МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ І ПРИКЛАДНИХ ТЕХНОЛОГІЙ
КАФЕДРА ПРИКЛАДНОЇ МАТЕМАТИКИ ТА КІБЕРБЕЗПЕКИ

ТЕОРІЯ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНА СТАТИСТИКА

Методичні рекомендації до самостійної роботи

для здобувачів вищої освіти
ступеня освіти «Бакалавр» спеціальностей
125 Кібербезпека та захист інформації,
111 Математика, 113 Прикладна математика

Вінниця
2023

УДК 519.2(075.4)
Т 338

Затверджено на засіданні вченої ради факультету
інформаційних і прикладних технологій
Донецького національного університету імені Василя Стуса
(протокол № 4 від 21 листопада 2023 р.)

Укладачі: Людмила ПОЛОВЕНКО, канд. пед. наук, доцент кафедри прикладної математики та кібербезпеки ДонНУ імені Василя Стуса;
Оксана ДАНИЛЬЧУК, канд. пед. наук, доцент кафедри прикладної математики та кібербезпеки ДонНУ імені Василя Стуса.

Рецензенти: Тарас ФУРМАН, канд. пед. наук, доцент кафедри маркетингу та бізнес-аналітики ДонНУ імені Василя Стуса;
Наталія ДОБРОВОЛЬСЬКА, канд. пед. наук, доцент кафедри обчислювальної техніки ВНТУ.

Т 338 Методичні рекомендації до самостійної роботи з дисципліни «Теорія ймовірностей та математична статистика» (для здобувачів вищої освіти ступеня освіти «Бакалавр» спеціальностей 125 Кібербезпека та захист інформації, 111 Математика, 113 Прикладна математика / укл. Л. П. Половенко, О М. Данильчук. Вінниця: ДонНУ імені Василя Стуса, 2023. 61с.

УДК 519.2(075.4

ЗМІСТ

ВСТУП

Знаменна річ, що наука (теорія ймовірності),
яка почалася з вивчення ігор, піднеслася
до найважливіших об'єктів людського пізнання.
П. Лаплас

«Теорія ймовірностей та математична статистика» входить до циклу дисциплін професійної та практичної підготовки здобувачів вищої освіти спеціальностей 125 Кібербезпека та захист інформації, 111 Математика, 113 Прикладна математика ступеня вищої освіти «Бакалавр».

Тематика курсу включає основи комбінаторики, теорії ймовірностей, теорії оцінювання невідомих параметрів, перевірки статистичних гіпотез, елементів кореляційно-регресійного аналізу. Розглядаються технологічні прийоми і способи комп'ютерної реалізації статистичної обробки даних. Питання цієї дисципліни мають широке прикладне застосування.

Мета вивчення дисципліни полягає у формуванні системи знань та умінь, аналітично-дослідницьких компетентностей, які необхідні сучасним фахівцям з кібербезпеки та прикладної математики для опанування та використання технологій і методів ймовірнісного та статистичного аналізу в інформаційній та кібербезпеці, проведенні прикладних досліджень.

Основними завданнями дисципліни є: формування у здобувачів базових знань з основ застосування ймовірнісно-статистичного апарату для розв'язування теоретичних і практичних задач у професійній діяльності, а також розвитку логічного, алгоритмічного мислення під час виявлення та дослідження закономірностей, яким підпорядковуються реальні соціальні, економічні та інформаційні процеси на основі певних статистичних даних та в умовах невизначеності.

У методичних рекомендаціях запропоновані допоміжні матеріали для опанування питань, які виносяться на самостійне опрацювання за кожною з 15 тем відповідно до силабусу дисципліни «Теорія ймовірностей та математична статистика».

Для зручності користування методичними рекомендаціями на початку кожної теми наведені методичні рекомендації до вивчення конкретних тем, приклади розв'язування задач, питання для самостійного вивчення, перелік індивідуальних завдань.

Для закріплення самостійно опрацьованого матеріалу здобувач вищої освіти може використати питання для самоконтролю, які запропоновані після кожної теми.

ТЕМА 1
ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ. ДІЇ З МНОЖИНАМИ

Немає принципів, є події;
немає законів, є обставини;
людина високого польоту
сама пристосовується до подій і обставин,
щоб керувати ними.
Оноре де Бальзак

Основні поняття: подія та випробування

Послідовність операцій, виконуваних із додержанням певного комплексу умов, називають експериментом (випробуванням, спостереженням).

Подією в теорії ймовірності називають довільний наслідок або результат будь-якого випробування, спостереження, стохастичного експерименту (який можна повторити будь-яку кількість разів), який може настати (відбутися, здійснитися), або не настати, тобто результат випробування не можна точно передбачити. Позначають події великими латинськими літерами A, B, C тощо, описують словами в лапках або фігурних дужках.

Для того, щоб результат експерименту був подією, має бути можливою повторюваність експерименту за незмінних умов.

Події поділяються на достовірні, неможливі та випадкові.

Якщо внаслідок експерименту, здійснюваного з додержанням певного комплексу умов, певна подія обов'язково настає, то вона називається **достовірною.**

Подія називається **неможливою,** якщо внаслідок експерименту, проведеного з додержанням певного комплексу умов, вона не настає ніколи.

Подія називається **випадковою,** якщо за певного комплексу умов внаслідок експерименту вона може настати або не настати залежно від дії численних дрібних факторів, урахувати які дослідник не в змозі.

Елементарні події позначаються ω_i ($i = 1, 2, 3,\ldots$).

Приклад 1.1. Монету підкидають один раз. Визначити елементарні події цього експерименту.

Розв'язання. Можливі такі елементарні випадкові події:

ω_1 = г (монета випаде гербом);

ω_2 = ц (монета випаде цифрою).

Приклад 1.2. Монету підкидають тричі. Визначити елементарні події цього експерименту.

Розв'язання. Триразове підкидання монети – це одна спроба. Елементарними випадковими подіями будуть:

ω_1 = ггг (тричі випаде герб);

ω_2 = ццц (тричі випаде цифра);

ω_3 = ггц (герб випаде двічі);

ω_4 = гцг (цифра випаде двічі);

ω_5 = цгг (герб випаде один раз);

ω_6 = гцц (цифра випаде один раз);

Отже, цьому експерименту відповідають вісім елементарних подій.

Приклад 1.3. Задано дві множини цілих чисел $\Omega_1=\{1,2,3\}$, $\Omega_2=\{1,2,3,4\}$. Із кожної множини навмання беруть по одному числу. Визначити елементарні події цього експерименту – появу пари чисел.

Розв'язання. Елементарними випадковими подіями будуть:

$\omega_1 = 1; 1;$	$\omega_5 = 2; 1;$	$\omega_9 = 3; 1;$
$\omega_2 = 1; 2;$	$\omega_6 = 2; 2;$	$\omega_{10} = 3; 2;$
$\omega_3 = 1; 3;$	$\omega_7 = 2; 3;$	$\omega_{11} = 3; 3;$
$\omega_4 = 1; 4;$	$\omega_8 = 2; 4;$	$\omega_{12} = 3; 4.$

Отже, цьому експерименту відповідають дванадцять елементарних подій.

Випадкова подія називається **складеною,** якщо її можна розкласти на прості (елементарні) події. Складені випадкові події також позначають великими латинськими літерами.

Приклад 1.4. Задано множину чисел $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Навмання із цієї множини беруть одне число. Побудувати такі випадкові події:

1) з'явиться число, кратне 2;

2) число, кратне 3;

3) число, кратне 5.

Розв'язання. Ці випадкові події будуть складеними. Позначимо їх відповідно A, B, C. Тоді $A=\{2, 4, 6, 8, 10, 12\}$; $B=\{3, 6, 9, 12\}$; $C=\{5,10\}$.

Елементарні випадкові події $\omega_i \in A$, $\omega_j \in B$, $\omega_k \in C$, які належать відповідно складеним випадковим подіям A, B, C, тобто є елементами цих множин, називають **елементарними подіями, які сприяють появі кожної із зазначених подій** внаслідок проведення експерименту (ω_i сприяють появі події A, ω_j – події B, ω_k – події C).

Кожному експерименту (спробі) з випадковими результатами (наслідками) відповідає певна множина Ω елементарних подій ω_i, кожна з яких може відбутися (настати) внаслідок його проведення: $\omega_i \in\Omega$. Цю множину називають **простором елементарних подій.**

Приклад 1.5. Гральний кубик, кожна грань якого позначена певною цифрою від 1 до 6, підкидають один раз. Під час цього на грані випадає одна із зазначених цифр. Побудувати простір елементарних подій для цього експерименту (множину Ω) і такі випадкові події:

1) A – випаде число, кратне 2;

2) B – випаде число, кратне 3.

Розв'язання. Оскільки кубик має шість граней, то внаслідок експерименту може випасти одна із цифр від 1 до 6.

Отже, $\Omega = \{1, 2, 3, 4, 5, 6\}$; 1) $A = \{2, 4, 6\}$; 2) $B = \{3, 6\}$.

Приклад 1.6. Монету підкидають чотири рази. Побудувати простір елементарних подій для цього експерименту і такі випадкові події:

1) A – герб випаде двічі;

2) B – герб випаде не менш як тричі.

Розв'язання. Шуканий простір елементарних подій:

Ω = {гггг, гггр, ггрг, гргг, рргг, ггрр, гррг, гргр, ргрг, рргг, рггр, гррр, ргрр, рргр, рррг, рррр}.

1) A = {ггрр, рргг, гргр, ргрг, гррг, рггр};

2) B = {гггг, гггр, ггрг, гргг, рггг).

Операції над подіями

Сумою або об'єднанням подій *A* та *B* називається така подія C, яка полягає у настанні події A, або події B, або подій A та B одночасно.

Отже, якщо Ω_A, Ω_B, Ω відповідно множини окремих реалізацій події A, B та суми цих подій, то $\Omega = \Omega_A \cup \Omega_B$.

Суму подій A та B будемо позначати: $C=A+B$, або $C=A \cup B$.

Сумою або об'єднанням будь-якої кількості подій $A_1, A_2, \ldots, A_n$ називається подія C, яка полягає в настанні хоч би однієї з цих подій :$C = \sum_{i=1}^{n} \mathrm{A_i}$ *або* $C = \bigcap_{i=1}^{n} \mathrm{A_i}$.

Приклад 1.7. Стрілець робить один постріл у мішень, поділену на три області. Позначимо:

подія A_1 – влучення в першу область;

подія A_2 – влучення в другу область;

подія A_3 – влучення в третю область;

подія A_4 –немає влучення у мішень;

подія B – влучення в першу або другу області;

подія D – влучення хоча б в одну область мішені.

Тоді маємо $B = A_1 \cup A_2$; $D = A_1 \cup A_2 \cup A_3$. Зазначимо, що події A_1, A_2, A_3 та A_4 – несумісні.

Добутком подій *A* та *B* називається така подія C, яка полягає в одночасному виконанні події A і події B відповідно.

Отже, якщо Ω_A, Ω_B, Ω відповідно множини окремих реалізацій подій A, B та добутку цих подій, то $\Omega = \Omega_A \cap \Omega_B$. Добуток подій A та B будемо позначати $C=A \cdot B=A \times B$.

Добуток подій $A_1, A_2, \ldots A_n$ полягає в одночасному настанні всіх n подій: $C = \prod_{i=1}^{n} \mathrm{A_i}$ або $C = \bigcap_{i=1}^{n} A_i$.

Різницею подій *A* і *B* називається подія C, яка полягає в тому, що подія A відбувається, а подія B – ні.

Отже, діаграми Ейлера введених операцій над подіями такі:

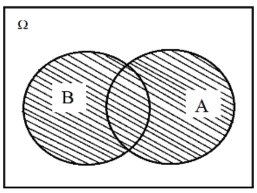


Рис. 1.1. $C=A+B$

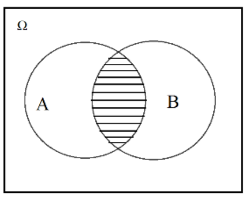


Рис. 1.2. $C=A \times B$

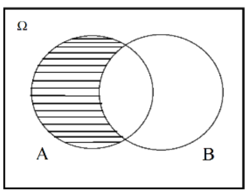


Рис. 1.3. $C=A \backslash B$

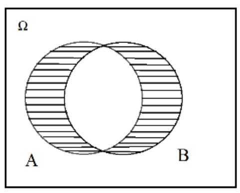


Рис. 1.4. $C=A \Delta B$

Якщо $A \cap B \neq \varnothing$, то випадкові події A і B називають **сумісними.**

Якщо $A \cap B = \varnothing$, то такі випадкові події A і B називають **несумісними.**

Якщо $A_1 \cup A_2 \cup A_3 \cup \ldots \ldots \cup A_n = \bigcup_{i=1}^{n} A_i = \Omega$, то такі випадкові події утворюють **повну групу,** а саме: внаслідок експерименту якась із подій A_i обов'язково настане.

Дві несумісні випадкові події, що утворюють повну групу, називають **протилежними.**

Протилежною подією до події ***A*** називається така подія $\bar{A}$, яка полягає в тому, що подія *A* не відбувається.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Ω_1=**{1, 2, 2},** Ω_2=**{1, 2, 2, 4}.** Утворити впорядковані пари чисел так, що перше з Ω_1, а друге з Ω_2. Описати простір елементарних подій цього експерименту.

Задача 2. Два стрільця стріляють по мішені. Нехай подія *A* полягає в тому, що перший стрілець влучив, подія *B* – другий стрілець влучив. Описати словами події $A \cup B$, $A \cap B$, $\bar{A}$, $\bar{B}$.

Задача 3. Зобразити на діаграмі Ейлера–Венна подію $(A/B) \cup (B/A)$.

Задача 4. Нехай *A*, *B*, *C* – довільні події. Зобразити на діаграмі Ейлера–Венна та записати символами дій подію «відбулося не більше 1 події з *A*, *B* та *C*».

Задача 5. Підкидають 2 кубики. Подія – кількість очок, що випали на кубиках. Описати ПЕП та події *A*={Обидва рази не більше трьох очок}, *B*={хоча б раз менше двох очок, $A \cup B$, $A \cap B$, $\bar{A}$.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що називається експериментом?
2. Яка подія називається випадковою? Навести приклади.
3. Яка подія називається елементарною; складеною випадковою подією? Навести приклади.
4. Які події називаються рівноможливими?
5. Що називається простором елементарних подій? Навести приклади.
6. Що називається сумою двох випадкових подій *A* і *B*?
7. Що називається добутком двох випадкових подій *A* і *B*?
8. Що називається різницею двох випадкових подій *A* і *B*?

ТЕМА 2
ОСНОВНІ ПОНЯТТЯ КОМБІНАТОРИКИ

Існує клас задач, у яких для обчислення n і m використовуються елементи комбінаторики: перестановка, розміщення та комбінації.

Перестановкою із n елементів називають такі впорядковані множини з n елементів, які різняться між собою порядком їх розміщення.

Кількість таких упорядкованих множин обчислюється за формулою:

$$P_n = n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \ldots \cdot (n-1) \cdot n. \quad (2.1)$$

Приклад 2.1. Скільки п'ятизначних чисел можна записати, використовуючи п'ять різних цифр (крім нуля)?

Розв'язання. Сполуки, що утворюють із п'яти різних цифр п'ятизначні числа, можуть відрізнятися лише порядком цифр, тому такі сполуки будуть представленням із 5 елементів. Їх кількість буде P_5=5!=1·2·3·4·5=120.

Розміщенням із n елементів по m $(0 \le m \le n)$ називаються такі впорядковані множини, кожна із яких містить m елементів і які відрізняються між собою порядком розташування цих елементів.

Кількість таких множин обчислюється за формулою:

$$A_n^m = n(n-1)(n-2) \ldots (n-m-1) \text{ або } A_n^m = \frac{n!}{(n-m)!}. \quad (2.2)$$

Приклад 2.2. Здобувачі другого курсу, відповідно до навчального плану, вивчають 10 дисциплін. На один день можна планувати заняття з 4 дисциплін. Скількома способами можна скласти розклад занять на один день?

Розв'язання. Усі можливі розклади занять на один день – це сполуки з 10 по 4, які можуть відрізнятися дисциплінами або їх порядком, тобто ці сполуки – розміщення. Кількість таких розміщень, згідно з формулою (2.2) буде:

$$A_{10}^4 = 10 \cdot 9 \cdot 8 \cdot 7 = 5040.$$

Комбінаціями з n елементів по m $(0 \le m \le n)$ називаються невпорядковані множини з m елементів, які різняться між собою хоча б одним елементом.

Кількість таких множин:

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n!}{m!(n-m)!}. \quad (2.3)$$

Розглянемо найважливіші властивості чисел:

1) $C_n^k = C_n^{n-k}$;

2) $C_{n+1}^k = C_n^k + C_n^{k-1}$;

3) $\sum_{k=0}^{n} C_n^k = 2^n$.

Через це числа виду C_n^k називають **біноміальними коефіцієнтами.**

Приклад 2.3. У шафі міститься 10 однотипних деталей, 6 із яких є стандартними, а решта бракованими. Навмання із шафи беруть чотири деталі. Обчислити ймовірність таких випадкових подій:

A – усі чотири деталі виявляться стандартними;

B – усі чотири деталі виявляться бракованими;

D – із чотирьох деталей виявляться дві стандартними і дві бракованими.

Розв'язання. Кількість усіх елементарних подій множини Ω:

$$n = C_{10}^4 = \frac{10!}{4! \cdot 6!} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4} = 210;$$

кількість елементарних подій, що сприяють події A:

$$m_1 = C_6^4 = \frac{6!}{2! \cdot 4!} = \frac{5 \cdot 6}{1 \cdot 2} = 15;$$

кількість елементарних подій, що сприяють появі B:

$$m_2 = C_4^4 = \frac{4!}{4! \cdot 0!} = 1;$$

кількість елементарних подій, що сприяють появі D:

$$m_3 = C_6^2 \cdot C_4^2 = 15 \cdot 6 = 90.$$

Обчислимо ймовірності цих подій:

$$P(A) = \frac{m_1}{n} = \frac{C_6^4}{C_{10}^4} = \frac{15}{210} = \frac{1}{14};$$

$$P(B) = \frac{m_2}{n} = \frac{C_4^4}{C_{10}^4} = \frac{1}{210};$$

$$P(D) = \frac{m_3}{n} = \frac{C_6^2 \cdot C_4^2}{C_{10}^4} = \frac{90}{210} = \frac{3}{7}.$$

Комбінаторика вивчає способи підрахунку числа можливих вибірок із деякої скінченої множини згідно із заданими правилами вибору.

Правило суми. Якщо предмет A може бути вибраним m способами, а предмет B іншими n способами (тобто ці вибори взаємовиключні), та вибір одного з предметів або A, або B може бути виконаним $m+n$ способами.

Узагальнене правило суми. Якщо предмет A_1 може бути вибраним m_1 способами, предмет A_2 іншими m_2 способами, предмет A_r іншими m_r способами, то вибір одного з предметів або A_1, або A_2 …, або A_r може бути виконаний $\sum_{i=1}^{r} m_i$ способами.

Правило добутку. Якщо предмет A може бути вибраний m способами і після кожного такого вибору предмет B може бути вибраний n способами (незалежно від того, який предмет A обрано), то сумісний вибір A і B, тобто впорядкована пара (A,B) може бути виконано $n \times m$ способами.

Узагальнене правило добутку. Якщо предмет A_1 можна вибрати m_1 способами, предмет A_2 – m_2 способами, …, предмет A_r – m_r способами, причому вибір кожного з них не впливає на вибір іншого, то вибір впорядкованої системи предметів $(A_1, A_2, \ldots, A_r)$ може бути виконаний $m_1 \cdot m_2 \cdot \ldots \cdot m_r$ способами.

Після формалізації умови (з'ясування того, які множини нам потрібні і яким способом вони утворюються) для вибору m необхідних елементів з n наявних різних елементів потрібно скористатись однією з чотирьох формул, поданих у таблиці 2.1.

Таблиця 2.1

Таблиця комбінаторних формул

Операції / Множини	**Безповторні**	**Повторні**
Впорядковані	$A_n^m = \dfrac{п!}{(п-m)!}$	n^m
Невпорядковані	$C_n^m = \dfrac{n!}{m!\,(n-m)!}$	$C_{n+m-1}^m = \dfrac{(n+m-1)!}{m!\,(n-1)!}$

Приклад 2.4. У кошику 4 яблука першого сорту та 5 яблук другого сорту. Навмання беруть 2 яблука. Знайти імовірність того, що будуть взяті яблука різних сортів.

Розв'язання. Нехай подія *A* – навмання взяті 2 яблука різних сортів. Усього яблук 9, із них сполучень по 2 буде C_9^2, тобто кількість усіх можливих наслідків $n = C_9^2$. Події *A* будуть сприяти сполуки, утворені з пар, елементами яких будуть яблука різних сортів. Згідно з принципом добутку кількість таких пар буде дорівнювати $m = C_4^1 \cdot C_5^1$. Використовуючи класичне означення ймовірності, одержимо шукану ймовірність події *A*:

$$P(A) = \frac{m}{n} = \frac{C_4^1 \cdot C_5^1}{C_9^2} = 4 \cdot 5 : \frac{9!}{2! \cdot 7!} = 20 \cdot \frac{2}{8 \cdot 9} = \frac{5}{9}.$$

Приклад 2.5. Із пункту *A* в пункт *Б* веде 3 дороги, а з пункту *Б* до пункту *В* веде 4 дороги. Скількома способами можна добратись із *A* до *В* через *Б*?

Розв'язання. Вибір дороги з *Б* до *В* не залежить від вибору з *A* до *Б*. Вибрати треба пару доріг з {1, 2, 3} – 3 способи та з {1, 2, 3, 4} – 4 способи. Тому за правилом добутку отримаємо $3 \cdot 4 = 12$ способами.

Приклад 2.6. Із пункту *A* в пункт *Б* можна дістатись або через *К* (2 дороги) або через *Л* (5 доріг). Скількома способами можна добратись із *A* до *Б*?

Розв'язання. У цій ситуації треба вибрати одну дорогу з *A* до *Б*: або через *К*, або через *Л*. Вибір, через який пункт – взаємовиключний. Тому за правилом суми з *A* до *Б* можна добратись $2 + 5 = 7$ способами.

Приклад 2.7. Набираючи номер телефону, абонент забув дві останні цифри і набрав їх навмання. Яка ймовірність того, що він набрав потрібні цифри?

Розв'язання. Нехай подія *A* – абонент набрав потрібні цифри. Кількість різних за складом або послідовністю набору пар цифр не менша за кількість розміщень із 10 елементів за двома, що складає загальну кількість елементарних випадків $n = A_{10}^2 = \frac{10!}{8!} = 10 \cdot 9 = 90$, із яких лише один набір цифр є сприятливим для події *A*. Зазначимо, що неврахованими залишилися десять пар однакових чисел: 00, 11, 22, …, 99. Тому загальне число *n* дорівнюватиме 100, і шукана ймовірність $P(A) = \frac{1}{100} = 0{,}01$.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. У групі 21 здобувач. Скількома способами можна вибрати в цій групі трьох студентів для проходження виробничої практичної підготовки на трьох підприємствах?

Задача 2. На вершину гори веде вісім стежок. Скількома способами турист може піднятися на гору і спуститися з неї? Дати відповідь на це саме запитання, якщо підйом і спуск потрібно здійснити різними стежками.

Задача 3. На чергування у студентському гуртожитку може піти або здобувач з кімнати № 1, у якій проживають чотири здобувачі, або здобувач з кімнати № 2, у якій проживають п’ять здобувачів. Скількома способами можна вибрати одного здобувача на чергування в гуртожитку?

Задача 4. Скільки існує способів розміщення на полиці шести книжок?

Задача 5. Комісія складається з голови, його заступника і ще п’яти членів. Скількома способами сім членів комісії можуть розподілити між собою обов’язки?

Задача 6. На зборах присутні 30 осіб. Скількома способами можна обрати президію зборів у складі трьох осіб?

Задача 7. Скільки трицифрових чисел можна записати за допомогою дев’яти цифр 1, 2, ... , 9, якщо цифри в запису числа можуть повторюватися?

Задача 8. Скільки різних слів, включно з беззмістовними, можна дістати, переставляючи букви у слові математика?

Задача 9. Скільки різних натуральних чисел можна утворити з цифр 0, 1, 2, 3, 4, якщо кожне число містить кожну з цих цифр не більше одного разу?

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що називається перестановкою із n елементів, розміщенням із n елементів по m, комбінацією із n елементів по m?
2. Відомо, що $A \cap B \neq \emptyset$. Чому дорівнює $P(A \cup B)$?

ТЕМА 3
КЛАСИЧНЕ ТА ГЕОМЕТРИЧНЕ ОЗНАЧЕННЯ ЙМОВІРНОСТІ

Вірогідне нам завжди здається неймовірним...
Еріх Марія Ремарк

Класичне означення ймовірності

Імовірністю випадкової події A називається невід'ємне число $P(A)$, що дорівнює відношенню числа елементарних подій m $(0 \leq m \leq n)$, які сприяють появі A, до кількості всіх елементарних подій n простору Ω:

$$P(A) = \frac{m}{n}.$$

Класичне означення ймовірності актуальне лише тоді, коли m та n скінчені, усі елементарні наслідки рівноможливі.

Приклад 3.1. У ящику міститься 15 однотипних деталей, із яких 6 браковані, а решта – стандартні. Навмання з ящика береться одна деталь. Яка ймовірність того, що вона буде стандартною?

Розв'язання. Число всіх рівноможливих елементарних подій для цього експерименту: $n = 15$.

Нехай A – подія, що полягає в появі стандартної деталі. Число елементарних подій, що сприяють появі випадкової події A, дорівнює дев'яти ($m = 9$). Згідно з класичним означенням маємо:

$$P(A) = \frac{m}{n} = \frac{9}{15} = \frac{3}{5}.$$

Приклад 3.2. Гральний кубик підкидають один раз. Яка ймовірність того, що на грані кубика з'явиться число, кратне 3?

Розв'язання. Число всіх елементарних подій для цього експерименту: $n = 6$. Нехай B – поява на грані числа, кратного 3. Число елементарних подій, що сприяють появі B, дорівнює двом ($m = 2$).

Отже:

$$P(B) = \frac{m}{n} = \frac{2}{6} = \frac{1}{3}.$$

Властивості ймовірності:

1. Ймовірність кожної події є відповідним числом з інтервалу [0; 1], тобто $0 \leq P(A) \leq 1$.
2. Ймовірності еквівалентних подій рівні, тобто, якщо $A = B$, то $P(A)=P(B)$.
3. Ймовірність достовірної події рівна одиниці, тобто $P(\Omega) = 1$.
4. Ймовірність неможливої події рівна нулю, тобто $P(\emptyset) = 0$.
5. Ймовірності події A та протилежної події $\bar{A}$ задовольняють співвідношення: $P(\bar{A}) = 1 - P(A)$.

Геометричне означення ймовірності

Якщо множина усіх елементарних наслідків нескінчена і, як наслідок, займає деяку область G, а події A сприяє лише частина $g \in G$, то обчислення імовірності події A виконують згідно з геометричним означенням ймовірності.

Ймовірність випадкової події *A* дорівнює відношенню міри *A* до міри Ω:

$$P(A) = \frac{m(A)}{m(\Omega)}.$$

Приклад 3.3. По трубопроводу між пунктами A і B, відстань між якими 2 км, перекачують нафту. Яка ймовірність того, що пошкодження через певний час роботи трубопроводу станеться на ділянці довжиною 100 м?

Розв'язання. Простір елементарних подій $\Omega = \{0 \le l \le 2 \text{ км}\}$, тоді $\text{A} = \{0 \le l_1 \le 0{,}1 \text{ км}\}$, $(A \subset \Omega)$.

Згідно з геометричним означенням маємо $P(A) = \frac{m(A)}{m(\Omega)} = \frac{l_1}{l} = \frac{0{,}1}{2} = \frac{1}{20}$.

Приклад 3.4. Задана множина $\Omega = (0 \le x \le e,\ 0 \le y \le 1)$. Яка ймовірність того, що навмання взяті два числа (x; y) утворять координати точки, яка влучить в область $A = (1 \le x \le e,\ 0 \le y \le lnx$?

Розв'язання. Множини Ω і *A* зображені на рис. 3.1.

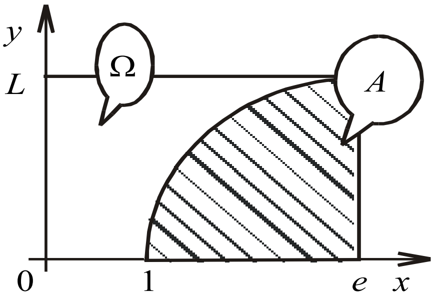


Рис. 3.1. Геометрична інтерпретація множин Ω і *A*

$$P\ (A) = \frac{m(A)}{m(\Omega)} = \frac{S_A}{S_\Omega} = \frac{\int_1^e \ln x dx}{e} = \frac{x \ln x\ |_1^e - \int_1^e dx}{e} = \frac{x \ln x\ |_1^e - x\ |_1^e}{e} = \frac{e - e + 1}{e} = \frac{1}{e}.$$

Приклад 3.5. Два туристичні пароплави повинні причалити до одного причалу. Час прибуття обох пароплавів рівноможливий протягом доби. Визначити імовірність того, що одному з пароплавів доведеться чекати звільнення причалу, якщо час стоянки першого пароплава дорівнює одній годині, а другого – двом годинам.

Розв'язання. Нехай X та Y – час прибуття пароплавів. Можливі значення X та Y: $0 \le X \le 24$; $0 \le Y \le 24$. Сприятливі значення: $Y - X \le 1$; $X - Y \le 2$. Побудуємо цю область (рис. 3.2). Відношення площі заштрихованої фігури $m(g)$ до площі квадрата, сторона якого дорівнює 24 згідно геометричного означення, дорівнює шуканій імовірності:

$$p = \frac{24 \cdot 24 - 0{,}5 \cdot 23 \cdot 23 - 0{,}5 \cdot 22 \cdot 22}{24 \cdot 24} = 0{,}121.$$

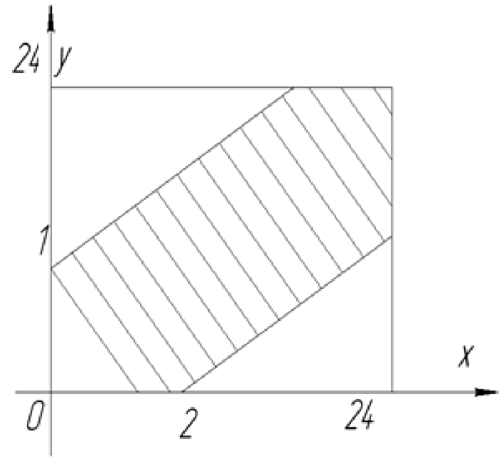


Рис. 3.2. Область сприятливих подій

ПИТАННЯ ДО САМОСТІЙНОГО ВИВЧЕННЯ

1. Статистична ймовірність.
2. Аксіоми теорії ймовірностей.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Назвати протилежні події до таких подій: A – випадіння 2 гербів під час підкидання 2 монет; B – поява білої кульки під час виймання однієї кульки з урни, в якій 2 білі, 3 чорні і 4 червоні кульки; C – три попадання під час 3 пострілів; D – хоча б одне попадання під час 5 пострілів; E – не більше двох попадань під час 5 пострілів; F – виграш першого гравця у грі в шахи.

Задача 2. У ящику 15 деталей, з них 6 – браковані. Взяли навмання 1 деталь. Яка ймовірність, що деталь стандартна?

Задача 3. Маємо 3 лампочки, кожна з яких може перегоріти. Описати ПЕП цього експерименту та події $A = \{$перегоріло не більше однієї лампочки$\}$, $B = \{$перегоріло не менше двох лампочок$\}$. Обчислити ймовірності $P(A)$, $P(B)$.

Задача 4. На площині накреслені два концентричні кола, радіуси яких 15 і 20 см відповідно. Знайти ймовірність того, що точка, кинута на вдачу у великий круг, попаде в кільце, утворене побудованими колами. Припускається, що ймовірність попадання точки у плоску фігуру пропорційна площі цієї фігури і не залежить від її розташування відносно великого круга.

Задача 5. У ящику знаходяться кульки пронумеровані від 1 до 200. Навмання дістають одну кульку. Яка ймовірність того, що це число буде кратне числам a та b (N – номер варіанта)?

1–5 варіанти: $a = 15 - N,\ b = -22N$;

6–10 варіанти: $a = 7 - N,\ b = N$;

11–15 варіанти: $a = 13 - N,\ b = 7N$.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Дати класичне означення ймовірності випадкової події.
2. Що називається відносною частотою випадкової події?
3. Що називається геометричною ймовірністю?
4. Що таке статистична ймовірність?

ТЕМА 4
НЕЗАЛЕЖНІСТЬ ПОДІЙ.
ФОРМУЛИ ПОВНОЇ ЙМОВІРНОСТІ ТА БАЙЄСА

Ймовірність – ступінь вірогідності
і відрізняється від неї,
як частина від цілого.
Я. Бернуллі

Випадкові A та B називають **залежними,** якщо ймовірність появи однієї з них залежить від появи або непояви другої події. Якщо ймовірність появи однієї події не залежить від появи або непояви другої, то такі події називають **незалежними.**

Ймовірність події B, обчислена за умови появи події A, називають **умовною ймовірністю події *B* за умови, що подія *A* відбулася,** і позначають:

$$P(B\backslash A) \quad \text{або} \quad P_A(B).$$

Приклад 4.1. В урні 10 куль: 3 білі і 7 чорних. Навмання беруть дві кулі. Яка ймовірність того, що друга куля буде чорною?

Розв'язання. Нехай подія A – взята біла куля; подія B – взята чорна. Якщо кулю, яку взяли першою, **повертають** до урни, то ймовірність появи другої кулі **не залежить** від того, яка взята перша куля. Якщо перша куля не повертається до урни, то ймовірність другої події залежить від результату першого випробування.

Якщо першою взяли білу кулю, то в урні залишилося 2 білі кулі та 7 чорних, тому $P_A(B) = \frac{7}{9}$. Якщо першою взяли чорну кулю, то в урні залишилося 3 білі кулі та 6 чорних, тому $P_A(B) = \frac{6}{9} = \frac{2}{3}$. Отже, ймовірність події B залежить від появи або непояви події A.

Ймовірність об'єднання двох випадкових несумісних подій дорівнює сумі їх імовірностей:

$$P(A \cup B) = P(A) + P(B). \tag{4.1}$$

Приклад 4.2. Імовірність влучення стрільцем у першу область мішені дорівнює 0,45, у другу область – 0,35, у третю – 0,15. Знайти ймовірність того, що при одному пострілі стрілець влучить у першу або другу області мішені.

Розв'язання. Позначимо за подію A_1 – влучення у першу область мішені; за подію A_2 – влучення у другу область мішені. Під час одного пострілу події A_1 та A_2 несумісні, тому ймовірність влучення в першу або другу області мішені буде:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) = 0{,}45 + 0{,}35 = 0{,}8.$$

Приклад 4.3. Проводиться стрільба по деякій області D мішені, яка містить три зони, які не перетинаються між собою. Ймовірність попадання в зону I: $P(A_1) = \frac{5}{100}$, в зону II: $P(A_2) = \frac{10}{100}$, в зону III: $P(A_3) = \frac{17}{100}$. Яка ймовірність попадання в область D?

Розв'язання. Подія A – попадання в область D. За формулою (4.1) маємо:

$$P(\mathrm{A}_1) + P(\mathrm{A}_2) + P(\mathrm{A}_3) = \frac{5}{100} + \frac{10}{100} + \frac{17}{100} = \frac{32}{100}.$$

Випадкові події утворюють **повну групу,** якщо під час кожного повторення випробування повинна відбутися хоча б одна з них.

Приклад 4.4. Нехай X – число очок, які випадають на верхній грані гральної кістки. Утворюють повну групу:

а) події X=1, X=2, X=3, X=4, X=5, X=6;

б) події X – парне, Х – непарне;

в) події X=1; 1<X<6, X=6.

Сума ймовірностей повної групи випадкових подій дорівнює одиниці:

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1. \quad (4.2)$$

Приклад 4.5. Фірма має три склади. Група товарів прибуває на один зі складів. Ймовірність призначення товарів для першого складу p_1=0,4, а для другого p_2 = 0,35. Визначити ймовірність p_3 того, що ці товари призначені для третього складу.

Розв'язання. Розглянуті події єдиноможливі і несумісні. Тому сума їх ймовірностей дорівнює одиниці: $p_1 + p_2 + p_3 = 0{,}4 + 0{,}35 + p_3 = 1$, звідси:

$$p_3 = 1 - (0{,}4 + 0{,}35) = 025.$$

Дві протилежні події A **та** $\overline{\mathrm{A}}$ **утворюють повну групу,** тому виконується рівність $P(A + \overline{\mathrm{A}}) = 1$, з якої одержуємо формулу **знаходження ймовірності протилежної події:**

$$P(\overline{\mathrm{A}}) = 1 - P(\mathrm{A}). \quad (4.3)$$

Приклад 4.6. Ймовірність одержати повідомлення від певної особи протягом доби дорівнює 0,25. Знайти ймовірність того, що повідомлення протягом доби від цієї особи не буде одержано.

Розв'язання. Позначимо за подію A – повідомлення від цієї особи протягом доби надійде. За умовою задачі виконується співвідношення $P(A) = 0{,}25$. Протилежна подія $\overline{\mathrm{A}}$ означає, що протягом доби від цієї особи повідомлення не надійде. За формулою (4.3) одержимо $P(\bar{A}) = 1 - 0{,}25 = 0{,}75$.

Якщо випадкова подія A може з'явитись лише сумісно з однією із несумісних між собою подій $B_1, B_2, \ldots B_n$, що утворюють повну групу, тоді імовірність події A обчислюється за **формулою повної ймовірності:**

$$P(A) = \sum_{k=1}^{n} P(B_k) \cdot P_{B_k}(A). \quad (4.4)$$

Приклад 4.7. У першому ящику 20 деталей, з яких 15 стандартні. У другому ящику 10 деталей, із яких 9 стандартні. Із другого ящика беруть навмання одну деталь і перекладають її до першого ящика. Знайти ймовірність того, що взята після цього навмання деталь із першого ящика стандартна.

Розв'язання. Позначимо такі події: A – з першого ящика взято стандартну деталь; B_1 – з другого ящика переклали до першого стандартну деталь; B_2 – з другого ящика переклали до першого нестандартну деталь. Згідно з умовою задачі, з першого ящика можна взяти деталь лише після того, як здійсниться подія B_1 або подія B_2. Події B_1 та B_2 несумісні, а подія A може з'явитись лише сумісно з однією із них. Тому для знаходження ймовірності події A можна використати формулу повної імовірності (4.4), яка у цьому випадку набуде вигляду:

$$P(A) = P(\mathrm{B}_1)P_{B_1}(A) + P(\mathrm{B}_2)\mathrm{P}_{\mathrm{B}_2}(A).$$

Знайдемо потрібні ймовірності:

$P(\mathrm{B}_1) = \frac{9}{10}$; $P(\mathrm{B}_2) = \frac{1}{10}$; $P_{\mathrm{B}_1}(A) = \frac{16}{21}$; $P_{\mathrm{B}_2}(A) = \frac{15}{21}$.

Підставимо ці значення у формулу $P(A)$, отримаємо:

$$P(A) = \frac{9}{10}\cdot\frac{16}{21} + \frac{1}{10}\cdot\frac{15}{21} = \frac{144+15}{210} = \frac{53}{70}.$$

Порівняння ймовірностей $P(B_к)$ та $P_A(B_к)$ дає змогу **переоцінити ймовірність гіпотези** за умови, що подія A з'явилася.

$$P_{\mathrm{A}}(\mathrm{B}_{\mathrm{к}}) = \frac{\mathrm{P}(\mathrm{B}_{\mathrm{к}})\mathrm{P}_{\mathrm{B}_{\mathrm{к}}}(\mathrm{A})}{\mathrm{P}(\mathrm{A})}.$$

$$P_{\mathrm{A}}(\mathrm{B}_{\mathrm{к}}) = \frac{\mathrm{P}(\mathrm{B}_{\mathrm{к}})\mathrm{P}_{\mathrm{B}_{\mathrm{к}}}(\mathrm{A})}{\sum_{\mathrm{к}=1}^{n}\mathrm{P}(\mathrm{B}_{\mathrm{к}})\mathrm{P}_{\mathrm{B}_{\mathrm{к}}}(\mathrm{A})}, \quad k = 1, 2, 3, \dots, n.$$

Ці формули називають **формулами Байєса.**

Приклад 4.8. Деталі, виготовленні цехом заводу, надходять для перевірки їх стандартності до одного з двох контролерів. Ймовірність того, що деталь надійде до першого контролера, дорівнює 0,6, а до другого – 0,4. Ймовірність того, що придатна деталь буде визнана стандартною першим контролером, дорівнює 0,94, а другим – 0,98. Придатна деталь під час перевірки визнана стандартною. Знайти ймовірність того, що деталь перевіряв перший контролер.

Розв'язання. Позначимо такі події: A – придатна деталь визнана стандартною; B_1 – деталь перевіряв перший контролер; B_2 – деталь перевіряв другий контролер. За умовою задачі отримаємо:

$P(B_1)$=0,6; $P(B_2)$=0,4;

$P_{\mathrm{B}_1}(A) = 0{,}94$; $P_{\mathrm{B}_2}(A) = 0{,}98$.

За формулою Байєса при $к = 1$ отримаємо:

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A)} = \frac{0{,}6\cdot 0{,}94}{0{,}6\cdot 0{,}94 + 0{,}4\cdot 0{,}98} = 0{,}59.$$

Зазначимо, що до появи події A імовірність $P(B_1) = 0{,}6$, а після появи події A ймовірність перевірки деталі першим контролером $P_A(B_1) = 0{,}59$ поменшала.

Відповідь: $P_A(B_1) = 0{,}59$.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. У магазин надійшла партія взуття одного фасону і розміру, але різного кольору. Партія містить 40 пар чорного кольору, 26 – коричневого, 22 – червоного і 12 пар синього. Коробки із взуттям виявились невідсортовані за кольором. Яка ймовірність того, що навмання взята коробка виявиться із взуттям червоного або синього кольору?

Задача 2. У банку працює 10 співробітників, 8 із яких є консультантами. Знайти ймовірність того, що серед навмання вибраних двох співробітників хоча б один буде консультантом.

Задача 3. До мінімаркету з п’ятьма відділами прибував товар до одного з них. Ймовірність призначення товару для першого відділу p_1 = 0,15, для другого p_2 = 0,25, для третього p_3 = 0,2, а для четвертого p_4 = 0,1. Знайти ймовірність p_5 того, що цей товар призначений для п’ятого відділу.

Задача 4. У графіку руху потягів на дільниці є 120 колій для вантажних потягів. З цієї дільниці на станцію прибувають за розбіркою 80 потягів. Знайти ймовірність прибуття двох розбіркових потягів по двох сусідніх колій.

Задача 5. Ймовірність виготовлення стандартного виробу даним станком дорівнює 0,9. Ймовірність появи виробу першого ґатунку серед стандартних виробів становить 0,8. Визначити ймовірність виготовлення виробу першого ґатунку даним станком.

Задача 6. У групі з 10 здобувачів, які прийшли на іспит, 3 підготовлені відмінно, 4 – добре, 2 – посередньо і 1 – погано. В екзаменаційних білетах є 20 питань. Здобувач, який підготовлений відмінно, може відповісти на всі 20 питань, який підготовлений добре – на 16, посередньо – на 10, погано – на 5. Викликаний навмання здобувач відповів на три довільно задані питання. Знайти ймовірність того, що цей здобувач підготовлений: а) відмінно; б) погано.

Задача 7. На трьох автоматизованих лініях виготовляють однакові деталі, причому 40 % – на першій лінії, 30 % – на другій та 30 % – на третій. Ймовірність виготовлення стандартної деталі для цих ліній становить відповідно 0,9, 0,95 та 0,95. Виготовлені деталі надходять на склад. Яка ймовірність того, що навмання взята деталь стандартна?

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Які події називаються сумісними? Несумісними?
2. Сформулювати теорему додавання для двох сумісних подій.
3. Сформулювати теорему додавання для двох несумісних подій.
4. Які події називаються залежними? Незалежними?
5. Яка ймовірність називається умовною?
6. Сформулювати теорему множення для двох залежних подій.
7. Сформулювати теорему множення для двох незалежних подій.
8. Формула для обчислення ймовірності появи принаймні однієї з групи несумісних подій.
9. Які події називаються гіпотезами?
10. Які події утворюють повну групу?
11. Сформулювати теорему для обчислення повної ймовірності.
12. Формула Байєса та її застосування.

ТЕМА 5
ГРАНИЧНІ ТЕОРЕМИ У СХЕМІ БЕРНУЛЛІ ТА ЇХ НАСЛІДКИ

Науку цю у давнині
з азартних ігор почали,
віками потім розвивали,
розумні люди на Землі,
щоб Ви ЇЇ застосували.

Закон великих чисел стверджує, що частота $\frac{m}{n}$ події *A* буде скільки завгодно близькою до її ймовірності p, якщо число випробувань *n* необмежено зростає.

Якщо в кожному незалежному випробуванні ймовірність настання події *A* одна й та сама і дорівнює $p(0 < p < 1)$, тобто *P(A)* = *p*, $P(\bar{A}) = q = 1 - p$, і не залежить від номера випробування, то такі випробування називаються **схемою незалежних випробувань, схемою Бернуллі.**

Ймовірність того, що в *n* повторних незалежних випробуваннях, у кожному з яких ймовірність появи випадкової події *A* рівна *p* $(0 < p < 1)$, ця подія відбудеться рівно *m* разів, знаходиться за **формулою Бернуллі:**

$$P_n(m) = C_n^m p^m q^{n-m}, \tag{5.1}$$

де $q = 1 - p$ – ймовірність непояви події *A* в кожному випробуванні.

Приклад 5.1. Якою повинна бути ймовірність влучення під час одного пострілу, щоб під час чотирьох пострілів $P(x = 0) = P(x = 1)$?

Розв'язання. Використовуючи формулу Бернуллі, маємо $\mathrm{P}_4(k = 0) = \mathrm{P}_4(k = 1)^4$або $C_4^0 p^0 (1 - p)^4 = C_4^1 p^1 (1 - p)^3$. Звідси 1 – *p* = 4*p*. Отже, $p =$ 0,2.

Зауваження 1.

Ймовірність появи події *A* в *n* випробуваннях схеми Бернуллі *менше **m** разів* знаходять за формулою:

$$P_n(k < m) = P_n(0) + P_n(1) + P_n(2) + \cdots + P_n(m - 1). \tag{5.2}$$

Ймовірність появи події *A* *не менше **m** разів* можна знайти за формулою:

$$P_n(k \geq m) = P_n(m) + P_n(m + 1) + P_n(m + 2) + \cdots + P_n(n)$$

або за формулою:

$$\mathrm{P}_n(k \geq m) = 1 - \sum_{k=0}^{m-1} P_n(k). \tag{5.3}$$

Ймовірність появи події *A* *хоча б один раз* у ***n** випробуваннях* доцільно знаходити за формулою:

$$P_n(1 \leq m \leq n) = 1 - q^n. \tag{5.4}$$

Приклад 5.2. Прилад складено з 10 блоків, надійність кожного з них 0,8. Блоки можуть виходити з ладу незалежно один від одного. Знайти ймовірність того, що:

а) відмовлять два блоки;

б) відмовить хоча б один блок;

в) відмовлять не менше двох блоків.

Розв'язання. Позначимо за подію A відмову блока. Тоді ймовірність події A за умовою буде:

$$P(A) = p = 1 - 0{,}8 = 0{,}2, \text{ тому } \quad q = 1 - p = 1 - 0{,}2 = 0{,}8.$$

Згідно з умовою задачі $n = 10$. Використовуючи формулу Бернуллі та зауваження 1, одержимо:

а) $\mathrm{P}_{10}(2) = C_{10}^{2}\mathrm{p}^2 p^8 = C_{10}^{2}(0{,}2)^2(0{,}8)^8 = 0{,}302$;

б) $\mathrm{P}_{10}(1 \le m \le 10) = 1 - P_{10}(0) = 1 - C_{10}^{0}(0{,}2)^0(0{,}8)^{10} = 0{,}8926$;

в) $\mathrm{P}_{10}(2 \le m \le 10) = 1 - (P_{10}(0) + \mathrm{P}_{10}(1)) = 1 - (C_{10}^{0}(0{,}2)^0(0{,}8)^{10} +$
$+\mathrm{C}_{10}^{1}(0{,}2)^1(0{,}8)^9) = 0{,}6244$.

Число m_0, за якого при заданому n відповідає максимальна біноміальна ймовірність $P_n(m_0)$, називається **найімовірнішим числом появи події *A*.**

Найімовірніше число *m*₀ задовольняє систему нерівностей:

$$np - q \le m_0 \le np + p \quad \text{або} \quad (n+1)p - 1 \le m_0 \le (n+1)p. \qquad (5.5)$$

Якщо $np - q$ – неціле, то є одне значення m_0, якщо $np - q$ – ціле, то таких значень є два, які будуть відрізнятися між собою на 1:

$$m_1 = (n+1)p - 1 \quad \text{та} \quad m_2 = (n+1)p.$$

Приклад 5.3. Частка виробів вищого ґатунку на даному підприємстві становить 26 %. Чому дорівнює найімовірніше число виробів вищого ґатунку у випадково відібраній партії з 68 виробів?

Розв'язання. За умовою маємо, що $p = 0{,}26$, $n = 68$, а $q = 1 - p = 1 -$
$-0{,}26 = 0{,}74$:

$$68 \cdot 0{,}26 - 0{,}74 \le m_0 \le 68 \cdot 0{,}26 + 0{,}26.$$
$$16{,}94 \le m_0 \le 17{,}94.$$

Отже, $m_0 = 17$.

Приклад 5.4. Під час нового технологічного процесу 80 % усієї виготовленої продукції має найвищу якість. Знайти найбільш ймовірне число виготовлених виробів найвищої якості серед 250 виготовлених виробів.

Розв'язання. Позначимо шукане число m_0, тоді отримаємо:

$$np - q \le m_0 \le np + q.$$

За умовою маємо, що $n = 250, p = 0{,}8, q = 0{,}2$, тому: $199{,}8 \le m_0 \le 200{,}8$.

Але m_0 повинно бути цілим числом, тому $m_0 = 200$.

Якщо ймовірність появи події A в кожному випробуванні дорівнює p, то кількість n випробувань, які необхідно здійснити, щоб зі ймовірністю P можна було стверджувати, що подія A **з'явиться хоча б один раз,** знаходять за формулою:

$$n > \frac{ln(1-P)}{ln(1-p)}. \qquad (5.6)$$

Приклад 5.5. За годину автомат виготовляє 20 деталей. За скільки годин ймовірність виготовлення хоча б однієї бракованої деталі буде не менша 0,952, якщо ймовірність браку будь-якої деталі дорівнює 0,01?

Розв'язання. Застосовуючи формулу (5.6), знайдемо спочатку таку кількість виготовлених деталей, щоб зі ймовірністю $p = 0{,}952$ можна було стверджувати про наявність хоча б однієї бракованої деталі, якщо ймовірність браку за умовою $p = 0{,}01$:

$$n \geq \frac{ln(1-0{,}952)}{ln(1-0{,}01)} = \frac{ln\,0{,}048}{ln\,0{,}99} \approx 300.$$

Отже, за час $t = \frac{300}{20} = 15$ (годин) автомат зі ймовірністю 0,952 виготовить хоча б одну браковану деталь.

Теорема Пуассона

Якщо $n \to \infty$ і $p \to 0$ так, що $n \to \lambda, 0 < \lambda < \infty$, то:

$$P_n(m) = C_n^m p^m q^{n-m} \to \frac{\lambda^m}{m!} e^{-\lambda}, \text{де } \lambda = np \qquad (5.7)$$

для будь-якого постійного $m = 0, 1, 2, \ldots$

Ймовірність появи події A m разів у n випробуваннях схеми Бернуллі можна знайти за наближеною **формулою Пуассона:** $P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}$, де $\lambda = np$.

Формулу доцільно застосовувати при великих n та малих p.

Приклад 5.6. На факультеті 730 здобувачів. Яка ймовірність того, що 1 вересня є днем народження одночасно трьох здобувачів?

Розв'язання. Ймовірність того, що днем народження окремого студента є 1 вересня, дорівнює $p = \frac{1}{365}$. Застосовуючи формулу (5.7), де: $n = 730, p = \frac{1}{365}$, $m = 3, \lambda = np = 730 \cdot \frac{1}{365} = 2$, дістаємо:$P_{730}(3) = \frac{2^3}{3!} e^{-2} = 0{,}1805$.

Отже, в середньому у 18 випадках зі 100 на 1 вересня припадає день народження 3 здобувачів факультету.

Локальна теорема Муавра–Лапласа

Якщо у схемі Бернуллі кількість випробувань n достатньо велика, *ймовірність* p появи події A в усіх випробуваннях *однакова*, то ймовірність появи події A m разів може бути знайдена за наближеною формулою:

$$P_n(m) = \frac{1}{\sqrt{npq}} \varphi(x_m), \;\; \text{де } \; x_m = \frac{m-np}{\sqrt{npq}}. \qquad (5.8)$$

Формулу доцільно використовувати при $n > 100$ та $npq > 20$.

Приклад 5.7. Ймовірність виготовлення деталі вищого ґатунку на цьому верстаті дорівнює 0,4. Знайти наближено ймовірність того, що серед навмання взятих 26 деталей половина виявиться вищого ґатунку.

Розв'язання. З умови маємо $p = 0{,}4, \; q = 0{,}6, \; n = 26, \; m = 13$. Тоді:

$$np = 26 \cdot 0{,}4 = 10{,}4, npq = 10{,}4 \cdot 0{,}6 = 6{,}24, \sqrt{npq} = \sqrt{6{,}24} \approx 2{,}5,$$

$$m - np = 13 - 10{,}4 = 2{,}6, \text{ тому } x = \frac{m-np}{\sqrt{npq}} = \frac{2{,}6}{2{,}5} = 1{,}04.$$

Значення $\varphi(x)$ при $x = 1{,}04$ знаходимо з таблиці $\varphi(\mathrm{x}) = \varphi(1{,}04) = 0{,}2323$.

$$P_{26}(13) = \frac{\phi(\mathrm{x})}{\sqrt{npq}} = \frac{0{,}2323}{2{,}5} = 0{,}093.$$

Інтегральна теорема Муавра–Лапласа

Якщо у схемі Бернуллі в кожному із n незалежних випробувань подія A

може з'явитися з постійною імовірністю p, яка відмінна від нуля і одиниці ($0 \leq p \leq 1$), тоді ймовірність $P_n(m_1 \leq m \leq m_2)$ появи події A не менше m_1 та не більше m_2, приблизно дорівнює визначеному інтегралу:

$$P_n(m_1 \leq m \leq m_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{z^2}{2}} dz = \Phi(x_2) - \Phi(x_1), \tag{5.9}$$

де $x_2 = \frac{m_2 - np}{\sqrt{npq}}, x_1 = \frac{m_1 - np}{\sqrt{npq}}$.

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ – інтегральна функція Лапласа, значення якої табульовані.

Приклад 5.8. Гральний кубик кидають 800 разів. Яка ймовірність того, що кількість очок, кратна трьом, з'явиться не менше 260 та не більше 274 разів?

Розв'язання. Для знаходження ймовірності $\mathrm{P}_{800}(260 \leq m \leq 274)$:

$$x_2 = \frac{274 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = \frac{22}{40} = 0{,}55, \qquad x_1 = \frac{260 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = -0{,}5,$$

$$\mathrm{P}_{800}(260 \leq m \leq 274) = \Phi(0{,}55) - \Phi(-0{,}5) = \Phi(0{,}55) + \Phi(0{,}5) =$$
$$= 0{,}2988 + 0{,}1915 = 0{,}4003.$$

Відповідь: $\mathrm{P}_{800}(260 \leq m \leq 274) = 0{,}4003$.

Теорема Я. Бернуллі

Якщо у n незалежних випробуваннях імовірність p появи події A однакова, і подія A з'явилася m разів, то для будь-якого додатного числа a виконується рівність:

$$\lim_{n \to \infty} P\left(\left|\frac{m}{n} - p\right| \geq a\right) = 0, \tag{5.10}$$

тобто границя ймовірності відхилення відносної частоти $\frac{m}{n}$ події A від її ймовірності на величину, що більше або дорівнює a, дорівнює нулю.

«Омана гравця»

Якщо ви підкинули монету 10 разів і отримали 10 «гербів», чи збільшилась від цього ймовірність, що на 11-му кидку випаде «цифра»? Або чи буде ймовірність виграшу в лотерею комбінації з шести послідовних чисел від 1 до 6 нижча, ніж у будь-якій іншій комбінації?

Не ставайте жертвою «помилки гравця»! Те, що сталось, уже ніяк не впливає на результат незалежної події. Ніяк. *Ніколи.* У по-справжньому випадковій лотереї, що розігрується, ймовірність випадання будь-якого конкретного числа точно така сама, як будь-якого іншого. Немає жодної закономірності, відповідно до якої числа, ті, що рідко випадали в минулому, повинні частіше випадати в майбутньому.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Ймовірність знаходження в кожному прибулому потязі вагонів на задане призначення – 0,2. Визначити ймовірність того, що в трьох із п’яти потягів, які прибувають протягом однієї години, будуть вагони на задане призначення.

Задача 2. Знайти ймовірність того, що у п’яти незалежних випробуваннях подія A відбудеться:

а) рівно 4 рази;

б) не менше 4 разів, якщо в кожному випробуванні ймовірність появи події становить 0,8.

Відповідь: а) 0,4096; б) 0,73728.

Задача 3. На автомобільному заводі у звичному режимі роботи з конвеєра сходить 100 000 автомобілів. Ймовірність бракованого автомобіля дорівнює 0,0001. Знайти ймовірність того, що з конвеєра зійшло 5 бракованих автомобілів.

Задача 4. Банк обслуговує 100 клієнтів, від кожного з яких може надійти вимога на проведення фінансової операції на наступний день зі ймовірністю 0,4. Знайти найімовірніше число вимог клієнтів кожного дня.

Задача 5. Завод випускає в середньому 4 % нестандартних виробів. Яка ймовірність того, що кількість нестандартних виробів у партії з 4 000 штук не більше 170?

Задача 6. Яка ймовірність того, що під час 10 000 незалежних підкидань монети «герб» випаде 5 000 разів?

Задача 7. Фірма відправила на базу 1 000 якісних виробів. Ймовірність того, що вироби в дорозі пошкодяться, дорівнює 0,002. Знайти ймовірність того, що на базу прибуде 5 пошкоджених виробів.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Яка послідовність випробувань утворює схему Бернуллі?
2. Як можна знайти найбільш імовірнісне значення числа появ події A у схемі Бернуллі?
3. У яких випадках доцільно використовувати граничні теореми у схемі Бернуллі?
4. Коли доцільно застосовувати формули Пуассона, локальну або інтегральну формули Муавра–Лапласа?
5. Як визначаються і які мають властивості локальна та інтегральна функції Лапласа?
6. Як формулюється теорема Бернуллі і який вона має наслідок?

ТЕМА 6
ВИПАДКОВІ ВЕЛИЧИНИ, ЇХ ТИПИ

Істинна логіка нашого світу –
правильний підрахунок ймовірностей.
М. М. Моисеєв

Випадкові події – події, наслідком яких є поява деякого числа, заздалегідь невідомого.

Випадковою величиною (ВВ) називають таку змінну величину, яка внаслідок випробування може набувати певних значень з відповідними ймовірностями.

Випадкова подія – якісна характеристика випробування, а випадкова величина – кількісна характеристика випробування.

Випадкові величини бувають дискретними та неперервними.

Дискретною випадковою величиною (ДВВ) називається випадкова величина, що набуває скінчене число значень (відокремлені ізольовані одне від одного числові значення з відповідними ймовірностями) з множини, елементи якої можна пронумерувати.

Неперервною випадковою величиною (НВВ) називається випадкова величина, можливі значення якої неперервно заповнюють деякий інтервал (тобто величина, яка може набути будь-якого числового значення з деякого скінченого або нескінченого інтервалу (a, b). Кількість можливих значень такої величини є нескінченна.

Законом розподілу випадкової величини є співвідношення, що встановлює зв'язок між можливими значеннями випадкової величини і відповідними їм ймовірностями.

Закон розподілу може бути задано функціями:

- **інтегральною функцією розподілу** $F(x)$ $F(x) = P(X < x)$;
- **функцією щільності розподілу** $f(x)$ або диференціальною функцією $f(x) = P(X = x)$;
- **в аналітичній формі для *дискретної* змінної функція розподілу** $P(x)$ матиме вигляд, наприклад:

$$F(x) = \begin{cases} 0{,}05, & x \le 1; \\ 0{,}15, & x \le 2; \\ 0{,}35, & x \le 3; \\ 0{,}60, & x \le 4; \\ 0{,}90, & x \le 5; \\ 1{,}00, & x \le 6. \end{cases}$$

Дискретною випадковою величиною (ДВВ) називається випадкова величина, що набуває скінченої кількості значень (відокремлені ізольовані одне від одного числові значення з відповідними ймовірностями) з множини, елементи якої можна пронумерувати.

Приклад 6.1. Експерт із банківського кредитування наголосив, що протягом місяця фірма A ліквідує свою заборгованість зі ймовірністю 0,8; фірма B – зі ймовірністю 0,9; а фірма C – зі ймовірністю 0,7. Скласти закон розподілу випадкової величини X – кількості фірм, які ліквідують заборгованість протягом місяця.

Розв'язання. ДВВ X може набувати таких значень: $x_1 = 0, x_2 = 1, x_3 = 2$, $x_4 = 3$.

Знайдемо їх ймовірності:

$P(X = 0) = 0{,}2 \cdot 0{,}1 \cdot 0{,}3 = 0{,}006.$

$P(X = 1) = 0{,}8 \cdot 0{,}1 \cdot 0{,}3 + 0{,}2 \cdot 0{,}9 \cdot 0{,}3 + 0{,}2 \cdot 0{,}1 \cdot 0{,}7 = 0{,}092.$

$P(X = 2) = 0{,}8 \cdot 0{,}9 \cdot 0{,}3 + 0{,}8 \cdot 0{,}1 \cdot 0{,}7 + 0{,}2 \cdot 0{,}9 \cdot 0{,}7 = 0{,}398.$

$P(X = 3) = 0{,}8 \cdot 0{,}9 \cdot 0{,}7 = 0{,}504.$

Запишемо закон розподілу:

x_i	0	1	2	3
p_i	0,006	0,092	0,398	0,504

За законом нормування перевіримо:

$$p_1 + p_2 + p_3 + p_4 = 0{,}006 + 0{,}092 + 0{,}398 + 0{,}504 = 1.$$

Числові характеристики дискретних випадкових величин

Найбільш часто використовують три числові характеристики:

1) математичне сподівання;
2) дисперсія;
3) середнє квадратичне відхилення.

Математичне сподівання та його властивості

Математичним сподіванням ДВВ X називається число, яке дорівнює сумі добутків усіх можливих значень X на відповідні їм ймовірності.

Математичне сподівання ДВВ X позначають $M(X)$ або m_X, тобто:

$$M(X) = \sum_{k=1}^{n} x_k \cdot p_k.$$

Якщо X приймає нескінченну кількість значень, то:

$$M(X) = \sum_{k=1}^{\infty} x_k \cdot p_k.$$

Основні властивості математичного сподівання:

1. Математичне очікування сталої величини дорівнює самій сталій:

$$M(C) = C, \text{ де } C = const.$$

2. Сталий множник можна виносити за знак математичного очікування:

$$M(CX) = C \cdot M(X).$$

3. Математичне очікування добутку взаємно незалежних дискретних випадкових величин дорівнює добутку математичних очікувань співмножників:

$$M(X_1 \cdot X_2 \cdot \ldots \cdot X_n) = M(X_1) \cdot M(X_2) \cdot \ldots \cdot M(X_n).$$

4. Математичне очікування суми ДВВ дорівнює сумі математичних очікувань значень:

$$M(X_1 + X_2 \cdot +..+X_n) = M(X_1) + M(X_2)+\ldots+M(X_n).$$

Ймовірнісний зміст математичного очікування: $M(X)$ – середнє арифметичне зважене всіх її можливих значень (математичне сподівання ДВВ X характеризує середнє значення випадкової величини (ВВ) X із врахуванням ймовірностей його можливих значень). Тобто під математичним сподіванням розуміють центр розподілу ВВ.

$M(X)$ – число на числовій осі, в околі якого розташовані всі можливі значення ВВ.

Приклад 6.2. Незалежні випадкові величини X і Y розподілені за таким законом:

X	5	2	4
P	0,6	0,1	0,3

Y	8	10
P	0,8	0,2

Знайти математичне сподівання ВВ $X \cdot Y$.

Розв'язання

$$M(X) = 5 \cdot 0{,}6 + 2 \cdot 0{,}1 + 4 \cdot 0{,}3 = 4{,}4.$$

$$M(Y) = 8 \cdot 0{,}8 + 10 \cdot 0{,}2 = 8{,}4.$$

$$M(\mathrm{X} \cdot Y) = M(X) \cdot M(Y) = 4{,}4 \cdot 8{,}4 = 36{,}96.$$

Приклад 6.3. Знайти математичне сподівання суми числа очок, які можуть з'явитися під час кидання двох гральних кубиків.

Розв'язання. Нехай кількість очок, які можуть з'явитися на першому кубику, – X, на другому – Y. Тоді імовірність кожного з цих значень буде $\frac{1}{6}$.

$$M(X) = M(Y) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

$$M(X + Y) = M(X) + M(Y) = \frac{7}{2} + \frac{7}{2} = 7.$$

Відповідь: $M(X + Y) = 7$.

Дисперсія та її властивості

Дисперсією ДВВ X називається число, яке дорівнює математичному сподіванню квадрата відхилення ДВВ X від її математичного сподівання.

Дисперсією величини X позначають $D(X)$ або D_X:

$$D(X) = M\left((X - M(X))^2\right).$$

Дисперсія ДВВ X дорівнює різниці між математичним сподіванням квадрата ВВ X та квадрата її математичного сподівання:

$$D(X) = M(X^2) - \left(M(X)\right)^2, \qquad \text{де } M(X^2) = \sum_{i=1}^{n} x_i^2 \cdot p_i,$$

$$D(X) = \sum_{i=1}^{n} (x_i - M(x))^2 \cdot p_i.$$

Дисперсія має такі **властивості:**

1. Дисперсія завжди набуває невід'ємних значень: $D(X) \geq 0$.

2. Дисперсія сталої величини дорівнює нулю: $D(C) = 0$.

3. Сталий множник можна виносити за знак дисперсії, підносячи його до квадрата: $D(CX) = C^2 \cdot D(X)$.

4. Дисперсія алгебраїчної суми ДВВ X та Y дорівнює сумі їх дисперсії $D(X \pm Y) = D(X) + D(Y)$.

5. Дисперсія суми декількох незалежних дискретних випадкових величин дорівнює сумі дисперсій значень:

$$D(X_1 + X_2 + \ldots + X_n) = D(X_1) + D(X_2) + \ldots + D(X_n).$$

Приклад 6.4. Знайти дисперсію випадкової величини X, що задана законом:

X	–5	0	4	5
P	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

Розв'язання

$M(X) = -5 \cdot \frac{1}{8} + 0 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 5\frac{1}{8} = 1$; $M^2(X) = 1^2 = 1$.

Тепер знаходимо X^2, тобто $M(X^2)$:

X^2	25	0	16	25
P	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

$M(X^2) = 25 \cdot \frac{1}{8} + 0 \cdot \frac{1}{2} + 16 \cdot \frac{1}{4} + 25\frac{1}{8} = \frac{82}{8}$, тому $D(X) = \frac{82}{8} - 1 = \frac{74}{8} = 9{,}25$.

Середнє квадратичне відхилення ДВВ

Середнє квадратичне відхилення дорівнює квадратному кореню з дисперсії і позначається $\sigma(X) = \sigma_X = \sqrt{D(X)}$.

Основна властивість середнього квадратичного відхилення: якщо $X =$ $= X_1 + X_2 + \ \ldots + \ X_n$ де $X_1, X_2, \ldots, X_n$ – взаємно незалежні випадкові величини, то $\sigma(X) = \sqrt{= \sigma^2(X_1) + \sigma^2(X_2) + \ \ldots + \sigma^2(\,X_n)}$.

Приклад 6.5. Майстерня з виготовлення шоколаду має договори на постачання своєї продукції з трьома магазинами та двома кафе. Ймовірність виконання договору одним магазином становить 0,9, а одним кафе – 0,8. Знайти середню кількість постачальників сировини, які виконають договори.

Розв'язання. Нехай X – кількість магазинів, Y – кількість кафе, що виконають договори. Складемо закони розподілу цих ДВВ:

Для випадкової величини X маємо: $p = 0{,}9$, $q = 0{,}1$.

Можливі значення X: $x_0 = 0,\ x_1 = 1,\ x_2 = 2,\ x_3 = 3$.

Ймовірності обчислюємо за формулою Бернуллі:

$p_0 = C_3^0 0{,}9^0 0{,}1^3. = 0{,}001$;

$p_1 = C_3^1 0{,}9^1 0{,}1^2. = 0{,}027$;

$p_2 = C_3^2 0{,}9^2 0{,}1^1. = 0{,}243$;

$p_3 = C_3^3 0{,}9^3 0{,}1^0. = 0{,}729$.

Перевірка: $p_0 + p_1 + p_2 + p_3 = 0{,}001 + 0{,}027 + 0{,}243 + 0{,}729 = 1$.

Закон розподілу випадкової величини X має вигляд:

x_i	0	1	2	3
p_i	0,001	0,027	0,243	0,729

Тоді за формулою математичного очікування, маємо:

$M(X) = 0 \cdot 0{,}001 + 1 \cdot 0{,}027 + 2 \cdot 0{,}243 + 3 \cdot 0{,}729 = 2{,}7.$

Для випадкової величини Y маємо: $p = 0{,}8$, $q = 0{,}2$.

Можливі значення Y: $y_0 = 0$, $y_1 = 1$, $y_2 = 2$.

Ймовірності обчислюємо за формулою Бернуллі:

$p_0 = C_2^0 0{,}8^0 0{,}2^2 = 0{,}04.$

$p_1 = C_2^1 0{,}8^1 0{,}2^1 = 0{,}32.$

$p_2 = C_2^2 0{,}8^2 0{,}2^0 = 0{,}64.$

Перевірка: $p_0 + p_1 + p_2 = 0{,}04 + 0{,}32 + 0{,}64 = 1.$

Закон розподілу випадкової величини Y має вигляд:

y_i	0	1	2
p_i	0,04	0,32	0,64

Тоді за формулою математичного очікування маємо:

$$M(Y) = 0 \cdot 0{,}04 + 1 \cdot 0{,}32 + 2 \cdot 0{,}64 = 1{,}6.$$

За ймовірнісним змістом математичного очікування середнє число постачальників наближено дорівнює $M(X+Y)$, тобто $M(X+Y) = 1{,}6 + 2{,}7 = 4{,}3.$

Отже, середнє число постачальників дорівнює 4–5.

Приклад 6.6. На першому сегменті ринку дохід з рівними ймовірностями може становити 100 млн грн за умови доброго розпродажу продукції, і 80 млн грн – за середнього. На другому сегменті ринку очікується стабільний дохід у розмірі 93 млн грн. Однак існує ймовірність (0,2) того, що попит різко впаде, і дохід стане 78 млн грн. Необхідно вибрати сегмент, оптимальний з погляду результативності та ризику.

Розв'язання. Спочатку знайдемо математичне очікування доходів під час роботи на кожному із сегментів ринку:

$M(X) = 0{,}5 \cdot 100 + 0{,}5 \cdot 80 = 90$ (млн грн),

$M(Y) = 0{,}8 \cdot 93 + 0{,}2 \cdot 78 = 90$ (млн грн).

Визначимо розкид результатів для кожного з варіантів:

$M(X^2) = 0{,}5 \cdot 100^2 + 0{,}5 \cdot 80^2 = 8200.$

$D(X) = 8200 - 8100 = 100.$ $\sigma(X) = \sqrt{100} = 10$ (млн грн).

$M(Y^2) = 0{,}8 \cdot 93^2 + 0{,}2 \cdot 78^2 = 8136.$

$D(Y) = 8136 - 8100 = 36.$ $\sigma(Y) = \sqrt{36} = 6$ (млн грн).

Отже, перший сегмент більш ризикований за другий (10 млн > 6 млн). Потрібно орієнтуватися на роботу у другому сегменті ринку, оскільки тут такий самий очікуваний результат, як і на першому, але ризик менший.

Резервне зберігання

Вам необхідно організувати зберігання даних впродовж року. Один диск має ймовірність збою 1 на 1 000 000 000. Другий коштує 20 % від ціни першого, але в його випадку ймовірність збою – 1 на 2 000. Який диск вам варто купити?

Якщо ви вирішите для надійності використовувати три дешеві диски (на кожному з яких будуть ті самі дані), то втратите дані, тільки якщо всі три вийдуть із ладу. Ймовірність того, що це трапиться, дорівнює $\left(\frac{1}{2000}\right)^3 = \frac{1}{8000000000}$.

Ризик втрати даних під час збереження їх резервних копій на трьох дешевших дисках здається значно нижчим, ніж у випадку з одним дорогим диском. А заплатите за всі три дешеві диски $0{,}2y + 0{,}2y + 0{,}2y$ всього 60 % від вартості дорогого (y).

ПИТАННЯ ДЛЯ САМОСТІЙНОГО ВИВЧЕННЯ

Двовимірна випадкова величина та її чисельні характеристики.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Пристрій складається з трьох елементів, що працюють незалежно. Ймовірність відмови кожного елемента в одному досліді дорівнює 0,1. Скласти закон розподілу числа елементів, що відмовили в одному досліді.

Задача 2. У партії з 10 деталей є 8 стандартних. Навмання відібрано дві деталі. Скласти закон розподілу числа стандартних деталей серед відібраних.

Задача 3. Статистика за регіоном свідчить, що 5 % працездатних мешканців регіону є безробітними. Навмання вибирають трьох мешканців. Скласти закон розподілу ймовірностей випадкової величини X – кількості безробітних осіб серед навмання відібраних трьох мешканців регіону.

Задача 4. Два бомбардувальники по черзі скидають бомби на об'єкт до першого влучення. Ймовірність влучення для першого бомбардувальника дорівнює 0,7, а для другого – 0,8. Першим починає скидати бомби перший бомбардувальник. Знайти перші чотири члени закону розподілу дискретної випадкової величини X – кількості скинутих бомб обома бомбардувальниками.

Задача 5. Пристрій складається з 1 000 елементів, що працюють незалежно. Ймовірність відмови будь-якого елемента за час t дорівнює 0,002. Знайти перші три члени закону розподілу ДВВ X – відмов елементів за час t.

Задача 7. ДВВ X може приймати лише два значення: x_1 і та x_2, причому $x_1 < x_2$. Відомі ймовірність $p_1 = 0{,}2$ появи можливого значення x_2, математичне сподівання $M(X) = 3{,}8$ та дисперсія $D(X) = 0{,}16$. Знайти закон розподілу цієї випадкової величини.

Задача 8. Розглядають три варіанти виробництва нових товарів. Можливі прибутки від реалізації проєктів у різних ситуаціях, а також ймовірності їх настання подано в таблиці 6.1.

Таблиця 6.1

Таблиця реалізації проєктів

Характеристика ситуації	Можливий дохід (тис. грн)	Ймовірність настання ситуації
Проєкт А		
Песимістична	100	0,1
Найбільш ймовірна	250	0,6
Оптимістична	400	0,3
Проєкт В		
Песимістична	80	0,25
Найбільш ймовірна	300	0,5
Оптимістична	500	0,25
Проєкт С		
Песимістична	90	0,2
Найбільш ймовірна	300	0,6
Оптимістична	400	0,2

Вибрати варіант, оптимальний із погляду результативності та ризику.

Задача 9. Випадкова величина X – кількість підприємців із кожних десяти, які декларують не весь товар під час перетину кордону, розподілена за таким законом:

x_i	0	1	2	3	4	5	6	7	8	9	10
p_i	0,28	0,15	0,1	0,1	0,09	0,09	0,07	0,05	0,04	0,02	0,01

Знайти числові характеристики випадкової величини X.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що називають випадковою величиною?
2. Які види випадкових величин бувають?
3. Сформулювати закон розподілу випадкової величини.
4. Які способи завдання закону розподілу бувають?
5. Що таке умова нормування? З яких причин вона виникає?
6. Навести приклади ДВВ.
7. Які є способи задання законів розподілу ДВВ?
8. Які ви знаєте числові характеристики ДВВ? Як вони обчислюються?
9. Які властивості має математичне очікування ДВВ?
10. Який ймовірнісний зміст математичного очікування ДВВ?

ТЕМА 7
ЗАКОНИ РОЗПОДІЛУ ДИСКРЕТНИХ ВИПАДКОВИХ ВЕЛИЧИН

Біноміальним називається закон розподілу ДВВ, яка приймає значення зі ймовірностями, що обчислюються за схемою Бернуллі. Тобто біноміальний розподіл має випадкова величина, яка дорівнює кількості «успіхів» у серії з n незалежних випробувань, у кожному з яких «успіх» відбувається зі ймовірністю p.

Отже, біноміальний закон розподілу можна задати таблицею:

x_i	0	1	$\dots$	k	$\dots$	n
p_i	q^n	$C_n^1 pq^{n-1}$	$\dots$	$C_n^k p^k q^{n-k}$	$\dots$	p^n

$$q^n + C_n^1 pq^{n-1} + \dots + C_n^k p^k q^{n-k} + \dots + p^n = (p+q)^n = 1^n = 1.$$

$$M(x) = np, D(x) = npq.$$

Приклад 7.1. Ймовірність погашення банківського кредиту кожним клієнтом становить 0,8. Знайти математичне очікування випадкової величини X– кількості клієнтів серед вибраних 10, які своєчасно і в повному обсязі повернуть кредити банку.

Розв'язання. Оскільки своєчасне погашення кредиту одним клієнтом не залежить від того, чи поверне кредит інший, і ймовірність своєчасного погашення кредиту кожним клієнтом є однакова, то маємо послідовність випробувань за схемою Бернуллі. У цьому випадку $n = 10, p = 0{,}8$ і $M(X) = 10 \cdot 0{,}8 = 8$, тобто в середньому 8 із 10-ти клієнтів погасять кредити своєчасно.

Якщо у схемі незалежних повторних випробувань n велике і p прямує до нуля, то біноміальний розподіл апроксимується розподілом Пуассона.

Дискретна випадкова величина має **розподіл Пуассона,** якщо вона набуває зліченної множини значень $m = 1,2,3, \dots$ зі ймовірностями:

$$P(X = m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}, \tag{7.1}$$

де $\lambda = np, (\lambda > 0)$.

Отже, пуассонівський закон розподілу ВВ заданий таблицею 7.1:

Таблиця 7.1

Пуассонівський закон розподілу

x_i	0	1	$\dots$	m	$\dots$
p_i	$e^{-\lambda}$	$\lambda \cdot e^{-\lambda}$	$\dots$	$\frac{\lambda^m}{m!} \cdot e^{-\lambda}$	$\dots$

Цей розподіл описує кількість подій, які настають в однакові проміжки часу за умови, що ці події відбуваються незалежно одна від одної зі сталою інтенсивністю. Закон Пуассона називають *законом рідкісних явищ:*

$$M(X) = D(X) = np.$$

Приклад 7.2. Електронна пошта банку підтримує зв'язки з сотнею абонентів. Ймовірність того, що за одиницю часу на електронну пошту надійде повідомлення від абонента, становить 0,02. Написати закон розподілу величини X– кількості надходження сигналів від абонентів. Яка з подій є найбільш ймовірною?

– B – за одиницю часу надійдуть сигнали від одного абонента,

– C – за одиницю часу надійдуть сигнали від трьох абонентів.

Розв'язання. У цьому випадку проводиться $n = 100$ випробувань за схемою Бернуллі. Випадкова величина X може набувати значень:

$x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3, \ldots, \; x_{100} = 100.$

Ймовірність події A – надходження сигналу від одного абонента є мала, а число $n = 100$ є велике і $\lambda = 100 \cdot 0{,}02 = 2$, тому відповідні ймовірності обчислюються за формулою Пуассона:

$p_0 = P_{100}(0) \approx e^{-2} \approx 0{,}1353;$

$p_1 = P_{100}(1) \approx 2e^{-2} \approx 0{,}2707;$

$p_3 = P_{100}(3) \approx \frac{4}{3}e^{-2} \approx 0{,}1804;$

$p_4 = P_{100}(4) \approx \frac{2}{3}e^{-2} \approx 0{,}0902. \ldots$

$$p_{100} = P_{100}(100) = 1 - \sum_{i=0}^{99} P_{100}\,(i).$$

Закон розподілу описаної в задачі ДВВ X записуємо у формі таблиці 7.2:

Таблиця 7.2

x_i	0	1	2	3	4	…
p_i	0,1353	0,2707	0,2707	0,1804	0,0902	…

Із наведеної таблиці видно, що $p(B) = 0{,}2707$ і $p(C) = 0{,}180$, тобто більш ймовірно, що сигнали надійдуть від одного абонента, ніж від трьох.

Для цього розподілу складено таблиці ймовірностей щодо різних значень $\lambda(0{,}1 - 5)$.

Геометричний закон розподілу має частота настання події у схемі незалежних повторних випробувань, якщо вони проводяться до першого настання події. Ймовірності можливих значень випадкової величини X визначається залежністю:

$$P(X = m) = pq^{m-1}, m = 1,2,3, \ldots,$$

p – ймовірність настання події в кожному випробуванні.

Цей закон називають *геометричним законом розподілу*.

У табличній формі геометричний закон розподілу має вигляд:

x_i	1	2	3	4	…
p_i	p	pq	pq^2	pq^3	…

Під час перевірки закону нормування використовується формула суми нескінченно спадної геометричної прогресії:

$$M(X) = \frac{1}{p}, \;\; D(X) = \frac{1-p}{p^2}.$$

Серед ДВВ лише геометричному закону притаманна властивість відсутності післядії.

Гіпергеометричний розподіл описує ймовірність появи m елементів з певною властивістю серед n елементів, взятих із сукупності N елементів, яка містить саме k елементів такої властивості:

$$p(X=m)=\frac{C_k^m\cdot C_{N-k}^{n-m}}{C_N^n},$$

де $m=1,2,\ldots,n; k\geq n$.

$$M(X)=\frac{kn}{N},\quad D(X)=\frac{nk(N-k)(N-n)}{N^2(N-1)}.$$

Цей закон розподілу застосовується в задачах статистичного контролю якості та в суміжних галузях.

Приклад 7.3. Ймовірність того, що з n деталей, які випадково вибрано з партії обсягом N, m виявляться дефектними, має гіпергеометричний закон розподілу (k – кількість дефектних деталей у партії).

Зі зменшенням відношення $\frac{n}{N}$ гіпергеометричний розподіл наближається до біноміального з параметрами n і $p=\frac{k}{N}$. Дуже часто гіпергеометричний розподіл апроксимується розподілом Пуассона.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Відомо, що серед готівкової маси 0,5 % купюр є непридатними до наступного використання. Скласти закон розподілу величини X – кількості нестандартних купюр серед двох навмання взятих, а також знайти $M(X)$ та $\sigma(X)$.

Задача 2. Фірма відвантажила споживачеві 60 000 виробів. Ймовірність того, що під час транспортування один виріб буде пошкоджений, становить 0,0001. Випадкова величина X – кількість пошкоджених виробів – розподілена за законом Пуассона. Знайти математичне очікування X.

Задача 3. У супермаркет відправлено 3 000 скляних пляшок мінеральної води «Боржомі». Ймовірність того, що під час транспортування пляшка буде розбита, дорівнює 0,001. Знайти:

а) математичне очікування випадкової величини X – кількості розбитих пляшок;

б) дисперсію X – кількості розбитих пляшок.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Коли закон розподілу ДВВ називається біноміальним? Чому?
2. Якими формулами виражаються числові характеристики випадкової величини, що має розподіл Пуассона?
3. Де застосовуються геометричний та гіпергеометричний закони розподілу ймовірностей?
4. Навести приклади ДВВ, розподіленої за гіпергеометричним законом розподілу ймовірностей.

ТЕМА 8
ЗАКОНИ РОЗПОДІЛУ НЕПЕРЕРВНИХ ВИПАДКОВИХ ВЕЛИЧИН

...де на поверхні проходить гра випадковості,
там сама ця випадковість
завжди виявляється підпорядкованою
внутрішнім, прихованим законом.
Уся справа тільки в тому, щоб
відкрити ці закони.

Неперервною випадковою величиною (НВВ) називається випадкова величина, можливі значення якої неперервно заповнюють деякий інтервал (тобто величина, яка може набути будь-якого числового значення з деякого скінченого або нескінченого інтервалу (a, b). Кількість можливих значень такої величини є нескінченна.

Властивості інтегральної функції розподілу ймовірностей НВВ:

1. Значення інтегральної функції розподілу ймовірностей належать відрізку $[0; 1]$: $0 \leq F(x) \leq 1$.
2. $F(x)$ – неспадна функція, тобто $F(x_2) \geq F(x_1)$, якщо$x_2 > x_1$.
3. Ймовірність того, що X набуде одного конкретного значення, дорівнює 0: $p(x = x_0) = 0$ для будь-якого значення x_0.
4. Ймовірність того, що випадкова величина *X* набуде значення з проміжку $[a; b]$, дорівнює приросту її функції розподілу ймовірностей на цьому проміжку: $p(a \leq X \leq b) = F(b) - F(a)$.
5. Якщо можливі значення випадкової величини X належать інтервалу $(a; b)$, то: $F(x) = 0$ при $x \leq a$ і $F(x) = 1$ при $x \geq b$.

Приклад 8.1. Випадкову величину *X* – витрати електроенергії на фірмі в кВт за годину роботи – задано функцією розподілу ймовірностей:

$$F(x) = \begin{cases} 0, x \leq 2, \\ 0{,}5\text{x} - 1, \ 2 < x \leq 4, \\ 1, x > 4. \end{cases}$$

Обчислити ймовірність $p(2{,}5 \leq X \leq 5)$.

Розв'язання. За властивістю 4 отримаємо:

$p(2{,}5 \leq X \leq 5) = F(5) - F(2{,}5). = 1 - (1{,}25 - 1) = 0{,}75.$

Щільність розподілу (диференціальна функція розподілу) має **властивості:**

1. На всій числовій осі щільність розподілу – невід'ємна функція, тобто $f(x) \geq 0$ для $x\, є(-\infty; +\infty)$.
2. Невласний інтеграл від щільності розподілу в межах від $-\infty$ до $+\infty$ дорівнює одиниці (*умова нормування*): $\int_{-\infty}^{+\infty} f(x)dx = 1$.
3. Остання властивість геометрично означає, що вся площа фігури, обмежена кривою розподілу і віссю абсцис, дорівнює одиниці.
4. Ймовірність того, що неперервна випадкова величина *X* набуде значення з інтервалу $(a; b)$, обчислюється за формулою: $p(a < X < b) = \int_a^b f(x)dx$.

5. Ймовірність потрапляння випадкової величини X на інтервал $(a;b)$ обчислюється як площа криволінійної трапеції, обмеженої зверху графіком кривої $f(x)$, знизу – відрізком $[a;b]$ осі абсцис, зліва і справа – відрізками прямих $x=a$ та $x=b$.

Приклад 8.2. Дано густину розподілу неперервної випадкової величини *X*:

$$F(x)=\begin{cases}0, x\le 1,\\ \mathrm{a}\sqrt{\mathrm{x}}, \quad 1<x\le 9,\\ 1, x>9.\end{cases}$$

Визначити число a.

Розв'язання. За умовою нормування $\int_1^9 a\sqrt{x}\,dx=1$.

Отже, $a\int_1^9 \sqrt{x}\,dx=a\cdot\frac{2}{3}\cdot\left(9^{\frac{3}{2}}-1^{\frac{3}{2}}\right)=1,\ \frac{2}{3}a(27-1)=1,$ звідси $a=\frac{3}{52}$.

Математичне очікування НВВ X, яка має щільність ймовірностей $f(x)$, визначається формулою:

$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx,$$

за умови, що такий невласний інтеграл збіжний. Якщо ж інтеграл розбіжний, то математичне очікування не визначене.

Дисперсією НВВ X називається число:

$$D(X)=M(X-M(X))^2=\int_{-\infty}^{+\infty}(x-M(X))^2\cdot f(x)dx,$$

за умови, що невласний інтеграл збіжний.

Зокрема, виконується формула:

$$D(X)=M(X^2)-(M(X))^2.$$

Середнє квадратичне відхилення визначається так: $\sigma(X)=\sqrt{D(X)}$.

Приклад 8.3. Знайти математичне очікування ВВ, яка задана функцією розподілу $F(x)=\begin{cases}0, x\le 0,\\ \frac{x^2}{9}, \quad 0<x\le 3,\\ 1, x>3.\end{cases}$

Розв'язання. Знайдемо диференціальну функцію розподілу:

$$f(x)=\begin{cases}0, x\le 0,\\ \dfrac{2x}{9}, \quad 0<x\le 3,\\ 0, x>3.\end{cases}$$

$$M(X)=\int_0^3 x\cdot\frac{2x}{9}dx=\frac{2}{9}\cdot\frac{3^3}{3}-\frac{2}{9}\cdot\frac{0^3}{3}=2.$$

Закони розподілу неперервної випадкової величини

Неперервна випадкова величина *X* називається рівномірно розподіленою на відрізку $[a;b]$, якщо усі її можливі значення належать цьому відрізку і її щільність розподілу ймовірностей на цьому відрізку стала, тобто:

$$f(x)=\begin{cases}\dfrac{1}{b-a}, x\notin[a,b],\\ 0, x\notin[a,b].\end{cases}$$

Числові характеристики розподілу:

Математичне очікування обчислюється за формулою:

$$M(X) = \frac{a+b}{2}.$$

Дисперсія обчислюється за формулою:

$$D(X) = \frac{(b-a)^2}{12}.$$

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. У деякій агенції кожне замовлення виконується частинами незалежно у трьох відділах. Ймовірність того, що якийсь відділ не виконає свою частину роботи вчасно, становить 0,2. Скласти закон розподілу числа відділів, що не вклалися у термін виконання цього замовлення.

Відповідь:

x_i	0	1	2	3
p_i	0,512	0,384	0,096	0,008

Задача 2. Статистика свідчить, що 9 % працездатних мешканців регіону є безробітними. Навмання вибирають 4 мешканців. Написати закон розподілу ймовірностей випадкової величини X – кількості безробітних осіб серед навмання вибраних 4 мешканців регіону. Знайдіть:

а) ймовірність того, що не більше ніж двоє осіб серед чотирьох вибраних мешканців регіону є безробітними;

б) найімовірніше число безробітних осіб серед 4 вибраних мешканців регіону.

Задача 3. Робітник виготовляє певний тип деталей. Ймовірність виготовлення бракованої деталі дорівнює 0,01. Написати закон розподілу випадкової величини X – кількості бракованих деталей серед 150 виготовлених (за формулою Пуассона). Знайти ймовірність того, що серед виготовлених деталей виявиться не більше 2 бракованих.

Задача 4. Випадкову величину X задано функцією розподілу:

$$F(x) = \begin{cases} 0, & x \le 0; \\ x^3, & 0 < x \le 1; \\ 1, & x > 1. \end{cases}$$

Знайти щільність розподілу ймовірностей, математичне сподівання та дисперсію цієї випадкової величини. Побудувати графік щільності $f(x)$.

Задача 5. Випадкова величина X задана інтегральною функцією:

$$F(x) = \begin{cases} 0, & x \le -1; \\ \frac{3}{4}x + \frac{3}{4}, & -1 < x \le 4; \\ 1, & x > 4. \end{cases}$$

Знайти ймовірність того, що X прийме значення з проміжку $\left(0; \frac{1}{3}\right)$.

Задача 6. Два стрільця роблять по одному пострілу. Ймовірність влучити для першого стрільця становить 0,6, для другого – 0,8. Випадкова величина X – кількість влучень в мішень. Побудувати графік функції розподілу.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Яка випадкова величина називається неперервною?
2. Як обчислюється ймовірність попадання значень НВВ в певний інтервал?
3. Навести властивості функції розподілу ймовірностей.
4. Якою рівністю визначається зв'язок між функцією розподілу $F(x)$ та густиною розподілу ймовірностей $f(x)$?
5. Як знайти функцію розподілу $F(x)$, якщо відома щільність розподілу $f(x)$?
6. Що таке «крива розподілу»?
7. Які особливості графіка щільності розподілу ймовірностей НВВ?
8. Які вам відомі числові характеристики НВВ? Як вони обчислюються?
9. Охарактеризувати основні властивості числових характеристик НВВ.

ТЕМА 9
ГРАНИЧНІ ТЕОРЕМИ ТЕОРІЇ ЙМОВІРНОСТЕЙ

Явища всього світу підпорядковуються
універсальному закону, який можна назвати
законом великих чисел.
Він полягає в тому,
що якщо спостерігати досить велику
кількість подій однієї і тієї ж природи,
які залежать і від постійних величин,
і від причин, що змінюються регулярно ...,
то між цими кількостями проявляються
майже постійні співвідношення.
С. Пуассон

У разі великого числа експериментів, що здійснюється для вивчення певної випадкової події або випадкової величини, середній їх результат практично перестає бути випадковим і може передбачатися з великою надійністю. Сукупність теорем, у яких стверджується, що існує зв'язок між середнім арифметичним достатньо великого числа випадкової величини і середнім арифметичним їх математичного сподівання в теорії імовірностей, має назву **закон великих чисел.**

Накопичений досвід вказує, що явища, які мають ймовірність, надто близьку до одиниці, майже обов'язково відбуваються. Тому **однією із задач теорії ймовірностей є встановлення закономірностей, що відбуваються зі ймовірностями, близькими до одиниці**. Закономірності такого вигляду називаються ***граничними теоремами теорії ймовірностей.***

Нерівність Чебишова

Перша форма: якщо випадкова величина X невід'ємна і має скінченне математичне очікування, то $p(X > 1) \leq M(X)$.

Нерівність Чебишова

Якщо випадкова величина X набуває тільки невід'ємні значення і має математичне сподівання $M(X)$, то для $\forall\ \varepsilon > 0$ виконується рівність:

$$P(X \geq \varepsilon) \leq \frac{M(X)}{\varepsilon}$$

Приклад 9.1. Сума всіх вкладів у відділенні банку складає 2 млн грн; ймовірність того, що випадково взятий вклад не перевищить 10 тис. грн, дорівнює 0,6. Що можна сказати про кількість вкладників?

Розв'язання. Нехай X – розмір випадкового взятого вкладу, n – числове значення кількості вкладників, $M(X) = \frac{2000}{n}$ (тис. грн) – середній розмір вкладу:

$$P(X \leq 10) \geq 1 - \frac{M(X)}{10n} = 1 - \frac{2000}{10n} = 1 - \frac{200}{n};$$

$$P(X \leq 10) = 0{,}6;$$

$$0{,}6 \geq 1 - \frac{200}{n};\ \ \frac{200}{n} \geq 0{,}4;\ \ n \leq 500.$$

Друга форма: якщо випадкова величина X має скінченне математичне очікування та дисперсію, то для довільного $\varepsilon > 0$ справедлива нерівність Чебишова:

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

Остання нерівність у разі малого числа $\varepsilon > 0$ дає *оцінку знизу* ймовірності того, що величина X набуде значення, досить близького до її математичного очікування.

Для *оцінки ймовірності зверху* використовують **нерівність Маркова.**

Якщо у невід'ємної випадкової величини X є математичне очікування $M(X)$, то при певному додатному ε виконується нерівність:

$$p(X < \varepsilon) \geq 1 - \frac{M(X)}{\varepsilon}; \qquad (9.1)$$

$$P(|X - M(X)| > \varepsilon) \leq \frac{D(X)}{\varepsilon^2}.$$

Приклад 9.2. Сума всіх вкладів у банках становить 2 000 у. о., а ймовірність того, що випадково взятий вклад не перевищує 100 у. о., дорівнює 0,8. Що можна сказати про кількість вкладників цього банку?

Розв'язання. Нехай X – величина випадково вибраного вкладу, N – кількість усіх вкладників. Тоді з умовою задачі випливає, що $M(X) = \frac{2000}{N}$. Оскільки ймовірність $(X < 100) = 0{,}8$, і за нерівністю Маркова (9.1) $p(X < \varepsilon) \geq 1 - \frac{M(X)}{\varepsilon}$, то $0{,}8 \geq 1 - \frac{2000}{100N}$. Звідси $\frac{2000}{100N} \geq 0{,}2$. Отже, отримаємо: $N \leq 100$.

Зауваження 1. Нерівність Чебишова має для практики обмежене значення, оскільки часто дає грубу, а іноді і тривіальну (не викликаючи інтересу) оцінку.

Зауваження 2

- Для випадкової величини $X = m$, яка має біноміальний закон розподілу з математичним сподіванням $a = M(X) = np$ і дисперсію $D(X) = npq$:

$$P(|m - np| \leq \varepsilon) \geq 1 - \frac{npq}{\varepsilon^2}.$$

- Для частоти $\frac{m}{n}$ події в n незалежних експериментах, у кожному з яких вона може відбуватися з однією і тією ж ймовірністю $a = M\left(\frac{m}{n}\right) = p$ і мати дисперсію $D\left(\frac{m}{n}\right) = \frac{pq}{n}$:

$$P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2}.$$

Приклад 9.3. Ймовірність здачі вчасно всіх екзаменів здобувачем дорівнює 0,7. За допомогою нерівності Чебишова оцінити ймовірність того, що частка здобувачів, які вчасно здали всі екзамени із 2 000 здобувачів знаходиться в межах від 0,66 до 0,74; $\varepsilon = 0{,}04$.

Розв'язання. $n = 2000,\ \varepsilon = 0{,}04,\ p = 0{,}7,\ q = 0{,}3.$

$$P\left(\left|\frac{m}{2000} - 0{,}7\right| \leq 0{,}04\right) \geq 1 - \frac{0{,}7 \cdot 0{,}3}{2000 \cdot (0{,}04)^2} = 1 - \frac{0{,}21}{3{,}2} = \frac{299}{320} \approx 0{,}9344.$$

Приклад 9.4. Ймовірність виходу з автомата стандартної деталі дорівнює 0,96. Оцінити за допомогою нерівності Чебишова ймовірність того, що кількість бракованих серед 2 000 деталей знаходиться в межах від 60 до 100 (включно). Уточнити ймовірність цієї ж події за допомогою інтегральної теореми Муавра–Лапласа.

Розв'язання. За умовою ймовірність того, що деталь бракована: $p = 1 - 0{,}96 = 0{,}04$. Кількість бракованих деталей має біноміальний закон розподілу, а його межі – 60 і 100: $a = M(X) = np = 2000 \cdot 0{,}04 = 80$. Тому оцінка ймовірності шуканої події буде:

$$P(60 \le m \le 100) = P(-20 \le m - 80 \le 20) = P(|m - 80| \le 20);$$

$$P(|m - 80| \le 20) \ge 1 - \frac{2000 \cdot 0{,}96 \cdot 0{,}04}{20^2} = 0{,}808.$$

А тепер застосовуємо інтегральну теорему Муавра–Лапласа:

$$P(|m - 80| \le 20) \approx \Phi(x_2) - \Phi(x_1);$$

$$x_2 = \frac{100 - 80}{\sqrt{2000 \cdot 0{,}04 \cdot 0{,}96}} = \frac{20}{8{,}8};\ \ x_1 = \frac{60 - 80}{8{,}8} = -\frac{20}{8{,}8};$$

$$P(|m - 80| \le 20) \approx \Phi(-2{,}27) - 2\Phi(2{,}27) = 2 \cdot 0{,}4881 = 0{,}9762.$$

Ми отримали різні результати, це пояснюється тим, що нерівність Чебишова дає лише нижню границю оцінки ймовірності шуканої події для ∀ випадкової величини, а інтегральна теорема дає достатньо точне значення самої імовірності P.

Правило трьох сигм

Приклад 9.5. Оцінити ймовірність того, що відхилення ∀ випадкової величини від її математичного сподівання не більше трьох середніх квадратичних відхилень.

Розв'язання. Використовуючи функцію Лапласа: $P(|x - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right)$, знайдемо імовірність того, що нормально розподілена випадкова величина відхилення від $a = M(X)$ на $\sigma, 2\sigma, 3\sigma$.

$$P(|x - a| < \sigma) = 2\Phi\left(\frac{\sigma}{\sigma}\right) = 2\Phi(1) = 2 \cdot 0{,}3413 = 0{,}6826;$$

$$P(|x - a| < 2\sigma) = 2\Phi\left(\frac{\sigma}{\sigma}\right) = 2\Phi(2) = 2 \cdot 0{,}4772 = 0{,}9544;$$

$$P(|x - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) = 2 \cdot 0{,}46869 = 0{,}9972.$$

Звідси випливає правило трьох сигм: якщо випадкова величина X має нормальний розподіл, то відхилення цієї випадкової величини від її математичного сподівання за абсолютною величиною не перевищує потроєне середнє квадратичне відхилення (3σ).

Теорема Чебишова

Якщо $X_1,\ X_2, \ldots,\ X_n, \ldots$ – попарно незалежні випадкові величини, причому дисперсії їх рівномірно обмежені (не перевищують постійного числа C), то яке б не було мале додатне число $\varepsilon > 0$, ймовірність нерівності $\left|\frac{X_1,\ X_2,\dots,\ X_n}{n} - \right.$ $\left. - \frac{M(X_1)+\dots+M(X_n)}{n}\right| < \varepsilon$ буде як завгодно близька до одиниці, якщо число випадкових величин достатньо велике:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| < \varepsilon\right) \text{ або } P(|X - M(X)| < \varepsilon) \geq 1 - \frac{C}{n\varepsilon^2}.$$

Наслідок 1. Якщо $X_1,\ X_2,\ \dots,\ X_n, \dots$ – попарно незалежні випадкові величини, мають одне й те саме математичне сподівання a, і якщо дисперсії цих величин рівномірно обмежені, то яке б не було мале число $\varepsilon > 0$, ймовірність нерівності $\left|\frac{X_1+\dots+\ X_n}{n} - a\right| < \varepsilon$ буде як завгодно близька до одиниці, якщо число випадкових величин достатньо велике:

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \dots X_n}{n} - a\right| < \varepsilon\right) = 1.$$

Приклад 9.6. Унаслідок фінансової перевірки 500 банків було з'ясовано, що середній борг кредиторів на 1 млн грн більший від середнього боргу попередньої перевірки. Чи можна це констатувати як випадковість, якщо середнє відхилення боргу кредиторів перед банком 3 млн грн?

Розв'язання. $C = 3$ млн грн, $n = 500$, $\varepsilon = 1$ млн грн.

$$P\left(\left|\frac{1}{500}\sum_{i=1}^{500} X_i - \frac{1}{500}\sum_{i=1}^{500} M(X_i)\right| < 1\right) \geq 1 - \frac{3}{500 \cdot 1^2} = 1 - \frac{3}{500} = 1 - 0{,}006 =$$
$$= 0{,}994.$$

Оскільки ця ймовірність дуже велика, то відхилення боргу можна вважати випадковим.

Для випробувань, проведених за схемою Бернуллі, закон великих чисел формулюється як **теорема Бернуллі:**

Нехай m – число появи події A в n незалежних випробуваннях Бернуллі, $W_n(A) = \frac{m}{n}$ – відносна частота цієї події, $p(A) = p$ – ймовірність появи події A в одному випробуванні. Тоді для будь-якого $\varepsilon > 0$ відносна частота події A наближається до її ймовірності p:

$$\lim_{n\to\infty} p(|W_n(A) - p| < \varepsilon) = 1.$$

Із теореми Бернуллі випливає, що коли число випробувань n є досить велике, то для будь-якого числа $\varepsilon > 0$ подія є практично достовірною, тобто відносна частота події A має властивість стійкості.

Приклад 9.7. Ймовірність виготовити стандартну деталь робітником дорівнює 0,95. Контролю підлягає 400 деталей. Оцінити ймовірність відхилення відносної частоти появи стандартної деталі $W(A)$ від ймовірності 0,95 не більш ніж на величину 0,02.

Розв'язання. $p = 0{,}95,\ q = 0{,}05,\ n = 400,\ \varepsilon = 0{,}02.$

$P(|W(A) - 0{,}95| \le 0{,}02) \ge 1 - \frac{0{,}95 \cdot 0{,}05}{400 \cdot (0{,}02)^2} = 1 - \frac{0{,}0475}{0{,}16} = 1 - 0{,}2969 = 0{,}7031.$

Теорема Пуассона (якщо ймовірність події в кожному експерименті різна). Якщо послідовність незалежних експериментів імовірність появи події A в кожному експерименті дорівнює p_k, то:

$$\lim_{n\to\infty}\left(\left|\frac{m}{n} - \frac{p_1 + p_2 + \ldots + p_n}{n}\right| < \varepsilon\right) = 1.$$

Центральна гранична теорема являє собою групу теорем, які присвячені встановленню умов, за яких виникає нормальний закон розподілу. Серед цих теорем важливе місце належить теоремі Ляпунова.

Теорема Ляпунова

Якщо $X_1,\ X_2, \ldots,\ X_n$ – незалежні випадкові величини, у кожної з яких існує математичне сподівання $M(X_i) = a$, а дисперсія $D(X_i) = \sigma^2$, абсолютний центральний момент третього порядку $M(|X_i - a_i|^3) = m_i$ і $\lim\limits_{n\to\infty} \frac{\sum_{i=1}^n m_i}{(\sum_{i=1}^n \sigma_i^2)^{\frac{3}{2}}} = 0$, то закон розподілу суми $Y_n = X_1 +\ X_2 + \ldots + X_n$ при $n \to \infty$ необмежено наближається до нормального з математичним сподіванням $\sum_{i=1}^n a_i$ і дисперсією $\sum_{i=1}^n \sigma_i^2$.

На практиці центральна гранична теорема переважно використовується в тому випадку, коли доданки X_i мають однаковий розподіл.

Наслідок із центральної граничної теореми. Нехай $X_1,\ X_2, \ldots,\ X_n \ldots$ – послідовність незалежних однаково розподілених випадкових величин зі скінченними математичними очікуваннями $a = M(X_n)$, дисперсією $\sigma^2 = D(X_n)$ і $Y_n = \sum_{i=1}^n X_i,\ \ M(Y_n) = na,\ \ \sigma^2(Y_n) = D(Y_n) = n\sigma^2$. Тоді для будь-якого x: $lim_{n\to\infty}\, p\,(\frac{Y_n - M(Y_n)}{\sigma(Y_n)} < x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}\,dz = 0{,}5 + \Phi(x)$.

Приклад 9.8. У касі банківської установи залишилась сума $d = 3500$ у. о. У черзі за отриманням грошей стоять 20 осіб. Сума X, яку потрібно виплатити окремій особі, – випадкова величина з математичним очікуванням $M(X) = 150$ у. о. і середнім квадратичним відхиленням $\sigma(X) = 60$ у. о. Знайти ймовірність того, що суми d не вистачить для виплати грошей усім особам, які стоять у черзі.

Розв'язання. Використовуючи центральну граничну теорему для однаково розподілених доданків X_i при великому n ($n = 20$ можна практично вважати великим), випадкова величина $Y_n = \sum_{i=1}^n X_i$, де X_i – сума, яку потрібно виплатити i-й особі, має приблизно нормальний розподіл із параметрами:

$M(Y_n) = nM(X)$, $D(X_n) = nD(X)$, $\sigma(Y_n) = \sqrt{n}\sigma(X)$

або $M(Y_n) = 20 \cdot 150 == 3000$, $\sigma(Y_n) = \sqrt{20} \cdot 60 \approx 268$.

Тоді отримаємо:

$$P(Y_n > 3500) = 1 - P(Y_n \leq 3500) \approx 1 - 0{,}5 - \Phi(\frac{3500 - 3000}{268}) =$$

$$= 0{,}5 - \Phi(1{,}87) \approx 0{,}0307.$$

Отже, зі ймовірністю, близькою до 0,0307, наявної в касі суми грошей не вистачить для виплати всім охочим.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Протягом деякий проміжку часу на біржі спостерігався відносно стабільний курс валют. На основі даних біржової статистики на цей період було складено таку таблицю можливих змін курсу валют:

Можлива зміна курсу, %	–1	–0,5	0	0,5	1
Ймовірність зміни	0,1	0,3	0,5	0,005	0,005

За допомогою нерівності Чебишова оцінити ймовірність того, що $p|X - -M(X)| < 0{,}54$.

Задача 2. Відомо, що 75 % продукції, яка виробляється підприємством, – вищого ґатунку. Оцінити ймовірність того, що кількість виробів першого ґатунку серед 200 000 виготовлених буде відрізнятись від математичного очікування цього числа не більше, ніж на 2000 шт.

Задача 3. Середньодобове споживання електроенергії в населеному пункті дорівнює 15 000 кВт·год. Визначити ймовірність того, що споживання електроенергії в цьому населеному пункті протягом доби перевищить 40 000 кВт·год.

Задача 4. Середні витрати води в населеному пункті становлять 50 000 л за день. Оцінити ймовірність того, що в цьому населеному пункті протягом одного певного дня витрати води не перевищать 150 000 л.

Задача 5. Майстерня обслуговує 100 комп’ютерів. Ймовірність того, що кожен зі 100 комп’ютерів витримає гарантійний термін роботи, становить 0,9. За допомогою нерівності Чебишова оцінити ймовірність випадкової події, яка полягає в тому, що більше ніж 85 і менше ніж 95 комп’ютерів витримають гарантійний термін роботи.

Задача 6. Ймовірність вчасної реалізації продукції дорівнює 0,4. Оцінити ймовірність того, що у 100 незалежно реалізованих одиницях продукції відхилення відносної частоти реалізації ймовірності 0,4 за абсолютним значенням буде не менше 0,1.

Задача 7. Ймовірність дозрівання кукурудзяного стебла з трьома качанами дорівнює 75 %. Оцінити ймовірність того, що серед 3 000 стебел частка з трьома качанами буде за абсолютною величиною відрізнятись від ймовірності дозрівання такого стебла не більше ніж на 0,02?

Задача 8. Із 5 000 електронних виробів було обстежено 500 шт., відібраних випадковим способом. Серед них виявилось 10 бракованих. Прийнявши частку бракованих виробів серед відібраних за ймовірність виготовлення бракованого виробу, оцінити ймовірність того, що у всій партії виявиться бракованих виробів не менше 1 % і не більше 3 %.

Задача 9. Перевіряється партія комп’ютерів. Зі ймовірністю 0,01 комп’ютер може мати дефект у відеокарті і, незалежно від цього, зі ймовірністю 0,02 – в оперативній пам’яті. В яких межах буде міститися практично достовірне число бракованих комп’ютерів у партії з 1 000 штук, якщо за ймовірність практичної достовірності приймаємо 0,997?

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Які закономірності називаються «граничними теоремами теорії ймовірностей»?
2. У чому полягає суть «закону великих чисел»?
3. Який факт виражає нерівність Чебишова?
4. У яких процесах відіграє важливу роль «закон великих чисел»?
5. Які умови встановлює центральна гранична теорема?
6. Що стверджує центральна гранична теорема і коли вона застосовується на практиці?

ТЕМА 10
ОСНОВНІ ПОНЯТТЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ

*Теорія ймовірностей – математична теорія,
яка лежить в основі всієї статистичної теорії,
і є також відправною точкою теоретичних
побудов при вивченні випадкових процесів.*
М. Барлетт

Предмет математичної статистики полягає в розробці методів збору та обробки статистичних даних для одержання наукових та практичних висновків.

Основні поняття математичної статистики:

– генеральна сукупність – це вся сукупність об'єктів, які досліджуються;

– вибірка або вибіркова сукупність – це об'єкти, довільно або випадково відібрані з генеральної сукупності для дослідження;

– обсяг (об'єм) сукупності – це кількість об'єктів цієї сукупності.

Приклад 10.1. Плоди дерева (200 шт.) обстежують на наявність специфічного смаку цього сорту.

Для цього відбирають 10 штук. Тут 200 – обсяг генеральної сукупності, 10 – обсяг вибірки. Вибірка повинна бути репрезентативною (представницькою), тобто правильно відображати ті властивості генеральної сукупності, які вивчаються. Досліджувані явища мають бути масовими. Лише тоді статистичні дані будуть достовірними.

Основні задачі, які розв'язує математична статистика:

1) вказати способи збору та групування (якщо даних дуже багато) статистичних відомостей;

2) визначити закон розподілу випадкової величини або системи випадкових величин за статистичними даними;

3) визначити невідомі параметри розподілу;

4) перевірити правдоподібність припущень про закон розподілу випадкової величини, про форму зв'язку між випадковими величинами або про значення параметра, який оцінюють.

Вибірковою сукупністю (вибіркою) називається сукупність випадково взятих об'єктів.

Генеральною називається сукупність об'єктів, із яких зроблено вибірку.

Об'ємом сукупності (вибіркової або генеральної) називається кількість об'єктів цієї сукупності.

Інколи, а взагалі дуже часто дані, які спостерігають, записують у вигляді таблиці 10.1:

Таблиця 10.1

X_i	5	10	13	17	21	25
n_i	8	5	1	10	13	20

Таке зведення називається **рядом варіант** або **простим статистичним рядом.**

Дані будемо називати **варіантами,** числову характеристику варіант – **ознакою.** Додатне число, що вказує, скільки разів та чи інша варіанта зустрічається в таблиці даних, називається **частотою.** Кількісні ознаки елементів генеральної сукупності можуть бути одновимірними і багатовимірними, дискретними і неперервними.

Коли реалізується вибірка, кількісна ознака, наприклад, X, набуває конкретних числових значень ($X = x_i$), які називають **варіантою.**

Зростаючий числовий варіант називають **варіаційним.** Кожна варіанта вибірки може бути спостереженою n_i разів, число n_i називають **частотою варіанти** x_i, тобто $n = \sum_{i=1}^{k} n_i$ де k – кількість варіант, що різняться числовим значенням, n – обсяг вибірки.

Статистичний розподіл вибірки встановлює зв'язок між рядом варіант, що зростає або спадає, і відповідними частотами.

Він може бути представлений таблицею 10.2:

Таблиця 10.2

x_i	x_1	x_2	...	x_m
n_i	n_1	n_2	...	n_m

$$n = n_1 + n_2 + \cdots + n_m.$$

Різниця чисел $R = X_{max} - X_{min}$ називається **розмахом варіант.**

Відношення частоти n_i варіанти x_i до обсягу вибірки n називають її відносною частотою і позначають через W_i:

$$W_i = \frac{n_i}{n}.$$

Для кожної вибірки виконується рівність $\sum_{i=1}^{k} W_i = 1$.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що є предметом математичної статистики?
2. Вказати основні завдання математичної статистики.
3. Що називають генеральною та вибірковою сукупністю, об'ємом цих сукупностей?
4. Навести приклади генеральної сукупності та вибірки.
5. Способи відбору статистичного матеріалу (проілюструвати на прикладах).
6. Що називається варіантою, варіаційним рядом?
7. Що таке частота, відносна частота варіант?

ТЕМА 11
СТАТИСТИЧНИЙ РОЗПОДІЛ ВИБІРКИ.

Дискретний статистичний розподіл вибірки – перелік варіант варіаційного ряду (такі частинні інтервали варіант, які розміщені у зростаючій послідовності, утворюють інтервальний варіаційний ряд) і відповідних їм частот, або відносних частот.

Полігон частот і відносних частот

Дискретний статистичний розподіл вибірки (емпірична функція $F^*(x)$) можна зобразити графічно у вигляді ламаної лінії, відрізки якої сполучають координати точок $(x_i; n_i)$ або $(x_i; W_i)$. У першому випадку ламану лінію називають **полігоном частот,** у другому **полігоном відносних частот.**

Емпіричною функцією розподілу (або функцією розподілу вибірки, або **комулятою**) називають функцію $F^*(x)$, яка визначає для кожного значення x частоту події $X < x$. Математично це має вигляд:

$$F^*(x) = \frac{n_x}{n},$$

де n_x – число варіант, менших x; n – об'єм вибірки.

Отже, щоб знайти, наприклад, $F^*(x_3)$, треба кількість варіант, що менший x_3, поділити на об'єм вибірки, тобто:

$$F^*(x_3) = \frac{n_1 + n_2}{n}.$$

Інтегральну функцію розподілу $F(x)$ генеральної сукупності називають **теоретичною функцією розподілу.** Вона відрізняється від емпіричної функції розподілу $F^*(x)$ тим, що визначає ймовірність події $X < x$, а не частоту цієї події. З теореми Бернуллі випливає, що частота $F^*(x) = \frac{n_x}{n}$ події $X < x$ прямує до імовірності $F(x) = P(X < x)$ цієї події. Тому $F(x)$ та $F^*(x)$ мало відрізняються одна від одної. Доцільно використовувати $F^*(x)$ для наближеного представлення функції розподілу $F(x)$ генеральної сукупності.

Емпірична функція розподілу вибірки слугує для оцінки теоретичної функції розподілу генеральної сукупності.

Приклад 11.1. Знайти емпіричну функцію за заданим розподілом вибірки:

x_i	2	5	10
n_i	14	16	30

Розв'язання. Знаходимо об'єм вибірки: $n = 14 + 16 + 30 = 60$.

Найменша варіанта дорівнює 2, тому $F^*(x) = 0$, при $x \leq 2$.

Значення $X < 5$, а особливо $x_1 = 2$, спостерігалось 14 разів, тому $F^*(x) = \frac{14}{60} \approx 0{,}2$ при $2 < x \leq 5$.

Значення $X < 10$, а особливо $x_1 = 2$, $x_2 = 5$, виконувалось $14 + 16 = 30$ разів, тому $F^*(x) = \frac{30}{60} = 0{,}5$, при $5 < x \leq 10$.

Оскільки $x = 10$ – найбільша варіанта, то $F^*(x) = 1$, при $x > 10$.

Шукана емпірична функція:

$$F^*(x) = \begin{cases} 0, \text{ якщо } x \le 2; \\ 0{,}2, \text{ якщо } 2 < x \le 5; \\ 0{,}5, \text{ якщо } 5 < x \le 10; \\ 1, \text{ якщо } x > 10. \end{cases}$$

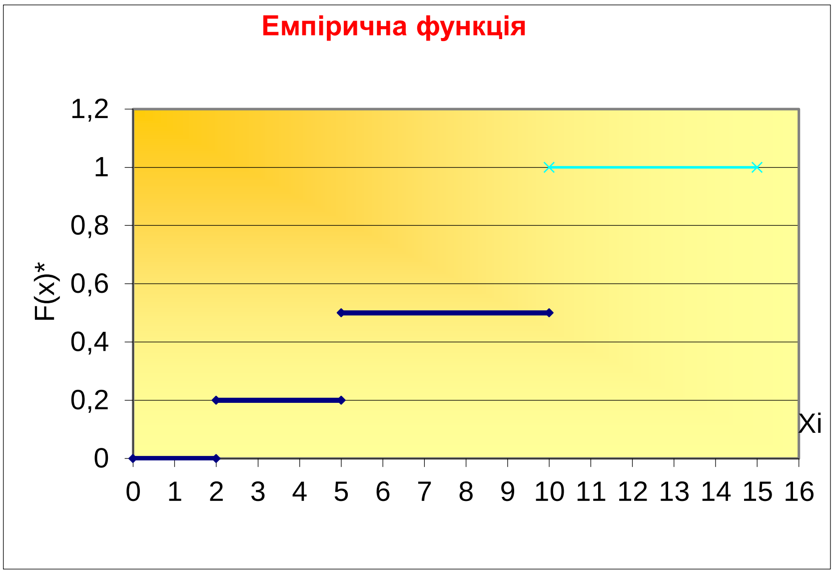


Рис. 11.1. Графік емпіричної функції

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

За заданим дискретним статистичним розподілом вибірки потрібно знайти полігон частот і відносних частот.

x_i	–6	–4	–2	2	4	6
n_i	5	10	15	20	40	10
W_i	0,05	0,1	0,15	0,2	0,4	0,1

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що називається емпіричною функцією розподілу?
2. Властивості емпіричної функції розподілу та її графік.
3. Що називається полігоном частот і відносних частот?
4. Що називається гістограмою частот і відносних частот?

ТЕМА 12
СТАТИСТИЧНІ ОЦІНКИ ПАРАМЕТРІВ РОЗПОДІЛУ

Числові характеристики

Вибіркова середня величина $\bar{x}_B$ дискретного статистичного розподілу вибірки $\bar{x}_B = \frac{1}{n}\sum x_i \cdot n_i$, де x_i – варіанта варіаційного ряду вибірки; n_i – частота цієї варіанти; n – об'єм вибірки ($n = \sum n_i$). Якщо всі варіанти з'являються у вибірці менше одного раза, тобто $n = 1$, то $\bar{x}_B = \frac{1}{n}\sum x_i$.

Відхилення варіант – різниця $(x_i - \bar{x}_B) \cdot n_i$. Сума відхилень усіх варіант варіаційного ряду вибірки завжди дорівнює нулю, тобто $\sum(x_i - \bar{x}_B) \cdot n_i = 0$.

Мода (M_0^*) дискретного статистичного розподілу вибірки – варіанта, що має найбільшу частоту появи.

Медіана (Me^*) дискретного статистичного розподілу вибірки – варіанта, що поділяє варіаційний ряд на дві частини, рівні за кількістю варіант.

Дисперсія вибірки – це середнє арифметичне квадратів відхилень варіант відносно $\bar{x}_B$, яке обчислюється за формулою:

$$D_B = \frac{1}{n}\sum(x_i - \bar{x}_B)^2 n_i \text{ або } D_B = \frac{1}{n}\sum x_i^2 n_i - (\bar{x}_B)^2.$$

Середнє квадратичне відхилення вибірки σ_B:

$$\sigma_B = \sqrt{D_B}.$$

Розмах (R). Для грубого оцінювання варіант відносно $\bar{x}_B$ застосовується величина, яка дорівнює різниці між найбільшою x_{max} і найменшою x_{min} варіантами варіаційного ряду. $R = x_{max} - x_{min}$.

Формула для розрахунку інтервалу $k = \frac{x_{\max} - x_{\min}}{1 + 3{,}322 \cdot \lg n}$.

Коефіцієнт варіації *V*: $V = \frac{\sigma_B}{\bar{x}_B} \cdot 100$ %.

Перелік часткових і відповідних їм частот, або відносних частот, називають **інтервальним статистичним розподілом вибірки.**

h	x_1–x_2	x_2–x_3	x_3–x_4	x_4–x_5	x_5–x_6	…	$x_{k-1} - x_k$
n_i	n_1	n_2	n_3	n_4	n_5	…	n_k
W_i	W_1	W_2	W_3	W_4	W_5	…	W_k

де $h = x_i - x_{i-1}$ – є довжиною часткового i-го інтервалу.

Коефіцієнт асиметрії: $As^* = \frac{\mu_3^*}{\sigma_B^3}$.

Якщо варіанти статистичного розподілу вибірки симетрично розміщені відносно $\bar{x}_B$, то $As^* = 0$, оскільки $\mu_3^* = 0$.

При $As^* < 0$ варіанти статистичного розподілу $x_i < \bar{x}_B$ переважають варіанти $x_i > \bar{x}_B$. Така асиметрія від'ємна.

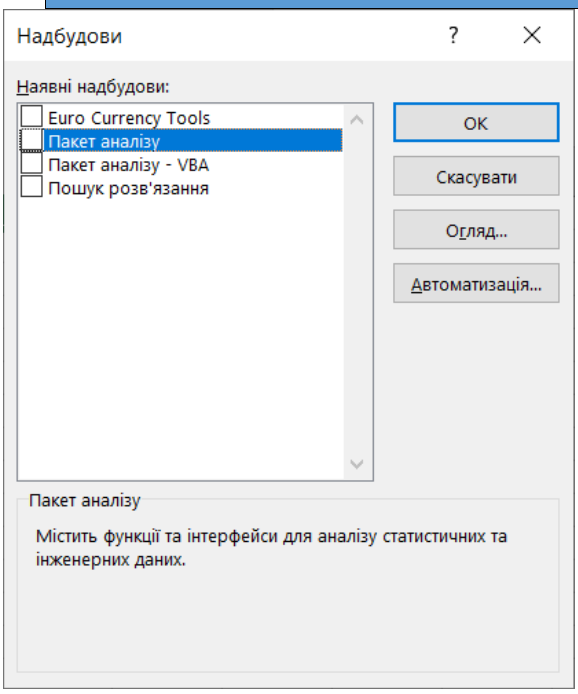
При $As^* > 0$ варіанти $x_j > \bar{x}_B$ переважають варіанти $x_j < \bar{x}_B$, і така асиметрія додатна.

Ексцес: $E_S^* = \frac{\mu_4^*}{\sigma_B^4} - 3$.

E_s^* зазвичай використовується під час дослідження неперервних ознак генеральних сукупностей, оскільки він оцінює крутизну закону розподілу неперервної випадкової величини, порівняно з нормальним. Для нормального закону розподілу $E_s^* = 0$.

Для розрахунків показників описової статистики можна використати статистичні функції MS Excel:

Обсяг вибірки =COUNT()	**Дисперсія = VAR()**
Середнє = AVERAGE ()	**Ст. відхилення =CTDEV()**
Мода =MODE()	**Асиметрія =SCEW()**
Медіана = MEDIAN()	**Ексцес =KURT()**
Розподіли розраховуються за допомогою функції = FREQUENCY ()	



Показники Мір центральної тенденції і мір мінливості вибірки можна обчислити за допомогою:

– пакету Аналіз даних;
– розділу Описова статистика.

Якщо ця надбудова не встановлена, то виконати наступну послідовність дій:

Файл /
Параметри /
Надбудови /
Пакет аналізу /

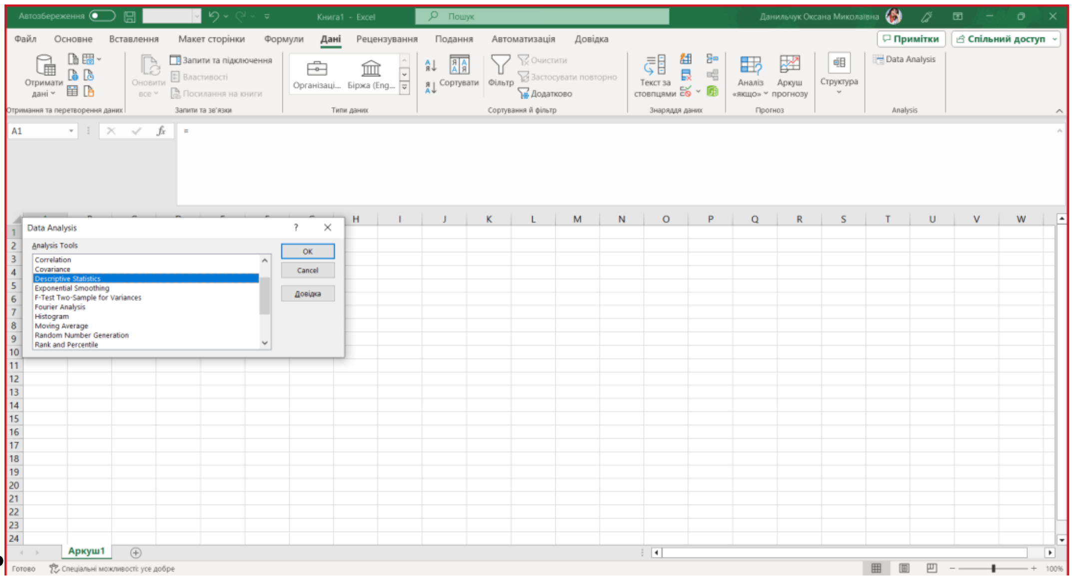


Рис. 12.1. Розділ «Описова статистика»

Гістограмою частот називають ступінчату фігуру, яка складається з прямокутників, основами яких є часткові інтервали варіант довжиною $h = x_k - -x_{k-1}$, а висоти дорівнюють $\frac{n_k}{h}$ (щільність частоти).

$h = x_k - x_{k-1}$ – інтервал варіант; $\frac{n_k}{h}$ – висота (щільність частоти).

Гістограмою відносних частот називають ступінчату фігуру, яка складається з прямокутників, основами яких є часткові інтервали варіант, а висоти дорівнюють відношенню $\frac{W_k}{h}$ (щільність частоти).

Площа гістограми частот дорівнює об’єму вибірки, а площа гістограми частоти – одиниці.

Від способу обрання ширини класу інтервалів h залежить виразність гістограми. Якщо значення h досить мале, то гістограма містить багато випадкового. Якщо значення h досить велике, то в гістограмі зникають індивідуальні особливості вибірки.

Альтернативою гістограми для розподілу частот є полігон частот. Для побудови цього графіка над серединою кожного інтервалу варіант відкладається точка на висоті, відповідній частоті цього інтервалу. Після цього ці точки сполучаються відрізками прямих.

Під час дослідження практичних проблем часто вибірка має значну кількість варіант.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Завдання 1. Побудувати гістограму частот за заданим розподілом вибірки та обчислити числові характеристики вибірки:

$h = 5$	2–7	7–12	12–17	17–22	22–27
n_i	5	10	25	6	4
W_i	0,1	0,2	0,5	0,12	0,08

Завдання 2. Для заданого інтервального статистичного розподілу вибірки побудувати $F^*(x)$:

h	0–10	10–20	20–30	30–40	40–50	50–60
n_i	5	15	20	25	30	5

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Згрупований розподіл вибірки, загальна схема його побудови.
2. Що називають згрупованим розподілом накопичених частот вибірки; накопичених відносних частот вибірки?
3. Що називають згрупованим розподілом щільності частот; щільності відносних частот?
4. Що називають інтервальним статистичним розподілом вибірки?
5. Які рекомендації до вибору числа класів інтервалів?
6. Що називають полігоном частот і відносних частот?
7. Що називають гістограмою частот і відносних частот?

ТЕМА 13
ДОВІРЧИЙ ІНТЕРВАЛ ТА ДОВІРЧІ ЙМОВІРНОСТІ

Щоб статистичні оцінки давали найкращі наближення параметрів, вони повинні задовольняти певні вимоги.

Нехай θ^* – статистична оцінка невідомого параметра θ теоретичного розподілу.

Розглянемо вибірки оцінок $\theta_1^*,\ \theta_2^*, \dots, \theta_m^*$. Оцінку θ^* можна розглядати як випадкову величину, а числа $\theta_1^*,\ \theta_2^*, \dots, \theta_m^*$ – як її можливі значення.

Якщо числа $\theta_k^* (k = 1, \dots, m)$ будуть більші за значення θ, тоді оцінка θ^* дає наближене значення θ з надлишком. У цьому випадку математичне сподівання випадкової величини θ^* буде більше за θ, $M(\theta^*) > \theta$.

Якщо θ^* дає оцінку θ з недостачею, тоді $M(\theta^*) < \theta$.

Отже, використання статистичної оцінки, математичне сподівання якої не дорівнює параметру θ, призводить до математичних похибок.

Ефективною називають таку статистичну оцінку θ^*, яка за заданого об'єму n має найменшу можливу дисперсію.

Обґрунтованою називають статистичну оцінку, яка при $n \to \infty$ прямує за ймовірністю до оцінюваного параметра.

Якщо дисперсія незсунутої оцінки при $n \to \infty$ прямує до нуля, то оцінка буде і обґрунтованою.

Інтервальною називають оцінку, яка визначається двома числами – кінцями інтервалу. Різницю між статистичною оцінкою θ^* та її оцінювальним параметром θ, взята за абсолютним значенням, називають точністю оцінки, а саме: $|\theta^* - \theta| < \delta$, де δ є точністю оцінки.

Ймовірність, із якою береться нерівність, називається **надійністю.**

$$P(|\theta^* - \theta| < \delta) = \gamma \text{ або } P(\theta^* - \delta < \theta < \theta^* + \delta) = \gamma.$$

Інтервал $[\theta^* - \delta;\ \theta^* + \delta]$, що покриває оцінювальний параметр θ генеральної сукупності з заданою надійністю γ, називають **довірчим.**

Надійність (довірча ймовірність) оцінки параметра $\boldsymbol{\theta}$ за $\boldsymbol{\theta^*}$:

$\gamma = P(|\theta - \theta^*| < \delta)$ з якою виконується нерівність $|\theta - \theta^*| < \delta$.

Найчастіше число γ задається наперед, і залежно від обставин воно дорівнює 0,95, або 0,99, або 0,999.

Інтервал $(\theta^* - \delta;\ \theta^* + \delta)$ містить невідомий параметр θ генеральної сукупності.

Інтервал $(\theta^* - \delta;\ \theta^* + \delta)$ називають **довірчим,** якщо він покриває невідомий параметр θ із заданою надійністю γ.

Нехай кількісна ознака X генеральної сукупності розподілена за нормальним законом, середнє квадратичне відхилення σ відомо. Треба знайти довірчий інтервал, що покриває математичне сподівання α генеральної сукупності із заданою надійністю γ.

Довірчий інтервал дорівнює $\bar{x}_B - \frac{t\cdot\sigma}{\sqrt{n}} < \alpha < \bar{x}_B + \frac{t\cdot\sigma}{\sqrt{n}}$ тобто з надійністю γ довірчий інтервал покриває невідомий параметр α:

$$\left(\bar{x}_B - \frac{t\cdot\sigma}{\sqrt{n}};\ \bar{x}_B + \frac{t\cdot\sigma}{\sqrt{n}}\right).$$

Точність оцінки буде $\delta = \frac{t\cdot\sigma}{\sqrt{n}}$. Число t визначається із рівності $\Phi(t) = \frac{\gamma}{2}$ з використанням таблиці значень інтегральної функції Лапласа.

У разі зростанні об'єму вибірки n число σ зменшується, а це означає, що точність оцінки збільшується.

Коли надійність γ збільшується, функція $\Phi(t)$ зростає, і згідно з її властивістю зростає t, і як наслідок, зростає σ.

Отже, збільшення надійності оцінки зменшує її точність.

Приклад 13.1. Випадкова величина розподілена за нормальним законом з параметром σ = 2. Зроблена вибірка об'єму n = 25, з надійністю γ = 0,95. Знайти довірчий інтервал невідомого параметра α цього розподілу.

Розв'язання. Із рівності $\Phi(t) = \frac{\gamma}{2}$, $\Phi(t) = 0{,}475$, $t = 1{,}96$, тоді точність оцінки буде: $\delta = \frac{t\cdot\sigma}{\sqrt{n}} = \frac{2\cdot 1{,}96}{\sqrt{25}} = 0{,}784$.

Отже, довірчий інтервал буде $(\bar{x}_B - 0{,}784;\ \bar{x}_B + 0{,}784)$. Якщо $\bar{x}_B = 2{,}8$, то з надійністю 95 % інтервал (2,016; 3,584) покриває параметр α = 2,8 з точністю до 0,8.

Довірчий інтервал дорівнює $\bar{x}_B - \frac{x\cdot\sigma_\Gamma}{\sqrt{n}} < \alpha < \bar{x}_B + \frac{x\cdot\sigma_\Gamma}{\sqrt{n}}$

Величину $\frac{x\cdot\sigma_\Gamma}{\sqrt{n}}$ називають **точністю оцінки,** або **похибкою вибірки.**

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Задача 1. Вимірявши 40 випадково відібраних після виготовлення деталей, знайшли вибіркову середню, що дорівнює 15 см. Із надійністю γ = 0,99 побудувати довірчий інтервал для середньої величини всієї партії деталей, якщо генеральна дисперсія дорівнює 0,09 см^2.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що називають статистичною оцінкою?
2. Що називають точковою статистичною оцінкою?
3. Яку статистичну оцінку називають спроможною?
4. Яку статистичну оцінку називають сильно спроможною?
5. Як побудувати довірчий інтервал із заданою надійністю?

ТЕМА 14
СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

Статистичними називають гіпотези, в яких йдеться про вигляд розподілу генеральної сукупності або про параметри відомих розподілів.

Разом із припущеною гіпотезою завжди можна розглядати протилежну їй гіпотезу. Якщо припущена гіпотеза була відхилена, тоді використовується протилежна гіпотеза.

Основною (нульовою) називають припущену гіпотезу її позначають H_0.

Альтернативною (конкурентною) називають гіпотезу, що суперечить основній, її позначають H_1.

Під час перевірки статистичної гіпотези за даними випадкової вибірки можна зробити хибний висновок. Якщо за висновком буде відкинута правильна гіпотеза, то кажуть, що це **похибка першого роду.** Якщо за висновком буде відкинута неправильна гіпотеза, то кажуть, що це **похибка другого роду.**

Ймовірність здійснити похибку першого роду позначають α і називають **рівнем значущості.** Найчастіше рівень значущості приймають рівним 0,05 або 0,01. Статистичним критерієм узгодження перевірки гіпотези (або просто критерієм) називають випадкову величину k, розподіл якої (точний або наближений) відомий і яка застосовується для перевірки основної гіпотези.

Критерій узгодження Пірсона $(\boldsymbol{\chi^2})$ (хі-квадрат)

Критерій узгодження Пірсона (χ^2) (хі-квадрат) ефективно використовують для перевірки гіпотези про розподіл генеральної сукупності. Якщо вибірка має розподіл, потрібно з рівнем значущості α перевірити основну гіпотезу H_0: генеральна сукупність розподілена нормально. Критерієм перевірки цієї гіпотези беруть випадкову величину (χ^2), яка у різних випробуваннях набуває різних, наперед невідомих значень.

Критичне значення цієї випадкової величини залежить від рівня значущості α та степенів вільності її розподілу k:

$$\chi^2{}_{kp} = \chi^2(\alpha, k).$$

Ці критичні значення табульовані для різних α і k.

Для розподілу генеральної сукупності за нормальним законом степінь вільності буде $k = m - 3$, де m – кількість варіантів вибірки або часткових інтервалів варіант.

Правило Пірсона. Щоб при заданому рівні значущості α перевірити основну гіпотезу H_0, якщо генеральна сукупність розподілена нормально, потрібно:

1) обчислити теоретичні частоти n'_k для варіант вибірки;

2) обчислити спостереження значення критерія χ^2 за формулою:

$$\chi^2_{\text{сп}} = \sum_{k=1}^{m} \frac{(n_k - n'_k)}{n'_k};$$

3) знайти степінь вільності χ^2;

4) знайти з таблиці критичну точку, χ^2_{kp}, яка відповідає заданому рівню значущості α та степенів вільності k;

5) порівняти $\chi^2_{\text{сп}}$ та χ^2_{kp} і зробити висновок:

- якщо $\chi^2_{\text{сп}} < \boldsymbol{\chi^2_{kp}}$, то гіпотезу H_0 треба прийняти;
- якщо $\chi^2_{\text{сп}} > \boldsymbol{\chi^2_{kp}}$, то гіпотезу H_0 треба відхилити.

Приклад 14.1. При рівні значущості $\alpha = 0{,}05$ перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні та теоретичні частоти:

n_k	6	13	38	74	106	85	30	14
n'_k	3	14	42	82	99	76	37	13

Розв'язання. Кількість варіант вибірки $m = 8, k = 8 - 3 = 5$. З таблиці критичних точок розподілу χ^2 (α, k), для $\alpha = 0{,}05$ та $k = 5$ знаходимо $\chi^2_{kp} = 11{,}1$. Для обчислення $\chi^2_{\text{сп}}$ використаємо розрахункову таблицю:

n_k	n'_k	$n_k - n'_k$	$(n_k - n'_k)$	$\dfrac{(n_k - n'_k)^2}{n'_k}$
6	3	3	9	3
13	14	–1	1	0,07
38	42	–4	16	0,38
74	82	–8	64	0,78
106	99	7	49	0,49
85	76	9	81	1,07
30	37	–7	49	1,32
14	13	1	1	0,08
				$\chi^2_{\text{сп}} = 7{,}19$

Отже, $\chi^2_{\text{сп}} = 7{,}19 < 11{,}1 = \chi^2_{\text{кр}}$, тому за правилом Пірсона гіпотезу H_0 треба прийняти. Так, дані вибірки узгоджуються з гіпотезою H_0, тому що розбіжність емпіричних та теоретичних частот незначна.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

За заданим інтервальним статистичним розподілом випадкової величини X отримали:

$h = 0{,}5$	1 – 1,5	1,5 – 2	2 – 2,5	2,5 – 3	3 – 3,5	3,5 – 4	4 – 4,5
n_i	10	20	50	35	28	15	12

при рівні значущості $\alpha = 0{,}01$ перевірити правильність H_0 про нормальний закон розподілу ознаки X.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Що таке статистична гіпотеза?
2. Як називають нульову (основну) гіпотезу?
3. Як називають конкуруючу (альтернативну) гіпотезу?

4. Коли виникає помилка першого роду?
5. Коли виникає помилка другого роду?
6. Що таке рівень значущості?
7. Що називають статистичним критерієм перевірки?

ТЕМА 15
ЗАСТОСУВАННЯ МЕТОДІВ МАТЕМАТИЧНОЇ СТАТИСТИКИ

ПИТАННЯ ДЛЯ САМОСТІЙНОГО ВИВЧЕННЯ

1. Перевірка гіпотези про нормальний розподіл.
2. Перевірка гіпотези про рівномірний розподіл.
3. Перевірка гіпотези про показниковий розподіл.
4. Перевірка гіпотези про біноміальний розподіл.
5. Перевірка гіпотези про розподіл Пуассона.

ЗАВДАННЯ ДЛЯ САМОСТІЙНОГО ВИКОНАННЯ

Вибірка обсягом 80 осіб має середнє арифметичне $X = 100$ і стандартне відхилення $s_x = 5{,}6$. Необхідно оцінити довірчий інтервал середньої генеральної сукупності μ на рівні значущості 0,05.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Як за допомогою критерію Пірсона перевірити гіпотезу про нормальний розподіл генеральної сукупності?
2. Як визначити теоретичні частоти під час перевірки гіпотези про нормальний розподіл генеральної сукупності?
3. Як за допомогою критерію Пірсона перевірити гіпотезу про нормальний розподіл, якщо вибірка задана інтервальним розподілом?
4. Як визначити кількість ступенів вільності під час перевірки гіпотези про нормальний розподіл за критерієм Пірсона?
5. Як за допомогою критерію перевірити гіпотезу про рівномірний розподіл генеральної сукупності?

СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

Основна література

1. Бідюк П. І., Данилов В. Я., Жиров О. Л. Прикладна статистика: навч. посіб. / Київ: КПІ ім. Ігоря Сікорського. 2023. 186 с.

2. Васильків І. М. Основи теорії ймовірностей і математичної статистики: навч. посібник. Львів: ЛНУ імені Івана Франка, 2020. 184 с.

3. Гулівата І. О., Гусак Л. П., Радзіховська Л. М. Вища та прикладна математика: теорія ймовірностей: навчальний посібник. Вінниця: Видавничо-редакційний відділ ВТЕІ КНТЕУ, 2018. 205 с.

4. Найко Д. А, Шевчук О. Ф. Теорія ймовірностей та математична статистика: навч. посіб. ВНАУ. Вінниця: ТОВ «Твори», 2020. 384 с.

5. Руденко В. М. Математична статистика: навчальний посібник. Київ: Центр учбової літератури, 2022. 304 с.

6. Теорія ймовірностей: навч. посіб. для студ. спеціальності 121 «Інженерія програмного забезпечення» / КПІ ім. Ігоря Сікорського; укл.: О. В. Барабаш, А. П. Мусієнко, О. В. Свинчук / Київ: КПІ ім. Ігоря Сікорського, 2021. 193 с.

7. Теорія ймовірностей та математична статистика: навчальний посібник / В. В. Барковський, Н. В. Барковська, О. К. Лопатін / Київ: Видавництво: Центр учбової літератури, 2021. 424 с.

Допоміжна література

8. Збірник задач з теорії ймовірностей: навч. посібник / П. І. Каленюк, П. А. Костробій, Ю. К. Рудавський та ін.; за ред. П. І. Каленюка. Львів: Видавництво Львівської політехніки, 2012. 248 с.

9. Кармелюк Г. І. Теорія ймовірностей та математична статистика: посібник з розв'язування задач. Київ: Центр учбової літератури, 2007. 575 с.

10. Карташов М. В. Імовірність, процеси, статистика. Київ: ВПЦ «Київський університет», 2007. 494 с.

11. Крикун І. Г. Теорія ймовірностей та математична статистика: навчальний посібник / Миколаїв: НУК, 2017. 148 с.

12. Медведєва М. І. Теорія ймовірностей і математична статистика із застосуванням інформаційних технологій: навч. посібник / М. І. Медведєва та ін. Донецьк: ДонНУ, 2017. 331 с.

13. Наумова М. А. Теорія ймовірностей: навчальний посібник. Вінниця: ТОВ «Нілан-ЛТД», 2018. 192 с.

14. Половенко Л. П., Мерінова С. В. Ймовірнісно-статистичні методи у виявленні кіберінцидентів: інноваційні підходи. *Наука і техніка сьогодні.* 2023. № 12(26). С. 737–750.

URL: http://perspectives.pp.ua/index.php/nts/article/view/6967/7007

15. Прикладна математика на основі MathCAD: навчальний посібник / В. Г. Дзісь, О. В. Левчук, О. М. Дячинська. Вінниця: ВНАУ, 2020. 378 с.

16. Теорія ймовірностей та математична статистика: навч. посібник / Х. Т. Дрогомирецька, О. М. Рибицька, О. З. Слюсарчук та ін. Львів: Видавництво Львівської політехніки, 2012. 396 с.

Інформаційні ресурси в мережі Інтернет

17. Аналіз даних та статистичне виведення на мові R. Онлайн-курс. URL: https://courses.prometheus.org.ua/courses/IRF/Stat101/2016_T3/about

18. Електронний посібник з теорії ймовірностей та математичної статистики. URL: http://lib.lntu.info/books/knit/vm/2011/11-47/

19. Теорія ймовірностей онлайн. URL: https://yukhym.com/uk/vipadkovi-podiji.html