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$

де $k = \min(n, M)$. Це є ряд розподілу ймовірностей кількості виграшних квитків серед n квитків, які навмання вибираються з N квитків, серед яких є M виграшних, а решта $N - M$ невиграшних.

6. Рівномірний абсолютно неперервний розподіл ймовірностей на обмеженому відрізку $[a; b]$. На практиці часто використовують так званий рівномірний неперервний розподіл ймовірностей на відрізку $[a; b]$, де $-\infty < a < b < \infty$.

Опис щільності рівномірного розподілу ймовірностей на відрізку $[a; b]$, $-\infty < a < b < +\infty$, подається у вигляді:

$$f(x) = \begin{cases} c, \textit{ коли } x \in [a; b], c > 0; \\ 0, \textit{ коли } x \overline{\in} [a; b]. \end{cases}$$

Графік функції $f(x)$ подано на Рис. 3.16.4.

Виходячи з вимоги $\int\limits_{-\infty}^{\infty} f(x)\ dx = 1$, знаходимо сталу c:

$$\int\limits_{-\infty}^{\infty} f(x)\ dx = \int\limits_{-\infty}^{a} 0\cdot dx + \int\limits_{a}^{b} c\, dx + \int\limits_{b}^{\infty} 0\cdot dx = c\int\limits_{a}^{b} dx = c(b-a) = 1.$$

Звідси

$$c = \frac{1}{b-a}.$$

Нагадаємо, що в геометричному тлумаченні $\int\limits_{a}^{b} f(x)\ dx$ – це площа криволінійної трапеції, обмеженої зверху кривою $y = f(x)$ ($f(x) \geq 0$ за будь-якого x), знизу – прямою $y = 0$, ліворуч – прямою $x = a$, праворуч – прямою $x = b$. Таким чином, площа, що лежить над віссю Ox нижче графіка функції $y = f(x)$, повинна дорівнювати 1.

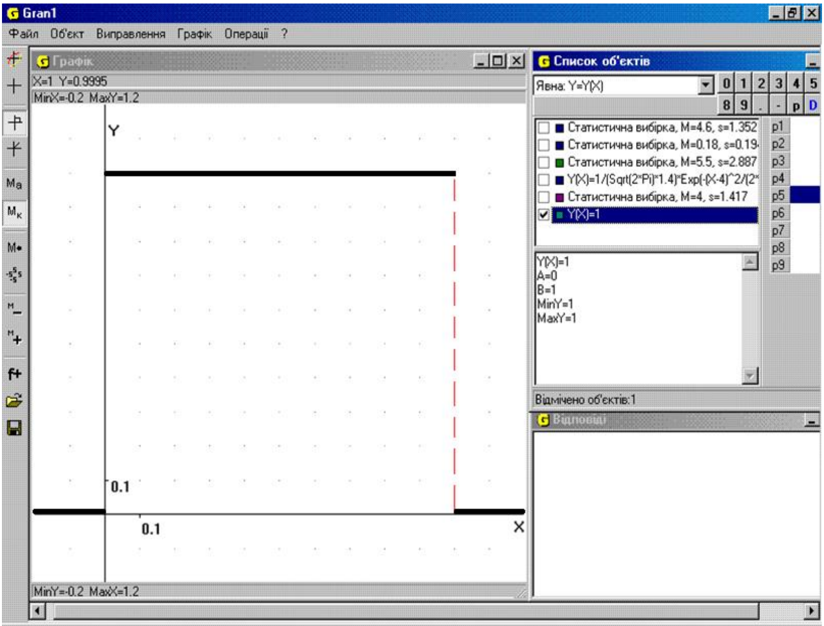


Рис. 3.16.4

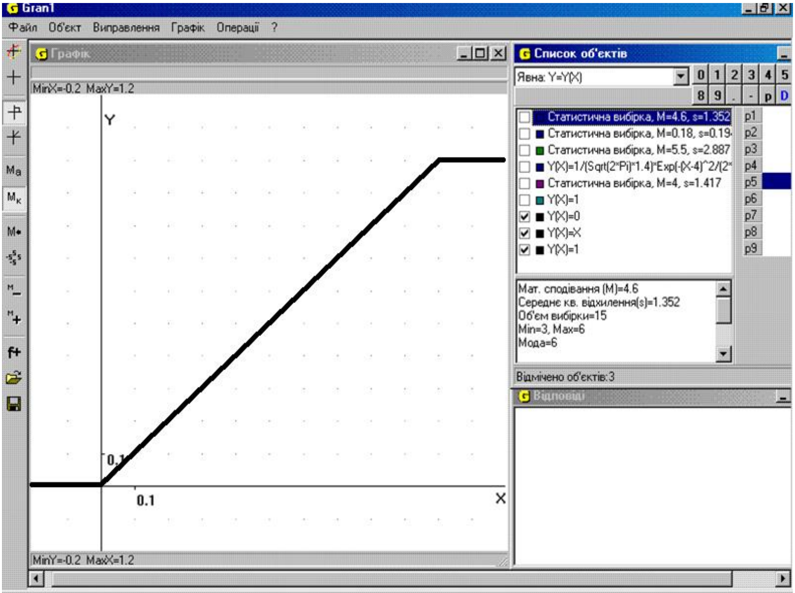


Рис. 3.16.5

Тому сталу c можна було визначити, виходячи з вимоги, що площа прямокутника, обмеженого прямими $x=a$, $x=b$, $y=0$, $y=c$, повинна дорівнювати 1, тобто $c(b-a)=1$. Опис функції $F(x)$ розподілу ймовірностей на $(R^1, \mathscr{B}(R^1))$ у випадку рівномірного абсолютно неперервного розподілу ймовірностей на відрізку $[a;b]$ набуває вигляду

$$F(x)=\begin{cases} 0, & \text{коли} \quad x\le a; \\ \int\limits_a^x f(x)\,dx=\dfrac{x-a}{b-a}, & \text{коли} \quad a\le x\le b; \\ 1, & \text{коли} \quad b\le x. \end{cases}$$

Графік функції $F(x)$ подано на Рис. 3.16.5.

Обчислимо тепер ймовірність попадання на відрізок $[\alpha;\beta]\subset(-\infty;+\infty)$ за умови рівномірного розподілу ймовірностей на відрізку $[a;b]$ (вважаючи, що $\Omega=R^1$, $S=\mathscr{B}(R^1)$).

Враховуючи, що $P([\alpha;\beta])=\int\limits_\alpha^\beta f(x)dx=F(\beta)-F(\alpha)$, дістаємо

$$P([\alpha\,;\,\beta]) = \begin{cases} \dfrac{\beta-\alpha}{b-a}, & \text{коли} \quad a \le \alpha \le \beta \le b; \\ \dfrac{\beta-a}{b-a}, & \text{коли} \quad \alpha \le a \le \beta \le b; \\ \dfrac{b-\alpha}{b-a}, & \text{коли} \quad a \le \alpha \le b \le \beta; \\ 1, & \text{коли} \quad \alpha \le a \le b \le \beta; \\ 0, & \text{коли} \quad \alpha \le \beta \le a \quad \textit{або} \quad b \le \alpha \le \beta. \end{cases}$$

Пригадуючи геометричне задання ймовірності, помічаємо, що там мався на увазі саме рівномірний розподіл імовірностей на множині G, проте розглядалися тільки випадки типу $a \le \alpha \le \beta \le b$ (див. §3.6).

Вправи для самостійного виконання

3.16.1. Перевірити, чи правильні твердження:

1. Якщо дискретний розподіл ймовірностей рівномірний, то простір Ω елементарних подій скінченний.

2. Твердження, обернене до 1, є правильним.

3. Якщо в просторі Ω елементарних подій зчисленна кількість елементів, то відповідний розподіл ймовірностей не може бути рівномірним.

4. Якщо многокутник дискретного розподілу ймовірностей є відрізком, то цей розподіл рівномірний.

5. Твердження, обернене до 4, є правильним.

6. Біноміальний розподіл ймовірностей завжди пов'язаний із скінченним простором елементарних подій.

7. Розподіл Пуассона завжди пов'язаний із зчисленним простором елементарних подій.

8. В розподілі Пуассона: $P(\{m\}) = \dfrac{a^m}{m!}e^{-a}$, $m \in \{0, 1, 2, \ldots, n, \ldots\}$, ймовірності $P(\{m\})$ як завгодно малі, коли числа m досить великі.

9. В разі розподілу Пуассона послідовність ймовірностей $P(\{m\}) = \dfrac{a^m}{m!}e^{-a}$, $m = 0,\ 1,\ 2,\ \ldots$, є спадаючою за будь-якого $a > 0$.

10. Коли n великі, тоді біноміальний розподіл ймовірностей близький до розподілу Пуассона

11. Геометричний розподіл ймовірностей утворюється із членів геометричної прогресії з першим членом p і знаменником прогресії q.

12. Геометрична прогресія із твердження 11 завжди є збіжною, а її сума дорівнює 1.

13. Рівномірний неперервний розподіл ймовірностей є абсолютно неперервним.

14. Рівномірний неперервний розподіл ймовірностей може бути сингулярним.

15. Щільність рівномірного розподілу ймовірностей є сталою функцією на $(-\infty;+\infty)$.

16. Функція рівномірного неперервного розподілу ймовірностей є лінійною на певному проміжку $\langle a\,;b\rangle$, а на будь-якому іншому проміжку $\langle c\,;d\rangle$, що не перетинається з $\langle a\,;b\rangle$, ця функція тотожно дорівнює 0 або 1.

3.16.2. Стосовно розподілів ймовірностей, вказаних в 3.16.1.1, 3.16.1.8, 3.16.1.11, 3.16.1.16, знайти центр розсіювання і дисперсію.

3.17. Одновимірний нормальний розподіл ймовірностей

Одним із найбільш часто використовуваних на практиці, окрім рівномірного розподілу ймовірностей, є так званий *нормальний розподіл імовірностей*.

Нормальним розподілом імовірностей на множині $\Omega = R^1$, $S = \mathscr{B}(R^1)$, називають абсолютно неперервний розподіл, щільність якого подається у вигляді

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}. \qquad (3.17.1)$$

Графік щільності нормального розподілу ймовірностей подано на Рис. 3.17.1. Цю криву називають *кривою Гаусса*.

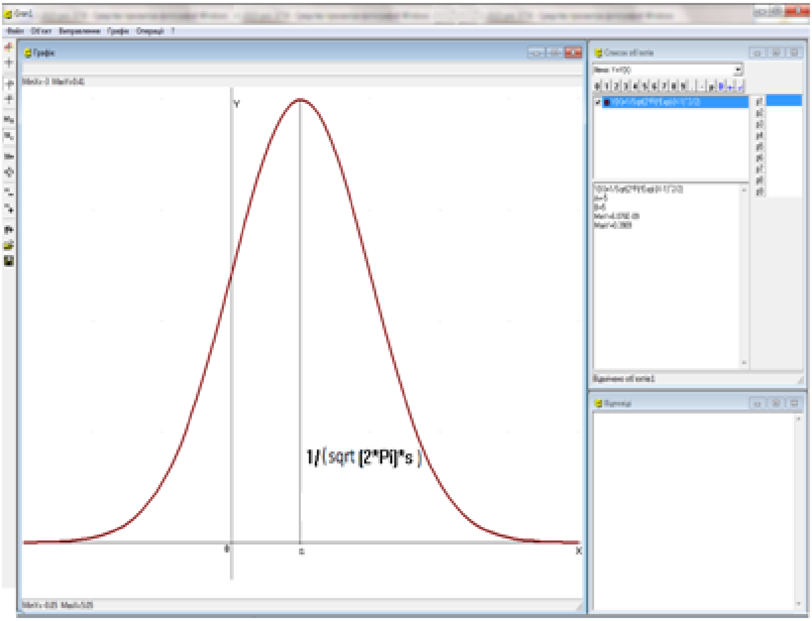


Рис. 3.17.1

Згідно з (3.17.1) графік щільності нормального розподілу ймовірностей симетричний відносно прямої $x = a$. Із зміною параметра a форма графіка не змінюється, графік зміщується так, що абсциса найвищої точки на графіку дорівнює a (Рис. 3.17.2). Виявляється, що параметр a є абсцисою центра нормального розподілу ймовірностей із щільністю розподілу ймовірностей (3.17.1).

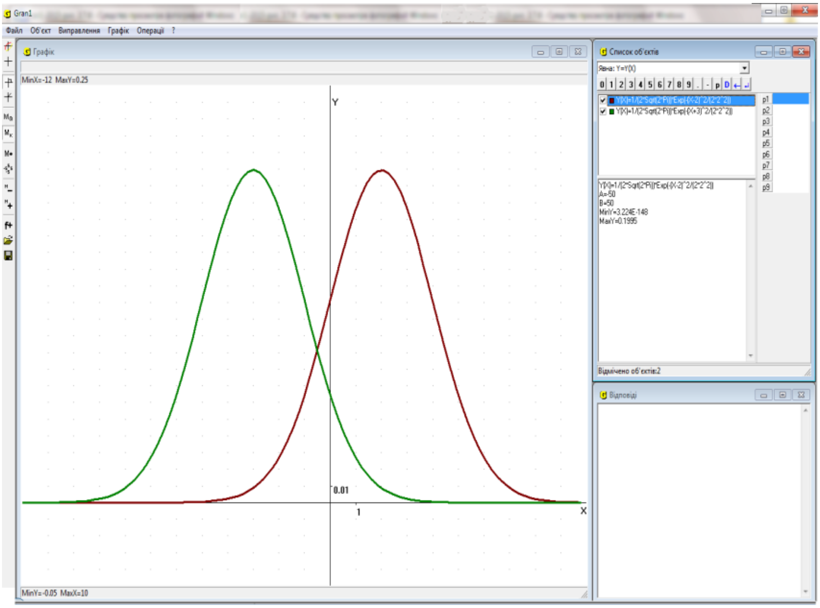


Рис. 3.17.2

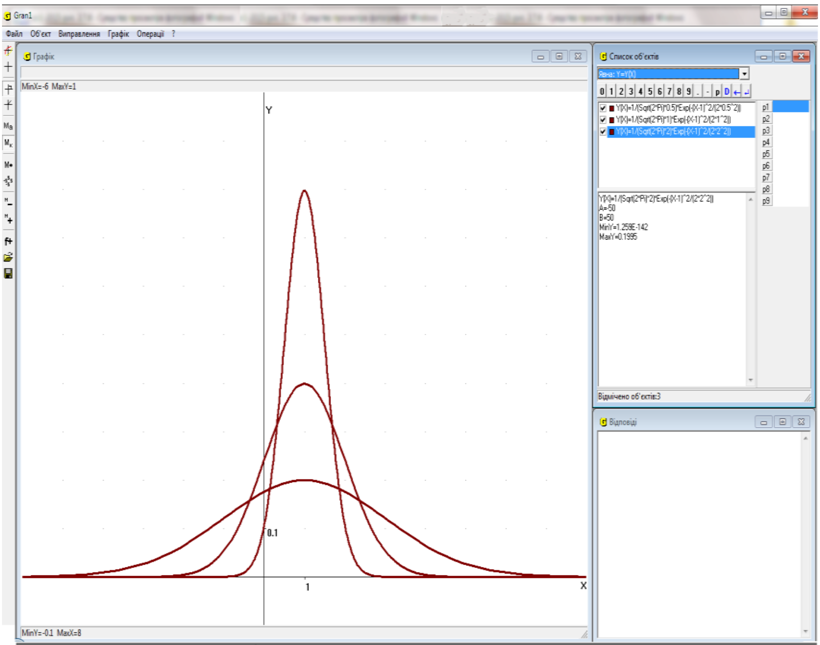


Рис. 3.17.3

За допомогою параметра σ характеризують розсіювання ймовірностей навколо центра розподілу ймовірностей. Справді, чим менше σ, тим вище піднімається вершина кривої $y = f(x)$ (Рис. 3.17.3) і, навпаки, чим більше σ, тим нижче вершина кривої $y = f(x)$. Оскільки площа під графіком щільності завжди повинна дорівнювати 1, то коли σ мале, над одним і тим самим відрізком $[x; x+\Delta x]$, досить віддаленим від центра розсіювання, площа під графіком щільності буде меншою, ніж площа під графіком щільності, коли σ велике. Отже, коли σ мале, вся маса ймовірностей зосереджується ближче до центра розподілу ймовірностей, ніж коли σ велике, і ймовірність попадання на проміжок $[\alpha; \beta]$ скінченної довжини, значно віддалений від центра розсіювання, досить мала, коли σ мале.

Покажемо, що стосовно функції $f(x)$, заданої за рівністю (3.17.1), задовільняються властивості щільності розподілу ймовірностей. Справді, оскільки $\sigma > 0$, то $0 \le f(x) \le \frac{1}{\sqrt{2\pi}\sigma}$, і, таким чином, перша властивість щільності задовільняється.

Покажемо, що стосовно функції $f(x)$ задовільняється й друга властивість щільності розподілу ймовірностей:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Для цього обчислимо

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}}\,dx,$$

або після заміни змінних $\frac{x-a}{\sqrt{2}\sigma} = t$

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-t^2}\sqrt{2}\sigma\,dt = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}\,dt = \frac{2}{\sqrt{\pi}}\int_{0}^{\infty} e^{-t^2}\,dt. \quad (3.17.2)$$

Стосовно функції e^{-t^2} первісної у скінченних виразах не існує. Тому для обчислення останнього інтеграла скористаємось таким прийомом. Позначимо

$$I = \int_{0}^{\infty} e^{-t^2}\,dt.$$

Тоді

$$I^2 = \int_{0}^{\infty} e^{-t^2}\,dt \int_{0}^{\infty} e^{-u^2}\,du = \int_{0}^{\infty}\left\{\int_{0}^{\infty} e^{-(t^2+u^2)}\,dt\right\}du.$$

Останній повторний інтеграл є подвійним інтегралом

$$\iint\limits_{R^+} e^{-(t^2+u^2)}dtdu\,, \qquad (3.17.3)$$

де $R^+ = \{(t,u) \mid t \ge 0, u \ge 0\}$ (Рис. 3.17.4).

Переходячи в (3.17.3) до полярних координат $(t = \rho\cos\varphi,\ u = \rho\sin\varphi,\ dtdu = \rho d\rho d\varphi)$, одержимо

$$I^2 = \iint\limits_{R^+} e^{-\rho^2}\rho d\rho d\varphi\,, \qquad (3.17.4)$$

де $R^+ = \{(\varphi,\ \rho) \mid 0 \le \varphi \le \dfrac{\pi}{2},\ 0 \le \rho < +\infty\}$(Рис. 3.17.4).

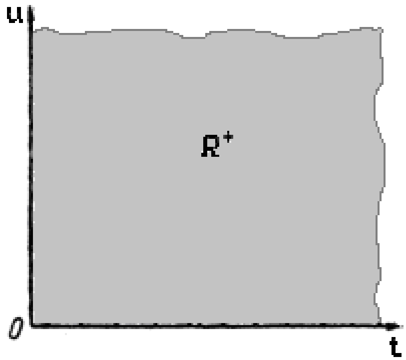


Рис. 3.17.4

Переходячи в (3.17.4) від подвійного інтеграла до повторного, знаходимо

$$I^2 = \int_0^{\frac{\pi}{2}}\left\{\int_0^{\infty} e^{-\rho^2}\rho d\rho\right\}d\varphi = \int_0^{\frac{\pi}{2}}d\varphi\int_0^{\infty} e^{-\rho^2}\rho d\rho = \frac{\pi}{2}\left(-\frac{e^{-\rho^2}}{2}\right)\Bigg|_0^{\infty} = \frac{\pi}{4}\,.$$

Таким чином,

$$I = \int_0^{\infty} e^{-t^2}dt = \frac{\sqrt{\pi}}{2}\,. \qquad (3.17.5)$$

Повертаючись до (3.17.2) і підставляючи замість $\int\limits_0^{\infty} e^{-t^2}dt$ його значення, дістаємо, що стосовно функції

$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}$ задовільняється і друга властивість щільності розподілу ймовірностей.

Інтеграл (3.17.5) називається *інтегралом Ейлера–Пуассона*. Він відіграє важливу роль в теорії ймовірностей.

Доведемо, що параметри a і σ є відповідно абсцисою центра розподілу і середнім квадратичним відхиленням $\sigma=\sqrt{D}$ нормального розподілу імовірностей із щільністю (3.17.1).

Справді, обчислюючи координату (абсцису) центра нормального розподілу ймовірностей із щільністю (3.17.1), одержуємо

$$x_c=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}xe^{-\frac{(x-a)^2}{2\sigma^2}}dx\ ,$$

або після заміни змінних $\dfrac{x-a}{\sqrt{2}\sigma}=t$ з врахуванням (3.17.5)

$$x_c=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(\sqrt{2}\sigma t+a)e^{-t^2}\sqrt{2}\sigma dt=\frac{\sqrt{2}\sigma}{\sqrt{\pi}}\int_{-\infty}^{\infty}te^{-t^2}dt+\frac{a}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-t^2}dt=$$

$$=\frac{\sqrt{2}\sigma}{\sqrt{\pi}}\cdot 0+\frac{a}{\sqrt{\pi}}\cdot\sqrt{\pi}=a\ .$$

Обчислюючи дисперсію D нормального розподілу ймовірностей із щільністю (3.17.1), знаходимо

$$D=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx\ ,$$

або після заміни змінних $\dfrac{x-a}{\sqrt{2}\sigma}=t$

$$D=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}2\sigma^2t^2e^{-t^2}\sigma\sqrt{2}dt=\frac{1}{\sqrt{\pi}}4\sigma^2\int_{0}^{\infty}t^2e^{-t^2}dt\ .$$

Інтегруючи частинами

$$(u=t,\ dv=te^{-t^2}dt,\ du=dt,\ v=-\frac{e^{-t^2}}{2})\ ,$$

дістаємо

$$D=\frac{1}{\sqrt{\pi}}4\sigma^2\left(-\frac{te^{-t^2}}{2}\Bigg|_0^{\infty}+\frac{1}{2}\int_0^{\infty}e^{-t^2}dt\right)=\frac{2}{\sqrt{\pi}}\sigma^2\int_0^{\infty}e^{-t^2}dt=\sigma^2\ .$$

Тут $\lim\limits_{t\to\infty}\left(-\dfrac{te^{-t^2}}{2}\right)=\lim\limits_{t\to\infty}\dfrac{-t}{2e^{t^2}}=0$ (за правилом Лопіталя), а за формулою (3.17.5) $\int\limits_0^\infty e^{-t^2}dt=\dfrac{\sqrt{\pi}}{2}$.

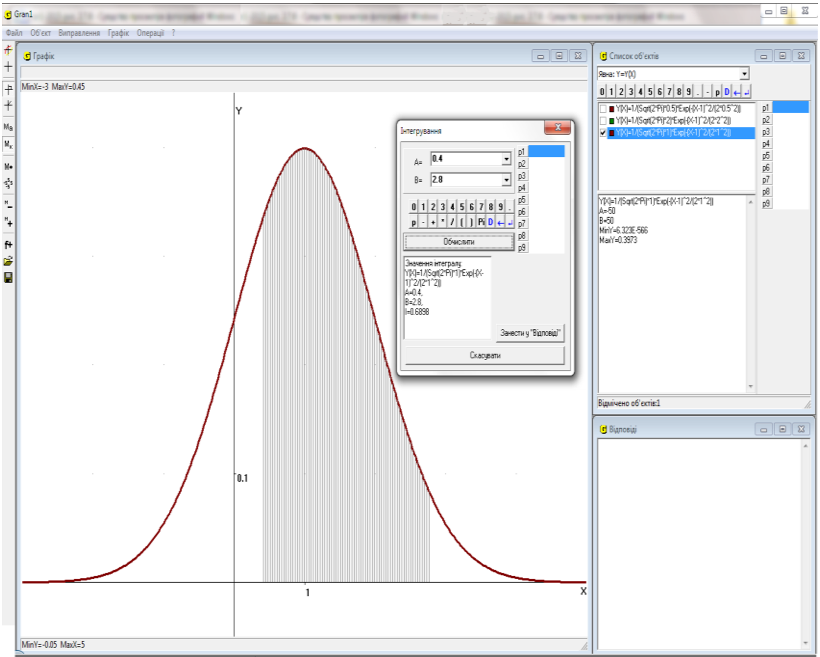


Рис. 3.17.5

Таким чином, якщо відомо, що розподіл ймовірностей нормальний, то щоб повністю його описати (за допомогою щільності розподілу ймовірностей $f(x)$), досить знати абсцису a центра розподілу ймовірностей і середнє квадратичне відхилення $\sigma=\sqrt{D}$.

Якщо $a=0$, $\sigma=1$, то говорять, що *нормальний розподіл ймовірностей центрований і нормований (або стандартний)*.

За нормального розподілу ймовірностей ймовірність попадання на відрізок $[\alpha;\beta]$ обчислюється за формулою

$$P([\alpha,\ \beta])=\frac{1}{\sqrt{2\pi}\sigma}\int\limits_\alpha^\beta e^{-\frac{(x-a)^2}{2\sigma^2}}dx. \qquad (3.17.6)$$

Хоч стосовно функції $e^{-\frac{(x-a)^2}{2\sigma^2}}$ первісної в скінчених виразах і не існує, проте інтеграл виду (3.17.6) легко обчислюється за допомогою програмних засобів типу GRAN1 (Рис. 3.17.5) чи інших математичних програм.

Виявляється, що коли розподіл ймовірностей нормальний на R^1, $S=\mathscr{B}(R^1)$, тоді ймовірність попадання на відрізок $[a-3\sigma; a+3\sigma]$ практично дорівнює 1 (Рис. 3.17.6). Тому на практиці можна вважати, що ймовірність попадання за межі відрізка $[a-3\sigma; a+3\sigma]$ дорівнює нулю (*правило трьох сигм*).

Виконавши в (3.17.6) заміну змінних $\frac{x-a}{\sigma}=t$, ймовірність $P[(\alpha;\beta])$ можна подати в дещо іншому вигляді

$$P[(\alpha;\beta])=\int\limits_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\left(\int\limits_{\frac{\alpha-a}{\sigma}}^{0}e^{-\frac{t^2}{2}}dt+\int\limits_{0}^{\frac{\beta-a}{\sigma}}e^{-\frac{t^2}{2}}dt\right)=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\frac{\beta-a}{\sigma}}e^{-\frac{t^2}{2}}dt-\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{\frac{\alpha-a}{\sigma}}e^{-\frac{t^2}{2}}dt\,. \qquad (3.17.7)$$

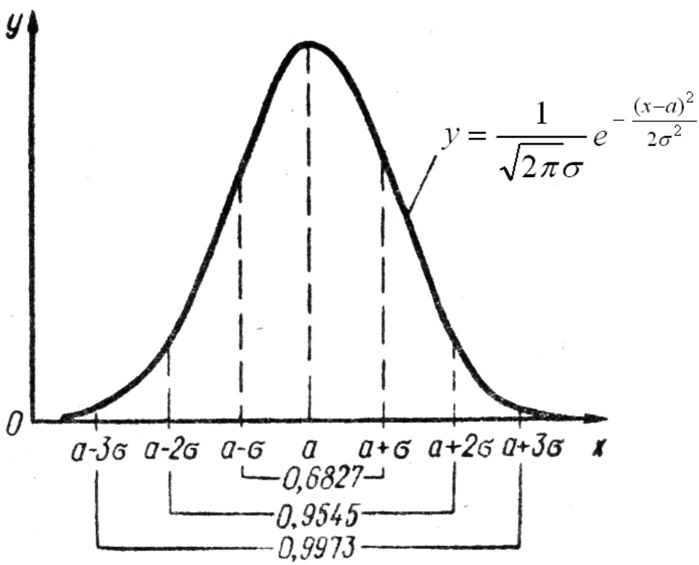


Рис. 3.17.6

Як і раніше, $\int e^{-\frac{t^2}{2}}dt$ у точних виразах знайти не вдається. Якщо немає програмних засобів типу GRAN1 чи інших, за допомогою яких можна обчислити інтеграли в (3.17.7), то для наближеного обчислення значень функції

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt\ ,$$

яка називається *функцією Лапласа*, використовують спеціальні таблиці (див. Додаток 2).

Використовуючи функцію Лапласа, рівність (3.17.7) можна подати у вигляді

$$P([\alpha,\ \beta]) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \qquad (3.17.8)$$

Легко бачити, що функція Лапласа непарна, тобто $\Phi(-x) = -\Phi(x)$.

Справді, $\Phi(-x) = \frac{1}{\sqrt{2\pi}}\int_0^{-x} e^{-\frac{t^2}{2}}dt$, і після заміни $t=-u$ дістанемо

$$\Phi(-x) = -\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{u^2}{2}}du = -\Phi(x)\,.$$

Тому, якщо відрізок $[\,\alpha\,;\beta\,]$ симетричний відносно центра розподілу ймовірностей, тобто $\alpha = a-\delta$, $\beta = a+\delta$, то

$$P\,(\,a-\delta\,;\,a-\delta\,) = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(\frac{-\delta}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

Використовуючи програму GRAN1 для обчислення значень функції Лапласа, стосовно нормального розподілу імовірностей (3.17.1) дістаємо

$$P\,(\,a-\sigma\,;\,a+\sigma\,) = 0.6827\,;$$
$$P\,(\,a-2\sigma\,;\,a+2\sigma\,) = 0.9545\,;$$
$$P\,(\,a-3\sigma\,;\,a+3\sigma\,) = 0.9973\,.$$

Зазначимо, що використовуючи програми типу GRAN1, доцільно $P([\,\alpha\,;\beta\,])$ обчислювати не за формулою (3.17.7), а безпосередньо за формулою $\int\limits_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{t^2}{2}}dt$, оскільки в останньому випадку звертатися до послуги обчислення визначеного інтеграла доведеться тільки один раз замість двох і відпадає необхідність виконувати операцію віднімання у формулі (3.17.7). Крім того, заощаджується час на зайве звернення до послуг програми для обчислення за формулою (3.17.7), а також час, який доведеться витрачати на запам'ятовування значень $\Phi\left(\frac{\beta - a}{\sigma}\right)$ і $\Phi\left(\frac{\alpha - a}{\sigma}\right)$ та виконання операції їх віднімання.

Нормальний розподіл ймовірностей широко використовується в математичній статистиці, теорії стріляння, для побудови таблиць стріляння, в теорії похибок вимірювань і наближених обчислень тощо.

Цей розподіл фактично вже зустрічався в локальній та інтегральній теоремах Муавра-Лапласа. Виявляється, що коли n досить великі, розподіл імовірностей $P_n(\mu = m)$ близький до нормального, причому абсциса центра розподілу ймовірностей різних чисел (від 0 до n) відбувань події A в серії з n незалежних випробувань за схемою Бернуллі дорівнює np, а середнє квадратичне відхилення дорівнює $\sqrt{npq}$. Саме це стверджується в локальній асимптотичній теоремі Муавра-Лапласа. Многокутник розподілу ймовірностей $P_n(\mu = m)$ (коли $p = 1/2$, $q = 1/2$), вже коли $n = 8$, дуже нагадує графік щільності нормального розподілу імовірностей (Рис. 3.17.7).

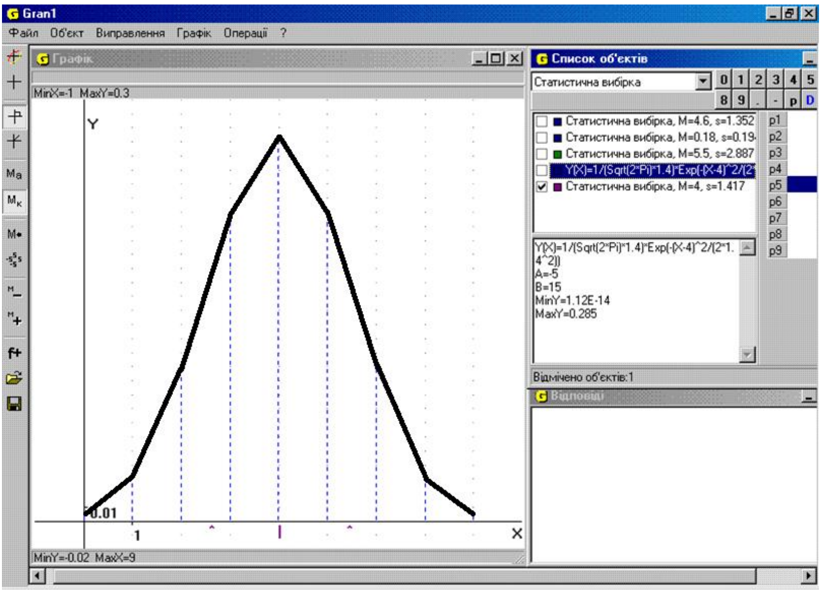


Рис. 3.17.7

В інтегральній теоремі Муавра-Лапласа також обчислюється ймовірність попадання на відрізок $[m_1; m_2]$ за умови, що розглядається нормальний розподіл імовірностей з абсцисою центра розподілу ймовірностей np і середнім квадратичним відхиленням $\sqrt{npq}$.

Приклад 3.17.1. На заняттях з фізкультури учні виконують стрибки в довжину. Ймовірності різних довжин стрибків учнів розподілені за нормальним розподілом імовірностей, в якого центр $a = 3\,м$, середнє квадратичне відхилення $\sigma = 1\,м$. Знайти ймовірність того, що довжина стрибка навмання взятого учня становитиме від $2\,м$ до $4\,м$.

За формулою (3.17.8) знаходимо

$$P([2;4]) = \Phi\left(\frac{4-3}{1}\right) - \Phi\left(\frac{2-3}{1}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1).$$

Скориставшись програмою GRAN1, одержимо

$$\Phi(1) = \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-\frac{t^2}{2}} dt = 0.3413,$$

Отже, шукана ймовірність дорівнює

$$P([2;4]) = 2 \cdot 0.3413 = 0.6826.$$

Вправи для самостійного виконання

3.17.1. Перевірити, чи правильні твердження:

1. Нормальний розподіл ймовірностей є абсолютно неперервним розподілом.

2. Якщо щільність розподілу ймовірностей $f(x) = \dfrac{e^{2x}}{\pi e} \cdot e^{-x^2}$, то цей розподіл є нормальним.

3. Якщо нормальний розподіл ймовірностей задано через щільність розподілу, то відразу можна записати, чому дорівнює координата центра розсіювання і дисперсія.

4. Щільність нормального розподілу ймовірностей:

1) є монотонною функцією;
2) змінюється монотонно на двох проміжках;
3) існує єдина точка, в якій щільність нормального розподілу ймовірностей набуває екстремального значення.

5. Якщо $f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$, то:

1) $f(x)$ зростає на проміжку $(-\infty; a)$;
2) $f(x)$ спадає на проміжку $(a; +\infty)$;

3) $f(a) = \max\limits_{(-\infty;+\infty)} f(x) = \dfrac{1}{\sqrt{2\pi}\sigma}$.

6. Інтеграл $\int\limits_0^x e^{-t^2}\, dt$ можна обчислити в скінченних виразах.

7. Кожен нормальний розподіл ймовірностей є центрованим і нормованим. Якщо $f(x)$ – щільність центрованого нормального розподілу ймовірностей, то $f(x)$ парна функція.

8. Якщо нормальний розподіл ймовірностей центрований, то відповідна функція розподілу ймовірностей є:
1) парною;
2) непарною;
3) ні парною, ні непарною.

9. Функція Лапласа є парною функцією.

10. Якщо розподіл ймовірностей нормальний з центром розсіювання a і дисперсією σ^2, то:
1) величина $P([a-\sigma;\ a+\sigma])$ залежить від a і σ;
2) величина $P([a-\sigma;\ a+\sigma])$ може не залежати від a і σ;
3) величина $P([a-3\sigma;\ a+3\sigma])$ залежить від a і σ;
4) величина $P([a-3\sigma;\ a+3\sigma])$ не залежить від a і σ;

11. За досить великих n біноміальний розподіл ймовірностей близький до нормального з параметрами $a = np$ і $\sigma = \sqrt{npq}$.

12. Використовуючи функцію Лапласа, можна обчислити ймовірність $P([\alpha;\ \beta])$ за будь якого нормального розподілу ймовірностей і будь якого відрізка $[\alpha;\ \beta]$.

3.18. Двохвимірний нормальний розподіл ймовірностей

Щільність двохвимірного нормального розподілу імовірностей на R^2, $S=\mathscr{B}(R^2)$, в загальному випадку подається у вигляді (Рис. 3.18.1)

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left(\frac{(x-m_1)^2}{\sigma_1^2}-2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right)}.$$

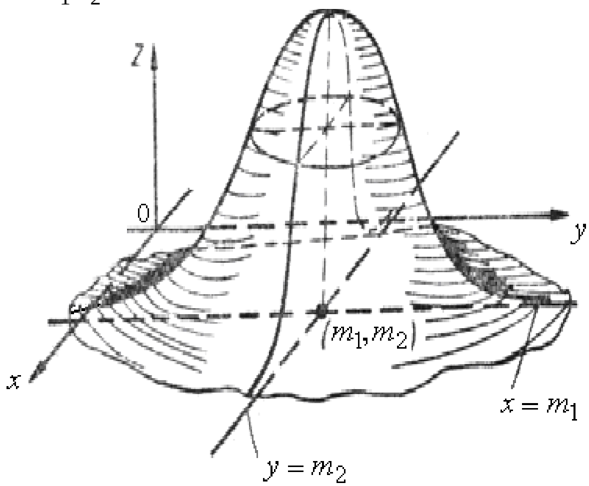


Рис. 3.18.1

Для з'ясування змісту параметрів знайдемо $f_1(x)$, поклавши $\frac{x-m_1}{\sigma_1}=u$, $\frac{y-m_2}{\sigma_2}=v$. Тоді

$$f_1(x)=\int\limits_{-\infty}^{\infty}f(x,y)dy=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int\limits_{-\infty}^{\infty}e^{-\frac{1}{2(1-r^2)}(u^2-2ruv+v^2)}\sigma_2 dv=$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-r^2}}\int\limits_{-\infty}^{\infty}e^{-\frac{1}{2(1-r^2)}(u^2-u^2r^2+u^2r^2-2ruv+v^2)}dv=$$

$$=\frac{1}{2\pi\sqrt{1-r^2}\,\sigma_1}e^{-\frac{u^2}{2}}\int\limits_{-\infty}^{\infty}e^{-\frac{1}{2(1-r^2)}(v-ru)^2}dv=$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-r^2}}e^{-\frac{u^2}{2}}\int\limits_{-\infty}^{\infty}e^{-\frac{(v-ru)^2}{2(1-r^2)}}dv.$$

Нехай $\frac{v - ru}{\sqrt{2}\sqrt{1-r^2}} = t$. Тоді $v = t\sqrt{2}\sqrt{1-r^2} + ru$, $dv = \sqrt{2}\sqrt{1-r^2}\,dt$, і таким чином

$$f_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{u^2}{2}} \int_{-\infty}^{\infty} e^{-t^2} dt = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{u^2}{2}} \sqrt{\pi} = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}.$$

Аналогічно $f_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-m_2)^2}{2\sigma_2^2}}$.

Отже, m_1 – координата (абсциса) центра нормального розподілу ймовірностей із щільністю розподілу ймовірностей $f_1(x)$, σ_1 – середнє квадратичне відхилення стосовно цього розподілу. Аналогічно, m_2 – координата (ордината) центра нормального розподілу ймовірностей із щільністю розподілу ймовірностей $f_2(y)$, σ_2 – середнє квадратичне відхилення стосовно цього розподілу.

Таким чином, центр двохвимірного нормального розподілу ймовірностей знаходиться в точці (m_1, m_2), σ_1^2 – дисперсія розсіювання ймовірностей навколо центра вздовж осі Ox, σ_2^2 – вздовж осі Oy.

Через параметр r характеризується залежність розподілів із щільностями $f_1(x)$ та $f_2(y)$. Якщо $r = 0$, тоді $f(x, y) = f_1(x) f_2(y)$, і таким чином розподіли виявляються незалежними. Якщо $r \neq 0$, тоді $f(x, y) \neq f_1(x) f_2(y)$, а тому розподіли залежні.

Вправи для самостійного виконання

3.18.1. Перевірити, чи правильні твердження:

1. Двохвимірний нормальний розподіл ймовірностей не є абсолютно неперервним.
2. Якщо відомо, що одновимірний розподіл ймовірностей є нормальним, то щоб задати цей розподіл, достатньо задати центр розсіювання і дисперсію цього розподілу.
3. Твердження 2 є правильним, коли слово “одновимірний” замінити словом “двохвимірний”.
4. Двохвимірний нормальний розподіл ймовірностей цілком визначається за двома одновимірними нормальними розподілами.
5. Параметри r, m_1, σ_1, m_2, σ_2 двохвимірного нормального розподілу ймовірностей – це будь-які фіксовані дійсні числа.

РОЗДІЛ IV. ВИПАДКОВІ ВЕЛИЧИНИ

4.1. Поняття випадкової величини

Нехай задано дві множини *V* і *W* елементів деякої природи. Відповідністю *X* між множинами *V* і *W* називають довільну фіксовану підмножину декартового добутку множин *V* і *W*, тобто $X \subset V \times W$. В разі коли $(v, w) \in X$, тоді кажуть, що елемент $w \in W$ поставлено у відповідність елементові $v \in V$ за відповідністю *X*. Множину *Q* тих елементів $w \in W$, що поставлені у відповідність елементові $v \in V$ за відповідністю *X*, називають *образом* елемента $v \in V$ за відповідністю *X* і пишуть $Q = X(v)$, $v \in V$, $Q = X(v) \subset W$. В цьому образі $Q = X(v) \subset W$ може міститися довільна кількість елементів із *W*, зокрема він може бути порожнім. *Образом множини* $G \subset V$ за відповідністю *X* називають об'єднання образів всіх елементів множини *G*, тобто $X(G) = \bigcup\limits_{v \in G} X(v)$. *Прообразом елемента* $w \in W$ за відповідністю *X* називають множину всіх елементів із *V*, яким поставлено у відповідність елемент $w \in W$. Таку множину елементів із *V* позначають через $X^{-1}(w)$, $X^{-1}(w) \subset V$. *Прообразом множини* $Q \subset W$ елементів із *W* за відповідністю *X* називають об'єднання прообразів елементів із *Q*, тобто $X^{-1}(Q) = \bigcup\limits_{w \in Q} X^{-1}(w)$.

Приклад 4.1.1. Нехай є дві множини
$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ і $W = \{w_1, w_2, w_3, w_4, w_5, w_6\}$.
Тоді множина пар

$$\begin{gathered}\{(v_1, w_1), (v_1, w_2), (v_1, w_3), (v_1, w_4), (v_1, w_5), (v_1, w_6), \\ (v_2, w_1), (v_2, w_2), (v_2, w_3), (v_2, w_4), (v_2, w_5), (v_2, w_6), \\ (v_3, w_1), (v_3, w_2), (v_3, w_3), (v_3, w_4), (v_3, w_5), (v_3, w_6), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ (v_7, w_1), (v_7, w_2), (v_7, w_3), (v_7, w_4), (v_7, w_5), (v_7, w_6)\}\end{gathered}$$

буде декартовим добутком $V \times W$ множин *V* і *W* (Рис. 4.1.1).

За множиною пар
$X = \{(v_1, w_1), (v_2, w_2), (v_3, w_3), (v_4, w_3), (v_5, w_4), (v_6, w_5), (v_7, w_5)\}$
визначається відповідність *X* між множинами *V* і *W*. Тут

$$\begin{gathered}X(v_1) = \{w_1\},\ X(v_2) = \{w_2\},\ X(v_3) = \{w_3\},\ X(v_4) = \{w_3\}, \\ X(v_5) = \{w_4\},\ X(v_6) = \{w_5\},\ X(v_7) = \{w_5\}, \\ X^{-1}(w_1) = \{v_1\},\ X^{-1}(w_2) = \{v_2\},\ X^{-1}(w_3) = \{v_3, v_4\}, \\ X^{-1}(w_4) = \{v_5\},\ X^{-1}(w_5) = \{v_6, v_7\},\ X^{-1}(w_6) = \varnothing.\end{gathered}$$

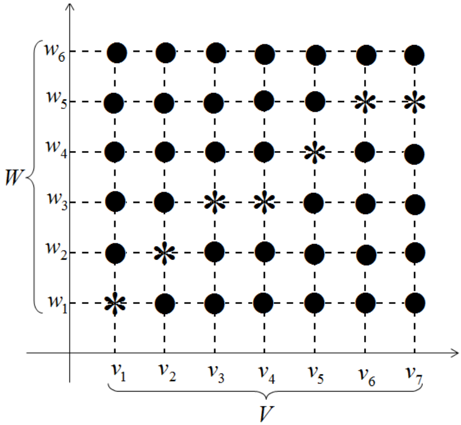


Рис. 4.1.1

В разі, коли за відповідністю X кожному елементові v із V ставиться у відповідність єдиний елемент w із W, тоді говорять, що відповідність X однозначна. Таку відповідність називають *відображенням* множини V в множину W або *функцією* і пишуть $X: V \to W$, або $X(v) = w$, $v \in V$, $w \in W$. Якщо $X(v_1) \neq X(v_2)$, коли $v_1 \neq v_2$, тоді говорять, що відображення взаємно однозначне і пишуть $X: V \leftrightarrow W$. Множини V і W в такому разі називають *еквівалентними*, а відповідність X^{-1} є відображенням $X^{-1}: W \to V$, оберненим до відображення X: $V \to W$.

Зауважимо, що якщо відповідність X не взаємнооднозначна, то не обов'язково за довільного $G \subset V$ має місце рівність $X^{-1}(X(G)) = G$.

Наприклад (див. Рис. 4.1.1), в розглянутому прикладі 4.1.1 $X(v_6) = w_5$, $X^{-1}(w_5) = X^{-1}(X(v_6)) = \{v_6, v_7\} \neq \{v_6\}$. Аналогічно не завжди має місце рівність $X(X^{-1}(Q)) = Q$ за довільного $Q \subset W$.

Якщо відповідність X однозначна, то за довільних $w_1 \in W$ і $w_2 \in W$, $w_1 \neq w_2$, прообрази елементів w_1 і w_2 не перетинаються, тобто $X^{-1}(w_1) \cap X^{-1}(w_2) = \varnothing$. Так само, якщо $Q_1 \subset W$, $Q_2 \subset W$ – дві непорожні підмножини із W такі, що $Q_1 \cap Q_2 = \varnothing$, то $X^{-1}(Q_1) \cap X^{-1}(Q_2) = \varnothing$, коли X є відображенням, тобто однозначною відповідністю. Якщо відображення X взаємно однозначне, тобто відображення X^{-1}, як і X, однозначне, то в образах двох непорожніх множин $G_1 \subset V$ і $G_2 \subset V$, в яких немає спільних елементів, тобто $G_1 \cap G_2 = \varnothing$, також немає спільних елементів, тобто $X(G_1) \cap X(G_2) = \varnothing$.

Приклад 4.1.2. Учневі пропонується випадковим чином обрати приз. Для цього він повинен навмання обрати одну із 12 кульок, серед яких 2 зелених, 3 червоних, 3 білих, 2 блакитних, 1 жовта, 1 чорна.

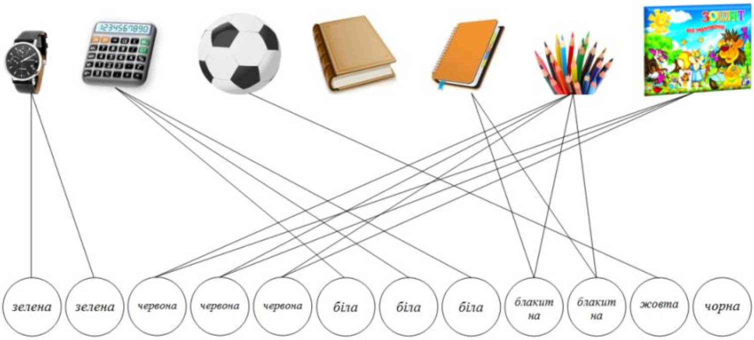


Рис. 4.1.2

Відомо, що серед призів є годинник, калькулятор, м’яч, книга, блокнот, олівці, альбом.

Якщо учень вибере зелену кульку, він одержить годинник,
якщо учень вибере червону кульку, він одержить олівці і альбом,
якщо учень вибере білу кульку, він одержить калькулятор,
якщо учень вибере блакитну кульку, він одержить олівці і блокнот,
якщо учень вибере жовту кульку, він одержить м’яч,
якщо учень вибере чорну кульку, він не одержує приз.

Нехай *V*={зелена, червона, біла, блакитна, жовта, чорна} – множина варіантів вибрати кульку якогось із вказаних шести кольорів, *W*={годинник, калькулятор, м’яч, книга, блокнот, олівці, альбом} – множина призів. Тоді одержимо

X({зелена})={годинник},

X({червона})={олівці, альбом},

X({біла})={калькулятор},

X({блакитна})={блокнот, олівці},

X({жовта})={м’яч},

X({чорна})=$\varnothing$.

Тут

$$X(\{\text{червона}\})\bigcap X(\{\text{блакитна}\})=\{\text{олівці, альбом}\}\bigcap$$
$$\bigcap\{\text{блокнот, олівці}\}=\{\text{олівці}\}\neq\varnothing,$$

перетини всіх інших пар образів елементів множини *V* порожні.

Прообразами елементів множини *W* в даному прикладі будуть:

X^{-1}({годинник})={зелена},

X^{-1}({калькулятор})={біла},
X^{-1}({м’яч})={жовта},
X^{-1}({книга})=$\varnothing$,
X^{-1}({блокнот})={блакитна},
X^{-1}({олівці})={червона, блакитна},
X^{-1}({альбом})={червона}.

В даному разі

X^{-1}({олівці})$\bigcap$ X^{-1}({блокнот})={блакитна},
X^{-1}({олівці})$\bigcap$ X^{-1}({альбом })={червона},

перетини всіх інших пар прообразів елементів множини W порожні.

В розглядуваному випадку відповідність X між елементами множин V та W неоднозначна. Тому X не є відображенням (функцією).

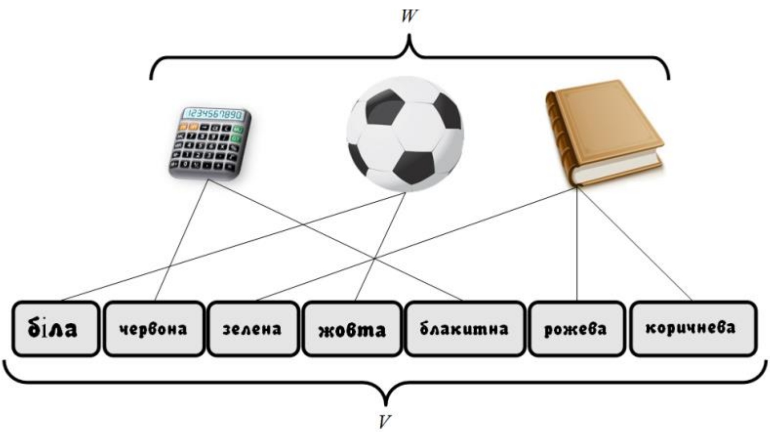


Рис. 4.1.3

Приклад 4.1.3. Учневі пропонується випадковим чином обрати собі один із трьох призів – калькулятор, м’яч, книга. Для цього він повинен навмання вибрати одну із семи кульок, серед яких біла, червона, зелена, жовта, блакитна, рожева, коричнева. Якщо учень вибере червону або блакитну кульку, він одержить калькулятор; якщо учень вибере білу або жовту кульку, він одержить м’яч; якщо учень вибере зелену, або рожеву, або коричневу кульку, він одержить книгу (див. Рис. 4.1.3).

Нехай V={біла, червона, зелена, жовта, блакитна, рожева, коричнева} – множина варіантів вибрати одну кульку якогось із вказаних семи кольорів, W={калькулятор, м'яч, книга} – множина призів.

Тоді одержимо:
X({біла})=X({жовта})={м’яч},

X^{-1}({м’яч})={біла, жовта},
X({червона})=X({блакитна})={калькулятор},
X^{-1}({калькулятор})={червона, блакитна},
X({зелена})=X({рожева})=X({коричнева})={книга},
X^{-1}({книга})={зелена, рожева, коричнева},
X({біла, жовта})={м’яч},
X({червона, блакитна})={калькулятор},
X({біла, жовта} $\cup$ {червона, блакитна})={м’яч} $\cup$
$\cup$ {калькулятор}={м’яч, калькулятор},
X^{-1}({м’яч, книга})= X^{-1}({м’яч}) $\cup$ X^{-1}({книга})=
={біла, жовта} $\cup$ {зелена, рожева, коричнева}=
={біла, жовта, зелена, рожева, коричнева},
X^{-1}({калькулятор, книга})= X^{-1}({калькулятор}) $\cup$
$\cup$ X^{-1}({книга})=
={червона, блакитна} $\cup$ {зелена, рожева, коричнева}=
={червона, блакитна, зелена, рожева, коричнева},
X^{-1}({калькулятор, м’яч})= X^{-1}({калькулятор}) $\cup$ X^{-1}({м’яч})=
={червона, блакитна} $\cup$ {біла, жовта}=
={червона, блакитна, біла, жовта},
X^{-1}({м’яч, книга}) $\cap$ X^{-1}({калькулятор, книга})=
={біла, жовта, зелена, рожева, коричнева} $\cap$ {червона, блакитна, зелена, рожева, коричнева}={зелена, рожева, коричнева}=
= X^{-1}({книга}),
X^{-1}({м’яч, книга}) $\cap$ X^{-1}({калькулятор, м’яч})=
={біла, жовта, зелена, рожева, коричнева} $\cap$
$\cap$ {червона, блакитна, біла, жовта}={біла, жовта}= X^{-1}({м’яч}),
X^{-1}({калькулятор, книга}) $\cap$ X^{-1}({калькулятор, м’яч})=
={червона, блакитна, зелена, рожева, коричнева} $\cap$
$\cap$ {червона, блакитна, біла, жовта}={червона, блакитна}=
= X^{-1}({калькулятор}).

Легко бачити, що в прообразах будь яких двох різних підмножин із W, в яких немає спільних елементів (підмножини не перетинаються), також немає спільних елементів (прообрази не перетинаються).

В даному прикладі відповідність X між елементами множин V={біла, червона, зелена, жовта, блакитна, рожева, коричнева} та W={калькулятор, м'яч, книга} однозначна, оскільки кожному елементові із множини V ставиться у відповідність єдиний елемент із множини W. Тому відповідність X є відображенням (функцією). Разом з тим відповідність X^{-1}, обернена до X, не є однозначною, оскільки в множині W є елементи, яким за правилом X^{-1} ставиться у відповідність не один, а кілька елементів:

X^{-1}({м'яч})={біла, жовта}, X^{-1}({калькулятор})={червона, блакитна}, X^{-1}({книга})={зелена, рожева, коричнева}.

Тому відповідність X^{-1} не є відображенням (функцією).

Нехай задано ймовірнісний простір (Ω, S, P_n^*) і на множині Ω елементарних подій задано дійсну функцію $X(E), E \in \Omega$, за якою кожному елементові E множини Ω ставиться у відповідність деяке дійсне число $x = X(E)$. Це число $x = X(E)$ (чи відповідна точка на числовій осі) називається *образом* елемента E. Не виключається, що одне і те саме число x ставиться у відповідність кільком елементарним подіям (чи навіть нескінченній їх кількості, якщо, наприклад, задана функція набуває одного і того самого значення, тобто стала, на деякому відрізку).

Множина всіх точок (елементарних подій) $E \in \Omega$, яким поставлено у відповідність на числовій осі одну і ту саму точку x, називається *прообразом* точки x. Прообраз точки x позначають символами $X^{-1}(x)$.

Очевидно, щоб знайти прообраз точки x, потрібно знайти множину всіх розв'язків рівняння $X(E) = x$.

Образом деякої множини $A \subset \Omega$ називають об'єднання образів всіх елементів множини A, тобто $\bigcup\limits_{E \in A} X(E)$, яке позначають через $X(A)$. Зокрема $X(\Omega) = \Omega_X$ – це множина значень функції $X(E)$, $E \in \Omega$.

Прообразом деякої множини G *точок на числовій осі* називають об'єднання прообразів всіх елементів множини G, тобто $\bigcup\limits_{x \in G} X^{-1}(x)$, яке позначають через $X^{-1}(G)$.

Зокрема $X^{-1}((-\infty; x))$ – це множина розв'язків нерівності $X(E) < x$, тобто $\{E \mid X(E) \in (-\infty, x)\}$.

Нехай S_X – деяка сукупність підмножин G множини $\Omega_X = X(\Omega)$, стосовно якої задовільняються вимоги 1_s-3_s:

1_s. $\Omega_X \in S_X$;

2_s. Якщо $G \in S_X$, то і $\overline{G} = \Omega_X \setminus G \in S_X$;

3_s. Якщо $G_i \in S_X$, $i = 1, 2, \ldots,$ то і $\bigcup\limits_i G_i \in S_X$.

Функцію $X(E)$, $E \in \Omega$, називають S / S_X*-вимірною функцією* або S / S_X*-випадковою величиною*, якщо за будь-якої підмножини $G \in S_X$ виявляється $X^{-1}(G) \in S$.

Оскільки значення $X(E)$ належить до множини $G \in S_X$ тоді й тільки тоді, коли $E \in X^{-1}(G) \in S$, то статистична ймовірність (відносна частота) попадання значень $X(E)$ в множину $G \in S_X$ дорівнює статистичній ймовірності (відносній частоті) попадання елементарних подій $E \in \Omega$ в множину $X^{-1}(G) \in S$, тобто $P_{nX}^*(G) = P_n^*(X^{-1}(G))$, $G \in S_X$, $X^{-1}(G) \in S$.

Зауважимо, що стосовно P_{nX}^* задовільняються умови 1_p-3_p відносно S_X. Тому $(\Omega_X, S_X, P_{nX}^*)$ є *ймовірнісним простором*. В такому разі говорять, що (за заданого S_X) *ймовірнісний простір* $(\Omega_X, S_X, P_{nX}^*)$ *породжується (генерується) через* S/S_X*-випадкову величину* X *за ймовірнісним простором* (Ω, S, P_n^*).

Зауважимо, що коли $S_X = \{\varnothing, \Omega_X\}$, то тоді будь-яка функція $X(E)$, $E \in \Omega$, буде S/S_X-вимірною за будь-якого ймовірнісного простору (Ω, S, P). В такому разі, якщо не існує $S_X \neq \{\varnothing, \Omega_X\}$, стосовно якої функція $X(E)$, $E \in \Omega$, є S/S_X-вимірною, то стосовно такої функції можна лише стверджувати, що її значення попадають в Ω_X з імовірністю 1, та в $\varnothing$ – з імовірністю 0, що зрештою має місце і стосовно будь-якої іншої функції $X(E)$, $E \in \Omega$.

Зауважимо також, що коли S – найширша сукупність підмножин множини Ω, $S = \{A \mid A \subset \Omega\}$, тобто в S містяться будь які підмножини множини Ω, зокрема і порожня $\varnothing$, то тоді будь яка дійсна функція, задана на Ω, буде S/S_X-вимірною за будь якої сукупності S_X підмножин множини Ω_X, стосовно якої задовільняються вимоги 1_s-3_s.

Надалі, якщо не виникатиме непорозумінь і не вимагатимуться уточнення, S/S_X-випадкову величину іноді будемо називати просто випадковою величиною. Однак завжди матимуться на увазі відповідні ймовірнісні простори (Ω, S, P) і (Ω_X, S_X, P_X) чи (Ω, S, P_n^*) і $(\Omega_X, S_X, P_{nX}^*)$, а отже S/S_X-випадкова величина.

Приклад 4.1.4. Нехай

$$\Omega = \{"1","2","3","4","5","6"\},\ H_1 = \{"1","2","3","4"\},\ H_2 = \{"5"\},$$

$$H_3 = \{"6"\},\ P_n^*(H_1) = 0.10,\ P_n^*(H_2) = 0.30,\ P_n^*(H_3) = 0.60,$$

$$S = \{\varnothing, H_1, H_2, H_3, H_1 + H_2, H_1 + H_3, H_2 + H_3, H_1 + H_2 + H_3 = \Omega\}.$$

Тоді (Ω, S, P_n^*) – ймовірнісний простір.

Нехай на множині $\Omega = \{"1","2","3","4","5","6"\}$ задана функція $X(E)$ наступним чином: $X(E) = 1$, коли $E \in \{"1","2","3","4"\}$; $X(E) = 2$, коли $E \in \{"5"\}$; $X(E) = 3$, коли

$E \in \{"6"\}$.

Якщо елементарні події "1", "2", "3", "4", "5", "6" подати як точки на числовій осі Ox, то графічно зазначену залежність можна подати так, як показано на Рис. 4.1.4.

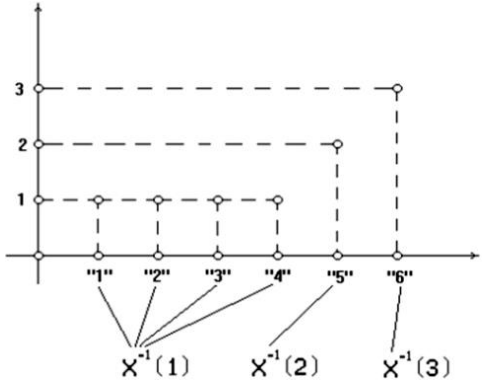


Рис. 4.1.4

Таким чином $X(\Omega) = \Omega_X = \{1, 2, 3\}$.

Розглянемо таку сукупність S_X підмножин множини $X(\Omega)$:

$$S_X = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Очевидно, стосовно цієї сукупності задовільняються вимоги $1_s - 3_s$.

Оскільки одне і те саме число 1 поставлено у відповідність елементарним подіям і "1", і "2", і "3", і "4", прообразом точки 1 є множина $X^{-1}(\{1\}) = H_1 = \{"1","2","3","4"\} \in S$. Це означає, що статистична ймовірність (відносна частота) того, що функція $X(E)$ набуває значення 1, дорівнює статистичній ймовірності (відносній частоті) попадання в підмножину $H_1 = \{"1","2","3","4"\} \in S$, тобто дорівнює 0.10 (бо значення 1 функція $X(E)$ набуває тоді, коли на верхній грані кубика випадає одна з цифр "1", або "2", або "3", або "4").

Аналогічно прообразом точки 2 є одноелементна множина $X^{-1}(\{2\}) = H_2 = \{"5"\} \in S$, тому статистична ймовірність того, що $X(E)$ набуває значення 2, дорівнює статистичній ймовірності випадання цифри "5" на верхній грані кубика, тобто дорівнює 0.30. Прообразом точки 3 є одноелементна множина $X^{-1}(\{3\}) = H_3 = \{"6"\} \in S$, а тому статистична ймовірність того, що функція $X(E)$ набуває значення 3, дорівнює статистичній ймовірності випадання цифри "6" на верхній грані кубика, тобто дорівнює 0.60.

Тепер легко знайти статистичні ймовірності попадання значень $X(E)$ в різні підмножини G множини $X(\Omega)=\Omega_X=\{1, 2, 3\}$, $G\in S_X$. Очевидно, статистична ймовірність попадання значень $X(E)$:

– у підмножину $\{1,2\}\in S_X$ така сама, як статистична ймовірність попадання елементарних подій E в підмножину

$$X^{-1}(\{1,2\})=X^{-1}(\{1\})\bigcup X^{-1}(\{2\})=\{"1","2","3","4"\}\bigcup\{"5"\}=$$
$$=\{"1","2","3","4","5"\}\in S,$$

тобто дорівнює 0.40;

– в підмножину $\{1,3\}\in S_X$ така сама, як в підмножину

$$X^{-1}(\{1,3\})=X^{-1}(\{1\})\bigcup X^{-1}(\{3\})=\{"1","2","3","4"\}\bigcup\{"6"\}=$$
$$=\{"1","2","3","4","6"\}\in S,$$

тобто дорівнює 0.70;

– в підмножину $\{2,3\}\in S_X$ така сама, як в підмножину

$$X^{-1}(\{2,3\})=X^{-1}(\{2\})\bigcup X^{-1}(\{3\})=\{"5"\}\bigcup\{"6"\}=\{"5","6"\}\in S,$$

тобто дорівнює 0.90;

– в підмножину $\{1,2,3\}\in S_X$ така сама, як в підмножину

$$X^{-1}(\{1, 2, 3\})=X^{-1}(\{1\})\bigcup X^{-1}(\{2\})\bigcup X^{-1}(\{3\})=$$
$$=\{"1","2","3","4"\}\bigcup\{"5"\}\bigcup\{"6"\}=\{"1","2","3","4","5","6"\}=\Omega\in S,$$

тобто дорівнює 1.00.

Як бачимо, прообрази всіх підмножин $G\in S_X$ належать до простору подій S, тобто $X^{-1}(G)\in S$, коли $G\in S_X$, а тому функція $X(E)$, $E\in\Omega$, є S/S_X-вимірною, тобто S/S_X-випадковою величиною.

Таким чином на сукупності

$$S_X=\{\varnothing, \{1\},\{2\},\{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$$

підмножин множини $X(\Omega)=\{1, 2, 3\}$ визначено ймовірнісну міру $P^*_{nX}(G)$, $G\in S_X$. Очевидно стосовно S_X задовільняються вимоги 1_s-3_s, а стосовно $P^*_{nX}(G)$, $G\in S_X$, задовільняються вимоги 1_p-3_p відносно S_X:

1_p. $P^*_{nX}(G)\ge 0$, $G\in S_X$.

2_p. Якщо $G_i\in S_X$, $i=1,2,...,$ $G_iG_j=\varnothing$, коли $i\neq j$, то $P^*_{nX}(\bigcup_i G_i)=\sum_i P^*_{nX}(G_i)$.

3_p. $P^*_{nX}(\Omega_X)=1$.

Тому $(\Omega_X, S_X, P^*_{nX})$ є ймовірнісним простором.

Зауважимо, що коли в наведеному прикладі замість розглянутої сукупності S_X обрати сукупність

$$S_X^{(1)} = \{\varnothing, \{1,2\}, \{3\}, \Omega_X\}, \text{ або } S_X^{(2)} = \{\varnothing, \{1\}, \{2,3\}, \Omega_X\},$$

$$\text{або } S_X^{(3)} = \{\varnothing, \{2\}, \{1,3\}, \Omega_X\} \text{ тощо,}$$

кожна з яких входить до S_X, тобто $S_X^{(i)} \subset S_X$, то очевидно розглядувана функція $X(E)$, $E \subset \Omega$, буде і $S/S_X^{(1)}$-вимірною, і $S/S_X^{(2)}$-вимірною, і $S/S_X^{(3)}$-вимірною і т.д., а також S/S_X-вимірною. Це надає можливість обирати різні сукупності S_X залежно від конкретних умов і потреб.

Часто як S_X обирають найширшу сукупність $\tilde{S}_X$ підмножин множини Ω_X таку, щоб за довільного $G \in \tilde{S}_X$ було $X^{-1}(G) \in S$. В такий спосіб за заданими ймовірнісним простором (Ω, S, P) і функцією $X(E)$, $E \subset \Omega$, однозначно визначається і множина $\Omega_X = X(\Omega)$ – образ множини Ω за відображенням *X*, і сукупність $\tilde{S}_X$ підмножин множини Ω_X, і ймовірнісна міра P_X на сукупності $\tilde{S}_X$, тобто ймовірнісний простір $(\Omega_X, \tilde{S}_X, P_X)$. В такому разі говорять, що ймовірнісний простір $(\Omega_X, \tilde{S}_X, P_X)$ породжується через S/S_X-випадкову величину X за ймовірнісним простором (Ω, S, P) і пишуть $(\Omega, S, P) \xrightarrow{X} (\Omega_X, \tilde{S}_X, P_X)$.

Разом з тим для практичних застосувань іноді буває зручніше обрати деяку сукупність $S_X \subset \tilde{S}_X$, стосовно якої задовільняються вимоги 1_s-3_s, і за довільного $G \subset S_X$ матиме місце $X^{-1}(G) \in S$, хоч S_X не найширша і не єдина сукупність серед усіх, прообрази елементів яких належать до *S*. Через таку багатоваріантність не лише не звужуються, а навпаки, розширюються можливості практичних застосувань розглянутих теоретичних положень.

Приклад 4.1.5. Нехай ймовірнісний простір (Ω, S, P_n^*) той самий, що і в прикладі 4.1.4, а на множині $\Omega = \{"1","2","3","4","5","6"\}$ задано функцію $X(E)$ наступним чином: $X(E) = 1$, в разі $E \in \{"1","3","5"\}$, тобто коли на верхній грані кубика випадає непарна цифра, і $X(E) = 2$, в разі $E \in \{"2","4","6"\}$, тобто коли на верхній грані кубика випадає парна цифра (Рис. 4.1.5).

В такому разі $X(\Omega) = \Omega_X = \{1, 2\}$.

Розглянемо таку найширшу сукупність S_X підмножин множини $X(\Omega) = \{1,2\}$: $S_X = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$.

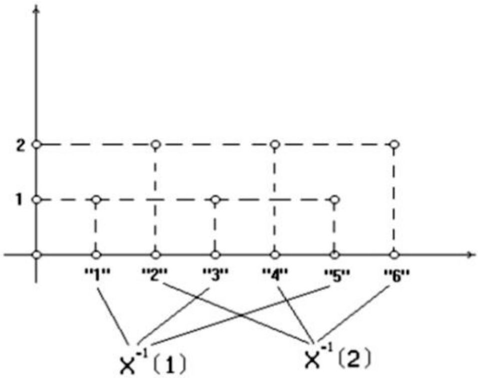


Рис. 4.1.5

Оскільки $X^{-1}(\{1\}) = \{"1","3","5"\} \overline{\in} S$, так само, як і $X^{-1}(\{2\}) = \{"2","4","6"\} \overline{\in} S$, то в даному прикладі функція $X(E)$, задана на множині $\Omega = \{"1","2","3","4","5","6"\}$, не є S/S_X-вимірною відносно так заданих S_X і (Ω, S, P), тобто не є S/S_X-випадковою величиною. За так заданої функції $X(E)$, $E \in \Omega$, на заданому ймовірнісному просторі $(\Omega,\ S,\ P_n^*)$ і так заданого S_X неможливо визначити статистичну ймовірність (відносну частоту)

$$P_{nX}^*(\{1\}) = P_n^*(X^{-1}(\{1\})) = P_n^*(\{"1","3","5"\}) \text{ чи}$$

$$P_{nX}^*(\{2\}) = P_n^*(X^{-1}(\{2\})) = P_n^*(\{"2","4","6"\}),$$

бо

$$X^{-1}(\{1\}) = \{"1","3","5"\} \overline{\in} S \text{ і } X^{-1}(\{2\}) = \{"2","4","6"\} \overline{\in} S,$$

тобто множини $X^{-1}(\{1\})$ і $X^{-1}(\{2\})$ виявляються невимірними відносно заданої на S ймовірнісної міри $P_n^*(A)$, $A \in S$.

Зауважимо, що коли в останньому прикладі розглянути таку сукупність S_{1X} підмножин множини $X(\Omega)$: $S_{1X} = \{\varnothing, \{1,2\}\}$, тоді функція $X(E)$, $E \in \Omega$, виявляється S/S_{1X}-вимірною, бо $X^{-1}(\varnothing) = \varnothing \in S$, $X^{-1}(\{1,2\}) = \Omega \in S$ і

$$P_{nX}^*(\varnothing) = P_n^*(X^{-1}(\varnothing)) = P_n^*(\varnothing) = 0,$$

$$P_{nX}^*(\{1,2\}) = P_n^*(X^{-1}(\{1,2\})) = P_n^*(\Omega) = 1.$$

Зауважимо також, що не завжди буває необхідним і практично прийнятним знаходити статистичні ймовірності попадання в будь які підмножини множини $X(\Omega)$ (див. приклад 1.8.2).

Часто *за заданої дійсної функції* $X(E)$ на просторі елементарних подій Ω замість $X(\Omega)$ як Ω_X розглядають простір $R^1=(-\infty;\infty)$, покладаючи $X^{-1}(x)=\varnothing$ в разі, коли значення (точка) $x\in R^1$ не поставлене у відповідність жодній елементарній події $E\in\Omega$.

Якщо в сукупності $S_X=\mathscr{B}(R^1)$, яка породжується за числовими проміжками (див. §1.5, §2.6), зокрема містяться всі проміжки $(-\infty;x)$, $x\in(-\infty;\infty)$, і за довільного $x\in(-\infty;\infty)$ $X^{-1}((-\infty;\ x))\in S$, тоді функцію $X(E)$ називають S-вимірною або випадковою величиною стосовно ймовірнісного простору (Ω,S,P) та так обраного простору S_X підмножин із R^1.

В такому разі до сукупності $S_X=\mathscr{B}(R^1)$ входять довільні числові проміжки та об’єднання довільної не більш ніж зчисленної кількості таких проміжків, доповнення таких об’єднань до R^1, скінченні та зчисленні множини тощо. Кожна відкрита множина із R^1 входить до S_X, оскільки може бути подана як сума скінченної або зчисленної кількості інтервалів. Кожна замкнена множина належить до S_X, оскільки є доповненням відкритої множини до R^1. Зокрема кожна скінченна множина виду $\{x_1,x_2,...,x_k\}$ належить до S_X (див. §2.6).

Виявляється, що коли прообраз будь-якої множини виду $(-\infty;\ x)\in S_X$, $x\in R^1$, належить до S, тобто якщо $X^{-1}((-\infty;\ x))\in S$, $x\in R^1$, то і прообраз будь-якої підмножини $G\in S_X$ належить до S, тобто $X^{-1}(G)\in S$, $G\in S_X$.

Зауважимо, що насправді в означенні S-вимірної функції мається на увазі S/S_X-вимірна функція, хоч сукупність $S_X=\mathscr{B}(R^1)$ явно і не вказується.

Оскільки значення $X(E)$ належить до множини $G\in S_X$ лише тоді, коли $E\in X^{-1}(G)\in S$, то статистична ймовірність $P^*_{nX}(G)$ попадання значень $X(E)$ в множину $G\in S_X$ дорівнює статистичній ймовірності $P^*_n(X^{-1}(G))$ попадання елементарних подій $E\in\Omega$ в множину $X^{-1}(G)\in S$, тобто $P^*_{nX}(G)=P^*_n(X^{-1}(G))$. Це стосується і будь-якої ймовірнісної міри $P(A)$, $A\in S$, заданої на S, та ймовірнісної міри $P_X(G)$, $G\in S_X$, породжуваної через S/S_X-вимірну функцію X за мірою $P(A)$, $A\in S$, заданою на S, тобто $P_X(G) = P(X^{-1}(G))$, $G \in S_X$, $X^{-1}(G) \in S$.

Часто означення випадкової величини подають в слідуючий спосіб.

Нехай (Ω, S, P) – деякий ймовірнісний простір, $(R^1, \mathscr{B}(R^1))$ – числова пряма із системою борелівських множин $\mathscr{B}(R^1)$.

Означення. Дійсна функція $X(E)$, $E \in \Omega$, називається S*-вимірною* або *випадковою величиною*, заданою на ймовірнісному просторі (Ω, S, P), якщо $X^{-1}(B) \in S$ за довільного $B \in \mathscr{B}(R^1)$.

Разом з тим насправді тут мається на увазі S / S_X-вимірна функція, де $S_X = \mathscr{B}(R^1)$.

Якщо $(\Omega, S) = (R^n, \mathscr{B}(R^n))$ *за деякого* $n \in N$, *то* $\mathscr{B}(R^n)$*-вимірні функції називаються борелівськими.*

Виходячи з ймовірнісного простору (Ω, S, P), можна побудувати ймовірнісний простір $(\Omega_X = R^1, S_X = \mathscr{B}(R^1), P_X)$, поклавши

$$P_X(G) = P(X^{-1}(G)),\ G \in \mathscr{B}(R^1).$$

Таким чином, ймовірнісний простір

$$(\Omega_X = R^1,\ S_X = \mathscr{B}(R^1),\ P_X)$$

індукується через випадкову величину X за ймовірнісним простором (Ω, S, P).

Через ймовірнісну міру P_X на $(R^1, \mathscr{B}(R^1))$

$$P_X(G) = P(X^{-1}(G)),\ G \in \mathscr{B}(R^1),$$

визначається *розподіл ймовірностей* на множині $\Omega_X = R^1$ значень випадкової величини X (на $(R^1, \mathscr{B}(R^1))$).

Приклад 4.1.6. Нехай $\Omega = \{1, 2, 3, 4, 5, 6\}$, $S = \{\varnothing,\ \Omega,\ \{1, 2, 3, 4\},\ \{5\},\ \{6\},\ \{1, 2, 3, 4, 5\},\ \{1, 2, 3, 4, 6\},\ \{5, 6\}\}$, а $X(E) = 1$, коли $E \in \{1, 2\}$, $X(E) = 2$, коли $E \in \{3, 4\}$, $X(E) = 3$, коли $E \in \{5, 6\}$. Тоді $\Omega_X = \{1, 2, 3\}$.

Знайдемо сукупність $\widetilde{S}_X$ підмножин G множини Ω_X таку, що матиме місце $X^{-1}(G) \in S$ за кожного $G \in \widetilde{S}_X$, тобто $\tilde{S}_X = \{G \mid G \subset \Omega_X,\ X^{-1}(G) \in S\}$.

Оскільки

$$X^{-1}(1)=\{1,2\}\,\overline{\in}\, S\,,\ \ X^{-1}(2)=\{3,4\}\,\overline{\in}\, S\,,\ \ X^{-1}(3)=\{5,6\}\in S\,,$$

$$X^{-1}(\{1,2\})=X^{-1}(1)\cup X^{-1}(2)=\{1,2,3,4\}\in S\,,$$

то $\tilde{S}_X=\{\varnothing,\Omega_X,\{1,2\},\{3\}\}$.

Таким чином, дана функція є $S/\tilde{S}_X$-вимірною, тобто $S/\tilde{S}_X$-випадковою величиною. Стосовно даної функції існує єдиний нетривіальний простір $S_X=\tilde{S}_X$, за якого функція $X(E)$, $E\in\Omega$, є S/S_X-вимірною. В розглядуваному випадку $S_X=\tilde{S}_X$ не співпадає із найширшою сукупністю підмножин множини Ω_X.

Приклад 4.1.7. Нехай $\Omega=\bigcup\limits_{i=1}^{6}H_i$, $H_iH_j=\varnothing$, $i\neq j$ (Рис. 4.1.6), простір подій $S=\{A\,|\ A=\bigcup\limits_{i\in I}H_i, I\subset\{1,2,...,6\}\}=$ $=\{\varnothing,H_1,...,H_6,H_1+H_2,...,\Omega\}$ – породжений за поділом множини Ω на підмножини H_i, $i\in\overline{1,6}$, $P_n^*(H_i)$, $i\in\overline{1,6}$, – задані.

Нехай на множині Ω задана кусково-стала функція $X_1(E)$, $E\in\Omega$: $X_1(\Omega)=\Omega_{X_1}=\{x_1,x_2,x_3,x_4,x_5\}$, яка на кожній із множин H_i набуває сталого значення (Рис. 4.1.6).

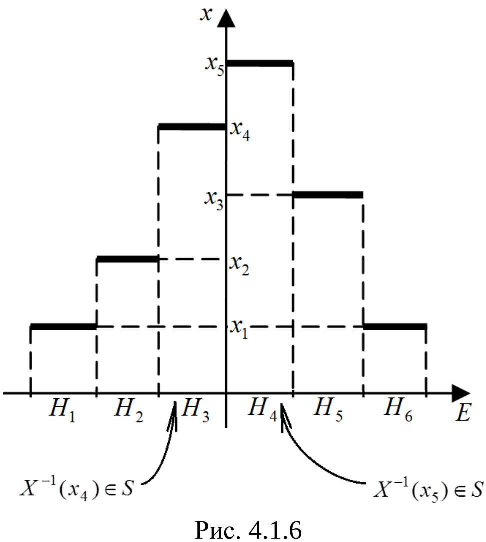


Рис. 4.1.6

Оскільки прообрази всіх елементів x_1, x_2, x_3, x_4, x_5

множини Ω_{X_1}: $X_1^{-1}(x_1)=H_1\bigcup H_6$, $X_1^{-1}(x_2)=H_2$, $X_1^{-1}(x_3)=H_5$, $X_1^{-1}(x_4)=H_3$, $X_1^{-1}(x_5)=H_4$ входять до породженої за підмножинами H_1, H_2, H_3, H_4, H_5, H_6 сукупності S підмножин множини Ω, то можна визначити статистичні ймовірності попадання в множини $\{x_1\}$, $\{x_2\}$, $\{x_3\}$, $\{x_4\}$, $\{x_5\}$, а тому і в будь-які підмножини множини Ω_{X_1}, звідки слідує, що функція $X_1(E)$, $E\in\Omega$, буде S/S_{X_1}-вимірною, тобто S/S_{X_1}-випадковою величиною за будь-якої сукупності S_{X_1}, стосовно якої задовільняються вимоги 1_s-3_s.

Нехай S_{X_1} – найширша сукупність підмножин множини Ω_{X_1}, тобто до S_{X_1} входить всі підмножини множини Ω_{X_1} (разом з $\varnothing$ і Ω_{X_1}).

Тоді, оскільки $X_1^{-1}(G)\in S$ за довільного $G\in S_{X_1}$, $G\subset\Omega_{X_1}$, то така функція $X_1(E)$, $E\in\Omega$, є S/S_{X_1}-вимірною, тобто S/S_{X_1}-випадковою величиною.

Очевидно, існують і інші сукупності S_{X_1} підмножин множини Ω_{X_1}, за яких функція $X_1(E)$ буде S/S_{X_1}-вимірна, наприклад,

$$S_{X_1}=\{\varnothing,\{x_1\},\{x_2\}+\{x_3\}+\{x_4\}+\{x_5\},\Omega_{X_1}\},$$

$$S_{X_1}=\{\varnothing,\{x_1\}+\{x_2\}+\{x_3\},\{x_4\}+\{x_5\},\Omega_{X_1}\}\text{ і т.д.}$$

Приклад 4.1.8. Нехай, як і в попередньому прикладі, $\Omega=\bigcup\limits_{i=1}^{6}H_i$, $H_iH_j=\varnothing$, $i\neq j$, $S=\{A\mid A=\bigcup\limits_{i\in I}H_i, I\subset\{1,2,...,6\}\}=$ $=\{\varnothing,H_1,...,H_6,H_1+H_2,...,\Omega\}$ – породжена за поділом множини Ω на підмножини H_i, $i\in\overline{1,6}$, $H_iH_j=\varnothing$, коли $i\neq j$, $\bigcup\limits_{i=1}^{6}H_i=\Omega$, $P_n^*(H_i)$, $i\in\overline{1,6}$, – задані.

Нехай на множині Ω задана кусково-стала функція $X_2(E)$, $E\in\Omega$, графік якої подано на Рис. 4.1.7 (функція $X_2(E)$ не набуває сталих значень на кожній із множин H_i). В розглядуваному випадку $X_2(\Omega)=\Omega_{X_2}=\{x_1,x_2,x_3,x_4,x_5\}=\Omega_{X_1}$, тобто множина Ω_{X_2} значень функції $X_2(E)$, $E\in\Omega$, така сама, як і множина Ω_{X_1} значень функції $X_1(E)$, $E\in\Omega$, із прикладу 4.1.7.

Нехай як і раніше, S_{X_2} – найширша сукупність підмножин множини Ω_{X_2}, тобто в S_{X_2} входять всі підмножини множини Ω_{X_2} (разом з $\varnothing$ і Ω_{X_2}).

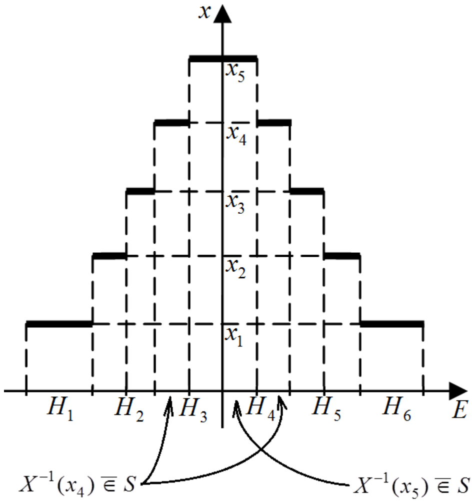


Рис. 4.1.7

Оскільки прообрази елементів x_2, x_3, x_4, x_5 складаються з інтервалів, які не входять до сукупності S, породженої за поділом множини Ω на підмножини H_1, H_2, H_3, H_4, H_5, H_6, то неможливо визначити статистичні ймовірності попадання в множини $\{x_2\}$, $\{x_3\}$, $\{x_4\}$, $\{x_5\}$, які входять до S_{X_2}, а тому функція $X_2(E)$ не є S/S_{X_2}-вимірною за так заданого S_{X_2}.

Разом з тим, якщо обрати S_{X_2} так:

$$S_{X_2} = \{\varnothing, \{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_1\} \bigcup \{x_2, x_3\}, \{x_1\} \bigcup \{x_4, x_5\},$$
$$\{x_2, x_3\} \bigcup \{x_4, x_5\}, \{x_1\} \bigcup \{x_2, x_3\} \bigcup \{x_4, x_5\} = \Omega_{X_2}\},$$

тоді прообрази всіх елементів так заданого S_{X_2} належатимуть до S, а тому в такому разі можна визначити статистичні ймовірності попадання в підмножини $G \in S_{X_2}$:

$$P^*_{nX_2}(\{x_1\}) = P^*_n(X_2^{-1}(\{x_1\})) = P^*_n(H_1 \bigcup H_6) = P^*_n(H_1) + P^*_n(H_6),$$
$$P^*_{nX_2}(\{x_2, x_3\}) = P^*_n(X_2^{-1}(\{x_2, x_3\})) = P^*_n(H_2 \bigcup H_5) = P^*_n(H_2) + P^*_n(H_5),$$
$$P^*_{nX_2}(\{x_4, x_5\}) = P^*_n(X_2^{-1}(\{x_4, x_5\})) = P^*_n(H_3 \bigcup H_4) = P^*_n(H_3) + P^*_n(H_4).$$

Очевидно, існують і інші сукупності S_{X_2} підмножин множини Ω_{X_2}, за яких функція $X_2(E)$ буде S/S_{X_2}-вимірна. Якщо позначити $\{x_1\}=\widetilde{H}_1$, $\{x_2,x_3\}=\widetilde{H}_2$, $\{x_4,x_5\}=\widetilde{H}_3$, тоді задаючи сукупності S_{X_2} у вигляді $S_{X_2}=\{\varnothing,\widetilde{H}_1,\widetilde{H}_2+\widetilde{H}_3,\Omega\}$, $S_{X_2}=\{\varnothing,\widetilde{H}_1+\widetilde{H}_2,\widetilde{H}_3,\Omega\}$ і т.д., будемо щоразу одержувати S/S_{X_2}-вимірну функцію $X_2(E)$, $E\in\Omega$.

Приклад 4.1.9. Нехай $\Omega=\bigcup\limits_{i=1}^{k}H_i$, $H_iH_j=\varnothing$, коли $i\neq j$, S породжена за поділом множини Ω на підмножини H_i, $P_n^*(H_i)$ задані. Нехай задано функцію $X(E)$, $E\in\Omega$, $\Omega_X=X(\Omega)$, $H_{iX}=X(H_i)$. Якщо відображення $X\colon\ \Omega\to\Omega_X$ взаємнооднозначне, тоді як S_X можна обрати сукупність, породжену за образами H_{iX} множин H_i, $H_{iX}=X(H_i)$. Зрозуміло, що в такому разі за довільного $G\in S_X$, $G=\bigcup\limits_{i\in I}H_{iX}$, $I\subset\{1,2,...,k\}$, буде

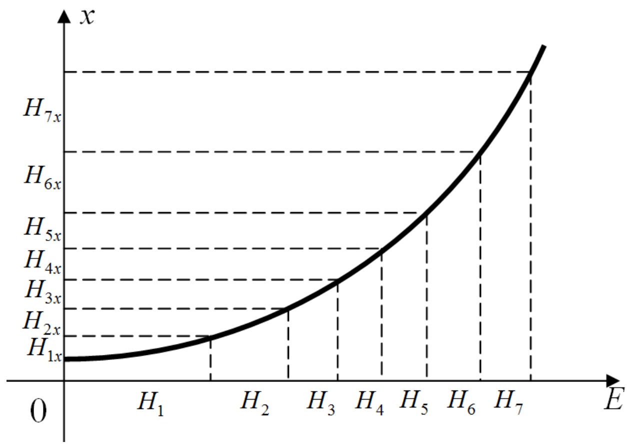


Рис. 4.1.8

$$X^{-1}(G)=X^{-1}\left(\bigcup_{i\in I}H_{iX}\right)=\bigcup_{i\in I}X^{-1}(H_{iX})=\bigcup_{i\in I}H_i\in S,$$

і тому функція $X(E)$, $E\in\Omega$, за так заданих S і S_X буде S/S_X-вимірна, тобто S/S_X-випадкова величина (Рис. 4.1.8).

В розглядуваному випадку

$$P_{nX}^{*}(G)=P_{n}^{*}(X^{-1}(G))=P_{n}^{*}\left(\bigcup_{i\in I}H_i\right)=\sum_{i\in I}P_{n}^{*}(H_i)\,.$$

Очевидно, виходячи з так побудованого простору S_X, як і раніше, можна побудувати і інші сукупності $\widetilde{S}_X$ підмножин множини Ω_X, стосовно яких задовільнятимуться вимоги 1_s-3_s, і функція $X(E)$ також буде $S/\widetilde{S}_X$ -вимірна.

Приклад 4.1.10. Нехай $\Omega=[0;1]$, S – σ-алгебра вимірних за Лебегом множин із $[0;1]$. Якщо $P(A)=m(A)$, де $A\in S$, $m(A)$ – міра Лебега множини A, то (Ω, S, P) – ймовірнісний простір.

Нехай G – невимірна за Лебегом підмножина відрізка $[0;1]$. Розглянемо функцію

$$X(E)=\begin{cases}1, \text{ коли } E\in G;\\ -1, \text{ коли } E\in\overline{G}=[0;1]\setminus G.\end{cases}$$

Тоді

$$X^{-1}((-\infty;x))=\{E\mid X(E)\in(-\infty;x)\}=\begin{cases}\varnothing, \text{ коли } x\le -1;\\ \overline{G}, \text{ коли } -1<x\le 1;\\ [0;1], \text{ коли } 1<x.\end{cases}$$

Отже, $X^{-1}((-\infty;x))=\{E\mid X(E)\in(-\infty;x)\}=\overline{G}\,\overline{\in}\,S$, коли $x\in(-1;1]$, і таким чином функція $X(E)$ не є випадковою величиною на ймовірнісному просторі (Ω, S, P), оскільки прообраз борелівської множини $(-\infty;x)$, коли $x\in(-1;1]$, не є вимірною множиною, тобто не належить до σ-алгебри S вимірних за Лебегом множин із $[0;1]$.

Можна показати, що для того, щоб $X(E)$ була випадковою величиною, необхідно й достатньо, щоб за будь-яких $x\in R^1$ мало місце включення

$$X^{-1}((-\infty;x))\in S\,.$$

Це випливає з того, що за системою Q множин виду $(-\infty;c)=\{x\mid x<c,\, c\in R^1\}$ породжується σ-алгебра $\mathscr{B}(R^1)$, $\sigma(Q)=\mathscr{B}(R^1)$, (бо довільний інтервал $[x_1;x_2)$ можна подати як різницю множин $(-\infty;x_2)$ та $(-\infty;x_1)$, тобто

$[x_1; x_2) = (-\infty; x_2) \setminus (-\infty; x_1))$. Часто співвідношення $X^{-1}((-\infty; x)) \in S$ покладають в основу означення S-вимірної функції (випадкової величини).

Нехай (Ω, S, P) – ймовірнісний простір, X і Y – $S / \mathscr{B}(R^1)$ випадкові величини, визначені на ньому, $\Omega_X = R^1$, $\Omega_Y = R^1$, $S_X = \mathscr{B}(R^1)$, $S_Y = \mathscr{B}(R^1)$, a, b, c – дійсні числа. Тоді

а) $cX(E)$; б) $X(E) + c$; в) $|X(E)|$; г) $X^2(E)$;

д) $1/X(E)$; е) $X(E) \pm Y(E)$; є) $X(E)Y(E)$; ж) $X(E)/Y(E)$

також $S / \mathscr{B}(R^1)$-випадкові величини (в останньому випадку припускається, що $P(\{E | Y(E) \neq 0\}) = 1)$.

Доведемо це.

а) коли $c > 0$, тоді (Рис. 4.1.9)

$$\{E | \; cX(E) \in (-\infty; x)\} = \{E | \; cX(E) < x\} = \{E | \; X(E) < \frac{x}{c}\} = X^{-1}((-\infty; \frac{x}{c})) \in S;$$

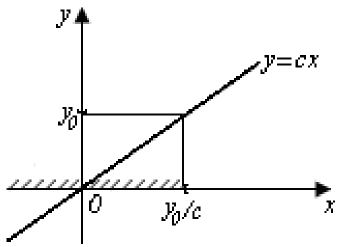


Рис. 4.1.9

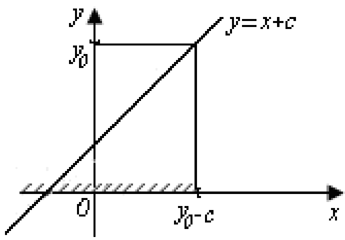


Рис. 4.1.10

коли $c < 0$, тоді

$$\{E | \; cX(E) \in (-\infty; x)\} = \{E | \; cX(E) < x\} = \{E | \; X(E) > \frac{x}{c}\} = X^{-1}((\frac{x}{c}; \infty)) \in S;$$

оскільки $(-\infty; \frac{x}{c})$ та $(\frac{x}{c}; \infty)$ – борелівські множини, X – $S / \mathscr{B}(R^1)$-випадкова величина, то $X^{-1}((-\infty; \frac{x}{c})) \in S$, $X^{-1}((\frac{x}{c}; \infty)) \in S$, отже cX – випадкова величина;

б) легко бачити, що (Рис. 4.1.10)

$$\{E | \; X(E) + c < x\} = \{E | \; X(E) < x - c\} = X^{-1}((-\infty; x - c)) \in S.$$

Отже $X+c\ -\ S/\mathscr{B}(R^1)$-випадкова величина.

Зрозуміло, що коли $X(E)=b=\text{const}$, то $X(E)$ – $S/\mathscr{B}(R^1)$-випадкова величина, оскільки

$$\{E \mid X(E)\in(-\infty;x)\}=\{E \mid X(E)<x\}=\begin{cases}\varnothing\in S, & \text{коли}\quad x\le \text{b};\\ \Omega\in S, & \text{коли}\quad b<x.\end{cases}$$

Звідси та з а) і б) випливає: якщо $X\in S/\mathscr{B}(R^1)$-випадкова величина, то $aX+b$ – також $S/\mathscr{B}(R^1)$-випадкова величина;

в) очевидно (Рис. 4.1.11)

$\{E \mid |X(E)|<x\}=\varnothing\in S$, коли $x\le 0$;

$\{E \mid |X(E)|<x\}=\{E \mid X(E)>-x\}\bigcap\{E \mid X(E)<x\}\in S$, коли $x>0$;

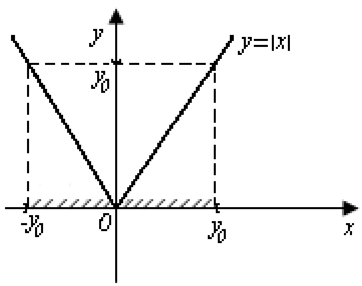


Рис.4.1.11

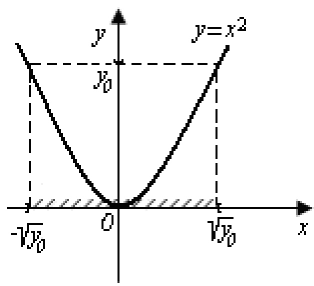


Рис.4.1.12

г) аналогічно до попереднього (Рис. 4.1.12)

$$\{E \mid (X(E))^2<x\}=\begin{cases}\varnothing\in S, & \text{коли}\quad x<0;\\ \{E \mid X(E)|<\sqrt{x}\}\in S, & \text{коли}\quad x\ge 0.\end{cases}$$

д) стосовно $\dfrac{1}{X(E)}$ одержимо (Рис. 4.1.13)

$$\{E \mid \frac{1}{X(E)}<x\}=\begin{cases}\{E \mid X(E)<0\}\bigcup\{E \mid X(E)>\dfrac{1}{x}\}\in S, & \text{коли}\quad x>0;\\ \{E \mid X(E)<0\}\bigcap\{E \mid X(E)>\dfrac{1}{x}\}\in S, & \text{коли}\quad x<0;\\ \{E \mid X(E)<0\}\in S, & \text{коли}\quad x=0.\end{cases}$$

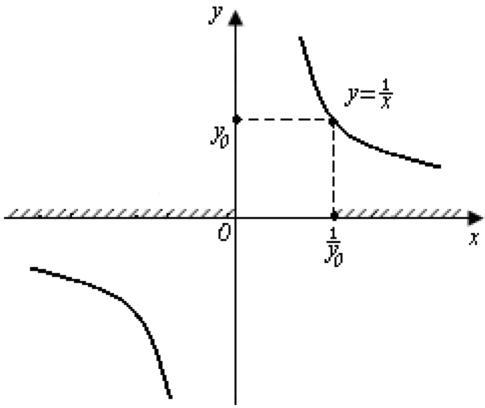


Рис. 4.1.13

е) Нехай X і Y – $S/\mathscr{B}(R^1)$-випадкові величини. Пронумеруємо всі раціональні числа $r_1, r_2, \ldots$ Тоді за будь-якого дійсного x

$$\{E \mid X(E)+Y(E)<x\} = \bigcup_{k=1}^{\infty}\{E \mid X(E)<x-r_k\}\bigcap\{E \mid Y(E)<r_k\} \in S .$$

Справді, якщо $E_0 \in \{E \mid X(E)+Y(E)<x\}$, то $Y(E_0)<x-X(E_0)$. Оскільки множина раціональних чисел всюди щільна в R^1, то знайдеться таке r_k, що $Y(E_0)<r_k<x-X(E_0)$, тобто $Y(E_0)<r_k$, $X(E_0)<x-r_k$.

Таким чином, E_0 належить принаймні до однієї із множин у правій частині рівності, а отже і до їх об'єднання.

Якщо E_0 входить до об'єднання множин у правій частині рівності, то за деякого k

$$X(E_0)<x-r_k , \qquad Y(E_0)<r_k , \qquad X(E_0)+Y(E_0)<x ,$$

тобто E_0 належить до множини $\{E \mid X(E)+Y(E)<x\}$. Звідси випливає правильність наведеної рівності і те, що $X(E)+Y(E)$ – $S/\mathscr{B}(R^1)$-випадкова величина.

Оскільки $X(E)-Y(E)=X(E)+(-1)\cdot Y(E)$ і, згідно з а), $-1\cdot Y(E)$ є $S/\mathscr{B}(R^1)$-випадковою величиною, то, згідно з е), $X(E)-Y(E)$ є також $S/\mathscr{B}(R^1)$-випадковою величиною.

З а), б), е), зокрема випливає:

$$\{E \mid X(E)<Y(E)\}=\{E \mid X(E)-Y(E)<0\}\in S ,$$
$$\{E \mid X(E)\le Y(E)\}=\{E \mid X(E)-Y(E)\le 0\}\in S ,$$

$$\{E \mid X(E) > Y(E)\} = \{E \mid X(E) - Y(E) > 0\} \in S,$$
$$\{E \mid X(E) = Y(E)\} = \{E \mid X(E) - Y(E) = 0\} \in S.$$

є) Оскільки $XY = \frac{1}{2}(X+Y)^2 - X^2 - Y^2$, то з а), г), е) випливає, що XY – $S/\mathscr{B}(R^1)$-випадкова величина.

ж) З рівності $X/Y = X \cdot \frac{1}{Y}$, враховуючи д) і є), дістаємо, що $\frac{X}{Y}$ – $S/\mathscr{B}(R^1)$-випадкова величина.

Якщо $\varphi(x)$ – борелівська функція, а $X(E)$ – $S/\mathscr{B}(R^1)$-випадкова величина, то складена функція $\varphi(X(E)) = \psi(E)$ є також $S/\mathscr{B}(R^1)$-випадковою величиною.

Справді, в разі $B \in \mathscr{B}(R^1)$ буде

$$\psi^{-1}(B) = \{E \mid \varphi(X(E)) \in B\} = \{E \mid X(E) \in \varphi^{-1}(B)\} = X^{-1}(\varphi^{-1}(B)) \in S,$$

оскільки $\varphi^{-1}(B) \in \mathscr{B}(R^1)$.

Таким чином в разі, коли X – $S/\mathscr{B}(R^1)$-випадкова величина, то функції

$$X^n; \ |X|; \ X^+ = \begin{cases} X, & \text{коли } X \ge 0, \\ 0, & \text{коли } X < 0, \end{cases} \quad X^- = \begin{cases} 0, & \text{коли } X \ge 0, \\ -X, & \text{коли } X < 0, \end{cases}$$

є також випадковими величинами, оскільки функції x^n, $|x|$, x^+, x^- борелівські.

Приклад 4.1.11. Нехай на (Ω, S, P) $X(E) = 0$. Тоді X – $S/\mathscr{B}(R^1)$-випадкова величина. Справді

$$X^{-1}((-\infty; x)) = \begin{cases} \varnothing, & \text{коли } x \le 0; \\ \Omega, & \text{коли } x > 0, \end{cases}$$

і таким чином, за будь-якого x

$$X^{-1}((-\infty; x)) \in S.$$

Приклад 4.1.12. Якщо S – система всіх підмножин множини Ω, то будь-яка дійсна функція $X(E)$ на Ω є S/S_X-випадковою величиною.

Вправи для самостійного виконання

4.1.1. Перевірити, чи правильні твердження:

1. Стосовно будь-якої функції $X(E)$, $E \in \Omega$, за кожної елементарної події $E \in \Omega$ існує єдиний образ.

2. За будь-якої функції $X(E)$, $E \in \Omega$, в прообразі $X^{-1}(x)$ числа $x \in R$ міститься лише один елемент.

3. Якщо функція $X(E)$, $E \in \Omega = [a;b]$, є зростаючою, то в прообразі $X^{-1}(x)$ міститься лише один елемент за кожного $x \in X([a;b])$.

4. Якщо функція $X(E)$ визначена на просторі Ω елементарних подій, то вона є випадковою величиною.

5. Твердження, обернене до 4, є правильним.

6. Дійсна функція $X(E)$, $E \in \Omega$, є S/S_X-вимірною або S/S_X-випадковою величиною, коли множина розв'язків нерівності $X(E) < x$, тобто $X^{-1}((-\infty; x))$, є подією за будь-якого числа x, а сукупність $S_X = \mathscr{B}(R^1)$.

7. Якщо простір подій S не є найширшим простором за даного простору Ω елементарних подій, то існують функція $X(E)$, $E \in \Omega$, і S_X такі, що $X(E)$ не є S/S_X-випадковою величиною.

8. Випадкова величина – це будь-яка функція, визначена на просторі елементарних подій.

9. Якщо $X(E)$ – дійсна функція, визначена на просторі елементарних подій, то $X(E)$ – випадкова величина, коли множина розв'язків рівняння $X(E) = x$ (де x – відоме, а E – невідоме) є подією: а) за деякого $x \in R$; б) за будь-якого $x \in R$.

10. Якщо $X(E)+Y(E)$ – випадкова величина, то $X(E)$ та $Y(E)$ – випадкові величини.

11. Якщо $X(E) \pm Y(E)$ – випадкові величини, то $X(E)$ та $Y(E)$ – випадкові величини.

12. Існує ймовірнісний простір (Ω, S, P), стосовно якого будь-яка дійсна функція $X(E)$, $E \in \Omega$, є випадковою величиною.

4.1.2. Нехай $\Omega = \{\Gamma, Ц\}$. Побудувати ймовірнісний простір (Ω, S, P_n^*), функцію $X(E)$, $E \in \Omega$, і S_X такі, що: а) $X(E)$ є S/S_X-випадковою величиною; б) $X(E)$ не є S/S_X-випадковою величиною.

4.1.3. Нехай задано дискретний поточковий розподіл статистичних ймовірностей на скінченній множині $\Omega = \{x_1, x_2, ..., x_k\}$ – $P_n^*(\{x_i\})$, $i \in \overline{1, k}$, причому кожна елементарна подія $E = x_i$ включена до простору S як подія: $\{E\} = \{x_i\} \in S$.

1. Довести, що функція $X(E) = P_n^*(E)$, $E \in \Omega = \{x_1, x_2, ..., x_k\}$, є

S/S_X-випадковою величиною за довільної сукупності підмножин множини $\Omega_X = X(\Omega)$, стосовно якої задовільняються вимоги 1_s-3_s.

2. Чи є твердження 1 правильним, коли множина $\{x_i\}$ не є подією за деякого $i \in \overline{1, k}$?

4.1.4. Нехай задано неперервний розподіл узагальнених статистичних ймовірностей з щільністю розподілу $f(x)$, $x \in R$. Чи є $f(x)$ S/S_X -випадковою величиною? Якщо так, то за яких S і S_X ?

4.1.5. Чи є S/S_X -випадковою величиною функція розподілу статистичних ймовірностей? Якщо так, то за яких S і S_X ?

4.1.6. Навести приклади дискретних та неперервних просторів Ω елементарних подій, відповідних імовірнісних просторів, функцій $X(E)$, $E \in \Omega$, та сукупностей S і S_X підмножин множин відповідно Ω і $\Omega_X = X(\Omega)$ таких, що $X(E)$ є S/S_X -випадковими величинами, і таких, що $X(E)$ не є S/S_X -випадковими величинами.

4.2. Розподіли статистичних ймовірностей на множинах значень випадкових величин та їх числові характеристики

Нехай задано ймовірнісний простір (Ω, S, P) і функцію $X(E)$, $E \in \Omega$, яка є S/S_X-вимірною, де S_X – деяка сукупність підмножин множини $\Omega_X = X(\Omega)$, стосовно якої задовільняються вимоги 1_s-3_s.

Якщо побудовано ймовірнісний простір $(\Omega_X, S_X, P_{nX}^*)$ *, який генерується через* S/S_X *-випадкову величину X за ймовірнісним простором* (Ω, S, P_n^*) *та* S_X *, і тим самим за будь-якого* $G \in S_X$ *визначена* $P_{nX}^*(G)$ *, тоді говорять, що задано розподіл статистичних ймовірностей (відносних частот) на множині* Ω_X *значень* S/S_X *-випадкової величини X .*

Такий розподіл може бути поточковим на скінченній множині $\Omega_X = \{x_1, x_2, ..., x_k\}$ чи нескінченній множині $\Omega_X = \{x_1, x_2, x_3, ...\}$, поінтервальним на неперервній обмеженій множині виду $\Omega_X = [a;b]$, $-\infty < a < b < \infty$, чи неперервним або мішаним (див. Розділ II, Розділ III).

У випадку дискретного поточкового розподілу статистичних ймовірностей на множині Ω_X значень S/S_X-випадкової величини $X(E)$, $E \in \Omega$, цей розподіл можна описати за допомогою ряду розподілу статистичних ймовірностей чи функції $F_{nX}^*(x)$ дискретного поточкового розподілу статистичних ймовірностей.

Зауважимо, що у випадку, коли множина $\Omega_X = \{x_1, x_2, ..., x_k\}$, де $x_1 < x_2 < ... < x_k$, скінченна, але число k дуже велике, наприклад $k = 10^{1000\,000}$, а $P_{nX}^*(\{x_i\}) = P_n^*(X^{-1}(\{x_i\}))$ виявляються рівними між собою або досить близькими, то на практиці доцільно вважати, що $\Omega_X = [x_1; x_k + \varepsilon)$, де $\varepsilon > 0$ досить мале число, і поділити Ω_X на деяку практично прийнятну кількість m підмножин $H_i = [a_{i-1}; a_i)$, $a_i - a_{i-1} = h > 0$, $i \in \overline{1,m}$, $a_0 = x_1$, $a_m = x_k + \varepsilon$, побудувати сукупність $\tilde{S}_X = \{G \mid G = \bigcup_{i \in I} H_i,\ I \subset \{1,2,...,m\}\}$, підмножин множини Ω_X, породжену за підмножинами H_i, $i \in \overline{1,m}$, і наближено вважати, що на множині значень $S/\tilde{S}_X$-випадкової величини X задано поінтервальний розподіл узагальнених статистичних ймовірностей $P_n^*(H_i)$, $i \in \overline{1,m}$, причому $P_n^*(H_i)$ рівні між собою за всіх $i \in \overline{1,m}$.

У випадку поінтервального розподілу статистичних ймовірностей на множині Ω_X значень S/S_X-випадкової величини $X(E)$, $E\in\Omega$, цей розподіл можна описати за допомогою щільності $f_X(x)$ чи функції $F_X(x)$ поінтервального розподілу статистичних ймовірностей (див. §2.1 - §2.7).

Як і раніше, можна обчислити деякі числові характеристики поточкового чи поінтервального, чи неперервного, в разі якщо $S_X=\mathscr{B}([a_0;a_m))$ – система борелівських підмножин із проміжка $[a_0;a_m)$, розподілу статистичних ймовірностей на множині Ω_X значень S/S_X-випадкової величини $X(E)$, $E\in\Omega$.

Коли розподіл статистичних ймовірностей дискретний, тоді величину $m_{nX}^*=\frac{1}{n}\sum_{i=1}^{n}x_{cn\,i}=\sum_{i=1}^{k}x_iP_{nX}^*(\{x_i\})$ називають центром розсіювання статистичних ймовірностей на множині Ω_X значень S/S_X-випадкової величини $X(E)$, $E\in\Omega$, або ж *статистичним математичним сподіванням* S/S_X*-випадкової величини* $X(E)$, $E\in\Omega$, яке позначають також через $M_n^*[X]$. Відповідно, величину $D_n^*[X]=\sum_{i=1}^{k}(x_i-M_n^*[X])^2P_{nX}^*(\{x_i\})$ називають *дисперсією розподілу статистичних ймовірностей на множині* Ω_X *значень* S/S_X*-випадкової величини* X.

Коли задано щільність $f_X(x)$ розподілу узагальнених статистичних ймовірностей на множині Ω_X значень S/S_X-випадкової величини $X(E)$, $E\in\Omega$, тоді $M_n^*[X]$ і $D_n^*[X]$ визначають за формулами

$$M_n^*[X]=\int_{-\infty}^{\infty}x\,f_X(x)\,dx,$$

$$D_n^*[X]=\int_{-\infty}^{\infty}(x-M_n^*[X])^2\,f_X(x)\,dx.$$

Приклад 4.2.1. Нехай задано ймовірнісний простір (Ω,S,P_n^*) і X – індикатор (або характеристична випадкова величина) події $A\in S$, тобто

$$X(E)=\begin{cases}1, & \text{коли } E\in A,\\ 0, & \text{коли } E\in\overline{A},\end{cases}$$

Тоді $\Omega_X=\{0,1\}$. Як сукупність S_X підмножин множини Ω_X розглядатимемо таку:

$$S_X = \{G \mid G \subset \Omega_X \ , X^{-1}(G) \in S\} = \{\varnothing, \{0\}, \{1\}, \{0,1\}\}\,.$$

Очевидно $X(E)$, $E \in \Omega$, S/S_X-вимірна функція. В даному разі ймовірнісна міра $P^*_{nX}(G)$, $G \in S_X$, на елементах сукупності S_X набуває значень:

$P^*_{nX}(\varnothing) = P^*_n(X^{-1}(\varnothing)) = P^*_n(\varnothing) = 0$,

$P^*_{nX}(\{0\}) = P^*_n(X^{-1}(\{0\})) = P^*_n(\overline{A})$,

$P^*_{nX}(\{1\}) = P^*_n(X^{-1}(\{1\})) = P^*_n(A)$,

$P^*_{nX}(\{0,1\}) = P^*_n(X^{-1}(\{0,1\})) = P^*_n(X^{-1}(\{0\}) \bigcup X^{-1}(\{1\})) =$

$= P^*_n(\overline{A} \bigcup A) = P^*_n(\Omega) = 1$.

Ряд розподілу статистичних ймовірностей на множині значень S/S_X-випадкової величини X в розглядуваному випадку набуває вигляду

x_i	0	1
$P^*_{nX}(\{x_i\})$	$P^*_n(\overline{A})$	$P^*_n(A)$

Функція розподілу узагальнених статистичних ймовірностей на множині $\tilde{\Omega}_X = R^1 \supset \Omega_X$, $\tilde{S}_X = \mathscr{B}(R^1) \supset S_X$, до якої входить множина Ω_X значень розглядуваної S/S_X-випадкової величини X (за умови $\Omega_X = R^1,\ S_X = \mathscr{B}(R^1)$) набуває вигляду (див. §2.1 - §2.7)

$$F_X(x) = P^*_{nX}((-\infty; x) \bigcap \Omega_X) = \sum_{x_i < x} P^*_{nX}(\{x_i\}) =$$

$$= \begin{cases} 0, \text{ коли } x \le 0, \\ P^*_{nX}(\{0\}) = P^*_n(\overline{A}), \text{ коли } 0 < x \le 1, \\ P^*_{nX}(\{0\}) + P^*_{nX}(\{1\}) = P^*_n(\overline{A}) + P^*_n(A) = 1, \text{ коли } 1 < x. \end{cases}$$

Графік функції $F_X(x)$ розподілу статистичних ймовірностей на множині значень розглядуваної S/S_X-випадкової величини X подано на Рис. 4.2.1.

Координата $x_c = m^*_{nX}$ центра розсіювання статистичних ймовірностей на множині значень розглядуваної S/S_X-випадкової величини X (статистичне математичне сподівання $M^*_n[X]$ або середнє значення S/S_X-випадкової величини X) дорівнює:

$$x_c = M^*_n[X] = 0 \cdot P^*_n(\overline{A}) + 1 \cdot P^*_n(A) = P^*_n(A)\,.$$

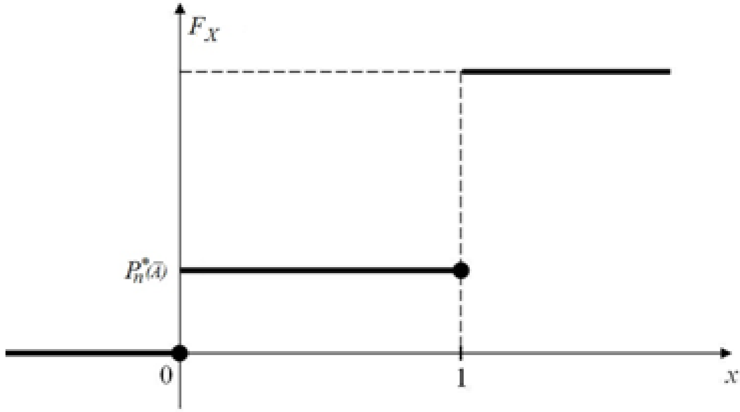


Рис. 4.2.1

Дисперсія розподілу статистичних ймовірностей на множині значень розглядуваної S/S_X-випадкової величини X дорівнює

$$D_n^*[X] = \sum_{i=1}^{k} (x_i - M_n^*[X])^2 P_{nX}^*(\{x_i\}),$$

тобто

$$D_n^*[X] = (0 - P_n^*(A))^2 \cdot P_n^*(\overline{A}) + (1 - P_n^*(A))^2 \cdot P_n^*(A) = (P_n^*(A))^2 \cdot P_n^*(\overline{A}) +$$
$$+ (P_n^*(\overline{A}))^2 \cdot P_n^*(A) =$$
$$= P_n^*(A) \cdot P_n^*(\overline{A})(P_n^*(A) + P_n^*(\overline{A})) = P_n^*(A) P_n^*(\overline{A}).$$

Надалі, якщо не виникатиме непорозумінь чи неоднозначностей, іноді S/S_X-випадкову величину називатимемо просто випадковою величиною, опускаючи символи S/S_X-, разом з тим щоразу маючи на увазі S/S_X-випадкові величини стосовно відповідних сукупностей S і S_X підмножин множин Ω і $\Omega_X = X(\Omega)$ відповідно, стосовно кожної з яких задовільняються вимоги 1_s-3_s.

Приклад 4.2.2. Нехай $\Omega = [-1;1) \subset R^1$ поділено на досить велику (однак практично прийнятну) кількість $2k$ досить дрібних проміжків $H_i = [a_{i-1}; a_i)$, однакової міри (довжини) $m(H_i) = \dfrac{1}{k}$, де $a_{i-1} = -1 + \dfrac{1}{k}(i-1)$, $i \in \overline{1, 2k}$. Як простір подій S розглядатимемо всеможливі об'єднання $\bigcup_{i \in I} H_i$, $I \subset \{1, 2, ..., 2k\}$ разом з порожньою множиною, тобто

$S=\{A \mid \ A=\bigcup_{i\in I} H_i,\ I\subset\{1,\ 2,\ \dots,\ 2k\}\}$, $\qquad \varnothing\in S$, $\qquad \bigcup_{i\in I} H_i \in S$,

$I\subset\{1,2,...,2k\}$, $\Omega=\bigcup_{i=1}^{2k} H_i \in S$. Нехай проведено досить велику серію із n випробувань, в результаті яких виявилось $P_n^*(H_i)=\dfrac{1}{2k}$ і тому

$$f_n^*(E)=\frac{P_n^*(H_i)}{m(H_i)}=\frac{1}{2},\ E\in H_i,$$

за всіх $i\in\overline{1,2k}$ (див. §2.2-2.4), тобто $f_n^*(E)=\dfrac{1}{2}$, $E\in[\,-1\,;1\,]$.

Тоді можна вважати, що ймовірнісна міра P_n^* задана на S через щільність розподілу статистичних ймовірностей відносно геометричної міри (довжини) підмножин множини Ω (складених із проміжків H_i):

$$f(E)=\begin{cases}\dfrac{1}{2}, \text{ коли } E\in[-1;1],\\ 0, \text{ коли } E\,\overline{\in}\,[-1;1].\end{cases}$$

В розглядуваному випадку узагальнені статистичні ймовірності попадання в підмножини (складені із проміжків H_i) $\Delta_i=\bigcup_{H_i\subset\Delta_i} H_i$ і $\Delta_j=\bigcup_{H_i\subset\Delta_j} H_i$ однакової геометричної міри (довжини)

$$m(\Delta_i)=\sum_{H_i\subset\Delta_i} m(H_i)=m(\Delta_j)=\sum_{H_i\subset\Delta_j} m(H_i)$$

виявляються однаковими

$$P(\Delta_i)=\int_{\Delta_i} f(E)P(dE)=\frac{1}{2}m(\Delta_i)=$$

$$=P(\Delta_j)=\int_{\Delta_j} f(E)P(dE)=\frac{1}{2}m(\Delta_j)$$

Це означає, що розподіл узагальнених статистичних ймовірностей на множині $\Omega=[-1;1]$ за проміжками H_i є рівномірним. Графік функції $f(E)$ подано на Рис. 4.2.2.

Нехай на множині $\Omega=[-1;1]$ задано функцію $X(E)$, $E\in\Omega$: $X(E)=1-|E|$, $E\in[-1;1]$, графік якої подано на Рис. 4.2.3, і нехай $\Omega_X=X(\Omega)$ – образ множини Ω за відображення X, $H_{iX}=X(H_i\bigcup H_{2k-i+1})$ – образи підмножин $(H_i\bigcup H_{2k-i+1})\subset\Omega$, $i\in\overline{1,k}$, $H_i\bigcap H_{2k-i+1}=\varnothing$. Тут $X^{-1}(H_{iX})=H_i\bigcup H_{2k-i+1}$, $i\in\overline{1,k}$.

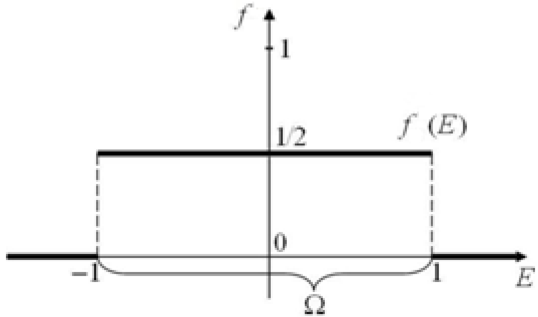


Рис. 4.2.2

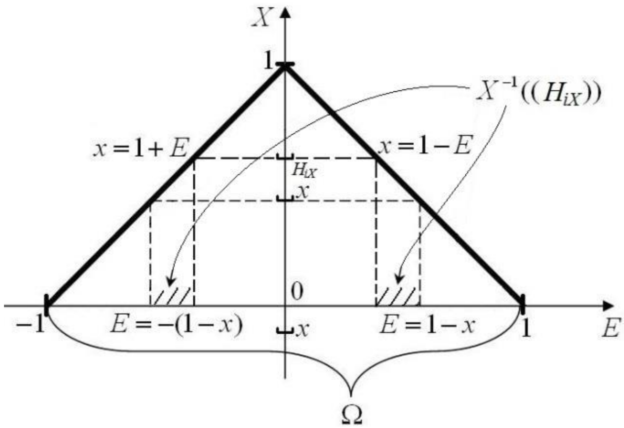


Рис. 4.2.3

Очевидно

$$\Omega_X = X(\Omega) = X\left(\bigcup_{i=1}^{k}(H_i \cup H_{2k-i+1})\right) = \bigcup_{i=1}^{k} X(H_i \cup H_{2k-i+1}),$$

тобто

$$\Omega_X = \bigcup_{i=1}^{k} X(H_i \cup H_{2k-i+1}) = \bigcup_{i=1}^{k} H_{iX} ,$$

причому $X(H_i \cup H_{2k-i+1}) \bigcap X(H_j \cup H_{2k-j+1}) = \varnothing$, коли $i \neq j$.

Розглянемо сукупність S_X підмножин множини Ω_X, до якої разом з порожньою множиною $\varnothing$ віднесемо всеможливі об’єднання $\bigcup_{i \in I} H_{iX}$, $I \subset \{1, 2, ..., k\}$, підмножин $H_{iX} \subset \Omega_X$, $i \in \overline{1,k}$, тобто $S_X = \{G \mid G = \bigcup_{i \in I} H_{iX}, I \subset \{1, 2, ..., k\}\}$. Оскільки за довільного $I \subset \{1, 2, ..., k\}$

$$X^{-1}\left(\bigcup_{i\in I} H_{iX}\right)=\bigcup_{i\in I} X^{-1}(H_{iX})=\bigcup_{i\in I}(H_i\cup H_{2k-i+1})\in S,$$

то функція $X(E)$ за так заданих S і S_X виявляється S/S_X-вимірною, тобто S/S_X-випадковою величиною.

В даному разі за довільного $G=\bigcup_{i\in I} H_{iX}\in S_X$, $I\subset\{1,2,...,k\}$, природно покласти (див. §2.2, §2.3)

$$P_{nX}^*(G)=P_{nX}^*\left(\bigcup_{i\in I} H_{iX}\right)=P_{nX}^*\left(X^{-1}\left(\bigcup_{i\in I} H_{iX}\right)\right)=P_{nX}^*(X^{-1}(G))=$$

$$=P_n^*\left(\bigcup_{i\in I} X^{-1}(H_{iX})\right)=P_n^*\left(\bigcup_{i\in I}\left(H_i\cup H_{2k-i+1}\right)\right)=\sum_{i\in I} P_n^*(H_i\cup H_{2k-i+1})=$$

$$=\sum_{i\in I}\int_{H_i\cup H_{2k-i+1}} f(E)m(dE),$$

$$f_X(x)=\begin{cases}\dfrac{P_{nX}^*(H_{iX})}{m(H_{iX})}, & \text{коли } x\in H_{iX},\ i\in\overline{1,k},\\[2ex] 0, & \text{коли } x\,\overline{\in}\,\Omega_X=\bigcup\limits_{i=1}^{k} H_{iX}.\end{cases}$$

Оскільки за заданого відображення $X(E)=1-|E|$, $E\in\Omega=[-1;1]$, буде $\Omega_X=X(\Omega)=[0;1]$, прообраз $X^{-1}(H_{iX})$ проміжка $H_{iX}\subset\Omega_X$ міри (довжини) $m(H_{iX})$ складається з двох відповідних проміжків H_i, кожен тієї самої міри (довжини), що і H_{iX}, за яких $X(H_i\cup H_{2k-i+1})=H_{iX}$, причому за заданого відображення $m(H_{iX})=m(H_i)$, $m(X^{-1}(H_{iX}))=m(H_i\cup H_{2k-i+1})==2\cdot m(H_i)$, то статистична ймовірність попадання в проміжок H_{iX} вдвічі більша, ніж статистична ймовірність попадання в проміжок H_i, а тому

$$f_X(x)=\begin{cases}1, & \text{коли } x\in[0;1],\\ 0, & \text{коли } x\,\overline{\in}\,[0;1].\end{cases}$$

Графік функції $f_X(x)$ подано на Рис. 4.2.4.

Якщо тепер вважати, що образ множини Ω є $\tilde{\Omega}_X=R^1$, покладаючи $X^{-1}(x)=\varnothing$, якщо число $x\in R^1$ не поставлено у відповідність жодному елементові $E\in\Omega$, а як $\tilde{S}_X$ розглядати сукупність $\mathscr{B}(R^1)$ борелівських підмножин із R^1 (див. §2.6),

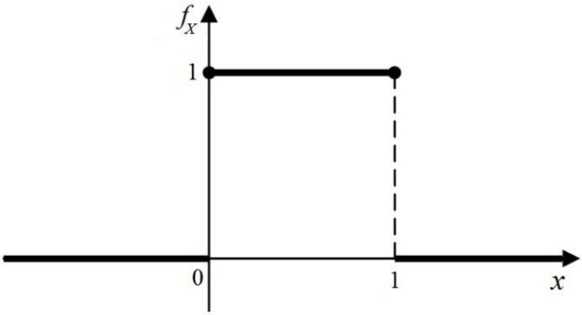


Рис. 4.2.4

тоді очевидно буде $S_X \subset \tilde{S}_X$.

Продовжимо міру P^*_{nX} із S_X на $\tilde{S}_X$, поклавши за довільного $G = (-\infty;\, x) \in \tilde{S}_X$ (див. §2.2-§2.7)

$$P_X(G) = P_X(G \cap \Omega_X) = P_X\left(G \cap \left(\bigcup_{i=1}^{k} H_{iX}\right)\right) =$$

$$= P_X\left(\bigcup_{i=1}^{k}(G \cap H_{iX})\right) = \sum_{i=1}^{k} P_X(G \cap H_{iX}) = \int_G f_X(x)dx\,.$$

Тоді за довільного $G \in S_X$ одержимо

$$P_X(G) = P^*_{nX}(G)\,.$$

Таким чином можна знайти функцію $F_X(x) = P_X((-\infty;\, x))$ розподілу узагальнених статистичних (гіпотетичних) ймовірностей на множині $\tilde{\Omega}_X = X(\Omega)$, $(-\infty;\, x) \overline{\in} S_X$, $(-\infty;\, x) \in \tilde{S}_X = \mathscr{B}(R^1)$ (див. §2.4-§2.7, §3.2-§3.5). В розглядуваному прикладі одержимо:

$$F_X(x) = P_X((-\infty;\, x)) = P_X\left(\bigcup_{i=1}^{k}((-\infty;\, x) \cap H_{iX})\right) =$$

$$= \sum_{i=1}^{k} P_X((-\infty;x) \cap H_{iX}) = \int_{(-\infty,x)} f_X(x)dx = \begin{cases} 0, \text{ коли } x \le 0, \\ x, \text{ коли } 0 \le x \le 1, \\ 1, \text{ коли } 1 \le x. \end{cases}$$

Справді, оскільки $X^{-1}((-\infty;\, x)) = \varnothing$, коли $x \le 0$, бо від'ємні значення x не поставлені у відповідність ніяким точкам E із $\Omega = [-1;1]$, то за $x \le 0$ буде

$$P_X((-\infty;\, x) \cap \bigcup_{i=1}^{k} H_{iX}) = P(X^{-1}((-\infty;\, x) \cap \bigcup_{i=1}^{k} H_{iX})) = P(\varnothing) = 0;$$

коли $0 < x \le 1$, тоді

$$X^{-1}((-\infty;x))=[-1;-(1-x))\bigcup(1-x;1],$$

$$m(X^{-1}((-\infty;x)))=m([-1;-(1-x))\bigcup(1-x;1])=m([-1;-(1-x))+$$

$$+m((1-x;1])=(-(1-x)-(-1))+(1-(1-x))=x+x=2x,$$

тому, коли $0<x\le 1$, одержуємо

$$P_X((-\infty;x)\bigcap\bigcup_{i=1}^{k}H_{iX})\approx\int\limits_{X^{-1}((-\infty;x))\bigcap X^{-1}\left(\bigcup_{i=1}^{k}H_{iX}\right)}f(x)dx=$$

$$=\frac{1}{2}\cdot m\left(X^{-1}\left((-\infty;x)\bigcap X^{-1}\left(\bigcup_{i=1}^{k}H_{iX}\right)\right)\right)=$$

$$=\frac{1}{2}m(([-1;-(1-x))\bigcup(1-x;1])\bigcap[-1;1])=\frac{1}{2}\cdot 2x=x;$$

коли $x>1$, тоді $X^{-1}((-\infty;x))=[-1;1]=\Omega$, тому за таких x $P_X((-\infty;x))=1$.

Графік функції $F_X(x)$ подано на Рис. 4.2.5.

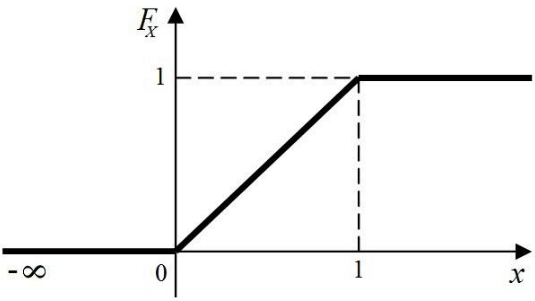


Рис. 4.2.5

Координата $x_c=m^*_{nX}$ центра розсіювання статистичних ймовірностей на множині $\Omega_X=[0;1]$ значень розглядуваної випадкової величини X (статистичне математичне сподівання $M^*_n[X]$ або середнє значення випадкової величини X) дорівнює

$$x_c=M^*_n[X]=\int\limits_{\Omega_X}xf_X(x)dx=\int\limits_{[0;1]}x\cdot 1\cdot dx=1\cdot\int\limits_0^1 xdx=\left.\frac{x^2}{2}\right|_0^1=\frac{1}{2}.$$

Дисперсія розподілу статистичних ймовірностей на множині Ω_X значень розглядуваної випадкової величини X дорівнює

$$D_n^*[X] = \int\limits_{\Omega_X} (x - M_n^*[X])^2 f_X(x) dx,$$

тобто

$$D_n^*[X] = \int\limits_0^1 \left(x - \frac{1}{2}\right)^2 \cdot 1 \cdot dx = \frac{1}{3}\left(x - \frac{1}{2}\right)^3 \Bigg|_0^1 = \frac{1}{3}\left(1 - \frac{1}{2}\right)^3 - \frac{1}{3}\left(0 - \frac{1}{2}\right)^3 = \frac{1}{12}.$$

Зауважимо, що як Ω_X завжди можна розглядати множину $(-\infty;\infty)$, не дивлячись на те, що статистичні ймовірності попадання значень випадкової величини X за межі деякого обмеженого проміжка $[a;b] \subset (-\infty;\infty) = \Omega_X$ дорівнюють нулеві.

Якщо розподіл статистичних ймовірностей неперервний і заданий через щільність $f_X(x)$, рівну нулеві за межами деякого проміжка $[a;b]$, то

$$P_X((-\infty;\infty)) = \int\limits_{-\infty}^{\infty} f_X(x)dx = \int\limits_a^b f_X(x)dx = P_{nX}^*([a;b]) = 1,$$

$$M_n^*[X] = \int\limits_{-\infty}^{\infty} x\, f_X(x)dx = \int\limits_a^b x\, f_X(x)dx \text{ і т.д.}$$

Якщо ж множина $\Omega_X = \{x_1, x_2, ..., x_k\}$ значень випадкової величини X скінченна, то

$$P_X((-\infty;\infty)) = \sum_{x_i \in (-\infty;\, \infty)} P_{nX}^*(\{x_i\}) = \sum_{x_i \in \{x_1, \dots, x_n\}} P_{nX}^*(\{x_i\}) = 1,$$

$$M_n^*[X] = \sum_{x_i \in \{x_1, \dots, x_n\}} x_i\, P_{nX}^*(\{x_i\}).$$

Якщо множина Ω_X обмежена, а k дуже велике (див. приклад 1.8.2), тоді на практиці доцільно розглядати множину Ω_X як неперервну і наближено обчислювати числові характеристики розподілу статистичних ймовірностей на множині Ω_X значень випадкової величини X за правилами, за якими обчислюють числові характеристики розподілів статистичних ймовірностей на неперервних множинах точок, попередньо визначивши щільність $f_X(x)$ розподілу статистичних ймовірностей на множині Ω_X значень випадкової величини X.

Надалі, якщо не буде сказано інше, матимемо на увазі $S/\mathscr{B}(R^1)$-випадкові величини, опускаючи позначення $S/\mathscr{B}(R^1)$ відповідних просторів подій.

Нехай $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$, ймовірнісна міра P_X на $(R^1, \mathscr{B}(R^1))$ визначається за функцією розподілу

$$F_X(x) = P_X((-\infty;x)) = P(X^{-1}((-\infty;x))).$$

Стосовно функції $F_X(x)$ розподілу ймовірностей на множині значень випадкової величини X задовільняються такі самі властивості, як і стосовно будь якої іншої функції одновимірного розподілу ймовірностей (див. Розділ 2).

Якщо множина значень функції розподілу $F_X(x)$ скінченна або зчисленна, то розподіл ймовірностей на множині Ω_X випадкової величини X може бути поточковим або поінтервальним (див.§2.1, §2.3, §2.5, §2.6, §2.7). В разі поінтервального розподілу ймовірностей розглядається неперервна множина Ω_X значень випадкової величини X типу $\Omega_X=[a;b)$ (чи типу $\Omega_X=(a;b]$), яка ділиться на k інтервалів $H_{Xi}=[a_{i-1};a_i)$ (чи $H_{Xi}=(a_{i-1};a_i]$), $i\in 1,2,...,k$, таких, що $a_0=a$, $a_k=b$, $a_i-a_{i-1}=h=\dfrac{b-a}{k}$, стосовно кожного з яких визначена ймовірнісна міра $P_X([a_{i-1};a_i))$ (чи $P_X((a_{i-1};a_i])$), а також щільність поінтервального розподілу ймовірностей на множині Ω_X за інтервалами H_{Xi}, $i\in\overline{1,k}$,

$$f_X(x)=\begin{cases}\dfrac{P_X([a_{i-1};a_i))}{h}, & \text{коли } x\in[a_{i-1};a_i),\ i\in\overline{1,k},\\ 0, & \text{коли } x\,\overline{\in}\,[a_0;a_k).\end{cases}$$

Слід зазначити, що хоч функції розподілу $F_X(x)$ за деяких поточкових і поінтервальних розподілів можуть виявитись однаковими (див. §2.7), різниця між такими розподілами ймовірностей суттєва.

В разі, коли множина Ω_X значень випадкової величини X дискретна (скінченна або зчисленна), тоді міра P_X буде зосереджена на не більш ніж зчисленній множині точок $x_1, x_2, \ldots$ і набуватиме вигляду

$$P_X(B)=\sum_{x_k\in B}P_X(\{x_k\})=\sum_{x_k\in B}P(X^{-1}(x_k)),$$

де

$$P_X(\{x_k\})=P(X^{-1}(x_k))=\Delta F_X(x_k)=F_X(x_k+0)-F_X(x_k).$$

В такому разі випадкову величину $X(E)$ можна подати у вигляді

$$X(E)=\sum_{i=1}^{\infty}x_i I_{X^{-1}(x_i)}(E),$$

де

$$I_{X^{-1}(x_i)}(E)=\begin{cases}1, & \text{коли } E\in X^{-1}(x_i),\\ 0, & \text{коли } E\,\overline{\in}\,X^{-1}(x_i),\end{cases}$$

індикаторна функція множини $X^{-1}(x_i)$.

Якщо функція розподілу $F_X(x)$ неперервна, то розподіл ймовірностейна множині Ω_X значень випадкової величини X називається *неперервним* (див. §3.2). Зрозуміло, що функція $F_X(x)$ може бути неперервною лише у випадку, коли множина Ω_X значень випадкової величини X неперервна.

Розподіл ймовірностей на множині Ω_X значень випадкової величини X називається *абсолютно неперервним* (див. §3.2), якщо функція розподілу $F_X(x)$ абсолютно неперервна, тобто якщо існує невід'ємна борелівська функція $f_X(x)$ така, що

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt\ ,\ \ x \in R^1 .$$

Цю функцію $f_X(x)$ називають *щільністю* розподілу ймовірностей на множині $\Omega_X = R_1$ значень випадкової величини X (в загальному випадку мають на увазі інтеграл Лебега).

Якщо функція $f_X(x)$ неперервна, то функція

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

диференційовна і

$$\frac{d}{dx} F_X(x) = f_X(x) .$$

В загальному випадку з рівності $F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$ випливає, що рівність $\frac{d}{dx} F_X(x) = f_X(x)$ виконується майже всюди стосовно міри Лебега, тобто

$$m\{x \,|\, \frac{d}{dx} F_X(x) \text{ не існує або } \frac{d}{dx} F_X(x) \neq f_X(x)\} = 0 .$$

Слід зазначити, що за похідною $\frac{d}{dx} F_X(x)$ можна відновити тільки абсолютно неперервну складову функції $F_X(x)$, оскільки дискретна й сингулярна складові під час диференціювання «безслідно зникають».

Нехай на ймовірнісному просторі (Ω, S, P) визначено дискретну випадкову величину X із скінченною множиною можливих значень $\Omega_X = X(\Omega) = \{x_1, x_2, \ldots, x_m\}$. До сукупності S_X включимо всі підмножини множини Ω_X разом з порожньою.

Кожному із можливих значень x_i випадкової величини X поставимо у відповідність подію $\{x_i\}$. Очевидно, що

$$P_X(\{x_i\}) = P(X^{-1}(x_i)),$$

$$X^{-1}(x_i) \cap X^{-1}(x_j) = \varnothing, \text{ коли } x_i \neq x_j,$$

$$P_X(\Omega_X) = P_X(\{x_1\} + \{x_2\} + \ldots + \{x_m\}) =$$

$$= \sum_{i=1}^{m} P_X(\{x_i\}) = P(X^{-1}(X(\Omega))) = P(\Omega) = 1.$$

Очевидно (Ω_X, S_X, P_X) є ймовірнісним простором.

Таблицю видуТабл. 4.2.1

Табл. 4.2.1

x_i	x_1	x_2	$\ldots$	x_m
$P_X(\{x_i\})$	$P_X(\{x_1\})$	$P_X(\{x_2\})$	$\ldots$	$P_X(\{x_m\})$

називають *рядом поточкового розподілу* ймовірностей на дискретній множині Ω_X значень випадкової величини X.

В механічній інтерпретації ряд поточкового розподілу ймовірностей є аналогом розподілу одиничної маси вздовж осі абсцис так, що на точку з абсцисою x_i припадає маса $P_X(\{x_i\})$, причому $\sum_{i=1}^{m} P_X(\{x_i\}) = 1$.

Ряд поточкового розподілу ймовірностей на множині $\Omega_X = \{x_1, x_2, ..., x_m\}$ можна подати графічно. В такому разі вздовж осі абсцис відкладають можливі значення x_i, $i = 1, 2, \ldots, m$, випадкової величини X, а вздовж осі ординат проти кожної абсциси – відповідні ймовірності (Рис. 4.2.6).

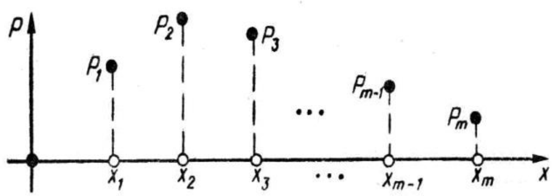


Рис. 4.2.6

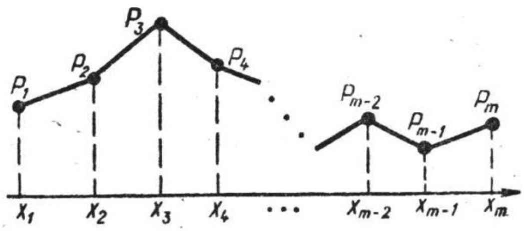


Рис. 4.2.7

Якщо точки $(x_1, P_X(\{x_1\}), \ldots, (x_m, P_X(\{x_m\})$ нанести на координатну площину і сполучити відрізками прямих, то одержиться так званий *многокутник розподілу ймовірностей* (Рис. 4.2.7), який є однією з форм описування розподілу ймовірностей на множині значень одновимірної випадкової величини із скінченною множиною значень (див. §2.1).

Приклад 4.2.3. Нехай випробування полягає в тому, що в деякому квадраті G, довжина сторони якого 22, навмання вибирається точка. Кожній точці квадрата поставимо у взаємно однозначну відповідність елементарну подію E. Тоді множина точок квадрата ототожнюється з множиною Ω елементарних подій. Кожній точці квадрата, віддаленій від його центра C не більше, ніж на 1, поставимо у відповідність число 10 (Рис. 4.2.8); кожній точці квадрата, віддаленій від центра більше ніж на 1, але не більше ніж на 2, поставимо у відповідність число 9; точкам, віддаленим від центра більше ніж на 2, але не більше ніж на 3, поставимо у відповідність число 8 і т.д.; точкам, віддаленим більше ніж на 9, але не більше ніж на 10, поставимо у відповідність число 1; нарешті, точкам, віддаленим більше ніж на 10, поставимо у відповідність число 0. В такий спосіб задано відображення X двохвимірної нескінченної і неперервної множини Ω на одновимірну скінченну множину:

$$\Omega_X = X(\Omega) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}.$$

Тут

$$X^{-1}(10) = \{M \mid M \in G,\ 0 \le \rho(M, C) \le 1\};$$

$$X^{-1}(9) = \{M \mid M \in G,\ 1 \le \rho(M, C) \le 2\};$$

....

$$X^{-1}(0) = \{M \mid M \in G,\ 10 \le \rho(M, C)\},$$

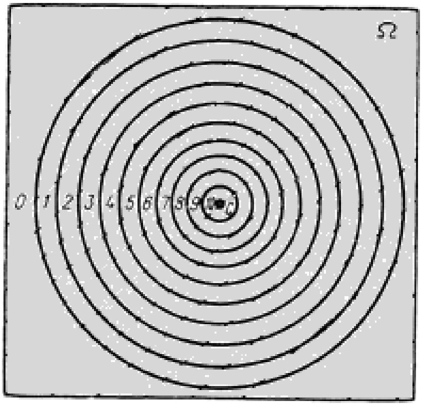


Рис.4.2.8

де через $\rho(M,C)$ позначено відстань між точками M і C.

Як події із простору подій S ймовірнісного простору (Ω, S, P) розглядатимемо вимірні за Лебегом підмножини множини Ω, а ймовірнісну міру P множини $A \in S$, $A \subset \Omega$, задамо як $P(A) = \dfrac{m(A)}{m(\Omega)}$ (геометричне задання імовірнісної міри) (див. §3.6). Тоді згідно з геометричним заданням імовірності,

$$P_X(\{i\}) = P(X^{-1}(i)) = \frac{m(X^{-1}(i))}{m(G)}.$$

Обчисливши останні ймовірності за всіх можливих значень $i \in \Omega_X = X(\Omega) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ і звівши їх у відповідну таблицю, дістанемо ряд поточкового розподілу ймовірностей на множині значень розглядуваної випадкової величини (Табл. 4.2.2)

Табл. 4.2.2

x_i	0	1	2	3	4
$P_X(\{x_i\})$	$\frac{22^2-\pi 10^2}{22^2}$	$\frac{\pi 10^2-\pi 9^2}{22^2}$	$\frac{\pi 9^2-\pi 8^2}{22^2}$	$\frac{\pi 8^2-\pi 7^2}{22^2}$	$\frac{\pi 7^2-\pi 6^2}{22^2}$

5	6	7	8	9	10
$\frac{\pi 6^2-\pi 5^2}{22^2}$	$\frac{\pi 5^2-\pi 4^2}{22^2}$	$\frac{\pi 4^2-\pi 3^2}{22^2}$	$\frac{\pi 3^2-\pi 2^2}{22^2}$	$\frac{\pi 2^2-\pi 1^2}{22^2}$	$\frac{\pi 1^2}{22^2}$

Приклад 4.2.4. Випробування полягає в одночасному підкиданні трьох монет. Множина Ω елементарних подій складається з восьми рівноймовірних елементів: E_1 – ГГГ, E_2 – ГГЦ, E_3 – ГЦГ, E_4 – ГЦЦ, E_5 – ЦГГ, E_6 – ЦГЦ, E_7 – ЦЦГ, E_8 –

ЦЦЦ (Г – монета впала догори гербом, Ц – цифрою). Причому $P(E_i)=\frac{1}{8}$ за всіх $i\in\{1,\ 2,\ ...,\ 8\}$ (як імовірності добутків із трьох незалежних в сукупності подій, ймовірність кожної з яких дорівнює $\frac{1}{2}$).

Кожній елементарній події поставимо у відповідність кількість випадань герба на трьох монетах. Тоді (Рис. 4.2.9) $X(E_1)=3$, $X(E_2)=2$, $X(E_3)=2$, $X(E_4)=1$, $X(E_5)=2$, $X(E_6)=1$, $X(E_7)=1$, $X(E_8)=0$, тобто $\Omega_X=X(\Omega)=\{0,\ 1,\ 2,\ 3\}$. Як елементи сукупності S будемо розглядати будь які підмножини множини Ω. Як елементи сукупності S_X також будемо розглядати будь які підмножини множини Ω_X. Тоді за властивістю адитивності ймовірності (або за формулою Бернуллі) одержимо

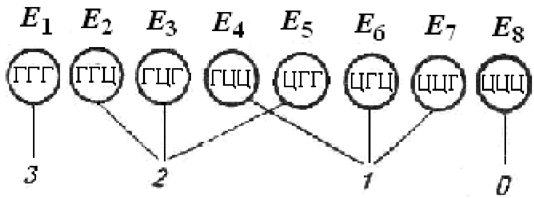


Рис. 4.2.9

$$P_X(\{0\})=P(X^{-1}(0))=P(\{E_8\})=1/8\ ,$$
$$P_X(\{1\})=P(X^{-1}(1))=P(\{E_4,\ E_6,\ E_7\})=3/8\ ,$$
$$P_X(\{2\})=P(X^{-1}(2))=P(\{E_2,\ E_3,\ E_5\})=3/8\ ,$$
$$P_X(\{3\})=P(X^{-1}(3))=P(\{E_1\})=1/8\ .$$

Приклад 4.2.5. Випробування полягає в тому, що вісім разів підкидається монета і фіксується кількість підкидань, в яких монета упала гербом догори.

Очевидно, що множиною можливих значень розглядуваної випадкової величини X є множина

$$\Omega_X=X(\Omega)=\{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8\}\ .$$

За формулою Бернуллі

$$P_X(\{x_i\})=C_8^{x_i}\left(\frac{1}{2}\right)^{x_i}\left(\frac{1}{2}\right)^{8-x_i}=\frac{C_8^{x_i}}{256}\ .$$

Надаючи x_i значень $0, 1, 2, \dots, 8$, дістанемо ряд розподілу ймовірностей (Табл. 4.2.3):

Табл. 4.2.3

x_i	0	1	2	3	4	5	6	7	8
$P_X(\{x_i\})$	$\frac{1}{256}$	$\frac{8}{256}$	$\frac{28}{256}$	$\frac{56}{256}$	$\frac{70}{256}$	$\frac{56}{256}$	$\frac{28}{256}$	$\frac{8}{256}$	$\frac{1}{256}$

Многокутник розподілу ймовірностей на множині значень даної випадкової величини зображено на Рис. 4.2.10.

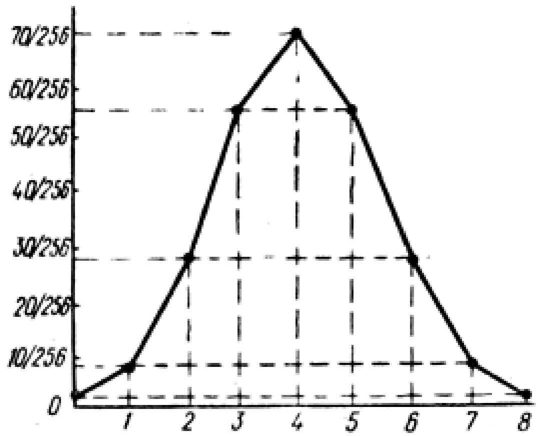


Рис. 4.2.10

Знаючи ряд поточкового розподілу ймовірностей на дискретній множині $\Omega_X = \{x_1, x_2, ..., x_k\}$ значень випадкової величини X, можна побудувати відповідну функцію розподілу ймовірностей:

$$F_X(x) = P_X((-\infty;\, x)) = \sum_{x_i \in (-\infty;\, x)} P_X(\{x_i\}).$$

В механічному тлумаченні за значенням $F_X(x)$ функції розподілу ймовірностей визначається загальна маса, що припадає на проміжок $(-\infty; x)$, за умови, що вздовж осі абсцис деяким чином розподілено одиничну масу.

Рис. 4.2.11

Наприклад, якщо вздовж осі абсцис розподілено одиничну масу так, як показано на Рис. 4.2.11, то значення $F_X(x_1)$ є маса, що припадає на проміжок $(-\infty; x_1)$; значення $F_X(x_2)$ – маса, що припадає на проміжок $(-\infty; x_2)$. Оскільки в цьому прикладі вся одинична маса розподілена на відрізку $[a; b]$, то замість проміжків $(-\infty; x_1)$, $(-\infty; x_2)$ і т.д. можна розглядати проміжки $[a; x_1)$, $[a; x_2)$ і т.д. Очевидно, щоб дістати масу, що припадає на проміжок $(-\infty; x_2)$, треба до маси, що припадає на проміжок $(-\infty; x_1)$, додати масу, що припадає на проміжок $[x_1; x_2)$.

Таким чином, за функцією розподілу $F_X(x)$ визначають, як із зростанням x накопичується (акумулюється) все більша і більша маса, яка проте обмежена зверху одиницею.

Приклад 4.2.6. Задано ряд поточкового розподілу ймовірностей на множині значень випадкової величини X (Табл. 4.2.4)

Табл. 4.2.4

x_i	1	2	3
p_i	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Побудувати відповідну функцію розподілу ймовірностей $F_X(x)$.

Нанесемо на координатну площину xOy точки (x_i, p_i), а саме $(1, \frac{1}{4})$, $(2, \frac{1}{2})$, $(3, \frac{1}{4})$ (Рис. 4.2.12).

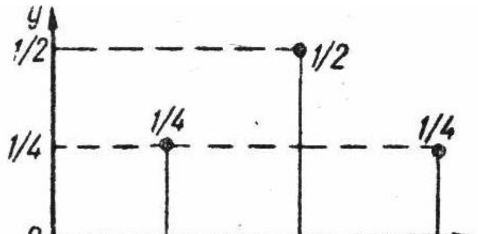


Рис. 4.2.12

Виходячи з означення функції розподілу ймовірностей, дістанемо, що за будь-якої точки x, що лежить в інтервалі $(-\infty; 1]$, буде

$$F_X(x) = 0,$$

оскільки за $x \in (-\infty; 1]$

$$F_X(x) = P_X((-\infty; x)) = \sum_{x_i \in (-\infty;\, x)} P_X(\{x_i\}) = 0,$$

(до суми не входить жоден відмінний від нуля доданок).

За довільної точки $x \in (1\,; 2]$ буде

$$F_X(x) = \frac{1}{4},$$

оскільки за таких x

$$F_X(x) = P_X((-\infty; x)) = \sum_{x_i \in (-\infty;\, x)} P_X(\{x_i\}) = P_X(\{1\}) = \frac{1}{4},$$

Якщо точка x лежить в інтервалі $(2\,; 3]$, то $F_X(x) = \dfrac{3}{4}$, оскільки за таких x буде

$$F_X(x) = P_X((-\infty; x)) = \sum_{x_i \in (-\infty;\, x)} P_X(\{x_i\}) = P_X(\{1\}) + P_X(\{2\}) = \frac{1}{4} + \frac{2}{4}.$$

Аналогічно за будь-якого x з інтервалу $(3; \infty)$ буде

$$F_X(x) = P_X((-\infty; x)) = P_X(\{1\}) + P_X(\{2\}) + P_X(\{3\}) = \frac{1}{4} + \frac{2}{4} + \frac{1}{4} = 1.$$

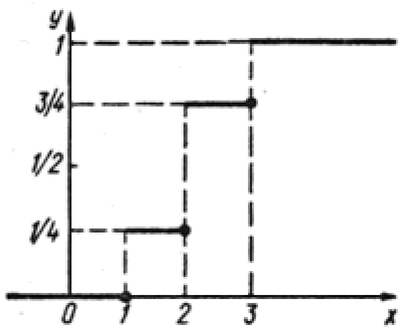


Рис.4.2.13

Таким чином

$$F_X(x) = \begin{cases} 0, & \text{коли} \quad x \le 1; \\ 1/4, & \text{коли} \quad 1 < x \le 2; \\ 3/4, & \text{коли} \quad 2 < x \le 3; \\ 1, & \text{коли} \quad 3 < x. \end{cases}$$

Графік функції $F_X(x)$ зображено на Рис. 4.2.13.

Якщо у функції розподілу ймовірностей є розриви першого роду, але множина її значень не є дискретною, то відповідний розподіл ймовірностей називається *мішаним*.

В механічній інтерпретації це означає, що частина одиничної маси розподілена вздовж осі Ox на деякому проміжку неперервно, а решта поділена на частинки, які містяться в окремих точках. Наприклад (див. Рис. 4.2.14а), $\frac{1}{2}$ одиничної маси розподілена рівномірно на відрізку $[0;1]$ із щільністю розподілу $f_1(x) = \frac{1}{2}$ і через неї визначається перша частина функції розподілу

$$F_1(x) = \begin{cases} 0, & \text{коли} \quad x \le 0; \\ (1/2)x, & \text{коли} \quad 0 < x \le 1; \\ 1/2, & \text{коли} \quad x > 1; \end{cases}$$

а ще $\frac{1}{2}$ одиничної маси міститься в точці $x = \frac{1}{2}$ і через неї визначається друга частина функції розподілу ймовірностей

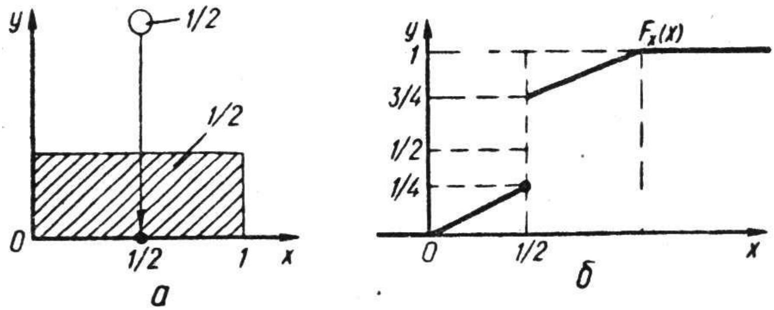


Рис. 4.2.14

$$F_2(x) = \begin{cases} 0, & \text{коли} \quad x \le \frac{1}{2}; \\ \frac{1}{2}, & \text{коли} \quad x > \frac{1}{2}. \end{cases}$$

Графік функції розподілу ймовірностей $F_X(x) = F_1(x) + F_2(x)$ в розглядуваному випадку набуває вигляду, поданого на Рис. 4.2.14б.

Одним з прикладів, коли виникає необхідність розглядати мішаний розподіл імовірностей, може бути такий. Прилад випробовується протягом часу t_0, $t_0 > 0$. Припускається, що прилад досить надійний і ймовірність його виходу з ладу в момент часу $\tau \in [0; t_0]$ невелика (середній час безвідмовної роботи приладу значно більший, ніж t_0). Випадкова величина τ – момент виходу з ладу приладу або час його безвідмовної роботи. Функція $F_\tau(t)$ розподілу ймовірностей на множині значень випадкової величини τ набуває вигляду

$$F_\tau(t) = P_\tau((-\infty; t)),$$

де $P_\tau((-\infty; t))$ ймовірність того, що відмови не трапиться в момент $\tau \in [t; +\infty)$, або, що те саме, відмова трапиться в момент $\tau \in (-\infty; t)$. Очевидно $F_\tau(t)$ неперервна скрізь, крім точки t_0, оскільки момент виходу з ладу не може наступити пізніше, ніж в момент t_0, а тому, коли $t > t_0$, тоді

$$F_\tau(t) = P_\tau((-\infty; t)) = 1.$$

За неперервного розподілу ймовірностей на множині Ω_X значень випадкової величини X ймовірність будь-якого окремого значення a дорівнює нулю:

$$P_X(\{a\}) = 0.$$

Справді,

$$P_X(\{a\}) = \lim_{b \to a} P_X([a; b)) = \lim_{b \to a} (F_X(b) - F_X(a)) = 0,$$

оскільки $F_X(x)$ неперервна функція.

Проте за неперервного розподілу ймовірностей на множині Ω_X значень випадкової величини X рівність $P_X(\{a\}) = 0$ не означає, що подія $\{a\}$ є неможливою. Така ситуація виникає і в разі геометричного задання ймовірностей. Влучення в будь-яку окрему точку мішені не є неможливою подією, але “площа” (міра Лебега) окремої точки дорівнює нулю.

Якщо розподіл ймовірностей на множині Ω_X значень випадкової величини X неперервний, то очевидно, що ймовірності $P_X([a; b))$, $P_X((a; b])$, $P_X((a; b))$, $P_X([a; b])$ рівні між собою.

Приклад 4.2.7. Розподіл імовірностей на множині Ω_X значень випадкової величини X визначається за функцією розподілу ймовірностей (див. Рис. 4.2.13)

$$F_X(x) = \begin{cases} 0, & \text{коли } x \le 1; \\ 1/4, & \text{коли } 1 < x \le 2; \\ 3/4, & \text{коли } 2 < x \le 3; \\ 1, & \text{коли } 3 < x. \end{cases}$$

Знайти ймовірність того, що значення випадкової величини X лежатиме в інтервалі [1.5; 3.5).

За формулою

$$P_X([a; b)) = F_X(b) - F_X(a) \qquad (4.2.1)$$

одержимо

$$P_X([1.5;\,3.5)) = F_X(3.5) - F_X(1.5) = 1 - \frac{1}{4} = \frac{3}{4}\,.$$

Цей результат випливає також з того, що в розглядуваному випадку подія $[1.5;\,3.5)$ відбувається, якщо відбувається принаймні одна з двох подій $(\{2\})$ або $(\{3\})$, оскільки $\{2\}\bigcup\{3\}\subset[1.5;3.5)$ (див. Рис. 4.2.12).

Приклад 4.2.8. Опис функції розподілу ймовірностей на множині Ω_X значень випадкової величини X подано у вигляді $F_X(x) = a + b\,\mathrm{arctg}\,(\frac{x}{2})$, $-\infty < x < \infty$ (розподіл Коші). Знайти сталі a і b, побудувати графік функції $F_X(x)$, знайти $P_X(X\in[\alpha;\,\beta))$.

Використовуючи властивості функції розподілу ймовірностей і враховуючи, що $\lim\limits_{x\to\infty}\mathrm{arctg}\frac{x}{2} = \frac{\pi}{2}$ і $\lim\limits_{x\to-\infty}\mathrm{arctg}\frac{x}{2} = -\frac{\pi}{2}$, дістаємо систему рівнянь

$$\begin{cases} 0 = a + b(-\frac{\pi}{2}); \\ 1 = a + b(\frac{\pi}{2}), \end{cases}$$

звідки $a = \frac{1}{2}$, $b = \frac{1}{\pi}$. Отже, $F_X(x) = \frac{1}{2} + \frac{1}{\pi}\mathrm{arctg}\frac{x}{2}$. Графік функції $F_X(x)$ зображено на Рис. 4.2.15.

За формулою (4.2.1) дістаємо

$$P_X([\alpha;\,\beta)) = F_X(\beta) - F_X(\alpha) =$$

$$= \left(\frac{1}{2} + \frac{1}{\pi}\mathrm{arctg}\frac{\beta}{2}\right) - \left(\frac{1}{2} + \frac{1}{\pi}\mathrm{arctg}\frac{\alpha}{2}\right) = \frac{1}{\pi}\left(\mathrm{arctg}\frac{\beta}{2} - \mathrm{arctg}\frac{\alpha}{2}\right).$$

Легко бачити, що опис щільності $f_X(x)$ розглядуваного розподілу ймовірностей набуває вигляду (Рис. 4.2.16):

$$f_X(x) = \frac{1}{\pi}\frac{2}{4 + x^2}\,,\ x\in(-\infty;\,\infty)\,.$$

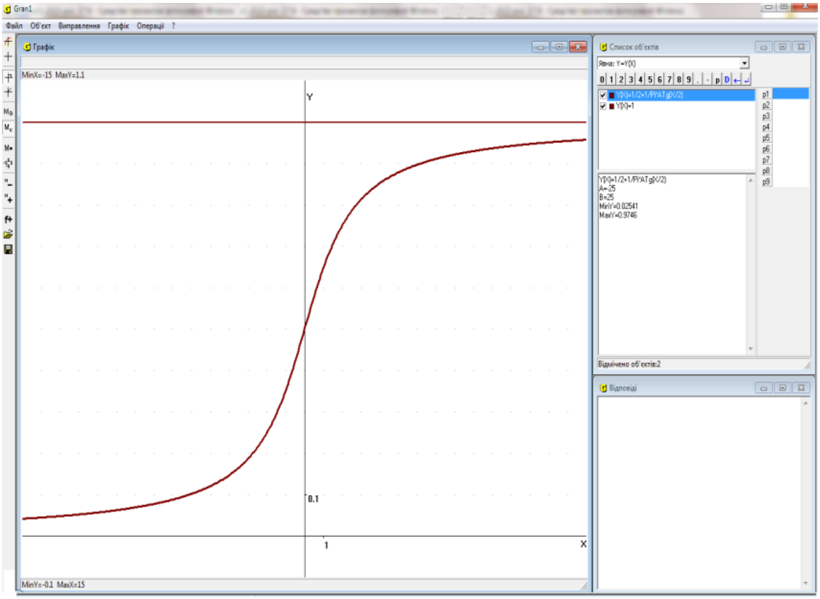


Рис. 4.2.15

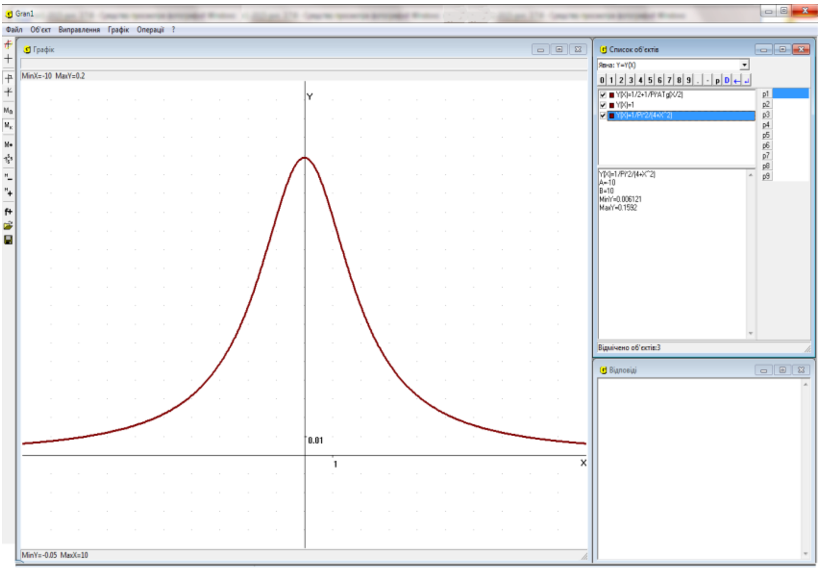


Рис. 4.2.16

Приклад 4.2.9. Знайти аналітичний вираз щільності розподілу ймовірностей, графік якої подано Рис. 4.2.17. Знайти ймовірність того, що значення випадкової величини X із заданим розподілом імовірностей лежатимуть у відрізку $[0.5; 1.5]$.

Виходячи з геометричного тлумачення властивостей щільності розподілу ймовірності, можна зробити висновок, що $h=1$. Тому

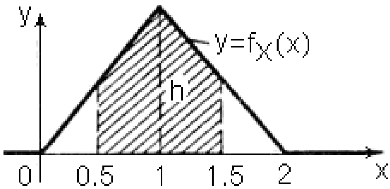


Рис. 4.2.17

$$f_X(x)=\begin{cases} 0, & \text{коли} \quad x \le 0; \\ x, & \text{коли} \quad 0 \le x \le 1; \\ 2-x, & \text{коли} \quad 1 \le x \le 2; \\ 0, & \text{коли} \quad 2 \le x, \end{cases}$$

оскільки кутовий коефіцієнт прямої, що проходить через точки $(0, 0)$ і $(1, 1)$, дорівнює $+1$, а кутовий коефіцієнт прямої, що проходить через точки $(1, 1)$ і $(2, 0)$, дорівнює -1. За заданих умов імовірність попадання на відрізок $[0.5; 1.5]$ дорівнює

$$\int_{0.5}^{1.5} f_X(x)\, dx = \frac{3}{4}.$$

Вправи для самостійного виконання

4.2.1. Перевірити, чи правильні твердження:

1. Кожній випадковій величині відповідає єдина функція розподілу ймовірностей на множині значень цієї величини.
2. Якщо випадкова величина визначена на дискретному просторі Ω елементарних подій, то множина її значень дискретна.
3. Твердження, обернене до 2, є правильним.
4. Дискретній множині значень випадкової величини відповідає кусково стала функція розподілу ймовірностей.
5. Якщо функція розподілу ймовірностей на множині значень випадкової величини не є кусково сталою, то множина значень випадкової величини неперервна.
6. Кожній неперервній множині значень випадкової величини відповідає неперервна щільність розподілу ймовірностей.
7. Вигляд функції $F_X(x)$ розподілу ймовірностей на множині Ω_X значень випадкової величини X залежить від структури подій із

простору S_X.

8. Кожна неперервна функція розподілу ймовірностей є абсолютно неперервною.

9. Функція розподілу ймовірностей на неперервній множині значень випадкової величини є скрізь диференційовною.

10. Кожна випадкова величина цілком визначається за рядом розподілу ймовірностей на множині її значень.

11. Стосовно кожної випадкової величини існує многокутник розподілу ймовірностей на множині її значень.

12. Якщо на множині значень випадкової величини розподіл ймовірностей мішаний, то множина значень функції $F_X(x)$ розподілу ймовірностей не є ні неперервною, ні дискретною.

4.2.2. Довести, що коли стосовно функції $F_X(x)$ задовільняються умови:

1) $F_X(x)$ неспадна на $(-\infty; +\infty)$;

2) $F_X(x)$ неперервна зліва в кожній точці із $(-\infty; +\infty)$;

3) $\lim\limits_{x \to -\infty} F_X(x) = 0$ і $\lim\limits_{x \to +\infty} F_X(x) = 1$,

то існує ймовірнісний простір (Ω, S, P) і випадкова величина $X(E),\ E \in \Omega$, такі, що $F_X(x)$ є функцією розподілу ймовірностей на множині значень випадкової величини $X(E)$.

4.3. Прості випадкові величини

Нехай задано ймовірнісний простір (Ω, S, P_n^*).

S/S_X-вимірну функцію $X = X(E)$, $E \in \Omega$, називають *простою випадковою величиною стосовно ймовірнісного простору* (Ω, S, P_n^*), коли множина значень цієї функції скінченна, тобто $X(\Omega) = \Omega_X = \{x_1, x_2, ..., x_m\}$, $m \in N$, де числа x_k попарно різні, і крім того $X^{-1}(\{x_i\}) \in S$ за довільного $x_i \in \Omega_X$, $i \in \overline{1, m}$.

Із наведеного означення випливає, що як S_X розглядається найширша сукупність підмножин множини Ω_X, а саме разом з порожньою підмножиною і множиною Ω_X до S_X входять всі одноелементні, всі двохелементні, всі триелементні, і т.д., всі (m-1)-елементні підмножини множини Ω_X, тобто $S_X = \{G \mid G = \bigcup\limits_{i \in I} \{x_i\},\ I \subset \{1, 2, ..., m\}\}$.

Зауважимо, що проста випадкова величина є S/S_X-вимірною функцією за будь-якої сукупності S_X підмножин множини Ω_X, стосовно якої задовільняються вимоги 1_s-3_s.

Прості випадкові величини надзвичайно важливі, оскільки за їх допомогою можна досліджувати й випадкові величини з нескінченними множинами значень.

За будь-якої $S/\mathscr{B}(R^1)$-випадкової величини X знайдеться послідовність простих випадкових величин $X_1, X_2, \ldots$ таких, що $|X_n| < |X|$ і $X_n(E) \underset{n \to \infty}{\to} X(E)$ за всіх $E \in \Omega$. Якщо крім того $X(E) \geq 0$, то знайдеться зростаюча послідовність простих випадкових величин $X_1, X_2, \ldots$ таких, що $X_n(E) \to X(E)$ за всіх $E \in \Omega$.

Справді, якщо $X(E) \geq 0$, то досить покласти

$$X_n(E) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{X^{-1}\left(\left[\frac{k-1}{2^n}; \frac{k}{2^n}\right)\right)}(E) + nI_{X^{-1}([n;\infty))}(E),$$

де $I_G(E)$ – індикаторна функція множини G, тобто

$$I_G(E) = \begin{cases} 1, & \text{коли } E \in G; \\ 0, & \text{коли } E \overline{\in} G. \end{cases}$$

Якщо X – довільна випадкова величина, то оскільки $X = X^+ - X^-$, можна окремо розглянути X^+ і X^-.

Зауважимо, що коли множина Ω_X обмежена, а m дуже велике (див. приклад 1.8.2), то на практиці множину Ω_X значень простої випадкової величини X доцільно розглядати як неперервну множину $\Omega_X = [a_0 ; a_k)$, де $a_0 = x_1$, $a_k = x_m + \varepsilon$, $\varepsilon > 0$ деяке досить мале число, і *наближено замінювати просту випадкову величину X іншою простою випадковою величиною* $\widetilde{X}$ з множиною значень $\Omega_{\widetilde{X}} = \{a_0, a_1, a_2, ..., a_{k-1}\}$ з ймовірностями

$$\widetilde{P}^*_{n\widetilde{X}}(\{a_{i-1}\}) = \widetilde{P}^*_{nX}([a_{i-1}; a_i)) = P^*_n(X^{-1}([a_{i-1}; a_i))),\ i \in \overline{1, k},$$

де $a_i = a_{i-1} + h$, $h = \dfrac{a_k - a_0}{k}$, k практично прийнятне, хоч можливо і досить велике, число.

В такому разі як $S_{\widetilde{X}}$ доцільно обрати сукупність підмножин множини $\Omega_{\widetilde{X}}$, в якій разом з порожньою множиною $\varnothing$ містяться всі одноелементні підмножини множини $\Omega_{\widetilde{X}}$, всі двохелементні, всі триелементні, і т.д., всі $(k-1)$-елементні, і нарешті k-елементна множина $\Omega_{\widetilde{X}}$, тобто як $S_{\widetilde{X}}$ обрати найширшу сукупність підмножин множини $\Omega_{\widetilde{X}}$, поклавши $S_{\widetilde{X}} = \{G \mid G = \bigcup\limits_{i \in I} \{a_i\},\ I \subset \{0, 1, 2, ..., k-1\}\}$.

Слід зауважити, що не обов'язково $\Omega_{\widetilde{X}} \subset \Omega_X$ і $S_{\widetilde{X}} \subset S_X$, однак обов'язково $\Omega_{\widetilde{X}} \subset \bigcup\limits_{i=1}^{k} [a_{i-1}; a_i)$,

$$P_{\widetilde{X}}(\{a_{i-1}\}) = P_X([a_{i-1}; a_i)) = P(X^{-1}([a_{i-1}; a_i))),\ i \in \overline{1, k}.$$

Приклади простих випадкових величин

1. Якщо задано ймовірнісний простір (Ω, S, P) і функція $X(E) = c = const$, $E \in \Omega$, то таку функцію називають *сталою випадковою величиною*. Множина значень сталої випадкової величини складається з одного елемента c, тобто $\Omega_X = X(\Omega) = \{c\}$. Тоді $X^{-1}(\{c\}) = \Omega$, а в найширшій сукупності S_X підмножин множини Ω_X міститься лише два елементи: $S_X = \{\varnothing, \Omega_X\}$.

2. Якщо задано ймовірнісний простір (Ω, S, P), $A \in S$ – випадкова подія і

$$X(E) = \begin{cases} 1, & \text{коли } E \in A, \\ 0, & \text{коли } E \in \overline{A}, \end{cases}$$

то таку функцію називають *індикатором події* A. Множина значень цієї функції складається з двох елементів 1 і 0, тобто $\Omega_X = X(\Omega) = \{0,1\}$, причому

$$X^{-1}(\{0\}) = \overline{A} \in S\,,\ X^{-1}(\{1\}) = A \in S\,,$$

$$P^*_{nX}(\{0\}) = P^*_n(X^{-1}(0)) = P^*_n(\overline{A})\,,\ P^*_{nX}(\{1\}) = P^*_n(X^{-1}(1)) = P^*_n(A)\,.$$

Тому $X(E)$, $E \in \Omega$ – проста випадкова величина відносно ймовірнісного простору (Ω, S, P) і $S_X = \{\varnothing, \{0\}, \{1\}, \{0,1\} = \Omega_X\}$.

3. Нехай $X(E)$, $E \in \Omega$, – кількість очок на грані кубика, якою кубик падає догори в результаті однократного підкидання,

$$\Omega = \{"1","2","3","4","5","6"\},\ X\{"k"\} = k\,,\ k \in \overline{1,6}\,,$$

і кожна множина $\{"k"\} \subset \Omega$ є подією, $k \in \overline{1,6}$, $P^*_n(\{"k"\}) = \dfrac{1}{6}$.
Тоді $\Omega_X = \{1,2,3,4,5,6\}$, $X^{-1}(\{i\}) = \{"i"\} \in S$, тому як S_X можна обрати $S^*_X = \{G \mid G \subset \Omega_X\}$ – найширшу сукупність підмножин множини Ω_X. Очевидно така функція $X(E)$, $E \in \Omega$, є S/S^*_X-вимірною, тобто S/S^*_X-випадковою величиною відносно (Ω, S, P) і S^*_X, а оскільки Ω_X скінченна множина, то $X(E)$, $E \in \Omega$, – проста випадкова величина.

4. Кількість $k_m(A)$ відбувань події A в серії із m випробувань – проста $\hat{S}/S_X$-випадкова величина відносно простору $(\hat{\Omega}_m, \hat{S}, \hat{P}^*_n)$ (див. §2.10), і найширшої сукупності S_X підмножин множини Ω_X =$\{0,\ 1,\ 2,\ \ldots,\ m-1,\ m\}$ значень випадкової величини *X*. Якщо в даному випадку всю серію із *m* випробувань розглядати як одне випробування, а події $\overline{A}_i$, A_j, $i \in I_1$, $j \in I_2$, $I_1 \subset \{1,2,...,m\}$, $I_2 \subset \{1,2,...,m\}$, $I_1 \bigcap I_2 = \varnothing$, $I_1 \bigcup I_2 = \{1,2,...,m\}$, статистично незалежні в сукупності (див. §2.10), тоді можна ввести ймовірнісний простір $(\widetilde{\Omega}, \widetilde{S}, \widetilde{P}^*_n)$, де $\widetilde{\Omega} = \Omega^m = \Omega \times \Omega \times ... \times \Omega$, сукупність $\widetilde{S}$ породжена за системою паралелепіпедів виду

$$\prod_{i \in I_1} \overline{A}_i \prod_{j \in I_2} A_j\,,\ I_1 \subset \{1,2,...,m\}\,,\ I_2 \subset \{1,2,...,m\}\,,$$

$$I_1 \bigcap I_2 = \varnothing\,,\ I_1 \bigcup I_2 = \{1,2,...,m\}\,,$$

$$\widetilde{E} = (E_1, E_2, ..., E_m) \in \widetilde{\Omega}\,,\ \widetilde{\Omega} = \bigcup_{i=0}^{r} H_i\,,\ r = 2^m - 1\,,$$

$X(\widetilde{E}) = k_m = m(I_2)$ – кількість множників A_l в добутках

$\prod_{i\in I_1}\overline{A}_i\prod_{j\in I_2}A_j$, H_i добутки виду (див. §2.10):

$$\begin{aligned}
H_0 &= \overline{A}_1\overline{A}_2\ldots\overline{A}_{m-2}\overline{A}_{m-1}\overline{A}_m\,,\\
H_1 &= \overline{A}_1\overline{A}_2\ldots\overline{A}_{m-2}\overline{A}_{m-1}A_m\,,\\
H_2 &= \overline{A}_1\overline{A}_2\ldots\overline{A}_{m-2}A_{m-1}\overline{A}_m\,,\\
H_3 &= \overline{A}_1\overline{A}_2\ldots\overline{A}_{m-2}A_{m-1}A_m\,,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
H_{2^m-1} &= A_1A_2\ldots A_{m-2}A_{m-1}A_m\,,\\
X^{-1}(k_m) &= B_{m,k_m}\in\hat{S}\,.
\end{aligned}\tag{4.3.1}$$

Кожна множина $H_i\subset\tilde{\Omega}$ є подією з імовірнісною мірою $\tilde{P}_n^*(H_i)=(P_n^*(A))^{k_m}(1-P_n^*(A))^{m-k_m}$, $k_m\in\overline{0,m}$, де k_m – кількість множників A_i в добутках H_i. Тому (див. формулу 2.10.3)

$$P_n^*(X^{-1}(k_m))=C_m^{k_m}(P_n^*(A))^{k_m}(1-P_n^*(A))^{m-k_m}\,,\ k_m\in\overline{0,m}\,.$$

5. Нехай задано ймовірнісний простір (Ω,S,P_n^*). Статистична ймовірність $P_m^*(A)=\dfrac{k_m(A)}{k_m(\Omega)}=\dfrac{k_m(A)}{m}$ події A, $A\in S$, визначена за результатами серії із m випробувань – проста $\tilde{S}/S_X$-випадкова величина $X(E)$ стосовно ймовірнісного простору $(\tilde{\Omega},\tilde{S},\tilde{P}_n^*)$ із прикладу 4 і найширшої сукупності S_X підмножин множини $\Omega_X=\left\{\dfrac{0}{m},\dfrac{1}{m},\dfrac{2}{m},\ldots,\dfrac{m-1}{m},\dfrac{m}{m}\right\}$ значень, яких може набувати величина $P_m^*(A)$, якщо $(\tilde{\Omega},\tilde{S},\tilde{P}_n^*)$, $\tilde{\Omega}$, $\tilde{S}$, $\tilde{P}_n^*(H_i)$, $\tilde{E}\in\tilde{\Omega}$, такі самі, як і в прикладі 4, а $X(\tilde{E})=\dfrac{k_m}{m}$, $\tilde{E}\in H_i$, де k_m – кількість множників A_i в добутках H_i, $i\in\overline{0,2^m-1}$ (див. формули 4.3.1).

Якщо $X(E)$ і $Y(E)$, $E\in\Omega$, – прості випадкові величини стосовно ймовірнісного простору (Ω,S,P_n^*), то їх сума $Z(E)=X(E)+Y(E)$, різниця $Z(E)=X(E)-Y(E)$, добуток $Z(E)=X(E)\cdot Y(E)$, частка $Z(E)=X(E)/Y(E)$, $E\in\Omega$, – прості випадкові величини стосовно того самого ймовірнісного простору (Ω,S,P_n^*).

Справді, нехай $\Omega_X=X(\Omega)=\{x_1,x_2,\ldots,x_m\}$ множина значень простої випадкової величини X, $\Omega_Y=Y(\Omega)=\{y_1,y_2,\ldots,y_r\}$ – множина значень простої випадкової величини Y. Зрозуміло, що кількість значень величини $Z(E)=X(E)+Y(E)$, $E\in\Omega$, не перевищує $m\cdot r$. Крім того, якщо випадкова величина $X(E)$ набуває значення x_j на множині $X^{-1}(x_j)\in S$, а випадкова величина $Y(E)$ набуває значення y_i на множині $Y^{-1}(y_i)\in S$, то на множині $X^{-1}(x_j)\bigcap Y^{-1}(y_i)\in S$ випадкова величина $Z(E)$ набуває значення x_j+y_i.

Нехай $\Omega_Z=\{z_1,z_2,...,z_l\}$, $l\le m\cdot r$, – множина значень величини $Z(E)=X(E)+Y(E)$.

Тоді $Z^{-1}(z_k)=\bigcup\limits_{x_j+y_i=z_k}(X^{-1}(x_j)\bigcap Y^{-1}(y_i))\in S$.

Отже $Z(E)$ – випадкова величина, а оскільки множина її значень скінченна, то $Z(E)$ – проста випадкова величина стосовно ймовірнісного простору (Ω,S,P_n^*).

Очевидно, що коли розглянути сукупність S_Z підмножин множини $\Omega_X=\{z_1,z_2, ..., z_k\}$, породжену за поділом множини Ω_Z на підмножини $\{z_1\}$, $\{z_2\}$, …, $\{z_k\}$, тобто $S_Z=\{G\,\big|\; G=\bigcup\limits_{i\in I}\{z_i\},\ I\subset\{1,2,...,k\}\}$, то за довільного $G\in S_Z$ буде $Z^{-1}(G)\in S$.

Аналогічно розглядаються випадки $Z(E)=X(E)-Y(E)$, $Z(E)=X(E)\cdot Y(E)$, $Z(E)=X(E)/Y(E)$, $(Y(E)\ne 0)$.

За методом математичної індукції легко довести, що сума і добуток довільної скінченної кількості простих випадкових величин $X_k(E)$, $E\in\Omega$, стосовно одного і того самого ймовірнісного простору (Ω,S,P_n^*), $k\in\overline{1,m}$, також є простою випадковою величиною стосовно того самого ймовірнісного простору (Ω,S,P_n^*).

Узагальненням цього є твердження про те, що *коли* $X_k(E)$ – *прості випадкові величини стосовно одного і того самого ймовірнісного простору* (Ω,S,P_n^*), $k\in\overline{1,m}$, *то будь-яка дійсна функція* $f(X_1(E), X_2(E), ..., X_m(E)), E\in\Omega$, *також є простою випадковою величиною стосовно того самого ймовірнісного простору* (Ω,S,P_n^*), *за умови, що* $(X_1(E), ..., X_m(E))$ *належить до області визначення* $D(f)$ *функції* f *за кожного* $E\in\Omega$. *Зокрема, якщо* $X(E)$ – *проста випадкова величина стосовно ймовірнісного простору* (Ω,S,P_n^*), *то* $X^m(E)$, $\sqrt[m]{X(E)}$ (*в разі парного* m $X(E)$ *має бути невід'ємною*), $e^{X(E)}$, $\ln X(E)$ (*коли*

$X(E) > 0$ *за* $E \in \Omega$), $\sin X(E)$, $|X(E)| = \sqrt{X^2(E)}$, $\cos X(E)$ *тощо також є простими випадковими величинами стосовно того самого ймовірнісного простору* (Ω, S, P_n^*).

Приклад 4.3.1. Нехай $\Omega = [0;1)$, $X(E)$ – індикатор події $A = [0; 0.5) \subset \Omega = [0;1)$, $Y(E)$ – індикатор події $B = [0.25; 1)$.

Тоді $X^{-1}(1) = A = [0; 0.5)$, $X^{-1}(0) = \overline{A} = [0.5; 1)$;

$Y^{-1}(1) = B = [0.25; 1)$, $Y^{-1}(0) = \overline{B} = [0; 0.25)$.

Якщо $Z(E) = X(E) + Y(E)$, то

а) $Z(E) = 2$, коли $E \in X^{-1}(1)$ і $E \in Y^{-1}(1)$ тобто

$Z^{-1}(2) = X^{-1}(1) \bigcap Y^{-1}(1) = [0; 0.5) \bigcap [0.25; 1) = [0.25; 0.5)$.

б) $Z(E) = 1$, коли $E \in X^{-1}(1) \bigcap Y^{-1}(0)$ або $E \in X^{-1}(0) \bigcap Y^{-1}(1)$, тобто

$$Z^{-1}(1) = (X^{-1}(1) \bigcap Y^{-1}(0)) \bigcup (X^{-1}(0) \bigcap Y^{-1}(1)) =$$
$$= ([0; 0.5) \bigcap ([0; 0.25)) \bigcup ([0.5; 1) \bigcap ([0.25; 1)) = ([0; 0.25) \bigcup [0.5; 1).$$

в) $Z(E) = 0$, коли $E \in X^{-1}(0)$ і $E \in Y^{-1}(0)$, тобто

$$Z^{-1}(0) = X^{-1}(0) \bigcap Y^{-1}(0) = [0.5; 1) \bigcap [0;\ 0.25) = \varnothing.$$

Таким чином,

$$Z(E) = X(E) + Y(E) = \begin{cases} 2, & \text{коли } E \in [0.25; 0.5), \\ 1, & \text{коли } E \in [0; 0.25) \bigcup [0.5; 1). \end{cases}$$

Прості випадкові величини $X(E)$ та $Y(E)$ називаються *незалежними відносно ймовірнісної міри* P (зокрема відносно статистичної ймовірності P_n^*), коли за будь-яких чисел $a \in X(\Omega) = \Omega_X$ та $b \in Y(\Omega) = \Omega_Y$ події $X^{-1}(a)$ та $Y^{-1}(b)$ є незалежними відносно міри P з відповідного ймовірнісного простору (Ω, S, P), тобто

$$P(X^{-1}(a) \bigcap Y^{-1}(b)) = P(X^{-1}(a)) \cdot P(Y^{-1}(b)).$$

В іншому разі *випадкові величини* X та Y називаються *залежними* відносно міри P.

Приклад 4.3.2. 1. Нехай $X(E)$ та $Y(E)$ з прикладу 4.3.1.

Якщо $P([a; b)) = b - a$, $[a; b) \subset [0; 1)$, то

$P(X^{-1}(1) \bigcap = Y^{-1}(1)) = 0.5 - 0.25 \neq P(X^{-1}(1)) \cdot P(Y^{-1}(1)) = 0.5 \cdot 0.75 = 0.375.$

Це означає, що випадкові величини $X(E)$ та $Y(E)$ відносно так заданої ймовірнісної міри P залежні.

2. Нехай P визначена так, що $P([0.25; 0.5)) = 1$. Тоді

$$P([0; 0.5)) = P([0.25; 1)) = 1,\ P([0; 0.25)) = P([0.5; 1)) = 0,$$
$$P(X^{-1}(1) \bigcap Y^{-1}(1)) = 1 = P(X^{-1}(1)) \cdot P(Y^{-1}(1)) = 1 \cdot 1,$$

$$P(X^{-1}(1)\bigcap Y^{-1}(0))=0=P(X^{-1}(1))\cdot P(Y^{-1}(0))=1\cdot 0,$$
$$P(X^{-1}(0)\bigcap Y^{-1}(1))=0=P(X^{-1}(0))\cdot P(Y^{-1}(1))=0\cdot 1,$$
$$P(X^{-1}(0)\bigcap Y^{-1}(0))=0=P(X^{-1}(0))\cdot P(Y^{-1}(0))=0\cdot 0.$$

Це означає, що випадкові величини $X(E)$ та $Y(E)$ незалежні відносно так заданої ймовірнісної міри P.

Таким чином, *залежність чи незалежність випадкових величин суттєво визначається за ймовірнісною мірою* P.

Випадкові величини $X_1(E), X_2(E), ..., X_m(E)$ називаються *попарно незалежними відносно ймовірнісної міри* P, коли будь які дві з них є незалежними між собою відносно тієї самої ймовірнісної міри P.

Приклад 4.3.3. 1. Якщо $X(E)$ і $Y(E)$ та P з прикладу 4.3.2.2, а $Z(E)=1$, коли $E\in\Omega$, то легко бачити, що $X(E)$, $Y(E)$ та $Z(E)$ – попарно незалежні прості випадкові величини, оскільки стала випадкова величина та будь-яка випадкова величина незалежні, бо $Z^{-1}(1)=\Omega$.

2. Нехай проведено n випробувань, подія A_k означає відбування події A у k-му випробуванні, події A_k виявились практично незалежними (статистично незалежними) в сукупності відносно деякої ймовірнісної міри P (див. §2.10). Тоді індикатори цих подій $X_k(\widetilde{E})=X_{A_k}(\widetilde{E})$, $k\in\overline{1,m}$, $\widetilde{E}\in\widetilde{\Omega}$, практично попарно незалежні (статистично незалежні) стосовно міри P, оскільки $X_k^{-1}(1)=A_k$, $X_k^{-1}(0)=\overline{A}_k$, $k\in\overline{1,m}$ і тому
$$P(X_k^{-1}(a)\bigcap X_i^{-1}(b))=P(X_k^{-1}(a))\cdot P(X_i^{-1}(b)),$$
коли $k\neq i$, $a\in\{0,1\}$ і $b\in\{0,1\}$.

Вправи для самостійного виконання

4.3.1. Перевірити, чи правильні твердження:

1. Якщо функція $X(E)$, $E\in\Omega$, набуває скінченну кількість значень, то вона є простою випадковою величиною.
2. Твердження, обернене до 1, правильне.
3. Якщо функція $X(E)$, $E\in\Omega$, така, що $P_n^*(\{E \mid E\in\Omega, X(E)\neq 1\})=0$, то ця функція є випадковою величиною.
4. Функція $X(E)$, $E\in\Omega$, з твердження 3 є простою випадковою величиною.
5. Якщо існують події $A_k\subset\Omega, k\in\overline{1,s}$, такі що $A_iA_j=\varnothing$, коли $i\neq j$, $\sum\limits_{k=1}^{s}A_k=\Omega$ і $X(E)=c_k$, коли $E\in A_k$, $k\in\overline{1,s}$, то $X(E)$, $E\in\Omega$, – проста випадкова величина.

6. Якщо $X_1(E) + X_2(E), E \in \Omega,$ – проста випадкова величина, то $X_1(E)$ і $X_2(E)$ – прості випадкові величини.

7. Твердження, обернене до 6, є правильним.

8. За будь-якої простої випадкової величини $X(E)$ стосовно ймовірнісного простору (Ω, S, P_n^*) функція $Y(E) = \sqrt{X(E)}$ є простою випадковою величиною стосовно того самого ймовірнісного простору.

9. Випадкові величини $X_1(E)$ та $X_2(E)$, $E \in \Omega$, незалежні відносно ймовірнісної міри P, коли

$$P(X_1^{-1}(a) \bigcap X_2^{-1}(a)) = P(X_1^{-1}(a)) \cdot P(X_2^{-1}(a))$$

за деякого числа a.

10. Залежність випадкових величин визначається за ймовірнісною мірою, тобто відносно однієї міри випадкові величини можуть бути залежними, а відносно іншої – незалежними.

11. Існує ймовірнісний простір, за якого будь-які дві випадкові величини є незалежними.

12. Прості випадкові величини $X_1(E)$ та $X_2(E)$ стосовно одного і того самого ймовірнісного простору (Ω, S, P_n^*) є незалежними відносно ймовірнісної міри P_n^* тоді і тільки тоді, коли вони *майже сталі*, тобто існують сталі c_i, за яких $P_n^*(\{E \in \Omega \mid X_i(E) \neq c_i\}) = 0$, $i \in \overline{1,2}$.

4.3.2. Нехай $X_1(E)$ – індикатор події $A = [0;1] \subset R^1 = \Omega$, а $X_2(E)$ – індикатор події $B = [1;2] \subset \Omega$. Знайти: суму, різницю, добуток і частку $X_1(E)$ та $X_2(E)$.

4.4. Числові характеристики розподілів статистичних ймовірностей на множинах значень простих випадкових величин

Нехай задано ймовірнісний простір (Ω, S, P_n^*) та просту випадкову величину $X(E)$, $E \in \Omega$, стосовно цього простору (Ω, S, P_n^*), множина значень якої $\Omega_X = X(\Omega) = \{x_1, x_2, \ldots, x_m\}$. Тоді за означенням *статистичним математичним сподіванням* $M_n^*(X)$ та *статистичною дисперсією* $D_n^*(X)$ розподілу статистичних ймовірностей на множині значень цієї випадкової величини є числа:

$$M_n^*[X] = \sum_{k=1}^{m} x_k P_{nX}^*(\{x_k\}) = \sum_{k=1}^{m} x_k P_n^*(X^{-1}(x_k)), \quad (3.4.1)$$

$$D_n^*[X] = M_n^*[(X - M_n^*[X])^2] = \sum_{k=1}^{m} (x_k - M_n^*[X])^2 P_{nX}^*(\{x_k\}) =$$

$$= \sum_{k=1}^{m} (x_k - M_n^*[X])^2 P_n^*(X^{-1}(x_k)). \quad (3.4.2)$$

Число $M_n^*[X]$ називають також *середнім статистичним значенням* простої випадкової величини X, або центром розсіювання ймовірностей на множині значень випадкової величини X, а число $D_n^*[X]$ – *мірою розсіювання* (або *скупченості*) ймовірностей на множині значень простої випадкової величини навколо середнього статистичного значення (центра розсіювання).

Очевидно, $D_n^*[X] \le (\max_{1 \le k \le m} x_k - \min_{1 \le k \le m} x_k)^2$.

Приклад 4.4.1.

1. Якщо $X(E) = c$, $E \in \Omega$, – стала випадкова величина, то $\Omega_X = X(\Omega) = \{c\}$, $X^{-1}(c) = \Omega$, $P_n^*(X^{-1}(c)) = P_n^*(\Omega) = 1$ і тому $M_n^*[c] = c \cdot 1 = c$, $D_n^*[c] = M_n^*[(c-c)^2] = M_n^*[0] = 0$.

2. Якщо $X_A(E)$, $E \in \Omega$, – індикатор події A, то $\Omega_X = X(\Omega) = \{1, 0\}$, $X^{-1}(1) = A$, $X^{-1}(0) = \overline{A}$, $P_n^*(X^{-1}(1)) = P_n^*(A) = p$, $P_n^*(X^{-1}(0)) = P_n^*(\overline{A}) = 1 - p$. Тому

$$M_n^*[X_A] = 1 \cdot p + 0 \cdot (1-p) = p,$$

$$D_n^*[X_A] = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = p(1-p)(1-p+p) = p(1-p).$$

Основні властивості статистичного математичного сподівання та статистичної дисперсії простої випадкової величини:

1. Статистичне математичне сподівання сталої дорівнює *цій сталій* (див. приклад 4.4.1):

$$M_n^*[c] = c\ .$$

2. Статистична дисперсія сталої дорівнює нулеві (див. приклад 4.4.1):

$$D_n^*[c] = 0\,.$$

3. Сталу можна виносити за знак статистичного математичного сподівання простої випадкової величини:

$$M_n^*[cX] = cM_n^*[X]\,.$$

Справді, якщо $\Omega_X = X(\Omega) = \{x_1,\ x_2,\ ...,\ x_s\}$, а $Y(E) = cX(E)$, $E \in \Omega$, то за $c \neq 0$

$$\Omega_Y = Y(\Omega) = \{cx_1,\, cx_2,\, ...,\, cx_s\},\ Y^{-1}(cx_i) = X^{-1}(x_i)\ ,$$

$$P_n^*(Y^{-1}(cx_i)) = P_n^*(X^{-1}(x_i))\ ,\ i = \{1, 2, ..., s\}\,,$$

тому

$$M_n^*[Y] = M_n^*[cX] = \sum_{i=1}^{s} cx_i P_n^*(X^{-1}(x_i)) = c\sum_{i=1}^{s} x_i P_n^*(X^{-1}(x_i)) = cM_n^*[X]\,.$$

4. За знак статистичної дисперсії стала виноситься в квадраті:

$$D_n^*\big[cX\big] = c^2 D_n^*\big[X\big].$$

Справді, враховуючи властивість 3, одержимо:

$$D_n^*\big[cX\big] = M_n^*[(cX - M_n^*[cX])^2] = M_n^*[(cX - cM_n^*[X])^2] =$$
$$= M_n^*[c^2(X - M_n^*[X])^2] = c^2 M_n^*[(X - M_n^*[X])^2] = c^2 D_n^*\big[X\big].$$

5. Статистичне математичне сподівання суми простих випадкових величин відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) *дорівнює сумі їх статистичних математичних сподівань*:

$$M_n^*[X + Y] = M_n^*[X] + M_n^*[Y]\,.$$

Справді, нехай

$$Z(E) = X(E) + Y(E),\ E \in \Omega\,,$$

$$\Omega_X = X(\Omega) = \{x_1,\ x_2,\ ...,\ x_s\},\ \ \Omega_Y = Y(\Omega) = \{y_1,\ y_2,\ ...,\ y_m\},$$

$$X^{-1}(x_j) \neq \varnothing,\ j \in \overline{1,\, s}\,,\ \ Y^{-1}(y_i) \neq \varnothing\,,\ i \in \overline{1,\, m}\,.$$

Зрозуміло, що $\bigcup_{j=1}^{s} X^{-1}(x_j) = \Omega$, $\bigcup_{i=1}^{m} Y^{-1}(y_i) = \Omega$,

$$(\bigcup_{j=1}^{s} X^{-1}(x_j)) \bigcap (\bigcup_{i=1}^{m} Y^{-1}(y_i)) = \bigcup_{j=1}^{s}\bigcup_{i=1}^{m}(X^{-1}(x_j) \bigcap Y^{-1}(y_i)) = \Omega\,.$$

Серед множин $X^{-1}(x_j) \bigcap Y^{-1}(y_i)$, $i \in \overline{1,\, m}$, $j \in \overline{1,\, s}$, можуть бути порожні, а також такі, за яких $x_j + y_i$ однакові між собою. Знайдемо серед значень $x_j + y_i$ всі попарно різні значення.

Позначимо їх через z_k, $k \in \overline{1, r}$, де $r \le m \cdot s$. Тоді

$$Z^{-1}(z_k) = \bigcup_{x_j + y_i = z_k} X^{-1}(x_j) \cap Y^{-1}(y_i), \ k \in \overline{1, r}, \ \bigcup_{k=1}^{r} Z^{-1}(z_k) = \Omega .$$

Оскільки

$$X^{-1}(x_j) \cap X^{-1}(x_{j_1}) = \varnothing, \text{ коли } j \ne j_1,$$

$$Y^{-1}(y_i) \cap Y^{-1}(y_{i_1}) = \varnothing, \text{ коли } i \ne i_1,$$

то

$$(X^{-1}(x_j) \cap Y^{-1}(y_i)) \cap (X^{-1}(x_{j_1}) \cap Y^{-1}(y_{i_1})) = \varnothing,$$

коли $i \ne i_1$ або $j \ne j_1$.

Тому

$$P_n^*(Z^{-1}(z_k)) = P_n^*(\bigcup_{x_j + y_i = z_k} X^{-1}(x_j) \cap Y^{-1}(y_i)) = \sum_{x_j + y_i = z_k} P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)),$$

Визначаючи $M_n^*[Z]$, враховуючи, що $P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)) = 0$, коли $X^{-1}(x_j) \cap Y^{-1}(y_i) = \varnothing$, одержимо

$$M_n^*[Z] = \sum_{k=1}^{r} z_k P_n^*(Z^{-1}(z_k)) = \sum_{k=1}^{r} z_k P_n^*(\bigcup_{x_j + y_i = z_k} (X^{-1}(x_j) \cap Y^{-1}(y_i))) =$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{s} (x_j + y_i) P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)) =$$

$$= \sum_{i=1}^{m} y_i \sum_{j=1}^{s} P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)) + \sum_{j=1}^{s} x_j \sum_{i=1}^{m} P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)) =$$

$$= \sum_{i=1}^{m} y_i P_n^*(\bigcup_{j=1}^{s} X^{-1}(x_j) \cap Y^{-1}(y_i)) + \sum_{j=1}^{s} x_j P_n^*(\bigcup_{i=1}^{m} X^{-1}(x_j) \cap Y^{-1}(y_i)) =$$

$$= \sum_{i=1}^{m} y_i P_n^*(Y^{-1}(y_i) \cap (\bigcup_{j=1}^{s} X^{-1}(x_j))) + \sum_{j=1}^{s} x_j P_n^*(X^{-1}(x_j) \cap (\bigcup_{i=1}^{m} Y^{-1}(y_i))) =$$

$$= \sum_{i=1}^{m} y_i P_n^*(Y^{-1}(y_i)) + \sum_{j=1}^{s} x_j P_n^*(X^{-1}(x_j)) = M_n^*[Y] + M_n^*[X].$$

За методом математичної індукції можна довести, що

$$M_n^*[\sum_{i=1}^{k} X_i] = \sum_{i=1}^{k} M_n^*[X_i], \ k \in N .$$

6. Статистичне математичне сподівання лінійної комбінації простих випадкових величин X та Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) *дорівнює цій самій лінійній комбінації їх статистичних математичних сподівань*:

$$M_n^*[aX + bY] = aM_n^*[X] + bM_n^*[Y] .$$

Ця властивість випливає із властивостей 3 та 5. За методом математичної індукції можна показати, що

$$M_n^*[\sum_{i=1}^{k} a_i X_i] = \sum_{i=1}^{k} a_i M_n^*[X_i].$$

Зокрема

$$M_n^*[aX + bY + c] = aM_n^*[X] + bM_n^*[Y] + c,$$

$$M_n^*[\sum_{i=1}^{k} a_i X_i + b] = \sum_{i=1}^{k} a_i M_n^*[X_i] + b.$$

7. Якщо $X[E] \geq 0$, $E \in \Omega$, то $M_n^*[X] \geq 0$.

Ця властивість випливає з означення статистичного математичного сподівання.

8. Якщо $X(E) \geq Y(E)$, $E \in \Omega$, де $X(E)$, $Y(E)$ – прості випадкові величини відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*), то $M_n^*[X] \geq M_n^*[Y]$.

Ця властивість випливає з властивостей 6 та 7, оскільки

$$M_n^*[X - Y] = M_n^*[X] - M_n^*[Y] \geq 0.$$

9. Статистичне математичне сподівання добутку незалежних простих випадкових величин X та Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) *дорівнює добуткові їх математичних сподівань*:

$$M_n^*[X \cdot Y] = M_n^*[X] \cdot M_n^*[Y].$$

Справді, нехай $Z = X \cdot Y$, $\Omega_X = X(\Omega) = \{x_1, x_2, ..., x_s\}$, $\Omega_Y = Y(\Omega) = \{y_1, y_2, ..., y_m\}$, $\Omega_Z = Z(\Omega) = \{z_1, z_2, ..., z_r\}$, $r \leq m \cdot s$.

Тоді

$$M_n^*[Z] = \sum_{k=1}^{r} z_k P_n^*(Z^{-1}(z_k)) = \sum_{k=1}^{r} z_k P_n^*(\bigcup_{x_j \cdot y_i = z_k} (X^{-1}(x_j) \cap Y^{-1}(y_i))) =$$

$$= \sum_{k=1}^{r} z_k \sum_{x_j y_i = z_k} P_n^*(X^{-1}(x_j) \cap Y^{-1}(y_i)) =$$

$$= \sum_{k=1}^{r} z_k \sum_{x_j y_i = z_k} P_n^*(X^{-1}(x_j)) P_n^*(Y^{-1}(y_i)) =$$

$$= \sum_{k=1}^{r} \sum_{x_j y_i = z_k} x_j y_i P_n^*(X^{-1}(x_j)) P_n^*(Y^{-1}(y_i)) =$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{s} x_j y_i \, P_n^*(X^{-1}(x_j)) P_n^*(Y^{-1}(y_i)) =$$

$$= \sum_{j=1}^{s} x_j P_n^*(X^{-1}(x_j)) \cdot \sum_{i=1}^{m} y_i P_n^*(Y^{-1}(y_i)) = M_n^*[X] \cdot M_n^*[Y].$$

10. Статистична дисперсія суми незалежних простих *випадкових величин* X *та* Y *відносно одного і того самого ймовірнісного простору* (Ω, S, P_n^*) *дорівнює сумі їх дисперсій*:

$$D_n^*[X+Y] = D_n^*[X] + D_n^*[Y].$$

Справді, за означенням статистичної дисперсії

$$D_n^*[X+Y] = M_n^*[(X+Y-M_n^*[X+Y])^2] =$$
$$= M_n^*[((X - M_n^*[X]) + (Y - M_n^*[Y]))^2] =$$
$$= M_n^*[(X - M_n^*[X])^2 + 2(X - M_n^*[X])(Y - M_n^*[Y]) + (Y - M_n^*[Y])^2] =$$
$$= M_n^*[(X - M_n^*[X])^2 + 2M_n^*[(X - M_n^*[X])(Y - M_n^*[Y])] +$$
$$+ M_n^*[(Y - M_n^*[Y])^2] = D_n^*[X] + D_n^*[Y],$$

оскільки

$$M_n^*[(X - M_n^*[X])(Y - M_n^*[Y])] =$$
$$= M_n^*[X \cdot Y - YM_n^*[X] - XM_n^*[Y] + M_n^*[X] \cdot M_n^*[Y]] = M_n^*[X] \cdot M_n^*[Y] -$$
$$- M_n^*[Y] \cdot M_n^*[X] - M_n^*[X] \cdot M_n^*[Y] + M_n^*[X] \cdot M_n^*[Y] = 0,$$

бо X і Y незалежні прості випадкові величини.

11. Статистична дисперсія лінійної комбінації (*з коефіцієнтами* a *і* b) *незалежних простих випадкових величин* X *та* Y *відносно одного і того самого ймовірнісного простору* (Ω, S, P_n^*) *дорівнює лінійній комбінації їх статистичних дисперсій з коефіцієнтами* a^2 *і* b^2:

$$D_n^*[aX + bY] = a^2 D_n^*[X] + b^2 D_n^*[Y].$$

Ця властивість випливає з властивостей 10 та 4.

Дана властивість задовільняється стосовно лінійної комбінації будь-якої скінченної кількості попарно незалежних випадкових величин X_j відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*), $j \in \overline{1, r}$:

$$D_n^*\left[\sum_{j=1}^{r} a_j X_j\right] = \sum_{j=1}^{r} a_j^2 \, D_n^*[X_j].$$

12. Якщо X *– невід'ємна проста випадкова величина відносно ймовірнісного простору* (Ω, S, P_n^*), *число* $\varepsilon > 0$ *і множина* $A_\varepsilon = \{E \mid E \in \Omega, X(E) \ge \varepsilon\}$, *то* A_ε *є подією, стосовно якої задовільняється нерівність П.Л. Чебишова*:

$$P_n^*(A_\varepsilon) \le \frac{1}{\varepsilon} M_n^*[X].$$

Справді, якщо

$$\Omega_X = X(\Omega) = \{x_1, x_2, \ldots, x_m\}, \text{ де } 0 \le x_1 < x_2 < \ldots < x_m,$$

то за фіксованого $\varepsilon > 0$ знайдемо найменший номер n_0, за якого $x_{n_0} \ge \varepsilon$. Тоді $x_k \ge \varepsilon$, коли $k \ge n_0$, а множина

$A_\varepsilon = \bigcup_{k=n_0}^{m} X^{-1}(x_k)$ є подією, бо $X^{-1}(x_k)$ події. Тому

$$M_n^*[X] = \sum_{k=1}^{m} x_k P_n^*(X^{-1}(x_k)) \geq \sum_{k=n_0}^{m} x_k P_n^*(X^{-1}(x_k)) \geq \varepsilon \sum_{k=n_0}^{m} P_n^*(X^{-1}(x_k)) =$$

$$= \varepsilon P_n^*(\bigcup_{k=n_0}^{m}(X^{-1}(x_k)) = \varepsilon P_n^*(A_\varepsilon).$$

Отже,

$$M_n^*[X] \geq \varepsilon P_n^*(A_\varepsilon), \text{ тобто } P_n^*(A_\varepsilon) \leq \frac{1}{\varepsilon} M_n^*[X].$$

Зауваження. Подію A_ε, яка полягає в тому, що випадкова величина X набуває значення, не меншого ніж ε, іноді позначають також $(X \geq \varepsilon)$, а тому нерівність П.Л. Чебишова часто записують у вигляді $P_n^*(X \geq \varepsilon) \leq \frac{1}{\varepsilon} M_n^*[X]$.

Зокрема

$$P_n^*(| X - M_n^*[X]| \geq \varepsilon) =$$

$$= P_n^*((X - M_n^*[X])^2 \geq \varepsilon^2) \leq \frac{1}{\varepsilon^2} M_n^*[(X - M_n^*[X])^2] = \frac{1}{\varepsilon^2} D_n^*[X].$$

Нерівність

$$P_n^*(| X - M_n^*[X]| \geq \varepsilon) \leq \frac{D_n^*[X]}{\varepsilon^2}$$

також називають нерівністю П.Л. Чебишова.

Вправи для самостійного виконання

4.4.1. Перевірити, чи правильні твердження:

1. Стосовно кожної простої випадкової величини відносно ймовірнісного простору (Ω, S, P_n^*) існує її статистичне математичне сподівання і статистична дисперсія, які можна обчислити за єдиним способом.

2. Якщо $X(E)$, $E \in \Omega$, – проста випадкова величина відносно ймовірнісного простору (Ω, S, P_n^*), то $D_n^*[X] = M_n^*[X^2] - (M_n^*[X])^2$.

3. Якщо $X(E)$ – проста випадкова величина відносно ймовірнісного простору (Ω, S, P_n^*) і $P_n^*(\{E \mid E \in \Omega, X(E) \neq c\}) = 0$ за деякої константи c, то $M_n^*[X] = c$.

4. Якщо $X(E)$ – з твердження 3, то $D_n^*[X] = 0$.

5. За будь яких простих випадкових величин X та Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*)

$M_n^*[XY] = M_n^*[X] \cdot M_n^*[Y]$.

6. За будь-яких простих випадкових величин X та Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) $M_n^*[X \pm Y] = M_n^*[X] \pm M_n^*[Y]$.

7. Якщо $M_n^*[X] \geq M_n^*[Y]$, то $X \geq Y$.

8. За будь яких простих випадкових величин X та Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) $D_n^*[X + Y] = D_n^*[X] + D_n^*[Y]$.

9. Нерівність Чебишова задовільняється стосовно будь якої простої випадкової величини.

4.4.2. Знайти статистичне математичне сподівання і статистичну дисперсію розподілу статистичних ймовірностей на множині значень випадкової величини X, якщо:

1. $X(E_i) = i$, коли $E_i = "i" \in \{"1", "2", "3", "4", "5", "6"\} = \Omega$, а $P_n^*(E_i) = \dfrac{1}{6}$, $i \in \overline{1,\ 6}$.

2. $\Omega_X = X(\Omega) = \{0, 1, 2, 3, 4, 5\}$, $P_n^*(x_i) = C_m^{x_i} (\dfrac{1}{2})^{x_i} (\dfrac{1}{2})^{m - x_i}$, $x_i \in \Omega_X$, $m = 5$, $i \in \{0, 1, 2, 3, 4, 5\}$.

3. $\Omega_X = X(\Omega) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $P_n^*(x_i) = \dfrac{a^{x_i}}{x_i!} e^{-a}$, $x_i \in \Omega_X$, $i = \overline{0,\ 10}$.

4.4.3. Використовуючи властивості статистичного математичного сподівання, довести, що коли прості випадкові величини X і Y відносно одного і того самого ймовірнісного простору (Ω, S, P_n^*) незалежні, то і випадкові величини $X + a$ і $Y + b$ відносно того самого ймовірнісного простору незалежні, де a і b – довільні сталі дійсні числа.

4.5. Функції випадкового аргумента

Нехай X – випадкова величина, задана на ймовірнісному просторі (Ω, S, P), де $\Omega = R^1$, $S = \mathscr{B}(R^1)$, P – ймовірнісна міра, задана на S, $\psi(x): R^1 \to R^1$ – борелівська функція. Тоді $Y = \psi(X)$ є також випадковою величиною стосовно простору (Ω, S, P).

Дійсно, якщо

$$\Omega_X = X(\Omega) = R^1,\ S_X = \mathscr{B}(R^1),\ P_X(G) = P(X^{-1}(G)),\ G \in S_X,$$

а ψ – деяка борелівська функція, то $Y = \psi(X)$ є S_X-вимірною функцією, тобто такою, що за довільного $B \in \mathscr{B}(R^1)$, $\psi^{-1}(B) \in S_X$, а оскільки $X^{-1}(\psi^{-1}(B)) = X^{-1}(Y^{-1}(B)) \in S$, $B \in S_Y = \mathscr{B}(R^1)$, то $Y = \psi(X)$ – випадкова величина, де S_Y – простір подій із ймовірнісного простору

$$(\Omega_Y, S_Y, P_Y),\ \Omega_Y = Y(\Omega_X) = \psi(\Omega_X),\ S_Y = \mathscr{B}(R^1),$$
$$P_Y(B) = P_X(Y^{-1}(B)) = P(X^{-1}(Y^{-1}(B)))$$

за довільного $B \in S_Y = \mathscr{B}(R^1)$.

Якщо на $(R^1, \mathscr{B}(R^1))$ визначено ймовірнісну міру $P_X(B)$,

$$P_X(B) = P(\{E \mid X(E) \in B\}) = P(X^{-1}(B)),\ B \in \mathscr{B}(R^1),$$

то покладаючи

$$P_Y(B) = P_X(\psi^{-1}(B)) = P(X^{-1}(\psi^{-1}(B))),$$

можна ввести також міру $P_Y(B)$, за якою описується розподіл імовірностей на множині значень випадкової величини Y, тобто на $(\Omega_Y, S_Y) = (R^1, \mathscr{B}(R^1))$.

Функція розподілу ймовірностей на множині значень випадкової величини Y має вигляд

$$F_Y(y) = P_Y((-\infty; y)) = P_X(\psi^{-1}((-\infty; y))).$$

Якщо розподіл ймовірностей на множині $\Omega_X = \{x_1, x_2, ..., x_k\}$ значень випадкової величини X поточковий, то

$$F_Y(y) = P_Y((-\infty; y)) = P_X(\psi^{-1}((-\infty; y))) = \sum_{x_i \in \psi^{-1}(-\infty; y)} P_X(\{x_i\}) = \sum_{x_i \in \psi^{-1}(-\infty; y)} p_i,$$

де x_i – окремі значення випадкової величини X, $p_i = P_X(\{x_i\})$ – імовірності цих значень (див. § 2.5).

Якщо множина Ω_X неперервна типу $\Omega_X = [a; b)$ (чи $\Omega_X = (a, b]$) і розподіл ймовірностей на множині Ω_X

поінтервальний за інтервалами $[a_{i-1};\, a_i)$ (чи $(a_{i-1};\, a_i])$, $i \in \overline{1,\, k}$, $a_0 = a$, $a_k = b$, $a_i - a_{i-1} = h > 0$ за всіх $i \in \overline{1, k}$, $[a_{i-1};\, a_i) \bigcap [a_{j-1};\, a_j) = \varnothing$ (чи $(a_{i-1};\, a_i] \bigcap (a_{j-1};\, a_j] = \varnothing$), коли $i \neq j$, $P([a_{i-1};\, a_i)) = p_i$ (як і $P((a_{i-1};\, a_i]) = p_i$), а функція $Y = \psi(X)$ набуває сталих значень на кожному інтервалі $[a_{i-1};\, a_i)$ (чи $(a_{i-1};\ a_i]$), так що $\Omega_Y = \{y_1, y_2, ... y_m\}$, тоді

$$F_Y(y) = P_Y((-\infty; y)) = P_X(\psi^{-1}(-\infty\,; y)) =$$
$$= \sum_{[a_{i-1}; a_i) \subset \psi^{-1}((-\infty; y))} P_X([a_{i-1};\, a_i)) = \sum_{[a_{i-1}; a_i) \subset \psi^{-1}((-\infty; y))} p_i$$

(чи $F_Y(y) = P_Y((-\infty;\, y)) = P_X(\psi^{-1}(-\infty\,; y)) = \sum\limits_{(a_{i-1}; a_i] \subset \psi^{-1}(-\infty; y)} P_X((a_{i-1};\, a_i])$).

В разі, коли множина

$$\Omega_Y = \psi(\Omega_X) = \psi(\bigcup_{i=1}^{k}[a_{i-1};\ a_i)) = \bigcup_{i=1}^{k}\psi([a_{i-1};\, a_i))$$

неперервна, тоді для визначення функції $F_Y(y)$ потрібно визначити ймовірності $P_Y(\psi([a_{i-1};\, a_i))) = P_X(\psi^{-1}(\psi([a_{i-1};\ a_i])))$ попадання в інтервали $\psi([a_{i-1};\, a_i))$ і далі визначати функцію $F_Y(y)$ поінтервального розподілу ймовірностей за інтервалами $\psi([a_{i-1};\, a_i))$ за правилами, наведеними в § 2.6.

В разі, коли розглядаються інтервали типу $(a_{i-1}, a_i]$, функція $F_Y(y)$ визначається аналогічно до попереднього.

Якщо розподіл ймовірностей на множині значень випадкової величини X абсолютно неперервний і відома щільність $f_X(x)$ цього розподілу ймовірностей, то

$$F_Y(y) = P_Y((-\infty;\ y)) = P_X(\psi^{-1}((-\infty;\ y))) = \int\limits_{\psi^{-1}((-\infty; y))} f_X(x)\ dx\,.$$

Зауважимо, що прообразом $\psi^{-1}((-\infty;\, y))$ множини виду $(-\infty;\, y)$ є не що інше, як множина розв'язків нерівності $\psi(x) < y$, прообразом окремої точки y є множина розв'язків рівняння $\psi(x) = y$, прообразом інтервала $[y_1;\, y_2)$ є множина розв'язків нерівності $y_1 \le \psi(x) < y_2$ або системи нерівностей $\{y_1 \le \psi(x)$, $\psi(x) < y_2\}$.

Якщо існує невід'ємна борелівська функція $f_Y(y)$ така, що $F_Y(y) = \int\limits_{-\infty}^{y} f_Y(y)\, dy$, то випадкова величина Y абсолютно неперервна, а $f_Y(y)$ – щільність розподілу ймовірностей на множині значень випадкової величини Y. Якщо $f_Y(y)$ неперервна, то $F_Y(y)$ диференційовна і

$$f_Y(y) = \frac{d}{dy} F_Y(y) .$$

Приклад 4.5.1. Задано випадкову величину X з рівномірним розподілом імовірностей на відрізку $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, що описується через щільність розподілу ймовірностей

$$f_X(x) = \begin{cases} \frac{1}{\pi}, & \text{коли } |x| \le \frac{\pi}{2}; \\ 0, & \text{коли } |x| > \frac{\pi}{2}. \end{cases}$$

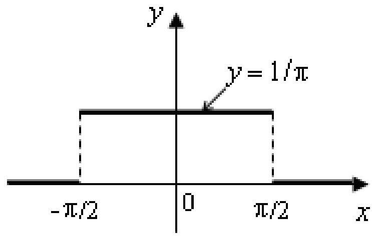


Рис. 4.5.1

Знайти розподіл імовірностей на множині значень випадкової величини $Y = \psi(X) = \cos(X)$.

Графік функції $f_X(x)$ подано на Рис. 4.5.1.

Визначаючи прообрази множин $(-\infty; y)$ за різних $y \in (-\infty; \infty)$ та ймовірності попадання значень X у такі прообрази, дістанемо (Рис. 4.5.2)

$$F_Y(y) = P_Y((-\infty;\ y)) = P_X(\psi^{-1}((-\infty;\ y))) = 0 \text{, коли } y \le -1 ,$$

оскільки прообраз $\psi^{-1}((-\infty; y))$, де $\psi(x) = \cos x$, за $y \le -1$ буде порожнім (значення $y = \cos x$ не можуть бути меншими за -1).

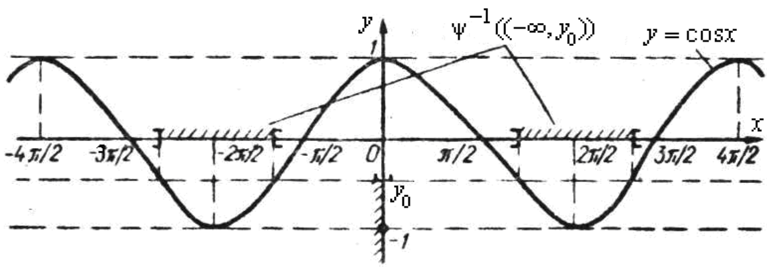


Рис. 4.5.2

Якщо $y > -1$, але $y \le 0$, то прообраз $\psi^{-1}((-\infty; y))$ множини $(-\infty;\, y)$ за відображення $y = \cos x$ вже не буде порожнім, проте

$\int\limits_{\psi^{-1}((-\infty;\, y))} f_X(x)\, dx = 0$, оскільки $f_X(x) = 0$, коли $x \in \psi^{-1}((-\infty;\, y))$, $y \in (-1;\, 0]$ (див. Рис. 4.5.2).

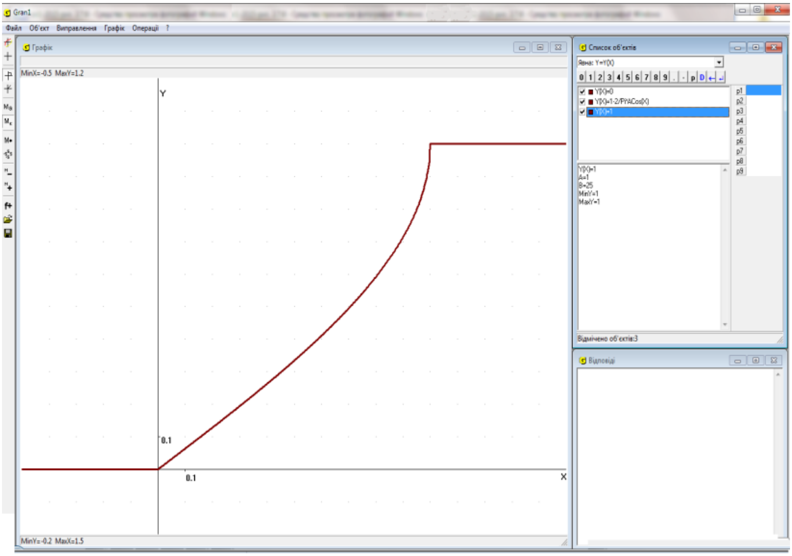


Рис. 4.5.3

Продовжуючи міркування, дістанемо

$$F_Y(y) = \begin{cases} 0, & \text{коли } y \le 0; \\ 1 - \dfrac{2}{\pi}\arccos y, & \text{коли } 0 \le y \le 1; \\ 1, & \text{коли } 1 \le y. \end{cases}$$

Графік функції $F_Y(y)$ зображено на Рис. 4.5.3. Графік щільності розподілу ймовірностей

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} \dfrac{2}{\pi} \dfrac{1}{\sqrt{1 - y^2}}, & \text{коли } y \in (0;\, 1); \\ 0, \ \text{коли } y \,\overline{\in}\, (0; 1) \end{cases}$$

зображено на Рис. 4.5.4.

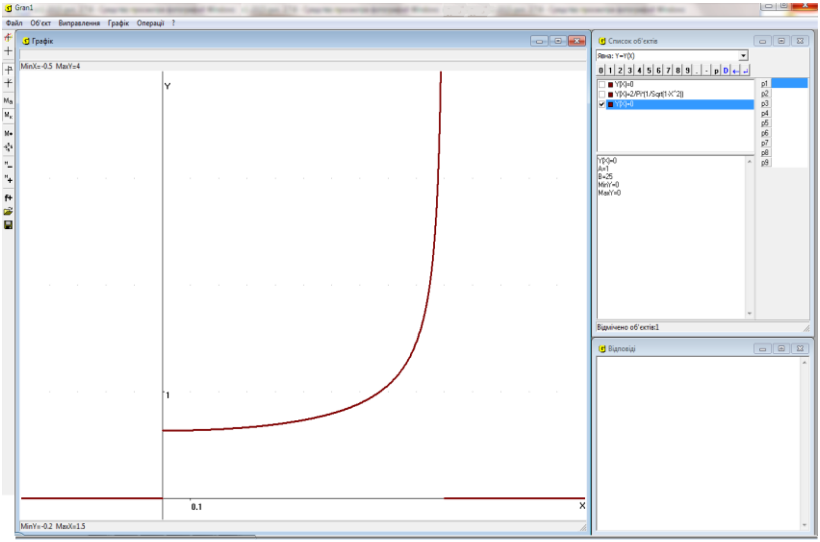


Рис. 4.5.4

Вправи для самостійного виконання

4.5.1. Перевірити, чи правильні твердження:

1. Якщо $\psi(x)$ – довільна дійсна функція, а $X(E)$ – випадкова величина, то $\psi(X(E))$ – випадкова величина.

2. Твердження 1 є правильним, коли $\psi(x)$ – неперервна на R^1 дійсна функція.

3. Якщо $\psi: R^1 \to R^1$ – неперервна функція, то вона є борелівською функцією.

4. Якщо множина Ω_X значень випадкової величини $X(E)$ – дискретна, то й множина Ω_Y значень випадкової величини $\psi(X(E))$ також дискретна.

5. Якщо розподіл ймовірностей на множині Ω_X значень випадкової величини $X(E)$ – абсолютно неперервний, а функція $\psi: R^1 \to R^1$ неперервна, то і розподіл ймовірностей на множині Ω_Y значень випадкової величини $\psi(X(E))$ – абсолютно неперервний.

6. Якщо розподіл ймовірностей на множині Ω_X значень випадкової величини $X(E)$ – неперервний, а $\psi: R^1 \to R^1$ – борелівська функція, то і розподіл ймовірностей на множині Ω_Y значень випадкової величини $\psi(X(E))$ – неперервний.

7. Якщо розподіл ймовірностей на множині значень випадкової величини $X(E)$ – дискретний, то розподіл ймовірностей на множині

значень випадкової величини $Y(E)=\psi(X(E))$ також дискретний.

4.5.2. Нехай на множині значень випадкової величини $X(E)$, $E\in\Omega$, розподіл ймовірностей дискретний і заданий через ряд розподілу:

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

1. Довести, що за будь-якої функції $\psi(x)$, в області визначення якої містяться усі точки x_k, $k\in\overline{1,n}$, розподіл ймовірностей на множині значень функції $Y(E)=\psi(X(E))$, $E\in\Omega$, дискретний, та знайти цей розподіл.

2. Чи є правильним твердження 1, якщо в області визначення функції $\psi(x)$ не міститься принаймні одне значення x_k випадкової величини $X(E)$.

4.5.3. 1. Для заданого розподілу ймовірностей на множині значень дискретної випадкової величини $X(E)$, $E\in\Omega$,

x_i	-1	0	1	2	3
p_i	0.1	0.2	0.1	0.45	0.15

знайти (якщо це можливо) розподіл ймовірностей на множині значень випадкової величини $Y(E)=\psi(X(E))$, $E\in\Omega$, коли:

1) $\psi(x)=x^2$;

2) $\psi(x)=\sqrt{x}$;

3) $\psi(x)=x(x^2-1)(x^2-5x+6)$;

4) $\psi(x)=\begin{cases}\dfrac{|x|}{x}, & \text{коли } x\neq 0,\\ 0, & \text{коли } x=0;\end{cases}$

5) $\psi(x)=[x]$ – ціла частина x;

6) $\psi(x)=\{x\}$ – дробова частина x;

7) $\psi(x)=\cos\pi x$;

8) $\psi(x)=\ln(x^2-5x+6)$.

2. Знайти функцію $F_Y(y)$ розподілу ймовірностей на множині значень випадкової величини $Y(E)=\psi(X(E))$, $E\in\Omega$.

4.6. Випадкові вектори. Розподіли ймовірностей на множинах значень випадкових векторів

Часто на практиці доводиться аналізувати об'єкти, що описуються не однією, а кількома випадковими величинами або й більш загальними випадковими характеристиками. Прикладами таких випадкових об'єктів є випадкові "точки", випадкові вектори, випадкові функції, випадкові процеси. В зв'язку з цим доцільно ввести поняття випадкового елемента загальної природи.

Означення. Нехай (Ω, S) і (W, Q) – два вимірні простори. Функція $X = X(E)$, яка визначена на Ω і набуває значень із W, називається S/Q-*вимірною функцією* або *випадковим елементом* із значеннями в W, якщо за довільного $B \in Q$

$$\{E \mid X(E) \in B\} = X^{-1}(B) \in S .$$

Випадкові елементи із значеннями із W називають також W-*значними випадковими величинами*.

Якщо $(W, Q) = (R^1, \mathscr{B}(R^1))$, то випадковий елемент є випадковою величиною.

Нехай $(W, Q) = (R^n, \mathscr{B}(R^n))$. Тоді випадковий елемент $X(E)$ є випадковою "точкою" в R^n. Якщо X_k – проекція точки X із R^n на k-ту координатну вісь, то $X(E)$ можна подати у вигляді

$$X(E) = (X_1(E),\ X_2(E),\ \dots\ , X_n(E)).$$

Тут $X_i(E)$, $i = \overline{1, n}$, є звичайними випадковими величинами.

Справді, за довільного $B \in \mathscr{B}(R^1)$

$$\{E \mid X_k(E) \in B\} = X_k^{-1}(B) =$$

$$= \{E \mid X_1(E) \in R^1, \dots, X_{k-1}(E) \in R^1, X_k(E) \in B, X_{k+1}(E) \in R^1, \dots , X_n(E) \in R^1\} =$$

$$= \{E \mid X(E) \in (R^1 \times \dots \times R^1 \times B \times R^1 \times \dots \times R^1)\} \in S ,$$

оскільки множина

$$R^1 \times \dots \times R^1 \times B \times R^1 \times \dots \times R^1 \in \mathscr{B}(R^n) .$$

Означення. Будь-який скінченний впорядкований набір із n випадкових величин $(X_1(E),\ X_2(E),\ \dots,\ X_n(E))$, $E \in \Omega$, називається n-*вимірним випадковим вектором на ймовірнісному просторі* $(\Omega,\ S,\ P)$.

Іноді випадковий вектор називають також *системою випадкових величин* $X_1,\ X_2,\ \dots\ , X_n$.

Згідно з наведеним означенням будь-який випадковий елемент із значеннями в R^n (R^n-значна випадкова величина) є n-вимірним випадковим вектором.

Правильне й обернене твердження:
будь-який випадковий вектор

$$X(E)=(X_1(E),\ X_2(E),\ \dots\ ,X_n(E))$$

є випадковим елементом в R^n.

Справді, якщо $B_k \in \mathscr{B}(R^1)$, $k=\overline{1,\ n}$, то

$$\{E \mid X(E)\in(B_1\times B_2\times\ldots\times B_n)\}=\bigcap_{k=1}^{n}\{E \mid X_k(E)\in B_k\}=\bigcap_{k=1}^{n} X_k^{-1}(B_k)\in S.$$

Проте найменша σ-алгебра, в якій містяться множини $B_1\times B_2\times\ldots\times B_n$, збігається з $\mathscr{B}(R^n)$. Отже, за довільного $B\in\mathscr{B}(R^n)$ множина $\{E \mid X(E)\in B\}=X^{-1}(B)$ належить S.

Означення. Нехай X є n-вимірним випадковим вектором на ймовірнісному просторі $(\Omega,\ S,\ P)$. Задання ймовірнісної міри P_X на $(R^n,\ \mathscr{B}(R^n))$, визначеної за рівністю

$$P_X(B)=P(\{E \mid X(E)\in B\})=P(X^{-1}(B)),\ B\in\mathscr{B}(R^n),$$

є заданням *розподілу ймовірностей на множині значень n-вимірного випадкового вектора X*.

Надалі запис $x^{(1)}\le x^{(2)}$ означає, що $x_i^{(1)}\le x_i^{(2)}$, $i\in\overline{1,\ n}$.

Множину

$$[a;\,b)=\{x \mid a\le x<b\}=\{(x_1,\ x_2,\ \ldots,x_n) \mid a_1\le x_1<b_1,\ \ldots,a_n\le x_n<b_n\}$$

називають n-*вимірним інтервалом* (n-інтервалом).

Означення. Функцією розподілу ймовірностей на множині $\Omega_X=X(\Omega)$ значень випадкового вектора X називається функція

$$F_X(x)=P_X((-\infty;\ x))=P(\{E \mid X(E)\in(-\infty;\ x)\})=P(X^{-1}(-\infty;x)),$$

$x\in R^n$, тобто

$$F_X(x_1,x_2,\ldots,x_n)=$$
$$=P(\{E \mid (X_1(E),\ X_2(E),\ \ldots,X_n(E))\in(-\infty;\ x_1)\times(-\infty;\ x_2)\times\ldots\times(-\infty;\ x_n)\})=$$
$$=P(\{E \mid X_1(E)\in(-\infty;\ x_1),\ X_2(E)\in(-\infty;\ x_2),\ \ldots,X_n(E)\in(-\infty;\ x_n)\}).$$

Стосовно функції розподілу ймовірностей задовільняються такі властивості.

1. $0\le F_X(x)\le 1$.

2. $F_X(x)$ – неспадна, тобто якщо $x^{(1)}<x^{(2)}$, то $F_X(x^{(1)})\le F_X(x^{(2)})$.

3. Якщо за деякого k $x_k\to-\infty$, то $\lim\limits_{x_k\to-\infty}F_X(x)=0$.

4. Якщо $x_1\to\infty$, $x_2\to\infty$, $\ldots$, $x_n\to\infty$, то $\lim\limits_{x\to\infty}F_X(x)=1$.

5. $F_X(x)$ неперервна зліва, тобто якщо $x^{(k)} < x$ і $\lim\limits_{k\to\infty} x^{(k)} = x$, то

$$\lim_{k\to\infty} F_X(x^{(k)}) = F_X(x).$$

Доведення цих властивостей аналогічні наведеним раніше (див. Розділ 2, §2.5-§2.7). Якщо $[a;b)$ є n-інтервалом, тобто $[a;\, b) = [a_1;\, b_1) \times [a_2;\, b_2) \times \ldots \times [a_n;\, b_n)$, то знаючи функцію розподілу ймовірностей $F_X(x)$ і використовуючи введений раніше різницевий оператор

$$\Delta_{a_i b_i} F_X(x) = F_X(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots, x_n) - F_X(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n),$$

знаходимо

$$P_X([a;b)) = \Delta_{a_1 b_1} \Delta_{a_2 b_2} \ldots \Delta_{a_n b_n} F_X(x).$$

Наприклад, за $n = 2$

$$P_X([a_1;\, b_1) \times [a_2;\, b_2)) =$$
$$= F_X(b_1,\, b_2) - F_X(a_1,\, b_2) - F_X(a_2,\, b_1) + F_X(a_1,\, a_2).$$

Таким чином, за функцією розподілу ймовірностей $F_X(x)$ однозначно визначається міра $P_X(B)$ на n-інтервалах.

Якщо B – довільна борелівська множина з R^n, то

$$P_X(B) = \inf_{\bigcup\limits_k [a_k;b_k) \supset B} \Big(\sum_{\bigcup\limits_k [a_k;b_k) \supset B} P_X([a^{(k)};\, b^{(k)})) \Big),$$

де $\inf$ – точна нижня грань на всіх можливих покриттях множини B n-інтервалами, $(B \subset \bigcup\limits_k [a^{(k)},\, b^{(k)}))$ (порівняйте з §2.2, §2.6).

Таким чином, через функцію розподілу ймовірностей $F_X(x)$ однозначно задається розподіл ймовірностей $P_X(B)$ на множині значень випадкового вектора X. До того ж справджується таке твердження.

Нехай $F_X(x)$, $x \in R^n$, – довільна функція, стосовно якої задовільняються умови 1–5.

Тоді на $\mathscr{B}(R^n)$ можна, і притому єдиним способом, визначити міру $P_X(B)$ таку, що

$$P_X([a;b)) = \Delta_{a_1 b_1} \Delta_{a_2 b_2} \ldots \Delta_{a_n b_n} F_X(x).$$

Довільна функція $F_X(x)$, стосовно якої задовільняються умови 1–5, називається *функцією розподілу ймовірностей на множині значень випадкового вектора* X, а через міру $P_X(B)$

задається розподіл імовірностей, відповідний функції розподілу $F_X(x)$.

Якщо множина $\Omega_X = X(\Omega)$ значень випадкового вектора дискретна, тобто скінченна або зчисленна:

$$\Omega_X = X(\Omega) = \{x^{(1)}, x^{(2)}, \ldots, x^{(k)}, \ldots\},$$

то розподіл імовірностей на множині значень цього вектора можна задати, вказавши набір чисел

$$p_k = P_X(\{x^{(k)}\}) = P(\{E \mid E \in X^{-1}(x^{(k)})\}) = P(X^{-1}(x^{(k)})),$$

таких, що $p_k \geq 0$, $\sum\limits_k p_k = 1$. Тоді за довільної множини B з $\mathscr{B}(R^n)$

$$P_X(B) = \sum_{x^{(k)} \in B} p_k .$$

Нехай, наприклад, (x_j, y_i), $i \in \overline{1, m}$, $j \in \overline{1, l}$, – множина значень двохвимірного випадкового вектора (X, Y) (Рис. 4.6.1). Тоді розподіл ймовірностей можна задати за таблицею виду Табл. 4.6.1. Ймовірність p_{ij} дорівнює ймовірності того, що елементарна подія E належатиме повному прообразу точки (x_j, y_i) за відображенням X вихідної множини Ω елементарних подій на R^2.

Табл. 4.6.1

$y_i \backslash x_j$	x_1	x_2	$\ldots$	x_l
y_1	p_{11}	p_{12}	$\ldots$	p_{1l}
y_2	p_{21}	p_{22}	$\ldots$	p_{2l}
$\vdots$	$\vdots$	$\vdots$	$\vdots\ \vdots\ \vdots$	$\vdots$
p_{1m}	p_{m1}	p_{m2}	$\ldots$	p_{ml}

Механічний аналог дискретного двохвимірного розподілу ймовірностей полягає в тому, що одинична маса певним чином розподіляється на множині точок $\{(x_1, y_1), (x_2, y_1), \ldots, (x_l, y_m)\}$ так, що на точку (x_j, y_i) припадає маса p_{ij}.

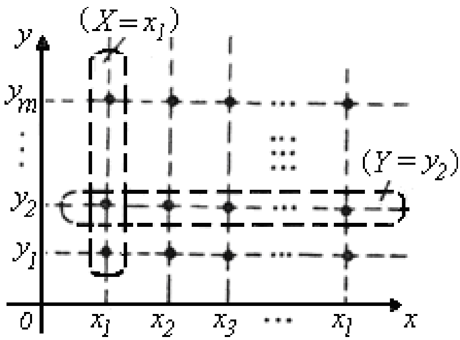


Рис. 4.6.1

Оскільки в результаті випробування кожна з випадкових величин X і Y обов'язково набуває одного із своїх можливих значень, то

$$P_Y(\{y_i\}) = P_{(X,Y)}(\{y_i\})(\{x_1\}) + (\{x_2\}) + \ldots + (\{x_l\})) =$$

$$= P_{(X,Y)}((\{y_i\})(\{x_1\})) + P_{(X,Y)}((\{y_i\})(\{x_2\})) + \ldots + P_{(X,Y)}((\{y_i\})(\{x_l\})) =$$

$$= p_{i1} + p_{i2} + \ldots + p_{il} = \sum_{j=1}^{l} p_{ij} \,,$$

$$\sum_{i=1}^{m} P_Y(\{y_i\}) = \sum_{i=1}^{m}\sum_{j=1}^{l} p_{ij} = 1 .$$

Аналогічно

$$P_X(\{x_j\}) = \sum_{i=1}^{m} p_{ij} \,, \;\; \sum_{j=1}^{l} P_X(\{x_j\}) = \sum_{j=1}^{l}\sum_{i=1}^{m} p_{ij} = 1 .$$

Таким чином, за дискретним розподілом імовірностей на множині значень двохвимірного випадкового вектора (X,Y), можна знайти ймовірності $P_X(\{x_j\}) = \sum_{i=1}^{m} p_{ij}$, $j = 1, 2, \ldots, l$, і $P_Y(\{y_i\}) = \sum_{j=1}^{l} p_{ij}$, $i = 1, 2, \ldots, m$, тобто розподіли ймовірностей на множинах значень кожної з випадкових величин X і Y.

Стосовно функції розподілу ймовірностей $F_{(X,Y)}(x, y)$ одержимо:

$$F_{(X,Y)}(x, y) = P_{(X,Y)}((-\infty; x) \times (-\infty; y)) = \sum_{x_j < x}\sum_{y_i < y} p_{ij} \,.$$

В геометричній інтерпретації $F_{(XY)}(x_0, y_0)$ є ймовірністю того, що випадкова точка (X,Y) лежить у нескінченному прямому куті з вершиною $(x_0,\ y_0)$ і сторонами, паралельними до осей координат, який розміщений строго зліва від прямої $x = x_0$ і строго нижче від прямої $y = y_0$ (точки межових прямих $x = x_0$ і $y = y_0$ до розглядуваного кута не включаються) (Рис. 4.6.2).

В механічній інтерпретації через значення функції розподілу $F_{(X,Y)}(x_0,\ y_0)$ визначають, яка загальна маса лежить строго зліва від прямої $x = x_0$ і строго нижче від прямої $y = y_0$, тобто яка маса припадає на множину $W = \{(x,\ y) \mid x < x_0,\ y < y_0\}$ або на перетин множин

$$\{(x,\ y) \mid x \in (-\infty;\ x_0),\ y \in (-\infty;\ \infty)\}$$

і

$$\{(x;\ y) \mid x \in (-\infty;\ \infty),\ y \in (-\infty;\ y_0)\}$$

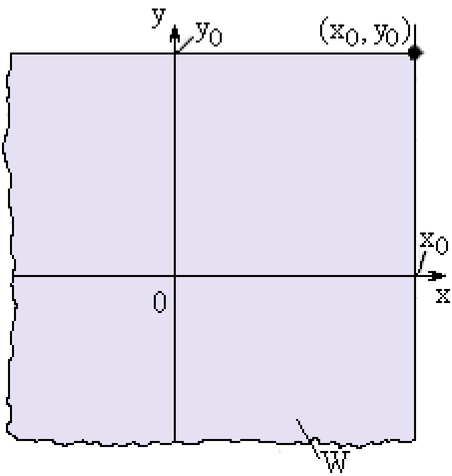


Рис. 4.6.2

(див. Рис. 4.6.2), за умови, що на площині xOy деяким чином розподілена одинична маса.

Якщо відома функція розподілу ймовірностей $F_{(X,Y)}(x,\ y)$, то ймовірність того, що значення випадкового вектора (X,Y) лежать у прямокутнику

$$D = \{(x,\ y) \mid x_1 \le x < x_2,\ y_1 \le y < y_2\},$$

обчислюється за формулою

$$P_{(X,Y)}(D) = \Delta_{x_1 x_2} \Delta_{y_1 y_2} F_{(X,Y)}(x,\ y) =$$

$$= F_{(X,Y)}(x_2, y_2) - F_{(X,Y)}(x_1, y_2) - F_{(X,Y)}(x_2, y_1) + F_{(X,Y)}(x_1, y_1) \quad (4.6.1)$$

Приклад 4.6.1. Розподіл імовірностей на скінченній множині значень R^2-значної випадкової величини (X,Y), поданий у вигляді в Табл. 4.6.2. Побудувати функцію розподілу ймовірностей $F_{(XY)}(x, y)$. Знайти ймовірність $P_{(XY)}((x, y) \in R)$, де

$$R = \{(x, y) \mid 0.5 \le x < 2.5;\ 0.5 \le y < 1.5\}.$$

Табл. 4.6.2

y_i \ x_j	1	2	3
0	0.2	0.1	0.2
1	0.1	0.1	0.3

Враховуючи означення функції розподілу ймовірностей, дістаємо (Рис. 4.6.3)

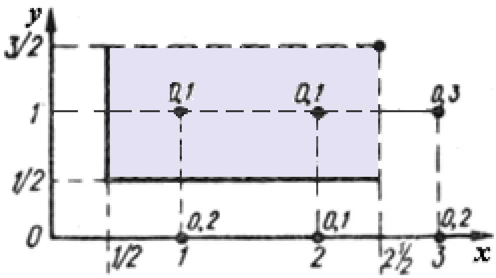


Рис. 4.6.3

$$F_{(X,Y)}(x, y) = \begin{cases} 0, & \text{коли } x \le 1 \text{ або } y \le 0; \\ 0.2, & \text{коли } 1 < x \le 2 \text{ і } 0 < y \le 1; \\ 0.3, & \text{коли } 1 < x \le 2 \text{ і } 1 < y \text{ або } 2 < x \le 3 \text{ і } 0 < y \le 1; \\ 0.5 & \text{коли } 2 < x \le 3 \text{ і } 1 < y \text{ або } 3 < x \text{ і } 0 < y \le 1; \\ 1, & \text{коли } 3 < x \text{ і } 1 < y. \end{cases}$$

На Рис. 4.6.4 зображено ступіньчату поверхню $z = F_{(XY)}(x, y)$.

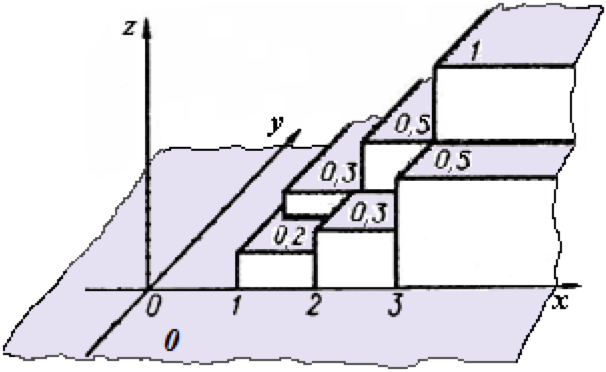


Рис. 4.6.4

Обчислюючи ймовірність $P_{(XY)}(R)$ за формулою (4.6.1), знаходимо

$$P_{(XY)}(R) = F_{(XY)}(2.5;1.5) - F_{(XY)}(2.5;\,0.5) - F_{(XY)}(0.5;1.5) + F_{(XY)}(0.5;\,0.5) =$$
$$= 0.5 - 0.3 - 0.0 + 0.0 = 0.2\,.$$

Якщо треба обчислити ймовірність попадання випадкової точки $(X,\,Y)$ в довільну область D, то таку ймовірність можна знайти наближено, апроксимуючи (наближаючи) область D деякою сукупністю прямокутників (Рис. 4.6.5) (порівняйте з §2.2, §2.6).

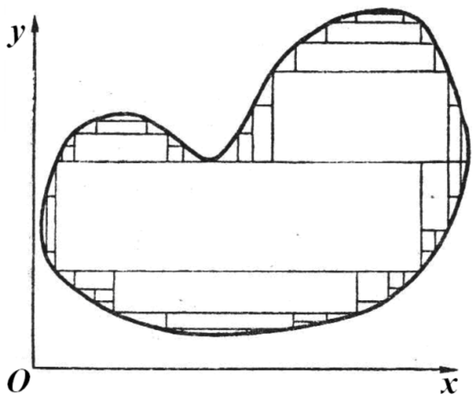


Рис. 4.6.5

Якщо існує така невід'ємна борелівська функція $f_X(x)$, що за довільного $B \in \mathscr{B}(R^n)$

$$P_X(B) = \int_B f_X(x)\,dx,$$

де

$$x = (x_1,\ x_2,\ \dots,\ x_n),\ dx = dx_1 dx_2 \dots dx_n,$$

$$\int_B f_X(x)\,dx = \int\int_B \dots \int f_X(x_1, x_2, \dots, x_n)\,dx_1 dx_2 \dots dx_n,$$

то розподіл імовірностей $P_X(B)$, $B \in \mathscr{B}(R^n)$, називається *абсолютно неперервним*. Функція $f_X(x)$ називається *щільністю розподілу* ймовірностей на множині значень випадкового вектора X (R^n-значної випадкової величини).

В даному разі буде

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx,$$

тобто

$$F_{(X_1, X_2,\dots, X_n)}(x_1, x_2,\dots, x_n) =$$
$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\dots\int_{-\infty}^{x_n} f_{(X_1,X_2,\dots, X_n)}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

Якщо $f_X(x_1,\ x_2,\ \dots,\ x_n)$ неперервна, то стосовно $F_X(x_1,\ x_2,\ \dots,\ x_n)$ існує n-та мішана похідна

$$\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\dots\frac{\partial}{\partial x_n} F_X(x_1,\ x_2,\ \dots,\ x_n) = f_X(x_1,\ x_2,\ \dots,\ x_n).$$

В механічній інтерпретації (в R^2) $f_{(X,Y)}(x,\ y)$ є щільністю в точці $(x,\ y)$ одиничної маси, розподіленої певним чином на площині *xOy*,

$$P_{(X,Y)}(G) = \iint_G f_{(X,\,Y)}(x,\ y)\ dxdy$$

є масою, що припадає на область G, коли на площині *xOy* розподілена деяка маса із щільністю $f_{(X,Y)}(x, y)$.

В геометричній інтерпретації $P_{(X,Y)}(G) = \iint_G f_{(X,\,Y)}(x,\ y)\ dxdy$ є об'єм під поверхнею щільності над областю G (Рис. 4.6.6).

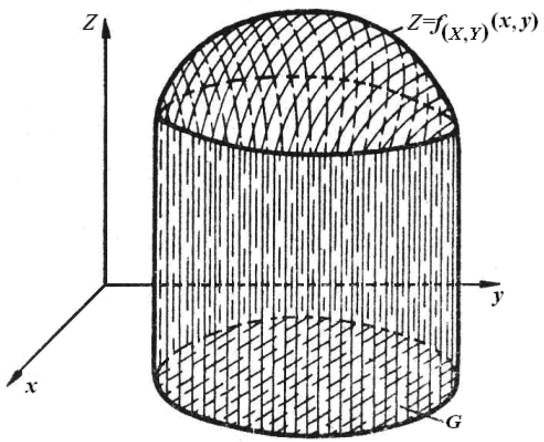


Рис. 4.6.6

Вправи для самостійного виконання

4.6.1. Перевірити, чи правильні твердження:

1. Якщо $X(E)$ – випадковий елемент, то $X(E)$ – випадкова величина.
2. Твердження, обернене до 1, є правильним.
3. Кожен випадковий вектор є випадковим елементом.
4. Щоб випадковий елемент $X(E),\, E \in \Omega$, був випадковим вектором, необхідно й досить щоб мало місце включення

$$X(E) = (X_1(E), X_2(E), ..., X_n(E)) \in R^n,$$

за кожного $E \in \Omega$ і $X_i(E),\ E \in \Omega,\ i = \overline{1, n}$, були випадковими величинами.

5. Кожному випадковому вектору відповідає єдиний розподіл ймовірностей.
6. За будь-яких точок $x^{(1)} = (x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)})$ та $x^{(2)} = (x_1^{(2)}, x_2^{(2)}, ..., x_n^{(2)})$ з простору R^n завжди або $x^{(1)} \le x^{(2)}$, або $x^{(1)} \ge x^{(2)}$.
7. Твердження 6 є правильним тоді і тільки тоді, коли $n = 1$.
8. Кожному випадковому вектору відповідає єдина функція розподілу ймовірностей.
9. Функція розподілу ймовірностей цілком визначається за своїми характеристичними властивостями.
10. Через функцію розподілу ймовірностей на множині значень випадкового вектора визначається єдина міра на $\mathscr{B}(R^n)$.
11. Кожен розподіл ймовірностей на множині значень випадкового вектора є дискретним або неперервним.
12. Якщо розподіл ймовірностей на множині значень випадкового вектора неперервний, то він і абсолютно неперервний.
13. Твердження, обернене до 12, є правильним.

14. Кожному розподілу ймовірностей відповідає щільність розподілу.

15. Щільність розподілу ймовірностей є невід’ємною та неперервною функцією.

4.7. Залежні і незалежні випадкові величини відносно ймовірнісної міри P

Нехай $X=(X_1,\ X_2,\ \dots,X_n)$ – випадковий вектор з функцією розподілу ймовірностей $F_X(x)=F_X(x_1,\ x_2,\ \dots,\ x_n)$. Розглянемо вектор

$$\tilde{X}=\{X_1,\ X_2,\ \dots,X_s\},\ s<n.$$

Тоді функція розподілу $F_{\tilde{X}}(\tilde{x})$ виражається через $F_X(x)$ так:

$$F_{\tilde{X}}(x_1, x_2, \dots, x_s)=P_{\tilde{X}}((-\infty;x_1)\times(-\infty;x_2)\times\dots\times(-\infty;x_s))=$$
$$P_X((-\infty;x_1)\times\dots\times(-\infty;x_s)\times(-\infty;\infty)\times\dots\times(-\infty;\infty))=$$
$$=F_X(x_1, x_2, \dots, x_s, \infty, \dots, \infty).$$

Якщо $\tilde{B}$ – борелівська множина в R^s, а B – множина всіх точок $x=(x_1,\ x_2,\ \dots,\ x_n)$ з R^n, таких що $(x_1,\ x_2,\ \dots,\ x_s)\in\tilde{B}$, то очевидно, що множини

$$\{E\,|\ \tilde{X}(E)\in\tilde{B}\}=\tilde{X}^{-1}(\tilde{B})\ \text{ і }\ \{E\,|\ X(E)\in B\}=X^{-1}(B)$$

співпадають. Тому

$$P_{\tilde{X}}(\tilde{B})=\{E\,|\ \tilde{X}(E)\in\tilde{B}\}=P_X(\{E\,|\ X(E)\in B\})=P_X(B).$$

Якщо $f_X(x)$ – щільність розподілу ймовірностей на множині значень випадкового вектора X, то

$$P_{\tilde{X}}(\tilde{B})=P_X(B)=$$

$$=\int\underset{\tilde{B}}{\dots}\int\Big(\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}f_X(x_1,\ \dots,x_s,\ x_{s+1},\ \dots,\ x_n)dx_{s+1}\ \dots\ dx_n\Big)\,dx_1\ \dots\ dx_s.$$

Отже, існує функція

$$f_{\tilde{X}}(x_1,\ x_2,\ \dots,\ x_s)=\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}f_X(x_1,\ \dots,\ x_s,\ x_{s+1},\ \dots,\ x_n)dx_{s+1}\ \dots\ dx_n$$

така, що

$$P_{\tilde{X}}(\tilde{B})=\int\int\underset{\tilde{B}}{\dots}\int f_{\tilde{X}}(x_1,x_2,\ \dots,\ x_s)\,dx_1dx_2\dots dx_s\ ,$$

тобто розподіл $P_{\tilde{X}}(\tilde{B})$ абсолютно неперервний із щільністю $f_{\tilde{X}}(x_1,x_2,\ \dots,\ x_s)$, якщо ймовірності на множині значень R^n-значної випадкової величини X розподілені із щільністю $f_X(x_1,x_2,\ \dots,\ x_n)$.

Множину B у розглядуваному випадку називають *циліндричною множиною* в R^n з основою $\tilde{B}$ в R^s.

Нехай, наприклад, розглядаються R^2-значні випадкові величини (двохвимірні випадкові вектори). Стосовно щільності $f_{(X,Y)}(x,y)$ двохвимірного розподілу ймовірностей задовільняються такі властивості.

1. $f_{(X,Y)}(x,y) \ge 0$ за будь-яких x і y. Ця властивість випливає з означення щільності розподілу ймовірностей в R^n.

2. $\iint\limits_{R^2} f_{(X,Y)}(x,\ y)\,dxdy = \int\limits_{-\infty}^{\infty}\left\{\int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,\ y)\,dy\right\}dx = 1,$

де $R^2 = \{((x,\ y)\mid\ x \in (-\infty;\ \infty), y \in (-\infty;\ \infty)\}$.

Справді,

$$\iint\limits_{R^2} f_{(X,Y)}(x,\ y)\,dxdy = \int\limits_{-\infty}^{\infty}\left\{\int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,\ y)\,dy\right\}dx = F_{(X,Y)}(+\infty;\ +\infty) = 1\cdot$$

В механічній інтерпретації ця властивість означає, що загальна маса, розподілена на всій площині *xOy*, дорівнює одиниці.

В геометричній інтерпретації це означає, що об'єм тіла, яке лежить під поверхнею $z = f_{(X,Y)}(x,y)$ над площиною $z = 0$, повинен дорівнювати одиниці.

3. $\int\limits_{-\infty}^{\infty} f_{(X,\,Y)}(x,y)\,dy = f_X(x)$, де $f_X(x)$ – щільність одновимірного розподілу ймовірностей на множині значень випадкової величини X, що є складовою випадкового вектора $(X,\ Y)$.

В геометричній інтерпретації $f_X(x_0) = \int\limits_{-\infty}^{\infty} f_{(X,\,Y)}(x_0,\ y)\,dy$ є площа перерізу площиною $x = x_0$ тіла, обмеженого поверхнями $z = f_{(X,Y)}(x,y)$ і $z = 0$ (Рис. 4.7.1).

4. $\int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,y)dx = f_Y(y)$, де $f_Y(y)$ – щільність одновимірного розподілу ймовірностей на множині значень випадкової величини Y, що є складовою вектора $(X,\ Y)$.

В геометричній інтерпретації $f_Y(y_0) = \int\limits_{-\infty}^{\infty} f_{(X,\,Y)}(x,\ y)\,dx$ є площа перерізу площиною $y = y_0$ тіла, обмеженого поверхнями $z = f_{(X,Y)}(x,y)$ і $z = 0$ (Рис. 4.7.2).

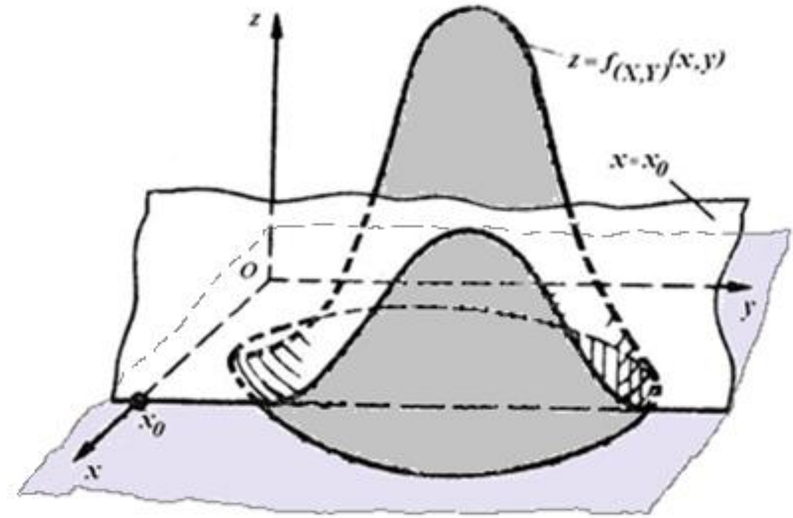


Рис. 4.7.1

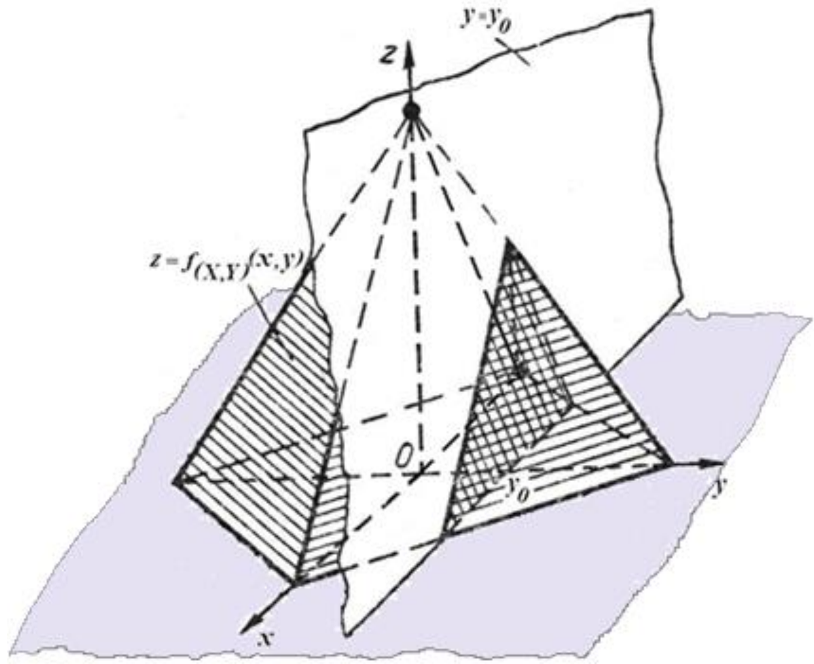


Рис. 4.7.2

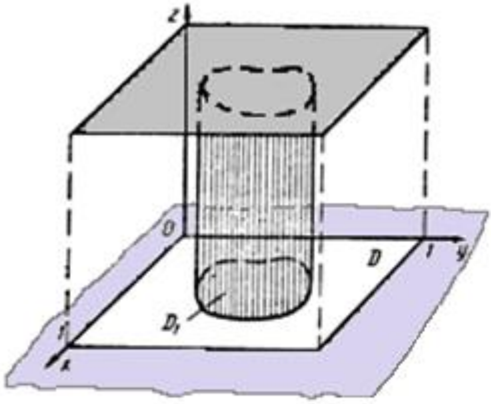


Рис. 4.7.3

Приклад 4.7.1. Розглянемо рівномірний двохвимірний розподіл імовірностей в прямокутнику $D=\{(x,y)\,\big|\ x\in[0;1], y\in[0;1]\}$ (Рис. 4.7.3). Щільність рівномірного розподілу ймовірностей в області D визначається за рівностями

$$f_{(X,Y)}(x,y)=\begin{cases} c, & \text{коли}\quad (x,y)\in D; \\ 0, & \text{коли}\quad (x,y)\,\overline{\in}\, D. \end{cases}$$

Враховуючи властивість 2 щільності розподілу ймовірностей $f_{(X,Y)}(x,y)$ та її геометричну інтерпретацію ($\iint\limits_D f_{(X,Y)}(x,\ y)\,dxdy$ є об'єм тіла під поверхнею $f_{(X,Y)}(x,y)$ над областю D, див. Рис. 4.7.3), слід зробити висновок, що у випадку рівномірного розподілі ймовірностей в області D стала c повинна дорівнювати $\dfrac{1}{\text{пл.}\,D}$, щоб об'єм $\iint\limits_D f_{(X,Y)}(x,\ y)\,dxdy$ дорівнював одиниці.

Якщо область $D_1 \subset D$ (множина D_1 є підмножиною множини D) (Рис. 4.7.4), то у випадку рівномірного розподілу ймовірностей в області D

$$P_{(X,\,Y)}(D_1)=\frac{\text{пл}\,.\,D_1}{\text{пл}\,.\,D}$$

(див. §3.6, геометричне задання ймовірності). Тут D_1 – випадкова подія, яка полягає в тому, що випадкова точка $(X,\ Y)$ попадає в область D_1.

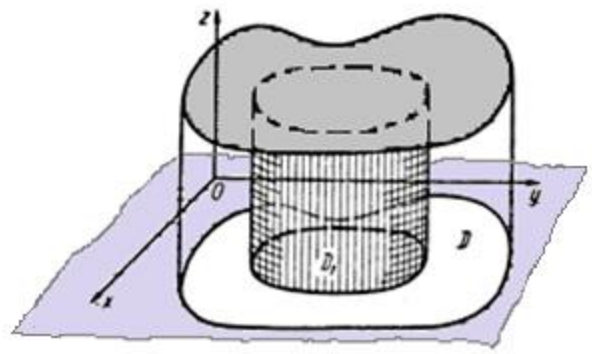


Рис. 4.7.4

Використовуючи формули

$$F_X(x) = F_{(X,Y)}(x; +\infty), \quad F_Y(y) = F_{(X,Y)}(+\infty; y),$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)\,dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)\,dx, \tag{4.7.1}$$

можна, знаючи розподіл ймовірностей на множині значень випадкового вектора (X, Y), знайти розподіли ймовірностей на множинах значень окремих його випадкових складових.

Природно, виникає питання про обернену задачу: чи можна, знаючи розподіли ймовірностей на множинах значень випадкових величин $X_1, X_2, \ldots, X_n$, знайти розподіл ймовірностей на множині значень випадкового вектора $(X_1, X_2, \ldots, X_n)$.

Виявляється, що в загальному випадку це зробити неможливо. Щоб повністю охарактеризувати розподіл ймовірностей на множині значень випадкового вектора $(X_1, X_2, \ldots, X_n)$, недосить знати розподіли ймовірностей на множинах значень кожної з його випадкових координат, треба знати ще так звані умовні розподіли ймовірностей.

Означення. Випадкові величини $X_1, X_2, \ldots, X_n$ називаються *незалежними* відносно ймовірнісної міри P, заданої на R^n, $S = \mathscr{B}(R^n)$, якщо якими б не були $x_1, x_2, \ldots, x_n$, події $X_1^{-1}((-\infty; x_1))$, $X_2^{-1}((-\infty; x_2))$, ..., $X_n^{-1}((-\infty; x_n))$ незалежні в сукупності, тобто

$$P\left(\prod_{i\in I} X_i^{-1}((-\infty; x_i))\right) = \prod_{i\in I} P(X_i^{-1}((-\infty; x_i))), \; I \subset \{1, 2, \ldots, n\}, \quad I \neq \varnothing .$$

Нехай X і Y – два випадкові вектори із значеннями в R^s і R^l відповідно,

$$X = (X_1, \ldots, X_s), \; Y = (Y_1, \ldots, Y_l).$$

Функцію $F_{(X,Y)}(x,y)$ розподілу ймовірностей на множині значень складеного вектора $(X,Y)=(X_1, \dots, X_s, Y_1, \dots, Y_l)$

$F_{(X,Y)}(x, y)=F_{(X,Y)}(x_1, ..., x_s, y_1, ..., y_l)=P_{(X,Y)}((-\infty; x)\times(-\infty; y))$,

де $x=(x_1, x_2, \dots, x_s)$, $y=(y_1, \dots, y_l)$, називають *функцією сумісного розподілу* ймовірностей на декартовому добутку множин значень випадкових векторів X і Y.

Випадкові вектори X і Y називаються *незалежними* відносно ймовірнісної міри $P_{(X,Y)}$, заданої на $R^s\times R^l$, $S_{(X,Y)}=\mathscr{B}(R^s\times R^l)$, якщо за довільних борелівських множин $B_X\in\mathscr{B}(R^s)$, $B_Y\in\mathscr{B}(R^l)$ випадкові події $\{E \mid X(E)\in B_X\}$ і $\{E \mid Y(E)\in B_Y\}$ незалежні, тобто

$$P_{(X,Y)}(\{E \mid X(E)\in B_X\}\cap\{E \mid Y(E)\in B_Y\})=$$

$$=P_{(X,Y)}(X^{-1}(B_X)\cap Y^{-1}(B_Y))=$$

$$=P_X(\{E \mid X(E)\in B_X\})P_Y(\{E \mid Y(E)\in B_Y\})=$$

$$=P_X(X^{-1}(B_X))P_Y(Y^{-1}(B_Y)).$$

Зокрема, якщо $B_X=\{x \mid x\in(-\infty;x_0)\}$, $B_Y=\{y \mid y\in(-\infty;y_0)\}$, то з умови незалежності відповідних подій випливає, що

$$F_{(X,Y)}(x,y)=F_X(x)F_Y(y).$$

Правильне й обернене твердження: якщо $F_{(X,Y)}(x,y)=F_X(x)F_Y(y)$, то випадкові вектори X і Y – незалежні.

Справді, нехай $F_{(X,Y)}(x,y)=F_X(x)F_Y(y)$ і $F_Y(y)>0$. Тоді розподіли в R^s

$$\frac{P_{(X,Y)}(E \mid X(E)\in B, Y(E)<y)}{F_Y(y)}=\frac{P_{(X,Y)}(X^{-1}(B)\cap Y^{-1}((-\infty; y))}{F_Y(y)} \text{ і } P_X(B)$$

описуються через однакові функції розподілу ймовірностей $F_X(x)$, а тому співпадають, оскільки через функцію розподілу ймовірностей однозначно визначається міра $P_X(B)$ на множині значень випадкового вектора X.

Отже,

$$P(\{E \mid X(E)\in B, Y(E)<y\})=P_X(B)F_Y(y), \text{ коли } B\in\mathscr{B}(R^s).$$

Ця рівність виконується й тоді, коли $F_Y(y)=0$.

Нехай $P_X(B)>0$. Тоді стосовно розподілів на R^l

$$\frac{P_{(X,Y)}(E \mid X(E)\in B, Y(E)\in A)}{P_X(B)}=\frac{P_{(X,Y)}(X^{-1}(B)\cap Y^{-1}(A))}{P_X(B)} \text{ і } P_Y(A)$$

буде одна і та сама функція розподілу $F_Y(y)$ ймовірностей на R^l, отже розподіли співпадають. Таким чином, коли $P_X(B)>0$, тоді

$$P_{(X,Y)}(\{E \mid X(E)\in B, Y(E)\in A\})=$$
$$=P_{(X,Y)}(X^{-1}(B)\cap Y^{-1}(A))=P_X(B)P_Y(A).$$

Очевидно, що ця рівність зберігається за всіх $B\in\mathscr{B}(R^s)$, $A\in\mathscr{B}(R^l)$.

Міркуючи аналогічно, можна показати, що для того, щоб випадкові величини $X_1, X_2, \ldots, X_n$ були незалежними, необхідно й достатньо, щоб за всіх $(x_1, x_2, \ldots, x_n)\in R^n$ виконувалась рівність

$$F_X(x_1, x_2, \ldots, x_n)=F_{X_1}(x_1)F_{X_2}(x_2)\ldots F_{X_n}(x_n).$$

Нехай тепер вектори X і Y незалежні з абсолютно неперервними розподілами ймовірностей, тобто існують невід'ємні функції $f_X(x)$, $x\in R^s$, $f_Y(y)$, $y\in R^l$, такі, що

$$P_X(A)=\int_A f_X(x)\,dx,\ A\in\mathscr{B}(R^s),\ P_Y(B)=\int_B f_Y(y)\,dy,\ B\in\mathscr{B}(R^l).$$

Тоді

$$P_{(X,Y)}(\{E \mid X(E)\in A, Y(E)\in B\})=P_{(X,Y)}(X^{-1}(A)\cap Y^{-1}(B))=$$
$$=P_X(\{E \mid X(E)\in A\})P_Y(\{E\mid Y(E)\in B\})=P_X(X^{-1}(A))P_Y(Y^{-1}(B))=$$
$$=\int_A f_X(x)\,dx\int_B f_Y(y)\,dy=\iint_{A\times B} f_X(x)f_Y(y)\,dxdy.$$

Оскільки розподіли

$$P_1(C)=P_{(X,Y)}((X,Y)\in C) \text{ і } P_2(C)=\iint_{\{(x,y)\in C\}} f_X(x)f_Y(y)dxdy$$

співпадають на множинах виду

$$C=A\times B,\quad A\in\mathscr{B}(R^s),\quad B\in\mathscr{B}(R^l),$$

то вони співпадають на всіх борелівських множинах в R^{s+l}.

Отже, якщо щільності розподілу ймовірностей на множинах значень незалежних випадкових векторів X і Y є $f_X(x)$ і $f_Y(y)$ відповідно, то щільністю сумісного розподілу ймовірностей на декартовому добутку множин значень цих векторів є $f_{(X,Y)}(x,y)=f_X(x)f_Y(y)$, $x\in R^s$, $y\in R^l$.

Нехай $X_1, X_2, \ldots, X_n$ – випадкові величини на ймовірнісному просторі (Ω, S, P), B – борелівська множина в R^n. Тоді

$$\{E \mid (X_1(E), X_2(E), \ldots, X_n(E))\in B\}\in S.$$

Справді, розглянемо спочатку випадок, коли $B=\prod_{i=1}^{n}[\,a_i\,;b_i)$.

Одержуємо

$$\{E\mid (X_1(E),X_2(E),...,X_n(E))\in B\}=$$
$$=\{E\mid a_1\le X_1(E)<b,...,a_n\le X_n(E)<b_n=$$
$$=\bigcap_{i=1}^{n}\{E\mid a_i\le X_i(E)<b_i\}\in S\ ,$$

і таким чином, множини $\{E\mid a_i\le X_i(E)<b_i\}=\bigcap_{i=1}^{n}X_i^{-1}([a_i\,;b_i))$ є випадковими подіями.

Нехай Q – система тих підмножин $B\subset R^n$, за яких

$$\{E\mid (X_1(E),X_2(E),...,X_n(E))\in B\}\in S\ .$$

Розглянемо відображення $X:\Omega\to R^n$, $X=(X_1,\ X_2,\ \dots,X_n)$. Із співвідношення між образами і прообразами випливає, що Q є σ-алгеброю. Проте в Q містяться всі паралелепіпеди, а отже і найменша σ-алгебра, породжена за системою паралелепіпедів, тобто в Q міститься всі борелівські множини, що й треба було довести.

Вправи для самостійного виконання

4.7.1. Перевірити, чи правильні твердження:

1. Через розподіл ймовірностей на множині значень випадкового вектора цілком визначаються розподіли ймовірностей на множинах значень випадкових величин, що є координатами даного випадкового вектора.
2. Якщо випадкові величини $X_1,X_2,...,X_n$ незалежні відносно ймовірнісної міри P, то вони й попарно незалежні, відносно тієї самої ймовірнісної міри.
3. Твердження, обернене до 2, правильне.
4. Координати будь-якого випадкового вектора є попарно незалежними випадковими величинами.
5. Випадкові величини X і Y незалежні тоді й тільки тоді, коли $F_{(X,Y)}(x,y)=F_X(x)\cdot F_Y(y)$.
6. Твердження 5 правильне й стосовно випадкових векторів X та Y.
7. Якщо розподіли ймовірностей на множинах значень випадкових векторів X та Y абсолютно неперервні із щільностями $f_X(x)$ та $f_Y(y)$, то щільність сумісного розподілу ймовірностей на декартовому добутку множин значень цих векторів визначається за рівністю

$$f_{(X,Y)}(x,y)=f_X(x)\cdot f_Y(y).$$

4.8. Функції кількох випадкових аргументів

Нехай задано ймовірнісний простір (Ω, S, P) і відображення X: $\Omega \to \Omega_X$, за яким кожному елементові $E \in \Omega$ ставиться у відповідність елемент із деякої множини Ω_X. Нехай Ω_X – образ множини Ω за відображенням X, а S_X – сукупність підмножин множини Ω_X, стосовно якої задовільняються вимоги 1_s-3_s:

1_s. $\Omega_X \in S_X$.

2_s. Якщо $G \in S_X$, то і $\overline{G} = \Omega_X \setminus G \in S_X$.

3_s. Якщо $G_i \in S_X$, $i = 1, 2, \ldots$, то і $\bigcup_i G_i \in S_X$.

Якщо стосовно кожного $G \in S_X$ виявляється $X^{-1}(G) \in S$, тоді *функція* X *називається* S/S_X*-вимірною або* S/S_X*-випадковою величиною* відносно простору (Ω, S, P) та S_X.

Природно покласти $P_X(G) = P(X^{-1}(G))$, $G \in S_X$, оскільки значення $x = X(E)$ належить до множини $G \subset \Omega_X$ тоді і тільки тоді, коли $E \in X^{-1}(G)$ (див. §4.1).

В такому разі говорять, що міра $P_X(G)$, $G \in S_X$, породжується (генерується) через S/S_X-випадкову величину X за мірою $P(A)$, $A \in S$, заданою на S.

В такий спосіб утворюється ймовірнісний простір (Ω_X, S_X, P_X), який (за заданого S_X) породжується (генерується) через S/S_X-випадкову величину X за простором (Ω, S, P).

Якщо за відображенням (функцією) X кожному елементові $E \in \Omega$ ставиться у відповідність пара дійсних чисел (тобто точка із двохвимірного дійсного координатного простору R^2) і стосовно такої функції виконуються всі наведені умови, тоді функція X називається R^2-значною S/S_X-випадковою величиною (або 2-вимірним випадковим вектором). Аналогічно вводяться поняття R^3-значної, …, R^n-значної S/S_X-випадкових величин (або 3-вимірних, …, n-вимірних випадкових векторів).

Нехай X– R^1-значна випадкова величина відносно простору (Ω, S, P) і S_X, і нехай на Ω_X задано функцію $Z = \psi(X)$, за якою кожному x із Ω_X ставиться у відповідність дійсне число. Нехай Ω_Z образ множини Ω_X за відображення Z, S_Z –

сукупність підмножин множини Ω_Z, стосовно якої задовільняються вимоги 1_s-3_s.

Якщо за кожного $Q \in S_Z$ $Z^{-1}(Q) \in S_X$, тоді функція Z буде S_X / S_Z-вимірною, тобто S_X / S_Z-випадковою величиною відносно ймовірнісного простору (Ω_X, S_X, P_X) і S_Z.

За таких умов природно покласти

$$P_Z(Q) = P_X(Z^{-1}(Q)) = P(X^{-1}(Z^{-1}(Q))).$$

В такий спосіб утворюється ймовірнісний простір (Ω_Z, S_Z, P_Z), який (із врахуванням заданого S_Z) породжується (генерується) через випадкову величину Z за ймовірнісним простором (Ω_X, S_X, P_X).

Якщо X R^2-значна, R^3-значна і т.д. S/S_X-випадкова величина, тобто функція $Z = \psi(X)$ залежить від двох, трьох і т.д. випадкових аргументів ($Z = \psi(X_1, X_2)$, $Z = \psi(X_1, X_2, X_3)$ і т.п.), міркування залишаються цілком аналогічними до попереднього.

Як Ω_Z часто розглядають множину $(-\infty; \infty)$ всіх дійсних чисел, вважаючи прообраз деякої точки $z \in R^1$ порожнім, якщо така точка не поставлена у відповідність ніяким $x \in \Omega_X$.

Як сукупність S_Z підмножин множини $\Omega_Z = (-\infty; \infty)$ найчастіше розглядають σ-алгебру $\mathscr{B}(R^1)$ борелівських множин числової прямої, породжену за числовими проміжками та їх скінченними сумами. Зокрема до сукупності S_Z включаються довільні числові проміжки та об'єднання не більш ніж зчисленної кількості таких проміжків, доповнення таких об'єднань до R^1, скінченні та зчисленні множини тощо (див. §3.1).

В останньому випадку розподіл ймовірностей на множині Ω_Z значень випадкової величини Z може бути описаний за допомогою функції розподілу ймовірностей

$$F_Z(z) = P_Z((-\infty;\ z)) = P_X(Z^{-1}((-\infty;\ z))).$$

Неперервний розподіл ймовірностей може бути описаний також за допомогою щільності розподілу ймовірностей $f_Z(z)$ (див. §3.2-§3.5). В разі необхідності можуть бути обчислені числові характеристики розподілу ймовірностей на множині Ω_Z значень випадкової величини Z цілком аналогічно до попереднього як у випадку дискретного, так і у випадку неперервного розподілу ймовірностей.

Нехай $f(x): R^n \to R^1$ – борелівська функція. Клас таких функцій досить широкий, до нього, зокрема, входять усі неперервні функції. Справді, якщо $f(x)$ неперервна, то

$$f^{-1}((-\infty;\, c)) = \{x \mid f(x) \in (-\infty; c)\} = \{x \mid f(x) < c\}$$

є відкритою множиною, а тому належить до σ-алгебри борелівських множин.

Теорема. Нехай $\psi(x)$ – борелівська функція, $\psi(x): R^n \to R^1$, а $X_1(E), X_2(E), \ldots, X_n(E)$ – випадкові величини на ймовірнісному просторі (Ω, S, P). Тоді $\psi(X_1(E), X_2(E), \ldots, X_n(E))$ – випадкова величина.

Справді, оскільки функція $\psi(x)$ – борелівська, то за довільного $c \in R^1$

$$\psi^{-1}((-\infty\,;\, c)) \in \mathscr{B}(R^n)\,.$$

Тоді

$$\{E \mid \psi(X_1(E), X_2(E), \ldots, X_n(E)) \in (-\infty; c)\} =$$

$$= \{E \mid (X_1(E), X_2(E), ..., X_n(E)) \in \psi^{-1}((-\infty; c))\} \in S\,.$$

Нехай в деякій області $A \subset R^s$ задано неперервні диференційовні функції

$$\begin{gathered} y_1 = g_1(x_1,\, x_2,\, \ldots, x_s), \\ y_2 = g_2(x_1,\, x_2,\, \ldots, x_s),\, ...,\; y_s = g_s(x_1,\, x_2,\, \ldots, x_s), \end{gathered} \tag{4.8.1}$$

причому існує єдиний розв’язок цієї системи рівнянь:

$$x_1 = g_1^{-1}(y_1, y_2, \ldots, y_s)\,,\; x_2 = g_2^{-1}(y_1,\, y_2,\, \ldots, y_s)\,,\; \ldots,$$
$$x_s = g_s^{-1}(y_1,\, y_2,\, \ldots, y_s)\,.$$

Нехай C – образ області A за перетворенням (4.8.1) в просторі значень $y_1, y_2, \ldots, y_s$, якобіан $J = \left|\dfrac{\partial x_j}{\partial y_i}\right|$ відмінний від нуля. Тоді правильне твердження: якщо $f_X(x_1, x_2, \ldots, x_s)$ щільність розподілу ймовірностей на множині значень випадкового вектора X, то

$$f_y(y_1, y_2, ... y_s) =$$

$$= f_X(g_1^{-1}(y_1, y_2, ... y_s),\, g_2^{-1}(y_1, y_2, ... y_s),\, ...,\, g_s^{-1}(y_1, y_2, ... y_s))J$$

є щільністю розподілу ймовірностей на множині значень випадкового вектора $Y = (g_1(X),\, g_2(X),\, \ldots, g_s(X))$.

Справді, якщо $B \in \mathscr{B}(R^s)$ – борелівська множина, $B \subset A$, B' – образ B за перетворенням (4.8.1), то

$$P_X(B) = \int\limits_B f_X(x_1, x_2, \dots, x_s) dx_1 dx_2 \dots dx_s =$$

$$= \int\limits_{B'} f_X(g_1^{-1}(y_1, y_2, \dots, y_s), \dots, g_s^{-1}(y_1, y_2, \dots, y_s)) J\, dy_1 dy_2 \dots dy_s =$$

$$= \int\limits_{B'} f_Y(y_1, y_2, \dots, y_s) dy_1 dy_2 \dots dy_s = P_Y(\{E \mid Y(E) \in B'\}) = P_Y(Y^{-1}(B')),$$

що й треба було довести.

Нехай $h(x)$, $g(y)$, $x \in R^s$, $y \in R^l$, – борелівські функції із значеннями в R^{s_1}, R^{l_1}, а $X(E)$ та $Y(E)$ – випадкові вектори відповідно R^s- та R^l-значні. Тоді $X' = h(X)$ та $Y' = g(Y)$ – випадкові вектори.

Теорема. Якщо X і Y незалежні, то X' і Y' також незалежні.

Справді, нехай A_1, B_1 – довільні борелівські множини в R^{s_1}, R^{l_1}. Тоді

$$\{E \mid X'(E) \in A_1, Y'(E) \in B_1\} = \{E \mid h(X(E)) \in A_1, g(Y(E)) \in B_1\} =$$

$$= \{E \mid X(E) \in h^{-1}(A_1), Y(E) \in g^{-1}(B_1)\},$$

де $h^{-1}(A_1)$, $g^{-1}(B_1)$ – борелівські множини. Отже,

$$P_{(X',Y')}\{E \mid X'(E) \in A_1, Y'(E) \in B_1\} =$$

$$= P_{(X,Y)}(\{E \mid X(E) \in h^{-1}(A_1), Y(E) \in g^{-1}(B_1)\}) =$$

$$= P_X(\{E \mid X(E) \in h^{-1}(A_1)\}) P_Y(\{E \mid Y(E) \in g^{-1}(B_1)\}) =$$

$$= P_{X'}(\{E \mid X'(E) \in A_1\}) P_{Y'}(\{E \mid Y'(E) \in g^{-1} B_1\}),$$

що й треба було довести.

Приклад 4.8.1. Знайти розподіл ймовірностей на множині значень випадкової величини $Z = \psi(X, Y) = X^2 + Y^2$, якщо розподіл ймовірностей на множині значень двохвимірної випадкової величини (X, Y) рівномірний у крузі радіуса R з центром у початку координат і описується через щільність розподілу ймовірностей (Рис. 4.8.1)

$$f_{(X,Y)}(x, y) = \begin{cases} \dfrac{1}{\pi R^2}, & \text{коли } x^2 + y^2 \le R^2; \\ 0, & \text{коли } x^2 + y^2 > R^2. \end{cases}$$

Тут покладається $\Omega_{(X,Y)} = R^2$, $S_{(X,Y)} = \mathscr{B}(R^2)$.

Враховуючи, що поверхня $z = x^2 + y^2$ є параболоїдом обертання (Рис. 4.8.2), дістаємо $P_Z((-\infty; z)) = 0$, коли $z \le 0$, оскільки прообрази множин $(-\infty; z)$ за відображенням $z = x^2 + y^2$ порожні, коли $z \le 0$. Як Ω_Z розглядається R^1, як $S_Z - \mathscr{B}(R^1)$. Оскільки функція $z = x^2 + y^2$ неперервна, а отже борелівська, то за довільного $G \in \mathscr{B}(R^1)$ буде $Z^{-1}(G) \in \mathscr{B}(R^2)$.

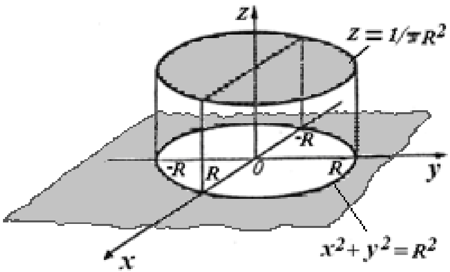


Рис. 4.8.1

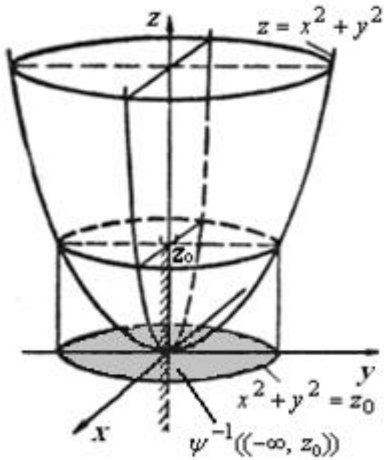


Рис. 4.8.2

Якщо $0 < z$, але $z \le R^2$, то прообразом множини $(-\infty;\, z) \in \mathscr{B}(R^1)$ за відображенням $z = x^2 + y^2$ є круг радіуса $\sqrt{z}$ з центром в початку координат, тобто $\psi^{-1}((-\infty;\, z)) = \{(x, y) \mid\; x^2 + y^2 < z\} \in \mathscr{B}(R^2)$ (див. Рис. 4.8.2). Ймовірність того, що пари координат двохвимірної випадкової величини (X, Y) належатимуть такому прообразу, дорівнює

$$\iint\limits_{Z^{-1}((-\infty;\; z))} f_{(X,Y)}(x,\, y)\; dxdy = \frac{1}{\pi R^2}\pi z$$

як об'єм під поверхнею щільності ймовірності над прообразом $Z^{-1}((-\infty;\; z))$, тобто над кругом радіуса $\sqrt{z}$ з центром в початку координат (Рис. 4.8.3).

Коли $z > R^2$, тоді прообраз $\psi^{-1}((-\infty; z))$ множини $(-\infty; z)$ також буде круг радіуса $\sqrt{z}$, але оскільки поза кругом радіуса R функція $f_{(X,Y)}(x,\, y) = 0$ (Рис. 4.8.4), то за умови $z > R^2$

імовірність того, що пари координат випадкової величини (X,Y) належатимуть прообразу $\psi^{-1}((-\infty;\,z))$, дорівнює

$$\iint\limits_{\psi^{-1}((-\infty,\,z))} f_{(X,Y)}(x,\,y)\,dxdy = 1\,.$$

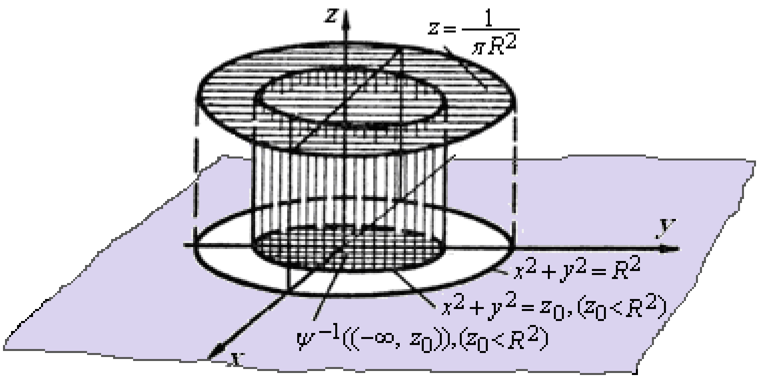


Рис. 4.8.3

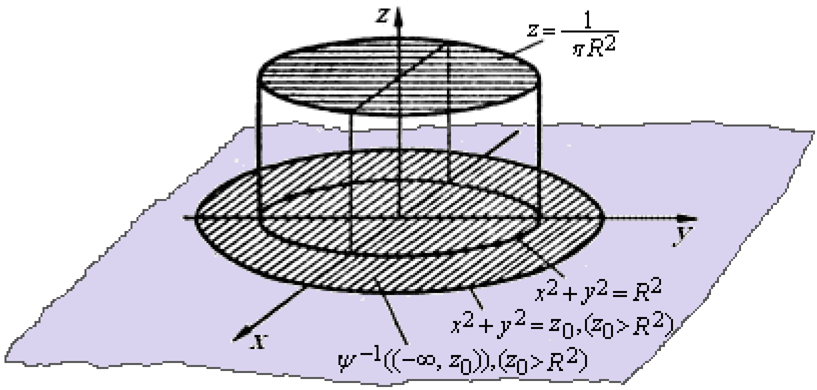


Рис. 4.8.4

Таким чином,

$$F_Z(z)=\begin{cases}0, & \text{коли } z\le 0;\\ \dfrac{z}{R^2}, & \text{коли } 0\le z\le R^2;\\ 1, & \text{коли } R^2\le z.\end{cases}$$

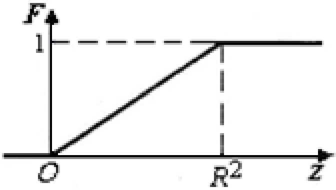


Рис. 4.8.5

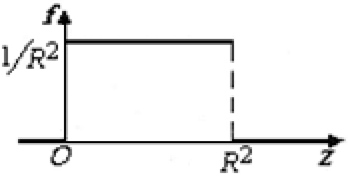


Рис. 4.8.6

Графік функції $F_Z(z)$ зображено на Рис. 4.8.5. Графік щільності $f_Z(z)$ розподілу ймовірностей на множині значень випадкової величини Z:

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} \dfrac{1}{R^2}, \text{ коли } z \in [0; R^2]; \\ 0, \text{ коли } z \overline{\in} [0\ ; R^2]. \end{cases}$$

зображено на Рис. 4.8.6.

Розглянемо окремо випадок, коли треба знайти розподіл імовірностей на множині значень суми $X+Y$ одновимірних випадкових величин X і Y за умови, що розподіл імовірностей на множині значень пари випадкових величин (X, Y) (випадкового вектора (X, Y)) описується через щільність розподілу ймовірностей $f_{(X,Y)}(x, y)$.

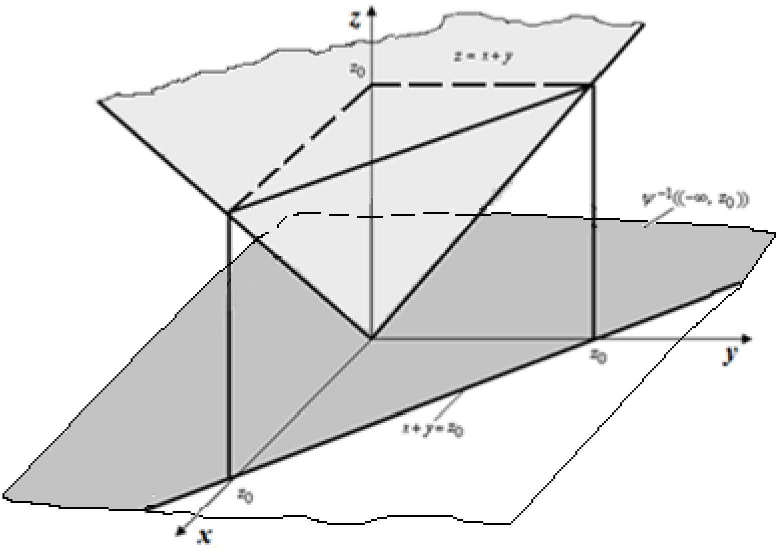


Рис. 4.8.7

Прообразом $\psi^{-1}((-\infty; z))$ множини $(-\infty; z)$ за відображення $z=\psi(x,y)=x+y$ є півплощина площини xOy, в точках якої виконується нерівність $x+y<z$ (Рис. 4.8.7).

Таким чином,

$$F_Z(z)=P_{(X;Y)}(\psi^{-1}((-\infty; z)))=\iint\limits_{\psi^{-1}((-\infty; z))} f_{(X,Y)}(x, y)dxdy=$$

$$=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-x} f_{(X,Y)}(x,\ y)dy\right\}dx.$$

Стосовно щільності розподілу ймовірностей дістаємо

$$f_Z(z)=\frac{d}{dz}F_{(Z)}(z)=\frac{d}{dz}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-x} f_{(X,Y)}(x,\ y)dy\right\}dx=\int_{-\infty}^{\infty} f_{(X,Y)}(x,\ z-x)dx.$$

Якщо змінити порядок інтегрування, то

$$F_Z(z)=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-y} f_{(X,Y)}(x,\ y)dx\right\}dy,$$

$$f_Z(z)=\frac{d}{dz}F_{(Z)}(z)=\frac{d}{dz}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{z-y} f_{(X,Y)}(x,\ y)dx\right\}dy=\int_{-\infty}^{\infty} f_{(X,Y)}(z-y, y)dy.$$

Якщо випадкові величини X і Y незалежні, то

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx, \tag{4.8.2}$$

або

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy \tag{4.8.3}$$

Функція $f_Z(z)$ в такому разі називається *згорткою функцій* $f_X(x)$, $f_Y(y)$.

Приклад 4.8.2. Опис щільності двохвимірного розподілу ймовірностей подано у вигляді

$$f_{(X,Y)}(x,y)=\begin{cases} c, & \text{коли } (x,\ y)\in D; \\ 0, & \text{коли } (x,\ y)\overline{\in} D, \end{cases}$$

де $D=\{(x,\ y)\mid\ x\in[0;1],\ y\in[0;1]\}$ (див. Рис. 4.8.8).

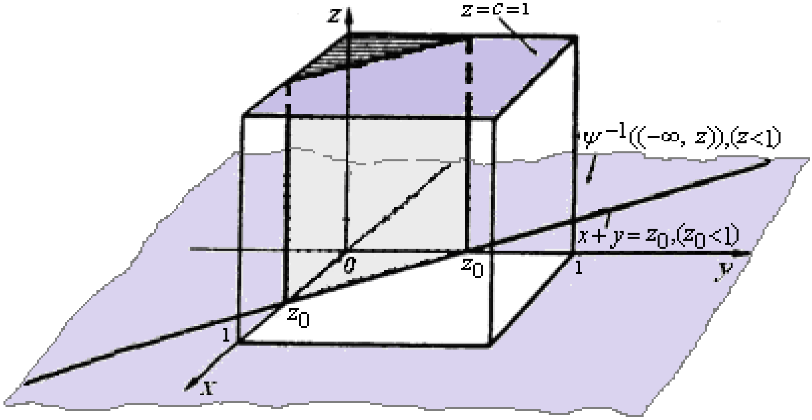


Рис. 4.8.8

Знайти функцію розподілу ймовірностей на множині значень випадкової величини $Z = X + Y$.

Враховуючи попередні міркування, одержуємо:

$$F_Z(z) = P_{(X,Y)}(\psi^{-1}((-\infty; z))) = \iint\limits_{\psi^{-1}((-\infty;\, z))} f_{(X,Y)}(x,y)dxdy = c \iint\limits_{\psi^{-1}((-\infty;\, z))\cap D} dxdy$$

.

Оскільки $c = 1$, то остаточно дістаємо (див. Рис. 4.8.8)

$$F_Z(z) = \begin{cases} 0, & \text{коли} \quad z \le 0; \\ \dfrac{z^2}{2}, & \text{коли} \quad 0 \le z \le 1; \\ 1 - \dfrac{(2-z)^2}{2}, & \text{коли} \quad 1 \le z \le 2; \\ 1, & \text{коли} \quad 2 \le z. \end{cases}$$

Графік функції $F_Z(z)$ подано на Рис. 4.8.9.

Щільність $f_Z(z)$ можна знайти як похідну $\dfrac{d}{dz}F_Z(z)$ або як згортку функцій $f_X(x)$, $f_Y(y)$, використовуючи одну з формул (4.8.2) або (4.8.3), оскільки випадкові величини X і Y в даному випадку незалежні.

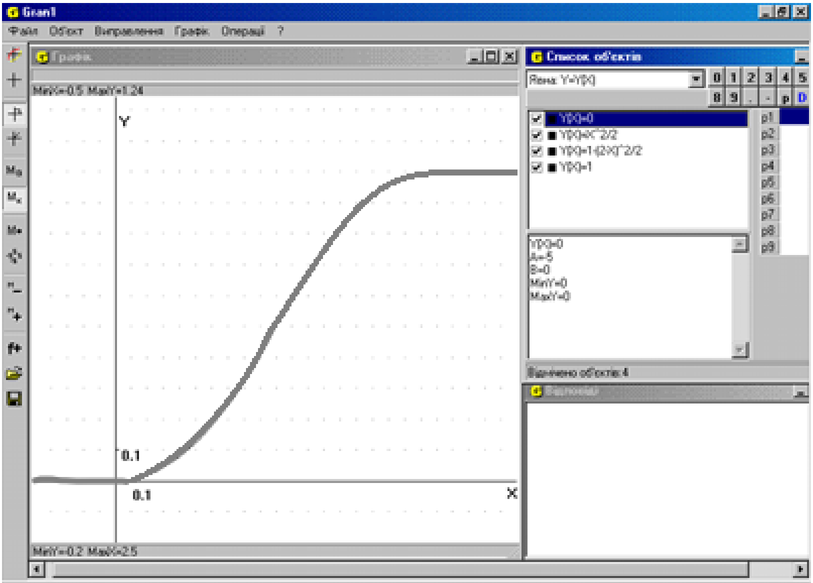


Рис. 4.8.9

Беручи до уваги те, що

$$f_X(x) = \begin{cases} 1, \text{ коли } x \in [0;\ 1]; \\ 0, \text{ коли } x \overline{\in} [0;\ 1], \end{cases} \qquad f_Y(y) = \begin{cases} 1, \text{ коли } y \in [0;\ 1]; \\ 0, \text{ коли } y \overline{\in} [0;\ 1], \end{cases}$$

знаходимо $f_X(x) f_Y(z-x) = 1 \cdot 1$, якщо $0 \le x \le 1$ і $0 \le z - x \le 1$.

Множину точок $(x,\ z)$, в яких виконуються нерівності $0 \le x \le 1$, $0 \le z - x \le 1$, зображено на Рис. 4.8.10.

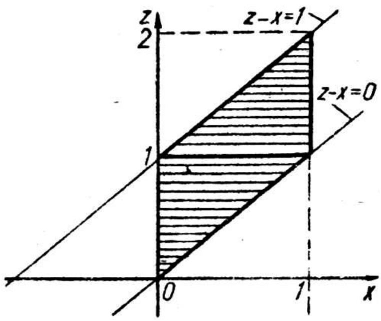


Рис. 4.8.10

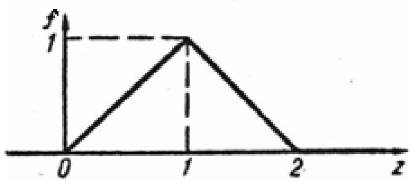


Рис. 4.8.11

Тепер знаходимо

$$f_Z(z)=\int\limits_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx=\begin{cases}\int\limits_0^z 1\cdot 1\cdot dx=z, \text{ коли } z\in[0;\ 1];\\ \int\limits_{z-1}^{1} 1\cdot 1\cdot dx=2-z, \text{ коли } z\in[1;\ 2];\\ 0, \qquad \text{коли } z\,\overline{\in}\,[0\ ;\ 2].\end{cases}$$

Графік щільності $f_Z(z)$ подано на Рис. 4.8.11.

Вправи для самостійного виконання

4.8.1. Перевірити, чи правильні твердження:

1. Кожна функція $y=f(x)$, $x\in R^n$, є борелівською.

2. Якщо функція $y=f(x)$, $x\in R^n$, неперервна, то вона борелівська.

3. Якщо $f:R^n\to R^1$ – борелівська функція, то вона неперервна.

4. Якщо $y=f(x),\ x\in R^n,\ y\in R^1$, а $X(E),\ E\in\Omega$, – випадковий R^n-значний вектор, то $Y(E)=f(X(E))$ – випадкова величина.

5. Твердження 4 є правильним, коли f – борелівська функція.

6. Якщо $X_i(E),\ i\in\overline{1,n}$, – випадкові величини, а $\psi:R^n\to R^1$ – довільна функція, то $\psi(X_1(E),X_2(E),...,X_n(E))$ – випадкова величина.

7. Якщо $h(X(E))$ і $g(Y(E))$ – незалежні випадкові величини, то такими є й випадкові величини $X(E)$ і $Y(E)$.

8. Якщо X_1 та X_2 незалежні R^s-значні випадкові вектори, а $y=f(x),\ x\in R^s,\ y\in R^l$, то $Y_1=f(X_1)$ і $Y_2=f(X_2)$ – незалежні R^l-значні випадкові вектори.

9. Якщо $f_X(x_1,x_2)$ – щільність розподілу ймовірностей на множині значень випадкового вектора $X=(X_1,X_2)$, $y_1=g_1(x_1,x_2)$, $y_2=g_2(x_1,x_2)$ – довільні функції, то $f_Y(y_1,y_2)=f_X(g_1^{-1}(y_1,y_2),\ g_2^{-1}(y_1,y_2))$ щільність розподілу ймовірностей на множині значень випадкового вектора $Y=(g_1(X),\ g_2(X))$.

10. Розподіл ймовірностей на множині значень суми $X+Y$ одновимірних випадкових величин X і Y з абсолютно неперервними розподілами ймовірностей завжди можна знайти за допомогою згортки щільностей $f_X(x)$ та $f_Y(y)$.

11. Твердження 10 є правильним, коли X та Y незалежні випадкові величини.

12. За будь-яких щільностей $f_X(x)$ та $f_Y(y)$ розподілів ймовірностей на множинах значень випадкових величин X та Y існує їх згортка.

13. Через згортку щільностей $f_X(x)$ та $f_Y(y)$ розподілів ймовірностей на множинах значень випадкових величин X та Y визначається спільний розподіл $f_{(X,Y)}(x,y)$ на декартовому добутку множин значень випадкових величин X та Y.

14. Існують функції $f_X(x)$ та $f_Y(y)$, за яких згорткою є функція

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(x)\cdot f_Y(z-x)\,dx .$$

15. Твердження 14 є правильним за будь-яких функцій $f_X(x)$ та $f_Y(y)$.

4.9. Математичне сподівання випадкової величини

Щоб більш повно охарактеризувати розподіл ймовірностей і мати змогу порівнювати за деякими ознаками розподіли ймовірностей, доцільно ввести певні *числові характеристики* розподілу ймовірностей на множині значень випадкової величини X. Іноді розподіл імовірностей буває невідомим, а деякі числові характеристики розподілу ймовірностей можна знайти (хоча б наближено). Тоді за такими характеристиками можна до певної міри охарактеризувати і весь розподіл імовірностей.

Однією з найбільш важливих числових характеристик розподілу ймовірностей на множині значень випадкової величини X є *математичне сподівання* або *середнє значення випадкової величини* X, яке позначається через $M[X]$.

Нехай (Ω, S, P) – ймовірнісний простір, X – проста випадкова величина з множиною можливих значень

$$\Omega_X = X(\Omega) = \{x_1,\ x_2,\ \ldots,\ x_k\},$$

тобто $X(E) = \sum_{i=1}^{k} x_i I_{X^{-1}(x_i)}(E)$, $I_{X^{-1}(x_i)}(E)$ – індикаторна функція множини $X^{-1}(x_i)$.

Математичним сподіванням простої випадкової величини X називається число

$$M[X] = \sum_{i=1}^{k} x_i P_X(\{x_i\}) = \sum_{i=1}^{k} x_i P(X^{-1}(x_i)) = \sum_{i=1}^{k} x_i p_i\ .$$

Якщо множина $X(\Omega)$ нескінченна (але зчисленна), то за означенням

$$M[X] = \sum_{i=1}^{\infty} x_i P_X(\{x_i\}) = \sum_{i=1}^{\infty} x_i P(X^{-1}(x_i)) = \sum_{i=1}^{\infty} x_i p_i\ ,$$

коли ряд $\sum_{i=1}^{\infty} x_i p_i$ збігається.

Якщо останній ряд не збігається, то говорять, що математичне сподівання випадкової величини X не існує.

Приклад 4.9.1. Нехай A – випадкова подія,

$$I_A(E) = \begin{cases} 1, & \text{коли} \quad E \in A; \\ 0, & \text{коли} \quad E \overline{\in} A, \end{cases}$$

індикаторна функція, яка є простою випадковою величиною.

Тоді $M[I_A(E)] = P(A)$.

Приклад 4.9.2. Знайти математичне сподівання числа очок, що випадають на верхній грані правильного однорідного ігрового кубика в результаті однократного підкидання, якщо ймовірності випадання на верхній грані будь якого числа очок із множини $\{1, 2, 3, 4, 5, 6\}$ однакові.

Позначимо це число очок, яке є простою випадковою величиною, через X. Тоді розподіл ймовірностей на множині значень даної випадкової величини набуває вигляду, поданого в Табл. 4.9.1.

За означенням математичного сподівання одержуємо

$$M[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{1}{6}(1+2+3+4+5+6) = 3.5.$$

Табл. 4.9.1

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Зміст математичного сподівання випадкової величини можна простежити на такому прикладі.

Нехай виконують n пострілів в мішень. Число здобутих очок в результаті кожного пострілу є окремим спостереженим значенням випадкової величини X. Для оцінювання якості стріляння в цілому слід розглядати не результат кожного окремого пострілу, а середній результат досить великої серії пострілів. Середня кількість очок на один постріл – це саме те значення (центр розсіювання ймовірностей), навколо якого групуються окремі значення випадкової величини X, які дістали в окремих випробуваннях, і за яким характеризують якість стріляння. Таким чином, за допомогою математичного сподівання випадкової величини можна дати орієнтовну оцінку середнього результату серії випробувань, тоді як в ряді розподілу ймовірностей вказується, яких значень може набувати випадкова величина та з якими ймовірностями.

Якщо в серії з n випробувань випадкова величина X набуває своїх можливих значень $x_1, x_2, \ldots, x_k$ відповідно $n_1, n_2, \ldots, n_k$ разів, причому $n_1 + n_2 + \ldots + n_k = n$, то середнім значенням є

$$\frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n} = x_1 \frac{n_1}{n} + x_2 \frac{n_2}{n} + \cdots + x_k \frac{n_k}{n} =$$

$$= x_1 P_n^*(\{x_1\}) + x_2 P_n^*(\{x_2\}) + \ldots + x_k P_n^*(\{x_k\}),$$

де $P_n^*(\{x_i\})$ – статистичні ймовірності подій $\{x_i\}$ в даній серії випробувань. Оскільки вірогідно, що статистичні ймовірності $P_n^*(\{x_i\})$ за великої кількості випробувань близькі до ймовірностей $p_i = P_X(\{x_i\})$ (одержаних в результаті узагальнення статистичних даних), то вірогідно, що середнє значення випадкової величини X, здобуте за результатами великої серії випробувань, близьке до математичного сподівання $M[X]$.

Як відомо, за дискретного розподілу деякої маси вздовж осі абсцис такого, що в точці з абсцисою x_i знаходиться маса p_i, координату центра мас вказаної системи мас обчислюють як частку від ділення суми $\sum_{i=1}^{k} x_i p_i$ статичних моментів розглядуваної системи мас відносно початку координат на суму мас $\sum_{i=1}^{k} p_i$. (Під статичним моментом маси p_i відносно точки O розуміють добуток маси p_i на відстань від точки, в якій поміщено масу p_i, до точки O).

Отже,

$$x_c = \frac{\sum_{i=1}^{k} x_i p_i}{\sum_{i=1}^{k} p_i}.$$

Якщо ймовірності p_i інтерпретувати як маси, то загальна сума мас (імовірностей) дорівнюватиме 1, і таким чином в механічній інтерпретації математичне сподівання випадкової величини X є не що інше, як абсциса центра мас загальної одиничної маси, розподіленої дискретно вздовж осі абсцис так, що на точку з абсцисою x_i припадає маса (ймовірність) p_i.

Приклад 4.9.3. Нехай є два дискретних розподіли ймовірностей (Табл. 4.9.2, 4.9.3).

Табл. 4.9.2

x_i	x_1	x_2	x_3	x_4	x_5
p_i	0.01	0.01	0.96	0.01	0.01

Табл. 4.9.3

x'_i	x'_1	x'_2	x'_3	x'_4	x'_5	x'_6
p'_i	0.01	0.01	0.48	0.48	0.01	0.01

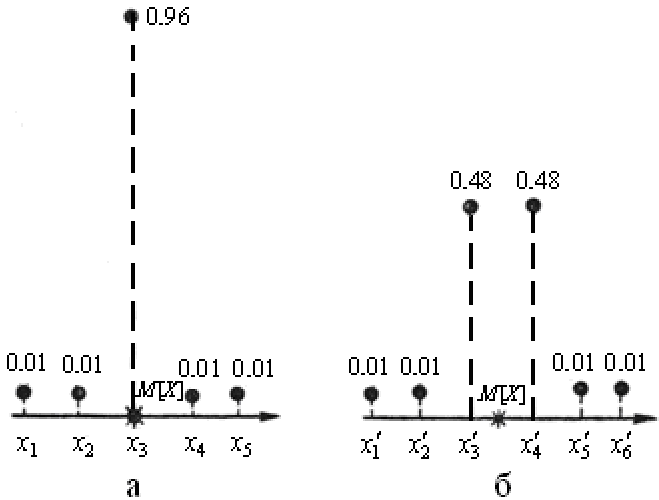


Рис. 4.9.1

Розглядаючи ці розподіли (Рис. 4.9.1, а, б), можна помітити, що в результаті проведення випробувань стосовно випадкової величини з розподілом імовірностей, поданим в Табл. 4.9.2, слід сподіватися, що в результаті випробування випадкова величина X набуде значення x_3, а стосовно випадкової величини з розподілом, поданим в Табл. 4.9.3 – одного із значень x'_3 або x'_4, на які припадають найбільші ймовірності і які найближчі до математичного сподівання. Цим і пояснюється назва *математичне сподівання випадкової величини* X. Зазначимо, що не обов'язково $M[X] \in X(\Omega)$, тобто математичне сподівання $M[X]$ може і не належати до множини $\Omega_X = X(\Omega)$ можливих значень випадкової величини X. Точку, абсциса якої дорівнює математичному сподіванню, називають *центром розсіювання ймовірностей* на множині можливих значень випадкової величини X.

Приклад 4.9.4. Знайти математичне сподівання випадкової величини, ймовірності на множині можливих значень якої розподілені за законом Пуассона (див. Розділ 2, п. 2.10).

За означенням математичного сподівання випадкової величини знаходимо

$$M[X]=\sum_{i=0}^{\infty}x_i p_i=\sum_{i=0}^{\infty}i\,\frac{a^i}{i!}e^{-a}=\sum_{i=1}^{\infty}i\,\frac{a^i}{i!}e^{-a}=$$

$$=e^{-a}\sum_{i=1}^{\infty}\frac{a^i}{(i-1)!}=e^{-a}a\sum_{i=1}^{\infty}\frac{a^{i-1}}{(i-1)!}=e^{-a}a\sum_{i=0}^{\infty}\frac{a^i}{i!}=e^{-a}ae^{a}=a\,.$$

Математичне сподівання простої випадкової величини X є не що інше, як інтеграл Лебега від S-вимірної функції $X(E)$ за мірою P, стосовно якого поряд з $M[X]$ використовуються також позначення

$$\int_{\Omega}X(E)P(dE)\,,\ \int_{\Omega}XdP\,.$$

Нехай $X=X(E)$ – невід'ємна випадкова величина. Розглянемо послідовність $\{X_n\}$, $n=1,\ 2,\ \ldots$, простих невід'ємних випадкових величин

$$X_n(E)=\sum_{k=1}^{n2^n}\frac{k-1}{2^n}I_{X^{-1}\left(\left[\frac{k-1}{2^n};\frac{k}{2^n}\right)\right)}(E)+nI_{X^{-1}([n;\infty))}(E)\,.$$

Ця послідовність зростаюча, тобто $X_{n+1}(E)\geq X_n(E)$, обмежена зверху: $X_n(E)<X(E)$, і якщо $n\to\infty$, то $X_n(E)\to X(E)$ за будь-якого $E\in\Omega$, ($\frac{n}{2^n}\to 0$, коли $n\to\infty$).

Можна показати, що $M[X_n]\leq M[X_{n+1}]$, а тому існує границя $\lim\limits_{n\to\infty}M[X_n]$, яка може набувати й значення $+\infty$.

Причому ця границя не залежить від вибору апроксимуючої послідовності X_n, тобто якщо монотонно зростаюча послідовність $\{X_n(E)\}$ простих випадкових величин збігається до $X(E)$ і деяка інша монотонно зростаюча послідовність $\{Y_m(E)\}$ простих випадкових величин також збігається до $X(E)$ за будь-якого $E\in\Omega$, то $\lim\limits_{n\to\infty}M[X_n]=\lim\limits_{m\to\infty}M[Y_m]$.

Інтегралом Лебега від невід'ємної випадкової величини X або її *математичним сподіванням* називається величина

$$M[X]=\lim_{n\to\infty}M[X_n]\,.$$

Нехай $X(E)$ – довільна випадкова величина, $X(E)=X^{+}(E)-X^{-}(E)$.

Говорять, що *математичне сподівання* випадкової величини X *існує* або *визначене*, якщо принаймні одна з величин $M[X^+]$ або $M[X^-]$ скінченна. Тут за означенням

$$M[X] = M[X^+] - M[X^-].$$

Математичне сподівання $M[X]$ називають також *інтегралом Лебега від функції X за ймовірнісною мірою P* і позначають $\int\limits_{\Omega} X(E)P(dE)$.

Функція $X(E)$ називається *інтегровною* або *сумовною*, якщо $M[X]$ скінченне, тобто якщо $M[X^+] < \infty$ і $M[X^-] < \infty$. Оскільки $X^+ - X^- = |X|$, то скінченність $M[X]$ означає, що $M[|X|] < \infty$, тобто що має бути скінченним інтеграл $\int\limits_{\Omega} |X(E)|P(dE)$.

Теорема (про монотонну збіжність). Якщо $\{X_n\}$ – монотонно зростаюча послідовність випадкових величин, обмежена зверху, і $X_n \underset{n\to\infty}{\to} X$, то $\{M[X_n]\}$ – монотонно зростаюча послідовність і $M[X_n] \underset{n\to\infty}{\to} M[X]$.

Якщо $(\Omega, S) = (R^1, \mathscr{B}(R^1))$, а μ – міра Лебега, то інтеграл $\int\limits_{R^1} X(t)\mu(dt)$ позначають $\int\limits_{R^1} X(t)dt$ або $\int\limits_{-\infty}^{\infty} X(t)dt$, або $(L)\int\limits_{-\infty}^{\infty} X(t)dt$ (щоб показати відмінність цього інтеграла від інтеграла Рімана $(R)\int\limits_{-\infty}^{\infty} X(t)dt$).

Зазначимо, що кожна функція, інтегровна за Ріманом, інтегровна й за Лебегом і обидва інтеграли співпадають. Кожна обмежена функція інтегровна за Лебегом на сегменті $[a;b]$. Щоб обмежена функція була інтегровною за Ріманом на відрізку $[a;b]$, необхідно й достатньо, щоб вона була неперервною майже всюди на $[a;b]$.

Якщо міра μ (Лебега-Стільтьєса) визначається за деякою функцією розподілу $G(t)$, то інтеграл $\int\limits_{R^1} X(t)\mu(dt)$ називають також *інтегралом Лебега-Стільтьєса* і позначають

$(L-S)\int\limits_{R^1} X(t)G(dt)$ на відміну від інтеграла Рімана-Стільтьєса $(R-S)\int\limits_{R^1} X(t)G(dt)$.

Якщо $M[X]$ визначене, то визначене також і математичне сподівання $M[XI_A]$ за довільного $A\in S$. Для $M[XI_A]$ або $\int\limits_{\Omega} XI_A\,dP$ використовують також позначення $\int\limits_A X\,dP$ або $M[X;A]$. Інтеграл $\int\limits_A X\,dP$ називається *інтегралом Лебега від функції* X *на множині* A *за мірою* P.

Якщо μ – n-вимірна міра Лебега-Стільтьєса, $A=[a_1;b_1)\times\ldots\times[a_n;b_n)$, то замість $\int\limits_A X\,d\mu$ пишуть

$$\int\limits_{a_1}^{b_1}\ldots\int\limits_{a_n}^{b_n} X(t_1,\ \ldots,t_n)\mu(dt_1\ldots dt_n).$$

Якщо μ – міра Лебега, то замість $\mu(dt_1\ldots dt_n)$ пишуть $dt_1\ldots dt_n$.

Нехай X – випадкова величина, стосовно якої визначено $M[X]$. Тоді визначена й функція множин

$$L(A)=\int\limits_A XdP=\int\limits_{\Omega} XI_A dP=M[XI_A],\ A\in S.$$

Це випливає з того, що

$$(XI_A)^+=X^+I_A\le X^+,\ (XI_A)^-=X^-I_A\le X^-.$$

Функція $L(A)$ є зчисленно-аддитивною, тобто якщо $A=\bigcup\limits_k A_k$, $A_k\in S$, $k\in N$, і $A_k\bigcap A_i=\varnothing$, $k\ne i$, то

$$L(A)=M(XI_A)=M[XI_{\bigcup\limits_k A_k}]=M\left[\sum\limits_k XI_{A_k}\right]=\sum\limits_k M[XI_{A_k}]=\sum\limits_k L(A_k).$$

Таким чином, якщо $M[X]$ визначено, то $L(A)=M[XI_A]$ є так званою *мірою із знаком*, тобто зчисленно-аддитивною функцією множин, яку можна подати у вигляді $L=L^+-L^-$, де принаймні одна з мір L^+ або L^- скінченна.

Крім того функція множин $L(A)$ абсолютно неперервна щодо міри P, тобто якщо $P(A)=0$, то $L(A)=0$, $A\in S$.

Справді, якщо X – невід'ємна проста випадкова величина, $X(E)=\sum_{i=1}^{k} x_i I_{X^{-1}(\{x_i\})}(E)$ і $P(A)=0$, то

$$L(A)=M[XI_A]=\sum_{i=1}^{k} x_i P(X^{-1}(\{x_i\})\cap A)=0 .$$

Якщо $\{X_i\}$ – монотонно зростаюча послідовність простих випадкових величин таких, що $X_i \underset{i\to\infty}{\to} X$, то за теоремою про монотонну збіжність

$$L(A)=M[XI_A]=\lim_{i\to\infty} M[X_i I_A]=0,$$

оскільки $M[X_i I_A]=0$ за будь-якого $i\geq 1$ і $A\in S$, за якого $P(A)=0$. Таким чином, інтеграл Лебега $L(A)=\int_A XdP=\int_\Omega XI_A dP$ є мірою зі знаком, абсолютно неперервною щодо міри P (позначається $L<<P$).

Теорема (Радона-Никодима). Нехай (Ω, S) – вимірний простір, P – σ-скінченна міра і L – міра зі знаком (тобто $L=L^+-L^-$, де принаймні одна з мір L^+ або L^- скінченна), яка є абсолютно неперервною щодо міри P.

Тоді існує S-вимірна функція $f(E)$ із значеннями в $\overline{R}^1=[-\infty;\infty]$ така, що

$$L(A)=\int_A f(E)P(dE),\ A\in S .$$

З точністю до множини P-міри нуль функція $f(E)$ єдина, тобто якщо існує інша S-вимірна функція $f_1(E)$ така, що

$$L(A)=\int_A f_1(E)P(dE),\ A\in S,$$

то

$$P(\{E\mid f(E)\neq f_1(E)\})=0 .$$

Якщо L – міра, то $f(E)$ набуває значень в $[0;\infty]$.

Таку функцію $f(E)$ називають *похідною Радона-Никодима* або *щільністю міри* L щодо міри P і позначають $\frac{dL}{dP}$ або $\frac{dL}{dP}(E)$.

Зазначимо, що коли деяка *властивість виконується з точністю до множини P-міри нуль*, тобто якщо існує така множина $V\in S$, що $P(V)=0$, і дана властивість виконується за кожної точки E з множини $\Omega\setminus V$, то говорять, що ця властивість виконується «*P-майже напевне*» («*майже напевне*» – м.н.). «*P-майже всюди*» («*майже всюди*» – м.в.).

Нехай $(R^1, \mathscr{B}(R^1), P_X)$ – ймовірнісний простір, що індукується за випадковою величиною X за простором (Ω, S, P), а $P_X(G) = P(X^{-1}(G))$, $G \in \mathscr{B}(R^1)$, – ймовірнісна міра на $(R^1, \mathscr{B}(R^1))$, $F_X(x)$ – відповідна функція розподілу ймовірностей.

Тоді, якщо випадкова величина X проста,

$$X(E) = \sum_{i=1}^{n} x_i I_{X^{-1}(\{x_i\})}(E) \text{, то}$$

$$M[X] = \sum_{i=1}^{n} x_i P(X^{-1}(\{x_i\})) = \sum_i x_i \Delta F_X(x_i) =$$

$$= \int_{R^1} x P_X(dx) = \int_{R^1} x F_X(dx) = \int_{-\infty}^{\infty} x F_X(dx),$$

де

$$\Delta F_X(x_i) = F_X(x_i + 0) - F_X(x_i).$$

В загальному випадку

$$M[X] = \int_{R^1} x P_X(dx) = \int_{R^1} x F_X(dx) = \int_{-\infty}^{\infty} x F_X(dx).$$

Якщо функція розподілу $F_X(x)$ абсолютно неперервна, тобто існує невід'ємна борелівська функція $f_X(x)$ (щільність розподілу ймовірностей на множині значень випадкової величини X) така, що

$$F_X(x) = \int_{-\infty}^{x} f_X(x) dx,$$

то

$$M[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx.$$

Таким чином, математичним сподіванням випадкової величини X з розподілом імовірностей, щільність якого $f_X(x)$, називається число

$$M[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx.$$

В механічній інтерпретації, як і раніше, математичне сподівання є абсцисою центра мас загальної одиничної маси, розподіленої вздовж осі абсцис із щільністю $f_X(x)$.

Точку, абсциса якої дорівнює математичному сподіванню $M[X]$, називають *центром розсіювання ймовірностей на множині значень випадкової величини* X.

Приклад 4.9.5. Знайти математичне сподівання неперервної випадкової величини X з рівномірним розподілом імовірностей на множині Ω_X її значень – відрізку $[a;b]$.

Одержуємо

$$M[X] = \int_{-\infty}^{\infty} x\, f_X(x)dx = \int_{-\infty}^{a} x \cdot 0 \cdot dx + \int_{a}^{b} x \frac{1}{b-a} dx + \int_{b}^{\infty} x \cdot 0 \cdot dx = \frac{1}{b-a} \cdot \frac{x^2}{2}\bigg|_a^b = \frac{a+b}{2}.$$

Таким чином, математичне сподівання випадкової величини X з рівномірним розподілом імовірностей на множині Ω_X значень випадкової величини X – відрізку $[a;b]$ знаходиться посередині цього відрізка (Рис. 4.9.2). Цей результат очевидний, якщо мати на увазі механічну інтерпретацію математичного сподівання.

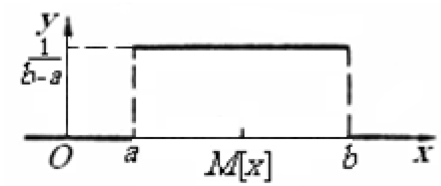


Рис. 4.9.2

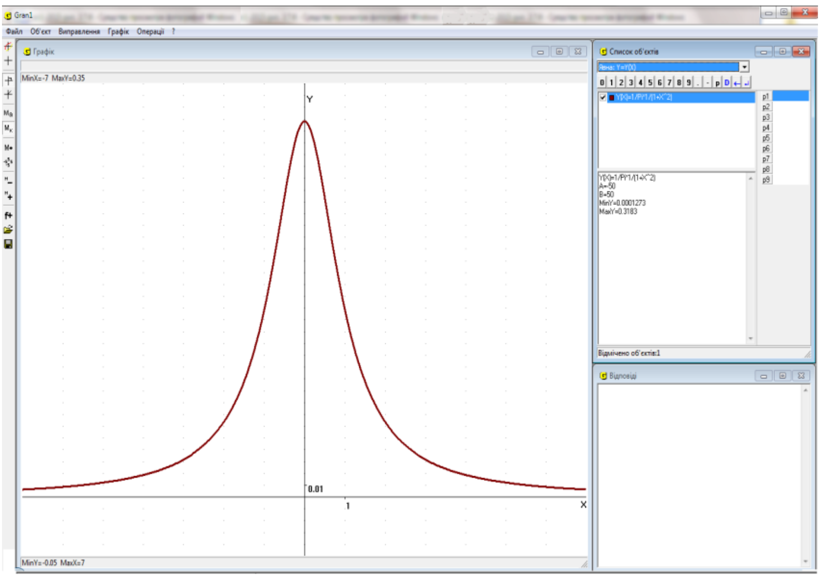


Рис. 4.9.3

Приклад 4.9.6. Щільність розподілу ймовірностей на множині значень випадкової величини X має вигляд (розподіл Коші, Рис. 4.9.3)

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \ x \in (-\infty; \infty).$$

Оскільки

$\frac{1}{\pi}\int\limits_{-\infty}^{\infty} x\frac{dx}{1+x^2} = 0,$ то в даному випадку $M[X=0]$.

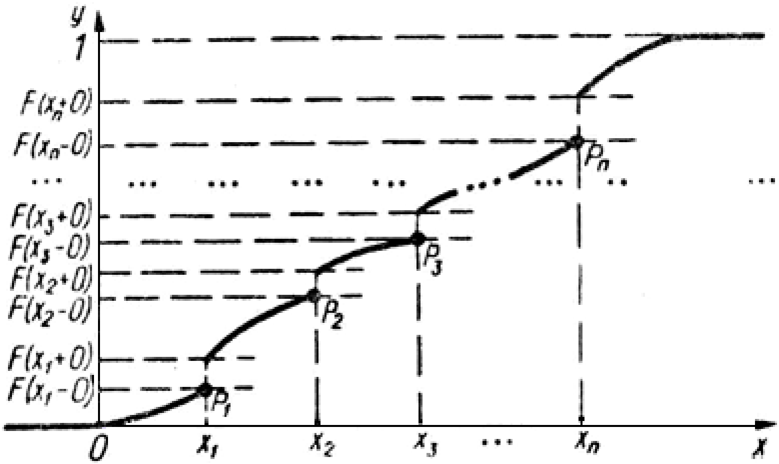


Рис. 4.9.4

Якщо розподіл імовірностей не є неперервним, тобто у функції розподілу $F_X(x)$ в деяких точках $x_1, x_2, \ldots, x_n$ є розриви $F_X(x_i - 0) - F_X(x_i) = p_i \neq 0$ (Рис. 4.9.4), причому $\sum\limits_{i=1}^{n} p_j < 1$, і існує невід'ємна обмежена функція $\tilde{f}_X(x)$ така, що

$$F_X(x) = \sum_{x_i < x} p_i + \int\limits_{-\infty}^{x} \tilde{f}_X(x)\,dx,$$

то

$$M[X] = \sum_{i=1}^{n} x_i \Delta F_X(x_i) + \int\limits_{-\infty}^{\infty} x\,\tilde{f}_X(x)\,dx = \sum_{i=1}^{n} x_i p_i + \int\limits_{-\infty}^{\infty} x\,\tilde{f}_X(x)\,dx.$$

В механічній інтерпретації в даному випадку частина p одиничної маси розподілена на дискретній множині точок

$x_1, x_2, \ldots, x_n$ так, що на точку x_i припадає маса $P_X(\{x_i\}) = p_i$, $\sum_{i=1}^{n} p_i = p$, а решта $1-p$ одиничної маси розподілена вздовж осі абсцис неперервно зі щільністю $\tilde{f}_X(x)$. В такому разі ймовірність попадання на проміжок $[a;b)$ обчислюватиметься за формулою

$$P_X([a,b)) = \sum_{x_i \in [a,b)} P_X(\{x_i\}) + \int_{[a,b)} \tilde{f}_X(x)dx = F_X(b) - F_X(a).$$

В механічній інтерпретації координатою центра мас дискретно розподіленої частини p від одиничної маси буде $\tilde{x}_1 = \dfrac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i}$, а координатою центра мас неперервно розподіленої частини $q = 1-p$ від одиничної маси буде

$$\tilde{x}_2 = \frac{\int_a^b x\, \tilde{f}_X(x)\, dx}{\int_a^b \tilde{f}_X(x)\, dx} = \frac{\int_a^b x\, \tilde{f}_X(x)\, dx}{1-p}.$$

Центр мас усієї одиничної маси з вказаним мішаним розподілом вздовж осі абсцис буде знаходитися в точці

$$\frac{\tilde{x}_1 p + \tilde{x}_2(1-p)}{p + (1-p)} = \tilde{x}_1 p + \tilde{x}_2(1-p) = \sum_{i=1}^{n} x_i p_i + \int_a^b x\, \tilde{f}_X(x)dx.$$

Із властивостей інтеграла Лебега випливають такі властивості математичного сподівання.

1. Математичне сподівання сталої дорівнює цій сталій, тобто якщо $X = c$, то $M[X] = c$.

2. Якщо X – інтегровна невід'ємна випадкова величина, то $M[X] \ge 0$.

3. Якщо $|X| \le c$, то і $M[X] \le c$.

Це випливає з так званої теореми про середнє: якщо стосовно вимірної функції X на вимірній множині Ω задовільняється нерівність $a \le X(E) \le b$, то

$$aP(\Omega) \le \int_{\Omega} X(E)\, P(dE) \le bP(\Omega).$$

4. Якщо X – інтегровна випадкова величина, то $M[aX] = a\,M[X]$.

5. Якщо випадкові величини X і Y інтегровні і $X \le Y$, то $M[X] \le M[Y]$.

6. Якщо $M[X]$ існує, то $|\,M[X]| \le M[|\,X\,|]$.

7. Якщо X і Y інтегровні випадкові величини, то $X+Y$ також інтегровна і

$$M[X+Y] = M[X] + M[Y].$$

8. Якщо $M[X]$ існує, то за кожного $A \in S$ також існує $M[XI_A]$. Якщо $M[X]$ скінченне, то $M[XI_A]$ також скінченне.

9. Якщо X і Y – інтегровні випадкові величини, то за будь-яких сталих a і b випадкова величина $aX+bY$ також інтегровна і

$$M[aX+bY] = aM[X] + bM[Y].$$

10. Якщо інтегровні випадкові величини X і Y незалежні, то

$$M[XY] = M[X]M[Y].$$

11. Якщо $\{X_i\}$ – зростаюча послідовність невід'ємних інтегровних випадкових величин $X_i \underset{i\to\infty}{\to} X$ і $\sup\limits_i M[X_i] < \infty$, то X також інтегровна і $\lim\limits_{i\to\infty} M[X_i] = M[X]$.

12. Якщо $\{X_i\}$ – послідовність невід'ємних інтегровних випадкових величин, то $M\left[\sum\limits_{i=1}^{\infty} X_i\right] = \sum\limits_{i=1}^{\infty} M[X_i]$ (досить покласти $Y_i = \sum\limits_{k=1}^{i} X_k$ і розглянути $\lim\limits_{i\to\infty} M[Y_i]$).

13. Якщо X – невід'ємна випадкова величина і $a > 0$, то виконується нерівність Чебишова

$$P(X \ge a) \le \frac{1}{a} M[X].$$

Для доведення досить розглянути випадкову величину

$$Y = \begin{cases} 0, & \text{коли } X < a; \\ a, & \text{коли } X \ge a. \end{cases}$$

Оскільки $X \ge Y$, то $M[X] \ge M[Y] = a\,P(X \ge a)$.

Приклад 4.9.7. Нехай $X(E)$, $E \in \Omega$, невід'ємна випадкова величина стосовно ймовірнісного простору (Ω, S, P), причому існує послідовність чисел x_i, $i \in N$, така, що $0 = x_0 < x_1 < \ldots < x_n < x_{n+1} \to +\infty$, коли $n \to \infty$, $x_{k+1} - x_k < \delta$, $k \in N$, і

$$\sum_{k=1}^{\infty} x_k \cdot P(E \mid x_{k-1} \le X(E) < x_k\}) = \sum_{k=1}^{\infty} x_k P(X^{-1}([x_{k-1}; x_k))) < +\infty.$$

1. Довести, що за будь-якої іншої послідовності чисел $\bar{x}_n$, $n \in N$, такої, що $\tilde{x}_0 = 0, \tilde{x}_n < \tilde{x}_{n+1} \to +\infty$, коли $n \to \infty$, і $\tilde{x}_{k+1} - \tilde{x}_k < \tilde{\delta}$, $k \in N$, правильні нерівності:

1) $\sum\limits_{k=1}^{\infty} \tilde{x}_{k-1} P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) \le \sum\limits_{k=1}^{\infty} x_k \; P(\{E \mid x_{k-1} \le X(E) < x_k\})$,

тобто

$$\sum_{k=1}^{\infty} \tilde{x}_{k-1} P(X^{-1}([\tilde{x}_{k-1}; \tilde{x}_k))) \le \sum_{k=1}^{\infty} \tilde{x}_k P(X^{-1}([\tilde{x}_{k-1}; x_k)));$$

2) $\sum\limits_{k=1}^{\infty} \tilde{x}_k P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) = \sum\limits_{k=1}^{\infty} \tilde{x}_{k-1} P(X^{-1}([\tilde{x}_{k-1}; \tilde{x}_k))) < +\infty$;

3) $0 \le \inf\limits_{(x_i)} \sum\limits_{k=1}^{\infty} x_k \; P(\{E \mid x_{k-1} \le X(E) < x_k\}) -$

$$-\sup_{(x_i)} \sum_{k=1}^{\infty} x_{k-1} \; P(\{E \mid x_{k-1} \le X(E) < x_k\}) \le \delta$$

або

$$0 \le \inf_{(x_i)} \sum_{k=1}^{\infty} x_k P(X^{-1}([x_{k-1}; x_k))) - \sup_{(x_i)} \sum_{k=1}^{\infty} \tilde{x}_{k-1} P(X^{-1}([x_{k-1}; x_k))) \le \delta \cdot$$

2. Невід'ємну *випадкову величину* $X(E)$, $E \in \Omega$, називають *інтегровною за Лебегом* відносно міри P, коли

$$\inf_{(x_i)} \sum_{k=1}^{\infty} x_k \; P(\{E \mid x_{k-1} \le X(E) < x_k\}) = \sup_{(x_i)} \sum_{k=1}^{\infty} x_{k-1} \; P(\{E \mid x_{k-1} \le X(E) < x_k\})$$

або

$$\inf_{(x_i)} \sum_{k=1}^{\infty} x_k P(X^{-1}([x_{k-1}; x_k))) = \sup_{(x_i)} \sum_{k=1}^{\infty} x_{k-1} P(X^{-1}([x_{k-1}; x_k))),$$

а спільне значення лівої і правої частин останньої рівності називають інтегралом Лебега від функції $X(E)$ відносно міри P і позначають $\int\limits_{\Omega} X(E)\,dP$. Довести, що $X(E)$ є інтегровною за Лебегом, коли існує принаймні одна послідовність x_i, $i \in N$, стосовно якої задовільняються вказані вище умови:
$0 = x_0$, $x_n < x_{n+1} \to +\infty$, $x_{k+1} - x_k < \delta$, $n \in N$, і

$$\sum_{k=1}^{\infty} x_k P(E \mid x_{k-1} \le X(E) < x_k\}) = \sum_{k=1}^{\infty} x_k P(X^{-1}([x_{k-1}; x_k))) < +\infty$$

Розв'язування

1. Насамперед зауважимо, що коли

$$A_k = \{E \in \Omega \mid \; x_{k-1} \le X(E) < x_k\} = X^{-1}([x_{k-1}; x_k))$$

і аналогічно $\tilde{A}_k = (\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) = X^{-1}([\tilde{x}_{k-1}; \tilde{x}_k))$, тоді події A_k (а також $\tilde{A}_k$) попарно несумісні і через їх об'єднання вичерпується простір Ω. Тому за будь-якого $E \in \Omega$ існує єдиний номер $k = k(E)$, (номер $\tilde{k} = \tilde{k}(E)$), за якого $\{E \mid E \in A_k\} = = X^{-1}([x_{k-1}; x_k))$, $\{E \mid E \in \tilde{A}_k\} = X^{-1}([\tilde{x}_{k-1}; \tilde{x}_k)))$, а тому $x_{k-1} \le X(E) < x_k$, $(\tilde{x}_{k-1} \le X(E) < \tilde{x}_k)$.

Звідси випливає, що

$$\sum_{k=1}^{\infty} \tilde{x}_{k-1} I_{\tilde{A}_k}(E) \le X(E) \le \sum_{k=1}^{\infty} x_k I_{A_k}(E), E \in \Omega,$$

а тому за будь-якого $n \in N$ існує $m = m(n) \in N$, таке, що

$$\sum_{k=1}^{\infty} \tilde{x}_{k-1} I_{\tilde{A}_k}(E) \le \sum_{k=1}^{m} x_k I_{A_k}(E), \;\; E \in \Omega.$$

Через останні суми визначаються прості випадкові величини такі, що

$$M[\sum_{n=1}^{n} \bar{x}_{k-1} I_{\bar{A}_k}(E)] = \sum_{n=1}^{n} \bar{x}_{k-1} \cdot P(\bar{A}_k) \le M[\sum_{k=1}^{m} x_k I_{A_k}(E)] =$$

$$= \sum_{k=1}^{m} x_k P(A_k) \le \sum_{k=1}^{\infty} x_k P(x_{k-1} \le X < x_k) < +\infty.$$

Враховуючи довільність числа $n \in N$, одержуємо:

$$\sum_{k=1}^{\infty} x_{k-1} P(\{E \mid x_{k-1} \le X(E) < x_k\}) \le \sum_{k=1}^{\infty} x_k P(\{E \mid x_{k-1} \le X(E) < x_k\}).$$

Цим доведено нерівність 1).

З нерівності 1), враховуючи, що $0 < x_k - x_{k-1} < \delta$, дістанемо нерівність:

$$4) \; \sum_{k=1}^{\infty} \tilde{x}_k P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) \le \sum_{k=1}^{\infty} \tilde{x}_k P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) +$$

$$+ \sum_{k=1}^{\infty} (\tilde{x}_k - \tilde{x}_{k-1}) P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) <$$

$$< \sum_{k=1}^{\infty} \tilde{x}_{k-1} P(\{E \mid \tilde{x}_{k-1} \le X(E) < \tilde{x}_k\}) + \delta < +\infty.$$

Цим доведено нерівність 2).

Оскільки

$$\inf_{(x_n)} \sum_{k=1}^{\infty} x_k P(\{E \mid x_{k-1} \le X(E) < x_k\}) - \sup_{(x_n)} \sum_{k=1}^{\infty} x_{k-1} P(\{E \mid x_{k-1} \le X(E) < x_k\}) \le$$

$$\le \sum_{k=1}^{\infty} x_k \; P(\{E \mid x_{k-1} \le X(E) < x_k\}) - \sum_{k=1}^{\infty} x_{k-1} P(\{E \mid x_{k-1} \le X(E) < x_k\}),$$

то з нерівності 4) випливає права частина нерівності 3). Ліва частина нерівності 3) випливає з нерівності 1).

2. Коли задовільняються умови твердження 2, тоді має місце нерівність 3), де в силу твердження 1 можна вважати, що число $\delta > 0$ – довільне. Тому можна спрямувати δ до нуля і дістати з нерівності 3) інтегровність функції $X(E)$ і формулу для обчислення її математичного сподівання:

$$M[X] = \int_{\Omega} X(E)\, dP .$$

Вправи для самостійного виконання

4.9.1. Перевірити, чи правильні твердження:

1. Стосовно кожної простої випадкової величини існує математичне сподівання.

2. Стосовно кожної дискретної випадкової величини існує математичне сподівання.

3. Математичне сподівання випадкової величини – це її середнє значення.

4. Математичне сподівання можна тлумачити і як статичний момент маси, розподіленої на множині Ω, і як центр цієї маси.

5. Кожна проста випадкова величина $X(E),\ E \in \Omega$, інтегровна за Лебегом-Стілтьєсом.

6. Кожна інтегровна за Лебегом-Стілтьєсом випадкова величина проста.

7. Якщо $h \in (0;1)$, $h_{n+1} < h_n \to 0$, коли $n \to \infty$, $X(E)$, $E \in \Omega$, – невід'ємна випадкова величина, $A_{n,k} = \{E \mid E \in \Omega, (k-1)h_n \le X(E) < kh_n\}$,

$$X_n(E) = \sum_{k=1}^{\infty} (k-1)\, h_n \; I_{A_{n,k}}(E) \text{ і } Y_n(E) = \sum_{k=1}^{\infty} kh_n \; I_{A_{n,k}}(E),$$

то: 1) $X_n(E) \le X_{n+1}(E) \le X(E) \le Y_{n+1}(E) \le Y_n(E),\ n \in N$;

2) $X_n(E) \to X(E)$ та $Y_n(E) \to X(E),\ (n \to \infty),\ E \in \Omega$;

3) $nI_{(X \ge n)}(E) \le \sum\limits_{k=[n h_n^{-1}]+1}^{\infty} (k-1)\, h_n \; I_{((k-1)h_n \le X < kh_n)}(E)$;

4) Якщо $\sum\limits_{k=1}^{\infty} (k-1)\, h_n \, P(\{E \mid (k-1)h_n \le X(E) < kh_n\}) < +\infty$, то

$$\lim_{n \to \infty} nP(\{E \mid X(E) \ge n\}) = 0;$$

5) Якщо $\sum_{k=1}^{\infty}(k-1)P(\{E \mid (k-1)h_n \le X(E) < kh_n\}) < \infty$ за якогось $h_n \in (0;1)$, то $\sum_{k=1}^{\infty}(k-1)P(\{E \mid (k-1)h_n^* \le X(E) < kh_n^*\}) < \infty$ за будь-якого $h_n^* \in (0;1)$, $h_{n+1}^* < h_n^* \to 0$, коли $n \to \infty$.

6) Якщо $S_*(X, h_n) = \sum_{k=1}^{\infty}(k-1)h_n\ P(\{E \mid (k-1)h_n \le X(E) < kh_n\})$ нижня інтегральна сума функції $X(E)$, а

$$S^*(X, h_n^*) = \sum_{k=1}^{\infty} k\, h_n^*\ P(\{E \mid (k-1)h_n^* \le X(E) < kh_n^*\})$$

верхня інтегральна сума функції $X(E)$, то

$$S_*(X, h_n) \le S^*(X, h_n^*) \le S_*(X, h_n^*) + h_n^*.$$

7) Якщо $S_*(X, h_n) = \sum_{k=1}^{\infty}(k-1)h_n\ P(\{E \mid (k-1)h_n \le X(E) < k\, h_n\}) < \infty$, то існує скінченна границя $\lim_{n\to\infty} S_*(X, h_n)$, яка не залежить від послідовності чисел $h_n \in (0;1)$, що прямує до 0.

8) Якщо $f(x) = x, x \in R^1$, то

$$S^*(X,\ h_n) = S^*(f^*,\ h_n) = \sum_{k=1}^{\infty}(k-1)h_n\ \ \tilde{P}(E \mid (k-1)h_n \le f < k\,h_n),$$

де $\tilde{P}$ – ймовірність, визначена на σ-алгебрі $\mathscr{B}(R^1)$ і породжена за функцією $F_X(x) = P_X(-\infty;\ x)$ розподілу ймовірностей на множині значень випадкової величини X.

8. Випадкова величина $X(E),\ E \in \Omega$, інтегровна за ймовірнісною мірою P тоді й тільки тоді, коли функція $f(x) = x,\ x \in R$, інтегровна за мірою $\tilde{P}$, породженою через випадкову величину X.

9. Стосовно будь-якої випадкової величини $X(E),\ E \in \Omega$,

$$M[X] = \int_{\Omega} X(E)\, dP = \int_{-\infty}^{\infty} x\, d\, F, \text{ де } F(x) = \tilde{P}((-\infty; x)).$$

10. Твердження 9 правильне лише стосовно інтегровної випадкової величини $X(E),\ E \in \Omega$.

11. Якщо існує $M[X]$, то існує й $M[|X|]$.

12. Твердження, обернене до 11, правильне.

13. Якщо за $n \to \infty$ $X_n(E) \to X(E)$ за всіх $E \in \Omega$, то

$$M[X_n] \to M(X).$$

14. $M[X]$ існує тоді й тільки тоді, коли існує $M[X\, I_A]$ за будь якої події A.

15. За математичним сподіванням породжується нова міра на просторі подій S, що є абсолютно неперервною відносно міри P.

16. За кожної міри зі знаком існує щільність її розподілу.

17. Стосовно кожної випадкової величини, відносно якої існує математичне сподівання, правильна рівність: $M[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx$.

18. За кожного розподілу ймовірностей на множині значень випадкової величини існує центр розсіювання.

19. За кожного розподілу ймовірностей на множині значень інтегровної випадкової величини існує центр розсіювання ймовірностей.

20. Твердження 19 є правильним стосовно абсолютно неперервної випадкової величини.

4.9.2. Довести, що коли $M[|X_1 - X_2|] = 0$, то випадкові величини майже напевно однакові, тобто $P(\{E \mid E \in \Omega,\ X_1(E) \neq X_2(E)\}) = 0$.

4.10. Математичне сподівання функції від випадкової величини

Нехай $Z=\psi(X)$ S_X/S_Z-випадкова величина, тобто S_X/S_Z-вимірна функція випадкового аргумента X, де S_Z і S_X простори подій із відповідних ймовірнісних просторів (Ω_Z, S_Z, P_Z), (Ω_X, S_X, P_X), $\Omega_Z=\psi(\Omega_X)$ – образ множини Ω_X за відображенням $Z=\psi(X)$, і потрібно знайти математичне сподівання $M[Z]$ випадкової величини Z або, що те саме, координату $z_c=M[Z]$ центра розподілу (чи розсіювання) ймовірностей на множині Ω_Z значень випадкової величини Z.

Розглянемо спочатку випадок, коли X проста випадкова величина, тобто множина Ω_X значень випадкової величини X скінченна: $\Omega_X=\{x_1, x_2, ..., x_k\}$. Тоді і випадкова величина Z буде проста із скінченною множиною значень $\Omega_Z=\{z_1, z_2, ..., z_m\}$, де $m\le k$. Знайшовши ймовірності $P_Z(\{z_i\})=P_X(\psi^{-1}(\{z_i\}))$, за означенням математичного сподівання простої випадкової величини дістанемо

$$M[Z]=\sum_{i=1}^{m} z_i P_Z(\{z_i\})=\sum_{i=1}^{m} z_i P_X(\psi^{-1}(\{z_i\})).$$

Оскільки $\psi^{-1}(\{z_i\})\cap\psi^{-1}(\{z_j\})=\varnothing$, коли $i\ne j$, $\bigcup_{i=1}^{m}\psi^{-1}(\{z_i\})=$ $=\psi^{-1}(\Omega_Z)=\Omega_X=\{x_1, x_2, ..., x_k\}$, $z_i=\psi(x_j)$, коли $x_j\in\psi^{-1}(\{z_i\})$, то

$$M[Z]=\sum_{i=1}^{m} z_i P_Z(\{z_i\})=\sum_{i=1}^{m} z_i P_X(\psi^{-1}(\{z_i\}))=$$

$$=\sum_{i=1}^{m} z_i\left(\sum_{x_j\in\psi^{-1}(\{z_i\})} P_X(\{x_j\})\right)=\sum_{j=1}^{k}\psi(x_j)P_X(\{x_j\}).$$

Таким чином, математичне сподівання $M[Z]$ випадкової величини $Z=\psi(X)$ в разі простої випадкової величини X можна знайти за двома способами.

1-й спосіб. Визначити розподіл ймовірностей на множині $\Omega_Z=\{z_1, z_2, ..., z_m\}$ значень випадкової величини Z, тобто спочатку знайти ймовірності $P_Z(\{z_i\})$, $i\in\overline{1,m}$, і далі знайти математичне сподівання $M[Z]$ за формулою

$$M[Z]=\sum_{i=1}^{m} z_i P_Z(\{z_i\}). \tag{4.10.1}$$

2-й спосіб. Не визначаючи розподіл ймовірностей на множині $\Omega_Z = \{z_1, z_2, ..., z_m\}$ значень випадкової величини Z, знайти математичне сподівання $M[Z]$ за формулою

$$M[Z] = \sum_{j=1}^{k} \psi(x_j) P_X(\{x_j\}). \tag{4.10.2}$$

Нехай тепер задана щільність $f_X(x)$ розподілу ймовірностей на множині $\Omega_X = (-\infty; \infty)$ значень випадкової величини X.

Нехай разом з тим множина Ω_Z значень функції $Z = \psi(X)$ скінченна (функція $z = \psi(x)$ кусково стала), $\Omega_Z = \{z_1, z_2, ..., z_m\}$, причому $\psi^{-1}(\{z_i\})$, $i \in \overline{1, m}$, є множиною, відносно якої існує інтеграл $\int\limits_{\psi^{-1}(\{z_i\})} f_X(x) dx = P_X(\psi^{-1}(\{z_i\}))$, $i \in \overline{1, m}$.

Тоді

$$M[Z] = \sum_{i=1}^{m} z_i P_Z(\{z_i\}) = \sum_{i=1}^{m} z_i P_X(\psi^{-1}(\{z_i\})) =$$

$$= \sum_{i=1}^{m} z_i \int\limits_{\psi^{-1}(\{z_i\})} f_X(x) dx = \sum_{i=1}^{m} \int\limits_{\psi^{-1}(\{z_i\})} z_i f_X(x) dx =$$

$$= \sum_{i=1}^{m} \int\limits_{\psi^{-1}(\{z_i\})} \psi(x) f_X(x) dx = \int\limits_{\Omega_X} \psi(x) f_X(x) dx.$$

Отже у випадку, коли множина Ω_X значень випадкової величина X неперервна і відома щільність $f_X(x)$ розподілу ймовірностей на множині Ω_X, а випадкова величина $Z = \psi(X)$ проста із скінченною множиною значень $\Omega_Z = \{z_1, z_2, ..., z_m\}$, математичне сподівання $M[Z]$ випадкової величини Z можна знайти за двома способами.

1-й спосіб. Визначити розподіл ймовірностей на множині $\Omega_Z = \{z_1, z_2, ..., z_m\}$ значень випадкової величини Z, тобто знайти ймовірності $P_Z(\{z_i\})$, $i \in \overline{1, m}$, і далі знайти математичне сподівання $M[Z]$ за формулою

$$M[Z] = \sum_{i=1}^{m} z_i P_Z(\{z_i\}).$$

2-й спосіб. Не визначаючи розподіл ймовірностей на множині $\Omega_Z = \{z_1, z_2, ..., z_m\}$ значень випадкової величини Z, знайти математичне сподівання $M[Z]$ за формулою

$$M[Z] = \int\limits_{\Omega_X} \psi(x) f_X(x) dx.$$

Нехай тепер множина значень випадкової величини X $\Omega_X = (-\infty;\infty)$, простір подій $S_X = \mathscr{B}(R^1)$, в який входять всі проміжки виду $(-\infty; x)$, $x \in (-\infty;\infty)$, і нехай задана щільність $f_X(x)$ розподілу ймовірностей на множині значень Ω_X, множина Ω_Z значень випадкової величини $Z = \psi(X)$ є проміжком $\Omega_Z = [z_0; z_m)$, функція $Z = \psi(X)$ борелівська (див. §4.1, §4.2), а простір подій $S_Z = \mathscr{B}(R^1)$ і за будь-якого $[\alpha;\beta) \subset \Omega_Z$ існує інтеграл

$$\int\limits_{\psi^{-1}([\alpha;\beta))} f_X(x)dx = P_X(\psi^{-1}([\alpha;\beta))) .$$

Нагадаємо, що *функція* $z = \psi(x)$, $z \in R^1$, $x \in R^n$, *називається борелівською, якщо прообраз* $\psi^{-1}(A)$ *будь якої борелівської множини* $A \in \mathscr{B}(R^1)$ *знову є борелівська множина із* $\mathscr{B}(R^n)$, *тобто* $\psi^{-1}(A) \in \mathscr{B}(R^n)$, *коли* $A \in \mathscr{B}(R^1)$.

Поділимо проміжок $[z_0; z_m)$ на досить дрібні проміжки $[z_{i-1}; z_i)$, $i \in 1,...,k$, однакової довжини $z_i - z_{i-1} = \dfrac{z_k - z_0}{k} = h$, де $z_k = z_m$, і розглянемо випадкову величину $\tilde{Z} = \tilde{\psi}(X)$, яка набуває значень z_{i-1}, коли $z = \psi(x)$ знаходиться на проміжку $[z_{i-1}; z_i)$, тобто $\tilde{\psi}(x) = z_{i-1}$, коли $x \in \psi^{-1}([z_{i-1};\, z_i))$, $i \in \overline{1,k}$.

Тоді $P_{\tilde{Z}}(\{z_i\}) = P_Z([z_{i-1};\, z_i)) = P_X(\psi^{-1}([z_{i-1}; z_i)))$, $i \in \overline{1,k}$.

В даному разі $\Omega_{\tilde{Z}} = \{z_0, z_1, z_2, ..., z_k\}$, в $S_{\tilde{Z}}$ разом з множинами $\varnothing$ і $\Omega_{\tilde{Z}}$ включаються всеможливі об'єднання підмножин $\{z_i\} \subset \Omega_{\tilde{Z}}$ із одного, із двох, із трьох, і т.д., із $(k-1)$ доданків, $\tilde{Z}^{-1}(\{z_{i-1}\}) = Z^{-1}([z_{i-1}; z_i))$. Очевидно функція $\tilde{Z} = \tilde{\psi}(X)$ $S_X / S_{\tilde{Z}}$-вимірна, тобто $\tilde{Z}$ – проста $S_X / S_{\tilde{Z}}$-випадкова величина відносно просторів (Ω_X, S_X, P_X) і $S_{\tilde{Z}}$.

В такий спосіб функція $\psi(x)$ наближено замінюється кусково сталою функцією $\tilde{\psi}(x)$, яка набуває сталих значень z_{i-1} на прообразах $\psi^{-1}([z_{i-1}; z_i))$ відрізків $[z_{i-1}; z_i)$.

Міркуючи аналогічно до попереднього, одержимо

$$M[\tilde{Z}] = \sum_{i=1}^{k} z_{i-1} P_{\tilde{Z}}(\{z_i\}) = \sum_{i=1}^{k} z_{i-1} P_Z([z_{i-1}; z_i)) =$$

$$=\sum_{i=1}^{k} z_{i-1}P_X(\psi^{-1}([z_{i-1};z_i))) = \sum_{i=1}^{k} z_{i-1} \int\limits_{\psi^{-1}([z_{i-1};z_i))} f_X(x)dx =$$

$$=\sum_{i=1}^{k} \int\limits_{\psi^{-1}([z_{i-1};z_i))} z_{i-1} f_X(x)dx =$$

$$=\sum_{i=1}^{k} \int\limits_{\psi^{-1}([z_{i-1};\, z_i))} \tilde{\psi}(x) f_X(x)dx = \int\limits_{\Omega_X} \tilde{\psi}(x) f_X(x)dx\,. \qquad (4.10.3)$$

Якщо за необмеженого збільшення k існує границя суми $\sum\limits_{i=1}^{k} z_{i-1}P_Z([z_{i-1};\, z_i))$, то її позначають через $\int\limits_{\Omega_Z} zP_Z(dz)$ і називають інтегралом від функції $\varphi(z)=z$ на множині Ω_Z за мірою P_Z, або математичним сподіванням випадкової величини $Z=\psi(X(E))$, яке позначають також через $M[Z]$.

Із збільшенням k $\int\limits_{\Omega_X}\tilde{\psi}(x) f_X(x)dx$ прямуватиме до інтеграла $\int\limits_{\Omega_X}\psi(x) f_X(x)dx$.

Аналогічно границю суми $\sum\limits_{i=1}^{k}\psi(x_i)P_X(\psi^{-1}([z_{i-1};z_i)))$, $x_i\in\psi^{-1}([z_{i-1};z_i))$, за необмеженого збільшення k (припускається, що така границя існує) позначають через $\int\limits_{\Omega_X}\psi(x)P_X(\psi^{-1}(dz))$.

Таким чином

$$M[Z]=\int\limits_{\Omega_Z} zP_Z(dz)=\int\limits_{\Omega_X}\psi(x)P_X(\psi^{-1}(dz)) = \int\limits_{\Omega_X}\psi(x) f_X(x)dx\,. \qquad (4.10.4)$$

Отже, математичне сподівання функції $Z=\psi(X)$ випадкового аргумента X можна знайти за двома способами.

1-й спосіб. Враховуючи, що $Z=\psi(X)$, і виходячи із заданої ймовірнісної міри P_X, за якою описується розподіл ймовірностей на множині Ω_X значень випадкової величини X, визначити ймовірнісну міру $P_Z([\alpha;\beta))=P_X(\psi^{-1}([\alpha;\beta)))$, $[\alpha;\beta)\subset\Omega_Z$, за якою описується розподіл ймовірностей на множині Ω_Z значень випадкової величини Z, і далі знайти математичне сподівання $M[Z]$ випадкової величини Z за формулою

$$M[Z]=\int\limits_{\Omega_Z} zP_Z(dz)\,. \qquad (4.10.5)$$

2-й спосіб. Не визначаючи ймовірнісної міри P_Z, за якою описується розподіл ймовірностей на множині Ω_Z значень випадкової величини Z, знайти математичне сподівання $M[Z]$ випадкової величини Z за формулою

$$M[Z] = \int_{\Omega_X} \psi(x) P_X(\psi^{-1}(dz)) . \qquad (4.10.6)$$

Зазначимо, що в наведених міркуваннях розмірність множини Ω_X значень S/S_X-випадкової величини X ніде не враховувалась, тому наведені твердження залишаються правильними і для R^n-значних S/S_X-випадкових величин X за довільного скінченного $n \ge 1$.

Зауважимо, що коли відомі щільності $f_Z(z)$ та $f_X(x)$ розподілів ймовірностей відповідно на множинах Ω_Z та Ω_X, тоді формули (4.10.5) та (4.10.6) набувають вигляду

$$M[Z] = \int_{\Omega_Z} z P_Z(dz) = \int_{\Omega_Z} z f_Z(z) dz , \qquad (4.10.7)$$

та

$$M[Z] = \int_{\Omega_X} \psi(x) P_X(\psi^{-1}(dz)) = \int_{\Omega_X} \psi(x) f_X(x) dx . \qquad (4.10.8)$$

Якщо X R^n-значна випадкова величина, тобто $X = (X_1, X_2, ..., X_n)$, тоді $x = (x_1, x_2, ..., x_n)$, $dx = dx_1 dx_2 ... dx_n$, $\Omega_X \subset R^n$, запис $\int_{\Omega_X} \psi(x) f_X(x) dx$ означає інтеграл відповідної розмірності.

Приклад 4.10.1. Нехай $Z = \psi(X, Y)$ – функція двох випадкових аргументів, задана на квадраті

$$\Omega_{(X,Y)} = \{(x, y) \mid x \in [-3;3], y \in [-3;3]\} ,$$

залежністю між координатами *x*, *y*, *z*:

$$z = \psi(x, y) = \begin{cases} h\left(1 - \frac{|x|}{3}\right), \text{ коли } x \in [-3;3],\ y \in [-|x|;|x|], \\ h\left(1 - \frac{|y|}{3}\right), \text{ коли } y \in [-3;3],\ x \in [-|y|;|y|], \end{cases}$$

$S_{(X,Y)}$ – сукупність підмножин множини $\Omega_{(X,Y)}$, породжена за прямокутниками виду $[a;b) \times [c;d)$, де $a \in [-3;3]$, $b \in [-3;3]$, $a \le b$, $c \in [-3;3]$, $d \in [-3;3]$, $c \le d$, $\Omega_Z = (-\infty;\infty)$, тобто $S_{(X,Y)} = \mathscr{B}(R^2)$, $S_Z = \mathscr{B}(R^1)$ – σ-алгебра борелівських множин на числовій прямій $(-\infty;\infty)$.

Довести, що Z – $S_{(X,Y)} / S_Z$-випадкова величина відносно просторів $(\Omega_{(X,Y)}, S_{(X,Y)}, P_{(X,Y)})$ і S_Z, якщо на множині $\Omega_{(X,Y)}$ значень випадкового вектора (X,Y) ймовірності розподілені рівномірно, тобто щільність розподілу ймовірностей

$$f_{(X,Y)}(x,y) = \begin{cases} \dfrac{1}{36}, & \text{коли } (x,y) \in \Omega_{(X,Y)}, \\ 0, & \text{коли } (x,y) \overline{\in} \Omega_{(X,Y)}. \end{cases}$$

Знайти $M[Z]$ за двома способами.

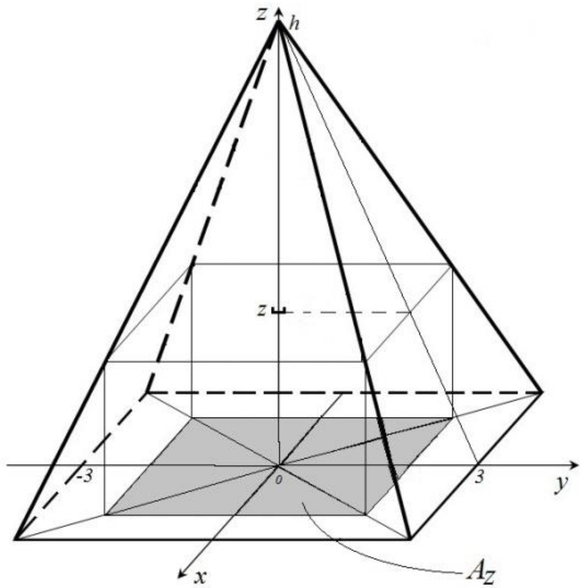


Рис. 4.10.1

Очевидно, що коли $z \le 0$, то $\psi^{-1}((-\infty; z)) = \varnothing \in S$. Коли $0 < z \le h$, то $\psi^{-1}((-\infty; z)) = \Omega_{(X,Y)} \setminus A_z \in S_{(X,Y)} = \mathscr{B}(R^2)$, де A_Z – квадрат із сторонами, паралельними до сторін квадрата $\Omega_{(X,Y)}$ і такий, що є проекцією на площину $z = 0$ квадрата, що отримується в перерізі фігури $z = \psi(x,y)$, $(x,y) \in \Omega_{(X,Y)}$, площиною, перпендикулярною до осі Oz і віддаленою від початку координат на віддаль z (див. Рис. 4.10.1). Коли $h < z$, тоді $\psi^{-1}((-\infty; z)) = \Omega_{(X,Y)} \in S_{(X,Y)}$. Отже, функція $Z = \psi(X,Y)$ $S_{(X,Y)} / S_Z$-вимірна, тобто випадкова величина.

З Рис. 4.10.2 видно, що коли $|x| = 2$, $y \in [-2; 2]$, тоді

$z_1 = h\left(1-\dfrac{2}{3}\right)$; якщо $|y|=2$, $x\in[-2;\,2]$, тоді $z_1 = h\left(1-\dfrac{2}{3}\right)$.

Отже прообразом точки $z_1 = h\left(1-\dfrac{2}{3}\right)$ буде множина точок

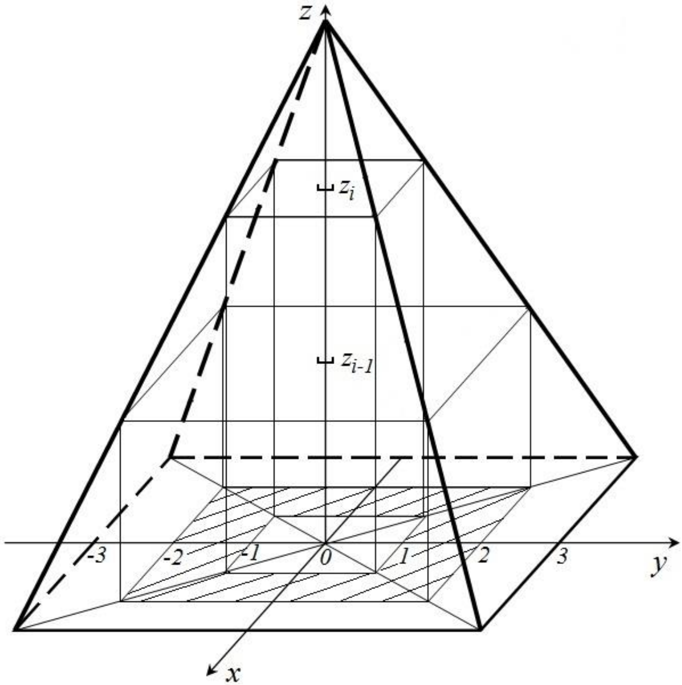


Рис. 4.10.2

$$\psi^{-1}(z_1) = \psi^{-1}\left(h\left(1-\frac{2}{3}\right)\right) = \{(x,y)\mid\ |x|=2,\ y\in[-2;2]\}\cup$$
$$\cup\{(x,y)\mid |y|=2,\ x\in[-2;2]\}\,.$$

В геометричному тлумаченні $\psi^{-1}(z_1) = \psi^{-1}\left(h\left(1-\dfrac{2}{3}\right)\right)$ є множина точок на сторонах квадрата з центром в початку координат, із сторонами, паралельними до осей координат *Ox* та *Oy*, і довжиною сторони, рівною 4 (Рис. 4.10.3).

Аналогічно прообразом точки $z_2 = h\left(1-\dfrac{1}{3}\right)$ буде множина точок

$$\psi^{-1}(z_2) = \psi^{-1}\left(h\left(1-\frac{1}{3}\right)\right) = \{(x,y)\mid |x|=1,\ y\in[-1;1]\}\cup$$

$$\bigcup\{(x,y)\big|\,|y|=1, x\in[-1;1]\}.$$

В геометричному тлумаченні $\psi^{-1}(z_2)=\psi^{-1}\left(h\left(1-\frac{1}{3}\right)\right)$ є множина точок на сторонах квадрата з центром в початку координат із сторонами, паралельними до осей координат Ox та Oy, і довжиною сторони, рівною 2 (Рис. 4.10.3).

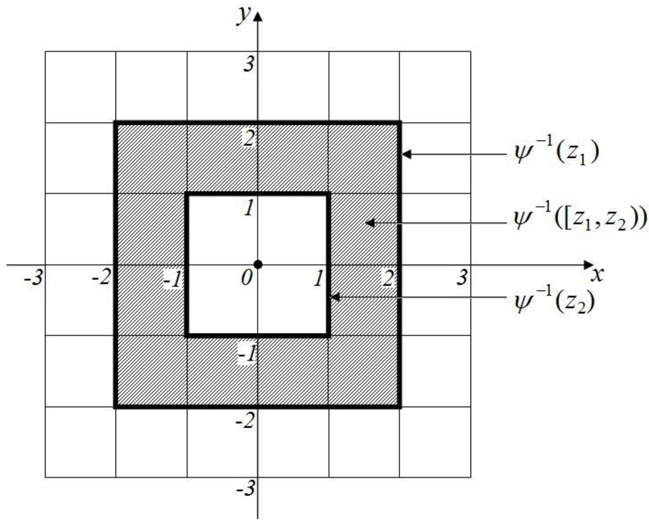


Рис. 4.10.3

Тоді прообразом відрізка $[z_1;z_2)$ буде множина точок (Рис. 4.10.2, 4.10.3):

$$\psi^{-1}([z_1;z_2))=\{(x,y)\big|\,|x|\in[1;2),\ y\in[-|x|;|x|]\}\cup$$
$$\cup\{(x,y)\big|\,|y|\in[1;2),\ x\in[-|y|;|y|]\}.$$

В даному разі $z_2-z_1=h\left(1-\frac{1}{3}\right)-h\left(1-\frac{2}{3}\right)=\frac{2}{3}h-\frac{1}{3}h=\frac{1}{3}h$, а на множині $\Omega_{(X,Y)}=\{(x,y)\big|\,x\in[-3;3], y\in[-3;3]\}$ значення Z змінюються від значення 0, коли $|x|=3$, $y\in[-3;3]$ або $|y|=3$, $x\in[-3;3]$, тобто в точках на сторонах квадрата з центром в початку координат, сторонами, паралельними до осей координат Ox та Oy і довжиною сторони 6 (Рис. 4.10.3), до значення h в точці (0, 0). Таким чином $\Omega_Z=[0;h]$.

Нехай тепер множину значень Ω_Z поділено на частини $[z_0;z_1)$, $[z_1;z_2)$, $[z_2;z_3)$, де $z_0=0$, $z_1=h\cdot\frac{1}{3}$, $z_2=h\cdot\frac{2}{3}$, $z_3=h$.

Тоді за формулами (4.10.5), (4.10.6) одержимо

$$M[Z]\approx z_0P_Z([z_0;z_1))+z_1P_Z([z_1;z_2))+z_2P_Z([z_2;z_3))=$$

$$= z_0 P_X(\psi^{-1}[z_0; z_1))) + z_1 P_X(\psi^{-1}[z_1; z_2))) + z_2 P_X(\psi^{-1}[z_2; z_3))),$$

де $\psi^{-1}([z_0; z_1))$ – множина точок між квадратами із довжинами сторін 6 та 4 (Рис. 4.10.3), $\psi^{-1}([z_1; z_2))$ – множина точок між квадратами із довжинами сторін 4 і 2 (Рис. 4.10.3), $\psi^{-1}([z_2; z_3))$ – множина точок в квадраті із стороною 2 (Рис. 4.10.3).

Суму вигляду

$$\sum_{i=1}^{k} z_{i-1} P_X(\psi^{-1}[z_{i-1}; z_i)) \tag{4.10.9}$$

називають інтегральною сумою Лебега-Стільтьєса стосовно функції $Z = \psi(X)$ на множині Ω_X за мірою P_X. В розглядуваному прикладі в такій сумі є три доданки, а випадкова величина Z є функцією від R^2-значної випадкової величини X, тобто від випадкової пари (X,Y): $Z = \psi(X,Y)$.

Якщо тепер в розглядуваному прикладі накрити $\Omega_{(X,Y)}$ сіткою прямокутників із довжинами сторін Δx і Δy і перебирати такі прямокутники вздовж горизонталей – спочатку найнижчу, потім слідкуючу за нею, обчислювати значення функції $z = \psi(x_j, y_i)$ в одній із точок (x_j, y_i) кожного такого прямокутника, множити на ймовірнісну міру такого прямокутника (ймовірність попадання в прямокутник) і знаходити суму всіх так знайдених величин, то одержимо той самий результат, що і раніше, поданий однак у вигляді $\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \psi(x_j, y_i) P_{(X,Y)}(\Delta_{ij})$, де Δ_{ij} – прямокутник, що знаходиться на i-й горизонталі і j-й вертикалі, m_1 – кількість горизонталей, m_2 – кількість вертикалей.

Суму вигляду

$$\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \psi(x_j, y_i) P_{(X,Y)}(\Delta_{ij})$$

називають інтегральною сумою Рімана стосовно функції $\psi(x, y)$ на множині Ω_X за мірою $P_{(X,Y)}$.

Якщо в розглядуваному прикладі розглянути сітку із квадратиків із стороною довжини 1 (Рис. 4.10.3) і пронумерувати рядки квадратиків від 1 до 6, а також стовпчики від 1 до 6, а квадратик в i-му рядкові і j-му стовпчикові позначати через Δ_{ij}, тоді одержимо інтегральну суму

$$\sum_{i=1}^{6}\sum_{j=1}^{6} \psi(x_j, y_i) P_{(X,Y)}(\Delta_{ij}).$$

Очевидно,

$$\psi^{-1}([z_0, z_1)) = \Delta_{11} + \Delta_{12} + \Delta_{13} + \Delta_{14} + \Delta_{15} + \Delta_{16} + \Delta_{21} + \Delta_{26} + \Delta_{31} + \\ + \Delta_{36} + \Delta_{41} + \Delta_{46} + \Delta_{51} + \Delta_{56} + \Delta_{61} + \Delta_{62} + \Delta_{63} + \Delta_{64} + \Delta_{65} + \Delta_{66},$$

а функція $\tilde{\psi}(x, y)$ у всіх вказаних квадратиках набуває одного і того самого значення z_0,

$$\psi^{-1}([z_1; z_2)) = \Delta_{22} + \Delta_{23} + \Delta_{24} + \Delta_{25} + \Delta_{32} + \Delta_{35} + \Delta_{42} + \Delta_{45} + \\ + \Delta_{52} + \Delta_{53} + \Delta_{54} + \Delta_{55},$$

а функція $\tilde{\psi}(x, y)$ у всіх вказаних квадратиках набуває одного і того самого значення z_1,

$$\psi^{-1}([z_2; z_3)) = \Delta_{33} + \Delta_{34} + \Delta_{43} + \Delta_{44},$$

а функція $\tilde{\psi}(x, y)$ у всіх вказаних квадратиках набуває одного і того самого значення z_2.

Тому

$$z_0 P_{(X,Y)}(\psi^{-1}([z_0, z_1))) + z_1 P_{(X,Y)}(\psi^{-1}([z_1, z_2))) + z_2 P_{(X,Y)}(\psi^{-1}([z_2, z_3))) =$$
$$= z_0 P_{(X,Y)}(\Delta_{11} + \Delta_{12} + \Delta_{13} + \Delta_{14} + \Delta_{15} + \Delta_{16} + \Delta_{21} + \Delta_{26} + \Delta_{31} +$$
$$+ \Delta_{36} + \Delta_{41} + \Delta_{46} + \Delta_{51} + \Delta_{56} + \Delta_{61} + \Delta_{62} + \Delta_{63} + \Delta_{64} + \Delta_{65} + \Delta_{66}) +$$
$$+ z_1 P_{(X,Y)}(\Delta_{22} + \Delta_{23} + \Delta_{24} + \Delta_{25} + \Delta_{32} + \Delta_{35} + \Delta_{42} + \Delta_{45} + \Delta_{52} + \Delta_{53} +$$
$$+ \Delta_{54} + \Delta_{55}) + z_2 P_{(X,Y)}(\Delta_{33} + \Delta_{34} + \Delta_{43} + \Delta_{44}) =$$
$$= \sum_{i=1}^{6} \sum_{j=1}^{6} \psi(x_j, y_i) P_{(X,Y)}(\Delta_{ij}),$$

причому в кожному доданку з номерами i, j $(x_j, y_i) \in \Delta_{ij}$.

Подібні міркування покладені в основу другого способу обчислення математичного сподівання випадкової величини Z (див. формули 4.10.7, 4.10.8)

$$M[Z] = \int_{\Omega_X} \psi(x) P_X(\psi^{-1}(dz)) = \int_{\Omega_X} \psi(x) f_X(x) dx.$$

Опис функції розподілу ймовірностей на множині Ω_Z значень випадкової величини Z в розглядуваному прикладі набуває вигляду (див. Рис. 4.10.1):

$$F_Z(z) = P_Z((-\infty; z)) = P_{(X,Y)}(Z^{-1}((-\infty; z))) =$$

$$
=\begin{cases} 0, \text{ коли } z \le 0, \\ \dfrac{36-36\left(1-\frac{z}{h}\right)^2}{36} = 1-\left(1-\dfrac{z}{h}\right)^2, \text{ коли } 0 \le z \le h,. \\ 1, \text{ коли } h \le z. \end{cases}
$$

Звідси випливає, що розподіл ймовірностей на множині Ω_Z значень випадкової величини Z неперервний і опис щільності цього розподілу набуває вигляду

$$
f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} 0, \text{ коли } z \overline{\in} [0,h], \\ \dfrac{2}{h}\left(1-\dfrac{z}{h}\right), \text{ коли } z \in [0,h]. \end{cases}
$$

Обчислюючи $M[Z]$ за формулою 4.10.7

$$
M[Z] = \int\limits_{\Omega_Z} z f_Z(z) dz = \int\limits_{-\infty}^{\infty} z f_Z(z) dz ,
$$

в розглядуваному випадку одержуємо

$$
M[Z] = \int\limits_0^h z \cdot \frac{2}{h}\left(1-\frac{z}{h}\right) dz = \frac{2}{h}\left(\frac{z^2}{2} - \frac{z^3}{3h}\right)\Bigg|_0^h = \frac{2}{h}\left(\frac{h^2}{2} - \frac{h^2}{3}\right) = \frac{1}{3} h .
$$

Обчислюючи $M[Z]$ за другим способом (за формулою 4.10.8)

$$
M[Z] = \iint\limits_{\Omega_{(X,Y)}} \psi(x,y) f_{(X,Y)} dxdy ,
$$

в розглядуваному прикладі одержуємо

$$
M[Z] = \iint\limits_{\Omega_{(X,Y)}} \psi(x,y) \frac{1}{36} dxdy,
$$

Враховуючи, що піраміду, подану на Рис. 4.10.1, можна подати як об'єднання чотирьох рівновеликих пірамід, основи яких лежать між діагоналями основи вихідної піраміди, одержимо

$$
M[Z] = 4 \cdot \frac{1}{36} \int\limits_0^3 \left\{ \int\limits_{-x}^{x} h\left(1-\frac{x}{3}\right) dy \right\} dx = \frac{4}{36} \int\limits_0^3 h\left(1-\frac{x}{3}\right) \cdot 2x dx =
$$

$$= \frac{8}{36} h \left(\frac{x^2}{2} - \frac{x^3}{9} \right) \Bigg|_0^3 = \frac{8}{36} h \left(\frac{9}{2} - \frac{27}{9} \right) = \frac{8}{36} h \left(\frac{81 - 54}{18} \right) = \frac{8}{36} h \cdot \frac{27}{18} =$$

$$= \frac{8}{36} \cdot h \cdot \frac{3}{2} = \frac{4}{12} h = \frac{1}{3} h .$$

Розглянемо окремо випадок, коли $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$, $\Omega_Y = R^1$, $S_Y = \mathscr{B}(R^1)$.

Нехай $Y = \psi(X)$ – функція випадкового аргумента X і треба знайти математичне сподівання $M[Y]$ випадкової величини Y за умови, що розподіл імовірностей на множині значень випадкової величини X відомий.

Розглянемо спочатку випадок, коли множина Ω_X значень випадкової величини X є дискретною скінченною множиною значень $\Omega_X = \{x^{(1)}, x^{(2)}, \ldots, x^{(k)}\}$.

Очевидно тоді й множина значень випадкової величини Y буде скінченна $\Omega_Y = \{y_1, y_2, \ldots, y_m\}$, де $m \le k$.

Знайшовши ймовірності $P_Y(\{y_i\}) = P_X(\psi^{-1}(y_i))$, дістанемо

$$M[Y] = \sum_{i=1}^{m} y_i P_Y(\{y_i\}) = \sum_{i=1}^{m} y_i P_X(\psi^{-1}(y_i)) .$$

Оскільки

$$\psi^{-1}(y_i) \cap \psi^{-1}(y_j) = \varnothing, \text{ коли } i \ne j ,$$

$$\bigcup_{i=1}^{m} \psi^{-1}(y_i) = \{x^{(1)}, x^{(2)}, \ldots, x^{(k)}\}, \; y_i = \psi(x^{(j)}), \text{ коли } x^{(j)} \in \psi^{-1}(y_i) ,$$

то

$$M[Y] = \sum_{i=1}^{m} y_i P_X(\psi^{-1}(y_i)) =$$

$$= \sum_{i=1}^{m} y_i \left(\sum_{x^{(j)} \in \psi^{-1}(y_i)} P_X(\{x^{(j)}\}) \right) = \sum_{j=1}^{k} \psi(x^{(j)}) P_X(\{x^{(j)}\}) .$$

Наприклад, якщо опис дискретного одновимірного розподілу імовірностей (Рис. 4.10.4) поданий в Табл. 4.10.1.

Табл. 4.10.1

$x^{(k)}$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$	$x^{(7)}$
p_k	p_1	p_2	p_3	p_4	p_5	p_6	p_7

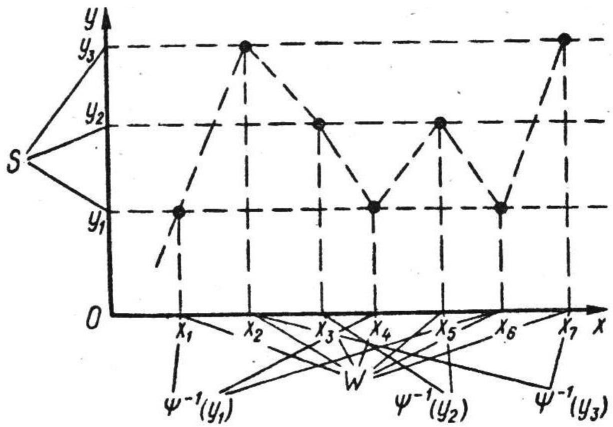


Рис. 4.10.4

$$y_1 = \psi(x^{(1)}) = \psi(x^{(4)}) = \psi(x^{(6)}),\ y_2 = \psi(x^{(3)}) = \psi(x^{(5)}),$$
$$y_3 = \psi(x^{(2)}) = \psi(x^{(7)}),$$

то

$$P_Y(\{y_1\}) = P_X(\psi^{-1}(y_1)) = P_X(\{x^{(1)},\ x^{(4)},\ x^{(6)}\}) =$$
$$= P_X(\{x^{(1)}\}) + P_X(\{x^{(4)}\}) + P_X(\{x^{(6)}\}) = p_1 + p_4 + p_6,$$
$$P_Y(\{y_2\}) = P_X(\psi^{-1}(y_2)) = P_X(\{x^{(3)},\ x^{(5)}\}) = p_3 + p_5,$$
$$P_Y(\{y_3\}) = P_X(\psi^{-1}(y_3)) = P_X(\{x^{(2)},\ x^{(7)}\}) = p_2 + p_7.$$

Таким чином

$$M[Y] = \sum_{i=1}^{3} y_i P_Y(\{y_i\}) = \sum_{i=1}^{3} y_i P_X(\psi^{-1}(y_i)) = y_1 P_X(\{x^{(1)},\ x^{(4)},\ x^{(6)}\}) +$$
$$+ y_2 P_X(\{x^{(3)},\ x^{(5)}\}) + y_3 P_X(\{x^{(2)},\ x^{(7)}\}) = y_1(P_X(\{x^{(1)}\}) + P_X(\{x^{(4)}\}) +$$
$$+ P_X(\{x^{(6)}\})) + y_2(P_X(\{x^{(3)}\}) + P_X(\{x^{(5)}\})) + y_3(P_X(\{x^{(2)}\}) + P_X(\{x^{(7)}\})) =$$
$$= y_1 P_X(\{x^{(1)}\}) + y_3 P_X(\{x^{(2)}\}) + y_2 P_X(\{x^{(3)}\}) + y_1 P_X(\{x^{(4)}\}) + y_2 P_X(\{x^{(5)}\}) +$$
$$+ y_1 P_X(\{x^{(6)}\}) + y_3 P_X(\{x^{(7)}\}) = \psi(x^{(1)}) P_X(\{x^{(1)}\}) + \psi(x^{(2)}) P_X(\{x^{(2)}\}) +$$
$$+ \psi(x^{(3)}) P_X(\{x^{(3)}\}) + \psi(x^{(4)}) P_X(\{x^{(4)}\}) + \psi(x^{(5)}) P_X(\{x^{(5)}\}) +$$
$$+ \psi(x^{(6)}) P_X(\{x^{(6)}\}) + \psi(x^{(7)}) P_X(\{x^{(7)}\}) = \sum_{i=1}^{7} \psi(x^{(i)}) P_X(\{x^{(i)}\}).$$

Отже, математичне сподівання $M[Y]$ випадкової величини $Y = \psi(X)$ у випадку простої випадкової величини X можна знайти за двома способами.

І спосіб. Визначити розподіл імовірностей на множині значень випадкової величини Y, тобто спочатку знайти ймовірності $P_Y(\{y_i\})$ і далі знайти математичне сподівання $M[Y]$ за формулою

$$M[Y] = \sum_{i=1}^{m} y_i P_Y(\{y_i\}).$$

II спосіб. Не визначаючи розподіл імовірностей на множині значень випадкової величини Y, знайти математичне сподівання $M[Y]$ за формулою

$$M[Y] = \sum_{x^{(i)} \in X(\Omega)} \psi(x^{(i)}) P_X(\{x^{(i)}\}),$$

де $X(\Omega)$ – множина можливих значень випадкової величини X.

Зазначимо, що в наведених міркуваннях розмірність випадкової величини X не враховувалась, а тому зроблені висновки залишаються правильними і стосовно R^n-значних випадкових величин X.

Теорема (про заміну змінних під знаком інтеграла Лебега). Нехай (Ω, S) і (W, Q) – вимірні простори, $X = X(E)$ – S/Q-вимірна функція із значеннями в W.

Нехай P – ймовірнісна міра на (Ω, S), P_X – ймовірнісна міра на (W, Q), що індукується за випадковою величиною $X = X(E)$ із простору (Ω, S, P):

$$P_X(A) = P(\{E \mid X(E) \in A\}) = P(X^{-1}(A)),\ A \in Q. \qquad (4.10.10)$$

Тоді за довільної Q-вимірної функції $y = \psi(x)$,

$$\int_A \psi(x) P_X(dx) = \int_{X^{-1}(A)} \psi(X(E)) P(dE),\ A \in Q \qquad (4.10.11)$$

(в тому розумінні, що коли існує один з інтегралів, то визначений і інший, і вони співпадають).

Справді, нехай $A \in Q$, $\psi(x) = I_B(x)$, $B \in Q$. Тоді рівність (4.10.11) набуває вигляду

$$P_X(AB) = P(X^{-1}(A) \cap X^{-1}(B)). \qquad (4.10.12)$$

Ця рівність випливає з (4.10.10) і того, що $X^{-1}(A) \cap X^{-1}(B) = X^{-1}(A \cap B)$. Згідно з рівністю (4.10.12), рівність (4.10.11) правильна стосовно невід'ємних простих функцій, а отже, за теоремою про монотонну збіжність, правильна і стосовно довільних невід'ємних S/Q-вимірних функцій.

Якщо $(W, Q) = (R^1, \mathscr{B}(R^1))$, $X = X(E)$ – випадкова величина з розподілом імовірностей P_X на множині Ω_X її значень,

$\psi(x)$ – борелівська функція, та існує один з інтегралів $\int_A \psi(x)P_X(dx)$ або $\int_{X^{-1}(A)} \psi(X(E))P(dE)$, то

$$\int_A \psi(x)P_X(dx) = \int_{X^{-1}(A)} \psi(X(E))P(dE).$$

Зокрема, коли $A = R^1$, тоді

$$M[\psi(X(E))] = \int_\Omega \psi(X(E))P(dE) = \int_{R^1} \psi(x)P_X(dx).$$

Міра P_X однозначно визначається за функцією розподілу F_X. Тому інтеграл Лебега $\int_{R^1} \psi(x)P_X(dx)$ часто позначають $\int_{R^1} \psi(x)F_X(dx)$ або $\int_{R^1} \psi(x)dF_X(x)$ і називають *інтегралом Лебега-Стільтьєса* (за мірою, що відповідає функції розподілу $F_X(x)$).

Нехай існує невід'ємна інтегровна борелівська функція $f_X(x)$ така, що $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$, де інтеграл розглядається як інтеграл Лебега за мірою Лебега на множині $(-\infty; x)$.

Тоді

$$M[\psi(X(E))] = \int_{-\infty}^{\infty} \psi(x) f_X(x)dx,$$

де інтеграл розглядається як інтеграл Лебега від функції $\psi(x) f_X(x)$ за мірою Лебега.

Справді, якщо $\psi(x) = I_B(x)$, $B \in \mathscr{B}(R^1)$, то рівність

$$\int_B \psi(x)P_X(dx) = \int_{X^{-1}(B)} \psi(X(E))P(dE)$$

набуває вигляду

$$P_X(B) = \int_B f_X(x)\,dx = \int_{X^{-1}(B)} P(dE),\ B \in \mathscr{B}(R^1).$$

Ця рівність випливає з того, що за функцією розподілу $F_X(x)$ породжується єдина ймовірнісна міра $P_X(B)$, $B \in \mathscr{B}(R^1)$, така, що за довільних x_1, x_2

$$P_X([x_1; x_2)) = F_X(x_2) - F_X(x_1);$$

а також з того, що

$$F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)\,dx\,.$$

В загальному випадку доведення аналогічне доведенню теореми про заміну змінних під знаком інтеграла Лебега.

Нехай, наприклад, X – деяка R^n-значна випадкова величина (n-вимірний випадковий вектор); $n \geq 1$, $y = \psi(x) : R^n \to R^1$ – обмежена борелівська функція така, що $y_0 \leq \psi(x) < y_{m+1}$, де y_0 і y_{m+1} – деякі скінченні числа, тобто всі значення випадкової величини $Y(E) = \psi(X(E))$ знаходяться у межах від y_0 до y_{m+1}. Поділимо інтервал $[y_0; y_{m+1})$ точками поділу $y_1, y_2, \ldots, y_m$ на досить дрібні частини $[y_0; y_1),\ [y_1; y_2), \ldots, [y_{m-1}; y_m),\ [y_m; y_{m+1})$.

Введемо до розгляду випадкову величину $\tilde{Y}^{(m)}$, вважаючи, що випадкова величина $\tilde{Y}^{(m)}$ набуває значення y_i, якщо значення випадкової величини Y лежить в інтервалі $[y_i; y_{i+1})$.

В такий спосіб дістають просту випадкову величину $\tilde{Y}^{(m)}$ з множиною значень $y_0, y_1, \ldots, y_m$, причому

$$P_Y(\tilde{Y}^{(m)} = y_i) = P_X(\psi^{-1}([y_i; y_{i+1})))\,.$$

Тут випадкова величина Y може бути як дискретною, так і неперервною.

Математичне сподівання $M[\tilde{Y}^{(m)}]$ випадкової величини $\tilde{Y}^{(m)}$ можна знайти за формулою

$$M[Y^{(m)}] = \sum_{i=0}^{m} y_i P_y([y_i; y_{i+1})) = \sum_{i=0}^{m} y_i P_X(\psi^{-1}([y_i; y_{i+1})))\,. \quad (4.10.13)$$

Якщо тепер спрямувати довжину $d = \max\limits_{1 \leq i \leq m}(y_{i+1} - y_i)$ найдовшого з інтервалів $[y_i; y_{i+1})$ до нуля, то границя

$$M[Y] = \lim_{d \to 0} \sum_{i=0}^{m} y_i\, P_X(\psi^{-1}([y_i; y_{i+1})))$$

і буде математичним сподіванням випадкової величини Y.

Зазначимо, що $\sum\limits_{i=0}^{m} y_i\, P_X(\psi^{-1}([y_i; y_{i+1})))$ є інтегралом Лебега $\int\limits_{R^n} \tilde{\psi}_m(x) P_X(dx)$ від функції $\tilde{\psi}_m(x)$ за мірою P_X, де $\tilde{\psi}_m(x) = y_i$, коли $x \in \psi^{-1}([y_i; y_{i+1}))$, $i = 0, 1, \ldots, m$, а

$$\lim_{d\to 0}\int_{R^n}\tilde{\psi}(x)P_X(dx)=\lim_{d\to 0}\sum_{i=0}^{m} y_i P_X(\psi^{-1}([y_i;y_{i+1})$$

є інтегралом Лебега $\int_{R^n}\psi(x)P_X(dx)$ від функції $\psi(x)$ на множині R^n за мірою P_X.

Розглянемо окремо випадок, коли випадкова величина X є R^1-значною неперервною, а розподіл ймовірностей на множині значень описується за допомогою функції розподілу $F_X(x)$.

Вважатимемо, що існує деякий скінченний відрізок $[a;b]$ такий, що $F_X(a)=0$, $F_X(b)=1$, тобто ймовірність того, що значення випадкової величини X лежатимуть за межами відрізка $[a;b]$, дорівнює нулю. Крім того, нехай значення випадкової величини $Y=\psi(X)$ не виходять за межі інтервала $[y_0;y_{m+1})$ (Рис. 4.10.5).

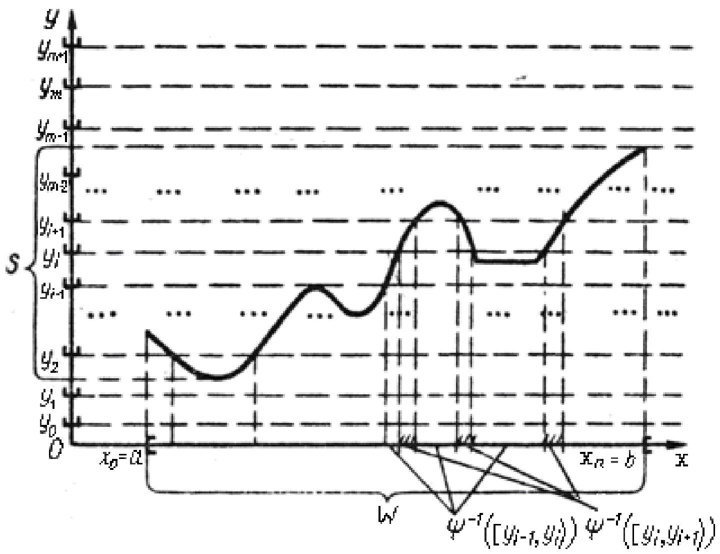


Рис. 4.10.5

Тоді, враховуючи (4.10.13), знайдемо

$$M[\tilde{Y}^m]=\sum_{i=0}^{m} y_i P_X(\psi^{-1}([y_i;\ y_{i+1})))=\sum_{i=0}^{m} y_i\left(\sum_{(\alpha_k;\beta_k)\subset\psi^{-1}([y_i;y_{i+1}))} P_X((\alpha_k;\beta_k))\right),$$

де $(\alpha_k;\beta_k)$ – інтервали, об’єднання яких вичерпує множину $\psi^{-1}([y_i;y_{i+1}))$, з якої вилучено точки α_k і β_k, тобто

$$\bigcup_{(\alpha_k;\beta_k)\subset\psi^{-1}([y_i;y_{i+1}))}(\alpha_k;\beta_k)=\psi^{-1}([y_i;y_{i+1}))\setminus\bigcup_k\{\alpha_k\}\setminus\bigcup_k\{\beta_k\}.$$

Як уже зазначалося, коли випадкова величина X неперервна, то ймовірності того, що значення цієї випадкової величини лежатимуть в інтервалах $[\alpha_k;\beta_k]$, $[\alpha_k;\beta_k)$, $(\alpha_k;\beta_k]$, $(\alpha_k;\beta_k)$, однакові.

Оскільки $\psi^{-1}([y_i,y_{i+1}))\cap\psi^{-1}([y_j,y_{j+1}))=\varnothing$, коли $i\neq j$, то

$$\bigcup_{i=0}^{m}\bigcup_{(\alpha_k;\beta_k)\subset\psi^{-1}([y_i;y_{i+1}))}(\alpha_k;\beta_k)=\bigcup_{k=1}^{l}(\alpha_k;\beta_k)=$$

$$=\bigcup_{i=0}^{m}\psi^{-1}([y_i;y_{i+1}))\setminus\bigcup_{k=1}^{l}\{\alpha_k\}\setminus\bigcup_{k=1}^{l}\{\beta_k\},$$

де l – загальна кількість інтервалів $(\alpha_k;\beta_k)$ в об’єднанні $\bigcup_{i=0}^{m}\psi^{-1}([y_i,y_{i+1}))$. Тому

$$M[\tilde{Y}^m]=\sum_{i=0}^{m}y_i\left(\sum_{(\alpha_k,\beta_k)\subset\psi^{-1}([y_i,y_{i+1}))}P_X((\alpha_k,\beta_k))\right)=\sum_{k=1}^{l}\tilde{\psi}_m(x_k)P_X(X\in((\alpha_k;\beta_k))=$$

$$=\sum_{k=1}^{l}\tilde{\psi}_m(x_k)(F_X(\beta_k)-F_X(\alpha_k)),\ x_k\in[\alpha_k;\beta_k].$$

Нехай $d=\max(y_{i+1}-y_i)$ прямує до нуля. Тоді

$$M[Y]=\lim_{d\to 0}M\{\tilde{Y}^{(m)}]=\lim_{d\to 0}\sum_{k=1}^{l}\tilde{\psi}_m(x_k)(F_X(\beta_k)-F_X(\alpha_k)),$$

$$x_k\in[\alpha_k;\beta_k] \tag{4.10.14}$$

Якщо границя $M[Y]$ існує, то вона і є математичним сподіванням випадкової величини Y.

$\lim\limits_{d\to 0}\sum\limits_{k=1}^{l}\tilde{\psi}_m(x_k)(F_X(\beta_k)-F_X(\alpha_k))$ називається *інтегралом Лебега-Стільтьєса від функції* $\psi(x)$ *на відрізку* $[a;b]$ за мірою, що породжується за функцією $F_X(x)$, і позначається

$\int\limits_a^b \psi(x)\, dF_X(x)$.

Оскільки $F_X(\beta_k) - F_X(\alpha_k) = 0$, коли $(\alpha_k;\ \beta_k) \bigcap [a; b] = \varnothing$, то

$$\int\limits_a^b \psi(x)\, dF_X(x) = \int\limits_{-\infty}^{\infty} \psi(x)\, dF_X(x)\,.$$

Якщо функція $F_X(x)$ абсолютно неперервна й диференційовна майже всюди на $[a;\ b]$, то (4.10.14) можна подати у вигляді

$$M[Y] = \lim_{d\to 0} M\{\tilde{Y}^{(m)}] = \lim_{d\to 0} \sum_{k=1}^{l} \tilde{\psi}_m(x_k)(F'_X(\xi_k)\, m([\alpha_k;\ \beta_k]),$$
$$\xi_k \in [\alpha_k;\ \beta_k]),$$

де $m([\alpha_k; \beta_k])$ – міра Лебега проміжка $[\alpha_k; \beta_k]$, тобто інтеграл Лебега-Стільтьєса $\int\limits_a^b \psi(x)\, dF_X(x)$ в цьому випадку дорівнюватиме інтегралу Лебега від функції $\psi(x) F'_X(x)$ на відрізку $[a;b]$ за мірою Лебега.

Нехай $\Phi(x)$ – деяка неперервна зліва функція з обмеженою зміною, задана на відрізку $[a;b]$; $a = x_0 < x_1 < x_2 < \ldots < x_m = b$ – деякий поділ відрізка $[a;b]$ на часткові відрізки $[x_{i-1}; x_i)$; $\psi(x)$ – довільна функція; ξ_i – довільна точка з відрізка $[x_{i-1}; x_i]$. Тоді якщо існує границя

$$\lim_{\max(x_i - x_{i-1})\to 0} \sum_{i=0}^{m} \psi(\xi_i)[\Phi(x_i) - \Phi(x_{i-1})],$$

яка не залежить ні від способу поділу відрізка $[a;b]$ на часткові відрізки $[x_{i-1}; x_i)$, ні від вибору точок $\xi_i \in [x_{i-1}; x_i)$, то вона називається *інтегралом Рімана-Стільтьєса від функції* ψ *за функцією* Φ і позначається $\int\limits_a^b \psi(x)\, d\Phi(x)$.

Як відомо, якщо функція $\psi(x)$ неперервна на відрізку $[a;b]$, то її інтеграл Рімана-Стільтьєса за функцією Φ існує і співпадає з відповідним інтегралом Лебега-Стільтьєса.

Якщо функція $\Phi(x)$ диференційовна, то інтеграл Рімана-Стільтьєса $\int_a^b \psi(x)\, d\Phi(x)$ від функції ψ за функцією Φ співпадає з інтегралом Рімана $\int_a^b \psi(x)\, \Phi'(x) dx$ від функції $\psi(x)\Phi'(x)$ на відрізку $[a;b]$.

Таким чином, якщо множина значень випадкової величини X неперервна, то математичне сподівання випадкової величини $Y = \psi(X)$ можна знайти за двома способами.

I спосіб. Знайти розподіл імовірностей на множині $\Omega_Y = \psi(\Omega_X)$ значень випадкової величини Y, тобто знайти функцію розподілу ймовірностей $F_Y(y) = P_Y((-\infty; y))$ або щільність розподілу ймовірностей $f_Y(y)$ і далі знайти математичне сподівання за формулою

$$M[Y] = \int_{-\infty}^{\infty} y dF_Y(y), \tag{4.10.15}$$

або за формулою (якщо існує щільність $f_Y(y)$)

$$M[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy.$$

II спосіб. Не визначаючи розподіл ймовірностей на множині значень випадкової величини $Y = \psi(x)$, знайти математичне сподівання $M[Y]$ за формулою

$$M[Y] = \int_{-\infty}^{\infty} \psi(x)\, dF_X(x),$$

або (якщо відома щільність розподілу ймовірностей $f_X(x)$) за формулою

$$M[Y] = \int_{-\infty}^{\infty} \psi(x)\, f_X(x) dx, \tag{4.10.16}$$

де $F_X(x)$ і $f_X(x)$ – відповідно функція розподілу та щільність розподілу ймовірностей на множині значень випадкової величини X.

Формули (4.10.15) і (4.10.16) правильні і тоді, коли випадкова величина X є R^n-значною (Рис. 4.10.6).

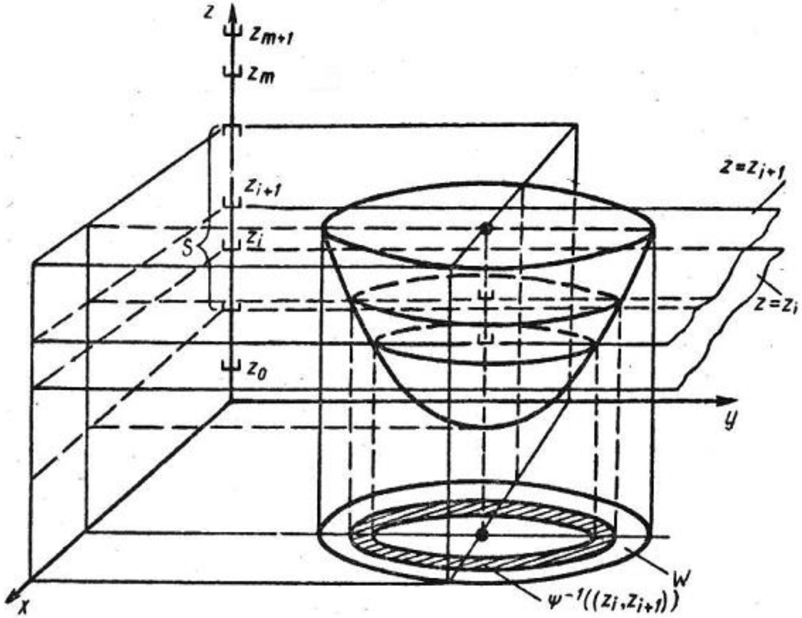


Рис. 4.10.6

Якщо множина значень випадкової величини $Y = \psi(X)$ або випадкової величини X не є неперервною, проте у функцій розподілу ймовірностей на множинах значень цих величин недискретні множини значень, то можна окремо розглянути дискретно розподілену ймовірність і окремо – неперервно розподілену ймовірність.

Якщо функція розподілу ймовірностей $F_X(x)$ є кусково-сталою (функцією стрибків, тобто міра $P_X(A)$, що породжується за функцією $F_X(x)$, є дискретною мірою), то інтеграл Лебега-Стільтьєса

$$\int_{-\infty}^{\infty} \psi(x)\, dF_X(x)$$

зводиться до суми

$$\sum_i \psi(x_i)\, p_i \ ,$$

де x_i – точки розриву функції $F_X(x)$, а p_i – стрибки функції $F_X(x)$ у точках x_i, тобто ймовірності $P_X(\{x_i\}) = p_i$.

Якщо $F_X(x)$ є сумою функції стрибків і абсолютно неперервної функції, то інтеграл Лебега-Стільтьєса за мірою P_X, що породжується за функцією $F_X(x)$, набуває вигляду

$$\sum_i \psi(x_i)\, p_i + \int_a^b \psi(x)\, F'(x) dx,$$

де $\int_a^b \psi(x)\, F'(x) dx$ – інтеграл за мірою Лебега від функції $\psi(x)\, F'(x)$.

Таким чином, математичне сподівання випадкової величини $Y = \psi(x)$ можна подати як інтеграл Лебега-Стільтьєса

$$M[Y] = \int_{-\infty}^{\infty} \psi(x)\, dF_X(x).$$

Якщо $Y = g(X)$, де X – випадковий вектор, і існує щільність $f_X(x_1, x_2, \ldots, x_n)$ n-вимірного розподілу ймовірностей на множині значень випадкового вектора X, то

$$M[g(X)] = \int_{R^n} g(x) f_X(x) dx$$

за довільної борелівської функції $g(x)$, $x \in R^n$, стосовно якої одна з частин останньої рівності має зміст.

Якщо, наприклад, $g(X) = X_i$ – i-та координата вектора X і відома щільність $f_X(x_1, x_2, \ldots, x_n)$, то

$$M[g(X)] = M[X_i] =$$

$$= \int_{R^n} x_i f_X(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) dx_1 \ldots dx_{i-1} dx_i dx_{i+1} \ldots dx_n =$$

$$= \int_{-\infty}^{\infty} x_i \left(\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n \right) dx_i = \int_{-\infty}^{\infty} x_i f_{X_i}(x_i) dx_i.$$

Точка $C(M[X_1], \ldots, M[X_n])$ називається *центром розсіювання* ймовірностей на множині значень випадкового вектора X. Якщо $X \in R^s$, то під $M[X]$ розуміють $(M[X_1], M[X_2], \ldots, M[X_s])$.

Нехай X і Y – незалежні випадкові вектори з функціями розподілу ймовірностей $F_X(X)$, $F_Y(Y)$, $x \in R^s$, $y \in R^l$, $g(x, y)$ – дійсна борелівська функція на R^{s+l}. Тоді вираз стосовно $M[g(x, y)]$ можна подати у вигляді

$$M[g(x, y)] = \int_{R^s \times R^l} g(x, y) F_X(dx) F_Y(dy),$$

причому інтеграл можна звести до повторного інтеграла

$$M[g(x,y)] = \int\limits_{R^s} \left(\int\limits_{R^l} g(x,y) F_Y(dy) \right) F_X(dx).$$

В такому разі спочатку обчислюють математичне сподівання випадкової величини $g(x,Y)$, де x – фіксоване, а потім здобутий результат розглядають як функцію від x. Таким чином

$$M[g(X,\,Y)] = M[M[g(x,\,Y)]|_{x=X}].$$

Розглянемо кілька прикладів.

Приклад 4.10.2. Задано функцію випадкового аргумента $Y = \psi(X)$, де (Рис. 4.10.7)

$$\psi(x) = \begin{cases} 0.5, & \text{коли} \quad x \le 0.5; \\ 1-x, & \text{коли} \quad x \ge 0.5. \end{cases}$$

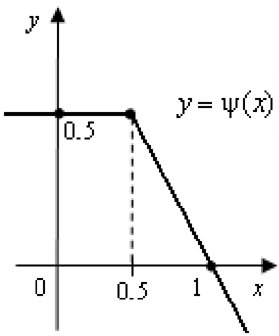


Рис. 4.10.7

Знайти математичне сподівання $M[Y]$ випадкової величини Y, якщо на множині значень випадкової величини X розподіл ймовірностей рівномірний на проміжку $[0;1]$.

Розв'яжемо задачу за двома способами.

I спосіб. Знайдемо розподіл імовірностей на множині значень випадкової величини Y:

$$F_Y(y) = P_Y((-\infty;\,y)) = P_X(\psi^{-1}((-\infty;\,y))).$$

Оскільки прообразом множини $(-\infty;y)$ за заданого відображення $y = \psi(x)$ є множина $(1-y;+\infty)$, коли $y \le 0.5$, і множина $(-\infty;\infty)$, коли $y \le 0.5$, а також враховуючи, що на множині значень випадкової величини X розподіл ймовірностей рівномірний на відрізку $[0;1]$ (Рис. 4.10.8), дістаємо

$$F_Y(y)=\begin{cases}0, \text{ коли } y\le 0;\\ y, \text{коли } 0\le y\le 0.5;\\ 1, \text{ коли } 0.5<y.\end{cases}\quad \frac{d}{dy}F_Y(y)=\begin{cases}0, \text{ коли } y\,\overline{\in}\,[0;\ 0.5];\\ 1, \text{ коли } y\in(0;\ 0.5).\end{cases}$$

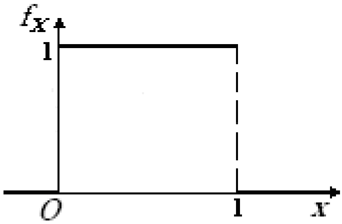


Рис. 4.10.8

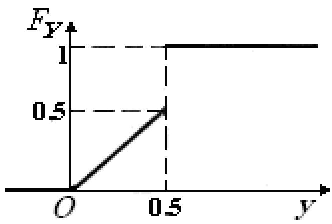


Рис. 4.10.9

Графік функції $F_Y(y)$ зображено на Рис. 4.10.9.

В механічній інтерпретації – половина одиничної маси зосереджена в точці $y=0.5$, інша половина розподілена рівномірно на відрізку $[0\,;0.5]$.

Таким чином

$$M[Y]=0.5\cdot 0.5+\int_0^{0.5} yF_Y'(y)\,dy=0.25+\left.\frac{y^2}{2}\right|_0^{0.5}=0.25+0.125=0.375\,.$$

II спосіб. За формулою (4.10.16) з врахуванням вигляду функцій $\psi(x)$ і $f_X(x)$ дістаємо

$$M[Y]=\int_0^{0.5}0.5\cdot 1dx+\int_{0.5}^{1}(1-x)\cdot 1dx=0.5x\Big|_0^{0.5}+\left.\left(x-\frac{x^2}{2}\right)\right|_{0.5}^{1}=$$

$$=0.5\cdot 0.5+1-\frac{1}{2}-\left(0.5-\frac{(0.5)^2}{2}\right)=0.375\,.$$

Приклад 4.10.3. Задано функцію $Z=|X|+|Y|$, де (X,Y) – R^2-значна випадкова величина з рівномірним розподілом імовірностей у квадраті

$$G=\{(x,y)\big|\ |x|+|y|\le 1\}$$

(Рис. 4.10.10). Знайти математичне сподівання $M[Z]$.

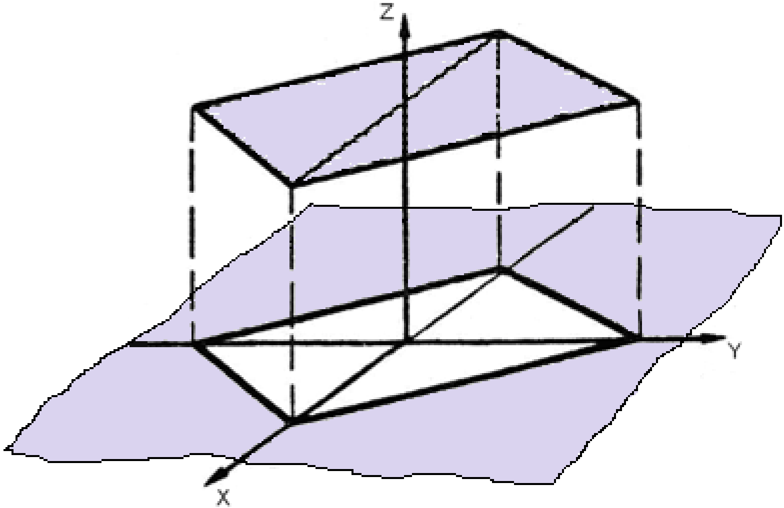


Рис. 4.10.10

І спосіб. Знайдемо розподіл імовірностей на множині значень випадкової величини Z:

$$F_Z(z) = P_Z((-\infty; z)) = P_{(X,Y)}(\psi^{-1}((-\infty; z))),$$

де $z = \psi(x, y) = |x| + |y|$, а

$$\psi^{-1}((-\infty; z)) = \{(x, y) \mid \psi(x, y) \in (-\infty; z)\} = \{(x, y) \mid |x| + |y| < z\}.$$

Очевидно, що прообраз множини $(-\infty; z)$ порожній, коли $z \le 0$, і є квадратом (Рис. 4.10.11):

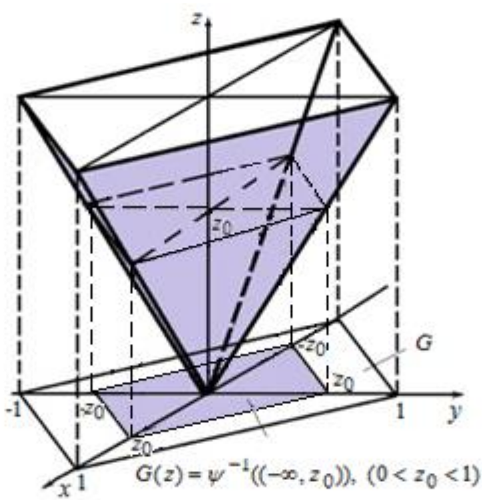


Рис. 4.10.11

$$G(z)=\{(x,y)\big|\ \ |x|+|y|<z\},\text{ коли } z>0.$$

Площа квадрата $G(z)$ дорівнює $2z^2$. Площа квадрата G дорівнює 2. Звідси, коли $G(z)\subset G$, тобто $z\le 1$, тоді

$$P_{(X,Y)}((\psi^{-1}((-\infty;z)))=\iint\limits_{G(z)} f_{(X,Y)}(x,y)\,dxdy=z^2.$$

Таким чином,

$$F_Z(z)=\begin{cases}0,\text{ коли } z\le 0;\\ z^2,\text{коли } 0\le z\le 1;\\ 1,\text{ коли } 1\le z.\end{cases}\qquad f_Z(z)=\begin{cases}0,\text{ коли } z\,\overline{\in}\,[0;1];\\ 2z,\text{ коли } z\in(0;1).\end{cases}$$

Графіки функцій $F_Z(z)$, $f_Z(z)$ зображено на Рис. 4.10.12. Стосовно $M[Z]$ одержуємо:

$$M[Z]=\int_0^1 zf_Z(z)\,dz=2\frac{z^3}{3}\bigg|_0^1=\frac{2}{3}.$$

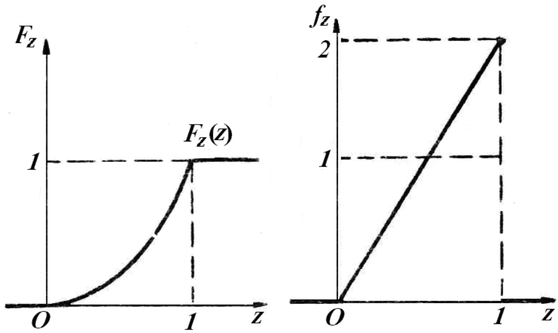


Рис. 4.10.12

II спосіб. Обчислюючи $M[Z]$ безпосередньо за формулою (4.10.16) і враховуючи, що $f_{(X,Y)}(x,y)=\frac{1}{2}$ у всіх точках квадрата G, дістаємо

$$M[Z]=\iint\limits_G(|x|+|y|)\,f_{(X,Y)}(x,y)\,dxdy=$$

$$=4\iint\limits_{G^+}(x+y)\frac{1}{2}\,dxdy=4\cdot\frac{1}{2}\int_0^1\left\{\int_0^{1-x}(x+y)dy\right\}dx=$$

$$=2\int_0^1\left(\left(xy+\frac{y^2}{2}\right)\bigg|_0^{1-x}\right)dx=2\int_0^1(x(1-x)+\frac{1}{2}(1-x)^2)dx=$$

$$=2\left(\frac{x^2}{2}-\frac{x^3}{3}-\frac{1}{2}\frac{(1-x)^3}{3}\right)\Bigg|_0^1=2\left(\frac{1}{2}-\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}\right)=\frac{2}{3},$$

де

$$G^+=\{(x,y)\,|\;x\geq 0, y\geq 0,\; x+y\leq 1\}.$$

Розглянемо ще дві числові характеристики розподілу ймовірностей на множині значень випадкової величини X, які деякою мірою характеризують положення центра розсіювання – моду й медіану.

Модою випадкової величини X з абсолютно неперервним розподілом ймовірностей на множині Ω_X її значень називається точка $Mo[X]$ максимуму щільності розподілу ймовірностей на множині Ω_X значень випадкової величини X (Рис. 4.10.13). Коли щільність розподілу ймовірностей досягає максимуму в кількох точках, тоді розподіл називається *полімодальним*, а коли тільки в одній точці, тоді *унімодальним*.

У випадку дискретного розподілу ймовірностей на множині Ω_X значень випадкової величини X модою називається така точка $Mo[X]$, у достатньо малому околі якої число $Mo[X]$ є найбільш імовірним значенням випадкової величини X.

Оскільки ймовірність будь-якого окремого значення x випадкової величини X з неперервною множиною значень дорівнює нулю, $P_X(X=x)=0$, то тут найбільш імовірного значення не існує.

Медіаною неперервного розподілу ймовірностей на множині Ω_X значень випадкової величини X називається таке її значення $Me[X]$, за якого

$$P_X((-\infty; Me[X]))=P_X((Me[X];+\infty))=\frac{1}{2}.$$

В геометричній інтерпретації медіана – це абсциса точки, в якій площа під графіком щільності розподілу ймовірностей над віссю абсцис ділиться навпіл (Рис. 4.10.14).

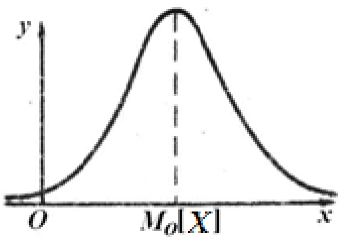


Рис. 4.10.13

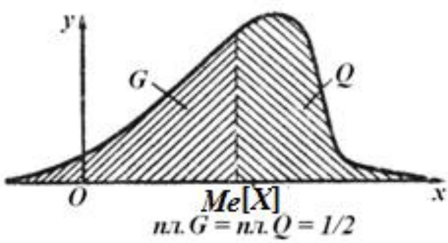


Рис. 4.10.14

Приклад 4.10.4. Знайти $Mo[X]$ (найбільш імовірне значення) випадкової величини X з біноміальним розподілом імовірностей (див.§2.10).

Розглянемо відношення

$$\frac{P_n(\mu = m+1)}{P_n(\mu = m)} = \frac{C_n^{m+1} p^{m+1} q^{n-m-1}}{C_n^m p^m q^{n-m}} = \frac{n-m}{m+1} \cdot \frac{p}{q}.$$

Якщо це відношення більше від одиниці, то $P_n(\mu = m+1) \geq P_n(\mu = m)$. Отже, ймовірності $P_n(\mu = m)$ із зростанням m зростають доти, доки виконується нерівність $\frac{P_n(\mu = m+1)}{P_n(\mu = m)} > 1$, тобто доки $np > mp + q + mq$, або, оскільки $p + q = 1$, доки $m < np - q$.

Аналогічно, ймовірності $P_n(\mu = m)$ із зростанням m спадають, коли $m > np - q$. Таким чином, коли $np - q$ не є цілим числом, тоді найбільш імовірним є значення $m = [np - q] + 1$, де через $[np - q]$ позначено цілу частину числа $np - q$. Коли $np - q$ є цілим числом, тоді найбільш імовірних значень буде два, а саме $m_1 = np - q$ і $m_2 = m_1 + 1 = np - q + 1$, оскільки коли $m = np - q$, буде

$$\frac{P_n(\mu = m+1)}{P_n(\mu = m)} = 1, \text{ тобто } P_n(\mu = m+1) = P_n(\mu = m).$$

Приклад 4.10.5. На деякий обслуговуючий пристрій у випадкові моменти часу надходять запити на обслуговування. Проміжок часу між двома сусідніми запитами є випадковою величиною τ із щільністю розподілу ймовірностей на множині її значень $f_\tau(t) = \frac{1}{T} e^{-\frac{t}{T}}$, $t \geq 0$ (так званий *експоненціальний розподіл ймовірностей*, який широко застосовується в теорії масового обслуговування). Знайти математичне сподівання, моду і медіану розглядуваного розподілу ймовірностей.

Враховуючи, що одинична ймовірність розподілена на інтервалі $[0; \infty)$, одержуємо

$$M[\tau] = \int_0^\infty t \frac{1}{T} e^{-\frac{t}{T}} dt.$$

Інтегруючи частинами:

$$u = t, \; dv = e^{-\frac{t}{T}} dt, \; du = dt, \; v = -Te^{-\frac{t}{T}},$$

дістанемо

$$\frac{1}{T} \int_0^\infty t e^{-\frac{t}{T}} dt = \frac{1}{T} \left(-tTe^{-\frac{t}{T}} \Big|_0^\infty + T \int_0^\infty e^{-\frac{t}{T}} dt \right) = -t e^{-\frac{t}{T}} \Big|_0^\infty - T e^{-\frac{t}{T}} \Big|_0^\infty.$$

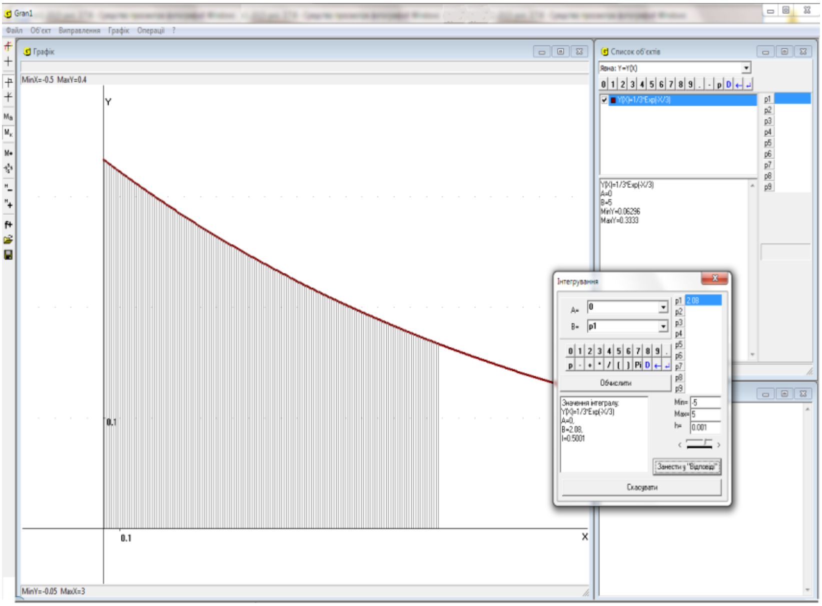


Рис. 4.10.15

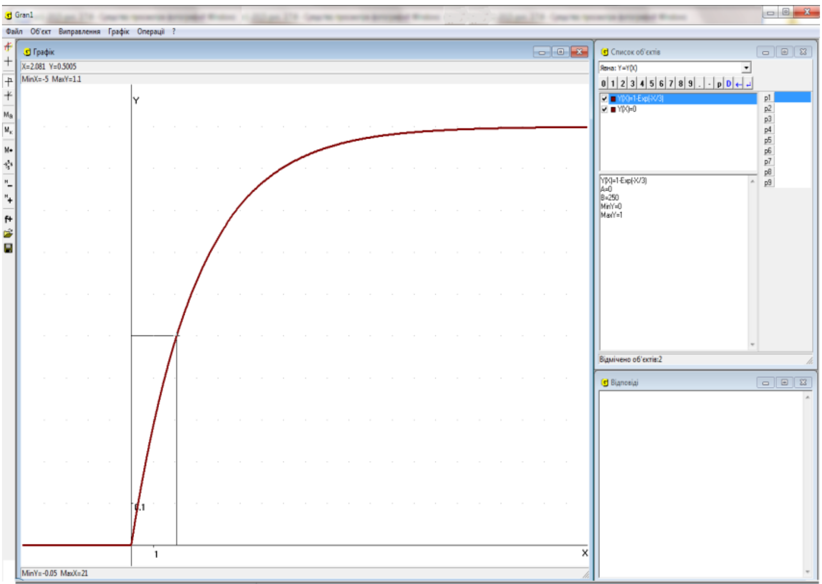


Рис. 4.10.16

Оскільки $\lim\limits_{t\to\infty}\dfrac{t}{e^{t/T}}=0$ (за правилом Лопіталя), то використовуючи формулу Ньютона-Лейбніца для обчислення розглядуваного інтеграла, остаточно одержимо $M[\tau\}=T$. Таким чином T є математичним сподіванням проміжку часу між двома сусідніми запитами.

Із графіка щільності розподілу ймовірностей (Рис. 4.10.15) випливає, що щільність $f_\tau(t)$ в точці $t=0$ набуває максимуму, який дорівнює $\dfrac{1}{T}$. Отже, $Mo[\tau]=0$.

Щоб обчислити медіану розглядуваного розподілу ймовірностей, знайдемо спочатку функцію розподілу ймовірностей $F_\tau(t)$:

$$F_\tau(t)=\int_0^t\frac{1}{T}e^{-\frac{t}{T}}dt=1-e^{-\frac{t}{T}},\ t\ge 0\,.$$

Графік функції $F_\tau(t)$ зображено на Рис. 4.10.16. Прирівнюючи $F_\tau(t)$ до $1/2$, можна знайти точку, строго зліва від якої лежить маса ймовірності, що дорівнює $1/2$. Оскільки розподіл неперервний, то в цій точці ділиться одинична ймовірність на дві рівні частини. Отже, з рівності $F_\tau(t)=1-e^{-\frac{t}{T}}=\dfrac{1}{2}$ дістаємо $e^{-\frac{t}{T}}=1/2$, $-\dfrac{t}{T}=\ln\left(\dfrac{1}{2}\right)$, $t=T\ \ln 2$.

Таким чином, $Me[\tau]=T\ln 2$.

Зазначимо, що в разі дискретного розподілу ймовірностей медіана не завжди має зміст. Так, коли випадкова величина X набуває значення 0 і 1 відповідно з імовірностями 0.1 і 0.9, тоді не існує точки, ймовірність попадання лівіше від якої дорівнює ймовірності попадання правіше від неї.

Вправи для самостійного виконання

4.10.1. Перевірити, чи правильні твердження:

1. Коли $M[X]=0$, тоді $X(E)=0$ за всіх $E\in\Omega$.
2. Коли $M[|X|]=0$, тоді $X(E)=0$.
3. Коли $M(X)\ge 0$, тоді й $X(E)\ge 0$ на Ω.
4. Коли $|X(E)|\le c$ майже всюди на Ω, тоді існує $M[X]$ і $M[X]\le c$.
5. Коли існує $M[cX]$ за деякої сталої c, тоді існує й $M[X]$.
6. Коли $M[X]\le M[Y]$, тоді $X\le Y$.
7. Коли $X(E)\le Y(E)$, тоді $M[X]\le M[Y]$.

8. Стосовно будь-якої випадкової величини X правильна нерівність $|M[X]| \leq M[|X|]$.

9. Коли існує $M[X+Y]$, тоді $M[X+Y] = M[X] + M[Y]$.

10. Коли існують скінченні $M[X \pm Y]$, тоді

$$M[aX+bY] = aM[X] + bM[Y]$$

за будь-яких сталих a і b.

11. Коли $M[XY] = M[X]M[Y]$, то випадкові величини X і Y незалежні.

12. Коли за всіх $E \in \Omega$, $X_i(E) \to X(E)$, за $i \to +\infty$, то $M[X_i] \to M[X]$.

13. За довільного $a > 0$ стосовно будь-якої інтегровної випадкової величини X правильна нерівність $P_X(E \mid -a \leq X(E) \leq a) \leq \frac{1}{a} M(|X|)$.

14. Коли X – випадкова величина з функцією розподілу ймовірностей на множині її значень $F_X(x) = P_X((-\infty; x))$, а $y = \psi(x)$ неперервна функція, тоді

$$M[\psi(X)] = \int_{\Omega} \psi(X(E))\, P(dE) = \int_{-\infty}^{\infty} \psi(x)\, dF_X .$$

15. Коли X – випадкова величина з щільністю розподілу ймовірностей на множині її значень $f_X(x)$, $x \in R$, а $y = \psi(x)$ – інтегровна функція, тоді $M[\psi(X)] = \int_{-\infty}^{\infty} \psi(x)\, f_X(x)\, dx$.

16. За кожного розподілу ймовірностей на множині значень випадкового вектора існує центр розсіювання ймовірностей.

17. Стосовно довільних випадкових векторів X та Y і борелівської функції $g(x,y)$ правильна рівність $M[g(X,Y)] = M[M[g(x,Y)]]$, де у виразі $M[g(x,Y)]$ елемент x вважається фіксованим.

18. За кожного розподілу ймовірностей на множині значень випадкової величини існує мода.

19. За кожного абсолютно неперервного розподілу ймовірностей на множині значень випадкової величини існує мода, причому єдина.

20. Якщо множина значень випадкової величини неперервна, то існує медіана розподілу ймовірностей на множині її значень.

21. В разі розподілу ймовірностей на множині значень дискретної випадкової величини може не існувати ні моди, ні медіани.

4.11. Моменти випадкових величин

Нехай (Ω, S, P) – імовірнісний простір; X – випадкова величина, визначена на Ω; $m \geq 0$ – ціле число.

Величина $M[X^m]$, в разі, коли вона визначена, називається *моментом* m*-го порядку випадкової величини* X.

Це означає, що існує також величина $M[|X|^m]$, яка називається *абсолютним моментом* m*-го порядку.* Іноді такі моменти називають *початковими моментами* m*-го порядку.*

Центральним моментом m*-го порядку* називається величина $M[(X - M[X])^m]$. *Абсолютним центральним моментом* m*-го порядку* називається величина $M[|X - M[X]|^m]$.

Нехай $\varphi(x)$ – деяка невід'ємна функція випадкової величини X. В разі, коли існує $M[\varphi(X)]$, тоді за довільного $\varepsilon > 0$, за якого $\varphi(\varepsilon) > 0$, правильна нерівність

$$P(\{E \mid \varphi(X(E)) \geq \varphi(\varepsilon)\}) \leq \frac{M[\varphi(X)]}{\varphi(\varepsilon)}.$$

Справді, оскільки

$$\varphi(X(E)) \geq \varphi(X(E)) \cdot I_{\{E|\varphi(X(E)) \geq \varphi(\varepsilon)\}}(E) \geq \varphi(\varepsilon) \cdot I_{\{E|\varphi(X(E)) \geq \varphi(\varepsilon)\}}(E),$$

то за властивостями математичного сподівання

$$M[\varphi(X)] \geq M[\varphi(X) I_{\{E|\varphi(X(E)) \geq \varphi(\varepsilon)\}}] \geq M[\varphi(\varepsilon) \cdot I_{\{E|\varphi(X(E)) \geq \varphi(\varepsilon)\}}] =$$

$$= \varphi(\varepsilon) M[I_{\{E|\varphi(X(E)) \geq \varphi(\varepsilon)\}}] = \varphi(\varepsilon) P(\{E \mid \varphi(X(E)) \geq \varphi(\varepsilon)\}).$$

Звідси

$$P(\{E \mid \varphi(X(E)) \geq \varphi(\varepsilon)\}) \leq \frac{M[\varphi(X)]}{\varphi(\varepsilon)}.$$

Цю нерівність називають *узагальненою нерівністю П.Л. Чебишова*.

Якщо X – невід'ємна випадкова величина, то в разі $\varphi(X) = X$, дістанемо нерівність П.Л. Чебишова

$$P_X([\varepsilon; \infty)) \leq \frac{M[X]}{\varepsilon}.$$

В разі $\varphi(X) = |X|^r$, $r > 0$, дістанемо нерівність

$$P(\{E \mid |X^r(E)| \geq \varepsilon\}) = P_{|X|^r}((-\infty; -\varepsilon] \cup [\varepsilon; \infty)) \leq \frac{M[|X|^r]}{\varepsilon^r},$$

яку називають *нерівністю Маркова*.

Крім математичного сподівання або моменту першого порядку важливою числовою характеристикою розподілу ймовірностей на множині значень випадкової величини X є *центральний момент другого порядку*, який називається *дисперсією* випадкової величини X і позначається $D[X]$:

$$D[X] = M[(X - M[X])^2].$$

Оскільки $(X - M[X])^2$ – невід'ємна випадкова величина, то $D[X]$ завжди визначена, коли визначене $M[X]$. Вона може набувати і значення $+\infty$.

Величина $\sigma[X] = \sqrt{D[X]}$ називається *середнім квадратичним відхиленням* або *стандартним відхиленням* випадкової величини X.

Величини $D[X]$ і $\sigma[X]$ є характеристиками того, наскільки щільно групуються ймовірності на множині значень випадкової величини X біля математичного сподівання $M[X]$ (навколо центра розсіювання), оскільки, як слідує із узагальненої нерівності *П.Л. Чебишова*:

$$P_X(|X - M[X]| \geq \varepsilon) = P_X((X - M[X])^2 \geq \varepsilon^2) \leq \frac{M[(X - M[X])^2]}{\varepsilon^2} = \frac{D[X]}{\varepsilon^2}.$$

Цю нерівність також називають *нерівністю П.Л. Чебишова.*

Необхідність введення числової характеристики розсіювання ймовірностей на множині значень випадкових величин навколо центра розсіювання (математичного сподівання) можна простежити на таких прикладах.

Розглянемо випадкові величини X і Y з розподілами ймовірностей (Табл. 4.11.1, 4.11.2).

Табл. 4.11.1

x_i	-0.01	0.01
p_i	0.5	0.5

Табл. 4.11.2

y_i	-100	100
p_i	0.5	0.5

Математичні сподівання випадкових величин X і Y дорівнюють відповідно

$$M[X] = -0.01 \cdot 0.5 + 0.01 \cdot 0.5 = 0,$$

$$M[Y] = -100 \cdot 0.5 + 100 \cdot 0.5 = 0,$$

тобто вони однакові. Проте розподіли ймовірностей на множинах значень випадкових величин X та Y суттєво відрізняються: випадкова величина X набуває значень, які мало відрізняються від її математичного сподівання (середнього значення) $M[X]$, тобто ймовірності на множині значень випадкової величини X досить скупчені біля центра розсіювання, а значення випадкової величини Y суттєво відрізняються від її математичного сподівання $M[Y]$, тобто ймовірності на множині значень випадкової величини Y досить сильно віддалені від центра розсіювання.

В механічній інтерпретації дисперсія є моментом інерції відносно центра мас системи мас із загальною одиничною масою, розподіленою вздовж осі абсцис за правилом P_X.

Приклад 4.11.1. Нехай значеннями випадкової величини X є числа, що випадають на верхній грані в результаті підкидання кубика. Знайти дисперсію випадкової величини X, якщо всі грані кубика рівноймовірні.

Розподіл ймовірностей на множині значень випадкової величини X в розглядуваному випадку набуває вигляду, поданого в Табл. 4.11.3.

Табл. 4.11.3

x_i	"1"	"2"	"3"	"4"	"5"	"6"
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$M[X]=3.5$. За означенням $D[X]$ знаходимо

$$D[X]=\sum_{i=1}^{6}(x_i-3.5)^2 p_i=\frac{1}{6}((1-3.5)^2+(2-3.5)^2+(3-3.5)^2+$$

$$+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2)=\frac{35}{12}.$$

Приклад 4.11.2. Розглянемо два розподіли ймовірностей, тобто дві випадкові величини X_1 і X_2 (Рис. 4.11.1). В одному випадку (Рис. 4.11.1, *а*)

$$D[X_1]=\sum_{i=1}^{6}(x_i-M[X])^2 p_i,$$

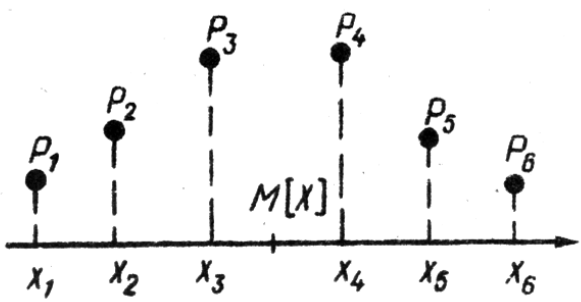


Рис. 4.11.1, *а*

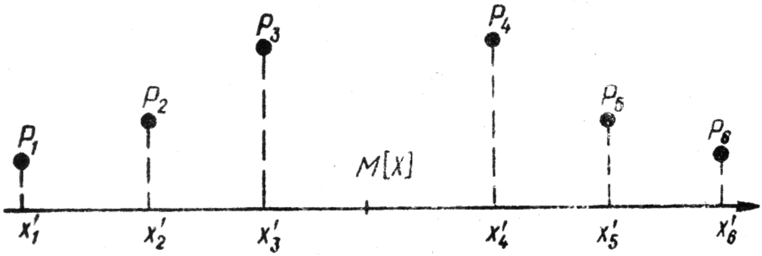


Рис. 4.11.1, *б*

а в другому (Рис. 4.11.1, *б*)

$$D[X_2]=\sum_{i=1}^{6}(x_i'-M[X])^2\,p_i\,.$$

Оскільки $|x_i'-M[X]|>|x_i-M[X]|$ за всіх $i=1,\,2,\,3,\,4,\,5,\,6$, тобто в другому випадку ймовірності $p_1,\,p_2,\,p_3,\,p_4,\,p_5,\,p_6$ більше розсіяні навколо точки $x=M[X]$, ніж у першому випадку (Рис. 4.11.1, *б*), то очевидно, що

$$D[X_2]>D[X_1].$$

Приклад 4.11.3. Знайти дисперсію рівномірного розподілу ймовірностей на множині Ω_X значень випадкової величини X – інтервалі $[a;b]$.

За означенням $D[X]$ знаходимо

$$D[X]=\int_{-\infty}^{\infty}(x-M[X])^2\,f_X(x)\,dx=\int_a^b(x-\frac{a+b}{2})^2\,\frac{1}{b-a}\,dx=$$

$$=\frac{1}{b-a}\cdot\frac{1}{3}(x-\frac{a+b}{2})^3\Big|_a^b=\frac{1}{3(b-a)}((\frac{b-a}{2})^3-(\frac{a-b}{2})^3)=\frac{(b-a)^2}{12}\,.$$

Отже, чим довший відрізок $[a;b]$, на якому рівномірно розподілена одинична ймовірність, тим більша дисперсія. Це ще одне свідчення того, що дисперсія є однією з характеристик розсіювання ймовірностей.

Розглянемо деякі *властивості початкових і центральних моментів*.

1. Коли $D[X]=0$, тоді $P_X\,(X=M[X])=1$.

Справді, в даному разі за нерівністю П.Л. Чебишова, $P_X\,(|X-M[X]|>\varepsilon)=0$ за будь-якого $\varepsilon>0$.

Звідси $P_X\,(|X-M[X]|=0)=1$.

2. $D[aX+b]=M\,[(aX+b-M\,[aX+b])^2\,]=$

$=M\,[a^2(X-M\,[X])^2\,]=a^2D[X]$, коли $M[|X|]<+\infty$.

Зокрема, $D\,[b]=0$, $D\,[aX]=a^2D\,[X]$.

3. $D\,[X]=M\;[(X-M\,[X])^2\,]=M\,[X^2-2XM\,[X]+M^2\,[X]]=$

$$= M\ [X^2] - 2M[X]M[X] + M^2[X] = M[X^2] - M^2[X].$$

Звідси

$$M[X^2] = M^2[X] + D[X].$$

Оскільки $D[X] \ge 0$, то

$$M[X^2] \ge M^2[X], \quad |M[X]| \le \sqrt{M[X^2]}, \quad M[X^2] \ge D[X].$$

4. Нехай X і Y – випадкові величини, $M[X^2] < \infty$, $M[Y^2] < \infty$. Тоді виконується *нерівність Коші-Буняковського*

$$|M[XY]| \le \sqrt{M[X^2]}\sqrt{M[Y^2]}.$$

Справді, квадратний тричлен (щодо t)

$$M[(X + tY)(X + tY)] = M[X^2] + 2tM[XY] + t^2 M[Y^2] \ge 0$$

невід'ємний, а тому його дискримінант повинен бути недодатним:

$$4M^2[XY] - 4M[X^2]M[Y^2] \le 0.$$

Звідси й випливає нерівність Коші-Буняковського.

5. З нерівності Коші-Буняковського випливає *нерівність Шварца*

$$M[(X+Y)^2] = M[X^2] + 2M[XY] + M[Y^2] \le$$

$$\le M[X^2] + 2\sqrt{M[X^2]}\sqrt{M[Y^2]} + M[Y^2] = (\sqrt{M[X^2]} + \sqrt{M[Y^2]})^2.$$

6. $M[(X - c)^2] \ge D[X] = M[(X - M[X])^2]$,

тобто $\min\limits_{c \in R} M[(X - c)^2]$ досягається, коли $c = M[X]$, за умови $|M[X]| < +\infty$.

Справді,

$$M[(X - c)^2] = M[(X - M[X] + M[X] - c)^2] =$$

$$= M[(X - M[X])^2 + 2(X - M[X])(M[X] - c) + (M[X] - c)^2] =$$

$$= D[X] + (M[X] - c)^2 > D[X], \text{ коли } M[X] \ne c.$$

7. Нехай $p > 1$, $q > 1$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, $M[|X|^p] < \infty$, $M[|Y|^p] < \infty$.

Тоді виконується *нерівність Гельдера*

$$M[|XY|] \le (M[|X|^p])^{1/p}(M[|Y|^q])^{1/q} .$$

Справді, розглянемо графік кривої $y = x^{p-1}$ або $x = y^{\frac{1}{p-1}} = y^{q-1}$ (Рис. 4.11.2).

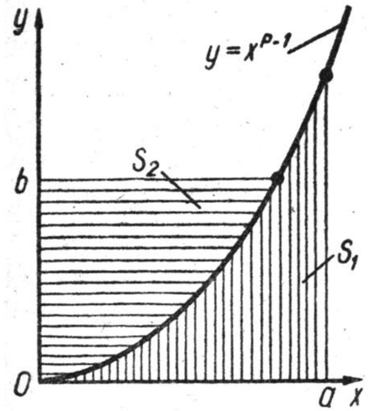


Рис. 4.11.2

Нехай $S_1 = \int\limits_0^a x^{p-1}dx = \frac{1}{p}a^p$, $S_2 = \int\limits_0^b y^{q-1}dy = \frac{1}{q}b^q$.

Тоді $S_1 + S_2 = \frac{1}{p}a^p + \frac{1}{q}b^q \ge ab$.

Поклавши $a = \frac{|X|}{(M[|X|^p])^{1/p}}$, $b = \frac{|Y|}{(M[|Y|^q])^{1/q}}$, дістанемо

$$M\left[\frac{|X|}{(M[|X|^p])^{1/p}} \cdot \frac{|Y|}{(M[|Y|^q])^{1/q}}\right] \le$$

$$\le M\left[\frac{1}{p}\frac{|X|^p}{M[|X|^p]} + \frac{1}{q}\frac{|Y|^q}{M[|Y|^q]}\right] = \frac{1}{p} + \frac{1}{q} = 1.$$

Звідси й випливає правильність твердження.

Зокрема, в разі $p = 2$, $q = 2$, дістанемо нерівність Коші-Буняковського.

8. Нехай X – випадкова величина, $\varphi(x)$ – опукла донизу функція. Причому існують скінченні $M[X]$ та $M[\varphi(X)]$. Тоді виконується *нерівність Ієнсена*

$$M[\varphi(X)] \ge \varphi(M[X]) .$$

Справді, коли $\varphi(x)$ опукла донизу функція, то це означає, що яким би не було x_0, знайдеться така лінійна функція $l(x)$, що виконуватиметься нерівність (Рис. 4.11.3)

$$\varphi(x)-\varphi(x_0)\ge l(x)-l(x_0).$$

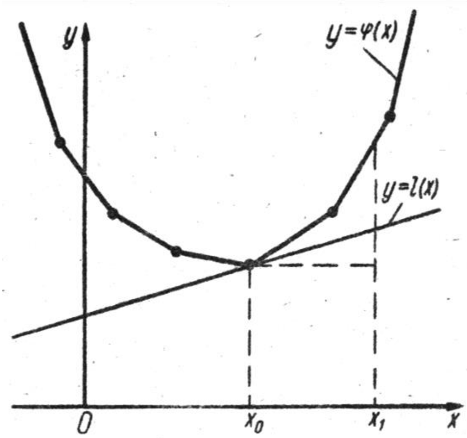


Рис. 4.11.3

Таким чином,

$$\varphi(X)-\varphi(M[X])\ge l(X)-l(M[X])$$

і

$$M[\varphi(X)-\varphi(M[X])]\ge M[l(X)-l(M[X])].$$

Оскільки математичне сподівання від лінійної функції дорівнює тій самій лінійній функції від математичного сподівання, тобто

$$M[aX+b]=aM[X]+b,$$

то

$$M[l(X)-l(M[X])]=0.$$

Отже,

$$M[\varphi(X)]\ge\varphi(M[X]).$$

Зокрема, якщо $\varphi(x)=x^2$, то $M[X^2]\ge(M[X])^2$; якщо $\varphi(x)=|x|^r$, $r\ge1$, то

$$M[|X|^r]\ge|M[X]|^r.$$

9. Якщо $1<s<t$, то виконується *нерівність Ляпунова*

$$(M[|X|^s])^{1/s}\le(M[|X|^t])^{1/t}.$$

Справді, за нерівністю Ієнсена

$$M[|X^s|^r]\ge(M[|X^s|])^r,\ r>1.$$

Нехай $r = \frac{t}{s}$. Тоді

$$M[|X^s|^{\frac{t}{s}}] \geq (M[|X^s|])^{\frac{t}{s}},$$

тобто

$$(M[|X|^t])^{\frac{1}{t}} \geq (M[|X^s|])^{\frac{1}{s}}.$$

Приклад 4.11.4. Нехай розподіл імовірностей на множині значень випадкової величини X такий, як подано в Табл. 4.11.4.

Табл. 4.11.4

x_i	-1	1
p_i	0.5	0.5

Тоді

$$M[X \cdot X] = M[X^2] = (-1)^2 \cdot 0.5 + (1)^2 \cdot 0.5 = 1,$$

а $M[X] = -1 \cdot 0.5 + 1 \cdot 0.5 = 0$, тобто $M[X^2] \neq M^2[X]$.

Приклад 4.11.5. Знайдемо математичне сподівання і дисперсію випадкового числа μ появ події A в серії з n незалежних випробувань за умови, що в кожному випробовуванні подія A відбувається з імовірністю p і не відбувається з імовірністю $q = 1 - p$. Випадкову величину $\tilde{\mu}_n$ (число появ події A в серії з n незалежних випробувань) можна подати у вигляді суми випадкових величин

$$\tilde{\mu}_n = \sum_{i=1}^{n} \mu_i,$$

де μ_i – число появ події A в i-му випробуванні. Очевидно, що випадкові величини μ_i за будь-якого i можуть набувати лише двох значень: 0, коли подія A в i-му випробуванні не відбувається, і 1, якщо подія A в i-му випробуванні відбувається. Ряд розподілу ймовірностей на множині значень випадкової величини μ_i за будь-якого i набуває вигляду (Табл. 4.11.5).

Табл. 4.11.5

0	1
q	p

Оскільки випробовування незалежні, то й величини μ_i між собою незалежні. Таким чином,

$$M[\mu_i] = 0 \cdot q + 1 \cdot p = p,$$

$$D[\mu_i] = (0 - p)^2 q + (1 - p)^2 p = p^2 q + q^2 p = pq(p + q) = pq.$$

Тому

$$M[\tilde{\mu}_n] = M[\sum_{i=1}^{n}\mu_i] = \sum_{i=1}^{n}M[\mu_i] = p + p + \cdots + p = np\ ,$$

$$D[\tilde{\mu}_n] = D[\sum_{i=1}^{n}\mu_i] = pq + pq + \cdots + pq = npq\ ,$$

$$\sigma[\tilde{\mu}_n] = \sqrt{D\,[\tilde{\mu}_n]} = \sqrt{npq}\ .$$

Враховуючи формулу Бернуллі, одержимо

$$M\ [\tilde{\mu}_n] = \sum_{k=0}^{n}kC_n^k\, p^k q^{n-k} = \sum_{k=0}^{n}kC_n^k\, p^k (1-p)^{n-k}\ .$$

Тоді

$$\sum_{k=0}^{n}kC_n^k\, p^k q^{n-k} = \sum_{k=0}^{n}kC_n^k\, p^k (1-p)^{n-k} = np\ ,\ p \in [0\,;1]\,.$$

Аналогічно

$$\sum_{k=0}^{n}(k-np)^2\, C_n^k\, p^k (1-p)^{n-k} = np(1-p)\ ,\ p \in [0\,;1]\,.$$

Коли $\varphi(x)$ неперервна функція, $(p = x,\ x \in [0;1])$, тоді

$$M\,[\varphi(\frac{\tilde{\mu}_n}{n})](x) = M\,[\varphi(P_n^*(A))](x) = \sum_{k=0}^{n}\varphi(\frac{k}{n})C_n^k\, x^k (1-x)^{n-k} = B_n(x)\ .$$

Поліноми $B_n(x)$ називають *поліномами Бернштейна.*

Приклад 4.11.6. Знайти математичне сподівання й дисперсію статистичної ймовірності $P_n^*(A)$, яку дістають за результатами серії з n незалежних випробувань, в кожному з яких подія A може відбутися з імовірністю p і не відбутися з імовірністю $q = 1 - p$.

Оскільки $P_n^*(A) = \dfrac{\tilde{\mu}_n}{n}$, де $\tilde{\mu}_n$ – випадкова величина, визначена в попередньому прикладі, то за теоремами про числові характеристики

$$M[P_n^*(A)] = \frac{1}{n}M[\tilde{\mu}_n] = \frac{1}{n}np = p\ ,$$

$$D[P_n^*(A)] = \frac{1}{n^2}D[\tilde{\mu}_n] = \frac{1}{n^2}npq = \frac{pq}{n}\ ,$$

$$\sigma[P_n^*(A)] = \sqrt{D[P_n^*(A)]} = \sqrt{\frac{pq}{n}}\ .$$

Таким чином, із збільшенням числа випробувань $D[P_n^*(A)]$ зменшується і за досить великих n може стати як завгодно малою, а це означає, що за досить великих n значення випадкової величини $P_n^*(A)$, здобуті в різних серіях з n незалежних випробувань, в більшості своїй мало між собою відрізняються і зосереджуються в досить малому околі деякого значення p, яке є математичним сподіванням випадкової величини $P_n^*(A)$ (див. властивість 1 початкових і центральних моментів).

Приклад 4.11.7. Обчислити дисперсію дискретної випадкової величини X, ймовірності на множині можливих значень якої розподілені за законом Пуассона.

Для обчислення $D[X]$ знайдемо спочатку $M[X^2]$:

$$M[X^2] = \sum_{i=0}^{\infty} i^2 \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=1}^{\infty} i^2 \frac{a^i}{i!} = e^{-a} \sum_{i=1}^{\infty} i \frac{a^i}{(i-1)!} =$$

$$= e^{-a} a \sum_{i=1}^{\infty} (i-1+1) \frac{a^{i-1}}{(i-1)!} = e^{-a} a \left(\sum_{i=1}^{\infty} (i-1) \frac{a^{i-1}}{(i-1)!} + \sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-1)!} \right) =$$

$$= e^{-a} a \left(\sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-2)!} + e^a \right) = e^{-a} a \left(a \sum_{i=2}^{\infty} \frac{a^{i-2}}{(i-2)!} + e^a \right) =$$

$$= e^{-e} a (a e^a + e^a) = a^2 + a.$$

Оскільки за даного розподілу ймовірностей

$$M[X] = \sum_{i=0}^{\infty} i \frac{a^i}{i!} e^{-a} = e^{-a} \sum_{i=1}^{\infty} i \frac{a^i}{i!} = e^{-a} \sum_{i=1}^{\infty} a \frac{a^{i-1}}{(i-1)!} = e^{-a} a \sum_{i=0}^{\infty} \frac{a^i}{i!} = e^{-a} a e^a = a,$$

то

$$D[X] = M[X^2] - M^2[X] = a^2 + a - a^2 = a.$$

Приклад 4.11.8. Знайти початкові моменти k-го порядку випадкової величини X з *показниковим розподілом* імовірностей, щільність якого

$$f_X(x) = \lambda e^{-\lambda x}, \ x \ge 0, \ \lambda > 0 \text{ – фіксоване число.}$$

Одержуємо:

$$M[X^k] = \lambda \int_0^{\infty} x^k e^{-\lambda x} dx,$$

або після заміни змінних $\lambda x = t$, $x = \dfrac{t}{\lambda}$,

$$M[X^k]=\frac{1}{\lambda^k}\int_0^\infty t^k e^{-t}dt=\frac{1}{\lambda^k}\Gamma(k+1),$$

де $\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}dt$ – так звана Г-функція (гамма-функція).

Інтегруючи частинами, дістаємо

$$\int_0^\infty t^k e^{-t}dt=-t^k e^{-t}\Big|_0^\infty+k\int_0^\infty t^{k-1}e^{-t}dt=k(k-1)\int_0^\infty t^{k-2}e^{-t}dt=\cdots$$

$$\ldots=k(k-1)(k-2)\cdots 2\cdot 1\cdot\int_0^\infty e^{-t}dt=k!$$

Таким чином за даного розподілу ймовірностей

$$M[X^k]=\frac{1}{\lambda^k}\Gamma(k+1)=\frac{k!}{\lambda^k}.$$

Зокрема $M[X]=\frac{1}{\lambda}$, $M[X^2]=\frac{2}{\lambda^2}$. Звідси

$$D[X]=M[X^2]-(M[X])^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2},\ \sigma[X]=\frac{1}{\lambda}.$$

Вправи для самостійного виконання

4.11.1 Перевірити, чи правильні твердження:

1. Стосовно кожної випадкової величини визначений її початковий момент m-го порядку за будь-якого $m\in N$.

2. Момент $M[|X|^m]$ визначений тоді і тільки тоді, коли визначений початковий момент m-го порядку.

3. Початковий момент m-го порядку – це те саме, що й абсолютний початковий момент m-го порядку.

4. Якщо існує скінченний початковий момент m-го порядку, то існує скінченний центральний момент m-го порядку.

5. Центральний момент m-го порядку не існує тоді і тільки тоді, коли не існує абсолютний центральний момент m-го порядку.

6. Центральний момент m-го порядку є скінченним тоді і тільки тоді, коли є скінченним абсолютний центральний момент m-го порядку.

7. Узагальнена нерівність Чебишова є правильною за будь-якої невід'ємної функції $\varphi(x)$.

8. Нерівність Маркова є наслідком нерівності Чебишова.

9. Нерівність Чебишова є частинним випадком нерівності Маркова.

10. Дисперсія – це центральний момент другого порядку.

11. Стандартне відхилення випадкової величини – це її середнє квадратичне відхилення.

12. $D[X]=0$ тоді і тільки тоді, коли випадкова величина є сталою.

13. Якщо $M[X^{2k}]=0$ за деякого $k \in N$, то $M[X^{m}]=0$ за будь-якого $m \in N$.

14. Якщо $M[X^2]<\infty$ і $M[Y^2]<\infty$, то й $M[XY]<\infty$.

15. Нерівність Коші-Буняковського є правильною стосовно будь-яких випадкових величин X і Y.

16. Нерівність Шварца є наслідком нерівності Коші-Буняковського.

17. Якщо дисперсія $D(X)$ – скінченна, то за будь-якого $c \in R$ скінченним є $M[(X-c)^2]$, причому $D[X]=\min\limits_{c} M[(X-c)^2]$.

18. Нерівність Гельдера є узагальненням нерівності Коші-Буняковського.

19. НерівністьІєнсена правильна за будь-якої опуклої донизу функції.

20. Нерівність Ляпунова є наслідком нерівностіІєнсена.

21. Якщо стосовно випадкової величини існує скінченний момент m-го порядку, то існує скінченний момент будь-якого k-го порядку, де $k<m$.

4.12. Деякі зв'язки стохастики та геометрії

Розглянемо випадкову величину X з рівномірним розподілом ймовірностей на множині Ω_X її значень, причому множина Ω_X – неперервна, обмежена з обмеженою мірою $0 < m(\Omega_X) < \infty$, простір подій $S_X = \mathscr{B}(R^1)$ (див. §2.6), ймовірнісна міра P_X задана через щільність $f_X(x)$ розподілу ймовірностей на Ω_X:

$$f_X(x) = \begin{cases} \dfrac{1}{m(\Omega_X)}, & \text{коли} \quad x \in \Omega_X, \\ 0, & \text{коли} \quad x \overline{\in} \Omega_X. \end{cases}$$

Нехай на множині Ω_X задано функцію $Z = \psi(X)$ від випадкового аргумента X таку, що $\Omega_z = [z_0; z_k] \subset (-\infty; \infty)$ – обмежена множина, простір подій $S_Z = \mathscr{B}(R^1)$, функція $z = \psi(x)$ S_X / S_Z-вимірна (тобто борелівська), інтегровна на Ω_X. Знайшовши розподіл ймовірностей на множині Ω_Z значень випадкової величини Z, обчислимо математичне сподівання $M[Z]$ випадкової величини Z (див. формулу (4.10.5))

$$M[Z] = \int\limits_{\Omega_Z} z P_Z(dz).$$

З іншого боку (див. формулу (4.10.6))

$$M[Z] = \int\limits_{\Omega_X} \psi(x) f_X(x) dx.$$

Оскільки $f_X(x)$ стала на Ω_X і набуває значень $\dfrac{1}{m(\Omega_X)}$, коли $x \in \Omega_X$, то

$$M[Z] = \frac{1}{m(\Omega_X)} \int\limits_{\Omega_X} \psi(x) dx,$$

і число $M[Z]$ називають *середнім значенням* підінтегральної функції $\psi(x)$ на множині Ω_X. Звідси

$$\int\limits_{\Omega_X} \psi(x) dx = m(\Omega_X) \cdot M[Z] = m(\Omega_X) \int\limits_{\Omega_Z} z P_Z(dz) \qquad (4.12.1).$$

Таким чином має місце *теорема про середнє значення* функції $\psi(X)$:

Інтеграл $\int\limits_{\Omega_X}\psi(x)dx$ *від функції* $\psi(x)$ *на множині* Ω_X *дорівнює мірі* $m(\Omega_X)$ *множини* Ω_X, *помноженій на середнє значення* $M[Z]$ *функції* $\psi(x)$, *зважене за ймовірнісною мірою* P_Z, *що породжується через* S_X / S_Z *-випадкову величину* $Z=\psi(X)$ *на множині* Ω_Z *значень випадкової величини Z за ймовірнісною мірою* P_X, *заданою на* Ω_X *через сталу щільність* $f_X(x)=\dfrac{1}{m(\Omega_X)}$.

Розглянемо приклади.

Площа трикутника.

Графік функції $y=\psi(x)$ подано на Рис. 4.12.1. Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкового аргумента (теорему про середнє), виведемо формулу для обчислення площі трикутника.

Правильні наступні рівності: $\dfrac{\overline{A}\overline{O}}{AO}=\dfrac{h-y}{h}$, $\dfrac{\overline{B}\overline{O}}{BO}=\dfrac{h-y}{h}$.

Враховуючи, що $OQ=\overline{O}\overline{B}$, $PO=\overline{A}\overline{O}$, одержуємо

$$QB=OB-OQ=OB\left(1-\left(1-\frac{y}{h}\right)\right)=OB\cdot\frac{y}{h}.$$

$$AP=AO-PO=AO\left(1-\left(1-\frac{y}{h}\right)\right)=AO\cdot\frac{y}{h}.$$

Опис функції розподілу ймовірностей на множині значень випадкової величини *Y* набуває вигляду (див. Рис. 4.12.1):

$$F_Y(y)=\begin{cases}0, & \text{коли } y\le 0,\\ \dfrac{y}{h}(AO+OB)\cdot\dfrac{1}{AO+OB}=\dfrac{y}{h}, & \text{коли } 0\le y\le h,\\ 1, & \text{коли } h\le y.\end{cases}$$

Відповідно до функції розподілу ймовірностей $F_Y(y)$ знайдемо щільність $f_Y(y)$ розподілу ймовірностей на множині значень випадкової величини *Y*:

$$f_Y(y)=\frac{d}{dy}F_Y(y)=\begin{cases}0,\ y\,\overline{\in}\,[0;h],\\ \dfrac{1}{h},\ y\in[0;h].\end{cases}$$

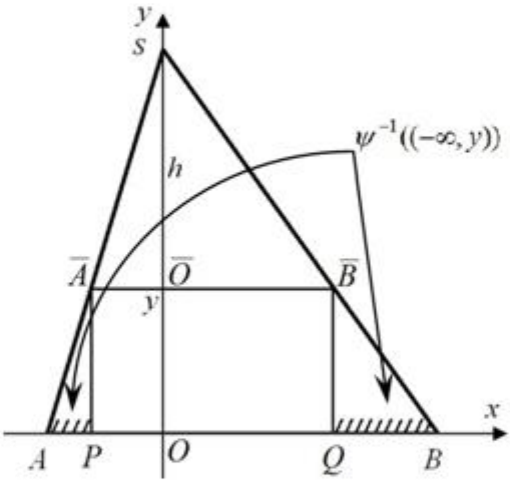


Рис. 4.12.1

Обчислимо тепер математичне сподівання випадкової величини Y

$$M[Y] = \int_0^h y f_Y(y) dy = \frac{1}{h} \cdot \frac{y^2}{2}\bigg|_0^h = \frac{h}{2}.$$

За формулою (4.10.6):

$$M[Y] = \int_{AB} \psi(x) f_X(x) dx = \frac{1}{m(AB)} \int_{AB} \psi(x) dx,$$

звідки

$$\int_{AB} \psi(x) dx = M[Y] \cdot m(AB),$$

тобто

$$\int_{AB} \psi(x) dx = \frac{h}{2} \cdot m(AO + OB) = \frac{h}{2} \cdot m(AB).$$

Таким чином, *площа трикутника* ASB дорівнює *половині добутку довжини його основи на висоту*.

Якщо трикутник такий, що його вершина проектується за межі його основи (Рис. 4.12.2), тоді досить продовжити основу трикутника так, щоб вершина трикутника проектувалася на продовжену основу, і розглянути новий трикутник $AB'C$ з так продовженою основою, отриманий з трикутника ABC продовженням його основи (Рис. 4.12.2). Тоді одержимо

$$S_{AB'C} = \frac{1}{2} |AB'| \cdot h = \frac{1}{2}(AB + BB')h =$$

$$= S_{ABC} + S_{BB'C} = \frac{1}{2} AB \cdot h + \frac{1}{2} BB' \cdot h,$$

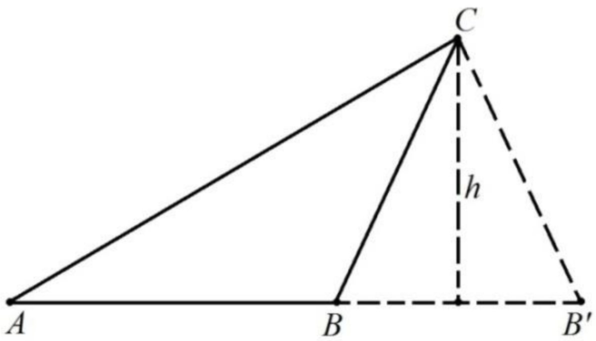


Рис. 4.12.2

де перший доданок – шукана площа, другий доданок – площа трикутника з основою BB', доповнюючого трикутник ABC з основою AB до трикутника $AB'C$ з основою AB'.

Таким чином формула для обчислення площі трикутника ABC залишається правильною, хоч безпосередньо до трикутника ABC застосувати наведені раніше міркування неможливо, оскільки залежність між змінними x і y, за допомогою якої описується множина точок на сторонах трикутника, в розглядуваному випадку виявляється неоднозначною.

Площа трапеції.

Графік функції $y=\psi(x)$ подано на Рис. 4.12.3. Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкового аргумента (теорему про середнє), виведемо формулу для обчислення площі трапеції.

З Рис. 4.12.3 видно, що $\dfrac{h-y}{h}=\dfrac{\bar{b}_2}{b_2}=\dfrac{\bar{b}_1}{b_1}$, звідки

$$\bar{b}_1=b_1\left(1-\frac{y}{h}\right), \qquad \bar{b}_2=b_2\left(1-\frac{y}{h}\right),$$

$$b_2-b_2\left(1-\frac{y}{h}\right)=b_2\frac{y}{h}, \qquad b_1-b_1\left(1-\frac{y}{h}\right)=b_1\frac{y}{h}.$$

Опис функції розподілу ймовірностей на множині значень випадкової величини Y в даному прикладі набуває вигляду (див. Рис. 4.12.3):

$$F_Y(y)=\begin{cases}0, & y\le 0,\\ \dfrac{y}{h}\cdot\dfrac{b_1+b_2}{b_1+b_2+a}, & 0\le y\le h,\\ 1, & h\le y.\end{cases}$$

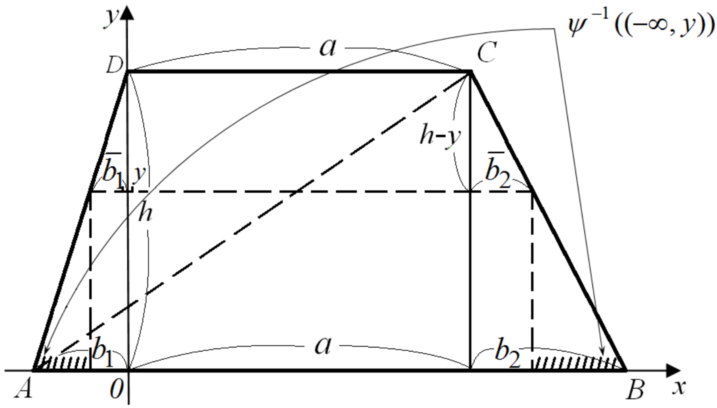


Рис. 4.12.3

Розподіл ймовірностей на множині Ω_Y значень випадкової величини Y в даному випадку мішаний: частина ймовірності $\dfrac{a}{b_1+b_2+a}$ зосереджена в точці $y=h$, інша частина $\dfrac{b_1+b_2}{b_1+b_2+a}$ розподілена на проміжку $[0;h]$ рівномірно із щільністю

$$\tilde{f}_Y(y)=\begin{cases}0, \text{ коли } y \overline{\in}[0;h],\\ \dfrac{1}{h}\cdot\dfrac{b_1+b_2}{b_1+b_2+a}, \text{ коли } y\in[0;h],\end{cases}$$

Тому

$$M[Y]=\int_0^h y\tilde{f}_Y(y)dy+h\cdot\frac{a}{b_1+b_2+a}=\left.\frac{y^2}{2}\right|_0^h\cdot\frac{1}{h}\cdot\frac{b_1+b_2}{b_1+b_2+a}+$$

$$+h\cdot\frac{a}{b_1+b_2+a}=\frac{h}{2}\cdot\frac{b_1+b_2}{b_1+b_2+a}+\frac{h}{2}\cdot\frac{2a}{b_1+b_2+a}=\frac{h}{2}\cdot\frac{(b_1+b_2+a)+a}{b_1+b_2+a}.$$

Оскільки

$$M[Y]=\int_{AB}\psi(x)f_X(x)dx=\int_{AB}\psi(x)\frac{1}{b_1+b_2+a}dx=\int_{AB}\psi(x)\frac{1}{m(AB)}dx,$$

то

$$\int_{AB}\psi(x)dx=M[Y]\cdot m(AB)=\frac{h}{2}\cdot\frac{(b_1+b_2+a)+a}{b_1+b_2+a}\cdot(b_1+b_2+a)=$$

$$=\frac{h}{2}\cdot((b_1+b_2+a)+a).$$

Таким чином, *площа трапеції* дорівнює *половині добутку суми довжин основ трапеції на висоту*.

Площа еліпса.

Задано функцію $y = b\sqrt{1-\dfrac{x^2}{a^2}}$, $x \in [-a\,;\,a]$ (рівняння верхньої половини еліпса, Рис. 4.12.4).

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкового аргумента (теорему про середнє), вивести формулу для обчислення площі еліпса з півосями a і b.

Вважаючи, що

$$f_X(x) = \begin{cases} \dfrac{1}{2a},\ x \in [-a\,;\,a], \\ 0,\ x \overline{\in} [-a\,;\,a], \end{cases}$$

одержимо (див. Рис. 4.12.4):

$$F_Y(y) = \begin{cases} 0,\ y \le 0, \\ 1 - \sqrt{1-\dfrac{y^2}{b^2}}, \quad y \in [0;\,b], \end{cases} \qquad f_Y(y) = \begin{cases} 0,\ y \overline{\in} [0\,;\,b], \\ \dfrac{\dfrac{y}{b^2}}{\sqrt{1-\dfrac{y^2}{b^2}}},\ y \in [0\,;\,b]. \end{cases}$$

Тому $M[Y] = \displaystyle\int_0^b y \cdot \dfrac{\dfrac{y}{b^2}}{\sqrt{1-\dfrac{y^2}{b^2}}}\,dy = \int_0^b \dfrac{\left(\dfrac{y}{b}\right)^2}{\sqrt{1-\left(\dfrac{y}{b}\right)^2}}\,dy.$

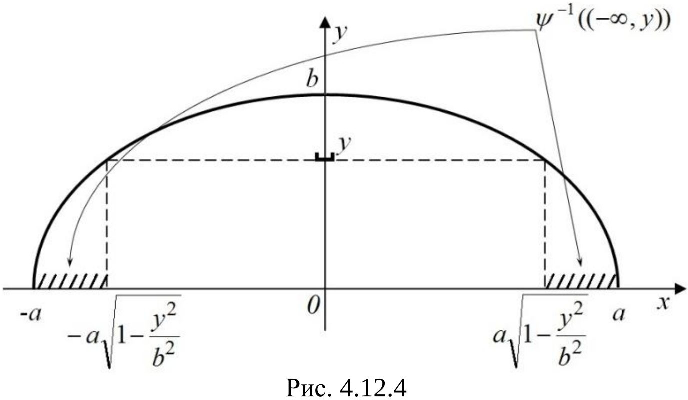


Рис. 4.12.4

Інтегруючи частинами: $y=u$, $dy=du$, $\dfrac{\frac{y}{b^2}dy}{\sqrt{1-\left(\frac{y}{b}\right)^2}}=dv$,

$v=-\sqrt{1-\left(\frac{y}{b}\right)^2}$, одержимо

$$M[Y]=-y\sqrt{1-\left(\frac{y}{b}\right)^2}\Bigg|_0^b+\int_0^b\sqrt{1-\left(\frac{y}{b}\right)^2}dy=$$

$$=\int_0^b\frac{1}{\sqrt{1-\left(\frac{y}{b}\right)^2}}dy-\int_0^b\frac{\left(\frac{y}{b}\right)^2}{\sqrt{1-\left(\frac{y}{b}\right)^2}}dy.$$

Звідки $2\int_0^b\frac{\left(\frac{y}{b}\right)^2}{\sqrt{1-\left(\frac{y}{b}\right)^2}}dy=b\int_0^b\frac{d\left(\frac{y}{b}\right)}{\sqrt{1-\left(\frac{y}{b}\right)^2}}=b\arcsin\left(\frac{y}{b}\right)\Bigg|_0^b=b\frac{\pi}{2}$.

Отже,

$$M[Y]=\int_0^b\frac{\left(\frac{y}{b}\right)^2}{\sqrt{1-\left(\frac{y}{b}\right)^2}}dy=b\frac{\pi}{4},$$

Тому (за формулою 4.12.1)

$$\int_{-a}^{a}\psi(x)dx=\int_{-a}^{a}b\sqrt{1-\left(\frac{x}{a}\right)^2}dx=M[Y]\cdot m((-a;a))=b\cdot\frac{\pi}{4}\cdot 2a=\frac{\pi ab}{2}.$$

Таким чином, площа еліпса з півосями a та b обчислюється за формулою

$$S_{еліпса}=\pi ab.$$

Якщо $a=b$, тоді еліпс стає колом радіуса a, яким обмежується круг площею

$$S_{круга}=\pi a^2.$$

Об'єм конуса.

Нехай задано функцію $z=\psi(x,y)$ на множині G, $(x,y)\in G$, причому в графічному поданні поверхнею, стосовно точок (x,y,z) якої задовільняється рівність $z=\psi(x,y)$, є поверхня конуса з основою G і висотою h.

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкових аргументів (теорему про середнє), виведемо формулу, за якою здійснюється обчислення об'єму конуса.

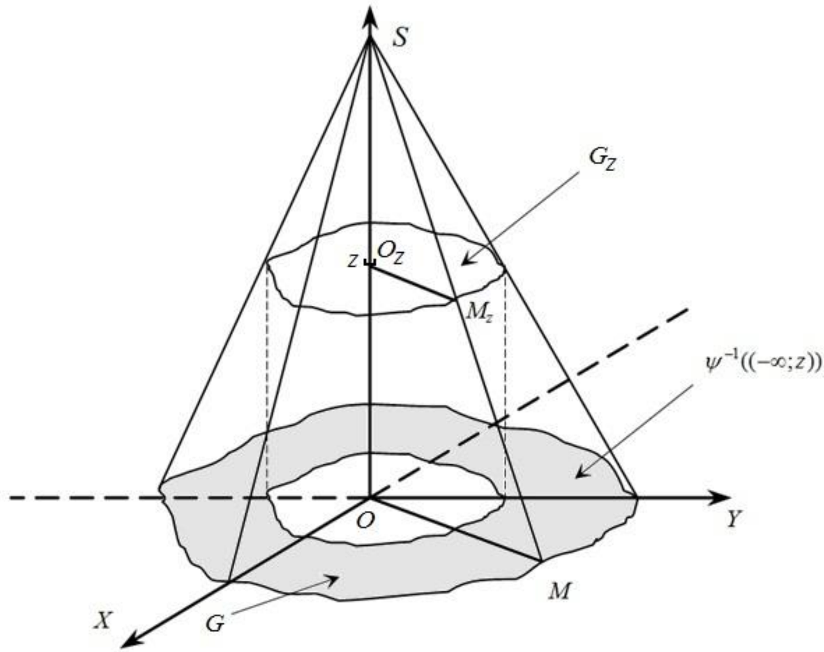


Рис. 4.12.5

Нехай потрібно обчислити об'єм конуса з висотою h і основою G, міра (площа) якої $m(G)$ відома, а бічна поверхня конуса описується за рівнянням $z=\psi(x,y)$, $(x,y)\in G$, $z\in[0;h]$. Нехай G_z – частина площини всередині конуса, паралельної до основи конуса, яка перетинає конус на висоті z, $0\le z\le h$ (Рис. 4.12.5). Із подібності трикутників OSM і O_zSM_z (Рис. 4.12.5) слідує

$$m(G_z)=\left(\frac{h-z}{h}\right)^2 m(G)=\left(1-\frac{z}{h}\right)^2 m(G),\ z\in[0;h].$$

Будемо вважати, що на множині G задано рівномірний розподіл ймовірностей із щільністю

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{m(G)}, & \text{коли } (x,y)\in G,\\ 0, & \text{коли } (x,y)\overline{\in} G\end{cases}$$

і розглянемо випадкову величину $Z=\psi(X,Y)$, де (X,Y) випадковий вектор із множиною значень G.

Тоді опис функції розподілу ймовірностей на множині значень випадкової величини Z

$$F_Z(z)=P_Z((-\infty;z))=P_{(X,Y)}(\psi^{-1}((-\infty;z)))$$

в розглядуваному випадку набуде вигляду (див. Рис. 4.12.5):

$$F_Z(z)=\begin{cases}0, & \text{коли } z\le 0,\\ \dfrac{1}{m(G)}(m(G)-m(G_Z))=1-\left(1-\dfrac{z}{h}\right)^2, & \text{коли } z\in[0;h];\\ 1, & \text{коли } z\ge h,\end{cases}$$

звідки

$$f_Z(z)=\frac{d}{dz}F_Z(z)=\begin{cases}0, & \text{коли } z\overline{\in}[0;h],\\ \dfrac{2}{h}\left(1-\dfrac{z}{h}\right), & \text{коли } z\in[0;h].\end{cases}$$

Обчислюючи тепер $M[Z]$, одержимо

$$M[Z]=\int_0^h zf_Z(z)dz=\int_0^h z\cdot\frac{2}{h}\left(1-\frac{z}{h}\right)dz=\frac{2}{h}\left(\frac{z^2}{2}-\frac{z^3}{3h}\right)\Bigg|_0^h=\frac{h}{3}.$$

Враховуючи, що

$$M[Z]=\int_{\Omega_Z} zP_Z(dz)=\int_{\Omega_Z} zf_Z(z)dz=\iint_{\Omega}\psi(x,y)P_{(X,Y)}(\psi^{-1}(dz))=$$
$$=\iint_{\Omega}\psi(x,y)\,f_{(X,Y)}(x,y)dxdy,$$

де $\Omega_Z=Z(\Omega)$ – образ множини Ω за відображенням $Z=\psi(x,y)$, $(x,y)\in\Omega$, а також, що в розглядуваному випадку $\Omega=G$, $\Omega_Z=\{z\mid z\in[0;h]\}$,

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{m(G)}, & \text{коли } (x,y)\in G,\\ 0, & \text{коли } (x,y)\overline{\in} G,\end{cases}$$

одержимо

$$M[Z]=\int_0^h zf_Z(z)dz=\iint_G \psi(x,y)\frac{1}{m(G)}dxdy,$$

звідки $\iint_G \psi(x,y)dxdy=m(G)\cdot M[Z]=m(G)\cdot\frac{h}{3}$.

Отже, об’єм конуса дорівнює третині добутку площі основи на висоту.

Зокрема, якщо в основі конуса круг радіуса R, тобто $G=\{(x,y)\,|\; x^2+y^2\le R^2\}$, тоді об’єм такого конуса дорівнюватиме

$$V=\iint_G \psi(x,y)dxdy=\frac{1}{3}\pi R^2 h.$$

Якщо G – круг радіуса R з центром в початку координат, $G=\{(x,y)\,|\; x^2+y^2\le R^2\}$, а конус прямий (вершина конуса проектується в центр круга), тоді функція $\psi(x,y)$ набуватиме вигляду

$$\psi(x,y)=h(1-\frac{1}{R}\sqrt{x^2+y^2}),\ (x,y)\in G$$

і за попереднім

$$\iint_G h(1-\frac{1}{R}\sqrt{x^2+y^2})dxdy=\frac{1}{3}\pi R^2 h.$$

Якщо вершина конуса проектується за межі його основи G (Рис. 4.12.6), тоді попередні міркування незастосовні, оскільки залежність між змінними x і y та z виявляється неоднозначною – одній і тій самій точці на площині xOy на поверхні конуса може ставитись у відповідність не одна точка. В такому випадку досить розглянути область G_1 таку, що $G\subset G_1$, і крім того вершина конуса проектується в область G_1 (Рис. 4.12.6). Тоді отримаємо

$$V_{конуса}=\frac{1}{3}hm(G_1)=\frac{1}{3}hm(G+(G_1\setminus G))=\frac{1}{3}hm(G)+\frac{1}{3}hm(G_1\setminus G),$$

де перший доданок – шуканий об’єм, другий доданок – об’єм конуса, доповнюючого конус з основою G до конуса з основою G_1.

Таким чином, формула для обчислення об’єму конуса залишається правильною і тоді, коли вершина конуса проектується за межі його основи.

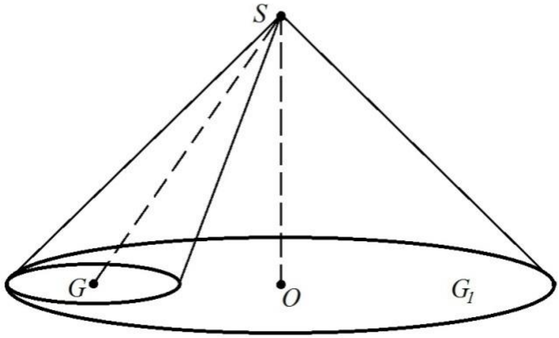


Рис. 4.12.6

Об'єм зрізаного конуса.

На множині $G \subset R^2$ задано функцію $z = \psi(x, y)$, $(x, y) \in G$. У графічному поданні поверхнею, стосовно точок (x, y, z) якої задовільняється рівність $z = \psi(x, y)$, є поверхня зрізаного конуса з нижньою основою G, верхньою основою G_1 і висотою h.

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкових аргументів (теорему про середнє), виведемо формулу для обчислення об'єму зрізаного конуса.

Нехай конус з основою G і вершиною S зрізано на висоті h (Рис. 4.12.7), x – висота конуса, відрізаного від конуса з висотою h+x з тією самою основою, що і у зрізаного конуса.

Тоді за попереднім (див. Рис. 4.12.7):

$$m(G_z) = \left(\frac{h + x - z}{h + x}\right)^2 m(G) = \left(1 - \frac{z}{h + x}\right)^2 m(G).$$

Будемо вважати, що на множині G задано рівномірний розподіл ймовірностей із щільністю

$$f_{(X,Y)} = \begin{cases} \dfrac{1}{m(G)}, & \text{коли } (x, y) \in G, \\ 0, & \text{коли } (x, y) \overline{\in} G. \end{cases}$$

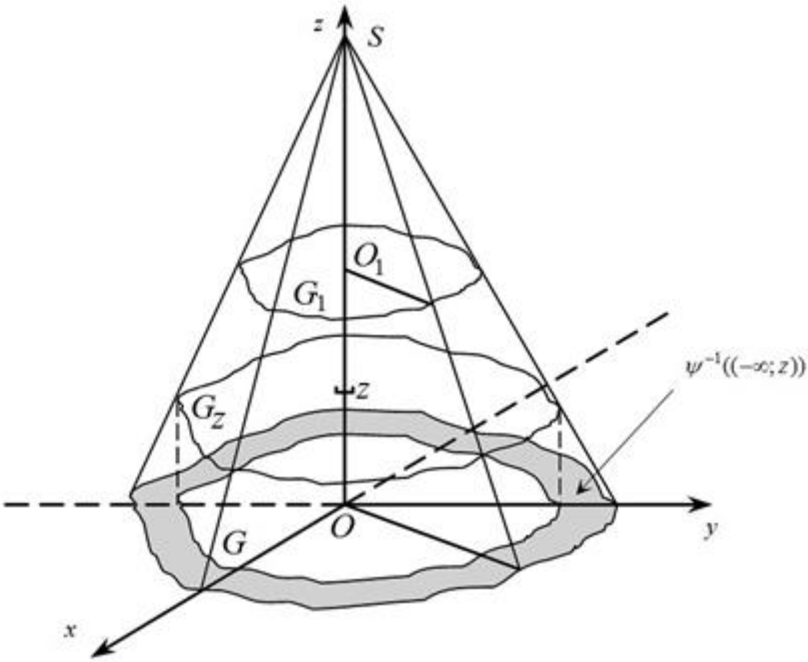


Рис. 4.12.7

Тоді опис функції розподілу ймовірностей на множині значень випадкової величини Z в розглядуваному випадку набуває вигляду

$$F_Z(z) = \begin{cases} 0, \text{ коли } z \le 0, \\ 1 - \left(1 - \dfrac{z}{h+x}\right)^2, \text{ коли } z \in [0;\, h), \\ 1, \text{ коли } z > h. \end{cases}$$

В розглядуваному випадку розподіл ймовірностей на множині Ω_Z значень випадкової величини Z виявляється мішаним, оскільки функція $F_z(z)$ змінюється неперервно, коли z змінюється від 0 до h, а в точці $z = h$ набуває відмінного від нуля приросту, рівного імовірності попадання в точку $z = h$:

$$1 - \left(1 - \left(1 - \frac{h}{h+x}\right)^2\right) = \left(1 - \frac{h}{h+x}\right)^2 .$$

Щільність розподілу на множині $[0;h]$ неперервно розподіленої частини одиничної ймовірності дорівнює

$$\widetilde{f}_Z(z) = \begin{cases} 0, \text{ коли } z \overline{\in} [0;\, h], \\ \dfrac{2}{h+x}\left(1 - \dfrac{z}{h+x}\right), \text{ коли } z \in [0;\, h]. \end{cases}$$

Центр розсіювання за такого розподілу ймовірностей на множині значень випадкової величини Z (математичне сподівання $M[Z]$) обчислюється за формулою

$$M[Z] = \int_0^h z\widetilde{f}_Z(z)dz + h\cdot\left(1 - \frac{h}{h+x}\right)^2 = \int_0^h z\frac{2}{h+x}\left(1-\frac{z}{h+x}\right)dz +$$

$$+h\left(1-\frac{h}{h+x}\right)^2 = \frac{2}{h+x}\left(\frac{z^2}{2} - \frac{1}{h+x}\cdot\frac{z^3}{3}\right)\Bigg|_0^h + h\frac{x^2}{(h+x)^2} =$$

$$= \frac{2}{h+x}\left(\frac{3z^2(h+x) - 2z^3}{2\cdot 3\cdot (h+x)}\right)\Bigg|_0^h + h\frac{x^2}{(h+x)^2} = \frac{1}{3}\cdot\frac{(3h^2(h+x) - 2h^3)}{(h+x)^2} +$$

$$+h\frac{x^2}{(h+x)^2} = \frac{1}{3}\cdot\frac{3h^3 + 3h^2x - 2h^3 + 3hx^2}{(h+x)^2} = \frac{1}{3}\cdot\frac{h^3 + 3h^2x + 3hx^2}{(h+x)^2}.$$

Враховуючи формулу 4.12.1, стосовно об'єму зрізаного конуса з висотою *h*, відрізаного від конуса з висотою (*h*+*x*) і з основою *G*, одержуємо:

$$V = \iint_G \psi(x,y)dxdy = M(Z)m(G) = \frac{1}{3}\cdot\frac{h^3 + 3h^2x + 3hx^2}{(h+x)^2}m(G).$$

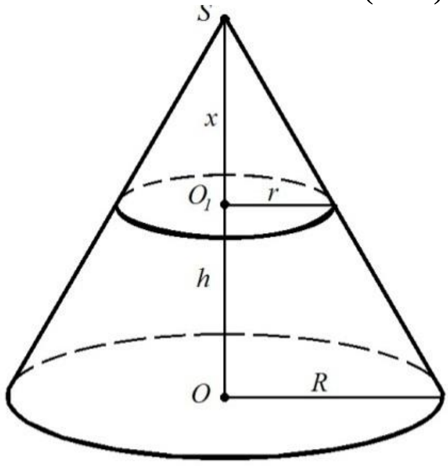


Рис. 4.12.8

Зокрема, якщо зрізаний конус отриманий відтинанням частини прямого кругового конуса так, що основами зрізаного конуса є круги радіусів R і r (Рис. 4.12.8), тоді $\dfrac{h+x}{R}=\dfrac{x}{r}$, звідки $x=\dfrac{hr}{R-r}$.

Тоді

$$V=\frac{1}{3}\left(\frac{h^3+3h^2x+3hx^2}{(h+x)^2}\right)m(G)=\frac{1}{3}hm(G)\left(\frac{h^2+3hx+3x^2}{(h+x)^2}\right)=$$

$$=\frac{1}{3}hm(G)\frac{h^2+3h\dfrac{hr}{R-r}+3\dfrac{h^2r^2}{(R-r)^2}}{\left(h+\dfrac{hr}{R-r}\right)^2}=$$

$$=\frac{1}{3}hm(G)\frac{h^2+3\dfrac{h^2r}{R-r}+3\dfrac{h^2r^2}{(R-r)^2}}{\left(\dfrac{hR-hr+hr}{R-r}\right)^2}=$$

$$=\frac{1}{3}hm(G)\frac{h^2+3\dfrac{h^2r}{R-r}+3\dfrac{h^2r^2}{(R-r)^2}}{\dfrac{h^2R^2}{(R-r)^2}}=$$

$$=\frac{1}{3}hm(G)\frac{h^2(R-r)^2+3h^2r(R-r)+3h^2r^2}{h^2R^2}=\frac{1}{3}hm(G)\frac{1}{R^2}(R^2-$$

$$-2rR+r^2+3rR-3r^2+3r^2)=\frac{1}{3}hm(G)\frac{1}{R^2}(R^2+rR+r^2)=$$

$$=\frac{1}{3}h\pi R^2\cdot\frac{1}{R^2}(R^2+rR+r^2)=\frac{1}{3}h\pi(R^2+rR+r^2).$$

Об’єм піраміди.

Задано функцію $z=\psi(x,y)$ на множині G, $(x,y)\in G$, де G многокутник з довільною скінченною кількістю сторін, не обов’язково опуклий.

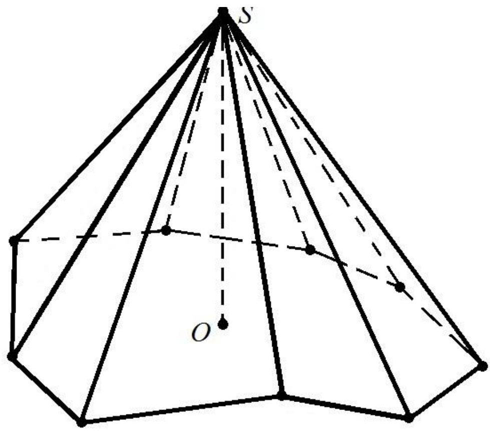


Рис. 4.12.9

В графічному поданні поверхнею, через точки (x, y, z) якої задовільняється рівність $z = \psi(x, y)$, є бічна поверхня піраміди з основою G і висотою h.

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкових аргументів (теорему про середнє), виведемо формулу для обчислення об'єму піраміди.

Оскільки піраміду можна розглядати як конус, твірна якого проходить через вершину *S*, а напрямною лінією є ламана, якою обмежується многокутник, який лежить в основі піраміди (Рис. 4.12.9), то всі міркування стосовно конуса залишаються застосовними і стосовно піраміди без будь яких змін.

Тому $V_{піраміди} = \frac{1}{3} S_{основи} \cdot h$.

Об'єм параболоїда обертання.

Задано функцію $z = h\left(1 - \left(\frac{x^2}{r^2} + \frac{y^2}{r^2}\right)\right)$, $x^2 + y^2 \le r^2$.

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкового аргумента (теорему про середнє), вивести формулу для обчислення об'єму параболоїда обертання.

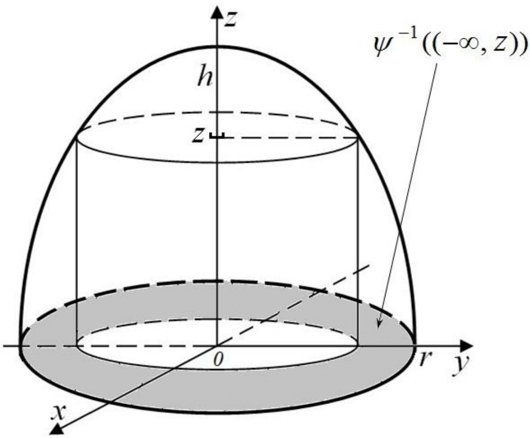


Рис. 4.12.10

Враховуючи, що

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{\pi r^2}, \text{ коли } (x,y)\in G,\\ 0, \quad \text{коли } (x,y)\overline{\in} G\end{cases}$$

а також, що коли $x=0$, тоді

$$z=h\left(1-\frac{y^2}{r^2}\right), \text{ звідки } \frac{y^2}{r^2}=1-\frac{z}{h},\ y=\pm r\sqrt{1-\frac{z}{h}},$$

одержимо (див. Рис. 4.12.10)

$$F_Z(z)=\begin{cases}0, \text{ коли } z\le 0,\\ \dfrac{1}{\pi r^2}\left(\pi r^2-\pi r^2\left(1-\dfrac{z}{h}\right)\right)=\dfrac{z}{h}, \text{ коли } z\in[0;h],\\ 1, \text{ коли } z>h.\end{cases}$$

$$f_Z(z)=\frac{d}{dz}F_Z(z)=\begin{cases}0, \text{ коли } z\overline{\in}[0;h],\\ \dfrac{1}{h} \text{ коли } z\in[0;h].\end{cases}$$

$$M[Y]=\int_0^h z\cdot\frac{1}{h}dz=\frac{h}{2}.$$

Враховуючи формулу 4.12.1, стосовно об'єму V вказаного параболоїда одержуємо

$$V=\iint_G h\left(1-\left(\frac{x^2}{r^2}+\frac{y^2}{r^2}\right)\right)dxdy=\frac{h}{2}\cdot\pi r^2=\frac{\pi r^2 h}{2}.$$

Об'єм еліпсоїда.

Задано функцію $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$, $x\in[-a;a]$, $y\in[-b;b]$, $z\in[-c;c]$.

Використовуючи зв'язок між 1-м та 2-м способами обчислення математичного сподівання функції випадкового аргумента (теорему про середнє), виведемо формулу для обчислення об'єму еліпсоїда.

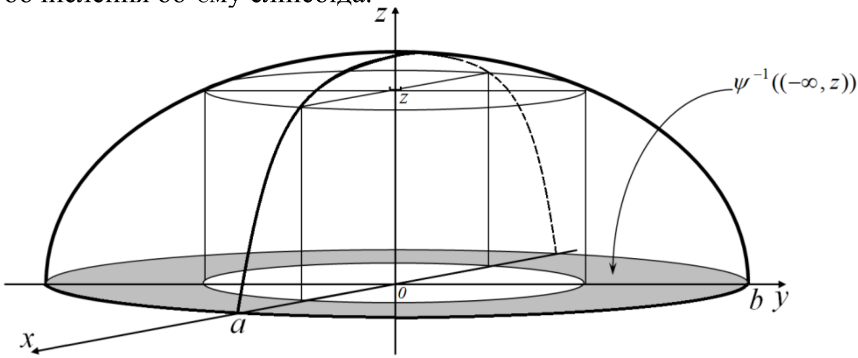


Рис. 4.12.11

Враховуючи, що рівнянням верхньої половини поверхні еліпсоїда (Рис. 4.12.11) є вираз $z=c\sqrt{1-\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)}$, а також, що за $x=0$ буде $y=\pm b\sqrt{1-\frac{z^2}{c^2}}$; за $y=0$ буде $x=\pm a\sqrt{1-\frac{z^2}{c^2}}$, і що в основі розглядуваної фігури лежить еліпс з півосями a та b, а в перетині еліпсоїда площиною, перпендикулярною до осі Oz і віддаленою від початку координат на віддаль z, також отримується еліпс з півосями $a\sqrt{1-\frac{z^2}{c^2}}$ та $b\sqrt{1-\frac{z^2}{c^2}}$, одержуємо (див. Рис. 4.12.11)

$$F_Z(z)=\begin{cases}0, & z\le 0,\\ \frac{1}{\pi ab}\left(\pi ab-\pi ab\sqrt{1-\frac{z^2}{c^2}}\sqrt{1-\frac{z^2}{c^2}}\right)=1-\left(1-\frac{z^2}{c^2}\right)=\frac{z^2}{c^2}, & z\in[0;c],\\ 1, & z\ge c,\end{cases}$$

$$f_Z(z)=\begin{cases}0, z\overline{\in}[0;c],\\ \dfrac{2z}{c^2}, z\in[0;c],\end{cases}$$

$$M[Z]=2\int_0^c\frac{z^2}{c^2}dz=2\frac{z^3}{3c^2}\Bigg|_0^c=\frac{2}{3}c\,.$$

Отже, $\iint_G \psi(x,y)dxdy = M[Z]\cdot m(G)=\dfrac{2}{3}c\cdot\pi ab$ – об'єм половини еліпсоїда, звідки

$$V_{еліпсоїда}=\frac{4}{3}\pi abc\,.$$

Якщо $a=b=c=r$, тоді еліпсоїд стає кулею, об'єм якої

$$V_{кулі}=\frac{4}{3}\pi r^3\,.$$

Об'єм тора.

Круг радіуса r переміщується так, що його центр переміщується вздовж кола радіуса *R*, $(R>r)$, в площині, перпендикулярній до площини круга (Рис. 4.12.13). Знайти об'єм тіла, яке обмежується поверхнею, що описується колом, яким обмежується переміщуваний вказаним чином круг (об'єм тора) (Рис. 4.12.12).

Нехай круг радіуса *r* розміщено в площині *yOz* так, що його центр лежить на осі *Oy* і віддалений від початку координат на віддаль *R* (Рис. 4.12.13).

Нехай круг переміщується так, що його центр переміщується в площині *xOy* вздовж кола радіуса *R* з центром в початку координат. Тоді через точки горизонтального діаметра круга описується в площині *xOy* кільце з внутрішнім радіусом $R-r$ і зовнішнім радіусом $R+r$, (Рис. 4.12.13), а через точки кола, що обмежує вказаний круг, описується поверхня, яка називається тором (Рис. 4.12.12).

Нехай на множині *G* точок цього кільця задано функцію $Z=\psi(x,y)$ таку, що в графічному поданні поверхнею, через координати точок на якій задовільняється рівність $z=\psi(x,y)$, є частина поверхні тора, яка лежить над площиною *xOy*.

Легко бачити, що $m(G)=\pi(R+r)^2-\pi(R-r)^2=4\pi Rr$. Нехай на множині *G* задано рівномірний розподіл ймовірностей із щільністю

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{m(G)}, & \text{коли } (x,y)\in G,\\ 0, & \text{коли } (x,y)\overline{\in} G.\end{cases}$$

і задано випадкову величину $Z=\psi(X,Y)$, де (X,Y) – випадкова точка із множини G.

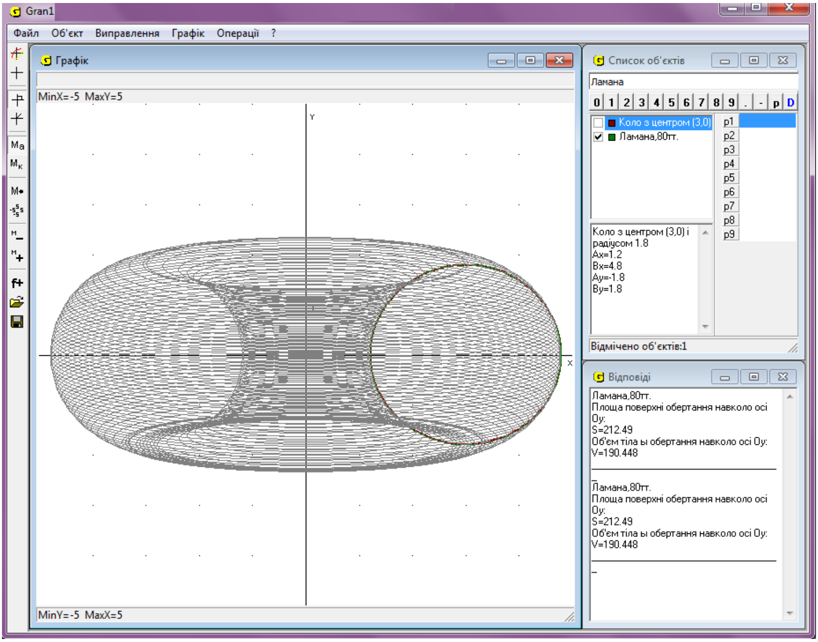


Рис. 4.12.12

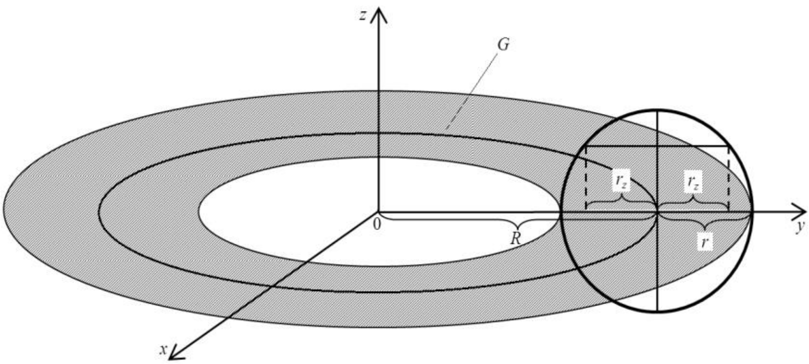


Рис. 4.12.13

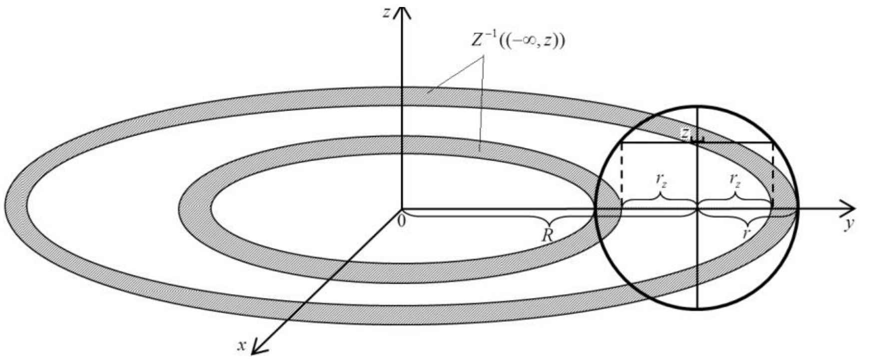


Рис. 4.12.14

Щоб обчислити математичне сподівання $M[Z]$ випадкової величини Z, знайдемо спочатку $F_Z(z)$:

$$F_Z(z) = P_Z((-\infty; z)) = P_{(X,Y)}(Z^{-1}((\infty; z))) =$$

$$= \begin{cases} 0, & \text{коли } z \le 0, \\ \dfrac{m(Z^{-1}((-\infty; z)))}{m(G)}, & \text{коли } 0 \le z \le r, \\ 1, & \text{коли } r \le z. \end{cases}$$

Очевидно в разі $0 \le z \le r$ (Рис. 4.12.14)

$$m(Z^{-1}((-\infty, z))) = \pi(R+r)^2 - \pi(R+r_z)^2 + \pi(R-r_z)^2 - \pi(R-r)^2,$$

де $r_Z = \sqrt{r^2 - z^2}$, тобто

$$m(Z^{-1}((-\infty; z))) = 4\pi Rr - 4\pi Rr_Z \ ,$$

звідки

$$F_Z(z) = \begin{cases} 0, & \text{коли } z \le 0, \\ 1 - \dfrac{r_Z}{r}, & \text{коли } 0 \le z \le r, \\ 1, & \text{коли } r \le z, \end{cases}$$

тобто

$$F_Z(z) = \begin{cases} 0, & \text{коли } z \le 0, \\ 1 - \sqrt{1 - \dfrac{z^2}{r^2}}, & \text{коли } 0 \le z \le r, \\ 1, & \text{коли } r \le z. \end{cases}$$

Звідси

$$f_Z(z)=\begin{cases}0, \text{ коли } z\,\overline{\in}[0;r],\\ \dfrac{\frac{1}{2}\cdot 2z}{r^2\sqrt{1-\dfrac{z^2}{r^2}}}=\dfrac{\frac{z}{r}}{r\sqrt{1-\left(\dfrac{z}{r}\right)^2}}, \text{ коли } z\in[0;r].\end{cases}$$

Тоді

$$M[Z]=\int_0^r z\frac{\frac{z}{r}}{r\sqrt{1-\left(\dfrac{z}{r}\right)^2}}dz=\int_0^r z\frac{\left(\dfrac{z}{r}\right)}{\sqrt{1-\left(\dfrac{z}{r}\right)^2}}d\left(\frac{z}{r}\right).$$

Інтегруючи частинами: $u=z$, $du=dz$, $dv=\dfrac{\left(\dfrac{z}{r}\right)d\left(\dfrac{z}{r}\right)}{\sqrt{1-\left(\dfrac{z}{r}\right)^2}}$,

$v=-\sqrt{1-\left(\dfrac{z}{r}\right)^2}$, одержимо

$$M[Z]=-z\sqrt{1-\left(\frac{z}{r}\right)^2}\Bigg|_0^r+\int_0^r\sqrt{1-\left(\frac{z}{r}\right)^2}dz=$$

$$=\int_0^r\frac{1-\dfrac{z^2}{r^2}}{\sqrt{1-\dfrac{z^2}{r^2}}}dz=r\int_0^r\frac{d\left(\dfrac{z}{r}\right)}{\sqrt{1-\dfrac{z^2}{r^2}}}-\int_0^r\frac{\dfrac{z^2}{r^2}}{\sqrt{1-\dfrac{z^2}{r^2}}}dz=$$

$$=r\arcsin\left(\frac{z}{r}\right)\Bigg|_0^r-\int_0^r\frac{\dfrac{z^2}{r^2}}{\sqrt{1-\dfrac{z^2}{r^2}}}dz=\frac{\pi r}{2}-\int_0^r\frac{\dfrac{z^2}{r^2}}{\sqrt{1-\dfrac{z^2}{r^2}}}dz\,.$$

Звідси $M[Z]=\int\limits_0^r \frac{\frac{z^2}{r^2}}{\sqrt{1-\frac{z^2}{r^2}}}dz=\frac{\pi r}{4}$.

Таким чином об’єм тіла під поверхнею тора і над площиною *xOy* (об’єм верхньої половини тора) дорівнює

$$V_1=\iint\limits_G \psi(x,y)dxdy=M[Z]\cdot m(G)=\frac{\pi r}{4}\cdot 4\pi Rr=\pi R\cdot \pi r^2.$$

Звідси стосовно повного об’єму тора одержуємо:

$$V_{тора}=2V_1=2\pi R\cdot \pi r^2.$$

Виведемо формули для обчислення об’єму зовнішньої та внутрішньої відносно циліндра $x^2+y^2=R^2$ частин тора.

Зовнішня частина тора (Рис. 4.12.15) описується через точки правої половинки круга, показаного на Рис. 4.12.13. Цього разу ймовірності розподілені рівномірно на множині G_1:

$$G_1=\{(x,y)\,|\,R^2\le x^2+y^2\le (R+r)^2\},$$
$$m(G_1)=\pi(R+r)^2-\pi R^2=\pi(2Rr+r^2).$$

Нехай на множині G_1 задано функцію $z=\psi(x,y)$ таку, що в графічному поданні поверхнею, через координати точок якої задовольняється рівність $z=\psi(x,y)$, є частина поверхні тора, що лежить над кільцем G_1. Нехай $Z_1=\psi(X,Y)$ – випадкова величина, задана на множині $\Omega=G_1$.

Для функції розподілу ймовірностей на множині значень випадкової величини Z_1 одержуємо:

$$F_{Z_1}(z)=P_{Z_1}((-\infty;z))=P_{(X,Y)}(Z_1^{-1}((\infty;z)))=$$
$$=\begin{cases}0,\text{ коли } z\le 0,\\ \dfrac{m(Z_1^{-1}((-\infty;z)))}{m(G_1)},\text{ коли } 0\le z\le r,\\ 1,\text{ коли } r\le z,\end{cases}$$

або з врахуванням конкретних даних (див. Рис. 4.12.15):

$$F_{Z_1}(z)=\begin{cases}0, \text{ коли } z\le 0,\\ \dfrac{\pi(R+r)^2-\pi(R+r_z)^2}{\pi(R+r)^2-\pi R^2}, \text{ коли } 0\le z\le r,\\ 1, \text{ коли } z\ge r;\end{cases}$$

тобто

$$F_{Z_1}(z)=\begin{cases}0, \text{ коли } z\le 0,\\ \dfrac{2\pi Rr+\pi r^2-2\pi Rr_z-\pi r_z^2}{2\pi Rr+\pi r^2}, \text{ коли } 0\le z\le r,\\ 1, \text{ коли } z\ge r;\end{cases}$$

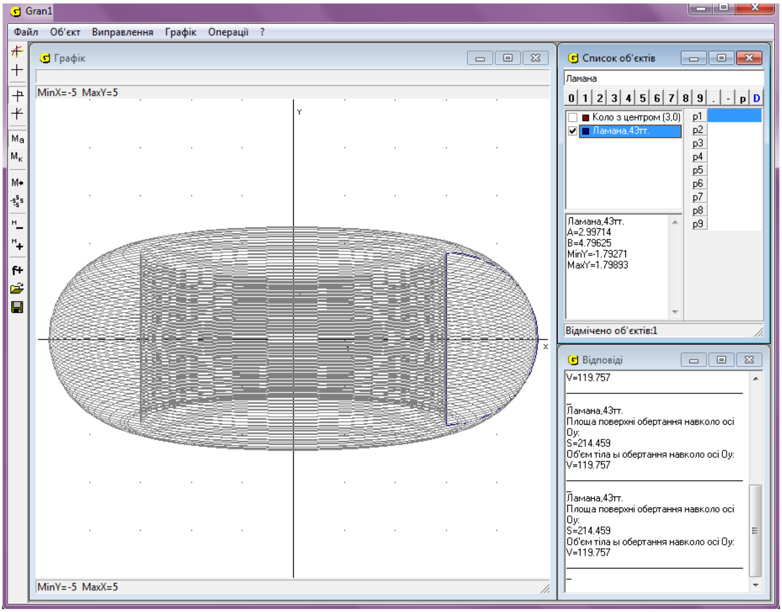


Рис. 4.12.15

звідки

$$F_{Z_1}(z)=\begin{cases}0, \text{ коли } z\le 0,\\ 1-\dfrac{2\pi Rr_z+\pi r_z^2}{2\pi Rr+\pi r^2}, \text{ коли } 0\le z\le r,\\ 1, \text{ коли } z\ge r;\end{cases}$$

або, враховуючи, що $r_z=\sqrt{r^2-z^2}$,

$$F_{Z_1}(z)=\begin{cases}0, \text{ коли } \ z\le 0,\\ 1-\dfrac{2\pi R\sqrt{r^2-z^2}+\pi(r^2-z^2)}{2\pi Rr+\pi r^2}, \text{ коли } \ 0\le z\le r,\\ 1, \text{ коли } \ z\ge r.\end{cases}$$

Обчислюючи тепер щільність $f_{Z_1}(z)$ розподілу ймовірностей на множині значень випадкової величини Z_1, одержимо

$$f_{Z_1}(z)=\begin{cases}0, \text{ коли } z\,\overline{\in}[0;r],\\ \dfrac{1}{2\pi Rr+\pi r^2}\left(2\pi R\dfrac{z}{\sqrt{r^2-z^2}}+2\pi z\right), \text{ коли } z\in[0;r];\end{cases}$$

тобто

$$f_{Z_1}(z)=\begin{cases}0, \text{ коли } z\,\overline{\in}[0;r],\\ \dfrac{1}{r+\dfrac{r^2}{2R}}\left(\dfrac{z}{\sqrt{r^2-z^2}}+\dfrac{z}{R}\right), \text{ коли } z\in[0;r].\end{cases}$$

Обчислюючи математичне сподівання випадкової величини Z_1, одержимо:

$$M[Z_1]=\int_0^r z\cdot\frac{1}{r+\dfrac{r^2}{2R}}\left(\frac{z}{\sqrt{r^2-z^2}}+\frac{z}{R}\right)dz=$$

$$=\frac{2R}{2rR+r^2}\left(\int_0^r\frac{z^2}{\sqrt{r^2-z^2}}\,dz+\int_0^r\frac{z^2}{R}\,dz\right).$$

Обчислимо окремо $\int_0^r\frac{z^2}{\sqrt{r^2-z^2}}dz$. Інтегруючи частинами: $z=u$, $\frac{zdz}{\sqrt{r^2-z^2}}=dv$, $dz=du$, $-\sqrt{r^2-z^2}=v$, одержимо

$$\int_0^r\frac{z^2}{\sqrt{r^2-z^2}}dz=-z\sqrt{r^2-z^2}\Big|_0^r+\int_0^r\sqrt{r^2-z^2}\,dz=$$

$$=0+\int_0^r\frac{r^2}{\sqrt{r^2-z^2}}dz-\int_0^r\frac{z^2}{\sqrt{r^2-z^2}}dz,$$

звідки

$$\int_0^r \frac{z^2}{\sqrt{r^2-z^2}}dz = \frac{1}{2}\int_0^r \frac{r^2}{\sqrt{r^2-z^2}}dz = \ = \frac{r^2}{2}\arcsin\left(\frac{z}{r}\right)\Bigg|_0^r = \frac{r^2}{2}\cdot\frac{\pi}{2} = \frac{\pi r^2}{4}.$$

Таким чином

$$M[Z_1] = \frac{2R}{2rR + r^2}\left(\frac{\pi r^2}{4} + \frac{r^3}{3R}\right),$$

а тому стосовно об'єму зовнішньої частини тора одержуємо:

$$V_{зовн.ч.тора} = 2\cdot M[Z_1]\cdot m(G_1)$$

$$= 2\cdot\frac{2R}{2rR + r^2}\left(\frac{\pi r^2}{4} + \frac{r^3}{3R}\right)\cdot \pi(2Rr + r^2) =$$

$$= 4\pi R\left(\frac{\pi r^2}{4} + \frac{r^3}{3R}\right) = 2\pi R\cdot\frac{\pi r^2}{2} + \frac{4}{3}\pi r^3.$$

Обчислимо тепер об'єм внутрішньої частини тора (Рис. 4.12.16), яка описується через точки лівої половинки круга, показаного на Рис. 4.12.13. Цього разу ймовірності розподілені рівномірно на множині G_2:

$$G_2 = \{(x,y)\,|\,(R-r)^2 \le x^2 + y^2 \le R^2\},$$
$$m(G_2) = \pi R^2 - \pi(R-r)^2 = \pi(2Rr - r^2).$$

Нехай на множині G_2 задано функцію $z = \psi(x,y)$ таку, що в графічному поданні поверхнею, через координати точок якої задовільняється рівність $z = \psi(x,y)$, є частина поверхні тора, що лежить над кільцем G_2. Нехай $Z_2 = \psi(X,Y)$ – випадкова величина, задана на множині $\Omega = G_2$.

Стосовно функції розподілу ймовірностей на множині значень випадкової величини Z_2 одержуємо:

$$F_{Z_2}(z) = P_{Z_2}((-\infty;\, z)) = P_{(X,Y)}({Z_2}^{-1}((-\infty;\, z))) =$$

$$= \begin{cases} 0, \text{ коли } z \le 0, \\ \dfrac{m(Z_2^{-1}((-\infty, z)))}{m(G_2)}, \text{ коли } 0 \le z \le r, \\ 1, \text{ коли } z \le r, \end{cases}$$

або з врахуванням конкретних даних:

$$F_{Z_2}(z)=\begin{cases}0, \ \text{коли} \quad z\le 0,\\ \dfrac{\pi(R-r_z)^2-\pi(R-r)^2}{\pi R^2-\pi(R-r)^2}, \ \text{коли} \quad 0\le z\le r,\\ 1, \ \text{коли} \quad z\ge r;\end{cases}$$

тобто

$$F_{Z_2}(z)=\begin{cases}0, \ \text{коли} \quad z\le 0,\\ \dfrac{2\pi Rr-\pi r^2-2\pi Rr_z+\pi r_z^2}{2\pi Rr-\pi r^2}, \ \text{коли} \quad 0\le z\le r,\\ 1, \ \text{коли} \quad z\ge r;\end{cases}$$

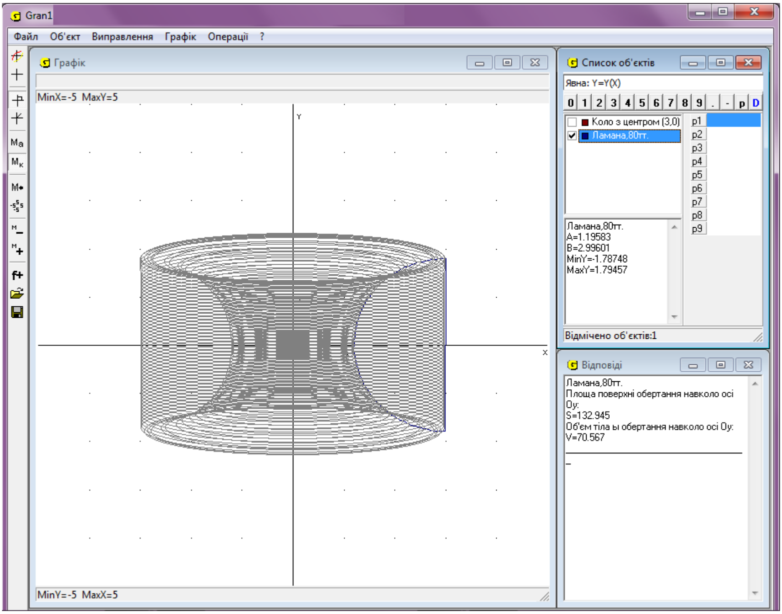


Рис. 4.12.16

звідки

$$F_{Z_2}(z)=\begin{cases}0, \ \text{коли} \quad z\le 0,\\ 1-\dfrac{2\pi Rr_z-\pi r_z^2}{2\pi Rr-\pi r^2}, \ \text{коли} \quad 0\le z\le r,\\ 1, \ \text{коли} \quad z\ge r;\end{cases}$$

або, враховуючи, що $r_z = \sqrt{r^2 - z^2}$

$$F_{Z_2}(z) = \begin{cases} 0, \text{ коли } z \le 0, \\ 1 - \dfrac{2\pi R\sqrt{r^2 - z^2} - \pi(r^2 - z^2)}{2\pi Rr - \pi r^2}, \text{ коли } 0 \le z \le r, \\ 1, \text{ коли } z \ge r. \end{cases}$$

Обчислюючи тепер щільність $f_{Z_2}(z)$ розподілу ймовірностей на множині значень випадкової величини Z_2, одержимо

$$f_{Z_2}(Z) = \begin{cases} 0, \text{коли } z \overline{\in} [0; r], \\ \dfrac{1}{2\pi Rr - \pi r^2}\left(2\pi R\dfrac{z}{\sqrt{r^2 - z^2}} - 2\pi z\right), \text{коли } z \in [0; r]; \end{cases}$$

Тобто

$$f_{Z_2}(Z) = \begin{cases} 0, \text{коли } z \overline{\in} [0; r], \\ \dfrac{1}{r - \dfrac{r^2}{2R}}\left(\dfrac{z}{\sqrt{r^2 - z^2}} - \dfrac{z}{R}\right), \text{коли } z \in [0; r]. \end{cases}$$

Обчислюючи математичне сподівання випадкової величини Z_2, одержимо:

$$M[Z_2] = \int_0^r z \cdot \frac{1}{r - \dfrac{r^2}{2R}}\left(\frac{z}{\sqrt{r^2 - z^2}} - \frac{z}{R}\right)dz =$$

$$= \frac{2R}{2rR - r^2}\left(\int_0^r \frac{z^2}{\sqrt{r^2 - z^2}}\,dz - \int_0^r \frac{z^2}{R}\,dz\right).$$

Оскільки, як було вже показано: $\int_0^r \dfrac{z^2}{\sqrt{r^2 - z^2}}\,dz = \dfrac{\pi r^2}{4}$, одержуємо

$$M[Z_2] = \frac{2R}{2rR - r^2}\left(\frac{\pi r^2}{4} - \frac{r^3}{3R}\right).$$

Отже,

$$V_{вн.ч.тора} = 2 \cdot M[Z_2] \cdot m(G_2) =$$

$$= 2 \cdot \frac{2R}{2rR - r^2}\left(\frac{\pi r^2}{4} - \frac{r^3}{3R}\right) \cdot \pi(2Rr - r^2) =$$

$$= 4\pi R\left(\frac{\pi r^2}{4} - \frac{r^3}{3R}\right) = 2\pi R \cdot \frac{\pi r^2}{2} - \frac{4}{3}\pi r^3.$$

Вправи для самостійного виконання

Використовуючи зв'язок між першим та другим способом обчислення математичного сподівання функції випадкового аргумента (теорему про середнє):

4.12.1. Вивести формулу для обчислення площі паралелограма.

4.12.2. Вивести формулу для обчислення площі кругового сегмента радіуса *R* і висотою *h*.

4.12.3. Вивести формулу для обчислення об'єму зрізаної піраміди.

4.12.4. Вивести формулу для обчислення площі круга радіуса r.

4.12.5. На множині $G = \{(x, y) \mid x^2 + y^2 \le R^2\}$ задано функцію $z = \sqrt{R^2 - (x^2 + y^2)}$. Вивести формулу для обчислення об’єму кулі.

4.12.6. Вивести формулу для обчислення об'єму кульового сегмента радіуса *R* і висотою *h*.

4.13. Кореляція випадкових величин

Нехай X і Y – дві випадкові величини, математичні сподівання яких скінченні. Тоді

$$D[X+Y]=M[(X+Y-M[X+Y])^2]=M[((X-M[X])+$$

$$+(Y-M[Y]))^2]=M[(X-M[X])^2]+2M[(X-M[X])(Y-M[Y])]+$$

$$+M[(Y-M[Y])^2]=D[X]+D[Y]+2M[(X-M[X])(Y-M[Y])].$$

Величину $M[(X-M[X])(Y-M[Y])]$ називають *коваріацією випадкових величин* X і Y і позначають $\operatorname{cov}[X, Y]$. Часто цю величину називають також *кореляційним моментом* випадкових величин X і Y і позначають $K[X, Y]$.

Зауважимо, що

$$K[X,Y]=M[(X-M[X])(Y-M[Y])]=$$

$$=M[XY-XM[Y]-YM[X]+M[X]M[Y]]=$$

$$=M[XY]-M[X]M[Y]-M[Y]M[X]+M[X]M[Y]=M[XY]-M[X]M[Y].$$

Якщо випадкові величини X і Y незалежні, то за властивостями математичного сподівання

$$K[X,Y]=M[(X-M[X])(Y-M[Y])]=$$

$$=M[XY-XM[Y]-YM[X]+M[X]M[Y]]=$$

$$=M[XY]-M[X]M[Y]-M[Y]M[X]+M[X]M[Y]=$$

$$=M[X]M[Y]-M[X]M[Y]-M[Y]M[X]+M[X]M[Y]=$$

$$=M[(X-M[X])]M[(Y-M[Y])]=0.$$

Зокрема з останніх рівностей випливає, що коли випадкові величини X і Y незалежні, то і випадкові величини $X-M[X]$ та $Y-M[Y]$ також незалежні.

Отже, якщо випадкові величини X і Y незалежні, то дисперсія суми $X+Y$ дорівнює сумі дисперсій:

$$D[X+Y]=D[X]+D[Y].$$

Для суми $X_1+X_2+...+X_n$ випадкових величин X_i, $i=\overline{1, n}$,

$$D\left[\sum_{i=1}^{n} X_i\right]=\sum_{i=1}^{n} D[X_i]+2\sum_{i<j} K[X_i, X_j],$$

де $K(X_i, X_j)=K[X_j, X_i]$ – кореляційний момент пари випадкових величин X_i і X_j. Враховуючи, що

$K[X_i,\ X_i] = D[X_i]$, $D\left[\sum_{i=1}^{n} X_i\right]$ можна подати в більш компактному вигляді:

$$D\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} K[X_i,\ X_j].$$

Матриця, складена з елементів $K_{ij} = K[X_i,\ X_j]$, називається *кореляційною матрицею*. Її подання записують у вигляді

$$\begin{pmatrix} K_{11} & K_{12} \ldots K_{1n} \\ K_{21} & K_{22} \ldots K_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ K_{n1} & K_{n2} \ldots K_{nn} \end{pmatrix}.$$

Якщо $K_{ij} = 0$, коли $i \neq j$, то *випадкові величини* $X_1, X_2, ..., X_n$ називаються *попарно некорельованими*. В разі попарно некорельованих випадкових величин так само, як і в разі попарно незалежних,

$$D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D[X_i].$$

Стосовно дисперсії лінійної функції $\sum_{i=1}^{n} a_i X_i + b_i$ випадкових величин $X_1, X_2, ..., X_n$ знаходимо

$$D\left[\sum_{i=1}^{n} a_i X_i + b\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K[X_i,\ X_j].$$

Кореляційний момент є характеристикою зв'язку між випадковими величинами в тому розумінні, що за незалежних випадкових величин X і Y виконується рівність $K[X,Y] = 0$. Якщо $K[X,Y] \neq 0$, то випадкові величини X і Y залежні. Проте з того, що $K[X,Y] = 0$, незалежність випадкових величин X і Y не випливає.

З означення $K[X,Y]$ випливає, що кореляційний момент (коваріація) є характеристикою не тільки зв'язку між випадковими величинами, але і відповідного розсіювання ймовірностей. Справді, якщо хоч одна з випадкових величин мало відрізняється від свого математичного сподівання, то коваріація буде малою, як би не були зв'язані між собою випадкові величини X і Y.

Щоб упевнитися в тому, що незалежність випадкових величин X і Y не випливає з некорельованості, розглянемо такий приклад.

Приклад 4.13.1. Нехай імовірності попадання в точки $(-1,0)$, $(0,-1)$, $(1,0)$, $(0,1)$ дорівнюють $1/4$ кожна (Рис. 4.13.1).

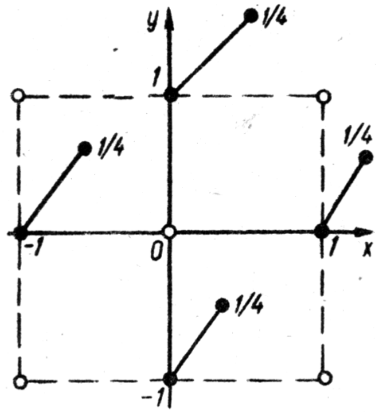


Рис 4.13.1

В такому разі двохвимірний розподіл імовірностей подається у вигляді, наведеному в Табл. 4.13.1.

Табл. 4.13.1

y_i \ x_j	-1	0	1
-1	0	$\frac{1}{4}$	0
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{4}$	0

Знайдемо безумовні розподіли ймовірностей на множинах значень випадкових величин X та Y (Табл. 4.13.2, 4.13.3).

Табл. 4.13.2

x_j	-1	0	1
$P_X(x_j)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Табл. 4.13.3

y_i	-1	0	1
$P_Y(y_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Очевидно, що рівності

$$\mathrm{P}_X(x_j)\mathrm{P}_Y(y_i) = p_{ij} = P_{XY}(\{(x_j; y_i)\})$$

за всіх i та j не справджуються. Отже, випадкові величини X і Y, що є компонентами даної системи двох випадкових величин (X, Y), залежні.

Знайдемо кореляційний момент даного розподілу ймовірностей стосовно системи (X, Y).

Для цього спочатку визначимо математичні сподівання $M[X]$ і $M[Y]$ випадкових величин X і Y:

$$M[X] = \sum_{j=1}^{3} x_j \mathrm{P}_X(x_j) = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0,$$

$$M[Y] = \sum_{i=1}^{3} y_i \mathrm{P}_Y(y_i) = -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0.$$

Тоді

$$K[X, Y] = \sum_{i=1}^{3}\sum_{j=1}^{3}(x_j - M[X])(y_i - M[Y])p_{ij} = (-1)\cdot(-1)\cdot 0 + 0\cdot(-1)\cdot\frac{1}{4} +$$

$$+1\cdot(-1)\cdot 0 + (-1)\cdot 0\cdot\frac{1}{4} + 0\cdot 0\cdot 0 + 1\cdot 0\cdot\frac{1}{4} + (-1)\cdot 1\cdot 0 + 0\cdot 1\cdot\frac{1}{4} + 1\cdot 1\cdot 0 = 0.$$

Таким чином $K[X, Y] = 0$, хоч випадкові величини X і Y залежні.

Досить часто для характеристики зв’язку між випадковими величинами X і Y використовують *коефіцієнт кореляції*

$$r[X, Y] = \frac{K[X, Y]}{\sigma[X]\sigma[Y]}.$$

Через коефіцієнт кореляції характеризується тільки зв'язок між випадковими величинами, тоді як через $K[X, Y]$ крім зв’язку між величинами X і Y характеризується ще й величина розсіювання ймовірностей. Коефіцієнт кореляції є безрозмірною величиною.

Зазначимо, що $K[X, Y] = M[(X - M[X])(Y - M[Y])]$ можна розглядати як скалярний добуток величин $(X - M[X])$ і $(Y - M[Y])$. Тоді $K[X, X] = D[X]$, $K[Y, Y] = D[Y]$ – скалярні добутки елементів $(X - M[X])$ чи $(Y - M[Y])$ самих на себе, тобто квадрати норм елементів $(X - M[X])$ і $(Y - M[Y])$, а $\sigma[X]$ і $\sigma[Y]$ – норми елементів $(X - M[X])$ і $(Y - M[Y])$. Стосовно введеного таким чином скалярного добутку елементів

$(X - M[X])$ і $(Y - M[Y])$ та їх норм виконується нерівність Коші-Буняковського

$$K[X, Y] \le \sigma[X]\sigma[Y].$$

Отже $|r[X, Y]| \le 1$. Коефіцієнт кореляції є аналогом косинуса кута між векторами $(X - M[X])$ і $(Y - M[Y])$.

Якщо між випадковими величинами X і Y існує лінійна залежність $Y = aX + b$, то $|r(X, Y)| = 1$, причому коли $a > 0$, то $r[X, Y] = 1$; коли $a < 0$, то $r[X, Y] = -1$. Справді, оскільки $M[Y] = aM[X] + b$, $D[Y] = a^2 D[X]$, то

$$r[X, Y] = \frac{M[(X - M[X])(Y - M[Y])]}{\sigma[X]\sigma[Y]} =$$

$$= \frac{M[(X - M[X])\,a\,(X - M[X])]}{\sigma[X]|a|\sigma[X]} =$$

$$= \frac{a}{|a|}\frac{M[(X - M[X])^2]}{(\sigma[X])^2} = \frac{a}{|a|}.$$

Виявляється, що й навпаки, коли $|r[X, Y]| = 1$, то між випадковими величинами X і Y існує лінійна залежність виду $Y = aX + b$, причому $a > 0$, коли $r[X, Y] = 1$, і $a < 0$, коли $r[X, Y] = -1$ (оскільки в нерівності Коші-Буняковського рівність досягається тільки тоді, коли між величинами X і Y існує лінійна залежність).

Приклад 4.13.2. Опис щільності розподілу ймовірностей на множині значень випадкового вектора (X, Y) подано у вигляді

$$f_{(X,Y)}(x, y) = \begin{cases} \dfrac{1}{\pi R^2}, \text{ коли } (x, y) \in W; \\ 0, \quad \text{коли } (x, y) \overline{\in} W, \end{cases}$$

де $W = \{(x, y) \mid x^2 + y^2 \le R^2\}$. Знайти математичні сподівання $M[X]$, $M[Y]$ і кореляційний момент $K[X, Y]$.

Одержуємо:

$$M[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-R}^{R} x \frac{2}{\pi R^2}\sqrt{R^2 - x^2}\,dx =$$

$$= -\frac{1}{\pi R^2}\int_{-R}^{R}(R^2 - x^2)^{1/2} d(R^2 - x^2) = -\frac{1}{\pi R^2}\cdot\frac{2}{3}(R^2 - x^2)^{3/2}\Big|_{-R}^{R} = 0.$$

Аналогічно $M[Y]=0$. Таким чином, центром розсіювання ймовірностей на множині значень випадкового вектора (X, Y) в розглядуваному випадку є точка $C(0,0)$. Цей результат очевидний, якщо враховувати механічну інтерпретацію розподілу ймовірностей.

Стосовно $K[X, Y]$ одержимо

$$K[X, Y]=\iint_{R^2}(x-M[X])(y-M[Y])\,f_{(X,Y)}(x,y)dxdy=\iint_{W} xy\frac{1}{\pi R^2}dxdy=$$

$$=\frac{1}{\pi R^2}\int_{-R}^{R}x\left\{\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}ydy\right\}dx=\frac{1}{\pi R^2}\int_{-R}^{R}x\left\{\frac{y^2}{2}\Bigg|_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}\right\}dx=$$

$$=\frac{1}{2\pi R^2}\int_{-R}^{R}x\{(R^2-x^2)-(R^2-x^2)\}dx=0.$$

Таким чином, $K[X, Y]=0$, хоч, як легко перевірити, випадкові величини X і Y залежні, бо

$$f_X(x)\cdot f_Y(y)=\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dy\cdot\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dx\neq f_{(X,Y)}(x,y)$$

за всіх $(x,y)\in R^2$.

Тут

$$f_X(x)=\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dy=\begin{cases}\dfrac{2\sqrt{R^2-x^2}}{\pi R^2}, & \text{коли } x\in[-R;\,R],\\ 0, & \text{коли } x\overline{\in}[-R;\,R],\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dx=\begin{cases}\dfrac{2\sqrt{R^2-y^2}}{\pi R^2}, & \text{коли } y\in[-R;\,R],\\ 0, & \text{коли } y\overline{\in}[-R;\,R],\end{cases}$$

$$f_{(X,Y)}(x,y)=\begin{cases}\dfrac{1}{\pi R^2}, & \text{коли } x^2+y^2\le R^2,\\ 0, & \text{коли } x^2+y^2>R^2.\end{cases}$$

Вправи для самостійного виконання

4.13.1 Перевірити, чи правильні твердження:

1. За будь-яких випадкових величин X та Y існує коваріація $\operatorname{cov}[X,Y]$.

2. Кореляційний момент $K[X,Y]$ випадкових величин X та Y – це коваріація $\operatorname{cov}[X,Y]$.

3. За будь-яких випадкових величин X та Y правильна рівність

$$K[X,Y] = M[XY] - M[X] \cdot M[Y].$$

4. $K[X,Y] = 0$ тоді й тільки тоді, коли величини X та Y незалежні.

5. Дисперсія випадкової величини X – це кореляційний момент $K[X, X]$.

6. За будь-яких випадкових величин X_k, $k \in \overline{1,n}$, існує відповідна кореляційна матриця.

7. Якщо випадкові величини X_k, $k \in \overline{1,n}$, незалежні, то відповідна кореляційна матриця є діагональною.

8. $D[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} D[X_i]$ тоді й тільки тоді, коли величини X_i попарно некорельовані.

9. Коефіцієнт кореляції – це те саме, що й кореляційний момент.

10. Коефіцієнт кореляції може набувати будь-яких значень.

11. Стосовно коефіцієнта кореляції $r[X,Y]$ задовольняється рівність $|r[X,Y]| = 1$ тоді й тільки тоді, коли між випадковими величинами існує лінійна залежність.

4.14. Умовні розподіли ймовірностей

Нехай $X = X(E)$ і $Y = Y(E)$, $E \in \Omega$, – випадкові величини з одновимірними дискретними розподілами ймовірностей, причому X набуває значень з множини $Q_1 = \{x_1, x_2, ..., x_l\}$, Y набуває значень з множини $Q_2 = \{y_1, y_2, ..., y_m\}$, $Q = Q_1 \times Q_2$.

Згідно з формулою стосовно ймовірності добутку подій одержимо

$$P_{(X,Y)}(\{y_i\}\{x_j\}) = P_Y(\{y_i\})P_{X/Y}(\{x_j\}/\{y_i\}),$$

де $P_{X/Y}(\{x_j\}/\{y_i\})$ – умовна ймовірність події $(\{x_j\})$, знайдена за умови, що подія $\{y_i\}$ відбулася.

Таким чином, коли $P_Y(\{y_i\}) \neq 0$, одержуємо:

$$P_{X/Y}(\{x_j\}/\{y_i\}) = \frac{P_{(X,Y)}(\{x_j\}\{y_i\})}{P_Y(\{y_i\})} = \frac{p_{ij}}{\sum\limits_{j=1}^{l} p_{ij}}, \ i = \overline{1, m}, \ j = \overline{1, l}.$$

Якщо $P_Y(\{y_i\}) = 0$, то й $P_{(X,Y)}(\{y_i\}\{x_j\}) = 0$. Тому $P_{X/Y}(\{x_j\}/\{y_i\})$ в такому разі можна визначити довільно. Отже, за кожного x_j, $j = 1, 2, ..., l$, можна знайти відповідну умовну ймовірність $P_{X/Y}(\{x_j\}/\{y_i\})$ за умови, що випадкова величина Y набуває значення y_i, $i = 1, 2, ..., m$. Таким чином дістаємо так званий *умовний дискретний розподіл ймовірностей* на множині значень випадкової величини X за умови, що випадкова величина Y набуває значення y_i, $i = 1,2,..., m$. Числа $P_{X/Y}(\{x_j\}/\{y_i\})$, $i = 1,2,..., m$, $j = 1,2,..., l$, можна тлумачити як значення деякої випадкової величини $P_{X/Y}(\{x_j\}/Y)$, $i = 1,2,..., m$, яких вона набуває з імовірностями $P_Y(\{y_i\})$, $i = 1, 2, ..., m$. За формулою повної ймовірності

$$\sum_{i=1}^{m} P_{X/Y}(\{x_j\}/\{y_i\})P_Y(\{y_i\}) = P_X(\{x_j\}), \ j = 1,2,..., l.$$

Отже, $P_X(\{x_j\})$ є математичним сподіванням випадкової величини $P(\{x_j\}/Y)$.

Аналогічно можна дістати умовні розподіли ймовірностей на множині значень випадкової величини Y за умови, що випадкова величина X набуває значень x_j, $j = 1,2,..., l$:

$$P_{Y/X}(\{y_i\}/\{x_j\}) = \frac{P_{(X,Y)}(\{y_i\}\{x_j\})}{P_X(\{x_j\})} = \frac{p_{ij}}{\sum_{i=1}^{m} p_{ij}},$$

$$P_X(\{x_j\}) \neq 0, \ \ i = \overline{1,\ m}, \ \ \ j = \overline{1,\ l}\ .$$

В механічній інтерпретації $P_{X/Y}(\{x_j\}/\{y_i\})$ є частка маси, розподіленої на множині точок з Q вздовж прямої $y = y_i$, яка припадає на множину тих точок з Q, що лежать на прямій $x = x_j$. Аналогічно $P_{Y/X}(\{y_i\}/\{x_j\})$ є частка маси, розподіленої в точках з Q вздовж прямої $x = x_j$, яка припадає на множину тих точок з Q, що лежать на прямій $y = y_i$.

Якщо випадкові величини X і Y незалежні, то $P_{(X,Y)}(\{x_j\}\{y_i\}) = P_X(\{x_j\})P_Y(\{y_i\})$ за будь-яких i та j.

Випадкова величина Y не залежить від випадкової величини X, якщо за будь-якого x_j умовний розподіл імовірностей на множині значень випадкової величини Y співпадає з безумовним розподілом імовірностей на множині значень цієї самої випадкової величини, тобто якщо

$$P_{Y/X}(\{y_i\}/\{x_j\}) = P_Y(\{y_i\}) \text{ за будь-якого } x_j,$$

або

$$\frac{p_{ij}}{\sum_{i=1}^{m} p_{ij}} = \sum_{j=1}^{l} p_{ij},$$

тобто

$$p_{ij} = \sum_{i=1}^{m} p_{ij} \sum_{j=1}^{l} p_{ij}, \ \ i = \overline{1,\ m}, \ \ \ j = \overline{1,\ l}\ .$$

Можна показати, що коли випадкова величина Y не залежить від випадкової величини X, то і випадкова величина X не залежить від випадкової величини Y.

Приклад 4.14.1. Для виконання спортивної вправи учневі надається три спроби. Ймовірність того, що він правильно виконає вправу в кожній спробі, дорівнює $1/2$.

Нехай випадкова величина X – кількість правильно виконаних вправ у трьох спробах, випадкова величина Y – кількість неправильно виконаних вправ. Очевидно, що тут випадкові величини X і Y залежні, причому функціонально,

$$X + Y = 3, \text{ або } Y = 3 - X\ .$$

Тому сумісний розподіл імовірностей на декартовому добутку множин значень випадкових величин X і Y набуває вигляду (див. схеми Бернуллі і Пуассона, §2.10):

$$p_{ij} = P_{(X,Y)}((\{x_j\})(\{y_i\})) = 0 \text{, коли } x_j + y_i \neq 3 \text{,}$$

як імовірності неможливих подій,

$$p_{0,3} = P_{(X,Y)}((\{3\})(\{0\})) = P_X(\{3\})P_{Y/X}(\{0\}/\{3\}) \text{,}$$

де $P_X(\{3\}) = C_3^3 (1/2)^3 (1/2)^0 = 1/8$ (за формулою Бернуллі), а $P_{Y/X}(\{0\}/\{3\}) = 1$ як імовірність вірогідної події. Тому $p_{0,3} = 1/8$.

Аналогічно $p_{1,2} = 3/8$ і $p_{2,1} = 3/8$, $p_{3,0} = 1/8$.

За цим двохвимірним розподілом

y_i \ x_j	0	1	2	3
0	0	0	0	1/8
1	0	0	3/8	0
2	0	3/8	0	0
3	1/8	0	0	0

можна визначити ряди розподілу на множинах значень випадкових величин X і Y (Табл.4.14.1, 4.14.2).

Табл. 4.14.1

x_j	0	1	2	3
p_j	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Табл. 4.14.2

y_i	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\sum_{j=0}^{3} p_j = 1, \ \sum_{i=0}^{3} p_i = 1.$$

Приклад 4.14.2. Двоє стрільців стріляють в мішень так: спочатку стріляє один, і в разі, коли він влучить, тоді стріляє другий.

Нехай випадкова величина X – кількість влучень у першого стрільця, Y – кількість влучень у другого стрільця. Треба побудувати сумісний розподіл імовірностей на декартовому добутку множин значень випадкових величин X і Y, якщо перший стрілець виконує тільки один постріл і влучає з імовірністю 0.6, а другий стрілець влучає з імовірністю 0.8.

Очевидно, що кожна з випадкових величин X і Y можуть набувати тільки двох значень: $x = 0$, $x = 1$ і $y = 0$, $y = 1$. Якщо перший стрілець промахнеться, то в другого не буде жодного влучення. Отже,

$$p_{11} = P_{(X,Y)}((\{0\})(\{0\})) = P_X(\{0\})P_{Y/X}(\{0\}/\{0\}) = 0.4 \cdot 1 = 0.4 \text{,}$$

$$p_{21} = P_{(X,Y)}((\{0\})(\{1\})) = P_X(\{0\})P_{Y/X}(\{1\}/\{0\}) = 0.4 \cdot 0 = 0,$$

$$p_{12} = P_{(X,Y)}((\{1\})(\{0\})) = P_X(\{1\})P_{Y/X}(\{0\}/\{1\}) = 0.6 \cdot 0.2 = 0.12,$$

$$p_{22} = P_{(X,Y)}((\{1\})(\{1\})) = P_X(\{1\})P_{Y/X}(\{1\}/\{1\}) = 0.6 \cdot 0.8 = 0.48.$$

Матриця розподілу ймовірностей на декартовому добутку множин значень випадкових величин X і Y набуває вигляду (табл. 4.14.3),

Табл. 4.14.3

$y_j \backslash x_j$	0	1
0	0.40	0.12
1	0.00	0.48

Ряди розподілу ймовірностей на множинах значень випадкових величин X і Y подані відповідно в Табл. 4.14.4 та Табл. 4.14.5.

Табл. 4.14.4

x_j	0	1
p_j	0.4	0.6

Табл. 4.14.5

y_i	0	1
p_i	0.52	0.48

Умовні розподіли ймовірностей на множині значень випадкової величини X набувають вигляду, поданого в Табл. 4.14.6, 4.14.7.

Табл. 4.14.6

x_j	0	1
$P_{X/Y}(\{x_j\}/\{0\})$	$\frac{40}{52}$	$\frac{12}{52}$

Табл. 4.14.7

x_j	0	1
$P_{X/Y}(\{x_j\}/\{1\})$	0	1

Умовні розподіли ймовірностей на множині значень випадкової величини Y набувають вигляду, поданого в Табл. 4.14.8, 4.14.9.

Табл. 4.14.8

y_i	0	1
$P_{Y/X}(\{y_i\}/\{0\})$	1	0

Табл. 4.14.9

y_i	0	1
$P_{Y/X}(\{y_i\}/\{1\})$	$\frac{12}{60}$	$\frac{48}{60}$

Отже, оскільки умовні розподіли ймовірностей на множинах значень випадкових величин X і Y не співпадають з безумовними, випадкові величини X і Y залежні.

Приклад 4.14.3. Сумісний розподіл ймовірностей на декартовому добутку множин значень випадкових величин X і Y описано в Табл. 4.14.10.

Чи залежні дані випадкові величини X і Y?

Табл. 4.14.10

$y_j \backslash x_j$	0	1
1	$\frac{1}{20}$	$\frac{4}{20}$
2	$\frac{3}{20}$	$\frac{12}{20}$

Безумовні розподіли ймовірностей на множинах значень випадкових величин X і Y набувають вигляду, поданого відповідно в Табл. 4.14.11, 4.14.12.

Табл. 4.14.11

x_j	0	1
p_j	$\frac{4}{20}$	$\frac{16}{20}$

Табл. 4.14.12

y_i	1	2
p_i	$\frac{5}{20}$	$\frac{15}{20}$

Умовні розподіли ймовірностей на множині значень випадкової величини X набувають вигляду, поданого в Табл. 4.14.13, 4.14.14.

Табл. 4.14.13

x_j	0	1
$P_{X/Y}(\{x_j\}/\{1\})$	$\frac{1}{5}$	$\frac{4}{5}$

Табл. 4.14.14

x_j	0	1
$P_{X/Y}(\{x_j\}/\{2\})$	$\frac{3}{15}$	$\frac{12}{15}$

Таким чином, всі умовні розподіли ймовірностей на множині значень випадкової величини X співпадають з безумовним, а отже, випадкова величина X не залежить від випадкової величини Y.

Стосовно умовних розподілів ймовірностей на множині значень випадкової величини Y одержуємо (Табл. 4.14.15, 4.14.16):

Табл. 4.14.15

y_i	1	2
$P_{Y/X}(\{y_i\}/\{0\})$	$\frac{1}{4}$	$\frac{3}{4}$

Табл. 4.14.16

y_i	0	1
$P_{Y/X}(\{y_i\}/\{1\})$	$\frac{4}{16}$	$\frac{12}{16}$

Таким чином, умовні розподіли ймовірностей на множині значень випадкової величини Y теж співпадають з безумовним розподілом ймовірностей на множині значень цієї випадкової величини.

Приклад 4.14.4. Розподіл ймовірностей на множині значень R^2-значної випадкової величини описаний в Табл. 4.14.17.

Табл. 4.14.17

$y_j \backslash x_j$	-1	0	1
-1	0	$\frac{1}{4}$	0
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{4}$	0

Можна показати, що випадкові величини X і Y залежні. Проте після повороту осей координат на кут $\frac{\pi}{4}$, тобто після перетворення

$$X' = X\cos\frac{\pi}{4} + Y\sin\frac{\pi}{4},\ Y' = -X\sin\frac{\pi}{4} + Y\cos\frac{\pi}{4},$$

на декардовому добутку множин значень випадкових величин X' і Y' сумісний розподіл ймовірностей набуде вигляду, поданого в Табл. 4.14.18.

Табл. 4.14.18

$y'_i \backslash x'_j$	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$
$-\frac{\sqrt{2}}{2}$	$\frac{1}{4}$	0	$\frac{1}{4}$
0	0	0	0
$\frac{\sqrt{2}}{2}$	$\frac{1}{4}$	0	$\frac{1}{4}$

Оскільки за будь-яких x'_j і y'_i виконуються рівності $P_{(X',Y')}((\{x'_j\}),(\{y'_i\})) = P_{X'}(\{x'_j\})P_{Y'}(\{y'_i\})$, то випадкові величини X' і Y' незалежні. Зазначимо, що умовні ймовірності $P_{Y'/X'}(\{y'_i\}/\{0\})$ та $P_{X'/Y'}(\{x'_j\}/\{0\})$ невизначені, їх можна визначити довільно.

Таким чином в деяких випадках за рахунок перетворення системи координат можна перейти від залежних випадкових величин до незалежних, досліджувати які простіше.

Нехай (X,Y) – R^2-значна випадкова величина з випадковими координатами X і Y, причому двохвимірний розподіл ймовірностей на відповідній множині значень визначається за щільністю $f_{(X,Y)}(x,y)$.

Ймовірність того, що значення R^2-значної випадкової величини (X,Y) лежить в прямокутнику

$$R_\Delta(x_0,y_0)=\{(x,y)\mid\ x\in[x_0\,;\,x_0+\Delta x),\ y\in[y_0\,;\,y_0+\Delta y)\},$$

є

$$P_{(X,Y)}(R_\Delta(x_0,y_0))=P_{(X,Y)}(([x_0\,;x_0+\Delta x))\times([y_0\,;y_0+\Delta y)))\,.$$

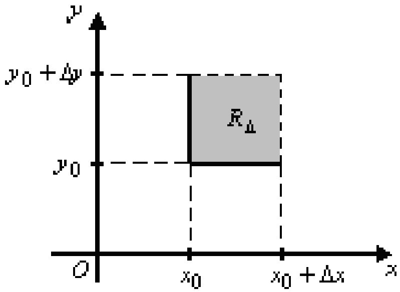


Рис. 4.14.1

Розглядаючи подію $(R_\Delta(x_0,y_0))$ як добуток подій $([x_0\,;\,x_0+\Delta x)\times(-\infty;\infty))$ і $((-\infty;\infty)\times[y_0\,;\,y_0+\Delta y))$, дістаємо

$$P_{(X,Y)}(([x_0\,;\,x_0+\Delta x))\times([y_0\,;\,y_0+\Delta y)))=$$

$$=P_X([x_0\,;\,x_0+\Delta x))\,P_{Y/X}(([y_0\,;\,y_0+\Delta y))\big/([x_0\,;\,x_0+\Delta x))).$$

Останнє співвідношення означає, що шукана ймовірність дорівнює добутку безумовної імовірності того, що значення випадкової величини X лежить в межах $[x_0\,;x_0+\Delta x)$, та умовної ймовірності того, що значення випадкової величини Y лежить в межах $[y_0\,;y_0+\Delta y)$, знайденої за умови, що подія $([x_0\,;x_0+\Delta x))$ відбулася, тобто за умови, що значення випадкової величини X лежить в межах $[x_0,x_0+\Delta x)$. Тоді в разі, коли $f_X(x)$ неперервна в точці x_0 і $f_X(x_0)\neq 0$, буде

$$P_{Y/X}(([y_0\,;\,y_0+\Delta y))\big/([x_0\,;\,x_0+\Delta x)))=$$

$$=\frac{P_{(X,Y)}(([x_0\,;\,x_0+\Delta x))\times([y_0\,;\,y_0+\Delta y)))}{P_{(X,Y)}(([x_0\,;x_0+\Delta x))\times((-\infty;\infty)))}=$$

$$=\frac{\int\limits_{x_0}^{x_0+\Delta x}\left\{\int\limits_{y_0}^{y_0+\Delta y} f_{(X,Y)}(x,y)dy\right\}dx}{\int\limits_{x_0}^{x_0+\Delta x}\left\{\int\limits_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy\right\}dx}=\frac{\int\limits_{x_0}^{x_0+\Delta x}\left\{\int\limits_{y_0}^{y_0+\Delta y} f_{(X,Y)}(x,y)dy\right\}dx}{\int\limits_{x_0}^{x_0+\Delta x} f_X(x)dx}\underset{\Delta x\to 0}{\to}$$

$$\underset{\Delta x\to 0}{\to}\frac{\int\limits_{y_0}^{y_0+\Delta y} f_{(X,Y)}(x_0,y)dy}{f_X(x_0)}=\int\limits_{y_0}^{y_0+\Delta y}\frac{f_{(X,Y)}(x_0,y)}{f_X(x_0)}dy=\int\limits_{y_0}^{y_0+\Delta y} f_{Y/X}(y/x_0)dy\,,$$

де $f_{Y/X}(y/x_0)$ – щільність умовного розподілу імовірностей на множині значень випадкової величини Y за умови, що випадкова величина X набула значення x_0, тобто за умови, що подія $(X=x_0)$ відбулася. Оскільки (x_0,y_0) довільна точка, то

$$f_{(X,Y)}(x,y)=f_X(x)f_{Y/X}(y/x)\,.$$

Аналогічно

$$f_{(X,Y)}(x,y)=f_Y(y)f_{X/Y}(x/y)\,.$$

Звідси

$$f_{Y/X}(y/x)=\frac{f_{(X,Y)}(x,y)}{f_X(x)}\,,\ \ f_{X/Y}(x/y)=\frac{f_{(X,Y)}(x,y)}{f_Y(y)}\,. \qquad (4.14.1)$$

Якщо щільність умовного розподілу ймовірностей на множині значень випадкової величини Y не залежить від того, якого значення набуває випадкова величина X, тобто якщо $f_{Y/X}(y/x)=f_Y(y)$, то випадкова величина Y не залежить від випадкової величини X.

Якщо випадкова величина Y не залежить від випадкової величини X, то і випадкова величина X не залежить від випадкової величини Y.

Якщо $f_{(X,Y)}(x,y)=f_X(x)f_Y(y)$, причому щільності $f_{XY}(x,y)$, $f_X(x)$ і $f_Y(y)$ – неперервні, то і $F_{(X,Y)}(x,y)=F_X(x)F_Y(y)$, і навпаки.

Справді, якщо

$$R_0=\{(x,y)\mid\ x\in(-\infty;x_0),\ y\in(-\infty;y_0)\}\,,$$

то

$$F_{(X,Y)}(x_0,y_0)=\iint\limits_{R_0} f_{(X,Y)}(x,y)dxdy=\int\limits_{-\infty}^{x_0}\left\{\int\limits_{-\infty}^{y_0} f_{(X,Y)}(x,y)dy\right\}dx=$$

$$= \int_{-\infty}^{x_0} \left\{ \int_{-\infty}^{y_0} f_X(x) f_Y(y) dy \right\} dx = \int_{-\infty}^{x_0} f_X(x) dx \int_{-\infty}^{y_0} f_Y(y) dy = F_X(x_0) F_Y(y_0) .$$

Оскільки точка (x_0, y_0) довільна, то стосовно незалежних випадкових величин X і Y остаточно знаходимо

$$F_{(X,Y)}(x,y) = F_X(x) F_Y(y) .$$

Навпаки, якщо

$$F_{(X,Y)}(x,y) = F_X(x) F_Y(y) \text{, то і } f_{(X,Y)}(x,y) = f_X(x) f_Y(y) .$$

Приклад 4.14.5. Нехай розподіл ймовірностей на множині значень R^2-значної випадкової величини (S,W) рівномірний у прямокутнику $D = \{(s,w) \mid |s| + |w| \le \sqrt{2}\}$ (Рис. 4.14.2), тобто

$$f_{(S,W)}(s,w) = \begin{cases} \frac{1}{4}, \text{ коли } (s,w) \in D; \\ 0, \text{ коли } (s,w) \overline{\in} D. \end{cases}$$

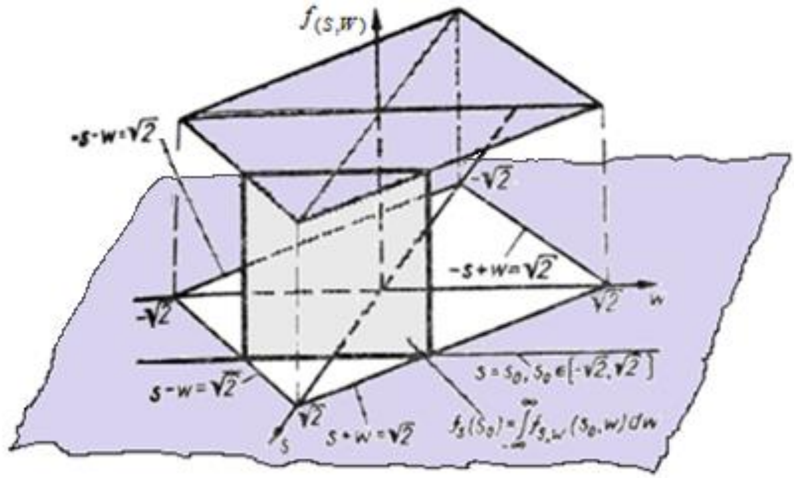


Рис. 4.14.2

Тоді

$$f_S(s) = \int_{-\infty}^{\infty} f_{(S,W)}(s,w) dw = \begin{cases} 0, \quad \text{коли } s \overline{\in} \left[-\sqrt{2}; \sqrt{2}\right]; \\ \frac{1}{2}(\sqrt{2}+s), \quad \text{коли } s \in \left[-\sqrt{2}; 0\right]; \\ \frac{1}{2}(\sqrt{2}-s), \text{ коли } s \in \left[0; \sqrt{2}\right]. \end{cases}$$

Аналогічно

$$f_W(w)=\int\limits_{-\infty}^{\infty} f_{(S,W)}(s,w)ds=\begin{cases} 0, & \text{коли } s\notin\left[-\sqrt{2}\,;\sqrt{2}\right];\\ \dfrac{1}{2}(\sqrt{2}+w), & \text{коли } s\in\left[-\sqrt{2}\,;0\right];\\ \dfrac{1}{2}(\sqrt{2}-w), & \text{коли } s\in\left[0\,;\sqrt{2}\right]. \end{cases}$$

Оскільки $f_S(s)f_W(w)\neq f_{S,W}(s,w)$, то випадкові величини S і W залежні.

Виконаємо тепер перетворення $X=S\cos\dfrac{\pi}{4}+W\sin\dfrac{\pi}{4}$, $Y=-S\sin\dfrac{\pi}{4}+W\cos\dfrac{\pi}{4}$, тобто повернемо систему координат на кут $\dfrac{\pi}{4}$.

Тоді (Рис. 4.14.3)

$$f_{(X,Y)}(x,y)=\begin{cases} 0, \text{ коли } (x,y)\notin D;\\ \dfrac{1}{4}, \text{ коли } (x,y)\in D, \end{cases}$$

де (в нових координатах)

$$D=\{(x,y)\,|\; x\in[-1;1],\; y\in[-1;1]\}.$$

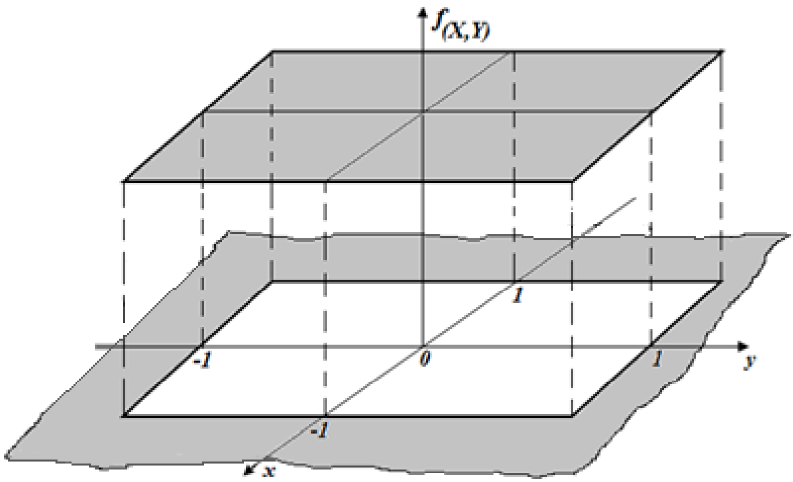


Рис. 4.14.3

Можна показати, що

$$f_X(x)=\begin{cases}\frac{1}{2}, \text{ коли } x\in[-1;1];\\ 0, \text{ коли } x \overline{\in} [-1;1],\end{cases} \qquad f_Y(y)=\begin{cases}\frac{1}{2}, \text{ коли } y\in[-1;1];\\ 0, \text{ коли } y \overline{\in} [-1;1].\end{cases}$$

Отже,

$$f_X(x)f_Y(y)=f_{(X,Y)}(x,y).$$

Таким чином, випадкові величини X і Y незалежні.
Визначаючи функцію розподілу ймовірностей, дістаємо

$$F_{(X,Y)}(x,y)=\int_{-\infty}^{x}\left\{\int_{-\infty}^{y} f_{(X,Y)}(x,y)dy\right\}dx=\int_{-\infty}^{x} f_X(x)dx\int_{-\infty}^{y} f_Y(y)dy=$$

$$=F_X(x)F_Y(y)=\begin{cases}0, & \text{коли } x\le -1 \text{ або } y\le -1;\\ \frac{1}{4}(x+1)(y+1), & \text{коли } -1\le x\le 1 \text{ і } -1\le y\le 1;\\ \frac{x}{2}+\frac{1}{2}, & \text{коли } -1\le x\le 1 \text{ і } 1\le y;\\ \frac{y}{2}+\frac{1}{2}, & \text{коли } 1\le x \text{ і } -1\le y\le 1;\\ 1, & \text{коли } 1\le x \text{ і } 1\le y,\end{cases}$$

Оскільки

$$F_X(x)=\begin{cases}0, & \text{коли } x\le 0;\\ \frac{x}{2}+\frac{1}{2}, & \text{коли } -1\le x\le 1;\\ 1, & \text{коли } 1\le x,\end{cases}$$

$$F_Y(y)=\begin{cases}0, & \text{коли } y\le 0;\\ \frac{y}{2}+\frac{1}{2}, & \text{коли } -1\le y\le 1;\\ 1, & \text{коли } 1\le y.\end{cases}$$

Поверхню $z=F_{(X,Y)}(x,y)$ стосовно розглядуваного розподілу ймовірностей зображено на Рис. 4.14.4.

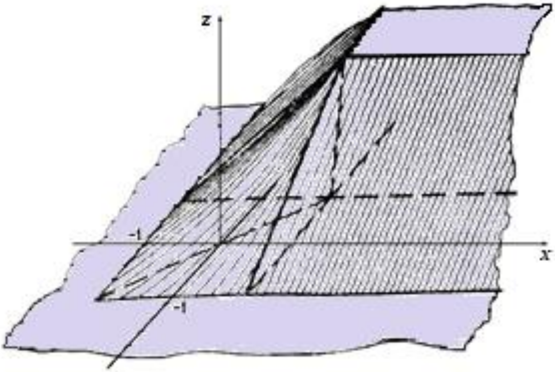


Рис. 4.14.4

Приклад 4.14.6. Опис щільності двохвимірного розподілу ймовірностей $f_{(X,Y)}(x,y)$ подано у вигляді

$$f_{(X,Y)}(x,y)=\begin{cases} c, & \text{коли} \quad (x,y)\in D; \\ 0, & \text{коли} \quad (x,y)\overline{\in} D, \end{cases}$$

де $D=\{(x,y)|\ x^2+y^2\le R^2\}$.

Знайти: сталу c, $f_X(x)$ і $f_Y(y)$, щільності умовних розподілів ймовірностей $f_{Y/X}(y/x)$ та $f_{X/Y}(x/y)$. З'ясувати, залежні чи ні випадкові величини X і Y. Знайти ймовірність того, що випадкова точка (X,Y) лежатиме в області

$$D_1=\left\{(x,y)|\ x^2+y^2\le\frac{R^2}{4}\right\}.$$

Поверхню $z=f_{(X,Y)}(x,y)$ стосовно даного випадку зображено на Рис. 4.14.5. На основі властивості щільності

$$\iint\limits_{R^2} f_{(X,Y)}(x,y)\,dxdy=1,$$

де $R^2=\{(x,\,y)|\ -\infty<x<\infty,-\infty<y<\infty\}$.

Враховуючи, що в геометричному тлумаченні подвійний інтеграл є об'єм тіла, що знаходиться під поверхнею $z=f_{(X,Y)}(x,y)$ і над площиною $z=0$, а також те, що об'єм тіла під поверхнею щільності повинен дорівнювати 1, одержуємо

$$\iint\limits_{R^2} f_{(X,Y)}(x,y)\,dxdy=c\pi R^2=1\,.$$

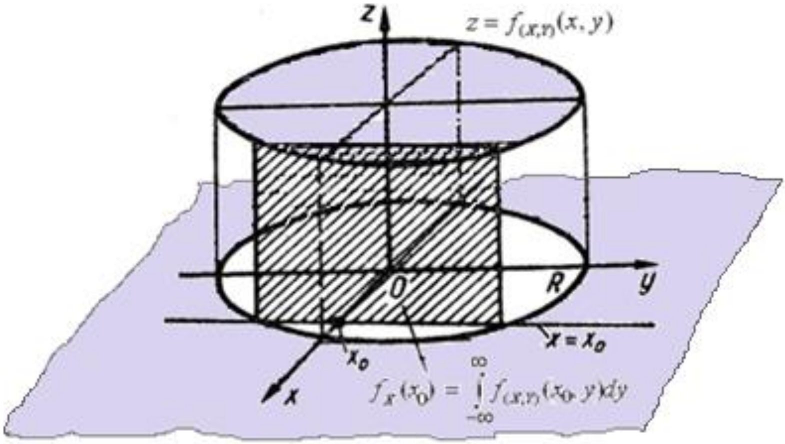


Рис. 4.14.5

Отже, $c = \dfrac{1}{\pi R^2}$.

Визначимо $f_X(x)$ і $f_Y(y)$:

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)dy = \begin{cases} 0, & \text{коли } x \overline{\in} [-R; R]; \\ \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \dfrac{1}{\pi R^2} dy = \dfrac{2\sqrt{R^2 - x^2}}{\pi R^2}, & \text{коли } x \in [-R; R], \end{cases}$$

де $y = -\sqrt{R^2 - x^2}$ – рівняння півкола, що лежить нижче від осі Ox, $y = +\sqrt{R^2 - x^2}$ – рівняння півкола, що лежить вище від осі Ox. Аналогічно стосовно $f_Y(y)$ одержимо

$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)dx = \begin{cases} 0, & \text{коли } y \overline{\in} [-R; R]; \\ \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} \dfrac{1}{\pi R^2} dx = \dfrac{2\sqrt{R^2 - y^2}}{\pi R^2}, & \text{коли } y \in [-R; R]. \end{cases}$$

За формулою (4.14.1) знаходимо

$$f_{Y/X}(y/x) = \frac{f_{(X,Y)}(x, y)}{f_X(x)} =$$

$$=\begin{cases} 0, & \text{коли } x\in[-R;R],\ y\,\overline{\in}\,[-\sqrt{R^2-x^2};\sqrt{R^2-x^2}]; \\ \dfrac{1}{\pi R^2}\dfrac{\pi R^2}{2\sqrt{R^2-x^2}}, & \text{коли } x\in[-R;R],\ y\in[-\sqrt{R^2-x^2};\sqrt{R^2-x^2}]; \\ \text{невизначена}, & \text{коли } x\,\overline{\in}\,[-R;R]. \end{cases}$$

Отже, $f_{Y/X}(y/x)\neq f_Y(y)$, щільність $f_{Y/X}(y/x)$ умовного розподілу ймовірностей на множині значень випадкової величини Y залежить від того, якого значення набуває випадкова величина X, тобто випадкові величини X і Y залежні.

Аналогічно стосовно $f_{X/Y}(x/y)$ знаходимо

$$f_{X/Y}(x/y)=\frac{f_{(X,Y)}(x,y)}{f_Y(y)}=$$

$$=\begin{cases} 0, & \text{коли } y\in[-R;R],\ x\,\overline{\in}\,[-\sqrt{R^2-x^2};\sqrt{R^2-x^2}]; \\ \dfrac{1}{2\sqrt{R^2-y^2}}, & \text{коли } y\in[-R;R],\ x\in[-\sqrt{R^2-y^2};\sqrt{R^2-y^2}]; \\ \text{невизначена}, & \text{коли } y\,\overline{\in}\,[-R;R]. \end{cases}$$

Обчислюючи, ймовірність того, що точка (X,Y) лежатиме в області D_1, $(D_1\subset D)$, одержимо:

$$P_{(X,Y)}((X,Y)\in D_1)=\iint\limits_{D_1} f_{(X,Y)}(x,y)dxdy=\frac{1}{\pi R^2}\cdot\pi\left(\frac{R}{2}\right)^2=\frac{1}{4}.$$

Нехай X і Y – відповідно R^s-значна і R^l-значна випадкові величини, $(s\geq 1,\ l\geq 1)$. Якщо $P_Y(Y=y)\neq 0$, то за означенням умовної ймовірності

$$P_{X/Y}(X\in B/Y=y)=\frac{P_{(X,Y)}(\{E\mid\ X(E)\in B,\ Y(E)=y\})}{P_Y(\{E\mid\ Y(E)=y\})}.$$

Нехай існує щільність $f_{(X,Y)}(x,y)$ сумісного розподілу ймовірностей на декартовому добутку множин значень випадкових величин X і Y і стосовно деякого куба $K_\varepsilon\subset R^l$ з центром в точці y і довжиною ребра 2ε

$$P_{(X,Y)}(Y\in K_\varepsilon)>0,\ \varepsilon>0.$$

Тоді

$$P_{X/Y}(X \in B/Y \in K_\varepsilon) = \frac{\int\limits_{K_\varepsilon}\int\limits_{B} f_{(X,Y)}(x,y)\,dxdy}{\int\limits_{K_\varepsilon}\int\limits_{R^s} f_{(X,Y)}(x,y)\,dxdy}.$$

Нехай $\varepsilon \to 0$. Тоді майже за всіх y (щодо міри Лебега в R^l) границя в лівій частині існує і дорівнює

$$\frac{\int\limits_{B} f_{(X,Y)}(x,y)\,dx}{\int\limits_{R^s} f_{(X,Y)}(x,y)\,dx} = \frac{\int\limits_{B} f_{(X,Y)}(x,y)\,dx}{f_Y(y)} = \int\limits_{B} \frac{f_{(X,Y)}(x,y)}{f_Y(y)}\,dx.$$

Цю рівність покладають в основу означення:

Умовною ймовірністю $P_{X/Y}(X \in B/Y = y)$ події $(X \in B)$ за умови (гіпотези) $Y = y$ називають величину

$$P_{X/Y}(X \in B/Y = y) = \frac{\int\limits_{B} f_{(X,Y)}(x,y)\,dx}{f_Y(y)} = \int\limits_{B} \frac{f_{(X,Y)}(x,y)}{f_Y(y)}\,dx = P_{X/Y}(B/y),$$

де

$$f_Y(y) = \int\limits_{R^s} f_{(X,Y)}(x,y)dx.$$

Якщо $f_Y(y) = 0$, то вважають, що $P_{X/Y}(X \in B/Y = y)$ дорівнює нулю. Оскільки

$$P_{X/Y}(X \in B/Y = y) = \int\limits_{B} \frac{f_{(X,Y)}(x,y)}{f_Y(y)}\,dx,$$

то функцію $\dfrac{f_{(X,Y)}(x,y)}{f_Y(y)}$ можна інтерпретувати як щільність розподілу ймовірностей на множині значень R^s-значної випадкової величини X за гіпотези $Y = y$. Таку щільність розподілу позначають $f_{X/Y}(x/y)$. Отже,

$$f_{X/Y}(x/y) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}.$$

Аналогічно

$$f_{Y/X}(y/x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}.$$

Звідси

$$f_{(X,Y)}(x,y) = f_X(x)\,f_{Y/X}(y/x) = f_Y(y)\,f_{X/Y}(x/y).$$

Формулу

$$f_{X/Y}(x/y) = \frac{f_X(x) f_{Y/X}(y/x)}{f_Y(y)} = \frac{f_X(x) f_{Y/X}(y/x)}{\int\limits_{R^S} f_X(x) f_{Y/X}(y/x) dx}$$

називають *формулою Байєса* стосовно апостеріорної щільності розподілу ймовірностей.

Умовна щільність $f_{X/Y}(x, y)$ є щільністю розподілу в R^S:

$$P_{X/Y}(A/y) = \int\limits_A f_{X/Y}(x/y) dx, \quad \int\limits_{R^S} f_{X/Y}(x/y) dx = 1.$$

Нехай (Ω, S, P) – ймовірнісний простір, $D = \{D_1, D_2, ...\}$, $D_i \in S$ – деякий скінченний або зчисленний поділ множини Ω на підмножини D_i, тобто $D_1 + D_2 + ... = \Omega$, $D_i \cap D_j = \varnothing$, коли $i \neq j$, $P(D_i) > 0$.

Нехай $A \in S$ – деяка подія, $P(A/D_i)$ – умовна ймовірність події A відносно події D_i. З набором імовірностей $P(A/D_i)$, $i = 1, 2, \ldots$, можна пов’язати випадкову величину $\sum\limits_i P(A/D_i) I_{D_i}(E)$ зі значеннями $P(A/D_i)$, коли $E \in D_i$, яких вона набуває з ймовірністю $P(D_i)$. Цю випадкову величину, пов’язану з поділом D, позначають $P(A/D)(E)$ і називають *умовною ймовірністю події A відносно поділу D*.

Якщо D – тривіальний поділ, що складається з однієї множини Ω, то

$$P(A/\Omega) = P(A).$$

Обчислюючи математичне сподівання випадкової величини $P(A/D)$, дістанемо формулу повної ймовірності:

$$M[P(A/D)] = \sum_i P(A/D_i) P(D_i) = P(A).$$

Нехай $Y = Y(E)$ – дискретна випадкова величина із множиною значень $y_1, y_2, ...$,

$$Y(E) = \sum y_j I_{Y^{-1}(y_j)}(E).$$

Поділ $D_Y = \{Y^{-1}(y_1), Y^{-1}(y_2), ...\}$ називається *поділом, що породжується за випадковою величиною Y*.

Умовну ймовірність $P(A/D_Y)$ називають *умовною ймовірністю події A щодо величини Y і позначають $P(A/Y)$*. Під $P(A/Y = y_i)$ розуміють $P(A/Y^{-1}(y_i))$.

Приклад 4.14.7. На множині $\Omega_{(X,Y)} = \bigcup_{i=1}^{2}\bigcup_{j=1}^{2} H_{i,j}$, де $H_{i,j} = (a_{j-1}; a_j] \times (c_{i-1}; c_i]$, $i \in \overline{1,2}$, $j \in \overline{1,2}$, $(a_0; a_1] = (0; \frac{1}{2}]$, $(a_1; a_2] = (\frac{1}{2}; 1]$, $(c_0; c_1] = (0; \frac{1}{2}]$, $(c_1; c_2] = (\frac{1}{2}; 1]$, задано двохвимірні поінтервальні розподіли ймовірностей за двохвимірними інтервалами H_{ij}:

а) $P_{(X,Y)}(H_{ij}) = 0.25$, коли $(i, j) \in \{1,2\} \times \{1,2\}$;

б) $\mathrm{P}_{(X,Y)}(H_{ij}) = 0.50$, коли $(i, j) \in \{(1,1)\} \bigcup \{(2,2)\}$,

$\mathrm{P}_{(X,Y)}(H_{ij}) = 0$, коли $(i, j) \in \{(1,2)\} \bigcup \{(2,1)\}$,

А як простори подій розглядаються

1. $S_{(X,Y)} = \mathscr{B}(R^2)$.

2. $S_{(X,Y)} = \{A \mid A = \bigcup_{i \in I}\bigcup_{j \in J} H_{ij},\ I \subset \{1,2\},\ J \subset \{1,2\}\}$.

За обох заданих розподілів а) і б) і обох просторів подій:

3. Визначити щільність $f_{(X,Y)}(x, y)$ поінтервального розподілу ймовірностей за двохвимірними інтервалами H_{ij} та побудувати її графік.

4. Визначити відповідні одновимірні поінтервальні розподіли ймовірностей за інтервалами $(a_{j-1}; a_j]$, $j \in \overline{1,2}$, $(c_{i-1}; c_i]$, $i \in \overline{1,2}$:

$P_X((a_{j-1}; a_j])$, $j \in \overline{1,2}$, $P_Y((c_{i-1}; c_i])$, $i \in \overline{1,2}$.

5. Перевірити, чи мають місце рівності

$P_{(X,Y)}(H_{ij}) = P_X((a_{j-1}; a_j])\ P_Y((c_{i-1}; c_i])$ за всіх $i \in \overline{1,2}$, $j \in \overline{1,2}$;

6. Визначити поінтервальні розподіли умовних ймовірностей

$P_X((a_{j-1}; a_j]/(c_{i-1}; c_i])$, $P_Y((c_{i-1}; c_i]/(a_{j-1}; a_j])$

за всіх $i \in \overline{1,2}$, $j \in \overline{1,2}$.

7. Визначити щільності $f_X(x)$ та $f_Y(y)$ відповідних поінтервальних розподілів ймовірностей за одновимірними інтервалами $(a_{j-1}; a_j]$, $j \in \overline{1,2}$, та $(c_{i-1}; c_i]$, $i \in \overline{1,2}$.

8. Перевірити, чи має місце рівність $f_{(X,Y)}(x, y) = f_X(x) f_Y(y)$ за всіх $(x, y) \in R^2$, $x \in R^1$, $y \in R^1$.

9. Визначити щільності $f_X(x/(c_{i-1}; c_i])$, $i \in \overline{1,2}$, та $f_Y(y/(a_{j-1}; a_j])$, $j \in \overline{1,2}$, умовних поінтервальних розподілів ймовірностей за інтервалами відповідно $(a_{j-1}; a_j]$, $j \in \overline{1,2}$, та $(c_{i-1}; c_i]$, $i \in \overline{1,2}$.

10.Визначити $F_{(X,Y)}(x,y)$ – функцію двохвимірного поінтервального розподілу ймовірностей за двохвимірними інтервалами H_{ij}, $i \in \overline{1,2}$, $j \in \overline{1,2}$;

11.Визначити $F_X(x)$ та $F_Y(y)$ – функції відповідних одновимірних поінтервальних розподілів ймовірностей за інтервалами $(a_{j-1}; a_j]$, $j \in \overline{1,2}$, та $(c_{i-1}; c_i]$, $i \in \overline{1,2}$ відповідно.

12.Перевірити, чи має місце рівність $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$ за всіх $(x,y) \in R^2$, $x \in R^1$, $y \in R^1$.

13.Обчислити числові характеристики заданого розподілу ймовірностей: $M[X]$, $M[Y]$, $D[X]$, $D[Y]$, $K(X,Y)$.

Як зміняться відповіді, коли як простір елементарних подій розглядати $\tilde{\Omega}_{(X,Y)} = R^2$ замість $\Omega_{(X,Y)} = \bigcup_{i=1}^{2}\bigcup_{j=1}^{2} H_{ij}$, а як простір подій розглядати $\tilde{S}_{(X,Y)} = \mathscr{B}(R^2)$ замість $S_{(X,Y)} = \{A \mid A = \bigcup_{i \in I}\bigcup_{j \in J} H_{ij},\quad I \subset \{1,2\},\ J \subset \{1,2\}\}$, а стосовно відповідних одновимірних розподілів ймовірностей розглядати простори елеменатрних подій $\tilde{\Omega}_X = R^1$, $\tilde{\Omega}_Y = R^1$ замість $\Omega_X = (0;1]$, $\Omega_Y = (0;1]$, а як простори подій розглядати $\tilde{S}_X = \mathscr{B}(R^1)$, $\tilde{S}_Y = \mathscr{B}(R^1)$ замість $S_X = \{A \mid A = \bigcup_{j \in J}(a_{j-1}; a_j],\ J \subset \{1,2\}\}$, $S_Y = \{A \mid A = \bigcup_{i \in I}(c_{i-1}; c_i],\ I \subset \{1,2\}\}$ відповідно.

Розв'язування

Стосовно усередненої щільності розподілу ймовірностей на множині $\Omega_{(X,Y)}$ одержуємо:

1. а) $$f_{(X,Y)}(x,y) = \begin{cases} 1, \text{ коли } (x,y) \in H_{i,j},\ i \in \overline{1,2},\ j \in \overline{1,2}, \\ 0, \text{ коли } (x,y) \overline{\in} H_{i,j},\ i \in \overline{1,2},\ j \in \overline{1,2}, \end{cases}$$

Графік усередненої щільності $f_{(X,Y)}(x,y)$ розподілу ймовірностей за двохвимірними інтервалами H_{ij}, $i \in \overline{1,2}$, $j \in \overline{1,2}$, стосовно випадку а) подано на Рис. 4.14.6.1 а).

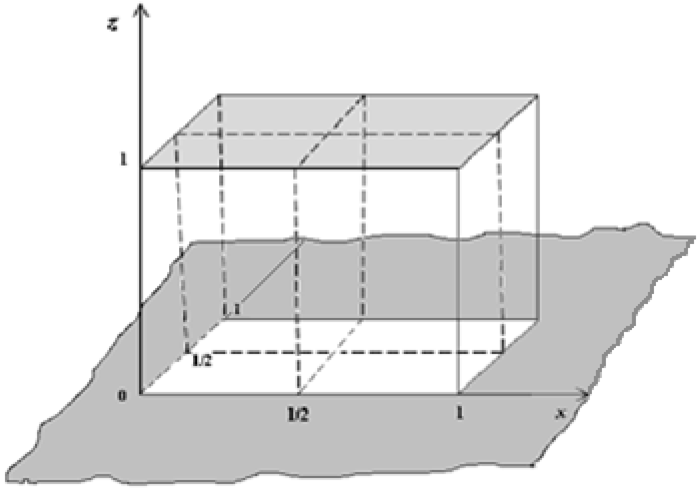


Рис. 4.14.6.1. а)

1. б) $f_{(X,Y)}(x,y)=\begin{cases}2, \text{ коли } (x,y)\in H_{1,1}\cup H_{2,2},\\ 0, \text{ коли } (x,y)\overline{\in}\, H_{1,1}\cup H_{2,2};\end{cases}$

Графік усередненої щільності $f_{(X,Y)}(x,y)$ розподілу ймовірностей за двохвимірними інтервалами H_{ij} , $i\in\overline{1,2}$, $j\in\overline{1,2}$, стосовно випадку б) подано на Рис. 4.14.6.1 б).

2. а) $P_X((a_{j-1};a_j])=0.50,\ \ j\in\overline{1,2};\ \ P_Y((c_{i-1};c_i])=0.50,\ \ i\in\overline{1,2};$

2. б) $P_X((a_{j-1};a_j])=0.50,\ \ j\in\overline{1,2};\ \ P_Y((c_{i-1};c_i])=0.50,\ \ i\in\overline{1,2};$

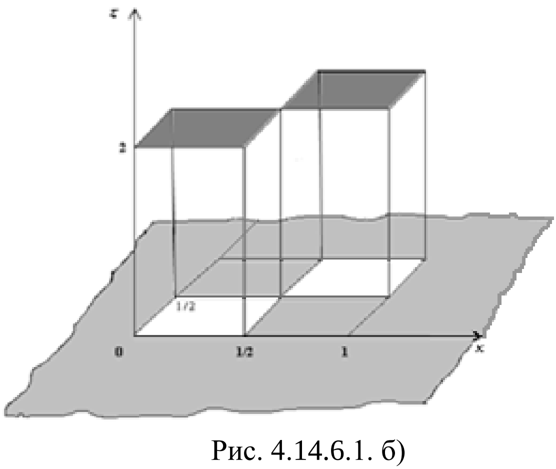


Рис. 4.14.6.1. б)

3. а) $P_X((a_{j-1};a_j])P_Y((c_{i-1};c_i])=P_{(X,Y)}((a_{j-1};a_j]\times(c_{i-1};c_i])$ за всіх $i\in\overline{1,2},\ j\in\overline{1,2}$;

3. б) $P_X((a_{j-1};a_j])P_Y((c_{i-1};c_i])\neq P_{(X,Y)}((a_{j-1};a_j]\times(c_{i-1};c_i])$ за всіх $i\in\overline{1,2},\ j\in\overline{1,2}$,

Наприклад

$P_X((a_{j-1};a_j])P_Y((c_{i-1};c_i])=0.50\cdot 0.50\neq P_{(X,Y)}((a_0;a_1]\times(c_1;c_2])=0$,

$(a_0;a_1]=(0;\frac{1}{2}],\ (c_1;c_2]=(\frac{1}{2};1]$.

4. а) $P_X((a_{j-1};a_j]/(c_{i-1};c_i])=\dfrac{P_{(X,Y)}((a_{j-1};a_j]\times(c_{i-1};c_i])}{P_Y((c_{i-1};c_i])}=\dfrac{1}{2}$ за всіх $i\in\overline{1,2},\ j\in\overline{1,2}$;

$$P_Y((c_{i-1};c_i]/(a_{j-1};a_j])=\frac{P_{(X,Y)}((a_{j-1};a_j]\times(c_{i-1};c_i])}{P_Y((a_{j-1};a_j])}=\frac{1}{2}$$

за всіх $i\in\overline{1,2},\ j\in\overline{1,2}$;

4. б) $P_X((a_0;a_1])/(c_0;c_1])=1$, $P_X((a_0;a_1])/(c_1;c_2])=0$;
$P_X((a_1;a_2])/(c_0;c_1])=0$, $P_X((a_1;a_2])/(c_1;c_2])=1$;
$P_Y((c_0;c_1]/(a_0;a_1]))=1$, $P_Y((c_0;c_1]/(a_1;a_2])=0$;
$P_Y((c_1;c_2]/(a_0;a_1]))=0$, $P_Y((c_1;c_2]/(a_1;a_2])=1$;

5. а) $f_X(x)=\begin{cases}1, & \text{коли } x\in(0;1],\\ 0, & \text{коли } x\,\overline{\in}\,(0;1];\end{cases}$ $f_Y(y)=\begin{cases}1, & \text{коли } y\in(0;1],\\ 0, & \text{коли } y\,\overline{\in}\,(0;1];\end{cases}$

5. б) $f_X(x)=\begin{cases}1, & \text{коли } x\in(0;1],\\ 0, & \text{коли } x\,\overline{\in}\,(0;1];\end{cases}$ $f_Y(y)=\begin{cases}1, & \text{коли } y\in(0;1],\\ 0, & \text{коли } y\,\overline{\in}\,(0;1];\end{cases}$

6. а) $f_X(x)f_Y(y)=f_{(X,Y)}(x,y)$ за всіх $(x,y)\in R^2$, $x\in R^1, y\in R^1$;

6. б) $f_X(x)f_Y(y)\neq f_{(X,Y)}(x,y)$ за всіх $(x,y)\in R^2$, $x\in R^1, y\in R^1$;

Наприклад: $x\in(0;\frac{1}{2}], y\in(\frac{1}{2};1]$, $f_{X,Y}(x,y)=0\neq f_X(x)f_Y(y)=1\cdot 1$.

7. а)

$$f_X(x/y\in(c_{i-1};c_i])=\begin{cases}1, & \text{коли } x\in(a_{j-1};a_j],\ \ j\in\overline{1,2},\ \ i\in\overline{1,2}\\ 0, & \text{коли } x\,\overline{\in}\,(a_{j-1};a_j],\ \ j\in\overline{1,2},\ \ i\in\overline{1,2}\end{cases}$$

$$f_Y(y/x\in(a_{j-1};a_j])=\begin{cases}1, & \text{коли } x\in(c_{i-1};c_i],\ \ i\in\overline{1,2},\ \ j\in\overline{1,2}\\ 0, & \text{коли } x\,\overline{\in}\,(c_{i-1};c_i],\ \ i\in\overline{1,2},\ \ j\in\overline{1,2}\end{cases}$$

7. б) $f_X(x/y\in(c_{i-1};c_i])=\begin{cases}2, & \text{коли } x\in(a_0;a_1],\ \ y\in(c_0;c_1],\\ 0, & \text{коли } x\,\overline{\in}\,(a_0;a_1],\ \ y\in(c_0;c_1],\end{cases}$

$$f_X(x/y\in(c_{i-1};c_i])=\begin{cases}2, & \text{коли } x\in(a_1;a_2],\ \ y\in(c_1;c_2],\\ 0, & \text{коли } x\,\overline{\in}\,(a_1;a_2],\ \ y\in(c_1;c_2],\end{cases}$$

$$f_Y(y/x\in(a_{j-1};a_j])=\begin{cases}2, \text{ коли } y\in(c_0;c_1],\ x\in(a_0;a_1],\\ 0, \text{ коли } y\,\overline{\in}\,(c_0;c_1],\ x\in(a_0;a_1],\end{cases}$$

$$f_Y(y/x\in(a_{j-1};a_j])=\begin{cases}2, \text{ коли } y\in(c_1;c_2],\ x\in(a_1;a_2],\\ 0, \text{ коли } y\,\overline{\in}\,(c_1;c_2],\ x\in(a_1;a_2];\end{cases}$$

8. Оскільки в разі $S_{(X,Y)}=\{A\,|\,A=\bigcup\limits_{i\in I}\bigcup\limits_{j\in J}H_{ij},$ $I\subset\{1,2\},\ J\subset\{1,2\}\}$, значення функції $F_{(X,Y)}(x,y)$ будуть визначені лише в вершинах квадратів H_{ij}, а у внутрішніх точках квадратів H_{ij} значення функції $F(x,y)$ будуть невизначені, оскільки за довільних $(x,y)\in R^2$ не завжди виконуватимуться вимоги $((-\infty;x)\times(-\infty;y))\bigcap\Omega\in S_{(X,Y)}$, і в такому разі визначити $F_{(X,Y)}(x,y)=P(((-\infty;x)\times(-\infty;y))\bigcap\Omega)$ буде неможливо, бо міра P визначена лише на множинах із $S_{(X,Y)}$, довизначимо функцію $F_{(X,Y)}(x,y)$, за всіх $(x,y)\in R^2$, поклавши $F(x,y)=P(G_*)$, $G_*\in S_{(X,Y)}$, $G_*=\bigcup\limits_{\cup H_{ij}\subset G\cap\Omega}\bigcup H_{ij}$, $G=(-\infty;x)\times(-\infty;y)$, (див. §2.2 - §2.6). В такому разі одержимо:

а) $$F_{(X,Y)}(x,y)=\begin{cases}0, \text{ коли } x\le\frac{1}{2} \text{ або } y\le\frac{1}{2},\\ 0.25, \text{ коли } \frac{1}{2}<x\le1 \text{ і } \frac{1}{2}<y\le1,\\ 0.50, \text{ коли } \frac{1}{2}<x\le1 \text{ і } 1<y \text{ або}\\ \qquad\qquad \frac{1}{2}<y\le1 \text{ і } 1<x,\\ 1, \text{ коли } 1\le x \text{ і } 1\le y;\end{cases}$$

Зауважимо, що значення функції $F_{(X,Y)}(x,y)=P_{(X,Y)}((-\infty;x)\times(-\infty;y))$, $(x,y)\in R^2$, визначені лише за таких $(x,y)\in R^2$, за яких $(-\infty;x)\times(-\infty;y))\bigcap\Omega\in S$. Зокрема, коли $S=\mathscr{B}(R^2)$, тоді значення $F(x,y)$ визначені за всіх $(x,y)\in R^2$.

Графік функції $F_{(X,Y)}(x,y)$ стосовно випадку а) подано на Рис. 4.14.7.2.а).

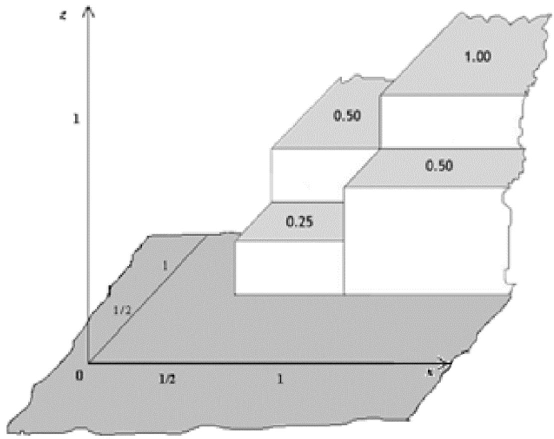


Рис 4.14.7.2.а)

$$8.\text{ б) } F_{(X,Y)}(x,y)=\begin{cases}0, \text{ коли } x\le\dfrac{1}{2} \text{ або } y\le\dfrac{1}{2},\\ 0.50, \text{ коли } \dfrac{1}{2}<x\le 1 \text{ і } \dfrac{1}{2}<y, \text{ або}\\ \qquad \text{коли } \dfrac{1}{2}<x \text{ і } \dfrac{1}{2}<y\le 1,\\ 1, \text{ коли } 1<x \text{ і } 1<y.\end{cases}$$

Графік функції $F_{(X,Y)}(x,y)$ стосовно випадку б) подано на Рис. 4.14.7.2.б).

Стосовно відповідних одновимірних поінтервальних розподілів за інтервалами $(0;\frac{1}{2}]$, $(\frac{1}{2};1]$ одержуємо

$$9.\text{ а) } F_X(x)=\begin{cases}0, \text{ коли } x\le\dfrac{1}{2},\\ 0.50, \text{ коли } \dfrac{1}{2}<x\le 1,\\ 1, \text{ коли } 1<x,\end{cases}$$

Рис. 4.14.7.2.б)

$$F_Y(y)=\begin{cases}0, \text{ коли } y\le\dfrac{1}{2},\\ 0.50, \text{ коли } \dfrac{1}{2}<y\le 1,\\ 1, \text{ коли } 1<y,\end{cases}$$

9. б) $$F_X(x)=\begin{cases}0, \text{ коли } x\le\dfrac{1}{2},\\ 0.50, \text{ коли } \dfrac{1}{2}<x\le 1,\\ 1, \text{ коли } 1<x,\end{cases}$$

$$F_Y(y)=\begin{cases}0, \text{ коли } y\le\dfrac{1}{2},\\ 0.50, \text{ коли } \dfrac{1}{2}<y\le 1,\\ 1, \text{ коли } 1<y,\end{cases}$$

10. а) $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$ за всіх $(x,y) \in R^2$, $x \in R^1$, $y \in R^1$;

10. б) $F_{(X,Y)}(x,y) \neq F_X(x)F_Y(y)$ за всіх $(x,y) \in R^2$, $x \in R^1$, $y \in R^1$;

Наприклад, коли $(x,y) \in (\frac{1}{2}, 1] \times (\frac{1}{2}, 1]$, тоді

$$F_{(X,Y)}(x,y) = 0.50 \neq 0.25 = F_X(x)F_Y(y), \ \ x \in (\frac{1}{2}; 1], \ y \in (\frac{1}{2}; 1];$$

Враховуючи вирази функцій $f_X(x)$, $f_Y(y)$, $f_{(X,Y)}(x,y)$, одержуємо:

11. а) $M[X] = \int\limits_{-\infty}^{\infty} x\, f_X(x)dx = \int\limits_0^1 x \cdot 1 \cdot dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2},$

$$M[Y] = \int\limits_{-\infty}^{\infty} y\, f_Y(y)dy = \int\limits_0^1 y \cdot 1 \cdot dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2},$$

$$D[X] = \int\limits_{-\infty}^{\infty} (x - M[X])^2 f_X(x)dx = \int\limits_0^1 (x - \frac{1}{2})^2 \cdot 1 \cdot dx = \frac{1}{3}(x - \frac{1}{2})^3\Big|_0^1 = \frac{1}{3} \cdot \frac{2}{8} = \frac{1}{12},$$

$$D[Y] = \int\limits_{-\infty}^{\infty} (y - M[Y])^2 f_Y(y)dy = \int\limits_0^1 (y - \frac{1}{2})^2 \cdot 1 \cdot dy = \frac{1}{3}(y - \frac{1}{2})^3\Big|_0^1 = \frac{1}{3} \cdot \frac{2}{8} = \frac{1}{12},$$

$$K[X,Y] = \iint\limits_{R^2} (x - M[X])(y - M[Y]) f_{(X,Y)}(x,y)dxdy = \int\limits_0^1\int\limits_0^1 (x - \frac{1}{2})(y - \frac{1}{2}) \cdot 1 \cdot dxdy =$$

$$= \frac{1}{2}(x - \frac{1}{2})^2\Big|_0^1 \cdot \frac{1}{2}(y - \frac{1}{2})^2\Big|_0^1 = \frac{1}{4}((\frac{1}{2})^2 - (-\frac{1}{2})^2)((\frac{1}{2})^2 - (-\frac{1}{2})^2) = 0.$$

11. б) $M[X]$, $M[Y]$, $D[X]$, $D[Y]$ будуть такі самі, як і в 11.а).

Обчислюючи $K[X,Y]$, одержуємо:

$$K[X,Y] = \int\limits_0^{\frac{1}{2}}\int\limits_0^{\frac{1}{2}} (x - \frac{1}{2})(y - \frac{1}{2})2dxdy + \int\limits_{\frac{1}{2}}^1\int\limits_{\frac{1}{2}}^1 (x - \frac{1}{2})(y - \frac{1}{2})2dxdy =$$

$$= 2 \cdot \frac{1}{2}(x - \frac{1}{2})^2\Big|_0^{\frac{1}{2}} \cdot \frac{1}{2}(y - \frac{1}{2})^2\Big|_0^{\frac{1}{2}} + 2 \cdot \frac{1}{2}(x - \frac{1}{2})^2\Big|_0^{\frac{1}{2}} \cdot \frac{1}{2}(y - \frac{1}{2})^2\Big|_0^{\frac{1}{2}} =$$

$$= 2 \cdot \frac{1}{2}(0 - (-\frac{1}{2})^2) \cdot \frac{1}{2}(0 - (-\frac{1}{2})^2) + 2 \cdot \frac{1}{2}((\frac{1}{2})^2 - (0)^2) \cdot \frac{1}{2}((\frac{1}{2})^2 - (0)^2) =$$

$$= 2\cdot\frac{1}{2}(-\frac{1}{4})\cdot\frac{1}{2}\cdot(-\frac{1}{4}) + 2\cdot\frac{1}{2}(\frac{1}{4})\cdot\frac{1}{2}\cdot(\frac{1}{4}) = \frac{1}{32}+\frac{1}{32}=\frac{1}{16}.$$

Таким чином у випадку б) випадкові величини X і Y залежні.

Коли як простір елементарних подій розглядати $\tilde{\Omega}_{(X,Y)} = R^2$, а як простір подій $\tilde{S}_{(X,Y)} = \mathscr{B}(R^2)$, тоді зміняться вирази стосовно $F_{(X,Y)}(x,y)$, а всі інші характеристики розподілів залишаться без змін.

У випадку а) одержимо:

$$F_{(X,Y)}(x,y) = \begin{cases} 0, \text{ коли } x \le 0 \text{ або } y \le 0, \\ xy, \text{ коли } x \in (0;1] \text{ і } y \in (0;1], \\ x, \text{ коли } x \in (0;1] \text{ і } 1 < y, \\ \mathrm{y}, \text{ коли } 1 < \mathrm{x} \text{ і } \mathrm{y} \in (0;1], \\ 1, \text{ коли } 1 < \mathrm{x} \text{ і } 1 < \mathrm{y}. \end{cases}$$

У випадку б) одержимо:

$$F_{(X,Y)}(x,y) = \begin{cases} 0, \text{ коли } x \le 0 \text{ або } y \le 0, \\ 2xy, \text{ коли } x \in (0;\frac{1}{2}] \text{ і } y \in (0;\frac{1}{2}], \\ 2x, \text{ коли } x \in (0;\frac{1}{2}] \text{ і } \frac{1}{2} < y, \\ 2y, \text{ коли } \frac{1}{2} < x \text{ і } y \in (0;\frac{1}{2}], \\ \frac{1}{2} + 2(x-\frac{1}{2})(y-\frac{1}{2}), \text{ коли } x \in (\frac{1}{2};1] \text{ і } y \in (\frac{1}{2};1], \\ \frac{1}{2} + 2(x-\frac{1}{2})\frac{1}{2}, \text{ коли } x \in (\frac{1}{2};1] \text{ і } 1 < y, \\ \frac{1}{2} + 2(y-\frac{1}{2})\cdot\frac{1}{2}, \text{ коли } 1 < x \text{ і } y \in (\frac{1}{2};1], \\ 1, \text{ коли } 1 < x \text{ і } 1 < y. \end{cases}$$

Вправи для самостійного виконання

4.14.1 Перевірити, чи правильні твердження:

1. За будь яких дискретних випадкових величин X та Y існує умовний розподіл ймовірностей на множині значень однієї випадкової величини за умови, що інша набуває фіксованого значення.

2. Величина $P_{X/Y}(\{x_j\}/\{y_j\})$ завжди визначена однозначно.

3. Завжди $P_{(X,Y)}(\{x_j\}\{y_j\}) = P_X(\{x_j\})P_Y(\{y_j\})$.

4. Випадкові величини X та Y незалежні тоді і тільки тоді, коли $P_{X/Y}(\{x_j\}/\{y_i\}) = P_X(\{x_j\})$ за будь-якого y_i.

5. За будь-яких випадкових величин X та Y з абсолютно неперервними розподілами ймовірностей на множинах їх значень спільний розподіл ймовірностей також абсолютно неперервний, причому

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy, \; f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dx .$$

6. За будь-яких випадкових величин X та Y з абсолютно неперервними розподілами ймовірностей на множинах їх значень існує абсолютно неперервний умовний розподіл ймовірностей на множині значень випадкової величини X за умови, що Y набуває значення y_0, $f_{X/Y}(x/y_0) = \dfrac{f_{(X,Y)}(x,y_0)}{f_Y(y_0)}$.

7. Стосовно випадкових величин X та Y з абсолютно неперервними розподілами ймовірностей на множинах їх значень $F_{(X,Y)}(x,y) = F_X(x) \cdot F_Y(y)$ за всіх (x,y) тоді і тільки тоді, коли $f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$ за майже всіх $x \in R$ та $y \in R$.

8. З формули Байєса стосовно апостеріорної щільності розподілу ймовірностей випливає формула Байєса стосовно апостеріорного розподілу ймовірностей: $P_{X/Y}(B/y) = \dfrac{\int_B f_X(x)\, f_{Y/X}(y/x)dx}{\int_{R^s} f_X(x)\, f_{Y/X}(y/x)dx}$, де $B \subset R^s$.

9. За будь-якого ймовірнісного простору (Ω, S, P) існує єдиний не більш ніж зчисленний поділ простору Ω на події $D_1, D_2, \ldots$.

10. За кожним поділом $D = \{D_1, D_2, \ldots\}$ простору Ω визначається єдина випадкова величина $X(E)$, стосовно якої $X(E) = P(A/D_i)$, коли $E \in D_i$.

11. Умовна ймовірність $P(A/D)(E)$ є простою випадковою величиною.

12. $P(A/D) = P(A)$ тоді і тільки тоді, коли в поділі D міститься лише одна подія.

4.15. Числові характеристики умовних розподілів ймовірностей

Нехай X – дискретна випадкова величина, що набуває значень $x_1, x_2, \ldots$, тобто

$$X(E) = \sum_i x_i I_{X^{-1}(x_i)}(E).$$

Нехай $D = \{D_1, D_2, \ldots\}$ – деякий поділ множини Ω на підмножини $D_1, D_2, \ldots$.

Умовним математичним сподіванням випадкової величини X *відносно поділу* D називають випадкову величину

$$M[X/D](E) = \sum_i x_i P(X^{-1}(x_i)/D)(E),$$

де

$$P(X^{-1}(x_i)/D)(E) = \sum_j P(X^{-1}(x_i)/D_j) I_{D_j}(E).$$

Якщо $E \in D_j$, тоді

$$M[X/D_j](E) = \sum_i x_i P(X^{-1}(x_i)/D_j) = \sum_i x_i \frac{P(X^{-1}(x_i)D_j)}{P(D_j)} = \frac{M[XI_{D_j}]}{P(D_j)},$$

тобто за всіх $E \in D_j$ випадкова величина $M[X/D](E)$ набуває одного і того самого значення $\sum_i x_i P(X^{-1}(x_i)/D_j)$ з ймовірністю $P(D_j)$.

Таким чином, до означення умовного математичного сподівання $M[X/D]$ можна підійти й так. Спочатку означити

$$M[X/D_j] = \sum_i x_i P(X^{-1}(x_i)/D_j),$$

а потім покласти

$$M[X/D](E) = \sum_j M[X/D_j] I_{D_j}(E).$$

Безпосередньо з означення випливають такі *властивості умовного математичного сподівання дискретної випадкової величини*.

1. $M[aX + bY/D] = aM[X/D] + bM[Y/D]$, $a - \text{const}$, $b - \text{const}$.
2. $M[X/\Omega] = M[X]$.
3. $M[c/D] = c$, $c - \text{const}$.
4. Якщо $X(E) = I_A(E)$, то $M[X/D] = P(A/D)$.

5. $M[M[X/D]] = M[X]$ (узагальнення формули повної ймовірності).

Справді,

$$M[M[X/D]] = M\left[\sum_i x_i P(X^{-1}(x_i)/D)\right] = \sum_i x_i M[P(X^{-1}(x_i)/D)] =$$

$$= \sum_i x_i P(X^{-1}(x_i)) = M[X].$$

Нехай $D = \{D_1, D_2, ...\}$ – деякий поділ простору Ω, $Y = Y(E)$ – дискретна випадкова величина, а D_Y – поділ, що породжується за випадковою величиною Y.

Випадкова величина Y називається *вимірною відносно поділу* D, якщо $D_Y \ll D$ (D дрібніше за D_Y), тобто $Y = Y(E)$ можна подати у вигляді

$$Y(E) = \sum_j y_j I_{Y^{-1}(y_j)}(E).$$

Іншими словами, Y вимірна відносно поділу D (D-вимірна) тоді і тільки тоді, коли вона набуває сталих значень на підмножинах D_j. Якщо, наприклад, $D = \{\Omega\}$ – тривіальний поділ, то Y є D-вимірною тільки тоді, коли $Y = \text{const}$.

Будь-яка дискретна випадкова величина Y вимірна відносно поділу D_Y. Якщо X і Y дискретні випадкові величини і поділ D_Y породжується за випадковою величиною Y, то $M[X/D_Y]$ називають *умовним математичним сподіванням* X *відносно* Y *і позначають* $M[X/Y]$.

6. Якщо $X \geq 0$, то $M[X/Y] \geq 0$. Зокрема, якщо $X_1 < X_2$, то $M[X_1/Y] \leq M[X_2/Y]$.

7. Якщо $X_1 < X_2 < ... X_n ...$ і $M[\lim\limits_{n\to\infty} X_n] < \infty$, то з імовірністю 1

$$\lim_{n\to\infty} M[X_n/Y] = M[\lim_{n\to\infty} X_n/Y].$$

Якщо X і Y незалежні дискретні випадкові величини, то, як випливає з означення, $M[X/Y]$ є сталою випадковою величиною:

$$M[X/Y] = M[X].$$

Якщо D – довільний поділ простору $\mathbf{\Omega}$ й дискретна випадкова величина X не залежить від поділу D, тобто X і I_{D_i} незалежні за будь-якого $D_i \in D$, то $M[X/D] = M[X]$.

Якщо Ω скінченна множина, то кожному поділу $D = \{D_1, D_2, ...\}$ відповідає алгебра $\alpha(D)$ підмножин Ω, і навпаки, будь-яка алгебра підмножин скінченного простору Ω

породжується за деяким поділом D, ($\mathscr{B} = \alpha(D)$). Тим самим між алгебрами підмножин і поділами скінченного простору Ω існує взаємно однозначна відповідність. У випадку скінченних множин поняття алгебри і σ-алгебри збігаються. Тому умовне математичне сподівання $M[X/D]$ можна означити як умовне математичне сподівання випадкової величини X відносно σ-алгебри $\sigma(D)$. Стосовно скінченних просторів $M[X / \mathscr{B}] = M[X / D]$ за означенням.

Зазначимо, що оскільки $P(A/D)$ – випадкова величина із значеннями $P(A/D_i)$, $(0 \le P(A/D_i) = \dfrac{P(AD_i)}{P(D_i)} \le 1,\ P(D_i) > 0)$, яких вона набуває з імовірностями $P(D_i)$, то ймовірність $P(P(A/D) \in B)$, де B – деяка борелівська множина з $[0;1]$, можна визначити так:

$$P(P(A/D) \in B) = \sum_{P(A/D_i) \in B} P(D_i).$$

Аналогічно, оскільки $M[X/D]$ – випадкова величина із значеннями $M[X/D_j]$, яких вона набуває з імовірностями $P(D_j)$, то

$$P(M[X/D] \in B) = \sum_{M[X/D_j] \in B} P(D_j),$$

а також за будь-якого $B \in \alpha(D)$

$$\sum_{D_j \subset B} M[X/D_j]\, P(D_j) = \sum_{D_j \subset B} \sum_i x_i P[X^{-1}(x_i)/D_j]\, P(D_j) =$$

$$= \sum_i x_i \sum_{D_j \subset B} P(X^{-1}(x_i)/D_j)\, P(D_j) = \sum_i x_i \sum_{D_j \subset B} P(X^{-1}(x_i) D_j) =$$

$$= \sum_i x_i P(X^{-1}(x_i) B) = M[X \cdot I_B] = M[M[X/D] I_B].$$

Зокрема,

$$M[X I_\Omega] = M[M[X/D] I_\Omega] = M[M[X/D]] = M[X].$$

Випадкова величина $P(A/D)$ вимірна відносно σ-алгебри $Q = \sigma(D)$, тому її також позначають $P(A/Q)$. Якщо X – довільна випадкова величина, стосовно якої визначено $M[X]$, $A \in S$ – випадкова подія така, що $P(A) > 0$, то величину

$$M[X/A] = \frac{M[X I_A]}{P(A)},\ P(A) > 0,$$

називають *умовним математичним сподіванням випадкової величини* X *відносно події* A.

Нехай (Ω, S, P) – імовірнісний простір, Q – деяка σ-алгебра, $Q \subseteq S$, X – невід'ємна випадкова величина. Невід'ємна випадкова величина $M\,[X/Q](E)$ називається *умовним математичним сподіванням невід'ємної випадкової величини* X *відносно* σ*-алгебри* Q, якщо:

1) $M[X/Q]$ є Q-вимірною;

2) за будь-якого $A \in Q$ виконується рівність

$$\int_A XdP = \int_A M[X/Q]dP .$$

Якщо X – довільна випадкова величина, то

$$M[X/Q] = M[X^+/Q] - M[X^-/Q],$$

причому на множині нульової міри, коли $M[X^+/Q] = M[X^-/Q] = \infty$, різниця $M[X^+/Q] - M[X^-/Q]$ визначається довільно, наприклад може дорівнювати нулю.

Оскільки $L(A) = \int_A XdP$, $A \in Q$, – міра на (Ω, Q), абсолютно неперервна відносно міри P (яка розглядається на (Ω, Q), $Q \subseteq S$), то за теоремою Радона-Никодима існує така невід'ємна Q-вимірна випадкова величина $M[X/Q]$, що

$$L(A) = \int_A M[X/Q]dP .$$

Таким чином $M[X/Q]$ існує і

$$\int_A XdP = \int_A M[X/Q]dP .$$

Слід звернути увагу на те, що $M[X/Q]$ визначається з точністю до множини міри нуль (див. теорему Радона-Никодима).

Якщо $Q = \{\varnothing, \Omega\}$ – тривіальна σ-алгебра, то

$$M[X/Q] = M[X].$$

В загальному випадку не можна покласти $M[X/Q] = M[X]$, оскільки випадкова величина X не обов'язково повинна бути Q-вимірною.

Умовне математичне сподівання $M[I_B/Q]$ випадкової величини I_B, $B \in S$, називається *умовною ймовірністю події* B *відносно* σ-алгебри Q, $Q \subseteq S$.

Таким чином, умовна ймовірність $P(B/Q)$, $B \in S$, є такою випадковою величиною, що

1) $P(B/Q)$ є Q-вимірною;

2) за довільного $A \in Q$ виконується рівність

$$P(A \bigcap B) = \int_A P(B/Q)\, dP .$$

Означення. Нехай X – випадкова величина і Q_Y – σ-алгебра, породжена за деяким випадковим елементом Y. Тоді $M[X/Q_Y]$, якщо воно існує, називається *умовним математичним сподіванням* X *відносно* Y і позначається $M[X/Y]$.

Умовна ймовірність події B відносно σ-алгебри Q_Y позначається $P(B/Y)$ і називається *умовною ймовірністю події* B відносно Y.

Оскільки $M[X/Y](E)$ Q_Y-вимірна, то знайдеться така борелівська функція $m(y)$ із значеннями в R^1, що за всіх $E \in \Omega$

$$m(Y(E)) = M[X/Y](E).$$

Функцію $m(y)$ позначають $M[X/Y = y]$ і називають *умовним математичним сподіванням випадкової величини* X *відносно події* $Y = y$ або *умовним математичним сподіванням випадкової величини* X *за умови* $Y = y$.

Таким чином,

$$\int_A X dP = \int_A M[X/Y]\, dP = \int_A m[Y]\, dP, \ A \in Q_Y .$$

Тому за теоремою про заміну змінних під знаком інтеграла Лебега

$$\int_{Y^{-1}(B)} m(Y)\, dP = \int_B m(y) P_Y(dy) ,$$

де P_Y – розподіл імовірностей на множині значень випадкового елемента Y. Отже, $m(y)$ – борелівська функція така, що за будь-якого $B \in \mathscr{B}(R^1)$

$$\int_{Y^{-1}(B)} X dP = \int_B m(y) P_Y(dy) .$$

Тоді $M[X/Y = y]$ можна означити й так.

Умовним математичним сподіванням випадкової величини X *за умови* $Y = y$ називається будь-яка борелівська функція $m(y)$, стосовно якої

$$\int_{Y^{-1}(B)} X dP = \int_B m(y) P_Y(dy) .$$

Згідно з цим означенням

$$\int\limits_{Y^{-1}(B)} XdP = \int\limits_{B} m(y)P_Y(dy) = \int\limits_{Y^{-1}(B)} m(Y)dP .$$

Функція $m(Y)$ є Q_Y-вимірною, і множинами $Y^{-1}(B)$ вичерпуються всі множини з Q_Y. Звідси випливає, що $m(Y)$ є математичним сподіванням $M[X/Y]$. Отже, знаючи $M[X/Y=y]$, можна відновити $M[X/Y]$ і, навпаки, знаючи $M[X/Y]$, можна знайти $M[X/Y=y]$.

Нехай X і Y – відповідно s-вимірний і l-вимірний випадкові вектори. Якщо відомий умовний розподіл імовірностей $f_{X/Y}(x/y)$, то

$$M[X/Y=y] = \int\limits_{-\infty}^{\infty} x f_{X/Y}(x/y)\,dx .$$

Якщо $M[|X|] < \infty$, то $M[X/Y=y]$ визначене за всіх $y \in \mathcal{N}_Y$, де $\mathcal{N}_Y = \{y \mid f_Y(y) = 0\}$.

Справді,

$$M[X/Y=y] = \int\limits_{-\infty}^{\infty} x \frac{f_{(X,Y)}(x,y)}{f_Y(y)}\,dx ,$$

причому

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} |x|\, f_{(X,Y)}(x,y)\,dxdy = \int\limits_{-\infty}^{\infty} |x|\, f_X(x)\,dx < \infty .$$

Тому $\int\limits_{-\infty}^{\infty} x f_{(X,Y)}(x,y)\,dx < \infty$ майже за всіх y.

Якщо замість значення y розглядати випадковий вектор Y, то дістанемо випадкові величини $f_{X/Y}(x/Y)$, $M[X/Y]$, які називаються відповідно *умовною щільністю розподілу ймовірностей і умовним математичним сподіванням випадкової величини* X *відносно випадкового вектора* Y. Те, що функція $f_{X/Y}(x/y)$ на множині $\{y \mid y \in \mathcal{N}_Y\}$ визначається довільно, не призводить до суперечностей, оскільки

$$P(\{Y \in \mathcal{N}_Y\}) = \int\limits_{\mathcal{N}_Y} f_Y(y)dy = 0 .$$

З цього означення $M[X/Y]$ випливає, що $M[X/Y] = g(Y)$, де $g(Y)$ – деяка борелівська функція вектора $y \in R^l$.

Ця функція має таку властивість. За довільної обмеженої борелівської функції $h(y)$, $Y \in R^1$,

$$M[h[Y]X] = M[h(Y)g(Y)].$$

Справді,

$$M[h(Y)g(Y)] = \int_{R^l} h(y)g(y)f_Y(y)dy = \int_{R^l} h(y)\left\{\int_{R^s} xf_{X/Y}(x/y)dx\right\} f_Y(y)dy =$$

$$= \int_{R^l}\int_{R^s} h(y)xf_{(X,Y)}(x,y)\,dxdy = M[h(Y)X].$$

Тут випадкова величина $g(Y)$ визначається однозначно, тобто якщо існують дві функції $g_1(y)$, $g_2(y)$ такі, що за всіх обмежених борелівських функцій $h(y)$ виконуються рівності

$$M[h(Y)X] = M[h(Y)g_1(Y)],\ M[h(Y)X] = M[h(Y)g_2(Y)],$$

то

$$P(g_1(Y) = g_2(Y)) = 1.$$

Справді, якщо зазначені рівності справджуються, то

$$M[h(Y)(g_1(Y) - g_2(Y))] = 0.$$

Якщо $h(y) = (q_1(y) - q_2(y))\gamma_m$, де $\gamma_m = 1$, коли $|q_1(y)| < M$, $|q_2(y)| < M$, і $\gamma_m = 0$ в протилежному разі, то

$$M[\gamma_m (g_1(Y) - g_2(Y))^2] = 0,$$

звідки

$$P(\gamma_m (g_1(Y) - g_2(Y))^2 = 0) = 1.$$

Нехай $M \to \infty$. Тоді

$$P(g_1(Y) = g_2(Y)) = 1.$$

Таким чином, враховуючи однозначність визначення $g(Y)$, можна дати таке означення.

Умовним математичним сподіванням $M[X/Y]$ випадкової величини X, $(M[|X|] < \infty)$, відносно випадкового вектора Y називають випадкову величину виду $g(Y)$, де $g(y)$ – деяка борелівська функція така, що за довільної обмеженої борелівської функції $h(y)$ виконується рівність

$$M[h(Y)X] = M[h(Y)g(Y)].$$

Це означення співпадає з означенням, даним раніше.

Здобутий результат можна узагальнити: якщо $f_{(X,Y)}(x,y)$ щільність сумісного розподілу ймовірностей випадкових векторів X і Y, то за довільної борелівської функції $\varphi(x)$, $x \in R^l$, за якої $M[\varphi(X)] < \infty$, виконується рівність

$$M[\varphi(X)/Y]=\int_{R^l}\varphi(x)f_{X/Y}(x/y)\,dx\,.$$

Справді, позначимо через $g(Y)$ праву частину останньої рівності. Тоді

$$M[h(Y)g(Y)]=\int_{R^s}h(y)g(y)f_Y(y)dy=\int_{R^s}\int_{R^l}h(y)\varphi(x)f_{X/Y}(x/y)f_Y(y)dydx=$$

$$=\int_{R^s}\int_{R^l}h(y)\varphi(x)f_{(X,Y)}(x,y)dydx=M[h(Y)\varphi(X)]\,.$$

Таким чином, стосовно випадкової величини $g(Y)$ задовільняється означення умовного математичного сподівання випадкової величини $\varphi(X)$.

Поклавши $h(y)=1$, дістанемо

$$M[M[X/Y]]=M[X]\,.$$

Якщо випадкові вектори X і Y незалежні, то

$$M[\varphi(X)/Y]=M[\varphi(X)]\,.$$

Нехай розглядаються випадкові величини Y, $X_1, X_2,..., X_n$, причому спостереженню підлягає вектор $X=(X_1, X_2,..., X_n)$, а за результатами спостережень необхідно оцінити величину Y з якомога меншою похибкою, тобто треба знайти таку функцію $g(x)=g(x_1, x_2, ..., x_n)$, щоб випадкова величина $g(X)$ якомога менше відрізнялася від випадкової величини Y.

Якщо за норму відхилення величин Y і $\overline{Y}$ взяти величину

$$\delta=(M[(Y-\overline{Y})^2])^{1/2},$$

то задача полягає в знаходженні такої функції $g(x_1, x_2, ..., x_n)$, за якої величина δ набуває найменшого значення.

Таким чином, ставиться наступна задача.

Нехай $M[Y^2]<\infty$. В класі H всіх борелівських функцій $h(x)=h(x_1, x_2, ..., x_n)$ на R^n, стосовно яких $M[h^2(X)]<\infty$, треба знайти таку функцію $g(x)$, щоб за всіх $h(x)\in H$ виконувалась умова

$$M[(Y-g(X))^2]\le M[(Y-h(X))^2].$$

Нехай L_2 – гільбертів простір всіх випадкових величин X, стосовно яких $M[X^2]<\infty$, із скалярним добутком $(X_1,X_2)=M[X_1\cdot X_2]$ і відповідно нормою $\|X\|^2=(X,X)$, H – лінійний простір в L_2, який складається з випадкових величин Z виду

$$Z=h(X)=h(X_1,\ X_2,\ ...,\ X_n)\,.$$

Тоді за довільного $h\in H$

$$\delta^2=M[(Y-h(X))^2]=M[(Y-h(X)+g(X)-g(X))^2]=$$

$$= M[(Y - g(X))^2] + 2M[(Y - g(X))(g(X) - h(X))] + M[(g(X) - h(X))^2].$$

Якщо

$$M[(Y - g(X))(g(X) - h(X))] = 0,$$

то

$$\delta^2 = M[(Y - g(X))^2] + M[(g(X) - h(X))^2] \geq M[(Y - g(X))^2].$$

Оскільки

$$M[(Y - g(X))(g(X) - h(X))] = M[(g(X) - h(X))]M[(Y - g(X))/X]$$

і $h \in H$ – довільне, то

$$M[(Y - g(X))/X] = 0.$$

Враховуючи, що

$$M[(Y - g(X))/X] = M[Y/X] - M[g(X)/X],$$

а

$$M[g(X)/X] = g(X)$$

(через те що $M[g(X)/X = x] = g(x)$), дістаємо

$$g(X) = M[Y/X].$$

Таким чином, найкращим наближенням випадкової величини Y є випадкова величина $\overline{Y} = g(X) = M[Y/X]$.

У геометричній інтерпретації $M[Y/X]$ є елементом з H, який найменше відхиляється від елемента X у зазначеному розумінні, тобто $M[Y/X]$ є ортогональною проекцією елемента Y на простір H. Функція $g(X) = M[Y/X]$ називається *функцією регресії величини Y на X*.

Функції регресії можна дати таке тлумачення.

Нехай між деякими фізичними величинами існує зв'язок $y = g(x_1, x_2, ..., x_n)$ і виконуються експерименти, в яких величина $Y = Y(X_1, X_2, ..., X_n)$ набуває випадкового значення $g(x_1, x_2, ..., x_n) + \Delta_y(x_1, x_2, ..., x_n)$ із незміщеною похибкою Δ_y, математичне сподівання якої дорівнює нулю, тобто $M[\Delta_y / X = x] = 0$ за всіх $x \in R^n$.

Тоді, з одного боку,

$$M[Y/X = x] = g(x),$$

а з іншого боку, за функцією $g(x_1, x_2, ..., x_n)$ визначається умовне середнє значення випадкової величини $Y = g(X_1, X_2, ..., X_n)$ за умови, що випадкові величини X_i набувають значень x_i, $i = 1, 2, ..., n$.

Вправи для самостійного виконання

4.15.1. Перевірити, чи правильні твердження:

1. Стосовно випадкової величини $P(A/D)(E)$ існує скінченне математичне сподівання за будь-якого поділу D простору Ω.

2. Завжди $0 < M[P(A/D)] < 1$.

3. За кожною випадковою величиною $Y(E)$, $E \in \Omega$, породжується деякий поділ D_Y простору Ω.

4. Якщо $Y(E)$ –випадкова величина з дискретною множиною значень, то твердження 3 правильне.

5. Умовна ймовірність події A відносно випадкової величини Y – це нова випадкова величина $P(A/Y)$, що набуває значень

$$P(A/(Y = y_i)) = P(A/Y^{-1}(y_i))$$

за будь-якого значення y_i величини Y.

6. Умовна ймовірність події A відносно випадкової величини Y існує завжди.

7. Умовне математичне сподівання випадкової величини $X(E)$, $E \in \Omega$, відносно поділу D простору Ω завжди існує і є скінченним числом.

8. Завжди

$$\sum_i x_i \sum_j P(X^{-1}(x_i)/D_j) Y_{D_j}(E) = \sum_j Y_{D_j}(E) \sum_i x_i P(X^{-1}(x_i)/D_j) \cdot$$

9. Кожна випадкова величина з дискретною множиною значень є вимірною відносно будь-якого поділу D простору Ω.

10. За будь-яких випадкових величин X та Y існує $M[X/Y]$ – умовне математичне сподівання X відносно Y.

11. Якщо $M[X/Y]$ існує, то воно є скінченним числом.

12. $M[X/Y]$ є випадковою величиною.

13. $M[X/Y]$ є простою випадковою величиною.

14. За довільної випадкової величини $X(E)$, $E \in \Omega$, і довільної події A існує умовне математичне сподівання випадкової величини X відносно події A.

15. За довільної випадкової величини $X(E)$, $E \in \Omega$, і довільної σ-алгебри $Q \subset S$ існує умовне математичне сподівання випадкової величини X відносно σ-алгебри Q.

16. За довільної події $A \in S$ і довільної σ-алгебри $Q \subset S$ існує умовна ймовірність події A відносно σ-алгебри Q.

17. За довільної випадкової величини X існує $P(A/X)$.

18. За будь-яких випадкових величин X та Y існує умовне математичне сподівання випадкової величини X за умови $Y = y$, де $y \in R$ – довільне фіксоване.

19. Для того, щоб знайти $M[X / Y]$, необхідно й достатньо знайти $M[X / Y = y]$ за деякого $y \in R$.

20. Умовне математичне сподівання випадкової величини X відносно випадкового вектора Y існує завжди і є певним числом.

21. За довільної випадкової величини X і довільного випадкового вектора Y існує борелівська функція $g(y)$ така, що $M[h(Y) \cdot X] = M[h(Y) \cdot g(Y)]$, де $h(y)$ – довільна борелівська функція.

22. За довільних випадкових векторів X та Y і довільної борелівської функції $\varphi(x)$ існує борелівська функція $g(y)$ така, що $M[h(Y) \cdot \varphi(X)] = M[h(Y) \cdot g(Y)]$, де $h(y)$ – довільна борелівська функція.

23. За будь-яких випадкових векторів X та Y правильна рівність

$$M[M[X / Y]] = M[X].$$

24. Завжди $M[\varphi(X) / Y] = M[\varphi(X)]$.

25. За будь-якої випадкової величини $Y \in L^2$ та випадкового вектора $X = (X_1, X_2, \ldots, X_n)$ існує борелівська функція $g(X) \in L^2$, стосовно якої $M[(Y - g(X))^2] \leq M[Y - h(X))^2]$, де $h(X) \in L^2$ – довільна борелівська функція.

26. Якщо $g(X)$ з твердження 25, то $g(X) = M[Y / X]$.

27. Твердження, обернене до 26, правильне.

4.16. Лінійна регресія

Нагадаємо, що умовним математичним сподіванням випадкової величини X з дискретним розподілом ймовірностей на множині Ω_X її значень за умови, що випадкова величина Y набула значення y (y – деяке можливе значення випадкової величини Y) називається сума добутків можливих значень величини X та їх умовних ймовірностей:

$$M[X / y] = \sum_{j=1}^{n} x_j P_{X/Y}(\{x_j\}/\{y\}),$$

де $P_{X/Y}(\{x_j\}/\{y\})$ умовна ймовірність події $\{x_j\}$ за умови, що мала місце подія $\{y\}$.

В разі випадкової величини X з абсолютно неперервним розподілом ймовірностей на множині Ω_X її значень

$$M[X / y] = \int_{-\infty}^{+\infty} x\, f_{X/Y}(x/y)\,dx,$$

де $f_{X/Y}(x/y)$ – щільність розподілу ймовірностей на множині Ω_X значень випадкової величини X за умови, що випадкова величина Y набула значення y.

Означення 1. *Функцією регресії випадкової величини X на випадкову величину Y* є функція від y, яка дорівнює умовному математичному сподіванню випадкової величини X за умови, що випадкова величина Y набула значення y, тобто

$$f(y) = M[X / y].$$

За даним означенням *функцією регресії випадкової величини Y на випадкову величину X* є функція від x, яка дорівнює умовному математичному сподіванню випадкової величини Y за умови, що випадкова величина X набула значення x, тобто

$$g(x) = M[Y / x].$$

Означення 2. Рівняння $x = f(y)$ (чи $y = g(x)$) називається *рівнянням регресії* X на Y (чи Y на X), а лінія на площині, яка відповідна цьому рівнянню, називається *лінією регресії*.

За лінією регресії Y на X (чи X на Y) оцінюють, як в середньому залежить Y від X (чи X від Y).

Приклад 4.16.1. Нехай X і Y незалежні випадкові величини, $M(X) = m_X$, $M(Y) = m_Y$. Тоді

$$g(x) = M[Y / x] = M(Y) = m_Y; \quad f(y) = M[X / y] = M(X) = m_X.$$

Лінії регресії, відповідні розглядуваному випадку, зображені на Рис. 4.16.1.

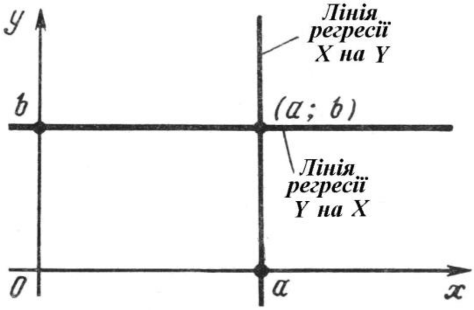


Рис. 4.16.1

Приклад 4.16.2. Нехай між випадковими величинами X і Y існує лінійна залежність: $Y = aX + b$, $a \neq 0$. Тоді функція регресії Y на X набуває вигляду

$$g(x) = M[Y / X = x] = M[aX + b / X = x] = ax + b.$$

Оскільки $X = \frac{1}{a}(Y - b)$, то функція регресії X на Y набуває вигляду

$$f(y) = M[X / Y = y] = M\left[\frac{1}{a}(Y - b) / Y = y\right] = \frac{1}{a}(y - b).$$

Отже, лінія регресії X на Y: $x = \frac{y - b}{a}$, тобто $y = ax + b$. Таким чином, у випадку лінійної залежності між випадковими величинами X і Y лінії регресії X на Y і Y на X співпадають, і ця спільна лінія є прямою.

Лінійна кореляція. Часто доводиться мати справу з *нечітко вираженими залежностями між випадковими величинами*. Такий зв'язок, наприклад, між опадами і врожаєм, між товщиною снігового покриття взимку і об'ємом стоку подальших повеней і т. ін. Тут кожному значенню однієї величини відповідає множина можливих значень іншої величини. Подібного роду залежності відносяться до *кореляційних залежностей*.

Означення 3. Кореляційна залежність між випадковими величинами X і Y називається *лінійною кореляцією*, якщо обидві функції регресії $f(y)$ і $g(x)$ є лінійними.

У цьому випадку обидві лінії регресії є прямими; вони називаються *прямими регресії*.

Виведемо рівняння прямої регресії випадкової величини Y на X, тобто знайдемо коефіцієнти лінійної функції $g(x) = ax + b$.

Позначимо $M[X]=m_X$, $M[Y]=m_Y$, $M[(X-m_X)^2]=\sigma_X^2$, $M[(Y-m_Y)^2]=\sigma_Y^2$. Використовуючи властивості математичного сподівання, знаходимо:

$$M[Y]=M[M[Y/X]]=M[g(X)]=M(aX+b)=aM[X]+b,$$

тобто $m_Y=am_X+b$, звідки $b=m_Y-am_X$.

Далі

$$M[XY]=M[Xg(X)]=M[aX^2+bX]=$$

$$=aM[X^2]+bM[X]=aM[X^2]+(m_Y-am_X)m_X,$$

звідки

$$a=\frac{K[X,Y]}{M[X^2]-m_X^2},$$

де

$$K[X,Y]=M[XY]-M[X]M[Y],$$

або, враховуючи, що $D[X]=M[X^2]-M^2[X]=\sigma_X^2$,

$$a=\frac{K[X,Y]}{\sigma_X^2}.$$

Отриманий коефіцієнт називається *коефіцієнтом регресії* Y *на* X і позначається через $\rho(Y/X)$:

$$\rho(Y/X)=\frac{K[X,Y]}{\sigma_X^2}.$$

Таким чином, рівняння лінії регресії Y на X набуває вигляду

$$y=\rho(Y/X)(x-m_X)+m_Y.$$

Аналогічно можна отримати рівняння лінії регресії X на Y

$$x=\rho(X/Y)(y-m_Y)+m_X,$$

де

$$\rho(X/Y)=\frac{K[X,Y]}{\sigma_Y^2}$$

коефіцієнт регресії X *на* Y.

Рівняння прямих регресії можна записати у більш симетричному вигляді, якщо скористатися *коефіцієнтом кореляції*,

$$r=\frac{M[XY]-M[X]M[Y]}{\sigma(X)\,\sigma(Y)},$$

або більш коротко

$$r = \frac{K[X,Y]}{\sigma_X \; \sigma_Y}.$$

Враховуючи цей коефіцієнт, дістанемо:

$$\rho(Y/X) = r\frac{\sigma_Y}{\sigma_X}; \qquad \rho(X/Y) = r\frac{\sigma_X}{\sigma_Y}, \qquad (4.16.1)$$

і тому рівняння прямих регресії набувають вигляду:

$$y - m_Y = r\frac{\sigma_Y}{\sigma_X}(x - m_X);$$

$$x - m_X = r\frac{\sigma_X}{\sigma_Y}(y - m_Y).$$

З рівнянь прямих регресії видно, що обидві ці прямі проходять через точку (m_X, m_Y) (центр розсіювання імовірностей); кутові коефіцієнти прямих регресії відповідно дорівнюють (позначення кутів див. на Рис. 4.16.2):

$$\mathrm{tg}\,\alpha = r\frac{\sigma_Y}{\sigma_X}; \qquad \mathrm{tg}\,\beta = (1/r)\frac{\sigma_Y}{\sigma_X}.$$

Оскільки $|r| \le 1$, то $|\mathrm{tg}\,\alpha| \le |\mathrm{tg}\,\beta|$. Це означає, що пряма регресії Y на X менше нахилена до осі абсцис, ніж пряма регресії X на Y. Чим ближче $|r|$ до одиниці, тим менший кут між прямими регресії. Ці прямі співпадають тоді і тільки тоді, коли $|r| = 1$.

Рис. 4.16.2

Коли $r = 0$, тоді прямі регресії задаються рівняннями $y = m_Y$, $x = m_X$.

У цьому випадку

$$M[Y/X = x] = m_Y = M(Y); \; M[X/Y = y] = m_X = M(X).$$

З формул (4.16.1) видно, що знаки коефіцієнтів регресії такі самі, як і знак коефіцієнта кореляції r, і пов'язані співвідношенням

$$\rho(Y/X)\,\rho(X/Y)=r^2.$$

Вправи для самостійного виконання

4.16.1. Перевірити, чи правильні твердження:

1. За будь-яких випадкової величини Y та випадкового вектора X існує функція регресії Y на X.
2. За будь-яких випадкових величин X та Y існують функції регресії однієї з цих величин на іншу.
3. Функція регресії – це те саме, що й рівняння регресії.
4. Лінія регресії – це множина точок на площині, через координати яких задовільняється рівняння регресії.
5. За будь-яких випадкових величин існує кореляційна залежність між ними.
6. За будь-яких випадкових величин X та Y існує рівняння лінії регресії Y на X та X на Y.
7. Якщо випадкові величини X та Y пов'язані лінійною залежністю, то існують коефіцієнти регресії $\rho(Y/X)$ та $\rho(X/Y)$.
8. За будь-яких випадкових величин X та Y існує коефіцієнт кореляції.
9. Коефіцієнт кореляції $r=0$ тоді й тільки тоді, коли $K[X,Y]=0$.
10. За будь-яких величин X та Y можна знайти коефіцієнт кореляції $r=\dfrac{K[X,Y]}{\sigma_X\sigma_Y}$ та записати рівняння прямих регресії:

$$y-M[Y]=r\frac{\sigma_Y}{\sigma_X}(x-M[X]) \text{ та } x-M[X]=r\frac{\sigma_X}{\sigma_Y}(y-M[Y]).$$

11. Коефіцієнт кореляції r завжди не перевищує 1 та не менший за -1.
12. Прямі регресії X на Y та Y на X співпадають тоді й тільки тоді, коли $|r|=1$.
13. Якщо $r=0$, то прямі регресії взаємно перпендикулярні.

4.17. Деякі важливі функції від випадкових величин з нормальними розподілами ймовірностей

Розглянемо розподіли ймовірностей на множинах значень деяких функцій від випадкових величин з нормальними розподілами ймовірностей на множинах їх значень, які використовуються під час обґрунтування окремих тверджень, опрацювання статистичних даних тощо.

1. Лінійна функція $Y = \psi(X) = aX + b$, $a \neq 0$, від випадкової величини X з нормальним розподілом ймовірностей на множині її значень із щільністю

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

є також випадковою величиною з нормальним розподілом ймовірностей на множині її значень, параметри якого:

$$M[Y] = aM[X] + b = am + b,\ \sigma[Y] = |a|\,\sigma[X] = |a|\,\sigma.$$

Розглянемо два випадки.

1-й випадок. Нехай $a > 0$. Тоді (Рис. 4.17.1)

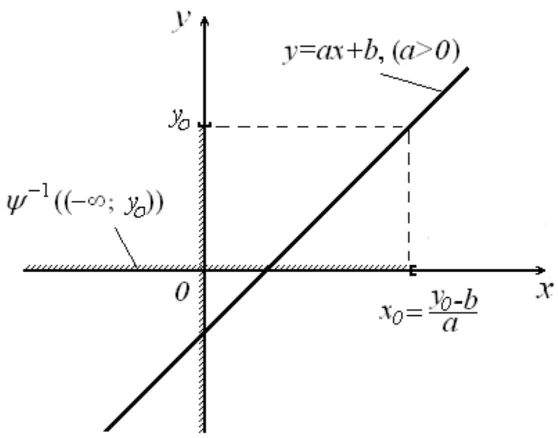


Рис. 4.17.1

$$F_Y(y) = P_Y((-\infty; y)) = P_X(\psi^{-1}((-\infty; y))) = \frac{1}{\sqrt{2\pi\sigma}} \int\limits_{-\infty}^{\frac{y-b}{a}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Звідси

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \frac{d}{dy} \int\limits_{-\infty}^{\frac{y-b}{a}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-m\right)^2}{2\sigma^2}} \cdot \frac{1}{a} = \frac{1}{\sqrt{2\pi}(a\sigma)} e^{-\frac{(y-(am+b))^2}{2(a\sigma)^2}} .$$

Таким чином, розподіл імовірностей на множині значень випадкової величини Y також нормальний, причому

$$M[Y] = aM[X] + b = am + b \, , \; \sigma(Y) = a\sigma[X] = a\sigma \, .$$

2-й випадок. Нехай $a < 0$. Дістаємо (Рис. 4.17.2)

$$F_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\frac{y-b}{a}}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \, ,$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-m\right)^2}{2\sigma^2}} \frac{1}{a} = \frac{1}{\sqrt{2\pi}(\sigma\,|\,a\,|)} e^{-\frac{(y-(a\,m+b))^2}{2(a\,\sigma)^2}} .$$

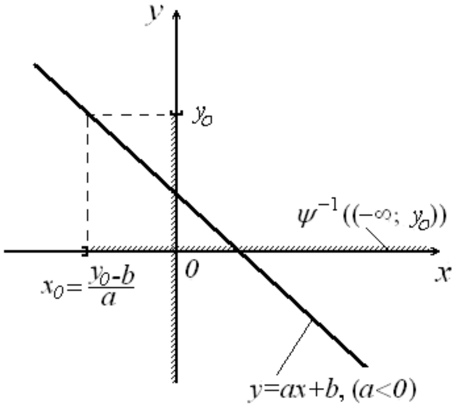


Рис. 4.17.2

Отже, розподіл імовірностей на множині значень лінійної функції $Y = aX + b$, $a \neq 0$, випадкового аргумента X з нормальним розподілом імовірностей на множині його значень є також нормальним, причому параметрами цього розподілу є

$$M[Y] = aM[X] + b = am + b \, , \; \sigma[Y] = |\,a\,|\,\sigma[X] = |\,a\,|\,\sigma \, .$$

2. Функція $W = \psi(X) = X^2$, якщо розподіл імовірностей на множині значень випадкової величини X нормальний з параметрами $m = 0$, $\sigma = 1$.

Одержуємо

$$F_W(w) = P_W((-\infty;\, w)) = P_X(\psi^{-1}((-\infty;\, w))) =$$

$$= \begin{cases} 0, & \text{коли } w \le 0; \\ \int\limits_{-\sqrt{w}}^{\sqrt{w}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{2}{\sqrt{2\pi}} \int\limits_{0}^{\sqrt{w}} e^{-\frac{t^2}{2}} dt, & \text{коли } w > 0, \end{cases}$$

$$f_W(w) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-\frac{w}{2}} \frac{1}{2\sqrt{w}} = \frac{1}{\sqrt{2\pi}} w^{-\frac{1}{2}} e^{-\frac{w}{2}}, & \text{коли } w > 0; \\ 0, & \text{коли } w \le 0. \end{cases}$$

3. Функція $Z_2 = \psi(X_1,\ X_2) = X_1^2 + X_2^2$, якщо X_1 і X_2 незалежні випадкові величини з нормальними розподілами ймовірностей на множинах їх значень з параметрами $a_{X_1} = 0$, $\sigma_{X_1} = 1$, $a_{X_2} = 0$, $\sigma_{X_2} = 1$.

Дістаємо

$$F_{Z_2}(z_2) = P_{Z_2}((-\infty;\, z_2)) = P_{(X_1,\, X_2)}(\psi^{-1}((-\infty;\, z_2))) =$$

$$= \begin{cases} 0, & \text{коли } z_2 \le 0; \\ \iint\limits_{D(z_2)} \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}} dx_1 dx_2 = \frac{1}{2\pi} \int\limits_{0}^{2\pi} d\varphi \int\limits_{0}^{\sqrt{z_2}} e^{-\frac{\rho^2}{2}} \rho\, d\rho, & \text{коли } z > 0, \end{cases}$$

де

$$D(z_2) = \psi^{-1}((-\infty;\, z_2)) = \{(x_1,\, x_2) \mid x_1^2 + x_2^2 < z_2,\ z_2 > 0\},$$

$$x_1 = \rho\cos\varphi,\quad x_2 = \rho\sin\varphi,\quad dx_1 dx_2 = \rho\, d\rho\, d\varphi .$$

Оскільки

$$\int\limits_{0}^{\sqrt{z_2}} e^{-\frac{\rho^2}{2}} \rho d\rho = 2 \int\limits_{0}^{\sqrt{z_2}} e^{-\left(\frac{\rho}{\sqrt{2}}\right)^2} \left(\frac{\rho}{\sqrt{2}}\right) d\left(\frac{\rho}{\sqrt{2}}\right) =$$

$$= 2 \int\limits_{0}^{\frac{\sqrt{z_2}}{\sqrt{2}}} e^{-t^2} t dt = \left(-e^{-t^2} \Bigg|_{0}^{\frac{\sqrt{z_2}}{\sqrt{2}}} \right) = (1 - e^{-\frac{z_2}{2}}),$$

то

$F_{Z_2}(z_2) = (1 - e^{-\frac{z_2}{2}})$, коли $z_2 > 0$; $f_{Z_2}(z_2) = \frac{1}{2} e^{-\frac{z_2}{2}}$, $z_2 > 0$.

Нехай тепер $Z_3 = \psi(X_1,\ X_2,\ X_3) = X_1^2 + X_2^2 + X_3^2$, де $X_1,\ X_2,\ X_3$ – незалежні випадкові величини з нормальними розподілами ймовірностей на множинах їх значень з параметрами $a_{X_i} = 0$, $\sigma_{X_i} = 1$, $i = 1,\ 2,\ 3$.

Тоді

$$F_{Z_3}(z_3) = \iiint\limits_{D(z_3)} \frac{1}{(\sqrt{2\pi})^3} e^{-\frac{x_1^2 + x_2^2 + x_3^2}{2}} dx_1 dx_2 dx_3 ,$$

де

$$D(z_3) = \{(x_1,\ x_2,\ x_3) \mid x_1^2 + x_2^2 + x_3^2 < z_3\} .$$

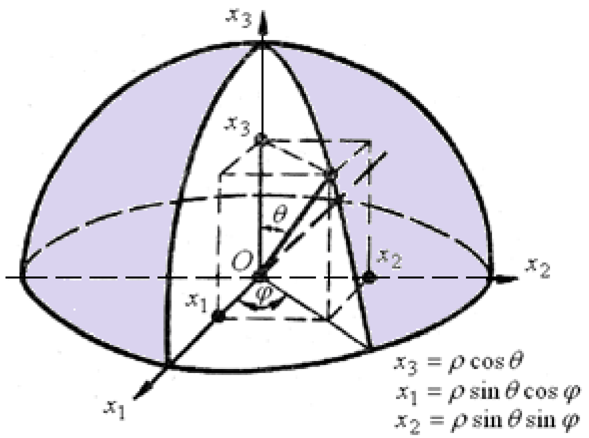


Рис. 4.17.3

Після переходу до сферичних координат (Рис. 4.17.3)

$$x_1 = \rho\sin\theta\cos\varphi ,\ \ x_2 = \rho\sin\theta\sin\varphi ,\ \ x_3 = \rho\cos\theta ,$$

$$dx_1 dx_2 dx_3 = \rho^2 \sin\theta d\rho d\theta d\varphi$$

дістанемо

$$F_{Z_3}(z_3) = \frac{1}{(\sqrt{2\pi})^3} \iiint\limits_{D(z_3)} e^{-\frac{\rho^2}{2}} \rho^2 \sin\theta d\rho d\theta d\varphi =$$

$$=\frac{1}{(\sqrt{2\pi})^3}\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\theta d\theta\int_0^{\sqrt{z_3}}e^{-\frac{\rho^2}{2}}\rho^2 d\rho=\frac{1}{(\sqrt{2\pi})^3}2\pi\cdot 2\cdot\int_0^{\sqrt{z_3}}e^{-\frac{\rho^2}{2}}\rho^2 d\rho\ ,$$

$$f_{Z_3}(z_3)=\frac{2}{\sqrt{2\pi}}e^{-\frac{z_3}{2}}z_3\frac{1}{2\sqrt{z_3}}=\frac{1}{\sqrt{2\pi}}z_3^{1/2}e^{-\frac{z_3}{2}}\ .$$

Цей розподіл імовірностей на множині значень випадкової величини Z_3 називають *розподілом Максвелла*. За допомогою величини Z_3 описують квадрат швидкості частинки в припущенні, що проекції швидкості на осі координат у тривимірному просторі є випадковими величинами зі стандартними нормальними розподілами ймовірностей на множинах їх значень.

Як було зазначено, коли випадкові величини X_1 і X_2 незалежні, то й випадкові величини $W_1=\psi_1(X_1)=X_1^2$, $W_2=\psi_2(X_2)=X_2^2$ незалежні. Тому щільність розподілу ймовірностей на множині значень суми незалежних випадкових величин X_1^2 і X_2^2 можна шукати як згортку щільностей $f_{W_1}(w_1)$ і $f_{W_2}(w_2)$:

$$f_{Z_2}(z_2)=\int_0^{z_2}f_{W_1}(w_1)f_{W_2}(z_2-w_1)dw_1=\int_0^{z_2}f_{W_1}(z_2-w_2)f_{W_2}(w_2)dw_2\ .$$

Таким чином $f_{Z_2}(z_2)$ за $(z_2>0)$ у прикладі 3 можна було знайти так:

$$f_{Z_2}(z_2)=\int_0^{z_2}\frac{1}{(\sqrt{2\pi})^2}(z_2-w_2)^{-\frac{1}{2}}e^{-\frac{z_2-w_2}{2}}w_2^{-\frac{1}{2}}e^{-\frac{w_2}{2}}dw_2=$$

$$=\frac{1}{(\sqrt{2\pi})^2}e^{-\frac{z_2}{2}}\int_0^{z_2}(z_2-w_2)^{-\frac{1}{2}}e^{\frac{w_2}{2}}e^{-\frac{w_2}{2}}w_2^{-\frac{1}{2}}dw_2=$$

$$=\frac{1}{(\sqrt{2\pi})^2}e^{-\frac{z_2}{2}}\int_0^{z_2}(z_2-w_2)^{-\frac{1}{2}}w_2^{-\frac{1}{2}}dw_2=\frac{1}{2\pi}e^{-\frac{z_2}{2}}\int_0^{z_2}\frac{dw_2}{\sqrt{z_2w_2-w_2^2}}=$$

$$=\frac{1}{2\pi}e^{-\frac{z_2}{2}}\int_0^{z_2}\frac{dw_2}{\sqrt{(\frac{z_2}{2})^2-(\frac{z_2}{2}-w_2)^2}}=\frac{1}{2\pi}e^{-\frac{z_2}{2}}\arcsin\frac{w_2-\frac{z_2}{2}}{\frac{z_2}{2}}\Bigg|_0^{z_2}=$$

$$=\frac{1}{2\pi}e^{-\frac{z_2}{2}}\pi=\frac{1}{2}e^{-\frac{z_2}{2}}.$$

Аналогічно, визначивши щільність розподілу ймовірностей на множинах значень випадкових величин $Z_2=X_1^2+X_2^2$ (див. п. 2.3) і $W_3=\psi_3(X_3)=X_3^2$ і враховуючи, що випадкові величини Z_2 і W_3 незалежні, стосовно $Z_3=Z_2+W_3$ знаходимо

$$f_{Z_3}(z_3)=\int_0^{z_3} f_{Z_2}(z_2)f_{W_2}(z_3-z_2)dz_2=$$

$$=\int_0^{z_3}\frac{1}{2}e^{-\frac{z_2}{2}}\frac{1}{\sqrt{2\pi}}(z_3-z_2)^{-\frac{1}{2}}e^{-\frac{z_3-z_2}{2}}dz_2=\frac{1}{2\sqrt{2\pi}}e^{-\frac{z_3}{2}}\int_0^{z_3}(z_3-z_2)^{-\frac{1}{2}}dz_2=$$

$$=\frac{1}{2\sqrt{2\pi}}e^{-\frac{z_3}{2}}\left(-2(z_3-z_2)^{\frac{1}{2}}\Big|_0^{z_3}\right)=\frac{1}{\sqrt{2\pi}}z_3^{1/2}e^{-\frac{z_3}{2}}.$$

За довільного $n\geq 1$ розподіл імовірностей на множині значень випадкової величини $Z_n=\sum_{i=1}^{n}X_i^2$, де X_i – незалежні випадкові величини зі стандартними нормальними розподілами ймовірностей на множинах їх значень, $m_{X_i}=0$, $\sigma_{X_i}=1$, називається χ_n^2*-розподілом з* n *ступенями вільності*. Саму величину Z_n також позначають через χ_n^2. Виявляється, що

$$f_{Z_n}(z_n)=\begin{cases}0, & \text{коли } z_n<0;\\ \dfrac{1}{\Gamma(\frac{n}{2})2^{n/2}}z_n^{\frac{n}{2}-1}e^{-\frac{z_n}{2}}, & \text{коли } z_n\geq 0.\end{cases}$$

За необмеженого збільшення числа доданків в сумі Z_n розподіл імовірностей на множині значень випадкової величини Z_n асимптотично наближається до нормального.

4. Розглянемо центральні моменти *k*-го порядку випадкової величини X з нормальним розподілом імовірностей на множині її значень з параметрами m і σ.

Одержуємо

$$M[(x-m)^k]=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-m)^k e^{-\frac{(x-m)^2}{2\sigma^2}}dx,$$

або після заміни $\dfrac{x-m}{\sigma}=t$

$$M[(X-m)^k]=\frac{\sigma^k}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^k e^{-\frac{t^2}{2}}dt\ .$$

Якщо k непарне, то $M[(X-m)^k]=0$. Якщо k парне, $k=2r$, $r=0,1,2,\ldots,$ то

$$M[(X-m)^{2r}]=\frac{2\sigma^{2r}}{\sqrt{2\pi}}\int_{0}^{\infty}t^{2r} e^{-\frac{t^2}{2}}dt\ ,$$

або після заміни $\dfrac{t^2}{2}=s$, $tdt=ds$, $t=\sqrt{2}\sqrt{s}$,

$$M[(X-m)^{2r}]=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\int_{0}^{\infty}s^{r-\frac{1}{2}}e^{-s}ds=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\Gamma\left(r+\frac{1}{2}\right).$$

Інтегруючи частинами, дістаємо

$$M[(X-m)^{2r}]=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\int_{0}^{\infty}s^{r-\frac{1}{2}}e^{-s}ds=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\int_{0}^{\infty}\left(r-\frac{1}{2}\right)s^{r-\frac{3}{2}}e^{-s}ds=\ldots=$$

$$=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\left(r-\frac{1}{2}\right)\left(r-\frac{3}{2}\right)\ldots\left(r-\left(r-\frac{3}{2}\right)\right)\int_{0}^{\infty}s^{\frac{1}{2}}e^{-s}ds\ .$$

Оскільки (після заміни $s=v^2$)

$$\int_{0}^{\infty}s^{\frac{1}{2}}e^{-s}ds=\int_{0}^{\infty}ve^{-v^2}2vdv=2\int_{0}^{\infty}v^2e^{-v^2}dv=2\left(-\frac{ve^{-v^2}}{2}\Bigg|_0^{\infty}+\frac{1}{2}\int_{0}^{\infty}e^{-v^2}dv\right)=\frac{\sqrt{\pi}}{2}\ ,$$

то

$$M[(X-m)^{2r}]=\frac{2^r\sigma^{2r}}{\sqrt{\pi}}\frac{2r-1}{2}\cdot\frac{2r-3}{2}\ldots\frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}=$$
$$=\sigma^{2r}(2r-1)(2r-3)\ldots3\cdot1=\sigma^{2r}(2r-1)!!$$

Зокрема, коли $r=1$, тоді

$$M[(X-m)^2]=\sigma^2\ .$$

5. Розглянемо двохвимірний нормальний розподіл ймовірностей з параметрами m_X, m_Y, σ_X, σ_Y, r (див. §3.18).

Обчислюючи $K[X,\,Y]$, одержимо

$$K[X,\,Y]=\iint_{R^2}(x-m_X)(y-m_Y)f_{(X,Y)}(x,y)dxdy=$$

$$=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma_X u\sigma_Y v e^{-\frac{1}{2(1-r^2)}(u^2(1-r^2)+(v-ru)^2)}\sigma_X\sigma_Y dudv=$$

$$=\frac{\sigma_X\sigma_Y}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}uve^{-\frac{u^2}{2}}e^{-\frac{(v-ru)^2}{2(1-r^2)}}dudv .$$

Зробимо заміну:

$$\frac{u}{\sqrt{2}}=s,\ \frac{v-ru}{\sqrt{2}\sqrt{1-r^2}}=w,\ u=\sqrt{2}s,\ v=\sqrt{2}\sqrt{1-r^2}w+\sqrt{2}rs,$$

$$\frac{\partial u}{\partial s}=\sqrt{2},\ \frac{\partial u}{\partial w}=0,\ \frac{\partial v}{\partial s}=\sqrt{2}r,\ \frac{\partial v}{\partial w}=\sqrt{2}\sqrt{1-r^2},$$

$$J=\begin{vmatrix}\frac{\partial u}{\partial s} & \frac{\partial u}{\partial w}\\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial w}\end{vmatrix}=\begin{vmatrix}\sqrt{2} & 0\\ \sqrt{2}r & \sqrt{2}\sqrt{1-r^2}\end{vmatrix}=2\sqrt{1-r^2} .$$

Тоді

$$K[X,\ Y]=\frac{\sigma_X\sigma_Y}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{2}s(\sqrt{2}\sqrt{1-r^2}w+\sqrt{2}rs)e^{-(s^2+w^2)}dsdw=$$

$$=\frac{\sigma_X\sigma_Y}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{2}s(\sqrt{2}\sqrt{1-r^2}we^{-s^2}e^{-w^2}dsdw+\frac{\sigma_X\sigma_Y}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}2rs^2e^{-s^2}e^{-w^2}dsdw=$$

$$=\frac{2r\sigma_X\sigma_Y}{\pi}\int_{-\infty}^{\infty}s^2e^{-s^2}ds\int_{-\infty}^{\infty}e^{-w^2}dw=\frac{2r\sigma_X\sigma_Y}{\pi}\cdot\frac{\sqrt{\pi}}{2}\cdot\sqrt{\pi}=r\sigma_X\sigma_Y .$$

Таким чином, r – коефіцієнт кореляції випадкових величин X і Y. Якщо $r=0$, то $f_{(X,Y)}(x,y)=f_X(x)f_Y(y)$ і тоді випадкові величини X і Y незалежні. Отже, за двохвимірного нормального розподілу ймовірностей із некорельованості випадкових величин X і Y випливає їх незалежність (хоч, як було показано, в загальному випадку це не так).

Знайдемо тепер

$$f_{Y/X}(y/x)=\frac{f_{(X,Y)}(x,y)}{f_X(x)}=\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-r^2}}e^{-\frac{\left(\frac{y-m_Y}{\sigma_Y}-r\frac{x-m_X}{\sigma_X}\right)^2}{2(1-r^2)}}=$$

$$= \frac{1}{\sqrt{2\pi}\sigma_Y \sqrt{1-r^2}} e^{-\frac{(y-m_Y - r\frac{\sigma_Y}{\sigma_X}(x-m_X))^2}{2(1-r^2)\sigma_Y^2}}.$$

Таким чином, $f_{Y/X}(y/x)$ є щільністю нормального розподілу ймовірностей на множині значень випадкової величини Y з параметрами $M[Y/X=x] = m_Y - r\frac{\sigma_Y}{\sigma_X}(x-m_X)$, $\sigma[Y/X=x] = \sigma_Y\sqrt{1-r^2}$, за умови, що випадкова величина X набуває значення x. Функція $g(x) = m_Y - r\frac{\sigma_Y}{\sigma_X}(x-m_X)$ є функцією регресії Y на X.

Аналогічно можна знайти щільність $f_{X/Y}(x/y)$ умовного розподілу ймовірностей на множині значень випадкової величини X за умови, що випадкова величина Y набуває значення y.

Лінії рівня $f_{(X,Y)}(x,y) = \text{const}$ в розглядуваному випадку є еліпси виду

$$\frac{(x-m_X)^2}{\sigma_X^2} - 2r\frac{(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2} = \lambda^2.$$

Перейдемо до нових координат s, w, поклавши

$$S = (X-m_X)\cos\alpha + (Y-m_Y)\sin\alpha,$$
$$W = -(X-m_X)\sin\alpha + (Y-m_Y)\cos\alpha,$$

або

$$X - m_X = S\cos\alpha - W\sin\alpha,$$
$$Y - m_Y = S\sin\alpha + W\cos\alpha.$$

Рівняння лінії рівня набуває вигляду

$$\frac{1}{\sigma_X^2}(s\cos\alpha - w\sin\alpha)^2 - \frac{2r}{\sigma_X\sigma_Y}(s\cos\alpha - w\sin\alpha)(s\sin\alpha + w\cos\alpha) +$$

$$+\frac{1}{\sigma_Y^2}(s\sin\alpha + w\cos\alpha)^2 = \lambda^2,$$

або

$$s^2\left(\frac{1}{\sigma_X^2}\cos^2\alpha - \frac{2r}{\sigma_X\sigma_Y}\sin\alpha\cos\alpha + \frac{1}{\sigma_Y^2}\sin^2\alpha\right) +$$

$$+w^2\left(\frac{1}{\sigma_X^2}\sin^2\alpha+\frac{2r}{\sigma_X\sigma_Y}\sin\alpha\cos\alpha+\frac{1}{\sigma_Y^2}\cos^2\alpha\right)+$$

$$+sw\left(-\frac{2\sin\alpha\cos\alpha}{\sigma_X^2}-\frac{2r}{\sigma_X\sigma_Y}(\cos^2\alpha-\sin^2\alpha)+\frac{2\sin\alpha\cos\alpha}{\sigma_Y^2}\right)=\lambda^2 .$$

Добираючи кут α так, щоб коефіцієнт біля sw дорівнював 0, дістаємо

$$\sin 2\alpha\left(\frac{1}{\sigma_Y^2}-\frac{1}{\sigma_X^2}\right)-\frac{2r}{\sigma_Y\sigma_X}\cos 2\alpha=0 ,$$

або

$$tg\,2\alpha=\frac{2r\sigma_X\sigma_Y}{\sigma_X{}^2-\sigma_Y{}^2} .$$

Очевидно, що

$$M[S]=0 ,\ \ M[W]=0 ,$$

$$\sigma_S^2=D[S]=M[S^2]=M[(X-m_X)^2]\cos^2\alpha+2M[(X-m_X)(Y-m_Y)]\times$$

$$\times\sin\alpha\cos\alpha+M[(Y-m_Y)^2]\sin^2\alpha=$$

$$=\sigma_X^2\cos^2\alpha+2r\sigma_X\sigma_Y\sin\alpha\cos\alpha+\sigma_Y^2\sin^2\alpha ,$$

$$\sigma_W^2=D[W]=\sigma_X^2\sin^2\alpha-2r\sigma_X\sigma_Y\sin\alpha\cos\alpha+\sigma_Y^2\cos^2\alpha .$$

Отже, після переходу до нової системи координат щільність розподілу ймовірностей набуває вигляду

$$f_{(S,W)}(s,w)=\frac{1}{2\pi\sigma_S\sigma_W}e^{-\frac{s^2}{2\sigma_S^2}-\frac{w^2}{2\sigma_W^2}}=$$

$$=\frac{1}{\sqrt{2\pi}\sigma_s}e^{-\frac{s^2}{2\sigma_S^2}}\frac{1}{\sqrt{2\pi}\sigma_W}e^{-\frac{w^2}{2\sigma_W^2}}=f_S(s)f_W(w) .$$

Таким чином після переходу до нової системи координат s, w випадкові величини S і W стають незалежними, хоч випадкові величини X і Y могли бути залежними (коли $r\neq 0$).

Зауважимо, що коли $r=0$, тоді $tg\,2\alpha=0$, тобто коли $r=0$, тоді випадкові величини X і Y незалежні, і тому немає потреби перетворювати систему координат, щоб перейти до незалежних випадкових величин.

Вправи для самостійного виконання

4.17.1. Перевірити, чи правильні твердження:

1. Якщо розподіл ймовірностей на множині значень випадкової величини X нормальний з параметрами m та σ, то й на множині

значень випадкової величини $Y = X + b$, $b = const$, розподіл ймовірностей такий самий.

2. Якщо розподіли ймовірностей на множинах значень випадкових величин X та Y нормальні, то й на множині значень випадкової величини $X + Y$ розподіл ймовірностей також нормальний.

3. Твердження, обернене до 2, є правильним.

4. Розподіл ймовірностей на множині значень випадкової величини $Y + X^2$ може бути нормальний, коли: а) на множині значень випадкової величини X розподіл ймовірностей нормальний; б) на множині значень випадкової величини X розподіл ймовірностей довільний.

5. Якщо розподіли ймовірностей на множинах значень випадкових величин X_1 та X_2 нормальні, то на множині значень випадкової величини $Z = X_1^2 + X_2^2$ розподіл ймовірностей не може бути нормальний.

6. Розподіл Максвелла – це розподіл ймовірностей суми квадратів трьох незалежних випадкових величин з нормальними розподілами ймовірностей на множинах їх значень.

7. Якщо розподіл ймовірностей на множині значень випадкової величини X нормальний з параметрами m та σ, то її центральні моменти k-го порядку: а) залежать від m і σ; б) не залежать від m та σ; в) залежать від m і не залежать від σ; г) залежать від σ і не залежать від m.

8. $f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$ тоді й тільки тоді, коли коефіцієнт кореляції r випадкових величин X та Y дорівнює нулю.

4.18. Поняття про випадкові процеси

Нехай T – деяка підмножина числової прямої. *Сукупність випадкових величин* $X=\{X_t, t\in T\}$ *називається випадковим процесом на множині* T *моментів часу*.

Якщо $T=\{t_1, t_2, ...\}$, то $X=\{X_{t_1}, X_{t_2}, ...\}$ називають *випадковим процесом з дискретним часом* або *випадковою послідовністю*. Якщо $T=[0;1]$, або $T=(-\infty;\infty)$, або $T=[0;\infty)$ і т. п., то $X=\{X_t, t\in T\}$ називають *випадковим процесом з неперервним часом*.

Тут за кожного фіксованого $t=t_0$ X_{t_0} є звичайною випадковою величиною. Якщо $X=\{X_t, t\in T\}$ – випадковий процес, то за кожного фіксованого $E\in\Omega$ функція $(X_t(E), t\in T)$ називається *реалізацією* або *траєкторією процесу*, що відповідає наслідку E. Розглянемо питання про визначення ймовірнісної міри в нескінченновимірному просторі.

Нехай X – випадковий процес з дискретним часом, тобто випадкова послідовність виду $X=\{X_{t_1}, X_{t_2}, ...\}$. Позначимо через R^∞ простір упорядкованих числових послідовностей виду $x=\{x_{t_1}, x_{t_2}, ...\}$, $-\infty<x_{t_i}<\infty$, $i=1, 2, ...$ Кожна реалізація випадкового процесу $X=\{X_{t_1}, X_{t_2}, ...\}$, яка визначається за зчисленним набором значень $\{x_{t_1}, x_{t_2}, ...\}$, яких набувають випадкові величини X_{t_i} в конкретному випробуванні, є деяка точка з простору R^∞.

Розглянемо будь-яку скінченну підмножину $\{x_{i_1}, x_{i_2}, ..., x_{i_m}\}$ координат точки $x\in R^\infty$. Такій підмножині координат точки x відповідатиме деяка точка $\tilde{x}$ в m-вимірному просторі $R_{i_1}\times R_{i_2}\times...\times R_{i_m}=R^m_{i_1 i_2 ... i_m}$. Точка $\tilde{x}$ є проекцією точки $x\in R^\infty$ на простір $R^m_{i_1 i_2 ... i_m}$. Множина W усіх точок з R^∞, яка проектується на дану множину $\tilde{W}\subset R^m_{i_1 i_2 ... i_m}$, називається *циліндричною множиною*. Якщо $\tilde{W}$ борелівська множина в $R^m_{i_1 i_2 ... i_m}$, то циліндрична множина $W\subset R^\infty$, яка проектується на $\tilde{W}$, називається *борелівською циліндричною множиною* в R^∞ з основою $\tilde{W}$.

Нехай I_k і B_k – відповідно інтервали $[a_k;b_k)$ і борелівські множини на k-й числовій прямій (з координатою x_k).

Розглянемо циліндричні множини

$$J(I_1\times \ldots \times I_n)=\{x \mid x=(x_1, x_2, \ldots),\ x_i\in I_i,\ i\in\overline{1,n}\},$$

$$J(B_1\times \ldots \times B_n)=\{x \mid x=(x_1, x_2, \ldots),\ x_i\in B_i,\ i\in\overline{1,n}\},$$

$$J(B^n)=\{x \mid x=(x_1, x_2, \ldots), x=(x_1, x_2, \ldots, x_n)\in B^n\},$$

де B^n – борелівські множини в R^n.

Кожна з циліндричних множин $J(B_1\times \ldots \times B_n)$ чи $J(B^n)$ може розглядатися також як циліндрична множина в R^{n+1}, R^{n+2}, ... , оскільки

$$J(B_1\times \ldots \times B_n)=J(B_1\times \ldots \times B_n\times R^1),$$

$$J(B^n)=J(B^n\times R^1)=J(B^{n+1}).$$

Сукупність множин, складена із скінченних сум циліндричних множин $J(I_1\times \ldots \times I_n)$, є алгеброю. Так само алгебрами є сукупності множин, складені з об'єднань циліндричних множин $J(B_1\times \ldots \times B_n)$. Система циліндричних множин $J(B^n)$ також є алгеброю. Виявляється, що σ-алгебри, породжені за такими системами циліндричних множин, співпадають. Ці σ-алгебри позначають $\mathscr{B}(R^\infty)$, а їх елементи називають борелівськими множинами в R^∞.

Якщо $\tilde{W}$ борелівська множина в $R^m_{i_1 i_2 \ldots i_m}$, а W відповідна їй циліндрична множина в R^∞, то щоб визначити ймовірність на борелівській циліндричній множині W, вважають, що

$$P(W)=P_{(X_{i_1},\ X_{i_2}, \ldots, X_{i_m})}(\tilde{W}),$$

де $P_{(X_{i_1},\ X_{i_2}, \ldots, X_{i_m})}(\tilde{W})$ – ймовірність, яка ставиться у відповідність множині W, згідно з розподілом ймовірностей на множині значень випадкового вектора $(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$, утвореного з відповідних компонент випадкового процесу $\{X_{t_1}, X_{t_2}, \ldots\}$.

Циліндричні множини в R^∞ природно вважати такими елементарними множинами, за значеннями ймовірностей попадання в які визначається ймовірнісна міра на довільних множинах з $\mathscr{B}(R^\infty)$.

Теорема (А.М. Колмогорова). Нехай $(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$ – довільний скінченний набір компонент випадкового процесу $(X_{t_1}, X_{t_2}, \ldots)$ і W – борелівська циліндрична множина в R^∞, що відповідає борелівській множині $\widetilde{W} \subset R^m_{i_1 i_2 \ldots i_m}$. Якщо покласти $P(W) = P_{(X_{i_1}, \ldots, X_{i_m})}(\widetilde{W})$, то на $\mathscr{B}(R^\infty)$ існує єдина імовірнісна міра, звуження якої на класі циліндричних множин в R^∞ є функція множини, означена за рівністю

$$P(W) = P_{(X_{i_1},\ X_{i_2}, \ldots, X_{i_m})}(\widetilde{W}) .$$

Цю теорему можна сформулювати інакше:
якщо P_1, P_2, ... – послідовність ймовірнісних мір на $(R^1, \mathscr{B}(R^1))$, $(R^2, \mathscr{B}(R^2))$, ..., причому ці міри узгоджуються, тобто $P_{n+1}(B \times R^1) = P_n(B)$, $B \in \mathscr{B}(R^n)$, то існує і притому єдина ймовірнісна міра P на $(R^\infty, \mathscr{B}(R^\infty))$ така, що за кожного $n = 1,\ 2,\ \ldots$

$$P(J_n(B)) = P_n(B),\ B \in \mathscr{B}(R^n) ,$$

де

$$J_n(B) = \{x \in R^\infty \mid (x_1,\ x_2, \ldots,\ x_n) \in B\},\ B \in \mathscr{B}(R^n) .$$

Нехай T – довільна множина на числовій прямій. Позначимо через R^T *простір дійсних функцій, визначених стосовно* $t \in T$.

Розглянемо циліндричні множини

$$J_{t_1 \ldots t_n}(I_1 \times\ \ldots\ \times I_n) = \{x \mid x_{t_i} \in I_i,\ \ i \in \overline{1,\ n}\},$$

$$J_{t_1 \ldots t_n}(B_1 \times\ \ldots\ \times B_n) = \{x \mid x_{t_i} \in B_i,\ \ i \in \overline{1,\ n}\},$$

$$J_{t_1 .. t_n}(B^n) = \{x \mid (x_{t_1},\ \ldots,\ x_{t_n}) \in B^n\},$$

де I_i – множини виду $[a_i ; b_i)$, B_i – борелівські множини на числовій прямій, B^n – борелівські множини в R^n.

Можна показати, що найменші σ-алгебри, що породжуються за розглядуваними циліндричними множинами, співпадають. Позначимо таку σ-алгебру через $\mathscr{B}(R^T)$.

Теорема. Нехай T – довільна незчисленна множина. Тоді довільна множина $A \in \mathscr{B}(R^T)$ буде такої структури: знайдеться не більш ніж зчисленна множина точок $t_1, t_2, \ldots$ з T і борелівська множина B з $\mathscr{B}(R^\infty)$ така, що $A = \{x \mid\ x = (x_{t_1}, x_{t_2}, \ldots) \in B\}$.

Таким чином, борелівська множина A з σ-алгебри $\mathscr{B}(R^T)$ визначається за обмеженнями, накладеними на функції

$x=(x_t),\ t\in T$, на не більш ніж зчисленній множині точок $t_1, t_2, \ldots$.

Звідси випливає, зокрема, що множини

$A_1=\{x \mid x_t < c$ за всіх $t\in[0;1]\}$,

$A=\{x \mid x_t = 0$ принаймні за одного $t\in[0;1]\}$,

$A=\{x \mid x_t$ неперервна в фіксованій точці $t_0\in[0;1]\}$,

що залежать від поведінки функцій на незчисленній множині точок, не обов'язково повинні бути борелівськими.

Нехай T – довільна множина індексів $t\in T$, R_t – числова пряма, відповідна індексу t. Розглянемо довільний скінченний набір $\tau=\{t_1, ..., t_n\}$ різних індексів, $t_i\in T$, $n\geq 1$, і нехай $P_{X_{t_1}\ldots X_{t_n}}$ – ймовірнісна міра на

$$(R^n_{t_1\ldots t_n}, \mathscr{B}(R^n_{t_1\ldots t_n})),\text{ де } R^n_{t_1\ldots t_n}=R_{t_1}\times\ldots\times R_{t_n}.$$

Говорять, що *сім'я ймовірнісних мір* $\{P_\tau\}$, де τ пробігає множину всіх скінченних невпорядкованих наборів, є *узгодженою*, якщо за довільних наборів $\tau=\{t_1, ..., t_n\}$ і $\sigma=\{s_1, ..., s_m\}$ таких, що $\sigma\subseteq\tau$

$$P_\sigma(\{(x_{s_1}, ..., x_{s_m}) \mid (x_{s_1}, ..., x_{s_m})\in B\})=$$
$$=P_\tau(\{(x_{t_1}, ..., x_{t_n}) \mid (x_{s_1}, ..., x_{s_m})\in B\})$$

за довільного $B\in\mathscr{B}(R^m_{s_1\ldots s_m})$.

Теорема А.М. Колмогорова (про продовження міри в $(R^T, \mathscr{B}(R^T))$). Нехай $\{P_\tau\}$ – сім'я узгоджених мір на $(R^n_{t_1\ldots t_n}, \mathscr{B}(R^n_{t_1\ldots t_n}))$. Тоді існує і притому єдина ймовірнісна міра P на $(R^T, \mathscr{B}(R^T))$ така, що

$$P(J_\tau(B))=P_\tau(B)$$

за всіх невпорядкованих наборів $\tau=\{t_1, ..., t_n\}$ різних індексів $t_i\in T$, $B\in\mathscr{B}(R^n_{t_1\ldots t_n})$, і $J_\tau(B)=\{x\in R^T \mid (x_{t_1}, ..., x_{t_n})\in B\}$.

Зауважимо, що відповідно до розглянутих раніше означень випадковий процес є випадковим елементом зі значеннями в R^T (чи в R^∞). Випадкові елементи зі значеннями в R^T називають також *випадковими функціями*.

Означення. Нехай $X=\{X_t,\ t\in T\}$ – випадковий процес. Ймовірнісна міра P_X на $(R^T, \mathscr{B}(R^T))$ така, що

$$P_X(B)=P(\{E \mid X(E)\in B\})=P(X^{-1}(B)),\ B\in\mathscr{B}(R^T),$$

називається *розподілом імовірностей випадкового процесу* X.

Розподіли ймовірностей

$$P_{t_1 \ldots t_n}(B) = P(\{E \mid (X_{t_1}, \ldots, X_{t_n}) \in B\}),$$

де $t_1 < t_2 < \ldots < t_n,\ t_i \in T$, називаються *скінченновимірними розподілами ймовірностей випадкового процесу*.

Функції

$$F_{X_{t_1} \ldots X_{t_n}}(x_1, \ldots, x_n) = P(\{E \mid X_{t_1} \in (-\infty; x_1), \ldots, X_{t_n} \in (-\infty; x_n)\}),$$

де $t_1 < t_2 < \ldots < t_n$, називаються *скінченновимірними функціями розподілу ймовірностей випадкового процесу*.

Випадковий процес вважається заданим, якщо за будь-якого набору параметрів $0 < t_1 < t_2 < \ldots < t_n$ задано многовимірний розподіл ймовірностей

$$F_{t_1 t_2 \ldots t_n}(x_1, x_2, \ldots, x_n) = P((X_{t_1} < x_1)(X_{t_2} < x_2)\ldots(X_{t_n} < x_n)),$$

причому ці розподіли між собою узгоджені, тобто якщо $k < n$, то

$$F_{t_1 \ldots t_k}(x_1, x_2, \ldots, x_k) = F_{t_1 t_2 \ldots t_n}(x_1, x_2, \ldots, x_k, +\infty, \ldots, +\infty).$$

Якщо випадкові величини $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ незалежні за будь-якого набору параметрів $0 < t_1 < t_2 < \ldots < t_n$, то випадковий процес називається *процесом з незалежними значеннями*. В такому разі

$$F_{t_1 t_2 \ldots t_n}(x_1, x_2, \ldots, x_n) = F_{t_1}(x_1) F_{t_2}(x_2) \ldots F_{t_n}(x_n).$$

Математичним сподіванням $M[X_t]$ випадкового процесу називається невипадкова (детермінована) функція $m(t)$, значення якої за фіксованого $t = t_0$ є математичним сподіванням $M[X_{t_0}]$ випадкової величини X_{t_0}.

Дисперсією $D[X_t]$ випадкового процесу X_t називається детермінована функція $\sigma^2(t) = \sigma^2[X_t] = M[(X_t - M(X_t])^2]$, значення якої за $t = t_0$ є дисперсією випадкової величини X_{t_0}.

Кореляційною (*автокореляційною*) функцією випадкового процесу X_t із скінченною дисперсією $\sigma^2(t)$ називається детермінована функція

$$B(t, s) = K[X_t, X_s] = M[(X_t - m(t))(X_s - m(s))],$$

значення якої за фіксованих $t = t_0$, $s = s_0$ є кореляційним моментом (коваріацією) двох випадкових величин X_{t_0} і X_{s_0}. Коли $t = s$, тоді

$$B(t, s) = B(t, t) = \sigma^2(t).$$

Функція

$$\rho(t, s) = \frac{B(t, s)}{\sigma(t)\sigma(s)}$$

називається *нормованою кореляційною функцією*. Якщо випадковий процес є процесом з незалежними значеннями, то $B(t,\ s) = 0$, $t \neq s$.

Випадковий процес називається *стаціонарним* (у вузькому розумінні), якщо його багатовимірні розподіли ймовірностей інваріантні щодо зсуву, тобто

$$F_{t_1+s, t_2+s, ..., t_n+s}(x_1,\ x_2,\ ...,\ x_n) = F_{t_1 t_2 ... t_n}(x_1,\ x_2,\ ...,\ x_n)\ .$$

Випадковий процес називається *стаціонарним у широкому розумінні*, якщо

$$M[X_t] = \text{const}\ ,\ \ \sigma^2[X_t] = \text{const}\ ,\ \ B(t,\ s) = R(s - t)\ ,\ \ s \geq t\ ,$$

де $R(u) = \rho(0,\ u)$.

Стаціонарний у вузькому розумінні процес є стаціонарним в широкому розумінні, але не навпаки.

Нехай деяка система в будь-який момент часу може знаходитись в одному із станів з деякої скінченної множини станів, яку називають *фазовим простором системи*, і нехай система може переходити з одного стану в інший в моменти часу 0, 1, 2, …, причому в кожен стан з будь-якого іншого система переходить випадково. Тоді дістаємо випадковий процес – послідовність випадкових величин $X_{t_0},\ X_{t_1},\ X_{t_2},\ ...,\ X_{t_n},\ ...$, кожна з яких може набувати одного із значень з множини $W = \{x_1,\ x_2,\ ...,\ x_m\}$.

Позначимо через $x(t_i) \in W$ стан, в якому система перебуває в момент t_i, $t_i = 0, 1, 2, ...$

Означення. Випадковий процес називається *ланцюгом Маркова*, якщо за всіх $n \geq 1,\ x(t_n) \in W$, виконується рівність

$$P(X_{t_{n+1}} = x(t_{n+1}) / X_{t_0} = x(t_0),\ ...,\ X_{t_n} = x(t_n)) =$$
$$= P(X_{t_{n+1}} = x(t_{n+1}) / X_{t_n} = x(t_n))$$

і крім того умовна ймовірність

$$P(X_{t_{n+1}} = x(t_{n+1}) / X_{t_0} = x(t_0),\ ...,\ X_{t_n} = x(t_n))$$

визначена, тобто $P(X_{t_0} = x(t_0),\ ...,\ X_{t_n} = x(t_n)) > 0$.

Функція $P_n(x_i,\ x_j),\ n = 0, 1, 2, ...,\ x_i \in W,\ x_j \in W$, називається *ймовірністю переходу ланцюга Маркова із стану* x_i *до стану* x_j на n-му кроці, якщо:

1) за будь-якого $n \geq 1$ і $x_i \in W$

$$\sum_{x_j \in W} P_n(x_i,\ x_j) = 1\ ,\ \ P_n(x_i,\ x_j) \geq 0\ ;$$

2) за всіх $x_i \in W$, за яких $P(X_{t_{n-1}} = x_i) > 0$,

$$P_n(x_i,\, x_j) = P(X_{t_n} = x_j \big/ X_{t_{n-1}} = x_i).$$

Ймовірність переходу завжди існує. Якщо $P(X_{t_{n-1}} = x_i) > 0$, то покладемо $P_n(x_i,\, x_j) = P(X_{t_n} = x_j \big/ X_{t_{n-1}} = x_i)$, і тоді умова 1) виконується. Якщо $P(X_{t_{n-1}} = x_i) = 0$, то $P_n(x_i,\, x_j)$ можна вибрати довільно так, щоб виконувалась умова 1).

Теорема. Якщо $P_n(x_i,\, x_j)$ – імовірність переходу ланцюга Маркова із стану x_i до стану x_j на n-му кроці, то за всіх $n > 0$ і $x(t_i) \in W,\ i = 0, 1, \ldots, n$, виконується рівність

$$P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_n} = x(t_n)) =$$
$$P(X_{t_0} = x(t_0))P_1(x(t_0),\, x(t_1)) \ldots P_n(x(t_{n-1}),\, x(t_n)).$$

Справді, в разі $n = 1$ одержуємо

$$P(X_{t_0} = x(t_0),\, X_{t_1} = x(t_1)) = P(X_{t_0} = x(t_0))P(X_{t_1} = x(t_1) \big/ X_{t_0} = x(t_0)) =$$
$$= P(X_{t_0} = x(t_0))P_1(x(t_0),\, x(t_1)),$$

якщо $P(X_{t_0} = x(t_0)) > 0$. В протилежному разі

$$P(X_{t_0} = x(t_0),\, X_{t_1} = x(t_1)) = 0 = P(X_{t_0} = x(t_0))P_1(x(t_0),\, x(t_1)).$$

Якщо $P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k)) > 0$, то на підставі означення ланцюга Маркова

$$P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k),\, X_{t_{k+1}} = x(t_{k+1})) =$$
$$= P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k))P_{k+1}(x(t_k),\, x(t_{k+1})),$$

звідки

$$P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k),\, X_{t_{k+1}} = x(t_{k+1})) =$$
$$= P(X_{t_0} = x(t_0))P_1(x(t_0),\, x(t_1)) \ldots P_k(x(t_{k-1}),\, x(t_k))P_{k+1}(x(t_k),\, x(t_{k+1}))$$

.

Якщо $P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k)) = 0$, то

$$P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_{k+1}} = x(t_{k+1})) = 0 =$$
$$= P(X_{t_0} = x(t_0),\, \ldots,\, X_{t_k} = x(t_k))P_{k+1}(x(t_k),\, x(t_{k+1})) =$$
$$= P(X_{t_0} = x(t_0))P_1(x(t_0),\, x(t_1)) \ldots P_{k+1}(x(t_k),\, x(t_{k+1})).$$

Знайдемо тепер імовірність $P_{n,m}(x_i,\, x_j)$ переходу системи в стан x_j на m-му кроці за умови, що на n-му кроці, $n < m$, система перебувала в стані x_i.

Одержуємо

$$P(X_{t_n} = x_i,\, X_{t_{n+1}} = x(t_{n+1}),\, \ldots,\, X_{t_{m-1}} = x(t_{m-1}),\, X_{t_m} = x_j) =$$

$$= P(X_{t_n} = x_i) P_{n+1}(x_i,\ x(t_{n+1})) P_{n+2}(x(t_{n+1}),\ x(t_{n+2}))\ \dots\ P_m(x(t_{m-1}),\ x_j)\,.$$

Якщо взяти суми стосовно всіх можливих значень $x(t_{n+1}),\ x(t_{n+2})),\ \dots,\ x(t_{m-1})$, то дістанемо (див. формулу повної ймовірності)

$$P(X_{t_n} = x_i,\ X_{t_m} = x_j) =$$

$$= P(X_{t_n} = x_i) \sum_{x(t_{n+1}) \in W} \dots \sum_{x(t_{m-1}) \in W} P_{n+1}(x_i, x(t_{n+1}))\ \dots\ P_m(x(t_{m-1}),\ x_j)$$

коли $P(X_{t_n} = x_i) > 0$.

Таким чином, імовірність переходу до стану x_j на m-му кроці за умови, що на n-му кроці, ($n < m$), система перебувала в стані x_i, визначається за формулою

$$P_{n,m}(x_i,\ x_j) = P(X_{t_m} = x_j \big/ X_{t_n} = x_i) =$$

$$= \sum_{x(t_{n+1}) \in W} \dots \sum_{x(t_{m-1}) \in W} P_{n+1}(x_i,\ x(t_{n+1}))\ \dots\ P_m(x(t_{m-1}),\ x_j)\,.$$

Якщо $x(t_0),\ \dots,\ x(t_n)$ такі, що $P(X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n)) > 0$, то

$$P(X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n),\ X_{t_m} = x_j) =$$

$$= \sum_{x(t_{n+1}) \in W} \dots \sum_{x(t_{m-1}) \in W} P(X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n)) \cdot$$

$$\cdot P_{n+1}(x(t_n),\ x(t_{n+1}))\ \dots\ P_m(x(t_{m-1}),\ x_j) =$$

$$= P(X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n)) P_{n,m}(x(t_n), x_j)\,,$$

тобто

$$P_{n,m}(x(t_n), x_j) = P(X_{t_m} = x_j \big/ X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n))\,.$$

Отже, за $n < m$

$$P(X_{t_m} = x_j \big/ X_{t_0} = x(t_0),\ \dots,\ X_{t_n} = x(t_n)) = P(X_{t_m} = x_j \big/ X_{t_n} = x(t_n))\,.$$

З наведених міркувань в разі $k < n < m$ випливає рівність

$$P_{k,m}(x_i,\ x_j) = \sum_{x(t_n) \in W} P_{k,n}(x_i,\ x(t_n)) P(x(t_n),\ x_j)\,.$$

Цю рівність називають *рівнянням Колмогорова-Чепмена*.

Означення. Якщо ймовірність $P_n(x_i,\ x_j)$ переходу ланцюга Маркова із стану x_i до стану x_j на n-му кроці не залежить від n, тобто за будь-якого n $P_n(x_i,\ x_j) = P(x_i,\ x_j)$, то ланцюг Маркова називається *однорідним*.

В однорідному ланцюгу Маркова ймовірність переходу до стану x_j на m-му кроці за умови, що система перебувала в стані x_i на n-му кроці, залежить лише від різниці $m-n$. Позначимо $P_{n,m}(x_i, x_j)$ через $P(x_i, x_j, m-n)$; через $\pi(1)$ позначимо матрицю $P(x_i, x_j) = P(x_i, x_j, 1)$, через $\pi(n)$ – матрицю $P(x_i, x_j, n)$. Тоді

$$\pi(n) = (\pi(1))^n .$$

Справді, з рівняння Колмогорова-Чепмена випливає

$$P(x_i, x_j, n) = \sum_{x_k \in W} P(x_i, x_k, n-1) P(x_k, x_j, 1),$$

тобто $\pi(n) = \pi(n-1)\pi(1)$. Звідси $\pi(n) = (\pi(1))^n$.

Теорема. Якщо в разі однорідного ланцюга Маркова за деякого i всі елементи матриці $\pi(i)$ додатні, то існують такі сталі числа p_s, $s = 1, 2, ..., m$, що

$$\lim_{i \to \infty} P(x_r, x_s, i) = p_s .$$

Справді

$$P(x_r, x_s, i) = \sum_{j=1}^{m} P(x_r, x_j, 1) P(x_j, x_s, i-1) \geq$$

$$\geq \min_{1 \leq j \leq m} P(x_j, x_s, i-1) \sum_{j=1}^{m} P(x_r, x_j, 1) = \min_{1 \leq j \leq m} P(x_j, x_s, i-1).$$

Оскільки нерівність виконується за довільного r, то й

$$\min_{1 \leq j \leq m} P(x_j, x_s, i) \geq \min_{1 \leq j \leq m} P(x_j, x_s, i-1),$$

тобто із зростанням i і за фіксованого s мінімальна з ймовірностей $P(x_j, x_s, i)$ не спадає.

Далі

$$P(x_r, x_s, i) = \sum_{j=1}^{m} P(x_r, x_j, 1) P(x_j, x_s, i-1) \leq$$

$$\leq \max_{1 \leq j \leq m} P(x_j, x_s, i-1) \sum_{j=1}^{m} P(x_r, x_j, 1) = \max_{1 \leq j \leq m} P(x_j, x_s, i-1).$$

Отже,

$$\max_{1\le j\le m} P(x_j,\, x_s,\, i) \le \max_{1\le j\le m} P(x_i,\, x_s,\, i-1)\,,$$

тобто максимальна з ймовірностей $P(x_j,\, x_s,\, i)$ із зростанням i і за фіксованого s не зростає.

Розглянемо різницю

$$P(x_r,\, x_s,\, i) - P(x_l,\, x_s,\, i) = \sum_{j=1}^{m} P(x_r,\, x_j,\, m) P(x_j,\, x_s,\, i-m) -$$

$$-\sum_{j=1}^{m} P(x_l,\, x_j,\, m) P(x_j,\, x_s,\, i-m) = \sum_{j=1}^{m} (P(x_r,\, x_j,\, m) - P(x_l,\, x_j,\, m)) P(x_j,\, x_s,\, i-m)\,.$$

Позначимо додатні різниці $P(x_r,\, x_j,\, m) - P(x_l,\, x_j,\, m)$ через $\beta_{rl}^{(j)}$, а абсолютні значення недодатних – через $\beta_{rl}'^{(j)}$.

Оскільки

$$\sum_{j=1}^{m} P(x_r,\, x_j,\, m) = \sum_{j=1}^{m} P(x_l,\, x_j,\, m) = 1\,,$$

то

$$\sum_{j=1}^{m} (P(x_r,\, x_j,\, m) - P(x_l,\, x_j,\, m)) = \sum_j \beta_{rl}^{(j)} - \sum_j \beta_{rl}'^{(j)} = 0\,.$$

Звідси

$$\sum_j \beta_{rl}^{(j)} = \sum_j \beta_{rl}'^{(j)} = h_{rl}\,.$$

За припущенням $P(x_r,\, x_s,\, i) > 0$ за всіх r і s, тому

$$h_{rl} = \sum_j \beta_{rl}^{(j)} < \sum_{s=1}^{m} P(x_r,\, x_s,\, i) = 1\,,$$

а отже й

$$0 \le h = \max_{1\le r,l\le m} h_{rl} < 1\,.$$

Тоді

$$|\, P(x_r,\, x_s,\, i) - P(x_l,\, x_s,\, i)\,| =$$

$$= \left| \sum_j \beta_{rl}^{(j)} P(x_j,\, x_s,\, i-m) - \sum_j \beta_{rl}'^{(j)} P(x_j,\, x_s,\, i-m) \right| \le$$

$$\le \left| \max_{1\le j\le m} P(x_j,\, x_s,\, i-m) \sum_j \beta_{rl}^{(j)} - \min_{1\le j\le m} P(x_j,\, x_s,\, i-m) \sum_j \beta_{rl}'^{(j)} \right| \le$$

$$\le h\left|\max_{1\le j\le m} P(x_j,\, x_s,\, i-m) - \min_{1\le j\le m} P(x_j,\, x_s,\, i-m)\right| \le$$

$$\le h \max_{1\le r,l\le m}\left|P(x_r,\, x_s,\, i-m) - P(x_l,\, x_s,\, i-m)\right|,$$

звідки

$$\max_{1\le r,l\le m}\left|P(x_r,\, x_s,\, i) - P(x_l,\, x_s,\, i)\right| \le h \max_{1\le r,l\le m}\left|P(x_r,\, x_s,\, i-m) - P(x_l,\, x_s,\, i-m)\right| \le$$

$$\le \ldots \le h^n \max_{1\le r,l\le m}\left|P(x_r,\, x_s,\, i-nm) - P(x_l,\, x_s,\, i-nm)\right|,$$

де $n = \left[\dfrac{i}{m}\right]$ – ціла частина дробу $\dfrac{i}{m}$.

Якщо $i \to \infty$, то й $n \to \infty$. Оскільки завжди

$$\left|P(x_r,\, x_s,\, i-nm) - P(x_l,\, x_s,\, i-nm)\right| \le 1,$$

то

$$\lim_{i\to\infty} \max_{1\le r,l\le m}\left|P(x_r,\, x_s,\, i) - P(x_l,\, x_s,\, i)\right| = 0 .$$

Можна показати, що $\sum\limits_{s=1}^{m} p_s = 1$.

Справді,

$$\sum_{s=1}^{m} p_s = \lim_{i\to\infty} \sum_{s=1}^{m} P(x_r,\, x_s,\, i) = 1 .$$

В фізичному тлумаченні зміст доведеної теореми можна подати так: імовірність того, що система перейде до стану x_s в деякий момент часу t, не залежить від того, в якому стані вона перебувала в далекому минулому.

Стани x_i, з яких система може переходити тільки до них самих, називаються *поглинаючими*. В разі поглинаючих станів $P(x_i,\, x_i) = 1$, $P(x_i,\, x_j) = 0$, коли $i \ne j$, а тому в разі ланцюгів Маркова з поглинаючими станами доведена теорема не справджується, оскільки не виконуються її умови $P(x_r,\, x_s,\, i) > 0$.

Якщо через e_i позначити вектор-рядок, в якого i-та координата дорівнює одиниці, а всі інші – нулю, то добуток $e_i \pi(1)$ є i-й рядок матриці $\pi(1)$, за яким визначається розподіл імовірностей переходів за один крок з i-го стану системи, а $e_i \pi(n)$ – розподіл імовірностей переходів за n

кроків з i-го стану системи. Якщо через $p(n)$ позначити $e_i\pi\ (n)$, то за умов доведеної теореми

$$p(n)=e_i\pi\ (n)=e_i\pi\ (n-1)\pi(1)=p(n-1)\pi(1),$$

і оскільки $\lim\limits_{n\to\infty} p(n)=\lim\limits_{n\to\infty} p(n-1)$, то стосовно граничного розподілу ймовірностей $p=p\pi(1)$, де $p=(p_1,\ p_2,\ ...,\ p_m)$ – набір граничних імовірностей.

Приклад 4.18.1. Нехай всі стани $x_1, x_2, ..., x_m$ системи рівноймовірні незалежно від того, в якому стані перебувала система на попередньому кроці, тобто $p_{ij}=\dfrac{1}{m}$ за всіх $i=1,\ 2,\ ...,\ m$, $j=1,\ 2,\ ...,\ m$. Тоді

$$\pi(2)=\pi(1)\pi(1)=\pi(1),\ \ \pi(n)=(\pi(1))^n=\pi(1),$$

тобто матриця переходів за будь-яку кількість кроків залишається такою самою, як і матриця переходу за один крок.

Той самий результат одержується і тоді, коли всі рядки матриці переходів однакові.

Якщо виконуються умови наведеної теореми, то коли k досить велике, рядки матриці $\pi(k)$ будуть практично однаковими і із зростанням n все менше різнитимуться між собою.

Марковський ланцюг називається *ланцюгом з поглинанням*, якщо в ньому є принаймні один поглинаючий стан і з кожного стану можливий перехід до поглинаючого стану (можливо не за один крок).

Приклад 4.18.2. Нехай на прямій блукає частинка, яка може знаходитися в одній з п'яти точок $x=0$, $x=1$, $x=2$, $x=3$, $x=4$. Імовірність переходу з будь-якої точки $x_i,\ i=1,\ 2,\ 3,$ до сусідньої точки ліворуч чи праворуч дорівнює $1/2$. З точок $x=0$ і $x=4$ частинка вийти не може. Тоді матриця ймовірностей переходів матиме вигляд, поданий в Табл. 4.18.1. Тут стани $x=0$ і $x=4$ є поглинаючими, $x=1$, $x=2$, $x=3$ – непоглинаючими.

Теорема. В разі Марковського ланцюга з поглинанням імовірність того, що процес закінчиться в одному з поглинаючих станів, із зростанням n прямує до 1.

Табл. 4.18.1

x_i \ x_j	0	1	2	3	4
0	1	0	0	0	0
1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
2	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0
3	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
4	0	0	0	0	1

Справді, нехай n_j – мінімальна кількість кроків, необхідна для переходу із стану x_j до якогось з поглинаючих станів, а p_j – ймовірність того, що система не перейде з x_j до поглинаючого стану. Оскільки за умовою з кожного непоглинаючого стану можливий перехід до поглинаючого, то $p_j < 1$. Нехай $n = \max_j n_j$, $p = \max_j p_j$. Оскільки ймовірність того, що система не перейде до поглинаючого стану за n кроків, не перевищує p, то ймовірність того, що система не перейде до поглинаючого стану за $2n$ кроків, не перевищує p^2. Оскільки $p < 1$, то $p^k \to 0$, і тому ймовірність того, що система не перейде до поглинаючого стану за kn кроків, прямує до нуля.

Поставимо такі запитання: 1) чому дорівнює ймовірність того, що процес закінчиться переходом у заданий поглинаючий стан? 2) скільки в середньому необхідно кроків на перехід до якогось з поглинаючих станів? 3) скільки в середньому разів процес проходить через кожен непоглинаючий стан?

Нехай процес починається з непоглинаючого стану x_i і нехай t_{ij} – середня кількість разів, які процес, що починається з x_i, проходить через x_j; із стану x_i до x_k процес переходить з імовірністю p_{ik}. Якщо x_k – поглинаючий стан, то процес не дійде до стану x_j, у протилежному разі він пройде через стан x_j в середньому t_{kj} разів. Отже,

$$t_{ij} = p_{i1}t_{1j} + p_{i2}t_{2j} + \dots + p_{i\,k-m}t_{k-m\,j}\,. \qquad (4.18.1)$$

Остання сума поширюється на всі непоглинаючі стани (m – кількість поглинаючих станів). Якщо разом з тим $i = j$, то до даної суми слід додати 1. Позначимо через Q матрицю, яка утворюється з матриці $\pi(1)$ викреслюванням усіх стовпців і рядків, що відповідають поглинаючим станам. Оскільки числа t_{ij} визначені лише стосовно непоглинаючих станів x_i та x_j, то квадратна матриця $T=(t_{ij}),\ i=\overline{1,\ k-m},\ j=\overline{1,\ k-m}$, буде того самого порядку, що й Q. Сума (4.18.1) є добутком i-го рядка матриці Q на j-й стовпчик матриці T. Якщо $i=j$, то до цієї суми слід додати 1, тобто відповідну компонентну одиничної матриці.

Таким чином,

$$T = QT + I\ ,$$

звідки $I = T(I-Q)$, $T=(I-Q)^{-1}$. Здобутий результат є відповіддю на запитання 3).

Нехай кожне проходження процесу через деякий непоглинаючий стан становить один крок. Тоді загальна кількість кроків, за які процес переходить в який-небудь непоглинаючий стан, дорівнює кількості кроків, необхідних для досягнення поглинаючого стану. Проте ця загальна кількість кроків дорівнює середній кількості проходжень процесу, що почався із стану x_i, через усі непоглинаючі стани, тобто сумі всіх елементів i-го рядка матриці T.

Якщо через t позначити вектор-стовпець з компонентами

$$t_i = \sum_{j=1}^{k-m} t_{ij} = \sum_{j=1}^{k-m}\ \sum_{r=1}^{k-m} p_{ir} t_{rj}\ ,$$

а через c – вектор-стовпець, усі компоненти якого дорівнюють 1, то вектор t можна подати у вигляді

$$t = (I-Q)^{-1} c\ .$$

Цей результат є відповіддю на запитання 2).

Нехай x_i – непоглинаючий стан, x_l – поглинаючий, b_{il} – імовірність того, що процес, який починається із стану x_i, перейде до стану x_l. Імовірність переходу з x_i до x_l за один крок дорівнює p_{il}. Якщо за перший крок процес перейде до стану x_k, то він або залишиться в стані x_k, якщо цей стан поглинаючий, або перейде до стану x_l з імовірністю b_{kl}, тобто

$$b_{il} = p_{il} + \sum_{r=1}^{k-m} p_{ir} b_{rl}\ , \tag{4.18.2}$$

де сума поширюється на всі непоглинаючі стани. Позначимо через R матрицю, яка утворюється з $\pi(1)$ викреслюванням рядків, що відповідають поглинаючим станам, і стовпців, що відповідають непоглинаючим станам. Оскільки b_{il} визначені стосовно непоглинаючого стану x_i та поглинаючого стану x_l, то матриця $B=(b_{ij})$ буде того самого порядку, що й матриця R. Таким чином (4.18.2) можна подати так:

$$B = R + QB,$$

звідки

$$B = (I - Q)^{-1} R = TR.$$

Цей результат є відповіддю на запитання 1).

Приклад 4.18.3. Нехай задано множину станів, за якими характеризуються перебування студентів у вузі з п'ятирічним терміном навчання: x_1 – першокурсник, x_2 – другокурсник, …, x_5 – п'ятикурсник, x_6 – закінчив вуз, x_7 – вчився, але не закінчив. Тоді матриця ймовірностей переходів у відповідному ланцюгу Маркова набуває вигляду (Табл. 4.18.2).

Табл. 4.18.2

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	q_1	r_1	0	0	0	0	p_1
x_2	0	q_2	r_2	0	0	0	p_2
x_3	0	0	q_3	r_3	0	0	p_3
x_4	0	0	0	q_4	r_4	0	p_4
x_5	0	0	0	0	q_5	r_5	p_5
x_6	0	0	0	0	0	1	0
x_7	0	0	0	0	0	0	1

Тут p_1 – імовірність відчислення, q_1 – імовірність залишитись на І курсі, r_1 – імовірність перейти на ІІ курс. Стани x_6 і x_7 – поглинаючі, з них система може переходити тільки в них самих.

Якщо внаслідок переставляння рядків і стовпців подати матрицю ймовірностей переходів у вигляді

$$P = \left(\frac{I_m \,|\, 0}{R \;|\, Q} \right), \qquad (4.18.3)$$

де I_m – одинична матриця розміру $m \times m$, яка відповідає поглинаючим станам; 0 – нульова матриця розміру $m \times (k-m)$; R – матриця розміру $(k-m)\times m$, Q – матриця розміру $(k-m)\times(k-m)$, а через τ_{ij} позначається випадковий час, що дорівнює сумарній кількості одиниць часу, протягом якого система перебувала в стані x_j за умови, що в початковий момент система перебувала в стані x_i, то $M[\tau_{ij}]$ можна знайти як елементи матриці

$$T = (I_{k-m} - Q)^{-1}.$$

Якщо $t_i = \sum_{j=1}^{k-m} \tau_{ij}$ – загальний час, протягом якого система перебуває в стані x_i до переходу в деякий поглинаючий стан, b_{ij} – імовірність переходу в поглинаючий стан x_j, $j = 1, 2, ..., l$, то стосовно ланцюгів Маркова з поглинаючими станами з матрицею ймовірностей переходів (4.18.3) $M[t_i]$ є компонентами вектора Tc, де c – вектор-стовпець, всі компоненти якого дорівнюють 1, b_{ij} – елементи матриці $B = TR$.

Нехай $p_i = 0{,}2$, $r_i = 0{,}7$, $q_i = 0{,}1$, рядки 6 і 7 і стовпці 6 і 7 переставляються на перше і друге місця. Тоді після переставляння рядків і стовпців матриця переходів набуде вигляду

$$P = \left(\frac{I \mid 0}{R \mid Q} \right) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.7 & 0.1 & 0 & 0 & 0 & 0 \\ 0.2 & 0 & 0.7 & 0.1 & 0 & 0 & 0 \\ 0.2 & 0 & 0 & 0.7 & 0.1 & 0 & 0 \\ 0.2 & 0 & 0 & 0 & 0.7 & 0.1 & 0 \\ 0.2 & 0 & 0 & 0 & 0 & 0.7 & 0.1 \end{bmatrix},$$

$$Q = \begin{pmatrix} 0.1 & 0 & 0 & 0 & 0 \\ 0.7 & 0.1 & 0 & 0, & 0 \\ 0 & 0.7 & 0.1 & 0 & 0 \\ 0 & 0 & 0.7 & 0.1 & 0 \\ 0 & 0 & 0 & 0.7 & 0.1 \end{pmatrix}, \quad R = \begin{pmatrix} 0.2 & 0.7 \\ 0.2 & 0 \\ 0.2 & 0 \\ 0.2 & 0 \\ 0.2 & 0 \end{pmatrix}$$

$$T=(I_5-Q)^{-1}=\begin{pmatrix}1.11 & 0 & 0 & 0 & 0\\ 0.86 & 1.11 & 0 & 0 & 0\\ 0.67 & 0.86 & 1.11 & 0 & 0\\ 0.52 & 0.67 & 0.86 & 1.11 & 0\\ 0.41 & 0.52 & 0.67 & 0.86 & 1.11\end{pmatrix}$$

$$t=Tc=\begin{pmatrix}1.11\\ 1.97\\ 2.64\\ 3.16\\ 3..57\end{pmatrix},\qquad TR=\begin{pmatrix}0.22 & 0.78\\ 0.40 & 0.60\\ 0.53 & 0.47\\ 0.63 & 0.37\\ 0.72 & 0.28\end{pmatrix}$$

Таким чином, за даних умов ймовірність закінчити вуз для першокурсника дорівнює 0.28, для другокурсника 0.37 і т. д.

Для кожного п’ятикурсника середній час перебування у вузі (до закінчення або відчислення) – 1.11 року, для четвертокурсника – 1.97, для третьокурсника – 2.64, для другокурсника – 3,16, для першокурсника – 3.57 року.

Приклад 4.18.4. Розглянемо процес випадкових блукань (див. приклад 4.18.2). Тоді

$$Q=\begin{pmatrix}0 & 1/2 & 0\\ 1/2 & 0 & 1/2\\ 0 & 1/2 & 0\end{pmatrix},$$

$$I-Q=\begin{pmatrix}1 & -1/2 & 0\\ -1/2 & 1 & -1/2\\ 0 & -1/2 & 1\end{pmatrix},\ T=(I-Q)^{-1}=\begin{pmatrix}3/2 & 1 & 1/2\\ 1 & 2 & 1\\ 1/2 & 1 & 3/2\end{pmatrix}.$$

Отже, якщо блукання починаються із стану $x=2$, то до закінчення процесу в середньому частинка один раз потрапить у точку $x=1$, два рази – в точку $x=2$, один раз – у точку $x=3$. Тоді

$$t=Tc=\begin{pmatrix}3/2 & 1 & 1/2\\ 1 & 2 & 1\\ 1/2 & 1 & 3/2\end{pmatrix}\cdot\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}=\begin{pmatrix}3\\ 4\\ 3\end{pmatrix}.$$

Отже, середня кількість кроків, які пройде частинка до поглинання, починаючи із стану $x=1$, дорівнює 3, із стану $x=2$ дорівнює 4, із стану $x=3$ дорівнює 3.

$$R=\begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{pmatrix},$$

$$B=TR=\begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix}\cdot\begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{pmatrix}=\begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix},$$

тобто якщо частинка починає блукати з точки $x=1$, то вона з імовірністю $3/4$ переходить у поглинаючий стан $x=0$ і з імовірністю $1/4$ – у поглинаючий стан $x=4$. Якщо блукання починається з точки $x=2$, то з однаковою ймовірністю $1/2$ частинка переходить у поглинаючі стани $x=0$ або $x=4$. Якщо блукання починаються з точки $x=3$, то з імовірністю $3/4$ процес закінчиться в точці $x=4$ і з імовірністю $1/4$ – в точці $x=0$.

Вправи для самостійного виконання

4.18.1. Перевірити чи правильні твердження:

1. Кожну випадкову величину можна вважати випадковим процесом на множині часу T, що складається з одного елемента.
2. Якщо X – випадковий вектор, то X – випадковий процес.
3. Твердження, обернене до 2, правильне.
4. Кожен випадковий процес є випадковою послідовністю.
5. Кожен випадковий процес є або випадковою послідовністю, або процесом з неперервним часом.
6. Кожному випадковому процесові відповідає безліч реалізацій.
7. Кожна реалізація випадкового процесу з дискретним часом є точкою з простору R^{∞}.
8. Простір R^{∞} – це сукупність всеможливих нескінченних послідовностей з попарно різними числами.
9. Кожна циліндрична множина простору R^{∞} складається з точок, в яких перші m координат є фіксованими числами, а всі інші координати змінюються в межах від $-\infty$ до $+\infty$.
10. В просторі R^{∞} існує єдина борелівська циліндрична множина з даною основою.
11. Основою кожної борелівської циліндричної множини є одна і та сама борелівська множина.
12. В просторі R^{T} міститься кожна дійсна функція.
13. $\mathscr{B}(R^{T})$ – це довільна σ-алгебра підмножин множини R^{T}.
14. Кожна сім’я ймовірнісних мір $\{P_{\tau}\}$ є узгодженою.

15. Кожен випадковий процес є випадковим елементом.

16. Кожен випадковий елемент є випадковою функцією.

17. Кожному випадковому процесові відповідає єдиний розподіл ймовірностей.

18. Розподіл ймовірностей довільного випадкового процесу є скінченновимірний.

19. Задання випадкового процесу здійснюється через задання системи скінченновимірних функцій розподілу ймовірностей.

20. Кожен випадковий процес є процесом з незалежними значеннями.

21. Існує математичне сподівання кожного випадкового процесу.

22. Математичне сподівання довільного випадкового процесу є а) фіксованим числом ; б) фіксованою випадковою функцією.

23. Існує дисперсія кожного випадкового процесу.

24. Кореляційна функція: а) існує за будь якого випадкового процесу ; б) може набувати довільних дійсних значень.

25. Нормована кореляційна функція може набувати довільних дійсних значень.

26. Кожен випадковий процес є стаціонарним.

27. Стаціонарність випадкового процесу рівносильна його стаціонарності в широкому розумінні.

28. Кожен випадковий процес є ланцюгом Маркова.

29. Ланцюг Маркова є випадковим процесом з дискретним часом.

30. Ланцюг Маркова є випадковим процесом із скінченною кількістю реалізацій.

31. Стан ланцюга Маркова – це його певна реалізація або траєкторія.

32. Стосовно довільного ланцюга Маркова завжди існує ймовірність $P_n(x_i, x_j)$ переходу із будь якого стану x_i в будь який стан x_j.

33. Будь який ланцюг Маркова є однорідним.

34. За довільного ланцюга Маркова ймовірність переходу до фіксованого стану x_j у деякий момент часу t не залежить від стану, в якому система перебувала в далекому минулому.

35. В однорідному ланцюгові Маркова можуть бути поглинаючі стани.

36. Кожен ланцюг Маркова є ланцюгом з поглинаннями.

РОЗДІЛ V. ЗАКОН ВЕЛИКИХ ЧИСЕЛ

5.1. Закон великих чисел стосовно статистичних ймовірностей

Нехай досить велику кількість m разів проведено n випробувань, в кожному з яких із деякої множини Ω навмання вибирається один елемент $E \in \Omega$ і фіксується, належить елемент E до множини $A \subset \Omega$, чи не належить, тобто виконується умова $E \in A$, чи ні, а виконується умова $E \in \Omega \setminus A = \overline{A}$.

Можна уявити (або здійснити реально), що серія із n випробувань проводиться слідуючим чином – із кожної з n однакових скриньок, в кожній з яких r білих і s чорних кульок, навмання діставали (з наступним поверненням в скриньку) одну кульку, і фіксували, відбулася подія A – кулька біла, чи ні. Порядок, в якому перебиралися всі n скриньок, ніяк не фіксувався – кожна скринька вибиралася навмання із ще не вибраних. Таке діставання однієї кульки із кожної із n скриньок здійснювали досить велику кількість m разів – m серій із n випробувань.

«Швидше за все» статистична ймовірність (відносна частота) $P_m^*(A)$, діставання білої кульки із будь якої скриньки практично не залежатиме від того, які кульки були дістані із вже вибраних раніше скриньок.

Зауважимо, що оскільки всі n скриньок однакові, то в кожному із $m \cdot n$ випробувань може бути використана одна і та сама скринька.

Будемо розглядати кожну із m серій із n випробувань вказаного типу як одне окреме випробування із множиною можливих наслідків $\tilde{\Omega} = \Omega^n = \Omega \times \Omega \times \ldots \times \Omega$. Елементарними наслідками $\tilde{E}$ такого випробування (окремої серії із n вказаних випробувань) є впорядковані набори виду $\tilde{E} = (E_1, E_2, \ldots, E_n)$, де $E_i \in \Omega$ – елементарна подія $E \in \Omega$, що відбувається в i-му випробуванні із розглядуваних n випробувань.

Позначимо через A_i подію, яка полягає в тому, що в i-му випробуванні із вказаних n випробувань виявилось, що $E \in A \subset \Omega$ – відбувається подія A, тобто (див. §2.10)

$$A_i = \{(E_1, E_2, \ldots, E_n) \mid E_i \in A, E_j \in \Omega, j \in \overline{1,n} \setminus \{i\}\}, i \in \overline{1,n}.$$

Нехай $I_1 \subset \{1, 2, \ldots, n\}$, $I_2 \subset \{1, 2, \ldots, n\}$, $I_1 \cap I_2 = \varnothing$, $I_1 \cup I_2 = \{1, 2, \ldots, n\}$, зокрема I_1 чи I_2 може бути порожнім, але якщо $I_1 = \varnothing$, то $I_2 = \{1, 2, \ldots, n\}$, і навпаки, якщо $I_2 = \varnothing$, то $I_1 = \{1, 2, \ldots, n\}$.

Очевидно, кожен окремий наслідок $\tilde{E} \in \tilde{\Omega}$ кожної серії із n випробувань є елементом однієї із множин (див. §2.10)

$$H_0 = \overline{A}_1 \overline{A}_2 ... \overline{A}_{n-2} \overline{A}_{n-1} \overline{A}_n ,$$
$$H_1 = \overline{A}_1 \overline{A}_2 ... \overline{A}_{n-2} \overline{A}_{n-1} A_n ,$$
$$H_2 = \overline{A}_1 \overline{A}_2 ... \overline{A}_{n-2} A_{n-1} \overline{A}_n ,$$
$$H_3 = \overline{A}_1 \overline{A}_2 ... \overline{A}_{n-2} A_{n-1} A_n ,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$H_{2^n-1} = A_1 A_2 ... A_{n-2} A_{n-1} A_n .$$

Як простір $\tilde{S}$ подій $\tilde{A} \subset \tilde{\Omega}$, $\tilde{A} \in \tilde{S}$, розглядатимемо сукупність $\tilde{S}$ підмножин множини $\tilde{\Omega}$, породжену за системою паралелепіпедів H_i, $i \in \overline{0, 2^n - 1}$, тобто

$$\tilde{S} = \{\tilde{A} \mid \ \tilde{A} = \bigcup_{i \in I} H_i, I \subset \{0, 1, 2, ..., 2^n - 1\}\} .$$

Нехай за результатами проведених m серій з n випробувань визначені

$$P_m^*(A_i),\ i \in \overline{1, n},\ P_m^*\left(\prod_{i \in I_1} \overline{A}_i \prod_{j \in I_2} A_j \right),$$

і нехай виявилось, що з достатньою точністю практично мають місце рівності

$$P_m^*\left(\prod_{i \in I_1} \overline{A}_i \prod_{j \in I_2} A_j \right) = \prod_{i \in I_1} P_m^*(\overline{A}_i) \prod_{j \in I_2} P_m^*(A_j) ,$$

за довільних $I_1 \subset \{1, 2, ..., n\}$, $I_2 \subset \{1, 2, ..., n\}$, $I_1 \bigcap I_2 = \varnothing$, $I_1 \bigcup I_2 \neq \varnothing$, тобто події $\overline{A}_i$, $i \in I_1$, A_j, $j \in I_2$, виявились практично незалежними в сукупності (статистично незалежними) відносно ймовірнісної міри P_m^* (а на підставі великої кількості раніше проведених випробувань можна сподіватися, що «швидше за все» так і буде).

Розглянемо відображення X множини $\tilde{\Omega} = \Omega^n$ в дійсний n-вимірний простір R^n за правилом: елементарній події $\tilde{E} = (E_1, E_2, ..., E_n) \in \tilde{\Omega}$ ставиться у відповідність точка, координата x_i якої набуває значення 1, коли $E_i \in A$ (тобто в i-му із вказаних n випробувань виявилось, що $E \in A \subset \Omega$), і x_i набуває значення 0, коли $E_i \overline{\in} A$ (тобто в i-му із n випробувань було вибрано $E \in \Omega$ таке, що $E \overline{\in} A$).

Образом множини $\tilde{\Omega}$ за так заданого відображення X буде множина $\tilde{\Omega}_X$, що складається із 2^n точок

$$M_0 = (0, 0, ..., 0, 0, 0),$$
$$M_1 = (0, 0, ..., 0, 0, 1),$$
$$M_2 = (0, 0, ..., 0, 1, 0),$$
$$M_3 = (0, 0, ..., 0, 1, 1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$M_{2^n - 1} = (1, 1, ..., 1, 1, 1),$$

Очевидно $X^{-1}(M_i) = H_i$, $i \in \{0, 1, 2, ..., 2^n - 1\}$, $M_i \in \tilde{\Omega}_X$, $H_i \in \tilde{S}$.

Розглянемо сукупність $\tilde{S}_X$ підмножин множини $\tilde{\Omega}_X$, до якої разом з множинами $\varnothing$ і $\tilde{\Omega}_X$ включаються всі одноелементні, всі двохелементні, і т.д., всі (2^n-1)-елементні підмножини множини $\tilde{\Omega}_X$, тобто $\tilde{S}_X = \{G \mid G = \bigcup_{i \in I} M_i, I \subset \{0,1,2,...,2^n - 1\}\}$. Очевидно, за довільного $G \subset \tilde{\Omega}_X$, $G \in \tilde{S}_X$, буде $X^{-1}(G) \in \tilde{S}$, тобто функція $X(\tilde{E})$, $\tilde{E} \in \tilde{\Omega}$, є $\tilde{S} / \tilde{S}_X$ -вимірною.

Таку функцію $X(\tilde{E})$, $\tilde{E} \in \tilde{\Omega}$, називають R^n-значною $\tilde{S} / \tilde{S}_X$ -випадковою величиною стосовно просторів $(\tilde{\Omega}, \tilde{S}, P_m^*)$ і $\tilde{S}_X$, або $\tilde{S} / \tilde{S}_X$ -випадковим вектором з координатами $X_1(\tilde{E})$, $X_2(\tilde{E})$, …, $X_n(\tilde{E})$, і пишуть $X(\tilde{E}) = (X_1(\tilde{E}),\ X_2(\tilde{E}),\ \ldots,\ X_n(\tilde{E}))$. Очевидно, в розглядуваному випадку кожна координата $X_i(\tilde{E})$ випадкового вектора $X(\tilde{E})$ може набувати лише одного із двох значень – або 0, або 1, тобто $X_i(\tilde{E}) \in \{0,1\}$. Окрім того кожна координата X_i $\tilde{S} / \tilde{S}_X$ -випадкового вектора X є випадковою величиною стосовно просторів $(\tilde{\Omega}, \tilde{S}, P_m^*)$ і $\tilde{S}_X$, оскільки

$$X_i^{-1}(\{0\}) = \{(E_1, ..., E_n) \in \tilde{\Omega} \mid E_i \in \overline{A}\} = \overline{A}_i \in \tilde{S}$$
$$X_i^{-1}(\{1\}) = \{(E_1, ..., E_n) \in \tilde{\Omega} \mid E_i \in A\} = A_i \in \tilde{S}.$$

Оскільки події $\overline{A_i}$, A_j, $i \in I_1$, $j \in I_2$, статистично незалежні в сукупності відносно міри P_m^*, то і випадкові величини X_i статистично незалежні між собою відносно міри P_{mX}^*, породженої за випадковим вектором X за мірою P_m^*.

Справді

$$P_{mX}^*((X_1, X_2, ..., X_n) = (x_1, x_2, ..., x_n)) =$$

$$= P_m^*(X^{-1}(\{(x_1, x_2, ..., x_n)\})) = P_m^*(H_i),$$

де $x_i \in \{0,1\}$, $x_i = 0$, коли $i \in I_1$, $x_i = 1$, коли $i \in I_2$, $(x_1, x_2, ..., x_n)$ – одна із точок M_i, яка відповідає наборові значень x_i, $i \in I_1 \bigcup I_2$, $I_1 \bigcup I_2 = \{1, 2, ..., n\}$, $H_i = X^{-1}(M_i)$, $i \in \{0, 1, 2, ..., 2^n - 1\}$.

Оскільки за довільних $H_i = \prod_{i \in I_1} \overline{A}_i \prod_{j \in I_2} A_j$, $i \in \overline{0, 2^n - 1}$, практично мають місце рівності

$$P_m^*(H_i) = P_m^*\left(\prod_{i \in I_1} \overline{A}_i \prod_{j \in I_2} A_j\right) = \prod_{i \in I_1} P_m^*(\overline{A}_i) \prod_{j \in I_2} P_m^*(A_j) =$$

$$= \prod_{i \in I_1} P_{mX}^*(X_i = 0) \prod_{i \in I_2} P_{mX}^*(X_i = 1),$$

то за довільних $I_1 \subset \{1, 2, ..., n\}$, $I_2 \subset \{1, 2, ..., n\}$ таких, що $I_1 \bigcap I_2 = \varnothing$, $I_1 \bigcup I_2 = \{1, 2, ..., n\}$, з достатньою точністю практично матимуть місце рівності

$$P_{mX}^*((X_1, X_2, ..., X_n) = (x_1, x_2, ..., x_n)) = \prod_{i=1}^{n} P_{mX}^*(X_i = x_i),$$

де $x_i = 0$, коли $i \in I_1$, $x_i = 1$, коли $i \in I_2$.

Таким чином випадкові величини X_i між собою практично незалежні (статистично незалежні) відносно міри P_{mX}^*, породженої за випадковим вектором *X* за мірою P_m^*.

Нехай тепер

$$Z(\widetilde{E}) = \sum_{i=1}^{n} X_i(\widetilde{E})$$

випадкова величина – функція від випадкових аргументів $X_1(\widetilde{E})$, $X_2(\widetilde{E})$, …, $X_n(\widetilde{E})$, $\widetilde{E} \in \widetilde{\Omega}$.

Оскільки $X_i(\widetilde{E}) = 1$, коли $E_i \in A$, тобто коли виявилось, що в *i*-му із вказаних *n* випробувань із множини Ω було вибрано елемент *E* такий, що $E \in A$, і таким чином в *i*-му випробуванні відбулася подія *A*, то $Z(\widetilde{E}) = \sum_{i=1}^{n} X_i(\widetilde{E})$ – кількість випробувань із вказаних *n*, в яких виявилось $E \in A$, тобто кількість відбувань події *A* у вказаній серії із *n* випробувань.

Якщо розглянути сукупність *S* підмножин множини Ω, породжену за подіями *A* і $\overline{A}$, тобто $S = \{\varnothing, A, \overline{A}, \Omega\}$, і на цій сукупності задати ймовірнісну міру $P_n^*(A)$ як відношення кількості із вказаних *n* випробувань, в яких виявилось $E \in A$, до кількості всіх n випробувань у такій серії, тоді отримаємо ймовірнісний простір (Ω, S, P_n^*) .

З іншого боку $Z(\widetilde{E}) = \sum_{i=1}^{n} X_i(\widetilde{E})$ – кількість відбувань події A в n випробуваннях, тому

$$\frac{1}{n} Z(\widetilde{E}) = \frac{1}{n}\sum_{i=1}^{n} X_i(\widetilde{E}) = P_n^*(A)(\widetilde{E}) .$$

Якщо *за результатами деякої* (*досить великої*) *кількості т серій із п випробувань визначено* розподіл статистичних ймовірностей на множині $\widetilde{\Omega}_X$ значень випадкового вектора X, а отже і на множинах значень випадкових величин X_i, і *виявилось*, що з достатньою точністю розподіли статистичних ймовірностей на множинах значень випадкових величин X_i практично мають вигляд

x_j	0	1
$P_{mX_i}^*(\{x_j\})$	q_i	p_i

тоді з достатньою точністю

$$M_m^*[X_i] = 0 \cdot q_i + 1 \cdot p_i = p_i ,$$

$$D_m^*[X_i] = (0 - p_i)^2 q_i + (1 - p_i)^2 p_i = p_i^2 q_i + q_i^2 p_i = p_i q_i (p_i + q_i) = p_i q_i .$$

Звідси з врахуванням того, що випадкові величини X_i статистично незалежні відносно ймовірнісної міри P_{mX}^*, стосовно випадкової величини $Y = \frac{1}{n} Z$ одержуємо

$$M_m^*[Y] = M_m^*\left[\frac{1}{n} Z\right] = M_m^*\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = \frac{1}{n}\sum_{i=1}^{n} p_i ,$$

$$D_m^*[Y] = D_m^*\left[\frac{1}{n} Z\right] = \frac{1}{n^2} D_m^*[Z] = \frac{1}{n^2} D_m^*\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} D_m^*[X_i] =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} p_i q_i \le \frac{1}{4n},$$

оскільки $p_i q_i = p_i (1 - p_i) \le \frac{1}{4}$ за всіх $i \in \{1, 2, ..., n\}$.

За нерівністю П.Л. Чебишова (див. §4.11)

$$P_{mY}^*(|Y - M_m^*[Y]| \ge \varepsilon) \le \frac{D_m^*[Y]}{\varepsilon^2}$$

в розглядуваному випадку одержуємо

$$P_{mY}^*(|Y - M_m^*[Y]| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$$

або

$$P_{mY}^*(|Y - M_m^*[Y]| < \varepsilon) > 1 - \frac{1}{4n\varepsilon^2},$$

тобто

$$P_{mP_n^*}^*\left(\left|P_n^*(A) - \frac{1}{n}\sum_{i=1}^{n} p_i\right| < \varepsilon\right) > 1 - \frac{1}{4n\varepsilon^2}$$

або

$$P_{mP_n^*}^*\left(\left|P_n^*(A) - \frac{1}{n}\sum_{i=1}^{n} P_m^*(A_i)\right| < \varepsilon\right) \geq 1 - \frac{1}{4n\varepsilon^2}.$$

де $P_m^*(A_i)$ – статистичні ймовірності подій A_i, визначені (спостережені) за результатами досить великої кількості m серій із n випробувань («швидше за все» всі $P_m^*(A_i)$ будуть досить близькі між собою, тобто практично рівні одне одному з достатньою точністю).

Таким чином за досить великих m і n статистичні ймовірності $P_n^*(A)$ події *A*, одержані за різними m серіями із n випробувань, в переважній більшості серій із статистичною ймовірністю (відносною частотою) $P_{mP_n^*}$, як завгодно близькою до 1 за досить великих n, групуються одна біля одної в ε-околі деякої точки $P^* \in [0;1]$ і в переважній більшості серій із n випробувань відрізняються одна від одної не більше, ніж на 2ε. За таку точку $P^* \in [0;1]$ можна взяти довільну опуклу лінійну комбінацію статистичних імовірностей $P_{ni}^*(A)$, одержаних за результатами *i*-тих серій серед m серій із n випробувань кожна, $i \in \overline{1,m}$:

$$\sum_{i=1}^{m} \alpha_i P_{ni}^*(A),\ \alpha_i \geq 0,\ \sum_{i=1}^{m} \alpha_i = 1.$$

Слід зауважити, що в різних наборах m серій із n випробувань кожна, можуть бути спостережені різні значення $P_m^*(A_i)$, а в кожній *i*-й серії із n випробувань можуть бути одержані різні значення $P_{ni}^*(A)$, $i \in \overline{1,m}$. Тому неможливо точно вказати єдине число, в досить малому ε-околі якого групуються всі числа $P_{ni}^*(A)$. Таких чисел, досить близьких між собою, може бути нескінченна кількість, а тому єдиного такого числа не існує.

Таким чином доведено закон великих чисел для статистичних ймовірностей: *за досить великих* n *практично всі спостережені значення статистичної ймовірності* $P_n^*(A)$ *випадкової події* A*, одержані за різними* m *серіями із* n *випробувань кожна, досить мало відрізняються одна від одної в переважній більшості серед* m *серій із* n *випробувань кожна,* і зосереджуються навколо деякого значення $P^*(A) \in [0;1]$. За таке значення можна взяти довільну опуклу лінійну комбінацію $\sum_{i=1}^{m} \alpha_i P_{ni}^*(A)$, $\alpha_i \geq 0$, $\sum_{i=1}^{m} \alpha_i = 1$, статистичних ймовірностей $P_{ni}^*(A)$, $i \in \overline{1,m}$, одержаних за результатами кожної серед m серій із n випробувань кожна. Найчастіше як $P^*(A)$ обирають середнє арифметичне статистичних ймовірностей $P_{ni}^*(A)$, $i \in \overline{1,m}$, тобто $P^*(A) = \dfrac{1}{m}\sum_{i=1}^{m} P_{ni}^*(A)$.

Це *число* $P^*(A)$ *є узагальненою статистичною ймовірністю випадкової події* A.

Очевидно стосовно $P^*(A)$ задовільняються всі вимоги (аксіоми) щодо статистичної ймовірності, зокрема вимоги $1_p - 3_p$, а тому $P^*(A)$ є ймовірнісною мірою.

Зауважимо, що узагальнені статистичні ймовірності завжди обчислюють з певною наперед заданою точністю Δ. Якщо вдається визначити $P^*(A)$ стосовно всіх подій $A \in S$ і в межах заданої точності узгодити Ω, S і P^*, то тим самим буде побудовано ймовірнісний простір (Ω, S, P^*). Стосовно цього ймовірнісного простору будуть правильними всі твердження, що стосуються розглянутих раніше ймовірнісних просторів (Ω, S, P_n^*), якщо в них замінити P_n^* на P^*.

Слід підкреслити, що $P^*(A)$, $A \in S$, – гіпотетична ймовірнісна міра, що вводиться на основі припущення, яке базується на статистичних даних – результатах m серій із n випробувань кожна.

Наведені факти цілком узгоджуються з думкою одного з творців теорії ймовірностей Я. Бернуллі, який писав: "*Що не дано нам вивести apriori* (*тобто наперед передбачити до проведення досліду*), *те принаймні можна дістати aposteriori*, *тобто з численних спостережень результатів подібних дослідів*. *Тому можна передбачати*, *що деяке явище згодом може відбутися у стількох саме випадках*, *у скількох воно* *раніше було відмічено як таке, що відбулося за подібних умов. Цей емпіричний спосіб визначення числа випадків за спостереженнями не новий і не незвичайний ..., тому усі дотримуються його у повсякденній практиці".*

Вправи для самостійного виконання

5.1.1. Перевірити, чи правильні твердження:

1. Якщо $X_k,\ k \in \overline{1, m}$, – попарно незалежні прості випадкові величини, то $Z_m = \left| \frac{1}{m} \sum_{k=1}^{m} X_k - \frac{1}{m} \sum_{k=1}^{m} M_n^*(X_k) \right|$ також є простою випадковою величиною.

2. Випадкова величина Z_m з попереднього твердження за досить великих m не може набувати досить великих значень.

3. Статистична ймовірність того, що Z_m набуває досить великих значень, є як завгодно малою, коли m досить велике число.

4. Значення середнього арифметичного попарно незалежних простих випадкових величин з однаковими статистичними математичними сподіваннями з великою статистичною ймовірністю можуть досить сильно відхилятися від середнього арифметичного математичних сподівань цих величин.

5. Статистична ймовірність якої завгодно близькості значень середнього арифметичного попарно незалежних простих випадкових величин з однаковими статистичними математичними сподіваннями до середнього арифметичного їх математичних сподівань мало відрізняється від 1, коли кількість цих випадкових величин досить велика.

6. Статистичні ймовірності $P_m^*(A)$ і $P_n^*(A)$ можуть відрізнятися одна від одної більше, ніж на 1.

7. Якщо за кожного із k випробувань, $k \in \overline{1, m}$, подія A відбувається практично з однією і тією самою статистичною ймовірністю $P_n^*(A)$, визначеною за результатами серії із n випробувань, то $P_{mX}^*(|X - P_n^*(A)| < 0{,}1) \geq 0.9$, коли $m \geq 250$, де $X = P_m^*(A)$.

5.1.2. Нехай за досить великою кількістю *m* серій із *n* випробувань кожна визначені статистичні ймовірності $P_{ni}^*(A)$ події A – випадання герба в результаті підкидання монети, які виявилися досить близькими між собою і практично з достатньою точністю рівними числу $\widetilde{P}_n^*(A) = \frac{1}{2}$.

Стосовно заданих чисел ε і δ знайти кількість *m* серій із *n* підкидань монети таку, що матиме місце

$$P^*_{mX}\left(\left|P^*_n(A)-\frac{1}{2}\right|<\varepsilon\right)>1-\delta \text{, якщо:}$$

1. $\varepsilon = 0.1, \quad \delta = 0.08$; 2. $\varepsilon = 0.05, \quad \delta = 0.02$; 3. $\varepsilon = 0.01, \quad \delta = 0.0001$.

В даному разі кожна *i*-та серія із *n* підкидань монети розглядається як одне випробування, в якому спостерігається значення $P^*_{ni\ \text{сп}}(A)$.

5.2. Закон великих чисел в формі П.Л. Чебишова

Одним з найважливіших для аналізу й обґрунтування висновків стосовно результатів дослідження різноманітних процесів і явищ є закон *великих чисел* – основа практичних застосувань методів теорії ймовірностей і математичної статистики.

Сутність закону великих чисел полягає в тому, що узагальнений (усереднений) результат великої серії випробувань може бути з великою мірою впевненості майже точно передбачений заздалегідь, хоч результат кожного окремого випробування наперед передбачити неможливо. Ця стійкість узагальнених результатів великої серії випробувань лежить в основі застосувань методів теорії ймовірностей і математичної статистики до розв'язування практичних задач.

Нехай, наприклад, потрібно визначити ймовірність того, що із навмання взятої ділянки землі буде зібрано врожай пшениці в межах від 30 до 32 ц з 1 га. Така подія є випадковою і може як відбутися, так і не відбутися в конкретному випробуванні. Проте тут неможливо ні перелічити всі можливі випадкові фактори, від яких може залежати чи не залежати відбування розглядуваної події, ні визначити вплив кожного з таких факторів (погодних умов, способу обробітку землі, догляду за рослинами, сусідства з іншими культурами, наявності різних видів тварин, комах, птахів і т. д.).

Разом з тим, якщо взяти дані за багато років про врожай, який збирали з розглядуваної ділянки за наближено однакових умов, створюваних для проростання й визрівання рослин, визначити середній урожай за всі роки, а також статистичні ймовірності того, що одержується врожай в межах, наприклад, від 10 до 12, від 12 до 14, від 14 до 16 і т.д. центнерів з 1 га, то з великою мірою впевненості можна буде сказати, в яких межах слід чекати врожай, а також наближено визначити ймовірність того, що буде зібрано врожай у цих межах.

Наведемо ще такий приклад. Тиск газу визначається сумарним впливом молекул, що зіткнулися з пластинкою одиничної площі за одиницю часу. Кількість зіткнень молекул з пластинкою за скінченний час та сила зіткнень в кожен момент часу набувають випадкових значень. Проте таких зіткнень настільки багато, що їх сумарний вплив на пластинку весь час майже однаковий (виявити різницю цих впливів у звичайних умовах практично неможливо).

Можна навести ще багато прикладів явищ, що вивчаються в фізиці, біології та інших науках, коли наслідки окремих випробувань є непередбачуваними, оскільки залежать від надзвичайно великої кількості випадкових факторів, вплив яких на наслідки окремого випробування врахувати неможливо.

Проте є певні закономірності, яким відповідають усереднені результати великої серії випробувань, тобто усереднений результат великої серії випробувань стає у певному розумінні передбачуваним і тому закономірним.

На практиці важливими є питання, наскільки великою повинна бути серія випробувань, щоб її усереднений результат можна було вважати коректною і надійною характеристикою всієї сукупності результатів випробувань і досліджуваного явища взагалі, і наскільки коректною і надійною характеристикою всієї сукупності результатів випробувань і досліджуваного явища взагалі є усереднений результат, здобутий за результатами окремих випробувань; з якою мірою впевненості можна стверджувати, що усереднений результат серії випробувань є коректною і надійною характеристикою всієї сукупності результатів випробувань і досліджуваного явища взагалі і т. п.

До закону великих чисел саме й належать твердження, в яких доводиться, що з імовірністю, як завгодно близькою до одиниці, відбудеться подія, яка полягає в тому, що усереднений результат досить великої серії випробувань, які виконуються в наближено однакових умовах, відхиляється від заздалегідь передбачуваного результату не далі, ніж за наперед вказані межі.

Перш ніж перейти до розгляду окремих форм закону великих чисел, розглянемо деякі види збіжності послідовностей випадкових величин.

Нехай $X_1, X_2, \ldots, X_n, \ldots$ − випадкові величини, задані на деякому ймовірнісному просторі (Ω, S, P).

Означення. Послідовність випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ називається *збіжною за ймовірністю* до випадкової величини X, якщо за будь-якого $\varepsilon > 0$

$$P(\{E \mid |X_n(E) - X(E)| \ge \varepsilon\}) \to 0 \Leftrightarrow$$

$$\Leftrightarrow P(\{E \mid |X_n(E) - X(E)| < \varepsilon\}) \to 1, \text{ коли } n \to \infty.$$

Іншими словами, послідовність випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ збігається за ймовірністю до випадкової величини X, якщо якими малими не були б додатні числа $\varepsilon > 0$ і $\delta > 0$, знайдеться настільки великий номер N, що за всіх $n > N$ виконується нерівність

$$P(\{E \mid |X_n(E) - X(E)| \ge \varepsilon\}) \le \delta \Leftrightarrow$$

$$\Leftrightarrow P(\{E \mid |X_n(E) - X(E)| < \varepsilon\}) > 1 - \delta,$$

тобто ймовірність події $\{E \mid |X_n(E) - X(E)| \ge \varepsilon\}$ є як завгодно малою, а ймовірність події $\{E \mid |X_n(E) - X(E)| < \varepsilon\}$ як завгодно близька до одиниці, коли n досить великі, $n > N$. Цей вид збіжності позначають $X_n \xrightarrow{P} X$.

Означення. Послідовність випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ називається *збіжною з ймовірністю* 1 або *збіжною майже напевно* до випадкової величини X, якщо

$$P(\{E \mid X_n(E) \not\to X(E)\} = 0$$

тобто якщо ймовірність події, яка визначається за елементарними подіями $E \in \Omega$, для яких $X_n(E)$ не збігається до $X(E)$, дорівнює нулю. Цей вид збіжності позначають

$$X_n \to X \text{ (P – м.н.), або } X_n \xrightarrow{м.н.} X \text{, або } X_n \xrightarrow{м.в.} X$$

(м.н. – майже напевне, м.в. – майже всюди).

Надалі іноді ймовірність події $\{E \mid |X_n(E) - X(E)| > \varepsilon\}$ і подібні будемо позначати $P(|X_n - X| > \varepsilon)$ тощо.

Означення. Послідовність випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ називається *збіжною в середньому порядку* α, $0 < \alpha < \infty$, до випадкової величини X, якщо

$$M[|X_n - X|^{\alpha}] \to 0 \text{, коли } n \to \infty.$$

Цей вид збіжності позначають $X_n \xrightarrow{L^{\alpha}} X$. Зокрема, коли $\alpha = 1$, то збіжність називають *збіжністю в середньому*, коли $\alpha = 2$, то збіжність називають *збіжністю в середньому квадратичному* (іноді цю збіжність позначають $X = l.\ i.\ m.\ X_n$).

Означення. Послідовність випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ називається *збіжною за розподілом ймовірностей* до випадкової величини X, якщо за будь-якої обмеженої неперервної функції $\varphi(x)$

$$M[\varphi(X_n)] \to M[\varphi(X)] \text{, коли } n \to \infty.$$

В такому разі кажуть, що F_{X_n} слабо збігається до F_X.

Цей вид збіжності позначають $X_n \xrightarrow{d} X$.

Виявляється, що умова $M[\varphi(X_n)] \to M[\varphi(X)]$ еквівалентна збіжності функцій розподілу $F_{X_n}(x)$ до функції $F_X(x)$ у кожній точці x, де $F_X(x)$ – неперервна. Цим пояснюється назва даного виду збіжності.

Розглянемо деякі властивості розглянутих видів збіжності.

1°. Якщо $X_n \xrightarrow{P} X$ і $\varphi(x)$ – неперервна функція, визначена на R^1, то $\varphi(X_n) \xrightarrow{P} \varphi(X)$.

Справді, оскільки $\varphi(x)$ рівномірно неперервна на будь-якій замкненій множині $|x| \le c$, тобто $|\varphi(x_1)-\varphi(x_2)| < \varepsilon$, як тільки $|x_1 - x_2| < \delta(\varepsilon)$, то враховуючи, що

$$B_n = \{E \,|\, |\varphi(X_n)-\varphi(X)| \ge \varepsilon\} = B_n \cdot (\{E \,|\, |X| \ge c\} + \{E \,|\, |X| < c\} \subset$$
$$\subset \{E \,|\, |X| \ge c\} + B_n \cdot \{E \,|\, |X| < c\} \cdot (\{E \,|\, |X_n - X| \ge \delta\} +$$
$$+ \{E \,|\, |X_n - X| < \delta\}) \subset \{E \,|\, |X| \ge c\} + \{E \,|\, |X_n - X| \ge \delta\} + \varnothing ,$$

коли $\delta > 0$ – досить мале,

$$P(|\varphi(X_n)-\varphi(X)| \ge \varepsilon) \le P(|X| \ge c) + P(|X_n - X| \ge \delta).$$

Проте ймовірність

$$P(|X| \ge c) = 1 - P(|X| < c) = 1 - F_{|X|}(c) \to 0, \ (c \to +\infty),$$

а тому за досить великого c стає як завгодно малою. Тому

$$\lim_{n\to\infty} P(|\varphi(X_n) - \varphi(X)| > \varepsilon) = 0 .$$

2°. Якщо задано m послідовностей випадкових величин $\{X_n^{(i)}\}$, $i=1,2,\ldots,m$, причому $X_n^{(i)} \xrightarrow{P} X^{(i)}$, $\Phi(x_1, x_2, \ldots, x_m)$ – неперервна функція, визначена на R^m, то

$$\Phi(X_n^{(1)},\ldots, X_n^{(m)}) \xrightarrow{P} \Phi(X^{(1)},\ldots, X^{(m)}) .$$

Доведення аналогічне до попереднього, воно випливає з нерівності

$$P(|\Phi(X_n^{(1)},\ldots, X_n^{(m)}) - \Phi(X^{(1)},\ldots, X^{(m)})| > \varepsilon) \le$$
$$\le \sum_{i=1}^{m} (P(|X^{(i)}| > c) + P(|X_n^{(i)} - X^{(i)}| > \delta)).$$

3°. Якщо $X_n \xrightarrow{P} X$ і $P(|X_n| < c) = 1$, коли $c > 0$ та всі $n \ge n_0$, то $\lim\limits_{n\to\infty} M[X_n] = M[X]$ і, крім того, $P(|X| < c) = 1$.

Справді, якщо $f(x)$ – неперервна функція така, що $f(x) = 0$, коли $|x| < c$, $f(x) > 0$, коли $|x| > c$, то враховуючи умови і властивість 1°, одержуємо:

$$P(f(X_n) = 0) = 1, \ f(X_n) \xrightarrow{P} f(X),$$
$$P(f(X) = 0) = 1, \ P(|X| < c) = 1.$$

Таким чином,

$$|M[X_n] - M[X]| \le M[|X_n - X|] \le \delta P(|X_n - X| < \delta) +$$
$$+ 2cP(|X_n - X| \ge \delta) \le \delta + 2cP(|X_n - X| > \delta).$$

Тому $\overline{\lim} |M[X_n] - M[X]| < \delta$ за як завгодно малого $\delta > 0$. Звідси, спрямовуючи δ до 0, дістанемо

$$\overline{\lim_{n\to\infty}} |M[X_n] - M[X]| = 0, \text{ тобто } \lim_{n\to\infty} M[X_n] = M[X].$$

4°. Якщо $X_n \xrightarrow{L^1} X$ (в середньому), то $X_n \xrightarrow{P} X$ (за ймовірністю).

Це випливає з нерівності Чебишова, згідно з якою за довільного $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \le \frac{1}{\varepsilon} M[|X_n - X|].$$

5°. Якщо $X_n \xrightarrow{L^2} X$, то $X_n \xrightarrow{L^1} X$ і, таким чином,

$$X_n \xrightarrow{P} X.$$

Це випливає з нерівності Ієнсена

$$(M[|X_n - X|])^2 \le M[|X_n - X|^2].$$

6°. Якщо $X_n \xrightarrow{P} X$, причому випадковим величинам X_i, X відповідають функції розподілу ймовірностей $F_{X_i}(x)$, $F_X(x)$, то послідовність функцій розподілу $F_{X_i}(x)$, $i = 1,2,...$, збігається до $F_X(x)$ в кожній точці неперервності функції $F_X(x)$.

Іншими словами, із збіжності за ймовірністю $X_n \xrightarrow{P} X$ послідовності випадкових величин X_1, X_2, … до випадкової величини X випливає збіжність $X_n \xrightarrow{d} X$ за розподілом ймовірностей.

Справді, нехай $F_X(x)$ − неперервна в точці x_0. Виберемо $x' < x_0$ і розглянемо події:

$$X < x',\ X_n < x_0,\ |X_n - X| \ge (x_0 - x').$$

Перша подія є частиною об'єднання другої та третьої. Звідси

$$F_X(x') = P_X(X < x') \le P_X(X_n < x_0) + P_X(|X_n - X| \ge (x_0 - x')).$$

Оскільки $X_n \xrightarrow{P} X$, то $\lim\limits_{n\to\infty} P_X(|X_n - X| \ge (x_0 - x')) = 0$. Тоді

$$F_X(x') \le P_X(X_n < x_0) = F_{X_n}(x_0), \text{ а тому}$$

$$\varliminf_{n\to\infty} F_{X_n}(x_0) \ge F_X(x').$$

Аналогічно, добираючи константу $x'' > x_0$, знаходимо

$$F_X(x'') \ge \varlimsup_{n\to\infty} F_{X_n}(x_0).$$

Таким чином,

$$F_X(x') \le \varliminf_{n\to\infty} F_{X_n}(x_0) \le \varlimsup_{n\to\infty} F_{X_n}(x_0) \le F_X(x'').$$

Оскільки $F_X(x)$ − неперервна в точці x_0, то

$$\lim_{x' \to x_0} F_X(x') = \lim_{x'' \to x_0} F_X(x'') = F_X(x_0) .$$

Отже,

$$\underline{\lim}_{n\to\infty} F_{X_n}(x_0) = \overline{\lim}_{n\to\infty} F_{X_n}(x_0) = \lim_{n\to\infty} F_{X_n}(x_0) = F_X(x_0) .$$

Зазначимо, що $X_n \xrightarrow{м.н.} X$ означає

$$P\left(\bigcap_{k=1}^{\infty}\bigcup_{s=1}^{\infty}\bigcap_{n=s}^{\infty}\left\{| X_n - X |\le \frac{1}{k}\right\}\right) = 1 .$$

Оскільки рівність $P\left(\bigcap_{k=1}^{\infty} A_k\right) = 1$ рівносильна тому, що $P(A_k) = 1$ за всіх k, то щоб мати збіжність $X_n \xrightarrow{м.н.} X$, необхідно й достатньо, щоб за всіх k виконувалася рівність

$$P\left(\bigcup_{s=1}^{\infty}\bigcap_{n=s}^{\infty}\left\{| X_n - X |\le \frac{1}{k}\right\}\right) = 1 .$$

Через те, що $\bigcap_{n=s}^{\infty}\left\{| X_n - X |\le \frac{1}{k}\right\}$ − зростаюча послідовність подій, то остання рівність виконується тоді й тільки тоді, коли за всіх k

$$\lim_{s\to\infty} P\left(\bigcap_{n=s}^{\infty}\left\{| X_n - X |\le \frac{1}{k}\right\}\right) = 1.$$

Отже, правильна теорема:

Теорема. Щоб послідовність випадкових величин X_n збігалася з ймовірністю 1 до випадкової величини X, необхідно й достатньо, щоб за довільного $\varepsilon > 0$ виконувалася рівність

$$\lim_{s\to\infty} P\left(\bigcap_{n=s}^{\infty}\left\{| X_n - X | \le \varepsilon\right\}\right) = 1 ,$$

що є рівносильною рівності

$$\lim_{s\to\infty} P\left(\bigcup_{n=s}^{\infty}\left\{| X_n - X | > \varepsilon\right\}\right) = 0 .$$

Оскільки $P\left(\bigcup_{n=s}^{\infty}\left\{| X_n - X | > \varepsilon\right\}\right) \le \sum_{n=s}^{\infty} P\left(\left\{| X_n - X | > \varepsilon\right\}\right)$, то з останньої теореми випливає, що для того щоб мала місце збіжність $X_n \xrightarrow{м.н.} X$, досить, щоб за довільного $\varepsilon > 0$ виконувалася умова

$$\sum_{n=1}^{\infty} P(\{|X_n - X| > \varepsilon\}) < \infty .$$

Для дослідження збіжності з ймовірністю 1 послідовностей випадкових величин важливою є така лема.

Лема (Бореля – Кантеллі). Нехай A_k – послідовність подій. Щоб у цій послідовності з ймовірністю 1 відбулися події A_k у скінченній кількості, достатньо, а у випадку незалежності A_k і необхідно, щоб виконувалась умова

$$\sum_{k=1}^{\infty} P(A_k) < \infty .$$

Справді, нехай $\sum_{k=1}^{\infty} P(A_k) < +\infty$, а подія C полягає в тому, що відбувається нескінченна кількість подій A_k. Тоді $C = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$, а тому $C \subset \bigcup_{k=n}^{\infty} A_k$ за довільного n. Звідси

$$P(C) \le \sum_{k=n}^{\infty} P(A_k) \to 0 \text{, коли } n \to \infty .$$

Достатність доведено.

Нехай події A_k незалежні і $P(C) = 0$. Тоді

$$\lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k \right) = 0 ,$$

тобто за деякого n

$$P\left(\bigcup_{k=n}^{\infty} A_k \right) < 1 ,$$

звідки

$$0 < P\left(\overline{\bigcup_{k=n}^{\infty} A_k} \right) = P\left(\bigcap_{k=n}^{\infty} \overline{A_k} \right) = \lim_{m\to\infty} \prod_{k=n}^{m} (1 - P(A_k)) \le \lim_{m\to\infty} e^{-\sum_{k=n}^{m} P(A_k)} = e^{-\sum_{k=n}^{\infty} P(A_k)} .$$

Отже, $\sum_{k=1}^{\infty} P(A_k) < \infty$.

Наслідок. Якщо $X_n \xrightarrow{P} X$, то знайдеться така послідовність X_{n_k}, що $X_{n_k} \xrightarrow{м.н.} X$.

Справді, за довільного k можна знайти таке n_k, що виконується нерівність

$$P\left(\left\{|X_{n_k}-X|>\frac{1}{k}\right\}\right)\le\frac{1}{k^2}.$$

Оскільки $\sum_{k=1}^{\infty}P\left(\left\{|X_{n_k}-X|>\frac{1}{k}\right\}\right)\le\sum_{k=1}^{\infty}\frac{1}{k^2}<\infty$, то за лемою Бореля-Кантеллі відбуватиметься лише скінченна кількість подій

$$\left\{|X_{n_k}-X|>\frac{1}{k}\right\}.$$

Таким чином, починаючи з деякого номера k з ймовірністю 1 виконується нерівність

$$|X_{n_k}-X|<\frac{1}{k}, \text{ тобто } X_{n_k}\xrightarrow{м.н.}X.$$

Розглянемо тепер теореми, що стосуються закону великих чисел. Це теореми, в яких стверджується, що за необмеженого збільшення числа доданків середнє арифметичне $\frac{1}{n}\sum_{i=1}^{n}X_i$ випадкових величин X_1, X_2, … збігається за ймовірністю до середнього арифметичного їх математичних сподівань $\frac{1}{n}\sum_{i=1}^{n}M[X_i]$.

Теорема Чебишова. Якщо дисперсії попарно незалежних випадкових величин X_i обмежені однією і тією самою сталою D, тобто $D[X_i]<D$ за всіх $i\in\{1,2,...,n,...\}$, то за як завгодно малого $\varepsilon>0$ ймовірність виконання нерівності $(|\overline{X}_n-\overline{M}|<\varepsilon)$, де $\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$, $\overline{M}=\frac{1}{n}\sum_{i=1}^{n}M[X_i]$, буде як завгодно близька до 1, коли n досить велике, тобто

$$\lim_{n\to\infty}P(|\overline{X}_n-\overline{M}|<\varepsilon)=1.$$

Справді, за умов теореми

$$M[\overline{X}_n]=M\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n}\sum_{i=1}^{n}M[X_i]=\overline{M},$$

$$D[\overline{X}_n]=D\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n^2}\sum_{i=1}^{n}D[X_i]\le\frac{1}{n^2}nD=\frac{D}{n}.$$

Тому, згідно з нерівністю Чебишова,

$$P(|\overline{X}_n-\overline{M}|\ge\varepsilon)\le\frac{D|\overline{X}_n|}{\varepsilon^2}\le\frac{D}{n\varepsilon^2},$$

звідки $P(|\overline{X}_n - \overline{M}| < \varepsilon) > 1 - \dfrac{D}{n\varepsilon^2} > 1 - \delta$, коли $n > \dfrac{D}{\varepsilon^2 \delta}$, де ε і δ як завгодно малі додатні числа, $(\varepsilon > 0, \delta > 0)$.

Враховуючи, що ймовірність будь-якої події не перевищує 1, дістанемо

$$\lim_{n \to \infty} P(|\overline{X}_n - \overline{M}| < \varepsilon) = 1.$$

Якщо за умов теореми Чебишова за всіх випадкових величин X_i математичні сподівання $M[X_i] = M$ однакові, то середнє арифметичне $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ випадкових величин X_i збігається за ймовірністю до їх спільного математичного сподівання, тобто, якими б малими не були $\varepsilon > 0$ і $\delta > 0$, за досить великих n виконується нерівність

$$P(|\overline{X}_n - M| < \varepsilon) > 1 - \delta.$$

Таким чином, згідно з теоремою Чебишова, середнє арифметичне $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ випадкових величин X_i практично не може набувати значень, які за досить великих n помітно відрізняються від числа $\overline{M} = \dfrac{1}{n}\sum_{i=1}^{n} M[X_i]$, хоч можливо, що деякі з випадкових величин X_i з досить великою ймовірністю можуть набувати значень, досить віддалених від $M[X_i]$.

Про близькість середнього арифметичного спостережених значень випадкової величини X, що можуть бути отримані в окремих незалежних випробуваннях, до математичного сподівання $M[X]$ вже йшлося, коли вводилося поняття математичного сподівання. Це пояснювалося міркуваннями емпіричного плану і носило інтуїтивний характер. Теорема Чебишова дає точну характеристику близькості значень випадкової величини X до $M[X]$ стосовно будь-якої випадкової величини.

Теорема Чебишова має велике практичне значення. Нехай, наприклад, вимірюється деяка фізична величина. Шуканим значенням вимірюваної величини вважають середнє арифметичне результатів кількох вимірювань, бо результат кожного окремого вимірювання є випадковим, оскільки залежить від впливу похибок приладу, температури повітря, освітлення, положення в просторі вимірювального приладу та спостерігача і т.п. Теорема Чебишова є теоретичним обґрунтуванням правомірності такого підходу до вимірювання значень величин на практиці.

Приклад 5.2.1. Використовуючи нерівність Чебишова, оцінити ймовірність того, що випадкова величина X відхилиться від свого математичного сподівання не більше ніж на $3\sigma[X]$, де $\sigma[X] = \sqrt{D[X]}$ – середнє квадратичне відхилення випадкової величини X.

За нерівністю Чебишова одержуємо

$$P(|X - M[X]| \geq 3\sigma[X]) \leq \frac{D[X]}{9\sigma^2[X]} = \frac{1}{9},$$

або

$$P(|X - M[X]| < 3\sigma[X]) > \frac{8}{9} \approx 0.89.$$

Ця оцінка придатна за будь-якого розподілу ймовірностей і тому виявляється іноді досить заниженою. Наприклад, за нормального розподілу ймовірностей

$$P(|X - M[X]| < 3\sigma[X]) > 0.997.$$

Узагальненням теореми Чебишова на випадок довільних випадкових величин є теорема Маркова.

Теорема Маркова. Якщо

$$D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} D\left[\sum_{i=1}^{n} X_i\right] \to 0, \text{ коли } n \to \infty,$$

то середнє арифметичне $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ випадкових величин X_i збігається за ймовірністю до середнього арифметичного їх математичних сподівань, тобто

$$\lim_{n\to\infty} P\left(\left| \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M[X_i] \right| < \varepsilon \right) = 1$$

за як завгодно малого $\varepsilon > 0$.

Доведення теореми аналогічне доведенню теореми Чебишова.

Однією з форм закону великих чисел є теорема Бернуллі, яка є окремим випадком теореми Чебишова.

Теорема Бернуллі. За необмеженого збільшення числа n незалежних випробувань, в кожному з яких подія A відбувається з ймовірністю $P(A) = p$ і не відбувається з ймовірністю $q = 1 - p$, статистична ймовірність $P_n^*(A)$ події A за наслідками n випробувань збігається за ймовірністю $\widetilde{P}$ до ймовірності $P(A) = p$ події A (де $\widetilde{P}$ – ймовірнісна міра, визначена на множині значень випадкової величини $P_n^*(A)$). Тобто, якими б малими не були $\varepsilon > 0$ і $\delta > 0$, за досить великих n буде виконуватися нерівність

$$\widetilde{P}(|P_n^*(A)-P(A)|<\varepsilon)>1-\delta,$$

що означає

$$\lim_{n\to\infty}\widetilde{P}(|P_n^*(A)-P(A)|<\varepsilon)=1.$$

Позначимо через μ_i кількість відбувань події A в i-му випробуванні $(i=1,2,...,n)$, а через μ_n – кількість відбувань події A в серії з n випробувань. Тоді $\mu_n=\sum_{i=1}^{n}\mu_i$, $P_n^*(A)=\dfrac{\mu_n}{n}$, $M[P_n^*(A)]=p$, $D[P_n^*(A)]=\dfrac{pq}{n}$. Оскільки $pq=p(1-p)\le\dfrac{1}{4}$, то, застосовуючи нерівність Чебишова до випадкової величини $P_n^*(A)$, одержуємо

$$\widetilde{P}(|P_n^*(A)-p|\ge\varepsilon)\le\frac{D[P_n^*(A)]}{\varepsilon^2}=\frac{pq}{n\varepsilon^2}\le\frac{1}{4n\varepsilon^2},$$

звідки

$$\widetilde{P}(|P_n^*(A)-P(A)|<\varepsilon)>1-\delta$$

за як завгодно малих $\varepsilon>0$ і $\delta>0$ і досить великих n.

Практичний зміст теореми Бернуллі полягає в тому, що за однакової ймовірності випадкової події A в усіх незалежних випробуваннях та за досить великого числа випробувань дуже малоймовірно (хоч і не виключено), що спостережувана статистична ймовірність $P_n^*(A)$ події A скільки-небудь помітно відрізнятиметься від імовірності $P(A)$ події A. Для посилення впевненості в тому, що здобуте за результатами серії із n незалежних випробувань значення $P_n^*(A)$ близьке до (можливо невідомої) ймовірності $P(A)$, слід провести кілька серій із n випробувань кожна і за наближене значення ймовірності $P(A)$ взяти середнє арифметичне здобутих значень випадкової величини $P_n^*(A)$ (див. §5.1).

Якщо випробування незалежні, але ймовірність події A не однакова в усіх випробуваннях, то ряд розподілу ймовірностей на множині значень випадкової величини μ_{ni} матиме вигляд, поданий в Табл. 5.2.1:

Табл. 5.2.1

μ_{nij}	0	1
p_{ij}	q_i	p_i

Тут $i=1,2$; $j=1,2,...,n$.

Тому

$$M[P_n^*(A)]=\frac{1}{n}\sum_{i=1}^{n}p_i\ ,\ D[P_n^*(A)]=\frac{1}{n^2}\sum_{i=1}^{n}p_iq_i\le\frac{1+1+...+1}{4n^2}\le\frac{1}{4n}$$

(оскільки $p_iq_i\le\frac{1}{4}$ за будь яких p_i та q_i таких, що $p_i+q_i=1$, $p_i\ge 0$, $q_i\ge 0$).

Застосовуючи знову нерівність Чебишова і міркуючи так, як і в ході доведення теореми Бернуллі, дістанемо *теорему Пуассона*, в якій стверджується, що за необмеженого збільшення числа незалежних випробувань таких, що подія A в i-му випробуванні відбувається з імовірністю p_i і не відбувається з імовірністю q_i, статистична ймовірність $P_n^*(A)$ події A в серії з n незалежних випробувань збігається за ймовірністю до середнього арифметичного ймовірностей події A в цій серії випробувань, тобто

$$\lim_{n\to\infty}\tilde{P}\left(\left|P_n^*(A)-\frac{1}{n}\sum_{i=1}^{n}p_i\right|<\varepsilon\right)=1$$

за як завгодно малого $\varepsilon>0$.

Приклад 5.2.2. Імовірність того, що навмання взята деталь якісна, дорівнює 0.8. Скільки треба переглянути деталей, щоб з імовірністю не меншою від 0.9 виявити процент якісних деталей, не допустивши похибку, більшу за 1%.

Позначимо через X_i кількість якісних деталей, виявлених в ході i-го випробування. Тоді розподіл імовірності на множині значень випадкової величини X_i набуватиме вигляду, поданого в Табл. 5.2.2:

Табл. 5.2.2

0	1
0.2	0.8

$$M[X_i]=0.2\cdot 0+0.8\cdot 1=0.8;$$

$$D[X_i]=(0-0.8)^2\cdot 0.2+(1-0.8)^2\cdot 0.8=0.16.$$

Статистична ймовірність появи якісних деталей серед n переглянутих є випадкова величина $\frac{1}{n}\sum_{i=1}^{n}X_i$. Використовуючи нерівність Чебишова, дістанемо

$$\tilde{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-0.8\right|\ge 0.01\right)\le\frac{0.16}{n(0.01)^2}.$$

Щоб імовірність $\tilde{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - 0.8\right| < 0.01\right)$ була не меншою від 0.9, треба, щоб виконувалась нерівність

$$\frac{0.16}{n\cdot(0.01)^2} < 0.1,$$

тобто

$$n > \frac{0.16}{0.1\cdot 0.0001} = 1600.$$

Закон великих чисел у розглянутій формі можна використати для доведення теореми Вейєрштрасса про рівномірне наближення неперервної функції поліномом.

Нехай виконується серія з n незалежних випробувань, в кожному з яких подія A відбувається з імовірністю $P(A) = x$, подія $\overline{A}$ – з імовірністю $P(\overline{A}) = 1 - x$, $x \in [0\,;1]$,

$$Y = f(P_n^*(A)) = f\left(\frac{1}{n}\mu_n\right),\ \mu_n \in \overline{0,\ n}\,.$$

Тоді

$$M[Y] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}.$$

Многочлен

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}$$

називають *многочленом Бернштейна* для функції $f(x)$.

Оскільки $f(x)$ рівномірно неперервна на $[0;1]$, то $|f(x_1) - f(x_2)| < \frac{\varepsilon}{2}$, як тільки $|x_1 - x_2| < \delta(\varepsilon)$. Крім того, $|f(x)| < c$ за деякого c і всіх $x \in [0;1]$. Тому, враховуючи, що $\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1$, дістанемо

$$|B_n(x) - f(x)| = \left|\sum_{k=0}^{n}\left(f\left(\frac{k}{n}\right) - f(x)\right) C_n^k x^k (1-x)^{n-k}\right| \le$$

$$\le \sum_{k=0}^{n}\left|f\left(\frac{k}{n}\right) - f(x)\right| C_n^k x^k (1-x)^{n-k} =$$

$$= \sum_{\left|\frac{k}{n} - x\right| \le \delta}\left|f\left(\frac{k}{n}\right) - f(x)\right| C_n^k x^k (1-x)^{n-k} + \sum_{\left|\frac{k}{n} - x\right| > \delta}\left|f\left(\frac{k}{n}\right) - f(x)\right| C_n^k x^k (1-x)^{n-k} \le$$

$$\leq \frac{\varepsilon}{2} + 2c \sum_{\left|\frac{k}{n}-x\right|>\delta} C_n^k x^k (1-x)^{n-k} = \frac{\varepsilon}{2} + 2cP(|P_n^*(A) - x| > \delta) \leq$$

$$\leq \frac{\varepsilon}{2} + 2c\frac{x(1-x)}{n\delta^2} \leq \frac{\varepsilon}{2} + 2c\frac{1}{4n\delta^2} < \varepsilon$$

(з врахуванням нерівності Чебишова) за досить великих n, як завгодно малих $\varepsilon > 0$, $\delta > 0$ і довільних $x \in [0;1]$.

Таким чином, доведено *теорему Вейєрштрасса*: послідовність поліномів $B_n(x)$ збігається до функції $f(x)$ рівномірно відносно $x \in [0;1]$.

Якщо в твердженні, що стосується закону великих чисел, замість збіжності за ймовірністю стверджується збіжність з імовірністю 1, то говорять, що виконується *посилений закон великих чисел*.

Перш ніж розглянути твердження, що стосуються посиленого закону великих чисел, розглянемо *нерівність Колмогорова*.

Нехай X_1, X_2, …, X_n – незалежні випадкові величини, $M[X_i]=0$, $D[X_i]<\infty$, і нехай $S_k = \sum_{i=1}^{k} X_i$. Тоді виконується нерівність

$$P(\max_k |S_k| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\sum_{k=1}^{n} D[X_k],\ \varepsilon > 0.$$

Справді, розглянемо події

$$A = \{\max_k |S_k| \geq \varepsilon\},\ A_1 = \{|S_1| \geq \varepsilon\},\ A_k = \{|S_1| < \varepsilon,\ \ldots,\ |S_{k-1}| < \varepsilon,$$
$$|S_k| \geq \varepsilon\},\ 2 \leq k \leq n.$$

Тоді

$$A = A_1 + A_2 + \ldots + A_n,\ M[S_n^2] \geq M[S_n^2 I_A] = \sum_{k=1}^{n} M[S_n^2 I_{A_k}].$$

Проте

$$M[S_n^2 I_{A_k}] = M[(S_k + (X_{k+1} + \ldots + X_n))^2 I_{A_k}] =$$
$$= M[S_k^2 I_{A_k}] + 2M[S_k(X_{k+1} + \ldots + X_n) I_{A_k}] +$$
$$+ M[(X_{k+1} + \ldots + X_n)^2 I_{A_k}] \geq M[S_k^2 I_{A_k}],$$

оскільки

$$M[S_k(X_{k+1} + \ldots + X_n) I_{A_k}] = M[S_k I_{A_k}] M[X_{k+1} + \ldots + X_n] = 0,$$

згідно з припущенням про незалежність випадкових величин і про те, що $M[X_i]=0$.

Тому

$$M[S_n^2] \ge \sum_{k=1}^{n} M[S_k^2 I_{A_k}] \ge \varepsilon^2 \sum_{k=1}^{n} P(A_k) = \varepsilon^2 P(A),$$

що й треба було довести.

Теорема (посилений закон великих чисел). Нехай $\{X_k\}$ – послідовність незалежних випадкових величин, стосовно яких $M[X_k]$ та $D[X_k]$ є скінченними числами.

Якщо $\sum_{k=1}^{\infty} \frac{D[X_k]}{k^2} < \infty$, то

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - M[X_i]) = 0\right) = 1.$$

Не обмежуючи загальності міркувань, можна вважати, що $M[X_i] = 0$. Введемо позначення $Z_n = \max_{m \le 2^n} \left| \sum_{k=1}^{m} X_k \right|$. Оскільки за $2^{n-1} \le m \le 2^n$

$$\left| \frac{1}{m} \sum_{k=1}^{m} X_k \right| \le \frac{1}{2^{n-1}} \max_{m \le 2^n} \left| \sum_{k=1}^{m} X_k \right| = \frac{1}{2^{n-1}} Z_n,$$

то за нерівністю Колмогорова

$$P(2^{-n} Z_n \ge \varepsilon) = P(Z_n \ge 2^n \varepsilon) \le \frac{1}{\varepsilon^2 2^{2n}} \sum_{k=1}^{2^n} D[X_k],$$

звідки

$$\sum_{n=1}^{\infty} P(2^{-n} Z_n > \varepsilon) \le \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{2^{2n}} \sum_{k=1}^{2^n} D[X_k] \le$$

$$\le \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} D[X_k] \sum_{2^k > k} \frac{1}{4^n} \le \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} D[X_k] \frac{\frac{1}{k^2}}{1 - \frac{1}{4}} < \infty .$$

На підставі леми Бореля-Кантеллі за досить великих n

$$P\left(\frac{1}{2^n} Z_n > \varepsilon\right) = 0,$$

а отже

$$P\left(\lim_{n\to\infty} \frac{1}{2^n} Z_n = 0\right) = 1,$$

звідки

$$P\left(\lim_{m\to\infty}\left|\frac{1}{m}\sum_{k=1}^{m}X_k\right|=0\right)=1.$$

Наслідок (теорема Бореля). Нехай виконується серія з n незалежних випробувань, в кожному з яких подія A відбувається з ймовірністю $P(A)=p$ і не відбувається з ймовірністю $P(\overline{A})=q=1-p$. Тоді

$$\tilde{P}(P_n^*(A)\to p)=1.$$

Справді, оскільки $P_n^*(A)=\frac{1}{n}\sum_{i=1}^{n}\mu_i$, де μ_i – число відбувань події A в i-му випробуванні, $M[\mu_i]=p$, $D[\mu_i]=pq$, то умови попередньої теореми виконуються, а тому

$$P_n^*(A)\xrightarrow{м.н.}P(A).$$

Останнім твердженням виправдовується статистичний підхід до визначення ймовірності за допомогою статистичної ймовірності. Однак оскільки стверджується збіжність $P_n^*(A)$ до $P(A)$ тільки з імовірністю 1, то статистичну ймовірність можна використовувати для наближеного обчислення ймовірності, але не для її логічного означення.

Вправи для самостійного виконання

5.2.1. Перевірити, чи правильні твердження:

1. Якщо за всіх $E\in\Omega$, $X_n(E)\to X(E)$, коли $n\to\infty$, то $X_n\xrightarrow{P}X$.
2. Якщо за всіх $E\in\Omega$, $X_n(E)\to X(E)$, коли $n\to\infty$, то $X_n\xrightarrow{м.н.}X$.
3. Твердження, обернене до 2, є правильним.
4. Якщо $X_n\xrightarrow{L^\alpha}X$, то $X_n\xrightarrow{м.н.}X$.
5. Якщо $X_n\xrightarrow{L^2}X$, то $X_n\xrightarrow{L^1}X$ і $X_n\xrightarrow{P}X$.
6. $X_n\xrightarrow{L^2}0\Leftrightarrow M[X_n]\to 0$ і $D[X_n]\to 0$.
7. Якщо $X_n\xrightarrow{P}X$ і $P(|X_n|\le c)=1$ за деякого $c>0$, то $X_n\xrightarrow{L^1}X$.
8. Якщо $X_n\xrightarrow{L^1}X$, то $M[X_n]\to M[X]$.
9. Якщо послідовність X_n є збіжною за ймовірністю, то такою є і послідовність $f(X_n)$ за будь-якої функції f.

10. Якщо послідовність X_n збігається за ймовірністю, то вона збігається за розподілом ймовірностей.

11. Твердження, обернене до 10, є правильним.

12. Якщо $P(\bigcap_{n=s}^{\infty}\{|X_n - X| < \varepsilon\}) \to 1,\ s \to \infty$, то $X_n \xrightarrow{м.н.} X$.

13. $\bigcup_{n=s}^{\infty}\{|X_n - X| > \varepsilon\} = \Omega - \bigcap_{n=s}^{\infty}\{|X_n - X| \le \varepsilon\}$.

14. Якщо $\sum_{n=1}^{\infty} P(\{|X_n - X| < \varepsilon\}) < +\infty$, то $X_n \xrightarrow{м.н.} X$.

15. Якщо $X_n \xrightarrow{P} X$, то $X_{nk} \xrightarrow{м.н.} X$ за будь-якої послідовності індексів $n_k \to \infty$.

16. Твердження, обернене до 15, є правильним.

17. За будь-якої послідовності випадкових величин X_n правильна рівність $\frac{1}{n}\sum_{k=1}^{n}\left(X_k - M[X_k]\right) \xrightarrow{P} 0$.

18. Твердження 17 є правильним, коли X_n є попарно незалежними випадковими величинами.

19. Твердження 17 є правильним, коли $\frac{1}{n} D[\sum_{k=1}^{n} X_k] \to 0$ за $n \to \infty$.

20. Теорема Маркова є узагальненням теореми Чебишова.

21. Теорема Бернуллі – це наслідок теореми Чебишова.

22. Нерівність А.М. Колмогорова є правильною за будь-яких випадкових величин X_k, $k \in \overline{1,n}$.

23. Якщо $\sum_{k=1}^{\infty} \frac{D(X_k)}{k^2} < +\infty$, то $\frac{1}{n}\sum_{k=1}^{n}(X_k - M[X_k]) \xrightarrow{м.н.} 0$.

24. Рівність $P(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(X_i - M(X_i)) = 0) = 1$ є правильною тоді й тільки тоді, коли $\frac{1}{n}\sum_{i=1}^{n}(X_i - M[X_i]) \xrightarrow{м.н.} 0$.

25. За будь-якої події A статистичні ймовірності $P_n^*(A)$ збігаються до ймовірності $P(A)$, коли $n \to \infty$.

5.3. Центральна гранична теорема

Центральна гранична теорема є однією з найбільш важливих теорем, що відносяться до закону великих чисел. Розглянуті раніше теореми стосувалися математичного сподівання усередненої випадкової величини, яка є середнім арифметичним великого числа випадкових величин X_i. Центральна гранична теорема стосується питання про розподіл ймовірностей на множині значень випадкової величини, яка є сумою великої кількості випадкових доданків. Сутність центральної граничної теореми полягає в тому, що коли випадкова величина X є сумою досить великого числа взаємно незалежних випадкових величин, кожна з яких не є переважаючою порівняно з іншими доданками, то розподіл ймовірностей на множині значень результуючої випадкової величини X близький до нормального і за необмеженого збільшення числа доданків цей розподіл ймовірностей необмежено наближається до нормального.

Центральна гранична теорема була встановлена в 1890 р. російським вченим О.М. Ляпуновим. Ця теорема є більш загальною за локальну та інтегральну асимптотичні теореми Муавра–Лапласа, які можна дістати як наслідки центральної граничної теореми.

Теорема 1 (Центральна гранична теорема). Нехай X_1, X_2, …, X_n, … – незалежні випадкові величини із скінченними третіми абсолютними центральними моментами $c_i^3 = M[|X_i - M[X_i]|^3]$, $i = 1,2,...$ Тоді якщо

$$\lim_{n\to\infty} \frac{\sqrt[3]{c_1^3 + c_2^3 + ... + c_n^3}}{\sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}} = 0, \qquad (5.3.1)$$

то коли $n \to \infty$, за довільного $x \in (-\infty; \infty)$ виконується співвідношення

$$\tilde{P}\left(\frac{\sum\limits_{i=1}^{n} X_i - \sum\limits_{i=1}^{n} M[X_i]}{\sqrt{\sum\limits_{i=1}^{n} \sigma_i^2[X_i]}} \in (-\infty; x) \right) \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \qquad (5.3.2)$$

З цього співвідношення слідує, що за досить великих n наближено можна вважати, що розподіл ймовірностей на множині значень випадкової величини $Y_n = \sum_{i=1}^{n} X_i$ нормальний з параметрами

$$M[Y_n] = \sum_{i=1}^{n} M[X_i],\ \sigma[Y_n] = \left(\sum_{i=1}^{n} \sigma^2[X_i]\right)^{\frac{1}{2}}.$$

Якщо розподіли ймовірностей на множинах значень випадкових величин X_i однакові з математичним сподіванням $M[X_i] = M$ і дисперсією $D[X_i] = \sigma^2[X_i] = \sigma^2$, причому існує абсолютний центральний момент $M[|X_i - M[X_i]|^3] = c^3$, то з (5.3.2) дістанемо

$$\tilde{P}\left(\frac{\sum_{i=1}^{n} X_i - nM}{\sigma\sqrt{n}} \in (-\infty, x)\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt,$$

оскільки

$$\lim_{n\to\infty} \frac{\sqrt[3]{c_1^3 + c_2^3 + ... + c_n^3}}{\sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}} = \lim_{n\to\infty} \frac{c\sqrt[3]{n}}{\sigma\sqrt{n}} = \frac{c}{\sigma} \lim_{n\to\infty} \frac{1}{n^{\frac{1}{6}}} = 0,$$

і, таким чином, умова (5.3.1) виконується.

Виявляється, що коли розподіли ймовірностей на множинах значень випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ однакові, то умову обмеженості третього центрального моменту можна зняти. В такому разі досить обмеженості $D[X_i]$. Таким чином стосовно незалежних випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ з однаковими розподілами ймовірностей на множинах їх значень правильною є центральна гранична теорема в такому формулюванні.

Теорема 2. Нехай $X_1, X_2, \ldots, X_n, \ldots$ – незалежні випадкові величини з однаковими розподілами ймовірностей на множинах їх значень, з спільним математичним сподіванням M і скінченною дисперсією σ^2. Тоді за умови $n \to \infty$ за довільних $x \in (-\infty; \infty)$

$$F_{Z_n}(x) = \tilde{P}\left(\frac{\sum_{i=1}^{n} X_i - nM}{\sigma\sqrt{n}} \in (-\infty;\ x)\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \quad (5.3.3)$$

Одна з найбільших загальних форм центральної граничної теореми була доведена О.М. Ляпуновим у 1900 р. Для її доведення О.М. Ляпунов розробив спеціальний метод характеристичних функцій.

Характеристичною функцією розподілу ймовірностей на множині значень випадкової величини X називається функція

$$g_X(t) = M[e^{itX}],$$

де $i = \sqrt{-1}$, $e^{itX} = \cos tX + i\sin tX$, $M[e^{itX}] = M[\cos tX] + iM[\sin tX]$.

Знаючи розподіл ймовірностей на множині значень випадкової величини X, можна знайти її характеристичну функцію за формулою

$$g_X(t) = \sum_{k=1}^{n} e^{itx_k} p_k \text{ або } g_X(t) = \sum_{k=1}^{\infty} e^{itx_k} p_k,$$

коли розподіл ймовірностей на множині значень випадкової величини X дискретний (поточковий), чи за формулою

$$g_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)dx \tag{5.3.4}$$

коли розподіл ймовірностей на множині значень випадкової величини X абсолютно неперервний.

Перетворення (5.3.4) називають *перетворенням Фур'є*. Як відомо, якщо $g_X(t)$ є результатом перетворення Фур'є, застосованого до $f_X(x)$, то $f_X(x)$ можна виразити через $g_X(t)$ за допомогою так званого оберненого *перетворення Фур'є*.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-itx} g_X(t)dt.$$

Приклад 5.3.1. Нехай розподіл ймовірностей на множині значень випадкової величини X нормальний зі щільністю

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Тоді

$$g_X(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2+2itx}{2}} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2+t^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} dx.$$

Можна переконатися, що похідна за змінною t від останнього інтеграла дорівнює нулю. Тому цей інтеграл не залежить від t, отже він дорівнює $\int\limits_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx = \sqrt{2\pi}$, а тому $g_X(t) = e^{-\frac{t^2}{2}}$.

Основні властивості характеристичної функції.

1°. Якщо $Y = aX$, то $g_Y(t) = g_X(at)$.

Справді, $g_Y(t) = M[e^{itY}] = M[e^{itaX}] = M[e^{i(at)X}] = g_X(at)$.

2°. Характеристична функція суми незалежних випадкових величин дорівнює добуткові характеристичних функцій доданків.

Справді, якщо $Y = X_1 + X_2 + ... + X_n$, причому випадкові величини X_i взаємно незалежні, тобто

$$f_{(X_1, X_2, ..., X_n)}(x_1, x_2, ..., x_n) = f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_n}(x_n),$$

то

$$g_Y(t) = M[e^{itY}] = M[e^{it(X_1 + X_2 + ... + X_n)}] = g_{X_1}(t) g_{X_2}(t) ... g_{X_n}(t).$$

Приклад 5.3.2. Використовуючи характеристичні функції, знайти композицію нормальних розподілів імовірностей $f_X(x)$ і $f_Y(y)$ відповідно з параметрами 0; σ_X і 0; σ_Y.

Подамо випадкову величину X у вигляді $X = \sigma_X U$, де U – нормована випадкова величина величина з нормальним розподілом ймовірностей, $M[U] = 0$, $\sigma[U] = 1$. Враховуючи приклад 5.3.1, дістаємо

$$g_X(t) = g_U(\sigma_X t) = e^{-\frac{(\sigma_X t)^2}{2}}.$$

Аналогічно $g_Y(t) = e^{-\frac{(\sigma_Y t)^2}{2}}$. Тоді стосовно $Z = X + Y$

$$g_Z(t) = e^{-\frac{\sigma_X^2 + \sigma_Y^2}{2} t^2}.$$

Остання функція є характеристичною функція розподілу ймовірностей на множині значень випадкової величини з нормальним розподілом ймовірностей з параметрами $M[Z] = 0$ і $\sigma[Z] = \sqrt{\sigma_X^2 + \sigma_Y^2}$.

Таким чином дістали композицію нормальних розподілів ймовірностей на множинах значень випадкових величин X і Y.

Доведення центральної граничної теореми для незалежних випадкових величин X_i з однаковими розподілами ймовірностей спирається на *теорему Б. Леві*:

Нехай Z_n, $n \in N$, – послідовність випадкових величин. Щоб послідовність розподілів $F_{Z_n}(x)$, $n \in N$, слабко збігалася до розподілу $F_Z(x)$, необхідно й досить, щоб характеристичні функції $g_{Z_n}(t)$ збігалися до $g_Z(t)$, ($n \to \infty$).

Нехай X_1, X_2, …, X_n, … – незалежні випадкові величини з одним і тим самим розподілом ймовірностей на множинах їх значень, математичним сподіванням m і дисперсією σ^2. Покажемо, що за необмеженого збільшення числа доданків розподіл імовірностей на множині значень випадкової величини

$$Y_n = X_1 + X_2 + ... + X_n$$

необмежено наближається до нормального.

Справді, $g_{Y_n}(t) = [q_X(t)]^n$. Подамо функцію $g_X(t)$ в околі точки $t = 0$ за формулою Маклорена у вигляді

$$g_X(t) = g_X(0) + g'_X(0)t + \left(\frac{g''_X(0)}{2} + \alpha(t) \right) t^2,$$

де $\alpha(t) \to 0$, коли $t \to 0$,

$$g_X(0) = \int_{-\infty}^{\infty} e^{i \cdot 0 \cdot x} f_X(x) dx = 1,$$

$$g'_X(0) = \int_{-\infty}^{\infty} ixe^{i \cdot 0 \cdot x} f_X(x) dx = iM[X] = im,$$

$$g''_X(0) = \int_{-\infty}^{\infty} (ix)^2 e^{i \cdot 0 \cdot x} f_X(x) dx = -\int_{-\infty}^{\infty} x^2 f_X(x) dx.$$

Якщо $m = 0$ (що не обмежує загальності міркувань), то $g''_X(0) = -\sigma^2$.

Отже,

$$g_X(t) = 1 - \left(\frac{\sigma^2}{2} - \alpha(t) \right) t^2.$$

Перейдемо від випадкової величини Y_n до нормованої випадкової величини $Z_n = \dfrac{Y_n}{\sigma\sqrt{n}}$. Тоді за умови $n \to \infty$ одержимо

$$g_{Z_n}(t) = g_{Y_n}\left(\frac{t}{\sigma\sqrt{n}} \right) = \left(g_X\left(\frac{t}{\sigma\sqrt{n}} \right) \right)^n =$$

$$= \left(1 - \left(\frac{\sigma^2}{2} - \alpha\left(\frac{t}{\sigma\sqrt{n}} \right) \right) \frac{t^2}{\sigma^2 n} \right)^n \to e^{-\frac{t^2}{2}}.$$

Звідси

$$\lim_{n\to\infty} g_{Z_n}(t) = e^{-\frac{t^2}{2}},$$

а тому, в силу теореми Б. Леві, дістаємо співвідношення (5.3.3).

Найбільш загальною необхідною й достатньою умовою, за якої залишається правильною центральна гранична теорема, є умова Ліндеберга: за будь-якого $\tau > 0$

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-m_k|>\tau B_n} |x - m_k|^2 f_{X_k}(x)dx = 0,$$

де $m_k = M[X_k]$, $B_n = \left(\sum_{k=1}^{n} D[X_k]\right)^{1/2}$, $f_{X_k}(x)$ – щільність розподілу ймовірностей на множині значень випадкової величини X_k.

На основі формули (5.3.3) можна за великих n наближено обчислювати ймовірності різних подій, пов’язаних із сумою n незалежних випадкових величин з однаковими розподілами ймовірностей на множинах їх значень, використовуючи нормальний розподіл ймовірностей.

Так, для ймовірності події $\left(\sum_{i=1}^{n} X_i \in [a\,;b]\right)$ в умовах останньої теореми одержуємо

$$\tilde{P}\left(\left(\frac{\sum_{i=1}^{n} X_i - nM}{\sigma\sqrt{n}}\right) \in [a\,;b]\right) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt.$$

Якщо перейти від випадкової величини $\sum_{i=1}^{n} X_i$ до нормованої випадкової величини (з математичним сподіванням, що дорівнює нулю, і середнім квадратичним відхиленням, що дорівнює 1), то одержимо

$$\tilde{P}\left(\sum_{i=1}^{n} X_i \in [a\,;b]\right) = \tilde{P}\left(\frac{\sum_{i=1}^{n} X_i - nM}{\sigma\sqrt{n}} \in \left[\frac{a - nM}{\sigma\sqrt{n}}\,;\frac{b - nM}{\sigma\sqrt{n}}\right]\right) \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{a-nM}{\sigma\sqrt{n}}}^{\frac{b-nM}{\sigma\sqrt{n}}} e^{-\frac{t^2}{2}} dt. \qquad (5.3.5)$$

За наявності комп'ютерних програмних засобів типу GRAN1 і ін. інтеграл

$$\frac{1}{\sqrt{2\pi}} \int\limits_{\frac{a-nM}{\sigma\sqrt{n}}}^{\frac{b-nM}{\sigma\sqrt{n}}} e^{-\frac{t^2}{2}} dt$$

досить легко визначається. Цей інтеграл можна також подати у вигляді

$$\tilde{\Phi}\left(\frac{b-nM}{\sigma\sqrt{n}}\right) - \tilde{\Phi}\left(\frac{a-nM}{\sigma\sqrt{n}}\right),$$

де $\tilde{\Phi}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \frac{1}{2} + \Phi(x)$ – функція розподілу ймовірностей на множині значень випадкової величини з центрованим і нормованим нормальним розподілом ймовірностей з параметрами $m=0$, $\sigma=1$, і скористатися таблицями значень функції Лапласа $\Phi(x)$ (див. додаток 2).

Слід зазначити, що коли у випадкових величин X_i різні математичні сподівання і дисперсії, то їх завжди можна нормувати, виконавши деяке лінійне перетворення так, щоб математичні сподівання і дисперсії здобутих в результаті випадкових величин були однаковими, причому $M=0$, $\sigma=1$.

Справді, нехай $M[X_i]$, $D[X_i]$ числові характеристики розподілів ймовірностей на множинах значень випадкових величин X_i, $i=1,2,...,n$. Кожній випадковій величині X_i поставимо у відповідність випадкову величину X_i' за правилом

$$X_i' = \frac{X_i - M[X_i]}{\sqrt{D[X_i]}}. \qquad (5.3.6)$$

Тоді

$$M[X_i'] = 0, \; D[X_i'] = 1.$$

Якщо тепер покласти $Y_n' = \frac{1}{\sqrt{n}}(X_1' + X_2' + ... + X_n')$, то згідно з центральною граничною теоремою,

$$\lim_{n\to\infty} \widetilde{P}(Y_n' \in [c\,;d]) = \frac{1}{\sqrt{2\pi}} \int_c^d e^{-\frac{t^2}{2}} dt\,.$$

Наведемо приклади застосувань центральної граничної теореми.

Нехай виконується вимірювання деякої фізичної величини. Необхідно опрацювати результати вимірювань.

В результаті будь-якого вимірювання отримується тільки наближене значення вимірюваної величини, оскільки результат вимірювання залежить від багатьох випадкових факторів (температура повітря, коливання приладу, вологість повітря та ін.). Під впливом кожного з цих факторів породжується „часткова похибка”. Проте оскільки таких факторів багато, то через їх сукупний вплив породжується вже помітна „сумарна похибка”. Розглядаючи сумарну похибку як суму великої кількості взаємно незалежних часткових похибок, можна стверджувати, що розподіл імовірностей на множині значень випадкової сумарної похибки близький до нормального. Практикою підтверджується цей висновок. Таким чином, в ході математичного опрацювання результатів вимірювань виходять із такого постулату: усереднена похибка великої кількості вимірювань є випадковою величиною з нормальним розподілом імовірностей на множині її значень.

Одним з наслідків центральної граничної теореми є ***інтегральна асимптотична теорема Муавра-Лапласа***. Нехай $\mu_n = \mu_n(A)$ – кількість відбувань події A в серії з n незалежних випробувань, в кожному з яких ймовірність події A однакова і дорівнює $P(A) = p \in (0;1)$, $q = 1 - p = P(\overline{A})$. Тоді за умови $n \to \infty$ за будь-якого $y \in (-\infty;\infty)$

$$F_n(y) = \widetilde{P}\left(\frac{\mu_n - np}{\sqrt{npq}} < y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt\,,$$

а тому за будь-яких чисел a і b, $a < b$,

$$\widetilde{P}\left(a \le \frac{\mu_n - np}{\sqrt{npq}} < b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt, \text{ коли } n \to \infty\,.$$

Дійсно, $\dfrac{\mu_n - np}{\sqrt{npq}} = \dfrac{1}{\sqrt{n}} \sum_{k=1}^{n} \dfrac{I_k - p}{\sqrt{pq}}$, де I_k – індикатор відбування події A у k-му випробуванні, а тому випадкові величини $X_k = \dfrac{I_k - p}{\sqrt{pq}}$ незалежні, $M(X_k) = 0$ і $D(X_k) = 1$. Отже, виконано всі умови центральної граничної теореми, в силу якої

$$\tilde{P}\left(\frac{\mu_n - np}{\sqrt{npq}} < y\right) = \tilde{P}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_n < y\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-t^2/2}dt, (n \to \infty).$$

З доведеної інтегральної теореми Муавра-Лапласа випливають наближені рівності:

$$\tilde{P}\left(\frac{\mu_n - np}{\sqrt{npq}} < y\right) \approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-t^2/2}dt.$$

та

$$\tilde{P}\left(a \le \frac{\mu_n - np}{\sqrt{npq}} < b\right) \approx \frac{1}{\sqrt{2\pi}}\int_{a}^{b} e^{-t^2/2}dt.$$

Виявляється, що коли n досить велике, то абсолютна похибка цих наближень не залежить відповідно від y та a і b.

Тому, вважаючи $y = \dfrac{m - np}{\sqrt{npq}}$, дістанемо

$$\tilde{P}(\mu_n < m) \approx \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y_{m,n}} e^{-t^2/2}dt, \tag{5.3.7}$$

де $y_{m,n} = \dfrac{m - np}{\sqrt{npq}}$, а тому

$$\tilde{P}(m_1 \le \mu_n \le m_2) \approx \frac{1}{\sqrt{2\pi}}\int_{y_{m_1,n}}^{y_{m_2,n}} e^{-t^2/2}dt, \tag{5.3.8}$$

де $y_{m_1,n} = \dfrac{m_1 - np}{\sqrt{npq}}$, $y_{m_2,n} = \dfrac{m_2 - np}{\sqrt{npq}}$.

В такому разі абсолютна похибка наближення (5.3.7) дорівнює

$$|R_{m,n}| \le \frac{|1-2p| \cdot |1-y_{m,n}^2|}{6\sqrt{npq}\sqrt{2\pi}} e^{-y_{m,n}^2/2} +$$

$$+\frac{(0.13 + 0.18|1-2p|)}{npq} + e^{-3\sqrt{npq}/2} \le \frac{H}{\sqrt{n}}$$

за умови, що $\sqrt{npg} \ge 5$, де $H > 0$ не залежить від m і n.

З рівності (5.3.8) випливає наступне твердження.

Локальна асимптотична теорема Муавра-Лапласа. Нехай $p = P(A)$ – імовірність події A, причому $0 < p < 1$. Тоді ймовірність того, що за умов схеми Бернуллі подія A в n

випробуваннях відбудеться рівно m разів, коли n досить велике, виражається наближено за формулою Лапласа

$$\tilde{P}_n(\mu_n = m) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{m-np}{\sqrt{npq}}\right),$$

де

$$q = 1 - p\,,\ \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,.$$

Враховуючи, що випадкова величина μ_n набуває тільки цілих значень, і обчислюючи ймовірність $\tilde{P}_n(\mu_n = m)$, дістанемо

$$\tilde{P}_n(\mu_n = m) = \tilde{P}_n(m \le \mu_n < m+1) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{m-np}{\sqrt{npq}}}^{\frac{m+1-np}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt\,.$$

Оскільки довжина відрізка $\left[\frac{m-np}{\sqrt{npq}};\frac{m+1-np}{\sqrt{npq}}\right]$ дорівнює $\frac{1}{\sqrt{npq}}$ і за досить великих n є досить малою, то всі значення функції $e^{-\frac{t^2}{2}}$ на цьому відрізку практично не відрізняються від значення в точці $t = \frac{m-np}{\sqrt{npq}}$, а тому за теоремою про середнє значення за досить великих n з досить незначною похибкою виконується рівність

$$\tilde{P}_n(\mu_n = m) = \tilde{P}_n(m \le \mu_n < m+1) \approx$$

$$\approx \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{m-np}{\sqrt{npq}}}^{\frac{m-np}{\sqrt{npq}}+\frac{1}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt \approx \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{npq}} e^{-\frac{(m-np)^2}{2npq}},$$

що стверджується в локальній асимптотичній теоремі Муавра–Лапласа.

Зазначимо, що коли відрізок $[m_1; m_2]$ симетричний щодо значення np, тобто $m_1 = np - \alpha$, $m_2 = np + \alpha$, то

$$\tilde{P}_n(np-\alpha \le \mu_n < np+\alpha) = \tilde{P}_n(|\mu_n - np| \le \alpha) \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-\frac{\alpha}{\sqrt{npq}}}^{\frac{\alpha}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt =$$

$$= 2\frac{1}{\sqrt{2\pi}} \int\limits_0^{\frac{\alpha}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt .$$

Звідси одержимо

$$\tilde{P}_n(|P_n^*(A) - p| \le \alpha) = \frac{2}{\sqrt{2\pi}} \int\limits_0^{\frac{n\alpha}{\sqrt{npq}}} e^{-\frac{t^2}{2}} dt \underset{n\to\infty}{\to} 1 \qquad (5.3.9)$$

Приклад 5.3.3. Імовірність p того, що виріб не пройде перевірку, а отже вважатиметься бракованим, дорівнює 0.2. Знайти ймовірність того, що серед 400 навмання відібраних вважатимуться бракованими від 70 до 100 виробів.

З умови випливає:

$$n = 400;\ m_1 = 70;\ m_2 = 100;\ p = 0.2;\ q = 0.8;$$

$$np = 400 \cdot 0.2 = 80;\ npq = 64;\ \sqrt{npq} = 8;$$

$$\frac{m_2 - np}{\sqrt{npq}} = \frac{100-80}{8} = 2.5;\ \ \frac{m_1 - np}{\sqrt{npq}} = \frac{70-80}{8} = -1.25 .$$

Таким чином за даних умов ймовірність того, що серед 400 виробів вважатимуться бракованими від 70 до 100 виробів, наближено дорівнює $\frac{1}{\sqrt{2\pi}} \int\limits_{-1.25}^{2.5} e^{-\frac{t^2}{2}} dt = 0.8882$.

Приклад 5.3.4. Імовірність відбування події A в кожному з n незалежних випробувань дорівнює $\frac{1}{2}$. Скільки випробувань треба провести, щоб з імовірністю 0.9 частота $P_n^*(A)$ відхилялася від імовірності $P(A)$ не більше ніж на $\alpha = 0.05$.

Застосовуючи формулу (5.3.9), знаходимо

$$\tilde{P}_n\left(\left|P_n^*(A) - \frac{1}{2}\right| \le 0.05\right) \approx \frac{2}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt = 0.9 ,$$

де $x = \alpha \frac{\sqrt{n}}{\sqrt{pq}} = 2 \cdot 0.05\sqrt{n} = 0.1\sqrt{n}$.

Розв’язавши за допомогою програми GRAN1 рівняння відносно x

$$\int_0^x \frac{2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 0.9,$$

знайдемо $x = 1.6$. Зауважимо, що в програмі GRAN1 передбачено можливість, використовуючи динамічні параметри P1, P2, …, P9, дібрати межі інтегрування a і b так, щоб інтеграл $\int_a^b f(x)dx$ набув наперед заданого значення c. Такий самий результат одержимо, використовуючи таблиці значень функції Лапласа. Отже, $0.1 \cdot \sqrt{n} = 1.6$; $\sqrt{n} = 16$, $n = 256$, тобто слід провести не менше 256 випробувань. Цю задачу можна розв’язувати і за допомогою нерівності Чебишова. Щоб виконувалася нерівність

$$\tilde{P}_n\left(\left|P_n^*(A) - \frac{1}{2}\right| < 0.05\right) \geq 0.9,$$

треба, щоб виконувалася нерівність

$$\tilde{P}_n\left(\left|P_n^*(A) - \frac{1}{2}\right| \geq 0.05\right) \leq \frac{\frac{1}{2} \cdot \frac{1}{2}}{n \cdot (0.05)^2} = \frac{0.25}{n \cdot 0.0025} \leq 0.1,$$

тобто повинно бути $n > 1000$. Як бачимо, результат досить завищений, чого й слід було сподіватися, оскільки в нерівності Чебишова не враховується конкретний вид розподілу ймовірностей. Ця нерівність виконується за будь-якого розподілу ймовірностей.

Розглянуті форми закону великих чисел можуть бути застосовані до розв’язування таких важливих для практичних застосувань задач:

1. Якщо проведено серію з n випробувань, то з якою ймовірністю усереднений результат цієї серії випробувань буде знаходитись в заданому околі теоретично очікуваного результату.

2. Якщо проведено серію з n випробувань, то в якому околі теоретично очікуваного результату буде із вказаною ймовірністю знаходитись усереднений результат?

3. Треба, щоб усереднений результат із заданою ймовірністю знаходився в заданому околі теоретично очікуваного результату. Скільки для цього треба провести випробувань?

Вправи для самостійного виконання

5.3.1. Перевірити, чи правильні твердження:

1. Якщо $M\left[\left|X-M[X]\right|^3\right]<\infty$, то і $\sigma[X]<+\infty$.

2. Якщо X_k, $k\in N$, – незалежні випадкові величини із скінченними дисперсіями, то $M[Y_n]=0$, а $D[Y_n]=1$, коли

$$Y_n=\frac{\sum_{k=1}^{n}(X_k-M[X_k])}{\sqrt{\sum_{k=1}^{n}D[X_k]}}.$$

3. Якщо Y_n з твердження 2, а $F_{Y_n}(x)$ – функція розподілу ймовірностей на множині значень випадкової величини Y_n, то

$$F_{Y_n}(x)\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt, \text{ коли } n\to\infty.$$

4. Якщо математичні сподівання і дисперсії незалежних випадкових величин X_k однакові, причому $D[X_k]<+\infty$, то

$$F_{Y_n}(x)\to\frac{1}{2\pi}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt, \text{ де } Y_n=\frac{1}{\sigma\sqrt{n}}\left(\sum_{k=1}^{n}X_k-nM\right),$$

а $M=M[X_k]$.

5. Стосовно кожної випадкової величини існує характеристична функція.

6. Характеристична функція завжди є перетворенням Фур'є відповідної щільності розподілу ймовірностей.

7. Якщо $g_X(t)=\int_{-\infty}^{\infty}e^{itx}f_X(x)dx$, то $f_X(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-itx}g_X(t)dt$.

8. Характеристична функція суми випадкових величин дорівнює сумі характеристичних функцій цих величин.

9. Стосовно будь-яких випадкових величин X_k, $k\in N$, правильна рівність

$$\tilde{P}\left(\frac{1}{\sigma\sqrt{n}}\left(\sum_{k=1}^{n}X_k-nM\right)\in[a,b]\right)\approx\frac{1}{\sqrt{2\pi}}\int_{a}^{b}e^{-\frac{t^2}{2}}dt.$$

10. Локальна асимптотична теорема Муавра-Лапласа є наслідком центральної граничної теореми.

11. $\tilde{P}\left(\left|P_n^*(A)-P(A)\right|\geq\varepsilon\right)\approx\frac{2}{\sqrt{2\pi}}\int_{\frac{n\varepsilon}{\sqrt{npq}}}^{\infty}e^{-\frac{t^2}{2}}dt\approx 0$, коли n досить велике.

РОЗДІЛ VI. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

6.1. Основні поняття й задачі математичної статистики

Досить часто на практиці доводиться вивчати процеси і явища, перебіг і прояви яких залежать від впливу численних факторів, випадкових за своєю природою, причому іноді важко з'ясувати, які саме фактори є суттєвими і які несуттєвими, їх взаємозв'язки і сукупний вплив на ті чи інші характеристики досліджуваного процесу, які залежності між цими характеристиками, як змінюються їх характеристики та взаємозв'язки під дією різних впливів, якими будуть характеристики за заданих умов і т.д.

У таких випадках для наближеного визначення певної характеристики досліджуваного явища здійснюється (якщо можливо, за одних і тих самих умов) досить велика серія спостережень чи випробувань, за результатами яких можна дістати наближені значення досліджуваних характеристик. Якщо у великій кількості випробувань виявляється певна стійкість здобутих значень, то можна говорити про певну закономірність, яка задовільняється на множині значень досліджуваної характеристики за заданих умов, і з великою мірою впевненості надалі передбачати результати аналогічних випробувань (спостережень).

Значення, що дістають в результаті кожного з таких спостережень, є випадковими і в сукупності називаються *статистичними даними.* Методи їх аналізу і встановлення закономірностей, які проявляються на множині статистичних даних, вивчаються в математичній статистиці. Разом з тим в математичній статистиці визначаються вимоги, яким повинен відповідати експеримент, щоб висновки, зроблені за його результатами, мали практичну цінність.

Таким чином, *математична статистика є наукою про планування експериментів та аналіз їх результатів*.

Методи математичної статистики безпосередньо пов'язані з ймовірнісними оцінками усереднених результатів серій випробувань. Теоретичною основою математичної статистики є теорія ймовірностей.

Слід сказати, що слово «статистика» має той самий корінь, що й слово «держава» (state) і спочатку це слово означало науку про управління. Поступово слово «статистика» стало означати збирання даних про державу, а пізніше взагалі збирання і аналіз даних.

До основних задач математичної статистики належать задачі про визначення розподілу ймовірностей на множині значень випадкової величини (чи випадкового вектора), визначення параметрів розподілу ймовірностей на основі статистичних даних.

Із задачами оцінювання параметрів розподілу ймовірностей за статистичними даними тісно пов’язані задачі про перевірку гіпотез. Для відповіді на питання, узгоджується чи ні із статистичними даними гіпотеза про те, що розподіл ймовірностей має певний вигляд або певні числові характеристики, звертаються до спеціальних методів підтвердження чи спростування гіпотез на основі відповідного аналізу статистичних даних. Крім того, в математичній статистиці розглядаються задачі про встановлення взаємозв’язків між кількома характеристиками об’єктів, узгодження спостережених значень з теоретично передбачуваними, деякі задачі теорії прийняття рішень тощо.

У подальшому вважатимемо, що спостереження (випробування) взаємно незалежні і що досліджується деяка випадкова величина X.

В результаті кожного i-го випробування із множини можливих значень випадкової величини X навмання (незалежно від спостерігача) вибирається одне єдине (спостережене) значення $x_{cn\,i}$.

Наприклад, якщо випробування полягає в підкиданні шестигранного кубика, а випадкова величина X – кількість очок, що випадає на верхній грані кубика, то множина можливих значень випадкової величини X містить шість елементів: $\Omega_X = \{1, 2, 3, 4, 5, 6\}$, тобто щоразу $x_{cn\,i} \in \Omega_X = \{1, 2, 3, 4, 5, 6\}$.

Скінченну множину спостережених значень $x_{cn1}, x_{cn2}, ..., x_{cn\,n}$ випадкової величини X називають *вибіркою об’єму* n.

В процесі розв’язування цілого ряду задач, виходячи з характеристик вибірки, роблять висновки про відповідні характеристики розподілу ймовірностей на множині значень досліджуваної випадкової величини. Наприклад, якщо 1000 разів підкидається гральний кубик, то в результаті кожного підкидання навмання вибирається один єдиний елемент із шестиелементної множини $\{1,2,3,4,5,6\}$. Не виключено, що в цій серії випробувань цифра 6 випадала 600 разів, цифра 5 – 200 разів, 4 – 100 разів, 3 – 50 разів, 2 – 30 разів, 1 – 20 разів. Тоді з великою мірою впевненості можна говорити, що кубик неоднорідний і в нього зміщений центр мас. Якщо перевіряється велика партія виробів і серед навмання відібраних з цієї партії m виробів (за досить великого m) виявиться певний відсоток бракованих виробів, то з великою мірою впевненості (яка залежить від m) можна говорити, що й у всій партії виробів такий самий відсоток бракованих. Проте постає питання: яким повинно бути m? Якщо перевірка виробу пов'язана з його демонтуванням, то це питання досить важливе для практики. Використовуючи методи математичної статистики, можна відповісти на такі питання.

Нехай проведено n випробувань, в результаті яких дістали вибірку $\{x_{cn1}, x_{cn2}, ..., x_{cnn}\}$. Елементи вибірки $\{x_{cn1}, x_{cn2}, ..., x_{cnn}\}$ називають *варіантами*, а впорядкований за зростанням набір варіант – *варіаційним рядом*. Якщо серед n спостережень значення x_1 спостерігалось n_1 разів, значення $x_2 - n_2$ разів, і т.д., значення $x_k - n_k$ разів, причому $n_1 + n_2 + ... + n_k = n$, то число $P_n^*(\{x_i\}) = \dfrac{n_i}{n}$ є статистичною ймовірністю події $\{x_i\}$ (тут припускається, що $\{x_i\}$ – події, див. §1.2, §1.6, §1.8).

Відповідність між окремими можливими значеннями випадкової величини X та їх статистичними ймовірностями, подану у вигляді Табл. 6.1.1, називають *рядом поточкового розподілу* статистичних ймовірностей за даної вибірки.

Табл. 6.1.1

Можливі значення випадкової величини	x_1	x_2	$\cdots$	
Статистичні ймовірності $P_n^*(\{x_i\})$	$P_n^*(\{x_1\})$	$P_n^*(\{x_2\})$	$\cdots$	$P_n^*(\{x_k\})$

Якщо на координатній площині побудувати точки $(x_i, P_n^*(\{x_i\}))$, $i \in \{1, 2, ..., k\}$, і сполучити їх відрізками прямих, то дістанемо *многокутник розподілу статистичних ймовірностей* або *полігон відносних частот* (Рис. 6.1.1).

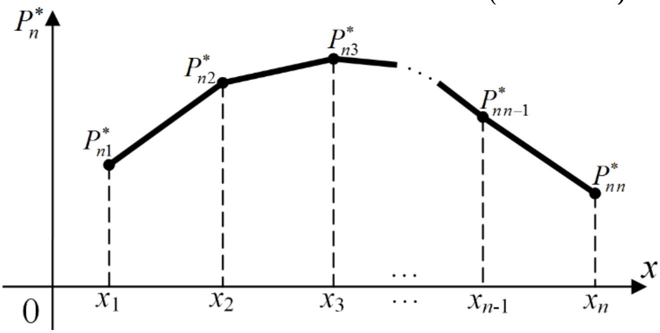


Рис. 6.1.1

Приклад 6.1.1. Побудувати ряд розподілу статистичних ймовірностей і многокутник розподілу статистичних ймовірностей (полігон відносних частот) стосовно вибірки із зафіксованих відхилень (вздовж лінії кидання) точки падіння спортивного снаряда від цілі:

$-20,\ 20, -10,\ 20,\ 10,\ 20, -50, -20, -10,\ 40, -20, -30, -10,\ 10,\ 20, -40,\ 50,$
$-10,\ 10,\ 50.$

Відповідний варіаційний ряд мнабуває вигляду
$-50, -50, -40, -30, -20, -20, -10, -10, -10,\ 10,\ 10,\ 10,\ 20,\ 20,\ 20, 20,\ 30,\ 40,$
$50, 50.$

Ряд дискретного поточкового розподілу статистичних ймовірностей подано в Табл. 6.1.2.

Табл. 6.1.2

x_i	-50	-40	-30	-20	-10	10	20	30	40	50
$P_n^*(\{x_i\})$	0.1	0.05	0.05	0.1	0.15	0.15	0.2	0.05	0.05	0.1

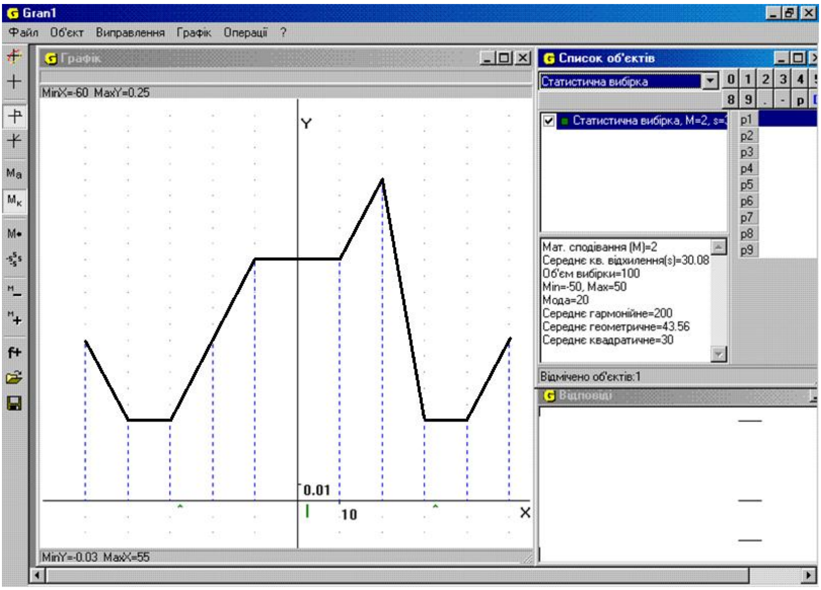


Рис. 6.1.2

Розподіл статистичних ймовірностей у розглянутому прикладі дещо нагадує нормальний розподіл ймовірностей (Рис. 6.1.2), проте оскільки вибірка не дуже велика, ця схожість не сильна. Насправді випадкові відхилення вздовж лінії кидання точки падіння спортивного снаряда від цілі розподіляються за нормальним розподілом ймовірностей.

Якщо спостерігається випадкова величина X з неперервною множиною значень, як це було в попередньому прикладі, або випадкова величина із скінченною множиною значень, але досить щільно розміщених на деякому проміжку $[a\,;b)$, і число спостережень досить велике, то й довжина вибірки може виявитися досить великою. Тоді подання статистичного матеріалу у вигляді Табл. 6.1.1 стає досить громіздким і практично недоцільним і непридатним. В такому разі проміжок $[a\,;b)$, в якому знаходяться всі спостережені значення випадкової величини X, ділять на практично прийнятну кількість часткових проміжків $[x_{i-1}\,;x_i)$, $i=\overline{1,k}$, однакової довжини h, $x_0=a$, $x_k=b$. Якщо a і b відповідно нижня й верхня межі проміжка, в якому знаходяться всі можливі значення досліджуваної випадкової величини, а k – кількість проміжків, то

$$h=\frac{b-a}{k}.$$

На практиці часто покладають $a=\min\limits_{1\le i\le n} x_{cn\,i}$, $b=\max\limits_{1\le i\le n} x_{cn\,i}+\varepsilon$, де $\varepsilon>0$ – деяке досить мале додатне число.

Стосовно кожного часткового проміжка $[x_{i-1};x_i)$ підраховують n_i – кількість спостережених значень, які лежать у цьому проміжку, і статистичну ймовірність $P_n^*([x_{i-1}\,;x_i))=\dfrac{n_i}{n}$.

Здобуті дані подають у вигляді Табл. 6.1.3, в якій вказують часткові проміжки і відповідні статистичні ймовірності.

Табл. 6.1.3

Частковий інтервал $I_i=([x_{i-1}\,;x_i)$	$[x_0+h;\,x_0+h)$	$[x_0+h;\,x_0+2h)$	…	$[x_0+(k-1)h;\,x_0+kh)$
Статистична ймовірність $P_n^*([x_{i-1}\,;x_i))$	$P_n^*([x_0\,;x_0+h))$	$P_n^*([x_0+h;\,x_0+2h))$	…	$P_n^*([x_0+(k-1)h;\,x_0+kh))$

Через таку таблицю задають *поінтервальний розподіл статистичних ймовірностей* (див. §2.2-§2.6).

Графік кусково сталої функції, який утворений з верхніх сторін прямокутників, основами яких є часткові інтервали $[x_{i-1}\,;x_i)$, $i=\overline{1,k}$, а висота i-го прямокутника добирається так, щоб його площа дорівнювала $P_n^*([x_{i-1}\,;x_i))$ (Рис. 6.1.3), називають *гістограмою*.

Для побудови гістограми вздовж осі абсцис відкладають часткові проміжки $[x_{i-1}\,;x_i)=[x_0+(i-1)h\,;x_0+ih)$, і на i-му проміжку будують прямокутник з висотою $\dfrac{P_n^*([x_{i-1}\,;x_i))}{h}$. Сума площ усіх таких прямокутників дорівнює 1.

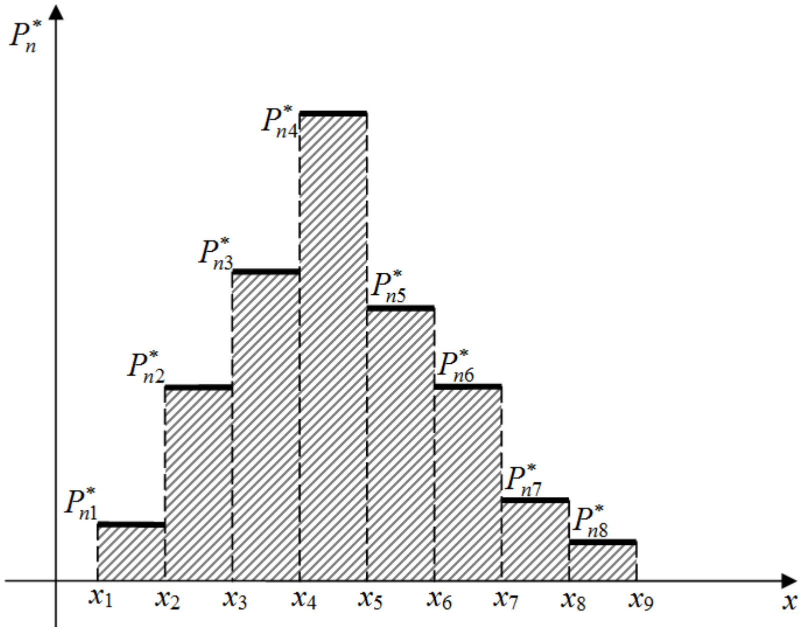


Рис. 6.1.3

Функцію

$$f_n^*(x) = \begin{cases} \dfrac{P_n^*([x_{i-1}\,;\,x_i))}{h}, & \text{коли} \quad x \in [x_{i-1}\,;\,x_i), \ i \in \overline{1,k}, \\ 0, & \text{коли} \quad x \,\overline{\in}\, [x_0\,;\,x_k), \end{cases}$$

називають щільністю поінтервального розподілу статистичних ймовірностей на множині $[x_0\,;x_k]$ за інтервалами $[x_{i-1}\,;x_i)$, $i \in \overline{1,\ k}$ (див. §2.2-§2.6). Графіком функції $f_n^*(x)$ є гістограма (Рис. 6.1.3).

На практиці число проміжків доцільно брати порядку 10–20. Іноді для визначення k використовують формулу Стерджеса $k = [1 + 3.322\ln(n)]$, де n – об'єм вибірки.

Приклад 6.1.2. Для визначення похибки вимірювального приладу зроблено 40 вимірювань, в яких зафіксовано похибки, подані в Табл. 6.1.4.

За цими даними побудувати поінтервальний розподіл статистичних ймовірностей та гістограму, якщо $k = 8$.

Поінтервальний розподіл статистичних ймовірностей стосовно розглядуваного прикладу подано в Табл. 6.1.5. Відповідну гістограму зображено на Рис. 6.1.4.

Табл. 6.1.4

i	1	2	3	4	5	6	7	8	9	10
x_i	-2.5	3	4	2	0.5	-1	2	4	-4	0
i	11	12	13	14	15	16	17	18	19	20
x_i	-0.5	-0.5	1	0.5	2.5	-0.5	2	1	-4	-2
i	21	22	23	24	25	26	27	28	29	30
x_i	-1	1.5	0.5	4	-1.5	-1	0	1	0	1
i	31	32	33	34	35	36	37	38	39	40
x_i	-1.5	1.5	0.5	0.5	-0.5	-1.5	-0.5	-1	2	0.5

Табл. 6.1.5

$[x_{i-1};x_i)$	$[-4;-3)$	$[-3;-2)$	$[-2;-1)$	$[-1;0)$	$[0;1)$	$[1;2)$	$[2;3)$	$[3;4{,}001)$
$P_n^*([x_{i-1};x_i))$	$\frac{2}{40}$	$\frac{1}{40}$	$\frac{4}{40}$	$\frac{9}{40}$	$\frac{9}{40}$	$\frac{6}{40}$	$\frac{5}{40}$	$\frac{4}{40}$

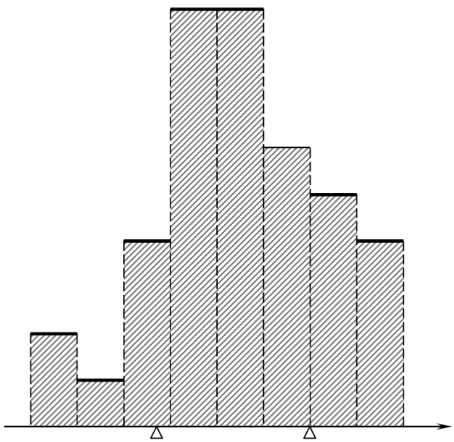


Рис. 6.1.4

Під час аналізу статистичних даних досить важливими є числові характеристики вибірки, за якими певною мірою можна охарактеризувати, де на числовій осі Ox знаходиться область, що охоплює множину можливих значень досліджуваної випадкової величини. До таких характеристик належать середнє арифметичне спостережених значень, середнє гармонійне, середнє геометричне, середнє квадратичне, мода, медіана, а також показники розсіювання (варіації) спостережених значень випадкової величини – розмах вибірки, середнє абсолютне відхилення, коефіцієнт варіації тощо.

Усі середні значення не менші за мінімальне і не більші за максимальне спостережені значення.

Середнє арифметичне спостережених значень обчислюється за формулою

$$\bar{x}_{арифм} = \frac{1}{n}\sum_{i=1}^{n} x_{cn\,i} = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \sum_{i=1}^{k} x_i \frac{n_i}{n} = \sum_{i=1}^{n} x_i P_n^*(\{x_i\}),$$

де n_i – кількість спостережень, в яких досліджувана випадкова величина X набувала одного із своїх можливих значень x_i.

Середнє гармонійне обчислюється за формулою

$$\bar{x}_{гарм} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{x_i}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_{cni}}},$$

де $x_{cn\,i}$ – спостережені значення випадкової величини, а n_i – кількість однакових значень $x_{cn\,i}$.

Так, якщо n_i, $i = \overline{1,\,k}$ – довжина деякого проміжка, а x_i – швидкість, з якою на цьому проміжку рухається деяке тіло, то середня швидкість, з якою це тіло пройде всі k проміжків, є середнім гармонійним значень x_i.

Середнє геометричне

$$\bar{x}_{геом} = \left(\prod_{i-1}^{n} x_{cni}\right)^{\frac{1}{n}}$$

використовується для усереднення набору дробів, темпів росту та ін. (Логарифм середнього геометричного дорівнює середньому арифметичному логарифмів спостережених значень $x_{cn\,i}$).

Середнє квадратичне

$$\bar{x}_{\sigma} = \frac{\sum_{i=1}^{k} x_i^2 n_i}{\sum_{i=1}^{k} n_i} = \sum_{i=1}^{k} x_i^2 P_n^*(\{x_i\})$$

використовується стосовно розрахунку показників розсіювання статистичних ймовірностей (дисперсії і середнього квадратичного відхилення). Нагадаємо в зв'язку з цим, що $D[X] = M[X^2] - M^2[X]$.

Слід зазначити, що коли стосовно одного й того самого варіаційного ряду визначені всі вказані середні, то

$$\bar{x}_{гарм} \le \bar{x}_{геом} \le \bar{x}_{арифм} \le \bar{x}_{\sigma}.$$

Крім вище згаданих середніх значень в статистичному аналізі іноді використовують також такі числові характеристики розподілу статистичних ймовірностей, як мода, медіана.

Мода – це варіанта, якій відповідає найбільша статистична ймовірність в ряді розподілу статистичних ймовірностей. Іншими словами, мода – це значення, яке найчастіше зустрічається у вибірці. Таких значень може бути кілька.

Медіана – це значення, частота попадання лівіше від якого дорівнює частоті попадання правіше від нього. Слід зазначити, що медіана має зміст не за будь-яких розподілів частот (наприклад, якщо статистичні імовірності на множині $\{x_1, x_2\}$ розподілені у співвідношенні 0.1 і 0.9, тоді медіана втрачає зміст).

Найпростішою характеристикою розсіювання спостережених значень є *розмах вибірки* – різниця між максимальним і мінімальним значеннями сукупності спостережених значень:

$$R = \max_{1 \le i \le n} x_{cn\,i} - \min_{1 \le i \le n} x_{cn\,i} .$$

Проте розмах вибірки не можна вважати задовільною оцінкою розсіювання, оскільки за його визначення враховуються тільки два крайні значення варіаційного ряду, а статистичні ймовірності проміжних варіант і особливості розподілу статистичних ймовірностей ніяк не враховуються.

Приклад 6.1.3. 20 навмання вибраних учнів виконують стрибки у висоту і зафіксовано такі результати (в сантиметрах) (Табл. 6.1.6).

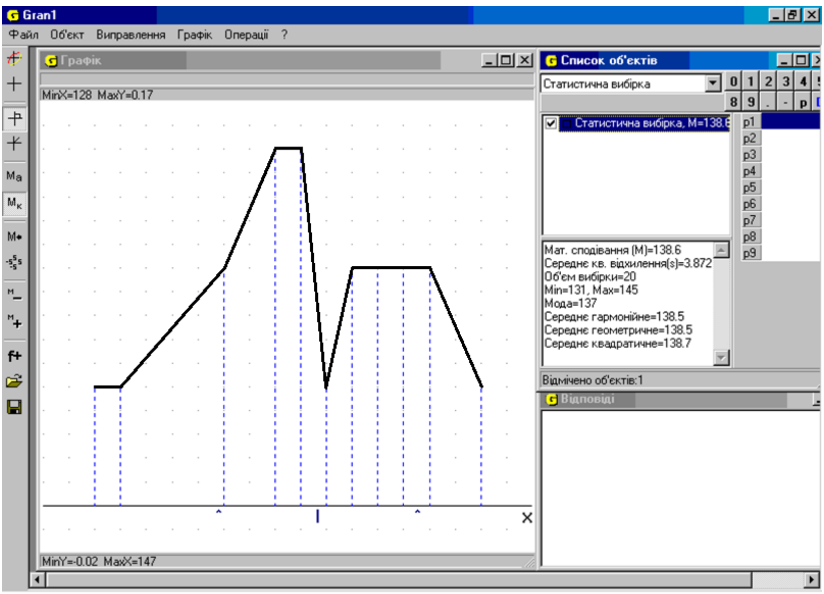


Рис. 6.1.5,а

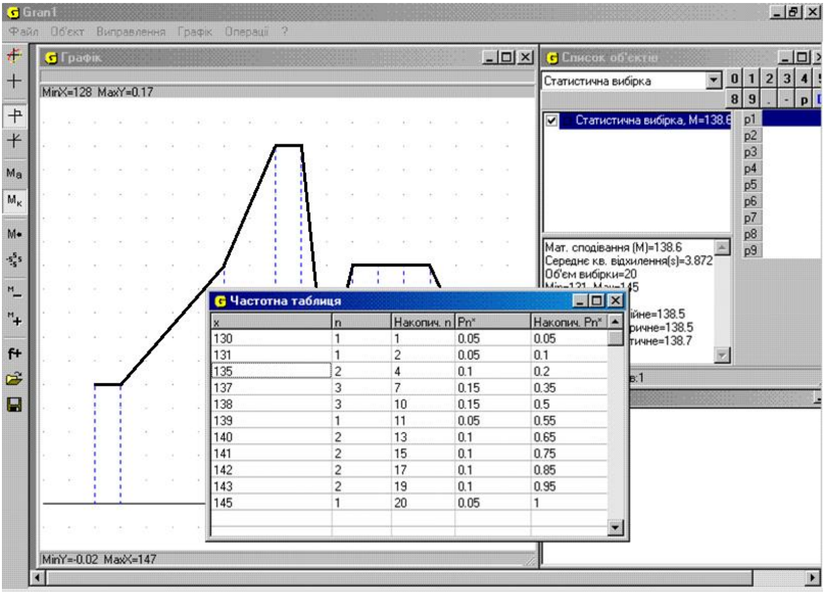


Рис. 6.1.5, б

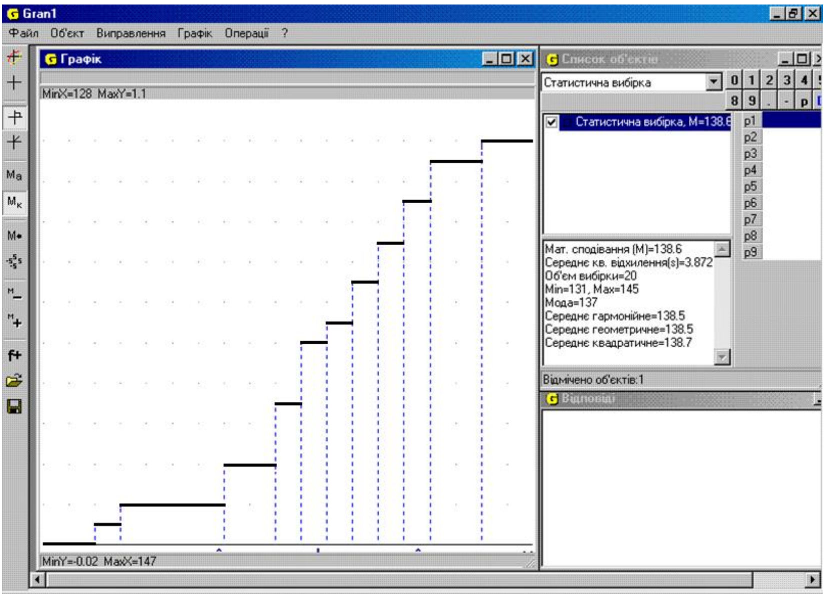


Рис. 6.1.6

Табл. 6.1.6

i	1	2	3	4	5	6	7	8	9	10
x_i	137	140	143	135	142	139	141	137	142	131
i	11	12	13	14	15	16	17	18	19	20
x_i	145	138	141	143	130	138	140	135	137	138

Ряд розподілу статистичних ймовірностей в даному разі набуває вигляду (Табл. 6.1.7):

Табл. 6.1.7

x_i	130	131	135	137	138	139	140	141	142	143	145
$P_n^*(\{x_i\})$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{1}{20}$

Многокутник розподілу статистичних ймовірностей (полігон відносних частот) подано на Рис. 6.1.5, а. Зазначимо, що за допомогою програмного засобу GRAN1 статистичні ймовірності (відносні частоти) окремих спостережених значень можна визначити за частотною таблицею (Рис. 6.1.5, б) або за многокутником розподілу статистичних ймовірностей (полігоном відносних частот), підвівши курсор у потрібну точку на графіку (Рис. 6.1.5, а).

На Рис. 6.1.6 подано графік функції розподілу статистичних ймовірностей та окремі числові характеристики розподілу статистичних ймовірностей стосовно вибірки, розглянутої в даному прикладі.

Вправи для самостійного виконання

6.1.1. Перевірити, чи правильні твердження:

1. Елементи вибірки (спостережені значення відповідної випадкової величини) можуть бути однаковими.
2. Об’єм вибірки дорівнює кількості різних спостережених значень.
3. Варіанти – це елементи вибірки.
4. Варіаційний ряд – це будь-яка сукупність спостережених значень.
5. Стосовно кожної вибірки можна скласти ряд дискретного поточкового розподілу статистичних ймовірностей.
6. Полігон відносних частот існує за будь-якого розподілу відносних частот.
7. Стосовно кожної вибірки можна скласти поінтервальний розподіл статистичних ймовірностей.
8. Кожний розподіл статистичних ймовірностей графічно можна подати за допомогою гістограми.
9. Гістограма – це графік щільності розподілу статистичних ймовірностей.
10. Стосовно кожної вибірки існують усі середні значення: арифметичне, гармонійне, геометричне і квадратичне.
11. Завжди $\bar{x}_{гарм} \le \bar{x}_{геом} \le \bar{x}_{арифм} \le \bar{x}_{\sigma}$.
12. Стосовно кожної вибірки існує: а) мода; б) медіана.

6.2. Статистичні наближення невідомих функції та щільності розподілу ймовірностей

Надалі вважатимемо, що $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$.

Нехай проведено n спостережень, в яких дістали варіанти $x_1, x_2, ..., x_n$.

Функцією розподілу узагальнених статистичних імовірностей на множині значень випадкової величини X називають узагальнену статистичну ймовірність попадання в множину $(-\infty; x)$ спостережених значень випадкової величини X , тобто.

$$F^*_{nX}(x) = P^*_{nX}((-\infty; x)) . \tag{6.2.1}$$

Згідно із законом великих чисел за досить великої кількості спостережень без великого ризику мати значну похибку статистичну ймовірність $P^*_{nX}((-\infty; x))$ можна вважати наближеним значенням ймовірності $P_X((-\infty; x))$, за яке приймається узагальнена статистична ймовірність попадання в множину $(-\infty; x)$ (див. §2.2-2.7). Отже, за досить великої кількості спостережень і досить малих часткових інтервалів $[x_{i-1}; x_i)$, $i \in \overline{1, k}$, таких, що $F^*_{nX}(x_0) = 0$, $F^*_{nX}(x_k) = 1$, $-\infty < x_0 < x_k < \infty$, функція розподілу узагальнених статистичних ймовірностей $F^*_{nX}(x)$ є досить добрим наближенням невідомої функції розподілу ймовірностей $F_X(x)$. Це саме стосується й гістограми, яка за досить великої кількості спостережень і досить малих часткових інтервалів $[x_{i-1}; x_i) = = (x_0 + (i-1)h; x_0 + ih))$, $i \in \overline{1, k}$ (див. §2.7), є досить добрим наближенням графіка функції $y = f_X(x)$, де $f_X(x)$ – невідома щільність розподілу ймовірностей на множині значень випадкової величини X .

Приклад 6.2.1. Перевірено 10 навмання вибраних класів. Кількість відмінників у кожному класі подано в Табл. 6.2.1.

Табл. 6.2.1

i	1	2	3	4	5	6	7	8	9	10
x_i	5	8	3	4	5	1	6	4	2	3

Побудувати функцію розподілу статистичних імовірностей на множині, елементами якої є значення кількості відмінників, та її графік.

Множина значень досліджуваної випадкової величини дискретна (кількість відмінників у навмання вибраному класі) і ця величина набуває лише скінченну кількість цілочисельних значень.

Згідно з означенням $F_{nX}^{*}(x)$ за наведеними даними спостережень одержуємо

$$F_{10X}^{*}(x)=\begin{cases}0, & \text{коли } x\le 1;\\ 0.1, & \text{коли } 1<x\le 2;\\ 0.2, & \text{коли } 2<x\le 3;\\ 0.4, & \text{коли } 3<x\le 4;\\ 0.6, & \text{коли } 4<x\le 5;\\ 0.8, & \text{коли } 5<x\le 6;\\ 0.9, & \text{коли } 6<x\le 8;\\ 1, & \text{коли } 8<x.\end{cases}$$

Тут через $F_{10X}^{*}(x)$ позначено функцію поточкового розподілу статистичних ймовірностей на $(R^1,\mathscr{B}(R^1))$ (відносну частоту попадання значень випадкової величини X на проміжкок $(-\infty;x)$) за будь-якого $x\in(-\infty;\infty)$ (див. §2.5).

Так, за $x=5.5$ серед усіх десяти проведених спостережень є 8 спостережень, в яких зафіксовано значення випадкової величини X менші, ніж 5.5, тобто які попадають на проміжок $(-\infty;5.5)$. Тому $F_{10X}^{*}(5.5)=P_{10}^{*}((-\infty;5.5))=\dfrac{8}{10}$.

Той самий результат дістанемо, коли взяти будь-яке x з проміжка $(5;6]$.

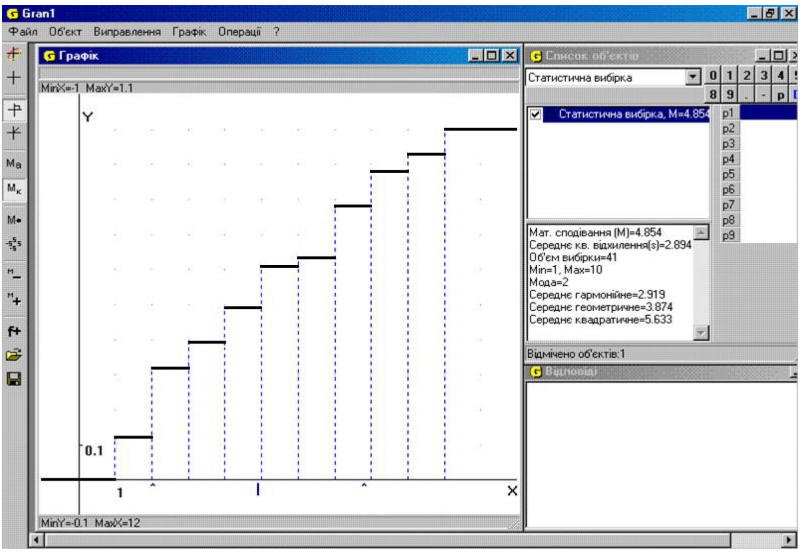


Рис. 6.2.1

Графік функції $F^*_{10X}(x)$ подано на Рис. 6.2.1.

Якщо множина Ω_X значень випадкової величини X неперервна, як $\tilde{\Omega}_X$ розглядається R^1, а як $\tilde{S}_X$ – $\mathscr{B}(R^1)$, тоді функцію розподілу статистичних ймовірностей будують за даними поінтервального розподілу, вважаючи розподіл статистичних ймовірностей в кожному окремому інтервалі $[x_{i-1}; x_i)$ рівномірним (див. §2.2-2.7). Значення функції розподілу узагальнених статистичних (гіпотетичних) імовірностей обчислюють в кінцях часткових інтервалів $[x_{i-1}; x_i) = (x_0 + (i-1)h; x_0 + ih))$, $i = 1, 2, ..., k$, покладаючи

$$F^*_{nX}(x_0) = 0,\ F^*_{nX}(x_0 + h) = P^*_{nX}([x_0; x_0 + h))$$

$$F^*_{nX}(x_0 + 2h) = P^*_{nX}([x_0; x_0 + h)) + P^*_{nX}([x_0 + h; x_0 + 2h)),\ \dots,$$

$$F^*_{nX}(x_0 + kh) = P^*_{nX}([x_0; x_0 + h)) + P^*_{nX}([x_0 + h; x_0 + 2h)) + ... +$$

$$+ P^*_{nX}([x_0 + (k-1)h; x_0 + kh))..$$

Побудувавши на площині точки

$$(x_0, 0),\ (x_0 + h, P^*_{nX}([x_0, x_0 + h))),$$

$$(x_0 + 2h, P^*_{nX}([x_0, x_0 + h)) + P^*_{nX}([x_0 + h, x_0 + 2h))),\ \dots,$$

$$(x_0 + kh, P^*_{nX}([x_0, x_0 + h)) + P^*_{nX}([x_0 + h, x_0 + 2h)) + ... +$$

$$+ P^*_{nX}([x_0 + (k-1)h, x_0 + kh)))$$

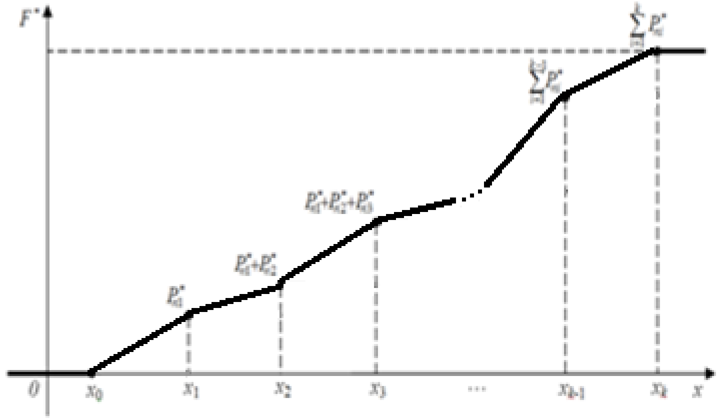


Рис. 6.2.2

і з’єднавши їх відрізками прямих, дістають графік функції $F^*_{nX}(x)$ розподілу узагальнених статистичних ймовірностей на множині $\tilde{\Omega}_X = R^1$ (Рис. 6.2.2), вважаючи, що $\tilde{S}_X = \mathscr{B}(R^1)$, а узагальнена статистична ймовірність $P^*_{nX}((-\infty; x)) = \int\limits_{-\infty}^{x} f^*_{nX}(x)dx$, де $f^*_{nX}(x)$ – щільність поінтервального розподілу статистичних ймовірностей на множині $[x_0; x_k)$ значень випадкової величини X за інтервалами $[x_{i-1}; x_i)$, $i \in \overline{1, k}$ (див. §2.2-§2.7), $\tilde{\Omega}_X = R^1$ $\tilde{S}_X = \mathscr{B}(R^1)$.

Якщо множина значень випадкової величини X дискретна, але ці значення розміщені досить щільно в деякому інтервалі $[a; b)$ і кількість можливих значень випадкової величини X дуже велика, то функцію розподілу ймовірностей такої випадкової величини також будують за відповідним поінтервальним розподілом, замінюючи в такий спосіб дискретний розподіл статистичних ймовірностей на множині значень випадкової величини X поінтервальним, також вважаючи рівномірним розподіл в кожному частковому інтервалі $[x_{i-1}; x_i)$, $\bigcup\limits_{i=1}^{k} [x_{i-1}; x_i) = [a; b)$ (див. §2.5, §6.1).

Зауважимо, що вимога щільності множини можливих значень випадкової величини X не є обов'язковою. Зауважимо також, що не всі можливі значення випадкової величини X повинні бути спостереженими. Якщо множина значень випадкової величини X неперервна, то це навіть неможливо за будь-якої кількості випробувань (спостережень).

Практично неможливо це і стосовно випадкової величини із скінченною, але дуже великою кількістю k різних можливих значень, наприклад, коли $k = 10^{10^{10}}$ і т.п. (див. §2.2-2.7 і ін.).

Зауважимо, що коли $\Omega_X = [x_0; x_k)$, $S_X = \{A \mid A = \bigcup\limits_{i \in I} [x_{i-1}; x_i), I \subset \{1, 2, \dots, k\}\}$, і інтервали $[x_{i-1}; x_k)$ поступово подрібнювати, спрямовуючи $h = x_i - x_{i-1}$ до нуля, залишаючи незмінною щільність $f^*_{nX}(x)$ поінтервального розподілу статистичних імовірностей за початковими інтервалами $[x_{i-1}; x_k)$, тоді графік кусково сталої функції $F^*_{nX}(x)$ поінтервального розподілу узагальнених статистичних ймовірностей поступово наближається до неперервного графіка функції $F^*_{nX}(x)$ неперервного розподілу узагальнених статистичних ймовірностей на неперервній множині точок (див. §2.7, Рис. 2.7.5а-Рис. 2.7.5д). На практиці часто покладають $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$, за заданим поінтервальним розподілом статистичних імовірностей знаходять значення функції $F^*_{nX}(x)$ в кінцях інтервалів, вважаючи що на інтервалі $[x_{i-1}; x_k)$ функція $F^*_{nX}(x)$ змінюється лінійно від значення $F^*_{nX}(x_{i-1})$ до значення $F^*_{nX}(x_i)$. На графіку функції $F^*_{nX}(x)$ в такому разі точки $(x_{i-1}, F^*_{nX}(x_{i-1}))$ і $(x_i, F^*_{nX}(x_i))$ з'єднують відрізками прямих

Приклад 6.2.2. За даними спостережень дістали такий поінтервальний розподіл статистичних ймовірностей на множині $\Omega_X = [0; 40)$ значень деякої випадкової величини X (Табл. 6.2.2):

Табл. 6.2.2

I_i	[0–5)	[5–10)	[10–15)	[15–20)	[20–25)	[25–30)	[30–35)	[35–40)
$P^*_n(I_i)$	0.01	0.04	0.12	0.28	0.30	0.16	0.08	0.01

Потрібно знайти значення функції розподілу узагальнених статистичних ймовірностей $F^*_{nX}(x)$ у кінцях інтервалів I_i і побудувати графік функції $y = F^*_{nX}(x)$ (див. §2.2-2.7).

Враховуючи означення функції $F^*_{nX}(x)$, дістаємо

$$F^*_{nX}(0) = 0.00; \; F^*_{nX}(5) = 0.01; \; F^*_{nX}(10) = 0.05;$$
$$F^*_{nX}(15) = 0.17; \; F^*_{nX}(20) = 0.45; \; F^*_{nX}(25) = 0.75;$$
$$F^*_{nX}(30) = 0.91; \; F^*_{nX}(35) = 0.99; \; F^*_{nX}(40) = 1.00.$$

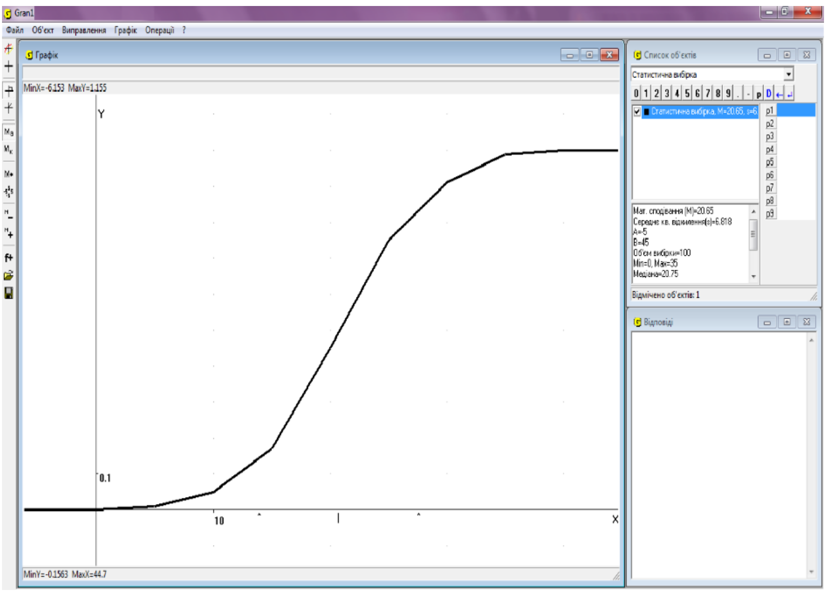


Рис. 6.2.3

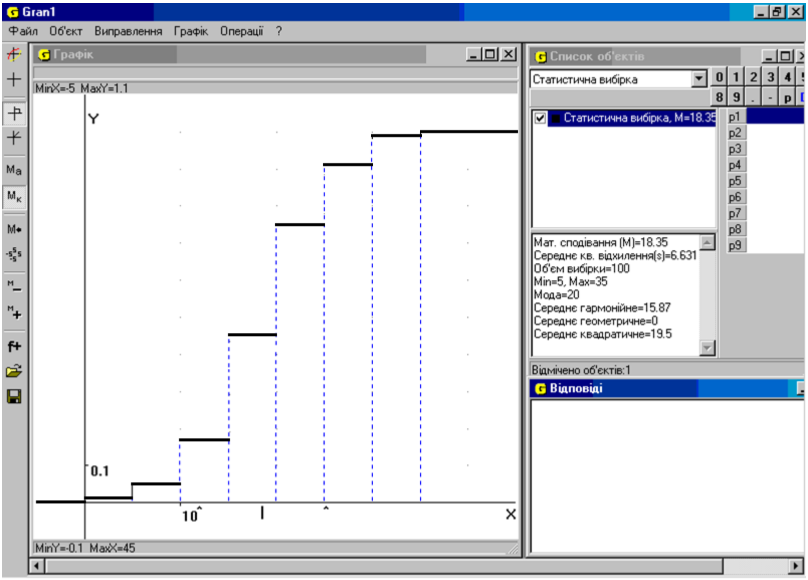


Рис. 6.2.4

Побудувавши точки $(x_i, F^*_{nX}(x_i))$ на площині xOy і сполучивши їх відрізками прямих, дістанемо графік функції $y = F^*_{nX}(x)$ (Рис. 6.2.3).

На основі властивостей функції $F^*_{nX}(x)$ вважатимемо, що $F^*_{nX}(x) = 0$, коли $x < 0$ і $F^*_{nX}(x) = 1$, коли $x > 40$. Тоді знайдемо $F^*_{nX}(x)$ за всіх $x \in (-\infty; \infty)$.

Якщо значеннями, яких набуває досліджувана випадкова величина всередині інтервалів $I_i = [x_0 + (i-1)h\,; x_0 + ih))$, нехтують або округлюють їх до наближених значень $x_0 + (i-1)h$, або якщо досліджувана випадкова величина X може набувати тільки значень $x_0 + (i-1)h$, тобто множина значень випадкової величини X дискретна, $S = \{A \mid A = \bigcup\limits_{i \in I} I_i, I \subset \overline{1{,}8}\}$, а значення $F^*_n(x)$ обчислюється за формулою $F^*_n(x) = P^*_n(\bigcup\limits_{A \in S,\, A \subset (-\infty\,;\, x)} A)$, тоді функція $F^*_{nX}(x)$ набуватиме сталих значень на проміжках I_i (Рис. 6.2.4) (див. §2.2-2.7).

Якщо h досить мале, коли функція $F_X(x)$ мало змінюється на проміжку I_i, графіки неперервної та кусково-сталої функції

$F_X(x)$ практично не відрізняються між собою (див. Рис. 2.7.5д). Разом з тим не слід змішувати дискретні і неперервні розподіли як ймовірностей, так і статистичних ймовірностей. Різниця між такими розподілами суттєва (див. §2.2-§2.7).

Наприклад, похідна від кусково-сталої функції $F_X(x)$ майже скрізь дорівнює нулеві, хоч така функція на інтервалах досить значної довжини набуває практично таких самих приростів, як і відповідна кусково-лінійна функція $F_X(x)$, похідна якої на інтервалах $[a_{i-1}\,;a_i)$ відмінна від нуля (див. §2.2-§2.7).

Вправи для самостійного виконання

6.2.1. Перевірити, чи правильні твердження:

1. Стосовно кожної вибірки можна побудувати функцію $F_X(x)$ розподілу узагальнених статистичних ймовірностей на множині значень відповідної випадкової величини X.

2. Завжди $F^*_{nX}(x) = \dfrac{1}{n}\,k_n((-\infty;x))$, де $k_n((-\infty;x))$ – кількість спостережених значень на проміжку $(-\infty;x)$, тобто кількість $x_{\text{сп}i}$, менших, ніж x.

3. За поточкового розподілу статистичних ймовірностей на скінченній множині точок відповідна функція розподілу кусково стала.

4. За поінтервального розподілу статистичних ймовірностей відповідна функція розподілу є неперервною.

5. Графік функції розподілу узагальнених статистичних ймовірностей завжди розривний.

6. Властивості функції розподілу узагальнених статистичних ймовірностей залежать від структури подій із простору подій S.

6.3. Статистичні оцінки параметрів розподілу ймовірностей

Надалі будемо вважати $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$.

Досить часто трапляються випадки, коли характер розподілу ймовірностей відомий, проте невідомі параметри цього розподілу. Наприклад, досліднику може бути відомо, що ймовірності розподілені за законом Пуассона $P_n(\mu_n = m) = \frac{a^m}{m!} e^{-a}$, а параметр a невідомий і підлягає наближеному визначенню на основі статистичних даних. Може бути відомо, що розподіл імовірностей на множині значень досліджуваної випадкової величини нормальний, проте його параметри a і σ невідомі і їх треба принаймні наближено визначити на основі статистичних даних і т.д.

Нехай проведено n спостережень за випадковою величиною X, в яких дістали значення $x_{cn1}, x_{cn2}, ..., x_{cn\,n}$ (не обов'язково різні), і треба оцінити параметр θ розподілу ймовірностей на множині значень випадкової величини X, причому характер цього розподілу відомий. Нехай θ_n^* є значення статистичної оцінки θ^* шуканого параметра, знайдене за результатами спостережень. Зрозуміло, що оцінка θ^* залежить від результатів спостережень, тобто є функцією цих результатів. Оскільки значення $x_{cn1}, x_{cn2}, ..., x_{cn\,n}$ наперед передбачити неможливо, то й значення θ_n^* оцінки θ^* наперед передбачити неможливо, θ^* є випадковою величиною, яка незалежно від спостерігача може набувати різних значень в різних серіях спостережень. Таким чином, θ_n^* є спостережене значення випадкової величини θ^*, знайдене за результатами n спостережинь.

Статистична оцінка θ^* параметра θ називається *незміщеною*, якщо $M[\theta^*] = \theta$, тобто якщо математичне сподівання випадкової величини θ^* дорівнює значенню параметра, яке оцінюється. Якщо $M[\theta^*] < \theta$, то оцінка для θ буде систематично занижуватись; якщо $M[\theta^*] > \theta$, то оцінка для θ буде систематично завищуватись. В обох випадках оцінка зміщена.

Оцінка θ^* називається *несуперечливою*, якщо за $n \to \infty$ θ_n^* збігається за ймовірністю до значення параметра θ, яке оцінюється.

Оцінка θ^* значення параметра Θ називається *ефективною*, якщо дисперсія $D[\theta^*]$ найменша порівняно з іншими оцінками цього параметра, знайденими на основі тих самих результатів спостережень.

Нехай спостерігається випадкова величина X, математичне сподівання якої m, а дисперсія D. Як статистичну оцінку математичного сподівання спостережуваної випадкової величини X за даними проведених n спостережень природно взяти середнє арифметичне спостережених значень

$$\bar{x}_n = \frac{x_{cn1} + x_{cn2} + \ldots + x_{cn\,n}}{n}, \qquad (6.3.1)$$

де x_{cni} – значення спостережуваної випадкової величини, одержане під час i-го спостереження.

Величину $\bar{x}_n$ часто називають *вибірковим середнім або статистичним середнім.*

Формулу для обчислення $\bar{x}_n$ можна подати у вигляді

$$\bar{x}_n = \sum_{i=1}^{k} x_i P_n^*(\{x_i\}), \qquad (6.3.2)$$

де x_i – різні можливі значення випадкової величини X, $P_n^*(\{x_i\})$ – відповідні їм статистичні ймовірності (відносні частоти), одержані за результатами n спостережень. Зауважимо, що не виключається, що якесь із можливих значень x_i не було спостережене і відповідна йому статистична ймовірність $P_n^*(\{x_i\})$ дорівнює нулеві.

Якщо через X_1 позначити випадкову величину X за першого спостереження, через X_2 – за другого спостереження і т.д., то $\bar{x}_n$ можна тлумачити як спостережене значення випадкової величини

$$\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n).$$

За властивостями числових характеристик розподілів ймовірностей на множинах значень випадкових величин дістаємо

$$M[\bar{X}_n] = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} M[X_i] = m,$$

оскільки у всіх випадкових величин X_i одне й те саме математичне сподівання

$$M[X_i] = M[X] = m, \quad i = 1, 2, \ldots, n.$$

Таким чином, оцінка $\overline{X}_n$ для математичного сподівання випадкової величини X незміщена.

За теоремою Чебишова *за необмеженого збільшення числа спостережень послідовність випадкових величин* $\overline{X}_n$ *збігається за ймовірністю до числа* $m = M[X]$, *тобто за досить великого* n *і як завгодно малих* $\varepsilon > 0$ *і* $\delta > 0$ *виконується нерівність*

$$P(|\overline{X}_n - m| < \varepsilon) > 1 - \delta ,$$

а це означає, що оцінка $\overline{X}_n$ стосовно математичного сподівання випадкової величини X *несуперечлива*.

Оскільки спостереження взаємно незалежні між собою, то за теоремами про числові характеристики розподілів ймовірностей на множинах значень функцій випадкових аргументів

$$D[\overline{X}_n] = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} D[X_i] = \frac{D}{n} .$$

Ефективність чи неефективність оцінки $\overline{X}_n$ залежить від того, як розподілені ймовірності на множині значень випадкової величини X. Виявляється, що коли розподіл ймовірностей нормальний, то оцінка $\overline{X}_n$ ефективна. За інших розподілів ймовірностей оцінка $\overline{X}_n$ може бути неефективною.

Розглянемо статистичну оцінку дисперсії розподілу ймовірностей на множині значень випадкової величини X. На перший погляд здається, що за таку оцінку слід взяти величину

$$\overline{D}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 .$$

Проте ця оцінка зміщена. Справді, враховуючи, що $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\overline{X}_n - m = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)$, одержуємо

$$\overline{D}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n}((X_i - m) - (\overline{X}_n - m))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}((X_i - m)^2 - 2(\overline{X}_n - m)(X_i - m) + (\overline{X}_n - m)^2) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 - 2(\overline{X}_n - m)\frac{1}{n}\sum_{i=1}^{n}(X_i - m) + (\overline{X}_n - m)^2 =$$

$$=\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2-2(\overline{X}_n-m)^2+(\overline{X}_n-m)^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2-$$

$$-(\overline{X}_n-m)^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2-\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-m)\right)^2=$$

$$=\frac{1}{n}\sum_{i=1}^{n}(X_i-m)^2-\frac{1}{n^2}\sum_{i=1}^{n}(X_i-m)^2-\frac{1}{n^2}2\sum_{i<j}(X_i-m)(X_j-m)=$$

$$=\frac{n-1}{n^2}\sum_{i=1}^{n}(X_i-m)^2-\frac{2}{n^2}\sum_{i<j}(X_i-m)(X_j-m)\ .$$

Звідси

$$M[\overline{D}_n]=\frac{n-1}{n}\sum_{i=1}^{n}M[(X_i-m)^2]-\frac{2}{n^2}\sum_{i<j}M[(X_i-m)(X_j-m)]=$$

$$=\frac{n-1}{n}\sum_{i=1}^{n}D[(X_i]-\frac{2}{n^2}\sum_{i<j}K[X_i,X_j]\ .$$

Однак $D[X_i]=D[X]=D$ за будь-якого $i=1,2,...,n$, а кореляційні моменти $K[X_i,X_j]$ дорівнюють нулю, оскільки спостереження взаємно незалежні, а отже й випадкові величини X_i, $i=1,2,...,n$, взаємно незалежні. Таким чином, $M[\overline{D}_n]=\frac{n-1}{n}D$, тобто оцінка $\overline{D}_n$ занижена і її використання спричинюватиме систематичну похибку.

Тому за оцінку дисперсії беруть не величину $\overline{D}_n$, яку називають *статистичною або вибірковою дисперсією*, а величину

$$\tilde{D}_n=\frac{n}{n-1}\overline{D}_n=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2\ .$$

Очевидно, що $M[\tilde{D}_n]=\frac{n}{n-1}M[\overline{D}_n]=D$, і таким чином оцінка $\tilde{D}_n$ для дисперсії D незміщена. Величину $\tilde{D}_n$ називають «*виправленою*» *статистичною дисперсією*. Очевидно, що за досить великих n оцінки $\overline{D}_n$ і $\tilde{D}_n$ практично не відрізняються між собою.

Формулу для обчислення спостереженого значення випадкової величини $\overline{D}_n$ можна подати у вигляді

$$\overline{D}_n = \sum_{i=1}^{k} (x_i - \overline{x}_n)^2 P_n^*(\{x_i\}), \qquad (6.3.3)$$

де x_i – можливі значення випадкової величини X, $P_n^*(\{x_i\})$ – відповідні їм статистичні ймовірності.

Оцінка $\overline{D}_n$ *несуперечлива*. Справді, $\overline{D}_n$ можна записати у вигляді

$$\overline{D}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2 .$$

Доданок $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ в цій формулі є середнім арифметичним n випадкових величин X_i^2, де X_i – випадкова величина X за i-го спостереженні. За теоремою Чебишова $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ збігається за імовірністю до $M[X^2]$; $\overline{X}_n^2$ збігається за ймовірністю до m^2, тому послідовність випадкових величин $\overline{D}_n$ збігається за ймовірністю до $M[X^2] - m^2 = D$, тобто $\overline{D}_n$ несуперечлива оцінка. Оскільки $\frac{n-1}{n}$ прямує до одиниці, коли $n \to \infty$, то й оцінка $\widetilde{D}_n$ також несуперечлива.

Аналогічно розглядають статистичні оцінки середнього квадратичного відхилення $\overline{\sigma}_n$, $\widetilde{\sigma}_n$, які називають *вибірковим середнім квадратичним відхиленням* і *«виправленим» середнім квадратичним відхиленням* відповідно та обчислюють як корінь квадратний із вибіркової дисперсії та виправленої статистичної дисперсії:

$$\overline{\sigma}_n = \sqrt{\overline{D}_n}\ ,\ \widetilde{\sigma}_n = \sqrt{\widetilde{D}_n}\ .$$

На основі статистичних даних досліджують також випадкові вектори; вивчають розподіли ймовірностей на множинах значень випадкових векторів та їх числові характеристики; зв’язки між координатами випадкових векторів і т.д.

Приклад 6.3.1. На 100 однакових за розмірами навмання вибраних ділянках землі з однаковою кількістю внесених добрив зібрано різні врожаї зерна. Результати проведених спостережень подано в Табл. 6.3.1. На основі поданого статистичного матеріалу треба визначити статистичні оцінки математичного сподівання і дисперсії досліджуваної випадкової величини (урожай).

Табл. 6.3.1

Урожай, ц/га	14	15	16	17	18	19	20
Кількість ділянок	6	10	18	28	20	12	6

За формулою (6.3.2) одержуємо

$$\bar{x}_{100} = 14 \cdot \frac{6}{100} + 15 \cdot \frac{10}{100} + 16 \cdot \frac{18}{100} + 17 \cdot \frac{28}{100} + 18 \cdot \frac{20}{100} +$$

$$+19 \cdot \frac{12}{100} + 20 \cdot \frac{6}{100} = \frac{84+150+228+476+360+228+120}{100} = 17.06\,.$$

Застосовуючи формулу (6.3.3), стосовно спостереженого значення випадкової величини $\bar{D}_n$ знаходимо

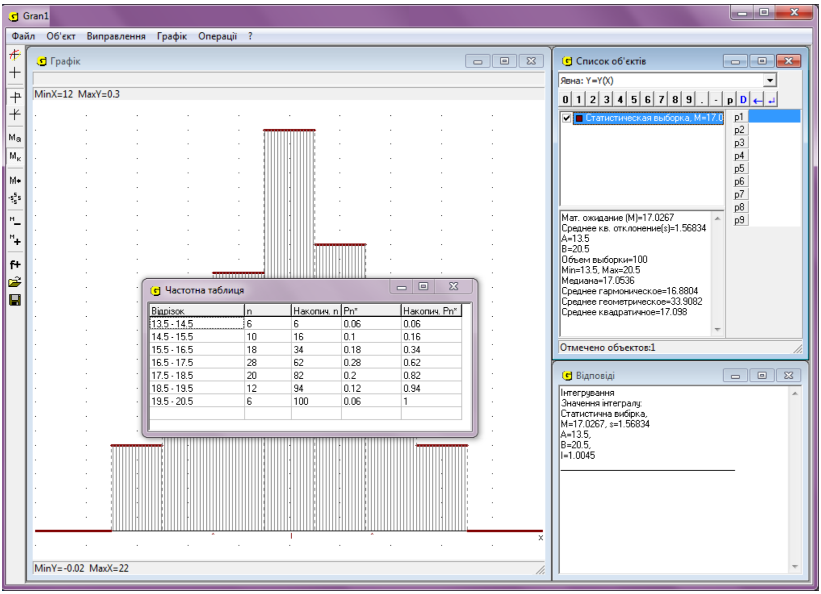


Рис. 6.3.1

$$\bar{D}_{100} = (-3.06)^2 \cdot \frac{6}{10} + (-2.06)^2 \cdot \frac{10}{100} + (-1.06)^2 \cdot \frac{18}{100} +$$

$$+(-0.06)^2 \cdot \frac{28}{100} + (0.94)^2 \cdot \frac{20}{100} + (1.94)^2 \cdot \frac{12}{100} + (2.94)^2 \cdot \frac{6}{100} =$$

$$= 9.36 \cdot 0.06 + 4.24 \cdot 0.10 + 1.12 \cdot 0.18 + 0 \cdot 0.28 + 088 \cdot 0.20 +$$

$$+3.76\cdot 0.12+8.64\cdot 0.06=2.33.$$

Далі дістаємо

$$\tilde{D}_{100}=\frac{n}{n-1}\overline{D}_{100}=\frac{100}{99}\cdot 2.33=1.01\cdot 2.33=2.35,\ \tilde{\sigma}_{100}=\sqrt{\tilde{D}_{100}}=1.536\ .$$

Для виконання обчислень розглянутого типу зручно скористатися послугами програми GRAN1 (Рис. 6.3.1).

Розрахунок прямих регресії. Нехай проведено n дослідів, результатами яких є наступні пари значень $(x_{cn\,i}, y_{cn\,i})$, $i=1,2,\ldots,n$. На практиці за наближені значення $M(X)$, $M(Y)$, $D(X)$ і $D(Y)$ беруть їх статистичні оцінки:

$$\tilde{m}_X=\bar{x}_n=\frac{1}{n}\sum_{i=1}^{n}x_{cn\,i}\ ;\ \tilde{m}_Y=\bar{y}_n=\frac{1}{n}\sum_{i=1}^{n}y_{cn\,i}\ ;$$

$$\tilde{D}_n[X]=\frac{1}{n-1}\sum_{i=1}^{n}(x_{cn\,i}-\bar{x}_n)^2\ ;\qquad \tilde{D}_n[Y]=\frac{1}{n-1}\sum_{i=1}^{n}(y_{cn\,i}-\bar{y}_n)^2\ .$$

Оцінкою $K[X,Y]$ є величина

$$\tilde{K}[X,Y]=\frac{1}{n-1}\sum_{i=1}^{n}(x_{cn\,i}-\bar{x}_n)(y_{cn\,i}-\bar{y}_n)\ .$$

Замінюючи величини $K[X,Y], \sigma_X, \sigma_Y$ їх статистичними оцінками $\tilde{K}[X,Y], \tilde{\sigma}_X=\sqrt{\tilde{D}_n[X]}, \tilde{\sigma}_Y=\sqrt{\tilde{D}_n[Y]}$, отримаємо наближені значення коефіцієнта кореляції і коефіцієнтів регресії:

$$r\approx\frac{\tilde{K}[X,Y]}{\tilde{\sigma}_X\tilde{\sigma}_Y}\ ;\ \rho(Y/X)\approx\frac{\tilde{K}[X,Y]}{\tilde{\sigma}_X^2}\ ;\ \rho(X/Y)\approx\frac{\tilde{K}[X,Y]}{\tilde{\sigma}_Y^2}\ ,$$

де $\dfrac{\tilde{K}[X,Y]}{\tilde{\sigma}_X\tilde{\sigma}_Y}$ і $\dfrac{\tilde{K}[X,Y]}{\tilde{\sigma}_X^2}$, $\dfrac{\tilde{K}[X,Y]}{\tilde{\sigma}_Y^2}$ – статистичні оцінки коефіцієнтів відповідно кореляції і регресії.

Підставляючи в рівняння регресії Y на X, тобто $y=\rho(Y/X)(x-m_X)+m_Y$, та рівняння регресії X на Y, тобто $x=\rho(X/Y)(y-m_Y)+m_X$, (див. §4.16) замість m_X, m_Y, $\rho(Y/X)$ і $\rho(X/Y)$ їх статистичні оцінки, отримаємо *вибіркові рівняння прямих регресій*:

$$y-\tilde{m}_Y=\frac{\tilde{K}[X,Y]}{\tilde{\sigma}_X^2}(x-\tilde{m}_X)\ ;\qquad x-\tilde{m}_X=\frac{\tilde{K}[X,Y]}{\tilde{\sigma}_Y^2}(y-\tilde{m}_Y)\ .$$

Приклад 6.3.2. Знайти вибіркове рівняння прямої регресії Y на X за даними $n=10$ спостережень.

Результати спостережень і відповідних обчислень подані в таблиці 6.3.2.

Табл. 6.3.2

$x_{cn\ i}$	$y_{cn\ i}$	$x_{cn\ i}-\widetilde{m}_X$	$(x_{cn\ i}-\widetilde{m}_X)^2$	$y_{cn\ i}-\widetilde{m}_Y$	$(x_{cn\ i}-\widetilde{m}_X)\cdot$ $\cdot(y_{cn\ i}-\widetilde{m}_Y)$
71	8.6	-4.5	20.25	-0.48	2.16
72	8.9	-3.5	12.25	-0.18	0.63
73	8.9	-2.5	6.25	-0.18	0.45
74	9.0	-1.5	2.25	0.08	-0.12
75	9.1	-0.5	0.25	0.02	-0.01
76	9.2	0.5	0.25	0.12	0.06
77	9.2	1.5	2.25	0.12	0.18
78	9.2	2.5	6.25	0.12	0.30
79	9.3	3.58	12.25	0.22	0.77
80	9.4	4.5	20.25	0.32	1.44
$\bar{x}_n=75.5$	$\bar{y}_n=9.08$		$\widetilde{\sigma}_X^2=9.17$		$\widetilde{K}[X,Y]=0.6$

Обчислюємо:

$$\frac{\widetilde{K}[X,Y]}{\widetilde{\sigma}_X^2}=\frac{0.65}{9.17}\approx 0.071\,.$$

Рівняння шуканої прямої має вигляд

$$y-9.08=0.071(x-75{,}5)\,,$$

або

$$y=0.071x+3.72\,.$$

Вправи для самостійного виконання

6.3.1. Перевірити, чи правильні твердження:

1. Коли кажуть, що характер розподілу ймовірностей на множині значень випадкової величини відомий, то це означає, що відомі:

а) поточковий це розподіл на скінченній множині $\Omega_X=\{x_1,\ x_2,\ ...,\ x_k\}$, чи поінтервальний на множині $\Omega_X=\bigcup\limits_{i=1}^{k}|[a_{i-1}\,;\,a_i)=[a_0\,;\,a_k)$;

б) дискретний цей розподіл, чи неперервний; абсолютно неперервний чи сингулярний;

в) функція розподілу чи щільність розподілу ймовірностей;

г) клас, до якого належить функція чи щільність розподілу ймовірностей;

д) усі параметри (числові характеристики) розподілу ймовірностей: математичне сподівання, дисперсія, центральні моменти m-го порядку тощо;

е) деякі параметри розподілу ймовірностей.

2. Стосовно кожного параметра θ розподілу ймовірностей на множині значень випадкової величини X можна дістати статистичну

оцінку θ^* за спостереженими значеннями випадкової величини X.

3. Оцінка θ^* є випадковою величиною, визначеною на сукупності всеможливих вибірок $(x_{cn\,1}, x_{cn\,2}, ..., x_{cn\,n})$ спостережених значень випадкової величини X, а тому $\theta^* = \theta^*(n)$ є послідовністю випадкових величин.

4. Кожна статистична оцінка θ^* параметра θ є незміщеною.

5. Для одного і того самого параметра за одними і тими самими спостереженими значеннями можна діставати статистичні оцінки, одні з яких систематично занижуються, а інші – систематично завищуються.

6. Кожна статистична оцінка є несуперечливою.

7. Ефективність статистичної оцінки θ^* тим вища, чим меншою є її дисперсія.

8. Вибіркове середнє – це те саме, що й статистичне середнє.

9. Вибіркове середнє є статистичною оцінкою математичного сподівання певної випадкової величини.

10. Якщо випадкова величина X спостерігається в серії з n незалежних спостережень (випробувань), то слова: “X_i – це випадкова величина X за i-го спостереження” означають, що X_i визначена на

$$\Omega^n = \{(E_1, E_2, ..., E_n) \mid E_i \in \Omega,\ i \in \overline{1,\ n}\},$$

причому $X_i(E_1, E_2, ..., E_n) = X(E)$, коли $E = E_i$, а інші координати $E_k, k \neq i$, можуть бути якими завгодно, якби тільки було $E_k \in \Omega$.

11. Якщо X_i – випадкова величина, визначена у твердженні 10, то

а) $X_i = X$ за всіх $i \in \overline{1, n}$;

б) $X_i = X$ принаймні за одного $i \in \overline{1, n}$;

в) $X_i \neq X$ за всіх $i \in \overline{1, n}$.

12. Статистичне середнє спостережених значень випадкової величини X є спостереженим значенням випадкової величини $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, де X_i з твердження 10.

13. Кожне значення $\overline{X}_n$ є значенням X.

14. Випадкова величина $\overline{X}_n$ є статистичною оцінкою математичного сподівання величини X.

15. Оцінка $\overline{X}_n$ є:

а) зміщеною; б) несуперечливою; в) ефективною.

16. Якщо $\overline{X}_n$ визначено як в 12, то $\overline{D}_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$ є:

а) фіксованим числом; б) випадковою величиною, визначеною на Ω, тобто там де визначена X; в) випадковою величиною, визначеною на Ω^n.

17. Випадкова величина $\overline{D}_n$ є статистичною оцінкою дисперсії розподілу ймовірностей на множині значень випадкової величини X .

18. Оцінка $\overline{D}_n$ є:

а) незміщеною; б) несуперечливою; в) ефективною.

19. За будь-якої випадкової величини X її статистична дисперсія $\overline{D}_n$ відрізняється від виправленої статистичної дисперсії $\tilde{D}_n$.

20. Статистична оцінка $\tilde{D}_n$ незміщена, а $\overline{D}_n$ – зміщена.

21. Статистична оцінка $\tilde{D}_n$ несуперечлива, а $\overline{D}_n$ – суперечлива.

22. Можна стверджувати, що коли n досить велике, то

$$\bar{x}_n = \tilde{m}_X \approx M[X], \quad \tilde{D}_n(X) \approx D(X), \quad \tilde{K}[X,Y] \approx K[X,Y]$$

у тому розумінні, що:

а) будь-яке значення в лівій частині рівності обов'язково досить мало відрізняється від значення в правій частині відповідної рівності;

б) окремі значення в лівій частині рівності можуть досить сильно відрізнятися від значення в правій частині відповідної рівності, проте ймовірність цього:

1) дорівнює нулю; 2) є як завгодно близькою до нуля, коли n досить велике.

23. Вибіркове рівняння регресії є наближеним рівнянням регресії.

6.4. Надійні інтервали. Надійна ймовірність

Нехай θ^* – статистична оцінка параметра θ. Природно постає питання, яка похибка може виникати, якщо замінити значення θ спостереженим значенням θ_0^* його статистичної оцінки θ^*, наскільки може відхилятися значення θ^* від істинного значення θ, чому дорівнює $P(|\theta^* - \theta| \le \alpha)$ – ймовірність того, що значення θ^* відхиляються від θ не більше, ніж на α. Якщо α досить мале, а ймовірність $P(|\theta^* - \theta| \le \alpha) = \beta$, то з імовірністю β можна вважати, що коли замінити невідоме значення параметра θ спостереженим значенням θ_0^* випадкової величини θ^*, похибка буде досить малою і не перевищуватиме α. Задача саме й полягає у знаходженні такого числа α, за якого з наперед заданою ймовірністю β виконувалась би нерівність $|\theta^* - \theta| \le \alpha$, тобто справджувалась рівність

$$P(|\theta^* - \theta| \le \alpha) = P(\theta - \alpha \le \theta^* \le \theta + \alpha) = P(\theta^* - \alpha \le \theta \le \theta^* + \alpha) = \beta. \quad (6.4.1)$$

Або ж навпаки – за заданою точністю α потрібно знайти, з якою ймовірністю буде виконуватися нерівність $|\theta^* - \theta| \le \alpha$.

Рівність (6.4.1) означає, що з імовірністю β невідоме значення параметра θ повинно знаходитися в інтервалі $(\theta_0^* - \alpha\,; \theta_0^* + \alpha)$.

Оскільки величина θ^* випадкова, то й інтервал $(\theta^* - \alpha\,; \theta^* + \alpha)$ є випадковим завдовжки 2α. Таким чином, якщо знайдено α, то тим самим визначено випадковий інтервал $(\theta^* - \alpha\,; \theta^* + \alpha)$, яким з імовірністю β накривається невідоме значення параметра θ.

Ймовірність β називають *надійною ймовірністю* (*надійністю*) оцінки θ^*, інтервал $(\theta^* - \alpha, \theta^* + \alpha)$ – *надійним інтервалом*, а його межі – *надійними межами*.

Задача знаходження надійного інтервала стосовно оцінки параметра θ з даною надійною ймовірністю β у загальному випадку є досить складною, оскільки розподіл імовірностей на множині значень випадкової величини θ^* невідомий і ймовірність

$$P(|\theta^* - \theta| \le \alpha) = P(\theta - \alpha \le \theta^* \le \theta + \alpha)$$

таким чином обчислити неможливо, до того ж проміжок $(\theta - \alpha\,; \theta + \alpha)$ невідомий, оскільки невідомий параметр θ.

Проте іноді за потреби знаходження ймовірності $P(|\theta^* - \theta| \le \alpha)$ можна вивести формулу, до якої параметр θ явно не входить. Так, в разі обчислення ймовірності того, що значення випадкової величини X з нормальним розподілом ймовірностей на множині її значень знаходитиметься в проміжку, симетричному щодо математичного сподівання, використовується формула, в яку не входить математичне сподівання. Цей факт і використовують в разі обчислення ймовірності

$$P(\theta - \alpha \le \theta^* \le \theta + \alpha) \text{ або } P(\theta^* - \alpha \le \theta \le \theta^* + \alpha).$$

Як відомо, коли розподіл імовірностей на множині значень випадкової величини X нормальний, то й розподіл імовірностей на множині значень випадкової величини $\overline{X}_n = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ (див. §6.3) також нормальний, причому

$$M[\overline{X}_n] = M[X] = m, \; \sigma[\overline{X}_n] = \frac{\sigma[X]}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}.$$

Тому якщо відоме середнє квадратичне відхилення випадкової величини X з нормальним розподілом ймовірностей на множині її значень, одержимо:

$$P(|\overline{X}_n - m| \le \alpha) = \frac{2}{\sqrt{2\pi}} \int\limits_0^{\frac{\alpha\sqrt{n}}{\sigma}} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int\limits_0^{t} e^{-\frac{x^2}{2}} dx,$$

де $t = \dfrac{\alpha\sqrt{n}}{\sigma}$. Визначивши параметр t так, щоб виконувалася рівність $P(|\overline{X}_n - m| \le \alpha) = \beta$, знайдемо інтервал $\left(\overline{X}_n - t\dfrac{\sigma}{\sqrt{n}};\; \overline{X}_n + t\dfrac{\sigma}{\sqrt{n}}\right)$, або $(\overline{X}_n - \alpha; \overline{X}_n + \alpha)$, яким із заданою надійністю накривається невідомий параметр m.

Отже, якщо розподіл ймовірностей на множині значень випадкової величини X нормальний і параметр σ відомий, то стосовно невідомого параметра m можна встановити надійний інтервал $(\overline{X}_n - \alpha; \overline{X}_n + \alpha)$, яким з надійністю β накривається невідоме m. В такому разі $\alpha = t\dfrac{\sigma}{\sqrt{n}}$, де t визначають так, щоб виконувалася рівність

$$\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx = \beta .$$

Наближений розв'язок рівняння $\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx = \beta$ може бути знайдений за допомогою програми GRAN1 (чи інших математичних програм) або за допомогою таблиці значень функції Лапласа $\Phi(x)$.

Приклад 6.4.1. Учні виконують стрибки у висоту. Висота стрибка навмання взятого учня є випадковою величиною з нормальним розподілом імовірностей, причому середнє квадратичне відхилення $\sigma = 5см$. За даними 100 спостережень здобута статистична оцінка для математичного сподівання досліджуваної випадкової величини $\overline{X}_n = 135см$. Знайти інтервал, в якому з ймовірністю 0.9 знаходиться математичне сподівання висоти стрибка учня (необхідно знайти таке t, щоб виконувалася рівність $\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx = 0.9$).

Скориставшись програмою GRAN1, дістанемо, що така рівність виконується, коли $t = 1.65$. Отже $\frac{\alpha \cdot \sqrt{100}}{5} = 1.65$, звідки $\alpha = 0.82$. Той самий результат одержимо, використовуючи таблицю значень функції Лапласа (див. додадок 2). Таким чином, з ймовірністю 0.9 невідоме математичне сподівання розглядуваної випадкової величини знаходиться в межах від 134.18 до 135.82 см.

Зауваження. Якщо стосовно випадкової величини з нормальним розподілом ймовірностей на множині її значень треба знайти кількість спостережень, які необхідно виконати, щоб дістати оцінку математичного сподівання, яка з надійністю β відхиляється від невідомого математичного сподівання не більше, ніж на α, то таке число спостережень визначається за формулою

$$t = \frac{\alpha\sqrt{n}}{\sigma}, \text{ тобто } n = \frac{t^2\sigma^2}{\alpha^2},$$

де t знаходять так, щоб виконувалась рівність $\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx = \beta$, що можна здійснити за допомогою програми GRAN1, або за таблицею значень функції Лапласа $\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx$ визначити t так, щоб виконувалась рівність $\Phi(t)=\frac{\beta}{2}$.

Приклад 6.4.2. Відомо, що зріст дванадцятирічної дитини є випадковою величиною з нормальним розподілом імовірностей, причому середнє квадратичне відхилення $\sigma=5\,\text{см}$. Скільки треба виконати спостережень, щоб знайти інтервал завширшки 4 см, яким з імовірністю 0.95 накривалось би невідоме математичне сподівання досліджуваної випадкової величини?

Скориставшись програмою GRAN1, дістанемо $\frac{2}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx=0.95$, коли $t=1.96$. За таблицею значень функції Лапласа для $\Phi(t)=0.4750$ дістанемо той самий результат. Оскільки $\alpha=2$, то $n=\frac{(1.96)^2\cdot 25}{4}\approx 24$.

Отже, за даних умов потрібно виконати 24 спостереження, щоб знайти інтервал завширшки 4 см, в якому з імовірністю 0.95 знаходиться невідоме математичне сподівання зросту навмання взятої дванадцятирічної дитини.

Якщо на множині значень випадкової величини X розподіл імовірностей нормальний, а величина σ невідома, то наведені міркування не приводять до мети, визначити надійний інтервал за розглянутими формулами неможливо. Якщо число спостережень досить велике, то замість σ можна використати його статистичну оцінку $\bar{\sigma}_n$. Однак, коли кількість спостережень невелика, така заміна може привести до значних похибок.

Аналогічно може бути побудований надійний інтервал і стосовно дисперсії.

Нехай виконано n незалежних спостережень за випадковою величиною X з невідомими параметрами $M[X]=m$, $D[X]=\sigma^2=D$ і стосовно дисперсії знайдено незміщену оцінку

$$\tilde{D}_n=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{X}_n)^2,$$

де $\bar{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$.

Величини $(X_i-\bar{X}_n)$ не є незалежними, оскільки $\bar{X}_n$ залежить від X_i. Проте відомо, що із збільшенням n розподіл ймовірностей на множині значень випадкової величини $\tilde{D}_n$ прямує до нормального. На практиці можна цей розподіл вважати нормальним вже за значень n порядку 20-30.

Параметри цього розподілу є

$$M[\overline{D}_n] = D\,,\ D[\overline{D}_n] = \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)}D^2\,,$$

де $\mu_4 = M[(X - M[X])^4]$ – четвертий центральний момент випадкової величини X.

Коли n досить великі, замість D можна використати оцінку $\tilde{D}_n$, а замість μ_4 – оцінку

$$\mu_{4,n}^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^4\ .$$

Слід мати на увазі, що за невеликих n така заміна може призвести до значних похибок, оскільки коли число спостережень невелике, моменти високих порядків визначаються досить неточно.

Побудуємо надійний інтервал стосовно середнього квадратичного відхилення.

Виявляється, що з надійністю β можна стверджувати, що в надійному інтервалі $(\tilde{\sigma}_n - \tilde{\sigma}_n q;\, \tilde{\sigma}_n + \tilde{\sigma}_n q)$ знаходиться невідоме значення параметра σ; точність оцінки $\delta = \tilde{\sigma}_n q$.

В додатку 8 наведена таблиця значень $q = q(\beta, n)$ за різних значень n і надійності β.

Приклад 6.4.3. Розподіл ймовірностей на множині значень ознаки X нормальний. Знайти надійний інтервал стосовно σ з надійністю $\beta = 0.95$, в разі, коли $n = 20$; $\tilde{\sigma}_n = 0.40$.

За надійності $\beta = 0.95$ і $n = 20$ знаходимо в таблиці (додаток 8) $q = 0.37$. Далі $\tilde{\sigma}_n q = 0.40 \cdot 0.37 \approx 0.15$. Таким чином межі надійного інтервалу $0.40 - 0.15 = 0.25$ і $0.40 + 0.15 = 0.25$. Отже, в надійному інтервалі $(0.25; 0.55)$ знаходиться значення σ з надійністю 0.95.

Якщо вид розподілу ймовірностей на множині значень випадкової величини X відомий, а невідомими є лише параметри розподілу, можна спробувати виразити μ_4 через D. Якщо, наприклад, розподіл імовірностей на множині значень випадкової величини X нормальний, то $\mu_4[X] = 3D^2[X] = 3D^2$, звідки

$$D[\tilde{D}_n] = \frac{3}{n}D^2 - \frac{n-3}{n(n-1)}D^2 = \frac{2}{n-1}D^2. \qquad (6.4.2)$$

Замінюючи невідоме значення параметра D його оцінкою, дістаємо

$$D[\tilde{D}_n] = \frac{2}{n-1}\overline{D}_n^2, \; \sigma[\tilde{D}_n] = \sqrt{\frac{2}{n-1}}\overline{D_n} . \qquad (6.4.3)$$

Якщо, наприклад, розподіл імовірностей на множині значень випадкової величини X рівномірний на відрізку $[a\,;b]$, то

$$\mu_4[X] = \frac{(b-a)^4}{80}, \; D[X] = \frac{(b-a)^2}{12},$$

тобто

$$\mu_4[X] = 1.8D^2[X],$$

$$D[\tilde{D}_n] = \frac{1}{n}1.8D^2 - \frac{n-3}{n(n-1)}D^2 = \frac{0.8n+1.2}{n(n-1)}D,$$

звідки

$$\sigma[\tilde{D}_n] = \sqrt{\frac{0.8n+1.2}{n(n-1)}}D .$$

Якщо вид розподілу ймовірностей на множині значень випадкової величини X невідомий, для наближеного визначення $\sigma[\tilde{D}_n]$ доцільно користуватися формулами (6.4.3), якщо немає підстав вважати, що цей розподіл різко відрізняється від нормального (близький до рівномірного або, навпаки, щільність розподілу імовірностей надто швидко зростає чи спадає).

Якщо орієнтовне значення $\sigma[\tilde{D}_n]$ тим або іншим способом знайдено, то надійний інтервал для дисперсії будується так само, як для математичного сподівання

$$I_\beta = (\tilde{D}_n - t_\beta \sigma[\tilde{D}_n]\,;\, \tilde{D}_n + t_\beta \sigma[\tilde{D}_n]),$$

де t_β знаходиться з рівності

$$P(|\tilde{D}_n - D| \le t_\beta \sigma[\tilde{D}_n]) = \beta ,$$

тобто з рівності

$$\frac{2}{\sqrt{2\pi}}\int_0^{t_\beta} e^{-\frac{t^2}{2}}\,dt = \beta ,$$

де β – надійна ймовірність оцінки $\tilde{D}_n$.

Розглянуті методи побудови надійних інтервалів для математичного сподівання, дисперсії й середнього квадратичного відхилення досить неточні. Для більш точного визначення надійних інтервалів треба заздалегідь знати вид розподілу ймовірностей на множині значень випадкової величини X, тоді як для розглянутих наближених методів це не обов'язково.

Точні методи використовують для оцінювання шуканих параметрів, які є функціями випадкових величин $X_1, X_2, \ldots, X_n$, проте розподіл імовірностей на множинах значень яких не залежить від значень невідомих параметрів, а залежить від числа n спостережень і виду розподілу ймовірностей на множині значень випадкової величини X. Такі оцінки відіграють важливу роль у математичній статистиці. Найбільш ґрунтовно вони вивчені для випадку, коли розподіл імовірностей на множині значень випадкової величини X нормальний.

Відомо, наприклад, що коли розподіл імовірностей на множині значень випадкової величини X нормальний, то розподіл імовірностей на множині значень випадкової величини

$$T_n = \sqrt{n}\,\frac{\overline{X}_n - m}{\sqrt{\tilde{D}_n}} \qquad (6.4.4)$$

є *розподілом Стьюдента* з $r = n-1$ ступенем вільності. Щільність цього розподілу ймовірностей подають у вигляді

$$s_{n-1}(t) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{(n-1)\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}, \qquad (6.4.5)$$

де $\Gamma(x)$ – гама-функція:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\,dt\,.$$

Відомо також, що розподіл імовірностей на множині значень випадкової величини

$$V_n = \frac{(n-1)\overline{D}_n}{D}$$

є χ^2-розподілом (хі-квадрат розподілом) з $n-1$ ступенями вільності, щільність якого подають у вигляді

$$k_{n-1}(t) = \begin{cases} \dfrac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}\, t^{\frac{n-1}{2}-1} e^{-\frac{t}{2}}, & \text{коли } t \ge 0; \\ 0, & \text{коли } t < 0. \end{cases}$$

Графік щільності $k_{n-1}(t)$ подано на Рис. 6.4.1

Нехай треба знайти надійний інтервал, в якому з імовірністю β знаходиться невідоме математичне сподівання досліджуваної випадкової величини X, тобто стосовно якого виконується умова

$$P(|\overline{X}_n - m| \le \alpha) = \beta .$$

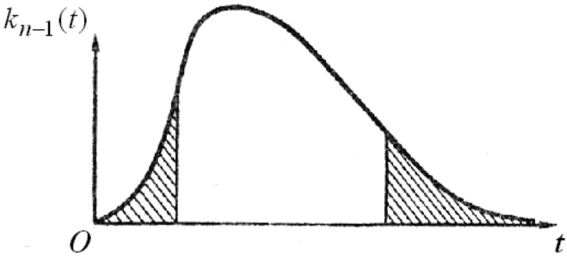


Рис. 6.4.1

Подамо останню рівність у вигляді

$$P\left(\frac{\sqrt{n}\,|\,\overline{X}_n - m\,|}{\sqrt{\tilde{D}_n}} \le \frac{\alpha}{\sqrt{\dfrac{\tilde{D}_n}{n}}} \right) = \beta \quad ,$$

або, враховуючи (6.4.4),

$$P\left(|\,T_n\,| \le \frac{\alpha}{\sqrt{\dfrac{\tilde{D}_n}{n}}} \right) = \beta .$$

Число t_β, за якого $P(|T_n| \le t_\beta) = \beta$ знаходять з умови

$$P(|T_n| \le t_\beta) = \int_{-t_\beta}^{t_\beta} s_{n-1}(t)dt = \beta ,$$

або, враховуючи, що $s_{n-1}(t)$ парна функція, з умови

$$2\int_0^{t_\beta} s_{n-1}(t)dt = \beta .$$

Значення функції

$$\Psi(x) = 2\int_0^x s_{n-1}(t)dt ,$$

обчислені за різних значень $n-1$ і x, зведені в таблицю (див. додаток 6).

Визначивши t_β за даними β і $n-1$, знайдемо $\alpha = t_\beta \sqrt{\dfrac{\tilde{D}_n}{n}}$.

Щоб побудувати надійний інтервал стосовно дисперсії, подамо випадкову величину $\tilde{D}_n$ у вигляді

$$\widetilde{D}_n = V_n \frac{D}{n-1},$$

де V_n – випадкова величина з χ^2-розподілом імовірностей.

Щоб вибрати інтервал I_β, в якому значення випадкової величини V_n знаходиться з імовірністю β, інтервал I_β добирають так, щоб імовірності попадання за його межі ліворуч і праворуч були однакові і дорівнювали

$$\frac{\alpha}{2} = \frac{1-\beta}{2}.$$

Використовуючи таблицю значень χ^2 (див. додаток 5) залежно від $r = n-1$ і α, знайдемо два значення χ_1^2 і χ_2^2, за яких

$$P(V_n < \chi_1^2) = \frac{\alpha}{2},\ P(V_n > \chi_2^2) = 1 - \frac{\alpha}{2},$$

де χ_1^2 і χ_2^2 – межі інтервалу I_β.

Оскільки нерівності

$$\frac{\widetilde{D}_n(n-1)}{\chi_1^2} < D,\ \frac{\widetilde{D}_n(n-1)}{\chi_2^2} > D$$

рівносильні нерівностям

$$V_n < \chi_1^2 \text{ і } V_n > \chi_2^2,$$

то інтервал

$$\left(\frac{\widetilde{D}_n(n-1)}{\chi_1^2}; \frac{\widetilde{D}_n(n-1)}{\chi_2^2} \right)$$

є шуканим надійним інтервалом стосовно дисперсії D.

Зазначимо, що коли X_0, X_1, … , X_r незалежні випадкові величини з нормальними розподілами ймовірностей з параметрами $M[X_i] = 0$, $\sigma[X_i] = 1$, то розподіл ймовірностей на множині значень випадкової величини

$$T_r = \frac{X_0}{\sqrt{\frac{1}{r}\sum_{i=1}^{r} X_i^2}}$$

є розподілом Стьюдента, щільність якого

$$f_{T_r}(x) = b_r \left(1 + \frac{x^2}{r}\right)^{-\frac{r+1}{2}},$$

де b_r – деяке число, що залежить від r, $b_r \to \frac{1}{\sqrt{2\pi}}$, коли $r \to \infty$.

Графік щільності $f_{T_r}(x)$ симетричний відносно прямої $x=0$ і нагадує графік щільності нормального розподілу ймовірностей.

Коли r необмежено збільшувати, щільність $f_{T_r}(x)$ наближатиметься до щільності нормального розподілу ймовірностей з параметрами $m=0$ і $\sigma=1$. Якщо $M[X_i]=m$, то $M[X_i-m]=0$, причому розподіл імовірностей на множині значень випадкової величини (X_i-m) залишається нормальним з параметрами 0 і 1. Таким чином, розподіл ймовірностей на множині значень випадкової величини

$$\frac{X_0-m}{\sqrt{\frac{1}{r}\sum_{i=1}^{r}(X_i-m)^2}}$$

є розподілом Стьюдента.

Оскільки (X_i-m)– випадкова величина з нормальним розподілом імовірностей, то

$$\frac{1}{r}\sum_{i=1}^{r}(X_i-m)^2=\frac{1}{r}\chi_r^2.$$

Отже, на множині значень випадкової величини

$$T_r=\frac{X_0}{\sqrt{\frac{1}{r}\chi_r^2}},$$

має місце розподіл імовірностей Стьюдента, де χ_r^2 – випадкова величина з χ^2-розподілом імовірностей. У додатку 6 наведено таблицю розподілу ймовірностей Стьюдента, за якою можна знайти $t_{r\,kp}$, за якого $P(|T_r|<t_{r\,kp})=1-\alpha$. За великих r, $(r>30)$, ці ймовірності визначаються за нормальним розподілом імовірностей.

Приклад 6.4.4. На фермі випробовували вплив вітамінів на приріст маси телят. З цією метою було оглянуто 20 телят одного віку. Середня маса їх виявилася рівною 340 кг, а відповідне “виправлене” середнє квадратичне відхилення – 20 кг.

Визначити: 1) надійний інтервал стосовно математичного сподівання m з надійністю 0.95; 2) надійний інтервал стосовно середнього квадратичного відхилення з тією самою надійністю.

Розв’язуючи задачу, виходити з припущення, що розподіл ймовірностей на множині значень досліджуваної величини нормальний.

Розв’язування. 1) Згідно умов задачі $\bar{x}_n=340$; $\tilde{\sigma}_n=20$;

$\beta = 0.95$; $n = 20$.

Користуючись розподілом Стьюдента, за надійності $\beta = 0.95$ і $n = 20$ знаходимо в таблиці (додаток 6) $t_\beta = 2.093$.

Отже $\delta = 2.093\frac{20}{\sqrt{20}} \approx 9.4$, звідки межі надійного інтервалу $340 - 9.4 = 330.6$ і $340 + 9.4 = 349.4$. Таким чином, надійним інтервалом $(330.6; 349.4)$ накривається значення m з надійністю 0.95.

Можна вважати, що в даному випадку справжня маса виміряна досить точно (відхилення близько $\frac{9.4}{340} \approx 0.03$).

2) Стосовно надійності $\beta = 0.95$ і $n = 20$ знаходимо в таблиці (додаток 7) $q = 0.37$. Далі $\tilde{\sigma}_n q = 20 \cdot 0.37 = 7.4$. Межі надійного інтервалу $20 - 7.4 = 12.6$ і $20 + 7.4 = 27.4$. Таким чином, $12.6 < \sigma < 27.4$, звідки можна зробити висновок, що σ визначено незадовільно (відхилення близько $\frac{\tilde{\sigma}_n q}{\tilde{\sigma}_n} = q \approx 0.4$ – майже половина). Щоб знайти вужчий надійний інтервал з такою самою надійністю, необхідно збільшити число n випробувань.

Примітка. Вище передбачалося, що $q < 1$. Якщо $q > 1$, то, враховуючи, що $\sigma > 0$, одержуємо $0 < \sigma < \tilde{\sigma}_n + \tilde{\sigma}_n q$. Значення q в цьому випадку визначаються за таблицею із додатку 8.

Приклад 6.4.5. Ознака X є випадковою величиною з нормальним розподілом ймовірностей. За вибіркою об'єму $n = 10$ знайдено “виправлене” середнє квадратичне відхилення $\tilde{\sigma}_n = 0.16$. Знайти надійний інтервал стосовно σ з надійністю 0.999.

За надійності $\beta = 0.999$ і $n = 10$ за таблицею із додатку 8 знаходимо $q = 1.80$.

Отже, шуканий надійний інтервал такий:

$$0 < \sigma < 0.16 + 0.16 \cdot 1.80,$$

або

$$0 < \sigma < 0.448.$$

Вправи для самостійного виконання

6.4.1. Перевірити, чи правильні твердження:

1. Стосовно кожної оцінки параметра θ розподілу ймовірностей на множині значень випадкової величини X і будь-якого числа β існує єдиний надійний інтервал, в якому значення параметра θ знаходиться з ймовірністю β.

2. Межі надійного інтервалу є:

а) фіксованими числами; б) значеннями певної випадкової величини.

3. Надійність оцінки θ^* – це ймовірність попадання значень випадкової величини θ^* у фіксований окіл відповідного значення θ.

4. У рівняння $P(\theta - \alpha \le \theta^* \le \theta + \alpha) = \beta$ відносно змінної α завжди існує розв’язок.

5. Якщо у рівняння $P(\theta - \alpha \le \theta^* \le \theta + \alpha) = \beta$ існує розв’язок відносно змінної α, то цей розв’язок єдиний.

6. Якщо X_k, $k \in N$, – випадкові величини з одним і тим самим нормальним розподілом ймовірностей на множинах їх значень і відомим параметром $\sigma = \sigma[X_k] = \sigma[X] = \sqrt{D[X]}$, то за спостереженими значеннями випадкової величини X невідомий параметр $m = M[X_k] = M[X]$ (однаковий для всіх X_k), можна

а) знайти з довільною точністю;

б) накрити надійним інтервалом як завгодно малої довжини з ймовірністю, як завгодно близькою до 1;

в) вважати таким, значення якого знаходиться в фіксованому інтервалі $(\bar{x}_n - \alpha;\ \bar{x}_n + \alpha)$, де $\bar{x}_n$ – фіксоване спостережене значення величини $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$.

г) вважати таким, значення якого з ймовірністю β знаходиться в інтервалі $(\bar{x}_n - \alpha;\ \bar{x}_n + \alpha)$, де $\bar{x}_n$ – навмання вибране спостережене значення величини $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$.

7. Точність оцінки α і надійність β завжди пов’язані рівністю

$$\frac{2}{\sqrt{2\pi}} \int_0^{\alpha\sqrt{n}/\sigma} e^{-t^2}\, dt = \beta.$$

8. Завжди $M[\tilde{D}_n] = D = D[X]$, а $D[\tilde{D}_n] = \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)} D^2$, де $\mu_4 = M[(X - M[X])^4]$ – четвертий центральний момент випадкової величини.

9. Стосовно будь-якої випадкової величини X правильна рівність:

а) $\mu_4[X] = 3D^2[X]$; б) $D[\tilde{D}_n] = \frac{2}{n-1} D^2$;

в) $\mu_4[X] = \dfrac{(b-a)^4}{80}$; г) $\mu_4[X] = 1.8 \cdot D^2[X]$;

д) $D[\tilde{D}_n] = \dfrac{0.8n+1.2}{n(n-1)} D^2$; е) $\sigma[\tilde{D}_n] = \sqrt{D[\tilde{D}_n]}$.

10. Надійний інтервал стосовно невідомої дисперсії випадкової величини X з нормальним розподілом ймовірностей на множині її значень знаходять так само, як і стосовно математичного сподівання.

11. Стосовно невідомої дисперсії точність оцінки α і надійність β пов’язані рівністю

$$\frac{2}{\sqrt{2\pi}} \int_0^{\alpha/\sigma[\overline{D}_n]} e^{-t^2/2} dt = \beta.$$

12. Якщо n досить велике, то щільність розподілу Стьюдента з $(n-1)$ ступенями вільності набуває вигляду:

$$S_{n-1}(t) \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \; .$$

13. Розподіл Стьюдента – це розподіл ймовірностей на множині значень випадкової величини

$$T_n = \sqrt{n} \, \frac{\overline{X}_n - m}{\sqrt{\tilde{D}_n}} \; .$$

14. Розподіл χ^2 – це розподіл ймовірностей на множині значень випадкової величини $V_n = \dfrac{(n-1)\tilde{D}_n}{D}$.

15. Стосовно невідомого математичного сподівання точність оцінки α і надійність β пов’язані за рівністю

$$2 \int_0^{\sqrt{n}\alpha/\sqrt{\tilde{D}_n}} S_{n-1}(t) dt = \beta \; .$$

16. Стосовно невідомої дисперсії випадкової величини X з нормальним розподілом ймовірностей надійний інтервал визначається як $\left(\dfrac{(n-1)\tilde{D}_n}{\chi_2^2}; \; \dfrac{(n-1)\tilde{D}_n}{\chi_1^2} \right)$, де χ_1^2 та χ_2^2 визначаються відповідно з умов:

$$\chi_1^2 < \chi_2^2\,, \quad \int_0^{\chi_1^2} k_{n-1}(t)dt = \frac{\alpha}{2}\,, \quad \int_0^{\chi_2^2} k_{n-1}(t)dt = 1 - \frac{\alpha}{2}\,.$$

6.5. Статистична перевірка гіпотез

Однією з основних задач математичної статистики є визначення розподілу ймовірностей або параметрів цього розподілу за статистичними даними. Така задача іноді ставиться так.

Розглядають деяку гіпотезу про те, що розподіл імовірностей є того чи іншого типу, або параметри розподілу набувають тих або інших значень. Задача полягає в тому, щоб на основі вивчення статистичних даних підтвердити правильність висунутої гіпотези чи спростувати її.

На основі статистичних даних знаходять розподіл статистичних імовірностей або статистичну оцінку шуканого параметра і оцінку (норму) відхилення U статистичного розподілу ймовірностей від гіпотетичного розподілу чи статистичної оцінки шуканого параметра від його гіпотетичного значення.

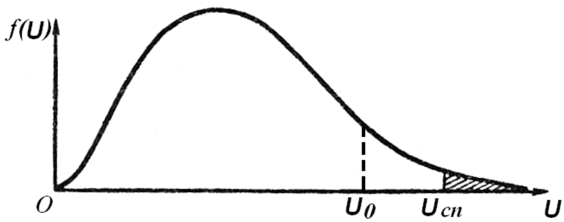


Рис. 6.5.1

Іноді вдається визначити розподіл ймовірностей на множині значень випадкової величини U. В такому разі можна визначити ймовірність $P(U \geq U_0)$ – ймовірність того, що випадкова величина U набуває значення не меншого, ніж деяке значення U_0. Якщо ця ймовірність дуже мала (Рис. 6.5.1), то подію $(U \geq U_0)$ слід вважати практично неможливою. Тому, якщо виявиться, що $U_{cn} \geq U_0$, то гіпотезу про близькість розподілу статистичних ймовірностей і гіпотетичного розподілу ймовірностей чи статистичної оцінки шуканого параметра від його гіпотетичного значення слід вважати такою, що не узгоджується із статистичними даними. Якщо ж виявиться, що $U_{cn} < U_0$, то вважають, що статистичний матеріал не суперечить розглядуваній гіпотезі про вигляд розподілу ймовірностей на множині значень випадкової величини, тобто гіпотеза узгоджується із статистичними даними.

Область S, в разі попадання в яку спостереженого значення гіпотеза відхиляється, називається *критичною.*

Рівнем значущості критерію перевірки статистичної гіпотези називається ймовірність того, що застосування критерію не приведе до відхилення гіпотези в разі її істинності. Іншими словами, рівень значущості – це ймовірність того, що в разі істинності гіпотези спостережене значення не знаходитиметься в критичній області.

Рівень значущості α часто задають наперед (досить близьким до 1), виходячи з особливостей задачі, і за рівнем значущості знаходять величину $U_{кр}$ так, щоб справджувалася рівність

$$P(U < U_{кр}) = \alpha .$$

Якщо за заданого α виконується нерівність $U_{сп} > U_{кр}$, то вважають, що розглядувана гіпотеза про вигляд розподілу ймовірностей не узгоджується із статистичними даними, оскільки ймовірність $P(U \ge U_{кр}) = 1 - \alpha$ досить мала, якщо α близьке до 1, а події, ймовірності яких не перевищують значення $1-\alpha$, вважаються неможливими. Імовірність (ризик) того, що в цьому випадку буде відхилена правильна гіпотеза, дорівнює $1-\alpha$. Якщо за статистичними даними виявляється $U_{сп} < U_{кр}$, то немає підстав відхиляти гіпотезу, оскільки статистичні дані не є суперечливими з гіпотезою про те, що розподіл імовірностей є саме того типу, про який йдеться в розглядуваній гіпотезі.

Якщо спостережене значення $U_{сп}$ знаходиться в критичній області, то гіпотеза відхиляється і приймається альтернативна гіпотеза, яка є запереченням основної гіпотези. Основною гіпотезою слід обирати таку гіпотезу, яку небажано помилково відхилити.

Критерії, за якими перевіряють, узгоджуються чи ні статистичні дані з розглядуваними гіпотезами, називають *критеріями узгодження.*

Одним з найбільш поширених критеріїв узгодження, за допомогою яких перевіряють гіпотезу про тип розподілу ймовірностей, є так званий *«критерій χ^2» Пірсона*.

Стосовно розподілу ймовірностей на неперервній множині $\Omega_X = [a_0\,; a_k)$ значень випадкової величини X критерій Пірсона застосовується так. За статистичними даними будують поінтервальний розподіл статистичних ймовірностей (Табл. 6.5.1) і крім того на основі гіпотетичного розподілу ймовірностей знаходять гіпотетичні ймовірності попадання значень досліджуваної випадкової величини в часткові інтервали I_i, $i = 1, 2, ..., k$.

Табл. 6.5.1

$I_i = [a_{i-1}\,; a_i)$	$[a_0\,; a_1)$	$[a_1\,; a_2)$	$\dots$	$[a_{k-1}\,; a_k)$
$P^*_{nX}(I_i)$	$P^*_{nX}([a_0\,; a_1))$	$P^*_{nX}([a_1\,; a_2))$	$\cdots$	$P^*_{nX}([a_{k-1}\,; a_k))$

Відхилення U статистичного розподілу від гіпотетичного визначають за формулою (це відхилення позначають через χ^2):

$$U = \chi^2 = n\sum_{i=1}^{k} \frac{(P_{nX}^{*}(I_i) - p_i)^2}{p_i}, \qquad (6.5.1)$$

або

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}, \qquad (6.5.2)$$

де n_i – кількість спостережених значень досліджуваної випадкової величини X, що знаходяться в *і-му* частковому інтервалі I_i, $i = 1, 2, ..., k$; n – кількість спостережень;

$$p_i = P_X(I_i) = P_X([a_{i-1}; a_i))$$

гіпотетичні ймовірності попадання в інтервали I_i. На множині значень випадкової величини U має місце χ^2-розподіл імовірностей. Цей χ^2-розподіл не залежить від невідомого розподілу ймовірностей досліджуваної випадкової величини, а залежить лише від числа r ступенів вільності. Число ступенів вільності r обчислюють за формулою $r = k - 1 - s$, де k – число часткових інтервалів I_i, s – число параметрів гіпотетичного розподілу, які оцінюються за статистичним матеріалом. Стосовно χ^2-розподілу складені спеціальні таблиці (див. додаток 5). За цими таблицями за даними r і α можна знайти $\chi^2_{кр}$, за якого $P(\chi^2 < \chi^2_{кр}) = \alpha$, де α досить близьке до 1, $0 < \alpha < 1$. Очевидно, коли $\chi^2_{сп} \geq \chi^2_{кр}$, то $P(\chi^2 \geq \chi^2_{сп}) \leq 1 - \alpha$ (Рис. 6.5.2), оскільки в цьому разі ймовірність того, що значення випадкової величини $U = \chi^2$ буде знаходитись в інтервалі $[\chi^2_{сп}; +\infty)$, не може бути більшою, ніж ймовірність того, що воно буде знаходитись в інтервалі $[\chi^2_{кр}; +\infty)$. Таким чином, якщо $\chi^2_{сп} \geq \chi^2_{кр}$, то $\chi^2_{сп}$ знаходиться в критичній області, а тому розглядувану гіпотезу про тип розподілу ймовірностей слід відхилити, оскільки вона за даного рівня значущості не узгоджується із статистичними даними.

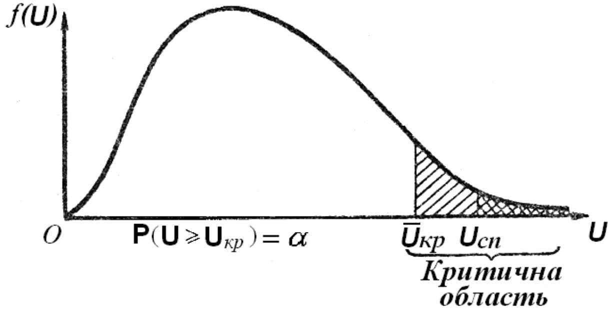


Рис. 6.5.2

Нехай поінтервальний розподіл статистичних ймовірностей на деякому проміжку $[a_0\,;a_k)$ визначається за таблицею 6.5.2 або за щільністю розподілу $f_n^*(x)$ (див. §2.2, §2.3).

Табл. 6.5.2

$[a_{i-1};a_i)$	$[a_0;a_1)$	$[a_1;a_2)$	...	$[a_{k-1};a_k)$
$P_n^*([a_{i-1};a_i))$	$P_n^*([a_0;a_1))$	$P_n^*([a_1;a_2))$	...	$P_n^*([a_{k-1};a_k))$

Часто на практиці виникає питання, чи не можна замінити функцію $f_n^*(x)$ деякою іншою функцією $f(x)$ такою, щоб числа $P_n^*([\alpha;\beta])=\int\limits_\alpha^\beta f_n^*(x)dx$ і $P([\alpha;\beta])=\int\limits_\alpha^\beta f(x)dx$ якомога менше відрізнялися між собою за будь яких проміжків $[\alpha;\,\beta]$.

Величину розходження (чи близькості) функцій $f_n^*(x)$ і $f(x)$ можна охарактеризувати числом:

$$\chi_{cn}^2=n\sum_{i=1}^{k}\frac{(P_n^*([a_{i-1};a_i))-P([a_{i-1};a_i)))^2}{P([a_{i-1};a_i))}=\sum_{i=1}^{k}\frac{(m_i-np_i)^2}{np_i}=\frac{1}{n}\sum_{i=1}^{k}\frac{(m_i-np_i)^2}{p_i},$$

де $m_i=n\,P_n^*([a_{i-1};a_i))=n\int\limits_{a_{i-1}}^{a_i}f_n^*(x)\,dx$ – абсолютна частота попадання спостережених значень в проміжок $[a_{i-1};a_i)$,

$$p_i = P([a_{i-1}; a_i)) = \int_{a_{i-1}}^{a_i} f(x)\,dx \in [0;1],\ i \in \overline{1,k}\ .$$

Підраховане так значення χ^2_{cn} за результатами серії із n спостережень називають *експериментальним значенням величини* χ^2 (читається: *хі–квадрат спостережене або хі–квадрат експериментальне*).

Разом з тим, якщо справді функція $f_n^*(x)$ практично співпадає з функцією $f(x)$, то тоді за досить великих n на множині значень кожної з величин $P_n^*([a_{i-1}; a_i))$ розподіл статистичних ймовірностей, близький до нормального з параметрами p_i і $\sqrt{\dfrac{p_i q_i}{n}} = \sqrt{\dfrac{p_i(1-p_i)}{n}}$, а на множині значень кожної з величин

$$\frac{P_n^*([a_{i-1}; a_i)) - p_i}{\sqrt{\dfrac{p_i q_i}{n}}}$$

розподіл статистичних ймовірностей близький до нормального з параметрами 0 і 1.

Крім того

$$\chi^2_{cn} = \sum_{i=1}^{k} \frac{(P_n^*([a_{i-1}; a_i)) - p_i)^2}{\dfrac{1}{n} p_i} = \sum_{i=1}^{k} \frac{(P_n^*([a_{i-1}; a_i)) - p_i)^2}{\dfrac{1}{n} p_i q_i} q_i \le$$

$$\le \sum_{i=1}^{k} \left(\frac{P_n^*([a_{i-1}; a_i)) - p_i}{\sqrt{\dfrac{p_i q_i}{n}}} \right)^2 = \chi^2\ .$$

Як відомо, розподіл ймовірностей на множині значень випадкової величини $\chi^2 = \sum_{i=1}^{k} Y_i^2$, де всі Y_i – випадкові величини, на множинах значень яких розподіли статистичних ймовірностей нормальні з параметрами 0 і 1, називають χ^2*– розподілом з* $(k-1)$ *ступенями вільності* (див. §4.17).

Слід підкреслити, що χ^2-розподіл не залежить від того, яка функція $f_n^*(x)$ аналізується і якою функцією $f(x)$

пропонується замінити функцію $f_n^*(x)$, а залежить лише від числа ступенів вільності $(k-1)$ (отже від числа доданків в сумі $\chi^2 = \sum_{i=1}^{k} Y_i^2$, а значить від кількості інтервалів $[a_{i-1}; a_i)$).

Якщо провести досить велику кількість m випробувань, в кожному з яких визначати значення χ^2, то за так одержаних значень можна скласти поінтервальний розподіл статистичних ймовірностей отримання різних значень величини χ^2, та (аналогічно до $f_n^*(x)$) побудувати функцію $f_{m\chi^2}^*(t)$, $t \in (-\infty\,;\infty)$, за якою описуватиметься щільність розподілу статистичних ймовірностей на множині значень випадкової величини χ^2 (χ^2 - розподіл з $(k-1)$ ступенями вільності), чи (аналогічно до $F_n^*(x)$) побудувати функцію $F_{m\chi^2}^*(t)$, $t \in (-\infty\,;\infty)$, розподілу статистичних ймовірностей на множині значень величини χ^2 (див. §2.2-§2.7).

Після цього можна обчислювати ймовірності $P_{m\chi^2}^*([\alpha;\beta))$ попадання значень χ^2 в довільні інтервали $[\alpha;\beta)$, як і раніше:

$$P_{m\chi^2}^*([\alpha;\beta)) = \int_{\alpha}^{\beta} f_{m\chi^2}^*(t)dt = F_{m\chi^2}^*(\beta) - F_{m\chi^2}^*(\alpha),$$

де $f_{m\chi^2}^*(t)$ – щільність розподілу ймовірностей на множині значень випадкової величини χ^2 з $(k-1)$ ступенями вільності, визначена за результатами великої серії із m випробувань, $F_{m\chi^2}^*(t) = \int_{-\infty}^{x} f_{m\chi^2}^*(t)dt$ – відповідна функція розподілу ймовірностей на множині значень випадкової величини χ^2 (див. §2.2-§2.7).

Знаючи $f_{m\chi^2}^*(t)$ (чи $F_{m\chi^2}^*(t)$), можна визначити, за якого значення $\chi^2_{\text{кр}}$ (читається: *хі-квадрат критичне*) ймовірність $P_{n\chi^2}^*((-\infty; \chi^2_{кр}))$ попадання спостережуваних значень величини χ^2 на проміжок $(-\infty; \chi^2_{\text{кр}})$ набуватиме наперед заданого значення γ:

$$P_{m\chi^2}^*((-\infty; \chi^2_{кр})) = P_{m\chi^2}^*(\chi^2 \le \chi^2_{\text{кр}}) = \gamma\,.$$

Якщо γ дорівнює, наприклад, 0.99, а подія, яка відбувається із ймовірністю 0.01, вважається практично неможливою, то таким чином попадання спостережуваного значення χ^2 правіше від значення $\chi^2_{кр}$ слід вважати практично неможливим. Тому, якщо спостережене значення $\chi^2_{сп}$ більше, ніж $\chi^2_{кр}$, такий результат вважається практично неможливим, а гіпотеза про те, що функція $f(x)$ є коректним наближенням функції $f_n^*(x)$, вважається такою, що не узгоджується з експериментальними даними, і відхиляється без великого ризику припуститися помилки, оскільки якщо справді $f_n^*(x)$ практично співпадає з $f(x)$, то в середньому в 99 випадках із 100 спостережуване значення $\chi^2_{сп}$ менше, ніж $\chi^2_{кр}$.

Число γ називають рівнем значущості оцінки $\chi^2_{кр}$. Значення $\chi^2_{кр}$ для різних γ і $r = k - 1$ знаходять за спеціальною таблицею (див. Додаток 5).

Таким чином, знайшовши $\chi^2_{сп}$, задавши рівень значущості $\gamma \in (0;1)$ (досить близький до 1) і визначивши $\chi^2_{кр}$, що відповідає заданому рівню значущості γ, співставляють значення $\chi^2_{сп}$ і $\chi^2_{кр}$. Якщо виявляється, що $\chi^2_{сп} > \chi^2_{кр}$, тоді вважають, що гіпотеза про те, що $f_n^*(x)$ справді практично співпадає з $f(x)$, не узгоджується з експериментальними даними.

Якщо ж $\chi^2_{сп} < \chi^2_{кр}$, то вважають, що гіпотеза про близькість функцій $f_n^*(x)$ і $f(x)$ не суперечить експериментальним даним за заданого рівня значущості γ. Найчастіше γ покладають рівним одному із значень: 0.95, 0.98, 0.99.

Слід зауважити, що, як видно з формули

$$\chi^2_{сп} = n\sum_{i=1}^{k} \frac{(P_n^*([a_{i-1};a_i)) - P([a_{i-1};a_i)))^2}{P([a_{i-1};a_i))},$$

за великих n числа $P_n^*([a_{i-1};a_i))$ і $P([a_{i-1};a_i))$ повинні бути досить близькі між собою для того, щоб виконувалась умова $\chi^2_{сп} < \chi^2_{кр}$. Тому за одних і тих самих значень $P_n^*([a_{i-1};a_i))$, $P([a_{i-1};a_i))$, γ, k та $\chi^2_{кр}$ гіпотеза про коректність заміни функції $f_n^*(x)$ функцією $f(x)$ залежно від n може виявитися як такою, що не узгоджується з експериментальними даними (за великих n), так і такою, що не суперечить тим самим даним (за невеликих n). В цьому є ще один із проявів закону великих чисел.

Описаний критерій перевірки несуперечливості гіпотези про близькість функцій $f_n^*(x)$ і $f(x)$ називають *критерієм Пірсона*.

Приклад 6.5.1. Середній урожай зерна з одного гектара є випадковою величиною. За даними 100 спостережень дістали поінтервальний розподіл статистичних ймовірностей, поданий в Табл. 6.5.3.

Табл. 6.5.3

Урожайність ц/га	13.5–14.5	14.5–15.5	15.5–16.5	16.5–17.5	17.5–18.5	18.5–19.5	19.5–20.5
Статистичні ймовірності	0.06	0.10	0.18	0.28	0.20	0.12	0.06

За допомогою критерію Пірсона перевірити, чи сумісні здобуті статистичні дані з гіпотезою про те, що розподіл імовірностей на множині значень досліджуваної випадкової величини нормальний, щільність якого

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot 1.5} e^{-\frac{(x-17)^2}{2\cdot(1.5)^2}} .$$

Рівень значущості α покласти рівним 0.95.

Скориставшись програмою GRAN1 або таблицею значень функції Лапласа $\Phi(x)$, знайдемо гіпотетичні ймовірності того, що значення досліджуваної випадкової величини попадають в задані часткові інтервали (див. Табл. 6.5.4):

$$P_X([13.5\,;14.5)) = 0.04; \quad P_X([14.5\,;15.5)) = 0.11;$$

$$P_X([15.5\,;1.5)) = 0.21; \quad P_X([16.5\,;17.5)) = 0.28;$$

$$P_X([17.5\,;18.5)) = 0.21; \quad P_X([18.5\,;19.5)) = 0.11;$$

$$P_X([19.5\,;20.5)) = 0.04 .$$

Табл. 6.5.4

Урожайність ц/га	13.3–14.5	14.5–15.5	15.5–16.5	16.5–17.5
Гіпотетичні ймовірності	0.04	0.11	0.21	0.28

17.5–18.5	18.5–19.5	19.5–20.5
0.21	0.11	0.04

Обчислюючи χ^2_{cn} за формулою (6.5.1), дістаємо $\chi^2_{cn} = 2.81$. Число ступенів вільності r в розглядуваному випадку дорівнює 6, оскільки число часткових інтервалів дорівнює 7, а за статистичним матеріалом не оцінюється жоден параметр.

За таблицею критичних точок розподілу ймовірностей на множині значень випадкової величини χ^2 за $\alpha = 0.95$ і $r = 6$ знайдемо $\chi^2_{кр} = 12.6$ (див. додаток 5).

Оскільки $\chi^2_{cn} < \chi^2_{кр}$, то за критерієм Пірсона наведені статистичні дані не суперечать гіпотезі про те, що розподіл імовірностей на множині значень досліджуваної випадкової величини нормальний зі щільністю розподілу ймовірності

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot 1.5} e^{-\frac{(x-17)^2}{2\cdot(1.5)^2}}.$$

Цілком аналогічно за критерієм Пірсона перевіряється і гіпотеза про дискретний поточковий розподіл ймовірностей на дискретній множині значень випадкової величини X.

А саме, будується ряд розподілу статистичних ймовірностей

x_i	x_1	x_2	...	x_k
$P^*_{nX}(\{x_i\})$	$P^*_{nX}(\{x_1\})$	$P^*_{nX}(\{x_2\})$	...	$P^*_{nX}(\{x_k\})$

а також ряд гіпотетичного розподілу ймовірностей

x_i	x_1	x_2	...	x_k
$P_X(\{x_i\})$	$P_X(\{x_1\})$	$P_X(\{x_2\})$	...	$P_X(\{x_k\})$

після чого обчислюється χ^2_{cn} за формулою

$$\chi^2_{cn} = n\sum_{i=1}^{k} \frac{(P^*_{nX}(\{x_i\}) - P_X(\{x_i\}))^2}{P_X(\{x_i\})} = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i}, \qquad (6.5.3)$$

де m_i – абсолютні частоти значень x_i, $p_i = P_X(\{x_i\})$, – гіпотетичні ймовірності значень x_i, $i = 1, 2, ..., k$.

Далі, як і раніше, за даним рівнем значущості визначається значення $\chi^2_{кр}$ і з'ясовується, чи задовільняється нерівність

$\chi_{cn}^2 \geq \chi_{\kappa p}^2$. Якщо нерівність задовільняється, то вважається, що гіпотеза про дискретний поточковий розподіл ймовірностей не узгоджується із статистичними даними. Якщо ж нерівність $\chi_{cn}^2 \geq \chi_{\kappa p}^2$ не задовільняється, то вважається, що гіпотеза про дискретний розподіл ймовірностей не суперечить статистичним даним.

В програмі Gran1 передбачено автоматичну перевірку за критерієм Пірсона гіпотез як про поінтервальний розподіл імовірностей на неперервній обмеженій множині точок виду $\Omega = [a; b)$, так і про поточковий розподіл ймовірностей на скінченній множині точок виду $\Omega_X = \{x_1, x_2, ..., x_k\}$.

Зауважимо, що, як видно із формул (6.5.1), (6.5.3), щоб значення χ_{cn}^2 за досить великих n не було дуже великим, потрібно, щоб ймовірності $P_{nX}^*(I_i)$ та $P_X(I_i)$ в поінтервальному розподілі ймовірностей чи ймовірності $P_{nX}^*(\{x_i\})$ та $P_X(\{x_i\})$ в поточковому розподілі ймовірностей були досить близькі між собою, тобто щоб розподіли статистичних та гіпотетичних ймовірностей справді були близькими.

В цьому полягає ще один прояв закону великих чисел.

Серед інших критеріїв узгодження розглянемо ще *критерій Колмогорова*. Відхилення статистичного розподілу ймовірностей від гіпотетичного цього разу обчислюють як

$$U = D = \max_x |F_{nX}^*(x) - F_X(x)|,$$

де $F_{nX}^*(x)$ – функція розподілу статистичних імовірностей, $F_X(x)$ – гіпотетична функція розподілу ймовірностей. Тут мається на увазі $\Omega_X = R^1$, $S_X = \mathscr{B}(R^1)$.

Табл. 6.5.5

λ	$P(\lambda)$	λ	$P(\lambda)$
0.0	1.000		
0.1	1.000	1.1	0.178
0.2	1.000	1.2	0.112
0.3	1.000	1.3	0.068
0.4	0.997	1.4	0.040
0.5	0.964	1.5	0.022
0.6	0.864	1.6	0.012
0.7	0.711	1.7	0.006
0.8	0.544	1.8	0.003
0.9	0.393	1.9	0.002
1.0	0.270	2.0	0.001

За спеціальною таблицею (Табл. 6.5.5) за різних λ можна знайти ймовірність $P(D\sqrt{n} \geq \lambda) = \beta$. Якщо ця ймовірність дуже мала, то подія $(D\sqrt{n} \geq \lambda)$ практично неможлива. За заданого рівня значущості $\alpha = 1 - \beta$ (досить близького до одиниці) за таблицею Табл. 6.5.5 можна знайти таке $\lambda_{кр}$, за якого

$$P(D\sqrt{n} \geq \lambda_{кр}) = \beta .$$

Якщо за статистичними даними $D_{cn}\sqrt{n} \geq \lambda_{кр}$, то $P(D\sqrt{n} \geq D_{cn}\sqrt{n}) \leq \alpha$, і таким чином гіпотезу про те, що функцією розподілу ймовірностей є $F_X(x)$, слід відхилити як таку, що не узгоджується із статистичними даними за заданого рівня значущості.

Якщо виявиться $D_{cn}\sqrt{n} < \lambda_{кр}$, то гіпотеза про розподіл ймовірностей за заданого рівня значущості не суперечить статистичним даним.

Є й інші методи статистичної перевірки гіпотез про розподіл імовірностей або про значення параметрів розподілу, а також інші критерії узгодження чи суперечливості гіпотез із статистичними даними.

Приклад 6.5.2. Між нулем і одиницею навмання вибирають число. В результаті 100 спостережень дістали поінтервальний розподіл статистичних ймовірностей стосовно спостережених випадкових значень (Табл. 6.5.6).

Табл. 6.5.6

Часткові інтервали	[0–0.1)	[0.1–0.2)	[0.2–0.03)	[0.3–0.4)	[0.4–0.5)	[0.5–0.6)	[0.6–0.7)	[0.7–0.8)	[0.8–0.9)	[0.9–1.0)
Статистичні ймовірності	0.08	0.12	0.07	0.10	0.15	0.09	0.06	0.13	0.09	0.11

Використовуючи критерій Колмогорова, перевірити, чи узгоджується із здобутими статистичними даними гіпотеза про те, що на множині значень досліджуваної випадкової величини розподіл імовірностей рівномірний на проміжку $[0\,;1]$. Рівень значущості α покласти рівним 0.9.

За статистичними даними знайдемо (скориставшись програмою GRAN1) значення функції розподілу статистичних ймовірностей на кінцях заданих інтервалів I_i (Рис. 6.5.3, *а*, *б*). Одержуємо

$F_X(0) = 0$; $F_n^*(0.1) = 0.08$; $F_n^*(0.2) = 0.20$; $F_n^*(0.3) = 0.27$;
$F_n^*(0.4) = 0.37$; $F_n^*(0.5) = 0.52$; $F_n^*(0.6) = 0.61$; $F_n^*(0.7) = 0.67$;
$F_n^*(0.8) = 0.80$; $F_n^*(0.9) = 0.89$; $F_n^*(1.0) = 1$.

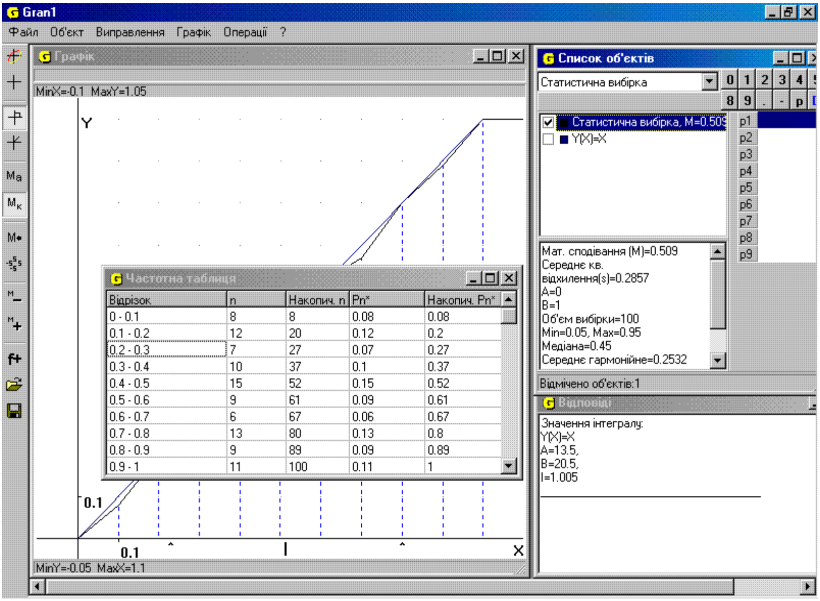


Рис. 6.5.3, а

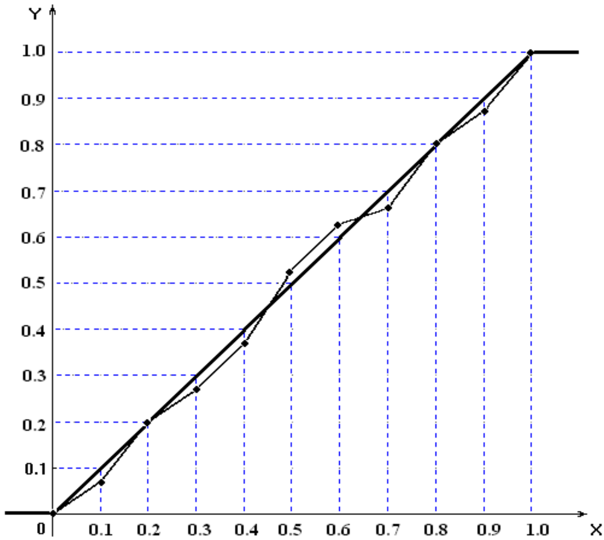


Рис. 6.5.3, б

На Рис. 6.5.3, б зображено графіки функцій $F_n^*(x)$ і $F_X(x)$ на відрізку $[0\,;1]$. Значення гіпотетичної функції розподілу

$$F_X(x) = \begin{cases} 0, & \text{коли } x \le 0; \\ x, & \text{коли } 0 \le x \le 1; \\ 1, & \text{коли } 1 \le x; \end{cases}$$

на кінцях тих самих інтервалів I_i такі:

$F_X(0) = 0$; $F_n^*(0.1) = 0.1$; $F_n^*(0.2) = 0.2$; $F_n^*(0.3) = 0.3$;

$F_n^*(0.4) = 0.4$; $F_n^*(0.5) = 0.5$; $F_n^*(0.6) = 0.6$; $F_n^*(0.7) = 0.7$;

$$F_n^*(0.8) = 0.8\,;\; F_n^*(0.9) = 0.9\,;\; F_n^*(1.0) = 1.0\,.$$

Знайдемо $D = \max\limits_{x \in [0,1]} |F_n^*(x) - F_X(x)|$. Зазначимо, що коли інтервали I_i досить малі, то замість $\max\limits_{x \in [0,1]} |F_n^*(x) - F_X(x)|$ можна наближено взяти найбільше значення різниці $|F_n^*(x) - F_X(x)|$ на кінцях інтервалів I_i. У розглядуваному прикладі

$$F_n^*(0) - F_X(0) = 0\,;\; F_n^*(0.1) - F_X(0.1) = -0.02\,;$$

$$F_n^*(0.2) - F_X(0.2) = 0.00\,;$$

$$F_n^*(0.3) - F_X(0.3) = -0.03\,;\; F_n^*(0.4) - F_X(0.4) = -0.03\,;$$

$$F_n^*(0.5) - F_X(0.5) = 0.02;\ F_n^*(0.6) - F_X(0.6) = 0.01;$$
$$F_n^*(0.7) - F_X(0.7) = -0.03;\ F_n^*(0.8) - F_X(0.8) = 0.00;$$
$$F_n^*(0.9) - F_X(0.9) = -0.01;\ F_n^*(1.0) - F_X(1.0) = 0$$

За всіх інших значень аргументу $x \in [0;1]$ значення різниці функцій $F_n^*(x)$ і $F_X(x)$ не перевищують знайдених. Тому

$$D_{cn} = \max_{x \in [0,1]} |F_n^*(x) - F_X(x)| = 0.03.$$

Таким чином, $D_{cn}\sqrt{n} = 0.03 \cdot \sqrt{100} = 0.3$. Оскільки стосовно заданого рівня значущості $\alpha = 0.9$, ($\beta = 1 - \alpha = 1 - 0.9 = 0.1$), за Табл. 6.5.5 $\lambda_{кр} \approx 1.2$ і $D_{cn}\sqrt{n} = 0.3 < 1.2$, то за критерієм Колмогорова за заданого рівня значущості $\alpha = 0.9$ гіпотеза про те, що на множині значень досліджуваної випадкової величини розподіл імовірностей рівномірний на відрізку $[0;1]$, не суперечить наведеним статистичним даним.

Зазначимо, що коли б число спостережень було набагато більшим, наприклад $n = 10000$, і дістали б той самий поінтервальний розподіл статистичних ймовірностей, то отримали б $D_{cn}\sqrt{n} = 0.03 \cdot \sqrt{10000} = 3$. У цьому разі, згідно з критерієм Колмогорова, гіпотеза про рівномірний розподіл імовірностей не узгоджувалася б із статистичними даними.

На відміну від критерію К. Пірсона, в разі застосування якого за статистичними даними можуть бути оцінені кілька невідомих параметрів гіпотетичного розподілу ймовірностей, через що відповідним чином зменшується число ступенів вільності r, в разі застосування критерію Колмогорова гіпотетичний розподіл імовірностей повинен бути повністю визначений разом з усіма його параметрами.

Крім розглянутих вище критеріїв перевірки гіпотез про розподіл імовірностей на множині значень випадкової величини X розглянемо ще деякі підходи до перевірки гіпотез про числові характеристики розподілів.

Нехай є дві незалежні випадкові величини X і Y з нормальними розподілами ймовірностей на множинах їх значень, в результаті спостережень за якими дістали значення $x_1, x_2, ..., x_n$ і $y_1, y_2, ..., y_n$. Треба перевірити, чи суперечить статистичним даним гіпотеза про те, що дисперсії $\sigma^2[X]$ і $\sigma^2[Y]$ рівні між собою (математичні сподівання тут невідомі).

Розглянемо відношення

$$F_{n,m}=\frac{\tilde{D}_n[X]}{\tilde{D}_m[Y]}=\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2}{\frac{1}{m-1}\sum_{k=1}^{m}(Y_k-\overline{Y}_m)^2}.$$

Ця величина не залежить від невідомих параметрів нормальних розподілів на множинах значень випадкових величин X_i і Y_k. Розподіл імовірностей на множині значень випадкової величини

$$F_{n_1,n_2}=\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}X_i^2}{\frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2}X_i^2},$$

де випадкові величини X_i незалежні з нормальними розподілами ймовірностей на множинах їх значень з параметрами $M[X_i]=0$, $\sigma[X_i]=1$, називають *розподілом Фішера* з n_1 і n_2 ступенями вільності.

Якщо $M[X_i]=m$, то на множині значень випадкової величини

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\overset{\circ}{X}_i{}^2}{\frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2}\overset{\circ}{X}_i{}^2},\text{ де } \overset{\circ}{X}_i=X_i-m,$$

має місце розподіл Фішера.

Значення ймовірностей $P(F_{n,m}\geq F_{n_1,n_2 кр})=\alpha$ можна визначити за спеціальними таблицями (див. додаток 7). Якщо дисперсії $D[X]$ і $D[Y]$ рівні між собою, то відношення їх оцінок повинно бути близьким до 1. Тому вважають, що гіпотеза не узгоджується із статистичними даними, коли

$$\frac{\tilde{D}_n[X]}{\tilde{D}_m[Y]}<c_1 \text{ або } \frac{\tilde{D}_n[X]}{\tilde{D}_m[Y]}>c_2,\text{ де } c_1<1,\ c_2>1.$$

Числа c_1 і c_2 визначають за таблицями розподілу ймовірностей Фішера, враховуючи рівень значущості α.

Якщо математичні сподівання $M[X]$ і $M[Y]$ відомі, то гіпотеза про рівність дисперсій перевіряється аналогічно до попереднього, вважаючи, що

$$\tilde{D}_n[X]=\frac{1}{n}\sum_{i=1}^{n}(X_i-M[X])^2\ ,\ \tilde{D}_m[Y]=\frac{1}{m}\sum_{i=1}^{m}(Y_i-M[Y])^2\ .$$

Розглянемо гіпотезу про рівність математичних сподівань $M[X]$ і $M[Y]$, якщо відомо, що розподіли ймовірностей на множинах значень випадкових величин X і Y нормальні з однаковими дисперсіями, проте параметри розподілів невідомі.

Очевидно, що

$$D[\overline{X}_n-\overline{Y}_n]=\frac{\sigma^2[X]}{n}+\frac{\sigma^2[Y]}{m}=\sigma^2\ \frac{m+n}{mn},$$

$$\sqrt{D[\overline{X}_n-\overline{Y}_n]}=\sigma\sqrt{\frac{m+n}{mn}}\ .$$

Отже, на множині значень випадкової величини

$$\frac{\overline{X}_n-\overline{Y}_m}{\sqrt{D[\overline{X}_n-\overline{Y}_n]}}=\frac{\overline{X}_n-\overline{Y}_m}{\sigma}\sqrt{\frac{mn}{m+n}}$$

розподіл імовірностей нормальний з параметрами $m=0$, $\sigma=1$. Оскільки

$$\tilde{D}_n[X]=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2\ ,\ \tilde{D}_m[Y]=\frac{1}{m-1}\sum_{i=1}^{m}(Y_i-\overline{Y}_m)^2\ ,$$

а

$$(n-1)\frac{\tilde{D}_n[X]}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}_n}{\sigma}\right)^2,\ (m-1)\frac{\tilde{D}_m[Y]}{\sigma^2}=\sum_{i=1}^{m}\left(\frac{Y_i-\overline{Y}_n}{\sigma}\right)^2$$

випадкові величини з χ^2-розподілами ймовірностей на множинах їх значень з $n-1$ і $m-1$ ступенями вільності, то

$$(n-1)\frac{\tilde{D}_n[X]}{\sigma^2}+(m-1)\frac{\tilde{D}_m[Y]}{\sigma^2}$$

випадкова величина з χ^2-розподілом ймовірностей на множині її значень з $n+m-2$ ступенями вільності.

Зрозуміло, що

$$M\left[\frac{n-1}{n+m-2}\overline{D}_n[X]+\frac{m-1}{n+m-2}\overline{D}_n[Y]\right]=$$

$$=\frac{n-1}{n+m-2}\sigma^2+\frac{m-1}{n+m-2}\sigma^2=\sigma^2.$$

Отже

$$\frac{1}{n+m-2}[\tilde{D}_n[X](n-1)+\tilde{D}_m[Y](m-1)]$$

є незміщеною оцінкою стосовно σ^2. Таким чином, величина

$$T=\frac{\bar{X}_n-\bar{Y}_m}{\sqrt{((n-1)\tilde{D}_n[X]+(m-1)\tilde{D}_m[Y])/(n+m-2)}}\sqrt{\frac{mn}{m+n}}$$

не залежить від невідомих $M[X]$, $M[Y]$ та σ^2 і на множині її значень має місце розподіл імовірностей Стьюдента з $n+m-2$ ступенями вільності.

Отже, якщо $|T|>t_{кр}(n+m-2)$, де $t_{кр}(n+m-2)$ – критичне значення за розподілом Стьюдента за заданого рівня значущості, то гіпотеза про рівність математичних сподівань не узгоджується із статистичними даними і повинна бути відхилена.

Якщо дисперсії $\sigma^2[X]$ і $\sigma^2[Y]$ відомі, то оцінка

$$\frac{\bar{X}_n-\bar{Y}_m}{\sqrt{D[\bar{X}_n]-D[\bar{Y}_m]}}$$

замінюється на

$$\frac{\bar{X}_n-\bar{Y}_m}{\sqrt{\frac{1}{n}\sigma^2[X]+\frac{1}{m}\sigma^2[Y]}}.$$

Розподіл імовірностей на множині значень останньої випадкової величини нормальний. Тому гіпотеза $M[X]=M[Y]$ відхиляється, якщо за рівня значущості α

$$\frac{\bar{X}_n-\bar{Y}_m}{\sqrt{\frac{1}{n}\sigma^2[X]+\frac{1}{m}\sigma^2[Y]}}>t_{кр},$$

де $t_{кр}$ – критичне значення стосовно нормального розподілу ймовірностей з параметрами $m=0$, $\sigma=1$, тобто

$$P(|T|\geq t_{кр})=1-\alpha.$$

Наведені критерії перевірки гіпотез про рівність математичних сподівань і дисперсій є точними і можуть бути використані як в разі малих, так і в разі великих об'ємів вибірки (якщо розподіли імовірностей на множинах значень досліджуваних випадкових величин нормальні). Коли об'єм вибірки великий, тоді можна вважати, що $\bar{X}_n$ і $\bar{Y}_m$ – випадкові величини з нормальними розподілами ймовірностей, тому розглянуті критерії перевірки гіпотез можуть бути використані й тоді, коли розподіли ймовірностей на множинах значень досліджуваних випадкових величин відмінні від нормального.

Є методи для перевірки й інших типів гіпотез, зокрема для багатовимірних розподілів ймовірностей, некількісних (якісних) ознак досліджуваних явищ тощо.

Вправи для самостійного виконання

6.5.1. Перевірити, чи правильні твердження:

1. За кожною вибіркою значень випадкової величини X можна сформулювати гіпотезу стосовно близькості розподілу ймовірностей (чи окремих параметрів розподілу) на множині значень цієї величини до якогось стандартного розподілу (чи окремих параметрів цього розподілу).

2. Гіпотеза з твердження 1 завжди повинна бути пов'язана з певною статистичною оцінкою (нормою, відхиленням) U.

3. Відхилення з твердження 2 є випадковою величиною з відомим розподілом ймовірностей.

4. Якщо розподіл ймовірностей на множині значень відхилення U відомий, то стосовно нього можна знайти критичну область.

5. Критерій Пірсона ґрунтується на тому, що:

а) відхилення U статистичного розподілу від гіпотетичного визначається за рівністю

$$U = \chi^2 = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i};$$

б) відхилення U завжди є випадковою величиною з χ^2-розподілом ймовірностей на множині її значень;

в) χ^2-розподіл залежить від гіпотетичного розподілу;

г) χ^2-розподіл не залежить від числа ступенів вільності;

д) стосовно χ^2-розподілу існують таблиці значень, за якими, знаючи r і χ^2_{cn}, можна знайти $\beta = P(\chi^2 \geq \chi^2_{cn})$, а знаючи r і β, можна знайти $\chi^2_{кр}$ з рівняння $P(\chi^2 < \chi^2_{kp}) = \alpha$;

е) якщо $\chi^2_{cn} \geq \chi^2_{kp}$, то гіпотеза про розподіл ймовірностей на множині значень випадкової величини X не узгоджується із наявними статистичними даними. В іншому разі ця гіпотеза не суперечить наявним статистичним даним.

6. Критерій Пірсона застосовний лише для неперервних випадкових величин.

7. Критерій Колмогорова ґрунтується на тому, що:

а) відхилення функції розподілу статистичних ймовірностей $F^*_{nX}(x)$ від гіпотетичної функції розподілу ймовірностей $F_X(x)$ визначається за рівністю:

$$U = D = \sup_x | F_{nX}^*(x) - F_X(x) |;$$

б) існує таблиця, за якою, знаючи рівень значущості α (близький до одиниці), можна знайти λ_{kp} таке, що $P(D\sqrt{n} \geq \lambda_{kp}) = 1 - \alpha$;

в) якщо за статистичними даними $D_{cn}\sqrt{n} \leq \lambda_{kp}$, то гіпотезу про те, що функцією розподілу ймовірностей на множині значень випадкової величини X є $F_X(x)$, відхиляють, а в іншому разі вважають, що ця гіпотеза не суперечить статистичним даним.

8. Для перевірки гіпотези про рівність дисперсій розподілів ймовірностей на множинах значень двох випадкових величин X та Y, коли такі розподіли нормальні (з невідомими математичними сподіваннями та дисперсіями), можна використати розподіл Фішера.

9. Для перевірки гіпотези про рівність математичних сподівань випадкових величин X та Y з нормальними розподілами ймовірностей на множинах їх значень (з невідомими параметрами, проте однаковими дисперсіями) можна використовувати розподіл Стьюдента на множині значень випадкової величини

$$T = \frac{(\overline{X}_n - \overline{Y}_m)\sqrt{mn/(m+n)}}{\sqrt{((n-1)\tilde{D}_n[x] + (m-1)\tilde{D}_m[y])/(n+m-2)}}.$$

10. Величина χ^2 залежить від порівнюваних функцій $f_n^*(x)$ і $f(x)$.

11. Величина χ^2 залежить від рівня значущості – числа γ і числа ступенів вільності r.

12. Гіпотеза про близькість функцій $f_n^*(x)$ і $f(x)$ не узгоджується з експериментальними даними, якщо $\chi^2_{\text{експ}} > \chi^2_{\text{кр}}$.

13. Критерій Пірсона можна застосовувати для перевірки гіпотез про дискретні розподіли статистичних ймовірностей.

6.5.2. Скориставшись програмою Gran1, перевірити за критерієм Пірсона гіпотезу про близькість функцій $f_n^*(x)$ і $f(x)$, якщо:

1. $f(x) = \begin{cases} 1, \text{ коли } x \in [0;1], \\ 0, \text{ коли } x \overline{\in} [0;1], \end{cases}$ $\gamma = 0.95$,

а $f_n^*(x)$ задана за таблицею

$[a_{i-1}; a_i)$	$[0; 0.2)$	$[0.2; 0.4)$	$[0.4; 0.6)$	$[0.6; 0.8)$	$[0.8; 1.0)$
$f_n^*(x)$	0.90	1.00	1.10	0.85	1.15

2. $f(x)=\begin{cases}1-|x|, & \text{коли} \quad x\in[-1;1], \\ 0, & \text{коли} \quad x\overline{\in}[-1;1],\end{cases}$ $\gamma = 0.95$,

а $f_n^*(x)$ задана за таблицею

$[a_{i-1};a_i)$	$[-1;-0.6)$	$[-0.6;\ -0.2)$	$[-0.2;\ 0.2)$	$[0.2;0.6)$	$[0.6;\ 1)$
$f_n^*(x)$	0.2	0.6	0.9	0.6	0.2

6.6. Поняття про метод статистичних випробувань (метод Монте-Карло)

Досить часто виникають задачі про визначення характеристик деяких процесів, перебіг яких не є детермінованим і залежить від випадкових факторів. Для таких задач розробляють спеціальні *імітаційні моделі* і імітуючи (як правило з використанням комп'ютера) більш чи менш точно перебіг реальних процесів, виконують спостереження за перебігом імітованих процесів. *Усереднені результати спостережень використовують для наближеного визначення шуканих характеристик досліджуваних процесів*. У цьому й полягає суть *методу статистичних випробувань*, який називають також *методом Монте-Карло.* Імітацію перебігу недетермінованих реальних процесів за допомогою імітаційних моделей називають *статистичним моделюванням*.

У деяких випадках метод Монте-Карло є єдиним, за яким можна дістати наближені розв'язки задач, які не можна проаналізувати за іншими аналітичними чи чисельними методами. Якщо, наприклад, треба спроектувати мережу зв'язку (Рис. 6.6.1) так, щоб відсоток відмов запитам на обслуговування не перевищував заданого, коефіцієнт використання каналів зв'язку був не меншим заданого, час чекання в черзі на обслуговування не перевищував заданого тощо, коли відомі розподіли ймовірностей на множинах випадкових моментів часу, в які надходять запити на обслуговування у кінцеві пункти, розподіли ймовірностей на множинах довжин випадкових проміжків часу обслуговування запитів, то метод Монте-Карло цілком можна застосувати для розв'язування такої задачі. Метод Монте-Карло можна використати і в разі розв'язування багатьох задач, які детерміновані і стосовно розв'язування яких відомі інші чисельні або аналітичні методи.

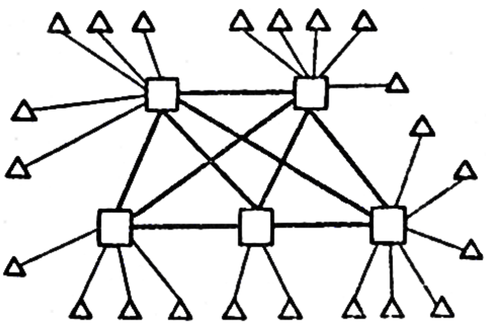


Рис. 6.6.1

Приклад 6.6.1. Нехай потрібно обчислити площу деякої досить складної плоскої фігури G (Рис. 6.6.2,а). Вибравши відповідний масштаб, побудуємо квадрат

$$W = \{(x, y) \mid 0 \le x \le 1;\ 0 \le y \le 1\},$$

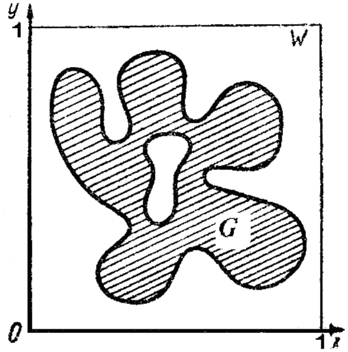


Рис. 6.6.2, а)

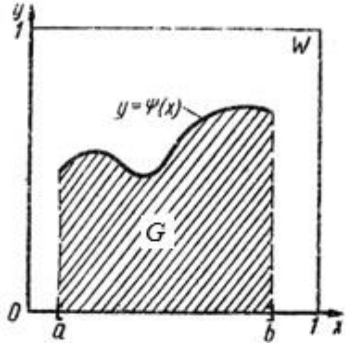


Рис. 6.6.2, б)

яким повністю охоплюється розглядувана фігура. У множині W навмання вибиратимемо точки (x, y) так, щоб будь-які з них були *рівноможливі*, тобто щоб розподіл статистичних ймовірностей на множині W за досить великих n був *рівномірним*.

Проведемо досить велику серію випробувань, щоразу навмання вибираючи випадкову точку (x, y) з множини W, і знайдемо статистичну ймовірність $P^*_{n\,(X,Y)}(G)$ попадання точки (x, y) в область G, що дорівнює відношенню числа випробувань, в яких точка (x, y) виявилась в області G, до числа всіх випробувань.

Внаслідок статистичних випробувань неможливо перебрати всі точки множини W, проте за досить великої вибірки можна скласти майже точне уявлення про те, яку частину відносно міри (площі) множини W становить міра (площа) підмножини G. Така частина за досить великих n практично дорівнюватиме

$$P^*_{n\,(X,Y)}(G) = \frac{k_n(G)}{k_n(W)} \approx \frac{m(G)}{m(W)}, \qquad (6.6.1)$$

де $k_n(G)$ – кількість точок (x, y), що попали в множину G, $k_n(W)$ – кількість всіх точок (x, y), навмання вибраних у множині W. Тому наближено можна вважати

$$m(G) = m(W) \cdot P^*_{n\,(X,Y)}(G).$$

Розглянутий метод наближеного обчислення площі фігури можна застосовувати до наближеного обчислення визначеного інтеграла $\int_a^b \psi(x)dx$ (Рис. 6.6.2, б).

Цей самий інтеграл можна обчислити за методом Монте-Карло інакше. Розглянемо випадкову величину $Y=\psi(X)$, де аргумент X – випадкова величина з рівномірним розподілом статистичних імовірностей на проміжку $[a;b]$, щільність якого можна подати у вигляді (див. §2.2, §2.7, §4.12)

$$f_{nX}^*(x)=\begin{cases}\dfrac{1}{b-a}, & \text{коли } x\in[a;b],\\ 0, & \text{коли } x\overline{\in}[a;b],\end{cases}$$

Тоді за досить великих n з достатньою точністю має місце рівність

$$M_n^*[Y]=\int_{-\infty}^{\infty}\psi(x)f_{nX}^*(x)dx=\int_a^b\psi(x)\frac{1}{b-a}dx=\frac{1}{b-a}\int_a^b\psi(x)dx.$$

Звідси

$$\int_a^b\psi(x)dx=(b-a)M_n^*[Y].$$

Остання формула є однією із форм теореми про середнє значення функції $\psi(x)$ (див. §4.12, формула 4.12.1).

Оскільки наближене значення інтеграла $\int_a^b\psi(x)dx$ можна знайти також за формулою прямокутників $\int_a^b\psi(x)dx\approx\sum_{i=1}^{n}\psi(\tilde{x}^{(i)})h$, де $h=\frac{b-a}{n}$, $x^{(i+1)}=x^{(i)}+h$, $i=0,1,2,...,n$; $x_0=a$, $x_0+nh=b$, $\tilde{x}^{(i)}$ – деяка точка з відрізка $[x^{(i)};x^{(i+1)}]=[x^{(i)};x^{(i)}+h]$, то наближено значення математичного сподівання $M[Y]$ випадкової величини $Y=\psi(X)$ з рівномірним розподілом статистичних ймовірностей на множині значень випадкової величини X (на відрізку $[a;b)$), з достатньою точністю за досить великих n можна знайти за формулою $M[Y]=\frac{1}{b-a}\int_a^b\psi(x)dx\approx\frac{1}{n}\sum_{i=1}^{n}\psi(\tilde{x}_i)$.

Наближене значення $M[Y]$ можна дістати і так: виконати серію n випробувань, в результаті кожного з яких спочатку вибрати навмання аргумент $x_{i\,\text{сп}}\in[a;b)$, а потім дістати спостережене значення $y_{i\,\text{сп}}=\psi(x_{i\,\text{сп}})$ випадкової величини Y, де $x_{i\ \text{сп}}$, $i \in \overline{1,n}$, – спостережені значення випадкової величини X з рівномірним розподілом імовірностей на відрізку $[a\,;b)$, за якими здобуто значення $y_{i\ \text{сп}} = \psi(x_{i\ \text{сп}})$, $i \in \overline{1,n}$.

Тоді $M[Y] \approx \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} y_{i\,\text{сп}}$, і оскільки

$$M[Y] \approx \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} y_{i\ \text{сп}} = \frac{1}{n}\sum_{i=1}^{n} \psi(x_{i\ \text{сп}}),$$

то (див. формулу 4.12.1)

$$\int_a^b \psi(x)dx = M[Y]\cdot m([a,b)) \approx \frac{b-a}{n}\sum_{i=1}^{n}\psi(x_{i\ \text{сп}})$$

(порівняйте з формулою прямокутників для наближеного обчислення визначеного інтеграла $\int_a^b \psi(x)dx \approx \sum_{i=1}^{n}\psi(\tilde{x}^{(i)})h$, де $h = \frac{b-a}{n}$).

Розглянутий метод можна майже без змін перенести на обчислення визначеного інтеграла будь-якої кратності

$$\int_G \psi(x)dx = \frac{m(G)}{n}\sum_{i=1}^{n}\psi(x_{i\ \text{сп}}),$$

де $x = (x_1, x_2, ..., x_n) \in R^n$, $dx = dx_1 dx_2 ... dx_n$,

$$\int_G \psi(x)dx = \iint_G ... \int \psi(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n,$$

$m(G)$ – міра множини G.

Приклад 6.6.2. (задача Бюффона).

На дощату підлогу навмання кидається голка завдовжки l. Потрібно знайти узагальнену статистичну ймовірність того, що голка падає на щілину між двома сусідніми дошками, якщо ширина дошки t, а розподіл узагальнених статистичних ймовірностей на множині пар

$$W = \{(x,\varphi) \mid -\frac{t}{2} \le x \le \frac{t}{2};\ 0 \le \varphi \le \pi\},$$

за результатами досить великої серії з n випробувань виявився з достатньою точністю практично рівномірним, де x відхилення центра голки від найближчої щілини (прямої MN), φ – кут, утворений голкою з цією щілиною (Рис. 6.6.3, а). Кут відлічується від додатного (зліва направо) напрямку прямої MN проти годинникової стрілки. Тоді положення голки відносно прямої MN можна охарактеризувати парою чисел (x,φ).

Отже, є двохвимірна випадкова величина (X,Φ) з рівномірним розподілом ймовірностей на множині W її значень. Голка перетне щілину (пряму MN), якщо $|x| < \frac{l}{2}\sin\varphi$, тобто $-\frac{l}{2}\sin\varphi \le x \le \frac{l}{2}\sin\varphi$ (див. Рис. 6.6.3, а).

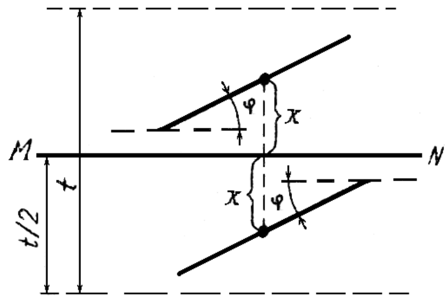


Рис. 6.6.3, а

Нехай G – подія, яка полягає в тому, що голка перетне пряму MN. Множина W всіх можливих пар (x,φ) та множина пар (x,φ), що належать до множини G, зображені на Рис. 6.6.3, *б*.

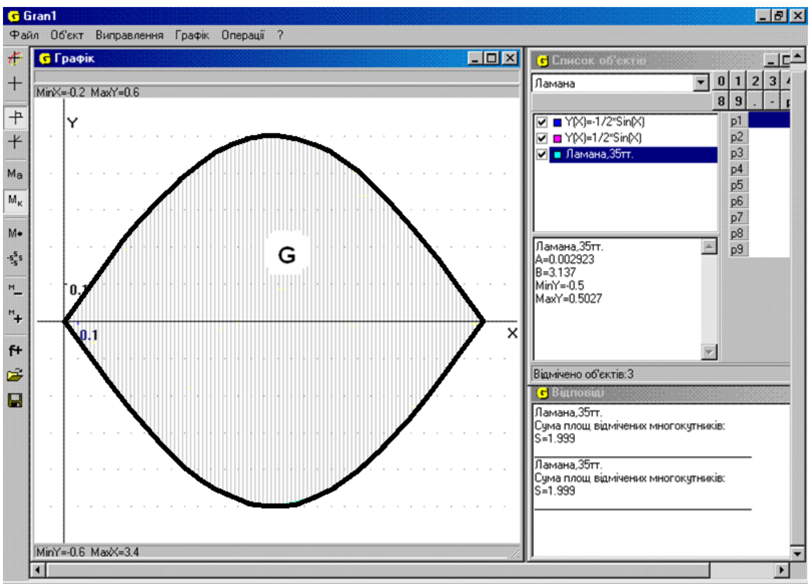


Рис. 6.6.3, б

Площа фігури G дорівнює

$$2\int_0^{\pi}\frac{l}{2}\sin\varphi d\varphi = -l\cos\varphi\,|_0^{\pi} = 2l\ ,$$

а площа фігури W дорівнює $t\pi$. Тоді, враховуючи що

$$P^*_{n\,(X,\Phi)}(G) = \frac{k_n(G)}{k_n(W)} \approx \frac{m(G)}{m(W)}$$

(див. формулу 6.6.1), дістаємо $P^*_{n\,(X,\Phi)}(G) \approx \frac{2l}{t\pi}$.

Якщо провести досить велику серію випробувань, у кожному з яких фіксувати, перетинає чи ні голка щілину *MN*, то відношення числа m випробувань, в яких голка перетинає щілину, до числа n всіх випробувань, є статистична ймовірність події G, яка за досить великих n мало відрізняється від числа $\frac{2l}{t\pi}$, і тому за досить великого числа випробувань наближено можна вважати $\frac{m}{n} = \frac{2l}{t\pi}$. З останньої рівності зокрема можна одержати наближене значення числа $\pi = \frac{2nl}{mt}$.

Отже, використовуючи метод Монте-Карло, можна наближено знайти число π. Надійність оцінки $\frac{m}{n}$ числа $\frac{2l}{t\pi}$ за заданих надійних меж і, навпаки, надійні межі за заданої надійності оцінки можна визначити, виходячи з міркувань, що стосуються закону великих чисел.

Слід зазначити, що на практиці метод Монте-Карло застосовують тоді, коли не вимагається висока точність результату, оскільки в протилежному разі доведеться виконати занадто великі серії випробувань, що іноді здійснити досить важко. Відомо, наприклад, що для того щоб збільшити точність в 10 разів, тобто дістати ще одну правильну цифру в результаті, треба збільшити кількість випробувань приблизно в 100 разів (це випливає з теорем про числові характеристики функцій випадкових аргументів). Зрозуміло, що дуже великої точності досягти нелегко навіть за допомогою комп'ютера. Проте багато практично важливих задач можна розв'язати з достатньою точністю, застосовуючи метод статистичних випробувань.

Розглянемо задачу про визначення довжини кривої, зображеної на Рис. 6.6.4, *а.* За звичайними методами знайти довжину цієї кривої досить важко. Скористаємося методом Монте-Карло і задачею Бюффона. Якщо уявити, що пофарбована голка падала на підлогу багато разів, залишаючи щоразу на підлозі слід, то зрештою ці сліди зіллються в деяку лінію досить складної конфігурації. Якщо кидати голку n разів і вона перетне щілину m разів, то можна вважати $\frac{m}{n}=\frac{2l}{t\pi}$. Крім того, довжина L лінії, що утворилась із слідів голки в n випробуваннях, дорівнює nl.

Звідси

$$L=\frac{mt\pi}{2} \tag{6.6.2}$$

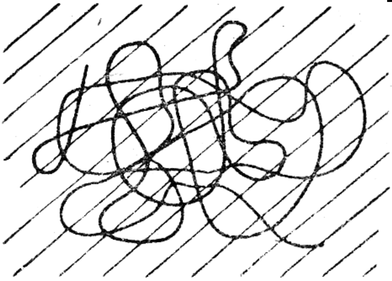


Рис. 6.6.4, а

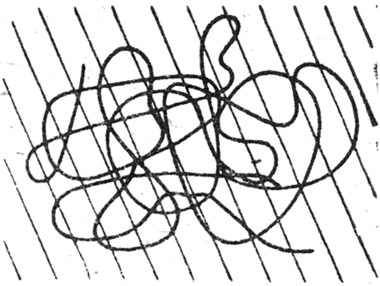


Рис. 6.6.4, б

Формулу 6.6.2 використовують так. На задану лінію накладають розлінований прозорий папір з нанесеними на нього паралельними лініями, відстань між якими t, і підраховують кількість перетинів із заданою лінією. За формулою (6.6.2) знаходять величину L. Різним чином розміщуючи набір паралельних ліній (Рис. 6.6.4, б), кілька разів визначають довжину L, а за остаточний результат беруть середнє арифметичне із здобутих щоразу значень.

Застосовуючи метод статистичних випробувань на практиці, треба мати можливість діставати випадкову точку M із заданим розподілом імовірностей на деякій множині W. Цей розподіл може бути дискретним чи неперервним залежно від того, який реальний процес вивчається.

Засіб, за допомогою якого дістають значення випадкової величини X із заданим одновимірним розподілом імовірностей, називають *генератором* (або датчиком) випадкових чисел. В результаті кожного звернення до такого датчика дістають деяке значення випадкової величини X. Кожне звернення до датчика можна тлумачити як проведення випробування, а результати всіх цих випробувань – як статистичний матеріал. Датчики випадкових чисел можуть бути реалізовані у вигляді деяких фізичних приладів, а також у вигляді деяких алгоритмів для знаходження значень випадкової величини, що вивчається.

Нехай задано випадкову величину X з дискретним розподілом імовірностей на множині її значень (Табл. 6.6.1).

Табл. 6.6.1

x_i	x_1	x_2	...	x_n
$P_X(\{x_i\})$	p_1	p_2	$\dots$	p_n

Один з найпростіших датчиків випадкових чисел з цим розподілом імовірностей реалізовують так. Вибирають досить велике число N таке, щоб числа Np_i, $i = 1, 2, ..., n$, були цілими і на $m_1 = Np_1$ картках пишуть число x_1, на $m_2 = Np_2$ картках пишуть число x_2 і т.д., на $m_n = Np_n$ картках пишуть число x_n. Потім змішують усі картки і навмання вибирають будь-яку з них, повертаючи щоразу вийняту картку. Якщо елементарні події, відповідні кожній окремій картці, рівноймовірні між собою, то число x_1 вибирається з імовірністю p_1, x_2 – з імовірністю p_2 і т. д., x_n – з імовірністю p_n. Таким чином дістають деякий датчик значень випадкової величини із заданим дискретним розподілом ймовірностей (Табл. 6.6.1).

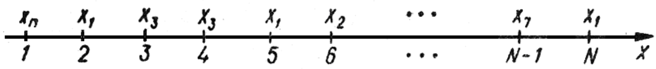


Рис. 6.6.5

Цей самий датчик можна реалізувати і так. Вибирають N точок на осі Ox (Рис. 6.6.5) і будь-які $m_1 = Np_1$ з цих точок позначають мітками x_1, $m_2 = Np_2$ точок – мітками x_2 і т. д., $m_n = Np_n$ точок – мітками x_n.

Якщо є можливість реалізувати вибір одного з чисел $1,2,...,N$ так, щоб відповідні їм елементарні події були рівноймовірні, то маючи певне число з множини $\{1,2,...,N\}$ і знаючи відповідну йому мітку, дістають відповідне значення випадкової величини X. В такому разі значення x_1 з'являється з ймовірністю p_1; x_2 – з ймовірністю p_2 і т. д.; x_n – з імовірністю p_n.

Наприклад, якщо точки з абсцисами 1, 2, ..., 10 на осі Ox позначити мітками відповідно x_1, x_2, x_1, x_2, x_2, x_1, x_2, x_2, x_2, x_1 (Рис. 6.6.6) і вибирати одну з точок 1, 2, ..., 10 так, щоб вони були рівноймовірні між собою, то значення x_1 з'являтиметься з ймовірністю 0.4, а значення x_2 – з ймовірністю 0.6.

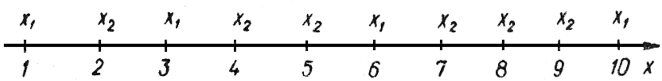


Рис. 6.6.6

Оскільки розміщення міток $x_1, x_2, ..., x_n$ не відіграє ніякої ролі, важливо тільки, щоб кожна з них повторювалась необхідну кількість разів, то мітки x_i доцільно розміщувати в порядку зростання індекса i. В такому разі якщо навмання вибирається значення γ з множини $\{1,2,...,N\}$ і значення $1,2,...,N$

рівноймовірні між собою, то значенням γ, стосовно яких задовільняється умова $\gamma \le p_1 N$, ставиться у відповідність число x_1; значенням γ, стосовно яких задовільняється умова $p_1 N < \gamma \le p_1 N + p_2 N$, ставиться у відповідність число x_2 і т.д.; значенням γ, стосовно яких задовільняється умова

$$N \sum_{k=1}^{i-1} p_k < \gamma \le N \sum_{k=1}^{i} p_k \,, \qquad (6.6.3)$$

ставиться у відповідність число x_i. Отже, треба мати можливість діставати в результаті випробування одну з N рівноможливих елементарних подій, тобто треба мати датчик рівноймовірних між собою чисел з множини $\{1, 2, ..., N\}$. Маючи такий датчик, легко реалізувати датчик з будь-яким дискретним розподілом імовірностей.

Датчик з дискретним розподілом імовірностей легко реалізувати також, маючи датчик випадкових чисел з рівномірним розподілом імовірностей на відрізку $[0\,;1]$.

Нехай відрізок $[0\,;1]$ довільним чином поділено на деякі підмножини $\Delta_1, \Delta_2, ..., \Delta_n$ так, що довжина всіх відрізків, з яких складено підмножину Δ_i, $i = 1, 2, ..., n$, відноситься до довжини всього відрізка $[0\,;1]$ як p_i (Рис. 6.6.7). Визначимо на відрізку $[0\,;1]$ функцію $y = \varphi(x) = x_i$, коли $x \in \Delta_i$. Тоді за рівномірного розподілу ймовірностей на відрізку $[0\,;1]$ буде

$$P_Y(\{x_i\}) = P_X(\varphi^{-1}(x_i)) = P_X(\Delta_i) = p_i \,.$$

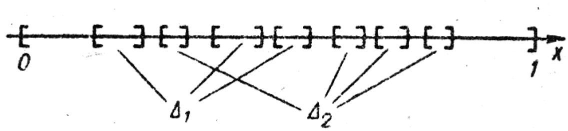


Рис. 6.6.7

Тобто значення x_i з'являтимуться з ймовірністю p_i, $i = 1, 2, ..., n$. Оскільки за рівномірного розподілу ймовірностей на відрізку $[0\,;1]$ ймовірність попадання в деяку вимірну підмножину Δ_i залежить тільки від міри цієї підмножини і не залежить від її структури та розміщення щодо відрізка $[0\,;1]$, то доцільно, як і в дискретному випадку, розміщувати спочатку всі проміжки множини Δ_1, потім всі проміжки множини Δ_2 і т.д. Тоді значенню $\gamma^* \in [0\,;1]$, яке дістають за допомогою датчика випадкових чисел з рівномірним розподілом ймовірностей на відрізку $[0\,;1]$, за якого задовільняється умова

$$\sum_{k=1}^{i-1} p_k < \gamma^* \le \sum_{k=1}^{i} p_k \ , \qquad (6.6.4)$$

відповідає число x_i.

Приклад 6.6.3. Побудувати датчик випадкових чисел з дискретним розподілом ймовірностей за законом Пуассона

$$P(\mu = m) = \frac{a^m}{m!} e^{-a} .$$

Оскільки $\frac{a^m}{m!} e^{-a} \to 0$, коли $m \to 0$ (як було показано раніше, $\sum_{k=0}^{m} \frac{a^k}{k!} e^{-a} \to 1$, коли $m \to \infty$, див. §3.16), то за досить великого $m = N$ з певною точністю можна вважати, що

$$\frac{a^N}{N!} e^{-a} = 0 \ , \ \sum_{m=0}^{N} \frac{a^m}{m!} e^{-a} = 1 .$$

Таким чином, за скінченну кількість кроків (не більшу від N) буде знайдено найменше число i таке, що виконуватиметься умова

$$\gamma^* \le \sum_{s=0}^{i} \frac{a^s}{s!} e^{-a} .$$

Отже, датчик випадкових чисел 0,1,2,... з розподілом ймовірностей Пуассона може бути реалізований так. Спочатку за допомогою датчика випадкових чисел з рівномірним розподілом ймовірностей на відрізку $[0\,;1]$ дістають число $\gamma^* \in [0\,;1]$. Потім поступово (один за одним) обчислюють доданки $\frac{a^s}{s!} e^{-a}$ (починаючи з $s = 0$ і збільшуючи s щоразу на 1) і додають їх до суми попередніх доданків $\sum_{k=0}^{s-1} \frac{a^k}{k!} e^{-a}$. Після кожного додавання до попереднього значення суми нового доданка $\frac{a^s}{s!} e^{-a}$ перевіряють, чи виконується умова

$$\gamma^* \le \sum_{k=0}^{s} \frac{a^k}{k!} e^{-a} .$$

Якщо умова не виконується, то обчислюють наступний доданок і знову повторюють попередні дії. Якщо вказана умова виконується (вперше), то здобуте значення s і є шуканим. Оскільки $\gamma^* \in [0\,;1]$, а $\sum_{k=0}^{N} \frac{a^k}{k!} e^{-a} = 1$ (за досить великого N, відповідного заданій точності обчислень), то через скінченне число кроків умова буде виконуватись. Розглянутий обчислювальний процес легко здійснити за допомогою комп'ютера.

Розглянемо тепер датчик випадкових чисел з наперед заданим неперервним розподілом ймовірностей, вважаючи, що датчик випадкових чисел з рівномірним розподілом ймовірностей на відрізку $[0\,;1]$ вже побудований.

Оскільки за даною щільністю розподілу ймовірностей $f_X(x)$ завжди можна знайти функцію розподілу ймовірностей $F_X(x)$, то вважатимемо, що з самого початку задано функцію розподілу ймовірностей $F_X(x)$. Виберемо числа a і b так, щоб із заданою точністю обчислень $\varepsilon > 0$ можна було вважати, що

$$F_X(a) = P_X(-\infty; a) = 0, \text{ коли } F_X(a) < \varepsilon ,$$

$$F_X(b) = P_X(-\infty; b) = 1, \text{ коли } F_X(b) > 1 - \varepsilon .$$

Тоді можна вважати, що вся одинична ймовірність (з точністю до 2ε) розподілена на відрізку $[a\,;b]$. Поділимо відрізок $[a\,;b]$ на кілька інтервалів завдовжки Δx точками поділу $x_0 = a, x_1, x_2, ..., x_n = b$, $x_{i+1} = x_i + \Delta x$.

Кожному з таких проміжків $[x_i\,;x_{i+1})$ поставимо у відповідність значення x_i (лівий кінець проміжка), вважаючи, що в результаті випробування дістанемо значення x_i, якщо навмання вибрана точка $x \in [a\,;b]$ буде знаходитись в інтервалі $[x_i\,;x_{i+1})$ (Рис. 6.6.8) (тут Δx може бути допустимою точністю обчислення значень x). Отже, кожне число x_i дістаємо з імовірністю $F_X(x_{i+1}) - F_X(x_i)$, що дорівнює ймовірності попадання в інтервал $[x_i\,;x_{i+1})$ одержуваного в результаті експерименту випадкового значення $x \in [a\,;b]$.

Таким чином дістанемо датчик випадкових чисел з дискретним розподілом імовірностей (Табл. 6.6.2).

Табл. 6.6.2

x_i	x_0	x_1	$\ldots$	x_{n-1}
$P_X(\{x_i\})$	$F_X(x_1)-F_X(x_0)$	$F_X(x_2)-F_X(x_1)$	$\ldots$	$F_X(x_n)-F_X(x_{n-1})$

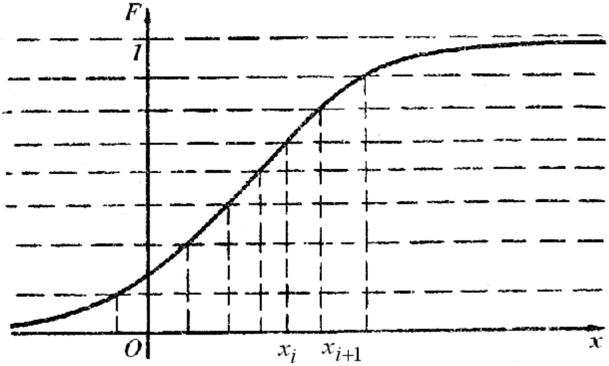


Рис. 6.6.8

За такого розподілу ймовірностей з ймовірністю $F_X(x_{i+1})-F_X(x_i)$ одержуватимемо випадкове значення x_i, якщо буде виконуватись умова

$$\sum_{k=0}^{i-1}(F_X(x_{k+1})-F_X(x_k))<\gamma^* \le \sum_{k=0}^{i}(F_X(x_{k+1})-F_X(x_k)),$$

або

$$F_X(x_i)-F_X(x_0)<\gamma^*<F_X(x_{i+1})-F_X(x_0).$$

Якщо врахувати, що $x_0=a$ і $F_X(x_0)=F_X(a)=0$, то остання умова набуває вигляду

$$F_X(x_i)<\gamma^* \le F_X(x_{i+1}).$$

Отже, діставши γ^*, знайдемо таке x_i, за якого

$$F_X(x_i)<\gamma^* \le F_X(x_{i+1}).$$

Переходячи до границі, коли $\Delta x \to 0$, з останньої умови дістанемо, що числу $\gamma^* \in [0;1]$, здобутому за допомогою датчика випадкових чисел з рівномірним розподілом імовірностей на відрізку $[0\,;1]$, слід поставити у відповідність таке число x, за якого

$$F_X(x)=\gamma^*. \qquad (6.6.5)$$

Розв'язуючи це рівняння відносно x, дістанемо значення випадкової величини X з неперервним розподілом ймовірностей, що описується за функцією розподілу ймовірностей

$$F_X(x) = \int\limits_{-\infty}^{x} f_X(x)dx.$$

Оскільки функція $F_X(x)$ неперервна, то рівняння (6.6.5) завжди має розв'язок.

Розглянутий метод знаходження значення випадкової величини X називають *методом обернення функції розподілу ймовірностей.*

Приклад 6.6.3. Побудувати датчик випадкових чисел з експоненціальним розподілом ймовірностей на множині її значень, щільність якого $f_\tau(t) = \lambda e^{-\lambda t}$, $t \geq 0$.

Знайдемо функцію експоненціального розподілу ймовірностей

$$F_\tau(t) = \int\limits_0^t f_\tau(t)dt = \int\limits_0^t \lambda e^{-\lambda t}dt = (-e^{-\lambda t})\Big|_o^t = 1 - e^{-\lambda t}.$$

Згідно з рівністю (6.6.5)

$$1 - e^{-\lambda t} = \gamma^*,$$

звідки

$$e^{-\lambda t} = 1 - \gamma^*,\ -\lambda t = \ln(1-\gamma^*),\ t = -\frac{1}{\lambda}\ln(1-\gamma^*). \qquad (6.6.6)$$

Отже, щоб дістати значення випадкової величини з експоненціальним розподілом ймовірностей на множині її значень, досить знайти значення γ^* випадкової величини з рівномірним розподілом ймовірностей на відрізку $[0\,;1]$, а потім обчислити число t за формулою (6.6.6).

Розглянутий датчик випадкових чисел широко використовується для імітації (за допомогою комп'ютера) функціонування деяких систем масового обслуговування. Так, проміжок часу між двома послідовними запитами на обслуговування є випадковою величиною з експоненціальним розподілом імовірностей на множині її значень. Математичне сподівання цієї величини є $\frac{1}{\lambda}$, а тому λ можна тлумачити як середню інтенсивність, з якою надходять запити (кількість запитів за одиницю часу). Використовуючи датчик випадкових чисел (6.6.6), можна дістати моменти часу, в які надходять запити на обслуговування, а саме, вважаючи, що попередній запит надійшов у момент $t = t_i$ і звертаючись до датчика випадкових чисел (6.6.6), дістанемо, що наступний запит надійде в момент $t_{i+1} = t_i + \tau$, де τ – щойно знайдене за (6.6.6) випадкове значення. Після звернення до датчика випадкових чисел дістають новий проміжок часу до наступного запиту, а отже, і момент його надходження, оскільки момент надходження попереднього запиту вже визначено.

Час обслуговування запиту на обслуговуючому пристрої також може бути випадковим (можливо, з розподілом імовірностей того самого типу). Знаходячи за допомогою відповідного датчика випадкових чисел час обслуговування чергового запиту, можна на розглянутій математичній моделі імітувати функціонування обслуговуючої системи, обчислити її пропускні характеристики, процент запитів, які дістають відмову в обслуговуванні, час простоювання обслуговуючих пристроїв тощо. Імітація функціонування такої обслуговуючої системи може бути здійснена за допомогою комп'ютера. Отже, за короткий час, не будуючи реальної обслуговуючої системи, можна до деякої міри охарактеризувати таку систему, імітуючи її функціонування відповідно до розглянутої математичної моделі.

Крім розглянутих вище існують і деякі інші способи знаходження випадкових чисел. Наприклад, дістаючи досить велику кількість випадкових чисел і беручи їх середнє арифметичне, дістають значення випадкової величини, розподіл ймовірностей на множині значень якої, згідно з центральною граничною теоремою, можна вважати нормальним.

Датчик випадкових чисел з рівномірним розподілом ймовірностей можна здійснити різними способами.

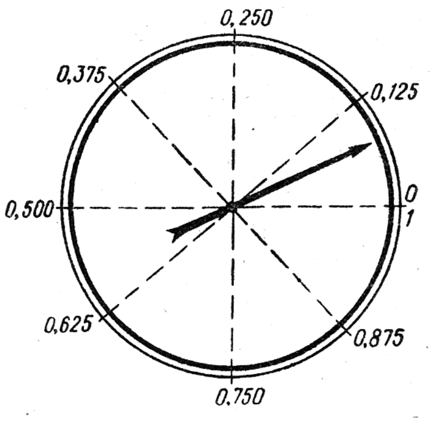


Рис. 6.6.9

Наприклад, це може бути урна з пронумерованими від 1 до n кульками, за допомогою якої можна дістати n рівноймовірних між собою чисел. Датчиком випадкових чисел може бути рулетка – диск, який вільно обертається навколо центра. Диск розкручують і дають вільно обертатися до повної зупинки. Позначка, нанесена на диск, після зупинки може вказувати на будь-яке число з інтервалу $[0\,;2\pi R)$ (Рис. 6.6.9).

Таким чином дістають окремі значення випадкової величини X з рівномірним неперервним розподілом ймовірностей на інтервалі $[0\,;2\pi R)$ (за умови, що диск симетричний і вільно обертається). Датчиком випадкових чисел може бути й спеціальна таблиця випадкових чисел.

На практиці, виконуючи обчислення за допомогою комп’ютера, часто випадкові числа дістають за спеціальними досить швидкодіючими алгоритмами (такі числа називають *псевдовипадковими)* або за допомогою спеціальних фізичних приладів.

Щоб дістати випадкове «значення» двохвимірної випадкової величини, наприклад з рівномірним розподілом ймовірностей на множині

$$W = \{(x, y) \mid x \in [0\,;1],\; y \in [0\,;1]\}$$

знаходять послідовно два випадкові значення за допомогою датчика випадкових чисел з одновимірним розподілом ймовірностей на відрізку $[0\,;1]$.

Є багато методів для знаходження псевдовипадкових чисел з різними розподілами ймовірностей (особливо з рівномірним розподілом). Сьогодні у всі мови програмування високого рівня поряд з іншими стандартними функціями ($\sin(x)$, $\cos(x)$, $\log(x)$, $\exp(x)$ і т. д.) включається і датчик випадкових чисел (функція $rnd(x)$), що є одним із свідчень про те, наскільки важливими є для практики і як широко застосовуються методи теорії ймовірностей та математичної статистики.

Вправи для самостійного виконання

6.6.1. Перевірити, чи правильні твердження:

1. За допомогою імітаційної моделі можна отримати точні характеристики відповідного реального процесу.
2. Метод Монте-Карло – це метод дослідження реальних процесів за допомогою імітаційних моделей.
3. Статистичне моделювання – це те саме, що й метод Монте-Карло.
4. Метод Монте-Карло застосовний лише для недетермінованих задач.
5. За методом Монте-Карло для обчислення визначеного інтеграла отримується така сама точність, що й за формулою прямокутників.
6. За методом Монте-Карло можна обчислити число π з будь-якою точністю.
7. Генератор випадкових чисел – це засіб одержання значень, що:

а) є довільними і ні від чого не залежними;

б) є значеннями фіксованої випадкової величини із заданим одновимірним розподілом ймовірностей на множині її значень.

8. За будь-якої випадкової величини з дискретною множиною значень можна побудувати генератор її значень.
9. За допомогою будь-якого генератора випадкових чисел можна

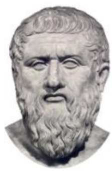
отримати лише скінченну кількість чисел.

10. Генератор випадкових чисел можна побудувати лише стосовно випадкової величини з дискретною множиною значень.

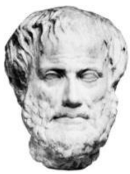
11. Псевдовипадкові числа є результатом звернення до генератора випадкових чисел.

Історичні відомості


Демокріт


Платон

Проблеми дослідження випадкових явищ хвилювали людство ще у сивій давнині. Так, видатні давньогрецькі філософи *Демокріт* (460-370 р.р. до н.е.), *Платон* (428-348р.р. до н.е.), *Арістотель* (384-322 р.р. до н.е.) та інші розмірковували над неминучими та випадковими явищами. Вельми загальні висловлювання щодо випадкових явищ робили мислителі стародавніх Китаю, Індії та Близького Сходу.


Арістотель

Про випадкові експерименти з підкиданням гральних кубиків писали у 1220-1250 роках *Річард де Форніваль* (1200-1250), у 1487 році *Лука Пачолі* (1445-1515), у 1526 році *Джироламо Кардано* (1501-1576), у 1556 році *Ніколло Тарталья* (1449-1557), а також видатний італійський вчений *Галілео Галілей* (1564-1642), який, мабуть, першим розглянув випадкові похибки, що виникають у деяких експериментах.


Л. Пачолі

Поняття ідеалізованого випадкового експерименту ввів у 1918 році відомий німецький математик *Ріхард Мізес* (1883-1953), а множину Ω, як модель множини наслідків випадкового випробування, і випадкову подію, як підмножину простору Ω, ввів у 1936 році видатний російський математик *Андрій Миколайович Колмогоров* (1903-1987).


Д. Кардано

До кінця XIX століття не виникало питання про необхідність строгого математичного означення «випадкової події». З давніх давен математики оперували з випадковими подіями (знаходили суми подій, добутки та протилежні до них події), спираючись лише на інтуїтивне, часто досить суб’єктивне уявлення про них. Лише після того, як теорія ймовірностей досягла досить високого рівня застосувань, виникла потреба в її логічному обґрунтуванні.

Н. Тарталья

Г. Галілей

Р. Мізес

А.М. Колмогоров

Д. Гільберт

С.Н. Бернштейн

Про це свідчить, зокрема, одна із славетних математичних проблем (шоста проблема, пов'язана з аксіоматичним обґрунтуванням математичних теорій, включаючи теорію ймовірностей) видатного німецького математика *Давіда Гільберта* (1862-1943), сформульованих ним на II Міжнародному конгресі математиків (Париж, 1900 р.). У 1917 році відомий російський математик *Сергій Натанович Бернштейн* (1880-1968) навів чітке математичне тлумачення випадкової події як елемента так званої булевої алгебри.

Основні вимоги 1_s-3_s до сукупності S подій ввів у 1936 році Андрій Миколайович Колмогоров.

Галілео Галілей розумів сутність поняття ймовірності, як числа, від якого мало відхиляються найбільша кількість відносних частот спостережених випадкових похибок вимірювання певної величини.

У роботах англійських математиків *Джона Граунта* (1620-1675) і *Вільяма Петті* (1623-1687), опублікованих у 1662 та 1682 роках і присвячених політичній арифметиці, для дослідження випадкових соціальних явищ мабуть уперше було суттєво використано поняття відносної частоти випадкової події.

Сутність поняття статистичної ймовірності розкрита у 1693 році видатним швейцарським математиком *Якобом Бернуллі* (1654-1705), який писав: «*Що не дано нам вивести apriori* (тобто наперед передбачити до проведення досліду), *те принаймні можна дістати aposteriori*, *тобто з численних спостережень результатів подібних дослідів*. Тому *можна передбачати, що деяке явище згодом зможе відбутися у стількох саме випадках, у скількох воно раніше було відмічене як таке, що відбулося за подібних умов…*».

Уявлення про правила обчислення ймовірності суми подій є вже в роботах, опублікованих в 1663 році відомим англійським астрономом *Едмондом Галлеєм* (1656-1742). У роботах, здійснених *Якобом Бернуллі* у 1693 році та *Миколою Бернуллі* (1687-1759) у 1718 році, сформульовано правила обчислення ймовірності протилежної події та суми подій.

Д. Граунт

Термін «ймовірність» мабуть першим використав відомий філософ *Бенедикт Спіноза* (1632–1677) у назві однієї із своїх робіт: «Допис про математичну ймовірність», проте ніякого означення ймовірності у цій роботі нема.

В. Петті

У своїй роботі «Книга про гру в кості» *Джероламо Кардано* кілька разів пропонував розглянути відношення m/n, яке іноді називають «класичним означенням ймовірності», проте важливість цього відношення він не помітив.

Я. Бернуллі

Недосконала форма «класичного означення» ймовірності введена у 1693 році *Якобом Бернуллі* у трактаті «Мистецтво припущень». Чітке «класичне означення» ймовірності з вимогою рівноможливості усіх шансів ввів у 1812 р. видатний французький вчений *П'єр Лаплас* (1749–1827) у трактаті «Аналітична теорія ймовірностей».

Е. Галлей

Видатні швейцарські вчені *Даніїл Бернуллі* (1700–1782) та *Леонард Ейлер* (1707–1782) зробили значний внесок у застосування теорії ймовірностей у демографії. Вони фактично започаткували основи сучасної демографії.

Статистичний та емпіричний підхід до формування поняття ймовірності розвинув англійський математик *Рональд Фішер* (1890–1962) та німецький математик *Ріхард Мізес*.

Б. Спіноза

Після виходу книги відомого французького математика *Жозефа Бертрана* (1822–1900) «Числення ймовірностей», опублікованої у 1899 році, математики звернули увагу на недосконалість існуючих означень ймовірності і на необхідність логічного обґрунтування теорії ймовірностей.

П. Лаплас

Д. Бернуллі

Л. Ейлер

Р. Фішер

Ж. Бертран

У. Феллер

Визначальні властивості ймовірностей $1_P - 3_P$ ввів у 1936 році *Андрій Миколайович Колмогоров*. Тому властивості 1_S-3_S та 1_P-3_P називають *системою аксіом теорії ймовірностей А.М. Колмогорова*.

Відомий американський математик *Уільям Феллер* (1906–1970) вважав, що у 30-х роках XX століття лише у Радянському Союзі теорія ймовірностей розвивалася як справжня математична наука, що стало можливим після створення А.М. Колмогоровим аксіоматичної теорії ймовірностей. Завдяки цьому теорія ймовірностей перейшла від етапу напівмістичних міркувань, які переважали ще у 20-х роках XX ст., до сучасного етапу розвитку.

Чітке формулювання теореми про ймовірність суми подій:

$$P(A+B) = P(A) + P(B) - P(AB)$$

міститься у роботі англійського математика *Томаса Байєса* (1702-1761), опублікованій лише у 1763 році. Т. Байєс по суті ввів також поняття умовної ймовірності і визначив, що

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B).$$

Поняття незалежності подій введено англійським математиком *Абрахамом Муавром* (1667-1773), який також чітко сформулював правило знаходження ймовірності добутку подій.

Формулу повної ймовірності першим у 1795 році навів *П'єр Лаплас* у своєму трактаті «Дослід філософії теорії ймовірностей». У цьому трактаті він вперше навів і формулу, яку зараз називають *формулою Байєса*, проте швидше за все Байєс не знав цієї формули.

Поняття випадкової величини фактично використовували вже у 1654 році засновники теорії ймовірностей видатні французькі математики *Блез Паскаль* (1623-1662) та *П'єр Ферма* (1601-1665). Випадкові похибки вимірювань, введені *Галілео Галілеєм* у першій

Т. Байєс

А. Муавр

Б. Паскаль

П. Ферма

С. Пуассон

К. Гаусс

половині XVII століття, також є одним з перших прикладів випадкових величин.

Поняття простої випадкової величини введено у 1832 році відомим французьким математиком *Симоном Пуассоном* (1781-1840) в роботі «Про ймовірність середніх результатів спостережень», проте сам термін «випадкова величина» вперше зустрічається у роботах видатного російського математика *Пафнутія Львовича Чебишова* (1821-1894).

Строге означення випадкової величини ввів наприкінці 20-х років XX століття *Андрій Миколайович Колмогоров*.

Уявлення про розподіл ймовірностей на певній (досить простій) множині виникли ще у сивій давнині. Проте чітке розуміння сутності поняття функції розподілу ймовірностей та класифікація розподілів сформувалося лише наприкінці XIX століття і на початку XX століття після побудови строгої теорії міри (теорії вимірювання довжин, площ, об'ємів, мас тощо) завдяки працям відомих французьких математиків *Каміла Жордана* (1838-1892), *Еміля Бореля* (1871-1956) та *Анрі Лебега* (1875-1941). *Олександр Михайлович Ляпунов* (1857-1918) у 1900 році першим ввів сучасне означення функції розподілу ймовірностей. Лише після того, як *Андрій Миколайович Колмогоров* побудував аксіоматичну теорію ймовірностей, стало можливим введення строгих означень, зокрема й означення функції розподілу ймовірностей.

Окремі випадки неперервних та абсолютно неперервних розподілів ймовірностей першими досліджували англійський математик *Абрахам Муавр* у 1733 році, німецький математик *Карл Гаусс* (1777-1855) у 1809 році та французький математик *П'єр Лаплас* у 1812 році. Важливі результати, пов'язані з функціями розподілу ймовірностей, належать російському математику *Пафнутію Львовичу Чебишову* та його


К. Жордан

учням *Маркову Андрію Андрійовичу* (1856-1922) і *Ляпунову Олександру Михайловичу*.

Класифікація функцій розподілу ймовірностей на дискретні, неперервні, абсолютно неперервні та сингулярні зустрічається у дослідженнях багатьох математиків, які у XX столітті розвивали загальну теорію міри, фундаментальні основи якої належать французькому математику *Анрі Лебегу*.


Е. Борель

Значні результати, пов’язані з розподілами ймовірностей, належать відомим українським математикам *Йосипові Іллічу Гіхману* (1918-1985), *Володимиру Семеновичу Королюку* (1925), *Анатолію Володимировичу Скороходу* (1930-2011), *Михайлу Йосиповичу Ядренку* (1932-2004) та іншим.


А. Лебег

Значний вплив на розвиток основ теорії ймовірностей мав трактат видатного нідерландського вченого *Християна Гюйгенса* (1629-1695) «Про розрахунки в азартних іграх», виданий у 1657 році. У цьому трактаті вперше вводиться поняття математичного сподівання стосовно деяких простих випадкових величин, множини значень яких складаються з двох або трьох чисел.


П.Л. Чебишов

Микола Бернуллі у 1709 році вивчав математичні сподівання простих випадкових величин з довільною скінченною множиною значень.

Середнє квадратичне відхилення увійшло в математику завдяки працям *Галілео Галілея*, *Абрахама Муавра*, *Леонарда Ейлера*, *Карла Гаусса* та *П’єра Лапласа*.


О.М. Ляпунов

Математичне сподівання і дисперсію довільної випадкової величини наприкінці XIX століття систематично вивчали *Пафнутій Львович Чебишов* та його учні *Андрій Андрійович Марков* та *Олександр Михайлович Ляпунов*.

П.Л. Чебишов першим довів властивості математичного сподівання суми та добутку і


А.А. Марков


В.С. Королюк


Й.І. Гіхман


А.В. Скороход


М.Й. Ядренко


Х. Гюйгенс

дисперсії суми випадкових величин, проте опубліковані ці доведення лише у 1913 році у відомому підручнику *А.А. Маркова* «Числення ймовірностей». У цьому підручнику вперше введено поняття незалежних випадкових величин.

Повторні незалежні випробування та біноміальні ймовірності розглядалися ще на початку розвитку теорії ймовірностей.

Якоб Бернуллі був одним з перших, хто систематично досліджував біноміальні ймовірності. Тому його ім’ям названо формулу для обчислення біноміальних ймовірностей, а схему повторних незалежних випробувань, в кожному з яких ймовірність події A не змінюється, названо *схемою Бернуллі*.

Джероламо Кардано ще у першій половині XVI століття мав уявлення про закон великих чисел. Наприкінці XVII століття у роботах *Едмунда Галлея* з теорії ймовірностей зустрічаються міркування, що нагадують закон великих чисел. *Якоб Бернуллі* наприкінці XVII століття використав біноміальні ймовірності для обґрунтування закону великих чисел, проте суттєве узагальнення і обґрунтоване доведення цього закону навів у 1867 році *Пафнутій Львович Чебишов*. Саму назву «закон великих чисел» ввів у 1837 році відомий французький математик *Симон Пуассон*.

Важливі результати, пов’язані з законом великих чисел, належать французькому математику *Емілю Борелю* та російським математикам *Андрію Миколайовичу Колмогорову*, *Олександру Яковичу Хінчину* (1894-1959), *Валерію Івановичу Глівенку* (1897-1940) та *Борису Володимировичу Гнєденку* (1912-1995).

Біноміальні ймовірності використовують для проведення так званого статистичного (або вибіркового) контролю якості продукції, що випускається у великих обсягах. Пріоритет у такому практичному застосуванні теорії ймовірностей належить відомому російському математику українського походження *Михайлу Васильовичу Остроградському* (1801-1862), який у 1846 році опублікував роботу, присвячену такому застосуванню.

О.Я. Хінчин

Рівномірний дискретний розподіл ймовірностей фактично використовувався ще на початку виникнення стохастичних ідей та методів.

Б.В. Гнеденко

Геометричний розподіл ймовірностей пов’язаний із запропонованою *Якобом Бернуллі* схемою послідовних випробувань для випадку, коли ці випробування проводяться доти, поки не відбудеться дана подія.

Свій розподіл *Симон Пуассон* запропонував у 1837 році. Він також довів теорему про наближення $P(B_{n,m}) \approx e^{-a_n} \dfrac{a_n^m}{m!}$.

М.В. Остроградський

Рівномірний неперервний розподіл ймовірностей вперше згадується у 1692 році у перекладі книги *Християна Гюйгенса* «Про розрахунки в азартних іграх». Цей розподіл суттєво використовував французький вчений *Жорж Бюффон* (1707-1788).

Ж. Бюффон

Систематично і ефективно рівномірний неперервний розподіл ймовірностей використовував *П’єр Лаплас*. Він у 1809 році за допомогою випадкових величин X з рівномірними розподілами ймовірностей на відрізку $[-h; h]$ довів інтегральну граничну теорему, яку задовго до нього у 1722 році довів за $p = \dfrac{1}{2}$ *Абрахам Муавр* за допомогою власної локальної граничної теореми.

К. Пірсон

Поняття ймовірнісних просторів фактично використовували вже засновники теорії ймовірностей, проте чіткі відповідні означення з’явилися лише у XX столітті після створення сучасних теорії міри та теорії ймовірностей.

Е. Пірсон

Формули, пов’язані з обчисленням

Е. Нейман

В.Я. Буняковський

А. Кетле

Ф. Гальтон

У. Госсет (Стьюдент)

кількостей різноманітних розміщень, перестановок, сполучень вивчаються у розділі математики, який називається «Комбінаторика». Тому і згадані формули називають комбінаторними. Засновниками комбінаторики були *Блез Паскаль* і *П'єр Ферма*. Значний вклад у розвиток комбінаторики внесли *Готфрід Лейбніц*, *Якоб Бернуллі*, *Леонард Ейлер* і багато інших вчених.

Велику роль у розповсюдженні ідей теорії ймовірностей та математичної статистики в Росії та Україні відіграли видатні російські математики українського походження В.Я. Буняковський (1804-1889) та М.В. Остроградський.

У фундаментальній праці В.Я. Буняковського "Основи математичної теорії ймовірностей" окрім суто математичних задач була розв'язана також задача стосовно формування російської термінології, яка залишилася майже незмінною до наших часів.

Праці М.В. Остроградського з теорії ймовірностей були присвячені розв'язуванню важливих практичних задач.

Як вважає відомий російський математик Б.В. Гнєденко, захоплення теорією ймовірностей у першій чверті XIX ст. привело до величезної кількості праць, пов'язаних із застосуванням цієї теорії до різних проблем природничих наук та суспільного життя. Багато з цих застосувань були мало обґрунтованими і сприймалися математиками як "математичні скандали". Тому це захоплення змінилося глибоким розчаруванням і повним скептицизмом щодо застосувань теорії ймовірностей до наукового пізнання світу. Подальший розвиток теорії ймовірностей потребував уточнення основних її положень. Потрібно було встановити предмет теорії ймовірностей, область її застосувань, вивчити та підсилити її специфічні методи досліджень. Велику роботу

Ж. Бертран

А. Пуанкаре

у цьому напрямку провів видатний російський математик П.Л. Чебишов.

П.Л. Чебишов залишив помітний вклад у багатьох розділах математики, зокрема, і у теорії ймовірностей, де він узагальнив закон великих чисел, довів так звану центральну граничну теорему для суми незалежних випадкових величин та отримав багато інших результатів. Його курс теорії ймовірностей, що він читав у Петербурзькому університеті, відрізняється чіткістю формулювань та обґрунтованістю доведень тверджень.

На думку видатного російського математика А.М. Колмогорова завдяки П.Л. Чебишову була створена російська математична школа, яка стала найкращою в світі в багатьох розділах математики, зокрема і в теорії ймовірностей.

Серед найвідоміших математиків, які були учнями П.Л. Чебишова, слід назвати А.А. Маркова та О.М. Ляпунова, які стали видатними математиками саме завдяки своїм дослідженням з теорії ймовірностей.

Книга А.А. Маркова “Числення ймовірностей”, перше видання якої відбулося у 1900 р., а четверте – у 1924 р., на протязі багатьох років була найкращою серед тих, за якими навчалися російські математики. В цій книзі, зокрема, розкривається, в якому розумінні статистична ймовірність $P_n^*(A)$ близька до ймовірності $P(A)$ за великих n: ймовірність значного відхилення $P_n^*(A)$ від $P(A)$ є близькою до нуля, проте це не означає, що значні відхилення неможливі за великих n. Це означає, що $\lim\limits_{n\to\infty} P_n^*(A) \neq P(A)$ у класичному розумінні границі послідовності.

Величезні досягнення в теорії ймовірностей майже не застосовуються в математичній статистиці наприкінці XIX ст. Це, зокрема, видно з важливих праць бельгійського математика А. Кетле (1796-1874) та англійських математиків Ф. Гальтона (1822-1911) і К. Пірсона (1857-1936). Ця недоречність була усунута на початку XX ст. завдяки працям багатьох математиків. Зокрема, значну роль відіграли представники англо-американської школи У. Госсет (Стьюдент) (1876-1937), Е. Пірсон (1895-1980) і Е. Нейман (1894-1981), завдяки працям яких була створена теорія статистичної перевірки гіпотез.

Наприкінці XIX ст. французький математик Ж. Бертран (1822-1900) навів ряд парадоксів, пов'язаних з теорією ймовірностей, а видатний французький математик А. Пуанкаре (1854-1912) узагальнив ці парадокси, де підкреслювалися нечіткість та неточність деяких означень понять теорії ймовірностей, а отже поставала необхідність відповідних уточнень. Зробити це стало можливим завдяки аксіоматичному методу, який на початку XX століття пронизує багато галузей математики.

У XX столітті теорія ймовірностей поступово теж перетворюється у строгу аксіоматичну теорію. Це відбулося завдяки працям багатьох математиків. Так, англійський математик Р. Фішер розвинув статистичний або емпіричний підхід до формування поняття ймовірності. Поняття простору елементарних подій ввів німецький математик Р. Мізес. Російський математик С.Н. Бернштейн у 1917 р. надрукував роботу "Досвід аксіоматичного обґрунтування теорії ймовірностей", а у 1927 р. видав книгу "Теорія ймовірностей", в якій описав власну аксіоматичну теорію ймовірностей. Ця книга вважається однією з найкращих серед творів світової літератури з теорії ймовірностей.

Але дійсно вирішальним етапом розвитку теорії ймовірностей стала праця А.М. Колмогорова "Основні поняття теорії ймовірностей" (1937), у якій він описав свою аксіоматику теорії ймовірностей і після якої теорія ймовірностей зайняла рівноправне місце серед інших математичних дисциплін.

Великі досягнення в розбудові теорії ймовірностей та математичної статистики мали також російські математики О.Я. Хінчин (1894-1959), Є.Є. Слуцький (1880-1948), Б.В. Гнєденко та багато інших, а також українські математики Й.І. Гіхман (1918-1985), Ю.М. Єрмольєв (1936), І.М. Коваленко (1935), В.С.Королюк (1925), В.С.Михалевич (1930-1994), А.В. Скороход (1930-2011), М.Й. Ядренко (1932-2004) та інші. Методи досліджень теорії ймовірностей сьогодні широко використовуються в теорії масового обслуговування, теорії надійності, теорії ігор, статистичній фізиці, кібернетиці, математичній статистиці, інформатиці.

Математична статистика – це наука, де вивчаються проблеми виявлення структури ймовірнісно-статистичних моделей досліджуваних явищ за даними експериментальних спостережень.

На базі теорії ймовірностей і математичної статистики ґрунтується багато досліджень у сучасній біології, медицині, сільському господарстві, військових науках, педагогіці, мовознавстві і т. д.

Внаслідок стрімкого розвитку комп’ютеризованих інформаційно-комунікаційних технологій з’явилися можливості широкого використання їх для розв’язування прикладних проблем на основі теорії ймовірностей та математичної статистики, зокрема на основі статистичного моделювання.

Додаток 1

Елементи комбінаторики

На практиці часто доводиться розглядати задачі, в яких із деякої скінченної множини елементів вибираються певні підмножини і треба з'ясувати взаємозв’язки між такими підмножинами, визначити їх кількість, властивості тощо. Оскільки в таких задачах йдеться про ті чи інші комбінації елементів, то відповідну область математики *називають комбінаторикою (теорією сполук).* Розглянемо такий приклад.

Із п’яти літер розрізної азбуки утворено слово ЗІРКА. Серед цих літер одна за однією без повернення назад вибираються три літери. Скільки всіх трилітерних слів можна утворити з даних п’яти літер? В даному разі словом вважається будь-який набір з трьох літер, розміщених у певному порядку, незалежно від того, має слово зміст, чи ні.

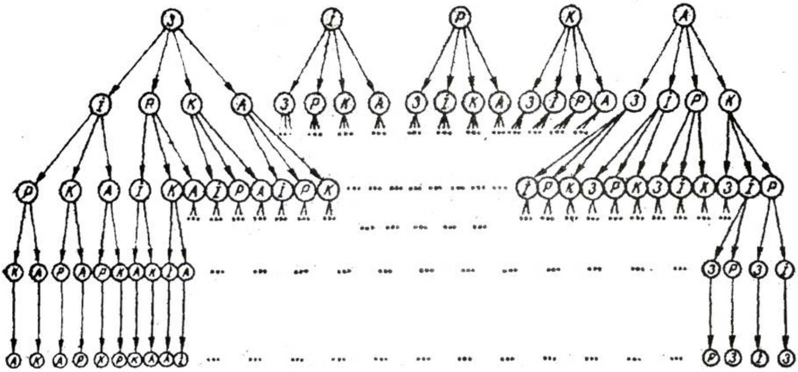


Рис. Д1.1

Очевидно, що слова, утворені з одних і тих самих літер, але розміщених у різному порядку, будуть різними. Якщо кожна вибрана літера не повертається назад, то з п’яти різних літер можна утворити $5 \cdot 4 \cdot 3$ різних трилітерних слів, оскільки для вибору першої літери існує п’ять можливих варіантів, для вибору другої – чотири, для вибору третьої – три (Рис. Д1.1). На рисунку Д1.1 показано граф, у вершинах якого розміщені зазначені літери.

Якщо пройти від якоїсь з вершин першого рівня (їх п’ять) до якоїсь з вершин третього рівня (їх 60) одним із вказаних стрілками можливих шляхів, то з літер, через які пройде такий шлях, і утвориться одне з можливих слів.

В даному прикладі з п'ятиелементної множини вибираються триелементні підмножини, причому такі підмножини вважаються різними, навіть якщо вони утворені з одних і тих самих елементів, але розміщених в різних порядках.

m-елементні підмножини n-елементної множини називаються *впорядкованими,* якщо вони визначаються за набором елементів та за порядком їх розміщення. Якщо порядок розміщення елементів несуттєвий, тобто підмножини визначаються лише за набором елементів незалежно від порядку, в якому вони розміщені, то такі підмножини називаються *невпорядкованими.*

Якщо є дві різні невпорядковані підмножини**,** то принаймні в одній з них знайдеться елемент, який не входить до іншої.

m-елементні впорядковані підмножини n-елементної множини називають розміщеннями з n елементів по m елементів.

Число m-елементних впорядкованих підмножин n-елементної множини (розміщень з n елементів по m елементів) позначають A_n^m. Узагальнюючи міркування наведеного прикладу, одержуємо

$$A_n^m = n(n-1)(n-2)\ldots(n-(m-1)) =$$
$$= \frac{n(n-1)\ldots(n-(m-1))(n-m)\ldots 2\cdot 1}{(n-m)\ldots 2\cdot 1} = \frac{n!}{(n-m)!}.$$

Якщо у наведеному прикладі кожну вибрану літеру зафіксувати і повернути назад, а потім наступну літеру вибирати з тих самих п’яти літер, то кількість можливих варіантів вибрати літеру залишатиметься однаковою як під час вибору першої, так і під час вибору другої та третьої літер. Тоді кількість трилітерних слів дорівнюватиме $5\cdot 5\cdot 5 = 5^3$.

m-елементні впорядковані множини, *утворені послідовним вибором елементів з однієї й тієї самої множини, називаються розміщеннями з повтореннями з n елементів по m елементів*.

Число таких розміщень позначають R_n^m і обчислюють за формулою

$$R_n^m = n^m$$

Якщо розглянути n-елементні впорядковані підмножини n-елементної множини (розміщення з n елементів по n елементів), то утворяться сполуки з одних і тих самих n елементів, лише розміщених в різних порядках. Такі n-елементні впорядковані підмножини n-елементної множини (розміщення з n елементів по n елементів) називають *перестановками з n елементів.* Число таких перестановок позначають P_n і обчислюють за формулою

$$P_n = A_n^n = n(n-1)(n-2)\ldots 2\cdot 1 = n!$$

Якщо в наведеному прикладі з даних п’яти літер утворювати не трилітерні, а п’ятилітерні слова, то кількість таких слів дорівнюватиме $5\cdot 4\cdot 3\cdot 2\cdot 1$, тобто $P_5 = 5!$

Узагальнення наведених міркування приводить до основного *правила комбінаторики* (*правила множення*): *якщо є скінченні множини* $E_1, E_2, \ldots, E_k$ *, причому в множині* E_i *міститься* n_i *елементів,* $i = 1, 2, \ldots, k$ *, то в декартовому добутку* $E_1 \times E_2 \times \ldots \times E_k$ *міститься* $n_1 n_2 \ldots n_k$ *впорядкованих наборів* $(x_1, x_2, \ldots, x_k)$ *таких, що елемент* x_1 *пробігає множину* E_1*, елемент* x_2 *– множину* E_2*, і т. д., елемент* x_k *– множину* E_k.

В наведених прикладах перша літера вибиралась з п'ятиелементної множини, друга – з чотириелементної, третя – з триелементної, четверта – з двохелементної, п'ята – з одноелементної.

m*-елементні невпорядковані підмножини* n*-елементної множини називають комбінаціями з* n *елементів по* m *елементів.* Число таких комбінацій позначають C_n^m .

Якщо елементи невпорядкованої m-елементної підмножини розмістити в усяких можливих порядках, то дістанемо $P_m = m!$ впорядкованих m-елементних підмножин. Таким чином, з усіх невпорядкованих m-елементних підмножин можна дістати всі впорядковані m-елементні підмножини.

Отже

$$C_n^m P_m = A_n^m$$

звідки

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n \cdot (n-1) \ldots (n-(m-1))}{m!} = \frac{n!}{m!(n-m!)}$$

Очевидно, що $C_n^m = C_n^{n-m}$, оскільки з вибором кожної m-елементної невпорядкованої множини однозначно визначається єдина $(n-m)$-елементна невпорядкована підмножина, доповнююча m-елементну підмножину до n-елементної множини. Тому кількість m-елементних невлорядкованих підмножин в n-елементній множині дорівнює кількості $(n-m)$-елементних невпорядкованих підмножин в тій самій n-елементній множині. Природно вважати, що $C_n^0 = C_n^n = 1$ і $0! = 1$, оскільки n-елементна невпорядкована підмножина n-елементної множини єдина.

Розглянемо розклад бінома $(a+b)^n$ за степенями a і b. Кілька перших розкладів набувають вигляду

$$(a+b)^0 = 1,$$

$$(a+b)^1 = a + b,$$

$$(a+b)^2 = (a+b)(a+b) =$$
$$= aa + ba + ab + bb,$$

$$(a+b)^3 = (aa+ba+ab+bb)(a+b) =$$
$$= aaa+baa+aba+bba+aab+bab+abb+bbb,$$
$$(a+b)^4 = aaaa+baaa+abaa+bbaa+aaba+baba+$$
$$+abba+bbba+aaab+baab+abab+bbab+$$
$$+aabb+babb+abbb+bbbb.$$

Таким чином, щоб, наприклад, у розкладі $(a+b)^4$ дістати доданки, до яких входять добутки a^2b^2, треба з чотирьох місць, на яких можуть бути множники a або b, вибрати два місця, на них записати літеру a, а на решту місць записати літеру b. Отже, доданків, в яких містяться двічі множник a і двічі множник b, буде стільки, скількома способами з множини з чотирьох місць 1, 2, 3, 4 можна вибрати невпорядковану підмножину з двох місць (порядок елементів у двохелементній підмножині несуттєвий, оскільки, якщо вказано для літери a підмножину місць {2, 3} чи {3, 2}, то це означатиме, що літера a буде записана на другому і на третьому місці в обох випадках). Таким чином, доданків, в які входять двічі множник a, а решту разів множник b, буде C_4^2.

Узагальнюючи наведені міркування, дістаємо

$$(a+b)^n = C_n^0 a^0 b^n + C_n^1 a^1 b^{n-1} + C_n^2 a^2 b^{n-2} + \ldots$$
$$\ldots + C_n^{n-1} a^{n-1} b^1 + C_n^n a^n b^0 = \sum_{k=0}^{n} C_n^k a^k b^{n-k}.$$

Цю формулу називають формулою *бінома Ньютона*, а коефіцієнти C_n^k – біноміальними коефіцієнтами.

З формули бінома Ньютона випливає (досить покласти $a = 1$, $b = 1$):

$$C_k^0 + C_n^1 + \cdots + C_n^{n-1} + C_n^n = \sum_{k=0}^{n} C_n^k = 2^n$$

Це означає, що в n-елементній множині є всього 2^n невпорядкованих підмножин (разом з порожньою 0-елементною і n-елементною підмножинами).

Розглянемо деякі властивості чисел C_n^m

1°. $\sum\limits_{i=0}^{k} C_n^i C_m^{k-i} = C_{n+m}^k$.

Справді, якщо розглянути біном $(1+x)^{m+n} = (1+x)^m (1+x)^n$, то коефіцієнтом біля x^k у лівій частині буде C_{n+m}^k, а в правій частині $C_n^0 C_m^k + C_n^1 C_m^{k-1} + C_n^2 C_m^{k-2} + \cdots + C_n^k C_m^0$.

Цю рівність можна обґрунтувати й так. Якщо з $(n+m)$ елементної множини, в якій є n елементів з певною ознакою і m елементів без такої ознаки, вибирати k елементні підмножини, то всіх таких підмножин буде C_{n+m}^{k}, причому підмножин таких, що i елементів будуть з вказаною ознакою, а решта $k-i$ будуть без такої ознаки, буде $C_n^i C_m^{k-i}$, оскільки для утворення таких підмножин досить з n-елементної множини елементів із вказаною ознакою вибрати i-елементну підмножину і до неї приєднати $(k-i)$ елементну підмножину, вибрану з m-елементної множини елементів без вказаної ознаки. Очевидно, що першу підмножину можна вибрати C_n^i способами, і до кожної такої підмножини можна приєднати C_m^{k-i} підмножин із другої множини. Оскільки серед C_{n+m}^{k} підмножин можуть бути такі, що підмножина елементів із вказаною ознакою порожня або містить один елемент, або два і т.д., то $\sum_{i=0}^{k} C_n^i C_m^{k-i} = C_{n+m}^k$.

2°. $C_{n+1}^k = C_n^k + C_n^{k-1}$.

Справді, якщо один з елементів в $(n+1)$-елементній множині з певною ознакою, а решта n елементів без такої ознаки, то всі можливі k-елементні підмножини можна поділити на дві групи. До першої групи віднести всі k-елементні множини, в яких немає вказаного елемента. Число таких підмножин C_n^k До другої групи віднести всі k-елементні підмножини, в яких є вказаний елемент з такою ознакою і ще $k-1$ елементів без ознаки. Число таких підмножин дорівнює C_n^{k-1}. Отже , число всіх k-елементних підмножин дорівнює $C_n^k + C_n^{k-1}$. Число всіх k-елементних підмножин в $(n+1)$-елементній множині дорівнює C_{n+1}^k. Звідси $C_{n+1}^k = C_n^k + C_n^{k-1}$.

Числа C_n^m за різних натуральних n і m зручно розмістити у вигляді таблиці.

$$
\begin{array}{ccccccc}
 & & & C_0^0 & & & \\
 & & C_1^0 & & C_1^1 & & \\
 & C_2^0 & & C_2^1 & & C_2^2 & \\
C_3^0 & & C_3^1 & & C_3^2 & & C_3^3 \\
\end{array}
$$

…………..…………………………

Обчисливши числа C_n^k, $n=0, 1, ...,$ $k=0, 1, ...n$, дістанемо таблицю

$$\begin{array}{ccccccccccc} &&&&&1&&&&& \\ &&&&1&&1&&&& \\ &&&1&&2&&1&&& \\ &&1&&3&&3&&1&& \\ &1&&4&&6&&4&&1& \\ 1&&5&&10&&10&&5&&1 \end{array}$$

Цю таблицю називають *трикутником Паскаля*, хоч вона була частково відома в Індії ще в II ст. до н. е. За $n=8$ вона наводиться в працях китайського математика Чжу Шіцзе (XII-XIII ст.); за $n=9$ - у ал-Каші (XIV—XV ст.). Зустрічається вона також у працях європейських вчених П. Ахіана (1527), Е Л. Штіфеля (1544). Найбільшого поширення таблиця набула в XVII ст. у зв’язку з роботами Б. Паскаля.

Розглянемо ще такий приклад.

В автобусі, де є 20 місць, одна за однією займають місця три групи пасажирів, причому в першій групі є вісім пасажирів, у другій групі - сім пасажирів, в третій - п’ять пасажирів.

Скільки може бути варіантів знайти місця в автобусі для цих трьох груп пасажирів?

Очевидно, є C_{20}^8 варіантів вибрати місця для першої групи пасажирів (порядок, в якому зайняті 8 місць, несуттєвий). Для другої групи із семи пасажирів може бути C_{12}^7 варіантів обрати 7 місць із 12, що залишилися, за кожного варіанту вибору восьми місць для першої групи пасажирів. Для третьої групи пасажирів за кожного із $C_{20}^8 C_{12}^7$ варіантів, якими перші дві групи вибрали 15 місць, є C_5^5 варіантів обрати 5 місць із п’яти, що залишились.

Таким чином, враховуючи правило, множення, дістаємо, що є $C_{20}^8 C_{12}^7 C_5^5$ варіантів обрати місця для вказаних трьох груп пасажирів із наявних 20 місць, тобто є

$$C_{20}^8 C_{12}^7 C_5^5 = \frac{20!}{8!*12!} * \frac{12!}{7!*5!} * \frac{5!}{5!*0!} = 99768240$$

варіантів розмістити три групи із 8, 7 та 5 пасажирів на 20 місцях.

В загальному випадку n-елементну множину G можна поділити на k підмножин так, щоб в першій з них було m_1 елементів, в другій - m_2 елементів і т.д., в k-й - m_k елементів, причому $m_1+m_2+...+m_k=n$, і в різних підмножинах не було спільних елементів, $C_n(m_1, m_2, ..., m_k)$ способами, де

$$C_n(m_1, m_2, ..., m_k) = C_n^{m_1} C_{n-m_1}^{m_2} C_{n-m_1-m_2}^{m_3} ... C_{n-m_1-m_2-...-m_{k-1}}^{m_k} =$$

$$= \frac{n!}{m_1!(n-m_1)!} \cdot \frac{(n-m_1)!}{m_2!(n-m_1-m_2)!} ... \frac{(n-m_1-m_2-...-m_{k-1})!}{m_k!(n-m_1-m_2-...-m_k)!}.$$

Тобто

$$C_n(m_1, m_2, ..., m_k) = \frac{n!}{m_1! m_2! \ldots m_k!}.$$

Розглянуті n-елементні сполуки називають перестановками з повтореннями, оскільки елемент першої групи з m_1 елементів повторюється m_1 разів, елемент другої групи повторюється m_2 разів і т. д. В даному разі вважається, що в кожній групі всі елементи однакові або мають якусь спільну ознаку. Наприклад, в першій групі m_1 літер «а», в другій m_2 літер «б» і т. д.

Числа $C_n(m_1, m_2, ..., m_k)$ є коефіцієнтами узагальненої формули бінома Ньютона

$$(a_1 + a_2 + \ldots + a_k)^m =$$

$$= \sum_{m_1+m_2+\ldots+m_k=n} \frac{n!}{m_1! m_2! \ldots m_k!} a_1^{m_1} a_2^{m2} \ldots a_k^{m_k},$$

де сума поширюється на всі можливі набори цілих невід'ємних чисел $m_1, m_2, ..., m_k$ таких, що $m_1 + m_2 + \ldots + m_k = n$. Останню формулу називають поліноміальною.

Слід зазначити, що стосовно перестановок з повтореннями є суттєвим те, що елемент першої групи (a_1) повторюється m_1 разів, елемент другої групи (a_2) - m_2 разів і т. д., елемент k-ї групи (a_k) – m_k разів.

Якщо множину всіх m-елементних сполук, складених із елементів деякої n-елементної множини, поділити на класи еквівалентності, відносячи до одного класу всі сполуки однакового складу, то такі класи еквівалентності називають комбінаціями з повтореннями з n елементів по m елементів. Іншими словами, комбінаціями з n елементів по m елементів з повтореннями називають групи із m елементів, причому ці елементи належать до одного з n типів. Число комбінацій з n елементів по m елементів з повтореннями обчислюється за формулою

$$f_n^m = C_{n+m-1}^{n-1} = \frac{(n+m-1)!}{(n-1)! m!}.$$

Справді кожна комбінація повністю визначається, якщо сказано, скільки елементів кожного з n типів до неї входить. Поставимо у відповідність кожній комбінації набір символів 1 і «:» (двокрапка) таким чином. Напишемо підряд стільки одиниць, скільки елементів першого типу входить до комбінації, потім поставимо двокрапку і після неї напишемо стільки одиниць, скільки елементів другого типу входить до комбінації і т. д. Двокрапка може бути і першим символом, якщо немає елементів першого типу, кілька двокрапок можуть бути записані підряд, якщо до комбінації не входять елементи відповідних типів. Усього буде $n-1$ двокрапок, розставлених між m елементами. Отже треба серед $m+n-1$ місць вибрати $n-1$ місць для двокрапок. Число таких варіантів, як відомо, дорівнює $C_{m+n-1}^{n-1} = C_{m+n-1}^{m}$.

Приклад. Скількома способами можна розмістити шість однакових папок у трьох різних ящиках письмового стола, якщо в кожен ящик можна вмістити всі папки?

Тут розглядаються шестиелементні сполуки з елементів трьох "типів": $Я_1, Я_2, Я_3$. Порядок розміщення папок у ящиках несуттєвий. Тому шукане число способів дорівнює числу комбінацій з повтореннями з трьох елементів по 6, тобто $f_3^6 = C_8^6 = 28$. Зазначимо, що за великих n обчислення значення $n!$ стає проблематичним. У таких випадках використовують асимптотичну формулу Стірлінга, згідно з якою

$$n! \approx n^n e^{-n} \sqrt{2\pi n}\ .$$

Відносна похибка обчислень за цією формулою лежить у межах від $1 - n^{-\frac{1}{12n+1}}$ до $1 - n^{-\frac{1}{12n}}$.

Додаток 2. Таблиця значень функції Лапласа: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$:

x	0	1	2	3	4	5	6	7	8	9
0,0	0.00000	0,00399	0,00798	0,01197	0,01595	0,01994	0,02392	0,02790	0,03188	0,03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46582	46638	46712	46784	46856	46926	46995	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3,0	49865									
3,1	49903									
3,2	49931									
3,3	49952									
3,4	49986									
3,5	49977									
3,6	49984									
3,7	49989									
3,8	49993									
3,9	49995									
4,0	49996									

Додаток 3. Значення функції $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0068	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	2058	0,0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0005	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Додаток 4. Значення $\dfrac{a_n^m e^{-a_n}}{m!}$

m \ a_n	0,1	0,2	0,3	0,4	0,5	0,6
0	0,904837	0,818731	0,740818	0,670320	0,606531	0,548812
1	904084	163746	222245	268128	303265	329287
2	004524	016375	033337	053626	075816	098786
3	000151	001091	003334	007150	012636	019757
4	000004	000055	000250	000715	001580	002964
5		000002	000015	000057	000158	000356
6			000001	000004	000013	000035
7					000001	000003
m \ a_n	0,7	0,8	0,9	1,0	2,0	3,0
0	0,496585	0,449329	0,406570	0,367879	0,135335	0,049787
1	347610	359463	365913	367879	270671	149361
2	121663	143785	164661	183940	270671	224042
3	028388	038343	049398	061313	180447	224042
4	004968	007669	011115	015328	090224	168031
5	000695	001227	002001	003066	036089	100819
6	000081	000164	000300	000511	012030	050409
7	000008	000019	000039	000073	003437	021604
8		000002	000004	000009	000859	008101
9				000001	000191	002701
10					000038	000810
11					000007	000221
12					000001	000055
13						000013
14						000002
15						000001

Додаток 5. Значення $\chi^2_{кр}$, за яких задовільняється рівність $P(\chi^2_r \le \chi^2_{кр}) = \alpha$, залежно від r і α

r \ α	0.01	0.02	0.05	0.1	0.2	0.3	0.5	0.7	0.8	0.9	0.95	0.98	0.99	0.999
1	0.000	0.001	0.004	0.016	0.064	0.148	0.455	1.074	1.642	2.71	3.84	5.41	6.64	10.83
2	0.020	0.040	0.103	0.211	0.446	0.713	1.386	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	0.115	0.185	0.352	0.584	1.005	1.424	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	0.297	0.429	0.711	1.064	1.649	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	0.554	0.752	1.145	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.05
6	0.872	1.134	1.635	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.5
7	1.239	1.564	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.3
8	1.646	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.1	26.1
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.7	27.9
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.2	23.2	29.6
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.6	24.7	31.3
12	3.57	4.18	5.23	6.30	7.81	9.03	11.34	14.01	15.81	18.55	21.0	24.1	26.2	32.9
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.4	22.5	27.7	34.6
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.1	23.7	26.9	29.1	36.1
15	5.23	5.98	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.3	25.0	28.3	30.6	37.7
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.5	23.5	26.3	29.6	32.0	39.3
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.6	24.8	27.6	31.0	33.4	40.8
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.6	22.8	26.0	28.9	32.3	34.8	42.3
19	7.63	8.57	10.11	11.65	13.62	15.35	18.34	21.7	23.9	27.2	30.1	33.7	36.2	43.8
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.8	25.0	28.4	31.4	35.0	37.6	45.3
21	8.90	9.92	11.59	13.24	15.44	17.18	20.3	23.9	26.2	29.6	32.7	36.3	38.9	46.8
22	9.54	10.60	12.34	14.04	16.31	18.10	21.3	24.9	27.3	30.8	33.9	37.7	40.3	78.3
23	10.20	11.29	13.09	14.85	17.19	19.02	22.3	26.0	28.4	32.0	35.2	39.0	41.6	49.7
24	10.86	11.99	13.86	15.66	18.06	19.94	23.3	27.1	29.6	33.2	36.4	40.3	43.0	51.2
25	11.62	12.70	14.61	16.47	18.94	20.9	24.3	28.2	30.7	34.4	37.7	41.7	44.3	52.6
26	12.20	13.41	15.38	17.29	19.82	21.8	25.3	29.2	31.8	35.6	38.9	42.9	45.6	54.1
27	12.88	14.12	16.15	18.11	20.07	22.7	26.3	30.3	32.9	36.7	40.1	44.1	47.0	55.5
28	13.56	14.85	16.93	18.94	21.6	23.6	27.3	31.4	34.0	37.9	41.3	45.4	48.3	56.9
29	14.26	15.57	17.71	19.77	22.5	24.6	28.3	32.5	35.1	39.1	42.6	46.7	49.6	58.3
30	14.95	16.31	18.49	20.6	23.4	25.5	29.3	33.5	36.2	40.3	43.8	48.0	50.9	59.7

Додаток 6. Значення t_β, за яких задовільняється рівність $P(\,|T_n|<t_\beta\,)=2\int_0^{t_\beta} s_{n-1}(t)\,dt=\beta$, залежно від β і $n-1$

$n-1$ \ β	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99	0,999
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,08	6,31	12,71	31,80	63,70	636,6
2	142	289	445	617	0,816	1,061	1,336	1,886	2,92	4,30	6,96	9,92	31,60
3	137	277	424	584	765	0,978	1,250	1,638	2,35	3,18	4,54	5,84	12,94
4	134	271	414	569	741	941	1,190	1,533	2,13	2,77	3,75	4,60	8,61
5	132	267	408	559	727	920	1,156	1,476	2,02	2,57	3,36	4,03	6,86
6	131	265	404	553	718	906	1,134	1,440	1,943	2,45	3,14	3,71	5,96
7	130	263	402	549	711	896	1,119	1,415	1,895	2,36	3,00	3,50	5,40
8	130	262	399	546	706	889	1,108	1,397	1,860	2,31	2,90	3,36	5,04
9	129	261	398	543	703	883	1,100	1,383	1,833	2,26	2,82	3,25	4,78
10	129	260	397	542	700	879	1,093	1,372	1,812	2,23	2,76	3,17	4,59
11	129	260	396	540	697	876	1,088	1,363	1,796	2,20	2,72	3,11	4,49
12	128	259	395	539	695	873	1,083	1,356	1,782	2,18	2,68	3,06	4,32
13	128	259	394	538	694	870	1,079	1,350	1,771	2,16	2,65	3,01	4,22
14	128	258	393	537	692	868	1,076	1,345	1,761	2,14	2,62	2,98	4,14
15	128	258	393	536	691	866	1,074	1,341	1,753	2,13	2,60	2,95	4,07
16	128	258	392	535	690	865	1,071	1,337	1,746	2,12	2,58	2,92	4,02
17	128	257	392	534	689	863	1,069	1,333	1,740	2,11	2,57	2,90	3,96
18	127	257	392	534	689	862	1,067	1,330	1,734	2,10	2,55	2,88	3,92
19	127	257	391	533	688	861	1,066	1,328	1,729	2,09	2,54	2,86	3,88
20	127	257	391	533	687	860	1,064	1,325	1,725	2,09	2,53	2,84	3,85
21	127	257	391	532	686	859	1,063	1,323	1,721	2,08	2,52	2,83	3,82
22	127	256	390	532	686	858	1,061	1,321	1,717	2,07	2,51	2,82	3,79
23	127	256	390	532	685	858	1,060	1,319	1,714	2,07	2,50	2,81	3,77
24	127	256	390	531	685	857	1,059	1,318	1,711	2,06	2,49	2,80	3,74
25	127	256	390	531	684	856	1,058	1,316	1,708	2,06	2,48	2,79	3,72
26	127	256	390	531	684	856	1,058	1,315	1,706	2,06	2,48	2,78	3,71
27	127	256	389	531	684	855	1,057	1,314	1,703	2,05	2,47	2,77	3,69
28	127	256	389	530	683	855	1,056	1,313	1,701	2,05	2,47	2,76	3,67
29	127	256	389	530	683	854	1,055	1,311	1,699	2,04	2,46	2,76	3,66
30	127	256	389	530	683	854	1,055	1,310	1,697	2,04	2,46	2,75	3,65
40	126	255	388	529	681	851	1,050	1,303	1,684	2,02	2,42	2,70	3,55
60	126	254	387	527	679	848	1,046	1,296	1,671	2,00	2,39	2,66	3,46
120	126	254	386	526	677	845	1,041	1,289	1,658	1,980	2,36	2,62	3,37

Додаток 7. Значення $F_{n_1,n_2,p}$, які відповідні ймовірності $p = P\{F_{n_1,n_2} > F_{n_1,n_2,p}\}$, коли на множині значень випадкової величини F_{n_1,n_2} має місце F-розподіл з n_1 і n_2 ступенями вільності

$p = 0{,}05$										
n_2 \ n_1	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,21	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,56	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,52
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Додаток 8. Таблиця значень $q = q(\beta, n)$

n \ β	0,95	0,99	0,999
5	1,37	2,67	5,64
6	1,09	2,01	3,88
7	0,92	1,62	2,98
8	0,80	1,38	2,42
9	0,71	1,20	2,06
10	0,65	1,08	1,80
11	0,59	0,98	1,60
12	0,55	0,90	1,45
13	0,52	0,83	1,33
14	0,48	0,78	1,23
15	0,46	0,73	1,15
16	0,44	0,70	1,07
17	0,42	0,66	1,01
18	0,40	0,63	0,96
19	0,39	0,60	0,92
20	0,37	0,58	0,88
25	0,32	0,49	0,73
30	0,28	0,43	0,63
35	0,26	0,38	0,56
40	0,24	0,35	0,50
45	0,22	0,32	0,46
50	0,21	0,30	0,43
60	0,188	0,269	0,38
70	0,174	0,245	0,34
80	0,161	0,226	0,31
90	0,151	0,211	0,29
100	0,143	0,198	0,27
150	0,115	0,160	0,221
200	0,099	0,136	0,185
250	0,089	0,120	0,162

Література

1. ***Боровков А.А***. Теория вероятностей. - М.: Наука, 1976. – 288 с.
2. ***Вулих Б. З***. Краткий курс теории функций вещественной переменной. – М.: Наука, 1965. – 304 с.
3. ***Гихман И. И., Скороход А. В., Ядренко М. И.*** Теория вероятностей и математическая статистика. – К.: Вища шк., 1988. – 440 с.
4. ***Гнеденко Б. В.*** Курс теории вероятностей. – М.: Наука, 1988. – 448 с.
5. ***Жалдак М.І., Кузьміна Н.М., Михалін Г.О.*** Теорія імовірностей і математична статистика. Збірник вправ і задач. К.: НПУ імені М.П. Драгоманова. 2019. – 842с.
6. ***Кемени Дж., Снелл Дж.*** Введение в конечную математику. – М.: Изд-во иностр. лит., 1963. – 486 с.
7. ***Коваленко И. Н., Филиппова А. А.*** Теория вероятностей и математическая статистика. – М.: Высш. Шк., 1973. – 368 с.
8. ***Колемаев В. А., Староверов О. В., Турундаевский В. Б.*** Теория вероятностей и математическая статистика. – М.: Высш. шк., 1991. – 400 с.
9. ***Колмогоров А. Н., Фомин С. В.*** Элементы теории функций и функционального анализа. – М.: Наука, 1968. – 496 с.
10. ***Колмогоров А.Н.*** Основные понятия теории вероятностей. - М.: Наука, 1974. – 120 с.
11. ***Крамер Г.*** Математические методы статистики. – М.: Мир, 1975. – 648 с.
12. ***Лоэв М.*** Теория вероятностей. – М.: Изд-во иностр. лит., 1962. – 720 с.
13. ***Майстров Д.Е.*** Теория вероятностей. Исторический очерк. - М.: Наука, 1967. – 320 с.
14. ***Погорєлов О.В.*** Геометрія: Планіметрія: Підруч. для 7-9 кл. загальноосвіт. навч. закл. – 7-е вид. – К.: Школяр, 2004. – 240 с.
15. ***Погорєлов О.В.*** Геометрія: Стереометрія: Підруч. для 10-11 кл. серед. шк. – 6-те вид. – К.: Освіта, 2001. – 128 с.
16. ***Сборник*** задач по теории вероятностей, математической статистике и теории случайных функций / Под ред. А. А. Свешникова. – М.: Наука, 1970. – 556 с.
17. ***Толстов Г.П.*** Мера и интеграл. – М.: Наука, 1976. – 392 с.
18. ***Уилкс С.*** Математическая статистика. – М.: Наука, 1967. – 632 с.

19. ***Феллер В.*** Введение в теорию вероятностей и ее приложения: В 2 т. – М.: Мир, 1964. – Т. 1. – 499 с.; 1967. – Т. 2. – 752 с.
20. ***Чистяков В. П.*** Курс теории вероятностей. – М.: Наука, 1987. – 240 с.
21. ***Ширяев А. Н.*** Вероятность. – М.: Наука, 1989. – 640 с.
22. ***Шметтерер Л.*** Введение в математическую статистику. – М., Наука, 1976. – 520 с.

Предметний та іменний покажчик

Зміст

РОЗДІЛ III. Ймовірності

Навчальне видання

Мирослав Іванович Жалдак
Наталія Миколаївна Кузьміна
Геннадій Олександрович Михалін

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Підручник для студентів фізико-математичних та інформатичних спеціальностей педагогічних університетів
Видання четверте, доповнене

За загальною редакцією *М.І. Жалдака*

Комп'ютерний набір *Н.П. Франчук, Є.В. Малюх*

Підписано до друку 26.03.2019 р. Формат 60x84/16.
Папір офісний. Гарнітура Times New Roman.
Ум. др. арк. 43,59. Обл.-вид. арк. 32,77
Наклад 1000 прим. Зам. № 005.
Віддруковано з оригіналів.

Видавництво Національного педагогічного університету імені М.П. Драгоманова. 01601, м. Київ-30, вул. Пирогова, 9
Свідоцтво про реєстрацію ДК № 1101 від 29.10.2002. (044) 234-75-87
Віддруковано в друкарні Національного педагогічного університету імені М.П. Драгоманова (044) 239-30-26