М.І. Жалдак, Г.О. Михалін
НПУ ім. М.П.Драгоманова

Про коректність введення понять “випадкова подія”, “ймовірність” та “випадкова величина” у шкільному курсі математики

1. Вступ. Вже біля 10 років однією з тем шкільного курсу математики є “Початки теорії ймовірностей”. Методиці навчання цієї теми присвячена досить велика кількість публікацій, в тому числі навчальних посібників та підручників. Пропоновані підходи найрізноманітніші: від того, що процес навчання повинен відповідати принципу історизму (учні повинні відчути, осягнути тисячорічний процес становлення теорії ймовірностей) і до того, що шкільний курс теорії ймовірностей повинен відповідати сучасному баченню цієї теорії.

На жаль, принцип історизму у багатьох шкільних (і не тільки) підручниках і посібниках зводиться до подання навчального матеріалу так, як це робили років 200 тому, аргументуючи це авторитетом великих математиків Я. Бернуллі, Л. Ейлера, П. Лапласа, П.Л. Чебишова, А.А. Маркова, А. М.Колмогорова та багатьох інших творців теорії ймовірностей.

Дана стаття присвячена аналізу коректності введення у шкільному курсі математики найважливіших стохастичних понять: “випадкова подія”, “ймовірність випадкової події”, “випадкова величина”.

Будемо виходити з того, що кожне тлумачення (зокрема й означення) певного поняття (об’єкта) або відношення між поняттями (об’єктами) – це визначення його характеристичних властивостей за допомогою вже розтлумачених понять та відношень (це можуть бути основні, тобто неозначувані, поняття та відношення або вже означені).

Означення певного об’єкта (поняття) можна розуміти як його ідентифікацію, тобто встановлення його істотних (найістотніших) ознак.

Введення означення відбувається шляхом розкриття і формулювання за певними правилами певного змісту у формі суджень. Такі шляхи можуть бути пов’язані з вказуванням:

- найближчих родових і видових ознак;
- правил утворення поняття (об’єкта);
- операцій вимірювання певної величини.

Виходячи з такого тлумачення сутності поняття “означення”, проаналізуємо, як у шкільних (та інших) посібниках вводяться поняття “випадкова подія”, “ймовірність події” та “випадкова величина”. При цьому не конкретизуватимемо, у яких саме підручниках та посібниках наведено відповідне тлумачення чи означення, оскільки кожен вчитель математики здатен сам вирішити, як стосується наведений аналіз тих джерел, якими користується він та його учні.

2. Поняття випадкової події. У багатьох навчальних посібниках та підручниках пишуть, що “поняття події належить до неозначуваних (основних) понять теорії ймовірностей; під подією розуміють кожне явище, про яке можна сказати, що воно відбувається або не відбувається”. Після таких слів у деяких шкільних підручниках вводяться: “**Означення.** Випадковою подією називають подію, яка може відбутися або не відбутися під час здійснення певного випробування… Під випробуванням розуміють ті умови, в результаті яких відбувається подія”. “**Означення.** Масовими називають однорідні події, що спостерігаються за певних умов, які можуть бути відтворені необмежену кількість разів. Масовими вважаються і ті події, для яких відповідні випробування не можна відтворити, але є можливість спостерігати аналогічні випробування у великій кількості”. “**Означення.** Множина подій утворює повну групу подій, якщо внаслідок кожного випробування хоч одна з цих подій напевне виконується”. “**Означення.** Події A_1, A_2, …, A_n називаються попарно несумісними у даному випробуванні, якщо ніякі дві з них не можуть відбуватися разом”. “Поняття рівноможливих подій первісне, неозначуване. Під рівноможливими розуміють такі події, кожна з яких не має ніяких переваг у появі частіше за іншу під час багаторазових випробувань, що проводяться за однакових умов”. “Багато теоретичних тверджень теорії ймовірностей будується на розгляді множини подій, які мають всі три властивості: утворюють повну групу подій, є несумісними і рівноможливими. Події такої множини називають елементарними”.

Стосовно наведеного виникає багато запитань:

- Чим відрізняється первісне поняття “подія” від означуваного поняття “випадкова подія”?
- Що таке однорідні події та певні умови? Чи насправді певні умови можуть бути відтвореними необмежену кількість разів? І що таке необмежена кількість разів?
- Чи може якась подія повної групи подій ніколи не виконуватися або виконуватися в результаті кожного випробування?
- У даному випробуванні несумісні події не можуть відбутися разом. А в іншому випробуванні – можуть?
- Чи можна з рівноможливих і попарно несумісних подій, що утворюють повну групу, утворити нову повну групу подій, рівноможливих і попарно несумісних? Події цієї групи також називаються елементарними? Чим одні елементарні події відрізняються від інших?

2.1. Перелік запитань, що виникають при читанні багатьох навчальних посібників, можна продовжувати досить довго. Головним недоліком наведених тлумачень та означень випадкових подій і відношень між ними є те, що всі ці тлумачення і означення важко назвати *математичними моделями* реальних випадкових подій, оскільки усі вони не є сукупностями відомих учням математичних понять і тверджень, що характеризують відповідні реальні випадкові події.

Побудову шкільного курсу теорії ймовірностей на основі таких тлумачень випадкових подій можна уподобити побудові шкільного курсу геометрії, у якому вивчали б не математичні моделі реальних об'єктів, а самі об'єкти, наприклад, не геометричне поняття кулі, а земну кулю, дощову краплю, футбольний м'яч та багато інших реальних об'єктів, математичною моделлю яких є поняття "куля".

2.2. Разом з тим, враховуючи сучасні підходи до побудови теорії ймовірностей, можна легко, навіть у рамках шкільного курсу математики будувати *математичні моделі* реальних випадкових подій, які називають *ймовірнісними моделями*. Як конкретно будувати такі моделі, про це йдеться у багатьох статтях і посібниках. Основна ідея полягає в тому, що побудову ймовірнісної моделі починають з випадкового (стохастичного) експерименту – спостереження за певним реальним випадковим явищем. Можливі результати цього експерименту називають *елементарними подіями*, а *множину усіх таких результатів* (елементарних подій) позначають Ω і називають *простором елементарних подій*. Ця множина Ω є *початковою моделлю* згаданого експерименту або реального випадкового явища. В залежності від поставленої мети початкова модель може неодноразово уточнюватися, проте кожного разу вона залишається математичною, оскільки пов'язана з математичними поняттями "множина" та "елемент множини" і відношеннями ("належності" та "неналежності") між цими поняттями.

В рамках початкової моделі Ω можна тлумачити поняття *випадкової події* A як певну підмножину елементарних подій з множини Ω. Випробування (окреме проведення експерименту) можна тлумачити як вибір навмання окремого елемента з деякої множини елементів Ω, кожному з яких поставлено у взаємно однозначну відповідність елементарну подію E з простору Ω. Якщо при цьому виявиться, що елементарна подія $E \in A$ (E сприяє події A), то це й означає відбування події A у даному випробуванні.

2.3. Не завжди зручно, а іноді і неможливо (див. наприклад, [1]) вважати кожну підмножину $A \subset \Omega$ випадковою подією. Тому бажано у процесі побудови ймовірнісної моделі вказати *характеристичні* або *основні властивості*, що цілком визначають сукупність S *випадкових подій*:

1_s. $\Omega \in S$, тобто вірогідна подія завжди є подією.

2_s. Якщо A є подією з простору S, то S містить і подію $\overline{A} = \Omega \setminus A$, протилежну до A ($\overline{A} = \Omega \setminus A$ часто позначають $\Omega - A$).

3_s. Якщо A_k, $k \in N$, є подіями з простору S, то S містить і суму цих подій, тобто $\bigcup_k A_k \in S$ ($\bigcup_k A_k$ часто позначають також $\sum_k A_k$).

Доцільно підкреслити, що умови 1_s-3_s нагадують умови, які задовольняють фігури, що мають довжину, площу чи об'єм. Так, якщо S – це сукупність тих частин відрізка $[0;1]$, що мають довжину, то $[0;1] \in S$; для кожної частини $A \in S$, частина $\overline{A} = [0,1] \setminus A$ також має довжину, а тому $\overline{A} \in S$; якщо частини A_k мають довжину, то і $\bigcup_k A_k$ також має довжину.

Отже, умови 1_s-3_s – це правила побудови сукупності випадкових подій. Виходячи з однієї і тієї ж сукупності Ω елементарних подій, можна побудувати багато різних сукупностей S випадкових подій (якщо Ω містить більше одного елемента), що задовольнятимуть умови 1_s-3_s. Серед таких сукупностей є така, коли кожна підмножина множини Ω результатів випадкового експерименту вважається випадковою подією; проте не завжди останній варіант можливий для побудови ймовірнісної моделі, а досить часто хоч і можливий, проте недоцільний (див., наприклад, [2]).

Пару (Ω, S), тобто простір Ω елементарних подій разом з простором подій S, іноді називають *уточненою моделлю відповідного стохастичного експерименту*. Можливість вибору уточненої моделі серед усіх існуючих забезпечує гнучкість та широту застосувань таких моделей.

3. Поняття ймовірності випадкової події. Досить велика кількість навчальних посібників з теорії ймовірностей присвячена побудові цієї теорії на основі так званого "класичного означення ймовірності".

3.1. Наведемо один з найгірших прикладів такого означення: "Ймовірністю випадкової події називається відношення кількості подій, які сприяють цій події, до кількості всіх рівноможливих несумісних подій, які утворюють повну групу подій під час певного випробування".

Зауважимо, що перед тим, як ввести наведене "означення", у відповідному шкільному посібнику формулюють довгу низку тлумачень та "означень", пов'язаних з поняттям випадкової події (див. початок пункту **2**), які на думку авторів посібника допомагають зрозуміти наведене означення ймовірності, проте насправді тільки заплутують сутність поняття "ймовірність випадкової події", оскільки залишаються без відповіді такі запитання (до наведеного означення ймовірності):

- Що таке подія, яка сприяє цій події?
- Скільки існує таких подій, що сприяють даній випадковій події, і як їх визначити?
- Яке відношення мають всі рівноможливі несумісні події, що утворюють повну групу, до події, ймовірність якої визначається?
- Чи завжди це означення можна застосувати?

В деяких посібниках "класичне означення" ймовірності $P(A)$ події A вводять акуратніше, вважаючи, що усі елементарні події простору Ω є рівноможливими, а $P(A) = \frac{m}{n}$, де m – кількість елементарних подій, що сприяють події A, а n – кількість усіх елементарних подій простору Ω.

Окрім того, що залишається невизначеним поняття рівноможливості елементарних подій, виникає питання, як знайти ймовірність події A, коли множина елементарних подій нескінченна. Також виявляється, що далеко не для кожного випадкового експерименту, для якого простір елементарних подій Ω скінченний, можна вважати елементарні події рівноможливими (наприклад при біноміальному розподілі ймовірностей). Отже "класичне означення" ймовірності не розкриває сутності поняття ймовірності випадкової події і тому некоректне.

Вагомим аргументом використання "класичного означення" ймовірності автори багатьох посібників вважають те, що цим означенням користувалися творці теорії ймовірностей – великі математики. Проте справа в тому, що творцям математичних (і не тільки) теорій часто не потрібні строгі означення. Більше того, народження принципово нової теорії навіть неможливе в якихось строгих рамках. Великі математики тому й вважаються великими, що здатні побачити і побудувати струнку математичну теорію серед хаосу різноманітних математичних фактів, на перший погляд, ніяк не пов'язаних між собою, здобутих їхніми попередниками протягом багатьох років, іноді століть і навіть тисячоліть.

Стрункість теорії досягається лише на заключних етапах її розвитку. Тоді вона стає доступною для вивчення не тільки студентами університетів, а й учнями середньої школи. Саме учням потрібні чіткі і зрозумілі означення, до яких класичне означення аж ніяк не відноситься.

3.2. Розуміючи (або ні) недосконалість "класичного означення" ймовірності, автори багатьох посібників наводять вже традиційно ще два означення ймовірності: статистичне та геометричне, які часто ніяк не пов'язують ні одне з одним, ні з класичним означенням ймовірності, а разом з тим вважають, що доведення різних властивостей ймовірностей достатньо провести лише для "класичних ймовірностей". Тому таке навчання теорії ймовірностей не лише не розвиває мислення студентів та учнів, але часто навіть шкідливе.

Проаналізуємо "*означення статистичної ймовірності*", що наводиться в багатьох посібниках: для n проведених випробувань підраховують кількість $m = m_n(A)$ тих випробувань, у яких подія A відбулася, і число $P_n^*(A) = \dfrac{m}{n}$ називають *відносною частотою події* A, а число $P(A)$, навколо якого *скупчуються* (стабілізуються) відносні частоти $P_n^*(A)$ при великій кількості n випробувань, називають *ймовірністю*, або *статистичною ймовірністю* події A. Іноді після таких слів роблять висновок, що $P(A) = \lim\limits_{n\to\infty} P_n^*(A)$.

Такі означення і ймовірності, і статистичної ймовірності не витримують жодної критики.

Стосовно наведених означень виникають запитання:

- Як розуміти слова "$P_n^*(A)$ скупчуються навколо $P(A)$"? Як перевірити це скупчування?
- Якщо "скупчуваність" означає існування границі $\lim\limits_{n\to\infty} P_n^*(A)$, то звідки випливає існування цієї границі?
- Чи якось пов'язані між собою скупчуваності $P_n^*(A)$ і $P_n^*(B)$ для двох різних подій та для більшої кількості подій?

Без відповідей на поставлені запитання наведене означення насправді не є означенням, а лише певним інтуїтивним тлумаченням ймовірності.

Окрім того можна навести приклади, коли відносні частоти $P_n^*(A)$ із збільшенням n не стабілізуються або ж стабілізуються, проте не виключено, що біля числа, далекого від ймовірності події A.

Нехай, наприклад, є кубик (можливо неоднорідний), на двох гранях якого нанесені цифри 0, а на інших чотирьох – цифри 1. Тоді не виключено, що при як завгодно великому числі випробувань – підкидань одного і того самого кубика буде отримано наступні серії наслідків таких випробувань:

1) $000...00...$ (на верхній грані кубика весь час випадає цифра нуль) – $P_n^*(\{0\})$ стабілізується біля значення 1;

2) $111...11...$ на верхній грані кубика весь час випадає цифра 1 – $P_n^*(\{0\})$ стабілізується біля значення 0;

3) $010101...0101...$(на верхній грані кубика почергово випадають цифри 0 і 1) – $P_n^*(\{0\})$ стабілізується біля значення 0,5;

4) $011000111111...$(на верхній грані кубика цифри 0 і 1 випадають у відповідності з правилами: 1-й раз випадає 0, потім двічі 1, далі набір цифр довжиною, що дорівнює кількості всіх попередніх, причому всі цифри такого набору однакові і протилежні до останньої в попередній серії, тобто, 1-ша цифра 0, потім 2 цифри 1, потім три цифри 0, потім 6 цифр 1, потім 12 цифр 0 і т.д. Тоді $P_n^*(\{0\})$ не стабілізується біля якогось числа і змінюється із збільшенням n між значеннями $\frac{1}{3}$ та $\frac{2}{3}$, коливаючись між ними і збільшуючись від $\frac{1}{3}$ до $\frac{2}{3}$, потім зменшуючись від $\frac{2}{3}$ до $\frac{1}{3}$, потім знову збільшуючись від $\frac{1}{3}$ до $\frac{2}{3}$ і т. д.

Такі приклади можна продовжити. Це говорить про те, що навіть у випадку, коли значення $P_n^*(A)$ із збільшенням n скупчуються (стабілізуються) біля якогось числа, не можна стверджувати, що в інших серіях випробувань $P_n^*(A)$ із збільшенням n будуть стабілізуватися біля того самого числа чи взагалі стабілізуватися. Звідси випливає некоректність статистичного означення ймовірності.

3.3. Разом з тим наведене "статистичне означення ймовірності" містить раціональну частину – означення відносної частоти або, що те саме, статистичної ймовірності події A.

Справа в тому, що поняття ймовірнісної моделі (уявлення про яку треба сформувати у тих, хто вивчає курс теорії ймовірностей) потрібне для того, щоб зрозуміти:

- математичну сутність випадкової події (про це йшлося у пунктах 2.2 і 2.3);
- закономірності, пов'язані з тим, наскільки часто можуть відбуватися випадкові події у довгій серії випробувань;
- як за спостереженими значеннями частоти відбування певної випадкової події зробити досить точний прогноз щодо можливостей подальшого відбування цієї події.

Статистичні ймовірності (відносні частоти) $P_n^*(A)$ характеризують середню можливість відбування події A з простору S у кожному з n проведених випробувань. Разом з тим виявляється, що *часто ці статистичні ймовірності (відносні частоти) досить добре характеризують можливість відбування події A у кожному випробуванні*, а не тільки в уже проведених. Число $P_n^*(A)$ називають *статистичною ймовірністю* відповідної події $A \in S$, а трійку (Ω, S, P_n^*), тобто простір Ω елементарних подій, разом з простором подій S та статистичними ймовірностями $P_n^*(A)$, $A \in S$, називають *статистичною моделлю відповідного стохастичного експерименту* або ймовірнісним простором.

3.4. Безпосередньо з означення *статистичної ймовірності* $P_n^*(A)$, $A \in S$, випливають її *основні властивості*:

1_p. $P_n^*(A) \geq 0$ – для кожної події $A \in S$.

2_p. $P_n^*(\sum_i A_i) = \sum_i P_n^*(A_i)$, коли події $A_i \in S$ попарно несумісні, тобто $A_i A_j = \varnothing$, коли $i \neq j$.

3_p. $P_n^*(\Omega) = 1$.

Властивості 1_p-2_p нагадують відповідні властивості площі, об'єму, довжини, маси тощо. Всі ці величини мають узагальнену назву – *міра.* Якщо така міра задовольняє ще і властивість 3_p, тоді вона називається *імовірнісною мірою* (або просто ймовірністю).

З основних властивостей 1_p-3_p статистичної ймовірності випливають усі інші її властивості (див., наприклад, [2]).

3.5. На практиці часто виявляється, що статистичні ймовірності (відносні частоти) $P_n^*(A)$ починають досить добре характеризувати можливість відбування подій $A \in S$ у кожному випробуванні лише після того, як їх дещо змінити: одні трохи збільшити, інші – зменшити, а деякі залишити без зміни. Такий процес можна назвати процесом визначення ймовірностей $P(A)$ подій A з простору S. Цей процес проводять різними шляхами. Головним вважають те, щоб числа $P(A)$, $A \in S$, задовольняли *основні властивості ймовірності*:

1_p. $P(A) \geq 0$ для кожної події $A \in S$.

2_p. $P(\sum_i A_i) = \sum_i P(A_i)$ для довільних попарно несумісних подій $A_i \in S$, тобто таких, що $A_i A_j = \varnothing$, коли $i \neq j$.

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

При цьому числа $P(A)$ називають ймовірностями подій $A \in S$. Трійку (Ω, S, P), тобто простір елементарних подій Ω, разом з простором подій S та ймовірнісною мірою $P(A)$, $A \in S$, називають *ймовірнісною моделлю* (або *ймовірнісним простором*) даного *стохастичного експерименту*. Такі моделі намагаються будувати так, щоб ймовірності $P(A)$ характеризували (можливо наближено, проте досить точно) можливість відбування кожної реальної події (модель якої – подія A є елементом простору подій S) у кожному випробуванні, а також частоту відбування цієї події у кожній (досить довгій) серії з n випробувань: у такій серії випробувань згадана подія A відбудеться приблизно $[n \cdot P(A)]$ разів (якщо $P(A)$ визначена досить вдало).

У такому разі *ймовірнісну модель* називають *ефективною*.

Іноді виявляється, що після побудови простору елементарних подій Ω і простору подій S не можна завершити побудову ймовірнісної моделі (Ω, S, P) так, щоб ймовірність P задовольняла властивості 1_p-3_p (див., наприклад, [1]), тобто може виявитися, що не всі елементи вже побудованої сукупності S підмножин множини Ω можна вважати подіями. В такому разі виникає необхідність уточнення простору S або пере визначення ймовірнісної міри P.

Лише *після побудови ймовірнісної моделі* (Ω, S, P) *даного стохастичного експерименту можна вважати*:

- результатами (або наслідками) експерименту (елементарними подіями) – елементи Ω;
- випадковими подіями – елементи S;
- ймовірностями випадкових подій – числа $P(A)$, $A \in S$.

Усі інші підмножини простору Ω, що не входять до S, не вважаються випадковими подіями, оскільки неможливо визначити їх ймовірності у побудованому ймовірнісному просторі (Ω, S, P).

Отже, для одного і того самого стохастичного експерименту можна побудувати, взагалі кажучи, багато різних імовірнісних моделей. У рамках шкільного курсу математики учитель може будувати одну ймовірнісну модель певного експерименту (на його думку найпростішу), а сильнішим учням пропонувати будувати інші моделі та перевіряти їх ефективність, реалізуючи таким чином принцип індивідуалізації навчання.

Якщо побудовано ймовірнісний простір (Ω, S, P) і тим самим визначено ймовірнісну міру для кожного $A \in S$, ($A \subset \Omega$), тоді говорять, що задано розподіл ймовірностей на множині Ω (оскільки визначено ймовірність $P(A)$ для кожного $A \in S$, $A \subset \Omega$).

3.6. Виявляється, що ефективність ймовірнісної моделі (Ω, S, P) не можна перевірити, залишаючись у рамках цієї моделі. Потрібна побудова нової ймовірнісної моделі $(\tilde{\Omega}_n, \tilde{S}_n, \tilde{P}_n)$ "нового експерименту", для якого кожне випробування складається з n послідовних випробувань, пов'язаних із "попереднім експериментом". Подія $A \in S$ вважається фіксованою, а результатом випробування (нового) вважають набір з n чисел: $(x_1, ..., x_n)$, де кожне число $x_k = 1$, коли у k-му випробуванні ("попереднього експерименту") подія $A \in S$ відбувається, і $x_k = 0$, коли у цьому k-му випробуванні подія A не відбувається.

Отже, $\tilde{\Omega}_n = \{(x_1, ..., x_n) : x_k \in \{0,1\}, k \in \overline{1,n}\}$, а новий простір подій $\tilde{S}_n$ визначається умовою, що $\{(x_1, ..., x_n)\} \in \tilde{S}_n$ для кожної елементарної події $(x_1, ..., x_n) \in \tilde{\Omega}_n$.

Ймовірність $\tilde{P}_n$ визначається умовою $\tilde{P}_n(\{(x_1, ..., x_n)\}) = p^m (1-p)^{n-m}$, де $p = P(A)$, а m – кількість одиниць серед n чисел x_k, $x_k \in \{0,1\}$, $k \in \overline{1,n}$.

За таких умов ймовірність $\tilde{P}_n(B_{n,m})$ події $B_{n,m}$ – "відбування події A в n випробуваннях m разів", обчислюється за *формулою Бернуллі*

$$\tilde{P}_n(B_{n,m}) = C_n^m p^n (1-p)^{n-m}, m \in \overline{0,n}, \text{ де } C_n^m = \frac{n(n-1)\cdot \ldots \cdot (n-m+1)}{m!}.$$

Ймовірності $\tilde{P}_n(B_{n,m})$ називають *біноміальними*, оскільки вони обчислюється за тими ж правилами, що й члени розкладу бінома Ньютона:

$$(p + (1-p))^n = \sum_{m=0}^{n} C_n^m p^m (1-p)^{n-m} = \sum_{m=0}^{n} \tilde{P}_n(B_{n,m}) = 1.$$

Кожну подію $B_{n,m}$ можна вважати такою, що "у серії з n послідовних випробувань статистична ймовірність $P_n^*(A)$ події A виявилася рівною $\frac{m}{n}$". *Тому кожну біноміальну ймовірність* $\tilde{P}_n(B_{n,m})$ *можна тлумачити як ймовірність того, що статистична ймовірність* $P_n^*(A)$ *виявилася рівною* $\frac{m}{n}$.

У зв'язку з цим виникає питання: яке з усіх можливих значень $P_n^*(A) \in \{0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n}{n}\}$ є найбільш ймовірним і саме тому заслуговує на те, щоб це значення вибрати при побудові ймовірнісної моделі стохастичного експерименту.

Легко бачити (див., наприклад, [2]), що таким значенням є число $\frac{m_0}{n}$, де $m_0 = [(n+1)p]$. При цьому $p + \frac{p-1}{n} \le \frac{m_0}{n} < p + \frac{p}{n}$. А тому при великих n число $\frac{m_0}{n}$ близьке до p.

Отже, *серед усіх можливих значень статистичної ймовірності* $P_n^*(A) \in \{0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n}{n}\}$ *найбільш ймовірним (стосовно ймовірності* $\tilde{P}_n$*) є значення* $\frac{m_0}{n} = \frac{[(n+1)p]}{n}$, *яке для великих* n *близьке до числа* $p = P(A) \in (0;1)$.

Це є одна з *найпростіших форм закону великих чисел*.

Згідно з таким законом у випадку, коли число $p = P(A)$ невідоме (або треба перевірити, наскільки вдало це число визначене), знайти його наближення (або здійснити перевірку вдалості вибору) можна так:

1) провести кілька (наприклад, r) серій по n випробувань;
2) для кожної з проведених серій знайти значення статистичної ймовірності $P_n^*(A)$;
3) серед усіх знайдених значень $P_n^*(A)$ вибрати те, яке повторюється найчастіше, а коли таких значень кілька знайти їх середнє арифметичне;
4) вибране значення можна вважати наближенням ймовірності $p = P(A)$;
5) якщо вибране значення значно відрізняється від раніше визначеної ймовірності $p = P(A)$, то ефективність побудованої ймовірнісної моделі (Ω, S, P) сумнівна.

4. Поняття випадкової величини. Оскільки у багатьох навчальних посібниках фактично не вводиться поняття ймовірнісної моделі (ймовірнісного простору) (Ω, S, P) (у найкращому випадку згадується простір Ω елементарних подій), то наведені у цих посібниках означення випадкової величини насправді не є означеннями, а лише нечіткими тлумаченнями на інтуїтивному рівні.

Означення типу "випадковою величиною називають будь-яку числову функцію, визначену на просторі Ω елементарних подій" настільки ж некоректне, як і, наприклад, "означення" – "диференційовною називають будь-яку числову функцію, визначену на проміжку $(a;b)$".

Не кращим є й означення типу: "випадковою називають таку величину, яка внаслідок випробування може набути лише одне числове значення, заздалегідь невідоме і зумовлене випадковими причинами", оскільки не зрозуміло, яким чином пов'язані такі величини з результатами випробувань і що це за випадкові причини, які зумовлюють значення величини.

Разом з тим, навіть у шкільному курсі математики поняття випадкової величини можна ввести на достатньому науковому рівні, коректно і водночас доступно.

4.1. У школі можна обмежитися вивченням так званих простих випадкових величини.

Нехай задано ймовірнісний простір (Ω,S,P). Числову функцію $X(E)$, $E\in\Omega$, називають *простою випадковою величиною* (стосовно (Ω,S,P)), якщо множина її значень скінченна і для кожного числа X множина розв'язків рівняння $X(E)=x$ є подією з простору S.

Множину розв'язків рівняння $X(E)=x$ (відносно E) позначають $X^{-1}(x)$.

Якщо множина $\Omega_X=X(\Omega)$ значень функції $X(E)$, $E\in\Omega$, скінченна, то $\Omega_X=\{x_1,x_2,...,x_m\}$, де числа x_k, $k\in\overline{1,m}$, попарно різні, а число $m\in N$ – фіксоване. Згідно з означенням простої випадкової величини $X^{-1}(x_k)\in S$, а тому існує $P_X(x_k)=P(X^{-1}(x_k))$ – *ймовірність кожного значення* $x_k\in\Omega_X$.

Це означає, що якщо побудувати довільну сукупність S_X підмножин із Ω_X, яка задовільнятиме вимоги 1_S-3_S, то для будь-якого $V\in S_X$, $V\subset\Omega_X$, буде визначена ймовірність $P_X(V)=P(X^{-1}(V))$. В такий спосіб випадкова величина X генерує із ймовірнісного простору (Ω,S,P) в новий ймовірнісний простір (Ω_X,S_X,P_X).

Тут приховані можливості використання внутріпредметних зв'язків, пов'язані з розв'язуванням рівнянь і нерівностей та їх систем, визначенням областей задання функцій, множини їх значень тощо.

На практиці часто буває корисним трохи інше означення простої *випадкової величини* (стосовно (Ω,S,P)): це така числова функція $X(E)$, $E\in\Omega$, для якої простір Ω є сумою попарно несумісних подій A_k, $k\in\overline{1,n}$, на кожній з яких $X(E)$ є сталою, тобто $X(E)=c_k$, коли $E\in A_k$, $k\in\overline{1,n}$.

У цьому означенні числа c_k не обов'язково попарно різні. Зручність другого означення та його еквівалентність першому означенню проілюстровано у роботі [3].

Легко довести, що сума , різниця, добуток та частка (коли знаменник не дорівнює нулеві) простих випадкових величин також є простою випадковою величиною.

Існують числові функції, що визначені на просторі елементарних подій Ω, набувають лише скінченну кількість значень $x_1,...,x_m$ і водночас не є випадковими величинами (див. наприклад, [2] та [3]).

Щоб функція була простою випаковою величиною, важливо не лише те, що вона визначена на просторі елементарних подій Ω і має скінченну множину значень, а й те, що ймовірність кожного значення цієї величини можна визначити, виходячи з відповідного ймовірнісного простору (Ω,S,P):

$$P_X(\{x_k\})=P(X^{-1}(x_k)=p_k\ge 0\,,\ \sum_{k=1}^{m}p_k=1,\ \text{де}\ x_k\in\Omega_X\,,\ k\in\overline{1,m}\,,$$

а тому для довільного $V\subset\Omega_X$

$$P_X(V)=\sum_{x_i\in V}P_X(x_i)=P(X^{-1}(V))=\sum_{x_i\in V}P(X^{-1}(x_i)),$$

оскільки $P_X(V)$ за даних умов буде визначеним, коли $V\in S_X$ для будь-якої сукупності S_X підмножин множини Ω_X, що задовольняє вимоги 1_S-3_S.

4.2. Для кожної простої випадкової величини $X(E)$, $E\in\Omega$, на основі відповідної ймовірнісної моделі (Ω,S,P) можна знайти не тільки ймовірність $P_X(x_i)$ кожного її значення, а й ймовірність $P_X((-\infty;x))$ того, що "значення цієї величини попадає у проміжок $(-\infty;x)$" для будь-якого фіксованого числа x:

$$P_X((-\infty;x))=P(\{E\in\Omega:X(E)<x\}).$$

Справді, якщо множина значень X складається з чисел $x_1,...,x_m$, то множина $\{E\in\Omega:X(E)<x\}=\sum_{x_i<x}\{E\in\Omega:X(E)=x_i\}=\sum_{x_i<x}X^{-1}(x_i)$, тобто є сумою скінченної кількості подій $X^{-1}(x_i)\in S$, для яких $x_i<x$. Тому ця сума за властивістю подій 3_s сама є подією з простору S. Враховуючи попарну несумісність подій $X^{-1}(x_i)$ –доданків суми, маємо за властивістю ймовірності 2_p, що

$$P_X((-\infty;x))=\sum_{x_i<x}P_X(x_i)=\sum_{x_i<x}P(\{E\in\Omega:X(E)<x\})=\sum_{x_i<x}P(X^{-1}(x_i)).$$

Виходячи з останнього, можна узагальнити поняття випадкової величини на випадок числової функції $X(E)$, $E\in\Omega$, множина значень якої не обов'язково скінченна.

Нехай задано ймовірнісний простір (Ω, S, P). Тоді числову функцію $X(E)$, $E \in \Omega$, називають *випадковою величиною* (стосовно (Ω, S, P)), якщо для будь-якого числа x множина розв'язків нерівності $X(E) < x$ є подією з простору S, тобто $X^{-1}((-\infty, x)) \in S$, $x \in R^1$.

Множину розв'язків нерівності $X(E) < x$ (відносно E) позначають часто $X^{-1}((-\infty, x))$. (Тут прихована ще одна можливість використання внутріпредметних зв'язків, пов'язана з розв'язуванням нерівностей виду $X(E) < x$).

Зауважимо, що будь-яка функція $X(E)$, $E \in \Omega$, є випадковою величиною (стосовно (Ω, S, P)), тоді й тільки тоді, коли кожна підмножина простору Ω є подією з простору S, а це лише одна з багатьох можливостей.

Для випадкової величини X за допомогою імовірнісної моделі (Ω, S, P) можна знайти ймовірність $P_X((-\infty; x))$ того, що "ця величина набуде значення з наперед заданого проміжку $(-\infty; x)$":

$$P_X((-\infty; x)) = P(\{E \in \Omega : X(E) < x\}) = P(X^{-1}((-\infty, x))),$$

а також ймовірності $P_X([a;b)) = P_X((-\infty; b)) - P_X((-\infty; a))$ для будь-яких проміжків $[a;b)$.

4.3. Важливість простих випадкових величин полягає у тому, що за їх допомогою можна як завгодно добре наблизити будь-яку випадкову величину (стосовно того самого ймовірнісного простору (Ω, S, P)). Це можна здійснити, наприклад, так.

Нехай $X(E)$, $E \in \Omega$, довільна фіксована випадкова величина (стосовно (Ω, S, P)). Для кожного фіксованого $n \in N$ поділимо проміжок $[-n; n)$ на $2n^2$ проміжків $[-n + \frac{k-1}{n}; -n + \frac{k}{n}]$, $k \in \overline{1, 2n^2}$.

Введемо прості випадкові величини

$$X_{n,k}(E) = \begin{cases} 1, & \text{коли } -n + \frac{k-1}{n} \le X(E) < -n + \frac{k}{n}, \\ 0, & \text{в іншому разі,} \end{cases} \quad \text{та} \quad X_n(E) = \sum_{k=1}^{2n^2} \left(-n + \frac{k-1}{n}\right) X_{n,k}(E), \; E \in \Omega.$$

Легко бачити, що для будь-якої точки $E_0 \in \Omega$ існує таке число n_0, що $| X_n(E_0) - X(E_0) | \le \frac{1}{n}$, коли $n \ge n_0$.

Отже, $X_n(E_0) \approx X(E_0)$, а похибку наближення можна зробити як завгодно малою за рахунок вибору великого числа n.

Саме тому у шкільному курсі математики можна обмежитися розглядом лише простих величин.

5. Числові характеристики розподілу ймовірностей на множині значень простої випадкової величини. Для простої випадкової величини $X(E)$, $E \in \Omega$, з множиною значень $\{x_1, ..., x_m\}$ легко ввести такі важливі числові характеристики, що характеризують розподіл ймовірностей на множині значень випадкової величини:

математичне сподівання (координата центра розсіювання ймовірностей на множині значень простої випадкової величини X):

$$M[X] = \sum_{i=1}^{m} x_i P_X(x_i) = \sum_{i=1}^{m} x_i P(X^{-1}(x_i)) \tag{1}$$

дисперсія:

$$D[X] = M[(X - M[X])^2] = \sum_{i=1}^{m} (x_i - M[X])^2 P_X(x_i) \tag{2}$$

та *середнє квадратичне відхилення:*

$$\sigma = \sqrt{D[X]}.$$

Поняття математичного сподівання і дисперсії можна визначити дещо іншим способом (проте еквівалентним наведеному способу (див. наприклад, [3])): якщо події $A_i \in S$, $i \in \overline{1, n}$, попарно несумісні, $\sum_{i=1}^{n} A_i = \Omega$ і $X(E) = c_i$, коли $E \in A_i$, $i \in \overline{1, n}$, то

$$M[X] = \sum_{i=1}^{n} c_i P(A_i), \tag{1*}$$

$$D[X] = \sum_{i=1}^{m} (C_i - M[X])^2 P(A_i). \tag{2*}$$

Це означення може бути використане для вивчення внутріпредметних зв'язків, пов'язаних з переставною, сполучною та розподільною властивостями суми.

5.1. З наведених означень $M[X]$ та $D[X]$ легко випливають *властивості математичного сподівання та дисперсії* (див. наприклад, [2], [3]). Серед них:

1. $M[X] = c$, а $D[X] = 0$, коли $X(E) = c$, $E \in \Omega$.

2. $M[\sum_{k=1}^{n} a_k X_k] = \sum_{k=1}^{n} a_k M[X_k]$, коли $X_k(E)$, $E \in \Omega$, – прості випадкові величини, а a_k, $k \in \overline{1,n}$, – довільні сталі.

3. $M[X \cdot Y] = M[X] \cdot M[Y]$, коли $X[E]$ та $Y[E]$, $E \in \Omega$, прості випадкові величини, що є *незалежними*, тобто такими, що $P(X^{-1}(x) \cdot Y^{-1}(y)) = P(X^{-1}(x)) \cdot P(Y^{-1}(y))$ для будь-яких чисел x та y.

4. $D[\sum_{k=1}^{n} a_k X_k] = \sum_{k=1}^{n} a_k^2 D[X_k]$, коли прості випадкові величини $X_k(E)$, $E \in \Omega$, $k \in \overline{1,n}$, попарно незалежні, а a_k, $k \in \overline{1,n}$, – довільні сталі.

5. Якщо $X(E)$, $E \in \Omega$, проста випадкова величина (стосовно (Ω, S, P)), то для будь-якого $\varepsilon > 0$ множина $\{E \in \Omega : | X(E) - M[X] | \geq \varepsilon\}$ є подією з простору S (її коротко позначають $| X - M[X] | \geq \varepsilon$) і має місце *нерівність Чебишова*:

$$P(| X - M[X] | \geq \varepsilon) \leq \frac{D[X]}{\varepsilon^2} \tag{3}$$

або

$$P(| X - M[X] | < \varepsilon) \geq 1 - \frac{D[X]}{\varepsilon^2}. \tag{4}$$

Із нерівності Чебишова видно, *навіщо потрібні математичне сподівання та дисперсія*. З неї випливає, що коли дисперсія $D[X]$ мала , то малоймовірно, щоб значення випадкової величини X сильно відрізнялися від її математичного сподівання $M[X]$.

Тепер зрозуміло, чому число $M[X]$ названо математичним сподіванням: це така точка, біля якої в основному скупчуються спостережені значення випадкової величини X, коли проводять серію послідовних випробувань. А дисперсія $D[X]$ – це міра згаданої скупченості: чим менше $D[X]$, тим більшою є скупченість (менше розсіювання).

5.2. Наприкінці пункту **3.6** наведено найпростішу форму закону великих чисел. За допомогою нерівності Чебишова можна подати загальнішу форму цього закону. Для цього розглянемо послідовність ймовірнісних моделей (Ω, S, P_k), $k \in N$, у кожній з яких простори Ω елементарних подій та простори подій S не залежать від k (однакові для всіх k).

Для фіксованої події $A \in S$ позначимо $p_k = P_k(A)$, $k \in N$, і утворимо нову ймовірнісну модель $(\tilde{\Omega}_n, \tilde{S}_n, \tilde{P}_n)$, для якої $\tilde{\Omega}_n$ і $\tilde{S}_n$ визначено у пункті **3.6**, а ймовірність $\tilde{P}_n$ визначимо рівністю $\tilde{P}_n(\{(x_1, ..., x_n)\}) = a_1 \cdot a_2 \cdot ... \cdot a_n$, де $a_k = p_k$, коли $x_k = 1$, і $a_k = (1 - p_k)$, коли $x_k = 0$.

Тоді серед подій простору $\tilde{S}_n$ є події $A_k = \{(x_1, ..., x_n) \in \tilde{\Omega}_n : x_k = 1\}$ – "відбування події A у k-му випробуванні", $k \in \overline{1,n}$.

При цьому $\tilde{P}_n(A_k) = p_k$, $k \in \overline{1,n}$, і події A_k, $k \in \overline{1,n}$, є *незалежними*, тобто ймовірність добутку подій дорівнює добутку ймовірностей цих подій, коли події-співмножники попарно різні та вибрані довільно серед подій A_k, $k \in \overline{1,n}$.

Розглянемо прості випадкові величини (стосовно ймовірнісного простору $(\tilde{\Omega}_n, \tilde{S}_n, \tilde{P}_n)$):

$$X_k(E) = \begin{cases} 1, & \text{коли } E \in A_k, \\ 0, & \text{коли } E \notin A_k, \end{cases} \qquad Y_n(E) = \frac{1}{n} \sum_{k=1}^{n} X_k(E),\ E \in \tilde{\Omega}_n.$$

Легко бачити, що $Y_n(E)$ набуває значення $\frac{m}{n}$ тоді й тільки тоді, коли $E \in B_{n,m}$, де $B_{n,m}$ – "це відбування події A m разів в n послідовних випробуваннях" (див. п. **3.6**). Отже, $Y_n(E)$ – це $P_n^*(A)$ – проста випадкова величина з можливими значеннями $\frac{m}{n}$, $m \in \overline{0,n}$. Тому можна позначати $Y_n(E)$ через $P_n^*(E)$.

Згідно з означенням математичного сподівання та дисперсії маємо $M[X_k] = \tilde{P}_n(A_k) = p_k$, $D[X_k] = p_k(1 - p_k)$, а тому за властивостями 2 і 4 $M[Y_n] = M[P_n^*] = \frac{1}{n} \sum_{k=1}^{n} p_k$, а $D[Y_n] = \frac{1}{n^2} \sum_{k=1}^{n} p_k(1 - p_k) \leq \frac{1}{4n}$.

Тепер з нерівності Чебишова (3) або (4) випливає відповідно:

$$\tilde{P}_n(| P_n^* - \frac{1}{n} \sum_{k=1}^{n} p_k | \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2} \tag{5}$$

або

$$\tilde{P}_n(| P_n^* - \frac{1}{n} \sum_{k=1}^{n} p_k | < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}. \tag{6}$$

У випадку, коли $p_k = p = P(A)$ для всіх $k \in N$, нерівності (5) та (6) набувають вигляду:

$$\tilde{P}_n(|P_n^* - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2} \qquad (5^*)$$

або

$$\tilde{P}_n(|P_n^* - p| < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}. \qquad (6^*)$$

Нерівності (5*) і (6*) складають сутність *закону великих чисел для статистичних ймовірностей*: якщо n досить велике, то малоймовірно (стосовно ймовірності $\tilde{P}_n$), щоб значення статистичної ймовірності $P_n^*(A)$ (як випадкової величини стосовно $(\tilde{\Omega}_n, \tilde{S}_n, \tilde{P}_n)$) сильно відрізнялися від ймовірності $p = P(A)$.

Згідно з цим законом, якщо провести кілька серій по $n = 100$ випробувань, то швидше за все, не менше ніж у 75% цих серій $P_n^*(A)$ відрізнятиметься від $P(A)$ менше, ніж на $\varepsilon = 0{,}1$. А якщо провести кілька серій по $n = 10000$ випробувань, то, швидше за все, не менше ніж у 99% цих серій $P_n^*(A)$ відрізнятиметься від $P(A)$ менше, ніж на $\varepsilon = 0{,}05$. Саме така поведінка $P_n^*(A)$ надає впевненість в ефективності побудованої ймовірнісної моделі (Ω, S, P). Якщо така поведінка не спостерігається, то виникає сумнів щодо ефективності побудованої ймовірнісної моделі (Ω, S, P).

Моделі (Ω, S, P_k) можна розглядати як послідовність наближень до бажаної ймовірнісної моделі (Ω, S, P). У роботі [4] показано, що таке наближення здійснюється (у певному розумінні) тоді й тільки тоді, коли існує границя $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_k$. Точніше : для того, щоб існувало число $p = P(A)$, для якого $\tilde{P}_n(|P_n^* - p| \geq \varepsilon) \to 0$, коли $\varepsilon > 0$ –довільне фіксоване, а $n \to \infty$, необхідно й досить, щоб $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_k = p$.

Разом з тим доведена можливість існування статистичних моделей (Ω, S, P_k), для яких $P_k(A) = p_k = P_k^*(A)$, а послідовність чисел $\frac{1}{n}\sum_{k=1}^{n} p_k$, $n \in N$, не має границі. Значення членів такої послідовності навіть можуть *повільно коливатися* і бути як завгодно близькими до будь-якого числа з проміжку $(0;1)$.

Також доведено, що існування границі $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_k^*(A) = p$ зовсім не гарантує існування границі $\lim_{k\to\infty} P_k^*(A)$.

6. Висновки. 1. Теорія ймовірностей, як один з розділів сучасної математики, вивчає, як будувати математичні моделі, а саме ймовірнісні моделі, і застосовувати їх, а також досліджувати ефективність цих моделей. Цього і треба навчати студентів – майбутніх вчителів та учнів (на прикладах найпростіших ймовірнісних моделей).

2. Кожна побудована ймовірнісна модель може бути ефективною для розв'язування одних задач і неефективною для інших. Висновки стосовно реальних випадкових подій (а не їх моделей), зроблені за допомогою навіть ефективної ймовірнісної моделі, не повинні бути категоричними, оскільки кожна ймовірнісна модель не може охопити усі можливі випадки. У кращому випадку зроблені висновки є "правильними швидше за все" (що не виключає випадків, коли ці висновки неправильні, причому завжди існує ризик цього).

3. Для кожної ймовірнісної моделі будь-яка випадкова подія є певною множиною можливих наслідків даного (фіксованого) випадкового експерименту, тобто певною підмножиною A відповідного цьому експерименту простору Ω елементарних подій. Проте не завжди кожна підмножина $A \subset \Omega$ обов'язково є випадковою подією. Для побудови ефективної ймовірнісної моделі не завжди зручно вважати випадковими подіями будь-які підмножини простору Ω, а іноді це навіть неможливо.

Коли Ω містить більше одного елемента, існує багато сукупностей S підмножин $A \subset \Omega$, які можна вважати випадковими подіями. Серед них лише одна така, коли кожна підмножина $A \subset \Omega$ є випадковою подією.

Головним при побудові простору подій S є те, щоб виділені випадкові події задовольняли основні умови 1_s-3_s, а також потреби, пов'язані з розв'язанням певних задач.

Лише після побудови ймовірнісного простору (Ω, S, P), тобто ймовірнісної моделі стохастичного експерименту, можна сказати, які саме підмножини $A \subset \Omega$ є випадковими подіями, а які ні; лише тоді можна порівнювати між собою випадкові події та їх ймовірності.

4. Способів задання ймовірнісних мір (ймовірностей випадкових подій) при побудові ймовірнісної моделі дуже багато. Лише один з них пов'язаний з так званим "класичним означенням ймовірності", проте можливості застосування такого задання імовірнісної міри досить обмежені навіть у випадку скінченних просторів Ω елементарних подій.

Головним є те, щоб побудовані ймовірності $P(A)$, $A \in S$, задовольняли основні умови $1_P - 3_P$, а також потреби, пов'язані з розв'язуванням певних задач.

5. Статистичні ймовірності (відносні частоти) значно корисніші ніж "класичні ймовірності". Без них неможливо побудувати жодну ефективну ймовірнісну модель.

Статистична ймовірність виступає або *як ймовірність* – фіксована міра можливості відбування події $A \in S$ у кожному випробуванні (і тоді ймовірнісна міра $P_n^*(A)$ визначена на просторі подій S, тобто $A \in S$), або як проста випадкова величина, яка набуває усіх можливих значень відносних частот фіксованої події A (і тоді P_n^* визначена на просторі $\widetilde{\Omega}_n$, кожна елементарна подія $E = (x_1, ..., x_n)$ якого характеризує відбування $(x_k = 1)$ або невідбування $(x_k = 0)$ події A у k-му випробуванні). Тому у шкільному курсі математики основну увагу слід приділяти статистичним ймовірностям та простим випадковим величинам.

6. Кожна випадкова величина є числовою функцією, визначеною на певному просторі Ω елементарних подій, проте не кожна числова функція, визначена на Ω, є випадковою величиною. Головним є те, що кожна випадкова величина $X(E)$, $E \in \Omega$, пов'язана з певною ймовірнісною моделлю (Ω, S, P), що дозволяє знайти ймовірність попадання значень цієї величини у будь-який проміжок $[-\infty; b)$: множина розв'язків нерівності $X(E) < b$ повинна бути подією з простору S для кожного числа b.

Кожну випадкову величину можна як завгодно точно наблизити за допомогою відповідної простої випадкової величини, тому у школі основну увагу слід приділити саме простим випадковим величинам.

7. Ефективність ймовірнісної моделі іноді можна перевірити за допомогою інших ймовірнісних моделей, з якими пов'язані біноміальні ймовірності та закон великих чисел.

8. На відміну від сучасних дедуктивних курсів теорії ймовірностей шкільний курс доцільно будувати індуктивно, допомагаючи учням самим відкривати основні результати. Зокрема, основні властивості подій: 1_s-3_s та ймовірностей: 1_p-3_p слід подавати не як аксіоми, що невідомо звідки і навіщо з'явилися, а як зручні і прості правила розв'язування великого класу задач. В ідеалі, ці правила учні (під керівництвом учителя) відкривають самі.

9. Існує безліч різноманітних імовірнісних моделей, проте на практиці найчастіше застосовуються лише деякі, найбільш ретельно досліджені. Серед них є ймовірнісні моделі, що визначаються такими розподілами ймовірностей: рівномірний дискретний, біноміальний, Пуассона, рівномірний неперервний, показниковий, нормальний, Коші. Застосування таких моделей у школі доцільно лише за умови використання відповідних комп'ютерних програм, наприклад, GRAN1.

Взагалі, використання комп'ютерних засобів при вивченні теорії ймовірностей повинно бути систематичним і таким, щоб повністю вивільнити учнів від рутинних та громіздких операцій.

ЛІТЕРАТУРА

1. Жалдак М.І., Михалін Г.О. Деякі властивості ймовірнісних моделей стохастичних експериментів // Комп'ютерно-орієнтовані системи навчання. Збірник наукових праць. – К.: Комп'ютер в школі та сім'ї. – 2001. – Вип. 3. – с. 49-68.
2. Жалдак М.І., Михалін Г.О. Елементи стохастики з комп'ютерною підтримкою. – К.: Шкільний світ, 2006. – 120 с.
3. Михалін Г.О., Стогній О.В. Статистичні ймовірності. Прості випадкові величини. Закон великих чисел. // Комп'ютерно-орієнтовані системи навчання. Науковий часопис НПУ ім. М.П. Драгоманова. Серія №2. – К.: НПУ ім. М.П. Драгоманова, 2006. – №4(11). – с. 163-170.
4. Жалдак М.І., Михалін Г.О., Деканов С.Я. Одне узагальнення поняття границі функції та деякі його застосування // Комп'ютерно-орієнтовані системи навчання. Науковий часопис НПУ ім. М.П. Драгоманова. Серія №2. – К.: НПУ ім. М.П. Драгоманова, 2007. – №5(12). – с. 3-9.