

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

КУРС ЛЕКЦІЙ

Навчальний посібник

Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за спеціальностями 126 «Інформаційні системи та технології»,
121 «Інженерія програмного забезпечення»

Укладач: О.В. Гавриленко

Електронне мережне навчальне видання

Київ
КПІ ім. Ігоря Сікорського
Київ-2025

УДК 004.896 (075)

A64

Укладач: *Гавриленко Олена Валеріївна*, к.ф.-м.н, доцент, доцент кафедри інформаційних систем та технологій, факультет інформатики та обчислювальної техніки.

Рецензент: *Онищенко Вікторія Валеріївна*, д.т.н., професор кафедри інформаційних систем та технологій, факультет інформатики та обчислювальної техніки

Відповідальний редактор: *Жураковська Оксана Сергіївна*, к.т.н., доцент, доцент кафедри інформаційних систем та технологій, факультет інформатики та обчислювальної техніки

*Гриф надано Методичною радою КПП ім. Ігоря Сікорського
(протокол № 6 від 10.04.2025 р.)
за поданням Вченої ради факультету/навчально-наукового інституту
(протокол № 9 від 24.03.2025 р.)*

Теорія ймовірностей і математична статистика. Курс лекцій.
A64 [Електронний ресурс]: навч. посіб. для здобувачів ступеня бакалавра за спец. 126 Інформаційні системи та технології, для здобувачів ступеня магістра за спец. 126 – Інформаційні системи та технології, 121 – Інженерія програмного забезпечення / КПП ім. Ігоря Сікорського; уклад.: О.В. Гавриленко. – Електрон. текст. дані (1 файл). – Київ : КПП ім. Ігоря Сікорського, 2025. – 357 с.

Навчальний посібник призначений для студентів спеціальності 126 «Інформаційні системи та технології» кафедри інформаційних систем та технологій та спеціальності 121 «Інженерія програмного забезпечення» кафедри інформатики та програмної інженерії всіх форм навчання. У посібнику наведено лекції з дисципліни «Теорія ймовірностей і математична статистика» засвоєння курсів «Теорія ймовірностей і математична статистика» (для спеціальності 126 «Інформаційні системи та технології») та «Теорія ймовірностей» (для спеціальності 121 «Інженерія програмного забезпечення»). Метою автора було розвинути теоретичні навички з даної дисципліни для їх застосування у розв'язанні задач та в інженерній практиці.

УДК 004.896 (075)

Реєстр. № НП 23/24-165. Обсяг 14,86 авт. арк.
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
проспект Берестейський, 37, м. Київ, 03056
<https://kpi.ua>

Свідectво про внесення до Державного реєстру видавців, виготовлювачів
і розповсюджувачів видавничої продукції ДК № 5354 від 25.05.2017 р.

© КПП ім. Ігоря Сікорського, 2025

ЗМІСТ

ВСТУП	4
ЛЕКЦІЯ 1	6
ЛЕКЦІЯ 2	24
ЛЕКЦІЯ 3	45
ЛЕКЦІЯ 4	61
ЛЕКЦІЯ 5	76
ЛЕКЦІЯ 6	93
ЛЕКЦІЯ 8	135
ЛЕКЦІЯ 9	162
ЛЕКЦІЯ 10	186
ЛЕКЦІЯ 11	210
ЛЕКЦІЯ 12	223
ЛЕКЦІЯ 13	242
ЛЕКЦІЯ 14	263
ЛЕКЦІЯ 15	277
ЛЕКЦІЯ 16	293
ЛЕКЦІЯ 17	306
ЛЕКЦІЯ 18	322
ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ	353
ЛІТЕРАТУРА	355

ВСТУП

Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Дисципліна «Теорія ймовірностей і математична статистика» («Теорія ймовірностей») має на меті вивчення основ математичної теорії ймовірностей, ймовірнісних процесів, вивчення ймовірнісно-статистичного матеріалу, вміння розв'язувати різноманітні задачі згідно класичної схеми, способи перерахунку елементів скінчених множин, вміння застосовувати набуті знання до розв'язку прикладних задач, які виникають на практиці, вміння будувати математичні моделі реальних процесів.

По закінченню вивчення дисципліни студент набуває такі навички та вміння:

- володіння основною термінологією дисципліни, вміння пояснити зміст базових понять і розділів;
- вміння класифікувати стандартні задачі за ознаками, вміння розв'язувати їх;
- володіння навичками роботи з літературою з дисципліни, знання основних підручників, довідників, таблиць. Вміння знайти у літературі необхідну інформацію (спосіб розв'язання задач);
- знання основних прийомів розв'язання стандартних задач теорії ймовірностей (комбінаторні методи, методи, пов'язані з основними теоремами, метод характеристичних функцій, методи розрахунків характеристик дискретних та неперервних розподілів, вміння користуватися біноміальним розподілом та його граничними випадками, нормальним розподілом, законом великих чисел, центральною граничною теоремою тощо)
- вміння проводити стандартні статистичні розрахунки вручну та з використанням комп'ютерних програм, знання основних статистичних процедур: розрахунків середніх, дисперсій, моментів, коефіцієнтів кореляції.

Компетентності

Загальні компетентності:

- ЗК-1. Здатність до абстрактного мислення, аналізу та синтезу.

Спеціальні (фахові, предметні) компетентності:

- ФК-13. Здатність проводити обчислювальні експерименти, порівнювати результати експериментальних даних і отриманих рішень.

Програмні результати навчання

- ПРН-1. Знати лінійну та векторну алгебру, диференціальне та інтегральне числення, теорію функцій багатьох змінних, теорію рядів, диференціальні рівняння для функції однієї та багатьох змінних, операційне числення, теорію ймовірностей та математичну статистику в обсязі,

необхідному для розробки та використання інформаційних систем, технологій та інфокомунікацій, сервісів та інфраструктури організації.

- ПРН-2. Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проектування і використання інформаційних систем та технологій.

Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Пререквізити:

- Спеціальні розділи математики.

Постреквізити:

- Ймовірнісні моделі та статистичне оцінювання в інформаційно-управляючих системах.

ЛЕКЦІЯ 1

ВИПАДКОВІ ПОДІЇ ТА ЇХ ЙМОВІРНОСТІ

Теорія ймовірностей — математична наука, яка вивчає закономірності випадкових явищ при їх масовому повторенні.

Серед явищ, що відбуваються навколо нас, важко назвати такі, які б не були в тій чи іншій мірі піддані впливу випадкових факторів. Проте дослідження великої кількості одноманітних явищ показує, що в них прояв того чи іншого випадкового фактору підлягає певній закономірності або стійкості. На встановлення цих закономірностей і спрямовані методи теорії ймовірностей.

Виникнення теорії ймовірностей відноситься до середини 17-го сторіччя і пов'язане з іменами видатних французьких математиків П. Ферма (1601-1665), Б. Паскаля (1623-1662) та видатного швейцарського математика Я. Бернуллі (1654-1705), в працях яких вже фігурують такі важливі поняття, як ймовірність випадкової події і математичне сподівання випадкової величини.

Проте низький рівень розвитку природознавства того часу, а також відсутність соціального замовлення на відповідні задачі стали причиною того, що основні поняття і методи теорії ймовірностей розвивалися головним чином на розв'язанні задач азартних ігор, демографії і страхової справи.

Лише кінець 18-го – початок 19-го сторіччя стали часом ґрунтовного застосування теорії ймовірностей до розв'язання актуальних проблем природознавства, виробничих і технологічних процесів, військової справи, економіки, статистики і інших галузей, що привело до необхідності створення розвинутого аналітичного апарата досліджень.

Велика роль в розробці аналітичних методів теорії ймовірностей належить англійському математику А. де Муавру (1667-1754), французьким математикам П. Лапласу (1749-1827) і С. Пуассону (1781-1840), німецькому математику К. Гаусу (1777-1855).

На сучасному етапі визначний внесок у розвиток теорії ймовірностей і математичної статистики, розробку нових напрямів і методів зробили відомі вчені – українські академіки Б.В. Гнеденко (1913-1996), Й.І. Гіхман (р.нар.1918), В.С. Королюк (р.нар.1925), А.В. Скороход (р.нар.1930), І.М. Коваленко (р.нар.1935).

1.1. Стохастичне випробування і події. Види подій

Поняття події відноситься до основних понять, на яких базується теорія ймовірностей. Під *подією* будемо розуміти всякий факт, який в результаті випробування (дослідку, експерименту, спостереження) може відбутися або не відбутися. При цьому в поняття «випробування» входять не

лише ті, які проводить людина, але й явища, що відбуваються в природі незалежно від людини.

В теорії ймовірностей розглядаються *стохастичні* (випадкові, ймовірнісні) *випробування*, тобто такі, які можна повторити будь-яку кількість разів, але результат яких при кожному повторенні наперед невідомий, тобто випадковий. Тому *теорія ймовірностей* визначається ще як математичний аналіз стохастичного випробування (рис. 1.1).

<i>Випробування</i>	<i>Події</i>
1. Кидання грального кубика	Випала парна кількість очок
2. Оцінювання якості виробу	Виріб виявився бракованим
3. Звернення до системи продажу авіаквитків	Одержано квиток
4. Випробування автоматичної системи на надійність протягом часу t	Система вийшла з ладу
5. Вимірювання фізичної величини приладом	Знайдено значення величини із заданою точністю

Рис. 1.1. Приклади випробувань та подій

Кожне стохастичне випробування має деяку множину $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ всіх можливих результатів або наслідків, які не розкладаються на простіші. Множина Ω утворює так звану *множину* (або *простір*) *елементарних наслідків* (подій) ω_i , якщо ці наслідки є взаємовиключаючими і результатом випробування завжди є один і тільки один наслідок. Множина Ω може бути скінченною $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ або нескінченною $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$.

В теорії ймовірностей довільні події позначаються літерами $A, B, C, \dots$ або літерами з індексами, наприклад, $A_1, A_2, \dots, A_k, \dots$, інакше A_i ($i = 1, 2, \dots, k$). Зміст події, виражений словами або математичними залежностями, заключається в фігурні дужки. Наприклад, якщо з партії виробів навмання вибирається один для оцінки якості, то можлива подія – $A = \{\text{вибрано бракований виріб}\}$.

Всі події підрозділяються на три види: достовірні, неможливі і випадкові.

Достовірною називається подія, яка обов'язково відбудеться в даному випробуванні. Достовірну подію позначатимемо Ω .

Наприклад, якщо з партії, яка складається лише з стандартних виробів, навмання вибрано три вироби, то подія $A = \{\text{всі вироби є стандартними}\}$ – достовірна.

Неможливою називається подія, яка напевно не відбудеться в даному випробуванні. Неможливу подію позначатимемо $\emptyset$.

Так, подія $B = \{\text{серед вибраних виробів є нестандартні}\}$ – неможлива в умовах випробування з попереднього прикладу.

Випадковою називається подія, яка в даному випробуванні може відбутися або не відбутися в залежності від впливу різноманітних випадкових факторів.

Приклади випадкових подій: поява шести очок при підкиданні грального кубика, відмова блоку системи на протязі певного часу t при дослідженні надійності системи та ін.

В наведених прикладах, крім події, вказано також випробування, наслідком якого є ця подія.

Оскільки довільна подія A є наслідком деякого стохастичного випробування, а простір Ω — множина всіх можливих елементарних наслідків випробування, то подія A входить в простір Ω або, іншими словами, подія A є підмножиною множини Ω , що позначається $A \subset \Omega$. При цьому ті елементарні наслідки ω_i з простору Ω , при яких подія A відбувається, тобто наслідки, які входять до складу події A ($\omega_i \in A$), називаються *сприятливими* події A .

Приклад 1.1. Із двозначних чисел, що не перевищують 20, навмання вибирається одне число. Описати простір елементарних наслідків Ω і події $A = \{\text{вибране число ділиться на } 5\}$; $B = \{\text{вибране число — просте}\}$; $C = \{\text{вибране число — парне}\}$.

Розв'язання. В результаті випробування може бути вибране будь-яке двозначне число, не більше 20, отже, простір елементарних наслідків $\Omega = \{10, 11, 12, \dots, 20\}$. З усіх наслідків із простору Ω події A сприяють наслідки 10, 15, 20, отже, $A = \{10, 15, 20\}$. Аналогічно $B = \{11, 13, 17, 19\}$ і $C = \{10, 11, 12, 14, 16, 18, 20\}$.

Відповідь. $A = \{10, 15, 20\}$, $B = \{11, 13, 17, 19\}$, $C = \{10, 11, 12, 14, 16, 18, 20\}$.

1.2. Операції над подіями

Об'єднання подій

Подія C є *об'єднанням* A і B ($C = A \cup B$), якщо подія C складається з усіх елементарних подій, що входять в подію A та з усіх елементарних подій, що входять в B . Причому, якщо елементарна подія входить як в A , так і в B , то вона входить в подію C один раз. Тому, що C є підмножиною Ω , то C складається з різних елементарних подій.

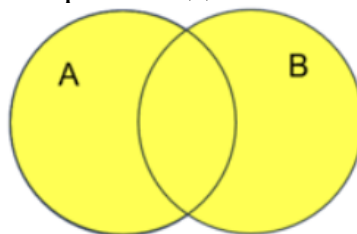


Рис. 1.2. $A \cup B$

Наслідок: Внаслідок випробування подія C настає тоді і тільки тоді, коли настала подія A або подія B , або та елементарна подія, що входить і в A і в B одночасно.

Вводиться операція об'єднання скінченної чи нескінченно-зліченої кількості подій.

Нехай I – скінченна чи нескінченно-злічена множина індексів, тоді

$$C = \bigcup_{i \in I} A_i.$$

C складається з усіх елементарних подій, що входять в A_i , і якщо елементарна подія входить в різні A_i , більш ніж один раз то в C лише один раз. C настає тільки тоді, коли настала хоча б якась A_i .

Елементарні тотожності об'єднання подій

1. $A \cup A = A$;
2. $A \cup \Omega = \Omega$;
3. $A \cup V = A$.

Перетин подій

Подія C є **перетином** A і B ($C = A \cap B$ або $C = A \cdot B$), якщо подія C складається з усіх елементарних подій, що входять і в A , і в B одночасно.

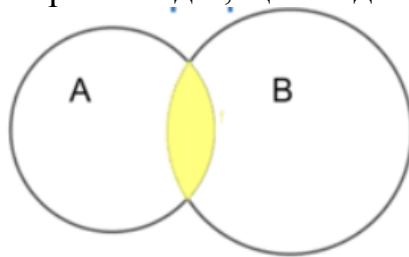


Рис. 1.3. $A \cap B$

Наслідок 1. Подія C настає тоді і тільки тоді, коли настали і A , і B .

Наслідок 2. Подія C настає тоді і тільки тоді, коли настали всі події A_i :

$$C = \bigcap_{i \in I} A_i.$$

Елементарні тотожності перетину подій

1. $A \cap A = A$;
2. $A \cap \Omega = A$;
3. $A \cap V = V$.

Закони об'єднання та перетину подій

1. $A \cup B = B \cup A$; $A \cap B = B \cap A$;
2. $A \cup (B \cap C) = (A \cup B) \cap C$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Різниця подій

Подія $C = A/B$ зветься **різницею подій** A і B якщо вона складається з усіх подій, що входять в A і не входять в B .

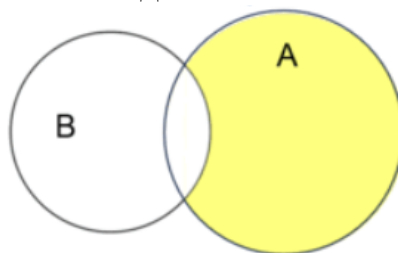


Рис. 1.4. A/B

Протилежна подія

Протилежна подія $\bar{A}$ – це така подія, яка відбувається тоді, коли не відбувається вказана подія A : $\bar{A} \cup A = \Omega$, $\bar{A} \cap A = \{\emptyset\}$.

A і B **несумісні**, якщо $A \cap B = \{\emptyset\}$

Наслідок: Якщо A і B несумісні, то вони ніколи не можуть наставати як результат 1 випробування.

Закони де Моргана

$$\overline{A \cup B} = \bar{A} \cap \bar{B};$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B};$$

$$\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \bar{A}_i \quad \overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \bar{A}_i.$$

Складна подія зветься **достовірною** і позначається як $\mathcal{U}$, якщо вона обов'язково настає внаслідок випробування. Протилежною подією до достовірної є неможлива.

$$\mathcal{U} = \Omega; V = \bar{\mathcal{U}} = \{\emptyset\}.$$

1.3. Поняття ймовірності події. Класична ймовірнісна схема і класичне означення ймовірності подій

Розв'язання прикладу 1.1 свідчить про те, що різні події характеризуються певною мірою можливості їх появи в результаті випробування. Так, зокрема, в прикладі 1.1 можливість появи події C більша, ніж події A або події B . Такою мірою можливості появи події є її ймовірність. Це поняття також відноситься до основних базових понять теорії ймовірностей.

Ймовірність події A позначається $P(A)$ (від *probabilitas* (лат.) – ймовірність). За одиницю її виміру природно прийняти ймовірність достовірної події Ω , тобто $P(\Omega) = 1$. Тоді всяка інша подія A – можлива, але не достовірна, буде характеризуватись ймовірністю, меншою за одиницю. Зрозуміло, що ймовірність неможливої події, яка в даному випробуванні не відбувається ні за яких умов, дорівнює нулю.

Отже, справедливі умови (1.1):

$$P(\Omega) = 1; P(\emptyset) = 0; P(A) \in (0; 1). \quad (1.1)$$

Класична ймовірнісна схема стосується подій найпростішого виду, які називають **випадками** або **шансами**. Це події, які утворюють повну групу, несумісні і рівноможливі.

Відповідно n подій $A_1, A_2, \dots, A_n$ називаються **попарно несумісними** (або просто несумісними), якщо поява однієї з них в даному випробуванні виключає появу всіх інших, тобто добуток двох будь-яких з цих подій в одному випробуванні є неможливою подією:

$$A_i \cdot A_j = \emptyset, (i, j = 1, 2, \dots, n; i \neq j).$$

Наприклад, при випробуванні надійності двох приладів на протязі часу t події $A = \{\text{один прилад виявився надійним}\}$; $B = \{\text{обидва прилади виявилися надійними}\}$ – несумісні.

Події $A_1, A_2, \dots, A_n$ вважаються **рівноможливими** в даному випробуванні, якщо за умовами випробування немає підстав вважати появу однієї з них більш можливою, ніж поява кожної з решти цих подій.

Наприклад, при передачі в однакових умовах по каналу зв'язку трьох сигналів однакової довжини події $A_1 = \{\text{перший сигнал передано правильно}\}$, $A_2 = \{\text{другий сигнал передано правильно}\}$ і $A_3 = \{\text{третій сигнал передано правильно}\}$ є рівноможливими.

Події $A_1, A_2, \dots, A_n$ утворюють **повну групу подій**, якщо принаймні одна з них неодмінно відбудеться в даному випробуванні, тобто їх сума є достовірною подією:

$$\sum_{i=1}^n A_i = \Omega.$$

Зокрема, попарно несумісні події $A_1, A_2, \dots, A_n$ утворюють **повну групу несумісних подій**, якщо в даному випробуванні неодмінно відбудеться одна і тільки одна з них, тобто їх сума є достовірною подією, а добуток будь-яких двох з цих подій є подія неможлива:

$$\sum_{i=1}^n A_i = \Omega; A_i \cdot A_j = \emptyset, (i, j = 1, 2, \dots, n; i \neq j).$$

Наприклад, при випробуванні надійності двох приладів на протязі часу t події $A_1 = \{\text{обидва прилади виявились надійними}\}$, $A_2 = \{\text{перший прилад виявився надійним, другий – ні}\}$, $A_3 = \{\text{другий прилад виявився надійним, перший – ні}\}$, $A_4 = \{\text{жоден прилад не пройшов випробування}\}$ складають повну групу несумісних подій.

Дві події, які утворюють повну групу несумісних подій, називають **протилежними** подіями, і позначають A і $\bar{A}$.

Наприклад, при подачі запиту до системи продажу авіаційних білетів на придбання білета на даний рейс події $A = \{\text{одержано білет}\}$ і $\bar{A} = \{\text{одержано відмову}\}$ є протилежними.

Ймовірністю $P(A)$ події A називають відношення числа m сприятливих цій події наслідків до загального числа n всіх рівноможливих наслідків випробування, що утворюють повну групу несумісних подій (1.2):

$$P(A) = \frac{m}{n}. \quad (1.2)$$

Приклад 1.2. З партії виробів, яка містить 20 стандартних і 6 нестандартних виробів, навмання вибирається 1 виріб. Знайти ймовірність того, що він нестандартний.

Розв'язання. Оскільки в партії всього 26 виробів, і будь-який з них може бути вибраний, то випробування має 26 рівноможливих наслідків, отже, $n = 26$. Сприятливими для події $A = \{\text{вибрано нестандартний виріб}\}$ є 6 наслідків, що відповідають вибору одного з нестандартних виробів, отже, $m = 6$. За формулою (1.2)

$$P(A) = \frac{6}{26} = \frac{3}{13}.$$

Відповідь. $P(A) = \frac{3}{13}$

Тут і в подальшому вибір навмання у випробуванні будемо розуміти як рівноможливість наслідків цього випробування.

1.4. Застосування формул комбінаторики для обчислення ймовірностей в класичній схемі

В багатьох випробуваннях класичної схеми обчислення загальної кількості елементарних наслідків і кількості сприятливих наслідків для подій, що відбуваються в цих випробуваннях, зв'язане з різними схемами вибору певного числа k елементів з n різних елементів деякої вихідної множини $M = \{m_1, m_2, \dots, m_n\}$. В цих випадках для обчислення ймовірності події застосовуються формули комбінаторики. Наведемо основні принципи та формули комбінаторики.

Комбінації з n елементів по k

Комбінаціями (сполученнями) з n елементів $\{m_1, m_2, \dots, m_n\}$ по k називаються всі підмножини множини M , що містять k елементів ($k = 0, 1, \dots, n$) і відрізняються між собою принаймні одним елементом.

Наприклад, з множини 3-х чисел $M = \{1, 2, 3\}$ можна утворити такі комбінації по 2 елемента: $\{1, 2\}$; $\{1, 3\}$; $\{2, 3\}$.

Число всіх комбінацій з n елементів по k позначається C_n^k і обчислюється за формулою (1.3):

$$C_n^k = \frac{n!}{k!(n-k)!}, \quad (1.3)$$

де $n!$ — добуток n перших натуральних чисел $n! = 1 \cdot 2 \cdot \dots \cdot n$, зокрема, за означенням прийнято $0! = 1$.

Розміщення з n елементів по k

Розміщеннями з n елементів по k називаються всі підмножини множини M , що містять k елементів ($k = 0, 1, \dots, n$) і відрізняються між собою принаймні одним елементом або порядком їх розташування.

Наприклад, з множини 3-х чисел $M = \{1, 2, 3\}$ можна утворити такі розміщення по 2 елемента: $\{1, 2\}$; $\{2, 1\}$; $\{1, 3\}$; $\{3, 1\}$; $\{2, 3\}$; $\{3, 2\}$.

Число розміщень з n елементів по k позначається A_n^k і обчислюється за формулою (1.4):

$$A_n^k = \frac{n!}{(n-k)!}. \quad (1.4)$$

Перестановки з n елементів.

Перестановками з n елементів $\{m_1, m_2, \dots, m_n\}$ називаються всі множини, що утворюються з множини M перестановкою будь-яких її елементів, тобто відрізняються між собою лише порядком розташування елементів.

Наприклад, з множини 3-х чисел $M = \{1, 2, 3\}$ можна утворити такі перестановки: $\{1, 2, 3\}$; $\{1, 3, 2\}$; $\{2, 1, 3\}$; $\{2, 3, 1\}$; $\{3, 1, 2\}$; $\{3, 2, 1\}$.

Число перестановок з n елементів позначається P_n і обчислюється за формулою (1.5):

$$P_n = n!. \quad (1.5)$$

Перестановка з повтореннями з n елементів

Перестановки з повтореннями з n елементів, серед яких є k_1 елементів першого типу, k_2 елементів другого типу, ..., k_m елементів m -го типу, обчислюються за формулою (1.6):

$$P_n(k_1, k_2, \dots, k_m) = \frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_m!}. \quad (1.6)$$

Розміщення з повтореннями з n елементів по k

Розміщення з повтореннями з n елементів по k по називається довільна впорядкована підмножина з k елементів з множини M (елементи не обов'язково різні). Число розміщень з n елементів по k знаходиться за формулою (1.7):

$$\overline{A}_n^k = n^k. \quad (1.7)$$

Комбінація з повтореннями з n елементів по k

Комбінацією з повтореннями з n елементів по k називається довільна впорядкована підмножина з k елементів з множини M (елементи не обов'язково різні). Число комбінацій з n елементів по k знаходиться за формулою (1.8):

$$\overline{C}_n^k = \frac{(n+k-1)!}{k! \cdot (n-1)!}. \quad (1.8)$$

Принцип добутку

Якщо послідовно виконується k дій, причому першу дію можна виконати n_1 способами, другу – n_2 способами і, нарешті, k -у дію n_k способами, то всі k дій разом можна виконати $n_1 \cdot n_2 \cdot \dots \cdot n_k$ способами.

Принцип суми

Якщо множини $M_1, M_2, \dots, M_k$ містять відповідно $n_1, n_2, \dots, n_k$ елементів, причому добуток цих множин є пуста множина:

$$M_1 \cap M_2 \cap \dots \cap M_k = \emptyset,$$

то їх сума містить $n_1 + n_2 + \dots + n_k$ елементів.

1.5. Приклади застосування формул комбінаторики для обчислення ймовірностей в класичній схемі

Приклад 1.3. Комплект з 30-ти однотипних радіодеталей містить 20% нестандартних. Для включення в схему навімання вибирається 4 деталі. Знайти ймовірність того, що будуть відібрані лише стандартні деталі.

Розв'язання. За умовою задачі комплект містить 24 стандартні і 6 нестандартних деталей. Випробування полягає в виборі 4-х деталей із 30, причому кожна вибірка визначається складом деталей і не залежить від порядку їх розташування у вибірці. Тому загальна кількість n наслідків випробування обчислюється за формулою (1.3):

$$n = C_{30}^4 = \frac{30 \cdot 29 \cdot 28 \cdot 27}{4!} = 27405.$$

Позначимо подію: $A = \{\text{всі вибрані деталі стандартні}\}$. Сприятливими цій події будуть лише ті наслідки, в яких вибір 4-х деталей проводиться з 24-х стандартних, тобто

$$m = C_{24}^4 = \frac{24 \cdot 23 \cdot 22 \cdot 21}{4!} = 10626.$$

Ймовірність події A обчислюється за формулою (1.2):

$$P(A) = \frac{10626}{27405} = 0,39.$$

Відповідь. $P(A) = 0,39$

Приклад 1.4. *Авіазавод одержав 16 однотипних агрегатів, 9 з яких виготовлені заводом N_1 , а інші – заводом N_2 . Для ремонту літака навмання відібрано 7 агрегатів. Знайти ймовірність того, що 5 з них виготовлені заводом N_1 .*

Розв’язання. За умовою задачі випробування полягає в виборі семи агрегатів із 16-ти. Загальна кількість n наслідків обчислюється за формулою (1.3):

$$n = C_{16}^7 = \frac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10}{7!} = 11440.$$

Позначимо через A подію {вибрано 5 агрегатів заводу N_1 і 2 – заводу N_2 }.

Кількість варіантів, якими 5 агрегатів можуть бути вибрані з 9-ти, виготовлених заводом N_1 , обчислюється за формулою

$$m_1 = C_9^5 = \frac{9 \cdot 8 \cdot 7 \cdot 6}{4!}.$$

Кількість варіантів, якими 2 агрегати можуть бути вибрані з 7-ми, виготовлених заводом N_2 , обчислюється за формулою

$$m_2 = C_7^2 = \frac{7 \cdot 6}{2!} = 21.$$

На основі принципу добутку кількість m наслідків, сприятливих події B , дорівнює

$$m = m_1 \cdot m_2 = 2646,$$

а ймовірність події B

$$P(A) = \frac{2646}{11440} \approx 0,23.$$

Відповідь. $P(A) \approx 0,23$.

Приклад 1.5. *В аеропорту здійснили посадку 5 літаків, серед яких два Ту-154. Літаки випадковим чином розподіляються по стоянках, розташованих в одному ряду. Знайти ймовірність того, що обидва Ту-154 розташуються поряд, якщо: а) кількість стоянок – 8; б) кількість стоянок – 5.*

Розв’язання. а) За умовою задачі випробування полягає в розміщенні п’яти літаків на восьми стоянках. Загальне число n варіантів такого розміщення обчислюється за формулою (1.4), оскільки ці варіанти відрізняються порядком розташування літаків на стоянках. Тому

$$n = A_8^5 = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 6720.$$

Позначимо через A подію {літаки Ту-154 будуть розташовані поруч}.

Кількість варіантів m_1 розташування двох літаків поряд у випадку восьми стоянок знаходиться безпосереднім підрахунком і дорівнює

$$m_1 = 7 \cdot 2 = 14.$$

При цьому 3 інші літаки розташовуються в довільному порядку на 6-ти стоянках, що залишились вільними. Кількість варіантів m_2 такого розташування також обчислюється за формулою (1.4):

$$m_2 = A_6^3 = 6 \cdot 5 \cdot 4 = 120.$$

Кількість m наслідків, сприятливих події C , згідно з принципом добутку дорівнює:

$$m = m_1 \cdot m_2 = 1680,$$

а ймовірність події A обчислюється за формулою (1.2):

$$P(A) = \frac{1680}{6720} = 0,25.$$

б) Загальна кількість n варіантів розташування п'яти літаків на п'яти стоянках обчислюється за формулою (1.8):

$$n = 5! = 120.$$

Кількість варіантів m_1 розташування двох літаків поряд на п'яти стоянках дорівнює

$$m_1 = 4 \cdot 2 = 8,$$

а кількість варіантів m_2 розташування інших трьох літаків на трьох вільних стоянках також обчислюється за формулою (1.5):

$$m_2 = P_3 = 3! = 6.$$

Згідно з принципом добутку кількість наслідків m , сприятливих події A , дорівнює

$$m = m_1 \cdot m_2 = 48,$$

а ймовірність події A

$$P(A) = \frac{48}{120} = 0,4.$$

Відповідь. а) $P(A) = 0,25$; б) $P(A) = 0,4$.

Із означення та всіх розглянутих прикладів випливає, що ймовірність події – це цілком визначене число, причому зовсім не випадкове. Це – не випадкова (детермінована) характеристика випадкової події.

Приклад 1.6. *Задано двозначне число, цифри якого різні. Дехто навмання назвав двозначне число. Знайти ймовірність, що це число і є задуманим.*

Розв'язання. Введемо подію: $A = \{\text{Назване число і є задуманим}\}$.

Кількість сприятливих подій (тих, що ми виберемо два однакових числа з різними цифрами) – $m = 81$. Ми маємо обрати з одного набору і другого. Тому кількість загальних подій (і сприятливих, і несприятливих) –

$$n = 81 \cdot 90 = 7290.$$

За формулою (1.2):

$$P(A) = \frac{81}{7290} = \frac{1}{90}.$$

Відповідь. $P(A) = \frac{1}{90}$.

1.6. Частість настання подій

Нехай простір елементарних подій складається з обмеженої кількості елементарних подій $\Omega = \{\omega_1, \dots, \omega_S\}$:

$$\Omega = \bigcup_{i=1}^S \{\omega_i\}.$$

F – позначення множини всіх подій що є всіма можливими різними підмножинами Ω , включаючи неможливу подію $V = \{\emptyset\}$.

Кількість елементів – $|F| = 2^S$.

Беремо довільну подію A , $A \in F$, n разів повторили випробування, для якого побудований простір Ω . n_A – кількість випробувань, в кожному з яких настала подія A .

Частістю настання події A в n довільних випробувань є $W_n = \frac{n_A}{n}$ (випробування зветься довільним, якщо на нього не накладаються умови).

Частість ще називають *відносною частотою*.

Властивості частоті

1. Частість належить сегменту від 0 до 1:

$$0 \leq W_n(A) \leq 1;$$

2. Частість достовірної події дорівнює одиниці

$$W_n(\Omega) = 1;$$

3. Якщо $A = \bigcup_{i=1}^K A_i$, $\forall i \neq j$, $A_i \cap A_j = \{\emptyset\}$ (попарно несумісні), то

$$W_n(A) = \sum_{i=1}^K W_n(A_i).$$

Доведення. Нехай провели випробування внаслідок якого настала подія A . За означенням об'єднання це означає, що в цьому випробуванні настала якась подія A_i ($i \in \overline{1; K}$). Кількість випробувань в яких вона настала – n_{A_i} . Оскільки A_i – попарно несумісні, то в цьому випробуванні не настала жодна інша подія A_j ($j \neq i$). Звідси $n_A = \sum_{i=1}^K n_{A_i} \Rightarrow \frac{n_A}{n} = \frac{\sum_{i=1}^K n_{A_i}}{n}$.

Теоретична ймовірність як формальна наука побудована для випадків, що описують такі випробування, для яких повинно виконуватись умова (1.9):

$$\exists \lim_{n \rightarrow +\infty} W_n(A) = \frac{n_A}{n} = P(A). \quad (1.9)$$

і ця границя однакова для будь-якої серії випробувань і називається *ймовірністю настання події A* .

1.7. Геометрична ймовірність

Основні поняття

Класичне означення ймовірності передбачає, що кількість елементарних наслідків скінченна. Якщо множина всіх елементарних наслідків нескінченна, застосовують геометричне означення ймовірності.

Геометрична ймовірність є розділом теорії ймовірностей, який стосується обчислення ймовірностей подій, пов'язаних із геометричними об'єктами, такими як точки, відрізки, площі та об'єми. Вона використовується в задачах, де ймовірність визначається через вимірювання довжин, площ або об'ємів.

Нехай простір елементарних подій Ω можна подати у вигляді деякого геометричного образу на прямій, площині або в просторі, а множину елементарних подій для події A – як частину цього геометричного образу (відрізок, площа, об'єм та ін.), то ймовірність події A визначається як відношення мір цих множин (формула (1.10)):

$$P(A) = \frac{m(A)}{m(\Omega)}. \quad (1.10)$$

Дана ймовірність називається *геометричною ймовірністю*.

Ймовірність на відрізку: Якщо на відрізку $\Omega = [a, b]$ випадковим чином вибирається точка, ймовірність того, що ця точка потрапить на певний підвідрізок $A = [c, d]$, обчислюється за формулою (рис. 1.5) (формула (1.11)):

$$P(A) = P(\text{точка потрапить на } [c, d]) = \frac{L_A}{L_\Omega} = \frac{d-c}{b-a}, \quad (1.11)$$

де L_A, L_Ω – довжини відповідних відрізків.

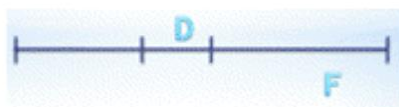


Рис. 1.5. Схема «Ймовірність на відрізку»

Ймовірність на площині: Якщо в прямокутнику Ω площі S_Ω випадковим чином вибирається точка, ймовірність того, що ця точка потрапить в певну область A з площею S_A , визначається як (рис. 1.6) (формула (1.12)):

$$P(A) = P(\text{точка потрапить в область } A) = \frac{S_A}{S_\Omega}. \quad (1.12)$$



Рис. 1.6. Схема «Ймовірність на площині»

Ймовірність в об'ємі: Якщо в просторі Ω об'ємом V_Ω випадковим чином вибирається точка, ймовірність того, що ця точка потрапить в певну область A об'єму V_A , обчислюється за формулою (рис. 1.7) (формула (1.13)):

$$P(A) = P(\text{точка потрапить в область } A) = \frac{V_A}{V_\Omega}. \quad (1.13)$$

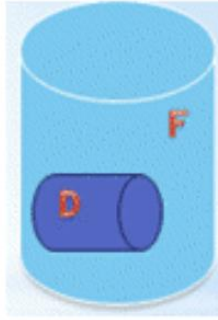


Рис. 1.7. Схема «Ймовірність в об'ємі»

Приклади застосування та обчислення

Задача Бюффона. Розглянемо площину, на якій проведені паралельні лінії, що розташовані на відстані d одна від одної. Кидаємо голку довжини l ($l \leq d$) на цю площину випадковим чином. Необхідно знайти ймовірність того, що голка перетне одну з ліній.

Розв'язання.

Геометричне моделювання

Позначимо кут між голкою і лініями як θ , де θ змінюється від 0 до π .

Позначимо відстань від центра голки до найближчої лінії як x , де $0 \leq x \leq \frac{d}{2}$.

Ймовірність перетину

Голка перетне лінію, якщо $x \leq \frac{l}{2} \cdot \sin\theta$.

Обчислення ймовірності

Ймовірність $P(A)$ перетину лінії можна знайти, взявши подвійний інтеграл відносно x і θ :

$$P(A) = \frac{2}{\pi \cdot d} \cdot \int_0^\pi \int_0^{\frac{l}{2} \sin\theta} dx d\theta = \frac{2}{\pi \cdot d} \cdot \int_0^\pi \left(\frac{l}{2} \cdot \sin\theta \right) d\theta = \frac{2 \cdot l}{\pi \cdot d}.$$

Приклад

Розглянемо випадок, коли $l = d$. У цьому випадку ймовірність того, що голка перетне лінію, дорівнює:

$$P(A) = \frac{2 \cdot d}{\pi \cdot d} = \frac{2}{\pi} \approx 0,6366.$$

Задача Бюффона є цікавим прикладом задачі з геометричної ймовірності і демонструє, як геометричні міркування можуть бути використані для обчислення ймовірностей у випадкових експериментах.

Ймовірність попадання в мішень. Якщо є кругова мішень радіусом R і внутрішнє коло радіусом r , ймовірність того, що випадково кинута точка потрапить у внутрішнє коло, обчислюється як:

$$P(A) = \left(\frac{r}{R} \right)^2.$$

Приклад 1.7. На відрізку довжини 20 см навмання взято дві точки. Яка ймовірність того, що відстань між ними не перевищує 10 см?

Розв'язання. Простір елементарних подій $\Omega = \{0 \leq l \leq 20 \text{ см}\}$, тоді $A = \{0 \leq l \leq 10 \text{ см}\}$ ($A \subset \Omega$). Згідно з (2.2) маємо:

$$P(A) = \frac{L_A}{L_\Omega} = \frac{10}{20} = \frac{1}{2}.$$

Відповідь. $P(A) = \frac{1}{2}$.

Приклад 1.8. Дві подруги Олена і Тетяна домовились зустрітися в певному місці у проміжку часу від 16 до 17 годин, а також про те, що той, хто прийде першим, чекатиме на другого протягом 15 хвилин. Знайти ймовірність того, що зустріч відбудеться.

Розв'язання. Подія $A = \{\text{зустріч відбудеться}\}$. Позначимо довжину часового проміжку $T = 1$ година, а моменти приходу кожної особи – x, y . Тоді подія A відбудеться за умови $|x - y| \leq \frac{1}{4}$, (різниця між моментами приходу Олени і Тетяни не перевищує 15 хв.), де $0 \leq x \leq 1, 0 \leq y \leq 1$. Зобразимо ці умови на площині в системі координат XOY (рис. 2.1).

Скориставшись геометричним означенням ймовірності, дістанемо:

$$P(A) = \frac{S_A}{S_\Omega} = \frac{7}{16}.$$

де S_A – площа шестикутника (заштрихована область на рис.1), а S_Ω – площа квадрата, яка відповідає часу $T = 1$ годині.

$$S_1 = 1 - 2 \cdot \frac{1}{2} \cdot \left(\frac{3}{4}\right)^2 = \frac{7}{16} \text{ (кв. од.)}, S_2 = 1 \text{ (кв. од.)}.$$

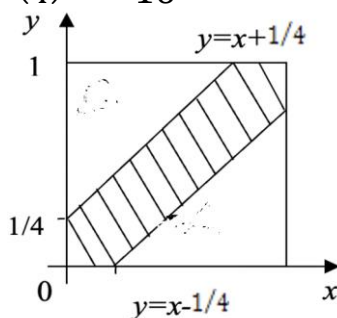


Рис. 1.8. Графік до задачі про зустріч (приклад 1.8)

Відповідь. $P(A) = \frac{7}{16}$.

Приклад 1.9. Задана множина $\Omega = (0 \leq x \leq 1, 0 \leq y \leq 1)$. Яка ймовірність того, що навмання взяті два числа (x, y) утворять координати точки, яка влучить в область $A = (0 \leq x \leq 1, 0 \leq y \leq x^2)$?

Розв'язання.

$$P(A) = \frac{S_A}{S_\Omega} = \frac{1}{3},$$

де S_A – площа криволінійної трапеції (заштрихована область на рис. 1.9), яка обчислюється за формулою: $S_A = \int_0^1 x^2 dx = \frac{1}{3}$ (кв. од.), а S_Ω – площа квадрата, яка дорівнює $S_\Omega = 1 \cdot 1 = 1$ (кв. од.).

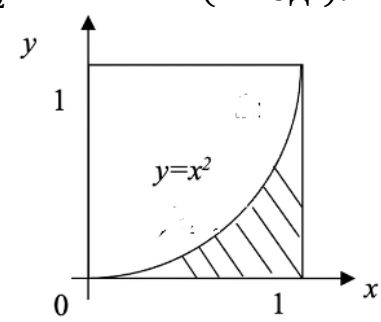


Рис. 1.9. Графік до прикладу 1.9

Відповідь. $P(A) = \frac{1}{3}$.

Приклад 1.10. В прямокутник зі сторонами a та b випадково кидають точку. Потрібно знайти ймовірність того, що ця точка потрапить в коло радіусом r , вписане в цей прямокутник.

Розв'язання.

$$P(A) = \frac{S_A}{S_\Omega} = \frac{\pi \cdot r^2}{a \cdot b}.$$

Відповідь. $P(A) = \frac{\pi \cdot r^2}{a \cdot b}$.

Приклад 1.11. В куб зі стороною a випадково кидають точку. Потрібно знайти ймовірність того, що ця точка потрапить у вписану кулю радіусом $r = \frac{a}{2}$.

Розв'язання.

$$P(A) = \frac{V_A}{V_\Omega} = \frac{\frac{4}{3} \cdot \pi \cdot r^3}{a^3} = \frac{\frac{4}{3} \cdot \pi \cdot \left(\frac{a}{2}\right)^3}{a^3} = \frac{\frac{\pi}{6} \cdot a^3}{a^3} = \frac{\pi}{6}.$$

Відповідь. $P(A) = \frac{\pi}{6}$.

Приклад 1.12. В круг вписано квадрат. У круг навмання кидають точку. Знайти ймовірність того, що вона виявиться всередині квадрата.

Розв'язання. Для кращого уявлення виконаємо графічний рисунок до задачі

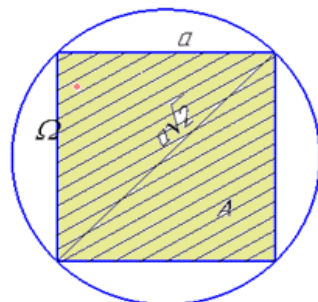


Рис. 1.10. Рисунок до прикладу 1.12

Нехай сторона квадрата A дорівнює a , тоді його площа – $S_A = a^2$. Діаметр круга Ω , що описаний навколо квадрата дорівнює діагоналі квадрата рівний $d = a \cdot \sqrt{2}$, тоді площа цього круга:

$$S_{\Omega} = \frac{\pi \cdot a^2}{2}.$$

Оскільки квадрат знаходиться всередині круга $A \subset \Omega$, то ймовірність того, що навмання кинута точка в круг виявиться всередині квадрата обчислюється за формулою:

$$P(A) = \frac{S_A}{S_{\Omega}} = \frac{2}{\pi} \approx 0,64.$$

Відповідь. $P(A) \approx 0,64$.

Приклад 1.13. Два судна повинні підійти до одного причалу. Поява суден – незалежні події, рівноможливі протягом доби. Знайти ймовірність того, що одному з суден доведеться чекати звільнення причалу, якщо час стоянки першого судна – 1 год, а другого – 2 год.

Розв’язання. Нехай перше судно прийшло до причалу в u год і стоїть 1 годину, а друге судно прийшло до того ж причалу в v год і стоїть 2 години. За простір елементарних подій виберемо

$$\Omega = \{(u, v): 0 \leq u \leq 24; 0 \leq v \leq 24\}.$$

Звідси $S_A = 24^2 = 576$ (площа). Тоді подія $A = \{\text{судно чекає звільнення причалу}\}$ відповідає множині (заштрихована фігура на рис. 1.11)

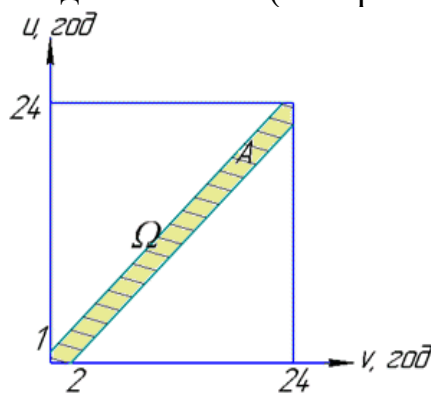


Рис. 1.11. Рисунок до прикладу 1.13

$$A = \{(u, v): 1 \leq u - v \leq 2; 0 \leq u \leq 24; 0 \leq v \leq 24\},$$

звідси її площа:

$$S_A = 24^2 - (0,5 \cdot 23^2 + 0,5 \cdot 22^2) = 69,5.$$

Оскільки $A \subset \Omega$, то ймовірність того, що одному з суден доведеться чекати звільнення причалу знайдемо за формулою:

$$P(A) = \frac{S_A}{S_{\Omega}} = \frac{69,5}{576} \approx 0,12.$$

Відповідь. $P(A) \approx 0,12$.

Приклад 1.14. На площину, на якій зображені паралельні прямі, що розташовані одна від одної на відстані 6 см, навмання кинули круг радіусом 1 см. Знайти ймовірність того, що круг не перетне жодної прямої. По

замовчуванню розуміється, що ймовірність попадання точки на відрізок пропорційна довжині відрізка і не залежить від його розташування.

Розв'язання: Нехай $D = 6$ см – відстань між паралельними прямими; $d = 2 \cdot r = 2$ см – діаметр круга, який кинули на площину. Схематично це можна візуалізувати наступним чином:

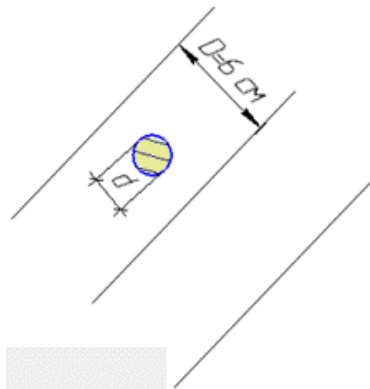


Рис. 1.12. Рисунок до прикладу 1.14

Ймовірність того, що круг не перетне жодної прямої знаходимо наступним чином:

$$P(A) = \frac{L_A}{L_\Omega} = \frac{D-d}{D} = \frac{2}{3} \approx 0,67,$$

де $D-d = 4$ см – найкоротша відстань від кола (яким обмежений круг) до паралельних прямих. Як бачимо така ймовірність є протилежною до того, що навімання кинутий круг перетне хоча б одну пряму:

$$P(\bar{A}) = 1 - P(A) = \frac{1}{3}.$$

Відповідь. $P(A) = \frac{1}{3}$.

Приклад 1.15. Навмання взято два додатних числа x і y , кожне з яких не перевищує одиниці. Знайти ймовірність того, що сума $x + y$ не перевищує одиниці, а добуток $x \cdot y$ не менше 0,09.

Розв'язання: За умовою задачі маємо

$$\begin{cases} x + y \leq 1, \\ x \cdot y \geq 0,09, \end{cases}$$

а також обмеження на змінні

$$0 \leq x \leq 1; 0 \leq y \leq 1.$$

Розв'язавши систему нерівностей отримаємо:

$$0,1 \leq x \leq 0,9; 0,1 \leq y \leq 0,9.$$

За простір елементарних подій виберемо

$$\Omega = \{(x, y): 0 \leq x \leq 1; 0 \leq y \leq 1\}.$$

Звідси $S_\Omega = 1$ (площа).

Тоді подія

$$A = \begin{cases} 0 \leq x + y \leq 1, \\ 0,09 \leq x \cdot y \leq 1 \end{cases}$$

відповідає множині (заштрихована фігура на рис. 1.15)

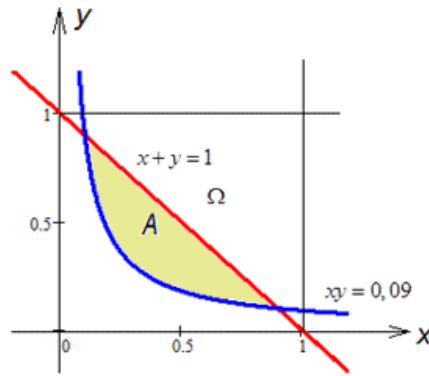


Рис. 1.13. Рисунок до прикладу 1.15

Побудова рисунків до задач на геометричну ймовірність дає повніше розуміння, що потрібно знайти. Якість рисунку ролі не грає, головне зміст, тому всюди де потрібно будуйте графічні схемки до умов задач.

$A = \{(x, y): 0 \leq x + y \leq 1; 0,09 \leq x \cdot y \leq 1; 0 \leq x \leq 1; 0 \leq y \leq 1\}$
з відповідними межами області A :

$$\begin{aligned} 0,1 &\leq x \leq 0,9; \\ \frac{0,09}{x} &\leq y \leq 1 - x. \end{aligned}$$

На основі меж змінних можемо знайти площу між кривими через визначений інтеграл:

$$S_A = \int_{0,1}^{0,9} \left(1 - x - \frac{0,09}{x} \right) dx \approx 0,2023.$$

Оскільки $A \subset \Omega$, то ймовірність шуканої події знайдемо за формулою частки:

$$P(A) = \frac{S_A}{S_\Omega} = 0,2023.$$

Відповідь. $P(A) = 0,2023$.

ЛЕКЦІЯ 2

ПАРАДОКСИ В ТЕОРІЇ ЙМОВІРНОСТЕЙ. ЙМОВІРНІСНИЙ ПРОСТІР. УМОВНА ЙМОВІРНІСТЬ. НЕЗАЛЕЖНІ ПОДІЇ. НЕСУМІСНІ ПОДІЇ

2.1. Парадокси в теорії ймовірностей

Парадокс (*Παράδοξος* – *несподіваний; дивний*) в широкому сенсі – висловлювання, думка, міркування, яке розходиться із загальноприйнятою думкою і здається нелогічним або таким, що суперечить здоровому глузду (найчастіше лише при поверхневому розумінні).

У логіці «парадоксом» називають формально-логічні суперечності, які виникають при збереженні логічної правильності міркування. Парадокс виникає, коли два взаємовиключних (суперечать) судження виявляються в рівній мірі доказовими.

Парадокси в теорії ймовірностей можуть демонструвати, як інтуїція людей часто виявляється неправильною в контексті математичних ймовірностей. Ось кілька відомих парадоксів у теорії ймовірностей.

Парадокс Бертрана

Парадокс Бертрана – проблема класичного визначення теорії ймовірностей. Жозеф Бертран описав парадокс у своїй роботі *Calcul des probabilités* (1888) в якості прикладу того, що ймовірність не може бути чітко визначена, поки не визначений механізм або метод вибору випадкової величини.

Парадокс

Розглянемо рівносторонній трикутник, вписаний в коло. Навмання вибирається хорда кола. Яка ймовірність того, що обрана хорда довша сторони трикутника?

Розв'язання

Бертран запропонував три рішення, очевидно вірних, але дають різний результат.

1. Метод «випадкових кінців»: навімання виберемо дві точки на колі і проведемо через них хорду. Щоб порахувати шукану ймовірність, уявімо, що трикутник повернутий так, що одна з його вершин збігається з кінцем хорди. Зауважимо, що якщо інший кінець хорди лежить на дузі між двома іншими вершинами трикутника, то довжина хорди більше сторони трикутника. Довжина розглянутої дуги дорівнює третині довжини кола, слідує класичним визначенням, шукана ймовірність дорівнює $\frac{1}{3}$ (рис. 2.1).

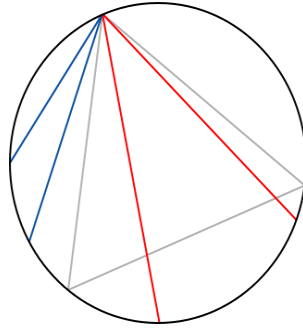


Рис 2.1. Випадкові хорди, обрані за методом 1; червоні – довші сторони трикутника, сині – коротші.

2. Метод «випадкового радіусу»: зафіксуємо радіус кола, навмання виберемо точку на радіусі. Побудуємо хорду, перпендикулярну зафіксованим радіусу, що проходить через обрану точку. Для знаходження шуканої ймовірності уявімо, що трикутник повернутий так, що одна з його сторін перпендикулярна зафіксованим радіусу. Хорда довше сторони трикутника, якщо її центр ближче до центру, ніж точка перетину трикутника з зафіксованим радіусом. Сторона трикутника ділить навпіл радіус, отже ймовірність вибрати хорду довше сторони трикутника $\frac{1}{2}$ (рис. 2.2).

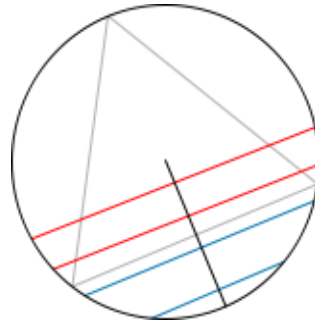


Рис 2.2. Випадкові хорди, обрані за методом 2

3. Метод «випадкового центру»: виберемо навмання довільну точку всередині кола і побудуємо хорду з центром у вибраній точці. Хорда довше сторони рівностороннього трикутника, якщо обрана точка знаходиться всередині кола, вписаного в трикутник. Площа вписаного кола є $\frac{1}{4}$ від площі більшого, значить, вихідна ймовірність дорівнює $\frac{1}{4}$ (рис. 2.3).

$$r = \frac{\sqrt{3}}{6} \cdot a, R = \frac{\sqrt{3}}{3} \cdot a.$$

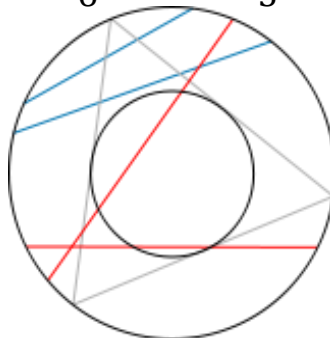


Рис 2.3. Випадкові хорди, обрані за методом 3

Парадокс де Мере

Історія парадоксу

Мабуть, Лейбніц вперше розказав про те, як відомий французький гравець де Мере по дорозі в свій маєток в Пуату зустрів Блеза Паскаля і поставив перед ним дві задачі. Ці задачі Паскаль обговорював у 1654 р. у своєму листуванні з П'єром де Ферма, що проживав у Тулузі. Обидва вчених дійшли однакового результату, що вельми обрадувало Паскаля. В своєму листі він відмічав: «Я бачу, що істина однакова і в Тулузі, і в Парижі» Ейнштейн Оре, професор Йельського університету, правда, стверджує, що парадокси, приписувані де Мере, насправді були відомі набагато раніше, справа лише в тому, що Паскаль про них не знав. Невірно й те, що шевальє був пристрасним гравцем. Парадокси цікавили його скоріш теоретично, і тому його не задовольнило, що Паскаль «всього-навсього» розв'язав задачу, підтвердивши правильну відповідь, яку де Мере вже знав. Шевальє не зміг зрозуміти, як пояснюється парадокс.

Парадокс

При чотирьох підкиданнях одного грального кубика ймовірність того, що як мінімум один раз випаде 1, більша $\frac{1}{2}$. В той же час при 24 підкиданнях двох кубиків ймовірність того, що як мінімум один раз випадуть дві 1 одночасно, менша $\frac{1}{2}$. Це здається дивним, оскільки шанси отримати одну 1 в шість раз більші, ніж шанси випадання двох 1, а 24 якраз в 6 раз більше 4.

Пояснення парадоксу

Прості підрахунки показують, що при k підкиданнях одного кубика потрібна ймовірність дорівнює

$$1 - \left(\frac{5}{6}\right)^k,$$

що менше $\frac{1}{2}$ при $k = 3$ і більше $\frac{1}{2}$ при $k = 4$. Так само друга потрібна ймовірність дорівнює

$$1 - \left(\frac{35}{36}\right)^k,$$

що менше $\frac{1}{2}$ при $k = 24$ і більше $\frac{1}{2}$, починаючи з $k = 25$. Отже, «критичне значення» для одного кубика дорівнює 4, а для двох дорівнює 25. Цей безумовно правильний розв'язок не задовольнив де Мере, оскільки відповідь він вже знав, але з розв'язку так і не зрозумів, чому відповідь не узгоджується з «правилом пропорційності критичних значень», стверджуючим, що коли ймовірність зменшується в 6 раз, то критичне значення зростає в 6 раз ($4:6 = 24:36$). Абрахам де Муавр (1667-1754) у своїй книзі «Доктрина шансів», опублікованій в 1718 р., довів, що «правило пропорційності критичних значень» недалеко від істини. Так, якщо p – ймовірність деякої події, то критичне значення k можна знайти з рівняння

$$(1 - p)^x = \frac{1}{2}.$$

Тоді критичне значення буде найменшим цілим числом, більшим за x .

Оскільки

$$x = -\frac{\ln(2)}{\ln(1-p)} = \frac{\ln(2)}{(p + \frac{p^2}{2} + \dots)},$$

то при можливості знехтувати значенням $\frac{p^2}{2}$ «правило пропорційності критичних значень» близьке до істини. Сам де Муавр використовував наближену формулу $x \approx \frac{\ln 2}{p}$ для дослідження питань, пов'язаних з Лондонською лотереєю. В цьому випадку значення p було $\frac{1}{32}$, а для $p = \frac{1}{32}$ точне значення $x = 22,135$, а наближена формула дає 22,08. Парадокс де Мере виникає тому, що при $p = \frac{1}{6}$ величина $\frac{p^2}{2}$ і інші доданки в знаменнику ще не досить малі. Отже, «правило пропорційності критичних значень» є правилом, вірним асимптотично, і помилка від його застосування росте з ростом p . Це й є справжнім розв'язком парадоксу.

Зауваження

Типове неправильне розв'язання задачі де Мере йде аж до Кардано. Він міркував так: ймовірність отримати пару одиниць дорівнює $\frac{1}{32}$, значить, щоб з ймовірністю $\frac{1}{2}$ отримати пару одиниць як мінімум один раз, треба зробити рівно 18 підкидань. Якщо так міркувати, то при підкиданні кубика більше 36 раз така ймовірність стане більшою 1!.

Парадокс роздачі подарунків; травми, спричинені кіньми

Історія парадоксу

Парадокс роздачі подарунків розглядався в книзі Ремона де Монмора, опублікованій в Парижі в 1708 р.

Парадокс

Декілька чоловік вирішили зробити один одному подарунки таким чином. Кожен з них приносить подарунок. Подарунки складаються разом, змішуються і випадково розподіляються серед учасників. Парадоксально, але ймовірність того, що ніхто не отримає свій власний подарунок, менша $\frac{1}{2}$ (крім випадку, коли учасників двоє, і ця ймовірність дорівнює $\frac{1}{2}$).

Пояснення парадоксу

Обчислення показують, що

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!}.$$

Якщо збирається щонайменше 6 чоловік, то $p_n \approx \frac{1}{e} \approx 0,3679$ з точністю до чотирьох знаків після крапки. Ймовірність конкретного співпадіння

дорівнює $\frac{1}{n}$ і прямує до 0 при збільшенні n . Цей парадокс підтверджує, що «по крапельці – море, по стеблинці – стіг».

Зауваження

1. Розглянемо дещо іншу проблему. Нехай подарунки розподіляються так, що кожен чоловік може отримати будь-який подарунок з однаковою ймовірністю незалежно від розподілу інших подарунків. Нехай подія A полягає в тому, що конкретний чоловік не отримає подарунку. Тоді ймовірність A дорівнює

$$q_n = \frac{(n-1)^n}{n^n} = \left(1 - \frac{1}{n}\right)^n$$

і прямує так само до $\frac{1}{e}$. Розглянемо тепер випадок, коли кількість людей n не обов'язково співпадає з кількістю подарунків m . В цьому випадку шукана ймовірність дорівнює

$$q_n = \frac{(n-1)^m}{n^m} = \left(1 - \frac{1}{n}\right)^m.$$

Якщо $\frac{m}{n}$ прямує до λ при $n \rightarrow +\infty$, то ця ймовірність прямує до $e^{-\lambda}$.

Узагальнюючи цей результат, отримаємо: ймовірність того, що конкретна людина отримає рівно k подарунків, прямує до $\frac{\lambda^k \cdot e^{-\lambda}}{k!}$ при $n \rightarrow +\infty$. Ми отримали добре відомий розподіл Пуассона. Повертаючись до «парадоксу розподілу подарунків», отримаємо, що кількість людей, яким дістануться їх власні подарунки, також описується законом Пуассона з параметром λ .

2. Поняття розподілу Пуассона вперше з'явилося в книзі Симеона Деніса Пуассона. В роботі «Дослідження про ймовірність судових вироків з кримінальних та цивільних справ», опублікованій в 1837 р., розділ 81 присвячений застосуванням теорії ймовірностей до судової практики. Можливості широкого застосування і велика важливість пуассонівського розподілу залишилися незрозумілими в середині минулого сторіччя; більш того, про нього майже повністю забули. Але після 1894 р. цей розподіл використовували при вивченні дивного феномену. За 20 років між 1875 і 1894 роками в 14 різних кавалерійських корпусах німецької армії була зібрана статистика трагічних випадків, при яких солдат загинув від удара копита коня. За даними 280 спостережень (280 спостережень = 20 років · 14 корпусів), загинуло 196 солдат, тобто в середньому за рік $\lambda = 0,7$, якби кількість трагічних випадків описувалася розподілом Пуассона з параметром 0,7, то можна було б очікувати, що при 280 спостереженнях в 139 випадках смертей немає, в 97 – 1 смерть, в 34 випадках – 2 смерті і т.д. А що показала статистика? В дійсності дані були 140, 91, 32 і т.д. Практика і теорія були настільки узгоджені, що навряд чи можна було чекати більшого.

Парадокс Монті Холла

Ситуація

Уявіть, що ви берете участь у телевізійній грі. Перед вами три двері. За одними з них знаходиться автомобіль, а за іншими двома – кози. Ви вибираєте одну з дверей. Ведучий (який знає, що за якими дверима) відкриває одні з двох інших дверей, за якими точно стоїть коза. Потім ведучий запитує вас, чи хочете ви змінити свій вибір на інші двері.

Питання

Чи варто змінювати свій вибір?

Розв'язок

Так, варто змінити свій вибір. При початковому виборі ймовірність вибрати автомобіль становить $\frac{1}{3}$, а ймовірність вибрати козу – $\frac{2}{3}$. Після того як ведучий відкриває двері з козою, ймовірність того, що автомобіль знаходиться за одними з двох інших дверей, залишається $\frac{2}{3}$. Отже, зміна вибору підвищує ваші шанси на виграш до $\frac{2}{3}$.

Парадокс дня народження

Ситуація

У групі з 23 осіб якова ймовірність того, що хоча б двоє людей мають однаковий день народження?

Питання

Наскільки велика ця ймовірність?

Розв'язання

Ймовірність того, що хоча б двоє з 23 людей мають однаковий день народження, становить приблизно 50,7%. Це значно більше, ніж інтуїтивно здається багатьом людям. Для групи з 57 осіб ця ймовірність перевищує 99%.

Парадокс двох конвертів

Ситуація

Перед вами два конверти. В одному з них міститься вдвічі більша сума, ніж в іншому. Ви вибираєте один конверт і відкриваєте його, щоб побачити суму X . Після цього вам пропонують обміняти конверт на інший.

Питання

Чи варто обмінювати конверт?

Розв'язання

На перший погляд здається, що обмін конвертів завжди вигідний, оскільки середнє очікуване значення при обміні становить $0,5 \cdot 2 \cdot X + 0,5 \cdot 0,5 \cdot X = 1,25 \cdot X$, що більше, ніж X . Однак, цей парадокс виникає через неправильне розуміння ймовірностей і очікуваних значень. В обох конвертах знаходиться однаковий розподіл сум, і обмін не змінює очікуваної вартості вашого вибору.

Парадокс Сімпсона

Ситуація

Парадокс Сімпсона демонструє, як загальні тенденції в даних можуть змінюватися, коли дані поділяються на групи.

Питання

Як це може відбуватися?

Розв'язання

Розглянемо дві групи пацієнтів, які отримують два різних лікування. Загалом, лікування А може виглядати більш ефективним. Проте, якщо розбити пацієнтів на підгрупи (наприклад, за віком або ступенем захворювання), може виявитися, що лікування В є ефективнішим у кожній підгрупі. Це відбувається через те, що пропорції пацієнтів у підгрупах впливають на загальну статистику.

Парадокс Санкт-Петербурга

Ситуація

Гра в казино пропонує вам наступний експеримент: ви підкидаєте монету доти, доки не випаде орел. Виграш дорівнює 2^n доларів, де n – кількість підкидань.

Питання

Скільки ви готові заплатити за участь у цій грі?

Розв'язання

Математичне очікування виграшу в цій грі є нескінченним, оскільки

$$\sum_{i=1}^n \left(\frac{1}{2^n} \cdot 2^n \right) = +\infty.$$

Проте в реальному житті люди готові заплатити лише обмежену суму, оскільки ймовірність виграти велику суму дуже мала, а психологічні та фінансові обмеження грають роль.

2.2. Ймовірнісний простір

Ймовірнісний простір – поняття, що ввів А.М. Колмогоров в 30-х роках ХХ століття для формалізації поняття ймовірності, але дало початок бурхливому розвитку теорії ймовірностей як строгої математичної дисципліни.

Ймовірнісний простір – це трійка $(\Omega, \mathcal{F}, P)$, де:

Ω – простір елементарних подій;

$\mathcal{F}$ – сігма-алгебра підмножин множини Ω (випадкових подій);

P – ймовірнісна міра або ймовірність, тобто сігма-адитивна скінченна міра, така, що $P(\Omega) = 1$.

Нагадаємо, що *σ -алгебра (сігма-алгебра)* — алгебра множин, замкнена щодо операції зліченого об'єднання.

Поняття сігма-алгебри має важливе значення для визначення мір множин, в математичному аналізі та теорії ймовірностей.

Міра множини — спільна назва різних типів узагальнень понять евклідової довжини, площі плоских фігур та n -вимірного об'єму для загальніших просторів, тощо.

Якщо зворотнє не вказане явно, то зазвичай мається на увазі зліченно-адитивна міра.

Нехай задано простір X (в нашому випадку $X = \Omega$) з виділеним класом підмножин $\mathcal{F}$, замкненим щодо скінчених перетинів та об'єднань. Функція $\mu: \mathcal{F} \rightarrow [0, +\infty)$ (в нашому випадку в якості функції μ розглядаємо функцію $P: \mathcal{F} \rightarrow [0, 1]$) називається **скінчено-адитивною мірою**, якщо вона задовольняє наступним умовам:

1. $\mu(\emptyset) = 0$;
2. Якщо X_i ($i = 1, 2, \dots, k$) $\subset \mathcal{F}$ — скінченне сімейство множин, що попарно не перетинаються $X_i \cap X_j, i \neq j$, то справедлива рівність (2.1):

$$\mu\left(\bigcup_{i=1}^k X_i\right) = \sum_{i=1}^k \mu(X_i). \quad (2.1)$$

В нашому випадку $X_i = A_i$, де A_i — це подія з простору елементарних подій Ω . Таким чином, отримуємо (2.2):

$$P\left(\bigcup_{i=1}^k A_i\right) = \sum_{i=1}^k P(A_i), \quad (2.2)$$

де $P(A_i)$ обчислюється за формулою (1.2) минулої лекції (класичне означення ймовірності).

Нехай задано простір X (в нашому випадку $X = \Omega$) з виділеним класом підмножин $\mathcal{F}$, замкненим щодо скінчених перетинів та об'єднань. Функція $\mu: \mathcal{F} \rightarrow [0, +\infty)$ (в нашому випадку в якості функції μ розглядаємо функцію $P: \mathcal{F} \rightarrow [0, 1]$) називається **злічено-адитивною мірою**, якщо вона задовольняє наступним умовам:

1. $\mu(\emptyset) = 0$;
2. Якщо X_i ($i = 1, 2, \dots, k, \dots$) $\subset \mathcal{F}$ — скінченне сімейство множин, що попарно не перетинаються $X_i \cap X_j, i \neq j$, то справедлива рівність (2.3):

$$\mu\left(\bigcup_{i=1}^{+\infty} X_i\right) = \sum_{i=1}^{+\infty} \mu(X_i). \quad (2.3)$$

В нашому випадку $X_i = A_i$, де A_i — це подія з простору елементарних подій Ω . Таким чином отримуємо (2.4):

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i), \quad (2.4)$$

де $P(A_i)$ обчислюється за формулою (1.9) минулої лекції (геометрична ймовірність).

Примітки

- Довільна злічено-адитивна міра є скінчено-адитивною, але не навпаки,
- Якщо міра всього простору скінчена, тобто $\mu(X) < +\infty$, то така міра називається скінченою. В протилежному випадку міра нескінчена.

Зауваження

- Елементарні події (елементи множини Ω) — це результати випадкового експерименту, з яких в даному експерименті відбулася лише одна конкретна подія,
- Кожна випадкова подія (елементи $\mathcal{F}$) — це підмножина множини Ω ,
- Вимога, що $\mathcal{F}$ є сігма-алгеброю підмножин Ω , дозволяє, зокрема, говорити про ймовірність випадкової події, об'єднання зліченої кількості

випадкових подій, а також про ймовірність доповнення будь-якої події (протилежної події до даної).

Приклад 2.1. Розглянуто експеримент з підкиданням монетки. Задати ймовірнісний простір.

Розв'язання. Простір елементарних подій – $\mathcal{U} = \Omega = \{0,1\}$, сігма-алгебра підмножин множини Ω – $\mathcal{F} = \{\{0\}, \{1\}, \{0,1\}, \{\emptyset\}\}$, ймовірності: $P(\{0\}) = 0,5$; $P(\{1\}) = 0,5$; $P(\{0,1\}) = 1$; $P(\{\emptyset\}) = 0$.

2.3. Умовні ймовірності

Часто доводиться розглядати ймовірність настання події A внаслідок випробування, якщо відомо що у цьому випробуванні вже відбулась подія B , що має додатну ймовірність.

Цю ймовірність називають *умовною ймовірністю* та позначають $P(A/B)$.

До того, як дати загальне формальне означення умовної ймовірності, розглянемо найпростішу ситуацію.

Нехай простір елементарних подій Ω складається з m рівноймовірних елементарних подій, настанню події B сприяє t елементарних подій, події $A \cap B = A \cdot B$ сприяє r елементарних подій (рис. 2.4).

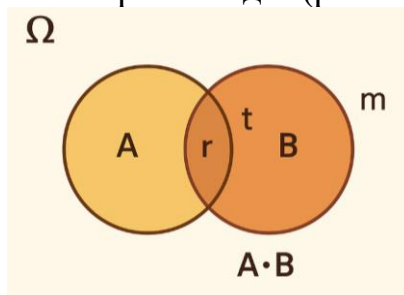


Рис. 2.4. Перетин множин A та B

Зауважимо, що згідно класичному означенню ймовірності

$$P(B) = \frac{t}{m}, P(A \cap B) = \frac{r}{m}.$$

Визначимо умовну ймовірність $P(A/B)$. Відомо, що у цьому випробуванні подія B вже відбулась і, таким чином, множина усіх можливих у випробуванні елементарних подій звузилась до t елементарних подій. Внаслідок рівноможливості, ймовірність кожної елементарної події, що належить B , тепер можна вважати рівною $\frac{1}{t}$. Події A сприяє тепер тільки r елементарних подій. Внаслідок рівноможливості природньо визначити

$$P(A/B) = \frac{r}{t} = \frac{\frac{r}{m}}{\frac{t}{m}} = \frac{P(A \cap B)}{P(B)}.$$

Розглянемо тепер більш загальну ситуацію, коли простір Ω не обов'язково складається з рівно можливих елементарних подій. Нехай проведено n випробувань і в n_B з них відбулась подія B , а в $n_{A \cap B}$

випробуваннях відбулася подія $A \cap B$. Визначимо умовну частість настання події A за умови, що відбулась подія B :

$$W_n(A/B) = \frac{n_{A \cap B}}{n_B} = \frac{\frac{n_{A \cap B}}{n}}{\frac{n_B}{n}} = \frac{W_n(A \cap B)}{W_n(B)}.$$

Раніше вже відмічалось, що частість настання події у певному розумінні збігається до ймовірності за кількості випробувань, що прямує до нескінченності.

Сказане дозволяє дати наступне загальне означення умовної ймовірності.

Нехай $(\Omega, \mathcal{F}, P)$ – ймовірнісний простір, $A \in \mathcal{F}, B \in \mathcal{F}, P(B) \neq 0$.

Умовною ймовірністю події A за умови, що відбулася подія B , називається величина, що обчислюється за формулою (2.5):

$$P(A/B) = \frac{P(A \cap B)}{P(B)}. \quad (2.5)$$

При $P(B) = 0$ умовна ймовірність $P(A/B)$ – невизначена.

Можна переконатися, що умовні ймовірності задовольняють наступним **властивостям**:

1. $P(A/B) \geq 0$;
2. $P(\Omega/B) = 1$;
3. $P(B/B) = 1$;
4. Якщо $\{A_i\}$ – послідовність парно несумісних випадкових подій ($A_i \cap A_j = \emptyset, i \neq j$), то

$$P\left(\bigcup_{i=1}^{+\infty} (A_i/B)\right) = \sum_{i=1}^{+\infty} P(A_i/B).$$

Доведення. Властивості 1) – 3) перевірте самостійно.

Доведемо 4). Маємо:

$$\begin{aligned} P\left(\bigcup_{i=1}^{+\infty} (A_i/B)\right) &= P\left(\left(\bigcup_{i=1}^{+\infty} A_i\right)/B\right) = \frac{P((\bigcup_{i=1}^{+\infty} A_i) \cap B)}{P(B)} = \frac{P(\bigcup_{i=1}^{+\infty} (A_i \cap B))}{P(B)} \\ &= \frac{\sum_{i=1}^{+\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{+\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{+\infty} P(A_i/B). \end{aligned}$$

Якщо позначити через $\mathcal{F}_B$ σ -алгебру усіх множин виду $A \cap B$, то з властивостей 1) – 4) виходить, що $(B, \mathcal{F}_B, P(B))$ – ймовірнісний простір.

З означення умовної ймовірності виходить (2.6)

$$P(A \cap B) = P(B) \cdot P(A/B) = P(A) \cdot P(B/A), \quad (2.6)$$

якщо $P(A) > 0, P(B) > 0$.

Співвідношення (2.6) має назву *формула множення ймовірностей двох подій*.

Вона узагальнюється на випадок довільної кількості k множників.

Теорема 2.1. Нехай є система випадкових подій $A_1, A_2, \dots, A_k$, і нехай

$$P(A_1, A_2, \dots, A_k) > 0,$$

тоді справедлива рівність (2.7)

$$P(A_1, \dots, A_k) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/(A_1 \cap A_2)) \cdot \dots \cdot P(A_k/(A_1 \cap \dots \cap A_{k-1})) \quad (2.7)$$

Доведення. Перш за все зауважимо, що з

$$\bigcap_{i=1}^{k-1} A_i \subseteq \bigcap_{i=1}^{k-2} A_i \subseteq \dots \subseteq A_1 \cap A_2 \subseteq A_1$$

та з $P(A_1, A_2, \dots, A_k) > 0$ виходить, що

$$P\left(\bigcap_{i=1}^{k-s} A_i\right) > 0, s = \overline{1; k-1}.$$

Отже усі умовні ймовірності $P(A_n / \bigcap_{i=1}^{n-1} A_i)$ визначено.

Далі використовуємо метод математичної індукції. Раніше встановлено, що формула вірна при $k = 2$. Нехай вона вірна для добутку $(k-1)$ випадкових подій, тобто

$$\begin{aligned} P(A_1, \dots, A_{k-1}) &= P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/(A_1 \cap A_2)) \cdot \dots \\ &\cdot P(A_{k-1}/(A_1 \cap \dots \cap A_{k-2})). \end{aligned}$$

Позначимо $B = A_1 \cap \dots \cap A_{k-1}$, так що $P(A_1, A_2, \dots, A_{k-1}) = P(B)$.

Тоді

$$\begin{aligned} P(A_1, A_2, \dots, A_k) &= P(A_k \cap B) = P(B) \cdot P(A_k/B) = \\ &= P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/(A_1 \cap A_2)) \cdot \dots \\ &\cdot P(A_k/(A_1 \cap \dots \cap A_{k-1})). \end{aligned}$$

Таким чином, формула (2.7) доведена.

Формула (2.7) називається *формулою множення ймовірностей k подій*.

Приклад 2.2. На шістьох картках намальовані літери, з яких можна зіставити слово КАРЕТА. Картки перегортуються літерами до столу та добре перемішуються. Далі витягуються три картки по одній. Треба обчислити ймовірність того, що в порядку витягування ми одержимо слово РАК.

Розв'язання. Введемо події:

$A = \{\text{витягнути картку з літерою Р}\};$

$B = \{\text{витягнути картку з літерою А}\};$

$C = \{\text{витягнути картку з літерою К}\}.$

Нас цікавить ймовірність добутку цих подій РАК. За формулою множення ймовірностей маємо

$$P(\text{РАК}) = P(A) \cdot P(B/A) \cdot P(C/(A \cap B)) = \frac{1}{6} \cdot \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{60}.$$

Відповідь. $P(\text{РАК}) = \frac{1}{60}$

2.4. Незалежні та залежні події

Нехай $(\Omega, \mathcal{F}, P)$ – ймовірнісний простір. Випадкові події A і B ($A \in \mathcal{F}, B \in \mathcal{F}$) називаються *незалежними*, якщо поява однієї події призводить до появи другої події.

Якщо ця умова не виконується, то події називаються *залежними*.

Таке означення незалежних подій, на перший погляд, має чисто формальний характер. Якщо, однак, відмовитись від розгляду випадкових подій нульової ймовірнісної міри, введене поняття незалежності має близьке до природних уявлень тлумачення. Доведемо наступне твердження.

Теорема 2.2. Нехай $P(B) > 0$. Випадкові події A і B незалежні тоді й тільки тоді, коли $P(A/B) = P(A)$ (поява події B не змінює ймовірності події A).

Доведення. Дійсно, якщо A і B незалежні та $P(B) > 0$, то

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A).$$

Якщо ж відомо, що

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A),$$

то $P(A \cap B) = P(A) \cdot P(B)$, а це означає, що A і B незалежні.

За допомогою теореми 2.2 та формули (2.6) для двох незалежних подій A і B отримуємо формулу (2.8):

$$P(A \cap B) = P(A) \cdot P(B). \quad (2.8)$$

Таким чином, для подій ненульової ймовірнісної міри маємо еквівалентне означення незалежності: дві події A і B *незалежні*, коли умовна ймовірність настання однієї з них, за умови, що друга відбулась, дорівнює безумовній ймовірності першої події.

Мають місце наступні **властивості незалежних подій**:

Якщо A і B незалежні, то події A і $\bar{B}$, $\bar{A}$ і B , $\bar{A}$ і $\bar{B}$ також незалежні.

Доведення: Доведемо, що A і $\bar{B}$ незалежні (решту довести самостійно). Зауважимо, що $(A \cap \bar{B}) \cup (A \cap B) = A$, $P(A \cap \bar{B}) + P(A \cap B) = P(A)$.

Тому

$$P(A \cap \bar{B}) = P(A) - P(A \cap B).$$

Якщо A і B незалежні, то

$P(A \cap \bar{B}) = P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = P(A) \cdot P(\bar{B})$.
Отже, A і $\bar{B}$ незалежні. Що й треба було довести.

Довести самостійно, що у випадку, коли A і B незалежні та $P(A) > 0$, $P(B) > 0$, то A і B обов'язково сумісні.

Випадкові події $A_1, A_2, \dots, A_n$, *незалежні у сукупності*, якщо для будь-якої підмножини $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$ для довільного $i, 1 \leq i \leq n$, та для довільного $1 \leq k \leq n$ виконується (2.9):

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k}). \quad (2.9)$$

Іншими словами, події $A_1, A_2, \dots, A_n$ *незалежні в сукупності*, якщо ймовірність одночасного настання будь-якої комбінації цих подій дорівнює добутку їх індивідуальних ймовірностей.

Отже, формули (2.8) та (2.9) називаються *формулами добутку ймовірностей незалежних (незалежних у сукупності) подій*.

Добуток ймовірностей залежних подій обчислюється за формулою (2.6) та (2.7).

Приклад 2.3. Розглянемо два незалежні кидки симетричних монет. Нехай події A і B означають випадання «орла» на першій та другій монетах відповідно. Перевірити, чи є події A і B незалежними.

Розв'язання. Ймовірності подій:

$$P(A) = P(B) = \frac{1}{2}.$$

Ймовірність одночасного настання подій A і B :

$$\begin{aligned} P(A \cap B) &= P(\text{орел на першій монеті і орел на другій монеті}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = \\ &= P(A) \cdot P(B). \end{aligned}$$

Оскільки $P(A \cap B) = P(A) \cdot P(B)$, події A і B не є незалежними.

Відповідь. Події A і B є незалежними.

Приклад 2.4. Розглянемо кидок грального кубика. Нехай події A , B , і C означають випадання чисел, кратних 2, 3 та 4 відповідно. Перевірити, чи є події A , B , і C незалежними в сукупності.

Розв'язання. Ймовірності подій:

$$P(A) = \frac{3}{6} = \frac{1}{2}, P(B) = \frac{2}{6} = \frac{1}{3}, P(C) = \frac{1}{6}.$$

Ймовірність одночасного настання подій A , B , і C ($A \cap B \cap C = \{\text{число, кратне } 12\}$):

$$P(A \cap B \cap C) = 0 \neq P(A) \cdot P(B) \cdot P(C).$$

У цьому випадку події A , B , і C не є незалежними у сукупності, оскільки їх одночасне настання неможливе.

Відповідь. Події A , B , і C не є незалежними у сукупності.

Приклад 2.5. Кидання двох кубиків. Нехай подія A означає, що сума очок на обох кубиках дорівнює 7, а подія B означає, що на першому кубіку випало число 4. Чи є події A і B незалежними?

Розв'язання. Ймовірність події A :

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

Ймовірність події B :

$$P(B) = \frac{1}{6}.$$

Ймовірність одночасного настання подій A і B (перший кубик 4, другий 3):

$$P(A \cap B) = \frac{1}{36} = P(A) \cdot P(B).$$

Оскільки $P(B/A) \neq P(B)$, події A і B не є незалежними.

Відповідь. Події A і B є незалежними.

Приклад 2.6. У коробці є 5 червоних і 5 синіх кульок. Виймаємо послідовно дві кульки без повернення. Нехай подія A означає, що перша кулька червона, а подія B означає, що друга кулька червона. Чи є події A і B незалежними?

Розв'язання. Ймовірність події A :

$$P(A) = \frac{5}{10} = \frac{1}{2}.$$

Ймовірність події B за умови A (перша кулька червона, залишилося 4 червоні і 5 синіх):

$$P(B/A) = \frac{4}{9}.$$

Ймовірність події B :

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

$$P(B) = \frac{5}{10} \cdot \frac{4}{9} + \frac{5}{10} \cdot \frac{5}{9} = \frac{20}{90} + \frac{25}{90} = \frac{45}{90} = \frac{1}{2}.$$

Оскільки $P(B/A) \neq P(B)$, події A і B не є незалежними.

Відповідь. Події A і B є незалежними.

2.5. Несумісні та сумісні події

Несумісні події — це події, які не можуть відбутися одночасно.

Іншими словами, якщо одна подія відбулася, то інша подія обов'язково не відбудеться. У математичних термінах це означає, що перетин таких подій є порожньою множиною.

Дві події A і B називаються **несумісними** (або **взаємовиключними**), якщо:

$$A \cap B = \emptyset.$$

Тобто ймовірність одночасного настання подій A і B дорівнює нулю (2.10):

$$P(A \cap B) = 0. \quad (2.10)$$

Приклади несумісних подій

1. Кидок грального кубика.

Розглянемо кидок симетричного грального кубика. Нехай подія A означає випадання числа 2, а подія B — випадання числа 5.

$$A = \{2\}, B = \{5\}.$$

Події A і B є несумісними, оскільки вони не можуть відбутися одночасно:

$$A \cap B = \emptyset.$$

2. Вибір з коробки з кульками.

Розглянемо коробку, що містить 3 червоні та 2 сині кульки. Нехай подія A означає витягування червоної кульки, а подія B означає витягування синьої кульки.

$$A = \{\text{червона кулька}\}, B = \{\text{синя кулька}\}.$$

Події A і B є несумісними, оскільки витягування кульки не може одночасно бути і червоним, і синім:

$$A \cap B = \emptyset.$$

Властивості несумісних подій

1. Сума ймовірностей: Якщо події A і B несумісні, то ймовірність об'єднання подій A і B дорівнює сумі їх індивідуальних ймовірностей (2.12).

2. Несумісність трьох і більше подій: Події $A_1, A_2, \dots, A_n$ є несумісними, якщо жодна пара цих подій не може відбутися одночасно. Тобто для будь-яких $i \neq j$:

$$A_i \cap A_j = \emptyset.$$

Несумісні події відіграють важливу роль у теорії ймовірностей, особливо при моделюванні ситуацій, де одночасне настання певних подій неможливе. Вони допомагають краще зрозуміти структуру ймовірнісних експериментів та коректно оцінювати ймовірності різних комбінацій подій.

Якщо події можуть відбуватися одночасно, то вони називаються **сумісними**.

У математичних термінах це означає, що перетин таких подій не є порожньою множиною.

Дві події A і B називаються **сумісними**, якщо $A \cap B \neq \emptyset$. Тобто для сумісних подій справедлива нерівність (2.11):

$$P(A \cap B) > 0. \quad (2.11)$$

Властивості сумісних подій

1. Сума ймовірностей. Для двох сумісних подій A і B ймовірність об'єднання подій обчислюється за формулою включення-виключення:

2. Ймовірність перетину. Сумісні події мають ненульову ймовірність перетину:

Приклади сумісних подій

1. Кидок грального кубика.

Розглянемо кидок симетричного грального кубика. Нехай подія A означає випадання парного числа, а подія B означає випадання числа більше трьох.

$$A = \{2, 4, 6\}, B = \{4, 5, 6\}.$$

Перетин подій A і B :

$$A \cap B = \{4, 6\}.$$

Ймовірність одночасного настання подій A і B :

$$P(A \cap B) = P(\{4, 6\}) = \frac{2}{6} = \frac{1}{3}.$$

Оскільки $P(A \cap B) > 0$, події A і B є сумісними.

2. Вибір з коробки з кульками.

Розглянемо коробку, що містить 3 червоні, 2 сині та 1 зелену кульку. Витягуємо одну кульку. Нехай подія A означає витягування червоної кульки, а подія B означає витягування не синьої кульки.

$$A = \{\text{червона кулька}\}, B = \{\text{не синя кулька}\} = \{\text{червона кулька, зелена кулька}\}.$$

Перетин подій A і B :

$$A \cap B = \{\text{червона кулька}\}.$$

Ймовірність одночасного настання подій A і B :

$$P(A \cap B) = P(\{\text{червона кулька}\}) = \frac{3}{6} = \frac{1}{2}.$$

Оскільки $P(A \cap B) > 0$, події A і B є сумісними.

Сумісні події можуть відбутися одночасно, і їх аналіз допомагає зрозуміти ймовірнісні зв'язки між різними випадковими подіями. Властивості та приклади сумісних подій є основою для побудови більш складних ймовірнісних моделей та розрахунків.

2.6. Формула додавання ймовірностей

Ймовірність суми несумісних подій дорівнює сумі ймовірностей доданків (2.12):

$$P(A \cup B) = P(A) + P(B) \quad (2.12)$$

або у випадку k несумісних подій $A_1, A_2, \dots, A_k$ (2.13):

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_k). \quad (2.13)$$

Формули (2.12) та (2.13) називаються *формулами додавання ймовірностей несумісних подій*.

Якщо події $A_1, A_2, \dots, A_k$ у даному випробуванні утворюють повну групу несумісних подій, то сума їх ймовірностей дорівнює 1 (2.14):

$$P(A_1) + P(A_2) + \dots + P(A_k) = 1, \quad (2.14)$$

оскільки сума, які утворюють повну групу, є подією достовірною.

Зокрема, для протилежних подій A та $\bar{A}$ маємо (2.15):

$$P(A) + P(\bar{A}) = 1. \quad (2.14)$$

Проте, коли доданки сумісні, згадане співвідношення, взагалі кажучи, не має місця.

Теорема 2.3.

Розглянемо дві події A і B , які є сумісними. Тоді справедлива рівність (2.15)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (2.15)$$

Доведення. Представимо події A і B за аналогією з множинами за допомогою кіл Ейлера (рис. 2.5).

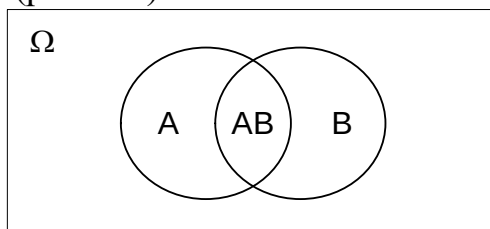


Рис. 2.5. Представлення подій A і B

Зауважимо, що $A \cup B$ можна зобразити як суму трьох несумісних доданків

$$A \cup B = (A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B)),$$

та $(A \cap B) \subseteq A$, $(A \cap B) \subseteq B$.

Звідси

$$\begin{aligned} P(A \cup B) &= P(A \setminus (A \cap B)) + P(A \cap B) + P(B \setminus (A \cap B)) = \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) = \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

Використовуючи метод математичної індукції, самостійно доведіть формулу додавання для довільної кількості доданків (2.16):

$$P\left(\bigcup_{i=1}^k A_i\right) = \sum_{i=1}^k P(A_i) - \sum_{i=1}^{k-1} \left(\sum_{j=i+1}^k P(A_i \cap A_j)\right) + \sum_{i=1}^{k-2} \left(\sum_{j=i+1}^{k-1} \left(\sum_{l=j+1}^k P(A_i \cap A_j \cap A_l)\right)\right) + \dots + (-1)^{k-1} \cdot P(A_1 \cap A_2 \cap \dots \cap A_k). \quad (2.16)$$

Формули (2.15) та (2.16) називається **формулами додавання ймовірностей сумісних подій**.

Приклад 2.7. З колоди в 32 карти навмання витягується одна карта. Треба обчислити ймовірність того, що це буде дама або піка.

Розв'язання. Вводимо події:

A =(витягнути даму),

B =(витягнути піку).

Нас цікавить ймовірність суми $A \cup B$. За формулою додавання ймовірностей маємо:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{4}{32} + \frac{8}{32} - \frac{1}{32} = \frac{11}{32}.$$

Відповідь. $P(A \cup B) = \frac{11}{32}$

Приклад 2.8. Кидання двох кубиків. Нехай подія A означає, що сума очок на обох кубиках дорівнює 7, а подія B означає, що на першому кубіку випало число 4. Обчислити $P(A \cup B)$.

Розв'язання. Ймовірність події A :

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

Ймовірність події B :

$$P(B) = \frac{1}{6}.$$

Ймовірність події $A \cap B$:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{36}.$$

Оскільки події є сумісними, ймовірність настання подій $A \cup B$ (перший кубик 4 або другий 3):

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

Відповідь. $P(A \cup B) = \frac{11}{36}$

Приклад 2.9. У коробці є 5 червоних і 5 синіх кульок. Виймаємо послідовно дві кульки без повернення. Нехай подія A означає, що перша кулька червона, а подія B означає, що друга кулька червона. Обчислити $P(A \cup B)$.

Розв'язання. Ймовірність події A :

$$P(A) = \frac{5}{10} = \frac{1}{2}.$$

Ймовірність події B (при цьому перша кулька червона, залишилося 4 червоні і 5 синіх або перша кулька синя, залишилося 5 червоні і 4 синіх).

$$P(B) = \frac{4}{9} \cdot \frac{5}{10} + \frac{5}{9} \cdot \frac{5}{10} = \frac{1}{2}.$$

Ймовірність події B за умови A (перша кулька червона, залишилося 4 червоні і 5 синіх):

$$P(B/A) = \frac{4}{9}.$$

Ймовірність події $A \cap B$:

$$P(A \cap B) = P(A) \cdot P(B/A) = \frac{5}{10} \cdot \frac{4}{9} = \frac{2}{9}.$$

Ймовірність події $A \cup B$:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{2} - \frac{2}{9} = \frac{7}{9}.$$

Дану задачу можна також розв'язати наступним чином. Розглянемо три випадки:

- $A \cap \bar{B}$ — 1-а кулька червона, друга синя;
- $\bar{A} \cap B$ — 1-а кулька синя, друга червона;
- $A \cap B$ — обидві кульки червоні.

$$P(A \cup B) = P(A \cap \bar{B}) + P(\bar{A} \cap B) + P(A \cap B) = \frac{5}{10} \cdot \frac{5}{9} + \frac{5}{10} \cdot \frac{5}{9} + \frac{5}{10} \cdot \frac{4}{9} = \frac{7}{9}.$$

Відповідь. $P(A \cup B) = \frac{7}{9}$.

Приклад 2.10. При підготовці до заліку перший студент вивчив 25 з 30 питань програми; другий студент вивчив 20 з 30 питань, а третій – 15 з 30. Для здачі заліку необхідно відповісти на 3 довільні питання програми. Знайти ймовірність того, що лише один студент здасть залік.

Розв’язання. Щоб знайти ймовірність того, що лише один студент здасть залік, спочатку знайдемо ймовірності того, що кожен студент складе залік, а потім обчислимо ймовірності для кожного сценарію, коли тільки один студент здав.

Позначимо ймовірність того, що студент складе залік як $P_i(A)$ (де $i = 1, 2, 3$).

Ймовірності здачі заліку кожним студентом окремо

Перший студент

Перший студент вивчив 25 з 30 питань. Ймовірність того, що він відповість на 3 довільні питання правильно:

$$P_1(A) = \frac{C_{25}^3}{C_{30}^3} = \frac{25 \cdot 24 \cdot 23}{30 \cdot 29 \cdot 28} \approx 0,548.$$

Другий студент

Другий студент вивчив 20 з 30 питань. Ймовірність того, що він відповість на 3 довільні питання правильно:

$$P_2(A) = \frac{C_{20}^3}{C_{30}^3} = \frac{20 \cdot 19 \cdot 18}{30 \cdot 29 \cdot 28} \approx 0,299.$$

Третій студент

Третій студент вивчив 15 з 30 питань. Ймовірність того, що він відповість на 3 довільні питання правильно:

$$P_3(A) = \frac{C_{15}^3}{C_{30}^3} = \frac{15 \cdot 14 \cdot 13}{30 \cdot 29 \cdot 28} \approx 0,108.$$

Ймовірності нездачі заліку кожним студентом

Позначимо ймовірність того, що студент не складе залік як $Q_i(A)$ (де $i = 1, 2, 3$).

$$Q_i(A) = 1 - P_i(A).$$

$$Q_1(A) = 0,452; Q_2(A) = 0,701; Q_3(A) = 0,892.$$

Ймовірність того, що лише один студент складе залік

Для цього сценарію ми маємо три можливі випадки:

- Перший студент складає залік, а другий і третій не складають;
- Другий студент складає залік, а перший і третій не складають;
- Третій студент складає залік, а перший і другий не складають.

Перший випадок

Перший студент складає залік, а другий і третій не складають:

$$P(\text{лише 1-й здав}) = P_1(A) \cdot Q_2(A) \cdot Q_3(A) \approx 0,342.$$

Другий випадок

Другий студент складає залік, а перший і третій не складають:

$$P(\text{лише 2-й здав}) = Q_1(A) \cdot P_2(A) \cdot Q_3(A) \approx 0,121.$$

Третій випадок

Третій студент складає залік, а перший і другий не складають:

$$P(\text{лише 3-й здав}) = Q_1(A) \cdot Q_2(A) \cdot P_3(A) \approx 0,034.$$

Загальна ймовірність

Загальна ймовірність того, що лише один студент складе залік:

$$P(\text{лише один здав}) = P_1(A) \cdot Q_2(A) \cdot Q_3(A) + Q_1(A) \cdot P_2(A) \cdot Q_3(A) + Q_1(A) \cdot Q_2(A) \cdot P_3(A) = 0,497.$$

Відповідь. $P(\text{лише один здав}) = 0,497.$

Слід зазначити, що формули (2.15) та (2.16) також впливають з формули включення-виключення.

Формула включення-виключення

Формула включення-виключення є важливим інструментом в комбінаториці та теорії ймовірностей для обчислення кількості елементів у об'єднанні кількох множин або ймовірності об'єднання кількох подій. Вона враховує як перетини множин (подій), так і перекривання цих перетинів, щоб отримати точну кількість або ймовірність.

Формула включення-виключення для двох множин

Для трьох множин A і B кількість елементів в об'єднанні $A \cup B$ визначається як (2.17):

$$|A \cup B| = |A| + |B| - |A \cap B| \quad (2.17)$$

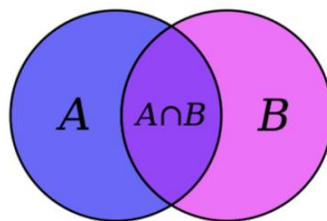


Рис. 2.6. Формула включення-виключення. Випадок двох множин

Формула включення-виключення для трьох множин

Для трьох множин A , B і C кількість елементів в об'єднанні $A \cup B \cup C$ визначається як (2.18):

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C| \quad (2.18)$$

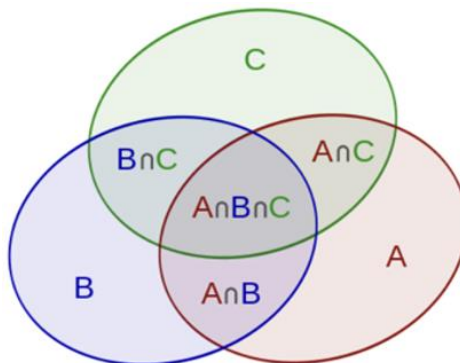


Рис. 2.7. Формула включення-виключення. Випадок двох множин

Загальна формула включення-виключення

Для n подій $A_1, A_2, \dots, A_n$ ймовірність їх об'єднання визначається як (2.19):

$$\bigcup_{i=1}^n A_i = \sum_{i=1}^n A_i - \sum_{1 \leq i < j \leq n} (A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} (A_i \cap A_j \cap A_k) + \dots + (-1)^{n+1} (A_1 \cap A_2 \cap \dots \cap A_n) \quad (2.19)$$

Приклад 2.11. Нехай є три множини A, B і C : $|A| = 10, |B| = 15, |C| = 20, |A \cap B| = 5, |B \cap C| = 7, |A \cap C| = 8, |A \cap B \cap C| = 3$. Чому дорівнює кількість елементів у множині $A \cup B \cup C$?

Розв'язання. Кількість елементів в об'єднанні $A \cup B \cup C$ дорівнює:

$$|A \cup B \cup C| = 10 + 15 + 20 - 5 - 7 - 8 + 3 = 28.$$

Відповідь. $|A \cup B \cup C| = 28$.

ЛЕКЦІЯ 3

ПРЕДСТАВЛЕННЯ ПОДІЙ ЗА ДОПОМОГОЮ ЕЛЕМЕНТАРНИХ ПОДІЙ. ЙМОВІРНІСТЬ ПОЯВИ ХОЧА Б ОДНІЄЇ ПОДІЇ. ОЦІНЮВАННЯ НАДІЙНОСТІ РОБОТИ ПРОСТИХ СХЕМ ЗА ДОПОМОГОЮ ФОРМУЛ ТЕОРІЇ ЙМОВІРНОСТЕЙ. ФОРМУЛА ПОВНОЇ ЙМОВІРНОСТІ. ФОРМУЛА БАЄСА

3.1. Представлення події за допомогою елементарних подій

Представлення події за допомогою елементарних подій є одним з основних концептів у теорії ймовірностей. *Елементарні події* є найпростішими можливими результатами експерименту, що не можуть бути розбиті на більш прості результати. Вони утворюють повний простір можливих результатів експерименту.

Основні поняття

Елементарна подія (ω): Це найпростіший можливий результат експерименту. Наприклад, у підкиданні монети елементарні події можуть бути «герб» і «решітка».

Простір елементарних подій (Ω): Це множина всіх можливих елементарних подій. Наприклад, для підкидання монети $\Omega = \{\text{Герб}, \text{Решітка}\}$.

Подія (A): Це будь-яка підмножина простору елементарних подій. Подія може складатися з однієї або кількох елементарних подій. Наприклад, подія A у підкиданні монети може бути «випаде герб», тобто $A = \{\text{Герб}\}$.

Приклад

Розглянемо приклад підкидання грального кубика. Простір елементарних подій буде:

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

Події можуть бути представлені як підмножини цього простору. Наприклад:

Подія A : Випаде парне число. Тоді $A = \{2, 4, 6\}$.

Подія B : Випаде число більше 4. Тоді $B = \{5, 6\}$.

Представлення подій

Будь-яка подія може бути представлена за допомогою операцій представлених в лекції 1, таких як об'єднання, перетин або різниця елементарних подій.

Об'єднання подій

Подія $A \cup B$ ($A + B$) означає, що відбудеться або подія A , або подія B , або обидві події одночасно. Наприклад, для подій $A = \{2, 4, 6\}$ і $B = \{5, 6\}$:

$$A + B = \{2, 4, 5, 6\}.$$

Перетин подій

Подія $A \cap B$ ($A \cdot B$) означає, що відбудеться як подія A , так і подія B .
Наприклад, для подій $A = \{2,4,6\}$ і $B = \{5,6\}$:

$$A \cdot B = \{6\}.$$

Різниця подій

Подія A/B ($A - B$) означає, що відбудеться подія A , але не подія B .
Наприклад, для подій $A = \{2,4,6\}$ і $B = \{5,6\}$:

$$A - B = \{2,4\}.$$

Представлення подій за допомогою елементарних подій дозволяє зрозуміти структуру і взаємозв'язки між подіями. Це фундаментальне поняття в теорії ймовірностей, яке лежить в основі багатьох інших концептів і методів аналізу.

Таким чином, якщо задана подія A та множина незалежних елементарних подій $\{A_1; A_2; \dots; A_n\}$, без настання кожної з яких неможливе настання події A , то її можна представити за допомогою операцій над даними елементарними подіями, таким чином, щоб утворені в результаті більш складні події були несумісними. Найчастіше використовуються операції перетину, об'єднання та доповнення.

Подібне представлення необхідно для можливості використання формул додавання ймовірностей несумісних подій та формул множення незалежних (незалежних у сукупності) подій, що значно спростить процес обчислення ймовірності події A .

Приклад 3.1. На модульній контрольній роботі студентам запропоновано 9 задач: 4 – першого рівня складності, 3 – другого і дві – третього рівня. Подія $A_i = \{\text{розв'язання } i\text{-ї задачі першого рівня, } i = 1,2,3,4\}$, $B_j = \{\text{розв'язання } j\text{-ї задачі першого рівня, } j = 1,2,3\}$, $C_k = \{\text{розв'язання } k\text{-ї задачі першого рівня, } k = 1,2\}$, $D = \{\text{одержання позитивної оцінки}\}$. Виразити подію D через події A_i , B_j , C_k , якщо для отримання позитивної оцінки необхідно розв'язати всі задачі першого рівня, принаймні дві другого рівня і принаймні одну третього рівня.

Розв'язання. Складемо допоміжні події

$$A = A_1 \cdot A_2 \cdot A_3 \cdot A_4,$$

де $A_i, i = 1,2,3,4$ – незалежні події.

$$B = B_1 \cdot B_2 \cdot \overline{B_3} + B_1 \cdot \overline{B_2} \cdot B_3 + \overline{B_1} \cdot B_2 \cdot B_3 + B_1 \cdot B_2 \cdot B_3,$$

де:

$B_j, j = 1,2,3$ – незалежні події;

$B_1 \cdot B_2 \cdot \overline{B_3}; B_1 \cdot \overline{B_2} \cdot B_3; \overline{B_1} \cdot B_2 \cdot B_3; B_1 \cdot B_2 \cdot B_3$ – несумісні події.

$$C = C_1 \cdot \overline{C_2} + \overline{C_1} \cdot C_2 + C_1 \cdot C_2,$$

де:

$C_k, k = 1,2$ – незалежні події;

$C_1 \cdot \overline{C_2}; \overline{C_1} \cdot C_2; C_1 \cdot C_2$ – несумісні події.

Кожен доданок в подіях B та C є несумісними подіями.

Отже, $D = A \cdot B \cdot C$.

Відповідь. $D = A \cdot B \cdot C$.

3.2. Ймовірність появи хоча б однієї події

Нехай проводиться n незалежних випробувань, у кожному з яких може відбутися подія A_i ($i = 1, 2, 3, \dots, n$) з ймовірністю $P(A_i) = p_i$ або протилежна подія $\bar{A}_i$ ($A_i \cup \bar{A}_i = \Omega, A_i \cap \bar{A}_i = \emptyset$) з ймовірністю $P(\bar{A}_i) = q_i$, ($p_i + q_i = 1$).

Ймовірність появи принаймні однієї з подій $A_1, A_2, \dots, A_n$, незалежних в сукупності, знаходиться за формулою (3.1):

$$P(A) = 1 - q_1 \cdot q_2 \cdot \dots \cdot q_n. \quad (3.1)$$

Якщо події $A_1, A_2, \dots, A_n$ мають однакову ймовірність p , то ймовірність настання принаймні однієї з цих подій дорівнює:

Тоді справедлива формула (3.2):

$$P(A) = 1 - q^n. \quad (3.2)$$

Приклад 3.2. Продано 10000 білетів лотереї. На один білет випадає виграш у 500 грн., на 100 білетів – виграші у 100 грн., на 200 білетів – виграші у 20 грн., на 300 білетів – виграші у 5 грн., інші білети – невикрашені. Дехто купив один білет. Знайти ймовірність того, що він виграв не менше 20 грн.

Розв’язання. Розглянемо подію $A = \{\text{виграти не менше 20 грн.}\}$. Це означає, що виграти можна або 20 грн., або 100 грн., або 500 грн. Введемо події:

$A_1 = \{\text{виграти 20 грн.}\}$, $A_2 = \{\text{виграти 100 грн.}\}$, $A_3 = \{\text{виграти 500 грн.}\}$.

Події A_1, A_2, A_3 – несумісні, тобто в наслідок експерименту може настати лише одна з них.

Тоді подію A можна представити у вигляді: $A = A_1 + A_2 + A_3$.

За формулою про додавання несумісних ймовірностей:

$$\begin{aligned} P(A) &= P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) = \\ &= \frac{200}{10000} + \frac{100}{10000} + \frac{1}{10000} = 0,0301. \end{aligned}$$

Або ж:

$$P(A) = 1 - \left(\frac{300}{10000} + \frac{9399}{10000} \right) = 0,0301.$$

Відповідь. $P(A) = 0,0301$.

Приклад 3.3. Кругова мішень має три зони: I, II, III. Ймовірність влучення у першу зону одним пострілом дорівнює 0,3; у другу – 0,5; у третю – 0,1. Знайти ймовірність промаху.

Розв’язання. Розглянемо події:

$A_i = \{\text{промах по } i\text{-ій зоні}\}$, $\bar{A}_i = \{\text{влучання в } i\text{-у зону}\}$, $i = 1, 2, 3$.

Тоді $\bar{A} = \bar{A}_1 + \bar{A}_2 + \bar{A}_3$, де $\bar{A}_1, \bar{A}_2, \bar{A}_3$ – влучення відповідно у першу, другу і третю зони. Ці події несумісні – поява однієї з них виключає появу інших, бо виконується тільки один постріл.

За формулою про додавання ймовірностей несумісних подій маємо

$$P(\bar{A}) = P(\bar{A}_1 + \bar{A}_2 + \bar{A}_3) = P(\bar{A}_1) + P(\bar{A}_2) + P(\bar{A}_3) = 0,3 + 0,5 + 0,1 = 0,9.$$

Подія A протилежна події $\bar{A}$. Тому $P(A) + P(\bar{A}) = 1$.

Таким чином $P(A) = 1 - 0,9 = 0,1$.

Відповідь. $P(A) = 0,1$.

Приклад 3.4. Виконуються три постріли у одну і ту ж мішень. Ймовірність влучення у мішень першим пострілом дорівнює 0,5; другим – 0,7; третім – 0,8. Знайти ймовірність того, що в результаті трьох пострілів у мішені буде рівно одна пробоїна.

Розв’язання. Розглянемо подію $A = \{\text{рівно одне влучення у мішень}\}$. Ця подія може здійснитись декількома способами, тобто вона розпадається на три несумісні варіанти: може бути влучення першим пострілом, промахи другим і третім; або влучення другим пострілом, промахи першим і третім; або влучення третім пострілом, промахи першим і другим.

Введемо події: $A_1; A_2; A_3 = \{\text{влучення у мішень першим, другим і третім пострілом відповідно}\}$, $\bar{A}_1; \bar{A}_2; \bar{A}_3 = \{\text{промах першим, другим і третім пострілами}\}$.

Тоді $A = A_1 \cdot \bar{A}_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot A_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot \bar{A}_2 \cdot A_3$.

Події $A_1; A_2; A_3; \bar{A}_1; \bar{A}_2; \bar{A}_3$ – незалежні, а події $A_1 \cdot \bar{A}_2 \cdot \bar{A}_3; \bar{A}_1 \cdot A_2 \cdot \bar{A}_3; \bar{A}_1 \cdot \bar{A}_2 \cdot A_3$ – несумісні.

За формулами про додавання і добуток ймовірностей

$$\begin{aligned} P(A) &= P(A_1 \cdot \bar{A}_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot A_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot \bar{A}_2 \cdot A_3) = \\ &= P(A_1 \cdot \bar{A}_2 \cdot \bar{A}_3) + P(\bar{A}_1 \cdot A_2 \cdot \bar{A}_3) + P(\bar{A}_1 \cdot \bar{A}_2 \cdot A_3) = \\ &= P(A_1) \cdot P(\bar{A}_2) \cdot P(\bar{A}_3) + P(\bar{A}_1) \cdot P(A_2) \cdot P(\bar{A}_3) + P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot P(A_3) \cdot \\ &\cdot P(A_3) = 0,5 \cdot 0,3 \cdot 0,2 + 0,5 \cdot 0,7 \cdot 0,2 + 0,5 \cdot 0,3 \cdot 0,8 = 0,22. \end{aligned}$$

Відповідь. $P(A) = 0,22$.

Приклад 3.5. За умов попередньої задачі знайти ймовірність того, що у мішені буде хоча б одна пробоїна.

Розв’язання. Введемо подію $B = \{\text{хоча б одне влучення у мішень}\}$. Ця подія означає, що у мішені може бути або одна пробоїна, або дві пробоїни, або три пробоїни. Тому подію B можна записати:

$$B = A_1 \cdot \bar{A}_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot A_2 \cdot \bar{A}_3 + \bar{A}_1 \cdot \bar{A}_2 \cdot A_3 + A_1 \cdot A_2 \cdot \bar{A}_3 + A_1 \cdot \bar{A}_2 \cdot A_3 + \bar{A}_1 \cdot A_2 \cdot A_3 + A_1 \cdot A_2 \cdot A_3.$$

Тепер, використовуючи формул про додавання і добуток ймовірностей, можна знайти ймовірність події B .

Але цю задачу простіше розв’язати, якщо перейти від прямої події B до протилежної події $\bar{B}$ – ні одного влучення.

Очевидно, що $\bar{B} = \bar{A}_1 \cdot \bar{A}_2 \cdot \bar{A}_3$. Події $\bar{A}_1; \bar{A}_2; \bar{A}_3$ – незалежні.

За формулою про добуток ймовірностей трьох незалежних подій:

$$P(\bar{B}) = 0,5 \cdot 0,3 \cdot 0,2 = 0,03.$$

Звідси $P(B) = 1 - P(\bar{B}) = 1 - 0,03 = 0,97$.

Відповідь. $P(A) = 0,97$.

Приклад 3.6. Гральний кубик підкидається чотири рази. Чому дорівнює ймовірність того, що цифра 3 з’явиться при цьому хоча б один раз?

Розв'язання. Ймовірність того, що при одному підкиданні з'явиться цифра 3, дорівнює $\frac{1}{6}$.

Тоді $q = 1 - p = 1 - \frac{1}{6} = \frac{5}{6}$. Згідно з (3.2) дістанемо:

$$P(A) = 1 - q^4 = 1 - \left(\frac{5}{6}\right)^4 = 1 - \frac{625}{1296} = \frac{671}{1296}.$$

Відповідь. $P(A) = \frac{671}{1296}$.

3.3. Оцінювання надійності роботи простих схем за допомогою формул теорії ймовірностей

Оцінювання надійності роботи простих схем за допомогою формул теорії ймовірностей є важливою задачею, особливо в інженерії та комп'ютерних науках. Умовна ймовірність, незалежні події, формула включення-виключення та інші концепції теорії ймовірностей допомагають аналізувати і прогнозувати роботу схем.

Основні поняття

Надійність (R) – ймовірність того, що система або компонент буде функціонувати безвідмовно протягом заданого часу.

Відмова (F) – ймовірність того, що система або компонент відмовлять протягом заданого часу.

Зв'язок між надійністю та відмовою обчислюється за формулою (3.3):

$$F = 1 - R. \quad (3.3)$$

Оцінювання надійності послідовних схем

У *послідовній схемі* всі компоненти підключені один за одним. Відмова будь-якого компонента призводить до відмови всієї системи.

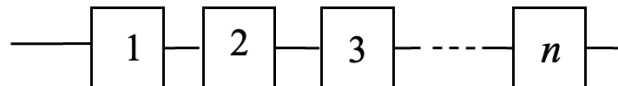


Рис. 3.1. Схема з послідовним з'єднанням

Якщо елементи системи з'єднані послідовно (рис. 3.1), і при цьому відомі ймовірності безвідмовної роботи кожного елемента p_i ($i = 1, \dots, n$), то надійність системи R , дістанемо з формули (3.4):

$$R = \prod_{i=1}^n p_i. \quad (3.4)$$

а відмову F – з формули (3.5):

$$F = 1 - \prod_{i=1}^n p_i. \quad (3.5)$$

Оцінювання надійності паралельних схем

У *паралельній схемі* всі компоненти підключені паралельно. Відмова всієї системи станеться тільки тоді, коли всі компоненти відмовлять.

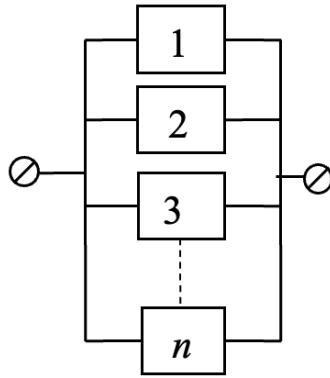


Рис. 3.2. Схема з паралельним з'єднанням

Якщо елементи системи з'єднані за схемою, наведеною на рис. 3.2, і при цьому відомі ймовірності безвідмовної роботи кожного елемента p_i ($i = 1, \dots, n$), то надійність роботи системи знайдемо за формулою (3.6):

$$R = 1 - \prod_{i=1}^n q_i, q_i = 1 - p_i, \quad (3.6)$$

а відмову F – з формули (3.7):

$$F = \prod_{i=1}^n q_i, q_i = 1 - p_i. \quad (3.7)$$

Приклад 3.7. Ділянку електричної схеми наведено на рис. 3.3. Події $A_i = \{\text{відмова елемента } a_i (i = 1, 2)\}$; $B = \{\text{відмова елемента } b\}$; $C_j = \{\text{відмова елемента } c_j (j = 1, 2)\}$; $D = \{\text{ділянка проводить струм}\}$. Виразити D та $\bar{D}$ через A_i, B, C_j .

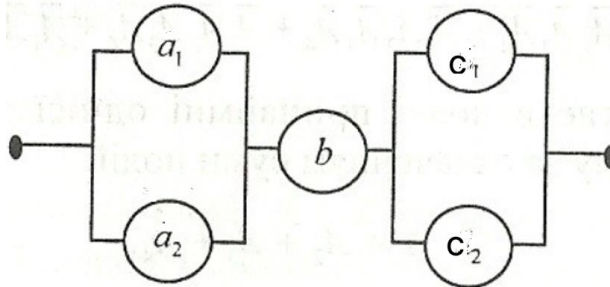


Рис. 3.3. Схема для прикладу

Розв'язання. Подія D відбудеться, коли не відмовить принаймні один з елементів a_i , елемент b і принаймні один з елементів c_j . Тобто

$$D = (A_1 \cdot \bar{A}_2 + A_2 \cdot \bar{A}_1 + \bar{A}_1 \cdot \bar{A}_2) \cdot \bar{B} \cdot (C_1 \cdot \bar{C}_2 + \bar{C}_1 \cdot C_2 + \bar{C}_1 \cdot \bar{C}_2),$$

$$\bar{D} = (A_1 \cdot A_2) \cdot \bar{B} \cdot (C_1 \cdot \bar{C}_2 + \bar{C}_1 \cdot C_2 + \bar{C}_1 \cdot \bar{C}_2) + (A_1 \cdot \bar{A}_2 + A_2 \cdot \bar{A}_1 + A_1 \cdot A_2) \cdot B \cdot (C_1 \cdot \bar{C}_2 + \bar{C}_1 \cdot C_2 + \bar{C}_1 \cdot \bar{C}_2) + (A_1 \cdot \bar{A}_2 + A_2 \cdot \bar{A}_1 + \bar{A}_1 \cdot \bar{A}_2) \cdot \bar{B} \cdot (C_1 \cdot C_2).$$

Відповідь. Представлено події D та $\bar{D}$.

Приклад 3.8. Електричні елементи з'єднані за схемами, наведеними на рис. 3.4 і 3.5.

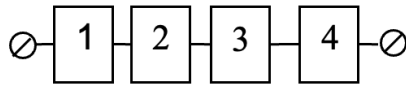


Рис. 3.4. Схема для прикладу (а)

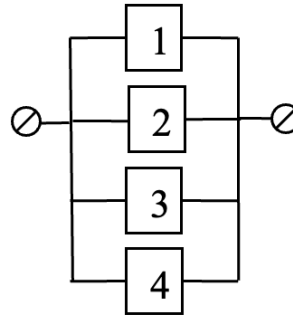


Рис. 3.5. Схема для прикладу (б)

Ймовірність того, що електричний елемент буде працювати при ввімкненні в електромережу наведених схем, є величиною сталою і дорівнює $p_i = 0,85$. Яка ймовірність того, що при ввімкненні в електромережу наведених схем у них буде електрострум?

Розв'язання. За відомим значенням p_i знаходимо

$$q_i = 1 - p_i = 1 - 0,85 = 0,15 \quad (i = 1, 2, 3, 4).$$

а) $R = \prod_{i=1}^4 p_i = (0,85)^4 = 0,522;$

б) $R = 1 - \prod_{i=1}^4 q_i = 1 - (0,15)^4 = 1 - 0,00051 = 0,99949.$

Відповідь. а) $R = 0,522$; б) $R = 0,99949$.

Приклад 3.9. Дано послідовну схему з трьома компонентами, надійності яких $R_1 = 0,9$; $R_2 = 0,8$; $R_3 = 0,95$. Визначити надійність всієї схеми.

Розв'язання. Надійність всієї послідовної схеми:

$$R = R_1 \cdot R_2 \cdot R_3 = 0,9 \cdot 0,8 \cdot 0,95 = 0,684.$$

Відповідь. $R = 0,684$.

Приклад 3.10. Дано паралельну схему з трьома компонентами, надійності яких $R_1 = 0,9$; $R_2 = 0,8$; $R_3 = 0,95$. Визначити надійність всієї схеми.

Розв'язання. Ймовірність відмови кожного компонента:

$$F_1 = 1 - R_1 = 0,1; F_2 = 1 - R_2 = 0,2; F_3 = 1 - R_3 = 0,05.$$

Ймовірність відмови всієї паралельної схеми:

$$F = F_1 \cdot F_2 \cdot F_3 = 0,1 \cdot 0,2 \cdot 0,05 = 0,001$$

Надійність всієї паралельної схеми:

$$R = 1 - F = 1 - 0,001 = 0,999.$$

Для визначення надійності електричних схем також можна використовувати формулу включення-виключення.

$$\begin{aligned} R &= R_1 + R_2 + R_3 - R_1 \cdot R_2 - R_1 \cdot R_3 - R_2 \cdot R_3 + R_1 \cdot R_2 \cdot R_3 = \\ &= 0,9 + 0,8 + 0,95 - 0,9 \cdot 0,8 - 0,9 \cdot 0,95 - 0,8 \cdot 0,95 + 0,9 \cdot 0,8 \cdot 0,95 = 2,65 - 2,335 + 0,684 = 0,999. \end{aligned}$$

Відповідь. $R = 0,999$.

3.4. Формула повної ймовірності

В практичній діяльності, як правило, зустрічаються події, що відбуваються не самі по собі в чистому вигляді, найчастіше вони відбуваються спільно з іншими подіями.

Нехай задано подію A та набір попарно несумісних подій $H_1, H_2, \dots, H_n$, кожна з яких може призвести до настання події A в рамках деякого експеримента. Таких, що $P(H_i) > 0$ для всіх i і $\bigcup_{i=1}^n H_i = \Omega$. Його називають *повною групою подій* або *розбиттям простору Ω* .

Події $H_1, H_2, \dots, H_n$ що утворюють повну групу подій, часто називають *гіпотезами*. При зручному виборі гіпотез для довільної події A можуть порівняно просто обчислюватися ймовірності $P(A/H_i)$ (умовна ймовірність події відбутися при виконанні «гіпотези» H_i) і $P(H_i)$ (ймовірність виконання «гіпотези» H_i). Як, використовуючи ці дані, обчислити ймовірність відбутися події A ?

Теорема 3.1. Нехай $H_1, H_2, \dots, H_n$ — повна група подій. Тоді ймовірність будь якої події A може бути обчислена за формулою (3.8):

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i). \quad (3.8)$$

Доведення. Розглянемо покрокове доведення

Початкове розбиття

Почнемо з розбиття простору Ω на взаємовиключні події H_i :

$$\Omega = H_1 \cup H_2 \cup \dots \cup H_n.$$

Об'єднання подій

Подію A можна представити як об'єднання її перетинів з подіями H_i :

$$A = (A \cap H_1) \cup (A \cap H_2) \cup \dots \cup (A \cap H_n).$$

Це справедливо, оскільки H_i утворюють повну систему подій.

Взаємовиключність

Оскільки події H_i взаємовиключні, то події $A \cap H_i$ також взаємовиключні. Тому ймовірність об'єднання подій $A \cap H_i$ дорівнює сумі їх ймовірностей:

$$P(A) = P((A \cap H_1) \cup (A \cap H_2) \cup \dots \cup (A \cap H_n)) = \sum_{i=1}^n P(A \cap H_i).$$

Умовна ймовірність

За визначенням умовної ймовірності

$$P(A/H_i) = \frac{P(A \cap H_i)}{P(H_i)},$$

тобто

$$P(A \cap H_i) = P(A/H_i) \cdot P(H_i).$$

Підстановка

Підставляємо $P(A \cap H_i)$ у суму:

$$P(A) = \sum_{i=1}^n P(A \cap H_i) = \sum_{i=1}^n P(A/H_i) \cdot P(H_i).$$

Що й треба було довести.

Формулу (3.8) називають *формулою повної ймовірності*.

Приклад 3.11. Маємо три урни. В першій урни є 6 білих і 2 чорні кульки, у другій – 3 білі і 4 чорні, в третій 4 білі і 3 чорні. Вибирає навмання будь-яку урну і виймаємо одну кульку. Знайти ймовірність того, що ця кулька біла.

Розв’язання. Розглянемо подію $A = \{\text{взято кульку білого кольору}\}$.

Сформулюємо гіпотези, що можуть призвести до настання події A в рамках проведеного експерименту:

$H_1 = \{\text{вибір першої урни}\}$; $H_2 = \{\text{вибір другої урни}\}$; $H_3 = \{\text{вибір третьої урни}\}$.

Гіпотези рівноможливі (з умов задачі), маємо:

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}.$$

Умовні ймовірності події A для цих гіпотез відповідно такі:

$$P(A/H_1) = \frac{6}{8} = \frac{3}{4}; \quad P(A/H_2) = \frac{3}{7}; \quad P(A/H_3) = \frac{4}{7}.$$

Таким чином:

$$P(A) = \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{3}{7} + \frac{1}{3} \cdot \frac{4}{7} = \frac{49}{84} \approx 0,583.$$

Відповідь. $P(A) \approx 0,583$.

3.5. Формула Баєса

Розглянемо подію A та сформульовані гіпотези H_i ($i = 1, 2, \dots, n$), які можуть призвести до настання даної події в рамках експерименту.

Формула Баєса використовується для обчислення умовної ймовірності настання гіпотези H_i за умови, що настання події A . Вона дозволяє знайти $P(H_i/A)$, знаючи $P(A/H_i)$.

Формула Баєса дуже корисна в різних областях, таких як статистика, машинне навчання та медицина.

Теорема 3.2. Для умовної ймовірності $P(A/H_i)$ справедлива рівність (3.9):

$$P(H_i/A) = \frac{P(A/H_i) \cdot P(H_i)}{P(A)}, \quad (3.9)$$

де:

$P(A/H_i)$ — ймовірність події A за умови, що відбулася подія H_i ;

$P(H_i/A)$ — ймовірність події H_i за умови, що відбулася подія A ;

$P(A)$ — апіорна ймовірність події A ;

$P(H_i)$ — ймовірність події H_i .

Доведення.

Визначення умовної ймовірності

За визначенням умовної ймовірності маємо:

$$P(A/H_i) = \frac{P(A \cap H_i)}{P(H_i)}$$

i

$$P(H_i/A) = \frac{P(A \cap H_i)}{P(A)}.$$

Перетин подій

Ймовірність перетину подій можна записати як:

$$P(A \cap H_i) = P(A/H_i) \cdot P(H_i).$$

Підставлення

Підставляючи $P(A \cap H_i)$ у формулу для $P(H_i/A)$:

$$P(H_i/A) = \frac{P(A/H_i) \cdot P(H_i)}{P(A)}.$$

Формула (3.9) *називається формулою Баєса*.

Формула Баєса є потужним інструментом для оновлення ймовірностей на основі нових даних. Вона широко використовується у багатьох галузях, таких як медицина, економіка, машинне навчання та багато інших.

Приклад 3.12. Три заводи виробляють одну й ту ж продукцію. При цьому 1-й завод виробляє 25%, 2-й завод — 35% і 3-й завод — 40% всієї виробленої продукції. Брак складає 5% від продукції 1-го заводу, 3% від продукції 2-го і 4% від продукції 3-го заводу. Вся продукція змішується і надходить до продажу. Знайти **а)** ймовірність придбати бракований виріб; **б)** умовну ймовірність того, що придбаний крам виготовлений 1-м заводом, якщо цей виріб бракований.

Розв’язання. Виріб вибирається навмання з усієї виробленої продукції. Розглянемо три гіпотези:

$$H_i = \{\text{виріб виготовлено на } i\text{-му заводі}\}, i = 1, 2, 3.$$

Ймовірності цих гіпотез дано

$$P(H_1) = 0,25; P(H_2) = 0,35; P(H_3) = 0,4.$$

Нехай $A = \{\text{виріб виявився бракованим}\}$.

Дано також умовні ймовірності

$$P(A/H_1) = 0,05; P(A/H_2) = 0,03; P(A/H_3) = 0,04.$$

а) За формулою повної ймовірності, отримуємо, що ймовірність події A дорівнює:

$$\begin{aligned} P(A) &= P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3) = \\ &= 0,25 \cdot 0,05 + 0,35 \cdot 0,03 + 0,4 \cdot 0,04 = 0,039. \end{aligned}$$

б) За формулою Баєса маємо:

$$P(H_1/A) = \frac{0,25 \cdot 0,05}{0,25 \cdot 0,05 + 0,35 \cdot 0,03 + 0,4 \cdot 0,04} \approx 0,32.$$

Відповідь. а) $P(A) = 0,039$; б) $P(H_1/A) \approx 0,32$.

Приклад 3.13. Два стрільці підкидають монетку і вибирають, хто з них стріляє по мішені (однією кулею). Перший стрілець влучає в мішень з ймовірністю 1, другий стрілець — з ймовірністю 0,00001. **а)** Яка ймовірність кулі влучити в мішень? **б)** Якщо куля влучила, то які ймовірності того що стріляв 1-й або 2-й стрілець?

Розв'язання. Сформулюємо дві гіпотези про експеримент:

$$H_1 = \{\text{стріляє 1-й стрілок}\}; H_2 = \{\text{стріляє 2-й стрілок}\}.$$

Апріорні (a'priori — «до дослід») ймовірності цих гіпотез однакові:

$$P(H_1) = P(H_2) = \frac{1}{2}.$$

Розглянемо подію $A = \{\text{куля потрапила в мішень}\}$. Відомо, що

$$P(A/H_1) = 1; P(A/H_2) = 0,00001.$$

а) Тому ймовірність кулі влучити в мішень

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0,00001 \approx 0,5.$$

б) Припустимо, що подія A відбулася. Яка тепер апостеріорна (a'posteriori — «після дослід») ймовірність кожної з гіпотез H_i ?

Зрозуміло, що перша з цих гіпотез набагато ймовірніша за другу (а саме, в 100000 раз). Дійсно,

$$P(H_1/A) = \frac{\frac{1}{2} \cdot 1}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0,00001} \approx 1; P(H_2/A) = \frac{\frac{1}{2} \cdot 0,00001}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0,00001} \approx 0.$$

Відповідь. а) $P(A) \approx 0,5$; б) $P(H_1/A) \approx 1$; $P(H_2/A) \approx 0$.

Приклад 3.14. Для пошуку космічного апарату, що приземлився, призначено 10 гелікоптерів, кожен з яких веде пошук в одному з двох районів, де апарат може знаходитися з ймовірностями 0,8 і 0,2. Як слід розподілити гелікоптери по районах пошуків, щоб ймовірність виявлення апарату була найбільшою, якщо кожен гелікоптер незалежно від інших знаходить апарат в районі пошуку з ймовірністю 0,3? Знайти ймовірність виявлення апарату при оптимальному розподілі гелікоптерів.

Розв'язання. Позначимо подію

$$A = \{\text{виявлення апарату}\}$$

і гіпотези

$$H_1 = \{\text{апарат знаходиться в першому районі}\};$$

$$H_2 = \{\text{апарат знаходиться в другому районі}\}.$$

За умовою задачі

$$P(H_1) = 0,8; P(H_2) = 0,2.$$

Нехай в перший район направлено m гелікоптерів, а в другий 10 — n . Подія A за умови виконання гіпотези H_1 полягає в виявленні апарату принаймні одним з m гелікоптерів, тому для обчислення умовної ймовірності $P(A/H_1)$ застосуємо формулу (3.2):

$$P(A/H_1) = 1 - (0,7)^m,$$

оскільки ймовірності знаходження апарату кожним з m гелікоптерів однакові і рівні $p = 0,3$, отже, $q = 0,7$.

Аналогічно $P(A/H_2) = 1 - (0,7)^{10-m}$, і за формулою повної ймовірності (3.8)

$$P(A) = 0,8 \cdot (1 - (0,7)^m) + 0,2 \cdot (1 - (0,7)^{10-m}) = 1 - 0,8 \cdot (0,7)^m - 0,2 \cdot (0,7)^{10-m}.$$

Далі потрібно знайти таке значення m , при якому ймовірність $P(A)$ буде максимальною. Прирівняємо похідну від $P(A)$ по m нулю:

$$\frac{dP(A)}{dm} = -0,8 \cdot (0,7)^m \cdot \ln(0,7) + 0,2 \cdot (0,7)^{10-m} \cdot \ln(0,7) = 0,$$

звідки

$$(0,7)^{2 \cdot m - 10} = 0,25.$$

або після логарифмування

$$2 \cdot m - 10 = \frac{\ln(0,25)}{\ln(0,7)} \approx 4,$$

звідки $m = 7$.

Отже, оптимальний розподіл гелікоптерів по районах пошуку: 7 – в перший район і 3 – в другий. При цьому найбільша ймовірність виявлення апарату

$$P(A) = 1 - 0,8 \cdot (0,7)^7 - 0,2 \cdot (0,7)^3 = 0,866.$$

Відповідь. 7 – в перший район і 3 – в другий; $P(A) = 0,866$.

Приклад 3.15. Проводиться випробування надійності приладу, який складається з двох вузлів. Вузли працюють або виходять з ладу незалежно один від одного. Надійності (ймовірності безвідмовної роботи за час t) першого і другого вузлів відомі і дорівнюють відповідно 0,9 і 0,95. Впродовж часу t прилад відмовив. Знайти ймовірність того, що до цього призвів: **а)** вихід з ладу 1-го вузла; **б)** вихід з ладу 2-го вузла; **в)** вихід з ладу обох вузлів.

Розв'язання. Позначимо подію $A = \{\text{за час } t \text{ прилад відмовив}\}$, а гіпотези H_1, H_2, H_3 – події, зазначені в умовах а)-в) задачі. Взагалі повна група гіпотез включає також гіпотезу $H_4 = \{\text{обидва вузли не вийшли з ладу}\}$, хоча сумісна поява подій A і H_4 є подією неможливою.

Ймовірності гіпотез обчислюються за формулою (2.8):

$$P(H_1) = 0,1 \cdot 0,95 = 0,095; P(H_2) = 0,9 \cdot 0,05 = 0,045;$$

$$P(H_3) = 0,1 \cdot 0,05 = 0,005; P(H_4) = 0,9 \cdot 0,95 = 0,855.$$

Оскільки подія A відбулась, то мала місце одна з гіпотез H_1, H_2, H_3 , причому $A = H_1 + H_2 + H_3$ (прилад відмовляє в разі появи будь-якої з цих гіпотез), тому

$$P(A) = P(H_1) + P(H_2) + P(H_3) = 0,145.$$

Далі за формулами Баєса (3.9) знаходимо апостеріорні ймовірності гіпотез $P(H_i/A)$, ($i = 1, 2, 3$) враховуючи, що всі умовні ймовірності

$P(H_i/A) = 1$, оскільки відмова приладу при появі будь-якої з гіпотез H_1, H_2, H_3 є подія достовірна:

$$P(H_1/A) = \frac{0,095}{0,145} = 0,655; P(H_2/A) = \frac{0,045}{0,145} = 0,31;$$

$$P(H_3/A) = \frac{0,005}{0,145} = 0,035.$$

Ймовірності гіпотез після того, як відбулася подія A , істотно зросли і ці апостеріорні ймовірності вже утворюють повну групу подій.

Відповідь. а) $P(H_1/A) = 0,655$; б) $P(H_2/A) = 0,31$; в) $P(H_3/A) = 0,035$.

Приклад 3.16. В продукції підприємства по виробництву електричних ламп число бракованих ламп серед будь-яких ста рівноможливе від 0 до 2. Знайти ймовірність того, що серед ста ламп не буде жодної бракованої, якщо з вибраних навмання десяти всі виявились придатними.

Розв'язання. Позначимо через A подію:

$$A = \{\text{серед 10-ти ламп всі придатні}\},$$

а через H_1, H_2, H_3 – гіпотези: серед ста ламп є відповідно 0, 1, 2 браковані.

За умовою задачі гіпотези рівноможливі:

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}.$$

Обчислимо умовні ймовірності події A для кожної з гіпотез:

$$P(A/H_1) = \frac{C_{100}^{10}}{C_{100}^{10}} = 1; P(A/H_2) = \frac{C_{99}^{10}}{C_{100}^{10}} = 0,9; P(A/H_3) = \frac{C_{98}^{10}}{C_{100}^{10}} = 0,81.$$

За формулою повної ймовірності (3.8)

$$P(A) = \frac{1}{3} \cdot (1 + 0,9 + 0,81) = 0,9,$$

а апостеріорна ймовірність гіпотези H_1 за формулою Байєса

$$P(H_1/A) = \frac{1}{3 \cdot 0,9} = 0,37.$$

Відповідь. $P(H_1/A) = 0,37$.

Приклад 3.17. Два стрільці незалежно один від одного виконали по одному пострілу у мішень. Ймовірність влучення у мішень першим стрільцем дорівнює 0,7; другим – 0,6. Після стрільби у мішені виявлено одну пробоїну. Знайти ймовірність того, що поцілив у мішень другий стрілець.

Розв'язання. Подія $A = \{\text{у мішені виявлено одну пробоїну}\}$.

Спочатку розглянемо можливі результати стрільби кожним стрільцем:

$A_1 = \{\text{перший стрілець поцілив у мішень}\}$; $\overline{A_1} = \{\text{перший стрілець не поцілив у мішень}\}$;

$A_2 = \{\text{другий стрілець поцілив у мішень}\}$; $\overline{A_2} = \{\text{другий стрілець не поцілив у мішень}\}$.

До проведення стрільби можливі такі гіпотези:

$H_1 = \{\text{обидва стрільці не поцілили у мішень}\}$;

$H_2 = \{\text{перший стрілець поцілив, другий не поцілив у мішень}\}$;

$H_3 = \{\text{перший стрілець не поцілив, другий – поцілив у мішень}\};$

$H_4 = \{\text{обидва стрільці поцілили у мішень}\}.$

Ймовірність цих подій:

$$P(H_1) = P(\overline{A_1} \cdot \overline{A_2}) = P(\overline{A_1}) \cdot P(\overline{A_2}) = 0,3 \cdot 0,7 = 0,12;$$

$$P(H_2) = P(A_1 \cdot \overline{A_2}) = P(A_1) \cdot P(\overline{A_2}) = 0,7 \cdot 0,4 = 0,28;$$

$$P(H_3) = P(\overline{A_1} \cdot A_2) = P(\overline{A_1}) \cdot P(A_2) = 0,3 \cdot 0,6 = 0,18;$$

$$P(H_4) = P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2) = 0,7 \cdot 0,6 = 0,42.$$

Умовні ймовірності події A після проведених пострілів для розглянутих гіпотез такі:

$$P(A/H_1) = 0; P(A/H_2) = 1; P(A/H_3) = 1; P(A/H_4) = 0.$$

Після проведених пострілів гіпотези H_1 і H_4 неможливі, а ймовірність гіпотези H_3 буде обчислюватися за формулою:

$$P(H_3/A) = \frac{P(H_3) \cdot P(A/H_3)}{\sum_{i=1}^4 P(H_i) \cdot P(A/H_i)} = \frac{0,18 \cdot 1}{0,28 \cdot 1 + 0,18 \cdot 1} = \frac{9}{23}.$$

Відповідь. Ймовірність того, що пробісна належить другому стрільцю дорівнює $\frac{9}{23}$.

Приклад 3.18. Один шах, якому страшенно наскучив придворний астролог з його брехливими віщуваннями, вирішив стратити його. Але, як людина не жорстока, він вирішив дати астрологу останній шанс. Астролог повинен розподілити по урнах 4 кулі: 2 білі та 2 чорні. Кат вибере навмання одну з урн і з неї витягне кулю. Якщо куля буде чорна, то астрологу відрубать голову, якщо ні – він залишиться живим. Як потрібно розподілити кулі в урнах, щоб забезпечити собі максимальну ймовірність залишитись живим.

Розв'язання. Вважатимемо, що астролог використає максимум 4 урни. Сформулюємо наступну подію: $A = \{\text{Астролог лишиться живим}\} = \{\text{Кат витягне білу кульку}\}.$

Настання цієї події можливо при наступних гіпотезах:

$H_1 = \{\text{Всього 4 урни: в кожній урні по 1 кульці}\};$

$H_2 = \{\text{Всього 3 урни: 1 урна – 2 чорні кульки, 2 урни – по 1 білій кульці}\};$

$H_3 = \{\text{Всього 3 урни: 1 урна – 2 білі кульки, 2 урни – по 1 чорній кульці}\};$

$H_4 = \{\text{Всього 3 урни: 1 урна – 1 біла кулька, 2 урна – 1 біла та 1 чорна кульки, 3 урна – 1 чорна кулька}\};$

$H_5 = \{\text{Всього 2 урни: 1 урна – 2 білі кульки, 2 урна – 2 чорні кульки}\};$

$H_6 = \{\text{Всього 2 урни: 1 урна – 1 біла та 1 чорна кульки, 2 урна – 1 біла та 1 чорна кульки}\};$

$H_7 = \{\text{Всього 2 урни: 1 урна – 1 біла кулька, 2 урна – 1 біла та 2 чорні кульки}\};$

$H_8 = \{\text{Всього 2 урни: 1 урна – 1 чорна кулька, 2 урна – 1 чорна та 2 білі кульки}\};$

$H_9 = \{\text{Всього 1 урна: 2 білі та 2 чорні кульки}\}.$

Обчислимо $P(A/H_i)$ ($i = 1, 2, \dots, 9$) за формулою повної ймовірності (3.8).

$$\begin{aligned}
P(A/H_1) &= \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 = \frac{1}{2}; \\
P(A/H_2) &= \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 = \frac{2}{3}; \\
P(A/H_3) &= \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 0 = \frac{1}{3}; \\
P(A/H_4) &= \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0 = \frac{1}{2}; \\
P(A/H_5) &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}; \\
P(A/H_6) &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}; \\
P(A/H_7) &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3}; \\
P(A/H_8) &= \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}; \\
P(A/H_9) &= 1 \cdot \frac{1}{2} = \frac{1}{2}.
\end{aligned}$$

Відповідь. Астрологу краще розмістити кульки в 3 урнах: 1 урна – 2 чорні кульки, 2 урни – по 1 білій кульці або в 2 урнах: 1 урна – 1 біла кулька, 2 урна – 1 біла та 2 чорні кульки.

Практичне застосування формули Баєса наведено в статті [27].

3.6. Наївний Баєсівський класифікатор

Наївний Баєсівський класифікатор (Naive Bayes classifier) — це проста й ефективна ймовірнісна модель для класифікаційних завдань, заснована на застосуванні формули Баєса з припущенням про незалежність ознак. Хоча припущення про незалежність ознак є досить сильним і рідко вірним у реальних задачах, наївний Баєсівський класифікатор часто показує хороші результати на практиці.

Основи Наївного Баєсівського класифікатора

Наївний Баєсівський класифікатор використовує формулу (3.10) для обчислення ймовірності належності зразка до певного класу. Для заданого зразка з ознаками $X = (x_1, x_2, \dots, x_n)$ і класу C_k , класифікатор обчислює апостеріорну ймовірність $P(C_k/X)$ за формулою (3.10).

$$P(C_k/X) = \frac{P(X/C_k) \cdot P(C_k)}{P(X)}, \quad (3.10)$$

де:

$P(C_k/X)$ — апостеріорна ймовірність класу C_k за умов спостережуваних ознак X ;

$P(X/C_k)$ — ймовірність спостережуваних ознак X за умов класу C_k ;

$P(C_k)$ — апріорна ймовірність класу C_k ;

$P(X)$ — ймовірність спостережуваних ознак X (нормалізуючий фактор).

Наївне припущення незалежності

Наївний Баєсівський класифікатор робить наївне припущення, що всі ознаки незалежні одна від одної за умов класу. Тобто:

$$P(X/C_k) = P((x_1, x_2, \dots, x_n)/C_k) = \prod_{i=1}^n P(x_i/C_k).$$

Тоді формула для апостеріорної ймовірності спрощується до (3.11):

$$P(C_k/X) = P(C_k) \cdot \prod_{i=1}^n P(x_i/C_k), \quad (3.11)$$

оскільки $P(X) = 1$, бо $X = \Omega$.

Виведення рішення

Для класифікації зразка X потрібно обчислити апостеріорні ймовірності для всіх можливих класів і вибрати клас з найбільшою ймовірністю (3.12):

$$\hat{C} = \arg \max_{C_k} P(C_k/X). \quad (3.12)$$

Приклад застосування

Припустимо, що у нас є набір даних про електронні листи, і ми хочемо класифікувати їх як «спам» або «не спам». Нехай ознаки $x_1, x_2, \dots, x_n$ представляють присутність або відсутність певних ключових слів у листі.

Обчислення апіорних ймовірностей

$$P(\text{спам}) = \frac{\text{кількість спам — листів}}{\text{загальна кількість листів}},$$
$$P(\text{не спам}) = \frac{\text{кількість не спам — листів}}{\text{загальна кількість листів}}.$$

Обчислення умовних ймовірностей для кожного ключового слова

$$P(x_i = 1/\text{спам}) = \frac{\text{кількість спам — листів з ключовим словом } x_i}{\text{кількість спам — листів}},$$
$$P(x_i = 1/\text{не спам}) = \frac{\text{кількість не спам — листів з ключовим словом } x_i}{\text{кількість не спам — листів}}.$$

Обчислення апостеріорної ймовірності для нового листа

$$P(\text{спам}/X) = P(\text{спам}) \cdot \prod_{i=1}^n P(x_i/\text{спам}),$$
$$P(\text{не спам}/X) = P(\text{не спам}) \cdot \prod_{i=1}^n P(x_i/\text{не спам}).$$

Класифікація листа

$$\hat{C} = \arg \max_{C \in \{\text{спам, не спам}\}} P(C/X).$$

Наївний Баєсівський класифікатор є потужним і ефективним інструментом для багатьох завдань класифікації, таких як фільтрація спаму, класифікація текстів і розпізнавання образів. Хоча припущення про незалежність ознак рідко виконується на практиці, цей класифікатор часто показує хороші результати завдяки своїй простоті і швидкості.

ЛЕКЦІЯ 4

ПОВТОРНІ НЕЗАЛЕЖНІ ВИПРОБУВАННЯ

Нагадаємо, що випробування називають *незалежними*, якщо ймовірність будь-якого результату випробування не залежить від того, які наслідки мали інші випробування.

Повторні незалежні випробування – це важливий концепт у теорії ймовірностей, який описує ситуації, коли одна й та сама випадкова подія повторюється багаторазово незалежно від попередніх результатів.



Рис. 4.1. Основні схеми для повторних незалежних випробувань

Окрім вищеназваних схем для розв'язання задач використовується Поліноміальна схема, яка є узагальненням схеми (формули) Бернуллі.

4.1. Формула Бернуллі. Найймовірніша кількість успіхів

У теорії ймовірності, формула Бернуллі дозволяє обчислити ймовірність успіхів у серії незалежних експериментів.

Теорема 4.1. Якщо ймовірність p настання події A в кожному випробуванні постійна, то ймовірність $P_n(m)$ того, що подія A наступить m раз в n незалежних випробуваннях, обчислюється за формулою (4.1):

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m} \quad (m = 1, 2, \dots, n). \quad (4.1)$$

Доведення. Нехай проводиться серія з n незалежних в сукупності випробувань. В кожному з цих випробувань подія A може з'явитись зі сталою ймовірністю p і не з'явитись з ймовірністю $q = 1 - p$. Така серія випробувань називається *схемою Бернуллі*.

Знайдемо ймовірність того, що в схемі Бернуллі подія A з'явиться рівно m разів ($0 \leq m \leq n$). Настання події A будемо називати «успіхом».

Позначимо через $P_n(m)$ ймовірність того, що в серії з n незалежних в сукупності випробувань подія A з'явиться рівно m разів. Нехай подія A_i

полягає у тому, що подія A з'явилась у i -му випробуванні. З теорем додавання і множення ймовірностей, дістанемо

$$\begin{aligned} P_n(m) &= P(A_1 \cdot A_2 \cdot \dots \cdot A_m \cdot \overline{A_{m+1}} \cdot \overline{A_{m+2}} \cdot \dots \cdot \overline{A_n}) + \dots \\ &+ P(\overline{A_1} \cdot \overline{A_2} \cdot \dots \cdot \overline{A_{n-m}} \cdot A_{n-m+1} \cdot A_{n-m+2} \cdot \dots \cdot A_n) = \\ &= P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_m) \cdot P(\overline{A_{m+1}}) \cdot P(\overline{A_{m+2}}) \cdot \dots \cdot P(\overline{A_n}) + \dots \\ &+ P(\overline{A_1}) \cdot P(\overline{A_2}) \cdot \dots \cdot P(\overline{A_{n-m}}) \cdot P(A_{n-m+1}) \cdot P(A_{n-m+2}) \cdot \dots \cdot P(A_n) = \\ &= \underbrace{p \cdot p \cdot \dots \cdot p}_m \cdot \underbrace{q \cdot q \cdot \dots \cdot q}_{n-m} + \dots + \underbrace{p \cdot p \cdot \dots \cdot p}_m \cdot \underbrace{q \cdot q \cdot \dots \cdot q}_{n-m}. \end{aligned}$$

Всього доданків $p^m \cdot q^{n-m}$ буде стільки, скількома способами можна з n випробувань вибрати m , а отже отримуємо (4.1).

Формула (4.1) називається **формулою Бернуллі**.

Формула Бернуллі широко використовується в задачах з дискретної ймовірності, наприклад:

- Задачі з монетами (ймовірність випадання певної кількості гербів при підкиданні монети кілька разів);
- Задачі з урнами (ймовірність витягнути певну кількість кульок певного кольору);
- Генетичні задачі (ймовірність певної кількості успадкувань ознаки).

Формула Бернуллі дає точні значення ймовірностей $P_n(m)$, проте при великих значеннях n вона приводить до громіздких обчислень.

Приклад 4.1. Підкидаємо три рази монету. Знайдіть ймовірність того, що випало рівно два герба.

Розв'язання. За формулою Бернуллі

$$P_3(2) = C_3^2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{3-2} = 3 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{8}.$$

Відповідь. $P_3(2) = \frac{3}{8}$.

Приклад 4.2. Яка ймовірність того, що при n підкиданнях грального кубика хоча б два рази з'явиться шістка (подія A)?

Розв'язання. Знайдемо ймовірність події $\bar{A}$ – шістка з'явилась менше двох раз.

$$P(\bar{A}) = P_n(0) + P_n(1).$$

Оскільки $p = \frac{1}{6}$, $q = \frac{5}{6}$, то

$$\begin{aligned} P(\bar{A}) &= C_n^0 \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^n + C_n^1 \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{5}{6}\right)^{n-1} = \left(\frac{5}{6}\right)^n + n \cdot \frac{5^{n-1}}{6^n} = \\ &= \frac{5^{n-1}}{6^n} \cdot (5 + n). \end{aligned}$$

$$\text{Тоді } P(A) = 1 - P(\bar{A}) = 1 - \frac{5^{n-1}}{6^n} \cdot (5 + n).$$

Відповідь. $P(A) = 1 - \frac{5^{n-1}}{6^n} \cdot (5 + n)$.

Приклад 4.3. Випадково зустрінутий на вулиці перехожий з ймовірністю 0,3 – шатен, з ймовірністю 0,4 – блондин, з ймовірністю 0,2 – брюнет і з ймовірністю 0,1 – рудий. Яка ймовірність того, що серед зустрінутих 6 людей **а)** не менше 4 блондинів, **б)** хоча б один рудий, **в)** 3 блондина і 3 шатена?

Розв'язання.

$$\text{а)} P_6(4 \leq m \leq 6) = C_6^4 \cdot p^4 \cdot q^2 + C_6^5 \cdot p^5 \cdot q^1 + C_6^6 \cdot p^6 \cdot q^0.$$

За умовою $p = 0,4$; $q = 0,6$.

Отже,

$$P_6(4 \leq m \leq 6) = 15 \cdot (0,4)^4 \cdot (0,6)^2 + 6 \cdot (0,4)^5 \cdot (0,6)^1 + 1 \cdot (0,4)^6 \cdot (0,6)^0 = 0,1792.$$

$$\text{б)} P_6(1 \leq m \leq 6) = 1 - P_6(0).$$

В цьому випадку, за умовою, $p = 0,1$; $q = 0,9$.

Звідки

$$P_6(1 \leq m \leq 6) = 1 - (0,1)^0 \cdot (0,9)^6 = 0,468559.$$

в) Розглянемо наступні події: $A = \{\text{зустрілись 3 блондина і 3 шатена}\}$; $B_i = \{i\text{-тий зустрінутий} - \text{блондин}\}$; $C_i = \{i\text{-тий зустрінутий} - \text{шатен}\}$ ($i = 1, 2, \dots, 6$).

Тоді

$$P(A) = P(B_1 \cdot B_2 \cdot B_3 \cdot C_4 \cdot C_5 \cdot C_6) + P(B_1 \cdot B_2 \cdot C_3 \cdot B_4 \cdot C_5 \cdot C_6) + \dots + P(C_1 \cdot C_2 \cdot C_3 \cdot B_4 \cdot B_5 \cdot B_6).$$

Доданків буде стільки, скількома способами можна вибрати трьох, наприклад, блондинів з шести чоловік, тобто C_6^3 . Кожен доданок дорівнює $(0,3)^3 \cdot (0,4)^3 \cdot (0,3)^0$.

Тоді

$$P(A) = 20 \cdot (0,3)^3 \cdot (0,4)^3 \cdot (0,3)^0 = 0,03456.$$

Відповідь. а) $P_6(4 \leq m \leq 6) = 0,1792$; б) $P_6(1 \leq m \leq 6) = 0,468559$; в) $P(A) = 0,03456$.

Приклад 4.4. Підводний човен атакує крейсер: випускає по ньому одну за одною 4 торпеди. Ймовірність попадання кожної дорівнює $\frac{3}{4}$. Кожна з торпед з однаковою ймовірністю може попасти в один з 10 відсіків крейсера. Якщо торпеди попали хоча б в два різних відсіки, то крейсер буде потоплено. Яка ймовірність цього (подія A)?

Розв'язання. Нехай події A_i полягають у тому, що в крейсер попало i торпед ($i = 2, 3, 4$). В цих випадках крейсер може бути потоплен.

Тоді

$$P(A_2) = P_4(2) = C_4^2 \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^2 = \frac{27}{128};$$

$$P(A_3) = P_4(3) = C_4^3 \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^1 = \frac{27}{64};$$

$$P(A_4) = P_4(4) = C_4^4 \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^0 = \frac{81}{256}.$$

Треба врахувати, що всі i торпед можуть попасти в один відсік; ймовірність цього $\frac{1}{10^{i-1}}$. Тоді крейсер не буде потоплено. Отже,

$$\begin{aligned} P(A) &= P(A_2) \cdot P(A/A_2) + P(A_3) \cdot P(A/A_3) + P(A_4) \cdot P(A/A_4) = \\ &= \frac{27}{128} \cdot \left(1 - \frac{1}{10^{2-1}}\right) + \frac{27}{64} \cdot \left(1 - \frac{1}{10^{3-1}}\right) + \frac{81}{256} \cdot \left(1 - \frac{1}{10^{4-1}}\right) \approx \\ &\approx 0,9236. \end{aligned}$$

Відповідь. $P(A) \approx 0,9236$.

Приклад 4.5. Ймовірність виготовлення на автоматичному верстаті стандартної деталі дорівнює 0,8. Знайти ймовірності можливого числа появи бракованих деталей серед 5 відібраних.

Розв'язання. Ймовірність виготовлення бракованої деталі $p = 1 - 0,8 = 0,2$. Шукані ймовірності знаходимо за формулою (4.1):

$$P_5(0) = C_5^0 \cdot (0,2)^0 \cdot (0,8)^5 = 0,32768;$$

$$P_5(1) = C_5^1 \cdot (0,2)^1 \cdot (0,8)^4 = 0,4096;$$

$$P_5(2) = C_5^2 \cdot (0,2)^2 \cdot (0,8)^3 = 0,2048;$$

$$P_5(3) = C_5^3 \cdot (0,2)^3 \cdot (0,8)^2 = 0,0512;$$

$$P_5(4) = C_5^4 \cdot (0,2)^4 \cdot (0,8)^1 = 0,0064;$$

$$P_5(5) = C_5^5 \cdot (0,2)^5 \cdot (0,8)^0 = 0,00032.$$

Отримані результати представимо у вигляді таблиці 4.1.

Таблиця 4.1. Результати для можливого числа бракованих деталей

m	0	1	2	3	4	5
$P_n(m)$	0,32768	0,4096	0,2048	0,0512	0,0064	0,00032

Згідно табл. 4.1, отримані ймовірності зобразимо графічно точками з координатами $(m, P_n(m))$. Поєднуючи ці точки, отримаємо багатокутник, або полігон, розподілу ймовірностей (рис. 4.1).

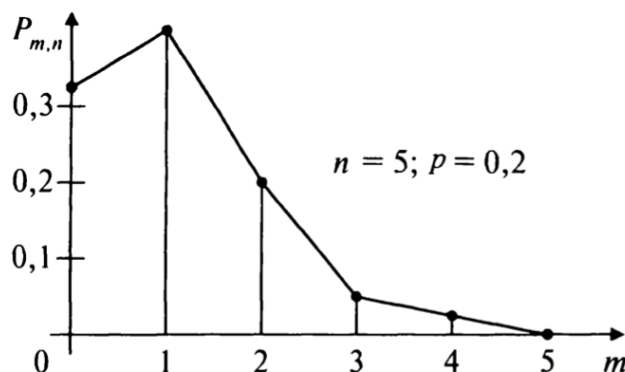


Рис. 4.1. Багатокутник розподілу при $n = 5$

Розглядаючи багатокутник розподілу ймовірностей (рис. 4.1), ми бачимо, що є такі значення m (в даному випадку, одне – $m_0 = 1$), що мають найбільшу ймовірність $P_n(m)$.

Відповідь. $P_5(0) = 0,32768; P_5(1) = 0,4096; P_5(2) = 0,2048; P_5(3) = 0,0512; P_5(4) = 0,0064; P_5(5) = 0,00032$.

Число m_0 настання події A в n незалежних випробуваннях називається *найймовірніше*, якщо ймовірність здійснення цієї події $P_n(m)$ буде, принаймні, не менше ймовірностей інших подій $P_n(m)$ при будь-якому n .

Теорема 4.2. Найймовірніше число m_0 появ події A в n випробуваннях визначається з нерівностей (4.2):

$$n \cdot p - q \leq m_0 \leq n \cdot p + p. \quad (4.2)$$

Доведення. Для знаходження m_0 запишемо систему нерівностей

$$\begin{cases} P_n(m_0) \geq P_n(m_0 + 1), \\ P_n(m_0) \geq P_n(m_0 - 1). \end{cases}$$

З першої нерівності, враховуючи формулу Бернуллі, матимемо

$$\frac{n!}{m_0! \cdot (n - m_0)!} \cdot p^{m_0} \cdot q^{n-m_0} \geq \frac{n!}{(m_0 + 1)! \cdot (n - m_0 - 1)!} \cdot p^{m_0+1} \cdot q^{n-m_0-1}.$$

Оскільки

$$(m_0 + 1)! = m_0! \cdot (m_0 + 1),$$

$$\text{то } \frac{1}{n-m_0} \cdot q \geq \frac{1}{m_0+1} \cdot p, \text{ або } (m_0 + 1) \cdot q \geq (n - m_0) \cdot p.$$

Остаточно $m_0 \geq n \cdot p - q$, враховуючи, що $p + q = 1$.

Аналогічно для другої нерівності $m_0 \geq n \cdot p + p$.

Об'єднуючи отримані рішення двох нерівностей, приходимо до подвійного нерівності:

$$n \cdot p - q \leq m_0 \leq n \cdot p + p.$$

Оскільки $n \cdot p + p - (n \cdot p - q) = p + q = 1$, число m_0 завжди існує, причому, якщо $n \cdot p + q$ – ціле число, то найймовірніших чисел буде два: $m_0 = n \cdot p - q$ і $m_0 = n \cdot p + p$.

Якщо ймовірність появи події A в кожному з n випробувань дорівнює p , то кількість випробувань, які необхідно здійснити, щоб з ймовірністю P можна було стверджувати, що подія A відбудеться хоча б один раз, обчислюється за формулою (4.3):

$$n \geq \frac{\ln(1-P(A))}{\ln(1-p)}. \quad (4.3)$$

Обчислення ймовірностей за формулою Бернуллі при великих значеннях n і m пов'язане з певними труднощами.

Приклад 4.6. Ймовірність відмовлення кожного приладу при його випробуванні складає 0,2. Скільки приладів потрібно випробувати, щоб з ймовірністю не менше 0,9 отримати хоча б один прилад, який відмовить.

Розв'язання. Нехай $A = \{\text{при } n \text{ випробувань хоча б один прилад відмовить}\}$.

Тоді $p = 0,2; P(A) = 0,9$. Одержимо за формулою (4.3):

$$n \geq \frac{\ln(0,1)}{\ln(0,8)} \approx 10,3; n \geq 11.$$

Відповідь. Необхідно випробувати не менше 11 приладів.

4.2. Локальна теорема Муавра-Лапласа

Локальна теорема Муавра-Лапласа є важливим результатом теорії ймовірностей. Вона апроксимує ймовірність появи події A рівно m разів в серії з n повторних незалежних випробувань (при великих значеннях параметра n).

Теорема 4.3. Якщо ймовірність p появи події A у кожному випробуванні стала і відмінна від 0 і 1, то ймовірність $P_n(m)$ того, що подія A відбудеться m разів в n випробуваннях для достатньо великих значень n , наближено обчислюється за формулою (4.4):

$$P_n(m) \approx \frac{1}{\sqrt{n \cdot p \cdot q}} \cdot \varphi(x), \quad (4.4)$$

де $\varphi(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{x^2}{2}}$ – функція Гауса, таблиця значень якої наведено в додатку 1 Гмурмана, де $x = \frac{m - n \cdot p}{\sqrt{n \cdot p \cdot q}}$. Чим більше n , тим точніше формула (4.4). Вона дає невелику похибку, якщо $n \cdot p \cdot q \geq 10$.

Доведення. Процес доведення має наступні етапи.

Формула Бернуллі

Ймовірність того, що випадкова величина A настане рівно m разів в серії з n експериментів обчислюється за формулою (4.1).

Формула Стірлінга

Для великих n використовуємо наближення Стірлінга для факторіалів:

$$n! \approx \sqrt{2 \cdot \pi \cdot n} \cdot \left(\frac{n}{e}\right)^n.$$

Застосуємо формулу Стірлінга до біноміального коефіцієнта:

$$C_n^m = \frac{n!}{m! \cdot (n - m)!} \approx \frac{\sqrt{2 \cdot \pi \cdot n} \cdot \left(\frac{n}{e}\right)^n}{\sqrt{2 \cdot \pi \cdot m} \cdot \left(\frac{m}{e}\right)^m \cdot \sqrt{2 \cdot \pi \cdot (n - m)} \cdot \left(\frac{n - m}{e}\right)^{n - m}}.$$

Спрощення

Спрощуємо це наближення:

$$C_n^m \approx \frac{\sqrt{2 \cdot \pi \cdot n} \cdot \left(\frac{n}{e}\right)^n}{\sqrt{2 \cdot \pi \cdot m} \cdot \left(\frac{m}{e}\right)^m \cdot \sqrt{2 \cdot \pi \cdot (n - m)} \cdot \left(\frac{n - m}{e}\right)^{n - m}}$$

Далі:

$$C_n^m \approx \frac{\sqrt{n} \cdot \left(\frac{n}{m}\right)^m \cdot \left(\frac{n}{n - m}\right)^{n - m}}{\sqrt{2 \cdot \pi \cdot m \cdot (n - m)}}.$$

Підставлення

Підставимо це у формулу для $P_n(m)$:

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m} \approx \frac{\sqrt{n} \cdot \left(\frac{n}{m}\right)^m \cdot \left(\frac{n}{n-m}\right)^{n-m}}{\sqrt{2 \cdot \pi \cdot m \cdot (n-m)}} \cdot p^m \cdot q^{n-m}.$$

Перепишемо в зручнішій формі:

$$P_n(m) \approx \frac{\sqrt{n}}{\sqrt{2 \cdot \pi \cdot m \cdot (n-m)}} \cdot \left(\frac{n^n}{m^m \cdot (n-m)^{n-m}}\right) \cdot p^m \cdot q^{n-m}.$$

Логарифмічне перетворення:

Враховуючи $m = n \cdot p + \Delta$, логарифмуючи 2, 3 та 4 множники, а також застосовуючи розкладання в ряд Тейлора, отримуємо:

$$P_n(m) \approx \frac{1}{\sqrt{2 \cdot \pi \cdot p \cdot q}} \cdot e^{-\left(\frac{\Delta^2}{2 \cdot n \cdot p \cdot q}\right)},$$

звідки й отримуємо формулу (4.4).

Властивості

1. Функція $\varphi(x)$ парна: $\varphi(x) = \varphi(-x)$;
2. Функція $\varphi(x)$ – монотонно спадна (практично можна вважати, що для $x > 4$: $\varphi(x) \approx 0$).

Значение функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
	Сотые доли x									
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3721	3605	3588	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3411	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1827	1804	1781	1759	1736
1,3	1714	1692	1669	1647	1626	1604	1582	1561	1540	1519

Рис. 4.2. Фрагмент таблиці функції Гауса

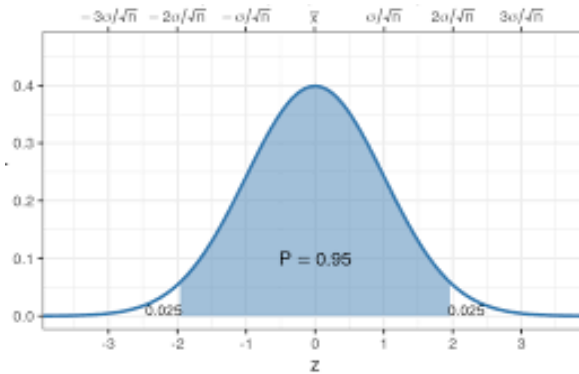


Рис. 4.3. Функція Гауса

Приклад 4.7. Перевірки податкової інспекції показали, що кожне друге мале підприємство порушує фінансову дисципліну. **1.** Знайти ймовірність того, що серед 800 малих підприємств району порушили фінансову дисципліну: 380 підприємств. **2.** Яке найімовірніше число малих підприємств, що порушили фінансову дисципліну?

Розв'язання. 1. За умовою $p = 0,5$. Число малих підприємств достатньо велике $n = 800$, тому застосуємо локальну формулу Муавра-Лапласа, крім того

$$n \cdot p \cdot q = 800 \cdot 0,5 \cdot (1 - 0,5) = 200.$$

Для цього знайдемо

$$x = \frac{380 - 800 \cdot 0,5}{\sqrt{800 \cdot 0,5 \cdot 0,5}} = -1,414.$$

За формулою (4,4) знайдемо:

$$P_{800}(380) \approx \frac{\varphi(-1,414)}{\sqrt{800 \cdot 0,5 \cdot 0,5}} \approx \frac{\varphi(1,414)}{\sqrt{200}} \approx \frac{0,1468}{\sqrt{200}} \approx 0,0106.$$

2. Тепер знайдемо найімовірніше число малих підприємств, що порушили фінансову дисципліну.

Маємо: $800 \cdot 0,5 - 0,5 \leq m_0 \leq 800 \cdot 0,5 + 0,5$, тобто $399,5 \leq m_0 \leq 400,5$. Звідси $m_0 = 400$.

Ймовірність цієї події, з-за того, що $x = 400 - 800 \cdot 0,5 = 0$, дорівнює

$$P_{800}(400) \approx \frac{\varphi(0)}{\sqrt{200}} \approx \frac{0,3989}{\sqrt{200}} \approx 0,0282.$$

Відповідь. 1. $P_{800}(380) \approx 0,0106$; $m = 400$; 2. $P_{800}(400) \approx 0,0282$.

4.3. Інтегральна теорема Муавра-Лапласа

Інтегральна теорема Муавра-Лапласа є результатом теорії ймовірностей, який дозволяє апроксимувати ймовірність появи події A від m_1 до m_2 разів в серії з n повторних незалежних випробувань.

Теорема 4.4. Якщо ймовірність p появи події A в кожному випробуванні стала і відмінна від 0 і 1, то ймовірність того, що число m появ події A в n незалежних випробуваннях буде міститься у границях від m_1 до

m_2 (включно), для достаточно великих n наближено обчислюється за формулою (4.5):

$$P_n(m_1 \leq m \leq m_2) \approx \Phi(x_2) - \Phi(x_1), \quad (4.5)$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{t^2}{2}} dt$ – функція Лапласа, таблиця значень якої наведена в додатку 2 Гмурмана, де $x_1 = \frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}}$, $x_2 = \frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}}$.

Доведення. Інтегральна теорема Муавра-Лапласа є наслідком центральної граничної теореми. Доведення буде надано пізніше.

Формула (4.5) дає невелику похибку для $n \cdot p \cdot q \geq 10$.

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0.00	0.0000	0.49	0.1879	0.98	0.3365	1.47	0.4292
0.01	0.0040	0.50	0.1915	0.99	0.3389	1.48	0.4306
0.02	0.0080	0.51	0.1950	1.00	0.3413	1.49	0.4319
0.03	0.0120	0.52	0.1985	1.01	0.3438	1.50	0.4332
0.04	0.0160	0.53	0.2019	1.02	0.3461	1.51	0.4345
0.05	0.0199	0.54	0.2054	1.03	0.3485	1.52	0.4357
0.06	0.0239	0.55	0.2088	1.04	0.3508	1.53	0.4370
0.07	0.0279	0.56	0.2123	1.05	0.3531	1.54	0.4382
0.08	0.0319	0.57	0.2157	1.06	0.3554	1.55	0.4394
0.09	0.0359	0.58	0.2190	1.07	0.3577	1.56	0.4406
0.10	0.0398	0.59	0.2224	1.08	0.3599	1.57	0.4418
0.11	0.0438	0.60	0.2257	1.09	0.3621	1.58	0.4429
0.12	0.0478	0.61	0.2291	1.10	0.3643	1.59	0.4441
0.13	0.0517	0.62	0.2324	1.11	0.3665	1.60	0.4452
0.14	0.0557	0.63	0.2357	1.12	0.3686	1.61	0.4463
0.15	0.0596	0.64	0.2389	1.13	0.3708	1.62	0.4474
0.16	0.0636	0.65	0.2422	1.14	0.3729	1.63	0.4484
0.17	0.0675	0.66	0.2454	1.15	0.3749	1.64	0.4495
0.18	0.0714	0.67	0.2486	1.16	0.3770	1.65	0.4505
0.19	0.0753	0.68	0.2517	1.17	0.3790	1.66	0.4515
0.20	0.0793	0.69	0.2549	1.18	0.3810	1.67	0.4525
0.21	0.0832	0.70	0.2580	1.19	0.3830	1.68	0.4535
0.22	0.0871	0.71	0.2611	1.20	0.3949	1.69	0.4545
0.23	0.0910	0.72	0.2642	1.21	0.3869	1.70	0.4554
0.24	0.0948	0.73	0.2673	1.22	0.3888	1.71	0.4564
0.25	0.0987	0.74	0.2703	1.23	0.3907	1.72	0.4573
0.26	0.1026	0.75	0.2734	1.24	0.3925	1.73	0.4582
0.27	0.1064	0.76	0.2764	1.25	0.3944	1.74	0.4591
0.28	0.1103	0.77	0.2794	1.26	0.3962	1.75	0.4599
0.29	0.1141	0.78	0.2823	1.27	0.3980	1.76	0.4608
0.30	0.1179	0.79	0.2852	1.28	0.3997	1.77	0.4616
0.31	0.1217	0.80	0.2881	1.29	0.4015	1.78	0.4625
0.32	0.1255	0.81	0.2910	1.30	0.4032	1.79	0.4633
0.33	0.1293	0.82	0.2939	1.31	0.4049	1.80	0.4641
0.34	0.1331	0.83	0.2967	1.32	0.4066	1.81	0.4649
0.35	0.1368	0.84	0.2995	1.33	0.4082	1.82	0.4656
0.36	0.1406	0.85	0.3023	1.34	0.4099	1.83	0.4664
0.37	0.1443	0.86	0.3051	1.35	0.4115	1.84	0.4671
0.38	0.1480	0.87	0.3078	1.36	0.4131	1.85	0.4678
0.39	0.1517	0.88	0.3106	1.37	0.4147	1.86	0.4686
0.40	0.1554	0.89	0.3133	1.38	0.4162	1.87	0.4693
0.41	0.1591	0.90	0.3159	1.39	0.4177	1.88	0.4699
0.42	0.1628	0.91	0.3186	1.40	0.4192	1.89	0.4706
0.43	0.1664	0.92	0.3212	1.41	0.4207	1.90	0.4713
0.44	0.1700	0.93	0.3238	1.42	0.4222	1.91	0.4719
0.45	0.1736	0.94	0.3264	1.43	0.4236	1.92	0.4726
0.46	0.1772	0.95	0.3289	1.44	0.4251	1.93	0.4732
0.47	0.1808	0.96	0.3315	1.45	0.4265	1.94	0.4738
0.48	0.1844	0.97	0.3340	1.46	0.4279	1.95	0.4744

Рис. 4.4. Фрагмент таблиці значень функції Лапласа

Графік функції Лапласа представлено на рис (4.3), а фрагмент таблиці значень на рис. 4.4.

Властивості

1. Функція Лапласа $\Phi(x)$ непарна ($\Phi(-x) = -\Phi(x)$);
2. Функція Лапласа монотонно зростаюча;
3. $\Phi(x) \rightarrow 0,5$, для $x \rightarrow +\infty$ (для $x \geq 5$ можна вважати, що $\Phi(x) = 0,5$).

Приклад 4.8. Ймовірність появи події в кожному з 100 незалежних випробувань стала і рівна $p = 0,8$. Знайти ймовірність того, що подія з'явиться не менше 75 раз і не більше 90 разів.

Розв'язання. Скористаємося інтегральною теоремою Муавра-Лапласа (4.5).

За умовою, $n = 100$; $p = 0,8$; $q = 0,2$; $m_1 = 75$; $m_2 = 90$. Обчислимо x_1 і x_2 :

$$x_1 = \frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}} = \frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = -1,25;$$
$$x_2 = \frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}} = \frac{90 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}} = 2,5.$$

Враховуючи, що функція Лапласа непарна, тобто $\Phi(-x) = -\Phi(x)$, отримаємо:

$$P_{100}(75 \leq m \leq 90) = \Phi(2,5) - \Phi(-1,25) = 0,4938 + 0,3944 = 0,8882.$$

Відповідь. $P_{100}(75 \leq m \leq 90) = 0,8882$.

Приклад 4.9. Ймовірність появи події в кожному з незалежних випробувань рівна 0,8. Скільки потрібно виконати випробувань, щоб з ймовірністю 0,9 можна було очікувати, що подія з'явиться не менше 75 разів?

Розв'язання. За умовою $p = 0,8$; $q = 0,2$; $m_1 = 75$; $m_2 = n$; $P_n(75 \leq m \leq n) = 0,9$.

Скористаємося інтегральною теоремою Лапласа (4.5).

Підставляючи дані завдання, отримаємо

$$0,9 = \Phi\left(\frac{n - n \cdot 0,8}{\sqrt{n \cdot 0,8 \cdot 0,2}}\right) - \Phi\left(\frac{75 - n \cdot 0,8}{\sqrt{n \cdot 0,8 \cdot 0,2}}\right)$$

або

$$0,9 = \Phi\left(\frac{\sqrt{n}}{2}\right) - \Phi\left(\frac{75 - n \cdot 0,8}{0,4 \cdot \sqrt{n}}\right).$$

Очевидно, число випробувань $n > 75$, тому $\frac{\sqrt{n}}{2} > \frac{\sqrt{75}}{2} \approx 4,33$.

Оскільки функція Лапласа зростає і $\Phi(4,33) \approx 0,5$, то можна покласти $\Phi\left(\frac{\sqrt{n}}{2}\right) = 0,5$. Отже,

$$0,9 = 0,5 - \Phi\left(\frac{75 - n \cdot 0,8}{0,4 \cdot \sqrt{n}}\right).$$

Таким чином, отримуємо рівняння (4.6)

$$\Phi\left(\frac{75 - n \cdot 0,8}{0,4 \cdot \sqrt{n}}\right) = -0,4. \quad (4.6)$$

За таблицею у рис. 4.4 знайдемо $\Phi(1,28) = 0,4$. Звідси з співвідношення (4.6), враховуючи, що функція Лапласа непарна, отримаємо

$$\frac{75 - n \cdot 0,8}{0,4 \cdot \sqrt{n}} = -1,28.$$

Розв'язавши це рівняння, як квадратне відносно $\sqrt{n}$, отримаємо $\sqrt{n} = 10$. Отже, шукане число випробувань $n = 100$.

Відповідь. $n = 100$.

Наслідки з інтегральної теореми Муавра-Лапласа

1. $P_n(|m - n \cdot p| \leq \varepsilon) \approx 2 \cdot \Phi\left(\frac{\varepsilon}{\sqrt{n \cdot p \cdot q}}\right)$;
2. $P_n\left(\alpha \leq \frac{m}{n} \leq \beta\right) \approx \Phi(z_2) - \Phi(z_1)$, де $z_1 = \frac{\alpha - p}{\sqrt{\frac{p \cdot q}{n}}}$, $z_2 = \frac{\beta - p}{\sqrt{\frac{p \cdot q}{n}}}$;
3. $P_n\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) \approx 2 \cdot \Phi\left(\frac{\varepsilon \sqrt{n}}{\sqrt{p \cdot q}}\right)$.

Доведення. Доведемо властивість 1.

$$\begin{aligned} P_n(|m - n \cdot p| \leq \varepsilon) &= P_n(n \cdot p - \varepsilon \leq m \leq n \cdot p + \varepsilon) \approx \\ &\approx \Phi\left(\frac{n \cdot p + \varepsilon - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right) - \Phi\left(\frac{n \cdot p - \varepsilon - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right) = \\ &= \Phi\left(\frac{\varepsilon}{\sqrt{n \cdot p \cdot q}}\right) - \Phi\left(\frac{-\varepsilon}{\sqrt{n \cdot p \cdot q}}\right) = \\ &= \Phi\left(\frac{\varepsilon}{\sqrt{n \cdot p \cdot q}}\right) + \Phi\left(\frac{\varepsilon}{\sqrt{n \cdot p \cdot q}}\right) = 2 \cdot \Phi\left(\frac{\varepsilon}{\sqrt{n \cdot p \cdot q}}\right). \end{aligned}$$

Наслідки 2 і 3 доводяться аналогічно.

Приклад 4.10. Ймовірність виготовити на заводі виріб найвищої якості дорівнює 0,9. Навмання беруть 200 виробів. Визначити межі, в яких перебуватиме відносна частота появи виробів найвищої якості з ймовірністю 0,96.

Розв'язання. За умовою задачі: $n = 200$; $p = 0,9$; $P_n\left(\left|\frac{m}{n} - 0,9\right| \leq \varepsilon\right) \approx 0,96$. Підставивши ці значення в наслідок 3, дістанемо:

$$\begin{aligned} P_n\left(\left|\frac{m}{n} - 0,9\right| \leq \varepsilon\right) &\approx 2 \cdot \Phi\left(\frac{\varepsilon \sqrt{200}}{\sqrt{0,9 \cdot 0,1}}\right) \approx 0,96; \\ 2 \cdot \Phi\left(\frac{\varepsilon \sqrt{200}}{\sqrt{0,9 \cdot 0,1}}\right) &\approx 0,96; \\ \Phi\left(\frac{\varepsilon \sqrt{200}}{\sqrt{0,9 \cdot 0,1}}\right) &\approx 0,48; \\ \varepsilon \cdot 27,14 &= 2,06; \\ \varepsilon &= 0,044. \end{aligned}$$

Тоді

$$\begin{aligned} \left| \frac{m}{n} - 0,9 \right| &\leq 0,044; \\ -0,044 &\leq \frac{m}{n} - 0,9 \leq 0,044; \\ 0,9 - 0,044 &\leq \frac{m}{n} \leq 0,9 + 0,044; \\ 0,856 &\leq \frac{m}{n} \leq 0,944. \end{aligned}$$

Відповідь. $0,856 \leq \frac{m}{n} \leq 0,944$.

4.4. Формула Пуассона

Формула Пуассона (4.7) використовується для обчислення ймовірності того, що певна кількість подій відбудеться за фіксований інтервал часу або в певній області простору, коли події відбуваються незалежно одна від одної і з постійною середньою швидкістю.

Теорема 4.5. Якщо ймовірність p появи події A у кожному випробуванні прямує до 0 ($p \rightarrow 0$) з одночасним необмеженим зростанням числа n випробувань ($n \rightarrow +\infty$), то ймовірність $P_n(m)$ того, що подія A відбудеться m разів в n незалежних випробувань, обчислюється за формулою (4.7):

$$P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!}, \quad (4.7)$$

де $\lambda = n \cdot p \leq 10$.

Доведення. Розглянемо випадковий процес (серію з n експериментів), у якому події відбуваються незалежно одна від одної, з постійною середньою швидкістю λ .

Ймовірність подій за малий інтервал часу Δt

Припустимо, що Δt є малим інтервалом часу. Тоді ймовірність того, що відбудеться одна подія в цьому інтервалі часу, пропорційна Δt :

$$P_n(m = 1 \text{ подія в } \Delta t) = p \approx \lambda \cdot \Delta t.$$

Ймовірність того, що події не відбудеться в Δt :

$$P_n(m = 0 \text{ подій в } \Delta t) = q = 1 - p \approx 1 - \lambda \cdot \Delta t.$$

Ймовірність того, що відбудеться більше ніж одна подія в Δt , настільки мала, що її можна знехтувати:

$$P_n(m \geq 2 \text{ подій в } \Delta t) \approx 0.$$

Розбиття інтервалу часу

Розділимо інтервал часу $[0; t]$ на n малих інтервалів $[0; \Delta t]$, де $\Delta t = \frac{t}{n}$.

Ймовірність k подій

Нехай k подій відбудеться за інтервал часу $[0; t]$. Ми можемо вважати, що ці k подій відбуваються у n малих інтервалах часу $[0; \Delta t]$. Тоді за формулою Бернуллі (4.1):

$$P_n(m) = C_n^m \cdot (\lambda \cdot \Delta t)^m \cdot (1 - \lambda \cdot \Delta t)^{n-m}.$$

Підставимо $\Delta t = \frac{t}{n}$:

$$P_n(m) = C_n^m \cdot \left(\lambda \cdot \frac{t}{n}\right)^m \cdot \left(1 - \lambda \cdot \frac{t}{n}\right)^{n-m}.$$

Границі при $n \rightarrow +\infty$:

Коли $n \rightarrow +\infty$, вираз $\left(1 - \lambda \cdot \frac{t}{n}\right)^{n-m}$ можна наблизити за допомогою експоненти:

$$\left(1 - \lambda \cdot \frac{t}{n}\right)^n \rightarrow e^{-(\lambda \cdot t)}.$$

Тоді:

$$\left(1 - \lambda \cdot \frac{t}{n}\right)^{-m} \rightarrow 1.$$

Таким чином, маємо:

$$P_n(m) = \lim_{n \rightarrow +\infty} \left(C_n^m \cdot \left(\lambda \cdot \frac{t}{n}\right)^m \cdot \left(1 - \lambda \cdot \frac{t}{n}\right)^{n-m} \right).$$

Комбінаторний вираз

Комбінаторний вираз для біноміального коефіцієнта:

$$C_n^m = \frac{n!}{k!(n-k)!}.$$

Коли $n \rightarrow +\infty$, ми можемо використовувати:

$$C_n^m \approx \frac{n^k}{k!}.$$

Тоді:

$$P_n(m) \approx \frac{(\lambda \cdot t)^m \cdot e^{-(\lambda \cdot t)}}{m!},$$

звідки підставивши $t = 1$, отримуємо (4.7).

Формулу Пуассона застосовують для подій, що рідко трапляються, таблиця значень якої наведена на рис 4.5.

m	λ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1638	0.2222	0.2681	0.3033	0.3293	0.3476	0.3596	0.3696
2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	-	-	0.0002	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	-	-	-	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	-	-	-	-	-	-	0.0001	0.0002	0.0003

Рис. 4.5. Таблиця значень $P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!}$

m	λ								
	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0
0	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009	0.0003	0.0001
1	0.3679	0.2707	0.1494	0.0733	0.0337	0.0149	0.0064	0.0027	0.0011
2	0.1839	0.2707	0.2240	0.1465	0.0842	0.0446	0.0223	0.0107	0.0055
3	0.0313	0.1804	0.2240	0.1954	0.1404	0.0892	0.0521	0.0286	0.0150
4	0.0153	0.0902	0.1680	0.1954	0.1755	0.1339	0.0912	0.0572	0.0337
5	0.0081	0.0361	0.1008	0.1563	0.1755	0.1606	0.1277	0.0916	0.0607
6	0.0005	0.0120	0.0504	0.1042	0.1462	0.1606	0.1490	0.1221	0.0911
7	0.0001	0.0034	0.0216	0.0595	0.1044	0.1377	0.1490	0.1396	0.1318
8	-	0.0009	0.0081	0.0298	0.0655	0.1033	0.1304	0.1396	0.1318
9	-	0.0002	0.0027	0.0132	0.0363	0.0688	0.1014	0.1241	0.1318
10	-	-	0.0008	0.0053	0.0181	0.0413	0.0710	0.0993	0.1180
11	-	-	0.0002	0.0019	0.0034	0.0225	0.0452	0.0722	0.0970
12	-	-	0.0001	0.0006	0.0013	0.0113	0.0264	0.0481	0.0728
13	-	-	-	0.0002	0.0005	0.0052	0.0142	0.0296	0.0504
14	-	-	-	0.0001	0.0002	0.0022	0.0071	0.0169	0.0324
15	-	-	-	-	-	0.0009	0.0033	0.0090	0.0194
16	-	-	-	-	-	0.0003	0.0014	0.0045	0.0109
17	-	-	-	-	-	0.0001	0.0006	0.0021	0.0058
18	-	-	-	-	-	-	0.0002	0.0009	0.0029
19	-	-	-	-	-	-	0.0001	0.0004	0.0014
20	-	-	-	-	-	-	-	0.0002	0.0006
21	-	-	-	-	-	-	-	0.0001	0.0003
22	-	-	-	-	-	-	-	-	0.0001

Рис. 4.6. Таблиця значень $P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!}$ (продовження)

Приклад 4.11. База отримує 10000 стандартних виробів. Транспортування цих виробів приводить до пошкодження у середньому 0,02% виробів. Знайти ймовірність того, що з 10000 виробів: **1.** буде пошкоджено: **а)** 3 вироби; **б)** хоча б 3 вироби; **2.** не буде пошкоджено: **а)** 9997 виробів; **б)** хоча б 9997 виробів.

Розв’язання. 1. а) Ймовірність того, що транспортування пошкодить виріб, дорівнює 0,0002. Оскільки p – мале, n – велике і $\lambda = n \cdot p = 10000 \cdot 0,0002 = 2 \leq 10$, ми можемо застосувати формулу Пуассона:

$$P_{10000}(3) = \frac{2^3 \cdot e^{-2}}{3!}.$$

Використовуючи таблицю рис. 4.5, отримаємо: $P_{10000}(3) = 0,18045$.

б) Ймовірність $P_{10000}(m \geq 3)$ знайдемо так:

$$\begin{aligned} P_{10000}(m \geq 3) &= 1 - P_{10000}(m < 3) \\ &= 1 - (P_{10000}(0) + P_{10000}(1) + P_{10000}(2)) \\ &= 1 - (0,13534 + 0,27067 + 0,27067) = 0,32332. \end{aligned}$$

2. а) У даному випадку $p = 1 - 0,0002 = 0,9998$. Потрібно знайти $P_{10000}(9997)$.

Оскільки ймовірність p велика, ми не можемо застосувати формулу Пуассона. Крім того, оскільки $n \cdot p \cdot q \approx 2 < 10$, ми не можемо скористатись інтегральною формулою Муавра – Лапласа. Але подія {не пошкоджено 9997 виробів з 10000} рівносильна події {пошкоджено 3 вироби з 10000}, ймовірність якої ми вже знайшли. Тому $P_{10000}(9997) = 0,18045$.

б) Подія {не пошкоджено хоча б 9997 з 10000 виробів} рівносильна події {пошкоджено не більше, ніж 3 вироби з 10000}, для якої $p = 0,0002$. Тому:

$$P_{10000}(m \leq 3) = P_{10000}(0) + P_{10000}(1) + P_{10000}(2) + P_{10000}(3) = 0,13534 + 0,27067 + 0,27067 + 0,18045 = 0,85713.$$

Відповідь. 1. а) $P_{10000}(3) = 0,18045$; б) $P_{10000}(m \geq 3) = 0,32332$; 2. а) $P_{10000}(9997) = 0,18045$; б) $P_{10000}(m \leq 3) = 0,85713$.

4.5. Поліноміальна схема

В *поліноміальній* (мультиноміальній) схемі здійснюється перехід від послідовності незалежних випробувань з двома наслідками (A та $\bar{A}$) до послідовності незалежних випробувань з k наслідками ($A_1, A_2, \dots, A_k$), які виключають одне одного. При цьому в кожному випробуванні події $A_1, A_2, \dots, A_k$ настають з ймовірностями $p_1, p_2, \dots, p_k$ відповідно. Тоді ймовірність $P_n(m_1, m_2, \dots, m_k)$ того, що в серії з n випробувань подія A_1 настане m_1 разів; подія A_2 настане m_2 разів;...; подія A_k настане m_k разів ($m_1 + m_2 + \dots + m_k = n$) визначається за формулою (4.8):

$$P_n(m_1, m_2, \dots, m_k) = \frac{n!}{m_1! \cdot m_2! \cdot \dots \cdot m_k!} \cdot p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_k^{m_k}. \quad (4.8)$$

Формула (4.8) отримано з урахуванням того, що подія $A = \{\text{в серії з } n \text{ випробувань подія } A_1 \text{ настане } m_1 \text{ разів; подія } A_2 \text{ настане } m_2 \text{ разів;...; подія } A_k \text{ настане } m_k \text{ разів}\}, (m_1 + m_2 + \dots + m_k = n)$ можна представити, як суму несумісних подій, ймовірності яких за теоремою про множення незалежних подій дорівнюють $p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_k^{m_k}$ ($p_1 + p_2 + \dots + p_k = 1$).

В частинному випадку $m_1 = m; m_2 = n - m; p_1 = p; p_2 = q$. Таким чином, отримуємо формулу Бернуллі (4.1).

Приклад 4.12. В урні міститься 10 кульок: 5 червоних, 3 зелених та 2 сині. З урни навмання беруть 5 кульок. Яка ймовірність того, що серед них будуть: 2 червоних, 2 зелених та 2 синя кульки?

Розв'язання. З умови задачі отримуємо: $m_1 = 2; m_2 = 2; m_3 = 1; n = 5; p_1 = 0,5; p_2 = 0,3; p_3 = 0,2$.

За формулою (4.8):

$$P_5(2, 2, 1) = \frac{5!}{2! \cdot 2! \cdot 1!} \cdot 0,5^2 \cdot 0,3^2 \cdot 0,2^1 = 0,135.$$

Відповідь. $P_5(2, 2, 1) = 0,135$.

ЛЕКЦІЯ 5

КОМПОЗИЦІЯ НЕЗАЛЕЖНИХ ВИПРОБУВАНЬ. ЕЛЕМЕНТАРНІ ФАКТИ З ТЕОРІЇ ІНФОРМАЦІЇ

5.1. Композиція двох випробувань

Композиція подій в теорії ймовірностей стосується операцій над подіями, які допомагають визначати нові події на основі вже існуючих.

Нехай (Ω_1, F_1, P_1) та (Ω_2, F_2, P_2) – ймовірнісні простори, що відповідають цим випробуванням, $\Omega_1 = \{E_1, E_2, \dots, E_{m_1}\}$, $\Omega_2 = \{Q_1, Q_2, \dots, Q_{m_2}\}$, F_k ($k = 1, 2$) складається з усіх підмножин множини Ω_k , а ймовірності складних подій $C_1 \subseteq \Omega_1$ і $C_2 \subseteq \Omega_2$ у випробуваннях

$$P_1(C_1) = \sum_{E_i \in C_1} P_1(E_i); P_2(C_2) = \sum_{Q_j \in C_2} P_2(Q_j);$$

задаються за допомогою ймовірностей елементарних подій $P_1(E_i)$, $i = 1, \dots, m_1$, та $P_2(Q_j)$, $j = 1, \dots, m_2$.

Зауважимо, що $P_1(\cdot)$ і $P_2(\cdot)$ – взагалі кажучи, різні ймовірнісні міри.

Композицією двох випробувань зветься складне випробування, що полягає у проведенні першого та другого випробування, елементарними подіями якого є всілякі можливі пари виду (E_i, Q_j) , $i = 1, \dots, m_1$; $j = 1, \dots, m_2$. Таким чином, простір елементарних подій композицій двох випробувань є прямий добуток просторів Ω_1 і Ω_2 , тобто

$$\Omega = \Omega_1 \times \Omega_2.$$

Приклад 5.1. Перше випробування полягає у підкиданні симетричної монети, $\Omega_1 = \{O, P\}$, друге випробування – підкидання грального кубика, $\Omega_2 = \{Q_1, Q_2, \dots, Q_6\}$. Побудувати простір композиції цих випробувань.

Розв’язання. Простір елементарних подій композиції цих двох випробувань –

$$\Omega = \{(O, Q_1), \dots, (O, Q_6), (P, Q_1), \dots, (P, Q_6)\}.$$

Як складні події для композиції розглядаються усі можливі підмножини множини Ω тобто усілякі теоретико-множинні суми пар виду (E_i, Q_j) (та неможлива подія $\emptyset$). Таким чином Ω і F для композиції побудовані.

Для повного завдання ймовірнісного простору композиції необхідно ще означити ймовірнісну міру $P(\cdot)$ на F . Але, взагалі кажучи, тільки позначенням $P_1(\cdot)$ і $P_2(\cdot)$ це зробити неможливо.

Відповідь. Композицію випробувань побудовано.

Розглянемо окремий випадок композиції.

Два випробування називаються **незалежними**, якщо для всіх пар ймовірність $P(\cdot)$ у композиційному просторі визначається співвідношенням (5.1):

$$P(E_i, Q_j) = P_1(E_i) \cdot P_2(Q_j). \quad (5.1)$$

Випробування *незалежні*, якщо неоднозначність результату кожного випробування визначена незв'язними між собою групами випадкових факторів, результат одного випробування не впливає на результат другого.

Ймовірність складної події C для композиції незалежних випробувань визначається з співвідношення (5.2):

$$P(C) = \sum_{(E_i, Q_j) \in C} P(E_i, Q_j) = \sum_{(E_i, Q_j) \in C} (P_1(E_i) \cdot P_2(Q_j)). \quad (5.2)$$

Ймовірність P , що так побудована, називається *прямим добутком* ймовірностей P_1 та P_2 та визначається з рівності (5.3):

$$P = P_1 \cdot P_2. \quad (5.3)$$

Загальна структура незалежних подій у композиційному просторі (Ω, F, P) має наступний вигляд.

Розглянемо подію $A \in F$

$$A = \bigcup_{l=1}^{s_1} \bigcup_{j=1}^{m_2} (E_l, Q_j).$$

Тобто внаслідок першого випробування відбувається одна з подій $E_{i_1}, E_{i_2}, \dots, E_{i_{s_1}}$, $s_1 \leq m_1$ (подія $A_1 = \bigcup_{l=1}^{s_1} E_{i_l} \in F_1$), а внаслідок другого випробування відбувається усе що завгодно, тобто одна з подій $Q_1, Q_2, \dots, Q_{m_2}$ (ймовірна подія $\Omega_2 = \bigcup_{j=1}^{m_2} Q_j$).

Нехай складна подія $B \in F$

$$B = \bigcup_{i=1}^{m_1} \bigcup_{t=1}^{s_2} (E_i, Q_t),$$

тобто внаслідок першого випробування відбувається усе що завгодно, тобто одна з подій $E_1, E_2, \dots, E_{m_2}$ (вірогідна подія $\Omega_1 = \bigcup_{i=1}^{m_1} E_i$), а внаслідок другого випробування відбувається одна з подій $Q_{j_1}, Q_{j_2}, \dots, Q_{j_{s_2}}$, $s_2 \leq m_2$ (подія $B_2 = \bigcup_{t=1}^{s_2} Q_{i_t} \in F_2$).

Можна коротко записати

$$A = A_1 \times \Omega_2, B = \Omega_1 \times B_2.$$

Переконаємось, що події A і B незалежні:

$$P(A \cap B) = P(A) \cdot P(B).$$

Дійсно,

$$\begin{aligned} P(A) &= \sum_{l=1}^{s_1} \sum_{j=1}^{m_2} P(E_{i_l}, Q_j) = \sum_{l=1}^{s_1} \sum_{j=1}^{m_2} (P_1(E_{i_l}) \cdot P_2(Q_j)) = \sum_{l=1}^{s_1} P_1(E_{i_l}) \cdot \\ &\quad \cdot \sum_{j=1}^{m_2} P_2(Q_j) = \sum_{l=1}^{s_1} P_1(E_{i_l}) = P_1(A_1). \end{aligned}$$

Аналогічно можна переконатись, що

$$P(B) = \sum_{t=1}^{s_2} P_2(Q_{j_t}) = P_2(B_2).$$

Очевидно,

$$A \cap B = \sum_{l=1}^{s_1} \sum_{t=1}^{s_2} (E_{il}, Q_{jt}) = A_1 \times B_2.$$

Тоді

$$P(A \cap B) = \sum_{l=1}^{s_1} \sum_{t=1}^{s_2} P(E_{il}, Q_{jt}) = \sum_{l=1}^{s_1} \sum_{t=1}^{s_2} (P(E_{il}) \cdot P(Q_{jt})) = \sum_{l=1}^{s_1} P(E_{il}) \cdot \sum_{t=1}^{s_2} P(Q_{jt}) = P(A) \cdot P(B) = P_1(A_1) \cdot P_2(B_2),$$

що і треба було довести.

Застосування композиції подій

1. Визначення складних подій: Композиція допомагає створювати складніші події на основі простіших подій. Наприклад, у випадкових експериментах ми можемо цікавитися як комбінаціями успіхів, так і їхніми взаємними виключеннями.

2. Моделювання систем: У багатьох задачах ймовірності, наприклад в аналізі надійності систем, використовують композицію подій для визначення ймовірності відмови системи (коли відмовляє один чи більше компонентів).

3. Умовні ймовірності: Часто цікаво знати ймовірність однієї події за умови, що сталася інша подія. Це також можна розглядати як композицію подій, де друга подія є умовою для першої.

Застосування в реальному житті

1. Аналіз ризиків: Композиції подій використовуються для обчислення ймовірностей різних сценаріїв ризиків, наприклад, об'єднання подій може показати ймовірність виникнення катастрофічного випадку при настанні кількох ризиків одночасно.

2. Теорія ігор: Композиції подій застосовуються для моделювання стратегій та визначення ймовірностей різних результатів гри.

5.2. Композиція n незалежних випробувань

Зовсім аналогічно будується схема композиції n незалежних випробувань.

Нехай здійснюється n випробувань із скінченими просторами елементарних подій

$$\Omega_i = \{E_1^i, E_2^i, \dots, E_{m_i}^i\}, i = \overline{1, n}.$$

Складними подіями у кожному i -му випробуванні є всілякі підмножини множини Ω_i , що визначає F_i . Задані на кожному F_i ймовірнісні міри $P_i(\cdot)$. І, таким чином, маємо n ймовірнісних просторів (Ω_i, F_i, P_i) , $i = \overline{1, n}$.

Результат спільного проведення n випробувань вважаємо результатом проведення одного складного випробування (**композиції**).

Композиційний простір Ω будуюмо як прямий добуток просторів Ω_i , $i = \overline{1, n}$,

$$\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n,$$

тобто у якості елементарних подій композиції беруться всілякі n -ки виду

$$(E_{j_1}^1, E_{j_2}^2, \dots, E_{j_n}^n); j_1 = \overline{1, m_1}; \dots; j_n = \overline{1, m_n}.$$

Складними подіями в композиційному просторі є теоретико-множинні суми вказаних n -ок.

Випробування називаються **незалежними**, якщо ймовірність кожної n -ки визначається співвідношенням (5.4):

$$P(E_{j_1}^1, E_{j_2}^2, \dots, E_{j_n}^n) = \prod_{i=1}^n P_i(E_{j_i}^i). \quad (5.4)$$

Ймовірності складних подій в композиційному просторі визначаються звичним для дискретних просторів способом – як сума ймовірностей n -ок, що належать даній складній події.

Побудована таким чином ймовірність P для подій композиційного простору Ω називається **прямим добутком ймовірностей** P_i і визначається з рівності (5.5):

$$P = P_1 \cdot P_2 \cdot \dots \cdot P_n. \quad (5.5)$$

В побудованому просторі (Ω, F, P) виділимо клас подій A , що називаються **прямокутниками**. Нехай $A_i \in F_i, i = \overline{1, n}$. Прямокутник (5.6):

$$A = A_1 \times A_2 \times \dots \times A_n \quad (5.6)$$

містить ті та тільки ті n -ки $E = (E_{j_1}^1, E_{j_2}^2, \dots, E_{j_n}^n)$, у яких $E_{j_i}^i \in A_i, i = \overline{1, n}$.

З означення ймовірності для композиції незалежних випробувань виходить, що ймовірність прямокутника A дорівнює:

$$\begin{aligned} P(A) &= \sum_{E \in A} P(E) = \sum_{E_{j_1}^1 \in A_1} \dots \sum_{E_{j_n}^n \in A_n} (P_1(E_{j_1}^1) \cdot \dots \cdot P_n(E_{j_n}^n)) \\ &= \sum_{E_{j_1}^1 \in A_1} P_1(E_{j_1}^1) \cdot \dots \cdot \sum_{E_{j_n}^n \in A_n} P_n(E_{j_n}^n) = P_1(A_1) \cdot P_2(A_2) \cdot \dots \cdot P_n(A_n). \end{aligned}$$

Загальна структура незалежних подій в композиційному просторі задається прямокутниками виду

$$A'_i = \Omega_1 \times \Omega_2 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n,$$

тобто подіями, які можна інтерпретувати наступним чином: в i -му випробуванні відбувається подія A_i , а в інших випробуваннях відбувається що завгодно.

З одержаного раніш співвідношення для ймовірності прямокутника маємо

$$P(A'_i) = P_i(A_i).$$

Тому що $A = A_1 \times A_2 \times \dots \times A_n = \bigcap_{i=1}^n A'_i$, то з того ж співвідношення виходить

$$P(\bigcap_{i=1}^n A'_i) = \prod_{i=1}^n P(A'_i),$$

тобто при різних i події A'_i незалежні.

Через те що у прямокутниках A'_i в якості A_i можуть фігурувати Ω_i , то ясно, що мова йде про незалежність у сукупності.

5.3. Елементарні факти з теорії інформації

Теорія інформації — це розділ математики, який досліджує процеси зберігання, перетворення і передачі інформації. Теорія інформації тісно

пов'язана з такими розділами математики як теорія ймовірностей і математична статистика. Вона пов'язана з інформаційною ентропією, комунікаційними системами, криптографією, корекцією помилок і іншими важливими областями.

Основні розділи теорії інформації — кодування джерела (стискаюче кодування) і каналне (завадостійке) кодування. Теорія інформації тісно пов'язана з інформаційною ентропією, комунікаційними системами, криптографією і іншими суміжними дисциплінами.

Аксіоми теорії інформації

- Інформація є лише там, де функціонують пристрої керування,
- Інформація зберігається і передається тільки на матеріальному носії,
- Інформація має ідеальний характер,
- Інформація має різні форми.

Базові закони теорії інформації

Закон 1: на отримання інформації будь-яка кібернетична система витрачає не менше деякої мінімальної кількості енергії.

Закон 2: кількість інформації, яку отримує кібернетична система в процесі розпізнавання після прийняття певного сигналу, дорівнює логарифму при основі m від кількості варіантів вибору, що передували розпізнаванню.

Закон 3: що меншою є ймовірність завершення якогось випробування з певним результатом, то більше інформації для будь-якої кібернетичної системи несе саме цей результат, і навпаки.

Закон 4: будь-які сигнали, отримані кібернетичною системою, впливають на цю систему.

Властивості інформації

Адекватність — відповідність відображуваному об'єктові або реальному стану об'єктивної дійсності при відсутності прихованих помилок у такій інформації.

Повнота — властивість інформації, що дозволяє характеризувати об'єкт вичерпним для споживача засобом, що й надає можливість ухвалювати на основі такої інформації управлінські рішення.

Релевантність — відповідність потребам споживача, що характеризує, наскільки інформація сприяє досягненню поставлених перед споживачем цілей і завдань.

Доступність — можливість одержання будь-яким споживачем.

Актуальність — відповідність інформації теперішньому моменту часу.

Коректність — властивість, що полягає в такому її зображенні, щоб інформація однозначно сприймалася всіма її споживачами.

Захищеність — неможливість несанкціонованого доступу й цілеспрямованого спотворення інформації.

Ергономічність — достатність обсягу і форми інформації для даного споживача.

Власна інформація

Математична теорія інформації ґрунтується на теорії ймовірності й статистиці, і вимірює інформацію за допомогою декількох кількостей інформації (quantities of information). Застосовувану в наступних формулах одиницю інформаційної ентропії визначає вибір логарифмічної основи. Найзвичнішою одиницею інформації є *біт*, що ґрунтується на двійковому логарифмі. До інших одиниць належать *нат*, що ґрунтується на натуральному логарифмі, та *гартлі*, що ґрунтується на десятьковому логарифмі. Надалі вираз $p \cdot \log(p)$ ($p \cdot \log_2(p)$; $p \cdot \lg(p)$ або $p \cdot \ln(p)$), коли p є нулем вважається рівним нулю. Це є виправданим, оскільки для будь-якої логарифмічної основи $\lim_{p \rightarrow 0+} (p \cdot \log(p)) = 0$.

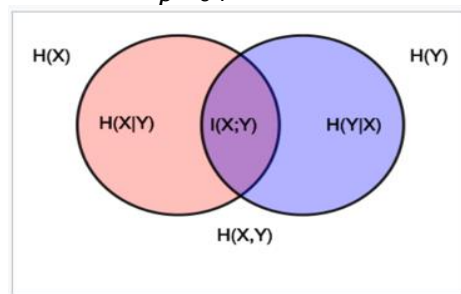


Рис. 5.1. Інформаційна діаграма

Міра інформаційного вмісту повідомлення, називається *власною інформацією* (*self-information*) або «*несподіваністю*»

Кількість власної інформації, яка міститься в ймовірнісній події, залежить лише від ймовірності цієї події: що меншою є її ймовірність, то більшою є власна інформація, пов'язана з отриманням інформації про те, що ця подія дійсно відбулася.

Міра власної інформації є додатною та адитивною. Якщо подія C є перетином двох незалежних подій A та B , то кількість інформації при оголошенні про те, що подія C сталася, дорівнює сумі кількостей інформації в оголошеннях про подію A та подію B відповідно:

$$I(A \cap B) = I(A \cdot B) = I(A) + I(B).$$

Із врахуванням цих властивостей, власною інформацією $I(x_n)$, пов'язаною з подією (повідомленням) x_n з ймовірністю $p(x_n)$, називається величина, що обчислюється з рівності (5.7):

$$I(x_n) = \log\left(\frac{1}{p(x_n)}\right) = -\log(p(x_n)). \quad (5.7)$$

Це визначення відповідає наведеним вище умовам. У формулі (5.4) не вказано основу логарифма: при застосуванні основи 2 одиницями $I(x_n)$ будуть біти, при застосуванні логарифма за основою e одиницею буде нат, для логарифма за основою 10 одиницею буде гартлі.

Як швидко пояснення, кількістю інформації, пов'язаною з випадінням 4 аверсів (або будь-якого конкретного виходу) в 4 послідовних підкиданнях монети, буде 4 біти (ймовірність $\frac{1}{16}$), а кількістю інформації, пов'язаною з

отриманням результату, відмінного від вказаного, буде 0,09 біт (ймовірність $\frac{15}{16}$).

Приклад 5.2. При підкиданні монети шансом «реверсу» є 0,5. Обчислити кількість інформації.

Розв'язання. Коли проголошується, що справді випав «реверс», то це дає кількість інформації

$$I(\text{«реверс»}) = \log_2 \left(\frac{1}{0,5} \right) = \log_2(2) = 1 \text{ біт інформації.}$$

Відповідь. $I(\text{«реверс»}) = 1$ біт.

Приклад 5.3. При викиданні правильного грального кубика ймовірність «четвірки» становить $\frac{1}{6}$. Обчислити кількість інформації.

Розв'язання. Коли проголошується, що випала «четвірка», то кількістю власної інформації є

$$I(\text{«четвірка»}) = \log_2 \left(\frac{1}{\frac{1}{6}} \right) = \log_2(6) = 2,585 \text{ бітів власної інформації.}$$

Відповідь. $I(\text{«четвірка»}) = 2,585$ бітів.

Приклад 5.4. При незалежному викиданні двох гральних кубиків обчислити кількість інформації, пов'язаної з подією {викидання 1 = «два» і викидання 2 = «чотири»}.

Розв'язання. При незалежному викиданні двох гральних кубиків кількість інформації, пов'язаної з {викидання 1 = «два» і викидання 2 = «чотири»}, дорівнює

$$\begin{aligned} I(\text{«викиданням 1 є два і викиданням 2 є чотири»}) = \\ \log_2 \left(\frac{1}{P(\text{викидання 1=«два» і викидання 2=«чотири»})} \right) = \log_2 \left(\frac{1}{\frac{1}{36}} \right) = \log_2(36) = \\ 5,170 \text{ бітів.} \end{aligned}$$

Також шукана кількість інформації дорівнює сумі окремих кількостей власної інформації, пов'язаних із {викидання 1 = «два»} і {викидання 2 = «чотири»}; а саме, $2,585 + 2,585 = 5,170$ бітів.

Відповідь. $(I\text{«викиданням 1 є два і викиданням 2 є чотири»}) = 5,170$ бітів.

Приклад 5.5. При незалежному викиданні двох гральних кубиків обчислити кількість інформації, пов'язаної з подією {Сумою двох гральних кубиків є п'ять}.

Розв'язання. В тій самій ситуації з двома гральними кубиками ми можемо розглядати інформацію, присутню в твердженні {Сумою двох гральних кубиків є п'ять}

$$\begin{aligned} I(\text{«Сумою викидів 1 та 2 є п'ять»}) = \log_2 \left(\frac{1}{P(\text{викиди 1 та 2 дають у сумі п'ять})} \right) = \\ \log_2 \left(\frac{1}{\frac{4}{36}} \right) = \log_2(9) = 3,17 \text{ бітів.} \end{aligned}$$

$P(\text{викиди 1 та 2 дають у сумі п'ять}) = \frac{4}{36}$, оскільки існує чотири варіанти з 36 можливих, щоби два кубики давали в сумі 5. Це показує, що складніші або неоднозначніші події теж можуть давати інформацію.

Відповідь. $I(\text{«Сумою викидів 1 та 2 є п'ять»}) = 3,17$ бітів.

Інформаційна ентропія

Розглянемо два ймовірнісних простори із скінченною кількістю елементарних подій (кожний задамо таблицею, у якій у верхньому рядку вказані елементарні події, у нижньому – ймовірності цих подій):

$$\left\{ \begin{matrix} x_1 & x_2 & x_3 & x_4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{matrix} \right\}, \left\{ \begin{matrix} x'_1 & x'_2 & x'_3 & x'_4 \\ 0,01 & 0,02 & 0,96 & 0,01 \end{matrix} \right\}.$$

Очевидно, що ступінь невизначеності результату випробування (до випробування) у першому ймовірнісному просторі вище, ніж у другому. Дійсно, до випробування у першому просторі ні одній з подій не можна віддати перевагу, а у другому майже напевно відбудеться подія x'_3 .

Виникає питання: чи можна якимось чином міряти ступінь невизначеності, що властива ймовірнісному простору. Першим, хто запропонував вираз для міри невизначеності, був К. Шеннон.

Нехай деякий ймовірнісний простір повідомлень Ω задано таблицею

$$\left\{ \begin{matrix} x_1 & x_2 & x_3 & \dots & x_n \\ p(x_1) & p(x_2) & p(x_3) & \dots & p(x_n) \end{matrix} \right\},$$

де:

$x_i (i = \overline{1, n})$ — повідомлення з цього простору;

$p(x_i)$ — ймовірність кожного можливого значення x_i , причому $\sum_{i=1}^n p(x_i) = 1$.

Таким чином, Ω може приймати лише окремі, чітко визначені значення, і кожне з цих значень має певну ймовірність, тому її надалі можна вважати *дискретною випадковою величиною*, яка полягає у отриманні якогось з повідомлень $x_i (i = \overline{1, n})$ або символів одного повідомлення з ймовірністю $p(x_i)$. Позначимо $\Omega = X$.

Для вимірювання ступеню невизначеності, що притаманна даному простору використовується інформаційна ентропія.

Інформаційна ентропія випадкової події — це математичне сподівання (середнє очікуване значення) її власної інформації (5.8):

$$H(X) = M(I(X)), \quad (5.8)$$

тобто, для дискретної випадкової величини X з ймовірностями $p(x_i)$ вона визначається з рівності (5.9):

$$H(X) = - \sum_{i=1}^n (p(x_i) \cdot \log_2(p(x_i))). \quad (5.9)$$

Пояснення

1. Інформаційна ентропія враховує всі можливі значення випадкової величини та їх ймовірності. Чим більше значень і чим рівномірніше розподілені ймовірності, тим вища ентропія.

2. Використання логарифма в основі 2 означає, що ентропія вимірюється в бітах. Це зручно для цифрових систем і кодування інформації.

3. Від'ємний знак гарантує, що значення ентропії є додатним або нульовим, оскільки ймовірності лежать в інтервалі від 0 до 1.

Приклад 5.6. Підкидається монета. Обчислити інформаційну ентропію.

Розв'язання. Якщо підкидається монета (ймовірності орла і решки рівні 0,5) обчислити інформаційну ентропію, ентропія дорівнює 1 біту, оскільки обидва результати є однаково ймовірними:

$$H(X) = -(0,5 \cdot \log_2(0,5) + 0,5 \cdot \log_2(0,5)) = 1 \text{ біт.}$$

Відповідь. $H(X) = 1$ біт.

Приклад 5.7. Підкидається шестигранний кубик. Обчислити інформаційну ентропію.

Розв'язання. Якщо підкидається шестигранний кубик (ймовірність кожного результату дорівнює $\frac{1}{6}$), ентропія дорівнює:

$$H(X) = -6 \cdot \left(\frac{1}{6} \cdot \log_2 \left(\frac{1}{6} \right) \right) \approx 2,585 \text{ бітів.}$$

Відповідь. $H(X) = 2,585$ бітів.

Ентропії притаманні такі простіші **властивості**:

1. Якщо ймовірнісний простір не має невизначеності, тобто якесь $p(x_i) = 1$, а інші дорівнюють нулеві, то ентропія дорівнює нулеві (очевидно);

2. Коли елементарні події рівноймовірні, тобто $p(x_i) = \frac{1}{n}$, $i = \overline{1, n}$, тоді ентропія приймає максимальне значення.

Для **доведення** треба вирішити задачу максимізації $H(X)$ за умови $\sum_{i=1}^n p(x_i) = 1$. Для цього використаємо метод невизначених множників Лагранжа: будемо шукати екстремум функції

$$F = -\sum_{i=1}^n (p(x_i) \cdot \log_2(p(x_i))) + \lambda \cdot (\sum_{i=1}^n (p(x_i)) - 1).$$

Диференціюємо цю функцію по p_i та прирівнюємо частинні похідні до нуля. Одержуємо систему рівнянь

$$-\log_2(p(x_i)) - \log_2(e) + \lambda = 0, \quad i = \overline{1, n}.$$

або

$$\log_2(p(x_i)) = \lambda - \log_2(e), \quad i = \overline{1, n}.$$

Звідси виходить, що екстремум (перевірте, що це максимум) досягається при рівних між собою $p(x_i)$. З умови нормування $\sum_{i=1}^n p(x_i) = 1$ маємо $p_i = \frac{1}{n}$, $i = \overline{1, n}$.

За одиницю невизначеності знову приймається біт.

Застосування

Інформаційна ентропія використовується в багатьох областях, включаючи:

Теорію інформації: Для вимірювання кількості інформації та оптимального кодування повідомлень.

Криптографію: Для оцінки безпеки шифрів та випадковості ключів.

Машинне навчання: Для оцінки невизначеності моделей та побудови дерев прийняття рішень.

Стиснення даних: Для визначення оптимальних алгоритмів стиснення.

Інформаційна ентропія є фундаментальною концепцією в розумінні передачі інформації та її ефективного кодування.

Інформаційна ентропія композиції випробувань

Розглянемо два ймовірнісних простори:

$$\Omega_1 = \begin{Bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ p_1 & p_2 & p_3 & \dots & p_n \end{Bmatrix}, \Omega_2 = \begin{Bmatrix} y_1 & y_2 & y_3 & \dots & y_m \\ q_1 & q_2 & q_3 & \dots & q_m \end{Bmatrix}.$$

Проведемо композицію двох випробувань. Для композиційного простору таблиця має наступний вигляд:

$$\Omega_1 \times \Omega_2 = \left\{ \begin{pmatrix} x_i & y_j \\ P(x_i, y_j) \end{pmatrix} \right\}, i = \overline{1, n}; j = \overline{1, m}.$$

Максимальна ентропія композиційного ймовірнісного простору досягається тоді, коли випробування незалежні (хоча б на основі якісного аналізу). Обчислимо ентропію композиційного простору для випадку незалежних випробувань:

$$\begin{aligned} H(\Omega_1 \times \Omega_2) &= \sum_{i=1}^n \sum_{j=1}^m (p_i \cdot q_j \cdot \log_2(p_i \cdot q_j)) = \\ &= - \sum_{i=1}^{s_1} \sum_{j=1}^{s_2} (p_i \cdot q_j \cdot \log_2(p_i)) - \\ &- \sum_{i=1}^{s_1} \sum_{j=1}^{s_2} (p_i \cdot q_j \cdot \log_2(q_j)) = - \sum_{i=1}^{s_1} (p_i \cdot \log_2(p_i)) - \\ &- \sum_{j=1}^{s_2} (q_j \cdot \log_2(q_j)) = H(\Omega_1) + H(\Omega_2). \end{aligned}$$

Таким чином, для композиції двох випробувань маємо (5.10):

$$0 \leq H(\Omega_1 \times \Omega_2) \leq H(\Omega_1) + H(\Omega_2), \quad (5.10)$$

або, враховуючи, що $\Omega_1 = X$; $\Omega_2 = Y$.

$$0 \leq H(X, Y) \leq H(X) + H(Y),$$

де $H(X, Y)$ – спільна ентропія.

Спільна ентропія

Спільну ентропію (joint entropy) двох дискретних випадкових змінних X та Y визначають як ентропію їхнього спільного розподілу (5.11):

$$H(X, Y) = M(-\log_2(p(x, y))) = - \sum_{x, y} (p(x, y) \cdot \log_2(p(x, y))). \quad (5.11)$$

Якщо X та Y є незалежними, то ця спільна ентропія є просто сумою їхніх окремих ентропій.

Умовна ентропія

За заданого конкретного значення випадкової змінної Y **умовну ентропію X за $Y = y$** визначено з рівності (5.12):

$$H(X/y) = M(-\log_2(p(x/y))) = -\sum_{x \in X} (p(x/y) \cdot \log_2(p(x/y))). \quad (5.12)$$

де $p(x/y) = \frac{p(x,y)}{p(y)}$ є умовною ймовірністю x за заданого y .

Умовну ентропію (conditional entropy) X за заданого Y , що також називають **ухильністю** (equivocation) X від Y , визначаються з рівності (5.13):

$$H(X/Y) = M(H(X/y)) = -\sum_{y \in Y} (p(y) \cdot \sum_{x \in X} (p(x/y) \cdot \log_2(p(x/y)))) = \sum_{x,y} \left(p(x,y) \cdot \log_2 \left(\frac{p(y)}{p(x,y)} \right) \right). \quad (5.13)$$

Вона використовує умовне математичне сподівання з теорії ймовірності.

Базовою властивістю умовної ентропії є те (5.14):

$$H(X/Y) = H(X, Y) - H(Y). \quad (5.14)$$

Приклад 5.8. Визначити інформаційну ентропію повідомлення, ймовірності появи символів якого такі: $p(x_1) = p(x_2) = p(x_3) = p(x_4) = 0,01$; $p(x_5) = 0,96$.

Розв'язання. Складемо випадкову величину X появи символів повідомлення (табл. 5.1):

Таблиця 5.1. Ряд розподілу випадкової величини X

x_i	x_1	x_2	x_3	x_4	x_5
$p(x_i)$	0,01	0,01	0,01	0,01	0,96

Підставляємо значення з табл. 5.1 у формулу (5.9) для обчислення ентропії. Розрахункова таблиця представлена в табл. 5.2.

Таблиця 5.2. Розрахункова таблиця

x_i	x_1	x_2	x_3	x_4	x_5	$H(X)$
$p(x_i)$	0,01	0,01	0,01	0,01	0,96	
$-(p(x_i) \cdot \log_2(p(x_i)))$	0,066439	0,066439	0,066439	0,066439	0,056538	0,322292

Відповідь. $H(X) = 0,322292$ біта.

Приклад 5.9. Визначити ентропію повідомлення «aaaabbcbddd».

Розв'язання. Складемо випадкову величину X появи символів повідомлення (табл. 5.3).

Всього літер – 10; $a - 4$; $b - 2$; $c - 1$; $d - 3$.

Тобто, ймовірності відповідно дорівнюють:

$$p(a) = 0,4; p(b) = 0,2; p(c) = 0,1; p(d) = 0,3.$$

Таблиця 5.3. Ряд розподілу випадкової величини X

x_i	a	b	c	d
$p(x_i)$	0,4	0,2	0,1	0,3

Підставляємо значення з табл. 5.3 у формулу (5.9) для обчислення ентропії. Розрахункова таблиця представлена в табл. 5.4.

Таблиця 5.4. Розрахункова таблиця до прикладу 5.9

x_i	a	b	c	d	$H(X)$
$p(x_i)$	0,4	0,2	0,1	0,3	
$-(p(x_i) \cdot \log_2(p(x_i)))$	0,528771	0,464386	0,332193	0,52109	1,846439

Відповідь. $H(X) = 1,846439$ біта.

Приклад 5.10. Визначити ентропію двійкового джерела інформації, якщо $p(0) = 0,1$, а $p(1) = 0,9$.

Розв'язання. Джерело представляло собою булеву функцію від однієї змінної.

Складемо випадкову величину X появи символів повідомлення (табл. 5.5).

Таблиця 5.5. Ряд розподілу випадкової величини X

x_i	0	1
$p(x_i)$	0,1	0,9

Підставляємо значення з табл. 5.5 у формулу (5.9) для обчислення ентропії. Розрахункова таблиця представлена в табл. 5.6.

Таблиця 5.6. Розрахункова таблиця

x_i	0	1	$H(X)$
$p(x_i)$	0,1	0,9	
$-(p(x_i) \cdot \log_2(p(x_i)))$	0,332193	0,136803	0,468996

Відповідь. $H(X) = 0,468996$ біта.

Приклад 5.11. Визначити ентропію джерела інформації, якщо $p(00) = 0,01$; $p(01) = 0,09$; $p(10) = 0,09$; $p(11) = 0,81$.

Розв'язання. Дане джерело інформації математично представляє собою булеву функцію від двох змінних.

Складемо випадкову величину X появи символів повідомлення (табл. 5.7).

Таблиця 5.7. Ряд розподілу випадкової величини X

x_i	00	01	10	11
$p(x_i)$	0,01	0,09	0,09	0,81

Підставляємо значення з табл. 5.7 у формулу (5.9) для обчислення ентропії. Розрахункова таблиця представлена в табл. 5.8.

Таблиця 5.8. Розрахункова таблиця

x_i	00	01	10	11	$H(X)$
$p(x_i)$	0,01	0,09	0,09	0,81	
$-(p(x_i) \cdot \log_2(p(x_i)))$	0,066439	0,312654	0,312654	0,246245	0,937991

Відповідь. $H(X) = 0,937991$ біта.

Приклад 5.12. Для двійкового джерела інформації визначити ймовірності появи «0» та «1», при яких ентропія максимальна.

Розв'язання. Складемо випадкову величину X появи символів повідомлення (табл. 5.9).

Таблиця 5.9. Ряд розподілу випадкової величини X

x_i	0	1
$p(x_i)$	p	$1 - p$

Підставляємо значення з табл. 5.9 у формулу (5.9) для обчислення ентропії.

$$H(X) = -p \cdot \log_2(p) - (1 - p) \cdot \log_2(1 - p).$$

Диференціюємо функцію по p .

$$\frac{dH(X)}{dp} = -\log_2(p) - \frac{1}{\ln(2)} + \log_2(1 - p) + \frac{1}{\ln(2)} = \log_2(1 - p) - \log_2(p).$$

Прирівнюємо до нуля.

$$\log_2(1 - p) - \log_2(p) = 0;$$

$$\log_2(1 - p) = \log_2(p);$$

$$1 - p = p;$$

$$p = 0,5.$$

Перевірка:

$$\frac{d^2H(X)}{dp^2} = -\frac{1}{(1-p) \cdot \ln(2)} - \frac{1}{p \cdot \ln(2)}.$$

Підставляємо $p = 0,5$ і бачимо, що $\frac{d^2H(X)}{dp^2} = -\frac{4}{\ln(2)} < 0$, отже, при $p = 0,5$ ентропія максимальна.

Відповідь. $p = 0,5$.

Приклад 5.13. Визначити ентропії $H(X)$, $H(Y)$, $H(X, Y)$, $H(X/Y)$, якщо задана матриця станів системи, яка об'єднує джерела X, Y :

$$P(X \cdot Y) = \begin{pmatrix} 0,3 & 0,1 & 0 \\ 0 & 0,2 & 0,2 \\ 0 & 0 & 0,2 \end{pmatrix}.$$

Розв'язання. Складемо випадкову величину Y появи символів повідомлення (табл. 5.10).

Таблиця 5.10. Ряд розподілу випадкової величини Y

y_i	y_1	y_2	y_3
$p(y_i)$	0,3	0,3	0,4

Підставляємо значення з табл. 5.10 у формулу (5.9) для обчислення ентропії.

Таблиця 5.11. Розрахункова таблиця 1

y_i	y_1	y_2	y_3	$H(Y)$
$p(y_i)$	0,3	0,3	0,4	

$-(p(y_i) \cdot \log_2(p(y_i)))$	0,5210897	0,5210897	0,5287712	1,57095059

Отже, $H(Y) = 1,57095059$ біта.

Аналогічно отримуємо

Таблиця 5.12. Ряд розподілу випадкової величини X

x_j	x_1	x_2	x_3
$p(x_j)$	0,4	0,4	0,2

Таблиця 5.13. Розрахункова таблиця 2

x_j	x_1	x_2	x_3	$H(X)$
$p(x_j)$	0,4	0,4	0,2	
$-(p(x_j) \cdot \log_2(p(x_j)))$	0,5287712	0,5287712	0,4643856	1,52192809

Отже, $H(X) = 1,52192809$ біта.

$H(X \cdot Y)$ обчислюємо за формулою (5.11).

Таблиця 5.14. Розрахункова таблиця 3

	y_1	y_2	y_2	$-(p(x_j, y_i) \cdot \log_2(p(x_j, y_i)))$			$H(X, Y)$
x_1	0,3	0,1	0	0,5210897	0,3321928	0	
x_2	0	0,2	0,2	0	0,4643856	0,4643856	
x_3	0	0	0,2	0	0	0,4643856	
				0,5210897	0,7965784	0,9287712	2,24643934

Отже, $H(X, Y) = 2,24643934$ біт.

З формули (5.14) отримаємо, що

$$H(X/Y) = 2,24643934 - 1,57095059 = 0,67558875 \text{ біта.}$$

Відповідь. $H(X) = 1,52192809$ біта; $H(Y) = 1,57095059$ біта; $H(X, Y) = 2,24643934$ біта; $H(X/Y) = 0,67558875$ біта.

5.4. Алгоритм формування комплексних послуг на основі визначення інформаційної ентропії

Для початку введемо основні означення.

Електронна публічна послуга – це послуга, яка надається громадянам, підприємствам та іншим організаціям через інформаційно-комунікаційні технології (ІКТ), такі як інтернет та інші електронні засоби зв'язку. Ці послуги забезпечують доступ до державних ресурсів, дозволяють подавати заяви, отримувати документи, здійснювати платежі тощо, без необхідності відвідувати державні установи особисто.

Комплексна електронна публічна послуга – електронна публічна послуга, в результаті якої суб'єкту звернення на підставі однієї заяви (звернення, запиту) на отримання декількох електронних публічних послуг одним або декількома суб'єктами надання публічних (електронних публічних) послуг або в автоматичному режимі програмними засобами інформаційно-телекомунікаційних систем (у тому числі з використанням

Єдиного державного веб-порталу електронних послуг) надаються декілька електронних публічних послуг.

На теперішній момент формування комплексних послуг в Україні здійснюється за принципом об'єднання за життєвою подією, або ситуацією. На відміну від комплексних послуг, під *портфелем послуг* будемо розуміти набір послуг, які об'єднуються на основі іншого принципу (окрім життєвої події або ситуації).

Ідея алгоритму [26] полягає у тому, що для кожної пари послуг визначається кількість інформації, яку одна послуга несе для іншої. Чим більше інформації, тим більше пов'язані послуги. Тобто вони можуть утворювати комплексну послугу.

1. Нехай є дві випадкові події:

$$X = \{\text{Користувач замовив послугу } i\};$$

$$Y = \{\text{Користувач замовив послугу } j\}.$$

2. Рахуємо коефіцієнт взаємозалежності Райського (5.15):

$$K_R = \frac{I(X,Y)}{H(X,Y)}, \quad (5.15)$$

де $I(X,Y)$ – взаємна інформація статистична функція двох випадкових величин, що описує кількість інформації, що міститься в одній випадковій величині відносно іншої (5.16):

$$I(X,Y) = H(X) + H(Y) - H(X,Y), \quad (5.16)$$

де $H(X)$, $H(Y)$, $H(X,Y)$ – інформаційні ентропії, що дорівнюють кількості інформації з послуг X , Y , та $X \cdot Y$ відповідно.

Ентропії та спільна ентропія обчислюються за формулами (5.9) та (5.11) відповідно.

В нашому випадку ймовірності $p(x_i), p(y_j), p(x_i, y_j)$ можна трактувати як відносну частоту замовлення клієнтами послуги i, j та одночасно обох послуг. Дана інформація повинна бути заздалегідь відома.

3. За властивостями ентропії видно, що коефіцієнт взаємозалежності Райського $1 \leq K_R < \infty$ (одиниці, якщо послуги залежні і прямує до нескінченості, якщо незалежні). Таким чином, чим ближче його значення до одиниці (не відрізняється від неї більше ніж на порядок), тим сильніше пов'язані між собою послуги i та j .

Вважатимемо, що якщо $K_R \leq \varepsilon$, то дві послуги i та j утворюють комплексну послугу, якщо більше, то не утворюють.

Взагалі кажучи, значення ε обирається експериментальним шляхом, в залежності від особливостей даних.

4. В разі, якщо маємо три послуги i, j та k , то користуємось правилом транзитивності:

В разі якщо послуги i та j є комплексними та послуги j та k також є комплексними, то послуги i та k теж є комплексними. Таким чином, i, j та k утворюють одну спільну комплексну послугу.

В разі якщо послуги i та j є комплексними та послуги i та k також є комплексними, то з цього не випливає, що послуги j та k теж є комплексними. Таким чином, i, j та k утворюють дві різні комплексні послуги.

Приклад 5.14. Нехай є дві окремі послуги. Встановити, чи можна їх об'єднати в одну комплексну послугу.

Розв'язання. Нехай є дві послуги X :

Таблиця 5.15. Ряд розподілу випадкової величини X

x_i	x_1	x_2	x_3
w_i	0,45	0,15	0,40

де:

x_i – документ, яку треба надати для отримання послуги X ;

w_i – відносна частота звернень по документ та послуга Y .

Таблиця 5.16. Ряд розподілу випадкової величини Y

y_k	y_1	y_2	y_3
w_k	0,35	0,35	0,30

де:

y_k – документ, яку треба надати для отримання послуги Y ;

w_k – відносна частота звернень по документ.

Одночасно послуга $X \cdot Y$ розглядається, як система двох випадкових величин:

Таблиця 5.17. Ряд розподілу випадкової величини $X \cdot Y$

	x_1	x_2	x_3
y_1	0,15	0,05	0,10
y_2	0,14	0,24	0,02
y_3	0,12	0,15	0,03

$$H(X) = -(0,45 \cdot \log_2(0,45) + 0,15 \cdot \log_2(0,15) + 0,40 \cdot \log_2(0,40)) = \\ = 0,39062395106447 + 0,054805219444381 + 0,30258831894641 = \\ = 0,74801748945526.$$

$$H(Y) = -(0,35 \cdot \log_2(0,35) + 0,35 \cdot \log_2(0,35) + 0,30 \cdot \log_2(0,30)) = \\ = 0,23108820773979 + 0,23108820773979 + 0,17271499274803 = \\ = 0,63489140822761.$$

$$H(X, Y) = -(0,15 \cdot \log_2(0,15) + 0,05 \cdot \log_2(0,05) + 0,10 \cdot \log_2(0,10) + 0,14 \cdot \log_2(0,14) + 0,24 \cdot \log_2(0,24) + 0,02 \cdot \log_2(0,02) + 0,12 \cdot \log_2(0,12) + 0,15 \cdot \log_2(0,15) + 0,03 \cdot \log_2(0,03)) = \\ = 0,054805219444381 + 0,011568910657988 + 0,030102999566398 + 0,049356579386504 + 0,11656745623924 + 0,0035436764027112 + 0,039229869422866 + 0,054805219444381 + 0,0059301503142701 = \\ = 0,36591008087874.$$

$$I(X, Y) = 0,74801748945526 + 0,63489140822761 - 0,36591008087874 = 1,0169988168041.$$

$$K_R = \frac{1,0169988168041}{0,36591008087874} = 2,7793681288085.$$

Отже, оскільки значення коефіцієнту досить близьке до одиниці, можемо вважати, що послуги X, Y утворюють комплексну послугу.

Відповідь. Послуги X, Y утворюють комплексну послугу.

ЛЕКЦІЯ 6

ВИПАДКОВІ ВЕЛИЧИНИ

6.1. Випадкові величини та закони їх розподілу

Означення та види випадкових величин

В практичній діяльності часто зустрічаються експерименти, випробування, досліди, результатами яких є чисельні значення. Наприклад, кількість замовлень на авіаквитки, що надходить до системи бронювання та продажу квитків впродовж часу t , не є сталою величиною і може приймати різні значення $0, 1, 2, \dots, n, \dots$ в залежності від впливу факторів випадкового характеру. Величина «кількість замовлень» відноситься до величин, які називаються випадковими. Вони дають кількісну оцінку результату випробування на відміну від випадкових подій, розглянутих попередньо, які характеризують результат випробування якісно.

Випадковою називається величина, яка в результаті випробування приймає те чи інше можливе значення, заздалегідь невідоме, яке змінюється від випробування до випробування і залежить від ряду випадкових факторів.

Випадкові величини позначаються великими літерами $X, Y, Z, \dots$, а їх можливі значення відповідними малими літерами з індексами. Наприклад, випадкова величина X , її можливі значення $x_1, x_2, \dots, x_n, \dots$

Застосовується також інше означення випадкової величини.

Випадковою величиною називається функція X , означена на множині наслідків $x_1, x_2, \dots, x_n, \dots$ даного випробування.

Наведемо приклади випадкових величин.

1. Кількість електричних ламп, що виходять з ладу в системі освітлення та сигналізації аеропорту на протязі доби, не є сталою і змінюється в залежності від якості ламп, умов експлуатації, рівня напруги в електромережі тощо. Ця випадкова величина має множину можливих значень $\{0, 1, 2, \dots, n, \dots\}$, яка теоретично може бути нескінченною.
2. Рівень напруги в електромережі аеропорту також не є сталою величиною і змінюється в залежності від режиму роботи електростанції, кількості споживачів, системи стабілізації тощо. Ця випадкова величина має множину можливих значень, які суцільно заповнюють деякий інтервал.

Випадкові величини бувають двох видів: дискретні і неперервні.

1. *Дискретні випадкові величини* – величини, які в результаті випробувань приймають окремі, ізольовані можливі значення, множина яких може бути скінченною або нескінченною. Можливі значення дискретної величини зображуються точками числової осі. Прикладами дискретних випадкових величин є кількість літаків в зоні диспетчера по керуванню повітряним рухом, кількість пасажирів на рейсі, кількість квитків, виданих на протязі зміни по запитам пасажирів системою продажу авіаквитків, число вузлів системи, які вийшли з ладу впродовж певного часу тощо.

2. *Неперервні випадкові величини* – величини, які в результаті випробувань приймають можливі значення, які суцільно заповнюють деякий інтервал числової осі, скінченний або нескінченний. Множина можливих значень неперервної випадкової величини нескінченна і незліченна. Прикладами неперервних величин є похибки вимірювань фізичних величин з допомогою приладів, час безвідмовної роботи окремих вузлів системи і всієї системи в цілому, відхилення геометричних розмірів виготовленої деталі від стандартних тощо.

Закон розподілу випадкової величини. Ряд розподілу

Для задання випадкової величини недостатньо перелічити всі її можливі значення, необхідно також вказати ймовірності, з якими ця величина приймає те чи інше можливе значення (для дискретної випадкової величини), або ймовірності, з якими випадкова величина попадає в деякий інтервал (для неперервної випадкової величини). Такі повні дані про випадкову величину дають так звані закони розподілу випадкової величини.

Законом розподілу випадкової величини називається залежність (таблиця, графік, функція тощо) між її можливими значеннями і відповідними їм імовірностями.

Найпростішою формою закону розподілу дискретної випадкової величини є *ряд розподілу*. Ряд розподілу являє собою таблицю, в першому рядку якої наведені всі можливі значення дискретної випадкової величини, а в другому – ймовірності, з якими випадкова величина приймає ці значення (табл. 6.1):

Таблиця 6.1. Ряд розподілу дискретної випадкової величини

x_i	x_1	x_2	$\dots$	x_{n-1}	x_n
p_i	p_1	p_2	$\dots$	p_{n-1}	p_n

Та обставина, що в результаті випробування випадкова величина X приймає певне можливе значення x_i , може розглядатися, як випадкова подія $\{X = x_i\}$. Оскільки в результаті випробування величина X приймає лише можливе значення x_i , то події $\{X = x_i\}$ ($0, 1, 2, \dots, n$) утворюють повну групу несумісних подій. Тому справедлива умова (6.1):

$$\sum_{i=1}^n p_i = 1. \quad (6.1)$$

де $p_i = P\{X = x_i\} = P(x_i)$ – ймовірність настання події $\{X = x_i\}$.

Приклад 6.1. Проводиться випробування надійності системи, яка складається з трьох працюючих незалежно приладів. Надійність (ймовірність безвідмовної роботи) першого приладу дорівнює 0,9, другого – 0,8, третього – 0,7. Побудувати ряд розподілу випадкової величини X – числа надійних приладів в системі.

Розв’язання. Випадкова величина X приймає можливі значення 0, 1, 2, 3. Позначимо через g_1, g_2, g_3 ймовірності безвідмовної роботи відповідно першого, другого, третього приладів, тоді за умовою задачі $g_1 = 0,9$; $g_2 = 0,8$; $g_3 = 0,7$, отже, ймовірності виходу з ладу приладів відповідно

дорівнюють $\overline{g}_1 = 0,1$; $\overline{g}_2 = 0,2$; $\overline{g}_3 = 0,3$. Застосувавши теореми додавання і множення ймовірностей, обчислимо ймовірності того, що випадкова величина X приймає можливі значення 0,1,2,3:

$$\begin{aligned} p_0 &= P\{X = 0\} = \overline{g}_1 \cdot \overline{g}_2 \cdot \overline{g}_3 = 0,006; \\ p_1 &= P\{X = 1\} = g_1 \cdot \overline{g}_2 \cdot \overline{g}_3 + g_2 \cdot \overline{g}_1 \cdot \overline{g}_3 + g_3 \cdot \overline{g}_1 \cdot \overline{g}_2 = 0,092; \\ p_2 &= P\{X = 2\} = g_1 \cdot g_2 \cdot \overline{g}_3 + g_1 \cdot g_3 \cdot \overline{g}_2 + g_2 \cdot g_3 \cdot \overline{g}_1 = 0,398; \\ p_3 &= P\{X = 3\} = g_1 \cdot g_2 \cdot g_3 = 0,504. \end{aligned}$$

Ряд розподілу випадкової величини X запишеться у вигляді (табл.6.2):

Таблиця 6.2. Ряд розподілу випадкової величини X

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

Для контролю обчислень перевіримо виконання умови (6.1):

$$\sum_{i=1}^n p_i = 0,006 + 0,092 + 0,398 + 0,504 = 1.$$

Геометричне зображення ряду розподілу називають *багатокутником розподілу*, для побудови якого на осі абсцис відкладають можливі значення випадкової величини X , а на осі ординат – відповідні їм ймовірності, після чого одержані точки з'єднують прямолінійними відрізками.

Багатокутник розподілу для ряду, одержаного в прикладі 6.1, наведено на рис. 6.1.

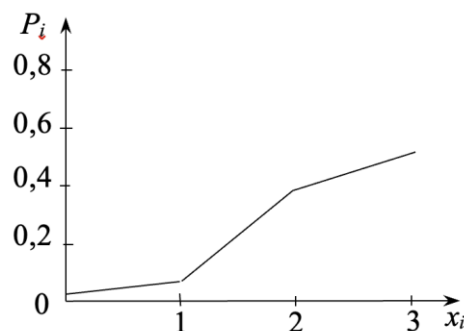


Рис. 6.1. Багатокутник розподілу

Відповідь. Ряд розподілу наведено в табл. 6.2.

6.2. Функція розподілу випадкової величини та її властивості

Ряд розподілу досить повно характеризує випадкову величину, проте побудувати його можна лише для дискретної випадкової величини, оскільки множина можливих значень неперервної випадкової величини – незліченна і, отже, їх не можна перелічити в ряді розподілу.

Загальною формою задання закону розподілу, яка застосовується як для дискретних, так і для неперервних випадкових величин, є функція розподілу, яку інколи також називають інтегральною функцією розподілу.

Функцією розподілу випадкової величини X називається функція $F(x)$, яка для кожного значення x дорівнює ймовірності того, що випадкова величина X прийме значення, менше за x , тобто (6.2):

$$F(x) = P\{X < x\}. \quad (6.2)$$

Побудуємо графік функції розподілу $F(x)$ для дискретної випадкової величини X , заданої рядом розподілу (табл. 6.1).

$$F(x) = P\{X < x\} = 0.$$
$$F(x) = P\{X < x\} = P\{X = x_1\} = p_1.$$

Тому, застосовуючи теорему додавання ймовірностей несумісних подій, одержимо:

$$F(x) = P\{X < x\} = P\{X = x_1\} + P\{X = x_2\} = p_1 + p_2.$$

$$F(x) = P\{X < x\} = P\{X = x_1\} + P\{X = x_2\} + \cdots + P\{X = x_{n-1}\} = p_1 + p_2 + \cdots + p_{n-1}.$$

Таким чином, *функція розподілу $F(x)$ для дискретної випадкової величини X* , заданої рядом розподілу (табл. 6.1), має такий аналітичний вираз (6.3):

$$F(x) = \begin{cases} 0, & \text{при } x \leq x_1; \\ p_1, & \text{при } x_1 < x \leq x_2; \\ p_1 + p_2, & \text{при } x_2 < x \leq x_3; \\ \dots & \\ p_1 + p_2 + \dots + p_{n-1}, & \text{при } x_{n-1} < x \leq x_n; \\ 1, & \text{при } x > x_n. \end{cases} \quad (6.3)$$

96

Як видно з рис. 6.2, графік функції розподілу $F(x)$ дискретної випадкової величини X є розривна східча лінія, стала в інтервалах між можливими значеннями випадкової величини, причому розмір стрибка функції $F(x)$ в точках x_i дорівнює ймовірності p_i , з якою випадкова величина приймає відповідне можливе значення x_i .

При збільшенні числа n можливих значень, які приймає випадкова величина X , довжини східців і розміри стрибків в точках розриву зменшуються, і графік функції $F(x)$ наближається до певної плавної неперервної кривої.

У випадку *неперервної випадкової величини*, у якій множина можливих значень на інтервалі $(a, b]$ незліченна, графік функції $F(x)$ є неперервною лінією, яка схематично зображена суцільною кривою на рис. 6.3.

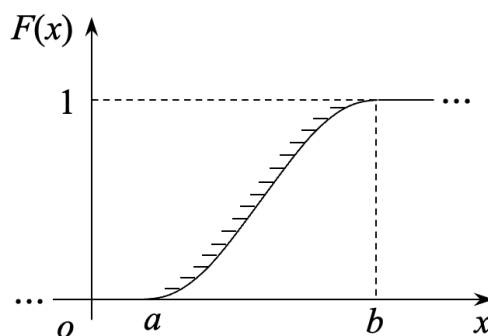


Рис. 6.3. Графік функції розподілу неперервної випадкової величини

Що ж стосується конкретної неперервної випадкової величини X , то її функція розподілу $F(x)$ повинна бути заданою аналітично або графічно. Наприклад, якщо неперервна випадкова величина X задана функцією розподілу $F(x)$:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 0; \\ \frac{x \cdot \sqrt{x}}{8}, & \text{при } 0 < x \leq 4 \\ 1, & \text{при } x > 4, \end{cases}$$

то графік функції $F(x)$ має вигляд, представлений на рис. 6.4.

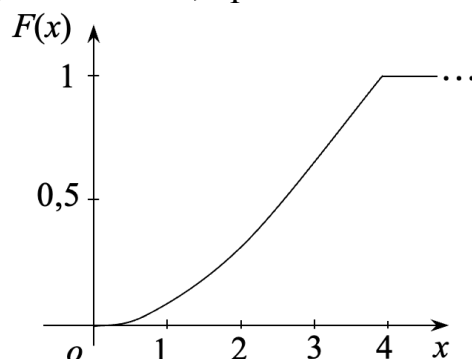


Рис. 6.4. Графік функції розподілу $F(x)$

Властивості функції розподілу випадкової величини

1. Функція розподілу приймає значення з відрізка $[0; 1]$, тобто справедлива нерівність (6.4):

$$0 \leq F(x) \leq 1. \quad (6.4)$$

2. $F(x)$ — неспадна функція, тобто при $x_2 \geq x_1$ справедлива нерівність (6.5):

$$F(x_2) \geq F(x_1). \quad (6.5)$$

3. Ймовірність того, що випадкова величина X в результаті випробування прийме значення з інтервалу $[\alpha; \beta)$, дорівнює приросту функції розподілу $F(x)$ на цьому інтервалі (6.6):

$$P\{\alpha \leq X < \beta\} = F(\beta) - F(\alpha). \quad (6.6)$$

Подібний канонічний запис інтервалу $\alpha \leq X < \beta$ має істотне значення у випадку дискретної випадкової величини.

У випадку неперервної випадкової величини факт включення чи не включення границь інтервалу не має значення.

З властивості 3 одержуємо такий важливий висновок:

4. Ймовірність того, що випадкова величина X в результаті випробування прийме одне конкретне можливе значення x_i , обчислюється за формулою (6.7):

$$P\{X = x_i\} = F(x_i + 0) - F(x_i). \quad (6.7)$$

Зокрема, якщо в точці x_i функція $F(x)$ неперервна, то виконується рівність (6.8):

$$P\{X = x_i\} = 0, \quad (6.8)$$

оскільки за означенням неперервної функції в точці x_i : $F(x_i + 0) = F(x_i)$.

Таким чином, не має сенсу розглядати ймовірність того, що неперервна випадкова величина прийме одне конкретне можливе значення, доцільно розглядати ймовірність її подання в деякий інтервал, нехай навіть досить малий.

5. Якщо випадкова величина X приймає всі можливі значення на інтервалі $[\alpha; \beta)$, то $F(x) = 0$ при $x \leq \alpha$ і $F(x) = 1$ при $x > \beta$.

З властивості 5 одержуємо такий важливий висновок:

6. Якщо випадкова величина X приймає можливі значення на всій числовій осі, то справедливі рівності (6.9):

$$F(-\infty) = 0; F(+\infty) = 1. \quad (6.9)$$

Приклад 6.2. Випадкова величина X має ряд розподілу (табл. 6.3)

Таблиця 6.3. Ряд розподілу випадкової величини X

x_i	1	2	3	4

p_i	0,006	0,092	0,398	0,504

Знайти функцію розподілу $F(x)$ та обчислити ймовірності подій: **а)** $\{X < 3\}$; **б)** $\{2 \leq X \leq 4\}$; **в)** $\{2 < X < 4\}$; **г)** $\{X = 3\}$; **д)** $\{X = 3,5\}$.

Розв'язання. Функція розподілу $F(x)$ будується за схемою (6.3).

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0,006, & \text{при } 1 < x \leq 2; \\ 0,098, & \text{при } 2 < x \leq 3; \\ 0,496, & \text{при } 3 < x \leq 4; \\ 1, & \text{при } x > 4. \end{cases}$$

а) Ймовірність події $\{X < 3\}$ обчислюється за формулою (6.6):

$$P\{X < 3\} = F(3) - F(-\infty) = 0,098 - 0 = 0,098.$$

б) Для обчислення ймовірності події $\{2 \leq X \leq 4\}$ застосовуємо формули (6.6) і (6.7):

$$\begin{aligned} P\{2 \leq X \leq 4\} &= P\{2 \leq X < 4\} + P\{X = 4\} \\ &= F(4) - F(2) + F(4 + 0) - F(4) = F(4 + 0) - F(2) \\ &= 1 - 0,006 = 0,994. \end{aligned}$$

в) $P\{2 < X < 4\} = P\{2 \leq X < 4\} - P\{X = 2\} = F(4) - F(2) - F(2 + 0) + F(2) = F(4) - F(2 + 0) = 0,496 - 0,098 = 0,398.$

г) $P\{X = 3\} = F\{3 + 0\} - F\{3\} = 0,498 - 0,098 = 0,398.$

д) $P\{X = 3,5\} = 0$, оскільки в точці $x = 3,5$ функція розподілу $F(x)$ неперервна.

Відповідь. а) $P\{X < 3\} = 0,098$; б) $P\{2 \leq X \leq 4\} = 0,994$; в) $P\{2 < X < 4\} = 0,398$; г) $P\{X = 3\} = 0,398$; д) $P\{X = 3,5\} = 0$.

Приклад 6.3. Неперервна випадкова величина X задана функцією розподілу

$$F(x) = \begin{cases} 0, & \text{при } x \leq -\frac{\pi}{2}; \\ A \cdot (1 + \sin(x)), & \text{при } -\frac{\pi}{2} < x \leq \frac{\pi}{2}; \\ 1, & \text{при } x > \frac{\pi}{2}. \end{cases}$$

1. Знайти коефіцієнт A ; **2.** Побудувати графік функції $F(x)$; **3.** Обчислити ймовірність того, що в результаті випробування X прийме значення з інтервалу $\left(-\frac{\pi}{4}; \frac{\pi}{4}\right)$.

Розв'язання. 1. Випадкова величина X приймає можливі значення на інтервалі $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$. Коефіцієнт A знайдемо за властивістю 4, згідно якої $F\left(\frac{\pi}{2}\right) = 1$, отже, $A = \frac{1}{2}$.

2. Графік функції $F(x)$ на інтервалі $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$ одержуємо з графіка функції $y = \sin(x)$ зсувом на одиницю в додатному напрямі осі ординат і стисканням уздовж цієї осі вдвічі.

Графік функції $F(x)$ поданий на рис. 6.5.

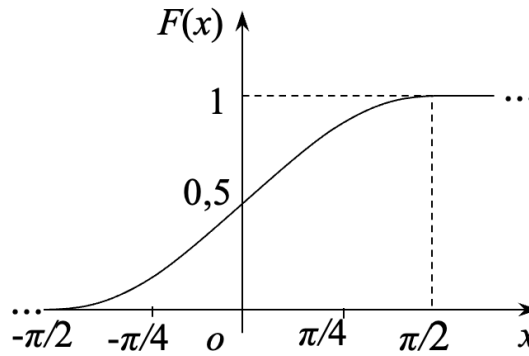


Рис. 6.5. Графік функції розподілу $F(x)$

3. Ймовірність того, що випадкова величина прийме можливе значення з інтервалу $\left(-\frac{\pi}{4}; \frac{\pi}{4}\right)$, обчислимо за формулою (6.6):

$$P\left\{-\frac{\pi}{4} \leq X < \frac{\pi}{4}\right\} = F\left(\frac{\pi}{4}\right) - F\left(-\frac{\pi}{4}\right) = \frac{1}{2} \cdot \left(1 + \frac{\sqrt{2}}{2}\right) - \frac{1}{2} \cdot \left(1 - \frac{\sqrt{2}}{2}\right) = \frac{\sqrt{2}}{2} \approx 0,707.$$

Відповідь. 1. $A = \frac{1}{2}$; 2. Рис. 6.5; 3. $P\left\{-\frac{\pi}{4} \leq X < \frac{\pi}{4}\right\} \approx 0,707$.

6.3. Щільність розподілу неперервної випадкової величини і її властивості

Щільність розподілу або щільність ймовірності $f(x)$ є однією з форм закону розподілу випадкової величини, який застосовується лише для неперервних величин. Термін «щільність ймовірності» походить із задач механіки: якщо в відношенні (6.10)

$$\frac{P\{x \leq X < x + \Delta x\}}{\Delta x} \quad (6.10)$$

відношення ймовірності $P\{x \leq X < x + \Delta x\}$ того, що в результаті випробування випадкова величина X потрапить в інтервал $[x; x + \Delta x)$, до довжини інтервалу Δx ймовірність $P\{x \leq X < x + \Delta x\}$ інтерпретувати як масу речовини, то дане відношення буде не що інше, як середня лінійна щільність (густина) речовини.

Враховуючи, що за формулою (6.6)

$$P\{x \leq X < x + \Delta x\} = F(x + \Delta x) - F(x),$$

представимо відношення (6.10) у вигляді:

$$\frac{P\{x \leq X < x + \Delta x\}}{\Delta x} = \frac{F(x + \Delta x) - F(x)}{\Delta x}$$

або, переходячи до границі при $\Delta x \rightarrow 0$, одержимо в правій частині похідну $F'(x)$ функції розподілу, тобто рівність (6.11):

$$\lim_{\Delta x \rightarrow 0} \left(\frac{P\{x \leq X < x + \Delta x\}}{\Delta x} \right) = \lim_{\Delta x \rightarrow 0} \left(\frac{F(x + \Delta x) - F(x)}{\Delta x} \right) = F'(x). \quad (6.11)$$

В лівій частині формули (6.11) знаходиться щільність ймовірності в точці x , яка позначається $f(x)$, тому отримуємо рівність (6.12):

$$f(x) = \lim_{\Delta x \rightarrow 0} \left(\frac{P\{x \leq X < x + \Delta x\}}{\Delta x} \right) = F'(x). \quad (6.12)$$

Щільністю розподілу або *щільністю ймовірності* $f(x)$ неперервної випадкової величини X називається перша похідна від її функції розподілу $F(x)$.

Функцію $f(x)$ інколи називають також *диференціальною функцією* розподілу випадкової величини.

Властивості щільності розподілу неперервної випадкової величини $f(x)$

1. Щільність ймовірності – невід’ємна функція: $f(x) \geq 0$.
2. Ймовірність того, що в результаті випробування неперервна випадкова величина прийме можливе значення з інтервалу $[\alpha; \beta)$, дорівнює визначеному інтегралу від щільності ймовірності на цьому інтервалі (6.13):

$$P\{\alpha \leq X < \beta\} = \int_{\alpha}^{\beta} f(x)dx. \quad (6.13)$$

Оскільки випадкова величина є неперервною факт включення чи не включення границь інтервалу не має значення.

Якщо крива щільності розподілу $f(x)$ має схематичний вигляд, представлений на рис. 6.6, то геометрично ймовірність $P\{\alpha \leq X < \beta\}$ дорівнює площі криволінійної трапеції, що спирається на відрізок $[\alpha; \beta)$ і обмежена зверху кривою $f(x)$.

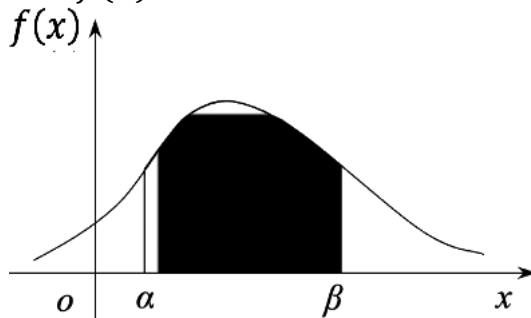


Рис.6.6. Графік функції щільності розподілу $f(x)$

3. Функція розподілу $F(x)$ виражається через щільність розподілу $f(x)$ формулою (6.14):

$$F(x) = \int_{-\infty}^x f(x)dx. \quad (6.14)$$

Геометрично значення функції розподілу $F(x)$ в точці x дорівнює площі, яка розташована лівіше точки x і обмежена кривою щільності розподілу і віссю OX (рис. 6.7).

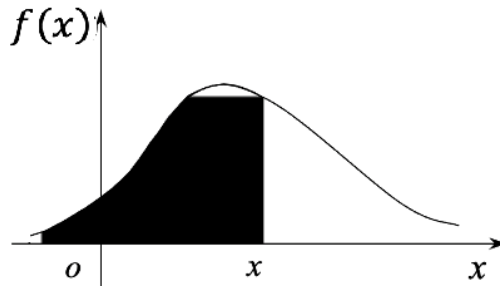


Рис. 6.7. Геометрична інтерпретація властивості 3

4. Якщо неперервна випадкова величина X приймає можливі значення на всій числовій осі, то виконується умова (6.15):

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \quad (6.15)$$

Якщо випадкова величина приймає всі можливі значення на обмеженому інтервалі $(a; b]$, то умова (6.15) набуває вигляду (6.16):

$$\int_a^b f(x)dx = 1. \quad (6.16)$$

Із щільністю ймовірності зв'язане важливе поняття елемента ймовірності, яке використовується для знаходження числових характеристик неперервної випадкової величини.

Із формули (6.16) з точністю до нескінченно малих вищих порядків випливає рівність:

$$P\{x \leq X < x + \Delta x\} = f(x) \cdot \Delta x = f(x)dx. \quad (6.17)$$

Величина $f(x)dx$ (або $f(x) \cdot \Delta x$) називається **елементом ймовірності** для точки x неперервної випадкової величини, оскільки вона дорівнює ймовірності того, що випадкова величина в результаті випробування потрапить в досить малий інтервал Δx .

Геометрично елемент ймовірності наближено дорівнює площі елементарної криволінійної трапеції з основою dx , прилеглою до точки x (рис. 6.8).

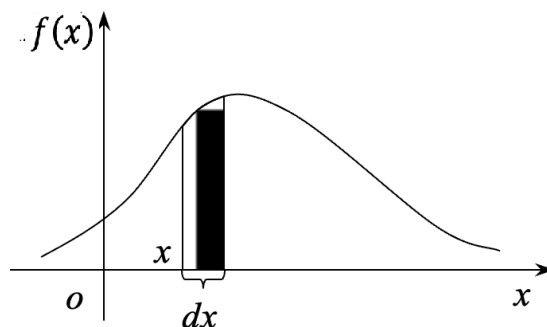


Рис. 6.8. Геометрична інтерпретація властивості 4

Приклад 6.4. Неперервна випадкова величина має щільність ймовірності:

$$f(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ A \cdot \ln(x), & \text{при } 1 < x \leq e^2; \\ 0, & \text{при } x > e^2. \end{cases}$$

Знайти: **1.** Сталій параметр A ; **2.** Функцію розподілу $F(x)$ (побудувати її графік); **3.** Ймовірність того, що в результаті випробування випадкова величина потрапить в інтервал $(e; e^2)$.

Розв'язання. 1. Параметр A знайдемо з умови (6.16):

$$\begin{aligned} \int_a^b f(x)dx &= A \cdot \int_1^{e^2} \ln(x)dx = \left| \begin{array}{l} u = \ln(x); dv = dx \\ du = \frac{1}{x} dx; v = x \end{array} \right| = \\ &= A \cdot \left((x \cdot \ln(x)) \Big|_1^{e^2} - x \Big|_1^{e^2} \right) = A \cdot (e^2 + 1) = 1. \end{aligned}$$

Звідки

$$A = \frac{1}{e^2 + 1}, f(x) = \frac{\ln(x)}{e^2 + 1}, \text{ при } 1 < x \leq e^2, f(x) = 0 \text{ поза цим інтервалом.}$$

2. Функція розподілу знаходиться за формулою (6.14):

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(x)dx = \frac{1}{e^2 + 1} \cdot \int_1^x \ln(x)dx = \frac{1}{e^2 + 1} \cdot (x \cdot \ln(x) - x) \Big|_1^x = \\ &= \frac{1}{e^2 + 1} \cdot (x \cdot \ln(x) - x + 1). \end{aligned}$$

Отже, функція розподілу $F(x)$ має вигляд:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ \frac{x \cdot \ln(x) - x + 1}{e^2 + 1}, & \text{При } 1 < x \leq e^2; \\ 1, & \text{при } x > e^2. \end{cases}$$

Графік $F(x)$ представлений на рис. 6.9.

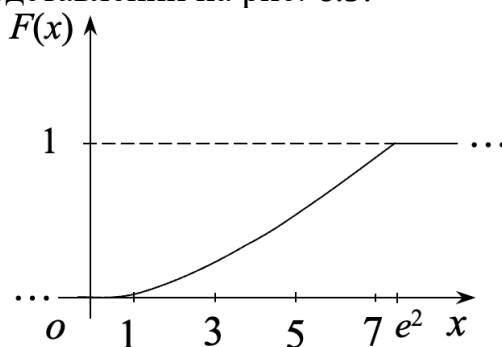


Рис. 6.9. Графік функції розподілу $F(x)$

3. Ймовірність події $\{e < X < e^2\}$ обчислимо за формулою (6.13):

$$\begin{aligned} P\{e < X < e^2\} &= \int_e^{e^2} f(x)dx = \frac{1}{e^2 + 1} \cdot \int_e^{e^2} \ln(x)dx \\ &= \frac{1}{e^2 + 1} \cdot (x \cdot \ln(x) - x) \Big|_e^{e^2} = \frac{e^2}{e^2 + 1} \approx 0,881. \end{aligned}$$

Відповідь. 1. $A = \frac{1}{e^2 + 1}$; 2. $F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ \frac{x \cdot \ln(x) - x + 1}{e^2 + 1}, & \text{при } 1 < x \leq e^2; \\ 1, & \text{при } x > e^2, \end{cases}$ 3. $P\{e <$

$X < e^2\} \approx 0,881.$

6.4. Мода та медіана випадкових величин

Ймовірнісне трактування

Модой $Mo(X)$ випадкової величини X називається її найбільш ймовірне можливе значення.

Для дискретної випадкової величини X , яка має ряд розподілу (табл. 6.1), $Mo(X)$ дорівнює тому можливому значенню, якого величина набуває з найбільшою ймовірністю. Так, для випадкової величини X , ряд розподілу якої побудовано у прикладі 6.1, мода $Mo(X) = 3$.

Для неперервної випадкової величини X , заданої щільністю ймовірності $f(x)$ на проміжку $(a; b)$, мода дорівнює тому значенню x , при якому $f(x)$ досягає максимуму або (у разі відсутності максимуму) набуває найбільшого значення на відрізку.

Якщо $f(x)$ набуває найбільшого значення в одній точці x з проміжку $(a; b)$, то розподіл випадкової величини X називається **унімодальним**, а якщо в кількох точках — **полімодальним**.

Медіаною $Me(X)$ випадкової величини X називається те значення x , для якого виконується рівність (6.18):

$$P(X < Me(X)) = P(X > Me(X)) = \frac{1}{2}. \quad (6.18)$$

Враховуючи формулу (6.6), з формули (6.18) отримуємо:

$$P(X < Me(X)) = F(Me(X)) - F(-\infty) = F(Me(X)) = \frac{1}{2}.$$

Таким чином, якщо випадкова величина задана своєю функцією розподілу, то для обчислення медіани варто використовувати формулу (6.19):

$$F(Me(X)) = \frac{1}{2}. \quad (6.19)$$

Медіана, як правило, застосовується для неперервної випадкової величини і геометрично подає абсцису точки, в якій пряма $x = Me(X)$ поділяє навпіл площу, обмежену кривою розподілу $f(x)$ і віссю OX .

У випадку дискретної випадкової величини медіана обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $[a; b)$ (**медіанному відрізку**) дорівнює $\frac{1}{2}$, згідно формули (6.19). Тоді справедлива рівність (6.20):

$$Me(X) = \frac{a+b}{2}. \quad (6.20)$$

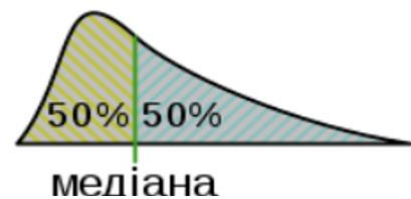
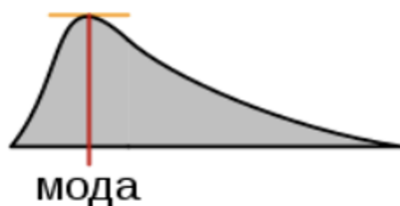


Рис. 6.10. Геометрична візуалізація моди та медіани функції густини ймовірностей

Приклад 6.5. Неперервна випадкова величина X має щільність ймовірності

$$f(x) = \begin{cases} -6 \cdot x^2 + 18 \cdot x - 12, & \text{при } x \in [1; 2); \\ 0, & \text{при } x \notin [1; 2). \end{cases}$$

Знайти моду та медіану цієї величини.

Розв'язання. Для визначення моди перевіримо функцію $f(x)$ на наявність екстремумів, для чого визначимо, при якому значенні x її похідна дорівнює нулю:

$$f'(x) = -12 \cdot x + 18 = 0; x = 1,5.$$

При переході через точку $x = 1,5$ похідна змінює знак з «+» на «-», тому в цій точці $f(x)$ має максимум, а отже, мода $Mo(X) = 1,5$.

Оскільки крива розподілу $f(x)$ симетрична відносно прямої $x = 1,5$ (рис. 6.10), то ця пряма поділяє навпіл площу, обмежену кривою $f(x)$ і віссю OX , а отже, $Me(X) = 1,5$.

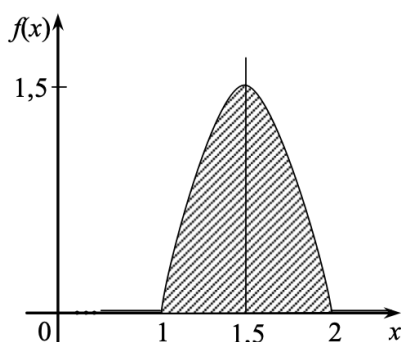


Рис.6.11. Графік функції щільності розподілу $f(x)$

Взагалі, якщо крива розподілу симетрична відносно прямої $x = a$ і має в точці a максимум, то $Mo(X) = Me(X) = a$.

Відповідь. $Mo(X) = Me(X) = 1,5$.

Приклад 6.6. Неперервна випадкова величина X має щільність ймовірності

$$f(x) = \begin{cases} A \cdot x^2, & \text{при } x \in [0; 2); \\ 0, & \text{при } x \notin [0; 2). \end{cases}$$

Знайти: **1.** Коефіцієнт A ; **2.** Моду і медіану величини X .

Розв'язання. 1. Коефіцієнт A знайдемо за властивістю 4 щільності ймовірності:

$$\int_a^b f(x) dx = A \cdot \int_0^2 x^2 dx = \frac{8 \cdot A}{3} = 1; A = \frac{3}{8}.$$

Тому щільність ймовірності має вигляд

$$f(x) = \frac{3}{8} \cdot x^2, \text{ при } x \in [0; 2).$$

Графік кривої розподілу зображено на рис. 6.11.

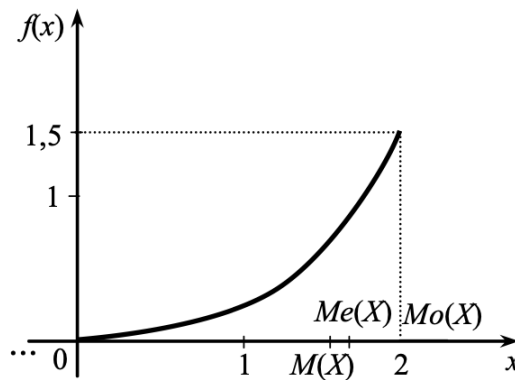


Рис.6.12. Графік функції щільності розподілу $f(x)$

2. На інтервалі $(0; 2)$ крива розподілу не має максимуму, тому мода $Mo(X) = 2$, оскільки при $x = 2$ функція $f(x)$ набуває найбільшого значення.

Медіану знайдемо за означенням (формула (6.18)):

$$P(X < Me(X)) = \frac{3}{8} \cdot \int_0^{Me(X)} x^2 dx = \frac{x^3}{8} \Big|_0^{Me(X)} = \frac{(Me(X))^3}{8} = \frac{1}{2},$$

звідки $Me(X) = \sqrt[3]{4} \approx 1,59$.

Відповідь. 1. $A = \frac{3}{8}$; 2. $Mo(X) = 2$; $Me(X) \approx 1,59$.

Приклад 6.7. Знайти моду і медіану дискретної випадкової величини X (табл. 6.4).

Таблиця 6.4. Ряд розподілу випадкової величини X

x_i	1	2	3	4
p_i	0,1	0,4	0,3	0,2

Розв'язання. Складемо функцію розподілу випадкової величини X :

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0,1, & \text{при } 1 < x \leq 2; \\ 0,5, & \text{при } 2 < x \leq 3; \\ 0,8, & \text{при } 3 < x \leq 4; \\ 1, & \text{при } x > 4. \end{cases}$$

$Mo(X) = 2$, оскільки 2-ці відповідає найбільше значення ймовірності.

$Me(X) = 2,5$ – за формулою (6.19), оскільки $F(x) = 0,5$, при $2 < x \leq$

3.

Відповідь. $Mo(X) = 2$; $Me(X) = 2,5$.

Приклад 6.8. Знайти моду і медіану дискретної випадкової величини X (табл. 6.5).

Таблиця 6.5. Ряд розподілу випадкової величини X

x_i	1	2	3	4
p_i	0,2	0,4	0,1	0,3

Розв'язання. Складемо функцію розподілу випадкової величини X :

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0,2, & \text{при } 1 < x \leq 2; \\ 0,6, & \text{при } 2 < x \leq 3; \\ 0,7, & \text{при } 3 < x \leq 4; \\ 1, & \text{при } x > 4. \end{cases}$$

$Mo(X) = 2$, оскільки 2-ці відповідає найбільше значення ймовірності.

$Me(X)$ немає – за формулою (6.19), оскільки $F(x) \neq 0,5$ на кожному з інтервалів.

Відповідь. $Mo(X) = 2$; $Me(X)$ немає.

Приклад 6.9. Знайти моду і медіану неперервної випадкової величини, заданої функцією розподілу

$$F(x) = \begin{cases} 0, & \text{при } x \leq -1; \\ \frac{x+1}{2}, & \text{при } -1 < x \leq 1; \\ 1, & \text{при } x > 1. \end{cases}$$

Розв'язання. Щільність розподілу отримуємо з формули (6.12).

$$f(x) = \begin{cases} 0, & \text{при } x \leq -1; \\ \frac{1}{2}, & \text{при } -1 < x \leq 1; \\ 0, & \text{при } x > 1. \end{cases}$$

Функція щільності набуває максимального значення на всьому інтервалі $(-1; 1]$. Таким чином, в даному розподілі модами є всі значення $-1 < x \leq 1$. Тобто він є полімодальним.

З формули (6.19)

$$\frac{Me(X) + 1}{2} = \frac{1}{2},$$

тобто $Me(X) = 0$.

Відповідь. $Mo(X)$ – весь інтервал $(-1; 1]$; $Me(X) = 0$.

Статистичне трактування

Для визначення моди та медіани у математичній статистиці введено декілька означень.

Статистична вибірка — це підмножина об'єктів, елементів або спостережень, відібрана з більшої сукупності (генеральної сукупності) для аналізу і вивчення. Вибірка використовується для того, щоб зробити висновки про всю генеральну сукупність на основі меншої кількості даних.

Варіаційний ряд — це впорядкований за зростанням або спаданням набір значень варіанти (спостережень) випадкової величини або статистичної вибірки (табл. 6.6). Він відображає розподіл частот значень досліджуваної величини, дозволяючи оцінювати розподіл даних та проводити різноманітні статистичні аналізи.

Таблиця 6.6. Варіаційний ряд

x_i	x_1	x_2	...	x_k

n_i	n_1	p_2	...	n_k

де:

x_i – елементи вибірки ($i = \overline{1, k}$);

n_i – частоти елементів вибірки;

$n = \sum_{i=1}^k n_i$ – об'єм вибірки.

Інтервальний варіаційний ряд — це спосіб подання статистичних даних, де всі спостереження розподіляються за інтервалами (табл. 6.7). Кожен інтервал включає певний діапазон значень змінної, а також частоти або частки, які показують, скільки спостережень належать до кожного інтервалу. Інтервальний варіаційний ряд використовується для групування даних і надання загальної картини їх розподілу.

Таблиця 6.7. Інтервальний варіаційний ряд

$[x_i; x_{i+1})$	$[x_1; x_2)$	$[x_2; x_3)$	...	$[x_m; x_{m+1})$
n_i	n_1	n_2	...	n_m

де:

$[x_i; x_{i+1})$ – інтервали ($i = \overline{1, m}$);

n_i – частоти інтервалів вибірки;

$n = \sum_{i=1}^m n_i$ – об'єм вибірки.

В статистиці **медіана** — це величина ознаки, що розташована посередині ранжованого ряду вибірки, тобто — це величина, що розташована в середині ряду величин, розташованих у зростальному або спадному порядку; в теорії ймовірності — характеристика розподілення випадкової величини.

Знаходження медіани:

1. Необхідно ранжувати статистичну вибірку;

2. Якщо n – непарне число, то $Me = x_{\frac{1+n}{2}}$ – середина вибірки;

3. Якщо n – парне число, то $Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$ – середнє арифметичне двох середніх елементів вибірки.

Модю називають значення ознаки, що має найбільшу частоту в статистичному ряду розподілу.

Спосіб обчислення моди залежить від того, в якому вигляді дано значення ознаки: дискретного чи інтервального ряду розподілу. В дискретних варіаційних рядах моду обчислюють без додаткових розрахунків за значенням варіанти з найбільшою частотою.

Тип	Опис	Приклад	Результат
Середнє арифметичне	Сума всіх значень вибірки поділена на кількість цих елементів вибірки: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$(1+2+2+3+4+7+9) / 7$	4
Медіана	Середнє значення, що відокремлює більшу половину і меншу половину вибірки	1, 2, 2, 3, 4, 7, 9	3
Мода	Значення, що зустрічається у вибірці найчастіше	1, 2, 2, 3, 4, 7, 9	2

Рис. 6.13. Порівняння різних загальних середніх значень для вибірки {1,2,2,3,4,7,9}

При розрахунку моди в інтервальному варіаційному ряду розподілу спочатку потрібно визначити модальний інтервал, в межах якого знаходяться мода, а потім значення модальної величини ознаки.

В інтервальному варіаційному ряду розподілу модою наближено вважають центральний варіант так званого модального інтервалу, тобто того інтервалу, який має найбільшу частоту. В межах інтервалу необхідно знайти те значення ознаки, яке є модою.

В інтервальних варіаційних рядах розподілів модою визначають за формулою (6.21):

$$Mo = x_{Mo} + i_{Mo} \cdot \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}, \quad (6.21)$$

де:

Mo – мода;

x_{Mo} – нижня межа модального інтервалу;

i_{Mo} – довжина модального інтервалу;

f_{Mo} – частота модального інтервалу;

f_{Mo-1} – частота інтервалу, що передуює модальному;

f_{Mo+1} – частота інтервалу, що є наступним за модальним.

Формула ґрунтується на припущенні, що відстані від нижньої межі модального інтервалу до моди і від моди до верхньої межі модального інтервалу прямо пропорційні різницям між чисельностями (частотами) модального інтервалу і інтервалів, що прилягають до нього.

В інтервальних варіаційних рядах розподілів медіану визначають за формулою. (6.22):

$$Me = x_{Me} + i_{Me} \cdot \frac{\frac{n}{2} - n_{Me-1}^H}{n_{Me}}, \quad (6.22)$$

де:

Me – медіана;

x_{Me} – нижня межа медіанного інтервалу;

i_{Me} – довжина медіанного інтервалу;

n – об'єм вибірки;

n_{Me-1}^H – накопичувана частота інтервалу, що передуює медіанному;

n_{Me} – частота медіанного інтервалу.

Приклад 6.10. Знайти моду і медіану вибірки, що задана інтервальним рядом (табл. 6.8).

Таблиця 6.8. Інтервальний варіаційний ряд

Інтервал добової потужності $[x_i; x_{i+1})$	Кількість заводів (частота) n_i	Кумулятивна (накопичувана) частота n_i^H
[20,25)	7	7
[25,35)	14	21
[35,40)	6	27

[40,45)	5	32
[45,50)	8	40
Разом n	40	—

Розв'язання. Обчислимо моду.

[25,35) – модальний інтервал, оскільки має найбільшу частоту.

$$Mo = 25 + 10 \cdot \frac{14-7}{(14-7)+(14-6)} \approx 30.$$

Обчислимо медіану.

[25,35) – медіанний інтервал, оскільки на цей інтервал припадає середина вибірки.

$$Me = 25 + 10 \cdot \frac{\frac{40}{2} - 7}{14} \approx 34.$$

ЛЕКЦІЯ 7

ОСНОВНІ ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИПАДКОВИХ ВЕЛИЧИН

Закон розподілу найбільш повно характеризує випадкову величину і дозволяє обчислювати ймовірності всіх подій, зв'язаних з випадковою величиною. Проте при розв'язанні багатьох практичних задач важливу роль відіграють числові значення, які дають деяку узагальнену, усереднену характеристику випадкової величини та її розподілу. Ці значення називаються **числовими характеристиками** випадкової величини. Основними з них є математичне сподівання, дисперсія, середнє квадратичне відхилення та моменти різних порядків.

Для їх знаходження не треба знати точні закони розподілу випадкових величин, встановлення яких часто пов'язане з значними труднощами. В той же час числові характеристики в багатьох випадках вичерпують наші потреби в даних про випадкову величину.

7.1. Математичне сподівання випадкової величини та його властивості

Розглянемо спочатку дискретну випадкову величину X , задану рядом розподілу (табл. 7.1):

Таблиця 7.1. Ряд розподілу дискретної випадкової величини

x_i	x_1	x_2	$\dots$	x_{n-1}	x_n
p_i	p_1	p_2	$\dots$	p_{n-1}	p_n

В табл. 7.1 сума всіх ймовірностей дорівнює одиниці:

$$\sum_{i=1}^n p_i = 1.$$

Математичним сподіванням $M(X)$ дискретної випадкової величини X називається сума добутків всіх її можливих значень на ймовірності, з якими випадкова величина приймає ці значення (7.1):

$$M(X) = \sum_{i=1}^n (x_i \cdot p_i). \quad (7.1)$$

Із означення випливає, що математичне сподівання є величина не випадкова (стала). Вона має такий **ймовірнісний зміст**: математичне сподівання при великій кількості випробувань наближено дорівнює середньому арифметичному можливих значень випадкової величини, які вона приймала в цих випробуваннях. Тому математичне сподівання $M(X)$ називають ще **середнім значенням** випадкової величини.

Перейдемо до обчислення математичного сподівання неперервної випадкової величини X , заданої щільністю ймовірності $f(x)$ на інтервалі $(a; b)$.

Розіб'ємо $(a; b]$, довільно на n частинних інтервалів з довжинами $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ і виберемо на кожному частинному інтервалі довільну точку $x_1, x_2, \dots, x_n$. Будемо вважати, що на i -му частинному інтервалі випадкова величина приймає стале можливе значення, що дорівнює x_i

($i = 1, 2, \dots, n$). Ймовірність того, що випадкова величина потрапить в інтервал Δx_i дорівнює $f(x_i) \cdot \Delta x_i$, тому по аналогії з математичним сподіванням дискретної випадкової величини (7.1) одержимо

$$M(X) \approx \sum_{i=1}^n (x_i \cdot f(x_i) \cdot \Delta x_i).$$

Наближена рівність виникає внаслідок припущення про сталість можливих значень випадкової величини на кожному частинному інтервалі. Для усунення цієї наближеності перейдемо до границі за умови, що довжина найбільшого частинного інтервалу прямує до нуля:

$$M(X) = \lim_{\max \Delta x_i \rightarrow 0} (\sum_{i=1}^n (x_i \cdot f(x_i) \cdot \Delta x_i)).$$

Або за означенням визначеного інтегралу від функції $x \cdot f(x)$ на відрізьку $[a; b]$, то отримаємо рівність (7.2):

$$M(X) = \int_a^b (x \cdot f(x)) dx. \quad (7.2)$$

Якщо неперервна випадкова величина задана щільністю ймовірності $f(x)$ при $x \in (-\infty; +\infty)$, то то отримаємо рівність (7.3):

$$M(X) = \int_{-\infty}^{+\infty} (x \cdot f(x)) dx. \quad (7.3)$$

Ясно, що в цьому випадку для існування математичного сподівання випадкової величини невласний інтеграл повинен бути збіжним. Це ж зауваження стосується і дискретної випадкової величини, у якій множина можливих значень нескінченна, і для якої формула (7.1) приймає вигляд (7.4):

$$M(X) = \sum_{i=1}^{+\infty} (x_i \cdot p_i). \quad (7.4)$$

Математичне сподівання цієї випадкової величини існує при умові збіжності ряду в правій частині.

Приклад 7.1. Випадкова величина X – число надійних приладів в системі має ряд розподілу (табл. 7.2).

Таблиця 7.2. Ряд розподілу випадкової величини X

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

Знайти $M(X)$.

Розв'язання. Математичне сподівання заданої дискретної величини обчислюється за формулою (7.1)

$$M(X) = 0 \cdot 0,006 + 1 \cdot 0,092 + 2 \cdot 0,398 + 3 \cdot 0,504 = 2,4.$$

Відповідь. $M(X) = 2,4$.

Приклад 7.2. Ймовірність події A дорівнює p . Знайти математичне сподівання кількості появ події A в одному випробуванні.

Розв'язання. Випадкова величина X — кількість появ події A в одному випробуванні приймає два можливі значення: $x_1 = 1$ (подія A відбулась) з ймовірністю p і $x_2 = 0$ (подія A не відбулась) з ймовірністю $q = 1 - p$, тому

$$M(X) = 1 \cdot p + 0 \cdot q = p.$$

Відповідь. Математичне сподівання кількості появ події A в одному випробуванні дорівнює ймовірності цієї події.

Приклад 7.3. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} \ln(x), & \text{при } x \in (1; e]; \\ 0, & \text{при } x \notin (1; e]. \end{cases}$$

Знайти математичне сподівання $M(X)$.

Розв'язання. Для знаходження математичного сподівання і використання в подальших прикладах обчислимо частинами інтеграл:

$$\int_1^e (x^n \cdot \ln(x)) dx = \left| \begin{array}{l} u = \ln(x); dv = x^n dx \\ du = \frac{dx}{x}; v = \frac{x^{n+1}}{n+1} \end{array} \right| = \frac{x^{n+1}}{n+1} \cdot \ln(x) \Big|_1^e - \frac{x^{n+1}}{(n+1)^2} \Big|_1^e =$$

$$\frac{n \cdot e^{n+1} + 1}{(n+1)^2}.$$

Тоді за формулою (7.2)

$$M(X) = \int_1^e (x \cdot \ln(x)) dx = \frac{e^2 + 1}{4} \approx 2,1.$$

Відповідь. $M(X) \approx 2,1$.

Властивості математичного сподівання випадкової величини

1. Якщо C – константа, то $M(C) = C$.

Доведення. Константу C можна розглядати, як випадкову величину, що приймає значення C з ймовірністю 1 (табл. 7.3).

Таблиця 7.3. Ряд розподілу випадкової величини C

x_i	$\bar{C}$	C
p_i	0	1

Тому $M(C) = \bar{C} \cdot 0 + C \cdot 1 = C$.

Для розгляду наступної властивості математичного сподівання дамо означення суми (різниці) випадкових величин.

Сумою (різницею) двох випадкових величин X та Y називається така випадкова величина, що приймає значення $x_i \pm y_j$ ($i = \overline{1, n}; j = \overline{1, m}$) з ймовірністю $p_{ij} = P\{(X = x_i)(Y = y_j)\}$.

2. Математичне сподівання суми (різниці) двох випадкових величин дорівнює сумі їх математичних сподівань:

$$M(X \pm Y) = M(X) \pm M(Y).$$

Доведення. Згідно означення суми (різниці) двох випадкових величин

$$\begin{aligned}
M(X \pm Y) &= \sum_{i=1}^n \sum_{j=1}^m ((x_i \pm y_j) \cdot p_{ij}) = \\
&= \sum_{i=1}^n \sum_{j=1}^m (x_i \cdot p_{ij}) \pm \sum_{i=1}^n \sum_{j=1}^m (y_j \cdot p_{ij}) = \sum_{i=1}^n \left(x_i \cdot \sum_{j=1}^m p_{ij} \right) \\
&\pm \sum_{j=1}^m \left(y_j \cdot \sum_{i=1}^n p_{ij} \right) = \sum_{i=1}^n (x_i \cdot p_i) \pm \sum_{j=1}^m (y_j \cdot p'_j) = M(X) \pm M(Y).
\end{aligned}$$

Для розгляду наступної властивості математичного сподівання дамо деякі означення.

Добутком двох випадкових величин X та Y називається така випадкова величина, що приймає значення $x_i \cdot y_j$ ($i = \overline{1, n}; j = \overline{1, m}$) з ймовірністю $p_{ij} = P\{(X = x_i)(Y = y_j)\}$.

Дві випадкові величини X і Y називаються **незалежними**, якщо закон розподілу однієї з них не залежить від того, які можливі значення приймає друга величина.

3. Математичне сподівання добутку двох незалежних випадкових величин і Y дорівнює добутку їх математичних сподівань:

$$M(X \cdot Y) = M(X) \cdot M(Y).$$

Доведення. Згідно означення добутку двох випадкових величин

$$\begin{aligned}
M(X \cdot Y) &= \sum_{i=1}^n \sum_{j=1}^m ((x_i \cdot y_j) \cdot p_{ij}) = \sum_{i=1}^n \sum_{j=1}^m ((x_i \cdot y_j) \cdot (p_i \cdot p'_j)) = \\
&= \sum_{i=1}^n \sum_{j=1}^m (x_i \cdot p_i) \cdot \sum_{i=1}^n \sum_{j=1}^m (y_j \cdot p'_j) = \sum_{i=1}^n (x_i \cdot p_i) \cdot \sum_{j=1}^m (y_j \cdot p'_j). \\
&= M(X) \cdot M(Y).
\end{aligned}$$

Методом математичної індукції властивість 3 поширюється на довільну кількість n незалежних випадкових величин $X_1, X_2, \dots, X_n$:

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

4. Сталій множник C можна виносити за знак математичного сподівання:

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

Доведення. Оскільки випадкова величина $C \cdot X$ приймає значення $C \cdot x_i$ ($i = \overline{1, n}$), то

$$M(C \cdot X) = \sum_{i=1}^n ((C \cdot x_i) \cdot p_i) = C \cdot \sum_{i=1}^n (x_i \cdot p_i) = C \cdot M(X).$$

5. Математичне сподівання відхилення випадкової величини X від її математичного сподівання $M(X)$ дорівнює нулю:

$$M(X - M(X)) = 0.$$

Доведення. Нехай константа C є математичним сподіванням $C = M(X)$. Тоді, враховуючи властивість 2, отримуємо

$$M(X - C) = M(X) - C = C - C = 0.$$

7.2. Дисперсія випадкової величини та її властивості.

Середнє квадратичне відхилення

Як уже відзначалося, математичне сподівання характеризує випадкову величину в середньому. Проте, випадкові величини, які мають одне і те ж математичне сподівання, можуть істотно відрізнятися законом розподілу. Отже, математичне сподівання недостатньо характеризує розподіл випадкової величини, так само, як, наприклад, однакова середня заробітна плата на двох підприємствах не дає уявлення про співвідношення низько і високооплачуваних категорій працівників цих підприємств.

Тому для характеристики випадкової величини важливо оцінити принаймні в середньому міру її розкиду (розсіювання) навколо математичного сподівання. За таку міру природно було б прийняти математичне сподівання відхилення випадкової величини. Проте за властивістю 5 ця величина дорівнює нулю і, отже, не може бути характеристикою розсіювання. Тому за міру розсіювання прийнята інша величина, яка називається дисперсією і поряд з математичним сподіванням відноситься до основних характеристик випадкової величини.

Дисперсією (розсіюванням) $D(X)$ випадкової величини X називається математичне сподівання квадрата її відхилення (7.5):

$$D(X) = M(X - M(X))^2. \quad (7.5)$$

Виходячи з формул (7.1) і (7.2) для математичного сподівання, одержуємо формули для обчислення дисперсії:

- для дискретної випадкової величини, заданої рядом розподілу (табл. 7.1) отримаємо рівність (7.6):

$$D(X) = \sum_{i=1}^n ((x_i - M(X))^2 \cdot p_i). \quad (7.6)$$

- для неперервної випадкової величини, заданої щільністю ймовірності $f(x)$ на інтервалі $(a; b)$ отримаємо рівність (7.7):

$$D(X) = \int_a^b ((x - M(X))^2 \cdot f(x)) dx. \quad (7.7)$$

Очевидними перетвореннями із застосуванням властивостей 1 і 2 математичного сподівання виразу для дисперсії (7.5) можна надати іншого вигляду:

$$\begin{aligned} D(X) &= M(X^2 - 2 \cdot X \cdot M(X) + (M(X))^2) = \\ &= M(X^2) - 2 \cdot M(X) \cdot M(X) + (M(X))^2 = M(X^2) - (M(X))^2. \end{aligned}$$

Отже, дисперсія випадкової величини X дорівнює різниці між математичним сподіванням квадрата випадкової величини і квадратом її математичного сподівання (7.8):

$$D(X) = M(X^2) - (M(X))^2. \quad (7.8)$$

Квадрат випадкової величини X , заданої рядом розподілу (табл. 7.1), є випадкова величина X^2 , яка приймає можливі значення x_i^2 з ймовірностями p_i ($i = \overline{1, n}$), оскільки, як тільки випадкова величина X приймає певне можливе значення, наприклад, x_1 , величина X^2 приймає можливе значення x_1^2 .

Тому для обчислення дисперсії дискретної випадкової величини X , крім формули (7.6), можна застосувати формулу (7.9):

$$D(X) = \sum_{i=1}^n (x_i^2 \cdot p_i) - (M(X))^2. \quad (7.9)$$

і, відповідно, для неперервної на відрізок $(a; b]$ випадкової величини, крім формули (7.7), формулу (7.10):

$$D(X) = \int_a^b (x^2 \cdot f(x)) dx - (M(X))^2. \quad (7.10)$$

Якщо неперервна випадкова величина задана щільністю ймовірності $f(x)$ при $x \in (-\infty; +\infty)$, то отримаємо рівність (7.11):

$$D(X) = \int_{-\infty}^{+\infty} (x^2 \cdot f(x)) dx - (M(X))^2. \quad (7.11)$$

Приклад 7.4. Випадкова величина X має ряд розподілу (табл. 7.4).

Таблиця 7.4. Ряд розподілу випадкової величини X

x_i	0	1	2	3
p_i	0,6	0,2	0,1	0,1

Знайти дисперсію $D(X)$.

Розв'язання. Математичне сподівання за формулою (7.1) дорівнює

$$M(X) = 0 \cdot 0,6 + 1 \cdot 0,2 + 2 \cdot 0,1 + 3 \cdot 0,1 = 0,7.$$

Знайдемо математичне сподівання квадрата цієї випадкової величини:

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

За формулою (7.9)

$$D(X) = 1,5 - 0,49 = 1,01.$$

Відповідь. $D(X) = 1,01$.

Приклад 7.5. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} \ln(x), & \text{при } x \in (1; e]; \\ 0, & \text{при } x \notin (1; e]. \end{cases}$$

Знайти дисперсію $D(X)$.

Розв'язання. Дисперсію обчислимо за формулою (7.8). Математичне сподівання $M(X)$ знайдено в прикладі 7.3: $M(X) = \frac{e^2+1}{4} \approx 2,1$.

Використовуючи формулу (7.2), одержимо

$$M(X^2) = \int_1^e (x^2 \cdot \ln(x)) dx = \frac{2 \cdot e^3 + 1}{9} \approx 4,57.$$

Отже,

$$D(X) \approx 4,57 - 4,41 \approx 0,16.$$

Відповідь. $D(X) \approx 0,16$.

Властивості дисперсії випадкової величини

1. Дисперсія сталої величини C дорівнює нулю:

$$D(C) = 0.$$

Доведення. За означенням дисперсії (7.5)

$$D(C) = M(C - M(C))^2 = M(C - C)^2 = M(0) = 0.$$

2. Сталу величину C можна виносити за знак дисперсії, підносячи її до квадрату:

$$D(C \cdot X) = C^2 \cdot D(X).$$

Доведення. За означенням дисперсії (7.5)

$$\begin{aligned} D(C \cdot X) &= M(C \cdot X - M(C \cdot X))^2 = M(C \cdot X - C \cdot M(X))^2 \\ &= C^2 \cdot M(X - M(X))^2 = C^2 \cdot D(X). \end{aligned}$$

3. Дисперсія суми двох незалежних випадкових величин X та Y дорівнює сумі дисперсій цих величин:

$$D(X \pm Y) = D(X) + D(Y).$$

Доведення. За формулою (7.8)

$$\begin{aligned} D(X \pm Y) &= M(X \pm Y)^2 - (M(X \pm Y))^2 = \\ &= M(X^2 \pm 2 \cdot X \cdot Y + Y^2) - (M(X) \pm M(Y))^2 = \\ &= M(X^2) \pm M(2 \cdot X \cdot Y) + M(Y^2) - (M(X))^2 \mp 2 \cdot M(X) \cdot M(Y) \\ &\quad - (M(Y))^2. \end{aligned}$$

Враховуючи властивості 3 та 4 математичного сподівання, отримуємо

$$\begin{aligned}
D(X \pm Y) &= M(X^2) \pm 2 \cdot M(X) \cdot M(Y) - (M(Y))^2 + M(Y^2) - (M(X))^2 \mp 2 \cdot \\
&\quad \cdot M(X) \cdot M(Y) - (M(Y))^2 = \\
&= (M(X^2) - (M(X))^2) + (M(Y^2) - (M(Y))^2) = \\
&= D(X) + D(Y).
\end{aligned}$$

4. Дисперсія суми випадкової величини X і сталої величини C дорівнює дисперсії випадкової величини X :

$$D(X + C) = D(X).$$

Доведення. За властивостями 1 та 3 дисперсії

$$D(X + C) = D(X) + D(C) = D(X) + 0 = D(X).$$

Ще однією характеристикою розсіювання випадкової величини є середнє квадратичне відхилення. Оскільки за означенням дисперсія має розмірність квадрата випадкової величини, то для характеристики розсіювання зручніше застосовувати арифметичне значення квадратного кореня з дисперсії. Ця величина називається *середнім квадратичним відхиленням* випадкової величини і позначається $\sigma(X)$ (7.12):

$$\sigma(X) = \sqrt{D(X)}. \quad (7.12)$$

Приклад 7.6. Випадкова величина X має ряд розподілу (табл. 7.5).

Таблиця 7.5. Ряд розподілу випадкової величини X

x_i	-0,1	x_2	0,2	0,4
p_i	0,3	0,1	p_3	p_4

1. Знайти x_2, p_3, p_4 , якщо відомі математичне сподівання $M(X) = 0,13$ та дисперсія $D(X) = 0,0341$. **2.** Скласти функцію розподілу випадкової величини X . **3.** Знайти ймовірність попадання цієї випадкової величини в інтервал $(-0,05; 0,2]$.

Розв'язання. 1. Складаємо систему рівнянь.

1-ше рівняння:

$$\begin{aligned}
M(X) &= -0,1 \cdot 0,3 + x_2 \cdot 0,1 + 0,2 \cdot p_3 + 0,4 \cdot p_4 \\
&= -0,03 + 0,1 \cdot x_2 + 0,2 \cdot p_3 + 0,4 \cdot p_4.
\end{aligned}$$

Отже, рівняння матиме вигляд:

$$\begin{aligned}
-0,03 + 0,1 \cdot x_2 + 0,2 \cdot p_3 + 0,4 \cdot p_4 &= 0,13; \\
0,1 \cdot x_2 + 0,2 \cdot p_3 + 0,4 \cdot p_4 &= 0,16; \\
x_2 + 2 \cdot p_3 + 4 \cdot p_4 &= 1,6.
\end{aligned}$$

2-ге рівняння:

$$\begin{aligned}
D(X) &= (-0,1)^2 \cdot 0,3 + x_2^2 \cdot 0,1 + (0,2)^2 \cdot p_3 + (0,4)^2 \cdot p_4 - (0,13)^2 = \\
&= 0,003 + 0,1 \cdot x_2^2 + 0,04 \cdot p_3 + 0,16 \cdot p_4 - 0,0169 \\
&= 0,1 \cdot x_2^2 + 0,04 \cdot p_3 + 0,16 \cdot p_4 - 0,0139.
\end{aligned}$$

Отже, рівняння матиме вигляд:

$$0,1 \cdot x_2^2 + 0,04 \cdot p_3 + 0,16 \cdot p_4 - 0,0139 = 0,0341;$$

$$\begin{aligned} 0,1 \cdot x_2^2 + 0,04 \cdot p_3 + 0,16 \cdot p_4 &= 0,048; \\ x_2^2 + 0,4 \cdot p_3 + 1,6 \cdot p_4 &= 0,48. \end{aligned}$$

3-тє рівняння:

$$0,3 + 0,1 + p_3 + p_4 = 1.$$

Отже, рівняння матиме вигляд:

$$p_3 + p_4 = 0,6.$$

Отримуємо наступну систему рівнянь.

$$\begin{cases} x_2 + 2 \cdot p_3 + 4 \cdot p_4 = 1,6; \\ x_2^2 + 0,4 \cdot p_3 + 1,6 \cdot p_4 = 0,48; \\ p_3 + p_4 = 0,6, \end{cases} \quad \begin{cases} x_2 - 2 \cdot p_3 = -0,8; \\ x_2^2 - 1,2 \cdot p_3 = -0,48; \\ p_4 = 0,6 - p_3. \end{cases}$$

Якщо помножити перше рівняння системи на $-0,6$ та додати до другого, отримаємо:

$$\begin{aligned} x_2^2 - 0,6 \cdot x_2 &= 0. \\ x_2 &= 0 \text{ або } x_2 = 0,6. \end{aligned}$$

Дана система має єдиний розв'язок: $x_2 = 0$; $p_3 = 0,4$; $p_4 = 0,2$, оскільки при $x_2 = 0,6$ значення ймовірностей p_3, p_4 не задовольняють третьому рівнянню системи.

Тоді остаточний ряд розподілу випадкової величини X матиме вигляд (табл. 7.6)

Таблиця 7.6. Ряд розподілу випадкової величини X

x_i	-0,1	0	0,2	0,4
p_i	0,3	0,1	0,4	0,2

2. Функція розподілу матиме вигляд (формула (6.3) минулої лекції):

$$F(x) = \begin{cases} 0, & \text{при } x \leq -0,1; \\ 0,3, & \text{при } -0,1 < x \leq 0; \\ 0,4, & \text{при } 0 < x \leq 0,2; \\ 0,8, & \text{при } 0,2 < x \leq 0,4; \\ 1, & \text{при } x > 0,4. \end{cases}$$

3. Обчислимо $P\{-0,05 < X \leq 0,2\}$ двома способами.

1 спосіб

За рядом розподілу випадкової величини X (табл. 7.6) обираємо значення x_i , які потрапляють в потрібний інтервал та додаємо їх ймовірності згідно теореми додавання ймовірностей несумісних подій. В результаті отримуємо:

$$P\{-0,05 < X \leq 0,2\} = P\{X = 0\} + P\{X = 0,2\} = 0,1 + 0,4 = 0,5.$$

2 спосіб

За формулою (6.6) минулої лекції

$$\begin{aligned} P\{-0,05 < X \leq 0,2\} &= P\{-0,05 \leq X < 0,2\} - P\{X = -0,05\} + P\{X = 0,2\} \\ &= \\ &= F(0,2) - F(-0,05) - F(-0,05 + 0) + F(-0,05) + F(0,2 + 0) \\ &\quad - F(0,2) = F(0,2 + 0) - F(-0,05 + 0) = 0,8 - 0,3 = 0,5. \end{aligned}$$

Відповідь. 1. $x_2 = 0; p_3 = 0,4; p_4 = 0,2;$ 2. $F(x) =$

$$\begin{cases} 0, & \text{при } x \leq -0,1; \\ 0,3, & \text{при } -0,1 < x \leq 0; \\ 0,4, & \text{при } 0 < x \leq 0,2; \\ 0,8, & \text{при } 0,2 < x \leq 0,4; \\ 1, & \text{при } x > 0,4; \end{cases}$$
 3. $P\{-0,05 < X \leq 0,2\} = 0,5.$

Приклад 7.7. Неперервну випадкову величину X задано функцією розподілу:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ A \cdot \sqrt[3]{x-1}, & \text{при } 1 < x \leq 9; \\ 1, & \text{при } x > 9. \end{cases}$$

Знайти: **а)** параметр A ; **б)** щільність ймовірності $f(x)$; **в)** числові характеристики $M(X), D(X), \sigma(X)$; **г)** ймовірність того, що випадкова величина X у результаті випробування набуде можливого значення з інтервалу $(1; 3)$.

Розв'язання. **а)** За властивістю 4 функції розподілу $F(9) = 2 \cdot A = 1$, звідки $A = \frac{1}{2}$.

б) Щільність ймовірності:

$$f(x) = F'(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ \frac{1}{6 \cdot \sqrt[3]{(x-1)^2}}, & \text{при } 1 < x \leq 9; \\ 1, & \text{при } x > 9. \end{cases}$$

в) Математичне сподівання обчислюємо за формулою (7.2):

$$\begin{aligned} M(X) &= \int_1^9 (x \cdot f(x)) dx = \int_1^9 \left(x \cdot \frac{1}{6 \cdot \sqrt[3]{(x-1)^2}} \right) dx = \left| \begin{matrix} t^3 = x-1 \\ dx = 3 \cdot t^2 dt \end{matrix} \right| \\ &= \frac{1}{2} \cdot \int_0^2 (t^3 + 1) dt = 3. \end{aligned}$$

Дисперсію обчислюємо за формулою (7.10), знову застосовуючи при інтегруванні заміну змінної $t^3 = x - 1$:

$$\begin{aligned} D(X) &= \int_1^9 (x^2 \cdot f(x)) dx - 3^2 \\ &= \int_1^9 \left(x^2 \cdot \frac{1}{6 \cdot \sqrt[3]{(x-1)^2}} \right) dx - 9 = \left| \begin{matrix} t^3 = x-1 \\ dx = 3 \cdot t^2 dt \end{matrix} \right| \\ &= \frac{1}{2} \cdot \int_0^2 (t^3 + 1)^2 dt - 9 \approx 5,143. \end{aligned}$$

$$\sigma(X) \approx 2,268.$$

г) Ймовірність того, що в результаті випробування X набуде можливого значення з інтервалу $(0; 1)$, знайдемо за формулою (6.6) минулої лекції:

$$P\{1 < X < 3\} = F(3) - F(1) = \frac{\sqrt[3]{2}}{2} \approx 0,63.$$

Відповідь. 1. $A = \frac{1}{2}$; 2. $f(x) = F'(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ \frac{1}{6 \cdot \sqrt[3]{(x-1)^2}}, & \text{при } 1 < x \leq 9; \\ 1, & \text{при } x > 9; \end{cases}$ 3. $M(X) = 3$; $D(X) \approx 5,143$; $\sigma(X) \approx 2,268$; 4. $P\{1 < X < 3\} \approx 0,63$.

7.3. Центральний та початковий моменти

Початковим моментом порядку k випадкової величини X (k – натуральне число) називається математичне сподівання величини X^k (7.13):

$$v_k(X) = M(X^k). \quad (7.13)$$

Для дискретної випадкової величини, заданої рядом розподілу (табл. 7.1), початковий момент k -го порядку згідно з означенням математичного сподівання обчислюється за формулою (7.14):

$$v_k(X) = \sum_{i=1}^n (x_i^k \cdot p_i). \quad (7.14)$$

Для неперервної випадкової величини X , що має щільність розподілу $f(x)$ при $x \in (-\infty; \infty)$, початковий момент k -го порядку виражається інтегралом (7.15):

$$v_k(X) = \int_{-\infty}^{+\infty} (x^k \cdot f(x)) dx. \quad (7.15)$$

Розглянемо вирази для початкових моментів першого і другого порядків, що дасть змогу пов'язати початкові моменти з основними числовими характеристиками випадкової величини — математичним сподіванням і дисперсією (7.16):

$$v_1(X) = M(X), v_2(X) = M(X^2). \quad (7.16)$$

Тоді формулу (7.12) для обчислення дисперсії можна записати у вигляді (7.17):

$$D(X) = v_2(X) - v_1^2(X). \quad (7.17)$$

Центральним моментом порядку k випадкової величини X називається математичне сподівання величини $(X - M(X))^k$ (7.18):

$$\mu_k(X) = M(X - M(X))^k. \quad (7.18)$$

Для дискретної випадкової величини з рядом розподілу (табл. 7.1) маємо рівність (7.19):

$$\mu_k(X) = \sum_{i=1}^n ((x_i - M(X))^k \cdot p_i). \quad (7.19)$$

Для неперервної випадкової величини зі щільністю розподілу $f(x)$ при $x \in (-\infty; +\infty)$ маємо рівність (7.20):

$$\mu_k(X) = \int_{-\infty}^{+\infty} ((x - M(X))^k \cdot f(x)) dx, \quad (7.20)$$

Зауважимо, що відхилення $X - M(X)$ — також випадкова величина, але приведена до центру (середнього значення) можливих значень випадкової величини X . Тому величину $\check{X} = X - M(X)$ називають ще центрованою випадковою величиною, звідки походить назва центральних моментів.

Розглянемо центральні моменти 1-го, 2-го, 3-го і 4-го порядків, найчастіше застосовувані на практиці, і встановимо їхній зв'язок із початковими моментами.

За властивістю 5 математичного сподівання отримуємо рівність (7.21):

$$\mu_1(X) = M(X - M(X)) = 0, \quad (7.21)$$

а за означенням дисперсії — рівність (7.22):

$$\mu_2(X) = M(X - M(X))^2 = D(X). \quad (7.22)$$

Прирівнюючи (7.17) і (7.22), одержимо рівність (7.23):

$$\mu_2(X) = v_2(X) - v_1^2(X). \quad (7.23)$$

Розкриваючи степінь $(X - M(X))^3$, дістаємо вираз центрального момента 3-го порядку через початкові моменти (7.24):

$$\mu_3(X) = M(X - M(X))^3 = v_3(X) - 3 \cdot v_2(X) \cdot v_1(X) + 2 \cdot v_1^3(X), \quad (7.24)$$

Аналогічно, розкривши степінь $(X - M(X))^4$, дістанемо вираз центрального момента 4-го порядку через початкові моменти (7.25):

$$\mu_4(X) = v_4(X) - 4 \cdot v_3(X) \cdot v_1(X) + 6 \cdot v_2(X) \cdot v_1^2(X) - 3 \cdot v_1^4(X). \quad (7.25)$$

В загальному випадку маємо формулу (7.26):

$$\mu_k = \sum_{s=1}^k (-1)^s \cdot C_k^s \cdot v_{k-s}(X) \cdot v_1^s(X). \quad (7.26)$$

Приклад 7.8. Знайти початкові і центральні моменти 1-го, 2-го, 3-го і 4-го порядків дискретної випадкової величини X — кількості надійних приладів у системі, ряд розподілу якої наведено в табл. 7.7.

Таблиця 7.7. Ряд розподілу випадкової величини X

x_i	0	1	2	3
p_i	0,6	0,2	0,1	0,1

Розв'язання. Математичні сподівання випадкової величини X та її квадрата обчислено у прикладі 7.4:

$$M(X) = 0,7; \quad M(X^2) = 1,5.$$

Отже, за формулами (7.16)

$$v_1(X) = 0,7; \quad v_2(X) = 1,5;$$

$$v_3(X) = 0^3 \cdot 0,6 + 1^3 \cdot 0,2 + 2^3 \cdot 0,1 + 3^3 \cdot 0,1 = 3,7;$$

$$v_4(X) = 0^4 \cdot 0,6 + 1^4 \cdot 0,2 + 2^4 \cdot 0,1 + 3^4 \cdot 0,1 = 9,9.$$

Обчислимо центральні моменти, використовуючи формули (7.21)-(7.25).

Дисперсію знайдено в прикладі 7.4: $D(X) = 1,01$, тому:

$$\mu_1(X) = 0; \quad \mu_2(X) = 1,01;$$

$$\mu_3(X) = 3,7 - 3 \cdot 1,5 \cdot 0,7 + 2 \cdot (0,7)^3 = 4,386 - 3,15 = 1,236;$$

$$\mu_4(X) = 9,9 - 4 \cdot 3,7 \cdot 0,7 + 6 \cdot 1,5 \cdot (0,7)^2 - 3 \cdot (0,7)^4 = 14,31 - 11,0803 = 3,2297.$$

Відповідь. $v_1(X) = 0,7; \quad v_2(X) = 1,5; \quad v_3(X) = 3,7; \quad v_4(X) = 9,9; \quad \mu_1(X) = 0; \quad \mu_2(X) = 1,01; \quad \mu_3(X) = 1,236; \quad \mu_4(X) = 3,2297.$

Приклад 7.9. Задано функцію щільності ймовірностей

$$f(x) = \begin{cases} 0, & \text{при } x \leq -3; \\ \frac{3 \cdot (x+3) \cdot (1-x)}{32}, & \text{при } -3 < x \leq 1; \\ 0, & \text{при } x > 1. \end{cases}$$

Обчислити початкові та центральні моменти 2-го та 3-го порядків.

Розв'язання. За формулою (7.15) обчислюємо початкові моменти 2-го та 3-го порядків.

$$\begin{aligned} v_2(X) &= \int_{-3}^1 (x^2 \cdot f(x)) dx = \int_{-3}^1 \left(x^2 \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx = \frac{3}{32} \cdot \\ &\quad \cdot \int_{-3}^1 (3 \cdot x^2 - 2 \cdot x^3 - x^4) dx = \frac{9}{5}; \\ v_3(X) &= \int_{-3}^1 (x^3 \cdot f(x)) dx = \int_{-3}^1 \left(x^3 \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx = \frac{3}{32} \cdot \\ &\quad \cdot \int_{-3}^1 (3 \cdot x^3 - 2 \cdot x^4 - x^5) dx = -\frac{17}{5}. \end{aligned}$$

За формулою (7.20) обчислюємо початкові моменти 2-го та 3-го порядків.

$$\begin{aligned} M(X) &= \int_{-3}^1 (x \cdot f(x)) dx = \int_{-3}^1 \left(x \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx = \frac{3}{32} \cdot \\ &\quad \cdot \int_{-3}^1 (3 \cdot x - 2 \cdot x^2 - x^3) dx = -1; \\ \mu_2(X) &= \int_{-3}^1 ((x+1)^2 \cdot f(x)) dx = \int_{-3}^1 \left((x+1)^2 \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx \\ &= \frac{3}{32} \cdot \int_{-3}^1 (3 + 4 \cdot x - 2 \cdot x^2 - 4 \cdot x^3 - x^4) dx = \frac{4}{5}; \\ \mu_3(X) &= \int_{-3}^1 ((x+1)^3 \cdot f(x)) dx = \int_{-3}^1 \left((x+1)^3 \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx \\ &= \frac{3}{32} \cdot \int_{-3}^1 (3 + 7 \cdot x + 2 \cdot x^2 - 6 \cdot x^3 - 5 \cdot x^4 - x^5) dx = 0. \end{aligned}$$

Відповідь. $v_2(X) = 1,8$; $v_3(X) = -3,4$; $\mu_2(X) = 0,8$; $\mu_3(X) = 0$.

Приклад 7.10. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} \ln(x), & \text{при } x \in (1; e]; \\ 0, & \text{при } x \notin (1; e]. \end{cases}$$

Обчислити початкові та центральні моменти 1-го, 2-го, 3-го та 4-го порядків.

Розв'язання. Для обчислення інтегралів використаємо наступну формулу:

$$\int_1^e (x^n \cdot \ln(x)) dx = \frac{n \cdot e^{n+1} + 1}{(n+1)^2}.$$

Тоді початкові моменти дорівнюватимуть:

$$v_1(X) = \frac{e^2 + 1}{4} \approx 2,1; v_2(X) = \frac{2 \cdot e^3 + 1}{9} \approx 4,57;$$

$$v_3(X) = \frac{3 \cdot e^4 + 1}{16} \approx 10,3; v_4(X) = \frac{4 \cdot e^5 + 1}{25} \approx 23,79.$$

Центральні моменти обчислюються за формулами (7.21)-(7.25). Дисперсію даної випадкової величини знайдено у прикладі 7.5.

$$\mu_1(X) = 0; \mu_2(X) \approx 0,16;$$

$$\mu_3(X) = 10,3 - 3 \cdot 4,57 + 2 \cdot (2,1)^3 \approx 0,031;$$

$$\mu_4(X) = 23,79 - 4 \cdot 10,3 + 6 \cdot 4,57 \cdot (2,1)^2 - 3 \cdot (2,1)^4 \approx -0,152.$$

Відповідь. $v_1(X) \approx 2,1; v_2(X) \approx 4,57; v_3(X) \approx 10,3; v_4(X) \approx 23,79;$
 $\mu_1(X) = 0; \mu_2(X) \approx 0,16; \mu_3(X) \approx 0,031; \mu_4(X) \approx -0,152.$

7.4. Коефіцієнти асиметрії та ексцесу

Коефіцієнт асиметрії (Skewness) — числова характеристика розподілу ймовірностей дійсної випадкової величини. Коефіцієнт асиметрії характеризує асиметрію розподілу випадкової величини відносно її середнього значення (рис. 7.1).

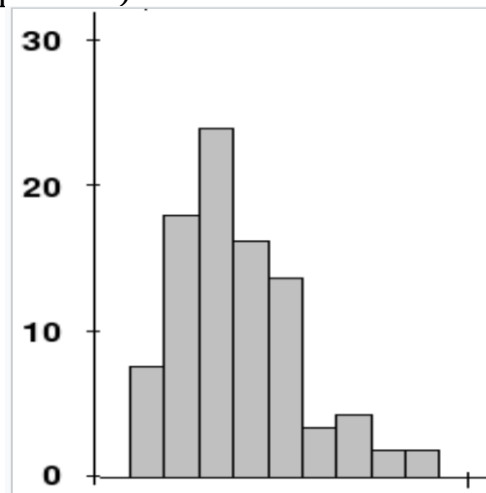


Рис. 7.1. Приклад експериментальних даних з ненульовою асиметрією

Асиметрією $As(X)$ (коефіцієнт асиметрії Фішера) теоретичного розподілу ймовірностей випадкової величини називають відношення центрального моменту третього порядку $\mu_3(X)$ до куба середнього квадратичного відхилення $\sigma^3(X)$ (7.27):

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)}, \quad (7.27)$$

$\mu_3(X)$ — центральний момент 3-го порядку;

$\sigma^3(X)$ — куб середньоквадратичного відхилення.

Властивості

1. Для еталонного (нормального) розподілу $\frac{\mu_3(X)}{\sigma^3(X)} = 0$; із чого випливає, що коефіцієнт асиметрії нормального розподілу дорівнює 0 (рис. 7.2);

2. Асиметрія додатна, якщо «довша частина» розподілу знаходиться праворуч від математичного сподівання; асиметрія від'ємна, якщо «довша частина» кривої знаходиться ліворуч від математичного сподівання;

3. На практиці, знак асиметрії визначають за положенням кривої відносно моди: якщо «довша» частина кривої знаходиться правіше моди, то асиметрія додатна, якщо лівіше — від'ємна (рис. 7.3).

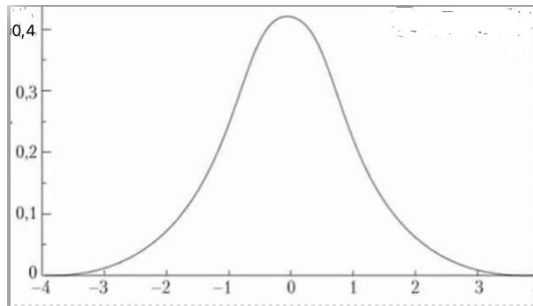


Рис. 7.2. Графік еталонного (нормального) розподілу

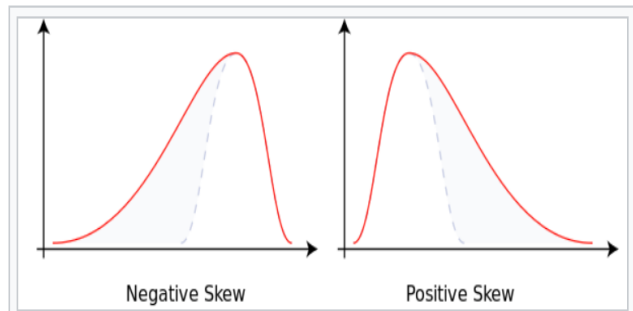


Рис. 7.3. Крива ліворуч має від'ємну асиметрію, крива праворуч — додатну асиметрію

Приклад 7.11. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} \ln(x), & \text{при } x \in (1; e]; \\ 0, & \text{при } x \notin (1; e]. \end{cases}$$

Обчислити коефіцієнт асиметрії для даної випадкової величини.

Розв'язання. В прикладі 7.10 обчислено $\mu_2(X) \approx 0,16$; $\mu_3(X) \approx 0,031$.

Обчислимо середньоквадратичне відхилення: $\sigma(X) = \sqrt{\mu_2(X)} \approx 0,4$.

Отже за формулою (7.27)

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} \approx \frac{0,031}{(0,4)^3} \approx 0,484.$$

Тобто «довша» частина кривої знаходиться правіше моди або математичного сподівання.

Відповідь. $As(X) \approx 0,484$.

Приклад 7.12. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} 0, & \text{при } x \leq -3; \\ \frac{3 \cdot (x + 3) \cdot (1 - x)}{32}, & \text{при } -3 < x \leq 1; \\ 0, & \text{при } x > 1. \end{cases}$$

Обчислити коефіцієнт асиметрії для даної випадкової величини.

Розв'язання. В прикладі 7.9 обчислено $\mu_2(X) \approx 0,8$; $\mu_3(X) = 0$.

Обчислимо середньоквадратичне відхилення: $\sigma(X) = \sqrt{\mu_2(X)} \approx 0,89$.

Отже за формулою (7.27)

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = 0.$$

Тобто крива є симетричною.

Правильність отриманих результатів можна побачити на графіку функції щільності (рис. 7.4).

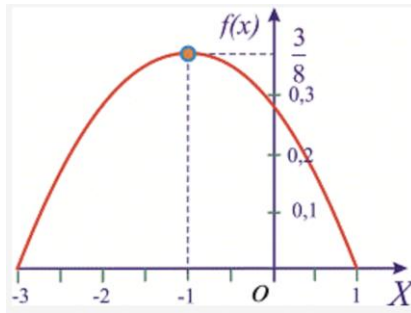


Рис. 7.4. Графік функції щільності $f(x)$

Відповідь. $As(X) = 0$

Коефіцієнт ексцесу (Kurtosis) — числова характеристика розподілу ймовірностей дійсної випадкової величини. Коефіцієнт ексцесу характеризує «крутість», тобто, стрімкість підвищення кривої розподілу у порівнянні з нормальною кривою.

Ексцесом $Es(X)$ теоретичного розподілу називають характеристику, що обчислюється за такою формулою (7.28):

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3, \quad (7.28)$$

$\mu_4(X)$ — центральний момент 4-го порядку;
 $\sigma^4(X)$ — квадрат дисперсії.

Властивості

1. Для еталонного (нормального) розподілу $\frac{\mu_4(X)}{\sigma^4(X)} = 3$; із чого випливає, що ексцес нормального розподілу дорівнює 0;
2. Якщо ексцес деякого розподілу відмінний від нуля, то крива щільності цього розподілу відрізняється від кривої щільності нормального розподілу: якщо ексцес додатний, то крива теоретичного має вищу та «гострішу» вершину ніж крива нормального; якщо ексцес від'ємний, то крива теоретичного має нижчу та «плоскішу» вершину ніж крива нормального. При цьому вважається, що нормальний і теоретичний розподіли мають однакові математичні сподівання та дисперсії;

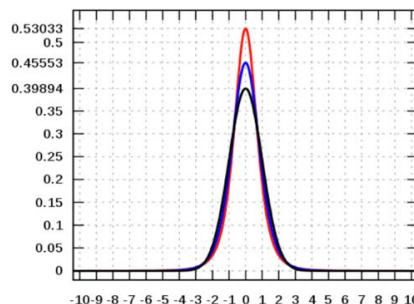


Рис. 7.5. Графік щільності розподілу для $Es(X) = +\infty$ (червона), $Es(X) = 2$ (синя) та $Es(X) = 0$ (чорна) криві

Приклад 7.13. Неперервна випадкова величина X має щільність розподілу

$$f(x) = \begin{cases} 0, & \text{при } x \leq -3; \\ \frac{3 \cdot (x+3) \cdot (1-x)}{32}, & \text{при } -3 < x \leq 1; \\ 0, & \text{при } x > 1. \end{cases}$$

Обчислити ексцес для даної випадкової величини.

Розв'язання: В прикладі 7.9 обчислено $\mu_2(X) \approx 0,8$, тобто $\sigma^2(X) = \mu_2(X) \approx 0,64$.

За формулою (7.20)

$$\begin{aligned} \mu_4(X) &= \int_{-3}^1 ((x+1)^4 \cdot f(x)) dx = \int_{-3}^1 \left((x+1)^4 \cdot \frac{3 \cdot (x+3) \cdot (1-x)}{32} \right) dx = \\ &= \left| \begin{array}{l} t = x+1; dx = dt \\ x = -3 \Rightarrow t = -2 \\ x = 1 \Rightarrow t = 2 \end{array} \right| = \frac{3}{32} \cdot \int_{-2}^2 (t^6 - 4 \cdot t^4) dt = \frac{48}{35} \approx 1,37. \end{aligned}$$

Отже за формулою (7.28),

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 \approx \frac{1,37}{0,4096} - 3 \approx 0,345.$$

Тобто крива теоретичного має вищу та «гострішу» вершину ніж крива нормального розподілу. Правильність отриманих результатів можна побачити на графіку функції щільності (рис. 7.1).

Відповідь. $Es(X) \approx 0,345$.

7.5. Квантиль

Квантиль — одна з числових характеристик випадкових величин, що здебільшого застосовується в математичній статистиці. Квантилі відсікають в межах ряду певну частину його членів.

Визначення

Нехай маємо ймовірнісний простір $(\Omega, \mathcal{F}, P)$, і P^X — ймовірнісна міра, що задає розподіл деякої випадкової величини X . Нехай зафіксовано $\alpha \in [0,1]$. Тоді α -квантилем (або квантилем рівня α) розподілу P^X називається число $x_\alpha \in R$, таке, що справедлива рівність (7.29):

$$P^X\{X \in (-\infty, x_\alpha]\} \equiv P(X \leq x_\alpha) = \alpha. \quad (7.29)$$

Зауваження

- Якщо розподіл неперервний, то α -квантиль однозначно задається рівнянням (7.30)

$$F(x_\alpha) = \alpha, \quad (7.30)$$

де $F(x)$ — функція розподілу P^X .

- Очевидно що для неперервних розподілів справедлива рівність (7.31):

$$P\left(x_{\frac{1-\alpha}{2}} \leq X \leq x_{\frac{1+\alpha}{2}}\right) = \alpha, \quad (7.31)$$

що широко використовується при побудові довірчих інтервалів

Види квантилів

Квантиль порядку $\alpha = 0,25$ називається *першим (або нижнім) квантилем*;

Кантиль порядку $\alpha = 0,5$ називається *медіаною (або другим квантилем)*;

Кантиль порядку $\alpha = 0,75$ називається *третім (або верхнім) квантилем*.

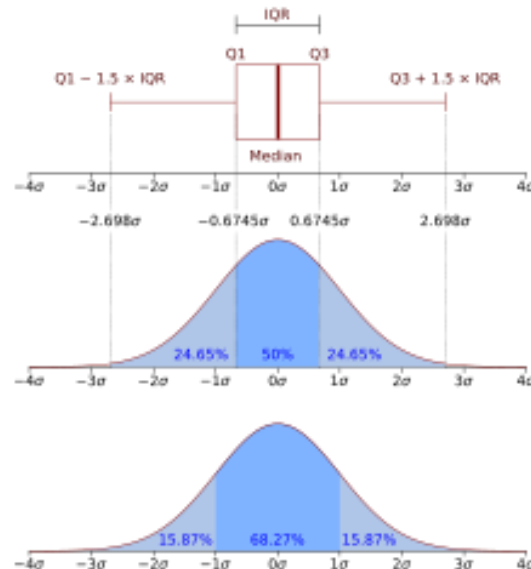


Рис. 7.6. Квартілі нормального розподілу

Інтерквартильним або *міжквартильним розмахом* (Interquartile range) називається різниця між третім і першим квантилем, тобто $x_{0,75} - x_{0,25}$. Інтерквартильний розмах є характеристикою розкиду розподілу величини. Разом медіана і інтерквартильний розмах можуть бути використані замість математичного сподівання і дисперсії у разі розподілів з великими викидами, або при неможливості обчислення останніх.

Дециль характеризує розподіл величин сукупності, при якій дев'ять значень дециля ділять її на десять рівних частин. Будь-яка з цих десяти частин становить 0,1 всієї сукупності. Так, перший дециль відокремлює 10% найменших величин, лежачих нижче дециля від 90% найбільших величин, лежачих вище дециля.

p-им перцентилем називають квантиль рівня $\alpha = \frac{p}{100}$. При цьому зазвичай розглядають перцентилі для цілих p , хоча дана вимога не обов'язкова. Відповідно, медіана є 50-м перцентилем, а перший і третій квартиль — 25-м і 75-м перцентилем. У цілому, поняття квантиль і перцентиль взаємозамінні, також, як і шкали числення ймовірності — абсолютна і процентна. Перцентилі також називаються процентилями або центилями.

Квантилі основних типів розподілів

Як зазначалося вище, квантилі як дискретних так і неперервних випадкових величин обчислюються з рівняння (7.30). Для цього необхідно знати функцію розподілу випадкової величини. Для неперервних величин, які мають досить складні функції розподілу, наперед розраховані значення квантилів. Деякі представлені на рис.7.7-7.9.

Ймовірність,%	99,99	99,90	99,00	97,72	97,50	95,00	90,00	84,13	50,00
Квантиль	3,715	3,090	2,326	2,000	1,960	1,645	1,282	1,000	0,000

Рис. 7.7. Квантилі стандартного нормального розподілу

	0.01	0.05	0.95	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.975	0.99
1	0.0002	0.0010	0.0039	0.0108	0.0242	0.0485	0.0750	0.1055	0.1395	0.1742	0.2090	0.2448	0.2815	0.3090
2	0.0044	0.0088	0.0128	0.0207	0.0440	0.0733	0.1027	0.1383	0.1808	0.2269	0.2748	0.3240	0.3737	0.4030
3	0.0148	0.0287	0.0419	0.0644	0.1084	0.1650	0.2242	0.2862	0.3520	0.4214	0.4942	0.5693	0.6454	0.6930
4	0.0291	0.0544	0.0798	0.1206	0.1948	0.2847	0.3797	0.4798	0.5839	0.6919	0.7994	0.9057	1.0103	1.0767
5	0.0443	0.0832	0.1165	0.1610	0.2425	0.2999	0.3655	0.4315	0.5018	0.5764	0.6544	0.7350	0.8185	0.8683
6	0.0571	0.1037	0.1401	0.1875	0.2735	0.3270	0.3912	0.4570	0.5261	0.5994	0.6760	0.7550	0.8363	0.8880
7	0.0685	0.1253	0.1651	0.2160	0.2960	0.3500	0.4142	0.4800	0.5491	0.6222	0.6994	0.7796	0.8619	0.9150
8	0.0786	0.1405	0.1825	0.2360	0.3150	0.3690	0.4332	0.5000	0.5691	0.6422	0.7194	0.7996	0.8819	0.9360
9	0.0877	0.1535	0.1975	0.2530	0.3320	0.3860	0.4502	0.5170	0.5861	0.6592	0.7364	0.8166	0.8989	0.9540
10	0.0960	0.1655	0.2115	0.2680	0.3470	0.4010	0.4652	0.5320	0.6011	0.6742	0.7514	0.8316	0.9139	0.9700
11	0.1036	0.1760	0.2230	0.2800	0.3590	0.4130	0.4772	0.5440	0.6131	0.6862	0.7634	0.8436	0.9259	0.9830
12	0.1105	0.1850	0.2330	0.2900	0.3690	0.4230	0.4872	0.5540	0.6231	0.6962	0.7734	0.8536	0.9359	0.9940
13	0.1168	0.1930	0.2410	0.3000	0.3790	0.4330	0.4972	0.5640	0.6331	0.7062	0.7834	0.8636	0.9459	1.0050
14	0.1225	0.2000	0.2480	0.3080	0.3870	0.4410	0.5052	0.5720	0.6411	0.7142	0.7914	0.8716	0.9539	1.0140
15	0.1277	0.2060	0.2540	0.3150	0.3940	0.4480	0.5122	0.5790	0.6481	0.7212	0.7984	0.8786	0.9609	1.0220
16	0.1324	0.2110	0.2590	0.3200	0.4000	0.4540	0.5182	0.5850	0.6541	0.7272	0.8044	0.8846	0.9669	1.0290
17	0.1367	0.2150	0.2630	0.3240	0.4040	0.4580	0.5222	0.5890	0.6581	0.7312	0.8084	0.8886	0.9709	1.0340
18	0.1406	0.2180	0.2660	0.3270	0.4070	0.4610	0.5252	0.5920	0.6611	0.7342	0.8114	0.8916	0.9739	1.0380
19	0.1442	0.2210	0.2690	0.3300	0.4100	0.4640	0.5282	0.5950	0.6641	0.7372	0.8144	0.8946	0.9769	1.0420
20	0.1475	0.2240	0.2720	0.3330	0.4130	0.4670	0.5312	0.5980	0.6671	0.7402	0.8174	0.8976	0.9799	1.0460
21	0.1507	0.2270	0.2750	0.3360	0.4160	0.4700	0.5342	0.6010	0.6701	0.7432	0.8204	0.9006	0.9829	1.0490
22	0.1538	0.2300	0.2780	0.3390	0.4190	0.4730	0.5372	0.6040	0.6731	0.7462	0.8234	0.9036	0.9859	1.0520
23	0.1568	0.2330	0.2810	0.3420	0.4220	0.4760	0.5402	0.6070	0.6761	0.7492	0.8264	0.9066	0.9889	1.0550
24	0.1597	0.2360	0.2840	0.3450	0.4250	0.4790	0.5432	0.6100	0.6791	0.7522	0.8294	0.9096	0.9919	1.0580
25	0.1625	0.2390	0.2870	0.3480	0.4280	0.4820	0.5462	0.6130	0.6821	0.7552	0.8324	0.9126	0.9949	1.0610
26	0.1653	0.2420	0.2900	0.3510	0.4310	0.4850	0.5492	0.6160	0.6851	0.7582	0.8354	0.9156	0.9979	1.0640
27	0.1680	0.2450	0.2930	0.3540	0.4340	0.4880	0.5522	0.6190	0.6881	0.7612	0.8384	0.9186	1.0009	1.0670
28	0.1707	0.2480	0.2960	0.3570	0.4370	0.4910	0.5552	0.6220	0.6911	0.7642	0.8414	0.9216	1.0039	1.0700
29	0.1734	0.2510	0.2990	0.3600	0.4400	0.4940	0.5582	0.6250	0.6941	0.7672	0.8444	0.9246	1.0069	1.0730
30	0.1760	0.2540	0.3020	0.3630	0.4430	0.4970	0.5612	0.6280	0.6971	0.7702	0.8474	0.9276	1.0099	1.0760
31	0.1786	0.2570	0.3050	0.3660	0.4460	0.5000	0.5642	0.6310	0.7001	0.7732	0.8504	0.9306	1.0129	1.0790
32	0.1812	0.2600	0.3080	0.3690	0.4490	0.5030	0.5672	0.6340	0.7031	0.7762	0.8534	0.9336	1.0159	1.0820
33	0.1838	0.2630	0.3110	0.3720	0.4520	0.5060	0.5702	0.6370	0.7061	0.7792	0.8564	0.9366	1.0189	1.0850
34	0.1863	0.2660	0.3140	0.3750	0.4550	0.5090	0.5732	0.6400	0.7091	0.7822	0.8594	0.9396	1.0219	1.0880
35	0.1888	0.2690	0.3170	0.3780	0.4580	0.5120	0.5762	0.6430	0.7121	0.7852	0.8624	0.9426	1.0249	1.0910
36	0.1913	0.2720	0.3200	0.3810	0.4610	0.5150	0.5792	0.6460	0.7151	0.7882	0.8654	0.9456	1.0279	1.0940
37	0.1938	0.2750	0.3230	0.3840	0.4640	0.5180	0.5822	0.6490	0.7181	0.7912	0.8684	0.9486	1.0309	1.0970
38	0.1962	0.2780	0.3260	0.3870	0.4670	0.5210	0.5852	0.6520	0.7211	0.7942	0.8714	0.9516	1.0339	1.1000
39	0.1987	0.2810	0.3290	0.3900	0.4700	0.5240	0.5882	0.6550	0.7241	0.7972	0.8744	0.9546	1.0369	1.1030
40	0.2011	0.2840	0.3320	0.3930	0.4730	0.5270	0.5912	0.6580	0.7271	0.8002	0.8774	0.9576	1.0399	1.1060
41	0.2035	0.2870	0.3350	0.3960	0.4760	0.5300	0.5942	0.6610	0.7301	0.8032	0.8804	0.9606	1.0429	1.1090
42	0.2059	0.2900	0.3380	0.3990	0.4790	0.5330	0.5972	0.6640	0.7331	0.8062	0.8834	0.9636	1.0459	1.1120
43	0.2082	0.2930	0.3410	0.4020	0.4820	0.5360	0.6002	0.6670	0.7361	0.8092	0.8864	0.9666	1.0489	1.1150
44	0.2106	0.2960	0.3440	0.4050	0.4850	0.5390	0.6032	0.6700	0.7391	0.8122	0.8894	0.9696	1.0519	1.1180
45	0.2129	0.2990	0.3470	0.4080	0.4880	0.5420	0.6062	0.6730	0.7421	0.8152	0.8924	0.9726	1.0549	1.1210
46	0.2152	0.3020	0.3500	0.4110	0.4910	0.5450	0.6092	0.6760	0.7451	0.8182	0.8954	0.9756	1.0579	1.1240
47	0.2175	0.3050	0.3530	0.4140	0.4940	0.5480	0.6122	0.6790	0.7481	0.8212	0.8984	0.9786	1.0609	1.1270
48	0.2198	0.3080	0.3560	0.4170	0.4970	0.5510	0.6152	0.6820	0.7511	0.8242	0.9014	0.9816	1.0639	1.1300
49	0.2221	0.3110	0.3590	0.4200	0.5000	0.5540	0.6182	0.6850	0.7541	0.8272	0.9044	0.9846	1.0669	1.1330
50	0.2244	0.3140	0.3620	0.4230	0.5030	0.5570	0.6212	0.6880	0.7571	0.8302	0.9074	0.9876	1.0699	1.1360

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0,1, & \text{при } 1 < x \leq 2; \\ 0,5, & \text{при } 2 < x \leq 3; \\ 0,8, & \text{при } 3 < x \leq 4; \\ 1, & \text{при } x > 4. \end{cases}$$

Тоді, з рівняння (7.30), $x_{0,5} = Me(X) = 2,5$.

Відповідь. $x_{0,5} = 2,5$.

Приклад 7.15. Знайти квантиль порядку 0,9 дискретної випадкової величини X (табл. 7.9).

Таблиця 7.9. Ряд розподілу випадкової величини X

x_i	1	2	3	5	6
p_i	0,1	0,3	0,2	0,3	0,1

Розв'язання. Складемо функцію розподілу випадкової величини X .

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0,1, & \text{при } 1 < x \leq 2; \\ 0,4, & \text{при } 3 < x \leq 3; \\ 0,6, & \text{при } 3 < x \leq 5; \\ 0,9, & \text{при } 5 < x \leq 6; \\ 1, & \text{при } x > 6. \end{cases}$$

Тоді, з рівняння (7.30), $x_{0,9} = 5,5$.

Відповідь. $x_{0,9} = 5,5$.

Приклад 7.16. Неперервна випадкова величина задана функцією розподілу

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ \frac{1}{2} \cdot \sqrt[3]{x-1}, & \text{при } 1 < x \leq 9; \\ 1, & \text{при } x > 9. \end{cases}$$

Знайти квантиль порядку 0,75.

Розв'язання. З рівняння (7.30)

$$\begin{aligned} \frac{1}{2} \cdot \sqrt[3]{x_{0,75} - 1} &= 0,75; \\ \sqrt[3]{x_{0,75} - 1} &= 1,5; \\ x_{0,75} - 1 &= 3,375; \\ x_{0,75} &= 4,375. \end{aligned}$$

Відповідь. $x_{0,9} = 4,375$.

7.6. Твірна функція та її застосування для знаходження числових характеристик дискретної випадкової величини

Твірна функція випадкової величини (generating function) — це аналітичний інструмент, який використовується для вивчення властивостей дискретної випадкової величини. Вона зберігає інформацію про всі моменти випадкової величини і може використовуватись для отримання різних характеристик розподілу цієї величини.

Побудова твірної функції для цілочислові випадкові величини (відомий ряд розподілу)

Відомо, що серед дискретних випадкових величин важливе значення в теорії ймовірностей займають такі, що набувають лише цілих невід'ємних значень $X = 0, 1, 2, 3, \dots$ з ймовірностями p_i ($i = 0, 1, \dots, n \dots$), інакше кажучи – *цілочислові випадкові величини*. Для дослідження законів розподілу цілочислових випадкових величин використовують ймовірнісну твірну функцію.

Ймовірнісною твірною функцією $\varphi(z)$ для цілочислові випадкової величини X наступний збіжний степеневий ряд (7.32):

$$\varphi(z) = \sum_{i=1}^{+\infty} (p_i \cdot z^i), \quad (7.32)$$

де z — довільний параметр ($0 < z < 1$).

Якщо випадкова величина має скінченну множину можливих значень n , то твірна функція приймає вигляд (7.33):

$$\varphi(z) = \sum_{i=1}^n (p_i \cdot z^i). \quad (7.33)$$

Властивості ймовірнісної твірної функції

1. Твірна функція $\varphi(z)$ визначена в кожній точці інтервалу $[-1; 1]$.
2. При $z = 1$ виконується співвідношення $\varphi(1) = \sum_{i=0}^{+\infty} p_i = 1$. Дане співвідношення є умовою нормування для дискретної випадкової величини.
3. Із залежності для твірної функції $\varphi(z)$ визначають ймовірність $P(X = i) = P(i)$ з рівності (7.34):

$$P(i) = \frac{1}{i!} \cdot \varphi^{(i)}(0), \quad (7.34)$$

де $\varphi^{(i)}(0)$ — i -та похідна від твірної функції $\varphi(z)$ при $z = 0$. Отже, знаючи аналітичний вираз для $\varphi(z)$, завжди можна визначити ймовірність будь-якого можливого значення $X = i$.

4. Математичне сподівання дискретної випадкової величини дорівнює (7.35):

$$M(X) = \varphi'(1). \quad (7.35)$$

Доведення. Похідна від твірної функції визначається співвідношенням

$$\varphi'(z) = \sum_{i=0}^{+\infty} (i \cdot p_i \cdot z^{i-1}).$$

При $z = 1$ похідна рівна математичному сподіванню

$$\varphi'(1) = \sum_{i=0}^{+\infty} (i \cdot p_i) = M(X) = v_1(x).$$

Звідси отримаємо $M(X) = \varphi'(1)$.

5. Дисперсія дискретної випадкової величини дорівнює (7.36):

$$D(X) = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2. \quad (7.36)$$

Доведення. Друга похідна твірної функції

$$\varphi''(z) = \sum_{i=0}^{+\infty} (i \cdot (i-1) \cdot p_i \cdot z^{i-2})$$

при $z = 1$ приймає значення

$$\begin{aligned} \varphi''(1) &= \sum_{i=0}^{+\infty} (i \cdot (i-1) \cdot p_i) = \varphi''(z) = \sum_{i=0}^{+\infty} (i^2 \cdot p_i) - \sum_{i=0}^{+\infty} (i \cdot p_i) = \\ &= v_2(x) - v_1(x). \end{aligned}$$

На основі цього виразу, дисперсія випадкової величини через значення похідних твірної функції при $z = 1$ подається таким чином:

$$\begin{aligned} D(X) &= v_2(x) - v_1^2(x) = v_2(x) - v_1(x) + v_1(x) - v_1^2(x) = \\ &= \varphi''(1) + \varphi'(1) - (\varphi'(1))^2. \end{aligned}$$

Формули (7.35) і (7.36) застосовуються для знаходження основних числових характеристик дискретних випадкових величин, які мають найбільш поширені закони розподілу.

6. Якщо випадкова величина набуває лише цілих невід'ємних значень $X = C, C+1, C+2, C+3, \dots$ з ймовірностями p_i ($i = 0, 1, \dots, n, \dots$), варто розглянути допоміжну випадкову величину Y , таку, що $X = Y + C$ і досліджувати її твірну функцію. Це необхідно для отримання числових характеристик дискретних випадкових величин різних типів за допомогою їх твірних функцій.

Приклад 7.19. Для дискретної випадкової величини X (табл. 7.10)

Таблиця 7.10. Ряд розподілу випадкової величини X

x_i	0	1	2	3	4
p_i	0,1	0,2	0,25	0,3	0,15

1. Побудувати твірну функцію; **2.** Обчислити математичне сподівання та дисперсію випадкової величини X .

Розв'язання. 1. За формулою (7.33)

$$\varphi(z) = 0,1 + 0,2 \cdot z + 0,25 \cdot z^2 + 0,3 \cdot z^3 + 0,15 \cdot z^4;$$

$$\varphi'(z) = 0,2 + 0,5 \cdot z + 0,9 \cdot z^2 + 0,6 \cdot z^3;$$

$$\mathbf{2.} \quad M(X) = \varphi'(1) = 0,2 + 0,5 + 0,9 + 0,6 = 2,2;$$

$$\varphi''(z) = 0,5 + 1,8 \cdot z + 1,8 \cdot z^2;$$

$$\varphi''(1) = 0,5 + 1,8 + 1,8 = 4,1;$$

$$D(X) = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2 = 4,1 + 2,2 - 4,84 = 1,46.$$

$$\text{Перевірка: } M(X) = 0 \cdot 0,1 + 1 \cdot 0,2 + 2 \cdot 0,25 + 3 \cdot 0,3 + 4 \cdot 0,15 = 2,2;$$

$$D(X) = 0^2 \cdot 0,1 + 1^2 \cdot 0,2 + 2^2 \cdot 0,25 + 3^2 \cdot 0,3 + 4^2 \cdot 0,15 - 4,84 = 1,46.$$

Відповідь. 1. $\varphi(z) = 0,1 + 0,2 \cdot z + 0,25 \cdot z^2 + 0,3 \cdot z^3 + 0,15 \cdot z^4$; 2. $M(X) = 2,2$; $D(X) = 1,46$.

Приклад 7.20. Для дискретної випадкової величини X (табл. 7.11)

Таблиця 7.11. Ряд розподілу випадкової величини X

x_i	1	2	3	4	5
p_i	0,1	0,2	0,25	0,3	0,15

Обчислити математичне сподівання.

Розв'язання. За властивістю 6 твірної функції очевидно, що $X = Y + 1$ (табл. 7.12).

Таблиця 7.12. Ряд розподілу випадкової величини Y

y_i	0	1	2	3	4
p_i	0,1	0,2	0,25	0,3	0,15

Тоді для величини Y формулою (7.33) будемо твірну функцію

$$\varphi(z) = 0,1 + 0,2 \cdot z + 0,25 \cdot z^2 + 0,3 \cdot z^3 + 0,15 \cdot z^4,$$

з якої знаходимо математичне сподівання:

$$\varphi'(z) = 0,2 + 0,5 \cdot z + 0,9 \cdot z^2 + 0,6 \cdot z^3;$$

$$M(Y) = \varphi'(1) = 0,2 + 0,5 + 0,9 + 0,6 = 2,2.$$

Таким чином, за властивостями 1 та 2 математичного сподівання

$$M(X) = M(Y) + 1 = 3,2.$$

Перевірка: $M(X) = 1 \cdot 0,1 + 2 \cdot 0,2 + 3 \cdot 0,25 + 4 \cdot 0,3 + 5 \cdot 0,15 = 3,2$.

Відповідь. $M(X) = 3,2$.

Побудова твірної функції для множини несумісних подій з відомими ймовірностями їх настання в серії незалежних випробувань (невідомий ряд розподілу)

Розглянемо варіант, коли маємо множину несумісних подій, які виконуються при незалежних випробуваннях з ймовірностями p_i ($i = 1, 2, \dots, n \dots$). Тоді твірна функція визначається із співвідношення (7.37):

$$\varphi(z) = \prod_{i=1}^{+\infty} (q_i + p_i \cdot z), \quad (7.37)$$

де $q_i = 1 - p_i$.

Формула (7.37) застосовується для швидшого знаходження ймовірності $P(k)$ (властивість 3 твірної функції) без необхідності побудови ряду розподілу випадкової величини.

Приклад 7.21. Два стрілки роблять по одному пострілу по мішені. Ймовірності попадання 1-го і 2-го стрілків відповідно дорівнюють: $p_1 = 0,7$; $p_2 = 0,8$. Побудувати твірну функцію для випадкової величини X – кількість попадань по мішені.

Розв'язання. Складемо ряд розподілу для випадкової величини X .

$$P(0) = q_1 \cdot q_2 = 0,3 \cdot 0,2 = 0,06;$$

$$P(1) = p_1 \cdot q_2 + q_1 \cdot p_2 = 0,7 \cdot 0,2 + 0,3 \cdot 0,8 = 0,38;$$

$$P(2) = p_1 \cdot p_2 = 0,7 \cdot 0,8 = 0,56.$$

Ряд розподілу випадкової величини X матиме вигляд (табл. 7.13):

Таблиця 7.13. Ряд розподілу випадкової величини X

x_i	0	1	2
p_i	0,06	0,38	0,56

Твірну функцію побудуємо двома способами.

1 спосіб

Оскільки випадкова величина цілочислова, то формулою (7.32)

$$\varphi(z) = 0,06 + 0,38 \cdot z + 0,56 \cdot z^2.$$

2 спосіб

Оскільки маємо множину з трьох несумісних подій: $A_1 = \{\text{жодного попадання}\}$, $A_2 = \{\text{рівно одне попадання}\}$, $A_3 = \{\text{рівно два попадання}\}$, які можуть настати в серії з двох експериментів: $\{1\text{-й стрілок зробив постріл}\}$, $\{2\text{-й стрілок зробив постріл}\}$. Ймовірності влучання стрілків $p_1 = 0,7$; $p_2 = 0,8$, а невлучання $q_1 = 0,3$; $q_2 = 0,2$. Тоді за формулою (7.37)

$$\varphi(z) = (0,3 + 0,7 \cdot z) \cdot (0,2 + 0,8 \cdot z) = 0,06 + 0,38 \cdot z + 0,56 \cdot z^2.$$

Відповідь. $\varphi(z) = 0,06 + 0,38 \cdot z + 0,56 \cdot z^2$.

ЛЕКЦІЯ 8

ОСНОВНІ ЗАКОНИ РОЗПОДІЛУ ДИСКРЕТНИХ ВИПАДКОВИХ ВЕЛИЧИН

Розглянемо основні, найбільш поширені в практичних застосуваннях розподіли випадкових величин та їх числові характеристики, обчислення яких для дискретних випадкових величин суттєво спрощується при застосуванні твірної функції для випадкової величини.

8.1. Біноміальний розподіл

Біноміальний розподіл — це один із базових дискретних розподілів ймовірностей, який описує кількість успіхів у послідовності з фіксованого числа незалежних випробувань, де кожне випробування має лише два можливі результати: успіх або невдача.

Біноміальний закон розподілу виникає в схемі Бернуллі, коли проводиться n незалежних випробувань, в кожному з яких випадкова подія A може або відбутися з сталою ймовірністю p , або не відбутися з ймовірністю $q = 1 - p$, а випадкова величина X — кількість появ події A в цих n випробуваннях.

Дискретна величина X приймає цілі невід'ємні значення $x_i = i$ ($i = 0, 1, 2, \dots, t, \dots, n$) з ймовірностями $p_i = P_n\{x_i = i\} = P_n(i)$, які обчислюються за формулою Бернуллі (8.1):

$$P_n(i) = C_n^i \cdot p^i \cdot q^{n-i}, \quad (8.1)$$

Для отримання формули (8.1) в формулі (4.1) лекції 4 підставляємо $m = i$.

Таким чином, величина X має ряд розподілу (табл. 8.1):

Таблиця 8.1. Ряд розподілу випадкової величини X , що має біноміальний розподіл

i	0	1	2	...	m	...	n
$P_n(i)$	q^n	$C_n^1 \cdot p \cdot q^{n-1}$	$C_n^2 \cdot p^2 \cdot q^{n-2}$	...	$C_n^m \cdot p^m \cdot q^{n-m}$	...	p^n

Ймовірності $P_n(i)$ дорівнюють відповідним членам розкладу n -го степеню бінома (8.2)

$$\sum_{i=0}^n P_n(i) = (q + p)^n = q^n + C_n^1 \cdot p \cdot q^{n-1} + C_n^2 \cdot p^2 \cdot q^{n-2} + \dots + C_n^m \cdot p^m \cdot q^{n-m} + \dots + p^n = \sum_{i=0}^n (C_n^i \cdot p^i \cdot q^{n-i}). \quad (8.2)$$

Очевидно, що біноміальний розподіл коректний, оскільки з (8.2) випливає, що $(q + p)^n = 1^n = 1$, звідки й походить назва цього розподілу.

Властивості біноміального розподілу

1. Кожне випробування повинно бути незалежним від інших, тобто результат одного випробування не впливає на результат іншого.

2. Якщо $p = 0,5$, біноміальний розподіл буде симетричним. Якщо $p \neq 0,5$, розподіл буде зміщеним в бік більш ймовірного результату (успіху або невдачі).

3. Біноміальний розподіл є цілочисловим дискретним розподілом.

4. Функція розподілу $F(x)$ випадкової величини будується за принципом, наведеним в лекції 6.

5. Математичне сподівання біноміального розподілу дорівнює (8.3):

$$M(X) = n \cdot p, \quad (8.3)$$

дисперсія дорівнює (8.4):

$$D(X) = n \cdot p \cdot q \quad (8.4)$$

і, відповідно, середнє квадратичне відхилення (8.5):

$$\sigma(X) = \sqrt{n \cdot p \cdot q}. \quad (8.5)$$

Доведення. Для знаходження основних характеристик $M(X), D(X), \sigma(X)$ біноміального розподілу використовується твірна функція (формула (7.33) лекції 7).

$$\varphi(z) = \sum_{i=1}^n \left((C_n^i \cdot p^i \cdot q^{n-i}) \cdot z^i \right) = \sum_{i=1}^n (C_n^i \cdot (p \cdot z)^i \cdot q^{n-i}),$$

яка у відповідності з формулою (8.1) матиме вигляд

$$\varphi(z) = (q + p \cdot z)^n.$$

У випадку біноміального розподілу для отримання твірної функції також можна використати формулу (7.37) лекції 7.

Знайдемо першу похідну по z твірної функції

$$\varphi'(z) = n \cdot (q + p \cdot z)^{n-1} \cdot p$$

та її значення при $z = 1$

$$M(X) = \varphi'(1) = n \cdot (q + p)^{n-1} \cdot p = n \cdot p.$$

Знайдемо другу похідну твірної функції

$$\varphi''(z) = n \cdot (n-1) \cdot (q + p \cdot z)^{n-2} \cdot p^2$$

та її значення при $z = 1$

$$\varphi''(1) = n \cdot (n-1) \cdot (q + p)^{n-2} \cdot p^2 = n \cdot (n-1) \cdot p^2.$$

Тоді одержимо дисперсію

$$D(X) = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2 = n \cdot (n-1) \cdot p^2 + n \cdot p - (n \cdot p)^2 = n \cdot p \cdot q.$$

Тоді, середньоквадратичне відхилення:

$$\sigma(X) = \sqrt{D(X)} = \sqrt{n \cdot p \cdot q}.$$

6. Математичне сподівання частоти $\frac{i}{n}$ появи події в серії з n незалежних випробувань, в кожному з яких подія настає зі сталою ймовірністю p , дорівнює самій ймовірності p (8.6).

$$M\left(\frac{i}{n}\right) = p, \quad (8.6)$$

а дисперсія (8.7)

$$D(X) = \frac{p \cdot q}{n}. \quad (8.7)$$

Доведення. Нехай деяка подія в серії з n випробувань настала i разів. Тоді частота настання події $\frac{i}{n}$ є випадковою величиною $\frac{X}{n}$, тобто $\frac{i}{n} = \frac{X}{n}$, де X – випадкова величина, що має біноміальний розподіл. Тоді за властивостями математичного сподівання та дисперсії отримуємо:

$$\begin{aligned} M\left(\frac{i}{n}\right) &= M\left(\frac{X}{n}\right) = \frac{1}{n} \cdot M(X) = \frac{n \cdot p}{n} = p; \\ D\left(\frac{i}{n}\right) &= D\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot D(X) = \frac{n \cdot p \cdot q}{n^2} = \frac{p \cdot q}{n}. \end{aligned}$$

7. *Мода біноміального розподілу співпадає з найімовірнішим числом i визначається з нерівності (8.8)*

$$n \cdot p - q \leq Mo(X) \leq n \cdot p + p. \quad (8.8)$$

Біноміальний закон розподілу широко використовується в теорії та практиці статистичного контролю якості продукції, при описі функціонування систем масового обслуговування, при моделюванні цін активів, теорії стрільби та в інших областях.

8. *Медіана біноміального розподілу обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $(a; b]$ дорівнює $\frac{1}{2}$, згідно формули (6.20) лекції 6 $(F(Me(X)) = \frac{1}{2})$.*

9. *Коефіцієнт асиметрії та ексцес біноміального розподілу відповідно дорівнюють (8.9):*

$$As(X) = \frac{1-2 \cdot p}{\sqrt{n \cdot p \cdot q}}; \quad Es(X) = \frac{1-6 \cdot p \cdot q}{n \cdot p \cdot q}. \quad (8.9)$$

Доведення.

$$\begin{aligned} \mu_3(X) &= M(X - M(X))^3 = n \cdot p \cdot q \cdot (1 - 2 \cdot p); \\ \sigma(X) &= \sqrt{n \cdot p \cdot q}; \\ \mu_4(X) &= M(X - M(X))^4 = 3 \cdot (n \cdot p \cdot q)^2 + n \cdot p \cdot q \cdot (1 - 6 \cdot p \cdot q); \\ As(X) &= \frac{\mu_3(X)}{\sigma^3(X)} = \frac{n \cdot p \cdot q \cdot (1 - 2 \cdot p)}{(\sqrt{n \cdot p \cdot q})^3} = \frac{1 - 2 \cdot p}{\sqrt{n \cdot p \cdot q}}; \\ Es(X) &= \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{3 \cdot (n \cdot p \cdot q)^2 + n \cdot p \cdot q \cdot (1 - 6 \cdot p \cdot q)}{(n \cdot p \cdot q)^2} - 3 = \frac{1 - 6 \cdot p \cdot q}{n \cdot p \cdot q}. \end{aligned}$$

Приклад 8.1. *За метеорологічних умов, що склалися в аеропорту, ймовірність своєчасного відправлення кожного рейсу дорівнює 0,6. Випадкова величина X – кількість своєчасно відправлених рейсів з трьох передбачених розкладом. Скласти ряд розподілу, знайти числові*

характеристики цієї випадкової величини та коефіцієнт асиметрії та ексцес.

Розв'язання. Випадкова величина X приймає можливі значення 0,1,2,3, а відповідні ймовірності обчислюються за формулою Бернуллі при $n = 3$; $p = 0,6$; $q = 0,4$:

$$p_0 = P_3(0) = (0,4)^3 = 0,064; p_1 = P_3(1) = C_3^1 \cdot (0,6) \cdot (0,4)^2 = 0,288;$$

$$p_2 = P_3(2) = C_3^2 \cdot (0,6)^2 \cdot (0,4) = 0,432; p_3 = P_3(3) = (0,6)^3 = 0,216.$$

Отже, випадкова величина має ряд розподілу (табл. 8.2).

Таблиця 8.2. Ряд розподілу

i	0	1	2	3
$P_n(i)$	0,064	0,288	0,432	0,216

Математичне сподівання і дисперсія обчислюються за формулами (8.3) і (8.4):

$$M(X) = n \cdot p = 1,8; D(X) = n \cdot p \cdot q = 0,72; \sigma(X) \approx 0,85.$$

Перевірка:

$$M(X) = 0 \cdot 0,064 + 1 \cdot 0,288 + 2 \cdot 0,432 + 3 \cdot 0,216 = 1,8;$$

$$M(X^2) = 0^2 \cdot 0,064 + 1^2 \cdot 0,288 + 2^2 \cdot 0,432 + 3^2 \cdot 0,216 = 3,96;$$

$$D(X) = 3,96 - 3,24 = 0,72.$$

Обчислимо коефіцієнт асиметрії та ексцес за формулами (8.9).

$$v_1(X) = M(X) = 1,8; v_2(X) = 3,96; v_3(X) = 9,576; v_4(X) = 24,696.$$

$$\mu_3(X) = v_3(X) - 3 \cdot v_2(X) \cdot v_1(X) + 2 \cdot v_1^3(X) = 9,576 - 21,384 + 11,664$$

$$= -0,144;$$

$$\mu_4(X) = v_4(X) - 4 \cdot v_3(X) \cdot v_1(X) + 6 \cdot v_2(X) \cdot v_1^2(X) - 3 \cdot v_1^4(X) =$$

$$= 24,696 - 68,9472 + 76,9824 - 31,4928 = 1,2384.$$

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} \approx -0,236; As(X) = \frac{1-2 \cdot p}{\sqrt{n \cdot p \cdot q}} \approx -0,236.$$

Це вказує на те, що даний біноміальний розподіл не є симетричним (зсув ліворуч від моди) в порівнянні з нормальним розподілом (де коефіцієнт асиметрії дорівнює 0).

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 \approx -0,611; Es(X) = \frac{1-6 \cdot p \cdot q}{n \cdot p \cdot q} \approx -0,611.$$

Це вказує на те, що даний біноміальний розподіл має меншу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Відповідь. $M(X) = 1,8$; $D(X) = 0,72$; $As(X) \approx -0,236$; $Es(X) \approx -0,611$.

Приклад 8.2. За даними прикладу 8.1. знайти математичне сподівання та дисперсію частоти (долі) своєчасного відправлення рейсів серед 3 запланованих, а також $Mo(X)$.

Розв'язання: За формулами (8.6) та (8.7) математичне сподівання та дисперсія дорівнюють:

$$M\left(\frac{i}{n}\right) = 0,6; D\left(\frac{i}{n}\right) = \frac{0,6 \cdot 0,4}{3} = 0,08.$$

За формулою (8.7)

$$0,6 \cdot 0,4 - 0,4 \leq Mo(X) \leq 0,6 \cdot 0,4 + 0,6;$$

$$1,4 \leq Mo(X) \leq 2,4;$$

$$Mo(X) = 2.$$

Отримане таким чином значення $Mo(X)$, узгоджується з результатом, який можна отримати з табл. 8.2.

Відповідь. $M\left(\frac{i}{n}\right) = 0,6$; $D\left(\frac{i}{n}\right) = 0,08$; $Mo(X) = 2$.

8.2. Розподіл Пуассона

Розподіл Пуассона виникає в схемі Бернуллі, коли число n виконуваних незалежних випробувань велике, ймовірність p появи події A в кожному випробуванні мала, а випадкова величина X — число появ події A в цих n випробуваннях ($n \rightarrow +\infty$).

В цьому випадку масових рідкісних подій випадкова величина X приймає цілі невід'ємні значення $x_i = i$ ($i = 0, 1, 2, \dots, n, \dots$) з ймовірностями $p_i = P_n\{x_i = i\} = P_n(i)$, які обчислюються за формулою Пуассона (8.10):

$$P_n(i) = \frac{\lambda^i}{i!} \cdot e^{-\lambda} \quad (\lambda = n \cdot p). \quad (8.10)$$

Для отримання формули (8.10) в формулі (4.7) лекції 4 підставляємо $m = i$.

Таким чином, величина X має ряд розподілу (табл. 8.3):

Таблиця 8.3. Ряд розподілу випадкової величини X , що має розподіл Пуассона

i	0	1	2	...	n	...
$P_n(i)$	$e^{-\lambda}$	$\lambda \cdot e^{-\lambda}$	$\frac{\lambda^2}{2!} \cdot e^{-\lambda}$	...	$\frac{\lambda^n}{n!} \cdot e^{-\lambda}$	...

Очевидно, що розподіл Пуассона коректний, оскільки

$$\sum_{i=0}^{+\infty} P_n(i) = e^{-\lambda} + \lambda \cdot e^{-\lambda} + \frac{\lambda^2}{2!} \cdot e^{-\lambda} + \dots + \frac{\lambda^n}{n!} \cdot e^{-\lambda} + \dots =$$

Оскільки $n \rightarrow +\infty$ використовуємо відомий розклад функції e^x в степеневий ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

$$= e^{-\lambda} \cdot \left(1 + \lambda + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} + \dots\right) = e^{-\lambda} \cdot e^{\lambda} = 1.$$

Властивості розподілу Пуассона

1. Події, які описуються розподілом Пуассона, повинні бути незалежними одна від одної.
2. Розподіл Пуассона часто використовується для моделювання рідкісних подій, наприклад, кількості автомобільних аварій на певній ділянці дороги за один день, кількості помилок у тексті або кількості клієнтів, що приходять до банку за годину.

3. Розподіл Пуассона є асиметричним, але з ростом значення λ він стає більш симетричним і наближається до нормального розподілу.

4. Розподіл Пуассона є цілочисловим дискретним розподілом.

5. Функція розподілу $F(x)$ випадкової величини в разі потреби будується за принципом, наведеним в лекції 6.

6. Математичне сподівання розподілу Пуассона дорівнює (8.11):

$$M(X) = \lambda, \quad (8.11)$$

дисперсія дорівнює (8.12):

$$D(X) = \lambda, \quad (8.12)$$

і, відповідно, середнє квадратичне відхилення (8.13):

$$\sigma(X) = \sqrt{\lambda}. \quad (8.13)$$

Доведення. Для знаходження основних числових характеристик $M(X), D(X), \sigma(X)$ розподілу Пуассона використовується твірна функція

$$\varphi(z) = \sum_{i=0}^{+\infty} \left(\left(\frac{\lambda^i}{i!} \cdot e^{-\lambda} \right) \cdot z^i \right) = e^{-\lambda} \cdot \sum_{i=0}^{+\infty} \left(\frac{(\lambda \cdot z)^i}{i!} \right).$$

Знову використовуючи відомий розклад функції e^x в степеневий ряд, представляємо твірну функцію у вигляді

$$\varphi(z) = e^{-\lambda} \cdot e^{\lambda \cdot z} = e^{\lambda \cdot (z-1)}.$$

Знайдемо першу похідну

$$\varphi'(z) = \lambda \cdot e^{\lambda \cdot (z-1)},$$

тоді математичне сподівання дорівнює

$$M(X) = \varphi'(1) = \lambda.$$

Знайдемо другу похідну

$$\begin{aligned} \varphi''(z) &= \lambda^2 \cdot e^{\lambda \cdot (z-1)}, \\ \varphi''(1) &= \lambda^2. \end{aligned}$$

Одержимо дисперсію

$$D(X) = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

і, відповідно, середнє квадратичне відхилення

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\lambda}.$$

Отже, характерною особливістю розподілу Пуассона є те, що його математичне сподівання збігається з дисперсією. Це робить його простим, але й більш обмеженим по своїм можливостям.

7. При $p \rightarrow 0; n \rightarrow +\infty; n \cdot p \rightarrow \lambda$ закон розподілу Пуассона є граничним випадком біноміального закону розподілу.

Поряд із «граничним» випадком біномного розподілу закон Пуассона може виникнути і в низці інших ситуацій. Показано, що для найпростішого

потоків подій кількість подій, що потрапляють на довільний відрізок часу, є випадковою величиною, що має розподіл Пуассона.

За законом Пуассона розподілені, наприклад, число народження четверні, кількість збоїв на автоматичній лінії, кількість відмов складної системи в «нормальному режимі», кількість «вимог на обслуговування», що надійшли в одиницю часу в системах масового обслуговування, та ін.

8. Якщо випадкова величина є сумою двох незалежних випадкових величин, розподілених кожна за законом Пуассона, то вона також розподілена за законом Пуассона.

Доведення. Розглянемо дві випадкові величини $\{X = i\}$ та $\{Y = j\}$, що мають розподіл Пуассона з параметрами λ_1 та λ_2 відповідно. Нехай є величина $\{Z = X + Y = k\}$.

$$\begin{aligned} P\{Z = k\} &= P\{X = i\} \cdot P\{Y = j\} = \sum_{s=m+k} \left(\left(\frac{\lambda_1^i}{i!} \cdot e^{-\lambda_1} \right) \cdot \left(\frac{\lambda_2^j}{j!} \cdot e^{-\lambda_2} \right) \right) = \\ &= e^{-(\lambda_1 + \lambda_2)} \cdot \sum_{k=i+j} \left(\frac{\lambda_1^i}{i!} \cdot \frac{\lambda_2^j}{j!} \right) = e^{-(\lambda_1 + \lambda_2)} \cdot \sum_{j=0}^k \left(\frac{\lambda_1^{k-j}}{(k-j)!} \cdot \frac{\lambda_2^j}{j!} \right) = \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \cdot \sum_{j=0}^k \left(k! \cdot \frac{\lambda_1^{k-j}}{(k-j)!} \cdot \frac{\lambda_2^j}{j!} \right) = \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \cdot \sum_{j=0}^k (C_k^j \cdot \lambda_1^{k-j} \cdot \lambda_2^j) = \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \cdot (\lambda_1 + \lambda_2)^k = \frac{\lambda^k}{k!} \cdot e^{-\lambda}; \end{aligned}$$

де $\lambda = \lambda_1 + \lambda_2$.

Тобто величина Z теж розподілена за законом Пуассона з параметром $\lambda = \lambda_1 + \lambda_2$.

9. Для розподілу Пуассона з параметром λ , мода визначається як (8.14):

$$Mo(X) = [\lambda] \text{ або } Mo(X) = [\lambda] - 1, \quad (8.14)$$

де $[\lambda]$ — це найбільше ціле число, що не перевищує λ (тобто ціла частина λ).

- Якщо λ — ціле число, то мода дорівнює самому цьому числу, тобто λ та $\lambda - 1$;

- Якщо λ — не ціле число, мода зазвичай дорівнює $[\lambda]$.

Доведення. Правомірність даного твердження можна побачити на рис. 8.1.

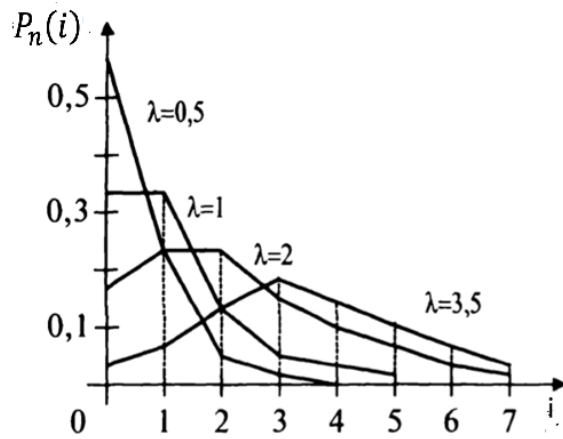


Рис. 8.1. Багатокутники розподілу випадкової величини, розподіленої за законом Пуассона з параметрами $\lambda = 0,5; 1; 2; 3,5$

10. Медіана розподілу Пуассона обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $(a; b]$ дорівнює $\frac{1}{2}$, згідно формули (6.20) лекції 6 $(F(Me(X)) = \frac{1}{2})$.

11. Коефіцієнт асиметрії та ексцес розподілу Пуассона відповідно дорівнюють (8.15):

$$As(X) = \frac{1}{\sqrt{\lambda}}; Es(X) = \frac{1}{\lambda}. \quad (8.15)$$

Доведення.

$$\begin{aligned} \mu_3(X) &= M(X - M(X))^3 = \lambda; \\ \sigma(X) &= \sqrt{\lambda}; \\ \mu_4(X) &= M(X - M(X))^4 = 3 \cdot (n \cdot p \cdot q)^2 + n \cdot p \cdot q \cdot (1 - 6 \cdot p \cdot q); \\ As(X) &= \frac{\mu_3(X)}{\sigma^3(X)} = \frac{\lambda}{(\sqrt{\lambda})^3} = \frac{1}{\sqrt{\lambda}}; \\ Es(X) &= \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{3 \cdot \lambda^2 + \lambda}{\lambda^2} - 3 = \frac{1}{\lambda}. \end{aligned}$$

Приклад 8.3. У відділі технічної підтримки компанії протягом кожної години в середньому надходить 4 дзвінки. Яка ймовірність того, що за наступну годину надійде: **1.** Рівно 6 дзвінків? **2.** Менше 3 дзвінків? **3.** Обчислити коефіцієнт асиметрії та ексцес випадкової величини.

Розв'язання. Нехай X — це випадкова величина, що позначає кількість дзвінків, які надійдуть протягом години. Оскільки дзвінки надходять незалежно один від одного, імовірність того, що дзвінок надійде в будь-який момент часу, постійна, тому можна припустити, що X має розподіл Пуассона з параметром $\lambda = 4$ (середня кількість дзвінків на годину).

1. Обчислимо ймовірність події $\{X = 6\}$ – {за годину надійде рівно 6 дзвінків}

За формула ймовірності для розподілу Пуассона (формула (8.9)):

$$P_n(6) = \frac{4^6 \cdot e^{-4}}{6!} = \frac{4096 \cdot 0,0183}{720} \approx 0,103.$$

2. Обчислимо ймовірність події $\{X < 3\}$ – {за годину надійде менше 3 дзвінків}.

$$P_n\{X < 3\} = P_n(0) + P_n(1) + P_n(2).$$

Для кожного значення k застосуємо формулу розподілу Пуассона:

Для $i = 0$:

$$P_n(0) = \frac{4^0 \cdot e^{-4}}{0!} = \frac{1 \cdot e^{-4}}{1} = e^{-4} \approx 0,0183.$$

Для $i = 1$:

$$P_n(1) = \frac{4^1 \cdot e^{-4}}{1!} = \frac{4 \cdot e^{-4}}{1} = 4 \cdot e^{-4} \approx 0,0732.$$

Для $i = 2$:

$$P_n(2) = \frac{4^2 \cdot e^{-4}}{2!} = \frac{16 \cdot e^{-4}}{2} = 8 \cdot e^{-4} \approx 0,1464.$$

Тепер знайдемо суму цих ймовірностей:

$$P_n\{X < 3\} = P_n(0) + P_n(1) + P_n(2) = 0,0183 + 0,0732 + 0,1464 \approx 0,2379.$$

Отже, ймовірність того, що за годину надійде менше 3 дзвінків, приблизно дорівнює 0,238.

3. За формулами (8.15) з властивості 9:

$$As(X) = \frac{1}{\sqrt{\lambda}} = \frac{1}{2}.$$

Це вказує на те, що даний розподіл Пуассона не є симетричним (зсув праворуч від моди) в порівнянні з нормальним розподілом (де коефіцієнт асиметрії дорівнює 0).

$$Es(X) = \frac{1}{\lambda} = \frac{1}{4}.$$

Це вказує на те, що даний розподіл Пуассона має більшу «піковість» або «випуклість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Відповідь: 1. $P_n(6) \approx 0,103$; 2. $P_n\{X < 3\} \approx 0,2379$; 3. $As(X) = \frac{1}{2}$; $Es(X) = \frac{1}{4}$.

Приклад 8.4. Система бронювання та продажу авіаційних білетів має 1000 периферійних пультів. Ймовірність надходження запиту з кожного пульта на протязі однієї хвилини дорівнює 0,002. Знайти **а)** середню кількість запитів; **б)** ймовірність надходження протягом однієї хвилини принаймні двох запитів; **в)** моду випадкової величини.

Розв'язання. За умовою $n = 1000$, $p = 0,002$, тому $\lambda = 2$.

а) Середня кількість (математичне сподівання) запитів, що надійде до системи на протязі однієї хвилини, $M(X) = \lambda = 2$.

б) Обчислимо ймовірність надходження протягом однієї хвилини принаймні двох запитів $\{2 \leq X \leq 1000\}$.

$$P_{1000}(0) = e^{-2} \approx 0,135; P(1) = 2 \cdot e^{-2} \approx 0,271$$

$$P_{1000}\{2 \leq X \leq 1000\} = 1 - P(0) - P(1) \approx 0,594.$$

в) За формулою(8.14):

$$Mo(X) = [\lambda] = 2; Mo(X) = [\lambda] - 1 = 1.$$

Відповідь. а) $M(X) = 2$; б) $P\{2 \leq X \leq 1000\} \approx 0,594$; в) $Mo(X) = 1$; $Mo(X) = 2$.

8.3. Геометричний розподіл

Геометричний розподіл виникає, коли незалежні випробування, в кожному з яких подія A може відбутися зі сталою ймовірністю p , проводяться до першого «невдалого» випробування (подія A не відбулась) і далі припиняються, а випадкова величина X — число проведених «вдалих» випробувань.

Дискретна величина X приймає цілі невід'ємні значення $x_i = i$ ($i = 0, 1, 2, \dots, n, \dots$) з ймовірностями $p_i = P_n\{x_i = i\} = P_n(i)$, що обчислюються за формулою (8.16):

$$P_n(i) = p^i \cdot q, \quad (8.16)$$

оскільки X приймає значення i , якщо в i випробуваннях подія A відбулась (ймовірність p^i), а в наступному — не відбулась (ймовірність q).

Тому X має нескінченний ряд розподілу (табл. 8.4).

Таблиця 8.4. Ряд розподілу випадкової величини X , що має геометричний розподіл

i	0	1	2	...	n	...
$P_n(i)$	q	$p \cdot q$	$p^2 \cdot q$	...	$p^n \cdot q$	...

Очевидно, що геометричний розподіл коректний, оскільки

$$\begin{aligned} \sum_{i=0}^{+\infty} P_n(i) &= p + p \cdot q + p^2 \cdot q + \dots + p^n \cdot q + \dots = \\ &= q \cdot (1 + p + p^2 + \dots + p^n + \dots) = q \cdot \frac{1}{1 - p} = \frac{q}{q} = 1. \end{aligned}$$

Властивості геометричного розподілу

1. Геометричний розподіл є цілочисловим дискретним розподілом.
2. Функція розподілу $F(x)$ випадкової величини в разі потреби будується за принципом, наведеним в лекції 6.
3. Математичне сподівання геометричного розподілу дорівнює (8.17):

$$M(X) = \frac{p}{q}, \quad (8.17)$$

дисперсія дорівнює (8.18):

$$D(X) = \frac{p}{q^2}, \quad (8.18)$$

і, відповідно, середнє квадратичне відхилення (8.19):

$$\sigma(X) = \frac{\sqrt{p}}{q}. \quad (8.19)$$

Доведення. Для знаходження основних числових характеристик $M(X), D(X), \sigma(X)$ застосовується твірна функція

$$\varphi(z) = \sum_{i=0}^{+\infty} (p^i \cdot q \cdot z^i) = q \cdot \sum_{i=0}^{+\infty} (p \cdot z)^i,$$

доданки якої є членами нескінченно спадної геометричної прогресії з першим членом $u_1 = 1$ і знаменником $b = p \cdot q < 1$. Сума членів такої геометричної прогресії обчислюється за формулою

$$S = \frac{u_1}{1-b} = \frac{1}{1-p \cdot q}, \text{ тому } \varphi(z) = \frac{q}{1-p \cdot z}.$$

Знайдемо першу похідну

$$\varphi'(z) = \frac{p \cdot q}{(1 - p \cdot z)^2},$$

тоді одержимо математичне сподівання

$$M(X) = \varphi'(1) = \frac{p \cdot q}{(1 - p)^2} = \frac{p}{q}.$$

Знайдемо другу похідну

$$\begin{aligned} \varphi''(z) &= \frac{2 \cdot p^2 \cdot q}{(1 - p \cdot z)^3}, \\ \varphi''(1) &= \frac{2 \cdot p^2 \cdot q}{(1 - p)^3} = \frac{2 \cdot p^2}{q^2}, \end{aligned}$$

тоді одержимо дисперсію

$$D(X) = \frac{2 \cdot p^2}{q^2} + \frac{p}{q} - \frac{p^2}{q^2} = \frac{p^2}{q^2} + \frac{p}{q} = \frac{p}{q^2}.$$

і, відповідно, середнє квадратичне відхилення

$$\sigma(X) = \frac{\sqrt{p}}{q}.$$

4. Геометричний розподіл також виникає, коли незалежні випробування, в кожному з яких подія A може відбутися зі сталою ймовірністю p , проводяться до першого «вдалого» випробування (подія A відбулась) і далі припиняються, а випадкова величина $\bar{X}$ — число проведених «невдалих» випробувань. Її ряд розподілу матиме вигляд (табл. 8.6):

Таблиця 8.6. Ряд розподілу випадкової величини $\bar{X}$

i	0	1	2	...	n	...
$P_n(i)$	p	$q \cdot p$	$q^2 \cdot p$	...	$q^n \cdot p$	...

Математичне сподівання геометричного розподілу дорівнює (8.20)-(8.22):

$$M(X) = \frac{q}{p}, \quad (8.20)$$

дисперсія дорівнює:

$$D(X) = \frac{q}{p^2}, \quad (8.21)$$

і, відповідно, середнє квадратичне відхилення:

$$\sigma(X) = \frac{\sqrt{q}}{p}. \quad (8.22)$$

Доведення. Коректність розподілу знову встановлюється за допомогою обчислення суми нескінченно спадної геометричної прогресії.

$$\begin{aligned} \sum_{i=0}^{+\infty} P_n(i) &= q + q \cdot p + q^2 \cdot p + \dots + q^n \cdot p + \dots = \\ &= p \cdot (1 + q + q^2 + \dots + q^n + \dots) = p \cdot \frac{1}{1 - q} = \frac{p}{p} = 1. \end{aligned}$$

Для знаходження основних числових характеристик $M(X), D(X), \sigma(X)$, на відміну від властивості 2, застосовується наступна твірна функція

$$\varphi(z) = \sum_{i=0}^{+\infty} (q^i \cdot p \cdot z^i) = p \cdot \sum_{i=0}^{+\infty} (q \cdot z)^i.$$

В іншому хід виведення формул аналогічний з наведеним у доведенні властивості 2.

5. Найбільш ймовірно, що перше випробування відразу буде «невдалим» (або, враховуючи властивість 3, – «вдалим»), тобто $M_0(X) = 0$.

Доведення. Дане твердження є очевидним, враховуючи той факт, що в геометричному розподілі значення, які дана набуває випадкова величина утворюють нескінченно спадну геометричну прогресію. Таким чином найбільша ймовірність відповідає саме значенню $X = 0$. Отже $M_0(X) = 0$.

6. Медіана геометричного розподілу обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $(a; b]$ дорівнює $\frac{1}{2}$, згідно формули (6.20) лекції 6 $(F(Me(X)) = \frac{1}{2})$.

7. Коефіцієнт асиметрії та ексцес випадкової величини X , що має геометричний розподіл, відповідно дорівнюють (8.23):

$$As(X) = \frac{1+p-4 \cdot p \cdot q}{\sqrt{p}}; Es(X) = \frac{6-14 \cdot q+21 \cdot q^2-13 \cdot q^3+4 \cdot q^4}{p} - 3. \quad (8.23)$$

Для випадкової величини $\bar{X}$ – (8.24):

$$As(X) = \frac{1+q-4 \cdot p \cdot q}{\sqrt{q}}; Es(X) = \frac{6-14 \cdot p+21 \cdot p^2-13 \cdot p^3+4 \cdot p^4}{q} - 3. \quad (8.24)$$

Доведення. Для випадкової величини X

$$\mu_3(X) = M(X - M(X))^3 = \frac{p}{q^3} \cdot (2 - 5 \cdot q + 4 \cdot q^2);$$

$$\sigma(X) = \frac{\sqrt{p}}{q};$$

$$\mu_4(X) = M(X - M(X))^4 = \frac{p}{q^4} \cdot (6 - 14 \cdot q + 21 \cdot q^2 - 13 \cdot q^3 + 4 \cdot q^4);$$

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = \frac{\frac{p}{q^3} \cdot (2 - 5 \cdot q + 4 \cdot q^2)}{\left(\frac{\sqrt{p}}{q}\right)^3} = \frac{2 - 5 \cdot q + 4 \cdot q^2}{\sqrt{p}}$$

$$= \frac{1 + p - 4 \cdot p \cdot q}{\sqrt{p}};$$

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{\frac{p}{q^4} \cdot (6 - 14 \cdot q + 21 \cdot q^2 - 13 \cdot q^3 + 4 \cdot q^4)}{\left(\frac{\sqrt{p}}{q}\right)^4} - 3$$

$$= \frac{6 - 14 \cdot q + 21 \cdot q^2 - 13 \cdot q^3 + 4 \cdot q^4}{p} - 3.$$

Для випадкової величини $\bar{X}$ коефіцієнт асиметрії та ексцес обчислюються аналогічно.

Практичне застосування має також розподіл випадкової величини X_1 — числа взагалі проведених випробувань. Ця величина приймає натуральні можливі значення $x_i = i$ ($i = 1, 2, \dots, n, \dots$) з ймовірностями $p_i = P_n\{x_i = i\} = P_n(i)$, що обчислюються за формулою (8.25):

$$P_n(i) = p^{i-1} \cdot q, \quad (8.25)$$

тобто X_1 має ряд розподілу (табл. 8.7):

Таблиця 8.7. Ряд розподілу випадкової величини X_1 , що має геометричний розподіл

i	1	2	3	...	n	...
$P_n(i)$	q	$p \cdot q$	$p^2 \cdot q$	...	$p^{n-1} \cdot q$	...

За властивістю 6 твірної функції розподіл випадкової величини X_1 є зміщеним на 1 по відношенню до геометричного розподілу величини X (табл. 8.4), тому її основні числові характеристики приймають вигляд (8.26)-(8.28):

$$M(X_1) = M(X + 1) = M(X) + 1 = \frac{p}{q} + 1 = \frac{1}{q}; \quad (8.26)$$

$$D(X_1) = D(X + 1) = D(X) = \frac{p}{q^2}; \quad (8.27)$$

$$\sigma(X_1) = \frac{\sqrt{p}}{q}; \quad (8.28)$$

Для випадкової величини X_1 мода дорівнює (8.29):

$$Mo(X_1) = 1. \quad (8.29)$$

Коефіцієнт асиметрії та ексцес (8.30):

$$As(X) = \frac{1-q}{\sqrt{p}}; Es(X) = 6 + \frac{q^2}{p}. \quad (8.30)$$

Формули (8.17)-(8.19) або (8.20)-(8.22) для знаходження основних числових характеристик застосовуються у випадку нескінченного ряду розподілу. Проте в деяких задачах умови випробувань передбачають обмеження ряду розподілу, тобто геометрично розподілена випадкова величина X_2 приймає можливі значення $x_i = i$ ($i = 0, 1, 2, \dots, n$). В цьому випадку випадкова величина X_2 приймає значення n , якщо в n -му випробуванні подія A не відбулась (ймовірність $p^n \cdot q$) або відбулась (ймовірність p^{n+1}), отже, згідно з теоремою додавання ймовірностей несумісних подій

$$P_n(n) = p^n \cdot q + p^{n+1} = p^n \cdot (p + q) = p^n,$$

а ряд розподілу цієї випадкової величини приймає вигляд (табл. 8.8):

Таблиця 8.8. Ряд розподілу випадкової величини X_2 , що має обмежений геометричний розподіл

i	0	1	2	...	n
$P_n(i)$	q	$p \cdot q$	$p^2 \cdot q$	...	p^n

Застосування формул (8.17)-(8.19) для знаходження основних числових характеристик цієї випадкової величини приводить до значних похибок, тому в цьому випадку доцільно використовувати загальні формули.

Приклад 8.5. Брак в продукції цеху по виробництву однотипних виробів складає 10%. Для оцінки якості великої партії виробів контролер навмання відбирає по одному виробу до появи першого бракованого. Знайти закони розподілу та основні числові характеристики випадкових величин: **а)** $X = \{\text{кількості відібраних доброякісних виробів}\}$; **б)** $X_1 = \{\text{загальної кількості перевірених виробів}\}$; **в)** $X_2 = \{\text{кількості відібраних доброякісних виробів}\}$, якщо в кожній партії контролер перевіряє не більше n виробів.

Розв'язання. **а)** Випадкова величина X має геометричний розподіл, для якого $p = 0,9, q = 0,1$ («невдалим» випробуванням вважається поява бракованого виробу). Тому ряд розподілу випадкової величини X будується у відповідності з табл. 8.4. Ряд розподілу наведено в табл. 8.9.

Таблиця 8.9. Ряд розподілу випадкової величини X

i	0	1	2	...	n	...
$P_n(i)$	0,1	0,09	0,081	...	$(0,9)^n \cdot (0,1)$	...

$$P_n(0) = q = 0,1; P_n(1) = p \cdot q = 0,09;$$

$$P_n(2) = p^2 \cdot q = 0,081; \dots;$$

$$P_n(n) = p^n \cdot q = (0,9)^n \cdot (0,1).$$

Основні числові характеристики обчислюються за формулами (8.17)-(8.19):

$$M(X) = \frac{p}{q} = 9; D(X) = \frac{p}{q^2} = 90; \sigma(X) = \frac{\sqrt{p}}{q} \approx 9,5.$$

б) Випадкова величина X_1 має геометричний розподіл, зміщений на 1, тому її ряд розподілу складається у відповідності з табл. 8.5 (дивись табл. 8.10).

Таблиця 8.10. Ряд розподілу випадкової величини X_1

i	1	2	3	...	n	...
$P_n(i)$	0,1	0,09	0,081	...	$(0,9)^{n-1} \cdot (0,1)$	...

$$P_n(1) = q = 0,1; P_n(2) = p \cdot q = 0,09;$$

$$P_n(3) = p^2 \cdot q = 0,081; \dots;$$

$$P_n(n) = p^{n-1} \cdot q = (0,9)^{n-1} \cdot (0,1).$$

Основні числові характеристики обчислюються за формулами (8.26)-(8.28):

$$M(X_1) = \frac{p}{q} = 10; D(X_1) = \frac{p}{q^2} = 90; \sigma(X_1) = \frac{\sqrt{p}}{q} \approx 9,5.$$

в) Випадкова величина X_2 має обмежений ряд розподілу (табл. 8.11) при $i = 5$:

Таблиця 8.11. Ряд розподілу випадкової величини X_2

i	0	1	2	3	4	5
$P_n(i)$	0,1	0,09	0,081	0,0729	0,06561	0,59049

$$P_n(0) = q = 0,1; P_n(1) = p \cdot q = 0,09;$$

$$P_n(2) = p^2 \cdot q = 0,081; P_n(3) = p^3 \cdot q = 0,0729;$$

$$P_n(4) = p^4 \cdot q = 0,06561; P_n(5) = p^5 \cdot q = 0,59049.$$

$$M(X_2) = 0 \cdot 0,1 + 1 \cdot 0,09 + 2 \cdot 0,081 + 3 \cdot 0,0729 + 4 \cdot 0,06561 + 5 \cdot 0,59049 \approx 3,686;$$

$$D(X_2) = 0^2 \cdot 0,1 + 1^2 \cdot 0,09 + 2^2 \cdot 0,081 + 3^2 \cdot 0,0729 + 4^2 \cdot 0,06561 + 5^2 \cdot 0,59049 - (3,689)^2 \approx 3,3;$$

$$\sigma(X_2) \approx 1,8.$$

Відповідь: а) табл. 8.7; $M(X) = 9$; $D(X) = 90$; $\sigma(X) \approx 9,5$; б) табл. 8.8; $M(X) = 10$; $D(X) = 90$; $\sigma(X) \approx 9,5$; в) табл. 8.9; $M(X) \approx 3,686$; $D(X) \approx 3,3$; $\sigma(X) \approx 1,8$.

Приклад 8.6. Стрілок стріляє по мішені з ймовірністю влучити 0,9. Випробування закінчується після першого ж промаху. Побудувати ряд розподілу випадкової величини X_1 – загальна кількість пострілів і знайти основні числові характеристики.

Розв'язок. Випадкова величина X_1 має геометричний розподіл і складається у відповідності з табл. 8.5. Ряд розподілу наведено в табл. 8.12.

Таблиця 8.12. Ряд розподілу випадкової величини X_1

i	1	2	3	...	n	...
$P_n(i)$	0,1	0,09	0,081	...	$(0,9)^{n-1} \cdot (0,1)$	...

$$P_n(1) = q = 0,1; P_n(2) = p \cdot q = 0,09;$$

$$P_n(3) = p^2 \cdot q = 0,081; \dots;$$

$$P_n(n) = p^{n-1} \cdot q = (0,9)^{n-1} \cdot (0,1).$$

Основні числові характеристики обчислюються за формулами (8.26)-(8.28):

$$M(X_1) = \frac{1}{q} = 10; D(X_1) = \frac{p}{q^2} = 90; \sigma(X_1) = \frac{\sqrt{p}}{q} \approx 9,5.$$

Відповідь: Табл. 8.10; $M(X_1) = 10$; $D(X_1) = 90$; $\sigma(X_1) \approx 9,5$.

8.4. Розподіл Паскаля

Розподіл Паскаля, також відомий як **негативний біноміальний розподіл**, описує ймовірність кількості невдач (або кількості випробувань), необхідних для досягнення певної кількості успіхів у послідовності незалежних випробувань, де кожне випробування має ймовірність успіху p .

Позначимо $P(A) = p$; $P(\bar{A}) = 1 - p = q$. Випробування повторюються до тих пір, поки подія A не з'явиться k разів, $k > 0$. Випадкова величина X – кількість невдалих випробувань до появи події k разів, включаючи і останнє випробування.

Ймовірність $p_m = P_n\{x_i = m\} = P_n(m)$ того, що подія A настане k разів і в результаті експериментів матиме m «невдалих» спроб дорівнює (8.31):

$$p_m = P_n(m) = C_{n-1}^m \cdot p^k \cdot q^m, \quad (8.31)$$

де $n = k + m$ – загальна кількість спроб.

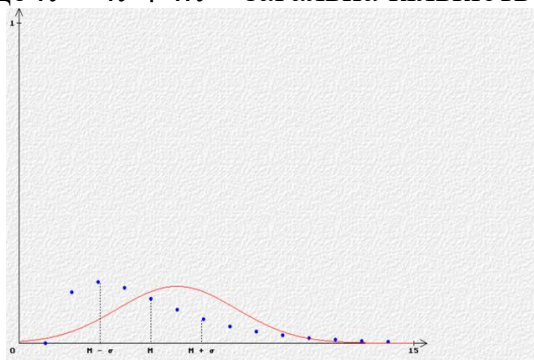


Рис. 8.2. Ілюстрація випадкової величини X – кількості збоїв при зчитуванні інформації з носія, що має розподіл Паскаля

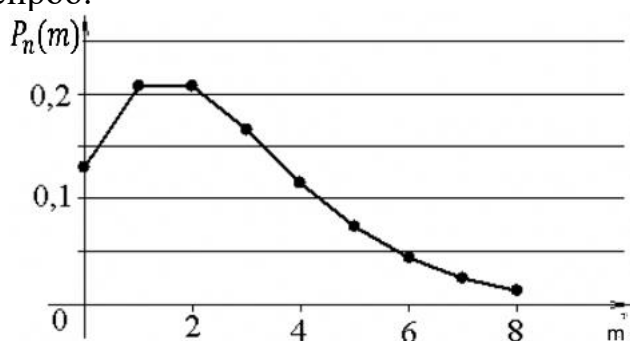


Рис. 8.3. Багатокутник розподілу Паскаля $k = 4, p = 0,6$

Властивості розподілу Паскаля

1. Розподіл Паскаля є цілочисловим дискретним розподілом.

2. Функція розподілу $F(x)$ випадкової величини в разі потреби будується за принципом, наведеним в лекції 6.

3. Математичне сподівання розподілу Паскаля дорівнює (8.32):

$$M(X) = \frac{k \cdot q}{p}, \quad (8.32)$$

дисперсія дорівнює (8.33):

$$D(X) = \frac{k \cdot q}{p^2}, \quad (8.33)$$

і, відповідно, середнє квадратичне відхилення (8.34):

$$\sigma(X) = \frac{\sqrt{k \cdot q}}{p}. \quad (8.34)$$

Доведення. Формули (8.32)-(8.34) отримано розглядаючи розподіл Паскаля як суму i геометричних розподілів. Слід зазначити, що в даному випадку геометричний розподіл має випадкова величина $\bar{X}$ – кількість «невдалих» спроб до першої появи події A (дивись властивість 3 геометричного розподілу). Її числові характеристики обчислюються за формулами (8.20)-(8.22).

4. Геометричний розподіл є частинним випадком геометричного розподілу.

Доведення. Нехай $k = 1$. Тоді з формули (8.31), враховуючи, що $n = m + 1$, матимемо:

$$P_n(m) = C_m^m \cdot p^1 \cdot q^m = p \cdot q^m.$$

В даному випадку випадкова величина X має m «невдалих» спроб до першої появи події A , а це і є ознака геометричного розподілу даної випадкової величини.

5. Якщо розподіл Паскаля має параметри k (кількість успіхів) та p (ймовірність успіху), моду можна обчислити за такою формулою (8.35):

$$Mo(X) = \left\lfloor \frac{(k-1) \cdot q}{p} \right\rfloor \text{ або } Mo(X) = \left\lfloor \frac{(k-1) \cdot q}{p} \right\rfloor - 1, \quad (8.35)$$

де $\lfloor \cdot \rfloor$ означає операцію округлення до найближчого меншого цілого числа.

- Якщо $\frac{(k-1) \cdot q}{p}$ — ціле число, то $Mo(X) = \frac{(k-1) \cdot q}{p}$ та $Mo(X) = \frac{(k-1) \cdot q}{p} - 1$.

- Якщо $\frac{(k-1) \cdot q}{p}$ — не ціле число, мода дорівнює $Mo(X) = \left\lfloor \frac{(k-1) \cdot q}{p} \right\rfloor$.

6. Медіана розподілу Паскаля обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $(a; b]$ дорівнює $\frac{1}{2}$, згідно формули (6.20) лекції 6 $(F(Me(X)) = \frac{1}{2})$.

7. Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл Паскаля, доцільніше обчислювати стандартним методом, наведеним в лекції 7.

Розподіл Паскаля часто використовується в моделюванні процесів, де кількість необхідних подій до досягнення певного результату є важливою, наприклад, у страхуванні, біології або в азартних іграх. Він також описує процес розповсюдження епідемій чи ланцюгових реакцій.

Приклад 8.7. Дехто кидає монету до тих пір, поки не випаде 3 герби. Обчислити: **1.** ймовірність того, що для цього знадобиться 5 невдач (тобто 5 разів випаде орел); **2.** моду та медіану випадкової величини.

Розв'язок: За умовою $k = 3$, $p = 0,5$, $m = 5$, $n = 3 + 5 = 8$.

1. Використовуємо формулу розподілу Паскаля (8.27):

$$P_8(5) = C_7^5 \cdot (0,5)^3 \cdot (0,5)^5 = \frac{7!}{5! \cdot 2!} \cdot (0,5)^8 = 21 \cdot \frac{1}{256} = \frac{21}{256} \approx 0,082.$$

2. За формулою (8.31)

$$Mo(X) = \frac{(k-1) \cdot q}{p} = 2 \text{ та } Mo(X) = \frac{(k-1) \cdot q}{p} - 1 = 1.$$

Коректність результатів можна побачити за допомогою ряду розподілу величини X (табл. 8.13).

Таблиця 8.13. Ряд розподілу випадкової величини X

i	0	1	2	3	4	5	...
$P_n(i)$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{5}{32}$	$\frac{15}{128}$	$\frac{21}{256}$	...

$$\begin{aligned} P_3(0) &= C_2^0 \cdot (0,5)^3 \cdot (0,5)^0 = \frac{1}{8}; P_4(1) = C_3^1 \cdot (0,5)^3 \cdot (0,5)^1 = \frac{3}{16}; \\ P_5(2) &= C_4^2 \cdot (0,5)^3 \cdot (0,5)^2 = \frac{3}{16}; P_6(3) = C_5^3 \cdot (0,5)^3 \cdot (0,5)^3 = \frac{5}{32}; \\ P_7(4) &= C_6^4 \cdot (0,5)^3 \cdot (0,5)^4 = \frac{15}{128}; P_8(5) = C_7^5 \cdot (0,5)^3 \cdot (0,5)^5 = \frac{21}{256}. \end{aligned}$$

Обчислимо медіану випадкової величини.

$$F(x) = \begin{cases} 0, & \text{де } x \leq 0; \\ \frac{1}{8}, & \text{де } 0 < x \leq 1; \\ \frac{5}{16}, & \text{де } 1 < x \leq 2; \\ \frac{1}{2}, & \text{де } 2 < x \leq 3; \\ \frac{21}{32}, & \text{де } 3 < x \leq 4; \\ \dots & \dots \\ 1, & \text{де } x > m. \end{cases}$$

$$Me(X) = 1,5.$$

Відповідь. 1. $P_8(5) \approx 0,082$; 2. $Mo(X) = 1$; $Mo(X) = 2$; $Me(X) = 1,5$.

8.5. Гіпергеометричний розподіл

Гіпергеометричний розподіл виникає, наприклад, в випробуваннях, коли з комплекту, який складається з N предметів, n з яких мають певну

властивість (наприклад, нестандартні, пофарбовані тощо), відбирається навмання m предметів (одноразово, або послідовно без повернення до комплекту), а випадкова величина X — кількість предметів з вказаною властивістю серед відібраних. В загальному випадку X приймає можливі значення $x_i = i$ ($i = 0, 1, 2, \dots, m$) з ймовірностями $p_i = P_n\{x_i = i\} = P_n(i)$, які обчислюються за формулою (8.36):

$$P_n(i) = \frac{C_n^i \cdot C_{N-n}^{m-i}}{C_N^m}, (i = 0, 1, 2, \dots, m). \quad (8.36)$$

Проте, деякі з подій $\{x_i = i\}$ можуть виявитись неможливими, зокрема при $m > n$, отже, відповідні ймовірності, обчислені за формулою (8.31), будуть рівні нулю. Тому при побудові ряду розподілу можливі значення випадкової величини X слід вибирати в межах від $m_1 = \max(0; m - N + n)$ до $m_2 = \min(n; m)$.

Властивості гіпергеометричного розподілу

1. Функція розподілу $F(x)$ випадкової величини будується за принципом, наведеним в лекції 6.
2. Знаходження основних числових характеристик гіпергеометричного розподілу — математичного сподівання $M(X)$ і дисперсії $D(X)$ безпосередньо за допомогою твірної функції

$$\varphi(z) = \sum_{i=0}^m \left(\frac{C_n^i \cdot C_{N-n}^{m-i}}{C_N^m} \right) \cdot z^i.$$

приводить до занадто складних обчислень, отже, гіпергеометричний розподіл дискретної випадкової величини X є прикладом розподілу, для якого твірна функція $\varphi(z)$ не є ефективним засобом знаходження числових характеристик.

В цьому випадку використовують відомі формули (7.1) та (7.8) лекції 7

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

або за формулами (8.37)

$$M(X) = m \cdot p; D(X) = m \cdot p \cdot q \cdot \frac{N-m}{N-1}, \text{ де } p = \frac{n}{N}. \quad (8.37)$$

Доведення. Довести самостійно формули (8.37).

3. Для випадкової величини, що має гіпергеометричний розподіл, моду можна знайти за наступною формулою (8.38):

$$Mo(X) = \left[\frac{(m+1) \cdot (n+1)}{N+2} \right], \quad (8.38)$$

де $[\cdot]$ означає операцію округлення до найближчого меншого цілого числа.

4. Гіпергеометричний розподіл можна розглядати, як модифікацію біноміального розподілу для випадку скінченної сукупності, що складається з N об'єктів, n з яких мають певну властивість.

При $N \rightarrow +\infty; n \rightarrow +\infty$, враховуючи, що $\frac{n}{N} \rightarrow p$ функція ймовірностей гіпергеометричного розподілу (формула (8.36)) прямує до аналогічної функції біноміального розподілу (формула (8.1)).

Доведення. Розпишемо формулу (8.36) згідно відповідних формул комбінаторики.

$$P_n(i) = \frac{C_n^i \cdot C_{N-n}^{m-i}}{C_N^m} = \frac{\left(\frac{n \cdot (n-1) \cdot \dots \cdot (n-i+1)}{i!} \right) \cdot \left(\frac{(N-n) \cdot (N-n-1) \cdot \dots \cdot (N-n-m+i-1)}{(m-i)!} \right)}{\left(\frac{N \cdot (N-1) \cdot \dots \cdot (N-m+1)}{m!} \right)}.$$

Після ділення чисельника та знаменника останнього виразу на N^m та переходячи до границі при $N \rightarrow +\infty; n \rightarrow +\infty$, $\frac{n}{N} \rightarrow p$ отримуємо формулу (8.1), де $q = 1 - \frac{n}{N}$. Що й треба було довести.

5. Медіана гіпергеометричного розподілу обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $(a; b]$ дорівнює $\frac{1}{2}$, згідно формули (6.20) лекції 6 $(F(Me(X)) = \frac{1}{2})$.

6. Коефіцієнт асиметрії та ексцес випадкової величини X , що має гіпергеометричний розподіл, доцільніше обчислювати стандартним методом, наведеним в лекції 7.

Гіпергеометричний розподіл широко використовується в практиці статистичного приймального контролю якості промислової продукції, у завданнях, пов'язаних з організацією вибірових обстежень, та інших областях.

Приклад 8.8. Партія, що складається з 10 виробів, містить 7 виробів вищого ґатунку. Навмання вибирається 6 виробів. Знайти ряд розподілу, основні числові характеристики та моду випадкової величини X – кількості виробів вищого ґатунку серед відібраних.

Розв'язання. Випадкова величина X має гіпергеометричний розподіл, в якому $N = 10; n = 7; m = 6; p = 0,7; q = 0,3$. Визначимо межі m_1 і m_2 можливих значень величини X : $m - N + n = 3$, тому $m_1 = 3$ (більше із чисел 0 і 3), а $m_2 = 6$ (менше із чисел 6 і 7). Ймовірність відповідних можливих значень обчислимо за формулою (8.31):

$$P_7(3) = \frac{C_7^3 \cdot C_3^3}{C_{10}^6} = \frac{1}{6}; \quad P_7(4) = \frac{C_7^4 \cdot C_3^2}{C_{10}^6} = \frac{1}{2};$$

$$P_7(5) = \frac{C_7^5 \cdot C_3^1}{C_{10}^6} = \frac{3}{10}; \quad P_7(6) = \frac{C_7^6 \cdot C_3^0}{C_{10}^6} = \frac{1}{30}.$$

Отже, ряд розподілу випадкової величини X має вигляд (табл. 8.14):
Таблиця 8.14. Ряд розподілу випадкової величини X

i	3	4	5	6
$P_n(i)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

Основні числові характеристики $M(X), D(X), \sigma(X)$ обчислимо за стандартними формулами:

$$M(X) = 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{2} + 5 \cdot \frac{3}{10} + 6 \cdot \frac{1}{30} = 4,2;$$

$$D(X) = 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{2} + 5^2 \cdot \frac{3}{10} + 6^2 \cdot \frac{1}{30} - (4,2)^2 = 0,56; \quad \sigma(X) = 0,748.$$

Для контролю застосуємо формули (8.37):

$$M(X) = 6 \cdot 0,7 = 4,2; \quad D(X) = 6 \cdot 0,7 \cdot 0,3 \cdot \frac{10 - 6}{10 - 1} = 0,56.$$

Також обчислимо моду $Mo(X)$:

$$Mo(X) = \left\lfloor \frac{(6 + 1) \cdot (7 + 1)}{10 + 2} \right\rfloor = 4.$$

Коректність результатів можна побачити за допомогою ряду розподілу величини X (табл. 8.14).

Відповідь. $M(X) = 4,2; D(X) = 0,56; \sigma(X) = 0,748; Mo(X) = 4.$

8.6. Уніформний розподіл

Уніформний розподіл або **дискретний рівномірний розподіл** — це розподіл ймовірностей, у якому всі можливі значення дискретної випадкової величини мають однакову ймовірність.

Іншими словами, якщо випадкова величина X може приймати одне з n можливих значень x_i , то ймовірність $p_i = P_n\{X = x_i\} = P_n(x_i)$ кожного з цих значень дорівнює (8.39):

$$P_n(x_i) = \frac{1}{n}. \quad (8.39)$$

Ряд розподілу уніформної величини матиме вигляд (табл. 8.15):

Таблиця 8.15. Ряд розподілу дискретної випадкової величини

x_i	x_1	x_2	$\dots$	x_{n-1}	x_n
$P_n(x_i)$	$\frac{1}{n}$	$\frac{1}{n}$	$\dots$	$\frac{1}{n}$	$\frac{1}{n}$

Властивості уніформного розподілу

1. Випадкова величина X може приймати n різних значень $x_1, x_2, \dots, x_n$ і події $\{X = x_i\}$ ($i = 1, 2, \dots, n$) є рівноможливими з ймовірностями $P_n(i) = \frac{1}{n}$.
2. Випадкова величина X не обов'язково є цілочисловою. В загальному випадку $x_i \in \mathcal{R}$.
3. Функція розподілу $F(x)$ випадкової величини будується за принципом, наведеним в лекції 6.
4. Уніформний розподіл дискретної випадкової величини X є прикладом розподілу, для якого твірна функція $\varphi(z)$

$$\varphi(z) = \frac{1}{n} \cdot \sum_{i=0}^n z^i = \frac{1}{n} \cdot \frac{z^{n+1} - 1}{1 - z}.$$

не є ефективним засобом знаходження числових характеристик.

Математичне сподівання (середнє значення) обчислюється за формулою (8.40):

$$M(X) = \frac{\sum_{i=1}^n x_i}{n}. \quad (8.40)$$

Дисперсія – за формулою (8.41):

$$D(X) = \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2. \quad (8.41)$$

Середньоквадратичне відхилення – за формулою (8.42):

$$\sigma(X) = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2}. \quad (8.42)$$

Доведення. Формулу для обчислення математичного сподівання (8.40) отримуємо безпосередньо з формули (7.1) лекції 7, враховуючи, що $p_i = \frac{1}{n}$.

Формулу для обчислення дисперсії (8.41) також отримуємо безпосередньо з формули (7.8) лекції 7, враховуючи, що $p_i = \frac{1}{n}$.

Формулу (8.42) отримуємо з формули (7.12) лекції 7.

5. Альтернативою для обчислення для обчислення математичного сподівання та дисперсії уніформного розподілу є (8.43):

$$M(X) = \frac{x_1 + x_n}{2}; D(X) = \frac{(x_n - x_1 + 1)^2 - 1}{12}, \quad (8.43)$$

де x_1, x_n – відповідно найменше та найбільше значення, що приймає випадкова величина X .

Доведення. В основі доведення лежать формула суми членів арифметичної прогресії та формула суми їх квадратів.

Довести самостійно формули (8.43).

6. Уніформний розподіл є багатомодальним (8.44).

$$Mo(X) = x_i \quad (i = 1, 2, \dots, n). \quad (8.44)$$

Доведення. Оскільки події $\{X = x_i\}$ ($i = 1, 2, \dots, n$) є рівноможливими, то всі значення x_i матимуть однакову ймовірність, тобто будуть модами.

7. Медіана уніформного розподілу, за умови, що n – парне число, дорівнює (8.45):

$$Me(X) = \frac{\frac{x_n + x_n}{2} + 1}{2} = \frac{x_1 + x_n}{2}. \quad (8.45)$$

Доведення. За формулою (6.20) лекції 6 медіана є серединою медіанного інтервалу. Для його визначення будемо функцію розподілу випадкової величини X .

$$F(x) = \begin{cases} 0, & \text{де } x \leq x_1; \\ \frac{1}{n}, & \text{де } x_1 < x \leq x_2; \\ \frac{2}{n}, & \text{де } x_2 < x \leq x_3; \\ \dots & \dots \\ \frac{\sum_{i=1}^{\frac{n}{2}} \frac{1}{n}}{n}, & \text{де } x_{\frac{n}{2}} < x \leq x_{\frac{n}{2}+1}; \\ \dots & \dots \\ \frac{\sum_{i=1}^n \frac{1}{n}}{n}, & \text{де } x > x_n. \end{cases} = \begin{cases} 0, & \text{де } x \leq x_1; \\ \frac{1}{n}, & \text{де } x_1 < x \leq x_2; \\ \frac{2}{n}, & \text{де } x_2 < x \leq x_3; \\ \dots & \dots \\ \frac{1}{2}, & \text{де } x_{\frac{n}{2}} < x \leq x_{\frac{n}{2}+1}; \\ \dots & \dots \\ 1, & \text{де } x > x_n. \end{cases}$$

Отже, $Me(X) = \frac{\frac{x_n + x_n}{2} + 1}{2} = \frac{x_1 + x_n}{2}$. Що й треба було довести.

З означення медіани та особливостей уніформного розподілу, медіану також інколи можна обчислювати за формулою $Me(X) = \frac{x_1 + x_n}{2}$.

Таким чином:

- При парному n уніформний розподіл гарантовано має медіану (формула 8.45);
- Якщо n – непарне число, медіана уніформного розподілу обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $[a; b)$ дорівнює $\frac{1}{2}$, згідно тієї ж формули (6.20) лекції 6.

8. Коефіцієнт асиметрії та ексцес випадкової величини X , що має гіпергеометричний розподіл, доцільніше обчислювати стандартним методом, наведеним в лекції 7.

Дискретний уніформний розподіл часто використовується у випадках, коли кожен можливий результат є рівноймовірним. Це може бути корисним при моделюванні ідеальних ігор на випадок, таких як підкидання монети, кидання гральної кістки, або вибір випадкового елемента з рівноймовірного набору елементів.

Приклад 8.9. Скласти ряд розподілу та обчислити основні числові характеристики моду та медіану випадкової величини X – можливі результати підкидання гральної кістки.

Розв’язання. Якщо підкидати правильну гральну кістку, можливі значення випадкової величини X (результат підкидання) — це випадання чисел 1,2,3,4,5,6.

Ймовірність випадання кожного значення буде:

$$P_6(1) = P_6(2) = P_6(3) = P_6(4) = P_6(5) = P_6(6) = \frac{1}{6}.$$

Ряд розподілу даної випадкової величини X матиме вигляд (табл. 8.16):

Таблиця 8.16. Ряд розподілу дискретної випадкової величини X

x_i	1	2	3	4	5	6
$P_n(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Математичне сподівання для такого випадку:

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

Аналогічно за формулою (8.43)

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

Дисперсія:

$$D(X) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} - (3,5)^2 \approx 2,92.$$

Аналогічно за формулою (8.43)

$$D(X) = \frac{(6 - 1 + 1)^2 - 1}{12} \approx 2,92.$$

За формулами (8.44) та (8.45):

$$Mo(X) = 1,2,3,4,5,6.$$

$$Me(X) = \frac{3 + 4}{2} = \frac{1 + 6}{2} = 3,5.$$

Відповідь. Табл. 8.16; $M(X) = 3,5$; $D(X) \approx 2,92$; $Mo(X) = 1,2,3,4,5,6$; $Me(X) = 3,5$.

Приклад 8.10. Скласти ряд розподілу та обчислити основні числові характеристики, моду, медіану, коефіцієнт асиметрії та ексцес випадкової величини X – можливі результати витягування картки із набору карток з літерами a, b, c, d .

Розв’язання. Якщо витягувати картки із набору карток з літерами a, b, c, d , можливі значення випадкової величини X — це випадання чисел 1,2,3,4 (для зручності поставимо кожній літері у відповідність цифру – її порядковий номер у абетці).

Ймовірність випадання кожного значення буде:

$$P_6(1) = P_6(2) = P_6(3) = P_6(4) = \frac{1}{4}.$$

Ряд розподілу даної випадкової величини X матиме вигляд (табл. 8.17):

Таблиця 8.17. Ряд розподілу дискретної випадкової величини X

x_i	1	2	3	4
$P_n(x_i)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Математичне сподівання для такого випадку:

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

Аналогічно за формулою (8.43)

$$M(X) = \frac{1 + 4}{2} = 2,5.$$

Дисперсія:

$$D(X) = \frac{1^2 + 2^2 + 3^2 + 4^2}{4} - (2,5)^2 = 1,25.$$

Аналогічно за формулою (8.45)

$$D(X) = \frac{(4 - 1 + 1)^2 - 1}{12} = 1,25.$$

$$Mo(X) = 1,2,3,4.$$

$$Me(X) = \frac{2 + 3}{2} = \frac{1 + 4}{2} = 2,5.$$

Обчислимо коефіцієнт асиметрії та ексцес.

$$v_1(X) = M(X) = 2,5; v_2(X) = 7,5; v_3(X) = 25,0; v_4(X) = 88,5.$$

$$\mu_3(X) = v_3(X) - 3 \cdot v_2(X) \cdot v_1(X) + 2 \cdot v_1^3(X) = 25,0 - 56,25 + 31,25 = 0;$$

$$\begin{aligned} \mu_4(X) &= v_4(X) - 4 \cdot v_3(X) \cdot v_1(X) + 6 \cdot v_2(X) \cdot v_1^2(X) - 3 \cdot v_1^4(X) = \\ &= 88,5 - 250 + 281,25 - 117,1875 = 2,5625. \end{aligned}$$

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = 0.$$

Це вказує на те, що даний уніформний розподіл є симетричним.

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = -1,36.$$

Це вказує на те, що уніформний розподіл має меншу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Відповідь. Табл. 8.17; $M(X) = 2,5$; $D(X) = 1,25$; $Mo(X) = 1,2,3,4$; $Me(X) = 2,5$; $As(X) = 0$; $Es(X) = -1,36$.

Приклад 8.11. У певний день продавець кольорової капусти сподівається на такий рівень продаж (кг) (табл. 8.16).

Таблиця 8.16. Ряд розподілу дискретної випадкової величини

продаж (кг)	50	100	200	300
ймовірність	0,1	0,4	0,4	0,1

Продавець купляє капусту по 1,5 грн., продає по 3 грн. У кінці дня залишок товару продає по 0,7 грн. перекупнику. Потрібно: побудувати ряд розподілу випадкової величини X – прибуток від продажу капусти, якщо вранці продавець закупив 130 кг; підрахувати математичне сподівання $M(X)$, дисперсію $D(X)$ і стандартне відхилення $\sigma(X)$.

Розв’язання. Нехай X – прибуток від продажу капусти. Якщо продавець закупив 130 кг капусти, розглянемо за допомогою сценарії залежно від рівня продажу.

Закупівельна ціна: $130 \cdot 1,5 = 195$ грн.

Ціна продажу: по 3 грн за кг.

Ціна залишку: по 0,7 грн за кг.

Відповідно до табл. 8.16 розглянемо наступні випадки:

1. Продаж 50 кг:

Прибуток від продажу: $50 \cdot 3 = 150$ грн.

Залишок капусти: $130 - 50 = 80$ кг.

Прибуток від залишку: $80 \cdot 0,7 = 56$ грн.

Загальний прибуток: $x_1 = 150 + 56 - 195 = 11$ грн.

Ймовірність: $p_1 = 0,1$.

2. Продаж 100 кг:

Прибуток від продажу: $100 \cdot 3 = 300$ грн.

Залишок капусти: $130 - 100 = 30$ кг.

Прибуток від залишку: $30 \cdot 0,7 = 21$ грн.

Загальний прибуток: $x_2 = 300 + 21 - 195 = 126$ грн.

Ймовірність: $p_2 = 0,4$.

3. Продаж 200 кг:

З 200 кг. капусти максимально можна продати 130 кг.

Прибуток від продажу: $130 \cdot 3 = 390$ грн.

Залишок капусти: немає залишку.

Загальний прибуток: $390 - 195 = 195$ грн.

4. Продаж 300 кг:

З 300 кг. капусти максимально можна продати 130 кг.

Прибуток від продажу: $130 \cdot 3 = 390$ грн.

Залишок капусти: немає залишку.

Загальний прибуток: $390 - 195 = 195$ грн.

Для випадків 200 кг і 300 кг, оскільки значення прибутків $x_3 = 195$ грн для них однакові, слід об'єднати їх ймовірності: $p_3 = 0,4 + 0,1 = 0,5$.

Побудуємо розподіл випадкової величини X (табл. 8.17):

Таблиця 8.17. Ряд розподілу випадкової величини X

x_i	11	126	195
p_i	0,1	0,4	0,5

Знайдемо математичне сподівання.

$$M(X) = 11 \cdot 0,1 + 126 \cdot 0,4 + 195 \cdot 0,5 = 149 \text{ грн.}$$

Знайдемо дисперсію.

$$D(X) = (11 - 149)^2 \cdot 0,1 + (126 - 149)^2 \cdot 0,4 + (195 - 149)^2 \cdot 0,5 = \\ = 1904,4 + 211,6 + 1058 = 3174 \text{ грн.}$$

Знайдемо стандартне відхилення.

$$\sigma(X) \approx 56,34 \text{ грн.}$$

Відповідь. табл. 8.17; $M(X) = 149$ грн; $D(X) = 3174$ грн; $\sigma(X) \approx 56,34$ грн.

ЛЕКЦІЯ 9

ОСНОВНІ РОЗПОДІЛИ НЕПЕРЕРВНИХ ВИПАДКОВИХ ВЕЛИЧИН

9.1. Рівномірний розподіл

Рівномірний розподіл неперервної випадкової величини X виникає в випробуваннях типу кидання намання точки на відрізок $[a; b]$ (X — відстань точки від границі a відрізка), або в випробуваннях, зв'язаних з округленням вимірювань фізичних величин за допомогою приладів (X — похибка округлень).

Неперервна випадкова величина X , яка приймає можливі значення з відрізка $[a; b]$, називається **рівномірно розподіленою**, якщо її щільність розподілу має стале значення на цьому відрізку:

$$f(x) = \begin{cases} C, & \text{де } x \in [a; b]; \\ 0, & \text{де } x \notin [a; b]. \end{cases}$$

Значення сталої величини C знаходиться з умови:

$$\int_a^b f(x) dx = C \cdot \int_a^b dx = C \cdot (b - a) = 1, \text{ звідки } C = \frac{1}{b-a}.$$

Отже, **щільність розподілу** рівномірно розподіленої на відрізок $[a; b]$ випадкової величини приймає вигляд (9.1):

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{де } x \in [a; b]; \\ 0, & \text{де } x \notin [a; b]. \end{cases} \quad (9.1)$$

Графік щільності ймовірності приведений на рис. 9.1.

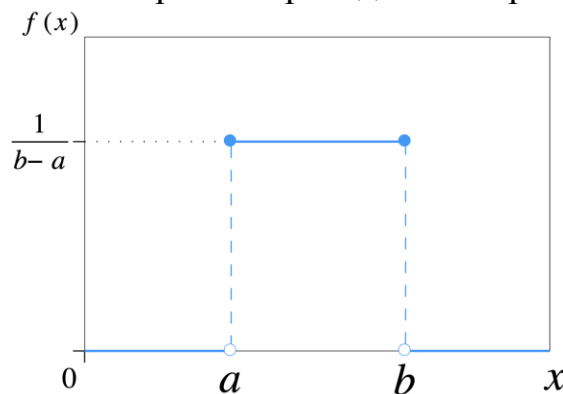


Рис. 9.1. Графік функції щільності $f(x)$ рівномірного розподілу

Властивості рівномірного розподілу

1. Функція розподілу $F(x)$ для випадкової величини X , що має рівномірний розподіл, має вигляд (9.2):

$$F(x) = \begin{cases} 0, & \text{де } x < a; \\ \frac{x-a}{b-a}, & \text{де } a \leq x \leq b; \\ 1, & \text{де } x > b. \end{cases} \quad (9.2)$$

Доведення. Функція розподілу $F(x)$ на відрізку $[a; b]$ знаходиться за стандартною формулою:

$$F(x) = \int_a^x f(x)dx = \frac{1}{b-a} \cdot \int_a^x dx = \frac{x-a}{b-a},$$

а за властивістю 4 функції розподілу $F(x) = 0$ при $x < a$ і $F(x) = 1$ при $x > b$, отже,

$$F(x) = \begin{cases} 0, & \text{де } x < a; \\ \frac{x-a}{b-a}, & \text{де } a \leq x \leq b; \\ 1, & \text{де } x > b. \end{cases}$$

Графік функції розподілу $F(x)$ поданий на рис. 9.2.

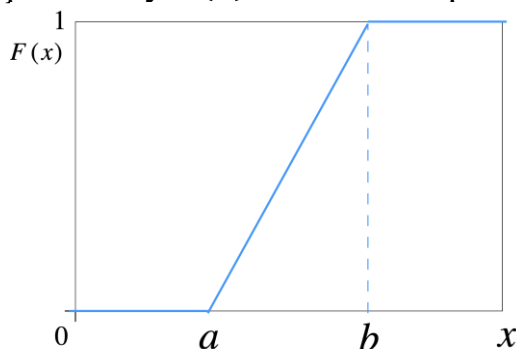


Рис. 9.2. Графік функції розподілу $F(x)$ рівномірного розподілу

2. Числові характеристики $M(X), D(X), \sigma(X)$ рівномірно розподіленої випадкової величини обчислюються за формулами (9.3)-(9.5):

$$M(X) = \frac{a+b}{2}; \quad (9.3)$$

$$D(X) = \frac{(b-a)^2}{12}; \quad (9.4)$$

$$\sigma(X) = \frac{b-a}{2\sqrt{3}}. \quad (9.5)$$

Доведення. Числові характеристики знаходяться відповідно за стандартними формулами для неперервних величин:

$$M(X) = \int_a^b (x \cdot f(x))dx = \frac{1}{b-a} \cdot \int_a^b xdx = \frac{b^2 - a^2}{2 \cdot (b-a)} = \frac{a+b}{2},$$

тобто математичне сподівання $M(X)$ співпадає з серединою відрізка $[a; b]$.

$$\begin{aligned} D(X) &= \int_a^b (x^2 \cdot f(x))dx - (M(X))^2 = \\ &= \frac{1}{b-a} \cdot \int_a^b x^2 dx - \left(\frac{a+b}{2}\right)^2 = \frac{b^3 - a^3}{3 \cdot (b-a)} - \frac{(a+b)^2}{4} = \\ &= \frac{b^2 + a \cdot b + a^2}{3} - \frac{b^2 + 2 \cdot a \cdot b + a^2}{4} = \frac{(b-a)^2}{12}; \end{aligned}$$

$$\sigma(X) = \sqrt{D(X)} = \frac{b-a}{2 \cdot \sqrt{3}}.$$

3. Ймовірність того, що в результаті випробування рівномірно розподілена випадкова величина X прийме можливе значення з інтервалу $(\alpha; \beta)$, який міститься у відрізку $[a; b]$, обчислюється за добре відомою формулою (9.6):

$$P\{\alpha < X < \beta\} = \frac{\beta - \alpha}{b - a}. \quad (9.6)$$

Доведення. За властивість 2 щільності розподілу

$$P\{\alpha < X < \beta\} = \int_{\alpha}^{\beta} f(x) dx = \frac{1}{b-a} \cdot \int_{\alpha}^{\beta} dx = \frac{\beta - \alpha}{b - a}.$$

4. Рівномірний розподіл є багатомодальним.

Доведення. Оскільки в рівномірному розподілі всі значення в інтервалі $[a, b]$ рівноймовірні, то будь-яке значення з інтервалу $[a, b]$ є його модою.

5. Медіана рівномірного розподілу дорівнює середині інтервалу $[a, b]$.

Доведення. Як можна побачити на рис. 9.2, графік щільності розподілу можна візуально розділити навпіл по вертикальній асимптоті, проведений через точку $x = \frac{a+b}{2}$. Отже, вона і буде медіаною рівномірного розподілу.

6. Коефіцієнт асиметрії рівномірного розподілу $As(X) = 0$. Експес рівномірного розподілу $Es(X) = -1,2$. Тобто графік щільності рівномірного розподілу є симетричним та має меншу «піковість» в порівнянні з нормальним розподілом (де експес дорівнює 0).

Доведення. Для визначення коефіцієнту асиметрії обчислимо центральний момент 3-го порядку за формулою (7.24) лекції 7.

$$\begin{aligned} v_1(X) &= \frac{a+b}{2}, \quad v_2(X) = \frac{b^2+a \cdot b+a^2}{3}. \\ v_3(X) &= \int_a^b (x^3 \cdot f(x)) dx = \frac{1}{b-a} \cdot \int_a^b x^3 dx = \frac{b^4 - a^4}{4 \cdot (b-a)} \\ &= \frac{(a+b) \cdot (a^2 + b^2)}{4} = \frac{a^3 + a \cdot b^2 + a^2 \cdot b + b^3}{4}. \\ \mu_3(X) &= v_3(X) - 3 \cdot v_2(X) \cdot v_1(X) + 2 \cdot v_1^3(X) \\ &= \frac{a^3 + a \cdot b^2 + a^2 \cdot b + b^3}{4} - 3 \cdot \frac{b^2 + a \cdot b + a^2}{3} \cdot \frac{a+b}{2} + 2 \cdot \frac{a^3 + 3 \cdot a^2 \cdot b + 3 \cdot a \cdot b^2 + b^3}{8} = 0. \end{aligned}$$

За означенням коефіцієнта асиметрії (7.27) лекції 7

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = 0.$$

Для визначення ексцесу обчислимо центральний момент 4-го порядку за формулою (7.24) лекції 7.

$$\begin{aligned} v_1(X) &= \frac{a+b}{2}, v_2(X) = \frac{b^2+a \cdot b+a^2}{3}, \\ v_3(X) &= \frac{a^3 + a \cdot b^2 + b^2 \cdot a + b^3}{4}, \\ v_4(X) &= \int_a^b (x^4 \cdot f(x)) dx = \frac{1}{b-a} \cdot \int_a^b x^4 dx = \frac{b^5 - a^5}{5 \cdot (b-a)}. \\ \mu_4(X) &= v_4(X) - 4 \cdot v_3(X) \cdot v_1(X) + 6 \cdot v_2(X) \cdot v_1^2(X) - 3 \cdot v_1^4(X) \\ &= \frac{(b-a)^4}{80}, \\ \sigma(X) &= \sqrt{D(X)} = \frac{b-a}{2 \cdot \sqrt{3}}. \end{aligned}$$

За означенням коефіцієнта асиметрії (7.27) лекції 7

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{\frac{(b-a)^4}{80}}{\frac{(b-a)^4}{144}} - 3 = 1,8 - 3 = -1,2.$$

7. Ентропія рівномірного розподілу дорівнює (9.7):

$$H(X) = \log_2(b-a). \quad (9.7)$$

Доведення. Ентропія рівномірного розподілу визначається за формулою:

$$\begin{aligned} H(X) &= - \int_{-\infty}^{+\infty} (f(x) \cdot \log_2(f(x))) dx = \\ &= - \int_a^b \left(\frac{1}{b-a} \cdot \log_2 \left(\frac{1}{b-a} \right) \right) dx = - \left(\frac{1}{b-a} \cdot \log_2 \left(\frac{1}{b-a} \right) \right) \cdot \int_a^b dx = \\ &= \log_2(b-a). \end{aligned}$$

Приклад 9.1. Автобуси «Політ» вирушають до аеропорту Бориспіль з інтервалом 30 хв. Час очікування автобуса на зупинці – випадкова рівномірно розподілена величина X . **1.** Знайти функцію розподілу, числові характеристики цієї випадкової величини, а також ймовірність того, що час очікування для пасажирів, який в випадковий момент підійшов до зупинки, не перевищить 5 хв. **2.** Обчислити інформаційну ентропію випадкової величини.

Розв'язання. 1. Випадкова величина X рівномірно розподілена на відрізьку $[0; 30]$, тому на цьому відрізьку щільність розподілу $f(x) = \frac{1}{30}$, а функція розподілу, згідно з формулою (9.2)

$$F(x) = \begin{cases} 0, & \text{де } x < 0; \\ \frac{x}{30}, & \text{де } 0 \leq x \leq 30; \\ 1, & \text{де } x > 30. \end{cases}$$

Числові характеристики знаходяться за формулами (9.3)-(9.5):

$$M(X) = 15; D(X) = 75; \sigma(X) \approx 8,66,$$

а ймовірність того, що час очікування не перевищить 5 хв. — за формулою (9.6):

$$P\{0 < X < 5\} = \frac{1}{6}.$$

2. За формулою (9.7)

$$H(X) = \log_2 (30) \approx 4,91 \text{ біта}.$$

Відповідь: 1. $M(X) = 15; D(X) = 75; \sigma(X) \approx 8,66; P\{0 < X < 5\} = \frac{1}{6};$ 2. $H(X) \approx 4,91$ біта.

Приклад 9.2. Ребро l кубу виміряне приблизно, причому $a \leq l \leq b$. Розглядаючи довжину ребра куба як неперервну випадкову величину X , розподілену рівномірно на $[a; b]$, знайти математичне сподівання та дисперсію об'єму кубу.

Розв'язання. Складемо функцію щільності для випадкової величини $X = l$.

$$f(X) = \begin{cases} \frac{1}{b-a}, & \text{де } l \in [a; b]; \\ 0, & \text{де } l \notin [a; b]. \end{cases}$$

Враховуючи, що $V_{\text{кубу}} = X^3 = l^3$, обчислимо числові характеристики для випадкової величини X^3 .

$$M(X^3) = \int_a^b \left(l^3 \cdot \frac{1}{b-a} \right) dx = \frac{l^4}{4 \cdot (b-a)} \Big|_a^b = \frac{b^4 - a^4}{4 \cdot (b-a)}.$$

$$M(X^6) = \int_a^b \left(l^6 \cdot \frac{1}{b-a} \right) dx = \frac{l^7}{7 \cdot (b-a)} \Big|_a^b = \frac{b^7 - a^7}{7 \cdot (b-a)}.$$

$$D(X^3) = M(X^6) - (M(X^3))^2 = \frac{b^7 - a^7}{7 \cdot (b-a)} - \left(\frac{b^4 - a^4}{4 \cdot (b-a)} \right)^2 =$$

$$= \frac{9 \cdot a^8 - 16 \cdot a^7 \cdot b - 16 \cdot a \cdot b^7 + 14 \cdot a^4 \cdot b^4 + 9 \cdot b^8}{112}.$$

Відповідь: $M(X^3) = \frac{b^4 - a^4}{4 \cdot (b-a)}; D(X^3) = \frac{b^7 - a^7}{7 \cdot (b-a)} - \left(\frac{b^4 - a^4}{4 \cdot (b-a)} \right)^2.$

9.2. Нормальний розподіл

Нормальний закон розподілу неперервної випадкової величини є одним із найбільш часто вживаних в практичних застосуваннях розподілів. Йому підлягають похибки вимірювань різних фізичних величин, розміри або маса виробів, які сходять з

поточної лінії, тощо. Взагалі, будь-яка випадкова величина, яка являє собою суму багатьох незалежних випадкових величин, кожна з яких відіграє незначну роль в утворенні суми, має нормальний розподіл.

Неперервна випадкова величина X називається розподіленою за **нормальним законом** (або **законом Гауса**) з параметрами a (в деяких підручниках позначається μ) і σ , якщо її **щільність розподілу** має вигляд (9.8):

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \quad (9.8)$$

для всіх $x \in (-\infty; +\infty)$.

Дамо коротку характеристику графіка щільності ймовірності $f(x)$ нормального розподілу, який називається **нормальною кривою** або **кривою Гауса**:

- Оскільки різниця $x - a$ входить в аналітичний вираз $f(x)$ (9.8) в квадраті, **нормальна крива симетрична відносно прямої $x = a$** ;

- Перша похідна $f'(x) = -\frac{x-a}{\sigma^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}}$ в точці $x = a$ перетворюється в нуль, а при переході через цю точку змінює знак з $+$ на $-$, отже, **в точці $x = a$ функція $f(x)$ має максимум $f(a) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}}$** ;

- Друга похідна $f''(x) = \frac{(x-a)^2 - \sigma^2}{\sigma^5 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}}$ в точках $x = a \pm \sigma$ перетворюється в нуль, а при переході через ці точки змінює знак, отже, в точках $x = a \pm \sigma$ графік функції має перегин

$$f(a \pm \sigma) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi} \cdot e};$$

- При $x \rightarrow \pm\infty$ $f(x) \rightarrow 0$, отже, вісь абсцис є лівою і правою асимптотою графіка функції $f(x)$.

Графік щільності ймовірності $f(x)$ поданий на рис. 9.3.

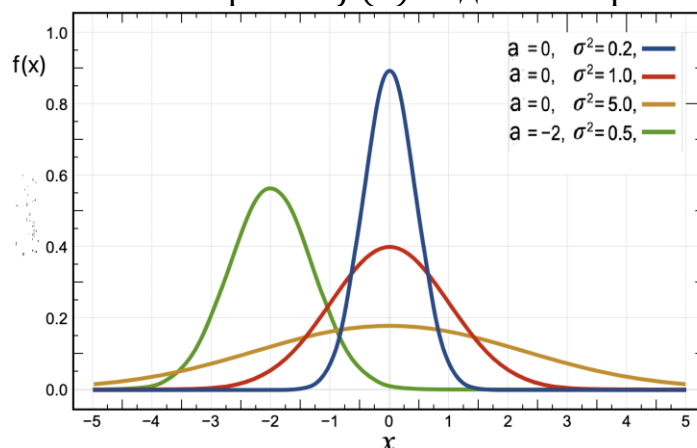


Рис. 9.3. Графік функції щільності $f(x)$ нормального розподілу

Властивості нормального розподілу

1. Функція розподілу $F(x)$ нормально розподіленої випадкової величини X має вигляд (9.9):

$$F(x) = \Phi\left(\frac{x-a}{\sigma}\right) + 0,5. \quad (9.9)$$

Доведення. З формули (9.9) отримуємо:

$$\begin{aligned} F(x) &= P\{X < x\} = P\{-\infty < X < x\} = \Phi\left(\frac{x-a}{\sigma}\right) - \Phi(-\infty) \\ &= \Phi\left(\frac{x-a}{\sigma}\right) + 0,5. \end{aligned}$$

Перевіримо, чи задовольняє функція $F(x)$ властивостям функції розподілу випадкової величини.

$$\text{При } \frac{x-a}{\sigma} \rightarrow -\infty \quad F(x) = -0,5 + 0,5 = 0; \quad \frac{x-a}{\sigma} \rightarrow +\infty \quad F(x) = 0,5 + 0,5 =$$

1.

Графік функції розподілу наведено на рис. 9.4.

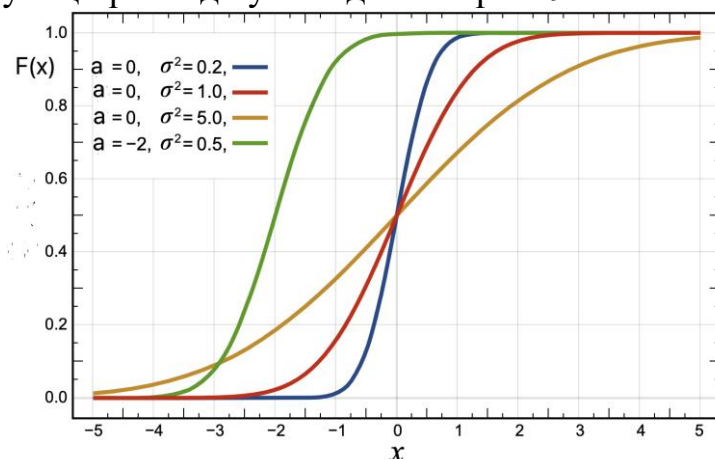


Рис. 9.4. Графік функції розподілу $F(x)$ нормального розподілу

2. Параметри нормального розподілу мають такий ймовірнісний зміст: у нормально розподіленої випадкової величини

$$M(X) = a; \quad D(X) = \sigma^2; \quad \sigma(X) = \sigma.$$

Доведення. Математичне сподівання випадкової величини X :

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} (x \cdot f(x)) dx = \int_{-\infty}^{+\infty} \left(x \cdot \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \right) dx = \\ &\left[\begin{array}{l} t = \frac{x-a}{\sigma \cdot \sqrt{2}} \\ x = t \cdot \sigma \cdot \sqrt{2} + a; dx = (\sigma \cdot \sqrt{2}) dt \end{array} \right] = \int_{-\infty}^{+\infty} \left((t \cdot \sigma \cdot \sqrt{2} + a) \cdot \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-t^2} \right) \cdot \\ &(\sigma \cdot \sqrt{2}) dt = \frac{\sigma \cdot \sqrt{2}}{\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} (t \cdot e^{-t^2}) dt + \frac{a}{\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} e^{-t^2} dt = 0 + \frac{a}{\sqrt{\pi}} \cdot \sqrt{\pi} = a. \end{aligned}$$

Перший інтеграл дорівнює нулю, як інтеграл від непарної функції; другий інтеграл – інтеграл Ейлера-Пуассона $\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$.

$$\begin{aligned}
D(X) &= \int_{-\infty}^{+\infty} ((x-a)^2 \cdot f(x)) dx = \int_{-\infty}^{+\infty} \left((x-a)^2 \cdot \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \right) dx \\
&= \left[\begin{array}{l} t = \frac{x-a}{\sigma \cdot \sqrt{2}} \\ x = t \cdot \sigma \cdot \sqrt{2} + a; dx = (\sigma \cdot \sqrt{2}) dt \end{array} \right] = \\
&= \int_{-\infty}^{+\infty} \left((t \cdot \sigma \cdot \sqrt{2})^2 \cdot \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-t^2} \right) \cdot (\sigma \cdot \sqrt{2}) dt = \\
&= \frac{2 \cdot \sigma^2}{\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} (t^2 \cdot e^{-t^2}) dt = -\frac{\sigma^2}{\sqrt{\pi}} \cdot \int_{-\infty}^{+\infty} t d(e^{-t^2}) =
\end{aligned}$$

Інтегруємо частинами.

$$\begin{aligned}
&= 0 + \frac{\sigma^2}{\sqrt{\pi}} \cdot \sqrt{\pi} = \sigma^2. \\
\sigma(X) &= \sqrt{D(X)} = \sigma.
\end{aligned}$$

3. Ймовірність того, що в результаті випробування нормально розподілена випадкова величина X прийме можливе значення з інтервалу $(\alpha; \beta)$, обчислюється за формулою (9.10):

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \quad (9.10)$$

Доведення. Ймовірність того, що в результаті випробування нормально розподілена випадкова величина X прийме можливе значення з інтервалу $(\alpha; \beta)$, обчислюється за стандартною формулою:

$$P\{\alpha < X < \beta\} = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} dx =$$

Застосовуємо заміну $t = \frac{x-a}{\sigma}$, одержуємо $dx = \sigma dt$ і межі інтегрування для змінної t від $\frac{\alpha-a}{\sigma}$ до $\frac{\beta-a}{\sigma}$, отже,

$$\begin{aligned}
&= \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} \cdot \sigma dt = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \left(\int_{\frac{\alpha-a}{\sigma}}^0 e^{-\frac{t^2}{2}} dt + \int_0^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt \right) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \\
&\quad \cdot \int_0^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_0^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right),
\end{aligned}$$

де $\Phi(x)$ – функція Лапласа (формула (4.5) лекції 4).

Що й треба було довести.

4. Ймовірність того, що відхилення нормально розподіленої випадкової величини X від її математичного сподівання за абсолютним значенням буде меншим заданого ε дорівнює (9.11):

$$P\{|X - a| < \varepsilon\} = 2 \cdot \Phi\left(\frac{\varepsilon}{\sigma}\right). \quad (9.11)$$

Доведення. З формули (9.10) маємо:

$$P\{|X - a| < \varepsilon\} = P\{a - \varepsilon < X < a + \varepsilon\} = \Phi\left(\frac{a + \varepsilon - a}{\sigma}\right) - \Phi\left(\frac{a - \varepsilon - a}{\sigma}\right) = 2 \cdot \Phi\left(\frac{\varepsilon}{\sigma}\right).$$

5. Правило «трьох сігм». Ймовірність того, що випадкова величина приймає значення з інтервалу $(a - 3 \cdot \sigma; a + 3 \cdot \sigma)$, близька до 1, отже, ця подія є практично достовірною. Виконання цього правила дає підставу вважати випадкову величину X розподіленою за нормальним законом.

Доведення. З формули (9.11) випливає:

При $\varepsilon = \sigma$ маємо $P\{|X - a| < \sigma\} = 2 \cdot \Phi(1) \approx 0,6827$;

При $\varepsilon = 2 \cdot \sigma$ маємо $P\{|X - a| < 2 \cdot \sigma\} = 2 \cdot \Phi(2) \approx 0,9545$;

При $\varepsilon = 3 \cdot \sigma$ маємо $P\{|X - a| < 3 \cdot \sigma\} = 2 \cdot \Phi(3) \approx 0,9973$.

Отже, дійсно, ймовірність того, що випадкова величина приймає значення з інтервалу $(a - 3 \cdot \sigma; a + 3 \cdot \sigma)$, близька до 1.

6. Мода та медіана випадкової величини X , що має нормальний розподіл, співпадають з її математичним сподіванням (9.12).

$$Mo(X) = Me(X) = a. \quad (9.12)$$

Доведення. З властивостей функції щільності $f(x)$ випливає, що $x_{max} = a$ (рис. 9.3). Отже $Mo(X) = a$.

Оскільки функції щільності $f(x)$ є симетричною відносно прямої $x = a$, то $Me(X) = a$.

7. Коефіцієнт асиметрії та ексцес випадкової величини X , що має нормальний розподіл дорівнюють: $As(X) = Es(X) = 0$.

Доведення. Оскільки функції щільності $f(x)$ є симетричною відносно прямої $x = a$, то $As(X) = 0$.

Для обчислення ексцесу, враховуємо, що для нормального розподілу всі непарні центральні моменти дорівнюють нулю, а парні моменти можуть бути виражені через дисперсію σ^2 . Зокрема, четвертий центральний момент дорівнює $\mu_4(X) = 3 \cdot \sigma^4$. Тобто

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = 0.$$

8. Інформаційна ентропія випадкової величини X , що має нормальний розподіл, дорівнюють (9.13):

$$H(X) = \frac{1}{2} \cdot \log_2(2 \cdot \pi \cdot e \cdot \sigma^2). \quad (9.13)$$

Доведення. Ентропія нормального розподілу визначається за формулою:

$$\begin{aligned}
H(X) &= - \int_{-\infty}^{+\infty} \left(f(x) \cdot \log_2(f(x)) \right) dx = \\
&= - \int_{-\infty}^{+\infty} \left(\frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \cdot \log_2 \left(\frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \right) \right) dx = \\
&= \frac{1}{2} \cdot \log_2(2 \cdot \pi \cdot e \cdot \sigma^2).
\end{aligned}$$

Приклад 9.3. Нормально розподілена випадкова величина X задана щільністю ймовірності

$$f(x) = \frac{1}{2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-3)^2}{8}}.$$

Знайти: **1.** Скласти функцію розподілу випадкової величини X ; **2.** Обчислити квантиль порядку 0,75; **3.** Інформаційну ентропію випадкової величини.

Розв'язання. За умовою. для випадкової величини X , $a = 3$, $\sigma = 2$.

1. За формулою (9.9)

$$F(x) = \Phi\left(\frac{x-3}{2}\right) + 0,5.$$

2. За означенням квантиля випадкової величини (формула (7.30) лекції 7):

$$\begin{aligned}
F(x_{0,75}) &= 0,75; \\
\Phi\left(\frac{x_{0,75}-3}{2}\right) + 0,5 &= 0,75; \\
\Phi\left(\frac{x_{0,75}-3}{2}\right) &= 0,25; \\
\frac{x_{0,75}-3}{2} &= 0,68; \\
x_{0,75} &= 4,36.
\end{aligned}$$

(значення функції $\Phi(x)$ взяті із таблиці значень функції Лапласа (рис. 4.4 лекції 4).

3. За формулою (9.13)

$$H(X) = \frac{1}{2} \cdot \log_2(2 \cdot \pi \cdot e \cdot \sigma^2) = \frac{1}{2} \cdot \log_2(8 \cdot \pi \cdot e) \approx 3,047 \text{ біта.}$$

Відповідь: 1. $F(x) = \Phi\left(\frac{x-3}{2}\right) + \frac{1}{2}$; 2. $x_{0,75} = 4,36$; 3. $H(X) \approx 3,047$ біта.

Приклад 9.4. За даними відділу технічного контролю 10% виробів підприємства має довжину, меншу 14,7 см, а 20% – довжину, більшу 15,2 см. Довжина виробів – нормально розподілена випадкова величина X . Знайти середній (номінальний) розмір виробів і його середнє квадратичне (стандартне) відхилення.

Розв'язання. За умовою задачі $P\{X < 14,7\} = 0,1$; $P\{X > 15,2\} = 0,2$. Згідно з формулою (9.10)

$$\begin{aligned}
P\{X < 14,7\} &= P\{-\infty < X < 14,7\} = \Phi\left(\frac{14,7 - a}{\sigma}\right) - \Phi(-\infty) = \\
&= \Phi\left(\frac{14,7 - a}{\sigma}\right) + 0,5 = 0,1,
\end{aligned}$$

звідки $\Phi\left(\frac{14,7-a}{\sigma}\right) = -0,4$.

Із таблиці значень функції Лапласа (рис. 4.4 лекції 4) за відомим значенням функції знаходимо

$$\frac{14,7-a}{\sigma} = -1,28 \text{ або } a - 1,28 \cdot \sigma = 14,7.$$

Аналогічно за формулою (9.9)

$$\begin{aligned} P\{X > 15,2\} &= P\{15,2 < X < +\infty\} = \Phi(+\infty) - \Phi\left(\frac{15,2-a}{\sigma}\right) = \\ &= 0,5 - \Phi\left(\frac{15,2-a}{\sigma}\right) = 0,2, \end{aligned}$$

звідки $\Phi\left(\frac{15,2-a}{\sigma}\right) = 0,3$; $\frac{15,2-a}{\sigma} = 0,84$ або $a + 0,84 \cdot \sigma = 15,2$.

Одержана система

$$\begin{cases} a - 1,28 \cdot \sigma = 14,7; \\ a + 0,84 \cdot \sigma = 15,2 \end{cases}$$

має розв'язок $a = 15$; $\sigma = 0,235$.

Відповідь: Середній розмір виробів дорівнює 15 см, а стандартне відхилення 0,235 см.

Приклад 9.5. Систематична похибка утримання висоти літаком складає ± 20 м, а випадкова похибка розподілена нормально з середнім квадратичним відхиленням 75 м. Для польоту літаку надано коридор висотою 100 м. Знайти ймовірність того, що політ буде відбуватись **а)** нижче; **б)** всередині; **в)** вище коридора, якщо літаку задана висота, відповідна середині коридора.

Розв'язання. Позначимо через X сумарну похибку утримання висоти. Її систематична складова ± 20 м, середнє квадратичне відхилення $\sigma = 75$ м. Оскільки літаку задана висота, відповідна середині коридора, для того, щоб політ відбувався нижче коридора, повинно бути $X < -50$, всередині коридора $-50 < X < 50$ і вище коридора $X > 50$.

За формулою (9.10) одержимо

$$\begin{aligned} \text{а)} \quad P\{X < -50\} &= P\{-\infty < X < -50\} = \Phi\left(\frac{-50-20}{75}\right) - \Phi(-\infty) = \\ &= \Phi(-0,93) + 0,5 = -0,3238 + 0,5 = 0,1762; \end{aligned}$$

$$\begin{aligned} \text{б)} \quad P\{-50 < X < 50\} &= \Phi\left(\frac{50-20}{75}\right) - \Phi\left(\frac{-50-20}{75}\right) = \Phi(0,4) - \Phi(-0,93) = \\ &= 0,1554 + 0,3238 = 0,4792; \end{aligned}$$

$$\begin{aligned} \text{в)} \quad P\{X > 50\} &= P\{50 < X < +\infty\} = \Phi(+\infty) - \Phi\left(\frac{50-20}{75}\right) = 0,5 - \Phi(0,4) = \\ &= 0,5 - 0,1554 = 0,3446. \end{aligned}$$

Очевидно, розглянуті події утворюють повну групу, тому їх сумарна ймовірність дорівнює одиниці.

Відповідь: а) $P\{X < -50\} = 0,1762$; б) $P\{-50 < X < 50\} = 0,4792$; в) $P\{X > 50\} = 0,3446$.

Приклад 9.6. Коробка з мармеладом пакується автоматично. В середньому маса однієї коробки становить 1,06 кг. Знайти середнє

квадратичне відхилення, якщо 5% коробок мають масу не більшу за 1 кг. (Вважається, що ваги коробок розподілені за нормальним законом).

Розв'язання. Оскільки 5% коробок з мармеладом важать менше, ніж 1 кг, а $M(X) = 1,06$, то $|X - 1,06| \geq 0,06$.

$$P\{|X - 1,06| \geq 0,06\} = 0,05,$$

тоді, ймовірність протилежної події дорівнює:

$$P\{|X - 1,06| < 0,06\} = 0,95.$$

За формулою (9.11)

$$P\{|X - 1,06| < 0,06\} = 2 \cdot \Phi\left(\frac{0,06}{\sigma}\right) = 0,95.$$

Розв'яжемо рівняння

$$2 \cdot \Phi\left(\frac{0,06}{\sigma}\right) = 0,95;$$

$$\Phi\left(\frac{0,06}{\sigma}\right) = 0,475;$$

$$\frac{0,06}{\sigma} = 1,96;$$

$$\sigma \approx 0,03.$$

Відповідь: $\sigma \approx 0,03$.

Приклад 9.7. Випадкова величина X має нормальний розподіл з математичним сподіванням $a = 25$. Ймовірність попадання випадкової величини X в інтервал $(20; 30)$ дорівнює 0,09. Чому дорівнює ймовірність попадання в інтервал $(35; 40)$?

Розв'язання. Щільність та функція розподілу випадкової величини мають вигляд:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-25)^2}{2 \cdot \sigma^2}}; F(x) = \Phi\left(\frac{x-25}{\sigma}\right) + 0,5.$$

$$20 < X < 30 \Rightarrow 20 - 25 < X - 25 < 30 - 25 \Rightarrow |X - 25| < 5.$$

$$P\{|X - 25| < 5\} = 2 \cdot \Phi\left(\frac{5}{\sigma}\right) = 0,09.$$

Розв'яжемо рівняння

$$2 \cdot \Phi\left(\frac{5}{\sigma}\right) = 0,09;$$

$$\Phi\left(\frac{5}{\sigma}\right) = 0,045;$$

$$\frac{5}{\sigma} = 0,11;$$

$$\sigma \approx 45.$$

Таким чином

$$F(x) = \Phi\left(\frac{x-25}{45}\right) + 0,5.$$

За формулою (9.10) одержимо

$$\begin{aligned} P\{35 < X < 40\} &= \Phi\left(\frac{40 - 25}{45}\right) - \Phi\left(\frac{35 - 25}{45}\right) = \\ &= \Phi\left(\frac{1}{3}\right) - \Phi\left(\frac{2}{9}\right) \approx 0,1293 - 0,0871 \approx 0,0422. \end{aligned}$$

Відповідь: $P\{35 < X < 40\} \approx 0,0422$.

9.3. Показниковий розподіл

Показниковий або експоненціальний розподіл неперервної випадкової величини має широке застосування в теорії надійності технічного обладнання для характеристики терміну безвідмовної роботи елементів та пристроїв і в теорії масового обслуговування для характеристики тривалості обслуговування або технологічних процесів.

Неперервна випадкова величина X називається розподіленою за *показниковим (експоненціальним) законом* з параметром λ , якщо її *щільність розподілу* (9.14)

$$f(x) = \begin{cases} \lambda \cdot e^{-(\lambda \cdot x)}, & \text{де } x > 0; \\ 0, & \text{де } x \leq 0. \end{cases} \quad (9.14)$$

Графік щільності ймовірності $f(x)$ поданий на рис. 9.5.

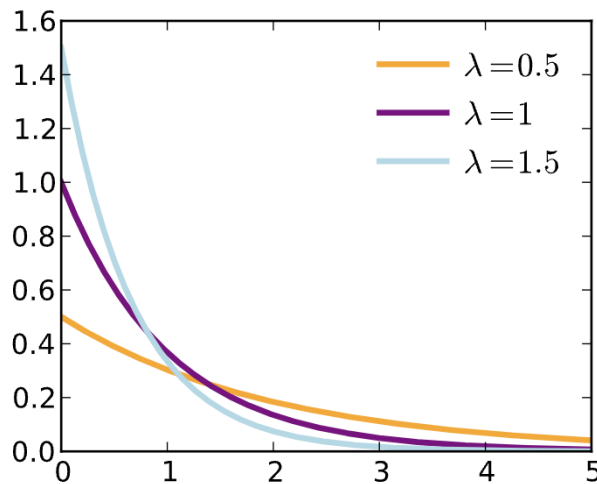


Рис. 9.5. Графік функції щільності $f(x)$ показникового розподілу

Властивості показникового розподілу

1. Функція розподілу $F(x)$ має вигляд (9.15):

$$F(x) = \begin{cases} 1 - e^{-(\lambda \cdot x)}, & \text{де } x > 0; \\ 0, & \text{де } x \leq 0. \end{cases} \quad (9.15)$$

Доведення. При $x > 0$

$$F(x) = \int_0^x f(x) dx = \lambda \cdot \int_0^x e^{-(\lambda \cdot x)} dx = -e^{-(\lambda \cdot x)} \Big|_0^x = 1 - e^{-(\lambda \cdot x)}.$$

При $x \leq 0$

$$F(x) = \int_{-\infty}^x f(x) dx = 0.$$

Графік функції розподілу $F(x)$ подані відповідно на рис. 9.6.

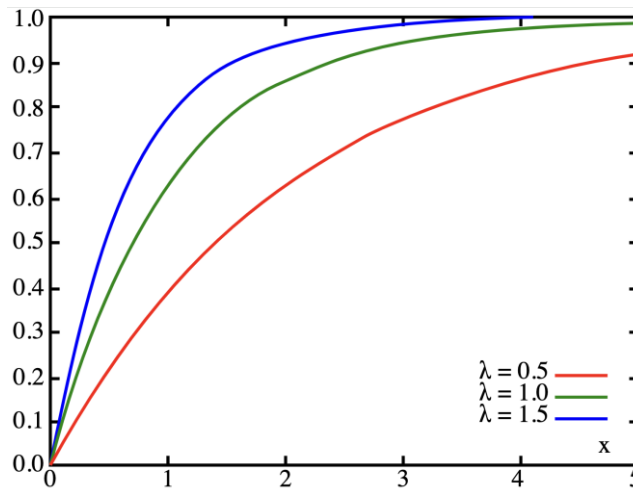


Рис. 9.6. Графік функції розподілу $F(x)$ показникового розподілу

2. Числові характеристики $M(X), D(X), \sigma(X)$ показниково розподіленої випадкової величини обчислюються за формулами (9.16)-(9.18):

$$M(X) = \frac{1}{\lambda}; \quad (9.16)$$

$$D(X) = \frac{1}{\lambda^2}; \quad (9.17)$$

$$\sigma(X) = \frac{1}{\lambda}. \quad (9.18)$$

Доведення. Математичне сподівання $M(X)$ обчислюється за стандартною формулою:

$$\begin{aligned} M(X) &= \lambda \cdot \int_0^{+\infty} (x \cdot e^{-(\lambda \cdot x)}) dx = \\ &= \lambda \cdot \lim_{b \rightarrow +\infty} \int_0^b (x \cdot e^{-(\lambda \cdot x)}) dx = \left[\begin{array}{l} u = x; du = dx \\ dv = e^{-(\lambda \cdot x)} dx; v = -\frac{e^{-(\lambda \cdot x)}}{\lambda} \end{array} \right] \\ &= \lambda \cdot \lim_{b \rightarrow +\infty} \left(-\frac{x \cdot e^{-(\lambda \cdot x)}}{\lambda} \Big|_0^b + \frac{1}{\lambda} \cdot \int_0^b e^{-(\lambda \cdot x)} dx \right) = \\ &= \lambda \cdot \lim_{b \rightarrow +\infty} \left(-\frac{b \cdot e^{-(\lambda \cdot b)}}{\lambda} - \frac{e^{-(\lambda \cdot x)}}{\lambda^2} \Big|_0^b \right) = \\ &= \lambda \cdot \lim_{b \rightarrow +\infty} \left(-\frac{b \cdot e^{-(\lambda \cdot b)}}{\lambda} - \frac{1}{\lambda^2} \cdot (e^{-(\lambda \cdot x)} - 1) \right) = \\ &= \lim_{b \rightarrow +\infty} \left(-b \cdot e^{-(\lambda \cdot b)} - \frac{1}{\lambda} \cdot (e^{-(\lambda \cdot b)} - 1) \right) = \frac{1}{\lambda}, \end{aligned}$$

$\lim_{b \rightarrow +\infty} (b \cdot e^{-(\lambda \cdot b)}) = 0$ за правилом Лопіталя.

Тобто математичне сподівання показникового розподілу є величина, обернена до параметра λ .

Дисперсію $D(X)$ обчислимо за стандартною формулою:

$$\begin{aligned}
 D(X) &= \lambda \cdot \int_0^{+\infty} (x^2 \cdot e^{-(\lambda \cdot x)}) dx - \frac{1}{\lambda^2} = \lambda \cdot \lim_{b \rightarrow +\infty} \int_0^b (x^2 \cdot e^{-(\lambda \cdot x)}) dx - \frac{1}{\lambda^2} = \\
 &= \left[\begin{array}{l} u = x^2; du = (2 \cdot x) dx \\ dv = e^{-(\lambda \cdot x)} dx; v = -\frac{e^{-\lambda \cdot x}}{\lambda} \end{array} \right] = \\
 &= \lambda \cdot \lim_{b \rightarrow +\infty} \left(-\frac{x^2 \cdot e^{-(\lambda \cdot x)}}{\lambda} \Big|_0^b + \frac{2}{\lambda} \cdot \int_0^b (x \cdot e^{-(\lambda \cdot x)}) dx \right) - \frac{1}{\lambda^2} = \\
 &= \lambda \cdot \lim_{b \rightarrow +\infty} \left(-\frac{b^2 \cdot e^{-(\lambda \cdot b)}}{\lambda} + \frac{2}{\lambda} \cdot \int_0^b (x \cdot e^{-(\lambda \cdot x)}) dx \right) - \frac{1}{\lambda^2} = \\
 &= \lim_{b \rightarrow +\infty} \left(-b^2 \cdot e^{-(\lambda \cdot b)} + 2 \cdot \int_0^b (x \cdot e^{-(\lambda \cdot x)}) dx \right) - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
 &= \frac{1}{\lambda^2}, \\
 \lim_{b \rightarrow +\infty} (b^2 \cdot e^{-(\lambda \cdot b)}) &= 0 \text{ за правилом Лопіталя, а } \int_0^b (x \cdot e^{-(\lambda \cdot x)}) dx = \frac{M(X)}{\lambda} = \frac{1}{\lambda^2}. \\
 \sigma(X) &= \sqrt{D(X)} = \frac{1}{\lambda}.
 \end{aligned}$$

3. Ймовірність того, що випадкова величина X , яка має показниковий розподіл, в результаті випробування прийме можливе значення з інтервалу $(a; b)$ при $a > 0$; $b > 0$, дорівнює (9.19):

$$P\{a < X < b\} = e^{-(\lambda \cdot a)} - e^{-(\lambda \cdot b)}. \quad (9.19)$$

Доведення. За стандартною формулою

$$P\{a < X < b\} = \int_a^b f(x) dx = \lambda \cdot \int_a^b e^{-(\lambda \cdot x)} dx = -e^{-(\lambda \cdot x)} \Big|_a^b = e^{-(\lambda \cdot a)} - e^{-(\lambda \cdot b)}.$$

4. Мода та медіана випадкової величини X , що має показниковий розподіл, відповідно дорівнюють (9.20):

$$Mo(X) = 0; Me(X) = \frac{\ln(2)}{\lambda}. \quad (9.20)$$

Доведення. Функція щільності показникової випадкової величини X

$$f(x) = \lambda \cdot e^{-(\lambda \cdot x)} = \frac{\lambda}{e^{(\lambda \cdot x)}}, \text{ де } x > 0,$$

досягає свого максимального значення $f_{max} = \lambda$ в точці $x_{max} = 0$ ($f(0) = \lambda$). Цей факт є очевидним, оскільки вираз $\frac{\lambda}{e^{\lambda \cdot x}} \rightarrow 0$ при $x \rightarrow +\infty$, а, тому немає необхідності досліджувати на екстремуми за допомогою похідних (також це видно з рис. 9.5). Таким чином, $Mo(X) = 0$.

З формули (6.18) лекції 6

$$\begin{aligned}
F(Me(X)) &= 0,5; \\
1 - e^{-(\lambda \cdot Me(X))} &= 0,5; \\
e^{-(\lambda \cdot Me(X))} &= 0,5; \\
-\lambda \cdot Me(X) &= \ln(0,5); \\
-\lambda \cdot Me(X) &= -\ln(2); \\
Me(X) &= \frac{\ln(2)}{\lambda}.
\end{aligned}$$

5. Коефіцієнт асиметрії випадкової величини X , що має показниковий розподіл, $As(X) = 2$. Ексцес випадкової величини X , що має показниковий розподіл, $Es(X) = 6$. Тобто показниковий розподіл не є симетричним (зсув праворуч від моди) в порівнянні з нормальним розподілом (де коефіцієнт асиметрії дорівнює 0). Також даний розподіл має більшу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Доведення. Для визначення коефіцієнту асиметрії обчислимо центральні моменти 3-го порядку. З формули (7.24) лекції 7

$$\begin{aligned}
\mu_3(X) &= M(X - M(X))^3 = M\left(X - \frac{1}{\lambda}\right)^3 = \lambda \cdot \int_0^{+\infty} \left(\left(x - \frac{1}{\lambda}\right)^3 \cdot e^{-(\lambda \cdot x)}\right) dx \\
&= \frac{2}{\lambda^3}.
\end{aligned}$$

Тоді коефіцієнт асиметрії дорівнює:

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = \frac{\frac{2}{\lambda^3}}{\frac{1}{\lambda^3}} = 2.$$

Для визначення ексцесу обчислимо центральні моменти 4-го порядку. З формули (7.25) лекції 7

$$\begin{aligned}
\mu_4(X) &= M(X - M(X))^4 = M\left(X - \frac{1}{\lambda}\right)^4 = \lambda \cdot \int_0^{+\infty} \left(\left(x - \frac{1}{\lambda}\right)^4 \cdot e^{-(\lambda \cdot x)}\right) dx \\
&= \frac{9}{\lambda^4}.
\end{aligned}$$

Тоді ексцес дорівнює:

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{\frac{9}{\lambda^4}}{\frac{1}{\lambda^4}} - 3 = 6.$$

6. Інформаційна ентропія випадкової величини X , що має показниковий розподіл, дорівнюють (9.21):

$$H(X) = \log_2\left(\frac{e}{\lambda}\right). \quad (9.21)$$

Доведення. Ентропія показникового розподілу визначається за формулою:

$$\begin{aligned}
 H(X) &= - \int_{-\infty}^{+\infty} (f(x) \cdot \log_2(f(x))) dx = \\
 &= -\lambda \cdot \int_0^{+\infty} (e^{-(\lambda \cdot x)} \cdot \log_2(\lambda \cdot e^{-(\lambda \cdot x)})) dx = \\
 &= -\lambda \cdot \int_0^{+\infty} (e^{-(\lambda \cdot x)} \cdot (\log_2(\lambda) + \log_2(e^{-(\lambda \cdot x)}))) dx = \\
 &= -\lambda \cdot \int_0^{+\infty} (e^{-(\lambda \cdot x)} \cdot (\log_2(\lambda) - (\lambda \cdot x) \cdot \log_2(e))) dx = \\
 &= -\lambda \cdot \log_2(\lambda) \cdot \int_0^{+\infty} e^{-(\lambda \cdot x)} dx + \lambda^2 \cdot \log_2(e) \cdot \\
 &\quad \cdot \int_0^{+\infty} (x \cdot e^{-(\lambda \cdot x)}) dx = \log_2(e) - \log_2(\lambda) = \log_2\left(\frac{e}{\lambda}\right).
 \end{aligned}$$

Приклад 9.8. Час обслуговування пасажирів в авіакасі – випадкова величина T , розподілена за показниковим законом з середнім значенням, рівним 5 хв. Знайти ймовірність того, що пасажир, який звернувся до каси буде обслуговуватись: **а)** від 2,5 до 5 хв.; **б)** більше 10 хв.

Розв’язання. **а)** За умовою задачі математичне сподівання (середнє значення) $M(T) = 5$, тому за формулою (9.16) параметр розподілу $\lambda = 0,2$. Ймовірність того, що час обслуговування пасажирів буде знаходитись в межах від 2,5 до 5 хв., обчислюється за формулою (9.19):

$$P\{2,5 < X < 5\} = e^{-(0,2 \cdot 2,5)} - e^{-(0,2 \cdot 5)} = e^{-0,5} - e^{-1} \approx 0,2386.$$

б) Ймовірність того, що час обслуговування буде більший 10 хв., також обчислюється за формулою (9.19):

$$P\{X > 10\} = P\{10 < X < +\infty\} = e^{-2} \approx 0,1353.$$

Відповідь. **а)** $P\{2,5 < X < 5\} \approx 0,2386$; **б)** $P\{X > 10\} \approx 0,1353$.

Якщо випадкова величина T з показниковим розподілом – тривалість безвідмовної роботи деякого елемента, а λ – інтенсивність відмов цього елемента, то функція розподілу $F(t) = 1 - e^{-(\lambda \cdot t)}$ ($\lambda > 0$) визначає ймовірність відмови елемента за час t . При цьому функція $R(t) = e^{-(\lambda \cdot t)}$ визначає ймовірність безвідмовної роботи елемента за час t і називається **функцією надійності**.

Приклад 9.9. Тривалість часу безвідмовної роботи елемента системи – випадкова величина T , розподілена за показниковим законом з функцією розподілу $F(t) = 1 - e^{-(0,01 \cdot t)}$. Знайти ймовірність того, що протягом доби елемент: **а)** відмовить; **б)** не відмовить.

Розв’язання. **а)** Розглянемо подію: $A = \{\text{елемент системи відмовить}\}$.

Ймовірність відмови $P(A)$ елемента на протязі доби дорівнює значенню функції розподілу $F(t)$ при $t = 24$ год.:

$$P(A) = F(24) = 1 - e^{-0,24} \approx 0,2134.$$

б) Розглянемо подію: $\bar{A} = \{\text{елемент системи не відмовить}\}$.

Ймовірність невідмови $P(\bar{A})$ елемента на протязі доби дорівнює значенню функції надійності $R(t)$ при $t = 24$ год.:

$$P(\bar{A}) = 1 - P(A) = R(t) = e^{-0,24} \approx 0,7866.$$

Відповідь. а) $P(A) \approx 0,2134$; б) $P(\bar{A}) \approx 0,7866$.

Приклад 9.10. Можна вважати, що час роботи електричної лампи до перегорання – це неперервна випадкова величина X , що має показниковий тип розподілу. Припускаючи, що з ймовірністю p лампа може перегоріти при включенні, знайти середній час роботи лампи (лампа вмикається лише один раз).

Розв’язання. Нехай T – момент часу, коли лампа перегорить.

$$P\{0 < X < T\} = 1 - e^{-\left(T \cdot \frac{1}{M(X)}\right)} = p;$$

$$1 - e^{-\left(T \cdot \frac{1}{M(X)}\right)} = p;$$

$$e^{-\left(T \cdot \frac{1}{M(X)}\right)} = 1 - p;$$

$$-\left(T \cdot \frac{1}{M(X)}\right) = \ln(q);$$

$$M(X) = -\frac{T}{\ln(q)}.$$

Відповідь. $M(X) = -\frac{T}{\ln(q)}.$

9.4. Логнормальний розподіл

Логнормальний розподіл — це ймовірнісний розподіл випадкової величини, логарифм якої має нормальний розподіл. Іншими словами, якщо випадкова величина X має логнормальний розподіл, то $\ln(X)$ буде мати нормальний розподіл.

Нехай розподіл випадкової величини X задається *щільністю ймовірності*, що має вигляд (9.22):

$$f(x) = \frac{1}{\sigma \cdot x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}}, \text{ де } x > 0; a > 0. \quad (9.22)$$

Тоді кажуть, що X має *логнормальний розподіл* з параметрами a і σ .

Графік щільності ймовірності $f(x)$ поданий на рис. 9.7.

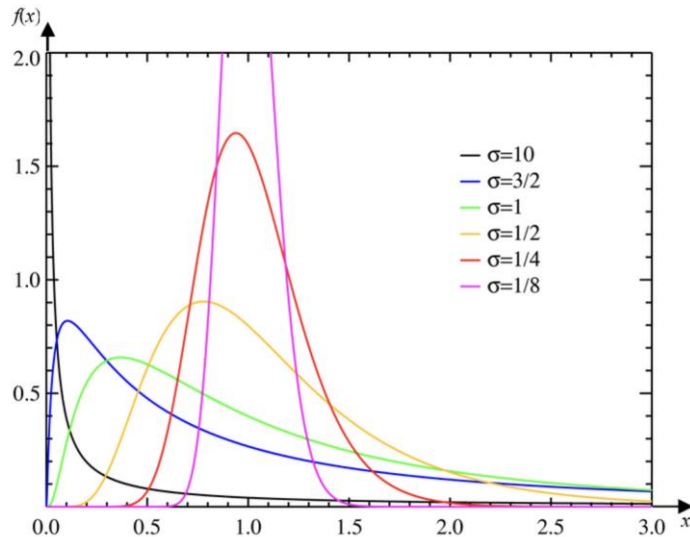


Рис. 9.7. Графік функції щільності $f(x)$ логнормального розподілу

Властивості логнормального розподілу

1. Функція розподілу $F(x)$ має вигляд (9.23):

$$F(x) = \Phi\left(\frac{\ln(x) - \ln(a)}{\sigma}\right) + 0,5. \quad (9.23)$$

Доведення. Оскільки при $x > 0$ нерівності $X < x$ та $\ln(X) < \ln(x)$ рівносильні, то функція розподілу логнормального розподілу співпадає з функцією розподілу нормального розподілу для випадкової величини $\ln(X)$, тобто

$$\begin{aligned} F(x) &= P\{\ln(X) < \ln(x)\} = P\{-\infty < \ln(X) < \ln(x)\} = \\ &= \Phi\left(\frac{\ln(x) - \ln(a)}{\sigma}\right) - \Phi(-\infty) = \\ &= \Phi\left(\frac{\ln(x) - \ln(a)}{\sigma}\right) + 0,5, \end{aligned}$$

Графік функції розподілу $F(x)$ подані відповідно на рис. 9.8.

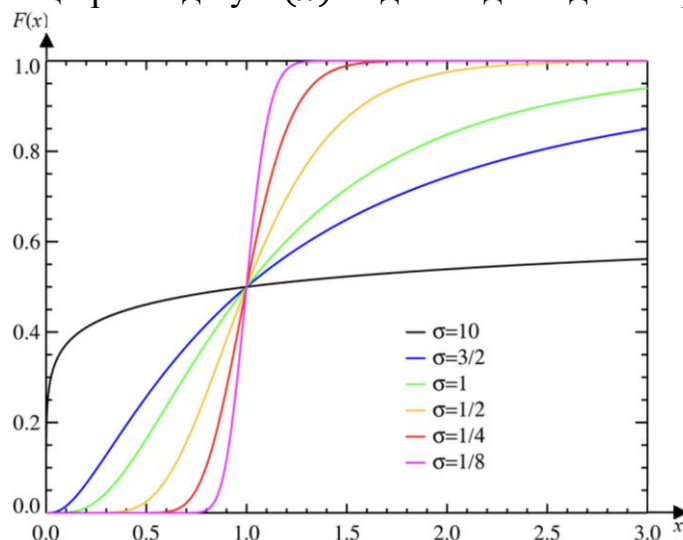


Рис. 9.8. Графік функції розподілу $F(x)$ показникового розподілу

2. Числові характеристики $M(X), D(X), \sigma(X)$ логнормально розподіленої випадкової величини обчислюються за формулами (9.24)-(9.26):

$$M(X) = a \cdot e^{\frac{\sigma^2}{2}}; \quad (9.24)$$

$$D(X) = a^2 \cdot e^{\sigma^2} \cdot (e^{\sigma^2} - 1); \quad (9.25)$$

$$\sigma(X) = a \cdot e^{\frac{\sigma^2}{2}} \cdot \sqrt{e^{\sigma^2} - 1}. \quad (9.26)$$

Доведення.

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} (x \cdot f(x)) dx = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_0^{+\infty} e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} dx \\ &= \left[\frac{\ln(x) - \ln(a)}{\sigma} = t; x = e^{(\ln(a) + t \cdot \sigma)} \right] = \\ &= \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} \cdot (\sigma \cdot e^{(\ln(a) + t \cdot \sigma)}) dt = \end{aligned}$$

Після спрощення

$$e^{-\left(\frac{t^2}{2} - \ln(a) - t \cdot \sigma\right)} = a \cdot e^{-\frac{1}{2} \cdot (t^2 - 2 \cdot t \cdot \sigma)} = a \cdot e^{-\frac{1}{2} \cdot ((t - \sigma)^2 - \sigma^2)} = a \cdot e^{-\frac{(t - \sigma)^2}{2}} \cdot e^{\frac{\sigma^2}{2}}$$

отримуємо

$$= a \cdot e^{\frac{\sigma^2}{2}} \cdot \left(\frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{(t - \sigma)^2}{2}} dt \right) = a \cdot e^{\frac{\sigma^2}{2}} \cdot 1 = a \cdot e^{\frac{\sigma^2}{2}}.$$

$$\text{Інтеграл } \int_{-\infty}^{+\infty} e^{-\frac{(t - \sigma)^2}{2}} dt = \left[u = t - \sigma \right] = \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du = \sqrt{2 \cdot \pi}.$$

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

Інтеграл

$$\begin{aligned} \int_{-\infty}^{+\infty} (x^2 \cdot f(x)) dx &= \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \int_0^{+\infty} \left(x \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} \right) dx = e^{(2 \cdot \ln(a) + 2 \cdot \sigma^2)} = \\ &= a^2 \cdot e^{(2 \cdot \sigma^2)}. \end{aligned}$$

обчислюється аналогічно інтегралу при обчисленні математичного сподівання (та сама заміна та спрощення степеня експоненти).

$$= a^2 \cdot e^{2 \cdot \sigma^2} - a^2 \cdot e^{\sigma^2} = a^2 \cdot e^{\sigma^2} \cdot (e^{\sigma^2} - 1).$$

$$\sigma(X) = \sqrt{D(X)} = a \cdot e^{\frac{\sigma^2}{2}} \cdot \sqrt{e^{\sigma^2} - 1}.$$

3. Ймовірність того, що в результаті випробування логнормально розподілена випадкова величина X прийме можливе значення з інтервалу $(\alpha; \beta)$, обчислюється за формулою (9.27):

$$P\{\alpha < X < \beta\} = \Phi\left(\frac{\ln(\beta) - \ln(a)}{\sigma}\right) - \Phi\left(\frac{\ln(\alpha) - \ln(a)}{\sigma}\right). \quad (9.27)$$

Доведення. Оскільки функція розподілу логнормального розподілу співпадає з функцією розподілу нормального розподілу для випадкової величини $\ln(X)$, то ймовірність того, що в результаті випробування логнормально розподілена випадкова величина X прийме можливе значення з інтервалу $(\alpha; \beta)$, обчислюється за стандартною формулою:

$$\begin{aligned} P\{\alpha < X < \beta\} &= F(\beta) - F(\alpha) = \\ &= 0,5 + \Phi\left(\frac{\ln(\beta) - \ln(a)}{\sigma}\right) - \left(0,5 + \Phi\left(\frac{\ln(\alpha) - \ln(a)}{\sigma}\right)\right) = \\ &= \Phi\left(\frac{\ln(\beta) - \ln(a)}{\sigma}\right) - \Phi\left(\frac{\ln(\alpha) - \ln(a)}{\sigma}\right) \end{aligned}$$

де $\Phi(x)$ – функція Лапласа (формула (4.5) лекції 4). Що й треба було довести.

4. *Мода та медіана випадкової величини X , що має логнормальний розподіл, відповідно дорівнюють (9.28):*

$$Mo(X) = a \cdot e^{-\sigma^2}; Me(X) = a. \quad (9.28)$$

Доведення. Обчислимо моду.

$$\begin{aligned} \text{Знайдемо похідну функції } f(x) &= \frac{1}{\sigma \cdot x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}}. \\ f'(x) &= -\frac{1}{\sigma \cdot x^2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} - \frac{\ln(x) - \ln(a)}{\sigma^3 \cdot x^2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} = \\ &= -\frac{\sigma^2 + \ln(x) - \ln(a)}{\sigma^3 \cdot x^2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}}. \end{aligned}$$

Знайдемо стаціонарні точки.

$$\sigma^2 + \ln(x) - \ln(a) = 0;$$

$$\ln(x) = -\sigma^2 + \ln(a)$$

$$x = a \cdot e^{-\sigma^2}.$$

Обчислимо 2-гу похідну:

$$\begin{aligned}
f''(x) &= - \frac{\sigma^3 \cdot x \cdot \sqrt{2 \cdot \pi} - 2 \cdot x \cdot \sigma^3 \cdot \sqrt{2 \cdot \pi} \cdot (\sigma^2 + \ln(x) - \ln(a))}{(\sigma^3 \cdot x^2 \cdot \sqrt{2 \cdot \pi})^2} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} + \frac{(\sigma^2 + \ln(x) - \ln(a)) \cdot (\ln(x) - \ln(a))}{\sigma^5 \cdot x^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} = \\
&= - \frac{1 - 2 \cdot (\sigma^2 + \ln(x) - \ln(a))}{\sigma^3 \cdot x^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}} + \frac{(\sigma^2 + \ln(x) - \ln(a)) \cdot (\ln(x) - \ln(a))}{\sigma^5 \cdot x^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x) - \ln(a))^2}{2 \cdot \sigma^2}}. \\
f''(a \cdot e^{-\sigma^2}) &= \\
&= - \frac{1 - 2 \cdot (\sigma^2 - \sigma^2)}{\sigma^3 \cdot (a \cdot e^{-\sigma^2})^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{\sigma^2}{2}} - \frac{(\sigma^2 - \sigma^2) \cdot \sigma^2}{\sigma^5 \cdot (a \cdot e^{-\sigma^2})^3 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{\sigma^2}{2}} = - \frac{e^{(2,5 \cdot \sigma^2)}}{\sigma^3 \cdot a^3 \cdot \sqrt{2 \cdot \pi}} < 0.
\end{aligned}$$

Тобто $x = a \cdot e^{-\sigma^2}$ – точка максимуму. Отже $Mo(X) = a \cdot e^{-\sigma^2}$.
Обчислимо медіану.

$$\begin{aligned}
F(Me(X)) &= 0,5; \\
\Phi\left(\frac{\ln(Me(X)) - \ln(a)}{\sigma}\right) + 0,5 &= 0,5; \\
\Phi\left(\frac{\ln(Me(X)) - \ln(a)}{\sigma}\right) &= 0; \\
\frac{\ln(Me(X)) - \ln(a)}{\sigma} &= 0; \\
\ln(Me(X)) - \ln(a) &= 0; \\
Me(X) &= a.
\end{aligned}$$

5. Чим менше значення параметру σ , то ближчими один до одного є значення моди, медіани та математичного сподівання, а крива розподілу – ближчою до симетрії. Якщо в нормальному законі параметр a виступає як середнє значення випадкової величини, то в логнормальному – як медіана.

6. Коефіцієнт асиметрії та ексцес випадкової величини X , що має логнормальний розподіл, відповідно дорівнюють (9.29):

$$As(X) = (e^{\sigma^2} + 2) \cdot \sqrt{e^{\sigma^2} - 1}; Es(X) = e^{(4 \cdot \sigma^2)} + 2 \cdot e^{(3 \cdot \sigma^2)} + 3 \cdot e^{(2 \cdot \sigma^2)} - 6. \quad (9.29)$$

Доведення. Довести формули (9.29) самостійно.

7. Інформаційна ентропія випадкової величини X , що має логнормальний розподіл, дорівнюють (9.30):

$$H(X) = \frac{1}{2} \cdot \ln(2 \cdot e \cdot \pi \cdot \sigma^2 \cdot a^2). \quad (9.30)$$

Доведення. Ентропія логнормального розподілу визначається за формулою:

$$\begin{aligned} H(X) &= - \int_{-\infty}^{+\infty} (f(x) \cdot \ln(f(x))) dx = \\ &= \int_0^{+\infty} \left(\frac{1}{\sigma \cdot x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x)-\ln(a))^2}{2 \cdot \sigma^2}} \right. \\ &\quad \cdot \ln \left(\frac{1}{\sigma \cdot x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x)-\ln(a))^2}{2 \cdot \sigma^2}} \right) \Bigg) dx = \\ &= \frac{1}{2} \cdot \ln(2 \cdot \pi \cdot e \cdot \sigma^2 \cdot a^2). \end{aligned}$$

Логнормальний розподіл використовують для опису розподілу доходів, банківських вкладів, цін активів, місячної заробітної плати, посівних площ під різні культури, довговічності виробів у режимі зношування та старіння тощо.

Приклад 9.11. Проведене дослідження показало, що вклади населення, розміщені в даному банку, можна описати випадковою величиною X , що має логнормальний розподіл з параметрами: $a = 530$; $\sigma^2 = 0,64$. Знайти: **а)** середній розмір вкладу; **б)** долю вкладників, розмір вкладу яких складає не менше 1000 грн.; **в)** моду та медіану випадкової величини X .

Розв'язання. а) Середній розмір вкладу – математичне сподівання, тому:

$$M(X) = 530 \cdot e^{0,32} \approx 730 \text{ (грн.)}$$

б) Доля вкладників, розмір вкладу яких не менше 1000 грн., обчислюється за формулою (9.27).

$$\begin{aligned} P\{X \geq 1000\} &= P\{1000 \leq X < +\infty\} = \Phi(+\infty) - \Phi\left(\frac{\ln(1000) - \ln(530)}{0,8}\right) \\ &= 0,5 - \Phi(0,7936) \approx 0,5 - 0,285 \approx 0,2148. \end{aligned}$$

в) За формулою (9.28)

$$Mo(X) = 530 \cdot e^{-0,64} \approx 280 \text{ (грн.)};$$

$$Me(X) = 530 \text{ (грн.)}$$

Відповідь. а) $M(X) \approx 730$ грн; б) $P\{X \geq 1000\} \approx 0,2148$; в) $Mo(X) \approx 280$ грн; $Me(X) = 530$ грн.

Приклад 9.12. Обчислити інформаційну ентропію випадкової величини X , що має логнормальний розподіл з параметрами: $a = 1$; $\sigma^2 = 1$, а також її коефіцієнт асиметрії та ексцес.

Розв'язання. За формулою (9.30)

$$H(X) = \frac{1}{2} \cdot \ln(2 \cdot e \cdot \pi \cdot 1^2 \cdot 1^2) = \frac{1}{2} \cdot \ln(2 \cdot e \cdot \pi) \approx 1,418.$$

За формулою (9.29)

$$As(X) = (e^1 + 2) \cdot \sqrt{e^1 - 1} = (e + 2) \cdot \sqrt{e - 1} \approx 6,185;$$

$$Es(X) = e^4 + 2 \cdot e^3 + 3 \cdot e^2 - 6 \approx 119,936.$$

Тобто даний розподіл не є симетричним (зсув праворуч від моди) в порівнянні з нормальним розподілом (де коефіцієнт асиметрії дорівнює 0).

Також даний розподіл має більшу «піковість» або «випуклість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Адекватність отриманих результатів можна побачити на рис. 9.9.

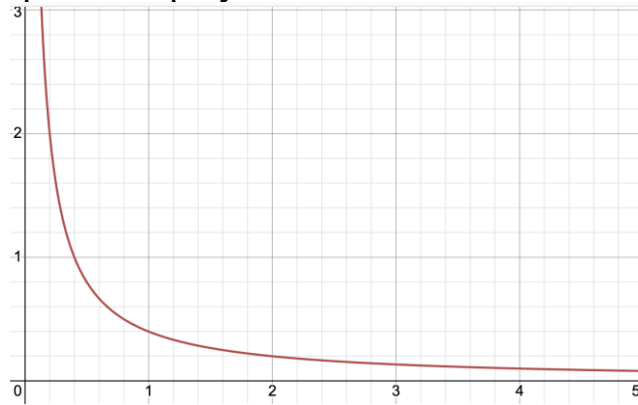


Рис. 9.9. Графік функції $f(x) = \frac{1}{x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(\ln(x))^2}{2}}$, де $x > 0$

Відповідь. $H(X) \approx 1,418$; $As(X) \approx 6,185$; $Es(X) \approx 119,936$.

ЛЕКЦІЯ 10

ОСНОВНІ РОЗПОДІЛИ НЕПЕРЕВНИХ ВИПАДКОВИХ ВЕЛИЧИН (ПРОДОВЖЕННЯ)

10.1. Степінь свободи

Степінь свободи (*degrees of freedom*) в розподілі випадкової величини — це параметр, який визначає кількість незалежних змінних або параметрів, які можуть бути вільно змінювані в межах певної статистичної моделі або розподілу.

Для статистичної вибірки степінь свободи обчислюється за формулою (10.1):

$$k = n - 1, \quad (10.1)$$

де n — кількість елементів у виборці або кількість інтервалів в групованому варіативному ряду.

Уявімо, що ми маємо вибрати три числа, середнє значення яких дорівнює 10. Прикладом можуть слугувати наступні трійки чисел: {9,10,11}, або {8,10,12}, або {5,10,15}. Як тільки було вибрано перші два числа — третє є фіксованим. Іншими словами, ми не можемо змінювати третє число. Лише перші два можуть змінювати свої значення так, аби середнє значення залишалося рівним 10.

В математичній статистиці при перевірці гіпотез про тип розподілу степінь свободи рахуватиметься за формулою (10.2):

$$k = n - m - 1, \quad (10.2)$$

де:

n — об'єм вибірки;

m — кількість параметрів, що описують розподіл.

Наприклад, нормальний розподіл має степінь свободи $k = n - 2 - 1 = n - 3$, оскільки даний розподіл залежить від 2-х параметрів: μ, σ . Для показникового розподілу степінь свободи $k = n - 1 - 1 = n - 2$, оскільки даний розподіл залежить від 1-го параметру λ .

Розуміння ступеня свободи допомагає в правильному інтерпретуванні статистичних тестів і побудові довірчих інтервалів.

10.2. Розподіл χ^2

Розподіл χ^2 (хі-квадрат) з k степенем свободи — це статистичний розподіл, який часто використовується в гіпотезах та тестах для оцінки відповідності між спостережуваними і очікуваними частотами або для перевірки незалежності у крос-табуляціях. Він є особливим випадком розподілу, що виникає як сума квадратів k незалежних нормально розподілених випадкових змінних Z_i ($i = 1, 2, \dots, k$), кожна з яких має середнє 0 і стандартне відхилення 1, тобто $N(0; 1)$.

$$\chi^2 = \sum_{i=1}^k Z_i^2. \quad (10.3)$$

Щільність ймовірності розподілу χ^2 має вигляд (10.4):

$$f(x) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \cdot \Gamma(\frac{k}{2})} \cdot x^{\frac{k}{2}-1} \cdot e^{-\frac{x}{2}}, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.4)$$

В формулі (10.4) $\Gamma(y) = \int_0^{+\infty} (e^{-t} \cdot t^{y-1}) dt$ – гамма-функція Ейлера.

Графік функції щільності розподілу χ^2 наведено на рис. 10.1.

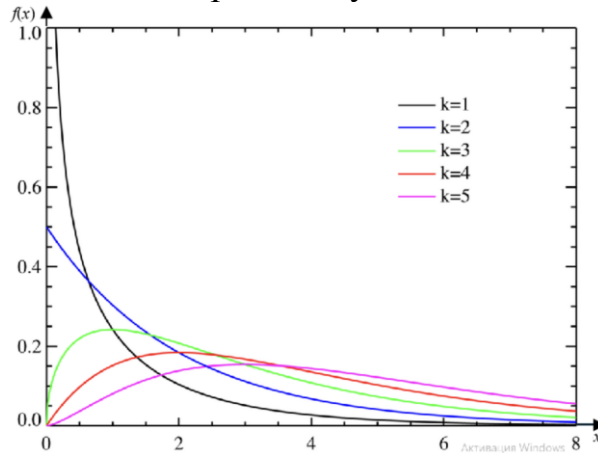


Рис. 10.1. Графік функції щільності $f(x)$ розподілу χ^2

Властивості розподілу χ^2

1. Функція розподілу $F(x)$ матиме вигляд (10.5):

$$F(x) = \begin{cases} \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.5)$$

Доведення. Перевіримо відповідність функції $F(x)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу χ^2 завжди монотонно зростає.

Графік функції розподілу для розподілу χ^2 наведено на рис. 10.2.

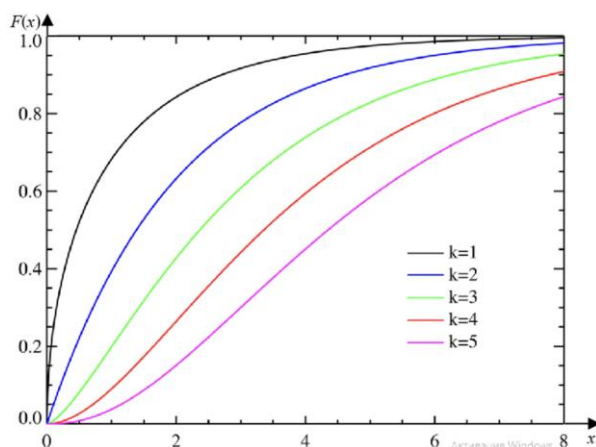


Рис. 10.2. Графік функції розподілу $F(x)$ розподілу χ^2

На рис. 10.2 наведено фрагмент таблиці критичних значень розподілу χ^2 для степеню свободи k та рівня значущості α .

k	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928

Рис. 10.3 Таблиця критичних значень розподілу χ^2

2. Числові характеристики $M(X)$, $D(X)$, $\sigma(X)$ випадкової величини, що має розподіл χ^2 , обчислюються за формулами (10.6):

$$M(X) = k; D(X) = 2 \cdot k; \sigma(X) = \sqrt{2 \cdot k}. \quad (10.6)$$

3. Мода та медіана випадкової величини X , що має розподіл χ^2 , співпадають з її математичним сподіванням (10.7).

$$Mo(X) = \begin{cases} 0, & \text{де } k < 2; \\ k - 2, & \text{де } k \geq 2; \end{cases} Me(X) \approx k \cdot \left(1 - \frac{2}{9 \cdot k}\right)^3. \quad (10.7)$$

4. Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл χ^2 , дорівнюють (10.8):

$$As(X) = \sqrt{\frac{8}{k}}, Es(X) = \frac{3}{k}. \quad (10.8)$$

Тобто, розподіл χ^2 має додатну (правосторонню) асиметрію та має більшу «піковість» або «випуклість» в порівнянні з нормальним розподілом.

Доведення. Обчислимо центральні моменти 3-го та 4-го порядків.

$$\mu_3(X) = 8 \cdot k; \mu_4(X) = 12 \cdot k \cdot (k + 1).$$

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = \frac{8 \cdot k}{(\sqrt{2 \cdot k})^3} = \sqrt{\frac{8}{k}}.$$

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{12 \cdot k \cdot (k + 1)}{(\sqrt{2 \cdot k})^4} - 3 = \frac{12 \cdot k \cdot (k + 1)}{4 \cdot k^2} - 3 = \frac{3}{k}.$$

Для розподілу χ^2 (формули (10.8)) очевидно, що $As(X) > 0$, тобто «довша» частина кривої знаходиться правіше моди, порівняно з графіком щільності нормального розподілу (де коефіцієнт асиметрії дорівнює 0).

Для розподілу χ^2 (формули (10.8)) очевидно, що $Es(X) > 0$, тобто графік функції щільності має більшу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

5. При степенях свободи $k > 30$ розподіл χ^2 наближається до стандартного нормального розподілу $N(0; 1)$.

Доведення. З попередньої властивості (формула (10.8)) видно, що при $k \rightarrow +\infty$ $As(X) = 0$; $Es(X) = 0$. Це свідчить про те, що розподіл χ^2 за рівнем асиметрії та «піковості» або «випуклості» дійсно стає подібним до нормального розподілу.

Визначення степеня свободи (k) для χ^2 -розподілу

Тест χ^2 для відповідності (Goodness-of-Fit Test):

Використовується для перевірки того, наскільки спостережувані частоти відрізняються від очікуваних частот на основі певного теоретичного розподілу.

Ступінь свободи визначається за формулою (10.1):

$$k = n - 1,$$

де n — кількість категорій або груп (в деяких випадках n — об'єм статистичної вибірки або n — кількість інтервалів в групованому статистичному ряду). Це тому, що одна з категорій залежить від інших, тому маємо $n - 1$ незалежних категорій.

Тест χ^2 для незалежності (Test of Independence):

Використовується для перевірки, чи існує залежність між двома категоріальними змінними.

Ступінь свободи визначається за формулою (10.9):

$$k = (r - 1) \cdot (c - 1), \quad (10.9)$$

де:

r — кількість рядків;

c — кількість стовпців у таблиці спряженості (крос-табуляції).

Це враховує той факт, що і ряди, і стовпці підпорядковуються обмеженням.

Тест χ^2 для перевірки однорідності (Test for Homogeneity):

Використовується для порівняння розподілів двох або більше груп.

Ступінь свободи також визначається за формулою (10.9),

$$k = (r - 1) \cdot (c - 1),$$

де:

r — кількість рядків;

c — кількість стовпців у крос-табуляції.

Використання розподілу χ^2

Тест χ^2 для відповідності: Використовується для перевірки того, чи відповідають спостережувані частоти певному теоретичному розподілу.

Тест χ^2 для незалежності: Використовується для перевірки незалежності двох категоріальних змінних у крос-табуляції.

Тест χ^2 для добротності підгонки: Застосовується для перевірки того, чи спостережувані дані відповідають певній моделі.

10.3. Розподіл Стьюдента

Розподіл Стьюдента (або *t-розподіл*) — це розподіл випадкової величини t , яка має вигляд (10.10):

$$t = \frac{Z}{\sqrt{\frac{\chi^2}{k}}}, \quad (10.10)$$

де:

Z — випадкова величина, розподілена за стандартним нормальним законом ($N(0; 1)$);

χ^2 — незалежна від Z випадкова величина, що має розподіл χ^2 з k степенями свободи.

Вона використовується для оцінки ймовірностей і перевірки гіпотез у випадках, коли розмір вибірки невеликий, а справжнє стандартне відхилення невідоме. Він названий на честь Вільяма Сілі Госсета, який використовував псевдонім «Стьюдент».

Щільність ймовірності розподілу Стьюдента має вигляд (10.11):

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \cdot \sqrt{\pi} \cdot k} \cdot \left(1 + \frac{x^2}{k}\right)^{-\left(\frac{k+1}{2}\right)}. \quad (10.11)$$

В формулі (10.11) $\Gamma(y) = \int_0^{+\infty} (e^{-t} \cdot t^{y-1}) dt$ — гамма-функція Ейлера.

Графік функції щільності розподілу Стьюдента наведено на рис.

10.4.

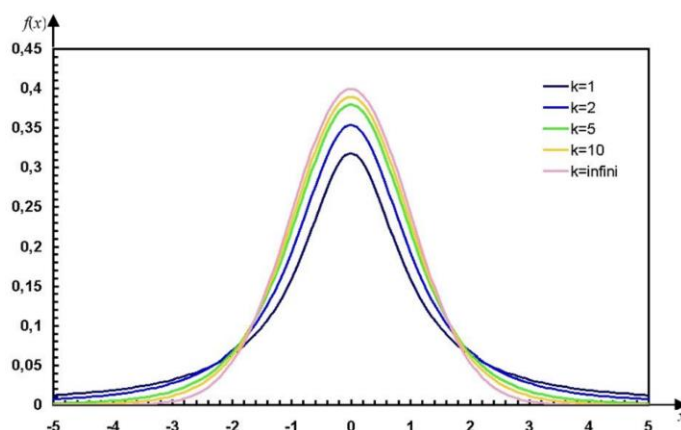


Рис. 10.4. Графік функції щільності $f(x)$ розподілу Стюдента

Властивості розподілу Стюдента

1. Функція розподілу $F(x)$ матиме вигляд (10.12):

$$F(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \cdot \sqrt{\pi \cdot k}} \cdot \int_0^{+\infty} \left(1 + \frac{x^2}{n}\right)^{-\left(\frac{k+1}{2}\right)} dx. \quad (10.12)$$

Доведення. Перевіримо відповідність функції $F(x; k)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу Стюдента завжди монотонно зростає.

Графік функції розподілу для розподілу Стюдента наведено на рис. 10.5.

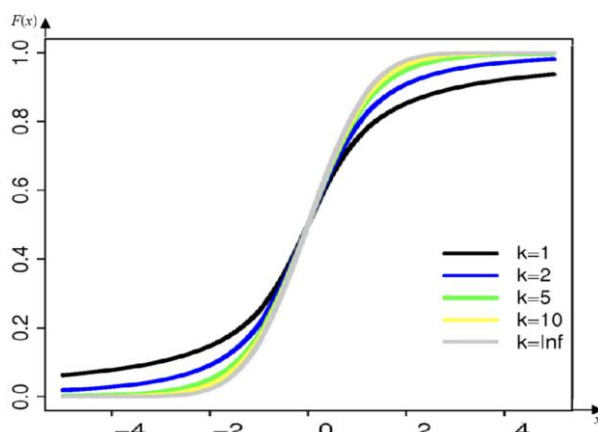


Рис. 10.5. Графік функції розподілу $F(x)$ розподілу Стюдента

На рис. 10.6 наведено таблицю критичних значень розподілу Стюдента для степеню свободи k та рівня значущості α .

Число ступенів свободи k	Рівень значимості α			
	0,10	0,05	0,01	0,001
1	6,31	12,70	63,70	637,00
2	2,92	4,30	9,92	31,60
3	2,35	3,18	5,84	12,90
4	2,13	2,78	4,60	8,61
5	2,01	2,57	4,03	6,86
6	1,94	2,45	3,71	5,96
7	1,89	2,36	3,50	5,40
8	1,86	2,31	3,36	5,04
9	1,83	2,26	3,25	4,78
10	1,81	2,23	3,17	4,59
11	1,80	2,20	3,11	4,44
12	1,78	2,18	3,05	4,32
13	1,77	2,16	3,01	4,22
14	1,76	2,14	2,98	4,14
15	1,75	2,13	2,95	4,07
16	1,75	2,12	2,92	4,01
17	1,74	2,11	2,90	3,96
18	1,73	2,10	2,88	3,92
19	1,73	2,09	2,86	3,88
20	1,73	2,09	2,85	3,85
21	1,72	2,08	2,83	3,82
22	1,72	2,07	2,82	3,79
23	1,71	2,07	2,81	3,77
24	1,71	2,06	2,80	3,74
25	1,71	2,06	2,79	3,72
26	1,71	2,06	2,78	3,71
27	1,71	2,05	2,77	3,69
28	1,70	2,05	2,76	3,66
29	1,70	2,05	2,76	3,66
30	1,70	2,04	2,75	3,65
40	1,68	2,02	2,70	3,55
60	1,67	2,00	2,66	3,46
120	1,66	1,98	2,62	3,37
∞	1,64	1,96	2,58	3,29

Рис. 10.6 Таблиця критичних значень розподілу Стюдента

2. Числові характеристики $M(X)$, $D(X)$, $\sigma(X)$ випадкової величини, що має розподіл Стюдента, обчислюються за формулами (10.13):

$$M(X) = 0; D(X) = \frac{k}{k-2}, k > 2; \sigma(X) = \sqrt{\frac{k}{k-2}}, k > 2. \quad (10.13)$$

3. Мода та медіана випадкової величини X , що має розподіл Стюдента, співпадають з її математичним сподіванням та дорівнюють: $Mo(X) = Me(X) = 0$.

Доведення. Правомірність цього факту легко можна побачити на рис. 10.4.

4. Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл Стюдента, дорівнюють (10.14):

$$As(X) = 0, k > 3, Es(X) = \frac{6}{k-4}, k > 4. \quad (10.14)$$

Доведення. Обчислимо центральні моменти 3-го та 4-го порядків.

$$\mu_3(X) = 0; \mu_4(X) = \frac{3 \cdot k^2}{(k-2) \cdot (k-4)}.$$

$$As(X) = \frac{\mu_3(X)}{\sigma^3(X)} = \frac{0}{\left(\sqrt{\frac{k}{k-2}}\right)^3} = 0.$$

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{\frac{3 \cdot k^2}{(k-2) \cdot (k-4)}}{\left(\sqrt{\frac{k}{k-2}}\right)^4} - 3 = \frac{3 \cdot (k-2)}{k-4} - 3 = \frac{6}{k-4}.$$

Дослідження коефіцієнту асиметрії та ексцесу показали, що загалом функція щільності розподілу Стюдента схожа на дзвоноподібну функцію щільності нормального розподілу, з тією відмінністю, що у розподілі Стюдента при малих степенях свободи вона трохи нижча і ширша (рис. 10.4).

5. При степенях свободи $k > 30$ розподіл Стюдента наближається до стандартного нормального розподілу $N(0; 1)$.

Доведення. З попередньої властивості (формула (10.14)) видно, що $As(X) = 0$, окрім того при $k \rightarrow +\infty$; $Es(X) = 0$. Це свідчить про те, що розподіл Стюдента за рівнем асиметрії та «піковості» або «випуклості» дійсно стає подібним до нормального розподілу.

Визначення степеня свободи (k) для розподілу Стюдента

Розподіл Стюдента залежить від кількості ступенів свободи k , які зазвичай обчислюються за формулою (10.1)

$$k = n - 1,$$

де n — це кількість спостережень у вибірці.

Цей ступінь свободи використовується для оцінки розподілу Стюдента, який підходить для ситуацій, коли розмір вибірки невеликий і стандартне відхилення генеральної сукупності невідоме.

Використання розподілу Стюдента

t -тест для однієї вибірки: Використовується для перевірки гіпотези про те, що середнє значення вибірки дорівнює певному значенню, коли стандартне відхилення генеральної сукупності невідоме і оцінюється за допомогою вибіркового стандартного відхилення.

t -тест для двох незалежних вибірок: Використовується для порівняння середніх значень двох незалежних вибірок, коли стандартні відхилення генеральних сукупностей невідомі.

Парний t -тест: Використовується для порівняння середніх значень двох залежних вибірок (наприклад, до і після експерименту) з невідомим стандартним відхиленням.

10.4. Розподіл Фішера-Снедекора

Розподілом Фішера-Снедекора (F-розподілом) називається розподіл випадкової величини (10.15):

$$F = \frac{\frac{x_1^2}{k_1}}{\frac{x_2^2}{k_2}}; \quad (10.15)$$

де χ_1^2 та χ_2^2 – випадкові величини, що мають розподіл χ^2 з степенями свободи k_1 та k_2 відповідно.

Щільність ймовірності розподілу Фішера-Снедекора має вигляд (10.16):

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{k_1+k_2}{2}\right) \cdot k_1^{\frac{k_1}{2}} \cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right) \cdot \Gamma\left(\frac{k_2}{2}\right)} \cdot x^{\frac{k_1}{2}-1} \cdot (k_1 \cdot x + k_2)^{-\left(\frac{k_1+k_2}{2}\right)}, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.16)$$

В формулі (10.16) $\Gamma(y) = \int_0^{+\infty} (e^{-t} \cdot t^{y-1}) dt$ – гамма-функція Ейлера.

Графік функції щільності розподілу Фішера-Снедекора наведено на рис. 10.7.

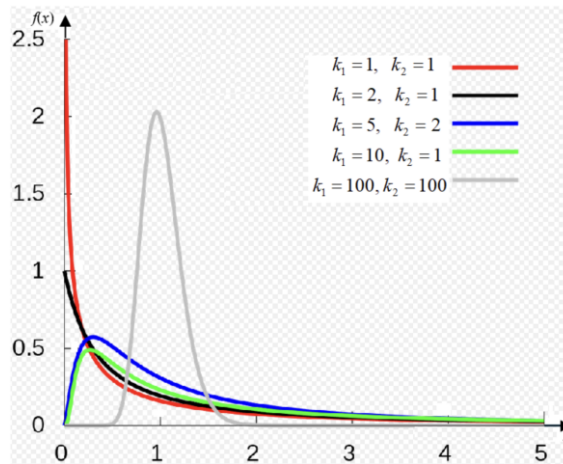


Рис. 10.7. Графік функції щільності $f(x)$ розподілу Фішера-Снедекора

Властивості розподілу Фішера-Снедекора

1. Функція розподілу $F(x)$ матиме вигляд (10.17):

$$F(x) = \begin{cases} \frac{\Gamma\left(\frac{k_1+k_2}{2}\right) \cdot k_1^{\frac{k_1}{2}} \cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right) \cdot \Gamma\left(\frac{k_2}{2}\right)} \cdot \int_{-\infty}^x x^{\frac{k_1}{2}-1} \cdot (k_1 \cdot x + k_2)^{-\left(\frac{k_1+k_2}{2}\right)} dx, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.17)$$

Доведення. Перевіримо відповідність функції $F(x)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу Фішера-Снедекора завжди монотонно зростає.

Графік функції розподілу для розподілу Фішера-Снедекора наведено на рис. 10.8.

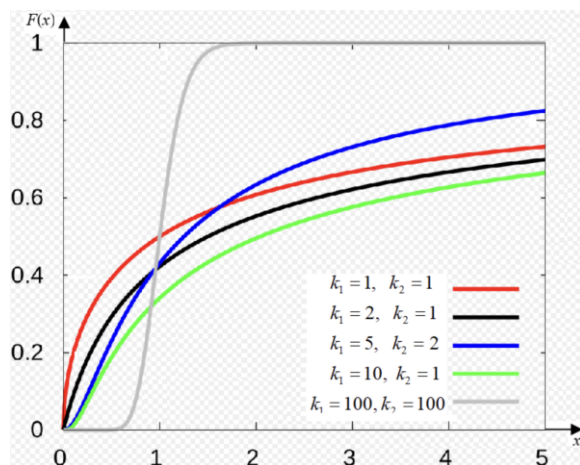


Рис. 10.8. Графік функції розподілу $F(x)$ розподілу Фішера-Снедекора

На рис. 10.9 наведено таблицю критичних значень розподілу Фішера-Снедекора для степенів свободи k_1, k_2 та рівня значущості α .

K_2	Число степенів свободи варіації для більшої дисперсії (K_1)									
	1	3	5	7	9	11	14	20	30	50
1	161	216	230	237	241	243	245	248	250	252
	4052	5403	5764	5928	6022	6082	6142	6208	6258	6302
3	10.3	9.28	9.01	8.88	8.81	8.76	8.71	8.66	8.62	8.58
	34.12	29.46	28.24	27.67	27.34	27.13	26.92	26.69	26.50	26.35
5	6.61	5.41	5.05	4.88	4.78	4.70	4.64	4.56	4.50	4.44
	16.26	12.06	10.97	10.45	10.15	9.96	9.77	9.55	9.38	9.24
7	5.59	4.35	3.97	3.79	3.68	3.60	3.52	3.44	3.38	3.32
	12.25	8.45	7.46	7.00	6.71	6.54	6.35	6.15	5.98	5.85
9	5.12	3.86	3.48	3.29	3.18	3.10	3.02	2.93	2.86	2.80
	10.56	6.99	6.06	5.62	5.35	5.18	5.00	4.80	4.64	4.51
11	4.84	3.59	3.20	3.01	2.90	2.82	2.74	2.65	2.57	2.50
	9.05	6.22	5.32	4.88	4.63	4.46	4.29	4.10	3.94	3.80
13	4.67	3.41	3.02	2.84	2.72	2.63	2.55	2.46	2.38	2.32
	9.07	5.74	4.86	4.44	4.19	4.02	3.85	3.67	3.51	3.37
15	4.54	3.29	2.90	2.70	2.59	2.51	2.43	2.33	2.25	2.18
	8.68	5.42	4.56	4.14	3.89	3.73	3.56	3.36	3.20	3.07
17	4.45	3.20	2.81	2.62	2.50	2.41	2.33	2.23	2.15	2.08
	8.40	5.18	4.34	3.93	3.68	3.52	3.35	3.16	3.00	2.89
19	4.38	3.13	2.74	2.55	2.43	2.34	2.26	2.15	2.07	2.00
	8.18	5.01	4.17	3.77	3.52	3.36	3.19	3.00	2.84	2.70
21	4.32	3.07	2.68	2.49	2.37	2.28	2.20	2.09	2.00	1.93
	8.02	4.87	4.04	3.65	3.40	3.24	3.07	2.88	2.72	2.58
23	4.28	3.03	2.64	2.45	2.32	2.24	2.14	2.04	1.96	1.88
	7.88	4.46	3.94	3.54	3.30	3.14	2.97	2.78	2.62	2.48
25	4.24	2.99	2.60	2.41	2.28	2.20	2.11	2.00	1.92	1.84
	7.77	4.68	3.86	3.46	3.21	3.05	2.89	2.70	2.54	2.40
27	4.21	2.96	2.57	2.37	2.25	2.16	2.08	1.97	1.88	1.80
	7.68	4.60	3.79	3.39	3.14	2.98	2.83	2.63	2.45	2.33
29	4.18	2.93	2.54	2.35	2.22	2.14	2.05	1.94	1.85	1.77
	7.60	4.54	3.73	3.33	3.08	2.92	2.77	2.57	2.41	2.27

Рис. 10.9 Таблиця критичних значень розподілу Фішера-Снедекора $\alpha = 0,95$ та $\alpha = 0,99$ (верхній та нижній рядок відповідно)

2. Числові характеристики $M(X), D(X), \sigma(X)$ випадкової величини, що має розподіл Фішера-Снедекора, обчислюються за формулами (10.18):

$$M(X) = \frac{k_2}{k_2 - 2} \quad (k_2 > 2); \quad D(X) = \frac{2 \cdot k_2^2 \cdot (k_1 + k_2 - 2)}{k_1 \cdot (k_2 - 2)^2 \cdot (k_2 - 4)} \quad (k_2 > 4);$$

$$\sigma(X) = \frac{k_2}{k_2-2} \cdot \sqrt{\frac{2 \cdot (k_1+k_2-2)}{k_1 \cdot (k_2-4)}} \quad (k_2 > 4). \quad (10.18)$$

3. Мода випадкової величини X , що має розподіл Фішера-Снедекора, дорівнює (10.19):

$$Mo(X) = \frac{k_1}{k_1-2} \cdot \frac{k_2+2}{k_2}. \quad (10.19)$$

4. Коефіцієнт асиметрії випадкової величини X , що має розподіл Фішера-Снедекора, дорівнює (10.20)

$$As(X) = \frac{2 \cdot (2 \cdot k_1 + k_2 - 2)}{k_2 - 6} \cdot \sqrt{\frac{2 \cdot (k_2 - 4)}{k_1 \cdot (k_1 + k_2 - 2)}} \quad (k_2 > 6). \quad (10.20)$$

Тобто, розподіл Фішера-Снедекора має додатну (правосторонню) асиметрію.

Доведення. Для розподілу Фішера-Снедекора (формула (10.20)) очевидно, що $As(X) > 0$, тобто «довша» частина кривої щільності знаходиться правіше моди, порівняно з графіком щільності нормального розподілу (де коефіцієнт асиметрії дорівнює 0).

5. При $n \rightarrow +\infty$ розподіл Фішера-Снедекора наближається до нормального розподілу.

Визначення степенів свободи (k_1, k_2) для розподілу Фішера-Снедекора

k_1, k_2 — це параметри, що визначають форму розподілу.

Степінь свободи k_1 зазвичай обчислюється за формулою (10.1)

$$k_1 = n - 1,$$

а k_2 — за формулою (10.21)

$$k_2 = n - s, \quad (10.21)$$

де:

n — загальна кількість спостережень;

s — кількість груп.

Використання розподілу Фішера-Снедекора

Аналіз варіацій (ANOVA): Розподіл Фішера-Снедекора використовується для перевірки гіпотез про рівність середніх у кількох групах. Він визначає, чи є значуща різниця між груповими середніми.

Регресійний аналіз: Розподіл Фішера-Снедекора застосовується для перевірки значущості регресійної моделі, зокрема для визначення, чи є включені предиктори (незалежні змінні) суттєвими для прогнозування залежної змінної.

Перевірка гіпотез: У тестах на рівність дисперсій (наприклад, тест Левена) використовується розподіл Фішера-Снедекора для перевірки, чи є дисперсії двох вибірок рівними.

10.5. Розподіл Парето

Розподіл Парето в теорії ймовірностей — двопараметрична сім'я абсолютно неперервних розподілів. Названий на честь італійського інженера з цивільного будівництва, економіста, і соціолога Вільфредо Парето.

Розподіл Парето — це степеневий розподіл ймовірностей, який використовується для описання соціальних, наукових, геофізичних, актуарних, та багатьох інших типів спостережуваних явищ.

Початково застосовувалася для описання розподілу багатства серед суспільства, що відповідає тенденції, що велика частина багатства зосереджена в руках невеликої частини населення людей. У розмовній версії розподіл Парето відомий як принцип Парето, або «правило 80-20», а також іноді може називатися «ефектом Матвія». Це правило стверджує що, наприклад, 80% багатства суспільства утримують 20% його населення.

Щільність ймовірності розподілу Парето має вигляд (10.22):

$$f(x) = \begin{cases} \frac{k}{x_0} \cdot \left(\frac{x_0}{x}\right)^{k+1}, & \text{де } x \geq x_0; \\ 0, & \text{де } x < x_0, \end{cases} \quad (10.22)$$

де:

x_0 — параметр масштабування, який характеризує розподіл Парето;

k — значення степеню, який є параметром форми для даного розподілу.

У теорії ймовірностей і статистиці **параметр форми** є числовим параметром із сімейства параметричних розподілів ймовірностей, який не є ні коефіцієнтом зсуву ні коефіцієнтом масштабу (ані їх функцією як, наприклад, параметр швидкості). Такий параметр повинен впливати на форму розподілу, а не просто зміщувати його (як це робить параметр розташування) або розтягувати/згортати (як це робить параметр масштабу).

Графік функції щільності розподілу Парето наведено на рис. 10.10.

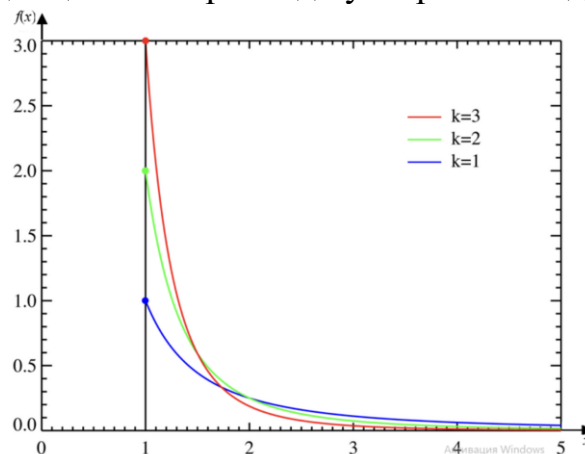


Рис. 10.10. Графік функції щільності $f(x)$ розподілу Парето при $x_0 = 1$

Властивості розподілу Парето

1. Функція розподілу $F(x)$ матиме вигляд (10.23):

$$F(x) = \begin{cases} 1 - \left(\frac{x_0}{x}\right)^k, & \text{де } x \geq x_0; \\ 0, & \text{де } x < x_0. \end{cases} \quad (10.23)$$

Доведення. Перевіримо відповідність функції $F(x)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу Парето завжди монотонно зростає.

Графік функції розподілу для розподілу Парето наведено на рис. 10.11.

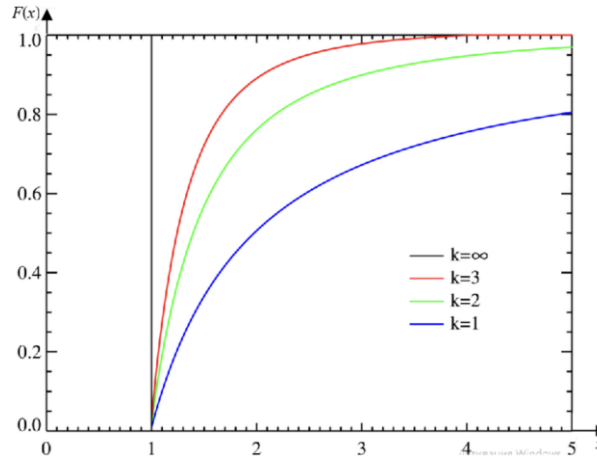


Рис. 10.11. Графік функції розподілу $F(x)$ розподілу Парето

2. Числові характеристики $M(X)$, $D(X)$, $\sigma(X)$ випадкової величини, що має розподіл Парето, обчислюються за формулами (10.24):

$$M(X) = \frac{k \cdot x_0}{k-1} \quad (k > 1); \quad D(X) = \frac{k \cdot x_0^2}{(k-1)^2 \cdot (k-2)} \quad (k > 2);$$

$$\sigma(X) = \frac{x_0}{k-1} \cdot \sqrt{\frac{k}{k-2}} \quad (k > 2). \quad (10.24)$$

Доведення. Обчислимо математичне сподівання:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} (x \cdot f(x)) dx = \int_{x_0}^{+\infty} \left(x \cdot \frac{k}{x_0} \cdot \left(\frac{x_0}{x}\right)^{k+1} \right) dx = \\ &= k \cdot \int_{x_0}^{+\infty} \left(\frac{x_0}{x}\right)^k dx = \lim_{b \rightarrow +\infty} \left(-\frac{k \cdot x_0}{k-1} \cdot \left(\frac{x_0}{x}\right)^{k-1} \right) \Big|_{x_0}^b = \\ &= \lim_{b \rightarrow +\infty} \left(-\frac{k \cdot x_0}{k-1} \cdot \left(\frac{x_0}{b}\right)^{k-1} + \frac{k \cdot x_0}{k-1} \right) = \frac{k \cdot x_0}{k-1}. \end{aligned}$$

Обчислимо дисперсію:

$$\begin{aligned}
M(X^2) &= \int_{-\infty}^{+\infty} (x^2 \cdot f(x)) dx = \int_{x_0}^{+\infty} \left(x^2 \cdot \frac{k}{x_0} \cdot \left(\frac{x_0}{x} \right)^{k+1} \right) dx = \\
&= \int_{x_0}^{+\infty} \left(\frac{x_0}{x} \right)^{k-1} dx = \lim_{b \rightarrow +\infty} \left(-\frac{k \cdot x_0^2}{k-2} \cdot \left(\frac{x_0}{x} \right)^{k-2} \right) \Big|_{x_0}^b = \\
&= \lim_{b \rightarrow +\infty} \left(-\frac{k \cdot x_0^2}{k-2} \cdot \left(\frac{x_0}{b} \right)^{k-2} + \frac{k \cdot x_0^2}{k-2} \right) = \frac{k \cdot x_0^2}{k-2}. \\
D(X) &= M(X^2) - (M(X))^2 = \frac{k \cdot x_0^2}{k-2} - \left(\frac{k \cdot x_0}{k-1} \right)^2 = \\
&= \frac{k \cdot x_0^2 \cdot (k^2 - 2 \cdot k + 1 - k^2 + 2 \cdot k)}{(k-1)^2 \cdot (k-2)} = \frac{k \cdot x_0^2}{(k-1)^2 \cdot (k-2)}.
\end{aligned}$$

Обчислимо середньоквадратичне відхилення:

$$\sigma(X) = \sqrt{D(X)} = \frac{x_0}{k-1} \cdot \sqrt{\frac{k}{k-2}}.$$

3. *Мода та медіана випадкової величини X , що має розподіл Парето, дорівнюють (10.25):*

$$Mo(X) = x_0; Me(X) = x_0 \cdot \sqrt[k]{2}. \quad (10.25)$$

Доведення. За формулою (10.22) $f(x) = \frac{k}{x_0} \cdot \left(\frac{x_0}{x} \right)^{k+1}$, де $x \geq x_0$.

Оскільки $\left(\frac{x_0}{x} \right)^{k+1} \rightarrow 0$ при $x \rightarrow +\infty$, то $f_{max} = \frac{k}{x_0}$, яке досягається при $x = x_0$. Отже $Mo(X) = x_0$.

Обчислимо медіану:

$$\begin{aligned}
F(Me(X)) &= 0,5 \\
1 - \left(\frac{x_0}{Me(X)} \right)^k &= 0,5; \\
\left(\frac{x_0}{Me(X)} \right)^k &= 0,5; \\
\frac{x_0}{Me(X)} &= \sqrt[k]{0,5}; \\
Me(X) &= \frac{x_0}{\sqrt[k]{0,5}} = x_0 \cdot \sqrt[k]{2}.
\end{aligned}$$

4. *Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл Парето, дорівнюють (10.26)*

$$As(X) = \frac{2 \cdot (k+1)}{k-3} \cdot \sqrt{\frac{k-2}{k}} \quad (k > 3); \quad Es(X) = \frac{6 \cdot (k^3 + k^2 - 6 \cdot k - 2)}{k \cdot (k-3) \cdot (k-4)} \quad (k > 4). \quad (10.26)$$

Тобто, розподіл Парето має додатну (правосторонню) асиметрію та має більшу «піковість» або «виуклість» в порівнянні з нормальним розподілом.

Доведення. Для розподілу Парето (формули (10.26)) очевидно, що $As(X) > 0$, тобто «довша» частина кривої знаходиться правіше моди, порівняно з графіком щільності нормального розподілу (де коефіцієнт асиметрії дорівнює 0).

Для розподілу Парето (формули (10.26)) очевидно, що $Es(X) > 0$, тобто графік функції щільності має більшу «піковість» або «випуклість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

5. Характеристична теорема. Припустимо, що $X_1, X_2, X_3, \dots$ є незалежні однаково розподілені випадкові величини, розподіл ймовірностей яких знаходиться в інтервалі $[x_m, +\infty)$ для деякого значення $x_m > 0$. Припустимо, що для всіх n , пара випадкових величин $\min\{X_1, \dots, X_n\}$ і $\frac{X_1 + \dots + X_n}{\min\{X_1, \dots, X_n\}}$ є незалежними. Тоді їх спільний розподіл буде розподілом Парето.

Використання розподілу Парето

Наступні приклади іноді розглядають як такі, що приблизно мають розподіл Парето:

- Розмір населених пунктів (небагато міст, багато селищ/сіл);
- Розподіл розмірів файлів в Інтернет-трафіку в якому використовується протокол ТСП (багато менших файлів, рідше великі);
- Частота помилок запису на жорсткому диску;
- Кластери конденсації Бозе-Ейнштейна близько абсолютного нуля;
- Величина запасів нафти в нафтових родовищах (не багато великих родовищ, і багато малих родовищ);
- Обсяг задач, які виносилися для вирішення на суперкомп'ютерах (декілька великих, багато малих);
- Нормалізована дохідність цін на окремі акції;
- Розміри частинок піску;
- Розмір метеоритів;
- Величина значних втрат унаслідок катастроф для певного роду бізнесу, генеральні зобов'язання, комерційні авто, і компенсація робітникам;
- В Гідрології розподіл Парето застосовується для моделювання надзвичайних подій таких як щорічні максимальні опади на добу і паводки у рік.

10.6. Розподіл Ерланга

Розподіл Ерланга — двопараметричне сімейство абсолютно неперервних розподілів, визначених для $x \in [0, +\infty)$. Параметрами розподілу є: $k \in N$ — параметр форми, $\lambda \in (0, +\infty)$ — коефіцієнт норми. Іноді використовується обернений параметр $\beta = \frac{1}{\lambda}$ — коефіцієнт масштабу.

Розподіл Ерланга — це розподіл суми k незалежних однаково експоненційно розподілених випадкових величин з параметром $\frac{1}{\lambda}$. Еквівалентним твердженням є те, що це розподіл часу до k -тої події пуассонівського процесу з параметром λ .

Розподіли Ерланга та Пуассона є взаємнодоповнюючими: розподіл Пуассона підраховує кількість подій, що відбудуться за фіксований проміжок часу, а розподіл Ерланга підраховує кількість часу до появи фіксованої кількості подій. При $k = 1$, розподіл Ерланга збігається з експоненційним розподілом. Розподіл Ерланга — окремий випадок гамма-розподілу з натуральними значеннями параметру форми k .

Щільність ймовірності розподілу Ерланга має вигляд (10.27):

$$f(x) = \frac{\lambda^k \cdot x^{k-1} \cdot e^{-(\lambda \cdot x)}}{(k-1)!}. \quad (10.27)$$

Графік функції щільності розподілу Ерланга наведено на рис. 10.12.

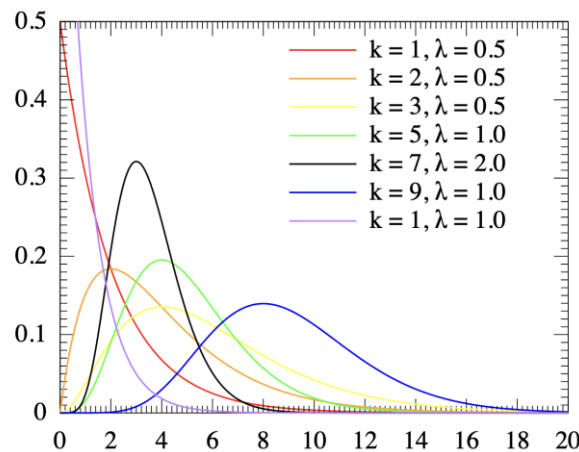


Рис. 10.12. Графік функції щільності $f(x)$ розподілу Ерланга

Властивості розподілу Ерланга

1. Функція розподілу $F(x)$ матиме вигляд (10.28):

$$F(x) = 1 - \sum_{s=1}^{k-1} \frac{1}{s!} \cdot e^{-(\lambda \cdot x)} \cdot (\lambda \cdot x)^s \quad (10.28)$$

Доведення. Перевіримо відповідність функції $F(x)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу Ерланга завжди монотонно зростає.

Графік функції розподілу для розподілу Ерланга наведено на рис. 10.13.

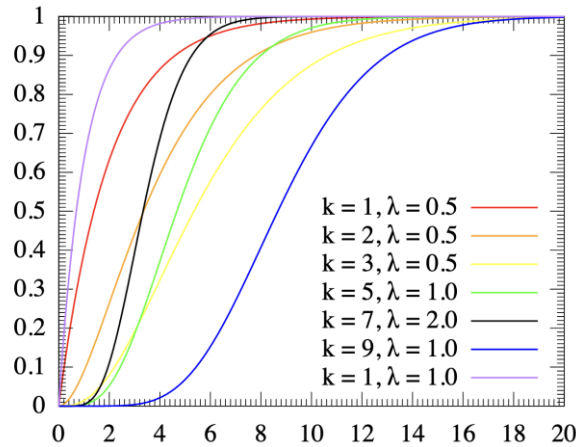


Рис. 10.13. Графік функції розподілу $F(x)$ розподілу Ерланга

2. Числові характеристики $M(X)$, $D(X)$, $\sigma(X)$ випадкової величини, що має розподіл Релея, обчислюються за формулами (10.29):

$$M(X) = \frac{k}{\lambda}; D(X) = \frac{k}{\lambda^2}; \sigma(X) = \frac{\sqrt{k}}{\lambda}. \quad (10.29)$$

Доведення. Обчислимо математичне сподівання:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} (x \cdot f(x)) dx = \int_0^{+\infty} \left(x \cdot \frac{\lambda^k \cdot x^{k-1} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} \right) dx = \\ &= \frac{\lambda^k}{(k-1)!} \cdot \int_0^{+\infty} (x^k \cdot e^{-(\lambda \cdot x)}) dx = \frac{\lambda^k}{(k-1)!} \cdot \frac{\Gamma(k+1)}{\lambda^{k+1}} = \frac{\lambda^k}{(k-1)!} \cdot \frac{k!}{\lambda^{k+1}} = \\ &= \frac{k}{\lambda}. \end{aligned}$$

Обчислимо дисперсію:

$$\begin{aligned} M(X^2) &= \int_{-\infty}^{+\infty} (x^2 \cdot f(x)) dx = \int_0^{+\infty} \left(x^2 \cdot \frac{\lambda^k \cdot x^{k-1} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} \right) dx = \\ &= \frac{\lambda^k}{(k-1)!} \cdot \int_0^{+\infty} (x^{k+1} \cdot e^{-(\lambda \cdot x)}) dx = \frac{\lambda^k}{(k-1)!} \cdot \frac{\Gamma(k+2)}{\lambda^{k+2}} \\ &= \frac{\lambda^k}{(k-1)!} \cdot \frac{(k+1)!}{\lambda^{k+2}} = \frac{k \cdot (k+1)}{\lambda^2}. \\ D(X) &= M(X^2) - (M(X))^2 = \frac{k \cdot (k+1)}{\lambda^2} - \left(\frac{k}{\lambda} \right)^2 = \frac{k \cdot (k+1 - k)}{\lambda^2} = \frac{k}{\lambda^2}. \end{aligned}$$

Обчислимо середньоквадратичне відхилення:

$$\sigma(X) = \sqrt{D(X)} = \frac{\sqrt{k}}{\lambda}.$$

3. Мода випадкової величини X , що має розподіл Ерланга, дорівнюють (10.30):

$$Mo(X) = \frac{k-1}{\lambda}. \quad (10.30)$$

Доведення. За формулою (10.27) $f(x) = \frac{\lambda^k \cdot x^{k-1} \cdot e^{-(\lambda \cdot x)}}{(k-1)!}$.

$$\begin{aligned} f'(x) &= \frac{\lambda^k \cdot (k-1) \cdot x^{k-2} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} - \frac{\lambda^{k+1} \cdot x^{k-1} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} = \\ &= \frac{\lambda^k \cdot x^{k-2} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} \cdot (k-1 - \lambda \cdot x). \end{aligned}$$

Стаціонарна точка:

$$\begin{aligned} k-1 - \lambda \cdot x &= 0; \\ x &= \frac{k-1}{\lambda}. \end{aligned}$$

Обчислимо $f''(x)$:

$$\begin{aligned} f''(x) &= \frac{\lambda^k \cdot (k-1) \cdot (k-2) \cdot x^{k-3} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} - \frac{\lambda^{k+1} \cdot (k-1) \cdot x^{k-2} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} - \\ &- \frac{\lambda^{(k+1)} \cdot (k-1) \cdot x^{k-2} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} + \frac{\lambda^{k+2} \cdot x^{(k-1)} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} = \\ &= \frac{\lambda^k \cdot x^{(k-3)} \cdot e^{-(\lambda \cdot x)}}{(k-1)!} \cdot ((k-1) \cdot (k-2) - 2 \cdot \lambda \cdot (k-1) \cdot x + \lambda^2 \cdot x^2). \\ f''\left(\frac{k-1}{\lambda}\right) &= \frac{\lambda^k \cdot \left(\frac{k-1}{\lambda}\right)^{k-3} \cdot e^{-(k-1)}}{(k-1)!} \cdot ((k-1) \cdot (k-2) - (k-1)^2) = \\ &= -\frac{\lambda^3 \cdot (k-1)^{k-3} \cdot e^{-(k-1)}}{(k-2)!} < 0. \end{aligned}$$

Отже точка $x = \frac{k-1}{\lambda}$ є точкою максимуму, тобто $Mo(X) = \frac{k-1}{\lambda}$.

4. Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл Ерланга, дорівнюють (10.31)

$$As(X) = \frac{2}{\sqrt{k}}; \quad Es(X) = \frac{6}{k}. \quad (10.31)$$

Тобто, розподіл Ерланга має додатну (правосторонню) асиметрію та має більшу «піковість» в порівнянні з нормальним розподілом.

Доведення. Обчислимо центральні моменти 3-го та 4-го порядків.

$$\begin{aligned} \mu_3(X) &= \frac{2 \cdot k}{\lambda^3}; \quad \mu_4(X) = \frac{3 \cdot k \cdot (k+2)}{\lambda^4}. \\ As(X) &= \frac{\mu_3(X)}{\sigma^3(X)} = \frac{\frac{2 \cdot k}{\lambda^3}}{\left(\frac{\sqrt{k}}{\lambda}\right)^3} = \frac{2}{\sqrt{k}}. \end{aligned}$$

$$Es(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = \frac{\frac{3 \cdot k \cdot (k+2)}{\lambda^4}}{\left(\frac{\sqrt{k}}{\lambda}\right)^4} - 3 = \frac{3 \cdot k \cdot (k+2)}{k^2} - 3 = \frac{6}{k}.$$

Для розподілу Ерланга (формули (10.31)) очевидно, що $As(X) > 0$, тобто «довша» частина кривої знаходиться правіше моди, порівняно з графіком щільності нормального розподілу (де коефіцієнт асиметрії дорівнює 0).

Для розподілу Ерланга (формули (10.31)) очевидно, що $Es(X) > 0$, тобто графік функції щільності має більшу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Використання розподілу Ерланга

Теорія черг: В моделюванні систем, де події або обслуговування розділяються на кілька етапів, розподіл Ерланга використовується для оцінки часу обслуговування.

Телекомунікації: Використовується для аналізу навантаження на телефонні мережі та обробки викликів.

Надійність систем: Описує час до відмови системи, яка складається з кількох компонентів.

10.8. Розподіл Релея

Розподіл Релея – це розподіл безперервної випадкової величини з використанням експоненти $e^{-(k \cdot x^2)}$ у функціях розподілу.

Щільність ймовірності розподілу Релея має вигляд (10.32):

$$f(x) = \begin{cases} \frac{x}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2 \cdot \sigma^2}\right)}, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.32)$$

Графік функції щільності розподілу Релея наведено на рис. 10.13.

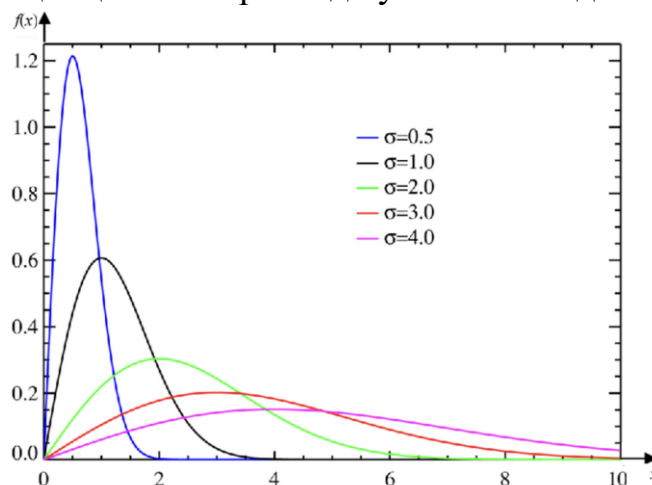


Рис. 10.13. Графік функції щільності $f(x)$ розподілу Релея

Властивості розподілу Релея

1. Функція розподілу $F(x)$ матиме вигляд (10.33):

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x^2}{2\sigma^2}\right)}, & \text{де } x \geq 0; \\ 0, & \text{де } x < 0. \end{cases} \quad (10.33)$$

Доведення. Перевіримо відповідність функції $F(x)$ основним властивостям функції розподілу.

Зростання: Функція розподілу ймовірностей $F(x)$ є зростаючою функцією, тобто з збільшенням x ймовірність також збільшується.

Граничні значення: При $x = 0$: $F(0) = 0$.

При $x \rightarrow +\infty$; $F(x) \rightarrow 1$.

Монотонність: Функція $F(x)$ для розподілу Релея завжди монотонно зростає.

Графік функції розподілу для розподілу Релея наведено на рис. 10.14.

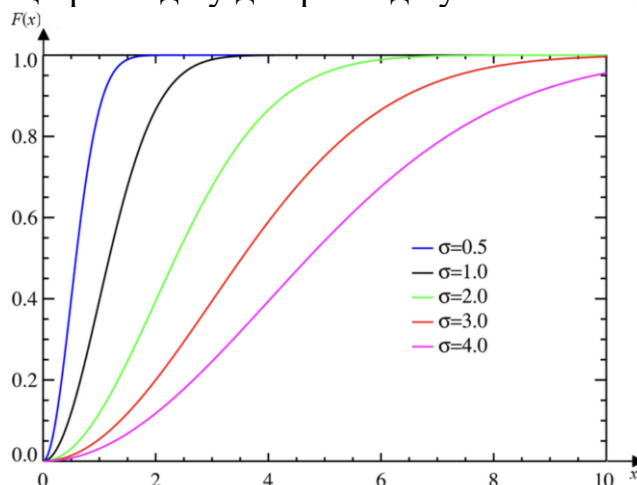


Рис. 10.14. Графік функції розподілу $F(x)$ розподілу Релея

2. Числові характеристики $M(X)$, $D(X)$, $\sigma(X)$ випадкової величини, що має розподіл Релея, обчислюються за формулами (10.34):

$$M(X) = \frac{\sqrt{2\pi}}{2} \cdot \sigma; D(X) = \frac{4-\pi}{2} \cdot \sigma^2; \sigma(X) = \sqrt{\frac{4-\pi}{2}} \cdot \sigma. \quad (10.34)$$

Доведення. Обчислимо математичне сподівання:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} (x \cdot f(x)) dx = \int_0^{+\infty} \left(x \cdot \frac{x}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} \right) dx = \\ &= \int_0^{+\infty} \left(\frac{x^2}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} \right) dx = \left[u = \frac{x}{\sigma}; x = u \cdot \sigma \right] = \sigma \cdot \int_0^{+\infty} (u^2 \cdot e^{-u^2}) du \\ &= (\sigma \cdot \sqrt{2\pi}) \cdot \int_0^{+\infty} \left(u^2 \cdot \left(\frac{1}{\sqrt{2\pi}} \cdot e^{-u^2} \right) \right) du = \frac{\sqrt{2\pi}}{2} \cdot \sigma. \end{aligned}$$

Обчислимо дисперсію:

$$\begin{aligned}
M(X^2) &= \int_{-\infty}^{+\infty} (x^2 \cdot f(x)) dx = \int_{x_0}^{+\infty} \left(x^2 \cdot \frac{x}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} \right) dx = \\
&= \int_0^{+\infty} \left(\frac{x^3}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} \right) dx = \left[y = \frac{x^2}{2\sigma^2}; x^2 = 2 \cdot y \cdot \sigma^2 \right] = \\
&\quad x dx = \sigma^2 dy \\
&= \int_0^{+\infty} \left(\frac{2 \cdot y \cdot \sigma^2}{\sigma^2} \cdot e^{-y} \cdot \sigma^2 \right) dy = 2 \cdot \sigma^2 \cdot \int_0^{+\infty} (y \cdot e^{-y}) dy = \\
&= \left[\begin{array}{l} u = y; du = dy \\ dv = e^{-y} dy; v = -e^{-y} \end{array} \right] \\
&= 2 \cdot \sigma^2 \cdot \lim_{b \rightarrow +\infty} \left(-(y \cdot e^{-y}) \Big|_0^b - e^{-y} \Big|_0^b \right) = 2 \cdot \sigma^2. \\
D(X) &= M(X^2) - (M(X))^2 = 2 \cdot \sigma^2 - \left(\frac{\sqrt{2 \cdot \pi}}{2} \cdot \sigma \right)^2 = 2 \cdot \sigma^2 - \frac{\pi}{2} \cdot \sigma^2 = \\
&= \frac{4 - \pi}{2} \cdot \sigma^2.
\end{aligned}$$

Обчислимо середньоквадратичне відхилення:

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\frac{4 - \pi}{2}} \cdot \sigma.$$

3. *Мода та медіана випадкової величини X , що має розподіл Релея, дорівнюють (10.35):*

$$Mo(X) = \sigma; Me(X) = \sigma \cdot \sqrt{\ln(4)}. \quad (10.35)$$

Доведення. За формулою (10.27) $f(x) = \frac{x}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)}$.

$$f'(x) = \frac{1}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} - \left(\frac{x}{\sigma^2}\right)^2 \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} = \frac{\sigma^2 - x^2}{\sigma^4} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)}.$$

Стационарна точка:

$$\begin{aligned}
\sigma^2 - x^2 &= 0; \\
x &= \sigma.
\end{aligned}$$

Обчислимо $f''(x)$:

$$\begin{aligned}
f''(x) &= -\frac{2 \cdot x}{\sigma^4} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)} - \frac{\sigma^2 - x^2}{\sigma^4} \cdot \frac{x}{\sigma^2} \cdot e^{-\left(\frac{x^2}{2\sigma^2}\right)}. \\
f''(\sigma) &= -\frac{2 \cdot \sigma}{\sigma^4} \cdot e^{-\left(\frac{\sigma^2}{2\sigma^2}\right)} - \frac{\sigma^2 - \sigma^2}{\sigma^4} \cdot \frac{\sigma}{\sigma^2} \cdot e^{-\left(\frac{\sigma^2}{2\sigma^2}\right)} = -\frac{2}{\sigma^3} \cdot e^{-0,5} < 0.
\end{aligned}$$

Отже точка $x = \sigma$ є точкою максимуму, тобто $Mo(X) = \sigma$.

Обчислимо медіану:

$$\begin{aligned}
F(Me(X)) &= 0,5 \\
1 - e^{-\left(\frac{(Me(X))^2}{2\sigma^2}\right)} &= 0,5;
\end{aligned}$$

$$e^{-\left(\frac{(Me(X))^2}{2 \cdot \sigma^2}\right)} = 0,5;$$

$$-\frac{(Me(X))^2}{2 \cdot \sigma^2} = -\ln(2);$$

$$Me(X) = \sqrt{2 \cdot \sigma^2 \cdot \ln(2)} = \sqrt{\sigma^2 \cdot \ln(4)} = \sigma \cdot \sqrt{\ln(4)}.$$

4. Коефіцієнт асиметрії та ексцес випадкової величини X , що має розподіл Релея, дорівнюють (10.36):

$$As(X) = 2 \cdot (\pi - 3) \cdot \sqrt{\frac{\pi}{(4-\pi)^3}}; Es(X) = \frac{8-6 \cdot (2-\pi)^2}{(4-\pi)^2}. \quad (10.36)$$

Тобто, розподіл Релея має додатну (правосторонню) асиметрію та має більшу «піковість» в порівнянні з нормальним розподілом.

Доведення. Для розподілу Релея (формули (10.36)) очевидно, що $As(X) > 0$, тобто «довша» частина кривої знаходиться правіше моди, порівняно з графіком щільності нормального розподілу (де коефіцієнт асиметрії дорівнює 0).

Для розподілу Релея (формули (10.36)) очевидно, що $Es(X) > 0$, тобто графік функції щільності має більшу «піковість» в порівнянні з нормальним розподілом (де ексцес дорівнює 0).

Використання розподілу Релея

Розподіл амплітуд сигналів: Розподіл Релея часто використовується для моделювання амплітуд радіосигналів у середовищах з великою кількістю відбиттів, таких як міські райони.

Аналіз вібрацій: Використовується в аналізі випадкових вібрацій та інших механічних систем.

Обробка зображень: Використовується в різних алгоритмах обробки зображень для моделювання розподілу інтенсивності пікселів у шумових умовах.

10.8. Інші розподіли неперервної випадкової величини

На практиці окрім перерахованих розподілів неперервних випадкових величин використовуються наступні:

- Бета-розподіл;
- Розподіл Вейбулла;
- Гамма-розподіл;
- Квадратичний розподіл;
- Розподіл Коші;
- Розподіл Лапласа;
- Лінійний розподіл;
- Логістичний розподіл;
- Розподіл Сосновського;

- Розподіл Кочрена;
- Ель-розподіл.

Приклад 10.1. Випадкова величина X задана своєю щільністю (рис. 10.15):

$$f(x) = \begin{cases} \frac{1}{\beta^{\alpha+1} \cdot \Gamma(\alpha+1)} \cdot x^{\alpha} \cdot e^{-\frac{x}{\beta}}, & \text{де } x \geq 0 \ (\alpha > -1, \beta > 0) \\ 0, & \text{де } x < 0, \end{cases}$$

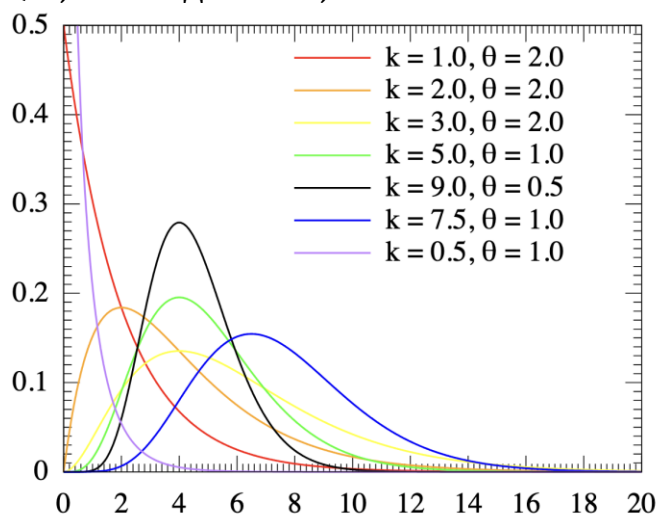


Рис. 10.15. Графік щільності $f(x)$ (Гамма розподіл $k = \alpha + 1$; $\theta = \beta$)
Знайти математичне сподівання, дисперсію та моду даної випадкової величини при $\alpha = 6, \beta = 7$.

Розв'язання.

Формули для гамма-розподілу:

Для гамма-розподілу з параметрами α (параметр форми) та β (параметр масштабу):

Математичне сподівання:

$$M(X) = (\alpha + 1) \cdot \beta.$$

Дисперсія:

$$D(X) = (\alpha + 1) \cdot \beta^2.$$

Мода:

$$Mo(X) = \alpha \cdot \beta, \text{ для } \alpha > 1.$$

Коефіцієнт асиметрії:

$$As(X) = \frac{2}{\sqrt{\alpha + 1}}.$$

Екссес:

$$Es(X) = \frac{6}{\alpha + 1}.$$

Обчислення для $\alpha = 6, \beta = 7$:

Математичне сподівання:

$$M(X) = 7 \cdot 7 = 49.$$

Дисперсія:

$$D(X) = 7 \cdot 7^2 = 7 \cdot 49 = 343.$$

Мода:

$$Mo(X) = 6 \cdot 7 = 42.$$

Коефіцієнт асиметрії:

$$As(X) = \frac{2}{\sqrt{7}} \approx 2,27.$$

Ексцес:

$$Es(X) = \frac{6}{7} \approx 0,86.$$

Відповідь: $M(X) = 49$; $D(X) = 343$; $Mo(X) = 42$; $As(X) \approx 2,27$; $Es(X) \approx 0,86$.

ЛЕКЦІЯ 11

ОПЕРАЦІЇ НАД ВИПАДКОВИМИ ВЕЛИЧИНАМИ. ФУНКЦІЇ ВИПАДКОВИХ АРГУМЕНТІВ

11.1. Арифметичні операції над випадковими величинами

Операції над випадковими величинами відіграють ключову роль у теорії ймовірностей і статистиці, дозволяючи будувати нові випадкові величини та вивчати їх властивості.

Арифметичні операції між випадковими величинами дозволяють створювати нові випадкові величини Z , аналізувати їх властивості та взаємозв'язки. Вони є фундаментальними для моделювання випадкових процесів та явищ у різних галузях науки і техніки.

Основні арифметичні операції над дискретними незалежними випадковими величинами:

1. Додавання: Якщо X і Y — випадкові величини, то сума $Z = X + Y$ також є випадковою величиною, де $z_{ij} = x_i + y_j$, $g_{ij} = p_i \cdot q_j$, де p_i — ймовірність, що випадкова величина X набуде значення x_i ; q_j — ймовірність, що випадкова величина Y набуде значення y_j ; g_{ij} — ймовірність, що випадкова величина X набуде значення x_i та випадкова величина Y набуде значення y_j .

Математичне сподівання:

$$M(Z) = M(X) + M(Y).$$

Дисперсія:

$$D(Z) = D(X) + D(Y).$$

2. Віднімання: Різниця $Z = X - Y$ також є випадковою величиною, де $z_{ij} = x_i - y_j$, $g_{ij} = p_i \cdot q_j$.

Математичне сподівання:

$$M(Z) = M(X) - M(Y).$$

Дисперсія:

$$D(Z) = D(X) + D(Y).$$

3. Множення на константу: Якщо C — константа, то $Z = C \cdot X$ — випадкова величина, де $z_i = C \cdot x_i$, а p_i лишається як і для випадкової величини X .

Математичне сподівання:

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

Дисперсія:

$$D(Z) = C^2 \cdot D(X).$$

4. Множення: Добуток $Z = X \cdot Y$ — випадкова величина, де $z_{ij} = x_i \cdot y_j$, $g_{ij} = p_i \cdot q_j$.

Математичне сподівання:

$$M(Z) = M(X \cdot Y) = M(X) \cdot M(Y).$$

Дисперсія:

$$D(Z) = D(X \cdot Y) = M(X^2) \cdot M(Y^2) - (M(X))^2 \cdot (M(Y))^2.$$

5. Ділення: Частка $Z = \frac{X}{Y}$ — випадкова величина, за умови, що $Y \neq 0$, де $z_{ij} = \frac{x_i}{y_j}$, $g_{ij} = p_i \cdot q_j$.

Формули для математичного сподівання та дисперсії частки випадкових величин досить складні, тому їх використання є недоречним. В даному випадку доцільніше використовувати стандартні формули, наведені в лекції 7.

6. Піднесення до n -го степеню: $Z = X^n$ — випадкова величина, де $z_i = x_i^n$, а p_i лишається як і для випадкової величини X .

Математичне сподівання:

$$M(Z) = M(X^n).$$

Дисперсія:

$$D(Z) = D(X^n).$$

7. Лінійні комбінації випадкових величин: Якщо X і Y — дві незалежні випадкові величини, а a і b — константи, то лінійна комбінація $Z = a \cdot X + b \cdot Y$ є також випадковою величиною.

Математичне сподівання:

$$M(Z) = M(a \cdot X + b \cdot Y) = a \cdot M(X) + b \cdot M(Y).$$

Дисперсія: $D(Z) = D(a \cdot X + b \cdot Y) = a^2 \cdot D(X) + b^2 \cdot D(Y)$.

Зауваження

1. Випадкова величина Z набуває лише тих значень, які є результатами перерахованих операцій;
2. Всі формули для обчислення математичного сподівання та дисперсії випадкової величини Z є наслідком з властивостей даних числових характеристик, які наведено в лекції 7;
3. Функція розподілу для випадкової величини Z будується за правилом, наведеним в лекції 6.

Приклад 11.1. Наведено закони розподілу двох незалежних випадкових величин (табл. 11.1, 11.2):

Таблиця 11.1. Ряд розподілу дискретної випадкової величини X

x_i	1	3	6
p_i	0,5	0,2	0,3

Таблиця 11.2. Ряд розподілу дискретної випадкової величини Y

y_j	-1	0	1
q_j	0,1	0,6	0,3

Знайти закони розподілу випадкових величин: **а)** $Z = X + Y$, **б)** $U = 2 \cdot X$.

Розв'язання. а) Складемо допоміжну таблицю за наступним принципом: $z_{ij} = x_i + y_j, g_{ij} = p_i \cdot q_j$ ($i, j = \overline{1,3}$) (табл. 11.3).

Таблиця 11.3. Розрахункова таблиця для прикладу 11.1

$X \backslash Y$		$y_1 = -1$	$y_2 = 0$	$y_3 = 1$
		$q_1 = 0,1$	$q_2 = 0,6$	$q_3 = 0,3$
$x_1 = 1$	$p_1 = 0,5$	0 0,05	1 0,3	2 0,15
$x_2 = 3$	$p_2 = 0,2$	2 0,02	3 0,12	4 0,06
$x_3 = 6$	$p_3 = 0,3$	5 0,03	6 0,18	7 0,09

В розрахунковій таблиці (табл. 11.3, згори ліворуч) наведено значення z_{ij} , які набуває випадкова величина Z . Причому, якщо її значення z_{ij}, z_{lt} ($i, j, l, t = \overline{1,3}$) співпадають, то їх ймовірності (табл. 11.3, знизу праворуч) додаються згідно теореми про додавання несумісних випадкових величин. В результаті маємо випадкову величину Z , яка набуває значень z_k ($k = \overline{1,8}$) з ймовірностями g_k .

Остаточний закон розподілу (табл. 11.4):

Таблиця 11.4. Ряд розподілу дискретної випадкової величини Z

z_k	0	1	2	3	4	5	6	7
g_k	0,05	0,3	0,17	0,12	0,06	0,03	0,18	0,09

Робимо перевірку: сума ймовірностей дорівнює одиниці.

Слід зазначити, що ряди розподілу для випадкових величин, які утворені додаванням, множенням та діленням двох випадкових величин X, Y будуються аналогічним чином.

б) Складемо допоміжну таблицю за наступним принципом: $u_i = 2 \cdot x_i$, а ймовірність p_i лишається без змін.

Остаточний закон розподілу (табл. 11.5):

Таблиця 11.5. Ряд розподілу дискретної випадкової величини U

u_i	2	6	12
p_i	0,5	0,2	0,3

Слід зазначити, що ряд розподілу для випадкової величини, яка утворена множенням величини X будується аналогічним чином.

Відповідь. а) табл. 11.4; б) табл. 11.5.

Приклад 11.2. Наведено закони розподілу двох незалежних випадкових величин (табл. 11.6):

Таблиця 11.6. Ряд розподілу дискретної випадкової величини X

x_i	0	2	4
p_i	0,5	0,2	0,3

Знайти закони розподілу випадкових величин: **а)** $Z = X \cdot X$, **б)** $U = X^2$.

Розв'язання. а) Складемо допоміжну таблицю за наступним принципом: $z_{ij} = x_i \cdot x_j, g_{ij} = p_i \cdot q_j$ ($i, j = \overline{1,3}$) (табл. 11.7).

Таблиця 11.7. Розрахункова таблиця для прикладу 11.2

$X \backslash Y$		$x_1 = 0$	$x_2 = 2$	$x_3 = 4$
		$p_1 = 0,5$	$p_2 = 0,2$	$p_3 = 0,3$
$x_1 = 0$	$p_1 = 0,5$	0 0,25	0 0,1	0 0,15
$x_2 = 2$	$p_2 = 0,2$	0 0,1	4 0,04	8 0,06
$x_3 = 4$	$p_3 = 0,3$	0 0,15	8 0,06	16 0,09

Остаточний закон розподілу (табл. 11.8):

Таблиця 11.8. Ряд розподілу дискретної випадкової величини Z

z_k	0	4	8	16
g_k	0,75	0,04	0,12	0,09

Робимо перевірку: сума ймовірностей дорівнює одиниці.

Слід зазначити, що ряди розподілу для випадкових величин, які утворені додаванням, множенням та діленням двох випадкових величин X, Y будуються аналогічним чином.

б) Складемо допоміжну таблицю за наступним принципом: $u_i = x_i^2$, а ймовірність p_i лишається без змін.

Остаточний закон розподілу (табл. 11.9):

Таблиця 11.9. Ряд розподілу дискретної випадкової величини U

u_i	0	4	16
p_i	0,5	0,2	0,3

Слід зазначити, що ряд розподілу для випадкової величини, яка утворена множенням величини X будується аналогічним чином.

Відповідь. а) табл. 11.8; б) табл. 11.9.

З прикладу 11.2 видно, що операції множення випадкової величини X на себе n разів та піднесення до n -го степеню не є тотожними. Аналогічно, сума n разів випадкової величини X та множення її на константу n також не є тотожними.

11.2. Функції одного випадкового аргументу.

Числові характеристики функції випадкового аргументу

Функції від випадкових величин — це математичні вирази, які перетворюють одну або кілька випадкових величин у нову випадкову величину. Це корисно в багатьох прикладних задачах, таких як статистичний аналіз, теорія ймовірностей, обробка сигналів, економіка тощо. Функції від випадкових величин широко використовуються для моделювання реальних процесів, оцінювання параметрів, статистичного аналізу та оптимізації.

Вищеописані операції над випадковими величинами (додавання, віднімання, лінійна комбінація випадкових величин, піднесення до степеню) є прикладами функцій від випадкових величин.

Розглянемо функцію від одного випадкового аргументу.

Нехай кожному можливому значенню випадкової величини X відповідає одне можливе значення випадкової величини Y , то Y називають *функцією випадкового аргументу* $X: Y = \varphi(X)$.

1. Нехай X – дискретна випадкова величина.

Якщо різним можливим значенням випадкової величини X відповідають різні значення функції Y , то ймовірність відповідних значень X і Y рівні між собою. Якщо різним можливим значенням випадкової величини X відповідають значення Y , серед яких є рівні між собою, то ймовірності рівних значень Y додаються.

Нехай задана дискретна випадкова величина X (табл. 11.10):

Таблиця 11.10. Ряд розподілу дискретної випадкової величини X

x_i	x_1	x_2	$\dots$	x_n	$\dots$
p_i	p_1	p_2	$\dots$	p_n	$\dots$

Нехай $Y = \varphi(X)$. Тоді справедливі формули (11.1)-(11.2):

$$M(\varphi(X)) = \sum_k \varphi(x_k) \cdot p_k, \quad (11.1)$$

$$D(\varphi(X)) = \sum_k \varphi^2(x_k) \cdot p_k - M^2(\varphi(X)). \quad (11.2)$$

2. Нехай X неперервна випадкова величина з щільністю розподілу $f(x)$, а $Y = \varphi(X)$.

Якщо $Y = \varphi(X)$ – диференційована строго зростаюча або строго спадна функція, обернена до якої $X = \psi(Y)$, то щільність розподілу $g(Y)$ випадкової величини Y знаходиться з рівності (11.3):

$$g(Y) = f(\psi(Y)) \cdot |\psi'(Y)|. \quad (11.3)$$

Якщо функція $Y = \varphi(X)$ в інтервалі можливих значень X не монотонна, то потрібно розбити цей інтервал на такі інтервали, на яких $\varphi(X)$ монотонна, знайти щільності розподілів $g_i(Y)$ для кожного з інтервалів монотонності, а потім подати $g(Y)$ у вигляді суми (11.4):

$$g(Y) = \sum_i g_i(Y). \quad (11.4)$$

Числові характеристики випадкової величини Y знаходяться за формулами (11.5)-(11.6):

$$M(\varphi(X)) = \int_{-\infty}^{+\infty} (\varphi(x) \cdot f(x)) dx, \quad (11.5)$$

$$D(\varphi(X)) = \int_{-\infty}^{+\infty} (\varphi^2(x) \cdot f(x)) dx - M^2(\varphi(X)). \quad (11.6)$$

Зокрема, якщо можливі значення X належать інтервалу (a, b) , то справедливі формули (11.7)-(11.8):

$$M(\varphi(X)) = \int_a^b (\varphi(x) \cdot f(x)) dx, \quad (11.7)$$

$$D(\varphi(X)) = \int_a^b (\varphi^2(x) \cdot f(x)) dx - M^2(\varphi(X)). \quad (11.8)$$

Для зручності функцію розподілу для випадкової величини Y позначимо як $G(y)$.

Приклад 11.2. Дискретна випадкова величина X задана рядом розподілу (табл. 11.11):

Таблиця 11.11. Ряд розподілу дискретної випадкової величини X

x_i	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$
p_i	0,1	0,5	0,4

Знайти розподіл функції $Y = \sin(X)$, $M(Y)$, $D(Y)$, $\sigma(Y)$.

Розв'язання. Знайдемо можливі значення Y :

$$y_1 = \sin(0) = 0, y_2 = \sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}, y_3 = \sin\left(\frac{\pi}{2}\right) = 1.$$

Одержимо ряд розподілу Y (табл. 11.12):

Таблиця 11.12. Ряд розподілу дискретної випадкової величини Y

y_i	0	$\frac{1}{\sqrt{2}}$	1
p_i	0,1	0,5	0,4

$$M(Y) = 0 \cdot 0,1 + \frac{1}{\sqrt{2}} \cdot 0,5 + 1 \cdot 0,4 = 0,755; M(Y^2) = 0 \cdot 0,1 + 0,5 \cdot 0,5 + 1 \cdot 0,4 = 0,65,$$

$$D(Y) = M(Y^2) - M^2(Y) = 0,65 - 0,755^2 = 0,08, \sigma(Y) = 0,28.$$

Відповідь. табл. 12.8; $M(Y) = 0,65$; $D(Y) = 0,08$; $\sigma(Y) = 0,28$.

Приклад 11.3. Дискретна випадкова величина X задана рядом розподілу (табл. 11.13):

Таблиця 11.13. Ряд розподілу дискретної випадкової величини X

x_i	-1	1	2
p_i	0,1	0,5	0,4

Знайти розподіл функції $Y = X^2$.

Розв'язання. Можливі значення Y : $y_1 = 1$, $y_2 = 1$, $y_3 = 4$. Значення $y_1 = y_2$ рівні між собою. Ймовірність цього значення

$$p_1 = 0,1 + 0,5 = 0,6.$$

Одержимо розподіл Y (табл. 11.14):

Таблиця 11.14. Ряд розподілу дискретної випадкової величини Y

y_i	1	4
p_i	0,6	0,4

Відповідь. табл. 11.14.

Приклад 11.4. Дана щільність розподілу $f(x)$ випадкової величини X , можливі значення якої належать інтервалу (a, b) . Знайти щільність розподілу $g(y)$ випадкової величини $Y = A \cdot X + B$.

Розв'язання. Функція $y = A \cdot x + B$ монотонна в інтервалі (a, b) , отже, має обернену функцію

$$\psi(y) = x = \frac{y-B}{A},$$

звідки

$$(\psi(y))' = \frac{1}{A}.$$

За формулою (11.3)

$$g(y) = \begin{cases} \left|\frac{1}{A}\right| \cdot f\left(\frac{y-B}{A}\right), & y \in (A \cdot a + B, A \cdot b + B); \\ 0, & y \notin (A \cdot a + B, A \cdot b + B). \end{cases}$$

Необхідно також зробити перевірку:

$$\int_{-\infty}^{+\infty} g(y) dy = 1.$$

Відповідь. $g(y) = \begin{cases} \left|\frac{1}{A}\right| \cdot f\left(\frac{y-B}{A}\right), & y \in (A \cdot a + B, A \cdot b + B); \\ 0, & y \notin (A \cdot a + B, A \cdot b + B). \end{cases}$

Приклад 11.5. Дана щільність розподілу $f(x)$ випадкової величини X , можливі значення якої належать інтервалу $(-\infty, +\infty)$. Знайти щільність розподілу випадкової величини $Y = X^2$.

Розв'язання. Функція $y = x^2$ спадна в інтервалі $(-\infty, 0)$ і в цьому інтервалі має обернену функцію

$$\psi_1(y) = x = -\sqrt{y},$$

звідки

$$(\psi_1(y))' = -\frac{1}{2\sqrt{y}}.$$

За формулою (11.3):

$$g_1(y) = \frac{1}{2\sqrt{y}} \cdot f(-\sqrt{y}).$$

Функція $y = x^2$ зростаюча в інтервалі $[0, +\infty)$ і в цьому інтервалі має обернену функцію

$$\psi_2(y) = x = \sqrt{y},$$

звідки

$$(\psi_2(y))' = \frac{1}{2\sqrt{y}}.$$

Отже,

$$g_2(y) = \frac{1}{2\sqrt{y}} \cdot f(\sqrt{y}).$$

За формулою (11.4)

$$g(y) = g_1(y) + g_2(y),$$

Тобто

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}} \cdot (f(-\sqrt{y}) + f(\sqrt{y})), & y \in (0, +\infty); \\ 0, & y \in (-\infty, 0]. \end{cases}$$

Необхідно також зробити перевірку:

$$\int_{-\infty}^{+\infty} g(y) dy = 1.$$

Відповідь. $g(y) = \begin{cases} \frac{1}{2\sqrt{y}} \cdot (f(-\sqrt{y}) + f(\sqrt{y})), & y \in (0, +\infty); \\ 0, & y \in (-\infty, 0]. \end{cases}$

Приклад 11.6. Випадкова величина X задана щільністю розподілу:

$$f(x) = \begin{cases} 0, & x \leq 0, \\ 0,5 \cdot x, & 0 < x \leq 2, \\ 0, & x > 2. \end{cases}$$

Знайти $G(y)$, $M(Y)$, $D(Y)$ для випадкової величини $Y = X^2$.

Розв'язання. За формулами (11.7), (11.8) одержимо $M(Y)$, $D(Y)$ без урахування функції щільності випадкової величини Y :

$$M(Y) = \int_0^2 (x^2 \cdot 0,5 \cdot x) dx = \left. \frac{x^4}{8} \right|_0^2 = 2,$$

$$M(Y^2) = \int_0^2 (x^4 \cdot 0,5 \cdot x) dx = \left. \frac{x^6}{12} \right|_0^2 = \frac{16}{3},$$

$$D(Y) = M(Y^2) - M^2(Y) = \frac{16}{3} - 2^2 = \frac{4}{3}.$$

Функція $y = x^2$ зростаюча в інтервалі $[0, +\infty)$. Тобто на цьому інтервалі вона має обернену функцію:

$$\psi(y) = x = \sqrt{y},$$

звідки

$$(\psi(y))' = \frac{1}{2\sqrt{y}},$$

отже

$$g(y) = \frac{1}{2\sqrt{y}} \cdot f(\sqrt{y}) = \frac{1}{2\sqrt{y}} \cdot \frac{1}{2} \cdot \sqrt{y} = \frac{1}{4} \text{ при } 0 < y \leq 4,$$

оскільки $y = x^2$, а $0 < x \leq 2$.

Остаточно щільність випадкової величини Y :

$$g(y) = \begin{cases} \frac{1}{4}, & y \in (0, 4]; \\ 0, & y \notin (0, 4]. \end{cases}$$

Необхідно також зробити перевірку:

$$\int_{-\infty}^{+\infty} g(y) dy = \int_0^4 \frac{1}{4} dy = 1.$$

Функція розподілу матиме вигляд:

$$G(y) = \begin{cases} 0, & \text{при } y \in (-\infty, 0]; \\ \frac{y}{4}, & \text{при } y \in (0, 4]; \\ 1, & \text{при } y \in (4, +\infty). \end{cases}$$

Маючи функцію щільності випадкової величини Y , безпосередньо отримаємо $M(Y)$, $D(Y)$

$$M(Y) = \int_0^4 \left(y \cdot \frac{1}{4}\right) dy = \frac{y^2}{8} \Big|_0^4 = 2,$$

$$M(Y^2) = \int_0^4 \left(y^2 \cdot \frac{1}{4}\right) dy = \frac{y^3}{12} \Big|_0^4 = \frac{16}{3},$$

$$D(Y) = M(Y^2) - M^2(Y) = \frac{16}{3} - 2^2 = \frac{4}{3}.$$

Відповідь. $G(y) = \begin{cases} 0, & \text{при } y \in (-\infty, 0]; \\ \frac{y}{4}, & \text{при } y \in (0, 4]; \\ 1, & \text{при } y \in (4, +\infty); \end{cases} \quad M(Y) = 2; D(Y) = \frac{4}{3}.$

12.3. Представлення функції розподілу та щільності системи двох випадкових величин за допомогою функції Дірака

Розглянемо окремо випадок, коли функція розподілу $G(y)$ має точки розриву $y_1, y_2, \dots, y_n$ зі стрибками $p_1, p_2, \dots, p_n$. Це означає існування значень випадкової величини η (збіжних з точками розриву $y_1, y_2, \dots, y_n$), яким відповідають ненульові ймовірності $p_1, p_2, \dots, p_n$. У даному випадку щільність розподілу ймовірностей в точках $y_1, y_2, \dots, y_n$ перетворюється на нескінченність. Математична ідеалізація цього явища спирається на використання *дельта-функції Дірака* $\delta(y)$, яка не є функцією у звичайному розумінні, а являє собою так звану узагальнену функцію.

Розглядатимемо $\delta(y)$ як похідну функції одиничного стрибка (рис. 11.1, 11.2).

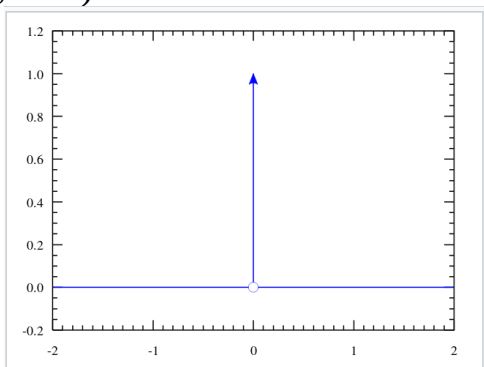


Рис.11.1. Схематичне зображення дельта-функції Дірака, як лінії, з якої виступає стрілка

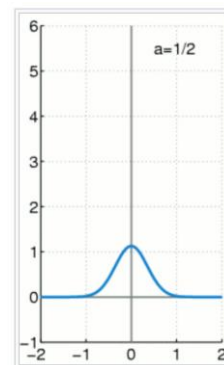


Рис. 11.2. Дельта-функція Дірака як границя послідовності гаусівських функцій розподілу

У класичному аналізі функція $\eta(y)$ не диференційована в точці $y = 0$, однак у теорії узагальнених функцій цього обмеження немає. Тоді справедлива рівність (11.9):

$$\delta(y) = \eta'(y), \text{ де } \eta(y) = \begin{cases} 1, & y > 0, \\ 0, & y \leq 0. \end{cases} \quad (11.9)$$

Для дельта-функції справедливі наступні **властивості**:

1. $\delta(y) = 0, \forall y \neq 0$;
2. $\int_{-\infty}^{+\infty} \delta(y) dy = 1$;
3. $y \cdot \delta'(y) = \delta(y)$;

4. $\delta(G(y)) = \sum_k \frac{\delta(y-y_k)}{|(G(y_k))'|}$, де y_k – нулі функції $G(y)$.

Зобразимо функцію $G(y)$ у вигляді (11.10):

$$G(y) = \tilde{G}(y) + \sum_{k=1}^n (p_k \cdot \eta(y - y_k)), \quad (11.10)$$

де $\tilde{G}(y)$ – неперервна («замкнена») функція. Для функції щільності одержуємо рівність (11.11):

$$g(y) = \tilde{g}(y) + \sum_{k=1}^n (p_k \cdot \delta(y - y_k)), \text{ де } \tilde{g}(y) = (\tilde{G}(y))'. \quad (11.11)$$

Приклад 11.7. За заданою щільністю розподілу

$$f(\xi_1) = \begin{cases} \frac{1}{2}, & \text{де } \xi_1 \in [-1; 1]; \\ 0, & \text{де } \xi_1 \notin [-1; 1]. \end{cases}$$

випадкової величини ξ_1 визначити функцію розподілу випадкової величини $\xi_2 = \varphi(\xi_1)$. Функцію $\xi_2 = \varphi(\xi_1)$ задано графічно на рис. 11.3:

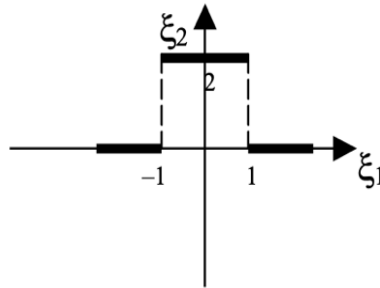


Рис. 11.3. Графік функції $\xi_2 = \varphi(\xi_1)$

Дати вираз для щільності $g(\xi_2)$ розподілу випадкової величини ξ_2 .

Розв'язання. Запишемо аналітично функцію, представлену на рис. 11.3.

$$\xi_2 = \begin{cases} 2, & \text{де } \xi_1 \in [-1; 1]; \\ 0, & \text{де } \xi_1 \notin [-1; 1]. \end{cases}$$

Випадок 1: $\xi_2 = 2$

З графіку та з попереднього аналітичного виразу видно, що $\xi_2 = 2$ для $\xi_1 \in [-1; 1]$. Тобто, ймовірність того, що $\xi_2 = 2$ дорівнюватиме ймовірності того, що $\xi_1 \in [-1; 1]$. Таким чином, враховуючи, що випадкова величина ξ_1 має неперервний рівномірний розподіл:

$$p\{\xi_2 = 2\} = p\{\xi_1 \in [-1; 1]\} = \int_{-1}^1 \left(\frac{1}{2}\right) dx = \frac{1+1}{2} = 1.$$

Отже, $\xi_2 = 2$ з ймовірністю 1.

Випадок 2: $\xi_2 = 0$

Оскільки $\xi_2 = 0$ виникає при $\xi_1 \notin [-1; 1]$, ймовірність того, що ξ_2 прийме це значення, буде дорівнювати ймовірності того, що $\xi_1 \notin [-1; 1]$. Випадкова величина ξ_1 має неперервний рівномірний розподіл, що означає, що $p\{\xi_1 \notin [-1; 1]\} = 1 - p\{\xi_1 \in [-1; 1]\} = 1 - 1 = 0$. Отже, $p(\xi_2 = 0) = 0$.

$$p_2(\xi_2) = \begin{cases} 0, & \text{де } \xi_2 = 0; \\ 1, & \text{де } \xi_2 = 2. \end{cases}$$

Виразимо ξ_1 через ξ_2 . Оскільки графік (рис. 11.3) має розриви, то представимо випадкову величину ξ_1 за допомогою дельта-функції Дірака.

$$\xi_1 = \psi(\xi_2) = \begin{cases} \delta(\xi_2), & \text{де } \xi_2 = 0; \\ \delta(\xi_2 - 2), & \text{де } \xi_2 = 2. \end{cases}$$

Тоді за формулою (11.11):

$$g(\xi_2) = \delta(\xi_2) \cdot 0 + \delta(\xi_2 - 2) \cdot 1 = \delta(\xi_2 - 2), \text{ де } \xi_2 = 2.$$

Це означає, що випадкова величина ξ_2 завжди дорівнює 0, тобто ймовірність того, що вона прийме будь-яке інше значення, крім 2, дорівнює 0.

Таким чином, функція $\delta(\xi_2)$ вказує на те, що щільність розподілу ξ_2 сконцентрована в одній точці, в даному випадку в точці $\xi_2 = 2$.

Перевірка:

$$\int_{-\infty}^{+\infty} g(\xi_2) d\xi_2 = \int_{-\infty}^{+\infty} \delta(\xi_2 - 2) d\xi_2 = 1.$$

Відповідь. $g(\xi_2) = \delta(\xi_2 - 2)$, де $\xi_2 = 2$.

12.4. Функції двох випадкових аргументів

Якщо кожній парі можливих значень випадкових величин X і Y відповідає одне можливе значення випадкової величини Z , то Z називають *функцією двох випадкових величин* X і Y : $Z = \varphi(X, Y)$.

Далі розглянемо як знайти розподіл суми $Z = X + Y$ за відомими розподілами доданків.

1. Нехай X і Y – дискретні незалежні випадкові величини. Можливі значення $Z = X + Y$ дорівнюють сумах кожного можливого значення X з кожним можливим значенням Y ; ймовірності можливих значень дорівнюють добуткам ймовірностей доданків.

2. Нехай X і Y – неперервні незалежні випадкові величини з щільностями розподілу $f_1(x)$ і $f_2(y)$, то щільність розподілу $g(z)$ суми $Z = X + Y$ за умови, що щільність розподілу хоча б одного з аргументів задана в інтервалі $(-\infty, +\infty)$ одною формулою, можна знайти за формулою (11.12):

$$g(z) = \int_{-\infty}^{+\infty} (f_1(x) \cdot f_2(z - x)) dx \quad (11.12)$$

або за формулою (11.13):

$$g(z) = \int_{-\infty}^{+\infty} (f_1(z - y) \cdot f_2(y)) dy. \quad (11.13)$$

Функція $g(z)$ утворена з функцій $f_1(x)$ і $f_2(y)$ за формулою (11.12) або (11.13) називається *згорткою* або *композицією* цих функцій.

Якщо можливі значення аргументів невід'ємні, то щільність розподілу $Z = X + Y$ знаходиться за формулою (11.14):

$$g(z) = \int_0^z (f_1(x) \cdot f_2(z - x)) dx \quad (11.14)$$

або за рівносильною формулою (11.15):

$$g(z) = \int_0^z (f_1(z - y) \cdot f_2(y)) dy. \quad (11.15)$$

Композиція незалежних нормально розподілених випадкових величин також має нормальний розподіл. Математичне сподівання і дисперсія цієї композиції відповідно дорівнюють сумах математичних сподівань і дисперсій доданків.

Приклад 11.8. Дискретні випадкові величини X і Y задані рядами розподілів (табл. 11.15, 11.16):

Таблиця 11.15. Ряд розподілу дискретної випадкової величини X

x_i	8	10	12
p_i	0,4	0,1	0,5

Таблиця 11.16. Ряд розподілу дискретної випадкової величини Y

y_i	1	2
p_i	0,2	0,8

Знайти розподіл випадкової величини $Z = X + Y$.

Розв'язання. Знайдемо можливі значення Z :

$$z_1 = 8 + 1 = 9, z_2 = 8 + 2 = 10, z_3 = 10 + 1 = 11, z_4 = 10 + 2 = 12, \\ z_5 = 12 + 1 = 13, z_6 = 12 + 2 = 14.$$

Оскільки випадкові величини X і Y незалежні, то події $X = 10$ і $Y = 1$ також незалежні, тому ймовірність їх одночасного настання (тобто ймовірність події $Z = 9$) за теоремою множення дорівнює $0,4 \cdot 0,2 = 0,08$, так само знаходимо: $P\{Z = 8 + 2 = 10\} = 0,4 \cdot 0,8 = 0,32$; $P\{Z = 10 + 1 = 11\} = 0,1 \cdot 0,2 = 0,02$; $P\{Z = 10 + 2 = 12\} = 0,1 \cdot 0,8 = 0,08$; $P\{Z = 12 + 1 = 13\} = 0,5 \cdot 0,2 = 0,1$, $P\{Z = 12 + 2 = 14\} = 0,5 \cdot 0,8 = 0,4$.

Одержимо ряд розподілу випадкової величини Z (табл. 11.17):

Таблиця 11.17. Ряд розподілу дискретної випадкової величини Z

z_i	9	10	11	12	13	14
p_i	0,08	0,32	0,02	0,08	0,1	0,4

Контроль: $0,08 + 0,32 + 0,02 + 0,08 + 0,1 + 0,4 = 1$.

Відповідь. табл. 11.17.

Приклад 11.9. Незалежні випадкові величини X і Y задані щільностями розподілів:

$$f_1(x) = \frac{1}{7} \cdot e^{-\frac{x}{2}} (0 \leq x < \infty), f_2(y) = \frac{1}{7} \cdot e^{-\frac{y}{2}} (0 \leq y < \infty)$$

Знайти щільність розподілу випадкової величини $Z = X + Y$.

Розв'язання. За формулою (11.12)

$$g(z) = \int_0^z \left(\frac{1}{2} \cdot e^{-\frac{x}{2}} \cdot \frac{1}{7} \cdot e^{-\left(\frac{z-x}{7}\right)} \right) dx = \frac{1}{14} \cdot e^{-\frac{z}{7}} \cdot \int_0^z e^{-\left(\frac{5}{14}x\right)} dx = \frac{1}{14} \cdot \\ e^{-\frac{z}{7}} \cdot \left(-\frac{14}{5} \cdot e^{-\left(\frac{5}{14}x\right)} \right) \Big|_0^z = \frac{1}{5} \cdot e^{-\frac{z}{7}} \cdot \left(1 - e^{-\left(\frac{5}{14}z\right)} \right).$$

Оскільки можливі значення X і Y невід'ємні, то $z > 0$, отже:

$$g(z) = \begin{cases} \frac{1}{5} \cdot e^{-\frac{z}{7}} \cdot \left(1 - e^{-\left(\frac{5}{14}z\right)} \right), & 0 \leq z < +\infty, \\ 0, & -\infty < z < 0. \end{cases}$$

$$\text{Перевірка: } \int_0^\infty \left(\frac{1}{5} \cdot e^{-\frac{z}{7}} - \frac{1}{5} \cdot e^{-\frac{z}{2}} \right) dz = \frac{7}{5} - \frac{2}{5} = 1.$$

Відповідь. $g(z) = \begin{cases} \frac{1}{5} \cdot e^{-\frac{z}{7}} \cdot \left(1 - e^{-\left(\frac{5}{14}z\right)} \right), & 0 \leq z < +\infty, \\ 0, & -\infty < z < 0. \end{cases}$

Приклад 11.10. Задані щільності розподілу незалежних випадкових величин X і Y :

$$f_1(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & x \notin [0,1], \end{cases} \quad f_2(y) = \begin{cases} 2 \cdot y, & y \in [0,1], \\ 0, & y \notin [0,1]. \end{cases}$$

Знайти щільність розподілу випадкової величини $Z = X + Y$.

Розв'язання. Розглянемо наступні випадки:

Якщо $z < 0$, то $z - x < 0$, тому $f_2(z - x) = 0$, отже $g(z) = 0$.

Якщо $z > 2$, то $z - x > z - 1 > 1$, тому $f_2(z - x) = 0$, отже $g(z) = 0$.

Нехай $1 < z \leq 2$. Підінтегральна функція $f_2(z - x)$ відмінна від нуля тільки при тих значеннях X , які задовольняють нерівності $0 \leq z - x \leq 1$, звідки. $z - 1 \leq x \leq z$.

Нехай $0 < z \leq 1$. Тоді

$$g(z) = 2 \cdot \int_0^z (z - x) dx = (2 \cdot z \cdot x - x^2)|_0^z = z^2.$$

Нехай $1 < z \leq 2$.

$$g(z) = 2 \cdot \int_{z-1}^1 (z - x) dx = (2 \cdot z \cdot x - x^2)|_{z-1}^1 = 2 \cdot z - z^2.$$

Об'єднуючи ці результати, дістанемо:

$$g(z) = \begin{cases} 0, & z \leq 0; \\ z^2, & 0 < z \leq 1; \\ 2 \cdot z - z^2, & 1 < z \leq 2; \\ 0, & z > 2. \end{cases}$$

Необхідно також зробити перевірку:

$$\begin{aligned} \int_{-\infty}^{+\infty} g(z) dz &= \int_0^1 z^2 dz + \int_1^2 (2 \cdot z - z^2) dz = \frac{1}{3} + \left(4 - \frac{8}{3} - 1 + \frac{1}{3} \right) = \frac{1}{3} + \frac{2}{3} \\ &= 1. \end{aligned}$$

Відповідь. $g(z) = \begin{cases} 0, & z \leq 0; \\ z^2, & 0 < z \leq 1; \\ 2 \cdot z - z^2, & 1 < z \leq 2; \\ 0, & z > 2. \end{cases}$

ЛЕКЦІЯ 12

СИСТЕМИ ВИПАДКОВИХ ВЕЛИЧИН

Упорядкований набір $(X_1, X_2, \dots, X_n)$ випадкових величин X_i ($i = 1, \dots, n$), які задані на одному й тому самому просторі елементарних подій Ω , називається *системою n випадкових величин* або *n -вимірним випадковим вектором*.

На багатовимірні випадкові величини поширюються майже без змін основні означення, які були розглянуті для одновимірної випадкової величини.

Розглянемо систему двох випадкових величин.

12.1. Двовимірні випадкові величини (системи двох випадкових величин)

Пара випадкових величин X і Y утворює *двовимірну випадкову величину* (X, Y) (*вектор*). Розподіл двовимірної випадкової величини задається функцією двох змінних (12.1):

$$F(x, y) = P(X < x; Y < y). \quad (12.1)$$

Геометрично рівність (12.1) можна тлумачити так: $F(x, y)$ – це ймовірність того, що випадкова точка (X, Y) попаде в нескінченний квадрат з вершиною (x, y) , розміщений лівіше і нижче цієї вершини (рис. 12.1).

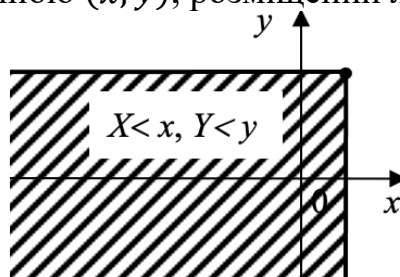


Рис. 12.1. Інтерпретація функції розподілу двовимірної випадкової величини

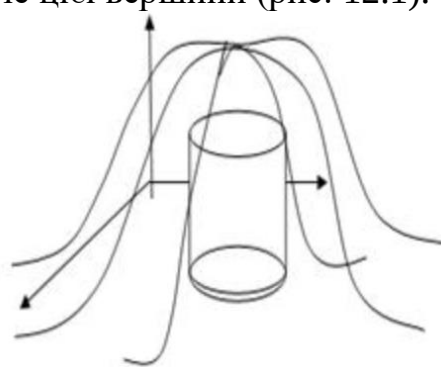


Рис. 12.2. Поверхня розподілу двовимірної випадкової величини

Властивості функції розподілу

- $0 \leq F(x, y) \leq 1$;
- Функція розподілу неперервна зліва по кожному аргументу:
 $\lim_{x \rightarrow x_1+0} F(x, y) = F(x_1, y)$; $\lim_{y \rightarrow y_1+0} F(x, y) = F(x, y_1)$;
- Функція $F(x, y)$ неспадна по кожному аргументу:
 $F(x_1, y) \leq F(x_2, y)$, якщо $x_1 < x_2$; $F(x, y_1) \leq F(x, y_2)$, якщо $y_1 < y_2$;
- $F(x, +\infty) = F_1(x)$ – функція розподілу компоненти X ,

$F(+\infty, y) = F_2(y)$ – функція розподілу компоненти Y ;
5. $F(-\infty, y) = 0$; $F(x, -\infty) = 0$; $F(-\infty, -\infty) = 0$; $F(+\infty, +\infty) = 1$.

Двовимірну випадкову величину (X, Y) називають **дискретною**, якщо множина значень, які вона може набути, є скінченною або зліченною. Для завдання дискретної двовимірної величини досить задати її можливі значення (x_i, y_j) і ймовірності кожного значення $p_{ij} = P\{X = x_i, Y = y_j\}$, $p_i = P\{X = x_i\}$, $q_j = P\{Y = y_j\}$ ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$).

Закон розподілу такої випадкової величини задається таблицею (табл. 12.1):

Таблиця 12.1. Матриця розподілу системи дискретних випадкових величин (X, Y)

X	Y				
	y_1	y_2	$\dots$	y_n	$\sum_{i=1}^m p_i$
x_1	p_{11}	p_{12}	$\dots$	p_{1n}	p_1
x_2	p_{21}	p_{22}	$\dots$	p_{2n}	p_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_m	p_{m1}	p_{m2}	$\dots$	p_{mn}	p_m
$\sum_{j=1}^n q_j$	q_1	q_2	$\dots$	q_n	1

В табл. 12.1 для ймовірностей p_i та q_j ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$) справедливі умови (12.2) та (12.3):

$$p_1 = \sum_{j=1}^n p_{1j}, p_2 = \sum_{j=1}^n p_{2j}, \dots, p_m = \sum_{j=1}^n p_{mj}, \quad (12.2)$$

$$q_1 = \sum_{i=1}^m q_{i1}, q_2 = \sum_{i=1}^m q_{i2}, \dots, q_n = \sum_{i=1}^m q_{in}. \quad (12.3)$$

Сума всіх ймовірностей, що знаходяться в таблиці, дорівнює одиниці. Додаючи ймовірності стовпців і рядків, одержимо закони розподілу компонент X і Y .

Закони розподілу випадкових величин X і Y можна записати (табл. 12.2, 12.3):

Таблиця 12.2. Ряд розподілу дискретної випадкової величини X

x_i	x_1	x_2	$\dots$	x_m
p_i	p_1	p_2	$\dots$	p_m

В табл. 12.2 ймовірності p_i обчислюються за формулами (12.2).

Таблиця 12.3. Ряд розподілу дискретної випадкової величини Y

y_j	y_1	y_2	$\dots$	y_n
q_j	q_1	q_2	$\dots$	q_n

В табл. 12.3 ймовірності q_j обчислюються за формулами (12.3).

Двовимірну випадкову величину (X, Y) називають **неперервною**, якщо її функція розподілу $F(x, y)$ має неперервну мішану похідну другого порядку: $F''_{xy} = F''_{yx}$.

Щільність ймовірностей $f(x, y)$ неперервної випадкової величини (X, Y) визначається формулою (12.4):

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \quad (12.4)$$

Тоді $F(x, y)$ можна подати у вигляді (12.5):

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy. \quad (12.5)$$

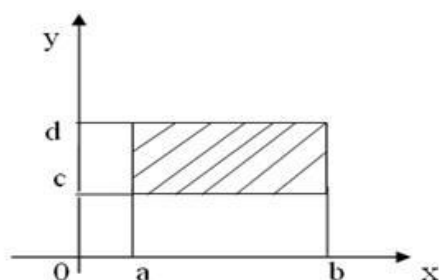
Властивості щільності розподілу

1. $f(x, y) \geq 0$ при всіх (x, y) ;
2. Для щільності $f(x, y)$ неперервної випадкової величини (X, Y) справедлива рівність (12.6):

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dy dx = 1; \quad (12.6)$$

3. Якщо D – довільна область у площині OXY , то ймовірність попадання в область D (приклад області наведено на рис. 12.3) дорівнює (12.7):

$$P\{(X, Y) \in D\} = \iint_D f(x, y) dx dy. \quad (12.7)$$



Задано двовимірний закон розподілу дискретних випадкових величин:

Y	X			
	1	2	3	Σ
3	0,35	0,20	0,25	0,80
4	0,10	0,05	0,05	0,20
Σ	0,45	0,25	0,30	1

б) функцію розподілу $F(X, Y)$;

X \ Y	$(-\infty; 1]$	$(1; 2]$	$(2; 3]$	$(3; +\infty)$
$(-\infty; 3]$	0	0	0	0
$(3; 4]$	0	0,35	0,55	0,8
$(4; +\infty)$	0	0,45	0,7	1

в) ймовірність події $P(X < 3; Y < 4)$

$$\begin{aligned}
 P(X < 3; Y < 4) &= 0,35 + 0,20 = 0,55 \\
 P(X < 3; Y < 4) &= P(-\infty < X < 3; -\infty < Y < 4) = \\
 &= F(3; 4) - F(-\infty; 4) - F(3; -\infty) + F(-\infty; -\infty) = \\
 &= 0,55 - 0 - 0 + 0 = 0,55
 \end{aligned}$$

Рис. 12.3. Приклад області D

Рис. 12.4. Принцип побудови матриці розподілу для системи (X, Y) та обчислення $P\{a \leq X < b; c \leq Y < d\}$

Якщо область $D = \{a \leq X < b; c \leq Y < d\}$, то справедлива рівність (12.8):

$$P\{(X, Y) \in D\} = F(b, d) - F(a, d) - F(b, c) + F(a, c). \quad (12.8)$$

Приклад побудови матриці розподілу наведено на рис. 12.4.

Знаючи щільність розподілу $f(x, y)$ двовимірної випадкової величини (X, Y) , можна знайти щільність розподілу $f_1(x)$ і $f_2(y)$ для її компонент X і Y .

Функції розподілу одновимірних випадкових величин X і Y мають вигляд (12.9):

$$\begin{aligned} F_1(x) &= F(x, +\infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(x, y) dx dy, \\ F_2(y) &= F(+\infty, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^y f(x, y) dx dy. \end{aligned} \quad (12.9)$$

Диференціюючи функції розподілу $F_1(x)$ і $F_2(y)$ відповідно по аргументам x і y , одержимо рівності (12.10):

$$f_1(x) = \frac{\partial F_1(x)}{\partial x} = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_2(y) = \frac{\partial F_2(y)}{\partial y} = \int_{-\infty}^{+\infty} f(x, y) dx. \quad (12.10)$$

Приклад 12.1. Задано дискретна двовимірна випадкова величина (табл. 12.4):

Таблиця 12.4. Ряд розподілу системи випадкових величин (X, Y)

X	Y			
	1	2	3	$\sum$
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6
5	0,2	0	0	0,2
$\sum$	0,3	0,3	0,4	1

Знайти закони розподілу компонент X і Y .

Розв'язання. Додаючи стовпці і рядки таблиці ймовірностей, дістанемо закони розподілу випадкових величин X і Y відповідно (табл. 12.5, 24.6).

Таблиця 12.5. Ряд розподілу випадкової величини X

x_i	0	2	5
p_i	0,2	0,6	0,2

Таблиця 12.6. Ряд розподілу випадкової величини Y

y_j	1	2	3
q_j	0,3	0,3	0,4

Відповідь. табл. 12.5; табл. 12.6.

Приклад 12.2. Задано функція розподілу двовимірної випадкової величини (X, Y) :

$$F(x, y) = \begin{cases} (1 - e^{-(3 \cdot x)}) \cdot (1 - e^{-(5 \cdot y)}), & x \geq 0, y \geq 0; \\ 0, & x < 0, y < 0. \end{cases}$$

Знайти щільність розподілу $f(x, y)$.

Розв'язання. Знаходимо

$$\frac{\partial F(x, y)}{\partial x} = \begin{cases} 3 \cdot e^{-(3 \cdot x)} \cdot (1 - e^{-(5 \cdot y)}), & x \geq 0, y \geq 0; \\ 0, & x < 0, y < 0. \end{cases}$$

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = \begin{cases} 15 \cdot e^{-(3 \cdot x)} \cdot e^{-(5 \cdot y)}, & x \geq 0, y \geq 0; \\ 0, & x < 0, y < 0. \end{cases}$$

Отже, за формулою (12.4):

$$f(x, y) = \begin{cases} 15 \cdot e^{-(3 \cdot x)} \cdot e^{-(5 \cdot y)}, & x \geq 0, y \geq 0; \\ 0, & x < 0, y < 0. \end{cases}$$

Відповідь. $f(x, y) = \begin{cases} 15 \cdot e^{-(3 \cdot x)} \cdot e^{-(5 \cdot y)}, & x \geq 0, y \geq 0; \\ 0, & x < 0, y < 0. \end{cases}$

12.2. Умовні закони розподілу

Умовним законом розподілу називається розподіл однієї випадкової величини, знайдений за умови, що інша випадкова величина системи набула деякого фіксованого значення.

а) Нехай (X, Y) – **дискретна двовимірна випадкова величина**.

Умовним законом розподілу дискретної випадкової величини X при фіксованому значенні $Y = y_j$ називається перелік можливих значень випадкової величини $X = x_i$ та відповідних їм умовних ймовірностей, обчислених при фіксованому значенні $Y = y_j$. У табличній формі запису цей умовний закон має такий вигляд (табл. 12.7):

Таблиця 12.7. Закон розподілу величини $X/Y = y_j$

x_i	x_1	x_2	...	x_m
$P(x_i/Y = y_j)$	$P(x_1/Y = y_j)$	$P(x_2/Y = y_j)$	...	$P(x_m/Y = y_j)$

Умовні ймовірності $P(x_i/Y = y_j)$ знаходяться за формулою (12.11):

$$P(x_i/Y = y_j) = \frac{p_{ij}}{q_j}, (i = 1, 2, \dots, m), \quad (12.11)$$

де q_j визначається з ряду розподілу компоненти Y (табл. 12.3).

Умовним законом розподілу дискретної випадкової величини Y при фіксованому значенні $X = x_i$ називається перелік можливих значень випадкової величини $Y = y_j$ та відповідних їм умовних ймовірностей, обчислених при фіксованому значенні $X = x_i$. У табличній формі запису цей умовний закон має такий вигляд (табл. 12.8):

Таблиця 12.8. Закон розподілу величини $Y/X = x_i$

y_j	y_1	y_2	...	y_n
$P(y_j/X = x_i)$	$P(y_1/X = x_i)$	$P(y_2/X = x_i)$	...	$P(y_n/X = x_i)$

Умовні ймовірності $P(y_j/X = x_i)$ знаходяться за формулою (12.12):

$$P(Y_j/X = x_i) = \frac{p_{ij}}{p_i}, (j = 1, 2, \dots, n), \quad (12.12)$$

де p_j визначається з ряду розподілу компоненти X (табл. 12.2).

б) Нехай (X, Y) – неперервна двовимірна випадкова величина.

Умовною щільністю $f(x/y)$ розподілу компоненти X при заданому значенні $Y = y$ називається відношення щільності розподілу двовимірної випадкової величини $f(x, y)$ до щільності розподілу $f_2(y)$ компоненти Y (12.13):

$$f(x/y) = \frac{f(x, y)}{f_2(y)}. \quad (12.13)$$

Умовною щільністю $f(y/x)$ розподілу компоненти Y при заданому значенні $X = x$ називається відношення щільності розподілу двовимірної випадкової величини $f(x, y)$ до щільності розподілу $f_1(x)$ компоненти X (12.14):

$$f(y/x) = \frac{f(x, y)}{f_1(x)}. \quad (12.14)$$

Приклад 12.3. Задано двовимірний закон розподілу дискретних випадкових величин (табл. 12.9):

Таблиця 12.9. Ряд розподілу системи випадкових величин (X, Y)

X	Y			
	1	2	3	$\sum$
3	0,15	0,2	0,25	0,6
4	0,1	0,1	0,2	0,4
$\sum$	0,25	0,3	0,45	1

Знайти: а) умовний закон розподілу випадкової величини X при $Y = 2$; б) умовний закон розподілу випадкової величини Y при $X = 3$.

Розв'язання. а) Знаходимо з формули (12.11):

$$P(X_1/Y = 2) = \frac{0,20}{0,30} = \frac{2}{3}, P(X_2/Y = 2) = \frac{0,10}{0,30} = \frac{1}{3}.$$

Таблиця 12.10. Закон розподілу величини $X/Y = 2$

x_i	3	4
$P(x_i/Y = 2)$	$\frac{2}{3}$	$\frac{1}{3}$

б) Знаходимо з формули (12.12):

$$P(Y_1/X = 3) = \frac{0,15}{0,60} = \frac{1}{4}, P(Y_2/X = 3) = \frac{0,20}{0,60} = \frac{1}{3}, P(Y_3/X = 3) = \frac{0,25}{0,60} = \frac{5}{12}.$$

Таблиця 12.11. Закон розподілу величини $Y/X = 3$

y_j	1	2	3
$P(y_j/X = 3)$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{5}{12}$

Відповідь. табл. 12.11.

Приклад 12.4. Двовимірна випадкова величина задана щільністю розподілу:

$$f(x, y) = \frac{2 \cdot \sqrt{2}}{\pi} \cdot e^{-(9 \cdot x^2 + 2 \cdot x \cdot y + y^2)}.$$

Знайти умовні закони розподілу компонент.

Розв'язання. Знайдемо умовні щільності розподілу компонент:

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \frac{2 \cdot \sqrt{2}}{\pi} \cdot \int_{-\infty}^{+\infty} e^{-(9 \cdot x^2 + 2 \cdot x \cdot y + y^2)} dy = \frac{2 \cdot \sqrt{2}}{\pi} \cdot e^{-(8 \cdot x^2)} \\ &\cdot \int_{-\infty}^{+\infty} e^{-(x+y)^2} d(x+y) = \frac{2 \cdot \sqrt{2}}{\sqrt{\pi}} \cdot e^{-(8 \cdot x^2)}, \end{aligned}$$

враховуючи, що інтеграл Пуассона $\int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{\pi}$.

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{+\infty} f(x, y) dx = \frac{2 \cdot \sqrt{2}}{\pi} \cdot \int_{-\infty}^{+\infty} e^{-(9 \cdot x^2 + 2 \cdot x \cdot y + y^2)} dx = \frac{2 \cdot \sqrt{2}}{\pi} \cdot e^{-\left(\frac{8}{9} \cdot y^2\right)} \cdot \frac{1}{3} \cdot \\ &\int_{-\infty}^{+\infty} e^{-\left(3 \cdot x + \frac{1}{3} \cdot y\right)^2} d\left(3 \cdot x + \frac{1}{3} \cdot y\right) = \frac{2 \cdot \sqrt{2}}{3 \cdot \sqrt{\pi}} \cdot e^{-\left(\frac{8}{9} \cdot y^2\right)}. \end{aligned}$$

Одержимо з формул з формул (12.13) та (12.14):

$$\begin{aligned} f(x/y) &= \frac{f(x, y)}{f_2(y)} = \frac{3}{\sqrt{\pi}} \cdot e^{-\left(3 \cdot x + \frac{1}{3} \cdot y\right)^2}, \\ f(y/x) &= \frac{f(x, y)}{f_1(x)} = \frac{1}{\sqrt{\pi}} \cdot e^{-(x+y)^2}. \end{aligned}$$

Відповідь. $f(x/y) = \frac{3}{\sqrt{\pi}} \cdot e^{-\left(3 \cdot x + \frac{1}{3} \cdot y\right)^2}$, $f(y/x) = \frac{1}{\sqrt{\pi}} \cdot e^{-(x+y)^2}$.

12.3. Залежні і незалежні випадкові величини

Дві дискретні випадкові величини X і Y називаються *незалежними*, якщо для кожної ймовірності p_{ij} в матриці розподілу системи випадкових величин (X, Y) (табл. 12.1) справедлива рівність (12.15):

$$p_{ij} = p_i \cdot q_j, \quad (12.15)$$

де p_i, q_j обчислюються за формулами (12.2), (12.3).

Дві неперервні випадкові величини X і Y називаються *незалежними*, якщо закон розподілу однієї з них не залежить від того, які можливі значення прийняла друга величина, і коли функція розподілу пари (X, Y) дорівнює добутку функцій розподілу компонент (12.16):

$$F(x, y) = F_1(x) \cdot F_2(y), \quad (12.16)$$

де $F_1(x), F_2(x)$ обчислюються за формулою (12.9).

В протилежному випадку, при невиконанні рівностей (12.15) або (12.16), випадкові величини X і Y називаються *залежними*.

З формули (12.4) одержимо (12.17):

$$f(x, y) = f_1(x) \cdot f_2(y), \quad (12.17)$$

де $f_1(x), f_2(y)$ обчислюються за формулою (12.10).

Приклад 12.5. Дана функція розподілу двовимірної випадкової величини (X, Y) :

$$F(x, y) = \begin{cases} 1 - 3^{-x} - 3^{-y} + 3^{-(x+y)}, & x \geq 0, y \geq 0; \\ 0, & x < 0 \text{ або } y < 0. \end{cases}$$

Знайти функції розподілів компонент. Перевірити незалежність випадкових величин.

Розв'язання. За властивістю функції розподілу $F(x, y)$ маємо:

$$F_1(x) = F(x, +\infty) = \lim_{y \rightarrow +\infty} (1 - 3^{-x} - 3^{-y} + 3^{-(x+y)}) = 1 - 3^{-x}, \quad x \geq$$

0.

Отже,

$$F_1(x) = \begin{cases} 1 - 3^{-x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

$$F_2(y) = F(+\infty, y) = \lim_{x \rightarrow +\infty} (1 - 3^{-x} - 3^{-y} + 3^{-(x+y)}) = 1 - 3^{-y}, \quad y \geq 0.$$

Отже,

$$F_2(y) = \begin{cases} 1 - 3^{-y}, & y \geq 0; \\ 0, & y < 0. \end{cases}$$

$$F(x, y) = F_1(x) \cdot F_2(y) = (1 - 3^{-x}) \cdot (1 - 3^{-y}) = 1 - 3^{-x} - 3^{-y} + 3^{-(x+y)} = F(x, y), \quad x \geq 0; y \geq 0.$$

Таким чином, випадкові величини X і Y незалежні.

Відповідь. X і Y незалежні; $F_1(x) = 1 - 3^{-x}$; $F_2(y) = 1 - 3^{-y}$, $x, y \geq 0$.

Приклад 12.6. Двовимірна випадкова величина (X, Y) задана щільністю розподілу:

$$f(x, y) = \frac{1}{\pi^2 \cdot (1+x^2) \cdot (1+y^2)}, \quad x, y \in R.$$

Довести, що X і Y незалежні.

Розв'язання. Маємо за формулами (12.8) та (12.9):

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{+\infty} \frac{1}{\pi^2 \cdot (1+x^2) \cdot (1+y^2)} dy = \frac{1}{\pi^2 \cdot (1+x^2)} \cdot \\ &\cdot \int_{-\infty}^{+\infty} \frac{1}{1+y^2} dy = \frac{1}{\pi^2 \cdot (1+x^2)} \cdot \arctg(y) \Big|_{-\infty}^{+\infty} = \\ &= \frac{1}{\pi \cdot (1+x^2)} \cdot \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = \frac{1}{\pi \cdot (1+x^2)}; \\ f_2(y) &= \int_{-\infty}^{+\infty} \frac{1}{\pi^2 \cdot (1+x^2) \cdot (1+y^2)} dx = \frac{1}{\pi^2 \cdot (1+y^2)} \cdot \\ &\cdot \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx = \frac{1}{\pi^2 \cdot (1+y^2)} \cdot \arctg(x) \Big|_{-\infty}^{+\infty} = \\ &= \frac{1}{\pi \cdot (1+y^2)} \cdot \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = \frac{1}{\pi \cdot (1+y^2)}. \end{aligned}$$

Отже, оскільки

$$f_1(x) \cdot f_2(y) = \frac{1}{\pi^2 \cdot (1+x^2) \cdot (1+y^2)} = f(x, y),$$

то X і Y незалежні.

Відповідь. X і Y незалежні.

Залежність між випадковими величинами X і Y називається **ймовірнісною** (*стохастичною* або *статистичною*), якщо залежність між випадковими величинами означає аналітичну залежність щільності умовного розподілу однієї з них від значень, яких набуває друга величини. Виявляється вона не лише у зміні умовних законів розподілу, а й у зміні умовних числових характеристик.

12.4. Нормальний закон розподілу двовимірної випадкової величини

З усіх законів розподілу системи двох випадкових величин найбільше поширення на практиці має нормальний розподіл.

Розглянемо спочатку нормальний розподіл для системи двох незалежних випадкових величин.

Нехай X і Y – нормально розподілені і незалежні випадкові величини, а відповідні їм щільності розподілення мають вигляд:

$$f_1(x) = \frac{1}{\sigma_1 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a_1)^2}{2 \cdot \sigma_1^2}}; \quad f_2(y) = \frac{1}{\sigma_2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(y-a_2)^2}{2 \cdot \sigma_2^2}}.$$

Отже, щільність розподілу системи (X, Y) , на підставі теореми добутку щільностей розподілу для випадку незалежних величин отримаємо у вигляді (12.18):

$$f(x, y) = f_1(x) \cdot f_2(y) = \frac{1}{2 \cdot \pi \cdot \sigma_1 \cdot \sigma_2} \cdot e^{-\frac{1}{2} \cdot \left(\frac{(x-a_1)^2}{\sigma_1^2} + \frac{(y-a_2)^2}{\sigma_2^2} \right)}. \quad (12.18)$$

Двовимірна випадкова величина (X, Y) має **нормальний закон розподілу**, якщо її щільність розподілу має вигляд (12.18).

Якщо центр розсіювання системи збігається з початком координат, то $a_1 = a_2 = 0$, отже, отримаємо (12.19):

$$f(x, y) = \frac{1}{2 \cdot \pi \cdot \sigma_1 \cdot \sigma_2} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right)}. \quad (12.19)$$

Вираз (12.19) називається **канонічною формою нормального розподілу на площині**.

Для дослідження виду поверхні розподілу необхідно застосувати метод перерізів. Перетинаємо поверхню розподілу площинами, паралельними координатній площині XOY . Проектуючи переріз на цю координатну площину, ми отримаємо сімейство подібних і однаково розміщених еліпсів зі спільним центром в початку координат. Для того, щоб переконатися в цьому, запишемо рівняння лінії перетину поверхні розподілу площиною

$$z = z_0 = \text{const}.$$

Очевидно, що сталому значенню z_0 функції (12.18) відповідає стале значення степеня, тобто: $\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = k^2$ ($k = \text{const}$).

Це рівняння є рівнянням проекції на координатну площину XOY лінії перерізу поверхні розподілу площиною $z = z_0$. Перетворивши рівняння до

вигляду $\frac{x^2}{k^2 \cdot \sigma_1^2} + \frac{y^2}{k^2 \cdot \sigma_2^2} = 1$ побачимо, що воно є рівнянням еліпса, головні півосі якого пропорційні σ_1, σ_2 і збігаються, відповідно, з осями OX, OY , а центр знаходиться в початку координат. Оскільки k може змінюватися від нуля до нескінченності, то ми маємо сімейство подібних і однаково розміщених еліпсів. Кожний еліпс із цього сімейства є геометричним місцем точок, де щільність розподілу $f(x, y)$ дорівнює постійній величині. Тому вони називаються еліпсами однакової щільності або еліпсами розсіювання. Спільні осі симетрії всіх еліпсів розсіювання називаються головними осями розсіювання. При $\sigma_1 = \sigma_2$ еліпси перетворюються в кола і розподіли називають круговими.

Перетинаючи поверхню розподілу площинами, паралельними координатній площині YOZ , ми будемо отримувати криві, подібні кривим нормального розподілу. Таким чином, поверхня розподілу має вигляд горба, вершина якого знаходиться на осі OZ (рис. 12.5).

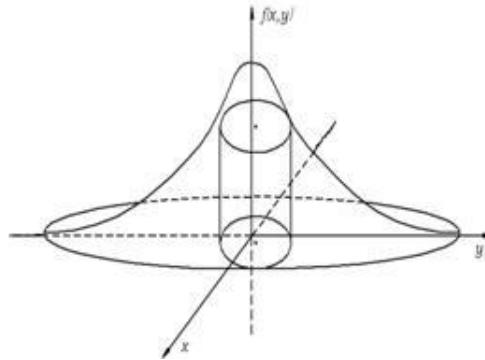


Рис. 12.5. Поверхня розподілу двовимірної нормальної випадкової величини (X, Y)

Поверхня розподілу (рис. 12.2) відрізняється від поверхні розподілу (рис. 12.4) тільки тим, що центр еліпсів розсіювання має координати (a_1, a_2) , а головні осі розсіювання паралельні, відповідно, осям координат, тобто поверхня розподілу (рис. 12.2) отримується паралельним перенесенням поверхні розподілу (рис. 12.4).

Обчислимо для розподілу (12.18) ймовірність попадання в прямокутник зі сторонами, паралельними осям координат. Прямокутник обмежений абсцисами a, b і ординатами c, d (рис. 12.3).

Застосовуючи загальну формулу для розрахунку ймовірності попадання випадкової точки в довільну область для вказаного випадку, запишемо початковий вираз для шуканої ймовірності у вигляді:

$$\begin{aligned}
P(a < X < b; c < Y < d) &= \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \\
&= \left(\int_a^b f_1(x) dx \right) \cdot \left(\int_c^d f_2(y) dy \right) = \\
&= \left(\int_a^b \left(\frac{1}{\sigma_1 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a_1)^2}{2 \cdot \sigma_1^2}} \right) dx \right) \cdot \\
&\cdot \left(\int_c^d \left(\frac{1}{\sigma_2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(y-a_2)^2}{2 \cdot \sigma_2^2}} \right) dy \right).
\end{aligned}$$

Для кожної з випадкових величин (X, Y) використаємо формулу (12.20) для обчислення ймовірностей за нормальним законом розподілу:

$$\begin{aligned}
P(a < X < b; c < Y < d) &= \left(\Phi \left(\frac{b-a_1}{\sigma_1} \right) - \Phi \left(\frac{a-a_1}{\sigma_1} \right) \right) \cdot \\
&\cdot \left(\Phi \left(\frac{d-a_2}{\sigma_2} \right) - \Phi \left(\frac{c-a_2}{\sigma_2} \right) \right), \quad (12.20)
\end{aligned}$$

де $\Phi(x)$ – функція Лапласа.

Якщо сторони прямокутника не паралельні осям координат (головним осям розсіювання), ця формула для ймовірності попадання в прямокутник не застосовуються.

Двовимірний нормальний розподіл допускає узагальнення на систему n випадкових величин $(X_1, X_2, X_3, \dots, X_n)$, а саме:

Якщо випадкові величини $(X_1, X_2, X_3, \dots, X_n)$ мають нормальний розподіл і незалежні між собою, то система випадкових величин $(X_1, X_2, X_3, \dots, X_n)$ має n -вимірний нормальний розподіл зі щільністю ймовірності (12.21):

$$f(X_1, X_2, X_3, \dots, X_n) = \frac{1}{(2 \cdot \pi)^{\frac{n}{2}} \cdot \sigma_1 \cdot \dots \cdot \sigma_n} \cdot e^{-\frac{1}{2} \left(\frac{(x_1-a_1)^2}{\sigma_1^2} + \dots + \frac{(x_n-a_n)^2}{\sigma_n^2} \right)}. \quad (12.21)$$

Розглянемо нормальний розподіл на площині для залежних величин. Щільність нормального розподілу для двовимірної випадкової величини (X, Y) записується формулою (12.22):

$$f(x, y) = \frac{1}{2 \cdot \pi \cdot \sigma_1 \cdot \sigma_2 \cdot \sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)} \left(\frac{(x-a_1)^2}{\sigma_1^2} - 2 \cdot \rho \cdot \frac{(x-a_1) \cdot (y-a_2)}{\sigma_1 \cdot \sigma_2} + \frac{(y-a_2)^2}{\sigma_2^2} \right)}, \quad (12.22)$$

де:

a_1, a_2 – математичні сподівання;

σ_1, σ_2 – середні квадратичні відхилення;

$\rho(X, Y)$ – коефіцієнт кореляції випадкових величин X і Y .

12.5. Числові характеристики двовимірної випадкової величини

Нехай (X, Y) – дискретна двовимірна випадкова величина.

Математичне сподівання, дисперсія і середнє квадратичне відхилення визначаються формулами (12.23)-(12.29):

$$M(X) = \sum_{i=1}^m (x_i \cdot p_i) = \sum_{i=1}^m \left(x_i \cdot \left(\sum_{j=1}^n p_{ij} \right) \right). \quad (12.23)$$

$$D(X) = M(X^2) - M^2(X) = \sum_{i=1}^m \left(x_i^2 \cdot \left(\sum_{j=1}^n p_{ij} \right) \right) - M^2(X). \quad (12.24)$$

$$\sigma(X) = \sqrt{D(X)}. \quad (12.25)$$

$$M(Y) = \sum_{j=1}^n (y_j \cdot q_j) = \sum_{j=1}^n \left(y_j \cdot \left(\sum_{i=1}^m p_{ij} \right) \right). \quad (12.26)$$

$$D(Y) = M(Y^2) - M^2(Y) = \sum_{j=1}^n \left(y_j^2 \cdot \left(\sum_{i=1}^m p_{ij} \right) \right) - M^2(Y). \quad (12.27)$$

$$\sigma(Y) = \sqrt{D(Y)}. \quad (12.28)$$

$$M(X, Y) = \sum_{i=1}^m \sum_{j=1}^n (x_i \cdot y_j \cdot p_{ij}). \quad (12.29)$$

Нехай (X, Y) – неперервна двовимірна випадкова величина з щільністю розподілу $f(x, y)$. Тоді справедливі формули (12.30)-(12.32):

$$M(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cdot f(x, y)) dx dy, \quad (12.30)$$

$$M(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y \cdot f(x, y)) dx dy, \quad (12.31)$$

$$M(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cdot y \cdot f(x, y)) dx dy. \quad (12.32)$$

Якщо $f(x, y)$ приймає значення в деякій області D площини OXY , то справедливі формули (12.30)-(12.35):

$$M(X) = \iint_D (x \cdot f(x, y)) dx dy, \quad (12.33)$$

$$M(Y) = \iint_D (y \cdot f(x, y)) dx dy, \quad (12.34)$$

$$M(X, Y) = \iint_D (x \cdot y \cdot f(x, y)) dx dy. \quad (12.35)$$

Аналогічно для дисперсії справедливі формули (12.36)-(12.37):

$$\begin{aligned} D(X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left((x - M(X))^2 \cdot f(x, y) \right) dx dy = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x^2 \cdot f(x, y)) dx dy - M^2(X), \end{aligned} \quad (12.36)$$

$$\begin{aligned} D(Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left((y - M(Y))^2 \cdot f(x, y) \right) dx dy = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y^2 \cdot f(x, y)) dx dy - M^2(Y). \end{aligned} \quad (12.37)$$

Якщо $f(x, y)$ приймає значення в деякій області D площини OXY , то справедливі формули (12.38)-(12.39):

$$D(X) = \iint_D \left((x - M(X))^2 \cdot f(x, y) \right) dx dy = \iint_D (x^2 \cdot f(x, y)) dx dy - M^2(X), \quad (12.38)$$

$$D(Y) = \iint_D \left((y - M(Y))^2 \cdot f(x, y) \right) dx dy = \iint_D (y^2 \cdot f(x, y)) dx dy - M^2(Y). \quad (12.39)$$

Точка $(M(X), M(Y))$ називається *центром розсіювання двовимірної випадкової величини*. Дисперсія випадкової величини (X, Y) характеризує розсіювання випадкової точки (X, Y) навколо точки $(M(X), M(Y))$ в напрямі OX і OY на площині OXY .

Для з'ясування наявності зв'язку та рівня залежності між випадковими величинами X і Y застосовують такі числові характеристики, як коваріацію та коефіцієнт кореляції.

Коваріацією (кореляційним моментом) випадкових величин X і Y називається число, яке обчислюється за формулою (12.40):

$$K(X, Y) = M((X - M(X)) \cdot (Y - M(Y))) = M(X, Y) - M(X) \cdot M(Y). \quad (12.40)$$

Нехай (X, Y) – дискретна двовимірна випадкова величина. Тоді справедлива формула (12.41):

$$K(X, Y) = \sum_{i=1}^m \sum_{j=1}^n (x_i \cdot y_j \cdot p_{ij}) - \sum_{i=1}^m (x_i \cdot p_i) \cdot \sum_{j=1}^n (y_j \cdot q_j). \quad (12.41)$$

Для неперервної двовимірної випадкової величини (X, Y) справедлива формула (12.42):

$$K(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cdot y \cdot f(x, y)) dx dy - \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cdot f(x, y)) dx dy \right) \cdot \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y \cdot f(x, y)) dx dy \right). \quad (12.42)$$

Якщо $f(x, y)$ приймає значення в деякій області D площини OXY , то справедлива формула (12.43):

$$K(X, Y) = \iint_D (x \cdot y \cdot f(x, y)) dx dy - \left(\iint_D (x \cdot f(x, y)) dx dy \right) \cdot \left(\iint_D (y \cdot f(x, y)) dx dy \right). \quad (12.43)$$

Коефіцієнтом кореляції між випадковими величинами X і Y називається число, що обчислюється за формулою (12.44):

$$\rho(X, Y) = \frac{K(X, Y)}{\sigma(X) \cdot \sigma(Y)}. \quad (12.44)$$

Для будь-яких випадкових величин $-1 \leq \rho(X, Y) \leq 1$.

Рівень корельованості між випадковими величинами встановлюється за допомогою шкали Чеддока (рис. 12.6).

Значення коефіцієнта кореляції (за модулем)	0,1 - 0,3	0,3 - 0,5	0,5 - 0,7	0,7 - 0,9	0,9 - 0,99
Характеристика сили зв'язку	слабкий	помірний	помітний	сильний	дуже сильний

Рис. 12.6. Шкала Чеддока виявлення сили зв'язку за значеннями парного коефіцієнту кореляції

Властивості

1. Коефіцієнт кореляції між випадковими величинами X і Y дорівнює ± 1 тоді і тільки тоді, коли X і Y зв'язані лінійною залежністю $Y = a \cdot X + b$, причому $\rho(X, Y) = 1$ при $a > 0$, $\rho(X, Y) = -1$ при $a < 0$;
2. Якщо випадкові величини X і Y незалежні, то $\rho(X, Y) = 0$. Якщо $\rho(X, Y) \neq 0$, то X і Y залежні випадкові величини.

Випадкові величини X і Y , для яких $\rho(X, Y) = 0$ називаються **некорельованими**.

Випадкові величини X і Y , для яких $\rho(X, Y) \neq 0$ називаються **корельованими**.

З некорельованості двох випадкових величин, взагалі кажучи, не випливає їх незалежність.

Для нормально розподілених випадкових величин некорельованість рівносильна незалежності.

Коефіцієнт кореляції служить для оцінки тісноти лінійного зв'язку між випадковими величинами X і Y : чим ближче модуль коефіцієнта кореляції до одиниці, тим зв'язок сильніший, чим ближче модуль коефіцієнта кореляції до нуля, тим зв'язок слабший.

Умовне математичне сподівання випадкової величини Y при $X = x$, тобто $M(Y/X) = \varphi(x)$ називається **функцією регресії випадкової величини Y відносно X** , або **регресією Y по X** .

Аналогічно $M(Y/X) = \varphi(y)$ називається **функцією регресії X відносно Y** , або **регресією X по Y** .

Регресія Y по X показує, як змінюється в середньому величина Y при зміні величини X .

Якщо обидві функції $\varphi(x)$ і $\varphi(y)$ лінійні, то X і Y зв'язані лінійною кореляційною залежністю.

Функція регресії може бути використана для прогнозування однієї з випадкових величин, якщо відоме значення іншої.

Приклад 12.7. Задано закон розподілу системи двох дискретних випадкових величин (X, Y) (табл. 12.12):

Таблиця 12.12. Ряд розподілу системи випадкових величин (X, Y)

$Y = y_j \backslash X = x_i$	5	6	7	$\sum$
2	0,1	0,1	0,2	0,4
4	0,01	0,02	0,17	0,2
6	0,3	0,05	0,05	0,4
$\sum$	0,41	0,17	0,42	1

Обчислити $M(X); D(X); \sigma(X); M(Y); D(Y); \sigma(Y); K(X, Y); \rho(X, Y); P(2 < X \leq 6; 5 \leq Y < 7)$.

Розв'язання. Числові характеристики обчислюємо за формулами (12.23)-(12.29):

$$M(X) = \sum_{i=1}^m (x_i \cdot p_i) = 5 \cdot 0,3 + 6 \cdot 0,5 + 7 \cdot 0,2 = 5,9;$$

$$M(X^2) = \sum_{i=1}^m (x_i^2 \cdot p_i) = 5^2 \cdot 0,3 + 6^2 \cdot 0,5 + 7^2 \cdot 0,2 = 35,3;$$

$$D(X) = M(X^2) - M^2(X) = 35,3 - 34,81 = 0,49;$$

$$\sigma(X) = \sqrt{D(X)} = 0,7;$$

$$M(Y) = \sum_{j=1}^n (y_j \cdot q_j) = 2 \cdot 0,4 + 4 \cdot 0,2 + 6 \cdot 0,4 = 4;$$

$$M(Y^2) = \sum_{j=1}^n (y_j^2 \cdot q_j) = 2^2 \cdot 0,4 + 4^2 \cdot 0,2 + 6^2 \cdot 0,4 = 19,2;$$

$$D(Y) = M(Y^2) - M^2(Y) = 19,2 - 16 = 3,2;$$

$$\sigma(Y) = \sqrt{D(Y)} \approx 1,79.$$

$$M(X, Y) = \sum_{i=1}^m \sum_{j=1}^n (x_i \cdot y_j \cdot p_{ij}) =$$

$$= 2 \cdot 5 \cdot 0,1 + 2 \cdot 6 \cdot 0,1 + 2 \cdot 7 \cdot 0,2 + 4 \cdot 5 \cdot 0,01 + 4 \cdot 6 \cdot 0,02 + 4 \cdot 7 \cdot 0,17 + 6 \cdot 5 \cdot 0,3 + 6 \cdot 6 \cdot 0,3 + 6 \cdot 7 \cdot 0,05 = 23,16;$$

$$K(X, Y) = M(X, Y) - M(X) \cdot M(Y) = 23,16 - 5,9 \cdot 4 = -0,44.$$

Для вимірювання тісноти кореляційного зв'язку обчислимо коефіцієнт кореляції

$$\rho(X, Y) = \frac{K(X, Y)}{\sigma(X) \cdot \sigma(Y)} = \frac{-0,44}{0,7 \cdot 1,79} \approx -0,35.$$

Випадкові величини X і Y залежні. Згідно шкали Чеддока (рис. 12.6) залежність між компонентами системи (X, Y) є помірною.

Знайдемо $P(5 < X \leq 7, 2 \leq Y < 7) = 0,1 + 0,02 + 0,2 + 0,17 = 0,49$.

Відповідь. $M(X) = 5,9; D(X) = 0,49; \sigma(X) = 0,7; M(Y) = 4; D(Y) = 3,2; \sigma(Y) \approx 1,79; K(X, Y) = -0,44; \rho(X, Y) \approx -0,35; P(2 < X \leq 6; 5 \leq Y < 7) = 0,49$.

Приклад 12.8. Неперервна двовимірна випадкова величина (X, Y) задана щільністю розподілу:

$$f(x, y) = \begin{cases} A \cdot x \cdot y, & (x, y) \in D, \\ 0, & (x, y) \notin D, \end{cases}$$

де область D – трикутник, обмежений прямими $x + y - 1 = 0, x = 0, y = 0$. Знайти число $A; M(X); D(X); \sigma(X); M(Y); D(Y); \sigma(Y); K(X, Y); \rho(X, Y)$.

Розв'язання. Оскільки $f(x, y)$ приймає значення в області D (рис. 12.7), то

$$\iint_D f(x, y) dx dy = 1.$$

Одержимо:

$$\begin{aligned} \iint_D (A \cdot x \cdot y) dx dy &= A \cdot \int_0^1 \left(x \cdot \left(\int_0^{1-x} y dy \right) \right) dx = \frac{A}{2} \cdot \int_0^1 (x - 2 \cdot x^2 + x^3) dx \\ &= \frac{A}{2} \cdot \left(\frac{x^2}{2} - \frac{2 \cdot x^3}{3} + \frac{x^4}{4} \right) \Big|_0^1 = \frac{A}{24} = 1. \end{aligned}$$

Звідси $A = 24$.

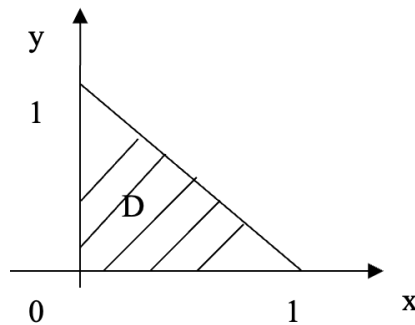


Рис. 12.7. Зображення області D

Отже,

$$f(x, y) = \begin{cases} 24 \cdot x \cdot y, & (x, y) \in D, \\ 0, & (x, y) \notin D. \end{cases}$$

За формулами (12.33)-(12.39), (12.43), (12.44)

$$\begin{aligned} M(X) &= \iint_D x \cdot f(x, y) dx dy = \\ &= 24 \cdot \int_0^1 \left(x^2 \cdot \left(\int_0^{1-x} y dy \right) \right) dx \\ &= 12 \cdot \int_0^1 (x^2 - 2 \cdot x^3 + x^4) dx = 12 \cdot \left(\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5} \right) \Big|_0^1 = \frac{2}{5}; \\ M(Y) &= \iint_D y \cdot f(x, y) dx dy = 24 \cdot \int_0^1 \left(y^2 \cdot \left(\int_0^{1-y} x dx \right) \right) dy = \frac{2}{5}; \end{aligned}$$

$$\begin{aligned}
M(X^2) &= \iint_D x^2 \cdot f(x, y) dx dy = \\
&= 24 \cdot \\
&\cdot \int_0^1 \left(x^3 \cdot \left(\int_0^{1-x} y dy \right) \right) dx \\
&= 12 \cdot \int_0^1 (x^3 - 2 \cdot x^4 + x^5) dx = 12 \cdot \left(\frac{x^4}{4} - \frac{2 \cdot x^5}{5} + \frac{x^6}{6} \right) \Big|_0^1 = \frac{1}{5};
\end{aligned}$$

$$M(Y^2) = \iint_D y^2 \cdot f(x, y) dx dy = 24 \cdot \int_0^1 \left(y^3 \cdot \left(\int_0^{1-y} x dx \right) \right) dy = \frac{1}{5};$$

$$D(X) = \frac{1}{5} - \frac{4}{25} = \frac{1}{25}, \sigma(X) = \frac{1}{5};$$

$$D(Y) = \frac{1}{5} - \frac{4}{25} = \frac{1}{25}, \sigma(Y) = \frac{1}{5};$$

$$\begin{aligned}
M(X, Y) &= \iint_D x \cdot y \cdot f(x, y) dx dy \\
&= 24 \cdot \\
&\cdot \int_0^1 \left(x^2 \cdot \left(\int_0^{1-x} y^2 dy \right) \right) dx \\
&= 8 \cdot \\
&\cdot \int_0^1 (x^2 - 3 \cdot x^3 + 3 \cdot x^4 - x^5) dx \\
&= 8 \cdot \left(\frac{x^3}{3} - \frac{3 \cdot x^4}{4} + \frac{3 \cdot x^5}{5} - \frac{x^6}{6} \right) \Big|_0^1 = \frac{2}{15};
\end{aligned}$$

$$K(X, Y) = M(X, Y) - M(X) \cdot M(Y) = \frac{2}{15} - \frac{2}{5} \cdot \frac{2}{5} = -\frac{2}{75};$$

$$\rho(X, Y) = \frac{K(X, Y)}{\sigma(X) \cdot \sigma(Y)} = -\frac{2}{3}.$$

Згідно шкали Чеддока (рис. 12.6) кореляційний зв'язок між компонентами системи (X, Y) є помітним.

Відповідь. $A = 24$; $M(X) = \frac{2}{5}$; $D(X) = \frac{1}{5}$; $\sigma(X) = \frac{1}{\sqrt{5}}$; $M(Y) = \frac{2}{5}$; $D(Y) = \frac{1}{5}$; $\sigma(Y) = \frac{1}{\sqrt{5}}$; $K(X, Y) = -\frac{2}{75}$; $\rho(X, Y) = -\frac{2}{3}$.

Приклад 12.9. Двовимірна випадкова величина (X, Y) задана щільністю розподілу:

$$f(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 \leq 1, \\ 0, & x^2 + y^2 > 1, \end{cases}$$

Обчислити коефіцієнт кореляції та зробити висновок про корельованість випадкових величин X, Y .

Розв'язання. З'ясуємо питання про залежність X, Y .

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{\pi} \cdot \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy = \frac{2}{\pi} \cdot \sqrt{1-x^2}, x \in [-1; 1];$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \frac{1}{\pi} \cdot \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx = \frac{2}{\pi} \cdot \sqrt{1-y^2}, y \in [-1; 1].$$

$$\frac{1}{\pi} \neq \frac{4}{\pi^2} \cdot \sqrt{(1-x^2) \cdot (1-y^2)}, x^2 + y^2 \leq 1,$$

отже $f(x, y) \neq f_1(x) \cdot f_2(y)$. Таким чином, X, Y – залежні.

Знайдемо чисельні характеристики:

$$M(X) = \int_{-\infty}^{+\infty} (x \cdot f_1(x)) dx = \frac{2}{\pi} \cdot \int_{-1}^1 (x \cdot \sqrt{1-x^2}) dx = -\frac{1}{\pi} \cdot \int_{-1}^1 \sqrt{1-x^2} d(1-x^2) = 0;$$

$$M(Y) = \int_{-\infty}^{+\infty} (y \cdot f_2(y)) dy = \frac{2}{\pi} \cdot \int_{-1}^1 (y \cdot \sqrt{1-y^2}) dy = -\frac{1}{\pi} \cdot \int_{-1}^1 \sqrt{1-y^2} d(1-y^2) = 0,$$

оскільки у кожному випадку підінтегральна функція непарна і проміжок інтегрування симетричний відносно початку координат.

$$M(X, Y) = \frac{1}{\pi} \cdot \int_{-1}^1 x dx \cdot \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} y dx = 0,$$

бо у внутрішньому інтегралі підінтегральна функція непарна і проміжок інтегрування симетричний відносно початку координат.

Таким чином:

$$K(X, Y) = M(X, Y) - M(X) \cdot M(Y) = 0.$$

Отже, $\rho(X, Y) = 0$, тобто випадкові величини X і Y некорельовані і залежні.

Відповідь. $\rho(X, Y) = 0$, випадкові величини X і Y некорельовані і залежні.

Приклад 12.10. Двовимірний випадковий величини (X, Y) задана матрицею розподілу (табл. 12.13):

Таблиця 12.13. Матриця розподілу величини (X, Y)

$X \backslash Y$	1	2	3	$\sum$
0	0,1	0,1	0,1	0,3
1	0,1	0,2	0,1	0,4
2	0,1	0,1	0,1	0,3
$\sum$	0,3	0,4	0,3	1

Обчислити коефіцієнт кореляції та зробити висновок про корельованість випадкових величин X, Y .

Розв'язання. З'ясуємо питання про залежність X, Y .

$$p_0 = 0,3; q_1 = 0,3, \text{ тобто } p_0 \cdot q_1 = 0,09; \text{ а } p_{01} = 0,1.$$

Тобто $p_{01} \neq p_0 \cdot q_1$. Отже X, Y – залежні.

Знайдемо чисельні характеристики:

$$M(X) = 0,4 + 0,6 = 1;$$

$$M(Y) = 0,3 + 0,8 + 0,9 = 2;$$

$$M(X, Y) = 0,1 + 0,4 + 0,3 + 0,2 + 0,4 + 0,6 = 2.$$

Таким чином:

$$K(X, Y) = M(X, Y) - M(X) \cdot M(Y) = 2 - 1 \cdot 2 = 0.$$

Отже, $\rho(X, Y) = 0$, тобто випадкові величини X і Y некорельовані і незалежні.

Відповідь. $\rho(X, Y) = 0$, випадкові величини X і Y некорельовані і незалежні.

ЛЕКЦІЯ 13

КОЕФІЦІЄНТИ КОРЕЛЯЦІЇ

13.1. Коефіцієнт вибіркової кореляції (коефіцієнт кореляції Пірсона)

Розглянемо статистичні вибірки об'єму n , які позначимо $X; Y$:

$$X = \{x_1, \dots, x_n\}; Y = \{y_1, \dots, y_n\}.$$

Частоти кожного елемента вибірок дорівнюють 1.

Незміщеними оцінками кореляційного моменту $K(X; Y)$ і дисперсій $\sigma^2(X)$ і $\sigma^2(Y)$ (лекція 12) є відповідно вибіркового кореляційного моменту (вибіркова коваріація) $K_B(X; Y)$ і виправлені вибіркові дисперсії, які обчислюються за формулою (13.1):

$$s^2(X) = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}_B)^2 \text{ і } s^2(Y) = \frac{1}{n-1} \cdot \sum_{i=1}^n (y_i - \bar{y}_B)^2. \quad (13.1)$$

Тоді, якщо відомі $(x_1, y_1); (x_2, y_2), \dots, (x_n, y_n)$ – фактичні значення змінних X і Y , то вибіркова коваріація між ними розраховується наступним чином (13.2):

$$K_B(X; Y) = \frac{1}{n-1} \cdot \sum_{i=1}^n ((x_i - \bar{x}_B) \cdot (y_i - \bar{y}_B)), \quad (13.2)$$

де $\bar{x}_B = \frac{\sum_{i=1}^n x_i}{n}$; $\bar{y}_B = \frac{\sum_{i=1}^n y_i}{n}$ – середні арифметичні відповідних величин.

Видно, що у виразі для $K_B(X; Y)$ присутній один зв'язок (добуток середніх), тому для нього, як і для вибірових дисперсій, число ступенів свободи дорівнює $n - 1$.

Той факт, що для отримання незміщеної оцінки s^2 в знаменнику вибіркової дисперсії довелося n замінити на $n - 1$, безпосередньо пов'язаний з тим, що величина відносно якої беруться відхилення, сама залежить від елементів вибірки. Якщо б у формулі вибіркової дисперсії були дві такі величини, то n потрібно було б замінити на $n - 2$ і т.д.

Кожна величина, яка залежить від елементів вибірки і є у формулі вибіркової дисперсії, називається *зв'язком*. Виявляється, що знаменник вибіркової дисперсії завжди дорівнює різниці між об'ємом вибірки n і числом зв'язків, накладених на цю вибірку. Число $k = n - l$ буде вважатися *числом ступенів свободи*.

Вибірковий коефіцієнт кореляції (коефіцієнт кореляції Пірсона) розраховується за формулою (13.3):

$$\rho_{\Pi}(X, Y) = \rho_B(X, Y) = \frac{K_B(X; Y)}{\sqrt{s^2(X) \cdot s^2(Y)}} = \frac{\sum_{i=1}^n (x_i - \bar{x}_B) \cdot (y_i - \bar{y}_B)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_B)^2 \cdot \sum_{i=1}^n (y_i - \bar{y}_B)^2}}. \quad (13.3)$$

Для зручності обчислень на практиці будемо використовувати формулу (13.4):

$$\rho_{\text{П}}(X, Y) = \frac{\sum_{i=1}^n (x_i \cdot y_i) - \frac{(\sum_{i=1}^n x_i) \cdot (\sum_{i=1}^n y_i)}{n}}{\sqrt{\left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right) \cdot \left(\sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n} \right)}}. \quad (13.4)$$

Дану формулу легко отримати з формули (13.3) розкривши дужки та замінивши: $\bar{x}_B = \frac{\sum_{i=1}^n x_i}{n}$; $\bar{y}_B = \frac{\sum_{i=1}^n y_i}{n}$.

При цьому оцінка дисперсії характеризує ступінь розкиду значень навколо їх середнього або варіабельність та визначається за формулою (13.5):

$$s^2(X) = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}_B)^2. \quad (13.5)$$

В загальному випадку для отримання незміщеної оцінки дисперсії суму квадратів необхідно поділити на число ступенів свободи. Так як вибірка вже використовувалася один раз для визначення вибіркового середнього вибірки X , то число накладених зв'язків в даному випадку дорівнює одиниці, а число ступенів свободи $k = n - 1$.

Однак, більш природно вимірювати ступінь розкиду значень змінних в тих же одиницях, в яких вимірюється і сама змінна. Цю задачу вирішує показник, що називається *середньоквадратичним відхиленням (стандартним відхиленням)* або *стандартною похибкою* і визначається співвідношенням (13.6):

$$s(X) = \sqrt{s^2(X)}. \quad (13.6)$$

Коефіцієнт кореляції Пірсона намагається встановити лінію, яка найкраще допасовується до набору даних із двох змінних, по суті викладаючи очікувані значення, а отриманий коефіцієнт кореляції Пірсона вказує, наскільки далеким від очікуваних значень є фактичний набір даних. Залежно від знаку нашого коефіцієнта кореляції Пірсона ми можемо отримати як від'ємну, так і додатну кореляцію, якщо якийсь зв'язок між змінними нашого набору даних існує (рис. 13.1).

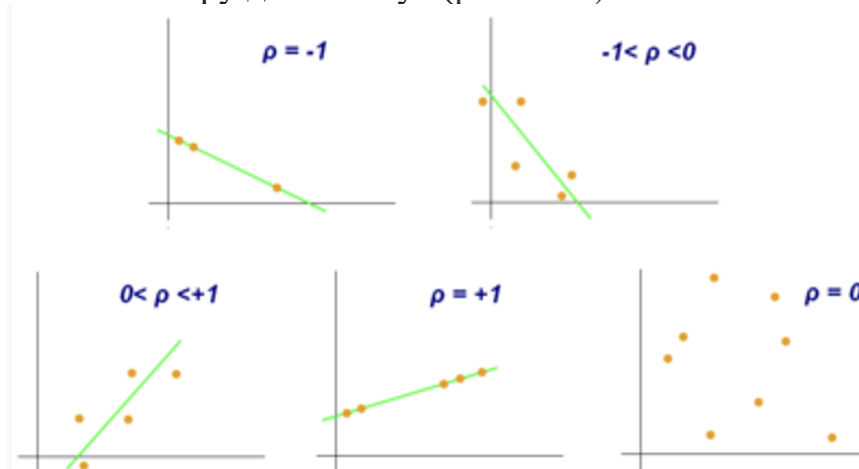


Рис. 13.1. Приклади діаграм розсіювання з різними значеннями коефіцієнту кореляції Пірсона

Під час інтерпретації коефіцієнта кореляції Пірсона необхідно враховувати такі моменти:

1. Коефіцієнт Пірсона можна використовувати для різних шкал (шкала відношень, інтервальна або порядкова) за винятком дихотомічної шкали.

Для якісної оцінки коефіцієнтів кореляції різні шкали, найбільш часто застосовується шкала Чеддока (табл. 13.1). Згідно неї в залежності від значень модуля коефіцієнта кореляції зв'язок може мати одну з оцінок:

Таблиця 13.1. Шкала Чеддока виявлення сили зв'язку за значенням парного коефіцієнта кореляції

Значення коефіцієнта кореляції (за модулем)	0,1 – 0,3	0,3 – 0,5	0,5 – 0,7	0,7 – 0,9	0,9 – 0,99
Характеристика сили зв'язку	слабкий	помірний	помітний	сильний	дуже сильний

2. Кореляційний зв'язок далеко не завжди означає зв'язок причинно-наслідковий. Інакше кажучи, якщо ми знайшли, припустімо, позитивну кореляцію між зростом і вагою в групі випробовуваних, то це зовсім не означає, що зріст залежить від ваги або навпаки (обидві ці ознаки залежать від третьої (зовнішньої) змінної, яка в цьому разі пов'язана з генетичними конституціональними особливостями людини).

3. $\rho_{\Pi}(X; Y) = 0$ може спостерігатися не тільки за відсутності зв'язку між x і y , а й у разі сильного нелінійного зв'язку (рис. 13.2 в).

У цьому разі від'ємна і позитивна кореляції врівноважуються і в результаті створюється ілюзія відсутності зв'язку.

За наявності прямого (позитивного) зв'язку між змінними (рис. 13.2 а) хмара розсіювання має більш-менш сплюснену еліптичну форму, довга вісь якої спрямована вправо і вгору. Іншими словами, у разі зростання значення однієї змінної є тенденція до збільшення іншої змінної.

У разі негативного зв'язку між змінними (рис. 13.2 б) довга вісь хмари розсіювання спрямована праворуч донизу, тобто збільшення значень однієї змінної відповідає закономірному зниженню значень іншої.

Зрештою, якщо хмара розсіювання має округлу форму (рис. 13.2 в), то можна припустити, що кореляція між змінними відсутня або, принаймні, вона дуже незначна.

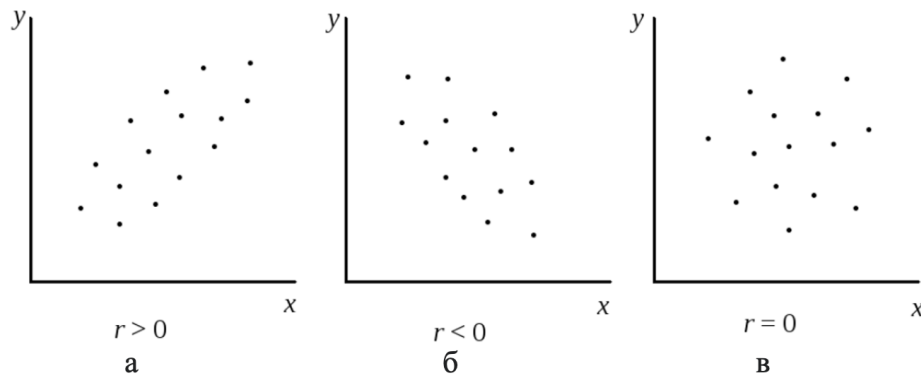


Рис. 13.2. Графічне представлення зв'язку між змінними (хмара розсіювання точок має різну форму залежно від характеру зв'язку)

Варто відмітити, що величина коефіцієнта кореляції не є доказом того, що між ознаками, що досліджуються, є причинно-наслідковий зв'язок, а являє собою оцінку ступеня взаємної узгодженості в змінах ознак. Для того, щоб встановити причинно-наслідкову залежність необхідним є аналіз якісної природи явищ.

Оскільки оцінка тісноти зв'язку за допомогою коефіцієнта кореляції проводиться, як правило, на основі більш чи менш обмеженої інформації про явище, що досліджується, то виникає питання: наскільки правомірним є висновок по вибірковим даним про наявність кореляційного зв'язку в тій генеральній сукупності, з якої була отримана вибірка?

Нехай X і Y мають нормальний розподіл. У цьому випадку при досить великому об'ємі вибірки n коефіцієнт $\rho_{\Pi}(X; Y)$ наближено дорівнює генеральному коефіцієнту $\rho(X; Y)$. Проте оцінити похибку, яка виникає при цьому, дуже важко. Це і не обов'язково, оскільки точне значення $\rho(X; Y)$ в розрахунках практично не використовується, а треба лише як показник наявності кореляції між Y і X . Вибірковий коефіцієнт кореляції $\rho_{\Pi}(X; Y)$ застосовується в основному для перевірки загальної гіпотези про наявність кореляції між спостережуваними величинами, не вдаючись у детальні оцінки сили цієї кореляції.

Вплив перетворення даних на $\rho_{\Pi}(X; Y)$:

1. Лінійні перетворення x і y типу $b \cdot x + a$ і $d \cdot y + c$ не змінять величину кореляції між x і y .
2. Лінійні перетворення x і y у разі $b < 0, d > 0$, а також у разі $b > 0$ і $d < 0$ змінюють знак коефіцієнта кореляції, не змінюючи його величини.

Достовірність (або, інакше, статистична значущість) коефіцієнта кореляції Пірсона може бути визначена різними способами:

- За таблицями критичних значень коефіцієнтів кореляції Пірсона. Якщо отримане в розрахунках значення $\rho_{\Pi}(X; Y)$ перевищує критичне (табличне) значення для даної вибірки, коефіцієнт Пірсона вважається статистично значущим. Число ступенів свободи в цьому випадку відповідає $k = n - 2$, де n – число пар порівнюваних значень (об'єм вибірки);

- За таблицями критичних значень коефіцієнтів кореляції Пірсона. Якщо отримане в розрахунках значення $\rho_{\Pi}(X; Y)$ перевищує критичне (табличне) значення для даної вибірки, коефіцієнт Пірсона вважається статистично значущим. Число ступенів свободи в цьому випадку відповідає $k = n - 2$, де n – число пар порівнюваних значень (об'єм вибірки);
- За коефіцієнтом Стюдента, який обчислюється як відношення коефіцієнта кореляції до його похибки.

Похибка коефіцієнта кореляції обчислюється за формулою (13.7):

$$\delta = \sqrt{\frac{1 - \rho_{\Pi}^2(X; Y)}{n - 2}}, \quad (13.7)$$

де:

δ – помилка коефіцієнта кореляції;

$\rho_{\Pi}(X; Y)$ – коефіцієнт кореляції Пірсона;

n – кількість порівнюваних пар.

Приклад 13.1. Нехай є два набори даних:

$$X = \{1; 2; 3; 4\}; Y = \{2; 1; 4; 5\}.$$

Встановити силу кореляційного зв'язку між ними.

Розв'язання. 1. Для обчислення коефіцієнта кореляції Пірсона отримаємо наступну розрахункову таблицю (табл. 13.2):

Таблиця 13.2. Розрахункова таблиця

	x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
	1	2	1	4	2
	2	1	4	1	2
	3	4	9	16	12
	4	5	16	25	20
Σ	10	12	30	46	36

2. Проводимо обчислення та підставляємо значення в формулу (13.4):

$$\Sigma x_i = 10; \Sigma x_i^2 = 30; \Sigma y_i = 12; \Sigma y_i^2 = 46; \Sigma(x_i \cdot y_i) = 36.$$

$$\rho_{\Pi}(X; Y) = \frac{36 - \frac{10 \cdot 12}{4}}{\sqrt{\left(30 - \frac{10^2}{4}\right) \cdot \left(46 - \frac{12^2}{4}\right)}} = \frac{6}{\sqrt{5 \cdot 10}} \approx 0,85.$$

3. Згідно шкали Чеддока між наборами даних є сильний додатний кореляційний зв'язок.

Відповідь. Між наборами даних є сильний додатний кореляційний зв'язок.

Приклад 13.2. 22 старшокласники були протестовані за двома тестами: РСК (рівень суб'єктивного контролю) і МдоУ (мотивація до успіху). Отримано такі результати (табл. 13.3):

Таблиця 13.3. Дані тестувань старшокласників

№	РСК (x_i)	МдоУ (y_i)	№	РСК (x_i)	МдоУ (y_i)
1	27	18	12	24	12
2	24	19	13	27	15
3	27	16	14	25	15
4	30	13	15	37	23
5	25	17	16	35	24
6	18	13	17	25	20
7	28	19	18	22	14
8	31	19	19	26	21
9	31	10	20	34	24
10	30	24	21	25	17
11	18	13	22	31	17

Перевірити гіпотезу про те, що для людей із високим рівнем інтернальності (бал РСК) характерний високий рівень мотивації до успіху (бал МдоУ).

Розв'язання. 1. Для обчислення коефіцієнта кореляції Пірсона отримаємо наступну розрахункову таблицю (табл. 13.4):

Таблиця 13.4. Розрахункова таблиця

x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
x_1	y_1	x_1^2	y_1^2	$x_1 \cdot y_1$
x_2	y_2	x_2^2	y_2^2	$x_2 \cdot y_2$
...	...	...	...	...
x_{22}	y_{22}	x_{22}^2	y_{22}^2	$x_{22} \cdot y_{22}$

$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum (x_i \cdot y_i)$

2. Проводимо обчислення та підставляємо значення в формулу (13.4):
 $\sum x_i = 600$; $\sum x_i^2 = 16864$; $\sum y_i = 383$; $\sum y_i^2 = 7025$; $\sum (x_i \cdot y_i) = 10689$.

$$\rho_{\text{П}}(X; Y) = \frac{10689 - \frac{600 \cdot 383}{22}}{\sqrt{\left(16864 - \frac{600^2}{22}\right) \cdot \left(7025 - \frac{383^2}{22}\right)}} \approx 0,58.$$

3. Визначимо статистичну значущість коефіцієнта кореляції Пірсона. Для цього обчислюємо похибку коефіцієнт Стюдента, як відношення коефіцієнта кореляції Пірсона до похибки.

$$\delta = \sqrt{\frac{1 - (0,58)^2}{22}} \approx 0,18; t_{\text{сп}} = \frac{\rho_{\text{П}}(X; Y)}{\delta} \approx 3,22.$$

У дод. 6 Гмурмана знаходимо стандартні значення коефіцієнта Стюдента для 1-го, 2-го і 3-го рівнів значущості ($\alpha = 0,05; 0,01; 0,001$) за числа ступенів свободи $k = n - 2 = 20$: $t_{\text{кр}} = 2,09; 2,85; 3,85$.

Відповідь: Кореляція між показниками тестів РСК і МдоУ є статистично значущою для 1-го і 2-го рівнів значущості.

13.2. Коефіцієнти рангової кореляції

Коефіцієнти рангової кореляції, такі як коефіцієнт рангової кореляції Спірмена та коефіцієнт рангової кореляції Кендалла (τ), вимірюють, до якої міри в разі збільшення однієї змінної інша змінна схильна збільшуватися, не вимагаючи, щоби це збільшення було подано лінійною залежністю. Якщо за збільшення однієї змінної інша зменшується, то коефіцієнти рангової кореляції будуть від'ємними. Ці коефіцієнти рангової кореляції часто розглядають як альтернативи коефіцієнту Пірсона, які використовують або для зменшення кількості обчислень, або для того, щоби зробити коефіцієнт менш чутливим до не нормальності в розподілах. Проте ця точка зору має мало математичних підстав, оскільки коефіцієнти рангової кореляції вимірюють інший тип зв'язку, ніж коефіцієнт кореляції Пірсона, і їх найкраще розглядати як показники іншого типу зв'язку, а не як альтернативну міру генерального коефіцієнту кореляції.

Щоб унаочнити природу рангової кореляції та її відмінність від лінійної кореляції, розглянемо наступні чотири пари чисел $(x; y)$: $(0; 1), (10; 100), (101; 500), (102; 2000)$.

В міру просування від кожної пари до наступної x збільшується, й те саме робить y . Цей взаємозв'язок ідеальний, у тому сенсі, що збільшення в x завжди супроводжується збільшенням в y . Це означає, що ми маємо ідеальну рангову кореляцію, й обидва коефіцієнти кореляції Спірмена та Кендалла дорівнюють 1, тоді як у цьому прикладі коефіцієнт кореляції Пірсона дорівнює 0,7544, вказуючи на те, що точки далеко не лежать на одній прямій. Так само, якщо y завжди зменшується, коли x збільшується,

коефіцієнти рангової кореляції становитимуть -1 , тоді як коефіцієнт кореляції Пірсона може бути або не бути близьким до -1 , залежно від того, наскільки близько до прямої лінії розташовані ці точки. Хоча в граничних випадках ідеальної рангової кореляції ці два коефіцієнти рівні (чи то обидва $+1$, чи обидва -1), зазвичай це не так, і тому значення цих двох коефіцієнтів неможливо порівнювати змістовно. Наприклад, для трьох пар $(1,1)$ $(2,3)$ $(3,2)$ коефіцієнт Спірмена дорівнює $\frac{1}{2}$, а коефіцієнт Кендалла дорівнює $\frac{1}{3}$.

Коефіцієнт рангової кореляції Спірмена

У багатьох випадках результати спостережень подаються не у вигляді кількісних вимірювань, а у вигляді бальних оцінок (рангів). Наприклад, студенти у групі можуть бути впорядковані по номерам за середнім балом в сесії, країни – за кількістю населення, учасники конкурсу – за зайнятим місцем тощо. При цьому інколи виникає можливість упорядкувати об'єкти дослідження за двома або більше показниками. У зв'язку з цим виникає задача дослідження кореляції цих показників.

Нехай n об'єктів дослідження, розташованих за рівнем якості, характеризуються парами рангів $(x_i; y_i)$ $i = 1, 2, \dots, n$. Потрібно з'ясувати рівень кореляції між двома ознаками, x та y . Для цього використовують *коефіцієнт рангової кореляції Спірмена*. Цей показник розраховують за формулою (13.8):

$$\rho_B(X, Y) = \rho_C(X, Y) = 1 - \frac{6 \cdot \sum_{i=1}^n d_i^2}{n \cdot (n^2 - 1)}, \quad (13.8)$$

де $d_i = x_i - y_i$ – різниця рангів для i -го об'єкту спостереження.

Це значення можна використати як наближення коефіцієнта кореляції, він є менш точним у порівнянні зі коефіцієнтом кореляції Пірсона, оскільки при його розрахунку не враховуються кількісні значення характеристик об'єктів, а лише їх порядок. Як і для коефіцієнта кореляції $\rho_P(X, Y)$, значення $\rho_C(X, Y)$ змінюється від -1 до 1 . Чим ближче абсолютне значення коефіцієнта рангової кореляції до одиниці, тим більш щільним є зв'язок між факторами.

У більш загальному випадку, якщо у вибірках X, Y присутні зв'язні ранги, тобто ранги елементів вибірки, які повторюються, коефіцієнт рангової кореляції Спірмена обчислюється за формулами (13.9) та (13.10):

$$\rho_C(X, Y) = 1 - \frac{6 \cdot \sum d_i^2 + T_X + T_Y}{n \cdot (n^2 - 1)}, \quad (13.9)$$

де

$$T_X = \frac{N_X^3 - N_X}{12} \text{ та } T_Y = \frac{N_Y^3 - N_Y}{12}. \quad (13.10)$$

В формулі (13.10) N_X та N_Y , відповідно, кількість рангів, що повторюються, у вибірках X та Y .

Для вибірки, яка не має повторюваних елементів, відповідний коефіцієнт із формули (13.9) дорівнює нулю, тобто отримуємо формулу (13.8).

Кроки для обчислення коефіцієнта кореляції Спірмена

Впорядкування даних: Присвоїти ранги кожному набору даних X і Y . Найменшому значенню присвоюється ранг 1, наступному — 2 і так далі. Якщо є однакові значення (збіги), їм присвоюється середній ранг (між наступними двома номерами рангів).

Обчислення різниць рангів: Для кожної пари даних $(x_i; y_i)$ знайти різницю між рангами d_i .

Зведення різниць до квадрату: Обчислити квадрат різниць рангів для кожної пари d_i^2 .

Обчислення коефіцієнтів T_X ; T_Y : В разі наявності зв'язних рангів, дані коефіцієнти обчислюються за формулою (13.9).

Обчислення коефіцієнта кореляції Спірмена: Підставити знайдені значення в формулу (13.7) і обчислити коефіцієнт.

Застосування

- Коефіцієнт рангової кореляції Спірмена використовується тоді, коли:
- Дані не мають нормального розподілу або містять викиди,
- Зв'язок між змінними не є лінійним, але очікується монотонний,
- Дані представлені у вигляді ранжувань.

Приклад 13.3. Нехай є два набори даних:

$$X = \{85; 95; 80; 70; 60\}; Y = \{80; 60; 90; 85; 70\}.$$

Встановити степінь кореляційного зв'язку між ними.

Розв'язання. Розставляємо ранги в обох наборах даних від найменшого до найбільшого.

Ранги для X : $R(X) = \{4; 5; 3; 2; 1\}$.

Ранги для Y : $R(Y) = \{3; 1; 5; 4; 2\}$.

Різниці рангів d_i : $d_i = \{1; 4; -2; -2; -1\}$.

Розставляємо ранги в обох наборах даних від найбільшого до найменшого.

Ранги для X : $R(X) = \{2; 1; 3; 4; 5\}$.

Ранги для Y : $R(Y) = \{3; 5; 1; 2; 4\}$.

Різниці рангів d_i : $d_i = \{-1; -4; 2; 2; 1\}$.

З чого можна зробити висновок, що d_i в обох випадках відрізнятимуться лише знаком, яким можна знехтувати при піднесенні до квадрату. Тобто спосіб ранжування можна обирати на власний розсуд.

Квадрати різниць d_i^2 : $d_i^2 = \{1, 16, 4, 4, 1\}$,

$$\sum d_i^2 = 26.$$

Коефіцієнти T_X ; T_Y : Оскільки зв'язних рангів немає, то $T_X = T_Y = 0$.

Обчислення коефіцієнта кореляції Спірмена (13.8):

$$\rho_c(X, Y) = 1 - \frac{6 \cdot 26}{5 \cdot 24} = 1 - 1,3 = -0,3.$$

Цей результат показує помірну від'ємну кореляцію між змінними X і Y .

Відповідь. Між наборами даних є помірний від'ємний кореляційний зв'язок.

Приклад 13.4. Нехай є два набори даних:

$$X = \{85; 95; 85; 70; 60\}; Y = \{80; 70; 90; 80; 60\}.$$

Встановити степінь кореляційного зв'язку між ними.

Розв'язання. Розставляємо ранги в обох наборах даних від найбільшого до найменшого. Також необхідно врахувати повтори елементів в обох вибірках.

Ранги для X : $R(X) = \{2,5; 1; 2,5; 4; 5\}$.

Ранги для Y : $R(Y) = \{2,5; 4; 1; 2,5; 5\}$.

Різниця рангів d_i : $d_i = \{0; -3; 1,5; 1,5; 0\}$.

Квадрати різниць d_i^2 : $d_i^2 = \{0; 9; 2,25; 2,25; 0\}$,

$$\sum d_i^2 = 13,5.$$

Коефіцієнти T_X ; T_Y : Оскільки обидві вибірки мають по два елементи, які повторюються, тобто $N_X = N_Y = 2$, то $T_X = T_Y = \frac{2^3 - 2}{12} = \frac{1}{2}$.

Обчислення коефіцієнта кореляції Спірмена: За формулою (13.9):

$$\rho_c(X, Y) = 1 - \frac{6 \cdot 13,5 + 0,5 + 0,5}{5 \cdot 24} \approx 1 - 0,683 \approx 0,317.$$

Цей результат показує помірну додатну кореляцію між змінними X і Y .

Відповідь. Між наборами даних є помірний додатний кореляційний зв'язок.

Коефіцієнт рангової кореляції Кендалла

Іншим показником, що характеризує узгодженість двох факторів, є коефіцієнт рангової кореляції Кендалла.

Для обчислення цього показника ранги N_x значень показника x розташовують у порядку зростання, при цьому для кожного значення рангу N_x фіксують ранг N_y відповідного значення показника y . Ідеальна кореляція між цими показниками буде спостерігатися у тому випадку, коли послідовності значень N_x та N_y будуть співпадати. Коефіцієнт рангової кореляції Кендалла дозволяє визначити міру відповідності цих послідовностей.

Для кожного значення N_y послідовно визначають кількість розташованих за ним рангів, що перевищують N_y , а також кількість рангів, менших, ніж N_y . Перша група рангів враховується зі знаком «+», їх суму позначимо P . Ранги другої групи враховуються зі знаком «-», нехай їх сума дорівнює Q .

Максимальне значення P досягається у випадку, коли ранги N_y значень фактору y співпадають з рангами N_x значень фактору x і кожна послідовність рангів співпадає з послідовністю натуральних чисел від 1 до n , розміщених у порядку зростання. Тоді після першої пари значень $N_x = 1$ та $N_y = 1$ кількість перевищень цих значень рангів буде дорівнювати $n - 1$, після другої пари $N_x = 2$ та $N_y = 2$ ця кількість буде дорівнювати $n - 2$ і так далі. Отже, у випадку, коли ранги x та y співпадають, а кількість рангів дорівнює n , маємо:

$$P_{max} = (n - 1) + (n - 2) + \dots + 3 + 2 + 1 = \frac{n \cdot (n - 1)}{2}.$$

Якщо послідовність рангів y має обернену тенденцію по відношенню до послідовності рангів x , то Q матиме таке ж максимальне значення по модулю:

$$|Q_{max}| = \frac{n \cdot (n - 1)}{2}.$$

Якщо ранги y не співпадають з рангами x , то знаходять суму $S = P + Q$. Відношення цієї суми до $P_{max} = |Q_{max}|$ дорівнює *коефіцієнту рангової кореляції Кендалла* (13.11):

$$\rho_B(X, Y) = \rho_K(X, Y) = \frac{2 \cdot S}{n \cdot (n - 1)} = \frac{4 \cdot P}{n \cdot (n - 1)} - 1 = 1 - \frac{4 \cdot Q}{n \cdot (n - 1)}. \quad (13.11)$$

Таким чином, для будь-яких двох пар даних (x_i, y_i) можна зробити такі висновки:

Узгоджена пара: Якщо $x_i < x_j, y_i < y_j$, або якщо $x_i > x_j, y_i > y_j$.

Кількість узгоджених пар дорівнює P .

Неузгоджена пара: Якщо $x_i < x_j, y_i > y_j$, або $x_i > x_j, y_i < y_j$.

Кількість неузгоджених пар дорівнює Q (береться зі знаком « $-$ »).

Нейтральна пара: Якщо $x_i = x_j$ або $y_i = y_j$ (такі пари не враховуються у підрахунках).

Отже, коефіцієнт кореляції Кендалла визначається як різниця між кількістю узгоджених пар і неузгоджених пар, нормована до кількості можливих пар.

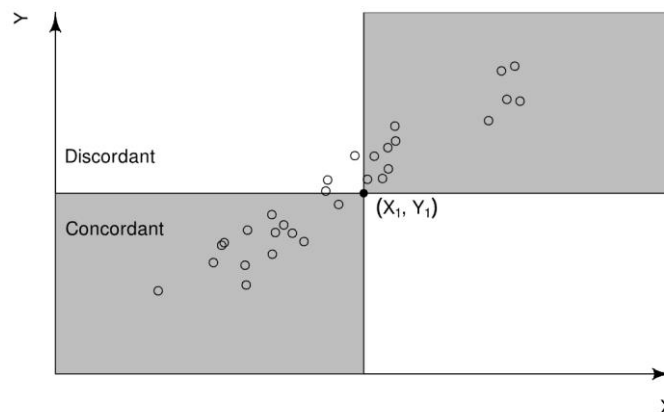


Рис. 13.5. Візуалізація узгоджених та неузгоджених з точкою (X_1, Y_1) пар

На рис. 13.5. усі точки в сірих прямокутниках є узгодженими, а всі точки в білих прямокутниках є неузгодженими з точкою (X_1, Y_1) . Загалом на графіку є $n = 30$ точок, які утворюють $C_{30}^2 = 435$ можливих пар, 395 з цих пар є узгодженими, 40 пар — неузгодженими, що дає коефіцієнт кореляції рангу Кендалла 0,816.

Приклад 13.5. Маємо дані для 10 аграрних компаній про врожайність картоплі у (ц/га) та кількість внесених на 1 га мінеральних добрив x (кг). Вихідні дані наведені у розрахунковій таблиці (табл. 13.5). З допомогою коефіцієнта рангової кореляції Кендалла виміряти щільність взаємозв'язку між цими показниками.

Таблиця 13.5. Розрахункова таблиця для визначення коефіцієнта рангової кореляції Кендалла

X	Y	Ранги		Підрахунок балів	
		N_x	N_y	« + »	« - »
138	218	1	1	9	0
175	240	2	3	7	1
190	232	3	2	7	0
196	280	4	6	4	2
200	260	5	4	5	0
235	310	6	9	1	3
250	290	7	7	2	1
260	278	8	5	2	0
275	300	9	8	1	0
290	320	10	10	—	—
$n = 10$				$P = 38$	$Q = -7$

Розв'язання. Розглянемо, як відбувається підрахунок балів.

Оскільки ранги x , тобто N_x , розташовані у порядку зростання, підрахунок балів здійснюємо, спостерігаючи за змінами N_y . Після першої пари 9 значень N_y більші 1 і жодного значення, меншого за 1. Тому у першому рядку стоїть 9 у стовпчику зі знаком « + » та 0 у стовпчику зі знаком « - ». Після другої пари, де значення $N_y = 3$, спостерігається 7 випадків, коли ранги у перевищують 3, і один випадок ($N_y = 2$), коли ранг менший 3. Відповідно, у другому рядку записані цифри 7 у стовпчику зі знаком « + » та 1 у стовпчику зі знаком « - ». Знайшовши суму елементів стовпчика зі знаком « + », отримуємо $P = 38$, підсумовуючи числа у стовпчику зі знаком « - » і враховуючи знак, отримуємо значення $Q = -7$. Тоді $S = P + Q = 38 - 7 = 31$.

Звідси знаходимо коефіцієнт кореляції Кендалла (13.11):

$$\tau(X, Y) = \frac{2 \cdot 31}{10 \cdot 9} = \frac{62}{90} \approx 0,69.$$

Отримане значення коефіцієнта рангової кореляції, згідно шкали Чеддока, свідчить про помітний кореляційний зв'язок між показниками x та y .

Відповідь. Між показниками є помітний додатний кореляційний зв'язок.

Приклад 13.6. У розрахунковій таблиці (табл. 13.6) наведено дані щодо погодинної оплати праці на підприємстві x (г.о.) та рівня плинності кадрів y (кількість працівників, що звільнилися за рік). Розрахувати коефіцієнт рангової кореляції Кендалла для цих даних.

Таблиця 13.6. Розрахункова таблиця для визначення коефіцієнта рангової кореляції Кендалла

X	Y	Ранги		Підрахунок балів	
		N_x	N_y	« + »	« - »
3	34	1	7	1	6
4	35	2	8	0	6
5	33	3	6	0	5
6	28	4	5	0	4
7	20	5	3	1	2
8	24	6	4	0	2
9	15	7	2	0	1
10	11	8	1	—	—
$n = 8$				$P = 2$	$Q = -26$

Розв'язання. Розрахункова таблиця для визначення коефіцієнта рангової кореляції Кендалла при протилежній спрямованості рангів має вигляд (табл. 13.6):

За даними розрахункової табл. 13.6 отримуємо $P = 2, Q = -26, S = P + Q = 2 - 26 = -24$.

Звідси знаходимо коефіцієнт рангової кореляції Кендалла (13.11):

$$\tau(X, Y) = \frac{2 \cdot (-24)}{8 \cdot 7} = \frac{4 \cdot 2}{8 \cdot 7} - 1 = 1 - \frac{4 \cdot 26}{8 \cdot 7} = -\frac{6}{7} \approx -0,86.$$

Отримане від'ємне значення коефіцієнта, за абсолютною величиною близьке до 1, свідчить про наявність досить сильного зворотного зв'язку між факторами x та y .

Відповідь. Між показниками є сильний від'ємний кореляційний зв'язок.

Зауважимо, що з допомогою коефіцієнтів рангової кореляції Спірмена та Кендалла можна вимірювати щільність взаємозв'язку, не лише між кількісними, але й якісними (атрибутивними) ознаками (стать, професія тощо), впорядкованими певним чином.

13.3. Коефіцієнт кореляції Фехнера

Коефіцієнт кореляції Фехнера, запропонований у 2-й половині XIX століття Г.Т. Фехнером, є найпростішою мірою зв'язку між двома змінними. Він ґрунтується на зіставленні двох психологічних ознак x_i і y_i , виміряних на одній і тій самій вибірці, за зіставленням знаків відхилень індивідуальних значень від середнього: $s(x_i - \bar{x}_B)$ і $s(y_i - \bar{y}_B)$. Висновок про кореляцію між

двома змінними роблять на підставі підрахунку числа збігів і розбіжностей цих знаків.

Коефіцієнт кореляції Фехнера обчислюється за формулою (13.12):

$$\rho_B(X, Y) = \rho_F(X, Y) = \frac{n_a - n_b}{n_a + n_b}, \quad (13.12)$$

де:

a — збіги знаків;

b — розбіжності знаків; n_a — кількість збігів;

n_b — кількість розбіжностей.

Приклад 13.7. Нехай x_i і y_i — дві ознаки, виміряні на одній і тій самій вибірці випробовуваних. За допомогою коефіцієнта кореляції Фехнера встановити рівень кореляційного зв'язку між ними.

Розв'язання. Для обчислення коефіцієнта Фехнера необхідно обчислити середні значення для кожної ознаки, а також для кожного значення змінної — знак відхилення від середнього (табл. 13.7):

Таблиця 13.7. Розрахункова таблиця для обчислення коефіцієнта кореляції Фехнера

№	x_i	y_i	$s(x_i - \bar{x}_B)$	$s(y_i - \bar{y}_B)$	Позначення
1	16	20	+	+	a
2	15	17	—	—	a
3	19	16	+	—	b
4	12	22	—	+	b
5	9	18	—	+	b
6	20	12	+	—	b
7	18	15	+	—	b
8	14	18	—	+	b
9	15	16	—	—	a
10	17	18	+	+	a

Згідно таблиці (13.7) обчислюємо вибіркові середні значення $\bar{x}_B = 15,5$; $\bar{y}_B = 17,2$. Тоді $n_a = 4$, $n_b = 6$. За формулою (13.12):

$$\rho_F(X, Y) = \frac{4 - 6}{4 + 6} = -\frac{1}{5} = -0,2.$$

Відповідь. Між показниками є слабкий від'ємний кореляційний зв'язок.

13.4. Розробка алгоритму формування портфеля послуг згідно з

принципом об'єднання на основі статистичного аналізу інформації про надання послуг

Детально даний алгоритм представлено в статті [26].

Математична постановка задачі

Нехай задано множину послуг $P = \{P_j\}$, $j = \overline{1, m}$. Кожна послуга P_j з множини P представлена набором значень:

$$P_j = (P_{j1}, \dots, P_{ji}, \dots, P_{jn}),$$

де:

P_{ji} – кількість звернень по послугу P_j ($j = \overline{1, m}$) за день i ($j = \overline{1, m}$);

n – кількість днів в періоді, за який здійснюється аналіз;

m – кількість послуг.

Необхідно на основі наявної інформації встановити взаємозв'язок між послугами множини P . На основі аналізу сили встановленого взаємозв'язку сформулювати рекомендації щодо зв'язування послуг в один або декілька портфелів послуг.

Теоретичне обґрунтування підходу

Кожна послуга P_j з множини P може розглядатись як дискретна випадкова величина, яка представлена своїм рядом розподілу. Тоді взаємозв'язок між цими величинами можна встановити на основі аналізу коефіцієнту парної кореляції Пірсона (13.3).

Сила кореляційних зв'язків між послугами встановлюється згідно шкали Чеддока (лекція 12). При формуванні рекомендацій щодо об'єднання послуг доцільно враховувати зв'язки, сила яких визначається діапазоном $0,7 \leq \rho_{\Pi}(P_k, P_j) \leq 1$.

Слід зазначити, що для отримання достовірних результатів аналізу необхідно встановлювати статистичний зв'язок між послугами, використовуючи тільки статистично значущі значення коефіцієнтів кореляції.

Під *статистично значущім кореляційним зв'язком* між парами послуг мається на увазі ті пари послуг, замовлення однієї з яких добре узгоджується із замовленням іншої, тобто сила зв'язку встановлюється не тільки кількісно (значення коефіцієнта кореляції), а й якісно (що обидві послуги замовлять ті самі люди).

Перевірити статистичну значущість розрахованого коефіцієнта кореляції можна, обчисливши значення t -критерія Стьюдента (13.13):

$$t_{\text{сп } k, j} = \rho_{\Pi}(P_k, P_j) \cdot \sqrt{\frac{n-2}{1-\rho_{\Pi}^2(P_k, P_j)}}, \quad (13.13)$$

де:

$\rho_{\Pi}(P_k, P_j)$ – коефіцієнт кореляції випадкових величин P_k, P_j ;

n – розмір вибірки (кількість днів в періоді).

Порівнюємо його із критичним табличним значенням цього критерія $t_{\text{кр}}$, яке береться із таблиці розподілу Стьюдента (дод. 6 Гмурмана) із

врахуванням заданого рівня значущості $\alpha = 0,05$, що є достатнім рівнем для отримання достовірних результатів, та кількості ступенів свободи $k = n - 2$.

Якщо розраховане значення $|t_{\text{сп } k, j}| > t_{\text{кр}}$, то відповідний коефіцієнт кореляції є статистично значущим.

Проведення статистичного аналізу для сформованого датасету

Для встановлення наявності зв'язку між кожною парою послуг побудовано матрицю парних коефіцієнтів кореляції, яка представлена на рис. 13.6. Враховуючи симетричність матриці відносно головної діагоналі, на рис. 13.6 наведено половину матриці. Значення коефіцієнтів, які відповідають помітним та сильним зв'язкам, виділено червоним кольором. Так, значення коефіцієнта кореляції для пари послуг 34 та 37 дорівнює 0,7, що згідно зі шкалою Чеддока, свідчить про наявність сильного статистичного зв'язку між цими послугами.

В загальному випадку вибір послуг для аналізу рекомендовано здійснювати за кількістю звернень, що перевищують вибіркове середнє. В прикладі вибрано 24 послуги з 84, кількість звернень за цими послугами складає 2173, або 91% від загальної кількості звернень.

	2	3	5	6	7	13	14	24	26	30	34	37	40	45	46	47	50	56	63	69	71	74	76	84
2		0	0,2	0,5	0	-0	0,2	-0	0,3	-0	-0	-0	0,2	-0	-0	-0	-0	-0	0	0,4	0,2	0	-0	-0
3			0,1	0	0,1	0,2	-0	-0	-0	0,1	-0	-0	-0	0,1	0,1	0,1	0	-0	-0	0,1	-0	0,1	0	0,2
5				0,3	-0	0,2	0,1	0,1	0,2	0,5	0,2	0,3	-0	0	-0	0,1	-0	0	0,4	0,4	0,1	0,2	-0	0,2
6					-0	0,2	0,1	0,1	0,3	-0	0,2	0,4	0,4	0,4	-0	0,2	-0	0,1	0,1	0,5	0,2	0,1	-0	0,4
7						0,1	0,2	0,1	0,1	-0	-0	-0	-0	-0	-0	0,3	-0	0,2	-0	-0	0,2	-0	0,2	0,4
13							0,1	-0	0,3	0,3	-0	0	0,3	0,1	-0	0,1	0,3	0,4	-0	0,1	0,1	0,1	0,1	0,4
14								0,1	0,1	0,1	-0	0,1	0,3	0,2	0	0,4	0,3	-0	0,5	0,3	0,3	0,2	0,5	-0
24									0,3	-0	0	0	-0	0,2	0,4	0,2	-0	0,3	0	0,3	0,4	0	-0	0,4
26										-0	-0	-0	-0	-0	0,2	0,2	-0	0,1	-0	0,4	0	0,1	-0	0,4
30											0,1	0,1	-0	-0	0	0,2	0,1	0,3	-0	0,1	0,2	0,3	0,2	-0
34												0,7	0,2	0,2	0,3	0,1	-0	-0	0,1	-0	0	0	-0	0
37													0,2	0,5	0	0,1	-0	0,1	0,4	0,1	0	-0	0	0,2
40														0,2	0,1	0,2	0,4	-0	0,1	-0	-0	0	0,2	-0
45															0,2	0,1	0,4	0,2	0,3	0,2	-0	-0	0,5	0,1
46																0,3	0	-0	-0	-0	0	0	-0	0,1
47																	-0	-0	-0	0,1	0,1	0	0,2	0,3
50																		0,1	0,2	0,1	-0	0,3	0,6	-0
56																			-0	0,1	0,5	0,2	0	0,3
63																				0,3	-0	0	0,1	-0
69																					0,4	-0	0,2	0
71																						0,2	-0	0,3
74																							-0	0
76																								-0
84																								

Рис. 13.6. Матриця парних коефіцієнтів кореляції

Для отримання достовірних результатів аналізу необхідно перевірити гіпотезу про статистичну значущість отриманих значень коефіцієнтів кореляції з рис. 13.6. Для цього вибирається табличне значення критерія Стюдента для $n = 24$ (ступінь свободи $k = 22$), $\alpha = 0,05$, яке дорівнює $t_{\text{сп}} = 2,07$.

На рис. 3 наведено матрицю розрахованих за формулою (13.13) значень t -критерія для перевірки статистичної значущості розрахованих коефіцієнтів кореляції (тут $n = 24$, r – коефіцієнт кореляції, значущість якого перевіряється).

Якщо розраховане значення $|t_{\text{сп}}| > t_{\text{кр}}$, то відповідний коефіцієнт кореляції є статистично значущим (в матриці на рис. 13.7 такі значення t -

критерія виділено червоним кольором).

	2	3	5	6	7	13	14	24	26	30	34	37	40	45	46	47	50	56	63	69	71	74	76	84	
2		0,1	1	2,4	0,1	0,8	0,9	0,3	1,3	0,8	0,8	0,8	0,7	0,1	1,1	0,9	1,1	0,6	0,1	2,2	0,9	0	0,3	0,2	
3				0,5	0	0,5	1,2	0,3	0,5	0,3	0,2	1,6	1,4	0,2	0,3	0,5	0,2	0	0,3	1,5	0,2	0,2	0,3	0,2	0,9
5					1,5	2,1	1	0,3	0,6	1	2,4	0,8	1,3	1	0,2	0,3	0,2	1,2	0,1	2,1	1,8	0,3	1	1,3	0,7
6						0,5	0,9	0,5	0,3	1,6	0,7	1	2,2	2	1,9	0,5	0,9	0,2	0,3	0,4	2,6	0,8	0,4	0,3	1,9
7							0,7	0,9	0,5	0,6	0,9	2,3	1	0,6	0,7	0,9	1,4	0,9	1,2	1,5	0,3	1	1,6	1	2,2
13								0,5	0,1	1,4	1,3	0,7	0,1	1,3	0,4	0,5	0,4	1,5	1,9	0,3	0,5	0,3	0,6	0,3	1,9
14									0,5	0,6	0,4	0,8	0,5	1,3	1,2	0,1	2,3	1,6	0,3	2,4	1,5	1,3	0,9	2,5	0,1
24										1,4	0,3	0,1	0,2	2,3	1,1	1,9	0,7	0,7	1,3	0,1	1,5	2,1	0	0,8	2,1
26											0,1	1,2	0,8	0,7	0,9	1	1,1	0,2	0,3	0,6	2,2	0,1	0,6	1,1	1,8
30												0,6	0,6	0,1	0,3	0,2	1	0,3	1,4	0	0,3	0,7	1,2	1	0,1
34													4,7	0,9	1,1	1,4	0,3	0,4	0,3	0,7	0,2	0	0	0,6	0,1
37														1	2,4	0,2	0,5	0,3	0,5	1,9	0,4	0,1	0,6	0,1	1,1
40															1,1	0,3	0,7	1,9	1,1	0,4	0,1	0,8	0	1,2	1
45																0,8	0,3	1,9	1	1,3	1,2	0,2	0,3	2,5	0,4
46																	1,3	0,2	0,6	1,2	0,3	0,1	0,1	0,6	0,4
47																		0,5	0,7	0,1	0,5	0,6	0,1	1,1	1,3
50																			0,5	1	0,3	1,2	1,4	3,3	0,8
56																				1	0,5	2,6	1	0,2	1,6
63																					1,3	1,1	0,2	0,6	1,6
69																						1,8	0,3	1	0,1
71																							1	0	1,3
74																								0,7	0,2
76																									1
84																									

Рис. 13.7. Матриця коефіцієнтів критерія Стьюдента

Згідно з даними, наведеними в матриці на рис. 13.7, можемо побачити, що сильний зв'язок послуг 37 та 34, є статистично значущим. Послуга 76 є помітно пов'язаною з послугою 50. Про це свідчить значення коефіцієнта кореляції 0,6 для цієї пари послуг, що видно із рис. 13.7. Цей зв'язок також є статистично значущим, про що свідчить значення *t*-критерія для цієї пари послуг на рис. 13.7.

Перелік послуг, пов'язаних із послугами 37 та 76 згідно проведеного аналізу, наведено в табл. 13.8, 13.9 відповідно.

Таблиця 13.8. Перелік послуг, пов'язаних із послугою 37 «Державна реєстрація інших (відмінних від права власності) речових прав на нерухоме майно»

№	Коефіцієнт кореляції	Номер послуги	Назва послуги
1	0,71143376	34	Державна реєстрація права власності на нерухоме майно
2	0,45735186	45	Надання відомостей з Державного земельного кадастру у формі копій документів, що створюються під час ведення Державного земельного кадастру.
3	0,429127614	6	Видача довідки до нотаріальної контори про реєстрацію місця проживання громадянина за даною адресою на день смерті (для оформлення спадщини та інше)

Таблиця 13.9. Перелік послуг, пов'язаних із послугою 76 «Призначення грошової компенсації вартості одноразової натуральної допомоги «пакунок малюка»»

№	Коефіцієнт кореляції	Номер послуги	Назва послуги
1	0,576879259	50	Надання державної допомоги при народженні дитини
2	0,470902426	14	Видача довідки про склад зареєстрованих у житловому приміщенні осіб
3	0,46632422	45	Надання відомостей з Державного земельного кадастру у формі копій документів, що створюються під час ведення Державного земельного кадастру

В табл. 13.8, 13.9 наведено значення коефіцієнтів кореляції, які є статистично значущими.

Для формування портфелів послуг на основі проведеного аналізу вибираються послуги, для яких встановлено наявність сильного статистично значущого зв'язку.

Для розглянутого прикладу в результаті формуються рекомендації щодо створення двох портфелів послуг – один включає послуги 34 і 37, а інший – послуги 76 та 50.

13.5. Встановлення рівня залежності та її статистичної значущості між прогнозованими та реальним курсами криптовалюти

Детально даний підхід представлено в статті [27].

Встановлення рівня залежності між прогнозованими та реальним курсами криптовалюти

Постановка задачі: Нехай задано пари статистичних вибірок (X, Y_i) , $i = \overline{1; 3}$ (X – реальні курси криптовалюти за певний проміжок часу; Y_i – прогнози курсу криптовалюти в аналогічні моменти часу, що зроблені за допомогою алгоритмів АУДСМ, ARIMA та Експоненційного згладжування відповідно).

Необхідно встановити рівень залежності між парами вибірок (X, Y_i) , враховуючи порядок слідування елементів у вибірці.

Обґрунтування: Загалом, для встановлення рівня залежності між двома вибірками використовується коефіцієнт кореляції Пірсона (13.3).

Оскільки для даного дослідження істотним є порядок слідування елементів у вибірках (реальний та прогнозований курси криптовалют розглядаються і порівнюються у відповідний момент часу), то рекомендовано використовувати коефіцієнт рангової кореляції, зокрема коефіцієнти Спірмена, яка обчислюється за формулами (13.9) та (13.10).

Слід зазначити, що коефіцієнт рангової кореляції Спірмена (13.9) є менш точними у порівнянні з коефіцієнтом кореляції Пірсона (13.3), оскільки при його розрахунку не враховуються кількісні значення елементів вибірок, а лише їх порядок. Це свідчить про необхідність враховувати коефіцієнт кореляції Пірсона при більш глибокому визначенні рівня залежності вибірок X та Y_i , $i = \overline{1; 3}$. Його встановлення відбувається за допомогою шкали Чеддока (лекція 12).

Запропонований підхід дає змогу оцінити наскільки тісний кореляційний зв'язок існує між реальними та прогнозованими курсами обраної криптовалюти за вказаний проміжок часу.

Результат: Для вибірок, результати представляються у наступному вигляді (табл. 13.10).

Таблиця 13.10. Аналіз кореляційного зв'язку між парами вибірок (X, Y_i) , $i = \overline{1; 3}$

Вибір-ка	Коефіцієнт рангової кореляції Спірмена $\rho_c(X, Y_i)$	Коефіцієнт рангової кореляції Кендалла $\rho_k(X, Y_i)$	Коефіцієнт кореляції Пірсона $\rho_p(X, Y_i)$	Сила зв'язку з вибіркою X
Y_1	0,97828283	0,91111111	0,99829258	Дуже сильний зв'язок як з точки зору подібності елементів вибірок, так і з точки зору врахування їх порядку слідування
Y_2	0,39040404	0,33333333	0,581244474	Помітний зв'язок з точки зору подібності елементів вибірок, та помірний з точки зору врахування їх порядку слідування
Y_3	0,09949495	0,06666667	0,58275805	Помітний зв'язок з точки зору подібності елементів вибірок, але майже відсутній з точки зору врахування їх порядку слідування

В даній таблиці коефіцієнти кореляції обчислюються за формулами (13.3), (13.9), (13.10), (13.11), а сила зв'язку з вибіркою X встановлюється за допомогою табл. 3.

З табл. 13.10 видно, що вибірка Y_1 має найсильніший кореляційний зв'язок з вибіркою X серед усіх розглянутих вибірок (дуже сильний зв'язок). Середній рівень зв'язку з вибіркою X має вибірка Y_2 , найгірші показники має вибірка Y_3 .

Встановлення статистичної значущості прогнозованих курсів криптовалюти

Постановка задачі: Нехай задана пара вибірок (X, Y_i) , $i = \overline{1; 3}$. Між ними відомий коефіцієнти рангової кореляції Спірмена, Кендалла або Пірсона.

Потрібно при рівні значущості α перевірити гіпотезу про значущість відповідного коефіцієнта кореляції.

Обґрунтування: Для розв'язання сформульованої задачі, використовувалися наступні правила, які входять до t -критерію Стюдента.

Правило 1. Для того, щоб при рівні значущості α перевірити гіпотезу про значущість коефіцієнтів рангової кореляції Спірмена необхідно обчислити критичне значення (13.14):

$$T_{кр i} = t_{кр} \cdot \sqrt{\frac{1 - \rho_c^2(X, Y_i)}{n - 2}}, \quad (13.14)$$

де критичне значення $t_{кр}$ знаходиться по таблиці розподілу Стюдента (дод. 6 Гмурмана) за рівнем значущості α та ступенем свободи $k = n - 2$, $\rho_c(X, Y_i)$ – коефіцієнти рангової кореляції Спірмена, $i = \overline{1; 3}$, n – об'єм вибірок X та Y_i .

Якщо $|\rho_c(X, Y_i)| \leq T_{кр i}$, то гіпотеза приймається, в противному разі – відхиляється, тобто між вибірками існує значущий кореляційний зв'язок.

Правило 2. Для того, щоб при рівні значущості α перевірити гіпотезу про значущість коефіцієнтів рангової кореляції Кендалла $\rho_K(X, Y_i)$, $i = \overline{1; 3}$, необхідно обчислити критичне значення (13.15):

$$T_{кр i} = T_{кр} = z_{кр} \cdot \sqrt{\frac{2 \cdot (2 \cdot n + 5)}{9 \cdot n \cdot (n - 1)}}, \quad (13.15)$$

де при заданому рівні значущості α необхідно відшукати критичну точку $z_{кр}$ з рівняння $\Phi(z_{кр}) = \frac{1 - \alpha}{2}$ (див. дод. 2 Гмурмана).

Якщо $|\rho_K(X, Y_i)| \leq T_{кр}$, то гіпотеза приймається, в противному разі – відхиляється, тобто між вибірками існує значущий кореляційний зв'язок.

Правило 3. Для того, щоб при рівні значущості α перевірити гіпотезу про значущість коефіцієнта кореляції Пірсона необхідно обчислити спостережуване значення (13.16):

$$t_{сп i} = \rho_{\Pi}(X, Y_i) \cdot \sqrt{\frac{n - 2}{1 - \rho_{\Pi}^2(X, Y_i)}}, \quad (13.16)$$

де:

n – об'єм вибірок X та Y_i ;

$\rho_{\Pi}(X, Y_i)$ – коефіцієнти кореляції Пірсона;

$i = \overline{1; 3}$.

Формула (13.16) є аналогічною до формули (13.13). Різниця полягає у позначенні вибірок.

Далі необхідно порівняти його із табличним критичним значенням цього критерія $t_{кр}(\alpha, k)$ (яке береться із таблиці розподілу Стюдента (дод. 6 Гмурмана) з урахуванням заданого рівня значущості α , що є достатнім рівнем для отримання достовірних результатів, та кількості ступенів свободи $k = n - 2$).

Якщо $|t_{сп i}| \leq t_{кр}(\alpha, k)$, то гіпотеза приймається, інакше – відхиляється, тобто між вибірками існує значущий кореляційний зв'язок.

Під **значущим кореляційним зв'язком** мається на увазі, що прогнозовані курси криптовалют, тобто вибірки Y_i добре узгоджуються з реальними курсами в аналогічні моменти часу, тобто з вибіркою X . Зокрема, якщо розглянути пару вибірок (X, Y_1) , в рамках запропонованого підходу можна простежити, на скільки прогноз курсу криптовалют з урахуванням дописів відомих людей в соціальних мережах корелюється з реальними курсами валют за аналогічний період часу. Це дає змогу оцінити рівень впливу цих дописів на курси криптовалют, що і є основною метою даного дослідження.

Результат: Для вибірок Y_i з урахуванням коефіцієнтів рангової кореляції, наведених в табл. 13.10, результати представляються у наступному вигляді (табл. 13.11-13.13):

Таблиця 13.11. Результати перевірки гіпотези про значущість коефіцієнта рангової кореляції Спірмена між парами вибірок (X, Y_i) при $\alpha = 0,05$ і $t_{кр}(0,05; 8) = 2,31$

Вибірка	Коефіцієнт рангової кореляції Спірмена $\rho_s(X, Y_i)$	Значення $T_{кр i}$	Висновки про прийняття або відкидання гіпотези
Y_1	0,97828283	0,16	значущий кореляційний зв'язок
Y_2	0,39040404	0,75	незначущий кореляційний зв'язок
Y_3	0,09949495	0,81	незначущий кореляційний зв'язок

Таблиця 13.12. Результати перевірки гіпотези про значущість коефіцієнта кореляції Кендалла між парами вибірок (X, Y_i) при $n = 10$; $\alpha = 0,05$ і $z_{кр}(0,05) = 1,96$

Вибірка	Коефіцієнт кореляції Кендалла $\rho_K(X, Y_i)$	Значення $T_{кр}$	Висновки про прийняття або відкидання гіпотези
Y_1	0,9111111	0,49	значущий кореляційний зв'язок
Y_2	0,3333333	0,49	незначущий кореляційний зв'язок
Y_3	0,0666667	0,49	незначущий кореляційний зв'язок

Таблиця 13.13. Результати перевірки гіпотези про значущість коефіцієнта кореляції Пірсона між парами вибірок (X, Y_i) при $\alpha = 0,05$ і $t_{кр}(0,05; 8) = 2,31$

Вибірка	Коефіцієнт кореляції Пірсона $\rho_P(X, Y_i)$	Значення $t_{сп i}$	Висновки про прийняття або відкидання гіпотези
Y_1	0,99829258	48,34	значущий кореляційний зв'язок
Y_2	0,581244474	2,02	незначущий кореляційний зв'язок
Y_3	0,58275805	2,03	незначущий кореляційний зв'язок

В даних таблицях коефіцієнти кореляції обчислюються за формулами (13.3), (13.9), (13.10) та (13.11), а критичні значення – за формулами (13.14), (13.15) та (13.16) відповідно.

З табл. 13.11-13.13 видно, що в парі вибірок (X, Y_1) переважно існує значущий кореляційний зв'язок, в той час, як в парах вибірок (X, Y_2) та (X, Y_3) кореляційний зв'язок є незначущим.

Більш докладно про статистичні критерії розглядатиметься в лекції 18.

ЛЕКЦІЯ 14

ЗАКОН ВЕЛИКИХ ЧИСЕЛ. ГРАНИЧНІ ТЕОРЕМИ ТЕОРІЇ ЙМОВІРНОСТЕЙ

Граничні теореми теорії ймовірностей умовно поділяють на дві групи. Перша група теорем, яку називають Законом Великих Чисел, встановлює стійкість середніх значень, тобто при великій кількості експериментів їх середній результат перестає бути випадковим і може бути передбаченим з достатньою точністю. Друга група теорем, яку називають Центральною Граничною Теоремою, встановлює умови, при яких закон розподілу суми великого числа випадкових величин необмежено наближається до нормального.

14.1. Закон Великих Чисел

Під законом великих чисел у широкому сенсі розуміють загальний принцип, згідно з яким, за формулюванням академіка О.М. Колмогорова, сукупна дія великої кількості випадкових факторів призводить (за деяких вельми загальних умов) до результату, майже не залежного від випадку. Іншими словами, при великому числі випадкових величин їхній середній результат перестає бути випадковим і може бути передбачений з великим ступенем визначеності.

Закон Великих Чисел — фундаментальний принцип теорії ймовірностей, який описує поведінку середніх значень великих наборів випадкових величин. Він вказує, що середнє значення великої кількості незалежних і однаково розподілених випадкових величин наближається до математичного сподівання цих величин зі збільшенням кількості спостережень.

Закон Великих Чисел є важливим результатом теорії ймовірностей, який описує стабілізацію середніх значень великих наборів випадкових даних. Він забезпечує теоретичну основу для багатьох методів статистичного аналізу і моделювання, дозволяючи робити надійні висновки на основі вибірових даних.

Інтуїтивне пояснення

Закон Великих Чисел можна інтуїтивно зрозуміти як принцип, згідно з яким випадкові коливання середніх значень врівноважуються зі збільшенням числа спостережень. Наприклад, якщо ми підкидаємо монету багато разів, відсоток випадків випадання орла наближатиметься до 50%.

Приклади застосування

Статистичне оцінювання: У вибірових дослідженнях середнє значення вибірки використовується для оцінювання середнього значення генеральної сукупності. Закон великих чисел гарантує, що ця оцінка буде точною при достатньо великій вибірці.

Страховання: Закон великих чисел є основою для страхових компаній, які використовують середні значення для прогнозування виплат по страхових випадках.

Гральні ігри: В азартних іграх закон великих чисел пояснює, чому казино завжди виграють у довгостроковій перспективі — середній виграш казино стабілізується навколо значення, що забезпечує прибуток.

До закону великих чисел належать:

1. Нерівність Маркова

Нерівність Маркова — один з основних результатів теорії ймовірностей, який дозволяє оцінити ймовірність того, що випадкова величина перевищить певне значення. Ця нерівність застосовується до не від'ємних випадкових величин і дозволяє робити корисні оцінки в різних задачах.

Теорема 14.1. Нехай X — не від'ємна випадкова величина, тобто $X \geq 0$, і нехай існує константа $A > 0$. Тоді виконується нерівність (14.1):

$$P(X \geq A) \leq \frac{M(X)}{A}, \quad (14.1)$$

або з нерівність (14.2):

$$P(X < A) \geq 1 - \frac{M(X)}{A}, \quad (14.2)$$

де $M(X)$ — математичне сподівання випадкової величини X .

Доведення. Доведення проведемо для дискретної випадкової величини X . Розташуємо її значення в порядку зростання наступним чином: частина значень $x_1, \dots, x_k$, які будуть не більшими за число A , а інша частина значень $x_{k+1}, \dots, x_n$, які будуть більшими за A , тобто $x_1 \leq A, \dots, x_k \leq A, x_{k+1} > A, \dots, x_n > A$ (рис. 14.1).

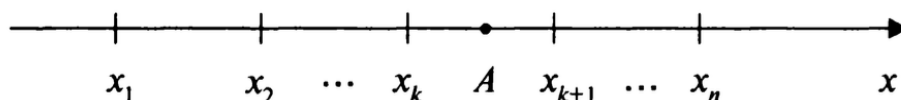


Рис. 14.1. Розташування значень випадкової величини X

Математичне сподівання дорівнює:

$$M(X) = x_1 \cdot p_1 + \dots + x_k \cdot p_k + x_{k+1} \cdot p_{k+1} + \dots + x_n \cdot p_n,$$

де:

$x_1, \dots, x_n$ — значення випадкової величини;

$p_1, \dots, p_k$ — ймовірності їх набуття.

Відкидаючи перші k невід'ємних доданків (оскільки всі $x_i \geq 0$, $i = \overline{1, n}$), отримаємо

$$x_{k+1} \cdot p_{k+1} + \dots + x_n \cdot p_n \leq M(X).$$

Замінюючи x_i на A матимемо:

$$A \cdot p_{k+1} + \dots + A \cdot p_n \leq M(X) \Rightarrow A \cdot (p_{k+1} + \dots + p_n) \leq M(X) \Rightarrow$$

$$p_{k+1} + \dots + p_n \leq \frac{M(X)}{A} \Rightarrow P(X \geq A) \leq \frac{M(X)}{A}.$$

Нерівність (14.2) отримується за допомогою властивостей протилежних подій.

Інтуїтивне пояснення

Нерівність Маркова (14.1) стверджує, що ймовірність того, що випадкова величина X перевищує певне значення A , не може бути більшою, ніж відношення математичного сподівання X до цього значення A . Іншими словами, чим більше A порівняно з $M(X)$, тим менша ймовірність того, що X перевищить A .

Нерівність Маркова можна застосовувати до будь-яких невід’ємних випадкових величин.

Приклад 14.1. Нехай X — випадкова величина, яка приймає значення числа від 0 до 10. Припустимо, що $M(X) = 4$. Оцінити ймовірність того, що X перевищує 8.

Розв’язання. За нерівністю Маркова (14.1).

$$P(X \geq 8) \leq \frac{M(X)}{8} = \frac{4}{8} = 0,5.$$

Це означає, що ймовірність того, що X перевищує 8, не може бути більшою за 0,5.

Відповідь. $P(X \geq 8) \leq 0,5$.

2. Нерівність Чебишева

Нерівність Чебишева — це важливий результат у теорії ймовірностей, який надає оцінку ймовірності того, що випадкова величина значно відхиляється від свого середнього значення. Вона є узагальненням нерівності Маркова і застосовується до випадкових величин з кінцевою дисперсією.

Теорема 14.2. Нехай X — випадкова величина з математичним сподіванням $M(X)$ і дисперсією $D(X)$. Тоді для будь-якого додатного числа $\varepsilon > 0$ виконується нерівність (14.3):

$$P(|X - M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}, \quad (14.3)$$

або нерівність (14.4):

$$P(|X - M(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}. \quad (14.4)$$

Доведення. Застосовуючи нерівність Маркова (14.1) до випадкової величини $X^* = (X - a)^2$ та поклавши $A = \varepsilon^2$, матимемо:

$$P(X^* \geq A) = P((X - a)^2 \geq \varepsilon^2) \leq \frac{M(X - a)^2}{\varepsilon^2}.$$

Оскільки з $(X - a)^2 \geq \varepsilon^2 \Rightarrow |X - a| \geq \varepsilon$, а $M(X - a)^2 = D(X)$, то отримуємо нерівність (14.3).

Нерівність (14.4) отримується за допомогою властивостей протилежних подій.

Інтуїтивне пояснення

Нерівність Чебишева (14.3) стверджує, що ймовірність того, що випадкова величина X відхилиться від свого математичного сподівання $M(X)$ щонайменше, ніж на ε , не перевищує $\frac{D(X)}{\varepsilon^2}$. Іншими словами, вона надає верхню межу ймовірності великих відхилень від середнього значення.

Нерівність Чебишева застосовна для багатьох випадкових величин. У формі (14.3) вона встановлює верхню границю, а у формі (14.4) – нижню межу ймовірності розглянутої події, що розглядається.

Приклад 14.2. Нехай X — випадкова величина з математичним сподіванням $M(X) = 50$ і стандартним відхиленням $\sigma = 5$. Оцінити ймовірність того, що X відхилиться від свого математичного сподівання на більше, ніж 3 стандартні відхилення (тобто на більше, ніж 15).

Розв’язання. За нерівністю Чебишева:

$$P(|X - 50| \geq 15) \leq \frac{25}{225} = \frac{1}{9} \approx 0,111.$$

Це означає, що ймовірність того, що X відхилиться від свого середнього значення на більше, ніж 15, не перевищує приблизно 11,1%.

Відповідь. $P(|X - 50| \geq 15) \leq \frac{1}{9}$.

3. Теорема Чебишева

Теорема Чебишева, також відома як **теорема Бернштейна-Чебишева**, є важливим результатом у теорії ймовірностей. Вона є узагальненням закону великих чисел і стверджує, що середнє значення великої кількості незалежних випадкових величин, які мають скінченну дисперсію, збігається до їх математичного сподівання в ймовірнісному сенсі.

Теорема 14.3. Нехай $X_1, X_2, \dots, X_n$ — послідовність попарно незалежних випадкових величин з скінченим математичним сподіванням $M(X_k) < +\infty$ і дисперсіями обмеженими однією і тією ж сталою $D(X_k) \leq C$. Тоді для будь-якого додатного числа $\varepsilon > 0$ справедлива рівність (14.5):

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \cdot \sum_{k=1}^n X_k - \frac{1}{n} \cdot \sum_{k=1}^n M(X_k)\right| < \varepsilon\right) = 1. \quad (14.5)$$

Доведення. За умовою теореми відомі $M(X_1), \dots, M(X_n)$ та $D(X_1) \leq C, \dots, D(X_n) \leq C$.

$$X = \frac{1}{n} \cdot \sum_{k=1}^n X_k.$$

$$M(X) = M\left(\frac{1}{n} \cdot \sum_{k=1}^n X_k\right) = \frac{1}{n} \cdot \sum_{k=1}^n M(X_k).$$

$$D(X) = D\left(\frac{1}{n} \cdot \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \cdot \sum_{k=1}^n D(X_k) \leq \frac{n \cdot C}{n^2} = \frac{C}{n}.$$

З нерівності Чебишева (14.3)

$$P\left(\left|\frac{1}{n} \cdot \sum_{k=1}^n X_k - \frac{1}{n} \cdot \sum_{k=1}^n M(X_k)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2} \geq 1 - \frac{C}{n \cdot \varepsilon^2}.$$

При $n \rightarrow +\infty$ отримуємо рівність (14.5).

Інтуїтивне пояснення

Теорема Чебишева стверджує, що середнє значення великої кількості незалежних однаково розподілених випадкових величин буде з великою ймовірністю дуже близьким до середнього значення математичних сподівань $M(X_k)$ цих величин. Це означає, що ймовірність того, що середнє значення вибірки сильно відхилиться від середнього значення $M(X_k)$, зменшується до нуля зі збільшенням розміру вибірки n .

Приклад 14.3. Нехай $X_1, X_2, \dots, X_n$ — незалежні однаково розподілені випадкові величини з математичними сподіваннями $M(X_k) = 10$ і дисперсіями $D(X_k) = 4$. Оцінити ймовірність того, що середнє значення вибірки відхилиться від 10 більше, ніж на 2, коли $n \rightarrow +\infty$.

Розв'язання. Використаємо нерівність Чебишева (14.3):

$$P\left(\left|\frac{1}{n} \cdot \sum_{k=1}^n X_k - 10\right| \geq 2\right) \leq \frac{4}{n \cdot 2^2} = \frac{1}{n} \rightarrow 0, n \rightarrow +\infty.$$

Це означає, що ймовірність того, що вибіркове середнє відхилиться від 10 більше, ніж на 2, є неможливою подією, що узгоджується з теоремою Чебишева (14.5).

Відповідь. Вибіркове середнє не відхилиться від 10 більше, ніж на 2.

4. Теорема Бернуллі

Теорема Бернуллі є спеціальним випадком Закону Великих Чисел для послідовності незалежних випробувань Бернуллі.

Теорема Бернуллі теоретично обґрунтовує властивість стійкості відносної частоти (при необмеженому збільшенні числа випробувань частота випадкової події збігається, за ймовірністю, до ймовірності події, якщо ймовірність події від випробування до випробування не змінюється і дорівнює p).

Теорема 14.4. Нехай m — число настання події A , при n послідовних незалежних випробуваннях, в кожному з яких ймовірність настання події A дорівнює p . Тоді для будь-якого як завгодно малого числа $\varepsilon > 0$ справедлива нерівність (14.6):

$$P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{p \cdot q}{n \cdot \varepsilon^2}, \quad (14.6)$$

яка при $n \rightarrow +\infty$ переходить у граничну рівність (14.7):

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \quad (14.7)$$

Доведення. Нерівність (14.6) отримано безпосередньо з нерівності Чебишева (14.4), покладаючи $X = \frac{m}{n}$ та враховуючи біноміальність розподілу.

$$M(X) = n \cdot p \cdot \frac{1}{n} = p; D(X) = n \cdot p \cdot q \cdot \frac{1}{n^2} = \frac{p \cdot q}{n}.$$

Тоді отримуємо нерівність (14.6).

При $n \rightarrow +\infty$ отримуємо рівність (14.7).

Сенс теореми Бернуллі полягає у тому, що при $n \rightarrow +\infty$ відносна частота (випадкова величина) гарантовано мало відрізняється від ймовірності (невипадкової величини). Тобто теж стає не випадковою величиною.

Інтуїтивне пояснення

Теорема Бернуллі стверджує, що якщо ми будемо проводити достатньо велику кількість незалежних випробувань, відносна частота успіхів (усереднене значення) наблизиться до справжньої ймовірності успіху p . Це означає, що у великій серії випробувань фактична частка успіхів буде дуже близькою до очікуваної ймовірності успіху.

Приклад 14.4. Монету підкидаєте 10000 разів. Нехай m – число випадань герба. Довести, що відносна частота випадання герба буде дуже близькою до ймовірності $p = 0,5$ зі збільшенням кількості підкидань n ($\varepsilon = 0,05$).

Розв'язання. Якщо ми оберемо $\varepsilon = 0,05$, то за наслідком 1 з інтегральної теореми Муавра-Лапласа (4.5) лекції 4:

$$P\left(\left|\frac{m}{10000} - 0,5\right| < 0,05\right) \approx 2 \cdot \Phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sqrt{p \cdot q}}\right) = 2 \cdot \Phi\left(\frac{0,05 \cdot \sqrt{10000}}{\sqrt{0,5 \cdot 0,5}}\right) = 2 \cdot \Phi(10) = 1.$$

Це узгоджується з теоремою Бернуллі (14.6)

$$P\left(\left|\frac{m}{10000} - 0,5\right| < 0,05\right) \geq 1 - \frac{0,5 \cdot 0,5}{10000 \cdot 0,05^2} \approx 0,99.$$

і означає, що з дуже високою ймовірністю частка гербів у 10000 підкиданнях буде в інтервалі від 0,45 до 0,55.

Відповідь. $P\left(\left|\frac{m}{10000} - 0,5\right| < 0,05\right) \geq 0,99$

Приклад 14.5. Знайти ймовірність того, що частота появи шістки в 10000 незалежних підкиданнях грального кубика відхиляється від ймовірності появи шістки за абсолютною величиною менше ніж на 0,01.

Розв'язання. За умовою $n = 10000$, $p = \frac{1}{6}$, $q = \frac{5}{6}$.

За формулою (14.6):

$$P\left(\left|\frac{m}{10000} - \frac{1}{6}\right| < 0,01\right) \geq 1 - \frac{\frac{1}{6} \cdot \frac{5}{6}}{10000 \cdot 0,01^2} \approx 0,86.$$

Відповідь. $P\left(\left|\frac{m}{10000} - \frac{1}{6}\right| < 0,01\right) \geq 0,86.$

Приклад 14.6. За даними служби перевезень аеропорту кількість затриманих за метеоумовами рейсів складає 7% від їх загальної щорічної кількості. Наступного року планується виконати 1400 рейсів. Застосовуючи теорему Бернуллі, **а)** оцінити ймовірність того, що в наступному році відносна частота затримки рейсів відхилиться від її ймовірності за абсолютною величиною менше ніж на 0,02; **б)** визначити, в

яких межах з ймовірністю, не меншою 0,96, буде знаходитись відносна частота затримки рейсів; **в)** знайти, скільки потрібно зробити рейсів, щоб з ймовірністю, не меншою 0,9, можна було б сподіватись, що абсолютна величина відхилення відносної частоти затримки рейсів від її ймовірності буде меншою 0,01.

Розв'язання. За умовою $n = 1400$; $p = 0,07$; $q = 0,93$.

а) За формулою (14.6) при $\varepsilon = 0,02$:

$$P\left(\left|\frac{m}{n} - p\right| < 0,02\right) \geq 1 - \frac{0,07 \cdot 0,93}{1400 \cdot (0,02)^2} \approx 0,884.$$

б) Щоб ймовірність була не меншою 0,96, достатньо в формулі (14.6) покласти

$$1 - \frac{0,07 \cdot 0,93}{1400 \cdot \varepsilon^2} = 0,96, \text{ звідки } \varepsilon = 0,034.$$

Отже, з ймовірністю, не меншою 0,96, має місце подія

$$\left|\frac{m}{n} - 0,07\right| < 0,034 \text{ або } 0,036 < \frac{m}{n} < 0,104.$$

в) За формулою (14.6)

$$P\left(\left|\frac{m}{n} - p\right| < 0,01\right) \geq 1 - \frac{0,07 \cdot 0,93}{n \cdot (0,01)^2} \geq 0,9, \text{ звідки } n \geq 6510.$$

Відповідь. а) $P\left(\left|\frac{m}{n} - p\right| < 0,02\right) \approx 0,884$; б) $0,036 < \frac{m}{n} < 0,104$; в) $n \geq 6510$.

14.2. Інші граничні теореми

1. Теорема Маркова

Теорема Маркова формулюється наступним чином:

Теорема 14.5. Нехай $X_1, X_2, \dots, X_n$ – послідовність випадкових величин з скінченним математичним сподіванням $M(X_k) < +\infty$ для яких виконується умова:

$$\lim_{n \rightarrow +\infty} \left(\frac{1}{n^2} \cdot D(\sum_{k=1}^n X_k) \right) = 0. \quad (14.7)$$

Тоді виконується рівність (14.5) теореми Чебишева:

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \cdot \sum_{k=1}^n X_k - \frac{1}{n} \cdot \sum_{k=1}^n M(X_k)\right| < \varepsilon\right) = 1.$$

Доведення: З рівності (14.7), враховуючи властивість про обмеженість збіжної послідовності, випливає, що

$$\frac{1}{n^2} \cdot D(\sum_{k=1}^n X_k) = \frac{1}{n^2} \cdot \sum_{k=1}^n D(X_k) \leq \frac{n \cdot C}{n^2} \Rightarrow D(X_i) \leq C.$$

Таким чином виконуються умови теореми Чебишева, тобто справедлива рівність (14.5).

2. Центральна Гранична Теорема

Центральна Гранична Теорема (ЦГТ) є ключовим результатом, який говорить, що сума (або середнє) великої кількості незалежних випадкових величин, навіть якщо вони не є нормально розподіленими, буде

наближатися до нормального розподілу при збільшенні кількості спостережень.

Класичне формулювання

Теорема 14.6. Нехай $X_1, X_2, \dots, X_n$ — послідовність взаємно незалежних випадкових величин з однаковими розподілами, які мають скінченне математичне сподівання $a = M(X_k)$ та скінченну дисперсію $\sigma^2 = D(X_k)$.

Нехай $S_n = \sum_{k=1}^n X_k$. Тоді справедливо (14.8):

$$\sqrt{n} \cdot \left(\frac{S_n}{n} - a \right) \rightarrow N(0,1), n \rightarrow +\infty, \quad (14.8)$$

а для довільних фіксованих α, β ($\alpha < \beta$) справедливо (14.9):

$$P \left(\alpha < \sqrt{n} \cdot \frac{\frac{S_n}{n} - a}{\sigma} < \beta \right) \rightarrow \Phi(\beta) - \Phi(\alpha), \quad (14.9)$$

де $\Phi(x)$ — функція Лапласа.

Доведення. Розглянемо основні кроки доведення.

1. Побудова сум та нормалізація:

Розглянемо суму S_n :

$$S_n = \sum_{k=1}^n X_k.$$

Математичне сподівання:

$$M(S_n) = n \cdot a.$$

Дисперсія:

$$D(S_n) = n \cdot \sigma^2.$$

Нормалізуємо цю суму

$$Z_n = \frac{S_n - n \cdot a}{\sqrt{n} \cdot \sigma} = \sqrt{n} \cdot \frac{\frac{S_n}{n} - a}{\sigma}.$$

2. Характеристична функція:

Для доведення використовуємо *характеристичну функцію випадкової величини X_k* , тобто математичне сподівання випадкової величини $e^{(i \cdot t \cdot X_k)}$:

$$\psi(X_k(t)) = M(e^{(i \cdot t \cdot X_k)}), \quad (14.10)$$

де:

t — параметр;

i — комплексна одиниця.

Для S_n характеристична функція дорівнює добутку характеристичних функцій для незалежних $X_k = X$:

$$\psi_{S_n}(t) = \left(\psi_X \left(\frac{t}{\sqrt{n} \cdot \sigma} \right) \right)^n.$$

3. Розклад у ряд Тейлора:

Характеристична функція однієї випадкової величини $X_k = X$:

$$\psi_X(u) = 1 + i \cdot \mu \cdot u - \frac{\sigma^2 \cdot u^2}{2} + o(u^2).$$

Для нормалізованої величини $\frac{t}{\sqrt{n} \cdot \sigma}$

$$\psi_X\left(\frac{t}{\sqrt{n} \cdot \sigma}\right) = 1 + i \cdot \mu \cdot \frac{t}{\sqrt{n} \cdot \sigma} - \frac{\sigma^2}{2} \cdot \left(\frac{t}{\sqrt{n} \cdot \sigma}\right)^2 + o\left(\left(\frac{t}{\sqrt{n} \cdot \sigma}\right)^2\right).$$

4. Добуток характеристичних функцій:

Підставляємо в характеристичну функцію для S_n :

$$\psi_{S_n}(t) = \left(1 + i \cdot \mu \cdot \frac{t}{\sqrt{n} \cdot \sigma} - \frac{\sigma^2}{2} \cdot \left(\frac{t}{\sqrt{n} \cdot \sigma}\right)^2 + o\left(\left(\frac{t}{\sqrt{n} \cdot \sigma}\right)^2\right)\right)^n.$$

При $n \rightarrow +\infty$ матимемо:

$$\psi_{S_n}(t) \approx e^{-\left(\frac{t^2}{2}\right)}.$$

5. Граничний розподіл:

Характеристична функція $\psi_{S_n}(t)$ відповідає $N(0,1)$. Отже, справедливий перехід (14.8). Що й треба було довести.

6. Отримання виразу (14.9):

Вираз (14.9) отримано з формули (6.6) лекції 6, враховуючи функцію розподілу стандартизованого розподілу $N(0,1)$, яка має вигляд:

$$F(x) = 0,5 + \Phi(x).$$

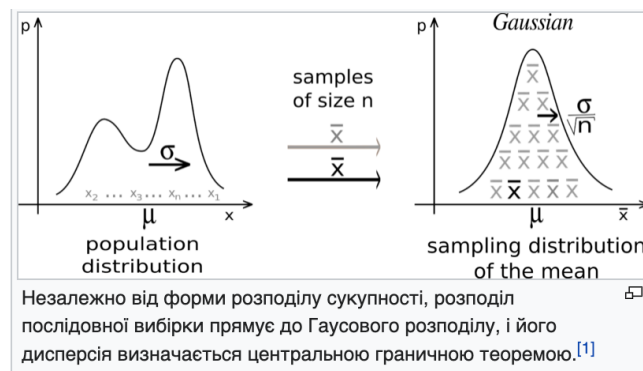


Рис. 14.1. Ілюстрація центральної граничної теореми

Інтерпретація і значення

Центральна Гранична Теорема пояснює поширення нормального закону розподілу, тобто: **якщо випадкова величина S_n формується під впливом багатьох незалежних факторів X_k ($k = \overline{1, n}$), кожен з яких здійснює на неї незначний вплив, то розподіл цієї величини близький до нормального розподілу при $n \rightarrow +\infty$.**

На практиці центральна гранична теорема реалізовується при $n = 40$.

Наслідок з центральної граничної теореми

Нехай $X_1, X_2, \dots, X_n, \dots$ – послідовність незалежних однаково розподілених випадкових величин зі скінченними математичними сподіваннями $a = M(X_k)$, дисперсією $\sigma^2(X_i) = D(X_i)$ і їх сума – $Y_n =$

$\sum_{k=1}^n X_k$, має математичне сподівання $M(Y_n) = n \cdot a$, та дисперсію $\sigma^2(Y_n) = D(Y_n) = n \cdot \sigma^2$. Тоді для будь-якого x справедлива рівність (14.10):

$$\lim_{n \rightarrow +\infty} \left(P \left(\frac{Y_n - M(Y_n)}{\sigma(Y_n)} < x \right) \right) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^x e^{-\frac{x^2}{2}} dx = 0,5 + \Phi(x). \quad (14.11)$$

Приклад 14.7. У касі деякої установи залишилася сума $d = 3500$ (грн). У черзі за одержанням грошей стоять 20 осіб. Сума X , яку потрібно виплатити окремій особі, – випадкова величина з математичним сподіванням $M(X) = 150$ (грн) і середнім квадратичним відхиленням $\sigma(X) = 60$ (грн). Знайти ймовірність того, що суми d не вистачить для виплати грошей усім особам, які стоять у черзі.

Розв’язання. На підставі центральної граничної теореми для однаково розподілених доданків X_i при великому n (а $n = 20$ практично можна вважати великим), випадкова величина $Y_n = \sum_{i=1}^n X_i$, де – сума, яку потрібно виплатити i -тій особі, має приблизно нормальний розподіл із параметрами:

$$M(Y_n) = n \cdot a; \sigma^2(Y_n) = D(Y_n).$$

або

$$M(Y_n) = 20 \cdot 150 = 3000; \sigma^2(Y_n) = n \cdot \sigma^2 = 20 \cdot 3600 = 72000; \sigma(Y_n) \approx 268.$$

$$\begin{aligned} P(Y_n > 3500) &= 1 - P(Y_n \leq 3500) \approx 1 - 0,5 - \Phi(1,87) = \\ &= 0,5 - \Phi\left(\frac{3500 - 3000}{268}\right) \approx 0,0307. \end{aligned}$$

Отже, з ймовірністю, близькою до 0,0307, наявної в касі суми грошей не вистачить для виплати всім бажаючим особам.

Відповідь. $P(Y_n > 3500) \approx 0,0307$.

Приклад застосування Центральної Граничної Теореми для прогнозування курсу валюти

Більш детально даний приклад розглянуто в статті [30].

Розглянемо вибірку курсу євро до австралійського долара. В якості статистичної вибірки X – вибірка значень курсу євро до австралійського долара за різні періоди з 1999-го року до кінця 2022-го. Об’єм вибірки n .

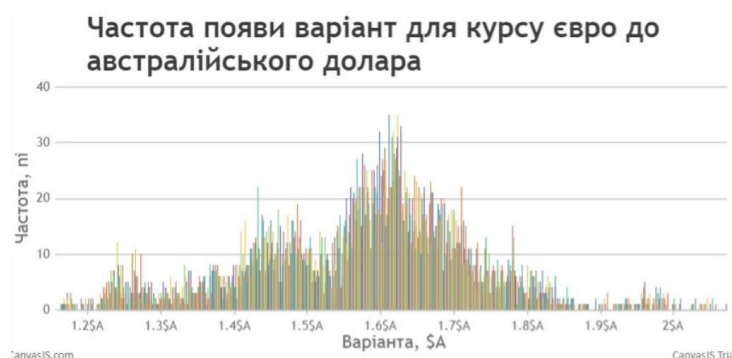


Рис.14.2. Детальний графік частот для курсу євро по відношенню до австралійського долара

Для дослідження було обрано саме австралійський долар, бо він є досить стабільною валютою і розмах його значень за вказаний період не є великим. Наприклад українська гривня має досить великий розмах за аналогічний період. Тому, цю валюту не рекомендується використовувати для даного дослідження.

Оскільки вибірка значень курсів валют є великою ($n = 756$), розмах вибірки X не є великим (дані не є розрідженими), а також на курс валюти впливають багато незалежних факторів, то Центральна Гранична Теорема є застосовною. Отже, згідно з теоремою можемо припустити, що генеральна сукупність, з якої зроблено вибірку X прямує до нормального розподілу.

Взагалі кажучи дану гіпотезу слід перевіряти за допомогою статистичних критеріїв.

Будемо вважати дану генеральну сукупність, випадковою величиною Y . Тоді можемо скласти функцію щільності для Y . Як відомо, для нормального розподілу вона має вигляд:

$$f(y) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(y-a)^2}{2 \cdot \sigma^2}}.$$

За допомогою методу моментів:

$$\begin{aligned} a &= \bar{x}_B = 1,577; \\ \sigma^2 &= s^2 = 0,0226; \\ \sigma &= 0,1503, \end{aligned}$$

де:

$\bar{x}_B$ – вибіркове середнє вибірки X ;

s^2 – виправлена дисперсія вибірки X (буде розглянуто в наступних лекціях).

Тоді

$$f(y) = \frac{1}{0,1503 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(y-1,577)^2}{2 \cdot 0,0226}} = \frac{1}{0,1503 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(y-1,577)^2}{0,0452}}.$$

За рис. 14.3 бачимо, що найбільша щільність ймовірності є для $Mo(Y) = a = 1,577$, тобто дане значення і буде прогнозом наступного значення для досліджуваної вибірки (наведено дані з 1999-го року по 14.11.2022). Перевіримо офіційне (фактичне) значення курсу на дату 15.11.2022 (рис. 14.4).

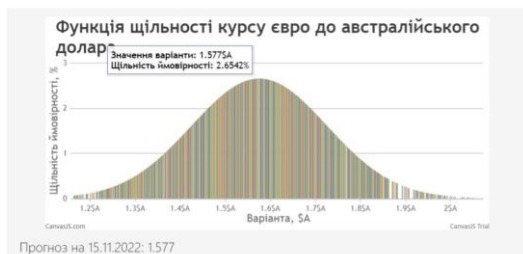


Рис. 14.3 Графік функції щільності розподілу значень курсу євро по відношенню австралійського до долара

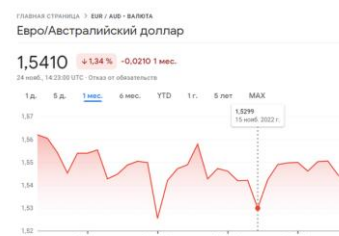


Рис. 14.4 Графік фактичних значень курсу євро по відношенню до австралійського долара

Похибка склала $\delta = 1,5770 - 1,5299 = 0,0471$, тобто прогноз має достатню точність і даний метод можна застосовувати для прогнозування стабільних курсів валют.

3. Теорема Ляпунова

Теорема Ляпунова є однією з важливих граничних теорем теорії ймовірностей, яка розширює умови Центральної Граничної Теорема (ЦГТ) для незалежних, але не обов'язково однаково розподілених випадкових величин. Вона встановлює умови, за яких сума незалежних випадкових величин наближається до нормального розподілу.

Класичне формулювання

Теорема 14.7. Якщо для послідовності попарно незалежних випадкових величин $X_1, X_2, \dots, X_n$ можна знайти таке число $\sigma > 0$, що

$$\lim_{n \rightarrow +\infty} \frac{\sum_{i=1}^n M|X_i - M(X_i)|^{2+\sigma}}{(\sqrt{\sum_{i=1}^n D(X_i)})^{2+\sigma}} = 0,$$

то існує границя (14.12):

$$\lim_{n \rightarrow +\infty} P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n M(X_i)}{\sqrt{\sum_{i=1}^n D(X_i)}} \rightarrow N(0, \sigma^2), n \rightarrow +\infty. \quad (14.12)$$

Інтерпретація і значення

Теорема Ляпунова є важливим узагальненням Центральної Граничної Теорема, оскільки вона не вимагає однаковості розподілів випадкових величин. Умова Ляпунова дозволяє встановити нормальний розподіл суми незалежних випадкових величин, за умови, що відповідні моменти вищого порядку мають обмежений вплив.

14.3. Інтегральна теорема Муавра-Лапласа

Повернемось до доведення теореми Муавра-Лапласа (теорема 4.4 лекції 4).

Теорема 4.4. Якщо ймовірність p появи події A в кожному випробуванні стала і відмінна від 0 і 1, то ймовірність того, що число m появ події A в n незалежних випробуваннях буде міститись у границях від m_1 до m_2 (включно), для достатньо великих n наближено дорівнює (формула (4.5) лекції 4):

$$P_n(m_1 \leq m \leq m_2) \approx \Phi(x_2) - \Phi(x_1),$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{t^2}{2}} dt$ – функція Лапласа, таблиця значень якої наведена в додатку 2 Гмурмана, де $x_1 = \frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}}$, $x_2 = \frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}}$.

Доведення. Вважатимемо, що розглянута випадкова величина має біноміальний розподіл.

Визначення біноміального розподілу:

$$P(X = m) = C_n^m \cdot p^m \cdot q^{n-m}.$$

Математичне очікування та дисперсія: Математичне очікування $M(X) = n \cdot p$ і дисперсія $D(X) = n \cdot p \cdot q$ є основними характеристиками для нормальної апроксимації.

Центральна гранична теорема: За умов великих n (зокрема, $n \cdot p > 5$ і $n \cdot q > 5$), згідно з центральною граничною теоремою, сума великої кількості незалежних випадкових величин (кожна з яких має біноміальний розподіл) буде наближена нормальним розподілом.

Апроксимація ймовірностей: Нехай S_n — сума n незалежних випадкових величин $\{X = m\}$, що розподілені за біноміальним законом. Згідно з центральною граничною теоремою, стандартний вигляд величини $Z = \frac{S_n - n \cdot p}{\sqrt{n \cdot p \cdot q}}$ буде асимптотично нормально розподілений: $Z \approx N(0,1)$.

Тому ймовірність $P_n(m_1 \leq m \leq m_2) = P(m_1 \leq X \leq m_2)$ можна оцінити через інтеграл від нормального розподілу:

$$P(m_1 \leq X \leq m_2) \approx P\left(\frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}} \leq Z \leq \frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right).$$

Остаточний результат: Цей підхід дає остаточну формулу для наближеної ймовірності:

$$P\left(\frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}} \leq Z \leq \frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right) \approx \Phi\left(\frac{m_2 - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right) - \Phi\left(\frac{m_1 - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right).$$

Теорема Муавра-Лапласа використовується для уточнення результатів отриманих, за допомогою нерівності Чебишева.

Приклад 14.8. Нехай X — випадкова величина з математичним сподіванням $M(X) = 50$ і стандартним відхиленням $\sigma = 5$. Оцінити ймовірність того, що X відхилиться від свого математичного сподівання менше, ніж на 3 стандартні відхилення (записати довірчий інтервал).

Розв'язання. За нерівністю Чебишева (14.4):

$$P(|X - 50| < 15) \geq 1 - \frac{25}{225} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0,889.$$

Це означає, що довірчий інтервал для шуканої ймовірності — $[0,889; 1]$.

Уточнимо результат за теоремою Муавра-Лапласа (формула (4.5) лекції 4):

$$\begin{aligned} & |X - 50| < 15; \quad 35 < X < 65. \\ & x_1 = \frac{35 - 50}{5} = -3, \quad x_2 = \frac{65 - 50}{5} = 3 \\ & P(|X - 50| < 15) = P(35 < X < 65) \approx \Phi(3) - \Phi(-3) = 2 \cdot \Phi(3) = \\ & \quad = 2 \cdot 0,49865 = 0,9973. \end{aligned}$$

Це пояснюється тим, що інтегральна теорема враховує нормальний розподіл і є точнішою, ніж нерівність Чебишева, яка задає лише верхню або нижню межі ймовірності.

З чого робимо висновок, що наближене значення для шуканої ймовірності 0,9973 попадає в довірчий інтервал $[0,889; 1]$.

Відповідь. $[0,889; 1]$.

Приклад 14.9. В інкубаторі вилуплюється в середньому курчат 70% від кількості закладених яєць. Скільки потрібно закласти яєць, щоб з ймовірністю не меншою від 0,95 очікувати, що відхилення вилуплених курчат від математичного сподівання не перевищувало 50 (за абсолютною величиною)? Розв'язати задачу за допомогою **а)** нерівності Чебишева; **б)** інтегральної теореми Муавра-Лапласа.

Розв'язання. З умови задачі: $p = 0,7$; $q = 0,3$; $\varepsilon = 50$.

$$M(X) = 0,7 \cdot n; D(X) = 0,21 \cdot n.$$

а) З нерівності Чебишева (14.4):

$$P(|X - 0,7 \cdot n| < 50) \geq 1 - \frac{0,21 \cdot n}{2500} \geq 0,95.$$

$$1 - \frac{0,21 \cdot n}{2500} \geq 0,95 \Rightarrow \frac{0,21 \cdot n}{2500} \leq 0,05 \Rightarrow n \leq 595.$$

б) За наслідком з інтегральної теореми Муавра-Лапласа (формула (4.5) лекції 4):

$$P(|X - 0,7 \cdot n| < 50) \approx 2 \cdot \Phi\left(\frac{50}{\sqrt{0,21 \cdot n}}\right) \geq 0,95.$$

$$\Phi\left(\frac{50}{\sqrt{0,21 \cdot n}}\right) \geq 0,475.$$

$$\frac{50}{\sqrt{0,21 \cdot n}} \geq 1,96.$$

$$\sqrt{0,21 \cdot n} \leq 25,51.$$

$$n \leq 3099.$$

В даному випадку точнішою виявилася нерівність Чебишева. Цей факт легко перевірити наступним чином:

$$\text{При } n = 600 \quad P(|X - 420| < 50) \geq 1 - \frac{126}{2500} \approx 0,9496 < 0,95;$$

$$\text{При } n = 3099 \quad P(|X - 2169,3| < 50) \geq 1 - \frac{650,79}{2500} \approx 0,7397 < 0,95;$$

$$\text{При } n = 595 \quad P(|X - 416,5| < 50) \geq 1 - \frac{124,95}{2500} \approx 0,95002 > 0,95;$$

$$\text{При } n = 100 \quad P(|X - 70| < 50) \geq 1 - \frac{21}{2500} \approx 0,9916 > 0,95;$$

$$\text{При } n = 1 \quad P(|X - 0,7| < 50) \geq 1 - \frac{0,21}{2500} \approx 0,999916 > 0,95;$$

Відповідь. $n \leq 595$.

ЛЕКЦІЯ 15

ВИПАДКОВІ ПРОЦЕСИ, МАРКІВСЬКІ ПРОЦЕСИ, ПОТОКИ ПОДІЙ. ТЕОРІЯ МАСОВОГО ОБСЛУГОВУВАННЯ

15.1. Основні поняття

Випадкова функція (випадковий процес) $X(t)$ – це функція, яка в результаті експерименту може прийняти той або інший заздалегідь невідомий конкретний вигляд. Цей конкретний вигляд, який приймає випадкова функція в результаті експерименту, називається *реалізацією випадкової функції (випадкового процесу)*.

Переріз випадкової функції – це випадкова величина при фіксованому t .

Одновимірним законом розподілу випадкової величини $X(t)$ називається закон розподілу $f(x, t)$ перерізу $x(t)$ випадкової функції.

Випадкова функція $X(t)$ називається *нормальною*, якщо закон розподілу системи будь-якої кількості числа n її перерізів є n -вимірний нормальний закон.

Математичним сподіванням випадкової функції $X(t)$ називається не випадкова функція $M_x(t)$, яка при кожному t є математичне сподівання відповідного перерізу випадкової функції (15.1):

$$M_x(t) = M(X(t)). \quad (15.1)$$

Дисперсія випадкової функції $X(t)$ є не випадкова величина $D_x(t)$, яка при будь-якому значенні змінної t дорівнює дисперсії відповідного перерізу випадкової функції $X(t)$ (15.2):

$$D_x(t) = D(X(t)). \quad (15.2)$$

Середнє квадратичне відхилення $\sigma_x(t)$ випадкової функції $X(t)$ дорівнює (15.3):

$$\sigma_x(t) = \sqrt{D_x(t)}. \quad (15.3)$$

Математичне сподівання випадкової функції характеризує середню траєкторію всіх можливих її реалізацій, а середнє квадратичне відхилення – розкид реалізацій відносно середньої траєкторії.

Кореляційною функцією випадкового процесу $X(t)$ називається не випадкова функція (15.4):

$$K_x(t_1, t_2) = M\left((X(t_1) - M_x(t_1)) \cdot (X(t_2) - M_x(t_2))\right), \quad (15.4)$$

яка характеризує ступень тісноти лінійної залежності між перерізами, а також розкид цих перерізів відносно математичного сподівання.

Нормованою кореляційною функцією випадкового процесу $X(t)$ називається функція $\rho_x(t_1, t_2)$, яка обчислюється за формулою (15.5):

$$\rho_x(t_1, t_2) = \frac{K_x(t_1, t_2)}{\sigma_x(t_1) \cdot \sigma_x(t_2)}. \quad (15.5)$$

Приклад 15.1. Випадкова функція $X(t)$ є ступінчата знакозмінна функція (рис. 15.1), яка через одиничні інтервали приймає значення: $+1, -1$. Положення ступінчатої функції відносно початку відліку випадково; випадкова величина T , що характеризує зсув першої точки зміни знаку відносно початку координат є випадковою величиною, розподіленою рівномірно в інтервалі $(0,1)$. Знайти математичне сподівання $M_x(t)$, дисперсію $D_x(t)$ і кореляційну функцію $K_x(t_1, t_2)$.

Розв'язання. Розглянемо переріз випадкової функції $X(t)$. Він може з однаковою ймовірністю попасти як на ділянку, де випадкова функція дорівнює 1, так і на ділянку, де вона рівна -1 . Тому ряд розподілу для любого перерізу має вигляд:

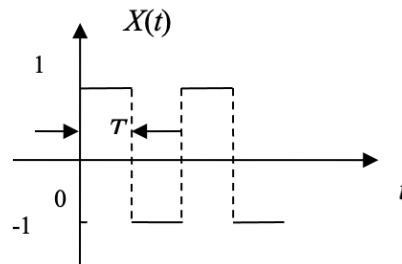


Рис.15.1. Графік функції $X(t)$

Ряд розподілу випадкової функції наведено в табл. 15.1:

Таблиця 15.1. Ряд розподілу $X(t)$

$x_i(t)$	-1	1
$p_i(t)$	0,5	0,5

Числові характеристики:

$$M_x(t) = (-1) \cdot 0,5 + 1 \cdot 0,5 = 0,$$

$$D_x(t) = (-1)^2 \cdot 0,5 + 1^2 \cdot 0,5 = 1$$

$$K_x(t, t - \tau) = M(X(t) \cdot X(t - \tau)).$$

Оскільки, добуток $X(t) \cdot X(t + \tau)$ може приймати тільки два значення 1 або -1 , то $M(X(t) \cdot X(t + \tau)) = 1 \cdot p_1 + (-1) \cdot (1 - p_1) = 2 \cdot p_1 - 1$, де p_1 – ймовірність того, що точки t і $t + \tau$ попадуть на ділянки, в яких $X(t)$ та $X(t + \tau)$ мають однаковий знак. Через рівномірності розподілу зсуву T можна перенести початок відліку до лівого кінця тої ділянки, на якій знаходиться точка t_i вважати, що точка t рівномірно розподілена в інтервалі $(0,1)$. Тоді ймовірність p_1 є ймовірність того, $t + \tau$ попадає в деякий з інтервалів, $(2n, 2n + 1)$; $n = 0, \pm 1, \pm 2, \dots$

Обчислимо цю ймовірність для різних значень τ .

При $0 \leq \tau \leq 1$ точка $t + \tau$ може попасти в інтервал $(0; 1)$, або в інтервал $(1; 2)$.

Тому

$$p_1 = p(t + \tau < 1) = p(t + \tau < 1) = p(t < 1 - \tau) = 1 - \tau.$$

При $1 \leq \tau \leq 2$ точка $(t + \tau)$ може попасти в інтервал $(1; 2)$, або в інтервал $(2; 3)$.

Тому

$$p_1 = p(t + \tau > 2) = p(t < 2 - \tau) = 1 - (2 - \tau) = \tau - 1.$$

Якщо продовжити ці міркування, то

$$p_1 = \begin{cases} -1 - (\tau - 2 \cdot n), 2 \cdot n < \tau < 2 \cdot n + 1, \\ (\tau - 2 \cdot n) - 1, 2 \cdot n + 1 < \tau < 2 \cdot n + 2. \end{cases}$$

$$K_x(t, t + \tau) = k_x(\tau) = \begin{cases} 4 \cdot n + 1 - 2 \cdot \tau, 2 \cdot n < \tau < 2 \cdot n + 1, \\ 2 \cdot \tau - (4 \cdot n + 3), 2 \cdot n + 1 < \tau < 2 \cdot n + 2. \end{cases}$$

Графік кореляційної функції $K_x(t, t + \tau)$ представлено на рис. 8.2.

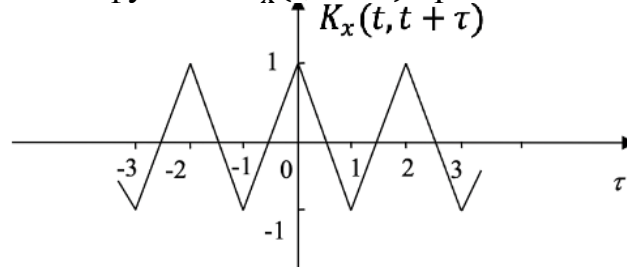


Рис. 15.2. Графік кореляційної функції $K_x(t, t + \tau)$

Відповідь. $M_x(t) = 0$; $D_x(t) = 1$; $K_x(t, t + \tau) = \begin{cases} 4 \cdot n + 1 - 2 \cdot \tau, 2 \cdot n < \tau < 2 \cdot n + 1, \\ 2 \cdot \tau - (4 \cdot n + 3), 2 \cdot n + 1 < \tau < 2 \cdot n + 2. \end{cases}$

Випадковий процес є процес переходу системи зі стану в стан під впливом випадкових факторів.

Система X називається **системою з дискретним станом**, якщо вона має злічену (в частковому випадку – скінчену) множину можливих станів $X_1, X_2, \dots, X_n$ і перехід з одного стану в інший відбувається стрибком. Можливі стани системи X зображуються, за допомогою графіків станів, на якому стани системи зображуються прямокутниками, а можливі переходи системи зі стану в стан – стрілками, які з'єднують відповідні прямокутники. Для того, щоб описати випадковий процес – в системі з дискретними станами, використовують ймовірності станів $p_1(t), p_2(t), \dots, p_n(t)$, де $p_k(t)$ ($k = 1, \dots, n$) – ймовірність того, що в момент t система знаходиться в стані X_k причому $\sum_{k=1}^n p_k(t) = 1$.

Випадковий процес, який відбувається в системі X , називається **процесом з дискретним часом**, якщо перехід системи зі стану в стан можливий тільки у визначені моменти часу $t_1, t_2, \dots, t_n$

Якщо переходи можливі в будь-який момент часу, то процес називається **процесом з неперервним часом**.

Потік подій – це послідовність подій, які відбуваються один за одним у випадкові моменти часу.

Щільністю (інтенсивністю) потоку подій називається середнє число подій в одиницю часу.

Потік подій називається **потокм без післядії**, якщо ймовірність появи на будь-якому проміжку часу того або іншого числа подій не залежить від того, яке число подій попало на інші проміжки, які не перетинаються з даним проміжком. Наприклад, потік пасажирів, які входять у метро.

Потік подій називається *ординарним*, якщо ймовірність появи на елементарному проміжку часу Δt двох або більш подій мала в порівнянні з ймовірністю появи однієї події, тобто події з'являються на проміжку Δt поодиночки, а не групами. Наприклад, потік потягів, які під'їжджають до станції, ординарне, а потік вагонів – ні.

Потік подій називається *регулярним*, якщо події сліднують одна за однієї через рівні визначені проміжки часу. Наприклад, потік виробів на конвеєрі.

Потік подій називається *стаціонарним*, якщо його ймовірності характеристики не залежать від часу. Наприклад, потік автомобілів на міській вулиці в час пік, протягом доби цей потік нестаціонарний.

Ординарний потік подій без післядії називається *пуассонівським потоком*. Якщо пуассоновський потік стаціонарне, то він називається *стаціонарним пуассонівським* або *найпростішим*.

Випадковий процес з дискретними станами називається *марківським*, якщо його ймовірнісні характеристики в майбутньому залежать тільки від того, в якому стані цей процес знаходиться в теперішній момент часу і не залежить від того, яким чином цей процес відбувався в минулому (майбутнє залежить від минулого тільки через теперішнє).

Якщо процес марківським, то всі потоки подій, які переводять систему зі стану до стану, є пуассонівськими.

15.2. Система рівнянь Колмогорова

Система рівнянь Колмогорова описує еволюцію ймовірностей станів у стохастичних процесах, особливо у процесах із марковською властивістю. Вона широко використовується в теорії масового обслуговування та інших областях.

Якщо процес є марківським, то для ймовірностей станів $p_1(t), p_2(t), \dots, p_n(t)$ можна скласти систему диференціальних рівнянь (рівняння Колмогорова). Для цього використовується розмічений граф станів системи, на якому проти кожної стрілки, що з'єднує стани, проставлена інтенсивність потоку подій. Інтенсивність потоку подій, яка переводить систему зі стану X_i до стану X_j позначається λ_{ij} .

Правило складання рівнянь Колмогорова

В лівій частині кожного з рівнянь стоїть похідна ймовірності i -го стану $\frac{dp_i(t)}{dt}$.

В правій частині – стільки членів, скільки стрілок зв'язано безпосередньо з даним станом; якщо стрілка веде до цього стану, член має знак « + », якщо веде з цього стану, то член має знак « - ». Кожен доданок дорівнює добутку інтенсивності потоку подій, який переводить систему за даній стрілки, на ймовірність того стану, з якого стрілка виходить.

Останнім рівнянням системи є умова:

$$\sum_{i=1}^n p_i = 1.$$

Число рівнянь в системі може бути зменшено на одиницю за рахунок останнього рівняння (якесь з рівнянь може бути лінійною комбінацією іншого).

Початкові умови для інтегрування системи відображають стан системи в початковий момент.

Якщо система при $t = 0$ була в стані X_i , то $p_i(0) = 1$, $p_j(0) = 0$, $i \neq j$

Граничним режимом для системи X називається випадковий процес, який утворюється в системі при $t \rightarrow \infty$. Він характеризується тим, що система в даний момент часу перестає залежати від часу (має сталі ймовірності).

Оскільки граничні ймовірності станів є сталі, то в лівій частині маємо нулі, а система диференціальних рівнянь перетворюється на систему алгебраїчних рівнянь.

Регулярним потоком подій називається потік, в якому події слідуєть одна за одною через строго визначені проміжки часу.

Приклад 15.2. Записати рівняння Колмогорова та систему рівнянь граничних станів для графа на рис. 15.3.

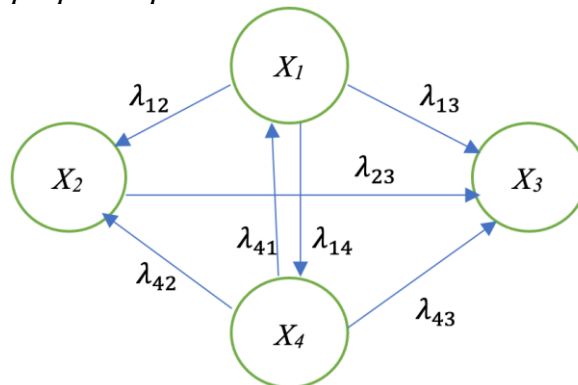


Рис. 15.3. Граф станів процесу

Розв'язання. Згідно правила система рівнянь Колмогорова для рис. 15.3 має вигляд (15.6):

$$\begin{cases} \frac{dp_1(t)}{dt} = -(\lambda_{12} + \lambda_{13} + \lambda_{14}) \cdot p_1(t) + \lambda_{41} \cdot p_4(t); \\ \frac{dp_2(t)}{dt} = \lambda_{12} \cdot p_1(t) + \lambda_{42} \cdot p_4(t) - \lambda_{23} \cdot p_2(t); \\ \frac{dp_3(t)}{dt} = \lambda_{13} \cdot p_1(t) + \lambda_{23} \cdot p_2(t) + \lambda_{43} \cdot p_4(t); \\ \frac{dp_4(t)}{dt} = \lambda_{14} \cdot p_1(t) - (\lambda_{41} + \lambda_{42} + \lambda_{43}) \cdot p_4(t); \\ p_1(t) + p_2(t) + p_3(t) + p_4(t) = 1. \end{cases} \quad (15.6)$$

Для системи X , розмічений граф станів якої представлено на рис. 15.3, система алгебраїчних рівнянь, яка визначає граничний режим, буде мати вигляд (15.7):

$$\begin{cases} 0 = -(\lambda_{12} + \lambda_{13} + \lambda_{14}) \cdot p_1 + \lambda_{41} \cdot p_4; \\ 0 = \lambda_{12} \cdot p_1 + \lambda_{42} \cdot p_4 - \lambda_{23} \cdot p_2; \\ 0 = \lambda_{13} \cdot p_1 + \lambda_{23} \cdot p_2 + \lambda_{43} \cdot p_4; \\ 0 = \lambda_{14} \cdot p_1 - (\lambda_{41} + \lambda_{42} + \lambda_{43}) \cdot p_4; \\ p_1 + p_2 + p_3 + p_4 = 1, \end{cases} \quad (15.7)$$

де $p_1(t) = p_1, p_2(t) = p_2, p_3(t) = p_3, \dots, p_n(t) = p_n$.

Відповідь: Системи (15.6) та (15.7).

15.3. Процес «загибелі та розмноження»

Процес «загибелі та розмноження» є частиною теорії марковських процесів. Він широко використовується для моделювання численних систем, де елементи можуть народжуватися або гинути з певними ймовірностями.

Основні поняття

Визначення

Процес «загибелі та розмноження» (birth-death process) – це вид марковського процесу, в якому стан системи змінюється шляхом додавання (народження) або видалення (смерті) одиниці.

Ймовірності

λ_{n-1n} – ймовірність народження (перехід від стану $n - 1$ до n);

λ_{nn-1} – ймовірність смерті (перехід від стану n до $n - 1$).

Стаціонарний розподіл

Для визначення стаціонарного розподілу необхідно вирішити рівняння балансу потоків для кожного стану системи.

Процес «загибелі та розмноження» має вигляд на рис. 15.4:

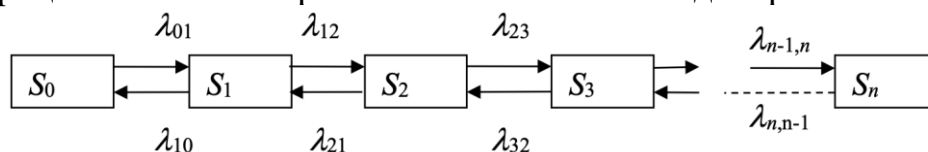


Рис. 15.4. Граф станів процесу «загибелі та розмноження»

Застосування

Процес «загибелі та розмноження» використовується в різних галузях, таких як:

- *Популяційна динаміка:* Моделювання росту та зменшення популяції;
- *Чергові системи:* Аналіз черг в обслуговуючих системах;
- *Епідеміологія:* Моделювання розповсюдження інфекційних захворювань.

Система рівнянь граничних станів

Записуючи рівняння для граничних ймовірностей. Маємо систему алгебраїчних рівнянь (15.8):

$$\begin{cases} \lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1; \\ \lambda_{12} \cdot p_1 = \lambda_{21} \cdot p_2; \\ \lambda_{23} \cdot p_2 = \lambda_{32} \cdot p_3; \\ \dots \\ \lambda_{n-1n} \cdot p_{n-1} = \lambda_{nn-1} \cdot p_n; \\ p_1 + p_2 + p_3 + \dots + p_n = 1. \end{cases} \quad (15.8)$$

Тоді ймовірності $p_1, p_2, p_3, \dots, p_n$ обчислюються зі співвідношень (15.9):

$$\begin{aligned} p_0 &= \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12} \cdot \lambda_{01}}{\lambda_{21} \cdot \lambda_{10}} + \frac{\lambda_{23} \cdot \lambda_{12} \cdot \lambda_{01}}{\lambda_{32} \cdot \lambda_{21} \cdot \lambda_{10}} + \dots + \frac{\lambda_{n-1n} \cdot \dots \cdot \lambda_{23} \cdot \lambda_{12} \cdot \lambda_{01}}{\lambda_{nn-1} \cdot \dots \cdot \lambda_{32} \cdot \lambda_{21} \cdot \lambda_{10}} \right)^{-1}; \\ p_1 &= \frac{\lambda_{01}}{\lambda_{10}} \cdot p_0; \quad p_2 = \frac{\lambda_{12} \cdot \lambda_{01}}{\lambda_{21} \cdot \lambda_{10}} \cdot p_0; \quad p_3 = \frac{\lambda_{23} \cdot \lambda_{12} \cdot \lambda_{01}}{\lambda_{32} \cdot \lambda_{21} \cdot \lambda_{10}} \cdot p_0; \quad \dots; \\ p_n &= \frac{\lambda_{n-1n} \cdot \dots \cdot \lambda_{23} \cdot \lambda_{12} \cdot \lambda_{01}}{\lambda_{nn-1} \cdot \dots \cdot \lambda_{32} \cdot \lambda_{21} \cdot \lambda_{10}} \cdot p_0. \end{aligned} \quad (15.9)$$

Приклад 15.3. Процес «загибелі та розмноження» представлено графом (рис. 15.5). Знайти граничні ймовірності станів.

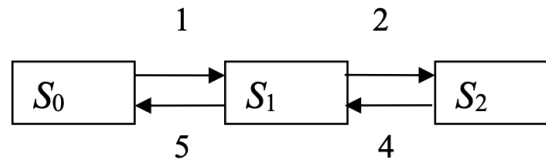


Рис. 15.5. Граф станів процесу «загибелі та розмноження»

Розв’язання. За формулами (15.9) отримуємо:

$$\begin{aligned} p_0 &= \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12} \cdot \lambda_{01}}{\lambda_{21} \cdot \lambda_{10}} \right)^{-1} = \left(1 + \frac{1}{5} + \frac{2 \cdot 1}{4 \cdot 5} \right)^{-1} = \frac{10}{13}; \\ p_1 &= \frac{\lambda_{01}}{\lambda_{10}} \cdot p_0 = \frac{1}{5} \cdot \frac{10}{13} = \frac{2}{13}; \\ p_2 &= \frac{\lambda_{12} \cdot \lambda_{01}}{\lambda_{21} \cdot \lambda_{10}} \cdot p_0 = \frac{2 \cdot 1}{4 \cdot 5} \cdot \frac{10}{13} = \frac{1}{13}. \end{aligned}$$

Відповідь. $p_0 = \frac{10}{13}; p_1 = \frac{2}{13}; p_2 = \frac{1}{13}.$

15.4. Циклічний процес

Циклічні процеси в контексті марківських процесів можуть розглядатися як тип процесів, де система повертається до початкового стану через певний період часу. Марківські процеси мають властивість, що майбутній стан системи залежить лише від її поточного стану, а не від попередніх станів (властивість марківськості).

Циклічні процеси в марківських ланцюгах

Для циклічних процесів у марківських ланцюгах важливо розглянути періодичність системи.

Періодичність

Стан i має період d , якщо повернення до стану i можливе лише через кратні d кроків.

Ланцюг називається *періодичним*, якщо всі стани мають один і той же період.

Аперіодичні ланцюги

Ланцюг є *аперіодичним*, якщо $d = 1$, тобто система може повернутися до будь-якого стану через будь-яку кількість кроків.

Стаціонарний розподіл:

- Розподіл ймовірностей по станах, який не змінюється з часом;
- Існує для аперіодичних та ергодичних марківських ланцюгів.

Циклічні марківські процеси

Ергодичні ланцюги

- Ланцюги, які є як аперіодичними, так і незвідними;
- Вони мають єдиний стаціонарний розподіл.

Застосування

Моделювання систем з регулярними повтореннями (наприклад, обслуговування клієнтів, процеси виробництва тощо).

Граф станів циклічного процесу має вигляд (рис. 15.6).

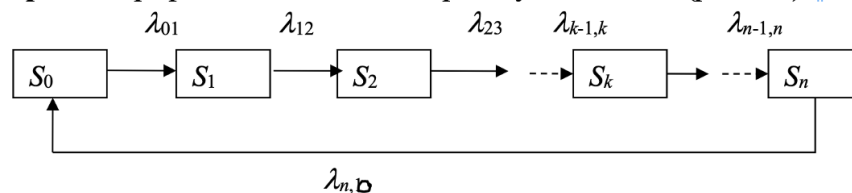


Рис. 15.6. Граф станів циклічного процесу

Система алгебраїчних рівнянь для цього графу має вигляд (15.10):

$$\begin{cases} \lambda_{01} \cdot p_0 = \lambda_{n0} \cdot p_n; \\ \lambda_{12} \cdot p_1 = \lambda_{01} \cdot p_0; \\ \lambda_{23} \cdot p_2 = \lambda_{12} \cdot p_1; \\ \dots \\ \lambda_{n-1,n} \cdot p_{n-1} = \lambda_{n0} \cdot p_n; \\ p_1 + p_2 + p_3 + \dots + p_n = 1. \end{cases} \quad (15.10)$$

Тоді ймовірності $p_1, p_2, p_3, \dots, p_n$ обчислюються зі співвідношень (15.11):

$$\begin{aligned} p_0 &= \left(1 + \lambda_{01} \cdot \left(\frac{1}{\lambda_{12}} + \frac{1}{\lambda_{23}} + \frac{1}{\lambda_{34}} + \dots + \frac{1}{\lambda_{n0}} \right) \right)^{-1}; \\ p_1 &= \frac{\lambda_{01}}{\lambda_{12}} \cdot p_0; \quad p_2 = \frac{\lambda_{01}}{\lambda_{23}} \cdot p_0; \quad p_3 = \frac{\lambda_{01}}{\lambda_{34}} \cdot p_0; \quad \dots; \\ p_n(t) &= \frac{\lambda_{01}}{\lambda_{n0}} \cdot p_0. \end{aligned} \quad (15.11)$$

Циклічні процеси в марківських системах є важливим інструментом для моделювання та аналізу систем з регулярними періодичними змінами. Вивчення таких процесів дозволяє краще розуміти та передбачати поведінку складних динамічних систем.

Приклад 15.4. Циклічний процес представлено графом (рис. 15.7). Знайти граничні ймовірності станів.

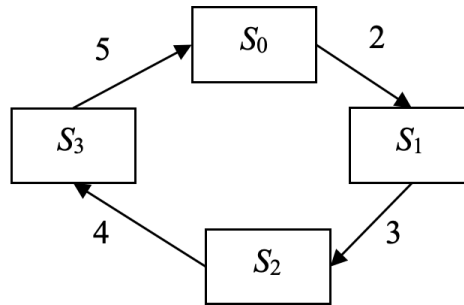


Рис. 15.7. Граф станів циклічного процесу

Розв'язання. За формулами (15.11) отримуємо:

$$p_0(t) = \left(1 + \lambda_{01} \cdot \left(\frac{1}{\lambda_{12}} + \frac{1}{\lambda_{23}} + \frac{1}{\lambda_{30}} \right) \right)^{-1} = \left(1 + 2 \cdot \left(\frac{1}{3} + \frac{1}{4} + \frac{1}{5} \right) \right)^{-1} = \frac{30}{77};$$

$$p_1(t) = \frac{\lambda_{01}}{\lambda_{12}} \cdot p_0(t) = \frac{2}{3} \cdot \frac{30}{77} = \frac{20}{77}; p_2(t) = \frac{1}{2} \cdot \frac{30}{77} = \frac{15}{77}; p_3(t) = \frac{2}{5} \cdot \frac{30}{77} = \frac{12}{77}.$$

Відповідь. $p_0(t) = \frac{30}{77}; p_1(t) = \frac{20}{77}; p_2(t) = \frac{15}{77}; p_3(t) = \frac{12}{77}.$

15.5. Матриця переходів

Припустимо, що проводиться серія експериментів з можливими результатами $X_1, X_2, \dots, X_n$, які називаються *станами*.

Введемо позначення:

$p_i^{(0)}$ – ймовірність того, що ми починаємо в стані X_i ;

p_{ij} – ймовірність того, що в результаті експеримента система перейшла зі стану X_i в стан X_j ;

$p_i^{(1)}$ – ймовірність того, що результатом експерименту буде стан X_i ;

$$p_i^{(1)} = p_1^{(0)} \cdot p_{i1} + p_2^{(0)} \cdot p_{i2} + p_3^{(0)} \cdot p_{i3} + \dots + p_n^{(0)} \cdot p_{in}.$$

Очевидно, що

$$p_{i1} + p_{i2} + p_{i3} + \dots + p_{in} = 1.$$

Матриця вигляду (15.12):

$$T = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \dots & p_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & p_{n3} & \dots & p_{nn} \end{bmatrix} \quad (15.12)$$

називається *матрицею переходу*. Позначимо її T .

Приклади задач

15.4. Побудувати граф переходів марківського процесу, що заданий матрицею переходів.

$$T = \begin{bmatrix} 0,5 & 0 & 0,5 & 0 & 0 \\ 0,25 & 0,5 & 0,25 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 0,5 & 0,5 \\ 0 & 0 & 0 & 0,5 & 0,5 \end{bmatrix}.$$

Розв'язання. Граф станів матиме вигляд (рис.15.8):

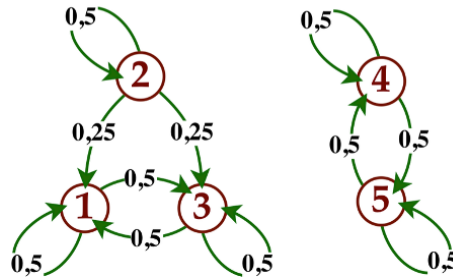


Рис. 15.8. Граф станів

Відповідь. Рис. 15.8.

Теорема 15.1. Нехай

$$p^{(0)} = (p_1^{(0)}; p_2^{(0)}; p_3^{(0)}; \dots; p_n^{(0)})$$

та

$$p^{(1)} = (p_1^{(1)}; p_2^{(1)}; p_3^{(1)}; \dots; p_n^{(1)}),$$

тоді справедлива рівність (15.13):

$$(p_1^{(1)}; p_2^{(1)}; p_3^{(1)}; \dots; p_n^{(1)}) = (p_1^{(0)}; p_2^{(0)}; p_3^{(0)}; \dots; p_n^{(0)}) \cdot \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \dots & p_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & p_{n3} & \dots & p_{nn} \end{bmatrix}. \quad (15.13)$$

Приклад 15.5. Погода класифікується як ясна, помірно похмура та похмура.

1. Якщо погода ясна, то з ймовірністю 0,5 вона буде ясною, з ймовірністю 0,4 – помірно похмурою, з ймовірністю 0,1 – похмурою;

2. Якщо погода помірно похмура, то з ймовірністю 0,3 вона буде ясною, з ймовірністю 0,5 – помірно похмурою, з ймовірністю 0,2 – похмурою;

3. Якщо погода похмура то з ймовірністю 0,2 вона буде ясною, з ймовірністю 0,4 – помірно похмурою, з ймовірністю 0,4 – похмурою.

1. Якщо ймовірність ясної погоди в неділю дорівнює 0,6, а ймовірність помірно похмурої – 0,4, то яка ймовірність, що погода в понеділок буде ясною?

2. Яка ймовірність що у вівторок погода буде помірно похмурою?

Розв'язання. 1. Згідно умови

$$p^{(0)} = (0,6; 0,4; 0),$$

$$T = \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,5 & 0,2 \\ 0,2 & 0,4 & 0,4 \end{bmatrix}.$$

Відповідно

$$p^{(1)} = (0,6; 0,4; 0) \cdot \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,5 & 0,2 \\ 0,2 & 0,4 & 0,4 \end{bmatrix} = (0,42; 0,44; 0,14)$$

та ймовірність, що в понеділок погода буде ясною дорівнює 0,42.

2. Нехай $p_1^{(2)}$ – ймовірність того, що погода у вівторок буде ясною, $p_2^{(2)}$ – ймовірність того, що погода у вівторок буде помірно похмурою, $p_3^{(2)}$ – ймовірність того, що погода у вівторок буде похмурою.

Нехай $p^{(2)} = (p_1^{(2)}, p_2^{(2)}, p_3^{(2)})$.

$$p^{(2)} = (0,42; 0,44; 0,14) \cdot \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,5 & 0,2 \\ 0,2 & 0,4 & 0,4 \end{bmatrix} = (0,37; 0,444; 0,186)$$

та ймовірність, що у вівторок погода буде помірно похмурою дорівнює 0,444.

Відповідь. 1. Ймовірність, що в понеділок погода буде ясною дорівнює 0,42; 2. Ймовірність, що у вівторок погода буде помірно похмурою дорівнює 0,444.

Теорема 15.2. Нехай $p_i^{(m)}$ – ймовірність того, що результатом m -го експерименту буде стан X_i та

$$p^{(m)} = (p_1^{(m)}; p_2^{(m)}; p_3^{(m)}; \dots; p_n^{(m)}).$$

Тоді для довільного цілого числа m справедлива рівність (15.14):

$$p^{(m)} = p^{(0)} \cdot T^m. \quad (15.14)$$

Ланцюги Маркова також застосовуються в оцінці майбутніх продажів. Наприклад, зробивши опитування серед покупців тієї чи іншої марки автомобіля про їхній наступний вибір, можна скласти матрицю T .

Приклад 15.6. У процесі опитування власників автомобілів трьох американських марок: марки А, марки В, марки С, їм було поставлено питання про те, яку торгову марку вони вибрали б для наступної покупки.

1. Серед власників автомобілів марки А 20% сказали, що виберуть знову цю ж марку, 50% сказали, що вони перейшли б на марку В, а 30% заявили, що віддали б перевагу марці С;

2. Серед власників автомобілів марки В 20% сказали, що перейдуть на марку А, натомість 70% заявили, що придбали б знову автомобіль марки В, а 10% заявили, що наступного разу віддали б перевагу марці С;

3. Серед власників автомобілів 30% відповіли, що перейшли б на марку А, 30% сказали, що перейшли б на марку В, а 40% заявили, що залишилися б вірними тій самій марці С.

1. Якщо хтось придбав автомобіль марки А, то яка ймовірність, що його другою машиною буде автомобіль марки С?

2. Якщо при покупці першої машини покупець підкинув монету, вибираючи між автомобілями марки В і С, то яка ймовірність, що його третьою машиною стане автомобіль марки В?

Розв'язання. 1. Матриця переходу має вигляд:

$$T = \begin{bmatrix} 0,2 & 0,5 & 0,3 \\ 0,2 & 0,7 & 0,1 \\ 0,3 & 0,3 & 0,4 \end{bmatrix}.$$

Для відповіді на перше питання матимемо $p^{(0)} = (1, 0, 0)$, тому:

$$p^{(1)} = (1, 0, 0) \cdot \begin{bmatrix} 0,2 & 0,5 & 0,3 \\ 0,2 & 0,7 & 0,1 \\ 0,3 & 0,3 & 0,4 \end{bmatrix} = (0,2; 0,5; 0,3).$$

Ймовірність, що друга машина буде марки С дорівнює 0,3.

2. Для відповіді на друге питання потрібно знайти

$$T^2 = \begin{bmatrix} 0,23 & 0,54 & 0,23 \\ 0,21 & 0,62 & 0,17 \\ 0,24 & 0,48 & 0,28 \end{bmatrix}.$$

Для другого питання $p^{(0)} = (0; 0,5; 0,5)$.

$$p^{(2)} = (0; 0,5; 0,5) \cdot \begin{bmatrix} 0,23 & 0,54 & 0,23 \\ 0,21 & 0,62 & 0,17 \\ 0,24 & 0,48 & 0,28 \end{bmatrix} = (0,225; 0,55; 0,225).$$

Отже, ймовірність, що другий автомобіль буде марки А, дорівнює 0,225.

Відповідь. 1. Ймовірність, що друга машина буде марки С дорівнює 0,3; 2. Ймовірність, що другий автомобіль буде марки А, дорівнює 0,225.

15.6. Класифікація станів і ланцюгів Маркова

Властивості ланцюга Маркова та відповідно способи розв'язання пов'язаних з його аналізом задач істотно залежать від того, з яких саме станів він складається. У зв'язку з цим наведемо тут прийняту класифікацію станів скінченного ланцюга Маркова та обумовлену нею класифікацію самих ланцюгів Маркова.

Кажуть, що **стан досягається зі стану** якщо існує таке число кроків k , що $p_{ij}^{(k)} > 0$, тобто з додатною ймовірністю ланцюг Маркова за кроків переходить із стану в стан (включаючи випадок).

Множина станів $C \subset E$ називається **поглинаючою** (або **замкненою**), якщо кожний стан, який не входить в C , є недосяжний з жодного стану, який їй належить, тобто, коли $p_{ij}^{(k)} = 0$ для всіх та таких, що i — належить, а j — не належить.

Звідси, зокрема, випливає, що окремий розгляд станів, які входять у замкнену множину, знову приводить до ланцюга Маркова, який можна вивчати незалежно.

Якщо один стан утворює замкнену множину, то він називається *поглинаючим станом*.

Стан називається *зворотним*, якщо ймовірність повернення для нього дорівнює одиниці, тобто $p_{ij}^{(k)} = 1$, і *незворотним*, якщо ця ймовірність менша від одиниці, тобто $p_{ij}^{(k)} < 1$.

Теорема 15.3. *Стан ланцюга Маркова є зворотним тоді і тільки тоді, коли середнє число повернень в нього нескінченне, і що з будь-якого зворотного стану не можна досягнути жодного незворотного стану. Крім цього, для того, щоб стан був незворотним необхідно і достатньо, щоб виконувалася умова (15.15):*

$$\sum_{k=1}^{+\infty} p_{ij}^{(k)} < +\infty. \quad (15.15)$$

Стан називається *періодичним* з періодом $\tau > 1$, якщо $p_{ij}^{(k)} = 0$, коли система не може повернутися в стан за час, відмінний від τ ; $2 \cdot \tau$; $3 \cdot \tau$; ...). Стан називається *неперіодичним*, якщо такого не існує.

Стан називається *ергодичним*, якщо він зворотний і неперіодичний.

Наведені означення різних типів станів марковського ланцюга становлять основу класифікації самих ланцюгів Маркова.

Ланцюг Маркова називається *незвідним*, якщо він не містить замкнених множин станів, відмінних від множини E всіх станів.

З даного означення випливає, що незвідний скінченний ланцюг Маркова складається лише із зворотних станів, кожний з яких досягається з будь-якого іншого стану ланцюга. Отже, виходячи з означення замкненої (поглинаючої) множини станів, робимо висновок, що будь-яка така множина може розглядатися як незвідний скінченний ланцюг Маркова.

Ланцюг Маркова називається *ергодичним*, якщо він незвідний і всі його стани є ергодичними.

Властивість ергодичності для однорідного ланцюга Маркова в термінах перехідних ймовірностей означає, що існує таке ціле число $k_0 \geq 1$, що всі елементи матриці P^{k_0} є додатними.

Зауважимо, що при аналізі структури ланцюга Маркова слід розрізняти визначене вище поняття поглинаючої (замкненої) множини станів від множини поглинаючих станів. Очевидно, в останньому випадку ланцюг Маркова вже не може бути незвідним, оскільки не всі стани є взаємодосяжними (з поглинаючого стану жодний інший стан не досягається). Зрозуміло також, що такий ланцюг мусить містити незворотні стани, бо якщо поглинаючий стан є досяжним для деякого стану, то система

вже не може повернутися в стан з ймовірністю одиниця (а тому стан незворотний). Якщо в склад складного ланцюга Маркова входять і замкнені множини станів, то, як вже відзначалося, їх можна окремо аналізувати як незвідний ланцюг Маркова. У зв'язку з цим важливим є розгляд марковських ланцюгів, які містять лише поглинаючі та незворотні стани.

Ланцюг Маркова називається *поглинаючим* (*ланцюгом з поглинанням*), якщо він складається лише з поглинаючих та незворотних станів.

15.7. Система масового обслуговування

Система масового обслуговування (СМО) – це математична модель, яка описує процеси надходження заявок на обслуговування, їх очікування в черзі і процеси обслуговування. Такі системи широко використовуються в різних сферах, від телефонних мереж і комп'ютерних систем до банків і транспортних систем.

Основні компоненти СМО

1. *Заявки (клієнти)* – це об'єкти, які потребують обслуговування. Вони можуть надходити до системи випадково або детерміновано.
2. *Система обслуговування* – це один або кілька обслуговуючих пристроїв (каналів), які обробляють заявки.
3. *Черга* – це місце, де заявки очікують на обслуговування, якщо всі обслуговуючі пристрої зайняті.
4. *Вхідний потік заявок* – Процес надходження заявок в систему. Він може бути випадковим або детермінованим.
5. *Час обслуговування* – час, необхідний для обробки однієї заявки.
6. *Вихідний потік заявок* – заявки, які вже були обслуговані і покинули систему.

Класифікація СМО

СМО можуть класифікуватися за різними ознаками, наприклад, за типом вхідного потоку, розподілом часу обслуговування, кількістю каналів обслуговування, наявністю або відсутністю черг і т.д.

СМО діляться на системи з відмовами і системи з чеканням. В *системі з відмовами* заява, яка поступає до системи в момент, коли всі канали обслуговування зайняті, отримує відмову і залишає систему. В *системі з чеканням* така заява не залишає систему, а стає в чергу, поки не звільнено буде який-небудь канал. Час чекання може бути як необмеженим, так і обмеженим.

Основні моделі СМО

1. Модель $M/M/1$:

Вхідний потік — Пуассонівський (M).

Час обслуговування — експоненціальний (M).

Один обслуговуючий канал (1).

Використовується для найпростіших систем із необмеженою чергою.

2. Модель $M/M/c$:

c – кількість обслуговуючих каналів.

Вхідний потік і час обслуговування залишаються Пуассонівськими та експоненціальними відповідно.

3. Модель $M/G/1$:

Вхідний потік — Пуассонівський.

Час обслуговування має довільний розподіл (G).

Один обслуговуючий канал.

4. Модель $G/G/1$:

Загальна модель, у якій і вхідний потік (G) і час обслуговування (G) мають довільні розподіли.

Використовується для складніших систем.

5. Модель $M/M/1/K$:

Пуассонівський потік, експоненціальне обслуговування.

Черга обмежена K місцями.

6. Модель $M/M/+\infty$:

Необмежена кількість каналів.

Вхідний потік Пуассонівський.

Основні характеристики СМО

Середня кількість заявок в системі (L): Включає як заявки в черзі, так і ті, що обслуговуються.

Середній час перебування заявки в системі (W): Включає час очікування в черзі і час обслуговування.

Середня кількість заявок в черзі (L_q): Тільки заявки, які очікують в черзі.

Середній час очікування в черзі (W_q): Час, який заявка проводить в черзі до початку обслуговування.

Ймовірність відмови ($P(n)$): Ймовірність того, що заявка не зможе бути обслугована через брак вільних обслуговуючих каналів (важливо для систем з обмеженою кількістю місць в черзі або без черги).

Зокрема:

Для моделі $M/M/1$ середній час очікування обчислюється формулою (15.16):

$$W_q = \frac{\rho}{\mu \cdot (1 - \rho)}, \quad (15.16)$$

де $\rho = \frac{\lambda}{\mu}$ — інтенсивність завантаження системи.

Для моделі $M/M/1$ ймовірність відмови обчислюється за формулою (15.17):

$$P(n) = (1 - \rho) \cdot \rho^n. \quad (15.17)$$

Для моделі $M/M/1$ середня кількість заявок у системі обчислюється за формулою (15.18):

$$L = \frac{\rho}{1 - \rho}. \quad (15.18)$$

Приклад застосування СМО

Прикладом системи масового обслуговування може бути кол-центр, де клієнти телефонують для отримання підтримки. *Клієнти* — це заявки,

оператори – це обслуговуючі канали, і якщо всі оператори зайняті, клієнти потрапляють в чергу. СМО дозволяє моделювати такі системи і оцінювати ефективність обслуговування, знаходити оптимальну кількість операторів для мінімізації часу очікування і уникнення перевантаження системи. Зокрема:

- $M/M/1$ – Обслуговування клієнтів на касі супермаркету;
- $M/M/c$ – Колл-центр із c операторами;
- $M/G/1$ – Система з довільним часом обслуговування, наприклад, склад із різними типами запитів.

Системи масового обслуговування є важливим інструментом для аналізу і оптимізації процесів обслуговування в різних сферах діяльності. Вони дозволяють розробляти ефективні стратегії управління чергами, підвищувати якість обслуговування і задоволеність клієнтів.

ЛЕКЦІЯ 16

ВИБІРКОВИЙ МЕТОД

16.1. Генеральна сукупність та вибірка. Варіаційний ряд.

Статистичний розподіл вибірки

Математична статистика — наука про методи збору, систематизації, обробки і використання для наукових і практичних висновків статистичних даних, здобутих у результаті випробувань із випадковими наслідками. Ці дані є можливими значеннями деякої випадкової величини (досліджуваної ознаки) X , наприклад похибки вимірювань фізичної величини, відхилення розмірів виробу від стандарту, процента зайнятості крісел на авіарейсах, прибутку підприємства тощо.

При цьому досліджувані ознаки можуть бути дискретними (набувати окремих ізольованих значень) або неперервними (набувати будь-яких можливих значень з деякого інтервалу).

При дослідженні ознаки X зібрати та обробити всі її можливі значення здебільшого буває неможливо через їх занадто великий об'єм, а також у випадках, коли, наприклад, аналізуючи відповідність стандарту геометричних розмірів виробів, доводиться виконувати масові вимірювання або, аналізуючи міцність виробів, руйнувати їх. Тому з усієї множини N даних про ознаку X , яка називається *генеральною сукупністю* об'ємом n , випадковим чином відбирається досить представницька (репрезентативна) її частина, що містить n даних і називається *вибірковою сукупністю*, або *вибіркою* об'єму n . При цьому об'єм вибірки n має бути суттєво меншим за об'єм генеральної сукупності N , але дані вибірки мають достатньо повно відбивати властивості ознаки X генеральної сукупності.

Різні можливі значення x_i ознаки X , які потрапили до вибірки, називаються *варіантами*, а система варіант $x_1, x_2, \dots, x_k$, розміщена в зростаючому порядку, називається *варіаційним рядом*.

Якщо варіанти $x_1, x_2, \dots, x_k$ спостерігаються у вибірці відповідно $n_1, n_2, \dots, n_k$ раз, то значення $n_1, n_2, \dots, n_k$ називаються *частотами варіант*, і для них виконується умова $\sum_{i=1}^k n_i = n$, ($i = \overline{1, k}$), а відношення

$$\omega_1 = \frac{n_1}{n}, \omega_2 = \frac{n_2}{n}, \dots, \omega_k = \frac{n_k}{n}$$

називаються *відносними частотами* варіант, і для них виконується умова $\sum_{i=1}^k \omega_i = 1$,

Таблиця, в першому рядку якої наведено впорядковані варіанти x_i (варіаційний ряд), а в другому — відповідні їм частоти n_i (або відносні частоти ω_i), називається дискретним статистичним розподілом вибірки (табл. 16.1 або табл. 16.2).

Таблиця 16.1. Варіативний ряд

x_i	x_1	x_2	...	x_{k-1}	x_k
n_i	n_1	n_2	...	n_{k-1}	n_k

Таблиця 16.2. Варіативний ряд у відносних частотах

x_i	x_1	x_2	...	x_{k-1}	x_k
ω_i	ω_1	ω_2	...	ω_{k-1}	ω_k

Якщо кількість варіант у вибірці надто велика або досліджувана ознака X є неперервною випадковою величиною, то будується інтервальний статистичний розподіл вибірки, аналогічний дискретному, у першому рядку якого замість дискретного варіаційного ряду записується *інтервальний ряд*

$$[x_1; x_2), [x_2; x_3), \dots, [x_l; x_{l+1}).$$

Кількість інтервалів l для вибірки об'ємом n орієнтовно вважається такою, що дорівнює $\sqrt{n}$ (із округленням до цілого значення), а розмір (крок) h інтервалу обчислюється за формулою (16.1):

$$h = \frac{x_k - x_1}{l}, \quad (16.1)$$

де:

x_1 і x_k — відповідно найменша і найбільша варіанти у вибірці;

l — кількість інтервалів.

Частоти n_j варіант, які потрапили в кожний частинний інтервал $[x_j; x_{j+1})$, ($j = \overline{1, l}$) записуються в другий рядок табл. 16.3.

Таблиця 16.3. Інтервальний статистичний ряд

$[x_j; x_{j+1})$	$[x_1; x_2)$	$[x_2; x_3)$	...	$[x_l; x_{l+1})$
n_j	n_1	n_2	...	n_l

Якщо межа двох частинних інтервалів збігається зі значенням варіанти, то частота цієї варіанти або ділиться порівну між цими двома інтервалами, або в усіх таких випадках цілком заноситься у правий (або лівий) від варіанти інтервал.

Правило Стерджеса — емпіричне правило визначення рекомендованої кількості інтервалів, на які розбивається діапазон зміни випадкової величини, що спостерігається, при побудові гістограми щільності її розподілу. Названо на ім'я американського статистика Герберта Стерджеса (Herbert Arthur Sturges, 1882-1958).

Кількість інтервалів визначається як:

$$l = 1 + [\log_2(n)], \quad (16.2)$$

де:

n — загальна кількість спостережень величини;

$\log_2(n)$ — логарифм за основою 2.

Часто зустрічається записаним через десятковий логарифм:

$$l = 1 + [3,332 \cdot \lg(n)], \quad (16.3)$$

де:

n — загальна кількість спостережень величини;

$lg(n)$ – логарифм за основою 10.

Також зустрічається записаним через натуральний логарифм:

$$l = 1 + [3,322 \cdot \ln(n)]. \quad (16.4)$$

де:

n – загальна кількість спостережень величини;

$\ln(n)$ – натуральний логарифм.

Приклад 16.1. У результаті вибіркового аналізу добової виручки авіакомпанії дістали вибірку обсягом $n = 40$ (млн грн):

0,87 0,94 0,98 0,90 0,90 0,87 0,85 0,87
0,90 0,94 0,87 0,87 0,82 0,90 0,94 0,90
0,85 0,85 0,87 0,94 0,81 0,82 0,87 0,97
0,90 0,94 0,85 0,81 0,87 0,85 0,90 0,82
0,98 0,90 0,94 0,82 0,97 0,81 0,85 0,87

Скласти: а) варіаційний ряд; б) інтервальний розподіл.

Розв’язання. а) Випишемо різні значення варіант, які потрапили у вибірку:

0,87; 0,94; 0,98; 0,90; 0,85; 0,82; 0,81; 0,97.

Розмістивши їх у порядку зростання, дістанемо дискретний варіаційний ряд:

0,81; 0,82; 0,85; 0,87; 0,90; 0,94; 0,97; 0,98.

Обчислимо частоту кожної варіанти із варіаційного ряду і складемо таблицю, відповідну табл. 16.1. Дістанемо дискретний статистичний розподіл вибірки (табл. 16.4).

Таблиця 16.4. Варіативний ряд

x_i	0,81	0,82	0,85	0,87	0,90	0,94	0,97	0,98
n_i	3	4	6	9	8	6	2	2

б) За обсягом вибірки $n = 40$ визначаємо за формулою (16.2) орієнтовну кількість $l = 6$ частинних інтервалів в інтервальному статистичному розподілі, а за формулою (16.1) обчислюємо крок інтервалу

$$h = \frac{0,98 - 0,81}{6} = 0,03$$

Тоді інтервальний варіаційний ряд запишеться у вигляді: [0,81; 0,84), [0,84; 0,87), [0,87; 0,90), [0,90; 0,93), [0,93; 0,96), [0,96; 0,99).

Тоді інтервальний статистичний розподіл вибірки набере вигляду (табл. 16.5):

Таблиця 16.5. Інтервальний варіативний ряд 1

$[x_j; x_{j+1})$	[0,81; 0,84)	[0,84; 0,87)	[0,87; 0,90)	[0,90; 0,93)	[0,93; 0,96)	[0,96; 0,99)
n_j	7	6	9	8	6	4

Якщо частоти варіант, які співпали з межами інтервалів 0,87 та 0,90 розподілити порівну між суміжними інтервалами (табл. 16.6), то:

Таблиця 16.6. Інтервальний варіативний ряд 2

$[x_j; x_{j+1})$	[0,81; 0,84)	[0,84; 0,87)	[0,87; 0,90)	[0,90; 0,93)	[0,93; 0,96)	[0,96; 0,99)
n_j	7	10	9	4	6	4

Відповідь. а) табл. 16.4; б) табл. 16.5.

16.2. Полігон та гістограма. Емпірична функція розподілу

Для створення наочного уявлення про статистичні розподіли застосовуються полігон та гістограма частот.

Полігоном частот (або **відносних частот**) називається ламана лінія, відрізки якої сполучають на площині точки з координатами (x_i, n_i) (або (x_i, ω_i)). Полігон застосовується також для графічного зображення інтервальних статистичних розподілів вибірки. У цьому випадку за абсциси x_i точок беруться центри частинних інтервалів.

Гістограма застосовується для графічного зображення інтервальних статистичних розподілів. Для побудови гістограми частот (або відносних частот) на осі абсцис відкладаються відрізки, що дорівнюють довжині (кроку) h частинних інтервалів, і на цих відрізках як на основах будуються прямокутники з висотами $\frac{n_j}{h}$ (або $\frac{\omega_j}{h} = \frac{n_j}{n \cdot h}$).

Оскільки величини $\frac{n_j}{h} \left(\frac{\omega_j}{h} \right)$ є щільностями частот (відносних частот) на відповідних інтервалах, то при достатньо великих обсягах вибірки n (і відповідно малих h) гістограма може бути достатньо близьким статистичним аналогом щільності ймовірності досліджуваної ознаки X у генеральній сукупності.

Приклад 16.2. Побудувати полігон відносних частот для дискретного статистичного розподілу (табл. 16.2) та гістограму відносних частот для інтервального розподілу (табл. 16.3).

Розв'язання. Обчислимо відносні частоти $\omega_i = \frac{n_i}{n}$ для дискретного розподілу і щільності відносних частот $\frac{\omega_j}{h} = \frac{n_j}{n \cdot h}$ для інтегрального розподілу при $n = 40$, $h = 0,03$. Дістанемо відповідно розподіли (табл. 16.7 і 16.8).

Таблиця 16.7. Варіативний ряд у відносних частотах

x_i	0,81	0,82	0,85	0,87	0,90	0,94	0,97	0,98
ω_i	0,075	0,10	0,15	0,225	0,20	0,15	0,05	0,05

Таблиця 16.8. Інтервальный варіативний ряд

$[x_j; x_{j+1})$	[0,81; 0,84)	[0,84; 0,87)	[0,87; 0,90)	[0,90; 0,93)	[0,93; 0,96)	[0,96; 0,99)
$\frac{\omega_j}{h}$	5,83	8,33	7,5	3,33	5	3,33

Полігон та гістограму відносних частот побудовано відповідно на рис. 16.1 і 16.2.

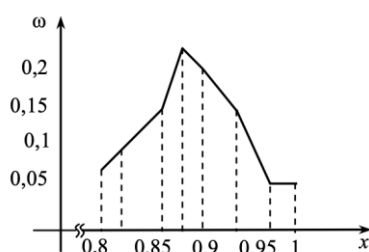


Рис. 16.1. Полігон

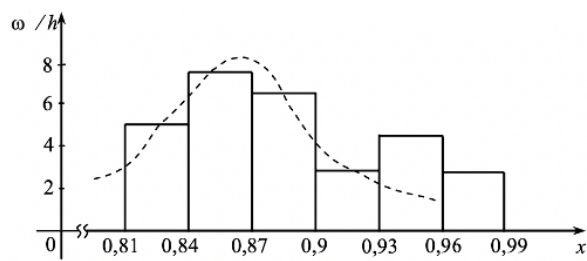


Рис. 16.2. Гістограма

Відповідь. Рис. 16.1 і 16.2.

За загальним виглядом гістограми можна скласти певне уявлення про графік щільності розподілу ознаки X у генеральній сукупності, наприклад припустити, що він має вигляд, зображений на рис. 16.2 пунктирною лінією (при згладженні деяких сплесків частот, викликаних, можливо, малим обсягом вибірки).

Розподіл ознаки у варіаційному ряду за накопиченими частотами (частинами) зображується за допомогою кумуляти. Кумулята або кумулятивна крива, на відміну від полігону, будується за накопиченими частотами або частотами. У цьому на осі абсцис поміщають значення ознаки, але в осі ординат — накопичені частоти чи частоти.

Кумулята — графічне порівняння двох або більше варіаційних розподілів з рівними чи нерівними інтервалами. Вона будується за кумулятивним розподілом накопичених частот, при цьому використовуються праві кінці інтервалів (рис. 16.3).

Огіва — це різновид кумулятивного розподілу. Вона є дзеркальним відображенням кумуляти. На осі ординат відкладаємо межі інтервалів, по осі абсцис — накопичені частоти (рис. 16.4).

Приклад 16.3. Для наступних статистичних даних (табл. 16.9) побудувати кумуляту та огіву.

Таблиця 16.9. Вхідні дані

Вибірка	Частота	Накопичувана частота
0,1	0	0
0,3	5	5
0,5	6	11
0,7	18	29
0,9	16	45
1,1	12	57
1,3	10	67
1,5	1	68
1,7	2	70

Розв’язання. На рис. 16.3 та 16.4 наведено відповідно графіки кумуляти та огіви.

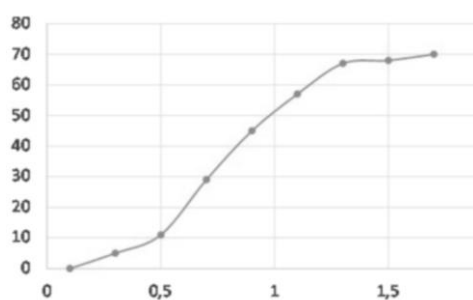


Рис. 16.3. Кумулята

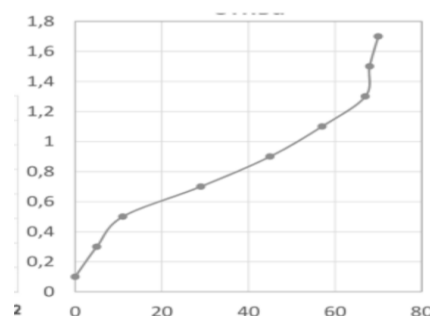


Рис. 16.4. Огіва

Відповідь. Рис. 16.3 і 16.4.

Перейдемо до розгляду емпіричної функції розподілу, яка утворюється за даними вибірки і є наближеною оцінкою реальної або так званої теоретичної функції розподілу $F(x)$ ознаки X генеральної сукупності.

Емпіричною функцією статистичного розподілу називається функція $F^*(x)$ яка для кожного значення x дорівнює відносній частоті події $X < x$, тобто (16.5):

$$F^*(x) = \frac{n_{x_i}}{n}, \quad (16.5)$$

де n_{x_i} — сумарна частота всіх варіант x_i , менших за x .

Для дискретного статистичного розподілу (табл. 16.1) функція $F^*(x)$ набуває таких значень:

$F^*(x) = 0$, при $x \leq x_1$, оскільки в цьому випадку немає варіант, менших за x ;

$F^*(x) = \frac{n_1}{n}$, при $x_1 < x \leq x_2$;

$F^*(x) = \frac{n_1+n_2}{n}$, при $x_2 < x \leq x_3$;

.....
 $F^*(x) = \frac{n_1+n_2+\dots+n_{k-1}}{n}$, при $x_{k-1} < x \leq x_k$;

$F^*(x) = \frac{n_1+n_2+\dots+n_k}{n} = 1$, при $x > x_k$, оскільки в цьому випадку додаються всі варіанти.

Графік функції $F^*(x)$ будується аналогічно графіку функції розподілу випадкової величини. Для неперервної ознаки X генеральної сукупності емпірична функція будується за інтервальним статистичним розподілом. Із цією метою обчислюються значення функції $F^*(x)$ у кількох точках, за які найчастіше беруть межі інтервалів. Знайдені точки $(x_i, F^*(x_i))$ наносять на графік і сполучають прямолінійними відрізками.

Якщо x_1 — ліва межа першого інтервалу, а x_{m+1} — права межа останнього інтервалу в інтервальному варіаційному ряду, то $F^*(x) = 0$ при $x \leq x_1$ і $F^*(x) = 1$ при $x > x_{m+1}$.

Приклад 16.4. Побудувати емпіричну функцію розподілу $F^*(x)$ за дискретним (табл. 16.2) та інтервальним (табл. 16.3) статистичними розподілами вибірки.

Розв'язання. Для обчислення значень емпіричної функції скористаємось дискретним статистичним розподілом із відносними частотами ω_i (табл. 16.5), який знайдено у прикладі 16.2. Значення функції $F^*(x)$ обчислюються простим нагромадженням відносних частот:

$$F^*(x) = \begin{cases} 0, & \text{при } x \leq 0,81; \\ 0 + 0,075 = 0,075, & \text{при } 0,81 < x \leq 0,82; \\ 0,075 + 0,1 = 0,175, & \text{при } 0,82 < x \leq 0,85; \\ 0,175 + 0,15 = 0,325, & \text{при } 0,85 < x \leq 0,87; \\ 0,325 + 0,225 = 0,55, & \text{при } 0,87 < x \leq 0,90; \\ 0,55 + 0,2 = 0,75, & \text{при } 0,90 < x \leq 0,94; \\ 0,75 + 0,15 = 0,9, & \text{при } 0,94 < x \leq 0,97; \\ 0,9 + 0,05 = 0,95, & \text{при } 0,97 < x \leq 0,99; \\ 0,95 + 0,05 = 1, & \text{при } x > 0,99. \end{cases}$$

Графік функції $F^*(x)$ подано на рис. 16.5 (східчаста розривна лінія).

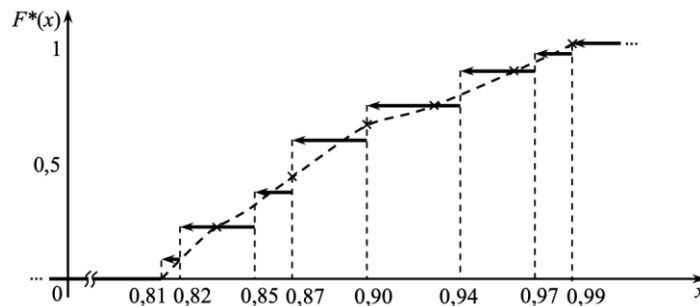


Рис.16.5. Емпірична функція розподілу

Для побудови функції $F^*(x)$ за інтервальним статистичним розподілом (табл. 16.3) обчислимо значення функції на межах інтервалів:

$x = 0,81, F^*(x) = 0$, оскільки варіант, менших за 0,81, у вибірці немає;

$x = 0,84, F^*(x) = \frac{7}{40} = 0,175$, оскільки варіанти, менші за 0,84, містяться в першому інтервалі, і їх кількість 7;

$x = 0,87, F^*(x) = \frac{7+10}{40} = 0,425$;

$x = 0,90, F^*(x) = \frac{7+10+9}{40} = 0,65$;

$x = 0,93, F^*(x) = \frac{7+10+9+4}{40} = 0,75$;

$x = 0,96, F^*(x) = \frac{7+10+9+4+6}{40} = 0,9$;

$x = 0,99, F^*(x) = \frac{7+10+9+4+6+4}{40} = 1$.

Нанесемо знайдені точки на графік рис. 16.5 і сполучимо їх відрізками прямої. Дістанемо графік функції $F^*(x)$, зображений на рис. 16.5 пунктирною лінією.

Відповідь. Рис. 16.5.

16.3. Мода та медіана статистичної вибірки

Більш докладно про моду та медіану розповідалося в лекції 6 паралельно з модою та медіаною випадкової величини.

Мода статистичної вибірки — це значення, яке найчастіше зустрічається у вибірці даних.

Особливості моди:

- Мода визначається як значення або категорія з найбільшою частотою.

Вибірка може мати:

- Одну моду (унімодальний розподіл),
- Дві моди (бімодальний розподіл),
- Більше двох мод (мультимодальний розподіл).

Приклад 16.5. Обчислити моду вибірки X : 2,4,4,6,7,4,8,9.

Розв'язання. Значення 4 зустрічається тричі, тоді як інші значення з'являються лише раз.

$$Mo(X) = 4.$$

Відповідь. Мода вибірки дорівнює 4.

Переваги використання моди:

- Можна застосовувати як для кількісних, так і для якісних даних;
- Є стійкою до викидів (аномально великих або малих значень).

Недоліки моди:

- Мода може бути неунікальною (багато мод);
- Вона не завжди відображає центр розподілу, як це роблять середнє значення чи медіана.

Застосування моди:

- Аналіз категоріальних даних (наприклад, найбільш популярний товар;
- Визначення найбільш часто зустрічаючогося значення в наборах кількісних даних.

Мода є простою, але ефективною характеристикою центральної тенденції, особливо в практичних застосуваннях.

Медіана статистичної вибірки — це значення, яке ділить упорядковану вибірку навпіл: 50% спостережень менші або рівні медіані, а 50% — більші або рівні медіані.

Властивості медіани:

- Стійкість до викидів: на відміну від середнього арифметичного, медіана менш чутлива до аномально великих або малих значень;
- Використовується як міра центральної тенденції для кількісних даних.

Спосіб обчислення медіани:

1. Впорядкувати значення у вибірці у зростаючому порядку.

2. Залежно від кількості значень у вибірці:

- Непарна кількість значень: медіана — це значення, що стоїть посередині;
- Парна кількість значень: медіана — це середнє арифметичне двох центральних значень.

Приклад 16.6. Обчислити медіану вибірок X та Y : 3,5,7,9,11 та 2,4,6,8 відповідно.

Розв'язання. Вибірка X має непарну кількість значень, отже, центральне значення дорівнює 7.

$$Me(X) = 7.$$

Вибірка Y має парну кількість значень, отже, центральні значення дорівнюють 4 та 6.

$$Me(Y) = \frac{4 + 6}{2} = 5.$$

Відповідь. Мода вибірки X дорівнює 7; мода вибірки Y дорівнює 5.

Переваги медіани:

- Використовується, коли дані містять викиди або асиметричний розподіл;
- Зручна для рангових і порядкових даних.

Недоліки медіани:

- Може бути менш точною, якщо розподіл даних симетричний (у такому випадку середнє значення краще відображає центр);
- Не враховує всі значення у вибірці.

Застосування медіани:

- Економіка: наприклад, медіанний дохід показує типову заробітну плату, ігноруючи викиди з дуже великими значеннями;
- Медицина: аналіз тривалості життя пацієнтів;
- Соціальні дослідження: оцінка середніх показників на основі неповних або асиметричних даних.

Медіана — це зручна і наочна характеристика розподілу даних, особливо у випадках, коли середнє значення є занадто чутливим до аномалій.

16.4. Часові ряди

Часові ряди — це набір даних, що відображають значення певної змінної через регулярні проміжки часу. Важливою особливістю часових рядів є те, що порядок спостережень має значення, оскільки він відображає зміну показника у часі (рис. 16.6, 16.7).

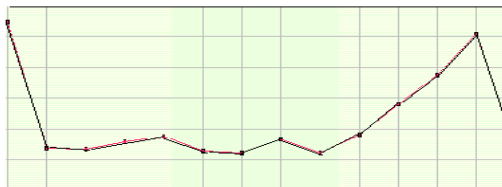


Рис. 16.6. Фрагмент часового ряду за сезонний період

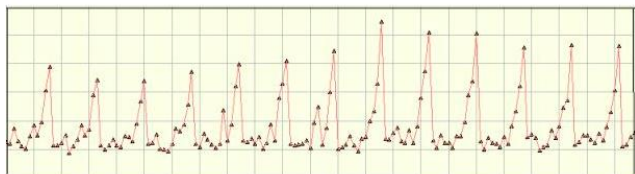


Рис. 16.7. Фрагмент часового ряду за 12 місяців

Основні поняття та компоненти часових рядів

Тренд (Trend) — це довгострокова тенденція, яка відображає загальний напрямок зміни даних (зростання, спад або стабільність) протягом тривалого часу. Наприклад, довгострокове зростання ВВП або поступове збільшення населення.

Сезонність (Seasonality) — це регулярні коливання, що повторюються з певною періодичністю (рік, квартал, місяць, тиждень) і пов'язані з певними факторами. Наприклад, збільшення продажів під час свят або сезонні коливання температури.

Циклічність (Cyclical Component) – це довгострокові коливання, які можуть тривати кілька років і пов'язані з економічними циклами, наприклад, економічні підйоми та спади. Циклічні коливання часто важко передбачити і вони можуть мати непостійний період.

Шум (Noise або Random Component) – це випадкові коливання в часовому ряді, які не піддаються прогнозуванню і відображають вплив непередбачуваних факторів. Вони додають «шуму» до загальної картини і ускладнюють аналіз.

Аналіз часових рядів

Аналіз часового ряду здійснюється з метою:

- визначення природи ряду;
- прогнозування майбутніх значень ряду.

У процесі визначення структури і закономірностей часового ряду передбачається виявлення: шумів і викидів, тренду, сезонної компоненти, циклічної компоненти. Визначення природи тимчасового ряду може бути використано як своєрідна розвідка даних. Знання аналітика про наявність сезонної компоненти необхідно, наприклад, для визначення кількості записів вибірки, яке повинно брати участь в побудові прогнозу.

Шуми і викиди будуть детально обговорюватися нижче. Вони ускладнюють аналіз часового ряду. Існують різні методи визначення та фільтрації викидів, що дають можливість виключити їх з метою більш якісного Data Mining.

Основою для прогнозування (однієї з основних задач Data Mining) служить історична інформація, що зберігається в базі даних у вигляді часових рядів.

$$y_t = tr_t + s_t + c_t + r_t \text{ або } y_t = tr_t + s_t + r_t, \quad (16.6)$$

де:

- tr_t – тренд;
- s_t – сезонна складова;
- c_t – циклічна складова;
- r_t – випадкова складова (шум).

Приклади часових рядів

Економічні дані: Валовий внутрішній продукт (ВВП) за квартал, рівень безробіття щомісяця, індекс цін на акції щодня.

Фінансові дані: Щоденні ціни на акції, обсяги торгівлі, валютні курси.

Метеорологічні дані: Щоденні температури, місячні опади, швидкість вітру.

Соціальні дані: Щорічний приріст населення, рівень злочинності за місяць, кількість користувачів соціальних мереж за рік.

16.5. Перехід від часових рядів до статистичних вибірок

Для контролю якості прогнозів (зокрема курсу криптовалют), необхідно сформулювати датасет з наявної статистичної інформації (табл. 16.10). Детально даний алгоритм представлено в статтях [28-29].

Таблиця 16.10. Значення курсу обраної криптовалюти за вказаний проміжок часу

Момент часу	Реальне значення курсу	Прогнозоване значення курсу
...	...	...
t_1	x_1	y_1
t_2	x_2	y_2
...	...	...
t_k	x_k	y_k
...	...	...

В таблиці 16.10 наведено реальні курси обраної криптовалюти $x_1, x_2, \dots, x_k$ в обраний проміжок часу $[t_1; t_k]$, які можна взяти з сайту крипто біржі Binance, прогнозовані курси криптовалюти $y_1, y_2, \dots, y_k$ отримано за допомогою обраного алгоритму прогнозування.

Таким чином, реальні та прогнозовані курси криптовалюти представляють собою часові ряди $X = (x_1, x_2, \dots, x_k)$ та $Y = (y_1, y_2, \dots, y_k)$ відповідно за період часу $[t_1; t_k]$.

Найчастіше часовий ряд є послідовністю, взятою на рівновіддалених точках в часі, які йдуть одна за одною. Таким чином, він є послідовністю даних дискретного часу. Тобто, для аналізу часового ряду параметр часу виступає одним із визначальних чинників. Це відрізняє часовий ряд від звичайної випадкової вибірки, де індекси вводять лише для зручності ідентифікації. Принциповою відмінністю часового ряду від простих статистичних сукупностей є:

- по-перше, рівні часового ряду є залежними. Інакше кажучи, якщо майбутні значення змінної можна визначити, то вони є функцією від минулих значень цієї змінної;
- по-друге, рівні часового ряду неоднаково розподілені. Закон розподілу ймовірностей цих випадкових величин і, зокрема, їхні математичні сподівання та дисперсії можуть залежати від часу.

З іншого боку, якщо абстрагуватися від того, що означають величини x_i та y_i , $i = \overline{1; k}$, але при цьому продовжувати враховувати порядок їх слідування, то можна перейти до статистичних вибірок $X = (x_1, x_2, \dots, x_k)$ та $Y = (y_1, y_2, \dots, y_k)$, де k – об'єм вибірок, частоти елементів вибірок дорівнюють 1.

Варто підкреслити, що принциповим є той факт, що обидві вибірки було отримано з реальних та прогнозованих курсів криптовалюти за однаковий період часу і в ті самі моменти часу. Це дає змогу обґрунтувати правомірність переходу від часових рядів до статистичних вибірок і, таким чином, коректно обчислювати середню відносну похибку прогнозування *MAPE*.

$$MAPE = \frac{1}{k} \cdot \sum_{i=1}^k \left(\frac{|x_i - y_i|}{x_i} \right) \cdot 100\%, \quad (16.7)$$

де:

x_i – елементи вибірки X ;

y_i – елементи вибірки Y , $i = \overline{1; k}$;

k – об'єм вибірок X та Y .

Також слід зазначити, що $MAPE$ для добре побудованого прогнозу має наближатися до нуля, але аналітик сам має змогу встановити його допустиму верхню межу згідно наступної шкали (табл. 16.11).

Таблиця 16.11. Рівень адекватності моделі прогнозування

$MAPE$	Точність прогнозу
менше 10%	Висока
10%– 20%	Добра
20%– 40%	Задовільна
40%– 50%	Погана
більше 50%	Незадовільна

Загалом, процес перевірки якості прогнозу складається з наступних кроків:

Крок 1. Встановлення верхньої допустимої межі $MAPE$.

Крок 2. Створення статистичних вибірок $X = (x_1, x_2, \dots, x_k)$ та $Y = (y_1, y_2, \dots, y_k)$ з реальних та прогнозованих курсів криптовалюти відповідно за період часу $[t_1; t_k]$.

Крок 3. Обчислення $MAPE$ для даних вибірок.

Крок 4. Прийняття рішення щодо точності прогнозу. Якщо значення $MAPE$ менше верхньої допустимої межі, то прогноз вважається якісним. В протилежному випадку потрібно збільшити об'єм вибірок або змінити алгоритм прогнозування за яким було отримано вибірку Y .

Перераховані кроки представлено на рис. 16.8:



Рис. 16.8. Схема перевірки якості прогнозу

Окрім вищесказаного, перехід до статистичного підходу дає змогу застосовувати для оцінки якості отриманого прогнозу класичні методи статистичного аналізу:

- обчислення основних точкових характеристик вибірок, таких як вибіркове середнє, вибіркове середньоквадратичне відхилення;
- обчислення коефіцієнти кореляції (зокрема коефіцієнти рангової кореляції Спірмена або Кендала, оскільки потрібно враховувати порядок слідування елементів вибірок);
- застосування статистичних критеріїв, зокрема t -критерію, що дає змогу перевірити гіпотезу про значущість кореляційного зв'язку між

виборками, а також критерію χ^2 –Пірсона для перевірки гіпотези про тип розподілу генеральної сукупності (див. лекцію 18).

Усі перераховані методи статистичного аналізу можуть бути використані не лише для контролю якості прогнозів, а й для побудови неперервних функцій, які є щільностями відповідного типу розподілу. За допомогою цих функцій також можна отримати потрібні прогнози.

ЛЕКЦІЯ 17

ТОЧКОВІ ТА ІНТЕРВАЛЬНІ ОЦІНКИ ПАРАМЕТРІВ РОЗПОДІЛУ

17.1. Статистичні оцінки параметрів розподілу

Параметри досліджуваної ознаки генеральної сукупності X , з якої зроблено вибірку $x = (x_1, x_2, \dots, x_n)$ є величинами сталими, але їхні числові значення невідомі. Ці параметри оцінюються параметрами вибірки.

Нехай θ — невідомий параметр (наприклад, математичне сподівання, дисперсія тощо) досліджуваної ознаки X .

Вибірковою оцінкою параметра θ називається число θ^* , знайдене за даними вибірки x або за статистичним розподілом, залежне від $x_1, x_2, \dots, x_n$, і таке, що наближено дорівнює оцінюваному параметру: $\theta^* \approx \theta$.

Ця оцінка, яка визначається у даній вибірці одним числом, точкою, називається *точковою*.

Оскільки θ^* є випадковою величиною, то вона може бути зміщеною і незміщеною. Коли математичне сподівання цієї оцінки точно дорівнює оцінюваному параметру θ , тобто $M(\theta^*) = \theta$, то оцінка θ^* називається *незміщеною*.

Якщо $M(\theta^*) \neq \theta$, то статистична оцінка θ^* називається *зміщеною* відносно параметра θ генеральної сукупності.

Різниця $\theta^* - \theta = \delta$ називається *зміщенням* статистичної оцінки θ^* .

Точкова статистична оцінка θ^* називається *спроможною*, якщо, у разі необмеженого збільшення обсягу вибірки, θ^* наближається до оцінюваного параметра θ (17.1):

$$\lim_{n \rightarrow +\infty} P\{|\theta^* - \theta| < \delta\} = 1. \quad (17.1)$$

Оцінюваний параметр може мати кілька точкових незміщених статистичних оцінок. Точкова статистична оцінка називається *ефективною*, коли при заданому обсязі вибірки вона має мінімальну дисперсію.

Для невідомого математичного сподівання $M(X)$ досліджуваної ознаки (генеральної сукупності) X зміщеною точковою оцінкою є *вибіркова середня*, яка позначається $\bar{x}_B$ і обчислюється за формулою (17.2):

$$\bar{x}_B = \frac{1}{n} \cdot \sum_{i=1}^k (x_i \cdot n_i). \quad (17.2)$$

Вона задовольняє всі три перелічені умови. Для невідомої дисперсії генеральної сукупності $D(X)$ спроможною оцінкою є *вибіркова дисперсія D_B* , обчислювана за формулою (17.3):

$$D_B = \frac{1}{n} \cdot \sum_{i=1}^k ((x_i - \bar{x}_B)^2 \cdot n_i). \quad (17.3)$$

Проте ця оцінка є зміщеною. Тому для оцінки невідомої дисперсії застосовують так звану *виправлену дисперсію*, яку позначають s^2 і обчислюють за формулою (17.4):

$$s^2 = \frac{n}{n-1} \cdot D_B = \frac{1}{n-1} \cdot \sum_{i=1}^k ((x_i - \bar{x}_B)^2 \cdot n_i). \quad (17.4)$$

Оцінкою для середнього квадратичного відхилення генеральної сукупності є *виправлене вибіркове середнє квадратичне відхилення* s .

Для неперервної ознаки X , для якої побудовано інтервальний статистичний розподіл, точкові оцінки невідомих математичного сподівання і дисперсії також знаходять за формулами (17.2) і (17.3), в які замість варіант x_i підставляють середини (центри) x_i^* інтервалів.

Приклад 17.1. Знайти точкові оцінки математичного сподівання, дисперсії і середнього квадратичного відхилення добової виручки авіакомпанії за даними вибірки з прикладу 16.1 лекції 16.

Розв'язання. За дискретним статистичним розподілом (табл. 16.2 лекції 16) обчислюємо вибіркиму середню $\bar{x}_B$, виправлену вибіркиму дисперсію s^2 і виправлене вибіркове середнє квадратичне відхилення σ_B , використовуючи формули (17.1)-(17.4):

$$\begin{aligned} \bar{x}_B &= \frac{1}{40} \cdot (0,81 \cdot 3 + 0,82 \cdot 4 + 0,85 \cdot 6 + 0,87 \cdot 9 + 0,90 \cdot 8 + 0,94 \cdot 6 + \\ &\quad 0,97 \cdot 2 + 0,98 \cdot 2) = 0,8845; \\ s^2 &= \frac{1}{39} \cdot ((0,81 - 0,8845)^2 \cdot 3 + (0,82 - 0,8845)^2 \cdot 4 + (0,85 - 0,8845)^2 \cdot 6 + \\ &\quad (0,87 - 0,8845)^2 \cdot 9 + (0,90 - 0,8845)^2 \cdot 8 + (0,94 - 0,8845)^2 \cdot 6 + \\ &\quad (0,97 - 0,8845)^2 \cdot 2 + (0,98 - 0,8845)^2 \cdot 2) \approx 0,0024; \\ s &\approx 0,049. \end{aligned}$$

Знайдемо оцінки $\bar{x}_B$, s^2 і σ_B за інтервальним статистичним розподілом (табл. 16.3 лекції 16), замінюючи у формулах (17.2) і (17.3) x_i центрами частинних інтервалів x_i^* . Для цього перейдемо спочатку від інтервального статистичного розподілу до дискретного, який набирає такого вигляду (табл. 17.1).

Таблиця 17.1. Дискретний варіативний ряд

x_i^*	0,825	0,855	0,885	0,915	0,945	0,975
n_i^*	7	10	9	4	6	4

Тоді використовуючи формули (17.1)-(17.4) отримаємо:

$$\begin{aligned} \bar{x}_B &= \frac{1}{40} \cdot (0,825 \cdot 7 + 0,855 \cdot 10 + 0,885 \cdot 9 + 0,915 \cdot 4 + 0,945 \cdot 6 + \\ &\quad 0,975 \cdot 4) = 0,888; \\ s^2 &= \frac{1}{39} \cdot ((0,825 - 0,888)^2 \cdot 7 + (0,855 - 0,888)^2 \cdot 10 + (0,885 - 0,888)^2 \cdot 9 + \\ &\quad (0,915 - 0,888)^2 \cdot 4 + (0,945 - 0,888)^2 \cdot 6 + (0,975 - 0,888)^2 \cdot 4) = \\ &\quad 0,0023; \\ \sigma_B &= 0,048. \end{aligned}$$

Відповідь. $\bar{x}_B = 0,888$; $s^2 = 0,0023$; $s = 0,048$.

17.2. Метод моментів

Метод моментів ґрунтується на тому, що невідомі параметри теоретичного розподілу (розподілу генеральної сукупності) визначаються із рівнянь, які добуваються прирівнюванням важливіших числових

характеристик (моментів) теоретичного розподілу відповідним числовим характеристикам емпіричного розподілу.

Так, нехай заданий, наприклад, вид теоретичного розподілу, який визначається невідомим параметром. Для знаходження одного параметра необхідне одне рівняння відносно даного параметра. Для цього використовується момент 1-го порядку (математичне сподівання) теоретичного розподілу і відповідна числова характеристика емпіричного розподілу – вибіркове середнє.

Знаходимо математичне сподівання:

$$M(X) = \int_{-\infty}^{+\infty} x d(F(x, \theta)) = M(\theta)$$

і вибіркове середнє:

$$\overline{x_B} = \frac{1}{n} \cdot \sum_{i=1}^k (x_i \cdot n_i).$$

Порівнюючи їх, одержуємо рівняння для визначення оцінки θ^* невідомого параметра (17.5):

$$M(\theta) = \overline{x_B}. \quad (17.5)$$

Для знаходження оцінок двох невідомих параметрів θ_1, θ_2 звичайно беруть математичне сподівання і дисперсію теоретичного розподілу та відповідні їм числові характеристики емпіричного розподілу – вибіркове середнє і вибіркovu дисперсію. Одержують два рівняння (17.6):

$$M(\theta_1, \theta_2) = \overline{x_B}, \sigma^2(\theta_1, \theta_2) = s^2. \quad (17.6)$$

В деяких підручниках, зокрема Гмурмана та Кремера, в системі (17.6) дисперсію теоретичного розподілу (генеральну дисперсію) оцінюють за допомогою вибіркової дисперсії D_B .

Розв'язуючи цю систему, знаходять відповідні оцінки θ_1^*, θ_2^* . Оцінки методу моментів звичайно є *слушними*, однак за ефективністю вони не є «найкращими». Тим не менш, метод моментів часто використовується на практиці, оскільки приводить до порівняно простих обчислень.

Приклад 17.2. Генеральна сукупність має експоненціальний розподіл із щільністю розподілу $f(x, \lambda) = \lambda \cdot e^{-(\lambda \cdot x)}$ ($x \geq 0$), де – невідомий параметр. Вибіркові значення $x_1, x_2, \dots, x_n$, узяті із однієї і тієї ж генеральної сукупності. Знайти методом моментів оцінку λ^* невідомого параметра розподілу.

Розв'язання. Визначимо математичне сподівання експоненціального розподілу:

$$M(\lambda) = \lambda \cdot \int_0^{+\infty} (x \cdot e^{-(\lambda \cdot x)}) dx = \frac{1}{\lambda}.$$

Далі по вибірці знаходимо вибіркове середнє емпіричного розподілу:

$$\overline{x_B} = \frac{1}{n} \cdot \sum_{i=1}^k (x_i \cdot n_i).$$

Із рівняння

$$M(\lambda) = \overline{x_B} = \frac{1}{\lambda}$$

знаходимо оцінку параметра:

$$\lambda^* = \frac{1}{\overline{x_B}}.$$

Відповідь. $\lambda^* = \frac{1}{\overline{x_B}}.$

Приклад 17.3. Вибірка $x_1, x_2, \dots, x_n$, одержана із генеральної сукупності з нормальним розподілом із щільністю розподілу:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}}, (-\infty < x < +\infty),$$

де a і σ – невідомі параметри розподілу. Знайти методом моментів оцінки a^* і σ^* невідомих параметрів a і σ .

Розв’язання. Визначимо математичне сподівання і дисперсію розподілу:

$$a = \int_{-\infty}^{+\infty} (x \cdot f(x)) dx, \sigma^2 = \int_{-\infty}^{+\infty} ((x-a)^2 \cdot f(x)) dx.$$

За даними вибірки знаходимо вибіркові числові характеристики $\overline{x_B}$, s і прирівнюємо їх до відповідних числових характеристик теоретичного розподілу:

$$a^* = \overline{x_B}, \sigma^* = s,$$

що є точковими оцінками відповідних параметрів.

Відповідь. $a^* = \overline{x_B}, \sigma^* = s.$

17.3. Метод максимальної правдоподібності

Одним із найбільш універсальних методів одержання оцінок параметрів розподілів генеральної сукупності є метод *максимальної правдоподібності*.

Оснoву методa складає функція правдоподібності, яка виражає щільність ймовірності (ймовірність) сумісної появи результатів вибірки $x_1, x_2, \dots, x_n$:

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta) \quad (17.7)$$

або:

$$L(x_1, x_2, \dots, x_n, \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \cdot \dots \cdot p(x_n, \theta) = \prod_{i=1}^n p(x_i, \theta). \quad (17.8)$$

Згідно з методом максимальної правдоподібності за оцінку невідомого параметра приймається таке його значення, яке максимізує функцію $L(x_1, x_2, \dots, x_n, \theta)$. Величина, при якій функція правдоподібності досягає максимального значення, називається *оцінкою максимальної правдоподібності*.

Природність такого підходу до визначення статистичних оцінок впливає із смислу функції правдоподібності, яка при кожному фіксованому значенні є *мірою правдоподібності* одержання вибірки. Оцінка є такою, що вибірка, яка одержана у результаті спостережень, є найбільш вірогідною. Знаходження оцінки спрощується, якщо максимізувати не саму функцію $L(x_1, x_2, \dots, x_n, \theta)$, а $\ln(L(x_1, x_2, \dots, x_n, \theta))$, оскільки максимум обох функцій досягається при одному і тому ж значенні.

Для одержання оцінки максимальної правдоподібності треба розв'язати рівняння:

$$\frac{d(\ln(L(x_1, x_2, \dots, x_n, \theta)))}{d\theta} = 0.$$

Якщо потрібно оцінити не один, а декілька параметрів $\theta_1, \theta_2, \dots, \theta_m$, то оцінка максимальної правдоподібності цих параметрів знаходиться із системи рівнянь:

$$\frac{\partial(\ln(L(x_1, x_2, \dots, x_n, \theta_1, \theta_2, \dots, \theta_m)))}{\partial \theta_i} \quad (i = 1, \dots, m).$$

Перевага методу максимальної правдоподібності полягає у тому, що для широкого класу розподілів він приводить до оцінок, які є слушними, асимптотично ефективними, мають асимптотично нормальний розподіл і, якщо для параметра існує ефективна оцінка, то рівняння правдоподібності має єдиний розв'язок, який співпадає з нею. Однак оцінка максимальної правдоподібності може виявитись зсуненою.

Приклад 17.4. Генеральна сукупність має розподіл Пуассона.

$$P_n(X = x_i) = \frac{\lambda^{x_i}}{x_i!} \cdot e^{-\lambda},$$

де n – кількість випробувань у кожній серії, x_i – кількість появ події у i -й серії ($i = 1, \dots, n$). Знайти оцінку λ^* невідомого параметра λ по вибірці.

Розв'язання. Складемо функцію правдоподібності.

$$\begin{aligned} L(x_1, x_2, \dots, x_n, \lambda) &= p(x_1, \lambda) \cdot p(x_2, \lambda) \cdot \dots \cdot p(x_n, \lambda) = \\ &= \left(\frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda}\right) \cdot \dots \cdot \left(\frac{\lambda^{x_n}}{x_n!} \cdot e^{-\lambda}\right) = \frac{\lambda^{\sum x_i}}{(x_1!) \cdot \dots \cdot (x_n!)} \cdot e^{-(n \cdot \lambda)}. \end{aligned}$$

Визначимо логарифм цієї функції

$$\ln(L(x_1, x_2, \dots, x_n, \lambda)) = \sum (x_i \cdot \ln(\lambda)) - n \cdot \lambda - \ln((x_1!) \cdot \dots \cdot (x_n!))$$

Прирівнюючи похідну цієї функції по до нуля, одержуємо рівняння правдоподібності

$$\frac{d(\ln(L(x_1, x_2, \dots, x_n, \lambda)))}{d\lambda} = \frac{\sum x_i}{\lambda} - n = 0.$$

Розв'язуючи це рівняння відносно λ , знаходимо:

$$\lambda^* = \frac{\sum x_i}{n} = \bar{x}_B.$$

Оскільки

$$\frac{d^2(\ln(L(x_1, x_2, \dots, x_n, \lambda)))}{d\lambda^2} = -\frac{\sum x_i}{\lambda^2} < 0,$$

то $\lambda^* = \bar{x}_B$ є точкою максимуму функції $\ln(L(x_1, x_2, \dots, x_n, \lambda))$.

З вищенаведеного одержуємо, що є оцінка максимальної правдоподібності параметра розподілу Пуассона.

Відповідь. $\lambda^* = \bar{x}_B$.

Приклад 17.5. Вибірка $x_1, x_2, \dots, x_n$ одержана із генеральної сукупності, розподіленої за нормальним законом з відповідними параметрами.

Знайдемо оцінки a^* і σ^* невідомих параметрів a і σ методом максимальної правдоподібності.

Розв'язання. Записуємо функцію правдоподібності:

$$L(x_1, x_2, \dots, x_n, a, \sigma^2) = \prod_{i=1}^n \left(\frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x_i - a)^2}{2 \cdot \sigma^2}} \right) = \frac{1}{(\sigma^2)^{\frac{n}{2}} \cdot (2 \cdot \pi)^{\frac{n}{2}}} \cdot e^{-\frac{\sum_{i=1}^n (x_i - a)^2}{2 \cdot \sigma^2}}.$$

Логарифмуючи, отримаємо:

$$\ln(L(a, \sigma)) = -\frac{n}{2} \cdot (\ln(2 \cdot \pi) + \ln(\sigma^2)) - \frac{1}{2 \cdot \sigma^2} \cdot \sum_{i=1}^n (x_i - a)^2.$$

Диференціюючи $\ln(L(a, \sigma))$ по a і σ^2 одержуємо систему рівнянь:

$$\begin{cases} \frac{1}{\sigma^2} \cdot \sum_{i=1}^n (x_i - a) = 0, \\ -\frac{n}{2 \cdot \sigma^2} + \left(\frac{1}{2 \cdot \sigma^4} \right) \cdot \sum_{i=1}^n (x_i - a)^2 = 0. \end{cases}$$

Звідки знаходимо оцінки:

$$a^* = \overline{x_B}, \sigma^* = s.$$

Ці оцінки співпадають з оцінками методу моментів. Вони обидві є слушні та незміщеними.

Відповідь. $a^* = \overline{x_B}, \sigma^* = s.$

Приклад 17.6. Вибірка $x_1, x_2, \dots, x_n$ одержана із генеральної сукупності, розподіленої за біноміальним законом з відповідним параметром. Знайти оцінку p^* параметра p методом максимальної правдоподібності.

Розв'язання. Записуємо функцію правдоподібності:

$$\begin{aligned} L(x_1, x_2, \dots, x_n, p) &= \prod_{i=1}^n (C_m^{x_i} \cdot p^{x_i} \cdot (1-p)^{m-x_i}) \\ &= (C_m^{x_1} \cdot \dots \cdot C_m^{x_n}) \cdot p^{(x_1 + \dots + x_n)} \cdot (1-p)^{(n \cdot m - (x_1 + \dots + x_n))}. \end{aligned}$$

Тоді

$$\ln(L(x_1, x_2, \dots, x_n, p)) = \ln(C_m^{x_1} \cdot \dots \cdot C_m^{x_n}) + (x_1 + \dots + x_n) \cdot \ln(p) + (n \cdot m - (x_1 + \dots + x_n)) \cdot \ln(1-p).$$

Знайдемо першу похідну по p :

$$\frac{d(\ln(L(x_1, x_2, \dots, x_n, p)))}{dp} = \frac{x_1 + \dots + x_n}{p} - \frac{n \cdot m - (x_1 + \dots + x_n)}{1-p}.$$

Прирівнюючи останню рівність до нуля матимемо:

$$p = \frac{x_1 + \dots + x_n}{n \cdot m} = \frac{\overline{x_B}}{m}.$$

Оскільки

$$\frac{d^2(\ln(L(x_1, x_2, \dots, x_n, p)))}{dp^2} = -\frac{x_1 + \dots + x_n}{p^2} - \frac{n \cdot m - (x_1 + \dots + x_n)}{(1-p)^2} < 0,$$

то $p^* = \frac{\bar{x}_B}{m}$ є точкою максимуму функції $\ln(L(x_1, x_2, \dots, x_n, p))$.

Відповідь. $p^* = \frac{\bar{x}_B}{m}$.

17.4. Метод найменших квадратів

Сутність *методу найменших квадратів* полягає у тому, що оцінки невідомих параметрів розподілу $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ визначаються із умови мінімізації суми квадратів відхилень вибірових даних x_i ($i = 1, \dots, n$) від оцінки, що визначається. Наприклад, знайдемо оцінку за методом найменших квадратів для генеральної середньої. Згідно з цим методом оцінку знаходимо з умови:

$$s(\theta) = \sum_{i=1}^n (x_i - \theta)^2 \rightarrow \min. \quad (17.9)$$

Використовуючи необхідну умову екстремуму функції, прирівнюємо до нуля похідну

$$\frac{ds}{d\theta} = -2 \cdot \sum_{i=1}^n (x_i - \theta) = 0 \text{ або } \sum_{i=1}^n x_i - n \cdot \theta = 0.$$

Звідки:

$$\theta^* = \bar{x}_B, \quad (17.10)$$

Отже, оцінка генеральної середньої є вибірова середня $\bar{x}_B$

Метод найменших квадратів має широке застосування у практиці статистичних досліджень, оскільки, по-перше, не потребує знання закону розподілу вибірових даних; по-друге, для нього досить добре розроблений математичний апарат чисельної реалізації.

Метод найменших квадратів застосовується у моделях кореляційного і регресійного аналізу.

17.5. Метод мінімуму χ^2

Нехай $x_1, x_2, \dots, x_n$ – вибірка незалежних спостережень над випадковою величиною X , розподіл якої належить класу розподілів $F(x, \theta)$, який залежить від невідомого параметра $\theta = (\theta_1, \theta_2, \dots, \theta_m)$. Вибірка одержана при деяких конкретних значеннях цього параметра.

Припустимо, множина значень X розбита на m інтервалів $\Delta_1, \dots, \Delta_m$, які не перетинаються. Позначимо через n_i число спостережень у вибірці, які потрапили у інтервал Δ_i . Якщо множина значень X скінченна, тобто величина приймає лише скінченне число значень, то можна вважати, що Δ_i – одноточкова множина. Таким чином, проведено групування результатів спостережень, у результаті чого одержано інтервальний варіаційний ряд.

Позначимо через $p_i(\theta) = F(\Delta_i, \theta)$, $i = 1, \dots, m$ ймовірність попадання значень випадкової величини X у i -й інтервал. Складемо величину χ^2 :

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - n \cdot p_i(\theta))^2}{n \cdot p_i(\theta)}. \quad (17.11)$$

Формула (17.11) служить оцінкою параметра на основі даної вибірки. Оскільки ймовірності $p_i(\theta)$ є функціями вибірових значень, то і величина χ^2 є функцією вибірки.

Оцінка $\theta(x_1, x_2, \dots, x_n)$ називається *оцінкою за методом мінімуму χ^2* , якщо вона одержана мінімізацією функції χ^2 .

Основні кроки методу мінімуму χ^2

Крок 1. Збір даних: Нехай маємо n спостережень випадкової величини X , які поділені на k інтервалів (або класів). Позначимо частоту спостережень у кожному інтервалі як $O_i = n_i$ (спостережувана частота).

Крок 2. Розрахунок теоретичних частот: Нехай $n \cdot p_i(\theta)$ – це ймовірність того, що випадкова величина X потрапить у i -й інтервал відповідно до теоретичного розподілу з параметрами θ . Тоді теоретичну частоту для i -го інтервалу можна знайти як $E_i = n \cdot p_i(\theta)$.

Крок 3. Обчислення статистики χ^2 : Статистика χ^2 обчислюється за формулою (17.11).

Крок 4. Мінімізація χ^2 : Параметри θ оцінюються шляхом мінімізації функції χ^2 .

Оптимальні значення параметрів – це ті, при яких χ^2 досягає мінімуму.

Для мінімізації функції χ^2 , зокрема, можна скористатися алгоритмом градієнтного спуску або функцією оптимізації з бібліотеки Python, такої як SciPy (рис. 17.1).

Метод мінімуму χ^2 часто застосовується у задачах статистичного моделювання, аналізу даних та оцінювання параметрів розподілів. Він дозволяє знизити вплив випадкових похибок і досягти більш точних результатів.

Приклад 17.7. Нехай ми маємо емпіричні дані, які поділені на k інтервалів. Необхідно оцінити параметри θ для нормального розподілу $N(a, \sigma^2)$.

Розв’язання. Наведемо поетапний хід розв’язку:

Збір даних: Припустимо, що ми маємо $n = 100$ спостережень і $k = 10$ інтервалів. Спостережувані частоти для кожного інтервалу такі:

[8; 12; 15; 14; 10; 9; 12; 8; 6; 6].

Розрахунок теоретичних частот: Припустимо, що початкові оцінки параметрів нормального розподілу a та σ дорівнюють 0 і 1 відповідно. Визначаємо теоретичні частоти для кожного інтервалу.

Обчислення статистики χ^2 : Обчислюємо $\chi^2(\theta)$ за формулою (17.11).

Мінімізація χ^2 : Мінімізуємо функцію $\chi^2(\theta)$ відносно параметрів a та σ для отримання оптимальних значень параметрів.

```
python
Копировать код

import numpy as np
from scipy.optimize import minimize
from scipy.stats import norm

# Спостережувані частоти
O = np.array([8, 12, 15, 14, 10, 9, 12, 8, 6, 6])

# Границі інтервалів (наприклад)
a = np.linspace(-3, 3, 11)

# Функція для обчислення теоретичних частот
def expected_frequencies(mu, sigma):
    E = []
    for i in range(len(a) - 1):
        E.append(100 * (norm.cdf(a[i+1], loc=mu, scale=sigma) - norm.cdf(a[i], loc=mu, scale=sigma)))
    return np.array(E)

# Функція для обчислення  $\chi^2$ 
def chi2(params):
    mu, sigma = params
    E = expected_frequencies(mu, sigma)
    return np.sum((O - E) ** 2 / E)

# Початкові оцінки параметрів
initial_params = [0, 1]

# Мінімізація
result = minimize(chi2, initial_params, method='BFGS')
mu_opt, sigma_opt = result.x

print(f"Оцінка параметра  $\mu$ : {mu_opt}")
print(f"Оцінка параметра  $\sigma$ : {sigma_opt}")
```

Рис. 17.1. Приклад мінімізації функції $\chi^2(\theta)$ в Python

Відповідь. Оптимальні значення a^* та σ^* .

17.6. Інтервальні статистичні оцінки.

Довірча ймовірність та довірчий інтервал

Точкові оцінки параметрів розподілу не дають можливості зробити висновки про їхню точність та надійність, оскільки є по суті випадковими величинами. Щоб мати уявлення про точність та надійність оцінки, у математичній статистиці користуються довірчими ймовірностями та довірчими інтервалами.

Інтервал $(-\theta^*; \theta^*)$ називається *довірчим* для оцінюваного параметра θ з довірчою ймовірністю (надійністю) γ , якщо він покриває цей параметр θ з ймовірністю, не меншою за γ , тобто

$$P\{-\theta^* < \theta < \theta^*\} \geq \gamma.$$

Межі $-\theta^*$ і θ^* довірчого інтервалу називаються довірчими межами для оцінюваного параметра θ . Статистична оцінка, що визначається двома числами — кінцями інтервалів, називається інтервальною. *Надійність γ* вибирається, як правило, достатньо великою, наприклад, 0,9; 0,95 або 0,99.

17.7. Схема побудови довірчого інтервалу для невідомого параметра a нормально розподіленої ознаки X при відомому параметрі σ

Довірчий інтервал матиме вигляд (17.12):

$$\left(\bar{x}_B - t \cdot \frac{\sigma}{\sqrt{n}}; \bar{x}_B + t \cdot \frac{\sigma}{\sqrt{n}} \right), \quad (17.12)$$

де:

$\bar{x}_B$ — вибіркова середня;

n — об'єм вибірки;

t — розв'язок рівняння $\Phi(t) = \frac{\gamma}{2}$ при заданій надійності γ .

Для функції Лапласа $\Phi(x)$ складено таблицю значень (дод. 2 Гмурмана), за допомогою якої за заданим значенням $\frac{\gamma}{2}$ функції знаходять значення аргументу t .

Величина $\delta = t \cdot \frac{\sigma}{\sqrt{n}}$ називається *похибкою*.

Приклад 17.8. Маємо такі дані про розміри виручки (у тис. грн) на 30 випадково вибраних рейсах:

42 24 49 67 45 27 39 21 58 40
28 78 44 66 20 62 70 81 7 68
94 76 63 88 65 14 46 20 72 91

1. Побудувати інтервальний статистичний розподіл із довжиною кроку $h = 20$ (тис. грн).

2. Припускаючи, що виручка розподілена за нормальним законом, знайти з надійністю $\gamma = 0,95$ довірчий інтервал для невідомого математичного сподівання, якщо середнє квадратичне відхилення $\sigma = 25$ (тис. грн).

Розв'язання. 1. Інтервальний статистичний розподіл буде таким (табл. 17.2):

Таблиця 17.2. Інтервальний варіативний ряд

$[x_i; x_{i+1})$	$[7; 27)$	$[27; 47)$	$[47; 67)$	$[67; 87)$	$[87; 107)$
n_i	6	8	6	7	3

Для визначення $\bar{x}_B$ необхідно побудувати дискретний статистичний розподіл, що має такий вигляд (табл. 17.3):

Таблиця 17.3. Дискретний варіативний ряд

x_i^*	17	37	57	77	97
n_i	6	8	6	7	3

де x_i^* — середини відповідних інтервалів.

Тоді за формулою (17.1)

$$\bar{x}_B = \frac{1}{30} \cdot (17 \cdot 6 + 37 \cdot 8 + 57 \cdot 6 + 77 \cdot 7 + 97 \cdot 3) = 52,3.$$

2. Для побудови довірчого інтервалу із заданою надійністю $\gamma = 0,95$ необхідно знайти t : $\Phi(t) = \frac{\gamma}{2} = 0,475$. Звідси $t = 1,96$ (дод. 2 Гмурмана).

Знаходимо межі інтервалу:

$$\begin{aligned} \bar{x}_B - t \cdot \frac{\sigma}{\sqrt{n}} &= 52,3 - 1,96 \cdot \frac{25}{\sqrt{30}} = 43,4, \\ \bar{x}_B + t \cdot \frac{\sigma}{\sqrt{n}} &= 52,3 + 1,96 \cdot \frac{25}{\sqrt{30}} = 61,2. \end{aligned}$$

Відповідь. 1. табл. 17.2; 2. з ймовірністю, не меншою за 0,95, середнє значення виручки на даному рейсі міститься в межах від 43,4 до 61,2 тис. грн.

Приклад 17.9. Визначити об'єм вибірки n нормально розподіленої випадкової величини, за якого похибка $\delta = 0,01$ при обчисленні математичного сподівання гарантується з ймовірністю 0,999, якщо середнє квадратичне відхилення $\sigma = 5$.

Розв'язання. Оскільки $\delta = \frac{t \cdot \sigma}{\sqrt{n}}$, то звідси $n = \frac{t^2 \cdot \sigma^2}{\delta^2}$. Значення t знаходимо з рівняння $\Phi(t) = \frac{\gamma}{2} = \frac{0,999}{2} = 0,4995$. За таблицею функції Лапласа (дод. 2 Гмурмана) знаходимо $t = 3,4$, отже, $n = 2890000$.

Відповідь. $n = 2890000$.

17.8. Схема побудови довірчого інтервалу для невідомого параметра α нормально розподіленої ознаки X при відомому параметрі s

Оскільки при розв'язуванні задач математичної статистики ми, як правило, маємо в розпорядженні лише вибірку, тобто параметр σ невідомий, то для знаходження наближених меж довірчого інтервалу замість σ використовується його вибіркова точкова оцінка s . При великих обсягах вибірки така заміна приводить до цілком прийнятних похибок.

При малих вибірках для оцінювання математичного сподівання якщо невідоме значення середнього квадратичного відхилення, застосовується випадкова величина t , що має розподіл Стюдента з $k = n - 1$ ступенями свободи та рівнем значущості $\alpha = 1 - \gamma$ (дод. 6 Гмурмана).

При цьому для кожного γ ($0 < \gamma < 1$) існує єдине значення t_γ , яке визначається за таблицею (дод. 3 Гмурмана) з для надійності γ та об'єму вибірки n .

Обчисливши за даним статистичним розподілом $\bar{x}_B$, s^2 і визначивши значення t_γ , будемо довірчий інтервал (17.13):

$$\left(\bar{x}_B - t_\gamma \cdot \frac{s}{\sqrt{n}}; \bar{x}_B + t_\gamma \cdot \frac{s}{\sqrt{n}} \right). \quad (17.13)$$

Приклад 17.10. Побудувати довірчий інтервал з надійністю $\gamma = 0,99$ для математичного сподівання максимальної швидкості (м/с) літака, припускаючи, що вона має нормальний розподіл, за даними 20 вимірювань:

504,5 485,2 512,0 497,1 502,4 488,3 491,9 489,3 502,7 509,2
514,3 502,3 497,8 511,3 515,7 506,3 499,2 497,4 485,2 507,8

Розв'язання. Для побудови довірчого інтервалу знайдемо середнє вибіркове і виправлену дисперсію.

$$\bar{x}_B = \frac{1}{20} \cdot (504,5 + 485,2 + 512,0 + 497,1 + 502,4 + 488,3 + 491,9 + 489,3 + 502,7 + 509,2 + 514,3 + 502,3 + 497,8 + 511,3 + 515,7 + 506,3 + 499,2 + 497,4 + 485,2 + 507,8) \approx 501;$$

$$s^2 = \frac{1}{19} \cdot ((504,5 - 501)^2 + (485,2 - 501)^2 + (512,0 - 501)^2 + (497,1 - 501)^2 + (502,4 - 501)^2 + (488,3 - 501)^2 + (491,9 - 501)^2 + (489,3 - 501)^2 + (502,7 - 501)^2 + (509,2 - 501)^2 + (514,3 - 501)^2 + (502,3 - 501)^2 + (497,8 - 501)^2 + (511,3 - 501)^2 + (515,7 - 501)^2 + (506,3 - 501)^2 + (499,2 - 501)^2 + (497,4 - 501)^2 + (485,2 - 501)^2 + (507,8 - 501)^2) \approx 89,6.$$

Відповідно до таблиці (дод. 3 Гмурмана) за даною надійністю $\gamma = 0,99$ і об'ємом вибірки $n = 20$ знаходимо значення $t_\gamma = 2,861$. Обчислимо межі довірчого інтервалу:

$$\begin{aligned}\bar{x}_B - t_\gamma \cdot \frac{s}{\sqrt{n}} &= 501 - \frac{2,861 \cdot \sqrt{89,6}}{\sqrt{20}} = 495; \\ \bar{x}_B + t_\gamma \cdot \frac{s}{\sqrt{n}} &= 501 + \frac{2,861 \cdot \sqrt{89,6}}{\sqrt{20}} = 507.\end{aligned}$$

Відповідь. З ймовірністю, не меншою за $\gamma = 0,99$, можна стверджувати, що математичне сподівання максимальної швидкості даного літака міститься в інтервалі (495; 507).

17.9. Схема 1 побудови довірчого інтервалу для невідомого параметра σ нормально розподіленої ознаки X при відомому параметрі s

При невідомому математичному сподіванні довірчий інтервал із надійністю γ для дисперсії $D(X)$ нормального розподілу має вигляд (17.14):

$$\left(\frac{k \cdot s^2}{\chi^2(k; \alpha_1)}; \frac{k \cdot s^2}{\chi^2(k; \alpha_2)} \right), \quad (17.14)$$

а для середнього квадратичного відхилення дістаємо (17.15):

$$\left(\frac{\sqrt{k \cdot s^2}}{\chi(k; \alpha_1)}; \frac{\sqrt{k \cdot s^2}}{\chi(k; \alpha_2)} \right), \quad (17.15)$$

де:

α_1, α_2 — рівні значущості;

$k = n - 1$ — ступінь свободи випадкової величини $\chi^2(k; \alpha)$.

Додатні числа $\chi^2(k; \alpha_1)$ та $\chi^2(k; \alpha_2)$ визначаються з рівностей (17.16):

$$F(\chi^2(k; \alpha_1)) = \alpha_1 = \frac{1-\gamma}{2}; \quad F(\chi^2(k; \alpha_2)) = \alpha_2 = \frac{1+\gamma}{2}, \quad (17.16)$$

де $F(\chi^2(k; \alpha))$ — функція розподілу випадкової величини $\chi^2(k; \alpha)$ з k ступенями свободи і рівнем значущості $\alpha \in [0,1]$.

Відповідне значення $\chi^2(k; \alpha)$ шукається як значення аргументу, при якому функція $F(\chi^2(k; \alpha))$ набуде значення α (дод. 5 Гмурмана) або за таблицею квантилів відповідного розподілу (рис. 17.2).

	0,01	0,025	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,975	0,99
1	0,0002	0,0010	0,0039	0,0158	0,0642	0,1485	0,2750	0,4549	0,7083	1,0742	1,6424	2,7055	3,8415	5,0239	6,6349
2	0,0201	0,0506	0,1026	0,2107	0,4463	0,7133	1,0217	1,3863	1,8326	2,4079	3,2189	4,6052	5,9915	7,3778	9,2103
3	0,1148	0,2158	0,3518	0,5844	1,0052	1,4237	1,8692	2,3660	2,9462	3,6649	4,6416	6,2514	7,8147	9,3484	11,3449
4	0,2971	0,4844	0,7107	1,0636	1,6488	2,1947	2,7528	3,3567	4,0446	4,8784	5,9886	7,7794	9,4877	11,1433	13,2767
5	0,5543	0,8312	1,1455	1,6103	2,3425	2,9999	3,6555	4,3515	5,1319	6,0644	7,2893	9,2364	11,0705	12,8325	15,0863
6	0,8721	1,2373	1,6354	2,2041	3,0701	3,8276	4,5702	5,3481	6,2108	7,2311	8,5581	10,6446	12,5916	14,4494	16,8119
7	1,2390	1,6899	2,1673	2,8331	3,8223	4,6713	5,4932	6,3458	7,2832	8,3834	9,8032	12,0170	14,0671	16,0128	18,4753
8	1,6465	2,1797	2,7326	3,4895	4,5936	5,5274	6,4226	7,3441	8,3505	9,5245	11,0301	13,3616	15,5073	17,5345	20,0902
9	2,0879	2,7004	3,3251	4,1682	5,3801	6,3933	7,3570	8,3428	9,4136	10,6564	12,2421	14,6837	16,9190	19,0228	21,6660
10	2,5582	3,2470	3,9403	4,8652	6,1791	7,2672	8,2955	9,3418	10,4732	11,7807	13,4420	15,9872	18,3070	20,4832	23,2093
11	3,0535	3,8157	4,5748	5,5778	6,9887	8,1479	9,2373	10,3410	11,5298	12,8987	14,6314	17,2750	19,6751	21,9200	24,7250
12	3,5706	4,4038	5,2260	6,3038	7,8073	9,0343	10,1820	11,3403	12,5838	14,0111	15,8120	18,5493	21,0261	23,3367	26,2170
13	4,1069	5,0088	5,8919	7,0415	8,6339	9,9257	11,1291	12,3398	13,6356	15,1187	16,9848	19,8119	22,3620	24,7356	27,6882
14	4,6604	5,6287	6,5706	7,7895	9,4673	10,8215	12,0785	13,3393	14,6853	16,2221	18,1508	21,0641	23,6848	26,1189	29,1412
15	5,2293	6,2621	7,2609	8,5468	10,3070	11,7212	13,0297	14,3389	15,7332	17,3217	19,3107	22,3071	24,9958	27,4884	30,5779
16	5,8122	6,9077	7,9616	9,3122	11,1521	12,6243	13,9827	15,3385	16,7795	18,4179	20,4651	23,5418	26,2962	28,8454	31,9999
17	6,4078	7,5642	8,6718	10,0852	12,0023	13,5307	14,9373	16,3382	17,8244	19,5110	21,6146	24,7690	27,5871	30,1910	33,4087
18	7,0149	8,2307	9,3905	10,8649	12,8570	14,4399	15,8932	17,3379	18,8679	20,6014	22,7595	25,9894	28,8693	31,5264	34,8053
19	7,6327	8,9065	10,1170	11,6509	13,7158	15,3517	16,8504	18,3377	19,9102	21,6891	23,9004	27,2036	30,1435	32,8523	36,1909
20	8,2604	9,5908	10,8508	12,4426	14,5784	16,2659	17,8088	19,3374	20,9514	22,7745	25,0375	28,4120	31,4104	34,1696	37,5662
21	8,8972	10,2829	11,5913	13,2396	15,4446	17,1823	18,7683	20,3372	21,9915	23,8578	26,1711	29,6151	32,6706	35,4789	38,9322
22	9,5425	10,9823	12,3380	14,0415	16,3140	18,1007	19,7288	21,3370	23,0307	24,9390	27,3015	30,8133	33,9244	36,7807	40,2894
23	10,1957	11,6886	13,0905	14,8480	17,1865	19,0211	20,6902	22,3369	24,0689	26,0184	28,4288	32,0069	35,1725	38,0756	41,6384
24	10,8564	12,4012	13,8484	15,6587	18,0618	19,9432	21,6525	23,3367	25,1063	27,0960	29,5533	33,1962	36,4150	39,3641	42,9798

Рис. 17.2. Фрагмент таблиці квантилів розподілу $\chi^2(k; \alpha)$

Слід зазначити, що в даній таблиці, випадкова величина, що має розподіл типу χ^2 з k ступенями свободи і рівнем значущості α , не перевищує значення $\chi^2(k; \alpha)$ з ймовірністю α і перевищує його з ймовірністю $1 - \alpha$. Даний факт впливає прямо з означення квантилів. Це означає, що наприклад рівню значущості $\alpha = 0,1$ та степеню свободи $k = 10$ відповідає $\chi^2(k; 1 - \alpha) = \chi^2(10; 0,9) = 15,9872$.

На рис. 17.3. наведено приклад розрахунку квантилів розподілу $\chi^2(k; \alpha)$ за допомогою відповідної функція в програмному забезпеченні для статистичного аналізу (у Python). Деякі розраховані квантелі представлено на рис. 17.4.

```
python
import scipy.stats as stats

# Задати рівні значущості та ступені свободи
alphas = [0.10, 0.05, 0.01]
degrees_of_freedom = [1, 5, 10, 20, 30, 40, 50]

# Розрахувати квантили
quantiles = {}
for df in degrees_of_freedom:
    quantiles[df] = {}
    for alpha in alphas:
        quantile = stats.chi2.ppf(1 - alpha, df)
        quantiles[df][alpha] = quantile

# Вивести результати
for df in degrees_of_freedom:
    print(f"Ступені свободи: {df}")
    for alpha in alphas:
        print(f"Квантиль для рівня значущості {alpha}: {quantiles[df][alpha]:.4f}")
```

Рис. 17.3. Приклад розрахунку квантилів розподілу $\chi^2(k; \alpha)$ у Python

	alpha=0.1	alpha=0.05	alpha=0.01
df			
1	2.705544	3.841459	6.634897
2	4.605170	5.991465	9.210340
3	6.251389	7.814728	11.344867
4	7.779440	9.487729	13.276704
5	9.236357	11.070498	15.086272
6	10.644641	12.591587	16.811894
7	12.017036	14.067140	18.475307
8	13.361566	15.507313	20.090235
9	14.683657	16.918978	21.665994
10	15.987179	18.307038	23.209251
11	17.275013	19.675138	24.725025
12	18.549348	21.026070	26.217000
13	19.811929	22.362032	27.688250
14	21.064140	23.684791	29.141238
15	22.307130	24.995790	30.577914
16	23.541856	26.296228	31.999927
17	24.769037	27.587112	33.408664
18	25.989420	28.869299	34.805306
19	27.203575	30.143527	36.190869

Рис. 17.4. Отримані результати (квантелі)

Приклад 17.11. Під час вибіркового аналізу максимального завантаження літака при певній кількості пального, дістали такі дані, т:

78 80 85 88 89 90 88 93
88 85 93 89 80 90 85 89

Припускаючи, що максимальне завантаження літака має нормальний розподіл, знайти довірчий інтервал для дисперсії та середнього квадратичного відхилення з надійністю $\gamma = 0,95$.

Розв'язання. Для побудови довірчих інтервалів необхідно знайти s^2 . Для цього обчислимо значення $\bar{x}_B$.

$$\bar{x}_B = \frac{1}{16} \cdot (78 \cdot 1 + 80 \cdot 2 + 85 \cdot 3 + 88 \cdot 3 + 89 \cdot 3 + 90 \cdot 2 + 92 \cdot 1 + 93 \cdot 1) = 86,8,$$

$$s^2 = \frac{1}{15} \cdot ((78 - 86,8)^2 \cdot 1 + (80 - 86,8)^2 \cdot 2 + (85 - 86,8)^2 \cdot 3 + (88 - 86,8)^2 \cdot 3 + (89 - 86,8)^2 \cdot 3 + (90 - 86,8)^2 \cdot 2 + (92 - 86,8)^2 \cdot 1 + (93 - 86,8)^2 \cdot 1) = 19.$$

За формулами (17.16) дістаємо:

$$F(\chi^2(15, \alpha_1)) = \frac{1 - 0,95}{2} = 0,025; F(\chi^2(15, \alpha_2)) = \frac{1 + 0,95}{2} = 0,975.$$

За таблицею значень величини $\chi^2(k, \alpha)$ (дод. 5 Гмурмана) з $k = n - 1 = 15$ та $\alpha_1 = 0,025$, $\alpha_2 = 0,975$ знаходимо $\chi_1^2(0,025; 15) \approx 27,5$, $\chi_2^2(0,975; 15) \approx 6,26$ і за формулами (17.14) та (17.15) дістаємо довірчі інтервали відповідно для дисперсії та середнього квадратичного відхилення $(10,4; 45,5)$ і $(3,2; 6,7)$.

Відповідь. $(10,4; 45,5)$ і $(3,2; 6,7)$.

17.10. Схема 2 побудови довірчого інтервалу для невідомого параметра σ нормально розподіленої ознаки X при відомому параметрі s

Інтервальною оцінкою з надійністю γ середнього квадратичного відхилення σ нормально розподіленої кількісної ознаки X по виправленому середньому квадратичному відхиленню s є довірчий інтервал (17.17):

$$\begin{aligned} s \cdot (1 - q) < \sigma < s \cdot (1 + q) \text{ при } q < 1, \\ 0 < \sigma < s \cdot (1 + q) \text{ при } q > 1, \end{aligned} \quad (17.17)$$

де q знаходиться за таблицею в дод. 4 Гмурмана за заданими n та γ .

Приклад 17.12. Під час вибіркового аналізу максимального завантаження літака при певній кількості пального, дістали такі дані, т:

78 80 85 88 89 90 88 93
88 85 93 89 80 90 85 89

Припускаючи, що максимальне завантаження літака має нормальний розподіл, знайти довірчий інтервал для середнього квадратичного відхилення з надійністю $\gamma = 0,95$.

Розв'язання. У прикладі 17.11. наведено обчислення s . З дод. 4 Гмурмана отримуємо $q = 0,44$. Згідно (17.17) довірчий інтервал для середнього квадратичного відхилення $(2,9; 6,3)$.

Відповідь. (2,9; 6,3).

17.11. Схема побудови довірчого інтервалу для невідомої ймовірності p біноміально розподіленої ознаки X

Інтервальною оцінкою з надійністю γ невідомої ймовірності p біноміального розподілу по відносній частоті w є довірчий інтервал (17.18):

$$p_1 < p < p_2. \quad (17.18)$$

В нерівності (17.18) отримуємо (17.19):

$$\begin{aligned} p_1 &= \frac{n}{t^2 + n} \cdot \left(w + \frac{t^2}{2 \cdot n} - \sqrt{\frac{w \cdot (1 - w)}{n} + \left(\frac{t}{2 \cdot n} \right)^2} \right), \\ p_2 &= \frac{n}{t^2 + n} \cdot \left(w + \frac{t^2}{2 \cdot n} + \sqrt{\frac{w \cdot (1 - w)}{n} + \left(\frac{t}{2 \cdot n} \right)^2} \right), \end{aligned} \quad (17.19)$$

де:

n – загальна кількість випробувань;

m – число появ події;

$w = \frac{m}{n}$ – відносна частота;

t — розв’язок рівняння $\Phi(t) = \frac{\gamma}{2}$ при заданій надійності γ .

Зауваження: при великих значеннях n (порядку сотні) в нерівності (17.19) отримуємо (17.20):

$$p_1 = w - t \cdot \sqrt{\frac{w \cdot (1 - w)}{n}}, \quad p_2 = w + t \cdot \sqrt{\frac{w \cdot (1 - w)}{n}}. \quad (17.20)$$

Приклад 17.13. Проводиться 60 незалежних випробувань. В них подія A з’являється 15 разів. Отримати інтервальну оцінку з надійністю $\gamma = 0,95$ невідомої ймовірності p .

Розв’язання. Згідно умови задачі $n = 60, m = 15 \Rightarrow w = 0,25$.

$$\Phi(t) = \frac{0,95}{2} = 0,475, \text{ тобто } t = 1,96 \text{ (див. дод. 2 Гмурмана).}$$

Згідно (17.20) довірчий інтервал для ймовірності p – (0,16; 0,37).

Відповідь. (0,16; 0,37)

17.12. Порівняння точкових та інтервальних оцінок

Вибір між точковими та інтервальними оцінками залежить від контексту та цілей аналізу:

- інтервальні оцінки зазвичай кращі, оскільки вони дають не просто число, а межі можливого значення, що робить їх надійнішими;
- точкові оцінки корисні, коли потрібно швидко отримати приблизне значення.

Результати їх порівняння наведено в табл. 17.4.

Таблиця 17.4. Порівняння точкових та інтервальних статистичних оцінок

Критерій	Точкова оцінка	Інтервальна оцінка
Визначення	Один числовий показник, що оцінює параметр розподілу	Діапазон значень, у якому з певною

		ймовірністю міститься істинний параметр
Приклад	Середнє значення вибірки як оцінка математичного сподівання	Довірчий інтервал для середнього
Точність	Дає лише одну можливу оцінку параметра, може бути неточною	Вказує можливий діапазон значень, що підвищує надійність
Ризик похибки	Високий, оскільки не враховує варіативність вибірки	Контролює рівень довіри (наприклад, 95% довірчий інтервал).
Використання	Швидкі оцінки параметра в експериментах і дослідженнях.	Аналіз точності та надійності оцінки.

ЛЕКЦІЯ 18

ПЕРВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

18.1. Основні поняття

Оскільки висновки про параметри і закони розподілу ознаки X генеральної сукупності робляться на основі обмеженого вибіркового матеріалу, вони не можуть бути категоричними і мають сприйматись лише як деякі припущення. Такі припущення про закони розподілу або їхні параметри називаються статистичними гіпотезами. Наприклад, статистичною гіпотезою є припущення: ознака X генеральної сукупності розподілена за нормальним законом.

Задача статистичної перевірки гіпотези полягає в побудові за вибірковими даними статистичного критерію (критерію узгодження), який дає змогу прийняти або відхилити висунуту гіпотезу.

Статистичною називають гіпотезу щодо типу невідомого розподілу або щодо параметрів відомого розподілу.

Висунуту гіпотезу H_0 називається *основною* або *нульовою*.

Конкуруючою (*альтернативною*) називають гіпотезу H_1 , яка суперечить нульовій.

Розрізняють гіпотези, які містять одне або більше припущень:

- *Простою* називають гіпотезу, що містить одне припущення;
- *Складною* називають гіпотезу, що містить скінчену або нескінчену кількість простих гіпотез.

Число α також називається *ймовірністю помилки 1-го роду*, яка полягає у відкиданні вірної нульової гіпотези.

Число β називається *ймовірністю помилки 2-го роду*, яка полягає у прийнятті невірної нульової гіпотези.

Статистичним критерієм називається випадкова величина K який призначений для перевірки гіпотез.

Спостережуваним (емпіричне) значенням називають те значення $K_{\text{сп}}$, яке обчислено по вибіркам.

Критичною областю називають сукупність значень критерію, при яких нульову гіпотезу відкидають.

Областю прийняття гіпотези називають сукупність значень критерію, при яких нульову гіпотезу приймають.

Критичною точкою називають число $K_{\text{кр}}$, яке відрізняють критичну область від області прийняття гіпотез.

Правосторонньою називається критична область, що визначається нерівністю $K > K_{\text{кр}}$, де $K_{\text{кр}}$ – додатне число.

Лівосторонньою називається критична область, що визначається нерівністю $K < K_{\text{кр}}$, де $K_{\text{кр}}$ – від'ємне число.

Двосторонньою називається критична область, що визначається нерівністю $|K| > K_{кр}$.

Для визначення критичної області задається рівень значущості та шукають критичні точки з наступних умов:

1. Для правосторонньої області (18.1):

$$P(K > K_{кр}) = \alpha \quad (K_{кр} > 0); \quad (18.1)$$

2. Для лівосторонньої області (18.2):

$$P(K < K_{кр}) = \alpha \quad (K_{кр} < 0); \quad (18.2)$$

3. Для двосторонньої симетричної області (18.3):

$$P(K > K_{кр}) = \frac{\alpha}{2}, P(K < -K_{кр}) = \frac{\alpha}{2} \quad (K_{кр} > 0). \quad (18.3)$$

Потужністю критерію називають ймовірність попадання в критичну область, за умовою, що справедлива конкуруюча гіпотеза.

Типи статистичних критеріїв

1. В яких в умові задано випадкову величину з відомим законом розподілу. Нульова гіпотеза полягає у рівності певних числових характеристик генеральної сукупності (математичне сподівання, дисперсія, коефіцієнт кореляції, тощо) або їхня рівність наперед заданому гіпотетичному значенню.

2. В яких в умові дано вибірку, взятую з певної генеральної сукупності. Нульова гіпотеза полягає у припущенні щодо типу розподілу всієї генеральної сукупності.

Етапи перевірки статистичної гіпотези

1. Формулювання нульової та альтернативної гіпотез.
2. Розрахунок необхідних статистичних характеристик.
3. Вибір рівня значущості α .
4. Вибір та обчислення розрахункового значення критерію для перевірки гіпотез. Критичне значення визначається за відповідною таблицею.
5. Порівняння розрахункового і табличного значень критерію, на основі чого робиться висновок про прийняття чи відхилення гіпотези.

ПЕРЕВІРКА СТАТИСТИЧНОЇ ГІПОТЕЗИ ЩОДО ПАРАМЕТРІВ ВІДОМОГО РОЗПОДІЛУ

Для зручності, надалі генеральні сукупності в статистичних критеріях будуть позначатися великими латинськими літерами, а вибірки з них – малими. Частоти елементів вибірок вважатимемо рівними 1.

18.2. Одновибірковий t -критерій (критерій Стюдента)

Одновибірковий t -критерій Стюдента застосовується для перевірки нульової гіпотези про рівність математичного сподівання генеральної сукупності деякому відомому значенню.

Правило 1

Збір даних: з генеральної сукупності зібрати дані вибірки $x = \{x_1, \dots, x_n\}$ і розрахувати її середнє значення $\bar{x}_B$ та стандартне відхилення s .

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = a_0$ – середнє значення генеральної сукупності дорівнює певному значенню a_0 ;
- альтернативна гіпотеза $H_1: M(X) \neq a_0$ – середнє значення генеральної сукупності не дорівнює a_0 .

Розрахунок спостережуваного значення: спостережуване значення розраховується за формулою (18.4):

$$t_{\text{сп}} = \frac{\sqrt{n} \cdot (\bar{x}_B - a_0)}{s}, \quad (18.4)$$

де:

$\bar{x}_B$ — середнє значення вибірки;

a_0 — відоме середнє значення популяції;

s — стандартне відхилення вибірки, n — об'єм вибірки (малий).

Визначення критичного значення: критичне значення $t_{\text{двост.кр}}$ визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n - 1$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $|t_{\text{сп}}| > t_{\text{двост.кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $|t_{\text{сп}}| < t_{\text{двост.кр}}$, не відхиляти нульову гіпотезу.

Правило 2

Збір даних: аналогічно до правила 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = a_0$ – середнє значення генеральної сукупності дорівнює певному значенню a_0 ;
- альтернативна гіпотеза $H_1: M(X) > a_0$ – середнє значення генеральної сукупності більше, ніж a_0 .

Розрахунок спостережуваного значення: спостережуване значення розраховується за формулою (18.4).

Визначення критичного значення: критичне значення $t_{\text{прав.кр}}$ визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n - 1$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $t_{\text{сп}} > t_{\text{прав.кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $t_{\text{сп}} < t_{\text{прав.кр}}$, не відхиляти нульову гіпотезу.

Правило 3

Збір даних: аналогічно до правила 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = a_0$ – середнє значення генеральної сукупності дорівнює певному значенню a_0 ;

- альтернативна гіпотеза $H_1: M(X) < a_0$ – середнє значення генеральної сукупності менша, ніж a_0 .

Розрахунок спостережуваного значення: спостережуване значення розраховується за формулою (18.4)

Визначення критичного значення: критичне значення $t_{\text{прав.кр}}$ визначається з таблиці розподілу Ст'юдента для заданого рівня значущості α і ступеня свободи $k = n - 1$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $t_{\text{сп}} < -t_{\text{прав.кр}}$, відхилити нульову гіпотезу;

- якщо розраховане значення $t_{\text{сп}} > -t_{\text{прав.кр}}$, не відхиляти нульову гіпотезу.

Одновибірковий t -критерій Ст'юдента є потужним інструментом для порівняння середнього значення вибірки з відомим середнім значенням популяції. Його використання дозволяє зробити висновки про статистичну значущість відмінностей у даних, що може бути корисним у різних сферах, таких як медицина, соціологія, економіка тощо.

Приклад 18.1. Необхідно при рівні значущості $\alpha = 0,05$ перевірити, чи відрізняється середнє значення ваги певної групи (сукупності) людей від 70 кг. Маємо вибірку з 10 людей, середнє значення ваги яких становить 72 кг, а стандартне відхилення – 5 кг. Альтернативна гіпотеза $H_1: M(X) \neq a_0$.

Розв'язання. Застосуємо одновибірковий t -критерій Ст'юдента при рівні значущості. Скористаємося правилом 1 даного критерію.

Збір даних: $\bar{x}_B = 72$, $a_0 = 70$, $s = 5$, $n = 10$.

Формулювання гіпотез: $H_0: M(X) = 70$, $H_1: M(X) \neq 70$.

Розрахунок спостережуваного значення: $t_{\text{сп}} = \frac{\sqrt{10} \cdot (72 - 70)}{5} \approx 1,265$.

Визначення критичного значення: для рівня значущості $\alpha = 0,05$ і ступенів свободи $k = 10 - 1 = 9$, критичне значення дорівнює $t_{\text{двост.кр}}(9; 0,05) \approx 2,262$.

Прийняття рішення: Розраховане значення $t_{\text{сп}} \approx 1,265$ менше критичного значення $t_{\text{двост.кр}} \approx 2,262$. Отже, немає достатніх підстав відхилити нульову гіпотезу. Середнє значення вибірки не відрізняється статистично значущо від 70 кг.

Відповідь. Нульова гіпотеза приймається.

18.3. Критерій порівняння двох середніх генеральних сукупностей, дисперсії яких відомі

Критерій порівняння двох середніх генеральних сукупностей з відомими дисперсіями для випадку двох незалежних вибірок x ; y великих об'ємів ($n, m \geq 30$) використовується для порівняння їх середніх значень при відомих дисперсіях генеральних сукупностей. Він допомагає визначити,

чи є статистично значуща різниця між середніми значеннями двох груп (сукупностей).

Правило 1

Збір даних: з нормально розподілених генеральних сукупностей, які мають дисперсії $D(X), D(Y)$ відповідно, зібрати дані вибірок $x = \{x_1, \dots, x_n\}; y = \{y_1, \dots, y_m\}$ ($n, m \geq 30$) і розрахувати їх середні значення $\bar{x}_B, \bar{y}_B$.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) \neq M(Y)$ – середні значення двох генеральних сукупностей не рівні.

Розрахунок спостережуваного значення: спостережуване значення для даного критерію розраховується за формулою (18.5):

$$z_{\text{сп}} = \frac{\bar{x}_B - \bar{y}_B}{\sqrt{\frac{D(X)}{n} + \frac{D(Y)}{m}}}, \quad (18.5)$$

де:

$\bar{x}_B, \bar{y}_B$ — середні значення вибірок;

$D(X), D(Y)$ — дисперсії генеральних сукупностей;

n, m — об'єми вибірок ($n, m \geq 30$).

Визначення критичного значення: критичне значення $z_{\text{кр}}$ для заданого рівня значущості α визначається з рівняння (18.6):

$$\Phi(z_{\text{кр}}) = \frac{1-\alpha}{2}, \quad (18.6)$$

де $\Phi(z_{\text{кр}})$ – функція Лапласа (див. дод. 2 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $|z_{\text{сп}}| > z_{\text{кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $|z_{\text{сп}}| < z_{\text{кр}}$, не відхиляти нульову гіпотезу.

Правило 2

Збір даних: аналогічно до правила 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) > M(Y)$ – середнє значення генеральної сукупності X більше, ніж середнє значення генеральної сукупності Y .

Розрахунок спостережуваного значення: спостережуване значення для даного критерію розраховується за формулою (18.5).

Визначення критичного значення: критичне значення $z_{\text{кр}}$ для заданого рівня значущості α визначається з рівняння (18.7):

$$\Phi(z_{\text{кр}}) = \frac{1-2\cdot\alpha}{2}, \quad (18.7)$$

де $\Phi(z_{\text{кр}})$ – функція Лапласа (див. дод. 2 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $z_{\text{сп}} > z_{\text{кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $z_{\text{сп}} < z_{\text{кр}}$, не відхиляти нульову гіпотезу.

Правило 3

Збір даних: аналогічно до правила 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) < M(Y)$ – середнє значення генеральної сукупності X менше, ніж середнє значення генеральної сукупності Y .

Розрахунок спостережуваного значення: спостережуване значення для даного критерію розраховується за формулою (18.5).

Визначення критичного значення: критичне значення $z_{\text{кр}}$ для заданого рівня значущості α визначається за формулою (18.7).

Прийняття рішення:

- якщо розраховане значення $z_{\text{сп}} < -z_{\text{кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $z_{\text{сп}} > -z_{\text{кр}}$, не відхиляти нульову гіпотезу.

Приклад 18.2. Необхідно перевірити при рівні значущості $\alpha = 0,05$, чи відрізняються середні значення балів двох груп студентів на екзамені. Генеральні дисперсії $D(X) = 6, D(Y) = 5$. Дані для двох вибірок з цих груп:

Вибірка x (група 1): $\bar{x}_B = 85, n = 36$;

Вибірка y (група 2): $\bar{y}_B = 80, m = 35$.

Альтернативна гіпотеза $H_1: M(X) \neq M(Y)$.

Розв’язання. Скористаємося правилом 1 даного критерію.

Збір даних: $\bar{x}_B = 85, n = 36, \bar{y}_B = 80, m = 35, D(X) = 6, D(Y) = 5$.

Формулювання гіпотез: $H_0: M(X) = M(Y), H_1: M(X) \neq M(Y)$.

Розрахунок спостережуваного значення: $z_{\text{сп}} = \frac{85-80}{\sqrt{\frac{6}{36} + \frac{5}{35}}} \approx 8,99$

Визначення критичного значення: Для рівня значущості $\alpha = 0,05$, критичне значення $z_{\text{кр}} \approx 1,96$.

Прийняття рішення: Розраховане значення $z_{\text{сп}} \approx 8,99$ більше критичного значення $z_{\text{кр}} \approx 1,96$. Отже, ми відхиляємо нульову гіпотезу. Є статистично значуща різниця між середніми значеннями балів двох груп студентів.

Відповідь. Нульова гіпотеза не приймається.

18.4. Двовибірковий t -критерій для незалежних вибірок

Двовибірковий t -критерій для незалежних вибірок використовується для порівняння середніх значень двох нормально розподілених генеральних сукупностей, з невідомими але однаковими дисперсіями, якщо відомі вибіркові середні з двох незалежних вибірок $x; y$ малих об’ємів ($n, m < 30$),

зроблених з них. Він допомагає визначити, чи є статистично значуща різниця між середніми значеннями двох груп.

Правило 1

Збір даних: зібрати дані для двох незалежних вибірок $x = \{x_1, \dots, x_n\}$; $y = \{y_1, \dots, y_m\}$ ($n, m < 30$) і розрахувати їхні середнє значення $\bar{x}_B$, $\bar{y}_B$ та стандартне відхилення s_x , s_y .

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) \neq M(Y)$ – середні значення двох генеральних сукупностей не рівні.

Розрахунок спостережуваного значення: спостережуване значення для двовибіркового t -критерію розраховується за формулою (18.8):

$$t_{\text{сп}} = \frac{\bar{x}_B - \bar{y}_B}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \cdot \frac{(n-1) \cdot s_x^2 + (m-1) \cdot s_y^2}{n+m-2}}}, \quad (18.8)$$

де:

$\bar{x}_B, \bar{y}_B$ — середні значення вибірок;

n, m — об'єми вибірок;

s_x, s_y — стандартні відхилення вибірок.

Визначення критичного значення: критичне значення $t_{\text{двост.кр}}$ визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n + m - 2$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $|t_{\text{сп}}| > t_{\text{двост.кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $|t_{\text{сп}}| < t_{\text{двост.кр}}$, не відхиляти нульову гіпотезу.

Правило 2

Збір даних: аналогічно правилу 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) > M(Y)$ – середнє значення генеральної сукупності X більше, ніж середнє значення генеральної сукупності Y .

Розрахунок спостережуваного значення: спостережуване значення для двовибіркового t -критерію розраховується за формулою (18.8).

Визначення критичного значення: критичне значення $t_{\text{прав.кр}}$ визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n + m - 2$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $t_{\text{сп}} > t_{\text{прав.кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $t_{\text{сп}} < t_{\text{прав.кр}}$, не відхиляти нульову гіпотезу.

Правило 3

Збір даних: аналогічно правилу 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(X) = M(Y)$ – середні значення двох генеральних сукупностей рівні;
- альтернативна гіпотеза $H_1: M(X) < M(Y)$ – середнє значення генеральної сукупності X менше, ніж середнє значення генеральної сукупності Y .

Розрахунок спостережуваного значення: спостережуване значення для двовибіркового t -критерію розраховується за формулою (18.8).

Визначення критичного значення: критичне значення $t_{\text{прав.кр}}$ визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n + m - 2$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо розраховане значення $t_{\text{сп}} < -t_{\text{прав.кр}}$, відхилити нульову гіпотезу;
- якщо розраховане значення $t_{\text{сп}} > -t_{\text{прав.кр}}$, не відхиляти нульову гіпотезу.

Двовибірковий t -критерій для незалежних вибірок є корисним інструментом для порівняння середніх значень двох груп. Він допомагає зробити висновки про значущість відмінностей між групами, що може бути корисним у багатьох дослідницьких сферах, таких як психологія, медицина, соціальні науки тощо.

Приклад 18.3. *Необхідно перевірити при рівні значущості $\alpha = 0,05$, чи відрізняються середні значення балів двох груп студентів на екзамені. Дані для двох незалежних вибірок:*

Вибірка x (група 1): $\bar{x}_B = 85, s_x = 5, n = 15$;

Вибірка y (група 2): $\bar{y}_B = 80, s_y = 6, m = 15$.

Альтернативна гіпотеза $H_1: M(X) \neq M(Y)$

Розв'язання. За правилом 1 даного критерію:

Збір даних: $\bar{x}_B = 85, s_x = 5, n = 15, \bar{y}_B = 80, s_y = 6, m = 15$.

Формулювання гіпотез: $H_0: M(X) = M(Y), H_1: M(X) \neq M(Y)$.

Розрахунок спостережуваного значення:

$$t_{\text{сп}} = \frac{85 - 80}{\sqrt{\left(\frac{1}{15} + \frac{1}{15}\right) \cdot \frac{14 \cdot 25 + 14 \cdot 36}{15 + 15 - 2}}} \approx 2,49.$$

Визначення критичного значення: Для рівня значущості $\alpha = 0,05$ і ступенів свободи $k = 15 + 15 - 2 = 28$, критичне значення можна знайти в таблиці розподілу Стюдента $t_{\text{двост.кр}} \approx 2,048$.

Прийняття рішення: Розраховане значення $t_{\text{сп}} \approx 2,49$ більше критичного значення $t_{\text{двост.кр}} \approx 2,048$. Отже, ми відхиляємо нульову гіпотезу. Є статистично значуща різниця між середніми значеннями балів двох груп студентів.

Відповідь. Нульова гіпотеза не приймається.

18.5. Двовибірковий t -критерій для залежних вибірок

Двовибірковий t -критерій для залежних вибірок використовується для порівняння середніх значень двох залежних вибірок x ; y , тобто коли кожен елемент однієї вибірки пов'язаний з конкретним елементом іншої вибірки. Це критерій часто використовується для аналізу даних до і після певного втручання або для парних спостережень.

Правило

Збір даних: зібрати дані для двох залежних вибірок $x = \{x_1, \dots, x_n\}$; $y = \{y_1, \dots, y_n\}$ і розрахувати різниці $d_i = x_i - y_i$ ($i = \overline{1, n}$) для кожної пари спостережень. З елементів d_i складаємо вибірку.

Формулювання гіпотез:

- нульова гіпотеза $H_0: M(d) = 0$ – середнє значення вибірки d дорівнює нулю (середні значення двох x ; y рівні);
- альтернативна гіпотеза $H_1: M(d) \neq 0$ – середнє значення вибірки d не дорівнює нулю (середні значення двох вибірок x ; y не рівні).

Розрахунок спостережуваного значення: спостережуване значення для двовибіркового t -критерію для залежних вибірок розраховується за формулою (18.9):

$$t_{\text{сп}} = \frac{\overline{d_B} \cdot \sqrt{n}}{s_d}, \quad (18.9)$$

де:

$\overline{d_B} = \frac{\sum_{i=1}^n d_i}{n}$ — середнє значення різниць парних спостережень;

$s_d = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \overline{d_B})^2}$ — стандартне відхилення різниць;

n — кількість пар спостережень.

Визначення критичного значення: критичне значення $t_{\text{двост.кр}}$, визначається з таблиці розподілу Стюдента для заданого рівня значущості α і ступеня свободи $k = n - 1$.

Прийняття рішення:

- якщо розраховане значення $|t_{\text{сп}}| > t_{\text{двост.кр}}$ (або знаходиться в критичній області), відхилити нульову гіпотезу;
- якщо розраховане значення $|t_{\text{сп}}| < t_{\text{двост.кр}}$, не відхиляти нульову гіпотезу.

Двовибірковий t -критерій для залежних вибірок є важливим інструментом для аналізу парних спостережень. Він дозволяє зробити висновки про ефективність втручання або методики, що використовується, і допомагає визначити, чи є різниця між двома залежними вибірками статистично значущою.

Приклад 18.4. Перевірити ефективність нової методики навчання для рівня значущості $\alpha = 0,05$. Наведено дані балів до та після застосування методики для 10 студентів (табл. 18.1):

Таблиця 18.1. Дані балів для студентів

Студент	1	2	3	4	5	6	7	8	9	10
До	70	68	75	72	69	74	73	71	72	70
Після	75	70	78	74	73	76	77	74	73	71

Розв'язання. Скористаємося двовибірковим t -критерієм для залежних вибірок.

Обчислення різниць: складаємо вибірку d за принципом:

$$d_i = x_{i \text{ після}} - x_{i \text{ до}}.$$

$$d: 5, 2, 3, 2, 4, 2, 4, 3, 1, 1.$$

Формулювання гіпотез: $H_0: M(d) = 0$, $H_1: M(d) \neq 0$.

Середнє значення різниць $\bar{d}_B$:

$$\bar{d}_B = \frac{5 + 2 + 3 + 2 + 4 + 2 + 4 + 3 + 1 + 1}{10} = 2,7$$

Стандартне відхилення різниць s_d :

$$s_d = \frac{1}{9} \cdot ((5 - 2,7)^2 + (2 - 2,7)^2 + (3 - 2,7)^2 + (2 - 2,7)^2 + (4 - 2,7)^2 + (2 - 2,7)^2 + (4 - 2,7)^2 + (3 - 2,7)^2 + (1 - 2,7)^2 + (1 - 2,7)^2) \approx 1,33$$

Розрахунок спостережуваного значення:

$$t_{\text{сп}} = \frac{2,7 \cdot \sqrt{10}}{1,33} \approx 6,43.$$

Визначення критичного значення: для рівня значущості $\alpha = 0.05$ і ступенів свободи $k = 10 - 1 = 9$, критичне значення $t_{\text{двост.кр}} \approx 2,262$.

Прийняття рішення: Розраховане значення $t_{\text{сп}} \approx 6,43$ більше критичного значення $t_{\text{двост.кр}} \approx 2,262$. Отже, ми відхиляємо нульову гіпотезу. Є статистично значуща різниця між середніми значеннями балів до і після застосування методики.

Відповідь. Нульова гіпотеза не приймається.

18.6. F -критерій (критерій Фішера)

Критерій Фішера, також відомий як **F -тест** або **дисперсійний аналіз (ANOVA)**, використовується для порівняння дисперсій двох або більше генеральних сукупностей X та Y , за відомими значення виправлених вибірових дисперсій s_x^2 і s_y^2 для вибірок x ; y , зроблених з них. Критерій

Фішера особливо корисний при аналізі різниць між групами в експериментах.

Правило 1

Збір даних: задані дві вибірки $x = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}$; $Y = (y_1, \dots, y_m)$, $y_i \in \mathbb{R}$. Вибірки зроблено з генеральних сукупностей X і Y , які є нормальними. Позначимо через σ_X^2 і σ_Y^2 дисперсії генеральних сукупностей X і Y , а через s_x^2 і s_y^2 ($s_x^2 > s_y^2$) — вибіркові оцінки дисперсій σ_X^2 і σ_Y^2 :

$$s_x^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}_B)^2, \quad s_y^2 = \frac{1}{m-1} \cdot \sum_{i=1}^m (y_i - \bar{y}_B)^2,$$

де $\bar{x}_B = \frac{1}{n} \cdot \sum_{i=1}^n x_i$, $\bar{y}_B = \frac{1}{m} \cdot \sum_{i=1}^m y_i$ — вибіркові середні вибірок x і y .

Формулювання гіпотез:

- нульова гіпотеза $H_0: \sigma_X^2 = \sigma_Y^2$ — дисперсії двох генеральних сукупностей рівні;
- альтернативна гіпотеза $\sigma_X^2 > \sigma_Y^2$ — дисперсія генеральної сукупності менша, ніж дисперсія генеральної сукупності

Розрахунок спостережуваного значення: спостережуване значення обчислюється за формулою (18.10).

$$F_{\text{сп}} = \frac{s_x^2}{s_y^2}. \quad (18.10)$$

має розподіл з $n - 1$ і $m - 1$ степенями свободи.

Визначення критичного значення: критичне значення $F_{\text{кр}}$, визначається з таблиці розподілу Фішера (див. дод. 7 Гмурмана) для заданого рівня значущості α і ступенів свободи $k_x = n - 1$, $k_y = m - 1$.

Прийняття рішення:

- якщо розраховане значення $F_{\text{сп}} > F_{\text{кр}}$ (або знаходиться в критичній області), відхилити нульову гіпотезу;
- якщо розраховане значення $F_{\text{сп}} < F_{\text{кр}}$, не відхиляти нульову гіпотезу.

Правило 2

Збір даних: аналогічно правилу 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \sigma_X^2 = \sigma_Y^2$ — вибіркові оцінки дисперсій двох генеральних сукупностей рівні;
- альтернативна гіпотеза $\sigma_X^2 \neq \sigma_Y^2$ — дисперсія генеральної сукупності менша, ніж дисперсія генеральної сукупності

Розрахунок спостережуваного значення: спостережуване значення обчислюється за формулою (18.10).

Визначення критичного значення: критичне значення $F_{\text{кр}}$, визначається з таблиці розподілу Фішера (див. дод. 7 Гмурмана) для заданого рівня значущості $\frac{\alpha}{2}$ і ступенів свободи $k_x = n - 1$, $k_y = m - 1$.

Прийняття рішення:

- якщо розраховане значення $F_{\text{сп}} > F_{\text{кр}}$ (або знаходиться в критичній області), відхилити нульову гіпотезу;
- якщо розраховане значення $F_{\text{сп}} < F_{\text{кр}}$, не відхиляти нульову гіпотезу.

Критерій Фішера (F -критерій) є важливим інструментом для порівняння дисперсій та визначення значущих відмінностей між групами. Він широко використовується в аналізі даних та експериментальних дослідженнях для виявлення впливу різних факторів на результати.

Приклад 18.5. У студентській групі з 10 осіб проводилося тестування з двох навчальних дисциплін (табл. 18.2). Чи є відмінності в ступені однорідності показників успішності між результатами тестів? Рівень значущості $\alpha = 0,05$. Альтернативна гіпотеза $\sigma_X^2 > \sigma_Y^2$.

Розв'язання. За правилом 1 даного критерію:

Збір даних: Розглянемо дві вибірки: тест x і тест y (табл. 18.2):

Таблиця 18.2. Результати тестування

№	ПІБ	Тест x	Тест y
1	Андрєєв Л.	90	41
2	Валєєва І.	29	49
3	Енокян А.	39	56
4	Іванов П.	79	64
5	Сидорова М	88	72
6	Петров А.	53	65
7	Попова Л.	34	63
8	Шестаков І.	40	87
9	Еховий М.	75	77
10	Якушева А.	79	62

$$n = 10, m = 10, \bar{x}_B = 60,6, \bar{y}_B = 63,6.$$

Розрахувавши виправлені дисперсії для вибірок x і y , отримуємо:

$$s_x^2 = 572,83; s_y^2 = 174,04.$$

Формулювання гіпотез: $H_0: \sigma_X^2 = \sigma_Y^2$ – нульова гіпотеза;

$$\sigma_X^2 > \sigma_Y^2 \text{ – альтернативна гіпотеза.}$$

Розрахунок спостережуваного значення: За формулою (18.10) для розрахунку F -критерію знаходимо:

$$F_{\text{сп}} = \frac{572,83}{174,04} = 3,29.$$

Визначення критичного значення: По таблиці критичних значень для критерію Фішера при рівні значущості $\alpha = 0,05$ та ступенях свободи в обох випадках рівних $k_x = k_y = 10 - 1 = 9$ знаходимо $F_{\text{кр}} = 3,18$.

Прийняття рішення: Оскільки $F_{\text{сп}} > F_{\text{кр}}$, в термінах статистичних гіпотез можна стверджувати, що H_0 (гіпотеза про подібність) може бути відкинута на рівні 5%, а приймається в цьому випадку гіпотеза H_1 . Дослідник може

стверджувати, що за ступенем однорідності певного показника є відмінність між вибірками з двох класів.

Відповідь. Нульова гіпотеза не приймається.

18.7. Перевірка гіпотези про значущість вибіркового коефіцієнту кореляції

Правило 1

Збір даних: нехай із генеральної сукупності двовимірної випадкової величини (X, Y) , розподіленої за нормальним законом, здійснено вибірку об'єму n і за даними вибірки $(x; y)$ ($x = \{x_1, \dots, x_n\}; y = \{y_1, \dots, y_n\}$) обчислено вибірковий коефіцієнт кореляції Пірсона $\rho_B(x, y) = \rho_P(x, y)$, який виявився відмінним від нуля.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \rho_\Gamma(X, Y) = 0$, коефіцієнт кореляції Пірсона для двовимірної генеральної сукупності дорівнює нулю;
- альтернативна гіпотеза $H_1: \rho_\Gamma(X, Y) \neq 0$, коефіцієнт кореляції Пірсона для двовимірної генеральної сукупності не дорівнює нулю.

Розрахунок спостережуваного значення:

За критерій перевірки нульової гіпотези беруть випадкову величину (18.11):

$$t_{\text{сп}} = \rho_B(x, y) \cdot \sqrt{\frac{n-2}{1-\rho_B^2(x, y)}}. \quad (18.11)$$

Визначення критичного значення: за даними таблиці критичних точок розподілу Стюдента для двосторонньої області при заданому рівні значущості α і числі ступенів свободи $k = n - 2$ необхідно відшукати критичну точку $t_{\text{кр}}$ (див. дод. 6 Гмурмана).

Прийняття рішення:

- якщо $|t_{\text{сп}}| < t_{\text{кр}}$, то нульову гіпотезу приймають;
- якщо $|t_{\text{сп}}| > t_{\text{кр}}$, то нульову гіпотезу відхиляють.

Якщо нульова гіпотеза буде відхилена, то це означає, що вибірковий коефіцієнт кореляції суттєво відрізняється від нуля, тобто значущий, і випадкові величини X і Y – корельовані. Якщо нульова гіпотеза буде прийнята, то вибірковий коефіцієнт кореляції мало відрізняється від нуля і випадкові величини X і Y – некорельовані.

Правило 2

Збір даних: нехай із генеральної сукупності двовимірної випадкової величини (X, Y) , розподіленої за нормальним законом, здійснено вибірку об'єму n і за даними вибірки $(x; y)$ ($x = \{x_1, \dots, x_n\}; y = \{y_1, \dots, y_n\}$) обчислено вибірковий коефіцієнт рангової кореляції Спірмена $\rho_B(x, y) = \rho_S(x, y)$, який виявився відмінним від нуля.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \rho_\Gamma(X, Y) = 0$, коефіцієнт рангової кореляції Спірмена для двовимірної генеральної сукупності дорівнює нулю;

- альтернативна гіпотеза $H_1: \rho_{\Gamma}(X, Y) \neq 0$, коефіцієнт кореляції Спірмена для двовимірної генеральної сукупності не дорівнює нулю.

Розрахунок спостережуваного значення:

За критерій перевірки нульової гіпотези беруть випадкову величину (18.12):

$$t_{\text{сп}} = \rho_{\text{В}}(x, y). \quad (18.12)$$

Визначення критичного значення: за даними таблиці критичних точок розподілу Стюдента для двосторонньої області при заданому рівні значущості α і числі ступенів свободи $k = n - 2$ необхідно відшукати критичну точку $t_{\text{кр}}$ (див. дод. 6 Гмурмана).

Також необхідно розрахувати $T_{\text{кр}}$ (18.13):

$$T_{\text{кр}} = t_{\text{кр}} \cdot \sqrt{\frac{1 - \rho_{\text{В}}^2(x, y)}{n - 2}}. \quad (18.13)$$

Прийняття рішення:

- якщо $|t_{\text{сп}}| < T_{\text{кр}}$, то нульову гіпотезу приймають;
- якщо $|t_{\text{сп}}| > T_{\text{кр}}$, то нульову гіпотезу відхиляють.

Якщо нульова гіпотеза буде відхилена, то це означає, що вибіркового коефіцієнта кореляції суттєво відрізняється від нуля, тобто значущий, і випадкові величини X і Y – корельовані. Якщо нульова гіпотеза буде прийнята, то вибіркового коефіцієнта кореляції мало відрізняється від нуля і випадкові величини X і Y – некорельовані.

Правило 3

Збір даних: нехай із генеральної сукупності двовимірної випадкової величини (X, Y) , розподіленої за нормальним законом, здійснено вибірку об'єму n і за даними вибірки $(x; y)$ ($x = \{x_1, \dots, x_n\}$; $y = \{y_1, \dots, y_n\}$) обчислено вибіркового коефіцієнта рангової кореляції Кендалла $\rho_{\text{В}}(x, y) = \rho_{\text{К}}(x, y)$, який виявився відмінним від нуля.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \rho_{\Gamma}(X, Y) = 0$, коефіцієнт рангової кореляції Кендалла для двовимірної генеральної сукупності дорівнює нулю;
- альтернативна гіпотеза $H_1: \rho_{\Gamma}(X, Y) \neq 0$, коефіцієнт рангової кореляції Кендалла для двовимірної генеральної сукупності не дорівнює нулю.

Розрахунок спостережуваного значення:

За критерій перевірки нульової гіпотези беруть випадкову величину (18.14):

$$t_{\text{сп}} = \rho_{\text{В}}(x, y). \quad (18.14)$$

Визначення критичного значення: за даними таблиці критичних точок розподілу Стюдента для двосторонньої області при заданому рівні значущості α необхідно відшукати критичну точку $z_{\text{кр}}$ з рівняння $\Phi(z_{\text{кр}}) = \frac{1 - \alpha}{2}$ (див. дод. 2 Гмурмана).

Також необхідно розрахувати $T_{кр}$ (18.15):

$$T_{кр} = z_{кр} \cdot \sqrt{\frac{2 \cdot (2 \cdot n + 5)}{9 \cdot n \cdot (n - 1)}}. \quad (18.15)$$

Прийняття рішення:

- якщо $|t_{сп}| < T_{кр}$, то нульову гіпотезу приймають;
- якщо $|t_{сп}| > T_{кр}$, то нульову гіпотезу відхиляють.

Якщо нульова гіпотеза буде відхилена, то це означає, що вибіркового коефіцієнта кореляції суттєво відрізняється від нуля, тобто значущий, і випадкові величини X і Y – корельовані. Якщо нульова гіпотеза буде прийнята, то вибіркового коефіцієнта кореляції мало відрізняється від нуля і випадкові величини X і Y – некорельовані.

Приклад 18.6. За допомогою коефіцієнта кореляції Пірсона оцінити значущість кореляційного зв'язку між успішністю виконання тестових завдань з фізики (X) і математики (Y) учнями загальноосвітньої школи. Фрагмент результатів тестування наведено в табл. 18.3. Рівень значущості $\alpha = 0,01$.

Розв'язання. Скористаємося правилом 1 про значущість коефіцієнту кореляції Пірсона.

Збір даних: Розглянемо дві вибірки: тест x і тест y (табл. 18.3).

Таблиця 18.3. Фрагмент результатів тестування

№	ПІБ	Тест x	Тест y
1	Жураковська Л.	22	18
2	Радченко І.	25	22
3	Коган А.	18	20
4	Онищенко П.	24	25
5	Митник М.	14	15
6	Андрієць А.	20	18
7	Лісовенко Л.	16	20
8	Мельник І.	15	13
9	Нечитайло М.	24	18
10	Чирва А.	20	15
11	Гамкало С.	11	10
12	Борисюк О.	11	13

Між вибірками x та y коефіцієнт кореляції $\rho_B(x, y) = 0,78$, а $n = 12$.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \rho_T(X, Y) = 0$;
- альтернативна гіпотеза $H_1: \rho_T(X, Y) \neq 0$.

Розрахунок спостережуваного значення: за формулою (18.11)

$$t_{сп} = 0,78 \cdot \sqrt{\frac{10}{1 - 0,78^2}} \approx 3,94.$$

Визначення критичного значення: $t_{кр} = 3,17$ при рівні значущості $\alpha = 0,01$ та степені свободи $k = 10$.

Прийняття рішення: $|t_{сп}| > t_{кр}$, отже нульова гіпотеза відхиляється. Тобто, між результатами тесту X і тесту Y значущий існує кореляційний зв'язок.

Інтерпретація результатів наведені на рис. 18.1.

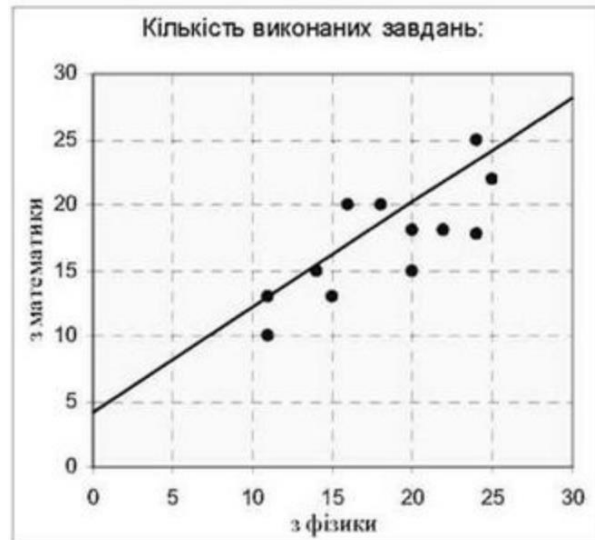


Рис 18.1. Інтерпретація результатів

Відповідь. Нульова гіпотеза не приймається.

Приклади практичного застосування даних критеріїв наведено в лекції 13. В якості генеральних сукупностей в них розглядалися статистика замовлення множини електронних послуг, а також курси обраної криптовалюти. Вибірками були відповідно розглядалися статистика замовлення множини електронних послуг, а також курси обраної криптовалюти за 30 та за 10 днів відповідно.

18.8. Порівняння виправленої вибіркової дисперсії з гіпотетичною генеральною дисперсією нормальної сукупності

Нехай генеральна сукупність X розподілена нормально, причому генеральна дисперсія хоч і невідома, але є підстави припускати, що вона дорівнює гіпотетичному (передбачуваному) значенню σ_0^2 . Насправді σ_0^2 встановлюється виходячи з попереднього досвіду чи теоретично.

Правило 1

Збір даних: нехай із генеральної сукупності X вилучено вибірку $x = \{x_1, \dots, x_n\}$ об'єму n і для неї знайдено виправлену вибірку дисперсію s^2 з $k = n - 1$ ступенями свободи. Потрібно по виправленій дисперсії при заданому рівні значущості перевірити нульову гіпотезу, яка полягає в тому, що генеральна дисперсія аналізованої сукупності дорівнює гіпотетичному значенню σ_0^2 .

Формулювання гіпотез:

- нульова гіпотеза $H_0 : \sigma^2 = \sigma_0^2$ – дисперсія генеральної сукупності дорівнює гіпотетичному значенню;

- альтернативна гіпотеза $H_1: \sigma^2 > \sigma_0^2$ – дисперсія генеральної сукупності більша за деяке гіпотетичне значення.

Розрахунок спостережуваного значення: як критерій перевірки нульової гіпотези приймемо випадкову величину (18.16):

$$\chi_{\text{сп}}^2 = \frac{(n-1) \cdot s^2}{\sigma_0^2}. \quad (18.16)$$

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 за заданим рівнем значущості α і числом ступенів свободи $k = n - 1$ знайти критичну точку $\chi_{\text{кр}}^2$ (див. дод. 5 Гмурмана).

Прийняття рішення:

- якщо $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 > \chi_{\text{кр}}^2$, то нульову гіпотезу відхиляють.

Правило 2

Збір даних: аналогічно правилу 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \sigma^2 = \sigma_0^2$ – дисперсія генеральної сукупності дорівнює гіпотетичному значенню;
- альтернативна гіпотеза $H_1: \sigma^2 \neq \sigma_0^2$ – дисперсія генеральної сукупності не дорівнює гіпотетичному значенню.

Розрахунок спостережуваного значення: критерій перевірки нульової гіпотези обчислюється за формулою (18.16).

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{\text{лів.кр}}^2$ за заданим рівнем значущості $1 - \frac{\alpha}{2}$ і числом ступенів свободи $k = n - 1$ (див. дод. 5 Гмурмана) та $\chi_{\text{прав.кр}}^2$ за заданим рівнем значущості $\frac{\alpha}{2}$ і числом ступенів свободи $k = n - 1$.

Прийняття рішення:

- якщо $\chi_{\text{лів.кр}}^2 < \chi_{\text{сп}}^2 < \chi_{\text{прав.кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 < \chi_{\text{лів.кр}}^2$, або $\chi_{\text{сп}}^2 > \chi_{\text{прав.кр}}^2$, то нульову гіпотезу відхиляють.

Правило 3

Збір даних: аналогічно правилу 1.

Формулювання гіпотез:

- нульова гіпотеза $H_0: \sigma^2 = \sigma_0^2$ – дисперсія генеральної сукупності дорівнює гіпотетичному значенню;
- альтернативна гіпотеза $H_1: \sigma^2 < \sigma_0^2$ – дисперсія генеральної сукупності менша за деяке гіпотетичне значення.

Розрахунок спостережуваного значення: критерій перевірки нульової гіпотези обчислюється за формулою (18.16).

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{\text{кр}}^2$ за заданим рівнем значущості $1 - \alpha$ і числом ступенів свободи $k = n - 1$ (див. дод. 5 Гмурмана).

Прийняття рішення:

- якщо $\chi_{\text{сп}}^2 > \chi_{\text{кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$, то нульову гіпотезу відхиляють.

Приклад 18.7. З нормальної генеральної сукупності вилучено вибірку обсягу $n = 13$ і по ній знайдено виправлену вибірккову дисперсію $s^2 = 14,6$. Потрібно при рівні значущості $\alpha = 0,01$ перевірити нульову гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 12$, прийнявши в якості конкуруючої гіпотези $H_1: \sigma^2 > 12$.

Розв'язання. Скористаємося правилом 1 даного критерію.

Збір даних: нормальна генеральна сукупність, $n = 13$, $s^2 = 14,6$, $\alpha = 0,01$.

Формулювання гіпотез: $H_0: \sigma^2 = \sigma_0^2 = 12$, $H_1: \sigma^2 > 12$.

Розрахунок спостережуваного значення: знайдемо спостережуване значення критерію:

$$\chi_{\text{сп}}^2 = \frac{(n-1) \cdot s^2}{\sigma_0^2} = \frac{(13-1) \cdot 14,6}{12} = 14,6.$$

Визначення критичного значення: за рівнем значущості $\alpha = 0,01$ і числом ступенів свободи $k = n - 1 = 13 - 1 = 12$ знаходимо критичну точку $\chi_{\text{кр}}^2 = 26,2$.

Прийняття рішення: оскільки $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$ — немає підстав відкинути нульову гіпотезу. Іншими словами, різниця між виправленою дисперсією та гіпотетичною генеральною дисперсією — незначна.

Відповідь. Нульова гіпотеза приймається.

ПЕРЕВІРКА СТАТИСТИЧНОЇ ГІПОТЕЗИ ЩОДО ТИПУ НЕВІДОМОГО РОЗПОДІЛУ

Критерієм згоди називають критерій перевірки гіпотези про вигляд невідомого закону розподілу.

Є ряд критеріїв згоди: χ^2 -Пірсона, Колмогорова, Смірнова та інші.

18.9. Критерій χ^2 -Пірсона

Критерій χ^2 -Пірсона (хі-квадрат Пірсона) використовується для перевірки гіпотез про розподіл випадкових величин. Цей критерій дозволяє визначити, чи є статистично значущі відмінності між очікуваними та спостережуваними частотами в одній або кількох категоріях.

Перевірка гіпотези про нормальний розподіл генеральної сукупності

Правило

Збір даних: нехай за вибіркою об'єму n отримано інтервальний статистичний ряд (табл. 18.4) із частковими інтервалами однакової довжини

$$\sum_{i=1}^l n_i = n.$$

Таблиця 18.4. Інтервальний статистичний ряд

$[x_i; x_{i+1})$	$[x_1; x_2)$	$[x_2; x_3)$	...	$[x_l; x_{l+1})$
n_i	n_1	n_2	...	n_l

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за нормальним законом;

- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за нормальним законом.

Розрахунок спостережуваного значення: критерієм перевірки цієї гіпотези беруть випадкову величину $\chi_{\text{сп}}^2$ (18.17):

$$\chi_{\text{сп}}^2 = \sum_{i=1}^l \frac{(n_i - n'_i)^2}{n'_i}, \quad (18.17)$$

де:

m – число часткових інтервалів в статистичному ряді;

n_i ($i = 1, \dots, l$) – їх *частоти (емпіричні частоти)*;

n'_i – *теоретичні частоти*, обчислені за припущенням, що генеральна сукупність розподілена нормально (число появ в n випробуваннях значень нормально розподіленої випадкової величини X із i -го часткового інтервалу).

Щоб знайти теоретичні частоти нормального розподілу, необхідно:

1. Знайти середини x_i^* ($i = 1, \dots, l$) часткових інтервалів даного статистичного ряду (усереднені варіанти);

2. Обчислити вибірку середню $\bar{x}_B$ та вибіркве середнє квадратичне відхилення σ^* за даним розподілом вибірки, беручи в якості варіант середини часткових інтервалів (18.18), (18.19):

$$\bar{x}_B^* = \frac{1}{n} \cdot \sum_{i=1}^l (n_i \cdot x_i^*), \quad (18.18)$$

$$\sigma_B^* = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^l (n_i \cdot (x_i^*)^2) - \left(\frac{1}{n} \cdot \sum_{i=1}^l (n_i \cdot x_i^*) \right)^2}; \quad (18.19)$$

3. Нормалізувати випадкову величину X , тобто перейти до величини $Z = \frac{X - \bar{x}_B^*}{\sigma_B^*}$ і обчислити кінці інтервалів (z_i, z_{i+1}) (18.20):

$$z_i = \frac{x_i - \bar{x}_B^*}{\sigma_B^*}, z_{i+1} = \frac{x_{i+1} - \bar{x}_B^*}{\sigma_B^*}, \quad (18.20)$$

покладаючи при цьому $z_i = -\infty$, $z_{i+1} = +\infty$;

4. Обчислити теоретичні ймовірності $P\{Z \in (z_i, z_{i+1})\} = P_i$ потрапляння величини Z в інтервал (z_i, z_{i+1}) (18.21):

$$P_i = \Phi(z_{i+1}) - \Phi(z_i); \quad (18.21)$$

5. Знайти теоретичні частоти (18.22):

$$n'_i = n \cdot P_i. \quad (18.22)$$

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{\text{кр}}^2$ (див. дод. 5 Гмурмана) за заданим рівнем значущості α і числом ступенів свободи (18.23):

$$k = l - r - 1 = l - 3, \quad (18.23)$$

де:

l – кількість інтервалів;

r – кількість параметрів, які описують відповідний тип розподілу (для нормального розподілу та рівномірного розподілів $r = 2$).

Прийняття рішення:

- якщо $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 > \chi_{\text{кр}}^2$, то нульову гіпотезу відхиляють.

Приклад 18.8. За даним інтервальним розподілом вибірки об'єму $n = 100$ при рівні значущості $\alpha = 0,05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл генеральної сукупності (табл.18.5).

Таблиця 18.5. Інтервальний варіативний ряд

$[x_i, x_{i+1})$	[3; 8)	[8; 13)	[13; 18)	[18; 23)	[23; 28)	[28; 33)	[33; 38)
n_i	6	8	15	40	16	8	7

Розв'язання. Скористаємося критерієм χ^2 -Пірсона.

Збір даних: Побудуємо статистичний розподіл, варіантами якого є середини даних інтервалів (табл. 18.6):

Таблиця 18.6. Допоміжний варіативний ряд

x_i^*	5,5	10,5	15,5	20,5	25,5	30,5	35,5
n_i	6	8	15	40	16	8	7

За формулами (18.18), (18.19) обчислимо вибірккову середню та вибірккове середнє квадратичне відхилення: $\bar{x}_B^* \approx 20,7$; $\sigma_B^* \approx 7,28$.

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за нормальним законом;
- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за нормальним законом.

Розрахунок спостережуваного значення: використовуючи формули (18.21), (8.22), обчислимо теоретичні частоти і n'_i . Для цього складемо розрахункову таблицю (табл. 18.7).

Таблиця 18.7. Розрахункова таблиця (частина 1)

i	x_i	x_{i+1}	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	P_i	n'_i
1	3	8	$-\infty$	-1,74	-0,5000	-0,4591	0,0409	4
2	8	13	-1,74	-1,06	-0,4591	-0,3554	0,1037	10
3	13	18	-1,06	-0,37	-0,3554	-0,1443	0,2111	21
4	18	23	-0,37	0,32	-0,1443	0,1255	0,2698	27
5	23	28	0,32	1,00	0,1255	0,3413	0,2158	22
6	28	33	1,00	1,69	0,3413	0,4545	0,1132	11
7	33	38	1,69	$+\infty$	0,4545	0,5000	0,0455	5
$\sum$							1	100

Обчислимо спостережуване значення критерію $\chi_{\text{сп}}^2$, для чого складемо розрахункову таблицю (табл. 18.8).

Таблиця 18.8. Розрахункова таблиця (частина 2)

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	6	4	2	4	1,0000
2	8	10	-2	4	0,4000
3	15	21	-6	36	1,7143
4	40	27	13	169	6,2593
5	16	22	-6	36	1,6364
6	8	11	-3	9	0,8181
7	7	5	2	4	0,4000
$\sum$	100	100			12,2281

Отже, $\chi^2_{\text{сп}} \approx 12,2281$.

Визначення критичного значення: за даними таблиці критичних точок розподілу χ^2 при рівні значущості $\alpha = 0,05$ та числі ступенів вільності $k = 4$ знаходимо критичну точку $\chi^2_{\text{кр}} = 9,5$.

Прийняття рішення: оскільки $\chi^2_{\text{сп}} > \chi^2_{\text{кр}}$, то гіпотезу про нормальний розподіл генеральної сукупності X відкидаємо. Це означає, що емпіричні та теоретичні частоти відрізняються істотно, тобто дані спостережень не узгоджуються з гіпотезою про нормальний розподіл генеральної сукупності.

Відповідь. Нульова гіпотеза не приймається.

Слід зазначити, що критерій Пірсона використовується також для перевірки гіпотез про інші типи розподілів генеральної сукупності (біноміальний, експоненційний тощо). Для цього в формулі (18.21) P_i шукається в залежності від типу розподілу за формулою (18.24):

$$P_i = F(z_{i+1}) - F(z_i). \quad (18.24)$$

Дана формула отримується з формули (6.6) лекції 6, де $\alpha = z_i$; $\beta = z_{i+1}$.

Перевірка гіпотези про показниковий розподіл генеральної сукупності

Правило

Збір даних: нехай за вибіркою об'єму n отримано інтервальний статистичний ряд (табл. 18.4) із частковими інтервалами однакової довжини.

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за показниковим законом;
- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за показниковим законом.

Розрахунок спостережуваного значення: критерієм перевірки цієї гіпотези беруть випадкову величину $\chi^2_{\text{сп}}$ (18.17).

Щоб знайти теоретичні частоти показникового розподілу, необхідно:

1. Знайти середини x_i^* ($i = 1, \dots, l$) часткових інтервалів даного статистичного ряду (усереднені варіанти);
2. За формулою (18.18) обчислити вибірку середню $\bar{x}_B^*$ за даним розподілом вибірки, беручи в якості варіант середини часткових інтервалів;
3. Застосовувати в якості оцінки параметра λ показникового розподілу величину (18.25):

$$\lambda^* = \frac{1}{x_B^*}; \quad (18.25)$$

4. Обчислити теоретичні ймовірності $P\{X \in (x_i, x_{i+1})\} = P_i$ потрапляння величини X в інтервал (x_i, x_{i+1}) (18.26):

$$P_i = e^{-x_i} - e^{-x_{i+1}}; \quad (18.26)$$

5. Знайти теоретичні частоти за формулою (18.22).

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{кр}^2$ (див. дод. 5 Гмурмана) за заданим рівнем значущості α і числом ступенів свободи (18.27):

$$k = l - r - 1 = l - 2, \quad (18.27)$$

де:

l – кількість інтервалів;

r – кількість параметрів, які описують відповідний тип розподілу (для показникового розподілу та розподілу Пуассона $r = 1$).

Прийняття рішення:

- якщо $\chi_{сп}^2 < \chi_{кр}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{сп}^2 > \chi_{кр}^2$, то нульову гіпотезу відхиляють.

Перевірка гіпотези про біноміальний розподіл генеральної сукупності

Правило

Збір даних: нехай проведено n серій дослідів. Кожна серія складається з m експериментів, в кожному з яких ймовірність появи деякої події A є сталою і дорівнює p . Емпіричний розподіл дискретної випадкової величини $X = \{\text{кількості появ події } A \text{ в } m \text{ експериментах}\}$ задано варіативним рядом (табл. 18.9).

Таблиця 18.9. Варіативний ряд

x_i	0	1	...	m
n_i	n_0	n_1	...	n_m

В табл. 18.9:

n_i – частота появи події A рівно i разів, ($i = 0, 1, \dots, m$);

$$\sum_{i=0}^m n_i = n;$$

m – максимальна кількість появ події A в одному досліді.

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за біноміальним законом;

- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за біноміальним законом.

Розрахунок спостережуваного значення: критерієм перевірки цієї гіпотези беруть випадкову величину $\chi^2_{\text{сп}}$ (18.17).

Щоб знайти теоретичні частоти біноміального розподілу, необхідно:

1. Обчислити за формулою Бернуллі ймовірності P_i настання події A рівно i разів ($i = 0, 1, \dots, m$) (18.28):

$$P_i = C_m^i \cdot p^i \cdot q^{m-i}; \quad (18.28)$$

2. Знайти теоретичні частоти за формулою (18.22).

Зауваження

У випадку, коли параметр p невідомий, обчислюємо його точкову оцінку (див. лекцію 17), за формулою (18.29)

$$p \approx \frac{\bar{x}_B}{m}, \quad (18.29)$$

де:

$\bar{x}_B$ – вибіркове середнє вибірки (табл. 18.9);

m – кількість експериментів в серії.

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi^2_{\text{кр}}$ (див. дод. 5 Гмурмана) за заданим рівнем значущості α і числом ступенів свободи (18.30), за умови, що значення p – невідоме:

$$k = m - r - 1 = m - 2, \quad (18.30)$$

де:

m – максимальна кількість появ події A а одному досліді;

r – кількість параметрів, які описують відповідний тип розподілу (для біноміального розподілу з невідомою ймовірністю p появ події $r = 1$).

Якщо значення p – задане, то справедлива формула (18.31):

$$k = m - r - 1 = m - 1, \quad (18.31)$$

де:

m – максимальна кількість появ події A а одному досліді;

r – кількість параметрів, які описують відповідний тип розподілу (для біноміального розподілу з відомою ймовірністю p появ події A $r = 0$).

Прийняття рішення:

- якщо $\chi^2_{\text{сп}} < \chi^2_{\text{кр}}$, то нульову гіпотезу приймають;
- якщо $\chi^2_{\text{сп}} > \chi^2_{\text{кр}}$, то нульову гіпотезу відхиляють.

Перевірка гіпотези про рівномірний розподіл генеральної сукупності

Правило

Збір даних: нехай за вибіркою об'єму n отримано інтервальний статистичний ряд (табл. 18.4) із частковими інтервалами однакової довжини.

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за рівномірним законом;
- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за рівномірним законом.

Розрахунок спостережуваного значення: критерієм перевірки цієї гіпотези беруть випадкову величину $\chi_{\text{сп}}^2$ (18.17).

Щоб знайти теоретичні частоти рівномірного розподілу, необхідно:

1. Знайти середини x_i^* ($i = 1, \dots, l$) часткових інтервалів даного статистичного ряду (усереднені варіанти);
2. Обчислити вибірку середню $\bar{x}_B^*$ та вибіркове середнє квадратичне відхилення σ^* за даним розподілом вибірки, беручи в якості варіант середини часткових інтервалів (18.18), (18.19);
3. Оцінити параметри a^*, b^* рівномірного розподілу за формулою (18.32):

$$a^* = \bar{x}_B^* - \sigma^* \cdot \sqrt{3}; \quad b^* = \bar{x}_B^* + \sigma^* \cdot \sqrt{3}; \quad (18.32)$$

4. Записати функцію щільності рівномірного розподілу (18.33):

$$f(x) = \begin{cases} \frac{1}{b^* - a^*}, & \text{якщо } x \in [a^*, b^*]; \\ 0, & \text{якщо } x \notin [a^*, b^*]; \end{cases} \quad (18.33)$$

5. Обчислити теоретичні ймовірності $P\{X \in (x_i, x_{i+1})\} = P_i$ потрапляння величини X в інтервал (x_i, x_{i+1}) (18.34):

$$P_1 = \frac{1}{b^* - a^*} \cdot (x_1 - a^*), \quad P_i = \frac{1}{b^* - a^*} \cdot (x_{i+1} - x_i), \quad P_l = \frac{1}{b^* - a^*} \cdot (b^* - x_{l-1}), \quad (18.34)$$

де $i = 2, \dots, l - 1$.

6. Знайти теоретичні частоти за формулою (18.22).

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{\text{кр}}^2$ (див. дод. 5 Гмурмана) за заданим рівнем значущості α і числом ступенів свободи (18.23).

Прийняття рішення:

- якщо $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 > \chi_{\text{кр}}^2$, то нульову гіпотезу відхиляють.

Перевірка гіпотези про розподіл Пуассона генеральної сукупності

Правило

Збір даних: нехай проведено серію з n дослідів. Кожна серія складається з m експериментів, в кожному з яких ймовірність появи деякої події A є сталою. Емпіричний розподіл дискретної випадкової величини $X = \{\text{кількості появ події } A \text{ в } m \text{ експериментах}\}$ задано варіативним рядом (табл. 18.9).

Формулювання гіпотез: потрібно при рівні значущості α перевірити нульову гіпотезу:

- нульова гіпотеза H_0 : генеральна сукупність розподілена за законом Пуассона

- альтернативна гіпотеза H_1 : генеральна сукупність не розподілена за законом Пуассона.

Розрахунок спостережуваного значення: критерієм перевірки цієї гіпотези беруть випадкову величину $\chi_{\text{сп}}^2$ (18.17).

Щоб знайти теоретичні частоти розподілу Пуассона, необхідно:

1. За формулою (18.18) обчислити вибірку середню $\bar{x}_B$ за даним розподілом вибірки, беручи в якості варіант середини часткових інтервалів.
2. Застосовувати в якості оцінки параметра λ показникового розподілу величину (18.35):

$$\lambda^* = \bar{x}_B; \quad (18.35)$$

3. Обчислити теоретичні ймовірності $P\{X \in (x_i, x_{i+1})\} = P_i$ потрапляння величини X в інтервал (x_i, x_{i+1}) за формулою Пуассона (18.36):

$$P_i = P_n(i) = \frac{\lambda^i}{i!} \cdot e^{-\lambda}; \quad (18.36)$$

4. Знайти теоретичні частоти за формулою (18.22).

Визначення критичного значення: за таблицею критичних точок розподілу χ^2 знайти критичну точку $\chi_{\text{кр}}^2$ (див. дод. 5 Гмурмана) за заданим рівнем значущості α і числом ступенів свободи (18.27).

Прийняття рішення

- якщо $\chi_{\text{сп}}^2 < \chi_{\text{кр}}^2$, то нульову гіпотезу приймають;
- якщо $\chi_{\text{сп}}^2 > \chi_{\text{кр}}^2$, то нульову гіпотезу відхиляють.

18.10. Критерій узгодження Колмогорова

Критерій Колмогорова (також відомий, як *критерій узгодження Колмогорова-Смірнова*) є одним з основних і найбільш широко використовуваних непараметричних методів у силу своєї достатньої чутливості до розходжень у досліджуваних розподілах.

Критерій Колмогорова призначений для зіставлення:

- емпіричного розподілу з теоретичним розподілом;
- одного емпіричного розподілу з іншим.

Критерій дозволяє знайти точку, у якій сума накопичених розбіжностей між двома розподілами є найбільшою, і оцінити ймовірність цієї розбіжності.

Обмеження для коректного застосування критерію:

1. Вибірка досить велика (більше 50 спостережень);
2. Класи інтервалів повинні бути впорядковані по зростанню або убутанню деякої ознаки. Вони обов'язково повинні відображати її направлену зміну.

Правило 1

Збір даних: нехай за вибіркою об'єму n отримано інтервальний статистичний ряд (табл. 16.3 лекції 16) із частковими інтервалами однакової довжини.

Формулювання гіпотез:

- нульова гіпотеза H_0 : вибірка походить з заданого теоретичного розподілу;

- альтернативна гіпотеза H_1 : вибірка не походить з заданого теоретичного розподілу.

Обчислення емпіричної функції розподілу: побудова емпіричної функції розподілу $F_n(x)$ на основі вибірових даних.

Обчислення теоретичної функції розподілу: побудова теоретичної функції розподілу $F(x)$ для заданого типу розподілу.

Обчислення статистики Колмогорова-Смирнова: знаходження максимальної різниці між емпіричною та теоретичною функціями розподілу (18.37):

$$D_{\text{сп}} = \max |F_n(x) - F(x)|. \quad (18.37)$$

Визначення критичного значення: для обчислення критичного значення $D_{\text{кр}}$ використовується таблиця критичних значень для критерію Колмогорова-Смирнова (табл. 18.10) або відповідна функція в програмному забезпеченні для статистичного аналізу.

Таблиця 18.10. Таблиця критичних значень для критерію Колмогорова-Смирнова

n	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
[1; 10]	0,409	0,455	0,531
[11; 20]	0,271	0,294	0,334
[21; 30]	0,222	0,240	0,271
[31; 40]	0,187	0,202	0,288
[41; 50]	0,162	0,175	0,198
[51; 60]	0,144	0,155	0,175
[61; 70]	0,131	0,141	0,159
[71; 80]	0,120	0,129	0,146
[81; 90]	0,112	0,121	0,137
[91; 100]	0,105	0,114	0,127

Табл. 18.10 отримано засобами Python (рис. 18.2).

```
python
Копіювати код

import numpy as np
from scipy.stats import kstest

# Генерація вибірки з нормального розподілу
np.random.seed(0)
data = np.random.normal(loc=0, scale=1, size=100)

# Застосування критерію Колмогорова-Смирнова для перевірки нормальності
statistic, p_value = kstest(data, 'norm')

print(f"Статистика Колмогорова-Смирнова: {statistic}")
print(f"p-значення: {p_value}")

# Перевірка гіпотези при рівні значущості 0.05
alpha = 0.05
if p_value < alpha:
    print("Нульова гіпотеза відхиляється: вибірка не походить з нормального розподілу")
else:
    print("Немає підстав відхилити нульову гіпотезу: вибірка може походити з нормального розподілу")
```

Рис. 18.2. Приклад застосування в Python

Для $n > 100$ використовуються такі асимптотичні формули:

$$\text{Для } \alpha = 0,10: D_{\text{кр}} = \frac{1,22}{\sqrt{n}};$$

$$\text{Для } \alpha = 0,05: D_{\text{кр}} = \frac{1,36}{\sqrt{n}};$$

$$\text{Для } \alpha = 0,01: D_{\text{кр}} = \frac{1,63}{\sqrt{n}}.$$

Прийняття рішення:

- якщо $D_{\text{сп}} > D_{\text{кр}}$, то нульова гіпотеза відхиляється;
- якщо $D_{\text{сп}} < D_{\text{кр}}$, то нульова гіпотеза не відхиляється.

Правило 2

Збір даних: аналогічно до правила 1

Формулювання гіпотез:

- нульова гіпотеза H_0 : вибірка походить з заданого теоретичного розподілу;
- альтернативна гіпотеза H_1 : вибірка не походить з заданого теоретичного розподілу.

Обчислення емпіричної функції розподілу: побудова емпіричної функції розподілу $F_n(x)$ на основі вибірових даних.

Обчислення теоретичної функції розподілу: побудова теоретичної функції розподілу $F(x)$ для заданого типу розподілу.

Обчислення статистики Колмогорова-Смирнова: знаходження максимальної різниці між емпіричною та теоретичною функціями розподілу (18.38):

$$\lambda_{\text{сп}} = \sqrt{n} \cdot \max |F_n(x) - F(x)|. \quad (18.38)$$

Визначення критичного значення: для обчислення критичного значення $\lambda_{\text{кр}}$ використовується таблиця критичних значень для критерію Колмогорова-Смирнова (табл. 18.11) або відповідна функція в програмному забезпеченні для статистичного аналізу.

Таблиця 18.11. Таблиця критичних значень для критерію Колмогорова-Смирнова

Рівень значущості α	0,40	0,30	0,20	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
Критичні значення $\lambda_{\text{кр}}$	0,89	0,97	1,07	1,22	1,36	1,48	1,63	1,77	1,95	2,03

В табл. 18.11 значення $\lambda_{\text{кр}}$ отримано з рівняння (18.39):

$$P(\lambda_{\text{кр}}) = \alpha, \quad (18.39)$$

де $P(\lambda_{\text{кр}})$ обчислюється з рівності (18.40):

$$P(\lambda_{\text{кр}}) = 1 - \sum_{k=-\infty}^{+\infty} ((-1)^k \cdot e^{-(2 \cdot k^2 \cdot \lambda_{\text{кр}}^2)}). \quad (18.40)$$

Це пов'язано із тим, що практично доведено, що якою б не була теоретична функція розподілу $F(x)$ неперервної випадкової величини X при необмеженій кількості спостережень ($n \rightarrow +\infty$) ймовірність нерівності $P(\lambda_{\text{сп}} \geq \lambda)$ прямує до границі (18.40).

Прийняття рішення:

- якщо $\lambda_{\text{сп}} > \lambda_{\text{кр}}$, то нульова гіпотеза відхиляється;
- якщо $\lambda_{\text{сп}} < \lambda_{\text{кр}}$, то нульова гіпотеза не відхиляється.

Критерій Колмогорова простіше критерію χ^2 -Пірсона, тому його охоче застосовують на практиці. Однак, цей критерій можна застосовувати тільки у випадку, коли гіпотетичний розподіл $F(x)$ разом з усіма параметрами, що в нього входять. Такий випадок порівняно нечасто зустрічається на практиці. Звичайно з теоретичних міркувань відомо тільки загальний вид функції $F(x)$. При застосуванні критерію Пірсона ця обставина враховується відповідним зменшенням числа степенів вільності розподілу χ^2 . Критерій Колмогорова такого узгодження не передбачає.

Приклад 18.9. На основі вибірки (табл. 18.12)

Таблиця 18.12. Інтервальний варіативний ряд

$[x_i, x_{i+1})$	$[0; 5)$	$[5; 10)$	$[10; 15)$	$[15; 20)$	$[20; 25)$	$[25; 30)$	$[30; 35)$
n_i	30	16	7	3	2	1	1

побудовано експоненціальний закон розподілу часу обслуговування 60 заявок на АТС, що визначається функцією розподілу $F(x) = \begin{cases} 1 - e^{-(0,14 \cdot x)}, & x \geq 0; \\ 0, & x < 0. \end{cases}$

Перевірити узгодженість теоретичних і емпіричних частот за критерієм Колмогорова. Рівень значущості $\alpha = 0,05$.

Розв'язання. За критерієм Колмагорова:

Збір даних: за вибіркою об'єму 60 отримано інтервальний статистичний ряд (табл. 18.11) із частковими інтервалами однакової довжини.

Формулювання гіпотез:

- нульова гіпотеза H_0 : вибірка узгоджується з заданим експоненційним розподілом;
- альтернативна гіпотеза H_1 : вибірка не узгоджується з заданим експоненційним розподілом.

Обчислення емпіричної функції розподілу: по заданому статистичному розподілу побудуємо емпіричну функцію розподілу $F_n(x) = \frac{n_x}{n}$, де n_x – кількість значень ознаки, які не перевищують величину X .

$$F_n(x) = 0, \text{ при } -\infty < x < 0;$$

$$F_n(x) = \frac{30}{60} = 0,5, \text{ при } 0 \leq x < 5;$$

$$F_n(x) = \frac{46}{60} = 0,77, \text{ при } 5 \leq x < 10;$$

$$F_n(x) = \frac{53}{60} = 0,88, \text{ при } 10 \leq x < 15;$$

$$F_n(x) = \frac{56}{60} = 0,93, \text{ при } 15 \leq x < 20;$$

$$F_n(x) = \frac{58}{60} = 0,97, \text{ при } 20 \leq x < 25;$$

$$F_n(x) = \frac{59}{60} = 0,99, \text{ при } 25 \leq x < 30;$$

$$F_n(x) = \frac{60}{60} = 1, \text{ при } 30 \leq x < 35;$$

$$F_n(x) = \frac{60}{60} = 1, \text{ при } 35 \leq x < +\infty.$$

Очевидно, що два останні інтервали можна об'єднати.

Обчислення теоретичної функції розподілу: значення теоретичної функції розподілу для кожного з інтервалів обчислюємо за формулою:

$$F(x) = \begin{cases} 1 - e^{-(0,14 \cdot x)}, & x \geq 0; \\ 0, & x < 0 \end{cases} \text{ для правого кінця проміжку.}$$

Наприклад,

$$F(x) = 1 - e^{-(0,14 \cdot 5)} = 0,503, \text{ при } 0 < x \leq 5;$$

$$F(x) = 1 - e^{-(0,14 \cdot 10)} = 0,753, \text{ при } 5 < x \leq 10.$$

Обчислення статистики Колмогорова-Смирнова: Обчислення проведемо в табл. 18.13.

Таблиця 18.13. Розрахунок величини $D_{\text{сп}}$

$[x_i, x_{i+1})$	n_i	$F_n(x)$	$F(x)$	$ F_n(x) - F(x) $
$(-\infty, 0)$	0	0	0	0
$[0; 5)$	30	0,50	0,503	0,003
$[5; 10)$	16	0,77	0,753	0,017
$[10; 15)$	7	0,88	0,878	0,002
$[15; 20)$	3	0,93	0,939	0,009
$[20; 25)$	2	0,97	0,969	0,001
$[25; 30)$	1	0,98	0,985	0,005
$[30; 35)$	1	1	0,993	0,007
$[35; +\infty)$	0	1	1	0

Найбільше значення в останньому стовпці і є спостережуваним значенням. Тобто $D_{\text{сп}} = 0,017$ ($\lambda_{\text{сп}} = 0,017 \cdot \sqrt{60} = 0,132$).

Визначення критичного значення: згідно таблиці критичних значень (табл. 18.10) при рівні значущості $\alpha = 0,05$ та $n = 60$ критичне значення $D_{\text{кр}} = 0,155$, а згідно табл. 18.11 $\lambda_{\text{кр}} = 1,36$.

Прийняття рішення: оскільки $D_{\text{сп}} < D_{\text{кр}}$ ($\lambda_{\text{сп}} < \lambda_{\text{кр}}$), то за критерієм Колмогорова гіпотезу про розподіл часу обслуговування заявок на АТС за експоненціальним законом варто визнати не суперечною експерименту.

Відповідь. Нульова гіпотеза приймається.

18.11. Метод спрямлюючих діаграм

Для перевірки гіпотези про нормальність розподілу генеральної сукупності за допомогою *методу спрямлюючих діаграм*, скористаємося $Q - Q$ діаграмою (*quantile-quantile plot*).

$Q - Q$ діаграма дозволяє оцінити, наскільки розподіл даних відповідає нормальному розподілу, порівнюючи квантілі спостережуваних даних з квантілями теоретичного нормального розподілу.

Правило

Побудова розрахункової таблиці. Побудувати розрахункову таблицю (табл. 18.14).

Таблиця 18.14. Розрахункова таблиця для методу спрямлюючих діаграм

1	2	3	4	5	6	7
Номер інтервалу	Правий кінець інтервалу	Частота	Накопичувана частота	Відносна накопичувана частота	Відносна накопичувана частота (%)	Квантил i
i	x_i	n_i	$\sum_{i=1}^m n_i$	$P_i = \frac{\sum_{i=1}^m n_i}{n}$	$P_i = \frac{\sum_{i=1}^m n_i}{n} \cdot 100\%$	u_{P_i}

Стовпець 6 обчислюється з міркувань зручності, в разі, коли цього вимагає таблиця квантелів.

Побудова $Q-Q$ діаграми. Для побудови $Q-Q$ діаграми необхідно побудувати в прямокутній системі координат точки $(x_i; u_{P_i})$ та пряму (18.41):

$$y = x, \quad (18.41)$$

яка будується по точках x_i (див. рис. 18.3).

В деякій літературі, зокрема в Гмурмані, рекомендується пряму представляти у вигляді (18.42):

$$y = \frac{x-a}{\sigma}, \quad (18.42)$$

де a, σ підбираються за допомогою точкових оцінок параметрів нормального розподілу $a = \bar{x}_B, \sigma = s$.

```
import numpy as np
import scipy.stats as stats
import matplotlib.pyplot as plt

# Генерація даних
np.random.seed(0)
normal_data = np.random.normal(loc=0, scale=1, size=1000) # Нормально розподілені
non_normal_data = np.random.exponential(scale=1, size=1000) # Ненормально розподілені

# Створення Q-Q діаграми для нормально розподілених даних
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
stats.probplot(normal_data, dist="norm", plot=plt)
plt.title('Q-Q діаграма для нормальних даних')

# Створення Q-Q діаграми для ненормально розподілених даних
plt.subplot(1, 2, 2)
stats.probplot(non_normal_data, dist="norm", plot=plt)
plt.title('Q-Q діаграма для ненормальних даних')

plt.tight_layout()
plt.show()
```

Рис. 18.3. Приклад застосування в Python

Прийняття рішення:

- якщо дані розподілені нормально, то точки на $Q-Q$ діаграмі будуть лежати на прямій лінії $y = x$;
- якщо дані розподілені ненормально, то точки значно відхиляються від неї (рис. 18.4, 18.5).

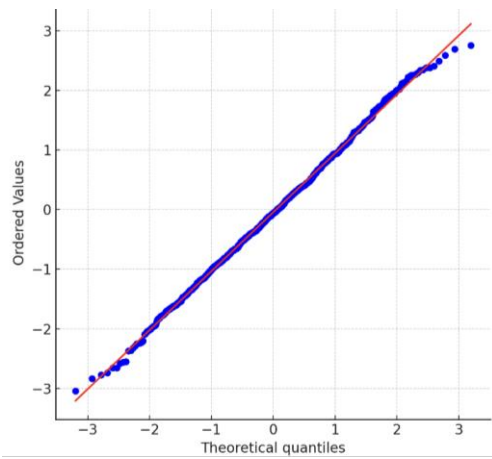


Рис. 18.4. $Q - Q$ діаграма для нормальних даних

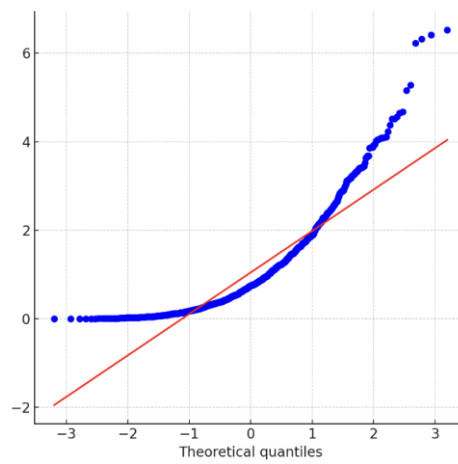


Рис. 18.5. $Q - Q$ діаграма для ненормальних даних

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ

1. Випробування, події,
2. Класичне означення ймовірності,
3. Частість настання подій.
4. Геометричне означення ймовірності,
5. Операції над подіями,
6. Класифікація подій,
7. Умовні ймовірності,
8. Теореми додавання та множення ймовірностей,
9. Формула повної ймовірності,
10. Формула Баєса,
11. Формула Бернуллі,
12. Найймовірніше число,
13. Формула Пуассона,
14. Локальна теорема Муавра-Лапласа,
15. Інтегральна теорема Муавра-Лапласа,
16. Випадкова величина. Класифікація випадкових величин.
17. Засоби представлення випадкової величини. Дискретна і неперервна випадкова величини,
18. Функція розподілу випадкової величини,
19. Функція щільності випадкової величини,
20. Мода, медіана випадкової величини,
21. Математичне сподівання, дисперсія,
22. Початкові моменти,
23. Центральні моменти,
24. Коефіцієнт асиметрії, ексцес,
25. Твірна функція,
26. Біноміальний розподіл,
27. Геометричний розподіл,
28. Пуассонівський розподіл,
29. Гіпергеометричний розподіл,
30. Розподіл Паскаля,
31. Уніформний розподіл,
32. Рівномірний розподіл,
32. Експоненційний розподіл,
34. Нормальний розподіл,
35. Логнормальний розподіл,
36. Розподіл «Хі квадрат»,
37. Розподіл Стюдента,
38. Розподіл Фішера-Снедекора,
39. Розподіл Парето,

40. Розподіл Ерланга,
41. Розподіл Релея,
42. Операції над випадковими величинами,
43. Функція від випадкової величини,
44. Композиція двох випадкових величин,
45. Система двох випадкових величин,
46. Умови незалежності компонент системи двох випадкових величин,
47. Шкала Чеддока,
48. Умовні закони розподілу системи двох випадкових величин,
49. Кореляційний момент (коваріація),
50. Коефіцієнт кореляції між двома випадковими величинами,
51. Коефіцієнт вибіркової кореляції Пірсона,
52. Коефіцієнт вибіркової рангової кореляції Спірсена,
53. Коефіцієнт вибіркової рангової кореляції Кендалла,
54. Коефіцієнт вибіркової кореляції Фехнера,
55. Нерівність Маркова,
56. Нерівність Чебишева,
57. Теорема Чебишева,
58. Теорема Ляпунова,
59. Наслідки із інтегральної теореми Муавра-Лапласа,
60. Ланцюги Маркова,
61. Схема «загибелі та розмноження»,
62. Циклічна схема,
63. Система рівнянь Колмогорова,
64. Система масового обслуговування,
65. Поняття вибірки та варіаційного ряду,
66. Полігон, гістограма, кумулята, огіва,
67. Точкові оцінки параметрів розподілу:
68. Метод моментів,
69. Метод максимальної правдоподібності,
70. Метод найменших квадратів,
71. Квантиль. Приклад розрахунку квантеля,
72. Інтервальні оцінки,
73. Статистичні гіпотези. Основні поняття,
74. Критерій Пірсона,
75. Критерій Колмогорова,
76. Критерій Стьюдента,
77. Критерій Фішера,
78. Побудова $Q - Q$ діаграми,
79. Типи статистичних критеріїв,
80. Етапи перевірки статистичних критеріїв.

ЛІТЕРАТУРА

1. Барковський В.В. та ін. // Математика для економістів. Теорій ймовірностей та математична статистика. / К.: Національна академія управління, 1999.
2. Гіхман І. І., Скороход А. В., Ядренко М. І. // Теорія ймовірностей та математична статистика. / Київ, Вища школа, 1979.
3. Дороговцев А. Я, Сільвестров Д. С., Скороход А. В., Ядренко М. Й. // Теорія ймовірностей (збірник задач). / Київ, Вища школа, 1977.
4. Денисюк В.П., Бобков В.М., Погребецька Т.А., Репета В.К. // Вища математика. Ч4. Теорій ймовірностей і математична статистика / К: вид-во «НАУ-друк», 2009, 256 с.
5. Жлуктенко В.І., Наконечний С.І. // Теорія ймовірностей та математична статистика Ч.1. Теорія ймовірностей. / К.: КНЕУ, 2000.
6. Жлуктенко В.І., Наконечний С.І. // Теорія ймовірностей та математична статистика Ч.2. Математична статистика / К.: КНЕУ, 2000.
7. Карташов М. В. // Імовірність, процеси, статистика: Посібник. / К: Видавничо-поліграфічний центр «Київський університет», 494 с., 2006.
8. Коваленко І. М., Гнеденко Б. В. // Теорія ймовірностей. / Київ, Вища школа, 1990.
9. Листопад В.В., Островська О.В. // Практикум з теорії ймовірностей із застосуванням інформаційно-комунікаційних технологій [Електронний ресурс]: навчальний посібник. / К.: НУХТ, 2016, 103 с.
10. Павлов О.А., Гавриленко О.В., Рибачук Л.В. // Навчальний посібник з дисципліни «Теорія ймовірностей, імовірнісні процеси та математична статистика». Курс лекцій. Частина 1 для студентів спеціальності 126 «Інформаційні системи та технології» [Електронний ресурс]: навчальний посібник. / Київ: КПІ, 154 с., 2021,
11. Павлов О.А., Гавриленко О.В., Жданова О.Г. // Навчальний посібник з дисципліни «Теорія ймовірностей, імовірнісні процеси та математична статистика». Курс лекцій. Частина 2 для студентів спеціальності 126 «Інформаційні системи та технології» [Електронний ресурс]: навчальний посібник. / Київ: КПІ, 72 с., 2022.
12. Павлов О.А., Гавриленко О.В. // Навчальний посібник з дисципліни «Теорія ймовірностей, імовірнісні процеси та математична статистика». Курс лекцій. Частина 3 для студентів спеціальності 126 «Інформаційні системи та технології» [Електронний ресурс]: навчальний посібник. / Київ: КПІ, 111 с., 2022.
13. Гавриленко О.В. // Навчальний посібник з дисципліни «Теорія ймовірностей, імовірнісні процеси та математична статистика». Розрахунково-графічна робота. Для студентів спеціальностей за спеціальностями 126 «Інформаційні системи та технології», 121 «Інженерія

програмного забезпечення» [Електронний ресурс]:навчальний посібник. / Київ: КПІ, 81 с., 2023.

14. Гавриленко О.В. // Навчальний посібник з дисципліни «Теорія ймовірностей, ймовірнісні процеси та математична статистика». Практикум. Частина 1 для студентів спеціальності 126 «Інформаційні системи та технології» [Електронний ресурс]:навчальний посібник. / Київ: КПІ, 140 с., 2021

15. Гавриленко О.В. // Навчальний посібник з дисципліни «Теорія ймовірностей, ймовірнісні процеси та математична статистика». Практикум. Частина 2 для студентів спеціальності 126 «Інформаційні системи та технології» [Електронний ресурс]:навчальний посібник. / Київ: КПІ, 130 с., 2023.

16. Скороход А.В // Теорія ймовірностей. Збірник задач. / К.: Вища школа, 1975.

17. Скороход. А.В. // Елементи теорії ймовірностей та теорії випадкових процесів» / К.: Вища школа, 1975.

18. Шефтель З.Г. // Теорія ймовірностей. / К.:Вища школа, 1994.

19. Братійчук М.С., Чечельницький О.А. // Лекції зі стохастики. Ймовірність. Статистика. Випадкові процеси» / Київ: електронні публікації на сайті факультету комп'ютерних наук та кібернетики, 395 с., 2021.

20. Галицька І.Є., Жданова О.Г., Кузнецов В.М. // Методичні вказівки до виконання індивідуальних завдань з дисципліни «Ймовірнісні процеси і математична статистика в автоматизованих системах». / К.: Політехніка, 44 с., 2002.

21. Sheldon Ross // A First Course in Probability / 10th ed. Pearson, pp. 528, 2018.

22. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying E. Ye. // Probability and Statistics for Engineers and Scientists / 9th ed. Pearson, pp. 816, 2016.

23. William Mendenhall, Robert J. Beaver, Barbara M. Beaver // Introduction to Probability and Statistics / 14th ed. Brooks/Cole, Cengage Learning, pp. 784, 2012.

24. Robert V. Hogg, Elliot A. Tanis // Probability and Statistical Inference / 8th ed. Prentice Hall, pp. 648, 2009.

25. Paul Newbold, William L. Carlson, Betty Thorne // Statistics for Business and Economics / 8th ed. Pearson, pp. 994, 2012.

26. О.В. Гавриленко, О.С. Жураковська, А.В. Коган, Р.М. Матвійчук, А.М. Піскун, Ю.І. Хавікова, О.А. Халус // Принцип формування портфелю публічних (адміністративних) послуг на основі аналізу статистичної інформації // Східно-Європейський журнал передових технологій, №3(117) 2022. DOI: <https://doi.org/10.15587/1729-4061.2022.260136>.

27. Sergii Telenyk, Grzegorz Nowakowski, Olena Gavrilenko, Mykhailo Miahkyi, Olena Khalus // An analysis of the influence of famous people's posts

on social networks on the cryptocurrency exchange rate / Bulletin of the polish academy of sciences technical sciences, Vol. 72(4), 2024. DOI: [10.24425/bpasts.2024.150117](https://doi.org/10.24425/bpasts.2024.150117).

28. Petro Bidyuk, Olena Gavrilenko, Mykhailo Myagkyi // The algorithm for predicting the cryptocurrency rate taking into account the influence of posts of a group of famous people in social networks // System research and information technologies, No.2, 2023. DOI: <https://doi.org/10.20535/SRIT.2308-8893.2023.2.02>.

29. Gavrilenko, O., & Myagkyi, M. // Analysis of communities and groups in social networks as a significant factor of influence on cryptocurrency rates. / Innovative technologies and scientific solutions for industries, 1 (27), 2024, pp. 18–25. DOI: <https://doi.org/10.30837/ITSSI.2024.27.018>.

30. Гавриленко О.В., Новаківська К.Д., Шумейко О.А. // Вибір найбільш впливових економічних факторів для прогнозування курсу долару США. / Вісник НТУ №54. – К.: НТУ, 2022. с.26-35. DOI: [10.33744/2308-6645-2022-4-54-026-035](https://doi.org/10.33744/2308-6645-2022-4-54-026-035).

ІНФОРМАЦІЙНІ РЕСУРСИ

1. Moodle <https://do.ipk.kpi.ua/course/view.php?id=5461>.

2. «Електронний Кампус КПІ» <https://ecampus.kpi.ua>.