УКРАЇНСЬКА ДЕРЖАВНА АКАДЕМІЯ

ЗАЛІЗНИЧНОГО ТРАНСПОРТУ

ЕЛЕМЕНТИ ТЕОРІЇ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНОЇ СТАТИСТИКИ В УПРАВЛІННІ ПРОЦЕСАМИ ПЕРЕВЕЗЕНЬ

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

Харків 2011

УДК 519.2

ББК 22.17
Е503

Рекомендовано Міністерством освіти і науки України як навчальний посібник для студентів вищих навчальних закладів (№ 1/11-11500 від 17.12.10 р.).

Авторський колектив:

Т.В. Бутько, Р.В. Вовк, Н.Г. Панченко, А.П. Рибалко

Рецензенти:

А.А. Янцевич, доктор фізико-математичних наук, професор (Харківський національний університет ім. В.Н. Каразіна)

Г.І. Нечаєв, доктор технічних наук, професор, академік транспортної академії України (Східноукраїнський національний університет ім. В.Даля)

І.В. Жуковицький, доктор технічних наук, професор (Дніпропетровський національний університет залізничного транспорту ім. академіка В. Лазаряна)

Е503 Елементи теорії ймовірностей і математичної статистики в управлінні процесами перевезень: Навч. посібник / Т.В. Бутько, Р.В. Вовк, Н.Г. Панченко та ін. – Харків: УкрДАЗТ, 2011. — 308 с., табл. 25, рис. 42.
ISBN 978-966-2033-55-7

Навчальний посібник призначений для систематичного вивчення дисципліни теорії ймовірностей і математичної статистики студентами інженерно-транспортних спеціальностей вищих навчальних закладів. Особливу увагу приділено можливостям застосування математичного апарату до розв’язання конкретних практичних задач з галузі експлуатації залізничного транспорту.

УДК 519.2
ББК 22.17

ISBN 978-966-2033-55-7

Навчальний посібник

Бутько Тетяна Васильївна
Вовк Руслан Володимирович
Панченко Наталія Георгієвна
та ін.

ЕЛЕМЕНТИ ТЕОРІЇ ЙМОВІРНОСТЕЙ
І МАТЕМАТИЧНОЇ СТАТИСТИКИ В УПРАВЛІННІ
ПРОЦЕСАМИ ПЕРЕВЕЗЕНЬ

Відповідальний за випуск Рибалко А.П.

Редактор Еткало О.О.

Підписано до друку 15.12.09 р.

Формат паперу 60х84 1/16. Папір писальний.
Умовн.-друк.арк. 8,25. Тираж 300. Замовлення №

Видавець та виготовлювач Українська державна академія залізничного транспорту,
61050, Харків-50, майдан Фейєрбаха, 7.
Свідотство суб’єкта видавничої справи ДК № 2874 від 12.06.2007 р.

ЕЛЕМЕНТИ ТЕОРІЇ ЙМОВІРНОСТЕЙ
І МАТЕМАТИЧНОЇ СТАТИСТИКИ В
УПРАВЛІННІ ПРОЦЕСАМИ ПЕРЕВЕЗЕНЬ

Навчальний посібник

Харків 2011

Зміст

ВСТУП

В основу навчального посібника покладено курс лекцій, що читався протягом кількох років на факультеті управління процесами перевезень Української державної академії залізничного транспорту. Він призначений для вивчення дисципліни теорії ймовірностей і математичної статистики студентами спеціальності “Організація перевезень і управління на залізничному транспорті”.

Сферами діяльності майбутніх спеціалістів будуть: оперативне регулювання та планування при організації роботи транспортних систем; аналіз, прогнозування та здійснення контролю роботи транспортних підприємств; розв’язання інших складних проблем управління з метою підвищення ефективності роботи транспортних комплексів.

Найчастіше процеси в транспортних системах мають імовірнісну природу, зокрема: процеси навантаження та вивантаження при обслуговуванні клієнтів на вантажних станціях, процеси формування пасажирських та вантажних потягів, прибуття пасажирів до квиткових кас, виникнення відмов у роботі деякого обладнання тощо. Особливості функціонування ринку транспортних послуг в умовах коливання також потребують застосування теоретико-імовірнісних та статистичних методів. Суть їх полягає в забезпеченні прийняття науково обґрунтованих управлінських рішень за умов невизначеності. Прикладне застосування теорії ймовірностей і математичної статистики дає змогу досліджувати та моделювати відповідні процеси, що приводить до подальшої автоматизації роботи на транспорті.

Метою даного посібника є системне ознайомлення студентів з відповідним математичним апаратом, демонстрація можливостей застосування його до розв’язання конкретних практичних задач, розвиток теоретико-імовірнісної інтуїції спеціаліста, формування у студентів навичок управління в невизначених умовах.

Імовірнісні методи використовуються, коли явища зумовлені надто великим числом факторів, що робить вивчення їх класичними математичними методами практично неможливим.

Виявляється, що навіть випадкові події в масі підлягають певним об’єктивним закономірностям. Теорія ймовірностей – наука, що вивчає закономірності, притаманні масовим випадковим явищам. Імовірнісний підхід дозволяє відтворити ці закономірності в абстрактній формі, тобто побудувати їх математичні моделі. Також теорія ймовірностей служить для обґрунтування математичної статистики, методи якої відіграють виключно важливу роль у розв’язанні широкого кола проблем практичного характеру при управлінні перевізним процесом на залізничному транспорті.

Даний навчальний посібник ознайомлює студентів з теоретичними основами теорії ймовірностей та математичної статистики, можливостями застосування імовірнісно-статистичного апарату для розв’язання практичних задач. Цей курс є базовим для вивчення спеціальних дисциплін, що викладатимуться профілюючими кафедрами, при виконанні курсових проектів, дипломних та магістерських кваліфікаційних робіт. У зв’язку з цільовим призначенням курсу значна частина прикладів взята з галузі експлуатації залізничного транспорту.

Навчальний посібник складається з дев’яти розділів: розділи 1–6 присвячені вивченню теорії ймовірностей, розділи 7–9 – елементам математичної статистики. Нумерація формул є індексаційною пороздільною.

Теоретичні поняття і факти викладено в посібнику у повному обсязі. Особливу увагу слід звертати на підкреслені формули, таблиці та виділені зауваження. Викладення теоретичного матеріалу супроводжується розв’язанням прикладів, що розглядаються в дисциплінах “Управління експлуатаційною роботою та якістю перевезень”, “Управління вантажною та комерційною роботою”, “Інформаційні системи” тощо. Це дозволяє студентам краще осмислити та навчитися застосовувати теоретичні знання.

Вивчення теорії ймовірностей обов’язково повинно супроводжуватись розв’язанням задач. Задачі теорії ймовірностей часто мають нематематичну постановку, тому для їх розв’язання треба розвивати навички теоретико-імовірнісного підходу. Тільки опрацювання отриманих знань за допомогою виконання завдань дозволить навчитися розробляти коректні математичні моделі випадкових явищ. Для отримання відповідних навичок студенту необхідно самостійно розв’язувати задачі, які містяться в кінці кожного підрозділу. Ці задачі мають відповіді та, в деяких випадках, вказівки щодо розв’язання.

Надзвичайно важливими для самоконтролю засвоєння матеріалу курсу є теоретичні та тестові питання до розглянутої теми. Їх можна знайти в кінці кожного розділу, причому до тестових питань наведено відповіді.

Для розв’язання задач з теорії ймовірностей і математичної статистики необхідно користуватися таблицями деяких спеціальних функцій та основних розподілів імовірностей. Для зручності вони наведені в додатках.

Оскільки курс спрямований на відповідну спеціальність, студентам факультету управління процесами перевезень цей посібник рекомендується використовувати як базовий, але, звичайно, бажано також користуватись іншими джерелами. Це дозволить поглибити теоретичні знання, вдосконалити навички розв’язання задач та розширити уявлення про застосування ймовірнісно-статистичних методів на практиці. Список рекомендованої літератури можна знайти наприкінці посібника.

Розділ 1. Випадкові події. Теореми теорії ймовірностей

1.1. Випадкові події та їх класифікація

Подія називається *випадковою*, якщо за певного комплексу умов у результаті експерименту вона може настати або не настати. Випадкові події позначають великими латинськими літерами:

$$A, B, C, \ldots; A_1, A_2, A_3, \ldots; B_1, B_2, B_3, \ldots.$$

Приклад. Випадковими подіями є:

1) наявність або відсутність у касі квитка на даний потяг;

2) наявність вагона, що потребує ремонту, у потязі, який прибув на станцію;

3) відмова у роботі деякого обладнання;

4) випадання герба на монеті, яку кинуто.

Випадкові події в масі спостережень підпорядковані певним характерним лише для них невипадковим законам. Математична наука, що вивчає закономірності масових подій, називається *теорією ймовірностей*. Науку, що використовує теорію ймовірностей для обробки числових одиниць інформації як наслідків експерименту, називають *математичною статистикою*. Послідовність операцій, виконуваних з додержанням певного комплексу умов, називають *експериментом.* Наслідок будь-якого експерименту називають *подією*.

Якщо в результаті експерименту, здійснюваного з додержанням певного комплексу умов, певна подія обов'язково настає, то вона називається *достовірною*. Достовірна подія позначається символом Ω («омега»).

Приклад. Усі вагони потяга, який підлягає розформуванню, обов'язково пройдуть першу розділову стрілку сортувальної гірки – це достовірна подія.

Подія називається *неможливою*, якщо в результаті експерименту, проведеного з додержанням певного комплексу умов, вона не настає ніколи. Неможлива подія позначається символом Ø (порожня множина).

Приклад. Одночасне зайняття одним потягом двох колій – це неможлива подія.

Безпосередні наслідки експерименту називають *елементарними* подіями. Елементарні події розглядаються як нерозкладні та взаємовиключні події. Елементарні події позначаються $\omega_i \, (i = 1,2,3,...)$.

Приклад. Монету підкидають один раз. Визначити елементарні події.

Розв'язання. $\omega_1 = \{Г\} = \{монета\ випаде\ гербом\}$,

$\omega_2 = \{Ц\} = \{монета\ випаде\ цифрою\}$.

Випадкова подія називається *складною,* якщо її можна розкласти на прості (елементарні) події.

Приклад. Задано множину чисел $\Omega = \{1,2,3,...,10\}$. Навмання із цієї множини беруть одне число. Побудувати подію $A = \{з'явиться\ парне\ число\}$.

Розв'язання. $A = \{2,4,6,8,10\}$.

Кожному експерименту з випадковими результатами відповідає певна множина Ω елементарних подій $\omega_i \, (i = 1,2,3,...)$, кожна з яких може відбуватися внаслідок його проведення $(\omega_i \in \Omega)$. Таку множину називають *простором елементарних подій.*

Приклад. Монету підкидають три рази. Побудувати: а) простір елементарних подій для цього експерименту; б) випадкову подію $A = \{герб\ випаде\ двічі\}$; в) випадкову подію $B = \{герб\ випаде\ не\ менш\ \ як\ двічі\}$.

Розв'язання. а) $\Omega = \{ГГГ; ГГЦ; ГЦГ; ЦГГ; ГЦЦ; ЦГЦ; ЦЦГ; ЦЦЦ\}$;

б) $A = \{ГГЦ; ЦГГ; ГЦГ\}$; в) $B = \{ГГЦ; ЦГГ; ГЦГ; ГГГ\}$.

Завдання

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

Відповідь. $\Omega = \{11;12;21;22;31;32;41;42\}$.

2. Гральний кубик підкидається один раз. Побудувати:

а) простір елементарних подій для цього експерименту;

б) подію $A = \{\textit{випаде число, кратне } 2\}$;

в) подію $B = \{\textit{випаде число, кратне } 3\}$.

<u>Відповідь.</u> $\Omega = \{1;2;3;4;5;6\}$; $A = \{2;4;6\}$; $B = \{3;6\}$.

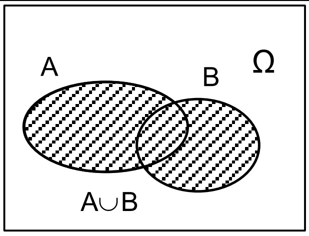
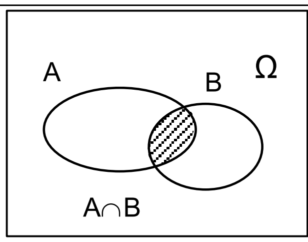
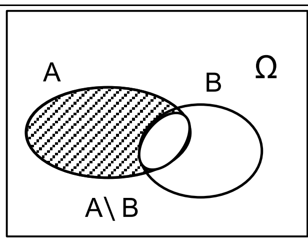
3. Відбувається експеримент з одночасного кидання монети та грального кубика. Вважаючи подією сукупний результат випадання монети та кубика, побудувати простір елементарних подій Ω для цього експерименту.

<u>Відповідь.</u> $\Omega = \{Г1; Г2; Г3; Г4; Г5; Г6; Ц1; Ц2; Ц3; Ц4; Ц5; Ц6\}$.

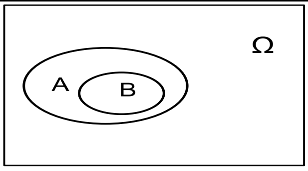
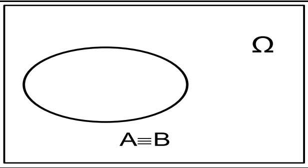
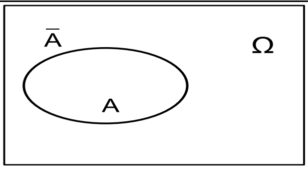
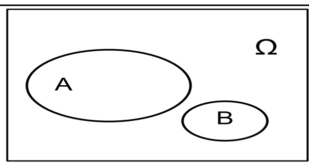
1.2. Операції над подіями

Означення операцій над випадковими подіями та їх схематичне зображення наведені у табл. 1.1. Види подій з точки зору співвідношення між ними наведені у табл. 1.2.

Таблиця 1.1

Математична операція	Схематичне зображення
1. *Додавання*. Сумою двох подій A і B називається така подія $C = A \cup B$ $(C = A + B)$, яка полягає в появі принаймні однієї з подій A або B	
2. *Множення*. Добутком двох подій A і B називається така подія $C = A \cap B$ $(C = AB)$, яка полягає в сумісній появі A і B	
3. *Віднімання*. Різницею двох подій A і B називається така подія $C = A \setminus B$ $(C = A - B)$, яка полягає в появі події A і одночасним ненастанням події B	

Таблиця 1.2

Означення	Схематичне зображення
1. Якщо подія A обов'язково відбувається, коли відбувається подія B, то будемо називати подію A *наслідком* події B $(A \supset B$ *або* $B \subset A)$	
2. Якщо подія A є наслідком події B, а подія B є наслідком події A, то ці події *збігаються*	
3. Подія, зміст якої полягає в тому, що подія A не відбувалася, називається подією, *протилежною* до A і позначається $\overline{A}$. Подія $\overline{A}$ складається з тих подій Ω, що не увійшли в A	
4. Події A і B називаються *несумісними*, якщо одночасне їх настання неможливе, тобто $A \cap B = \varnothing$ $(AB = \varnothing)$	

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

Приклади

1. До потрібного пункту можна доїхати потягами двох призначень. Наявність квитків у касі на один потяг (подія A) не виключає наявності квитків на другий потяг (подія B). Як наслідок, події A і B − сумісні.

2. У групі вагонів, які стоять на станції, є вагони на вантажний двір, завод і фабрику. Вагон, призначений на вантажний двір – подія A, на завод – подія B, на фабрику – подія C. Події A, B, C − несумісні.

Події називаються *єдиноможливими*, якщо поява внаслідок експерименту однієї і тільки однієї з них є достовірною подією.

Приклад. Із трьох локомотивів, що працюють на коліях формування потягів, жодний незайнятий – подія A; тільки один зайнятий – подія B; тільки два зайняті – подія C; три зайняті – подія D. Події A, B, C, D − єдиноможливі.

Випадкові події $A_1, A_2, ..., A_n$ утворюють *повну групу подій,* якщо:

1) вони попарно несумісні, тобто $A_i \cap A_j = \varnothing$ при $i \neq j$;

2) настання хоча б однієї з подій $A_1, A_2, ..., A_n$ є достовірною подією, тобто $A_1 \cup A_2 \cup A_3 \cup ... \cup A_n = \bigcup_{i=1}^{n} A_i = \Omega$.

Операції додавання, множення та віднімання випадкових подій мають такі властивості:

1. $A \cup A = A$; $A \cap A = A$.
2. $A \cup B = B \cup A$; $A \cap B = B \cap A$.
3. $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$.
4. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$; $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.
5. $A \cup \Omega = \Omega$; $A \cap \Omega = A$; $A \cup \varnothing = A$; $A \cap \varnothing = \varnothing$.
6. $\overline{A} = \Omega \setminus A$; $\overline{\Omega} = \varnothing$.
7. $A \cup (A \cap \overline{B}) = A$; $\overline{A \cup B} = \overline{A} \cap \overline{B}$; $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

Завдання

1. Задано множину цілих чисел $\Omega = \{1,2,3,...,15\}$. Навмання з цієї множини беруть одне число. Побудувати випадкові події: а) A={узяте число кратне 2}; б) B={узяте число кратне 3}. Визначити $A \cup B, A \cap B, A \setminus B$.

<u>Відповідь.</u> $A = \{2;4;6;8;10;12;14\}$; $B = \{3;6;9;12;15\}$.
$A \cup B = \{2;3;4;6;8;9;10;12;14;15\}$; $A \cap B = \{6;12\}$;
$A \setminus B = \{2;4;8;10;14\}$.

2. Подія B є частинним випадком події A. Чому дорівнюють їх сума та добуток?

<u>Відповідь.</u> $A + B = A$; $AB = B$.

3. Довести, що $\overline{\overline{A}\,\overline{B}} = A + B$ і $\overline{\overline{C} + \overline{D}} = CD$.

4. Довести, що $A + B = A + \overline{A}B$.

1.3. Елементи комбінаторики

Комбінаторика – самостійний розділ математики, що вивчає різні способи поєднань деяких елементів.

Принципи комбінаторики

1. Якщо елемент A можна вибрати n способами, а елемент B − m несумісними способами, то вибрати або A, або B можна $n+m$ способами.
2. Якщо елемент A можна вибрати n способами, а елемент B при довільному виборі A − m способами, то вибір пари елементів і A, і B може бути здійснений $n \times m$ способами.

Множину називають *упорядкованою*, якщо при її побудові істотним є порядок розміщення елементів. У противному разі множину називають *невпорядкованою*.

Перестановки

Нехай маємо множину S, що складається з n різних елементів $s_1, s_2, ..., s_n$. Якщо переставляти елементи місцями різними способами, не вилучаючи їх і не додаючи інших, отримаємо ряд послідовностей, які відрізняються між собою порядком розташування елементів:

$$s_1, s_2, s_3, ..., s_n;\ s_2, s_1, s_3, ..., s_n; ...;\ s_n, ..., s_3, s_2, s_1.$$

Кожна з цих послідовностей відрізняється від усіх інших перестановкою принаймні двох елементів і називається *перестановкою* з n елементів. Кількість *перестановок* з n різних елементів визначається за формулою:

$$P_n = 1 \cdot 2 \cdot ... \cdot (n-1) \cdot n = n! \qquad (1.1)$$

Якщо серед елементів $s_1, s_2, s_3, ..., s_n$ множини є однакові елементи, причому s_1 зустрічається n_1 разів, s_2 − n_2 разів, s_k − n_k разів, причому $n_1 + n_2 + ... + n_k = n$, то загальна кількість різних *перестановок з повтореннями* обчислюється за формулою

$$P_n(n_1, n_2, ..., n_k) = \frac{n!}{n_1! \cdot n_2! \cdot ... \cdot n_k!}. \qquad (1.2)$$

Комбінації

Нехай маємо множину S, яка складається з n різних елементів. Будь-яка підмножина, що містить k елементів $(k = 0,1,2,...,n)$ називається *комбінацією* з даних n елементів по k елементів. Різними комбінаціями вважаються такі, що відрізняються принаймні одним елементом, а порядок до уваги не береться.

Кількість комбінацій з n по k елементів позначається символом C_n^k $(0 \le k \le n)$ і обчислюється за формулою

$$C_n^k = \frac{n!}{k!(n-k)!}. \qquad (1.3)$$

За домовленістю вважається, що $0! = 1$. В частинних випадках маємо: $C_n^0 = 1$ і $C_n^n = 1$. Крім того, справедливі такі властивості кількості комбінацій:

$$C_n^m = C_n^{n-m}; \quad C_n^m + C_n^{m+1} = C_{n+1}^{m+1}; \quad C_n^0 + C_n^1 + C_n^2 + ... + C_n^n = 2^n. \qquad (1.4)$$

Якщо в комбінаціях з n елементів по k деякі з елементів (або всі) можуть бути однакові, то такі комбінації називаються *комбінаціями з повтореннями* з n елементів по k. Кількість комбінацій з повтореннями з n елементів по k обчислюється за формулою

$$\overline{C}_n^k = C_{n+k-1}^k = \frac{(n+k-1)!}{k!(n-1)!}. \qquad (1.5)$$

Розміщення

Кожна впорядкована множина, яка містить k елементів $(0 \le k \le n)$ множини S, що складається з n елементів, називається *розміщенням* з n елементів по k елементів.

Кількість можливих *розміщень з n елементів по k* елементів, що відрізняються між собою або елементами, або порядком їх розміщення, позначається A_n^k і визначається за формулою

$$A_n^k = \frac{n!}{(n-k)!}. \qquad (1.6)$$

Якщо в розміщеннях з n елементів по k деякі з елементів (або всі) можуть бути однакові, то такі розміщення називають *розміщеннями з повтореннями* з n елементів по k елементів. Кількість розміщень з повтореннями з n елементів по k елементів обчислюється за формулою

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

Приклади

1. Розклад одного дня складається з шести уроків з різних дисциплін. Знайти кількість розкладів, що можна скласти на цей день.

Розв’язання. За формулою (1.1) отримаємо $6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720$.

2. В шаховому турнірі беруть участь 18 студентів. Скільки партій має бути зіграно, якщо кожна пара учасників повинна зіграти одну партію?

Розв’язання. За формулою (1.3) обчислюємо

$$C_{18}^2 = \frac{18!}{2! \cdot 16!} = \frac{17 \cdot 18}{2} = 153.$$

3. Розклад одного дня складається з п’ять різних уроків. Знайти кількість розкладів, які можна скласти з 11 дисциплін.

Розв’язання. За формулою (1.6) отримаємо

$$A_{11}^5 = \frac{11!}{6!} = 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 = 55440.$$

4. У конкурсі беруть участь 10 кінофільмів, що проходять по п’яти номінаціях. Знайти кількість варіантів розподілу призів, якщо по кожній номінації встановлено: а) різні призи; б) однакові призи.

Розв’язання: а) за формулою (1.7) $\overline{A}_{10}^{5} = 10^5 = 100000$; б) за формулою (1.5) $\overline{C}_{10}^{5} = C_{10+5-1}^{5} = C_{14}^{5} = \frac{10 \cdot 11 \cdot 12 \cdot 13 \cdot 14}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 2002$.

5. Скільки існує семизначних чисел, складених з цифр 4,5,6, якщо цифра 4 повторюється три рази, а цифри 5 і 6 − по два рази?

Розв’язання. За формулою (1.2) отримаємо

$$P_7(3;2;2) = \frac{7!}{3! \cdot 2! \cdot 2!} = 210.$$

Завдання

1. Скількома способами можна розкласти на полиці чотири різні книги?

Відповідь. 24.

2. Скількома способами можна скласти букет з п’яти квітів, якщо є квіти трьох сортів?

Відповідь. 21.

3. Електричка має вісім вагонів. Скільки існує способів розміщення чотирьох пасажирів, якщо в кожному вагоні повинно бути не більш одного пасажира?

Відповідь. 1680.

4. З чотирьох першокурсників, п’ятьох другокурсників та шести третьокурсників треба обрати три студенти на конференцію. Скількома способами це можна зробити, якщо серед обраних повинні бути студенти різних курсів?

Відповідь. 120.

5. Скількома способами можна розставити на полиці сім різних книг, щоб окремі три стояли поряд?

Відповідь. 720.

6. Десять крісел поставлено в ряд: а) скількома способами на них можуть сісти дві людини; б) скількома способами вони можуть сісти поряд; в)скількома способами вони можуть сісти так, щоб між ними було хоча б одне крісло?

Відповідь: а) 90; б)18; в) 72.

7. Скільки чотиризначних чисел можна скласти, використавши цифри 1,2,3,4,5, якщо: а) жодна цифра не повторюється більш одного разу; б) повтор цифр допустимий; в) число має бути непарним і повторів бути не повинно.

Відповідь: а) 120; б) 625; в) 72.

8. В одного учня є 6 різних книжок для обміну, а в іншого − 17. Скількома способами вони зможуть здійснити обмін: а) книга на книгу; б) дві книги на дві книги?

Відповідь: а) 102; б) 2040.

9. В урні міститься 10 білих і 6 чорних кульок. Скількома способами можна взяти 5 кульок так, щоб серед них було: а) 5 чорних; б) 3 білі та 2 чорні?

Відповідь: а) 6; б) 1800.

10. Скількома способами можна розподілити 20 видів продукції, якщо на перший склад потрібно 9, на другий – 3, на третій – 8 видів продукції?

Відповідь. 27713400.

11. Дано три цифри: 1; 2; 3. Визначити, скільки дев'ятизначних чисел можна скласти з цифр за умови, що в кожне з цих чисел цифра 1 входить тричі, цифра 2 – чотири рази, і цифра 3 – двічі.

Відповідь. 1260.

12. У партії з 20 виробів міститься 16 стандартних виробів. Навмання вибирається 6 виробів. Визначити, скількома способами можна вибрати при цьому чотири стандартні вироби і два нестандартні.

Відповідь. 10920.

13. Скількома способами можна з колоди в 32 карти взяти 10 карт так, щоб 8 з них були однієї масті?

Відповідь. $4 \cdot C_{24}^{2}$.

14. З 52 делегатів конференції треба обрати президію з 5 чоловік і делегацію з 3 чоловік. Скількома способами можна здійснити вибір, якщо: а) члени президії можуть входити до складу делегації; б) члени президії не повинні входити до складу делегації?

Відповідь: а) $C_{52}^{5} \cdot C_{52}^{3}$; б) $C_{52}^{5} \cdot C_{47}^{3}$.

15. Скількома способами можна розділити 15 різних предметів між трьома особами так, щоб кожна особа одержала по 5 предметів?

Відповідь. $C_{15}^{5} \cdot C_{10}^{5}$.

16. На залізничній станції є n світлофорів. Скільки різних сигналів можна подати за їхньою допомогою, якщо кожний світлофор має два сигнали – червоний і зелений?

Відповідь. 2^{n}.

17. Два кур'єри повинні віднести 10 листів. Скількома способами вони можуть розподілити між собою цю роботу?

Відповідь. $2^{10} = 1024$.

18. В магазині є гвоздики трьох кольорів. Скільки різних букетів по п'ять квітів можна скласти?

Відповідь. $\overline{C}_3^{5} = C_7^{5} = 21$.

18. На станції працюють 1000 працівників. Довести, що принаймні двоє з них мають однакові ініціали.

1.4. Класичне означення ймовірності

Ймовірністю випадкової події A називається невід'ємне число $P(A)$, що дорівнює відношенню числа елементарних подій m $(0 \le m \le n)$, які сприяють події A, до кількості всіх рівноможливих несумісних елементарних подій, що утворюють повну групу:

$$P(A) = \frac{m}{n}. \tag{1.8}$$

Властивості ймовірності події:

1) $0 \le P(A) \le 1$;
2) $P(\Omega) = 1$ − ймовірність достовірної події дорівнює 1; (1.9)
3) $P(\varnothing)=0$ – ймовірність неможливої події дорівнює 0.

Зауваження. Використання класичної моделі ймовірності є обмеженим через те, що весь простір елементарних подій не завжди вдається розбити на рівноможливі події.

Приклади

1. На кожній із п'яти однакових карток записано одну з літер: Я, П, Г, Т,О. Яка ймовірність того, що: а) три картки, взятих навмання, утворять слово «ТОП»; б) картки, навмання розкладені в рядок, утворять слово «ПОТЯГ»?

Розв'язання:

а) A={отримання слова «ТОП»}. За класичним означенням ймовірності

$$P(A)=\frac{m}{n}=\frac{1}{A_5^3}=\frac{1}{3\cdot4\cdot5}=\frac{1}{60};$$

б) $B=${отримання слова «ПОТЯГ»}.

$$P(A)=\frac{m}{n}=\frac{1}{P_5}=\frac{1}{5!}=\frac{1}{120}.$$

2. На картках написано літери Ж, І, К, Н, Т, дві літери В і три літери А. Знайти ймовірність того, що картки, навмання розкладені в рядок, утворять слово «ВАНТАЖІВКА».

Розв'язання. A={отримання слова «ВАНТАЖІВКА»}.

$$P(A)=\frac{m}{n}=\frac{P_3\cdot P_2}{P_{10}}=\frac{3!\cdot2!}{10!}=\frac{1}{302400}.$$

3. Дев'ять пасажирів навмання розміщуються у трьох вагонах. Обчислити ймовірність таких випадкових подій: а) A – у кожному вагоні виявиться по три пасажири; б) B– в один з вагонів сядуть 4, у другий – 3 і в третій – 2 пасажири.

Розв'язання:

а) $P(A)=\dfrac{C_9^3\cdot C_6^3\cdot C_3^3}{3^9}=\dfrac{1}{3^9}\cdot\left(\dfrac{9!}{3!\cdot6!}\cdot\dfrac{6!}{3!\cdot3!}\cdot1\right)=\dfrac{1680}{3^9}\approx0{,}085;$

б) $P(B)=\dfrac{C_9^4\cdot C_5^3\cdot C_2^2\cdot3!}{3^9}=\dfrac{1}{3^9}\cdot\left(\dfrac{9!}{4!\cdot5!}\cdot\dfrac{5!}{3!\cdot2!}\cdot1\cdot3!\right)=\dfrac{7560}{3^9}\approx0{,}384.$

Завдання

1. Монета підкидається 10 раз. Яка ймовірність того, що герб з'явиться чотири рази?

Відповідь. $\dfrac{C_{10}^{4}}{2^{10}} = 0{,}2051.$

2. Числа 1,2,3,4 написані на чотирьох однакових картках. Навмання послідовно по одній вибирають три картки й розкладають їх у рядок. Яка ймовірність того, що при цьому утвориться парне трицифрове число?

Відповідь. $\dfrac{2 \cdot A_{3}^{2}}{A_{4}^{3}} = 0{,}5.$

3. Гральний кубик підкидається один раз, а монета чотири рази. Знайти ймовірність таких подій: а) A − на гральному кубику з'явиться число, кратне двом, і герб при цьому випаде двічі; б) B − на гральному кубику з'явиться число, кратне трьом, і герб при цьому випаде тричі.

Відповідь: а) $\dfrac{3}{16}$; б) $\dfrac{1}{12}$.

4. За умовою лотереї «Спортлото 6 з 45» учасник, відгадавши чотири, п'ять, шість видів спорту з відібраних при випадковому розіграші 6 видів спорту з 45, отримує грошовий приз. Знайти ймовірність того, що будуть угадані: а) всі шість чисел; б) чотири числа; в) одне число.

Відповідь: а) $\dfrac{1}{C_{45}^{6}}$; б) $\dfrac{C_{6}^{4} \cdot C_{39}^{2}}{C_{45}^{6}}$; в) $\dfrac{C_{6}^{1} \cdot C_{39}^{5}}{C_{45}^{6}}$.

5. З урни, де міститься 10 сигнальних прапорців (4 зе-лених та 6 червоних), навмання було вибрано 7 прапорців. Визначити ймовірність того, що серед вибраних є 2 зелених.

Відповідь. 0,3.

6. У ліфт на першому поверсі дев'ятиповерхового будинку ввійшли чотири пасажири, кожен з яких може вийти незалежно один від одного на будь-якому поверсі з другого по дванадцятий. Знайти ймовірність того, що всі пасажири: а) вийдуть на п'ятому поверсі; б) на одному поверсі.

Відповідь: а) 0,00015; б) 0,00137.

7. У групі з 50 вагонів, що подають під навантаження, шість – пошкоджені. Групу навмання поділили навпіл та відправили до двох вагонних депо. Знайти ймовірність того, що: а) всі пошкоджені вагони потраплять до одного вагонного депо; б) вагонні депо отримають однакову кількість пошкоджених вагонів; в) одне вагонне депо отримає два, а інше – чотири пошкоджених вагони.

Відповідь: а) $\frac{C_6^6 \cdot C_{44}^{19} + C_6^0 \cdot C_{44}^{25}}{C_{50}^{25}} \approx 0{,}02$; б) $\frac{C_6^3 \cdot C_{44}^{22}}{C_{50}^{25}} \approx 0{,}057$;

в) $\frac{C_6^2 \cdot C_{44}^{23} + C_6^4 \cdot C_{44}^{21}}{C_{50}^{25}} \approx 0{,}005$.

8. 10 чоловік: а) сідають за круглий стіл; б) стають в чергу за білетами. Знайти ймовірність того, що певні дві особи випадково опиняться поруч.

Відповідь: а) $\frac{10 \cdot 2! \cdot 8!}{10!}$; б) $\frac{9 \cdot 2! \cdot 8!}{10!}$.

1.5. Статистичне означення ймовірності

Відносною частотою $w(A)$ випадкової події A називається відношення кількості експериментів m, при яких подія A спостерігалася, до загальної кількості n проведених експериментів:

$$w(A) = \frac{m}{n}. \tag{1.10}$$

Очевидно, для відносної частоти виконується нерівність: $0 \le w(A) \le 1$.

Ймовірністю випадкової події називається число, біля якого групуються значення відносних частот цієї події $w_i(A)$, коли необмежено збільшується кількість експериментів:

$$P(A) = \lim_{n \to \infty} w(A). \tag{1.11}$$

Перевага статистичного означення ймовірності полягає в тому, що в його основі лежить експеримент. При цьому маємо й недоліки: необхідність великої кількості експериментів для впевненого визначення ймовірності.

Завдання

При стрільбі з гвинтівки по мішені відносна частота влучення дорівнює 0,95. Знайти число влучень, якщо було здійснено 20 пострілів.

Відповідь. 19.

1.6. Геометричне означення ймовірності

Класичним означенням імовірності неможливо користуватися, коли кількість елементарних подій експерименту нескінченна, а також коли неможливо подію розбити на рівноможливі, несумісні елементарні події.

Якщо множина Ω є неперервною і квадровною, то ймовірність події $A (A \subset \Omega)$ обчислюється за формулою

$$P(A) = \frac{mes(A)}{mes(\Omega)}. \qquad (1.12)$$

Приклад. Внаслідок буревію стався обрив проводу телефонної мережі на ділянці між двома залізничними станціями, відстань між якими складає 30 км. Знайти ймовірність того, що обрив стався між кілометровими стовпчиками з позначками 10-й та 15-й кілометр?

Розв'язання. Нехай відрізок l являється частиною відрізка L. На відрізку L навмання відмічається точка x. Ймовірність того, що $x \in l$, пропорційна довжині відрізка l і не залежить від його розміщення відносно відрізка L (рис. 1.1).

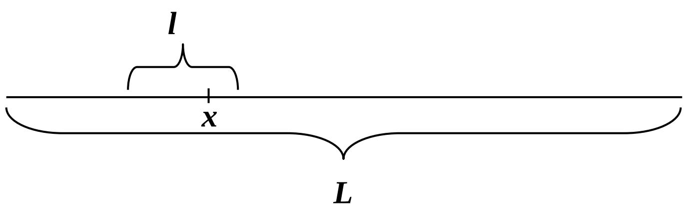


Рис. 1.1

Ймовірність попадання x на відрізок l обчислюється за формулою (1.12)

$$P = \frac{\textit{довжина}\, l}{\textit{довжина}\, L} = \frac{15 - 10}{30} = \frac{1}{6}.$$

Завдання

1. По нафтопроводу між пунктами А і В, відстань між якими 3 км, перекачують нафту. Яка ймовірність того, що пошкодження через деякий час T роботи станеться на виділеній ділянці довжиною 150 м?

Відповідь. 0,05.

2. Дві подруги домовились про зустріч біля фонтану між 10 та 11 годинами. Кожна подруга приходить у призначене місце незалежно одна від другої в будь-який момент часу та чекає протягом указаної години. Прийшовши на зустріч, подруга чекає не більше 10 хвилин і йде не пізніше 11 годин. Знайти ймовірність того, що подруги зустрінуться.

Відповідь. $\frac{11}{36}$.

3. У мішень, яка має вигляд кола, вписано квадрат. По ній здійснюється один постріл. Вважається при цьому, що влучення в коло мішені є подією достовірною. Яка ймовірність того, що куля влучить в квадрат.

Відповідь. $\frac{2}{\pi}$.

4. Відстань між пунктами А і В автобус проходить за τ хвилин, а пішохід за t хвилин $(t > \tau)$. Інтервал руху автобуса T хвилин. У деякий момент часу пішохід виходить із А. Знайти ймовірність події: дорогою пішохода дожене черговий автобус.

Відповідь. $\begin{cases} \dfrac{t-\tau}{T} & \textit{при} \quad t-\tau \le T, \\ 1 & \textit{при} \quad t-\tau \ge T. \end{cases}$

5. Стрижень довжиною l розламано у двох будь-яких обраних точках. Знайти ймовірність того, що з отриманих відрізків можна скласти трикутник.

Відповідь. $\frac{1}{4}$.

6. На відрізку $[0;2]$ випадково беремо дві точки з абсцисами x, y. Знайти ймовірність того, що відстань між ними буде менше одиниці.

Відповідь. $\frac{3}{4}$.

7. Два судна повинні підійти до одного причалу. Появи суден – незалежні випадкові події, рівноможливі протягом доби. Знайти ймовірність того, що одному з суден доведеться чекати звільнення причалу, якщо час стоянки першого судна – одна година, а другого – дві години.

Відповідь. 0,121.

Вказівка. $\Omega = \{(x, y) : 0 \le x \le 24; 0 \le y \le 24\}$;
$A = \{(x, y) \in \Omega : \ x \le y \le x + 1; y \le x \le y + 2\}$.

1.7. Теорема додавання ймовірностей

Нагадаємо, що *несумісними* подіями називаються такі події, коли настання однієї з них повністю виключає настання іншої.

Теорема 1.1. Ймовірність появи однієї з двох несумісних подій, байдуже якої, дорівнює сумі ймовірностей цих подій:

$$\underline{P(A+B) = P(A) + P(B).} \tag{1.13}$$

Наслідок. Ймовірність появи однієї з декількох попарно несумісних подій, байдуже якої, дорівнює сумі ймовірностей цих подій:

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n). \tag{1.14}$$

Зокрема, для трьох подій:

$$P(A+B+C) = P(A) + P(B) + P(C). \tag{1.15}$$

Приклад. Кожен студент групи проходить практику на одній із чотирьох залізничних станцій. Таким чином, на першій станції проходять практику сім студентів, на другій, третій і четвертій – дев'ять, п'ять та два студенти відповідно. Знайти ймовірність того, що перші три студенти, що з'явились на залік з практики, проходили практику на одній станції.

Розв'язання. A={три студенти з першої станції}; B={три студенти з другої станції}; C={три студенти з третьої станції}; D={три студенти з четвертої станції}.

$$P(A)=\frac{C_7^3}{C_{23}^3};\ P(B)=\frac{C_9^3}{C_{23}^3};\ P(C)=\frac{C_5^3}{C_{23}^3};\ P(D)=0 \text{ (неможлива подія).}$$

Події A, B, C, D є несумісними. Таким чином, за формулою (1.14) обчислюємо:

$$P(A+B+C+D)=P(A)+P(B)+P(C)+P(D)=\frac{C_7^3+C_9^3+C_5^3}{C_{23}^3}=\frac{129}{1771}\approx 0{,}073.$$

Теорема 1.2. Сума ймовірностей подій $A_1, A_2, ..., A_n$, що утворюють повну групу, дорівнює одиниці:

$$\underline{P(A_1)+P(A_2)+...+P(A_n)=1.} \qquad (1.16)$$

Приклад. Куплено лотерейний квиток. Відомо, що ймовірність виграшу 0,05. Знайти ймовірність програшу.

Розв'язання. A={квиток виграшу}; B={квиток програшу}. Інших подій не існує. Таким чином, $P(A)+P(B)=1 \Rightarrow P(B)=1-P(A)=1-0{,}05=0{,}95.$

Протилежними називають дві єдиноможливі події, що утворюють повну групу. Таким чином, з теореми 1.2 випливає теорема 1.3.

Теорема 1.3.

$$\underline{P(A)+P(\overline{A})=1.} \qquad (1.17)$$

Наслідок. $\underline{P(A)=1-P(\overline{A})}.$

Зауваження. Якщо ймовірність однієї з двох протилежних подій позначити p, то ймовірність іншої події позначають q. Таким чином отримаємо

$$p + q = 1. \tag{1.18}$$

1.8. Залежні та незалежні випадкові події. Умовна ймовірність. Теорема добутку ймовірностей

Події A і B називаються *незалежними*, якщо ймовірність однієї, будь-якої з них, не залежить від того, відбувається чи ні інша подія. У протилежному випадку події A і B називають *залежними*.

Приклади

1. До сортувальної станції примикають три дільниці. Подія прибуття на станцію в даний момент часу потяга на розформування з однієї із дільниць не залежить від того, чи прибули в цей момент часу такі потяги з інших дільниць. Це незалежні події. Вважається, що кількість колій у парку приймання дозволяє прийняти одночасно потяги з усіх дільниць.

2. У групі з 10 вагонів три вагони не можна використовувати для навантаження. Навмання беруть один вагон, а потім, не повертаючи його, другий. Подія A – узятий з першого разу вагон непридатний для навантаження. Подія B – взятий з другого разу вагон також непридатний для навантаження. Ймовірність появи події B за умови, що відбулася подія A, дорівнює: $P(B) = \frac{3-1}{10-1} = \frac{2}{9}$. Якщо узятий з першого разу вагон виявиться придатним для навантаження, то $P(B) = \frac{3}{10-1} = \frac{3}{9}$. Таким чином, ймовірність появи події B залежить від того, відбулася чи ні подія A. Як наслідок, події A і B залежні.

Умовною ймовірністю $P_A(B)$ називають ймовірність події B, обчисленої за припущенням, що подія A вже відбулася:

$$P_A(B)=\frac{P(AB)}{P(A)}, \quad P(A)>0. \tag{1.19}$$

Аналогічно

$$P_B(A)=\frac{P(AB)}{P(B)}, \quad P(B)>0. \tag{1.20}$$

Властивості умовної ймовірності:

1) $P_A(B)=0$, якщо $A\cap B=\varnothing$;

2) $P_A(B)=1$, якщо $A\cap B=A$;

3) у решті випадків $0<P_A(B)<1$.

Теорема 1.4. Ймовірність сумісної появи двох подій дорівнює добутку ймовірності однієї з них на умовну ймовірність іншої, яка обчислюється за умови, що перша подія вже відбулася:

$$\underline{P(AB)=P(A)\cdot P_A(B)} \quad \text{або} \quad \underline{P(AB)=P(B)\cdot P_B(A)}. \tag{1.21}$$

Наслідок. Ймовірність сумісної появи декількох подій дорівнює добутку ймовірності однієї з них на умовні ймовірності всіх інших, причому ймовірність кожної наступної події обчислюється за припущенням, що всі попередні події вже відбулися:

$$P(A_1A_2A_3...A_n)=P(A_1)\cdot P_{A_1}(A_2)\cdot P_{A_1A_2}(A_3)\cdot...\cdot P_{A_1A_2...A_{n-1}}(A_n). \tag{1.22}$$

Зокрема, для трьох подій

$$P(ABC)=P(A)\cdot P_A(B)\cdot P_{AB}(C). \tag{1.23}$$

Приклад. В коробці знаходяться чотири білі, три сині та дві чорні кульки. Навмання послідовно беруть три кульки. Яка ймовірність того, що перша кулька – біла, друга – синя, третя – чорна?

Розв'язання. $A=\{$перша витягнута кулька біла$\}$; $B=\{$друга витягнута кулька синя$\}$; $C=\{$третя витягнута кулька чорна$\}$. Тоді за формулою (1.23) отримаємо:

$$P(ABC)=P(A)\cdot P_A(B)\cdot P_{AB}(C)=\frac{4}{9}\cdot\frac{3}{8}\cdot\frac{2}{7}=\frac{1}{21}.$$

Нагадаємо, що події A і B називаються *незалежними*, якщо:

$$P_B(A)=P(A) \quad \text{або} \quad P_A(B)=P(B). \tag{1.24}$$

Тоді з формули (1.21) випливає теорема 1.5.

Теорема 1.5. Ймовірність сумісного настання двох незалежних подій дорівнює добутку ймовірностей цих подій:

$$\underline{P(A\cdot B)=P(A)\cdot P(B).} \tag{1.25}$$

Декілька подій називаються *незалежними у сукупності*, якщо незалежні кожні дві з них і незалежна кожна подія та всі можливі добутки останніх.

Наслідок. Ймовірність сумісної появи декількох подій, незалежних в сукупності, дорівнює добутку ймовірностей цих подій:

$$P(A_1A_2A_3...A_n)=P(A_1)\cdot P(A_2)\cdot P(A_3)\cdot...\cdot P(A_n). \tag{1.26}$$

Зокрема, для трьох подій формула (1.26) має вигляд

$$P(ABC)=P(A)\cdot P(B)\cdot P(C). \tag{1.27}$$

Приклади

1. Гральний кубик і монету підкидають по одному разу. Яка ймовірність того, що при цьому на грані кубика випаде парне число, а на монеті герб?

Розв’язання. $A=\{$поява парного числа на грані кубика$\}$; $B=\{$поява герба$\}$. Випадкові події A і B незалежні між собою. Отже:

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

За формулою (1.25) обчислюємо:

$$P(A\cdot B)=P(A)\cdot P(B)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}.$$

2. Три працівники станції складають іспит з правил технічної експлуатації залізниць. Ймовірність того, що перший складе іспит – 0,9, для другого та третього працівників ймовірності дорівнюють відповідно 0,85 і 0,75. Обчислити ймовірності подій: а) три працівники складуть іспит; б) три працівники не складуть іспит; в) два працівники складуть іспит.

<u>Розв'язання:</u> а) $A=\{$три працівники складуть іспит$\}$. Позначимо A_1, A_2, A_3 − випадкові події, які полягають у тому, що перший, другий і третій працівники відповідно складуть іспити. Тоді, $\overline{A_1}, \overline{A_2}, \overline{A_3}$ − відповідно не складуть. Відомо, що $P(A_1)=0{,}9;$ $P(A_2)=0{,}85;$ $P(A_3)=0{,}75.$ Знаходимо ймовірності протилежних подій:

$$P(\overline{A_1})=1-P(A_1)=1-0{,}9=0{,}1;$$
$$P(\overline{A_2})=1-P(A_2)=1-0{,}85=0{,}15;$$
$$P(\overline{A_3})=1-P(A_3)=1-0{,}75=0{,}25.$$

Тоді $P(A)=P(A_1A_2A_3)=P(A_1)\cdot P(A_2)\cdot P(A_3)=0{,}9\cdot 0{,}85\cdot 0{,}75\approx 0{,}574$;

б) $B=\{$три працівники не складуть іспит$\}$.

$$P(B)=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=0{,}1\cdot 0{,}15\cdot 0{,}25=0{,}003785\,;$$

в) $C=\{$два працівники складуть іспит$\}$. За формулами (1.13) та (1.25)

$$P(C)=P(A_1A_2\overline{A_3})+P(A_1\overline{A_2}A_3)+P(A_1\overline{A_2}A_3)+P(\overline{A_1}A_2A_3)=$$
$$=P(A_1)\cdot P(A_2)\cdot P(\overline{A_3})+P(A_1)\cdot P(\overline{A_2})\cdot P(A_3)+P(\overline{A_1})\cdot P(A_2)\cdot P(A_3)=$$
$$=0{,}9\cdot 0{,}85\cdot 0{,}25+0{,}9\cdot 0{,}15\cdot 0{,}75+0{,}1\cdot 0{,}85\cdot 0{,}75=0{,}356.$$

Нагадаємо, що дві події називаються *сумісними*, якщо поява однієї з них не виключає появу іншої.

Теорема 1.6. Ймовірність появи хоча б однієї з двох сумісних подій дорівнює сумі ймовірностей цих подій без ймовірності їх сумісної появи:

$$\underline{P(A+B)=P(A)+P(B)-P(AB).} \qquad (1.28)$$

Приклади

1. Знайти ймовірність того, що навмання взяте двозначне число буде кратне 2, або 5, або 2 і 5 одночасно?

Розв'язання. $A = \{$двозначне число, кратне 2$\}$. $B = \{$двозначне число, кратне 5.$\}$ Події A і B сумісні і незалежні. За формулою (1.28) обчислюємо

$$P(A+B) = \frac{45}{90} + \frac{18}{90} - \frac{9}{90} = 0{,}6.$$

2. Зі 100 гальмових башмаків є три несправних. Навмання взяли два башмаки. Знайти ймовірність того, що принаймні один башмак буде несправним.

Розв'язання. A_i = $\{$несправність i-го башмака$\}$, $i = 1{,}2$. За формулою (1.28)

$$P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1 A_2) = P(A_1) + P(A_2) - P(A_1) \cdot P_{A_1}(A_2) =$$
$$= \frac{3}{100} + \frac{3}{100} - \frac{3}{100} \cdot \frac{2}{99} = 0{,}0594.$$

Завдання

1. Студент шукає потрібну йому формулу у трьох книгах. Ймовірність того, що потрібна формула знаходиться у першій, другій й третій книзі, відповідно дорівнює 0,6; 0,7 і 0,8. Знайти ймовірність того, що потрібна формула знаходиться принаймні у двох книгах.

Відповідь. 0,788.

2. З урни, в якій 2 білих та 3 чорних кулі, два гравці по черзі виймають по одній кульці, не повертаючи їх назад. Виграє той, хто раніше витягне білу кулю. Знайти ймовірність того, що виграє перший гравець.

Відповідь. 0,6.

1.9. Ймовірність появи випадкової події принаймні один раз при n незалежних спробах

Проводиться n незалежних експериментів, у кожному з яких може відбутися подія A_i з ймовірністю $P(A_i) = p_i$ або подія $\overline{A_i}$ $(A_i \cap \overline{A_i} = \varnothing,$ $A_i \cup \overline{A_i} = \Omega)$ з ймовірністю $P(\overline{A_i}) = q_i$ $(p_i + q_i = 1)$, $i = 1,2,3,...,n.$ Нехай B − поява події A_i хоча б один раз при n незалежних спробах, тобто ця подія може з'явитися або один раз, або двічі, тричі і так далі, включаючи всі n раз $(i = 1,2,...,n)$. Тоді ймовірність події B обчислюється за формулою

$$P(B) = 1 - P(\overline{A_1}) \cdot P(\overline{A_2}) \cdot ... \cdot P(\overline{A_n}) \qquad (1.29)$$

або

$$\underline{P(B) = 1 - q_1 \cdot q_2 \cdot ... \cdot q_n.} \qquad (1.30)$$

Якщо $P(A_i) = p_i = p = const$, то $q_i = q = const$ $(i = 1,2,...,n)$, тоді

$$\underline{P(B) = 1 - q^n.} \qquad (1.31)$$

Приклад. Прилад складається з чотирьох елементів, що працюють незалежно один від одного. Ймовірність того, що перший елемент не вийде з ладу під час роботи приладу, є величиною сталою і дорівнює 0,95. Для другого, третього і четвертого елементів ця ймовірність дорівнює відповідно 0,9; 0,87; 0,85. Яка ймовірність того, що під час роботи приладу не вийде з ладу хоча б один елемент?

Розв'язання. Ймовірності того, що елементи не вийдуть із ладу, відповідно дорівнюють:

$$p_1 = 0{,}95;\ p_2 = 0{,}9;\ p_3 = 0{,}87;\ p_4 = 0{,}85.$$

Ймовірності того, що ці елементи вийдуть з ладу, дорівнюватимуть відповідно:

$$q_1 = 1 - p_1 = 1 - 0{,}95 = 0{,}05;\ q_2 = 1 - p_2 = 1 - 0{,}9 = 0{,}1;$$
$$q_3 = 1 - p_3 = 1 - 0{,}87 = 0{,}13;\ q_4 = 1 - p_4 = 1 - 0{,}85 = 0{,}15.$$

За формулою (1.30) обчислюємо

$$P(B) = 1 - q_1 \cdot q_2 \cdot q_3 \cdot q_4 = 1 - 0{,}05 \cdot 0{,}1 \cdot 0{,}13 \cdot 0{,}15 \approx 0{,}9999.$$

Завдання

1. В ящику вісім червоних та п’ять синіх ґудзиків. Навмання беруть два ґудзики. Знайти ймовірність того, що ґудзики будуть одного кольору?

Відповідь. $\frac{19}{39}$.

2. Для навантаження використовуються платформи, піввагони та криті вагони. Платформа використовується з ймовірністю 0,9, піввагон – 0,8 і критий вагон – 0,7. Знайти ймовірність того, що буде використано: а) всі три типи вагонів; б) тільки два типи; в) тільки один тип вагонів.

Відповідь: а) 0,504; б) 0,398; в) 0,092.

3. У подачі з 16 вагонів є три критих навантажених вагони. Яка ймовірність того, що всі криті навантажені вагони стоять у хвості подачі? (Прийнято, що всі 16 вагонів завантажені).

Відповідь. 0,0018.

4. Ймовірність безвідмовної роботи блока, що входить у систему впродовж певного часу дорівнює 0,9. Для надійності роботи системи встановлюється такий же блок, що буде знаходитись у резерві. Яка ймовірність безвідмовної роботи системи з урахуванням резервного блока?

Відповідь. 0,99.

5. Ймовірність появи в поїзді вагонів на вантажний двір $p_1 = 0{,}2$, на контейнерну площадку $p_2 = 0{,}3$, на промислове підприємство $p_3 = 0{,}4$. Знайти ймовірність появи в поїзді вагонів: а) на всі три пункти; б) на два пункти; в) на один пункт; г) хоча б на один пункт.

Відповідь: а) 0,024; б) 0,188; в) 0,452; г) 0,664.

6. Вагони з проміжної станції забирають збірним потягом та вивізним локомотивом протягом доби. Збірний потяг їх вивозить

з імовірністю $p_1 = 0{,}6$. Якщо він їх не вивіз, то вивізний локомотив забере їх з імовірністю $p_2 = 0{,}9$. Яка ймовірність того, що після вивізного локомотива вагони не залишаться на станції протягом доби.

Відповідь. 0,96.

7. Студент при підготовці до іспиту з 30 білетів вивчив білети з номерами з 1 по 5 і з 26 по 30. Знайти ймовірність того, що студент витягне вивчений білет, якщо відомо, що на іспиті він витягнув білет з номером, не перевищуючим 20?

Відповідь. $\frac{1}{4}$.

8. Гральний кубик підкидається чотири рази. Чому дорівнює ймовірність того, що цифра 5 з'явиться при цьому хоча б один раз?

Відповідь. $\frac{671}{1296}$.

9. Два мисливці стріляють у вовка. Кожен робить по одному пострілу. Для першого мисливця ймовірність влучання у ціль 0,7, для другого – 0,9. Знайти ймовірність того, що вони влучать у вовка хоча б один раз. Як зміниться результат, якщо мисливці зроблять по два постріли?

Відповідь. 0,97; 0,9991.

10. Гардеробниця одночасно видала номерки трьом чоловікам, що здали свої капелюхи. Після цього вона переплутала всі капелюхи і повісила їх навмання. Знайти ймовірність таких подій: а) кожен отримає власного капелюха; б) двоє отримають свої капелюхи; в) один отримає свій капелюх; г) жоден не отримає свого капелюха.

Відповідь: а) $\frac{1}{6}$; б) 0; в) $\frac{1}{2}$; г) $\frac{1}{3}$.

11. Два стрільці стріляють у мішень по черзі до першого влучення. Ймовірність влучення в мішень для першого стрілка дорівнює 0,2; для другого – 0,3. Яка ймовірність того, що перший зробить більше пострілів?

Вказівка. Скористатись формулою для суми геометричної прогресії.

Відповідь. 0,45.

12. Два пункти сполучаються кількома лініями зв’язку. Ймовірність пошкодження кожної з них протягом часу Т дорівнює 0,8. Заміна будь-якої пошкодженої лінії може бути проведена лише після пошкодження всіх ліній. Скільки потрібно провести ліній, щоб ймовірність функціонування зв’язку між пунктами протягом часу Т була більш ніж 0,99?

Відповідь. $n \geq 3$.

13. Урна містить дев’ять білих і одну чорну кулю. Знайти ймовірність того, що при 10 вийманнях з поверненнями буде витягнута принаймні одна чорна куля. Скільки разів треба виймати по одній кулі з поверненням, щоб імовірність появи хоча б однієї чорної кулі була не менша 0,9?

Відповідь. $1-(0{,}9)^{10} \approx 0{,}651$; 22.

1.10. Формула повної ймовірності. Формула Байєса

Одним з наслідків застосування теорем додавання і добутку є формула повної ймовірності та формула Байєса.

Нагадаємо, що події $H_1, H_2, ..., H_n$ утворюють повну групу, якщо:

$$H_i \cdot H_j = \varnothing,\ i \neq j \qquad \text{і} \qquad \sum_{i=1}^{n} H_i = \Omega.$$

Теорема 1.7 *(Формула повної ймовірності).* Якщо подія A може настати тільки при умові настання однієї з подій (гіпотези) $H_1, H_2, ..., H_n$, які утворюють повну групу, то ймовірність події A дорівнює сумі добутків ймовірностей кожної з подій $H_1, H_2, ..., H_n$ на відповідну умовну ймовірність події A, тобто

$$P(A) = \sum_{i=1}^{n} P(H_i) \cdot P_{H_i}(A). \tag{1.32}$$

Приклад. До станції примикають дві під'їзні колії. Щодоби з однієї з них відправляються вагони у двох подачах по 30 вагонів, а з другої − в чотирьох подачах по 20 вагонів. У кожній подачі з першої під'їзної колії п'ять, а з другої три вагони недовантажені. Знайти ймовірність того, що вибраний навмання вагон є недовантажений.

<u>Розв'язання.</u> $A = \{$Вагон є недовантаженим$\}$. Гіпотези $H_1 = \{$Вагон вибраний з подачі з першої під'їзної колії$\}$; $H_2 = \{$Вагон вибраний з подачі з другої під'їзної колії$\}$. Тоді $P(H_1) = \frac{2}{2+4} = \frac{1}{3}$; $P(H_2) = \frac{4}{2+4} = \frac{2}{3}$. Умовна ймовірність того, що не довантажили вагон на першій під'їзній колії $P_{H_1}(A) = \frac{5}{30} = \frac{1}{6}$, а на другій $P_{H_2}(A) = \frac{3}{20}$. За формулою (1.32) обчислюємо ймовірність $P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) = \frac{1}{3} \cdot \frac{1}{6} + \frac{2}{3} \cdot \frac{3}{20} = \frac{14}{90}$.

Наслідком формули (1.32) є теорема 1.8.

Теорема 1.8. *Формула Байєса (теорема гіпотез).* Нехай події $H_1, H_2, ..., H_n$ утворюють повну групу. Тоді умовна ймовірність події $H_k \ (k = 1,2,...,n)$ за умовою, що подія A відбулася, обчислюється за *формулою Байєса*:

$$P_A(H_k) = \frac{P(H_k) \cdot P_{H_k}(A)}{P(A)}, \tag{1.33}$$

де $P(A) = P(H_1) \cdot P_{H_1}(A) + P(H_2) \cdot P_{H_2}(A) + ... + P(H_n) \cdot P_{H_n}(A)$ − формула повної ймовірності.

Формула (1.33) дозволяє переоцінити ймовірності гіпотез $H_k \ (k = 1,2,...,n)$ за результатами експерименту, тобто знайти апостеріорні ймовірності $P_A(H_k)$ та порівняти їх з апріорними (прийнятими до експерименту).

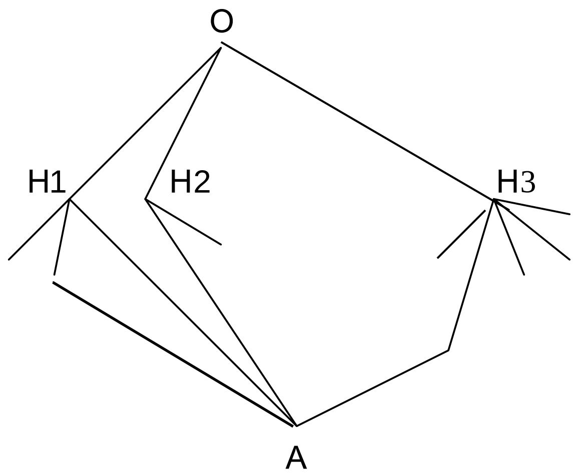
Приклад. У подачі є декілька вагонів, серед яких можуть знаходитися чотиривісна платформа з ймовірністю $p_1 = 0{,}3$, чотиривісний піввагон з ймовірністю $p_2 = 0{,}5$ і шестивісний піввагон з ймовірністю $p_3 = 0{,}2$. При огляді їх встановили, що ймовірність пошкодження чотиривісного вагона дорівнює 0,01, а шестивісного – 0,014. Знайти ймовірність того, що пошкодженим є шестивісний піввагон.

Розв’язання. Висунуті гіпотези: $H_1 = \{$Вибраний навмання вагон є чотиривісним$\}$; $H_2 = \{$Вибраний навмання вагон є шестивісним$\}$. Подія $A = \{$Вагон пошкоджений$\}$. Тоді $P(H_1) = 0{,}3 + 0{,}5 = 0{,}8;\ P(H_2) = 0{,}2.$ За формулою Байєса

$$P_A(H_2) = \frac{P(H_2)\cdot P_{H_2}(A)}{P(H_1)\cdot P_{H_1}(A) + P(H_2)\cdot P_{H_2}(A)} = \frac{0{,}2\cdot 0{,}014}{0{,}8\cdot 0{,}01 + 0{,}2\cdot 0{,}014} = \frac{7}{27}.$$

Завдання

1. На рисунку зображено схему дирекції залізничних перевезень. З пункту O в пункт A відправили туристсько-екскурсійний потяг, обираючи шлях навмання. Знайти ймовірність того, що потяг потрапить в пункт A.



Відповідь. $\frac{41}{120}$.

2. На станції прибувають піввагони, платформи і криті вагони з ймовірностями відповідно 0,35; 0,4; 0,25. При огляді їх встановлено, що ймовірність пошкодження піввагона дорівнює 0,015, платформи – 0,01, критого вагона – 0,02. Знайти ймовірність того, що пошкодженим є піввагон.

Відповідь. 0,55.

3. На головний матеріальний склад залізниці надходять автонавантажувачі від трьох постачальників у відношенні 1:4:5. 98%, 88% та 92% автонавантажувачів, що надходять від 1, 2 та 3-го постачальників відповідно, не потребують ремонту протягом гарантійного строку. Визначити: а) ймовірність того, що прийнятий автонавантажувач на склад не потребує ремонту протягом гарантійного строку; б) від якого постачальника правдоподібніше надійшов автонавантажувач, якщо він підлягає ремонту протягом гарантійного строку?

Відповідь: а) 0,91; б) від другого постачальника.

4. З 30 екзаменаційних білетів студент вивчив тільки 20. Яким йому вигідно зайти до іспиту: першим або другим?

Відповідь. Однаково.

5. Двоє стрільців стріляють у мішень незалежно один від одного, кожен робить по одному пострілу. Ймовірність влучення в мішень при одному пострілі для першого стрільця – 0,7; для другого – 0,8. Після пострілів у мішені виявлено одну пробоїну. Яка ймовірність того, що вона належить: а)першому стрільцю; б) другому стрільцю?

Відповідь: а) 0,37; б) 0,63.

6. Пасажир для придбання квитка може звернутися до однієї з чотирьох кас на одній станції. Відповідні ймовірності дорівнюють 0,2; 0,3; 0,4;. 0,5. Імовірність того, що до моменту появи пасажира в касі буде квиток, дорівнює відповідно 0,6; 0,7; 0,8; 0,9. Пасажир звернувся до однієї з кас і купив квиток. Яка ймовірність того, що квиток пасажир придбав у першій касі?

Відповідь. 0,1091.

7. Маємо три групи ящиків. До першої групи належить п’ять ящиків, у кожному з яких шість стандартних і три браковані однотипні вироби, до другої групи – вісім ящиків, у кожному з яких п’ять стандартних і п’ять бракованих виробів, а до третьої – два ящики, у кожному з яких три стандартні і сім бракованих виробів. Із довільно вибраного ящика три навмання взяті вироби виявились стандартними. Яка ймовірність того, що вони були взяті з ящика, який належить третій групі?

Відповідь. $\frac{7}{787}$.

8. Деталь може надійти для обробки на перший верстат із імовірністю 0,2, на другий верстат – із імовірністю 0,3 і на третій – із імовірністю 0,5. При обробці деталі на першому верстаті ймовірність допустити брак – 0,01, на другому і третьому верстатах ця ймовірність відповідно дорівнює 0,05 і 0,08. Оброблені деталі вміщують в одну шухляду. Навмання взята деталь виявилась бракованою. Яка ймовірність того, що її обробляв перший верстат?

Відповідь. $\frac{20}{57}$.

9. Чотири робітники виготовляють однотипні вироби. При цьому продуктивність праці цих робітників задовольняє таке відношення 2:1,5:4:2,5. Відомо, що частка браку для першого, другого, третього та четвертого робітників дорівнює відповідно 1,5; 2,8; 2 та 4,5%. Після робочої зміни вироби вміщують в один бункер. Навмання взятий виріб із бункера виявився стандартним. Яка ймовірність, що його виготовив перший або третій робітник?

Відповідь. 0,6.

10. В ящику міститься 20 тенісних м’ячів, із них 12 нових і 8, які були в користуванні. Із ящика беруть навмання два м’ячі і після закінчення гри повертають у ящик. Після цього із ящика навмання вибирають знову два м’ячі для наступної гри. Обчислити ймовірності випадкових подій: а) два м’ячі, що останніми вийняли із ящика, ще не були в користуванні; б) два м’ячі, що останніми вийняли із ящика, вже були в користуванні.

Відповідь:

а) $\frac{C_{12}^2 C_{10}^2 + C_{12}^1 C_8^1 C_{11}^2 + C_8^2 C_{12}^2}{C_{20}^2 C_{20}^2}$; б) $\frac{C_{12}^2 C_8^2 + C_{12}^1 C_8^1 C_7^2 + C_8^2 C_6^2}{C_{20}^2 C_{20}^2}$.

11. В групі 25 студентів, які складають іспит з математики, із них 5 підготовлено відмінно, 10 – добре, 9 – задовільно і 6 – незадовільно. В екзаменаційних тестах міститься 10 запитань. Відмінно підготовлений студент може відповісти на 10 запитань, добре підготовлений – на 7 запитань, задовільно підготовлений –

на 5 запитань і незадовільно підготовлений – на 3 запитання. Навмання викликаний студент відповів на три запропоновані йому запитання. Знайти ймовірність того, що це був студент: а) відмінно підготовлений; б) незадовільно підготовлений.

Відповідь: а) 0,57; б) 0,06.

12. Працівники залізниці, що відповідають за безпеку руху потягів, розділені на три категорії: 20, 50 і 30% відповідно. Ймовірність того, що один із залізничників здійснить порушення протягом року, відповідно дорівнює 0,01; 0,015; 0,02. Навмання перевірений працівник два роки підряд з п'яти років роботи порушив правила безпеки руху. Яка ймовірність того, що він належить до: а) першої категорії; б) третьої категорії?

Відповідь: а) $\approx 0{,}08$; б) $\approx 0{,}47$.

13. Пасажир для придбання квитка може звернутися в одну із трьох кас. Ймовірність того, що він звернеться до тієї чи іншої каси, залежить від їх місцезнаходження і відповідно дорівнює p_1, p_2, p_3. Ймовірність того, що до моменту приходу пасажира квитки, які є у касі, будуть розпродані, дорівнює для першої каси – P_1, для другої – P_2, для третьої – P_3. Пасажир підійшов до однієї з кас і придбав квиток. Знайти ймовірність того, що це була третя каса.

Відповідь. $\dfrac{p_3 \cdot (1-P_3)}{p_1 \cdot (1-P_1) + p_2 \cdot (1-P_2) + p_3 \cdot (1-P_3)}$.

14. Стрілець A влучає в мішень з імовірністю $p_1 = 0{,}6$, стрілець B – з імовірністю $p_2 = 0{,}5$, а стрілець C – з імовірністю $p_3 = 0{,}4$. Стрільці зробили залп по мішені, після чого було виявлено дві пробоїни. Що більш імовірно: стрілець C влучив чи ні?

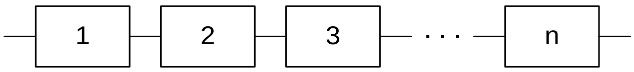
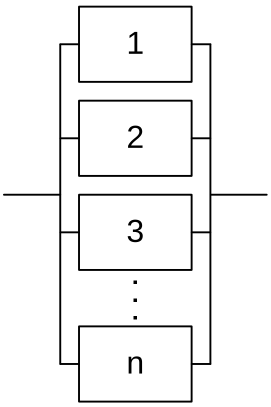
Відповідь. Ймовірність влучення стрільця C дорівнює $\frac{10}{19} > \frac{1}{2}$.

1.11. Використання формул теорії ймовірностей для оцінювання надійності роботи простих систем

Розглянемо системи, що складаються з деякої кількості елементів з відомими показниками надійності – ймовірностями безвідмовної роботи. За допомогою формул теорії ймовірностей можна визначити надійність системи в цілому. В залежності від того, яким чином елементи об'єднані в систему, використовуються різні формули.

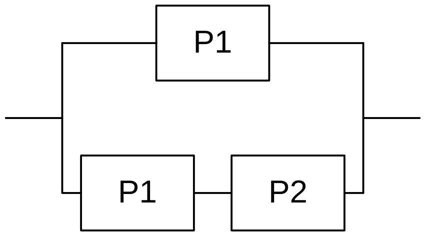
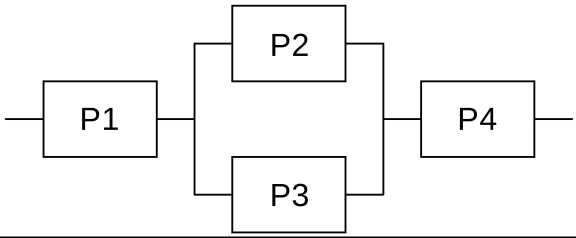
У табл. 1.3 наведено системи з послідовним та паралельним способом зв'язку елементів, коли відомі ймовірності роботи кожного елемента $p_i, i = 1, 2, ..., n.$

Таблиця 1.3

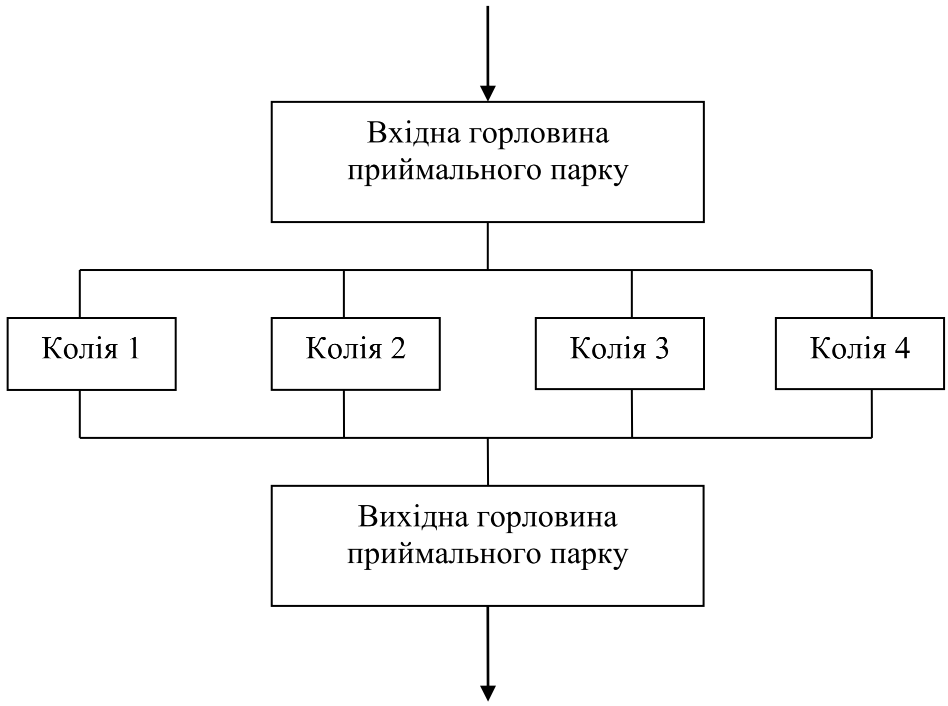
<table>
<tr><th>Схема системи</th><th>Надійність роботи системи, R</th></tr>
<tr><td>1.

</td><td>$R = p_1 \cdot p_2 \cdot ... \cdot p_n.$</td></tr>
<tr><td>2.

</td><td>$R = 1 - q_1 \cdot q_2 \cdot ... \cdot q_n;$
$q_i = 1 - p_i, \ i = 1, 2, ..., n.$</td></tr>
</table>

За допомогою цих формул обчислюються надійності систем з більш складною структурою.

Приклади

<table>
<tr><th>Схема системи</th><th>Надійність роботи системи, R</th></tr>
<tr><td>1.
</td><td>$R = 1 - q_1(1 - p_1 p_2)$,
$q_1 = 1 - p_1$.</td></tr>
<tr><td>2.
</td><td>$R = p_1 p_4 (1 - q_2 q_3)$;
$q_2 = 1 - p_2, q_3 = 1 - p_3$.</td></tr>
</table>

Приклад. Структурна схема приймального парку дільничної станції має такий вигляд:



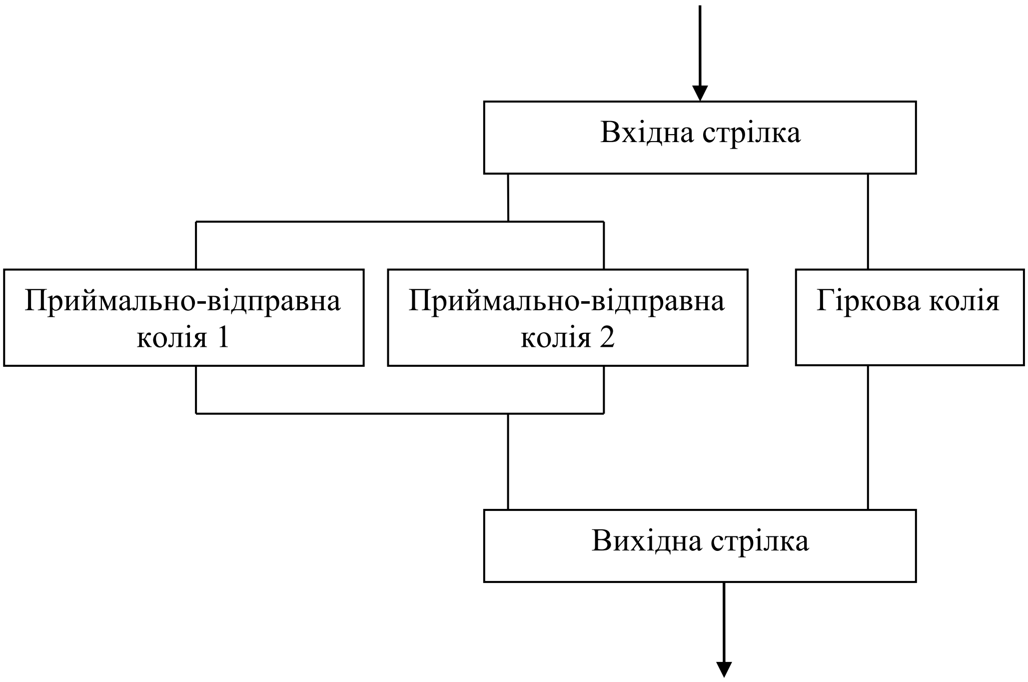
Імовірності безвідмовної роботи елементів системи дорівнюють: $p_1 = 0{,}89$ для вхідної горловини, $p_2 = 0{,}9$; $p_3 = 0{,}95$; $p_4 = 0{,}8$; $p_5 = 0{,}92$ для приймальних колій 1, 2, 3 і 4 відповідно, $p_6 = 0{,}87$ для вихідної горловини. Обчислити надійність системи.

Розв’язання. Надійність системи обчислюється за формулою:

$$R = p_1 \cdot \left(1-(1-p_2)\cdot(1-p_3)\cdot(1-p_4)\cdot(1-p_5)\right)\cdot p_6 =$$
$$= 0{,}89 \cdot \left(1 - 0{,}1\cdot 0{,}05 \cdot 0{,}2 \cdot 0{,}08\right)\cdot 0{,}87 \approx 0{,}774.$$

Завдання

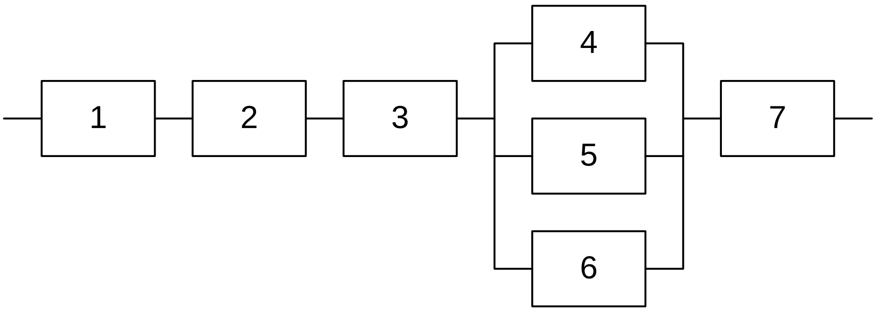
1. Розрахувати надійність системи, схему якої та імовірності безвідмовної роботи її елементів наведено нижче.



Назва елемента	Вхідна стрілка	ПВК 1	ПВК 2	ГК	Вихідна стрілка
Імовірність безвідмовної роботи	0,7	0,92	0,91	0,98	0,8

Відповідь. 0,56.

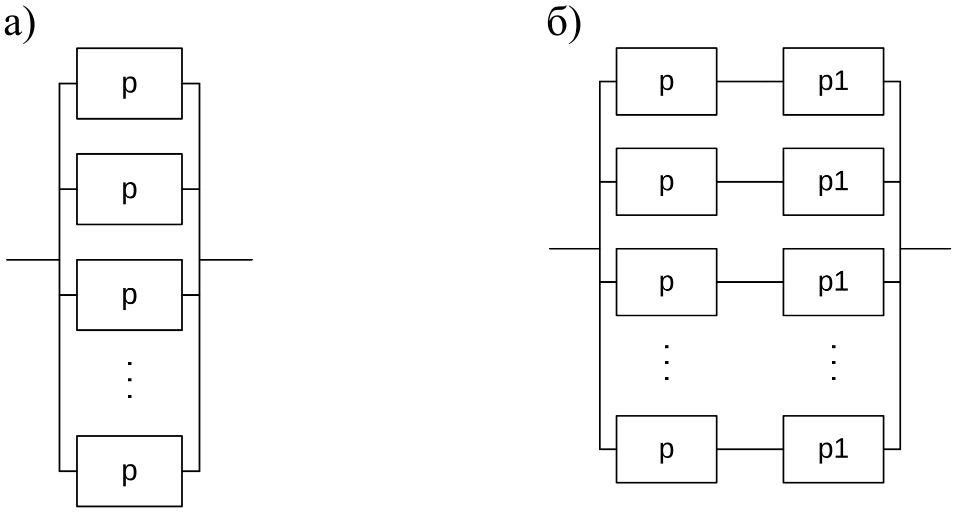
2. Електролампочки з’єднані за схемою:



Імовірність того, що електролампочка не вийде з ладу при ввімкненні системи в електричну мережу, є величиною сталою і дорівнює 0,8. Яка ймовірність того, що в електричній схемі при ввімкненні її в електричну мережу потече електричний струм?

Відповідь. 0,4063.

3. Для підвищення надійності вагонного уповільнювача він дублюється $(n-1)$ іншими такими ж уповільнювачами.



Надійність кожного уповільнювача дорівнює p. Скільки треба взяти уповільнювачів, щоб підвищити надійність системи до заданої R_1?

Відповідь: а) $n \geq \dfrac{\ln(1-R_1)}{\ln(1-p)}$; б) $n \geq \dfrac{\ln(1-R_1)}{\ln(1-pp_1)}$.

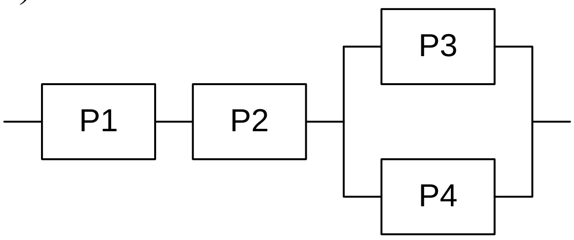
4. Залізничний потяг має n вагонів, кожний з яких з імовірністю p має дефект. Всі вагони оглядають незалежно один від одного два оглядачі вагонів. Перший з них установлює дефект

з імовірністю p_1, другий – з імовірністю p_2. Якщо в жодному з вагонів не визначено дефекту, потяг відправляють. Знайти ймовірність того, що відправлено потяг, у якому є хоча б один дефектний вагон.

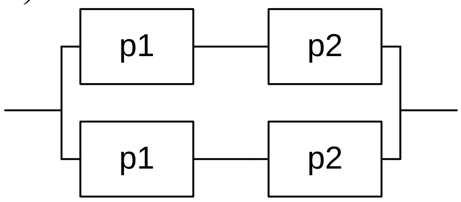
<u>Відповідь.</u> $1-\left[1-p(1-p_1)(1-p_2)\right]^n$.

5. Визначити надійність схем, якщо надійність кожного елемента дорівнює p_i і відмови елементів незалежні:

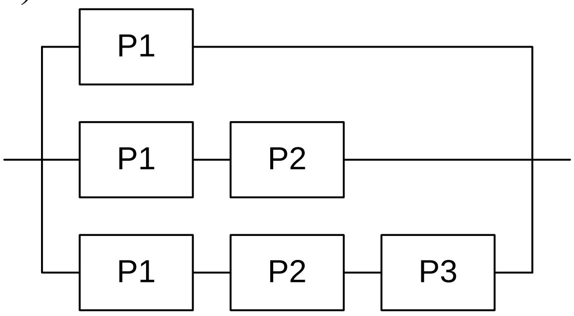
а)



б)



в)



Питання до теми

1. Що є предметом теорії ймовірностей? Які задачі теорії ймовірностей?
2. Що називається подією? Яка подія називається випадковою?
3. Яка подія називається достовірною? Неможливою? Наведіть приклади.
4. Які події називаються елементарними? Що називається простором елементарних подій? Наведіть приклади.
5. Що називається сумою двох подій? Наведіть приклади.
6. Що називається добутком двох подій? Наведіть приклади.

7. Що називається різницею двох подій? Наведіть приклади.

8. Які події називаються протилежними? Наведіть приклади.

9. Які події називаються сумісними? Несумісними? Наведіть приклади.

10. Які події називаються незалежними? Залежними? Наведіть приклади.

11. Що називається повною групою подій?

12. Що вивчає комбінаторика? Які її основні принципи?

13. Що таке перестановки? Чому дорівнює кількість перестановок з повтореннями і без? Наведіть приклади застосування перестановок.

14. Що таке комбінації? Чому дорівнює кількість комбінацій з повтореннями і без? Які властивості комбінацій ви знаєте? Наведіть приклади застосування комбінацій.

15. Що таке розміщення? Чому дорівнює кількість розміщенны з повтореннями і без? Наведіть приклади застосування розміщень.

16. Наведіть класичне означення ймовірності. Які її властивості?

17. Наведіть статистичне означення ймовірності. Які його переваги та недоліки?

18. Наведіть геометричне означення ймовірності. В яких випадках воно застосовується?

19. Чому дорівнює ймовірність суми двох (кількох) несумісних подій?

20. Чому дорівнює ймовірність суми подій, що утворюють повну групу?

21. Чому дорівнює ймовірність протилежної події? Наведіть приклади.

22. Чому дорівнює ймовірність добутку двох (кількох) подій?

23. Чому дорівнює ймовірність добутку двох (кількох) незалежних подій?

24. Чому дорівнює ймовірність суми двох сумісних подій?

25. Як знайти ймовірність появи випадкової події хоча б один раз в кількох незалежних експериментах?

26. Які події можуть служити гіпотезами?

27. Наведіть формулу повної ймовірності. Коли вона застосовується?

28. Яка формула дозволяє переоцінити ймовірності гіпотез після проведення експерименту?

Тестові питання

1. Чому дорівнює ймовірність достовірної події?

А	Б	В	Г	Д
0	1	-1	1/2	Інша відповідь

2. Чому дорівнює ймовірність неможливої події?

А	Б	В	Г	Д
0	1	-1	1/2	Інша відповідь

3. Чому дорівнює ймовірність випадіння герба при одноразовому киданні монети?

А	Б	В	Г	Д
0	1	-1	1/2	Інша відповідь

4. Чому дорівнює ймовірність випадіння шести очок при однократному киданні грального кубика?

А	Б	В	Г	Д
0	1	-1	1/2	Інша відповідь

5. Ймовірність безвідмовної роботи приладу протягом часу T дорівнює 0,7. Яка ймовірність його відмови протягом часу T?

А	Б	В	Г	Д
1	0,7	0,5	0,3	Інша відповідь

6. На картках написано букви А, І, К, Р. Яка ймовірність викласти з них слово ІКРА?

А	Б	В	Г	Д
$\frac{1}{4}$	$\frac{1}{4^4}$	$\frac{1}{4!}$	4	Інша відповідь

7. На картках написано букви А, І, К, Р. Яка ймовірність викласти з них слово РІК?

А	Б	В	Г	Д
$\frac{1}{A_4^3}$	$\frac{1}{P_4}$	$\frac{1}{P_3}$	$\frac{1}{A_4^3}$	Інша відповідь

8. Три студенти можуть скласти іспит з ймовірностями 0,8; 0,5 і 0,6 відповідно. Яка ймовірність того, що іспит складуть всі троє?

А	Б	В	Г	Д
0	0,24	0,024	2,4	Інша відповідь

9. Три студенти можуть скласти іспит з ймовірностями 0,8; 0,5 і 0,6 відповідно. Яка ймовірність того, що іспит складе хоча б один?

А	Б	В	Г	Д
$0{,}8+0{,}5+0{,}6$	$0{,}8\cdot0{,}5\cdot0{,}6$	$1-0{,}8\cdot0{,}5\cdot0{,}6$	$1-0{,}2\cdot0{,}5\cdot0{,}4$	Інша відповідь

10. Ймовірність наявності квитків до пункту призначення в першій касі дорівнює 0,7, а в другій - 0,9. Яка ймовірність наявності квитків хоча б в одній касі?

А	Б	В	Г	Д
$0{,}7+0{,}9$	$0{,}7\cdot0{,}9$	$0{,}7+0{,}9-0{,}7\cdot0{,}9$	$1-0{,}7\cdot0{,}9$	Інша відповідь

11. У двох урнах міститься по 4 білих та 6 чорних кульок. З першої урни навмання дістали одну кульку та поклали до другої.

Потім з другої урни навмання дістали одну кульку. Яка ймовірність того, що обидві кульки були чорними?

А	Б	В	Г	Д
$\frac{6}{10}\cdot\frac{6}{10}$	$\frac{6}{10}\cdot\frac{7}{11}$	$\frac{6}{10}\cdot\frac{6}{11}$	$\frac{6}{10}\cdot\frac{7}{10}$	Інша відповідь

12. У двох урнах міститься по 4 білих та 6 чорних кульок. З першої урни навмання дістали одну кульку та поклали до другої. Яка ймовірність дістати з другої урни чорну кульку?

А	Б	В	Г	Д
$\frac{6}{10}\cdot\frac{7}{11}$	$\frac{4}{10}\cdot\frac{7}{11}+\frac{6}{10}\cdot\frac{6}{11}$	$\frac{6}{10}\cdot\frac{7}{11}+\frac{4}{10}\cdot\frac{6}{11}$	$\frac{4}{10}\cdot\frac{6}{11}$	Інша відповідь

<u>Відповіді:</u> 1.Б; 2.А; 3.Г; 4.Д; 5.Г; 6.В; 7.А; 8.Б; 9.Г; 10.В; 11.Б; 12.В.

Розділ 2. Повторні незалежні випробування

2.1. Формула Бернуллі. Найімовірніше число появи випадкової події

Якщо кожне випробування має лише два несумісні наслідки (події) зі сталими ймовірностями p і q (випадкова подія A відбувається з імовірністю p, випадкова подія A не відбувається з імовірністю q, $p+q=1$), то їх називають випробуваннями *за схемою Бернуллі*. Простір елементарних подій для одного випробування містить дві елементарні події, а для n випробувань за схемою Бернуллі − 2^n елементарних подій.

Імовірність того, що в результаті n незалежних випробувань за схемою Бернуллі подія A з'явиться k раз, обчислюється за *формулою Бернуллі*

$$\underline{P_n(k) = C_n^k p^k q^{n-k},} \tag{2.1}$$

де $C_n^k = \dfrac{n!}{k!(n-k)!}$, p − ймовірність появи події A в кожному випробуванні, $q = 1-p$ − ймовірність появи події $\overline{A}$ в кожному випробуванні.

Приклад. Імовірність знаходження вагонів на дане призначення в кожному потязі, що прибув на станцію, дорівнює $p = 0{,}2.$ Визначити ймовірність того, що в трьох із п'яти потягів, будуть вагони на дане призначення.

<u>Розв'язання.</u> За формулою Бернуллі (2.1) обчислюємо:

$$P_5(3) = C_5^3 \cdot 0{,}2^3 \cdot (1-0{,}2)^{5-3} = \frac{5!}{3!\cdot 2!} \cdot 0{,}2^3 \cdot 0{,}8^2 = 0{,}051.$$

Ймовірність того, що в результаті n незалежних випробувань подія A з'явиться від k_1 до k_2 раз, обчислюється за формулою

$$\underline{P_n(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} C_n^k p^k q^{n-k}.} \tag{2.2}$$

Оскільки

$$P_n(0 \le k \le n) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1, \qquad (2.3)$$

отримаємо:

$$P_n(0 \le k \le k_i) = 1 - \sum_{k=k_i+1}^{n} C_n^k p^k q^{n-k}, \qquad (2.4)$$

$$P_n(k_i \le k \le n) = 1 - \sum_{k=0}^{k_i-1} C_n^k p^k q^{n-k}. \qquad (2.5)$$

Приклад. У парку приймання шість колій, ймовірність зайнятості кожної з них потягами, що прибувають, дорівнює $p = 0{,}7$. Знайти ймовірність того, що в даний момент часу в парку зайняті потягами не більше двох колій.

Розв'язання. Шукана ймовірність – сума ймовірностей таких подій: жодна колія не зайнята, тільки одна колія зайнята, тільки дві колії зайняті. Ймовірність кожної з них обчислимо за формулою Бернуллі (2.1):

$$P_6(0) = C_6^0 \cdot (0{,}7)^0 \cdot (1 - 0{,}7)^{6-0} = 0{,}000729;$$

$$P_6(1) = C_6^1 \cdot (0{,}7)^1 \cdot (1 - 0{,}7)^{6-1} = 0{,}010206;$$

$$P_6(2) = C_6^2 \cdot (0{,}7)^2 \cdot (1 - 0{,}7)^{6-2} = 0{,}059435.$$

Таким чином, за формулою (2.2) отримаємо:

$$P_6(0 \le k \le 2) = P_6(0) + P_6(1) + P_6(2) = 0{,}00729 + 0{,}010206 + 0{,}059435 = 0{,}07037.$$

Найімовірнішим числом появи випадкової події A в результаті n незалежних випробувань за схемою Бернуллі називається таке число k_0, для якого ймовірність $P_n(k_0)$ є не меншою за ймовірність $P_n(k)$ будь-якого іншого наслідку серії випробувань, тобто: $P_n(k_0) \ge P_n(k), \forall k.$

Для знаходження k_0 розв'яжемо систему рівнянь:

$$\begin{cases} P_n(k_0) \geq P_n(k_0+1), \\ P_n(k_0) \geq P_n(k_0-1). \end{cases}$$

За формулою Бернуллі (2.1) отримаємо:

$$\begin{cases} C_n^{k_0} \cdot p^{k_0} \cdot q^{n-k_0} \geq C_n^{k_0+1} \cdot p^{k_0+1} \cdot q^{n-k_0-1}, \\ C_n^{k_0} \cdot p^{k_0} \cdot q^{n-k_0} \geq C_n^{k_0-1} \cdot p^{k_0-1} \cdot q^{n-k_0+1}, \end{cases}$$

тобто

$$\begin{cases} \dfrac{n!}{k_0!(n-k_0)!} \cdot p^{k_0} \cdot q^{n-k_0} \geq \dfrac{n!}{(k_0+1)!(n-k_0-1)!} \cdot p^{k_0+1} \cdot q^{n-k_0-1}, \\ \dfrac{n!}{k_0!(n-k_0)!} \cdot p^{k_0} \cdot q^{n-k_0} \geq \dfrac{n!}{(k_0-1)!(n-k_0+1)!} \cdot p^{k_0-1} \cdot q^{n-k_0+1}. \end{cases}$$

Після спрощення маємо:

$$\begin{cases} \dfrac{q}{n-k_0} \geq \dfrac{p}{k_0+1}, \\ \dfrac{p}{k_0} \geq \dfrac{q}{n-k_0+1}; \end{cases} \Rightarrow \begin{cases} k_0(p+q) \geq pn - q, \\ pn + p \geq k_0(p+q); \end{cases}$$

Звідси отримаємо формулу для знаходження найімовірнішого числа k_o:

$$\underline{np - q \leq k_0 \leq np + p.} \qquad (2.6)$$

Зауваження. Якщо $np + p$ – ціле число, то найімовірніших числа два $k_0 = np + p$ і $k_0' = np - q$, в противному випадку – одне.

Приклад. Імовірність того, що студент складе іспит з математики, 0,7. Нехай є група з 8 студентів. Знайти найімовірнішу кількість членів цієї групи, які складуть іспит з математики. Обчислити відповідну ймовірність.

Розв'язання. За умовою $n = 8; p = 0{,}7; q = 0{,}3$. За формулою (2.6):

$$8\cdot 0{,}7-0{,}3\le k_0\le 8\cdot 0{,}7+0{,}7$$
$$5{,}3\le k_0\le 6{,}3$$
$$k_0=6.$$

Отже, $P_8(6)=C_8^6(0{,}7)^6(0{,}3)^2=0{,}2965$.

Приклад. Скільки разів треба підкинути гральний кубик, щоб найімовірніша кількість випадань трійки дорівнювала 10?

Розв’язання. За умовою $p=\dfrac{1}{6}\Rightarrow q=\dfrac{5}{6}$. За формулою (2.6):

$$n\cdot\frac{1}{6}-\frac{5}{6}\le 10\le n\cdot\frac{1}{6}+\frac{1}{6};$$
$$n-5\le 60\le n+1;$$
$$59\le n\le 65.$$

Відповідь. Від 59 до 65 разів включно.

Завдання

1. Висадка з вагона приміського потяга в середньому по станції А складає приблизно 60% від загальної кількості пасажирів, що прибували до даної станції. Яка ймовірність того, що з п’яти пасажирів на станції А вийдуть: а) чотири пасажири; б) не менш як чотири пасажири?

Відповідь: а) 0,259; б) 0,337.

2. У подачі порожніх вагонів кожний з них потребує очищення з ймовірністю $p=0{,}1$. Яка ймовірність того, що в подачі з шести вагонів буде підлягати очищенню не більше двох вагонів?

Відповідь. 0,98.

3. Під час тестування з математики студент повинен дати правильні відповіді на п’ять запитань. Імовірність того, що він відповість на одне запитання, у середньому дорівнює 0,8. Щоб скласти тест, студентові необхідно відповісти не менш ніж на три питання. Знайти ймовірність того, що студент складе тест.

Відповідь. 0,942.

4. Імовірність виготовлення робітником деталі відмінної якості становить 0,75. Яка ймовірність того, що серед 6 виготовлених деталей робітником хоча б одна буде відмінної якості? Знайти найімовірніше число виготовлених робітником деталей відмінної якості й обчислити ймовірність цього числа.

Відповідь. $P_6(k \geq 1) = 0{,}9998;$ $k_0 = 5;$ $P_6(5) = 0{,}356.$

5. У локомотивному депо є 12 локомотивів для пасажирського руху. Імовірність того, що під состав вагонів вийде локомотив, у середньому дорівнює 0,85. Знайти ймовірність того, що депо працюватиме в нормальному режимі, якщо для цього потрібно, не менш ніж 9 локомотивів.

Відповідь. 0,687.

6. У партії однотипних деталей кількість стандартних і бракованих відносяться, як 5:2. Навмання з партії беруть вісім деталей. Яка ймовірність того, що серед них стандартних виявиться шість? Знайти найімовірніше число появи стандартних деталей серед семи навмання взятих і обчислити відповідну ймовірність.

Відповідь. $\frac{250000}{823543};$ $k_0 = 6;$ $P_8(6) = \frac{250000}{823543}.$

7. У кожному із семи ящиків міститься по шість стандартних і чотири бракованих однотипних деталей. Навмання з кожного ящика беруть по одній деталі. Обчислити ймовірність того, що серед семи взятих деталей стандартних буде: а) три; б) не менш як три; в) не більш як три.

Відповідь: а) 0,194; б) 0,981; в) 0,096.

2.2. Локальна та інтегральна формули Муавра-Лапласа

Припустимо, що нам потрібно обчислити ймовірність $P_n(k)$ появи події A при $n = 500, k = 200$. За формулою Бернуллі отримаємо:

$$P_{500}(200) = C_{500}^{200} \cdot (p)^{200} \cdot (q)^{300} = \frac{500!}{200! \cdot 300!}(p)^{200}(q)^{300}.$$

Очевидно, що при великих n обчислення $P_n(k)$ за формулою Бернуллі технічно складно. Тому в цих випадках користуються наближеними формулами для обчислення $P_n(k)$ і формули називаються *асимптотичними* і визначаються локальною та інтегральною теоремами Муавра-Лапласа, теоремою Пуассона.

Теорема 2.1 *(Локальна теорема Муавра-Лапласа.)* Якщо ймовірність появи випадкової події A в кожному з n незалежних експериментів є величиною сталою і дорівнює $p\,(0 < p < 1)$, то для великих значень n і k імовірність того, що випадкова подія A настане k раз, обчислюється за такою асимптотичною формулою:

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \tag{2.7}$$

де $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ – *функція Гаусса,* що обчислена при $x = \frac{k - np}{\sqrt{npq}}$.

Функція Гаусса є табульованою, її значення наведено в дод. 1.

Зауваження. Чим більше n, тим точніше наближена формула (2.7). Наближені значення ймовірності $P_n(k)$ (2.7) на практиці використовуються як точні, коли $npq \geq 9$.

Властивості функції Гаусса:

1) $\varphi(x)$ визначена на всій осі абсцис і $\varphi(x) > 0$;

2) $\varphi(x)$ є парною функцією: $\varphi(-x) = \varphi(x)$;

3) $\lim\limits_{x \to \pm\infty} \varphi(x) = 0$;

4) $\varphi_{\max} = \varphi(0) = \frac{1}{\sqrt{2\pi}}$ – максимум функції Гаусса;

5) $\left(\pm 1; \frac{1}{\sqrt{2\pi e}}\right)$ – точки перегину графіка функції Гаусса.

Графік функції зображено на рис. 2.1.

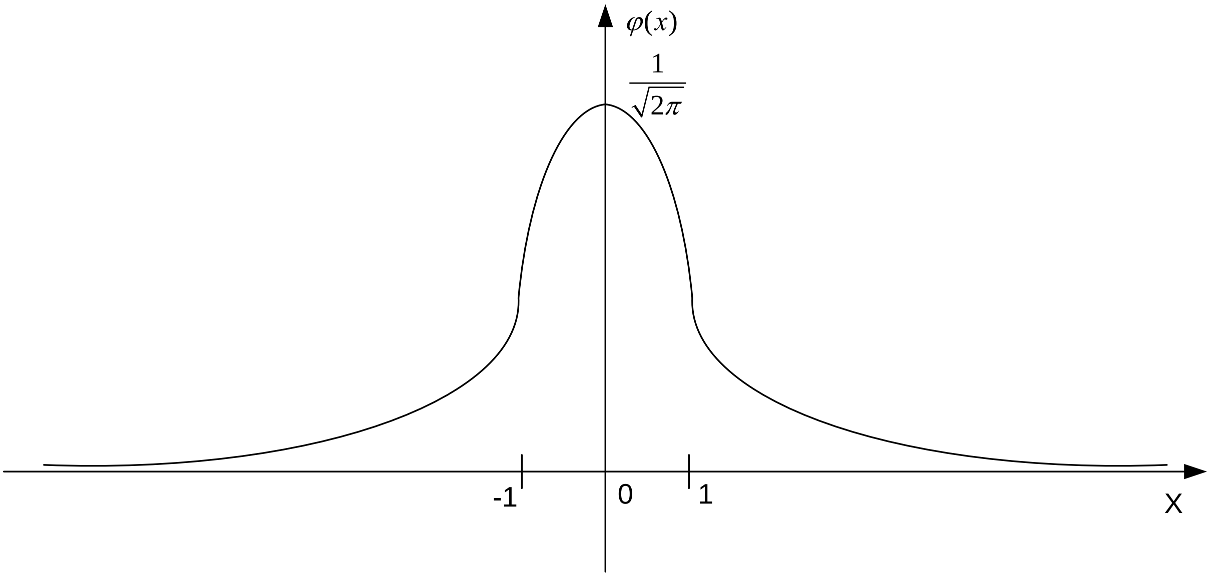


Рис. 2.1. Графік функції Гаусса

Зауважимо, що при розв'язанні задач дотримуються таких правил:

1) оскільки функція $\varphi(x)$ є парною, для від'ємних значень x користуються дод. 1, враховуючи, що $\varphi(-x)=\varphi(x)$;

2) для значень $|x|\geq 4$ вважають $\varphi(x)=0$.

Приклад. У кожній з подач на під'їзну колію ймовірність появи шестивісного піввагона $p=0{,}2$. Знайти ймовірність появи восьми шестивісних піввагонів у 100 подачах.

<u>Розв'язання.</u> За умовою задачі маємо: $n=100$; $k=8$; $p=0{,}2$; $q=1-0{,}2=0{,}8$. Оскільки $npq=16\geq 9$, то слід скористатися локальною формулою Муавра-Лапласа:

$$P_{100}(8)=\frac{1}{\sqrt{100\cdot 0{,}2\cdot 0{,}8}}\cdot\varphi(x).$$

Обчислимо значення $x=\frac{k-np}{\sqrt{npq}}=\frac{8-100\cdot 0{,}2}{4}=-3.$ Оскільки функція $\varphi(x)$ є парною, то $\varphi(-3)=\varphi(3)$ і за дод. 1 знаходимо: $\varphi(-3)=\varphi(3)=0{,}0044$. Тому $P_{100}(8)=\frac{1}{4}\cdot 0{,}0044=0{,}0011$.

Теорема 2.2 (*Інтегральна теорема Муавра-Лапласа*). Якщо ймовірність появи випадкової події A в кожному з n незалежних експериментів є величиною сталою і дорівнює $p\,(0<p<1)$, то для великих значень n імовірність появи випадкової події від k_1 до k_2 раз (включно) обчислюється за такою асимптотичною формулою:

$$P_n(k_1 \le k \le k_2) \approx \Phi(x_2) - \Phi(x_1), \tag{2.8}$$

де

$$x_2 = \frac{k_2 - np}{\sqrt{npq}}; \quad x_1 = \frac{k_1 - np}{\sqrt{npq}}, \tag{2.9}$$

а

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

є *функцією Лапласа*, значення якої наведено в дод. 2.

Формула (2.8) називається інтегральною формулою Муавра-Лапласа. Чим більш n, тим точніша ця формула. За умовою $npq \ge 9$ наближена формула (2.8) використовується як точна.

Властивості функції Лапласа:

1) $\Phi(x)$ визначена на всій осі абсцис;
2) $\Phi(x)$ є непарною функцією, тобто $\Phi(-x) = -\Phi(x)$;
3) $\Phi(x)$ є функцією неспадною;
4) $\Phi(0) = 0$; $\Phi(+\infty) = 0{,}5$; $\Phi(-\infty) = -0{,}5$;
5) $(0;0)$ – точка перегину графіку функції Лапласа.

Графік функції $\Phi(x)$ зображено на рис. 2.2.

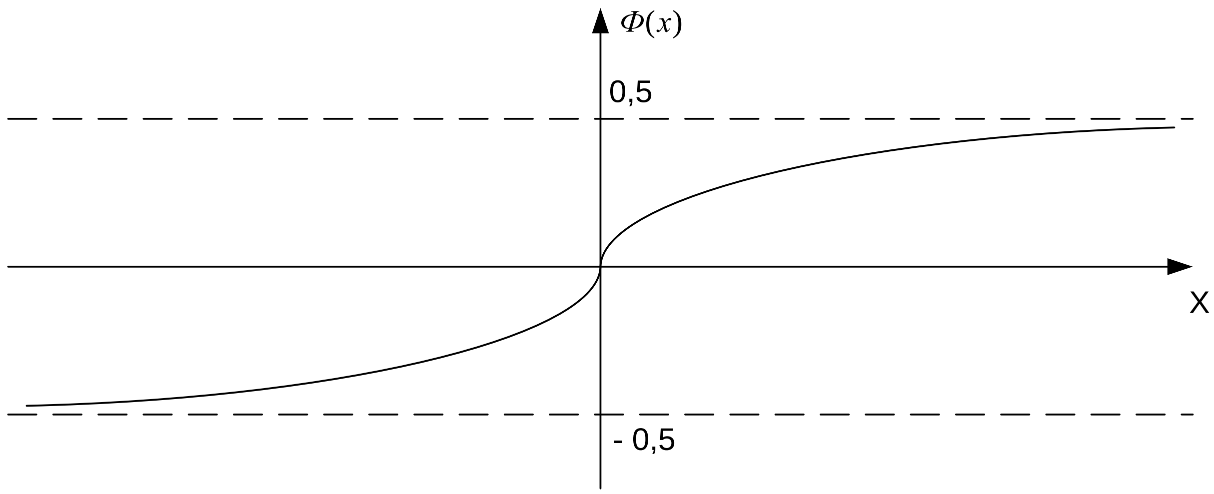


Рис.2.2. Графік функції Лапласа

Зауважимо, що при розв'язанні задач дотримуються таких правил:

1) оскільки функція Лапласа є непарною, для від'ємних значень x користуються дод. 2, враховуючи, що $\Phi(-x) = -\Phi(x)$;

2) для значень $|x| > 5$ вважають $\Phi(x) = 0{,}5$.

Приклад. Кубик підкинули 800 разів. Яка ймовірність того, що число очок, кратне трьом, випаде не менш 260 і не більше 274 разів?

<u>Розв'язання.</u> За умовою $n = 800;\ k_1 = 260; k_2 = 274;\ p = \frac{1}{3}; q = \frac{2}{3}$. За формулою (2.9) обчислюємо:

$$x_1 = \frac{260 - 800 \cdot \frac{1}{3}}{\sqrt{800 \cdot \frac{1}{3} \cdot \frac{2}{3}}} = -\frac{6{,}6667}{13{,}3333} = -0{,}5; \quad x_2 = \frac{274 - 800 \cdot \frac{1}{3}}{\sqrt{800 \cdot \frac{1}{3} \cdot \frac{2}{3}}} = \frac{7{,}3333}{13{,}3333} = 0{,}55.$$

За таблицею для функції $\Phi(x)$ (дод. 2) знаходимо:

$$\Phi(-0{,}5) = -\Phi(0{,}5) = -0{,}1915; \quad \Phi(0{,}55) = 0{,}2088.$$

Отже, за формулою (2.8):

$$P_{400}(260 \le k \le 274) = \Phi(0{,}55) - \Phi(-0{,}5) = 0{,}2088 + 0{,}1915 = 0{,}4003.$$

За допомогою формули (2.8) можна оцінити відхилення відносної частоти $w(A)$ (1.10) від імовірності p випадкової події A.

Наслідок *(З інтегральної теореми Муавра-Лапласа).* Якщо p – імовірність появи випадкової події A в кожному експерименті за схемою Бернуллі $(0 < p < 1)$ і $w(A)$ – відносна частота появи цієї події при n експериментах, то при великих значеннях n можна використовувати формулу

$$\underline{P\left(|w(A) - p| \le \varepsilon\right) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right),} \qquad (2.10)$$

де $\varepsilon > 0$ і є малою величиною.

Дійсно, за формулою (2.8) та (1.10) отримаємо:

$$P\left(|w(A) - p| \le \varepsilon\right) = P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) = P\left(-\varepsilon \le \frac{m}{n} - p \le \varepsilon\right) = P\left(-\varepsilon \le \frac{m - np}{n} \le \varepsilon\right) =$$

$$= P\left(-\varepsilon\sqrt{\frac{n}{pq}} \le \frac{m - np}{\sqrt{npq}} \le \varepsilon\sqrt{\frac{n}{pq}}\right) = \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Приклад. Імовірність того, що пасажир запізниться до відправлення потяга дорівнює 0,1. Було перевірено 400 пасажирів. Чому дорівнює ймовірність того, що абсолютна величина відхилення відносної частоти запізнення пасажирів від ймовірності $p = 0{,}1$ становить $\varepsilon = 0{,}001$?

<u>Розв'язання.</u> За умовою: $n = 400; p = 0{,}1; q = 0{,}9; \varepsilon = 0{,}001$. За формулою (2.10):

$$P\left(|w(A) - 0{,}1| \le 0{,}001\right) = 2\Phi\left(0{,}001\sqrt{\frac{400}{0{,}1 \cdot 0{,}9}}\right) = 2\Phi\left(0{,}07\right) = 0{,}0558.$$

Завдання

1. Фабрика випускає 75% виробів 1-го сорту. Із партії готових виробів навмання беруть 400 деталей. Обчислити ймовірність таких випадкових подій: а) виробів 1-го сорту

виявиться 290 шт.; б) виробів 1-го сорту виявиться 300 шт.; в) виробів 1-го сорту виявиться 320 шт.

Відповідь: а) 0,024; б) 0,046; в) 0,003.

2. У 400 потягах, що прибувають на розформування за 15 діб на станцію *A,* ймовірність появи місцевих вагонів $p = 0{,}2$. Знайти ймовірність того, що місцеві вагони прибувають у 80 потягах.

Відповідь. 0,05.

3. Відомо, що 60% пасажирів від загальної кількості пасажирів потяга даного напрямку складає пільговий контингент. Яка ймовірність того, що: а) з 10 пасажирів виявиться 6 пасажирів пільгового контингенту; б) з 200 пасажирів виявиться 120 пасажирів пільгового контингенту?

Відповідь: а) 0,251; б) 0,058.

4. Завод випускає в середньому 4% нестандартних виробів. Яка ймовірність того, що число нестандартних виробів у партії з 4000 шт. не більше 170?

Відповідь. 0,8.

5. Імовірність того, що покупець, який завітав до магазину взуття, здійснить покупку, дорівнює в середньому 0,1. Яка ймовірність того, що із 900 покупців, що завітали до магазину, здійснять покупку: а) 90 покупців; б) від 100 до 180 покупців?

Відповідь: а) 0,044; б) 0,134.

6. Імовірність появи випадкової події в кожному зі 100 незалежних експериментів дорівнює 0,3. З якою ймовірністю можна стверджувати, що відносна частота появи випадкової події знаходиться в межах $[0{,}2;0{,}4]$.

Відповідь. 0,971.

7. Імовірність появи випадкової події в кожному з 800 незалежних експериментів є величиною сталою і дорівнює 0,7. Яким має бути значення $\varepsilon > 0$, щоб $P(|w(A) - p| \le \varepsilon) = 0{,}99$?

Відповідь. 0,042.

8. За статистичними даними біля 87% пасажирів потяга Харків-Москва прямують до кінцевої зупинки. Знайти

ймовірність того, що з 1000 пасажирів відносна частота того, що пасажири прямують до кінцевої зупинки: а) міститься в межах від 0,9 до 0,95; б) відрізняється від імовірності цієї події не більш як на 0,04 (за абсолютною величиною).

Відповідь: а) 0,0024; б) 0,9998.

9. До дирекції залізничних перевезень направлено 50 молодих спеціалістів. Згодом навмання було відібрано 200 працівників цієї дирекції, серед яких 5 виявились молодими спеціалістами. Визначити з імовірністю 0,9 кількість працівників даної дирекції.

Відповідь. $980 \le N \le 4082$.

2.3. Формула Пуассона

Точність асимптотичних формул Муавра-Лапласа для великих значень n (числа повторних незалежних експериментів за схемою Бернуллі) знижується з наближенням p до нуля. Тому в цьому випадку користуються іншою асимптотичною формулою – формулою Пуассона.

Теорема 2.3 *(Теорема Пуассона).* Якщо $n \to +\infty,\ p \to 0$ так, що $np \to \lambda = const$, то $\forall k = 0,1,\ldots$

$$P_n(k) = C_n^k p^k q^{n-k} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \qquad (2.11)$$

де $\lambda = np.$

Дійсно, за формулою Бернуллі:

$$P_n(k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k} =$$

$$= \frac{(n-k+1)(n-k+2)(n-k+3)\ldots(n-1)n}{k!} \cdot p^k \cdot (1-p)^{n-k}.$$

Оскільки $\lambda = np,$ то $p = \dfrac{\lambda}{n},\ q = 1 - p = 1 - \dfrac{\lambda}{n},$ отримаємо:

$$P_n(k)=\frac{(n-k+1)(n-k+2)(n-k+3)\ldots(n-1)n}{k!}\cdot\left(\frac{\lambda}{n}\right)^k\cdot\left(1-\frac{\lambda}{n}\right)^{n-k}=$$

$$=\frac{\lambda^k}{k!}\cdot\left(1-\frac{k-1}{n}\right)\cdot\left(1-\frac{k-2}{n}\right)\cdot\ldots\cdot\left(1-\frac{1}{n}\right)\cdot 1\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-k}. \quad (2.12)$$

Оскільки

$$\lim_{n\to+\infty}\left(1-\frac{k-1}{n}\right)=\lim_{n\to+\infty}\left(1-\frac{k-2}{n}\right)=\lim_{n\to+\infty}\left(1-\frac{k-3}{n}\right)=\ldots=\lim_{n\to+\infty}\left(1-\frac{1}{n}\right)=1;$$

$$\lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda};\ \lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^{-k}=1,$$

то з формули (2.12) отримаємо:

$$\lim_{n\to+\infty}P_n(k)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}.$$

Таким чином, ймовірність $P_n(k)$ появи випадкової події k раз $(0\le k\le n)$ обчислюється за асимптотичною формулою

$$\underline{P_n(k)\cong\frac{\lambda^k}{k!}\cdot e^{-\lambda},} \quad (2.13)$$

де $\lambda=np,$ яка називається *формулою Пуассона*.

Значення функції $P_n(k,\lambda)=\dfrac{\lambda^k}{k!}\cdot e^{-\lambda}$ визначаються за таблицею, наведеною в дод. 3, за заданим k і обчисленим значенням $\lambda=np.$

<u>Зауваження.</u> Якщо ймовірність p стала та мала $(p<0{,}1)$, кількість експериментів n велика, $\lambda=np\to const$ та $\underline{npq<9}$, то наближене значення ймовірності $P_n(k)$ обчислюється за формулою (2.13) і використовується як точне.

Приклад. При огляді потягів у парку відправлення в кожному з них з імовірністю $p=0{,}05$ є вагони, що вимагають ремонту. Щодоби зі станції відправляються 120 потягів. Яка ймовірність того, що у 12 з них є пошкоджені вагони?

Розв'язання. Оскільки $p = 0{,}05\ (p < 0{,}1)$, $n = 120$, $q = 0{,}95$ та $npq = 120 \cdot 0{,}05 \cdot 0{,}95 = 5{,}7 < 9$, то за формулою Пуассона (2.13) отримаємо:

$$\lambda = np = 120 \cdot 0{,}05 = 6; \quad P_{120}(12) = \frac{6^{12}}{12!} \cdot e^{-6} \approx 0{,}0003.$$

Із формули (2.13) випливає:

$$P_n(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} P_n(k) = \sum_{k=k_1}^{k_2} \frac{\lambda^k}{k!} \cdot e^{-\lambda}. \tag{2.14}$$

Приклад. На станції працює 300 робітників. Імовірність того, що протягом доби працівник захворіє, дорівнює у середньому 0,02. Яка ймовірність того, що протягом доби на станції захворіє: а) 5 осіб; б) не більш як 3 особи?

Розв'язання. За умовою задачі $n = 300$; $k = 5$; $p = 0{,}02$; $q = 1 - 0{,}02 = 0{,}98$. Оскільки $npq = 300 \cdot 0{,}02 \cdot 0{,}98 = 5{,}88 < 9$, в даному випадку слід скористатися формулою Пуассона. Обчислюємо $\lambda = np = 300 \cdot 0{,}02 = 6$ і за даними дод. 3 отримаємо ймовірності:

а) $P_{300}(5) = \dfrac{6^5}{5!} \cdot e^{-6} = 0{,}160623$;

б) $P_{300}(0 \le k \le 3) = P_{300}(0) + P_{300}(1) + P_{300}(2) + P_{300}(3) =$

$$= \frac{6^0}{0!} \cdot e^{-6} + \frac{6^1}{1!} \cdot e^{-6} + \frac{6^2}{2!} \cdot e^{-6} + \frac{6^3}{3!} \cdot e^{-6} =$$

$$= 0{,}002479 + 0{,}014873 + 0{,}044618 + 0{,}089235 = 0{,}151205.$$

Завдання

1. Підручник надруковано тиражем 10000 екземплярів. Імовірність того, що підручник зброшуровано неправильно, дорівнює 0,0001. Яка ймовірність того, що в тиражі 5 бракованих книг?

Відповідь. 0,003.

2. Завод відправляє на базу 1500 виробів. Імовірність пошкодження виробу в дорозі дорівнює 0,002. Яка ймовірність того, що в дорозі буде пошкоджено виробів: а) рівно 4; б) менше 4; в) більше 4?

Відповідь: а) 0,17; б) 0,65; в) 0,65.

3. Імовірність того, що пасажир запізниться до відправлення потяга, дорівнює 0,01. Знайти найімовірніше число пасажирів, що запізняться, з 800 пасажирів та ймовірність цього числа.

Відповідь. $k_0 = 8; P_{800}(8) = 0,14.$

4. Ймовірність того, що будь-який абонент подзвонить на комутатор протягом години, становить 0,01. Телефонна станція обслуговує 800 абонентів. Яка ймовірність того, що протягом години зателефонують п'ять абонентів?

Відповідь. 0,09.

Питання до теми

29. Наведіть приклади повторних незалежних випробувань.
30. Що називається схемою Бернуллі? Яка формула служить її моделлю?
31. Як знайти найімовірніше число успіхів в схемі Бернуллі? Чи є воно єдиним?
32. Коли неможливо скористатися формулою Бернуллі? Як називаються формули, що використовуються в цьому випадку?
33. Що можна обчислити за допомогою локальної теореми Муавра-Лапласа? Наведіть відповідну формулу. В якому випадку вона використовується?
34. Які властивості функції Гаусса вам відомі? Як знайти її значення?
35. Що обчислюють за допомогою інтегральної теореми Лапласа? Наведіть відповідну формулу. За якої умови вона використовується як точна?
36. Які властивості функції Лапласа ви знаєте? Як знайти її значення?
37. Як оцінити відхилення відносної частоти випадкової події від її ймовірності?

38. Що дозволяє обчислити теорема Пуассона? Наведіть відповідну формулу. В якому випадку вона використовується?

39. Як обрати асимптотичну формулу для обчислення потрібної ймовірності?

40. Які табульовані функції вам відомі?

Тестові питання

1. Яка з даних формул є формулою Бернуллі?

А	Б	В	Г	Д
$P_n(k)=C_n^k p^k$	$P_n(k)=p^k q^{n-k}$	$P_n(k)=C_n^k p^k q^{n-k}$	$P_n(k)=np^k$	Інша відпо-відь

2. Ймовірність випадіння трьох гербів при підкиданні монети чотири рази дорівнює:

А	Б	В	Г	Д
$\frac{3}{4}$	$3\cdot(0,5)^5$	$(0,5)^3$	$4\cdot(0,5)^5$	Інша відповідь

3. Ймовірність зайнятості колії сортувального парку складає 0,6. Ймовірність того, що з п'яти колій парку зайнятими є дві, дорівнює:

А	Б	В	Г	Д
$C_5^2\cdot(0,6)^2$	$C_5^2\cdot(0,6)^2\cdot(0,4)^3$	$(0,6)^2\cdot(0,4)^3$	$5\cdot(0,6)^2$	Інша відповідь

4. Теореми Муавра-Лапласа використовуються за умови:

А	Б	В	Г	Д
$npq=9$	$npq<9$	$npq\geq 9$	$\forall n,p,q$	Інша відповідь

5. За допомогою інтегральної теореми Муавра-Лапласа не можна обчислити:

А	Б	В	Г	Д
$P_n(k)$	$P_n(k \le k_2)$	$P_n(k_1 \le k \le k_2)$	$P_n(k \ge k_1)$	Інша відповідь

6. Ймовірність помилкового набору номера складає 0,003. Щоб знайти ймовірність того, що з 1000 абонентів двоє наберуть номер невірно, потрібно скористатися:

А	Б	В	Г	Д
формулою Бернуллі	локальною формулою Муавра-Лапласа	інтегральною формулою Муавра-Лапласа	формулою Пуассона	Інша відповідь

7. З наведених нижче функцією Гаусса є:

А	Б	В	Г	Д
$P(k,\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$	$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$	$P_n(k) = C_n^k p^k q^{n-k}$	$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$	Жодна

<u>Відповіді:</u> 1.В; 2.Г; 3.Б; 4. В; 5.А; 6.Г; 7.Б.

Розділ 3. Випадкові величини

3.1. Поняття випадкової величини

Випадковою називається величина, що внаслідок експерименту набуває певного значення, заздалегідь невідомо, якого.

Прикладами випадкових величин є: кількість пасажирів, що звернулись за квитками в касу протягом години; кількість вагонів, що надійшли на вантажний двір протягом доби; час очікування пасажиром поїзда метрополітену; вага навантажених вагонів, що прибувають на вантажний двір.

Дискретною (перервною) називається випадкова величина, що набуває окремих ізольованих значеннь з певними ймовірностями. Множина значень дискретної випадкової величини (ДВВ) або скінченна, або нескінченна, але злічена (елементи її можна занумерувати натуральними числами).

Неперервною називається випадкова величина, всі можливі значення якої заповнюють деякий проміжок числової осі. На відміну від дискретної, значення неперервної випадкової величини (НВВ) перелічити неможливо.

З наведених вище прикладів, очевидно, перший і другий відповідають дискретним випадковим величинам, а третій і четвертий – неперервним .

Випадкові величини зазвичай позначаються великими буквами латинського алфавіту $X, Y, Z\ldots$, а їх значення – відповідними малими літерами $x, y, z\ldots$

3.2. Закон розподілу дискретної випадкової величини

Найбільш повним, вичерпним описом випадкової величини є її закон розподілу ймовірностей. Про випадкову величину кажуть, що вона розподілена за даним законом або підлегла йому.

Законом розподілу випадкової величини називається відповідність між можливими значеннями випадкової величини та їх імовірностями.

Закон розподілу дискретної випадкової величини X може бути заданий:

- у вигляді таблиці (рядом розподілу);
- графічно (полігоном);
- аналітично (формулою).

Найчастіше закон розподілу ДВВ X задається *в табличному вигляді*:

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

(3.1)

де у першому рядку перелічені в порядку зростання всі можливі значення $x_1,\dots,x_n$ випадкової величини X, в другому – відповідні їм імовірності $p_1,\dots,p_n$. Оскільки події $X = x_1, \dots, X = x_n$ утворюють повну групу, то $\sum_i p_i = p_1 + \dots + p_n \equiv 1$. Формула (3.1) також називається *рядом розподілу* ДВВ X.

При *графічному задаванні* ДВВ на прямокутній системі координат будують точки з координатами $(x_i; p_i)$ та послідовно (у порядку зростання номерів) з'єднують відрізками. Отриману ламану називають *багатокутником розподілу* або *полігоном* (рис. 3.1).

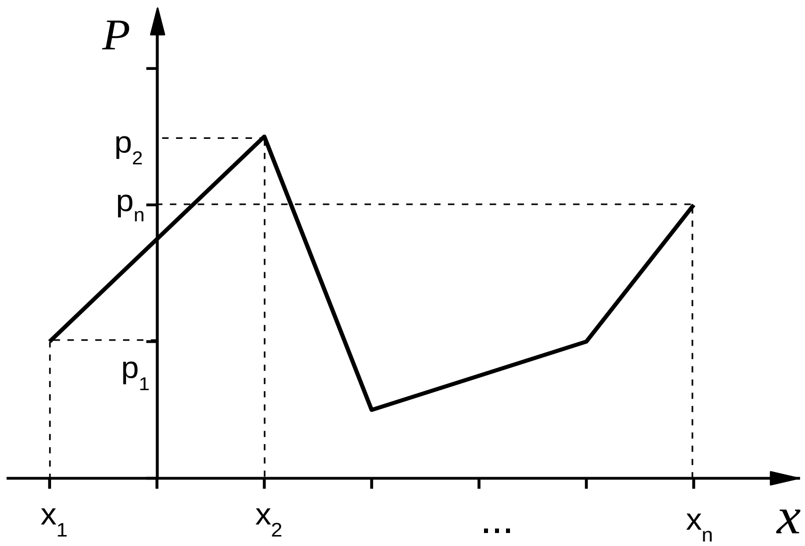


Рис. 3.1

При *аналітичному задаванні* ДВВ відповідність між значеннями x_i та їх імовірностями p_i задається формулою виду:

$$p_i = P(X = x_i) = \varphi(x_i),$$

де $\varphi(x_i)$ − деяка функція.

Зауваження. Зазвичай аналітично задаються спеціальні (стандартні) розподіли, що будуть розглянуті у наступному розділі.

Модою M_0 дискретної випадкової величини X називається її найімовірніше значення. З геометричної точки зору, мода є абсцисою такої точки багатокутника розподілу (полігона), ордината якої максимальна.

Приклад. У парку приймання 3 колії. Ймовірність зайнятості кожної з них потягами, що прибувають, $p = 0{,}8$. Знайти розподіл числа зайнятих колій. Визначити моду M_0.

Розв'язання. Позначимо через X дискретну випадкову величину, що дорівнює числу колій, зайнятих потягами, що прибувають. Можливі значення X: $x_1 = 0; x_2 = 1; x_3 = 2; x_4 = 3$. Ймовірності p_i значень x_i знайдемо за формулою Бернуллі $P_n(k) = C_n^k p^k q^{n-k}$, де $n = 3;\ p = 0{,}8;\ q = 1 - p = 0{,}2$. Отримаємо:

$$p_1 = p(x_1) = P_3(0) = C_3^0 p^0 q^3 = 1 \cdot 1 \cdot (0{,}2)^3 = 0{,}008;$$

$$p_2 = p(x_2) = P_3(1) = C_3^1 p^1 q^2 = 3 \cdot (0{,}8) \cdot (0{,}2)^2 = 0{,}096;$$

$$p_3 = p(x_3) = P_3(2) = C_3^2 p^2 q^1 = 3 \cdot (0{,}8)^2 \cdot 0{,}2 = 0{,}384;$$

$$p_4 = p(x_4) = P_3(3) = C_3^3 p^3 q^0 = 1 \cdot (0{,}8)^3 \cdot 1 = 0{,}512.$$

Складемо розподіл ДВВ X в табличному вигляді:

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

Перевірка: $\sum_i p_i =$ 0,512+0,384+0,096+0,008=1.

Для даного розподілу мода дорівнює $M_0 = 3$. Багатокутник розподілу побудуємо за точками з координатами $(x_i; p_i)$ (рис.3.2).

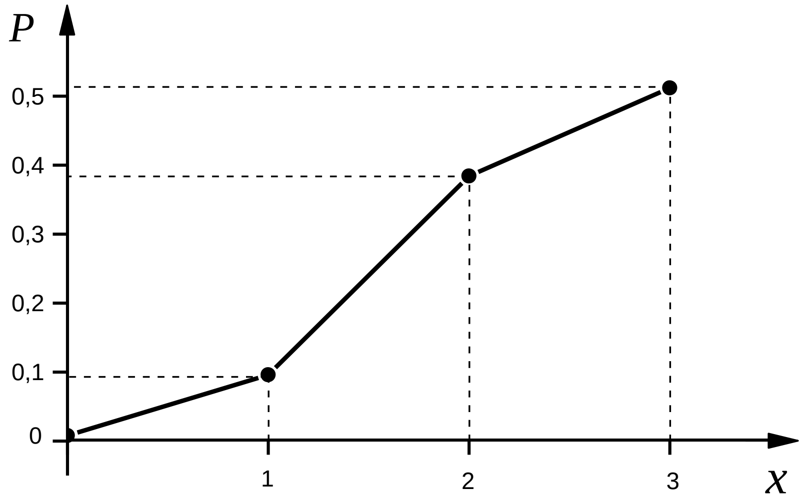


Рис. 3.2

Даний розподіл може бути заданий і аналітично, а саме формулою

$$P(X = k) = C_3^k \cdot (0{,}8)^k \cdot (0{,}2)^{3-k}, \quad k = 0{,}1{,}2{,}3.$$

Розподіли такого типу називаються *біноміальними* (див. п. 4.1).

Приклад. На шляху руху потяга є чотири світлофори. Кожен з них з ймовірністю 0,5 або дозволяє, або забороняє потягу подальший рух. Побудувати розподіл ймовірностей числа світлофорів, що були пройдені до першої зупинки, та визначити моду M_0.

Розв'язання. Нехай X – ДВВ, що дорівнює кількості світлофорів, що були пройдені до першої зупинки. Можливі значення X: $x_1 = 0$; $x_2 = 1$; $x_3 = 2$; $x_4 = 3$; $x_5 = 4$. Ймовірності p_i значень x_i знаходимо за формулою:

$$p_i = p(X = x_i) = \begin{cases} p \cdot (1-p)^{i-1} & при \quad i = \overline{1,4}; \\ (1-p)^4 & при \quad i = 5, \end{cases}$$

де $p = 0{,}5$ – ймовірність зупинки потяга на світлофорі. Розподіл ймовірностей ДВВ X запишемо у вигляді таблиці:

x_i	0	1	2	3	4
p_i	0,5	0,25	0,125	0,0625	0,0625

Перевірка: $\sum_i p_i = 0{,}5+0{,}25+0{,}125+0{,}0625+0{,}0625=1$.

З розподілу X випливає: найімовірніше, що потяг зупиниться вже на першому світлофорі. Тобто мода ДВВ X дорівнює $M_0 = 0$. Багатокутник розподілу X зображено на рис. 3.3.

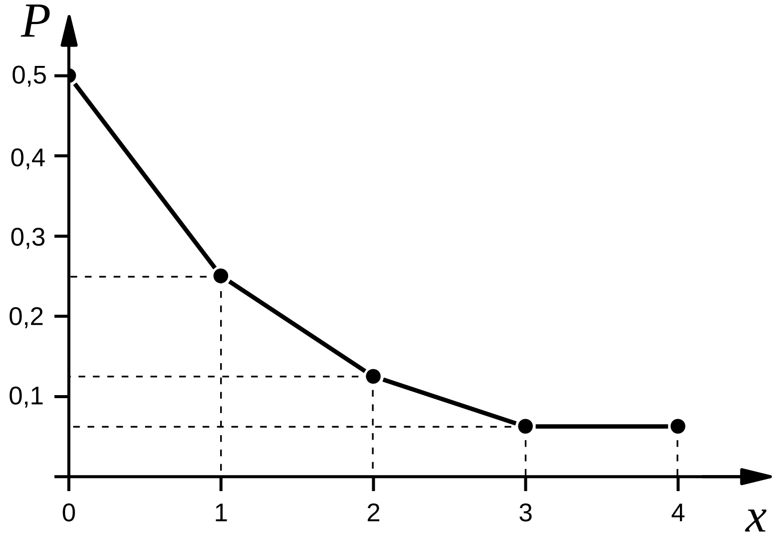


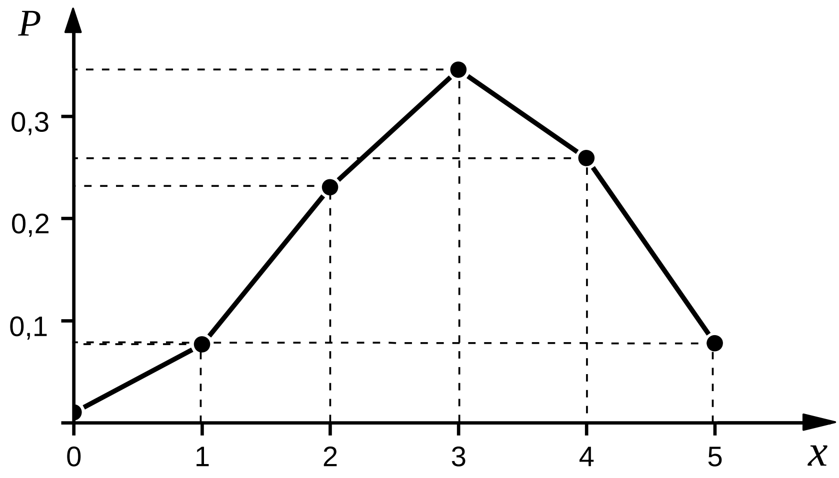
Рис. 3.3

Завдання

1. У кожному поїзді, що прибуває для розформування, з ймовірністю $p = 0{,}6$ можуть бути вагони на дане призначення. Протягом години на станцію прибуло $n = 5$ поїздів для розформування. Скласти закон розподілу кількості поїздів, в яких є вагони на дане призначення. Побудувати багатокутник розподілу. Визначити моду.

<u>Відповідь.</u> $P(X = k) = C_5^k \cdot (0{,}6)^k \cdot (0{,}4)^{5-k}$, $k = \overline{0{,}5}$. $M_0 = 3$.

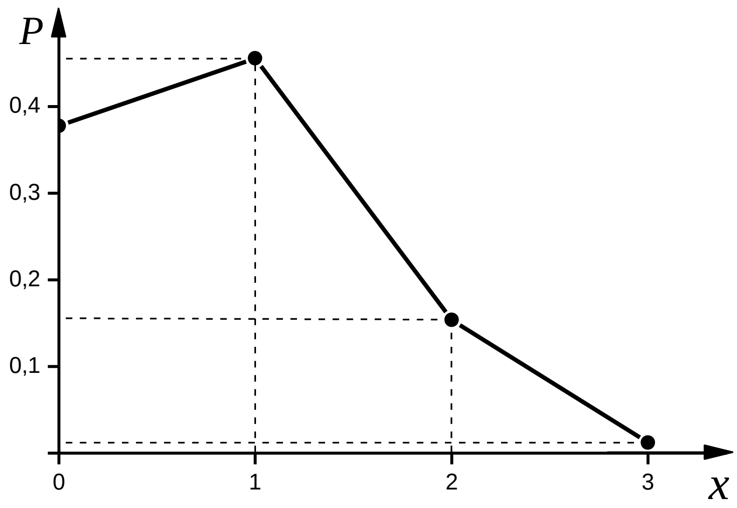
x_i	0	1	2	3	4	5
p_i	0,01024	0,0768	0,2304	0,3456	0,2592	0,07776



2. Пристрій складається з трьох вузлів, ймовірності безвідмовної роботи яких протягом часу T дорівнюють 0,9, 0,7 та 0,6 відповідно. Скласти розподіл кількості вузлів, що відмовили протягом часу T. Побудувати багатокутник розподілу. Визначити моду.

x_i	0	1	2	3
p_i	0,378	0,456	0,154	0,012

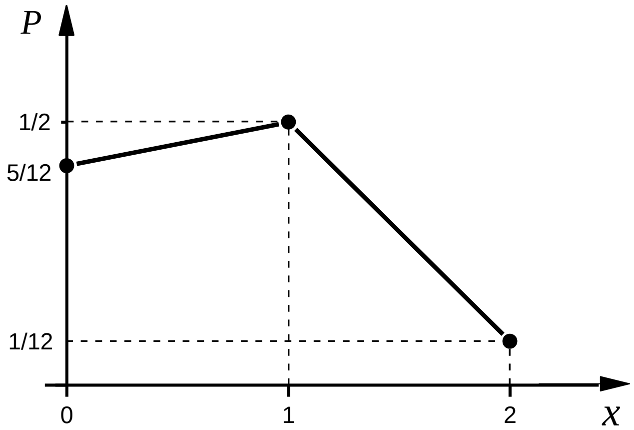
Відповідь. $M_0 = 1$



3. Студент вивчив сім з дев’яти тем екзамену. В білеті міститься три питання з різних тем. Скласти закон розподілу числа невідомих питань в білеті, побудувати полігон.

Відповідь. $P(X=k)=\dfrac{C_2^k\cdot C_7^{3-k}}{C_9^3}$, $k=0,1,2.$

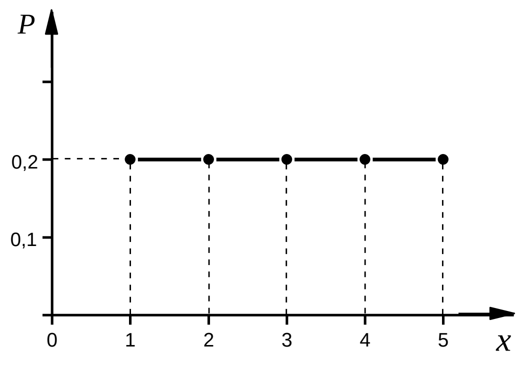
x_i	0	1	2
p_i	$\dfrac{C_7^3}{C_9^3}=\dfrac{5}{12}$	$\dfrac{C_2^1\cdot C_7^2}{C_9^3}=\dfrac{1}{2}$	$\dfrac{C_2^2\cdot C_7^1}{C_9^3}=\dfrac{1}{12}$



4. Один з п’яти приладів є несправним. Щоб його виявити, навмання перевіряють по одному приладу до появи несправного. Скласти розподіл кількості перевірених приладів в табличному та графічному вигляді, знайти моду.

x_i	1	2	3	4	5
p_i	0,2	0,2	0,2	0,2	0,2

Відповідь.



$M_0=1;2;3;4;5$.

5. Імовірність спізнення потяга на станцію дорівнює 0,1. За годину прибуває чотири потяги. Скласти закон розподілу числа потягів, що прибули вчасно протягом години.

<u>Відповідь.</u> $P(X=k)=C_4^k\cdot(0{,}9)^k\cdot(0{,}1)^{4-k},\quad k=0{,}1{,}2{,}3{,}4.$

x_i	0	1	2	3	4
p_i	0,0001	0,0036	0,0486	0,2916	0,6561

6. В офісі працюють два секретарі. Ймовірності припущення ними помилки при підготовці одного документа дорівнюють 0,1 та 0,2 відповідно. З трьох документів два були підготовлені першим, один – другим секретарем. Знайти розподіл кількості безпомилково підготовлених документів.

<u>Відповідь.</u>

x_i	0	1	2	3
p_i	0,002	0,044	0,306	0,648

7. З 10 пристроїв 2 мають приховані дефекти. Знайти закон розподілу числа справних пристроїв серед трьох навмання обраних.

<u>Відповідь.</u> $P(X=k)=\dfrac{C_8^k\cdot C_2^{3-k}}{C_{10}^3},\ k=1{,}2{,}3.$

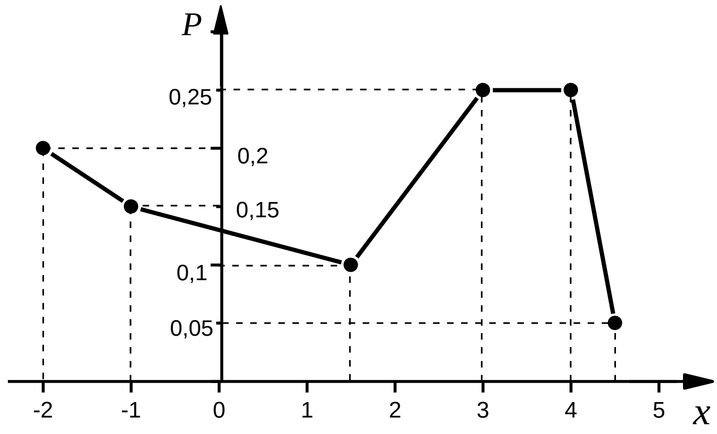
x_i	1	2	3
p_i	$\dfrac{C_8^1\cdot C_2^2}{C_{10}^3}=\dfrac{1}{15}$	$\dfrac{C_8^2\cdot C_2^1}{C_{10}^3}=\dfrac{7}{15}$	$\dfrac{C_8^3}{C_{10}^3}=\dfrac{7}{15}$

8. Ймовірність влучення стрільцем в ціль дорівнює 0,8. Стрільба ведеться до першого враження цілі. Скласти розподіл кількості здійснених пострілів, якщо стрілець має п’ять патронів.

Відповідь. $P(X=k)=\begin{cases} 0{,}8\cdot(0{,}2)^{k-1} & при \quad k=\overline{1,4}; \\ (0{,}2)^4 & при \quad k=5. \end{cases}$

x_i	1	2	3	4	5
p_i	0,8	0,16	0,032	0,0064	0,0016

9. Відновити ряд розподілу ДВВ X за її полігоном та визначити моду.



x_i	-2	-1	1,5	3	4	4,5
p_i	0,2	0,15	0,1	0,25	0,25	0,05

Відповідь. $M_0 = 3;4$.

10. Відновити ряд розподілу ДВВ X, заданої таблицею

x_i	-3	-1	0	1	3
p_i	0,18	$2k$	0,31	0,24	k

та визначити її моду.

Відповідь. $k = 0{,}9$; $M_0 = 0$.

x_i	-3	-1	0	1	3
p_i	0,18	0,18	0,31	0,24	0,09

3.3. Функція розподілу дискретних випадкової величини

Окрім закону розподілу, однозначно описує розподіл ймовірностей випадкової величини її функція розподілу.

Функцією розподілу ймовірностей випадкової величини X або *інтегральною функцією* називається ймовірність того, що в результаті експерименту X набуде значення, меншого за аргумент x, тобто функція

$$F(x) = P(X < x), \qquad x \in R. \tag{3.2}$$

Зауваження. Дане означення функції розподілу є спільним як для дискретних, так і для неперервних випадкових величин. З означення безпосередньо випливають нижченаведені властивості $F(x)$.

Властивості функції розподілу $F(x)$:

1) як імовірність, $0 \le F(x) \le 1$;

2) функція розподілу є неспадною: якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$;

3) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1$;

4) імовірність попадання випадкової величини X в півінтервал $[\alpha; \beta)$ дорівнює різниці значень функції розподілу в цих точках:

$$P(\alpha \le X < \beta) = F(\beta) - F(\alpha). \tag{3.3}$$

Наведені вище властивості є спільними для функцій розподілу як дискретних, так і неперервних випадкових величин. Крім того, існує і ряд суттєвих відмінностей. Зараз ми переходимо до вивчення функції розподілу для дискретних випадкових величин.

Для ДВВ X функція розподілу має вигляд

$$F(x) = P(X < x) = \sum_{x_i < x} p_i, \qquad x \in R, \tag{3.4}$$

тобто $F(x)$ дорівнює сумі ймовірностей значень X, що менші за x (лежать зліва від x). Таким чином, функція розподілу ймовірностей ДВВ є *кусково-сталою*, а тому графік має *ступінчастий вигляд*.

Приклад. Рухаючись з гірки на колію сортувального парку, відчеп проходить чотири стрілочні переводи, на кожному з яких він може зупинитись. Число стрілочних переводів X, що пройдені відчепом до зупинки, задається рядом розподілу:

x_i	0	1	2	3	4
p_i	0,05	0,1	0,2	0,3	0,35

Побудувати функцію розподілу ДВВ X. Знайти ймовірність того, що відчеп зупиниться не менше, ніж на двох стрілочних переводах.

Розв'язання. Функція розподілу ймовірностей X задається формулою

$$F(x) = P(X < x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ 0{,}05 & \text{при} \quad 0 < x \le 1, \\ 0{,}15 & \text{при} \quad 1 < x \le 2, \\ 0{,}35 & \text{при} \quad 2 < x \le 3, \\ 0{,}65 & \text{при} \quad 3 < x \le 4, \\ 1 & \text{при} \quad x > 4, \end{cases}$$

а її графік зображений на рис. 3.4. Ймовірність того, що зупинок буде не менше двох, знайдемо за формулою (3.3):

$$P(X \ge 2) = P(2 \le X < +\infty) = F(+\infty) - F(2) = 1 - 0{,}15 = 0{,}85.$$

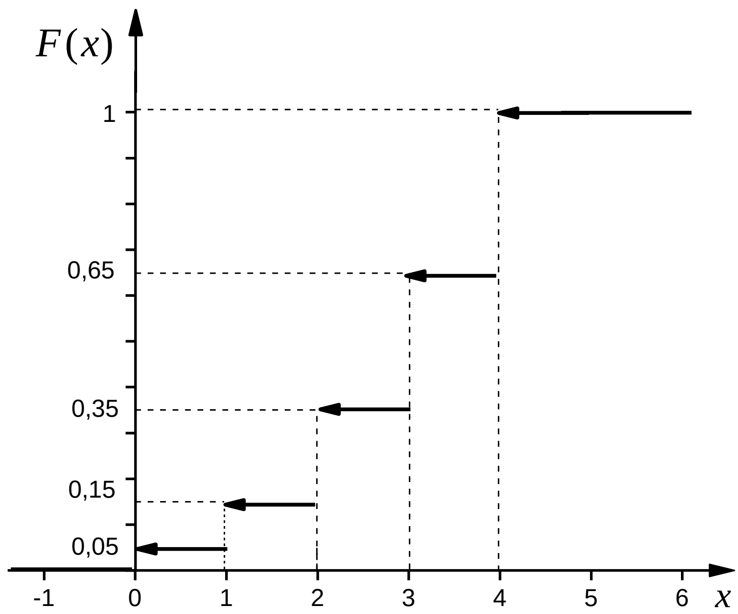


Рис. 3.4

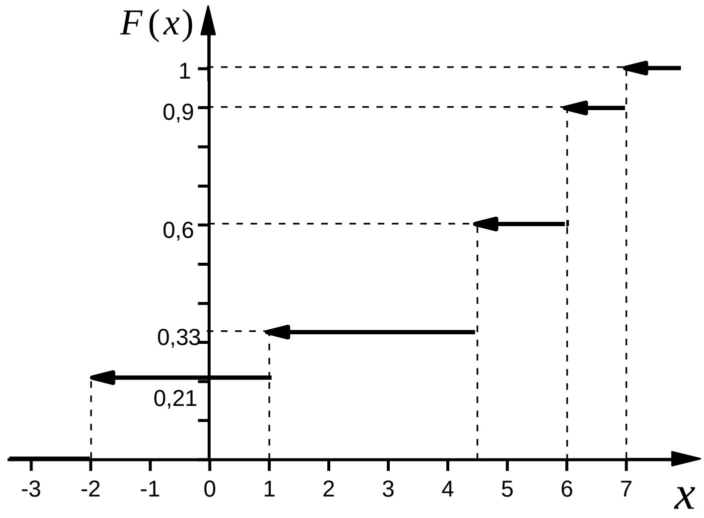
Завдання

1. ДВВ X задана законом розподілу в табличній формі:

x_i	-2	1	4,5	6	7
p_i	0,21	0,12	?	0,3	0,1

Знайти: невідому ймовірність, функцію розподілу $F(x)$, ймовірності $P(X<6)$, $P(X \le 6)$.

Відповідь. $p_3 = 0{,}27$; $F(x) = \begin{cases} 0 & \text{при} \quad x \le -2, \\ 0{,}21 & \text{при} \quad -2 < x \le 1, \\ 0{,}33 & \text{при} \quad 1 < x \le 4{,}5, \\ 0{,}6 & \text{при} \quad 4{,}5 < x \le 6, \\ 0{,}9 & \text{при} \quad 6 < x \le 7, \\ 1 & \text{при} \quad x > 7. \end{cases}$



$$P(X<6)=0{,}6;\ P(X\le 6)=0{,}9.$$

2. Побудувати функцію розподілу ДВВ X, заданої рядом розподілу

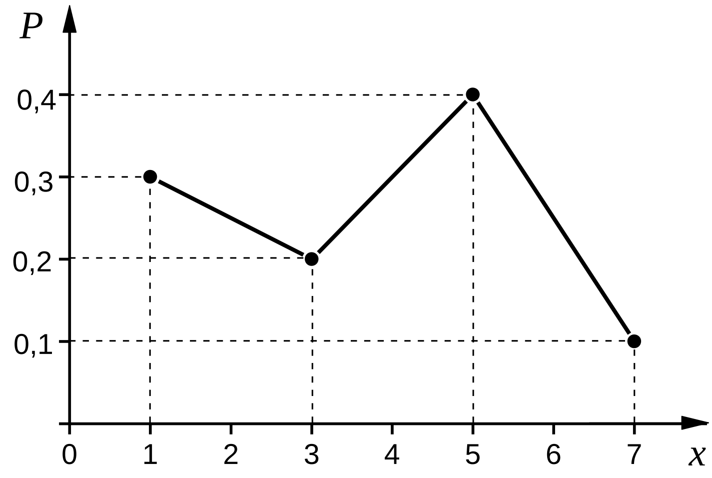
x_i	-3	0	1	2	5
p_i	k	$0{,}2$	$0{,}4$	$0{,}1$	$3k$

Знайти ймовірності того, що випадкова величина набуде: а) додатного значення; б) невід'ємного значення; в) значення не менше 1 і не більше 4.

<u>Відповідь.</u> $k=0{,}075$; $F(x)=\begin{cases} 0 & \text{при} \quad x\le -3, \\ 0{,}075 & \text{при} \quad -3<x\le 0, \\ 0{,}275 & \text{при} \quad 0<x\le 1, \\ 0{,}675 & \text{при} \quad 1<x\le 2, \\ 0{,}775 & \text{при} \quad 2<x\le 5, \\ 1 & \text{при} \quad x>5. \end{cases}$

а) $P(X>0)=0{,}725$; б) $P(X\ge 0)=0{,}925$; в) $P(1\le X\le 4)=0{,}5$.

3. Побудувати функцію розподілу випадкової величини, заданої своїм багатокутником розподілу:



Обчислити ймовірності $P(X<5)$, $P(X\ge 5)$.

Відповідь. $$F(x)=\begin{cases}0 & \text{при} \quad x\le 1,\\ 0{,}3 & \text{при} \quad 1<x\le 3,\\ 0{,}5 & \text{при} \quad 3<x\le 5,\\ 0{,}9 & \text{при} \quad 5<x\le 7,\\ 1 & \text{при} \quad x>7.\end{cases}$$

$P(X<5)=P(X\ge 5)=0{,}5.$

4. Функція розподілу ДВВ X має вигляд:

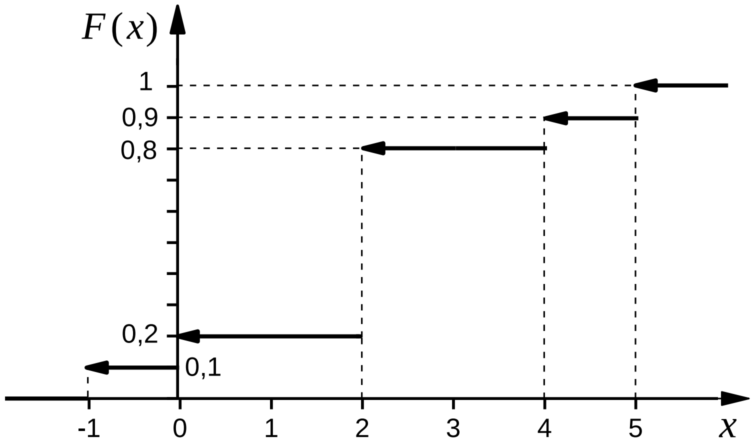
$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ 0{,}1 & \text{при} \quad 0<x\le 1,\\ 0{,}3 & \text{при} \quad 1<x\le 4,\\ 0{,}45 & \text{при} \quad 4<x\le 7,\\ 0{,}65 & \text{при} \quad 7<x\le 10,\\ 1 & \text{при} \quad x>10.\end{cases}$$

Побудувати ряд розподілу ДВВ. Обчислити ймовірності того, що X є: а) не меншою за 5; б) не меншою за 3 і меншою за 7.

x_i	0	1	4	7	10
p_i	0,1	0,2	0,15	0,2	0,35

Відповідь. а) $P(X\ge 5)=0{,}55$; б) $P(3\le X<7)=0{,}15$.

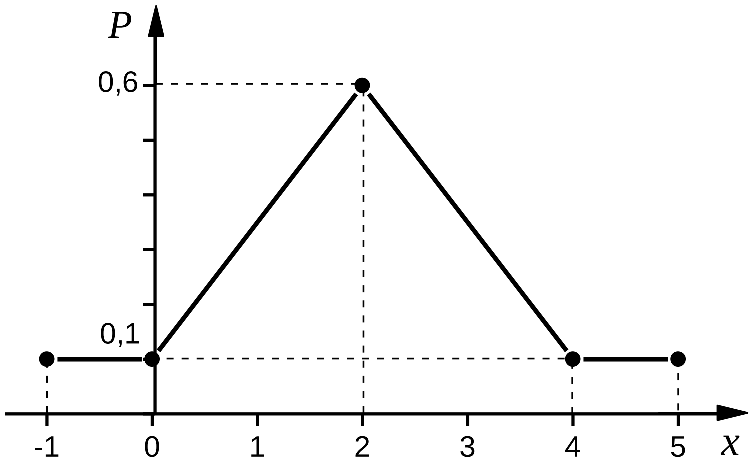
5. Графік функції розподілу ДВВ X має вигляд:



Побудувати ряд розподілу та полігон.

<u>Відповідь.</u>

x_i	-1	0	2	4	5
p_i	0,1	0,1	0,6	0,1	0,1



3.4. Дії над випадковими величинами

Для простоти будемо розглядати дискретні випадкові величини. Всі означення легко можна поширити на випадок неперервних величин.

Добутком ДВВ X *на число* C називається ДВВ CX, значення якої дорівнюють добутку кожного можливого значення x_i ДВВ X на число C, а ймовірності цих значень дорівнюють p_i:

CX	Cx_1	Cx_2	$\ldots$	Cx_n
P	p_1	p_2	$\ldots$	p_n

Наприклад, розподіли ДВВ X та відповідної ДВВ $3X$ мають вигляд:

X	-1	0	2
P	0,1	0,5	0,4

$3X$	-3	0	6
P	0,1	0,5	0,4

Дві випадкові величини називаються *незалежними*, якщо закон розподілу однієї з них не залежить від того, яких можливих значень набула інша.

Сумою випадкових величин X та Y називається випадкова величина $X+Y$, значення якої дорівнюють сумам $x_i + y_j$ кожного можливого значення X і кожного можливого значення Y, а відповідні ймовірності p_{ij} обчислюють за правилом:

а) $p_{ij} = p_i \cdot q_j$ у випадку незалежних випадкових величин;

б) $p_{ij} = p_i \cdot p_{x_i}(y_j)$, коли випадкові величини – залежні.

Якщо значення деяких сум виявляться рівними для різних пар (i, j), то для знаходження ймовірності цього значення відповідні ймовірності треба додати. Наприклад, якщо $x_1 + y_3 = x_2 + y_2$, то ймовірність появи цього значення дорівнює $p_{13} + p_{22}$.

Звичайно, це означення легко поширити на випадок багатьох доданків.

Приклад. Знайти суму незалежних ДВВ X та Y, що задані рядами розподілу:

X	-1	0	2
P	0,1	0,5	0,4

Y	3	4
P	0,7	0,3

Розв'язання. По-перше, знайдемо, яких значень набуває випадкова величина $X+Y$ та відповідні цим значенням ймовірності:

$x_1+y_1=-1+3=2$, $\quad p_{11}=p_1\cdot q_1=0{,}1\cdot 0{,}7=0{,}07$;

$x_1+y_2=-1+4=3$, $\quad p_{12}=p_1\cdot q_2=0{,}1\cdot 0{,}3=0{,}03$;

$x_2+y_1=0+3=3$, $\quad p_{21}=p_2\cdot q_1=0{,}5\cdot 0{,}7=0{,}35$;

$x_2+y_2=0+4=4$, $\quad p_{22}=p_2\cdot q_2=0{,}5\cdot 0{,}3=0{,}15$;

$x_3+y_1=2+3=5$, $\quad p_{31}=p_3\cdot q_1=0{,}4\cdot 0{,}7=0{,}28$;

$x_3+y_2=2+4=6$, $\quad p_{32}=p_3\cdot q_2=0{,}4\cdot 0{,}3=0{,}12$.

Оскільки $x_1+y_2=x_2+y_1=3$, то при формування ряду розподілу $X+Y$ цьому значенню відповідатиме ймовірність $p_{12}+p_{21}=0{,}03+0{,}35=0{,}38$. Остаточно розподіл суми даних випадкових величин має вигляд:

$X+Y$	2	3	4	5	6
P	0,07	0,38	0,15	0,28	0,12

Перевірка: $\sum_i p_i =$0,07+0,38+0,15+0,28+0,12=1.

Добутком випадкових величин X та Y називається випадкова величина $X\cdot Y$, значення якої дорівнюють добуткам $x_i\cdot y_j$ кожного можливого значення X на кожне можливе значення Y, а відповідні ймовірності p_{ij} обчислюють за правилом:

а) $p_{ij}=p_i\cdot q_j$ у випадку незалежних випадкових величин;

б) $p_{ij}=p_i\cdot p_{x_i}(y_j)$, коли випадкові величини – залежні.

Якщо значення деяких добутків рівні для різних пар (i,j), то для знаходження ймовірності цього значення відповідні ймовірності треба додати. Наприклад, якщо $x_1\cdot y_3=x_2\cdot y_2$, то ймовірність появи цього значення дорівнює $p_{13}+p_{22}$.

Це означення легко узагальнити на випадок багатьох множників.

Приклад. Знайти добуток незалежних ДВВ X та Y, що задані рядами розподілу:

X	-1	0	2
P	0,1	0,5	0,4

Y	3	4
P	0,7	0,3

<u>Розв’язання.</u> Знайдемо, яких значень набуває випадкова величина $X \cdot Y$ та відповідні цим значенням ймовірності:

$x_1 \cdot y_1 = -1 \cdot 3 = -3$, $\quad p_{11} = p_1 \cdot q_1 = 0{,}1 \cdot 0{,}7 = 0{,}07$;

$x_1 \cdot y_2 = -1 \cdot 4 = -4$, $\quad p_{12} = p_1 \cdot q_2 = 0{,}1 \cdot 0{,}3 = 0{,}03$;

$x_2 \cdot y_1 = 0 \cdot 3 = 0$, $\quad p_{21} = p_2 \cdot q_1 = 0{,}5 \cdot 0{,}7 = 0{,}35$;

$x_2 \cdot y_2 = 0 \cdot 4 = 0$, $\quad p_{22} = p_2 \cdot q_2 = 0{,}5 \cdot 0{,}3 = 0{,}15$;

$x_3 \cdot y_1 = 2 \cdot 3 = 6$, $\quad p_{31} = p_3 \cdot q_1 = 0{,}4 \cdot 0{,}7 = 0{,}28$;

$x_3 \cdot y_2 = 2 \cdot 4 = 8$, $\quad p_{32} = p_3 \cdot q_2 = 0{,}4 \cdot 0{,}3 = 0{,}12$.

Оскільки $x_2 \cdot y_1 = x_2 \cdot y_2 = 0$, то цьому значенню відповідає ймовірність $p_{21} + p_{22} = 0{,}35 + 0{,}15 = 0{,}5$. Таким чином, розподіл добутку $X \cdot Y$ випадкових величин X та Y має вигляд:

$X \cdot Y$	-4	-3	0	6	8
P	0,03	0,07	0,5	0,28	0,12

Перевірка: $\sum_i p_i =$0,03+0,07+0,5+0,28+0,12=1.

К-м степенем випадкової величини X називається випадкова величина X^k, значення якої дорівнюють x_i^k, а відповідні ймовірності дорівнюють p_i:

X^k	x_1^k	x_2^k	$\cdots$	x_n^k
P	p_1	p_2	...	p_n

Якщо значення степенів різних значень співпадають, то відповідні ймовірності треба додати. Наприклад, якщо $x_1^2 = x_3^2$, то ймовірність появи цього значення дорівнює $p_1 + p_3$.

Приклад. Знайти розподіл X^2, якщо ДВВ X задано таблицею:

X	-1	0	1	3
P	0,4	0,3	0,2	0,1

Розв’язання. Перелічимо усі можливі значення випадкової величини X^2 з відповідними ймовірностями:

$x_1{}^2 = (-1)^2 = 1, \qquad p_1 = 0{,}4;$

$x_2{}^2 = 0^2 = 0, \qquad p_2 = 0{,}3;$

$x_3{}^2 = 1^2 = 1, \qquad p_3 = 0{,}2;$

$x_4{}^2 = 3^2 = 9, \qquad p_4 = 0{,}1.$

Оскільки $x_1{}^2 = x_3{}^2 = 1$, цьому значенню відповідає ймовірність $p_1 + p_3 = 0{,}4 + 0{,}2 = 0{,}6$. Звідси розподіл X^2 має вигляд:

X^2	0	1	9
P	0,3	0,6	0,1

Завдання

1. Число вагонів призначенням на вантажний двір в поїзді, що прибуває, є випадковою величиною X, яка розподілена за законом

X	0	1	2	3	4
P	0,1	0,15	0,3	0,35	0,1

а число вагонів призначенням на контейнерний майданчик є випадковою величиною Y, яка має розподіл

Y	0	1	2
P	0,2	0,55	0,25

Скласти розподіл загального числа $X + Y$ вагонів призначенням на вантажний двір та на контейнерний майданчик.

Відповідь.

$X+Y$	0	1	2	3	4	5	6
P	0,02	0,085	0,1675	0,2725	0,2875	0,1425	0,025

2. На станції завантажуються піввагони та платформи. Кількість повновантажних піввагонів, відправлених із станції, є випадковою величиною X , яка має розподіл:

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

Кількість повновантажних платформ також є випадковою величиною Y з розподілом:

Y	0	1	2	3	4
P	0,05	0,2	0,35	0,25	0,15

Складти розподіл загального числа $X+Y$ повновантажних вагонів, що завантажуються. Знайти також розподіли $X \cdot Y$ та X^2.

Відповідь.

$X+Y$	1	2	3	4	5	6	7	8
P	0,02	0,095	0,21	0,25	0,225	0,13	0,055	0,015

$X \cdot Y$	0	1	2	3	4	6	8	9	12	16
P	0,05	0,08	0,2	0,14	0,185	0,145	0,08	0,05	0,055	0,015

X^2	1	4	9	16
P	0,4	0,3	0,2	0,1

3.5. Числові характеристики дискретної випадкової величини

Закон розподілу та функція розподілу повністю визначають випадкову величину, але вони не завжди відомі. Проте існують певні, пов'язані з випадковою величиною значення, яких виявляється достатньо для розв'язання широкого кола задач. Такі величини називаються *числовими характеристиками випадкової величини*.

3.5.1. Математичне сподівання

Математичним сподіванням $M(X)$ *ДВВ* X називається сума добутків значень випадкової величини на їх імовірності:

$$M(X) = x_1 \cdot p_1 + \ldots + x_n \cdot p_n = \sum_i x_i p_i . \qquad (3.5)$$

Імовірнісний зміст математичного сподівання полягає в тому, що $M(X)$ наближено дорівнює середньому арифметичному спостережених значень випадкової величини.

Приклад. У сортувальному парку чотири колії. Кількість колій, зайнятих у даний момент потягами, є випадкова величина X, що розподілена за законом

X	0	1	2	3	4
P	0,05	0,20	0,35	0,25	0,15

Знайти середню кількість зайнятих колій сортувального парку.

Розв'язання. Середню кількість зайнятих колій знайдемо як математичне сподівання числа зайнятих колій. За формулою (3.5) маємо:

$$M(X) = 0 \cdot 0{,}05 + 1 \cdot 0{,}20 + 2 \cdot 0{,}35 + 3 \cdot 0{,}25 + 4 \cdot 0{,}15 = 2{,}25 .$$

Таким чином, у середньому зайнятими є 2,25 умовних колій.

Приклад. Кількість вагонів-рефрижераторів, які прибувають на станцію під навантажування, є випадкова величина X, яка набуває значення $x_1 = 5$ з імовірністю $p_1 = 1/3$ і значення x_2 з імовірністю p_2. Знайти значення x_2 і p_2, якщо відоме математичне сподівання $M(X) = 13$.

Розв'язання. Оскільки $p_1 + p_2 = 1$, то $p_2 = 1 - p_1 = 1 - \frac{1}{3} = \frac{2}{3}$. За формулою математичного сподівання (3.5)

$$M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 \quad \Leftrightarrow \quad 5 \cdot \frac{1}{3} + x_2 \cdot \frac{2}{3} = 13 .$$

Звідси $x_2 = \left(13 - \frac{5}{3}\right) \cdot \frac{3}{2} = 17$.

Відповідь. $x_2 = 17$; $p_2 = 2/3$.

Властивості математичного сподівання $M(X)$:

1) $M(C) = C, \quad C = const$;

2) $M(CX) = CM(X)$;

3) $M(X+Y) = M(X) + M(Y)$;

3') $M(X_1 + \ldots + X_k) = M(X_1) + \ldots + M(X_k)$;

4) $M(XY) = M(X) \cdot M(Y)$ для незалежних X, Y.

4') $M(X_1 \ldots X_k) = M(X_1) \cdot \ldots \cdot M(X_k)$ для попарно незалежних $X_1, \ldots, X_k$.

Приклад. ДВВ X та Y − незалежні і задані рядами розподілу:

X	-1	0	2
P	0,1	0,5	0,4

Y	3	4
P	0,7	0,3

Знайти $M(3 + 2X - Y + XY)$.

Розв'язання. По-перше, знаходимо математичні сподівання ДВВ X і Y:

$$M(X) = \sum_i x_i \cdot p_i = (-1) \cdot 0{,}1 + 0 \cdot 0{,}5 + 2 \cdot 0{,}4 = 0{,}7;$$

$$M(Y) = \sum_j y_j \cdot q_j = 3 \cdot 0{,}7 + 4 \cdot 0{,}3 = 3{,}3.$$

Тепер скористаємось властивостями математичного сподівання:

$$M(3 + 2X - Y + XY) = M(3) + 2M(X) - M(Y) + M(X) \cdot M(Y) =$$
$$= 3 + 2 \cdot 0{,}7 - 3{,}3 + 0{,}7 \cdot 3{,}3 = 3{,}41.$$

Приклад. У парк приймання сортувальної станції для переробки прибувають потяги, що спочатку обробляються двома паралельно працюючими бригадами оглядачів, а потім розформовуються на гірці. Середнє число потягів, що обробляються кожною з двох бригад протягом часу *T*, дорівнює $M(X) = 1{,}5$, а середнє число потягів, що розформовуються на гірці, дорівнює $M(Y) = 2{,}3$. Знайти середню загальну кількість потягів у парку $M(2X + Y)$.

Розв'язання. $M(2X + Y) = 2M(X) + M(Y) = 2 \cdot 1{,}5 + 2{,}3 = 5{,}3$.

3.5.2. Дисперсія та середнє квадратичне відхилення. Коефіцієнт варіації

Очевидно, що існує багато випадкових величин з однаковими математичними сподіваннями. Таким чином, важливим стає питання, як далеко від математичного сподівання знаходяться реальні значення ДВВ. Для контролю *розсіювання* значень вводять нижченаведені характеристики.

Відхиленням випадкової величини X від її математичного сподівання називається випадкова величина $X - M(X)$. Випадкова величина $X_0 = X - M(X)$ називається *центрованою*, оскільки її математичне сподівання дорівнює нулю.

Дисперсією $D(X)$ випадкової величини X називається математичне сподівання квадрата відхилення X від її математичного сподівання:

$$\underline{D(X) = M\left[X - M(X)\right]^2}. \tag{3.6}$$

Імовірнісний зміст дисперсії полягає в тому, що $D(X)$ характеризує розсіювання значень випадкової величини навколо її математичного сподівання $M(X)$.

Якщо скористатися властивостями математичного сподівання, нескладно вивести формулу, що іноді спрощує обчислення дисперсії:

$$D(X) = M\left[X - M(X)\right]^2 = M\left(X^2 - 2X \cdot M(X) + M^2(X)\right) =$$
$$= M\left(X^2\right) - 2M(X) \cdot M(X) + M^2(X) = M\left(X^2\right) - 2M^2(X) + M^2(X) = M\left(X^2\right) - M^2(X).$$

Таким чином,

$$\underline{D(X) = M\left(X^2\right) - M^2(X)}. \tag{3.7}$$

Приклад. Розподіл числа вагонів у відчепах на сортувальній станції має вигляд:

Число вагонів у відчепах	1	2	3	4	5
Ймовірність	0,6	0,15	0,11	0,09	0,05

Знайти математичне сподівання і дисперсію числа X вагонів у відчепах.

Розв’язання. За формулою (3.5) знаходимо математичне сподівання:

$$M(X)=\sum_i x_i\cdot p_i=1\cdot 0{,}6+2\cdot 0{,}15+3\cdot 0{,}11+4\cdot 0{,}09+5\cdot 0{,}05=1{,}84.$$

Обчислимо дисперсію двома способами. За означенням (3.6)

$$D(X)=\sum_i [x_i-M(X)]^2\cdot p_i=(1-1{,}84)^2\cdot 0{,}6+(2-1{,}84)^2\cdot 0{,}15+$$

$$+(3-1{,}84)^2\cdot 0{,}11+(4-1{,}84)^2\cdot 0{,}09+(5-1{,}84)^2\cdot 0{,}05=1{,}4944\,.$$

Щоб скористатись формулою (3.7), знайдемо розподіл ДВВ X^2 (п.3.4):

X^2	1	4	9	16	25
P	0,6	0,15	0,11	0,09	0,05

і математичне сподівання

$$M(X^2)=1\cdot 0{,}6+4\cdot 0{,}15+9\cdot 0{,}11+16\cdot 0{,}09+25\cdot 0{,}05=4{,}88.$$

Звідси, $D(X)=M\left(X^2\right)-M^2(X)=4{,}88-(1{,}84)^2=1{,}4944\,.$

З означення дисперсії і вже відомих властивостей математичного сподівання випливають властивості дисперсії.

Властивості дисперсії $D(X)$:

1) $D(X)\ge 0$;
2) $D(C)=0$, $\quad C=const$;
3) $D(CX)=C^2D(X)$;
4) $D(X+Y)=D(X-Y)=D(X)+D(Y)$ для незалежних X,Y.
 4') $D(C+X)=D(X)$
 4'') $D(X_1+\ldots+X_k)=D(X_1)+\ldots+D(X_k)$ для попарно незалежних $X_1,\ldots,X_k$.

Зауваження. Дисперсія має розмірність, що дорівнює квадрату розмірності випадкової величини. Інколи на практиці виявляється зручнішим мати оцінку розсіювання в тих же одиницях, що і сама випадкова величина.

Середнім квадратичним відхиленням називається величина

$$\sigma(X) = \sqrt{D(X)}, \tag{3.8}$$

що як і дисперсія характеризує розсіювання, але вимірюється в тих же одиницях, що і сама випадкова величина.

Коефіцієнтом варіації V *випадкової величини* X називається величина

$$V = \frac{\sigma(X)}{M(X)}, \tag{3.9}$$

що характеризує відносне відхилення від середнього значення. Ця додаткова характеристика зазвичай використовується для порівняння розсіювань випадкових величин, що мають різні одиниці вимірювання.

Приклад. Задані дисперсії двох незалежних випадкових величин: $D(X)=1$, $D(Y)=2{,}5$. Знайти $D(5X+Y-4)$, $\sigma(X-Y)$.

Розв'язання. Скористаємось властивостями дисперсії:

$$D(5X+Y-4)=5^2\cdot D(X)+1^2\cdot D(Y)+D(-4)=25\cdot 1+2{,}5+0=27{,}5;$$

$$\sigma(X-Y)=\sqrt{D(X-Y)}=\sqrt{D(X)+D(Y)}=\sqrt{1+2{,}5}\approx 1{,}87.$$

Завдання

1. Знайти числові характеристики ДВВ X, заданої законом розподілу

X	-1	0	1	3
P	0,4	0,3	0,2	0,1

Відповідь. $M(X)=0{,}1$, $D(X)=1{,}49$, $\sigma(X)\approx 1{,}22$, $V(X)\approx 12{,}2$.

2. На одному із складів залізниці добове надходження матеріалу у вагонах є випадкова величина X, яка набуває значення $x_1=0$ з імовірністю $p_1=1/6$, $x_2=1$ з імовірністю $p_2=1/2$ і значення x_3 з імовірністю p_3. Знайти x_3 і p_3, якщо відоме математичне сподівання $M(X)=1{,}5$.

Відповідь. $x_3=3$; $p_3=1/3$.

3. За статистикою служби автодорожніх пригод, в робочі дні на даній ділянці дороги в інтервалі часу від 17 до 19 годин можуть відбутися 0, 1, 2 або 3 автомобільні катастрофи з імовірностями 0,89; 0,07; 0,035 та 0,005 відповідно. Знайти середню кількість катастроф у даний період часу.

Відповідь. $M(X) = 0{,}155$.

4. Кількість вагонів на дане призначення в кожному поїзді, що прибуває, є випадкова величина X, яка розподілена за законом:

X	0	1	2	3	4
P	0,05	0,15	0,35	0,3	0,15

Знайти середню кількість вагонів $M(X)$, дисперсію $D(X)$.

Відповідь. $M(X)$=2,35; $D(X)$=1,1275.

5. На станції завантажуються чотири вагони. Кількість навантажених вагонів, відправлених із станції, є випадкова величина X, яка має такий розподіл:

X	0	1	2	3	4
P	0,0625	0,250	0,375	0,250	0,0625

Знайти числові характеристики X.

Відповідь. $M(X)$=2; $D(X)$=1; $\sigma(X)$=1; $V(X) = 2$.

6. Кількість чотиривісних піввагонів, які можуть зустрітися в групі з шести вагонів, є випадкова величина X, яка має такий розподіл:

X	0	1	2	3	4	5	6
P	0,04	0,2	0,35	0,2	0,15	0,04	0,02

Знайти математичне сподівання і середнє квадратичне відхилення ДВВ X.

Відповідь. $M(X)$=2,42; $\sigma(X) \approx 1{,}299$.

7. Розподіл кількості вагонів у групах, що прибувають на дану колію сортувального парку, має вигляд:

Величина групи в умовних вагонах	5	8	10	15	20
Ймовірність	0,4	0,25	0,1	0,2	0,05

Визначити середню величину групи та дисперсію.

Відповідь. $M(X)=9;\quad D(X)=20$.

8. Число вагонів призначенням на вантажний двір у поїздах, що прибувають з першого напрямку, є випадковою величиною X, розподіленою за законом

X	0	1	2	3	4	5
P	0,03	0,15	0,35	0,3	0,15	0,02

а з другого напрямку − випадковою величиною Y, яка розподілена за законом

Y	0	1	2	3
P	0,25	0,35	0,3	0,1

Знайти середню кількість $M(X+Y)$ вагонів призначенням на вантажний двір.

Відповідь. $M(X+Y)=3{,}7$.

9. Число піввагонів, відправлених із станції протягом часу T, є випадковою величиною X, математичне сподівання якої $M(X)=3$, а дисперсія − $D(X)=0{,}9$. Число відправлених платформ також є випадковою величиною Y, яка має розподіл

Y	0	1	2	3
P	1/9	2/9	2/9	4/9

Знайти середню загальну кількість відправлених піввагонів і платформ $M(X+Y)$ та її дисперсію $D(X+Y)$. (Вважається, що випадкові величини X і Y є незалежними.)

Відповідь. $M(X+Y)=5;\ D(X+Y)\approx 2$.

10. У транзитному парку сортувальної станції є дві колії. На першу колію прибувають поїзди з двох напрямків, а на другу − з трьох напрямків. Кількість поїздів, що прибувають з кожного напрямку на першу колію, є випадковою величиною X,

дисперсія якої відома: $D(X)=2$. Кількість поїздів, які прибувають з кожного напрямку на другу колію, є випадковою величиною Y з дисперсією $D(Y)=1$. Визначити середнє квадратичне відхилення $\sigma(2X+3Y)$ загальної кількості поїздів, що прибувають до парку. (Прибуття поїздів з різних напрямків вважати незалежними.)

Відповідь. $\sigma(2X+3Y)=\sqrt{17}\approx 4{,}123$.

3.6. Неперервні випадкові величини

Ми переходимо до вивчення неперервних випадкових величин. Нагадаємо, що на відміну від дискретної, значення неперервної випадкової величини заповнюють деякий, скінченний або нескінченний, проміжок числової осі, а тому не можуть бути перелічені.

Як було зазначено вище (див. п. 3.3), *функція розподілу НВВ* X визначається так саме, як і для дискретної випадкової величини, а саме формулою

$$F(x)=P(X<x), \qquad x\in R ,$$

але має ряд особливих властивостей.

Неперервною називається випадкова величина X, функція розподілу якої $F(x)$ є неперервною і має неперервну похідну всюди, крім скінченого числа точок.

Щільністю розподілу ймовірностей НВВ X називається похідна від функції розподілу НВВ X, тобто функція

$$f(x)=F'(x), \qquad x\in R. \tag{3.10}$$

З означення щільності розподілу $f(x)$ випливає формула відшукання функції розподілу за її щільністю:

$$F(x)=\int_{-\infty}^{x} f(z)\,dz, \qquad x\in R. \tag{3.11}$$

Функція розподілу ймовірностей $F(x)$ також називається *інтегральною функцією НВВ* X, а щільність розподілу $f(x)$ – *диференціальною*.

Розподіл неперервної випадкової величини задається $F(x)$ або $f(x)$, так як зв'язок між ними однозначний.

Графік щільності розподілу $f(x)$ називається *кривою розподілу.*

Зауваження 1. Щільність розподілу $f(x)$ існує лише для неперервних випадкових величин.

Зауваження 2. Графік функції розподілу $F(x)$ НВВ X завжди є неперервною кривою, в той час як графік щільності розподілу $f(x)$ може бути розривним.

Приклад. НВВ X задана своєю функцією розподілу

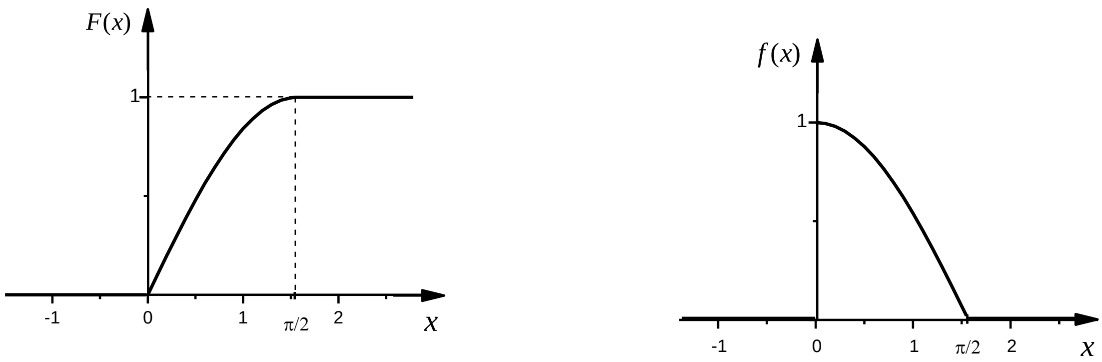
$$F(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ \sin x & \text{при} \quad 0 < x \le \pi/2, \\ 1 & \text{при} \quad x > \pi/2. \end{cases}$$

Знайти функцію щільності розподілу і побудувати графіки $F(x)$ і $f(x)$.

Розв'язання. Згідно з означенням (3.10) щільності розподілу:

$$f(x) = F'(x) = \begin{cases} 0 & \text{при} \ x \le 0, \\ \cos x & \text{при} \ 0 < x \le \pi/2, \\ 0 & \text{при} \ x > \pi/2. \end{cases} = \begin{cases} \cos x & \text{при} \quad x \in (0; \pi/2], \\ 0 & \text{при} \quad x \notin (0; \pi/2]. \end{cases}$$

Функція щільності розподілу має одну точку розриву. Графіки $F(x)$ і $f(x)$ зображено на рис. 3.5.



а) б)

Рис. 3.5

Приклад. НВВ X задана своєю щільністю розподілу

$$f(x)=\begin{cases}\frac{3}{4}(1-x^2) & \text{при} \quad x\in(-1;1], \\ 0 & \text{при} \quad x\notin(-1;1].\end{cases}$$

Знайти функцію розподілу і побудувати графіки $F(x)$ і $f(x)$.

<u>Розв'язання.</u> Згідно з (3.11), щоб знайти $F(x)$ потрібно зінтеґрувати щільність

$$f(x)=\begin{cases}\frac{3}{4}(1-x^2) & \text{при } x\in(-1;1], \\ 0 & \text{при } x\notin(-1;1].\end{cases} = \begin{cases}0 & \text{при} \quad x\le -1, \\ \frac{3}{4}(1-x^2) & \text{при} \quad -1<x\le 1, \\ 0 & \text{при} \quad x>1.\end{cases}$$

Отримаємо:

при $x\le -1$, $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{x} 0dz=0$;

при $-1<x\le 1$,

$$F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{-1} 0dz+\frac{3}{4}\int\limits_{-1}^{x}(1-z^2)dz=0+\frac{3}{4}\cdot\left(z-\frac{z^3}{3}\right)\Bigg|_{-1}^{x}=\frac{3}{4}x-\frac{x^3}{4}+\frac{1}{2};$$

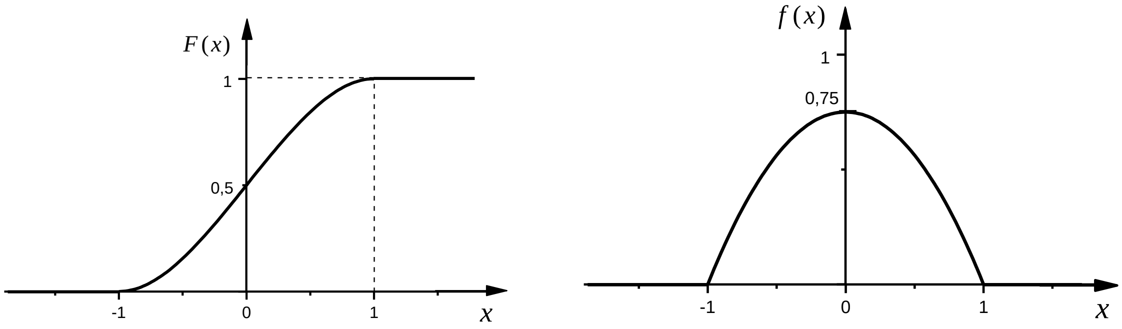
при $x>1$, $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\frac{3}{4}\int\limits_{-1}^{1}(1-z^2)dz=\frac{3}{4}\cdot\left(z-\frac{z^3}{3}\right)\Bigg|_{-1}^{1}=1.$

Остаточно функція розподілу має вигляд:

$$F(x)=\begin{cases}0 & \text{при} \quad x\le -1, \\ \frac{1}{4}\left(3x-x^3+2\right) & \text{при} \quad -1<x\le 1, \\ 1 & \text{при} \quad x>1.\end{cases}$$

Зауважимо, що перший і третій рядки можна було заповнити автоматично, маючи на увазі властивості функції розподілу.

У даному випадку обидві функції $F(x)$ і $f(x)$ є неперервними. Графіки $F(x)$ і $f(x)$ зображено на рис. 3.6.



а) б)

Рис. 3.6

В п. 3.3 вже обговорювались загальні властивості функції розподілу $F(x)$. Ми поповнимо цей перелік характерними для неперервних випадкових величин властивостями $F(x)$ і наведемо властивості $f(x)$.

Властивості функції розподілу $F(x)$ для НВВ:

1) $0 \le F(x) \le 1$;

2) якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$;

3) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1$;

4) імовірність попадання випадкової величини X в півінтервал $[\alpha;\beta)$ дорівнює різниці значень функції розподілу в цих точках:

$$P(\alpha \le X < \beta) = F(\beta) - F(\alpha);$$

5) імовірність того, що НВВ X набуде деяке фіксоване значення x_0 дорівнює нулю: $P(X = x_0) = 0$;

6) імовірності попадання *НВВ* X у проміжки $(\alpha;\beta), (\alpha;\beta], [\alpha;\beta)$ та $[\alpha;\beta]$ є рівними:

$$\underline{P(\alpha < X < \beta) = P(\alpha \le X < \beta) = P(\alpha \le X \le \beta) = P(\alpha < X \le \beta)} \ . \qquad (3.12)$$

Властивості функції щільності розподілу $f(x)$:

1) $f(x) \ge 0$;

2) $\underline{\int\limits_{-\infty}^{+\infty} f(x)dx = 1}$; (3.13)

3) $P(\alpha < X < \beta) = \int\limits_{\alpha}^{\beta} f(x)dx.$ (3.14)

Зауваження. З геометричної точки зору властивість (3.13) означає, що площа фігури, розташованої під графіком диференціальної функції, дорівнює одиниці.

Приклад. НВВ X задана своєю щільністю розподілу

$$f(x) = \begin{cases} ax^2 & \text{при} \quad x \in (0;2], \\ 0 & \text{при} \quad x \notin (0;2]. \end{cases}$$

Визначити параметр a та ймовірності попадання X в сегмент $[0;1]$, інтервал $(1;3)$, півінтервал $(-1;1,5]$.

Розв'язання. Для знаходження параметра скористаємось властивістю (3.15) щільності розподілу:

$$\int\limits_{-\infty}^{+\infty} f(x)dx = \int\limits_{0}^{2} ax^2 dx = a \cdot \left. \frac{x^3}{3} \right|_0^2 = a \cdot \frac{2^3}{3} = \frac{8a}{3} = 1 \qquad \Rightarrow \qquad a = \frac{3}{8}.$$

Тоді функція щільності розподілу має вигляд

$$f(x) = \begin{cases} \dfrac{3}{8}x^2 & \text{при} \quad x \in (0;2], \\ 0 & \text{при} \quad x \notin (0;2], \end{cases}$$

її графік зображено на рис. 3.7.

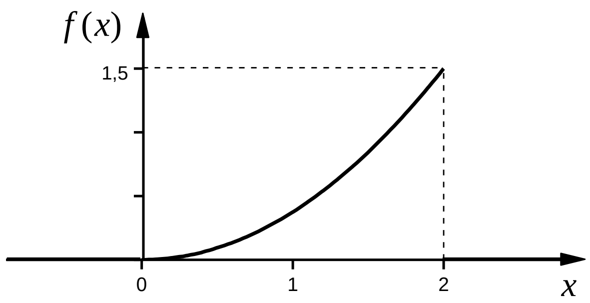


Рис. 3.7

Для обчислення ймовірностей скористаємось формулами (3.12) і (3.14):

$$P(0 \le X \le 1) = \int_0^1 f(x)dx = \int_0^1 \frac{3}{8} \cdot x^2 dx = \frac{3}{8} \cdot \frac{x^3}{3}\Bigg|_0^1 = \frac{1}{8};$$

$$P(1 < X < 3) = \int_1^3 f(x)dx = \int_1^2 \frac{3}{8} \cdot x^2 dx + \int_2^3 0 \cdot dx = \frac{x^3}{8}\Bigg|_1^2 = 1 - \frac{1}{8} = \frac{7}{8};$$

$$P(-1 < X \le 1{,}5) = \int_{-1}^{1{,}5} f(x)dx = \int_{-1}^{0} 0 \cdot dx + \int_0^{1{,}5} \frac{3}{8} \cdot x^2 dx = \frac{x^3}{8}\Bigg|_0^{1{,}5} = \frac{1}{8} \cdot \left(\frac{3}{2}\right)^3 = \frac{27}{64}.$$

Приклад. Графік диференціальної функції неперервної випадкової величини X зображено на рис. 3.8.

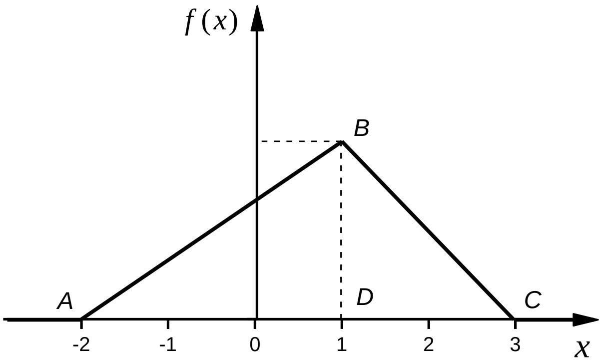


Рис. 3.8

Записати функцію щільності розподілу формулою, знайти функцію розподілу та обчислити ймовірності $P(-1 < X < 1)$, $P(0 \le X \le 2)$.

Розв'язання. За властивістю (3.13) диференціальної функції площа трикутника ABC дорівнює одиниці. З цієї умови знайдемо невідому ординату вершини $B(1; y)$:

$$S_{\Delta ABC} = \frac{1}{2} \cdot AC \cdot BD = \frac{1}{2} \cdot 5 \cdot y = 1 \qquad \Rightarrow \; y = \frac{2}{5} \qquad \Rightarrow \; B\left(1; \frac{2}{5}\right).$$

Диференціальна функція $f(x)$ є лінійною як на проміжку $[-2;1]$**,** так і на проміжку $[1;3]$, тобто:

$$f(x) = k_1 x + b_1 \quad \text{при} \quad x \in [-2;1];$$
$$f(x) = k_2 x + b_2 \quad \text{при} \quad x \in [1;3].$$

Пряма $y = k_1 x + b_1$ проходить через точки $A(-2;0)$ і $B\left(1;\frac{2}{5}\right)$:

$$\frac{y-0}{x+2} = \frac{\frac{2}{5}-0}{1+2} \quad \Rightarrow \quad y = \frac{2}{15}(x+2).$$

Пряма $y = k_2 x + b_2$ проходить через точки $C(3;0)$ і $B\left(1;\frac{2}{5}\right)$:

$$\frac{y-0}{x-3} = \frac{\frac{2}{5}-0}{1-3} \quad \Rightarrow \quad y = -\frac{1}{5}(x-3).$$

Таким чином, щільність розподілу НВВ X має вигляд:

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le -2, \\ \frac{2}{15}(x+2) & \text{при} \quad -2 < x \le 1, \\ \frac{1}{5}(3-x) & \text{при} \quad 1 < x \le 3, \\ 0 & \text{при} \quad x > 3. \end{cases}$$

За формулою (3.11) знайдемо функцію розподілу:

при $x \in (-2;1]$:

$$F(x) = \int_{-2}^{x} \frac{2}{15}(z+2)\,dz = \frac{2}{15}\int_{-2}^{x}(z+2)\,dz = \frac{2}{15}\cdot\frac{(z+2)^2}{2}\Bigg|_{-2}^{x} = \frac{(x+2)^2}{15};$$

при $x \in (1;3]$:

$$F(x) = \int_{-2}^{1}\frac{2}{15}(z+2)dz + \int_{1}^{x}\frac{1}{5}(3-z)\,dz = \frac{3}{5} + \frac{1}{5}\cdot\int_{1}^{x}(3-z)\,dz =$$

$$= \frac{3}{5} + \frac{1}{5}\cdot\left(-\frac{(3-z)^2}{2}\right)\Bigg|_{1}^{x} = 1 - \frac{(x-3)^2}{10}.$$

Остаточно інтегральна функція має вигляд:

$$F(x)=\begin{cases}0 & \text{при} \quad x\le -2,\\ \dfrac{1}{15}(x+2)^2 & \text{при} \quad -2<x\le 1,\\ 1-\dfrac{1}{10}(x-3)^2 & \text{при} \quad 1<x\le 3,\\ 1 & \text{при} \quad x>3,\end{cases}$$

її графік зображено на рис. 3.9.

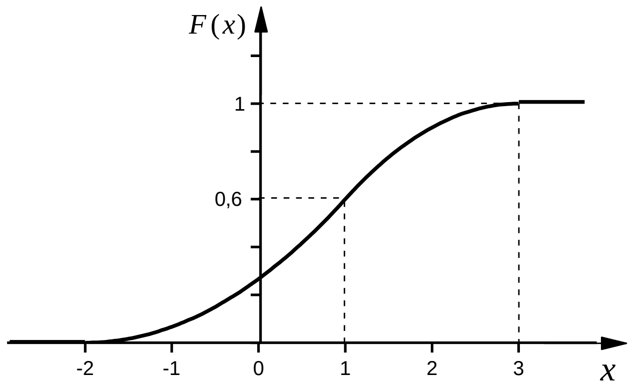


Рис. 3.9

За допомогою функції розподілу обчислюємо ймовірності:

$$P(-1<X<1)=F(1)-F(-1)=\frac{3}{5}-\frac{1}{15}(-1+2)^2=\frac{8}{15},$$

$$P(0\le X\le 2)=F(2)-F(0)=1-\frac{1}{10}(2-3)^2-\frac{1}{15}(0+2)^2=\frac{19}{30}.$$

Завдання

1. Час очікування, хв, пасажиром електропоїзда метрополітену на деякій станції є неперервною випадковою величиною X, щільність розподілу якої дорівнює

$$f(x)=\begin{cases}C & \text{при} \quad x\in(0;5],\\ 0 & \text{при} \quad x\notin(0;5].\end{cases}$$

Знайти невідомий параметр C, функцію розподілу та ймовірність того, що час очікування пасажиром електропоїзда не перевищить 3 хв.

Відповідь. $C=\dfrac{1}{5}$; $F(x)=\begin{cases}0 & \text{при } x\le 0,\\ \dfrac{x}{5} & \text{при } 0<x\le 5,\\ 1 & \text{при } x>5;\end{cases}$ $P(X\le 3)=0{,}6$.

2. Інтервал часу, год, між моментами прибуття потягів на сортувальну станцію є неперервною випадковою величиною T, диференціальна функція якої має вигляд $f(t)=\begin{cases}ae^{-4t} & \text{при } t\ge 0,\\ 0 & \text{при } t<0.\end{cases}$ Знайти невідомий параметр a, функцію розподілу та ймовірність того, що інтервал часу між прибуттям двох потягів буде не більше півгодини.

Відповідь. $a=4$; $F(t)=\begin{cases}1-e^{-4t} & \text{при } t\ge 0,\\ 0 & \text{при } t<0;\end{cases}$ $P(T\le 0{,}5)\approx 0{,}865$.

3. Час T, год, роботи маневрових локомотивів на позиції 3 контролера машиніста є неперервною випадковою величиною, функція розподілу якої має вигляд:

$$F(t)=\begin{cases}1-\left(28{,}125t^2+7{,}5t+1\right)\cdot e^{-7{,}5\cdot t} & \text{при } t\ge 0,\\ 0 & \text{при } t<0.\end{cases}$$

Знайти щільність розподілу НВВ T та обчислити ймовірність того, що час роботи на позиції 3 складатиме від 15 до 30 хв.

Відповідь. $f(t)=\begin{cases}210{,}9375\cdot t^2\cdot e^{-7{,}5\cdot t} & \text{при } t\ge 0,\\ 0 & \text{при } t<0.\end{cases}$

$P(0{,}25\le T\le 0{,}5)\approx 0{,}434$.

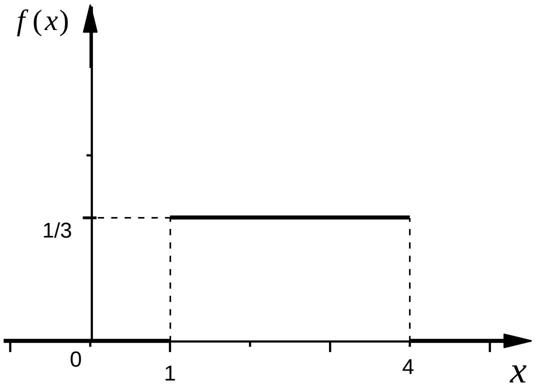
4. Інтегральна функція НВВ X має вигляд

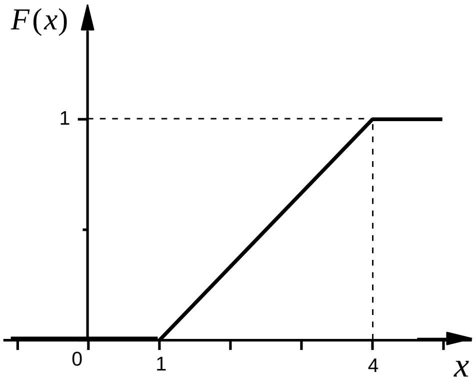
$$F(x)=\begin{cases}0 & \text{при } x\le 1,\\ a(x-1) & \text{при } 1<x\le 4,\\ 1 & \text{при } x>4.\end{cases}$$

Визначити щільність розподілу та побудувати графіки $f(x)$ і $F(x)$. Знайти ймовірності попадання НВВ X у проміжки $(0;2]$, $(2;3)$, $[3;7]$.

Відповідь. $a=\frac{1}{3}$. $f(x)=\begin{cases}1/3 & \text{при } x\in(1;4],\\ 0 & \text{при } x\notin(1;4].\end{cases}$

$P(0<X\le 2)=P(2<X<3)=P(3\le X\le 7)=\frac{1}{3}$.



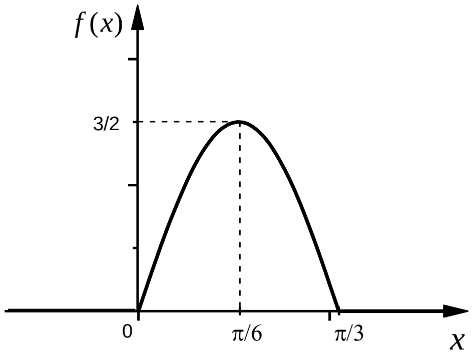


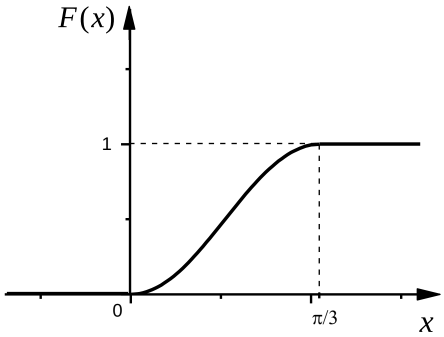
5. Дана функція $f(x)=\begin{cases}a\sin 3x & \text{при } x\in(0;\pi/3],\\ 0 & \text{при } x\notin(0;\pi/3].\end{cases}$

При якому значенні параметра a функція $f(x)$ може служити щільністю розподілу? Знайти інтегральну функцію, побудувати графіки $f(x)$ і $F(x)$. Обчислити ймовірності попадання НВВ X у проміжки $\left(0;\frac{\pi}{6}\right)$, $\left[\frac{\pi}{6};\frac{\pi}{4}\right]$.

Відповідь. $a=1{,}5$. $F(x)=\begin{cases}0 & \text{при } x<0,\\ \frac{1-\cos 3x}{2} & \text{при } 0<x\le\pi/3,\\ 1 & \text{при } x>\pi/3.\end{cases}$

$P\left(0<X<\frac{\pi}{6}\right)=\frac{1}{2}$; $P\left(\frac{\pi}{6}\le X\le\frac{\pi}{4}\right)=\frac{\sqrt{2}}{4}$.





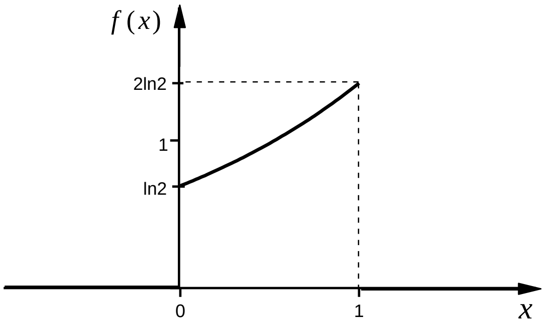
6. НВВ X задана своєю функцією розподілу

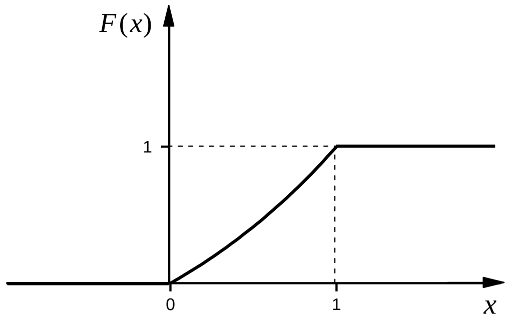
$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ 2^x-1 & \text{при} \quad 0<x\le 1,\\ 1 & \text{при} \quad x>1.\end{cases}$$

Знайти диференціальну функцію, побудувати графіки $f(x)$ і $F(x)$. Обчислити ймовірності $P(-0{,}5<X<0{,}5)$, $P(0{,}5\le X\le 1)$.

<u>Відповідь.</u> $f(x)=\begin{cases}2^x\ln 2 & \text{при} \quad x\in(0;1],\\ 0 & \text{при} \quad x\notin(0;1].\end{cases}$

$P(-0{,}5<X<0{,}5)=\sqrt{2}-1$; $P(0{,}5\le X\le 1)=2-\sqrt{2}$.





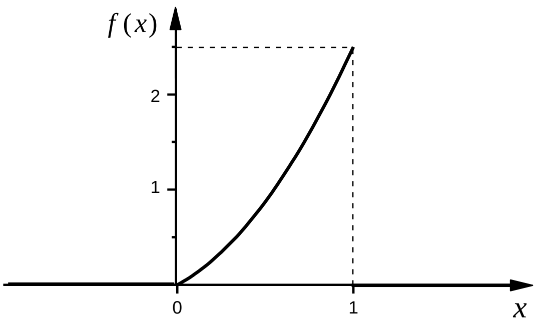
7. Щільність розподілу випадкової величини X має вигляд

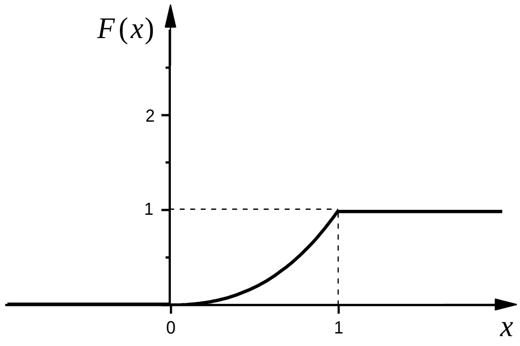
$$f(x)=\begin{cases}kx^2+x & \text{при} \quad x\in(0;1],\\ 0 & \text{при} \quad x\notin(0;1].\end{cases}$$

Знайти невідомий параметр, інтегральну функцію та $P(0\le X\le 0{,}5)$. Побудувати графіки функцій $f(x)$ і $F(x)$.

<u>Відповідь.</u> $k=1{,}5$. $F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ \dfrac{x^2+x^3}{2} & \text{при} \quad 0<x\le 1,\\ 1 & \text{при} \quad x>1.\end{cases}$

$P(X\le 0{,}5)=0{,}1875$.





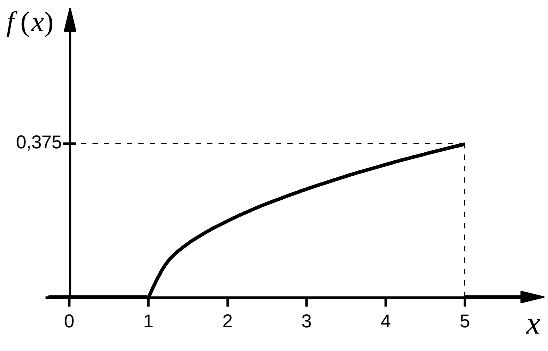
8. Щільність розподілу має вигляд

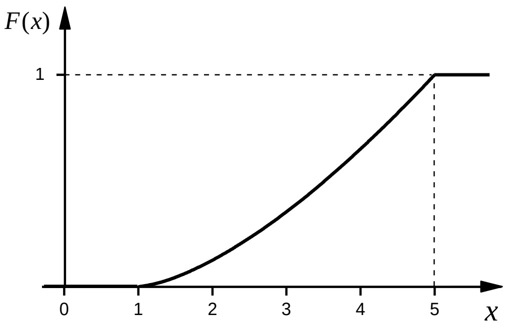
$$f(x)=\begin{cases} a\sqrt{x-1} & \text{при} \quad x\in(1;5], \\ 0 & \text{при} \quad x\notin(1;5]. \end{cases}$$

Знайти невідомий параметр, інтегральну функцію та побудувати графіки $f(x)$ і $F(x)$. Обчислити ймовірність $P(1<X<2)$.

<u>Відповідь.</u> $a=\dfrac{3}{16}$. $F(x)=\begin{cases} 0 & \text{при } x\le 1, \\ \dfrac{1}{8}\sqrt{(x-1)^3} & \text{при } 1<x\le 5, \\ 1 & \text{при } x>5. \end{cases}$

$P(1<X<2)=\dfrac{1}{8}$.





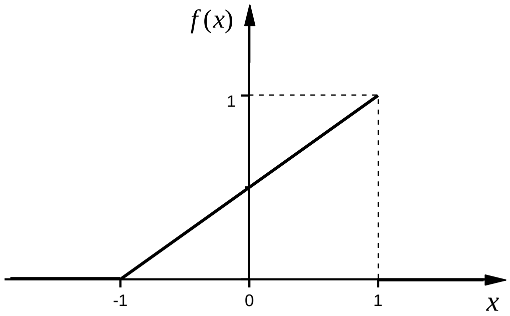
9. Інтегральна функція НВВ X має

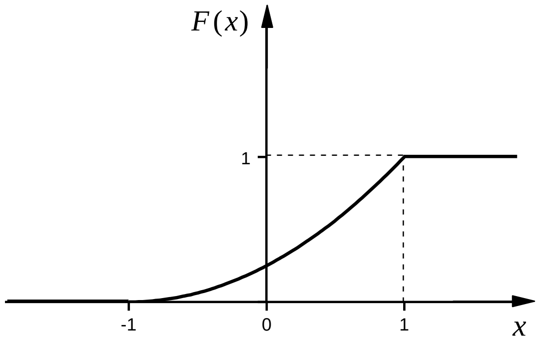
$$F(x)=\begin{cases} 0 & \text{при} \quad x\le -1, \\ k(x+1)^2 & \text{при} \ -1<x\le 1, \\ 1 & \text{при} \quad x>1. \end{cases}$$

Знайти диференціальну функцію, побудувати графіки $f(x)$ і $F(x)$. Обчислити ймовірності $P(-1 < X \le 0), P(0 \le X < 1)$.

Відповідь. $k = \frac{1}{4}$. $f(x) = \begin{cases} \frac{x+1}{2} & \text{при } x \in (-1;1], \\ 0 & \text{при } x \notin (-1;1]. \end{cases}$

$P(-1 < X \le 0) = 0{,}25; \quad P(0 \le X < 1) = 0{,}75$.





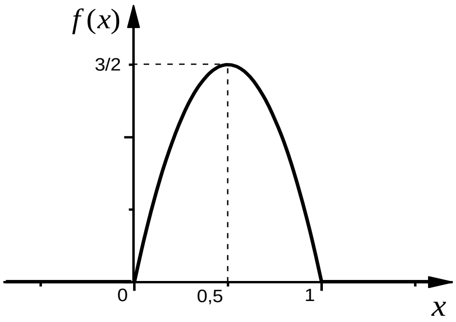
10. Задано диференціальну функцію НВВ X

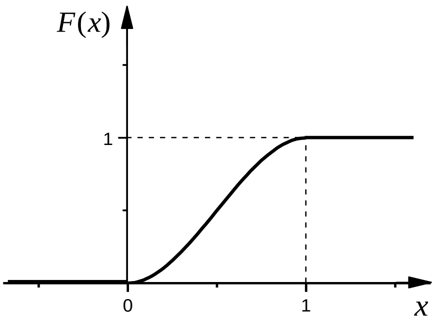
$$f(x) = \begin{cases} a\left(x - x^2\right) & \text{при} \quad x \in (0;1], \\ 0 & \text{при} \quad x \notin (0;1]. \end{cases}$$

Знайти функцію розподілу та побудувати графіки $f(x)$ і $F(x)$. Обчислити ймовірності попадання X у проміжки $\left(-\frac{1}{2};\frac{1}{2}\right), \left[\frac{1}{2};\frac{3}{2}\right), \left[\frac{1}{4};\frac{3}{4}\right]$.

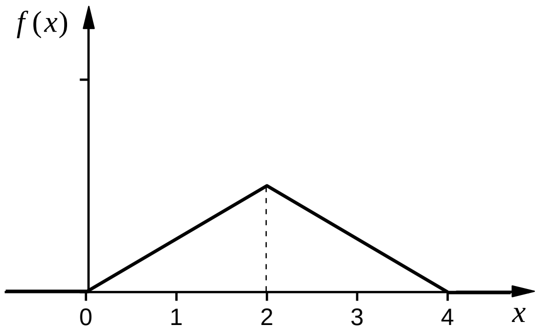
Відповідь. $a = 6$. $F(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ 3x^2 - 2x^3 & \text{при} \quad 0 < x \le 1, \\ 1, & \text{при} \quad x > 1. \end{cases}$

$P\left(-\frac{1}{2} < X < \frac{1}{2}\right) = P\left(\frac{1}{2} \le X < \frac{3}{2}\right) = \frac{1}{2}; \quad P\left(\frac{1}{4} \le X \le \frac{3}{4}\right) = \frac{11}{64}$.





11. Графік диференціальної функції НВВ X має вигляд:

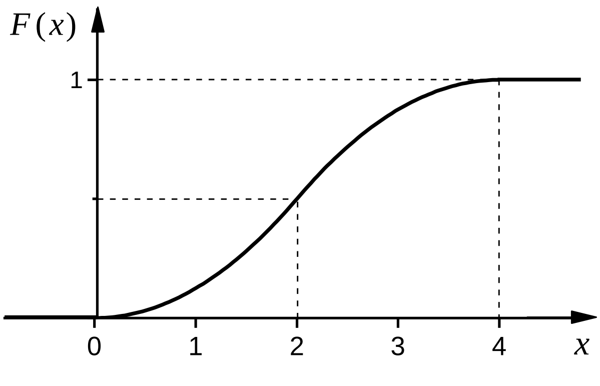


Знайти аналітичні вирази функцій $f(x)$ і $F(x)$, побудувати графік інтегральної функції. Обчислити ймовірності $P(0 < X < 1)$, $P(2 \le X < 4)$.

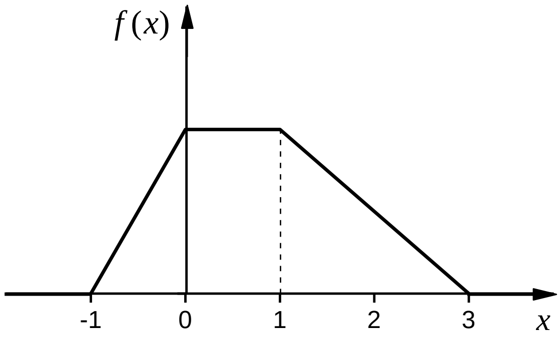
<u>Відповідь.</u> $$f(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ \frac{1}{4}x & \text{при} \quad 0 < x \le 2, \\ \frac{1}{4}(4-x) & \text{при} \quad 2 < x \le 4, \\ 0 & \text{при} \quad x > 4. \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ \frac{x^2}{8} & \text{при} \quad 0 < x \le 2, \\ 1 - \frac{1}{8}(x-4)^2 & \text{при} \quad 2 < x \le 4, \\ 1 & \text{при} \quad x > 4. \end{cases}$$

$P(0 < X < 1) = 0{,}125$; $P(2 \le X < 4) = 0{,}5$.



12. НВВ X задана графіком щільності розподілу:

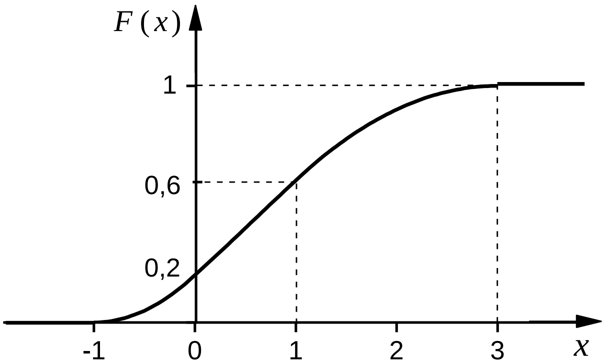


Знайти функцію розподілу, побудувати її графік. Обчислити ймовірності $P(0 < X < 1)$, $P(2 \le X \le 4)$.

Відповідь. $$f(x) = \begin{cases} 0 & \text{при} \quad x \le -1, \\ 0{,}4 \cdot (x+1) & \text{при} \quad -1 < x \le 0, \\ 0{,}4 & \text{при} \quad 0 < x \le 1, \\ 0{,}2 \cdot (3-x) & \text{при} \quad 1 < x \le 3, \\ 0 & \text{при} \quad x > 3. \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{при} \quad x \le -1, \\ 0{,}2 \cdot (x+1)^2 & \text{при} \quad -1 < x \le 0, \\ 0{,}2 + 0{,}4 \cdot x & \text{при} \quad 0 < x \le 1, \\ 1 - 0{,}1 \cdot (x-3)^2 & \text{при} \quad 1 < x \le 3, \\ 1 & \text{при} \quad x > 3. \end{cases}$$

$$P(0 < X < 1) = 0{,}4; \quad P(2 \le X \le 4) = 0{,}1.$$



3.7. Числові характеристики неперервної випадкової величини

Поняття математичного сподівання, дисперсії і середнього квадратичного відхилення та коефіцієнта варіації визначаються і мають зміст такий же, як і для дискретних випадкових величин. Але, звичайно, у зв'язку з нескінченною множиною можливих значень НВВ, формули обчислення числових характеристик є інтегральними.

Математичне сподівання $M(X)$ *НВВ* X визначається формулою

$$M(X)=\int_{-\infty}^{+\infty} x\,f(x)\,dx. \tag{3.15}$$

Дисперсія $D(X)$ *НВВ* X обчислюється за формулою

$$D(X)=\int_{-\infty}^{+\infty} x^2 f(x)dx-\left(M(X)\right)^2. \tag{3.16}$$

Незмінними залишаються формули для *середнього квадратичного відхилення*

$$\sigma(X)=\sqrt{D(X)} \tag{3.17}$$

і *коефіцієнта варіації*

$$V=\frac{\sigma(X)}{M(X)}\ . \tag{3.28}$$

Приклад. Функція розподілу випадкової величини X надходження порожніх вагонів до пункту завантаження має вигляд:

$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ ax & \text{при} \quad 0<x\le 1,\\ 1 & \text{при} \quad x>1.\end{cases}$$

Знайти щільність розподілу ймовірностей і числові характеристики X.

Розв’язання. Згідно з означенням (3.10), щільність заданого розподілу ймовірностей, має вигляд:

$$f(x) = F'(x) = \begin{cases} a & \text{при} \quad 0 < x \le 1, \\ 0 & \text{при} \quad x \le 0 \quad \text{і} \quad x > 1. \end{cases}$$

За властивістю (3.13) щільності розподілу знаходимо невідомий параметр:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^1 adx = ax\Big|_0^1 = a = 1.$$

Таким чином,

$$f(x) = \begin{cases} 1 & \text{при} \quad 0 < x \le 1, \\ 0 & \text{при} \quad x \le 0 \quad \text{і} \quad x > 1. \end{cases}$$

За формулами (3.15)−(3.18) знаходимо математичне сподівання, дисперсію, середнє квадратичне відхилення та коефіцієнт варіації:

$$M(X) = \int_0^1 xdx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}; \qquad D(X) = \int_0^1 x^2dx - \left(\frac{1}{2}\right)^2 = \frac{x^3}{3}\Big|_0^1 - \frac{1}{4} = \frac{1}{12};$$

$$\sigma(X) = \sqrt{D(X)} = \frac{1}{2\sqrt{3}}; \qquad V = \frac{\sigma(X)}{M(X)} = \frac{2}{2\sqrt{3}} = \frac{1}{\sqrt{3}}.$$

Модою НВВ X називається таке її значення M_0, при якому щільність розподілу набуває максимального значення:

$$\underline{f(M_0) = f_{\max}}. \tag{3.19}$$

Медіаною НВВ X називається таке її значення Me, для якого з однаковою ймовірністю значення X буде менше чи більше Me:

$$\underline{P(X < Me) = P(X > Me) = 0{,}5}. \tag{3.20}$$

З геометричної точки зору пряма $x = Me$ ділить пополам площу фігури, обмеженої кривою розподілу.

Приклад. Знайти моду і медіану неперервної випадкової величини X, заданої своєю диференціальною функцією

$$f(x)=\begin{cases}\frac{3}{4}\left(4x-x^2-3\right) & \text{при } x\in(1;3],\\ 0 & \text{при } x\notin(1;3].\end{cases}$$

<u>Розв’язання.</u> Моду M_0 знаходимо з умови (3.19). По-перше, знаходимо похідні першого та другого порядку:

$$f'(x)=\frac{3}{4}(4-2x)=\frac{3}{2}(2-x); \qquad f''(x)=\frac{3}{2}\cdot(-1)=-\frac{3}{2}.$$

За першою похідною знайдемо стаціонарну точку: $f'(x)=0$ при $x=2$. Оскільки друга похідна в цій точці від’ємна $f''(1{,}5)=-\frac{3}{2}<0$, то $x=2$ є точкою максимуму функції. Таким чином, мода НВВ X дорівнює $M_0=2$.

Медіану Me знайдемо з умови (3.20).

$$P(X<m)=\frac{3}{4}\int_0^m\left(4x-x^2-3\right)dx=\frac{3}{4}\cdot\left(2x^2-\frac{x^3}{3}-3x\right)\Bigg|_0^m=\frac{3}{2}m^2-\frac{1}{4}m^3-\frac{9}{4}m;$$

Для знаходження m складаємо і розв’язуємо рівняння:

$$\frac{3}{2}m^2-\frac{1}{4}m^3-\frac{9}{4}m=\frac{1}{2};$$
$$m^3-6m^2+9m-2=0;$$
$$(m-2)(m^2-4m+1)=0;$$
$$m_1=2;\quad m_2=2+\sqrt{3};\quad m_3=2-\sqrt{3}.$$

Лише перше з цих значень належить проміжку $(1;3]$ розподілу НВВ X. Таким чином медіана даної випадкової величини дорівнює $Me=2$. На рис. 3.10 зображено отримані моду та медіану даної випадкової величини.

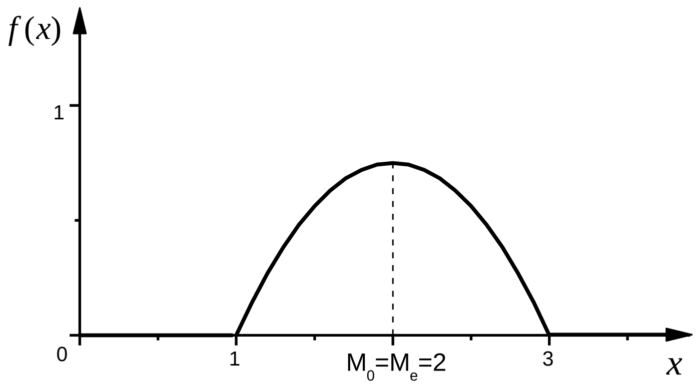


Рис. 3.10

Завдання

1. Час очікування, хв, пасажиром приміського поїзда даного напряму є неперервною випадковою величиною X, розподіл якої задано інтегральною функцією

$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ \dfrac{x}{10} & \text{при} \quad 0<x\le 10,\\ 1 & \text{при} \quad x>10.\end{cases}$$

Знайти середній час очікування $M(X)$ і середнє квадратичне відхилення $\sigma(X)$.

Відповідь. $M(X)=5$ хв, $\sigma(X)\approx 2{,}89$ хв.

2. Інтервал часу, хв, між пасажирами, що стають в чергу за квитками, є неперервною випадковою величиною T, щільність розподілу якої дорівнює

$$f(t)=\begin{cases}0{,}2\cdot e^{-0{,}2t} & \text{при} \quad t\ge 0,\\ 0 & \text{при} \quad t<0.\end{cases}$$

Знайти середній інтервал часу між прибуттям пасажирів $M(T)$, середнє квадратичне відхилення $\sigma(T)$ і коефіцієнт варіації.

Відповідь. $M(T)=\sigma(T)=5$ хв; $V(T)=1$.

3. НВВ X задана своєю щільністю розподілу

$$f(x)=\begin{cases}4-2x & \text{при} \quad x\in(1;2],\\ 0 & \text{при} \quad x\notin(1;2].\end{cases}$$

Знайти числові характеристики НВВ X.

Відповідь. $M(X)=\frac{4}{3}; D(X)=\frac{1}{18}; \sigma(X)=\frac{1}{3\sqrt{2}}; V(X)=\frac{1}{4\sqrt{2}}$.

4. НВВ X задана своєю диференціальною функцією

$$f(x)=\begin{cases}x+ax^2 & \text{при} \ x\in(1;2],\\ 0 & \text{при} \ x\notin(1;2].\end{cases}$$

Знайти невідомий параметр, математичне сподівання та середнє квадратичне відхилення.

Відповідь. $a=-\frac{3}{14}$; $M(X)=\frac{257}{168}\approx 1{,}53; \sigma(X)\approx 0{,}28$.

5. Знайти математичне сподівання НВВ X, заданої функцією розподілу

$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ e^x-1 & \text{при} \quad 0<x\le \ln 2,\\ 1 & \text{при} \quad x>\ln 2.\end{cases}$$

Відповідь. $M(X)=2\ln 2-1$.

6. Функція розподілу НВВ X має вигляд

$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ \sqrt[3]{x} & \text{при} \quad 0<x\le 1,\\ 1 & \text{при} \quad x>1.\end{cases}$$

Знайти числові характеристики X.

Відповідь. $M(X)=\frac{1}{4}; \ D(X)=\frac{9}{112}; \ \sigma(X)=\frac{3}{4\sqrt{7}}; V(X)=\frac{3}{\sqrt{7}}$.

7. Знайти моду і медіану НВВ X, заданої щільністю розподілу

$$f(x)=\begin{cases}\frac{3}{4}(1-x^2) & \text{при} \quad x\in(-1;1],\\ 0 & \text{при} \quad x\notin(-1;1].\end{cases}$$

Відповідь. $M_0 = Me = 0$.

8. Знайти моду, математичне сподівання, дисперсію, середнє квадратичне відхилення і коефіцієнт варіації неперервної випадкової величини X, заданої своєю диференціальною функцією

$$f(x)=\begin{cases}\frac{1}{9}\left(6x-2x^2\right) & \text{при} \ x\in(0;3],\\ 0 & \text{при} \ x\notin(0;3].\end{cases}$$

Відповідь. $M_0 = 1{,}5$; $M(X)=1{,}5;$ $D(X)=0{,}45;$ $\sigma(X)\approx 0{,}67; V(X)\approx 0{,}45$.

9. Щільність розподілу X дорівнює

$$f(x)=\begin{cases}kx-0{,}25x^3 & \text{при} \ x\in(0;2],\\ 0 & \text{при} \ x\notin(0;2].\end{cases}$$

Знайти невідомий параметр, моду і медіану НВВ X.

Відповідь. $k=1;$ $M_0=\frac{2}{\sqrt{3}};$ $Me=\sqrt{4-\sqrt{8}}$.

10. Знайти моду та медіану випадкової величини, заданої щільністю

$$f(x)=\begin{cases}0 & \text{при} \quad x\le -1,\\ x+1 & \text{при} \ -1<x\le 0,\\ 1-x & \text{при} \quad 0<x\le 1,\\ 0 & \text{при} \quad x>1.\end{cases}$$

Відповідь. $M_0 = Me = 0$.

3.8. Початкові і центральні моменти. Асиметрія і ексцес

Тут ми розглянемо деякі додаткові числові характеристики випадкових величин.

При наявності великих, але малоймовірних значень випадкової величини, виявляється недостатньо математичного сподівання. Щоб краще врахувати вплив цих значень, доцільно розглядати математичне сподівання не тільки самої випадкової величини, а й цілої додатної степені випадкової величини. Для цього вводиться нижченаведена числова характеристика випадкової величини.

Початковим моментом порядку k випадкової величини X називається математичне сподівання величини X^k:

$$\underline{\nu_k = M(X^k)}. \tag{3.21}$$

Для дискретної випадкової величини

$$\nu_k = \sum_i x_i^k \cdot p_i = x_1^k \cdot p_1 + x_2^k \cdot p_2 + \ldots,$$

для неперервної – $\nu_k = \int\limits_{-\infty}^{+\infty} x^k \cdot f(x)dx.$

Згідно з означенням математичне сподівання випадкової величини дорівнює початковому моменту першого порядку:

$$\underline{\nu_1 = M(X)}. \tag{3.22}$$

Оскільки $\nu_2 = M(X^2)$, то в термінах початкових моментів дисперсія виражається у вигляді

$$\underline{D(X) = \nu_2 - \nu_1^2}. \tag{3.23}$$

На практиці важливими стають також моменти $[X - M(X)]$ – відхилення випадкової величини X від математичного сподівання.

Центральним моментом порядку k випадкової величини X називається математичне сподівання k-ї степені відхилення $[X - M(X)]^k$:

$$\underline{\mu_k = M\left(\left[X - M(X)\right]^k\right)} \qquad (3.24)$$

Для дискретної випадкової величини ця формула має вигляд

$$\mu_k = \sum_i (x_i - m_x)^k \cdot p_i = (x_1 - m_x)^k \cdot p_1 + (x_2 - m_x)^k \cdot p_2 + \ldots,$$

для неперервної – $\mu_k = \int\limits_{-\infty}^{+\infty} (x - m_x)^k \cdot f(x)dx$.

З означення центральних моментів випливає, що для будь-якої випадкової величини:

$$\mu_1 = M\left[(X - M(X))\right] = 0, \qquad \mu_2 = M\left[(X - M(X))^2\right] = D(X).$$

З наведених означень нескладно встановити співвідношення між початковими і центральними моментами:

$$\underline{\mu_2 = \nu_2 - \nu_1^2};$$
$$\underline{\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3}; \qquad (3.25)$$
$$\underline{\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4}.$$

Саме цими формулами зазвичай користуються для обчислення центральних початкових моментів замість означення.

Зауважимо, що дисперсія випадкової величини дорівнює центральному моменту другого порядку:

$$\underline{D(X) = \mu_2.} \qquad (3.26)$$

Якщо розподіл є симетричним відносно математичного сподівання, то всі центральні моменти непарних порядків дорівнюють нулю: $\mu_1 = \mu_3 = \mu_5 = \ldots = 0$.

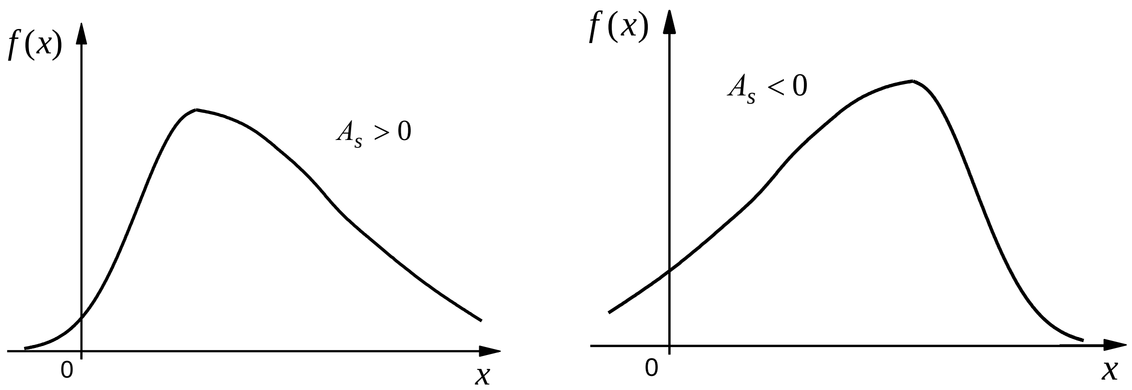
Зауваження. Розглянуті моменти називаються *теоретичними,* на відміну від обчислених за даними спостережень, що називаються *емпіричними.* Емпіричні розподіли та характеристики вивчає математична статистика (див. п. 7.7).

Через теоретичні центральні моменти визначаються асиметрія і ексцес теоретичного розподілу. Ці числові характеристики вводяться для визначення властивостей кривої розподілу (графіка щільності розподілу). Зазвичай вони використовуються для порівняння з нормальним розподілом, для якого крива розподілу має $A_s = E_k = 0$ (див. п. 4.7).

Асиметрією теоретичного розподілу називається відношення центрального моменту третього порядку до куба середнього квадратичного відхилення:

$$A_s = \frac{\mu_3}{\sigma^3}. \quad (3.27)$$

Для будь-якого симетричного відносно математичного сподівання розподілу асиметрія дорівнює нулю: $A_s = 0$. При $A_s > 0$ графік щільності розподілу «витягнутий» направо, при $A_s < 0$ – наліво (рис. 3.11).



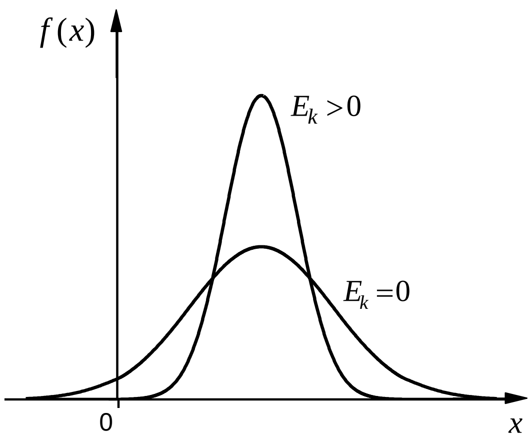
а) б)

Рис. 3.11

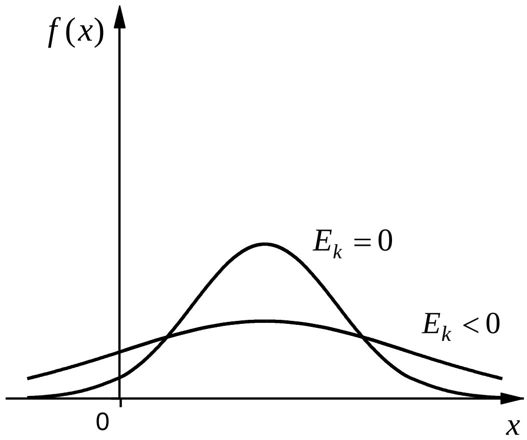
Ексцесом теоретичного розподілу називається величина

$$E_k = \frac{\mu_4}{\sigma^4} - 3. \quad (3.28)$$

Крива нормального розподілу має нульовий ексцес: $E_k = 0$. При $E_k > 0$ криві розподілів є більш «гостровершинними» за нормальну криву, а при $E_k < 0$ – більш «плосковершинними» (рис. 3.12).



а)



б)

Рис. 3.12

Зауваження. Аналогічний зміст мають асиметрія і ексцес емпіричного розподілу, що будуть розглянуті при вивченні математичної статистики (див. п.7.7) .

Приклад. Дискретна випадкова величина X задана рядом розподілу

x_i	-1	0	2	4	5
p_i	0,1	0,3	0,3	0,1	0,2

Знайти початкові і центральні моменти до четвертого порядку, асиметрію і ексцес. Визначити математичне сподівання, дисперсію, середнє квадратичне відхилення.

Розв'язання. Знаходимо початкові моменти за формулою (3.21):

$$\nu_1 = \sum_i x_i \cdot p_i = (-1)\cdot 0{,}1 + 0\cdot 0{,}3 + 2\cdot 0{,}3 + 4\cdot 0{,}1 + 5\cdot 0{,}2 = 1{,}9;$$

$$\nu_2 = \sum_i x_i^2 \cdot p_i = (-1)^2\cdot 0{,}1 + 0^2\cdot 0{,}3 + 2^2\cdot 0{,}3 + 4^2\cdot 0{,}1 + 5^2\cdot 0{,}2 = 7{,}9;$$

$$\nu_3 = \sum_i x_i^3 \cdot p_i = (-1)^3\cdot 0{,}1 + 0^3\cdot 0{,}3 + 2^3\cdot 0{,}3 + 4^3\cdot 0{,}1 + 5^3\cdot 0{,}2 = 33{,}7;$$

$$\nu_4 = \sum_i x_i^4 \cdot p_i = (-1)^4\cdot 0{,}1 + 0^4\cdot 0{,}3 + 2^4\cdot 0{,}3 + 4^4\cdot 0{,}1 + 5^4\cdot 0{,}2 = 155{,}5.$$

За формулою (3.22) математичне сподівання дорівнює $M(X) = \nu_1 = 1{,}9$.

Перший центральний момент завжди дорівнює нулю: $\mu_1 = 0$.

Для знаходження інших центральних моментів скористаємось формулами (3.25):

$\mu_2 = \nu_2 - \nu_1^2 = 7{,}9 - (1{,}9)^2 = 4{,}29$;

$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 = 33{,}7 - 3 \cdot 1{,}9 \cdot 7{,}9 + 2 \cdot (1{,}9)^3 = 2{,}388$;

$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4 = 155{,}5 - 4 \cdot 1{,}9 \cdot 33{,}7 + 6 \cdot (1{,}9)^2 \cdot 7{,}9 - 3 \cdot (1{,}9)^4 =$
$= 31{,}3977$.

Звичайно, цей результат можна перевірити за означенням (3.24).

Згідно з (3.26) дисперсія дорівнює $D(X) = \mu_2 = 4{,}29$. Як наслідок:

$$\sigma = \sqrt{D(X)} = \sqrt{4{,}29} \approx 2{,}07;\ \sigma^3 \approx (2{,}07)^3 = 8{,}8803;\ \sigma^4 \approx (2{,}07)^4 = 18{,}4041.$$

За формулами (3.27), (3.28) обчислюємо асиметрію і ексцес:

$$A_s = \frac{\mu_3}{\sigma^3} = \frac{2{,}388}{8{,}8803} = 0{,}26 \quad \text{і} \quad E_k = \frac{\mu_4}{\sigma^4} - 3 = \frac{31{,}3977}{18{,}4041} - 3 = -1{,}29.$$

Таким чином графік щільності даного розподілу є несиметричним (а саме, «витягнутим» направо) та «більш плоским» за нормальну криву.

Приклад. Неперервна випадкова величина X задана своєю диференціальною функцією

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le -1, \\ x+1 & \text{при} \quad -1 < x \le 0, \\ 1-x & \text{при} \quad 0 < x \le 1, \\ 0 & \text{при} \quad x > 1. \end{cases}$$

Знайти початкові і центральні моменти до четвертого порядку, асиметрію і ексцес.

Розв’язання. Початкові моменти знайдемо за формулою (3.21).

$$\nu_1 = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \int_{-1}^{0} x(x+1)dx + \int_{0}^{1} x(1-x)dx = \left(\frac{x^3}{3}+\frac{x^2}{2}\right)\Bigg|_{-1}^{0} + \left(\frac{x^2}{2}-\frac{x^3}{3}\right)\Bigg|_{0}^{1} =$$

$$= -\left(-\frac{1}{3}+\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)=0;$$

$$\nu_2 = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx = \int_{-1}^{0} x^2(x+1)dx + \int_{0}^{1} x^2(1-x)dx = \left(\frac{x^4}{4}+\frac{x^3}{3}\right)\Bigg|_{-1}^{0} + \left(\frac{x^3}{3}-\frac{x^4}{4}\right)\Bigg|_{0}^{1} =$$

$$= -\left(\frac{1}{4}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)=\frac{1}{6};$$

$$\nu_3 = \int_{-\infty}^{+\infty} x^3 \cdot f(x)dx = \int_{-1}^{0} x^3(x+1)dx + \int_{0}^{1} x^3(1-x)dx = \left(\frac{x^5}{5}+\frac{x^4}{4}\right)\Bigg|_{-1}^{0} + \left(\frac{x^4}{4}-\frac{x^5}{5}\right)\Bigg|_{0}^{1} =$$

$$= -\left(-\frac{1}{5}+\frac{1}{4}\right)+\left(\frac{1}{4}-\frac{1}{5}\right)=0;$$

$$\nu_4 = \int_{-\infty}^{+\infty} x^4 \cdot f(x)dx = \int_{-1}^{0} x^4(x+1)dx + \int_{0}^{1} x^4(1-x)dx = \left(\frac{x^6}{6}+\frac{x^5}{5}\right)\Bigg|_{-1}^{0} + \left(\frac{x^5}{5}-\frac{x^6}{6}\right)\Bigg|_{0}^{1} =$$

$$= -\left(\frac{1}{6}-\frac{1}{5}\right)+\left(\frac{1}{5}-\frac{1}{6}\right)=\frac{1}{15}.$$

Через початкові моменти знаходимо центральні за формулою (3.24):

$$\mu_2 = \nu_2 - \nu_1^2 = \frac{1}{6};$$

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3 = 0;$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4 = \frac{1}{15}.$$

Непарні центральні моменти невипадково виявились нульовими: даний розподіл дійсно є симетричним відносно математичного сподівання $M(X) = \nu_1 = 0$ (див. рис. 3.13). З тієї ж причини асиметрія також дорівнює нулю:

$$A_s = \frac{\mu_3}{\sigma^3} = 0.$$

Оскільки $D(X) = \sigma^2 = \mu_2 = \frac{1}{6}$, ексцес має величину

$$E_k = \frac{\mu_4}{\sigma^4} - 3 = \frac{1/15}{(1/6)^2} - 3 = -0{,}06.$$

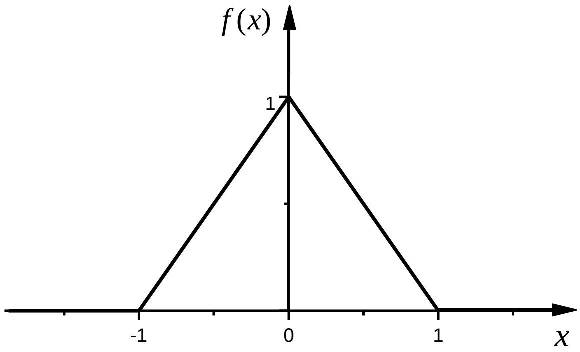


Рис. 3.13

Завдання

1. Протягом часу T на сортувальну станцію надійшло 3 подачі вагонів. Кількість подач, що мають вагони призначенням на вантажний двір розподілена за законом:

x_i	0	1	2	3
p_i	0,216	0,432	0,288	0,064

Знайти асиметрію і ексцес розподілу випадкової величини X.

<u>Відповідь.</u> $A_s \approx 0{,}236$; $E_k \approx -2{,}36$.

2. Дискретна випадкова величина X задана своїм рядом розподілу:

x_i	1	3	5	7	9
p_i	0,1	0,4	0,2	0,2	0,1

Знайти початкові і центральні моменти до четвертого порядку, асиметрію і ексцес. Чому дорівнюють математичне сподівання і дисперсія?

<u>Відповідь.</u> $\nu_1 = 4{,}6$; $\nu_2 = 26{,}6$; $\nu_3 = 177{,}4$; $\nu_4 = 1293{,}8$.

$\mu_1 = 0$; $\mu_2 = 5{,}44$; $\mu_3 = 4{,}992$; $\mu_4 = 64{,}55$. $A_s = 0{,}394$; $E_k = -0{,}82$.

$M(X) = \nu_1 = 4{,}6$; $D(X) = \mu_2 = 5{,}44$.

3. Дискретна випадкова величина X задана своїм рядом розподілу:

x_i	2	4	6	8
p_i	0,4	0,3	0,2	0,1

Знайти початкові і центральні моменти до четвертого порядку, асиметрію і ексцес.

Відповідь. $\nu_1 = 4$; $\nu_2 = 20$; $\nu_3 = 116$; $\nu_4 = 752$. $\mu_1 = 0$; $\mu_2 = 4$; $\mu_3 = 4{,}8$; $\mu_4 = 35{,}2$. $A_s = 0{,}6$; $E_k = -0{,}8$.

4. Неперервна випадкова величина X задана своєю диференціальною функцією

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ x & \text{при} \quad 0 < x \le 1, \\ (2 - x) & \text{при} \quad 1 < x \le 2, \\ 0 & \text{при} \quad x > 2. \end{cases}$$

Знайти початкові і центральні моменти до четвертого порядку, асиметрію і ексцес.

Відповідь. $\nu_1 = 1$; $\nu_2 = 7/6$; $\nu_3 = 3/2$; $\nu_4 = 31/15$.
$\mu_1 = 0$; $\mu_2 = 1/16$; $\mu_3 = 0$; $\mu_4 = 1/15$. $A_s = 0$; $E_k = -0{,}6$.

5. Неперервна випадкова величина X задана своєю диференціальною функцією

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le 0, \\ 1{,}5 \cdot x^2 & \text{при} \quad 0 < x \le 1, \\ 1{,}5 \cdot (2 - x)^2 & \text{при} \quad 1 < x \le 2, \\ 0 & \text{при} \quad x > 2. \end{cases}$$

Знайти математичне сподівання, середнє квадратичне відхилення, початкові і центральні моменти до четвертого порядку, асиметрію і ексцес.

Відповідь. $M(X) = 1$; $\sigma(X) \approx 0{,}316$; $\nu_1 = 1$; $\nu_2 = 1{,}1$; $\nu_3 = 1{,}3$; $\nu_4 = 1{,}628$.
$\mu_1 = 0$; $\mu_2 = 0{,}1$; $\mu_3 = 0$; $\mu_4 = 0{,}028$. $A_s = 0$; $E_k = -0{,}143$.

Питання до теми

41. Що таке випадкова величина? Чим відрізняються дискретні і неперервні випадкові величини? Наведіть приклади.

42. Що називається законом розподілу випадкової величини? Який вигляд може мати закон розподілу дискретної випадкової величини?

43. Чим задається розподіл неперервної випадкової величини?

44. Як визначається функція розподілу ймовірностей випадкової величини?

Чи однозначно задає вона випадкову величину?

45. Які властивості функції розподілу є спільними для ДВВ і НВВ, які − ні?

46. Які властивості мають графіки функцій розподілу ДВВ і НВВ? Що називають кривою розподілу?

47. Що таке щільність розподілу ймовірностей? У яких випадкових величин її немає? Наведіть властивості щільності розподілу.

48. Чи можна знайти функцію розподілу за щільністю та навпаки? Наведіть формули.

49. Яку функцію називають диференціальною? Інтегральною?

50. Які числові характеристики випадкових величин ви знаєте?

51. Який ймовірнісний зміст математичного сподівання? Наведіть формули для обчислення математичного сподівання ДВВ і НВВ.

52. Чи може математичне сподівання бути від'ємним? Чому дорівнює математичне сподівання сталої? Які ще властивості $M(X)$ вам відомі?

53. Який ймовірнісний зміст дисперсії? Середнього квадратичного відхилення? Наведіть формули для обчислення дисперсії ДВВ і НВВ.

54. Чи може дисперсія бути від'ємною? Середнє квадратичне відхилення? Наведіть властивості дисперсії.

55. Як обчислити коефіцієнт варіації? Що він характеризує?

56. Як визначається мода для ДВВ і НВВ?

57. Чи існує медіана для ДВВ? Для НВВ? Наведіть означення медіани.

58. Наведіть означення початкових та центральних моментів. Як через них виразити математичне сподівання? Дисперсію? Середнє квадратичне відхилення?

59. В яких випадках непарні центральні моменти дорівнюють нулю?

60. Для обчислення яких характеристик використовуються початкові та центральні моменти?

61. Як обчислити асиметрію та ексцес і що вони характеризують?

62. Коли асиметрія дорівнює нулю? Ексцес?

Тестові питання

1. В урні 12 кульок, з яких 8 − червоного кольору, інші − зелені. Дискретна випадкова величина X дорівнює кількості зелених кульок серед п'яти навмання взятих з урни. Яких значень може набувати X ?

А	Б	В	Г	Д
1,2,3,4,5	0,1,2,3,4,5	0,1,2,3,4	1,2,3,4,5,6,7,8	Інша відповідь

2. Дискретна випадкова величина має розподіл

X	1	2	3
P	0,1	0,2	?

Знайти невідому ймовірність.

А	Б	В	Г	Д
0,3	0,1	1	0,7	Інша відповідь

3. Закон розподілу ДВВ X , що задано таблицею

X	-3	0	6
P	0,1	0,6	0,3

Математичне сподівання ДВВ X дорівнює:

А	Б	В	Г	Д
1,5	2,1	1,1	-0,3	Інша відповідь

4. Математичне сподівання випадкової величини X дорівнює $M(X)=2$. Знайти $M(3X+1)$.

А	Б	В	Г	Д
6	7	2	1	Інша відповідь

5. Дисперсії незалежних випадкових величин X і Y дорівнюють $D(X)=3$ і $D(Y)=5$. Знайти $D(X-Y)$.

А	Б	В	Г	Д
3	5	-2	2	Інша відповідь

6. Закон розподілу дискретної випадкової величини задано таблицею:

X	-1	0	2	4	5
P	0,1	0,2	0,15	0,25	0,3

Чому дорівнює функція розподілу $F(x)$ при від'ємних $-1<x\le 0$?

А	Б	В	Г	Д
1	0	0,1	0,2	Інша відповідь

7. Чому дорівнює ймовірність того, що неперервна випадкова величина X набуде деякого значення x_0?

А	Б	В	Г	Д
1	0	0,1	0,2	Інша відповідь

8. Чому дорівнює ймовірність того, що випадкова величина набуде значення з інтервалу $(\alpha;\beta)$?

А	Б	В	Г	Д
$f(\alpha)-f(\beta)$	$F(\beta)+F(\alpha)$	$f(\beta)-f(\alpha)$	$F(\beta)-F(\alpha)$	Інша відповідь

9. Диференціальна функція може бути знайдена за формулою:

А	Б	В	Г	Д
$f(x)=F'(x)$	$f(x)=\int\limits_{-\infty}^{x}F(z)dz$	$F(x)=\int\limits_{-\infty}^{x}f(z)dz$	$F(x)=f'(x)$	Інша відповідь

10. Інтегральна функція може бути знайдена за формулою:

А	Б	В	Г	Д
$f(x)=F'(x)$	$f(x)=\int\limits_{-\infty}^{x}F(z)dz$	$F(x)=\int\limits_{-\infty}^{x}f(z)dz$	$F(x)=f'(x)$	Інша відповідь

11. Функція розподілу неперервної випадкової величини має вигляд:

$$F(x)=\begin{cases}0 & \text{при} \quad x\le 0,\\ x^4 & \text{при} \quad 0<x\le 1,\\ 1 & \text{при} \quad x>1.\end{cases}$$

Тоді щільність розподілу визначається формулою:

А	Б	В
$f(x)= =\begin{cases}4x & \text{при } x\in(0;1],\\ 0 & \text{при } x\notin(0;1].\end{cases}$	$f(x)= =\begin{cases}4x^3 & \text{при } x\in(0;1],\\ 0 & \text{при } x\notin(0;1].\end{cases}$	$f(x)= =\begin{cases}\dfrac{x^5}{5} & \text{при } x\in(0;1],\\ 0 & \text{при } x\notin(0;1].\end{cases}$

Г	Д
$f(x)= =\begin{cases}x^5 & \text{при } x\in(0;1],\\ 0 & \text{при } x\notin(0;1].\end{cases}$	Інша відповідь

12. Дискретна випадкова величина X задана таблицею:

X	-2,5	1,1	3	7,9
P	0,17	0,21	0,43	0,19

Яку з характеристик неможливо обчислити?

А	Б	В	Г	Д
$M(X)$	μ_1	M_0	Me	Інша відповідь

<u>Відповіді:</u> 1.В; 2.Г; 3.А; 4.Б; 5.Д; 6.В; 7.Б; 8.Г; 9.А ; 10.В; 11.Б; 12.Г.

Розділ 4. Основні закони розподілу випадкових величин

Деякі розподіли ймовірностей є дуже важливими, оскільки відповідні випадкові величини часто використовуються для моделювання різних процесів, розв'язання широкого кола практичних задач. У даному розділі ми вивчимо:

● дискретні розподіли: біноміальний, Пуассона, геометричний та гіпергеометричний;

● неперервні розподіли: рівномірний, показниковий (експоненціальний), нормальний та Ерланга.

Крім того, ми розглянемо гамма-розподіл, розподіли χ^2 («хі-квадрат») і Стьюдента.

4.1. Біноміальний розподіл

Нехай проводиться n незалежних випробувань, в кожному з яких деяка подія A (успіх) може з'явитись з імовірністю p і не з'явитись з імовірністю $q=1-p$. Нехай дискретна випадкова величина X дорівнює кількості успіхів (появ події A) у цій серії випробувань. ДВВ X може набувати значень $x_1=0, x_2=1, \ldots, x_{n+1}=n$, імовірності яких $p_i \ (i=1,\ldots,n+1)$ обчислюються за формулою Бернуллі (див. п. 2.1).

Біноміальним називається розподіл імовірностей, що описує кількість успіхів у схемі незалежних випробувань Бернуллі:

$$\underline{P(X=k)=P_n(k)=C_n^k p^k q^{n-k}}, \qquad k=0,1\ldots n, \qquad (4.1)$$

де n − кількість незалежних випробувань, у кожному з яких певна подія A (успіх) може з'явитись з імовірністю p і не з'явитись з імовірністю $q=1-p$; k − кількість успіхів. Числа $n \in Z_+$ − кількість випробувань і $0 \le p \le 1$ − ймовірність успіху називаються *параметрами біноміального розподілу.*

Наприклад:

1) кількість X влучень стрільцем при здійсненні $n=5$ пострілів, якщо ймовірність влучання $p=0{,}9$, розподілена за біноміальним законом

$$P(X=k)=P_5(k)=C_5^k \cdot (0{,}9)^k \cdot (0{,}1)^{5-k}, \qquad k=0,1,2,3,4,5;$$

2) кількість X «гербів», що випадуть при киданні трьох монет $(n=3,\ p=0{,}5)$, розподілена за законом

$$P(X=k)=P_3(k)=C_3^k\cdot(0{,}5)^k\cdot(0{,}5)^{3-k}, \qquad k=0{,}1{,}2{,}3.$$

За допомогою формул розд.3 отримаємо властивості біноміального розподілу, наведені в табл. 4.1.

Таблиця 4.1

Біноміальний розподіл ймовірностей	$P(X=k)=C_n^k p^k q^{n-k}$, $k=0{,}1\ldots n$
Параметри	$n\in Z_+$ − кількість випробувань; $0\le p\le 1$ − ймовірність успіху
Математичне сподівання	$M(X)=np$
Дисперсія	$D(X)=npq=np(1-p)$
Середнє квадратичне відхилення	$\sigma(X)=\sqrt{npq}=\sqrt{np(1-q)}$
Коефіцієнт варіації	$V(X)=\sqrt{\dfrac{q}{np}}$

Приклад. У подачі порожніх вагонів з імовірністю $p=0{,}2$ кожний з них вимагає очищення. Знайти розподіл кількості вагонів, що вимагають очищення в подачі з чотирьох вагонів.

Розв’язання. Подачі вагонів розглядаємо як незалежні випробування. Нехай X − число вагонів, що вимагають очищення в даній подачі. X є дискретною випадковою величиною, розподіленою за біноміальним законом з параметрами $p=0{,}2$ та $n=4$:

$$P(X=k)=P_4(k)=C_4^k(0{,}2)^k(0{,}8)^{4-k}, \quad k=0{,}1{,}2{,}3{,}4.$$

Розподіл ДВВ X надамо у вигляді таблиці:

x_i	0	1	2	3	4
p_i	0,4096	0,4096	0,1536	0,0256	0,0016

Перевірка: $\sum_i p_i=$0,4096+0,4096+0,1536+0,0256+0,0016=1.

Приклад. Знайти числові характеристики кількості надходження критих порожніх вагонів в 10 подачах на під'їзну колію, якщо ймовірність надходження критих порожніх вагонів в кожній подачі дорівнює $0{,}3$.

Розв'язання. Подачі вагонів розглядаємо як серію з $n=10$ незалежних випробувань, у кожному з яких імовірність успіху (надходження критих порожніх вагонів) стала і дорівнює $p=0{,}3$. Тому середня кількість надходження критих порожніх вагонів дорівнює $M(X)=n\cdot p=10\cdot 0{,}3=3$, а дисперсія − $D(X)=npq=10\cdot 0{,}3\cdot 0{,}7=2{,}1$.

Завдання

1. Довести формули для математичного сподівання і дисперсії випадкової величини, розподіленої за біноміальним законом (див. табл. 4.1).

Вказівка. Скористатись формулами (3.5)−(3.7).

2. На шляху руху поїзда стоять чотири світлофори, кожний з яких з імовірністю 0,5 дозволяє або забороняє подальший рух. Побудувати розподіл ймовірностей числа X світлофорів, що будуть пройдені поїздом без зупинки.

Відповідь. $P(X=k)=P_4(k)=C_4^k\cdot(0{,}5)^4$, $k=0{,}1{,}2{,}3{,}4$.

x_i	0	1	2	3	4
p_i	0,0625	0,25	0,325	0,25	0,0625

3. У сортувальному парку п'ять колій, кожна з яких може виявитися зайнятою в даний момент з імовірністю 0,7. Знайти розподіл випадкової величини X, що дорівнює числу зайнятих в даний момент колій у сортувальному парку. Знайти середню та найімовірнішу кількість зайнятих колій.

Відповідь. $P(X=k)=P_5(k)=C_5^k\cdot(0{,}7)^k\cdot(0{,}3)^{5-k}$, $k=0{,}1{,}2{,}3{,}4{,}5$.

x_i	0	1	2	3	4	5
p_i	0,00243	0,02835	0,1323	0,3087	0,36015	0,16807

$M(X)=3{,}5$; $M_0=4$.

4. Протягом години на станцію прибуло три потяги для розформування, у кожному з яких з імовірністю 0,2 можуть бути вагони на дане призначення. Скласти закон розподілу числа потягів, у яких є вагони на дане призначення.

Відповідь. $P(X=k)=C_3^k\cdot(0{,}2)^k\cdot(0{,}8)^{3-k}$, $k=0{,}1{,}2{,}3$.

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

5. На сортувальну станцію, що обслуговує промисловий район, кожної доби надходить 20 подач із завантаженими вагонами. Імовірність наявності недовантажених вагонів у кожній подачі дорівнює 0,2. Знайти математичне сподівання і середнє квадратичне відхилення загальної кількості подач з недовантаженими вагонами.

Відповідь. $M(X)=4$, $\sigma(X)\approx 1{,}8$.

4.2. Закон розподілу Пуассона. Найпростіший потік подій

Нехай реалізується схема Бернуллі з великою кількістю випробувань, в яких успіх є малоймовірним. Формула Бернуллі в цьому випадку є неприйнятною, і відповідні ймовірності обчислюються за асимптотичною формулою Пуассона (2.14) (див. також п.2.3).

Закон розподілу Пуассона ДВВ X задається формулою

$$\underline{P(X=k)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}},\qquad \text{де}\quad \underline{\lambda=np},\quad k=0{,}1{,}2\ldots \tag{4.2}$$

і описує число X малоймовірних подій у великій кількості випробувань. Число $\lambda>0$ називається *параметром розподілу Пуассона*.

Властивості розподілу Пуассона наведені в табл. 4.2; графік розподілу Пуассона при різних значеннях параметра λ зображено на рис. 4.1.

Таблиця 4.2

Розподіл ймовірностей Пуассона	$P(X=k)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}$, $\lambda=np$ $k=0,1,2\ldots$
Параметр	$\lambda>0$
Математичне сподівання	$M(X)=\lambda$
Дисперсія	$D(X)=\lambda$
Середнє квадратичне відхилення	$\sigma(X)=\sqrt{\lambda}$
Коефіцієнт варіації	$V(X)=\frac{1}{\sqrt{\lambda}}$

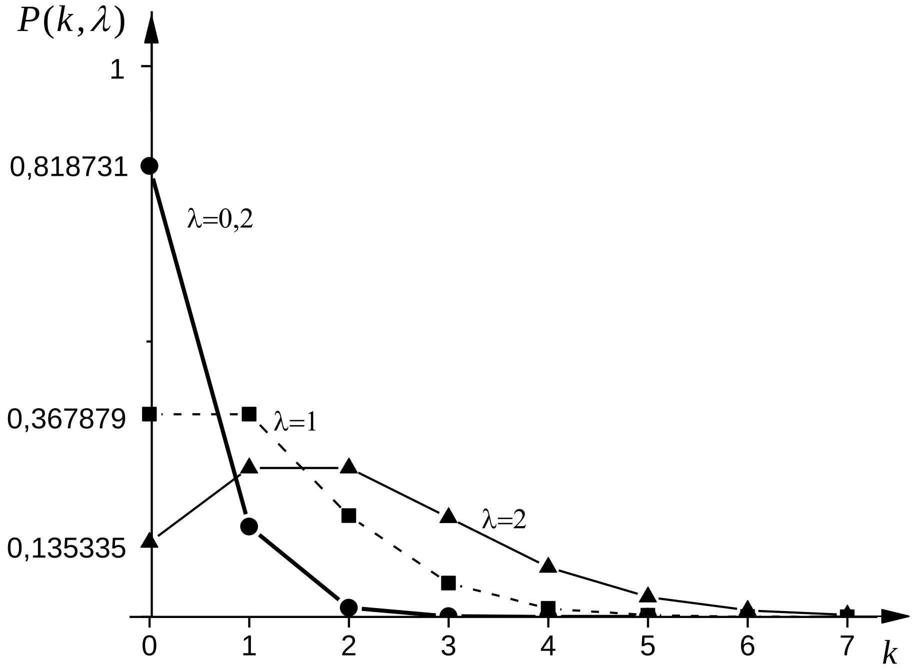


Рис. 4.1. Розподіл Пуассона

Зауваження. Нагадаємо, що значення $P(k,\lambda)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}$ при заданих k і λ можна знайти за спеціальними таблицями (див. дод.3).

Приклад. Кожної доби станція відправляє в потягах 5000 ваг. Імовірність того, що на шляху руху вагон потребує ремонту, дорівнює 0,0002. Знайти розподіл і числові характеристики числа вагонів, що вийдуть з ладу з технічних причин. Яка ймовірність того, що жоден вагон протягом доби не буде вимагати ремонту на шляху руху?

Розв’язання. Оскільки пошкодження вагона є малоймовірною подією, то потреба у ремонті розподілена за законом Пуассона. Нехай ДВВ X дорівнює числу вагонів, що вийдуть з ладу протягом доби, тоді X розподілена за законом Пуассона з параметром $\lambda = np = 5000 \cdot 0{,}0002 = 1$:

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \frac{1}{e \cdot k!}.$$

Математичне сподівання і дисперсія числа пошкоджених протягом доби вагонів дорівнюють $M(X) = D(X) = \lambda = 1$. Ймовірність того, що жодного пошкодженого вагона не виявиться протягом доби, знайдемо за таблицею дод.3 (в нашому випадку $k = 0,\ \lambda = 1$):

$$P(0) = \frac{1}{e \cdot 0!} \approx 0{,}367879.$$

Надзвичайно важливим для практичних задач є застосування закону Пуассона до так званого найпростішого потоку подій.

Потоком подій називається послідовність подій, що настають у випадкові моменти часу.

Найпростішим називається потік, що має властивості стаціонарності, відсутності післядії та ординарності.

Стаціонарність виражає незмінність імовірнісного режиму потоку протягом часу: ймовірність появи рівно k подій протягом часу t є функцією, що залежить лише від k і t. (Звідси середня кількість подій, що з’являються в одиницю часу, є сталою величиною.)

Відсутність післядії означає, що кількості подій у проміжки часу, що не перетинаються, не залежать одна від іншої.

Ординарність виражає практичну неможливість одночасної появи двох чи більше подій (тобто ймовірність появи більше однієї події за малий проміжок часу Δt нехтовно мала у порівнянні з імовірністю появи тільки однієї події).

Інтенсивністю потоку λ називається середня кількість подій, що з’являються в одиницю часу.

Найпростіший потік також називається *пуассонівським потоком,* тому що ймовірність появи k подій найпростішого потоку протягом часу t описується законом Пуассона з параметром λt:

$$P_t(X=k)=\frac{(\lambda t)^k}{k!}\cdot e^{-\lambda t}, \tag{4.3}$$

де λ − *інтенсивність потоку*.

Зауваження 1. Ми розглянули опис найпростішого потоку за допомогою дискретного розподілу (4.3) кількості подій у заданий проміжок часу. Пуассонівський потік також можна характеризувати неперервним, а саме, показниковим (експоненціальним) розподілом інтервалів між подіями (див. п.4.6).

Зауваження 2. На практиці інколи складно встановити властивості, що дозволять вважати потік найпростішим. Тому були знайдені деякі інші умови, за яких потік можна вважати близьким до пуассонівського. Наприклад, користуються ознакою: якщо потік є сумою великої кількості незалежних стаціонарних потоків, кожен з яких має зовсім незначний вплив на суму, то сумарний потік є найпростішим.

Наведемо приклади найпростіших (пуассонівських) потоків:

а) надходження викликів на телефонну станцію. Якщо в середньому надходить $\lambda=2$ виклики за хвилину, то ймовірність того, що за час *t, хв* надійде k викликів, дорівнює

$$P_t(k)=\frac{(2t)^k}{k!}\cdot e^{-2t};$$

б) надходження потягів до сортувального парку. Нехай середня інтенсивність прибуття $\lambda=3$ *потяг/год*. ДВВ X − кількість потягів, що прибудуть за час *t, год*, − розподілена за законом

$$P_t(X=k)=\frac{(3t)^k}{k!}\cdot e^{-3t};$$

в) надходження пасажирів до білетної каси. Нехай в середньому за годину до каси звертаються 12 *пас*. Тоді число X заявок, що надійдуть за час *t, хв*, є розподіленим за законом

Пуассона $P_t(X=k)=\frac{(0{,}2t)^k}{k!}\cdot e^{-0{,}2t}$ з інтенсивністю $\lambda=\frac{12}{60}=0{,}2$ *пас/хв*.

Таким чином, формула Пуассона служить *математичною моделлю* найпростішого потоку випадкових подій.

Приклад. По магістралі проходить пуассонівський потік автомобілів з інтенсивністю $\lambda=6$ *маш/хв*. Для переходу магістралі пішоходу необхідно 15 *с*. Обчислити ймовірність очікування пішоходом вільної магістралі.

Розв'язання. Очікування пішоходом вільної магістралі подія, що є протилежною до події відсутності машин в період $t=15\,с=0{,}25\,хв$. Ймовірність того, що протягом цього часу не проїде жодної машини обчислюється за формулою (4.3):

$$P(0)=\frac{(\lambda t)^k}{k!}\cdot e^{-\lambda t}\bigg|_{k=0}=\frac{(6\cdot 0{,}25)^0}{0!}\cdot e^{-6\cdot 0{,}25}=e^{-1{,}5}\approx 0{,}223.$$

Звідси ймовірність $P_{оч}$ того, що пішохід змушений буде чекати вільної магістралі, дорівнює $P_{оч}=1-P(0)=1-0{,}223=0{,}777$.

Приклад. На сортувальну станцію, що обслуговує промисловий район з трьома підприємствами, надходять подачі із завантаженими вагонами. Середній час очікування простою партії вагонів під завантаженням на кожному підприємстві складає 4 *год*, а надходження подач вагонів на станцію підлягає пуассонівському закону розподілу. Знайти середню кількість подач, що надійдуть протягом 2 *год*. Знайти ймовірність того, що за півтори години на станцію надійдуть: а) рівно одна подача; б) хоча б одна подача; в) не більше двох подач.

Розв'язання. Число подач, що надходять на станцію з одного підприємства за годину, дорівнює $\lambda_1=\frac{1}{4}=0{,}25$ *потяг/год*, а загальна кількість подач за годину дорівнює $\lambda=3\lambda_1=0{,}75$ *потяг/год*.

Нехай ДВВ X − число подач, що надходять за час *t*, тоді X підлягає закону розподілу Пуассона (4.3) з інтенсивністю $\lambda=0{,}75$ *потяг/год*. Тому середня кількість подач, що поступлять протягом $t=2$ годин, дорівнює математичному сподіванню $M(X)=\lambda\cdot t=0{,}75\cdot 2=1{,}5$ умовних потяги.

Ймовірності того, що за $t = 1{,}5$ *год* на станцію прибуде:

а) рівно одна подача:

$$P_{t=1,5}(k=1) = \frac{(0{,}75 \cdot 1{,}5)^1}{1!} \cdot e^{-0{,}75 \cdot 1{,}5} \approx 1{,}125 \cdot 0{,}325 \approx 0{,}366;$$

б) хоча б одна подача:

$$P_{t=1,5}(k \geq 1) = 1 - P_{t=1,5}(0) = 1 - \frac{1{,}125^0}{0!} \cdot e^{-1{,}125} \approx 1 - 0{,}325 = 0{,}675;$$

в) не більше двох подач:

$$P_{t=1,5}(k \leq 2) = P_{t=1,5}(0) + P_{t=1,5}(1) + P_{t=1,5}(2) \approx 0{,}325 + 0{,}366 + 0{,}1898 = 0{,}8808.$$

Завдання

1. Освітлення сортувального парку станції здійснюється 500 електричними лампочками. Протягом місяця 1% лампочок виходить з ладу незалежно одна від одної. Якому закону розподілу підлягає випадкова величина X, яка дорівнює числу лампочок, що вийшли з ладу? Яка ймовірність, що протягом місяця вийдуть з ладу: а) дві лампочки; б) не менше двох лампочок?

<u>Відповідь.</u> ДВВ X розподілена за законом Пуассона з параметром $\lambda = 5$; $P(2) \approx 0{,}08$; $P(k \geq 2) = 1 - \left(P(0) + P(1)\right) \approx 0{,}96$.

2. До підприємства побутового обслуговування надходять в середньому 120 заявок за добу. Скласти закон розподілу числа X заявок на обслуговування, що надійдуть протягом години. Знайти ймовірність того, що протягом години до підприємства надійде не менше трьох викликів.

<u>Відповідь.</u> Закон розподілу ДВВ X: $P(X = k) = \dfrac{5^k}{k!} \cdot e^{-5}$. $P(k \geq 3) \approx 0{,}88$.

3. Протягом години комутатор отримує в середньому 60 викликів. Випадкова величина X дорівнює числу викликів, що надходять за час t. Знайти закон розподілу ДВВ X та ймовірність того, що за 30 с відсутності телефоністки жодного виклику не надійде.

<u>Відповідь.</u> ДВВ X розподілена за законом Пуассона (4.3) з інтенсивністю $\lambda = 1$ (викликів за хвилину). $P_{t=0,5}(0) \approx 0{,}61$.

4. У середньому до залізничної каси звертаються з приводу придбання квитків 60 *пас/год*. Враховуючи, що такі особи утворюють найпростіший потік, обчислити ймовірність того, що за 3 *хв* до каси надійдуть: а) 3 пасажири; б) не більш як 3 пасажири.

Відповідь: а) 0,224042; б) 0,647232.

5. Довідкове бюро на залізничній станції відповідає на запитання, що надходять до нього в реальному масштабі часу. Протягом 1 *хв* бюро дає відповіді на 4 запитання. Беручи до уваги, що потік запитань є найпростішим, обчислити ймовірності таких подій: а) за 2 *хв* бюро відповість на 5 запитань; б) за 2 *хв* бюро дасть відповіді від двох до шести запитань.

Відповідь: а) 0,091604; б) 0,3103551.

6. У години «пік» через пропускний автомат станції метро за 1 *с* проходить у середньому 1 пасажир. Потік пасажирів вважають найпростішим. Яка ймовірність того, що за 5 *с* через пропускний автомат станції метро пройдуть: а) 4 пасажири; б) від одного до чотирьох пасажирів?

Відповідь: а) 0,175467; б) 0,433755.

7. На АЗС за кожну хвилину надходять у середньому два автомобілі для заправлення пальним. Потік автомобілів для заправлення вважають найпростішим. Яка ймовірність того, що за 3 *хв* на АЗС для заправлення надійде: а) один автомобіль; б) не більш як три автомобілі?

Відповідь: а) 0,0148725; б) 0,151203718.

8. У транзитний парк сортувальної станції щогодини прибувають два потяги, кожний з яких у середньому простоює в парку одну годину. Визначити ймовірність приймання потягів у парк без затримок, якщо в ньому існує 4 колії, а прибуття потягів підкоряється найпростішому потоку з параметром $\lambda = 2$. (Затримок через відсутність приймання потягів не буде, якщо протягом однієї години їх прибуде не більше 4).

Відповідь. 0,85712346.

4.3. Геометричний розподіл

Нехай проводиться серія випробувань, у кожному з яких подія A (успіх) може з'явитись з імовірністю p і не з'явитись з імовірністю $q=1-p$. Випробування зупиняються при першому успіху. Нехай дискретна випадкова величина X дорівнює кількості випробувань, необхідних для появи події A. Множина можливих значень $x_1=1; x_2=2; x_3=3\ldots$ − необмежена, а ймовірності дорівнюють $p_1=p; p_2=qp; p_3=q^2p\ldots.$

Геометричний розподіл описує кількість спроб, що необхідні для отримання першого успіху в схемі незалежних випробувань Бернуллі. Геометричний закон розподілу дискретної випадкової величини X визначається формулою

$$\underline{P(X=k)=pq^{k-1}=p(1-p)^{k-1},} \qquad k=1,2,3\ldots, \tag{4.4}$$

де p − ймовірність успіху ($q=1-p$). Число $0<p<1$ − ймовірність успіху − є єдиним *параметром геометричного розподілу.*

Геометричний розподіл називається також *розподілом Фаррі.* Основні властивості геометричного розподілу наведені в табл. 4.3.

Таблиця 4.3

Геометричний розподіл ймовірностей	$P(X=k)=pq^{k-1}=p(1-p)^{k-1}$ $k=1,2,3\ldots$
Параметр	$0<p<1$ − ймовірність успіху
Математичне сподівання	$M(X)=\dfrac{1}{p}$
Дисперсія	$D(X)=\dfrac{q}{p^2}=\dfrac{1-p}{p^2}$
Середнє квадратичне відхилення	$\sigma(X)=\dfrac{\sqrt{q}}{p}=\dfrac{\sqrt{1-p}}{p}$
Коефіцієнт варіації	$V(X)=\sqrt{1-p}$

Приклади дискретних випадкових величин, розподілених за геометричним законом:

1) кількість пострілів X, що потрібно зробити стрільцю для враження мішені, якщо ймовірність влучання $p=0{,}7$, розподілена за законом

$$P(X=k)=0{,}7\cdot(0{,}3)^{k-1}, \qquad k=1,2,3\ldots;$$

2) кількість X підкидань кістки, що необхідно зробити до першого випадіння «шістки» $\left(p=\dfrac{1}{6}\right)$, має розподіл:

$$P(X=k)=\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{k-1}, \qquad k=1,2,3\ldots$$

Зауваження. Інколи геометричний закон розподілу визначається як опис кількості невдач $Y=X-1$ до першого успіху. Очевидно, що відповідний розподіл має вигляд:

$$P(Y=k)=pq^k=p(1-p)^k, \qquad k=0,1,2\ldots$$

4.4. Гіпергеометричний розподіл

Почнемо з класичного модельного прикладу. Нехай в партії з N виробів знаходиться M стандартних. Навмання дістають n виробів, не повертаючи їх (тобто формулу Бернуллі застосовувати неможливо). Нехай ДВВ X − кількість стандартних виробів серед n навмання обраних. Ймовірність того, що серед обраних виробів буде m стандартних визначаються формулою

$$P(X=m)=\frac{C_M^m\cdot C_{N-M}^{n-m}}{C_N^n}, \tag{4.5}$$

де $m=\max(0,M+n-N),\ldots,\min(M,n)$; $C_n^k=\dfrac{n!}{k!(n-k)!}$. Таким чином, значення, що набуває ДВВ X, залежать від чисел N, M, n.

Гіпергеометричний розподіл ймовірностей задається формулою (4.5) і визначає кількість успіхів у вибірці без повернення із скінченної сукупності. *Параметрами* розподілу є числа N, M, n.

Наведемо приклади ДВВ, розподілених за гіпергеометричним законом.

1. В урні міститься 7 білих і 6 чорних кульок. Число X білих кульок серед 5 навмання взятих з урни визначається гіпергеометричним законом з параметрами $N=7+6=13,\ M=7, n=5$:

$$P(X=m)=\frac{C_7^m\cdot C_6^{5-m}}{C_{13}^5}, \qquad m=0,\ldots,5.$$

(В даному випадку $\max(0, M+n-N)=0;\ \min(M,n)=5$.)

2. В поїзді 50 *ваг*, з яких 4 − спеціального призначення, інші − універсальні. Для перевірки обирають навмання 10 *ваг*. Число X спеціальних вагонів серед перевірених визначається гіпергеометричним законом з параметрами $N=50,\ M=4, n=10$:

$$P(X=m)=\frac{C_4^m\cdot C_{46}^{10-m}}{C_{50}^{10}}, \qquad m=0,\ldots,4.$$

(Тут $\max(0, M+n-N)=0; \min(M,n)=4$.)

3. У бригаді 15 залізничників, з яких 11 − молоді спеціалісти. Кількість X молодих спеціалістів серед 7 навмання обраних з бригади розподілена за гіпергеометричним законом з параметрами $N=15, M=11, n=7$:

$$P(X=m)=\frac{C_{11}^m\cdot C_4^{7-m}}{C_{15}^7}, \qquad m=3,\ldots,7.$$

($\max(0, M+n-N)=3; \min(M,n)=7$.)

Основні властивості гіпергеометричного розподілу наведені в табл. 4.4.

Таблиця 4.4

Гіпергеометричний розподіл ймовірностей	$P(X=m)=\dfrac{C_M^m\cdot C_{N-M}^{n-m}}{C_N^n}$, $m=\max(0, M+n-N),\ldots,\min(M,n)$
Параметри	$N, M, n\in Z_+$
Математичне сподівання	$M(X)=\dfrac{Mn}{N}$
Дисперсія	$D(X)=\dfrac{Mn(N-M)(N-n)}{N^2(N-1)}$

Зауваження 1. Зазначимо, що можливі значення ДВВ, розподіленої за гіпергеометричним законом зазвичай простіше знаходити з сенсу задачі, ніж за формулою $X = m = \max(0, M + n - N), \ldots, \min(M, n)$.

Зауваження 2. При $n << N$ $(n < 0{,}1N)$ гіпергеометричний розподіл наближений до біноміального, тому в масових випробуваннях ігнорується умова, що елемент не повертається.

Приклад. У партії з 10 виробів є 2 з прихованим дефектом. Навмання обирають 3. Скласти розподіл числа бракованих виробів.

Розв'язання. Нехай X − кількість виробів з дефектом серед обраних. Можливі значення ДВВ X: $x_1 = 0, x_2 = 1, x_3 = 2$ (Дійсно, більше двох не може бути за умовою задачі. Якщо скористатися формулами, отримаємо той самий результат: $\max(0, M + n - N) = 0$; $\min(M, n) = 2$). ДВВ X розподілена за гіпергеометричним законом з параметрами $N = 10, M = 2, n = 3$:

$$P(X = m) = \frac{C_2^m \cdot C_8^{3-m}}{C_{10}^3}, \; m = 0{,}1{,}2.$$

Обчислимо ймовірності значень випадкової величини:

$$p_1 = P(X = 0) = \frac{C_2^0 \cdot C_8^3}{C_{10}^3} = \frac{C_8^3}{C_{10}^3} = \frac{7}{15};$$

$$p_2 = P(X = 1) = \frac{C_2^1 \cdot C_8^2}{C_{10}^3} = \frac{7}{15};$$

$$p_3 = P(X = 2) = \frac{C_2^2 \cdot C_8^1}{C_{10}^3} = \frac{C_8^1}{C_{10}^3} = \frac{1}{15}.$$

Розподіл ймовірностей ДВВ X запишемо у вигляді таблиці:

x_i	0	1	2
p_i	$\frac{7}{15}$	$\frac{7}{15}$	$\frac{1}{15}$

Перевірка: $\sum\limits_i p_i = \frac{7}{15} + \frac{7}{15} + \frac{1}{15} = 1.$

Завдання

1. Кожної доби на сортувальну станцію в середньому прибувають $N = 80$ потягів, з яких $M = 60$ − розбірні. За чотири години на станцію прибуло 12 потягів. Скласти розподіл імовірностей числа розбірних потягів у вигляді формули. Обчислити ймовірність того, що розбірних потягів прибуло 10.

<u>Відповідь.</u> $P(X = m) = \dfrac{C_{60}^{m} \cdot C_{20}^{12-m}}{C_{80}^{12}}, m = 0,\ldots,12$.

$$P(X = 10) = \frac{C_{60}^{10} \cdot C_{20}^{2}}{C_{80}^{12}} \approx 0,24.$$

2. Відомо, що протягом часу T на станцію в середньому надходить 100 *ваг*, причому 7 з них − цистерни. Скласти розподіл ймовірностей числа цистерн серед вагонів, що надійшли, у вигляді формули. Обчислити ймовірність того, що з 40 *ваг* поїзда, який прибув на станцію, виявилось 3 цистерни.

<u>Відповідь.</u> $P(X = m) = \dfrac{C_{7}^{m} \cdot C_{93}^{40-m}}{C_{100}^{40}}, \quad m = 0,\ldots,7$.

$$P(X = 3) = \frac{C_{7}^{3} \cdot C_{93}^{37}}{C_{100}^{40}} \approx 0,3.$$

4.5. Рівномірний розподіл

Рівномірним називається розподіл ймовірностей неперервної випадкової величини, щільність якого є сталою на сегменті $[a;b]$ її можливих значень і дорівнює нулю поза ним.

Рівномірний розподіл зазвичай позначають $U(a;b)$. Числа a і b (кінці сегмента) називаються *параметрами рівномірного розподілу* ймовірностей.

З означення розподілу та властивостей функцій розподілу і щільності (див. п. 3.6) нескладно вивести відомості про рівномірний розподіл, наведені в табл. 4.5. Графіки диференціальної та інтегральної функцій рівномірного розподілу $U(a;b)$ зображено на рис. 4.2.

Таблиця 4.5

Щільність розподілу $U(a;b)$	$f(x)=\begin{cases}\dfrac{1}{b-a} & \text{при } x\in[a;b],\\ 0 & \text{при } x\notin[a;b].\end{cases}$
Функція розподілу	$F(x)=\begin{cases}0 & \text{при } x<a,\\ \dfrac{x-a}{b-a} & \text{при } a\le x\le b,\\ 1 & \text{при } x>b.\end{cases}$
Параметри	$a,b\in R$
Математичне сподівання	$M(X)=\dfrac{a+b}{2}$
Дисперсія	$D(X)=\dfrac{(b-a)^2}{12}$
Середнє квадратичне відхилення	$\sigma(X)=\dfrac{b-a}{2\sqrt{3}}$
Імовірність попадання в $(\alpha;\beta)$	$P(\alpha<X<\beta)=\dfrac{\beta-\alpha}{b-a}$ при $a\le\alpha<\beta\le b$

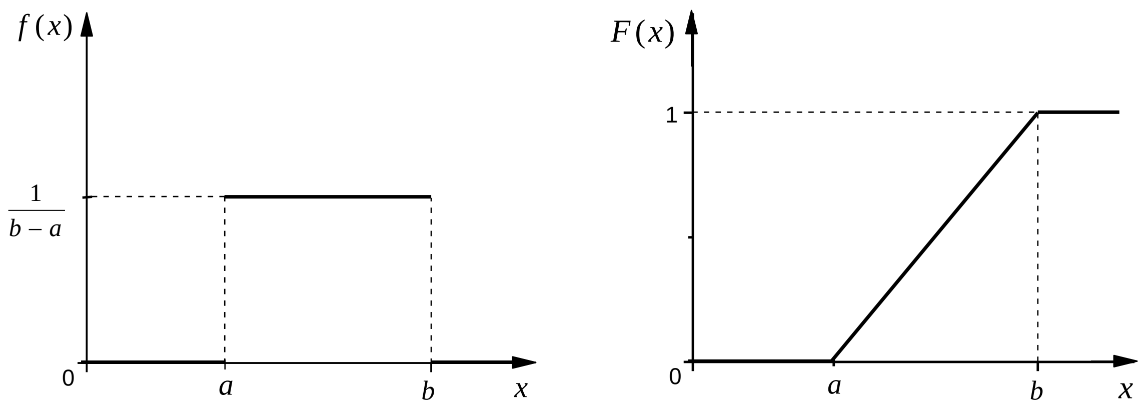


Рис. 4.2

Приклад. Функція щільності розподілу випадкової величини надходження порожніх вагонів до пункту завантаження має вигляд:

$$f(x)=\begin{cases}1 & \text{при } x\in[0;1],\\ 0 & \text{при } x\notin[0;1].\end{cases}$$

Виписати функцію розподілу та знайти числові характеристики.

Розв'язання. Оскільки щільність розподілу є сталою, вона відповідає рівномірному розподілу $U(a;b) = U(0;1)$ на сегменті $[a;b] = [0;1]$. Тому функція розподілу має вигляд:

$$F(x) = \begin{cases} 0 & \text{при} \quad x < 0, \\ x & \text{при} \quad 0 \le x \le 1, \\ 1 & \text{при} \quad x > 1. \end{cases}$$

Щоб знайти математичне сподівання, дисперсію і середнє квадратичне відхилення, скористаємось формулами, наведеними в табл. 4.5:

$$M(X) = \frac{a+b}{2} = \frac{0+1}{2} = \frac{1}{2};$$

$$D(X) = \frac{(b-a)^2}{12} = \frac{(1-0)^2}{12} = \frac{1}{12}; \qquad \sigma(X) = \sqrt{D(X)} = \frac{1}{2\sqrt{3}}.$$

Зауваження. Такий самий розподіл був розглянутий у прикладі п. 3.7. Легко бачити, що спеціальні формули для рівномірного розподілу значно спрощують обчислення числових характеристик.

Приклад. Електропоїзди метрополітену йдуть строго за розкладом. Інтервал руху складає 3 *хв*. Визначити середній час очікування електропоїзда пасажиром, середнє квадратичне відхилення і ймовірність того, що пасажир, який підійшов до зупинки, буде чекати менше 2 *хв.*

Розв'язання. При випадковому підході пасажирів до станції кожен з них з однаковою ймовірністю може опинитись у будь-якій точці інтервалу. Тому час очікування X є рівномірно розподіленою на інтервалі $[a,b] = [0,3]$ випадковою величиною. Тому середній час очікування потяга пасажиром дорівнює

$$M(X) = \frac{a+b}{2} = \frac{0+3}{2} = 1{,}5\ \text{хв},$$

а дисперсія та середнє квадратичне відхилення дорівнюють відповідно

$$D(X)=\frac{(b-a)^2}{12}=0{,}75\,хв^2; \quad \sigma(X)=\sqrt{D}\approx 0{,}66\,хв.$$

Ймовірність того, що пасажир буде чекати потяг менше 2 *хв* дорівнює

$$P(0\le X<2)=\frac{\beta-\alpha}{b-a}=\frac{2-0}{3-0}=\frac{2}{3}\approx 0{,}66.$$

Завдання

1. Довести формули для щільності рівномірного розподілу, функції розподілу та числових характеристик $U(a;b)$.

Вказівка. Скористатись формулами (3.15), (3.13), (3.17)−(3.20).

2. Маршрутне таксі рухається строго за розкладом, інтервал руху складає 10 *хв*. Записати вирази для функції розподілу і щільності. Знайти ймовірність того, що пасажир, який підійшов до зупинки, буде чекати не більше 6 *хв*.

Відповідь. $f(x)=\begin{cases}0{,}1 & \text{при} \quad x\in[0;10],\\ 0 & \text{при} \quad x\notin[0;10].\end{cases}$

$$F(x)=\begin{cases}0 & \text{при} \quad x<0,\\ 0{,}1\cdot x & \text{при} \quad 0\le x\le 10,\\ 1 & \text{при} \quad x>10.\end{cases}$$

$P(X\le 6)=0{,}6.$

3. Моторвагонні потяги відправляються зі станції з інтервалом $I=8$ *хв*. Визначити середній час очікування потяга пасажиром, середнє квадратичне відхилення та ймовірність того, що пасажир очікуватиме потяг більше 5 *хв*.

Відповідь. $M(X)=4\,хв$, $\sigma(X)\approx 2{,}31\,хв$, $P(X>5)\approx 0{,}375$.

4. Якщо виконується графік руху на маршруті, то середній час очікування пасажиром автобуса дорівнює 5 *хв*. Мінімальний час очікування – 0 *хв*. Відомо, що час очікування має рівномірний закон розподілу. Знайти його параметри. Яка ймовірність того, що пасажир очікуватиме автобус від 4 до 7 *хв*?

Відповідь. $a=0; b=10$. $P(4\le X\le 7)=0{,}3.$

5. Час надходження порожніх вагонів до пункту навантаження підлягає рівномірному закону з диференціальною функцією

$$f(x)=\begin{cases}1/4 & \text{при} \quad x\in[1;5],\\ 0 & \text{при} \quad x\notin[1;5].\end{cases}$$

Знайти математичне сподівання та дисперсію.

<u>Відповідь.</u> $M(X)=3,\ D(X)=4/3$.

6**.** Час очікування, *хв*, електропоїзда розподілений за законом $U(0;16)$. Знайти середній час очікування електропоїзда та ймовірність того, що пасажиру доведеться чекати електропоїзда не більше 10 *хв*.

<u>Відповідь.</u> $M(X)=8$ *хв*, $P(X\le 10)=0{,}625$.

4.6. Показниковий (експоненціальний) розподіл

Розподіл ймовірностей неперервної випадкової величини називається *показниковим (або експоненціальним)*, якщо він визначається щільністю

$$f(x)=\begin{cases}0 & \text{при} \quad x<0,\\ \lambda e^{-\lambda x} & \text{при} \quad x\ge 0,\end{cases} \qquad (\lambda>0) \tag{4.6}$$

число $\lambda>0$ називається *параметром показникового розподілу*. Показниковий (експоненціальний) розподіл зазвичай позначається $E(\lambda)$.

За допомогою відповідних формул розділу 3 нескладно отримати відомості про показниковий розподіл $E(\lambda)$ з параметром λ (табл. 4.6).

Графіки диференціальної та інтегральної функцій експоненціального розподілу $E(\lambda)$ при різних значеннях параметра λ зображено на рис. 4.3 та 4.4 відповідно.

Таблиця 4.6

Щільність розподілу $E(\lambda)$	$f(x)=\begin{cases}0 & \text{при } x<0,\\ \lambda e^{-\lambda x} & \text{при } x\ge 0,\end{cases}\quad \lambda>0$
Функція розподілу	$F(x)=\begin{cases}0 & \text{при } x<0,\\ 1-e^{-\lambda x} & \text{при } x\ge 0,\end{cases}\quad \lambda>0$
Параметр	$\lambda>0$
Математичне сподівання	$M(X)=\frac{1}{\lambda}$
Дисперсія	$D(X)=\frac{1}{\lambda^2}$
Середнє квадратичне відхилення	$\sigma(X)=\frac{1}{\lambda}$
Коефіцієнт варіації	$V(X)=1$
Імовірність попадання в $(\alpha;\beta)$	$P(\alpha<X<\beta)=e^{-\lambda\alpha}-e^{-\lambda\beta}$ при $0\le\alpha<\beta$
Мода	$Mo=0$
Медіана	$Me=\frac{\ln 2}{\lambda}$

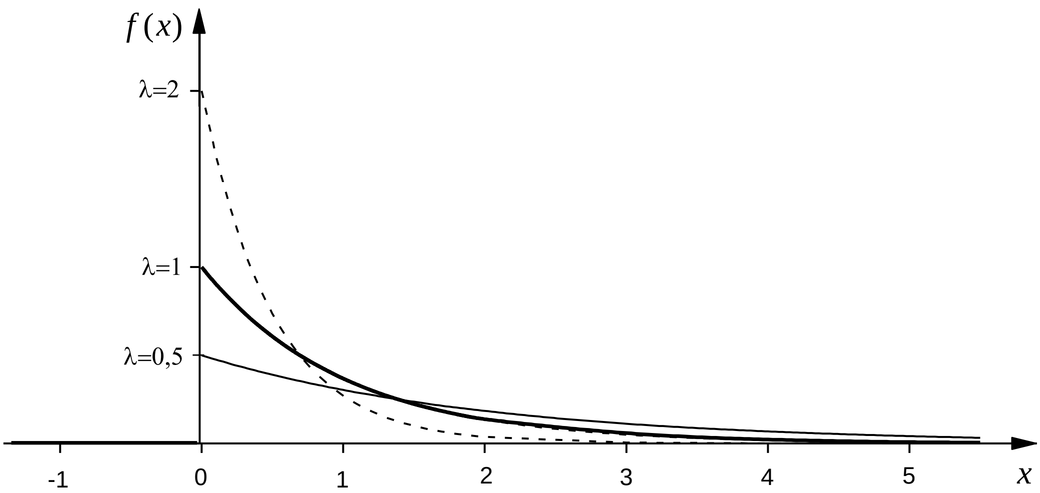


Рис. 4.3. Щільність експоненціального розподілу

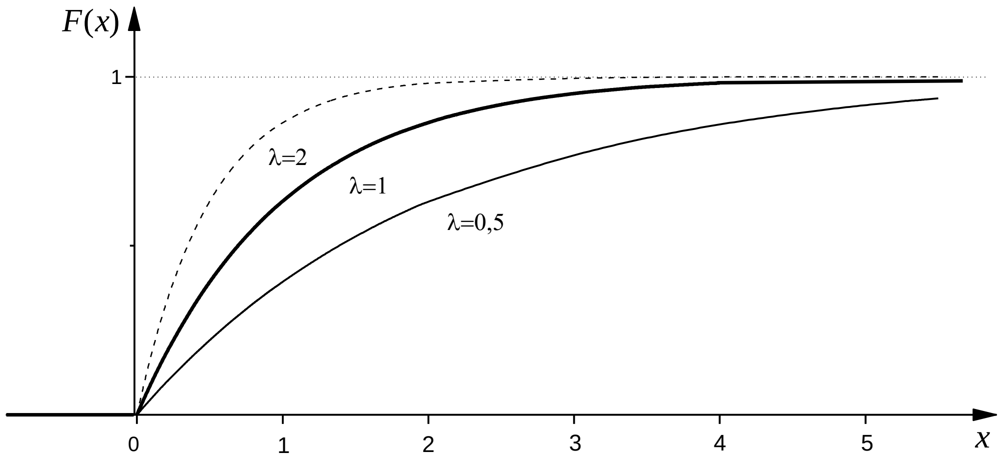


Рис. 4.4. Функція експоненціального розподілу

Зауваження 1. На практиці часто параметр показникового розподілу знаходять за відомим середнім значенням випадкової величини.

Зауваження 2. Характерне співвідношення $M^2(X) = D(X)$ між матема-тичним сподіванням і дисперсією показникового розподілу інколи стає вирішальним для висунення гіпотези про розподіл.

Зауваження 3. Ми розглянули опис найпростішого потоку за допомогою дискретного розподілу (4.3) кількості подій в заданий проміжок часу (див. п.4.2). Розподіл інтервалів між подіями найпростішого потоку є показниковим.

Зауваження 4. Експоненціальний закон є частинним випадком закону розподілу Ерланга та гамма-розподілу (див. п.п. 4.8, 4.9).

Приклад. Інтервал часу, *хв*, між пасажирами, що стають в чергу за квитками в залізничну касу, є випадковою величиною T, щільність розподілу якої дорівнює

$$f(t) = \begin{cases} 0 & \text{при} \quad t < 0, \\ 0{,}1 \cdot e^{-0{,}1 \cdot t} & \text{при} \quad t \ge 0. \end{cases}$$

Касирові потрібно замінити стрічку у касовому апараті, на що витрачається 2 *хв*. Яка ймовірність того, що за цей час не утвориться черга? Визначити середній інтервал часу між пасажирами.

Розв'язання. З формули щільності $f(t)$ дістанемо параметр показникового розподілу $\lambda = 0{,}1$ *пас/хв*. Те, що за 2 *хв* не утвориться черга в касу означає, що час очікування чергового пасажира перевищить 2 *хв*. Ймовірність цієї події дорівнює:

$$P(T > 2) = F(+\infty) - F(2) = 1 - (1 - e^{-0{,}1 \cdot 2}) = e^{-0{,}2} \approx 0{,}818\,.$$

Середній інтервал часу між пасажирами дорівнює $M(T) = \dfrac{1}{\lambda} = 10$ *хв*.

Приклад. Час ремонту та обслуговування автомобіля після однієї поїздки є випадковим і має показниковий закон розподілу. Практика свідчить, що в поточному сезоні на ремонт і обслуговування автомобіля після однієї поїздки витрачалось в середньому 0,1 *год*. Яка ймовірність того, що під час чергової поїздки цей час складатиме не більше 10 *хв*?

Розв'язання. За умовою математичне сподівання випадкової величини часу ремонту та обслуговування автомобіля дорівнює $M(T) = 0{,}1\,год = 6\,хв$. Оскільки T має показниковий закон розподілу, то параметр цього розподілу дорівнює $\lambda = \dfrac{1}{M(T)} = \dfrac{1}{6} \approx 0{,}17$, а функція розподілу − $F(t) = 1 - e^{-0{,}17t}$, $t \geq 0$. Ймовірність того, що під час чергової поїздки цей час складатиме не більше 10 *хв*:

$$P(T \leq 10) = P(0 \leq T \leq 10) = F(10) \approx 1 - e^{-0{,}17 \cdot 10} = 1 - e^{-1{,}7} \approx 1 - 0{,}18 = 0{,}82\,.$$

Якщо T − неперервна випадкова величина, що виражає тривалість часу безвідмовної роботи якого-небудь елемента, а λ − інтенсивність відмов (середня кількість відмов в одиницю часу), тоді T можна вважати випадковою величиною, розподіленою за показниковим розподілом з параметром λ. При цьому ймовірність відмови за час t визначається функцією розподілу $F(t) = 1 - e^{-\lambda t}$ $(t \geq 0)$, а ймовірність безвідмовної роботи елемента за час t − *функцією надійності*

$$\underline{R(t) = e^{-\lambda t}, \quad t \geq 0\,.} \qquad (4.7)$$

Приклад. Імовірність безвідмовної роботи елемента розподілено за показниковим законом $f(t) = 0{,}02e^{-0{,}02t}$, $t \ge 0$. Знайти ймовірність безвідмовної роботи елемента протягом 50 *год*.

Розв'язання. Використовуючи функцію надійності (4.7) отримуємо:

$$P = R(50) = e^{-50\cdot 0{,}02} = e^{-1} \approx 0{,}368 .$$

Завдання

1. Довести формули для функції $F(x)$ показникового розподілу та числових характеристик $E(\lambda)$ (див. табл. 4.6).

Вказівка. Скористатись формулами (3.13), (3.17)−(3.20).

2. Час розформування складу через гірку T − випадкова величина, що підлягає експоненціальному закону. Середня інтенсивність розформування дорівнює $\lambda = 5$ *потяг/год*. Скласти функцію розподілу і визначити ймовірності, що час розформування буде: а) менше півгодини; б) більше 5 хвилин, але менше 25 *хв*.

Відповідь. $F(t) = 1 - e^{-5t}$; $P(T < 0{,}5) \approx 0{,}92$; $P(0{,}08 < T < 0{,}4) \approx 0{,}53$.

3. Інтервали між прибуттям потягів на сортувальну станцію розподілені за показниковим законом. Інтенсивність надходження потягів на станцію дорівнює $\lambda = 4$ *потяг/год*. Знайти щільність розподілу; обчислити ймовірність того, що інтервал між прибуттям двох потягів буде не більше 20 хв.

Відповідь. $f(t) = \begin{cases} 4e^{-4t}, & \text{при } t \ge 0 \\ 0, & \text{при } t < 0 \end{cases}$. $P(T < 1/3) \approx 0{,}74$.

4. Товарний касир обслуговує клієнтів із середньою інтенсивністю шість чоловік за годину. Знайти ймовірність того, що даний клієнт буде обслуговуватися менше 15 *хв*, якщо час обслуговування T розподілений за показниковим законом.

Відповідь. $P(T < 0{,}25) \approx 0{,}78$.

5. На розігрів у тепляку партії вагонів з вантажами певної в'язкості за нормою потрібно 4 *год*. Насправді, у зв'язку з порушеннями термінів доставки вантажу, що змерзається за провиною залізниці, час розігріву T коливається від 4 до 10 *год* і підлягає експоненціальному розподілу. Необхідно визначити ймовірність того, що розігрів забере: а) більше 4, але менше 5 *год* ; б) більше 10 *год*.

Відповідь. $\lambda = \frac{1}{4}$ партії за годину; $P(4 < T < 5) = P(T > 10) \approx 0{,}082$.

6. Пасажири, що стають у чергу за квитками в залізничну касу, утворюють найпростіший потік, в якому інтервал часу, *хв*, між моментами прибуття пасажирів є випадковою величиною T з показниковим законом розподілу:

$$f(t) = \begin{cases} 0{,}5e^{-0{,}5t} & \text{при } t \geq 0, \\ 0 & \text{при } t < 0. \end{cases}$$

Знайти математичне сподівання $M(T)$. Обчислити ймовірність того, що протягом 3 *хв* жоден пасажир не з'явиться.

Відповідь. $M(T) = 2$ *хв*. $P(T > 3) \approx 0{,}223$.

7. Середній час, *год*, простою вагонів під накопиченням на вантажній станції складає 40 *год*. Відомо, що простій має експоненціальний розподіл. Виписати диференціальну та інтегральну функції розподілу. Знайти ймовірність того, що накопичення вагонів триватиме не більше 2 діб.

Відповідь. $f(t) = \begin{cases} 0{,}025 \cdot e^{-0{,}025t} & \text{при } t \geq 0, \\ 0 & \text{при } t < 0. \end{cases}$ $P(T < 48) \approx 0{,}7$.

4.7. Нормальний розподіл

Нормальний закон розподілу має виняткове значення в теорії ймовірностей. Він є найбільш поширеним у природі, господарстві і на транспорті. Крім того, за достатньо типових умов, до нормального закону наближаються інші закони, тобто він є граничним законом для багатьох інших (див. розд.6). Зокрема нормальному закону розподілу підлягають: кількість вагонів у поїздах, що прибувають на розформування; погонне навантаження; тривалість технічного обслуговування у парках; ходові швидкості руху поїздів на ділянках та ін.

Нормальним розподілом з параметрами a і σ називається розподіл імовірностей неперервної випадкової величини, що задається функцією щільності

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}, \qquad x\in R,\ a\in R,\ \sigma>0. \tag{4.8}$$

Імовірнісний зміст параметрів цього розподілу: $a\in R$ − математичне сподівання, $\sigma>0$ − середнє квадратичне відхилення. Нормальний розподіл з параметрами a і σ зазвичай позначають $N(a;\sigma)$.

Графік щільності розподілу $f(x)$ нормального розподілу називається *нормальною кривою або кривою Гаусса.*

Властивості нормальної кривої:

1) графік $f(x)$ є симетричним відносно вертикалі $x=a$:

$$f(a+x)=f(a-x);$$

2) функція $f(x)$ має максимум у точці $x=a$:

$$\max f(x)=f(a)=\frac{1}{\sigma\sqrt{2\pi}};$$

3) точки перегину нормальної кривої:

$$\left(a\pm\sigma;\frac{1}{\sigma\sqrt{2\pi e}}\right);$$

4) параметр a відповідає зсуву нормальної кривої вздовж осі OX: при $a>0$ − направо, при $a<0$ − наліво (див. рис. 4.5);

5) параметр σ відповідає за величину піку: при зменшенні σ нормальна крива витягується, при збільшенні − стягується (див. рис. 4.6).

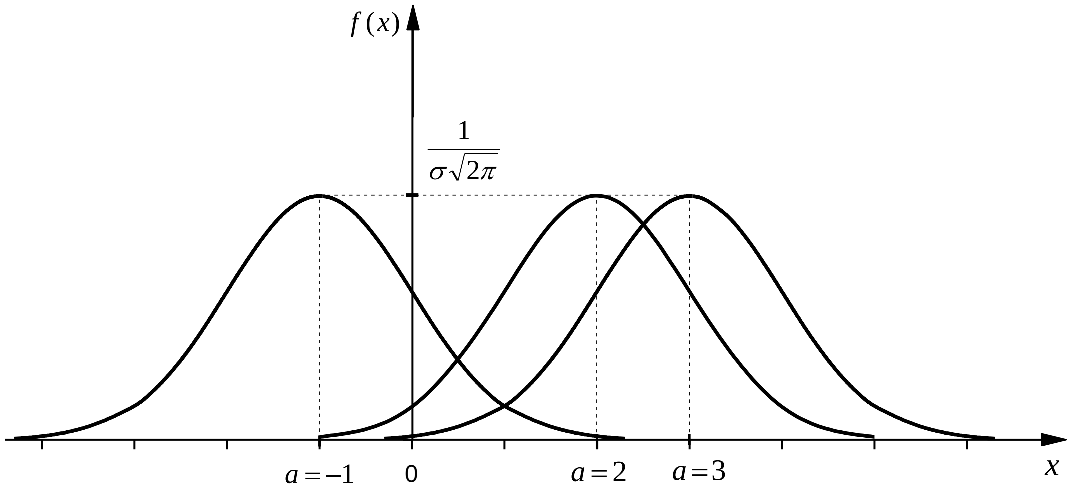


Рис. 4.5

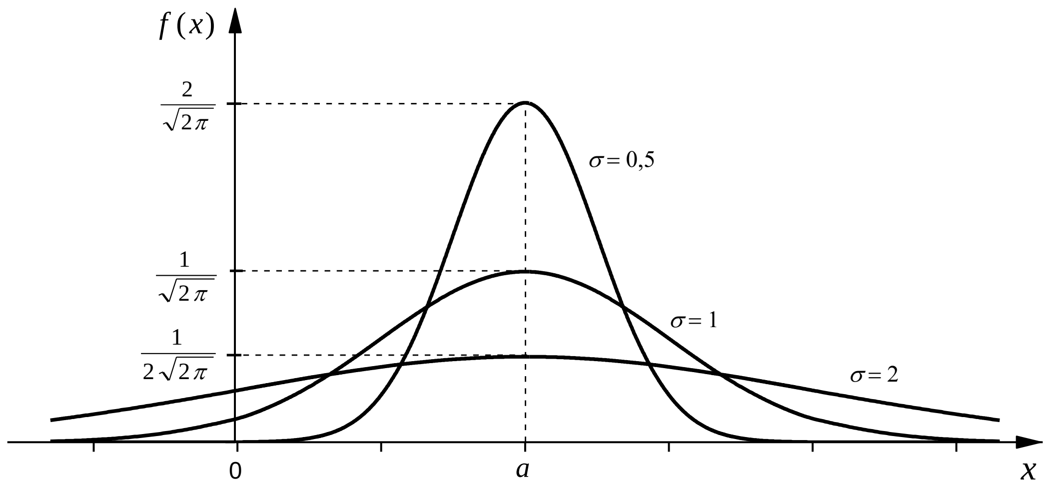


Рис. 4.6

Інтегральна функція нормального розподілу має вигляд

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{(z-a)^2}{2\sigma^2}}\,dz\,, \qquad x\in R. \tag{4.9}$$

Приклад. Коливання часу T, *хв*, проходження поїздом перегону підлягає нормальному закону розподілу із середнім значенням $M(T)=15$ *хв* і $\sigma=2$ *хв*. Записати вирази для диференціальної та інтегральної функцій, побудувати графік щільності розподілу.

<u>Розв'язання.</u> НВВ T розподілена за законом $N(a;\sigma) = N(15;2)$. За формулами (4.8), (4.9) обчислимо :

$$f(t) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(t-15)^2}{8}}, \qquad F(t) = \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{(z-15)^2}{8}} dz .$$

Графік нормальної кривої в даному випадку зображений на рис. 4.7.

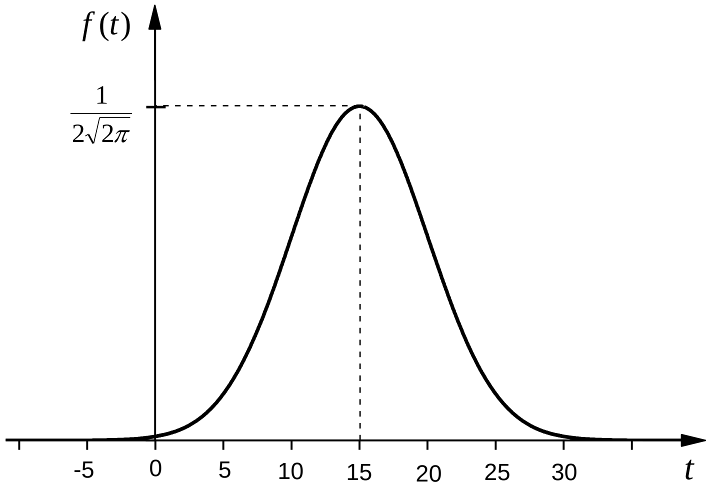


Рис. 4.7

Особливе значення має нормальний розподіл $N(0;1)$ з параметрами $a=0$, $\sigma=1$, який називається *нормованим нормальним розподілом.* У цьому випадку щільністю розподілу є *функція Гаусса*

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in R, \tag{4.10}$$

графік якої зображено на рис. 4.8. Нагадаємо (див. розд.2), що значення $\varphi(x)$ знаходять за спеціальними таблицями (дод. 1), враховуючи її властивості.

Властивості функції Гаусса $\varphi(x)$:

1) $\varphi(x)$ визначена і додатна на всій осі;

2) $\varphi(x)$ є парною функцією: $\varphi(-x) = \varphi(x)$;

3) $\varphi(x) \to 0$ при $x \to \pm\infty$.

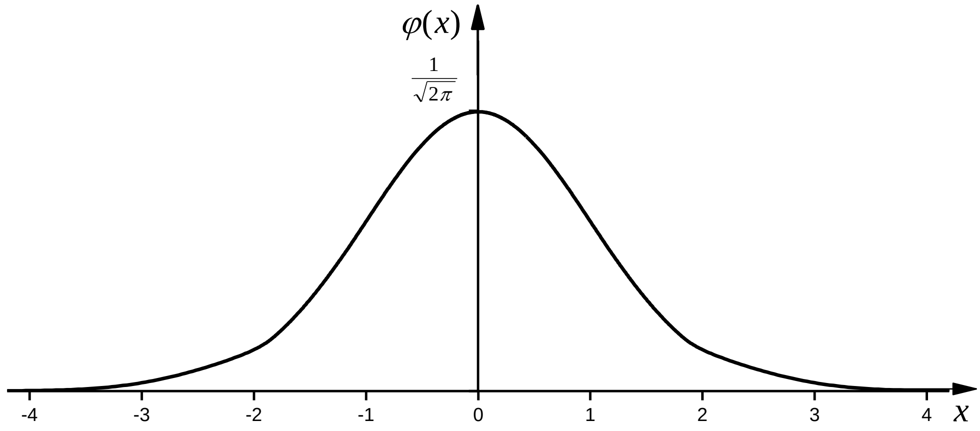


Рис. 4.8. Графік функції Гауса

Інтегральна функція нормованого нормального розподілу має вигляд

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz \;, \qquad x \in R.$$

Як щільність, функція Гаусса задовольняє умову $\int\limits_{-\infty}^{+\infty} \varphi(x)dx = 1.$ Крім того, вона є парною: $\varphi(-x) = \varphi(x)$. З цього випливає, що

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{z^2}{2}} dz = \int_{-\infty}^{0} \varphi(x)dx = \frac{1}{2} \int_{-\infty}^{+\infty} \varphi(x)dx = 0{,}5.$$

Як наслідок, отримуємо вираз для функції розподілу $N(0;1)$:

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{z^2}{2}} dz = 0{,}5 + \Phi(x),$$

де

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{z^2}{2}} dz \;, \qquad x \in R \tag{4.11}$$

– *функція Лапласа*, графік якої зображено на рис. 4.9. Значення $\Phi(x)$ знаходять за спеціальними таблицями (дод. 2), враховуючи її властивості.

Властивості функції Лапласа $\Phi(x)$:

1) $\Phi(x) \in (-0{,}5;0{,}5)$;
2) $\Phi(x)$ є непарною функцією: $\Phi(-x) = -\Phi(x)$;
3) $\Phi(x) \to \pm 0{,}5$ при $x \to \pm\infty$.

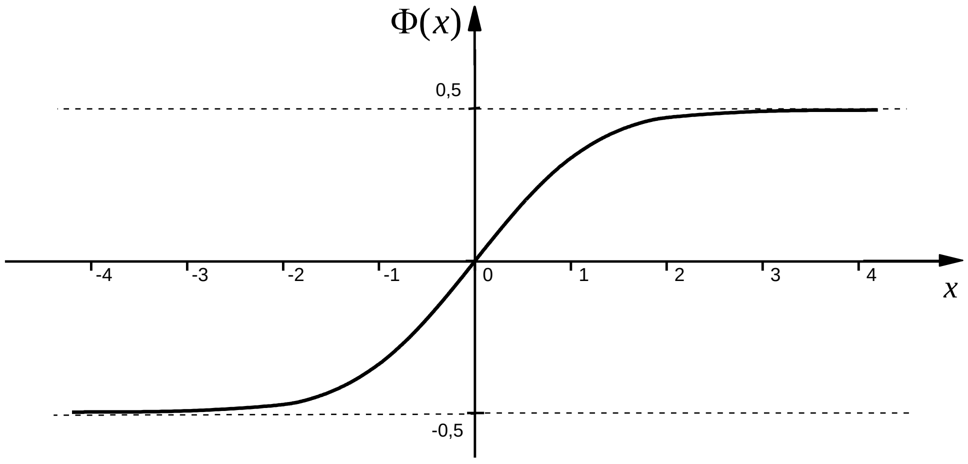


Рис.4.9. Графік функції Лапласа

Імовірність попадання нормально розподіленої випадкової величини X в інтервал $(\alpha;\beta)$:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right), \qquad (4.12)$$

де a і σ − параметри розподілу X, $\Phi(x)$ − функція Лапласа.

Приклад. Число вагонів у поїзді, що прибуває на розформування, є випадковою величиною, що розподілена за нормальним законом з параметрами $a = 50$, $\sigma = 5$. Виписати функцію щільності розподілу та обчислити ймовірності того, що величина рухомого складу: а) не перевищить 55 *ваг*; б) складатиме від 40 до 50 *ваг*; в) перевищить 53 *ваг*.

Розв'язання. Щільність нормального розподілу $N(a;\sigma) = N(50;5)$ має вигляд

$$f(x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-50)^2}{50}}, \quad x \in R.$$

Потрібні ймовірності знаходимо за формулою (4.12):

а) $P(X \le 55) = P(0 \le X \le 55) = \Phi\left(\dfrac{55-50}{5}\right) - \Phi\left(\dfrac{0-50}{5}\right) =$

$$= \Phi(1) - \Phi(-10) = 0{,}3413 + 0{,}5 = 0{,}8413;$$

б) $P(40 \le X \le 50) = \Phi\left(\dfrac{50-50}{5}\right) - \Phi\left(\dfrac{40-50}{5}\right) = \Phi(0) - \Phi(-2) = 0{,}4773;$

в) $P(X > 53) = P(53 < X < +\infty) = \Phi(+\infty) - \Phi\left(\dfrac{53-50}{5}\right) =$

$$= 0{,}5 - \Phi(0{,}6) = 0{,}5 - 0{,}2258 = 0{,}2742.$$

Імовірність того, що величина $|X-a|$ відхилення нормально розподіленої випадкової величини X від її математичного сподівання a не перебільшує δ, знаходиться за формулою

$$\underline{P(|X-a|<\delta) = 2\Phi\left(\frac{\delta}{\sigma}\right).} \qquad (4.13)$$

У частинному випадку $\delta = 3\sigma$ отримаємо рівність:

$$\underline{P(|X-a|<3\sigma) = 2\Phi(3) \approx 0{,}9973,} \qquad (4.14)$$

сенс якої полягає в тому, що протилежне є практично неможливим. Співвідношення (4.14) називається *правилом трьох сигм:* якщо НВВ X розподілена нормально, то абсолютна величина відхилення X від її математичного сподівання не перебільшує 3σ.

Приклад. Вважається, що відхилення довжини виробу від стандарту є випадковою величиною, що розподілена за нормальним законом. Стандартна довжина дорівнює $a = 50$ *см*, середнє квадратичне відхилення – $\sigma = 0{,}5$ *см*.

Знайти ймовірність того, що відхилення від стандарту: а) не перевищить 0,7 *см*; б) перевищить 1 *см*. Яку точність довжини виготовленого виробу можна гарантувати з імовірністю 0,9?

<u>Розв'язання.</u> Для знаходження потрібних ймовірностей скористаємось формулою (4.13):

а) $P\left(|X-50|<0{,}7\right)=2\Phi\left(\dfrac{0{,}7}{0{,}5}\right)=2\Phi(1{,}4)\approx 2\cdot 0{,}4192=0{,}8384$;

б) $P\left(|X-50|>1\right)=1-P\left(|X-50|<1\right)=1-2\Phi\left(\dfrac{1}{0{,}5}\right)=1-2\Phi(2)\approx$

$\approx 1-0{,}9544=0{,}0456.$

Потрібно знайти допустиму похибку $\delta>0$ таку, що:

$$P\left(|X-50|<\delta\right)=0{,}9.$$

За формулою (4.13) отримаємо:

$$2\Phi\left(\frac{\delta}{0{,}5}\right)=2\Phi(2\cdot\delta)=0{,}9;$$

$$\Phi(2\cdot\delta)=0{,}45.$$

За дод.2 знаходимо відповідний аргумент функції Лапласа:

$$2\cdot\delta=1{,}65 \quad\Rightarrow\quad \delta=0{,}825.$$

Таким чином, з імовірністю 0,9 можна гарантувати, що похибка довжини виробу не перевищить $\delta=0{,}825$.

Згідно з означеннями мода та медіана нормально розподіленої випадкової величини співпадають з математичним сподіванням: $M_0=Me=a$. Асиметрія та ексцес нормального розподілу дорівнюють нулю, тому чим більші асиметрія і ексцес теоретичного розподілу, тим більше він відрізняється від нормального.

Основні властивості нормального розподілу наведено у табл. 4.7.

Таблиця 4.7

Щільність розподілу $N(a;\sigma)$	$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$
Функція розподілу $N(a;\sigma)$	$F(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot\int_{-\infty}^{x} e^{-\frac{(z-a)^2}{2\sigma^2}}dz$
Параметри	$a\in R,\ \sigma>0$
Математичне сподівання	$M(X)=a$
Дисперсія	$D(X)=\sigma^2$
Середнє квадратичне відхилення	$\sigma(X)=\sigma$
Імовірність попадання в $(\alpha;\beta)$	$P(\alpha<X<\beta)=\Phi\left(\dfrac{\beta-a}{\sigma}\right)-\Phi\left(\dfrac{\alpha-a}{\sigma}\right)$
Оцінка для відхилення	$P(\lvert X-a\rvert<\delta)=2\Phi\left(\dfrac{\delta}{\sigma}\right)$
Правило трьох сигм	$P(\lvert X-a\rvert<3\sigma)=0{,}9973$
Мода	$M_0=a$
Медіана	$Me=a$
Асиметрія	$A_s=0$
Ексцес	$E_k=0$

Завдання

1. За допомогою методів диференціального числення провести повне дослідження функції щільності нормального розподілу та впевнитись у справедливості властивостей нормальної кривої.

2. Час формування потягів T підлягає нормальному закону розподілу, заданому графіком щільності розподілу (рис. 4.10).

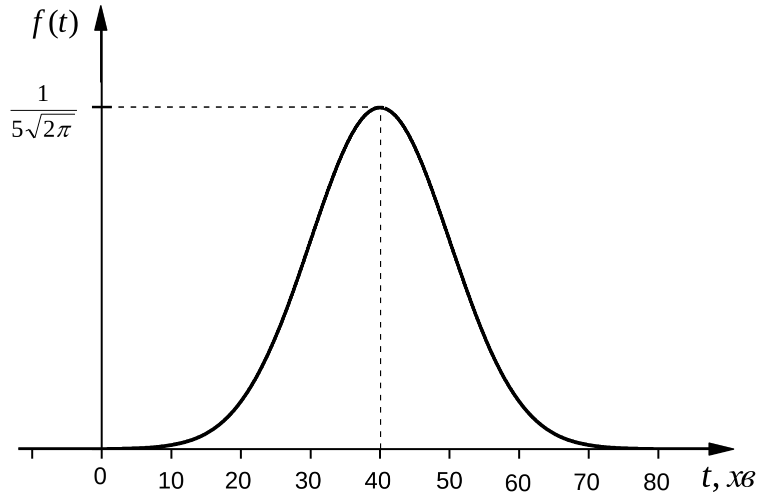


Рис. 4.10

Який закон розподілу має НВВ T? Записати аналітичний вираз функції щільності та обчислити ймовірність часу формування від 35 до 45 *хв*.

Відповідь. $N(5;40)$. $f(t) = \frac{1}{5\sqrt{2\pi}} \cdot e^{-\frac{(t-40)^2}{50}}$. $P(35 < T < 45) \approx 0{,}6826$.

3. Вага вагонів, т, що прибувають на вантажний двір, є випадковою величиною Q, розподіл якої задано щільністю:

$$f(q) = \frac{1}{4\sqrt{2\pi}} \cdot e^{-\frac{(q-65)^2}{32}}.$$

Визначити числові характеристики випадкової величини Q, обчислити ймовірність появи вагонів вагою менше 60 т.

Відповідь. $M(Q) = 65$т; $\sigma(Q) = 4$т; $D(Q) = 16$; $P(Q < 60) \approx 0{,}12$.

4. Період накопичення поїзда на сортувальній колії − випадкова величина T, що розподілена за нормальним законом. Середній час накопичення дорівнює $M(T) = 6$ *год*, $\sigma = 1$ *год*. Виписати щільність розподілу та знайти ймовірність того, що період накопичення перевищить 7 *год*.

<u>Відповідь.</u> $f(t)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(t-6)^2}{2}}$, $P(T>7)\approx 0{,}1587$.

5. Розподіл погонних навантажень P потягів підлягає нормальному закону. Середнє значення погонного навантаження дорівнює $M(P)=4$ т/пог.м, середнє квадратичне відхилення − $\sigma=0{,}6$ т/пог.м. Виписати щільність розподілу та визначити ймовірність того, що погонне навантаження буде більше 4,5 т/пог.м.

<u>Відповідь.</u> $f(p)=\frac{5}{3\sqrt{2\pi}}\cdot e^{-\frac{(p-4)^2}{0{,}72}}$; $P(P>4{,}5)\approx 0{,}2033$.

6. Тривалість непродуктивних затримок, *год*, при перевезеннях на даній ділянці залізниці підлягає нормальному закону $N(2{,}5;0{,}8)$. Виписати щільність розподілу та обчислити ймовірність, що непродуктивна затримка не перевищить 3 *год*.

<u>Відповідь.</u> $f(t)=\frac{5}{4\sqrt{2\pi}}\cdot e^{-\frac{(t-2{,}5)^2}{1{,}28}}$, $P(T\le 3)\approx 0{,}735$.

7. Число вагонів, що прибувають на промислову станцію, має нормальний розподіл із середнім квадратичним відхиленням $\sigma=3$ і середнім значенням $a=32$ *ваг/доб.* Обчислити ймовірність прибуття на станцію від 30 до 35 *ваг*.

<u>Відповідь.</u> $P(30<T<35)\approx 0{,}5867$.

8. Помилка X вимірювання ваги вантажу розподілена за нормальним законом з $a=0$ і середньою квадратичною похибкою $\sigma(X)=100$ г. Знайти ймовірність того, що: а) помилка вимірювання не перевищить 50 г; б) вимірювання буде виконано з помилкою, більшою за 70 г.

<u>Відповідь:</u> а) $P\left(|X|<50\right)\approx 0{,}383$; б) $P\left(|X|>70\right)\approx 0{,}484$.

9. Діаметр деталі − випадкова величина, розподілена за нормальним законом. Стандартний діаметр деталі дорівнює $a=15$ см, а середнє квадратичне відхилення − $\sigma=0{,}2$ см. Яка

ймовірність того, що у навмання перевіреної деталі відхилення від стандартного діаметра не перевищить 0,15 см? Яку точність виготовлення можна гарантувати з імовірністю 0,8?

Відповідь. $P\left(|X-15|<0{,}15\right)\approx 0{,}5468$.

$$\delta=0{,}258\ \left(P\left(|X-15|<0{,}258\right)=0{,}8\right).$$

4.8. Закон розподілу Ерланга. Потоки Ерланга

Розподілом Ерланга порядку k називається розподіл суми k незалежних випадкових величин, кожна з яких розподілена за показниковим законом з (однаковим) параметром λ.

Закон розподілу Ерланга порядку k визначається щільністю розподілу

$$f(x)=\begin{cases}0 & \text{при } \ x<0,\\ \dfrac{\lambda^k}{(k-1)!}\cdot x^{k-1}\cdot e^{-\lambda x} & \text{при } \ x\geq 0;\end{cases} \tag{4.15}$$

параметрами розподілу Ерланга є інтенсивність (показникової складової) $\lambda>0$, кількість складових $k\in N$.

Зауваження. При $k=1$ розподіл Ерланга співпадає з показниковим. При великому значенні k розподіл Ерланга наближається до нормального і може бути заміненим на останній. Нарешті, розподіл Ерланга є частинним випадком гамма-розподілу (див. п. 4.9).

Функцію розподілу Ерланга порядку k знаходять за формулою

$$F(x)=\begin{cases}0 & \text{при } \ x<0,\\ \displaystyle\int_0^x \frac{\lambda^k}{(k-1)!}\cdot z^{k-1}\cdot e^{-\lambda z}dz & \text{при } \ x\geq 0,\end{cases} \tag{4.16}$$

а числові характеристики − за формулами:

$$M(X)=\frac{k}{\lambda};\quad D(X)=\frac{k}{\lambda^2};\quad \sigma(X)=\frac{\sqrt{k}}{\lambda};\quad V(X)=\frac{1}{\sqrt{k}}\le 1. \qquad (4.17)$$

Найважливішим застосуванням розподілу Ерланга є моделювання потоків Ерланга, що утворюються шляхом просіювання найпростішого (пуассонівського) потоку (див. п. 4.2).

Потоком Ерланга порядку k називається потік, що утворюється з найпростішого потоку зберіганням кожної k-ї точки та відкиданням решти.

Очевидно, що найпростіший потік можна розглядати як потік Ерланга 1-го порядку. Параметр $k \in N$ служить мірою післядії: при $k=1$ післядія повністю відсутня (найпростіший потік), а при $k \to \infty$ потік Ерланга наближається до регулярного потоку зі сталими інтервалами.

При розв'язанні практичних задач зазвичай переходять до нормованої випадкової величини $T=\frac{X}{k}$, тобто від інтенсивності складової λ до інтенсивності ерлангівського потоку в цілому $\mu=\frac{\lambda}{k}$. Отриманий закон називається *нормованим розподілом Ерланга порядку* k; його характеристики нескладно отримати з (4.15), (4.16) відповідною заміною.

Дані про нормований розподіл Ерланга порядку k наведемо в табл.4.8 (μ − інтенсивність ерлангівського потоку в цілому).

Для отримання виразу функції розподілу Ерланга порядку k доводиться обчислювати відповідний інтеграл методом інтегрування частинами. Кілька найбільш поширених прикладів функції розподілу та щільності розподілу Ерланга порядку k з інтенсивністю μ наведені в табл. 4.9.

Графіки щільності розподілу Ерланга при фіксованому значенні $\mu=1$ та різних значеннях параметра k зображено на рис. 4.11.

Таблиця 4.8

Щільність нормованого розподілу Ерланга	$f(t)=\begin{cases}0 & \text{при } t<0,\\ \dfrac{(\mu k)^k}{(k-1)!}\cdot t^{k-1}\cdot e^{-\mu kt} & \text{при } t\ge 0.\end{cases}$
Функція розподілу	$F(t)=\begin{cases}0 & \text{при } t<0,\\ \int\limits_0^t \dfrac{(\mu k)^k}{(k-1)!}\cdot \tau^{k-1}\cdot e^{-\mu k\tau} & \text{при } t\ge 0.\end{cases}$
Параметри	$\mu>0;\ k\in N$
Математичне сподівання	$M(T)=\dfrac{1}{\mu}$
Дисперсія	$D(T)=\dfrac{1}{\mu^2\cdot k}$
Середнє квадратичне відхилення	$\sigma(T)=\dfrac{1}{\mu\cdot\sqrt{k}}$
Коефіцієнт варіації	$V(T)=\dfrac{1}{\sqrt{k}}\le 1$

Таблиця 4.9

k	Щільність розподілу $f(t)\quad(t\ge 0)$	Функція розподілу $F(t)\quad(t\ge 0)$
1	$f(t)=\mu\cdot e^{-\mu t}$	$F(t)=1-e^{-\mu t}$
2	$f(t)=4\mu^2\cdot t\cdot e^{-2\mu t}$	$F(t)=1-\left(1+2\mu t\right)\cdot e^{-2\mu t}$
3	$f(t)=\dfrac{27}{2}\cdot\mu^3\cdot t^2\cdot e^{-3\mu t}$	$F(t)=1-\dfrac{1}{2}\cdot\left(9\mu^2t^2+6\mu t+2\right)\cdot e^{-3\mu t}$
4	$f(t)=\dfrac{128}{3}\cdot\mu^4\cdot t^3\cdot e^{-4\mu t}$	$F(t)=1-\dfrac{1}{3}\cdot\left(32\mu^3t^3+24\mu^2t^2+12\mu t+3\right)\cdot e^{-4\mu t}$

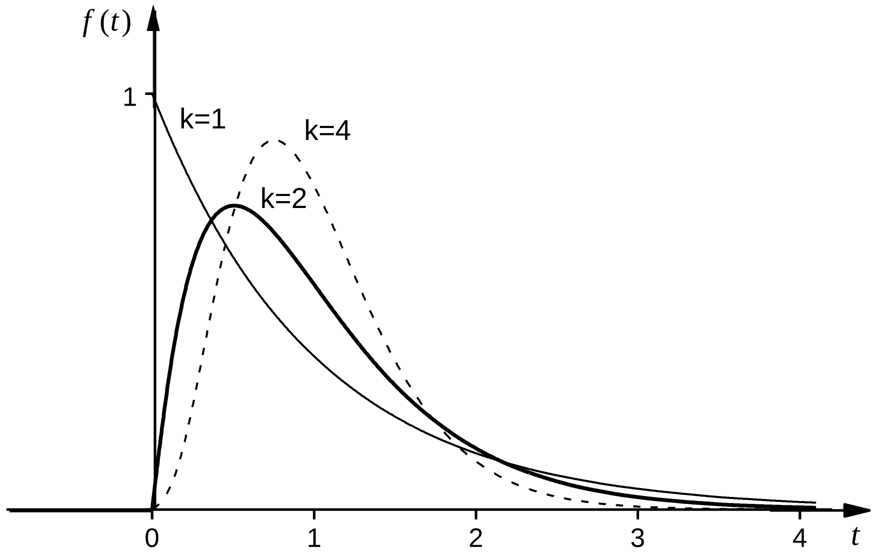


Рис. 4.11. Щільність розподілу Ерланга

Графіки функції розподілу Ерланга при фіксованому значенні $\mu = 1$ та різних значеннях параметра k зображено на рис. 4.12.

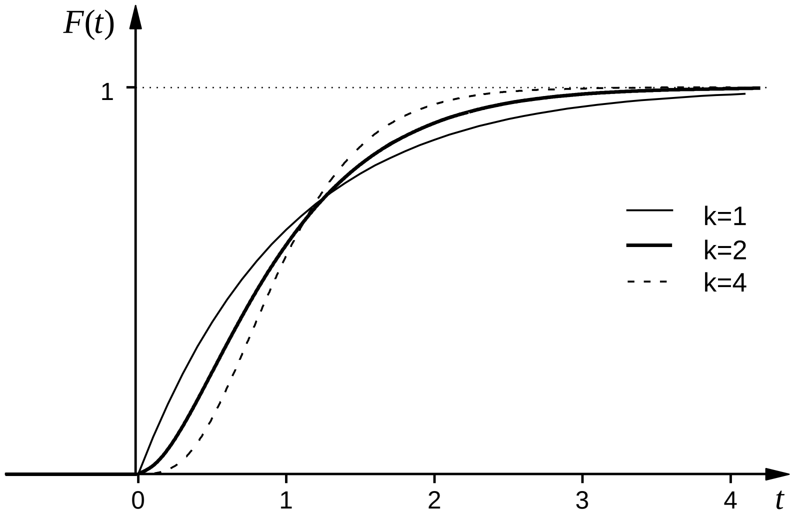


Рис. 4.12. Функція розподілу Ерланга

Приклад. Час простою T вагонів під накопиченням на сортувальній станції розподілений за законом Ерланга 3-го порядку. Середній час простою складає 3,5 *год*. Виписати вираз для щільності розподілу та функції розподілу, обчислити числові характеристики випадкової величини T. Знайти ймовірність того, що час простою: а) не перевищить 4 *год*; б) триватиме від 3 до 4 *год*; в) перевищить 5 *год*.

Розв’язання. За умовою математичне сподівання дорівнює $M(T)=3{,}5\,год$, звідси знаходимо параметр $\mu=\dfrac{1}{M(T)}=\dfrac{1}{3{,}5}\approx 0{,}288$ та обчислюємо інші числові характеристики:

$$D(T)=\frac{1}{(0{,}288)^2\cdot 3}\approx 4{,}083;\quad \sigma(T)=\sqrt{4{,}083}\approx 2{,}021;\quad V(T)=\frac{1}{\sqrt{3}}\approx 0{,}577\,.$$

Щільність розподілу часу простою має вигляд:

$$f(t)=\frac{(0{,}288\cdot 3)^3}{2}\cdot t^2\cdot e^{-0{,}288\cdot 3\cdot t}=0{,}322\cdot t^2\cdot e^{-0{,}864\cdot t}\quad (t\ge 0),$$

а функція розподілу:

$$F(t)=1-\frac{1}{2}\cdot\left(9\cdot(0{,}288)^2\cdot t^2+6\cdot 0{,}288\cdot t+2\right)\cdot e^{-3\cdot 0{,}288\cdot t}=$$
$$=1-\left(0{,}373\cdot t^2+0{,}864\cdot t+1\right)\cdot e^{-0{,}864\cdot t}\qquad (t\ge 0).$$

Імовірність того, що час простою не перевищить 4 *год*, дорівнює:

$$P(T\le 4)=F(4)=1-\left(0{,}373\cdot 4^2+0{,}864\cdot 4+1\right)\cdot e^{-0{,}864\cdot 4}\approx 0{,}671\,.$$

Імовірність того, що час простою триватиме від 3 до 4 *год*:

$$P(3\le T\le 4)=F(4)-F(3)=$$
$$=\left(0{,}373\cdot 3^2+0{,}864\cdot 3+1\right)\cdot e^{-0{,}864\cdot 3}-\left(0{,}373\cdot 4^2+0{,}864\cdot 4+1\right)\cdot e^{-0{,}864\cdot 4}\approx$$
$$\approx 0{,}192.$$

Імовірність того, що час простою перевищить 5 *год*:

$$P(T>5)=F(+\infty)-F(5)=1-\left(1-(0{,}373\cdot 5^2+0{,}864\cdot 5+1)\cdot e^{-0{,}864\cdot 5}\right)\approx 0{,}195\,.$$

Завдання

1. Довести формули табл. 4.9 для щільності та функції розподілу Ерланга в частинних випадках.

2. Маса поїзд, т, що відправляється зі станції даного вузла, є випадковою величиною Q, розподіленою за законом Ерланга 2-го порядку з параметром $\mu=0{,}001$. Необхідно: виписати функції щільності та розподілу; знайти середню масу $M(Q)$ поїзда та

середнє квадратичне відхилення $\sigma(Q)$; обчислити ймовірність того, що маса поїзда не перевищить 2000 тон.

Відповідь. $f(q) = 0{,}000004 \cdot q \cdot e^{-0{,}002 \cdot q}$, $F(q) = 1 - (1 + 0{,}002q) \cdot e^{-0{,}002 \cdot q}$ $(q \geq 0)$; $M(Q) = 1000$ т, $\sigma(Q) = 707{,}107$ т; $P(Q \leq 2000) = F(2000) \approx 0{,}91$.

3. Час T, *год*, роботи маневрових локомотивів на позиції 1 контролера машиніста розподілений за законом Ерланга 3-го порядку. Середній час роботи складає 0,8 *год*. Виписати функцію щільності розподілу та функцію розподілу, числові характеристики випадкової величини T. Знайти ймовірність того, що час роботи на позиції 1 триватиме від 30 *хв* до 1 *год*.

Відповідь. $f(t) = 26{,}337 \cdot t^2 \cdot e^{-3{,}75 \cdot t}$;
$F(t) = 1 - (7{,}03 \cdot t^2 + 3{,}75 \cdot t + 1) \cdot e^{-3{,}75 \cdot t}$.
$M(T) = 0{,}8$; $D(T) \approx 0{,}21$; $\sigma(T) \approx 0{,}46$; $V(T) \approx 0{,}707$. $P(0{,}5 \leq T \leq 1) \approx 0{,}434$.

4. Час обслуговування поїзда бригадою оглядачів у парку відправлення сортувальної станції розподілений за законом Ерланга з параметром $k = 4$. Середній технологічний час огляду поїзда складає 33 *хв*. Виписати функції щільності та розподілу. Обчислити ймовірність того, що час огляду не перевищить 45 *хв*.

Відповідь. $f(t) = \dfrac{128}{3} \cdot (1{,}82)^4 \cdot t^3 \cdot e^{-4 \cdot 1{,}82 \cdot t} = 468{,}14 \cdot t^3 \cdot e^{-7{,}28 \cdot t}$ $(t \geq 0)$;
$F(t) = 1 - (64{,}3 \cdot t^3 + 26{,}5 \cdot t^2 + 7{,}28 \cdot t + 1) \cdot e^{-7{,}28 \cdot t}$ $(t \geq 0)$; $P(T \leq 0{,}75) \approx 0{,}79$.

5. Час формування T поїзда підлягає закону розподілу Ерланга 2-го порядку із середньою інтенсивністю $\mu = 3$ *потяг/год*. Записати вираз для щільності та функції розподілу, знайти числові характеристики випадкової величини T. Яка ймовірність того, що час формування: а) перевищить 30 *хв*; б) буде тривати від 15 до 30 *хв*?

Відповідь. $f(t) = 36 \cdot t \cdot e^{-6t}$; $F(t) = 1 - (1 + 6t) \cdot e^{-6t}$ $(t \geq 0)$.
$M(T) = 20\, хв$, $D(T) \approx 3{,}33\, хв$, $\sigma(T) \approx 1{,}142\, хв$, $V(T) \approx 0{,}707$.
$P(T > 0{,}5) \approx 0{,}2$; $P(0{,}25 \leq T \leq 0{,}5) \approx 0{,}36$.

6. Розподіл інтервалів між поїздами, що прибувають на станцію на розформування, підлягає ерлангівському розподілу з параметром $k = 2$. Гірковий технологічний інтервал дорівнює

12 *хв*. Виписати вирази для щільності та функції розподілу. Визначити ймовірність того, що інтервал між поїздами, що прибувають, складатиме від 5 до 15 *хв*.

Відповідь. $f(t) = 100 \cdot t \cdot e^{-10t}$; $F(t) = 1 - (1 + 10t) \cdot e^{-10t}$ $(t \geq 0)$.

$P(0{,}08 \leq T \leq 0{,}25) \approx 0{,}49$.

4.9. Гамма-розподіл

Гамма-розподіл є найбільш загальним розподілом інтервалів, що включає в себе показниковий та ерлангівський розподіли.

Гамма-розподіл визначається щільністю

$$f(x) = \begin{cases} 0, & x < 0 \\ \dfrac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, & x \geq 0 \end{cases}, \tag{4.18}$$

де $\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx$ − гамма-функція Ейлера, числа $\alpha > 0$ і $\lambda > 0$ − *параметри гамма-розподілу.*

Зауваження 1. При $k \in N$ гамма-функція дорівнює $\Gamma(k) = (k-1)!$, тому в цьому випадку гамма-розподіл перетворюється на розподіл Ерланга. Нарешті, при $k = 1$ знову отримаємо показниковий розподіл.

Зауваження 2. Застосування гамма-розподілу доволі складне як для моделювання потоків, так і для побудови наближених аналітичних моделей. Тому відповідні потоки зазвичай апроксимуються розподілами Ерланга.

4.10. Розподіл χ^2

Цей розподіл відіграє важливу роль при розгляданні різних питань статистики.

Нехай X_i, $i = \overline{1,k}$ − незалежні випадкові величини, що розподілені за нормованим нормальним законом $N(0;1)$.

Законом розподілу χ^2 *(«хі квадрат») з* k *степенями вільності* називається розподіл суми квадратів $\chi^2 = \sum_{i=1}^{k} X_i$ нормально розподілених випадкових величин. Єдиним *параметром* цього розподілу є число k степенів вільності.

Щільність цього розподілу має вигляд

$$f(x) = \begin{cases} 0, & x \le 0; \\ \dfrac{1}{2^{k/2} \cdot \Gamma\left(\dfrac{k}{2}\right)} \cdot x^{\frac{k}{2}-1} \cdot e^{-x/2}, & x > 0, \end{cases} \qquad (4.19)$$

де $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$ − гамма-функція. Чим більше число степенів вільності, тим ближче розподіл χ^2 до нормального.

Зауваження. При наявності (лінійних) співвідношень між випадковими величинами X_i число степенів вільності зменшується. Так, в статистиці ми розглядатимемо випадок $\sum_{i=1}^{n} X_i = n\overline{X}$, при цьому число степенів вільності дорівнює $k = n - 1$.

4.11. Розподіл Стьюдента

Нехай Z − нормально розподілена випадкова величина з $a = 0$, $\sigma = 1$ (нормована), а V − незалежна від Z випадкова величина, що розподілена за законом χ^2 з k степенями вільності.

Розподіл випадкової величини

$$T = \frac{Z}{\sqrt{V/k}} \qquad (4.20)$$

називається t−розподілом або *розподілом Стьюдента з* k *степенями вільності*.

При збільшенні числа k степенів вільності розподіл Стьюдента швидко наближається до нормального.

Щільність розподілу Стьюдента

$$S_k(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k}\cdot\Gamma\left(\frac{k}{2}\right)}\cdot\left[1+\frac{x^2}{k}\right]^{-\frac{k+1}{2}}, \qquad x \in R \qquad (4.21)$$

де $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1}e^{-x}dx$ − гамма-функція.

Випадкові величини виду (4.20) часто використовуються в математичній статистиці.

Питання до теми

1. Які розподіли випадкових величин ви знаєте? Які з них належать до дискретних, які – до неперервних?
2. Який розподіл імовірностей служить моделлю схеми Бернуллі? Випишіть формулу, що задає відповідний розподіл. Наведіть приклади застосування.
3. Скільки параметрів має біноміальний закон розподілу, який зміст вони мають?
4. Якого виду набувають формули для обчислення числових характеристик біноміального розподілу?
5. Які випадкові величини розподілені за законом Пуассона? Яка формула задає цей розподіл?
6. Скільки параметрів має розподіл Пуассона? Який його зміст?
7. Як знайти числові характеристики випадкової величини, розподіленої за законом Пуассона?
8. Що називають найпростішим потоком випадкових подій? Як ще називається такий потік і чому? Які ознаки найпростішого потоку вам відомі?
9. Що описує та якою формулою задається геометричний розподіл? Наведіть приклади.
10. Що можна сказати про множину значень геометрично розподіленої випадкової величини?
11. Як обчислюються числові характеристики випадкової величини, розподіленої за геометричним законом?
12. Який розподіл називається гіпергеометричним? Наведіть формулу розподілу ймовірностей та приклади її застосування.

13. Які параметри гіпергеометричного розподілу? Як за ними знайти математичне сподівання і дисперсію?

14. Як визначається рівномірний розподіл? Які числа є його параметрами? Наведіть приклади.

15. Як знайти числові характеристики випадкової величини, розподіленої за рівномірним законом? Чи може математичне сподівання бути від'ємним при рівномірному розподілі?

16. Як обчислити ймовірність попадання в заданий інтервал рівномірно розподіленої випадкової величини? Наведіть приклади використання різних формул (спеціальної для рівномірного розподілу та загальної для неперервних випадкових величин).

17. Який розподіл називається показниковим? Скільки параметрів він має? Як позначають експоненціальний закон?

18. Чому дорівнюють числові характеристики рівномірно розподіленої випадкової величини? Які їх властивості найчастіше використовуються?

19. Чи може математичне сподівання бути від'ємним при експоненціальному розподілі?

20. Які характеристики не залежать від конкретного виду показникового розподілу?

21. Як визначається та для чого використовується функція надійності?

22. Яка випадкова величина називається нормально розподіленою?

23. Скільки параметрів має нормальний розподіл? Який їх імовірнісний зміст?

24. Як називають графік щільності нормального розподілу? Які його властивості ви знаєте?

25. Який з нормальних розподілів називається нормованим? Яка функція служить його щільністю?

26. Яке походження має правило трьох сигм? Який його сенс?

27. Чому дорівнюють асиметрія та ексцес нормального розподілу?

28. Що описує та як визначається розподіл Ерланга порядку k? Які його параметри? Як він пов'язаний з іншими розподілами (показниковим, нормальним, гамма-розподілом)?

29. Що називають нормованим розподілом Ерланга? Які його числові характеристики?

30. Як визначається гамма-розподіл? Які його частинні випадки ви знаєте?

31. Що описують розподіли χ^2 і Стьюдента? Що служить параметром цих розподілів? Де вони використовуються?

Тестові питання

1. Дискретна випадкова величина X розподілена за біноміальним законом з параметрами $n=10; p=0{,}4$. Чому дорівнює математичне сподівання?

А	Б	В	Г	Д
0,4	0,04	$\frac{1}{4}$	4	Інша відповідь

2. Монету кидають 8 раз. Чому дорівнює дисперсія числа випадіння гербів?

А	Б	В	Г	Д
2	4	6	8	Інша відповідь

3. Параметром розподілу Пуассона є число:

А	Б	В	Г	Д
$\lambda \in R$	$\lambda \in N$	$\lambda > 0$	$\lambda \geq 0$	Інша відповідь

4. Розподіл випадкової величини X задано формулою

$$P(X=k)=\frac{2^k}{k!}\cdot e^{-2}.$$

Чому дорівнює математичне сподівання X?

А	Б	В	Г	Д
$\frac{1}{2}$	2	-2	$2k$	Інша відповідь

5. Яке твердження є справедливим для випадкової величини X, розподіленої за законом Пуассона?

А	Б	В	Г	Д
$M(X)=\sigma(X)$	$\sigma(X)=V(X)$	$V(X)=M(X)$	$M(X)=D(X)$	Інша відповідь

6. Дискретна випадкова величина розподілена за законом: $P(X=k)=\frac{2}{3}\cdot\left(\frac{1}{3}\right)^{k-1}$, $k=1,2,3\ldots$ Цей закон має назву:

А	Б	В	Г	Д
біноміальний	Пуассона	геометричний	гіпергеометричний	Інша відповідь

7. Кількість перескладань, необхідних студенту N для отримання задовільної оцінки, розподілена за законом: $P(X=k)=0{,}4\cdot(0{,}6)^{k-1}$, $k=1,2,3\ldots$ Чому дорівнює середня кількість необхідних перескладань?

А	Б	В	Г	Д
1	2,5	2,4	0,4	Інша відповідь

8. У ящику знаходяться 9 пристроїв, третина з яких має прихований дефект. ДВВ X дорівнює числу пристроїв з прихованим дефектом серед п'яти навмання взятих. Які значення набуває ДВВ X?

А	Б	В	Г	Д
1,2,3,4,5	0,1,2,3,4,5	1,2,3	0,1,2,3	Інша відповідь

9. У подачі з 10 вагонів 4 потребують очищення. Яка ймовірність того, що серед трьох навмання перевірених вагонів лише один виявився таким, що потребує очищення?

А	Б	В	Г	Д
$\frac{C_4^1\cdot C_6^2}{C_{10}^3}$	$\frac{C_4^2\cdot C_6^1}{C_{10}^3}$	$\frac{C_4^3\cdot C_6^1}{C_{10}^4}$	$\frac{C_4^1\cdot C_6^3}{C_{10}^4}$	Інша відповідь

10. У групі 15 студентів, 8 з яких – дівчата. Який вигляд має розподіл кількості $X = m$ дівчат з 6 навмання обраних студентів?

А	Б	В	Г	Д
$\frac{C_6^m \cdot C_7^{8-m}}{C_{15}^8}$	$\frac{C_7^m \cdot C_8^{6-m}}{C_{15}^6}$	$\frac{C_8^m \cdot C_7^{6-m}}{C_{15}^6}$	$\frac{C_{15}^m \cdot C_8^{6-m}}{C_{23}^6}$	Інша відповідь

11. Випадкова величина розподілена за рівномірним законом на сегменті $[-2;4]$. Чому дорівнює математичне сподівання?

А	Б	В	Г	Д
0	1	2	3	Інша відповідь

12. Випадкова величина X має рівномірний закон розподілу на сегменті $[0;2]$. Чому дорівнює дисперсія?

А	Б	В	Г	Д
1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{12}$	Інша відповідь

13. Випадкова величина X розподілена за рівномірним законом $U(3;5)$. Чому дорівнює $P(1 < X < 2)$?

А	Б	В	Г	Д
0	1	$\frac{1}{2}$	2	Інша відповідь

14. Випадкова величина X розподілена за показниковим розподілом. Чому дорівнює функція розподілу при $x \geq 0$?

А	Б	В	Г	Д
0	$1 - e^{-\lambda x}$	$\lambda e^{-\lambda x}$	1	Інша відповідь

15. Випадкова величина розподілена за показниковим законом $E(\lambda)$. Чому дорівнює її коефіцієнт варіації?

А	Б	В	Г	Д
1	–1	λ	$\frac{1}{\lambda}$	Інша відповідь

16. Випадкова величина X розподілена за показниковим законом. Знайти її математичне сподівання, якщо відома щільність розподілу:

$$f(x)=\begin{cases}0 & \text{при} \quad x<0,\\ 5\cdot e^{-5x} & \text{при} \quad x\geq 0.\end{cases}$$

А	Б	В	Г	Д
$M(X)=5$	$M(X)=-5$	$M(X)=\frac{1}{5}$	$M(X)=\frac{1}{25}$	Інша відповідь

17. Випадкова величина розподілена за законом $E(0{,}1)$. Чому дорівнює її дисперсія?

А	Б	В	Г	Д
1	0,1	10	100	Інша відповідь

18. Випадкова величина X розподілена за експоненціальним законом. Знайти її середнє квадратичне відхилення, якщо функція розподілу має вигляд:

$$F(x)=\begin{cases}0 & \text{при} \quad x<0,\\ 1-e^{-4x} & \text{при} \quad x\geq 0.\end{cases}$$

А	Б	В	Г	Д
1	2	3	4	Інша відповідь

19. Щільність нормально розподіленої випадкової величини X має вигляд:

$$f(x)=\frac{1}{3\sqrt{2\pi}}\cdot e^{-\frac{(x-2)^2}{18}}, \quad -\infty<x<+\infty.$$

Чому дорівнюють її математичне сподівання та дисперсія?

А	Б	В	Г	Д
$M(X)=2$, $D(X)=3$	$M(X)=-2$, $D(X)=3$	$M(X)=2$, $D(X)=9$	$M(X)=-2$, $D(X)=9$	Інша відповідь

20. Який з наведених нормальних законів не має сенсу?

А	Б	В	Г	Д
$N(1;2)$	$N(1;-2)$	$N(-1;2)$	$N(0{,}1;0{,}2)$	Інша відповідь

21. Функція Лапласа є:

А	Б	В	Г	Д
сталою	додатною	парною	непарною	Інша відповідь

22. Який розподіл є розподілом Ерланга першого порядку?

А	Б	В	Г	Д
експонен-ціальний	нормальний	гамма-розподіл	Пуассона	Інша відповідь

23. Який розподіл є наближеним до розподілу Ерланга при великих k?

А	Б	В	Г	Д
експонен-ціальний	нормальний	гамма-розподіл	Пуассона	Інша відповідь

Відповіді: 1.Г; 2.А; 3.В; 4.Б; 5.Г; 6.В; 7.Б; 8.Г; 9.А; 10.В; 11.Б; 12.В; 13.А; 14.Б; 15.А; 16.В; 17.Г; 18.Д; 19.В; 20.Б; 21.Г; 22.А; 23.Б.

Розділ 5. Системи двох випадкових величин

5.1. Поняття багатовимірної випадкової величини

До цього часу ми розглядали випадкові величини, можливі значення яких складаються з одного числа. Такі випадкові величини називаються також одновимірними. Якщо випадкова величина визначається більше ніж одним числом, вона називається *багатовимірною*.

n*-вимірною випадковою величиною* $(X_1;\ldots;X_n)$ називається величина, значення якої визначаються набором з n чисел $(x_1;\ldots;x_n)$.

n-вимірна випадкова величина також називається *системою* n *випадкових величин* або $n-$*вимірним випадковим вектором.*

Кожна з випадкових величин $X_1,\ldots,X_n$ n-вимірної випадкової величини $(X_1;\ldots;X_n)$ називається її *складовою* або *компонентою.*

Багатовимірна випадкова величина є *дискретною,* якщо її компоненти є дискретними випадковими величинами. У *неперервної* багатовимірної випадкової величини компоненти неперервні.

Прикладами багатовимірних величин є:

1) координати $(x;y)$ точки площини – двовимірна випадкова величина;
2) координати $(x;y;z)$ точки простору − трихвимірна випадкова величина.

Ми розглянемо двовимірні випадкові величини.

Двовимірною випадковою величиною $(X;Y)$ називається величина, значення якої визначаються парою чисел $(x;y)$. Двовимірна випадкова величина також називається *системою двох випадкових величин або двовимірним випадковим вектором.*

5.2. Закон розподілу дискретної двовимірної випадкової величини

Двовимірна дискретна випадкова величина (ДДВВ) задається таблицею, в якій перелічені всі можливі пари чисел $(x_i;y_j)$ та відповідні ймовірності їх появи $p_{ij}=p(x_i;y_j)$ $(i=\overline{1,n};\ j=\overline{1,m})$.

Y	X			
	x_1	x_2	…	x_n
y_1	p_{11}	p_{21}	…	p_{n1}
y_2	p_{12}	p_{22}	…	p_{n2}
…	…	…	…	…
y_m	p_{1m}	p_{2m}	…	p_{nm}

Звичайно, ці ймовірності задовольняють умову $\sum\limits_{i,j} p_{ij} \equiv 1$.

За законом сумісного розподілу системи двох випадкових величин $(X;Y)$ можна знайти закони розподілів її складових X і Y.

Щоб знайти *закон розподілу компоненти* X ДДВВ $(X;Y)$ потрібно перелічити її значення x_i $(i=\overline{1,n})$, а ймовірності $p_i = p(x_i)$ цих значень знайти, сумуючи ймовірності в стовпчиках:

X	x_1	x_2	…	x_n
P	$p_{11}+p_{12}+\ldots+p_{1m}$	$p_{21}+p_{22}+\ldots+p_{2m}$	…	$p_{n1}+p_{n2}+\ldots+p_{nm}$

Аналогічно, щоб знайти *закон розподіл компоненти* Y, потрібно перелічити значення y_j $(j=\overline{1,m})$, а їх імовірності $p_j = p(y_j)$ знайти як суму в рядках:

Y	y_1	y_2	…	y_m
P	$p_{11}+p_{21}+\ldots+p_{n1}$	$p_{12}+p_{22}+\ldots+p_{n2}$	…	$p_{1m}+p_{2m}+\ldots+p_{nm}$

Числові характеристики компонент X і Y знаходять як звичайно для одновимірних дискретних випадкових величин, виходячи із законів розподілу X і Y (див. п.3.4).

Приклад. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X		
	-1	3	4
0	0,3	0,2	0,1
2	k	0,15	0,2

Знайти закони розподілів та числові характеристики її компонент.

Розв'язання. По-перше, за умовою нормування $\sum_{i,j} p_{ij} \equiv 1$ знайдемо невідому ймовірність появи пари чисел $(X = -1; Y = 2)$:

$$p(-1;2) = 1 - (0{,}3 + 0{,}2 + 0{,}1 + 0{,}15 + 0{,}2) = 0{,}05.$$

Тоді закон сумісного розподілу X і Y має вигляд:

Y	X		
	-1	3	4
0	0,3	0,2	0,1
2	0,05	0,15	0,2

Закон розподілу компоненти X (сумуємо ймовірності у стовпчиках):

X	-1	3	4
P	0,35	0,35	0,3

Закон розподілу компоненти Y (сумуємо ймовірності в рядках):

Y	0	2
P	0,6	0,4

Знаходимо математичне сподівання, дисперсію та середнє квадратичне відхилення одновимірних випадкових величин X та Y за формулами (3.5)–(3.8) розд. 3:

$$M(X) = (-1)\cdot 0{,}35 + 3\cdot 0{,}35 + 4\cdot 0{,}3 = 1{,}9;$$
$$D(X) = (-1)^2\cdot 0{,}35 + 3^2\cdot 0{,}35 + 4^2\cdot 0{,}3 - (1{,}9)^2 = 8{,}3 - 3{,}61 = 4{,}69;$$
$$\sigma(X) = \sqrt{4{,}69} \approx 2{,}17;$$
$$M(Y) = 0\cdot 0{,}6 + 2\cdot 0{,}4 = 0{,}8;$$
$$D(Y) = 0^2\cdot 0{,}6 + 2^2\cdot 0{,}4 - (0{,}8)^2 = 1{,}6 - 0{,}64 = 0{,}96;$$
$$\sigma(Y) = \sqrt{0{,}96} \approx 0{,}98.$$

Завдання

1. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X		
	1	3	6
2	0,1	k	0,1
10	0,05	0,39	0,06

Знайти закони розподілів та числові характеристики її компонент.

<u>Відповідь.</u> $k = 0{,}3$.

X	1	3	6
P	0,15	0,69	0,16

Y	2	10
P	0,5	0,5

$M(X) = 3{,}18$; $D(X) \approx 2{,}01$; $\sigma(X) \approx 1{,}42$. $M(Y) = 6$; $D(Y) = 16$; $\sigma(Y) = 4$.

2. Двовимірна випадкова величина $(X;Y)$ задана законом розподілу

Y	X		
	1	2	3
1	0,05	0,08	a
2	0,1	0,15	0,05
3	0,16	0,24	0,1

Знайти закони розподілів та числові характеристики її компонент.

<u>Відповідь.</u> $a = 0{,}07$.

X	1	2	3
P	0,31	0,47	0,22

Y	1	2	3
P	0,2	0,3	0,5

$M(X) = 1{,}91$; $D(X) \approx 0{,}52$; $\sigma(X) \approx 0{,}72$. $M(Y) = 2{,}3$; $D(Y) = 0{,}61$; $\sigma(Y) \approx 0{,}78$.

5.3. Функція розподілу двовимірної випадкової величини та її властивості

Двовимірна випадкова величина може бути задана своєю функцією розподілу.

Функцією розподілу двовимірної випадкової величини $(X;Y)$ називається функція двох змінних

$$\underline{F(x,y) = P(X < x; Y < y)}. \tag{5.1}$$

З геометричної точки зору значення функції розподілу $F(x,y)$ дорівнює ймовірності попадання значення випадкової величини в нескінченний (необмежений зліва і знизу) квадрант, права верхня вершина якого знаходиться в точці з координатами (x,y) (рис.5.1).

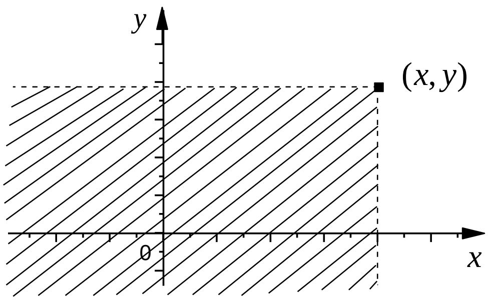


Рис. 5.1

Властивості функції розподілу $F(x,y)$:

1) $0 \le F(x,y) \le 1$;

2) $F(x,y)$ не спадає по обох аргументах:
якщо $x_2 > x_1$, то $F(x_2,y) \ge F(x_1,y)$; якщо $y_2 > y_1$, то $F(x,y_2) \ge F(x,y_1)$;

3) $F(x,y)$ задовольняє граничні умови:
а) $F(-\infty,-\infty) = 0$; $F(-\infty,y) = 0$; $F(x,-\infty) = 0$; $F(+\infty,+\infty) = 1$;
б) $F(x,+\infty) = F_1(x)$; $F(+\infty,y) = F_2(y)$,
тобто при $y = +\infty$ функція розподілу перетворюється на функцію розподілу $F_1(x)$ компоненти X, а при $x = +\infty$ − на функцію розподілу $F_2(y)$ компоненти Y;

4) імовірності попадання випадкової точки $(X;Y)$
а) в півсмугу $\{x_1 < X < x_2;\quad Y < y\}$:

$$P(x_1 \le X < x_2; y_1 \le Y < y_2) = F(x_2,y) - F(x_1,y);$$

б) в півсмугу $\{X < x;\quad y_1 < Y < y_2\}$:

$$P(x_1 \le X < x_2; y_1 \le Y < y_2) = F(x,y_2) - F(x,y_1);$$

в) в прямокутник $\{x_1 < X < x_2; \quad y_1 < Y < y_2\}$:

$$P(x_1 \le X < x_2; y_1 \le Y < y_2) = F(x_2, y_2) - F(x_1, y_2) - [F(x_2, y_1) - F(x_1, y_1)].$$

Наведене означення (5.1) і властивості функції розподілу $F(x, y)$ є спільними для дискретних і неперервних двовимірних випадкових величин.

5.4. Щільність розподілу ймовірностей неперервної двовимірної випадкової величини

Неперервна двовимірна випадкова величина (НДВВ) може бути заданою як своєю функцією розподілу, так і функцією щільності розподілу ймовірностей.

Щільністю сумісного розподілу ймовірностей НДВВ $(X; Y)$ називається друга мішана похідна функції розподілу:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \tag{5.2}$$

З геометричної точки зору функція $z = f(x, y)$ визначає деяку поверхню у просторі, яка називається *поверхнею розподілу.*

З означення щільності ймовірностей випливає, що функція розподілу $F(x, y)$ може бути відновленою за функцією $f(x, y)$:

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x, y) dx dy. \tag{5.3}$$

Приклад. НДВВ $(X; Y)$ задано функцією розподілу

$$F(x, y) = \begin{cases} \left(1 - e^{-x}\right) \cdot \left(1 - e^{-2y}\right) & \text{при} \quad x \ge 0, y \ge 0, \\ 0 & \text{при} \quad x < 0, y < 0. \end{cases}$$

Знайти щільність імовірностей $f(x, y)$.

Розв’язання. Скористаємось формулою (5.2). Спочатку обчислюємо частинну похідну по y:

$$\frac{\partial}{\partial y}F(x,y)=\left(1-e^{-x}\right)\cdot 2e^{-2y}, \qquad x\ge 0, y\ge 0,$$

Потім – частинну похідну по x від отриманого виразу:

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}=\frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}F(x,y)\right]=e^{-x}\cdot 2e^{-2y}=2e^{-x-2y}, \qquad x\ge 0, y\ge 0.$$

Перевіримо результат інтегруванням за формулою (5.3):

$$F(x,y)=\int\limits_0^y\int\limits_0^x 2e^{-x-2y}dxdy=2\int\limits_0^y e^{-2y}dy\int\limits_0^x e^{-x}dx=$$
$$=2\cdot\left(-\frac{1}{2}e^{-2y}+\frac{1}{2}e^0\right)\cdot\left(-e^{-x}+e^0\right)=\left(1-e^{-x}\right)\cdot\left(1-e^{-2y}\right), \qquad x\ge 0, y\ge 0.$$

Отже, щільність розподілу має вигляд:

$$f(x,y)=\begin{cases}2e^{-x-2y} & \text{при } x\ge 0, y\ge 0,\\ 0 & \text{при } x<0, y<0.\end{cases}$$

Властивості щільності ймовірностей $f(x,y)$:

1) $f(x,y)\ge 0$;

2) $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)dxdy\equiv 1$; (5.4)

3) імовірність попадання НДВВ $(X;Y)$ в область $D\subset R^2$ дорівнює

$$P((X,Y)\in D)=\iint\limits_D f(x,y)dxdy. \qquad (5.5)$$

Приклад. Система неперервних випадкових величин $(X;Y)$ розподілена в прямокутнику $D=\{(x;y):\ 0\le x\le 1; 0\le y\le 3\}$ з щільністю

$$f(x,y)=\begin{cases}axy & \text{при } (x;y)\in D,\\ 0 & \text{при } (x;y)\notin D.\end{cases}$$

Визначити параметр a та ймовірність попадання $(X;Y)$ в квадрат $K=\{(x;y):\ 0\le x\le 1; 0\le y\le 1\}$ (рис. 5.2).

Розв'язання. Побудуємо на площині область D, визначимо квадрат K (див. рис. 5.2).

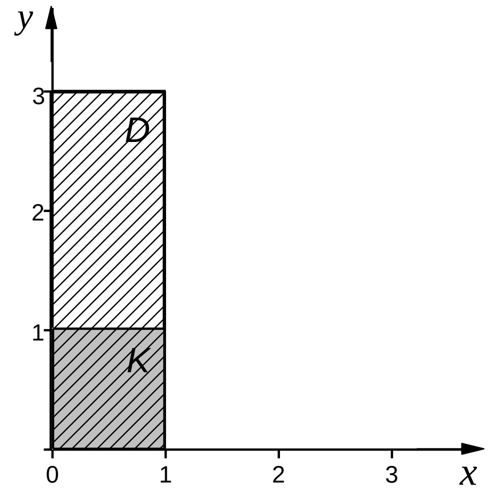


Рис. 5.2

Параметр a знаходимо з умови нормування (5.4):

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)dxdy = \iint\limits_{D} f(x;y)dxdy = \int\limits_{0}^{1}dx\int\limits_{0}^{3}axy\,dy = a\cdot\int\limits_{0}^{1}xdx\int\limits_{0}^{3}y\,dy = a\cdot\int\limits_{0}^{1}xdx\cdot\left(\frac{y^2}{2}\bigg|_0^3\right) =$$

$$= a\cdot\int\limits_{0}^{1}xdx\cdot\left(\frac{y^2}{2}\bigg|_0^3\right) = \frac{9}{2}a\cdot\int\limits_{0}^{1}xdx = \frac{9}{2}a\cdot\left(\frac{x^2}{2}\bigg|_0^1\right) = \frac{9}{4}a = 1 \qquad \Rightarrow \quad a = \frac{4}{9}.$$

Таким чином, щільність розподілу має вигляд:

$$f(x,y) = \begin{cases} \dfrac{4}{9}xy & \text{при } (x;y)\in D, \\ 0 & \text{при } (x;y)\notin D. \end{cases}$$

Імовірність попадання $(X;Y)$ в квадрат K обчислимо за формулою (5.5):

$$P((X,Y)\in K)=\iint\limits_K f(x,y)dxdy=\int_0^1 dx\int_0^1\left(\frac{4}{9}xy\right)dy=\frac{4}{9}\cdot\int_0^1 xdx\int_0^1 ydy=$$

$$=\frac{4}{9}\cdot\left(\left.\frac{x^2}{2}\right|_0^1\right)\cdot\left(\left.\frac{y^2}{2}\right|_0^1\right)=\frac{4}{9}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{9}.$$

Для дискретної двовимірної випадкової величини ми знаходили закони її компонент за законом сумісного розподілу. Аналогічно, для неперервної двовимірної випадкової величини за щільністю сумісного розподілу можна знайти щільності розподілів її компонент:

$$\underline{f_1(x)=\int_{-\infty}^{+\infty} f(x,y)dy} - \textit{щільність розподілу } X; \qquad (5.6)$$

$$\underline{f_2(y)=\int_{-\infty}^{+\infty} f(x,y)dx} - \textit{щільність розподілу } Y. \qquad (5.7)$$

<u>Зауваження.</u> Як і будь-які щільності розподілів імовірностей, щільності розподілів компонент задовольняють умови:

1) $f_1(x)\ge 0$; $\quad f_2(y)\ge 0$;

2) $\int_{-\infty}^{+\infty} f_1(x)dx\equiv 1$; $\quad \int_{-\infty}^{+\infty} f_2(y)dy\equiv 1$.

Приклад. Знайти щільності розподілів компонент неперервної двовимірної випадкової величини $(X;Y)$, заданої функцією $f(x,y)=2e^{-x-2y}, \quad x\ge 0, y\ge 0$.

<u>Розв’язання.</u> Згідно з формулами (5.6), (5.7)

$$f_1(x)=\int_0^{+\infty} 2e^{-x-2y}dy=2e^{-x}\cdot\frac{1}{2}=e^{-x}, \qquad x\ge 0.$$

$$f_2(y)=\int_0^{+\infty} 2e^{-x-2y}dx=2e^{-2y}, \qquad y\ge 0.$$

Завдання

1. Система двох випадкових величин задана своєю функцією розподілу:

$$F(x,y)=\left(\frac{1}{\pi}\operatorname{arctg}x+\frac{1}{2}\right)\cdot\left(\frac{1}{\pi}\operatorname{arctg}\frac{y}{4}+\frac{1}{2}\right).$$

Знайти щільність сумісного розподілу.

<u>Відповідь.</u> $f(x,y)=\dfrac{4}{\pi^2(x^2+1)(y^2+16)}$.

2. Випадковий вектор $(X;Y)$ розподілений рівномірно в трикутнику D, утвореному прямими $x=0,\ y=0$ і $x+y=-5$, тобто

$$f(x,y)=\begin{cases}C & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D.\end{cases}$$

Знайти невідомий параметр C та ймовірність попадання в прямокутник $\Pi=\{(x;y):\ -1\le x\le 0;-2\le y\le 0\}$.

<u>Відповідь.</u> $C=0{,}08$. $P((X;Y)\in\Pi)=0{,}16$.

3. Неперервна двовимірна випадкова величина $(X;Y)$ задана своєю щільністю

$$f(x,y)=\begin{cases}a\cdot\sin x\cdot\sin y & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D\end{cases}$$

в області $D=\{(x;y):\ 0\le x\le\pi;\ 0\le y\le\pi\}$. Знайти невідомий параметр a, щільності розподілів компонент.

<u>Відповідь.</u> $a=0{,}25$. $f_1(x)=0{,}5\cdot\sin x$; $f_2(y)=0{,}5\cdot\sin y$.

5.5. Умовні закони розподілу

Нагадаємо, що дві випадкові величини називаються *незалежними*, якщо закон розподілу однієї з них не залежить від того, які можливі значення набула інша. Тому, звичайно, для незалежних випадкових величин їх умовні розподіли збігаються з безумовними.

Для визначення незалежності випадкових величин використовують нижченаведені критерії.

Теорема 5.1. Випадкові величини X і Y незалежні тоді і тільки тоді, коли функція розподілу системи $(X;Y)$ дорівнює добутку функцій розподілу складових:

$$\underline{F(x,y)=F_1(x)\cdot F_2(y)}. \qquad (5.8)$$

Теорема 5.2. Випадкові величини X і Y незалежні тоді і тільки тоді, коли щільність сумісного розподілу системи $(X;Y)$ дорівнює добутку щільностей розподілу складових:

$$\underline{f(x,y)=f_1(x)\cdot f_2(y)}. \qquad (5.9)$$

5.5.1. Умовні закони розподілу дискретних двовимірних випадкових величин

Нехай $(X;Y)$ − дискретна двовимірна випадкова величина; x_i $(i=\overline{1,n})$ − можливі значення компоненти X, y_j $(j=\overline{1,m})$ − можливі значення компоненти Y.

Нехай $p(x_i\,|\,y_j)$ – імовірність того, що $X=x_i$ за умови, що $Y=y_j$. *Умовним розподілом* X *при* $Y=y_j$ називається розподіл

X	x_1	x_2	…	x_n
$P(X\,|\,y_j)$	$p(x_1\,|\,y_j)$	$p(x_2\,|\,y_j)$	…	$p(x_n\,|\,y_j)$

Нехай $p(y_j\,|\,x_i)$ – імовірність того, що $Y=y_j$ за умови, що $X=x_i$. *Умовним розподілом* Y *при* $X=x_i$ називається розподіл

Y	y_1	y_2	…	y_m
$P(Y\,|\,x_i)$	$p(y_1\,|\,x_i)$	$p(y_2\,|\,x_i)$	…	$p(y_m\,|\,x_i)$

Умовні ймовірності обчислюють за формулами:

$$\underline{p(x_i\,|\,y_j)=\frac{p(x_i;y_j)}{p(y_j)}}; \qquad \underline{p(y_j\,|\,x_i)=\frac{p(x_i;y_j)}{p(x_i)}}, \qquad (5.10)$$

де $p_{ij} = p(x_i; y_j)$ – імовірності появи пари $(x_i; y_j)$ (знаходиться з таблиці сумісного розподілу X і Y);

$p_i = p(x_i)$ – імовірність появи значення x_i (знаходиться з безумовного розподілу компоненти X);

$p_j = p(y_j)$ –імовірності появи y_j (знаходиться з безумовного розподілу компоненти Y).

Зауваження. Як і для будь-якого розподілу, сума ймовірностей умовного розподілу завжди дорівнює одиниці.

Умовним математичним сподіванням $M(X | Y = y_j)$ ДДВВ називається

$$M(X | Y = y_j) = \sum_{i=1}^{n} x_i \cdot p(x_i | y_j). \quad (5.11)$$

Умовним математичним сподіванням $M(Y | X = x_i)$ ДДВВ називається

$$M(Y | X = x_i) = \sum_{j=1}^{m} y_j \cdot p(y_j | x_i). \quad (5.12)$$

Приклад. ДДВВ $(X; Y)$ задано законом розподілу

Y	X		
	-1	0	2
1	0,1	0,2	0,05
3	0,15	0,3	0,2

Знайти:

а) умовний закон розподілу компоненти X при $Y = y_1 = 1$ та відповідне умовне математичне сподівання;

б) умовний закон розподілу компоненти Y при $X = x_2 = 0$ та відповідне умовне математичне сподівання.

Розв’язання: а) закон розподілу компоненти Y:

Y	1	3
P	0,35	0,65

Обчислимо умовні ймовірності X при $Y = 1$:

$$p(x=-1\,|\,y=1)=\frac{p(-1;1)}{p(1)}=\frac{0{,}1}{0{,}35}=\frac{2}{7};$$
$$p(x=0\,|\,y=1)=\frac{p(0;1)}{p(1)}=\frac{0{,}2}{0{,}35}=\frac{4}{7};$$
$$p(x=2\,|\,y=1)=\frac{p(2;1)}{p(1)}=\frac{0{,}05}{0{,}35}=\frac{1}{7}.$$

Умовний закон розподілу компоненти X при $Y=1$ має вигляд

X	-1	0	2
$P(X\,\|\,Y=1)$	$2/7$	$4/7$	$1/7$

Перевірка: $\frac{2}{7}+\frac{4}{7}+\frac{1}{7}\equiv 1$.

За отриманим умовним розподілом знаходимо умовне математичне сподівання компоненти X при $Y=1$:

$$M(X\,|\,Y=1)=(-1)\cdot\frac{2}{7}+0\cdot\frac{4}{7}+2\cdot\frac{1}{7}=0;$$

б) закон розподілу компоненти X:

X	-1	0	2
P	0,25	0,5	0,25

Обчислимо умовні ймовірності Y при $X=0$:

$$p(y=1\,|\,x=0)=\frac{p(0;1)}{p(0)}=\frac{0{,}2}{0{,}5}=0{,}4;$$
$$p(y=3\,|\,x=0)=\frac{p(0;3)}{p(0)}=\frac{0{,}3}{0{,}5}=0{,}6.$$

Умовний закон розподілу компоненти Y при $X=3$ має вигляд

Y	1	3
$P(Y\,\|\,X=0)$	0,4	0,6

Перевірка: $0{,}4+0{,}6\equiv 1$.

За отриманим умовним розподілом знаходимо умовне математичне сподівання компоненти Y при $X=3$:

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

5.5.2. Умовні закони розподілу неперервних двовимірних випадкових величин

Нехай $(X;Y)$ − неперервна двовимірна випадкова величина, $f(x,y)$ − щільність сумісного розподілу $(X;Y)$, $f_1(x)$ і $f_2(y)$ − щільності розподілів компонент X і Y відповідно.

Умовною щільністю розподілу компоненти X при заданому значенні $Y=y$ називається функція

$$\varphi(x\,|\,y)=\frac{f(x,y)}{f_2(y)}. \tag{5.13}$$

Умовною щільністю розподілу компоненти Y при заданому значенні $X=x$ називається функція

$$\psi(y\,|\,x)=\frac{f(x,y)}{f_1(x)}. \tag{5.14}$$

Зауваження. Умовні щільності задовольняють умови:

1) $\varphi(x\,|\,y)\ge 0$; $\psi(y\,|\,x)\ge 0$;

2) $\int\limits_{-\infty}^{+\infty}\varphi(x\,|\,y)dx\equiv 1$; $\int\limits_{-\infty}^{+\infty}\psi(y\,|\,x)dy\equiv 1$. (5.15)

Умовним математичним сподіванням $M(X\,|\,Y=y)$ НДВВ називається величина

$$M(X\,|\,Y=y)=\int\limits_{-\infty}^{+\infty}x\cdot\varphi(x\,|\,y)dx \tag{5.16}$$

Умовним математичним сподіванням $M(Y\,|\,X=x)$ НДВВ називається величина

$$M(Y\,|\,X=x)=\int\limits_{-\infty}^{+\infty}y\cdot\psi(y\,|\,x)dy \tag{5.17}$$

Приклад. Неперервний випадковий вектор $(X;Y)$ задано щільністю $f(x,y)=\begin{cases}\frac{1}{4}\cdot\sin x\cdot\sin y & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D,\end{cases}$

де $D=\{(x;y):\ 0\le x\le\pi;\ 0\le y\le\pi\}$. Знайти умовні закони розподілу та умовні математичні сподівання.

Розв'язання. По-перше, знаходимо закони розподілів компонент:

$$f_1(x)=\begin{cases}\frac{1}{4}\int\limits_0^{\pi}\sin x\sin y dy=\frac{1}{4}\cdot\sin x\cdot\left(-\cos\pi+\cos 0\right)=\frac{1}{2}\sin x & \text{при } x\in[0;\pi],\\ 0 & \text{при } x\notin[0;\pi];\end{cases}$$

$$f_2(y)=\begin{cases}\frac{1}{2}\sin y & \text{при } x\in[0;\pi],\\ 0 & \text{при } x\notin[0;\pi].\end{cases}$$

Звідси за формулами (5.13), (5.14) умовні щільності мають вигляд:

$$\varphi(x\,|\,y)=\begin{cases}\frac{1}{2}\cdot\sin x & \text{при } x\in[0;\pi],\\ 0 & \text{при } x\notin[0;\pi];\end{cases}\qquad \psi(y\,|\,x)=\begin{cases}\frac{1}{2}\cdot\sin y & \text{при } y\in[0;\pi],\\ 0 & \text{при } y\notin[0;\pi].\end{cases}$$

За допомогою формул (5.16), (5.17) одержимо умовні математичні сподівання:

$$M(X\,|\,Y=y)=\frac{1}{2}\cdot\int\limits_0^{\pi}x\sin x dx=\frac{\pi}{2};\qquad M(Y\,|\,X=x)=\frac{1}{2}\cdot\int\limits_0^{\pi}y\sin y dx=\frac{\pi}{2}.$$

Завдання

1. Незалежні випадкові величини X і Y мають показникові закони розподілу:

$$f_1(x)=\begin{cases}\lambda e^{-\lambda x} & \text{при } x\ge 0,\\ 0 & \text{при } x<0;\end{cases}\qquad f_2(y)=\begin{cases}\mu e^{-\mu y} & \text{при } y\ge 0,\\ 0 & \text{при } y<0.\end{cases}$$

Знайти щільність та функцію розподілу системи випадкових величин $(X;Y)$.

Відповідь. $f(x,y)=\begin{cases}\lambda\mu e^{-(\lambda x+\mu y)} & \text{при} \quad (x;y)\in[0;+\infty)\times[0;+\infty),\\ 0 & \text{при} \quad (x;y)\notin[0;+\infty)\times[0;+\infty);\end{cases}$

$$F(x,y)=\begin{cases}\left(1-e^{-\lambda x}\right)\cdot\left(1-e^{-\mu y}\right) & \text{при} \quad (x;y)\in[0;+\infty)\times[0;+\infty),\\ 0 & \text{при} \quad (x;y)\notin[0;+\infty)\times[0;+\infty).\end{cases}$$

2. Дискретний випадковий вектор $(X;Y)$ задано законом розподілу

Y	X	
	-2	1
-1	0,1	0,3
2	0,2	0,1
3	0,1	0,2

Знайти: а) умовний закон розподілу компоненти X при $Y=3$ та відповідне умовне математичне сподівання;

б) умовний закон розподілу компоненти Y при $X=-2$ та відповідне умовне математичне сподівання.

Відповідь.

X	-2	1
$P(X\mid Y=3)$	1/3	2/3

Y	-1	2	3
$P(Y\mid X=-2)$	0,25	0,5	0,25

$M(X\mid Y=3)=0$; $M(Y\mid X=-2)=1{,}5$.

3. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X		
	0	3	4
-2	0,21	0,13	0,07
1	0,05	0,04	0,11
2	0,19	0,06	0,14

Знайти: а) умовний закон розподілу компоненти X при $Y=1$;

б) умовний закон розподілу компоненти Y при $X=0$.

Відповідь.

X	0	3	4
$P(X\mid Y=1)$	0,25	0,2	0,55

Y	-2	1	2
$P(Y\mid X=0)$	7/15	1/9	19/45

$M(X\mid Y=1)=2{,}8$; $M(Y\mid X=0)=1/45$.

4. Знайти щільності розподілів компонент неперервної двовимірної випадкової величини $(X;Y)$, заданої функцією $f(x,y)=2e^{-x-2y}, \quad x\ge 0, y\ge 0$.

Відповідь. $f_1(x)=\begin{cases} e^{-x} & \text{при } x\ge 0,\\ 0 & \text{при } x<0;\end{cases}$ $\quad f_2(y)=\begin{cases} 2e^{-2y} & \text{при } y\ge 0,\\ 0 & \text{при } y<0.\end{cases}$

5. Неперервний випадковий вектор $(X;Y)$ задано щільністю

$$f(x,y)=\begin{cases}\cos x\cdot\cos y & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D,\end{cases}$$

де $D=\{(x;y):\ 0\le x\le \pi/2;\ 0\le y\le \pi/2\}$. Знайти умовні закони розподілу.

Відповідь. $\varphi(x\mid y)=\begin{cases}\cos x & \text{при } x\in\left[0;\dfrac{\pi}{2}\right],\\ 0 & \text{при } x\notin\left[0;\dfrac{\pi}{2}\right];\end{cases}$

$$\psi(y\mid x)=\begin{cases}\cos y & \text{при } y\in\left[0;\dfrac{\pi}{2}\right],\\ 0 & \text{при } y\notin\left[0;\dfrac{\pi}{2}\right].\end{cases}$$

5.6. Числові характеристики двовимірної випадкової величини. Кореляційний момент та коефіцієнт кореляції

Математичне сподівання добутку випадкових величин X і Y обчислюється за формулами

$$M(XY)=\sum_i\sum_j x_i\cdot y_j\cdot p_{ij}\,, \tag{5.18}$$

$$M(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot y\cdot f(x,y)dxdy \tag{5.19}$$

для дискретних і неперервних величин відповідно.

Крім числових характеристик компонент випадкового вектора, використовуються величини, що характеризують зв'язок між його компонентами.

Кореляційним моментом μ_{xy} або *коваріацією* $\mathrm{cov}(X;Y)$ випадкових величин X і Y називається математичне сподівання добутку відхилень X і Y:

$$\mu_{xy}=\mathrm{cov}(X;Y)=M\big[(X-M(X))\cdot(Y-M(Y))\big]. \tag{5.20}$$

На практиці для обчислення μ_{xy} зручніше використовувати формулу

$$\mu_{xy}=M(XY)-M(X)M(Y). \tag{5.21}$$

Зауваження. Іноді кореляційний момент визначають у термінах центрованих величин $X_0=X-M(X)$ і $Y_0=Y-M(Y)$: $\mu_{xy}=M\left(X_0\cdot Y_0\right)$.

Кореляційною матрицею випадкових величин X і Y називається симетрична матриця

$$K_{xy}=\begin{pmatrix} k_{xx} & k_{xy}\\ k_{yx} & k_{yy}\end{pmatrix}, \tag{5.22}$$

компоненти якої визначаються формулами:

$$k_{xy}=k_{yx}=\mu_{xy}; \qquad k_{xx}=D(X); \qquad k_{yy}=D(Y). \tag{5.23}$$

Коефіцієнтом кореляції r_{xy} випадкових величин X і Y називається величина

$$r_{xy} = \frac{\mu_{xy}}{\sigma(X)\cdot\sigma(Y)}, \tag{5.24}$$

де $\sigma(X), \sigma(Y)$ − середні квадратичні відхилення X і Y відповідно.

Теорема 5.3. Кореляційний момент двох незалежних випадкових величин дорівнює нулю.

Теорема 5.4. $|\mu_{xy}| \le \sqrt{D_x \cdot D_y}$.

Теорема 5.5. $|r_{xy}| \le 1$.

Випадкові величини X і Y називаються *корельованими*, якщо їх кореляційний момент (або коефіцієнт кореляції) не дорівнює нулю. Навпаки, X і Y – *некорельовані*, якщо

$$\mu_{xy} = r_{xy} = 0. \tag{5.25}$$

З наведених вище теорем випливають твердження:

• дві корельовані випадкові величини завжди є залежними. Але залежні випадкові величини можуть бути як корельованими, так і некорельованими;

• дві незалежні випадкові величини завжди є некорельованими. Але некорельовані можуть бути як залежними, так і незалежними.

Для дискретних випадкових величин формула обчислення кореляційного моменту (5.21) набуває вигляду:

$$\mu_{xy} = M(XY) - M(X)M(Y) = \sum_{i,j} x_i y_j p_{ij} - M(X)M(Y), \tag{5.26}$$

де $M(X), M(Y)$ − математичні сподівання X і Y відповідно. Як і середні квадратичні відхилення $\sigma(X), \sigma(Y)$ у формулі (5.24) для коефіцієнта кореляції вони обчислюються за безумовними законами розподілу X і Y.

Приклад. Розподіл дискретного випадкового вектора $(X;Y)$ задано таблицею:

Y	X		
	-1	2	3
0	0,3	0,1	0,2
1	0,05	0,2	0,15

Знайти кореляційний момент і коефіцієнт кореляції X і Y, скласти кореляційну матрицю. Встановити, чи є компоненти $(X;Y)$ залежними (незалежними).

Розв’язання. За законом розподілу двовимірної випадкової величини $(X;Y)$ знаходимо:

$$M(XY)=$$

$$=0\cdot(-1)\cdot 0{,}3+0\cdot 2\cdot 0{,}1+0\cdot 3\cdot 0{,}2+1\cdot(-1)\cdot 0{,}05+1\cdot 2\cdot 0{,}2+1\cdot 3\cdot 0{,}15=0{,}8.$$

Закон розподілу компоненти X:

X	-1	2	3
P	0,35	0,3	0,35

Числові характеристики X:

$M(X)=(-1)\cdot 0{,}35+2\cdot 0{,}3+3\cdot 0{,}35=1{,}3$;

$D(X)=(-1)^2\cdot 0{,}35+2^2\cdot 0{,}3+3^2\cdot 0{,}35-(1{,}3)^2=4{,}7-1{,}69=3{,}01$;

$\sigma(X)=\sqrt{3{,}01}\approx 1{,}73$.

Закон розподілу компоненти Y:

Y	0	1
P	0,6	0,4

Числові характеристики Y:

$M(Y)=0\cdot 0{,}6+1\cdot 0{,}4=0{,}4$;

$D(Y)=0^2\cdot 0{,}6+1^2\cdot 0{,}4-(0{,}4)^2=0{,}24$;

$\sigma(Y)=\sqrt{0{,}24}\approx 0{,}49$.

Кореляційний момент, кореляційну матрицю і коефіцієнт кореляції складових X та Y знаходимо за формулами (5.26), (5.22)−(5.24):

$$\mu_{xy} = M(XY) - M(X)M(Y) = 0{,}8 - 1{,}3 \cdot 0{,}4 = 0{,}28;$$

$$K_{xy} = \begin{pmatrix} 3{,}01 & 0{,}28 \\ 0{,}28 & 0{,}24 \end{pmatrix};$$

$$r_{xy} = \frac{\mu_{xy}}{\sigma(X)\cdot\sigma(Y)} = \frac{0{,}28}{1{,}73 \cdot 0{,}49} \approx 0{,}33.$$

Оскільки кореляційний момент (а тому і коефіцієнт кореляції) не дорівнюють нулю, компоненти X і Y заданого випадкового вектора є корельованими, а тому залежними.

Для неперервних випадкових величин, формула обчислення кореляційного моменту (5.21) набуває вигляду:

$$\mu_{xy} = M(XY) - M(X)M(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x \cdot y \cdot f(x,y)dxdy - M(X)M(Y), \quad (5.27)$$

де $$M(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x \cdot f(x,y)dxdy, \quad M(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y \cdot f(x,y)dxdy. \quad (5.28)$$

Для обчислення коефіцієнта кореляції використовується формула (5.24), де

$$\sigma(X) = \sqrt{D(X)}, \; \sigma(Y) = \sqrt{D(Y)}, \quad (5.29)$$

$$D(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 \cdot f(x,y)dxdy - M^2(X), \quad (5.30)$$

$$D(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2 \cdot f(x,y)dxdy - M^2(Y). \quad (5.31)$$

Приклад. Розподіл неперервного випадкового вектора $(X;Y)$ задано щільністю:

$$f(x,y) = \begin{cases} 1, & (x,y) \in [0;1]\times[0;1]; \\ 0, & (x,y) \notin [0;1]\times[0;1]. \end{cases}$$

Знайти кореляційний момент і коефіцієнт кореляції X і Y, кореляційну матрицю. Встановити, чи є компоненти $(X;Y)$ залежними (незалежними).

<u>Розв'язання.</u> По-перше, за формулами (5.28)-(5.31) обчислюємо числові характеристики компонент X і Y:

$$M(X)=\int_0^1\int_0^1 x dx dy=\int_0^1 x dx\int_0^1 dy=\int_0^1 x dx=\left.\frac{x^2}{2}\right|_0^1=\frac{1}{2};$$

$$M(Y)=\int_0^1\int_0^1 y dx dy=\int_0^1 y dy\int_0^1 dx=\int_0^1 y dy=\left.\frac{y^2}{2}\right|_0^1=\frac{1}{2};$$

$$D(X)=\int_0^1\int_0^1 x^2 dx dy-\left(\frac{1}{2}\right)^2=\int_0^1 x^2 dx\int_0^1 dy-\frac{1}{4}=\int_0^1 x^2 dx-\frac{1}{4}=\left.\frac{x^3}{3}\right|_0^1-\frac{1}{4}=\frac{1}{3}-\frac{1}{4}=\frac{1}{12};$$

$$D(Y)=\int_0^1\int_0^1 y^2 dx dy-\left(\frac{1}{2}\right)^2=\frac{1}{12};$$

$$\sigma(X)=\sqrt{D(X)}=\frac{1}{2\sqrt{3}};\quad \sigma(Y)=\sqrt{D(Y)}=\frac{1}{2\sqrt{3}}.$$

За формулою (5.19) обчислимо також

$$M(XY)=\int_0^1\int_0^1 x\cdot y\, dx dy=\int_0^1 x dx\int_0^1 y dy=\left(\left.\frac{x^2}{2}\right|_0^1\right)\cdot\left(\left.\frac{y^2}{2}\right|_0^1\right)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}.$$

Тоді за формулою (5.27) кореляційний момент X і Y дорівнює

$$\mu_{xy}=M(XY)-M(X)M(Y)=\frac{1}{4}-\frac{1}{2}\cdot\frac{1}{2}=0,$$

а тому і коефіцієнт кореляції X і Y також дорівнює нулю. Кореляційна матриця в даному випадку має вигляд:

$$K_{xy}=\begin{pmatrix}1/12 & 0\\ 0 & 1/12\end{pmatrix}.$$

Таким чином, компоненти заданого випадкового вектора є некорельованими. Щоб встановити, чи є вони незалежними, скористаємось Теоремою 5.2 (див. п. 5.5). За формулами (5.6), (5.7) щільності розподілів компонент дорівнюють

$$f_1(x)=\begin{cases}\int\limits_{-\infty}^{+\infty} f(x,y)dy=\int\limits_0^1 dy=1 & \text{при } x\in[0;1],\\ 0 & \text{при } x\notin[0;1];\end{cases}$$

$$f_2(y)=\begin{cases}\int\limits_{-\infty}^{+\infty} f(x,y)dx=\int\limits_0^1 dx=1 & \text{при } y\in[0;1],\\ 0 & \text{при } y\notin[0;1],\end{cases}$$

тому виконується $f(x,y)=f_1(x)\cdot f_2(y)$. Згідно з (5.9) компоненти X і Y даної системи випадкових величин $(X;Y)$ є незалежними.

Завдання

1. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X		
	1	3	6
2	0,1	0,3	0,1
10	0,05	0,39	0,06

Знайти кореляційний момент, коефіцієнт кореляції X і Y, кореляційну таблицю.

<u>Відповідь.</u> $\mu_{xy}=-0{,}08$; $r_{xy}\approx-0{,}014$; $K_{xy}=\begin{pmatrix}2{,}0076 & -0{,}08\\ -0{,}08 & 16\end{pmatrix}$.

2. Дискретна двовимірна випадкова величина $(X;Y)$ задана законом розподілу

Y	X		
	10	20	30
20	0,15	0,1	0,05
40	0,05	0,2	0,1
60	0	0,1	0,25

Знайти числові характеристики компонент, кореляційний момент, коефіцієнт кореляції X і Y, кореляційну таблицю.

Відповідь. $M(X)=22$; $M(Y)=41$; $D(X)=56$; $D(Y)=259$; $\sigma(X)\approx 7{,}48$; $\sigma(Y)\approx 16{,}09$; $\mu_{xy}\approx 67{,}4$; $r_{xy}\approx 0{,}56$; $K_{xy}=\begin{pmatrix}56 & 67{,}4\\ 67{,}4 & 259\end{pmatrix}$.

3. Неперервний випадковий вектор $(X;Y)$ задано щільністю:

$$f(x,y)=\begin{cases} axy & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D\end{cases}$$

у трикутнику D, утвореному прямими $x=0,\ y=0$ і $x+y-1=0$. Знайти невідомий параметр a, кореляційний момент і коефіцієнт кореляції X і Y.

Відповідь. $a=24$; $\mu_{xy}=2/75$; $r_{xy}=-2/3$.

4. Неперервний випадковий вектор $(X;Y)$ задано щільністю

$$f(x,y)=\begin{cases} \cos x\cdot\cos y & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D,\end{cases}$$

$D=\{(x;y):\ 0\le x\le \pi/2;\ 0\le y\le \pi/2\}$. Знайти числові характеристики компонент, кореляційний момент, кореляційну таблицю і коефіцієнт кореляції.

Відповідь. $M(X)=M(Y)=\dfrac{\pi}{2}-1$; $D(X)=D(Y)=\pi-3$; $\mu_{xy}=r_{xy}=0$; $K_{xy}=\begin{pmatrix}\pi-3 & 0\\ 0 & \pi-3\end{pmatrix}$.

5. Неперервна двовимірна випадкова величина $(X;Y)$ задана своєю щільністю

$$f(x,y)=\begin{cases} a\cdot\sin(x+y) & \text{при } (x,y)\in D,\\ 0 & \text{при } (x,y)\notin D\end{cases}$$

в області $D=\{(x;y):\ 0\le x\le \pi/2;\ 0\le y\le \pi/2\}$. Знайти невідомий параметр a, числові характеристики компонент, кореляційний момент, кореляційну таблицю і коефіцієнт кореляції.

Відповідь. $a = 0{,}5$. $M(X) = M(Y) = \pi/4$;

$D(X) = D(Y) = (\pi^2 + 8\pi - 32)/16$; $\mu_{xy} = \dfrac{8\pi - 16 - \pi^2}{16} \approx -0{,}046$;

$K_{xy} = \dfrac{1}{16} \cdot \begin{pmatrix} \pi^2 + 8\pi - 32 & 8\pi - 16 - \pi^2 \\ 8\pi - 16 - \pi^2 & \pi^2 + 8\pi - 32 \end{pmatrix}$; $r_{xy} \approx -0{,}2454$.

5.7. Лінійна регресія

Нехай $(X;Y)$ − двовимірна випадкова величина, компоненти якої X і Y − залежні випадкові величини. На практиці виникає необхідність знаходження наближеного виразу компоненти Y через X у вигляді лінійної залежності:

$$Y \cong f(X) = \alpha X + \beta, \qquad \alpha, \beta \in R.$$

Середньоквадратичною регресією Y на X називається така лінійна функція $f(X) = \alpha X + \beta$, для якої $M[Y - f(X)]^2 = \min$.

Теорема 5.6. Лінійна середня квадратична регресія Y на X має вигляд

$$\underline{f(X) = m_y + r_{xy} \cdot \frac{\sigma_y}{\sigma_x} \cdot (X - m_x),} \tag{5.32}$$

де $m_x = M(X)$, $m_y = M(Y)$ − математичні сподівання;

$\sigma_x = \sqrt{D(X)}$, $\sigma_y = \sqrt{D(Y)}$ − середні квадратичні відхилення;

$r_{xy} = \dfrac{\mu_{xy}}{\sigma_x \sigma_y}$ − коефіцієнт кореляції X і Y.

Відповідна функції (5.32) пряма

$$\underline{y - m_y = r_{xy} \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - m_x)} \tag{5.33}$$

називається *прямою середньоквадратичної регресії Y на X*; коефіцієнт $r_{xy} \cdot \dfrac{\sigma_y}{\sigma_x}$ − *коефіцієнтом регресії Y на X*.

Величину похибки, що допускається при заміні Y лінійною функцією $f(X)$ вигляду (5.32) характеризує величина

$$\sigma_y^2 \cdot \left(1 - r_{xy}^2\right) = M\left[Y - f(X)\right]^2,$$

що називається *остаточною дисперсією* величини Y відносно X.

Аналогічно, *пряма середньоквадратичної регресії X на Y* має вигляд

$$x - m_x = r_{xy} \cdot \frac{\sigma_x}{\sigma_y} \cdot (y - m_y), \tag{5.34}$$

коефіцієнт регресії X на Y дорівнює $r_{xy} \cdot \dfrac{\sigma_x}{\sigma_y}$, а *остаточна дисперсія* випадкової величини X відносно Y дорівнює $\sigma_x^2 \cdot \left(1 - r_{xy}^2\right)$.

З рівнянь (5.33) і (5.34) легко бачити, що обидві прямі проходять через точку з координатами $(m_x; m_y)$, яка називається *центром сумісного розподілу* X і Y.

<u>Зауваження.</u> Коли коефіцієнт кореляції $r_{xy} = \pm 1$, остаточна дисперсія дорівнює нулю. Таким чином, при $r_{xy} = \pm 1$ випадкові величини X і Y пов'язані лінійною залежністю. Більше того, в цьому випадку прямі регресії (5.33) і (5.34) збігаються.

Приклад. Розподіл неперервного випадкового вектора $(X; Y)$ задано щільністю:

$$f(x, y) = \begin{cases} a(x + y) & \text{при } (x, y) \in D, \\ 0 & \text{при } (x, y) \notin D; \end{cases}$$

область D обмежена лініями $x = 1, y = 0, y = x$. Знайти рівняння лінійної регресії Y на X та X на Y.

<u>Розв'язання.</u> Область D є трикутником (рис. 5.3). По-перше, знаходимо значення параметра a з умови нормування (5.4):

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)dxdy = a\iint_D (x + y)dxdy = a \cdot \left[\iint_D xdxdy + \iint_D ydxdy\right] =$$

$$= a\cdot\left[\int_0^1 x dx\int_0^x dy+\int_0^1 dx\int_0^x y\, dy\right]=a\cdot\left[\int_0^1 x dx\cdot\left(y\Big|_0^x\right)+\int_0^1 dx\cdot\left(\frac{y^2}{2}\Bigg|_0^x\right)\right]=$$

$$= a\cdot\left[\int_0^1 x^2 dx+\frac{1}{2}\cdot\int_0^1 x^2 dx\right]=\frac{3}{2}a\cdot\int_0^1 x^2 dx=\frac{3}{2}a\cdot\left(\frac{x^3}{3}\Bigg|_0^1\right)=\frac{a}{2}\equiv 1 \qquad \Rightarrow \quad a=2.$$

Звідси щільність розподілу має вигляд:

$$f(x,y)=\begin{cases}2(x+y) & \text{при} \quad (x,y)\in D,\\ 0 & \text{при} \quad (x,y)\notin D.\end{cases}$$

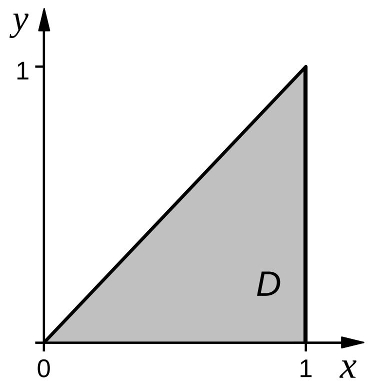


Рис. 5.3

За формулами (5.19), (5.28) знайдемо числові характеристики компонент даного випадкового вектора.

$$M(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot f(x,y)dxdy=2\iint_D x\cdot(x+y)dxdy=2\cdot\left[\iint_D x^2 dxdy+\iint_D xydxdy\right]=$$

$$=2\cdot\left[\int_0^1 x^2 dx\int_0^x dy+\int_0^1 xdx\int_0^x y\, dy\right]=2\cdot\left[\int_0^1 x^2 dx\cdot\left(y\Big|_0^x\right)+\int_0^1 xdx\cdot\left(\frac{y^2}{2}\Bigg|_0^x\right)\right]=$$

$$=2\cdot\left[\int_0^1 x^3 dx+\frac{1}{2}\cdot\int_0^1 x^3 dx\right]=3\cdot\int_0^1 x^3 dx=3\cdot\left(\frac{x^4}{4}\Bigg|_0^1\right)=\frac{3}{4}.$$

$$M(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\cdot f(x,y)dxdy=2\iint_D y\cdot(x+y)dxdy=2\cdot\left[\iint_D xydxdy+\iint_D y^2 dxdy\right]=$$

$$=2\cdot\left[\int_0^1 xdx\int_0^x ydy+\int_0^1 dx\int_0^x y^2\,dy\right]=2\cdot\left[\int_0^1 xdx\cdot\left(\frac{y^2}{2}\Big|_0^x\right)+\int_0^1 dx\cdot\left(\frac{y^3}{3}\Big|_0^x\right)\right]=$$

$$=2\cdot\left[\int_0^1\frac{x^3}{2}dx+\int_0^1\frac{x^3}{3}dx\right]=2\cdot\frac{5}{6}\cdot\int_0^1 x^3dx=\frac{5}{3}\cdot\left(\frac{x^4}{4}\Big|_0^1\right)=\frac{5}{12}.$$

$$M(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\cdot f(x,y)dxdy=2\iint_D xy\cdot(x+y)dxdy=2\cdot\left[\iint_D x^2ydxdy+\iint_D xy^2dxdy\right]=$$

$$=2\cdot\left[\int_0^1 x^2dx\int_0^x ydy+\int_0^1 xdx\int_0^x y^2\,dy\right]=2\cdot\left[\int_0^1 x^2dx\cdot\left(\frac{y^2}{2}\Big|_0^x\right)+\int_0^1 xdx\cdot\left(\frac{y^3}{3}\Big|_0^x\right)\right]=$$

$$=2\cdot\left[\int_0^1\frac{x^4}{2}dx+\int_0^1\frac{x^4}{3}dx\right]=2\cdot\frac{5}{6}\cdot\int_0^1 x^4dx=\frac{5}{3}\cdot\left(\frac{x^5}{5}\Big|_0^1\right)=\frac{1}{3}.$$

Помічаємо, що оскільки $M(XY)\neq M(X)\cdot M(Y)$, випадкові величини X і Y є залежними.

Кореляційний момент X і Y дорівнює

$$\mu_{xy}=M(XY)-M(X)M(Y)=\frac{1}{3}-\frac{3}{4}\cdot\frac{5}{12}=\frac{1}{48}.$$

Дисперсії компонент обчислюємо за формулами (5.30), (5.31):

$$D(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2\cdot f(x,y)dxdy-M^2(X)=2\iint_D x^2\cdot(x+y)dxdy-\left(\frac{3}{4}\right)^2=$$

$$=2\cdot\left[\int_0^1 x^3dx\int_0^x dy+\int_0^1 x^2dx\int_0^x y\,dy\right]-\frac{9}{16}=2\cdot\left[\int_0^1 x^3dx\cdot\left(y\big|_0^x\right)+\int_0^1 x^2dx\cdot\left(\frac{y^2}{2}\Big|_0^x\right)\right]-\frac{9}{16}=$$

$$=2\cdot\left[\int_0^1 x^4dx+\frac{1}{2}\cdot\int_0^1 x^4dx\right]-\frac{9}{16}=3\cdot\int_0^1 x^4dx-\frac{9}{16}=3\cdot\left(\frac{x^5}{5}\Big|_0^1\right)-\frac{9}{16}=\frac{3}{5}-\frac{9}{16}=\frac{3}{80}.$$

$$D(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2\cdot f(x,y)dxdy-M^2(Y)=2\iint_D y^2\cdot(x+y)dxdy-\left(\frac{5}{12}\right)^2=$$

$$= 2 \cdot \left[\int_0^1 x dx \int_0^x y^2 dy + \int_0^1 dx \int_0^x y^3 \, dy \right] - \frac{25}{144} = 2 \cdot \left[\int_0^1 x dx \cdot \left(\frac{y^3}{3} \bigg|_0^x \right) + \int_0^1 dx \cdot \left(\frac{y^4}{4} \bigg|_0^x \right) \right] - \frac{25}{144} =$$

$$= 2 \cdot \left[\int_0^1 \frac{x^4}{3} dx + \int_0^1 \frac{x^4}{4} dx \right] - \frac{25}{144} = 2 \cdot \frac{7}{12} \cdot \int_0^1 x^4 dx - \frac{25}{144} = \frac{7}{6} \cdot \left(\frac{x^5}{5} \bigg|_0^1 \right) - \frac{25}{144} =$$

$$= \frac{7}{30} - \frac{25}{144} = \frac{43}{720}.$$

Середні квадратичні відхилення компонент дорівнюють

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\frac{3}{80}} = \frac{\sqrt{3}}{4\sqrt{5}}; \qquad \sigma(Y) = \sqrt{D(Y)} = \sqrt{\frac{43}{720}} = \frac{\sqrt{43}}{12\sqrt{5}}.$$

За формулою (5.24) знайдемо коефіцієнт кореляції:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} = \frac{\frac{1}{48}}{\frac{\sqrt{3}}{4\sqrt{5}} \cdot \frac{\sqrt{43}}{12\sqrt{5}}} = \frac{5}{\sqrt{129}}.$$

Перш ніж скласти рівняння лінійної регресії, зручніше обчислити коефіцієнт регресії Y на X:

$$r_{xy} \cdot \frac{\sigma_y}{\sigma_x} = \frac{5}{\sqrt{129}} \cdot \frac{\frac{\sqrt{43}}{12\sqrt{5}}}{\frac{\sqrt{3}}{4\sqrt{5}}} = \frac{5}{9}.$$

Підставивши результати в (5.33), отримаємо рівняння прямої регресії Y на X:

$$y - \frac{5}{12} = \frac{5}{9} \cdot \left(x - \frac{3}{4} \right) \quad \Leftrightarrow \quad y = \frac{5}{9} x.$$

Аналогічно знаходимо рівняння прямої регресії X на Y:

$$x - \frac{3}{4} = \frac{15}{43} \cdot \left(y - \frac{5}{12} \right) \quad \Leftrightarrow \quad x = \frac{15}{43} y + \frac{26}{43}.$$

Завдання

1. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X		
	1	3	6
2	0,1	0,3	0,1
10	0,05	0,39	0,06

Знайти рівняння лінійної регресії.

Відповідь. $y = -0{,}04 \cdot x + 6{,}13$.

2. Розподіл системи неперервних випадкових величин $(X;Y)$ задано щільністю:

$$f(x,y) = \begin{cases} a, & (x,y) \in D; \\ 0, & (x,y) \notin D, \end{cases}$$

де D − трикутник, обмежений прямими $x = 0, y = 0$ і $y = x - 3$. Знайти невідомий параметр і рівняння лінійної регресії Y на X.

Відповідь. $a = 4{,}5$; $y = 0{,}5x - 1{,}5$.

5.8. Двовимірний нормальний закон розподілу

Як і одновимірні, значно поширеними є нормально розподілені двовимірні випадкові величини.

Нормальним законом розподілу на площині називається розподіл імовірностей випадкового вектора $(X;Y)$, заданий щільністю вигляду

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_{xy}^2}} \times$$
$$\times \exp\left(-\frac{1}{2(1-r_{xy}^2)} \cdot \left[\frac{(x-a_1)^2}{\sigma_x^2} + \frac{(y-a_2)^2}{\sigma_y^2} - 2r_{xy}\frac{x-a_1}{\sigma_x} \cdot \frac{y-a_2}{\sigma_y} \right] \right), \quad (5.35)$$

де $a_1, a_2, \sigma_x, \sigma_y, r_{xy}$ – параметри розподілу. Імовірнісний зміст параметрів:

$a_1, a_2 \in R$ – математичні сподівання X і Y відповідно;

$\sigma_x > 0, \sigma_y > 0$ – середні квадратичні відхилення X і Y;

r_{xy} – коефіцієнт кореляції X і Y.

Компоненти X і Y нормально розподіленої двовимірної випадкової величини $(X;Y)$ також розподілені нормально з параметрами a_1, σ_x і a_2, σ_y відповідно.

Зауваження. Для нормально розподіленої двовимірної випадкової величини некорельованість рівносильна незалежності.

Теорема 5.7. Компоненти X і Y нормально розподіленої двовимірної випадкової величини $(X;Y)$ пов'язані лінійною кореляційною залежністю.

Таким чином, графіки функцій регресії Y на X та X на Y збігаються з відповідними прямими середньоквадратичної регресії і мають вигляд:

$$\underline{y - a_2 = r_{xy} \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - a_1);} \qquad \underline{x - a_1 = r_{xy} \cdot \frac{\sigma_x}{\sigma_y} \cdot (y - a_2).} \tag{5.36}$$

Приклад. Кореляційна матриця системи нормально розподілених випадкових величин X і Y має вигляд $K_{xy} = \begin{pmatrix} 9 & -6 \\ -6 & 16 \end{pmatrix}$, а математичні сподівання дорівнюють $M(X) = 5$ і $M(Y) = -1$ відповідно. Виписати щільність розподілу. Знайти рівняння регресії.

Розв'язання. Згідно з (5.22), (5.23) діагональні елементи кореляційної матриці дорівнюють дисперсіям компонент випадкового вектора, тому

$$\sigma_x = \sqrt{D(X)} = \sqrt{9} = 3; \qquad \sigma_y = \sqrt{D(Y)} = \sqrt{16} = 4.$$

Крім того, з (5.22)–(5.23) знайдемо, що кореляційний момент і коефіцієнт кореляції дорівнюють відповідно

$$\mu_{xy} = -6; \qquad r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y} = \frac{-6}{3 \cdot 4} = -\frac{1}{2}.$$

За формулою (5.35) запишемо вираз для щільності ймовірностей:

$$f(x, y) = \frac{1}{12\pi\sqrt{3}} \cdot \exp\left(-\frac{2}{3} \cdot \left[\frac{(x-5)^2}{9} + \frac{(y+1)^2}{16} + \frac{(x-5)\cdot(y+1)}{12} \right] \right).$$

Рівняння лінійної регресії Y на X згідно з (5.36) має вигляд:

$$y + 1 = -\frac{1}{2} \cdot \frac{4}{3} \cdot (x-5) \qquad \Leftrightarrow \qquad y = -\frac{2}{3} \cdot x + \frac{7}{3},$$

рівняння регресії X на Y:

$$x - 5 = -\frac{1}{2} \cdot \frac{3}{4} \cdot (y+1) \qquad \Leftrightarrow \qquad x = -\frac{3}{8} y + \frac{37}{8}.$$

Завдання

1. Довести, що нормально розподілені некорельовані випадкові величини X і Y є незалежними.

Вказівка. Скористатись теоремою 5.2 п.5.5.

2. Задано щільності розподілів незалежних компонент X і Y нормально розподіленої двовимірної випадкової величини $(X; Y)$:

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-2)^2}{2}}; \qquad f_2(y) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(y+1)^2}{18}}.$$

Виписати щільність розподілу і кореляційну матрицю системи випадкових величин $(X; Y)$.

Відповідь. $K_{xy} = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}; \quad f(x, y) = \frac{1}{6\pi} \cdot \exp\left(-\frac{9(x-2)^2 + (y+1)^2}{18} \right).$

3. Система $(X; Y)$ нормально розподілених центрованих ($M(X) = M(Y) = 0$) випадкових величин має кореляційну матрицю $K_{xy} = \begin{pmatrix} 4 & 0{,}5 \\ 0{,}5 & 1 \end{pmatrix}$. Виписати щільність розподілу системи; знайти рівняння регресії Y на X і X на Y.

Відповідь. $f(x,y)=\dfrac{1}{\pi\sqrt{15}}\cdot\exp\left(-\dfrac{2}{15}\cdot\left[x^2+4y^2-xy\right]\right)$; $y=\dfrac{x}{8}$; $x=\dfrac{y}{2}$.

Питання до теми

1. Як визначається багатовимірна випадкова величина? Наведіть приклади.

2. Який вигляд має закон розподілу системи дискретних випадкових величин?

3. Як знайти закони розподілу компонент дискретного випадкового вектора? Наведіть приклад.

4. Як визначається функція розподілу двовимірної випадкової величини? Які її властивості?

5. Наведіть означення та властивості щільності сумісного розподілу неперервної двовимірної випадкової величини. Як за щільністю знайти функцію розподілу?

6. Як знайти щільності розподілів компонент ДНВВ? Які властивості вони мають?

7. Які критерії незалежності компонент випадкового вектора ви знаєте?

8. Як знайти умовні закони розподілу та умовні математичні сподівання компонент двовимірного дискретного вектора?

9. Як визначаються умовні щільності розподілів та умовні математичні сподівання компонент двовимірного неперервного вектора?

10. Які числові характеристики визначають зв'язок між компонентами випадкового вектора? Які їх властивості?

11. Як визначається кореляційна матриця?

12. Як пов’язані між собою властивості залежності (незалежності) та корельованості (некорельованості)?

13. Як визначається функція середньоквадратичної регресії? Як побудувати прямі регресій?

14. Який вигляд має двовимірний нормальний закон розподілу? Які властивості мають компоненти нормально розподіленого випадкового вектора?

Тестові питання

1. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X	
	0	1
2	0,4	k
3	0,25	0,15

Чому дорівнює параметр k?

А	Б	В	Г	Д
1	0,1	0,2	0,3	Інша відповідь

2. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X	
	0	1
2	0,1	0,2
3	0,3	0,4

Яка з таблиць задає розподіл компоненти Y?

<table>
<tr><th>А</th><th>Б</th><th>В</th><th>Г</th><th>Д</th></tr>
<tr>
<td><table><tr><td>Y</td><td>0</td><td>1</td></tr><tr><td>P</td><td>0,4</td><td>0,6</td></tr></table></td>
<td><table><tr><td>Y</td><td>2</td><td>3</td></tr><tr><td>P</td><td>0,3</td><td>0,7</td></tr></table></td>
<td><table><tr><td>Y</td><td>0</td><td>1</td></tr><tr><td>P</td><td>0,3</td><td>0,7</td></tr></table></td>
<td><table><tr><td>Y</td><td>2</td><td>3</td></tr><tr><td>P</td><td>0,4</td><td>0,6</td></tr></table></td>
<td>Інша відповідь</td>
</tr>
</table>

3. Як за щільністю сумісного розподілу знайти функцію розподілу?

А	Б	В	Г	Д
$f(x) = F'_x(x)$	$F(x) = \int\limits_{-\infty}^{x} f(x)dx$	$f(x,y) = F''_{xy}(x,y)$	$F(x,y) = \int\limits_{-\infty}^{y}\int\limits_{-\infty}^{x} f(x,y)dxdy$	Інша відповідь

4. Система двох випадкових величин $(X;Y)$ задана функцією розподілу:

$$F(x,y) = \frac{1}{2}\big(\sin x + \sin y - \sin(x+y)\big) \quad \text{при} \quad (x,y) \in \left[0;\frac{\pi}{2}\right] \times \left[0;\frac{\pi}{2}\right].$$

Яка формула відповідає щільності сумісного розподілу при $0 \le x \le \pi/2$, $0 \le y \le \pi/2$?

А	Б	В	Г	Д
$f(x,y) = 0{,}5\sin(x+y)$	$f(x,y) = 2\sin(x+y)$	$f(x,y) = 0{,}5$	$f(x,y) = \sin(x+y)$	Інша відповідь

5. Система двох випадкових величин $(X;Y)$ задана законом розподілу

Y	X	
	-1	1
-2	0,3	0,2
2	0,4	0,1

Чому дорівнює умовна ймовірність $p(X=1|Y=2)$?

А	Б	В	Г	Д
0,1	0,2	0,3	0,4	Інша відповідь

6. Яку умову задовольняє коефіцієнт кореляції X і Y?

А	Б	В	Г	Д
$\lvert r_{xy} \rvert \le 1$	$\lvert r_{xy} \rvert \ge 1$	$r_{xy} \le 1$	$r_{xy} \ge 1$	Інша відповідь

7. Коефіцієнтом регресії X на Y називається величина:

А	Б	В	Г	Д
$r_{xy} \cdot \sigma_x \cdot \sigma_y$	$\dfrac{\mu_{xy}}{\sigma_x \sigma_y}$	$r_{xy} \cdot \dfrac{\sigma_y}{\sigma_x}$	$r_{xy} \cdot \dfrac{\sigma_x}{\sigma_y}$	Інша відповідь

8. Яка матриця не може служити кореляційною матрицею?

А	Б	В	Г	Д
$\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}$	$\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$	Інша відповідь

Відповіді. 1.В; 2.Б; 3.Г; 4.А; 5.Б; 6.А; 7.Г; 8.В.

Розділ 6. Граничні теореми

Граничні теореми складають основу математичної статистики. Умовно їх можна поділити на дві групи.

Перша група, що називається *законом великих чисел (ЗВЧ)*, встановлює стійкість середніх значень: при великій кількості експериментів їх середній результат вже не є випадковим та може бути завбачений з достатньою точністю.

Друга група, що називається *центральною граничною теоремою (ЦГТ),* встановлює умови, за якими закон розподілу великої кількості випадкових величин необмежено наближається до нормального. Для доведення цих теорем використовується нерівність Чебишова.

6.1. Нерівність Чебишова

Теорема 6.1. Якщо випадкова величина X має обмежені математичне сподівання $M(X)$ і дисперсію $D(X)$, то для будь-якого $\varepsilon > 0$ справедлива нерівність Чебишова:

$$P\left\{ \left| X - M(X) \right| \geq \varepsilon \right\} \leq \frac{D(X)}{\varepsilon^2}. \tag{6.1}$$

Іноді нерівність виду (6.1) зручніше використовувати у вигляді:

$$P\left\{ \left| X - M(X) \right| < \varepsilon \right\} \geq 1 - \frac{D(X)}{\varepsilon^2}. \tag{6.2}$$

Приклад. За допомогою нерівності Чебишова дати оцінку того, що відхилення випадкової величини X від свого математичного сподівання буде менш ніж три середніх квадратичних відхилення, тобто менш ніж $3\sigma(X)$.

Розв'язання. $\varepsilon = 3\sigma(X)$. За формулою (6.2):

$$P\left\{ \left| X - M(X) \right| < 3\sigma(X) \right\} \geq 1 - \frac{\sigma^2(X)}{\left(3\sigma(X)\right)^2};$$

$$P\{\,|X-M(X)|<3\sigma(X)\}\ge 1-\frac{1}{9};$$
$$P\{\,|X-M(X)|<3\sigma(X)\}\ge 0{,}8889.$$

Якщо випадкова величина X має біноміальний розподіл з математичним сподіванням $M(X)=np$ та дисперсією $D(X)=npq$, то нерівність Чебишова має вигляд:

$$\underline{P\{|m-np|\ge\varepsilon\}\le\frac{npq}{\varepsilon^2},} \tag{6.3}$$

або
$$\underline{P\{|m-np|<\varepsilon\}\ge 1-\frac{npq}{\varepsilon^2}.} \tag{6.4}$$

Для відносної частоти $\frac{m}{n}$ події в n незалежних експериментах, в кожному з яких ця подія може відбуватися з імовірністю p, нерівність Чебишова набуває вигляду:

$$\underline{P\left\{\left|\frac{m}{n}-p\right|\ge\varepsilon\right\}\le\frac{pq}{n\varepsilon^2},} \tag{6.5}$$

або
$$\underline{P\left\{\left|\frac{m}{n}-p\right|<\varepsilon\right\}\ge 1-\frac{pq}{n\varepsilon^2}.} \tag{6.6}$$

Завдання

1. Залізнична станція споживає для опалення щоденно в середньому $350\,м^3$ газу. Яке можна очікувати споживання кількості газу за один день з імовірністю, не меншою ніж 0,96, якщо середнє квадратичне відхилення споживання газу становить $16\,м^3$?

<u>Відповідь.</u> Від 270 до $430\,м^3$.

2. Випадкова величина X має закон розподілу $N(-1;2)$. За допомогою нерівності Чебишова оцінити ймовірність $|X-a|<\varepsilon$, якщо $\varepsilon=4\sigma$.

<u>Відповідь.</u> 0,9375.

3. Імовірність появи випадкової події A в кожному із 500 незалежних експериментів є величиною сталою і дорівнює 0,8. Нехай X – число появ події A. Використовуючи нерівність Чебишова, оцінити ймовірність події $|X - M(X)| < \varepsilon$, якщо $\varepsilon = 10$.

Відповідь. 0,2.

4. Яке потрібно взяти значення ε у нерівності Чебишова, щоб $P\{|X - M(X)| < \varepsilon\} \approx 0{,}99$, коли відомо, що $D(X) = 9$.

Відповідь. 30.

5. Випадкова подія A може здійснитися при одному експерименті з імовірністю p. Експеримент повторили n раз. Яка ймовірність того, що при цьому виконується нерівність $np - 4\sqrt{npq} < m < np + 4\sqrt{npq}$ (m – число появ події A).

Відповідь. 0,9375.

6. За допомогою нерівності Чебишова знайти ймовірність того, що при киданні монети 500 раз кількість появи герба буде знаходитись в межах від 200 до 300.

Відповідь. 0,95.

6.2. Теорема Чебишова

Теорема 6.2. *(ЗВЧ у формі П.Л. Чебишова.)* Нехай задано n незалежних випадкових величин $X_1, X_2, ..., X_n$, які мають обмежені математичні сподівання $M(X_i)$ $(i = 1,2,...,n)$ і дисперсії яких $D(X_i)$ не перевищують деякої сталої $C > 0$, тобто $D(X_i) \le C$. Тоді для будь-якого наперед взятого числа $\varepsilon > 0$ імовірність того, що абсолютна величина відхилення середнього арифметичного цих величин

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{6.7}$$

від середнього арифметичного їх математичних сподівань

$$M\left(\overline{X}\right)=\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}=\frac{1}{n}\sum_{i=1}^{n}M(X_i) \qquad (6.8)$$

є меншою ε, прямуватиме до одиниці при необмеженому збільшенні числа n:

$$\lim_{n\to+\infty}P\left\{\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<\varepsilon\right\}=1, \qquad (6.9)$$

або

$$\lim_{n\to+\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}M(X_i)\right|<\varepsilon\right\}=1. \qquad (6.10)$$

Дійсно, застосуємо нерівність Чебишова (6.2) до випадкової величини $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$. Якщо $D(X_i)\le C, i=1,2,...,n$, то маємо

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=$$

$$=\frac{1}{n^2}\left[D(X_1)+D(X_2)+...+D(X_n)\right]\le\frac{1}{n^2}\left[C+C+...+C\right]=\frac{1}{n^2}\cdot nC=\frac{C}{n}$$

і користуючись формулою (6.2) отримаємо:

$$P\left\{\left|\overline{X}-M(\overline{X})\right|<\varepsilon\right\}\ge1-\frac{D(\overline{X})}{\varepsilon^2};$$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}M(X_i)\right|<\varepsilon\right\}\ge1-\frac{C}{n\varepsilon^2}. \qquad (6.11)$$

Тоді при $n\to+\infty$

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

Оскільки ймовірність не може бути більшою за одиницю, а нерівність є не строгою, отримаємо нерівність (6.10).

Підкреслимо *значення теореми Чебишова.* При великій кількості n незалежних випадкових величин $X_1, X_2, ..., X_n$ практично достовірно, що їх середнє арифметичне $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ – *випадкова* величина, що як завгодно мало відрізняється від *невипадкової* величини $M\left(\overline{X}\right) = \frac{1}{n}\sum_{i=1}^{n} M(X_i)$, тобто практично перестає бути випадковою величиною.

Наслідок. Якщо незалежні випадкові величини $X_1, X_2, ..., X_n$ мають однакове математичне сподівання $M(X_i) = a$ $(i = 1, 2, ..., n)$, а їх дисперсії обмежені деякою сталою $C > 0$ $\left(D(X_i) \le C,\ i = 1, 2, ..., n\right)$, то формула (6.10) набуває вигляду:

$$\lim_{n \to +\infty} P\left\{ \left| \frac{1}{n}\sum_{i=1}^{n} X_i - a \right| < \varepsilon \right\} = 1, \tag{6.12}$$

або

$$\frac{1}{n}\sum_{i=1}^{n} X_i \underset{n \to +\infty}{\overset{P}{\to}} a. \tag{6.13}$$

Приклад. Дисперсія кожного із 5000 незалежних величин, що мають один і той самий закон розподілу ймовірностей, не перевищує 4. Оцінити ймовірність того, що відхилення середнього арифметичного цих величин від середнього арифметичного їх математичних сподівань, взяте за абсолютною величиною, не перевищить 0,3.

Розв’язання. За формулою (6.11) отримаємо:

$$P\left\{ \left| \frac{1}{5000}\sum_{i=1}^{5000} X_i - \frac{1}{5000}\sum_{i=1}^{5000} M\left(X_i\right) \right| < 0{,}3 \right\} \ge 1 - \frac{4}{5000 \cdot 0{,}3^2} \ge 0{,}991.$$

Завдання

Скільки необхідно провести вимірів, щоб з імовірністю, не меншою ніж 0,9976 можна було стверджувати, що похибка середньої арифметичної результатів цих вимірів не перевищує 0,01, якщо вимір характеризується дисперсією 0,0008?

Відповідь. 3333.

6.3. Теорема Бернуллі

Теорема 6.3. *(ЗВЧ у формі Бернуллі.)* Якщо ймовірність появи випадкової події A в кожному з n незалежних експериментів є величиною сталою і дорівнює p, то при необмеженому збільшенні числа експериментів імовірність того, що абсолютна величина відхилення відносної частоти появи випадкової події $w(A)$ від імовірності p є меншою $\varepsilon > 0$, прямуватиме до одиниці:

$$\lim_{n\to+\infty} P\left\{\left|w(A) - p\right| < \varepsilon\right\} = 1, \tag{6.14}$$

тобто відносна частота $w(A)$ події A наближається за ймовірністю до ймовірності p події A:

$$w(A) \underset{n\to+\infty}{\overset{P}{\to}} p(A). \tag{6.15}$$

Нагадаємо, що нерівність Чебишова для теореми Бернуллі має вигляд:

$$\underline{P\left\{\left|w(A) - p\right| < \varepsilon\right\} \ge 1 - \frac{pq}{n\varepsilon^2}.} \tag{6.16}$$

Завдання

1. Знайти ймовірність того, що при киданні монети 800 раз відносна частота появи герба відхилиться від імовірності появи герба при одному киданні за модулем не більш ніж на 0,1.

Відповідь. 0,96875.

2. Імовірність того, що виготовлена робітником деталь є стандартною, 0,97. Контролю підлягає 500 деталей. Оцінити ймовірність відхилення відносної частоти появи стандартної деталі $w(A)$ від імовірності 0,97 не більш ніж на 0,02.

Відповідь. 0,8545.

3. Бензоколонка заправляє легкові та вантажні машини. Імовірність того, що проїжджаючий легковий автомобіль буде

заправлятися, дорівнює 0,4. Знайти межі, в яких з імовірністю не менш ніж 0,81 знаходиться відносна частота легкових автомобілів, що заправилися протягом двох годин, якщо за цей час заправилось всього 150 автомобілів.

Відповідь. $0{,}3084 < w(A) < 0{,}4917$.

4. У середньому 10% населення деякого регіону користується приміським електричним транспортом. Знайти ймовірність того, що серед 10000 мешканців міста рівень користування приміським електричним транспортом знаходиться в межах від 9 до 11%.

Відповідь. 0,91.

6.4. Центральна гранична теорема

Центральна гранична теорема (ЦГТ) належить до другої групи теорем, що установлює зв'язок між законом розподілу суми випадкової величини та його граничною формою – нормальним законом розподілу (див. п. 4.7).

Теорема 6.4. Нехай задано n незалежних випадкових величин $X_1, X_2, ..., X_n$, кожна із яких має один і той самий закон розподілу ймовірностей із $M(X_i) = a$, $D(X_i) = \sigma^2, i = 1,2,...,n.$ Тоді функція розподілу центрованої та нормованої суми цих випадкових величин при необмеженому збільшенні $n \to +\infty$ наближається до функції розподілу стандартної нормальної випадкової величини:

$$Z_n = \frac{\sum_{i=1}^{n} X_i - M\left(\sum_{i=1}^{n} X_i\right)}{\sqrt{D\left(\sum_{i=1}^{n} X_i\right)}} = \frac{\sum_{i=1}^{n} X_i - na}{\sigma\sqrt{n}}, \tag{6.17}$$

$$F_{Z_n}(x) = P\{Z_n < x\} \underset{n\to+\infty}{\to} F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \tag{6.18}$$

З відношення (6.17) випливає, що при достатньо великому n сума Z_n наближено розподілена за нормальним законом $N(0;1)$.

Тобто сума $S_n = X_1 + X_2 + ... + X_n$ наближено розподілена за нормальним законом $N(na; \sqrt{n}\sigma)$.

З формули (6.18) випливає наближена формула для обчислення ймовірності попадання суми великого числа n випадкових величин до заданого інтервалу:

$$P(\alpha < S_n < \beta) \approx \Phi\left(\frac{\beta - M(S_n)}{\sigma(S_n)}\right) - \Phi\left(\frac{\alpha - M(S_n)}{\sigma(S_n)}\right), \qquad (6.19)$$

де $S_n = \sum_{i=1}^{n} X_i$, $M(S_n) = na$, $\sigma(S_n) = \sqrt{D(S_n)} = \sigma\sqrt{n}$.

Як правило, ЦГТ використовують при $n > 10$.

Приклад. Незалежні випадкові величини X_i розподілені рівномірно на відрізку $[0;2]$. Знайти: а) закон розподілу випадкової величини $Y = \sum_{i=1}^{n} X_i$; б) $P(50 < Y < 80)$.

Розв'язання. Умови ЦГТ виконуються, тому випадкова величина Y має щільність розподілу $f(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \cdot e^{-\frac{(y-a_y)^2}{2\sigma_y}}$. Величини X_i розподілені рівномірно на відрізку $[0;2]$, тому $M(X_i) = \frac{0+2}{2} = 1; D(X_i) = \frac{(2-0)^2}{12} = \frac{1}{3}$. Як наслідок,

$$a_y = M\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} M(X_i) = 100 \cdot 1 = 100;$$

$$D(Y) = D\left(\sum_{i-1}^{n} X_i\right) = \sum_{i=1}^{n} D(X_i) = 100 \cdot \frac{1}{3} = \frac{100}{3};$$

$$\sigma_y = \frac{10}{\sqrt{3}}.$$

Звідси $f(y) = \frac{3}{10\sqrt{6\pi}} \cdot e^{-\frac{3(y-100)^2}{200}}$. За формулою (6.19) знаходимо

$$P(50 < Y < 80) = \Phi\left(\frac{80-100}{\frac{10}{\sqrt{3}}}\right) - \Phi\left(\frac{50-100}{\frac{10}{\sqrt{3}}}\right) =$$

$$= \Phi(-3{,}4641) - \Phi(-8{,}6603) = -\Phi(3{,}4641) + 0{,}5 = -0{,}4997 + 0{,}5 = 0{,}0003.$$

Наслідками ЦГТ є розглянуті в розд. 2 локальна та інтегральна теореми Муавра – Лапласа.

Завдання

Потяг складається з 50 ваг. Маса кожного з них є випадковою величиною X_i з математичним сподіванням $M(X_i) = 78\,\text{т}$ і середнім квадратичним відхиленням $\sigma(X_i) = 5\,\text{т}$, $i = 1,2,...,50$. Локомотив на даній ділянці може везти масу, не більшу за 4000 т. Якщо маса потяга перевищує допустиму, то необхідно причіплювати другий локомотив. Знайти ймовірність того, що одного локомотива не досить для перевезення потяга.

Відповідь. 0,0025.

Питання до теми

1. Які граничні теореми ви знаєте?
2. Який зміст закону великих чисел? Які форми закону великих чисел вам відомі?
3. Який результат доводить центральна гранична теорема? У чому полягає її прикладне значення?
4. Які з відомих вам результатів є наслідками центральної граничної теореми?

Розділ 7. Вибірки та їх основні характеристики

7.1. Предмет і задачі математичної статистики

Наука, яка займається розробкою методів збору й обробки дослідних даних з метою виявлення закономірностей масових випадкових явищ, називається *математичною статистикою*.

Предметом математичної статистики є вивчення випадкових величин (або випадкових подій, процесів) за результатами спостережень.

Вкажемо основні *задачі*, які розв'язує математична статистика:

1) розроблення методів збору та групування статистичних даних, одержаних дослідним шляхом;

2) визначення закону розподілу випадкової величини або системи випадкових величин за статистичними даними;

3) визначення невідомих параметрів розподілу;

4) перевірка правдоподібності гіпотез про закон розподілу випадкової величини, про форму зв'язку між випадковими величинами або про значення параметра, який оцінюють.

Отже, *основна задача* математичної статистики – розроблення методів аналізу статистичних даних у залежності від мети дослідження.

7.2. Генеральна та вибіркова сукупності

Нехай потрібно дослідити будь-яку якісну або кількісну ознаку, властиву великій групі однорідних об'єктів. Наприклад, якщо досліджують партію деталей, то якісною ознакою може бути стандартність або нестандартність кожної деталі, а кількісною ознакою – розмір деталі. Кількісні ознаки бувають неперервними та дискретними.

Перевірку сукупності деталей можна провести двома способами:

1) провести перевірку усіх деталей;
2) перевірити лише певну частину деталей.

Якщо дослідити усі деталі неможливо, тоді відбирають із усієї сукупності обмежене число деталей і перевіряють лише їх.

За результатами вивчення цієї невеликої частини деталей одержують з достатньою ймовірністю необхідну інформацію про всю партію деталей. Такий метод дослідження має назву *вибіркового*.

Уся сукупність об'єктів N, яку вивчають, називається *генеральною сукупністю.* Частина об'єктів n, що випадково відібрана з неї для перевірки, називається *вибірковою сукупністю*, або *вибіркою* ($n < N$). *Об'ємом* сукупності (вибіркової або генеральної) називається кількість об'єктів цієї сукупності (n – об'єм вибірки, N – об'єм генеральної сукупності). Наприклад, якщо з 1000 виробів для дослідження взято 50, тоді об'єм генеральної сукупності $N = 1000$, а об'єм вибірки $n = 50$.

Вибірки бувають *повторні* та *безповторні*. *Повторною* називається вибірка, при якій відібраний об'єкт повертається до генеральної сукупності перед відбором іншого об'єкта. Вибірка називається *безповторною*, якщо взятий об'єкт до генеральної сукупності не повертається. Найчастіше зустрічаються безповторні вибірки.

Вибірку можна ефективно використовувати для вивчення відповідної ознаки генеральної сукупності лише тоді, коли дані вибірки правильно відображають цю ознаку, тобто вибірка повинна бути *репрезентативною (представницькою).* Згідно із законом великих чисел теорії ймовірностей можна стверджувати, що вибірка буде репрезентативною лише тоді, коли її здійснюють випадково. Кожен об'єкт вибірки вважається відібраним випадково з генеральної сукупності, якщо усі її об'єкти мають однакову ймовірність потрапити до вибірки.

7.3. Способи відбору

Усі способи відбору, що використовують у практичній діяльності, можна поділити на два види:

1) відбір, який не потребує розділення генеральної сукупності на частини. До цього виду відбору належать:

- простий випадковий безповторний відбір;
- простий випадковий повторний відбір;

2) відбір, при якому генеральна сукупність розділяється на частини. До цього виду належать:

- типовий відбір (об'єкти відбирають не з усієї генеральної сукупності, а лише з її типових частин);
- механічний відбір (генеральна сукупність механічно поділяється на стільки частин, скільки має бути об'єктів у вибірці, та з кожної частини випадковим чином відбирають один об'єкт);
- серійний відбір (об'єкти із генеральної сукупності відбирають не по одному, а серіями, які і досліджують).

7.4. Варіаційний ряд розподілу

Нехай із генеральної сукупності взята вибірка об'єктів для вивчення ознаки X, яка набула значення x_1 – n_1 раз, значення x_2 – n_2 раз, … , x_m – n_m раз. Значення $x_1, x_2, ..., x_m$ називають *варіантами* ознаки X. Варіанти, що записані у зростаючому порядку, називають *варіаційним рядом.* Кількість спостережень варіант $n_1, n_2, ..., n_m$ називають *рядом частот*.

Нагадаємо, що сума усіх частот повинна дорівнювати об'єму вибірки:

$$\sum_{i=1}^{m} n_i = n. \tag{7.1}$$

Відношення частоти n_k варіанти x_k до об'єму вибірки n називається *відносною частотою* варіанти x_k і позначається $w_k = \frac{n_k}{n}$, причому

$$\sum_{i=1}^{m} w_i = 1. \tag{7.2}$$

Розглянемо таблицю, у верхньому рядку якої запишемо варіаційний ряд ознаки X, а в нижньому – частоти n_i (або відносні частоти w_i) кожного значення x_i. Така таблиця називається емпіричним законом розподілу (статистичним рядом розподілу) або *варіаційним рядом розподілу*, тобто:

x_i	x_1	x_2	$\dots$	x_m
n_i	n_1	n_2	$\dots$	n_m

або

x_i	x_1	x_2	$\dots$	x_m
w_i	$\frac{n_1}{n}$	$\frac{n_2}{n}$	$\dots$	$\frac{n_m}{n}$

де n – об'єм вибірки, $x_1 < x_2 < \ldots < x_m$.

Якщо число варіант досить велике або ж вивчається неперервна ознака X, то варіаційний ряд розподілу стає незручним для користування. У цьому випадку на основі варіаційного ряду розподілу складають *інтервальний ряд розподілу*. Для цього довжину $R = x_{\max} - x_{\min}$, яка називається *розмахом варіації*, ділять точками на декілька однакових (іноді – неоднакових) інтервалів, і відносять до кожного інтервалу відповідні варіанти та вказують відповідні суми частот цих варіант. Інтервальний ряд розподілу можна подати у такому вигляді:

інтервали	$x_1 \div x_2'$	$x_2' \div x_3'$	$x_3' \div x_4'$	$\dots$	$x_l' \div x_{l+1}'$
частота, n_i'	n_1'	n_2'	n_3'	$\dots$	n_l'

де n_i' – сума частот, що потрапили до i-го інтервалу. Для інтервального ряду виконується властивість:

$$\sum_{i=1}^{l} n_i' = n. \tag{7.3}$$

Приклад. У квитковій касі проведено $n = 50$ замірів часу обслуговування пасажирів, *хв*: 1; 2; 7; 5; 3; 4; 6; 10; 1; 3; 2; 1; 3; 3; 4; 5; 6; 4; 3; 10; 1; 8; 7; 1; 3; 2; 2; 5; 11; 6; 1; 2; 3; 3; 4; 6; 3; 3; 4; 1; 4; 3; 2; 7; 2; 3; 1; 11; 3; 5. Необхідно: а) побудувати варіаційний ряд розподілу для ознаки X – тривалість обслуговування кожного пасажира; б) інтервальний ряд розподілу, розділивши розмах варіації на 4 однакові інтервали.

Розв'язання.

а) для складання варіаційного ряду розподілу розмістимо варіанти у порядку зростання і обчислимо, скільки разів зустрічається кожна з них у проведеній вибірці (n_i):

x_i, *хв*	1	2	3	4	5	6	7	8	10	11
n_i	8	7	13	6	4	4	3	1	2	2

Перевірка: $\sum_i n_i = 8+7+13+6+4+4+3+1+2+2=50$;

б) розмах варіації $R = x_{\max} - x_{\min} = 11-1=10$, то обчислимо довжину кожного інтервалу $h=\frac{R}{4}=\frac{10}{4}=2{,}5.$ Інтервальний розподіл має вигляд:

інтервали	$1\div3{,}5$	$3{,}5\div6$	$6\div8{,}5$	$8{,}5\div11$
частота, n_i'	28	14	4	4

Перевірка: $\sum_i n_i' = 28+14+4+4=50.$

Зауваження. Якщо варіанта потрапляє на межу інтервалу, то її включають лише один раз до лівого чи правого інтервалу, керуючись певним принципом.

7.5. Емпірична функція розподілу та її властивості

Якщо задано статистичний розподіл частот деякої ознаки X, n – загальна кількість спостережень (об'єм вибірки), n_x – кількість спостережень, при яких спостерігалися ознаки X менше ніж x, тоді відносна частота події $X<x$ дорівнює $\frac{n_x}{n}$. Зі зміною x змінюється і відносна частота, тобто $\frac{n_x}{n}$ є функцією від x. Ця функція знаходиться емпіричним шляхом, тому її називають *емпіричною*.

Емпіричною функцією розподілу (функцією розподілу вибірки) називається функція $F^*(x)$, яка визначає для кожного значення x відносну частоту події $X < x$, тобто

$$F^*(x) = \frac{n_x}{n}, \tag{7.4}$$

де n_x – сумарна частота варіант, які менше від x, n – об'єм вибірки.

Інтегральна функція розподілу $F(x)$ генеральної сукупності у математичній статистиці називається *теоретичною функцією* розподілу. Вона відрізняється від емпіричної функції розподілу $F^*(x)$ тим, що визначає ймовірність події $X < x$ (тобто $F(x) = P(X < x)$), а не відносну частоту цієї події. Нагадаємо, що відносна частота $F^*(x) = \dfrac{n_x}{n}$ події $X < x$ при $n \to +\infty$ прямує до ймовірності $F(x) = P(X < x)$ цієї події. Тому $F(x)$ та $F^*(x)$ мало відрізняються одна від одної.

Властивості емпіричної функції $F^*(x)$:

1) $0 \le F^*(x) \le 1$;
2) $F^*(x)$ – неспадна функція;
3) $$F^*(x) = \begin{cases} 0 & \text{при } x \le x_{\min}, \\ 1 & \text{при } x > x_{\max}, \end{cases} \tag{7.5}$$

де $x_{\min}$ – найменша варіанта; $x_{\max}$ – найбільша варіанта.

Приклади

1. Знайти емпіричну функцію за даним розподілом вибірки:

x_i	1	2	3	4	5	6	7	8	10	11
n_i	8	7	13	6	4	4	3	1	2	2

Розв’язання. У нашому випадку $n = 50,\ x_{\min} = 1,\ x_{\max} = 11.$

За другою властивістю емпіричної функції розподілу $F^*(x) = 0$ при $x \le 1$. Значення $X < 2$, а саме $x_1 = 1$ зустрічається 8 раз, тому $F^*(x) = \dfrac{8}{50} = 0{,}16$ при $1 < x \le 2$. Значення $X < 3$, а саме $x_1 = 1,\ \ x_2 = 2$ зустрічається 8+7=15 раз, тому $F^*(x) = \dfrac{15}{50} = 0{,}3$ при $2 < x \le 3$. І так далі. Оскільки $x_{\max} = 11$, то $F^*(x) = 1$ при $x > 1$. Графік емпіричної функції розподілу зображено на рис. 7.1, а формула має вигляд:

$$F^*(x) = \begin{cases} 0 & \text{при} \quad x \le 1, \\ 0{,}16 & \text{при} \quad 1 < x \le 2, \\ 0{,}30 & \text{при} \quad 2 < x \le 3, \\ 0{,}56 & \text{при} \quad 3 < x \le 4, \\ 0{,}68 & \text{при} \quad 4 < x \le 5, \\ 0{,}76 & \text{при} \quad 5 < x \le 6, \\ 0{,}84 & \text{при} \quad 6 < x \le 7, \\ 0{,}90 & \text{при} \quad 7 < x \le 8, \\ 0{,}92 & \text{при} \quad 8 < x \le 10, \\ 0{,}96 & \text{при} \quad 10 < x \le 11, \\ 1 & \text{при} \quad x > 11. \end{cases}$$

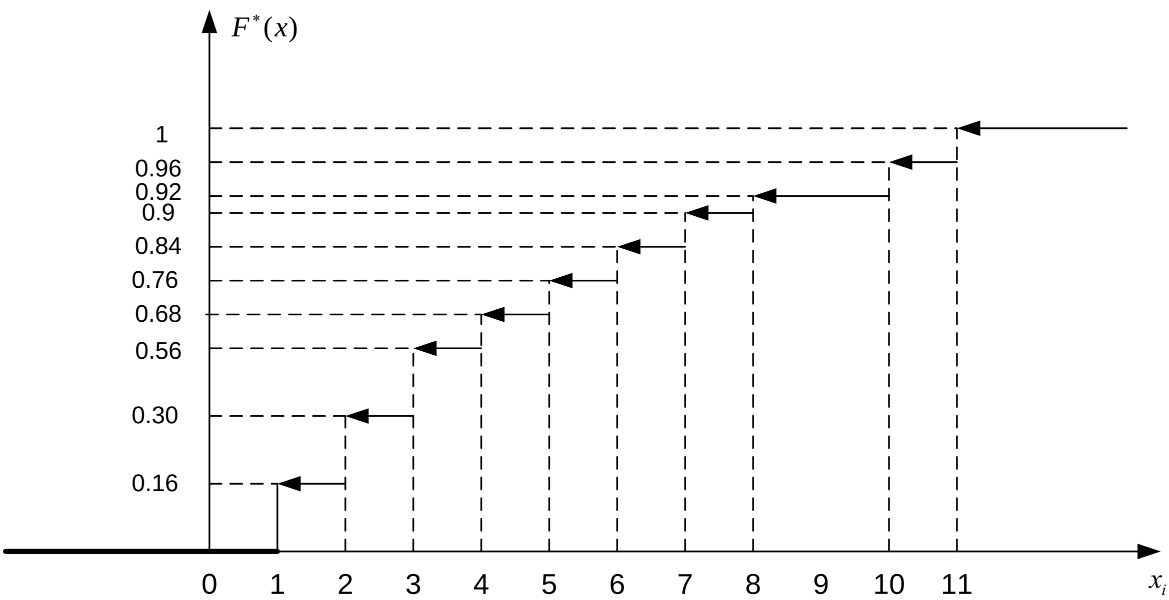


Рис.7.1. Графік емпіричної функції розподілу

2. Знайти емпіричну функцію за даним інтервальним рядом розподілу вибірки:

Інтервали	$1 \div 3,5$	$3,5 \div 6$	$6 \div 8,5$	$8,5 \div 11$
Частота, n_i	28	14	4	4
Накопичувана відносна частота, $\frac{n_i'}{n}$	$\frac{28}{50} = 0,56$	$\frac{28+14}{50} = 0,84$	$\frac{28+14+4}{50} = 0,92$	$\frac{28+14+4+4}{50} = 1$

<u>Розв'язання.</u> Записуємо у третій рядок даної таблиці накопичувану відносну частоту відповідно до кожного інтервалу та будуємо графік емпіричної функції $F^*(x)$ (рис 7.2).

Запишемо отриману функцію:

$$F^*(x) = \begin{cases} 0 & \text{при} \quad x \le 1, \\ 0,224(x-1) & \text{при} \quad 1 < x \le 3,5, \\ 0,112(x-3,5)+0,56 & \text{при} \quad 3,5 < x \le 6, \\ 0,032(x-6)+0,84 & \text{при} \quad 6 < x \le 8,5, \\ 0,032(x-8,5)+0,92 & \text{при} \quad 8,5 < x \le 11, \\ 1 & \text{при} \quad x > 11. \end{cases}$$

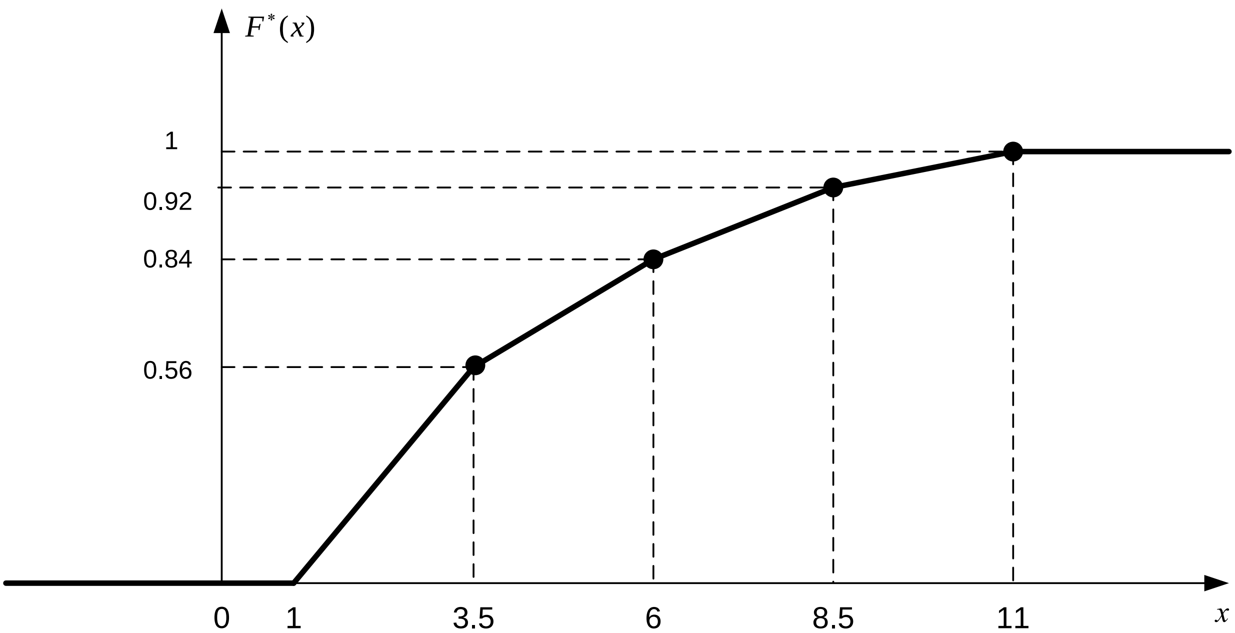


Рис.7.2. Графік емпіричної функції розподілу

Завдання

1. Вибірку задано у вигляді розподілу частот:

x_i	4	7	8	12
n_i	5	2	3	10

Знайти розподіл відносних частот, емпіричну функцію розподілу та побудувати її графік.

2. Маємо статистичний розподіл часу між потягами, що прибувають для переробки:

Інтервали $t_i \div t_{i+1}$, *год.*	0÷0,15	0,15÷0,3	0,3÷0,45	0,45÷0,6
Частота, n_i	30	20	16	12

Інтервали $t_i \div t_{i+1}$, *год.*	0,6÷0,75	0,75÷0,9	0,9÷1,2	1,2÷1,35
Частота, n_i	11	7	3	1

Знайти емпіричну функцію розподілу та побудувати її графік.

7.6. Графічне зображення статистичних розподілів. Полігон. Гістограма

Якщо в результаті вибірки отримали варіаційний ряд розподілу ознаки X, яку треба дослідити, тобто перелік варіант ознаки $x_1, x_2, ..., x_m$ та відповідних їм частот $n_1, n_2, ..., n_m$ (або відносних частот $w_1, w_2, ..., w_m$), то значення варіант та частот (або відносних частот) можна розглядати як координати точок $A_1(x_1;n_1), A_2(x_2;n_2), ..., A_m(x_m;n_m)$ або $A_1(x_1;w_1)$, $A_2(x_2;w_2), ..., A_m(x_m;w_m)$.

Полігоном частот називається ламана, відрізки якої з'єднують точки $(x_1;n_1), (x_2;n_2), ..., (x_m;n_m)$.

Полігоном відносних частот називається ламана, відрізки якої з'єднують точки $(x_1;w_1), (x_2;w_2), ..., (x_m;w_m)$.

Приклад. Побудувати полігон частот, полігон відносних частот за даним розподілом вибірки:

x_i	2	3	5	6
n_i	10	15	5	20

Розв'язання. Відкладаючи у системі координат точки $A_1(2;10)$, $A_2(3;15)$, $A_3(5;5)$, $A_4(6;20)$ та з'єднуючи їх відрізками прямих, одержимо полігон частот цієї вибірки (рис. 7.3).

Для побудови полігона відносних частот отримаємо варіаційний ряд розподілу відносних частот:

x_i	2	3	5	6
w_i	0,2	0,3	0,1	0,4

Відкладаємо в системі координат *xow* точки $A_1(2;0{,}2)$, $A_2(3;0{,}3)$, $A_3(5;0{,}1)$, $A_4(6;0{,}4)$ та з'єднуємо їх відрізками. Отримаємо полігон відносних частот (рис. 7.4).

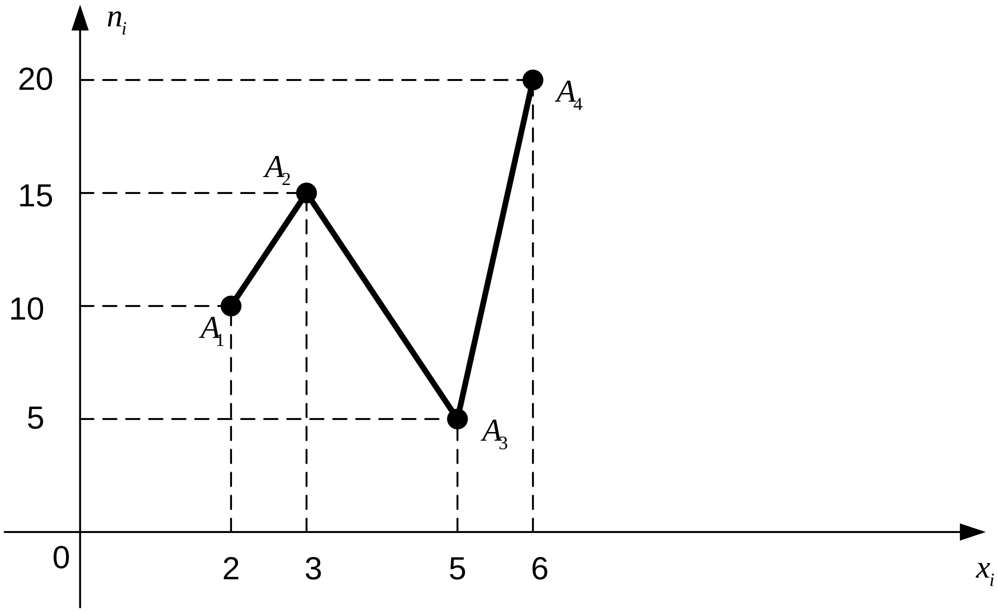


Рис. 7.3. Полігон частот

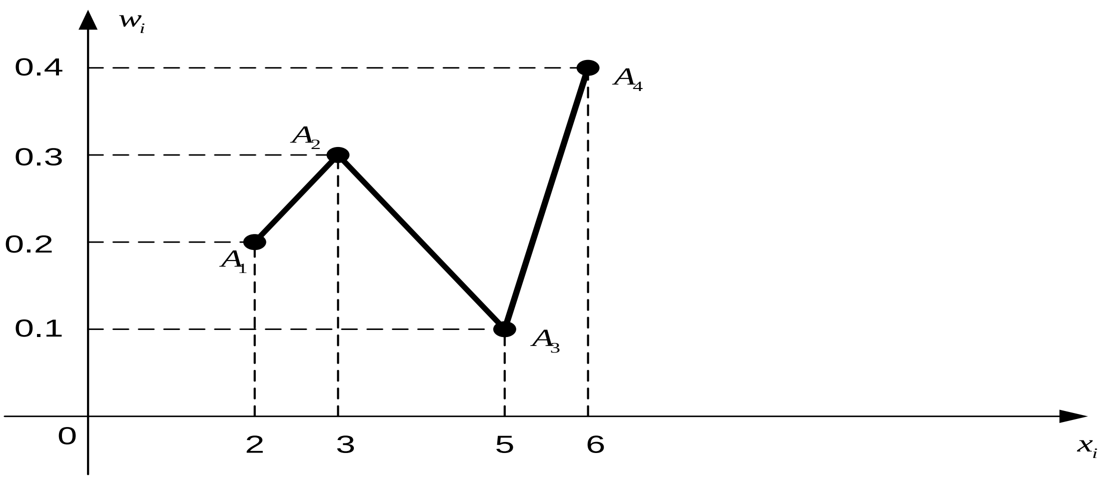


Рис.7.4. Полігон відносних частот

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

Гістограмою відносних частот називається східчаста фігура, яка складається з прямокутників, основами яких є часткові інтервали варіант, а висоти дорівнюють відношенню $\frac{w_k}{h}$ (щільність відносної висоти).

Площа гістограми частот дорівнює об'єму вибірки, а площа гістограми відносних частот – одиниці. Якщо в k-му відрізку $(k = 1,2,...)$ кількість варіант, що спостерігали, з урахуванням їх частот дорівнює n_k (або w_k), то будують прямокутник Π_k, основою якого буде k-й відрізок довжиною h, а висотою – $\frac{n_k}{h}$ (або $\frac{w_k}{h}$). Площа такого прямокутника дорівнює $h \cdot \frac{n_k}{h} = n_k$ (або $h \cdot \frac{w_k}{h} = w_k$). Тому площа всіх прямокутників буде дорівнювати: $S = \sum_{k=1}^{m} n_k = n$ (або у випадку гістограми відносних частот $S = \sum_{k=1}^{m} w_k = 1$).

Гістограму будують у випадку неперервної ознаки, що вивчається. Для цього інтервал, на якому знаходяться всі значення ознаки, що спостерігається, розбивають на декілька часткових інтервалів довжиною h і для кожного інтервалу знаходять суму частот n_i варіант, що потрапили в i-й інтервал. Довжину часткового інтервалу h потрібно вибирати так, щоб побудований ряд не був занадто громіздким і в той же час дозволяв виявити характерні риси зміни значень випадкової величини, що вивчається.

Розбиття інтервалу зміни значень ознаки випадкової величини на часткові інтервали здійснюється так:

1) визначається розмах варіації R:

$$\underline{R = x_{\max} - x_{\min}}; \tag{7.6}$$

2) число інтервалів s вибирається в залежності від числа спостережень n за формулою

$$s = [1 + 3{,}222 \lg n], \tag{7.7}$$

де $[1 + 3{,}222 \lg n]$ – ціла частина числа $1 + 3{,}222 \lg n$, n – об'єм вибірки;

3) визначається довжина часткового інтервалу за формулою

$$h = \frac{R}{s-1}; \tag{7.8}$$

4) за початок першого інтервалу рекомендується брати величину

$$x_{поч} = x_{\min} - 0{,}5h; \tag{7.9}$$

5) верхня границя останнього інтервалу $x_{кінц}$ повинна задовольняти вимогу:

$$(x_{кінц} - h) \le x_{\max} < x_{кінц}. \tag{7.10}$$

Проміжні інтервали отримують додаючи до кінця попереднього інтервалу довжину часткового інтервалу h.

Далі необхідно визначити інтервальні частоти, тобто кількість значень ознаки, що потрапили в кожний частковий інтервал. При цьому в інтервал включається значення більші або рівні нижній границі і менші верхньої границі.

Зауваження. Вигляд гістограми відносних частот дає уявлення про щільність імовірностей випадкової величини, яку досліджують. Гістограма відносних частот – це наближений графік щільності ймовірностей випадкової величини, тобто наближений графік диференціальної функції її розподілу.

Приклади

1. За розподілом вибірки, наданим у табл. 7.1, побудувати гістограми частот та відносних частот.

Таблиця 7.1

Номер інтервалу i	Частковий інтервал $x_i \div x_{i+1}$	Сума частот варіант інтервалу n_i
1	1÷5	50
2	5÷9	20
3	9÷13	12
4	13÷17	10
5	17÷21	8

Розв’язання. Довжина кожного часткового інтервалу дорівнює $h = x_{i+1} - x_i = 4$. Відповідно до кожного інтервалу обчислимо значення щільності частот $\frac{n_i}{h}$ та щільності відносних частот $\frac{w_i}{h}$. За отриманими даними побудуємо табл. 7.2, гістограму частот (рис. 7.5) та гістограму відносних частот (рис. 7.6).

Таблиця 7.2

Номер інтервалу i	Щільність частоти інтервалу $\frac{n_i}{h}$	Щільність відносної частоти інтервалу $\frac{w_i}{h}$
1	12,5	0,125
2	5,0	0,05
3	3,0	0,03
4	2,5	0,025
5	2,0	0,02

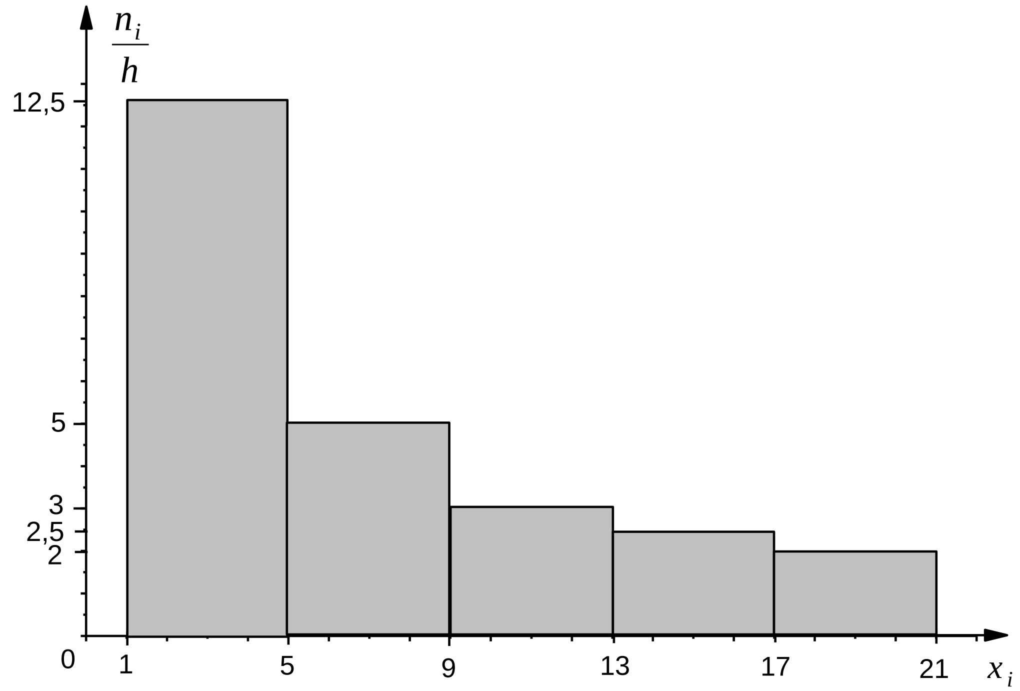


Рис.7.5. Гістограма частот

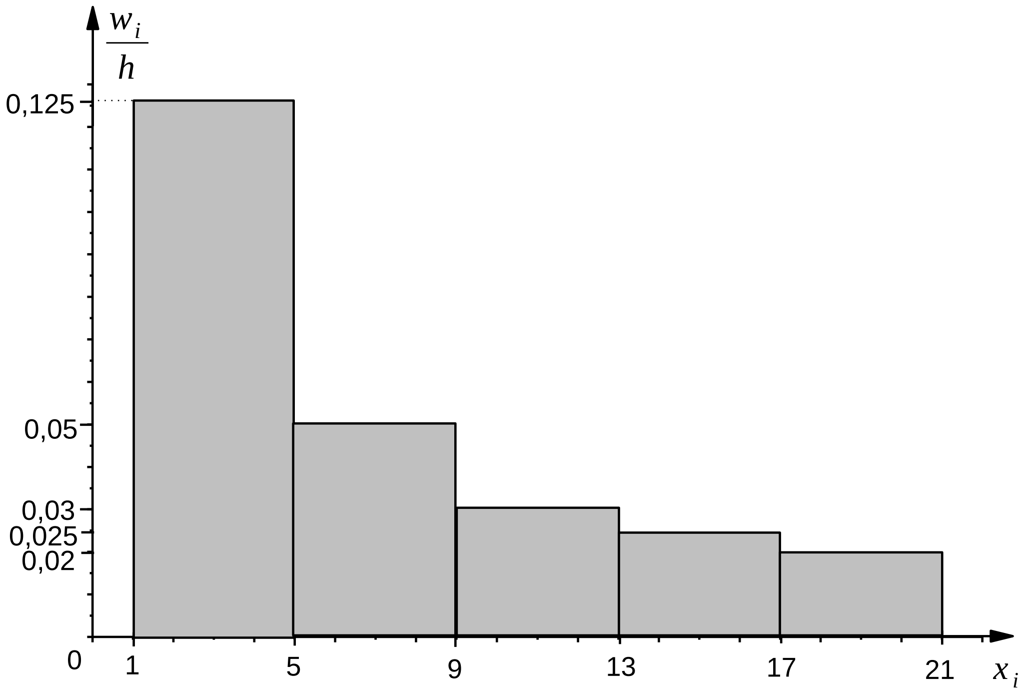


Рис.7.6. Гістограма відносних частот

2. Для вивчення попиту додаткової зупинки автобуса на заданому маршруті було спостережено щоденну кількість пасажирів, що користуються цією зупинкою протягом 100 днів (табл. 7.3). Побудувати гістограму відносних частот.

Таблиця 7.3

2	10	15	18	21	23	26	29	32	37
3	10	15	18	21	23	26	29	32	38
4	11	15	18	21	23	26	29	33	38
5	12	16	19	21	24	26	29	33	39
5	12	16	19	22	24	27	30	34	40
6	12	16	19	22	24	27	30	34	40
7	13	17	19	22	24	27	31	35	42
8	14	17	20	22	25	28	31	35	43
9	14	17	20	22	25	28	31	36	46
9	14	17	20	23	25	28	32	37	49

Розв’язання:

1) за формулою (7.6) обчислюємо розмах варіації: $R = 49 - 2 = 47$;
2) за (7.7) обчислюємо кількість інтервалів: $s = [1 + 3{,}222 \lg 100] = 7$;
3) за (7.8) отримаємо довжину часткового інтервалу: $h = \frac{47}{6} \approx 8$;
4) за (7.9) $x_{поч} = 2 - 0{,}5 \cdot 8 = -2$.

Для побудови гістограми заповнюємо табл. 7.4.

Таблиця 7.4

Номер інтервалу i	Частковий інтервал $x_i \div x_{i+1}$	Сума частот варіант інтервалу n_i	Відносна частота інтервалу w_i	Щільність відносної частоти інтервалу w_i / h
1	-2÷6	5	0,05	0,00625
2	6÷14	12	0,12	0,015
3	14÷22	27	0,27	0,03375
4	22÷30	30	0,30	0,0375
5	30÷38	17	0,17	0,0225
6	38÷46	7	0,07	0,00875
7	46÷54	2	0,02	0,0025
$\sum$		100	1	

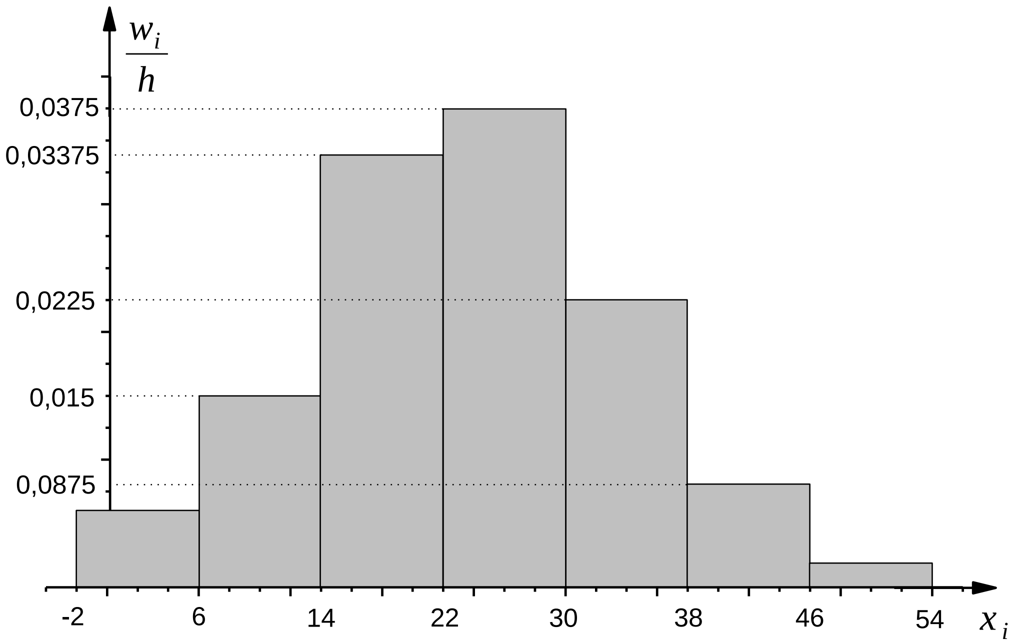


Рис. 7.7. Гістограма відносних частот розподілу

Завдання

1. Побудувати полігон частот, полігон відносних частот та емпіричну функцію розподілу за даним розподілом вибірки:

x_i	2	4	5	6	7	8	10
n_i	10	15	5	20	12	18	20

2. Побудувати: а) гістограми частот та відносних частот; б) емпіричну функцію розподілу за інтервальним розподілом вибірки:

Частковий інтервал, $x_i \div x_{i+1}$	1÷5	5÷9	9÷13	13÷17	17÷21
Сума частот варіант інтервалу, n_i	10	20	50	12	8

3. На телефонній станції проводилися спостереження за числом неправильних з'єднань за хвилину. Отримали такі результати: 3; 1; 3; 4; 2; 1; 2; 4; 6; 3; 0; 2; 2; 0; 2; 5; 4; 3; 3; 1; 4; 5; 2; 5; 1; 2; 1; 6; 3; 4. Для даної вибірки необхідно побудувати:

а) варіаційний ряд розподілу;

б) полігон частот;

в) полігон відносних частот;
г) гістограму частот;
д) гістограму відносних частот;
е) емпіричну функцію розподілу та її графік.

7.7. Числові характеристики вибіркової сукупності

Розглянутий варіаційний ряд та емпірична функція дають вичерпну характеристику статистичних даних. Проте іноді достатньо знати лише окремі ознаки даного варіаційного ряду. Числа, які є кількісним виразом таких ознак, називаються *числовими характеристиками вибірки.*

До основних числових характеристик вибірки належать: *вибіркова середня, вибіркова дисперсія та вибіркове середнє квадратичне відхилення*. Ці числові характеристики ознаки X проведеної вибірки є аналогами відповідно математичного сподівання, дисперсії та середнього квадратичного відхилення випадкової величини X .

Нехай задано варіаційний ряд розподілу ознаки X проведеної вибірки об'ємом n:

x_i	x_1	x_2	$\dots$	x_m
n_i	n_1	n_2	$\dots$	n_m

Вибірковою середньою ознаки X вибірки називається число $\overline{x}_в$, що обчислюється за формулою

$$\overline{x}_в = \frac{1}{n} \cdot \sum_{i=1}^{m} x_i \cdot n_i. \quad (7.11)$$

Іноді, коли частоти кожної варіанти, яка належить варіаційному ряду розподілу, дорівнюють одиниці, вибіркова середня дорівнює середньому арифметичному усіх варіант:

$$\overline{x}_в = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i. \quad (7.12)$$

Основні властивості вибіркової середньої:

1) при множенні усіх варіант вибірки на однаковий множник вибіркова середня також множиться на цей множник:

$$\frac{1}{n}\cdot\sum_{i=1}^{m} n_i\cdot\left(Cx_i\right)=\frac{C}{n}.\cdot\sum_{i=1}^{m} n_i\cdot x_i=C\cdot\overline{x}_в; \qquad (7.13)$$

2) якщо додати (відняти) до усіх варіант вибірки однакове число, то вибіркова середня зростає (зменшується) на це число:

$$\frac{1}{n}\cdot\sum_{i=1}^{m} n_i\cdot\left(x_i\pm C\right)=\frac{1}{n}.\cdot\sum_{i=1}^{m} n_i\cdot x_i\pm\frac{C}{n}\cdot\sum_{i=1}^{m} n_i=\overline{x}_в\pm C. \qquad (7.14)$$

Вибірковою дисперсією $D_в$ називається число, яке дорівнює середньому арифметичному квадратів відхилень варіант від вибіркової середньої:

$$\underline{D_в=\frac{1}{n}\cdot\sum_{i=1}^{m}\left(x_i-\overline{x}_в\right)^2\cdot n_i.} \qquad (7.15)$$

Підкреслимо, що вибіркову дисперсію зручніше обчислювати за формулою

$$\underline{D_в=\frac{1}{n}\cdot\sum_{i=1}^{m} x_i^2 n_i-\left(\overline{x}_в\right)^2.} \qquad (7.16)$$

Вибірковим середнім квадратичним відхиленням називається квадратний корінь з вибіркової дисперсії:

$$\underline{\sigma_в=\sqrt{D_в}.} \qquad (7.17)$$

Середнє геометричне $\overline{x}_{геом}$ – це значення, яке використовується головним чином для розрахунку середніх темпів зростання і приросту при вивченні динаміки певного процесу:

$$\underline{\overline{x}_{геом}=\sqrt[n]{x_1^{n_1}\cdot x_2^{n_2}\cdot\ldots\cdot x_m^{n_m}}.} \qquad (7.18)$$

Вибірка також характеризується допоміжними числовими характеристиками: коефіцієнтом варіації, модою та медіаною.

Коефіцієнт варіації обчислюється за формулою

$$\nu = \frac{\sigma_в}{\left|\bar{x}_в\right|} \cdot 100\% \qquad (7.19)$$

і використовується для порівняння величин розсіювання по відношенню до вибіркової середньої двох варіаційних рядів.

Модою M_0 називається та варіанта варіаційного ряду, яка має найбільшу частоту.

Медіаною M_e називається варіанта варіаційного ряду розподілу, яка рівновіддалена від його кінців. Якщо до варіаційного ряду розподілу входить парна кількість варіант, то за моду беруть середнє арифметичне двох сусідніх варіант, що знаходяться посередині варіаційного ряду.

Коефіцієнт асиметрії a_s – безрозмірна величина, яка характеризує ступінь схилення варіаційного ряду розподілу щодо його центра симетрії вправо або вліво та обчислюється за формулою

$$a_s = \frac{\mu_3}{\sigma_в^3}, \qquad (7.20)$$

де

$$\mu_3 = \frac{1}{n} \cdot \sum_{i=1}^{m} \left(x_i - \bar{x}_в \right)^3 \cdot n_i \; – \qquad (7.21)$$

центральний момент третього порядку. Залежно від напряму схилення коефіцієнт симетрії є додатним (при схиленні у правий бік) і від’ємним (при схиленні у лівий бік). На рис. 7.8 зображено графік щільності розподілу для таких випадків: а) розподіл симетричний; б) асиметрія лівостороння; в) асиметрія правостороння.

Вважається, що криві з абсолютною величиною коефіцієнта асиметрії $|a_s| > 0{,}5$ мають значне зміщення; якщо $|a_s| \le 0{,}25$ – асиметрія незначна.

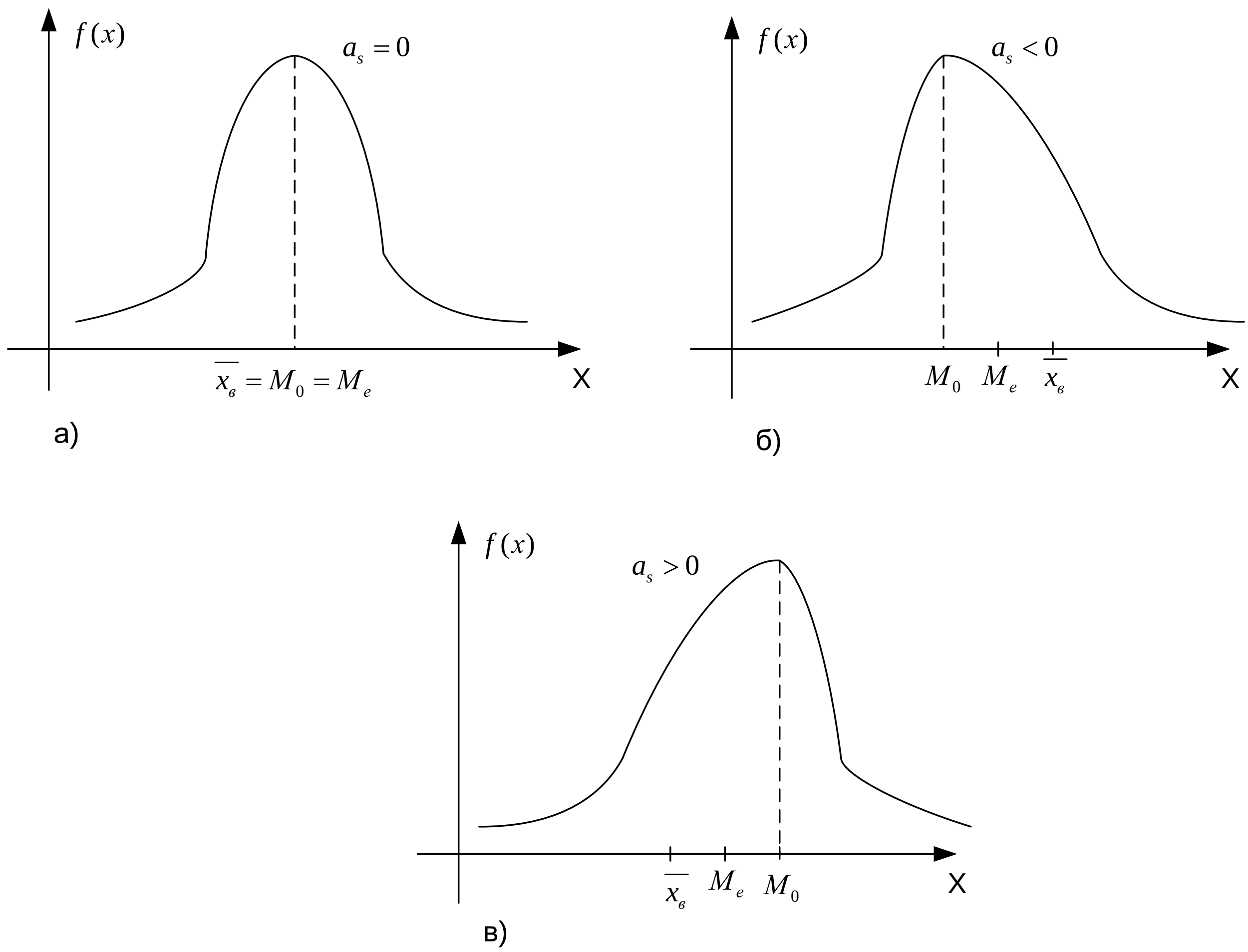


Рис. 7.8

Ексцесом вибірки e_k називається безрозмірна величина, яка виражає ступінь концентрації елементів вибірки навколо її середнього, обчислюється за формулою

$$e_k = \frac{\mu_4}{\sigma_в^4} - 3, \tag{7.22}$$

де

$$\mu_4 = \frac{1}{n} \cdot \sum_{i=1}^{m} \left(x_i - \overline{x}_в \right)^4 \cdot n_i \; - \tag{7.23}$$

центральний момент четвертого порядку.

Залежно від знака ексцесу при додатному його значенні графік щільності розподілу є високо- або гостровершинним, а при від’ємному значенні – низько- або плосковершинним. На рис. 7.9: $e_k = 0$ – нормально вершинний; $e_k > 0$ - високовершинний; $e_k < 0$ – низьковершинний.

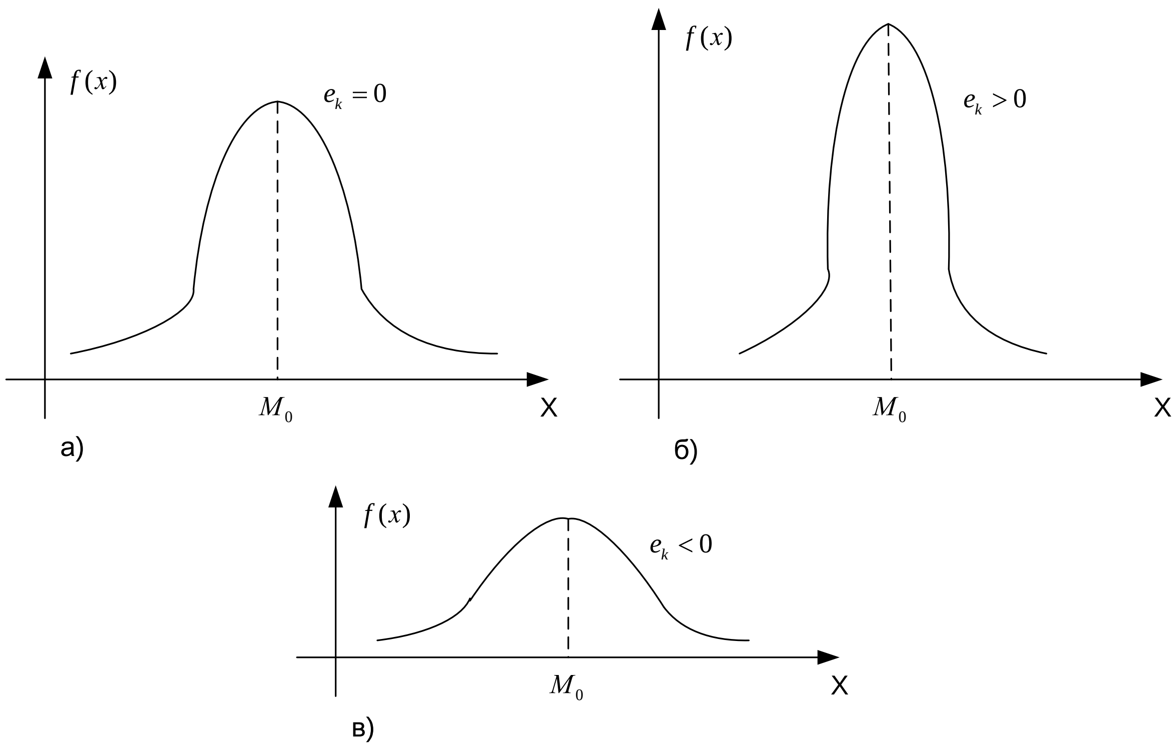


Рис. 7.9

Приклад. За даними розподілу числа вагонів у відчепі, що надходять на колії сортувального парку:

x_i, число вагонів у відчепі	1	2	3	4	5	6	7
n_i, число спостережень	35	40	32	20	10	4	1

Обчислити вибіркову середню $\bar{x}_в$; вибіркову дисперсію $D_в$; вибіркове середнє квадратичне відхилення $\sigma_в$; моду M_0; медіану M_e; коефіцієнт асиметрії a_s; ексцес вибірки e_k.

<u>Розв'язання.</u> Об'єм вибірки дорівнює $n = \sum n_i = 142$. За формулою (7.11) обчислюємо:

$$\bar{x}_в = \frac{1}{142} \cdot (1 \cdot 35 + 2 \cdot 40 + 3 \cdot 32 + 4 \cdot 20 + 5 \cdot 10 + 6 \cdot 4 + 7 \cdot 1) \approx 2{,}62.$$

Знаходимо вибіркову дисперсію $D_в$ за формулою (7.16):

$$D_в = \frac{1}{142} \cdot \left(1^2 \cdot 35 + 2^2 \cdot 40 + 3^2 \cdot 32 + 4^2 \cdot 20 + 5^2 \cdot 10 + 6^2 \cdot 4 + 7^2 \cdot 1\right) - (2{,}62)^2 \approx$$
$$\approx 1{,}9102$$

та вибіркове середнє квадратичне відхилення за (7.17):

$$\sigma_в = \sqrt{1{,}9102} \approx 1{,}3821.$$

Мода $M_0 = 2$; медіана $M_e = 4$.

За (7.21) та (7.20) обчислюємо коефіцієнт асиметрії:

$$\mu_3 = \frac{1}{142}((1-2{,}62)^3 \cdot 35 + (2-2{,}62)^3 \cdot 40 + (3-2{,}62)^3 \cdot 32 + (4-2{,}62)^3 \cdot 20 +$$
$$+(5-2{,}62)^3 \cdot 10 + (6-2{,}62)^3 \cdot 4 + (7-2{,}62)^3 \cdot 1) \approx 1{,}8963;$$

$$a_s = \frac{1{,}8963}{(1{,}3821)^3} \approx 0{,}7183.$$

За (7.23) та (7.22) визначимо ексцес:

$$\mu_4 = \frac{1}{142}((1-2{,}62)^4 \cdot 35 + (2-2{,}62)^4 \cdot 40 + (3-2{,}62)^4 \cdot 32 + (4-2{,}62)^4 \cdot 20 +$$
$$+(5-2{,}62)^4 \cdot 10 + (6-2{,}62)^4 \cdot 4 + (7-2{,}62)^4 \cdot 1) \approx 10{,}7827;$$

$$e_k = \frac{10{,}7827}{(1{,}3821)^4} - 3 \approx -0{,}0049.$$

Аналіз одержаних характеристик свідчить про те, що даний вибірковий матеріал є майже симетричним, незначно схиленим у правий бік $(a_s > 0)$ і низьковершинним $(e_k < 0)$.

Зауваження. Якщо вибірка задана у вигляді інтервального статистичного розподілу:

Частковий інтервал, $x_i \div x_{i+1}$	$x_1 \div x_2$	$x_2 \div x_3$	$\ldots$	$x_m \div x_{m+1}$
Частота варіант інтервалу, n_i	n_1	n_2	$\ldots$	n_m

то для обчислення $\overline{x}_в$; $D_в$; $\sigma_в$ треба побудувати варіаційний ряд вибірки:

z_i	z_1	z_2	$\ldots$	z_m
n_i	n_1	n_2	$\ldots$	n_m

де $z_i = \frac{x_i + x_{i+1}}{2}, i = 1,2,\ldots,m$ і для нього знайти числові характеристики.

У наступних розділах буде наведено спрощення обчислення наданих характеристик.

Завдання

1. Знайти $\overline{x}_в$; $D_в$; $\sigma_в$; M_0; M_e та коефіцієнт варіації ν ознаки X, яка характеризується вибіркою:

x_i	2	5	7	10
n_i	8	40	10	2

<u>Відповідь.</u> $\overline{x}_в = 5{,}1$; $D_в = 2{,}69$; $\sigma_в \approx 1{,}64$; $M_0 = 5$; $M_e = 6$; $\nu \approx 32{,}16\%$.

2. На станції спостерігалась нижченаведена тривалість очікування розформування составів у парку приймання:

Довжина інтервалу, x	$0 \div 10$	$10 \div 20$	$20 \div 30$	$30 \div 40$	$40 \div 50$	$50 \div 60$
Частота n_i	32,5	25,6	20,3	12,6	7,7	1,3

Знайти: $\overline{x}_в$; $D_в$; $\sigma_в$; M_0; M_e; коефіцієнт варіації ν; коефіцієнт асиметрії a_s; ексцес e_k.

<u>Відповідь.</u> $\overline{x}_в = 19{,}13$; $D_в \approx 176{,}24$; $\sigma_в \approx 13{,}28$; $M_0 = 5$; $M_e = 30$; $\nu \approx 69{,}42\%$; $a_s \approx 0{,}66$; $e_k \approx -0{,}52$.

Питання до теми

1 Що вивчає математична статистика? Як визначаються її основні задачі?

2 Що називається вибірковим методом? Які способи відбору ви знаєте?

3 Які типи вибірок ви знаєте?

4 Що таке варіаційний ряд розподілу? Як ще вони називаються?

5 Чому дорівнює сума частот вибірки? Сума відносних частот?

6 Як побудувати інтервальний ряд розподілу?

7 Що називається розмахом варіації?

8 Як визначається емпірична функція розподілу? Який вигляд має її графік? Наведіть основні властивості емпіричної функції розподілу.

9 Як можна графічно зобразити емпіричний розподіл?

10 Як побудувати гістограму? Чому дорівнює її площа? Про що говорить її форма?

11 Для чого використовується формула Стреджерса? Наведіть приклад її застосування.

12 Які числові характеристики вибіркової сукупності ви знаєте? Наведіть формули для обчислення та їх основні властивості. Наведіть приклади.

Тестові питання

1. Яка з наведених послідовностей є варіаційним рядом?

А	Б	В	Г	Д
2, 5, 1, 6, 17	10, 9, 8, 5, 1	2, 4, 7, 8, 10	3, 5, 7, 2, 4	Інша відповідь

2. Чому дорівнює сума відносних частот вибірки?

А	Б	В	Г	Д
0	1	n	$\frac{1}{n}$	Інша відповідь

3. Чому дорівнює сума частот вибірки, заданої статистичним розподілом

x_i	0	1	3	4	5	7	10
n_i	1	4	5	10	5	4	1

А	Б	В	Г	Д
30	20	10	1	Інша відповідь

4. Чому дорівнює об’єм вибірки, розподіл якої задано таблицею

x_i	0	8	13	15
n_i	13	20	17	10

А	Б	В	Г	Д
10	20	50	60	Інша відповідь

5. Чому чисельно дорівнює площа гістограми частот вибірки?

А	Б	В	Г	Д
n^2	n	1	0	Інша відповідь

6. Чому чисельно дорівнює площа гістограми відносних частот вибірки, заданої розподілом

x_i	0	1	3	4	5	7	10	11	12	14	15
n_i	1	4	5	10	15	14	21	18	15	12	8

А	Б	В	Г	Д
123	113	1	0	Інша відповідь

7. Що можна обчислити за допомогою формули Стреджерса?

А	Б	В	Г	Д
Площу гістограми	Розмах вибірки	Довжину частинних інтервалів	Кількість частинних інтервалів	Інша відповідь

8. Емпіричний розподіл має вигляд

x_i	0	1	5	7	10	11	15
n_i	1	4	15	14	21	18	8

Визначити моду.

А	Б	В	Г	Д
$M_0 = 0$	$M_0 = 7$	$M_0 = 10$	$M_0 = 15$	Інша відповідь

9. Емпіричний розподіл має вигляд

x_i	0	1	5	7	10	11	15
n_i	1	4	15	14	21	18	8

Визначити медіану.

А	Б	В	Г	Д
$M_e = 0$	$M_e = 7$	$M_e = 10$	$M_e = 15$	Інша відповідь

10. Статистичний розподіл має вигляд

x_i	-1	0	1
n_i	20	50	30

Чому дорівнює вибіркова середня?

А	Б	В	Г	Д
$\overline{x}_в = 0{,}1$	$\overline{x}_в = 1$	$\overline{x}_в = -1$	$\overline{x}_в = 0$	Інша відповідь

<u>Відповіді:</u> 1.В; 2.Б; 3.А; 4.Г; 5.Б; 6.В; 7.Г; 8.В; 9.Б; 10.А.

Розділ 8. Елементи теорії оцінок

8.1. Основні вимоги до статистичних оцінок

У практичних задачах часто виникає необхідність дослідити кількісну ознаку X генеральної сукупності, використовуючи результати вибірки. Наприклад, знайти наближені значення математичного сподівання $M(X)$, дисперсії $D(X)$, середнього квадратичного відхилення $\sigma(X)$, початкових або центральних моментів випадкової величини X. Іноді треба оцінити параметри закону розподілу X, який вдалось встановити за даними вибірки. Наприклад, відомо, що випадкова величина X розподілена у генеральній сукупності за нормальним законом, і необхідно оцінити (знайти наближені значення) параметрів a і σ, що повністю характеризують нормальний закон розподілу. Якщо випадкова величина X розподілена за законом Пуассона, то необхідно оцінити лише один параметр λ, яким цей розподіл визначається. Для рівномірно розподіленої випадкової величини X треба наближено знайти сегмент, що містить значення X. Невідомі параметри розподілів ознаки X генеральної сукупності знаходяться за даними вибірки.

Статистичною оцінкою невідомого параметра випадкової величини X генеральної сукупності (теоретичного розподілу X) називається функція від випадкових величин (даних вибірки), що спостерігаються. Щоб статистичні оцінки давали найкращі наближення параметрів, вони повинні задовольняти певні вимоги. Розглянемо ці вимоги.

Нехай θ^* є статистична оцінка невідомого параметра θ теоретичного розподілу. Припустимо, що за вибіркою об'єму n знайдена оцінка θ_1^*. При інших вибірках того ж об'єму одержимо деякі інші оцінки $\theta_2^*, \theta_3^*, ..., \theta_m^*$. Саму оцінку θ^* можна розглядати як випадкову величину, а числа $\theta_1^*, \theta_2^*, \theta_3^*, ..., \theta_m^*$ як її можливі значення. Якщо числа $\theta_k^* (k = 1,2,...,m)$ будуть більші значення θ, тоді оцінка θ^* дає наближене значення θ з *надлишком.* У цьому випадку $M(\theta^*) > \theta.$ Якщо θ^* дає оцінку θ з *недостачею,* тоді $M(\theta^*) < \theta.$ Таким чином, вимога $M(\theta^*) = \theta$ застерігає від систематичних похибок.

Статистична оцінка параметра θ називається *незміщеною,* якщо

$$M(\theta^*) = \theta. \qquad (8.1)$$

Оцінку θ^* називають *зміщеною*, якщо ця рівність не виконується.

Вимога про незміщеність оцінки θ^* є недостатньою тому, що значення θ^* можуть бути значно розсіяні навколо свого середнього значення, тобто дисперсія $D(\theta^*)$ може бути великою. Якщо $D(\theta^*)$ буде малою, тоді можливість припустити значну похибку буде малоймовірною. Тому для статистичної оцінки ставиться вимога про її ефективність.

Ефективною називається така статистична оцінка θ^*, яка при заданому об’ємі вибірки має найменшу можливу дисперсію.

При розгляді вибірки великого об’єму до статистичних оцінок висувають вимогу їх обґрунтованості. *Обґрунтованою* називається статистична оцінка, яка при $n \to +\infty$ прямує за ймовірністю до оцінюваного параметра, тобто

$$\forall \varepsilon > 0 \qquad P\left(\left|\theta - \theta^*\right| < \varepsilon\right) \xrightarrow[n \to +\infty]{} 1. \qquad (8.2)$$

8.2. Точкові оцінки для математичного сподівання та дисперсії

Нехай $X_1, X_2, ..., X_n$ – випадкові величини: X_1 – результат першого спостереження; X_2 – результат другого спостереження і т. д., причому випадкова величина X_i $(i = 1,2,..,n)$ має такий самий розподіл, як і випадкова величина X; конкретна вибірка $x_1, x_2, ..., x_n$ – це значення незалежних випадкових величин $X_1, X_2, ..., X_n$. Статистичною оцінкою θ^* (або просто оцінкою θ^*) параметра θ теоретичного розподілу називають його наближене значення, що залежить від даних вибірки. Очевидно, θ^* є значення деякої функції результатів спостережень над випадковою величиною, тобто

$$\theta^* = \theta^*(X_1, X_2, ..., X_n). \qquad (8.3)$$

Функцію значень спостережень називають *статистикою.*

Припустимо, що вивчається випадкова величина X з математичним сподіванням $M(X)=a$ і дисперсією $D(X)$, причому обидва параметри невідомі.

Точковою оцінкою називається статистика, що використовується як наближене значення параметра генеральної сукупності. Тобто точкова оцінка характеристики генеральної сукупності – це число, що знаходиться за результатами вибірки.

Нехай $x_1, x_2, ..., x_n$ – отримана вибірка в результаті n незалежних спостережень за випадковою величиною X. Щоб підкреслити випадковий характер величин $x_1, x_2, ..., x_n$ перепишемо їх у вигляді $X_1, X_2, ..., X_n$, де X_i – значення випадкової величини X в i–му спостереженні. Випадкові величини $X_1, X_2, ..., X_n$ можна розглядати як n незалежних екземплярів величини X. Тому $M(X_1)=...=M(X_n)=a;\quad D(X_1)=...=D(X_n)=D(X).$

За оцінку a математичного сподівання візьмемо вибіркове середнє

$$\overline{X}_в = \frac{1}{n}\sum_{i=1}^{n} X_i. \qquad (8.4)$$

Покажемо, що ця оцінка є обґрунтованою і незміщеною. Згідно з визначенням незміщеності (8.1) та властивостями математичного сподівання знайдемо

$$M[X_в] = M\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} M(X_i) = \frac{1}{n}\cdot na = a. \qquad (8.5)$$

Згідно із законом великих чисел (теорема Чебишова) при збільшенні n величина $\frac{1}{n}\sum_{i=1}^{n} X_i$ збігається за ймовірністю до a: $\lim\limits_{n\to+\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < \varepsilon\right\} = 1,$ тобто оцінка є обґрунтованою. Можна показати, що за нормальним законом розподілу оцінка $\overline{X}_в$ є також і ефективною.

Можна довести, що для вибіркової дисперсії $D_в = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}_в\right)^2}{n}$ виконується

$$M(D_в) = \frac{n-1}{n} \cdot D(X), \qquad (8.6)$$

тобто $M(D_в) \neq D(X)$. Таким чином вибіркова дисперсія є зміщеною оцінкою дисперсії $D(X)$. Щоб ліквідувати це зміщення, достатньо ввести поправку, помноживши $D_в$ на $\frac{n}{n-1}$:

$$\frac{n}{n-1} D_в = \frac{n}{n-1} \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}_в\right)^2}{n} = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_в\right)^2. \qquad (8.7)$$

Величину

$$s^2 = \frac{n}{n-1} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}_в\right)^2 \qquad (8.8)$$

називають *виправленою вибірковою дисперсією*. Вона є незміщеною оцінкою для $D(X)$.

Зауваження. Далі для простоти будемо заміняти великі X_i на малі x_i, середнє $\overline{X}_в$ на $\overline{x}_в$, але пам'ятати, що x_i є випадковими величинами. Отже, якщо в n незалежних випробуваннях випадкова величина X набуває можливих значень x_i, $i = 1,2,...,n$, то для визначення невідомого математичного сподівання і дисперсії слід використовувати наступні наближені оцінки:

$$\underline{\overline{x}_в = \frac{1}{n} \sum x_i n_i,} \qquad (8.9)$$

$$\underline{s^2 = \frac{n}{n-1} D_в = \frac{n}{n-1} \left(\frac{1}{n} \sum x_i^2 n_i - \left(\overline{x}_в\right)^2 \right),} \qquad (8.10)$$

$$\underline{s = \sqrt{s^2}.} \qquad (8.11)$$

Приклад. При визначенні середнього інтервалу подачі порожніх вагонів на об'єкт за добу було отримано такий розподіл:

Інтервали подачі, *хв*	$0 \div 12$	$12 \div 24$	$24 \div 36$	$36 \div 48$	$48 \div 60$	$60 \div 72$	$72 \div 84$	$84 \div 96$
Кількість поданих вагонів	24	27	18	17	9	3	1	1

Знайти точкові оцінки для математичного сподівання та дисперсії.

<u>Розв’язання.</u> Для обчислення точкових оцінок побудуємо додаткову таблицю:

z_i	6	18	30	42	54	66	78	90
n_i	24	27	18	17	9	3	1	1

Об’єм вибірки дорівнює $n = \sum n_i = 100$. Оцінку математичного сподівання знаходимо за формулою (8.9):

$$\bar{x}_в = \frac{1}{100}\left(6 \cdot 24 + 18 \cdot 27 + 30 \cdot 18 + 42 \cdot 17 + 54 \cdot 9 + 66 \cdot 3 + 78 \cdot 1 + 90 \cdot 1\right) = 27{,}36.$$

Оцінки дисперсії та середньоквадратичного відхилення обчислюємо згідно з (8.10) та (8.11):

$$s^2 = \frac{100}{99} \times$$

$$\times \left[\frac{6^2 \cdot 24 + 18^2 \cdot 27 + 30^2 \cdot 18 + 42^2 \cdot 17 + 54^2 \cdot 9 + 66^2 \cdot 3 + 78^2 \cdot 1 + 90^2 \cdot 1}{100} - \right.$$

$$\left. - (27{,}36)^2\right] = 347{,}869;$$

$$s = \sqrt{347{,}869} \approx 18{,}65.$$

Завдання

1. Задано статистичний розподіл випадкової величини X (тривалість безвідмовної роботи пристроїв):

Тривалість безвідмовної роботи, *год*	$0 \div 5$	$5 \div 10$	$10 \div 15$	$15 \div 20$	$20 \div 25$	$25 \div 30$
Число пристроїв	173	75	30	14	7	1

Вважаючи, що випадкова величина має показниковий розподіл, знайти ймовірність того, що тривалість безвідмовної роботи буде знаходитись в межах від 4 до 7 годин.

<u>Відповідь.</u> 0,2.

8.3. Розрахунок основних числових характеристик вибірки із застосуванням умовних варіант. Метод добутків

Нагадаємо, якщо дискретна (неперервна) випадкова величина задана рядом розподілу (щільністю розподілу $f(x)$), то *початковим моментом* s*-го порядку* для дискретної випадкової величини називається величина

$$M_s = \sum_{i=1}^{n} x_i^s \cdot p_i ; \tag{8.12}$$

для неперервної випадкової величини

$$M_s = \int_{-\infty}^{+\infty} x^s \cdot f(x) dx. \tag{8.13}$$

Із формул (8.12) і (8.13) видно, що математичне сподівання дорівнює першому початковому моменту, тобто

$$M(X) = M_1. \tag{8.14}$$

Центральним моментом s*-го порядку* для дискретної випадкової величини X називається величина

$$\mu_s = \sum_{i=1}^{n} (x_i - a)^s \cdot p_i , \tag{8.15}$$

а для неперервної випадкової величини –

$$\mu_s = \int_{-\infty}^{+\infty} (x - a)^s \cdot f(x) dx. \tag{8.16}$$

З формул (8.15) та (8.16) бачимо, що $\mu_1 = 0$, а дисперсія випадкової величини X дорівнює другому центральному моменту, тобто

$$D(X) = \mu_2. \tag{8.17}$$

Центральні та початкові моменти для випадкових величин пов’язують співвідношення:

$$\begin{aligned} \mu_2 &= M_2 - M_1^2; \\ \mu_3 &= M_3 - 3M_1M_2 + 2M_1^3; \\ \mu_4 &= M_4 - 4M_1M_3 + 6M_1^2M_2 - 3M_1^4. \end{aligned} \qquad (8.18)$$

Для варіаційного ряду, що містить рівновіддалені одна від одної варіанти, розрахунок числових характеристик вибірки спрощується, якщо застосовувати *умовні варіанти*, які визначаються формулою

$$u_i = \frac{x_i - C}{h}, \qquad (8.19)$$

де C – хибний нуль, значення якого вибирають найближчим до середини варіаційного ряду і бажано таким, що має найбільшу частоту; h – крок варіаційного ряду.

Використовуючи значення умовних варіант u_i, обчислимо *умовні моменти* чотирьох порядків за формулами:

$$M_1^* = \frac{1}{h}\sum u_i \cdot n_i; \; M_2^* = \frac{1}{n}\sum u_i^2 n_i; \; M_3^* = \frac{1}{n}\sum u_i^3 n_i; \; M_4^* = \frac{1}{n}\sum u_i^4 n_i. \qquad (8.20)$$

Тепер можна обчислити центральні моменти s-го порядку:

$$\begin{aligned} \mu_2 &= \left(M_2^* - \left(M_1^*\right)^2\right)h^2; \\ \mu_3 &= \left(M_3^* - 3M_1^*M_2^* + 2\left(M_1^*\right)^3\right)h^3; \\ \mu_4 &= \left(M_4^* - 4M_1^*M_3^* + 6M_2^*\left(M_1^*\right)^2 - 3\left(M_1^*\right)^4\right)h^4. \end{aligned} \qquad (8.21)$$

За наведеними умовними моментами першого і другого порядків визначаємо:

1) вибіркове середнє:

$$\bar{x}_в = M_1^* h + C; \qquad (8.22)$$

2) вибіркову дисперсію:

$$D_в = \mu_2 = \left(M_2^* - \left(M_1^*\right)^2\right)h^2; \qquad (8.23)$$

3) середньоквадратичне відхилення:

$$\sigma_в = \sqrt{D_в}; \qquad (8.24)$$

4) коефіцієнт асиметрії:

$$a_s = \frac{\mu_3}{\sigma_в^3}; \qquad (8.25)$$

5) ексцес:

$$e_k = \frac{\mu_4}{\sigma_в^4} - 3. \qquad (8.26)$$

Приклад. Знайти методом добутків вибіркове середнє, вибіркову дисперсію, асиметрію та ексцес за даною вибіркою об'ємом $n = 60$:

x_i	0	1	2	3	4	5	6
n_i	8	17	16	10	6	2	1

<u>Розв'язання.</u> За хибний нуль беремо $C = 2$, крок $h = 1$, $u_i = \frac{x_i - 2}{1}$. Заповнюємо табл. 8.1.

Таблиця 8.1

i	x_i	n_i	u_i	$n_i u_i$	$n_i u_i^2$	$n_i u_i^3$	$n_i u_i^4$
1	0	8	-2	-16	32	-64	128
2	1	17	-1	-17	17	-17	17
3	2	16	0	0	0	0	0
4	3	10	1	10	10	10	10
5	4	6	2	12	24	48	96
6	5	2	3	6	18	54	162
7	6	1	4	4	16	64	256
$\sum$		60		-1	117	95	669

За формулами (8.20), (8.21) обчислюємо умовні і центральні моменти:

$$M_1^* = \frac{1}{h}\sum u_i \cdot n_i = \frac{1}{60}(-1) \approx -0{,}017; \quad M_2^* = \frac{1}{n}\sum u_i^2 n_i = \frac{1}{60}\cdot 117 \approx 1{,}95;$$

$$M_3^* = \frac{1}{n}\sum u_i^3 n_i = \frac{1}{60}\cdot 95 \approx 1{,}583; \quad M_4^* = \frac{1}{n}\sum u_i^4 n_i = \frac{1}{60}\cdot 669 = 11{,}15;$$

$$\mu_2 = \left(1{,}95 - (-0{,}017)^2\right)\cdot 1^2 \approx 1{,}95;$$

$$\mu_3 = \left(1{,}583 - 3\cdot(-0{,}017)\cdot 1{,}95 + 2\cdot(-0{,}017)^3\right)\cdot 1^3 \approx 1{,}68;$$

$$\mu_4 = \left(11{,}15 - 4\cdot(-0{,}017)\cdot 1{,}583 + 6\cdot 1?95\cdot(-0{,}017)^2 - 3\cdot(-0{,}017)^4\right)\cdot 1^4 \approx 11{,}26.$$

За формулами (8.22)-(8.26) обчислимо характеристики:

$$\overline{x}_в = M_1^* h + C = -0{,}017 \cdot 1 + 2 = 1{,}983;$$

$$D_в = \mu_2 = 1{,}95; \quad \sigma_в = \sqrt{D_в} = \sqrt{1{,}95} \approx 1{,}396;$$

$$a_s = \frac{\mu_3}{\sigma_в^3} = \frac{1{,}68}{2{,}722} \approx 0{,}617; \quad e_k = \frac{\mu_4}{\sigma_в^4} - 3 = \frac{11{,}26}{3{,}8025} - 3 \approx -0{,}039.$$

Аналіз одержаних результатів свідчить про те, що даний вибірковий розподіл має значне зміщення у правий бік і є низьковершинним.

8.4. Вирівнювання варіаційних рядів. Метод моментів

На практиці число дослідів завжди обмежене, тому необхідно зважати на те, що будь – якому варіаційному ряду розподілу властиві у певній мірі риси випадковості. Випадковим є східчастий вид емпіричної функції $F^*(x)$ для неперервної випадкової величини; випадковою є також форма гістограми. Проте на практиці незручно користуватися негладкими лініями, тому найчастіше намагаються підібрати аналітичну форму, яка визначає лише істотні риси статистичного матеріалу. Така задача називається *вирівнюванням варіаційних рядів розподілу*. Найчастіше вирівнювання використовується для гістограми відносних частот. Задача вирівнювання зводиться до того, щоб замінити гістограму гладкою кривою, рівняння якої в подальшому можна використовувати як щільність розподілу $f(x)$. Параметри, що входять до виразу функції $f(x)$, підбирають такими, щоб аналітичний вираз, який вирівнює статистичний ряд, якнайкраще наближався до статистичних даних. При цьому користуються різними методами. Найбільш поширеним з них є *метод моментів*. Суть цього методу полягає в тому, щоб найважливіші характеристики (моменти): математичне сподівання, дисперсія тощо вирівнюваного та вирівнюючого розподілів збігалися.

Приклади

1. Випадкова величина X – швидкість співударяння вагонів на сортувальних коліях. Зроблено $n = 200$ спостережень.

Результати спостережень зведені в табл. 8.2. Побудувати гістограму і вирівняти варіаційний розподіл за допомогою нормального закону: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$.

Таблиця 8.2

Номер інтервалу i	Частковий інтервал $x_i \div x_{i+1}$	Сума частот інтервалу n_i	Відносна частота інтервалу w_i	Щільність відносної частоти w_i/h
1	$0{,}73 \div 0{,}81$	1	0,005	0,0625
2	$0{,}81 \div 0{,}89$	6	0,03	0,375
3	$0{,}89 \div 0{,}97$	16	0,08	1
4	$0{,}97 \div 1{,}05$	29	0,145	1,85
5	$1{,}05 \div 1{,}13$	38	0,19	2,375
6	$1{,}13 \div 1{,}21$	60	0,3	3,75
7	$1{,}21 \div 1{,}29$	29	0,145	1,85
8	$1{,}29 \div 1{,}37$	15	0,075	0,9375
9	$1{,}37 \div 1{,}45$	4	0,02	0,25
10	$1{,}45 \div 1{,}53$	2	0,01	0,125
$\sum$		200	1	

Розв’язання. За даними табл. 8.2. будуємо гістограму (рис. 8.1). Її вигляд указує на розподіл, близький до нормального.

За оцінку параметра математичного сподівання a беремо вибіркове середнє $\overline{x}_в = \frac{1}{200}\sum_{i=1}^{10} x_i n_i$, прийнявши за x_i середину відповідного інтервалу: $\overline{x}_в = 1{,}1324$. За оцінку дисперсії беремо вибіркову дисперсію: $\sigma_в^2 = \frac{1}{200}\sum_{i=1}^{10} x_i^2 n_i - \left(\overline{x}_в\right)^2 = 0{,}017 \Rightarrow \sigma_в \approx 0{,}13$. Отже, рівняння вирівнюючої кривої має вигляд: $f(x) = \frac{1}{0{,}13\sqrt{2\pi}} \cdot e^{-\frac{(x-1{,}1324)^2}{0{,}034}}$.

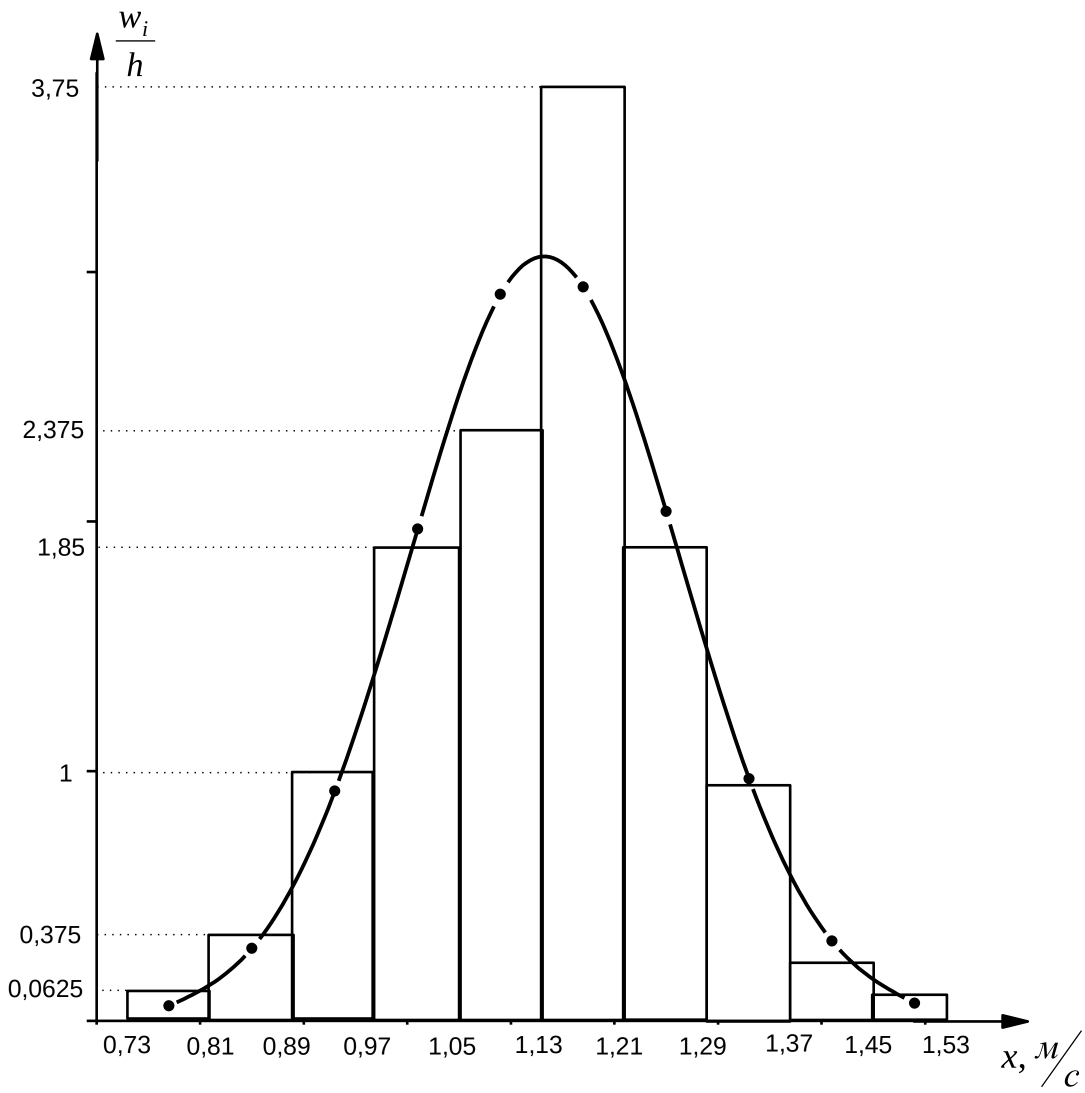


Рис. 8.1. Гістограма відносних частот

За даними табл. 8.3 будуємо криву щільності розподілу $f(x)$ (рис. 8.1).

Таблиця 8.3

Середина інтервалу x_i	0,77	0,85	0,93	1,01	1,09	1,17	1,25	1,33	1,41	1,49
$f(x_i)$	0,06	0,29	0,92	1,97	2,91	2,94	2,04	0,97	0,32	0,07

2. Випадкова величина X (час очікування чергового пасажира, який підійде до квиткової каси) має показниковий

розподіл $f(x)=\lambda e^{-\lambda x}$, де $x \ge 0$. Проведено $n=100$ спостережень. Результати спостережень надані в табл. 8.4. Побудувати гістограму і вирівняти варіаційний розподіл за допомогою показникового закону розподілу.

Таблиця 8.4

Номер інтервалу i	Частковий інтервал $x_i \div x_{i+1}$	Сума частот інтервалу n_i	Відносна частота інтервалу w_i	Щільність відносної частоти w_i/h
1	$0 \div 10$	35	0,35	0,035
2	$10 \div 20$	33	0,33	0,033
3	$20 \div 30$	13	0,13	0,013
4	$30 \div 40$	6	0,06	0,006
5	$40 \div 50$	5	0,05	0,005
6	$50 \div 60$	3	0,03	0,003
7	$60 \div 70$	2	0,02	0,002
8	$170 \div 80$	2	0,02	0,002
9	$80 \div 90$	1	0,01	0,001
$\sum$		100	1	

Розв’язання. За даними табл. 8.4 будуємо гістограму (рис. 8.2). Її вигляд указує на експоненціальний закон розподілу.

Параметр $\lambda=\dfrac{1}{M(X)}$ (див. п. 4.6). В якості оцінки математичного сподівання беремо (як в попередньому прикладі) вибіркове середнє $\bar{x}_в=\dfrac{1}{100}\sum_{i=1}^{9}x_i n_i$, прийнявши за x_i середину відповідного інтервалу:

$$\bar{x}_в=\frac{1}{100}(5\cdot 3{,}5+15\cdot 33+25\cdot 13+35\cdot 6+45\cdot 5+55\cdot 3+65\cdot 2+75\cdot 2+85\cdot 1)=$$
$$=19{,}6.$$

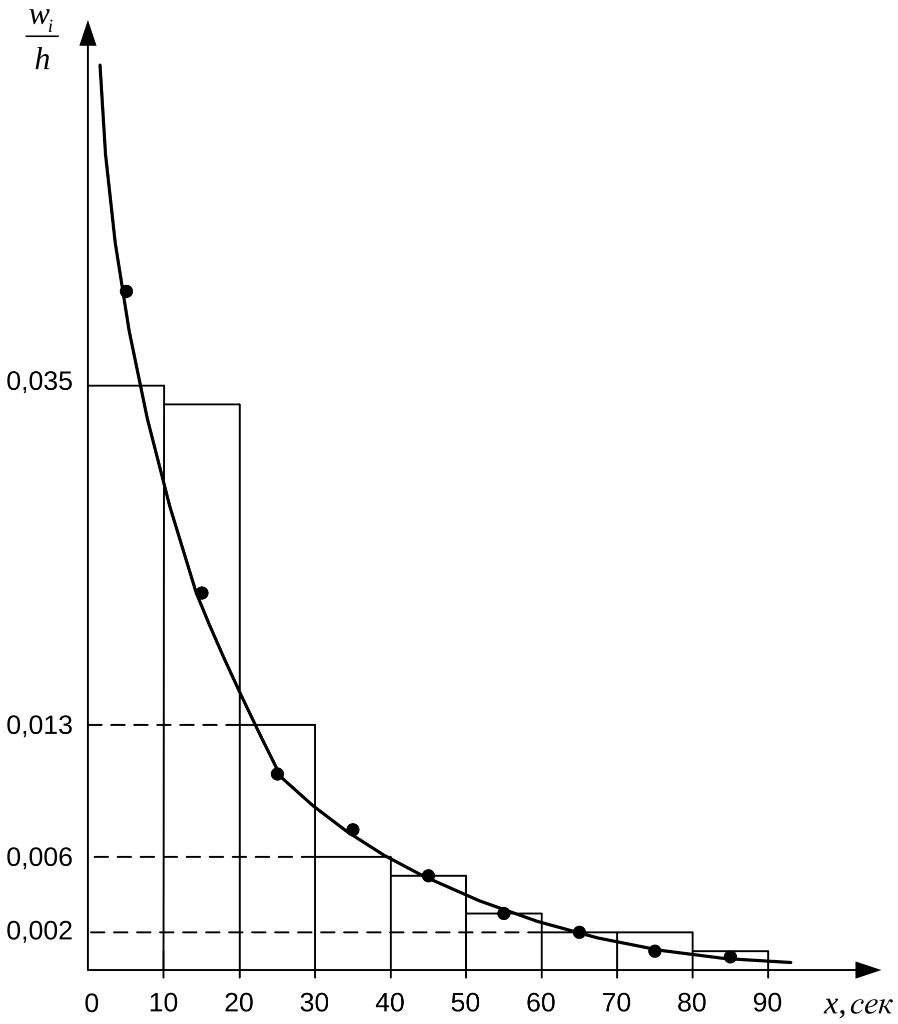


Рис. 8.2

Отримуємо наближене значення параметра $\lambda = \dfrac{1}{19{,}6} \approx 0{,}05$ і відповідну щільність розподілу $f(x) = \begin{cases} 0{,}05e^{-0{,}05x} & \text{при } x \geq 0, \\ 0 & \text{при } x < 0. \end{cases}$

За даним табл. 8.5 будуємо криву (рис 8.2).

Таблиця 8.5

Середина інтервалу x_i	5	15	25	35	45	55	65	75	85
$f(x_i)$	0,04	0,02	0,01	0,009	0,005	0,003	0,002	0,001	0,0007

Завдання

1. Випадкова величина X (похибка вимірювання дальності радіодалекоміром) підлягає рівномірному закону розподілу з невідомими параметрами a і b. За даними таблиці методом моментів знайти точкові оцінки невідомих параметрів a і b рівномірного розподілу. Побудувати гістограму відносних частот та графік функції щільності розподілу $f(x)$.

$x_i \div x_{i+1}$	$2\div 4$	$4\div 6$	$6\div 8$	$8\div 10$	$10\div 12$	$12\div 14$	$14\div 16$	$16\div 18$	$18\div 20$	$20\div 22$
n_i	21	16	15	26	22	14	21	22	18	25

<u>Відповідь.</u> $\overline{x}_в = 12{,}320{,}375$; $D(X) = 33{,}78$; $a = 2{,}265; b = 22{,}355$.

2. Випадкова величина X (швидкість входження у гірковий парк) має нормальний закон розподілу. За даними таблиці методом моментів знайти точкові оцінки невідомих параметрів. Побудувати гістограму і вирівняти варіаційний розподіл за допомогою нормального закону розподілу $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$.

Швидкість входження у гірковий парк x_i, $м/с$	1,95	2,01	2,07	2,13	2,19	2,25	2,31	2,37	2,43
Кількість випадків n_i	2	2	4	6	8	10	16	18	28

<u>Відповідь.</u> $a = 2{,}3; \sigma(X) \approx 0{,}2$.

8.5. Метод максимальної правдоподібності для знаходження оцінок параметрів розподілу

Крім методу моментів, існують і інші методи точкової оцінки невідомих параметрів. До них належить метод максимальної правдоподібності.

Дискретна випадкова величина

Нехай X – дискретна випадкова величина, яка в результаті n експериментів набула значення $x_1, x_2, ..., x_n$. Припустимо, що вид закону розподілу X відомий, але параметр θ, що входить до нього, невідомий. Позначимо $P(X = x_i) = p(x_i, \theta)$. *Функцією правдоподібності дискретної випадкової величини* X називається функція аргументу θ

$$L(x_1, x_2, ..., x_n, \theta) = p(x_1, \theta) \cdot p(x_2, \theta) \cdot ... \cdot p(x_n, \theta), \tag{8.27}$$

де $x_1, x_2, ..., x_n$ – вибіркові значення.

Суть методу максимальної правдоподібності полягає в тому, що за оцінку параметра θ беруть таке його значення θ^*, яке перетворює функцію правдоподібності $L(x_1, x_2, ..., x_n, \theta)$ (8.27) у максимум. Тоді функцію $\theta^* = \theta^*(x_1, x_2, ..., x_n)$ називають оцінкою *максимальної правдоподібності.* Оскільки функції L та $\ln L$ досягають максимуму при одному і тому ж значенні θ, то на практиці зручніше шукати екстремум функції $\ln L$. За правилом аналізу для знаходження максимуму $\ln L$ необхідно:

1) знайти похідну $\frac{d \ln L}{d\theta}$;

2) прирівняти похідну до нуля і знайти стаціонарну точку θ^*;

3) знайти $\left.\frac{d^2 \ln L}{d\theta^2}\right|_{\theta=\theta^*}$, якщо $\left.\frac{d^2 \ln L}{d\theta^2}\right|_{\theta=\theta^*} < 0$, то $\theta = \theta^*$ – точка максимуму.

Знайдену точку θ^* приймають за оцінку максимальної правдоподібності параметра θ. У випадку наявності двох параметрів θ_1 і θ_2 їх оцінки визначаються із двох сумісно розв’язаних рівнянь. Оцінки, одержані за методом максимальної правдоподібності, є обґрунтованими, асимптотично нормально розподіленими, мають найменшу дисперсію в порівнянні з іншими оцінками, але не завжди є незміщеними. Недолік методу полягає в необхідності використання громіздких обчислень.

Зауваження. Оцінки, знайдені за різними методами, можуть не збігатися.

Неперервна випадкові величина

Нехай X – неперервна випадкова величина, яка в результаті n експериментів набула значення $x_1, x_2, ..., x_n$. Припустимо, що вид закону розподілу $f(x,\theta)$ відомий, але невідомий параметр θ, що входить до нього.

Функцією правдоподібності для неперервної випадкової величини X називається функція

$$L(x_1, x_2, ..., x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot ... \cdot f(x_n, \theta), \qquad (8.28)$$

де $x_1, x_2, ..., x_n$ – вибіркові значення.

Оцінку максимальної правдоподібності невідомого параметра θ розподілу неперервної випадкової величини знаходять аналогічно до дискретної випадкової величини.

Завдання

1. Знайти методом максимальної правдоподібності оцінку параметра λ розподілу Пуассона $P(X = k) = \dfrac{\lambda^k \cdot e^{-\lambda}}{k!}$.

Відповідь. $\lambda \approx \bar{x}_в$.

2. Знайти за методом максимальної правдоподібності оцінку параметра p біноміального розподілу $P(X = k) = C_n^k p^k (1-p)^{n-k}$.

Відповідь. $p \approx \dfrac{1}{nk} \cdot \sum_{i=1}^{n} x_i$.

3. Знайти за методом максимальної правдоподібності оцінки параметрів a і σ нормального розподілу $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$.

Відповідь. $a \approx \bar{x}_в, \sigma^2 \approx \sigma_в^2$ (перша оцінка незміщена, а друга – зміщена).

8.6. Поняття про інтервальні оцінки

Обчислена на основі вибірки оцінка θ^* невідомого параметра θ є його наближеним значенням. Точкові оцінки значень невідомих параметрів не дають відповіді на запитання: на яку величину θ^* відрізняється від θ? Це є їх суттєвим недоліком. Одержання відповіді на це питання приводить до необхідності знаходження інтервалу $(\theta_н^*, \theta_в^*)$, в якому знаходиться точне значення оцінюваного параметра θ. Нижня і верхня межі інтервалу обчислюються за результатами спостережень і змінюються від вибірки до вибірки, тому є випадковими значеннями. При цьому ми не можемо гарантувати, що знайдений інтервал містить точне значення невідомого параметра, а можемо лише стверджувати цей факт з певною ймовірністю.

Вибравши ймовірність γ досить близькою до одиниці, можемо поставити задачу про знаходження за результатами спостережень такого інтервалу $(\theta_н^*, \theta_в^*)$, який містить з імовірністю γ невідомий параметр θ. Такий інтервал називають *довірчим* або *надійним*, а ймовірність γ – *надійністю оцінки*:

$$P(\theta_н^* < \theta < \theta_в^*) = \gamma. \qquad (8.29)$$

Довірчий інтервал частіше всього знаходять у вигляді $(\theta^* - \delta; \theta^* + \delta)$. Такий інтервал потрібен для того, щоб оцінити відхилення θ^* від θ. Іншими словами, треба вказати значення δ, щоб виконувалась нерівність $|\theta^* - \theta| < \delta$ з імовірністю γ. Величина δ називається *точністю оцінки*. Тобто формулу (8.29) можна записати у такому вигляді:

$$P(\theta^* - \delta < \theta < \theta^* + \delta) = \gamma. \qquad (8.30)$$

Іноді замість величини γ використовують величину $\alpha = 1 - \gamma$, яку називають *рівнем значущості*. Найчастіше α береться рівним 0,05; 0,01; 0,0001 і показує, з якою ймовірністю висновок про надійність є помилковим.

Зрозуміло, що збільшення надійності приводить до зменшення точності, а підвищення точності приводить до зменшення надійності. Одночасне підвищення точності і надійності можливе лише за рахунок збільшення об’єму вибірки.

8.7. Довірчі інтервали для оцінки математичного сподівання нормального розподілу

Розглянемо приклади побудови довірчих інтервалів для оцінки математичного сподівання a нормально розподіленої генеральної сукупності.

1. Довірчий інтервал для оцінки з надійністю γ математичного сподівання a при умові, що *параметр* σ *відомий*.

Нехай випадкова величина X, математичне сподівання a якої невідоме, в результаті n незалежних експериментів набула значення $X_1,\ldots,X_n$. Точкова оцінка для цього параметра обчислюється за формулою

$$\overline{X}=\frac{1}{n}\cdot\sum_{i=1}^{n}X_i. \tag{8.31}$$

Оскільки $\overline{X}$ є сумою великої кількості випадкових величин X_i, а відповідно до центральної граничної теореми закон розподілу величини X близький до нормального. Відомо, що коли випадкова величина X розподілена за нормальним законом, то вибіркова середня $\overline{X}_в$ також розподілена за нормальним законом. Використовуючи властивості математичного сподівання і дисперсії, знаходимо параметри розподілу $\overline{X}$:

$$M\left(\overline{X}\right)=M\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}M\left(X_i\right)=\frac{1}{n}\cdot na=a; \tag{8.32}$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\left(X_i\right)=\frac{1}{n^2}\cdot n\sigma^2=\frac{\sigma^2}{n}, \tag{8.33}$$

тобто

$$M\left(\overline{X}\right)=a;\ \sigma\left(\overline{X}\right)=\frac{\sigma}{\sqrt{n}}. \tag{8.34}$$

Знайдемо таку величину δ, щоб виконувалося співвідношення

$$P\left(\left|\overline{X}-a\right|<\delta\right)=\gamma, \tag{8.35}$$

де γ – задана надійність.

Згідно з властивістю (4.13) нормально розподіленої випадкової величини (див. п.4.7) і використовуючи формули (8.34), отримаємо

$$P\left(\left|\overline{X}-a\right|<\delta\right)=2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)=2\Phi(t), \tag{8.36}$$

де

$$t=\frac{\delta\sqrt{n}}{\sigma}, \tag{8.37}$$

$\Phi(t)$ – функція Лапласа. З (8.37) знаходимо точність оцінки

$$\delta=\frac{t\sigma}{\sqrt{n}}. \tag{8.38}$$

Тоді (8.36) набуває вигляду

$$P\left(\left|\overline{X}-a\right|<t\frac{\delta}{\sqrt{n}}\right)=2\Phi(t)=\gamma. \tag{8.39}$$

Замінюючи у формулі (8.39) $\overline{X}$ на $\overline{x}_в$, одержимо

$$\underline{P\left(\overline{x}_в-t\frac{\sigma}{\sqrt{n}}<a<\overline{x}_в+t\frac{\sigma}{\sqrt{n}}\right)=2\Phi(t)=\gamma.} \tag{8.40}$$

Зміст співвідношення (8.40) такий: з імовірністю γ можна стверджувати, що довірчий інтервал $\left(\overline{x}_в-t\frac{\sigma}{\sqrt{n}};\overline{x}_в+t\frac{\sigma}{\sqrt{n}}\right)$ містить невідомий параметр a, при цьому точність оцінки дорівнює $\delta=\frac{t\sigma}{\sqrt{n}}$. Параметр t знаходимо за допомогою дод. 2 як аргумент функції Лапласа, якому відповідає значення $\Phi(t)=\frac{\gamma}{2}$.

2. Довірчий інтервал для оцінки з надійністю γ математичного сподівання a при умові, що *параметр* σ *невідомий*.

Оцінка середнього квадратичного відхилення має вигляд

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2. \qquad (8.41)$$

Введемо допоміжну випадкову величину

$$T = \frac{\overline{X} - a}{\frac{s}{\sqrt{n}}}, \qquad (8.42)$$

де s – виправлене середнє квадратичне відхилення випадкової величини X, що обчислюється за формулою (8.41). Можна довести, що випадкова величина T має розподіл Стьюдента з $n-1$ ступенями вільності (див. п.4.11).

Переходимо у (8.35) від випадкової величини X до випадкової величини T:

$$P\left(\frac{\left|\overline{X} - a\right|}{\frac{s}{\sqrt{n}}} < \frac{\delta}{\frac{s}{\sqrt{n}}}\right) = \gamma, \qquad (8.43)$$

або

$$P\left(|T| < t_\gamma\right) = \gamma, \qquad (8.44)$$

де

$$t_\gamma = \frac{\delta\sqrt{n}}{s}. \qquad (8.45)$$

Параметр t_γ знаходимо за дод. 4 значень $t_\gamma = t(\gamma; n)$. За допомогою t_γ з (8.45) знаходимо точність оцінки

$$\delta = t_\gamma \frac{s}{\sqrt{n}}. \qquad (8.46)$$

Тоді формула (8.43) набуває вигляду

$$P\left(\bar{x}_в - t_\gamma \frac{s}{\sqrt{n}} < a < \bar{x}_в + t_\gamma \frac{s}{\sqrt{n}}\right) = \gamma. \qquad (8.47)$$

Зауваження. При великих значеннях $n\,(n \ge 30)$ розподіл Стьюдента наближається до стандартного нормального розподілу, тому значення t_γ можна знаходити в цьому випадку за допомогою таблиці значень функції Лапласа (див. дод. 2).

Приклади

1. Із генеральної сукупності проведено вибірку об'ємом $n = 20$, яка характеризується ознакою X, що підлягає нормальному закону розподілу з параметром $\sigma = 1{,}5$. Знайти довірчий інтервал для оцінки невідомого математичного сподівання a з надійністю $\gamma = 0{,}95$, якщо $\overline{x}_в = 12$.

Розв'язання. При заданій надійності за (8.40) маємо $2\Phi(t) = 0{,}95$, звідки $\Phi(t) = 0{,}475$. За дод. 2 отримаємо $t = 1{,}96$. Тоді $\delta = \dfrac{t\sigma}{\sqrt{n}} = \dfrac{1{,}96 \cdot 1{,}5}{\sqrt{20}} \approx 0{,}6577$. Отримаємо довірчий інтервал:

$$12 - 0{,}6577 < a < 12 + 0{,}6577; \quad 11{,}3423 < a < 12{,}6577.$$

Отже, шуканий довірчий інтервал має вигляд: $(11{,}3423; 12{,}6577)$.

2. Вибірку задано таблицею

x_i	1	3	5	7	9
n_i	1	3	4	15	2

Необхідно побудувати довірчий інтервал з надійністю $\gamma = 0{,}95$ для математичного сподівання a ознаки X, припустивши, що генеральна сукупність розподілена нормально.

Розв'язання. За формулами (7.11), (7.16) обчислюємо числові характеристики вибірки: $\overline{x}_в = 6{,}12$; $D_в = 3{,}5456$. Згідно з (8.9), (8.10) виправлена дисперсія та виправлене середнє квадратичне відхилення дорівнюють відповідно $s^2 = \dfrac{n}{n-1} \cdot D_в \approx 3{,}6933$ і $s \approx 1{,}9218$. За даною надійністю $\gamma = 0{,}95$ з дод. 4 знаходимо $t_\gamma = t(\gamma; n) = t(0{,}95; 25) = 2{,}064$. Використавши знайдені значення числових параметрів, обчислимо межі довірчого інтервалу за формулою (8.47):

$$\bar{x}_в - t_\gamma \frac{s}{\sqrt{n}} = 6{,}12 - 2{,}064 \cdot \frac{1{,}9218}{\sqrt{25}} \approx 5{,}3267;$$
$$\bar{x}_в + t_\gamma \frac{s}{\sqrt{n}} = 6{,}12 + 2{,}064 \cdot \frac{1{,}9218}{\sqrt{25}} \approx 6{,}9133.$$

Таким чином, довірчий інтервал має вигляд: $5{,}3267 < a < 6{,}9133$.

Завдання

1. Із генеральної сукупності проведено вибірку об'ємом $n = 25$, яка характеризується ознакою X, що підлягає нормальному закону розподілу з параметром $\sigma = 2$. Знайти довірчий інтервал для оцінки невідомого математичного сподівання a з надійністю $\gamma = 0{,}95$, якщо $\bar{x}_в = 10{,}3$.

Відповідь. $(9{,}516; 11{,}084)$.

2. Знайти мінімальний об'єм вибірки, при якому з надійністю $\gamma = 0{,}975$ точність оцінки математичного сподівання a генеральної сукупності за вибірковою середньою $\bar{x}_в$ буде $\delta = 0{,}3$, якщо $\sigma = 1{,}3$.

Відповідь. $n = 95$.

3. За вибіркою об'єму $n = 25$ з нормальної генеральної сукупності при $\bar{x}_в = 40{,}8$ і $s = 0{,}9$ оцінити невідоме математичне сподівання за допомогою надійного інтервалу з надійністю $0{,}95$.

Відповідь. $(40{,}428; 41{,}172)$.

4. За даними дев'яти незалежних рівноточних вимірів фізичної величини знайдено $\bar{x}_в = 24{,}217$ і $s = 4$. Оцінити істинне значення цієї величини з надійністю 0,95.

Відповідь. $21{,}137 < a < 27{,}297$.

5. Задано вибірку:

x_i	-2	1	2	3	4	5
n_i	2	1	2	2	2	1

Побудувати інтервал з надійністю $\gamma = 0{,}95$ для математичного сподівання a ознаки X, якщо генеральна сукупність розподілена нормально.

Відповідь. $0{,}3 < a < 3{,}7$.

6. Статистичний аналіз непродуктивних затримок при перевезеннях на деякій філії ВАТ *"В"* має вигляд

Інтервал, *год*	$0{,}4 \div 1{,}2$	$1{,}2 \div 2{,}0$	$2{,}0 \div 2{,}8$	$2{,}8 \div 3{,}6$	$3{,}6 \div 4{,}4$	$4{,}4 \div 5{,}2$	$5{,}2 \div 6{,}0$
Число випадків n_i	1	3	6	7	19	8	6

Вважаючи, що час непродуктивних затримок при перевезеннях розподілено за нормальним законом, знайти: а) моду M_0 і медіану Me; б) коефіцієнт асиметрії a_s і ексцес e_k; в) імовірність того, що час затримок знаходиться в межах від 2,5 до 4,5 год; г) межі, в яких з імовірністю 0,95 буде знаходитись середній час затримок.

Відповідь: а) M_0=4,0; Me=3,2; б) $a_s \approx -0{,}42$; $e_k \approx -0{,}37$; в) 0,2; г) (3,5;4,1).

7. Випадкова величина X (тривалість непродуктивних затримок при перевезеннях на філії *"К"*) має нормальний розподіл з відомим середнім квадратичним відхиленням $\sigma = 0{,}267$. Знайти довірчий інтервал для оцінки невідомого математичного сподівання a з надійністю $\gamma = 0{,}95$, якщо $\overline{x}_в = 4{,}3$, а об'єм вибірки $n = 24$.

Відповідь. (4,1932; 4,4068).

8.8. Довірчий інтервал для оцінки невідомого середнього квадратичного відхилення σ нормального розподілу

Довірчий інтервал для невідомого середнього квадратичного відхилення σ нормально розподіленої ознаки X генеральної сукупності будемо знаходити у вигляді

$$s - \delta < \sigma < s + \delta, \tag{8.48}$$

де s – виправлене вибіркове середнє квадратичне відхилення.

Нерівність (8.48) рівносильна нерівності:

$$s\left(1-\frac{\delta}{s}\right)<\sigma<s\left(1+\frac{\delta}{s}\right). \qquad (8.49)$$

Позначивши $\frac{\delta}{s}=q$, отримаємо

$$\underline{s(1-q)<\sigma<s(1+q).} \qquad (8.50)$$

Значення $q=q(\gamma,n)$ знаходять за дод. 5 в залежності від заданої надійності та об'єму вибірки.

Таким чином, для оцінки параметра σ нормально розподіленої ознаки X генеральної сукупності при відомому виправленому середньому квадратичному відхиленні s, об'ємі вибірки n та надійності γ використовують одну з двох нерівностей:

$$\begin{aligned} &1)\quad s(1-q)<\sigma<s(1+q) \quad \text{при} \quad q<1;\\ &2)\quad 0<\sigma<s(1+q) \quad \text{при} \quad q>1. \end{aligned} \qquad (8.51)$$

Приклади

1. Ознака X генеральної сукупності розподілена нормально. За вибіркою об'єму $n=25$ обчислене виправлене вибіркове середнє квадратичне відхилення $s=0{,}85$. Знайти довірчий інтервал, який покриває генеральне середнє квадратичне відхилення σ з надійністю 0,95.

Розв'язання. За дод. 5 за даними $\gamma=0{,}95$ і $n=25$ знаходимо $q=0{,}32$. За формулою (8.51) знаходимо $(q<1)$:

$$0{,}85(1-0{,}32)<\sigma<0{,}85(1+0{,}32);$$

$$0{,}578<\sigma<1{,}122.$$

Таким чином, довірчий інтервал має вигляд: $(0{,}578;1{,}122)$.

2. Кількісна ознака X генеральної сукупності розподілена нормально. За вибіркою об'єму $n=10$ знайдено виправлене вибіркове середнє квадратичне відхилення $s=0{,}2$. Знайти

довірчий інтервал, який покриває середнє квадратичне відхилення σ з надійністю 0,999.

Розв'язання. За дод. 5 і даними $\gamma = 0{,}999$, $n = 10$ знаходимо $q = 1{,}8$. Оскільки $q > 1$, згідно з (8.51) довірчий інтервал визначається нерівністю

$$0 < \sigma < 0{,}2(1 + 1{,}8);$$

$$0 < \sigma < 0{,}56.$$

Отже, шуканий довірчий інтервал $(0; 0{,}56)$.

Завдання

1. Ознака X (тривалість непродуктивних затримок при перевезеннях на філії *"L"*) розподілена нормально. За вибіркою об'єму $n = 24$ обчислене виправлене вибіркове середнє квадратичне відхилення $s = 0{,}27$. Знайти довірчий інтервал, який покриває генеральне середнє квадратичне відхилення з надійністю $\gamma = 0{,}95$.

Відповідь. (0,1836; 0,3564).

2. Випадкова величина X – тривалість роботи маневрових локомотивів на позиції 00Х контролера машиніста розподілена за нормальним законом. За вибіркою об'єму $n = 64$ та виправленим вибірковим середнім квадратичним відхиленням $s = 6{,}16$ знайти довірчий інтервал, який покриває генеральне середнє квадратичне відхилення з надійністю $\gamma = 0{,}95$.

Відповідь. (5,0; 7,3).

Питання до теми

1. Що називається статистичною оцінкою? Які вимоги ставляться до оцінки?
2. Які оцінки називаються точковими, які – інтервальними?
3. Які точкові оцінки ви знаєте?
4. Що служить оцінкою математичного сподівання? Чи задовольняє вона умову незміщеності?
5. Які точкові оцінки дисперсії ви знаєте? Чим вони відрізняються?

6. Які методи знаходження точкових оцінок невідомих параметрів розподілів вам відомі?

7. У чому полягає і для чого використовується метод добутків?

8. Для чого застосовується вирівнювання варіаційних рядів?

9. Що таке довірчий інтервал?

10. Що називається надійністю оцінки? Які значення зазвичай вона набуває?

11. Що називається рівнем значущості оцінки? Як він пов'язаний з точністю оцінки?

12. Як побудувати довірчий інтервал для оцінки математичного сподівання нормально розподіленої ознаки генеральної сукупності? Від чого залежить його вигляд?

13. Як визначається довірчий інтервал для середнього квадратичного відхилення?

Тестові питання

1. Статистична оцінка називається незміщеною, якщо:

А	Б	В	Г	Д
$M(\theta^*) < \theta$	$M(\theta^*) > \theta$	$M(\theta^*) = \theta$	$\theta^* \xrightarrow[n\to+\infty]{P} \theta$	Інша відповідь

2. Яка з наведених оцінок є оцінкою математичного сподівання?

А	Б	В	Г	Д
$\overline{x}_в$	$D_в$	s	s^2	Інша відповідь

3. Яка з наведених оцінок є зміщеною?

А	Б	В	Г	Д
$\overline{x}_в$	$D_в$	s	s^2	Інша відповідь

4. Статистичний розподіл задано таблицею

x_i	-1	1	2
n_i	15	15	10

Чому дорівнює оцінка математичного сподівання?

А	Б	В	Г	Д
-1	1	2	1/2	Інша відповідь

5. Як позначається надійність оцінки?

А	Б	В	Г	Д
α	β	γ	δ	Інша відповідь

6. Як позначається рівень значущості?

А	Б	В	Г	Д
α	β	γ	δ	Інша відповідь

7. Як позначається точність оцінки?

А	Б	В	Г	Д
α	β	γ	δ	Інша відповідь

8. Який з даних інтервалів є довірчим інтервалом для оцінки параметра σ нормально розподіленої ознаки X, якщо відомі $s = 0{,}5$, $q = 1{,}8$?

А	Б	В	Г	Д
$(0;0{,}9)$	$(0;1{,}4)$	$(-0{,}4;1{,}4)$	$(0{,}5;1{,}4)$	Інша відповідь

Відповіді: 1.В; 2.А; 3.Б; 4.Г; 5.В; 6. А; 7.Г; 8.Б.

Розділ 9. Перевірка статистичних гіпотез

9.1 Основні поняття

Нехай над випадковою величиною X проведено n випробувань: $x_1, x_2, ..., x_n$ – її можливі значення. Часто на практиці необхідно знати закон розподілу випадкової величини X. Якщо закон розподілу невідомий, але є підстави вважати, що він має вид $F(x)$(інтегральна функція) або $f(x)$ (щільність розподілу), то *висувають гіпотезу*: випадкова величина X розподілена за законом $F(x)$ або $f(x)$. Можливий випадок, коли закон розподілу відомий, але деякий параметр невідомий, тоді висувають гіпотезу про певне значення цього параметра. Можливі також інші гіпотези: про рівність параметрів двох або кількох розподілів, про однорідність вибірок та ін.

Статистичною гіпотезою називають будь-яке твердження про вид або властивості розподілу випадкової величини, що спостерігається. *Нульовою (основною)* називають висунуту гіпотезу H_0. *Конкуруючою (альтернативною)* називають гіпотезу H_1, яка суперечить нульовій. Наприклад, якщо H_0=“математичне сподівання спостережуваного розподілу дорівнює 5”, тоді H_1=“математичне сподівання спостережуваного розподілу не дорівнює 5”.

Вибір нульової гіпотези повністю покладається на дослідника і залежить від постановки задачі. Математична статистика пропонує лише методи перевірки статистичних гіпотез. Зміст методів перевірки такий: якщо приймається гіпотеза H_0, то відхиляється H_1, і навпаки. Гіпотезу перевіряють на підставі вибірки, одержаної з генеральної сукупності X. Через випадковість вибірки можуть виникнути помилки двох типів.

Помилку, яка полягає в тому, що гіпотеза H_0 відхиляється у випадку, коли вона справедлива, називають *помилкою першого роду*. Помилку, яка полягає в тому, що гіпотеза H_0 приймається у випадку, коли вона несправедлива, називають *помилкою другого роду*. Імовірність допущення помилки першого роду називають *рівнем значущості* і позначають через α. Якщо, наприклад, візьмемо $\alpha = 0{,}01$, то це означає, що в одному випадку зі ста є ризик зробити помилку першого роду (відхилити правильну гіпотезу).

Для перевірки нульової гіпотези використовують спеціальні статистичні критерії (критерій χ^2 Пірсона, критерій λ Колмогорова, t–критерій Стьюдента та ін.).

Статистичний критерій – це величина K, обчислена на основі спостережень, в залежності від значення якої нульову гіпотезу приймають або відхиляють.

Після вибору критерію K множину його можливих значень розбивають на дві підмножини: S – підмножина значень, при яких гіпотеза H_0 відхиляється, і $\overline{S}$ – підмножина значень, при яких гіпотеза H_0 приймається. Підмножина S називається *критичною областю*, а $\overline{S}$ – *областю прийняття гіпотези*.

Нехай для випадкової величини X можливими значеннями є $x_1, x_2, ..., x_n$. За даною вибіркою обчислюємо значення критерію K. Якщо це значення належить до підмножини S, то гіпотезу відхиляємо, якщо до $\overline{S}$, то гіпотезу приймаємо. Оскільки критерій K – одновимірна випадкова величина, то її можливі значення належать деякому інтервалу. Тому критична область і область прийняття також є інтервальними. Як наслідок, існують точки, які їх відокремлюють. Вони називаються *критичними точками* (або критичними межами). Критичні точки визначаються за таблицями розподілу критерію K і позначаються $K_{кр}$.

Розрізняють односторонні (право- та лівосторонні) та двосторонні критичні області. Критична область називається *правосторонньою*, якщо $K > K_{кр}$; *лівосторонньою* – якщо $K < K_{кр}$; *двосторонньою* – якщо $K < K_{кр1}$ або $K > K_{кр2}$.

Розглянемо випадок *правосторонньої* критичної області. Очевидно, що для її визначення треба знайти критичну точку. Для цього задаємо рівень значущості α і знаходимо критичну точку $K_{кр}$ таку, що $P(K > K_{кр}) = \alpha$. Для кожного критерію є відповідні таблиці, за якими знаходять $K_{кр}$. Коли критичну точку знайдено, за даними вибірки обчислюємо спостережуване значення критерію. Якщо $K_{спост} > K_{кр}$, то нульову гіпотезу відхиляють, якщо $K_{спост} < K_{кр}$, то підстав для відхилення нульової гіпотези немає.

Нульову гіпотезу H_0 обирають таку, для якої важливіше уникнути помилки першого роду. Імовірність помилки другого роду позначають через β. Тоді ймовірність протилежної події (відхилення нульової гіпотези у випадку, коли справедлива конкуруюча) позначають через $1-\beta$. Саме таку ймовірність має відсутність помилки другого роду у випадку правильності конкуруючої гіпотези. Зазначену ймовірність називають *потужністю критерію*. Тобто потужність критерію – це ймовірність того, що нульова гіпотеза відхиляється, якщо справедлива конкуруюча.

Критичну область треба будувати так, щоб потужність критерію була максимальною. Виконання цієї умови забезпечить мінімальну помилку другого роду. Слід пам'ятати, що помилки першого і другого роду є конкуруючими, тобто зменшення ймовірності допустити одну з них зумовлює збільшення ймовірності допустити іншу. Тому в кожному випадку слід обирати компромісне рішення.

Єдиним правильним шляхом зменшення можливих помилок є збільшення об'єму вибірок.

9.2. Критерії згоди. Критерій згоди χ^2 Пірсона

Розглянемо способи перевірки гіпотези про розподіл генеральної сукупності X. Нехай вона має невідомий розподіл $f(x)$ або $F(x)$. Розглянемо вибірку $x_1, x_2, ..., x_n$ об'єму n, яка має розподіл $f_0(x)$ або $F_0(x)$. На підставі аналізу висуваємо гіпотезу H_0: досліджувана генеральна сукупність розподілена за законом, визначеним функцією $f(x)$. Розподіл $f(x)$ – *теоретичний*, $f_0(x)$ – *емпіричний*. Якщо відхилення емпіричного розподілу від теоретичного мале, то гіпотезу H_0 приймаємо; якщо відхилення велике – гіпотезу треба відхилити. Для перевірки подібних гіпотез розроблено декілька критеріїв згоди. Найчастіше використовується так званий критерій згоди «хі – квадрат» (χ^2) Пірсона, суть якого полягає у порівнянні емпіричних і теоретичних частот.

Нехай множина усіх значень ознаки X розділена на s інтервалів. Позначимо через n_i частоту варіант, що потрапили до

i-го інтервалу, а через n_i' – теоретичні частоти, які обчислюються за формулою

$$n_i' = p_i \cdot n, \tag{9.1}$$

де p_i – імовірність того, що ознака X набуде значення з i-го інтервалу. Імовірності p_i можуть бути обчислені за допомогою функції розподілу $F(x)$:

$$p_i = P(\alpha_i \le X \le \beta_i) = F(\beta_i) - F(\alpha_i), \tag{9.2}$$

де α_i, β_i – кінці i-го інтервалу.

Статистика

$$\chi^2 = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'}, \tag{9.3}$$

де n_i – емпіричні (обчислені за вибіркою), а n_i' – теоретичні частоти називається критерієм Пірсона.

Незалежно від виду розподілу $F(x)$ дана випадкова величина при зростанні n швидко наближається до розподілу χ^2 з $k = s - 1 - r$ степенями вільності, де r – число параметрів розподілу генеральної сукупності, що оцінюються за вибіркою.

Чим менша величина χ^2, тим менше буде відрізнятись емпіричний розподіл $f_0(x)$ або $F_0(x)$ від теоретичного $f(x)$ або $F(x)$. Виберемо рівень значущості α близьким до нуля і знайдемо критичне значення критерію χ^2 за формулою $P(\chi^2 > \chi^2_{кр}) = \alpha$. За заданим рівнем значущості α і числом степенів вільності $k = s - 1 - r$, за таблицею критичних точок розподілу χ^2 (дод. 6) знаходимо критичну точку критерію $\chi^2_{кр}$. На основі відомих емпіричних і теоретичних частот обчислюємо спостережуване значення критерію $\chi^2_{спост}$ (9.3). Якщо $\chi^2_{спост} < \chi^2_{кр}$, а критична область є правосторонньою, то гіпотеза H_0 про вид розподілу приймається; якщо $\chi^2_{спост} > \chi^2_{кр}$ – гіпотезу H_0 відхиляємо.

Правило використання критерію χ^2 Пірсона

1. За формулою (9.3) обчислити $\chi^2_{спост}$ – вибіркове значення статистики критерію.

2. Вибрати рівень значущості критерію α та за таблицею χ^2-розподілу (дод. 6) знайти критичну точку $\chi^2_{кр}$.

3. Якщо $\chi^2_{спост} < \chi^2_{кр}$, то гіпотеза H_0 про вид розподілу приймається; якщо $\chi^2_{спост} > \chi^2_{кр}$, то гіпотеза H_0 відхиляється.

Зауваження. Необхідною умовою застосування критерію χ^2 Пірсона є наявність у кожному з інтервалів не менш ніж 5 спостережень (тобто $n_i \geq 5, i = 1,2,...,s$). Якщо в окремих інтервалах частота спостережень менш ніж 5, то число інтервалів треба зменшити шляхом об'єднання сусідніх інтервалів.

Завдання

1. При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо емпіричні і теоретичні частоти задані таблицею:

Емпіричні частоти n_i	6	12	16	40	13	8	5
Теоретичні частоти n_i'	4	11	15	43	15	6	6

Відповідь. $\chi^2_{спост}$=2,5; $\chi^2_{кр}$=9,5. Немає підстав відкинути гіпотезу про нормальний розподіл генеральної сукупності.

2. Зроблено аналіз маси поїзда, що відправляється зі станції вузла. При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо:

Емпіричні частоти n_i	98	115	86	63	21	15	9
Теоретичні частоти n_i'	97	111	81	60	29	16	13

Відповідь. Гіпотезу про нормальний розподіл генеральної сукупності приймаємо, $\chi^2_{спост} \approx 4{,}11$.

3. Зроблено аналіз простою на вантажних станціях. При рівні значущості $\alpha = 0{,}05$ перевірити гіпотезу про показниковий розподіл генеральної сукупності:

Емпіричні частоти n_i	185	76	32	14	5	4
Теоретичні частоти n_i'	188	75	31	13	6	3

Відповідь. Гіпотеза про показниковий розподіл приймається, $\chi^2_{спост} \approx 0{,}7$.

9.3. Методика обчислення критерію χ^2 для нормального розподілу

Нехай задано інтервальний ряд:

$x_i \div x_{i+1}$	$x_1 \div x_2$	$x_2 \div x_3$	…	$x_s \div x_{s+1}$
n_i	n_1	n_2	…	n_s

і вказано рівень значущості α. Необхідно перевірити гіпотезу H_0 про нормальний розподіл генеральної сукупності.

Етапи перевірки критерію

1. Параметри нормального закону розподілу a і σ^2 (математичне сподівання і дисперсія випадкової величини X) невідомі, тому визначаємо їх за вибіркою незміщеними й обґрунтованими оцінками $\overline{x}_в$ та «виправленою» вибірковою дисперсією s^2 відповідно. Якщо об’єм вибірки достатньо великий $(n > 30 - 40)$, то замість «виправленої» s^2 можна брати «звичайну» вибіркову дисперсію $D_в$.

2. Для обчислення ймовірностей p_i попадання випадкової величини X в інтервал $x_i \div x_{i+1}$ використовуємо функцію Лапласа (дод. 2):

$$p_i = P\{x_i \le X \le x_{i+1}\} = \Phi\left(\frac{x_{i+1} - a}{\sigma}\right) - \Phi\left(\frac{x_i - a}{\sigma}\right). \qquad (9.4)$$

3. Обчислюємо теоретичні частоти $n_i', i = 1,2,...,s$ за формулою (9.1).

4. Обчислюємо $\chi^2_{спост}$ за формулою (9.3).

5. Знаходимо число степенів вільності k: за вибіркою обчислено два параметри (a і σ), як наслідок

$$k = s - 1 - r = s - 3. \quad (9.5)$$

За таблицею χ^2 (дод. 6) знаходимо $\chi^2_{кр}$.

6. Якщо $\chi^2_{спост} < \chi^2_{кр}$, то гіпотезу H_0 про нормальний розподіл генеральної сукупності приймаємо; якщо $\chi^2_{спост} > \chi^2_{кр}$, то гіпотезу H_0 відхиляємо.

Приклад. При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл випадкової величини X (швидкість виходу відчепів з першої гальмової позиції гірки середньої потужності), заданої варіаційним рядом розподілу:

Швидкість виходу, $^{м}/_{c}, x_i \div x_{i+1}$	$3{,}00 \div 3{,}15$	$3{,}15 \div 3{,}30$	$3{,}30 \div 3{,}45$	$3{,}45 \div 3{,}60$	$3{,}60 \div 3{,}75$	$3{,}75 \div 3{,}90$	$3{,}90 \div 4{,}05$	$4{,}05 \div 4{,}20$	$4{,}20 \div 4{,}35$	$4{,}35 \div 4{,}50$
Число випадків n_i	2	8	18	32	34	52	20	18	10	6

<u>Розв'язання.</u> Параметри нормального закону розподілу a і σ^2 (математичне сподівання та дисперсія) випадкової величини X невідомі. Тому замінюємо їх оцінками за вибіркою – незміщеними та обґрунтованими оцінками відповідно вибірковим середнім $\bar{x}_в$ та «виправленою» вибірковою дисперсією s^2. Об'єм вибірки великий ($n = 200$), тому замість «виправленої» s^2 можна брати «звичайну» вибіркову дисперсію $D_в$. За формулами (7.11) та (7.16) отримаємо:

$$\bar{x}_в \approx 3{,}76 \; ; \; D_в \approx 0{,}0988 \; \Rightarrow \sigma_в \approx 0{,}31.$$

Враховуючи, що в емпіричному розподілу частота першого інтервалу ($n_1 = 2$) менш ніж 5, то доцільно об'єднати перший та другий інтервали. Отримаємо такий ряд розподілу:

Швидкість виходу, $м/с$, $x_i \div x_{i+1}$	$3{,}00 \div 3{,}30$	$3{,}30 \div 3{,}45$	$3{,}45 \div 3{,}60$	$3{,}60 \div 3{,}75$	$3{,}75 \div 3{,}90$	$3{,}90 \div 4{,}05$	$4{,}05 \div 4{,}20$	$4{,}20 \div 4{,}35$	$4{,}35 \div 4{,}50$
Число випадків n_i	10	18	32	34	52	20	18	10	6

Для обчислення ймовірностей p_i попадання випадкової величини X в інтервал $x_i \div x_{i+1}$ використовуємо функцію Лапласа за формулою (9.4):

$$p_i = P\left(x_i \le X \le x_{i+1}\right) = \Phi\left(\frac{x_{i+1} - a}{\sigma}\right) - \Phi\left(\frac{x_i - a}{\sigma}\right) =$$

$$= \Phi\left(\frac{x_{i+1} - 3{,}76}{0{,}31}\right) - \Phi\left(\frac{x_i - 3{,}76}{0{,}31}\right).$$

Так як випадкова величина X, що розподілена за нормальним законом з параметрами a і σ^2, визначена на $(-\infty;+\infty)$, то крайні інтервали в ряду розподілу заміняємо відповідно на $(-\infty;3{,}30)$ та $(4{,}35;+\infty)$. Тоді

$$p_1 = P\left(-\infty \le X \le 3{,}30\right) = \Phi\left(\frac{3{,}30 - 3{,}76}{0{,}31}\right) - \Phi(-\infty) = -0{,}4306 + 0{,}5 = 0{,}0694;$$

$$p_2 = P\left(3{,}30 \le X \le 3{,}45\right) = \Phi\left(\frac{3{,}45 - 3{,}76}{0{,}31}\right) - \Phi\left(\frac{3{,}30 - 3{,}76}{0{,}31}\right) =$$

$$= -0{,}3413 + 0{,}4306 = 0{,}0893.$$

Аналогічно знаходимо p_i $(i = 3,4,...,9)$. Для обчислення статистики χ^2 складемо табл. 9.1.

Таблиця 9.1

i	Інтервал $x_i \div x_{i+1}$	Емпіричні частоти n_i	Імовірності p_i	Теоретичні частоти $n'_i = n \cdot p_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	$-\infty \div 3{,}30$	10	0,0694	13,88	15,0544	1,0846
2	$3{,}30 \div 3{,}45$	18	0,0893	17,86	0,0196	0,0011
3	$3{,}45 \div 3{,}60$	32	0,1428	28,56	11,8336	0,4143
4	$3{,}60 \div 3{,}75$	34	0,1865	37,3	10,89	0,2920
5	$3{,}75 \div 3{,}90$	52	0,1856	37,12	221,4144	5,9648
6	$3{,}90 \div 4{,}05$	20	0,1528	30,56	111,5136	3,6490
7	$4{,}05 \div 4{,}20$	18	0,0958	19,16	1,3456	0,0702
8	$4{,}20 \div 4{,}35$	10	0,0491	9,82	0,0324	0,0033
9	$4{,}35 \div +\infty$	6	0,0287	5,74	0,0676	0,0118
$\sum$		200	1	200		11,4911

Отримаємо $\chi^2_{спост}$=11,4911. Так як кількість інтервалів $s = 9$, а нормальний закон розподілу визначається двома параметрами $r = 2$, то кількість степенів вільності $k = s - r - 1 = 9 - 2 - 1 = 6$. За даним рівнем значущості $\alpha = 0{,}05$ та кількістю степенів вільності $k = 6$ з дод. 6 отримаємо $\chi^2_{кр}$=12,6.

Так як $\chi^2_{спост} < \chi^2_{кр}$, то гіпотеза про нормальний закон розподілу випадкової величини X приймається.

Завдання

1. При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл ознаки X генеральної сукупності, заданої варіаційним рядом розподілу:

x_i	5	7	9	11	13	15	17	19	21
n_i	15	26	25	30	26	21	24	20	13

<u>Відповідь.</u> $\chi^2_{спост}$=22,2; $\chi^2_{кр}$=12,6. Гіпотеза про нормальний розподіл генеральної сукупності відхиляється.

2. При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл ознаки X генеральної сукупності, заданої варіаційним рядом розподілу:

x_i	8	16	40	72	36	18	10
n_i	6	18	36	76	39	18	7

<u>Відповідь.</u> $\chi^2_{спост}$=3,061; $\chi^2_{кр}$=13,3. Немає підстав відхиляти гіпотезу про нормальний розподіл генеральної сукупності.

3. Результати вимірювання швидкості x_i руху автомобілів на певній ділянці шляху наведено у вигляді статистичного розподілу:

x_i, *км/год*	60	65	70	75	80	85	90	95
n_i	2	5	6	7	9	8	6	7

Вважаючи, що швидкість руху автомобіля є випадковою величиною, перевірити за критерієм згоди Пірсона за рівнем значущості $\alpha = 0{,}01$ гіпотезу про нормальний закон розподілу.

<u>Відповідь.</u> $\chi^2_{спост}$=1,8284; $\chi^2_{кр}$=15,1. Немає підстав відхилити гіпотезу про нормальний розподіл генеральної сукупності.

4. Впродовж деякого часу було зафіксовано розпуск 100 составів, композиція і тривалість розпуску яких наведені у табл. 9.2:

Таблиця 9.2

№ составу	Число вагонів у составі	Число відчепів у составі	Тривалість розпуску, *хв.*
1	2	3	4
1	64	22	20,5
2	48	20	11
3	53	20	17
4	52	20	15
5	60	33	13,5
6	48	18	10,5
7	59	35	19
8	8	4	2,0

Продовження табл. 9.2

1	2	3	4
9	62	40	31
10	41	6	7,5
11	48	28	14
12	59	30	13,5
13	17	8	5
14	52	20	15
15	20	14	8
16	51	22	15
17	35	10	8
18	59	22	16
19	44	13	11
20	57	25	13,5
21	40	8	7,5

Продовження табл. 9.2

1	2	3	4
22	27	10	5,5
23	58	24	16
24	57	28	16,5
25	3	3	1,5
26	54	17	12,5
27	45	19	13
28	43	23	14,5
29	52	24	17
30	78	50	36
31	57	26	16
32	46	20	11
33	46	19	10
34	47	20	14
35	54	31	18,5
36	25	19	9
37	58	22	14
38	53	28	16,5
39	55	24	14,5
40	46	20	15
41	52	24	17
42	55	24	16,5
43	47	18	14,5
44	25	20	9
45	8	5	2,5
46	21	6	5
47	22	7	6,5
48	55	9	8
49	48	28	19
50	46	23	16,5
51	57	15	18
52	9	5	2,5
53	50	22	15
54	52	23	16
55	48	19	13,5
56	40	15	10

Продовження табл. 9.2

1	2	3	4
57	54	18	14
58	36	12	8
59	48	27	14
60	3	2	1
61	48	19	15
62	54	17	18
63	35	10	12
64	43	13	15
65	43	11	13,5
66	51	28	19
67	47	11	12
68	58	26	20
69	8	3	2
70	57	16	18
71	54	17	18
72	28	10	7,5
73	12	2	2,5
74	57	10	10,5
75	20	4	4
76	58	17	16,5
77	39	7	10
78	52	23	18,5
79	49	13	12
80	48	14	12,5
81	56	26	18
82	60	9	10
83	56	30	18,5
84	34	15	11,5
85	54	11	13
86	6	4	2
87	63	16	17
88	55	17	16
89	61	28	17
90	57	22	13,5
91	48	31	16

Продовження табл. 9.2

1	2	3	4
92	58	21	13
93	41	22	12,5
94	57	18	11
95	28	19	9
96	68	11	23

Продовження табл. 9.2

1	2	3	4
97	52	10	9
98	44	11	10
99	53	28	18
100	61	23	14

При рівні значущості $\alpha = 0{,}05$ за критерієм згоди Пірсона перевірити гіпотезу про нормальний розподіл випадкової величини: а) X (кількість вагонів у составі); б) Y (кількість відчепів у составі); в) Z (тривалість розпуску состава).

Відповідь: а) $\chi^2_{спост} \approx 38$, гіпотезу про нормальний розподіл генеральної сукупності X відхиляємо; б) $\chi^2_{спост} \approx 4{,}5$, немає підстав відхиляти гіпотезу про нормальний розподіл генеральної сукупності Y; в) $\chi^2_{спост} \approx 7{,}8$, гіпотеза про нормальний розподіл генеральної сукупності Z відхиляється.

5. Зроблено аналіз маси поїзда, що відправляється зі станції вузла. При рівні значущості $\alpha = 0{,}05$, використовуючи критерій згоди Пірсона, перевірити гіпотезу про нормальний розподіл генеральної сукупності.

Емпіричні частоти n_i	98	115	86	63	21	15	9
Теоретичні частоти n_i'	97	111	81	60	29	16	13

Відповідь. Гіпотезу про нормальний розподіл генеральної сукупності приймаємо, $\chi^2_{спост} \approx 4{,}1132$.

9.4. Методика обчислення χ^2 для показникового розподілу

Нехай задано інтервальний варіаційний ряд:

$x_i \div x_{i+1}$	$0 \div x_1$	$x_1 \div x_2$	$\ldots$	$x_{s-1} \div x_s$
n_i	n_1	n_2	$\ldots$	n_s

За критерієм згоди χ^2 необхідно перевірити гіпотезу про показниковий закон розподілу даного статистичного матеріалу.

Нагадаємо (див. п. 4.6), що щільність показникового розподілу має вигляд

$$f(x) = \begin{cases} 0 & \text{при } x < 0, \\ \lambda e^{-\lambda x} & \text{при } x \geq 0, \end{cases}$$

а оцінка параметра λ визначається рівністю $\lambda = \frac{1}{\overline{x}_в}$, де $\overline{x}_в = \frac{1}{n}\sum_{i=1}^{s} x_i^* \cdot n_i$ (x_i^* – середина i-го інтервалу).

Теоретична ймовірність p_i попадання випадкової величини X, розподіленої за показниковим законом, в i-й інтервал визначається за формулою

$$p_i = P(x_{i-1} < X < x_i) = \\ = F(x_i) - F(x_{I-1}) = 1 - e^{-\lambda x_i} - \left(1 - e^{-\lambda x_{i-1}}\right) = e^{-\lambda x_{i-1}} - e^{-\lambda x_i}, \quad i = 2,3,\ldots,s-1. \tag{9.6}$$

При цьому

$$p_1 = P(0 < X < x_1) = F(x_1) - F(0) = 1 - e^{-\lambda x_1} - (1-1) = 1 - e^{-\lambda x_1}; \tag{9.7}$$

$$p_s = P(x_{s-1} < X < x_s) = F(+\infty) - F(x_{s-1}) = 1 - \left(1 - e^{-\lambda x_{s-1}}\right) = e^{-\lambda x_{s-1}}. \tag{9.8}$$

Приклад. У результаті випробувань 300 елементів на тривалість роботи отримано інтервальний варіаційний ряд:

$x_i \div x_{i+1}$	$0 \div 5$	$5 \div 10$	$10 \div 15$	$15 \div 20$	$20 \div 25$	$25 \div 30$	$30 \div 35$
n_i	143	55	40	25	14	12	11

де $x_i \div x_{i+1}$ – інтервал часу, n_i – кількість відмов в i-му інтервалі.

Користуючись критерієм Пірсона, за рівнем значущості $\alpha = 0{,}01$, перевірити гіпотезу про показниковий закон розподілу ознаки X – часу безвідмовної роботи елементів.

Розв'язання. Середній час роботи елементів дорівнює:

$$\bar{x}_в = \frac{1}{300}\left(2{,}5\cdot 143 + 7{,}5\cdot 55 + 12{,}5\cdot 40 + 17{,}5\cdot 25 + 22{,}5\cdot 14 + 27{,}5\cdot 12 + 32{,}5\cdot 11\right) \approx$$
$$\approx 9{,}03,$$

звідки отримаємо оцінку параметра:

$$\lambda = \frac{1}{\bar{x}_в} = \frac{1}{9{,}03} \approx 0{,}11.$$

Обчислюємо за допомогою формул (9.6)-(9.8) значення ймовірностей:

$$p_1 = 1 - e^{-0{,}11\cdot 5} \approx 0{,}4231; \quad p_2 = e^{-0{,}11\cdot 5} - e^{-0{,}11\cdot 10} \approx 0{,}244;$$
$$p_3 = e^{-0{,}11\cdot 10} - e^{-0{,}11\cdot 15} \approx 0{,}1409; \quad p_4 = e^{-0{,}11\cdot 15} - e^{-0{,}11\cdot 20} \approx 0{,}0812;$$
$$p_5 = e^{-0{,}11\cdot 20} - e^{-0{,}11\cdot 25} \approx 0{,}0469; \quad p_6 = e^{-0{,}11\cdot 25} - e^{-0{,}11\cdot 30} \approx 0{,}027;$$
$$p_7 = e^{-0{,}11\cdot 30} \approx 0{,}0369.$$

Для обчислення статистики χ^2 складемо табл. 9.3.

Таблиця 9.3

i	Інтер-вал $x_i \div x_{i+1}$	Емпі-ричні частоти n_i	Імовір-ності p_i	Теоре-тичні частоти $n'_i = n \cdot p_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	$0 \div 5$	143	0,4231	126,93	258,2449	2,0345
2	$5 \div 10$	55	0,244	73,2	331,24	4,5251
3	$10 \div 15$	40	0,1409	42,27	5,1529	0,1219
4	$15 \div 20$	25	0,0812	24,36	0,4096	0,0168
5	$20 \div 25$	14	0,0469	14,07	0,0049	0,0003
6	$25 \div 30$	12	0,027	8,1	15,21	1,8778
7	$30 \div 35$	11	0,0369	11,07	0,0049	0,0004
$\sum$		300	1	300		8,5768

Отримаємо $\chi^2_{спост}$=8,5768. За рівнем значущості $\alpha = 0{,}01$ та кількістю степенів вільності $k = 7 - 2 = 5$ з дод. 6 отримаємо $\chi^2_{кр}$=15,1. Оскільки $\chi^2_{спост} <$ $\chi^2_{кр}$, то гіпотеза про показниковий розподіл даної сукупності приймається.

Завдання

1. На одному зі складів залізниці проаналізували добовий відпуск матеріалу A за 50 навмання обраних днів. Внаслідок аналізу отримали нижченаведені згруповані дані:

Добовий відпуск матеріалу, од.	$0 \div 10$	$10 \div 20$	$20 \div 30$	$30 \div 40$	$40 \div 50$	$50 \div 60$	$60 \div 70$	$70 \div 80$
Кількість днів з даним відпуском	18	12	8	5	3	2	1	1

Застосовуючи критерій χ^2 Пірсона, перевірити, чи узгоджуються ці емпіричні дані при рівні значущості $\alpha = 0{,}05$ із гіпотезою про показниковий розподіл добового відпуску матеріалу зі складу.

<u>Відповідь.</u> $\chi^2_{спост} \approx 0{,}18$, немає підстав відхиляти гіпотезу про показниковий розподіл випадкової величини.

2. Маємо статистичний розподіл часу між поїздами, що прибувають для переробки:

Інтервал $t_i \div t_{i+1}$, *год*	$0 \div 0{,}2$	$0{,}2 \div 0{,}4$	$0{,}4 \div 0{,}6$	$0{,}6 \div 0{,}8$	$0{,}8 \div 1{,}0$
Частота n_i	66	49	33	19	14

Інтервал $t_i \div t_{i+1}$, *год*	$1{,}0 \div 1{,}2$	$1{,}2 \div 1{,}4$	$1{,}4 \div 1{,}6$
Частота n_i	8	6	5

Користуючись критерієм χ^2 Пірсона, за рівнем значущості $\alpha = 0{,}05$ перевірити гіпотезу про те, що розподіл випадкової величини є показниковим.

<u>Відповідь.</u> $\chi^2_{спост} \approx 3{,}3153$, немає підстави для відмови від гіпотези про показниковий розподіл випадкової величини.

3. Задано статистичний аналіз простою вагонів на вантажних станціях:

Інтервал $x_i \div x_{i+1}$, *год*	$0{,}5 \div 34$	$34 \div 68$	$68 \div 102$	$102 \div 135$	$135 \div 170$
Число випадків n_i	185	76	32	14	5

Інтервал $x_i \div x_{i+1}$, *год*	$170 \div 203$	$203 \div 237$	$237 \div 270$	$270 \div 304$
Число випадків n_i	2	1	0	1

Користуючись критерієм χ^2 Пірсона, за рівнем значущості $\alpha = 0{,}05$ перевірити гіпотезу про те, що випадкова величина X (тривалість простою вагонів на вантажних станціях) розподілена за показниковим законом. Якщо гіпотеза приймається, то: а) записати функції щільності та розподілу випадкової величини X та побудувати їх графіки; б) обчислити ймовірність того, що час простою складе від 35 до 60 годин.

Відповідь. Гіпотеза приймається, оскільки $\chi^2_{спост} \approx 6{,}1828$;

а) $f(x) = \begin{cases} 0{,}024 \cdot e^{-0{,}024x} & \text{при } x \ge 0, \\ 0 & \text{при } x < 0; \end{cases}$ $\quad F(x) = \begin{cases} 1 - e^{-0{,}024x} & \text{при } x \ge 0, \\ 0 & \text{при } x < 0; \end{cases}$

б) $P(35 < X < 60) = 0{,}19$.

4. За даним статистичним аналізом простою на вантажних станціях перевірити за допомогою критерію Пірсона гіпотезу про показниковий розподіл генеральної сукупності (рівень значущості $\alpha = 0{,}05$).

Емпіричні частоти n_i	185	76	32	14	5	4
Теоретичні частоти n_i'	188	75	31	13	6	3

Відповідь. $\chi^2_{спост} \approx 0{,}6704$, гіпотеза про показниковий розподіл приймається.

9.5. Методика обчислення критерію χ^2 для розподілу Пуассона

Нехай задано статистичний ряд розподілу:

X	0	1	2	…	s
n_i	n_0	n_1	n_2	…	n_s

Необхідно знайти теоретичний розподіл Пуассона і перевірити за критерієм χ^2 Пірсона при заданому рівні значущості α ступінь узгодженості теоретичного та емпіричного розподілів.

За формулою $\bar{x}_в = \frac{1}{n}\sum_{i=1}^{s} x_i \cdot n_i$ знаходимо вибіркове середнє. Теоретичні ймовірності подій обчислюємо за формулами:

$$p_i = P(X = x_i) = \frac{\left(\bar{x}_в\right)^i \cdot e^{-\bar{x}_в}}{i!}, i = 0,1,2,\ldots,s-1;\ p_s = 1 - \sum_{i=0}^{s-1} p_i. \qquad (9.9)$$

Обчислюємо значення $n \cdot p_i\ (i = 0,1,\ldots,s)$, знаходимо $\chi^2_{спост} = \sum_{i=0}^{s} \frac{\left(n_i - np_i\right)^2}{np_i}$.

За даним числом степенів вільності $k = s-2$ за дод. 6 знаходимо $\chi^2_{кр}$ і перевіряємо правильність гіпотези про розподіл випадкової величини X за законом Пуассона. Якщо $\chi^2_{спост} < \chi^2_{кр}$, то при заданому рівні значущості α закон про пуассонівський розподіл приймається.

Приклад. У разі перевірки на нестандартність 300 ящиків з виробами отримали емпіричний розподіл:

x_i	0	1	2	3	4	5	6	7
n_i	15	71	75	68	39	17	10	5

За рівнем значущості $\alpha = 0{,}01$ перевірити гіпотезу про те, що випадкова величина X – кількість нестандартних виробів – розподілена за законом Пуассона.

Розв’язання. Знаходимо $\overline{x}_в \approx 2{,}5$. За формулами (9.9) обчислюємо:

$$p_i = P(X = x_i) = \frac{(2{,}5)^i \cdot e^{-2{,}5}}{i!}, i = 0,1,2,\dots,6; \qquad p_7 = 1 - \sum_{i=0}^{6} p_i .$$

А саме:

$$p_0 = P(X = 0) = \frac{(2{,}5)^0 \cdot e^{-2{,}5}}{0!} \approx 0{,}0821; \; p_1 = P(X = 1) = \frac{(2{,}5)^1 \cdot e^{-2{,}5}}{1!} \approx 0{,}2053;$$

$$p_2 = P(X = 2) = \frac{(2{,}5)^2 \cdot e^{-2{,}5}}{2!} \approx 0{,}2566; \; p_3 = P(X = 3) = \frac{(2{,}5)^3 \cdot e^{-2{,}5}}{3!} \approx 0{,}2138;$$

$$p_4 = P(X = 4) = \frac{(2{,}5)^4 \cdot e^{-2{,}5}}{4!} \approx 0{,}1336; \; p_5 = P(X = 5) = \frac{(2{,}5)^5 \cdot e^{-2{,}5}}{5!} \approx 0{,}0668;$$

$$p_6 = P(X = 6) = \frac{(2{,}5)^6 \cdot e^{-2{,}5}}{6!} \approx 0{,}0278;$$

$$p_7 = 1 - (0{,}0821 + 0{,}2053 + 0{,}2566 + 0{,}2138 + 0{,}1336 + 0{,}0668 + 0{,}0278) = 0{,}014.$$

Для обчислення статистики χ^2 складемо табл. 9.4.

Таблиця 9.4

i	x_i	Емпіричні частоти n_i	Імовірності p_i	Теоретичні частоти $n_i' = n \cdot p_i$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$
1	0	15	0,0821	24,63	92,7369	3,7652
2	1	71	0,2053	61,59	88,5481	1,4377
3	2	75	0,2566	76,98	3,9204	0,0509
4	3	68	0,2138	64,14	14,8996	0,2323
5	4	39	0,1336	40,08	1,1664	0,0291
6	5	17	0,0668	20,04	9,2416	0,4612
7	6	10	0,0278	8,34	2,7556	0,3304
8	7	5	0,014	4,2	0,64	0,1524
$\sum$		300	1	300		6,4592

Отримаємо $\chi^2_{спост}$=6,4592. За рівнем значущості $\alpha = 0{,}01$ та числом степенів вільності $k = 8 - 2 = 6$ з дод. 6 отримаємо $\chi^2_{кр}$=16,8. Оскільки $\chi^2_{спост} < \chi^2_{кр}$, то гіпотеза про те, що випадкова величина X розподілена за законом Пуассона приймається.

Завдання

1. У результаті перевірки 400 контейнерів зі скляними виробами встановлено, що кількість зіпсованих виробів X має емпіричний розподіл:

x_i	0	1	2	3	4	5	6	7
n_i	162	130	62	31	10	3	1	1

де x_i – кількість зіпсованих виробів в одному контейнері; n_i – кількість контейнерів, які вміщують x_i зіпсованих виробів. За рівнем значущості $\alpha = 0{,}01$ перевірити гіпотезу, що випадкова величина X – число зіпсованих виробів – розподілена за законом Пуассона.

Відповідь. Гіпотеза про розподіл Пуассона відхиляється, $\chi^2_{спост} \approx 16{,}229$.

Питання до теми

1. Що називається статистичною гіпотезою?
2. Яка гіпотеза називається нульовою, яка – конкуруючою? Наведіть приклади.
3. Які типи помилок можуть виникнути при перевірці статистичної гіпотези?
4. Як називається і позначається ймовірність помилки першого роду?
5. Як називається і позначається ймовірність помилки другого роду?
6. Що таке статистичний критерій?
7. Що таке потужність критерію?
8. Що таке критична область? З яких міркувань вона будується?
9. Яка величина називається критерієм Пірсона? Який його зміст?
10. Для чого використовується критерій Пірсона?
11. Як знайти критичні значення критерію $\chi^2_{кр}$? У яких випадках гіпотеза приймається, у яких – відхиляється?
12. Які етапи перевірки гіпотези за допомогою критерію Пірсона?

13. Які особливості має перевірка гіпотези за допомогою критерію Пірсона у випадках нормального, показникового та закону Пуассона? Наведіть відповідні формули.

Тестові питання

1. Як позначається помилка першого роду?

А	Б	В	Г	Д
α	β	γ	δ	Інша відповідь

2. Як позначається помилка другого роду?

А	Б	В	Г	Д
α	β	γ	δ	Інша відповідь

3. Чому дорівнює потужність критерію?

А	Б	В	Г	Д
α	β	$1-\alpha$	$1-\beta$	Інша відповідь

4. Гіпотезу про який закон розподілу можна перевірити за допомогою критерію Пірсона?

А	Б	В	Г	Д
Нормальний	Показниковий	Пуассона	Будь-який	Інша відповідь

5. Якою є критична область для критерію Пірсона?

А	Б	В	Г
Двосторонньою	Ліво-сторонньою	Право-сторонньою	Інша відповідь

<u>Відповіді:</u> 1.А; 2.Б; 3.Г; 4.Г; 5.В.

Додаток 1

Таблиця значень функції Гаусса $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$

	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	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3478	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	2813	2589	2565	2541	2516	2492	2468	2444

1,0	0,2420	2396	2371	2347	2323	2293	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	1646	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	1107
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0978	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	0,0540	0525	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	0279	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	0164	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0118	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

Продовження дод. 1

3,0	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	0014	0014	0013	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0003	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	0002	0000

Додаток 2

Таблиця значень функції Лапласа $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,50	0,1915	1,00	0,3413	1,50	0,4332	2,02	0,4783
0,01	0,0040	0,51	0,1950	1,01	0,3448	1,51	0,4345	2,04	0,4793
0,02	0,0080	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,06	0,4803
0,03	0,0120	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,08	0,4812
0,04	0,0160	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,10	0,4821
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,12	0,4830
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,14	0,4838
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,16	0,4846
0,08	0,0319	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,18	0,4854
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,20	0,4861
0,10	0,0398	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,22	0,4868
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,24	0,4875
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,26	0,4881
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,28	0,4887
0,14	0,0557	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,30	0,4893
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,32	0,4898
0,16	0,0636	0,66	0,2454	1,16	0,3770	1,66	0,4515	2,34	0,4904
0,17	0,0675	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,36	0,4909
0,18	0,0714	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,38	0,4913
0,19	0,0753	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,40	0,4918
0,20	0,0793	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,42	0,4922
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,44	0,4927
0,22	0,0871	0,72	0,2642	1,22	0,3883	1,72	0,4573	2,46	0,4931
0,23	0,0910	0,73	0,2673	1,23	0,3807	1,73	0,4582	2,48	0,4934
0,24	0,0948	0,74	0,2703	1,24	0,3925	1,74	0,4591	2,50	0,4938
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,52	0,4941
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,54	0,4945
0,27	0,1064	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,56	0,4948
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,58	0,4951
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,60	0,4953
0,30	0,1179	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,62	0,4956
0,31	0,1217	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,64	0,4959
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,66	0,4961
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,68	0,4963
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,70	0,4965

Продовження дод. 2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,72	0,4967
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,74	0,4969
0,37	0,1443	0,87	0,3076	1,37	0,4147	1,87	0,4693	2,76	0,4971
0,38	0,1480	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,78	0,4973
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,80	0,4974
0,40	0,1554	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,82	0,4976
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,84	0,4977
0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,86	0,4979
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,88	0,4980
0,44	0,1700	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,90	0,4981
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,92	0,4982
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,94	0,4985
0,47	0,1808	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,96	0,4985
0,48	0,1884	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,98	0,4986
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	3,00	0,49865
								3,20	0,49931
								3,40	0,49966
								3,60	0,499841
								3,80	0,499928
								4,00	0,499468
								4,50	0,499997
								5,00	0,499997

Додаток 3

Таблиця значень функції $P(k,\lambda) = \dfrac{\lambda^{k}}{k!} \cdot e^{-\lambda}$

k	λ								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	0,904837	0,818731	0,740818	0,670320	0,65310	0,548812	0,496585	0,449329	0,406570
1	090484	163746	222245	268128	303265	329287	347610	359463	365913
2	004524	016375	033337	353626	065816	098786	121663	143785	164661
3	000151	001092	003334	007150	012636	019757	028388	038343	049398
4	000004	000055	000250	000715	001580	002764	004968	007669	011115
5		000002	000015	000057	000158	000356	000696	001227	002001
6			000001	000004	000013	000036	000081	000164	000300
7					000001	000003	000008	000019	000039
8							000001	000002	000004

k	λ								
	1	2	3	4	5	6	7	8	9
0	0,367879	0,135335	0,049787	0,018316	0,006738	0,002479	0,000912	0,000335	0,000123
1	367879	270671	149361	073263	033690	014873	006383	002684	001111
2	183940	270671	224042	146525	084224	044618	022341	010735	004998
3	061313	180447	224042	195367	140374	089235	052129	028626	014994
4	015328	090224	168031	195367	175467	133853	091226	057252	033737
5	003066	036089	100819	156293	175467	160623	127717	091604	060727
6	000511	012030	050409	104196	146223	160623	149003	122138	091090
7	000073	003437	021604	059540	104445	137677	149003	139587	117126
8	000009	000859	008102	029770	065278	103258	130377	138587	131756
9	000001	000191	002701	013231	036266	068838	101405	124077	131756
10		000038	000810	005292	018138	041303	070983	099262	118580
11		000007	000221	001295	008242	022529	045171	072190	097020
12		000001	000055	000642	003434	011264	026350	048127	072765
13			000013	000197	001321	005199	014188	029616	050376
14			000003	000056	000472	002228	007094	016924	032384
15			000001	000015	000157	000891	003311	009026	019431
16				000004	000049	000334	001448	004513	010930
17				000001	000014	000118	000596	002124	005786
18					000004	000039	000232	000944	002893
19					000001	000012	000085	000397	001370
20						000004	000030	000159	000617
21						000001	000010	000061	000264
22							000003	000022	000108
23							000001	000008	000042
24								000003	000016
25								000001	000006
26									000002
27									000001

Додаток 4

Таблиця значень $t_\gamma = t(\gamma, n)$

n	γ 0.95	γ 0.99	γ 0.999	n	γ 0.95	γ 0.99	γ 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	2,729	3,600
9	2,31	2,36	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	1,001	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	150	1,980	2,617	3,374
18	2,11	2,90	3,97	200	1,960	2,576	3,291
19	2,10	2,88	3,92	250			

Додаток 5

Таблиця значень $q = q(\gamma, n)$

n	γ 0.95	γ 0.99	γ 0.999	n	γ 0.95	γ 0.99	γ 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

Додаток 6

Критичні точки розподілу χ^2

Число степенів вільності k	Рівень значущості α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63

Бібліографічний список

1. Агапов Г.И. Задачник по теории вероятностей. – М.: Высш. шк., 1986. – 80 с.

2. Акулиничев В.М., Кудрявцев В.А., Корешков А.Н. Математические методы в эксплуатации железных дорог. – М.: Транспорт, 1981. – 223 с.

3. Барковський В.В., Барковська Н.В., Лопатін О.К. Теорія ймовірностей та математична статистика: Навч. посібник. – К.: ЦУЛ, 2010. – 424 с.

4. Бабак В.П., Марченко Б.Г., Фриз М.Є. Теорія ймовірностей, випадкові процеси та математична статистика. – К.: Техніка, 2004. – 288 с.

5. Бобик О.І., Берегова Г.І., Копитко Б.І. Теорія ймовірностей і математична статистика: Підручник. – К.: ВД "Професіонал", 2007. – 560 с.

6. Бочаров П.П., Печенкин А.В. Теория вероятностей и математическая статистика. – М.: Гардарика, 1998.

7. Вентцель Е.С. Теория вероятностей. – М.: Наука, 1964. – 576 с.

8. Вентцель Е.С., Овчаров Л.А. Задачи и упражнения по теории вероятностей. – М.: Высш. шк., 2002.

9. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. – М.: Наука,1988.

10. Виленкин Н.Я. Комбинаторика. – М.: Наука, 1969. – 328 с.

11. Войтенко М.А. Руководство к решению задач по теории вероятностей. – М.: Изд-во ВЗФЭИ, 1988.

12. Вища математика: Збірник задач: у 2 ч. Ч.2 / П.П. Овчинников, П.С. Кропив'янський, С.П. Полушкін та ін.; За заг. ред. П.П. Овчинникова. – К.: Техніка, 2003. – 376 с.

13. Гихман И.И., Скороход А.В., Ядренко М.Н. Теория вероятностей и математическая статистика. – К.: Вища шк., 1979.

14. Гмурман В.Е. Теория вероятностей и математическая статистика. − М.: Высшее образование, 2009. − 479 с.

15. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. − М.: Высш. шк., 2006. − 405 с.

16. Гнеденко Б.В. Курс теории вероятностей. – М.: Наука, 1975. – 400 с.

17. Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах. − М.: Высшая школа, 2002. − Ч.2. − 416 с.

18. Донченко В.С., Сидоров М.В.-С., Шарапов М.М. Теорія ймовірностей та математична статистика. – К.: ВЦ “Академія”, 2009. – 288с.

19. Жлуктечко В. І., Наконечний С.І. Теорія ймовірностей і математична статистика: Навч.-метод. Посібник: у 2 ч. – К: КНЕУ, 2000. – ч.1. Теорія ймовірностей. – 304с.

20. Иванова В.М., Калинина В.Н. Математическая статистика. – М.: Высш. шк., 1981.

21. Кармелюк Г.І. Теорія ймовірностей та математична статистика: посібник з розв’язування задач: Навч. посібник. – К.: ЦУЛ, 2007. – 576 с.

22. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. – М.: ИНФРА-М, 2004.

23. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник. – М.: ЮНИТИ, 2009. – 551 с.

24. Мацкевич И.П., Свирид Г.П., Булдык Г.М. Сборник задач и упражнений по высшей математике. Теория вероятностей и математическая статистика. – Минск: Выщэйш. шк.,1996.

25. Мацкевич И.П., Свирид Г.П., Булдык Г.М. Высшая математика (Теория вероятностей и математическая статистика). – Минск: Выщэйш. шк., 1996.

26. Маяковський С.О., Гальків Л.І., Гринькевич О.С., Сорочак О.З. Статистика: Навч. посібник – Львів: Новий світ-2000, 2009. – 430 с.

27. Письменный Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам – М.: Айрис-пресс, 2008. – 288 с. – (Высшее образование).

28. Платонов Г.А., Файнберг М.А., Штильман М.С. Поезда, пассажиры и … математика. – М.: Транспорт, 1977. – 240 с.

29. Сборник задач по математике для втузов. Спец. курсы / Под ред. А.В.Ефимова и Б.П. Демидовича. – М.: Наука, 1986.

30. Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под ред. А.А. Свешникова. – М.: Наука, 1970. – 656 с.

31. Севастьянов Б.А. Курс теории вероятностей и математической статистики. – М.: Наука, 1982. – 256 с.

32. Севастьянов Б.А., Чистяков В.П., Зубков А.М. Сборник задач по теории вероятностей. – М.: Наука, 1980.

33. Соколов Г.А., Чистяков Н.А. Теория вероятностей. – М.: Экзамен, 2005.

34. Фадеева Л.Н. Теория вероятностей и математическая статистика: Курс лекций. – М.: Эксмо, 2006.

35. Чистяков В.П. Курс теории вероятностей: Учебник. – М.: Наука, 1987. – 240 с.

Іменний покажчик

Предметний покажчик