

МІЖРЕГІОНАЛЬНА
АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ



МАУП

ПРАКТИКУМ З ТЕОРІЇ ЙМОВІРНОСТЕЙ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

За редакцією Р. К. Чорнея

*Рекомендовано
Міністерством освіти і науки України
як навчальний посібник
для студентів вищих навчальних закладів*

Київ 2003

ББК 22.17я73
П69

Автори: *Р. К. Чорней*, канд. фіз.-мат. наук, доц.
О. Ю. Дюженкова, канд. фіз.-мат. наук, доц.
О. Б. Жильцов, канд. фіз.-мат. наук, доц.
Г. М. Торбін, канд. фіз.-мат. наук, доц.
І. І. Юртин, канд. фіз.-мат. наук, доц.

Рецензенти: *П. С. Кнопов*, д-р фіз.-мат. наук, проф.
О. В. Школьный, канд. фіз.-мат. наук

*Схвалено Вченою радою Міжрегіональної Академії
управління персоналом (протокол № 1 від 28.01.03)*

*Рекомендовано Міністерством освіти і науки України
(лист № 14/18.2-2479 від 27.12.02)*

Практикум з теорії ймовірностей та математичної статистики:

П69 Навч. посіб. для студ. вищ. навч. закл. / Р. К. Чорней,
О. Ю. Дюженкова, О. Б. Жильцов та ін.; За ред. Р. К. Чорнея. —
К.: МАУП, 2003. — 328 с.: іл. — Бібліогр.: с. 321–322.

ISBN 966-608-341-8

У навчальному посібнику подано теоретичний матеріал, розв'язки типових задач, задачі для самостійного розв'язування із вказівками та відповідями до них. Значну увагу приділено методам статистичної обробки експериментальних даних.

Для студентів нематематичних спеціальностей вищих навчальних закладів, а також для всіх, хто застосує теорію ймовірностей і статистичні методи при розв'язуванні практичних задач.

ББК 22.17я73

ISBN 966-608-341-8

© Р. К. Чорней, О. Ю. Дюженкова,
О. Б. Жильцов та ін., 2003
© Міжрегіональна Академія
управління персоналом (МАУП), 2003

ЧАСТИНА І

Випадкові події

Розділ 1. Випадкові події та операції над ними. Означення ймовірності

1.1. Випадкові події

Забезпечення певного комплексу умов називають *випробуванням* або *дослідом*, а можливий результат випробування — *подією*. Наприклад, підкидання монети є випробуванням, а випадання «герба» або «номіналу» — подіями. Події позначатимемо великими латинськими літерами: *A, B, C*.

Подію називають *випадковою*, якщо вона може відбутися або не відбутися в даному випробуванні.

Достовірною називають подію, яка обов'язково відбудеться в даному випробуванні.

Неможливою називають подію, яка точно не відбудеться в даному випробуванні.

Зауважимо, що будь-яка подія пов'язана з певним випробуванням.

Дві події називають *сумісними*, якщо поява однієї з них не виключає появи іншої в одному й тому самому випробуванні.

Дві події називають *несумісними*, якщо вони не можуть відбутися одночасно в одному й тому самому випробуванні.

Попарно несумісні випадкові події A_1, A_2, \dots, A_n утворюють *повну групу подій*, якщо внаслідок випробування одна з них обов'язково відбудеться. Наприклад, події «виграш», «програш» і «нічия» (для певного гравця) утворюють повну групу подій у випробуванні — гри в шахи двох суперників.

Елементарними подіями (наслідками) у певному випробуванні називають усі можливі результати цього випробування, які не можна розкласти на простіші. Множину всіх можливих елементарних подій ω називають *простором елементарних подій*, який позначають Ω . Наприклад, при підкиданні грального кубика простір елементарних подій утворюють події $\omega_i = \{\text{випаде } i \text{ очок}\}$, $i = 1, 2, 3, 4, 5, 6$.

Елементарні події, при появі яких відбувається певна подія, називають *сприятливими* для цієї події. Наприклад, при підкиданні грального кубика для події $A = \{\text{випаде непарне число очок}\}$ сприятливими є елементарні події $\omega_1, \omega_3, \omega_5$.

Кожну подію можна розглядати як деяку підмножину простору елементарних подій у даному випробуванні. Зокрема, подія $A = \Omega$ є достовірною, а подія $B = \emptyset$ — неможливою.

Приклад 1. Монету підкидають двічі. Для даного випробування описати простір елементарних подій.

Розв'язання. При двократному підкиданні монети можливі чотири елементарних наслідки:

$$(A, A); \quad (A, P); \quad (P, A); \quad (P, P),$$

де A — випадання аверса (зображення «герба»), P — випадання реверса (зображення «номіналу»). Очевидно, вони утворюють повну групу подій, тому

$$\Omega = \{(A, A); (A, P); (P, A); (P, P)\}$$

є простір елементарних подій даного випробування. □

1.2. Операції над подіями

Сумою двох випадкових подій A і B називають таку подію, яка полягає в появі хоча б однієї з подій A або B , і позначають $A + B$ (або $A \cup B$).

Сумою n випадкових подій A_1, A_2, \dots, A_n називають таку подію, яка полягає в появі принаймні однієї з цих подій (позначається $\bigcup_{i=1}^n A_i$).

Добутком двох випадкових подій A і B називають таку подію, яка полягає в сумісній появі обох подій A і B , і позначають $A \cdot B$ (або $A \cap B$).

Добутком n випадкових подій A_1, A_2, \dots, A_n називають таку подію, яка полягає в сумісній появі всіх цих подій (позначається $\bigcap_{i=1}^n A_i$).

Різницею двох випадкових подій A і B називають подію, яка полягає в тому, що відбувається подія A і не відбувається подія B , і позначають $A - B$ (або $A \setminus B$).

Подію \bar{A} називають *протилежною* до події A в даному випробуванні, якщо вона відбувається тоді, коли не відбувається подія A , тобто $\bar{A} = \Omega - A$. Очевидно, протилежні події несумісні й утворюють повну групу подій.

Приклад 1. У ящику містяться кульки білого та чорного кольору. Навмання з нього виймають одну кульку. Подія $A = \{\text{вийнято кульку білого кольору}\}$, подія $B = \{\text{вийнято кульку чорного кольору}\}$. Сумісні чи несумісні ці події?

Розв'язання. Ці події несумісні, тому що поява події A виключає можливість появи події B , і навпаки. У даному випробуванні події A і B є протилежними:

$$A = \bar{B}, \quad B = \bar{A}. \quad \square$$

Приклад 2. Підкидають два гральних кубики. Нехай події $A_i = \{\text{випаде } i \text{ очок на першому кубіку}\}$, $i = 1, 2, 3, 4, 5, 6$, $B_j = \{\text{випаде } j \text{ очок на другому кубіку}\}$, $j = 1, 2, 3, 4, 5, 6$. Виразити через A_i , B_j такі події:

- а) сума очок на двох кубиках дорівнює п'яти;
- б) випаде в сумі хоча б десять очок;
- в) випаде в сумі не більше трьох очок.

Розв'язання. а) Нехай $C_1 = \{\text{сума очок на двох кубиках дорівнює п'яти}\}$. Ця подія можлива лише тоді, коли на першому кубіку випаде i очок, а на другому — j очок так, щоб $i + j = 5$, тобто $i = 1, j = 4$, або $i = 2, j = 3$, або $i = 3, j = 2$, або $i = 4, j = 1$. Отже,

$$C_1 = A_1 \cdot B_4 + A_2 \cdot B_3 + A_3 \cdot B_2 + A_4 \cdot B_1.$$

б) Позначимо $C_2 = \{\text{випаде в сумі хоча б десять очок}\}$. Подія C_2 відбудеться тоді, коли на двох кубиках у сумі випаде або 10, або 11, або 12 очок, тобто $i = 4, j = 6$, або $i = 5, j = 5$, або $i = 6, j = 4$, або $i = 6, j = 5$, або $i = 5, j = 6$, або $i = 6, j = 6$. Тому

$$C_2 = A_4 \cdot B_6 + A_5 \cdot B_5 + A_6 \cdot B_4 + A_6 \cdot B_5 + A_5 \cdot B_6 + A_6 \cdot B_6.$$

в) Нехай $C_3 = \{\text{випаде в сумі не більше трьох очок}\}$. Оскільки найменша кількість очок, яка може випасти на кожному кубіку, дорівнює одиниці, то подія C_3 можлива лише тоді, коли сума очок на двох кубиках дорівнюватиме або двом, або трьом. Тому

$$C_3 = A_1 \cdot B_1 + A_1 \cdot B_2 + A_2 \cdot B_1. \quad \square$$

Приклад 3. Два стрільці роблять постріл у мішень по одному разу. Подія $A = \{\text{у мішень влучив перший стрілець}\}$, подія $B = \{\text{у мішень влучив другий стрілець}\}$. Виразити через A і B такі події: $C = \{\text{два влучення в мішень}\}$, $D = \{\text{жодного влучення в мішень}\}$, $E = \{\text{хоча б одне влучення в мішень}\}$, $F = \{\text{лише одне влучення в мішень}\}$.

Розв'язання. Простір елементарних подій складається з чотирьох подій

$$A \cdot B, \quad \bar{A} \cdot B, \quad A \cdot \bar{B}, \quad \bar{A} \cdot \bar{B}.$$

Подія C відбудеться тоді, коли обидва стрільці влучать у мішень. Тому вона є добутком двох подій A і B . Отже,

$$C = A \cdot B.$$

Подія D полягає в тому, що в мішень не влучить жодний стрілець, тобто не влучить ані перший (\bar{A}), ані другий (\bar{B}). Тому

$$D = \bar{A} \cdot \bar{B}.$$

Подія E відбудеться тоді, коли в мішень влучить хоча б один стрілець. Це може бути тоді, коли або обидва стрільці влучать у мішень, або перший влучить, другий не влучить, або перший не влучить, а другий влучить. Тому

$$E = AB + \bar{A}B + A\bar{B},$$

тобто

$$E = A + B.$$

Подія F полягає в тому, що перший стрілець влучить у мішень, а другий не влучить або другий влучить, а перший не влучить. Тому

$$F = A\bar{B} + \bar{A}B. \quad \square$$

1.3. Класичне означення ймовірності

Нехай усі елементарні наслідки рівноможливі.

Ймовірність події A дорівнює відношенню кількості елементарних наслідків, сприятливих цій події, до кількості всіх рівноможливих елементарних наслідків у даному випробуванні.

Ймовірність події A позначають $\mathbf{P}(A)$, тому за означенням

$$\mathbf{P}(A) = \frac{m}{n}, \quad (1)$$

де m — кількість елементарних наслідків, сприятливих події A ; n — кількість усіх елементарних наслідків у даному випробуванні.

З класичного означення ймовірності випливає, що

$$0 \leq \mathbf{P}(A) \leq 1,$$

причому $\mathbf{P}(A) = 0$, якщо $A = \emptyset$ — неможлива подія, і $\mathbf{P}(A) = 1$, якщо $A = \Omega$ — достовірна подія.

Приклад 1. В урні міститься 5 білих, 3 чорних і 4 червоних кульки. Навмання виймають одну. Знайти ймовірність того, що навмання вийнята кулька червона.

Розв'язання. Нехай подія $A = \{\text{вийнята з урни кулька червона}\}$. Загальна кількість кульок в урні — $5 + 3 + 4 = 12$, причому вийняти можна будь-яку з них з однаковою ймовірністю. Тому в даному випробуванні є 12 рівноможливих наслідків, тобто $n = 12$. Кількість подій, що сприяють події A , визначається кількістю червоних кульок, тобто $m = 4$. Отже, за означенням (1) імовірності

$$\mathbf{P}(A) = \frac{m}{n} = \frac{4}{12} = \frac{1}{3}. \quad \square$$

Приклад 2. Знайти ймовірність того, що вибране випадковим чином двозначне число ділиться на:

- а) 3;
- б) 5.

Розв'язання. У даному разі випробування полягає в тому, що вибирається випадковим чином двозначне число. Наслідком такого випробування є одне з чисел від 10 до 99. Оскільки таких чисел 90, то $n = 90$.

а) Нехай подія $A = \{\text{вибране двозначне число ділиться на 3}\}$. Оскільки кожне третє з 90 двозначних чисел ділиться на 3, то сприятливими для події A є 30 наслідків, тобто $m = 30$. Тоді за формулою (1) імовірність події A

$$\mathbf{P}(A) = \frac{30}{90} = \frac{1}{3}.$$

б) Нехай подія $B = \{\text{вибране двозначне число ділиться на 5}\}$. Загальна кількість наслідків випробування, як і в попередньому випадку, $n = 90$. Визначимо кількість чисел, які діляться на 5. Очевидно, що таких чисел буде $m = 18$ (кожне п'яте число ділиться на 5). Отже,

$$\mathbf{P}(B) = \frac{18}{90} = \frac{1}{5}. \quad \square$$

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

1.4. Геометричне означення ймовірності

Нехай множина всіх елементарних наслідків випробування нескінченна і утворює деяку множину Ω , усі елементарні наслідки рівноможливі, причому події A сприяють ті елементарні події, які утворюють множину $A \subseteq \Omega$. Тоді ймовірність події A дорівнює відношенню міри множини A до міри множини Ω , тобто

$$\mathbf{P}(A) = \frac{m(A)}{m(\Omega)}. \quad (2)$$

Мірою множини на прямій, площині, у просторі є відповідно довжина, площа, об'єм геометричної фігури, яку утворює ця множина.

Приклад 1. Два дійсних числа випадковим чином вибираються з інтервалу $[0; 5]$. Яка ймовірність того, що:

- а) сума двох чисел менша 4;
- б) добуток двох чисел більший 5;
- в) різниця двох чисел менша 2, а їх добуток більший 3.

Розв'язання. Позначимо через x перше число, вибране випадковим чином з інтервалу $[0; 5]$, а через y — друге число. Тоді $0 \leq x \leq 5$, $0 \leq y \leq 5$. Внаслідок нескінченної кількості таких дійсних чисел треба скористатися означенням геометричної ймовірності. У цьому разі множиною всіх можливих наслідків випробування є квадрат (рис. 1–3) зі стороною 5, площа якого дорівнює 25, тобто

$$m(\Omega) = 25.$$

- а) Нехай подія $A = \{\text{сума двох чисел менша 4}\}$. Тоді маємо

$$A = \{(x; y) : x + y < 4\},$$

або

$$A = \{y < 4 - x\}.$$

Отже, елементарні наслідки випробування, що сприяють події A , утворюють фігуру (заштриховану на рис. 1), площа якої

$$m(A) = \frac{1}{2} \cdot 4 \cdot 4 = 8.$$

Застосовуючи формулу (2) для геометричної ймовірності, отримуємо

$$\mathbf{P}(A) = \frac{m(A)}{m(\Omega)} = \frac{8}{25}.$$

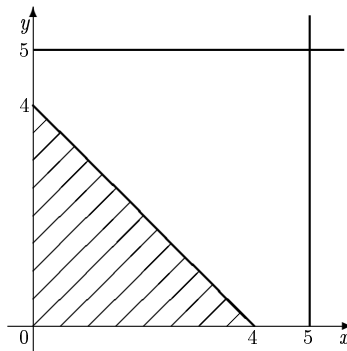


Рис. 1

б) Нехай подія $B = \{\text{добуток двох чисел більший } 5\}$. Тоді

$$B = \{(x; y) : xy > 5\},$$

або

$$B = \left\{ (x; y) : y > \frac{5}{x} \right\}.$$

Множині всіх наслідків, що сприяють події B , відповідає фігура ABC , заштрихована на рис. 2, де лінія AC — графік функції $y = \frac{5}{x}$.

Обчислимо площу цієї фігури за допомогою визначеного інтеграла:

$$m(B) = \int_1^5 \left(5 - \frac{5}{x}\right) dx = (5x - 5 \ln x) \Big|_1^5 = 20 - 5 \ln 5.$$

Отже,

$$\mathbf{P}(B) = \frac{m(B)}{m(\Omega)} = \frac{20 - 5 \ln 5}{25} \approx 0,478.$$

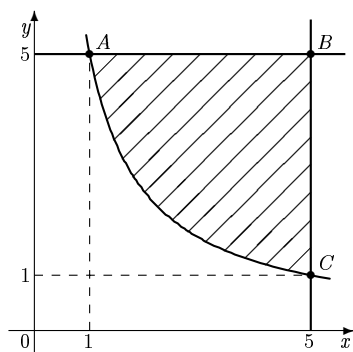


Рис. 2

в) Нехай подія $C = \{\text{різниця двох чисел менша 2, а їх добуток більший 3}\}$, тоді

$$C = \{(x; y): |x - y| < 2, xy > 3\},$$

звідки отримуємо нерівності

$$y < x + 2, \quad y > x - 2, \quad y > \frac{x}{3}.$$

Множина точок, координати яких задовольняють зазначені нерівності, утворює фігуру $ABCDE$ (рис. 3). Пряма AE задається рівнянням $y = x + 2$, пряма CD — рівнянням $y = x - 2$, а лінія $MEDN$ — графіком функції $y = \frac{3}{x}$. Абсциса точки M визначається з рівності

$\frac{3}{x} = 5$, тобто $x = \frac{3}{5}$. Ордината точки N визначається з рівності $y = \frac{3}{5}$. Абсцису точки E як точки перетину двох ліній $y = \frac{3}{x}$ і $y = x + 2$ визначаємо з рівняння $x + 2 = \frac{3}{x}$, звідки маємо $x = 1$. Аналогічно визначаємо абсцису точки D як точки перетину двох ліній $y = x - 2$ і $y = \frac{3}{x}$, тобто $x = 3$.

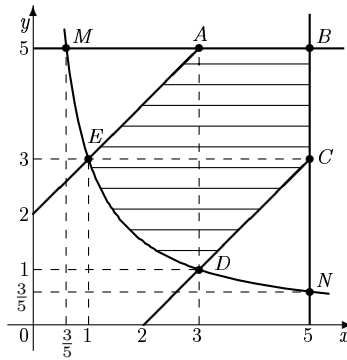


Рис. 3

Для визначення ймовірності події C обчислимо площу області $ABCDE$. Очевидно, що

$$S_{ABCDE} = S_{MBN} - S_{MAE} - S_{CND} = S_{MBN} - 2S_{CND},$$

оскільки $S_{MAE} = S_{CND}$. Тоді

$$\begin{aligned} S_{MBN} &= \int_{\frac{3}{5}}^5 \left(5 - \frac{3}{x}\right) dx = \\ &= (5x - 3 \ln x) \Big|_{\frac{3}{5}}^5 = 25 - 3 \ln 5 - \left(3 - 3 \ln \frac{3}{5}\right) = \\ &= 22 + 3 \ln 3 - 6 \ln 5; \end{aligned}$$

$$\begin{aligned}
 S_{CND} &= \int_3^5 \left(x - 2 - \frac{3}{x} \right) dx = \left(\frac{x^2}{2} - 2x - 3 \ln x \right) \Big|_3^5 = \\
 &= \frac{25}{2} - 10 - 3 \ln 5 - \left(\frac{9}{2} - 6 - 3 \ln 3 \right) = 4 + 3 \ln 3 - 3 \ln 5.
 \end{aligned}$$

Отже,

$$m(C) = 22 + 3 \ln 3 - 6 \ln 5 - 8 - 6 \ln 3 + 6 \ln 5 = 14 - 3 \ln 3,$$

а

$$\mathbf{P}(C) = \frac{14 - 3 \ln 3}{25} \approx 0,428. \quad \square$$

Приклад 2. Два студенти призначили зустріч у певному місці між третьою та четвертою годинами дня. Той, хто прийде перший, чекає іншого протягом 15 хв, після чого йде з місця зустрічі. Знайти ймовірність того, що зустріч відбудеться.

Розв'язання. Позначимо через x час приходу на місце зустрічі першого студента, а через y — другого. Припустимо також, що година, під час якої має відбутися зустріч, несуттєва, тобто студенти можуть зустрітися протягом однієї години. Тоді для x і y виконуються умови

$$0 \leq x \leq 1 \quad \text{і} \quad 0 \leq y \leq 1.$$

Нехай подія A полягає в тому, що зустріч відбулася. Це можливо лише тоді, коли різниця між часом приходу на місце зустрічі першого та другого студентів менша 15 хв, або $\frac{1}{4}$ год, тобто

$$|x - y| \leq \frac{1}{4}.$$

Звідси отримуємо нерівності

$$y \leq x + \frac{1}{4} \quad \text{та} \quad y \geq x - \frac{1}{4}.$$

Множина точок, координати яких задовольняють ці нерівності, утворює фігуру $ABCDE$, зображену на рис. 4.

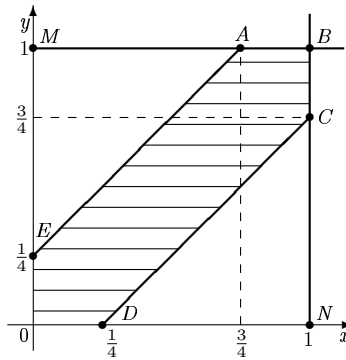


Рис. 4

Оскільки

$$S_{ABCDOE} = S_{MBNO} - 2S_{EMA},$$

причому

$$S_{MBNO} = 1, \quad S_{EMA} = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16},$$

то

$$S_{ABCDOE} = 1 - \frac{9}{16} = \frac{7}{16}.$$

Отже, ймовірність того, що зустріч відбудеться,

$$\mathbf{P}(A) = \frac{7}{16}.$$

□

1.5. Статистичне означення ймовірності

Оскільки класичне означення ймовірності передбачає, що всі елементарні наслідки випробування рівноможливі, що важко обґрунтувати, то розглядають ще й *статистичне означення ймовірності*.

Відносною частотою події A називають відношення кількості випробувань, у яких подія A відбулася, до кількості всіх проведених випробувань. Відносну частоту події A позначають $W(A)$. Тоді

$$W(A) = \frac{m}{n},$$

де m — кількість випробувань, у яких відбулася подія A ; n — кількість усіх проведених випробувань.

Число, навколо якого групуються значення частоти події A при великій кількості випробувань, називають ймовірністю події A :

$$P(A) = \lim_{n \rightarrow \infty} W(A).$$

Приклад 1. При перевірці готової продукції було виявлено 5 бракованих одиниць товару з 200 перевірених. Знайти відносну частоту бракованих одиниць товару.

Розв'язання. Нехай подія A полягає в тому, що виявлено браковану одиницю товару. Тоді відносна частота події A

$$W(A) = \frac{m}{n} = \frac{5}{200} = 0,025. \quad \square$$

Приклад 2. При стрільбі по мішені було виявлено, що відносна частота влучень дорівнює 0,85. Проведено 100 пострілів. Скільки пострілів були влучними?

Розв'язання. Нехай подія A полягає в тому, що постріл був влучним. Тоді з формули для відносної частоти події A отримуємо, що кількість влучних пострілів

$$m = n \cdot W(A) = 100 \cdot 0,85 = 85. \quad \square$$

Задачі до розділу 1

Задача 1. Два стрільці роблять по одному пострілу в мішень. Сумісні чи несумісні події A і B , якщо $A = \{\text{перший стрілець влучив у мішень}\}$, $B = \{\text{другий стрілець влучив у мішень}\}$.

Відповідь. Сумісні.

Задача 2. Перевірити, чи утворюють повну групу такі події:

а) $A = \{\text{випадання не менше трьох очок}\}$, $B = \{\text{випадання не більше трьох очок}\}$ у випробуванні — підкиданні грального кубика;

б) $A = \{\text{один промах}\}$, $B = \{\text{одне влучення}\}$, $C = \{\text{два влучення}\}$ у випробуванні, яке полягає в тому, що два стрільці роблять по одному пострілу в мішень;

в) $A = \{\text{випадання двох аверсів}\}$, $B = \{\text{випадання хоча б одного реверсу}\}$ у випробуванні — підкиданні двох монет.

Чи є протилежними події A і B у кожному з цих випробувань.

Відповідь. У всіх випадках події не утворюють повної групи подій і не є протилежними.

Задача 3. Описати простір елементарних подій для випробування, яке полягає в підкиданні грального кубика.

Відповідь. $\Omega = \{w_1; w_2; w_3; w_4; w_5; w_6\}$, де $w_i = \{\text{випаде } i \text{ очок}\}$.

Задача 4. Підкидають двічі гральний кубик. Нехай $A_i = \{\text{випаде } i \text{ очок при першому підкиданні}\}$, $B_i = \{\text{випаде } i \text{ очок при другому підкиданні}\}$. Виразити через A_i , B_i такі події:

$A = \{\text{обидва рази випаде парна кількість очок}\}$,

$B = \{\text{сума очок при двох підкиданнях дорівнює 6}\}$,

$C = \{\text{сума очок при двох підкиданнях більше 8}\}$,

$D = \{\text{обидва рази випаде однакова кількість очок}\}$.

Відповідь. $A = \bar{A}_1 \cdot \bar{A}_3 \cdot \bar{A}_5 \cdot \bar{B}_1 \cdot \bar{B}_3 \cdot \bar{B}_5$, $B = A_1B_5 + A_2B_4 + A_3B_3 + A_4B_2 + A_5B_1$, $C = A_3B_6 + A_4B_5 + A_4B_6 + A_5B_4 + A_5B_5 + A_5B_6 + A_6B_3 + A_6B_4 + A_6B_5 + A_6B_6$, $D = A_1B_1 + A_2B_2 + A_3B_3 + A_4B_4 + A_5B_5 + A_6B_6$.

Задача 5. Студент на екзамені відповідає на білет, у якому три питання. Нехай $A_i = \{\text{студент відповів на } i\text{-те питання}\}$. Виразити через A_i такі події:

$A = \{\text{студент відповів принаймні на два питання}\}$,

$B = \{\text{студент не відповів на жодне питання}\}$,

$C = \{\text{студент відповів тільки на одне питання}\}$.

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

Задача 6. Підкидають три монети. Для даного випробування записати простір елементарних подій і розкласти на елементарні такі події:

$$\begin{aligned} A &= \{\text{на двох монетах випаде «аверс»}\}, \\ B &= \{\text{на жодній монеті не випаде «реверс»}\}, \\ C &= \{\text{хоча б на одній монеті випаде «реверс»}\}. \end{aligned}$$

Відповідь. $\Omega = \{(A, A, A); (A, A, P); (A, P, A); (A, P, P); (P, A, A); (P, A, P); (P, P, A); (P, P, P)\}$, $A = \{(A, A, P); (A, P, A); (P, A, A)\}$, $B = \{(A, A, A)\}$, $C = \{(A, A, P); (A, P, A); (A, P, P); (P, A, A); (P, A, P); (P, P, A); (P, P, P)\}$.

Задача 7. Стрілець виконує чотири постріли в мішень. Нехай подія $A_i = \{\text{влучення в мішень при } i\text{-му пострілі}\}$. Виразити через A_i такі події:

$$\begin{aligned} A &= \{\text{три влучення}\}, \\ B &= \{\text{хоча б один промах}\}, \\ C &= \{\text{не більше одного влучення}\}, \\ D &= \{\text{хоча б одне влучення}\}. \end{aligned}$$

Відповідь. $A = A_1 A_2 A_3 \bar{A}_4 + A_1 A_2 \bar{A}_3 A_4 + A_1 \bar{A}_2 A_3 A_4 + \bar{A}_1 A_2 A_3 A_4$,
 $B = \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4$, $C = \bar{A}_2 \bar{A}_3 \bar{A}_4 + \bar{A}_1 A_2 \bar{A}_3 \bar{A}_4 + \bar{A}_1 \bar{A}_2 A_3 \bar{A}_4 + \bar{A}_1 \bar{A}_2 \bar{A}_3 A_4$,
 $D = \bar{A}_1 \bar{A}_2 \bar{A}_3 \bar{A}_4$.

Задача 8. Підкидають чотири рази монету. Розкласти на елементарні такі події:

$$\begin{aligned} A &= \{\text{два рази випаде «аверс»}\}, \\ B &= \{\text{випаде не більше одного «реверса»}\}. \end{aligned}$$

Відповідь. $A = \{(A, A, P, P); (A, P, A, P); (A, P, P, A); (P, A, A, P); (P, A, P, A); (P, P, A, A)\}$, $B = \{(A, A, A, A); (A, A, A, P); (A, A, P, A); (A, P, A, A); (P, A, A, A)\}$.

Задача 9. У класі навчається 10 хлопців і 12 дівчат. Знайти ймовірність того, що навмання вибраний учень хлопець.

Відповідь. $P(\text{«вибраний хлопець»}) = \frac{5}{11}$.

Задача 10. Знайти ймовірність того, що при двох підкиданнях монети:

- а) хоча б один раз випаде «реверс»;
- б) жодного разу не випаде «реверс».

Відповідь. а) $P = \frac{3}{4}$; б) $P = \frac{1}{4}$.

Задача 11. Визначити ймовірність того, що при підкиданні грального кубика випаде:

- а) парна кількість очок;
- б) непарна кількість очок;
- в) кількість очок більше чотирьох;
- г) кількість очок менше десяти;
- д) кількість очок більше шести.

Відповідь. а) $P = \frac{1}{2}$; б) $P = \frac{1}{2}$; в) $P = \frac{1}{3}$; г) $P = 1$; д) $P = 0$.

Задача 12. У круг навмання кинута точка. Яка ймовірність того, що вона потрапить у квадрат, вписаний у це коло?

Відповідь. $P = \frac{2}{\pi}$.

Задача 13. Абонент протягом години чекає телефонного дзвінка. Яка ймовірність того, що йому подзвонять протягом перших 15 хв?

Відповідь. $P = \frac{1}{4}$.

Задача 14. Студент і студентка домовились зустрітися в певному місці між 19 і 20 год. Якщо студент приходить перший, він чекає студентку 30 хв і йде з місця зустрічі. Якщо першою прийде студентка, вона чекає студента 10 хв і йде з місця зустрічі. Знайти ймовірність того, що зустріч відбудеться в призначеному місці.

Відповідь. $P = \frac{19}{36}$.

Задача 15. На площині проведено прями на однаковій відстані 15 см. На цю площину кидають монету діаметром 5 см. Знайти ймовірність того, що монета перетне одну з паралельних прямих.

Відповідь. $P = \frac{1}{3}$.

Задача 16. Два дійсних числа p і q випадковим чином вибираються з інтервалу $[-2; 2]$. Знайти ймовірність того, що квадратне рівняння $x^2 + px + q = 0$ має:

- а) дійсні корені;
- б) дійсні додатні корені;
- в) один корінь додатний, а інший від'ємний;
- г) хоча б один дійсний корінь з інтервалу $[-1; 1]$.

Відповідь. а) $P = \frac{7}{12}$; б) $P = \frac{1}{24}$; в) $P = \frac{1}{2}$; г) $P = \frac{25}{48}$.

Задача 17. При перевірці готової продукції було виявлено 7 бракованих одиниць товару із 140 перевірених. Знайти відносну частоту бракованих одиниць товару.

Відповідь. $W = 0,04$.

Задача 18. При стрільбі по мішені було виявлено, що відносна частота влучень дорівнює 0,9. Проведено 70 пострілів. Скільки пострілів були влучними?

Відповідь. $m = 63$.

Розділ 2. Елементи комбінаторики та їх застосування при обчисленні ймовірностей

При обчисленні ймовірностей подій досить часто потрібно підраховувати кількість елементарних подій (сприятливих деякій події або всіх можливих подій). Здебільшого це зумовлює великі труднощі, подолати які допомагає комбінаторика, що вивчає способи підрахунку кількості розміщень, перестановок, сполучень.

Перш ніж представити деталі, нагадаємо, що вираз $n!$ читається «*n*-факторіал» і означає добуток усіх натуральних чисел до n :

$$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n,$$

причому вважають, що $0! = 1$.

Розміщеннями із n елементів по m називають множини із m елементів, вибраних із n елементів, які можуть розрізнятися між собою як складом елементів, так і їх порядком. Наприклад, розміщеннями із трьох елементів по два будуть такі множини: $\{1; 2\}$, $\{1; 3\}$, $\{2; 1\}$, $\{2; 3\}$, $\{3; 1\}$, $\{3; 2\}$. Кількість усіх розміщень із n елементів по m визначають за формулою

$$A_n^m = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot (n - m + 1) = \frac{n!}{(n - m)!}. \quad (3)$$

Перестановками із n елементів називають множини із n елементів, що відрізняються лише їх порядком. Наприклад, перестановками із трьох елементів будуть такі множини: $\{1; 2; 3\}$, $\{1; 3; 2\}$, $\{2; 1; 3\}$, $\{2; 3; 1\}$, $\{3; 1; 2\}$, $\{3; 2; 1\}$. Кількість усіх перестановок із n елементів визначають так:

$$P_n = \prod_{k=1}^n k = 1 \cdot 2 \cdot \dots \cdot n = n! \quad (4)$$

Сполученнями із n елементів по m називають множини із m елементів, вибраних із n елементів, які розрізняються між собою тільки

складом елементів. Наприклад, сполученнями із трьох елементів по два будуть такі множини: $\{1; 2\}$, $\{1; 3\}$, $\{2; 3\}$. Кількість усіх сполучень із n елементів по m визначають за формулою

$$C_n^m = \frac{(n - m + 1) \cdot (n - m + 2) \cdot \dots \cdot n}{m!} = \frac{n!}{m! \cdot (n - m)!}. \quad (5)$$

Між переліченими поняттями існують такі співвідношення:

$$P_n = A_n^n = n!; \quad C_n^m = \frac{A_n^m}{P_m}; \quad C_n^m = C_n^{n-m}; \\ C_n^0 = C_n^n = 1; \quad C_n^1 = n; \quad C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n.$$

Приклад 1. За навчальним планом студенти протягом семестру вивчають 10 дисциплін. На кожний день плануються 4 пари з різних дисциплін. Скількома способами можна скласти розклад занять на один день.

Розв'язання. Усі можливі розклади занять на один день — це розміщення із 10 елементів по 4, оскільки вони можуть різнитися дисциплінами або порядком. Тому кількість способів складання розкладу на один день визначається за формулою (3), тобто

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

Приклад 2. Скільки п'ятизначних чисел можна утворити за допомогою цифр 1, 2, 3, 4, 5, якщо кожна з них у числі зустрічається лише один раз?

Розв'язання. Різні п'ятизначні числа, утворені за допомогою цифр 1, 2, 3, 4, 5, можна отримати лише перестановкою цих цифр у числі. Тому їх кількість визначається перестановкою з п'яти елементів. Згідно з формулою (4)

$$P_5 = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120. \quad \square$$

Приклад 3. У групі налічується 15 студентів. Скількома способами можна обрати:

- а) студентську раду в кількості трьох студентів;
- б) голову, заступника та секретаря студентської ради?

Розв'язання. а) У цьому разі студентська рада з трьох студентів, вибрана із 15 студентів групи, різниться лише складом. Порядок вибраних студентів не має значення. Тому кількість можливих виборів визначається кількістю сполучень із 15 елементів по 3, яка згідно з формулою (5) дорівнює

$$C_{15}^3 = \frac{15 \cdot 14 \cdot 13}{1 \cdot 2 \cdot 3} = 455.$$

б) У цьому разі студентська рада з трьох студентів, вибрана із 15 студентів групи, різниться не лише складом, а й тим, хто буде головою, заступником і секретарем. Порядок вибраних студентів має значення. Тому кількість можливих виборів визначається кількістю розміщень із 15 елементів по 3, яка згідно з формулою (3) дорівнює

$$A_{15}^3 = 15 \cdot 14 \cdot 13 = 2730. \quad \square$$

Приклад 4. В урні міститься 8 чорних і 5 білих кульок. Навмання виймають 4 кульки. Яка ймовірність того, що витягнули:

- а) 4 чорних кульки;
- б) 2 чорних і 2 білих кульки?

Розв'язання. а) Кількість кульок в урні — 13. Визначимо, скількома способами можна дістати 4 кульки з 13 (без повернення їх в урну). Оскільки при цьому не має значення порядок вийнятих кульок, кількість сполучень визначаємо за формулою (5), тобто

$$n = C_{13}^4 = \frac{13 \cdot 12 \cdot 11 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4} = 715$$

є кількість усіх елементарних наслідків у даному випробуванні.

Нехай подія A полягає в тому, що вийнято 4 чорних кульки. Ці 4 чорних кульки можна вийняти з урни з восьми чорних кульок. Тому кількість наслідків, сприятливих події A , визначається за формулою (5), а саме

$$m = C_8^4 = \frac{8 \cdot 7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3 \cdot 4} = 70.$$

Отже, імовірність того, що витягнули 4 чорних кульки,

$$\mathbf{P}(A) = \frac{m}{n} = \frac{70}{715} = \frac{14}{143} \approx 0,098.$$

б) У цій задачі кількість усіх можливих наслідків така сама, як і в попередньому пункті: $n = 715$. Нехай подія B полягає в тому, що вийнято 2 чорних і 2 білих кульки. Визначимо, скількома способами це можна зробити. Дві чорні кульки можна вийняти з восьми, які містяться в урні,

$$m_1 = C_8^2 = \frac{8 \cdot 7}{1 \cdot 2} = 28$$

способами, а 2 білих — із п'яти білих, що знаходяться в урні,

$$m_2 = C_5^2 = \frac{5 \cdot 4}{1 \cdot 2} = 10$$

способами. Тоді 2 білих і 2 чорних кульки можна вийняти

$$m = m_1 \cdot m_2 = 28 \cdot 10 = 280$$

способами. Тому кількість наслідків, що сприяє події B , $m = 280$.

Отже, імовірність події B

$$\mathbf{P}(B) = \frac{280}{715} = \frac{56}{143} \approx 0,39. \quad \square$$

Приклад 5. Шістнадцять варіантів контрольної роботи написані на окремих картках і розподіляються випадковим чином серед 14 студентів, які сидять в одному ряду. Кожний студент отримує одну картку. Знайти ймовірність того, що:

- а) варіанти 1 і 2 не будуть використані;
- б) варіанти 1 і 2 видадуть студентам, які сидять поруч.

Розв'язання. Маємо випробування розподілу 16 білетів серед 14 студентів. У цьому разі події відрізняються одна від одної не лише номерами варіантів, що розподіляються серед студентів, а й порядком розподілу. Тому такі сполучення називають розміщеннями, а кількість таких розміщень визначається за формулою (3):

$$n = A_{16}^{14} = \frac{16!}{2!}.$$

а) Позначимо через A подію, яка полягає в тому, що варіанти 1 і 2 залишаться нерозподіленими. Тоді інші 14 білетів розподіляться серед 14 студентів. Такі сполучення називають перестановками, а їх кількість визначається за формулою (4):

$$m = P_{14} = 14!$$

Отже, застосувавши класичну формулу ймовірності (1), матимемо

$$\mathbf{P}(A) = \frac{m}{n} = \frac{14!2!}{16!} = \frac{1}{15 \cdot 8} \approx 0,008.$$

б) Нехай подія B полягає в тому, що варіанти 1 і 2 видані студентам, які сидять поруч. У ряду із 14 місць є 13 пар сусідніх місць, причому в кожній парі варіанти можуть розподілятися двома способами:

$$m_1 = 13 \cdot 2 = 26.$$

Інші 14 варіантів білетів розподіляються між 12 студентами

$$m_2 = A_{14}^{12} = \frac{14!}{2!}$$

способами. Тому події B сприяють

$$m = m_1 \cdot m_2 = 26 \cdot \frac{14!}{2!}$$

наслідків.

Отже, ймовірність події B

$$\mathbf{P}(B) = \frac{26 \cdot 14! \cdot 2!}{2! \cdot 16!} = \frac{13}{120} \approx 0,108. \quad \square$$

Приклад 6. Комплект складається з восьми різних виробів, 3 з яких коштують по 4 грн, ще 3 — по 5 грн і 2 останніх — по 3 грн. Знайти ймовірність того, що взяті навмання 2 вироби коштують 7 грн.

Розв'язання. Вибір двох виробів з восьми є сполученням, кількість яких

$$n = C_8^2 = \frac{8 \cdot 7}{2} = 28.$$

Це загальна кількість наслідків у випробуванні — виборі двох виробів з восьми.

Нехай подія A полягає в тому, що вартість двох вибраних виробів становить 7 грн. Це можливо лише тоді, коли один виріб коштує 4 грн., а інший — 3 грн. Оскільки кількість виробів вартістю 4 грн. дорівнює 3, а вартістю 3 грн. — 2, то вибрати 2 вироби вартістю 7 грн. можна $m = 3 \cdot 2 = 6$ способами. Отже,

$$P(A) = \frac{m}{n} = \frac{6}{28} \approx 0,214. \quad \square$$

Приклад 7. У групі 10 хлопців і 5 дівчат, серед яких вибирають двох для участі в конференції. Яка ймовірність того, що:

- а) виберуть двох хлопців;
- б) виберуть хлопця й дівчину?

Розв'язання. У групі 15 осіб, серед яких 10 хлопців і 5 дівчат. Вибір двох осіб із 15 є сполученнями, кількість яких

$$n = C_{15}^2 = \frac{15 \cdot 14}{2} = 105.$$

Це загальна кількість наслідків у випробуванні — виборі двох осіб із 15.

а) Нехай подія $A = \{\text{вибрали двох хлопців}\}$. Загальна кількість наслідків, що сприяють події A , визначається кількістю виборів двох хлопців із 10:

$$m = C_{10}^2 = \frac{10 \cdot 9}{2} = 45.$$

Отже,

$$P(A) = \frac{m}{n} = \frac{45}{105} \approx 0,43.$$

б) Нехай подія $B = \{\text{вибрали хлопця й дівчину}\}$. Це можливо зробити

$$m = 10 \cdot 5 = 50$$

способами.

Отже,

$$P(B) = \frac{m}{n} = \frac{50}{105} \approx 0,48. \quad \square$$

Задачі до розділу 2

Задача 1. У студентську раду інституту обрано 8 студентів. Скількома способами можна обрати керівну групу у складі голови, заступника та секретаря?

Відповідь. 336 способів.

Задача 2. У бібліотеку одночасно зайшли чотири відвідувачі. Скількома способами вони можуть утворити чергу?

Відповідь. 24 способи.

Задача 3. У деканаті знаходиться 10 студентських книжок. Шестеро студентів, які зайшли до деканату, навмання беруть по одній студентській книжці. Яка ймовірність того, що:

- а) усі студенти взяли свої студентські книжки;
- б) 4 студенти взяли свої студентські книжки?

Відповідь. а) $P = \frac{1}{148200}$; б) $P = \frac{5}{2016} \approx 0,0025$.

Задача 4. З партії у 40 виробів, яка містить 20 % браку, навмання виймають 7 виробів. Знайти ймовірність того, що серед вийнятих виявиться 5 бракованих виробів.

Відповідь. $P = \frac{704816}{2330445} \approx 0,3$.

Задача 5. У туриста є десять однакових консервних банок, серед яких три банки — з тушкованим м'ясом, а сім — з рибою. Під час зливи етикетки відклеїлись. Яка ймовірність того, що дві навмання відкриті банки відрізняться вмістом?

Відповідь. $P = \frac{7}{15}$.

Задача 6. У гуманітарному класі 25 учнів, з яких семеро цікавляться математикою, десятеро — економічними дисциплінами, інші восьмеро — літературою. Знайти ймовірність того, що два навмання вибраних учні цікавитимуться однією дисципліною.

Відповідь. $P = \frac{47}{150} \approx 0,31$.

Задача 7. У вазі 5 троянд рожевого, 7 червоного та 3 білого кольору. Навмання дістають дві троянди. Яка ймовірність того, що вони будуть:

- а) одного кольору;
- б) різних кольорів?

Відповідь. а) $P = \frac{34}{105} \approx 0,32$; б) $P = \frac{71}{105} \approx 0,68$.

Задача 8. У кошику лежать 8 червоних і 2 зелених яблука. Знайти ймовірність того, що серед чотирьох узятих яблук будуть:

- а) усі червоні;
- б) 2 зелених яблука.

Відповідь. а) $P = \frac{1}{3}$; б) $P = \frac{2}{15}$.

Задача 9. Чотири білети в театр розігрують 5 хлопців і 7 дівчат. Знайти ймовірність того, що в театр підуть 2 хлопця і 2 дівчини.

Відповідь. $P = \frac{14}{33} \approx 0,4$.

Задача 10. Із десяти лотерейних білетів книжкової лотереї, що перебувають у продажу, два виграшні. Визначити ймовірність того, що серед куплених п'яти білетів:

- а) один виграшний;
- б) хоча б один виграшний.

Відповідь. а) $P = \frac{5}{9}$; б) $P = \frac{7}{9}$.

Задача 11. П'ять книжок, серед яких 2 підручники з математики, довільно розміщують на полиці. Яка ймовірність того, що ці 2 підручники стоятимуть поряд?

Відповідь. $P = \frac{2}{5}$.

Задача 12. В урні 9 білих і 6 чорних кульок. З урни дістають навмання 5 кульок. Знайти ймовірність того, що три з них — білі, а дві — чорні.

Відповідь. $P = \frac{60}{143} \approx 0,42$.

Задача 13. На дев'яти картках написано літери «м», «с», «к», «н», «е», «о», «і», «т», «о». Знайти ймовірність того, що, навмання викладаючи ці картки, ви отримаєте слово «економіст».

Відповідь. $P = \frac{1}{181440}$.

Задача 14. Три стрільці незалежно один від одного стріляють у мішень. Ймовірність влучення в мішень для першого дорівнює 0,7, для другого — 0,8, для третього — 0,9. Яка ймовірність того, що:

- а) хоча б один із них влучить у мішень;
- б) тільки двоє влучать у мішень;
- в) жоден не влучить у мішень?

Відповідь. а) $P = 0,994$; б) $P = 0,398$; в) $P = 0,006$.

Задача 15. З урни, у якій міститься 8 кольорових і 7 білих кульок, навмання дістають одну за одною дві кульки. Знайти ймовірність того, що вони кольорові, якщо:

- а) першу кульку повертають в урну, після чого дістають другу;
- б) першу кульку не повертають в урну.

Відповідь. а) $P = \frac{64}{225} \approx 0,28$; б) $P = \frac{4}{15}$.

Задача 16. У змаганнях з легкої атлетики беруть участь 6 учнів восьмого класу, 7 — дев'ятого та 8 — десятого. Знайти ймовірність того, що за жеребкуванням до першої пари бігунів потраплять двоє учнів з одного класу.

Відповідь. $P = \frac{32}{105} \approx 0,3$.

Розділ 3. Основні формули додавання і множення ймовірностей

Ймовірність суми двох довільних випадкових подій A і B дорівнює сумі їх ймовірностей без ймовірності їхнього добутку, тобто

$$\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cdot B).$$

Якщо події A і B несумісні, то $\mathbf{P}(A \cdot B) = 0$. Тоді

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

Якщо випадкові події A_1, A_2, \dots, A_n попарно несумісні, то ймовірність появи хоча б однієї з цих подій дорівнює сумі їх ймовірностей:

$$\mathbf{P}(A_1 + A_2 + \dots + A_n) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \dots + \mathbf{P}(A_n).$$

Сума ймовірностей випадкових подій A_1, A_2, \dots, A_n , які утворюють повну групу, дорівнює одиниці, тобто

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \dots + \mathbf{P}(A_n) = 1.$$

Зокрема, для протилежних подій A і \bar{A} виконується рівність

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

Випадкові події A та B називаються *залежними*, якщо ймовірність появи однієї з них залежить від появи або не появи іншої. В іншому випадку події A та B називаються *незалежними*.

Ймовірність події B , обчислена за умови появи події A , називають *умовною ймовірністю* події B (за умови появи події A) і позначають $\mathbf{P}(B/A)$ або $\mathbf{P}_A(B)$.

Якщо події A та B незалежні, то $\mathbf{P}(B/A) = \mathbf{P}(B)$, і навпаки, $\mathbf{P}(A/B) = \mathbf{P}(A)$.

Ймовірність добутку двох випадкових подій A і B дорівнює добутку ймовірності однієї з цих подій на умовну ймовірність іншої події за умови, що перша подія відбулася, тобто

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

Якщо події незалежні, то

$$\mathbf{P}(A \cdot B) = \mathbf{P}(A) \cdot \mathbf{P}(B).$$

Нехай є n незалежних випадкових подій A_1, A_2, \dots, A_n . Імовірність появи хоча б однієї з цих подій визначається за формулою

$$\mathbf{P}(A_1 + A_2 + \dots + A_n) = 1 - \mathbf{P}(\bar{A}_1) \cdot \mathbf{P}(\bar{A}_2) \cdot \dots \cdot \mathbf{P}(\bar{A}_n).$$

Приклад 1. Мішень складається з трьох областей. Імовірність влучення стрільцем у першу область мішені дорівнює 0,45, у другу область — 0,35, у третю — 0,15. Яка ймовірність того, що при одному пострілі стрілець влучить:

- а) у першу або в другу область;
- б) не влучить у мішень?

Розв'язання. Уведемо позначення:

подія $A_1 = \{\text{влучення в першу область}\}$;
подія $A_2 = \{\text{влучення в другу область}\}$;
подія $A_3 = \{\text{влучення в третю область}\}$;
подія $A_4 = \{\text{невлучення в мішень}\}$.

Тоді

$$\mathbf{P}(A_1) = 0,45; \quad \mathbf{P}(A_2) = 0,35; \quad \mathbf{P}(A_3) = 0,15.$$

- а) При одному пострілі події A_1, A_2, A_3 несумісні. Тому

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

б) Події A_1, A_2, A_3, A_4 попарно несумісні та утворюють повну групу випадкових подій. Тому

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \mathbf{P}(A_3) + \mathbf{P}(A_4) = 1.$$

Отже,

$$\mathbf{P}(A_4) = 1 - \mathbf{P}(A_1) - \mathbf{P}(A_2) - \mathbf{P}(A_3) = 1 - 0,45 - 0,35 - 0,15 = 0,05. \quad \square$$

Приклад 2. Для виконання одного завдання замовник звернувся до двох виконавців. Ймовірність того, що перший виконавець виконає замовлення, дорівнює 0,8, а другий — 0,9. Знайти ймовірність того, що завдання замовника буде виконано.

Розв'язання. Позначимо подію $A = \{\text{завдання для замовника буде виконано}\}$; $A_1 = \{\text{завдання буде виконано першим виконавцем}\}$; $A_2 = \{\text{завдання буде виконано другим виконавцем}\}$. Тоді

$$A = A_1 + A_2.$$

Оскільки виконання замовлення першим виконавцем не виключає виконання цього самого замовлення другим, події сумісні. Тоді

$$\mathbf{P}(A) = \mathbf{P}(A_1 + A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cdot A_2).$$

Однак події A_1 та A_2 незалежні, тому що ймовірність виконання замовлення першим виконавцем не залежить від того, чи виконав це завдання інший виконавець. Тому

$$\mathbf{P}(A_1 \cdot A_2) = \mathbf{P}(A_1) \cdot \mathbf{P}(A_2).$$

Отже,

$$\mathbf{P}(A) = 0,8 + 0,9 - 0,8 \cdot 0,9 = 0,98. \quad \square$$

Задачі до розділу 3

Задача 1. У першій урні міститься 7 білих і 3 чорних кульки; у другій — 5 білих і 5 чорних кульок; у третій — 4 білих і 6 чорних кульок. З кожної урни навмання виймають по одній кульці. Знайти ймовірність того, що серед вибраних кульок виявиться:

- а) лише одна біла кулька;
- б) дві білі кульки;
- в) три білі кульки;
- г) хоча б одна біла кулька.

Відповідь. а) $\mathbf{P} = 0,36$; б) $\mathbf{P} = 0,41$; в) $\mathbf{P} = 0,14$; г) $\mathbf{P} = 0,91$.

Задача 2. Знайти ймовірність того, що випадковим чином вибране двозначне число ділиться:

- а) або на 2, або на 3;
- б) на 2 і на 3.

Відповідь. а) $P = \frac{2}{3}$; б) $P = \frac{1}{6}$.

Задача 3. Групу із 30 студентів, серед яких 10 відмінників, випадковим чином розбивають на дві рівні підгрупи. Знайти ймовірність того, що в кожній підгрупі буде по 5 відмінників.

Відповідь. $P = \frac{1001}{3335} \approx 0,3$.

Задача 4. У групі 8 студентів чоловічої статі та 4 жіночої. Випадковим чином групу розбивають на чотири рівні частини. Знайти ймовірність того, що в кожній частині виявиться одна студентка.

Відповідь. $P = \frac{9}{55} \approx 0,16$.

Задача 5. Комплект із 20 виробів містить 40 % нестандартних. Двічі навмання із комплекту виймають по 6 виробів. Знайти ймовірність того, що після цих двох виймань у комплекті залишаться лише нестандартні вироби.

Відповідь. $P = \frac{1}{38760}$.

Задача 6. Перший стрілець може влучити в ціль з імовірністю 0,8, другий — з імовірністю 0,9, а третій — з імовірністю 0,85. Яка ймовірність того, що хоча б один стрілець влучить у ціль?

Відповідь. $P = 0,997$.

Задача 7. Робітник обслуговує одночасно три верстати. Імовірність порушення роботи протягом години для першого дорівнює 0,1, для другого — 0,15, для третього — 0,2. Яка ймовірність того, що:

- а) усі три верстати працюватимуть протягом години;
- б) хоча б один із них вийде з ладу?

Відповідь. а) $P = 0,712$; б) $P = 0,329$.

Розділ 4. Формула повної ймовірності. Формула Байєса

Якщо випадкова подія A може відбутися лише сумісно з однією із несумісних між собою подій B_1, B_2, \dots, B_n , що утворюють повну групу, то ймовірність появи події A обчислюється за *формулою повної ймовірності*

$$\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(B_k) \cdot \mathbf{P}(A/B_k). \quad (6)$$

В умовах теореми невідомо, з якою із несумісних подій B_1, B_2, \dots, B_n відбудеться подія A . Тому появу кожної з цих подій можна вважати гіпотезою, а $\mathbf{P}(B_k)$ — ймовірністю k -ї гіпотези.

Якщо в результаті проведеного випробування відбулася подія A , то умовна ймовірність $\mathbf{P}(B_k/A)$ може не дорівнювати $\mathbf{P}(B_k)$. Щоб отримати умовну ймовірність, використовують *формулу Байєса*:

$$\mathbf{P}(B_k/A) = \frac{\mathbf{P}(B_k) \cdot \mathbf{P}(A/B_k)}{\mathbf{P}(A)} = \frac{\mathbf{P}(B_k) \cdot \mathbf{P}(A/B_k)}{\sum_{k=1}^n \mathbf{P}(B_k) \cdot \mathbf{P}(A/B_k)}. \quad (7)$$

Приклад 1. У першому ящику 30 деталей, з яких 20 стандартних. У другому — 15 деталей, з яких 10 стандартних. З другого ящика беруть навмання одну деталь і перекладають її до першого ящика.

а) Знайти ймовірність того, що навмання взята після цього деталь з першого ящика стандартна.

б) Нехай відомо, що з першого ящика вийнято стандартну деталь. Знайти ймовірність того, що до першого ящика перекладено саме:

- стандартну деталь;
- нестандартну деталь.

Розв'язання. Уведемо такі події: $A = \{\text{з першого ящика взято стандартну деталь}\}$; $B_1 = \{\text{з другого ящика переклали до першого стандартну деталь}\}$; $B_2 = \{\text{з другого ящика переклали до першого нестандартну деталь}\}$.

а) За умовою задачі з першого ящика деталь виймається лише після того, як до нього покладуть деталь, узятую з другого ящика, тобто подія A відбудеться лише тоді, коли здійсниться подія B_1 або подія B_2 . Ці події несумісні, а подія A може відбутися лише сумісно з однією із них. Тому для знаходження ймовірності події A можна скористатися формулою повної ймовірності (6):

$$\mathbf{P}(A) = \mathbf{P}(B_1) \cdot \mathbf{P}(A/B_1) + \mathbf{P}(B_2) \cdot \mathbf{P}(A/B_2).$$

Знайдемо потрібні ймовірності:

$$\begin{aligned} \mathbf{P}(B_1) &= \frac{10}{15} = \frac{2}{3}; & \mathbf{P}(B_2) &= \frac{5}{15} = \frac{1}{3}; \\ \mathbf{P}(A/B_1) &= \frac{21}{31}; & \mathbf{P}(A/B_2) &= \frac{20}{31}. \end{aligned}$$

Отже,

$$\mathbf{P}(A) = \frac{2}{3} \cdot \frac{21}{31} + \frac{1}{3} \cdot \frac{20}{31} = \frac{62}{93} = \frac{2}{3}.$$

б) Якщо подія A вже відбулася, то умовна ймовірність того, що до першого ящика перекладено саме стандартну деталь, така:

$$\mathbf{P}(B_1/A) = \frac{\mathbf{P}(B_1) \cdot \mathbf{P}(A/B_1)}{\mathbf{P}(A)} = \frac{\frac{2}{3} \cdot \frac{21}{31}}{\frac{2}{3}} = \frac{21}{31},$$

а нестандартну деталь — така:

$$\mathbf{P}(B_2/A) = \frac{\mathbf{P}(B_2) \cdot \mathbf{P}(A/B_2)}{\mathbf{P}(A)} = \frac{\frac{1}{3} \cdot \frac{20}{31}}{\frac{2}{3}} = \frac{10}{31}. \quad \square$$

Приклад 2. В урні 10 кульок — 3 білих і 7 чорних. Навмання виймають дві кульки. Знайти ймовірність того, що друга кулька виявилася чорною, якщо:

- а) першу кульку повертали до урни;
- б) першу кульку не повертали до урни.

Розв'язання. Нехай подія A_i полягає в тому, що за i -м разом узято білу кульку, а подія B_i — за i -м разом узято чорну кульку.

а) Якщо кульку, яку взято першою, повертають до урни, то ймовірність появи другої чорної кульки не залежить від того, яку кульку взято першою. Тому

$$\mathbf{P}(B_2) = \frac{7}{10} = 0,7.$$

б) Якщо першу кульку не повертати, то ймовірність появи другої чорної кульки вже залежить від того, яку кульку було взято першою. Якщо першою взяли білу кульку, в урні залишилося 2 білих і 7 чорних кульок. Тому

$$\mathbf{P}(B_2/A_1) = \frac{7}{9}.$$

Якщо першою взяли чорну кульку, то

$$\mathbf{P}(B_2/B_1) = \frac{6}{9} = \frac{2}{3}.$$

Ймовірність того, що першою взяли білу кульку,

$$\mathbf{P}(A_1) = \frac{3}{10} = 0,3,$$

чорну кульку —

$$\mathbf{P}(B_1) = \frac{7}{10} = 0,7.$$

Оскільки події A_1 і B_1 несумісні та утворюють повну групу, то за формулою повної ймовірності (6) отримуємо

$$\begin{aligned}\mathbf{P}(B_2) &= \mathbf{P}(A_1) \cdot \mathbf{P}(B_2/A_1) + \mathbf{P}(B_1) \cdot \mathbf{P}(B_2/B_1) = \\ &= 0,3 \cdot \frac{7}{9} + 0,7 \cdot \frac{2}{3} = 0,7.\end{aligned}$$

□

Приклад 3. Деталі, виготовлені цехом заводу, потрапляють для перевірки їхньої стандартності до одного з двох контролерів. Ймовірність того, що деталь потрапить до першого контролера, дорівнює

0,6, до другого — 0,4. Імовірність того, що придатна деталь буде визнана стандартною першим контролером, дорівнює 0,94, другим — 0,98. Придатна деталь при перевірці визнана стандартною. Знайти ймовірність того, що деталь перевіряв:

- перший контролер;
- другий контролер.

Розв'язання. Позначимо такі події:

$A = \{\text{придатна деталь визнана стандартною}\};$

$B_1 = \{\text{деталь перевіряв перший контролер}\};$

$B_2 = \{\text{деталь перевіряв другий контролер}\}.$

Тоді згідно з умовою задачі

$$\begin{aligned} \mathbf{P}(B_1) &= 0,6; & \mathbf{P}(B_2) &= 0,4; \\ \mathbf{P}(A/B_1) &= 0,94; & \mathbf{P}(A/B_2) &= 0,98. \end{aligned}$$

Згідно з формулою Байеса (7) маємо

$$\begin{aligned} \mathbf{P}(B_1/A) &= \frac{\mathbf{P}(B_1) \cdot \mathbf{P}(A/B_1)}{\mathbf{P}(A)} = \frac{0,6 \cdot 0,94}{0,6 \cdot 0,94 + 0,4 \cdot 0,98} = \\ &= \frac{141}{239} \approx 0,59; \\ \mathbf{P}(B_2/A) &= \frac{\mathbf{P}(B_2) \cdot \mathbf{P}(A/B_2)}{\mathbf{P}(A)} = \frac{0,4 \cdot 0,98}{0,6 \cdot 0,94 + 0,4 \cdot 0,98} = \\ &= \frac{98}{239} \approx 0,41. \end{aligned}$$

□

Задачі до розділу 4

Задача 1. На двох автоматичних лініях виготовляють однакові деталі: на першій — 30 %, на другій — 70 %. Імовірність виготовлення стандартної деталі на першій лінії дорівнює 0,9, на другій — 0,5. Усі виготовлені на цих лініях деталі надходять на склад.

а) Знайти ймовірність того, що навмання вибрана зі складу деталь стандартна.

б) Навмання вибрана деталь, що виготовлена на одній із ліній, виявилася стандартною. Знайти ймовірність того, що вона виготовлена на першій лінії.

Відповідь. а) $P = 0,62$; б) $P = \frac{27}{62} \approx 0,4$.

Задача 2. У першій урні міститься 4 білих і 3 чорних кульки, а в другій — 3 білих і одна чорна кулька. Із першої урни навмання вийняли одну кульку і переклали її в другу. Знайти ймовірність того, що після перекладання навмання вийнята з другої урни кулька буде білою.

Відповідь. $P = \frac{5}{7}$.

Задача 3. В урні 8 кульок — 5 білих і 3 чорних. Навмання виймають дві кульки. Знайти ймовірність того, що друга кулька виявилася чорною, якщо:

- а) першу кульку повертали до урни;
- б) першу кульку не повертали до урни.

Відповідь. а) $P = \frac{3}{8}$; б) $P = \frac{3}{8}$.

Задача 4. Із 30 студентів групи, що прийшли на екзамен, 8 підготовлені відмінно, 10 — добре, 8 — задовільно, а решта — незадовільно. Програма екзамену включає 50 питань. Білет містить 3 питання. Студент, підготовлений відмінно, знає всі питання; добре — 40 питань; задовільно — 25 питань і незадовільно — 10 питань.

а) Знайти ймовірність того, що навмання викликаний студент відповість на всі 3 питання білета.

б) Студент відповів на всі питання. Знайти ймовірність того, що студент підготовлений:

- добре;
- задовільно;
- незадовільно.

Відповідь. а) $P = \frac{691}{1470} \approx 0,47$; б) $P\{\text{добре}\} = \frac{247}{691} \approx 0,36$; $P\{\text{задовільно}\} = \frac{46}{691} \approx 0,07$; $P\{\text{незадовільно}\} = \frac{6}{691} \approx 0,01$.

Задача 5. До центру статистичних досліджень надходить інформація з трьох пунктів: з першого — 50 %, з другого — 30 %, з третього — 20 % усієї інформації. Ймовірність допущення помилки при обробці статистичних даних у першому пункті дорівнює 0,1, у другому — 0,05, у третьому — 0,15. Яка ймовірність того, що отримана центром у даний момент часу інформація цілком правильна?

Відповідь. $P = 0,905$.

Задача 6. В одному класі 5 відмінників, у другому — 3 відмінники, а в третьому класі відмінників немає. З навмання взятого класу вибрали учня. Знайти ймовірність того, що він відмінник, якщо в кожному класі вчиться по 30 дітей.

Відповідь. $P = \frac{4}{45} \approx 0,09$.

Задача 7. Два економісти заповнюють документи, які складають у спільну папку. Ймовірність зробити помилку для першого економіста дорівнює 0,1, для другого — 0,2. Перший економіст заповнив 40 документів, другий — 60. Під час перевірки навмання взятий із папки документ виявився з помилкою. Знайти ймовірність того, що його склав перший економіст.

Відповідь. $P = 0,25$.

Задача 8. Робітник обслуговує три верстати. Ймовірність виготовлення бракованої деталі на першому верстаті дорівнює 0,04, на другому — 0,01, а на третьому — 0,07. Продуктивність праці першого верстата втричі перевищує продуктивність другого, а третього — удвічі менша від продуктивності другого.

а) Знайти ймовірність того, що навмання вибрана деталь, яка виготовлена робітником, бракована.

б) Узята навмання деталь виявилася бракованою. На якому верстаті найімовірніше вона виготовлена?

Відповідь. а) $P = \frac{11}{300} \approx 0,037$; б) на першому верстаті.

Задача 9. Імовірності того, що під час роботи комп'ютера станеться збій в арифметичному пристрої, у оперативній пам'яті або в пристрої введення співвідносяться як 2 : 1 : 3. Імовірності віднайти збій у цих пристроях дорівнюють відповідно 0,9; 0,75; 0,7. Знайти ймовірність знаходження збою в роботі комп'ютера.

Відповідь. $P = 0,775$.

Задача 10. У кошику міститься 20 яблук білого та 15 червоного кольору. Навмання виймають три яблука з кошика і вкладають замість них три яблука червоного кольору. Потім навмання виймають одне яблуко. Яка ймовірність того, що вийнято яблуко червоного кольору?

Відповідь. $P \approx 0,4776$.

Розділ 5. Незалежні повторні випробування. Формула Бернуллі

5.1. Формула Бернуллі

Нехай проводиться n незалежних випробувань, причому ймовірність появи події A в кожному випробуванні одна й та сама, тобто не залежить від її появи в інших випробуваннях. Таку серію повторних незалежних випробувань називають *схемою Бернуллі*. Прикладом повторних незалежних випробувань є підкидання монети певну кількість разів, стрільба по мішені тощо.

Нехай випадкова подія A може відбутися в кожному випробуванні з однаковою ймовірністю $\mathbf{P}(A) = p$ або не відбутися з ймовірністю $q = 1 - p$. Ймовірність того, що при n випробуваннях подія A відбудеться m разів, визначається за *формулою Бернуллі*:

$$\mathbf{P}_n(m) = C_n^m \cdot p^m \cdot q^{n-m}. \quad (8)$$

Ймовірність появи події A в n випробуваннях m разів, де число m перебуває між числами k_1 і k_2 , $0 \leq k_1 \leq k_2 \leq n$, знаходиться за формулою

$$\mathbf{P}_n(k_1 \leq m \leq k_2) = \mathbf{P}_n(k_1) + \mathbf{P}_n(k_1 + 1) + \dots + \mathbf{P}_n(k_2).$$

Ймовірність появи події A в n випробуваннях хоча б один раз

$$\mathbf{P}_n(1 \leq m \leq n) = 1 - q^n.$$

Найімовірніша кількість m_0 появи події A в n випробуваннях визначається з нерівностей

$$n \cdot p - q \leq m_0 \leq n \cdot p + p.$$

Якщо ймовірність появи події A в кожному випробуванні дорівнює p , то кількість n випробувань, які необхідно здійснити, щоб з ймовірністю \mathbf{P} можна було стверджувати, що подія A відбудеться хоча б один раз, обчислюють за формулою

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

Приклад 1. Прилад складено з 10 блоків, надійність кожного з яких 0,8. Блоки можуть виходити з ладу незалежно один від одного. Знайти ймовірність того, що:

- а) відмовлять два блоки;
- б) відмовить хоча б один блок;
- в) відмовлять не менше двох блоків.

Знайти найімовірнішу кількість блоків, що вийдуть з ладу.

Розв'язання. Позначимо як подію A відмову блока. Тоді

$$P(A) = p = 1 - 0,8 = 0,2.$$

Тому $q = 0,8$. За умовою задачі $n = 10$. Застосувавши формулу Бернуллі (8) та наслідки з неї, матимемо:

- а) $P_{10}(2) = C_{10}^2 p^2 q^8 = 45 \cdot (0,2)^2 \cdot (0,8)^8 \approx 0,302$;
- б) $P_{10}(1 \leq m \leq 10) = 1 - q^{10} = 1 - (0,8)^{10} \approx 0,893$;
- в) $P_{10}(2 \leq m \leq 10) = 1 - [P_{10}(0) + P_{10}(1)] =$
 $= 1 - C_{10}^0 \cdot (0,2)^0 \cdot (0,8)^{10} - C_{10}^1 \cdot (0,2)^1 \cdot (0,8)^9 \approx$
 $\approx 0,624$.

Найімовірнішу кількість блоків, що вийдуть з ладу, знайдемо з нерівностей

$$n \cdot p - q \leq m_0 \leq n \cdot p + p,$$

тобто

$$10 \cdot 0,2 - 0,8 \leq m_0 \leq 10 \cdot 0,2 + 0,2.$$

Отже,

$$m_0 = 2. \quad \square$$

Приклад 2. Оператор за хвилину відправляє 3 повідомлення. Через скільки хвилин імовірність відправлення оператором хоча б одного помилкового повідомлення буде не меншою 0,952, якщо ймовірність помилки оператора дорівнює 0,02?

Розв'язання. Позначимо як подію A помилку оператора. Визначимо, скільки повідомлень має відправити оператор, щоб з імовірністю 0,952 можна було стверджувати, що він зробить хоча б одну помилку, якщо ймовірність помилки оператора за умовою $p = 0,02$:

$$n > \frac{\ln(1 - 0,952)}{\ln(1 - 0,02)} > 150.$$

Отже, оператор має відправити понад 150 повідомлень. \square

5.2. Граничні теореми у схемі Бернуллі

Для наближеного обчислення ймовірності появи події A у n незалежних випробуваннях m разів схеми Бернуллі при великих n і малих p таких, що $np < 10$, доцільно використовувати *формулу Пуассона*:

$$P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda = np.$$

Значення функції P_n^m з формули Пуассона наведено в дод. 3.

Локальною функцією Лапласа називають функцію

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Інтегральною функцією Лапласа називають функцію

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

Значення функцій $\varphi(x)$ і $\Phi(x)$ наведено відповідно в дод. 2 і 1.

Локальна теорема Муавра — Лапласа. Якщо у схемі Бернуллі кількість випробувань n достатньо велика, а ймовірність p появи події A в усіх випробуваннях однакова та відмінна від нуля й одиниці, то ймовірність появи m разів події A можна знайти наближено за формулою

$$P_n(m) = \frac{1}{\sqrt{npq}} \varphi(x),$$

де $x = \frac{m - np}{\sqrt{npq}}$; $\varphi(x)$ — локальна функція Лапласа.

Інтегральна теорема Муавра — Лапласа. Якщо у схемі Бернуллі кількість випробувань n достатньо велика, а ймовірність p появи події A в усіх випробуваннях однакова та відмінна від нуля й одиниці, то ймовірність появи події A не менше m_1 і не більше m_2 разів можна знайти за наближеною формулою

$$\mathbf{P}_n(m_1 \leq m \leq m_2) = \Phi(x_2) - \Phi(x_1),$$

де $x_2 = \frac{m_2 - np}{\sqrt{npq}}$, $x_1 = \frac{m_1 - np}{\sqrt{npq}}$, $\Phi(x)$ — інтегральна функція Лапласа.

Теорема Бернуллі. Якщо в n незалежних випробуваннях ймовірність p появи події A однакова й подія A відбудеться m разів, то для будь-якого додатного числа $\varepsilon > 0$ виконується рівність

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{m}{n} - p \right| \geq \varepsilon \right) = 0,$$

тобто подія, для якої відхилення визначається формулою

$$\left| \frac{m}{n} - p \right| \geq \varepsilon$$

при великих значеннях n , майже неможлива. Тому протилежна подія

$$\left| \frac{m}{n} - p \right| < \varepsilon$$

майже достовірна:

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{m}{n} - p \right| < \varepsilon \right) = 1.$$

Якщо $m_1 = np - \varepsilon n$, $m_2 = np + \varepsilon n$, то

$$x_2 = \frac{np + \varepsilon n - np}{\sqrt{npq}} = \varepsilon \sqrt{\frac{n}{pq}}; \quad x_1 = \frac{np - \varepsilon n - np}{\sqrt{npq}} = -\varepsilon \sqrt{\frac{n}{pq}}$$

і тому

$$\mathbf{P}\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) = \mathbf{P}(np - \varepsilon n \leq m \leq np + \varepsilon n) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Ці формули доцільно застосовувати за умови $n > 100$, $npq > 20$.

Приклад 1. Підручник надруковано тиражем 90000 примірників. Імовірність неправильного брошурування підручника дорівнює 0,0001. Знайти ймовірність того, що тираж має 5 бракованих підручників.

Розв'язання. Брошурування підручника можна розглядати як випробування, що вкладається в схему Бернуллі. Кількість випробувань n велика, а ймовірність кожного випробування p незначна. Тому в цьому разі доцільно застосувати формулу Пуассона.

Згідно з умовою задачі маємо

$$n = 90000, \quad p = 0,0001.$$

Отже, при $\lambda = np = 9$ з дод. 3 маємо

$$\mathbf{P}_{90000}(5) = \frac{9^5}{5!} \cdot e^{-9} \approx 0,0607. \quad \square$$

Приклад 2. Гральний кубик підкидають 800 разів. Яка ймовірність того, що кількість очок, кратна трьом, з'явиться 267 разів?

Розв'язання. У цій задачі кількість випробувань $n = 800$ досить велика. Імовірність того, що при підкиданні кубика випаде число, кратне трьом, дорівнює $p = \frac{1}{3}$ і постійна для всіх випробувань, $q = 1 - p = \frac{2}{3}$. Тому для обчислення ймовірності скористаємося локальною теоремою Муавра — Лапласа. Для цього знайдемо

$$x = \frac{267 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = 0,025.$$

Отже,

$$\mathbf{P}_{800}(267) = \frac{3}{40} \cdot \varphi(0,025) = \frac{3}{40} \cdot 0,3988 \approx 0,03. \quad \square$$

Приклад 3. Гральний кубик підкидають 800 разів. Яка ймовірність того, що кількість очок, кратна трьом, з'явиться не менше 260 разів і не більше 274 разів?

Розв'язання. Для знаходження ймовірності $\mathbf{P}_{800}(260 \leq m \leq 274)$ застосуємо інтегральну теорему Муавра — Лапласа. Визначимо x_1 , x_2 :

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{260 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = \frac{20}{40} = -0,5;$$
$$x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{274 - 800 \cdot \frac{1}{3}}{\frac{40}{3}} = \frac{22}{40} = 0,55.$$

Отже, за дод. 1 знаходимо

$$\begin{aligned} \mathbf{P}_{800}(260 \leq m \leq 274) &= \Phi(0,55) - \Phi(-0,5) = 0,2088 + 0,1915 = \\ &= 0,4003. \end{aligned} \quad \square$$

Приклад 4. Імовірність появи події в кожному із 625 незалежних випробувань дорівнює 0,8. Знайти ймовірність того, що відносна частота появи події відхиляється від ймовірності за абсолютною величиною не більше ніж на 0,04.

Розв'язання. За умовою задачі

$$n = 625; \quad p = 0,8; \quad q = 0,2; \quad \varepsilon = 0,04.$$

Потрібно знайти $\mathbf{P} \left(\left| \frac{m}{n} - 0,8 \right| \leq 0,04 \right)$.

За теоремою Бернуллі маємо

$$\begin{aligned}\mathbf{P}\left(\left|\frac{m}{n} - 0,8\right| \leq 0,04\right) &= 2\Phi\left(0,04\sqrt{\frac{625}{0,8 \cdot 0,2}}\right) = 2\Phi(2,5) = \\ &= 2 \cdot 0,4938 = 0,9876.\end{aligned}\quad \square$$

Приклад 5. Імовірність появи деякої події в кожному з незалежних випробувань дорівнює 0,5. Знайти кількість випробувань n , при якій з імовірністю 0,7698 можна очікувати, що відносна частота появи події відхилиться від її ймовірності за абсолютною величиною не більше ніж на 0,02.

Розв'язання. За умовою задачі

$$p = 0,5; \quad q = 0,5; \quad \varepsilon = 0,02.$$

Потрібно знайти кількість випробувань n , для якої

$$\mathbf{P}\left(\left|\frac{m}{n} - 0,8\right| \leq 0,02\right) = 0,7698.$$

За теоремою Бернуллі маємо

$$\mathbf{P}\left(\left|\frac{m}{n} - 0,8\right| \leq 0,02\right) = 2\Phi\left(0,02\sqrt{\frac{n}{0,5 \cdot 0,5}}\right) = 0,7698.$$

Звідси

$$\Phi(0,04\sqrt{n}) = 0,3843.$$

Визначивши за дод. 1 аргумент інтегральної функції Лапласа, при якому ця функція дорівнює 0,3843, отримаємо рівняння

$$0,04\sqrt{n} = 1,2.$$

Отже,

$$n = \left(\frac{1,2}{0,04}\right)^2 = 900.\quad \square$$

Задачі до розділу 5

Задача 1. Оглядову лекцію мають прослухати 100 студентів. Ймовірність бути присутнім на цій лекції для кожного студента дорівнює 0,7. Знайти ймовірність того, що на лекцію прийде більше половини студентів.

Відповідь. $P_{100}(m > 50) \approx 0,99998$.

Задача 2. Ймовірність вчасної реалізації зі складу однієї пари взуття дорівнює 0,8. Знайти ймовірність того, що вчасно буде реалізовано не менше 75 пар, якщо на склад завезено 100 пар взуття. Знайти найімовірнішу кількість вчасно реалізованих пар взуття.

Відповідь. $P_{100}(m \geq 75) \approx 0,8943$; $m_0 = 80$.

Задача 3. Зі статистичних даних відомо, що ймовірність захворіти грипом під час епідемії для кожної особи дорівнює 0,1. Яка ймовірність того, що із 100 перевірених осіб хворими виявляться від 20 до 50 осіб?

Відповідь. $P_{100}(20 \leq m \leq 50) \approx 0,00043$.

Задача 4. У магазин зайшло вісім покупців. Ймовірність того, що будь-який з них не піде з магазину без покупки, дорівнює 0,4.

а) Знайти ймовірність того, що троє з них щось куплять.

б) Яка ймовірність того, що жоден з них нічого не купить?

Відповідь. а) $P_8(3) \approx 0,2787$; б) $P_8(0) \approx 0,0168$.

Задача 5. Знайти ймовірність того, що серед 100 перехожих буде не більше 40 брюнетів, якщо близько 30 % населення — брюнети.

Відповідь. $P_{100}(m \leq 40) \approx 0,9855$.

Задача 6. У процесі виробництва ймовірність дефектів у кожній партії продукції становить 0,1. Яка ймовірність того, що з десяти партій дефекти матимуть менше двох партій?

Відповідь. $P_{10}(m < 2) \approx 0,7361$.

Задача 7. Знайти ймовірність того, що серед 1000 новонароджених буде 480 дівчаток, якщо ймовірність народження хлопчика дорівнює 0,515.

Відповідь. $P_{1000}(480) \approx 0,024$.

Задача 8. Припустимо, що ймовірності народитися у будь-який з днів року однакові. Знайти ймовірність того, що серед 500 учнів школи:

- а) троє народилися 8-го березня;
- б) жоден не народився 1-го січня.

Відповідь. а) $P_{500}(3) \approx 0,1089$; б) $P_{500}(0) \approx 0,2541$.

Задача 9. Імовірність влучення в літак з гвинтівки при кожному пострілі дорівнює 0,001. Здійснюється 3000 пострілів. Знайти ймовірність того, що буде хоча б одне влучення.

Відповідь. $P_{3000}(m \geq 1) \approx 0,9502$.

Задача 10. Серед п'яти студентів проводиться психологічний тест на визначення типу характеру людини. Імовірність того, що для кожної людини буде правильно визначено тип характеру за результатами тестування, дорівнює 0,9. Знайти ймовірність того, що лише для трьох протестованих студентів буде правильно визначено тип характеру.

Відповідь. $P_5(3) = 0,0729$.

Задача 11. В одержаній партії текстильних виробів 0,6% браку. Яка ймовірність при випадковому відборі 1000 виробів виявити:

- а) шість бракованих виробів;
- б) хоча б один бракований виріб?

Відповідь. а) $P_{1000}(6) \approx 0,1606$; б) $P_{1000}(m \geq 1) \approx 0,9975$.

ЧАСТИНА II

Випадкові величини

Розділ 6. Дискретні випадкові величини

Випадковою величиною, пов'язаною з даним імовірнісним експериментом, називається величина, яка при кожному проведенні цього експерименту набуває певного числового значення, причому заздалегідь невідомо, якого саме. Випадкові величини позначатимемо грецькими літерами ξ, ψ, η, \dots (див. дод. 10) або великими літерами латинського алфавіту (див. дод. 9).

Множина називається *зліченною*, якщо між її елементами та елементами множини натуральних чисел можна встановити взаємно однозначну відповідність. Іншими словами, зліченна множина — це така нескінченна множина, елементи якої можна пронумерувати за допомогою множини натуральних чисел.

Випадкова величина називається *дискретною*, якщо множина її можливих значень є скінченною або зліченною множиною.

Законом розподілу випадкової величини називається довільна відповідність, що встановлює зв'язок між можливими значеннями випадкової величини та відповідними їм імовірностями.

У випадку дискретної випадкової величини закон розподілу найзручніше описувати за допомогою ряду розподілу — таблиці, де наведено всі можливі значення цієї випадкової величини та відповідні їм імовірності:

ξ	x_1	x_2	\dots	x_n
\mathbf{P}	p_1	p_2	\dots	p_n

причому $p_i = \mathbf{P}\{\xi = x_i\}$, $\sum_{i=1}^n p_i = 1$.

Ламана з вершинами в точках (x_i, p_i) називається *многокутником розподілу* (рис. 5).

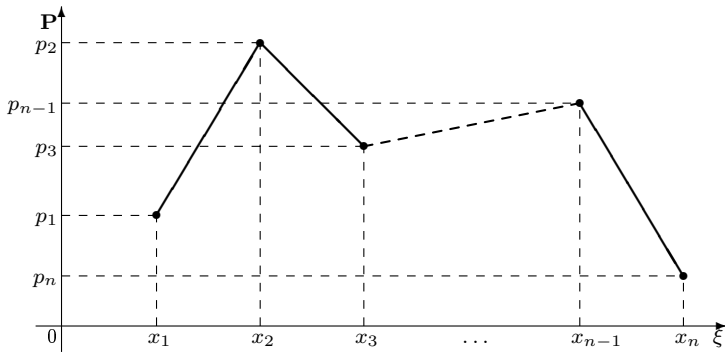


Рис. 5

Інший спосіб задання розподілу випадкової величини — зазначення її функції розподілу.

Функцією розподілу випадкової величини ξ називається функція $F(x)$, значення якої дорівнює ймовірності того, що випадкова величина ξ набуде значення, яке менше x , тобто

$$F(x) = \mathbf{P}\{\xi < x\}.$$

Функція розподілу має такі властивості.

Властивість 1. $\lim_{x \rightarrow +\infty} F(x) = 1$, $\lim_{x \rightarrow -\infty} F(x) = 0$.

Властивість 2. $F(x)$ — неспадна функція, тобто

$$x_1 < x_2 \quad \Rightarrow \quad F(x_1) \leq F(x_2).$$

Властивість 3. $F(x)$ — неперервна зліва функція, тобто для довільного $x \in \mathbb{R}$

$$\lim_{\varepsilon \rightarrow 0^+} F(x - \varepsilon) = F(x).$$

Якщо випадкова величина ξ дискретна, то

$$F(x) = \sum_{i: x_i < x} p_i, \quad (1)$$

де підсумовування виконується по всіх індексах i , для яких виконується умова $x_i < x$. Іншими словами, за рядом розподілу легко відновлюється функція розподілу. І навпаки, за функцією розподілу легко відновлюється ряд розподілу дискретної випадкової величини: *множина значень випадкової величини збігається з множиною точок $\{x_i\}$ розриву $F(x)$, а відповідні ймовірності p_i дорівнюють величинам стрибків функції розподілу в точках x_i* , тобто

$$p_i = \mathbf{P}\{\xi = x_i\} = \lim_{\varepsilon \rightarrow 0} F(x + \varepsilon) - F(x), \quad \varepsilon > 0.$$

Зауважимо, що з одним і тим самим експериментом може бути пов'язано багато різних випадкових величин.

Приклад 1. Експеримент полягає у підкиданні симетричного однорідного грального кубика.

Нехай випадкова величина ξ дорівнює кількості очок, що випали при одному підкиданні, а випадкова величина ψ — кількості шісток при одному підкиданні кубика.

Побудувати ряд розподілу, многокутник розподілу, функцію розподілу випадкової величини:

- а) ξ ;
- б) ψ .

Частина II. Випадкові величини

Розв'язання. а) Випадкова величина ξ може набувати значень 1, 2, 3, 4, 5 і 6 з однаковими ймовірностями $p = \frac{1}{6}$. Отже, ряд розподілу випадкової величини ξ має такий вигляд:

ξ	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

а многокутник розподілу зображено на рис. 6.

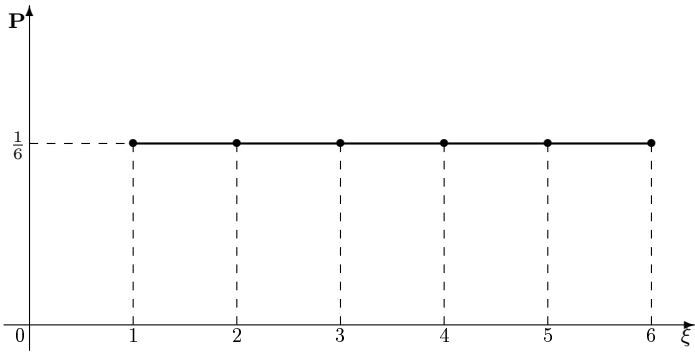


Рис. 6

Скориставшись формулою (1), отримаємо

$$F(x) = \begin{cases} 0, & x \leq 1; \\ \frac{1}{6}, & 1 < x \leq 2; \\ \frac{2}{6}, & 2 < x \leq 3; \\ \frac{3}{6}, & 3 < x \leq 4; \\ \frac{4}{6}, & 4 < x \leq 5; \\ \frac{5}{6}, & 5 < x \leq 6; \\ 1, & x > 6. \end{cases}$$

Графік функції розподілу $F(x)$ зображено на рис. 7.

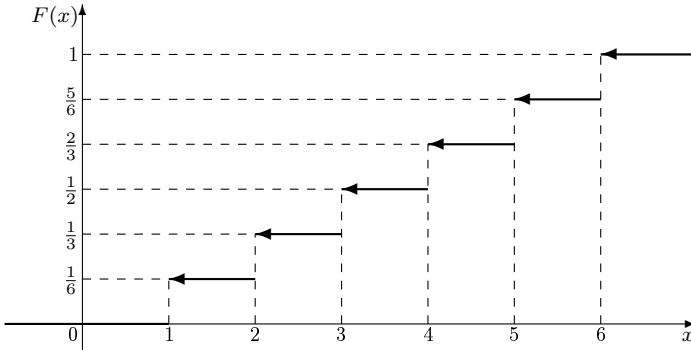


Рис. 7

б) Оскільки при одному підкиданні грального кубика шістька може або зовсім не випасти, або випасти лише один раз, випадкова величина ψ може набувати двох значень: 0 і 1. Неважко переконалися, що

$$\mathbf{P}\{\psi = 0\} = \frac{5}{6}; \quad \mathbf{P}\{\psi = 1\} = \frac{1}{6}.$$

Ряд розподілу випадкової величини ψ має такий вигляд:

ψ	0	1
\mathbf{P}	$\frac{5}{6}$	$\frac{1}{6}$

Многокутник розподілу випадкової величини ψ зображено на рис. 8.

Функція розподілу випадкової величини ψ має вигляд

$$F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{5}{6}, & 0 < x \leq 1; \\ 1, & x > 1. \end{cases}$$

Графік функції розподілу зображено на рис. 9. □

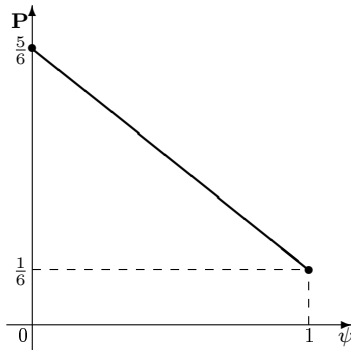


Рис. 8

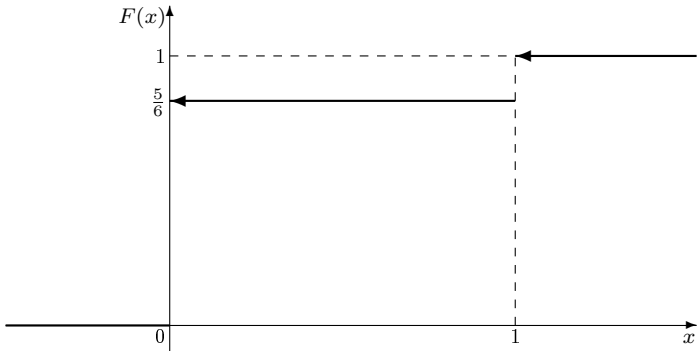


Рис. 9

Найважливішими є такі типи дискретних розподілів: рівномірний, біноміальний, показниковий, геометричний, гіпергеометричний.

Біноміальний розподіл (розподіл Бернуллі) — це розподіл випадкової величини, яка набуває значень $i \in \{0, 1, \dots, n\}$ з імовірностями

$$p_i = C_n^i p^i (1-p)^{n-i},$$

де $p \in (0; 1)$. Біноміальний розподіл має випадкова величина, яка дорівнює кількості «успіхів» у серії з n незалежних випробувань, у кожному з яких «успіх» відбувається з імовірністю p .

Рівномірний розподіл — це розподіл випадкової величини, яка набуває n різних значень з однаковими ймовірностями

$$p_i = \frac{1}{n}.$$

Показниковий розподіл (розподіл Пуассона) — це розподіл випадкової величини, яка набуває значень $k \in \{0, 1, 2, \dots, n\}$ з ймовірностями

$$P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

де $\lambda = np$, $n \rightarrow \infty$, $p \rightarrow 0$.

Геометричний розподіл — це розподіл випадкової величини, яка набуває значень $k \in \{1, 2, \dots\}$ з ймовірностями

$$p_k = pq^{k-1},$$

де $q = 1 - p$. Геометричний розподіл має випадкова величина, яка дорівнює кількості спроб до першого «успіху» в серії незалежних випробувань, у кожному з яких ймовірність успіху дорівнює p .

Гіпергеометричний розподіл — це розподіл випадкової величини, яка набуває значень $k \in \{0, 1, \dots, t\}$ з ймовірностями

$$p_k = \frac{C_n^k C_{N-n}^{m-k}}{C_N^m},$$

де $n \geq t$, $N \geq n$.

Гіпергеометричний розподіл має така випадкова величина. Нехай у ящику міститься N однакових за фізичними властивостями кульок, серед яких S білих і $(N - S)$ чорних. З ящика навмання виймається t кульок. Тоді випадкова величина ξ , яка дорівнює кількості білих кульок серед t вийнятих, має гіпергеометричний розподіл (рис. 10).

Приклад 2. Нехай випадкова величина ξ дорівнює кількості номерів, угаданих гравцем у лотереї «6 із 39». Побудувати ряд розподілу випадкової величини ξ . Знайти значення функції розподілу в точці $x = 3$.

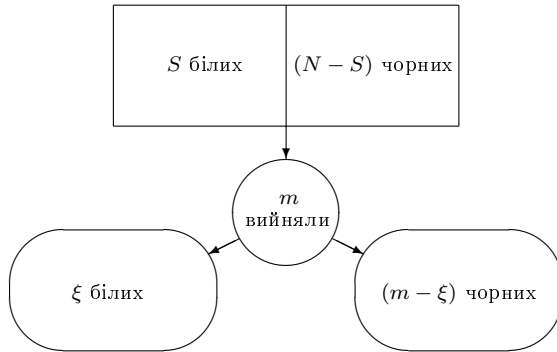


Рис. 10

Розв'язання. Випадкова величина ξ має гіпергеометричний розподіл. Справді, процес розіграшу можна змоделювати так: у лототроні (ящику) міститься $N = 39$ однакових за фізичними властивостями кульок, серед яких $S = 6$ «білих» (так можна називати кульки, які в результаті розіграшу випадуть) і $N - S = 33$ «чорних». Вважати-мемо, що кульки добре перемішані. Тому можна вважати, що з 39 кульок навмання виймається $m = 6$ шт. Тоді ξ — це кількість «білих» кульок серед шести вибраних (це номери, які загадані гравцем).

Очевидно, випадкова величина ξ набуває значень від 0 до 6, причому

$$\mathbf{P}\{\xi = i\} = \frac{C_6^i \cdot C_{33}^{6-i}}{C_{39}^6}.$$

Виконавши підрахунки для $i = 0, 1, 2, 3, 4, 5, 6$, отримаємо:

ξ	0	1	2	3	4	5	6
\mathbf{P}	$\frac{1107568}{3262623}$	$\frac{1424016}{3262623}$	$\frac{613800}{3262623}$	$\frac{109120}{3262623}$	$\frac{7920}{3262623}$	$\frac{198}{3262623}$	$\frac{1}{3262623}$

Значення функції розподілу $F(x)$ випадкової величини ξ у точці $x = 3$

$$\begin{aligned} F(3) &= \mathbf{P}\{\xi < 3\} = \mathbf{P}\{\xi = 0 \text{ або } \xi = 1 \text{ або } \xi = 2\} = \\ &= \mathbf{P}\{\xi = 0\} + \mathbf{P}\{\xi = 1\} + \mathbf{P}\{\xi = 2\}. \end{aligned}$$

Іншими словами, імовірність нічого не виграти, заповнивши один лотерейний квиток, дорівнює

$$\frac{1107568}{3262623} + \frac{1424016}{3262623} + \frac{613800}{3262623} = \frac{3145384}{3262623} \approx 0,9641. \quad \square$$

Задачі до розділу 6

Задача 1. Нехай ξ — випадкова величина, яка дорівнює кількості хлопчиків у сім'ї з трьома дітьми. Вважаючи народження хлопчика й дівчинки рівноймовірними подіями, знайти ряд розподілу і ймовірність того, що в сім'ї буде більше хлопчиків, аніж дівчаток.

Відповідь.
$$\begin{array}{c|c|c|c} \xi & 0 & 1 & 2 & 3 \\ \hline \mathbf{P} & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{array}; \mathbf{P}(\text{«хлопчиків більше»}) = \frac{1}{2}.$$

Задача 2. Нехай ξ — випадкова величина, яка дорівнює кількості випадань грані «6» при двох підкиданнях кубика. Знайти ряд розподілу, функцію розподілу випадкової величини ξ та її графік.

Відповідь.
$$\begin{array}{c|c|c|c} \xi & 0 & 1 & 2 \\ \hline \mathbf{P} & \frac{25}{36} & \frac{5}{18} & \frac{1}{36} \end{array}; F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{25}{36}, & 0 < x \leq 1; \\ \frac{35}{36}, & 1 < x \leq 2; \\ 1, & x > 2, \end{cases}$$

графік функції розподілу зображено на рис. 11.

Задача 3. Стрілець робить по одному пострілу по чотирьох мішенях. Випадкова величина ξ — кількість влучень. Вважаючи, що ймовірність влучення при одному пострілі дорівнює 0,7, знайти ряд розподілу випадкової величини ξ та ймовірність того, що влучень буде більше, ніж промахів.

Відповідь.
$$\begin{array}{c|c|c|c|c|c} \xi & 0 & 1 & 2 & 3 & 4 \\ \hline \mathbf{P} & 0,0081 & 0,0756 & 0,2646 & 0,4116 & 0,2401 \end{array}; \mathbf{P}(\text{«влучень більше»}) = 0,6517.$$

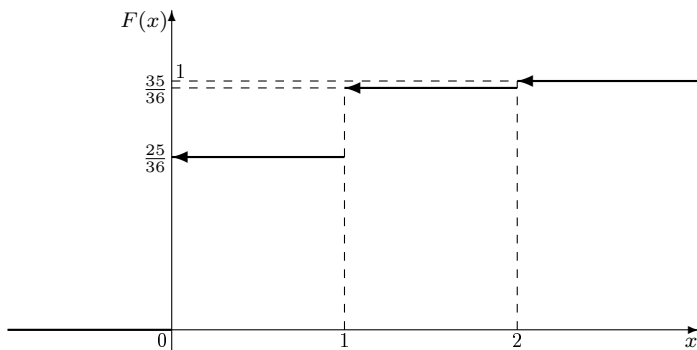


Рис. 11

Задача 4. Два стрільці роблять по одному незалежному пострілу в одну мішень. Імовірність влучення для першого стрільця дорівнює 0,8, а для другого — 0,5. Знайти ряд розподілу випадкової величини ξ — кількості влучень у мішень, а також імовірність того, що кількість влучень дорівнюватиме кількості промахів.

Відповідь. $\begin{array}{c|c|c|c} \xi & 0 & 1 & 2 \\ \hline \mathbf{P} & 0,1 & 0,5 & 0,4 \end{array}$; \mathbf{P} («кількість влучень дорівнює кількості промахів») = 0,5.

Задача 5. Студент підкидає монету до першого випадання аверса. Випадкова величина ξ — кількість підкидань. Знайти розподіл випадкової величини ξ . Яка ймовірність того, що студент змушений буде виконати більше 10 підкидань?

Відповідь. $\mathbf{P}(\xi = k) = \frac{1}{2^k}$, $k = 1, 2, \dots$; $\mathbf{P}(\xi > 10) = \frac{1}{1024}$.

Задача 6. Гральний кубик підкидається до першої появи п'ятірки. Випадкова величина ξ — кількість підкидань кубика. Знайти ряд розподілу випадкової величини ξ і найімовірнішу кількість підкидань.

Відповідь. $\mathbf{P}(\xi = k) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{k-1}$, $k = 1, 2, \dots$; $m_0 = 1$.

Задача 7. Студент вивчив 10 білетів з 25 і намагається скласти іспит. Випадкова величина ξ — кількість спроб для успішного складання. Знайти ряд розподілу випадкової величини ξ та ймовірність того, що студента відрахують за академзаборгованість (іспит дозволяється перескладати двічі).

Відповідь. $\mathbf{P}(\xi = k) = \frac{2}{5} \cdot \left(\frac{3}{5}\right)^{k-1}$, $k = 1, 2, \dots$; $\mathbf{P}(\xi > 3) = 0,216$.

Задача 8. Гравець в казино ставить «на червоне» доти, доки не виграє. Випадкова величина ξ — кількість спроб. Знайти ряд розподілу випадкової величини ξ , найімовірнішу кількість спроб і ту кількість спроб x , при якій виконується умова $\mathbf{P}(\xi \leq x) \geq 0,999$.

Вказівка. Вважати, що рулетка має 37 полів (від 0 до 36), серед яких 18 червоних.

Відповідь. $\mathbf{P}(\xi = k) = \frac{18}{37} \cdot \left(\frac{19}{37}\right)^{k-1}$, $k = 1, 2, \dots$; $m_0 = 1$; $x \geq 11$.

Задача 9. Студент знає 20 питань з 25. Навмання виймається 5 питань. Випадкова величина ξ — кількість питань, на які студент знає відповідь. Знайти ряд розподілу випадкової величини ξ та ймовірність складання іспиту (іспит вважається складеним, якщо студент дає відповідь більше ніж на половину питань).

Відповідь.

ξ	0	1	2	3	4	5
\mathbf{P}	$\frac{1}{3125}$	$\frac{4}{625}$	$\frac{32}{625}$	$\frac{128}{625}$	$\frac{256}{625}$	$\frac{1024}{3125}$

; $\mathbf{P}(\text{«складе»}) = \frac{2944}{3125} = 0,94208$.

Задача 10. У партії з 10 деталей 2 браковані. Навмання виймається 3 деталі. Нехай ξ — випадкова величина, яка дорівнює кількості бракованих деталей серед вибраних. Знайти ряд розподілу і функцію розподілу випадкової величини ξ .

Відповідь.
$$\begin{array}{c|c|c|c} \xi & 0 & 1 & 2 \\ \hline \mathbf{P} & \frac{7}{15} & \frac{7}{15} & \frac{1}{15} \end{array}; F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{7}{15}, & 0 < x \leq 1; \\ \frac{14}{15}, & 1 < x \leq 2; \\ 1, & x > 2. \end{cases}$$

Задача 11. Нехай ξ — випадкова величина, яка дорівнює кількості вгаданих номерів у лотереї «5 із 36». Знайти ряд розподілу випадкової величини ξ і ймовірність хоча б що-небудь виграти (виграші починаються з трьох угаданих номерів).

Відповідь.
$$\begin{array}{c|c|c|c|c|c|c} \xi & 0 & 1 & 2 & 3 & 4 & 5 \\ \hline \mathbf{P} & 0,45 & 0,42 & 0,12 & 0,01 & 0,0004 & 0,000003 \end{array};$$

\mathbf{P} («виграш») $\approx 0,013$.

Задача 12. Після тестування виявилось, що серед 15 студентів — один меланхолік, 5 флегматиків, 6 сангвініків і 3 холерики. З цієї групи навмання вибирають чотирьох студентів. Нехай ξ — випадкова величина, яка дорівнює кількості сангвініків серед вибраних студентів. Знайти ряд розподілу даної випадкової величини.

Відповідь.
$$\begin{array}{c|c|c|c|c|c} \xi & 0 & 1 & 2 & 3 & 4 \\ \hline \mathbf{P} & 0,47 & 0,41 & 0,11 & 0,01 & 0,0003 \end{array}.$$

Задача 13. Магазин отримав 10000 пляшок мінеральної води. Імовірність того, що під час перевезення пляшка виявиться розбитою, дорівнює 0,0004. Знайти ряд розподілу випадкової величини ξ , що характеризує кількість розбитих пляшок, і ймовірність того, що розбитих пляшок буде більше трьох.

Відповідь. $\mathbf{P}(\xi = k) = \frac{4^k}{k!}e^{-4}$, $k = 1, 2, \dots$; $\mathbf{P}(\xi \geq 4) \approx 0,5665$.

Задача 14. Підручник видано тиражем 10000 екземплярів. Імовірність того, що зброшурований екземпляр виявиться неякісним, дорівнює 0,001. Знайти ряд розподілу випадкової величини ξ , що характеризує кількість неякісно зброшурованих екземплярів, і ймовірність того, що бракованих екземплярів буде більше двох.

Відповідь. $P(\xi = k) = \frac{10^k}{k!}e^{-10}$, $k = 1, 2, \dots$; $P(\xi > 2) \approx 0,9972$.

Задача 15. При наборі тексту друкарка набирає неправильний символ з імовірністю 0,001. Знайти ймовірність того, що при наборі сторінки тексту (3000 символів) буде зроблено більше трьох помилок.

Відповідь. $P_{3000}(\xi > 3) \approx 0,3528$.

Задача 16. Імовірність влучення в муху на льоту з рогатки дорівнює 0,001. Виконується 2000 пострілів. Випадкова величина ξ дорівнює кількості вбитих мух. Побудувати ряд розподілу випадкової величини ξ і знайти ймовірність того, що буде вбито менше чотирьох мух.

Відповідь. $P(\xi = k) = \frac{2^k}{k!}e^{-2}$, $k = 1, 2, \dots$; $P(\xi < 4) \approx 0,8571$.

Задача 17. Автоматична телефонна станція обслуговує 10000 телефонних номерів. Імовірність того, що протягом 1 хв на АТС надійде виклик від абонента, дорівнює 0,0004. Знайти ряд розподілу випадкової величини X , яка дорівнює кількості викликів, які надійшли на АТС протягом 1 хв, та ймовірність того, що за цей час надійде хоча б один виклик.

Відповідь. $P(\xi = k) = \frac{4^k}{k!}e^{-4}$, $k = 1, 2, \dots$; $P(\xi \geq 1) \approx 0,9817$.

Розділ 7. Неперервні випадкові величини

Випадкова величина називається *неперервною*, якщо її функція розподілу є неперервною функцією.

Якщо ξ — неперервна випадкова величина, то ймовірність того, що ξ набуде деякого свого конкретного значення x , дорівнює 0 ($\mathbf{P}(\xi = x) = 0, x \in \mathbb{R}$). Тому на відміну від дискретної випадкової величини розподіл неперервної випадкової величини неможливо задати, зазначивши значення, яких вона набуває, та відповідні їм ймовірності. Зрозуміло також, що множина можливих значень неперервної випадкової величини незліченна.

Найпростішим і найпоширенішим способом задання розподілу неперервної випадкової величини є задання її функції розподілу $F_\xi(x)$. Тоді для будь-якого проміжку $(a; b]$ можна обчислити ймовірність

$$\mathbf{P}(\xi \in (a; b]) = F_\xi(b) - F_\xi(a).$$

Неперервні випадкові величини поділяються на два класи: *сингулярні* та *абсолютно неперервні*.

Випадкова величина ξ називається *абсолютно неперервною*, якщо існує така функція $f_\xi(x)$, для якої виконується рівність

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt.$$

При цьому функція $f_\xi(x)$ називається *щільністю розподілу випадкової величини ξ* (або *диференціальною функцією розподілу*).

Очевидно, що:

- 1) $f_\xi(x) \geq 0$;
- 2) $F'_\xi(x) = f_\xi(x)$ у кожній точці неперервності функції $f_\xi(x)$;
- 3) $\int_{-\infty}^{+\infty} f_\xi(x) dx = 1$;

4) $P(\xi \in (a; b]) = F_\xi(b) - F_\xi(a) = \int_a^b f_\xi(x) dx$. Іншими словами, імовірність того, що випадкова величина ξ набуде значення з проміжку $(a; b]$, чисельно дорівнює площі криволінійної трапеції, обмеженої графіком функції $y = f_\xi(x)$, віссю абсцис і прямими $x = a$, $x = b$.

Сингулярні випадкові величини, на відміну від абсолютно неперервних, неможливо задати за допомогою щільності розподілу.

Найважливішими є такі типи абсолютно неперервних розподілів: рівномірний, нормальний, експоненціальний.

Розподіл випадкової величини ξ називається *рівномірним* на відрізьку $[a; b]$, якщо його щільність має вигляд

$$f_\xi(x) = \begin{cases} 0, & x \notin [a; b]; \\ \frac{1}{b-a}, & x \in [a; b]. \end{cases}$$

Функція розподілу рівномірно розподіленої випадкової величини має вигляд

$$F_\xi(x) = \begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a < x \leq b; \\ 1, & x > b. \end{cases}$$

Графіки щільності та функції розподілу рівномірно розподіленої випадкової величини зображено відповідно на рис. 12 і 13.

Розподіл випадкової величини ξ називається *нормальним* з параметрами a , σ , якщо його щільність має вигляд

$$f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Функція $y = f_\xi(x)$ швидко спадає при $x \rightarrow \pm\infty$. Площа фігури, обмеженої графіком функції $y = f_\xi(x)$ і віссю Ox , дорівнює одиниці. Площі криволінійних трапецій над інтервалами

$$[a - \sigma; a + \sigma], \quad [a - 2\sigma; a + 2\sigma], \quad [a - 3\sigma; a + 3\sigma]$$

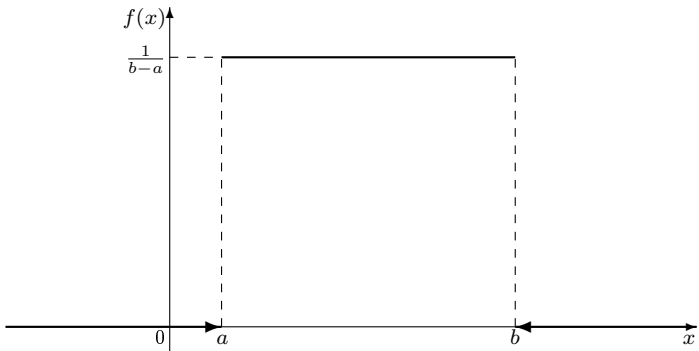


Рис. 12

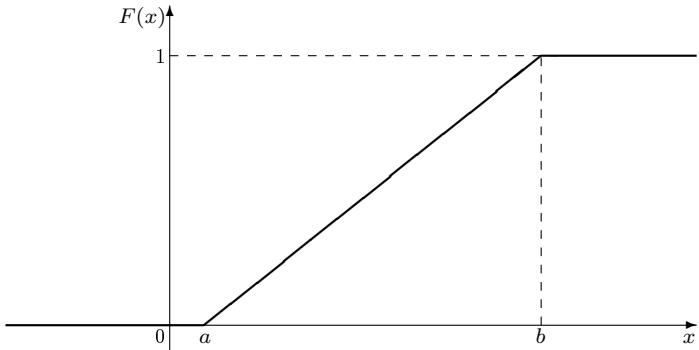


Рис. 13

дорівнюють відповідно

$$0,6827, \quad 0,9545, \quad 0,9973.$$

Отже, імовірність того, що нормально розподілена випадкова величина набуде значення з відрізка

$$[a - 3\sigma; a + 3\sigma],$$

дорівнює 0,9973. У цьому разі кажуть, що нормально розподілена випадкова величина ξ потрапляє в «трисигмовий» проміжок

$[a - 3\sigma; a + 3\sigma]$ з *практичною достовірністю*. На рис. 14 зображено графік функції щільності нормального розподілу з математичним сподіванням 0.

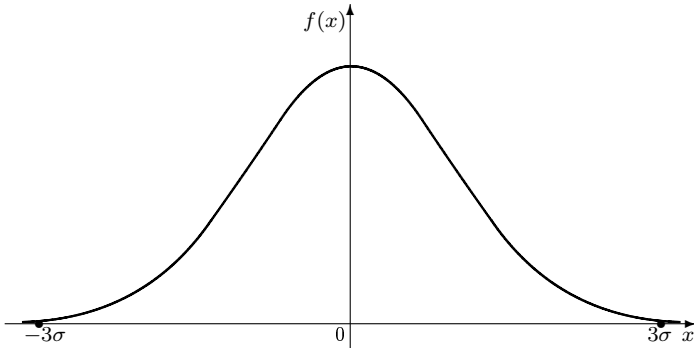


Рис. 14

Функція розподілу нормально розподіленої випадкової величини має вигляд

$$F_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt.$$

Ця функція не є елементарною. Тому її значення табульовані. Для зручності розглядають функцію Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

Функції $F_{\xi}(x)$ і $\Phi(x)$ пов'язані співвідношенням

$$F_{\xi}(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right).$$

Таблиці значень функції $\Phi(x)$ наведено в дод. 1. Зазначимо, що $\Phi(x)$ має такі властивості.

Властивість 1. $\Phi(-x) = -\Phi(x)$.

Властивість 2. При $x > 5$ можна вважати, що $\Phi(x) \approx 0,5$.

Імовірність того, що нормально розподілена випадкова величина набуде значення, яке належить відрізку $[\alpha; \beta]$,

$$\mathbf{P}(\xi \in [\alpha; \beta]) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

Імовірність того, що нормально розподілена випадкова величина набуде значень з проміжку $[a - \delta; a + \delta]$,

$$\mathbf{P}(|\xi - a| \leq \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

Нормальний розподіл відіграє надзвичайно важливу роль у теорії ймовірностей та її застосуваннях. Ця роль значною мірою пояснюється центральною граничною теоремою. Якщо відомо, що досліджувана випадкова величина є сумою великої кількості випадкових доданків, кожен з яких невеликою мірою впливає на всю суму, то можна вважати, що результуюча випадкова величина має нормальний розподіл. Наприклад, випадкова похибка вимірювання має, як правило, нормальний розподіл.

Розподіл випадкової величини ξ називається *експоненціальним (показниковим)* з параметром $\lambda > 0$, якщо його щільність має вигляд

$$f_{\xi}(x) = \begin{cases} 0, & x < 0; \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

Графік функції щільності експоненціального закону зображено на рис. 15.

Функція розподілу випадкової величини з експоненціальним розподілом має вигляд

$$F_{\xi}(x) = \begin{cases} 0, & x < 0; \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$$

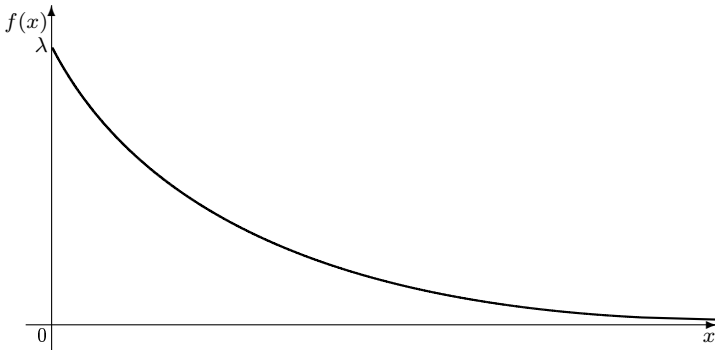


Рис. 15

Якщо $\alpha > 0$ і $\beta > 0$, то

$$\mathbf{P}(\xi \in [\alpha; \beta]) = e^{-\lambda\alpha} - e^{-\lambda\beta}.$$

Якщо $\alpha \leq 0$ і $\beta > 0$, то

$$\mathbf{P}(\xi \in [\alpha; \beta]) = \mathbf{P}(\xi \in [0; \beta]) = 1 - e^{-\lambda\beta}.$$

Показниковий розподіл часто зустрічається в теорії масового обслуговування (зокрема, для опису роботи телефонних станцій).

Приклад 1. Випадкову величину ξ задано функцією розподілу

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \sqrt{x}, & 0 < x \leq \frac{1}{4}; \\ 8x^2, & \frac{1}{4} < x \leq \frac{\sqrt{2}}{4}; \\ 1, & x > \frac{\sqrt{2}}{4}. \end{cases}$$

Переконайтеся, що величина ξ має щільність розподілу ймовірностей, і знайдіть її.

Розв'язання. У точках

$$x = 0; \frac{1}{4}; \frac{\sqrt{2}}{4}$$

правосторонні та лівосторонні границі функції $F_\xi(x)$ збігаються, з чого випливає, що $F_\xi(x)$ неперервна всюди. Похідна $F'_\xi(x)$ неперервна всюди, за винятком точок $x = 0; \frac{1}{4}; \frac{\sqrt{2}}{4}$. Отже, щільність розподілу ймовірностей $f_\xi(x)$ існує. Обчислимо односторонні границі $F'_\xi(x)$ у точках $x = 0; \frac{1}{4}; \frac{\sqrt{2}}{4}$:

$$\begin{aligned} F'_\xi(0-0) &= 0, & F'_\xi(0+0) &= \infty, \\ F'_\xi\left(\frac{1}{4}-0\right) &= 1, & F'_\xi\left(\frac{1}{4}+0\right) &= 4, \\ F'_\xi\left(\frac{\sqrt{2}}{4}-0\right) &= 4\sqrt{2}, & F'_\xi\left(\frac{\sqrt{2}}{4}+0\right) &= 0. \end{aligned}$$

Отже, точки $0; \frac{1}{4}; \frac{\sqrt{2}}{4}$ є точками розриву для $F'_\xi(x)$.

Таким чином,

$$f_\xi(x) = \begin{cases} 0, & x \leq 0; \\ \frac{1}{2\sqrt{x}}, & 0 < x \leq \frac{1}{4}; \\ 16x, & \frac{1}{4} < x \leq \frac{\sqrt{2}}{4}; \\ 0, & x > \frac{\sqrt{2}}{4} \end{cases}$$

(точки $0; \frac{1}{4}; \frac{\sqrt{2}}{4}$ можна включити в будь-який сусідній проміжок). □

Приклад 2. Дано функцію

$$f(x) = \begin{cases} 0, & x \leq 0; \\ ae^{-\alpha x}, & x > 0, \end{cases} \quad \alpha > 0.$$

При якому значенні сталої a функція $f(x)$ є щільністю розподілу ймовірностей деякої випадкової величини ξ ? Знайти функцію розподілу $F_\xi(x)$ величини ξ . Обчислити ймовірність того, що випадкова величина ξ потрапить у проміжок $[0; 1]$ двома способами: за допомогою щільності розподілу ймовірностей $f_\xi(x)$ і за допомогою функції розподілу $F_\xi(x)$.

Розв'язання. Передусім a має бути невід'ємним: $a \geq 0$. Для знаходження значення a запишемо умову

$$\int_{-\infty}^{+\infty} f_\xi(x) dx = 1$$

і перетворимо її:

$$\int_{-\infty}^{\infty} f_\xi(x) dx = a \int_0^{\infty} e^{-\alpha x} dx = -\frac{a}{\alpha} e^{-\alpha x} \Big|_0^{\infty} = \frac{a}{\alpha} = 1.$$

Отже, $a = \alpha$ і функція $f_\xi(x)$ має вигляд

$$f(x) = \begin{cases} 0, & x \leq 0; \\ \alpha e^{-\alpha x}, & x > 0, \end{cases} \quad \alpha > 0. \quad (2)$$

Знайдемо функцію розподілу $F_\xi(x)$ за формулою

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt.$$

Якщо $x \leq 0$, то

$$F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt = \int_{-\infty}^x 0 dt = 0,$$

якщо $x > 0$, то

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(t)dt(x) = \int_{-\infty}^0 0dt + \int_0^x \alpha e^{-\alpha t} dt = 1 - e^{-\alpha x}.$$

Отже,

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ 1 - \alpha e^{-\alpha x}, & x > 0, \end{cases} \quad \alpha > 0. \quad (3)$$

За формулою

$$\mathbf{P}(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f_{\xi}(x)dx$$

обчислимо ймовірність $\mathbf{P}(0 \leq x \leq 1)$:

$$\mathbf{P}(0 \leq x \leq 1) = \int_0^1 f_{\xi}(x)dx = \int_0^1 \alpha e^{-\alpha x} dx = 1 - \frac{1}{e^{\alpha}}.$$

Формула

$$\mathbf{P}(x_1 \leq x \leq x_2) = F_{\xi}(x_2) - F_{\xi}(x_1)$$

дає той самий результат:

$$\mathbf{P}(0 \leq x \leq 1) = F_{\xi}(1) - F_{\xi}(0) = 1 - \frac{1}{e^{\alpha}}.$$

Випадкові величини зі щільністю (2) чи з функцією розподілу (3) досить часто зустрічаються на практиці (показниковий закон розподілу). \square

Приклад 3. Випадкова величина ξ має нормальний розподіл з параметрами $a = 7$ і $\sigma = 2$. Порівняти ймовірності потрапляння випадкової величини ξ у проміжки $(3; 7]$ і $(-100; 1]$. Зазначити інтервал, у який випадкова величина ξ потрапляє з практичною достовірністю.

Розв'язання. Якщо випадкова величина ξ має нормальний розподіл, то

$$\mathbf{P}(\xi \in (\alpha; \beta]) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right).$$

Тому

$$\mathbf{P}(\xi \in (\alpha; \beta]) = \Phi\left(\frac{7 - 7}{2}\right) - \Phi\left(\frac{3 - 7}{2}\right) = \Phi(0) - \Phi(-2).$$

За таблицею значень функції $\Phi(x)$ (дод. 1) знаходимо:

$$\Phi(0) = 0; \quad \Phi(-2) = -\Phi(2) \approx -0,4772.$$

Тому

$$\mathbf{P}(\xi \in (3; 7]) \approx 0,4772.$$

Аналогічно,

$$\begin{aligned} \mathbf{P}(\xi \in (-100; 1]) &= \Phi\left(\frac{1 - 7}{2}\right) - \Phi\left(\frac{-100 - 7}{2}\right) = \\ &= \Phi(-3) - \Phi(-53,5) = -\Phi(3) + \Phi(53,5). \end{aligned}$$

За дод. 1 отримуємо:

$$\Phi(3) \approx 0,49865.$$

Значення $\Phi(53,5)$ у дод. 1 знайти не можна. Оскільки $53,5 > 5$, то з великою точністю можна вважати, що

$$\Phi(53,5) \approx 0,5.$$

Отже,

$$\mathbf{P}(\xi \in (-100; 1]) \approx -0,49865 + 0,5 = 0,00135.$$

Нормально розподілена випадкова величина з практичною достовірністю потрапляє в так званий трисигмовий інтервал

$$(a - 3\sigma; a + 3\sigma].$$

Отже,

$$\mathbf{P}(\xi \in (1; 12]) = 0,9973.$$

Пропонуємо читачеві самостійно пересвідчитись у правильності останньої рівності. \square

Задачі до розділу 7

Задача 1. Випадкова величина X має функцію розподілу

$$F(x) = \begin{cases} 0, & x \leq -4, \\ \frac{a(x+4)^2}{16}, & -4 < x \leq 0, \\ 1, & x > 0. \end{cases}$$

Знайти параметр a , аналітичний вираз для щільності, імовірність потрапляння випадкової величини X в інтервал $(-2; 5)$.

Відповідь. $a = 1$, $f(x) = \begin{cases} \frac{x+4}{8}, & x \in (-4; 0); \\ 0, & x \notin (-4; 0), \end{cases}$ $\mathbf{P}\{X \in (-2; 5)\} =$
 $= \frac{3}{4}$.

Задача 2. Випадкова величина X має функцію розподілу

$$F(x) = \begin{cases} 0, & x \leq 0, \\ a\sqrt{2x}, & 0 < x \leq 8, \\ 1, & x > 8. \end{cases}$$

Знайти параметр a , аналітичний вираз для щільності, імовірність потрапляння випадкової величини X в інтервал $(1; 6)$.

Відповідь. $a = \frac{1}{4}$, $f(x) = \begin{cases} \frac{1}{4\sqrt{2x}}, & x \in (0; 8); \\ 0, & x \notin (0; 8), \end{cases}$ $\mathbf{P}\{X \in (1; 6)\} =$
 $= \frac{2\sqrt{3} - \sqrt{2}}{4} \approx 0,51$.

Задача 3. При яких значеннях параметрів a і b функція

$$F(x) = \begin{cases} 0, & x \leq a, \\ \sin x, & a < x \leq b, \\ 1, & x > b \end{cases}$$

буде функцією розподілу деякої випадкової величини? Знайти щільність даної випадкової величини.

Відповідь. $a = 2\pi k$, $b = \frac{\pi}{2} + 2\pi k$, $k \in \mathbb{Z}$. Якщо, наприклад, $k = 0$,

$$\text{то } f(x) = \begin{cases} 0, & x \notin \left(0; \frac{\pi}{2}\right); \\ \cos x, & x \in \left(0; \frac{\pi}{2}\right). \end{cases}$$

Задача 4. При яких значеннях параметрів a і b функція

$$F(x) = \begin{cases} 0, & x \leq a, \\ \operatorname{tg} x, & a < x \leq b, \\ 1, & x > b \end{cases}$$

буде функцією розподілу деякої випадкової величини? Знайти щільність даної випадкової величини.

Відповідь. $a = \pi k$, $b = \frac{\pi}{4} + \pi k$, $k \in \mathbb{Z}$. Якщо, наприклад, $k = 0$, то

$$f(x) = \begin{cases} 0, & x \notin \left(0; \frac{\pi}{4}\right); \\ \frac{1}{\cos^2 x}, & x \in \left(0; \frac{\pi}{4}\right). \end{cases}$$

Задача 5. Щільність випадкової величини X задано графічно (рис. 16).

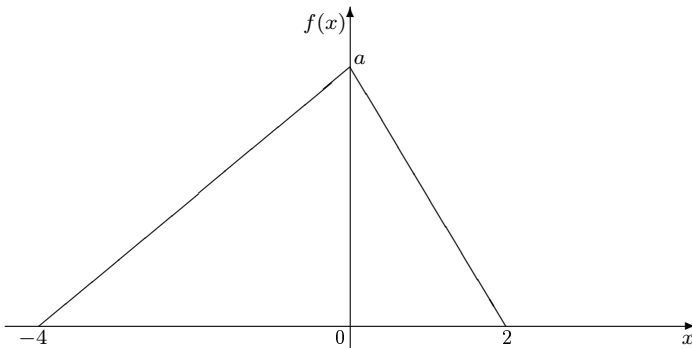


Рис. 16

Визначити значення параметра a й аналітичні вирази для $f(x)$ і $F(x)$. Знайти

$$\mathbf{P}\{X \in [0; 2]\}, \quad \mathbf{P}\{X \in [-2; 1]\}.$$

Відповідь. $a = \frac{1}{3}$, $f(x) = \begin{cases} 0, & x \notin (-4; 2); \\ \frac{1}{12}(x+4), & x \in [-4; 0]; \\ \frac{1}{6}(2-x), & x \in [0; 2], \end{cases}$

$$F(x) = \begin{cases} 0, & x \leq -4; \\ \frac{(x+4)^2}{24}, & -4 \leq x \leq 0; \\ 1 - \frac{(2-x)^2}{12}, & 0 \leq x \leq 2; \\ 1, & x \geq 2, \end{cases} \quad \mathbf{P}\{X \in [0; 2]\} = \frac{1}{3},$$

$$\mathbf{P}\{X \in [-2; 1]\} = \frac{3}{4}.$$

Задача 6. Щільність випадкової величини ξ задано графічно (рис. 17).

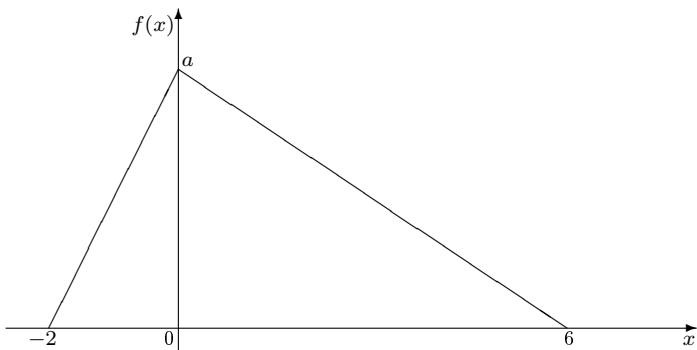


Рис. 17

Визначити значення параметра a й аналітичні вирази для $f(x)$ і $F(x)$. Знайти

$$\mathbf{P}\{\xi \in [6; 8]\}, \quad \mathbf{P}\{\xi \in [-2; 2]\}.$$

Відповідь. $a = \frac{1}{4}$, $f(x) = \begin{cases} 0, & x \notin [-2; 6]; \\ \frac{1}{8}(x+2), & x \in [-2; 0]; \\ \frac{1}{24}(6-x), & x \in [0; 6], \end{cases}$

$$F(x) = \begin{cases} 0, & x \leq -2; \\ \frac{(x+2)^2}{16}, & -2 < x \leq 0; \\ 1 - \frac{(6-x)^2}{48}, & 0 < x \leq 6; \\ 1, & x > 6, \end{cases} \quad \mathbf{P}\{\xi \in [6; 8]\} = 0,$$

$$\mathbf{P}\{\xi \in [-2; 2]\} = \frac{2}{3}.$$

Задача 7. При яких значеннях параметрів a та b функція

$$f(x) = \begin{cases} 0, & x \leq a, \\ \sin x, & a < x \leq b, \\ 0, & x > b \end{cases}$$

може бути щільністю деякої випадкової величини X ? Знайти функцію розподілу даної випадкової величини й імовірність потрапляння цієї випадкової величини в проміжок $\left[\frac{a+b}{2}; b\right]$.

Відповідь. Число a можна вибрати довільно з проміжку $\left[2\pi k; \frac{\pi}{2} + 2\pi k\right]$, $k \in \mathbb{Z}$, а число b визначити з умови $\cos a - \cos b = 1$.

$$\text{Якщо } a = 0, \text{ то } b = \frac{\pi}{2}. \text{ Тоді } F(x) = \begin{cases} 0, & x \leq 0; \\ 1 - \cos x, & x \in \left[0; \frac{\pi}{2}\right]; \\ 1, & x \geq \frac{\pi}{2}, \end{cases}$$

$$\mathbf{P} \left\{ X \in \left[\frac{\pi}{4}; \frac{\pi}{2} \right] \right\} = \frac{\sqrt{2}}{2}.$$

Задача 8. При яких значеннях параметрів a та b функція

$$f(x) = \begin{cases} 0, & x \leq a, \\ \operatorname{tg} x, & a < x \leq b, \\ 0, & x > b \end{cases}$$

може бути щільністю деякої випадкової величини X ? Знайти функцію розподілу даної випадкової величини й імовірність потрапляння цієї випадкової величини в проміжок $\left[a; \frac{a+b}{2} \right]$.

Відповідь. Число a можна вибрати довільно з проміжку $\left[\pi k; \frac{\pi}{2} + \pi k \right]$, $k \in \mathbb{Z}$, а число b визначити з умови $\frac{\cos a}{\cos b} = e$. Якщо $a = 0$,

$$\text{то } b = \arccos \frac{1}{e}. \text{ Тоді } F(x) = \begin{cases} 0, & x \leq 0; \\ -\ln(\cos x), & x \in \left[0; \arccos \frac{1}{e} \right]; \\ 1, & x \geq \arccos \frac{1}{e}, \end{cases}$$

$$\mathbf{P} \left\{ X \in \left[0; \frac{1}{2} \arccos \frac{1}{e} \right] \right\} = -\frac{1}{2} \ln \left(\frac{e+1}{2e} \right) \approx 0,19.$$

Задача 9. Випадкову величину ξ задано щільністю

$$f_{\xi}(x) = \begin{cases} 0, & x \notin \left[-\frac{\pi}{2}; \frac{\pi}{2} \right]; \\ a \cos x, & x \in \left[-\frac{\pi}{2}; \frac{\pi}{2} \right]. \end{cases}$$

Знайти параметр a , функцію розподілу $F_{\xi}(x)$, імовірність потрапляння випадкової величини ξ в інтервал $\left(0; \frac{\pi}{4} \right]$.

Відповідь. $a = \frac{1}{2}; F_{\xi}(x) = \begin{cases} 0, & x \leq -\frac{\pi}{2}; \\ \frac{1 + \sin x}{2}, & -\frac{\pi}{2} < x \leq \frac{\pi}{2}; \\ 1, & x > \frac{\pi}{2}; \end{cases}$

$\mathbf{P}\left(\xi \in \left(0; \frac{\pi}{4}\right]\right) = \frac{\sqrt{2}}{4}.$

Задача 10. Випадкову величину ξ задано щільністю

$$f_{\xi}(x) = \begin{cases} 0, & x \leq -2; \\ a, & -2 < x \leq 5; \\ 0, & x > 5. \end{cases}$$

Знайти параметр a , функцію розподілу $F_{\xi}(x)$, імовірність потрапляння випадкової величини ξ в інтервал $(0; 3]$.

Відповідь. $a = \frac{1}{7}; F_{\xi}(x) = \begin{cases} 0, & x \leq -2; \\ \frac{x+2}{7}, & -2 < x \leq 5; \\ 1, & x > 5; \end{cases} \mathbf{P}(\xi \in (0; 3]) =$
 $= \frac{3}{7}.$

Задача 11. Випадкова величина ξ має рівномірний закон розподілу на відрізку $[2; 5]$. Знайти аналітичні вирази для щільності та функції розподілу цієї випадкової величини. Побудувати їх графіки. Знайти ймовірність потрапляння випадкової величини в інтервал $(0; 3]$.

Відповідь. $f_{\xi}(x) = \begin{cases} 0, & x \notin [2; 5]; \\ \frac{1}{3}, & x \in [2; 5], \end{cases} F_{\xi}(x) = \begin{cases} 0, & x \leq 2; \\ \frac{x-2}{3}, & 2 < x \leq 5; \\ 1, & x > 5. \end{cases}$

Графіки щільності та функції розподілу зображено відповідно на рис. 18 і 19. $\mathbf{P}(\xi \in (0; 3]) = \frac{1}{3}.$

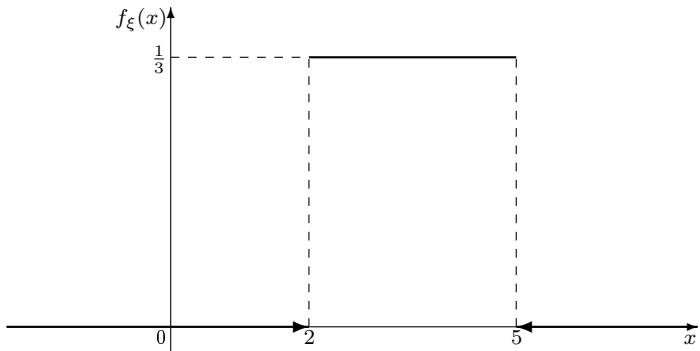


Рис. 18

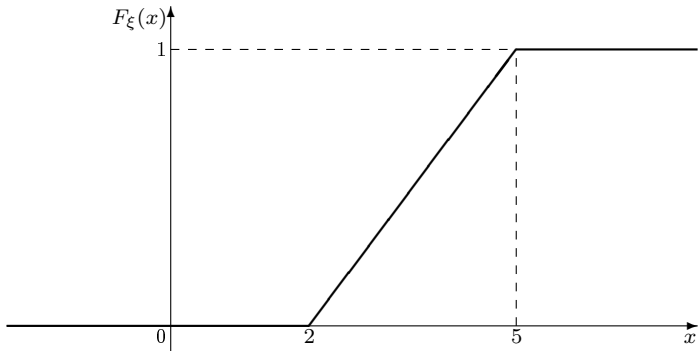


Рис. 19

Задача 12. Поїзди метро йдуть з інтервалом 2 хв. Вважаючи, що час ξ очікування поїзда на зупинці має рівномірний розподіл, знайти аналітичні вирази для щільності та функції розподілу цієї випадкової величини. Побудувати їх графіки. Знайти ймовірність того, що час очікування перевищуватиме 30 с.

Відповідь. $f_{\xi}(x) = \begin{cases} 0, & x \notin [0; 2]; \\ \frac{1}{2}, & x \in [0; 2], \end{cases}$ $F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x}{2}, & 0 < x \leq 2; \\ 1, & x > 2. \end{cases}$ Гра-

фіки щільності та функції розподілу зображено відповідно на рис. 20

i 21. $\mathbf{P}(\xi > 30 \text{ с}) = \frac{3}{4}$.

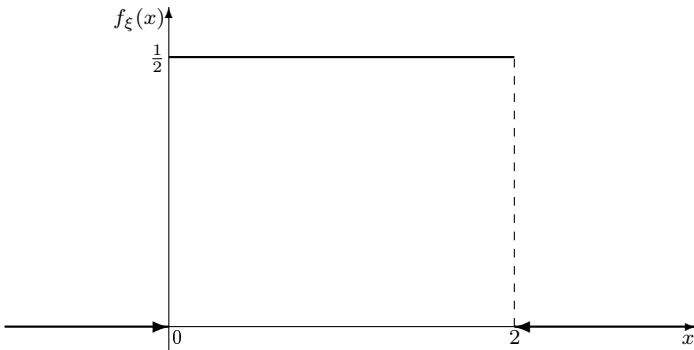


Рис. 20

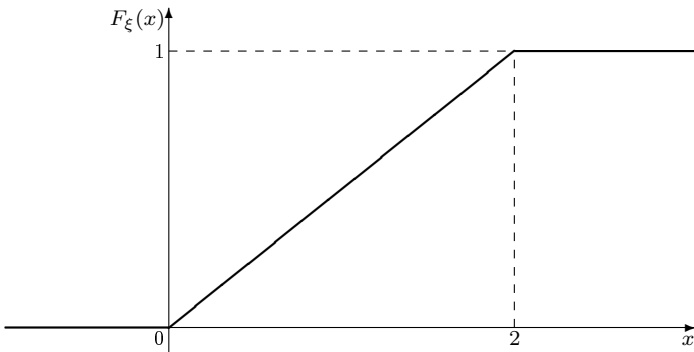


Рис. 21

Задача 13. Нормальний закон розподілу випадкової величини X задано функцією розподілу

$$F(x) = \frac{1}{4\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t+1)^2}{32}} dt.$$

Побудувати графік функції $f(x)$. Обчислити

$$\mathbf{P}\{-5 < X < 3\}, \quad \mathbf{P}\{|X + 1| < 12\}.$$

Відповідь. $a = -1$; $\sigma = 4$. Графік $f(x)$ утворюється з графіка щільності стандартного нормального розподілу шляхом стиску в 4 рази вздовж осі Oy і перенесення на 1 ліворуч уздовж осі Ox (рис. 22). $\mathbf{P}\{-5 < X < 3\} = 2 \cdot \Phi(1) \approx 0,6318$; $\mathbf{P}\{|X + 1| < 12\} = 2 \cdot \Phi(3) \approx 0,9973$.

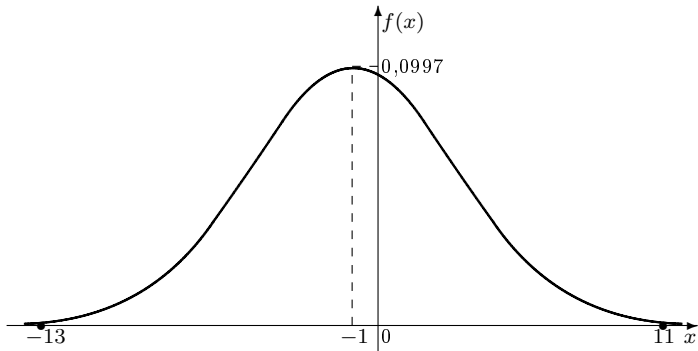


Рис. 22

Задача 14. Нормальний закон розподілу випадкової величини X задано функцією розподілу

$$F(x) = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-3)^2}{8}} dt.$$

Побудувати графік функції $f(x)$. Обчислити

$$\mathbf{P}\{1 < X < 5\}, \quad \mathbf{P}\{|X - 3| < 6\}.$$

Відповідь. $a = 3$; $\sigma = 2$. Графік $f(x)$ утворюється з графіка щільності стандартного нормального розподілу шляхом стиску в 2 рази вздовж осі Oy і перенесення на 3 праворуч уздовж осі Ox (рис. 23). $\mathbf{P}\{1 < X < 5\} \approx 0,6318$; $\mathbf{P}\{|X - 3| < 6\} \approx 0,9973$.

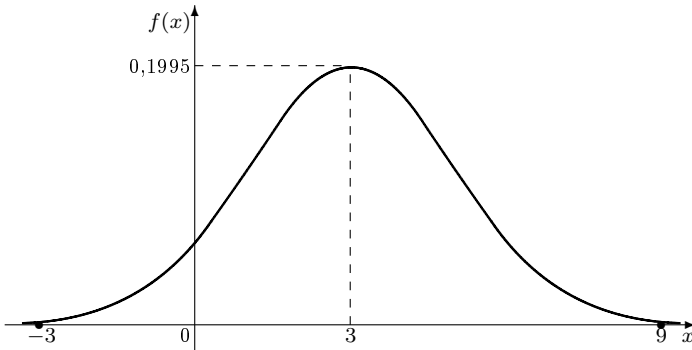


Рис. 23

Задача 15. Випадкова величина ξ має нормальний розподіл із параметрами a, σ . Записати вираз для щільності та функції розподілу випадкової величини ξ і знайти ймовірність того, що випадкова величина ξ набуде значення з проміжку $(\alpha, \beta]$, якщо:

- а) $a = 0, \sigma = 1, \alpha = 0, \beta = 5$;
- б) $a = 5, \sigma = 3, \alpha = 2, \beta = 8$;
- в) $a = 9, \sigma = 2, \alpha = 3, \beta = 15$.

Відповідь. а) $f_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, F_{\xi}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$

$\mathbf{P}(\xi \in (0; 5]) \approx 0,5$; б) $f_{\xi}(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x-5)^2}{18}}, F_{\xi}(x) = \frac{1}{3\sqrt{2\pi}} \times$
 $\times \int_{-\infty}^x e^{-\frac{(t-5)^2}{18}} dt, \mathbf{P}(\xi \in (2; 8]) \approx 0,6827$; в) $f_{\xi}(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-9)^2}{8}},$

$F_{\xi}(x) = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-9)^2}{8}} dt, \mathbf{P}(\xi \in (3; 15]) \approx 0,9973.$

Задача 16. Щільність нормально розподіленої випадкової величини ξ має вигляд $f_{\xi}(x) = c \cdot e^{-\frac{(x+3)^2}{8}}$. Знайти параметри c, a та σ . Порівняти ймовірності потрапляння випадкової величини ξ у проміжки $(-9; 3]$ і $(4; 100]$.

Відповідь. $c = \frac{1}{2\sqrt{2\pi}}$; $a = -3$; $\sigma = 2$; $\mathbf{P}(\xi \in (-9; 3]) > \mathbf{P}(\xi \in (4; 100])$.

Задача 17. Випадкова величина ξ має нормальний розподіл із параметрами a, σ . У кожному з наступних пунктів знайти інтервал, у який потрапляє випадкова величина ξ з практичною достовірністю (з імовірністю 0,9973):

- а) $a = 0, \sigma = 1$;
- б) $a = 5, \sigma = 2$;
- в) $a = 5, \sigma = 100$.

Відповідь. а) $[-3; 3]$; б) $[-1; 11]$; в) $[-295; 305]$.

Задача 18. Випадкова величина ξ має показниковий розподіл з параметром $\lambda = 3$. Знайти ймовірність того, що випадкова величина ξ набуде значення з проміжку $[\alpha; \beta]$, якщо:

- а) $\alpha = -2, \beta = -1$;
- б) $\alpha = 2, \beta = 3$;
- в) $\alpha = -2, \beta = 3$.

Відповідь. а) $\mathbf{P}(\xi \in [-2; -1]) = 0$; б) $\mathbf{P}(\xi \in [-3; -1]) \approx 0,002355$;
в) $\mathbf{P}(\xi \in [-2; 3]) \approx 0,9999$.

Задача 19. Випадкова величина ξ має експоненціальний розподіл. Імовірність того, що ця випадкова величина набуде значення з проміжку $[0; 5]$, дорівнює 0,7. Знайти ймовірність того, що ця випадкова величина набуде значення з проміжку $[7; 9]$.

Відповідь. $\mathbf{P}(\xi \in [7; 9]) \approx 0,0708$.

Розділ 8. Числові характеристики випадкових величин

До основних числових характеристик, які описують розподіл випадкової величини, належать *математичне сподівання, дисперсія, середнє квадратичне відхилення, мода, медіана*.

Математичним сподіванням дискретної випадкової величини ξ , яка набуває значень x_i з імовірностями p_i , називається число

$$M(\xi) = \sum_i x_i p_i. \quad (4)$$

Якщо множина значень випадкової величини ξ скінченна, то математичне сподівання є звичайною сумою добутків значень випадкової величини на відповідні ймовірності. Якщо множина значень випадкової величини ξ зліченна, то вимагатимемо абсолютної збіжності ряду (4).

Математичним сподіванням абсолютно неперервної випадкової величини ξ зі щільністю $f(x)$ називається число

$$M(\xi) = \int_{-\infty}^{\infty} x f(x) dx, \quad (5)$$

за умови абсолютної збіжності невластного інтеграла (5).

Математичне сподівання володіє такими властивостями.

Властивість 1. Якщо $P\{\xi = C\} = 1$, то $M(C) = C$.

Властивість 2. $M(C \cdot \xi) = C \cdot M(\xi)$.

Властивість 3. $M(\xi_1 + \xi_2) = M(\xi_1) + M(\xi_2)$.

Властивість 4. Якщо ξ_1 та ξ_2 незалежні випадкові величини, то

$$M(\xi_1 \cdot \xi_2) = M(\xi_1) \cdot M(\xi_2).$$

Дисперсією випадкової величини ξ називають математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання:

$$\mathbf{D}(\xi) = \mathbf{M}[\xi - \mathbf{M}(\xi)]^2. \quad (6)$$

Формула (6) еквівалентна такій формулі:

$$\mathbf{D}(\xi) = \mathbf{M}(\xi^2) - [\mathbf{M}(\xi)]^2. \quad (7)$$

Для дискретної випадкової величини формули (6) і (7) набувають відповідно такого вигляду:

$$\mathbf{D}(\xi) = \sum_i (x_i - \mathbf{M}(\xi))^2 p_i, \quad (8)$$

$$\mathbf{D}(\xi) = \sum_i x_i^2 p_i - [\mathbf{M}(\xi)]^2. \quad (9)$$

Для абсолютно неперервної випадкової величини формули (6) та (7) набувають відповідно такого вигляду:

$$\mathbf{D}(\xi) = \int_{-\infty}^{\infty} (x - \mathbf{M}(\xi))^2 f(x) dx, \quad (10)$$

$$\mathbf{D}(\xi) = \int_{-\infty}^{\infty} x^2 f(x) dx - [\mathbf{M}(\xi)]^2. \quad (11)$$

В означеннях вважається, що ряд (8) і невласний інтеграл (10) абсолютно збігаються.

Дисперсія має такі властивості:

Властивість 1. $\mathbf{D}(\xi) \geq 0$.

Властивість 2. Якщо $\mathbf{P}\{\xi = C\} = 1$, то $\mathbf{D}(C) = 0$.

Властивість 3. $\mathbf{D}(C \cdot \xi) = C^2 \cdot \mathbf{D}(\xi)$.

Властивість 4. Якщо ξ_1 та ξ_2 незалежні випадкові величини, то

$$\mathbf{D}(\xi_1 + \xi_2) = \mathbf{D}(\xi_1) + \mathbf{D}(\xi_2).$$

Середнім квадратичним відхиленням випадкової величини ξ називають число

$$\sigma(\xi) = \sqrt{\mathbf{D}(\xi)}.$$

Дисперсія та середнє квадратичне відхилення характеризують ступінь розсіювання значень випадкової величини навколо її математичного сподівання.

Модю $\mathbf{Mo}(\xi)$ дискретної випадкової величини називається найімовірніше її значення.

Розподіл називається *унімодальним*, якщо він має єдину моду (біноміальний, показниковий тощо).

Розподіл називається *полімодальним*, якщо він має більше однієї моди (рівномірний тощо).

Модю абсолютно неперервної випадкової величини називається точка максимуму щільності розподілу.

Медіаною випадкової величини називається таке число $\mathbf{Me}(\xi)$, для якого виконується умова

$$\mathbf{P}\{\xi < \mathbf{Me}(\xi)\} = \mathbf{P}\{\xi > \mathbf{Me}(\xi)\}.$$

Якщо ξ має дискретний розподіл, то, взагалі кажучи, $\mathbf{Me}(\xi)$ визначається неоднозначно. Наприклад, якщо ξ — кількість очок, які випали на гральному кубіку, то як $\mathbf{Me}(\xi)$ можна взяти довільне число з інтервалу (3; 4).

У табл. 1 наведено математичні сподівання, дисперсії та середні квадратичні відхилення найпоширеніших розподілів ймовірностей.

Приклад 1. Випадкова величина X має такий розподіл:

X	1	2	3
\mathbf{P}	0,2	0,6	0,2

Знайти числові характеристики розподілу випадкової величини X .

Розв'язання. Математичне сподівання випадкової величини X обчислимо за формулою

Таблиця 1

Числові характеристики найпоширеніших імовірнісних розподілів

№ п/п	Назва розподілу	$M(\xi)$	$D(\xi)$	$\sigma(\xi)$
I. Дискретні розподіли				
1	Рівномірний з параметром n	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\sqrt{\frac{n^2-1}{12}}$
2	Біноміальний з параметрами n і p	np	$np(1-p)$	$\sqrt{np(1-p)}$
3	Показниковий з параметром λ	λ	λ	$\sqrt{\lambda}$
4	Геометричний з параметром p	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{\sqrt{1-p}}{p}$
II. Неперервні розподіли				
5	Рівномірний з параметрами a, b	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{b-a}{2\sqrt{3}}$
6	Нормальний з параметрами a, σ	a	σ^2	σ
7	Показниковий з параметром a	$\frac{1}{a}$	$\frac{1}{a^2}$	$\frac{1}{a}$

$$\begin{aligned}
 M(X) &= \sum_i x_i p_i = x_1 p_1 + x_2 p_2 + x_3 p_3 = \\
 &= 1 \cdot 0,2 + 3 \cdot 0,6 + 5 \cdot 0,2 = 3.
 \end{aligned}$$

Дисперсію випадкової величини X знайдемо за формулою

$$\begin{aligned}
 D(X) &= \sum_i x_i^2 p_i - [M(X)]^2 = x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 - 3^2 = \\
 &= 1^2 \cdot 0,2 + 3^2 \cdot 0,6 + 5^2 \cdot 0,2 - 3^2 = 1,6.
 \end{aligned}$$

Середнє квадратичне відхилення випадкової величини X

$$\sigma(X) = \sqrt{D(X)} = \sqrt{1,6} \approx 0,2649.$$

Мода і медіана дорівнюють

$$Mo(X) = Me(X) = 3.$$

□

Приклад 2. Нехай випадкова величина ξ дорівнює кількості номерів, угаданих гравцем в лотереї «6 із 39».

Знайти математичне сподівання, дисперсію та середнє квадратичне відхилення розподілу випадкової величини ξ .

Розв'язання. Ряд розподілу даної випадкової величини побудовано у прикладі 2 розділу 6.

ξ	0	1	2	3	4	5	6
P	$\frac{1107568}{3262623}$	$\frac{1424016}{3262623}$	$\frac{613800}{3262623}$	$\frac{109120}{3262623}$	$\frac{7920}{3262623}$	$\frac{198}{3262623}$	$\frac{1}{3262623}$

За формулою

$$M(\xi) = \sum_i x_i p_i$$

обчислимо математичне сподівання випадкової величини ξ :

$$\begin{aligned} M(\xi) &= 0 \cdot \frac{1107568}{3262623} + 1 \cdot \frac{1424016}{3262623} + 2 \cdot \frac{613800}{3262623} + 3 \cdot \frac{109120}{3262623} + \\ &\quad + 4 \cdot \frac{7920}{3262623} + 5 \cdot \frac{198}{3262623} + 6 \cdot \frac{1}{3262623} = \\ &= \frac{12}{13} < 1. \end{aligned}$$

Отже, середня кількість угаданих номерів приблизно дорівнює одиниці.

За формулою

$$D(\xi) = \sum_i x_i^2 p_i - [M(\xi)]^2$$

знайдемо дисперсію випадкової величини ξ :

$$\begin{aligned} D(\xi) &= 0^2 \cdot \frac{1107568}{3262623} + 1^2 \cdot \frac{1424016}{3262623} + 2^2 \cdot \frac{613800}{3262623} + 3^2 \cdot \frac{109120}{3262623} + \\ &\quad + 4^2 \cdot \frac{7920}{3262623} + 5^2 \cdot \frac{198}{3262623} + 6^2 \cdot \frac{1}{3262623} - \left(\frac{12}{13}\right)^2 = \\ &= \frac{2178}{3211} \approx 0,6783. \end{aligned}$$

Середнє квадратичне відхилення випадкової величини ξ

$$\sigma(\xi) = \sqrt{\mathbf{D}(\xi)} = \sqrt{\frac{2178}{3211}} \approx 0,8236. \quad \square$$

Приклад 3. Випадкову величину X задано щільністю розподілу

$$f(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{2\sqrt{x}}, & 0 < x \leq \frac{1}{4}, \\ 16x, & \frac{1}{4} < x \leq \frac{\sqrt{2}}{4}, \\ 0, & x > \frac{\sqrt{2}}{4}. \end{cases}$$

Знайти числові характеристики даної випадкової величини.

Розв'язання. Математичне сподівання випадкової величини X обчислимо за формулою $\mathbf{M}(X) = \int_{-\infty}^{\infty} xf(x)dx$:

$$\begin{aligned} \mathbf{M}(X) &= \int_{-\infty}^0 x \cdot 0 dx + \int_0^{\frac{1}{4}} x \cdot \frac{1}{2\sqrt{x}} dx + \int_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} x \cdot 16x dx + \int_{\frac{\sqrt{2}}{4}}^{\infty} x \cdot 0 dx = \\ &= \frac{1}{2} \int_0^{\frac{1}{4}} \sqrt{x} dx + 16 \int_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} x^2 dx = \frac{1}{2} \cdot \frac{x^{\frac{3}{2}}}{\frac{3}{2}} \Big|_0^{\frac{1}{4}} + 16 \cdot \frac{x^3}{3} \Big|_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} = \\ &= \frac{\sqrt{2}}{6} - \frac{1}{24} \approx 0,1940. \end{aligned}$$

За формулою

$$\mathbf{D}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - [\mathbf{M}(\xi)]^2$$

знайдемо дисперсію випадкової величини X :

$$\begin{aligned}
 \mathbf{D}(X) &= \left(\int_{-\infty}^0 x^2 \cdot 0 dx + \int_0^{\frac{1}{4}} x^2 \cdot \frac{1}{2\sqrt{x}} dx + \int_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} x^2 \cdot 16x dx + \int_{\frac{\sqrt{2}}{4}}^{\infty} x^2 \cdot 0 dx \right) - \left(\frac{\sqrt{2}}{6} - \frac{1}{24} \right)^2 = \\
 &= \frac{1}{2} \int_0^{\frac{1}{4}} x\sqrt{x} dx + 16 \int_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} x^3 dx - \left(\frac{\sqrt{2}}{6} - \frac{1}{24} \right)^2 = \\
 &= \frac{1}{2} \cdot \frac{x^{\frac{5}{2}}}{\frac{5}{2}} \Big|_0^{\frac{1}{4}} + 16 \cdot \frac{x^4}{4} \Big|_{\frac{1}{4}}^{\frac{\sqrt{2}}{4}} - \frac{11}{192} + \frac{\sqrt{2}}{72} = \\
 &= \frac{\sqrt{2}}{72} - \frac{1}{240} \approx 0,0155.
 \end{aligned}$$

Середнє квадратичне відхилення випадкової величини X

$$\sigma(\xi) = \sqrt{\mathbf{D}(\xi)} = \sqrt{\frac{\sqrt{2}}{72} - \frac{1}{240}} \approx 0,1244. \quad \square$$

Задачі до розділу 8

Задача 1. Нехай ξ — випадкова величина, яка дорівнює кількості очок, які випадуть при підкиданні кубика. Знайти математичне сподівання, дисперсію та середнє квадратичне відхилення випадкової величини ξ .

Відповідь. $\mathbf{M}(\xi) = 3,5$; $\mathbf{D}(\xi) = \frac{35}{12}$; $\sigma(\xi) = \sqrt{\frac{35}{12}} \approx 1,71$.

Задача 2. У ящику лежить 5 пронумерованих кульок (номери змінюються від 1 до 5). Навмання виймається кулька. Випадкова вели-

чина ξ — номер цієї кульки. Знайти математичне сподівання, дисперсію та середнє квадратичне відхилення випадкової величини ξ .

Відповідь. $M(\xi) = 3$; $D(\xi) = 2$; $\sigma(\xi) = \sqrt{2} \approx 1,41$.

Задача 3. Симетрична монета підкидається 3 рази. Випадкова величина X — кількість «аверсів», які при цьому випадали. Знайти математичне сподівання та дисперсію випадкової величини X .

Відповідь. $M(X) = \frac{3}{2}$; $D(X) = \frac{3}{4}$.

Задача 4. Знайти числові характеристики випадкової величини X , яка дорівнює кількості хлопчиків у сім'ї з трьома дітьми (вважаючи народження хлопчика й дівчинки рівноймовірними подіями).

Відповідь. $M(X) = \frac{3}{2}$; $D(X) = \frac{3}{4}$; $\sigma(X) = \frac{\sqrt{3}}{2} \approx 0,87$.

Задача 5. Стрілець робить по одному пострілу по чотирьох мішенях. Знайти числові характеристики випадкової величини X , яка дорівнює кількості влучень, вважаючи, що ймовірність влучення при одному пострілі 0,7.

Відповідь. $M(X) = 2,8$; $D(X) = 0,84$; $\sigma(X) = \sqrt{0,84} \approx 0,92$.

Задача 6. Студент знає 20 питань з 25. Навмання виймається 5 питань. Знайти числові характеристики випадкової величини X , яка дорівнює кількості питань, на які студент знає відповідь.

Вказівка. Скористайтесь числовими характеристиками гіпергеометричного розподілу з параметрами m, n, N :

$$M(X) = \frac{mn}{N}, \quad D(X) = \frac{mn(N-m)(N-n)}{N^2(N-1)},$$

$$\sigma(X) = \sqrt{\frac{mn(N-m)(N-n)}{N^2(N-1)}}.$$

Відповідь. $M(X) = 4$; $D(X) = \frac{2}{3}$; $\sigma(X) = \sqrt{\frac{2}{3}} \approx 0,82$.

Задача 7. У ящику лежать кульки — 7 зелених і 3 жовті. Навмання вибирається кулька, фіксується її колір і кулька повертається в ящик. Експеримент повторюється 4 рази. Знайти математичне сподівання та дисперсію випадкової величини X , яка дорівнює кількості зелених кульок, які з'явилися при виїманні.

Відповідь. $M(X) = 2,8$; $D(X) = 0,84$; $\sigma(X) = \sqrt{0,84} \approx 0,92$.

Задача 8. У групі з 20 студентів є троє відмінників. Випадково вибирається 4 студенти. Знайти числові характеристики випадкової величини ξ , яка дорівнює кількості відмінників серед вибраних студентів.

Відповідь. $M(\xi) = 0,6$; $D(\xi) = \frac{204}{475} \approx 0,43$; $\sigma(\xi) = \sqrt{\frac{204}{475}} \approx 0,66$.

Задача 9. Серед 10 деталей є 2 браковані. Для контролю навмання вибирають 3 деталі. Знайти числові характеристики випадкової величини ξ , яка дорівнює кількості бракованих деталей серед вибраних.

Відповідь. $M(\xi) = 0,6$; $D(\xi) = \frac{28}{75} \approx 0,37$; $\sigma(\xi) = \sqrt{\frac{28}{75}} \approx 0,61$.

Задача 10. Знайти числові характеристики випадкової величини Y , яка дорівнює кількості вгаданих номерів у лотереї «5 із 36».

Відповідь. $M(Y) \approx 0,69$; $D(Y) \approx 0,53$; $\sigma(Y) \approx 0,73$.

Задача 11. Кубик підкидається доти, поки не випаде шістька. X — випадкова величина, яка дорівнює необхідній кількості підкидань. Знайти числові характеристики випадкової величини X і

$$\mathbf{P}\{|X - M(X)| < 3\sigma(X)\}.$$

Відповідь. $M(X) = 6$; $D(X) = 30$; $\sigma(X) = \sqrt{30} \approx 5,48$; $\mathbf{P}\{|X - M(X)| < 3\sigma(X)\} = \mathbf{P}\{1 < X < 22\} = 1 - \frac{1}{6^{22}} \approx 1$.

Задача 12. Гравець грає в рулетку до першого свого виграшу, кожного разу ставлячи на червоне й потроюючи ставку. X — випадкова величина, яка дорівнює кількості спроб, Y — випадкова величина, яка дорівнює виграшу гравця. Знайти числові характеристики випадкових величин X і Y , вважаючи, що гравець має необмежену кількість грошей і розмір ставки не обмежений правилами.

Відповідь. $M(X) = \frac{37}{18} \approx 2,06$; $D(X) = \frac{703}{324} \approx 2,17$; $\sigma(X) = \sqrt{\frac{703}{324}} \approx 1,47$. Випадкова величина $Y = a$, якщо $X = 1$, і $Y = 2a \cdot 3^{n-2} \cdot X$, якщо $X = n$, $n = 2, 3, \dots$, де a — розмір першої ставки; $M(Y) = +\infty$.

Задача 13. Магазин отримав 10000 пляшок мінеральної води. Імовірність того, що при перевезенні пляшка виявиться розбитою, дорівнює 0,0004. Знайти числові характеристики випадкової величини X , яка дорівнює кількості розбитих пляшок.

Відповідь. $M(X) = 4$; $D(X) = 4$; $\sigma(X) = 2$.

Задача 14. Підручник видано тиражем 10000 екземплярів. Імовірність того, що зброшурований екземпляр виявиться неякісним, дорівнює 0,001. Знайти числові характеристики випадкової величини, яка дорівнює кількості бракованих екземплярів.

Відповідь. $M(X) = 10$; $D(X) = 10$; $\sigma(X) = \sqrt{10} \approx 3,16$.

Задача 15. Автоматична телефонна станція обслуговує 10000 телефонних номерів. Імовірність того, що протягом 1 хв на АТС надійде виклик від абонента, дорівнює 0,0004. Знайти числові характеристики випадкової величини X , яка дорівнює кількості викликів, що надійшли на АТС протягом 1 хв.

Відповідь. $M(X) = 4$; $D(X) = 4$; $\sigma(X) = 2$.

Задача 16. Імовірність влучення в муху на льоту з рогатки дорівнює 0,001. Відбувається 2000 пострілів. Знайти числові характеристики випадкової величини Y , яка дорівнює кількості вбитих мух.

Відповідь. $M(Y) = 2$; $D(Y) = 2$; $\sigma(Y) = \sqrt{2} \approx 1,41$.

Задача 17. Випадкова величина ξ рівномірно розподілена на відріжку $[1; 11]$. Знайти числові характеристики випадкової величини ξ .

Відповідь. $M(X) = 6$; $D(X) = \frac{25}{3}$; $\sigma(X) = \sqrt{\frac{25}{3}} \approx 2,89$.

Задача 18. Випадкова величина ξ рівномірно розподілена на $[7; a]$, причому щільність на цьому проміжку дорівнює $\frac{1}{20}$. Зазначити значення параметра a і знайти числові характеристики випадкової величини ξ .

Відповідь. $a = 27$; $M(\xi) = 17$; $D(\xi) = \frac{100}{3}$; $\sigma(\xi) = \sqrt{\frac{100}{3}} \approx 5,77$.

Задача 19. Випадкова величина ξ задана щільністю

$$f(x) = \begin{cases} 0, & x \leq 0, \\ \frac{2}{e^{2x}}, & x > 0. \end{cases}$$

Знайти числові характеристики випадкової величини ξ .

Відповідь. $M(\xi) = \frac{1}{2}$; $D(\xi) = \frac{1}{4}$; $\sigma(\xi) = \frac{1}{2}$.

Задача 20. Випадкова величина ξ задана функцією розподілу

$$F(x) = \begin{cases} 0, & x \leq 2, \\ (x - 2)^2, & 2 < x \leq 3, \\ 1, & x > 3. \end{cases}$$

Знайти числові характеристики випадкової величини ξ .

Відповідь. $M(\xi) = \frac{8}{3}$; $D(\xi) = \frac{1}{18}$; $\sigma(\xi) = \frac{\sqrt{2}}{6} \approx 0,24$.

Задача 21. Щільність випадкової величини X задано графічно (рис. 24).

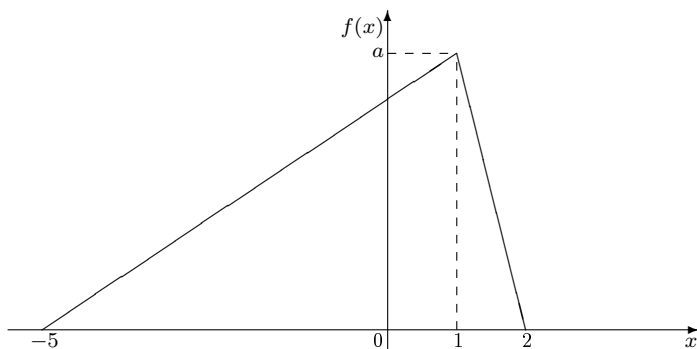


Рис. 24

Знайти значення параметра a та числові характеристики випадкової величини X .

Відповідь. $a = \frac{2}{7}$; $\mathbf{M}(X) = -\frac{2}{3}$; $\mathbf{D}(X) = \frac{43}{18} \approx 2,39$; $\sigma(X) = \sqrt{\frac{43}{18}} \approx 1,55$.

Задача 22. У середині круга радіуса R випадковим чином вибирається точка. Імовірність потрапляння точки в будь-яку область, що міститься в крузі, пропорційна площі цієї області. Знайти функцію розподілу та дисперсію випадкової величини ξ , яка дорівнює відстані від точки до центра круга.

Відповідь. $F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x}{2}, & 0 < x \leq R; \\ 1, & x > R, \end{cases} \quad \mathbf{D}(\xi) = \frac{R^2}{12}.$

Задача 23. Коробки з шоколадом пакуються автоматично, їх середня маса дорівнює 1,06 кг. Вважаючи, що маса коробок розподілена за нормальним законом, знайти стандартне відхилення, якщо 5% коробок мають масу менше 1 кг.

Відповідь. $\sigma = \frac{0,06}{2,58} \approx 0,023$.

Розділ 9. Системи двох випадкових величин

Якщо на одному і тому самому просторі елементарних подій Ω задано дві випадкові величини ξ і ψ , то вважають, що задано випадкову величину (ξ, ψ) (або випадковий двовимірний вектор).

При вивченні системи випадкових величин, узагалі кажучи, недостатньо знати інформацію про кожну випадкову складову. Необхідно враховувати ще й залежність між ними.

Функцією розподілу двовимірної випадкової величини (ξ, ψ) називається функція двох змінних

$$F(x, y) = \mathbf{P}\{\xi < x, \psi < y\}.$$

Функція розподілу геометрично визначає ймовірність потрапляння випадкової точки (ξ, ψ) у нескінченний квадрант з вершиною в точці (x, y) , що знаходиться лівіше й нижче від неї.

Функція розподілу двовимірної випадкової величини має такі властивості.

Властивість 1. $\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} F(x, y) = 1.$

Властивість 2. Для будь-якої системи випадкових величин

$$\lim_{x \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} F(x, y) = \lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F(x, y) = 0.$$

Властивість 3. а) Якщо $x_1 < x_2$, то

$$F(x_1, y) \leq F(x_2, y)$$

для всіх $y \in \mathbb{R}$.

б) Якщо $y_1 < y_2$, то

$$F(x, y_1) \leq F(x, y_2)$$

для всіх $x \in \mathbb{R}$.

Властивість 4. Якщо випадкова величина ξ має функцію розподілу $F_\xi(x)$, а випадкова величина ψ — функцію розподілу $F_\psi(x)$, то

$$\lim_{x \rightarrow +\infty} F(x, y) = F_\psi(y), \quad \lim_{y \rightarrow +\infty} F(x, y) = F_\xi(x).$$

Властивість 5. Імовірність потрапляння випадкової точки (ξ, ψ) у прямокутник зі сторонами, які паралельні координатним осям, обчислюється за формулою

$$P\{a \leq \xi < b; c \leq \psi < d\} = F(b, d) - F(a, d) - F(b, c) + F(a, c).$$

Як і в одновимірному випадку, двовимірні випадкові величини можна розділити на два класи: *дискретні* та *неперервні*.

Двовимірна випадкова величина (ξ, ψ) називається *дискретною*, якщо і випадкова величина ξ , і випадкова величина ψ є дискретними випадковими величинами. У цьому разі випадкову величину (ξ, ψ) зручно задавати за допомогою таблиці:

$\psi \backslash \xi$	ξ_1	ξ_2	\dots	ξ_n
ψ_1	p_{11}	p_{12}	\dots	p_{1n}
ψ_2	p_{21}	p_{22}	\dots	p_{2n}
\dots	\dots	\dots	\dots	\dots
ψ_m	p_{m1}	p_{m2}	\dots	p_{mn}

де p_{ij} — імовірності того, що одночасно виконуватимуться рівності $\psi = \psi_i, \xi = \xi_j$. Очевидно, при цьому

$$\sum_{i=1}^m \sum_{j=1}^n p_{ij} = 1.$$

Для дискретної випадкової величини (ξ, ψ) функція розподілу набуває вигляду суми

$$F(x, y) = \sum_{(i,j): \begin{cases} \xi_j < x; \\ \psi_i < y. \end{cases}} p_{ij},$$

тобто підсумовування виконується по всіх наборах (i, j) , для яких одночасно $\xi_j < x$ і $\psi_i < y$.

Користуючись наведеною вище таблицею, легко визначити розподіли кожної випадкової величини окремо за допомогою формул

$$\begin{aligned} \mathbf{P}\{\xi = \xi_j\} &= p_{1j} + p_{2j} + \dots + p_{mj}, & j &= 1, 2, \dots, n; \\ \mathbf{P}\{\psi = \psi_i\} &= p_{i1} + p_{i2} + \dots + p_{in}, & i &= 1, 2, \dots, m. \end{aligned}$$

Двовимірна випадкова величина називається *неперервною*, якщо ймовірність потрапляння випадкової точки (ξ, ψ) у довільну фіксовану точку дорівнює нулю, тобто

$$\mathbf{P}\{(\xi, \psi) = (x_0, y_0)\} = 0$$

для будь-яких (x_0, y_0) .

Випадкова величина (ξ, ψ) називається *абсолютно неперервною*, якщо існує така функція $f(x, y)$, для якої виконується рівність

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy.$$

Функція $f(x, y)$ називається *щільністю* розподілу двовимірної випадкової величини (ξ, ψ) .

Щільність $f(x, y)$ має такі властивості.

Властивість 1. $f(x, y) = F''_{xy}(x, y)$.

Властивість 2. $f(x, y) \geq 0$.

Властивість 3. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$.

Властивість 4. Імовірність потрапляння випадкової точки (ξ, ψ) у плоску область \mathcal{D}

$$\mathbf{P}\{(\xi, \psi) \in \mathcal{D}\} = \iint_{\mathcal{D}} f(x, y) dx dy.$$

Випадкові величини ξ і ψ називаються *незалежними*, якщо

$$F(x, y) = F_\xi(x) \cdot F_\psi(y).$$

Для абсолютно неперервних випадкових величин ця умова рівно-сильна умові

$$f(x, y) = f_\xi(x) \cdot f_\psi(y),$$

де $f_\xi(x)$ і $f_\psi(y)$ — щільність розподілу випадкових величин ξ і ψ відповідно.

Щільності $f_\xi(x)$ і $f_\psi(y)$ можна обчислити за такими формулами:

$$f_\xi(x) = \int_{-\infty}^{+\infty} f(x, y) dy = F'_\xi(x);$$
$$f_\psi(y) = \int_{-\infty}^{+\infty} f(x, y) dx = F'_\psi(y).$$

Якщо випадкові величини ξ і ψ залежні, то для знаходження закону розподілу двовимірної випадкової величини (ξ, ψ) недостатньо знати закони розподілу кожної з випадкових величин ξ і ψ .

Умовним законом розподілу називається розподіл однієї випадкової величини, знайдений за умови, що інша випадкова величина системи набула деякого фіксованого значення.

Умовний закон можна задавати як функцією розподілу $F(x/y)$, так і щільністю $f(x/y)$ (для абсолютно неперервного випадку), де

$$f(x/y) = \frac{f(x, y)}{f_\psi(y)}; \quad f(y/x) = \frac{f(x, y)}{f_\xi(x)}.$$

Отже,

$$f(x, y) = f_\xi(x) \cdot f(y/x) = f_\psi(y) \cdot f(x/y).$$

Умовна щільність має всі властивості щільності одновимірного закону розподілу.

Числові характеристики умовних законів розподілу обчислюють як і раніше, але використовуючи умовну щільність.

Для опису рівня залежності випадкових величин ξ і ψ використовують такі числові характеристики, як *коваріацію* та *коефіцієнт кореляції*.

Коваріацією (кореляційним моментом) системи випадкових величин називають число

$$\text{cov}(\xi, \psi) = \mathbf{M}[(\xi - \mathbf{M}\xi)(\psi - \mathbf{M}\psi)].$$

Якщо ξ і ψ дискретні, то

$$\text{cov}(\xi, \psi) = \sum_i \sum_j (x_i - \mathbf{M}\xi)(y_j - \mathbf{M}\psi) p_{ij}.$$

Якщо ξ і ψ неперервні, то

$$\text{cov}(\xi, \psi) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mathbf{M}\xi)(y - \mathbf{M}\psi) dx dy.$$

Коефіцієнтом кореляції системи випадкових величин ξ і ψ називають число

$$r(\xi, \psi) = \frac{\text{cov}(\xi, \psi)}{\sigma(\xi)\sigma(\psi)}.$$

Якщо ξ і ψ — незалежні випадкові величини, то

$$\text{cov}(\xi, \psi) = r(\xi, \psi) = 0.$$

Обернене твердження неправильне. Існують залежні випадкові величини, для яких коваріація та коефіцієнт кореляції дорівнюють нулю.

Для довільних випадкових величин

$$-1 \leq r(\xi, \psi) \leq 1.$$

Чим ближче $|r(\xi, \psi)|$ до одиниці, тим сильніший рівень залежності між ξ і ψ . Якщо $r(\xi, \psi) = 0$, то випадкові величини називаються *некорельованими*.

Приклад 1. Із коробки, у якій міститься три червоні й дві сині кульки, навмання без повернень виймають послідовно кульки до першої появи синьої кульки. Далі кульки виймають до першої появи червоної кульки.

Необхідно описати закон розподілу ймовірностей системи випадкових величин (X, Y) , де X — кількість кульок, узятих з коробки до першої появи синьої кульки; Y — кількість кульок, узятих з коробки до першої появи червоної кульки після того, як перша синя кулька була вийнята з коробки.

Скласти окремі закони розподілу для випадкових величин X і Y .

Розв'язання. Випадкова величина X може набувати значень 1, 2, 3, 4, а випадкова величина Y — значень 0, 1, 2. Обчислимо $p_{i,j}$, $i = 1, 4$, $j = 0, 1, 2$ — ймовірності того, що $X = i$, $Y = j$:

$$p_{11} = \mathbf{P}(X = 1, Y = 1) = \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{10};$$

$$p_{12} = \mathbf{P}(X = 1, Y = 2) = \frac{2}{5} \cdot \frac{1}{4} \cdot 1 = \frac{1}{10};$$

$$p_{21} = \mathbf{P}(X = 2, Y = 1) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{2}{10};$$

$$p_{22} = \mathbf{P}(X = 2, Y = 2) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot 1 = \frac{1}{10};$$

$$p_{31} = \mathbf{P}(X = 3, Y = 1) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{10};$$

$$p_{32} = \mathbf{P}(X = 3, Y = 2) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{10};$$

$$p_{41} = \mathbf{P}(X = 4, Y = 1) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot 0 = 0;$$

$$p_{42} = \mathbf{P}(X = 4, Y = 2) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot 0 = 0;$$

$$p_{10} = \mathbf{P}(X = 1, Y = 0) = 0;$$

$$p_{20} = \mathbf{P}(X = 2, Y = 0) = 0;$$

$$p_{30} = \mathbf{P}(X = 3, Y = 0) = 0;$$

$$p_{40} = \mathbf{P}(X = 4, Y = 0) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot 1 = \frac{1}{10}.$$

Розділ 9. Системи двох випадкових величин

Отримані значення запишемо в таблицю розподілу ймовірностей системи випадкових величин (X, Y) (табл. 2).

Таблиця 2

Y	X				$\sum_{i=1}^4 p_{ij}$
	1	2	3	4	
0	0	0	0	$\frac{1}{10}$	$\frac{1}{10}$
1	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	0	$\frac{6}{10}$
2	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	0	$\frac{3}{10}$
$\sum_{j=1}^3 p_{ij}$	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\sum_{i=1}^4 \sum_{j=1}^3 p_{ij} = 1$

Із отриманої табл. 2 досить легко записати безумовні закони розподілу для випадкових величин X і Y :

$$\begin{array}{c|c|c|c|c} X & 1 & 2 & 3 & 4 \\ \hline \mathbf{P} & \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10} \end{array}; \quad \begin{array}{c|c|c} Y & 0 & 1 & 2 \\ \hline \mathbf{P} & \frac{1}{10} & \frac{6}{10} & \frac{3}{10} \end{array}. \quad \square$$

Приклад 2. Задано таблицю розподілу двовимірної випадкової величини (X, Y) :

$Y \backslash X$	1	2	3
-1	0,2	0,1	0,3
0	0,15	0,15	0,1

Потрібно:

- а) знайти безумовні закони розподілу двовимірної випадкової величини (X, Y) ;
- б) знайти умовний закон розподілу X за умови, що $Y = -1$;
- в) знайти умовний закон розподілу Y за умови, що $X = 3$;
- г) з'ясувати, залежні чи ні випадкові величини X і Y .

Розв'язання. а) Обчислимо ймовірності $\mathbf{P}(X = i)$ ($i = 1, 2, 3$) і $\mathbf{P}(Y = j)$ ($j = -1, 0$):

$$\begin{aligned}\mathbf{P}(x = 1) &= 0,2 + 0,15 = 0,35; \\ \mathbf{P}(x = 2) &= 0,1 + 0,15 = 0,25; \\ \mathbf{P}(x = 3) &= 0,3 + 0,1 = 0,4; \\ \mathbf{P}(y = -1) &= 0,2 + 0,1 + 0,3 = 0,6; \\ \mathbf{P}(y = 0) &= 0,15 + 0,15 + 0,1 = 0,4.\end{aligned}$$

Запишемо безумовні закони розподілу X і Y :

X		1		2		3	;	Y		-1		0
\mathbf{P}		0,35		0,25		0,4		\mathbf{P}		0,6		0,4

б) Оскільки

$$\begin{aligned}\mathbf{P}(X = 1/Y = -1) &= \frac{0,2}{0,2 + 0,1 + 0,3} = \frac{1}{3}; \\ \mathbf{P}(X = 2/Y = -1) &= \frac{0,1}{0,2 + 0,1 + 0,3} = \frac{1}{6}; \\ \mathbf{P}(X = 3/Y = -1) &= \frac{0,3}{0,2 + 0,1 + 0,3} = \frac{1}{2},\end{aligned}$$

умовний закон розподілу X за умови, що $Y = -1$, матиме такий вигляд:

X		1		2		3	.
$\mathbf{P}(X/Y = -1)$		$\frac{1}{3}$		$\frac{1}{6}$		$\frac{1}{2}$	

в) Оскільки

$$\begin{aligned}\mathbf{P}(Y = -1/X = 3) &= \frac{0,3}{0,3 + 0,1} = \frac{3}{4}; \\ \mathbf{P}(Y = 0/X = 3) &= \frac{0,1}{0,3 + 0,1} = \frac{1}{4},\end{aligned}$$

умовний закон розподілу Y за умови, що $X = 3$, матиме такий вигляд:

Y	-1	0
$\mathbf{P}(Y/X = 3)$	$\frac{3}{4}$	$\frac{1}{4}$

г) Той факт, що безумовний закон розподілу величини X не збігається з умовним законом розподілу цієї величини, свідчить про те, що величини X і Y залежні. □

Приклад 3. Система випадкових величин (X, Y) має такий розподіл імовірностей:

$Y \backslash X$	0	1
0	$0,2$	$0,15$
1	$0,15$	$0,15$
2	$0,1$	$0,25$

Знайти:

- а) математичні сподівання $\mathbf{M}(X)$ і $\mathbf{M}(Y)$;
- б) дисперсії $\mathbf{D}(X)$ і $\mathbf{D}(Y)$.

Розв'язання. а) Щоб відшукати математичне сподівання $\mathbf{M}(X)$, можна скористатися формулою

$$\mathbf{M}(X) = \sum_i \sum_j x_i p_{ij},$$

звідки

$$\mathbf{M}(X) = 0 \cdot 0,2 + 0 \cdot 0,15 + 0 \cdot 0,1 + 1 \cdot 0,15 + 1 \cdot 0,15 + 1 \cdot 0,25 = 0,55.$$

Такий самий результат отримаємо, якщо запишемо безумовний закон розподілу випадкової величини X :

X	0	1
\mathbf{P}	$0,45$	$0,55$

і далі скористаємося формулою

$$\mathbf{M}(X) = \sum_i x_i p_i,$$

звідки

$$\mathbf{M}(X) = 0 \cdot 0,45 + 1 \cdot 0,55 = 0,55.$$

Аналогічно обчислюємо математичне сподівання

$$\mathbf{M}(Y) = 0 \cdot 0,2 + 0 \cdot 0,5 + 1 \cdot 0,15 + 1 \cdot 0,15 + 2 \cdot 0,1 + 2 \cdot 0,25 = 1,$$

або з безумовного закону розподілу

Y	0	1	2
P	0,35	0,3	0,35

отримуємо

$$\mathbf{M}(Y) = 0 \cdot 0,35 + 1 \cdot 0,3 + 2 \cdot 0,35 = 1.$$

б) Щоб відшукати дисперсію $\mathbf{D}(X)$, можна скористатися формулою

$$\mathbf{D}(X) = \sum_i \sum_j (x_i - \mathbf{M}(X))^2 p_{ij},$$

звідки

$$\begin{aligned} \mathbf{D}(X) &= (0 - 0,55)^2 \cdot 0,2 + (0 - 0,55)^2 \cdot 0,15 + (0 - 0,55)^2 \cdot 0,1 + \\ &\quad + (1 - 0,55)^2 \cdot 0,15 + (1 - 0,55)^2 \cdot 0,15 + (1 - 0,55)^2 \cdot 0,25 = \\ &= 0,0605 + 0,045375 + 0,03025 + 0,030375 + 0,030375 + \\ &\quad + 0,050625 = 0,2475. \end{aligned}$$

Такий самий результат отримаємо, якщо скористаємося безумовним законом розподілу ймовірностей випадкової величини X і формулою

$$\mathbf{D}(X) = \sum_i (x_i - \mathbf{M}(X))^2 p_i.$$

Отримаємо

$$\begin{aligned} \mathbf{D}(X) &= (0 - 0,55)^2 \cdot 0,45 + (1 - 0,55)^2 \cdot 0,55 = \\ &= 0,136125 + 0,111375 = 0,2475. \end{aligned}$$

Аналогічно обчислюється дисперсія $\mathbf{D}(Y)$. Скористаємося безумовним законом розподілу ймовірностей випадкової величини Y і формулою

$$\mathbf{D}(Y) = \sum_i (y_i - \mathbf{M}(Y))^2 p_i.$$

Отримаємо

$$\begin{aligned} \mathbf{D}(Y) &= (0 - 1)^2 \cdot 0,35 + (1 - 1)^2 \cdot 0,3 + (2 - 1)^2 \cdot 0,35 = \\ &= 0,35 + 0 + 0,35 = 0,7. \end{aligned} \quad \square$$

Приклад 4. По мішені проводиться один постріл. Імовірність влучення дорівнює p . Розглядається X — кількість влучень, Y — кількість промахів.

Побудувати функцію розподілу $F(x, y)$ двовимірної випадкової величини (X, Y) .

Розв'язання. Знайдемо ймовірності:

$$\begin{aligned} \mathbf{P}(X = 0/Y = 0) &= 0; & \mathbf{P}(X = 1/Y = 0) &= p; \\ \mathbf{P}(X = 0/Y = 1) &= 1 - p; & \mathbf{P}(X = 1/Y = 1) &= 0. \end{aligned}$$

Побудуємо таблицю розподілу ймовірностей системи випадкових величин (X, Y) :

$Y \backslash X$	0	1
0	0	p
1	$1 - p$	0

Функцію розподілу ймовірностей системи випадкових величин зручно подати у вигляді таблиці:

$F(x, y)$	$x \leq 0$	$0 < x \leq 1$	$x > 1$
$y \leq 0$	0	0	0
$0 < y \leq 1$	0	0	p
$y > 1$	0	$1 - p$	1

□

Приклад 5. Система дискретних випадкових величин (X, Y) задана таблицею розподілу ймовірностей (табл. 3).

Таблиця 3

Y	X				$\sum_{i=1}^4 p_{ij}$
	-10	-8	-6	-4	
10	0,023	0,027	0,05	0,1	0,2
20	0,05	0,1	0,025	0,025	0,2
30	0,05	0,05	0,025	0,025	0,15
40	0,027	0,023	0,05	0,35	0,45
$\sum_{j=1}^4 p_{ij}$	0,15	0,2	0,15	0,5	$\sum_{i=1}^4 \sum_{j=1}^4 p_{ij} = 1$

Обчислити математичні сподівання $\mathbf{M}(X)$ і $\mathbf{M}(Y)$, середні квадратичні відхилення $\sigma(X)$ і $\sigma(Y)$, кореляційний момент $\text{cov}(X, Y)$, коефіцієнт кореляції $r(X, Y)$, а також умовні математичні сподівання $\mathbf{M}(X/Y = 40)$ і $\mathbf{M}(Y/X = -8)$.

Розв'язання. Побудуємо безумовні закони розподілу X і Y :

X	-10	-8	-6	-4	Y	10	20	30	40
\mathbf{P}	0,15	0,2	0,15	0,5	\mathbf{P}	0,2	0,2	0,15	0,45

Обчислимо математичні сподівання за формулами

$$\mathbf{M}(X) = \sum_{i=1}^4 x_i p_i, \quad \mathbf{M}(Y) = \sum_{i=1}^4 y_i p_i.$$

Отже,

$$\mathbf{M}(X) = -1,5 - 1,6 - 0,9 - 2 = -5,$$

$$\mathbf{M}(Y) = 2 + 4 + 4,5 + 18 = 28,5.$$

Середнє квадратичне відхилення

$$\sigma(X) = \sqrt{\mathbf{D}(X)},$$

де $\mathbf{D}(X)$ — дисперсія випадкової величини X :

$$\begin{aligned} \mathbf{D}(X) &= (-10 + 5)^2 \cdot 0,15 + (-8 + 5)^2 \cdot 0,2 + (-6 + 5)^2 \cdot 0,15 + \\ &+ (-4 + 5)^2 \cdot 0,5 = 6,2. \end{aligned}$$

Отже,

$$\sigma(X) = \sqrt{6,2} \approx 2,49.$$

Аналогічно обчислюємо $\mathbf{D}(Y)$ і $\sigma(Y)$:

$$\begin{aligned} \mathbf{D}(Y) &= (10 - 28,5)^2 \cdot 0,2 + (20 - 28,5)^2 \cdot 0,2 + (30 - 28,5)^2 \cdot 0,15 + \\ &+ (40 - 28,5)^2 \cdot 0,45 = 136,4; \end{aligned}$$

$$\sigma(Y) = \sqrt{136,4} \approx 11,68.$$

Кореляційний момент

$$\text{cov}(X, Y) = \mathbf{M}(XY) - \mathbf{M}(X)\mathbf{M}(Y),$$

де

$$\begin{aligned} \mathbf{M}(XY) &= \sum_i \sum_j x_i y_j p_{ij} = \\ &= -2,3 - 80 \cdot 0,027 - 60 \cdot 0,05 - 40 \cdot 0,1 - \\ &- 200 \cdot 0,05 - 160 \cdot 0,1 - 120 \cdot 0,025 - \\ &- 80 \cdot 0,025 - 30 \cdot 0,05 - 240 \cdot 0,05 - \\ &- 180 \cdot 0,025 - 120 \cdot 0,025 - 400 \cdot 0,027 - \\ &- 320 \cdot 0,023 - 240 \cdot 0,05 - 160 \cdot 0,35 = \\ &= -149,62; \end{aligned}$$

$$\mathbf{M}(X)\mathbf{M}(Y) = -5 \cdot 28,5 = -142,5.$$

Отже,

$$\text{cov}(X, Y) = -149,62 - (-142,5) = -7,12.$$

Коефіцієнт кореляції

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Отже,

$$r(X, Y) = \frac{-7,12}{2,49 \cdot 11,68} = -0,24.$$

Оскільки $r(X, Y) \neq 0$, можемо зробити висновок про корельованість випадкових величин X і Y .

Для обчислення $\mathbf{M}(X/Y = 40)$ будемо умовний закон розподілу для випадкової величини X при $Y = 40$. Застосовуючи формулу

$$\mathbf{P}(X = x_i / Y = y_j) = \frac{\mathbf{P}(X = x_i, Y = y_j)}{\mathbf{P}(Y = y_j)},$$

маємо

X	-10	-8	-6	-4
$\mathbf{P}(X/Y = 40)$	$\frac{27}{450}$	$\frac{23}{450}$	$\frac{50}{450}$	$\frac{50}{450}$

Отже,

$$\begin{aligned} \mathbf{M}(X/Y = 40) &= \sum_{i=1}^4 x_i \mathbf{P}(X = x_i / Y = 40) = \\ &= \frac{1}{450} (-10 \cdot 27 - 8 \cdot 23 - 6 \cdot 50 - 4 \cdot 350) = \\ &= -\frac{2154}{450} \approx -4,79. \end{aligned}$$

Аналогічно обчислюємо $\mathbf{M}(Y/X = -8)$:

Y	10	20	30	40
$\mathbf{P}(Y/X = -8)$	$\frac{27}{200}$	$\frac{100}{200}$	$\frac{50}{200}$	$\frac{23}{200}$

$$\mathbf{M}(Y/X = -8) = \frac{1}{200} (270 + 2000 + 1500 + 920) = \frac{4690}{200} = 23,45. \quad \square$$

Приклад 6. Система випадкових величин (X, Y) має щільність розподілу ймовірностей

$$f(x, y) = a(xy + y^2),$$

де $(x, y) \in \mathcal{D} = \{(x, y): 0 \leq x \leq 1; 0,5x \leq y \leq x\}$.

Знайти значення константи a , математичні сподівання $\mathbf{M}(X)$ і $\mathbf{M}(Y)$, середні квадратичні відхилення $\sigma(X)$ і $\sigma(Y)$, кореляційний момент $\text{cov}(X, Y)$ і коефіцієнт кореляції $r(X, Y)$.

Розв'язання. Щоб відшукати константу a , скористаємося характеристичною властивістю функції щільності розподілу ймовірностей $f(x, y)$:

$$\begin{aligned} \iint_{\mathbb{R}^2} f(x, y) dx dy &= 1; \\ \iint_{\mathcal{D}} a(xy + y^2) dx dy &= a \iint_{\mathcal{D}} (xy + y^2) dx dy = a \int_0^1 \int_{0,5x}^x (xy + y^2) dx dy = \\ &= a \int_0^1 \left(\int_{0,5x}^x (xy + y^2) dy \right) dx = \\ &= a \int_0^1 \left(\frac{xy^2}{2} + \frac{y^3}{3} \right) \Big|_{0,5x}^x dx = \\ &= a \int_0^1 \left(\frac{x^3}{2} + \frac{x^3}{3} - \frac{0,25x^3}{2} - \frac{0,125x^3}{3} \right) dx = \\ &= a \int_0^1 \frac{4x^3}{6} dx = \frac{4a}{6} \cdot \frac{x^4}{4} \Big|_0^1 = \frac{a}{6}, \end{aligned}$$

звідки

$$\frac{a}{6} = 1,$$

тому

$$a = 6.$$

Функцію $f(x, y)$ можна записати так:

$$f(x, y) = 6(xy + y^2).$$

Математичне сподівання $\mathbf{M}(X)$ обчислюємо за формулою

$$\mathbf{M}(X) = \iint_{\mathbb{R}^2} xf(x, y) dx dy.$$

Отримуємо

$$\begin{aligned} \mathbf{M}(X) &= \int_0^1 \int_{0,5x}^x 6x(xy + y^2) dx dy = 6 \int_0^1 \left(\int_{0,5x}^x (x^2y + xy^2) dy \right) dx = \\ &= 6 \int_0^1 \left(\frac{x^2y^2}{2} + \frac{xy^3}{3} \right) \Big|_{0,5x}^x dx = \\ &= 6 \int_0^1 \left(\frac{x^4}{2} + \frac{x^4}{3} - \frac{0,25x^4}{2} - \frac{0,125x^4}{3} \right) dx = \\ &= 6 \int_0^1 \frac{4}{6} x^4 dx = 4 \cdot \frac{x^5}{5} \Big|_0^1 = \frac{4}{5} = 0,8. \end{aligned}$$

Аналогічно обчислюємо $\mathbf{M}(Y)$:

$$\mathbf{M}(Y) = \iint_{\mathbb{R}^2} yf(x, y) dx dy.$$

$$\begin{aligned} \mathbf{M}(Y) &= \int_0^1 \int_{0,5x}^x 6y(xy + y^2) dx dy = 6 \int_0^1 \left(\int_{0,5x}^x (xy^2 + y^3) dy \right) dx = \\ &= 6 \int_0^1 \left(\frac{xy^3}{3} + \frac{y^4}{4} \right) \Big|_{0,5x}^x dx = \end{aligned}$$

$$\begin{aligned}
 &= 6 \int_0^1 \left(\frac{x^4}{3} + \frac{x^4}{4} - \frac{0,125x^4}{3} - \frac{0,0625x^4}{4} \right) dx = \\
 &= 6 \int_0^1 \frac{6,3125}{12} x^4 dx = \frac{1}{2} \cdot 6,3125 \cdot \frac{x^5}{5} \Big|_0^1 = 0,63125.
 \end{aligned}$$

Середнє квадратичне відхилення

$$\sigma(X) = \sqrt{\mathbf{D}(X)},$$

де $\mathbf{D}(X)$ — дисперсія випадкової величини X , яка обчислюється за формулою

$$\mathbf{D}(X) = \iint_{\mathbb{R}^2} x^2 f(x, y) dx dy - (\mathbf{M}(X))^2.$$

Отже,

$$\begin{aligned}
 \mathbf{D}(X) &= 6 \int_0^1 \int_{0,5x}^x x^2 (xy + y^2) dx dy - 0,64 = \\
 &= 6 \int_0^1 \left(\frac{x^3 y^2}{2} + \frac{x^2 y^3}{3} \right) \Big|_{0,5x}^x dx - 0,64 = \\
 &= 6 \int_0^1 \left(\frac{x^5}{2} + \frac{x^5}{3} - \frac{0,25x^5}{2} - \frac{0,125x^5}{3} \right) dx - 0,64 = \\
 &= \frac{4x^6}{6} \Big|_0^1 - 0,64 = \frac{2}{3} - 0,64 = \frac{2}{75}; \\
 \sigma(X) &= \sqrt{\frac{2}{75}} \approx 0,163.
 \end{aligned}$$

Аналогічно обчислюємо $\sigma(Y)$:

$$\mathbf{D}(Y) = \iint_{\mathbb{R}^2} y^2 f(x, y) dx dy - (\mathbf{M}(Y))^2.$$

$$\begin{aligned}
 \mathbf{D}(Y) &= 6 \int_0^1 \int_{0,5x}^x y^2(xy + y^2) dx dy - 0,398 = \\
 &= 6 \int_0^1 \left(\frac{xy^4}{4} + \frac{y^5}{5} \right) \Big|_{0,5x}^x dx - 0,398 = \\
 &= 6 \int_0^1 \left(\frac{x^5}{4} + \frac{x^5}{5} - \frac{0,0625x^5}{4} - \frac{0,03125x^5}{5} \right) dx - 0,398 \approx \\
 &\approx 0,428; \\
 \sigma(X) &= \sqrt{0,428} \approx 0,654.
 \end{aligned}$$

Кореляційний момент

$$\text{cov}(X, Y) = \mathbf{M}[(X - \mathbf{M}(X))(Y - \mathbf{M}(Y))].$$

В інтегральному поданні маємо

$$\begin{aligned}
 \text{cov}(X, Y) &= \iint_{\mathbb{R}^2} xyf(x, y) dx dy - \mathbf{M}(X)\mathbf{M}(Y) = \\
 &= 6 \int_0^1 \int_{0,5x}^x xy(xy + y^2) dx dy - 0,8 \cdot 0,63125 = \\
 &= 6 \int_0^1 \left(\frac{x^2y^3}{3} + \frac{xy^4}{4} \right) \Big|_{0,5x}^x dx - 0,505 = \\
 &= 6 \int_0^1 \left(\frac{x^5}{3} + \frac{x^5}{4} - \frac{0,125x^5}{3} - \frac{0,0625x^5}{4} \right) dx - 0,505 = \\
 &= \frac{6,3125}{12} x^6 \Big|_0^1 - 0,505 \approx 0,021.
 \end{aligned}$$

За формулою

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

обчислимо коефіцієнт кореляції:

$$r(X, Y) = \frac{0,021}{0,163 \cdot 0,654} \approx 0,2.$$

Отже,

$$r(X, Y) = 0,2,$$

а це означає, що випадкові величини X і Y корельовані. □

Приклад 7. Система випадкових величин (X, Y) має рівномірний розподіл імовірностей в області

$$\mathcal{D} = \{(x, y) : x \in [0; 1], y \in [0; 1]\}.$$

За межами області \mathcal{D} щільність розподілу ймовірностей дорівнює нулю.

Записати функцію щільності розподілу ймовірностей $f(x, y)$, інтегральну функцію розподілу ймовірностей $F(x, y)$.

Розв'язання. Відповідно до умови задачі в загальному вигляді диференціальну функцію $f(x, y)$ можна записати так:

$$f(x, y) = \begin{cases} c, & (x, y) \in \mathcal{D}; \\ 0, & (x, y) \notin \mathcal{D}. \end{cases}$$

За характеристичною властивістю

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1$$

і враховуючи, що за межами області \mathcal{D} щільність розподілу ймовірностей дорівнює нулю, можна записати

$$\iint_{\mathcal{D}} f(x, y) dx dy = 1.$$

Геометрична інтерпретація останньої рівності така: об'єм паралелепіпеда, обмеженого знизу областю \mathcal{D} і зверху графіком функції $f(x, y)$, дорівнює одиниці. Оскільки система випадкових величин (X, Y) за умовою задачі має рівномірний розподіл, геометричну інтерпретацію можна деталізувати так: об'єм прямокутного паралелепіпеда, в основі якого лежить прямокутник $OABC$ (рис. 25), дорівнює одиниці.

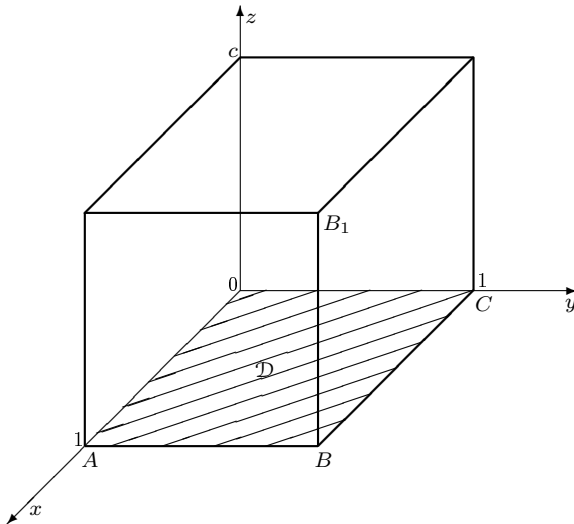


Рис. 25

Довжина відрізка BB_1 — висоти прямокутного паралелепіпеда — дорівнює значенню c у записі функції щільності $f(x, y)$.

Отже, очевидним є спосіб відшукування c :

$$V_{\text{парал}} = 1.$$

Оскільки

$$V_{\text{парал}} = S_{\text{осн}} \cdot H,$$

де $S_{\text{осн}}$ — площа основи паралелепіпеда (прямокутника $OABC$); H —

висота паралелепіпеда ($H = BB_1 = c$),

$$S_{OABC} = AO \cdot OC = 1 \cdot 1 = 1;$$

$$V_{\text{парал}} = 1 \cdot BB_1 = c.$$

Отже,

$$c = 1.$$

Функцію $f(x, y)$ можна записати так:

$$f(x, y) = \begin{cases} 1, & (x, y) \in \mathcal{D}; \\ 0, & (x, y) \notin \mathcal{D}. \end{cases}$$

За означенням функції розподілу ймовірності

$$F(x, y) = \mathbf{P}(X < x, Y < y).$$

З геометричної точки зору це ймовірність потрапляння системи випадкових величин $(X; Y)$ у нескінченний квадрант з вершиною в точці $(x; y)$ (рис. 26).

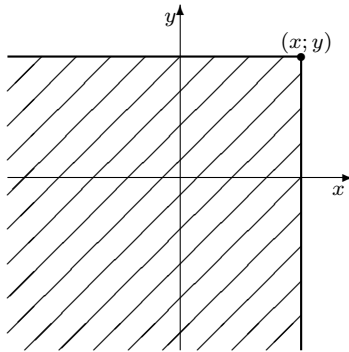


Рис. 26

Відповідно до умови задачі при $x \leq 0$ і $y \leq 0$

$$F(x, y) = 0,$$

оскільки за такої умови $f(x, y) = 0$.

Якщо $0 < x \leq 1$ і $0 < y \leq 1$, то для відшукування $F(x, y)$ необхідно знайти об'єм прямокутного паралелепіпеда з основою $OA'B'C'$ (рис. 27).

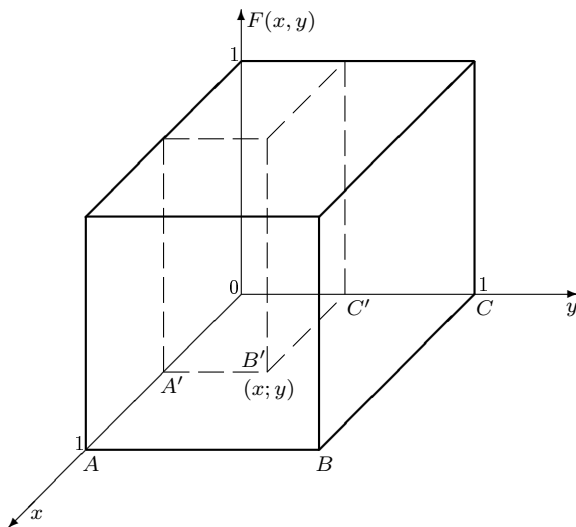


Рис. 27

Зазначимо, що т. B' має координати $(x; y)$, через які визначимо шуканий об'єм. Очевидно, що площа прямокутника $OA'B'C'$

$$S_{OA'B'C'} = xy.$$

Ураховуючи, що висота паралелепіпеда дорівнює одиниці, можна записати, що

$$F(x, y) = xy, \quad \text{якщо } 0 < x \leq 1 \text{ і } 0 < y \leq 1.$$

Зрозуміло, що у випадку $0 < x \leq 1$ і $y > 1$ (рис. 28) маємо

$$F(x, y) = x,$$

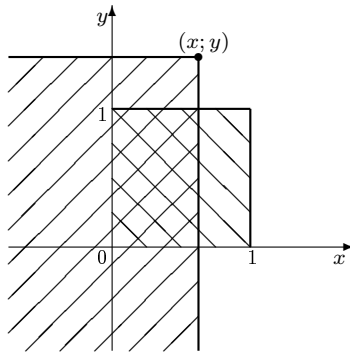


Рис. 28

а при $x > 1$ і $0 < y \leq 1$ (рис. 29)

$$F(x, y) = y.$$

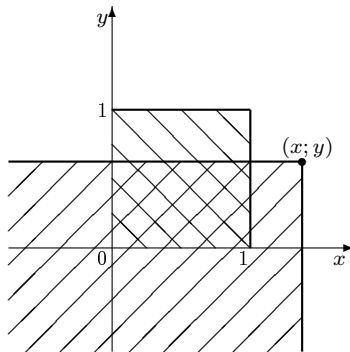


Рис. 29

Якщо $x > 1$ і $y > 1$ (рис. 30), то

$$F(x, y) = 1,$$

оскільки об'єм прямокутного паралелепіпеда з основою $OABC$ дорівнює одиниці.

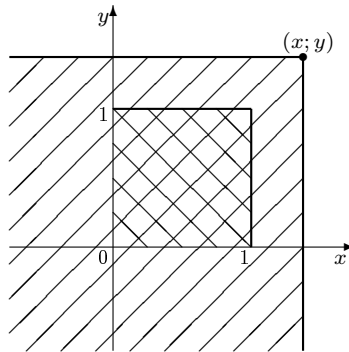


Рис. 30

Остаточно можна записати вираз для інтегральної функції розподілу ймовірностей:

$$F(x, y) = \begin{cases} 0, & x \leq 0 \text{ або } y \leq 0; \\ x, & 0 < x \leq 1 \text{ і } y > 1; \\ y, & x > 1 \text{ і } 0 < y \leq 1; \\ xy, & 0 < x \leq 1 \text{ і } 0 < y \leq 1; \\ 1, & x > 1 \text{ і } y > 1. \end{cases}$$

Такий самий результат можна отримати, не вдаючись до геометричних міркувань, а застосовуючи лише означення $F(x, y)$ і апарат інтегрального числення:

а) якщо $0 < x \leq 1$ і $0 < y \leq 1$, то

$$\begin{aligned} F(x, y) &= \mathbf{P}(X < x, Y < y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy = \\ &= \int_{-\infty}^0 \int_{-\infty}^0 0 dx dy + \int_0^x \int_0^y 1 dx dy = \int_0^x \int_0^y dx dy = xy; \end{aligned}$$

б) якщо $0 < x \leq 1$ і $y > 1$, то

$$\begin{aligned} F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy = \int_{-\infty}^0 \int_{-\infty}^0 0 dx dy + \int_0^x \int_0^y f(x, y) dx dy = \\ &= \int_0^x \int_0^1 1 dx dy + \int_0^x \int_1^y 0 dx dy = \int_0^x \int_0^1 dx dy = x; \end{aligned}$$

в) якщо $x > 1$ і $0 < y \leq 1$, то аналогічно отримуємо

$$F(x, y) = \int_0^1 \int_0^y dx dy = y;$$

г) якщо $x > 1$ і $y > 1$, то

$$\begin{aligned} F(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy = \\ &= \int_{-\infty}^0 \int_{-\infty}^0 0 dx dy + \int_0^1 \int_0^1 1 dx dy + \int_1^x \int_1^y 0 dx dy = 1, \end{aligned}$$

що й підтверджує правильність попереднього розв'язку. □

Приклад 8. Система випадкових величин (X, Y) має рівномірний розподіл ймовірностей в області

$$\mathcal{D} = \{(x, y) : |x| + |y| \leq 1\}.$$

За межами області \mathcal{D} щільність розподілу ймовірностей дорівнює нулю.

Записати вирази для диференціальної функції розподілу $f(x, y)$ та інтегральної функції $F(x, y)$.

Розв'язання. Область \mathcal{D} є квадратом $ABCD$ з вершинами на осях координат (рис. 31):

$$A(1; 0), \quad B(0; 1), \quad C(-1; 0), \quad D(0; -1).$$

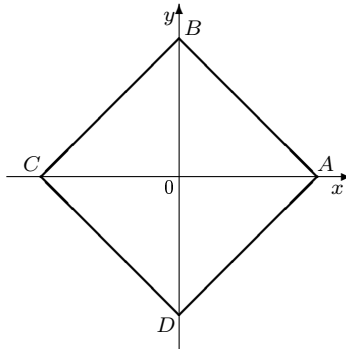


Рис. 31

Оскільки в межах області \mathcal{D} розподіл ймовірностей рівномірний, у загальному вигляді функція щільності запишеться так:

$$f(x, y) = \begin{cases} c, & (x, y) \in \mathcal{D}; \\ 0, & (x, y) \notin \mathcal{D}. \end{cases}$$

Значення c визначимо, використовуючи характеристичну властивість функції щільності $f(x, y)$, а саме:

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1.$$

Для даного прикладу можемо записати

$$\iint_{\mathcal{D}} c dx dy = 1.$$

Ураховуючи рівномірність розподілу та геометричну інтерпретацію останнього інтеграла, значення c знайдемо як довжину висоти

прямокутного паралелепіеда, основою якого є квадрат $ABCD$:

$$V_{\text{парал}} = S_{ABCD} \cdot H, \quad H = \frac{V_{\text{парал}}}{S_{ABCD}}.$$

Оскільки

$$S_{ABCD} = AB \cdot CD = \sqrt{2} \cdot \sqrt{2} = 2,$$

то

$$c = H = \frac{1}{2}.$$

Отже, функція $f(x, y)$ має вигляд

$$f(x, y) = \begin{cases} \frac{1}{2}, & (x, y) \in \mathcal{D}; \\ 0, & (x, y) \notin \mathcal{D}. \end{cases}$$

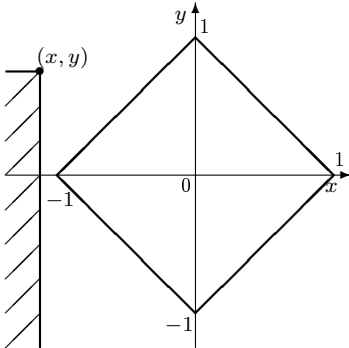
Функцію $F(x, y)$ визначатимемо, спираючись на означення інтегральної функції розподілу ймовірностей

$$F(x, y) = \mathbf{P}(X < x, Y < y)$$

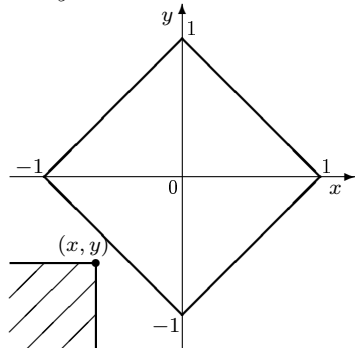
і геометричні міркування, пов'язані з подвійними інтегралами.

Щоб отримати вираз функції $F(x, y)$, необхідно розглянути такі варіанти розміщення точки $f(x, y)$ відносно області \mathcal{D} :

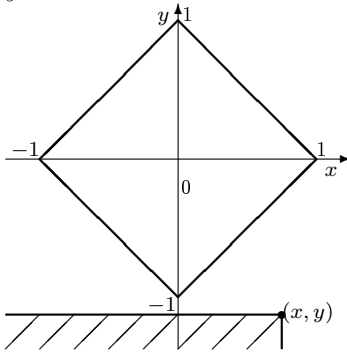
а) $x \leq -1$



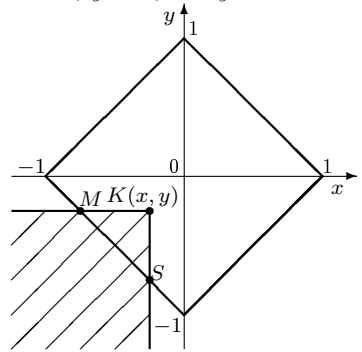
б) $x + y \leq -1$



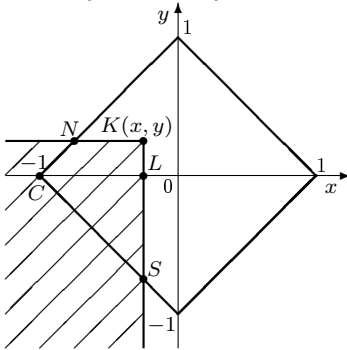
в) $y \leq -1$



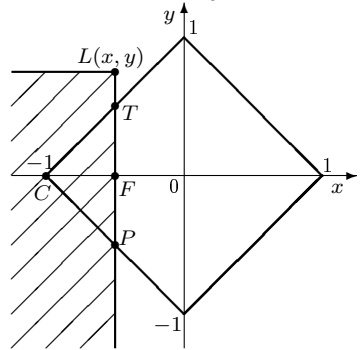
г) $x \leq 0, y \leq 0, x + y > -1$



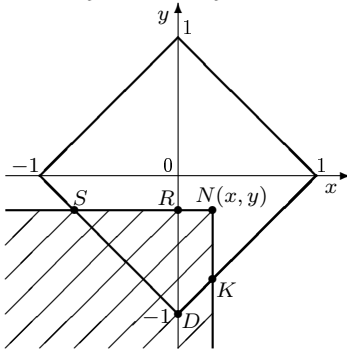
д) $x \leq 0, y > 0, x - y > -1$



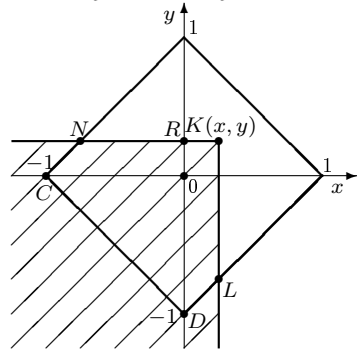
е) $-1 < x \leq 0, x - y \leq -1$



е) $x > 0, y \leq 0, x - y \leq 1$

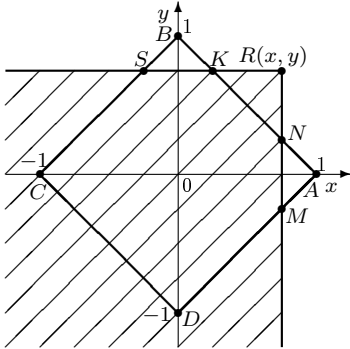


ж) $x > 0, y > 0, x + y \leq 1$

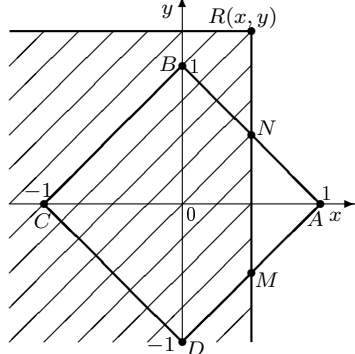


Розділ 9. Системи двох випадкових величин

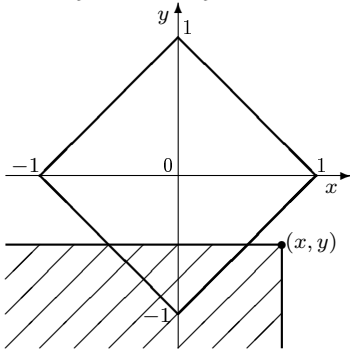
з) $x \leq 1, y \leq 1, x + y > 1$



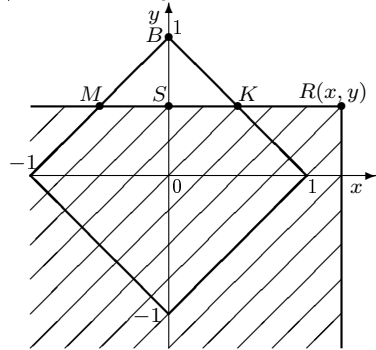
и) $0 < x \leq 1, y > 1$



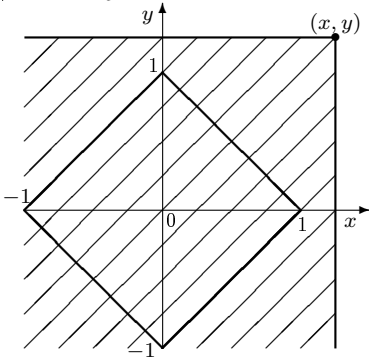
і) $-1 < y \leq 0, x - y > 1$



ї) $x > 1, 0 < y \leq 1$



й) $x > 1, y > 1$



Очевидно, що у випадках а), б) і в)

$$F(x, y) = 0,$$

оскільки згідно з умовою задачі поза межами області \mathcal{D}

$$f(x, y) = 0$$

і жодна частина області \mathcal{D} із нескінченним квадрантом не перетинається.

Для того щоб записати вираз $F(x, y)$ у випадку г), необхідно через координати точки K виразити об'єм прямої призми з основою MKS і висотою $\frac{1}{2}$:

$$V_{\text{призми}} = S_{MKS} \cdot H.$$

Рівняння прямої MS має вигляд $-x - y = 1$, або $y = -x - 1$. Якщо (x, y) — координати точки K , то довжину відрізка KS можна знайти як різницю ординати точок S і K , тобто

$$KS = y - (-x - 1) = x + y + 1.$$

Оскільки $MK = KS$, то

$$S_{MKS} = \frac{1}{2} \cdot MK \cdot KS = \frac{(x + y + 1)^2}{2}.$$

Отже,

$$V_{\text{призми}} = \frac{(x + y + 1)^2}{2} \cdot \frac{1}{2} = \frac{(x + y + 1)^2}{4}.$$

Остаточно для випадку г) функція $F(x, y)$ має вигляд

$$F(x, y) = \frac{(x + y + 1)^2}{4}.$$

У випадку д) необхідно обчислити об'єм прямої призми з основою $KNCS$ і висотою $\frac{1}{2}$, виразивши його через координати (x, y)

точки K . Для цього розіб'ємо фігуру $KNCS$ на дві: $LKNC$ і CLS . Довжина катетів $LS = LC$ рівнобедреного трикутника

$$LC = x - (-1) = x + 1.$$

Отже,

$$S_{CLB} = \frac{1}{2}(x+1)(x+1) = \frac{(x+1)^2}{2}.$$

Фігура $CLKN$ є трапецією з основами

$$CL = x + 1, \quad KN = x - y + 1$$

(використано рівняння прямої CB : $-x + y = 1$ і той факт, що точка N належить прямій CB) і висотою $KL = y$. Ураховуючи зазначене, маємо

$$\begin{aligned} S_{CLKN} &= \frac{x+1+x-y+1}{2} \cdot y = \frac{2x-y+2}{2} \cdot y; \\ V_{\text{призми}} &= \frac{1}{2} \cdot \left(\frac{(x+1)^2}{2} + \frac{2xy-y^2+2y}{2} \right) = \\ &= \frac{1}{4} \left((x+1)^2 + 2y(x+1) - y^2 \right) = \\ &= \frac{1}{4} \left((x+y+1)^2 - 2y^2 \right). \end{aligned}$$

Отже, для випадку д) функція $F(x, y)$ має вигляд

$$F(x, y) = \frac{1}{4} \left((x+y+1)^2 - 2y^2 \right).$$

Цілком аналогічна ситуація і у випадку е). Площа фігури $KNSD$ обчислюється як сума площ фігур $KNRD$ і DRS . При цьому використовуються координати точки $N(x, y)$, рівняння прямої AD : $x - y = 1$, факт належності точки K прямій AD і те, що SDR — рівнобедрений трикутник, а $DRNK$ — трапеція.

$$S_{DRS} = \frac{1}{2} \cdot DR \cdot RS, \quad DR = RS, \quad DR = y + 1.$$

Отже,

$$S_{DRS} = \frac{(y+1)^2}{2}.$$

$$S_{DRNK} = \frac{NK + RD}{2} \cdot NR, \quad RD = y+1, \quad NK = y-x+1, \quad NR = x.$$

Отже,

$$S_{DRNK} = \frac{y-x+1+y+1}{2} \cdot x = \frac{2y+2-x}{2} \cdot x;$$

$$\begin{aligned} V_{\text{призми}} &= \frac{1}{2} \cdot \left(\frac{(y+1)^2}{2} + \frac{2y+2-x}{2} \cdot x \right) = \\ &= \frac{1}{4} ((y+1)^2 + 2xy + 2x - x^2) = \frac{1}{4} ((x+y+1)^2 - 2x^2). \end{aligned}$$

Остаточно для випадку є) вираз функції $F(x, y)$ можна записати так:

$$F(x, y) = \frac{1}{4} ((x+y+1)^2 - 2x^2).$$

Розглядаючи випадок е), необхідно знайти об'єм призми з основою CPT . При цьому шуканий об'єм потрібно виразити через координати (x, y) точки L . Оскільки

$$S_{CPT} = \frac{1}{2} CF \cdot PT, \quad CF = x+1, \quad TP = 2TF = 2CF = 2(x+1),$$

площа трикутника CPT

$$S_{CPT} = \frac{1}{2}(x+1) \cdot 2(x+1) = (x+1)^2.$$

Отже,

$$V_{\text{призми}} = \frac{1}{2}(x+1)^2.$$

Вираз $F(x, y)$ для випадку е) буде таким:

$$F(x, y) = \frac{(x+1)^2}{2}.$$

Аналогічна ситуація у випадку і), лише x і y міняються ролями:

$$F(x, y) = \frac{(y + 1)^2}{2}.$$

Для того щоб записати вираз функції $F(x, y)$ у випадку з), необхідно виразити площу фігури $MNKSCD$ через координати (x, y) точки R . Зазначену площу обчислимо як різницю площ прямокутника $ABCD$ і двох трикутників AMN і KBS :

$$\begin{aligned} S_{ABCD} &= 2, \\ S_{ANM} &= \frac{1}{2} \cdot (1 - x) \cdot 2(1 - x) = (1 - x)^2, \\ S_{KBS} &= \frac{1}{2} \cdot (1 - y) \cdot 2(1 - y) = (1 - y)^2; \\ S_{MNKSCD} &= 2 - (1 - x)^2 - (1 - y)^2. \end{aligned}$$

Об'єм прямої призми з основою $MNKSCD$

$$V_{\text{призми}} = \frac{1}{2} \cdot S_{MNKSCD}.$$

Отже, у випадку з)

$$F(x, y) = \frac{1}{2}(2 - (1 - x)^2 - (1 - y)^2) = x + y - \frac{1}{2}(x^2 + y^2).$$

У випадку ж) необхідно обчислити об'єм прямої призми з основою $LKNCD$ через координати (x, y) точки K . Для цього знайдемо площу фігури $LKNCD$, розбивши її на три частини: $LKRD$, $ORNC$ і OCD :

$$\begin{aligned} S_{LKRD} &= \frac{y + 1 + y - x + 1}{2} \cdot x = \frac{2y + 2 - x}{2} \cdot x, \\ S_{ORNC} &= \frac{1 - y + 1}{2} \cdot y = \frac{2 - y}{2} \cdot y, \\ S_{OCD} &= \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}; \end{aligned}$$

$$\begin{aligned} S_{LKNCD} &= \frac{2y + 2 - x}{2} \cdot x + \frac{2 - y}{2} \cdot y + \frac{1}{2} = \\ &= \frac{1}{2}(2xy + 2x - x^2 + 2y - y^2 + 1) = x + y - \frac{(x - y)^2}{2} + \frac{1}{2}. \end{aligned}$$

Об'єм прямої призми з основою $LKNCD$

$$V_{\text{призми}} = \frac{1}{2} S_{LKNCD}.$$

Отже, у випадку ж)

$$F(x, y) = \frac{x + y}{2} - \frac{(x - y)^2}{4} + \frac{1}{4}.$$

У випадку и) необхідно обчислити об'єм прямої призми з основою $NBCDM$ через координати (x, y) точки R . Для цього знайдемо площу фігури $NBCDM$ як різницю площ квадрата $ABCD$ і трикутника ANM :

$$S_{ABCD} = 2,$$

$$S_{ANM} = (1 - x)^2;$$

$$S_{NBCDM} = 2 - (1 - x)^2.$$

Об'єм прямої призми з основою $NBCDM$

$$V_{\text{призми}} = \frac{1}{2} S_{NBCDM}.$$

Отже, у випадку и)

$$F(x, y) = 1 - \frac{(1 - x)^2}{2}.$$

Аналогічна ситуація у випадку і), лише x і y міняються ролями:

$$F(x, y) = 1 - \frac{(1 - y)^2}{2}.$$

Очевидно, що для випадку и)

$$x > 1, \quad y > 1$$

функція розподілу ймовірностей $F(x, y)$ дорівнює одиниці. \square

Приклад 9. Функція щільності розподілу ймовірностей системи випадкових величин (X, Y) має вигляд

$$f(x, y) = ae^{-2x^2+xy-y^2},$$

при $x \in \mathbb{R}$ і $y \in \mathbb{R}$.

Визначити константу a . Знайти $f(x)$, $f(y)$, $f(x/y)$, $f(y/x)$. Обчислити математичні сподівання $\mathbf{M}(X)$, $\mathbf{M}(Y)$, коефіцієнт кореляції $r(X, Y)$. З'ясувати чи корельовані випадкові величини X і Y .

Розв'язання. Значення константи a обчислимо, використовуючи характеристичну властивість функції щільності:

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1.$$

У нашому випадку

$$\begin{aligned} \iint_{\mathbb{R}^2} f(x, y) dx dy &= \iint_{\mathbb{R}^2} ae^{-2x^2+xy-y^2} dx dy = a \iint_{\mathbb{R}^2} e^{-2x^2+xy-y^2} dx dy = \\ &= a \iint_{\mathbb{R}^2} e^{-2x^2} \cdot e^{-(y^2-xy)} dx dy = \\ &= a \iint_{\mathbb{R}^2} e^{-2x^2} \cdot e^{-(y^2-2 \cdot \frac{1}{2}xy + \frac{x^2}{4} - \frac{x^2}{4})} dx dy = \\ &= a \iint_{\mathbb{R}^2} e^{-2x^2} \cdot e^{-(y-\frac{1}{2}x)^2 + \frac{x^2}{4}} dx dy = \\ &= a \iint_{\mathbb{R}^2} e^{-2x^2} \cdot e^{\frac{x^2}{4}} \cdot e^{-(y-\frac{x}{2})^2} dx dy = \\ &= a \int_{-\infty}^{+\infty} e^{-\frac{7}{4}x^2} dx \cdot \underbrace{\int_{-\infty}^{+\infty} e^{-(y-\frac{x}{2})^2} d\left(y - \frac{x}{2}\right)}_{=\sqrt{\pi}, \text{ як інтеграл Пуассона}} = \end{aligned}$$

$$\begin{aligned}
 &= a\sqrt{\pi} \int_{-\infty}^{+\infty} e^{-\frac{7}{4}x^2} dx = \\
 &= a\sqrt{\pi} \cdot \frac{2}{\sqrt{7}} \int_{-\infty}^{+\infty} e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} d\left(\frac{\sqrt{7}x}{2}\right) = \\
 &= \frac{2a\sqrt{\pi}}{\sqrt{7}} \int_{-\infty}^{+\infty} e^{-u^2} du = \frac{2a\sqrt{\pi}}{\sqrt{7}} \sqrt{\pi} = \frac{2a\pi}{\sqrt{7}}.
 \end{aligned}$$

Отже,

$$a \cdot \frac{2\pi}{\sqrt{7}} = 1,$$

звідки

$$a = \frac{\sqrt{7}}{2\pi}.$$

Функція щільності має вигляд

$$f(x, y) = \frac{\sqrt{7}}{2\pi} e^{-2x^2 + xy - y^2}, \quad x \in \mathbb{R}, \quad y \in \mathbb{R}.$$

Для того щоб знайти функцію щільності розподілу ймовірностей випадкової величини X , скористаємося формулою

$$f(x) = \int_{-\infty}^{+\infty} f(x, y) dy.$$

Отже,

$$\begin{aligned}
 f(x) &= \int_{-\infty}^{+\infty} \frac{\sqrt{7}}{2\pi} e^{-2x^2 + xy - y^2} dy = \frac{\sqrt{7}}{2\pi} e^{-2x^2} \int_{-\infty}^{+\infty} e^{-\left(y^2 - 2y \cdot \frac{1}{2}x + \frac{x^2}{4} - \frac{x^2}{4}\right)} dy = \\
 &= \frac{\sqrt{7}}{2\pi} e^{-2x^2} \int_{-\infty}^{+\infty} e^{-(y - \frac{x}{2})^2} \cdot e^{\frac{x^2}{4}} dy = \frac{\sqrt{7}}{2\pi} e^{-2x^2} \cdot e^{\frac{x^2}{4}} \int_{-\infty}^{+\infty} e^{-u^2} du =
 \end{aligned}$$

$$= \frac{\sqrt{7}}{2\pi} e^{-\frac{7x^2}{4}} \sqrt{\pi} = \frac{\sqrt{7}}{2\sqrt{\pi}} e^{-\frac{7x^2}{4}}.$$

Аналогічно визначаємо функцію $f(y)$:

$$\begin{aligned} f(y) &= \int_{-\infty}^{+\infty} f(x, y) dx = \frac{\sqrt{7}}{2\pi} \int_{-\infty}^{+\infty} e^{-2x^2+xy-y^2} dx = \\ &= \frac{\sqrt{7}}{2\pi} \int_{-\infty}^{+\infty} e^{-2x^2+xy} \cdot e^{-y^2} dx = \frac{\sqrt{7}}{2\pi} e^{-y^2} \int_{-\infty}^{+\infty} e^{-(2x^2-xy)} dx = \\ &= \frac{\sqrt{7}}{2\pi} e^{-y^2} \int_{-\infty}^{+\infty} e^{-\left(-\sqrt{2}x - \frac{y}{2\sqrt{2}}\right)^2 + \frac{y^2}{8}} dx = \\ &= \frac{\sqrt{7}}{2\pi} e^{-y^2} \int_{-\infty}^{+\infty} e^{-\left(-\sqrt{2}x - \frac{y}{2\sqrt{2}}\right)^2} \cdot e^{\frac{y^2}{8}} dx = \\ &= \frac{\sqrt{7}}{2\pi} e^{-\frac{7y^2}{8}} \cdot \frac{1}{\sqrt{2}} \int_{-\infty}^{+\infty} e^{-\left(-\sqrt{2}x - \frac{y}{2\sqrt{2}}\right)^2} d\left(-\sqrt{2}x - \frac{y}{2\sqrt{2}}\right) = \\ &= \frac{\sqrt{7}}{2\sqrt{2}\pi} e^{-\frac{7y^2}{8}} \cdot \sqrt{\pi} = \frac{\sqrt{7}}{2\sqrt{2}\pi} e^{-\frac{7y^2}{8}}. \end{aligned}$$

Для знаходження $f(x/y)$ використовуємо формулу

$$f(x/y) = \frac{f(x, y)}{f(y)}.$$

Отже,

$$f(x/y) = \frac{\frac{\sqrt{7}}{2\pi} e^{-2x^2+xy-y^2}}{\frac{\sqrt{7}}{2\sqrt{2}\pi} e^{-\frac{7y^2}{8}}} = \sqrt{\frac{2}{\pi}} e^{-2x^2+xy-\frac{y^2}{8}}.$$

Аналогічно шукаємо $f(y/x)$:

$$f(y/x) = \frac{f(x, y)}{f(x)} = \frac{\frac{\sqrt{7}}{2\pi} e^{-2x^2 + xy - y^2}}{\frac{\sqrt{7}}{2\sqrt{\pi}} e^{-\frac{7x^2}{4}}} = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{4} + xy - y^2}.$$

Обчислимо математичне сподівання $\mathbf{M}(X)$:

$$\begin{aligned} \mathbf{M}(X) &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x \frac{\sqrt{7}}{2\sqrt{\pi}} e^{-\frac{7x^2}{4}} dx = \frac{\sqrt{7}}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{7x^2}{4}} dx = \\ &= \frac{\sqrt{7}}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} dx = \\ &= \frac{\sqrt{7}}{2\sqrt{\pi}} \left(\int_{-\infty}^0 x e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} dx + \int_0^{+\infty} x e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} dx \right) = \\ &= \frac{\sqrt{7}}{2\sqrt{\pi}} \left(\lim_{A \rightarrow -\infty} \int_A^0 x e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} dx + \lim_{A \rightarrow +\infty} \int_0^A x e^{-\left(\frac{\sqrt{7}x}{2}\right)^2} dx \right) = 0. \end{aligned}$$

Аналогічно визначаємо $\mathbf{M}(Y)$:

$$\mathbf{M}(Y) = \int_{-\infty}^{+\infty} y f(y) dy = \int_{-\infty}^{+\infty} y \frac{\sqrt{7}}{2\sqrt{2\pi}} e^{-\frac{7y^2}{8}} dy = \frac{\sqrt{7}}{2\sqrt{2\pi}} \int_{-\infty}^{+\infty} y e^{-\frac{7y^2}{8}} dy = 0.$$

Отже,

$$\mathbf{M}(X) = \mathbf{M}(Y) = 0.$$

Коефіцієнт кореляції

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Знайдемо коваріацію $\text{cov}(X, Y)$:

$$\begin{aligned} \text{cov}(X, Y) &= \iint_{\mathbb{R}^2} (x - \mathbf{M}(X))(y - \mathbf{M}(Y))f(x, y) dx dy = \\ &= \iint_{\mathbb{R}^2} xy \cdot \frac{\sqrt{7}}{2\pi} e^{-2x^2 + xy - y^2} dx dy = \\ &= \frac{\sqrt{7}}{2\pi} \iint_{\mathbb{R}^2} x e^{-\frac{7}{4}x^2} \cdot y e^{-(y - \frac{1}{2}x)^2} dx dy = \\ &= \frac{\sqrt{7}}{2\pi} \int_{-\infty}^{+\infty} x e^{-\frac{7}{4}x^2} dx \cdot \int_{-\infty}^{+\infty} y e^{-(y - \frac{x}{2})^2} dy = \frac{\sqrt{7}}{2\pi} \cdot 0 = 0, \end{aligned}$$

тобто

$$r(X, Y) = 0,$$

що означає некорельованість випадкових величин X і Y . \square

Задачі до розділу 9

Задача 1. Система випадкових величин (X, Y) має рівномірний розподіл ймовірностей в області \mathcal{D} . Поза цією областю щільність розподілу ймовірностей системи випадкових величин дорівнює нулю.

Записати аналітичний вираз для функції розподілу ймовірностей $F(x, y)$ і функції щільності $f(x, y)$. Обчислити $\mathbf{M}(X)$, $\mathbf{M}(Y)$, якщо

$$\mathcal{D} = \{(x, y) : x \in [-1; 1], y \in [-1; 1]\}.$$

Відповідь.

$$F(x, y) = \begin{cases} 0, & x \leq -1 \text{ або } y \leq -1; \\ \frac{1}{4}(x+1)(y+1), & (x, y) \in \mathcal{D}; \\ \frac{1}{2}(y+1), & x > 1 \text{ і } -1 < y \leq 1; \\ \frac{1}{2}(x+1), & y > 1 \text{ і } -1 < x \leq 1; \\ 1, & x > 1 \text{ і } y > 1, \end{cases}$$

$$f(x, y) = \begin{cases} \frac{1}{4}, & (x, y) \in \mathcal{D}; \\ 0, & (x, y) \notin \mathcal{D}, \end{cases} \quad \mathbf{M}(X) = 0, \mathbf{M}(Y) = 0.$$

Задача 2. Система дискретних випадкових величин задана таблицею розподілу (табл. 4).

Таблиця 4

Y	X				$\sum_{i=1}^4 p_{ij}$
	5	10	15	20	
-10	0,012	0,038	0,2	0,1	0,35
-8	0,038	0,012	0,05	0,05	0,15
-6	0,05	0,05	0,012	0,038	0,15
-4	0,1	0,2	0,038	0,012	0,35
$\sum_{j=1}^4 p_{ij}$	0,2	0,3	0,3	0,2	$\sum_{i=1}^4 \sum_{j=1}^4 p_{ij} = 1$

Обчислити математичні сподівання $\mathbf{M}(X)$, $\mathbf{M}(Y)$, середнє квадратичне відхилення $\sigma(X)$, $\sigma(Y)$, коефіцієнт кореляції $r(X, Y)$, умовні математичні сподівання $\mathbf{M}(X/Y = -8)$, $\mathbf{M}(Y/X = 5)$.

Відповідь. $\mathbf{M}(X) = 12,5$; $\mathbf{M}(Y) = -7$; $\sigma(X) = 5,12$; $\sigma(Y) = 2,57$; $r(X, Y) = -0,513$; $\mathbf{M}(X/Y = -8) = 13,73$; $\mathbf{M}(Y/X = 5) = -5,62$.

Задача 3. Щільність розподілу ймовірностей системи випадкових величин (X, Y) має вигляд

$$f(x, y) = ae^{-4x^2 - 6xy - 10y^2}, \quad x \in \mathbb{R}, \quad y \in \mathbb{R}.$$

Знайти a , $\mathbf{M}(X)$, $\mathbf{M}(Y)$, $\sigma(X)$, $\sigma(Y)$, $r(X, Y)$.

Відповідь. $a = \frac{\sqrt{31}}{\pi}$, $\mathbf{M}(X) = 0$, $\mathbf{M}(Y) = 0$, $\sigma(X) = \frac{5}{31}$, $\sigma(Y) = \frac{2}{31}$, $r(X, Y) = 0$.

Розділ 10. Функції випадкових величин

Нехай ξ — випадкова величина, пов'язана з деяким імовірнісним експериментом, $y = \varphi(\xi)$ — числова функція, яка задовольняє такі вимоги:

- Функція $y = \varphi(x)$ визначена для всіх значень, яких набуває випадкова величина ξ ;
- для довільного значення $y_0 \in \mathbb{R}$ можна обчислити ймовірність $\mathbf{P}(\xi \in \varphi^{-1}(y_0))$, де

$$\varphi^{-1}(y_0) = \{x: \varphi(x) = y_0\},$$

тобто $\varphi^{-1}(y_0)$ — повний прообраз функції φ у точці y_0 .

Тоді $\eta = \varphi(\xi)$ — випадкова величина, яка набуває значення залежно від того, якого значення набула випадкова величина ξ . Якщо в результаті експерименту випадкова величина ξ набуває значення ξ_0 , то випадкова величина η набуває значення $\eta_0 = \varphi(\xi_0)$.

Якщо ξ — дискретна випадкова величина з рядом розподілу

$$\begin{array}{c|c|c|c|c} \xi & x_1 & x_2 & \cdots & x_n \\ \hline \mathbf{P} & p_1 & p_2 & \cdots & p_n \end{array},$$

то випадкова величина $\eta = \varphi(\xi)$ — дискретна випадкова величина з рядом розподілу

$$\begin{array}{c|c|c|c|c} \eta & y_1 & y_2 & \cdots & y_n \\ \hline \mathbf{P} & p_1 & p_2 & \cdots & p_n \end{array},$$

де $y_1 = \varphi(x_1)$, $y_2 = \varphi(x_2)$, ..., $y_n = \varphi(x_n)$. Якщо деякі значення y_i і y_j ($i \neq j$) рівні між собою, то в ряд розподілу записують це значення лише один раз, а ймовірність, що йому відповідає, дорівнює $p_i + p_j$.

Якщо ξ — неперервна випадкова величина, то випадкова величина $\eta = \varphi(\xi)$ може:

- бути дискретною випадковою величиною (якщо $y = \varphi(x)$ — ступінчаста функція, тобто розривна функція, яка набуває не більш як зліченну кількість значень);

- б) бути неперервною випадковою величиною, якщо $y = \varphi(x)$ — неперервна функція без інтервалів сталості;
- в) не бути ані дискретною, ані неперервною випадковою величиною (такі випадкові величини називаються випадковими величинами змішаного типу).

У випадку а) випадкова величина η набуває значень y_1, y_2, \dots з імовірностями

$$q_1 = \mathbf{P}(\xi \in \varphi^{-1}(y_1)), \quad q_2 = \mathbf{P}(\xi \in \varphi^{-1}(y_2)), \dots$$

У випадку б) для довільного інтервалу $(a; b)$ ми маємо визначити повний прообраз

$$\varphi^{-1}((a; b)) = \{x: \varphi(x) \in (a; b)\}$$

і ймовірність

$$\mathbf{P}(\eta \in (a; b)) = \mathbf{P}(\xi \in \varphi^{-1}((a; b))).$$

Функція розподілу випадкової величини η має вигляд

$$G_\eta(y) = \mathbf{P}(\eta < y) = \mathbf{P}(\xi \in \varphi^{-1}((-\infty; y))) = \int_{D_y} f_\xi(x) dx,$$

де

$$D_y = \varphi^{-1}((-\infty; y)) = \{x: \varphi(x) < y\},$$

$f_\xi(x)$ — щільність розподілу ймовірностей випадкової величини ξ .

Якщо ξ — абсолютно неперервна випадкова величина зі щільністю $f_\xi(x)$ і $y = \varphi(x)$ — диференційована строго монотонна функція з оберненою функцією $x = \psi(y)$, то щільність розподілу $g(y)$ випадкової величини η визначається з рівності

$$g_\eta(y) = f_\xi(\psi(y)) \cdot |\psi'(y)|.$$

При цьому функція розподілу $G_\eta(y)$ обчислюється за формулою

$$G_\eta(y) = \int_{-\infty}^y g_\eta(u) du.$$

Приклад 1. Випадкова величина ξ задана рядом розподілу

ξ	-2	-1	0	1	2	3
\mathbf{P}	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Знайти ряди розподілу таких випадкових величин:

- а) $\eta = 2\xi$;
 б) $\zeta = \xi^2$.

Розв'язання. а) Знайдемо можливі значення випадкової величини $\eta = 2\xi$:

$$\begin{aligned} \eta_1 = 2 \cdot (-2) = -4; & \quad \eta_2 = 2 \cdot (-1) = -2; & \quad \eta_3 = 2 \cdot 0 = 0; \\ \eta_4 = 2 \cdot 1 = 2; & \quad \eta_5 = 2 \cdot 2 = 4; & \quad \eta_6 = 2 \cdot 3 = 6. \end{aligned}$$

Різним значенням випадкової величини ξ відповідають різні значення випадкової величини η . Отже, ряд розподілу випадкової величини η має вигляд

η	-4	-2	0	2	4	6
\mathbf{P}	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

б) Знайдемо можливі значення випадкової величини $\zeta = \xi^2$:

$$\begin{aligned} \zeta_1 = (-2)^2 = 4; & \quad \zeta_2 = (-1)^2 = 1; & \quad \zeta_3 = 0^2 = 0; \\ \zeta_4 = 1^2 = 1; & \quad \zeta_5 = 2^2 = 4; & \quad \zeta_6 = 3^2 = 9. \end{aligned}$$

Оскільки

$$\zeta_1 = \zeta_5, \quad \zeta_2 = \zeta_4,$$

то ряд розподілу випадкової величини η має такий вигляд:

η	0	1	4	9
\mathbf{P}	$\frac{1}{6}$	$\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$	$\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$	$\frac{1}{6}$

□

Приклад 2. Випадкова величина ξ має рівномірний розподіл на відрізьку $[0; 9]$. Знайти закон розподілу випадкової величини $\eta = \varphi(\xi)$, якщо:

$$\begin{aligned} \text{а) } \varphi(x) &= \begin{cases} 0, & x \leq 1; \\ 5, & 1 < x \leq 6; \\ 7, & 6 < x \leq 9; \\ 2002, & x > 9; \end{cases} \\ \text{б) } \varphi(x) &= \sqrt{x}; \\ \text{в) } \varphi(x) &= \begin{cases} 0, & x \leq 0; \\ x^2, & 0 < x \leq 2; \\ 4, & 2 < x \leq 6; \\ x - 2, & x > 6. \end{cases} \end{aligned}$$

Розв'язання. а) Оскільки множина значень функції $y = \varphi(x)$ складається лише з чотирьох чисел, то випадкова величина η дискретна з можливими значеннями 0, 5, 7, 2002 і ймовірностями, що їм відповідають:

$$\begin{aligned} \mathbf{P}(\eta = 0) &= \mathbf{P}(\xi \in \varphi^{-1}(0)) = \mathbf{P}(\xi \in (-\infty; 1]); \\ \mathbf{P}(\eta = 5) &= \mathbf{P}(\xi \in \varphi^{-1}(5)) = \mathbf{P}(\xi \in (1; 6]); \\ \mathbf{P}(\eta = 7) &= \mathbf{P}(\xi \in \varphi^{-1}(7)) = \mathbf{P}(\xi \in (6; 9]); \\ \mathbf{P}(\eta = 2002) &= \mathbf{P}(\xi \in \varphi^{-1}(2002)) = \mathbf{P}(\xi \in (9; +\infty)). \end{aligned}$$

Оскільки випадкова величина ξ має рівномірний розподіл на $[0; 9]$, то її функція розподілу

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x}{9}, & 0 < x \leq 9; \\ 1, & x > 9. \end{cases}$$

Звідси отримуємо:

$$\mathbf{P}(\xi \in (-\infty; 1]) = F_{\xi}(1) - F_{\xi}(-\infty) = \frac{1}{9} - 0 = \frac{1}{9};$$

$$\begin{aligned} \mathbf{P}(\xi \in (1; 6]) &= F_{\xi}(6) - F_{\xi}(1) = \frac{6}{9} - \frac{1}{9} = \frac{5}{9}; \\ \mathbf{P}(\xi \in (6; 9]) &= F_{\xi}(9) - F_{\xi}(6) = \frac{9}{9} - \frac{6}{9} = \frac{3}{9} = \frac{1}{3}; \\ \mathbf{P}(\xi \in (9; +\infty)) &= F_{\xi}(+\infty) - F_{\xi}(9) = 1 - \frac{9}{9} = 0. \end{aligned}$$

Отже, ряд розподілу випадкової величини η має вигляд

η	0	5	7
\mathbf{P}	$\frac{1}{9}$	$\frac{5}{9}$	$\frac{1}{3}$

б) Оскільки $0 \leq \xi \leq 9$, то $0 \leq \sqrt{\xi} \leq 3$. Функція $y = \sqrt{x}$ на інтервалі $[0; 9]$ є строго зростаючою диференційованою, тому випадкова величина $\eta = \varphi(\xi)$ має неперервний розподіл, причому щільність $g_{\eta}(y)$ цього розподілу задовольняє рівність

$$g_{\eta}(y) = f_{\xi}(\psi(y)) \cdot |\psi'(y)|,$$

де $f_{\xi}(\cdot)$ — щільність розподілу ймовірностей випадкової величини ξ :

$$f_{\xi}(x) = \begin{cases} 0, & x \notin [0; 9]; \\ \frac{1}{9}, & x \in [0; 9], \end{cases}$$

$\psi(\cdot)$ — функція, обернена до функції $y = \sqrt{x}$, тобто

$$\psi(y) = y^2; \quad \psi'(y) = 2y \quad (\text{при } 0 \leq y \leq 3).$$

Тоді

$$g_{\eta}(y) = \begin{cases} 0, & y \notin [0; 3]; \\ \frac{1}{9} \cdot 2y = \frac{2}{9}y, & y \in [0; 3]. \end{cases}$$

Функція розподілу $G_{\eta}(y)$ випадкової величини η має вигляд

$$G_{\eta}(y) = \int_{-\infty}^y g(u) du = \begin{cases} 0, & y \leq 0; \\ \frac{y^2}{9}, & 0 < y \leq 3; \\ 1, & y > 3. \end{cases}$$

в) Функція $y = \varphi(x)$ є неперервною і неспадною, графік якої зображено на рис. 32.

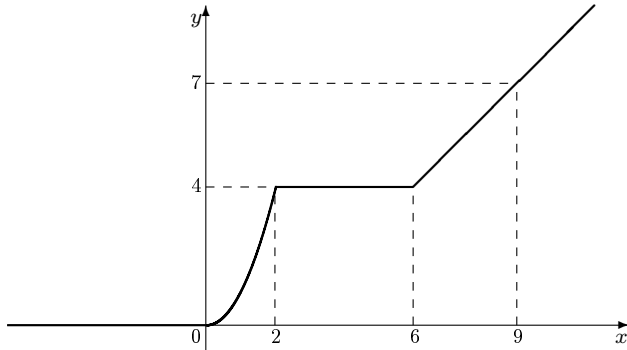


Рис. 32

Образом відрізка $[0; 9]$ при відображенні $y = \varphi(x)$ є відрізок $[0; 7]$. Оскільки ξ розподілена на $[0; 9]$, то випадкова величина η розподілена на $[0; 7]$. Якщо $y_0 \neq 4$ і $y_0 \in [0; 7]$, то прообразом точки y_0 є єдина точка $x_0 \in [0; 9]$. Тому

$$\mathbf{P}(\eta = y_0) = \mathbf{P}(\xi = x_0 = \varphi^{-1}(y_0)) = 0, \quad y_0 \neq 4.$$

Якщо $y_0 = 4$, то $\varphi^{-1}(4) = [2; 6]$. Тому

$$\mathbf{P}(\eta = 4) = \mathbf{P}(\xi \in [2; 6]) = F_\xi(6) - F_\xi(2) = \frac{6}{9} - \frac{2}{9} = \frac{4}{9}.$$

Якщо $y \in [0; 4)$, то

$$y = \varphi(x) = x^2 \quad \text{і} \quad x = \psi(y) = \sqrt{y} \quad (0 \leq y < 4), \quad |\psi'(y)| = \frac{1}{2\sqrt{y}}.$$

Оскільки

$$g_\eta(y) = f_\xi(\psi(y)) \cdot |\psi'(y)|,$$

де

$$f_\xi(x) = \begin{cases} 0, & x \notin [0; 9]; \\ \frac{1}{9}, & x \in [0; 9], \end{cases}$$

то

$$g_{\eta}(y) = \frac{1}{9} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{18\sqrt{y}}$$

при $y \in [0; 4)$.

Якщо $y \in (4; 7]$, то

$$y = \varphi(x) = x - 2.$$

Тоді

$$x = \psi(y) = y + 2, \quad |\psi'(y)| = 1$$

і

$$g_{\eta}(y) = f_{\xi}(\psi(y)) \cdot |\psi'(y)| = \frac{1}{9} \cdot 1 = \frac{1}{9}.$$

Отже,

$$g_{\eta}(y) = \begin{cases} 0, & y \notin [0; 7]; \\ \frac{1}{18\sqrt{y}}, & y \in [0; 4); \\ \text{не існує,} & y = 4; \\ \frac{1}{9}, & y \in (4; 7], \end{cases}$$

$$G_{\eta}(y) = \begin{cases} 0, & y \leq 0; \\ \int_0^y \frac{du}{18\sqrt{u}}, & 0 < y \leq 4; \\ \int_0^4 \frac{du}{18\sqrt{u}} + \mathbf{P}(y = 4) + \int_4^y \frac{1}{9} du, & 4 < y \leq 7; \\ 1, & y > 7; \end{cases} =$$

$$= \begin{cases} 0, & y \leq 0; \\ \frac{\sqrt{y}}{9}, & 0 < y \leq 4; \\ \frac{y+2}{9}, & 4 < y \leq 7; \\ 1, & y > 7. \end{cases}$$

Графік функції $G_\eta(y)$ зображено на рис. 33. □

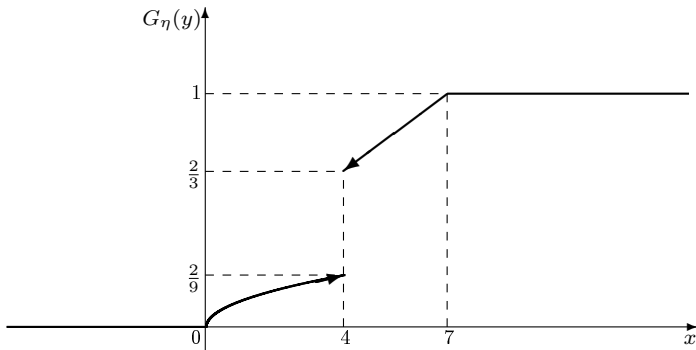


Рис. 33

Задачі до розділу 10

Задача 1. Випадкова величина ξ має такий ряд розподілу:

ξ	1	2	3	4
\mathbf{P}	0,1	0,2	0,3	0,4

Побудувати ряд розподілу випадкової величини η , якщо:

а) $\eta = \xi + 5$;

б) $\eta = |\xi - 2|$.

Відповідь. а)

η	6	7	8	9
\mathbf{P}	0,1	0,2	0,3	0,4

; б)

η	0	1	2
\mathbf{P}	0,2	0,4	0,4

Задача 2. Випадкова величина ξ має такий ряд розподілу:

ξ	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π
P	0,2	0,2	0,2	0,2	0,2

Побудувати ряд розподілу випадкової величини X , якщо:

а) $X = \sin(2\xi)$;

б) $X = \frac{4\xi}{\pi}$.

Відповідь. а) $\frac{X}{\mathbf{P}} \begin{array}{c|c|c|c} -1 & 0 & 1 & \\ \hline 0,2 & 0,6 & 0,2 & \end{array}$; б) $\frac{X}{\mathbf{P}} \begin{array}{c|c|c|c|c|c} 0 & 1 & 2 & 3 & 4 & \\ \hline 0,2 & 0,2 & 0,2 & 0,2 & 0,2 & 0,2 \end{array}$.

Задача 3. Випадкова величина X має такий ряд розподілу:

X	-3	-2	-1	0	1	2
P	0,25	0,05	0,2	0,1	0,3	0,1

Побудувати ряд розподілу випадкової величини Y , якщо:

а) $Y = X^3$;

б) $Y = |X|$.

Відповідь. а) $\frac{Y}{\mathbf{P}} \begin{array}{c|c|c|c|c|c} -27 & -8 & -1 & 0 & 1 & 8 & \\ \hline 0,25 & 0,05 & 0,2 & 0,1 & 0,3 & 0,1 & \end{array}$;

б) $\frac{Y}{\mathbf{P}} \begin{array}{c|c|c|c} 0 & 1 & 2 & 3 & \\ \hline 0,1 & 0,5 & 0,15 & 0,25 & \end{array}$.

Задача 4. Випадкова величина X має такий ряд розподілу:

X	-9	-4	-1	0	1	4	9	16
P	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Побудувати ряд розподілу випадкової величини Y , якщо:

а) $Y = \sqrt{|X|}$;

б) $Y = 2X + 1$.

Відповідь. а)
$$\begin{array}{c|c|c|c|c|c} Y & 0 & 1 & 2 & 3 & 4 \\ \hline \mathbf{P} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \end{array};$$

б)
$$\begin{array}{c|c|c|c|c|c|c|c|c} Y & -17 & -7 & -1 & 1 & 3 & 9 & 19 & 33 \\ \hline \mathbf{P} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{array}.$$

Задача 5. Випадкова величина ξ рівномірно розподілена на відріжку $[0; 4]$. Знайти розподіл випадкової величини $X = [\xi]$, де $[a]$ означає цілу частину числа a .

Відповідь.
$$\begin{array}{c|c|c|c|c} X & 0 & 1 & 2 & 3 \\ \hline \mathbf{P} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array}.$$

Задача 6. Випадкова величина ξ рівномірно розподілена на відріжку $[-5; 5]$. Знайти розподіл випадкової величини $X = \varphi(\xi)$, якщо

$$\varphi(x) = \begin{cases} 3, & x \leq 0; \\ 7, & 0 < x \leq 2; \\ 11, & x > 2. \end{cases}$$

Відповідь.
$$\begin{array}{c|c|c|c} X & 3 & 7 & 11 \\ \hline \mathbf{P} & \frac{1}{2} & \frac{1}{5} & \frac{3}{10} \end{array}.$$

Задача 7. Випадкова величина ξ рівномірно розподілена на відріжку $[0; 2]$. Знайти щільність і функцію розподілу випадкової величини $X = \xi^2$.

Відповідь.
$$f(x) = \begin{cases} \frac{1}{4\sqrt{x}}, & x \in [0; 4]; \\ 0, & x \notin [0; 4], \end{cases} \quad F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{\sqrt{x}}{2}, & 0 < x \leq 4; \\ 1, & x > 4. \end{cases}$$

Задача 8. Випадкова величина ξ рівномірно розподілена на відріжку $[1; 10]$. Знайти щільність і функцію розподілу випадкової величини $X = \sqrt{\xi - 1}$.

Відповідь. $f(x) = \begin{cases} \frac{2x}{9}, & x \in [0; 3]; \\ 0, & x \notin [0; 3], \end{cases} \quad F(x) = \begin{cases} 0, & x \leq 0; \\ \frac{x^2}{9}, & 0 < x \leq 3; \\ 1, & x > 3. \end{cases}$

Задача 9. Випадкова величина ξ має таку щільність розподілу:

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Знайти щільність розподілу випадкової величини $X = \frac{1}{\xi}$.

Відповідь. $f(x) = \frac{1}{\pi(1+x^2)}.$

Задача 10. Випадкова величина ξ має таку щільність розподілу:

$$f(x) = \begin{cases} xe^{-\frac{x^2}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Знайти щільність розподілу випадкової величини $X = e^{-\xi^2}$.

Відповідь. $f(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & x \in [0; 1]; \\ 0, & x \notin [0; 1]. \end{cases}$

Задача 11. Випадкова величина ξ має таку щільність розподілу:

$$f(x) = \begin{cases} \sin x, & x \in \left[0; \frac{\pi}{2}\right]; \\ 0, & x \notin \left[0; \frac{\pi}{2}\right]. \end{cases}$$

Знайти щільність розподілу випадкової величини $X = \xi^4$.

Відповідь. $f(x) = \begin{cases} \frac{\sin \sqrt[4]{x}}{4\sqrt[4]{x^3}}, & x \in \left[0; \frac{\pi^4}{16}\right]; \\ 0, & x \notin \left[0; \frac{\pi^4}{16}\right]. \end{cases}$

Задача 12. Випадкова величина ξ має таку щільність розподілу:

$$f(x) = \begin{cases} \frac{1}{2} \cos x, & x \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]; \\ 0, & x \notin \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]. \end{cases}$$

Знайти щільність розподілу випадкової величини $X = \xi^3$.

Відповідь. $f(x) = \begin{cases} \frac{\cos \sqrt[3]{x}}{6\sqrt[3]{x^2}}, & x \in \left[-\frac{\pi^3}{8}; 0\right) \cup \left(0; \frac{\pi^3}{8}\right]; \\ 0, & x \notin \left[-\frac{\pi^3}{8}; 0\right) \cup \left(0; \frac{\pi^3}{8}\right]. \end{cases}$

Задача 13. Випадкова величина ξ рівномірно розподілена на відрізьку $[-10; 10]$.

Знайти закон розподілу випадкової величини $X = \varphi(\xi)$, якщо

$$\varphi(x) = \begin{cases} 0, & x \leq 2; \\ x - 2, & 2 < x \leq 9; \\ 7, & x > 9. \end{cases}$$

Відповідь. Розподіл випадкової величини X є сумішшю дискретного та неперервного розподілів. $F(x) = \begin{cases} 0, & x \leq 0; \\ 0,6 + 0,05x, & 0 < x \leq 7; \\ 1, & x > 7. \end{cases}$

Задача 14. Випадкова величина ξ рівномірно розподілена на відрізьку $[0; 5]$.

Знайти закон розподілу випадкової величини $X = \varphi(\xi)$, якщо

$$\varphi(x) = \begin{cases} 3, & x \in [-2\pi; 2\pi]; \\ \ln(\sqrt{\cos x + 3}), & x \notin [-2\pi; 2\pi]. \end{cases}$$

Відповідь. Випадкова величина X має вироджений розподіл:

$$\mathbf{P}\{X = 3\} = 1.$$

Розділ 11. Граничні теореми теорії ймовірностей

Граничні теореми теорії ймовірностей можна умовно розбити на два класи: закони великих чисел і центральні граничні теореми.

Закони великих чисел математично описують стійкість середніх значень масових випадкових явищ. Історично першим законом великих чисел була така теорема:

Теорема 1 (теорема Я. Бернуллі). Нехай k — кількість появ події A в серії з n незалежних випробувань, у кожному з яких подія A відбувається з імовірністю p . Тоді для довільного $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \left| \frac{k}{n} - p \right| < \varepsilon \right\} = 1.$$

Іншими словами з імовірністю, як завгодно близькою до одиниці, можна стверджувати, що при досить великій кількості незалежних випробувань статистична частота $\frac{k}{n}$ появи події A як завгодно мало відрізняється від імовірності появи події A в одному випробуванні.

Ця теорема, зокрема, пояснює, чому при багаторазовому підкиданні симетричної монети кількість гербів становить приблизно половину від загальної кількості підкидань.

Узагальненням теореми 1 є така теорема:

Теорема 2 (теорема П. Л. Чебишева). Нехай ξ_i — послідовність попарно незалежних випадкових величин зі скінченними математичними сподіваннями $\mathbf{M}(\xi_i)$ і обмеженими дисперсіями $\mathbf{D}(\xi_i) \leq c_0$. Тоді для довільного $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n \mathbf{M}(\xi_i) \right| < \varepsilon \right\} = 1.$$

В обох теоремах суттєво використано нерівність Чебишева, яка, втім, має й самостійну цінність.

Лема 1 (нерівність Чебишева). Нехай ξ — довільна випадкова величина з дисперсією $\mathbf{D}(\xi)$ і математичним сподіванням $\mathbf{M}(\xi)$. Тоді для довільного $\varepsilon > 0$

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

Нерівність Чебишева можна записати також у формі

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

Крім законів великих чисел, описаних у теоремах 1 і 2, спостерігається ще одне досить цікаве явище, яке полягає в тому, що при великій кількості випадкових доданків, кожен з яких вносить лише невеликий внесок у загальну суму, розподіл кожного з випадкових доданків не впливає на сумарний результат. Більш точно твердження сформульовано в такій теоремі.

Теорема 3 (центральна гранична теорема). Нехай ξ_i — однаково розподілені випадкові величини з математичним сподіванням $\mathbf{M}(\xi_i) = a$ і дисперсією $\mathbf{D}(\xi_i) = \sigma^2$. Тоді при великому n розподіл суми $\xi_1 + \xi_2 + \dots + \xi_n = \xi$ близький до нормального розподілу.

Якщо результуюча випадкова величина ξ має математичне сподівання $\mathbf{M}(\xi)$ і дисперсію $\mathbf{D}(\xi)$, то з теореми 3 випливає, що

$$\mathbf{P}\{a \leq \xi \leq b\} \approx \Phi\left(\frac{b - \mathbf{M}(\xi)}{\sqrt{\mathbf{D}(\xi)}}\right) - \Phi\left(\frac{a - \mathbf{M}(\xi)}{\sqrt{\mathbf{D}(\xi)}}\right),$$

а отже,

$$\mathbf{P}\{\mathbf{M}(\xi) - \delta \leq \xi \leq \mathbf{M}(\xi) + \delta\} \approx 2\Phi\left(\frac{\delta}{\sqrt{\mathbf{D}(\xi)}}\right).$$

Із останніх рівностей випливає, що ймовірність p події можна оцінити, користуючись експериментальними даними. Справді, якщо в

серії з n незалежних випробувань досліджувана подія A відбувається k разів, то статистична частота появи події A дорівнює $\frac{k}{n}$. Ця частота є випадковою величиною ξ , для якої математичне сподівання

$$\mathbf{M}(\xi) = p,$$

а дисперсія

$$\mathbf{D}(\xi) = \frac{p(1-p)}{n}.$$

Тоді при досить великих значеннях n ($n \geq 30$) виконується рівність

$$\mathbf{P}\left\{\frac{k}{n} - \delta < p < \frac{k}{n} + \delta\right\} \approx 2\Phi\left(\frac{\delta}{\sqrt{\mathbf{D}(\xi)}}\right) = 2\Phi\left(\delta\sqrt{\frac{n}{p(1-p)}}\right).$$

Приклад 1. Використовуючи нерівність Чебишева, знайти ймовірність події A , яка полягає в тому, що випадкова величина X набуде значення, яке відрізнятиметься від математичного сподівання $\mathbf{M}(X)$ на величину, що не перевищує потроєного середнього квадратичного відхилення. Чи зміниться відповідь, якщо відомо, що випадкова величина X має нормальний розподіл?

Розв'язання. За нерівністю Чебишева

$$\begin{aligned} \mathbf{P}(A) &= \mathbf{P}\{|X - \mathbf{M}(X)| < 3\sigma(X)\} \geq 1 - \frac{\mathbf{D}(X)}{(3\sigma(X))^2} = \\ &= 1 - \frac{\mathbf{D}(X)}{9\mathbf{D}(X)} = \frac{8}{9}. \end{aligned}$$

Отже,

$$\mathbf{P}(A) \geq \frac{8}{9} \approx 0,89.$$

Якщо X має нормальний розподіл, то

$$\begin{aligned} \mathbf{P}(A) &= \mathbf{P}\{|X - \mathbf{M}(X)| < 3\sigma(X)\} = 2\Phi\left(\frac{3\sigma(X)}{\sigma(X)}\right) = 2\Phi(3) = \\ &= 0,9973. \end{aligned} \quad \square$$

Приклад 2. При виробництві дискет брак становить 1 %. Скільки дискет потрібно відібрати для перевірки якості, щоб з імовірністю 0,95 можна було стверджувати, що у випадковій вибірці дискет відсоток бракованих відрізняється від 1 % не більш як на 0,5 %?

Розв'язання. Кількість бракованих дискет є випадковою величиною

$$\xi = \xi_1 + \xi_2 + \dots + \xi_n,$$

де $\xi_1, \xi_2, \dots, \xi_n$ — незалежні однаково розподілені випадкові величини; ξ_i — випадкова величина, яка дорівнює кількості бракованих дискет при виготовленні однієї дискети, тобто ξ_i може набувати значення або 0, або 1 з імовірністю відповідно 0,99 і 0,01. Якщо n — досить велике число, то за центральною граничною теоремою розподіл випадкової величини ξ близький до нормального. Тому

$$\mathbf{P} \left\{ \left| \frac{k}{n} - p \right| < \delta \right\} \approx 2\Phi \left(\frac{\delta}{\sqrt{D(\xi)}} \right) = 2\Phi \left(\delta \sqrt{\frac{n}{p(1-p)}} \right),$$

де $\frac{k}{n}$ — частота браку; $p = 0,01$; $\delta = 0,005$; n — невідома кількість дискет. Число n потрібно вибрати таким, щоб

$$\mathbf{P} \left\{ \left| \frac{k}{n} - 0,01 \right| < 0,005 \right\} \approx 2\Phi \left(0,005 \sqrt{\frac{n}{0,01 \cdot 0,99}} \right) = 0,95.$$

Іншими словами,

$$\Phi \left(0,5 \sqrt{\frac{n}{99}} \right) = 0,475.$$

За таблицею значень функції Лапласа (дод. 1) знаходимо значення аргументу x таке, що $\Phi(x) = 0,475$:

$$x = 1,96.$$

Розв'язавши рівняння

$$0,5 \sqrt{\frac{n}{99}} = 1,96,$$

отримаємо

$$n \geq 99 \cdot \left(\frac{1,96}{0,5} \right)^2 \geq 1522. \quad \square$$

Приклад 3. При анонімному тестуванні виявилось, що 10% працівників підприємства зовсім не вживають спиртного. Випадкова величина X — кількість людей, які зовсім не вживають спиртного у випадковій вибірці з 900 робітників. Зазначити межі, у які потрапляє випадкова величина X з імовірністю 0,9.

Розв'язання. Згідно з центральною граничною теоремою можна вважати, що випадкова величина X має розподіл, близький до нормального. За умовою

$$p = 0,1; \quad n = 900.$$

Тоді

$$\mathbf{P} \left\{ \left| \frac{X}{900} - 0,1 \right| < \delta \right\} \approx 2\Phi \left(\delta \sqrt{\frac{900}{0,1 \cdot 0,9}} \right) = 0,9.$$

Отже,

$$\Phi(100\delta) = 0,45.$$

За таблицею значень функції Лапласа знаходимо значення аргументу x таке, що $\Phi(x) = 0,45$:

$$x = 1,65.$$

Звідси $100\delta = 1,65$ і $\delta = 0,0165$. Отже, з імовірністю 0,9 виконуватиметься умова

$$\mathbf{P} \left\{ \left| \frac{X}{900} - 0,1 \right| < 0,0165 \right\} \geq 0,9,$$

що рівносильно тому, що випадкова величина X потрапить в інтервал (75; 105) з імовірністю 0,9. \square

Приклад 4. Нехай x — випадково вибране число з відрізка $[0; 1]$,

$$x = \alpha_1(x)\alpha_2(x) \dots \alpha_k(x) \dots$$

є десятковий розклад числа x , $\alpha_k(x) \in \{0, 1, \dots, 9\}$. Нехай $N_7(x, k)$ — кількість цифр «7», які зустрілися в десятковому розкладі числа x до k -го місця включно. Знайти ймовірність того, що частота цифри «7», яка дорівнює

$$\lim_{k \rightarrow \infty} \frac{N_7(x, k)}{k},$$

відрізнятиметься від 0,1 не більше ніж на ε .

Розв'язання. $\xi = N_7(x, k)$ — випадкова величина, яка є сумою

$$\xi = \xi_1 + \xi_2 + \dots + \xi_n,$$

де $\xi_i = 0$, якщо $\alpha_i(x) \neq 7$, і $\xi_i = 1$, якщо $\alpha_i(x) = 7$. Іншими словами, ξ_i має такий розподіл:

$$\begin{array}{c|cc} \xi_i & 0 & 1 \\ \mathbf{P} & 0,9 & 0,1 \end{array}.$$

Легко побачити, що

$$\mathbf{M}(\xi_i) = 0,1; \quad \mathbf{M}(\xi_i^2) = 0,1.$$

Тому

$$\mathbf{D}(\xi_i) = \mathbf{M}(\xi_i^2) - \mathbf{M}^2(\xi_i) = 0,1 - 0,01 = 0,09.$$

Оскільки випадкові величини ξ_i незалежні та мають обмежені дисперсії й математичні сподівання, то виконується теорема Чебишева:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P} \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{n} \sum_{i=1}^n \mathbf{M}(\xi_i) \right| < \varepsilon \right\} = \\ = \lim_{n \rightarrow \infty} \mathbf{P} \left\{ \left| \frac{N_7(x, k)}{n} - 0,1 \right| < \varepsilon \right\} = 1. \end{aligned}$$

Отже, для довільного $\varepsilon > 0$ з імовірністю 1 частота цифри «7» у десятковому розкладі випадково обраного числа x відрізняється від 0,1 не більше ніж на ε . \square

Задачі до розділу 11

Задача 1. Використовуючи нерівність Чебишева, оцінити ймовірність того, що

$$|\xi - \mathbf{M}(\xi)| < 0,3,$$

якщо $\mathbf{D}(\xi) = 0,0025$. Оцінити ймовірність тієї самої події, якщо відомо, що ξ має нормальний розподіл.

Відповідь. За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 0,3\} \geq 0,97.$$

Якщо припустити, що ξ має нормальний розподіл, то

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 0,3\} \approx 1 \quad (\text{з точністю до } 0,0001).$$

Задача 2. Використовуючи нерівність Чебишева, оцінити ймовірність того, що

$$|\xi - \mathbf{M}(\xi)| > 0,1,$$

якщо $\sigma(\xi) = 0,4$. Оцінити ймовірність тієї самої події, якщо відомо, що ξ має нормальний розподіл.

Відповідь. За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| > 0,1\} < 16.$$

Якщо припустити, що ξ має нормальний розподіл, то

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| > 0,1\} \approx 0,8026.$$

Задача 3. Маса деталей, що виготовляються на верстаті, є випадковою величиною, середнє значення якої (математичне сподівання) дорівнює 1,2 кг. Дисперсія цієї величини дорівнює 0,012. Використовуючи нерівність Чебишева, оцінити ймовірність того, що:

а) відхилення маси деталі від її середнього значення за абсолютною величиною не перевищить 0,2;

б) маса деталі набуде значення від 1,18 до 1,22.

Відповідь. а) $P \geq 0,7$; б) $P \geq 0$.

Задача 4. Прилад складається з 10 елементів, що працюють незалежно. Імовірність відмови кожного елемента за певний час t дорівнює 0,05. Використовуючи нерівність Чебишева, оцінити ймовірність того, що абсолютна величина різниці між кількістю елементів, які відмовили, і середнім числом (математичним сподіванням) відмов за час t буде менше 2.

Відповідь. $P \geq 0,88$.

Задача 5. Із 1000 виробів навання відібрано 200. Після перевірки серед них виявлено 15 бракованих. Приймаючи частку бракованих виробів за ймовірність виготовлення браку, оцінити ймовірність того, що в усій партії бракованих виробів буде не більше 10 % і не менше 5 %.

Відповідь. $P \geq 0,889$.

Задача 6. У країні N проживає 50 млн людей. ВІЛ-тестування пройшли 1000 випадково відібраних громадян, серед яких виявилось 20 інфікованих. Приймаючи частку інфікованих з дослідженої групи за ймовірність того, що випадково обраний громадянин країни N інфікований, оцінити ймовірність того, що в даній країні відсоток ВІЛ-інфікованих становить не більше 3 % і не менше 1 %.

Відповідь. $P \geq 0,99996$.

Задача 7. Скільки виробів необхідно відібрати для перевірки якості продукції, щоб з імовірністю 0,95 можна було стверджувати, що частка бракованих деталей відрізняється від імовірності $p = 0,1$ випуску бракованої деталі не більше ніж на 0,02?

Відповідь. $n \geq 865$.

Задача 8. При виробництві поліетиленових пакетів брак становить 5 %. Скільки виробів потрібно відібрати для перевірки якості продукції, щоб з імовірністю 0,9 можна було стверджувати, що частка бракованих пакетів становить від 4 до 6 %?

Відповідь. $n \geq 1294$.

Задача 9. Випадкова величина ξ має такий закон розподілу:

ξ	3	5
\mathbf{P}	0,4	0,6

Використовуючи нерівність Чебишева, оцінити ймовірність того, що:

а) $|\xi - \mathbf{M}(\xi)| < 0,2$;

б) $|\xi - \mathbf{M}(\xi)| < 2$.

Обчислити точні ймовірності цих подій.

Відповідь. а) За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 0,2\} \geq 0.$$

Насправді,

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 0,2\} = 0.$$

б) За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 2\} \geq 0,76.$$

Насправді,

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 2\} = 1.$$

Задача 10. Випадкова величина ξ має такий закон розподілу:

ξ	1	5
\mathbf{P}	0,7	0,3

Використовуючи нерівність Чебишева, оцінити ймовірність того, що:

а) $|\xi - \mathbf{M}(\xi)| < 1$;

б) $|\xi - \mathbf{M}(\xi)| < 2$.

Обчислити точні ймовірності цих подій.

Відповідь. а) За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 1\} \geq 0.$$

Насправді,

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 1\} = 0.$$

б) За нерівністю Чебишева

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 2\} \geq 0,16.$$

Насправді,

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| < 2\} = 0,7.$$

Задача 11. Нехай x — випадково вибране число з відрізка $[0; 1]$,

$$x = \alpha_1(x)\alpha_2(x) \dots \alpha_k(x) \dots$$

є десятковий розклад числа x , $\alpha_k(x) \in \{0, 1, \dots, 9\}$. Нехай $N_5(x, k)$ — кількість цифр «5», які зустрілися в десятковому розкладі числа x до k -го місця включно. Знайти ймовірність того, що частота цифри «5», яка дорівнює

$$\lim_{k \rightarrow \infty} \frac{N_5(x, k)}{k},$$

перебуватиме в межах від 0,0999 до 0,1001.

Відповідь. $\mathbf{P} = 1$.

Задача 12. Нехай

$$x = \alpha_1(x)\alpha_2(x) \dots \alpha_k(x) \dots$$

є двійковий розклад числа $x \in [0; 1]$, $\alpha_k(x) \in \{0, 1\}$, $N_0(x, k)$ — кількість нулів, які зустрілися в двійковому розкладі числа x до k -го місця включно. Знайти ймовірність того, що частота цифри «0», яка дорівнює

$$\lim_{k \rightarrow \infty} \frac{N_0(x, k)}{k},$$

перебуватиме в межах від 0,49999 до 0,50001.

Відповідь. $\mathbf{P} = 1$.

ЧАСТИНА III

Математична статистика

Розділ 12. Елементи математичної статистики. Вибірковий метод

12.1. Вибірковий метод

Нехай для вивчення кількісної (дискретної чи неперервної) ознаки X із генеральної сукупності отримано вибірку x_1, x_2, \dots, x_n обсягом n .

Якщо записати елементи вибірки в порядку зростання, отримаємо *варіаційний ряд*.

Спостережувані **різні** значення x_i ознаки X називають *варіантами*, кількість значень однієї варіанти у вибірці — її *частотою* n_i (сума частот усіх варіант дорівнює обсягу вибірки), відношення частоти до обсягу вибірки — *відносною частотою* або *емпіричною ймовірністю* $w_i = \frac{n_i}{n}$ (сума відносних частот усіх варіант дорівнює одиниці).

Статистичним розподілом вибірки називають перелік варіант варіаційного ряду з відповідними їм частотами і/або відносними частотами. Статистичний розподіл також задають у вигляді послідовності замкнених справа напівінтервалів і відповідних їм частот і/або

відносних частот (частотою інтервалу вважають суму частот усіх варіант з даного інтервалу).

Щоб підкреслити зазначені відмінності в першому випадку говорять про *точковий*, а в другому — про *інтервальний* статистичний розподіл вибірки.

Приклад 1. Під час дослідження кількісної ознаки X із генеральної сукупності було отримано вибірку

$$4, 3, 6, 4, 7, 2, 5, 1, 2, 5, 4, 4, 3, 5, 6, 3, 4, 1, 3, 4.$$

Знайти обсяг вибірки, побудувати варіаційний ряд вибірки та її статистичний розподіл.

Розв'язання. Оскільки вибірка складається з 20 значень, то обсяг вибірки $n = 20$.

Побудуємо варіаційний ряд вибірки:

$$1, 1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 6, 6, 7.$$

У даній вибірці всього сім різних значень, тобто варіант:

$$1, 2, 3, 4, 5, 6, 7.$$

Знайдемо їх частоти:

$$n_1 = 2; \quad n_2 = 2; \quad n_3 = 4; \quad n_4 = 6; \quad n_5 = 3; \quad n_6 = 2; \quad n_7 = 1.$$

Запишемо шуканий статистичний розподіл:

$$\begin{array}{c|c|c|c|c|c|c} x_i & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline n_i & 2 & 2 & 4 & 6 & 3 & 2 & 1 \end{array}.$$

□

Приклад 2. Вибірка задана розподілом частот:

$$\begin{array}{c|c|c|c|c|c} x_i & 3 & 5 & 7 & 10 & 15 \\ \hline n_i & 2 & 4 & 7 & 4 & 3 \end{array}.$$

Знайти розподіл відносних частот.

Розв'язання. Знайдемо обсяг вибірки:

$$n = 2 + 4 + 7 + 4 + 3 = 20.$$

Визначимо відносні частоти:

$$w_1 = \frac{2}{20} = 0,1; \quad w_2 = \frac{4}{20} = 0,2; \quad w_3 = \frac{7}{20} = 0,35;$$
$$w_4 = \frac{4}{20} = 0,2; \quad w_5 = \frac{3}{20} = 0,15.$$

Отже, шуканий розподіл відносних частот має такий вигляд:

x_i	3	5	7	10	15
w_i	0,1	0,2	0,35	0,2	0,15

Контроль: $0,1 + 0,2 + 0,35 + 0,2 + 0,15 = 1$. □

Емпіричною функцією розподілу (функцією розподілу вибірки) називають функцію

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

де n_x — кількість елементів вибірки, менших від x (тобто сума частот усіх варіант, менших від x); n — обсяг вибірки.

Емпірична функція розподілу має такі властивості.

Властивість 1. Емпірична функція розподілу $F^*(x)$ невід'ємна і не перевищує одиниці при будь-яких значеннях x :

$$F^*(x) \in [0; 1].$$

Властивість 2. Емпірична функція розподілу $F^*(x)$ неспадна:

$$x < y \quad \Rightarrow \quad F^*(x) \leq F^*(y).$$

Властивість 3. Якщо x_1 — найменша варіанта, а x_n — найбільша, то емпірична функція розподілу $F^*(x) = 0$ при $x \leq x_1$ і $F^*(x) = 1$ при $x > x_n$.

Приклад 3. Знайти емпіричну функцію розподілу за даним розподілом вибірки:

x_i	3	5	7	10	15
n_i	2	4	7	4	3

Розв'язання. Знайдемо обсяг вибірки:

$$n = 2 + 4 + 7 + 4 + 3 = 20.$$

Оскільки найменша варіанта дорівнює трьом,

$$F^*(x) = 0$$

при всіх $x \leq 3$.

Значення $X < 5$, а саме значення $x_1 = 3$, спостерігалося двічі, тому

$$F^*(x) = \frac{2}{20} = 0,1$$

при $3 < x \leq 5$.

Значення $X < 7$, а саме $x_1 = 3$ і $x_2 = 5$, спостерігалися $2+4=6$ разів, тому

$$F^*(x) = \frac{6}{20} = 0,3$$

при $5 < x \leq 7$.

Значення $X < 10$, а саме $x_1 = 3$, $x_2 = 5$ і $x_3 = 7$, спостерігалися $2+4+7=13$ разів, тому

$$F^*(x) = \frac{13}{20} = 0,65$$

при $7 < x \leq 10$.

Значення $X < 15$, а саме $x_1 = 3$, $x_2 = 5$, $x_3 = 7$ і $x_4 = 10$, спостерігалися $2 + 4 + 7 + 4 = 17$ разів, тому

$$F^*(x) = \frac{17}{20} = 0,85$$

при $10 < x \leq 15$.

Оскільки $x_5 = 15$ — найбільша варіанта, то

$$F^*(x) = 1$$

при $x > 15$.

Отже, запишемо шукану емпіричну функцію:

$$F^*(x) = \begin{cases} 0 & \text{при } x \leq 3; \\ 0,1 & \text{при } 3 < x \leq 5; \\ 0,3 & \text{при } 5 < x \leq 7; \\ 0,65 & \text{при } 7 < x \leq 10; \\ 0,85 & \text{при } 10 < x \leq 15; \\ 1 & \text{при } x > 15. \end{cases}$$

Графік цієї функції зображено на рис. 34. □

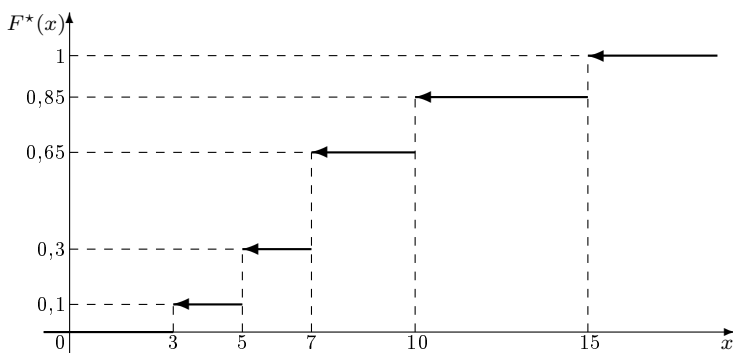


Рис. 34

Полігоном частот називають ламану, відрізки якої з'єднують точки

$$(x_1, n_1), (x_2, n_2), \dots, (x_k, n_k),$$

де x_i — варіанти вибірки; n_i — відповідні частоти, $i = 1, 2, \dots$

Полігоном відносних частот називають ламану, відрізки якої з'єднують точки

$$(x_1, w_1), (x_2, w_2), \dots, (x_k, w_k),$$

де x_i — варіанти вибірки; w_i — відповідні емпіричні ймовірності, $i = 1, 2, \dots$

При неперервному розподілі ознаки X у разі великої кількості спостережень увесь інтервал, у якому розміщені спостережені значення ознаки, як правило, розбивають на кілька частинних інтервалів однакової довжини h і знаходять n_i — суму частот варіант, що потрапили в i -й інтервал. Для вибору *оптимальної величини інтервалу* рекомендовано використовувати формулу

$$h \approx \frac{x_{\max} - x_{\min}}{1 + 3,2 \cdot \lg n},$$

де x_{\max} , x_{\min} — відповідно найбільше і найменше значення у вибірці; n — обсяг вибірки.

Якщо задано інтервальний статистичний розподіл вибірки, то для побудови полігону частот або емпіричних ймовірностей за даними вибірки з'єднують точки, абсцисами яких є значення середин частинних інтервалів, а ординатами — відповідні їм значення частот або емпіричних ймовірностей.

Гістограмою частот називають ступінчасту фігуру, що складається з прямокутників, основами яких є частинні інтервали довжиною h , а висоти дорівнюють відношенню $\frac{n_i}{h}$ (*щільність частоти*). Площа i -го частинного прямокутника

$$h \cdot \frac{n_i}{h} = n_i,$$

тобто сумі частот варіант, що потрапили в i -й частинний інтервал. Площа гістограми частот дорівнює сумі всіх частот, тобто обсягу вибірки n .

Гістограмою відносних частот називають ступінчасту фігуру, що складається з прямокутників, основами яких є частинні інтервали довжиною h , а висоти дорівнюють відношенню $\frac{w_i}{h}$ (*щільність відносної частоти*). Площа i -го частинного прямокутника

$$h \cdot \frac{w_i}{h} = w_i,$$

тобто сумі відносних частот варіант, що потрапили в i -й інтервал. Площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто одиниці.

Якщо задано точковий статистичний розподіл вибірки, то для побудови гістограми частот або емпіричних імовірностей за даними вибірки як частинні інтервали беруть інтервали, серединами яких є значення відповідних їм варіант.

Приклад 4. Побудувати полігон частот за даним розподілом вибірки:

x_i	1	2	3	4	5	6	7
n_i	5	1	8	3	2	2	6

Розв'язання. Відкладемо на осі абсцис значення варіант x_i , а на осі ординат — значення відповідних їм частот n_i . Послідовно з'єднуючи між собою точки (x_i, n_i) відрізками, отримуємо полігон частот (рис. 35). □

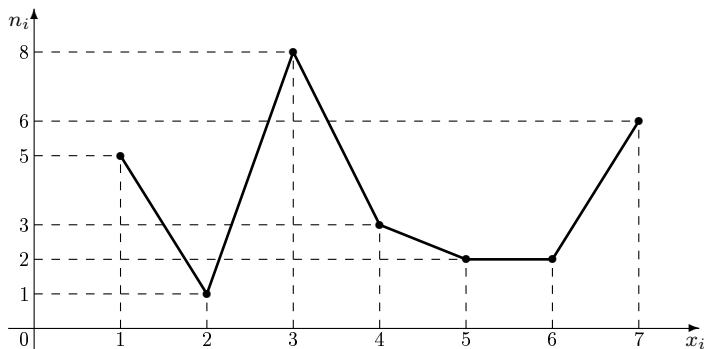


Рис. 35

Приклад 5. Вибірку задано інтервальним розподілом частот:

$(x_i; x_{i+1}]$	(1; 3]	(3; 5]	(5; 7]	(7; 9]	(9; 11]
n_i	13	9	5	16	7

Побудувати полігон відносних частот.

Розв'язання. Знайдемо обсяг вибірки:

$$n = 13 + 9 + 5 + 16 + 7 = 50.$$

Для побудови інтервального статистичного розподілу відносних частот визначимо відносні частоти:

$$\begin{aligned} w_1 &= \frac{n_1}{n} = \frac{13}{50} = 0,26; & w_2 &= \frac{n_2}{n} = \frac{9}{50} = 0,18; \\ w_3 &= \frac{n_3}{n} = \frac{5}{50} = 0,1; & w_4 &= \frac{n_4}{n} = \frac{16}{50} = 0,32; \\ w_5 &= \frac{n_5}{n} = \frac{7}{50} = 0,14. \end{aligned}$$

Побудуємо інтервальний статистичний розподіл відносних частот:

$(x_i; x_{i+1}]$	(1; 3]	(3; 5]	(5; 7]	(7; 9]	(9; 11]
w_i	0,26	0,18	0,1	0,32	0,14

Знайдемо середини частинних інтервалів:

$$\begin{aligned} x_1^{\text{сер}} &= \frac{x_1 + x_2}{2} = \frac{1 + 3}{2} = 2; & x_2^{\text{сер}} &= \frac{x_2 + x_3}{2} = \frac{3 + 5}{2} = 4; \\ x_3^{\text{сер}} &= \frac{x_3 + x_4}{2} = \frac{5 + 7}{2} = 6; & x_4^{\text{сер}} &= \frac{x_4 + x_5}{2} = \frac{7 + 9}{2} = 8; \\ x_5^{\text{сер}} &= \frac{x_5 + x_6}{2} = \frac{9 + 11}{2} = 10. \end{aligned}$$

Відкладемо на осі абсцис значення середин частинних інтервалів $x_i^{\text{сер}}$, а на осі ординат — значення відповідних їм відносних частот w_i . Послідовно з'єднуючи між собою точки $(x_i^{\text{сер}}, w_i)$ відрізками, отримуємо полігон відносних частот (рис. 36). \square

Приклад 6. Вибірку задано інтервальним розподілом частот:

$(x_i; x_{i+1}]$	(1; 2]	(2; 3]	(3; 4]	(4; 5]	(5; 6]	(6; 7]	(7; 8]
n_i	19	9	12	14	7	22	17

Побудувати гістограму частот.

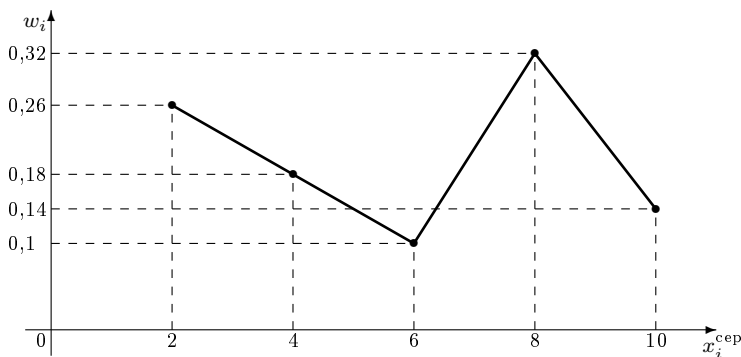


Рис. 36

Розв'язання. Знайдемо довжину інтервалів:

$$h = x_{i+1} - x_i \equiv 1.$$

Визначимо щільність частоти:

$$\begin{aligned} n_1^h &= \frac{n_1}{h} = \frac{19}{1} = 19; & n_2^h &= \frac{n_2}{h} = \frac{9}{1} = 9; & n_3^h &= \frac{n_3}{h} = \frac{12}{1} = 12; \\ n_4^h &= \frac{n_4}{h} = \frac{14}{1} = 14; & n_5^h &= \frac{n_5}{h} = \frac{7}{1} = 7; & n_6^h &= \frac{n_6}{h} = \frac{22}{1} = 22; \\ n_7^h &= \frac{n_7}{h} = \frac{17}{1} = 17. \end{aligned}$$

Далі побудуємо на осі абсцис задані частинні інтервали довжиною $h = 1$. Проведемо над цими інтервалами відрізки, що паралельні осі абсцис і розміщені від неї на відстанях, які дорівнюють відповідним щільностям частоти n_i^h . Краї цих відрізків з'єднаємо з віссю абсцис відрізками, паралельними осі ординат. Отриману гістограму частот зображено на рис. 37. \square

Приклад 7. Вибірку задано розподілом частот:

x_i	2	4	6	8	10	12	14
n_i	2	3	5	1	4	2	3

Побудувати гістограму відносних частот.

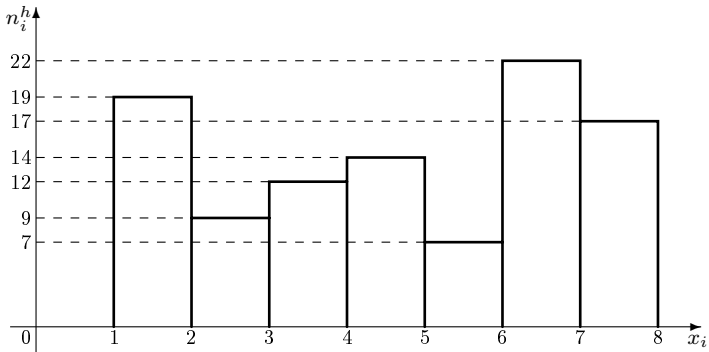


Рис. 37

Розв'язання. Щоб побудувати гістограму відносних частот, даний дискретний статистичний розподіл частот потрібно перетворити на інтервальний статистичний розподіл відносних частот (емпіричних імовірностей) і знайти щільність цих відносних частот. Для цього спочатку визначимо обсяг вибірки й емпіричні ймовірності:

$$n = 2 + 3 + 5 + 1 + 4 + 2 + 3 = 20,$$

$$w_1 = \frac{n_1}{n} = \frac{2}{20} = 0,1; \quad w_2 = \frac{n_2}{n} = \frac{3}{20} = 0,15;$$

$$w_3 = \frac{n_3}{n} = \frac{5}{20} = 0,25; \quad w_4 = \frac{n_4}{n} = \frac{1}{20} = 0,05;$$

$$w_5 = \frac{n_5}{n} = \frac{4}{20} = 0,2; \quad w_6 = \frac{n_6}{n} = \frac{2}{20} = 0,1;$$

$$w_7 = \frac{n_7}{n} = \frac{3}{20} = 0,15.$$

Знайдемо довжину частинних інтервалів, ці частинні інтервали й щільність відносних частот. Частинні інтервали визначимо з умови, що задані в дискретному статистичному розподілі варіанти мають бути серединами частинних інтервалів.

Отже, довжина частинних інтервалів

$$h = x_{i+1} - x_i \equiv 2,$$

шуканий інтервальний статистичний розподіл емпіричних імовірностей має такий вигляд:

$(x_i; x_{i+1}]$	$(1; 3]$	$(3; 5]$	$(5; 7]$	$(7; 9]$	$(9; 11]$	$(11; 13]$	$(13; 15]$
w_i	0,1	0,15	0,25	0,05	0,2	0,1	0,15

а щільності емпіричних імовірностей такі:

$$\begin{aligned}
 w_1^h &= \frac{w_1}{h} = \frac{0,1}{2} = 0,05; & w_2^h &= \frac{w_2}{h} = \frac{0,15}{2} = 0,075; \\
 w_3^h &= \frac{w_3}{h} = \frac{0,25}{2} = 0,125; & w_4^h &= \frac{w_4}{h} = \frac{0,05}{2} = 0,025; \\
 w_5^h &= \frac{w_5}{h} = \frac{0,2}{2} = 0,1; & w_6^h &= \frac{w_6}{h} = \frac{0,1}{2} = 0,05; \\
 w_7^h &= \frac{w_7}{h} = \frac{0,15}{2} = 0,075.
 \end{aligned}$$

Тепер, застосувавши міркування попереднього прикладу, можна легко побудувати шукану гістограму відносних частот (рис. 38). \square

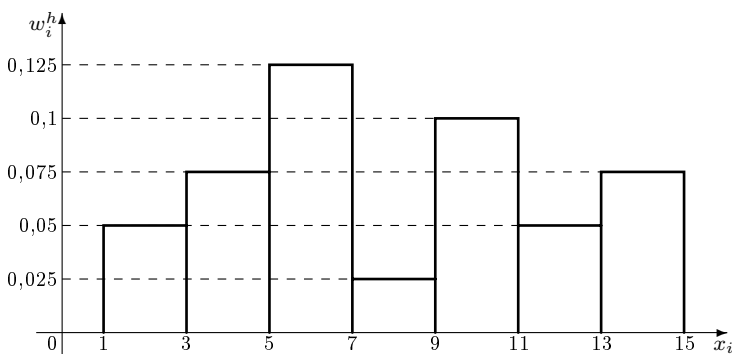


Рис. 38

Приклад 8. Дано вибірку: 0,1; 0,4; 0,23; 0,12; 0,35; 0,46; 0,11; 0,04; 0,51; 0,27; 0,31; 0,34; 0,09; 0,18; 0,49; 0,33; 0,3; 0,22; 0,14; 0,5; 0,41; 0,25; 0,48; 0,32; 0,29; 0,31; 0,31; 0,46; 0,44; 0,38; 0,39; 0,13; 0,47; 0,4; 0,53; 0,37; 0,16; 0,44; 0,39; 0,27; 0,25; 0,46; 0,2; 0,11; 0,32; 0,41; 0,48; 0,22; 0,35; 0,52.

За даними вибірки побудувати інтервальний статистичний розподіл, полігон і гістограму частот.

Розв'язання. Знайдемо обсяг вибірки: $n = 50$. Оскільки найбільше значення $x_{\max} = 0,53$, а найменше — $x_{\min} = 0,04$, то оптимальна величина частинного інтервалу

$$h = \frac{x_{\max} - x_{\min}}{1 + 3,2 \cdot \lg n} = \frac{0,53 - 0,04}{1 + 3,2 \cdot \lg 50} \approx \frac{0,49}{1 + 3,2 \cdot 1,69897} \approx 0,076.$$

Частинні інтервали почнемо знаходити із серединного інтервалу з умови, що він має розміщуватися на однаковій відстані від кінців варіаційного ряду (тобто від 0,04 і 0,53). Отже, середина серединного частинного інтервалу

$$x_{\text{сеп}} = \frac{x_{\max} + x_{\min}}{2} = \frac{0,53 + 0,04}{2} = 0,285.$$

Оскільки оптимальна величина частинного інтервалу $h = 0,076$, то серединний інтервал

$$\begin{aligned} (x_4; x_5] &= \left(x_{\text{сеп}} - \frac{h}{2}; x_{\text{сеп}} + \frac{h}{2} \right] = \\ &= \left(0,285 - \frac{0,076}{2}; 0,285 + \frac{0,076}{2} \right] = \\ &= (0,247; 0,323]. \end{aligned}$$

Тепер легко знайти решту частинних інтервалів (спочатку будемо рухатися в напрямку від серединного частинного інтервалу до першого, а потім — від серединного до останнього):

$$\begin{aligned} (x_3; x_4] &= (0,247 - 0,076; 0,247] = (0,171; 0,247], \\ (x_2; x_3] &= (0,171 - 0,076; 0,171] = (0,095; 0,171], \\ (x_1; x_2] &= (0,095 - 0,076; 0,095] = (0,019; 0,095], \\ (x_5; x_6] &= (0,323; 0,323 + 0,076] = (0,323; 0,399], \\ (x_6; x_7] &= (0,399; 0,399 + 0,076] = (0,399; 0,475], \\ (x_7; x_8] &= (0,475; 0,475 + 0,076] = (0,475; 0,551]. \end{aligned}$$

Визначивши частоти частинних інтервалів, отримаємо інтервальний статистичний розподіл вибірки

Номер інтервалу i	Частинний інтервал $(x_i, x_{i+1}]$	Сума частот варіант інтервалу n_i
1	(0,019; 0,095]	2
2	(0,095; 0,171]	7
3	(0,171; 0,247]	5
4	(0,247; 0,323]	11
5	(0,323; 0,399]	8
6	(0,399; 0,475]	10
7	(0,475; 0,551]	7

Контроль: $n = 2 + 7 + 5 + 11 + 8 + 10 + 7 = 50$.

Щоб побудувати полігон частот, знайдемо значення середин частинних інтервалів:

$$x_1^{\text{сер}} = \frac{x_1 + x_2}{2} = \frac{0,019 + 0,095}{2} = 0,057;$$

$$x_2^{\text{сер}} = \frac{x_2 + x_3}{2} = \frac{0,095 + 0,171}{2} = 0,133;$$

$$x_3^{\text{сер}} = \frac{x_3 + x_4}{2} = \frac{0,171 + 0,247}{2} = 0,209;$$

$$x_4^{\text{сер}} = \frac{x_4 + x_5}{2} = \frac{0,247 + 0,323}{2} = 0,285;$$

$$x_5^{\text{сер}} = \frac{x_5 + x_6}{2} = \frac{0,323 + 0,399}{2} = 0,361;$$

$$x_6^{\text{сер}} = \frac{x_6 + x_7}{2} = \frac{0,399 + 0,475}{2} = 0,437;$$

$$x_7^{\text{сер}} = \frac{x_7 + x_8}{2} = \frac{0,475 + 0,551}{2} = 0,513.$$

Гістограму частот побудуємо, знайшовши щільність частоти:

$$\begin{aligned} n_1^h &= \frac{n_1}{h} = \frac{2}{0,076} \approx 26; & n_2^h &= \frac{n_2}{h} = \frac{7}{0,076} \approx 92; \\ n_3^h &= \frac{n_3}{h} = \frac{5}{0,076} \approx 66; & n_4^h &= \frac{n_4}{h} = \frac{11}{0,076} \approx 145; \\ n_5^h &= \frac{n_5}{h} = \frac{8}{0,076} \approx 105; & n_6^h &= \frac{n_6}{h} = \frac{10}{0,076} \approx 132; \\ & & n_7^h &= \frac{n_7}{h} = \frac{7}{0,076} \approx 92. \end{aligned}$$

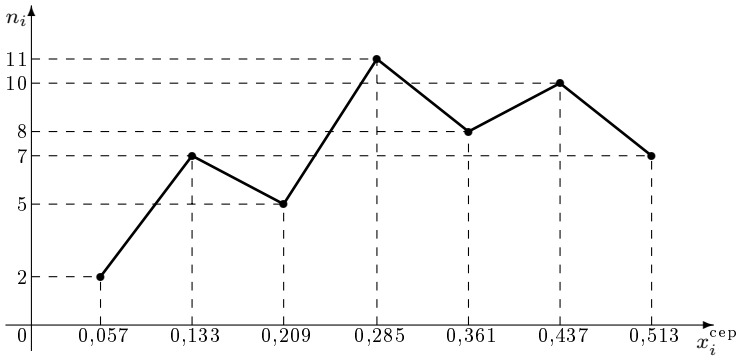


Рис. 39

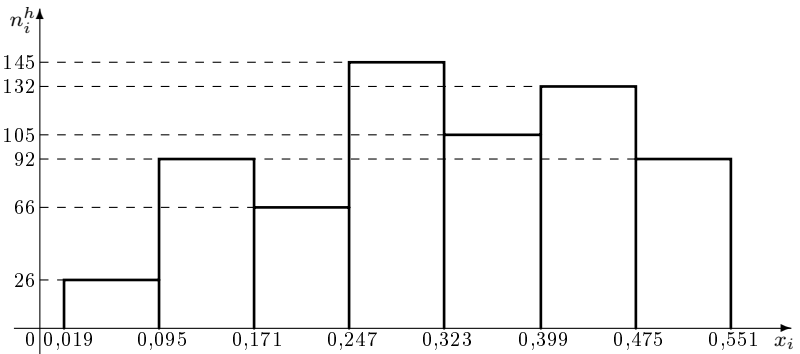


Рис. 40

Відклавши на осях координат усі необхідні точки, легко побудувати полігон і гістограму частот, які зображені відповідно на рис. 39 і 40. □

12.2. Числові характеристики вибірки

Середнє арифметичне значення вибірки називається *вибірковим середнім* \bar{x}_B і обчислюється за формулами

$$\bar{x}_B = \frac{\sum_{i=1}^k x_i n_i}{n} \quad \text{або} \quad \bar{x}_B = \sum_{i=1}^k x_i w_i,$$

де x_i — значення i -ї варіанти; n_i — частота i -ї варіанти; n — обсяг вибірки; k — кількість варіант у вибірці; w_i — відносна частота i -ї варіанти.

Середній квадрат відхилення значень елементів вибірки від вибіркового середнього називається *вибірковою дисперсією* D_B і обчислюється за формулами

$$D_B = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^2 n_i}{n} \quad \text{або} \quad D_B = \sum_{i=1}^k (x_i - \bar{x}_B)^2 w_i.$$

Після перетворень формули для знаходження вибіркової дисперсії дещо спрощуються:

$$D_B = \frac{\sum_{i=1}^k x_i^2}{n} - \bar{x}_B^2 \quad \text{або} \quad D_B = \sum_{i=1}^k x_i^2 w_i - \bar{x}_B^2,$$

тобто вибіркова дисперсія дорівнює різниці середнього квадрата елементів вибірки й квадрата вибіркового середнього.

Квадратний корінь з вибіркової дисперсії $\sigma = \sqrt{D_B}$ називається *середнім квадратичним відхиленням вибірки*.

Приклад 1. Задано статистичний розподіл вибірки:

x_i	1	3	4	7	10	12	15
n_i	5	2	12	7	4	3	2

Знайти вибіркове середнє, вибірккову дисперсію та середнє квадратичне відхилення вибірки.

Розв'язання. Обчислимо обсяг вибірки:

$$n = 5 + 2 + 12 + 7 + 4 + 3 + 2 = 35.$$

Знайдемо відповідно вибірккове середнє, вибірккову дисперсію та середнє квадратичне відхилення вибірки:

$$\begin{aligned} \bar{x}_B &= \frac{\sum_{i=1}^7 x_i n_i}{n} = \frac{1 \cdot 5 + 3 \cdot 2 + 4 \cdot 12 + 7 \cdot 7 + 10 \cdot 4 + 12 \cdot 3 + 15 \cdot 2}{35} = \\ &= \frac{214}{35} \approx 6,1; \\ D_B &= \frac{\sum_{i=1}^7 (x_i - \bar{x}_B)^2 n_i}{n} = \frac{\sum_{i=1}^7 x_i^2 n_i}{n} - \bar{x}_B^2 = \\ &= \frac{1^2 \cdot 5 + 3^2 \cdot 2 + 4^2 \cdot 12 + 7^2 \cdot 7 + 10^2 \cdot 4 + 12^2 \cdot 3 + 15^2 \cdot 2}{35} - \\ &\quad - \left(\frac{214}{35} \right)^2 = \frac{18604}{1225} \approx 15,2; \\ \sigma &= \sqrt{D_B} = \sqrt{\frac{18604}{1225}} \approx 3,9. \end{aligned} \quad \square$$

Якщо задано інтервальний статистичний розподіл вибірки, то вибірккове середнє, вибірккову дисперсію та середнє квадратичне відхилення вибірки шукають за допомогою такого статистичного розподілу: варіантами вважаються середини частинних інтервалів, а частоти або відносні частоти залишаються такими самими.

Приклад 2. Задано інтервальний статистичний розподіл вибірки:

$(x_i; x_{i+1}]$	(0; 2]	(2; 4]	(4; 6]	(6; 8]	(8; 10]	(10; 12]	(12; 14]
w_i	0,1	0,2	0,3	0,1	0,1	0,1	0,1

Знайти вибіркове середнє, вибіркєву дисперсїю та середнє квадратичнє вїдхилєння вибірки.

Розв'язання. Спочатку перетворимо даний інтервальний статистичний розподїл вибірки на точковий, знайшовши середини частинних інтервалів:

$$\begin{aligned}
 x_1 &= \frac{x_1 + x_2}{2} = \frac{0 + 2}{2} = 1; & x_2 &= \frac{x_2 + x_3}{2} = \frac{2 + 4}{2} = 3; \\
 x_3 &= \frac{x_3 + x_4}{2} = \frac{4 + 6}{2} = 5; & x_4 &= \frac{x_4 + x_5}{2} = \frac{6 + 8}{2} = 7; \\
 x_5 &= \frac{x_5 + x_6}{2} = \frac{8 + 10}{2} = 9; & x_6 &= \frac{x_6 + x_7}{2} = \frac{10 + 12}{2} = 11; \\
 & & x_7 &= \frac{x_7 + x_8}{2} = \frac{12 + 14}{2} = 13.
 \end{aligned}$$

Отже, ми отримали такий статистичний розподїл вибірки:

x_i	1	3	5	7	9	11	13
w_i	0,1	0,2	0,3	0,1	0,1	0,1	0,1

Знайдемо вїдповїдно вибіркєву середнє, вибіркєву дисперсїю та середнє квадратичнє вїдхилєння вибірки:

$$\begin{aligned}
 \bar{x}_B &= \sum_{i=1}^7 x_i w_i = 1 \cdot 0,1 + 3 \cdot 0,2 + 5 \cdot 0,3 + 7 \cdot 0,1 + 9 \cdot 0,1 + \\
 &\quad + 11 \cdot 0,1 + 13 \cdot 0,1 = 6,2; \\
 D_B &= \sum_{i=1}^7 (x_i - \bar{x}_B)^2 w_i = \sum_{i=1}^7 x_i^2 w_i - \bar{x}_B^2 = 1^2 \cdot 0,1 + 3^2 \cdot 0,2 + \\
 &\quad + 5^2 \cdot 0,3 + 7^2 \cdot 0,1 + 9^2 \cdot 0,1 + 11^2 \cdot 0,1 + 13^2 \cdot 0,1 - 6,2^2 = \\
 &= 12,96; \\
 \sigma &= \sqrt{D_B} = \sqrt{12,96} = 3,6.
 \end{aligned}$$

□

Медіаною (Me) називається значення середнього елемента варіаційного ряду. Якщо обсяг вибірки $n = 2m + 1$ непарний, то медіаною буде значення елемента варіаційного ряду з номером $m + 1$:

$$\text{Me} = x_{m+1}.$$

Якщо обсяг вибірки $n = 2m$ парний, то медіаною буде середнє значення елементів варіаційного ряду з номерами m і $m + 1$:

$$\text{Me} = \frac{x_m + x_{m+1}}{2}.$$

Якщо задано інтервальний статистичний розподіл вибірки, то спочатку знаходять *медіанний частинний інтервал*, тобто перший частинний інтервал, для якого сума частот усіх попередніх частинних інтервалів з даним включно перевищує половину обсягу вибірки. У цьому разі медіану знаходять за формулою

$$\text{Me} = x_{\text{Me}}^{\min} + h_{\text{Me}} \frac{\frac{1}{2}n - \sum_{i=1}^{\text{Me}-1} n_i}{n_{\text{Me}}},$$

де x_{Me}^{\min} — початок медіанного частинного інтервалу; h_{Me} — довжина медіанного частинного інтервалу; n — обсяг вибірки; n_{Me} — частота медіанного частинного інтервалу; $\text{Me} - 1$ — номер попереднього до медіанного частинного інтервалу.

Медіана має таку властивість: сума абсолютних величин відхилень елементів вибірки від медіани менша, ніж від будь-якої іншої величини:

$$\sum_{i=1}^n |x_i - \text{Me}| < \sum_{i=1}^n |x_i - a|, \quad a \neq \text{Me}.$$

Приклад 3. На одному з відрізків залізниці планується створити зупинку пасажирського поїзда. Розподіл населених пунктів з чисельністю їх населення наведено в таблиці.

На якому кілометрі залізниці розташований населений пункт, км	10	12	15	25	28	30	33
Чисельність населення, тис. чол.	5	2	3	10	1	4	6

На якому кілометрі залізниці потрібно розташувати цю зупинку, щоб сумарна відстань, яку покриватимуть потенційні пасажирів до цієї зупинки, була найменшою.

Розв'язання. Оскільки медіана має властивість, що сума абсолютних величин відхилень елементів вибірки від медіани менша, ніж від будь-якої іншої величини, то для розв'язання прикладу потрібно знайти медіану.

Спочатку визначимо обсяг вибірки:

$$n = 5 + 2 + 3 + 10 + 1 + 4 + 6 = 31.$$

Отже, серединою (середнім членом) варіаційного ряду буде елемент із номером 16: $Me = x_{16}$. Оскільки варіаційний ряд можна записати у вигляді

$$\underbrace{10, \dots, 10}_{5 \text{ разів}}, 12, 12, 15, 15, 15, \underbrace{25, \dots, 25}_{10 \text{ разів}}, 28, \underbrace{30, \dots, 30}_{4 \text{ рази}}, \underbrace{33, \dots, 33}_{6 \text{ разів}};$$

легко бачити, що $x_{16} = 25$, тобто зупинку слід розташувати на 25-му кілометрі залізниці. □

Модю (Mo) називається варіанта з найбільшою частотою. У випадку інтервального статистичного розподілу вибірки з однаковими за довжиною частинними інтервалами *модальний частинний інтервал* визначається за найбільшою частотою, а при різних за довжиною частинних інтервалах — за найбільшою щільністю $\max_i \frac{n_i}{h_i}$, де

n_i, h_i — відповідно частота і довжина i -го частинного інтервалу. У цьому разі моду знаходять за формулою

$$Mo = x_{Mo}^{\min} + h_{Mo} \frac{n_{Mo} - n_{Mo-1}}{2n_{Mo} - n_{Mo-1} - n_{Mo+1}},$$

де x_{Mo}^{\min} — початок модального частинного інтервалу; h_{Mo} — довжина модального частинного інтервалу; n_{Mo} — частота модального частинного інтервалу; n_{Mo-1} — частота попереднього до модального частинного інтервалу; n_{Mo+1} — частота наступного за модальним частинного інтервалу.

Варіаційним розмахом називається різниця між найбільшим і найменшим значеннями вибірки:

$$R = x_{\max} - x_{\min}.$$

Приклад 4. Задано інтервальний статистичний розподіл вибірки:

$(x_i, x_{i+1}]$	(10; 15]	(15; 20]	(20; 25]	(25; 30]	(30; 35]	(35; 40]	(40; 45]
n_i	7	4	5	1	12	3	18

Знайти медіану, моду, варіаційний розмах.

Розв'язання. Визначимо обсяг вибірки:

$$n = 7 + 4 + 5 + 1 + 12 + 3 + 18 = 50.$$

Медіанним частинним інтервалом буде п'ятий інтервал, оскільки це перший інтервал, для якого сума частот усіх попередніх частинних інтервалів з даним включно перевищує половину обсягу вибірки:

$$7 + 4 + 5 + 1 + 12 = 29 > \frac{n}{2} = \frac{50}{2} = 25.$$

Модальним частинним інтервалом буде останній інтервал, оскільки він має найбільшу частоту.

Знайдемо початок медіанного частинного інтервалу x_{Me}^{\min} , довжину медіанного частинного інтервалу h_{Me} , номер попереднього до медіанного частинного інтервалу $Me - 1$, частоту медіанного частинного інтервалу n_{Me} , початок модального частинного інтервалу x_{Mo}^{\min} ,

довжину модального частинного інтервалу $h_{\mathbf{Mo}}$, частоту модального частинного інтервалу $n_{\mathbf{Mo}}$, частоту попереднього до модального частинного інтервалу $n_{\mathbf{Mo}-1}$, частоту наступного за модальним частинного інтервалу $n_{\mathbf{Mo}+1}$, найбільше й найменше значення вибірки:

$$\begin{aligned} x_{\mathbf{Me}}^{\min} &= 30; & h_{\mathbf{Me}} &= 35 - 30 = 5; & \mathbf{Me} - 1 &= 4; \\ n_{\mathbf{Me}} &= 12; & x_{\mathbf{Mo}}^{\min} &= 40; & h_{\mathbf{Mo}} &= 45 - 40 = 5; \\ n_{\mathbf{Mo}} &= 18; & n_{\mathbf{Mo}-1} &= 3; & n_{\mathbf{Mo}+1} &= 0; \\ x_{\max} &= 45; & x_{\min} &= 10. \end{aligned}$$

Шукані медіана, мода та варіаційний розмах відповідно будуть такі:

$$\begin{aligned} \mathbf{Me} &= x_{\mathbf{Me}}^{\min} + h_{\mathbf{Me}} \frac{\frac{1}{2}n - \sum_{i=1}^{\mathbf{Me}-1} n_i}{n_{\mathbf{Me}}} = 30 + 5 \cdot \frac{\frac{1}{2} \cdot 50 - \sum_{i=1}^4 n_i}{12} = \\ &= \frac{100}{3} \approx 33,3; \end{aligned}$$

$$\begin{aligned} \mathbf{Mo} &= x_{\mathbf{Mo}}^{\min} + h_{\mathbf{Mo}} \frac{n_{\mathbf{Mo}} - n_{\mathbf{Mo}-1}}{2n_{\mathbf{Mo}} - n_{\mathbf{Mo}-1} - n_{\mathbf{Mo}+1}} = 40 + 5 \cdot \frac{18 - 3}{2 \cdot 18 - 3 - 0} = \\ &= \frac{465}{11} \approx 42,34; \end{aligned}$$

$$\mathbf{R} = x_{\max} - x_{\min} = 45 - 10 = 35. \quad \square$$

12.3. Метод добутків обчислення вибіркового середнього та вибіркової дисперсії

Нехай вибірку задано у вигляді розподілу рівновіддалених варіант і відповідних їм частот. У цьому разі вибіркове середнє та вибіркору дисперсію зручно знаходити *методом добутків* за формулами

$$\bar{x}_v = \mathbf{M}_1^* h + C, \quad \mathbf{D}_v = \left[\mathbf{M}_2^* - (\mathbf{M}_1^*)^2 \right] h^2,$$

де $\mathbf{M}_1^* = \frac{\sum_i n_i u_i}{n}$ — умовний момент першого порядку; $u_i = \frac{x_i - C}{h}$ — умовна варіанта; h — довжина кроку (тобто відстань між сусідніми варіантами); C — хибний нуль (значення варіанти, розміщеної

приблизно посередині варіаційного ряду); $M_2^* = \frac{\sum_i n_i u_i^2}{n}$ — умовний момент другого порядку.

Приклад 1. Знайти методом добутків вибіркове середнє та вибіркву дисперсію для такого розподілу вибірки обсягу $n = 100$:

x_i	24	28	32	36	40	44	48
n_i	6	13	20	45	11	3	2

Розв'язання. Складемо розрахункову табл. 5, а саме:

- 1) запишемо значення варіант у перший стовпець;
- 2) запишемо значення частот у другий стовпець, суму частот (обсяг вибірки) — у нижній його клітинці;
- 3) як хибний нуль C оберемо значення четвертої варіанти $x_4 = 36$, яка має найбільшу частоту (як C можна взяти значення будь-якої варіанти, розміщеної приблизно посередині стовпця), у клітинці третього стовпця, яка відповідає хибному нулю, запишемо 0, над ним послідовно записуємо -1 , -2 , -3 , а під ним -1 , 2 , 3 ;
- 4) добутки частот n_i на значення відповідних умовних варіант u_i запишемо у четвертий стовпець, окремо знайдемо суму всіх чисел четвертого стовпця, яку помістимо в нижню клітинку четвертого стовпця;
- 5) добутки частот n_i на значення квадратів відповідних умовних варіант u_i^2 , тобто $n_i u_i^2$, запишемо у п'ятий стовпець (зручніше перемножити числа кожного рядка третього й четвертого стовпців: $u_i \cdot n_i u_i = n_i u_i^2$), окремо знайдемо суму всіх чисел п'ятого стовпця, яку помістимо в нижню клітинку п'ятого стовпця;
- 6) добутки частот n_i на значення квадратів відповідних умовних варіант u_i , збільшених на одиницю, тобто $n_i (u_i + 1)^2$, запишемо в шостий стовпець, окремо знайдемо суму всіх чисел шостого стовпця, яку помістимо в нижню клітинку шостого стовпця.

Отже, ми отримали розрахункову табл. 5.

Таблиця 5

1	2	3	4	5	6
x_i	n_i	u_i	$n_i u_i$	$n_i u_i^2$	$n_i (u_i + 1)^2$
24	6	-3	-18	54	24
28	13	-2	-26	52	13
32	20	-1	-20	20	0
36	45	0	0	0	45
40	11	1	11	11	44
44	3	2	6	12	27
48	2	3	6	18	32
	$n = 100$		$\sum_i n_i u_i = -41$	$\sum_i n_i u_i^2 = 167$	$\sum_i n_i (u_i + 1)^2 = 185$

Для контролю обчислень скористаємося тотожністю

$$\sum_i n_i (u_i + 1)^2 \equiv \sum_i n_i u_i^2 + 2 \sum_i n_i u_i + n.$$

Контроль:

$$\sum_i n_i (u_i + 1)^2 = 185;$$

$$\sum_i n_i u_i^2 + 2 \sum_i n_i u_i + n = 167 + 2 \cdot (-41) + 100 = 185.$$

Збіг контрольних сум свідчить про правильність обчислень.

Обчислимо умовні моменти першого й другого порядку:

$$M_1^* = \frac{\sum_i n_i u_i}{n} = \frac{-41}{100} = -0,41; \quad M_2^* = \frac{\sum_i n_i u_i^2}{n} = \frac{167}{100} = 1,67.$$

Знайдемо крок (різницю між будь-якими двома сусідніми варіантами):

$$h = 28 - 24 = 4.$$

Визначимо шукані вибіркові середнє й дисперсію, ураховуючи, що хибний нуль (значення варіанти, яка має найбільшу частоту) $C = 36$:

$$\bar{x}_v = M_1^* h + C = -0,41 \cdot 4 + 36 = 34,36;$$

$$D_v = [M_2^* - (M_1^*)^2] h^2 = [1,67 - (-0,41)^2] \cdot 4^2 = 24,0304. \quad \square$$

Якщо варіанти вибірки не є рівновіддаленими, то інтервал, у якому розміщені всі варіанти вибірки, ділять на кілька однакових за довжиною частинних інтервалів (кожний з них має містити не менше 8–10 варіант). Потім знаходять середини частинних інтервалів, які й утворюватимуть послідовність рівновіддалених варіант. Як частоту кожної середини частинного інтервалу беруть суму частот варіант, які потрапили у відповідний частинний інтервал.

При обчисленні вибіркової дисперсії з метою зменшення похибки, зумовленої групуванням (особливо при малій кількості інтервалів), роблять *поправку Шеппарда*, а саме обчислюють дисперсію за формулою

$$D'_v = D_v - \frac{1}{12} h^2.$$

Приклад 2. Знайти методом добутків вибіркове середнє та вибіркову дисперсію для такого розподілу вибірки обсягу $n = 100$:

x_i	2	8	9	13	15	18	20	21	24	27
n_i	8	1	15	13	9	3	4	10	22	15

Розв'язання. Розіб'ємо інтервал $[2; 27]$ на такі п'ять частинних інтервалів довжиною $h = 5$:

$$[2; 7], \quad (7; 12], \quad (12; 17], \quad (17; 22], \quad (22; 27].$$

Новими варіантами будуть середини цих частинних інтервалів:

$$y_1 = \frac{2+7}{2} = 4,5; \quad y_2 = \frac{7+12}{2} = 9,5; \quad y_3 = \frac{12+17}{2} = 14,5;$$

$$y_4 = \frac{17+22}{2} = 19,5; \quad y_5 = \frac{22+27}{2} = 24,5.$$

Як частоти m_i варіант y_i візьмемо суму частот варіант, які потрапили у відповідний i -й інтервал:

$$\begin{aligned} m_1 &= n_1 = 8; & m_2 &= n_2 + n_3 = 1 + 15 = 16; \\ m_3 &= n_4 + n_5 = 13 + 9 = 22; & m_4 &= n_6 + n_7 + n_8 = 3 + 4 + 10 = 17; \\ & & m_5 &= n_9 + n_{10} = 22 + 15 = 37. \end{aligned}$$

Отже, ми отримали такий статистичний розподіл рівновіддалених варіант:

y_i	4,5	9,5	14,5	19,5	24,5
m_i	8	16	22	17	37

Як і в прикладі 1, складемо розрахункову табл. 6.

Таблиця 6

1	2	3	4	5	6
y_i	m_i	u_i	$m_i u_i$	$m_i u_i^2$	$m_i (u_i + 1)^2$
4,5	8	-2	-16	32	8
9,5	16	-1	-16	16	0
14,5	22	0	0	0	22
19,5	17	1	17	17	68
24,5	37	2	74	148	333
$n = 100$			$\sum_i m_i u_i = 59$	$\sum_i m_i u_i^2 = 213$	$\sum_i m_i (u_i + 1)^2 = 431$

Контроль:

$$\begin{aligned} \sum_i m_i (u_i + 1)^2 &= 431; \\ \sum_i m_i u_i^2 + 2 \sum_i m_i u_i + n &= 213 + 2 \cdot 59 + 100 = 431. \end{aligned}$$

Обчислимо умовні моменти першого й другого порядку:

$$\mathbf{M}_1^* = \frac{\sum_i m_i u_i}{n} = \frac{59}{100} = 0,59; \quad \mathbf{M}_2^* = \frac{\sum_i m_i u_i^2}{n} = \frac{213}{100} = 2,13.$$

Методом добутків знайдемо вибіркові середню й дисперсію, урахувавши, що хибний нуль $C = 14,5$:

$$\bar{x}_b = M_1^* h + C = 0,59 \cdot 5 + 14,5 = 17,45;$$

$$D_b = \left[M_2^* - (M_1^*)^2 \right] h^2 = [2,13 - 0,59^2] \cdot 5^2 = 44,5475.$$

Оскільки кількість частинних інтервалів мала (п'ять), скористаємося поправкою Шешпарда:

$$D'_b = D_b - \frac{1}{12} h^2 = 44,5475 - \frac{5^2}{12} \approx 42,46. \quad \square$$

12.4. Метод сум обчислення вибіркового середнього та вибіркової дисперсії

Нехай вибірку задано у вигляді розподілу рівновіддалених варіант і відповідних їм частот. У цьому разі вибіркоче середнє й вибіркочу дисперсію можна обчислити за формулами (див. підрозд. 12.3.)

$$\bar{x}_b = M_1^* h + C, \quad D_b = \left[M_2^* - (M_1^*)^2 \right] h^2.$$

При використанні *методу сум* умовні моменти першого й другого порядку обчислюють за формулами

$$M_1^* = \frac{d_1}{n}, \quad M_2^* = \frac{s_1 + 2s_2}{n},$$

де $d_1 = a_1 - b_1$, $s_1 = a_1 + b_1$, $s_2 = a_2 + b_2$. Отже, потрібно обчислити значення a_1 , a_2 , b_1 , b_2 . Як практично розрахувати ці числа, показано в прикладі 1.

Приклад 1. Знайти методом сум вибіркоче середнє й вибіркочу дисперсію за заданим статистичним розподілом вибірки обсягу $n = 100$:

x_i	52	58	64	70	76	82	88	94	100	106
n_i	4	8	10	13	18	21	14	9	2	1

Розв'язання. Складемо розрахункову табл. 7, а саме:

- 1) запишемо значення варіант у перший стовпець;
- 2) запишемо значення частот у другий стовпець, суму частот (обсяг вибірки) — у нижній його клітинці;
- 3) як хибний нуль C оберемо значення шостої варіанти $x_6 = 82$, яка має найбільшу частоту (як C можна взяти значення будь-якої варіанти, розміщеної приблизно посередині стовпця), у клітинках рядка, який відповідає хибному нулю, запишемо нулі, у четвертому стовпці над і під нулем — іще по одному нулю;
- 4) у порожніх клітинках третього стовпця, що розміщені над нулем (крім верхньої), зверху вниз запишемо послідовно нагромаджені частоти:

$$4, \quad 4 + 8 = 12, \quad 12 + 10 = 22, \quad 22 + 13 = 35, \quad 35 + 18 = 53.$$

Склавши всі нагромаджені частоти, отримаємо число $b_1 = 126$, яке запишемо у верхню клітинку третього стовпця. У порожніх клітинках третього стовпця, що розміщені під нулем (крім нижньої), знизу вверх запишемо послідовно нагромаджені частоти:

$$1, \quad 1 + 2 = 3, \quad 3 + 9 = 12, \quad 12 + 14 = 26.$$

Склавши всі нагромаджені частоти, отримаємо число $a_1 = 42$, яке запишемо в нижню клітинку третього стовпця;

- 5) аналогічно заповнимо четвертий стовпець, причому нагромаджуються частоти третього стовпця. Суму нагромаджених частот, розміщених над нулем, позначимо b_2 і запишемо у верхню клітинку четвертого стовпця. Суму нагромаджених частот, розміщених під нулем, позначимо a_2 і запишемо в нижню клітинку четвертого стовпця.

Отже, ми отримали розрахункову табл. 7.

Таблиця 7

1	2	3	4
x_i	n_i	$b_1 = 126$	$b_2 = 131$
52	4	4	4
58	8	12	16
64	10	22	38
70	13	35	73
76	18	53	0
82	21	0	0
88	14	26	0
94	9	12	16
100	2	3	4
106	1	1	1
	$n = 100$	$a_1 = 42$	$a_2 = 21$

Знайдемо d_1, s_1, s_2 :

$$d_1 = a_1 - b_1 = 42 - 126 = -84;$$

$$s_1 = a_1 + b_1 = 42 + 126 = 168;$$

$$s_2 = a_2 + b_2 = 21 + 131 = 152.$$

Обчислимо умовні моменти першого й другого порядку:

$$M_1^* = \frac{d_1}{n} = \frac{-84}{100} = -0,84;$$

$$M_2^* = \frac{s_1 + 2s_2}{n} = \frac{168 + 2 \cdot 152}{100} = 4,72.$$

Методом добутків знайдемо вибіркві середню та дисперсію, урахувавши, що хибний нуль $C = 82$, а крок (відстань між двома сусідніми варіантами) $h = 6$; отримаємо

$$\bar{x}_в = M_1^* h + C = -0,84 \cdot 6 + 82 = 76,96;$$

i

$$D_v = [M_2^* - (M_1^*)^2] h^2 = [4,72 - (-0,84)^2] \cdot 6^2 = 144,5184. \quad \square$$

Задачі до розділу 12

Задача 1. Під час дослідження кількісної ознаки X із генеральної сукупності було отримано вибірку

$$5, 7, 4, 6, 5, 5, 5, 7, 5, 6, 6, 5, 5, 6, 4, 6, 5, 6, 4, 5.$$

Знайти обсяг вибірки, побудувати варіаційний ряд вибірки та її статистичний розподіл.

Відповідь. $n = 20$. Варіаційний ряд вибірки: 4, 4, 4, 5, 5, 5, 5, 5, 5,

5, 5, 5, 6, 6, 6, 6, 6, 7, 7. Статистичний розподіл:

x_i	4	5	6	7
n_i	3	9	6	2

.

Задача 2. Вибірку задано розподілом частот:

$$\frac{x_i}{n_i} \left| \begin{array}{c|c|c|c|c|c|c|c} -3 & -1 & 0 & 4 & 8 & 13 & 17 \\ \hline 1 & 3 & 4 & 7 & 5 & 8 & 12 \end{array} \right.$$

Знайти розподіл відносних частот.

Відповідь.

x_i	-3	-1	0	4	8	13	17
w_i	0,025	0,075	0,1	0,175	0,125	0,2	0,3

.

Задача 3. Вибірку задано інтервальним розподілом частот:

$$\frac{x_i}{n_i} \left| \begin{array}{c|c|c|c|c} (0; 2] & (2; 4] & (4; 6] & (6; 8] & (8; 10] \\ \hline 5 & 14 & 16 & 11 & 4 \end{array} \right.$$

Знайти розподіл відносних частот.

Відповідь.

x_i	(0; 2]	(2; 4]	(4; 6]	(6; 8]	(8; 10]
w_i	0,1	0,28	0,32	0,22	0,08

.

Задача 4. Знайти емпіричну функцію розподілу за даним розподілом вибірки:

x_i	1	4	8	13	19	26	34
n_i	5	14	17	22	25	11	6

Відповідь. $F^*(x) = \begin{cases} 0 & \text{при } x \leq 1; \\ 0,05 & \text{при } 1 < x \leq 4; \\ 0,19 & \text{при } 4 < x \leq 8; \\ 0,36 & \text{при } 8 < x \leq 13; \\ 0,58 & \text{при } 13 < x \leq 19; \\ 0,83 & \text{при } 19 < x \leq 26; \\ 0,94 & \text{при } 26 < x \leq 34; \\ 1 & \text{при } x > 34. \end{cases}$

Задача 5. Побудувати полігон частот за даним розподілом вибірки:

x_i	3	5	8	14	16	20	22
n_i	32	41	18	56	23	12	47

Відповідь. Шуканий полігон частот зображено на рис. 41.

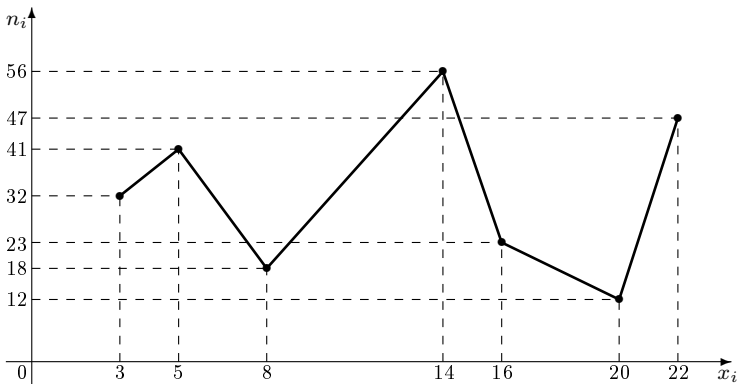


Рис. 41

Задача 6. Побудувати полігон відносних частот за даним розподілом вибірки:

x_i	1	7	11	12	18	26	31
w_i	0,15	0,2	0,1	0,15	0,1	0,05	0,25

Вказівка. Будувати так, ніби замість відносних частот зазначено частоти.

Відповідь. Шуканий полігон емпіричних імовірностей зображено на рис. 42.

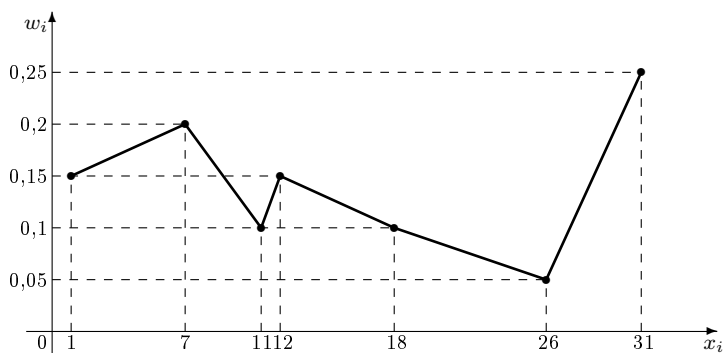


Рис. 42

Задача 7. Вибірку задано інтервальним розподілом частот:

$(x_i; x_{i+1}]$	(5; 10]	(10; 15]	(15; 20]	(20; 25]	(25; 30]	(30; 35]	(35; 40]
n_i	34	12	23	18	42	30	7

Побудувати гістограму частот.

Відповідь. Шукану гістограму частот зображено на рис. 43.

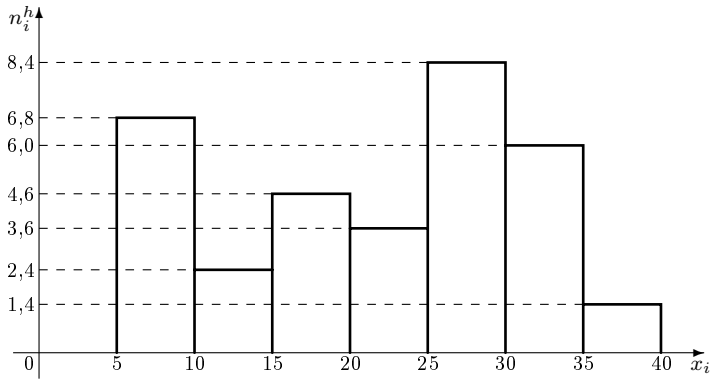


Рис. 43

Задача 8. Вибірку задано розподілом частот:

x_i	5	11	17	23	29	35	41
n_i	26	54	33	67	39	21	10

Побудувати гістограму відносних частот.

Відповідь. Шукану гістограму емпіричних імовірностей зображено на рис. 44.

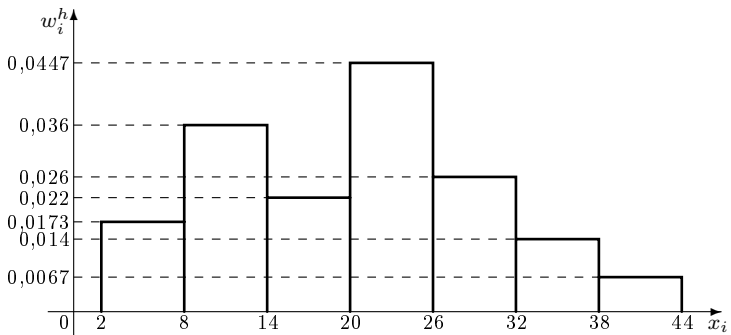


Рис. 44

Задача 9. Дано вибірку: 15; 18; 11; 12; 35; 46; 10; 41; 23; 27; 31; 34; 9; 8; 15; 13; 17; 22; 21; 44; 19; 14; 30; 25; 10; 14; 11; 18; 48; 20; 12; 32; 7; 13; 15; 41; 23; 19; 24; 28; 12; 9; 33; 17; 16; 8; 29; 15; 11; 40.

За даними вибірки побудувати інтервальний статистичний розподіл, полігон і гістограму частот.

Відповідь. Інтервальний статистичний розподіл вибірки такий:

Номер інтервалу i	Частинний інтервал $(x_i, x_{i+1}]$	Сума частот варіант інтервалу n_i
1	(5,1; 11,5]	11
2	(11,5; 17,9]	13
3	(17,9; 24,3]	10
4	(24,3; 30,7]	5
5	(30,7; 37,1]	5
6	(37,1; 43,5]	3
7	(43,5; 49,9]	3

Шукані полігон і гістограму частот зображено відповідно на рис. 45 і 46.

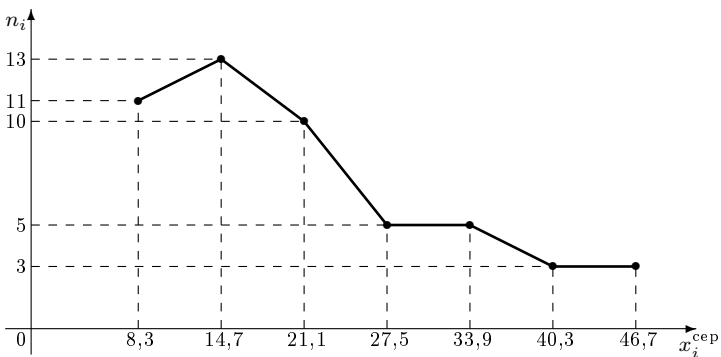


Рис. 45

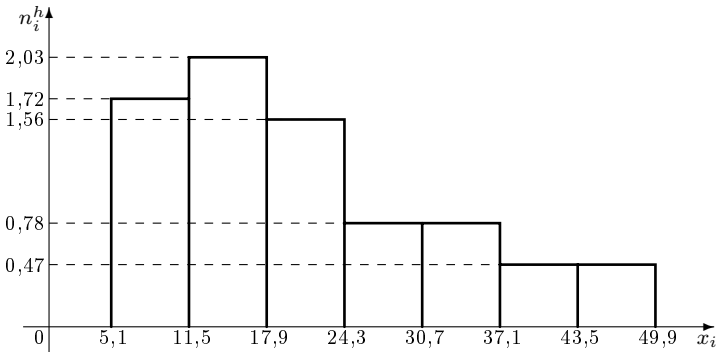


Рис. 46

Задача 10. Задано статистичний розподіл вибірки:

x_i	2	3	5	8	11	12	18
n_i	2	7	15	32	21	14	9

Знайти вибіркове середнє, вибіркєву дисперсію та середнє квадратичнє відхиленнє вибірки.

Відповідь. $\bar{x}_в = 9,17$; $D_в = 15,5811$; $\sigma \approx 3,95$.

Задача 11. Задано інтервальний статистичний розподіл вибірки:

$(x_i; x_{i+1}]$	(1; 3]	(3; 5]	(5; 7]	(7; 9]	(9; 11]	(11; 13]	(13; 15]
w_i	0,15	0,05	0,2	0,1	0,05	0,25	0,2

Знайти вибіркєву середнє, вибіркєву дисперсію та середнє квадратичнє відхиленнє вибірки.

Відповідь. $\bar{x}_в = 8,8$; $D_в = 17,76$; $\sigma \approx 4,214$.

Задача 12. На одному з відрізків автостради планується створити зупинку автобуса. Розподіл населених пунктів із чисельністю їх населеннє наведено в таблиці.

На якому кілометрі автостради розташований населений пункт, км	3	6	8	14	20	22	25
Чисельність населення, тис. чол.	4	3	5	6	2	3	1

На якому кілометрі автостради потрібно розташувати цю зупинку, щоб сумарна відстань, яку покриватимуть потенційні пасажирки до цієї зупинки, була найменшою.

Відповідь. На 11-му кілометрі.

Задача 13. Задано інтервальний статистичний розподіл вибірки:

$(x_i, x_{i+1}]$	(4; 8]	(8; 12]	(12; 16]	(16; 20]	(20; 24]	(24; 28]	(28; 32]
n_i	3	6	6	10	14	5	1

Знайти медіану, моду, варіаційний розмах.

Відповідь. $Me = 19$; $Mo = \frac{276}{13} \approx 21,23$; $R = 28$.

Задача 14. Знайти методом добутків вибіркове середнє та вибіркору дисперсію для такого розподілу вибірки обсягу $n = 100$:

x_i	40	45	50	55	60	65	70
n_i	12	18	20	24	14	7	5

Вказівка. Як хибний нуль вибрати значення четвертої варіанти: $C = x_4 = 55$.

Відповідь. $\bar{x}_B = 52,55$; $D_B = 65,7475$.

Задача 15. Знайти методом добутків вибіркове середнє та вибіркору дисперсію для такого розподілу вибірки обсягу $n = 100$:

x_i	5	7	10	11	14	15	18	20	22	25
n_i	6	8	11	18	22	11	7	7	5	5

Вказівка. Розбити весь інтервал на п'ять частинних інтервалів, як хибний нуль вибрати значення середини третього інтервалу. При визначенні вибіркової дисперсії урахувати поправку Шеппарда.

Відповідь. $\bar{x}_B = 14,08$; $D_B \approx 20,06$.

Задача 16. Знайти методом сум вибіркоче середнє та вибіркочу дисперсію за заданим статистичним розподілом вибірки обсягу $n = 100$:

x_i	42	49	56	63	70	77	84	91	98	105
n_i	2	5	8	19	26	14	10	8	5	3

Вказівка. Як хибний нуль вибрати значення п'ятої варіанти: $C = x_5 = 70$.

Відповідь. $\bar{x}_B = 72,45$; $D_B = 194,4075$.

Розділ 13. Статистичні оцінки параметрів розподілу

13.1. Точкові оцінки

Статистичною оцінкою Θ^* невідомого параметра Θ теоретичного розподілу називають функцію $f(X_1, X_2, \dots, X_n)$ від спостережуваних випадкових величин X_1, X_2, \dots, X_n .

Точковою називають статистичну оцінку, яка визначається одним єдиним числом $\Theta^* = f(x_1, x_2, \dots, x_n)$, де x_1, x_2, \dots, x_n — результати n спостережень над кількісною ознакою випадкової величини X (вибірка).

Незміщеною називають точкову оцінку Θ^* , математичне сподівання якої дорівнює оцінюваному параметру Θ при будь-якому обсягу вибірки:

$$M(\Theta^*) = \Theta.$$

Зміщеною називають точкову оцінку Θ^* , математичне сподівання якої відмінне від оцінюваного параметра Θ :

$$M(\Theta^*) \neq \Theta.$$

Незміщеною оцінкою генерального середнього (математичного сподівання) є вибіркове середнє

$$\bar{x}_B = \frac{\sum_{i=1}^k n_i x_i}{n},$$

де n_i — частота варіанти x_i ; x_i — варіанта вибірки; $n = \sum_{i=1}^k n_i$ — обсяг вибірки.

Зауваження 1. Якщо варіанти x_i вибірки є дуже великими або дуже малими (близькими до нуля) числами, то для спрощення розрахунків доцільно відняти (у разі великих від'ємних чисел — додати) від кожної варіанти одне й те саме число C (як C можна вибрати

будь-яке число, розміщене приблизно посередині варіаційного ряду), потім поділити (у разі близьких до нуля чисел — помножити) на одне й те саме число b (як b можна вибрати найбільше спільне кратне), тобто перейти до умовних варіант

$$u_i = \frac{x_i \pm C}{b} \quad (u_i = x_i b).$$

Тоді

$$\bar{x}_B = b \frac{\sum_{i=1}^k n_i u_i}{n} \mp C = b \bar{u}_B \mp C \quad \left(\bar{x}_B = \frac{\sum_{i=1}^k n_i u_i}{bn} = \frac{\bar{u}_B}{b} \right).$$

Зміщеною оцінкою генеральної дисперсії є вибіркова дисперсія

$$D_B = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n}.$$

Ця оцінка зміщена, оскільки

$$M(D_B) = \frac{n-1}{n} D_G.$$

Для обчислення вибіркової дисперсії можна скористатися більш зручною формулою

$$D_B = \overline{x^2}_B - \bar{x}_B^2 = \frac{\sum_{i=1}^k n_i x_i^2}{n} - \frac{\left(\sum_{i=1}^k n_i x_i \right)^2}{n^2}.$$

Зауваження 2. Якщо варіанти x_i вибірки є дуже великими або дуже малими (близькими до нуля) числами, то для спрощення розрахунків доцільно відняти (у разі великих від'ємних чисел — додати) від кожної варіанти одне й те саме число C (як C можна вибрати будь-яке число, розміщене приблизно посередині варіаційного ряду), потім поділити (у разі близьких до нуля чисел — помножити) на одне

й те саме число b (як b можна вибрати найбільше спільне кратне), тобто перейти до умовних варіант

$$u_i = \frac{x_i \pm C}{b} \quad (u_i = x_i b).$$

Тоді

$$\begin{aligned} \mathbf{D}_B &= b^2 \left(\frac{1}{n} \sum_{i=1}^k n_i u_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^k n_i u_i \right)^2 \right) = b^2 (\overline{u^2}_B - \bar{u}_B^2) \\ \left(\mathbf{D}_B &= \frac{\sum_{i=1}^k n_i u_i^2}{b^2 n} - \frac{\left(\sum_{i=1}^k n_i u_i \right)^2}{(bn)^2} = \frac{\overline{u^2}_B - \bar{u}_B^2}{b} \right). \end{aligned}$$

Незміщеною оцінкою генеральної дисперсії є виправлена вибіркова дисперсія

$$s^2 = \frac{n}{n-1} \mathbf{D}_B.$$

Поправку $\frac{n}{n-1}$ називають *поправкою Бесселя*.

Приклад 1. З генеральної сукупності отримано деяку вибірку обсягу $n = 100$:

x_i	50	80	90	110	120	140	170
n_i	12	7	6	6	15	24	30

Знайти незміщену оцінку генерального середнього.

Розв'язання. Незміщеною оцінкою генерального середнього \bar{x}_T є вибіркоче середнє \bar{x}_B . Оскільки варіанти вибірки є великими числами, перейдемо до умовних варіант

$$u_i = \frac{x_i - 110}{10}$$

(як C ми вибрали значення четвертої варіанти $x_4 = 110$, а як b — число 10, оскільки всі варіанти кратні десяти):

$$\begin{aligned}
 u_1 &= \frac{x_1 - 110}{10} = \frac{50 - 110}{10} = -6; & u_2 &= \frac{x_2 - 110}{10} = \frac{80 - 110}{10} = -3; \\
 u_3 &= \frac{x_3 - 110}{10} = \frac{90 - 110}{10} = -2; & u_4 &= \frac{x_4 - 110}{10} = \frac{110 - 110}{10} = 0; \\
 u_5 &= \frac{x_5 - 110}{10} = \frac{120 - 110}{10} = 1; & u_6 &= \frac{x_6 - 110}{10} = \frac{140 - 110}{10} = 3; \\
 & & u_7 &= \frac{x_7 - 110}{10} = \frac{170 - 110}{10} = 6.
 \end{aligned}$$

Тепер можна легко знайти вибіркове середнє:

$$\begin{aligned}
 \bar{x}_B &= b \cdot \frac{\sum_{i=1}^7 n_i u_i}{n} + C = 10 \cdot \frac{\sum_{i=1}^7 n_i u_i}{100} + 110 = \\
 &= \frac{12 \cdot (-6) + 7 \cdot (-3) + 6 \cdot (-2) + 6 \cdot 0 + 15 \cdot 1 + 24 \cdot 3 + 30 \cdot 6}{10} + \\
 &+ 110 = \frac{-72 + (-21) + (-12) + 0 + 15 + 72 + 180}{10} + 110 = \\
 &= \frac{162}{10} + 110 = 126,2. \quad \square
 \end{aligned}$$

Приклад 2. З генеральної сукупності отримано деяку вибірку обсягу $n = 100$:

x_i	0,002	0,004	0,006	0,008	0,01	0,012	0,014
n_i	7	29	35	12	9	5	3

Знайти зміщену оцінку генеральної дисперсії.

Розв'язання. Зміщеною оцінкою генеральної дисперсії D_Γ є вибіркова дисперсія D_B . Оскільки варіанти вибірки є малими близькими до нуля числами, перейдемо до умовних варіант

$$u_i = x_i \cdot 1000$$

(як b вибрано число 1000, оскільки в такому разі ми отримуємо цілі числа):

$$\begin{aligned} u_1 &= x_1 \cdot 1000 = 0,002 \cdot 1000 = 2; & u_2 &= x_1 \cdot 1000 = 0,004 \cdot 1000 = 4; \\ u_3 &= x_1 \cdot 1000 = 0,006 \cdot 1000 = 6; & u_4 &= x_1 \cdot 1000 = 0,008 \cdot 1000 = 8; \\ u_5 &= x_1 \cdot 1000 = 0,01 \cdot 1000 = 10; & u_6 &= x_1 \cdot 1000 = 0,012 \cdot 1000 = 12; \\ & & u_7 &= x_1 \cdot 1000 = 0,014 \cdot 1000 = 14. \end{aligned}$$

Тепер можна легко знайти вибірккову дисперсію:

$$\begin{aligned} \mathbf{D}_b &= \frac{\sum_{i=1}^7 n_i u_i^2}{b^2 n} - \frac{\left(\sum_{i=1}^7 n_i u_i\right)^2}{(bn)^2} = \\ &= \frac{7 \cdot 2^2 + 29 \cdot 4^2 + 35 \cdot 6^2 + 12 \cdot 8^2 + 9 \cdot 10^2 + 5 \cdot 12^2 + 3 \cdot 14^2}{1000^2 \cdot 100} - \\ &\quad - \left[\frac{7 \cdot 2 + 29 \cdot 4 + 35 \cdot 6 + 12 \cdot 8 + 9 \cdot 10 + 5 \cdot 12 + 3 \cdot 14}{1000 \cdot 100} \right]^2 = \\ &= \frac{28 + 232 + 1260 + 768 + 900 + 720 + 588}{100000000} - \\ &\quad - \left[\frac{14 + 116 + 210 + 96 + 90 + 60 + 42}{100000} \right]^2 = \\ &= 4496 \cdot 10^{-8} - 394384 \cdot 10^{-10} = 5,5216 \cdot 10^{-6}. \quad \square \end{aligned}$$

Приклад 3. За даними вибірки обсягу $n = 21$ знайдено вибірккову дисперсію $\mathbf{D}_b = 5$.

Знайти незміщену оцінку генеральної дисперсії.

Розв'язання. Незміщеною оцінкою генеральної дисперсії є виправлена дисперсія:

$$s^2 = \frac{n}{n-1} \mathbf{D}_b = \frac{21}{21-1} \cdot 5 = 5,25. \quad \square$$

Приклад 4. У результаті статистичних досліджень випадкової величини X отримано таку вибірку: 47, 45, 46, 46, 46, 45, 47, 44, 46, 45,

45, 46, 46, 44, 46, 48, 46, 46, 45, 46, 44, 46, 45, 47, 46, 46, 47, 46, 46, 48, 44, 46, 45, 46, 45, 46, 44, 47, 46, 46, 45, 47, 48, 44, 46, 46, 45, 46, 47, 45.

Знайти незміщені оцінки генерального середнього та генеральної дисперсії.

Розв'язання. Знайдемо обсяг вибірки: $n = 50$. Побудуємо статистичний розподіл вибірки:

x_i	44	45	46	47	48
n_i	6	11	23	7	3

Контроль: $n = 6 + 11 + 23 + 7 + 3 = 50$.

Незміщеною оцінкою генерального середнього є вибіркове середнє

$$\bar{x}_B = 46 + \frac{-2 \cdot 6 + (-1) \cdot 11 + 0 \cdot 23 + 1 \cdot 7 + 2 \cdot 3}{50} = 46 - \frac{10}{50} = 45,8.$$

Щоб знайти незміщену оцінку генеральної дисперсії — виправлену вибіркочну дисперсію, визначимо вибіркочну дисперсію й помножимо її на поправку Бесселя:

$$\begin{aligned} D_B &= \frac{(-2)^2 \cdot 6 + (-1)^2 \cdot 11 + 0^2 \cdot 23 + 1^2 \cdot 7 + 2^2 \cdot 3}{50} - \\ &\quad - \left[\frac{(-2) \cdot 6 + (-1) \cdot 11 + 0 \cdot 23 + 1 \cdot 7 + 2 \cdot 3}{50} \right]^2 = \\ &= \frac{54}{50} - \left[\frac{-10}{50} \right]^2 = 1,04; \\ s^2 &= \frac{50}{49} \cdot 1,04 = \frac{52}{49} \approx 1,06. \end{aligned}$$

Отже, незміщеними оцінками генерального середнього та генеральної дисперсії є

$$\bar{x}_r^* = 45,8; \quad D_r^* \approx 1,06. \quad \square$$

13.2. Метод моментів

Методом моментів знаходження точкових оцінок називають метод, при якому для обчислення невідомих параметрів заданого розподілу прирівнюють відповідні теоретичні та емпіричні моменти.

Якщо розподіл визначається одним параметром, то для знаходження його оцінки прирівнюють математичне сподівання до вибіркового середнього:

$$M(X) = \bar{x}_v,$$

а потім із цього рівняння визначають шукану точкову оцінку невідомого параметра.

Якщо розподіл визначається двома параметрами, то їх точкові оцінки знаходять із системи рівнянь

$$\begin{cases} M(X) = \bar{x}_v, \\ D(X) = D_v. \end{cases}$$

Лівими частинами цих рівнянь є математичне сподівання та дисперсія, які прирівнюють відповідно до вибіркового середнього та вибіркової дисперсії.

Приклад 1. Випадкова величина X (кількість бракованих деталей у партії товару) розподілена за законом Пуассона з параметром λ . У результаті статистичних досліджень отримано такий статистичний розподіл кількості бракованих деталей у $n = 1000$ партіях товару:

Кількість бракованих деталей	0	1	2	3	4	5	6
Кількість партій товару	505	284	131	63	12	3	2

Знайти методом моментів точкову оцінку невідомого параметра λ розподілу Пуассона.

Розв'язання. Оскільки розподіл Пуассона залежить лише від одного параметра, то прирівнюємо математичне сподівання та вибіркоче середнє. Математичне сподівання розподілу Пуассона дорівнює його параметру:

$$M(X) = \lambda,$$

а вибіркове середнє

$$\bar{x}_v = \frac{0 \cdot 505 + 1 \cdot 284 + 2 \cdot 131 + 3 \cdot 63 + 4 \cdot 12 + 5 \cdot 3 + 6 \cdot 2}{1000} = 0,81.$$

Отже, точковою оцінкою невідомого параметра λ розподілу Пуассона буде оцінка

$$\lambda^* = 0,81. \quad \square$$

Приклад 2. Випадкова величина X (зріст дорослої людини) розподілена за нормальним розподілом з параметрами a, σ . У результаті статистичних досліджень отримано такий статистичний розподіл зросту дорослих людей для $n = 1000$ осіб:

Зріст, см	Кількість осіб
(145; 155]	24
(155; 165]	112
(165; 175]	263
(175; 185]	322
(185; 195]	202
(195; 205]	66
(205; 215]	11

Знайти методом моментів точкову оцінку невідомих параметрів a, σ нормального розподілу.

Розв'язання. Перетворимо інтервальний статистичний розподіл на точковий, вибравши як варіанти середини частинних інтервалів:

Зріст, см	150	160	170	180	190	200	210
Кількість осіб	24	112	263	322	202	66	11

Оскільки нормальний закон розподілу залежить від двох параметрів, потрібно знайти вибіркове середнє та вибіркиму дисперсію, а потім прирівняти їх відповідно до математичного сподівання й дисперсії.

Перейдемо до умовних варіант

$$u_i = \frac{x_i - 180}{10}$$

(x_i — зріст людини):

$$\begin{aligned} u_1 &= \frac{150 - 180}{10} = -3; & u_2 &= \frac{160 - 180}{10} = -2; \\ u_3 &= \frac{170 - 180}{10} = -1; & u_4 &= \frac{180 - 180}{10} = 0; \\ u_5 &= \frac{190 - 180}{10} = 1; & u_6 &= \frac{200 - 180}{10} = 2; \\ u_7 &= \frac{210 - 180}{10} = 3. \end{aligned}$$

Знайдемо вибіркове середнє та вибіркву дисперсію:

$$\begin{aligned} \bar{u}_B &= \frac{-3 \cdot 24 - 2 \cdot 112 - 1 \cdot 263 + 0 \cdot 322 + 1 \cdot 202 + 2 \cdot 66 + 3 \cdot 11}{1000} = \\ &= -0,192; \end{aligned}$$

$$\bar{x}_B = 10 \cdot \bar{u}_B + 180 = 178,08;$$

$$\begin{aligned} D_B &= 10^2 \cdot [((-3)^2 \cdot 24 + (-2)^2 \cdot 112 + (-1)^2 \cdot 263 + 0^2 \cdot 322 + \\ &+ 1^2 \cdot 202 + 2^2 \cdot 66 + 3^2 \cdot 11] / 1000 - [-0,192]^2 = 145,5136. \end{aligned}$$

Оскільки параметр a нормального закону розподілу є математичним сподіванням, а параметр σ — середнім квадратичним відхиленням, то оцінками цих параметрів є

$$a^* = \bar{x}_B = 178,08; \quad \sigma^* = \sqrt{D_B} = \sqrt{145,5136} \approx 12,0629. \quad \square$$

13.3. Метод найбільшої правдоподібності

Метод найбільшої (максимальної) правдоподібності полягає в знаходженні максимуму функції одного або кількох оцінюваних параметрів.

Припустимо, що X — дискретна випадкова величина з відомим законом розподілу, однак нам невідомий параметр Θ , яким визначається цей закон розподілу. За даними вибірки x_1, x_2, \dots, x_n , отриманої

в результаті спостережень над випадковою величиною X , необхідно знайти точкову оцінку $\Theta^* = \Theta^*(x_1, x_2, \dots, x_n)$ параметра Θ .

Функцією правдоподібності дискретної випадкової величини називають функцію аргументу Θ

$$L(x_1, x_2, \dots, x_n; \Theta) = p(x_1; \Theta) \cdot p(x_2; \Theta) \cdot \dots \cdot p(x_n; \Theta),$$

де $p(x_i; \Theta)$ — імовірність того, що в результаті випробування випадкова величина X набуде значення x_i .

Оцінкою найбільшої або максимальної правдоподібності параметра Θ називають таке його значення Θ^* , при якому функція правдоподібності досягає свого максимуму.

Логарифмічною функцією правдоподібності називають функцію $\ln L$.

Оскільки функції L і $\ln L$ досягають свого максимуму при одному й тому самому значенні аргументу Θ , здебільшого зручніше знаходити максимум функції $\ln L$, а не L .

Якщо випадкова величина X неперервна, то відомою вважається щільність розподілу ймовірностей $f(x)$, а невідомим — параметр, від якого залежить ця щільність.

Функцією правдоподібності неперервної випадкової величини називають функцію аргументу Θ

$$L(x_1, x_2, \dots, x_n; \Theta) = f(x_1; \Theta) \cdot f(x_2; \Theta) \cdot \dots \cdot f(x_n; \Theta).$$

Оцінку найбільшої правдоподібності невідомого параметра розподілу неперервної випадкової величини шукають так само, як і в разі дискретної випадкової величини, а саме:

- 1) визначають похідну $\frac{d \ln L}{d\Theta}$ (або $\frac{dL}{d\Theta}$);
- 2) знаходять корені Θ_i^* рівняння $\frac{d \ln L}{d\Theta} = 0$ (або $\frac{dL}{d\Theta} = 0$). Ці рівняння називають *рівняннями правдоподібності*;
- 3) визначають другу похідну $\frac{d^2 \ln L}{d\Theta^2}$ (або $\frac{d^2 L}{d\Theta^2}$). Корінь Θ_i^* рівняння правдоподібності, для якого друга похідна від'ємна, беруть як оцінку Θ^* найбільшої правдоподібності параметра Θ .

Якщо параметр Θ двовимірний, тобто $\Theta = (\Theta_1, \Theta_2)$, то для відшукування максимуму функції правдоподібності складають і розв'язують систему рівнянь

$$\begin{cases} \frac{\partial \ln L}{\partial \Theta_1} = 0, \\ \frac{\partial \ln L}{\partial \Theta_2} = 0 \end{cases} \quad \text{або} \quad \begin{cases} \frac{\partial L}{\partial \Theta_1} = 0, \\ \frac{\partial L}{\partial \Theta_2} = 0. \end{cases}$$

Приклад 1. Випадкова величина X (кількість битого скляного посуду в одній упаковці) розподілена за законом Пуассона з невідомим параметром λ . У результаті статистичних досліджень отримано такий емпіричний розподіл кількості битого скляного посуду в $n = 1000$ упаковках:

x_i	0	1	2	3	4	5	6
n_i	554	324	98	19	3	1	1

Знайти методом найбільшої правдоподібності точкову оцінку невідомого параметра λ розподілу Пуассона.

Розв'язання. Оскільки випадкова величина X розподілена за законом Пуассона, то функція правдоподібності має вигляд

$$\begin{aligned} L &= \left(\frac{\lambda^0 e^{-\lambda}}{0!} \right)^{554} \left(\frac{\lambda^1 e^{-\lambda}}{1!} \right)^{324} \left(\frac{\lambda^2 e^{-\lambda}}{2!} \right)^{98} \left(\frac{\lambda^3 e^{-\lambda}}{3!} \right)^{19} \times \\ &\quad \times \left(\frac{\lambda^4 e^{-\lambda}}{4!} \right)^3 \left(\frac{\lambda^5 e^{-\lambda}}{5!} \right)^1 \left(\frac{\lambda^6 e^{-\lambda}}{6!} \right)^1 = \frac{\lambda^{600} e^{-1000\lambda}}{2^{133} \cdot 3^{25} \cdot 5^2}. \end{aligned}$$

Запишемо логарифмічну функцію правдоподібності:

$$\ln L = 600 \ln \lambda - 1000\lambda - \ln(2^{133} \cdot 3^{25} \cdot 5^2).$$

Визначимо похідну логарифмічної функції розподілу за λ :

$$\frac{d \ln L}{d\lambda} = \frac{600}{\lambda} - 1000.$$

Прирівнявши її до нуля, знайдемо єдиний корінь рівняння правдоподібності:

$$\lambda^* = 0,6.$$

Оскільки друга похідна логарифмічної функції правдоподібності

$$\frac{d^2 \ln L}{d\lambda^2} = -\frac{600}{\lambda^2} < 0$$

завжди від'ємна, точковою оцінкою максимальної правдоподібності параметра λ розподілу Пуассона буде

$$\lambda^* = 0,6. \quad \square$$

Приклад 2. Випадкова величина X (довжина деталі) розподілена за нормальним законом з невідомими параметрами a, σ . У результаті статистичних досліджень отримано такий емпіричний розподіл довжини $n = 1000$ деталей:

Довжина деталі, см	5,1	5,2	5,3	5,4	5,5	5,6	5,7
Кількість деталей	7	78	289	392	195	36	3

Знайти методом найбільшої правдоподібності точкову оцінку невідомих параметрів a, σ нормального розподілу.

Розв'язання. Щільність розподілу ймовірностей випадкової величини X

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Запишемо функцію правдоподібності:

$$L = \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,1-a)^2}{2\sigma^2}} \right)^7 \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,2-a)^2}{2\sigma^2}} \right)^{78} \times \\ \times \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,3-a)^2}{2\sigma^2}} \right)^{289} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,4-a)^2}{2\sigma^2}} \right)^{392} \times$$

$$\begin{aligned}
 & \times \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,5-a)^2}{2\sigma^2}} \right)^{195} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,6-a)^2}{2\sigma^2}} \right)^{36} \times \\
 & \times \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5,7-a)^2}{2\sigma^2}} \right)^3 = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^{1000} \times \\
 & \times \exp\{-7 \cdot (5,1-a)^2 + 78 \cdot (5,2-a)^2 + 289 \cdot (5,3-a)^2 + \\
 & + 392 \cdot (5,4-a)^2 + 195 \cdot (5,5-a)^2 + 36 \cdot (5,6-a)^2 + \\
 & + 3 \cdot (5,7-a)^2\} / [2\sigma^2] = \\
 & = (2\pi\sigma^2)^{-500} \cdot e^{-\frac{1000a^2 - 10762a + 28965,1}{2\sigma^2}}.
 \end{aligned}$$

Запишемо логарифмічну функцію правдоподібності:

$$\ln L = -500 \cdot \ln 2\pi\sigma^2 - \frac{1000a^2 - 10762a + 28965,1}{2\sigma^2}.$$

Визначимо частинні похідні логарифмічної функції розподілу за a і σ^2 :

$$\begin{aligned}
 \frac{d \ln L}{da} &= -\frac{2000a - 10762}{2\sigma^2}; \\
 \frac{d \ln L}{d\sigma^2} &= -\frac{500}{\sigma^2} + \frac{1000a^2 - 10762a + 28965,1}{2\sigma^4}.
 \end{aligned}$$

Привівнявши їх до нуля, знайдемо розв'язок системи рівнянь:

$$\begin{cases} -\frac{2000a - 10762}{2\sigma^2} = 0; \\ -\frac{500}{\sigma^2} + \frac{1000a^2 - 10762a + 28965,1}{2\sigma^4} = 0. \end{cases}$$

Розв'язком буде:

$$a = 5,381; \quad \sigma^2 = 0,009939.$$

Отже, точковими оцінками максимальної правдоподібності параметрів a , σ нормального розподілу є

$$a^* = 5,381; \quad \sigma^* = \sqrt{0,009939} \approx 0,1. \quad \square$$

13.4. Інтервальні оцінки

Інтервальною називають оцінку, яка визначається числовим інтервалом.

Довірчим називають інтервал (Θ_1, Θ_2) , у який із заданою надійністю α (імовірністю, близькою до одиниці) потрапляє оцінюваний параметр Θ :

$$P(\Theta_1 < \Theta < \Theta_2) = \alpha, \quad \alpha \rightarrow 1.$$

1. *Інтервальною оцінкою з надійністю α математичного сподівання a нормально розподіленої випадкової величини X за вибіркоvim середнім \bar{x}_B при відомому середньому квадратичному відхиленні σ називають довірчий інтервал*

$$\bar{x}_B - t \frac{\sigma}{\sqrt{n}} < a < \bar{x}_B + t \frac{\sigma}{\sqrt{n}},$$

де $t \frac{\sigma}{\sqrt{n}} = \delta$ — точність оцінки; t — значення аргументу функції Лапласа $\Phi(t)$ (дод. 1), при якому $\Phi(t) = \frac{\alpha}{2}$; n — обсяг вибірки; *при невідомому σ (та обсягу вибірки $n < 30$)*

$$\bar{x}_B - t_\alpha \frac{s}{\sqrt{n}} < a < \bar{x}_B + t_\alpha \frac{s}{\sqrt{n}},$$

де t_α знаходять за дод. 4 за заданими n і α ; s — «виправлене» вибіркoве середнє квадратичне відхилення.

2. *Інтервальною оцінкою з надійністю α середнього квадратичного відхилення σ нормально розподіленої випадкової величини X за «виправленим» вибіркovim середнім квадратичним відхиленням s називають довірчий інтервал*

$$\begin{aligned} s(1 - q) < \sigma < s(1 + q) & \quad \text{при} & \quad q < 1; \\ 0 < \sigma < s(1 + q) & \quad \text{при} & \quad q > 1, \end{aligned}$$

де q знаходять за дод. 5 за заданими n і α .

3. Інтервальною оцінкою з надійністю α невідомої ймовірності p біноміального розподілу за відносною частотою w називають довірчий інтервал

$$p_1 < p < p_2,$$

де

$$p_1 = \frac{n}{t^2 + n} \left(w + \frac{t^2}{2n} - t \sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right);$$
$$p_2 = \frac{n}{t^2 + n} \left(w + \frac{t^2}{2n} + t \sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right);$$

n — загальна кількість випробувань; t — значення аргументу функції Лапласа (дод. 1), при якому $\Phi(t) = \frac{\alpha}{2}$; $w = \frac{m}{n}$ — відносна частота; m — кількість появ події.

Зауваження 1. При великих значеннях n (кілька сотень) можна прийняти як межі довірчого інтервалу такі значення:

$$p_1 = w - t \sqrt{\frac{w(1-w)}{n}}; \quad p_2 = w + t \sqrt{\frac{w(1-w)}{n}}.$$

Приклад 1. Знайти довірчий інтервал з надійністю 0,95 для оцінки невідомого математичного сподівання a нормально розподіленої випадкової величини X , якщо дисперсія цієї випадкової величини $\sigma^2 = 16$, вибіркове середнє $\bar{x}_в = 15$, а обсяг вибірки $n = 25$.

Розв'язання. Потрібно визначити довірчий інтервал

$$\bar{x}_в - t \frac{\sigma}{\sqrt{n}} < a < \bar{x}_в + t \frac{\sigma}{\sqrt{n}}.$$

Для цього обчислимо значення t і σ . Із співвідношення

$$\Phi(t) = \frac{\alpha}{2} = \frac{0,95}{2} = 0,475$$

за дод. 1 знаходимо

$$t = 1,96.$$

Середнє квадратичне відхилення

$$\sigma = \sqrt{\sigma^2} = \sqrt{16} = 4.$$

Підставивши всі значення, отримаємо шуканий довірчий інтервал:

$$15 - 1,96 \cdot \frac{4}{\sqrt{25}} < a < 15 + 1,96 \cdot \frac{4}{\sqrt{25}},$$

тобто

$$13,432 < a < 16,568. \quad \square$$

Приклад 2. З конвеєра надходять електричні лампи. Яким має бути мінімальний розмір партії електроламп для того, щоб з надійністю 0,99 точність оцінки математичного сподівання a випадкової величини X , що характеризує тривалість горіння лампи, за вибірко-вим середнім становила $\varepsilon = 1$ год, якщо відомо середнє квадратичне відхилення випадкової величини X : $\sigma = 3$ год? Вважається, що випадкова величина X має нормальний закон розподілу.

Розв'язання. Оскільки довірчий інтервал для математичного сподівання a випадкової величини X обчислюється за формулою

$$\bar{x}_v - t \frac{\sigma}{\sqrt{n}} < a < \bar{x}_v + t \frac{\sigma}{\sqrt{n}},$$

то точність ε оцінки визначається так:

$$\varepsilon = t \frac{\sigma}{\sqrt{n}}.$$

Звідси отримуємо формулу для обчислення мінімального обсягу вибірки, який забезпечує задану точність оцінювання:

$$n = \min\{n: n > t^2 \frac{\sigma^2}{\varepsilon^2}\}.$$

Знайдемо t із співвідношення

$$\Phi(t) = \frac{0,99}{2} = 0,495$$

за дод. 1:

$$t = 2,58.$$

Отже, мінімальний обсяг вибірки

$$n = \min\{n: n > 2,58^2 \cdot \frac{3^2}{1^2} = 23,22\} = 24. \quad \square$$

Приклад 3. У результаті статистичних досліджень випадкової величини X отримано вибірку обсягу $n = 25$ із таким статистичним розподілом:

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

Знайти з надійністю $\alpha = 0,99$ інтервальну оцінку математичного сподівання a випадкової величини X за вибіркоvim середнім. Вважається, що випадкова величина X нормально розподілена.

Розв'язання. Перш ніж знайти довірчий інтервал для математичного сподівання a випадкової величини X за формулою

$$\bar{x}_B - t_\alpha \frac{s}{\sqrt{n}} < a < \bar{x}_B + t_\alpha \frac{s}{\sqrt{n}},$$

необхідно визначити вибіркoве середнє \bar{x}_B , «виправлене» вибіркoве середнє квадратичне відхилення s і t_α за дод. 4:

$$\bar{x}_B = \frac{\sum_{i=1}^7 n_i x_i}{n} = \frac{0 \cdot 1 + 1 \cdot 3 + 2 \cdot 4 + 3 \cdot 6 + 4 \cdot 5 + 5 \cdot 4 + 6 \cdot 2}{25} = 3,24;$$

$$s = \sqrt{\frac{n}{n-1} \mathbf{D}_B} = \sqrt{\frac{n}{n-1} \left(\frac{\sum_{i=1}^7 n_i x_i^2}{n} - [\bar{x}_B]^2 \right)} =$$

$$= \left[\frac{25}{25-1} \left(\frac{0^2 \cdot 1 + 1^2 \cdot 3 + 2^2 \cdot 4 + 3^2 \cdot 6 + 4^2 \cdot 5 + 5^2 \cdot 4 + 6^2 \cdot 2}{25} - 3,24^2 \right) \right]^{\frac{1}{2}} = \sqrt{\frac{25}{24} \cdot 2,5024} \approx 1,6145;$$

$$t_{\alpha} = t_{\alpha}(n, \alpha) = t_{0,99}(25; 0,99) = 2,797.$$

Шуканий довірчий інтервал

$$3,24 - 2,797 \cdot \frac{1,6145}{\sqrt{25}} < a < 3,24 + 2,797 \cdot \frac{1,6145}{\sqrt{25}},$$

тобто

$$2,337 < a < 4,143. \quad \square$$

Приклад 4. За даними вибірки обсягу $n = 20$ визначено «виправлене» вибіркове середнє квадратичне відхилення $s = 2$ нормально розподіленої випадкової величини X .

Знайти довірчий інтервал для середнього квадратичного відхилення σ випадкової величини X з надійністю $\alpha = 0,95$.

Розв'язання. Спочатку знайдемо за дод. 5 значення q :

$$q = q(n, \alpha) = q(20; 0,95) = 0,37.$$

Оскільки $q < 1$, довірчий інтервал для середнього квадратичного відхилення σ випадкової величини X з надійністю $\alpha = 0,95$ обчислюється за формулою

$$s(1 - q) < \sigma < s(1 + q),$$

тобто

$$1,26 < \sigma < 2,74. \quad \square$$

Приклад 5. У результаті статистичних досліджень випадкової величини X отримано вибірку обсягу $n = 15$ із таким статистичним розподілом:

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

Знайти з надійністю $\alpha = 0,999$ інтервальну оцінку середнього квадратичного відхилення σ випадкової величини X . Вважається, що випадкова величина X розподілена за нормальним законом.

Розв'язання. Спочатку знайдемо за дод. 5 значення q і обчислимо «виправлене» вибіркове середнє квадратичне відхилення s :

$$\begin{aligned}
 q &= q(n, \alpha) = q(15; 0,999) = 1,15; \\
 s &= \sqrt{\frac{n}{n-1} \mathbf{D}_B} = \sqrt{\frac{n}{n-1} \left(\frac{\sum_{i=1}^5 n_i x_i^2}{n} - \frac{\left(\sum_{i=1}^5 n_i x_i \right)^2}{n^2} \right)} = \\
 &= \sqrt{\frac{15}{14} \left(\frac{144}{15} - \frac{44^2}{15^2} \right)} = \sqrt{\frac{16}{15}} \approx 1,0328.
 \end{aligned}$$

Оскільки $q > 1$, довірчий інтервал для середнього квадратичного відхилення σ випадкової величини X з надійністю $\alpha = 0,999$ обчислюється за формулою

$$0 < \sigma < s(1 + q),$$

тобто

$$0 < \sigma < 2,22. \quad \square$$

Приклад 6. Проводяться незалежні випробування з однаковою, але невідомою ймовірністю p успіху.

Знайти довірчий інтервал для оцінки ймовірності p з надійністю $\alpha = 0,99$, якщо з 50 випробувань успішними були 12.

Розв'язання. Для того щоб знайти довірчий інтервал для оцінки ймовірності p , необхідно визначити межі цього інтервалу за формулами

$$p_1 = \frac{n}{t^2 + n} \left(w + \frac{t^2}{2n} - t \sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right);$$
$$p_2 = \frac{n}{t^2 + n} \left(w + \frac{t^2}{2n} + t \sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2} \right).$$

Знайдемо значення t із співвідношення

$$\Phi(t) = \frac{\alpha}{2} = 0,495,$$

скориставшись дод. 1:

$$t = 2,58;$$

відносна частота

$$w = \frac{12}{50} = 0,24.$$

Обсяг вибірки $n = 50$. Використавши ці значення, визначимо кінці довірчого інтервалу:

$$p_1 = \frac{50}{2,58^2 + 50} \left[0,24 + \frac{2,58^2}{2 \cdot 50} - 2,58 \sqrt{\frac{0,24(1-0,24)}{50} + \left(\frac{2,58}{2 \cdot 50}\right)^2} \right] \approx 0,12;$$
$$p_2 = \frac{50}{2,58^2 + 50} \left[0,24 + \frac{2,58^2}{2 \cdot 50} + 2,58 \sqrt{\frac{0,24(1-0,24)}{50} + \left(\frac{2,58}{2 \cdot 50}\right)^2} \right] \approx 0,42.$$

Отже, шуканий довірчий інтервал такий:

$$0,12 < p < 0,42.$$

□

Приклад 7. Під час заливання киплячої води в склянку виявилось, що із 1000 склянок тріснули або розбилися 24.

Знайти з надійністю $\alpha = 0,95$ інтервальну оцінку ймовірності пошкодження склянки при заливанні в неї киплячої води.

Розв'язання. Визначимо відносну частоту пошкодження склянки:

$$w = \frac{24}{1000} = 0,024.$$

Знайдемо значення t із співвідношення

$$\Phi(t) = \frac{\alpha}{2} = \frac{0,95}{2} = 0,475,$$

скориставшись дод. 1:

$$t = 1,96.$$

Ураховуючи, що обсяг вибірки досить великий ($n = 1000$), для обчислення значень кінців довірчого інтервалу оцінки ймовірності пошкодження склянки використаємо формули:

$$p_1 = w - t\sqrt{\frac{w(1-w)}{n}} = 0,024 - 1,96\sqrt{\frac{0,024(1-0,024)}{1000}} \approx 0,0145;$$

$$p_2 = w + t\sqrt{\frac{w(1-w)}{n}} = 0,024 + 1,96\sqrt{\frac{0,024(1-0,024)}{1000}} \approx 0,0335.$$

Отже, шуканий довірчий інтервал такий:

$$0,0145 < p < 0,0335. \quad \square$$

Задачі до розділу 13

Задача 1. Із генеральної сукупності отримано таку вибірку обсягу $n = 100$:

x_i	0,03	0,04	0,05	0,06	0,07	0,08	0,09
n_i	15	10	9	4	12	21	29

Знайти незміщену оцінку генерального середнього.

Вказівка. Перейдіть до умовних варіант $u_i = 100x_i - 6$.

Відповідь. $\bar{x}_B = 0,0667$.

Задача 2. Із генеральної сукупності отримано таку вибірку обсягу $n = 100$:

x_i	260	300	340	380	420	460	500
n_i	8	30	36	11	8	4	3

Знайти зміщену оцінку генеральної дисперсії.

Вказівка. Перейдіть до умовних варіант $u_i = \frac{x_i - 380}{40}$.

Відповідь. $D_B = 3020$.

Задача 3. За даними вибірки обсягу $n = 51$ знайдено вибіркочну дисперсію $D_B = 15$.

Знайти незміщену оцінку генеральної дисперсії.

Відповідь. $s^2 = 15,3$.

Задача 4. У результаті статистичних досліджень випадкової величини X отримано таку вибірку: 0,01; 0,03; 0,02; 0,02; 0,03; 0,04; 0,01; 0,05; 0,04; 0,03; 0,03; 0,03; 0,02; 0,03; 0,01; 0,04; 0,03; 0,02; 0,05; 0,03; 0,03; 0,04; 0,02; 0,02; 0,01; 0,03; 0,02; 0,03; 0,04; 0,02; 0,01; 0,03; 0,03; 0,03; 0,05; 0,04; 0,02; 0,02; 0,03; 0,03; 0,05; 0,03; 0,02; 0,03; 0,01; 0,03; 0,04; 0,02; 0,03; 0,03.

Знайти незміщені оцінки генерального середнього та генеральної дисперсії.

Вказівка. Перейдіть до умовних варіант $u_i = 100x_i - 3$.

Відповідь. $\bar{x}_B = 0,0282$, $s^2 = \frac{2869}{24500000} \approx 0,0001171$.

Задача 5. Випадкова величина X (кількість знайдених самородків у одного старателя за один день) розподілена за законом Пуассона з параметром λ . У результаті статистичних досліджень отримано такий статистичний розподіл кількості знайдених самородків у $n = 1000$ старателів:

Розділ 13. Статистичні оцінки параметрів розподілу

Кількість знайдених самородків	0	1	2	3	4	5	6
Кількість старателів	369	367	183	62	15	3	1

Знайти методом моментів точкову оцінку невідомого параметра λ розподілу Пуассона.

Відповідь. $\lambda^* = 1$.

Задача 6. Випадкова величина X (кількість бракованих одиниць товару) розподілена за законом Пуассона з невідомим параметром λ . У результаті статистичних досліджень отримано такий емпіричний розподіл кількості браку в $n = 1000$ одиницях товару:

x_i	0	1	2	3	4	5	6
n_i	82	208	261	217	136	68	28

Знайти методом найбільшої правдоподібності точкову оцінку невідомого параметра λ розподілу Пуассона.

Відповідь. $\lambda^* = 2,433$.

Задача 7. Випадкова величина X (маса тіла дорослої людини) розподілена за нормальним законом з параметрами a , σ . У результаті статистичних досліджень отримано такий статистичний розподіл маси тіла дорослих людей для $n = 1000$ осіб.

Маса тіла, кг	Кількість осіб
(45; 55]	57
(55; 65]	136
(65; 75]	223
(75; 85]	249
(85; 95]	191
(95; 105]	100
(105; 115]	36
(115; 125]	8

Знайти методом моментів точкову оцінку невідомих параметрів a , σ нормального розподілу.

Вказівка. Перетворити інтервальний статистичний розподіл на точковий і перейти до умовних варіант $u_i = \frac{x_i - 80}{10}$ (x_i — маса тіла людини).

Відповідь. $a^* = 78,65$; $\sigma^* \approx 15,18$.

Задача 8. Випадкова величина X (дальність польоту ракети) розподілена за нормальним законом із невідомими параметрами a , σ . У результаті статистичних досліджень отримано такий емпіричний розподіл дальності польоту $n = 1000$ ракет:

Дальність польоту, км	32	33	34	35	36	37	38
Кількість ракет	7	78	289	392	195	36	3

Знайти методом найбільшої правдоподібності точкову оцінку невідомих параметрів a , σ нормального розподілу.

Відповідь. $a^* = 35,3$, $\sigma^* \approx 0,991$.

Задача 9. Знайти довірчий інтервал з надійністю 0,999 для оцінки невідомого математичного сподівання a нормально розподіленої випадкової величини X , якщо дисперсія цієї випадкової величини $\sigma^2 = 25$, вибіркове середнє $\bar{x}_v = 3$, а обсяг вибірки $n = 36$.

Відповідь. $0,25 < a < 5,75$.

Задача 10. На одному й тому самому верстаті виготовляються деталі. Якою має бути мінімальна кількість деталей для того, щоб з надійністю 0,95 точність оцінки математичного сподівання a випадкової величини X , що характеризує довжину деталі, за вибіркочним середнім становила $\varepsilon = 3$ мм, якщо відомо середнє квадратичне відхилення випадкової величини X : $\sigma = 5$ мм? Вважається, що випадкова величина X має нормальний закон розподілу.

Відповідь. $n = 11$.

Задача 11. У результаті статистичних досліджень випадкової величини X отримано вибірку обсягу $n = 50$ із таким статистичним розподілом:

x_i	-3	-2	-1	0	1	2	3
n_i	2	4	8	11	11	9	5

Знайти з надійністю $\alpha = 0,999$ інтервальну оцінку математичного сподівання a випадкової величини X за вибіркоvim середнім. Вважається, що випадкова величина X нормально розподілена.

Відповідь. $-0,349 < a < 1,229$.

Задача 12. За даними вибірки обсягу $n = 25$ знайдено «виправлене» вибіркoве середнє квадратичне відхилення $s = 1$ нормально розподіленої випадкової величини X .

Знайти довірчий інтервал для середнього квадратичного відхилення σ випадкової величини X з надійністю $\alpha = 0,99$.

Відповідь. $0,27 < \sigma < 1,73$.

Задача 13. У результаті статистичних досліджень випадкової величини X отримано вибірку обсягу $n = 10$ із таким статистичним розподілом:

x_i	-1	0	1	2
n_i	1	3	4	2

Знайти з надійністю $\alpha = 0,99$ інтервальну оцінку середнього квадратичного відхилення σ випадкової величини X . Вважається, що випадкова величина X розподілена за нормальним законом.

Відповідь. $0 < \sigma < 1,872$.

Задача 14. Проводяться незалежні випробування з однаковою, але невідомою ймовірністю p успіху.

Знайти довірчий інтервал для оцінки ймовірності p з надійністю $\alpha = 0,95$, якщо з 100 випробувань успішними були 27.

Відповідь. $0,193 < p < 0,364$.

Задача 15. Під час випробовування сталі на міцність виявилось, що із 1000 сталевих прутів не витримали випробування 57.

Знайти з надійністю $\alpha = 0,999$ інтервальну оцінку ймовірності того, що сталевий прут не пройде випробування на міцність.

Відповідь. $0,0328 < p < 0,0812$.

Розділ 14. Елементи теорії регресії і кореляції

14.1. Рівняння прямої лінії регресії. Лінійна кореляція

Вибірковим рівнянням прямої лінії регресії Y на X (X і Y — випадкові величини, що спостерігаються) називають рівняння

$$\bar{y}_x - \bar{y} = r_{xy} \frac{\sigma_y}{\sigma_x} (x - \bar{x}),$$

де \bar{y}_x — умовне середнє; \bar{x} і \bar{y} — вибіркові середні випадкових величин X і Y відповідно; σ_x і σ_y — вибіркові середні квадратичні відхилення; r_{xy} — вибірковий коефіцієнт кореляції, причому

$$r_{xy} = \frac{\sum_{i=1}^n \sum_{j=1}^m w_{ij} x_i y_j - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y};$$

$$\bar{x}_y - \bar{x} = r_{xy} \frac{\sigma_x}{\sigma_y} (y - \bar{y}),$$

де w_{ij} — емпірична ймовірність появи значення (x_i, y_j) ; n — кількість спостережених варіант випадкової величини X ; m — кількість спостережених варіант випадкової величини Y .

Якщо обидві лінії регресії Y на X і X на Y — прямі, кореляцію називають *лінійною*.

Якщо дані спостережень над випадковими величинами X і Y задані кореляційною таблицею з рівновіддаленими варіантами, доцільно перейти до умовних варіант:

$$u_i = \frac{x_i - C_1}{h_1}; \quad v_j = \frac{y_j - C_2}{h_2},$$

де C_1 — «хибний нуль» варіант випадкової величини X ; h_1 — крок варіант випадкової величини X ($h_1 = x_{i+1} - x_i$); C_2 — «хибний нуль» варіант випадкової величини Y ; h_2 — крок варіант випадкової величини Y ($h_2 = y_{j+1} - y_j$).

У цьому разі вибірковий коефіцієнт кореляції

$$r_{xy} = \frac{\sum_{i=1}^n \sum_{j=1}^m w_{ij} u_i v_j - \bar{u} \cdot \bar{v}}{\sigma_u \sigma_v}.$$

Величини \bar{u} , \bar{v} , σ_u , σ_v знаходять за формулами

$$\begin{aligned} \bar{u} &= \sum_{i=1}^n w_i u_i; & w_i &= \sum_{k=1}^m w_{ik}; & \sigma_u &= \sqrt{u^2 - \bar{u}^2}; \\ \bar{v} &= \sum_{j=1}^m w_j v_j; & w_j &= \sum_{k=1}^n w_{kj}; & \sigma_v &= \sqrt{v^2 - \bar{v}^2}, \end{aligned}$$

а при великій кількості даних — методом добутків.

Знаючи ці значення, можна перейти назад до величин, які входять до рівняння регресії, за допомогою формул

$$\begin{aligned} \bar{x} &= \bar{u} h_1 + C_1; & \sigma_x &= \sigma_u h_1; \\ \bar{y} &= \bar{v} h_2 + C_2; & \sigma_y &= \sigma_v h_2. \end{aligned}$$

Приклад 1. Знайти вибіркове рівняння прямої лінії регресії Y на X за даними кореляційної таблиці (табл. 8).

Таблиця 8

X	Y					n_x
	12	15	18	21	24	
45	8	12	—	—	—	20
55	1	9	11	—	—	21
65	—	3	15	14	—	32
75	—	—	4	10	8	22
85	—	—	—	1	4	5
n_y	9	24	30	25	12	$n = 100$

Розв'язання. Перетворимо кореляційну таблицю:

Розділ 14. Елементи теорії регресії і кореляції

- уведемо умовні варіанти $u_i = \frac{x_i - 65}{10}$, $v_j = \frac{y_j - 18}{3}$;
- усі частоти замінимо емпіричними ймовірностями.

Отже, отримано табл. 9.

Таблиця 9

U	V					w_u
	-2	-1	0	1	2	
-2	0,08	0,12	—	—	—	0,2
-1	0,01	0,09	0,11	—	—	0,21
0	—	0,03	0,15	0,14	—	0,32
1	—	—	0,04	0,1	0,08	0,22
2	—	—	—	0,01	0,04	0,05
w_v	0,09	0,24	0,3	0,25	0,12	$w = 1$

Знайдемо \bar{u} , \bar{v} :

$$\begin{aligned} \bar{u} &= \sum_{i=1}^5 w_i u_i = \\ &= 0,2 \cdot (-2) + 0,21 \cdot (-1) + 0,32 \cdot 0 + 0,22 \cdot 1 + 0,05 \cdot 2 = -0,29; \end{aligned}$$

$$\begin{aligned} \bar{v} &= \sum_{j=1}^5 w_j v_j = \\ &= 0,09 \cdot (-2) + 0,24 \cdot (-1) + 0,3 \cdot 0 + 0,25 \cdot 1 + 0,12 \cdot 2 = 0,07. \end{aligned}$$

Визначимо $\overline{u^2}$, $\overline{v^2}$:

$$\begin{aligned} \overline{u^2} &= \sum_{i=1}^5 w_i u_i^2 = 0,2 \cdot (-2)^2 + 0,21 \cdot (-1)^2 + 0,32 \cdot 0^2 + 0,22 \cdot 1^2 + \\ &+ 0,05 \cdot 2^2 = 1,43; \end{aligned}$$

$$\begin{aligned}\bar{v}^2 &= \sum_{j=1}^5 w_j v_j^2 = 0,09 \cdot (-2)^2 + 0,24 \cdot (-1)^2 + 0,3 \cdot 0^2 + 0,25 \cdot 1^2 + \\ &+ 0,12 \cdot 2^2 = 1,33.\end{aligned}$$

Знайдемо σ_u , σ_v :

$$\begin{aligned}\sigma_u &= \sqrt{u^2 - \bar{u}^2} = \sqrt{1,43 - (-0,29)^2} \approx 1,16; \\ \sigma_v &= \sqrt{v^2 - \bar{v}^2} = \sqrt{1,33 - 0,07^2} \approx 1,15.\end{aligned}$$

Визначимо $\sum_{i=1}^5 \sum_{j=1}^5 w_{ij} u_i v_j$:

$$\begin{aligned}\sum_{i=1}^5 \sum_{j=1}^5 w_{ij} u_i v_j &= 0,08 \cdot (-2) \cdot (-2) + 0,12 \cdot (-2) \cdot (-1) + \\ &+ 0,01 \cdot (-1) \cdot (-2) + 0,09 \cdot (-1) \cdot (-1) + \\ &+ 0,11 \cdot (-1) \cdot 0 + 0,03 \cdot 0 \cdot (-1) + 0,15 \cdot 0 \cdot 0 + \\ &+ 0,14 \cdot 0 \cdot 1 + 0,04 \cdot 1 \cdot 0 + 0,1 \cdot 1 \cdot 1 + \\ &+ 0,08 \cdot 1 \cdot 2 + 0,01 \cdot 2 \cdot 1 + 0,04 \cdot 2 \cdot 2 = \\ &= 1,11.\end{aligned}$$

Знайдемо вибірковий коефіцієнт кореляції:

$$r_{xy} = r_{uv} = \frac{\sum_{i=1}^5 \sum_{j=1}^5 w_{ij} u_i v_j - \bar{u} \cdot \bar{v}}{\sigma_u \sigma_v} = \frac{1,11 - (-0,29) \cdot 0,07}{1,16 \cdot 1,15} \approx 0,85.$$

Визначимо \bar{x} , \bar{y} :

$$\begin{aligned}\bar{x} &= \bar{u} \cdot 10 + 65 = -0,29 \cdot 10 + 65 = 62,1; \\ \bar{y} &= \bar{v} \cdot 3 + 18 = 0,07 \cdot 3 + 18 = 18,21.\end{aligned}$$

Обчислимо σ_x , σ_y :

$$\begin{aligned}\sigma_x &= 10 \cdot \sigma_u \approx 10 \cdot 1,16 = 11,6; \\ \sigma_y &= 3 \cdot \sigma_v \approx 3 \cdot 1,15 = 3,45.\end{aligned}$$

Підставивши всі значення у вибіркоче рівняння лінійної регресії, отримаємо

$$\bar{y}_x - 18,21 = 0,85 \cdot \frac{3,45}{11,6}(x - 62,1),$$

тобто

$$\bar{y}_x = 0,25x + 2,56. \quad \square$$

14.2. Рівняння параболічної регресії. Параболічна кореляція

Вибірковим рівнянням кривої лінії регресії Y на X (X і Y — випадкові величини, що спостерігаються) називають рівняння

$$\bar{y}_x = f(x),$$

де $f(x)$ — деяка криволінійна функція. У цьому разі кореляцію називають *криволінійною*.

Частинним випадком вибіркового рівняння кривої лінії регресії є *вибіркоче рівняння параболічної регресії* Y на X :

$$\bar{y}_x = Ax^2 + Bx + C.$$

Невідомі параметри A , B , C рівняння параболічної регресії знаходять з такої системи рівнянь:

$$\left\{ \begin{array}{l} A \sum_{i=1}^m n_{x_i} x_i^4 + B \sum_{i=1}^m n_{x_i} x_i^3 + C \sum_{i=1}^m n_{x_i} x_i^2 = \sum_{i=1}^m n_{x_i} \bar{y}_{x_i} x_i^2, \\ A \sum_{i=1}^m n_{x_i} x_i^3 + B \sum_{i=1}^m n_{x_i} x_i^2 + C \sum_{i=1}^m n_{x_i} x_i = \sum_{i=1}^m n_{x_i} \bar{y}_{x_i} x_i, \\ A \sum_{i=1}^m n_{x_i} x_i^2 + B \sum_{i=1}^m n_{x_i} x_i + Cn = \sum_{i=1}^m n_{x_i} \bar{y}_{x_i}, \end{array} \right. \quad (1)$$

де n_{x_i} — частота варіанти x_i ; $\bar{y}_{x_i} = \frac{\sum_{j=1}^k n_{x_i y_j} y_j}{n_{x_i}}$ — умовне середнє; k — кількість спостережених варіант випадкової величини Y ; $n_{x_i y_j}$ —

частота варіанти (x_i, y_j) ; n — обсяг вибірки; m — кількість спостережених варіант випадкової величини X .

У цьому разі говорять про *параболічну кореляцію другого порядку*.

Для оцінки сили кореляції Y на X знаходять *вибіркове кореляційне відношення* — відношення міжгрупового вибіркового середнього квадратичного відхилення до загального вибіркового середнього квадратичного відхилення:

$$\eta_{yx} = \frac{\sigma_{\bar{y}_x}}{\sigma_y},$$

де

$$\sigma_{\bar{y}_x} = \sqrt{D_{\text{міжгр}}} = \sqrt{\frac{\sum_{i=1}^n n_{x_i} (\bar{y}_{x_i} - \bar{y})^2}{n}};$$

$$\sigma_y = \sqrt{D_{\text{в}}} = \sqrt{\frac{\sum_{j=1}^k n_{y_j} (y_j - \bar{y})^2}{n}}.$$

Аналогічно визначаються рівняння параболічної регресії X на Y і вибіркове кореляційне відношення η_{xy} X до Y .

Таблиця 10

X	Y					n_x
	2	3	5	7	8	
12	5	—	—	2	6	13
14	1	8	—	12	3	24
17	—	6	11	10	1	28
18	—	4	15	8	—	27
20	—	—	7	1	—	8
n_y	6	18	33	33	10	$n = 100$

Приклад 1. Знайти вибіркове рівняння параболічної регресії

$$\bar{y}_x = Ax^2 + Bx + C$$

за даними кореляційної таблиці (табл. 10).

Оцінити силу кореляційного зв'язку за допомогою вибіркового кореляційного відношення.

Розв'язання. Складемо розрахункову таблицю (табл. 11).

Таблиця 11

<i>X</i>	n_{x_i}	\bar{y}_{x_i}	$n_{x_i}x_i$	$n_{x_i}x_i^2$	$n_{x_i}x_i^3$	$n_{x_i}x_i^4$	$n_{x_i}\bar{y}_{x_i}$	$n_{x_i}\bar{y}_{x_i}x_i$	$n_{x_i}\bar{y}_{x_i}x_i^2$
12	13	5,538	156	1872	22464	269568	72	864	10368
14	24	5,583	336	4704	65856	921984	134	1876	26264
17	28	5,393	476	8092	137564	2338588	151	2567	43639
18	27	5,296	486	8748	157464	2834352	143	2574	46332
20	8	5,25	160	3200	64000	1280000	42	840	16800
Σ	100		1614	26616	447348	7644492	542	8721	143403

Підставивши числа з останнього рядка табл. 11 у систему рівнянь (1), отримуємо таку систему рівнянь відносно невідомих коефіцієнтів A, B, C :

$$\begin{cases} 7644492 A + 447348 B + 26616 C = 143403, \\ 447348 A + 26616 B + 1614 C = 8721, \\ 26616 A + 1614 B + 100 C = 542. \end{cases}$$

Розв'язавши цю систему рівнянь, отримуємо

$$A \approx -0,00434; \quad B \approx 0,08868; \quad C \approx 5,1434,$$

тобто рівняння параболічної регресії $\bar{y}_x = Ax^2 + Bx + C$ має вигляд

$$\bar{y}_x = -0,00434x^2 + 0,08868x + 5,1434.$$

Для того щоб знайти вибіркове кореляційне відношення η_{yx} , визначимо спочатку вибіркове середнє \bar{y} , вибіркове середнє квадратичне відхилення σ_y і міжгрупове середнє квадратичне відхилення $\sigma_{\bar{y}_x}$:

$$\bar{y} = \frac{\sum_{j=1}^5 n_{y_j} y_j}{n} = \frac{2 \cdot 6 + 3 \cdot 18 + 5 \cdot 33 + 7 \cdot 33 + 8 \cdot 10}{100} = 5,42;$$

$$\begin{aligned} \sigma_y &= \sqrt{\frac{\sum_{j=1}^k n_{y_j} y_j^2}{n} - \bar{y}^2} = \\ &= \sqrt{\frac{2^2 \cdot 6 + 3^2 \cdot 18 + 5^2 \cdot 33 + 7^2 \cdot 33 + 8^2 \cdot 10}{100} - 5,42^2} \approx 1,82; \end{aligned}$$

$$\begin{aligned} \sigma_{\bar{y}_x} &= \sqrt{\frac{\sum_{i=1}^n n_{x_i} (\bar{y}_{x_i} - \bar{y})^2}{n}} = \left(\frac{13(5,54 - 5,42)^2 + 24(5,58 - 5,42)^2}{100} + \right. \\ &\quad \left. + \frac{28(5,39 - 5,42)^2 + 27(5,3 - 5,42)^2 + 8(5,25 - 5,42)^2}{100} \right)^{\frac{1}{2}} \approx \\ &\approx 0,122. \end{aligned}$$

Отже, шукане вибіркове кореляційне відношення

$$\eta_{yx} = \frac{\sigma_{\bar{y}_x}}{\sigma_y} = \frac{0,122}{1,82} \approx 0,067. \quad \square$$

Задачі до розділу 14

Задача 1. Знайти вибіркове рівняння прямої лінії регресії Y на X за даними кореляційної таблиці (табл. 12).

Відповідь. $\bar{y}_x = -1,39x + 48,68$.

Задача 2. Знайти вибіркове рівняння параболічної регресії

$$\bar{y}_x = Ax^2 + Bx + C$$

Розділ 14. Елементи теорії регресії і кореляції

за даними кореляційної таблиці (табл. 13).

Оцінити силу кореляційного зв'язку за допомогою вибіркового кореляційного відношення.

Відповідь. $\bar{y}_x = 0,05x^2 - 0,62x + 18,39$; $\eta_{yx} \approx 0,0632$.

Таблиця 12

X	Y					n_x
	27	31	35	39	43	
5	—	—	—	3	5	8
7	—	—	2	8	4	14
9	—	3	18	7	—	28
11	1	6	11	3	—	21
13	10	13	6	—	—	29
n_y	11	22	37	21	9	$n = 100$

Таблиця 13

X	Y					n_x
	12	15	17	18	20	
3	—	—	15	—	—	15
4	—	12	3	15	—	30
6	9	3	—	5	11	28
8	7	1	—	2	9	19
9	3	—	—	—	5	8
n_y	19	16	18	22	25	$n = 100$

Розділ 15. Статистична перевірка статистичних гіпотез

Статистичною називають гіпотезу про вигляд невідомого розподілу або про параметри відомих розподілів.

Нульовою (основною) гіпотезою H_0 називають висунуту гіпотезу.

Альтернативною (конкуруючою) гіпотезою H_1 називають протилежну до нульової гіпотезу.

Проста гіпотеза складається з одного припущення.

Складна гіпотеза складається зі скінченної або нескінченної кількості припущень.

Похибка першого роду полягає у відхиленні правильної нульової гіпотези внаслідок її перевірки.

Рівень значущості α — це ймовірність похибки першого роду.

Похибка другого роду полягає у прийнятті неправильної нульової гіпотези внаслідок її перевірки. Ймовірність похибки другого роду позначають β .

Статистичним критерієм (або просто *критерієм*) називають випадкову величину K , яка використовується для перевірки нульової гіпотези.

Спостережуваним (емпіричним) значенням $K_{\text{спост}}$ є значення критерію, обчислене за вибірками.

Критична область — це сукупність значень критерію, при яких відхиляють нульову гіпотезу.

Областю прийняття гіпотези (областю допустимих значень) є сукупність значень критерію, при яких приймають нульову гіпотезу.

Основний принцип перевірки статистичних гіпотез такий: якщо спостережуване значення критерію належить критичній області, нульову гіпотезу відхиляють; якщо спостережуване значення критерію належить області прийняття гіпотези, гіпотезу приймають.

Критичними точками (межами) $k_{\text{кр}}$ називають такі, що відділяють критичну область від області прийняття гіпотези.

Правостороння критична область визначається нерівністю

$$K > k_{\text{кр}}.$$

Лівостороння критична область визначається нерівністю

$$K < k_{\text{кр}}.$$

Двостороння критична область визначається нерівностями

$$K < k_1, \quad K > k_2,$$

де $k_1 < k_2$. Зокрема, якщо критичні точки симетричні відносно нуля, то двостороння область визначається нерівностями

$$K < -k_{\text{кр}}, \quad K > k_{\text{кр}} \quad (\text{вважається, що } k_{\text{кр}} > 0)$$

або рівносильною нерівністю

$$|K| > k_{\text{кр}}.$$

Для отримання критичної області задають рівень значущості α і шукають критичні точки із таких співвідношень:

а) для правосторонньої критичної області

$$\mathbf{P}(K > k_{\text{кр}}) = \alpha \quad (k_{\text{кр}} > 0);$$

б) для лівосторонньої критичної області

$$\mathbf{P}(K < k_{\text{кр}}) = \alpha \quad (k_{\text{кр}} < 0);$$

в) для двосторонньої симетричної критичної області

$$\mathbf{P}(K < -k_{\text{кр}}) = \frac{\alpha}{2}, \quad \mathbf{P}(K > k_{\text{кр}}) = \frac{\alpha}{2} \quad (k_{\text{кр}} > 0).$$

Потужність критерію називають ймовірність потрапляння критерію в критичну область за умови, що правильною є конкуруюча гіпотеза. Іншими словами, потужність критерію — це ймовірність того, що нульова гіпотеза буде відхилена, якщо справджується конкуруюча гіпотеза.

15.1. Перевірка рівності вибіркового середнього гіпотетичному генеральному середньому

Нехай з нормальної генеральної сукупності з відомою дисперсією σ^2 отримано вибірку обсягу n .

Правило 1. Для того щоб при заданому рівні значущості α перевірити нульову гіпотезу $H_0: a = a_0$ про рівність генерального середнього a нормальної сукупності з відомою дисперсією σ^2 гіпотетичному (пропонованому) значенню a_0 при конкуруючій гіпотезі $H_1: a \neq a_0$, потрібно обчислити спостережуване значення критерію

$$U_{\text{спост}} = \frac{(\bar{x}_в - a_0)\sqrt{n}}{\sigma}$$

і за таблицею функції Лапласа (дод. 1) знайти критичну точку $u_{\text{кр}}$ двосторонньої критичної області з рівності

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2}.$$

Якщо $|U_{\text{спост}}| < u_{\text{кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $|U_{\text{спост}}| > u_{\text{кр}}$, нульову гіпотезу відхиляють.

Правило 2. При конкуруючій гіпотезі $H_1: a > a_0$ критичну точку правосторонньої критичної області знаходять з рівності

$$\Phi(u_{\text{кр}}) = \frac{1}{2} - \alpha.$$

Якщо $U_{\text{спост}} < u_{\text{кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $U_{\text{спост}} > u_{\text{кр}}$, нульову гіпотезу відхиляють.

Правило 3. При конкуруючій гіпотезі $H_1: a < a_0$ спочатку знаходять «допоміжну» критичну точку $u_{\text{кр}}$ за правилом 2, а потім вважають межею лівосторонньої критичної області

$$u'_{\text{кр}} = -u_{\text{кр}}.$$

Якщо $U_{\text{спост}} > -u_{\text{кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $U_{\text{спост}} < -u_{\text{кр}}$, нульову гіпотезу відхиляють.

Потужність критерію перевірки основної гіпотези $H_0: a = a_0$ про рівність генерального середнього a гіпотетичному значенню a_0 при відомому середньому квадратичному відхиленні σ знаходять залежно від вигляду альтернативної гіпотези.

При альтернативній гіпотезі $H_1: a > a_0$ для гіпотетичного значення генерального середнього $a = a_1 > a_0$ потужність правостороннього критерію

$$1 - \beta = \frac{1}{2} - \Phi(u_{\text{кр}} - \lambda),$$

де $u_{\text{кр}}$ знаходять з рівності $\Phi(u_{\text{кр}}) = \frac{1}{2} - \alpha$, $\lambda = \frac{(a_1 - a_0)\sqrt{n}}{\sigma}$. При різних значеннях a_1 функція потужності одностороннього критерію

$$\pi_1(a_1) = \frac{1}{2} - \Phi(u_{\text{кр}} - \lambda).$$

При альтернативній гіпотезі $H_1: a \neq a_0$ для гіпотетичного значення генерального середнього $a = a_1$ потужність двостороннього критерію

$$1 - \beta = 1 - [\Phi(u_{\text{кр}} - \lambda) + \Phi(u_{\text{кр}} + \lambda)],$$

де $u_{\text{кр}}$ знаходять з рівності $\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2}$, $\lambda = \frac{(a_1 - a_0)\sqrt{n}}{\sigma}$. При різних значеннях a_1 функція потужності одностороннього критерію

$$\pi_1(a_1) = 1 - [\Phi(u_{\text{кр}} - \lambda) + \Phi(u_{\text{кр}} + \lambda)].$$

Приклад 1. Із нормальної генеральної сукупності з відомим середнім квадратичним відхиленням $\sigma = 4$ отримано вибірку обсягу $n = 100$, за якою знайдено вибіркове середнє $\bar{x}_b = 29$. Потрібно при рівні значущості 0,01 перевірити нульову гіпотезу $H_0: a = a_0 = 30$ при конкуруючій гіпотезі $H_1: a \neq 30$.

Розв'язання. Знайдемо спочатку спостережуване значення критерію

$$U_{\text{спост}} = \frac{(\bar{x}_b - a_0)\sqrt{n}}{\sigma} = \frac{(29 - 30)\sqrt{100}}{4} = -2,5.$$

Оскільки за умовою конкуруюча гіпотеза має вигляд $H_1: a \neq a_0$, критична область — двостороння.

Обчислимо критичну точку з рівності

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2} = \frac{1 - 0,01}{2} = 0,495.$$

За таблицею функції Лапласа (дод. 1) знаходимо критичну точку:

$$u_{\text{кр}} \approx 2,58.$$

Оскільки $|U_{\text{спост}}| < u_{\text{кр}}$, підстав відхиляти основну гіпотезу немає. Іншими словами, вибіркоче та гіпотетичне генеральне середні різняться несуттєво. \square

Приклад 2. Із нормальної генеральної сукупності з відомим середнім квадратичним відхиленням $\sigma = 4,8$ отримано вибірку обсягу $n = 144$, за якою знайдено вибіркоче середнє $\bar{x}_в = 16$. Потрібно при рівні значущості $0,05$ перевірити нульову гіпотезу $H_0: a = a_0 = 15$ при конкуруючій гіпотезі:

- а) $H_1: a \neq 15$;
- б) $H_1: a > 15$;
- в) $H_1: a < 15$.

Крім того, необхідно знайти потужності правостороннього та двостороннього критеріїв.

Розв'язання. Обчислимо спочатку спостережуване значення критерію

$$U_{\text{спост}} = \frac{(\bar{x}_в - a_0)\sqrt{n}}{\sigma} = \frac{(16 - 15)\sqrt{144}}{4,8} = 2,5.$$

а) Скористаємося правилом 1. Знайдемо критичну точку з рівності

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2} = \frac{1 - 0,05}{2} = 0,475.$$

За таблицею функції Лапласа (дод. 1) визначимо критичну точку:

$$u_{\text{кр}} \approx 1,96.$$

Оскільки $|U_{\text{спост}}| > u_{\text{кр}}$, основна гіпотеза відхиляється. Іншими словами, вибіркоче та гіпотетичне генеральне середні різняться суттєво.

б) Скористаємося правилом 2. Знайдемо критичну точку з рівності

$$\Phi(u_{\text{кр}}) = \frac{1}{2} - \alpha = \frac{1}{2} - 0,05 = 0,45.$$

За таблицею функції Лапласа (дод. 1) визначимо критичну точку:

$$u_{\text{кр}} \approx 1,64.$$

Оскільки $U_{\text{спост}} > u_{\text{кр}}$, основна гіпотеза відхиляється. Іншими словами, вибіркоче та гіпотетичне генеральне середні різняться суттєво.

в) Скористаємося правилом 3. Критична точка буде такою самою, як і в пункті б), але з протилежним знаком:

$$u_{\text{кр}} \approx -1,64.$$

Оскільки $U_{\text{спост}} > u_{\text{кр}}$, підстав відхилити основну гіпотезу немає. Іншими словами, вибіркоче та гіпотетичне генеральне середні різняться несуттєво.

Тепер знайдемо потужності правостороннього та двостороннього критеріїв. Нагадаємо, що критичні точки у цих випадках різні та дорівнюють відповідно 1,64 і 1,96.

Знайдемо параметр λ , який входить в обидва рівняння для визначення потужності критеріїв:

$$\lambda = \frac{(a_1 - a_0)\sqrt{n}}{\sigma} = \frac{(16 - 15)\sqrt{144}}{4,8} = 2,5.$$

Отже, потужності відповідно правостороннього та двостороннього критеріїв будуть такі:

$$\begin{aligned}\pi_1^{\text{правост}}(16) &= \frac{1}{2} - \Phi(u_{\text{кр}} - \lambda) = \frac{1}{2} - \Phi(1,64 - 2,5) = \\ &= \frac{1}{2} + \Phi(0,86) = \frac{1}{2} + 0,3051 = 0,8051; \\ \pi_1^{\text{двост}}(16) &= 1 - [\Phi(u_{\text{кр}} - \lambda) + \Phi(u_{\text{кр}} + \lambda)] = \\ &= 1 - [\Phi(1,96 - 2,5) + \Phi(1,96 + 2,5)] = \\ &= 1 - [-\Phi(0,54) + \Phi(4,46)] \approx \\ &\approx 1 + 0,2054 - 0,5 = 0,7054.\end{aligned}$$

Іншими словами, імовірності того, що нульова гіпотеза буде відхилена, якщо правильною є конкуруюча гіпотеза, дорівнюють 0,8051 і 0,7054 відповідно для правостороннього та двостороннього критеріїв. \square

Якщо дисперсія генеральної сукупності невідома, як критерій перевірки основної гіпотези беруть випадкову величину

$$T = \frac{(\bar{x} - a_0)\sqrt{n}}{s},$$

де $s = \sqrt{\frac{n \sum_i n_i x_i^2 - \left(\sum_i n_i x_i\right)^2}{n(n-1)}}$ — «виправлене» вибіркове середнє квадратичне відхилення. Величина T має розподіл Стьюдента з $k = n - 1$ ступенями вільності.

Правило 1. Для того щоб при заданому рівні значущості α перевірити нульову гіпотезу $H_0: a = a_0$ про рівність невідомого генерального середнього a нормальної сукупності з невідомою дисперсією гіпотетичному (пропонованому) значенню a_0 при конкуруючій гіпотезі $H_1: a \neq a_0$, потрібно обчислити спостережуване значення критерію

$$T_{\text{спост}} = \frac{(\bar{x}_в - a_0)\sqrt{n}}{s}$$

і за таблицею критичних точок розподілу Стьюдента (дод. 6) при заданому рівні значущості α , розміщеному у верхній частині таблиці, і кількості ступенів вільності $k = n - 1$ знайти критичну точку $t_{\text{двост. кр}}(\alpha; k)$.

Якщо $|T_{\text{спост}}| < t_{\text{двост. кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $|T_{\text{спост}}| > t_{\text{двост. кр}}$, нульову гіпотезу відхиляють.

Правило 2. При конкуруючій гіпотезі $H_1: a > a_0$ за рівнем значущості α , розміщеним у нижній частині таблиці дод. 6, і при кількості ступенів вільності $k = n - 1$ знаходять критичну точку правосторонньої критичної області $t_{\text{правост. кр}}(\alpha; k)$.

Якщо $T_{\text{спост}} < t_{\text{правост. кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $T_{\text{спост}} > t_{\text{правост. кр}}$, нульову гіпотезу відхиляють.

Правило 3. При конкуруючій гіпотезі $H_1: a < a_0$ спочатку знаходять так звану «допоміжну» критичну точку (за правилом 2) $t_{\text{правост. кр}}(\alpha; k)$ і вважають межею лівосторонньої критичної області $t_{\text{лівост. кр}} = -t_{\text{правост. кр}}$.

Якщо $T_{\text{спост}} > -t_{\text{правост. кр}}$, немає підстав відхиляти нульову гіпотезу. Якщо $T_{\text{спост}} < -t_{\text{правост. кр}}$, нульову гіпотезу відхиляють.

Приклад 3. Для вибірки обсягу $n = 25$, отриманої з нормальної генеральної сукупності, знайдено вибіркове середнє $\bar{x} = 43$ і «виправлене» вибіркове середнє квадратичне відхилення $s = 4$. Потрібно при рівні значущості $\alpha = 0,01$ перевірити основну гіпотезу $H_0: a = a_0 = 45$ при альтернативній гіпотезі $H_1: a \neq 45$.

Розв'язання. Обчислимо спочатку спостережуване значення критерію

$$T_{\text{спост}} = \frac{(\bar{x} - a_0)\sqrt{n}}{s} = \frac{(43 - 45)\sqrt{25}}{4} = 2,5.$$

Оскільки альтернативна гіпотеза має вигляд $H_1: a \neq a_0$, то критична область — двостороння.

За таблицею критичних точок розподілу Стьюдента (дод. 6) при рівні значущості $\alpha = 0,01$ і кількості ступенів вільності

$$k = n - 1 = 25 - 1 = 24$$

знаходимо критичну точку

$$t_{\text{двост. кр}}(0,01; 24) = 2,80.$$

Оскільки $|T_{\text{спост}}| < t_{\text{двост. кр}}$, підстав відхиляти основну гіпотезу немає. Іншими словами, вибіркове та гіпотетичне генеральне середні різняться несуттєво. \square

Приклад 4. Контрольна маса буханки хліба, що випікається на заводі, має становити $a = a_0 = 1250$ г. Контрольні вимірювання 20 випадково відібраних буханок хліба виявили такі результати:

Маса, г	1240	1245	1250	1255	1260
Кількість буханок	3	5	6	4	2

Потрібно при рівні значущості 0,001 перевірити основну гіпотезу $H_0: a = a_0 = 1250$ при конкуруючій гіпотезі:

- а) $H_1: a \neq a_0$;
- б) $H_1: a > a_0$;
- в) $H_1: a < a_0$.

Розв'язання. Обчислимо спочатку вибіркове середнє \bar{x} і «виправлене» середнє квадратичне відхилення s :

$$\bar{x} = \frac{1240 \cdot 3 + 1245 \cdot 5 + 1250 \cdot 6 + 1255 \cdot 4 + 1260 \cdot 2}{20} = 1249,25;$$
$$s = \left(\frac{1240^2 \cdot 3 + 1245^2 \cdot 5 + 1250^2 \cdot 6 + 1255^2 \cdot 4 + 1260^2 \cdot 2}{19} - \frac{20}{19} \cdot 1249,25^2 \right)^{\frac{1}{2}} \approx 6,13.$$

Знайдемо спостережуване значення критерію:

$$T_{\text{спост}} = \frac{(\bar{x} - a_0)\sqrt{n}}{s} = \frac{(1249,25 - 1250)\sqrt{20}}{6,13} \approx -0,547.$$

а) Скористаємося правилом 1. За дод. 6 при рівні значущості $\alpha = 0,001$ і кількості ступенів вільності

$$k = n - 1 = 20 - 1 = 19$$

знайдемо критичну точку розподілу Стьюдента:

$$t_{\text{двост. кр}}(0,001; 19) = 3,88.$$

Оскільки $|T_{\text{спост}}| < t_{\text{двост. кр}}$, підстав відхилити основну гіпотезу немає. Іншими словами, вибіркове та гіпотетичне генеральне середні різняться несуттєво.

б) Скористаємося правилом 2. За дод. 6 при рівні значущості $\alpha = 0,001$ і кількості ступенів вільності

$$k = n - 1 = 20 - 1 = 19$$

знайдемо критичну точку розподілу Стьюдента:

$$t_{\text{правост. кр}}(0,001; 19) = 3,58.$$

Оскільки $T_{\text{спост}} < t_{\text{правост. кр}}$, підстав відхиляти основну гіпотезу немає. Іншими словами, вибіркове та гіпотетичне генеральне середні різняться несуттєво.

в) Скористаємося правилом 3. Критична точка буде такою самою, як і в пункті б), але з протилежним знаком:

$$t_{\text{лівост. кр}}(0,001; 19) = -3,88.$$

Оскільки $T_{\text{спост}} > t_{\text{лівост. кр}}$, підстав відхиляти основну гіпотезу немає. Іншими словами, вибіркове та гіпотетичне генеральне середні різняться несуттєво. \square

15.2. Перевірка рівності «виправленої» вибіркової дисперсії генеральній дисперсії

Нехай з нормальної генеральної сукупності отримано вибірку обсягу n , для якої знайдено «виправлену» вибіркору дисперсію s^2 .

Правило 1. Для того щоб при заданому рівні значущості α перевірити нульову гіпотезу $H_0: \sigma^2 = \sigma_0^2$ про рівність невідомої генеральної дисперсії σ^2 гіпотетичному (пропонованому) значенню σ_0^2 при конкуруючій гіпотезі $H_1: \sigma^2 > \sigma_0^2$, потрібно обчислити спостережуване значення критерію

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

і за таблицею критичних точок розподілу χ^2 (дод. 7), при заданому рівні значущості α і кількості ступенів вільності $k = n - 1$ знайти критичну точку $\chi_{\text{кр}}^2(\alpha; k)$.

Якщо $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2$, немає підстав відхиляти нульову гіпотезу. Якщо $\chi_{\text{спост}}^2 > \chi_{\text{кр}}^2$, нульову гіпотезу відхиляють.

Правило 2. При конкуруючій гіпотезі $H_1: \sigma^2 \neq \sigma_0^2$ знаходять ліву $\chi_{\text{лів. кр}}^2(1 - \alpha/2; k)$ і праву $\chi_{\text{прав. кр}}^2(\alpha/2; k)$ критичні точки.

Якщо $\chi_{\text{лів.кр}}^2 < \chi_{\text{спост}}^2 < \chi_{\text{прав.кр}}^2$, немає підстав відхиляти нульову гіпотезу. Якщо $\chi_{\text{спост}}^2 < \chi_{\text{лів.кр}}^2$ або $\chi_{\text{спост}}^2 > \chi_{\text{прав.кр}}^2$, нульову гіпотезу відхиляють.

Правило 3. При конкуруючій гіпотезі $H_1: \sigma^2 < \sigma_0^2$ знаходять критичну точку $\chi_{\text{кр}}^2(1 - \alpha; k)$.

Якщо $\chi_{\text{спост}}^2 > \chi_{\text{кр}}^2(1 - \alpha; k)$, немає підстав відхиляти нульову гіпотезу. Якщо $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2(1 - \alpha; k)$, нульову гіпотезу відхиляють.

Зауваження 1. Якщо кількість ступенів вільності $k > 30$, критичну точку $\chi_{\text{кр}}^2(\alpha; k)$ можна знайти з рівності Вілсона — Гільферті:

$$\chi_{\text{кр}}^2(\alpha; k) = k \left(1 - \frac{2}{9k} + z_\alpha \sqrt{\frac{2}{9k}} \right)^3,$$

де z_α визначають, використовуючи таблицю інтегральної функції Лапласа (дод. 1), з рівності

$$\Phi(z_\alpha) = \frac{1}{2} - \alpha.$$

Приклад 1. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 25$, для якої знайдено вибіркову «виправлену» дисперсію $s^2 = 12,3$. Перевірити при рівні значущості $\alpha = 0,01$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 12$, якщо альтернативна гіпотеза $H_1: \sigma_0^2 > 12$.

Розв'язання. Знайдемо спостережуване значення критерію:

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(25-1) \cdot 12,3}{12} = 24,6.$$

Оскільки альтернативна гіпотеза має вигляд $H_1: \sigma_0^2 > 12$, критична область правостороння (правило 1). За дод. 7 при рівні значущості $\alpha = 0,01$ і кількості ступенів вільності

$$k = n - 1 = 25 - 1 = 24$$

знаходимо критичну точку:

$$\chi_{\text{кр}}^2(0,01; 24) = 42,98.$$

Оскільки $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2$, немає підстав відхилити основну гіпотезу про рівність генеральної дисперсії гіпотетичному значенню $\sigma_0^2 \leq 12$. Іншими словами, різниця між «виправленою» вибірковою дисперсією $s^2 = 12,3$ і гіпотетичною генеральною дисперсією $\sigma_0^2 \leq 12$ незначуща. \square

Приклад 2. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 20$, для якої знайдено вибірку «виправлену» дисперсію $s^2 = 2,7$. Перевірити при рівні значущості $\alpha = 0,05$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 3$, якщо альтернативна гіпотеза:

- а) $H_1: \sigma_0^2 > 3$;
- б) $H_1: \sigma_0^2 \neq 3$;
- в) $H_1: \sigma_0^2 < 3$.

Розв'язання. Знайдемо спостережуване значення критерію:

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1) \cdot 2,7}{3} = 17,1.$$

а) Skorистаємося правилом 1. За дод. 7 при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = n - 1 = 20 - 1 = 19$$

знаходимо критичну точку

$$\chi_{\text{кр}}^2(0,05; 19) = 30,14.$$

Оскільки $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2$, немає підстав відхилити основну гіпотезу про рівність генеральної дисперсії гіпотетичному значенню $\sigma_0^2 \leq 3$. Іншими словами, різниця між «виправленою» вибірковою дисперсією $s^2 = 2,7$ і гіпотетичною генеральною дисперсією $\sigma_0^2 \leq 3$ незначуща.

б) Skorистаємося правилом 2. За дод. 7 при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = n - 1 = 20 - 1 = 19$$

знаходимо критичні точки:

$$\begin{aligned}\chi_{\text{лів. кр}}^2(1 - \alpha/2; k) &= \chi_{\text{лів. кр}}^2(0,975; 19) = 8,91; \\ \chi_{\text{прав. кр}}^2(\alpha/2; k) &= \chi_{\text{прав. кр}}^2(0,025; 19) = 32,85.\end{aligned}$$

Оскільки $\chi_{\text{лів. кр}}^2 < \chi_{\text{спост}}^2 < \chi_{\text{прав. кр}}^2$, немає підстав відхилити основну гіпотезу про рівність генеральної дисперсії гіпотетичному значенню $\sigma_0^2 = 3$. Іншими словами, різниця між «виправленою» вибірковою дисперсією $s^2 = 2,7$ і гіпотетичною генеральною дисперсією $\sigma_0^2 = 3$ незначуща.

в) Скористаємося правилом 3. За дод. 7 при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = n - 1 = 20 - 1 = 19$$

знаходимо критичну точку:

$$\chi_{\text{кр}}^2(1 - \alpha; k) = \chi_{\text{кр}}^2(0,95; 19) = 10,12.$$

Оскільки $\chi_{\text{спост}}^2 > \chi_{\text{кр}}^2(1 - \alpha; k)$, немає підстав відхилити основну гіпотезу про рівність генеральної дисперсії гіпотетичному значенню $\sigma_0^2 \geq 3$. Іншими словами, різниця між «виправленою» вибірковою дисперсією $s^2 = 2,7$ і гіпотетичною генеральною дисперсією $\sigma_0^2 \geq 3$ незначуща. \square

Приклад 3. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 100$:

x_i	5	6	7	8	9	10	11
n_i	9	15	19	20	17	13	7

Потрібно перевірити при рівні значущості $\alpha = 0,01$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 4$, якщо альтернативна гіпотеза $H_1: \sigma_0^2 > 4$.

Розв'язання. Спочатку за даними вибірки знайдемо «виправлену»

вибіркову дисперсію s^2 :

$$\bar{x} = \frac{5 \cdot 9 + 6 \cdot 15 + 7 \cdot 19 + 8 \cdot 20 + 9 \cdot 17 + 10 \cdot 13 + 11 \cdot 7}{100} = 7,88;$$
$$s^2 = \frac{5^2 \cdot 9 + 6^2 \cdot 15 + 7^2 \cdot 19 + 8^2 \cdot 20 + 9^2 \cdot 17 + 10^2 \cdot 13^2 + 11^2 \cdot 7}{100} - 7,88^2 = 2,9056.$$

Обчислимо спостережуване значення критерію:

$$\chi_{\text{спост}}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(100-1) \cdot 2,9056}{4} = 71,9136.$$

Оскільки альтернативна гіпотеза має вигляд $H_1: \sigma_0^2 > 12$, критична область правостороння (правило 1).

Знайдемо кількість ступенів вільності k :

$$k = n - 1 = 100 - 1 = 99.$$

Оскільки кількість ступенів вільності $k > 30$, критичну точку $\chi_{\text{кр}}^2(\alpha; k)$ знайдемо з рівності Вілсона — Гільферті:

$$\chi_{\text{кр}}^2(\alpha; k) = k \left(1 - \frac{2}{9k} + z_\alpha \sqrt{\frac{2}{9k}} \right)^3,$$

де z_α визначимо з рівності

$$\Phi(z_\alpha) = \frac{1}{2} - \alpha = \frac{1}{2} - 0,01 = 0,49,$$

використовуючи таблицю функції Лапласа (дод. 1):

$$z_\alpha \approx 2,33.$$

Отже,

$$\chi_{\text{кр}}^2(0,01; 99) = 99 \left(1 - \frac{2}{9 \cdot 99} + 2,33 \sqrt{\frac{2}{9 \cdot 99}} \right)^3 = 113,0187.$$

Оскільки $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2$, немає підстав відхилити основну гіпотезу про рівність генеральної дисперсії гіпотетичному значенню $\sigma_0^2 \leq 4$. Іншими словами, різниця між «виправленою» вибірковою дисперсією $s^2 \approx 2,9$ і гіпотетичною генеральною дисперсією $\sigma_0^2 \leq 4$ незначуща. \square

15.3. Перевірка рівності відносної частоти гіпотетичній імовірності

Нехай при достатньо великій кількості n незалежних випробувань, у кожному з яких імовірність p появи події стала, але невідома, знайдено відносну частоту $\frac{m}{n}$. Потрібно при заданому рівні значущості α перевірити основну гіпотезу, яка полягає в тому, що невідома ймовірність p дорівнює гіпотетичній імовірності p_0 .

Правило 1. Для того щоб при заданому рівні значущості α перевірити основну гіпотезу $H_0: p = p_0$ про рівність невідомої ймовірності p гіпотетичній імовірності p_0 при конкуруючій гіпотезі $H_1: p \neq p_0$, потрібно обчислити спостережуване значення критерію

$$U_{\text{спост}} = \frac{\left(\frac{m}{n} - p_0\right) \sqrt{n}}{\sqrt{p_0 q_0}}$$

і за таблицею функції Лапласа (дод. 1) знайти критичну точку $u_{\text{кр}}$ з рівності

$$\Phi(u_{\text{кр}}) = \frac{1 - \alpha}{2}.$$

Якщо $|U_{\text{спост}}| < u_{\text{кр}}$, немає підстав відхилити основну гіпотезу. Якщо $|U_{\text{спост}}| > u_{\text{кр}}$, основну гіпотезу відхиляють.

Правило 2. При альтернативній гіпотезі $H_1: p > p_0$ критичну точку правосторонньої критичної області знаходять з рівності

$$\Phi(u_{\text{кр}}) = \frac{1}{2} - \alpha.$$

Якщо $U_{\text{спост}} < u_{\text{кр}}$, немає підстав відхилити основну гіпотезу. Якщо $U_{\text{спост}} > u_{\text{кр}}$, основну гіпотезу відхиляють.

Правило 3. При альтернативній гіпотезі $H_1: p < p_0$ спочатку знаходять «допоміжну» критичну точку $u_{кр}$ за правилом 2, а потім вважають межею лівосторонньої критичної області

$$u'_{кр} = -u_{кр}.$$

Якщо $U_{спост} > -u_{кр}$, немає підстав відхилити основну гіпотезу. Якщо $U_{спост} < -u_{кр}$, основну гіпотезу відхиляють.

Зауваження 1. Задовільні результати отримуються в разі виконання нерівності $np_0q_0 > 9$.

Приклад 1. Для деякої події в результаті 100 незалежних випробувань було знайдено відносну частоту $\frac{m}{n} = 0,34$. При рівні значущості 0,05 потрібно перевірити основну гіпотезу $H_0: p = p_0 = 0,3$ при конкуруючій гіпотезі $H_1: p \neq 0,3$.

Розв'язання. Спочатку знайдемо q_0 :

$$q_0 = 1 - p_0 = 1 - 0,3 = 0,7,$$

тобто

$$np_0q_0 = 100 \cdot 0,3 \cdot 0,7 = 21 > 9.$$

Обчислимо спостережуване значення критерію:

$$U_{спост} = \frac{\left(\frac{m}{n} - p_0\right) \sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,34 - 0,3) \sqrt{100}}{\sqrt{0,3 \cdot 0,7}} \approx 0,87.$$

Оскільки за умовою альтернативна гіпотеза має вигляд $H_1: p \neq p_0$, критична область двостороння (правило 1).

Визначимо критичну точку $u_{кр}$ з рівності

$$\Phi(u_{кр}) = \frac{1 - \alpha}{2} = \frac{1 - 0,05}{2} = 0,475.$$

За дод. 1 знаходимо

$$u_{кр} \approx 1,96.$$

Оскільки $|U_{спост}| < u_{кр}$, немає підстав відхилити основну гіпотезу. Іншими словами, спостережувана відносна частота $\frac{m}{n} = 0,34$ незначуще відрізняється від гіпотетичної ймовірності $p_0 = 0,3$. \square

Приклад 2. Для деякої події в результаті 400 незалежних випробувань було знайдено відносну частоту $\frac{m}{n} = 0,44$. При рівні значущості 0,01 потрібно перевірити основну гіпотезу $H_0: p = p_0 = 0,5$ при конкуруючій гіпотезі:

- а) $H_1: p > 0,5$;
- б) $H_1: p < 0,5$.

Розв'язання. Спочатку знайдемо q_0 :

$$q_0 = 1 - p_0 = 1 - 0,5 = 0,5,$$

тобто

$$np_0q_0 = 400 \cdot 0,5 \cdot 0,5 = 100 > 9.$$

Обчислимо спостережуване значення критерію:

$$U_{\text{спост}} = \frac{\left(\frac{m}{n} - p_0\right) \sqrt{n}}{\sqrt{p_0q_0}} = \frac{(0,44 - 0,5) \sqrt{400}}{\sqrt{0,5 \cdot 0,5}} \approx -2,4.$$

а) Застосуємо правило 2. Визначимо критичну точку $u_{\text{кр}}$ з рівності

$$\Phi(u_{\text{кр}}) = \frac{1}{2} - \alpha = \frac{1}{2} - 0,01 = 0,49.$$

За дод. 1 знаходимо

$$u_{\text{кр}} \approx 2,33.$$

Оскільки $U_{\text{спост}} < u_{\text{кр}}$, немає підстав відхилити основну гіпотезу. Іншими словами, спостережувана відносна частота $\frac{m}{n} = 0,47$ незначуще відрізняється від гіпотетичної ймовірності $p_0 \leq 0,5$.

б) Застосуємо правило 3. Використовуючи дані пункту а), знаходимо критичну точку

$$u_{\text{кр}} \approx -2,33.$$

Оскільки $U_{\text{спост}} < u_{\text{кр}}$, основна гіпотеза відхиляється. Іншими словами, спостережувана відносна частота $\frac{m}{n} = 0,47$ значуще відрізняється від гіпотетичної ймовірності $p_0 \geq 0,5$. □

15.4. Перевірка гіпотези про нормальний розподіл за критерієм Пірсона

Нехай статистичний розподіл вибірки задано у вигляді послідовності рівновіддалених варіант і відповідних їм частот:

x_i	x_1	x_2	\cdots	x_N
n_i	n_1	n_2	\cdots	n_N

Необхідно за допомогою *критерію Пірсона* перевірити гіпотезу про нормальний розподіл генеральної сукупності.

Правило 1. Для того щоб при заданому рівні значущості α перевірити гіпотезу про нормальний розподіл генеральної сукупності, необхідно:

- 1) обчислити вибіркове середнє \bar{x}_B і вибіркове середнє квадратичне відхилення σ_B ;
- 2) визначити теоретичні частоти

$$n'_i = \frac{nh}{\sigma_B} \cdot \varphi(u_i),$$

де n — обсяг вибірки; h — крок (різниця між двома сусідніми варіантами); $\varphi(u)$ — диференціальна функція Лапласа (дод. 2);

$$u_i = \frac{x_i - \bar{x}_B}{\sigma_B};$$

- 3) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона. Для цього:
 - а) складають розрахункову таблицю (див. табл. 15), за якою знаходять спостережуване значення критерію Пірсона

$$\chi^2_{\text{спост}} = \sum_i \frac{(n_i - n'_i)^2}{n'_i};$$

б) за таблицею критичних точок розподілу χ^2 (дод. 7) при заданому рівні значущості α і кількості ступенів вільності $k = s - 3$ (s — кількість варіант вибірки) знаходять критичну точку $\chi_{кр}^2(\alpha; k)$ правосторонньої критичної області. Якщо $\chi_{спост}^2 < \chi_{кр}^2$, немає підстав відхиляти гіпотезу про нормальний розподіл генеральної сукупності. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво (випадково). Якщо $\chi_{спост}^2 > \chi_{кр}^2$, гіпотезу відхиляють. Іншими словами, емпіричні та теоретичні частоти різняться суттєво.

Зауваження 1. Малочисельні частоти вибірки ($n_i < 5$) слід об'єднувати; у цьому разі відповідні їм теоретичні частоти також потрібно додати. Якщо відбувалося об'єднання частот, то при визначенні кількості ступенів вільності за формулою $k = s - 3$ слід як s узяти кількість варіант вибірки, що залишилися після об'єднання частот.

Приклад 1. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про нормальний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 100$:

x_i	-5	-3	-1	1	3	5	7	9
n_i	5	9	13	18	21	18	10	6

Розв'язання. Спочатку обчислимо вибіркове середнє $\bar{x}_в$ і середнє квадратичне відхилення вибірки $\sigma_в$:

$$\begin{aligned} \bar{x}_в &= \frac{-5 \cdot 5 - 3 \cdot 9 - 1 \cdot 13 + 1 \cdot 18 + 3 \cdot 21 + 5 \cdot 18 + 7 \cdot 10 + 9 \cdot 6}{100} = \\ &= 2,3; \\ \sigma_в &= \left([5^2 \cdot 5 + 3^2 \cdot 9 + 1 \cdot 13 + 1 \cdot 18 + 3^2 \cdot 21 + 5^2 \cdot 18 + 7^2 \cdot 10 + \right. \\ &\quad \left. + 9^2 \cdot 6] / 100 - 2,3^2 \right)^{\frac{1}{2}} \approx 3,64. \end{aligned}$$

Визначимо теоретичні частоти, ураховуючи, що $n = 100$, $h = 2$,

Розділ 15. Статистична перевірка статистичних гіпотез

$\sigma_B = 3,64$, за формулою

$$n'_i = \frac{nh}{\sigma_B} \cdot \varphi\left(\frac{x_i - \bar{x}_B}{\sigma_B}\right).$$

Для цього складемо розрахункову табл. 14 (значення диференціальної функції Лапласа $\varphi(u)$ наведено в дод. 2).

Таблиця 14

i	x_i	$u_i = \frac{x_i - \bar{x}_B}{\sigma_B}$	$\varphi(u_i)$	$n'_i = \frac{nh}{\sigma_B} \cdot \varphi(u_i)$
1	-5	-2,01	0,0529	2,91
2	-3	-1,46	0,1374	7,56
3	-1	-0,91	0,2637	14,50
4	1	-0,36	0,3739	20,56
5	3	0,19	0,3918	21,54
6	5	0,74	0,3034	16,68
7	7	1,29	0,1736	9,55
8	9	1,84	0,0734	4,04

Порівняємо емпіричні та теоретичні частоти. Для цього обчислимо спостережуване значення критерію

$$\chi_{\text{спост}}^2 = \sum_i \frac{(n_i - n'_i)^2}{n'_i}$$

за допомогою розрахункової табл. 15.

Із табл. 15 знаходимо

$$\chi_{\text{спост}}^2 = 3,33.$$

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 3 = 8 - 3 = 5$$

знаходимо критичну точку правосторонньої критичної області:

$$\chi_{\text{кр}}^2(0,05; 5) = 11,07.$$

Таблиця 15

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	5	2,91	2,09	4,37	1,50
2	9	7,56	1,44	2,07	0,27
3	13	14,50	-1,50	2,25	0,16
4	18	20,56	-2,56	6,55	0,32
5	21	21,54	-0,54	0,29	0,01
6	18	16,68	1,32	1,74	0,10
7	10	9,55	0,45	0,20	0,02
8	6	4,04	1,96	3,84	0,95
Σ	100				$\chi^2_{\text{спост}} = 3,33$

Оскільки $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, гіпотезу про нормальний розподіл не відхиляємо. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво. \square

Нехай статистичний розподіл вибірки задано у вигляді послідовності інтервалів і відповідних їм частот:

$(x_i; x_{i+1}]$	$(x_1; x_2]$	$(x_2; x_3]$	\dots	$(x_N; x_{N+1}]$
n_i	n_1	n_2	\dots	n_N

Необхідно за допомогою *критерію Пірсона* перевірити гіпотезу про нормальний розподіл генеральної сукупності.

Правило 2. Для того щоб при заданому рівні значущості α перевірити гіпотезу про нормальний розподіл генеральної сукупності, необхідно:

- 1) обчислити вибіркове середнє \bar{x}_v і вибіркове середнє квадратичне відхилення σ_v , причому як варіанти x_i^* беруть середнє арифметичне кінців інтервалу:

$$x_i^* = \frac{x_i + x_{i+1}}{2};$$

- 2) пронормувати досліджувану випадкову величину X , тобто перейти до випадкової величини $Z = \frac{X - \bar{x}^*}{\sigma^*}$, і обчислити кінці інтервалів: $z_i = \frac{x_i - \bar{x}^*}{\sigma^*}$, причому найменше значення Z , тобто z_1 , вважають рівним $-\infty$, а найбільше, тобто z_{N+1} , — рівним ∞ ;

- 3) обчислити теоретичні частоти

$$n'_i = n \cdot \mathbf{P}_i,$$

де n — обсяг вибірки; $\mathbf{P}_i = \Phi(z_{i+1}) - \Phi(z_i)$ — імовірності потрапляння X в інтервали $(x_i; x_{i+1}]$; $\Phi(z)$ — функція Лапласа (дод. 1);

- 4) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона. Для цього:

- а) складають розрахункову таблицю (див. табл. 15), за якою знаходять спостережуване значення критерію Пірсона

$$\chi^2_{\text{спост}} = \sum_i \frac{(n_i - n'_i)^2}{n'_i};$$

- б) за таблицею критичних точок розподілу χ^2 (дод. 7) при заданому рівні значущості α і кількості ступенів вільності $k = s - 3$ (s — кількість інтервалів вибірки) знаходять критичну точку $\chi^2_{\text{кр}}(\alpha; k)$ правосторонньої критичної області. Якщо $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, немає підстав відхилити гіпотезу про нормальний розподіл генеральної сукупності. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво (випадково). Якщо $\chi^2_{\text{спост}} > \chi^2_{\text{кр}}$, гіпотезу відхиляють. Іншими словами, емпіричні та теоретичні частоти різняться суттєво.

Зауваження 2. Інтервали, які містять малочисельні частоти вибірки ($n_i < 5$), слід об'єднувати, а частоти цих інтервалів додавати.

Якщо відбувалося об'єднання інтервалів, при визначенні кількості ступенів вільності за формулою $k = s - 3$ слід як s узяти кількість інтервалів вибірки, що залишилися після об'єднання інтервалів.

Приклад 2. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про нормальний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 100$:

$(x_i; x_{i+1}]$	(3; 7]	(7; 11]	(11; 15]	(15; 19]	(19; 23]	(23; 27]	(27; 31]
n_i	6	16	19	17	15	14	13

Розв'язання. Спочатку обчислимо вибіркове середнє $\bar{x}_в$ і середнє квадратичне відхилення вибірки $\sigma_в$. Для цього перетворимо інтервальний статистичний розподіл вибірки на точковий:

x_i	5	9	13	17	21	25	29
n_i	6	16	19	17	15	14	13

Отже,

$$\begin{aligned} \bar{x}_в &= \frac{5 \cdot 6 - 9 \cdot 16 - 13 \cdot 19 + 17 \cdot 17 + 21 \cdot 15 + 25 \cdot 14 + 29 \cdot 13}{100} = \\ &= 17,52; \\ \sigma_в &= \left([5^2 \cdot 6 + 9^2 \cdot 16 + 13^2 \cdot 19 + 17^2 \cdot 17 + 21^2 \cdot 15 + 25^2 \cdot 14 + \right. \\ &\quad \left. + 29^2 \cdot 13] / 100 - 17,52^2 \right)^{\frac{1}{2}} \approx 7,19. \end{aligned}$$

Знайдемо нормовані інтервали $(z_i; z_{i+1}]$, урахувуючи, що вибіркове середнє $\bar{x}_в = 17,52$ і середнє квадратичне відхилення вибірки $\sigma_в = 7,19$. Для цього складемо розрахункову табл. 16 (лівий кінець першого інтервалу покладемо рівним $-\infty$, а правий кінець останнього інтервалу — рівним ∞).

Визначимо теоретичні ймовірності \mathbf{P}_i і теоретичні частоти:

$$n'_i = n \cdot \mathbf{P}_i.$$

Для цього складемо розрахункову табл. 17.

Розділ 15. Статистична перевірка статистичних гіпотез

Таблиця 16

i	x_i	x_{i+1}	$x_i - \bar{x}_B$	$x_{i+1} - \bar{x}_B$	$z_i = \frac{x_i - \bar{x}_B}{\sigma_B}$	$z_{i+1} = \frac{x_{i+1} - \bar{x}_B}{\sigma_B}$
1	3	7	-14,5	-10,5	$-\infty$	-1,46
2	7	11	-10,5	-6,5	-1,46	-0,90
3	11	15	-6,5	-2,5	-0,90	-0,35
4	15	19	-2,5	1,5	-0,35	0,21
5	19	23	1,5	5,5	0,21	0,76
6	23	27	5,5	9,5	0,76	1,32
7	27	31	9,5	13,5	1,32	∞

Таблиця 17

i	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	$\mathbf{P}_i = \Phi(z_{i+1}) - \Phi(z_i)$	$n'_i = n \cdot \mathbf{P}_i$
1	$-\infty$	-1,48	-0,5000	-0,4306	0,0694	6,94
2	-1,48	-0,92	-0,4306	-0,3212	0,1094	10,94
3	-0,92	-0,35	-0,3212	-0,1368	0,1844	18,44
4	-0,35	0,21	-0,1368	0,0832	0,2200	22,00
5	0,21	0,78	0,0832	0,2823	0,1991	19,91
6	0,78	1,34	0,2823	0,4099	0,1276	12,76
7	1,34	∞	0,4099	0,5000	0,0901	9,01
Σ					1	100

Порівняємо емпіричні та теоретичні частоти, використовуючи критерій Пірсона. Для цього обчислимо спочатку спостережуване значення критерію Пірсона за допомогою розрахункової табл. 18. Стівці 7 і 8 цієї таблиці введено для контролю обчислень за формулою

$$\chi^2_{\text{спост}} = \sum_i \frac{n_i^2}{n'_i} - n.$$

Таблиця 18

1	2	3	4	5	6	7	8
i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$	n_i^2	$\frac{n_i^2}{n'_i}$
1	6	6,94	-0,94	0,8836	0,1273	36	5,1873
2	16	10,94	5,06	25,6036	2,3404	256	23,4004
3	19	18,44	0,56	0,3136	0,0170	361	19,5770
4	17	22,00	-5,00	25,0000	1,1364	289	13,1364
5	15	19,91	-4,91	24,1081	1,2109	225	11,3009
6	14	12,76	1,24	1,5376	0,1205	196	15,3605
7	13	9,01	3,99	15,9201	1,7669	169	18,7569
Σ	100	100			$\chi^2_{\text{спост}} = 6,7194$		106,7194

Контроль: Оскільки

$$\sum_i \frac{n_i^2}{n'_i} - n = 106,7194 - 100 = 6,7194 = \chi^2_{\text{спост}},$$

обчислення проведено правильно.

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 3 = 7 - 3 = 4$$

знайдемо критичну точку правосторонньої критичної області:

$$\chi^2_{\text{кр}}(0,05; 4) = 9,49.$$

Оскільки $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, гіпотезу про нормальний розподіл не відхиляємо. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво. \square

15.5. Перевірка гіпотези про рівномірний розподіл

Нехай задано інтервальний статистичний розподіл вибірки:

$(x_i; x_{i+1}]$	$(x_0; x_1]$	$(x_1; x_2]$	\dots	$(x_{N-1}; x_N]$
n_i	n_0	n_1	\dots	n_{N-1}

причому n — обсяг вибірки.

Потрібно за допомогою критерію Пірсона перевірити гіпотезу про рівномірний розподіл генеральної сукупності.

Правило 1. Для того щоб перевірити гіпотезу про рівномірний розподіл досліджуваної випадкової величини X , тобто про щільність X :

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a; b], \\ 0, & x \notin [a; b], \end{cases}$$

необхідно:

- 1) оцінити параметри a і b — кінці інтервалу, у якому спостерігалися можливі значення X , за формулами (a^* і b^* — оцінки параметрів)

$$a^* = \bar{x}_B - \sqrt{3}\sigma_B, \quad b^* = \bar{x}_B + \sqrt{3}\sigma_B;$$

- 2) знайти щільність імовірності передбачуваного розподілу:

$$f(x) = \begin{cases} \frac{1}{b^* - a^*}, & x \in [a^*; b^*], \\ 0, & x \notin [a^*; b^*]; \end{cases}$$

- 3) визначити теоретичні частоти:

$$\begin{aligned} n'_1 &= n\mathbf{P}_1 = n \frac{1}{b^* - a^*} (x_1 - a^*); \\ n'_i &= n \frac{1}{b^* - a^*} (x_i - x_{i-1}), \quad i = 2, 3, \dots, N-1; \\ n'_N &= n \frac{1}{b^* - a^*} (b^* - x_{N-1}); \end{aligned}$$

- 4) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона, узявши кількість ступенів вільності $k = s - 3$, де s — кількість інтервалів у вибірці.

Приклад 1. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про рівномірний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 200$:

$(x_i; x_{i+1}]$	$(-4; -2]$	$(-2; 0]$	$(0; 2]$	$(2; 4]$	$(4; 6]$	$(6; 8]$	$(8; 10]$
n_i	28	27	34	23	25	31	32

Розв'язання. Спочатку обчислимо вибіркове середнє \bar{x}_B і середнє квадратичне відхилення вибірки σ_B . Для цього перетворимо інтервальний статистичний розподіл вибірки на точковий:

x_i	-3	-1	1	3	5	7	9
n_i	28	27	34	23	25	31	32

Отже,

$$\begin{aligned} \bar{x}_B &= \frac{-3 \cdot 28 - 1 \cdot 27 + 1 \cdot 34 + 3 \cdot 23 + 5 \cdot 25 + 7 \cdot 31 + 9 \cdot 32}{200} = \\ &= 3,11; \\ \sigma_B &= \left(\frac{3^2 \cdot 28 + 1 \cdot 27 + 1 \cdot 34 + 3^2 \cdot 23 + 5^2 \cdot 25 + 7^2 \cdot 31 + 9^2 \cdot 32}{200} - \right. \\ &\quad \left. - 3,11^2 \right)^{\frac{1}{2}} \approx 4,075. \end{aligned}$$

Знайдемо оцінки параметрів a і b рівномірного розподілу:

$$\begin{aligned} a^* &= \bar{x}_B - \sqrt{3}\sigma_B \approx 3,11 - \sqrt{3} \cdot 4,075 \approx -3,95; \\ b^* &= \bar{x}_B + \sqrt{3}\sigma_B \approx 3,11 + \sqrt{3} \cdot 4,075 \approx 10,17. \end{aligned}$$

Визначимо щільність рівномірного розподілу, що припускається:

$$f(x) = \begin{cases} \frac{1}{b^* - a^*}, & x \in [a^*; b^*], \\ 0, & x \notin [a^*; b^*] \end{cases} =$$

$$= \begin{cases} \frac{1}{10,17 + 3,95}, & x \in [-3,95; 10,17], \\ 0, & x \notin [-3,95; 10,17] \end{cases} \approx$$

$$\approx \begin{cases} 0,07, & x \in [-3,95; 10,17], \\ 0, & x \notin [-3,95; 10,17]. \end{cases}$$

Обчислимо теоретичні частоти:

$$n'_1 = n \cdot \frac{1}{b^* - a^*} \cdot (x_1 - a^*) = 200 \cdot 0,07 \cdot (-2 + 3,95) = 27,3;$$

$$n'_2 = n \cdot \frac{1}{b^* - a^*} \cdot (x_2 - x_1) = 200 \cdot 0,07 \cdot (0 + 2) = 28;$$

$$n'_3 = \dots = n'_6 = n'_2 = 28;$$

$$n'_7 = n \cdot \frac{1}{b^* - a^*} \cdot (b^* - x_6) = 200 \cdot 0,07 \cdot (10,17 - 8) = 30,38.$$

Порівняємо емпіричні та теоретичні частоти, використовуючи критерій Пірсона. Для цього обчислимо спочатку спостережуване значення критерію Пірсона за допомогою розрахункової табл. 19.

Таблиця 19

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	28	27,3	-0,7	0,49	0,02
2	27	28	1	1	0,04
3	34	28	-6	36	1,29
4	23	28	5	25	0,89
5	25	28	3	9	0,32
6	31	28	-3	9	0,32
7	32	30,38	-1,62	2,62	0,09
Σ	200				$\chi^2_{\text{спост}} = 2,97$

Із розрахункової табл. 19 отримуємо

$$\chi^2_{\text{спост}} = 2,97.$$

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 3 = 7 - 3 = 4$$

знаходимо критичну точку правосторонньої критичної області:

$$\chi_{\text{кр}}^2(0,05; 4) = 9,49.$$

Оскільки $\chi_{\text{спост}}^2 < \chi_{\text{кр}}^2$, гіпотезу про рівномірний розподіл не відхиляємо. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво. \square

15.6. Перевірка гіпотези про показниковий розподіл

Нехай задано інтервальний статистичний розподіл вибірки:

$(x_i; x_{i+1}]$	$(x_1; x_2]$	$(x_2; x_3]$	\dots	$(x_N; x_{N+1}]$
n_i	n_1	n_2	\dots	n_N

(n — обсяг вибірки).

Потрібно за допомогою критерію Пірсона перевірити гіпотезу про показниковий розподіл генеральної сукупності.

Правило 1. Для того щоб при рівні значущості α перевірити гіпотезу про показниковий розподіл досліджуваної випадкової величини X , необхідно:

- 1) знайти вибіркове середнє \bar{x}_b , узявши як варіанти середини інтервалів

$$x_i^* = \frac{x_i + x_{i+1}}{2};$$

- 2) узяти як оцінку параметра λ показникового розподілу величини, обернену до вибіркового середнього:

$$\lambda^* = \frac{1}{\bar{x}_b};$$

- 3) знайти ймовірності потрапляння випадкової величини X у частинні інтервали $(x_i; x_{i+1}]$ за формулою

$$\mathbf{P}_i = \mathbf{P}(x_i < X < x_{i+1}) = e^{-\lambda^* x_i} - e^{-\lambda^* x_{i+1}};$$

- 4) обчислити теоретичні частоти:

$$n'_i = n \cdot \mathbf{P}_i;$$

- 5) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона, узявши кількість ступенів вільності $k = s - 2$, де s — кількість інтервалів у вибірці.

Зауваження 1. Інтервали, які містять малочисельні частоти вибірки ($n_i < 5$), слід об'єднати, а частоти цих інтервалів додати. Якщо відбувалося об'єднання інтервалів, при визначенні кількості ступенів вільності за формулою $k = s - 2$ слід як s узяти кількість інтервалів вибірки, що залишилися після об'єднання інтервалів.

Приклад 1. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про показниковий розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 200$:

$(x_i; x_{i+1}]$	(0; 6]	(6; 12]	(12; 18]	(18; 24]	(24; 30]	(30; 36]	(36; 42]
n_i	115	51	18	9	4	2	1

Розв'язання. Оскільки ми маємо малочисельні частоти, то попередньо об'єднаємо їх. Отримаємо такий розподіл вибірки:

$(x_i; x_{i+1}]$	(0; 6]	(6; 12]	(12; 18]	(18; 24]	(24; 42]
n_i	115	51	18	9	7

Щоб знайти вибіркове середнє $\bar{x}_в$, перетворимо інтервальний статистичний розподіл на точковий:

x_i	3	9	15	21	33
n_i	115	51	18	9	7

Отже,

$$\bar{x}_B = \frac{3 \cdot 115 + 9 \cdot 51 + 15 \cdot 18 + 21 \cdot 9 + 33 \cdot 7}{200} = 7,47.$$

Знайдемо оцінку параметра показникового розподілу, що припускається:

$$\lambda^* = \frac{1}{\bar{x}_B} = \frac{1}{7,47} \approx 0,13.$$

Іншими словами, щільність показникового розподілу, що припускається, має вигляд

$$f(x) = \begin{cases} 0, & x < 0; \\ 0,13 \cdot e^{-0,13x}, & x \geq 0. \end{cases}$$

Обчислимо ймовірності потрапляння випадкової величини X у кожний з інтервалів за формулою

$$\mathbf{P}_i = \mathbf{P}(x_i < X < x_{i+1}) = e^{-\lambda^* x_i} - e^{-\lambda^* x_{i+1}}.$$

Отже,

$$\begin{aligned} \mathbf{P}_1 &= \mathbf{P}(0 < X < 6) = e^{-0,13 \cdot 0} - e^{-0,13 \cdot 6} \approx 0,5565; \\ \mathbf{P}_2 &= \mathbf{P}(6 < X < 12) = e^{-0,13 \cdot 6} - e^{-0,13 \cdot 12} \approx 0,2468; \\ \mathbf{P}_3 &= \mathbf{P}(12 < X < 18) = e^{-0,13 \cdot 12} - e^{-0,13 \cdot 18} \approx 0,1095; \\ \mathbf{P}_4 &= \mathbf{P}(18 < X < 24) = e^{-0,13 \cdot 18} - e^{-0,13 \cdot 24} \approx 0,0486; \\ \mathbf{P}_5 &= \mathbf{P}(24 < X < 42) = e^{-0,13 \cdot 24} - e^{-0,13 \cdot 42} \approx 0,0353. \end{aligned}$$

Знайдемо теоретичні частоти:

$$\begin{aligned} n'_1 &= n \cdot \mathbf{P}_1 = 200 \cdot 0,5565 = 111,3; \\ n'_2 &= n \cdot \mathbf{P}_2 = 200 \cdot 0,2468 = 49,36; \\ n'_3 &= n \cdot \mathbf{P}_3 = 200 \cdot 0,1095 = 21,9; \\ n'_4 &= n \cdot \mathbf{P}_4 = 200 \cdot 0,0486 = 9,72; \\ n'_5 &= n \cdot \mathbf{P}_5 = 200 \cdot 0,0353 = 7,06. \end{aligned}$$

Розділ 15. Статистична перевірка статистичних гіпотез

Порівнюємо емпіричні та теоретичні частоти, використовуючи критерій Пірсона. Для цього спочатку обчислимо спостережуване значення критерію Пірсона за допомогою розрахункової табл. 20.

Таблиця 20

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	115	111,3	-3,7	13,69	0,1230
2	51	49,36	-1,64	2,6896	0,0545
3	18	21,9	3,9	15,21	0,6945
4	9	9,72	0,72	0,5184	0,0533
5	7	7,06	0,06	0,0036	0,0005
Σ	200				$\chi^2_{\text{спост}} = 0,9258$

Із розрахункової табл. 20 отримуємо

$$\chi^2_{\text{спост}} = 0,9258.$$

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 2 = 5 - 2 = 3$$

знаходимо критичну точку правосторонньої критичної області:

$$\chi^2_{\text{кр}}(0,05; 3) = 7,81.$$

Оскільки $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, гіпотезу про показниковий розподіл не відхиляємо. Іншими словами, емпіричні та теоретичні частоти різняться несуттєво. \square

15.7. Перевірка гіпотези про біноміальний розподіл

Нехай проведено n дослідів. Кожний дослід складається із N незалежних випробувань, у кожному з яких імовірність появи деякої події A одна й та сама. Реєструється кількість появ події A в кожному досліді. У підсумку отримано статистичний розподіл дискретної

випадкової величини X , яка характеризує кількість появ події A (у першому рядку наведено кількість появ події A в одному досліді, а в другому рядку — частоту n_i , тобто кількість дослідів, у яких зареєстровано x_i появ події A):

x_i	0	1	...	N
n_i	n_0	n_1	...	n_N

Потрібно за допомогою критерію Пірсона перевірити гіпотезу про біноміальний закон розподілу дискретної випадкової величини X .

Правило 1. Для того щоб при рівні значущості α перевірити гіпотезу про біноміальний розподіл дискретної випадкової величини X , необхідно:

- 1) знайти за формулою Бернуллі

$$\mathbf{P}_N(i) = C_N^i p^i (1 - p)^{N-i}$$

імовірності $\mathbf{P}_N(i)$ появи події A рівно i раз у N випробуваннях ($i = 0, 1, \dots, s$, де s — максимальна кількість спостережених появ події A в одному досліді, тобто $s \leq N$);

- 2) знайти теоретичні частоти:

$$n'_i = n \cdot \mathbf{P}_N(i),$$

де n — кількість дослідів;

- 3) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона, поклавши кількість ступенів вільності $k = s - 1$ (вважається, що ймовірність p появи події A задана, тобто вона не оцінювалась за вибіркою і не об'єднувались малочисельні частоти).

Якщо ймовірність p була оцінена за вибіркою, то $k = s - 2$. Якщо, крім того, було об'єднано малочисельні частоти, то s — кількість варіант вибірки, що залишилися після об'єднання частот.

Приклад 1. Над подією A , імовірність появи якої дорівнює $0,3$, проведено $n = 100$ незалежних випробувань, кожне з яких складалося з $N = 7$ дослідів. Використовуючи критерій Пірсона, при рівні значущості $0,05$ перевірити, чи справджується гіпотеза про біноміальний розподіл випадкової величини X (кількість появ події A), якщо отримано таку вибірку:

x_i	0	1	2	3	4	5	6	7
n_i	4	23	31	23	11	5	2	1

Розв'язання. Ураховуючи, що

$$p = 0,3; \quad q = 1 - p = 1 - 0,3 = 0,7,$$

за формулою Бернуллі

$$\mathbf{P}_N(i) = C_N^i p^i q^{N-i}$$

обчислимо ймовірності $\mathbf{P}_N(i)$:

$$\begin{aligned} \mathbf{P}_7(0) &= C_7^0 0,3^0 0,7^7 \approx 0,0824; & \mathbf{P}_7(1) &= C_7^1 0,3^1 0,7^6 \approx 0,2471; \\ \mathbf{P}_7(2) &= C_7^2 0,3^2 0,7^5 \approx 0,3177; & \mathbf{P}_7(3) &= C_7^3 0,3^3 0,7^4 \approx 0,2269; \\ \mathbf{P}_7(4) &= C_7^4 0,3^4 0,7^3 \approx 0,0972; & \mathbf{P}_7(5) &= C_7^5 0,3^5 0,7^2 \approx 0,0250; \\ \mathbf{P}_7(6) &= C_7^6 0,3^6 0,7^1 \approx 0,0036; & \mathbf{P}_7(7) &= C_7^7 0,3^7 0,7^0 \approx 0,0002. \end{aligned}$$

Знайдемо теоретичні частоти $n'_i = n \cdot \mathbf{P}_N(i)$:

$$\begin{aligned} n'_0 &= 100 \cdot \mathbf{P}_7(0) = 8,24; & n'_1 &= 100 \cdot \mathbf{P}_7(1) = 24,71; \\ n'_2 &= 100 \cdot \mathbf{P}_7(2) = 31,77; & n'_3 &= 100 \cdot \mathbf{P}_7(3) = 22,69; \\ n'_4 &= 100 \cdot \mathbf{P}_7(4) = 9,72; & n'_5 &= 100 \cdot \mathbf{P}_7(5) = 2,5; \\ n'_6 &= 100 \cdot \mathbf{P}_7(6) = 0,36; & n'_7 &= 100 \cdot \mathbf{P}_7(7) = 0,02. \end{aligned}$$

Оскільки частоти $n_0 = 4$, $n_6 = 2$ і $n_7 = 1$ малочисельні (менше п'яти), об'єднаємо їх з іншими частотами, а саме:

$$n_1 = 23 + 4 = 27; \quad n_5 = 5 + 2 + 1 = 8.$$

Як теоретичні частоти, що відповідають об'єднаним частотам, візьмемо суму відповідних теоретичних частот:

$$n'_1 = 24,71 + 8,24 = 32,95; \quad n'_5 = 2,5 + 0,36 + 0,02 = 2,88.$$

Порівняємо емпіричні та теоретичні частоти за допомогою критерію Пірсона. Для цього складемо розрахункову табл. 21.

Таблиця 21

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	27	32,95	5,95	35,4025	1,0744
2	31	31,77	0,77	0,5929	0,0187
3	23	22,69	-0,31	0,0961	0,0042
4	11	9,72	-1,28	1,6384	0,1686
5	8	2,88	-5,12	26,2144	9,1022
Σ	100				$\chi^2_{\text{спост}} = 10,3681$

Із розрахункової табл. 21 отримуємо

$$\chi^2_{\text{спост}} = 10,3681.$$

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 1 = 5 - 1 = 4$$

знаходимо критичну точку правосторонньої критичної області:

$$\chi^2_{\text{кр}}(0,05; 4) = 9,49.$$

Оскільки $\chi^2_{\text{спост}} > \chi^2_{\text{кр}}$, гіпотезу про біноміальний розподіл відхиляємо. □

15.8. Перевірка гіпотези про розподіл Пуассона

Нехай задано точковий статистичний розподіл вибірки. Потрібно за допомогою критерію Пірсона перевірити гіпотезу про розподіл генеральної сукупності за законом Пуассона.

Правило 1. Для того щоб при рівні значущості α перевірити гіпотезу про те, що досліджувана випадкова величина X розподілена за законом Пуассона, необхідно:

- 1) знайти за заданим статистичним розподілом вибіркове середнє $\bar{x}_в$;
- 2) узяти як оцінку параметра λ розподілу Пуассона вибіркове середнє $\lambda^* = \bar{x}_в$;
- 3) знайти за формулою Пуассона (або за готовою таблицею дод. 3) імовірності \mathbf{P}_i появи рівно i подій у n випробуваннях ($i = 1, 2, \dots, r$, де r — максимальна кількість спостережених подій; n — обсяг вибірки);
- 4) визначити теоретичні частоти

$$n'_i = n \cdot \mathbf{P}_i;$$

- 5) порівняти емпіричні та теоретичні частоти за допомогою критерію Пірсона, узявши кількість ступенів вільності $k = s - 2$, де s — кількість варіант вибірки (якщо проводилося об'єднання малочисельних частот в одну групу, то s — кількість варіант, що залишилися після об'єднання частот).

Приклад 1. У $n = 1000$ перевірках партій товару реєструвалася кількість x_i неякісної продукції, внаслідок чого було отримано такий статистичний розподіл кількості x_i браку в n_i партіях товару:

x_i	0	1	2	3	4	5	6	7
n_i	242	349	234	107	43	21	3	1

Потрібно при рівні значущості 0,05 перевірити гіпотезу про те, що кількість неякісної продукції X розподілена за законом Пуассона.

Розв'язання. Спочатку знайдемо вибіркове середнє:

$$\begin{aligned}\bar{x}_B &= \frac{0 \cdot 242 + 1 \cdot 349 + 2 \cdot 234 + 3 \cdot 107 + 4 \cdot 43 + 5 \cdot 21 + 6 \cdot 3 + 7 \cdot 1}{1000} = \\ &= 1,44.\end{aligned}$$

Візьмемо як оцінку параметра λ розподілу Пуассона вибіркове середнє:

$$\lambda^* = \bar{x}_B = 1,44.$$

Отже, закон Пуассона, що припускається, має такий вигляд:

$$P_{1000}(i) = 1,44^i \cdot \frac{e^{-1,44}}{i!}.$$

Поклавши $i = 0, 1, \dots, 7$, обчислимо ймовірності $P_i = P_{1000}(i)$:

$$\begin{aligned}P_0 &= 1,44^0 \cdot \frac{e^{-1,44}}{0!} \approx 0,2369; & P_1 &= 1,44^1 \cdot \frac{e^{-1,44}}{1!} \approx 0,3412; \\ P_2 &= 1,44^2 \cdot \frac{e^{-1,44}}{2!} \approx 0,2456; & P_3 &= 1,44^3 \cdot \frac{e^{-1,44}}{3!} \approx 0,1179; \\ P_4 &= 1,44^4 \cdot \frac{e^{-1,44}}{4!} \approx 0,0424; & P_5 &= 1,44^5 \cdot \frac{e^{-1,44}}{5!} \approx 0,0122; \\ P_6 &= 1,44^6 \cdot \frac{e^{-1,44}}{6!} \approx 0,0029; & P_7 &= 1,44^7 \cdot \frac{e^{-1,44}}{7!} \approx 0,0006.\end{aligned}$$

Знайдемо теоретичні частоти $n'_i = n \cdot P_i$:

$$\begin{aligned}n'_0 &= 1000 \cdot P_0 = 236,9; & n'_1 &= 1000 \cdot P_1 = 341,2; \\ n'_2 &= 1000 \cdot P_2 = 245,6; & n'_3 &= 1000 \cdot P_3 = 117,9; \\ n'_4 &= 1000 \cdot P_4 = 42,4; & n'_5 &= 1000 \cdot P_5 = 12,2; \\ n'_6 &= 1000 \cdot P_6 = 2,9; & n'_7 &= 1000 \cdot P_7 = 0,6.\end{aligned}$$

Оскільки частоти $n_6 = 3$ і $n_7 = 1$ малочисельні (менше п'яти), об'єднаємо їх із частотою n_5 , а саме

$$n_5 = 21 + 3 + 1 = 25.$$

Розділ 15. Статистична перевірка статистичних гіпотез

Як теоретичну частоту, що відповідає об'єднаній частоті, візьмемо суму відповідних теоретичних частот:

$$n'_5 = 12,2 + 2,9 + 0,6 = 15,7.$$

Порівняємо емпіричні та теоретичні частоти за допомогою критерію Пірсона. Для цього складемо розрахункову табл. 22.

Таблиця 22

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
0	242	236,9	-5,1	26,01	0,1098
1	349	341,2	-7,8	60,84	0,1783
2	234	245,6	11,6	134,56	0,5479
3	107	117,9	10,9	118,81	1,0077
4	43	42,4	-0,6	0,36	0,0085
5	25	15,7	-9,3	86,49	5,5089
Σ	1000				$\chi^2_{\text{спост}} = 7,3611$

Із розрахункової табл. 22 отримуємо

$$\chi^2_{\text{спост}} = 7,3611.$$

За таблицею критичних точок розподілу χ^2 (дод. 7) при рівні значущості $\alpha = 0,05$ і кількості ступенів вільності

$$k = s - 1 = 6 - 2 = 4$$

знаходимо критичну точку правосторонньої критичної області:

$$\chi^2_{\text{кр}}(0,05; 4) = 9,49.$$

Оскільки $\chi^2_{\text{спост}} < \chi^2_{\text{кр}}$, немає підстав відхилити гіпотезу про розподіл випадкової величини X (кількість неякісного товару в партії) за законом Пуассона. \square

Задачі до розділу 15

Задача 1. Із нормальної генеральної сукупності з відомим середнім квадратичним відхиленням $\sigma = 3,2$ отримано деяку вибірку обсягу $n = 64$, за якою знайдено вибіркоче середнє $\bar{x}_в = 13$. Потрібно при рівні значущості $0,05$ перевірити нульову гіпотезу $H_0: a = a_0 = 12$ при конкуруючій гіпотезі $H_1: a \neq 12$.

Відповідь. Основна гіпотеза відхиляється.

Задача 2. Із нормальної генеральної сукупності з відомим середнім квадратичним відхиленням $\sigma = 3,3$ отримано вибірку обсягу $n = 121$, за якою знайдено вибіркоче середнє $\bar{x}_в = 10,6$. Потрібно при рівні значущості $0,01$ перевірити нульову гіпотезу $H_0: a = 10$ при конкуруючій гіпотезі

- а) $H_1: a \neq 10$;
- б) $H_1: a > 10$;
- в) $H_1: a < 10$.

Крім того, знайти потужності правостороннього та двостороннього критеріїв.

Відповідь. Основні гіпотези не відхиляються в усіх трьох випадках. Потужності правостороннього та двостороннього критеріїв дорівнюють відповідно $0,3721$ і $0,2824$.

Задача 3. Для вибірки обсягу $n = 20$, отриманої з нормальної генеральної сукупності, знайдено вибіркоче середнє $\bar{x} = 37$ і «виправлене» вибіркоче середнє квадратичне відхилення $s = 4,2$. Потрібно при рівні значущості $\alpha = 0,05$ перевірити основну гіпотезу $H_0: a = a_0 = 35$ при альтернативній гіпотезі $H_1: a \neq 35$.

Відповідь. Основна гіпотеза відхиляється.

Задача 4. Контрольна довжина деякої деталі має становити $a = a_0 = 12$ мм. Контрольні вимірювання 25 випадково відібраних деталей виявили такі результати:

Довжина, мм	11	11,5	12	12,5	13
Кількість деталей	6	8	7	3	1

Потрібно при рівні значущості 0,01 перевірити основну гіпотезу $H_0: a = a_0 = 12$ при конкуруючій гіпотезі:

- а) $H_1: a \neq a_0$;
- б) $H_1: a > a_0$;
- в) $H_1: a < a_0$.

Відповідь. а) і б) Основна гіпотеза не відхиляється; в) основна гіпотеза відхиляється.

Задача 5. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 31$, для якої знайдено вибіркочну виправлену дисперсію $s^2 = 44$. Перевірити при рівні значущості $\alpha = 0,01$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 25$, якщо альтернативна гіпотеза $H_1: \sigma_0^2 > 25$.

Відповідь. Основна гіпотеза відхиляється.

Задача 6. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 26$, для якої знайдено вибіркочну «виправлену» дисперсію $s^2 = 3,2$. Перевірити при рівні значущості $\alpha = 0,05$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 2$, якщо альтернативна гіпотеза:

- а) $H_1: \sigma_0^2 > 2$;
- б) $H_1: \sigma_0^2 \neq 2$;
- в) $H_1: \sigma_0^2 < 2$.

Відповідь. а) Основна гіпотеза відхиляється; б) і в) основна гіпотеза не відхиляється.

Задача 7. Із нормальної генеральної сукупності отримано вибірку обсягу $n = 200$:

x_i	-1	0	1	2	3	4	5
n_i	17	24	30	34	35	33	27

Потрібно перевірити при рівні значущості $\alpha = 0,01$ основну гіпотезу $H_0: \sigma^2 = \sigma_0^2 = 3$, якщо альтернативна гіпотеза $H_1: \sigma_0^2 > 3$.

Відповідь. Основна гіпотеза відхиляється.

Задача 8. Для деякої події в результаті 100 незалежних випробувань було знайдено відносну частоту $\frac{m}{n} = 0,34$. При рівні значущості 0,05 потрібно перевірити основну гіпотезу $H_0: p = p_0 = 0,25$ при конкуруючій гіпотезі $H_1: p \neq 0,25$.

Відповідь. Основна гіпотеза відхиляється.

Задача 9. Для деякої події в результаті 256 незалежних випробувань було знайдено відносну частоту $\frac{m}{n} = 0,375$. При рівні значущості 0,01 потрібно перевірити основну гіпотезу $H_0: p = p_0 = 0,4$ при конкуруючій гіпотезі:

а) $H_1: p > 0,4$;

б) $H_1: p < 0,4$.

Відповідь. Основна гіпотеза не відхиляється в обох випадках.

Задача 10. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про нормальний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 200$:

x_i	15	16	17	18	19	20	21	22
n_i	8	28	31	41	33	28	24	7

Відповідь. Гіпотеза про нормальний розподіл відхиляється.

Задача 11. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про нормальний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 200$:

$(x_i; x_{i+1}]$	(0; 3]	(3; 6]	(6; 9]	(9; 12]	(12; 15]	(15; 18]	(18; 21]
n_i	15	25	45	51	34	22	8

Відповідь. Гіпотеза про нормальний розподіл відхиляється.

Задача 12. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про рівномірний розподіл генеральної сукупності X , якщо з цієї сукупності отримано таку вибірку обсягу $n = 100$:

$(x_i; x_{i+1}]$	(4; 7]	(7; 10]	(10; 13]	(13; 16]	(16; 19]	(19; 22]	(22; 25]
n_i	12	9	21	8	11	19	20

Відповідь. Гіпотеза про рівномірний розподіл відхиляється.

Задача 13. Над подією A , ймовірність появи якої дорівнює 0,4, проведено $n = 200$ незалежних випробувань, кожне з яких складалося з $N = 10$ дослідів. Використовуючи критерій Пірсона, при рівні значущості 0,05 перевірити, чи справджується гіпотеза про біноміальний розподіл випадкової величини X (кількість появ події A), якщо отримано таку вибірку:

x_i	0	1	2	3	4	5	6	7	8	9
n_i	2	12	31	42	39	35	27	5	4	2

Вказівка. Слід об'єднати малочисельні частоти: n_0 об'єднати з n_1 , а n_8, n_9 і n_{10} — з n_7 .

Відповідь. Гіпотеза про біноміальний розподіл не відхиляється.

Задача 14. У $n = 1000$ перевірках партій товару реєструвалася кількість x_i неякісної продукції, внаслідок чого було отримано такий статистичний розподіл кількості x_i браку в n_i партіях товару:

x_i	0	1	2	3	4	5	6	7
n_i	427	363	154	41	9	3	2	1

Потрібно при рівні значущості 0,05 перевірити гіпотезу про те, що кількість неякісної продукції X розподілена за законом Пуассона.

Вказівка. Слід об'єднати малочисельні частоти n_6 , і n_7 з n_5 .

Відповідь. Гіпотеза про розподіл Пуассона не відхиляється.

Розділ 16. Елементи дисперсійного аналізу

16.1. Однофакторний дисперсійний аналіз для незв'язних вибірок

Нехай на кількісну нормально розподілену ознаку X діє деякий фактор F , який має p сталих рівнів F_1, F_2, \dots, F_p . На кожному рівні здійснено по q випробувань **над різними вибірками**. Результати спостережень — числа x_{ij} , де i — номер випробування, $i = 1, 2, \dots, q$; j — номер рівня фактора, $j = 1, 2, \dots, p$, — записують у вигляді таблиці (табл. 23).

Таблиця 23

Номер випробування i	Рівень фактора			
	F_1	F_2	\dots	F_p
1	x_{11}	x_{12}	\dots	x_{1p}
2	x_{21}	x_{22}	\dots	x_{2p}
\dots	\dots	\dots	\dots	\dots
q	x_{q1}	x_{q2}	\dots	x_{qp}
Групове середнє $\bar{x}_{гр}$	$\bar{x}_{гр. 1}$	$\bar{x}_{гр. 2}$	\dots	$\bar{x}_{гр. p}$

Задача полягає в такому: на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові.

Щоб розв'язати цю задачу, знаходять:

- а) *загальну суму* квадратів відхилень спостережуваних значень ознаки від загального вибіркового середнього:

$$S_{\text{заг}} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x}_в)^2;$$

- б) *факторну суму* квадратів відхилень групових середніх від загального вибіркового середнього (характеризує розсіювання

«між групами»):

$$S_{\text{факт}} = q \sum_{j=1}^p (\bar{x}_{\text{гр. } j} - \bar{x}_v)^2;$$

в) *залишкову суму* квадратів відхилень спостережуваних значень групи від свого групового середнього (характеризує розсіювання «всередині групи»):

$$S_{\text{зал}} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x}_{\text{гр. } j})^2.$$

Практично залишкову суму знаходять за формулою

$$S_{\text{зал}} = S_{\text{зар}} - S_{\text{факт}}.$$

Загальну та факторну суми зручніше обчислювати за такими формулами:

$$S_{\text{зар}} = \sum_{j=1}^p P_j - \frac{\left(\sum_{j=1}^p R_j\right)^2}{pq};$$

$$S_{\text{факт}} = \frac{\sum_{j=1}^p R_j^2}{q} - \frac{\left(\sum_{j=1}^p R_j\right)^2}{pq},$$

де $P_j = \sum_{i=1}^q x_{ij}^2$ — сума квадратів спостережуваних значень ознаки на рівні F_j ; $R_j = \sum_{i=1}^q x_{ij}$ — сума спостережуваних значень ознаки на рівні F_j .

Якщо спостережувані значення ознаки порівняно великі, то для спрощення обчислень віднімають від кожного спостережуваного значення одне й те саме число C , яке приблизно дорівнює загальному середньому. Якщо зменшені значення позначити

$$y_{ij} = x_{ij} - C,$$

то

$$S_{\text{заг}} = \sum_{j=1}^p Q_j - \frac{\left(\sum_{j=1}^p T_j\right)^2}{pq};$$

$$S_{\text{факт}} = \frac{\sum_{j=1}^p T_j^2}{q} - \frac{\left(\sum_{j=1}^p T_j\right)^2}{pq},$$

де $Q_j = \sum_{i=1}^q y_{ij}^2$ — сума квадратів зменшених значень ознаки на рівні F_j ; $T_j = \sum_{i=1}^q y_{ij}$ — сума зменшених значень ознаки на рівні F_j .

Поділивши факторну та залишкову суми на відповідну кількість ступенів вільності

$$k_{\text{факт}} = p - 1, \quad k_{\text{зал}} = p(q - 1),$$

знаходять факторну та залишкову дисперсії:

$$s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p - 1}; \quad s_{\text{зал}}^2 = \frac{S_{\text{зал}}}{p(q - 1)}.$$

Факторну та залишкову дисперсії порівнюють за *критерієм Фішера — Снедекора*:

$$F_{\text{спост}} = \frac{s_{\text{факт}}^2}{s_{\text{зал}}^2}.$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості α і кількості ступенів вільності

$$k_1 = k_{\text{факт}}, \quad k_2 = k_{\text{зал}}$$

знаходять критичну точку $F_{\text{кр}}(\alpha; k_1, k_2)$.

Якщо $F_{\text{спост}} < F_{\text{кр}}$, відмінність групових середніх вважають несуттєвою. Якщо $F_{\text{спост}} > F_{\text{кр}}$, цю відмінність вважають суттєвою.

Зауваження 1. Однофакторний дисперсійний аналіз вимагає не менше трьох рівнів фактора й не менше двох спостережень на кожному рівні, причому для всіх рівнів кількість спостережень має бути однаковою.

Зауваження 2. Результуюча ознака має бути нормально розподіленою в досліджуваній вибірці.

Зауваження 3. Якщо факторна дисперсія виявиться меншою від залишкової, одразу можна робити висновок про те, що немає підстав відхилити основну гіпотезу. Іншими словами, у даному разі подальші обчислення зайві.

Зауваження 4. Якщо спостережувані значення x_{ij} кратні деякому числу k (наприклад, десяткові дробі кратні числу 10^{-d} , де d — кількість цифр після коми), доцільно перейти до нових значень:

$$y_{ij} = \frac{x_{ij}}{k}.$$

При цьому факторна та залишкова дисперсії зменшаться у k^2 разів, але їх відношення не зміниться.

Приклад 1. Проведено по сім випробувань на кожному з чотирьох рівнів фактора над різними вибірками. Методом дисперсійного аналізу при рівні значущості 0,05 перевірити основну гіпотезу про рівність групових середніх. Припускається, що вибірки отримані з нормальних сукупностей з однаковими дисперсіями. Результати випробувань подано в табл. 24.

Розв'язання. Знайдемо загальне середнє:

$$\begin{aligned} \bar{x}_в = & (24 + 26 + 21 + 25 + 25 + 27 + 22 + 26 + 26 + 28 + 24 + 28 + \\ & + 29 + 29 + 25 + 28 + 31 + 30 + 27 + 29 + 32 + 32 + 29 + 30 + \\ & + 33 + 35 + 30 + 33)/(7 \cdot 4) = 28. \end{aligned}$$

Для спрощення розрахунків віднімемо від кожного спостережуваного значення x_{ij} загальне середнє $\bar{x}_в = 28$, тобто перейдемо до зменшених величин:

$$y_{ij} = x_{ij} - 28.$$

Таблиця 24

Номер випробування i	Рівень фактора			
	F_1	F_2	F_3	F_4
1	24	26	21	25
2	25	27	22	26
3	26	28	24	28
4	29	29	25	28
5	31	30	27	29
6	32	32	29	30
7	33	35	30	33
Групове середнє $\bar{x}_{гр}$	28,57143	29,57143	25,42857	28,42857

Складемо розрахункову табл. 25.

Таблиця 25

Номер випробування i	Рівень фактора								Підсумковий стовпчик
	F_1		F_2		F_3		F_4		
	y_{i1}	y_{i1}^2	y_{i2}	y_{i2}^2	y_{i3}	y_{i3}^2	y_{i4}	y_{i4}^2	
1	-4	16	-2	4	-7	49	-3	9	
2	-3	9	-1	1	-6	36	-2	4	
3	-2	4	0	0	-4	16	0	0	
4	1	1	1	1	-3	9	0	0	
5	3	9	2	4	-1	1	1	1	
6	4	16	4	16	1	1	2	4	
7	5	25	7	49	2	4	5	25	
$Q_j = \sum_{i=1}^7 y_{ij}^2$		80		75		116		43	$\sum_{j=1}^4 Q_j = 314$
$T_j = \sum_{i=1}^7 y_{ij}$	4		11		-18		3		$\sum_{j=1}^4 T_j = 0$
T_j^2	16		121		324		9		$\sum_{j=1}^4 T_j^2 = 470$

Використовуючи підсумковий стовпчик табл. 25, знаходимо загальну й факторну суми квадратів відхилень, ураховуючи, що кількість рівнів фактора $p = 4$, кількість випробувань на кожному рівні $q = 7$:

$$S_{\text{заг}} = \sum_{j=1}^p Q_j - \frac{\left(\sum_{j=1}^p T_j\right)^2}{pq} = 314 - 0 = 314;$$

$$S_{\text{факт}} = \frac{\sum_{j=1}^p T_j^2}{q} - \frac{\left(\sum_{j=1}^p T_j\right)^2}{pq} = \frac{470}{7} - 0 \approx 67,14.$$

Знайдемо залишкову суму квадратів відхилень:

$$S_{\text{зал}} = S_{\text{заг}} - S_{\text{факт}} = 314 - 67,14 = 246,86.$$

Визначимо кількість ступенів вільності:

$$k_{\text{факт}} = p - 1 = 4 - 1 = 3, \quad k_{\text{зал}} = p(q - 1) = 4 \cdot (7 - 1) = 24.$$

Обчислимо факторну та залишкову дисперсії:

$$s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p - 1} = \frac{67,14}{3} \approx 22,38;$$

$$s_{\text{зал}}^2 = \frac{S_{\text{зал}}}{p(q - 1)} = \frac{246,86}{24} \approx 10,29.$$

Порівняємо факторну та залишкову дисперсії за допомогою критерію Фішера — Снедекора. Для цього спочатку знайдемо спостережуване значення критерію:

$$F_{\text{спост}} = \frac{s_{\text{факт}}^2}{s_{\text{зал}}^2} = \frac{22,38}{10,29} \approx 2,175.$$

Ураховуючи, що кількість ступенів вільності чисельника $k_1 = 3$, а знаменника — $k_2 = 24$, при рівні значущості $\alpha = 0,05$ за дод. 8 знаходимо критичну точку:

$$F_{\text{кр}}(0,05; 3; 24) = 3,01.$$

Оскільки $F_{\text{спост}} < F_{\text{кр}}$, основну гіпотезу про рівність групових середніх не відхиляємо. Іншими словами, групові середні відрізняються не суттєво. \square

16.2. Однофакторний дисперсійний аналіз для зв'язних вибірок

Нехай на кількісну нормально розподілену ознаку X діє фактор F , який має p сталих рівнів F_1, F_2, \dots, F_p . На кожному рівні проведено випробування **над однією й тією самою вибіркою**, що складається з q елементів. Результати спостережень — числа x_{ij} , де i — номер випробування, $i = 1, 2, \dots, q$, j — номер рівня фактора F , $j = 1, 2, \dots, p$, — записують у вигляді таблиці (табл. 26).

Таблиця 26

Номер випробування i	Рівень фактора				Індивідуальні середні \bar{x}_i
	F_1	F_2	\dots	F_p	
1	x_{11}	x_{12}	\dots	x_{1p}	\bar{x}_1
2	x_{21}	x_{22}	\dots	x_{2p}	\bar{x}_2
\dots	\dots	\dots	\dots	\dots	\dots
q	x_{q1}	x_{q2}	\dots	x_{qp}	\bar{x}_q
Групове середнє $\bar{x}_{\text{гр}}$	$\bar{x}_{\text{гр. 1}}$	$\bar{x}_{\text{гр. 2}}$	\dots	$\bar{x}_{\text{гр. } p}$	

У цьому разі виникають дві задачі:

- 1) на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- 2) на рівні значущості α перевірити основну гіпотезу про рівність індивідуальних середніх за умови, що індивідуальні генеральні дисперсії хоч і невідомі, але однакові.

Щоб розв'язати ці задачі, знаходять:

- а) *загальну суму* квадратів відхилень спостережуваних значень ознаки від загального вибіркового середнього:

$$S_{\text{заг}} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x}_v)^2;$$

- б) *факторну суму* квадратів відхилень групових середніх від загального вибіркового середнього (характеризує розсіювання «між групами»):

$$S_{\text{факт}} = q \sum_{j=1}^p (\bar{x}_{\text{гр. } j} - \bar{x}_v)^2;$$

- в) *індивідуальну суму* квадратів відхилень індивідуальних середніх від загального вибіркового середнього (характеризує розсіювання «між індивідами»):

$$S_{\text{інд}} = p \sum_{i=1}^q (\bar{x}_i - \bar{x}_v)^2;$$

- г) *залишкову суму* за формулою

$$S_{\text{зал}} = S_{\text{заг}} - S_{\text{факт}} - S_{\text{інд}}.$$

Поділивши факторну, індивідуальну та залишкову суми на відповідну кількість ступенів вільності

$$k_{\text{факт}} = p - 1, \quad k_{\text{інд}} = q - 1, \quad k_{\text{зал}} = (p - 1)(q - 1),$$

знаходять факторну, індивідуальну та залишкову дисперсії:

$$s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p - 1}; \quad s_{\text{інд}}^2 = \frac{S_{\text{інд}}}{q - 1}; \quad s_{\text{зал}}^2 = \frac{S_{\text{зал}}}{(p - 1)(q - 1)}.$$

Нарешті, порівнюють факторну та індивідуальну дисперсії із залишковою дисперсією за *критерієм Фішера — Снедекора*:

$$F_{\text{спост}}^{\text{факт}} = \frac{s_{\text{факт}}^2}{s_{\text{зал}}^2}, \quad F_{\text{спост}}^{\text{інд}} = \frac{s_{\text{інд}}^2}{s_{\text{зал}}^2}.$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості α і кількості ступенів вільності ($k_{\text{факт}}$ — чисельник, $k_{\text{інд}}$ — чисельник, $k_{\text{зал}}$ — знаменник) знаходять критичні точки $F_{\text{кр}}^{\text{факт}}(\alpha; k_{\text{факт}}, k_{\text{зал}})$ і $F_{\text{кр}}^{\text{інд}}(\alpha; k_{\text{інд}}, k_{\text{зал}})$.

Якщо $F_{\text{спост}}^{\text{факт}} < F_{\text{кр}}^{\text{факт}}$, відмінність групових середніх вважають несуттєвою. Якщо $F_{\text{спост}}^{\text{факт}} > F_{\text{кр}}^{\text{факт}}$, відмінність групових середніх вважають суттєвою.

Якщо $F_{\text{спост}}^{\text{інд}} < F_{\text{кр}}^{\text{інд}}$, відмінність індивідуальних середніх вважають несуттєвою. Якщо $F_{\text{спост}}^{\text{інд}} > F_{\text{кр}}^{\text{інд}}$, відмінність індивідуальних середніх вважають суттєвою.

Зауваження 1. Однофакторний дисперсійний аналіз для зв'язних вибірок вимагає не менше трьох рівнів фактора й не менше двох елементів вибірки, ознаки яких вимірюються на всіх рівнях фактора.

Зауваження 2. Результуюча ознака має бути нормально розподіленою в досліджуваній вибірці.

Зауваження 3. Якщо факторна або індивідуальна дисперсії виявляться меншими від залишкової, одразу можна робити висновок про те, що немає підстав відхиляти основну гіпотезу. Іншими словами, у таких випадках подальші обчислення зайві.

Зауваження 4. Якщо спостережувані значення x_{ij} кратні деякому числу k (наприклад, десяткові дробі кратні числу 10^{-d} , де d — кількість цифр після коми), доцільно перейти до нових значень:

$$y_{ij} = \frac{x_{ij}}{k}.$$

При цьому факторна, індивідуальна та залишкова дисперсії зменшаться у k^2 разів, але їх відношення не зміниться.

Приклад 1. Проведено випробування на кожному з чотирьох рівнів фактора над однією й тією самою вибіркою з шести елементів. Методом дисперсійного аналізу при рівні значущості 0,01 перевірити основну гіпотезу про рівність групових середніх та основну гіпотезу про рівність індивідуальних середніх. Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в табл. 27.

Розділ 16. Елементи дисперсійного аналізу

Таблиця 27

Номер випробування i	Рівень фактора				Індивідуальні середні \bar{x}_i
	F_1	F_2	F_3	F_4	
1	5	11	7	13	9
2	3	9	8	12	8
3	5	12	5	14	9
4	6	11	6	17	10
5	4	8	5	11	7
6	7	15	5	17	11
Групове середнє $\bar{x}_{\text{гр}}$	5	11	6	14	

Розв'язання. Знайдемо загальне середнє:

$$\bar{x}_B = (5 + 11 + 7 + 13 + 3 + 9 + 8 + 12 + 5 + 12 + 5 + 14 + 6 + 11 + 6 + 17 + 4 + 8 + 5 + 11 + 7 + 15 + 5 + 17) / (6 \cdot 4) = 9.$$

Для обчислення загальної суми квадратів відхилень спостережуваних значень ознак від загального вибіркового середнього складемо розрахункову табл. 28, в яку замість значень x_{ij} запишемо значення

$$y_{ij} = (x_{ij} - \bar{x}_B)^2.$$

Таблиця 28

Номер випробування i	Рівень фактора				$(\bar{x}_i - \bar{x}_B)^2$
	F_1	F_2	F_3	F_4	
1	16	4	4	16	0
2	36	0	1	9	1
3	16	9	16	25	0
4	9	4	9	64	1
5	25	1	16	4	4
6	4	36	16	64	4
$(\bar{x}_{\text{гр}} - \bar{x}_B)^2$	16	4	9	25	$S_{\text{заг}} = 404$

За табл. 28 визначимо:

$$\begin{aligned}S_{\text{заг}} &= 404; \\S_{\text{факт}} &= q \sum_{j=1}^p (\bar{x}_{\text{гр. } j} - \bar{x}_{\text{в}})^2 = 6 \cdot (16 + 4 + 9 + 25) = 324; \\S_{\text{інд}} &= p \sum_{i=1}^q (\bar{x}_i - \bar{x}_{\text{в}})^2 = 4 \cdot (0 + 1 + 0 + 1 + 4 + 4) = 40.\end{aligned}$$

Залишкова сума квадратів відхилень

$$S_{\text{зал}} = S_{\text{заг}} - S_{\text{факт}} - S_{\text{інд}} = 404 - 324 - 40 = 40.$$

Знайдемо кількість ступенів вільності:

$$\begin{aligned}k_{\text{факт}} &= p - 1 = 4 - 1 = 3, \\k_{\text{інд}} &= q - 1 = 6 - 1 = 5, \\k_{\text{зал}} &= (p - 1)(q - 1) = 3 \cdot 5 = 15.\end{aligned}$$

Знайдемо факторну, індивідуальну та залишкову дисперсії:

$$\begin{aligned}s_{\text{факт}}^2 &= \frac{S_{\text{факт}}}{p - 1} = \frac{324}{3} = 108; \\s_{\text{інд}}^2 &= \frac{S_{\text{інд}}}{q - 1} = \frac{40}{5} = 8; \\s_{\text{зал}}^2 &= \frac{S_{\text{зал}}}{(p - 1)(q - 1)} = \frac{40}{15} = \frac{8}{3}.\end{aligned}$$

Порівняємо факторну та індивідуальну дисперсії із залишковою дисперсією за критерієм Фішера — Снедекора:

$$\begin{aligned}F_{\text{спост}}^{\text{факт}} &= \frac{s_{\text{факт}}^2}{s_{\text{зал}}^2} = \frac{108 \cdot 3}{8} = 40,5; \\F_{\text{спост}}^{\text{інд}} &= \frac{s_{\text{інд}}^2}{s_{\text{зал}}^2} = \frac{8 \cdot 3}{8} = 3.\end{aligned}$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості $\alpha = 0,01$ і кількості ступенів вільності ($k_{\text{факт}}$ — чисельник, $k_{\text{інд}}$ — чисельник, $k_{\text{зал}}$ — знаменник) знаходимо критичні точки:

$$F_{\text{кр}}^{\text{факт}}(0,01; 3; 15) = 5,42;$$

$$F_{\text{кр}}^{\text{інд}}(0,01; 5; 15) = 4,56.$$

Оскільки $F_{\text{спост}}^{\text{факт}} > F_{\text{кр}}^{\text{факт}}$, основну гіпотезу про рівність групових середніх відхиляємо. Іншими словами, відмінність групових середніх суттєва.

Оскільки $F_{\text{спост}}^{\text{інд}} < F_{\text{кр}}^{\text{інд}}$, основну гіпотезу про рівність індивідуальних середніх не відхиляємо. Іншими словами, відмінність індивідуальних середніх несуттєва. \square

16.3. Двофакторний дисперсійний аналіз для незв'язних вибірок

Нехай на кількісну нормально розподілену ознаку X діють два фактори F і G , які мають відповідно p і g сталих рівнів F_1, F_2, \dots, F_p і G_1, G_2, \dots, G_g . На кожній парі $(F_i; G_j)$ рівнів здійснено по q випробувань **над різними вибірками**. Результати спостережень — числа (x_{ijk}) , де i — номер випробування, $i = 1, 2, \dots, q$, j — номер рівня фактора F , $j = 1, 2, \dots, p$, k — номер рівня фактора G , $k = 1, 2, \dots, g$ — записують у вигляді таблиці (табл. 29).

У табл. 29 уведено такі позначення:

- $S_{jk} = \sum_{i=1}^q x_{ijk}$ — сума значень ознаки для вибірки, яка піддавалася дослідженню на рівні F_j фактора F і рівні G_k фактора G , $j = 1, 2, \dots, p$, $k = 1, 2, \dots, g$;
- $S_{F_j} = \sum_{k=1}^g S_{jk}$ — сума значень ознаки для вибірок, які піддавалися дослідженню на рівні F_j фактора F , $j = 1, 2, \dots, p$;

- $S_{G_k} = \sum_{j=1}^p S_{jk}$ — сума значень ознаки для вибірок, які піддавалися дослідженню на рівні G_k фактора G , $k = 1, 2, \dots, g$;
- $S = \sum_{j=1}^p S_{F_j} = \sum_{k=1}^g S_{G_k}$ — сума значень ознаки для всіх вибірок, які піддавалися дослідженню.

Таблиця 29

Рівні фактора G	Рівні фактора F				Загальна сума
	F_1	F_2	\dots	F_p	
G_1	$x_{111}, x_{211}, \dots, x_{q11}$	$x_{121}, x_{221}, \dots, x_{q21}$	\dots	$x_{1p1}, x_{2p1}, \dots, x_{qp1}$	S_{G_1}
Сума	S_{11}	S_{21}	\dots	S_{p1}	
G_2	$x_{112}, x_{212}, \dots, x_{q12}$	$x_{122}, x_{222}, \dots, x_{q22}$	\dots	$x_{1p2}, x_{2p2}, \dots, x_{qp2}$	S_{G_2}
Сума	S_{12}	S_{22}	\dots	S_{p2}	
\dots	\dots	\dots	\dots	\dots	\dots
G_g	$x_{11g}, x_{21g}, \dots, x_{q1g}$	$x_{12g}, x_{22g}, \dots, x_{q2g}$	\dots	$x_{1pg}, x_{2pg}, \dots, x_{qpg}$	S_{G_g}
Сума	S_{1g}	S_{2g}	\dots	S_{pg}	
Загальна сума	S_{F_1}	S_{F_2}	\dots	S_{F_p}	S

Крім цих позначень уведемо такі позначення для середніх:

$$\begin{aligned} \bar{x}_B &= \frac{1}{ppg} S; \\ \bar{x}_{\text{гр. } j}^F &= \frac{1}{qg} S_{F_j}, \quad j = 1, 2, \dots, p; \\ \bar{x}_{\text{гр. } k}^G &= \frac{1}{qp} S_{G_k}, \quad k = 1, 2, \dots, g; \\ \bar{x}_{\text{гр. } jk}^{FG} &= \frac{1}{q} S_{jk}, \quad j = 1, 2, \dots, p, \quad k = 1, 2, \dots, g. \end{aligned}$$

У цьому разі виникають три задачі:

- 1) для фактора F на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- 2) для фактора G на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- 3) на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градаціях фактора G , і навпаки.

Щоб розв'язати ці задачі, знаходять:

- а) *загальну суму* квадратів відхилень спостережуваних значень ознаки від загального вибіркового середнього:

$$S_{\text{заг}} = \sum_{k=1}^g \sum_{j=1}^p \sum_{i=1}^q (x_{ijk} - \bar{x}_B)^2;$$

- б) *факторну суму* квадратів відхилень групових середніх від загального вибіркового середнього для фактора F :

$$S_{\text{факт}}^F = qg \sum_{j=1}^p (\bar{x}_{\text{гр. } j}^F - \bar{x}_B)^2;$$

- в) *факторну суму* квадратів відхилень групових середніх від загального вибіркового середнього для фактора G :

$$S_{\text{факт}}^G = qp \sum_{k=1}^g (\bar{x}_{\text{гр. } k}^G - \bar{x}_B)^2;$$

- г) *факторну суму* квадратів відхилень групових середніх від загального вибіркового середнього для факторів F і G :

$$S_{\text{факт}}^{FG} = q \sum_{k=1}^g \sum_{j=1}^p (\bar{x}_{\text{гр. } jk}^{FG} - \bar{x}_B)^2 - S_{\text{факт}}^F - S_{\text{факт}}^G;$$

д) залишкову суму:

$$S_{\text{зал}} = S_{\text{заг}} - S_{\text{факт}}^F - S_{\text{факт}}^G - S_{\text{факт}}^{FG}.$$

Поділивши факторні та залишкову суми на відповідну кількість ступенів вільності

$$\begin{aligned} k_{\text{факт}}^F &= p - 1; & k_{\text{факт}}^G &= g - 1; \\ k_{\text{факт}}^{FG} &= (p - 1)(g - 1); & k_{\text{зал}} &= pg(q - 1), \end{aligned}$$

знаходять факторні та залишкову дисперсії:

$$\begin{aligned} (s_{\text{факт}}^F)^2 &= \frac{S_{\text{факт}}^F}{p - 1}; & (s_{\text{факт}}^G)^2 &= \frac{S_{\text{факт}}^G}{g - 1}; \\ (s_{\text{факт}}^{FG})^2 &= \frac{S_{\text{факт}}^{FG}}{(p - 1)(g - 1)}; & s_{\text{зал}}^2 &= \frac{S_{\text{зал}}}{pg(q - 1)}. \end{aligned}$$

Нарешті, порівнюють факторні дисперсії із залишковою дисперсією за критерієм Фішера — Снедекора:

$$F_{\text{спост}}^F = \frac{(s_{\text{факт}}^F)^2}{s_{\text{зал}}^2}, \quad F_{\text{спост}}^G = \frac{(s_{\text{факт}}^G)^2}{s_{\text{зал}}^2}, \quad F_{\text{спост}}^{FG} = \frac{(s_{\text{факт}}^{FG})^2}{s_{\text{зал}}^2}.$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості α і кількості ступенів вільності (k_1 — кількість ступенів вільності, що відповідає більшій дисперсії, k_2 — кількість ступенів вільності, що відповідає меншій дисперсії) знаходять критичні точки $F_{\text{кр}}^F(\alpha; k_1, k_2)$, $F_{\text{кр}}^G(\alpha; k_1, k_2)$ і $F_{\text{кр}}^{FG}(\alpha; k_1, k_2)$.

Якщо $F_{\text{спост}}^F < F_{\text{кр}}^F$, відмінність групових середніх квадратичних відхилень для фактора F вважають несуттєвою. Якщо $F_{\text{спост}}^F > F_{\text{кр}}^F$, відмінність групових середніх квадратичних відхилень для фактора F вважають суттєвою.

Якщо $F_{\text{спост}}^G < F_{\text{кр}}^G$, відмінність групових середніх квадратичних відхилень для фактора G вважають несуттєвою. Якщо $F_{\text{спост}}^G > F_{\text{кр}}^G$,

відмінність групових середніх квадратичних відхилень для фактора G вважають суттєвою.

Якщо $F_{\text{спост}}^{FG} < F_{\text{кр}}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважають несуттєвою. Якщо $F_{\text{спост}}^{FG} > F_{\text{кр}}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважають суттєвою.

Зауваження 1. Двофакторний дисперсійний аналіз для незв'язних вибірок вимагає не менше двох рівнів кожного фактора й не менше двох спостережень на кожній парі рівнів, причому для всіх пар рівнів кількість спостережень має бути однаковою.

Зауваження 2. Результуюча ознака має бути нормально розподіленою в досліджуваній вибірці.

Зауваження 3. Якщо деяка факторна дисперсія виявиться меншою від залишкової, одразу можна робити висновок про те, що немає підстав відхиляти основну гіпотезу. Іншими словами, у такому випадку для даного фактора подальші обчислення зайві.

Зауваження 4. Якщо спостережувані значення x_{ij} кратні деякому числу k (наприклад, десяткові дроби кратні числу 10^{-d} , де d — кількість цифр після коми), доцільно перейти до нових значень:

$$y_{ij} = \frac{x_{ij}}{k}.$$

При цьому факторні та залишкова дисперсії зменшаться у k^2 разів, але їх відношення не зміниться.

Приклад 1. У результаті проведення досліджень над різними вибірками за двома факторами, кожний з яких має два рівні, отримано деякі дані.

Методом двофакторного дисперсійного аналізу при рівні значущості 0,01 перевірити:

- для фактора F основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- для фактора G основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градаціях фактора G , і навпаки.

Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в таблиці (табл. 30).

Таблиця 30

Рівні фактора G	Рівні фактора F								Загальна сума
	F_1				F_2				
G_1	4	3	5	4	2	4	3	2	$S_{G_1} = 27$
Сума	$S_{11} = 16$				$S_{21} = 11$				
G_2	4	5	6	4	4	5	5	5	$S_{G_2} = 38$
Сума	$S_{12} = 19$				$S_{22} = 19$				
Загальна сума	$S_{F_1} = 35$				$S_{F_2} = 30$				$S = 65$

Розв'язання. Для даної задачі значення параметрів q, p, g такі:

$$q = 4; \quad p = 2; \quad g = 2.$$

За допомогою даних табл. 30 знайдемо \bar{x}_B ; $\bar{x}_{гр. j}^F, j = 1, 2; \bar{x}_{гр. k}^G, k = 1, 2; \bar{x}_{гр. jk}^{FG}, j = 1, 2, k = 1, 2$:

$$\begin{aligned} \bar{x}_B &= \frac{1}{ppg} S = \frac{1}{4 \cdot 2 \cdot 2} 65 = 4,0625; \\ \bar{x}_{гр. 1}^F &= \frac{1}{qg} S_{F_1} = \frac{1}{4 \cdot 2} 35 = 4,375; \\ \bar{x}_{гр. 2}^F &= \frac{1}{qg} S_{F_2} = \frac{1}{4 \cdot 2} 30 = 3,75; \\ \bar{x}_{гр. 1}^G &= \frac{1}{qp} S_{G_1} = \frac{1}{4 \cdot 2} 27 = 3,375; \end{aligned}$$

$$\begin{aligned}\bar{x}_{\text{гр. } 2}^G &= \frac{1}{qp} S_{G_1} = \frac{1}{4 \cdot 2} 38 = 4,75; \\ \bar{x}_{\text{гр. } 11}^{FG} &= \frac{1}{q} S_{11} = \frac{1}{4} 16 = 4; \\ \bar{x}_{\text{гр. } 12}^{FG} &= \frac{1}{q} S_{12} = \frac{1}{4} 19 = 4,75; \\ \bar{x}_{\text{гр. } 21}^{FG} &= \frac{1}{q} S_{21} = \frac{1}{4} 11 = 2,75; \\ \bar{x}_{\text{гр. } 22}^{FG} &= \frac{1}{q} S_{22} = \frac{1}{4} 19 = 4,75.\end{aligned}$$

Обчислимо:

- а) загальну суму квадратів відхилень спостережуваних значень ознаки від загального вибіркового середнього:

$$\begin{aligned}S_{\text{заг}} &= \sum_{k=1}^g \sum_{j=1}^p \sum_{i=1}^q (x_{ijk} - \bar{x}_B)^2 = (4 - 4,0625)^2 + (3 - 4,0625)^2 + \\ &+ (5 - 4,0625)^2 + (4 - 4,0625)^2 + (2 - 4,0625)^2 + \\ &+ (4 - 4,0625)^2 + (3 - 4,0625)^2 + (2 - 4,0625)^2 + \\ &+ (4 - 4,0625)^2 + (5 - 4,0625)^2 + (6 - 4,0625)^2 + \\ &+ (4 - 4,0625)^2 + (4 - 4,0625)^2 + (5 - 4,0625)^2 + \\ &+ (5 - 4,0625)^2 + (5 - 4,0625)^2 = \\ &= 18,9375;\end{aligned}$$

- б) факторну суму квадратів відхилень групових середніх від загального вибіркового середнього для фактора F :

$$\begin{aligned}S_{\text{факт}}^F &= qg \sum_{j=1}^p (\bar{x}_{\text{гр. } j}^F - \bar{x}_B)^2 = \\ &= 4 \cdot 2 \cdot [(4,375 - 4,0625)^2 + (3,75 - 4,0625)^2] = \\ &= 1,5625;\end{aligned}$$

- в) факторну суму квадратів відхилень групових середніх від загального вибіркового середнього для фактора G :

$$\begin{aligned} S_{\text{факт}}^G &= qp \sum_{k=1}^g (\bar{x}_{\text{гр. } k}^G - \bar{x}_B)^2 = \\ &= 4 \cdot 2 \cdot [(3,375 - 4,0625)^2 + (4,75 - 4,0625)^2] = \\ &= 7,5625; \end{aligned}$$

- г) факторну суму квадратів відхилень групових середніх від загального вибіркового середнього для факторів F і G :

$$\begin{aligned} S_{\text{факт}}^{FG} &= q \sum_{k=1}^g \sum_{j=1}^p (\bar{x}_{\text{гр. } jk}^{FG} - \bar{x}_B)^2 - S_{\text{факт}}^F - S_{\text{факт}}^G = \\ &= 4 \cdot [(4 - 4,0625)^2 + (2,75 - 4,0625)^2 + (4,75 - 4,0625)^2 + \\ &\quad + (4,75 - 4,0625)^2] - 1,5625 - 7,5625 = \\ &= 1,5625; \end{aligned}$$

- д) залишкову суму:

$$\begin{aligned} S_{\text{зал}} &= S_{\text{заг}} - S_{\text{факт}}^F - S_{\text{факт}}^G - S_{\text{факт}}^{FG} = \\ &= 18,9375 - 1,5625 - 7,5625 - 1,5625 = \\ &= 8,25. \end{aligned}$$

Поділивши факторні та залишкову суми на відповідну кількість ступенів вільності

$$\begin{aligned} k_{\text{факт}}^F &= p - 1 = 2 - 1 = 1; \\ k_{\text{факт}}^G &= g - 1 = 2 - 1 = 1; \\ k_{\text{факт}}^{FG} &= (p - 1)(g - 1) = 1 \cdot 1 = 1; \\ k_{\text{зал}} &= pg(q - 1) = 2 \cdot 2 \cdot (4 - 1) = 12, \end{aligned}$$

знайдемо факторні та залишкову дисперсії:

$$\begin{aligned} (s_{\text{факт}}^F)^2 &= \frac{S_{\text{факт}}^F}{p-1} = \frac{1,5625}{1} = 1,5625; \\ (s_{\text{факт}}^G)^2 &= \frac{S_{\text{факт}}^G}{g-1} = \frac{7,5625}{1} = 7,5625; \\ (s_{\text{факт}}^{FG})^2 &= \frac{S_{\text{факт}}^{FG}}{(p-1)(g-1)} = \frac{1,5625}{1} = 1,5625; \\ s_{\text{зал}}^2 &= \frac{S_{\text{зал}}}{pg(q-1)} = \frac{8,25}{12} = 0,6875. \end{aligned}$$

Нарешті, порівняємо факторні дисперсії із залишковою дисперсією за критерієм Фішера — Снедекора:

$$\begin{aligned} F_{\text{спост}}^F &= \frac{(s_{\text{факт}}^F)^2}{s_{\text{зал}}^2} = \frac{1,5625}{0,6875} \approx 2,27; \\ F_{\text{спост}}^G &= \frac{(s_{\text{факт}}^G)^2}{s_{\text{зал}}^2} = \frac{7,5625}{0,6875} = 11; \\ F_{\text{спост}}^{FG} &= \frac{(s_{\text{факт}}^{FG})^2}{s_{\text{зал}}^2} = \frac{1,5625}{0,6875} \approx 2,27. \end{aligned}$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості $\alpha = 0,01$ і кількості ступенів вільності знайдемо критичні точки:

$$\begin{aligned} F_{\text{кр}}^F(0,01; 1; 12) &= 9,33; \\ F_{\text{кр}}^G(0,01; 1; 12) &= 9,33; \\ F_{\text{кр}}^{FG}(0,01; 1; 12) &= 9,33. \end{aligned}$$

Оскільки $F_{\text{спост}}^F < F_{\text{кр}}^F$, відмінність групових середніх квадратичних відхилень для фактора F вважаємо несуттєвою.

Оскільки $F_{\text{спост}}^G > F_{\text{кр}}^G$, відмінність групових середніх квадратичних відхилень для фактора G вважаємо суттєвою.

Оскільки $F_{\text{спост}}^{FG} < F_{\text{кр}}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважаємо несуттєвою. \square

16.4. Двофакторний дисперсійний аналіз для зв'язних вибірок

Нехай на кількісну нормально розподілену ознаку X діють два фактори F і G , які мають відповідно p і g сталих рівнів F_1, F_2, \dots, F_p і G_1, G_2, \dots, G_g . На кожній парі $(F_i; G_j)$ рівнів здійснено по q випробувань **над однією й тією самою вибіркою**. Результати спостережень — числа $(x_{ijk}, \text{ де } i — \text{ номер випробування, } i = 1, 2, \dots, q, j — \text{ номер рівня фактора } F, j = 1, 2, \dots, p, k — \text{ номер рівня фактора } G, k = 1, 2, \dots, g) — \text{ записують у вигляді таблиці (табл. 31).$

Уведемо позначення, які будуть потрібні для подальших обчислень (деякі з них наведено в табл. 31):

$$S_{jk} = \sum_{i=1}^q x_{ijk}, \quad j = 1, 2, \dots, p, \quad k = 1, 2, \dots, g;$$

$$S_{iF_j} = \sum_{k=1}^g x_{ijk}, \quad i = 1, 2, \dots, q, \quad j = 1, 2, \dots, p;$$

$$S_{iG_k} = \sum_{j=1}^p x_{ijk}, \quad i = 1, 2, \dots, q, \quad k = 1, 2, \dots, g;$$

$$S_i = \sum_{j=1}^p S_{iF_j} = \sum_{k=1}^g S_{iG_k}, \quad i = 1, 2, \dots, q;$$

$$S_{F_j} = \sum_{i=1}^q S_{iF_j} = \sum_{k=1}^g S_{jk}, \quad j = 1, 2, \dots, p;$$

$$S_{G_k} = \sum_{i=1}^q S_{iG_k} = \sum_{j=1}^p S_{jk}, \quad k = 1, 2, \dots, g;$$

Розділ 16. Елементи дисперсійного аналізу

$$S = \sum_{i=1}^q S_i = \sum_{j=1}^p S_{F_j} = \sum_{k=1}^g S_{G_k}.$$

Таблиця 31

Рівні фактора F	Рівні фактора G	Номер елемента вибірки i				Загальна сума
		1	2	...	q	
F_1	G_1	x_{111}	x_{211}	...	x_{q11}	S_{11}
	G_2	x_{112}	x_{212}	...	x_{q12}	S_{12}

	G_g	x_{11g}	x_{21g}	...	x_{q1g}	S_{1g}
	Індивідуальні суми по F_1	S_{1F_1}	S_{2F_1}	...	S_{qF_1}	S_{F_1}
F_2	G_1	x_{121}	x_{221}	...	x_{q21}	S_{21}
	G_2	x_{122}	x_{222}	...	x_{q22}	S_{22}

	G_g	x_{12g}	x_{22g}	...	x_{q2g}	S_{2g}
	Індивідуальні суми по F_2	S_{1F_2}	S_{2F_2}	...	S_{qF_2}	S_{F_2}
...
F_p	G_1	x_{1p1}	x_{2p1}	...	x_{qp1}	S_{p1}
	G_2	x_{1p2}	x_{2p2}	...	x_{qp2}	S_{p2}

	G_g	x_{1pg}	x_{2pg}	...	x_{qpg}	S_{pg}
	Індивідуальні суми по F_p	S_{1F_p}	S_{2F_p}	...	S_{qF_p}	S_{F_p}
Індивідуальні суми		S_1	S_2	...	S_q	S

У цьому разі виникають чотири задачі:

- 1) для фактора F на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- 2) для фактора G на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- 3) на рівні значущості α перевірити основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градациях фактора G , і навпаки;
- 4) на рівні значущості α перевірити основну гіпотезу про рівність індивідуальних середніх за умови, що індивідуальні генеральні дисперсії хоч і невідомі, але однакові.

Щоб розв'язати ці задачі, знаходять такі суми квадратів відхилень:

$$\begin{aligned}
 S_{\text{заг}} &= \sum_{k=1}^g \sum_{j=1}^p \sum_{i=1}^q x_{ijk}^2 - \frac{1}{qpg} S^2; \\
 S_{\text{інд}} &= \frac{1}{pg} \sum_{i=1}^q S_i^2 - \frac{1}{qpg} S^2; \\
 S_{\text{факт}}^F &= \frac{1}{qg} \sum_{j=1}^p S_{F_j}^2 - \frac{1}{qpg} S^2; \\
 S_{\text{факт}}^G &= \frac{1}{qp} \sum_{k=1}^g S_{G_k}^2 - \frac{1}{qpg} S^2; \\
 S_{\text{інд}}^F &= \frac{1}{g} \sum_{j=1}^p \sum_{i=1}^q S_{iF_j}^2 - \frac{1}{qpg} S^2 - S_{\text{інд}} - S_{\text{факт}}^F; \\
 S_{\text{інд}}^G &= \frac{1}{p} \sum_{k=1}^g \sum_{i=1}^q S_{iG_k}^2 - \frac{1}{qpg} S^2 - S_{\text{інд}} - S_{\text{факт}}^G; \\
 S_{\text{факт}}^{FG} &= \frac{1}{q} \sum_{k=1}^g \sum_{j=1}^p S_{jk}^2 - \frac{1}{qpg} S^2 - S_{\text{факт}}^F - S_{\text{факт}}^G; \\
 S_{\text{інд}}^{FG} &= S_{\text{заг}} - S_{\text{інд}} - S_{\text{факт}}^F - S_{\text{факт}}^G - S_{\text{інд}}^F - S_{\text{інд}}^G - S_{\text{факт}}^{FG}.
 \end{aligned}$$

Поділивши обчислені суми на відповідну кількість ступенів вільності

$$\begin{aligned} k_{\text{інд}} &= q - 1; & k_{\text{інд}}^F &= (q - 1)(p - 1); \\ k_{\text{факт}}^F &= p - 1; & k_{\text{інд}}^G &= (q - 1)(g - 1); \\ k_{\text{факт}}^G &= g - 1; & k_{\text{факт}}^{FG} &= (p - 1)(g - 1); \\ k_{\text{заг}} &= qpg - 1; & k_{\text{інд}}^{FG} &= (q - 1)(p - 1)(g - 1), \end{aligned}$$

знаходять такі дисперсії:

$$\begin{aligned} s_{\text{інд}}^2 &= \frac{S_{\text{інд}}}{q - 1}; & (s_{\text{інд}}^F)^2 &= \frac{S_{\text{інд}}^F}{(q - 1)(p - 1)}; \\ (s_{\text{факт}}^F)^2 &= \frac{S_{\text{факт}}^F}{p - 1}; & (s_{\text{інд}}^G)^2 &= \frac{S_{\text{інд}}^G}{(q - 1)(g - 1)}; \\ (s_{\text{факт}}^G)^2 &= \frac{S_{\text{факт}}^G}{g - 1}; & (s_{\text{факт}}^{FG})^2 &= \frac{S_{\text{факт}}^{FG}}{(p - 1)(g - 1)}; \\ s_{\text{заг}}^2 &= \frac{S_{\text{заг}}}{qpg - 1}; & (s_{\text{інд}}^{FG})^2 &= \frac{S_{\text{інд}}^{FG}}{(q - 1)(p - 1)(g - 1)}. \end{aligned}$$

Нарешті, порівнюють дисперсії за *критерієм Фішера — Снедекора*:

$$\begin{aligned} F_{\text{спост}}^F &= \frac{(s_{\text{факт}}^F)^2}{(s_{\text{інд}}^F)^2}; & F_{\text{спост}}^G &= \frac{(s_{\text{факт}}^G)^2}{(s_{\text{інд}}^G)^2}; \\ F_{\text{спост}}^{FG} &= \frac{(s_{\text{факт}}^{FG})^2}{(s_{\text{інд}}^{FG})^2}; & F_{\text{спост}}^{\text{інд}} &= \frac{s_{\text{інд}}^2}{(s_{\text{інд}}^{FG})^2}. \end{aligned}$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості α і кількості ступенів вільності знаходять критичні точки:

$$\begin{aligned} F_{\text{кр}}^F(\alpha; k_{\text{факт}}^F, k_{\text{інд}}^F); & & F_{\text{кр}}^G(\alpha; k_{\text{факт}}^G, k_{\text{інд}}^G); \\ F_{\text{кр}}^{FG}(\alpha; k_{\text{факт}}^{FG}, k_{\text{інд}}^{FG}); & & F_{\text{кр}}^{\text{інд}}(\alpha; k_{\text{інд}}, k_{\text{інд}}^{FG}). \end{aligned}$$

Якщо $F_{\text{спост}}^F < F_{\text{кр}}^F$, відмінність групових середніх квадратичних відхилень для фактора F вважають несуттєвою. Якщо $F_{\text{спост}}^F > F_{\text{кр}}^F$,

відмінність групових середніх квадратичних відхилень для фактора F вважають суттєвою.

Якщо $F_{\text{спост}}^G < F_{\text{кр}}^G$, відмінність групових середніх квадратичних відхилень для фактора G вважають несуттєвою. Якщо $F_{\text{спост}}^G > F_{\text{кр}}^G$, відмінність групових середніх квадратичних відхилень для фактора G вважають суттєвою.

Якщо $F_{\text{спост}}^{FG} < F_{\text{кр}}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважають несуттєвою. Якщо $F_{\text{спост}}^{FG} > F_{\text{кр}}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважають суттєвою.

Якщо $F_{\text{спост}}^{\text{інд}} < F_{\text{кр}}^{\text{інд}}$, відмінність групових індивідуальних середніх квадратичних відхилень при різних градаціях факторів F і G вважають несуттєвою. Якщо $F_{\text{спост}}^{\text{інд}} > F_{\text{кр}}^{\text{інд}}$, відмінність групових індивідуальних середніх квадратичних відхилень при різних градаціях факторів F і G вважають суттєвою.

Зауваження 1. Двофакторний дисперсійний аналіз для зв'язних вибірок вимагає не менше двох рівнів кожного фактора й не менше двох елементів вибірки, ознаки яких вимірюються на всіх рівнях усіх пар факторів.

Зауваження 2. Результуюча ознака має бути нормально розподіленою в досліджуваній вибірці.

Зауваження 3. Якщо спостережувані значення x_{ij} кратні деякому числу k (наприклад, десяткові дробі кратні числу 10^{-d} , де d — кількість цифр після коми), доцільно перейти до нових значень:

$$y_{ij} = \frac{x_{ij}}{k}.$$

При цьому факторні та залишкова дисперсії зменшаться у k^2 разів, але їх відношення не зміниться.

Приклад 1. У результаті проведення досліджень над однією й тією самою вибіркою за двома факторами, кожний з яких має два рівні, отримано деякі дані.

Методом двофакторного дисперсійного аналізу при рівні значущості 0,01 перевірити:

- для фактора F основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- для фактора G основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;
- основну гіпотезу про рівність групових середніх за умов, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градаціях фактора G , і навпаки;
- основну гіпотезу про рівність індивідуальних середніх за умови, що індивідуальні генеральні дисперсії хоч і невідомі, але однакові.

Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в таблиці (табл. 32).

Таблиця 32

Рівні фактора F	Рівні фактора G	Номер елемента вибірки i				Загальна сума
		1	2	3	4	
F_1	G_1	17	13	16	13	$S_{11} = 59$
	G_2	11	12	16	10	$S_{12} = 49$
	Індивідуальні суми по F_1	$S_{1F_1} = 28$	$S_{2F_1} = 25$	$S_{3F_1} = 32$	$S_{4F_1} = 23$	$S_{F_1} = 108$
F_2	G_1	19	17	18	14	$S_{21} = 68$
	G_2	14	11	10	12	$S_{22} = 47$
	Індивідуальні суми по F_2	$S_{1F_2} = 33$	$S_{2F_2} = 28$	$S_{3F_2} = 28$	$S_{4F_2} = 26$	$S_{F_2} = 115$
Індивідуальні суми		$S_1 = 61$	$S_2 = 53$	$S_3 = 60$	$S_4 = 49$	$S = 223$

Розв'язання. Для даної задачі значення параметрів q, p, g такі:

$$q = 4; \quad p = 2; \quad g = 2.$$

Знайдемо значення сум S_{iG_k} , $i = 1, 2, 3, 4$, $k = 1, 2$; S_{G_k} , $k = 1, 2$, не наведених в табл. 32:

$$S_{1G_1} = \sum_{j=1}^2 x_{1j1} = 17 + 19 = 36;$$

$$S_{2G_1} = \sum_{j=1}^2 x_{2j1} = 13 + 17 = 30;$$

$$S_{3G_1} = \sum_{j=1}^2 x_{3j1} = 16 + 18 = 34;$$

$$S_{4G_1} = \sum_{j=1}^2 x_{4j1} = 13 + 14 = 27;$$

$$S_{1G_2} = \sum_{j=1}^2 x_{1j2} = 11 + 14 = 25;$$

$$S_{2G_2} = \sum_{j=1}^2 x_{2j2} = 12 + 11 = 23;$$

$$S_{3G_2} = \sum_{j=1}^2 x_{3j2} = 16 + 10 = 26;$$

$$S_{4G_2} = \sum_{j=1}^2 x_{4j2} = 10 + 12 = 22;$$

$$S_{G_1} = \sum_{i=1}^4 S_{iG_1} = 36 + 30 + 34 + 27 = 127;$$

$$S_{G_2} = \sum_{i=1}^4 S_{iG_2} = 25 + 23 + 26 + 22 = 96.$$

Обчислимо такі суми квадратів відхилень:

$$\begin{aligned}
 S_{\text{зар}} &= \sum_{k=1}^g \sum_{j=1}^p \sum_{i=1}^q x_{ijk}^2 - \frac{1}{qpg} S^2 = 17^2 + 13^2 + 16^2 + 13^2 + 11^2 + \\
 &\quad + 12^2 + 16^2 + 10^2 + 19^2 + 17^2 + 18^2 + 14^2 + 14^2 + 11^2 + \\
 &\quad + 10^2 + 12^2 - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 = \\
 &= 126,9375;
 \end{aligned}$$

$$\begin{aligned}
 S_{\text{інд}} &= \frac{1}{pg} \sum_{i=1}^q S_i^2 - \frac{1}{qpg} S^2 = \\
 &= \frac{1}{2 \cdot 2} \cdot (61^2 + 53^2 + 60^2 + 49^2) - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 = \\
 &= 24,6875;
 \end{aligned}$$

$$\begin{aligned}
 S_{\text{факт}}^F &= \frac{1}{qg} \sum_{j=1}^p S_{F_j}^2 - \frac{1}{qpg} S^2 = \frac{1}{4 \cdot 2} \cdot (108^2 + 115^2) - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 = \\
 &= 3,0625;
 \end{aligned}$$

$$\begin{aligned}
 S_{\text{факт}}^G &= \frac{1}{qp} \sum_{k=1}^g S_{G_k}^2 - \frac{1}{qpg} S^2 = \frac{1}{4 \cdot 2} \cdot (127^2 + 96^2) - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 = \\
 &= 60,0625;
 \end{aligned}$$

$$\begin{aligned}
 S_{\text{інд}}^F &= \frac{1}{g} \sum_{j=1}^p \sum_{i=1}^q S_{iF_j}^2 - \frac{1}{qpg} S^2 - S_{\text{інд}} - S_{\text{факт}}^F = \\
 &= \frac{1}{2} (28^2 + 25^2 + 32^2 + 23^2 + 33^2 + 28^2 + 28^2 + 26^2) - \\
 &\quad - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 - 24,6875 - 3,0625 = \\
 &= 11,6875;
 \end{aligned}$$

$$\begin{aligned}
 S_{\text{інд}}^G &= \frac{1}{p} \sum_{k=1}^g \sum_{i=1}^q S_{iG_k}^2 - \frac{1}{qpg} S^2 - S_{\text{інд}} - S_{\text{факт}}^G = \\
 &= \frac{1}{2} \cdot (36^2 + 30^2 + 34^2 + 27^2 + 25^2 + 23^2 + 26^2 + 22^2) -
 \end{aligned}$$

$$\begin{aligned}
 & - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 - 24,6875 - 60,0625 = \\
 & = 4,6875; \\
 S_{\text{факт}}^{FG} &= \frac{1}{q} \sum_{k=1}^g \sum_{j=1}^p S_{jk}^2 - \frac{1}{qpg} S^2 - S_{\text{факт}}^F - S_{\text{факт}}^G = \\
 &= \frac{1}{4} \cdot (59^2 + 49^2 + 68^2 + 47^2) - \\
 & \quad - \frac{1}{4 \cdot 2 \cdot 2} \cdot 223^2 - 24,6875 - 3,0625 - 60,0625 = \\
 &= 7,5625; \\
 S_{\text{інд}}^{FG} &= S_{\text{заг}} - S_{\text{інд}} - S_{\text{факт}}^F - S_{\text{факт}}^G - S_{\text{інд}}^F - S_{\text{інд}}^G - S_{\text{факт}}^{FG} = \\
 &= 126,9375 - 24,6875 - 3,0625 - 60,0625 - \\
 & \quad - 11,6875 - 4,6875 - 7,5625 = \\
 &= 15,1875.
 \end{aligned}$$

Поділивши обчислені суми на відповідну кількість ступенів вільності

$$\begin{aligned}
 k_{\text{інд}} &= q - 1 = 4 - 1 = 3; \\
 k_{\text{інд}}^F &= (q - 1)(p - 1) = (4 - 1)(2 - 1) = 3; \\
 k_{\text{факт}}^F &= p - 1 = 2 - 1 = 1; \\
 k_{\text{інд}}^G &= (q - 1)(g - 1) = (4 - 1)(2 - 1) = 3; \\
 k_{\text{факт}}^G &= g - 1 = 2 - 1 = 1; \\
 k_{\text{факт}}^{FG} &= (p - 1)(g - 1) = (2 - 1)(2 - 1) = 1; \\
 k_{\text{заг}} &= qpg - 1 = 4 \cdot 2 \cdot 2 - 1 = 15; \\
 k_{\text{інд}}^{FG} &= (q - 1)(p - 1)(g - 1) = (4 - 1)(2 - 1)(2 - 1) = 3,
 \end{aligned}$$

знайдемо такі дисперсії:

$$s_{\text{інд}}^2 = \frac{S_{\text{інд}}}{q - 1} = \frac{24,6875}{3} \approx 8,23;$$

$$\begin{aligned} (s_{\text{інд}}^F)^2 &= \frac{S_{\text{інд}}^F}{(q-1)(p-1)} = \frac{11,6875}{3} \approx 3,90; \\ (s_{\text{факт}}^F)^2 &= \frac{S_{\text{факт}}^F}{p-1} = \frac{3,0625}{1} = 3,0625; \\ (s_{\text{інд}}^G)^2 &= \frac{S_{\text{інд}}^G}{(q-1)(g-1)} = \frac{4,6875}{3} \approx 1,5625; \\ (s_{\text{факт}}^G)^2 &= \frac{S_{\text{факт}}^G}{g-1} = \frac{60,0625}{1} = 60,0625; \\ (s_{\text{факт}}^{FG})^2 &= \frac{S_{\text{факт}}^{FG}}{(p-1)(g-1)} = \frac{7,5625}{1} \approx 7,5625; \\ s_{\text{заг}}^2 &= \frac{S_{\text{заг}}}{qpg-1} = \frac{126,9375}{15} = 8,4625; \\ (s_{\text{інд}}^{FG})^2 &= \frac{S_{\text{інд}}^{FG}}{(q-1)(p-1)(g-1)} = \frac{15,1875}{3} = 5,0625. \end{aligned}$$

Нарешті, порівняємо дисперсії за критерієм Фішера — Снедекора:

$$\begin{aligned} F_{\text{спост}}^F &= \frac{(s_{\text{факт}}^F)^2}{(s_{\text{інд}}^F)^2} = \frac{3,0625}{3,90} \approx 0,79; \\ F_{\text{спост}}^G &= \frac{(s_{\text{факт}}^G)^2}{(s_{\text{інд}}^G)^2} = \frac{60,0625}{1,5625} = 38,44; \\ F_{\text{спост}}^{FG} &= \frac{(s_{\text{факт}}^{FG})^2}{(s_{\text{інд}}^{FG})^2} = \frac{7,5625}{5,0625} \approx 1,49; \\ F_{\text{спост}}^{\text{інд}} &= \frac{s_{\text{інд}}^2}{(s_{\text{інд}}^{FG})^2} = \frac{8,23}{5,0625} \approx 1,63. \end{aligned}$$

Далі за таблицею критичних точок розподілу Фішера — Снедекора (дод. 8) при заданому рівні значущості $\alpha = 0,01$ і кількості ступенів вільності знайдемо критичні точки:

$$F_{\text{кр}}^F(0,01; 1; 3) = 34,12;$$

$$F_{кр}^G(0,01; 1; 3) = 34,12;$$

$$F_{кр}^{FG}(0,01; 1; 3) = 34,12;$$

$$F_{кр}^{інд}(0,01; 3; 3) = 29,46.$$

Оскільки $F_{спост}^F < F_{кр}^F$, відмінність групових середніх квадратичних відхилень для фактора F вважаємо несуттєвою.

Оскільки $F_{спост}^G > F_{кр}^G$, відмінність групових середніх квадратичних відхилень для фактора G вважаємо суттєвою.

Оскільки $F_{спост}^{FG} < F_{кр}^{FG}$, відмінність групових середніх квадратичних відхилень для фактора F при різних градаціях фактора G вважаємо несуттєвою.

Оскільки $F_{спост}^{інд} < F_{кр}^{інд}$, відмінність групових індивідуальних середніх квадратичних відхилень при різних градаціях факторів F і G вважаємо несуттєвою. \square

Задачі до розділу 16

Задача 1. Проведено по сім випробувань на кожному з чотирьох рівнів фактора над різними вибірками. Методом дисперсійного аналізу при рівні значущості 0,01 перевірити основну гіпотезу про рівність групових середніх. Припускається, що вибірки отримані з нормальних сукупностей з однаковими дисперсіями. Результати випробувань подано в табл. 33.

Відповідь. Основна гіпотеза не відхиляється.

Задача 2. Проведено випробування на кожному з чотирьох рівнів фактора над однією й тією самою вибіркою з шести елементів. Методом дисперсійного аналізу при рівні значущості 0,05 перевірити основну гіпотезу про рівність групових середніх та основну гіпотезу про рівність індивідуальних середніх. Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в табл. 34.

Відповідь. Основна гіпотеза про рівність групових середніх не відхиляється. Основна гіпотеза про рівність індивідуальних середніх відхиляється.

Розділ 16. Елементи дисперсійного аналізу

Таблиця 33

Номер випробування i	Рівень фактора			
	F_1	F_2	F_3	F_4
1	12	11	10	13
2	15	12	13	15
3	16	13	14	17
4	19	15	16	19
5	20	16	18	22
6	21	18	20	23
7	23	20	21	24
Групове середнє $\bar{x}_{гр}$	18	15	16	19

Таблиця 34

Номер випробування i	Рівень фактора				Індивідуальні середні \bar{x}_i
	F_1	F_2	F_3	F_4	
1	22	23	22	21	22
2	25	24	22	24	23,75
3	19	20	22	25	21,5
4	14	17	20	24	18,75
5	20	21	21	21	20,75
6	14	15	19	23	17,75
Групове середнє $\bar{x}_{гр}$	19	20	21	23	

Задача 3. У результаті проведення досліджень над різними вибірками за двома факторами, кожний з яких має два рівні, отримано деякі дані.

Методом двофакторного дисперсійного аналізу при рівні значущості 0,05 перевірити:

- для фактора F основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- для фактора G основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градаціях фактора G , і навпаки.

Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в таблиці (табл. 35).

Таблиця 35

Рівні фактора G	Рівні фактора F								Загальна сума
	F_1				F_2				
G_1	7	5	6	4	5	8	7	6	$S_{G_1} = 48$
Сума	$S_{11} = 22$				$S_{21} = 26$				
G_2	4	3	5	4	7	5	6	5	$S_{G_2} = 39$
Сума	$S_{12} = 16$				$S_{22} = 23$				
Загальна сума	$S_{F_1} = 38$				$S_{F_2} = 49$				$S = 87$

Відповідь. Основна гіпотеза про рівність групових середніх для фактора F відхиляється. В інших випадках основна гіпотеза про рівність групових середніх не відхиляється.

Задача 4. У результаті проведення досліджень над однією й тією самою вибіркою за двома факторами, кожний з яких має два рівні, отримано деякі дані.

Методом двофакторного дисперсійного аналізу при рівні значущості 0,05 перевірити:

- для фактора F основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- для фактора G основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові;

- основну гіпотезу про рівність групових середніх за умови, що групові генеральні дисперсії хоч і невідомі, але однакові для фактора F при різних градаціях фактора G , і навпаки;
- основну гіпотезу про рівність індивідуальних середніх за умови, що індивідуальні генеральні дисперсії хоч і невідомі, але однакові.

Припускається, що результуюча ознака нормально розподілена в досліджуваній вибірці. Результати випробувань подано в таблиці (табл. 36).

Таблиця 36

Рівні фактора F	Рівні фактора G	Номер елемента вибірки i				Загальна сума
		1	2	3	4	
F_1	G_1	37	42	31	25	$S_{11} = 135$
	G_2	32	39	28	19	$S_{12} = 118$
	Індивідуальні суми по F_1	$S_{1F_1} = 69$	$S_{2F_1} = 81$	$S_{3F_1} = 59$	$S_{4F_1} = 44$	$S_{F_1} = 253$
F_2	G_1	45	47	35	26	$S_{21} = 153$
	G_2	33	35	29	22	$S_{22} = 119$
	Індивідуальні суми по F_2	$S_{1F_2} = 78$	$S_{2F_2} = 82$	$S_{3F_2} = 64$	$S_{4F_2} = 48$	$S_{F_2} = 272$
Індивідуальні суми		$S_1 = 147$	$S_2 = 163$	$S_3 = 123$	$S_4 = 92$	$S = 525$

Відповідь. Основна гіпотеза про рівність групових середніх для фактора F і фактора F при різних градаціях фактора G не відхиляється. В інших випадках основна гіпотеза відхиляється.

Додатки

Додаток 1

Таблиця значень функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,22	0,0871	0,44	0,1700	0,66	0,2454	0,88	0,3106
0,01	0,0040	0,23	0,0910	0,45	0,1736	0,67	0,2486	0,89	0,3133
0,02	0,0080	0,24	0,0948	0,46	0,1772	0,68	0,2517	0,90	0,3159
0,03	0,0120	0,25	0,0987	0,47	0,1808	0,69	0,2549	0,91	0,3186
0,04	0,0160	0,26	0,1026	0,48	0,1844	0,70	0,2580	0,92	0,3212
0,05	0,0199	0,27	0,1064	0,49	0,1879	0,71	0,2611	0,93	0,3238
0,06	0,0239	0,28	0,1103	0,50	0,1915	0,72	0,2642	0,94	0,3264
0,07	0,0279	0,29	0,1141	0,51	0,1950	0,73	0,2673	0,95	0,3289
0,08	0,0319	0,30	0,1179	0,52	0,1985	0,74	0,2704	0,96	0,3315
0,09	0,0359	0,31	0,1217	0,53	0,2019	0,75	0,2734	0,97	0,3340
0,10	0,0398	0,32	0,1255	0,54	0,2054	0,76	0,2764	0,98	0,3365
0,11	0,0438	0,33	0,1293	0,55	0,2088	0,77	0,2794	0,99	0,3389
0,12	0,0478	0,34	0,1331	0,56	0,2123	0,78	0,2823	1,00	0,3413
0,13	0,0517	0,35	0,1368	0,57	0,2157	0,79	0,2852	1,01	0,3438
0,14	0,0557	0,36	0,1406	0,58	0,2190	0,80	0,2881	1,02	0,3461
0,15	0,0596	0,37	0,1443	0,59	0,2224	0,81	0,2910	1,03	0,3485
0,16	0,0636	0,38	0,1480	0,60	0,2257	0,82	0,2939	1,04	0,3508
0,17	0,0675	0,39	0,1517	0,61	0,2291	0,83	0,2967	1,05	0,3531
0,18	0,0714	0,40	0,1554	0,62	0,2324	0,84	0,2995	1,06	0,3554
0,19	0,0753	0,41	0,1591	0,63	0,2357	0,85	0,3023	1,07	0,3577
0,20	0,0793	0,42	0,1628	0,64	0,2389	0,86	0,3051	1,08	0,3599
0,21	0,0832	0,43	0,1664	0,65	0,2422	0,87	0,3078	1,09	0,3621

Закінчення дод. 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,10	0,3643	1,47	0,4292	1,84	0,4671	2,21	0,4864	2,66	0,4961
1,11	0,3665	1,48	0,4306	1,85	0,4678	2,22	0,4868	2,68	0,4963
1,12	0,3686	1,49	0,4319	1,86	0,4686	2,23	0,4871	2,70	0,4965
1,13	0,3708	1,50	0,4332	1,87	0,4693	2,24	0,4875	2,72	0,4967
1,14	0,3729	1,51	0,4345	1,88	0,4699	2,25	0,4878	2,74	0,4969
1,15	0,3749	1,52	0,4357	1,89	0,4706	2,26	0,4881	2,76	0,4971
1,16	0,3770	1,53	0,4370	1,90	0,4713	2,27	0,4884	2,78	0,4973
1,17	0,3790	1,54	0,4382	1,91	0,4719	2,28	0,4887	2,80	0,4974
1,18	0,3810	1,55	0,4394	1,92	0,4726	2,29	0,4890	2,82	0,4976
1,19	0,3830	1,56	0,4406	1,93	0,4732	2,30	0,4893	2,84	0,4977
1,20	0,3849	1,57	0,4418	1,94	0,4738	2,31	0,4896	2,86	0,4979
1,21	0,3869	1,58	0,4429	1,95	0,4744	2,32	0,4898	2,88	0,4980
1,22	0,3888	1,59	0,4441	1,96	0,4750	2,33	0,4901	2,90	0,4981
1,23	0,3907	1,60	0,4452	1,97	0,4756	2,34	0,4904	2,92	0,4982
1,24	0,3925	1,61	0,4463	1,98	0,4761	2,35	0,4906	2,94	0,4984
1,25	0,3944	1,62	0,4474	1,99	0,4767	2,36	0,4909	2,96	0,4985
1,26	0,3962	1,63	0,4484	2,00	0,4772	2,37	0,4911	2,98	0,4986
1,27	0,3980	1,64	0,4495	2,01	0,4778	2,38	0,4913	3,00	0,4987
1,28	0,3997	1,65	0,4505	2,02	0,4783	2,39	0,4916	3,05	0,4989
1,29	0,4015	1,66	0,4515	2,03	0,4788	2,40	0,4918	3,10	0,49903
1,30	0,4032	1,67	0,4525	2,04	0,4793	2,41	0,4920	3,15	0,49918
1,31	0,4049	1,68	0,4535	2,05	0,4798	2,42	0,4922	3,20	0,49931
1,32	0,4066	1,69	0,4545	2,06	0,4803	2,43	0,4925	3,25	0,49942
1,33	0,4082	1,70	0,4554	2,07	0,4808	2,44	0,4927	3,30	0,49952
1,34	0,4099	1,71	0,4564	2,08	0,4812	2,45	0,4929	3,35	0,49960
1,35	0,4115	1,72	0,4573	2,09	0,4817	2,46	0,4931	3,40	0,49966
1,36	0,4131	1,73	0,4582	2,10	0,4821	2,47	0,4932	3,50	0,49977
1,37	0,4147	1,74	0,4591	2,11	0,4826	2,48	0,4934	3,60	0,49984
1,38	0,4162	1,75	0,4599	2,12	0,4830	2,49	0,4936	3,70	0,49989
1,39	0,4177	1,76	0,4608	2,13	0,4834	2,50	0,4938	3,80	0,499928
1,40	0,4192	1,77	0,4616	2,14	0,4838	2,52	0,4941	3,90	0,499952
1,41	0,4207	1,78	0,4625	2,15	0,4842	2,54	0,4945	4,00	0,499968
1,42	0,4222	1,79	0,4633	2,16	0,4846	2,56	0,4948	4,20	0,499987
1,43	0,4236	1,80	0,4641	2,17	0,4850	2,58	0,4951	4,40	0,4999946
1,44	0,4251	1,81	0,4649	2,18	0,4854	2,60	0,4953	4,60	0,4999979
1,45	0,4265	1,82	0,4656	2,19	0,4857	2,62	0,4956	4,80	0,4999992
1,46	0,4279	1,83	0,4664	2,20	0,4861	2,64	0,4959	5,00	0,4999997

Таблиця значень функції $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	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	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	989	973	957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013

Закінчення дод. 2

x	0	1	2	3	4	5	6	7	8	9
3,4	0,0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
4,0	0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
4,1	0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
4,2	0001	0001	0001	0001	0000	0000	0000	0000	0000	0000

Додаток 3

Таблиця значень функції Пуассона $P_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}$

$m \backslash \lambda$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
0	0,9048	8187	7408	6703	6065	5488	4966	4493	4066	3679
1	0905	1637	2222	2681	3033	3293	3476	3595	3659	3679
2	0045	0164	0333	0536	0758	0988	1217	1438	1647	1839
3	0002	0011	0033	0072	0126	0198	0284	0383	0494	0613
4	0000	0001	0003	0007	0016	0030	0050	0077	0111	0153
5	0000	0000	0000	0001	0002	0004	0007	0012	0020	0031
6	0000	0000	0000	0000	0000	0000	0001	0002	0003	0005

$m \backslash \lambda$	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2
0	0,3329	3012	2725	2466	2231	2019	1827	1653	1496	1353
1	3662	3614	3543	3452	3347	3230	3106	2975	2842	2707
2	2014	2169	2303	2417	2510	2584	2640	2678	2700	2707
3	0738	0867	0998	1128	1255	1378	1496	1607	1710	1804
4	0203	0260	0324	0395	0471	0551	0636	0723	0812	0902
5	0045	0062	0084	0111	0141	0176	0216	0260	0309	0361
6	0008	0012	0018	0026	0035	0047	0061	0078	0098	0120
7	0001	0002	0003	0005	0008	0011	0015	0020	0027	0034
8	0000	0000	0001	0001	0001	0002	0003	0005	0006	0009
9	0000	0000	0000	0000	0000	0000	0001	0001	0001	0002

Додатки

Закінчення дод. 3

$m \backslash \lambda$	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3
0	0,1225	1108	1003	0907	0821	0743	0672	0608	0550	0498
1	2572	2438	2306	2177	2052	1931	1815	1703	1596	1494
2	2700	2681	2652	2613	2565	2510	2450	2384	2314	2240
3	1890	1966	2033	2090	2138	2176	2205	2225	2237	2240
4	0992	1082	1169	1254	1336	1414	1488	1557	1622	1680
5	0417	0476	0538	0602	0668	0735	0804	0872	0940	1008
6	0146	0174	0206	0241	0278	0319	0362	0407	0455	0504
7	0044	0055	0068	0083	0099	0118	0139	0163	0188	0216
8	0011	0015	0019	0025	0031	0038	0047	0057	0068	0081
9	0003	0004	0005	0007	0009	0011	0014	0018	0022	0027
10	0001	0001	0001	0002	0002	0003	0004	0005	0006	0008
11	0000	0000	0000	0000	0000	0001	0001	0001	0002	0002

$m \backslash \lambda$	3	3,5	4	4,5	5	6	7	8	9	10
0	0,0498	0302	0183	0111	0067	0025	0009	0003	0001	0000
1	1494	1057	0733	0500	0337	0149	0064	0027	0011	0005
2	2240	1850	1465	1125	0842	0446	0223	0107	0050	0023
3	2240	2158	1954	1687	1404	0892	0521	0286	0150	0076
4	1680	1888	1954	1898	1755	1339	0912	0573	0337	0189
5	1008	1322	1563	1708	1755	1606	1277	0916	0607	0378
6	0504	0771	1042	1281	1462	1606	1490	1221	0911	0631
7	0216	0385	0595	0824	1044	1377	1490	1396	1171	0901
8	0081	0169	0298	0463	0653	1033	1304	1396	1318	1126
9	0027	0066	0132	0232	0363	0688	1014	1241	1318	1251
10	0008	0023	0053	0104	0181	0413	0710	0993	1186	1251
11	0002	0007	0019	0043	0082	0225	0452	0722	0970	1137
12	0001	0002	0006	0016	0034	0113	0263	0481	0728	0948
13	0000	0001	0002	0006	0013	0052	0142	0296	0504	0729
14	0000	0000	0001	0002	0005	0022	0071	0169	0324	0521
15	0000	0000	0000	0001	0002	0009	0033	0090	0194	0347
16	0000	0000	0000	0000	0000	0003	0014	0045	0109	0217
17	0000	0000	0000	0000	0000	0001	0006	0021	0058	0128
18	0000	0000	0000	0000	0000	0000	0002	0009	0029	0071
19	0000	0000	0000	0000	0000	0000	0001	0004	0014	0037
20	0000	0000	0000	0000	0000	0000	0000	0002	0006	0019
21	0000	0000	0000	0000	0000	0000	0000	0001	0003	0009
22	0000	0000	0000	0000	0000	0000	0000	0000	0001	0004

Додаток 4

Таблиця значень функції $t_\alpha = t(\alpha, n)$

$n \backslash \alpha$	0,95	0,99	0,999	$n \backslash \alpha$	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Додаток 5

Таблиця значень функції $q = q(\alpha, n)$

$n \backslash \alpha$	0,95	0,99	0,999	$n \backslash \alpha$	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Критичні точки розподілу Стьюдента

Кількість ступенів вільності k	Рівень значущості α (двостороння критична область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,71	31,82	63,66	318,29	636,58
2	2,92	4,30	6,96	9,92	22,33	31,60
3	2,35	3,18	4,54	5,84	10,21	12,92
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,02	2,57	3,36	4,03	5,89	6,87
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,41
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,02	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,97
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,72	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,50	3,79
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,73
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Рівень значущості α (одностороння критична область)					

Додаток 7

Критичні точки розподілу χ^2

Кількість ступенів вільності k	Рівень значущості α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,63	5,02	3,84	0,0039	0,00098	0,00016
2	9,21	7,38	5,99	0,103	0,051	0,020
3	11,34	9,35	7,81	0,352	0,216	0,115
4	13,28	11,14	9,49	0,711	0,484	0,297
5	15,09	12,83	11,07	1,15	0,831	0,554
6	16,81	14,45	12,59	1,64	1,24	0,872
7	18,48	16,01	14,07	2,17	1,69	1,24
8	20,09	17,53	15,51	2,73	2,18	1,65
9	21,67	19,02	16,92	3,33	2,70	2,09
10	23,21	20,48	18,31	3,94	3,25	2,56
11	24,73	21,92	19,68	4,57	3,82	3,05
12	26,22	23,34	21,03	5,23	4,40	3,57
13	27,69	24,74	22,36	5,89	5,01	4,11
14	29,14	26,12	23,68	6,57	5,63	4,66
15	30,58	27,49	25,00	7,26	6,26	5,23
16	32,00	28,85	26,30	7,96	6,91	5,81
17	33,41	30,19	27,59	8,67	7,56	6,41
18	34,81	31,53	28,87	9,39	8,23	7,01
19	36,19	32,85	30,14	10,12	8,91	7,63
20	37,57	34,17	31,41	10,85	9,59	8,26
21	38,93	35,48	32,67	11,59	10,28	8,90
22	40,29	36,78	33,92	12,34	10,98	9,54
23	41,64	38,08	35,17	13,09	11,69	10,20
24	42,98	39,36	36,42	13,85	12,40	10,86
25	44,31	40,65	37,65	14,61	13,12	11,52
26	45,64	41,92	38,89	15,38	13,84	12,20
27	46,96	43,19	40,11	16,15	14,57	12,88
28	48,28	44,46	41,34	16,93	15,31	13,56
29	49,59	45,72	42,56	17,71	16,05	14,26
30	50,89	46,98	43,77	18,49	16,79	14,95

Критичні точки розподілу F Фішера — Снедекора

		Рівень значущості $\alpha = 0,01$											
$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	11	12	
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6083	6107	
2	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40	99,41	99,42	
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45	14,37	
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89	
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47	
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67	
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11	
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71	
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40	
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16	
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62	3,55	
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,46	
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,43	3,37	
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,29	3,23	
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,24	3,17	
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12	
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07	
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09	3,03	
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	3,06	2,99	
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	3,02	2,96	
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,99	2,93	
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,96	2,90	
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,93	2,87	
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,91	2,84	
31	7,53	5,36	4,48	3,99	3,67	3,45	3,28	3,15	3,04	2,96	2,88	2,82	
32	7,50	5,34	4,46	3,97	3,65	3,43	3,26	3,13	3,02	2,93	2,86	2,80	
33	7,47	5,31	4,44	3,95	3,63	3,41	3,24	3,11	3,00	2,91	2,84	2,78	
34	7,44	5,29	4,42	3,93	3,61	3,39	3,22	3,09	2,98	2,89	2,82	2,76	

Рівень значущості $\alpha = 0,01$

$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	11	12
35	7,42	5,27	4,40	3,91	3,59	3,37	3,20	3,07	2,96	2,88	2,80	2,74
36	7,40	5,25	4,38	3,89	3,57	3,35	3,18	3,05	2,95	2,86	2,79	2,72
37	7,37	5,23	4,36	3,87	3,56	3,33	3,17	3,04	2,93	2,84	2,77	2,71
38	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,92	2,83	2,75	2,69
39	7,33	5,19	4,33	3,84	3,53	3,30	3,14	3,01	2,90	2,81	2,74	2,68
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,73	2,66
41	7,30	5,16	4,30	3,81	3,50	3,28	3,11	2,98	2,87	2,79	2,71	2,65
42	7,28	5,15	4,29	3,80	3,49	3,27	3,10	2,97	2,86	2,78	2,70	2,64
43	7,26	5,14	4,27	3,79	3,48	3,25	3,09	2,96	2,85	2,76	2,69	2,63
44	7,25	5,12	4,26	3,78	3,47	3,24	3,08	2,95	2,84	2,75	2,68	2,62
45	7,23	5,11	4,25	3,77	3,45	3,23	3,07	2,94	2,83	2,74	2,67	2,61
46	7,22	5,10	4,24	3,76	3,44	3,22	3,06	2,93	2,82	2,73	2,66	2,60
47	7,21	5,09	4,23	3,75	3,43	3,21	3,05	2,92	2,81	2,72	2,65	2,59
48	7,19	5,08	4,22	3,74	3,43	3,20	3,04	2,91	2,80	2,71	2,64	2,58
49	7,18	5,07	4,21	3,73	3,42	3,19	3,03	2,90	2,79	2,71	2,63	2,57
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78	2,70	2,63	2,56
55	7,12	5,01	4,16	3,68	3,37	3,15	2,98	2,85	2,75	2,66	2,59	2,53
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56	2,50
65	7,04	4,95	4,10	3,62	3,31	3,09	2,93	2,80	2,69	2,61	2,53	2,47
70	7,01	4,92	4,07	3,60	3,29	3,07	2,91	2,78	2,67	2,59	2,51	2,45
75	6,99	4,90	4,05	3,58	3,27	3,05	2,89	2,76	2,65	2,57	2,49	2,43
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55	2,48	2,42
85	6,94	4,86	4,02	3,55	3,24	3,02	2,86	2,73	2,62	2,54	2,46	2,40
90	6,93	4,85	4,01	3,53	3,23	3,01	2,84	2,72	2,61	2,52	2,45	2,39
95	6,91	4,84	3,99	3,52	3,22	3,00	2,83	2,70	2,60	2,51	2,44	2,38
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,43	2,37
125	6,84	4,78	3,94	3,47	3,17	2,95	2,79	2,66	2,55	2,47	2,39	2,33
150	6,81	4,75	3,91	3,45	3,14	2,92	2,76	2,63	2,53	2,44	2,37	2,31
175	6,78	4,73	3,90	3,43	3,12	2,91	2,74	2,61	2,51	2,42	2,35	2,29
200	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50	2,41	2,34	2,27
300	6,72	4,68	3,85	3,38	3,08	2,86	2,70	2,57	2,47	2,38	2,31	2,24
400	6,70	4,66	3,83	3,37	3,06	2,85	2,68	2,56	2,45	2,37	2,29	2,23
500	6,69	4,65	3,82	3,36	3,05	2,84	2,68	2,55	2,44	2,36	2,28	2,22
700	6,67	4,64	3,81	3,35	3,04	2,83	2,66	2,54	2,43	2,35	2,27	2,21
1000	6,66	4,63	3,80	3,34	3,04	2,82	2,66	2,53	2,43	2,34	2,27	2,20
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,25	2,18

		Рівень значущості $\alpha = 0,01$										
$k_2 \backslash k_1$	14	16	20	24	30	40	50	75	100	200	500	∞
1	6143	6170	6209	6234	6260	6286	6302	6324	6334	6350	6360	6366
2	99,43	99,44	99,45	99,46	99,47	99,48	99,48	99,48	99,49	99,49	99,50	99,50
3	26,92	26,83	26,69	26,60	26,50	26,41	26,35	26,28	26,24	26,18	26,15	26,13
4	14,25	14,15	14,02	13,93	13,84	13,75	13,69	13,61	13,58	13,52	13,49	13,46
5	9,77	9,68	9,55	9,47	9,38	9,29	9,24	9,17	9,13	9,08	9,04	9,02
6	7,60	7,52	7,40	7,31	7,23	7,14	7,09	7,02	6,99	6,93	6,90	6,88
7	6,36	6,28	6,16	6,07	5,99	5,91	5,86	5,79	5,75	5,70	5,67	5,65
8	5,56	5,48	5,36	5,28	5,20	5,12	5,07	5,00	4,96	4,91	4,88	4,86
9	5,01	4,92	4,81	4,73	4,65	4,57	4,52	4,45	4,41	4,36	4,33	4,31
10	4,60	4,52	4,41	4,33	4,25	4,17	4,12	4,05	4,01	3,96	3,93	3,91
11	4,29	4,21	4,10	4,02	3,94	3,86	3,81	3,74	3,71	3,66	3,62	3,60
12	4,05	3,97	3,86	3,78	3,70	3,62	3,57	3,50	3,47	3,41	3,38	3,36
13	3,86	3,78	3,66	3,59	3,51	3,43	3,38	3,31	3,27	3,22	3,19	3,17
14	3,70	3,62	3,51	3,43	3,35	3,27	3,22	3,15	3,11	3,06	3,03	3,00
15	3,56	3,49	3,37	3,29	3,21	3,13	3,08	3,01	2,98	2,92	2,89	2,87
16	3,45	3,37	3,26	3,18	3,10	3,02	2,97	2,90	2,86	2,81	2,78	2,75
17	3,35	3,27	3,16	3,08	3,00	2,92	2,87	2,80	2,76	2,71	2,68	2,65
18	3,27	3,19	3,08	3,00	2,92	2,84	2,78	2,71	2,68	2,62	2,59	2,57
19	3,19	3,12	3,00	2,92	2,84	2,76	2,71	2,64	2,60	2,55	2,51	2,49
20	3,13	3,05	2,94	2,86	2,78	2,69	2,64	2,57	2,54	2,48	2,44	2,42
21	3,07	2,99	2,88	2,80	2,72	2,64	2,58	2,51	2,48	2,42	2,38	2,36
22	3,02	2,94	2,83	2,75	2,67	2,58	2,53	2,46	2,42	2,36	2,33	2,31
23	2,97	2,89	2,78	2,70	2,62	2,54	2,48	2,41	2,37	2,32	2,28	2,26
24	2,93	2,85	2,74	2,66	2,58	2,49	2,44	2,37	2,33	2,27	2,24	2,21
25	2,89	2,81	2,70	2,62	2,54	2,45	2,40	2,33	2,29	2,23	2,19	2,17
26	2,86	2,78	2,66	2,58	2,50	2,42	2,36	2,29	2,25	2,19	2,16	2,13
27	2,82	2,75	2,63	2,55	2,47	2,38	2,33	2,26	2,22	2,16	2,12	2,10
28	2,79	2,72	2,60	2,52	2,44	2,35	2,30	2,23	2,19	2,13	2,09	2,06
29	2,77	2,69	2,57	2,49	2,41	2,33	2,27	2,20	2,16	2,10	2,06	2,03
30	2,74	2,66	2,55	2,47	2,39	2,30	2,25	2,17	2,13	2,07	2,03	2,01
31	2,72	2,64	2,52	2,45	2,36	2,27	2,22	2,14	2,11	2,04	2,01	1,98
32	2,70	2,62	2,50	2,42	2,34	2,25	2,20	2,12	2,08	2,02	1,98	1,96
33	2,68	2,60	2,48	2,40	2,32	2,23	2,18	2,10	2,06	2,00	1,96	1,93
34	2,66	2,58	2,46	2,38	2,30	2,21	2,16	2,08	2,04	1,98	1,94	1,91
35	2,64	2,56	2,44	2,36	2,28	2,19	2,14	2,06	2,02	1,96	1,92	1,89
36	2,62	2,54	2,43	2,35	2,26	2,18	2,12	2,04	2,00	1,94	1,90	1,87

Рівень значущості $\alpha = 0,01$

$k_2 \backslash k_1$	14	16	20	24	30	40	50	75	100	200	500	∞
37	2,61	2,53	2,41	2,33	2,25	2,16	2,10	2,03	1,98	1,92	1,88	1,85
38	2,59	2,51	2,40	2,32	2,23	2,14	2,09	2,01	1,97	1,90	1,86	1,84
39	2,58	2,50	2,38	2,30	2,22	2,13	2,07	1,99	1,95	1,89	1,85	1,82
40	2,56	2,48	2,37	2,29	2,20	2,11	2,06	1,98	1,94	1,87	1,83	1,80
41	2,55	2,47	2,36	2,28	2,19	2,10	2,04	1,97	1,92	1,86	1,82	1,79
42	2,54	2,46	2,34	2,26	2,18	2,09	2,03	1,95	1,91	1,85	1,80	1,78
43	2,53	2,45	2,33	2,25	2,17	2,08	2,02	1,94	1,90	1,83	1,79	1,76
44	2,52	2,44	2,32	2,24	2,15	2,07	2,01	1,93	1,89	1,82	1,78	1,75
45	2,51	2,43	2,31	2,23	2,14	2,05	2,00	1,92	1,88	1,81	1,77	1,74
46	2,50	2,42	2,30	2,22	2,13	2,04	1,99	1,91	1,86	1,80	1,76	1,73
47	2,49	2,41	2,29	2,21	2,12	2,03	1,98	1,90	1,85	1,79	1,74	1,71
48	2,48	2,40	2,28	2,20	2,12	2,02	1,97	1,89	1,84	1,78	1,73	1,70
49	2,47	2,39	2,27	2,19	2,11	2,02	1,96	1,88	1,83	1,77	1,72	1,69
50	2,46	2,38	2,27	2,18	2,10	2,01	1,95	1,87	1,82	1,76	1,71	1,68
55	2,42	2,34	2,23	2,15	2,06	1,97	1,91	1,83	1,78	1,71	1,67	1,64
60	2,39	2,31	2,20	2,12	2,03	1,94	1,88	1,79	1,75	1,68	1,63	1,60
65	2,37	2,29	2,17	2,09	2,00	1,91	1,85	1,77	1,72	1,65	1,60	1,57
70	2,35	2,27	2,15	2,07	1,98	1,89	1,83	1,74	1,70	1,62	1,57	1,54
75	2,33	2,25	2,13	2,05	1,96	1,87	1,81	1,72	1,67	1,60	1,55	1,52
80	2,31	2,23	2,12	2,03	1,94	1,85	1,79	1,70	1,65	1,58	1,53	1,49
85	2,30	2,22	2,10	2,02	1,93	1,83	1,77	1,69	1,64	1,56	1,51	1,47
90	2,29	2,21	2,09	2,00	1,92	1,82	1,76	1,67	1,62	1,55	1,49	1,46
95	2,28	2,20	2,08	1,99	1,90	1,81	1,75	1,66	1,61	1,53	1,48	1,44
100	2,27	2,19	2,07	1,98	1,89	1,80	1,74	1,65	1,60	1,52	1,47	1,43
125	2,23	2,15	2,03	1,94	1,85	1,76	1,69	1,60	1,55	1,47	1,41	1,37
150	2,20	2,12	2,00	1,92	1,83	1,73	1,66	1,57	1,52	1,43	1,38	1,33
175	2,19	2,10	1,98	1,90	1,81	1,71	1,64	1,55	1,50	1,41	1,35	1,30
200	2,17	2,09	1,97	1,89	1,79	1,69	1,63	1,53	1,48	1,39	1,33	1,28
300	2,14	2,06	1,94	1,85	1,76	1,66	1,59	1,50	1,44	1,35	1,28	1,22
400	2,13	2,05	1,92	1,84	1,75	1,64	1,58	1,48	1,42	1,32	1,25	1,19
500	2,12	2,04	1,92	1,83	1,74	1,63	1,57	1,47	1,41	1,31	1,23	1,16
600	2,11	2,03	1,91	1,82	1,73	1,63	1,56	1,46	1,40	1,30	1,22	1,15
700	2,11	2,03	1,90	1,82	1,72	1,62	1,55	1,45	1,39	1,29	1,21	1,14
800	2,10	2,02	1,90	1,81	1,72	1,62	1,55	1,45	1,39	1,29	1,20	1,13
1000	2,10	2,02	1,90	1,81	1,72	1,61	1,54	1,44	1,38	1,28	1,19	1,11
∞	2,08	2,00	1,88	1,79	1,70	1,59	1,52	1,42	1,36	1,25	1,15	1,00

Рівень значущості $\alpha = 0,05$												
$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	11	12
1	161	199	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17	2,13
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09
31	4,16	3,30	2,91	2,68	2,52	2,41	2,32	2,25	2,20	2,15	2,11	2,08
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	2,10	2,07
33	4,14	3,28	2,89	2,66	2,50	2,39	2,30	2,23	2,18	2,13	2,09	2,06
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	2,08	2,05
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	2,07	2,04
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	2,07	2,03

Рівень значущості $\alpha = 0,05$												
$k_2 \backslash k_1$	1	2	3	4	5	6	7	8	9	10	11	12
37	4,11	3,25	2,86	2,63	2,47	2,36	2,27	2,20	2,14	2,10	2,06	2,02
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02
39	4,09	3,24	2,85	2,61	2,46	2,34	2,26	2,19	2,13	2,08	2,04	2,01
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00
41	4,08	3,23	2,83	2,60	2,44	2,33	2,24	2,17	2,12	2,07	2,03	2,00
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	2,03	1,99
43	4,07	3,21	2,82	2,59	2,43	2,32	2,23	2,16	2,11	2,06	2,02	1,99
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	2,01	1,98
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	2,01	1,97
46	4,05	3,20	2,81	2,57	2,42	2,30	2,22	2,15	2,09	2,04	2,00	1,97
47	4,05	3,20	2,80	2,57	2,41	2,30	2,21	2,14	2,09	2,04	2,00	1,96
48	4,04	3,19	2,80	2,57	2,41	2,29	2,21	2,14	2,08	2,03	1,99	1,96
49	4,04	3,19	2,79	2,56	2,40	2,29	2,20	2,13	2,08	2,03	1,99	1,96
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95
55	4,02	3,16	2,77	2,54	2,38	2,27	2,18	2,11	2,06	2,01	1,97	1,93
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92
65	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,03	1,98	1,94	1,90
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93	1,89
75	3,97	3,12	2,73	2,49	2,34	2,22	2,13	2,06	2,01	1,96	1,92	1,88
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91	1,88
85	3,95	3,10	2,71	2,48	2,32	2,21	2,12	2,05	1,99	1,94	1,90	1,87
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,90	1,86
95	3,94	3,09	2,70	2,47	2,31	2,20	2,11	2,04	1,98	1,93	1,89	1,86
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89	1,85
125	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,96	1,91	1,87	1,83
150	3,90	3,06	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,82
175	3,90	3,05	2,66	2,42	2,27	2,15	2,06	1,99	1,93	1,89	1,84	1,81
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,84	1,80
300	3,87	3,03	2,63	2,40	2,24	2,13	2,04	1,97	1,91	1,86	1,82	1,78
400	3,86	3,02	2,63	2,39	2,24	2,12	2,03	1,96	1,90	1,85	1,81	1,78
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,77
600	3,86	3,01	2,62	2,39	2,23	2,11	2,02	1,95	1,90	1,85	1,80	1,77
700	3,85	3,01	2,62	2,38	2,23	2,11	2,02	1,95	1,89	1,84	1,80	1,77
800	3,85	3,01	2,62	2,38	2,23	2,11	2,02	1,95	1,89	1,84	1,80	1,76
1000	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1,80	1,76
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75

		Рівень значущості $\alpha = 0,05$										
$k_2 \backslash k_1$	14	16	20	24	30	40	50	75	100	200	500	∞
1	245	246	248	249	250	251	252	253	253	254	254	254
2	19,42	19,43	19,45	19,45	19,46	19,47	19,48	19,48	19,49	19,49	19,49	19,50
3	8,71	8,69	8,66	8,64	8,62	8,59	8,58	8,56	8,55	8,54	8,53	8,53
4	5,87	5,84	5,80	5,77	5,75	5,72	5,70	5,68	5,66	5,65	5,64	5,63
5	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,42	4,41	4,39	4,37	4,37
6	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,73	3,71	3,69	3,68	3,67
7	3,53	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,27	3,25	3,24	3,23
8	3,24	3,20	3,15	3,12	3,08	3,04	3,02	2,99	2,97	2,95	2,94	2,93
9	3,03	2,99	2,94	2,90	2,86	2,83	2,80	2,77	2,76	2,73	2,72	2,71
10	2,86	2,83	2,77	2,74	2,70	2,66	2,64	2,60	2,59	2,56	2,55	2,54
11	2,74	2,70	2,65	2,61	2,57	2,53	2,51	2,47	2,46	2,43	2,42	2,40
12	2,64	2,60	2,54	2,51	2,47	2,43	2,40	2,37	2,35	2,32	2,31	2,30
13	2,55	2,51	2,46	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,22	2,21
14	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13
15	2,42	2,38	2,33	2,29	2,25	2,20	2,18	2,14	2,12	2,10	2,08	2,07
16	2,37	2,33	2,28	2,24	2,19	2,15	2,12	2,09	2,07	2,04	2,02	2,01
17	2,33	2,29	2,23	2,19	2,15	2,10	2,08	2,04	2,02	1,99	1,97	1,96
18	2,29	2,25	2,19	2,15	2,11	2,06	2,04	2,00	1,98	1,95	1,93	1,92
19	2,26	2,21	2,16	2,11	2,07	2,03	2,00	1,96	1,94	1,91	1,89	1,88
20	2,22	2,18	2,12	2,08	2,04	1,99	1,97	1,93	1,91	1,88	1,86	1,84
21	2,20	2,16	2,10	2,05	2,01	1,96	1,94	1,90	1,88	1,84	1,83	1,81
22	2,17	2,13	2,07	2,03	1,98	1,94	1,91	1,87	1,85	1,82	1,80	1,78
23	2,15	2,11	2,05	2,01	1,96	1,91	1,88	1,84	1,82	1,79	1,77	1,76
24	2,13	2,09	2,03	1,98	1,94	1,89	1,86	1,82	1,80	1,77	1,75	1,73
25	2,11	2,07	2,01	1,96	1,92	1,87	1,84	1,80	1,78	1,75	1,73	1,71
26	2,09	2,05	1,99	1,95	1,90	1,85	1,82	1,78	1,76	1,73	1,71	1,69
27	2,08	2,04	1,97	1,93	1,88	1,84	1,81	1,76	1,74	1,71	1,69	1,67
28	2,06	2,02	1,96	1,91	1,87	1,82	1,79	1,75	1,73	1,69	1,67	1,65
29	2,05	2,01	1,94	1,90	1,85	1,81	1,77	1,73	1,71	1,67	1,65	1,64
30	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,70	1,66	1,64	1,62
31	2,03	1,98	1,92	1,88	1,83	1,78	1,75	1,70	1,68	1,65	1,62	1,61
32	2,01	1,97	1,91	1,86	1,82	1,77	1,74	1,69	1,67	1,63	1,61	1,59
33	2,00	1,96	1,90	1,85	1,81	1,76	1,72	1,68	1,66	1,62	1,60	1,58
34	1,99	1,95	1,89	1,84	1,80	1,75	1,71	1,67	1,65	1,61	1,59	1,57
35	1,99	1,94	1,88	1,83	1,79	1,74	1,70	1,66	1,63	1,60	1,57	1,56
36	1,98	1,93	1,87	1,82	1,78	1,73	1,69	1,65	1,62	1,59	1,56	1,55

Рівень значущості $\alpha = 0,05$												
$k_2 \backslash k_1$	14	16	20	24	30	40	50	75	100	200	500	∞
37	1,97	1,93	1,86	1,82	1,77	1,72	1,68	1,64	1,62	1,58	1,55	1,54
38	1,96	1,92	1,85	1,81	1,76	1,71	1,68	1,63	1,61	1,57	1,54	1,53
39	1,95	1,91	1,85	1,80	1,75	1,70	1,67	1,62	1,60	1,56	1,53	1,52
40	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51
41	1,94	1,90	1,83	1,79	1,74	1,69	1,65	1,61	1,58	1,54	1,52	1,50
42	1,94	1,89	1,83	1,78	1,73	1,68	1,65	1,60	1,57	1,53	1,51	1,49
43	1,93	1,89	1,82	1,77	1,72	1,67	1,64	1,59	1,57	1,53	1,50	1,48
44	1,92	1,88	1,81	1,77	1,72	1,67	1,63	1,59	1,56	1,52	1,49	1,48
45	1,92	1,87	1,81	1,76	1,71	1,66	1,63	1,58	1,55	1,51	1,49	1,47
46	1,91	1,87	1,80	1,76	1,71	1,65	1,62	1,57	1,55	1,51	1,48	1,46
47	1,91	1,86	1,80	1,75	1,70	1,65	1,61	1,57	1,54	1,50	1,47	1,46
48	1,90	1,86	1,79	1,75	1,70	1,64	1,61	1,56	1,54	1,49	1,47	1,45
49	1,90	1,85	1,79	1,74	1,69	1,64	1,60	1,56	1,53	1,49	1,46	1,44
50	1,89	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44
55	1,88	1,83	1,76	1,72	1,67	1,61	1,58	1,53	1,50	1,46	1,43	1,41
60	1,86	1,82	1,75	1,70	1,65	1,59	1,56	1,51	1,48	1,44	1,41	1,39
65	1,85	1,80	1,73	1,69	1,63	1,58	1,54	1,49	1,46	1,42	1,39	1,37
70	1,84	1,79	1,72	1,67	1,62	1,57	1,53	1,48	1,45	1,40	1,37	1,35
75	1,83	1,78	1,71	1,66	1,61	1,55	1,52	1,47	1,44	1,39	1,36	1,34
80	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,43	1,38	1,35	1,32
85	1,81	1,76	1,70	1,65	1,59	1,54	1,50	1,45	1,42	1,37	1,34	1,31
90	1,80	1,76	1,69	1,64	1,59	1,53	1,49	1,44	1,41	1,36	1,33	1,30
95	1,80	1,75	1,68	1,63	1,58	1,52	1,48	1,43	1,40	1,35	1,32	1,29
100	1,79	1,75	1,68	1,63	1,57	1,52	1,48	1,42	1,39	1,34	1,31	1,28
125	1,77	1,73	1,66	1,60	1,55	1,49	1,45	1,40	1,36	1,31	1,27	1,25
150	1,76	1,71	1,64	1,59	1,54	1,48	1,44	1,38	1,34	1,29	1,25	1,22
175	1,75	1,70	1,63	1,58	1,52	1,46	1,42	1,36	1,33	1,27	1,23	1,20
200	1,74	1,69	1,62	1,57	1,52	1,46	1,41	1,35	1,32	1,26	1,22	1,19
300	1,72	1,68	1,61	1,55	1,50	1,43	1,39	1,33	1,30	1,23	1,19	1,15
400	1,72	1,67	1,60	1,54	1,49	1,42	1,38	1,32	1,28	1,22	1,17	1,13
500	1,71	1,66	1,59	1,54	1,48	1,42	1,38	1,31	1,28	1,21	1,16	1,11
600	1,71	1,66	1,59	1,54	1,48	1,41	1,37	1,31	1,27	1,20	1,15	1,10
700	1,71	1,66	1,59	1,53	1,48	1,41	1,37	1,30	1,27	1,20	1,15	1,09
800	1,70	1,66	1,58	1,53	1,47	1,41	1,37	1,30	1,26	1,20	1,14	1,09
1000	1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08
∞	1,69	1,64	1,57	1,52	1,46	1,39	1,35	1,28	1,24	1,17	1,11	1,00

Додаток 9

Латинський алфавіт

<i>A</i>	<i>a</i>	а	<i>N</i>	<i>n</i>	ен
<i>B</i>	<i>b</i>	бе	<i>O</i>	<i>o</i>	о
<i>C</i>	<i>c</i>	це	<i>P</i>	<i>p</i>	пе
<i>D</i>	<i>d</i>	де	<i>Q</i>	<i>q</i>	ку
<i>E</i>	<i>e</i>	е	<i>R</i>	<i>r</i>	ер
<i>F</i>	<i>f</i>	еф	<i>S</i>	<i>s</i>	ес
<i>G</i>	<i>g</i>	ге (же)	<i>T</i>	<i>t</i>	те
<i>H</i>	<i>h</i>	ха (аш)	<i>U</i>	<i>u</i>	у
<i>I</i>	<i>i</i>	і	<i>V</i>	<i>v</i>	ве
<i>J</i>	<i>j</i>	йот (жі)	<i>W</i>	<i>w</i>	дубль-ве
<i>K</i>	<i>k</i>	ка	<i>X</i>	<i>x</i>	ікс
<i>L</i>	<i>l</i>	ель	<i>Y</i>	<i>y</i>	ігрек
<i>M</i>	<i>m</i>	ем	<i>Z</i>	<i>z</i>	зет

Додаток 10

Грецький алфавіт

<i>A</i>	α	альфа	<i>N</i>	ν	ню (ні)
<i>B</i>	β	бета	Ξ	ξ	ксі
<i>Г</i>	γ	гамма	<i>O</i>	<i>o</i>	омікрон
Δ	δ	дельта	<i>П</i>	π	пі
<i>E</i>	ϵ	епсілон	<i>P</i>	ρ	ро
<i>Z</i>	ζ	дзета	Σ	σ	сігма
<i>H</i>	η	ета	<i>T</i>	τ	тау
Θ	θ, ϑ	тета	Υ	υ	юпсілон (іпсілон)
<i>I</i>	ι	йота	Φ	φ	фі
<i>K</i>	κ	каппа	<i>X</i>	χ	хі
<i>L</i>	λ	лямбда	Ψ	ψ	псі
<i>M</i>	μ	мю (мі)	Ω	ω	омега

Список використаної та рекомендованої літератури

1. *Бешелев С. Д., Гурвич Ф. Г.* Математико-статистические методы экспертных оценок. — М.: Статистика, 1980. — 263 с.
2. *Вайнберг Дж., Шуменер Дж.* Статистика: Пер. с англ. — М.: Статистика, 1979. — 389 с.
3. *Вентцель Е. С.* Теория вероятностей. — М.: Физматгиз, 1963.
4. *Гмурман В. Е.* Руководство к решению задач по теории вероятностей и математической статистике: Учеб. пособие для студентов вузов. — 6-е изд., доп. — М.: Высш. шк., 2002. — 405 с.
5. *Гмурман В. Е.* Теория вероятностей и математическая статистика. — М.: Высш. шк., 1999.
6. *Гнеденко Б. В.* Курс теории вероятностей. — М.: Физматгиз, 1961.
7. *Гнеденко Б. В., Хинчин А. Я.* Элементарное введение в теорию вероятностей. — М.: Наука, 1976. — 168 с.
8. *Горбань С. Ф., Снижско Н. В.* Теория вероятностей и математическая статистика. — К.: МАУП, 1999. — 168 с.
9. *Громько Г. Л.* Статистика. — М.: Изд-во МГУ, 1981. — С. 3–166.
10. *Гурский Е. М.* Теория вероятностей с элементами математической статистики. — М.: Высш. шк., 1971.

Список використаної та рекомендованої літератури

11. *Жлуктенко В. І., Наконечний С. І.* Практикум з курсу «Теорія ймовірностей і математична статистика». — К.: КІНГ, 1991.
12. *Жлуктенко В. І., Наконечний С. І.* Теорія ймовірностей із елементами математичної статистики. — К.: НМК ВО, 1991.
13. *Карасев А. И.* Теория вероятности и математическая статистика. — М.: Статистика, 1977.
14. *Кимбл Г.* Как правильно пользоваться статистикой: Пер. с англ. — М.: Финансы и статистика, 1982. — 294 с.
15. *Кремер Н. Ш.* Теория вероятностей и математическая статистика. — М.: ЮНИТИ, 2000.
16. *Лбов Г. С.* Методы обработки разнотипных экспериментальных данных. — Новосибирск: Наука, 1981.
17. *Математическая статистика: Учебник.* — М.: Высш. шк., 1981. — 371 с.
18. *Мюллер П., Нойман Ф., Шторм Р.* Таблицы по математической статистике: Пер. с нем. — М.: Финансы и статистика, 1982. — 178 с.
19. *Окунь Я.* Факторный анализ: Пер. с польск. — М.: Статистика, 1974.
20. *Сидоренко Е. В.* Методы математической обработки в психологии. — СПб.: ООО «Речь», 2001. — 350 с.
21. *Справочник по теории вероятностей и математической статистике.* — К.: Наук. думка, 1978. — 582 с.
22. *Турчин В. М.* Математична статистика. Навч. посіб. — К.: Видавничий центр «Академія», 1999. — 240 с.
23. *Хастинг Н., Пикон Дж.* Справочник по статистическим распределениям: Пер. с англ. — М.: Статистика, 1980. — 95 с.
24. *Чистяков В. П.* Курс теории вероятностей. — М., 1978.
25. *Шефтель З. Г.* Теорія ймовірностей. — К., 1994.

Предметний покажчик

В

- Варіанта 157
- Варіаційний
 - розмах 176
 - ряд 157
- Вибіркове
 - кореляційне відношення 224
 - середнє 171
- Випадкова величина 49
 - дискретна 49, 96
 - неперервна 62, 97
 - абсолютно 62, 97
 - сингулярна 62
 - подія 3
- Випадкові величини
 - незалежні 98
 - некорельовані 99
- Випробування 3

Г

- Гіпотеза
 - альтернативна 228
 - конкуруюча 228
 - нульова 228
 - основна 228
 - проста 228
 - складна 228
 - статистична 228
- Гістограма
 - відносних частот 162
 - частот 162

Д

- Дисперсія 83
 - вибіркова 171
 - «виправлена» 195
- Добуток подій 5
- Дослід 3

З

- Закон розподілу 49
 - умовний 98
- Зліченна множина 49
- Значення
 - емпіричне 228
 - спостережуване 228

І

- Інтервал
 - довірчий 206
 - частинний
 - медіанний 174
 - модальний 175

Й

- Ймовірність
 - добутку 29
 - емпірична 157
 - суми 29
 - умовна 29

К

- Коваріація 99
- Коефіцієнт кореляції 99
- Кореляційний момент 99
- Кореляція

- криволінійна 223
- лінійна 219
- параболічна другого порядку 224
- Критерій 228
 - Пірсона 245, 248
 - Фішера — Снедекора 272, 277, 284, 293
 - статистичний 228
- Критичні
 - межі 228
 - точки 228
- М**
- Математичне сподівання
 - абсолютно неперервної випадкової величини 83
 - дискретної випадкової величини 83
- Медіана 85, 174
- Метод
 - добутків 177
 - максимальної правдоподібності 201
 - моментів 199
 - найбільшої правдоподібності 201
 - сум 182
- Многокутник розподілу 50
- Мода 175
 - випадкової величини
 - абсолютно неперервної 85
 - дискретної 85
- Н**
- Наслідок елементарний 4
- О**
- Область
 - допустимих значень 228
 - критична 228
 - двостороння 229
 - лівостороння 229
 - правостороння 228
 - прийняття гіпотези 228
- Означення ймовірності
 - геометричне 9
 - класичне 7
 - статистичне 14
- Оптимальна величина інтервалу 162
- Основний принцип перевірки статистичних гіпотез 228
- Оцінка
 - зміщена 193
 - генеральної дисперсії 194
 - інтервальна 206
 - ймовірності 207
 - математичного сподівання 206
 - середнього квадратичного відхилення 206
 - максимальної правдоподібності 202
 - найбільшої правдоподібності 202
 - незміщена 193
 - генерального середнього 193
 - генеральної дисперсії 195
 - статистична 193
 - точкова 193
- П**
- Перестановки 20
- Повна група подій 4
- Подія 3
 - достовірна 3
 - елементарна 4
 - сприятлива 4
 - неможлива 3
 - протилежна 5
- Події
 - залежні 29
 - незалежні 29
 - несумісні 4
 - сумісні 4
- Полігон
 - частот 161
 - відносних частот 161
- Поправка
 - Бесселя 195
 - Шеппарда 180

- Потужність критерію 229
- Похибка
 - другого роду 228
 - першого роду 228
- Практична достовірність 65
- Проміжок «трисигмовий» 64
- Простір елементарних подій 4
- Р**
- Регресія
 - криволінійна 223
 - лінійна 219
 - параболічна 223
- Рівень значущості 228
- Рівність Вілсона — Гільферті 238
- Рівняння
 - кривої лінії регресії 223
 - параболічної регресії 223
 - правдоподібності 202
 - прямої лінії регресії 219
- Різниця подій 5
- Розміщення 20
- Розподіл
 - Бернуллі 54
 - біноміальний 54
 - вибірки
 - інтервальний 158
 - статистичний 157
 - точковий 158
 - геометричний 55
 - гіпергеометричний 55
 - експоненціальний 66
 - нормальний 63
 - показниковий 55, 66
 - полімодальний 85
 - Пуассона 55
 - рівномірний 55, 63
 - унімодальний 85
- С**
- Середнє квадратичне відхилення 85
 - вибірки 171
- Сполучення 20
- Сума
 - загальна 270, 277, 283
 - залишкова 271, 277, 284
 - індивідуальна 277
 - подій 5
 - факторна 270, 277, 283
- Схема Бернуллі 40
- Т**
- Теорема
 - Бернуллі 43
 - Муавра — Лапласа
 - інтегральна 43
 - локальна 42
- Ф**
- Формула
 - Байеса 33
 - Бернуллі 40
 - повної ймовірності 33
 - Пуассона 42
- Функція
 - випадкової величини 135
 - Лапласа 65
 - інтегральна 42
 - локальна 42
 - правдоподібності
 - дискретної випадкової величини 202
 - логарифмічна 202
 - неперервної випадкової величини 202
 - розподілу 50, 95
 - вибірки 159
 - диференціальна 62
 - емпірична 159
- Ч**
- Частота 157
 - відносна 157
 - події 14
- Щ**
- Щільність
 - частоти 162
 - відносної частоти 162
 - розподілу 97
 - випадкової величини 62

Зміст

Частина I. Випадкові події	3
<i>Розділ 1.</i> Випадкові події та операції над ними.	
Означення ймовірності	3
1.1. Випадкові події (3). 1.2. Операції над подіями (5). 1.3. Класичне означення ймовірності (7). 1.4. Геометричне означення ймовірності (9). 1.5. Статистичне означення ймовірності (14).	
Задачі до розділу 1	15
<i>Розділ 2.</i> Елементи комбінаторики та їх застосування при обчисленні ймовірностей	20
Задачі до розділу 2	26
<i>Розділ 3.</i> Основні формули додавання і множення ймовірностей	29
Задачі до розділу 3	31
<i>Розділ 4.</i> Формула повної ймовірності. Формула Байеса	33
Задачі до розділу 4	36
<i>Розділ 5.</i> Незалежні повторні випробування.	
Формула Бернуллі	40
5.1. Формула Бернуллі (40). 5.2. Граничні теореми у схемі Бернуллі (42).	
Задачі до розділу 5	47
Частина II. Випадкові величини	49
<i>Розділ 6.</i> Дискретні випадкові величини	49
Задачі до розділу 6	57
<i>Розділ 7.</i> Неперервні випадкові величини	62
Задачі до розділу 7	72
<i>Розділ 8.</i> Числові характеристики випадкових величин	83
Задачі до розділу 8	89
<i>Розділ 9.</i> Системи двох випадкових величин	95
Задачі до розділу 9	133
<i>Розділ 10.</i> Функції випадкових величин	135
Задачі до розділу 10	142
<i>Розділ 11.</i> Граничні теореми теорії ймовірностей	147
Задачі до розділу 11	153

Частина III. Математична статистика	157
<i>Розділ 12. Елементи математичної статистики.</i>	
Вибірковий метод	157
12.1. Вибірковий метод (157). 12.2. Числові характеристики вибірки (171). 12.3. Метод добутоків обчислення вибіркового середнього та вибіркової дисперсії (177). 12.4. Метод сум обчислення вибіркового середнього та вибіркової дисперсії (182).	
Задачі до розділу 12	185
<i>Розділ 13. Статистичні оцінки параметрів розподілу</i>	193
13.1. Точкові оцінки (193). 13.2. Метод моментів (199). 13.3. Метод найбільшої правдоподібності (201). 13.4. Інтервальні оцінки (206).	
Задачі до розділу 13	213
<i>Розділ 14. Елементи теорії регресії і кореляції</i>	219
14.1. Рівняння прямої лінії регресії. Лінійна кореляція (219). 14.2. Рівняння параболічної регресії. Параболічна кореляція (223).	
Задачі до розділу 14	226
<i>Розділ 15. Статистична перевірка статистичних гіпотез</i>	228
15.1. Перевірка рівності вибіркового середнього гіпотетичному генеральному середньому (230). 15.2. Перевірка рівності «виправленої» вибіркової дисперсії генеральній дисперсії (237). 15.3. Перевірка рівності відносної частоти гіпотетичній імовірності (242). 15.4. Перевірка гіпотези про нормальний розподіл за критерієм Пірсона (245). 15.5. Перевірка гіпотези про рівномірний розподіл (252). 15.6. Перевірка гіпотези про показниковий розподіл (256). 15.7. Перевірка гіпотези про біноміальний розподіл (259). 15.8. Перевірка гіпотези про розподіл Пуассона (263).	
Задачі до розділу 15	266
<i>Розділ 16. Елементи дисперсійного аналізу</i>	270
16.1. Однофакторний дисперсійний аналіз для незв'язних вибірок (270). 16.2. Однофакторний дисперсійний аналіз для зв'язних вибірок (276). 16.3. Двофакторний дисперсійний аналіз для незв'язних вибірок (281). 16.4. Двофакторний дисперсійний аналіз для зв'язних вибірок (290).	
Задачі до розділу 16	300
<i>Додатки</i>	304
<i>Список використаної та рекомендованої літератури</i>	321
<i>Предметний покажчик</i>	323

In the educational manual necessary theoretical information and formulas, solutions of typical tasks are given, it also contains tasks for self solutions with instructions and answers for them. A great attention is paid for methods of statistic processing of experimental data.

It is meant for students of non-mathematical specialities in the higher educational establishments and also for all, who uses the probability theory and statistic methods in solutions of practical tasks.

Навчальне видання

Чорней Руслан Костянтинович

Дюженкова Ольга Юріївна

Жильцов Олексій Борисович та інші

**ПРАКТИКУМ З ТЕОРІЇ ЙМОВІРНОСТЕЙ
ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ**

Навчальний посібник

Educational edition

Chornei, Ruslan K.

Dyuzhenkova, Olga Yur.

Zhyl'tsov, Oleksiy B. and others

**PRACTICE BOOK OF PROBABILITY THEORY
AND MATHEMATICAL STATISTICS**

Educational manual

Відповідальний редактор *С. Г. Рогузко*

Редактор *Л. В. Логвиненко*

Коректор *Л. Г. Бурлакіна*

Комп'ютерне верстання *Д. А. Варварін*

Оформлення обкладинки *О. О. Стеценко*

Підп. до друку 22.09.03. Формат 60 × 84/16. Папір офсетний. Друк офсетний.
Ум. друк. арк. 19,06. Обл.-вид. арк. 15,7. Тираж 8000 пр. Зам. № 03-02

Міжрегіональна Академія управління персоналом (МАУП)
03039 Київ-39, вул. Фрометівська, 2, МАУП

*Свідоцтво про внесення до Державного реєстру
суб'єктів видавничої справи ДК № 8 від 23.02.2000*

АТЗТ “Видавничий центр “Друк”
03057 Київ-57, вул. Довженка, 10

Свідоцтво ДК № 108 від 4.07.2000