Національний технічний університет України «Київський політехнічний інститут»

Теорія ймовірностей та математична статистика

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний університет»

О. І. Кушлик-Дивульська, Н. В. Поліщук,

Б. П. Орел, $\boxed{\text{П. І. Штабалюк}}$

Теорія ймовірностей та математична статистика

Навчальний посібник

Рекомендовано Міністерством освіти і науки України
як навчальний посібник для студентів
технічних та економічних спеціальностей
вищих навчальних закладів

Київ
НТУУ «КПІ»
2014

УДК 519.2(075.8)
ББК 22.17я73
Т33

Рекомендовано Міністерством освіти і науки України
(Лист №1/11-10451 від 08.07.2014 р.)

Рецензенти:

М. В. Працьовитий, д-р фіз.-мат. наук, проф.,
заслужений діяч науки і техніки України,
Інститут математики НАН України

О. В. Капустян, д-р фіз.-мат. наук, проф.,
Київський національний університет імені Тараса Шевченка

Відповідальний редактор

С. Д. Івасишен, д-р фіз.-мат. наук, проф.,
Національний технічний університет України
«Київський політехнічний інститут»

Т33 **Теорія** ймовірностей та математична статистика: навч. посіб./ О. І. Кушлик-Дивульська, Н. В. Поліщук, Б. П. Орел, П. І. Штабалюк. – К: НТУУ «КПІ», 2014. – 212 с. − Бібліогр.: с.205. − 300пр.

ISBN 978-966-622-654-2

Розглянуто основи теорії ймовірностей та математичної статистики. Подано весь теоретичний матеріал, що відповідає навчальній програмі кредитного модуля «Теорія ймовірностей і математична статистика» напряму підготовки «Менеджмент». Наведено детальні алгоритми розв’язання всіх основних задач теорії ймовірностей та математичної статистики, надано чіткі інструкції застосування засобів стандартного програмного забезпечення у разі розв’язання задач із трудомісткими алгоритмами. Підібрано достатній масив завдань для аудиторної роботи та індивідуальних завдань для самостійної роботи. Наведено необхідний довідковий матеріал.

Для студентів економічних і технічних спеціальностей Національного технічного університету України «Київський політехнічний інститут», студентів інших навчальних закладів усіх форм навчання, що мають схожу програму підготовки, а також для всіх зацікавлених осіб.

УДК 519.2(075.8)
ББК 22.17я73

ISBN 978-966-622-654-2

Передмова

Теорія ймовірностей та математична статистика – математичні науки, які вивчають закономірності в масових випадкових явищах, – це складова теоретичної основи викладання багатьох економічних, соціологічних та спеціальних дисциплін.

Мета цього навчального посібника – ознайомити студентів економічних і технічних спеціальностей, а також усіх зацікавлених, з основними поняттями, методами, теоремами та формулами теорії ймовірностей і математичної статистики, допомогти їм набути первинних навичок застосування теоретичного матеріалу на практиці.

Навчальний посібник «Теорія ймовірностей та математична статистика» призначений для студентів денної та заочної форм навчання Видавничо-поліграфічного інституту.

Посібник складається з трьох розділів, кожен з яких розподілено за темами. Достатньо і доступно викладено основний теоретичний матеріал, передбачений програмою для економічних спеціальностей вищих навчальних закладів, який апробовано для підготовки студентів за напрямом 6.030601 «Менеджмент». До кожної теми наведено приклади розв'язання, що сприяє більш повному засвоєнню матеріалу.

Навчальний посібник містить завдання для самостійної роботи студентів: 13 наборів завдань, кожен з яких має 10 варіантів. У навчальний посібник включено завдання підсумкової модульної контрольної роботи та розрахункові завдання, які можуть бути використані як індивідуальні домашні чи контрольні завдання для студентів заочної форми навчання технічних та економічних спеціальностей. Також запропоновано контрольні тести, які дають можливість перевірити базові знання та готовність студента до практичного їх застосування. Додаток містить усі таблиці, необхідні для розв'язання задач.

У навчальному посібнику наведено розв'язання деяких практичних задач з теорії ймовірностей засобами Excel. Це дає можливість застосовувати алгоритми Excel також для практичного використання статистичних задач, трудомістких через значну кількість математичних обчислень. Викладання цього матеріалу розраховане на широке коло користувачів, які мають доступ до програмного забезпечення і початкові необхідні знання з теорії ймовірностей.

До більшої частини прикладів та задач, як для самостійної, модульної чи індивідуальної роботи, наведено відповіді.

Автори вдячні студентам Видавничо-поліграфічного інституту за спільну працю і підготовку задач для навчального посібника.

У текст видання внесено уточнення і виправлення. Автори вдячні співробітникам кафедри математичної фізики, кафедри теорії ймовірностей фізико-математичного факультету НТУУ «КПІ» та всім читачам, що висловили свої побажання та зауваження.

Автори вважають, що посібник буде корисний не тільки для студентів економічних і технічних спеціальностей різних форм навчання, але й для фінансистів, бізнесменів, співробітників податкової інспекції та страхових компаній, соціологів та політологів.

Частина I

Основні поняття теорії ймовірностей. Теореми додавання і множення ймовірностей. Послідовності випробувань

Розділ 1. Основні поняття теорії ймовірностей та комбінаторики

Вивчення курсу «Теорія ймовірностей» передбачає знання основних розділів курсу «Вища математика». Початковий розділ, в якому розглянуто означення основних понять теорії ймовірностей, вимагає знання алгебри висловлювань, теорії множин, комбінаторики.

1.1. Коротка історична довідка

Теорія ймовірностей – математична наука, що вивчає закономірності в масових випадкових явищах.

Основні поняття, методи, теореми та формули теорії ймовірностей ефективно використовують у науці, техніці, економіці, зокрема теоріях надійності та масового обслуговування, в плануванні та організації виробництва, страховій та податковій справах, соціології та політології, демографії та охороні здоров'я.

Теорія ймовірностей як наука зародилась в середині XVII ст., хоча перші роботи, в яких з'явились основні поняття теорії ймовірностей (XV – XVI ст.), виникли як спроби побудови теорії азартних ігор і належать таким видатним вченим, як Б. Спіноза, Дж. Кардано, Г. Галілей. Наступний етап розвитку теорії ймовірностей пов'язаний з роботами Б. Паскаля, П. Ферма, К. Гаусса, С. Пуассона, А. Муавра, П. Лапласа, Т. Байєса. Я. Бернуллі зробив перші теоретичні обґрунтування зібраних раніше матеріалів. Аксіоматичне означення ймовірності (1933) належить А. М. Колмогорову.

Вимоги різних технічних потреб суспільства дали поштовх до розвитку низки прикладних розділів теорії ймовірностей. Це статистична теорія зв'язку, теорія інформації, теорія масового обслуговування, теорія надійності, економетрія, математична статистика тощо.

Доречно навести слова засновника української школи теорії ймовірностей академіка Б. В. Гнєденка: «Теорія ймовірностей з дедалі більшою швидкістю почала посідати основні позиції в прикладній математиці й одночасно зміцнювати своє становище як одна з центральних математичних дисциплін» [3].

1.2. Елементи комбінаторики

Комбінаторикою називають розділ математики, що вивчає правила групування елементів множини. Ров'язання задач ґрунтується на основному правилі та принципах комбінаторики, конкретних способах підрахунку кількості підмножин, які можна утворити з певної множини елементів, що розглянемо далі.

1.2.1. Два основні принципи комбінаторики

Основне правило комбінаторики. Між скінченними множинами A та B можна встановити взаємно однозначну відповідність тоді і тільки тоді, коли вони мають однакову кількість елементів.

Основне правило комбінаторики дозволяє виконувати обчислення кількості елементів множини фактично зведенням до множини з відомою кількістю елементів або до множини, для якої кількість можна підрахувати дещо простіше. У більшості випадків його використання не акцентується, а приймається як дещо зрозуміле.

Значна кількість формул і теорем комбінаторики ґрунтуються на двох основних елементарних принципах, які називаються принципами суми та добутку.

Принцип суми. Нехай $n(A)$, $n(B)$ – кількості елементів скінченних неперетинних множин A та B відповідно. Тоді для множини $C = A \bigcup B$ кількість елементів

$$n(C) = n(A) + n(B). \tag{1.1}$$

Принцип добутку. Нехай потрібно послідовно виконати k дій, причому першу дію можна виконати n_1 способом, після чого другу дію – n_2 способами і так далі до k-ї дії, яку можна виконати n_k способами. Тоді всі k дій можна виконати $n_1 \cdot n_1 \cdot n_2 \cdot \ldots \cdot n_k$ способами.

1.2.2. Розміщення, перестановки, комбінації

Нехай M **–** множина, що містить n елементів.

Означення. **Розміщенням** із n елементів по k називають довільну впорядковану підмножину з k елементів із множини M.

Два розміщення вважають різними не тільки тоді, коли вони складаються з різних елементів, але й тоді, коли вони складаються з однакових елементів, але відрізняються порядком їх розташування.

Кількість розміщень з n елементів по k ($k \le n$) позначають A_n^k і обчислюють за формулою

$$A_n^k = \frac{n!}{(n-k)!} = n(n-1)(n-2)\cdot\ldots(n-k+1). \qquad (1.2)$$

Означення. **Розміщенням з повтореннями** з n елементів по k називають довільну впорядковану підмножину з k елементів із множини M (елементи не обов’язково різні).

Кількість розміщень із повтореннями з n елементів по k обчислюють за формулою

$$\overline{A_n^k} = n^k. \qquad (1.3)$$

Зауваження. $\overline{A_n^k}$ – це кількість способів, якими можна розкласти k різних предметів у n ящиків.

Приклад 1. Скласти всі двозначні числа з трьох цифр 3, 5, 7 (числа з однаковими цифрами не враховувати).

Розв’язання. Маємо 35, 37, 53, 57, 73, 75. Загальну кількість різних чисел можна підрахувати за формулою (1.2): $A_3^2 = 3\cdot 2 = 6.$

Приклад 2. У ліфт шістнадцятиповерхового будинку зайшло на першому поверсі 10 осіб. Скількома способами вони можуть вийти з ліфта, починаючи з другого поверху?

Розв’язання. Задача зводиться до розкладання 10 предметів у 15 ящиків. Кількість таких способів дорівнює $\overline{A_{15}^{10}} = 15^{10}$.

Означення. Розміщення з n елементів по n називають **перестановками.**

Різні перестановки відрізняються лише порядком розташування елементів.

Кількість перестановок з n елементів дорівнює добутку всіх натуральних чисел від 1 до n

$$P_n = 1 \cdot 2 \cdot \ldots \cdot n = n! \tag{1.4}$$

Означення. **Перестановками з повтореннями** називають розміщення з n елементів, які мають однотипні елементи.

Кількість різних перестановок з повторенням, які можна утворити з n елементів, серед яких є k_1 елементів першого типу, k_2 елементів другого типу, …, k_m елементів m-го типу,

$$P_n(k_1, k_2, \ldots, k_m) = \frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_m!}. \tag{1.5}$$

Приклад 3. Скількома способами можна скласти список з п’яти студентів?

Розв’язання. За формулою (1.4) отримаємо:

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

Приклад 4. Скільки різних «слів» можна утворити перестановкою літер у слові «математика»?

Розв’язання. Слово «математика» містить $n = 10$ літер. Літер одного типу: «м» – дві, «а» – три, «т» – дві, «е» – одна, «и» – одна, «к» – одна. За формулою (1.5) отримаємо:

$$P_{10}(2,\ 3,\ 2,\ 1,\ 1,\ 1) = \frac{10!}{2! \cdot 3! \cdot 2! \cdot 1! \cdot 1! \cdot 1!} = 151\,200.$$

Означення. **Комбінацією** (сполукою) з n елементів по k називають довільну підмножину з k елементів із множини M, яка відрізняється одна від одної принаймні одним елементом.

Порядок розташування елементів у комбінаціях не має значення. Кількість комбінацій з n елементів по k ($k \le n$) позначають C_n^k та обчислюють за формулою

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{1.6}$$

Означення. **Комбінацією з повтореннями** з n елементів по k називають довільну множину з k елементів із множини M (елементи не обов’язково різні).

Кількість комбінацій з повтореннями з n елементів по k обчислюють за формулою

$$\overline{\tilde{N}_n^k} = \tilde{N}_{n+k-1}^k = \frac{(n+k-1)!}{k!(n-1)!}. \tag{1.7}$$

Зауваження. $\overline{\tilde{N}_n^k}$ – це кількість способів, якими можна розкласти k однакових предметів по n ящиках.

Приклад 5. На конференції присутні 30 студентів. Скількома способами можна обрати президію у складі трьох осіб?

Розв'язання. Кількість способів дорівнює кількості комбінацій із 30 елементів по 3:

$$\tilde{N}_{30}^3 = \frac{30!}{3!\cdot 27!} = \frac{30\cdot 29\cdot 28}{1\cdot 2\cdot 3} = 4060.$$

Приклад 6. У кондитерській є 6 різних сортів тістечок. Скількома способами можна купити 8 тістечок?

Розв'язання. Шукане число

$$\overline{\tilde{N}_6^8} = \tilde{N}_{13}^8 = \tilde{N}_{13}^5 = \frac{13!}{5!\cdot 8!} = \frac{13\cdot 12\cdot 11\cdot 10\cdot 9}{1\cdot 2\cdot 3\cdot 4\cdot 5} = 1287.$$

1.3. Простір елементарних подій. Випадкові події та операції над ними

За допомогою теорії ймовірностей та математичної статистики в науці та техніці досліджують випадкові явища і події та їх закономірності.

1.3.1. Простір елементарних подій

Означення. Експерименти, які в незмінних умовах можна повторити будь-яку кількість разів, але результати яких непередбачувані, у теорії ймовірностей називають **випробуваннями**.

Означення. **Елементарною подією** називають найпростіший результат випробування в певному експерименті.

Означення. Сукупність усіх можливих елементарних подій випробування називають **простором елементарних** подій і позначають Ω.

Означення. Підмножина A простору елементарних подій називається **випадковою подією**.

1.3.2. Випадкові події та операції над ними

Перш ніж продовжити викладання понять теорії ймовірностей, потрібно розглянути деякі важливі співвідношення між випадковими подіями. Для випадкових подій вводять операції, які можна подати мовою операцій із множинами.

Означення. **Сумою** *(об'єднанням)* подій *A* і *B* називають подію $A+B$ (àáî æ $A\bigcup B$), яка настає тоді, коли настає принаймні одна з подій *A* чи *B* (елементарні події, сприятливі або для події *A*, або для події *B*, або для обох подій *A* і *B*).

Означення. **Добутком** *(перетином)* *A* і *B* називають подію AB (àáî æ $A\bigcap B$), яка настає тоді, коли настають обидві події *A* і *B* (елементарні події, які сприятливі і для події *A*, і для події *B*).

Означення. Події *A* і *B* називають **несумісними**, якщо $A\bigcap B=\varnothing$. Несумісні події не можуть відбуватися одночасно.

Означення. **Різницею** подій *A* і *B* називають подію $A\backslash B$, яка настає тоді, коли настає подія *A* і не настає подія *B*.

Означення. Подію $\overline{A}$ називають **протилежною** до події *A*, якщо $\overline{A}$ настає тоді, коли не настає *A* (сприятливими для $\overline{A}$ є всі елементарні події, які не входять в *A*).

Властивості операцій над подіями:

1. $A\bigcup\overline{A}=\Omega$, $A\bigcap\overline{A}=\varnothing$, $\overline{A}=\Omega\backslash A$.
2. $A\bigcup B=B\bigcup A$, $A\bigcap B=B\bigcap A$.
3. $(A\bigcup B)\bigcup C=A\bigcup(B\bigcup C)$, $(A\bigcap B)\bigcap C=A\bigcap(B\bigcap C)$.
4. $(A\bigcup B)\bigcap C=(A\bigcap C)\bigcup(B\bigcap C)$.
5. $A\bigcap B\bigcup C=(A\bigcup C)\bigcap(B\bigcup C)$.
6. $\overline{A\bigcup B}=\overline{A}\bigcap\overline{B}$, $\overline{A\bigcap B}=\overline{A}\bigcup\overline{B}$ (закони де Моргана).

1.4. Ймовірності подій

Ще до того, як було введене поняття ймовірності, що ґрунтувалось на відносній частоті настання події, під час дослідження азартних ігор використовувалось інше поняття ймовірності, яке називають зараз класичним.

Розглянемо класичне означення ймовірності, а також статистичне та геометричне поняття ймовірності.

1.4.1. Класичне означення ймовірності

Класичне означення ймовірності ввів Лаплас (1749–1827).

Означення. **Ймовірністю** події A називають відношення кількості результатів випробування, сприятливих для A, до кількості всіх рівноможливих і попарно несумісних наслідків випробування:

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

Побудова логічно повноцінної теорії ймовірностей ґрунтується на аксіоматичному визначенні випадкової події і її ймовірності. У системі аксіом, запропонованій О. М. Колмогоровим, невизначуваними поняттями є елементарна подія і ймовірність. Приведемо аксіоми, що визначають імовірність:

1. Кожній події A поставлено у відповідність невід'ємне дійсне число $P(A)$. Це число називається ймовірністю події A.

2. Ймовірність достовірної події дорівнює одиниці:

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

3. Ймовірність настання хоча б однієї із попарно несумісних подій дорівнює сумі ймовірностей цих подій.

Якщо події $A_1, A_2, \ldots$ попарно несумісні, то ймовірність суми подій дорівнює сумі ймовірностей цих подій:

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

Аксіому 3 називають аксіомою адитивності і для подій A і B, що є несумісні

$$P(A \cup B) = P(A) + P(B).$$

Аксіому додавання ймовірностей інколи називають «теоремою додавання» (для дослідів, що зводяться до схеми випадків, її можна довести).

Виходячи з цих аксіом, властивості ймовірностей і залежності між ними виводять як теореми.

Приклад 1. Знайти ймовірність того, що кількість очок, яка випадає на гральному кубику за одного підкидання, буде непарною (подія A).

Розв'язання. Простір елементарних подій $\Omega = \{1, 2, 3, 4, 5, 6\}$ містить $n = 6$ елементарних подій, подія $A = \{1, 3, 5\}$ містить $m = 3$ сприятливі події. Тоді

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

Приклад 2. З урни, в якій 12 білих та 8 чорних кульок, виймають навмання 3 кульки. Яка ймовірність того, що:

1) всі три кульки будуть чорними (подія *A*);

2) дві кульки будуть чорними та одна білою (подія *B*).

Розв'язання. 1. Кількість усіх елементарних подій дорівнює кількості комбінацій з 20 по 3, тобто

$$n = C_{20}^{3} = \frac{20!}{3!(20-3)!} = \frac{18 \cdot 19 \cdot 20}{1 \cdot 2 \cdot 3} = 1140\,.$$

Кількість випадків, які сприяють події *A*, становить $m = C_8^3 = \frac{8!}{3! \cdot 5!} = \frac{6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3} = 56$.

Тоді ймовірність виймання трьох чорних кульок

$$P(A) = \frac{C_8^3}{C_{20}^3} = \frac{56}{1140} = \frac{14}{285} \approx 0,049\,.$$

2. Як і в п. 1 $n = C_{20}^3 = 1140$. Знаходимо кількість сприятливих подій: дві чорні кульки можна взяти з восьми чорних $\tilde{N}_8^2 = \frac{8!}{2! \cdot 6!} = \frac{7 \cdot 8}{2} = 28$ способами, а одну білу з 12 кульок $\tilde{N}_{12}^1 = 12$ способами. За правилом множення (принцип добутку) $m = C_8^2 \cdot C_{12}^1$. Отже,

$$P(B) = \frac{m}{n} = \frac{C_8^2 \cdot C_{12}^1}{C_{20}^3} = \frac{28 \cdot 12}{1140} = \frac{84}{285} \approx 0,29.$$

1.4.2. Статистичне означення ймовірності

***Означення*. Відносною частотою** події називають відношення кількості випробувань, в яких подія настала *m* разів, до загальної кількості *n* фактично проведених випробувань:

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

Відносна частота $\frac{m}{n}$ прямує за ймовірністю до числа *p*, якщо для довільного $\varepsilon > 0$ ймовірність того, що $\left|\frac{m}{n} - p\right| < \varepsilon$ прямує до одиниці за $n \to \infty$, тобто

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

Відносна частота може бути обчислена після того, коли проведено серію експериментів, і, загалом кажучи, вона змінюється, якщо здійснити іншу серію з *n* експериментів або якщо змінити *n*. Однак практика показує, що за достатньо великих *n* для більшості таких серій експериментів відносна частота має властивість стійкості.

Означення. Число p називають **статистичною ймовірністю** статистично стійкої події, якщо відносна частота цієї події прямує за ймовірністю до числа p за $n \to \infty$, тобто $\frac{m}{n} \xrightarrow{p} p,\ n \to \infty.$

Зауважимо, що теорія ймовірностей має справу тільки зі статистично стійкими експериментами, і коли кажуть, що ймовірність деякої випадкової події дорівнює, наприклад 0,28, то практично це означає, що у середньому в кожних 100 дослідах ця подія відбувається 28 разів.

1.4.3. Геометричні ймовірності

Нехай Ω – деяка область на прямій, площині або в просторі, *A* – деяка частина області Ω. В області Ω навмання вибирають точку, вважаючи, що вибір точок області рівноможливий. Імовірність того, що вибрана точка належить *A*, визначається рівністю

$$P(A) = \frac{\text{mes } A}{\text{mes } \Omega},$$

де $\text{mes } A$, $\text{mes } \Omega$ – міра (довжина, площа, об’єм) A, Ω.

Приклад 3. У коло радіусом R вписано правильний трикутник. Визначити ймовірність того, що навмання поставлена в коло точка потрапить у трикутник.

Розв'язання. Нехай подія A – точка потрапляє у трикутник. Тоді

$$\text{mes } A = S_\Delta = \frac{3\sqrt{3}}{4}R^2, \ \text{mes } \Omega = \pi R^2.$$

За означенням геометричної ймовірності

$$P(A) = \frac{\text{mes } A}{\text{mes } \Omega} = \frac{3\sqrt{3}}{4\pi}.$$

1.5. Теореми додавання ймовірностей

Теореми додавання і множення ймовірностей використовують для визначення ймовірностей складних подій, які можна записати через інші події. Розглянемо теореми додавання для сумісних та несумісних подій.

1.5.1. Теорема додавання для несумісних подій

Означення. Дві події називають **несумісними**, якщо вони не можуть одночасно настати в одному досліді, іншими словами, настання однієї події виключає можливість настання другої.

Означення. Події $A_1, A_2, \ldots, A_k$ утворюють **повну групу** несумісних подій, якщо $A_i \cap A_j = \varnothing$ за $i \neq j$ $(i, \ j = 1, \ 2, \ \ldots, k)$, $\bigcup\limits_{i=1}^{k} A_i = \Omega$.

Теорема додавання для несумісних подій. *Якщо події A і B несумісні ($A \cap B = \varnothing$), причому відомі їх імовірності P(A) і P(B), то ймовірність суми цих подій дорівнює сумі їх імовірностей:*

$$P(A \cup B) = P(A) + P(B). \tag{1.9}$$

Доведення. Дійсно, нехай n – кількість усіх елементарних подій у певному досліді; m_1 – кількість елементарних подій, сприятливих події A; m_2 – кількість елементарних подій, сприятливих події B. Тоді настанню події $A \cup B$ сприяють $m_1 + m_2$ елементарних подій. Отже, за класичним означенням імовірності

$$P(A \bigcup B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B).$$

Наслідок 1. Якщо події A_1, A_2, …, A_k попарно несумісні, то ймовірність суми подій дорівнює сумі ймовірностей цих подій:

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

Наслідок 2. Ймовірність протилежної до A події $\overline{A}$:

$$P(\overline{A}) = 1 - P(A). \qquad (1.10)$$

Дійсно, оскільки $A \bigcup \overline{A} = \Omega$, то $P(A \bigcup \overline{A}) = P(\Omega) = 1$. З другого боку, $P(A \bigcup \overline{A}) = P(A) + P(\overline{A})$.

Отже, $P(A) + P(\overline{A}) = 1$, звідки $P(\overline{A}) = 1 - P(A)$.

Наслідок 3. Сума ймовірностей подій A_1, A_2, …, A_k, що утворюють повну групу подій, дорівнює одиниці:

$$P(A_1) + P(A_2) + \ldots + P(A_k) = 1.$$

Приклад 4. У партії з 20 деталей є 16 стандартних. Знайти ймовірність того, що серед навмання взятих трьох деталей виявиться принаймні одна стандартна.

Розв'язання. Нехай подія A_1: виявиться одна стандартна; подія A_2: виявиться дві стандартні; подія A_3: виявиться три стандартні деталі. Ці події попарно несумісні.

Нехай подія A: серед навмання взятих трьох деталей виявиться принаймні одна стандартна. Отже, $A = A_1 \cup A_2 \cup A_3$, і за формулою (1.9) маємо:

$$P(A) = P(A_1) + P(A_2) + P(A_3).$$

Обчислимо ймовірності подій A_1, A_2, A_3:

$$P(A_1) = \frac{C_{16}^1 \cdot C_4^2}{C_{20}^3} = \frac{8}{95}; \; P(A_2) = \frac{C_{16}^2 \cdot C_4^1}{C_{20}^3} = \frac{40}{95}; \; P(A_3) = \frac{C_{16}^3}{C_{20}^3} = \frac{28}{57}.$$

Отже, $P(A) = \frac{8}{95} + \frac{40}{95} + \frac{28}{57} = \frac{284}{285}$.

Інший спосіб. Під час розв'язування задач часто буває зручно переходити до протилежної події. Так, якщо подія A_0 – не виявиться жодної стандартної деталі, тоді подія A_0 протилежна до події A, і події A_0, A_1, A_2, A_3 утворюють повну групу попарно несумісних подій. Отже, будемо використовувати формулу (1.10). Обчисливши ймовірність $P(A_0) = \frac{C_4^3}{C_{20}^3} = \frac{1}{285}$, отримаємо $P(A) = 1 - \frac{1}{285} = \frac{284}{285}$.

1.5.2. Теорема додавання для сумісних подій

Нехай дві події A і B сумісні, причому відомі ймовірності цих подій $P(A)$, $P(B)$ та ймовірність їх сумісного настання $P(A \cap B)$.

Теорема додавання (для сумісних подій). *Імовірність настання принаймні однієї з двох сумісних подій A і B дорівнює сумі ймовірностей цих подій без імовірності їх спільного настання*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1.11}$$

Доведення. Дійсно, оскільки події A і B сумісні, то подія $A \cup B$ настане, якщо настане одна з трьох несумісних подій $A \cap \overline{B}$, $\overline{A} \cap B$ або $A \cap B$:

$$A \cup B = (A \cap \overline{B}) \cup (\overline{A} \cap B) \cup (A \cap B).$$

Подія A настане, якщо настане одна з двох несумісних подій $A \cap \overline{B}$ або $A \cap B$. Аналогічно, подія B настане, якщо настане одна з двох несумісних подій $\overline{A} \cap B$ або $A \cap B$.

Отже, $A = (A \cap \overline{B}) \cup (A \cap B)$; $B = (\overline{A} \cap B) \cup (A \cap B)$,
і за теоремою додавання ймовірностей несумісних подій (1.9) маємо:

$$P(A) = P(A \cap \overline{B}) + P(A \cap B) \Rightarrow P(A \cap \overline{B}) = P(A) - P(A \cap B);$$

$$P(B) = P(\overline{A} \cap B) + P(A \cap B) \Rightarrow P(\overline{A} \cap B) = P(B) - P(A \cap B).$$

Таким чином,

$$P(A\bigcup B)=P(A)-P(A\bigcap B)+P(B)-P(A\bigcap B)+P(A\bigcap B)=$$
$$=P(A)+P(B)-P(A\bigcap B).$$

Зауваження. Теорема може бути узагальнена на довільну скінченну кількість сумісних подій. Наприклад, для трьох сумісних подій

$$P(A\bigcup B\bigcup C)=P(A)+P(B)+P(C)-P(A\bigcap B)-P(A\bigcap C)-P(B\bigcap C)+$$
$$+P(A\bigcap B\bigcap C).$$

Приклад 5. В ящику 10 червоних і 6 синіх ґудзиків. Навмання виймають два ґудзики. Визначити ймовірність того, що ґудзики будуть одного кольору.

Розв'язання. Нехай подія A – ґудзики одного кольору, подія B – ґудзики червоні, подія C – ґудзики сині. Очевидно, $A=B\bigcup C$, і події B і C несумісні, тому за теоремою 1 $P(A)=P(B)+P(C)$. Знайдемо $P(B)$ та $P(C)$:

$$P(B)=\frac{C_{10}^2}{C_{16}^2}=\frac{10\cdot 9}{1\cdot 2}\cdot\frac{1\cdot 2}{16\cdot 15}=\frac{3}{8};\ P(C)=\frac{C_6^2}{C_{16}^2}=\frac{6\cdot 5}{1\cdot 2}\cdot\frac{1\cdot 2}{16\cdot 15}=\frac{1}{8}.$$

Отже, $P(A)=\frac{3}{8}+\frac{1}{8}=\frac{1}{2}.$

Розділ 2. Основні теореми теорії ймовірностей. Послідовності випробувань

Після того, як навчилися обчислювати ймовірності випадкових подій, користуючись різними формами означення ймовірності випадкової події, можна перейти на наступний рівень. Розглянемо задачі на обчислення ймовірностей складних подій з урахуванням їх сумісності чи несумісності, залежності чи незалежності. Такі задачі розв'язують за допомогою теорем додавання і множення ймовірностей. Особливу увагу має бути приділено класичній схемі незалежних випробувань – схемі Бернуллі та відповідним граничним теоремам.

2.1. Умовні ймовірності та незалежні події

Для теорем множення ймовірностей використовують поняття умовної ймовірності та залежних і незалежних подій.

2.1.1. Теореми множення ймовірностей

Як зазначалось, в основі означення ймовірності випадкової події лежить сукупність певних умов. Якщо ж ніяких інших обмежень, окрім цих умов, у разі обчислення ймовірності $P(A)$ не накладається, така ймовірність називається ***безумовною***. Якщо ж подія A відбувається за умови, що відбулась інша подія B, причому $P(B) > 0$, то ймовірність настання події A називають ***умовною*** і обчислюють за формулою

$$P\left(A \mid B\right) = \frac{P(A \cap B)}{P(B)}, \ \left(P(B) > 0\right). \tag{2.1}$$

Означення. Ймовірність настання події A за умови, що подія B вже відбулася, називають **умовною** і позначають $P\left(A \mid B\right)$.

Тоді $$P\left(A \mid B\right) = \frac{N(A \cap B)}{N(B)} = \frac{\dfrac{N(A \cap B)}{n}}{\dfrac{N(B)}{n}} = \frac{P(A \cap B)}{P(B)}, \ P(B) > 0,$$

де n – кількість наслідків стохастичного експерименту, $N(B)$, $N(A \cap B)$ – кількості наслідків стохастичного експерименту, що сприяють відповідно події B та перетину події A з подією B.

Означення. **Добутком** двох подій *A* і *B* називають подію *C*, яка полягає в сумісному настанні цих подій.

Означення. Подію *A* називають **незалежною** від події *B*, якщо ймовірність події *A* не залежить від того, чи відбулася подія *B* і справедлива рівність

$$P\left(A \mid B\right) = P(A). \tag{2.2}$$

Означення. Подію *A* називають **залежною** від події *B*, якщо ймовірність події *A* залежить від того, чи відбулася подія *B*, тобто $P\left(A \mid B\right) \neq P(A)$.

Теорема множення ймовірностей залежних подій. Розглянемо дві залежні події A і B, причому відомі ймовірності $P(A)$ і $P\left(B \mid A\right)$. Ймовірність суміщення цих подій обчислюють за теоремою:

Імовірність сумісного настання двох залежних подій дорівнює добутку ймовірності однієї з них на умовну ймовірність іншої, обчисленої за умови, що перша відбулася:

$$P(AB) = P(A) \cdot P\left(B \mid A\right)$$

або

$$P(AB) = P(B) \cdot P\left(A \mid B\right). \tag{2.3}$$

Наслідок. Якщо події A_i ($i = \overline{1,n}$) залежні, то

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_1\right) \cdot P\left(A_2 \mid A_1\right) \cdot P\left(A_3 \mid \left(A_1 \cap A_2\right)\right) \cdot \ldots \cdot P\left(A_n \mid \left(A_1 \cap A_2 \cap \cdots \cap A_{n-1}\right)\right),$$

тобто ймовірність сумісного настання декількох залежних подій дорівнює добутку ймовірності однієї з них на умовні ймовірності решти, причому ймовірність кожної наступної події обчислюється за припущення, що всі попередні події відбулися.

Зокрема, для трьох залежних подій *A, B, C* маємо:

$$P\left(A \cap B \cap C\right) = P(A) \cdot P(B \mid A) \cdot P(C \mid A \cap B). \tag{2.4}$$

Теорема множення ймовірностей незалежних подій. *Для двох незалежних подій A і B*

$$P(A \cap B) = P(A) \cdot P(B). \tag{2.5}$$

Означення. Події називають **незалежними в сукупності**, якщо ймовірність будь-якої події не залежить від настання довільної групи інших.

Події A_1, A_2, … A_n будуть незалежними в сукупності тоді і тільки тоді, коли імовірність добутку довільної їх комбінації дорівнює добутку відповідних імовірностей, тобто

$$P\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right) = P\left(A_{i_1}\right) \cdot P\left(A_{i_2}\right) \cdot \ldots \cdot P\left(A_{i_k}\right)$$

для будь-яких $k = 2,\ 3,\ \ldots n,\ \ 1 \le i_1 < i_2 < \ldots < i_k \le n.$

Якщо ця рівність виконується за $k = 2$, то події A_1, A_2, … A_n називають **попарно незалежними**. Попарно незалежні події можуть не бути незалежними в сукупності.

Приклад 1 (задача Бернштейна). На трьох гранях правильного тетраедра, зробленого з однорідного матеріалу, написано цифри 1, 2, 3, а на четвертій грані – всі три цифри 1, 2, 3. Позначимо A_i (i = 1, 2, 3) подію, яка полягає в тому, що підкинутий тетраедр впаде на грань, на якій є цифра i. Показати, що події A_1, A_2, A_3 попарно незалежні, але не незалежні в сукупності.

Розв'язання. Очевидно, що $P(A_i) = \dfrac{2}{4} = \dfrac{1}{2}\ (i = 1,\ 2,\ 3)$. Покажемо, що події A_1, A_2, A_3 попарно незалежні. Оскільки $P(A_1 \cap A_2) = \dfrac{1}{4}$ (подія A_1, A_2 настає тоді і тільки тоді, коли тетраедр падає на четверту грань), то $P\left(A_1 \cap A_2\right) = = P\left(A_1\right) \cdot P\left(A_2\right)$, тобто події A_1 та A_2 незалежні. Аналогічно незалежні події A_1 та A_3, A_2 та A_3. Але події A_1, A_2, A_3 не незалежні в сукупності, оскільки

$$P\left(A_1 \cap A_2 \cap A_3\right) = \frac{1}{4} \ne P\left(A_1\right) \cdot P\left(A_2\right) \cdot P\left(A_3\right) = \frac{1}{8}.$$

Приклад 2. Студент прийшов на екзамен, підготувавши лише 20 з 25 питань програми. Екзаменатор задав йому три запитання. Знайти ймовірність того, що студент знає відповіді на них.

Розв'язання. Нехай подія A – студент знає відповіді на всі три запитання; подія A_i (i = 1, 2, 3) – студент знає відповідь на i-те запитання. Тоді $A = A_1 \cap A_2 \cap A_3$ і за формулою умовної ймовірності

$$P(A) = P(A_1) \cdot P\left(A_2 \mid A_1\right) \cdot P\left(A_3 \mid A_1 \cap A_2\right).$$

Очевидно, що $P(A_1)=\frac{20}{25}$, $P\left(A_2 \mid A_1\right)=\frac{19}{24}$ (оскільки залишилось запитань 24, з яких студент знає 19), аналогічно $P\left(A_3 \mid A_1 \cap A_2\right)=\frac{18}{23}$. Підставивши у формулу, отримуємо:

$$P(A)=\frac{20}{25}\cdot\frac{19}{24}\cdot\frac{18}{23}=\frac{57}{115}=0{,}496.$$

2.1.2. Імовірність настання принаймні однієї події

Нехай у результаті n послідовних випробувань може настати n подій $A_1, A_2, \ldots, A_n$, незалежних в сукупності, ймовірності яких відомі $p_1, p_2, \ldots, p_n$. Тоді ймовірності протилежних подій $\overline{A_1}, \overline{A_2}, \ldots, \overline{A_n}$ дорівнюють

$$q_1=1-p_1,\ q_2=1-p_2,\ \ldots,\ q_n=1-p_n.$$

Теорема. *Ймовірність настання принаймні однієї з подій $A_1, A_2, \ldots, A_n$, незалежних в сукупності, обчислюють за формулою*

$$P(A)=1-q_1q_2\ldots q_n. \tag{2.6}$$

Наслідок. Якщо події $A_1, A_2, \ldots, A_n$ мають однакову ймовірність p, то ймовірність настання принаймні однієї з цих подій

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

Приклад 3. Ймовірність принаймні одного влучення в ціль за чотирьох пострілів дорівнює 0,9984. Знайти ймовірність влучення в ціль за одного пострілу.

Розв'язання. Нехай подія A – принаймні одне влучення в ціль за чотирьох пострілів, p – ймовірність влучення за одного пострілу, $q=1-p$ – ймовірність промаху за одного пострілу. Тоді $P(A)=1-q^4=0{,}9984$, тобто $q^4=0{,}0016$, $q=0{,}2$. Отже, $p=1-q=0{,}8$.

2.1.3. Надійність системи

Означення. **Надійністю** системи називають імовірність її безвідмовної роботи протягом певного часу t (гарантійний термін).

Системи складаються з n елементів, з'єднаних послідовно (рис. 2.1) або паралельно (рис. 2.2).

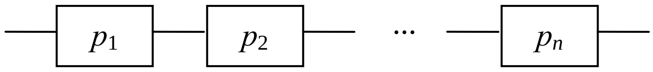


Рис. 2.1. Схема послідовного з'єднання елементів

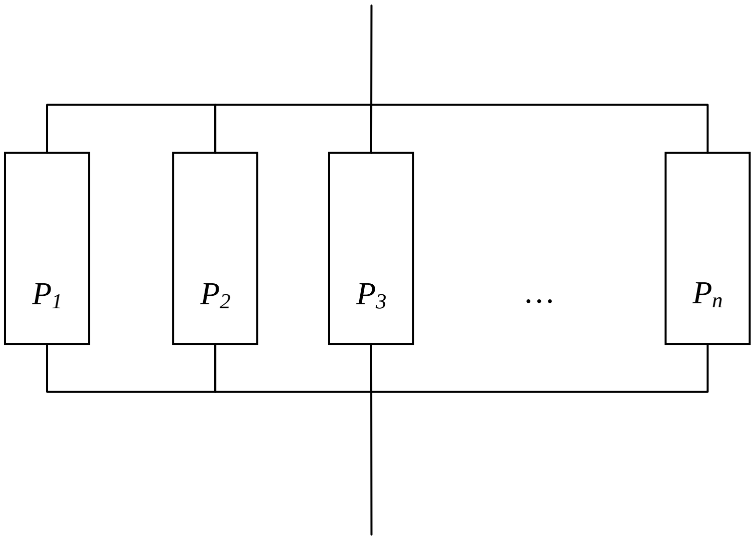


Рис. 2.2. Схема паралельного з'єднання елементів

У разі обчислення надійності систем потрібно виразити їх надійність через надійність елементів та блоків. Надійність елементів вважать відомою, оскільки вона пов'язана з технологією їх виготовлення та умовами експлуатації.

Позначимо p_k надійність k-го елемента, q_k – ймовірність виходу з ладу за час t k-го елемента, P – надійність блока.

Послідовне з'єднання незалежних елементів (рис. 2.1). Такий блок працюватиме безвідмовно лише той час, коли усі елементи працюватимуть безвідмовно. За теоремою множення ймовірностей незалежних подій імовірність безвідмовної роботи такого блока

$$P = p_1 \cdot p_2 \cdot \ldots \cdot p_n .$$

Паралельне з'єднання (рис. 2.2). Такий блок працюватиме безвідмовно, якщо принаймні один елемент не вийде з ладу, тому ймовірність P безвідмовної роботи

$$P = 1 - q_1 \cdot q_2 \cdot \ldots \cdot q_n .$$

Приклад. Прилад складено з двох елементів, з'єднаних послідовно, які працюють незалежно. Ймовірність відмови елементів дорівнює 0,05 та 0,08. Знайти ймовірність відмови приладу.

Розв'язання. Відмова приладу – подія, протилежна до його безвідмовної роботи. Ймовірності безвідмовної роботи елементів

$$p_1 = 1 - 0,05 = 0,95;\ \ p_2 = 1 - 0,08 = 0,92.$$

Ймовірність безвідмовної роботи такого блока (послідовне з'єднання)

$$P = p_1 \cdot p_2 = 0,95 \cdot 0,92 = 0,874.$$

Ймовірність відмови приладу

$$P = 1 - p_1 \cdot p_2 = 1 - 0,874 = 0,126$$

або $P = q_1 + q_2 - q_1 \cdot q_2 = 0,05 + 0,08 - 0,05 \cdot 0,08 = 0,126.$

2.2. Формули повної ймовірності та Байєса

Якщо подія *A* має складну структуру, то пряме визначення її ймовірності може виявитись надто важким завданням (наприклад, за класичним означенням з використанням комбінаторного аналізу). Виникає питання про можливість розглянути подію *A* з іншими подіями, ймовірності яких відомі, і на основі розглянутих подій знайти ймовірність події *A*. Виявляється, що в багатьох випадках це можливо, і відповідь на поставлене запитання дає формула повної ймовірності.

Теорема (формула повної ймовірності). *Нехай випадкова подія A може відбуватись лише сумісно з однією із несумісних між собою подій* $H_1,\ H_2,\ \ldots,H_n$, *що утворюють повну групу* $\left(\bigcup_{i=1}^{n} H_i = \Omega,\ H_i \cap H_j = \varnothing \ \text{çà}\ \ i \neq j\right)$, *які називають гіпотезами,* $P(H_i) > 0,\ \ (i = 1, 2, \ldots, n)$. *Тоді ймовірність події A обчислюють за формулою*

$$P(A) = \sum_{i=1}^{n} P(H_i) P\left(A \mid H_i\right). \tag{2.7}$$

Отже, ймовірність події *A* дорівнює сумі добутків імовірностей гіпотез та умовних імовірностей події *A* за цих гіпотез.

Доведення. Подію *A* можна подати у такому вигляді:

$$A = \Omega \cap A = \left(\bigcup_{i=1}^{n} H_i\right) \cap A = \bigcup_{i=1}^{n} (H_i \cap A).$$

Оскільки події H_i несумісні, то події $H_i \bigcap A$ також несумісні. За теоремою додавання для несумісних подій

$$P(A) = \sum_{i=1}^{n} P(H_i \bigcap A),$$

і за теоремою множення $P(H_i \bigcap A) = P(H_i)P(A \mid H_i)$. Отже,

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

Якщо до випробування відомі апріорні ймовірності гіпотез $P(H_1)$, $P(H_2)$, … , $P(H_n)$, а в результаті випробування настала подія A, і $P(A) > 0$, то, з урахуванням настання цієї події, умовні ймовірності гіпотез обчислюють за **формулою Байєса**:

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{P(A)}, \; i = 1, \; 2, \; \ldots, n. \tag{2.8}$$

Приклад. Два економісти заповнюють документи з власними прогнозами щодо курсу певних коштовних паперів, після чого їх складають у спільну папку. Ймовірність відхилення прогнозованого значення курсу від фактичного понад допустиму похибку для першого економіста дорівнює 0,1; для другого – 0,2. Перший економіст заповнив 40 документів, другий – 60. Під час перевірки навмання взятий із папки документ виявився з суттєвою похибкою. Визначити ймовірність того, що його склав перший економіст.

Розв'язання. Позначимо події H_1, H_2 – документ заповнив відповідно перший, другий економіст, A – навмання взятий із папки документ виявився з суттєвою похибкою. Тоді за умовою $P(H_1) = \frac{40}{100} = 0,4$; $P(H_2) = \frac{60}{100} = 0,6$; $p(A \mid H_1) = 0,1$; $P(A \mid H_2) = 0,2$. За формулою повної ймовірності

$$P(A) = P(H_1)P(A \mid H_1) + P(H_2)P(A \mid H_2) = 0,4 \cdot 0,1 + 0,6 \cdot 0,2 = 0,16.$$

За формулою Байєса отримаємо:

$$P(H_1 \mid A) = \frac{P(H_1)P(A \mid H_1)}{P(A)} = \frac{0,4 \cdot 0,1}{0,16} = 0,25.$$

Отже, ймовірність заповнення документа з суттєвою похибкою першим економістом

$$P\left(H_1 \mid A\right) = 0,25.$$

2.3. Послідовні незалежні випробування. Граничні теореми формули Бернуллі

2.3.1. Послідовні незалежні випробування. Формула Бернуллі

Означення. Якщо n випробувань проводити за однакових умов і ймовірність настання події A в усіх випробуваннях однакова та не залежить від настання або ненастання A в інших випробуваннях, таку послідовність незалежних випробувань називають ***схемою Бернуллі.***

Нехай проводиться скінченна кількість спроб n, в результаті яких може з'явитися подія A з певною ймовірністю p, причому ця подія не залежить від наслідків інших спроб. Такі спроби назвемо *незалежними відносно події* A.

Обчислимо ймовірність того, що в результаті проведення n незалежних спроб подія A настане рівно k разів, якщо в кожній із спроб вона настає із сталою ймовірністю $P(A) = p$ або не настає з ймовірністю $P(\overline{A}) = q$ $(p + q = 1)$. Позначимо шукану ймовірність $P_n(k)$, це означає, що в n спробах подія A з'явиться k разів. Зауважимо, що тут не вимагається, щоб подія A повторилася k разів у певній послідовності. Для розв'язання поставленої задачі за великих значень n і k безпосереднє застосування теорем додавання і множення ймовірностей приводить до громіздких розрахунків, тому зручніше користуватися формулою Бернуллі, виведення якої ми розпочнемо.

Ймовірність того, що подія A в n спробах настане рівно k разів, а в решті $n - k$ спроб настане протилежна подія $\overline{A}$, за теоремою множення ймовірностей незалежних подій дорівнює $p^k \cdot q^{n-k}$. При цьому подія A в n спробах може настати рівно k разів у різних комбінаціях, кількість яких C_n^k. Оскільки всі комбінації подій – події несумісні і не важливо, в якій послідовності настане подія A або подія $\overline{A}$, то, застосовуючи теорему додавання ймовірностей несумісних подій, отримаємо ***формулу Бернуллі:***

$$P_n(k) = C_n^k p^k q^{n-k}. \qquad (2.9)$$

Отже, ймовірність $P_n(k)$ того, що подія A настане k разів у n випробуваннях $(0 \le k \le n)$ обчислюють за ***формулою Бернуллі***:

$$P_n(k) = C_n^k p^k q^{n-k}.$$

Число k_0, за якого ймовірність $P_n(k_0)$ найбільша, називають ***найімовірнішим*** числом настання події A. Його визначають співвідношенням

$$np - q \le k_0 \le np + p$$

або

$$(n+1)p - 1 \le k_0 \le (n+1)p. \qquad (2.10)$$

Число k_0 має бути цілим. Якщо $(n+1)p-$ ціле число, то найбільшого значення ймовірність $P_n(k)$ досягає у двох випадках:

$$k_1 = (n+1)p - 1 \text{ та } k_2 = (n+1)p.$$

Зауваження. Ймовірність настання події A у n випробуваннях схеми Бернуллі *менше ніж m разів* знаходять за формулою

$$P_n(k < m) = P_n(0) + P_n(1) + \cdots + P_n(m-1).$$

Ймовірність настання події A *не менше ніж m разів* знаходять за формулою $P_n(k \ge m) = P_n(m) + P_n(m+1) + \cdots + P_n(n)$ або за такою формулою:

$$P_n(k \ge m) = 1 - \sum_{k=0}^{m-1} P_n(k).$$

Імовірність настання події A *принаймні один раз у n випробуваннях* доцільно обчислювати за формулою $P_n(1 \le m \le n) = 1 - q^n$.

2.3.2. Граничні теореми формули Бернуллі

Якщо проводять випробування за схемою Бернуллі та числа n і k великі, то обчислення ймовірності за формулою Бернуллі викликає певні труднощі. У таких випадках для обчислення цих імовірностей застосовують асимптотичні (наближені) формули, які випливають із локальної та інтегральної теорем Муавра–Лапласа та граничної тереми Пуассона. Назва «гранична» в обох випадках пов'язана з тим, що ці теореми встановлюють поведінку ймовірності $P_n(k)$ або $P_n(k_1 \le k \le k_2)$ за певних умов, до яких обов'язково входить умова $n \to \infty$.

З огляду на це за досить великих значень *n* та за малих p або q замість формули Бернуллі часто використовують наближені асимптотичні формули.

1. Формула Пуассона

Теорема (теорема Пуассона). *Якщо* $n \to \infty$ *і* $p \to 0$ *так, що* $np \to \lambda$, $0 < \lambda < \infty$, *ò î*

$$P_n(k) = C_n^k p^k q^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda} \tag{2.11}$$

для будь-якого сталого $k = 0, 1, 2, \cdots$.

Наслідок. Ймовірність настання події *A* k разів у *n* випробуваннях схеми Бернуллі знаходять за наближеною ***формулою Пуассона***

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \tag{2.12}$$

де $\lambda = np$.

Ця формула дає досить точне наближення за значень *p*, близьких до нуля ($p < 0{,}1$), тобто для подій, що рідко трапляються і для достатньо великих *n* ($npq \le 9$).

2. Локальна теорема Муавра–Лапласа

Означення. *Локальною функцією Лапласа* називають функцію вигляду

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

Основні властивості локальної функції Лапласа:

1. Парність функції: $\varphi(-x) = \varphi(x)$;

2. $\varphi(x)$ визначена для всіх дійсних значень змінної;

3. $\varphi(x) \to 0$ за $x \to \pm\infty$;

4. $\varphi_{\max} = \varphi(0) = \dfrac{1}{\sqrt{2\pi}}$.

Теорема (локальна теорема Муавра–Лапласа). *Якщо ймовірність p настання події A в кожній спробі стала і така, що* $0 < p < 1$*, то ймовірність числа* k *настання події в* n *спробах обчислюють за наближеною рівністю*

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{k-np}{\sqrt{npq}}\right), \text{ де } \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \qquad (2.13)$$

Формула (2.13) дає добре наближення, якщо *n* достатньо велике, *p* та *q* не дуже близькі до нуля, $npq > 9$.

3. Інтегральна теорема Муавра–Лапласа

Означення. *Інтегральною функцією Лапласа* називають функцію вигляду

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

Як правило, табулюється функція $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$, яка пов'язана із функцією $F(x)$ співвідношенням $F(x) = \frac{1}{2} + \Phi(x)$.

Основні властивості функції Лапласа $\Phi(x)$:

1. Непарність функції: $\Phi(-x) = -\Phi(x)$;

2. $\Phi(0) = 0$;

3. $\Phi(x) \approx 0{,}5$ для $x \geq 5$.

Теорема (інтегральна теорема Муавра–Лапласа). *Ймовірність того, що в n незалежних випробуваннях схеми Бернуллі, в яких подія A може відбутись з імовірністю p, подія A відбудеться не менше* k_1 *та не більше* k_2 *разів, наближено рівна*

$$P_n(k_1 \leq k \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right), \ \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt. \qquad (2.14)$$

Формула (2.14) дає добре наближення, якщо *n* достатньо велике, *p* та *q* не дуже близькі до нуля, *npq* > 9. Для всіх значень *x* ≥ 5 можна вважати $\Phi(x) \approx 0{,}5$.

4. Імовірність відхилення відносної частоти від сталої ймовірності в незалежних випробуваннях

Теорема (Я. Бернуллі). *Якщо в n незалежних випробуваннях імовірність p настання події A однакова і подія A настала m разів, то для будь-якого додатного числа* α *справедлива рівність*

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

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

Зауваження. З використанням інтегральної функції Лапласа формулу записують у вигляді

$$P\left(\left|\frac{m}{n}-p\right|\leq\alpha\right)\approx 2\hat{O}\left(\alpha\sqrt{\frac{n}{pq}}\right). \qquad (2.15)$$

Зауваження 1. Послідовність випробувань із різними ймовірностями

У практичній діяльності не завжди у *n* незалежних випробуваннях імовірності настання події *A* рівні, наприклад, вони дорівнюють $p_1,\ p_2,\ \ldots,p_n$, відповідно й імовірності ненастання події *A*: $q_1=1-p_1$, $q_2=1-p_2$, …, $q_n=1-p_n$. У цьому випадку використовують твірну функцію

$$\varphi_n(z)=\prod_{k=1}^{n}\left(q_k+p_k z\right).$$

Потрібна ймовірність $P_n(m)$ рівна коефіцієнту при z^m.

Зауваження 2. Проста течія подій

Означення. Течією подій називають послідовність таких подій, які настають у випадкові моменти часу.

Означення. Течію подій називають ***пуассонівською***, якщо вона:

1. *Стаціонарна* – залежить від кількості k настання події та часу t і не залежить від моменту свого початку.

2. Має *властивість відсутності післядії* – ймовірність настання події не залежить від настання або ненастання події раніше та не впливає на найближче майбутнє.

3. *Ординарна* – ймовірність настання більше однієї події в малий проміжок часу є величина нескінченно мала порівняно з імовірністю настання події один раз у цей проміжок часу.

Означення. Середнє число λ настання події A в одиницю часу називають ***інтенсивністю течії***.

Теорема (математична модель простої течії подій). *Для пуассонівської течії подій імовірність настання A k разів за час t знаходять за формулою*

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

де λ – інтенсивність течії.

Приклад 1. Монету підкинуто 6 разів. Знайти ймовірність того, що герб випаде 4 рази.

Розв'язання. За умовою задачі $n = 6$, $p = q = \frac{1}{2}$ та $k = 4$. За формулою Бернуллі $P_6(4) = C_6^4 \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^2 = \frac{6!}{2! \cdot 4!}\left(\frac{1}{2}\right)^6 = \frac{5 \cdot 6}{2} \cdot \frac{1}{64} = \frac{15}{64}$.

Приклад 2. Ймовірність народження хлопчика дорівнює 0,515, а дівчинки 0,485. У певній сім'ї шестеро дітей. Знайти ймовірність того, що серед них не більше двох дівчаток.

Розв'язання. За умовою $n = 6$, $p = 0{,}485$, $q = 0{,}515$. Шукана ймовірність дорівнює

$$P_6(0) + P_6(1) + P_6(2) = (0{,}515)^6 + C_6^1 \cdot 0{,}485 \cdot (0{,}515)^5 + C_6^2 \cdot (0{,}485)^2 \cdot (0{,}515)^4 =$$
$$= 0{,}018657 + 0{,}10542 + 0{,}2482 = 0{,}3723.$$

Приклад 3. Підручник надруковано тиражем 10 000 примірників. Ймовірність бракованого брошурування підручника дорівнює 0,0001. Знайти ймовірність того, що тираж має 5 бракованих підручників.

Розв'язання. Згідно з умовою задачі $n = 10\,000$ досить велике число, а ймовірність $p = 0{,}0001$ – мала, $\lambda = np = 1$. Застосовуючи формулу Пуассона (2.12), отримаємо

$$P_{10000}(5) \approx \frac{1^5}{5!} e^{-1} = \frac{1}{120 \cdot e} \approx 0{,}0031.$$

Приклад 4. За нового технологічного процесу 80 % виготовленої продукції має найвищу якість. Знайти найбільш імовірну кількість виготовлених виробів найвищої якості серед 250 виготовлених виробів.

Розв'язання. Потрібно використати співвідношення – формулу (2.10)

$$np - q \le m_0 \le np + p,$$

де $n = 250,\ p = 0{,}8,\ q = 0{,}2,$ тоді $199{,}8 \le m_0 \le 200{,}8,$ але m_0 має бути цілим числом, тому $m_0 = 200$.

Приклад 5. Імовірність успіху в кожному випробуванні дорівнює 0,25. Яка ймовірність того, що за 300 випробувань успішними будуть рівно 75 випробувань.

Розв'язання. За локальною формулою Муавра–Лапласа $P_n(k) \approx \approx \frac{1}{\sqrt{npq}}\varphi\left(\frac{k - np}{\sqrt{npq}}\right),$ де $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$ тоді за $n = 300,\quad k = 75,\quad p = 0{,}25$ отримуємо:

$$\frac{k - np}{\sqrt{npq}} = \frac{75 - 300 \cdot 0{,}25}{\sqrt{300 \cdot 0{,}25 \cdot 0{,}75}} = 0,\ P_{300}(75) \approx \frac{1}{7{,}5}\varphi(0) = \frac{0{,}3989}{7{,}5} \approx 0{,}0532.$$

Приклад 6. Імовірність виходу з ладу за час t одного приладу рівна 0,1. Визначити ймовірність того, що за час t із 100 приладів вийде з ладу від 6 до 18 приладів.

Розв'язання. Використовуємо інтегральну формулу Муавра–Лапласа:

$$P_n(k_1 \le k \le k_2) \approx \hat{\text{O}}\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \hat{\text{O}}\left(\frac{k_1 - np}{\sqrt{npq}}\right),\ \hat{\text{O}}(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt.$$

За умовою задачі $n = 100,\ k_2 = 18,\ k_1 = 6,\ p = 0{,}1;\ q = 1 - p = 0{,}9,$ тоді

$$P_{100}\left(6 \le k \le 18\right) \approx \hat{\text{O}}\left(\frac{18 - 100 \cdot 0{,}1}{\sqrt{100 \cdot 0{,}1 \cdot 0{,}9}}\right) - \hat{\text{O}}\left(\frac{6 - 100 \cdot 0{,}1}{\sqrt{100 \cdot 0{,}1 \cdot 0{,}9}}\right) = \hat{\text{O}}\left(\frac{8}{3}\right) - \hat{\text{O}}\left(-\frac{4}{3}\right) \approx$$

$$\approx \hat{\text{O}}\left(2{,}66\right) - \hat{\text{O}}\left(-1{,}33\right) = \hat{\text{O}}\left(2{,}66\right) + \hat{\text{O}}\left(1{,}33\right) = 0{,}49609 + 0{,}40824 = 0{,}90433.$$

Приклад 7. Імовірність того, що деталь нестандартна, дорівнює 0,2. Знайти ймовірність того, що серед випадково відібраних 300 деталей відносна частота натрапляння на нестандартні деталі відхилиться від імовірності 0,2 за абсолютною величиною не більше ніж на 0,01.

Розв'язання. За умовою задачі $n = 300,\ p = 0,2;\ q = 1 - p = 0,8;\ \alpha = 0,01,$ тоді потрібно знайти ймовірність $P\left(\left|\dfrac{m}{300} - 0,2\right| \le 0,01\right)$. За теоремою Бернуллі

$$P\left(\left|\frac{m}{300} - 0,2\right| \le 0,01\right) \approx 2\hat{\text{Ô}}\left(0,01\sqrt{\frac{300}{0,2 \cdot 0,8}}\right) = 2\hat{\text{Ô}}(0,43) = 0,328.$$

Зміст отриманого результату: якщо згідно з умовою задачі взяли достатньо велику кількість спроб, тобто по 300 деталей у кожній, то приблизно для 33 % цих спроб відхилення відносної частоти від сталої ймовірності 0,2 за абсолютною величиною не перевищить 0,01.

Приклад 8. Працівник обслуговує три верстати, що працюють незалежно один від одного. Ймовірність того, що протягом години перший верстат не потребуватиме уваги працівника, дорівнює 0,9, а для другого та третього верстата – 0,8 та 0,85 відповідно. Яка ймовірність того, що протягом години:

а) ні один верстат не потребуватиме уваги працівника;

б) усі три верстати потребуватимуть уваги працівника;

в) принаймні один верстат потребуватиме уваги працівника.

Розв'язання. Цей приклад можна розв'язати, використовуючи теореми множення та додавання ймовірностей. Розв'яжемо задачу з використанням твірної функції, яка в цьому разі матиме вигляд

$$\varphi_n(z) = \prod_{k=1}^{3}\left(q_k + p_k z\right) = (0,1 + 0,9z)(0,2 + 0,8z)(0,15 + 0,85z) =$$

$$= 0,003 + 0,056z + 0,329z^2 + 0,612z^3.$$

Тоді коефіцієнт при z^k $(k = 0,\ 1,\ 2,\ 3)$ рівний імовірності того, що протягом години уваги працівника не потребують k верстатів. Тож відповіді на запитання такі:

а) ймовірність того, що усі три верстати не потребують уваги працівника, дорівнює коефіцієнту при z^3, тобто $P_3(3) = 0,612$;

б) $P_3(0) = 0,003$;

в) $P_3(1 \le m \le 3) = 1 - P_3(0) = 1 - 0,003 = 0,997.$

Приклад 9. Середня кількість замовлень, що надходять до комбінату побутового обслуговування кожної години, дорівнює 3. Знайти ймовірність того, що протягом двох годин надійде 5 замовлень.

Розв'язання. Маємо просту течію подій з інтенсивністю $\lambda = 3$. За формулою $P_t(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ отримуємо $P_2(5) = \frac{(3 \cdot 2)^5}{5!} e^{-3 \cdot 2} = \frac{6^5}{120} e^{-6} \approx 0{,}16.$

Приклад 10. Скільки разів потрібно підкинути монету, щоб з імовірністю 0,6 можна було сподіватись, що відхилення відносної частоти випадання герба від імовірності $p = 0{,}5$ буде за модулем не більше 0,01?

Розв'язання. За умовою $\alpha = 0{,}01$; $p = 0{,}5$, отже $q = 0{,}5$, тому

$$P\left(\left|\frac{m}{n} - 0{,}5\right| \le 0{,}01\right) = 0{,}6$$

і за формулою (15): $2\hat{O}\left(\alpha\sqrt{\frac{n}{pq}}\right) = 2\hat{O}\left(0{,}01\sqrt{\frac{n}{0{,}5 \cdot 0{,}5}}\right) = 0{,}6,$ тобто $\hat{O}\left(0{,}02\sqrt{n}\right) = \ = 0{,}3,$ звідси $0{,}02\sqrt{n} = 0{,}84$, тобто $n = 1764$.

Частина II

Випадкові величини і процеси. Основні розподіли випадкових величин

Розділ 3. Одновимірні випадкові величини

У теорії ймовірностей поряд з поняттям випадкова подія і ймовірність одним з основних є поняття випадкова величина. Наприклад, час безвідмовної роботи деякого приладу, кількість випадань герба за багаторазового підкидання монети тощо.

Досі ми розглядали випадкові події та дії над ними і маємо базу для поглибленого вивчення випадкових експериментів. Поштовхом у цьому напрямі буде природне або штучне введення до розгляду випадкової величини, оскільки в теорії ймовірностей і математичній статистиці часто доцільно пов'язувати різноманітні випадкові події з дійсними числами і проводити відповідні обчислення, тому виникає нове поняття випадкової величини, про яке йтиметься далі.

3.1. Види випадкових величин та способи їх задання. Дискретні та неперервні випадкові величини

***Означення*. Випадковою величиною** називають таку величину, яка внаслідок випробування може набути лише одного числового значення, яке зумовлене результатом експерименту.

Отже, ***випадковою величиною***, пов'язаною з певним дослідом, називають величину, яка під час кожного проведенням досліду може набувати того чи того числового значення, залежно від випадку.

Між випадковими подіями і випадковими величинами є тісний зв'язок. Випадкова подія – це якісна характеристика випадкового результату досліду, а випадкова величина – його кількісна характеристика. Випадкові величини за певними властивостями поділяються на ***дискретні*** та ***неперервні***.

***Означення*. Дискретною випадковою величиною** (ДВВ) називають таку величину, яка внаслідок випробування може набути відокремлених, ізольованих одне від одного числових значень з відповідними ймовірностями.

Іншими словами, вона має таку властивість, що кожне з її можливих значень має окіл, який вже не містить жодного з інших значень цієї самої величини. Всі можливі значення дискретної випадкової величини можуть бути перенумеровані:

$$x_1, x_2, ..., x_n, ... \ .$$

Означення. **Неперервною випадковою величиною** (НВВ) називають величину, яка може набувати будь-якого числового значення з певного обмеженого інтервалу (a, b) або необмеженого інтервалу $(-\infty, +\infty)$. Наприклад, випадкова величина X – час безвідмовної роботи приладу, неперервна, оскільки її можливе значення $t > 0$.

Означення. Співвідношення, яке встановлює зв'язок між можливими значеннями випадкової величини і ймовірностями, з якими приймають ці значення, називають **законом розподілу ймовірностей випадкової величини**.

Для дискретної випадкової величини X закон розподілу може бути заданий таблично або графічно.

У першому випадку закон розподілу називають ***рядом розподілу ймовірностей*** випадкової величини X:

X	x_1	x_2	$\ldots$	x_n	$\ldots$
P	p_1	p_2	$\ldots$	p_n	$\ldots$

У першому рядку таблиці записують усі можливі значення випадкової величини, а в другому − відповідні їм імовірності. Оскільки події $\{X = x_1\}$, $\{X = x_2\}$, ..., $\{X = x_n\}$ становлять повну групу несумісних подій, то за теоремою додавання ймовірностей маємо:

$$\sum_{k=1}^{\infty} p_k = 1,$$

тобто сума ймовірностей усіх можливих значень випадкової величини дорівнює одиниці.

Графічна форма закону розподілу називається ***багатокутником розподілу***: по осі абсцис відкладаємо можливі значення x_k випадкової величини X, а по осі ординат – імовірності p_k цих значень; точки (x_k, p_k) послідовно з'єднуємо відрізками прямих.

Закон розподілу неперервної випадкової величини може бути заданий графічно або аналітично $p_k = f(x_k)$ (за допомогою формули). Табличне задання неможливе, оскільки ймовірність отримати будь-яке значення неперервної величини дорівнює нулю, що пов'язано не з неможливістю самої події (потрапляння в певну точку на числовій осі), а з нескінченно великою кількістю можливих випадків.

З огляду на це для неперервних випадкових величин (як, зрештою, і для дискретних) визначають імовірність потрапляння в деякий інтервал числової осі.

Імовірність потрапляння випадкової величини X в інтервал $[a, b]$ визначають як імовірність події $P(a \le X < b)$.

Для кількісного оцінювання закону розподілу випадкової величини (дискретної або неперервної) задають ***функцію розподілу ймовірностей випадкової величини,*** котру визначають як імовірність того, що випадкова величина X набуде значення, меншого від певного фіксованого числа x і позначають $F(x) = P(X < x)$ або $F(x) = P(-\infty < X < x)$.

Функцію розподілу $F(x)$ називають ***інтегральною функцією*** розподілу ймовірностей випадкової величини.

Знаючи функцію розподілу $F(x)$, можна обчислити ймовірність потрапляння випадкової величини у деякий інтервал $[a, b)$:

$$P(a < X < b) = F(b) - F(a). \tag{3.1}$$

Дійсно, випадкова подія $(X < b)$ – об'єднання двох несумісних подій $(X < a)$ і $(a \le X < b)$.

Отже, за теоремою додавання ймовірностей несумісних подій маємо:

$$P(X < b) = P(X < a) + P(a \le X < b),$$

звідки

$$P(a \le X < b) = P(X < b) - P(X < a),$$

або, враховуючи позначення, $P(a \le X < b) = F(b) - F(a)$.

Зауваження. Для дискретної випадкової величини X

$$F(x) = \sum_{x_k < x} p_k . \tag{3.2}$$

Властивості функцій розподілу

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

2. Функція розподілу неспадна: якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$.

3. Функція розподілу неперервна зліва: $\lim\limits_{x \to x_1 - 0} F(x) = F(x_1)$.

4. Імовірність потрапляння випадкової величини X у проміжок $[a,\ b)$:

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

5. $P(X \ge x) = 1 - F(x)$.

6. $F(-\infty) = 0, \quad F(+\infty) = 1$.

Зауваження. Неперервна випадкова величина X, що набуває значення в інтервалі $(a,\ b)$, має незліченну кількість можливих значень, тому набуття X певних значень, наприклад, $X = a$ або $X = b$, – події з нульовою ймовірністю. Це означає, що $P(X = a) = 0$ та $P(X = b) = 0$. З огляду на це справедливі

$$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b).$$

Тоді, згідно з зауваженням, для такої *неперервної випадкової величини* $F(x) = 0$ за $x \le a$; $F(x) = 1$ за $x \ge b$.

Графік її функції розподілу наведено на рис. 3.1.

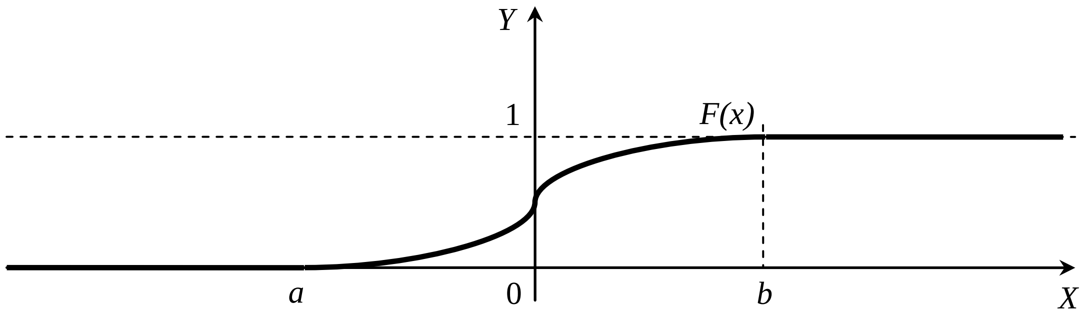


Рис. 3.1. Неперервна функція розподілу

Приклад 1. За трьох підкидань монети випадкова величина X − кількість випадань герба – може набувати значення $x_1 = 0,\ x_2 = 1,\ x_3 = 2,\ x_4 = 3$ із відповідними ймовірностями, які обчислимо за формулою Бернуллі:

$$p_1 = P_3(0) = C_3^0 \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^3 = \frac{1}{8},\ \ p_2 = P_3(1) = C_3^1 \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)^2 = \frac{3}{8};$$

$$p_3 = P_3(2) = C_3^2\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right) = \frac{3}{8},\ \ p_4 = P_3(3) = C_3^3\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^0 = \frac{1}{8}.$$

Ряд розподілу має такий вигляд:

X	0	1	2	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Побудуємо функцію розподілу випадкової величини X, заданої цим рядом розподілу:

якщо $x \le 0,$ $\quad F(x) = P(X < 0) = 0;$

якщо $0 < x \le 1,$ $\quad F(x) = P(X < 1) = P(X = 0) = \frac{1}{8};$

якщо $1 < x \le 2,$ $\quad F(x) = P(X < 2) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2};$

якщо $2 < x \le 3,$

$$F(x) = P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8};$$

якщо $x > 3,$

$$F(x) = P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1.$$

Приклад 2. Нехай функцію розподілу деякої неперервної випадкової величини X задано у вигляді

$$F(x) = \begin{cases} 0, & x \le 0; \\ a(1 - \cos x), & 0 < x \le \pi; \\ 1, & x > \pi. \end{cases}$$

Визначити значення коефіцієнта a.

Розв'язання. Значення коефіцієнта обчислимо, користуючись властивостями функції розподілу.

Оскільки функція неперервна зліва, то за $x = \pi$ маємо $a(1 - \cos\pi) = 1,$ звідки $a = \frac{1}{2}$. З другого боку, $F(\pi) - F(0) = 1,$ тобто $a(1 - \cos\pi) - a(1 - \cos 0) = 1$ і, відповідно, $2a = 1,$ $a = \frac{1}{2}.$

Закон розподілу ймовірностей неперервних випадкових величин може бути заданий також і ***щільністю розподілу.***

Нехай неперервна випадкова величина X задана неперервною і диференційовною функцією розподілу $F(x)$. Імовірність потрапляння цієї випадкової величини в деякий інтервал $(x,\ x+\Delta x)$ знайдемо на підставі співвідношення (3.3):

$$P\left(x < X < x+\Delta x\right) = F(x+\Delta x) - F(x),$$

тобто як приріст функції розподілу на цьому інтервалі.

Відношення $\dfrac{P\left(x < X < x+\Delta x\right)}{\Delta x}$ виражає середню ймовірність, яка припадає на одиницю довжини інтервалу.

Перейшовши до границі за $\Delta x \to 0$, отримаємо:

$$\lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = F'(x).$$

Означення. **Диференціальною функцією розподілу** або **щільністю розподілу ймовірностей** неперервної випадкової величини називають

$$f(x) = F'(x). \qquad (3.4)$$

Назва *щільність ймовірностей* випливає з рівності

$$f(x) = \lim_{\Delta x \to 0} \frac{P\left(X < x+\Delta x\right) - P\left(X < x\right)}{\Delta x}.$$

Властивості диференціальних функцій розподілу

1. $f(x) \geq 0,\ x \in R.$

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

3. $P\left(a \leq X < b\right) = \int\limits_{a}^{b} f(x)dx.$

4. Якщо всі можливі значення випадкової величини належать інтервалу $\left(a,\ b\right)$, то $\int\limits_{a}^{b} f(x)dx = 1.$

За відомою диференціальною функцією розподілу $f(x)$ знаходять інтегральну функцію розподілу $F(x) = \int\limits_{-\infty}^{x} f(t)dt.$

Геометричне тлумачення щільності розподілу випливає із формули $P\left(a \le X < b\right) = \int_{a}^{b} f(x)dx$: імовірність попадання випадкової величини X на проміжок $[a,\ b)$ обчислюють як площу криволінійної трапеції, обмеженої зверху графіком функції $y = f(x)$, знизу – відрізком $\left[a,\ b\right]$ осі абсцис, зліва і справа – відрізками прямих $x = a,\ x = b$.

3.2. Числові характеристики випадкових величин

Під час розв'язування практичних задач досить важко, а іноді й неможливо визначити функцію розподілу випадкової величини X, тому потрібно вміти характеризувати її розподіл через параметри, найважливіші з яких – математичне сподівання та дисперсія.

3.2.1. Математичне сподівання

Означення. **Математичним сподіванням** *дискретної випадкової величини* X називають суму добутків усіх можливих значень x_k випадкової величини X відповідних імовірностей p_k

$$M(X) = \sum_{k} x_k p_k. \tag{3.5}$$

Якщо при цьому множина можливих значень X нескінченна, то накладається умова абсолютної збіжності ряду.

Математичне сподівання називають центром розсіювання випадкової величини і воно відповідає середньому значенню випадкової величини.

Означення. **Математичним сподіванням** *неперервної випадкової величини* X, яка задана щільністю розподілу $f(x)$, називають

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx, \tag{3.6}$$

якщо цей інтеграл абсолютно збіжний. Якщо можливі значення неперервної випадкової величини X належать проміжку $\left[a,\ b\right]$, то

$$M(X) = \int_{a}^{b} xf(x)dx. \tag{3.7}$$

Означення, потрібні для розуміння властивостей числових характеристик випадкових величин.

1. *Добутком сталої величини* C та дискретної випадкової величини X називають дискретну випадкову величину CX, можливі значення якої рівні добутку сталої C та можливих значень X, а ймовірності цих значень дорівнюють імовірностям відповідних значень X.

2. Дві випадкові величини називають *незалежними*, якщо закон розподілу однієї з них не залежить від того, яких можливих значень набуває інша величина. Кілька випадкових величин називають взаємно незалежними, якщо закони розподілу будь-якої кількості цих величин не залежать від того, яких можливих значень набули інші величини.

3. *Добутком незалежних дискретних випадкових величин* X та Y називають випадкову величину XY, можливі значення якої дорівнюють добуткам кожного можливого значення X та кожного можливого значення Y, а ймовірності цих значень дорівнюють добуткам імовірностей можливих значень співмножників. Якщо деякі добутки $x_i y_i$ рівні між собою, то їх імовірності додаються.

4. *Сумою дискретних випадкових величин* X та Y називають випадкову величину $X + Y$, можливі значення якої дорівнюють сумам кожного можливого значення X та кожного можливого значення Y, а імовірності цих значень для незалежних величин X і Y дорівнюють добуткам імовірностей доданків; для залежних величин – добуткам імовірностей одного доданку та умовних імовірностей другого. Якщо деякі суми $x_i + y_j$ рівні між собою, то їх імовірності додаються.

Властивості математичного сподівання

1. Математичне сподівання сталої величини – стала $M(C) = C$.

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

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

Наслідки : *а*) $M(X_1 + X_2 + \ldots + X_n) = M(X_1) + M(X_2) + \ldots + M(X_n)$;

б) $M(X - Y) = M(X) - M(Y)$.

4. Математичне сподівання добутку двох незалежних випадкових величин $M(XY) = M(X)M(Y)$.

3.2.2. Дисперсія. Середнє квадратичне відхилення

Означення. **Дисперсією** випадкової величини називають математичне сподівання квадрата різниці випадкової величини і її математичного сподівання:

$$D(X) = M(X - M(X))^2. \tag{3.8}$$

Теорема. Дисперсія випадкової величини дорівнює різниці між математичним сподіванням квадрата цієї величини і квадратом її математичного сподівання:

$$D(X) = M(X^2) - \left(M(X)\right)^2. \tag{3.9}$$

Доведення. Згідно з означенням дисперсії та властивостями математичного сподівання маємо:

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

$$= M(X^2) - 2M(X)M(X) + (M(X))^2 = M(X^2) - 2(M(X))^2 + (M(X))^2 =$$

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

За теоремою зручно для обчислення дисперсії використовувати формули:

1) для дискретної випадкової величини

$$D(X) = \sum_k x_k^2 p_k - (M(X))^2; \tag{3.10}$$

2) для неперервної випадкової величини

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - (M(X))^2; \tag{3.11}$$

якщо її можливі значення належать відрізку $[a, b]$, тоді

$$D(X) = \int_a^b x^2 f(x)dx - (M(X))^2. \tag{3.12}$$

Розмірність дисперсії не збігається з розмірністю випадкової величини. Для того, щоб розмірності були однаковими, вводять поняття ***середнього квадратичного відхилення (стандартного відхилення)***.

***Означення*. Середнім квадратичним відхиленням** випадкової величини називають

$$\sigma_{\tilde{o}} = \sigma(\tilde{O}) = \sqrt{D(X)}. \tag{3.13}$$

Дисперсія і середнє квадратичне відхилення – міра розсіювання значень випадкової величини навколо її математичного сподівання. В економічних задачах дисперсію часто беруть за міру ризику, оскільки вона характеризує мінливість, розсіювання випадкової величини прибутку, зумовленої, наприклад, випадковими коливаннями цін на енергоносії.

Середнє квадратичне відхилення σ обчислюють, на відміну від дисперсії, у тій самій розмірності, що й випадкову величину. Саме це є причиною його широкого застосування для характеристики відхилень та ймовірної оцінки поведінки випадкової величини. Зокрема, середнє квадратичне відхилення має надзвичайно важливе значення для критеріальної характеристики так званого *принципу практичної впевненості*. ***Принцип практичної впевненості*** можна сформулювати так: *якщо ймовірність деякої події в певному досліді досить мала, то можна бути практично впевненим у тому, що за одноразового проведення досліду подія A не настане*. Що стосується підприємницької діяльності, *принцип практичної впевненості*, мабуть, можна сформулювати так: *можна бути майже впевненим, що запланований захід, дія чи прийняте рішення будуть здійснені, якщо ймовірність їх нездійснення та ризик досить малі.*

Властивості дисперсії

1. $D(C)=0,\ C$ – стала величина.

2. $D(CX)=C^2D(X)$.

3. Дисперсія суми незалежних випадкових величин дорівнює сумі їх дисперсій:

$$D(X+Y)=D(X)+D(Y).$$

Наслідок. Дисперсія різниці двох незалежних випадкових величин дорівнює сумі їх дисперсій:

$$D(X-Y)=D(X)+D(Y).$$

Приклад 1. За законом розподілу дискретної випадкової величини

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

знайти $D(X),\ \sigma(X)$.

Розв'язання. Знаходимо $M(X) = 1 \cdot 0{,}4 + 2 \cdot 0{,}5 + 3 \cdot 0{,}1 = 1{,}7$. Також $M(X^2) = 1 \cdot 0{,}4 + 4 \cdot 0{,}5 + 9 \cdot 0.1 = 3{,}3$. Тоді за формулою (3.9)

$$D(X) = M(X^2) - (M(X))^2 = 3{,}3 - 1{,}7^2 = 0{,}41,$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{0{,}41} = 0{,}64.$$

Приклад 2. Є можливість вибрати спосіб виробництва і реалізації двох наборів товарів широкого вжитку. За даними відділу маркетингу, яким були проведені дослідження ринку, можливий прибуток від виробництва і реалізації *X* і *Y* наведено в таблицях (*X*, *Y* – прибуток у грошових одиницях):

X	1000	1500	2000
P	0,5	0,3	0,2

Y	1000	1500	1750
P	0,4	0,4	0,2

Потрібно оцінити ступінь ризику і запропонувати рішення щодо випуску і реалізації одного із наборів товарів.

Розв'язання. Знаходимо математичне сподівання можливого прибутку для кожного варіанта:

$$M(X) = 1000 \cdot 0{,}5 + 1500 \cdot 0{,}3 + 2000 \cdot 0{,}2 = 1350;$$

$$M(Y) = 1000 \cdot 0{,}4 + 1500 \cdot 0{,}4 + 1750 \cdot 0{,}2 = 1350.$$

Обидва варіанти мають однакове математичне сподівання можливого прибутку. Оцінюємо ступінь ризику кожного варіанта:

$$D(X) = (1000 - 1350)^2 \cdot 0{,}5 + (1500 - 1350)^2 \cdot 0{,}3 + (2000 - 1350)^2 \cdot 0{,}2 =$$

$$= 61250 + 6750 + 84500 = 152500, \sigma(X) = 390{,}5.$$

$$D(Y) = (1000 - 1350)^2 \cdot 0{,}4 + (1500 - 1350)^2 \cdot 0{,}4 + (1750 - 1350)^2 \cdot 0{,}2 =$$

$$= 49000 + 9000 + 32000 = 90000, \ \sigma(Y) = 300.$$

Ступінь ризику, пов'язаний з виробництвом і реалізацією набору X більший, ніж набору Y. Варіант Y менш ризикований.

Приклад 3. Знайти числові характеристики випадкової величини X, яка задана функцією розподілу $F(x)=\begin{cases}0,\ x\le 0;\\ \dfrac{x^2}{25},\ 0<x\le 5;\\ 1,\ x>5.\end{cases}$

Розв'язання. Спочатку знайдемо диференціальну функцію розподілу, тобто щільність розподілу ймовірності за формулою $f(x)=F'(x)$:

$$f(x)=\begin{cases}\dfrac{2}{25}x,\ 0<x\le 5;\\ 0,\ x\notin(0,\ 5].\end{cases}$$

За формулою $M(X)=\int\limits_a^b xf(x)dx$ знаходимо математичне сподівання:

$$M(X)=\int\limits_0^5 x\cdot\frac{2}{25}xdx=\frac{2}{25}\cdot\frac{x^3}{3}\bigg|_0^5=\frac{2}{75}\left(5^3-0\right)=\frac{10}{3}.$$

Для дисперсії використовуємо формулу $D(X)=\int\limits_a^b x^2f(x)dx-(M(X))^2$:

$$D(X)=\int\limits_0^5 x^2\cdot\frac{2}{25}xdx-\left(\frac{10}{3}\right)^2=\frac{2}{25}\cdot\frac{x^4}{4}\bigg|_0^5-\frac{100}{9}=\frac{25}{18}.$$

Середнє квадратичне відхилення $\sigma(X)=\sqrt{D(X)}=\sqrt{\dfrac{25}{18}}=\dfrac{5}{3\sqrt{2}}=\dfrac{5\sqrt{2}}{6}\approx 1{,}17.$

3.2.3. Однаково розподілені взаємно незалежні випадкові величини

Нехай $X_1,\ X_2,\ \ldots,\ X_n$ – n взаємно незалежних однаково розподілених випадкових величин, які мають однакові характеристики. Позначимо через a і D відповідно їх математичне сподівання і дисперсію, $\sigma=\sqrt{D}$. Нехай

$$\overline{X}=\frac{X_1+X_2+\ldots+X_n}{n}$$

– це середнє арифметичне цих величин.

Тоді числові характеристики $\overline{X}$ такі:

1. $M(\overline{X}) = a.$

2. $D(\overline{X}) = \dfrac{D}{n}.$

3. $\sigma(\overline{X}) = \dfrac{\sigma}{\sqrt{n}}.$

З огляду на ці результати середнє арифметичне досить великої кількості взаємно незалежних випадкових величин має менше розсіювання, ніж кожна окрема величина. Тому середнє арифметичне кількох вимірювань беруть за наближене значення вимірюваної величини.

3.2.4. Початкові та центральні моменти, інші числові характеристики

***Означення*. Початковим моментом порядку** k (k **–** натуральне число) випадкової величини X називають математичне сподівання величини X^k:

$$\nu_k = M(X^k). \tag{3.14}$$

Зокрема, $\nu_1 = M(X)$, $\nu_2 = M(X^2)$, тоді формулу (3.9) для обчислення дисперсії можна записати так: $D(X) = \nu_2 - \nu_1^2$.

***Означення*. Центральним моментом порядку** k *(*k **–** натуральне число**)** випадкової величини X називають математичне сподівання величини $(X - M(X))^k$:

$$\mu_k = M(X - M(X))^k. \tag{3.15}$$

Зокрема, $\mu_1 = M(X - M(X)) = 0$, $\mu_2 = M(X - M(X))^2 = D(X)$, відповідно порівнюючи формули дисперсії маємо:

$$\mu_2 = \nu_2 - \nu_1^2.$$

Справедливі формули:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3;$$

$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

Для симетричного розподілу кожен центральний момент непарного порядку дорівнює нулю, тому кожен відмінний від нуля центральний момент непарного порядку характеризує ступінь асиметрії розподілу.

Означення. Величину, яка є безрозмірною і має вигляд

$$A_s = \frac{\mu_3}{\sigma^3} \tag{3.16}$$

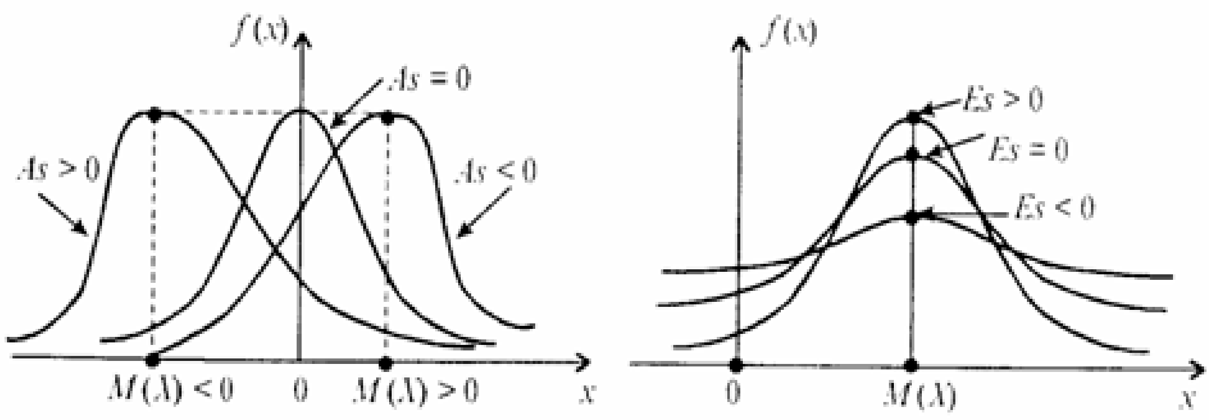
називають **асиметрією розподілу** (характеристика зсуву – вершини графіка щільності розподілу).

Означення. **Ексцесом** називають нормований центральний момент четвертого порядку

$$E_s = \frac{\mu_4}{\sigma^4} - 3, \tag{3.17}$$

що є характеристикою гостроверхості (плосковерхості) вершини графіка щільності розподілу.

Для наочності графіки щільності розподілу ймовірностей неперервної випадкової величини $f(x)$ за різних значень асиметрії A_s та ексцесу E_s наведено на рис. 3.2.



а *б*

Рис. 3.2. Графіки $f(x)$ за різних значень: *а* – асиметрії; *б* – ексцесу

Означення. **Модою** $M_o(X)$ розподілу неперервної випадкової величини X із щільністю розподілу $f(x)$ називають кожне значення x, за якого $f(x)$ має максимум. Розподіли, які мають одну моду, називають *унімодальними*.

***Означення*. Медіаною** $M_e(X)$ розподілу неперервної випадкової величини X називають можливе значення x, за якого пряма $x = M_e(X)$ ділить криволінійну трапецію, обмежену кривою розподілу і віссю Ox, на частини рівної площі.

Приклад 4. Випадкова величина X задана щільністю розподілу

$$f(x) = \begin{cases} -\frac{3}{4}x^2 + 6x - \frac{45}{4}, \ x \in (3,\ 5); \\ 0, \ x \notin (3,\ 5). \end{cases}$$

Знайти моду, математичне сподівання і медіану випадкової величини X.

Розв'язання. Запишемо щільність розподілу ймовірностей у вигляді $f(x) = -\frac{3}{4}(x-4)^2 + \frac{3}{4}$. Звідси видно, що за $x = 4$ щільність розподілу $f(x)$ досягає максимуму, отже і $M_o(X) = 4$. Крива симетрична відносно прямої $x = 4$, тому $M(X) = 4$, $M_e(X) = 4$.

Розділ 4. Основні закони розподілу дискретних та неперервних випадкових величин

У попередньому розділі ми загалом розглянули закони розподілу дискретних та неперервних випадкових величин. Деякі з розподілів більш поширені у застосуванні, тому зупинимось на них більш детально.

4.1. Основні закони розподілу дискретних випадкових величин, їх основні числові характеристики

4.1.1. Біномний розподіл

Проводять n однакових незалежних випробувань, у кожному з яких може відбутись подія A з імовірністю p $(0 < p < 1)$. Випадкова величина X – кількість разів настання події A. Ряд розподілу

X	0	1	2	$\ldots$	k	$\ldots$	n
P	$p_n(0)$	$p_n(1)$	$p_n(2)$	$\ldots$	$p_n(k)$	$\ldots$	$p_n(n)$

де $p_n(k) = C_n^k p^k q^{n-k}$, $q = 1 - p$, $(k = 0,\ 1,\ 2,\ \ldots, n)$.

Для біномного розподілу

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

Доведення. Нехай випадкова величина X_k набуває двох значень: 1, якщо подія A настала в k-му випробуванні, і 0, якщо подія A не настала в k-му випробуванні $(k = 1, 2, \ldots, n)$. Ряд розподілу для випадкової величини X_k має такий вигляд:

X_k	0	1
P	$1 - p$	p

Тоді $X = X_1 + X_2 + \ldots + X_n$. Визначимо числові характеристики для випадкової величини X_k: $M(X_k) = 0 \cdot (1-p) + 1 \cdot p = p$, $D(X_k) = M(X_k^2) - (M(X_k))^2 = 0^2 \cdot (1-p) + 1^2 \cdot p - p^2 = p - p^2 = p(1-p) = pq.$

Тоді $M(X) = M(X_1) + M(X_2) + \ldots + M(X_n) = np;$

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

Отже,

$$M(X) = np, \; D(X) = npq. \tag{4.1}$$

Приклад 1. В партії 10 % нестандартних деталей. Навмання вибрали 4 деталі. Знайти закон розподілу дискретної випадкової величини X – кількості нестандартних деталей серед чотирьох відібраних та обчислити $M(X)$, $D(X)$, $\sigma(X)$.

Розв'язання. Випадкова величина X може набувати значення $x_1 = 0,\; x_2 = 1,\; x_3 = 2,\; x_4 = 3,\; x_5 = 4.$ Очевидно, $n = 4,\; p = 0{,}1;\; q = 1 - p = 0{,}9.$ За формулою Бернуллі обчислимо ймовірності значень випадкової величини:

$$p_4(0) = q^4 = 0{,}9^4 = 0{,}6561;$$

$$p_4(1) = C_4^1 pq^3 = 4 \cdot 0{,}1 \cdot 0{,}9^3 = 0{,}2916;$$

$$p_4(2) = C_4^2 p^2 q^2 = \frac{4 \cdot 3}{1 \cdot 2} \cdot 0{,}1^2 \cdot 0{,}9^2 = 6 \cdot 0{,}01 \cdot 0{,}81 = 0{,}0486;$$

$$p_4(3) = C_4^3 p^3 q = 4 \cdot 0{,}1^3 \cdot 0{,}9 = 0{,}0036;$$

$$p_4(4) = p^4 = 0{,}1^4 = 0{,}0001.$$

Контроль обчислень: $0{,}6561 + 0{,}2916 + 0{,}0486 + 0{,}0036 + 0{,}0001 = 1.$

Ряд розподілу для випадкової величини X має такий вигляд:

X	0	1	2	3	4
P	0,6561	0,2916	0,0486	0,0036	0,0001

За формулами (4.1) $M(X) = np = 4 \cdot 0{,}1 = 0{,}4;$ $D(X) = npq = 4 \cdot 0{,}1 \cdot 0{,}9 = 0{,}36,$ відповідно $\sigma(X) = \sqrt{D(X)} = \sqrt{0{,}36} = 0{,}6.$

4.1.2. Розподіл Пуассона

Якщо в схемі незалежних повторних випробувань n досить велике, а p або $1 - p$ прямує до нуля, тоді біномний закон розподілу апроксимує *розподіл*

Пуассона, параметр якого $\lambda = np,$ причому $p \le 0{,}1$ або $p \ge 0{,}9$:

X	0	1	2	$\ldots$	m	$\ldots$
P	p_0	p_1	p_2	$\ldots$	p_m	$\ldots$

Відповідні ймовірності обчислюють за формулою Пуассона:

$$P(X = m) = \frac{\lambda^m}{m!}e^{-\lambda},\ \lambda > 0\ \ (m = 0,\ 1,\ 2,\ \ldots),$$

причому $\sum\limits_{m=0}^{\infty} \frac{\lambda^m}{m!}e^{-\lambda} = e^{-\lambda}\sum\limits_{m=0}^{\infty}\frac{\lambda^m}{m!} = e^{-\lambda}\cdot e^{\lambda} = 1.$

Знаходимо математичне сподівання та дисперсію:

$$M(X) = \sum_{m=0}^{\infty} m\frac{\lambda^m e^{-\lambda}}{m!} = \sum_{m=1}^{\infty} m\frac{\lambda^m e^{-\lambda}}{m(m-1)!} = \lambda e^{-\lambda}\sum_{m=1}^{\infty}\frac{\lambda^{m-1}}{(m-1)!} = \lambda e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!} = \left|\begin{matrix}\text{çàì ³í à}\\ n = m-1\end{matrix}\right| =$$

$$= \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda, \text{ тобто } M(X) = \lambda.$$

$$M(X^2) = \sum_{m=0}^{\infty} m^2\frac{\lambda^m e^{-\lambda}}{m!} = \sum_{m=1}^{\infty} m\frac{m\lambda^m e^{-\lambda}}{m(m-1)!} = \lambda\sum_{m=1}^{\infty} m\frac{\lambda^{m-1}e^{-\lambda}}{(m-1)!} =$$

$$= \lambda\sum_{m=1}^{\infty}\big((m-1)+1\big)\frac{\lambda^{m-1}e^{-\lambda}}{(m-1)!} = \lambda\left[\sum_{m=1}^{\infty}(m-1)\frac{\lambda^{m-1}e^{-\lambda}}{(m-1)!} + \sum_{m=1}^{\infty}\frac{\lambda^{m-1}e^{-\lambda}}{(m-1)!}\right] = \left|\begin{matrix}\text{заміна}\\ n = m-1\end{matrix}\right| =$$

$$= \lambda\left[\sum_{n=0}^{\infty} n\frac{\lambda^n e^{-\lambda}}{n!} + \sum_{n=0}^{\infty}\frac{\lambda^n e^{-\lambda}}{n!}\right] = \lambda(\lambda + e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}) = \lambda(\lambda + e^{-\lambda}e^{\lambda}) = \lambda(\lambda+1) = \lambda^2 + \lambda,$$

$$D(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Отже,

$$M(X) = \lambda,\ D(X) = \lambda,\ \sigma(X) = \sqrt{\lambda}. \qquad (4.2)$$

Розподіл Пуассона використовують у задачах статистичного контролю якості, в теорії надійності, теорії масового обслуговування, для прогнозування кількості вимог на виплату страхових компенсацій за рік, кількості дефектів в однакових виробах.

4.1.3. Геометричний розподіл

Проводять незалежні випробування, в кожному з яких може відбутись подія A з імовірністю p $(0 < p < 1)$. Дискретна випадкова величина X – кількість проведених випробувань до першого настання події A.

Ряд імовірностей цього розподілу – нескінченно спадна геометрична прогресія зі знаменником $q = 1 - p$, сума всіх членів якої дорівнює одиниці.

X	1	2	3	$\dots$	n	$\dots$
P	p	pq	pq^2	$\dots$	pq^{n-1}	$\dots$

Для даного розподілу

$$M(X) = \frac{1}{p},$$

$$D(X) = \frac{q}{p^2}.$$

Обчислимо математичне сподівання:

$$M(X) = 1p + 2pq + 3pq^2 + \dots + npq^{n-1} + \dots + = p(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots) =$$

$$= p(q + q^2 + q^3 + \dots + q^n + \dots)' = p\left(\frac{q}{1-q}\right)' = p\frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}.$$

Отже, $M(X) = \dfrac{1}{p}$.

Для $D(X) = M(X^2) - (M(X))^2 = M(X^2) - \dfrac{1}{p^2}$.

X^2	1	4	9	$\dots$	n^2	$\dots$
P	p	pq	pq^2	$\dots$	pq^{n-1}	$\dots$

$$M(X^2) = 1p + 4pq + 9pq^2 + \dots + n^2pq^{n-1} + \dots = p(1 + 4q + 9q^2 + \dots + n^2q^{n-1} + \dots).$$

Для обчислення суми в дужках розглянемо суму

$$1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots = \frac{1}{(1-q)^2},$$

звідси

$$q(1 + 2q + 3q^2 + \dots + nq^{n-1} + \dots) = \frac{q}{(1-q)^2}$$

або

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

Диференціюючи ліву та праву частини цієї рівності, отримуємо:

$$1+4q+9q^2+\ldots+n^2q^{n-1}+\ldots=\frac{1+q}{(1-q)^3}.$$

Отже, $M(X^2)=p\frac{1+q}{(1-q)^3}=\frac{1+q}{p^2}$,

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

Геометричний розподіл застосовують у різноманітних задачах статистичного контролю якості виробів, у теорії надійності та страхових розрахунках.

Зауваження. Можливий також випадок, коли перше значення випадкової величини – 0. Тоді ряд розподілу має вигляд

X	0	1	2	$\ldots$	n	$\ldots$
P	p	pq	pq^2	$\ldots$	pq^n	$\ldots$

Можна довести, що в даному випадку $M(X)=\frac{q}{p}$, $D(X)=\frac{q}{p^2}$.

Наприклад, розглянемо таку задачу.

В урні міститься 5 білих і 7 чорних куль. Випробування полягає у вийманні по одній кулі з поверненням. Розглянемо дві випадкові величини: X – кількість вийнятих чорних куль до першого виймання білої кулі; Y – кількість виймань, які закінчуються з першим вийманням білої кулі. Скласти ряди розподілу для цих випадкових величин.

Розв'язання. За умовою задачі та класичним означенням імовірності $p=\frac{5}{12}$ – ймовірність виймання білої кулі в разі одного виймання, тоді $q=1-p=\frac{7}{12}$.

Випадкова величина X набуде значення 0, якщо першою вийнята біла куля, тому $P(X=0)=p=\frac{5}{12}$. Також X = 1, якщо першою вийнята чорна куля, а другою – біла. Відповідно

$$P(X=1)=pq=\frac{5}{12}\cdot\frac{7}{12}=\frac{35}{144}.$$

Аналогічно, для $X = 2$ обчислюємо

$$P(X = 2) = pq^2 = \frac{5}{12} \cdot \left(\frac{7}{12}\right)^2 = \frac{245}{1728}$$

і так далі. Маємо ряд розподілу для випадкової величини *X*:

X	0	1	2	…	n	…
P	p	pq	pq^2	…	pq^n	…
	$\frac{5}{12}$	$\frac{5}{12} \cdot \frac{7}{12}$	$\frac{5}{12} \cdot \left(\frac{7}{12}\right)^2$	…	$\frac{5}{12} \cdot \left(\frac{7}{12}\right)^n$	…

Випадкова величина Y набуває найменшого значення 1, якщо першою вийняли білу кулю, оскільки виймання припиняється. Тоді $P(Y = 1) = p = \frac{5}{12}$.

Ряд розподілу для випадкової величини Y буде таким:

Y	1	2	3	…	n	…
P	p	pq	pq^2	…	pq^{n-1}	…
	$\frac{5}{12}$	$\frac{5}{12} \cdot \frac{7}{12}$	$\frac{5}{12} \cdot \left(\frac{7}{12}\right)^2$	…	$\frac{5}{12} \cdot \left(\frac{7}{12}\right)^{n-1}$	…

Приклад 2. Проводять багаторазові випробування певного елемента на надійність доти, доки елемент не відмовить. Знайти математичне сподівання і дисперсію дискретної випадкової величини *X* – кількості випробувань, які потрібно провести. Ймовірність відмови елемента в кожному випробуванні дорівнює 0,1.

Розв'язання. За умовою $p = 0{,}1;\ q = 1 - p = 0{,}9,$ тоді

$$M(X) = \frac{1}{p} = \frac{1}{0{,}1} = 10;$$

$$D(X) = \frac{q}{p^2} = \frac{0{,}9}{0{,}01} = 90.$$

4.1.4. Гіпергеометричний розподіл

Такий розподіл має вигляд

$$P(X=m)=\frac{C_k^m\cdot C_{N-k}^{n-m}}{C_N^n},\ m=0,\ 1,\ 2,\ \ldots,\ n,\ k\geq n.$$

Він визначає ймовірність отримання m елементів з певною властивістю серед n елементів, взятих із сукупності N елементів, яка містить k елементів саме з такою властивістю.

Цей розподіл використовують у багатьох задачах статистичного контролю якості.

Зауваження. Якщо обсяг вибірки n малий порівняно з обсягом N сукупності, тобто

$$\frac{n}{N}\leq 0{,}1;\ \frac{n}{k}\leq 0{,}1;\ \frac{n}{N-k}\leq 0{,}1,$$

то ймовірності у гіпергеометричному розподілі будуть близькими до відповідних імовірностей біномного розподілу з $p=\dfrac{k}{N}$.

У статистиці це означає, що розрахунки ймовірностей для вибірки без повторення мало відрізнятимуться від розрахунків імовірностей для повторної вибірки.

Для гіпергеометричного розподілу

$$M(X)=\frac{kn}{N},\ D(X)=\frac{nk(N-k)}{N^2}\cdot\frac{(N-n)}{(N-1)}. \tag{4.3}$$

Приклад 3. Партія містить 200 виробів, серед яких 25 бракованих. Для перевірки якості з партії відібрали 10 виробів. Якщо кількість бракованих виробів не перевищує одиниці, то партію приймають. Знайти ймовірність того, що партію буде прийнято. Визначити цю саму ймовірність, якщо апроксимувати гіпергеометричний розподіл біномним розподілом і законом розподілу Пуассона.

Розв'язання. Застосуємо формулу гіпергеометричного закону розподілу. Партію буде прийнято, якщо кількість бракованих серед відібраних десяти дорівнюватиме нулю або одиниці.

$$P\left(X\le 1\right)=P\left(X=0\right)+P\left(X=1\right)=\frac{C_{175}^{10}}{C_{200}^{10}}+\frac{C_{25}^{1}\cdot C_{175}^{9}}{C_{200}^{10}}\approx 0,638.$$

Обчислимо цю саму ймовірність за допомогою формули біномного закону розподілу,

$$p=\frac{25}{200}=\frac{1}{8};$$

$$P\left(X\le 1\right)=P\left(X=0\right)+P\left(X=1\right)=\left(\frac{7}{8}\right)^{10}+10\cdot\frac{1}{8}\cdot\left(\frac{7}{9}\right)^{9}\approx 0,639.$$

Обчислимо цю саму ймовірність за допомогою закону розподілу Пуассона:

$$\lambda=np=10\cdot\frac{1}{8}=1,25.$$

$$P\left(X\le 1\right)=P\left(X=0\right)+P\left(X=1\right)=e^{-1,25}+1,25e^{-1,25}\approx 0,644.$$

Як бачимо, похибки обчислення в разі апроксимації гіпергеометричного розподілу порівняно невеликі.

4.1.5. Поліномний розподіл

Цей розподіл має вигляд

$$P_n(X_1=m_1;\ X_2=m_2;\ \ldots,\ X_s=m_s)=\frac{n!}{m_1!\cdot m_2!\cdot\ldots\cdot m_s!}p_1^{m_1}\cdot p_2^{m_2}\cdot\ldots\cdot p_s^{m_s}.$$

Його застосовують тоді, коли внаслідок кожного із здійснених повторних випробувань може з’явитися s різних подій A_i з імовірностями p_i, причому $\sum_{i=1}^{s}p_i=1$. Його числові характеристики такі:

$$M(X_i)=np_i,\ D(X_i)=np_iq_i,\ q_i=1-p_i, i=1,\ \ldots,\ s.$$

Отже, поліномний розподіл є узагальненням біномного.

4.1.6. Дискретний рівномірний розподіл

Дискретна випадкова величина рівномірно розподілена, якщо вона набуває n значень з однаковими ймовірностями. Згідно з умовою нормування

$$\sum_{i=1}^{n}p_i=1,\ \ p_i=\frac{1}{n},\ i=1,2,...,n.$$

Прикладом рівномірно розподіленої випадкової величини є кількість очок, що випадає за одного підкидання грального кубика. Її закон розподілу має вигляд

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

Легко показати, що для рівномірно розподіленої випадкової величини

$$M(X)=\frac{n+1}{2},\ D(X)=\frac{n^2-1}{12}. \tag{4.4}$$

Подамо у вигляді таблиці залежності основних числових характеристик дискретних розподілів від параметрів цих розподілів.

№	Розподіл	$M(X)$	$D(X)$
1	Біномний	np	npq
2	Пуассонівський	λ	λ
3	Геометричний: *а*) $P(X=n)=pq^n,\ n=1,2,\dots$ *б*) $P(X=n)=pq^n,\ n=0,1,2,\dots$	 $\frac{1}{p}$ $\frac{q}{p}$	 $\frac{q}{p^2}$ $\frac{q}{p^2}$
4	Гіпергеометричний	$\frac{kn}{N}$	$\frac{nk(N-k)}{N^2}\cdot\frac{(N-n)}{(N-1)}$
5	Рівномірний	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$

Зауваженння. В EXCEL для обчислення ймовірності окремого значення біномного розподілу або значення випадкової величини за заданою ймовірністю використовують функції БИНОМРАСП і КРИТБИНОМ. Функцію БИНОМРАСП використовують для обчислення ймовірності в задачах із фіксованою кількістю випробувань, коли результатом будь-якого випробування може бути тільки «успіх» або «невдача»; випробування незалежні; ймовірність «успіху» – стала величина протягом усього експерименту. Ця функція використовує параметри і має запис БИНОМРАСП *(число_успехов, число_ис-*

пытаний, вероятность_успеха, интегральная), де *число_успехов* – кількість успішних випробувань; *число_испытаний* − кількість проведених незалежних випробувань, при цьому кількість випробувань і кількість успіхів – цілі числа; *вероятность_успеха* − ймовірність «успіху» кожного випробування; *интегральная* − логічне значення, яке визначає форму успіху. Якщо ця величина дорівнює 1 (ИСТИНА), функція БИНОМРАСП повертає інтегральну функцію розподілу, тобто обчислює ймовірність того, що кількість успішних випробувань не менша, ніж значення аргумента *число_успеха*; якщо логічне значення дорівнює 0 (ЛОЖЬ), тоді обчислюється значення функції щільності ймовірності, тобто ймовірність того, що кількість успішних випробувань точно дорівнює значенню аргумента *число_успеха*.

Функція КРИТБИНОМ обчислює найменше значення кількості успішних випробувань випадкової величини, для якого інтегральний біномний розподіл більший від заданої величини (критерію) або дорівнює їй. Цю функцію використовують у задачах, пов'язаних з контролем якості.

Аналогічно можна використовувати для відповідних розподілів випадкових величин функції ПУАССОН та ГИПЕРГЕОМЕТ із їх параметрами задання.

Приклад 4. По мішені виконують чотири незалежні постріли. Ймовірність влучення в разі одного пострілу дорівнює 0,25. Скласти ряд розподілу випадкової величини X – кількості влучень у мішень та обчислити його основні числові характеристики. Визначити функцію розподілу $F(x)$ та побудувати її графік.

Розв'язання. Випадкова величина X – кількість влучень у мішень – може набувати значення 0, 1, 2, 3, 4. Оскільки розглядані випробування задовольняють схему Бернуллі, то X має біномний закон розподілу. У даному разі $n=4$, $p=0,25$, $q=1-p=0,75$.

Для складання ряду розподілу використаємо функцію Excel категорії «Статистические» БИНОМРАСП (*число_успехов; число_испытаний; вероятность_успеха; интегральная*) із такими параметрами: *число_успехов* : k – змінна величина, яка набуває значення: 0, 1, 2, 3, 4; *число_испытаний* : 4 – кількість незалежних випробувань; *вероятность_успеха:* 0,25 – ймовірність успіху у кожному випробуванні; *интегральная:* 0 − для визначення ймовірності випадкової події $P(X=k)$.

Відповідні ймовірності, знайдені за допомогою цієї функції БИНОМРАСП, запишемо у вигляді ряду розподілу:

k	0	1	2	3	4	Сума
P	0,3164	0,4219	0,2109	0,0469	0,0039	1

В останньому стовпчику знайдена сума $\sum_{k=0}^{4} p_k$ для перевірки умови нормування.

Знайдемо основні числові характеристики розподілу цієї випадкової величини: математичне сподівання, дисперсію, середнє квадратичне відхилення.

Оскільки у цьому разі маємо справу із дискретною випадковою величиною, яка має біномний розподіл, основні числові характеристики можна обчислити так:

$$M(X) = np = 4 \cdot 0,25 = 1;$$

$$D(X) = npq = 4 \cdot 0,25 \cdot 0,75 = 0,75; \;\; \sigma(X) = \sqrt{0,75} = 0,866025.$$

У загальному випадку математичне сподівання дискретної випадкової величини обчислюють за формулою

$$M\left(X\right) = \sum_{k=0}^{4} x_k p_k = 0 \cdot 0,3164 + 1 \cdot 0,4219 + 2 \cdot 0,2109 + 3 \cdot 0,0469 + 4 \cdot 0,0039 = 1.$$

Для обчислення дисперсії знайдемо $M\left(X^2\right)$:

$$M\left(X^2\right) = \sum_{k=0}^{4} \left(x_k\right)^2 p_k = 0 \cdot 0,3164 + 1 \cdot 0,4219 + 4 \cdot 0,2109 + 9 \cdot 0,0469 + 16 \cdot 0,0039 = 1,75.$$

Звідси

$$D(X) = M(X^2) - \left(M(X)\right)^2 = 1,75 - 1^2 = 0,75.$$

Середнє квадратичне відхилення обчислимо за формулою

$$\sigma(X) = \sqrt{D(X)} = \sqrt{0,75} = 0,866025.$$

Для функції розподілу випадкової величини скористаємось функцією БИНОМРАСП (*число_успехов; число_испытаний; вероятность_успеха; интегральная*) з параметрами: *число_успехов*: k – змінна величина, яка набуває значення 0, 1, 2, 3, 4; *число_испытаний*: 4 – кількість незалежних випробувань; *вероятность_успеха:* 0,25 – ймовірність успіху кожного випробування; *интегральная:* 1 – для знаходження функції розподілу $F(x)$.

Відповідні значення, знайдені за допомогою даної функції, наведено в таблиці:

X	0	1	2	3	4
$F(x)$	0,3164	0,7383	0,9492	0,9961	1

Зауваження. У разі використання функції БИНОМРАСП для побудови функції розподілу $F(x)$ потрібно врахувати, що БИНОМРАСП повертає значення $P(X \le x)$, а не $P(X < x)$. Результат застосування табличного процесора Microsoft Excel для розв’язування *прикладу* 4 наведено на рис. 4.1.

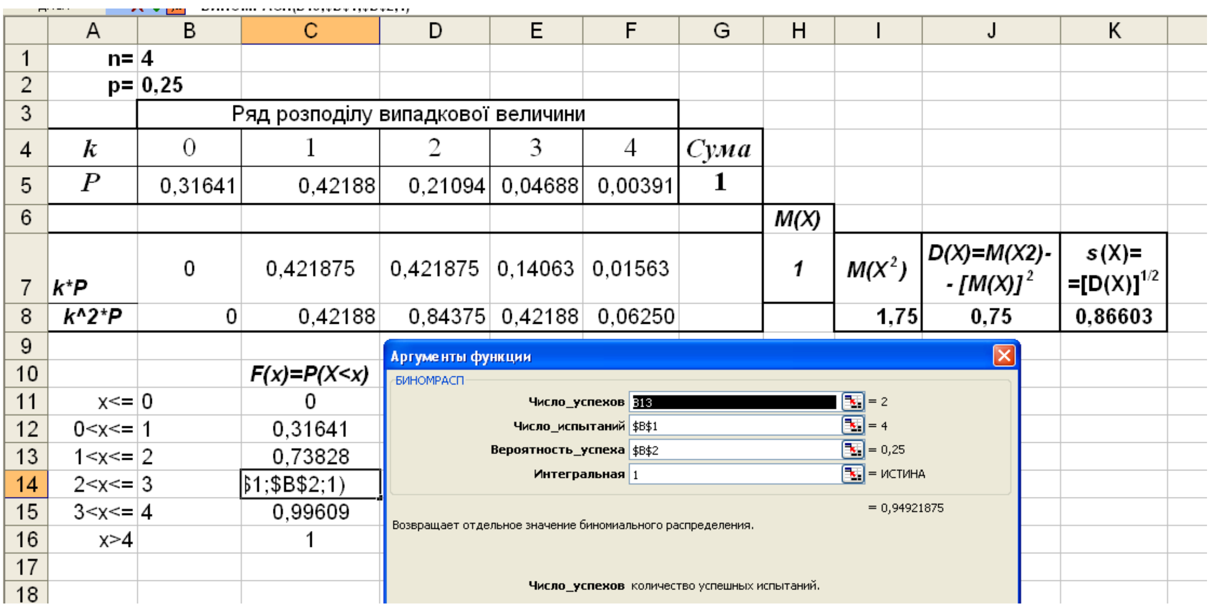


Рис. 4.1. Вигляд таблиці до прикладу 4 на моніторі

Функція розподілу $F(x)$ має такий вигляд:

$$F(x)=\begin{cases} 0, & \text{ÿêù î} \quad x \le 0; \\ 0,3164, & \text{ÿêù î} \quad 0 < x \le 1; \\ 0,7383, & \text{ÿêù î} \quad 1 < x \le 2; \\ 0,9492, & \text{ÿêù î} \quad 2 < x \le 3; \\ 0,9961, & \text{ÿêù î} \quad 3 < x \le 4; \\ 1, & \text{ÿêù î} \quad x > 4, \end{cases}$$

а її графік подано далі (рис. 4.2).

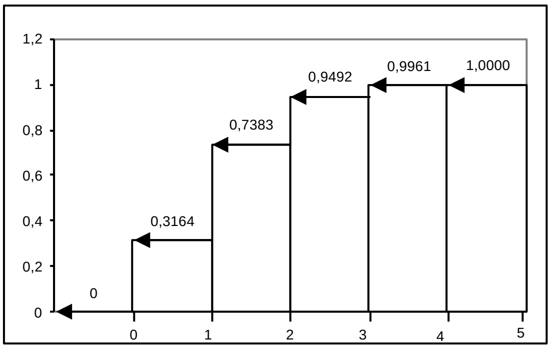


Рис. 4.2. Графік функції розподілу $F(x)$

4.2. Основні закони розподілу неперервних випадкових величин, їх основні числові характеристики

4.2.1. Рівномірний розподіл

Означення. Величина X розподілена рівномірно на проміжку $[a,\ b]$, якщо усі її можливі значення належать цьому проміжку і щільність розподілу її ймовірностей має такий вигляд:

$$f(x)=\begin{cases}\dfrac{1}{b-a},\ x\in[a,\ b];\\ 0,\ x\notin[a,\ b].\end{cases} \tag{4.5}$$

Для рівномірно розподіленої випадкової величини на проміжку $[a, b]$:

$$P(x_1<X<x_2)=\frac{x_2-x_1}{b-a},\ a\le x_1\le x_2\le b,$$

тобто ймовірність потрапляння X в інтервал $(x_1,\ x_2)$ дорівнює відношенню довжини цього інтервалу до довжини всього проміжку $[a,\ b]$.

Функція розподілу рівномірного закону розподілу має такий вигляд:

$$F(x)=\begin{cases}0,\ x<a;\\ \dfrac{x-a}{b-a},\ x\in[a,\ b];\\ 1,\ x>b.\end{cases}$$

Рівномірний закон розподілу легко моделювати. За допомогою функціональних перетворень із величин, розподілених рівномірно, можна отримувати величини з довільним законом розподілу. Графіки щільності ймовірності і функції розподілу наведено на рис. 4.3. та рис. 4.4.

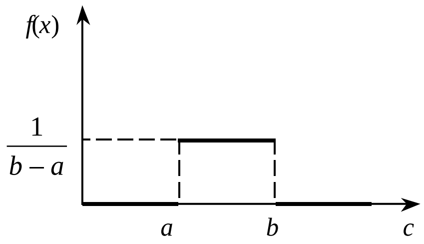


Рис. 4.3. Щільність розподілу ймовірності

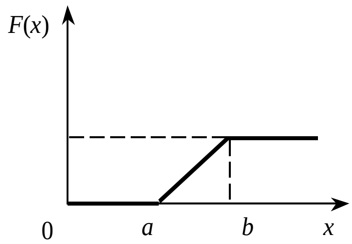


Рис. 4.4. Функція розподілу

Цей розподіл задовольняють похибки округлення різноманітних розрахунків.

Знайдемо числові характеристики цього розподілу:

$$M(X)=\int_a^b xf(x)dx=\int_a^b \frac{x}{b-a}dx=\frac{1}{b-a}\left.\frac{x^2}{2}\right|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{(b+a)(b-a)}{2(b-a)}=\frac{a+b}{2}.$$

Отже, $M(X)=\dfrac{a+b}{2}$, тобто $M(X)$ – середина відрізка $[a,b]$.

Для дисперсії $D(X)=\int_a^b x^2 f(x)dx-(M(X))^2=\int_a^b \dfrac{x^2}{b-a}dx-\left(\dfrac{a+b}{2}\right)^2=$

$$=\frac{1}{b-a}\left.\frac{x^3}{3}\right|_a^b-\frac{a^2+2ab+b^2}{4}=\frac{b^3-a^3}{3(b-a)}-\frac{a^2+2ab+b^2}{4}=$$

$$=\frac{4(a^2+ab+b^2)-3(a^2+2ab+b^2)}{12}=\frac{(b-a)^2}{12}.$$

Таким чином, $M(X)=\dfrac{a+b}{2}$,

$$D(X)=\frac{(b-a)^2}{12}. \tag{4.6}$$

Приклад 5. Рівномірно розподілена випадкова величина X задана щільністю розподілу

$$f(x)=\begin{cases}\dfrac{1}{2l},\ x\in(a-l,\ a+l);\\ 0,\ x\notin(a-l,\ a+l).\end{cases}$$

Знайти $M(X),\ D(X)$.

Розв'язання. За формулами (4.5), (4.6)

$$M(X)=\frac{a-l+a+l}{2}=a,\ D(X)=\frac{(a+l-a+l)^2}{12}=\frac{4l^2}{12}=\frac{l^2}{3}.$$

4.2.2. Показниковий розподіл

Означення. Випадкову величина X називають розподіленою за показниковим розподілом, якщо її щільність розподілу ймовірностей має такий вигляд:

$$f(x)=\begin{cases}\lambda e^{-\lambda x},\ x>0;\\ 0,\ x\le 0,\end{cases}$$

де $\lambda>0$ – параметр розподілу.

Показниковий розподіл задовольняють: час телефонної розмови, час ремонту техніки, час безвідмовної роботи комп'ютерної мережі.

Якщо випадкова величина X має показниковий розподіл з параметром λ, тоді

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

Дійсно, $M(X)=\lambda\int\limits_0^{+\infty}xe^{-\lambda x}dx=\lambda\lim\limits_{b\to\infty}\int\limits_0^b xe^{-\lambda x}dx=\begin{vmatrix}u=x & du=dx\\ dv=e^{-\lambda x}dx & v=-\frac{1}{\lambda}e^{-\lambda x}\end{vmatrix}=$

$$=\lambda\lim_{b\to\infty}\left(-\frac{1}{\lambda}xe^{-\lambda x}\Big|_0^b+\frac{1}{\lambda}\int_0^b e^{-\lambda x}dx\right)=-\lim_{b\to\infty}be^{-\lambda b}-\lim_{b\to\infty}\frac{1}{\lambda}e^{-\lambda x}\Big|_0^b=$$

$$=-\lim_{b\to\infty}\frac{1}{\lambda e^{\lambda b}}+\frac{1}{\lambda}=\frac{1}{\lambda}.$$

Для дисперсії обчислимо $M(X^2)=\lambda\int\limits_0^{+\infty}x^2e^{-\lambda x}dx=\lambda\lim\limits_{b\to\infty}\int\limits_0^b x^2e^{-\lambda x}dx=$

$$=\begin{vmatrix}u=x^2 & du=2xdx\\ dv=e^{-\lambda x}dx & v=-\frac{1}{\lambda}e^{-\lambda x}\end{vmatrix}=\lambda\lim_{b\to\infty}\left(-\frac{1}{\lambda}x^2e^{-\lambda x}\Big|_0^b+\frac{2}{\lambda}\int_0^b xe^{-\lambda x}dx\right)=$$

$$=-\lim_{b\to\infty}\frac{b^2}{e^{\lambda b}}+\frac{2}{\lambda}\cdot\frac{1}{\lambda}=\frac{2}{\lambda^2}.$$

$$D(X)=M(X)-(M(X))^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2},\ \text{тобто}\ D(X)=\frac{1}{\lambda^2}.$$

Отже, для показникового розподілу

$$M(X)=\frac{1}{\lambda},\ D(X)=\frac{1}{\lambda^2}. \qquad (4.7)$$

Зауваження. Якщо випадкова величина X розподілена за показниковим розподілом, то її інтегральна функція розподілу має такий вигляд:

$$F(x)=\begin{cases}1-e^{-\lambda x},\ x\ge 0;\\ 0,\ x<0,\end{cases}$$

відповідно

$$P\left(a<X<b\right)=\begin{cases}e^{-a\lambda}-e^{-b\lambda},\ a\ge 0;\\ 1-e^{-b\lambda},\ a<0,\ b>0;\\ 0,\ b<0.\end{cases}$$

Графіки диференціальної та інтегральної функцій показникового розподілу зображено на рис. 4.5 *а*, *б*.

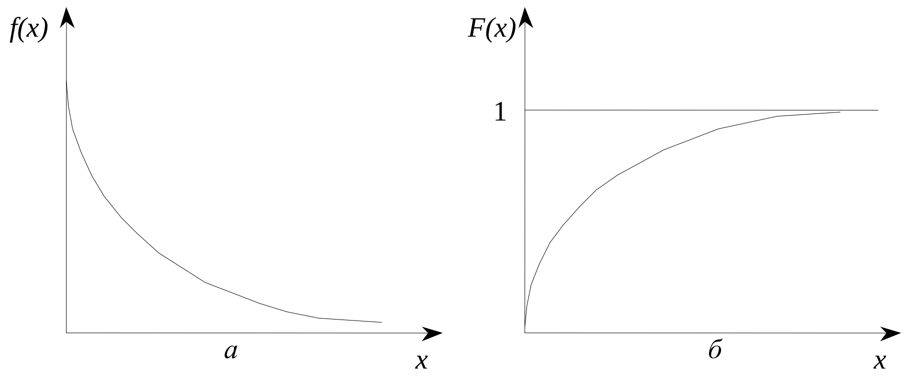


Рис. 4.5. Диференціальна та інтегральна функції показникового розподілу: *а* – $f(x)$; *б* – $F(x)$

Показниковий розподіл часто трапляється в теорії масового обслуговування, теорії надійності. Нехай t – час безвідмовної роботи деякого елемента, λ – інтенсивність відмов (середня кількість відмов за одиницю часу). Тоді час роботи елемента t можна вважати неперервною випадковою величиною, розподіленою за показниковим законом із функцією розподілу

$$F(x)=p(t<x)=1-e^{-\lambda x}\ \ (\lambda>0),$$

яка визначає ймовірність відмови елемента за час, менший від x.

Імовірність

$$R(x)=p(t\ge x)=e^{-\lambda x}$$

називають ***функцією надійності***, яка визначає ймовірність безвідмовної роботи елемента за час *x*.

Приклад 5. Час t роботи MP3-плеєра має показниковий розподіл. Знайти ймовірність його безвідмовної роботи протягом 600 годин, якщо середній час роботи плеєра – 400 годин.

Розв'язання. За умовою задачі математичне сподівання величини t дорівнює 400, тому $\lambda = \frac{1}{m} = \frac{1}{400}$. За формулою $R(x) = p(t \geq x) = e^{-\lambda x}$, маємо:

$$R(x = 600) = e^{-\frac{600}{400}} = e^{-1,5} = 0,2231.$$

Отже, ймовірність того, що MP3-плеєр працюватиме не менше 600 годин, дорівнює 22 %.

Приклад 6. Знайти числові характеристики випадкової величини, розподіленої за законом

$$f(x) = \begin{cases} 4e^{-4x}, \ x \geq 0; \\ 0, \ x < 0. \end{cases}$$

Розв'язання. Для показникового розподілу випадкової величини з параметром $\lambda = 4$ (4.7) маємо:

$$M(X) = \frac{1}{\lambda} = \frac{1}{4} = 0,25; \ D(X) = \frac{1}{\lambda^2} = \frac{1}{16} = 0,0625.$$

Приклад 7. Величина X розподілена за показниковим розподілом

$$f(x) = \begin{cases} 3e^{-3x}, \ x \geq 0; \\ 0, \ x < 0. \end{cases}$$

Знайти ймовірність того, що X потрапить в інтервал $(0,4; 1)$.

Розв'язання. Параметр розподілу $\lambda = 3$, тому

$$P(0,4 < X < 1) = e^{-0,4 \cdot 3} - e^{-1 \cdot 3} = e^{-1,2} - e^{-3} = 0,302 - 0,050 = 0,252.$$

Розділ 5. Нормальний закон розподілу та його значення у теорії ймовірностей. Граничні теореми теорії ймовірностей

У значній частині задач випадкова величина – це сума великої кількості малих доданків, що визначаються факторами, які діють незалежно один від одного. За таких обставин слід очікувати, що випадкова величина має нормальний розподіл або близький до нього. Отже, така особлива увага до цього розподілу цілком виправдана.

5.1. Нормальний закон розподілу, його основні характеристики

Означення. Випадкову величину X називають *розподіленою нормально*, якщо її диференціальна функція розподілу ймовірностей має вигляд

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}},$$

де $a,\ \sigma$ – параметри розподілу.

Графік функції $f(x)$ називають *нормальною кривою* або *кривою Гаусса*.

За $a=0$ та $\sigma=1$ нормальну криву називають *нормованою*, функція розподілу щільності має вигляд

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

тобто це функція Лапласа, табульована, зокрема в електронній формі.

Нормальний закон розподілу повністю визначається своїм математичним сподіванням та дисперсією (середнім квадратичним відхиленням):

$$M(X)=a,\ D(X)=\sigma^2,\ \sigma(X)=\sigma. \tag{5.1}$$

Справді, $M(X)=\int\limits_{-\infty}^{+\infty}xf(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}xe^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left|\begin{array}{l}t=\frac{x-a}{\sqrt{2}\sigma},\\ x=\sigma\sqrt{2}t+a\\ dx=\sigma\sqrt{2}dt\end{array}\right|=$

$$=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma\sqrt{2}t+a)e^{-t^2}\sigma\sqrt{2}dt=\frac{\sigma^2 2}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}te^{-t^2}dt+\frac{\sigma\sqrt{2}a}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-t^2}dt=$$

$=\frac{a}{\sqrt{\pi}}\sqrt{\pi}=a$, оскільки $\int_{-\infty}^{+\infty}te^{-t^2}dt=0$, як інтеграл від непарної функції з симетричними межами (за умови збіжності інтеграла $\int_{0}^{+\infty}te^{-t^2}dt$) і $\int_{-\infty}^{+\infty}e^{-t^2}dt=\sqrt{\pi}$, як інтеграл Пуассона.

Обчислимо дисперсію

$$D(X)=\int_{-\infty}^{+\infty}\left(x-M(X)\right)^2dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\begin{vmatrix} t=\dfrac{x-a}{\sqrt{2}\sigma} \\ x=\sigma\sqrt{2}t+a \\ dx=\sigma\sqrt{2}dt \end{vmatrix}=$$

$$=\frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty}t^2e^{-t^2}dt=\frac{\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty}t\cdot 2te^{-t^2}dt=\begin{vmatrix} u=t & du=dt \\ dv=2te^{-t^2}dt & v=-e^{-t^2} \end{vmatrix}=-\frac{\sigma^2}{\sqrt{\pi}}te^{-t^2}\Big|_{-\infty}^{+\infty}+$$

$$+\frac{\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-t^2}dt=\frac{\sigma^2}{\sqrt{\pi}}\sqrt{\pi}=\sigma^2.$$

Отже, $M(X)=a,\ D(X)=\sigma^2.$

Для нормального розподілу ймовірність потрапляння випадкової величини в інтервал $(\alpha,\ \beta)$ обчислюють за формулою

$$P(\alpha<X<\beta)=\hat{O}\left(\frac{\beta-a}{\sigma}\right)-\hat{O}\left(\frac{\alpha-a}{\sigma}\right), \tag{5.2}$$

де $\hat{O}(x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{x}e^{-\frac{t^2}{2}}dt$ – функція Лапласа.

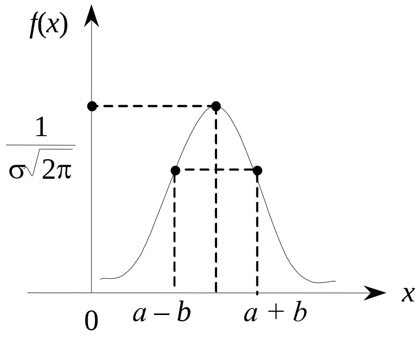
Інтегральна функція розподілу має вигляд

$$F(x)=\int_{-\infty}^{x}f(t)dt=\hat{O}\left(\frac{x-a}{\sigma}\right)-\hat{O}(-\infty)=\hat{O}\left(\frac{x-a}{\sigma}\right)+\frac{1}{2},$$

тобто

$$F(x)=\frac{1}{2}+\hat{O}\left(\frac{x-a}{\sigma}\right). \tag{5.3}$$

Графіки щільності розподілу ймовірностей та функції розподілу наведено відповідно на рис. 5.1 та рис. 5.2.



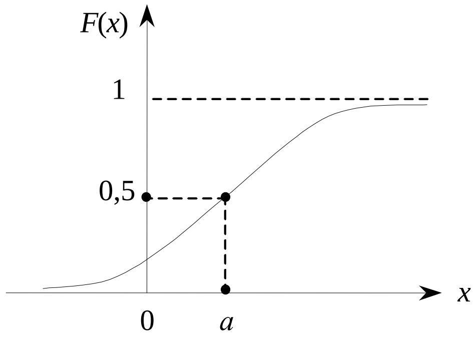


Рис. 5.1. Диференціальна функція нормального розподілу

Рис. 5.2. Інтегральна функція нормального розподілу

Зауваження. Нормально розподілену випадкову величину із параметрами розподілу a, σ часто позначають $N(a,\ \sigma)$.

5.2. Правило трьох сигм

Із рівності (5.2) випливає:

$$\text{Đ}\left(\left|X-a\right|\le\delta\right)=2\hat{\text{O}}\left(\frac{\delta}{\sigma}\right).$$

Зокрема, $P\left(\left|X-a\right|\le t\sigma\right)=2\hat{\text{O}}(t),\ t>0$.

Підставляючи послідовно в останню нерівність $t=1,\ 2,\ 3,$ за допомогою таблиці функції Лапласа отримують:

$$P\left(\left|X-a\right|\le\sigma\right)=2\hat{\text{O}}(1)=0,68;$$

$$P\left(\left|X-a\right|\le 2\sigma\right)=2\hat{\text{O}}(2)=0,95;$$

$$P\left(\left|X-a\right|\le 3\sigma\right)=2\hat{\text{O}}(3)=0,997.$$

Число 0,997 мало відрізняється від одиниці, тому подію з імовірністю 0,997 можна вважати майже достовірною. Рівність

$$P\left(\left|X-a\right|\le 3\sigma\right)=2\hat{\text{O}}(3)=0,997$$

називають ***«правилом*** 3σ***»***: *майже достовірно, що нормально розподілена випадкова величина може відхилятись від свого математичного сподівання не більше, ніж на потроєне середнє квадратичне відхилення.*

На практиці його використовують так: *якщо закон розподілу випадкової величини X невідомий, але* $|X - a| < 3\sigma$, *тоді можна припустити, що X розподілена нормально.*

Нормальний закон розподілу широко застосовують у математичній статистиці. Головна особливість нормального закону полягає в тому, що він є граничним законом, до якого наближуються інші закони розподілу за типових умов.

Приклад 1. Для нормально розподіленої випадкової величини X знайти $P(\alpha < X < \beta)$, якщо $a = 2,\ \sigma = 2,\ \alpha = 1,\ \beta = 6.$

Розв'язання. За формулою $P(\alpha < X < \beta) = \hat{O}\left(\frac{\beta - a}{\sigma}\right) - \hat{O}\left(\frac{\alpha - a}{\sigma}\right)$ маємо

$$P(1 < X < 6) = \hat{O}\left(\frac{6-2}{2}\right) - \hat{O}\left(\frac{1-2}{2}\right) = \hat{O}(2) - \hat{O}(-0,5) = \hat{O}(2) + \hat{O}(0,5) =$$

$$= 0,47725 + 0,19146 = 0,66871.$$

Приклад 2. Зріст студентів ВПІ розподілено за нормальним законом, причому $a = 175$ см, а $\sigma = 6$ см. Визначити ймовірність того, що принаймні один із 5 викликаних навмання студентів матиме зріст від 170 до 180 см.

Розв'язання. Позначимо подію A – із 5 викликаних студентів зріст принаймні одного належить проміжку (170, 180); тоді $\overline{A}$ – зріст усіх 5 викликаних студентів не належить проміжку.

Обчислимо:

$$P(170 < X < 180) = \hat{O}\left(\frac{180-175}{6}\right) - \hat{O}\left(\frac{170-175}{6}\right) = \hat{O}\left(\frac{5}{6}\right) - \hat{O}\left(-\frac{5}{6}\right) =$$

$$= 2\hat{O}\left(\frac{5}{6}\right) = 2\hat{O}(0,833) = 2 \cdot 0,2967 = 0,5934.$$

Отже ймовірність того, що зріст одного викликаного студента не належить проміжку (170, 180)

$$p = 1 - P(170 < X < 180) = 1 - 0,5934 = 0,4066.$$

За теоремою множення ймовірностей незалежних подій знаходимо ймовірність події $\overline{A}$:

$$P(\overline{A}) = (0,4066)^5 = 0,0111.$$

Отже, тоді ймовірність події *A* становить:

$$P(A) = 1 - P(\overline{A});$$

$$P(A) = 1 - 0{,}0111 = 0{,}9889.$$

Зауваження. Для обчислень, пов'язаних із нормальним законом розподілу, засобами Excel використовують функцію НОРМРАСП (*x*; *среднее*; *стандартное_отклонение*; *интегральная*), де *x* – значення, для якого будується розподіл; *«среднее»* − середнє арифметичне розподілу (математичне сподівання); «*стандартное_отклонение*» − стандартне відхилення розподілу; *«интегральная»* − логічне значення, яке визначає форму функції. Якщо *«интегральная»* = ЛОЖЬ, то функція НОРМРАСП використовує функцію щільності ймовірності, а в разі *«интегральная»* = ИСТИНА – інтегральну функцію і значення функції з геометричного погляду, дає площу криволінійної фігури, обмеженої кривою щільності розподілу ймовірностей, вертикальною прямою, яка проходить через точку *x* (ліворуч від цієї прямої) та віссю абсцис. Функцію НОРМСТРАСП використовують для стандартного нормального розподілу, для якого математичне сподівання дорівнює 0 і стандартне відхилення $\sigma = 1$.

5.3. Розподіли χ^2 та Стьюдента

5.3.1. Розподіл χ^2

Нехай $X_i (i = 1,\ 2,\dots, n)$ – нормально розподілені нормовані незалежні величини (їх математичне сподівання дорівнює 0, середні квадратичні відхилення дорівнюють 1). Тоді сума квадратів цих величин

$$\chi^2 = \sum_{i=1}^{n} X_i^2 \tag{5.4}$$

розподілена ***за законом*** χ^2 з $k = n$ степенями вільності.

Якщо величини X_i зв'язані одним лінійним співвідношенням, наприклад

$$\sum_{i=1}^{n} X_i = n\overline{X},$$

то кількість степенів вільності буде $k = n - 1$.

Диференціальна функція розподілу χ^2 має вигляд

$$f(x)=\begin{cases} 0,\ x\le 0; \\ \dfrac{1}{2^{\frac{k}{2}}\cdot \tilde{A}\left(\dfrac{k}{2}\right)}\cdot e^{-\frac{x}{2}}\cdot x^{\frac{k}{2}-1},\ x>0, \end{cases}$$

де $\tilde{A}(x)=\int\limits_0^\infty t^{x-1}e^{-t}dt$ – гамма-функція, $\Gamma(n+1)=n!$.

Розподіл χ^2 визначається параметром – кількість степенів вільності k. Якщо k зростає, розподіл χ^2 прямує до нормального розподілу дуже повільно.

5.3.2. Розподіл Стьюдента

Розподіл Стьюдента було введено 1908 року англійським статистом В. Госсетом, який працював на фабриці з виробництва пива. Ймовірнісно-статистичні методи використовувалися для ухвалення економічних і технічних рішень на цій фабриці, тому її керівництво забороняло В. Госсету публікувати наукові статті під своїм ім’ям. Проте він мав можливість публікуватися під псевдонімом Стьюдент. Історія Госсета-Стьюдента показує, що вже сто років тому менеджерам Великобританії була очевидна велика економічна ефективність ймовірносно-статистичних методів.

Нині розподіл Стьюдента − один з найбільш відомих розподілів серед використовуваних під час аналізу реальних даних. Його застосовують для оцінювання математичного сподіванння, прогнозного значення та інших характеристик за допомогою надійних інтервалів, під час перевірки гіпотез про значення математичних сподівань, коефіцієнтів регресійної залежності, гіпотез однорідності вибірок.

Якщо випадкова величина Y має стандартний розподіл ($N(0;\ 1)$), а випадкова величина X − розподіл χ^2 із k степенями вільності, то величина

$Z_k = Y\sqrt{\dfrac{k}{\chi_k^2}}$ характеризується ***t*-*розподілом або розподілом Стьюдента*** з k степенями вільності та диференціальною функцією розподілу

$$f(z) = \frac{\Gamma\left(\dfrac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\dfrac{k}{2}\right)}\left(1+\frac{z^2}{k}\right)^{-\frac{k+1}{2}}, \qquad -\infty < z < \infty.$$

У разі зростання k ***розподіл Стьюдента*** швидко наближується до нормального розподілу. Загалом крива щільності розподілу подібна за формою до кривої щільності нормального розподілу з математичним сподіванням 0 і дисперсією 1 з тією відмінністю, що у ***t-розподілу*** вона трохи нижча і ширша. За кількості степенів вільності, що прямує до нескінченості, ***t-розподіл*** прямує до нормального розподілу з математичним сподіванням 0 і дисперсією 1.

5.4. Закон великих чисел та центральна гранична теорема

У теорії ймовірностей, особливо в разі її застосування на практиці, важливу роль відіграють події з імовірностями, близькими до нуля або до одиниці, тому одна з основних задач теорії ймовірностей – встановлення закономірностей, що можуть виникати з імовірністю, близькою до одиниці, й особливо таких закономірностей, які виникають в результаті спільної дії великої кількості незалежних випадкових факторів. Закон великих чисел і є одним з найважливіших тверджень такого типу.

Під законом великих чисел у теорії ймовірностей розуміють групу теорем, кожна з яких встановлює умови, за яких середнє арифметичне випадкових величин має властивість стійкості. Тобто, середній результат дії великої кількості випадкових явищ практично перестає бути випадковим. У масі однорідних явищ випадкові особливості компенсуються і виявляються якісь закономірності, що служать предметом дослідження. Саме у виявленні загальних умов, що забезпечують статистичну стійкість середніх, полягає неминуча наукова цінність закону великих чисел.

Граничні теореми, які встановлюють відповідність між теоретичними та дослідними характеристиками випадкових подій, називають ***законом великих чисел.***

5.4.1. Нерівність і теорема Чебишова

Розглянемо спочатку нерівність Чебишова. Нехай випадкова величина X має скінченне математичне сподівання $M(X)$ і скінченну дисперсію $D(X)$. Тоді для будь-якого $\varepsilon > 0$ виконується ***нерівність Чебишова***

$$P\left(\left|X - M(X)\right| < \varepsilon\right) \geq 1 - \frac{D(X)}{\varepsilon^2}. \tag{5.5}$$

Перейшовши до протилежної події, маємо:

$$P\left(\left|X - M(X)\right| \geq \varepsilon\right) \leq \frac{D(X)}{\varepsilon^2}. \tag{5.6}$$

Наслідок (***правило*** 3σ). Для будь-якої випадкової величини X, яка має скінченне математичне сподівання та скінченну дисперсію, виконується нерівність

$$P\left(\left|X - M(X)\right| < 3\sigma\right) \geq 1 - \frac{\sigma^2}{(3\sigma)^2} = 1 - \frac{1}{9} = \frac{8}{9}. \tag{5.7}$$

Зауваження. Нерівність Чебишова на практиці має обмежене значення, оскільки часто дає грубу оцінку. Нехай, наприклад, $\varepsilon = \dfrac{\sigma}{2}$, тоді за нерівністю (5.6) матимемо:

$$P\left(\left|X - M(X)\right| \geq \frac{\sigma}{2}\right) \leq \frac{\sigma^2}{\sigma^2 / 4} = 4.$$

Але і так зрозуміло, що ймовірність не може бути більшою за одиницю. Якщо, тепер $\varepsilon = 10\cdot\sigma$, то

$$P\left(\left|X - M(X)\right| \geq 10\sigma\right) \leq \frac{\sigma^2}{100\sigma^2} = 0{,}01.$$

Це вже непогана оцінка для ймовірності.

Попри те, що на практиці цей варіант закону великих чисел застосовують нечасто, теоретичне значення нерівності Чебишова дуже велике.

Однією з фундаментальних теорем теорії ймовірностей є **теорема Чебишова.** *Ймовірність того, що абсолютне відхилення середнього арифметичного попарно незалежних випадкових величин від середнього арифметичного їх математичних сподівань не перевищить як завгодно малого додатного числа, буде як завгодно близька до одиниці, якщо кількість випадкових величин достатньо велика.*

Теорема Чебишова дає змогу обґрунтувати правило середнього, яким широко користуються в практиці вимірювань: за достатньо великої кількості вимірювань з імовірністю, як завгодно близькою до одиниці, можна вважати, що середнє арифметичне результатів вимірювань буде як завгодно мало відрізнятись від вимірюваної величини, тобто точно характеризує вимірювану величину.

Теорема Чебишова має велике теоретичне, практичне та методологічне значення. На цій теоремі ґрунтується вибірковий метод, який застосовують у статистиці. Його суть полягає в тому, що за порівняно невеликою випадковою вибіркою роблять висновок про всю сукупність досліджуваних об’єктів.

Означення. Послідовність випадкових величин $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ називається *збіжною за ймовірністю* до числа γ, якщо для довільного $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|X_n - \gamma\right| < \varepsilon\right) = 1.$$

Позначення збіжності за ймовірністю $X_n \xrightarrow{p} \gamma,\ n \to \infty.$

Теорема Чебишова (закон великих чисел). *Нехай* $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ *– послідовність попарно незалежних випадкових величин зі скінченними математичними сподіваннями* $M(X_i) = a_i < \infty$ *і дисперсіями, обмеженими однією і тією ж сталою* $D(X_i) \le C, i \in N$. *Тоді*

$$\frac{X_1 + X_2 + \cdots + X_n}{n} - \frac{a_1 + a_2 + \cdots + a_n}{n} \xrightarrow{p} 0,\ n \to \infty. \tag{5.8}$$

Наслідок. Нехай $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ – послідовність попарно незалежних випадкових величин з однаковими математичними сподіваннями $M(X_i) = a$ і дисперсіями, обмеженими однією і тією ж сталою $D(X_i) \le C,\ i \in N$. Тоді

$$\frac{X_1 + X_2 + \cdots + X_n}{n} - a \xrightarrow{p} 0,\ n \to \infty.$$

Отже, це твердження обґрунтовує правило середнього арифметичного. За виконаних n незалежних вимірювань, позбавлених систематичних помилок та здійснених з деякою гарантованою точністю, *за достатньо великої кількості вимірювань з імовірністю, як завгодно близькою до одиниці, середнє арифметичне результатів вимірювань буде як завгодно мало відрізнятися від вимірюваної величини.*

Теорема Бернуллі. Закон великих чисел у формі Бернуллі – безпосередній наслідок теореми Чебишова. Теорема Бернуллі відображає в математичній формі об'єктивно притаманну багатьом випадковим явищам властивість стійкості відносних частот. Ця теорема дає теоретичне обґрунтування заміни невідомої ймовірності події її частотою або статистичною ймовірністю, одержаною в n послідовних незалежних випробуваннях, що проводять за одного і того самого комплексу умов. Ця теорема поклала початок теорії ймовірностей як науки.

Теорема. *Нехай* μ_n *– кількість настання події A за n послідовних незалежних випробувань, у кожному з яких імовірність настання події A дорівнює p. Тоді*

$$\frac{\mu_n}{n} \xrightarrow{p} p, \ n \to \infty. \tag{5.9}$$

Теорема Маркова. *Нехай* $X_1, \ X_2, \ \ldots, \ X_n, \ \ldots$ *– послідовність випадкових величин зі скінченними математичними сподіваннями* $(M(X_i) = a_i < \infty)$*, для яких виконується умова* $\lim\limits_{n\to\infty} \dfrac{D(X_1 + X_2 + \cdots + X_n)}{n^2} = 0.$

Тоді

$$\frac{X_1 + X_2 + \cdots X_n}{n} - \frac{a_1 + a_2 + \cdots + a_n}{n} \xrightarrow{p} 0, \quad n \to \infty. \tag{5.10}$$

Твердження Маркова справедливе для будь-якого закону розподілу.

Зауваження. Велике практичне значення має теорема Пуассона, яка є узагальненням теореми Бернуллі на випадок, коли незалежні випробування відбуваються не обов'язково за однакових умов. Наприклад, часто доводиться перевіряти відповідність теоретично обчисленої ймовірності події фактичній відносній частоті цієї події. Для визначення відносної частоти потрібне багаторазове відтворення експерименту, та далеко не завжди це можна зробити в тих самих умовах. Однак перевірку можна здійснити, якщо вдається обчислити ймовірності для різних умов. У цьому випадку можна порівняти спостережену в досліді відносну частоту із середнім арифметичним обчислених імовірностей.

Закон великих чисел лежить в основі різних видів страхування (страхування життя людини на різні терміни, страхування майна тощо). У разі планування асортименту товарів широкого ужитку враховують попит на них у населення. У цьому попиті виявляється дія закону великих чисел.

5.4.2. Центральна гранична теорема

Граничні теореми, що встановлюють граничні закони розподілу випадкових величин, об’єднують загальною назвою – ***центральна гранична теорема.***

У групі теорем, що носять загальну назву *«центральна гранична теорема»,* йдеться про закономірності розподілу суми досить великої кількості випадкових величин. Різні форми центральної граничної теореми відрізняються одна від одної тими чи тими умовами, за яких виникає нормальний закон. Досить загальні умови має **теорема О. М. Ляпунова*:*** *якщо випадкова величина є сумою взаємно незалежних випадкових величин* X_1, X_2, …, X_n, *кожна з яких мало впливає на суму, то за наявності досить великої кількості доданків n закон розподілу їх суми стає як завгодно близьким до нормального, незважаючи на закон розподілу доданків.*

За теоремою Ляпунова вплив кожного окремого доданка на суму за великих n дуже малий, і в разі необмеженого збільшення кількості доданків закон розподілу їх суми необмежено наближується до нормального з математичним сподіванням і дисперсією, які дорівнюють сумам відповідних числових характеристик доданків.

Центральна гранична теорема пояснює велике поширення нормального закону розподілу і є теоретичною основою застосування нормального розподілу для багатьох практичних задач: *за широких припущень сума великої (але скінченної) кількості незалежних випадкових величин розподілена згідно із законом, близьким до нормального.* Наприклад, на відлагодженому виробництві якість продукції змінюється за нормальним законом унаслідок того, що виробнича похибка є результатом сумарної дії великої кількості випадкових величин. Окремим випадком центральної граничної теореми є інтегральна формула Лапласа.

Отже, і закон великих чисел, і центральна гранична теорема становлять теоретичну основу розв’язання багатьох задач, що виникають у практичній діяльності людини.

Отже, сформулюємо граничні теореми.

Теорема Ляпунова. *Якщо для послідовності попарно незалежних випадкових величин* $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ *можна знайти таке число* $\delta > 0$*, що*

$$\lim_{n\to\infty}\frac{\sum\limits_{k=1}^{n} M\left|X_k - M(X_k)\right|^{2+\delta}}{\left(\sqrt{\sum\limits_{k=1}^{n} D(X_k)}\right)^{2+\delta}} = 0,$$

то

$$\lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} M(X_k)}{\sqrt{\sum\limits_{k=1}^{n} D(X_k)}} < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt.$$

Зауваження. На практиці зазвичай найлегше перевірити умову Ляпунова для $\delta = 1$. Якщо послідовність випадкових величин задовольняє умову Ляпунова, то вона задовольняє також умову Ліндеберга. Обернене твердження неправильне.

Теорема Ліндеберга–Леві. *Якщо попарно незалежні випадкові величини* $X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$ *однаково розподілені і мають математичне сподівання а і дисперсію* σ^2*, то*

$$\lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - na}{\sigma\sqrt{n}} < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt.$$

Центральна гранична теорема пояснює достатньо широке застосування нормального закону розподілу: *якщо випадкова величина формується під впливом багатьох незалежних факторів, кожен з яких здійснює на неї незначний вплив, то розподіл цієї величини близький до нормального.*

Розділ 6. Багатовимірні випадкові величини

У попередніх розділах ми розглядали випадкові величини (дискретні та неперервні), їх закони розподілу та числові характеристики. У прийнятому теоретико-множинному трактуванні будь-яка випадкова величина X є функцією елементарної події, яка входить у простір елементарних подій. Але в прикладних задачах зазвичай доводиться розглядати не одну, а декілька випадкових величин, які одночасно вимірюються (спостерігаються) в одному й тому ж експерименті. При цьому з кожною елементарною подією пов'язана система (сукупність) двох чи більше випадкових величин. Наприклад, успішність студента характеризується системою n випадкових величин, $X_1, X_2, ..., X_n$ – оцінки, отримані ним у семестрі. Такі питання, а також функції випадкових аргументів розглянемо в цьому розділі.

6.1. Функція одного випадкового аргумента, її розподіл та числові характеристики

Означення. Якщо кожному можливому значенню випадкової величини X відповідає одне можливе значення випадкової величини Y, то Y називають ***функцією випадкового аргумента X***:

$$Y = \varphi(X).$$

Розглядають такі функції як дискретні, так і неперервні, на чому і зупинимось далі.

Нехай X – дискретна випадкова величина. Якщо різним можливим значенням випадкової величини X відповідають різні значення функції Y, то ймовірності відповідних значень X і Y різні. Якщо різним можливим значенням випадкової величини X відповідають значення Y, серед яких є рівні між собою, то ймовірності рівних значень Y додаються.

Нехай дискретна випадкова величина X задана рядом розподілу

X	x_1	x_2	$\dots$	x_n	$\dots$
P	p_1	p_2	$\dots$	p_n	$\dots$

і $Y = \varphi(X)$.

Тоді справедливі такі формули:

$$M\left(\varphi(X)\right)=\sum_k \varphi(x_k)p_k; \tag{6.1}$$

$$D\left(\varphi(X)\right)=\sum_k \varphi^2(x_k)p_k-\left(M(\varphi(X))\right)^2. \tag{6.2}$$

Нехай X – неперервна випадкова величина з щільністю розподілу $f(x)$, а $Y=\varphi(X)$. Якщо $y=\varphi(x)$ – диференційовна, строго зростаюча або строго спадна функція, обернена до якої має вигляд $x=\psi(y)$, то щільність розподілу $g(y)$ випадкової величини $Y=\varphi(X)$ обчислюють за формулою

$$g(y)=f(\psi(y))\cdot\left|\psi'(y)\right|. \tag{6.3}$$

Якщо функція $y=\varphi(x)$ в інтервалі можливих значень X не монотонна, то потрібно розбити цей інтервал на такі інтервали, на яких $\varphi(x)$ монотонна, знайти щільності розподілів $g_i(y)$ для кожного з інтервалів монотонності, а потім подати $g(y)$ у вигляді суми

$$g(y)=\sum_i g_i(y).$$

Числові характеристики випадкової величини Y обчислюють за формулами:

$$M\left((\varphi(X)\right)=\int_{-\infty}^{+\infty}\varphi(x)f(x)dx;$$

$$D\left(\varphi(X)\right)=\int_{-\infty}^{+\infty}\varphi^2(x)f(x)dx-\left(M(\varphi(X))\right)^2.$$

Зокрема, якщо можливі значення X належать інтервалу (a,b), то

$$M\left(\varphi(X)\right)=\int_a^b\varphi(x)f(x)dx;$$

$$D\left(\varphi(X)\right)=\int_a^b\varphi^2(x)f(x)dx-\left(M(\varphi(X))\right)^2.$$

Приклад 1. Дана дискретна випадкова величина X

X	0	$\frac{\pi}{6}$	$\frac{\pi}{2}$
P	0,3	0,5	0,2

та $Y=\sin X$. Знайти $M(Y),\ D(Y),\ \sigma(Y)$.

Розв'язання. Знаходимо розподіл випадкової величини Y:

$$y_1 = \sin 0 = 0,\ y_2 = \sin\frac{\pi}{6} = \frac{1}{2},\ y_3 = \sin\frac{\pi}{2} = 1.$$

Y	0	$\frac{1}{2}$	1
P	0,3	0,5	0,2

$$M(Y) = 0\cdot 0{,}3 + 0{,}5\cdot 0{,}5 + 1\cdot 0{,}2 = 0{,}45;$$

$$M(Y^2) = 0\cdot 0{,}3 + 0{,}25\cdot 0{,}5 + 1\cdot 0{,}2 = 0{,}325;$$

$$D(Y) = M(Y^2) - (M(Y))^2 = 0{,}325 - (0{,}45)^2 = 0{,}1225;$$

$$\sigma(Y) = \sqrt{D(Y)} = \sqrt{0{,}1225} = 0{,}35.$$

Приклад 2. Випадкова величина X задана щільністю розподілу

$$f(x) = \begin{cases} 0, & x \le 0 \ \text{àáî} \ \ x > 1; \\ 2x, & 0 < x \le 1; \end{cases} \quad Y = X^2.$$

Знайти $M(Y), D(Y)$.

Розв'язання.

$$M(Y) = \int_0^1 x^2\cdot 2x dx = 2\int_0^1 x^3 dx = \left.\frac{x^4}{2}\right|_0^1 = \frac{1}{2};$$

$$M(Y^2) = \int_0^1 x^4\cdot 2x dx = 2\int_0^1 x^5 dx = \left.\frac{x^6}{3}\right|_0^1 = \frac{1}{3};$$

$$D(Y) = M(Y^2) - (M(Y))^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

Приклад 3. Задана щільність розподілу $f(x)$ випадкової величини X, можливі значення якої належать інтервалу $(0, +\infty)$. Знайти щільність розподілу випадкової величини $Y = \ln X$.

Розв'язання. Функція $y = \ln x$ монотонна в інтервалі $(0, +\infty)$, отже має обернену функцію $x = e^y$, звідки $x' = e^y$. Тоді за формулою (6.3) отримаємо:

$$g(y) = e^y f(e^y),\ y \in (-\infty,\ +\infty).$$

Приклад 4. Дискретна випадкова величина X задана рядом розподілу

X	–1	1	2
P	0,1	0,5	0,4

Знайти розподіл функції $Y = X^2$.

Розв'язання. Можливі значення Y: $y_1 = 1,\ y_2 = 1,\ y_3 = 4$. Значення $y_1 = y_2$ рівні між собою. Ймовірність цього значення $p_1 = 0{,}1 + 0{,}5 = 0{,}6$. Одержимо розподіл Y:

Y	1	4
P	0,6	0,4

6.2. Функція двох випадкових аргументів

Якщо кожній парі можливих значень випадкових величин X і Y відповідає одне можливе значення випадкової величини Z, то Z називають ***функцією двох випадкових величин*** X і Y:

$$Z = \varphi(X,\ Y).$$

Закон розподілу суми $Z = X + Y$ ***за відомими розподілами доданків***

1. Нехай X і Y – дискретні незалежні випадкові величини. Можливі значення $Z = X + Y$ дорівнюють сумам кожного можливого значення X та кожного можливого значення Y; імовірності можливих значень дорівнюють добуткам імовірностей доданків.

2. Якщо X і Y – неперервні незалежні випадкові величини, які мають відповідно щільності розподілу $f_1(x)$ і $f_2(y)$, то щільність розподілу $g(z)$ суми $Z = X + Y$ за умови, що щільність розподілу принаймні одного з аргументів задана в інтервалі $(-\infty, +\infty)$ можна знайти за формулою

$$g(z) = \int_{-\infty}^{+\infty} f_1(x) f_2(z - x) dx$$

або за рівносильною формулою

$$g(z) = \int_{-\infty}^{+\infty} f_1(z - y) f_2(y) dy.$$

Функція $g(z)$, утворена з функцій $f_1(x)$ і $f_2(y)$ за наведеними формулами, називається ***згорткою*** або ***композицією*** цих функцій.

Якщо можливі значення аргументів не від'ємні, то щільність розподілу суми $Z = X + Y$ обчислюють за формулою

$$g(z) = \int_0^z f_1(x) f_2(z-x)dx$$

або за рівносильною формулою

$$g(z) = \int_0^z f_1(z-y) f_2(y)dy.$$

Зауваження. Сума незалежних нормально розподілених випадкових величин також має нормальний розподіл. Математичне сподівання і дисперсія цієї суми відповідно дорівнюють сумам математичних сподівань і дисперсій доданків.

Приклад 5. Для незалежних випадкових величин X і Y відомі щільності їх розподілів

$$f_1(x) = \frac{1}{3}e^{-\frac{x}{3}} \ (0 \le x < \infty), \ f_2(y) = \frac{1}{5}e^{-\frac{y}{5}} \ (0 \le y < \infty).$$

Знайти щільність розподілу випадкової величини $Z = X + Y$.

Розв'язання. За формулою $g(z) = \int_0^z f_1(x) f_2(z-x)dx$ обчислимо:

$$g(z) = \int_0^z \frac{1}{3}e^{-\frac{x}{3}} \cdot \frac{1}{5}e^{-\frac{(z-x)}{5}}dx = \frac{1}{15}e^{-\frac{z}{5}}\int_0^z e^{-\frac{2x}{15}}dx = \frac{1}{15}e^{-\frac{z}{5}}\left(-\frac{15}{2}e^{-\frac{2x}{15}}\right)\Bigg|_0^z = \frac{1}{2}e^{-\frac{z}{5}}\left(1 - e^{-\frac{2z}{15}}\right).$$

Оскільки можливі значення X і Y невід'ємні і $Z = X + Y$, то $z \ge 0$. Отже,

$$g(z) = \begin{cases} \frac{1}{2}e^{-\frac{z}{5}}\left(1 - e^{-\frac{2z}{15}}\right), \ 0 \le z < +\infty; \\ 0, \ -\infty < z < 0. \end{cases}$$

6.3. Двовимірні випадкові величини

Випадкові величини, які входять у систему (X_1, X_2, …, X_n), можуть бути як дискретними, так і недискретними (неперервними або мішаними).

Приклад 6. Експеримент полягає у підкиданні двох монет. Елементарною подією буде випадання герба і положення монет одна відносно одної. Нехай випадкова величина X – кількість гербів, які випали, а випадкова величина Y – відстань між монетами. У цьому разі X – дискретна випадкова величина, яка набуває значень 0, 1, 2, а Y – неперервна випадкова величина із значеннями в R.

Для наочності дослідження n-вимірних випадкових величин зручно користуватися геометричною інтерпретацією. Так, систему двох випадкових величин (X, Y) можна зобразити *випадковою точкою* на площині Oxy з координатами X та Y, або, що рівносильно, *випадковим вектором*, напрямленим з початку координат у точку (X, Y). Аналогічно, систему трьох випадкових величин (X, Y, Z) можна зобразити в тривимірному просторі як випадковий вектор, напрямлений з початку координат у точку (X, Y, Z).

Якщо вимірність величини більше від 3, то геометрична інтерпретація втрачає наочність, але користуватися геометричною термінологією зручно. Так, про систему випадкових величин $(X_1, X_2, \ldots, X_n)$ будемо казати як про випадкову точку в n-вимірному просторі R^n або як про випадковий вектор, напрямлений з початку координат у точку $(X_1, X_2, \ldots, X_n)$: $\overrightarrow{X} = (X_1, X_2, \ldots, X_n)$.

На n-вимірні випадкові величини поширюються майже без змін основні означення, які було розглянуто для одновимірної випадкової величини.

Очевидно, що властивості n-вимірної випадкової величини не обмежуються властивостями окремих складових випадкових величин; суттєві також зв’язки між складовими величинами.

Повна характеристика n-вимірної випадкової величини – її закон розподілу, що як і для окремих випадкових величин може мати різні форми: функція розподілу, щільність розподілу, таблиця ймовірностей окремих значень випадкового вектора тощо.

Для практики важливий випадок $n = 2$, яким ми обмежимося в подальшому.

Функцією розподілу ймовірностей (або сумісною функцією розподілу) двовимірної випадкової величини (X, Y) називається така функція двох аргументів x, y, яка визначає ймовірність сумісного настання подій

$$F(x, y) = P\left(X < x,\ Y < y\right).$$

Подія в дужках означає добуток подій: $(X < x, Y < y) = (X < x) \cap (Y < y)$.

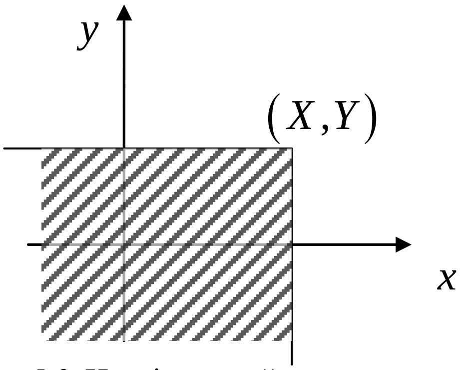


Рис. 5.3. Нескінченний прямокутник із вершиною в точці (*X, Y*)

Користуючись геометричною інтерпретацією двовимірної випадкової величини (*X, Y*) як випадкової точки на площині, можна дати геометричне тлумачення функції розподілу: це ймовірність влучення в нескінченний прямокутник із вершиною в точці (*X, Y*) (рис. 5.3), який лежить лівіше і нижче від неї.

Властивості функції розподілу двовимірної випадкової величини (*X, Y*) аналогічні властивостям одновимірної випадкової величини:

1. $0 \le F(x, y) \le 1.$

2. Неперервність зліва за кожним аргументом:

$$\lim_{x \to x_1 - 0} F(x, y) = F(x_1, y); \quad \lim_{y \to y_1 - 0} F(x, y) = F(x, y_1).$$

3. Функція $F(x, y)$ – неспадна функція за кожним аргументом:

$F(x_1, y) \le F(x_2, y)$, якщо $x_1 < x_2$ та $F(x, y_1) \le F(x, y_2)$, якщо $y_1 < y_2$.

4. $F(x, +\infty) = F_1(x)$ – функція розподілу компонента *X*, $F(+\infty, y) = F_2(y)$ – функція розподілу компонента Y (властивість встановлює природний зв'язок між функцією розподілу двовимірної випадкової величини і функціями розподілу $F_1(x)$, $F_2(y)$, які також називають частинними функціями одновимірних випадкових величин X і *Y*.

5. $F(-\infty, y) = 0$, $F(x, -\infty) = 0$, $F(-\infty, -\infty) = 0$, $F(+\infty, +\infty) = 1$.

6. Ймовірність потрапляння випадкової точки до прямокутника $(x_1 \le X \le x_2;\ y_1 \le Y \le y_2)$ обчислюють за формулою

$$P(x_1 < X < x_2;\ y_1 < Y < y_2) = (F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)).$$

6.3.1. Дискретні двовимірні випадкові величини

Означення. Двовимірну випадкову величину називають **дискретною,** якщо множина значень, яких вона може набути, скінченна або зліченна.

Означення. **Законом розподілу дискретної двовимірної випадкової величини** називають перелік можливих значень цієї величини (x_i, y_k) та їх імовірностей $p(x_i, y_k)$, $i=1, 2, \ldots, m$; $k=1, 2, \ldots, n$.

Закон розподілу такої випадкової величини задають таблицею:

$Y \backslash X$	x_1	x_2	…	x_m	$\sum$
y_1	p_{11}	p_{21}	…	p_{m1}	q_1
y_2	p_{12}	p_{22}	…	p_{m2}	q_2
…	…	…	…	…	…
y_n	p_{1n}	p_{2n}	…	p_{mn}	q_n
$\sum$	p_1	p_2	…	p_m	1

Причому $q_1=\sum_{i=1}^{m} p_{i1},\ q_2=\sum_{i=1}^{m} p_{i2},\ \ldots, q_n=\sum_{i=1}^{m} p_{in};\ \ p_1=\sum_{i=1}^{n} p_{1i},\ p_2=\sum_{i=1}^{n} p_{2i},\ \ldots,\ p_m=\sum_{i=1}^{n} p_{mi}.$

Події $(X=x_i,\ Y=y_k)$, $i=1, 2, \ldots, m$; $k=1, 2, \ldots, n$ утворюють повну групу, тому сума ймовірностей наведеної таблиці дорівнює 1. Додаючи ймовірності стовпчиків і рядків, одержують закони розподілу компонентів X і Y:

X	x_1	x_2	...	x_m
P	p_1	p_2	...	p_m

Y	y_1	y_2	...	y_n
P	q_1	q_2	...	q_n

6.3.2. Неперервні двовимірні випадкові величини

Означення. Двовимірну випадкову величину (X, Y) називають **неперервною,** якщо вона набуває значення в усіх точках деякої області площини Oxy, що рівносильно існуванню такої функції $f(x, y)$, за якої функцію розподілу $F(x, y)$ можна подати у вигляді

$$F(x,\ y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(x,\ y)dxdy.$$

Означення. Функцію $f(x, y)$ називають **щільністю розподілу ймовірностей** *двовимірної випадкової величини* (X, Y).

Якщо функція $f(x, y)$ неперервна, то

$$f(x,\ y) = \frac{\partial^2 F(x,\ y)}{\partial x \partial y}.$$

Властивості функції щільності розподілу ймовірностей:

1. $f(x,\ y) \geq 0$ за всіх x, y.

2. $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,\ y)dxdy = 1.$

3. Якщо D – будь-яка область у площині Oxy, то

$$P\big((X,\ Y) \in D\big) = \iint\limits_D f(x,\ y)dxdy.$$

Знаючи щільність розподілу $f(x, y)$ двовимірної випадкової величини $(X,\ Y)$, знаходять щільності розподілу $f_1(x)$ і $f_2(x)$ для її компонентів *X* та *Y*. Справді, функція розподілу випадкової величини X

$$F_1(x) = F(x,\ +\infty) = \int\limits_{-\infty}^{x} dx \int\limits_{-\infty}^{+\infty} f(x,\ y)dy.$$

Диференціюючи обидві частини рівності за *x*, отримаємо:

$$f_1(x) = \frac{\partial F_1(x)}{\partial x} = \int\limits_{-\infty}^{+\infty} f(x,\ y)dy.$$

Аналогічно, $F_2(y) = F(+\infty, y) = \int\limits_{-\infty}^{y} dy \int\limits_{-\infty}^{+\infty} f(x,\ y)dx$ та

$$f_2(y) = \int\limits_{-\infty}^{+\infty} f(x,\ y)dx.$$

Теорема. *Випадкові величини X та Y незалежні тоді і тільки тоді, коли функція розподілу пари* (X, Y) *дорівнює добутку функцій розподілу компонентів:*

$$F(x, y) = F_1(x) \cdot F_2(y).$$

Наслідок. Неперервні випадкові величини X і Y незалежні тоді і тільки тоді, коли щільність розподілу ймовірностей пари (X, Y) дорівнює добутку щільностей розподілу ймовірностей компонентів:

$$f(x,\ y) = f_1(x) \cdot f_2(y).$$

Приклад 7. Задано дискретну двовимірну випадкову величину:

$Y \backslash X$	1	2	3	$\sum$
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6
5	0,2	0	0	0,2
$\sum$	0,3	0,3	0,4	1

Знайти закони розподілу компонент X і Y.

Розв'язання. Додаючи стовпчики і рядки таблиці ймовірностей, отримують відповідні закони розподілу випадкових величин X і Y:

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

Y	0	2	5
P	0,2	0,6	0,2

Приклад 8. Функцію розподілу двовимірної випадкової величини (X, Y) задано формулою:

$$F(x,\ y)=\begin{cases}(1-e^{-4x})(1-e^{-2y}),\ x\ge 0,\ y\ge 0;\\ 0,\ x<0\ \text{ àáî }\ y<0.\end{cases}$$

Знайти щільність розподілу ймовірностей $f(x, y)$.

Розв'язання. Використовуємо формулу $f(x, y)=\dfrac{\partial^2 F(x, y)}{\partial x\partial y}$, тому обчислюємо $\dfrac{\partial F(x)}{\partial x}=\begin{cases}4e^{-4x}(1-e^{-2y}),\ x\ge 0,\ y\ge 0;\\ 0,\ x<0\ \text{ àáî }\ y<0,\end{cases}$ тоді

$$\frac{\partial^2 F(x,\ y)}{\partial x\partial y}=f(x,\ y)=\begin{cases}8e^{-4x}e^{-2y},\ x\ge 0,\ y\ge 0;\\ 0,\ x<0\ \text{ àáî }\ y<0.\end{cases}$$

Приклад 9. За відомою щільністю розподілу двовимірної випадкової величини $(X,\ Y)$

$$f(x,\ y)=\frac{20}{\pi^2(16+x^2)(25+y^2)}$$

знайти її функцію розподілу $F(x, y)$.

Розв'язання. Обчислення проводимо згідно з формулою

$$F(x,\ y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(x,\ y)dxdy.$$

Отже, $F(x,\ y)=\int_{-\infty}^{x}\int_{-\infty}^{y}\frac{20}{\pi^2(16+x^2)(25+y^2)}dxdy=\frac{20}{\pi^2}\int_{-\infty}^{x}\frac{dx}{4^2+x^2}\int_{-\infty}^{y}\frac{dy}{5^2+y^2}=$

$$=\frac{20}{\pi^2}\lim_{a\to-\infty}\left(\frac{1}{4}\operatorname{arctg}\frac{x}{4}\right)\Bigg|_a^x\cdot\lim_{b\to-\infty}\left(\frac{1}{5}\operatorname{arctg}\frac{y}{5}\right)\Bigg|_b^y=$$

$$=\frac{1}{\pi^2}\left(\operatorname{arctg}\frac{x}{4}-\lim_{a\to-\infty}\operatorname{arctg}\frac{a}{4}\right)\left(\operatorname{arctg}\frac{y}{5}-\lim_{b\to-\infty}\operatorname{arctg}\frac{b}{5}\right)=$$

$$=\frac{1}{\pi^2}\left(\operatorname{arctg}\frac{x}{4}+\frac{\pi}{2}\right)\left(\operatorname{arctg}\frac{y}{5}+\frac{\pi}{2}\right)=\left(\frac{1}{\pi}\operatorname{arctg}\frac{x}{4}+\frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}\frac{y}{5}+\frac{1}{2}\right).$$

Таким чином, функція розподілу має вигляд

$$F(x,y)=\left(\frac{1}{\pi}\operatorname{arctg}\frac{x}{4}+\frac{1}{2}\right)\left(\frac{1}{\pi}\operatorname{arctg}\frac{y}{5}+\frac{1}{2}\right).$$

6.3.3. Числові характеристики двовимірної випадкової величини. Коефіцієнт кореляції та його властивості

Нехай (X,Y) – дискретна двовимірна випадкова величина, для якої закон розподілу задано таблицею розподілу. Тоді, додаючи стовпчики і рядки таблиці ймовірностей, знаходять закони розподілу компонентів X і Y. Відповідно обчислюють їх числові характеристики.

Згідно з даними прикладу 1 п. 6.3.2 можемо визначити числові характеристики кожного компонента:

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

$$D(X)=M(X^2)-(M(X))^2=1\cdot 0{,}3+4\cdot 0{,}4+9\cdot 0{,}3-4=4{,}6-4=0{,}6;$$

$$\sigma(X)=\sqrt{D(X)}=\sqrt{0{,}6}=0{,}7746.$$

Аналогічні обчислення для компонента Y:

$$M(Y)=1\cdot 0{,}5+2\cdot 0{,}5=1{,}5;$$

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

$$\sigma(Y)=\sqrt{D(Y)}=\sqrt{0{,}25}=0{,}5.$$

Нехай (X, Y) – неперервна двовимірна випадкова величина з щільністю розподілу $f(x, y)$. Якщо $f_1(x)$ і $f_2(y)$ – щільності розподілу її компонентів X і Y, то за означенням математичного сподівання одновимірної випадкової величини

$$M(X)=\int_{-\infty}^{+\infty} xf_1(x)dx, \quad M(Y)=\int_{-\infty}^{+\infty} yf_2(y)dy.$$

Підставивши $f_1(x)=\int_{-\infty}^{+\infty} f(x,y)dy, \quad f_2(y)=\int_{-\infty}^{+\infty} f(x,y)dx$ отримаємо:

$$M(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x, y)dxdy, \quad M(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x, y)dxdy.$$

Якщо $f(x,y)$ набуває значення в деякій області D площини Oxy, то

$$M(X)=\iint_D xf(x, y)dxdy, \quad M(Y)=\iint_D yf(x, y)dxdy.$$

Означення. **Коваріацією** *(кореляційним моментом)* випадкових величин X і Y називається математичне сподівання добутку різниць випадкових величин і їх математичних сподівань:

$$K(X,Y)=\operatorname{cov}(X,Y)=M\left(\left(X-M(X)\right)\left(Y-M(Y)\right)\right).$$

Після розкриття дужок

$$K(X,Y)=M(XY)-M(X)M(Y).$$

З останньої формули випливає, що для дискретної двовимірної випадкової величини (X,Y)

$$K(X,Y)=\sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j p_{ij} - \sum_{i=1}^{m} x_i p_i \sum_{j=1}^{n} y_j q_j$$

та для неперервної двовимірної випадкової величини (X,Y)

$$K(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x, y)dxdy-\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x, y)dxdy\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x, y)dxdy.$$

Якщо $f(x, y)$ набуває значення в деякій області D площини Oxy, то

$$K(X,\ Y)=\iint_D xyf(x, y)dxdy-\iint_D xf(x, y)dxdy\iint_D yf(x, y)dxdy.$$

Означення. **Дисперсією** *(дисперсійною матрицею) двовимірної випадкової* *величини* (X, Y) *називається сукупність чотирьох чисел* d_{ik}, $(d_{ik} = d_{ki})$, яка є матрицею другого порядку:

$$D(X,Y)=\begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}.$$

Тоді, у разі двовимірної неперервної випадкової величини для елементів матриці

$$d_{11}=D(X)=\int_{-\infty}^{+\infty}(x-M(X))^2 f_1(x)dx=\int_{-\infty}^{+\infty}x^2 f_1(x)dx-(M(X))^2,$$

$$d_{22}=D(Y)=\int_{-\infty}^{+\infty}(y-M(Y))^2 f_1(y)dy=\int_{-\infty}^{+\infty}y^2 f_1(y)dy-(M(Y))^2,$$

$$d_{12}=d_{21}=K(X,Y),$$

або

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

$$D(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(y-M(Y))^2 f(x,y)dxdy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2 f(x,y)dxdy-(M(Y))^2.$$

Якщо $f(x, y)$ набуває значення в деякій області D площини O_{xy}, то

$$D(X)=\iint_D (x-M(X))^2 f(x,y)dxdy=\iint_D x^2 f(x,y)dxdy-(M(X))^2,$$

$$D(Y)=\iint_D (y-M(Y))^2 f(x,y)dxdy=\iint_D y^2 f(x,y)dxdy-(M(Y))^2.$$

Точка $(M(X), M(Y))$ називається *центром розсіювання* двовимірної випадкової величини.

Якщо випадкові величини X і Y незалежні, то $K(X,Y)=0$ і дисперсійна матриця є діагональною матрицею:

$$D(X,Y)=\begin{pmatrix} D(X) & 0 \\ 0 & D(Y) \end{pmatrix}.$$

Звідси випливає, що коваріація є характеристикою *зв'язку* між величинами системи (X,Y). Коваріація описує не тільки зв'язок між X і Y, а також і їх розсіювання. Дійсно, якщо принаймні одна з величин мало відрізняється від свого математичного сподівання, то коваріація буде малою, як би не були зв'язані між собою випадкові величини X і Y.

Істотний недолік коваріації – те, що її розмірність збігається з добутком розмірностей випадкових величин, тому замість коваріації використовують безрозмірну величину – *коефіцієнт кореляції*, який характеризує *тісноту зв'язку* між величинами X і Y.

Означення. **Коефіцієнтом кореляції** між випадковими величинами X і Y називається число

$$\rho = \rho(X,Y) = \frac{K(X,Y)}{\sigma(X)\sigma(Y)}.$$

Властивості коефіцієнта кореляції

Теорема 1. *Для будь-яких випадкових величин* $-1 \le \rho(X,Y) \le 1.$

Теорема 2. *Коефіцієнт кореляції між випадковими величинами* X *і* Y *дорівнює* ± 1 *тоді і тільки тоді, коли* X *і* Y *зв'язані лінійною залежністю* $Y = aX + b$, *причому* $\rho(X,Y) = 1$ *за* $a > 0$, $\rho(X,Y) = -1$ *за* $a < 0$.

Теорема 3. *Якщо випадкові величини* X *і* Y *незалежні, то* $\rho(X,Y) = 0.$

Наслідок. Якщо $\rho(X,Y) \ne 0,$ то X і Y залежні випадкові величини.

Означення. Випадкові величини X і Y, для яких $\rho(X,Y) = 0,$ називаються **некорельованими,** для яких $\rho(X, Y) \ne 0$ називаються **корельованими**.

Зауваження. З некорельованості двох випадкових величин, взагалі кажучи, не випливає їх незалежності. Для нормально розподілених випадкових величин некорельованість рівносильна незалежності.

Коефіцієнт кореляції служить для оцінювання тісноти лінійного зв'язку між випадковими величинами X і Y: що ближче модуль коефіцієнта кореляції до одиниці, то зв'язок сильніший, що ближче модуль коефіцієнта кореляції до нуля, то зв'язок слабший.

Приклад 10. Неперервна двовимірна випадкова величина (X, Y) задана щільністю розподілу

$$f(x, y) = \begin{cases} Axy, (x, y) \in D; \\ 0, (x, y) \notin D, \end{cases}$$

де область D обмежена прямими $x + y - 1 = 0,\ x = 0,\ y = 0$ (рис. 6.1). Знайти числа $A,\ M(X),\ M(Y),\ D(X),\ D(Y),\ K(X,Y),\ \rho(X,Y).$

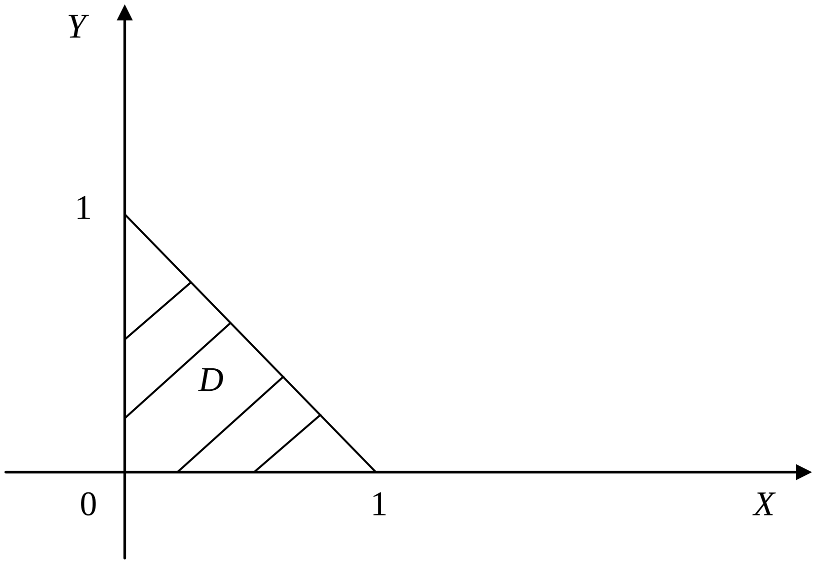


Рис. 6.1. Задання області *D*

Розв'язання. Оскільки $f(x,\, y)$ набуває значення в області *D*, то

$$\iint_D f(x,y)dxdy = 1.$$

Одержимо $\iint_D Axydxdy = A\int_0^1 xdx\int_0^{1-x} ydy = A\int_0^1 x\left.\frac{y^2}{2}\right|_0^{1-x} dx = \frac{A}{2}\int_0^1 x(1-x)^2 dx =$

$$= \frac{A}{2}\int_0^1 (x-2x^2+x^3)dx = \frac{A}{2}\left.\left(\frac{x^2}{2}-\frac{2x^3}{3}+\frac{x^4}{4}\right)\right|_0^1 = \frac{A}{24};\ \frac{A}{24}=1 \text{ і звідси } A=24.$$

Отже,

$$f(x,\, y) = \begin{cases} 24\, xy,\ (x,\, y)\in D; \\ 0,\ (x,\, y)\notin D. \end{cases}$$

Визначимо числові характеристики:

$$M(X) = \iint_D xf(x,\, y)dxdy = 24\int_0^1 x^2dx\int_0^{1-x} ydy = 24\int_0^1 x^2\left.\frac{y^2}{2}\right|_0^{1-x} = 12\int_0^1 x^2(1-x)^2dx =$$

$$= 12\int_0^1 (x^2-2x^3+x^4)dx = 12\left.\left(\frac{x^3}{3}-\frac{x^4}{2}+\frac{x^5}{5}\right)\right|_0^1 = \frac{2}{5};$$

$$M(Y) = \iint_D yf(x,\, y)dxdy = 24\int_0^x y^2dy\int_0^{1-y} xdx = \frac{2}{5};$$

$$M(X^2)=\iint_D x^2 f(x,y)dxdy=24\int_0^1 x^3dx\int_0^{1-x} ydy=24\int_0^1 x^3\left.\frac{y^2}{2}\right|_0^{1-x}dx=12\int_0^1 x^3(1-x)^2dx=$$

$$=12\int_0^1 (x^3-2x^4+x^5)dx=12\left.\left(\frac{x^4}{4}-\frac{2x^5}{5}+\frac{x^6}{6}\right)\right|_0^1=\frac{1}{5};$$

$$M(Y^2)=\iint_D y^2 f(x,y)dxdy=24\int_0^1 y^3dy\int_0^{1-y} xdx=\frac{1}{5};$$

$$D(X)=M(X^2)-(M(X))^2=\frac{1}{5}-\frac{4}{25}=\frac{1}{25};\ \sigma(X)=\frac{1}{5};$$

$$D(Y)=M(Y^2)-(M(Y))^2=\frac{1}{5}-\frac{4}{25}=\frac{1}{25};\ \sigma(Y)=\frac{1}{5};$$

$$M(XY)=\iint_D xyf(x,y)dxdy=24\int_0^1 x^2dx\int_0^{1-x} y^2dy=24\int_0^1 x^2\left.\frac{y^3}{3}\right|_0^{1-x}dx=8\int_0^1 x^2(1-x)^3dx=$$

$$=8\int_0^1 (x^2-3x^3+3x^4-x^5)dx=8\left.\left(\frac{x^3}{3}-\frac{3x^4}{4}+\frac{3x^5}{5}-\frac{x^6}{6}\right)\right|_0^1=\frac{2}{15};$$

$$K(X,Y)=M(XY)-M(X)M(Y)=\frac{2}{15}-\frac{2}{5}\cdot\frac{2}{5}=-\frac{2}{75};$$

$$\rho(X,Y)=\frac{K(X,Y)}{\sigma(X)\sigma(Y)}=-\frac{2}{3}.$$

6.3.4. Лінійна регресія

Означення. Пряма $y=\alpha_1x+\alpha_0$ називається *прямою регресії* (або прямою середньоквадратичної регресії) Y на X, якщо коефіцієнти α_1 і α_0 вибрано так, щоб середнє квадратичне відхилення $\alpha_1X+\alpha_0$ від Y було мінімальним:

$$M(Y-\alpha_1X-\alpha_0)^2=\min.$$

Аналогічно визначається пряма регресії X на Y.

Тоді коефіцієнти α_1 і α_0 набувають таких значень:

$$\alpha_1=\rho\frac{\sigma_y}{\sigma_x},\ \alpha_0=M(Y)-\alpha_1M(X).$$

Коефіцієнт α_1 називається коефіцієнтом регресії Y на X. Рівняння прямої регресії Y на X має вигляд

$$y=\rho\frac{\sigma_y}{\sigma_x}x+M(Y)-\rho\frac{\sigma_y}{\sigma_x}M(X)$$

або

$$y-M(Y)=\rho\frac{\sigma_y}{\sigma_x}(x-M(X)).$$

Аналогічно пряма регресії X на Y має вигляд

$$x-M(X)=\rho\frac{\sigma_x}{\sigma_y}(y-M(Y)).$$

Обидві прямі регресії проходять через точку $\left(M(X),M(Y)\right)$ – центр розсіювання двовимірної випадкової величини. Прямі регресії збігаються тоді і тільки тоді, коли $\rho=\pm1$.

Величина $(1-\rho^2)\sigma_y^2$ називається залишковою регресією Y на X. Вона характеризує величину похибки у разі заміни Y на $\alpha_1 X+\alpha_0$.

Приклад 11. Двовимірна дискретна випадкова величина $\left(X,Y\right)$ задана таблицею розподілу

$Y \backslash X$	1	2	4	$\sum$
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6
5	0,2	0	0	0,2
$\sum$	0,3	0,3	0,4	1

Знайти рівняння прямих регресії Y на X та X на Y.

Розв'язання. Додаючи стовпчики і рядки таблиці ймовірностей, знаходимо закони розподілу випадкових величин X і Y:

X	1	2	4
p	0,3	0,3	0,4

Y	0	2	5
p	0,2	0,6	0,2

Тоді

$$M(X)=1\cdot 0,3+2\cdot 0,3+4\cdot 0,4=2,5;$$

$$D(X)=M(X^2)-(M(X))^2=1\cdot 0,3+4\cdot 0,3+16\cdot 0,4-2,5^2=7,9-6,25=1,65;$$

$$\sigma_x=\sqrt{D(X)}=\sqrt{1,65}\cong 1,2845;$$

$$M(Y)=0\cdot 0,2+2\cdot 0,6+5\cdot 0,2=2,2;$$

$$D(Y)=M(Y^2)-(M(Y))^2=4\cdot 0,6+25\cdot 0,2-2,2^2=7,4-4,84=2,56;$$

$$\sigma_y=\sqrt{D(Y)}=\sqrt{2,56}=1,6;$$

$$M(XY)=1\cdot 0\cdot 0,1+2\cdot 0\cdot 0+4\cdot 0\cdot 0,1+1\cdot 2\cdot 0+2\cdot 2\cdot 0,3+4\cdot 2\cdot 0,3+$$
$$+1\cdot 5\cdot 0,2+2\cdot 5\cdot 0+4\cdot 5\cdot 0=4\cdot 0,3+8\cdot 0,3+5\cdot 0,2=4,6;$$

$$K(X,Y)=M(XY)-M(X)M(Y)=4,6-2,5\cdot 2,2=-0,9;$$

$$\rho(X,Y)=\frac{K(X,Y)}{\sigma_x\sigma_y}=\frac{-0,9}{1,2845\cdot 1,6}=-0,438;$$

$$\rho\cdot\frac{\sigma_y}{\sigma_x}=-0,438\cdot\frac{1,6}{1,2845}=-0,5456;\ \ \rho\frac{\sigma_x}{\sigma_y}=-0,438\cdot\frac{1,2845}{1,6}=-0,3516.$$

Рівняння прямої регресії Y на X:

$y-2,2=-0,5456\,(x-2,5)$ (пряма 1 на рис. 6.2).

Рівняння прямої регресії X на Y:

$x-2,5=-0,3516\,(y-2,2)$ або $x=-0,3516y+3,2735$ (пряма 2 на рис. 6.2).

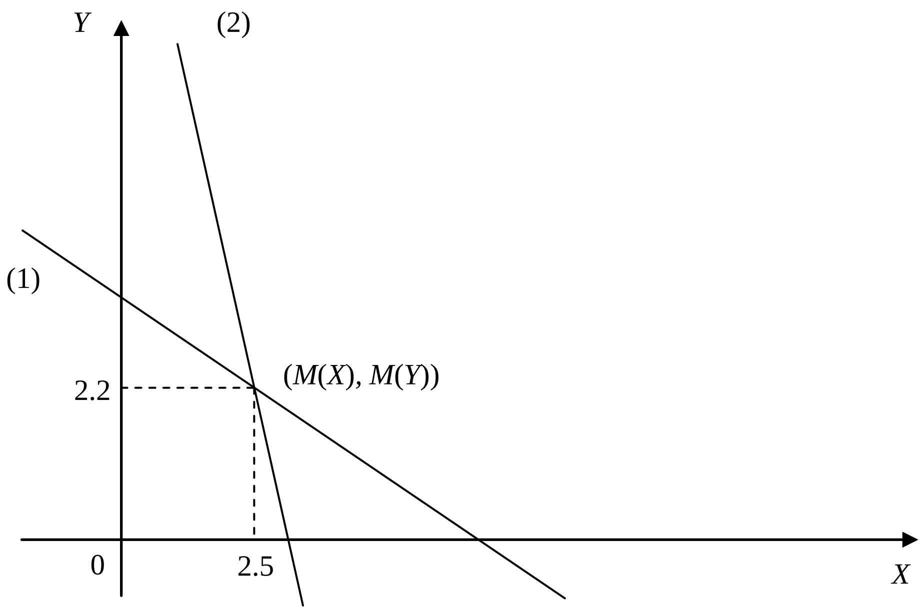


Рис. 6.2. Прямі регресії

6.4. Умовні закони розподілу

Поняття умовної ймовірності події розглядали в розділі 2 і ймовірність події A за умови, що відбулась подія B, обчислюється за формулою (2.1), якщо $P(B) \neq 0$.

6.4.1. Умовні закони розподілу компонентів дискретної двовимірної випадкової величини

Нехай (X, Y) – дискретна двовимірна випадкова величина, закон розподілу якої задають таблицею

$Y \backslash X$	x_1	x_2	$\dots$	x_m	$\sum$
y_1	p_{11}	p_{21}	$\dots$	p_{m1}	q_1
Y_2	p_{12}	p_{22}	$\dots$	p_{m2}	q_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_n	p_{1n}	p_{2n}	$\dots$	$p_{mò}$	q_n
$\sum$	p_1	p_2	$\dots$	p_m	1

Припустимо, що в результаті випробування величина Y набула значення $Y = y_j$, при цьому X набуде одного із своїх можливих значень $x_1,\ x_2, ..., x_m$. Позначимо умовну ймовірність того, що X набуде значення x_i за умови, що $Y = y_j$ через $p\left(x_i \mid y_j\right)$.

***Означення*. Умовним законом розподілу** дискретної випадкової величини X за фіксованого значення випадкової величини $Y = y_j$ називають множину можливих значень випадкової величини X та відповідних їм умовних імовірностей, обчислених за фіксованого значення $Y = y_j$, тобто розподіл

X	x_1	x_2	$\dots$	x_m
$p\left(X \mid y_j\right)$	$p\left(x_2 \mid y_j\right)$	$p\left(x_2 \mid y_j\right)$	$\cdots$	$p\left(x_m \mid y_j\right)$

Умовні ймовірності $p\left(x_i \mid y_j\right)$ обчислюють за формулою

$$p\left(x_i \mid y_j\right) = \frac{p_{ij}}{q_j} \quad (i = 1,\ 2,\ \dots,\ m,\ \ j = 1,\ 2,\ \dots,\ n).$$

Аналогічно визначається умовний закон розподілу компоненти Y за $X = x_i$

Y	y_1	y_2	$\dots$	y_n
$p(Y \mid x_i)$	$p(y_1 \mid x_i)$	$p(y_2 \mid x_i)$	$\dots$	$p(y_n \mid x_i)$

$$p\left(y_j \mid x_i\right) = \frac{p_{ij}}{p_i} \quad (i = 1,\ 2,\ \dots,\ m,\ j = 1,\ 2,\ \dots,\ n).$$

6.4.2. Умовні закони розподілу компонентів неперервної двовимірної випадкової величини

Нехай задана двовимірна неперервна випадкова величина (X, Y). Як і для двовимірної дискретної випадкової величини, для двовимірної неперервної випадкової величини розглядають умовні закони розподілу. Умовні закони розподілу для неперервних випадкових величин X, Y, що утворюють вектор (X, Y), визначаються умовними щільностями ймовірностей.

Означення. **Умовною щільністю** $\varphi(x \mid y)$ *розподілу компоненти* X *за даного значення* $Y = y$ називається відношення щільності розподілу двовимірної випадкової величини $f(x, y)$ до щільності розподілу $f_2(y)$ компоненти Y:

$$\varphi(x \mid y) = \frac{f(x,y)}{f_2(y)}.$$

Аналогічно визначається *умовна щільність компоненти* Y *за* $X = x$:

$$\psi(y \mid x) = \frac{f(x,y)}{f_1(x)}.$$

Приклад 12. Ця дискретна випадкова величина має розподіл, що визначається $(X,\ Y)$ таблицею розподілу

$Y \backslash X$	1	2	3	$\sum$
4	0,1	0,15	0,2	0,45
5	0,15	0,25	0,15	0,55
$\sum$	0,25	0,4	0,35	1

Знайти:

1) умовний закон розподілу випадкової величини X за $Y=4$;

2) умовний закон розподілу випадкової величини Y за $X=2$.

Розв'язання 1. Знаходимо

$$p\left(x_1\,|\,y_1\right)=\frac{p_{11}}{p_1}=\frac{0{,}10}{0{,}45}=\frac{2}{9};\quad p\left(x_2\,|\,y_1\right)=\frac{p_{21}}{p_1}=\frac{0{,}15}{0{,}45}=\frac{3}{9};$$

$$p\left(x_3\,|\,y_1\right)=\frac{p_{31}}{p_1}=\frac{0{,}20}{0{,}45}=\frac{4}{9}.$$

Отже,

X	1	2	3
$p\left(X\,|\,y_1\right)$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{4}{9}$

Розв'язання 2. Аналогічно обчислимо для випадкової величини Y за $X=2$.

$$p\left(y_1\,|\,x_2\right)=\frac{p_{21}}{q_2}=\frac{0{,}15}{0{,}4}=\frac{3}{8};\quad p\left(y_2\,|\,x_2\right)=\frac{p_{22}}{q_2}=\frac{0{,}25}{0{,}4}=\frac{5}{8}.$$

Отже,

Y	4	5
$p\left(Y\,|\,x_2\right)$	$\frac{3}{8}$	$\frac{5}{8}$

6.5. Приклади деяких важливих для практики двовимірних розподілів випадкових величин

Двовимірний біномний розподіл. Двовимірна випадкова величина $\left(X_1,\,X_2\right)$ має *двовимірний біномний розподіл* з параметрами $\left(n,\,p_1,\,p_2\right)$, $0<p_i<1,\ p_1+p_2=1$, якщо

$$P(X_1=k_1,\,X_2=k_2)=\frac{n!}{k_1!\cdot k_2!}\,p_1^{k_1}\cdot p_2^{k_2},\,k_1+k_2=n.$$

Двовимірний біномний розподіл – окремий випадок поліномного розподілу.

Поліномний розподіл – багатовимірний аналог біномного розподілу.

Двовимірний рівномірний розподіл в області *D*. *Двовимірний рівномірний розподіл* в області D визначається щільністю

$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{S_D}, & (x, y) \in D; \\ 0, & (x, y) \notin D, \end{cases}$$

де S_D – площа області *D*, тобто щільність сумісного розподілу зберігає стале значення в області *D*, якій належать усі можливі значення випадкового вектора (X, Y).

Двовимірний експоненціальний розподіл. *Означення.* Двовимірна випадкова величина (X, Y) називається *експоненціально розподіленою*, якщо її функція розподілу визначається формулою:

$$F_{X,Y}(x, y) = \begin{cases} 0, \ x < 0 \ \text{àáî} \ \ y < 0; \\ 1 - e^{-(\lambda_1+\lambda_{12})x} - e^{-(\lambda_2+\lambda_{12})y} + e^{-\lambda_1 x - \lambda_2 y - \lambda_{12}\max\{x,y\}}, \ x \ge 0, \ y \ge 0. \end{cases}$$

Двовимірний нормальний розподіл. Нехай *X* та *Y* − нормально розподілені і незалежні випадкові величини, задані щільностями розподілу ймовірностей

$$f_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-a_1)^2}{2\sigma_1^2}}; \qquad f_2(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(y-a_2)^2}{2\sigma_2^2}}.$$

Тоді на основі теореми множення щільностей маємо:

$$f(x, y) = \frac{1}{\sigma_1\sigma_2 \cdot 2\pi} e^{-\frac{1}{2}\left[\frac{(x-a_1)^2}{\sigma_1^2} + \frac{(y-a_2)^2}{\sigma_2^2}\right]}.$$

Якщо центр розсіювання збігається з початком координат, тобто $a_1 = a_2 = 0$, тоді маємо канонічну форму двовимірного нормального розподілу:

$$f(x, y) = \frac{1}{\sigma_1\sigma_2 \cdot 2\pi} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2}\right)}.$$

Фіксованому значенню $f(x, y) = C$ відповідає деяке стале значення показника степеня $\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = k^2$, де k − const, звідки отримаємо рівняння

$$\frac{x^2}{(k\sigma_1)^2} + \frac{y^2}{(k\sigma_2)^2} = 1,$$

яке називають рівнянням еліпса однакової щільності або еліпсом розсіювання. Осі симетрії називають головними осями розсіювання.

Нехай X та Y − *залежні нормально розподілені* випадкові величини. Тоді

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \; e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x-a_1)(y-a_2)}{\sigma_1\sigma_2} + \frac{(y-a_2)^2}{\sigma_2^2}\right]},$$

де $\rho = \rho(x, y) -$ коефіцієнт кореляції.

Рівняння

$$\frac{(x-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x-a_1)(y-a_2)}{\sigma_1\sigma_2} + \frac{(y-a_2)^2}{\sigma_2^2} = k^2$$

є рівнянням еліпса з центром у точці (a_1, a_2), а осі симетрії утворюють з віссю Ox кути, які визначаються з рівняння $\operatorname{tg}2\alpha = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}$.

Орієнтація еліпсів розсіювання відносно координатних осей перебуває у прямій залежності від значення коефіцієнта кореляції. Якщо X та Y некорельовані $(\rho = 0)$, то головні осі розсіювання паралельні осям координат.

Щільність розподілу для некорельованих нормальних випадкових величин

$$f(x, y) = \frac{1}{\sigma_1\sigma_2 \cdot 2\pi} \; e^{-\frac{1}{2}\left[\frac{(x-a_1)^2}{\sigma_1^2} + \frac{(y-a_2)^2}{\sigma_2^2}\right]},$$

тобто така, як і для незалежних нормально розподілених випадкових величин. Це означає, що для них поняття некорельованості і незалежності еквівалентні.

Умовні закони нормального розподілу

$$f(y/x)=\frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)dy}=\frac{1}{\sigma_2\cdot\sqrt{2\pi}\sqrt{1-\rho^2}}\quad e^{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left[y-a_2-\rho\frac{\sigma_2}{\sigma_1}(x-a_1)\right]^2};$$

$$f(x/y)=\frac{f(x,y)}{\int\limits_{-\infty}^{\infty} f(x,y)dx}=\frac{1}{\sigma_1\cdot\sqrt{2\pi}\sqrt{1-\rho^2}}\quad e^{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left[x-a_1-\rho\frac{\sigma_1}{\sigma_2}(y-a_2)\right]^2}.$$

Легко бачити, що ці вирази є щільностями нормального розподілу з математичними сподіваннями

$$M[Y\mid X=x]=a_2+\rho\frac{\sigma_2}{\sigma_1}(x-a_1)=y(x);$$

$$M[X\mid Y=y]=a_1+\rho\frac{\sigma_1}{\sigma_2}(y-a_2)=x(y),$$

і середніми квадратичними відхиленнями $\sigma(y/x)=\sigma_2\sqrt{1-\rho^2}$, $\sigma(x/y)=\sigma_1\sqrt{1-\rho^2}$.

Частина III

Елементи математичної статистики

Розділ 7. Елементи математичної статистики. Вибірковий метод

У теорії ймовірностей вивчають питання, пов'язані з випадковими подіями, їх імовірностями та числовими характеристиками. Але в більшості випадків, які трапляються на практиці, точне значення цих характеристик невідоме. Метою кожного наукового дослідження власне і є встановлення закономірностей явищ, які спостерігають, та використання їх у повсякденній практичній діяльності. З огляду на це постає питання про їх експериментальне визначення й узагальнення висновку у вигляді закону. Можна сказати, що предметом математичної статистики є розроблення методів збору та оброблення статистичних даних для одержання наукових та практичних висновків. Звичайним є використання статистичних методів у техніці, медицині, економіці, соціології, політології.

7.1. Основні задачі математичної статистики

Математичною статистикою називається наука, яка займається розробкою методів відбору, опису та аналізу дослідних даних з метою вивчення закономірностей випадкових масових явищ.

У свою чергу, встановлення цих закономірностей ґрунтується на вивченні методами теорії ймовірностей статистичних даних – результатів досліду або спостережень.

Отже, математична статистика вивчає методи, які дають змогу за результатами випробувань робити певні ймовірнісні висновки.

Основні задачі математичної статистики:

1) вказати способи збору та групування (якщо даних дуже багато) статистичної інформації;

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

3) оцінити невідомі параметри розподілу;

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

Можна сказати, що ***основне завдання математичної статистики*** – розроблення методів аналізу статистичних даних залежно від мети дослідження.

Методи математичної статистики ефективно використовують під час розв'язування багатьох наукових задач, задач щодо організації технологічного процесу, планування, управління та ціноутворення.

Математична статистика виникла (XVII ст.) та почала розвиватись паралельно з теорією ймовірностей. Подальшим розвитком (кінець XIX – початок XX ст.) математична статистика зобов'язана П. Л. Чебишову, А. А. Маркову, О. М. Ляпунову, а також К. Гауссу, Ф. Гальтону, К. Пірсону й іншим вченим.

У XX ст. найбільший вклад у математичну статистику зробили В. І. Романовський, Е. Е. Слуцький, А. Н. Колмогоров, Стьюдент (псевдонім В. Госсета), Е. Пірсон, Ю. Нейман, А. Вальд, А. В. Скороход.

7.2. Генеральна та вибіркова сукупності

Кожен об'єкт, який спостерігають, має декілька ознак. Розглядаючи лише одну ознаку кожного об'єкта, припускають, що інші ознаки несуттєві, тобто множина об'єктів однорідна.

Означення. Множину однорідних об'єктів називають **статистичною сукупністю. Вибірковою сукупністю (вибіркою)** називають сукупність випадково взятих об'єктів із статистичної сукупності.

Генеральною називають сукупність об'єктів, з яких зроблено вибірку. ***Обсягом сукупності (вибіркової або генеральної)*** називають кількість об'єктів цієї сукупності.

Вибірки бувають ***повторні*** або ***безповторні***. Альтернатива вибірки – *перепис* – обстеження, що мають на меті дослідження кожного елемента сукупності (генеральної сукупності), яку вивчають.

Означення. **Повторною** називають вибірку, за якої відібраний об'єкт повертають до генеральної сукупності перед відбором іншого об'єкта.

Означення. Вибірку називають **безповторною**, якщо взятий об'єкт до генеральної сукупності не повертається перед відбором наступного об'єкта.

Вибірка має бути ***репрезентативною***, тобто *представницькою* – її дані мають правильно відображати відповідну ознаку генеральної сукупності, яку вивчають.

Проста випадкова вибірка здійснюється за допомогою *простого випадкового відбору* (кожний елемент із однаковою ймовірністю може потрапити до вибірки).

Джерела даних у статистиці: вибіркові обстеження, спеціально поставлені експерименти і дані, що є результатом повсякденної роботи у бізнесі. Джерела даних бувають *первинними і вторинними.*

Первинні дані збираються спеціально для статистичного дослідження. Для цих даних є відомості про методи збирання, точність даних тощо.

Вторинні дані – це дані, використовувані у статистиці, але зібрані з іншою метою. Наприклад, робочі записи про діяльність фірм, офіційні статистичні звіти, журнали обліку успішності та відвідування занять студентами тощо.

Способи відбору:

1. Вибір, який не потребує розділення генеральної сукупності на частини: простий випадковий безповторний відбір та простий випадковий повторний відбір.

2. Вибір, за якого генеральна сукупність розділяється на частини (розшарований випадковий відбір): типовий, механічний та серійний відбори.

В економічних дослідженнях іноді використовують *комбінований відбір*. Наприклад, спочатку поділяють генеральну сукупність на серії однакового обсягу, навмання відбирають декілька серій і, нарешті, з кожної серії навмання беруть окремі об'єкти.

Найкращий спосіб здійснення простої вибірки – *використання випадкових вибіркових чисел*. Ці числа складаються з цифр від 0 до 9, генеруються випадково (імовірність розміщення від будь-якої цифри у будь-якому місці таблиці не більше й не менше імовірності розміщення будь-якої іншої цифри з десяти наведених цифр) і записуються до спеціальної таблиці. Використання таблиць випадкових чисел гарантує, що не буде скоєно систематичної помилки, тобто помилки, яка робить дані не репрезентативними.

Означення. **Вибіркою обсягу *n*** називають n незалежних випадкових величин $X_1,\ X_2,\ \ldots,\ X_n$, кожна з яких – копія випадкової величини X з функцією розподілу $F(x)$.

Числа $x_1,\ x_2,\ \ldots,\ x_n$ – реалізація вибірки обсягу *n.*

7.3. Статистичний розподіл вибірки. Емпірична функція розподілу. Полігон і гістограма

Дані у статистиці, отримані за допомогою спеціальних досліджень або із звичайних робочих записів у бізнесі, надходять до дослідника у вигляді неорганізованої маси незалежно від того, чи є вони даними із вибіркової сукупності, чи даними з генеральної сукупності. Тому постає питання щодо оброблення і впорядкування даних.

Значення x_i вибірки називають ***варіантами.*** Послідовність варіант, розміщених у порядку зростання, називають *варіаційним рядом*. Якщо при цьому x_i повторюється n_i разів $(i=1,\ 2,\ \ldots,\ k;\ n_1+n_2+\ldots+n_k=n)$, то число n_i називають ***абсолютною частотою*** варіанти x_i, а $\dfrac{n_i}{n}$ – ***відносною частотою*** варіанти x_i.

Означення. **Статистичним розподілом** вибірки називають перелік варіант та відповідних частот або відносних частот. Статистичний закон розподілу зручно задавати таблицею, що встановлює зв'язок між значеннями випадкової величини та їх частотами:

x_i	x_1	x_2	$\ldots$	x_k
n_i	n_1	n_2	$\ldots$	n_k

Закон розподілу такий: *в даній серії випробувань випадкова величина x_i набуває значення n_i разів (i = 1, 2, …, k).*

Часто статистичний закон розподілу задають аналогічною таблицею, що встановлює зв'язок між значеннями випадкової величини та її відносними частотами. Цей закон називають статистичним рядом відносних частот.

x_i	x_1	x_2	$\ldots$	x_k
$\dfrac{n_i}{n}$	$\dfrac{n_1}{n}$	$\dfrac{n_2}{n}$	$\ldots$	$\dfrac{n_k}{n}$

Зауваження. У теорії ймовірностей під розподілом дискретної випадкової величини розуміють відповідність можливих значень випадкової величини їх імовірностям, а в математичній статистиці під розподілом розуміють відповідність спостережених значень (варіантів) їх частотам (або відносними частотам).

***Означення*. Емпірична функція розподілу** має такий вигляд:

$$F_n^*(x)=\begin{cases}0, & x\le x_1;\\ \sum\limits_{i=1}^{k}\dfrac{n_i}{n}, & x_i\le x<x_{i+1};\\ 1, & x>x_k.\end{cases} \tag{7.1}$$

Функція $F_n^*(x)$ має всі властивості функції розподілу $F(x)$ випадкової величини X.

1. $0\le F_n^*(x)\le 1$.

2. $F_n^*(x)$ – неспадна функція.

3. $F_n^*(x)$ – неперервна зліва.

Функцію називають *емпіричною функцією розподілу*, оскільки її знаходять емпіричним (дослідним) способом.

Функцію розподілу $F(x)$ генеральної сукупності називають *теоретичною функцією розподілу.*

Теоретична функція розподілу визначає ймовірність події $(X<x)$, *а емпірична функція* $F^*(x)$ *– відносну частоту цієї події.*

Із закону великих чисел, зокрема з теореми Бернуллі, випливає, що відносна частота події $(X<x)$ збігається з імовірністю цієї події, тобто

$$\lim_{n\to\infty}P\left(\left|F^*(x)-F(x)\right|<\varepsilon\right)=1.$$

Іншими словами, для великих значень n емпірична функція розподілу наближено представляє теоретичну функцію розподілу генеральної сукупності.

Означення. **Полігоном частот** вибірки називають ламану з вершинами в точках $(x_i,\ n_i)$, $(i=1,\ 2,\ ...,\ k)$. **Полігоном відносних частот** вибірки називають ламану з вершинами в точках $\left(x_i,\ \frac{n_i}{n}\right)$, $(i=1,\ 2,\ ...,\ k)$.

Якщо X – неперервна випадкова величина, то її статистичний розподіл також подають у вигляді таблиць. Для оцінювання $f(x)$ за вибіркою $x_1,\ x_2, ...,\ x_n$ розбивають інтервал на частинні інтервали довжини h, у таблицю записують середину i-го інтервалу, n_i – кількість елементів вибірки i-го інтервалу. Прямокутники з основами h і висотами $\frac{n_i}{nh}$ у прямокутній системі координат утворюють фігуру, яку називають ***гістограмою*** вибірки.

Приклад 1. Побудувати емпіричну функцію розподілу за розподілом вибірки

x_i	0,2	0,4	0,8	1,0	1,2	1,6
n_i	5	8	12	13	10	2

Розв'язання. Обсяг вибірки $n=50$.

Шуканий розподіл відносних частот має вигляд

x_i	0,2	0,4	0,8	1,0	1,2	1,6
$\frac{n_i}{n}$	$\frac{5}{50}$	$\frac{8}{50}$	$\frac{12}{50}$	$\frac{13}{50}$	$\frac{10}{50}$	$\frac{2}{50}$

або остаточно

x_i	0,2	0,4	0,8	1,0	1,2	1,6
$\frac{n_i}{n}$	0,1	0,16	0,24	0,26	0,2	0,04

Тоді записуємо емпіричну функцію розподілу за формулою (7.1):

$$F^*(x)=\begin{cases} 0, & x\le 0,2; \\ \dfrac{5}{50}=0{,}1, & 0,2<x\le 0,4; \\ \dfrac{5+8}{50}=0{,}26, & 0,4<x\le 0,8; \\ \dfrac{5+8+12}{50}=0{,}5, & 0,8<x\le 1,0; \\ \dfrac{5+8+12+13}{50}=0{,}76, & 1,0<x\le 1,24; \\ \dfrac{5+8+12+13+10}{50}=0{,}96, & 1,2<x\le 1,6; \\ 1, & x>1,6. \end{cases}$$

7.4. Точкові оцінки невідомих параметрів розподілів

У багатьох випадках потрібно дослідити кількісну ознаку X генеральної сукупності, використовуючи результати вибірки. Дослідник має дані вибірки, одержані в результаті спостережень. Іноді з певних міркувань вдається встановити закон розподілу випадкової величини X. Тоді потрібно вміти оцінювати параметри цього розподілу.

Означення. **Статистичною оцінкою** невідомого параметра випадкової величини X генеральної сукупності називають функцію від випадкових величин (результатів вибірки), що спостерігаються.

Нехай $X_1,\ X_2,\ ...,\ X_n$ – вибірка з генеральної сукупності X із функцією розподілу $F(x,\ \theta)$, яка має відомий функціональний вигляд, але залежить від невідомого параметра θ.

Означення. Оцінка $\widehat{\theta}=h(X_1,\ X_2,\ ...,X_n)$ параметра θ називається ***незміщеною***, якщо $M\widehat{\theta}=\theta$.

Якщо $M\widehat{\theta}=\theta+b(\theta)$, то $b(\theta)$ називають зсувом оцінки $\widehat{\theta}$.

Означення. Послідовність оцінок $\widehat{\theta}_n=h(X_1,\ X_2,\ \cdots,X_n),\quad n=1,\ 2,...$ параметра θ називається ***слушною***, якщо $\widehat{\theta}_n\xrightarrow{p}\theta,\ n\to\infty.$

Означення. Послідовність оцінок $\widehat{\theta}_n = h(X_1,\ X_2,\ ...,\ X_n),\ n=1,\ 2,\ ...$ параметра θ називається **сильно слушною**, якщо $\widehat{\theta}_n \xrightarrow{p} \theta,\ n\to\infty$ з імовірністю 1.

Означення. Оцінка $\widehat{\theta}$ параметра θ називається ***ефективною***, якщо вона за даного обсягу вибірки має найменшу можливу дисперсію.

Незміщеною сильно слушною *оцінкою математичного сподівання* випадкової величини *X* є ***вибіркове середнє*** $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ і реалізацію цієї оцінки позначають

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i . \tag{7.2}$$

Вибіркове середнє для ***дискретного статистичного ряду*** обчислюють за формулою

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i,\ \text{де}\ \sum_{i=1}^{k} n_i = n.$$

Вибіркове середнє для ***інтервального статистичного ряду*** обчислюють за формулою

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i z_i,$$

де z_i – середина *i*-го інтервалу, $\sum_{i=1}^{k} n_i = n.$

Оцінкою дисперсії σ^2 випадкової величини *X* є ***вибіркова дисперсія***

$$\overline{s}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2, \tag{7.3}$$

яка є ***зміщеною*** оцінкою для σ^2.

Величина

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2,\ s^2 = \frac{n}{n-1}\overline{s}^2 \tag{7.4}$$

є ***незміщеною*** оцінкою дисперсії σ^2 випадкової величини *X*.

Якщо математичне сподівання *a* відоме, то *незміщеною сильно слушною оцінкою дисперсії* σ^2 випадкової величини *X* є оцінка

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - a\right)^2.$$

Вибіркову дисперсію для *дискретного статистичного ряду* обчислюють за формулою

$$\overline{s}^2 = \frac{1}{n}\sum_{i=1}^{k}\left(n_i x_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^{\,2} - \overline{x}^2,$$

відповідно

$$s^2 = \frac{n}{n-1}\overline{s}^2;\ s^2 = \frac{1}{n-1}\sum_{i=1}^{k}\left(n_i x_i - \overline{x}\right)^2 = \frac{1}{n-1}\sum_{i=1}^{k} n_i x_i^{\,2} - \overline{x}^2.$$

Вибіркову дисперсію для *інтервального статистичного ряду* обчислюють за формулою

$$\overline{s}^2 = \frac{1}{n}\sum_{i=1}^{k}\left(n_i z_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i z_i^{\,2} - \overline{x}^2,$$

відповідно

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k}\left(n_i z_i - \overline{x}\right)^2 = \frac{1}{n-1}\sum_{i=1}^{k} n_i z_i^{\,2} - \overline{x}^2.$$

Значення $\overline{s} = \sqrt{\overline{s}^2}$ називається ***вибірковим середнім квадратичним відхиленням,*** $s = \sqrt{s^2}$ ***– вибірковим виправленим середнім квадратичним відхиленням*** (s^2 – вибіркова виправлена дисперсія).

Зауваження. Властивістю ефективності часто нехтують, вибираючи в якості пріоритету раціональність обчислення оцінки:

1) якщо $\left(X_1,\ldots,X_n\right)$ – незалежна вибірка з генеральної сукупності з розподілом $N(a,\ \sigma^2)$, то вибіркове середнє $\overline{x}$ є ефективною оцінкою для математичного сподівання і при відомому математичному сподіванні вибіркова дисперсія s^2 є також ефективною оцінкою дисперсії σ^2;

2) відносна частота $\hat{\theta} = \dfrac{\mu_n}{n}$ є ефективною оцінкою ймовірності появи події в схемі Бернуллі;

3) Вибіркова середня закону Пуассона є ефективною оцінкою параметра λ.

Означення. **Медіаною** M_e називають варіанту, яка ділить варіаційний ряд на дві частини, рівні за кількістю варіант.

Для дискретних статистичних рядів

$$M_e = \begin{cases} x_m, \ n = 2m - 1; \\ \dfrac{x_m + x_{m+1}}{2}, \quad \text{çà} \ \ n = 2m. \end{cases}$$

Для інтервальних статистичних рядів

$$M_e = x_i + h_i \frac{\dfrac{n}{2} - \sum\limits_{j=1}^{i} n_j}{n_i},$$

де x_i – початок медіанного інтервалу (йому відповідає перша з нагромаджених частот, що більша від половини всіх спостережень), h_i – довжина i-го інтервалу, n_i – частота медіанного інтервалу.

Означення. **Модою** називають варіанту, яка найчастіше трапляється у вибірці.

Для дискретних статистичних рядів

$$M_0 = x_j, \text{ якщо } n_j = \max_i n_i.$$

Для інтервальних статистичних рядів

$$M_0 = x_i + h \frac{n_i - n_{i-1}}{2n_i - n_{i-1} - n_{i+1}},$$

де x_i – початок інтервалу із найбільшою частотою, n_i – частота i-го інтервалу.

Зауваження. Якщо кількість варіант досить велика, то для спрощення обчислень застосовують групування емпіричних даних. Інтервал, в якому лежать вибіркові дані, розбивають на k рівних частин довжини h. Якщо z_i – середина i-ї частини, n_i – кількість варіант, що потрапили в i-ту частину, то $\overline{x}, \ \overline{s}^2$ обчислюють за формулами

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i z_i; \ \ \overline{s}^2 = \frac{1}{n}\sum_{i=1}^{k} n_i z_i^2 - \overline{x}^2.$$

Для подальшого спрощення обчислень застосовують метод умовних варіант, які визначаються рівністю $u_i = \dfrac{z_i - C}{h}$, C – умовний нуль (новий початок відліку). Найбільшої простоти обчислень досягають тоді, коли за умовний нуль взяти варіанту, розміщену приблизно посередині варіаційного ряду і яка має найбільшу частоту. Легко знайти, що

$$\overline{x} = h\overline{u} + C,\ \overline{s}_x^{\,2} = h^2\overline{s}_u^{\,2}.$$

Якщо варіанти x_i вибірки рівновіддалені, тобто утворюють арифметичну прогресію з різницею h, то можна зразу вводити умовні варіанти, не застосовуючи методу групування.

Приклад 2. Обчислити незміщені оцінки математичного сподівання та дисперсії, вибіркову медіану та моду вибірки.

x_i	12	14	16	18	20	22
n_i	5	15	50	16	10	4

Розв'язання. Обсяг вибірки $n = 5+15+50+16+10+4 = 100$. Вибіркове середнє

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{1}{100}\left(5\cdot 12 + 15\cdot 14 + 50\cdot 16 + 16\cdot 18 + 10\cdot 20 + 4\cdot 22\right) = 16{,}46.$$

Щоб знайти значення $\overline{s}^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \overline{x}^2$ обчислюємо

$$\frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 = \frac{1}{100}\left(5\cdot 12^2 + 15\cdot 14^2 + 50\cdot 16^2 + 16\cdot 18^2 + 10\cdot 20^2 + 4\cdot 22^2\right) = 275{,}8,$$

і тоді $\overline{s}^2 = 275{,}8 - 16{,}46^2 = 4{,}8684$;

$$s^2 = \frac{n}{n-1}\overline{s}^2 = \frac{100}{99}\cdot 4{,}8684 = 4{,}92.$$

Медіана та мода вибірки $M_e = \dfrac{16+18}{2} = 17;\ M_0 = 16.$

Відповідь. $\overline{x} = 16{,}46;\ s^2 = 4{,}92;\ M_e = 17;\ M_0 = 16.$

Приклад 3. Обчислити вибіркове середнє та дисперсію, медіану та моду для вибірки

інтервал	$[2, 4)$	$[4, 6)$	$[6, 8)$	$[8, 10)$	$[10, 12)$
n_i	2	8	35	40	15

Розв'язання. Обсяг вибірки $n = 2+8+35+40+15 = 100$. Тоді

z_i	3	5	7	9	11
n_i	2	8	35	40	15

Знаходимо числові характеристики:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i z_i = \frac{1}{100}\left(2\cdot 3 + 8\cdot 5 + 35\cdot 7 + 40\cdot 9 + 15\cdot 11\right) = 8{,}16;$$

$$\overline{s}^2 = \frac{1}{n}\sum_{i=1}^{k} n_i z_i^2 - \overline{x}^2 = \frac{1}{100}\left(2\cdot 9 + 8\cdot 25 + 35\cdot 49 + 40\cdot 81 + 15\cdot 121\right) - 8{,}16^2 =$$

$$= 69{,}88 - 66{,}5856 = 3{,}2944;$$

$$s^2 = \frac{n}{n-1}\overline{s}^2 = \frac{100}{99}\cdot 3{,}29 = 3{,}33;$$

$$M_e = 8 + 2\frac{50-45}{40} = 8{,}25;\ M_0 = 8 + 2\frac{40-35}{80-35-15} = 8{,}33.$$

Відповідь. $\overline{x} = 8{,}16;\ s^2 = 3{,}33;\ M_e = 8{,}25;\ M_0 = 8{,}33.$

7.5. Метод моментів та метод максимальної правдоподібності

Метод моментів точкової оцінки невідомих параметрів розподілу полягає у прирівнюванні теоретичних моментів розподілу та відповідних емпіричних моментів того самого порядку. Цей метод запропонований К. Пірсоном.

7.5.1. Метод моментів

Нехай $\tilde{O}_1,\ \tilde{O}_2,\ \ldots,\tilde{O}_n$ − вибірка з генеральної сукупності, $x_1,\ x_2,\ \ldots,\ x_n$ – реалізація вибірки.

Означення. **Початковим емпіричним моментом порядку** k називається вираз

$$\grave{I}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \qquad (7.5)$$

зокрема $M_1 = \overline{x}$.

Означення. **Центральним емпіричним моментом порядку** k називається вираз

$$m_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^k, \tag{7.6}$$

зокрема $m_2 = \overline{s}^2$.

Означення. **Коефіцієнтом асиметрії** A_s називається відношення центрального емпіричного моменту третього порядку до куба середнього квадратичного відхилення:

$$A_s = \frac{m_3}{\sigma^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{s^3}. \tag{7.7}$$

Означення. **Ексцесом** E називається зменшене на три одиниці відношення центрального моменту четвертого порядку до четвертого степеня середнього квадратичного відхилення:

$$E = \frac{m_4}{\sigma^4} - 3. \tag{7.8}$$

Оцінка одного параметра. Припустимо, що нам відомий вигляд щільності розподілу $f(x, \theta)$ ознаки X, який визначається одним параметром θ. Розглянемо, як знайти точкову оцінку цього параметра. Для отримання оцінки одного параметра достатньо мати одне рівняння відносно цього параметра, тому прирівнюють початковий момент першого порядку $\nu_1 = M(X)$ і початковий емпіричний момент першого порядку $M_1 = \overline{x}$:

$$M(\theta) = \overline{x}. \tag{7.9}$$

Розв'язавши рівняння, одержують оцінку параметра θ, яка є функцією від $\overline{x}$, отже і від $x_1,\ x_2,\ ...,\ x_n$.

Приклад 4. Випадкова величина X − час роботи елемента, вона має показниковий розподіл з параметром λ. Отримано статистичний розподіл середнього часу роботи 200 елементів

x_i	2,5	7,5	12,5	17,5	22,5	27,5
n_i	133	45	15	4	2	1

де x_i − середній час роботи елемента в годинах, частота n_i − кількість елементів, які пропрацювали в середньому x_i годин. Знайти методом моментів точкову оцінку параметра λ.

Розв'язання. Прирівнявши теоретичний і емпіричний моменти першого порядку і враховуючи, що для показникового закону $M\left(X\right)=\dfrac{1}{\lambda}$, отримаємо $\lambda=\dfrac{1}{\overline{x}}$. Отже, точковою оцінкою параметра λ є $\lambda^*=\dfrac{1}{\overline{x}}$. Обчисливши

$$\overline{x}=\frac{1}{200}\sum_{i=1}^{6}n_i x_i=5,$$

одержимо $\lambda^*=\dfrac{1}{5}=0{,}2.$

Оцінка двох параметрів. Нехай щільність розподілу $f(x,\theta_1,\theta_2)$ генеральної сукупності X визначається двома невідомими параметрами θ_1,θ_2. Для їх визначення потрібно мати два рівняння. Прирівняємо теоретичний та емпіричний початкові моменти першого порядку $\nu_1=\nu_1^*$ і теоретичний та емпіричний центральні моменти другого порядку $m_2=m_2^*$. Розв'язавши систему

$$\begin{cases} M(\theta_1)=\overline{x}; \\ D(\theta_2)=\overline{s}^2, \end{cases} \qquad (7.10)$$

одержимо оцінку параметрів $\theta_1,\ \theta_2$, які є функціями від $x_1,\ x_2,\ ...,\ x_n$.

Приклад 5. Випадкова величина X − відхилення контрольованого параметра виробу від переглянутого і зміненого нормативу, що підлягає нормальному закону розподілу з параметрами a і σ. Отримано статистичний розподіл відхилення 200 виробів

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,2	2,3
n_i	6	9	26	25	30	26	21	24	20	8	5

Знайти методом моментів точкові оцінки параметрів a і σ.

Розв'язання. Враховуючи, що для нормального розподілу $\nu_1=M\left(X\right)=a$,

$m_2 = D(X) = \sigma^2$, і прирівнюючи відповідні теоретичні й емпіричні моменти $\nu_1 = \nu_1^*$, $m_2 = m_2^*$, отримаємо вирази для точкових оцінок:

$$a^* = \bar{x} = \frac{1}{200}\sum_{i=1}^{11} n_i x_i = 1{,}266;$$

$$(\sigma^2)^* = \frac{1}{200}\sum_{i=1}^{11} n_i x_i^2 - \left(\frac{1}{200}\sum_{i=1}^{11} n_i x_i\right)^2 = 0{,}25,$$

звідки $\sigma^* = 0{,}5$.

Отже, $a^* = 1{,}266$ та $\sigma^* = 0{,}5$.

7.5.2. Метод максимальної правдоподібності

Ідея методу максимуму правдоподібності полягає в тому, що за оцінку $\widehat{\theta}$ беруть таке значення параметра θ, для якого ймовірність отримання вже наявної вибірки максимальна.

Припустимо, що вигляд закону розподілу випадкової величини заданий, але параметр θ, яким визначається цей закон, невідомий. Потрібно знайти його точкову оцінку. Позначимо ймовірність того, що в результаті випробування випадкова величина X набуде значення x_i $(i = 1, 2, ..., n)$ через $p(x_i, \theta)$.

***Означення*. Функцією правдоподібності** дискретної випадкової величини X називають функцію аргумента θ:

$$L(x_1,\ x_2,\ ...,\ x_n,\ \theta) = p(x_1, \theta)\cdot p(x_2, \theta)\cdot\ldots\cdot p(x_n, \theta),$$

де $x_1,\ x_2,\ ...,\ x_n$ – вибіркові значення.

***Означення*. Оцінкою максимальної правдоподібності** параметра θ називається таке його значення $\widehat{\theta} = \widehat{\theta}(x_1,\ x_2,\ \ldots x_n)$, за якого функція правдоподібності досягає максимального значення.

Функцію $\ln L$ називають ***логарифмічною функцією правдоподібності***, а рівняння $\dfrac{d\ln L}{d\theta} = 0$ – ***рівнянням правдоподібності***. Якщо друга похідна

$$\frac{d^2 \ln L}{d\theta^2} < 0$$

за $\theta = \widehat{\theta}$, то $\widehat{\theta}$ – точка максимуму.

***Означення*. Функцією правдоподібності** неперервної випадкової величини X називають функцію аргумента θ:

$$L(x_1,\ x_2,\ ...,\ x_n,\ \theta) = f(x_1, \theta)\cdot f(x_2, \theta)\cdot \ldots \cdot f(x_n, \theta),$$

де $x_1,\ x_2,\ ...,\ x_n$ – фіксовані числа.

Якщо щільність розподілу $f(x, \theta_1, \theta_2)$ неперервної випадкової величини X визначається двома параметрами θ_1 і θ_2, то функція правдоподібності – це функція двох аргументів θ_1 і θ_2:

$$L(x_1,\ x_2,\ ...,\ x_n, \theta_1, \theta_2) = f(x_1, \theta_1, \theta_2)\cdot f(x_2, \theta_1, \theta_2)\cdot \ldots \cdot f(x_n, \theta_1, \theta_2),$$

де $x_1,\ x_2,\ ...,\ x_n$ – значення X, які спостерігались.

Приклад. Знайти оцінку максимальної правдоподібності параметрів $a,\ \sigma$ нормального розподілу $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$ за вибіркою $x_1,\ x_2,\ ...,\ x_n$.

Розв'язання. Складемо функцію правдоподібності

$$L(x_1,\ x_2,\ ...,\ x_n,\ a, \sigma) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} e^{-\frac{\sum_{i=1}^{n}(x_i-a)^2}{2\sigma^2}},$$

звідки $\ln L = -n\ln\sigma - \ln\left(\sqrt{2\pi}\right)^n - \dfrac{\sum_{i=1}^{n}(x_i-a)^2}{2\sigma^2}.$

Розв'яжемо систему

$$\begin{cases} \dfrac{\partial \ln L}{\partial a} = 0; \\ \dfrac{\partial \ln L}{\partial \sigma} = 0; \end{cases} \quad \begin{cases} \sum_{i=1}^{n} x_i - na = 0; \\ -n\sigma^2 + \sum_{i=1}^{n}\left(x_i - a\right)^2 = 0; \end{cases} \quad \begin{cases} a = \dfrac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}; \\ \sigma^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - a)^2 = \overline{s}^2. \end{cases}$$

Знаходимо частинні похідні другого порядку для визначення точки екстремуму:

$$\frac{\partial^2 \ln L}{\partial a^2} = -\frac{n}{\sigma^2};\ \frac{\partial^2 \ln L}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^{n}(x_i-a)^2}{\sigma^4};\ \frac{\partial^2 \ln L}{\partial a \partial \sigma} = -\frac{2\sum_{i=1}^{n}(x_i - a)}{\sigma^3}.$$

Тоді $A=\left.\dfrac{\partial^2 \ln L}{\partial a^2}\right|_{\substack{a=\overline{x}\\ \sigma=\overline{s}}}=-\dfrac{n}{\overline{s}^2}<0;\ \tilde{N}=\left.\dfrac{\partial^2 \ln L}{\partial \sigma^2}\right|_{\substack{a=\overline{x}\\ \sigma=\overline{s}}}=\dfrac{n}{\overline{s}^2}-\dfrac{3\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}{\overline{s}^4}=-\dfrac{2n}{\overline{s}^2};$

$$B=\left.\frac{\partial^2 \ln L}{\partial a \partial \sigma}\right|_{\substack{a=\overline{x}\\ \sigma=\overline{s}}}=-\frac{2\sum\limits_{i=1}^{n}(x_i-\overline{x})}{\overline{s}^3}=0, \text{ оскільки } \sum_{i=1}^{n}(x_i-\overline{x})=0.$$

Отже, $\Delta = AC-B^2=\left(-\dfrac{n}{\overline{s}^2}\right)\cdot\left(-\dfrac{2n}{\overline{s}^2}\right)=\dfrac{2n^2}{\overline{s}^4}>0$ і $A<0$, тому функція $\ln L$ в точці $(\overline{x},\ \overline{s}^2)$ має максимум. Для параметрів (a,σ^2) оцінкою максимальної правдоподібності є $(\overline{x},\ \overline{s}^2)$.

Зауваження. Метод максимуму правдоподібності оцінювання параметрів розподілу має низку важливих переваг: вони слушні, асимптотично нормально розподілені (за великих n їх розподіл близький до нормального) і мають найменшу дисперсію порівняно з іншими асимптотично нормальними оцінками. Цей метод найбільш повно використовує дані вибірки для оцінювання параметрів, тому він особливо корисний для малих обсягів вибірки. Однак його недолік полягає у тому, що він вимагає складних обчислень.

Розділ 8. Статистичні оцінки параметрів розподілу

Одна з основних задач математичної статистики – оцінювання параметрів розподілів випадкових величин. У попередній темі ми розглядали методи побудови і властивості точкових оцінок, тобто оцінок, які визначаються одним числом. Звичайно, в разі користування такими оцінками ми не можемо вказати їх точність, тобто наскільки вони відхиляються від істинних значень параметрів. Зрозуміло, що точкові оцінки залежать від обсягу вибірки. Зокрема, якщо обсяг вибірки малий, то точкова оцінка може суттєво відрізнятися від оцінюваного параметра. З огляду на це користуватися інтервальними оцінками, тобто такими оцінками, які визначаються двома числами − кінцями інтервалу.

8.1. Надійні інтервали

Якщо обсяг вибірки досить великий, то точкові оцінки задавольняють практичні потреби точності.

Нагадаємо, що ***точковими оцінками*** параметрів розподілу генеральної сукупності називають такі оцінки, які визначаються одним числом.

***Означення*. Інтервальною оцінкою** параметрів розподілу генеральної сукупності називають оцінку, яка визначається двома числами – кінцями інтервалу.

Нехай $\tilde{O}_1,\ \tilde{O}_2,\ ...,\tilde{O}_n$ – вибірка з генеральної сукупності, θ – невідомий параметр генеральної сукупності *X*. У разі інтервального оцінювання невідомого параметра θ шукають такі дві функції $h_1 = h_1(X_1,\ X_2,\ \ldots,\ X_n)$ і $h_2 = h_2(X_1,\ X_2,\ \ldots,\ X_n)$, для яких завжди $h_1 < h_2$ і за заданого $\gamma \in (0,\ 1)$ виконується умова

$$P\left(h_1 < \theta < h_2\right) \geq \gamma . \tag{8.1}$$

Тоді інтервал (h_1, h_2) називають ***надійним інтервалом***, γ ***– рівнем надійності***, а числа h_1 і h_2 – ***нижньою та верхньою межами надійності***.

Рівень надійності задають наперед, найчастіше його беруть 0,95; 0,99.

Надійний інтервал для математичного сподівання будь-якого розподілу легко знайти за допомогою центральної граничної теореми.

Метод надійних інтервалів у статистиці започаткований Р. Фішером і Ю. Нейманом.

Розглянемо деякі задачі на побудову надійних інтервалів.

8.1.1. Надійні інтервали для оцінювання математичного сподівання нормального розподілу за відомого σ

Нехай відомо, що випадкова величина X розподілена нормально і σ − її середнє квадратичне відхилення. Потрібно побудувати інтервальну оцінку для невідомого математичного сподівання $M(X)=a$. Точковою оцінкою для математичного сподівання є вибіркове середнє

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i.$$

Середнє вибіркове $\overline{x}$ різне для окремо взятих вибірок з генеральної сукупності, отже, його можна розглядати як випадкову величину $\overline{X}$, а значення x_i як однаково розподілені незалежні випадкові величини X_i ($i=\overline{1,n}$). Оскільки значення x_i незалежні, то

$$M\left(\overline{x}\right)=M\left[\frac{1}{n}\sum_{i=1}^{n}x_i\right]=\frac{1}{n}\sum_{i=1}^{n}M\left(X_i\right)=M\left(X\right);$$

$$D\left(\overline{x}\right)=D\left[\frac{1}{n}\sum_{i=1}^{n}x_i\right]=\frac{1}{n^2}\sum_{i=1}^{n}D\left(X_i\right)=\frac{D\left(X\right)}{n}.$$

Вважаємо, що $\sqrt{D\left(\overline{x}\right)}=\frac{\sqrt{D\left(X\right)}}{\sqrt{n}}=\frac{\sigma}{\sqrt{n}}$ − відома величина. Нерівність $\left|\overline{x}-a\right|<\varepsilon$ має виконуватись із заданою ймовірністю γ:

$$P\left(\left|\overline{x}-a\right|<\varepsilon\right)=\gamma$$

або, замінивши її еквівалентною нерівністю, отримаємо:

$$P\left(a-\varepsilon<\overline{x}<a+\varepsilon\right)=\gamma, \tag{8.2}$$

Пригадаємо, що для нормально розподіленої випадкової величини X з параметрами a і σ ймовірність попадання в інтервал (α, β) визначається формулою

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right),$$

де $\Phi(z)$ − функція Лапласа (табульована).

Тоді співвідношення (8.2) можна переписати так:

$$\Phi\left(\frac{a+\varepsilon-a}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{a-\varepsilon-a}{\sigma/\sqrt{n}}\right) = \Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{-\varepsilon}{\sigma/\sqrt{n}}\right) = 2\Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}}\right) = \gamma .$$

Позначивши $t = \dfrac{\varepsilon\sqrt{n}}{\sigma}$, маємо рівняння $2\Phi(t) = \gamma$ і остаточно отримаємо

$$P\left(\overline{x} - \frac{t\sigma}{\sqrt{n}} < a < \overline{x} + \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma. \tag{8.3}$$

Отже, побудований надійний інтервал

$$\left(\overline{x} - \frac{t\sigma}{\sqrt{n}},\ \overline{x} + \frac{t\sigma}{\sqrt{n}}\right) \tag{8.4}$$

містить невідомий параметр a (математичне сподівання) з імовірністю γ. Число t за заданого значення γ знаходимо із таблиці значень функції Лапласа.

Висновки:

1) зі збільшенням обсягу n вибірки число $\varepsilon = \dfrac{\sigma t}{\sqrt{n}}$ зменшується, тобто точність оцінки підвищується;

2) зростання надійності $\gamma = 2\Phi(t)$ веде до збільшення t, отже до зростання ε або до зниження точності.

Приклад 1. Нехай $\sigma = 2,\ n = 25,\ \gamma = 0{,}95$. Знайти надійний інтервал для a, якщо $\overline{x} = 5$.

Розв'язання. Для обчислення t використаємо рівняння

$$2\Phi(t) = \gamma \quad \Rightarrow \quad 2\Phi(t) = 0{,}95 \quad \Rightarrow \Phi(t) = 0{,}475 .$$

Із таблиці значень функції Лапласа знаходимо $t = 1{,}96$.

Отже, $\varepsilon = \frac{2 \cdot 1,96}{\sqrt{25}} = 0,784$ й отримаємо інтервал $\left(\overline{x} - 0,784;\ \ \overline{x} + 0,784\right)$ або $\left(4,216;\ \ 5,784\right)$.

Цей результат можна розуміти так: якщо зроблена достатньо велика кількість вибірок, то в 95 % випадків значення *a* належить знайденому інтервалу, а в 5 % це значення *a* може вийти за межі інтервалу.

8.1.2. Надійні інтервали для оцінювання математичного сподівання нормального розподілу за невідомого σ

У цьому випадку для малих обсягів вибірки $(n < 30)$ використовують ***розподіл Стьюдента***. Середнє вибіркове $\overline{x}$ і виправлене середнє квадратичне відхилення *s* є різними для окремо взятих вибірок з генеральної сукупності, отже, їх можна розглядати як випадкові величини $\overline{X}$ і S. За даними вибірки обсягу *n* можна побудувати випадкову величину

$$T = \frac{\overline{X} - a}{S / \sqrt{n}} \tag{8.5}$$

(її можливі значення позначимо через *t*), яка має ***розподіл Стьюдента*** з $n-1$ степенями вільності. Перевагою цього розподілу є те, що він визначається одним параметром − обсягом вибірки *n* і не залежить від невідомих параметрів a з σ. Щільність ***розподілу Стьюдента*** має вигляд

$$f(t, n) = \frac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\Gamma((n-1)/2)} \left(1 + t^2/(n-1)\right)^{-n/2}, \tag{8.6}$$

де $\Gamma(z) = \int\limits_0^{+\infty} u^{z-1} e^{-u} du$ − гамма-функція.

Оскільки $f(t, n)$ парна функція від t, то ймовірність виконання нерівності

$$\left| \frac{\overline{X} - a}{S / \sqrt{n}} \right| < t_\gamma$$

можна замінити рівносильною їй подвійною нерівністю вигляду

$$P\left\{ \overline{X} - t_\gamma \frac{S}{\sqrt{n}} < a < \overline{X} + t_\gamma \frac{S}{\sqrt{n}} \right\} = \gamma, \tag{8.7}$$

Для конкретної вибірки обсягу n випадкові величини $\overline{X}$ і S замінимо невипадковими $\overline{x}$ і s. Отже, використовуючи ***розподіл Стьюдента***, знайдемо надійний інтервал для оцінювання математичного сподівання нормального розподілу за невідомого σ:

$$\left(\overline{x}-t_{\gamma}\frac{s}{\sqrt{n}};\ \overline{x}+t_{\gamma}\frac{s}{\sqrt{n}}\right), \tag{8.8}$$

який покриває невідомий параметр a з надійністю γ, $t_{\gamma}=t(\gamma,k)$ − табульоване значення випадкової величини, розподіленої за ***законом Стьюдента***, $k=n-1$ – кількість степенів вільності.

Зауваження. У статистиці кількістю степенів вільності певної величини називають різницю між кількістю випробувань і кількістю величин, обчислених завдяки цим випробуванням.

Приклад 2. Нехай X − нормально розподілена випадкова величина, яка має такі параметри: $n=25$, $\overline{x}=20$, $s=0,4$. Знайти надійний інтервал для оцінювання математичного сподівання, якщо $\gamma=0,95$.

Розв'язання. Знаходимо t_{γ} з таблиці: $t(0{,}95;\ 24)=2{,}064$. Тоді

$$t_{\gamma}\frac{s}{\sqrt{n}}=2{,}064\cdot\frac{0{,}4}{5}\approx 0{,}165.$$

Отже, $(\overline{x}-0{,}165;\ \overline{x}+0{,}165)$ − надійний інтервал або $(19{,}835;\ 20{,}165)$.

Приклад 3. Для визначення середньої врожайності пшениці на площі 500 000 га проведено вибіркові вимірювання врожайності на 2 500 га. Результати вимірювань подано в таблиці:

Врожайність з 1 га (в центнерах)	$[21, 23)$	$[23,25)$	$[25, 27)$	$[27, 29)$	$[29, 31)$	$[31, 33)$
Кількість гектарів	100	300	500	700	600	300

Знайти середню врожайність з гектара і вибіркову дисперсію. Знайти ймовірність того, що середня врожайність з 500 000 га відрізняється від вибіркової середньої не більше ніж на 9 кг.

Розв'язання. Знайдемо середини інтервалів z_i і введемо умовну варіанту $C=28$, $h=2$.

№	z_i	n_i	u_i	n_iu_i	$n_iu_i^2$
1	22	100	–3	–300	900
2	24	300	–2	–600	1200
3	26	500	–1	–500	500
4	28	700	0	$A_1=-1400$	0
5	30	600	1	600	600
6	32	300	2	600	1200
				$A_2=1200$	
$\sum$	–	2500	–	–200	4400

$$\overline{u}=\frac{1}{n}\sum n_iu_i=-\frac{200}{2500}=-0,08;\ \frac{1}{n}\sum n_iu_i^2=\frac{4400}{2500}=1,76;$$

$$\overline{s}_u^2=\frac{1}{n}\sum n_iu_i^2-\overline{u}^2=1,76-(-0,08)^2=1,7536;$$

$$\overline{x}=h\overline{u}+C=-2\cdot 0,08+28=27,84;$$

$$\overline{s}_x^{\,2}=h^2\overline{s}_u^2=4\cdot 1,7536=7,0144;$$

$$s_x^{\,2}=\frac{n}{n-1}\overline{s}_x^2=\frac{2500}{2499}\cdot 7,0144=7,0172;\ s_x=2,649;$$

$$P\left(\left|\overline{x}-a\right|<\frac{s}{\sqrt{n}}t\right)=2\Phi(t).$$

За умовою задачі $\frac{s}{\sqrt{n}}t=0,09;\ t=\frac{0,09\sqrt{n}}{s}=\frac{0,09\cdot\sqrt{2500}}{2,649}=1,7,$ отже

$$P\left(\left|\overline{x}-a\right|<0,09\right)=2\Phi(1,7)=2\cdot 0,45543=0,91086.$$

8.1.3. Оцінка ймовірності (біномного розподілу) за відносною частотою

Нехай проводяться незалежні випробування з невідомою ймовірністю p настання події *A* в кожній спробі. Оцінимо невідому ймовірність p за відносною частотою $k_n=\frac{m}{n}$.

Зауваження. За точкову оцінку ймовірності *p* беруть відносну частоту

$k_n = \dfrac{m}{n}$ (n − кількість проведених випробувань, у яких подія A відбулася m разів). Ця оцінка незміщена. Дійсно, оскільки $M(m) = np,$ то

$$M(k_n) = \frac{M(m)}{n} = \frac{np}{n} = p,$$

тобто математичне сподівання оцінки дорівнює значенню оцінюваного параметра.

Дисперсія оцінки (оскільки $D(m) = npq$)

$$D(k_n) = \frac{D(m)}{n^2} = \frac{pq}{n},$$

звідки $\sigma_{k_n} = \sqrt{\dfrac{pq}{n}}.$

Розглянемо *інтервальну оцінку*.

Якщо n досить велике і $0 < p < 1,$ тоді для відносної частоти k_n можна користуватися формулою

$$P\left(\left|X - a\right| < \varepsilon\right) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

В нашому випадку

$$P\left\{\left|k_n - p\right| < \varepsilon\right\} = 2\Phi\left(\frac{\varepsilon}{\sigma_{k_n}}\right) = \gamma \quad \text{або} \quad P\left\{\left|k_n - p\right| < \varepsilon\right\} = 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) = \gamma\,.$$

Позначимо $t = \dfrac{\varepsilon\sqrt{n}}{\sqrt{pq}},$ де t визначається з рівняння $2\Phi(t) = \gamma.$

Отже, $| k_n - p | < \dfrac{t\sqrt{pq}}{\sqrt{n}}$ або

$$| k_n - p | < \frac{t\sqrt{p(1-p)}}{\sqrt{n}}. \tag{8.9}$$

Розв’яжемо останню нерівність (8.9) відносно $p.$

Якщо $k_n > p,$ то $k_n - p < \dfrac{t\sqrt{p(1-p)}}{\sqrt{n}}.$ Піднісши до квадрата, отримаємо:

$$k_n^2 - 2k_n p + p^2 < \frac{t^2}{n}(p - p^2) \quad \text{або} \quad \left(\frac{t^2}{n} + 1\right)p^2 - 2\left(k_n + \frac{t^2}{2n}\right)p + k_n^2 < 0.$$

Дискримінант тричлена нерівності, точніше

$$\frac{1}{4}\Delta=\left(k_n+\frac{t^2}{2n}\right)^2-k_n^2\left(\frac{t^2}{n}+1\right)=\frac{t^2}{n}\left(\frac{t^2}{4n}+k_n(1-k_n)\right)>0,\quad (0<k_n<1)$$

додатна, тому корені дійсні і різні.

Знаходимо менший корінь p_1:

$$p_1=\frac{n}{t^2+n}\left[k_n+\frac{t^2}{2n}-t\sqrt{\frac{k_n(1-k_n)}{n}+\left(\frac{t}{2n}\right)^2}\right] \tag{8.10}$$

і більший корінь p_2:

$$p_2=\frac{n}{t^2+n}\left[k_n+\frac{t^2}{2n}+t\sqrt{\frac{k_n(1-k_n)}{n}+\left(\frac{t}{2n}\right)^2}\right] \tag{8.11}$$

нерівності, тоді шуканий надійний інтервал

$$p_1<p<p_2. \tag{8.12}$$

Приклад 4. Знайти надійний інтервал для оцінки ймовірності p настання деякої події з надійністю $\gamma=0,95,$ якщо у 80 спробах ця подія настала 20 разів.

Розв'язання. За умовою $\gamma=0,95,$ $n=80,$ $m=20.$ Корінь рівняння $2\Phi(t)=\gamma$ знаходимо з табл. Д.2:

$$\Phi(t)=0,475,\quad t=1,96.$$

Обчислимо відносну частоту настання події $k_n=\frac{m}{n}=\frac{1}{4}=0,25.$

За формулами (8.10) і (8.11) знаходимо $p_1=0,167;$ $p_2=0,353.$

Отже, шуканий надійний інтервал $0,167<p<0,353.$

Зауваження. Із формули (8.4) можна знаходити мінімальний обсяг вибірки, за якого із вказаними надійністю і точністю виконується рівність $\overline{x}=a,$ якщо випадкова величина розподілена за нормальним законом із параметром $\sigma.$

Приклад 5. Знайти мінімальний обсяг вибірки, за якого із надійністю 0,95 точність оцінки математичного сподівання нормально розподіленої генеральної сукупності за вибірковою середньою дорівнює 0,2, якщо відоме середнє квадратичне відхилення генеральної сукупності $\sigma=1,5.$

Розв'язання. Із формули (8.4) видно, що точність оцінки дорівнює $\frac{t\sigma}{\sqrt{n}}$, яка за умовою дорівнює 0,2, тобто $\frac{t\sigma}{\sqrt{n}} = 0,2$, тому

$$n = \frac{t^2\sigma^2}{0,2^2}; \;\; \gamma = 0,95; \;\; \Phi(t) = 0,475.$$

Згідно з таблицею $t = 1,96$, тоді $n = \frac{1,96^2 \cdot 1,5^2}{0,2^2} = 217$ – мінімальний обсяг вибірки.

8.2. Вибіркові характеристики зв'язку

8.2.1. Обробка вибірки методом найменших квадратів

Нехай відома функціональна залежність між випадковими величинами X та Y вигляду

$$Y = f(X, \alpha_1, \alpha_2, \ldots, \alpha_m)$$

з невідомими параметрами $\alpha_1, \alpha_2, \ldots, \alpha_m$.

Наприклад, можна розглядати залежність між собівартістю продукції (ознака Y) та обсягом продукції (ознака X) деякої кількості однотипних підприємств.

Нехай унаслідок n незалежних випробувань отримано варіанти ознак Y та X, які оформлено у статистичній таблиці вигляду

Номери випробувань	1	2	…	k	…	n
X	x_2	x_2	…	x_k	…	x_n
Y	y_1	y_2	…	y_k	…	y_n

Для визначення оцінок параметрів функціональної залежності $\alpha_1, \alpha_2, \ldots, \alpha_m$ згідно з методом найменших квадратів найімовірніші значення параметрів $\alpha_1, \alpha_2, \ldots, \alpha_m$ мають давати мінімум функції

$$\Phi = \sum_{k=1}^{n} \left(y_k - f(x_k, \alpha_1, \alpha_2, \ldots, \alpha_m) \right)^2.$$

Якщо функція $f(x_k, \alpha_1, \alpha_2, \ldots, \alpha_m)$ має неперервні частинні похідні першого порядку відносно невідомих параметрів $\alpha_1, \alpha_2, \ldots, \alpha_m$, то необхідною умовою існування мінімуму функції Ф буде система m рівнянь з m невідомими

$$\frac{\partial \Phi}{\partial \alpha_k} = 0, \ k = 1, \ 2, \ \ldots, \ m.$$

Визначення функціональної залежності між випадковими величинами X та Y з використанням даних випробувань називають ***вирівнюванням емпіричних даних*** уздовж кривої $y = f(x, \alpha_1, \alpha_2, \ldots, \alpha_m)$.

8.2.2. Основні характеристики зв'язку

Нехай X і Y – випадкові величини, зв'язок між якими потрібно вивчити. В результаті n випробувань отримали n точок $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. Вибіркові середні та дисперсії позначимо $\overline{x}, \overline{y}, \overline{s}_x^{\,2}, \overline{s}_y^{\,2}$. За оцінку коефіцієнта кореляції $\rho = \rho(X, Y)$ можна взяти

$$r = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\overline{s}_x \overline{s}_y} = \frac{\sum\limits_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{n\overline{s}_x \overline{s}_y} \tag{8.13}$$

у разі невеликої кількості випробувань.

Аналогічно визначаються рівняння вибіркових прямих регресії: прямої регресії Y на X

$$y - \overline{y} = r\frac{\overline{s}_y}{\overline{s}_x}(x - \overline{x})$$

і прямої регресії X на Y

$$x - \overline{x} = r\frac{\overline{s}_x}{\overline{s}_y}(y - \overline{y}).$$

Прямі регресії доцільно шукати у тому разі, коли точки (x_i, y_i) $(i = 1, 2, \ldots, n)$ групуються навколо деякої прямої.

Якщо кількість випробувань велика, то для спрощення обчислень дані групують і застосовують метод умовних варіант. Вважатимемо, що дані уже згруповані і що пара чисел $(x_i,\ y_j)$ спостерігалась n_{ij} разів $(i=1,\ 2,\ \ldots,\ l;$ $j=1,\ 2,\ \ldots,\ k)$. Ці дані записують у вигляді кореляційної таблиці:

$Y \backslash X$	x_1	x_2	…	x_l	$\sum$
y_1	n_{11}	n_{21}	…	n_{l1}	n_1
y_2	n_{12}	n_{22}	…	n_{l2}	n_2
…	…	…	…	…	…
y_k	n_{1k}	n_{2k}	…	n_{lk}	n_k
$\sum$	m_1	m_2	…	m_l	n

де $m_i=\sum_{j=1}^{k}n_{ij};\ n_j=\sum_{i=1}^{l}n_{ij};\ \sum_{i=1}^{l}m_i=\sum_{j=1}^{k}n_j=n$ (n – кількість усіх спостережень). Тоді формула обчислення r набуває вигляду

$$r=\frac{\sum\limits_{i,j}n_{ij}x_iy_j-n\overline{x}\overline{y}}{n\overline{s}_x\overline{s}_y}.$$

Нехай $x_{i+1}-x_i=h_1(i=1,\ 2,\ \ldots,\ l-1);$ $y_{j+1}-y_j=h_2(j=1,\ 2,\ \ldots,\ k-1)$ – кроки таблиці (варіанти рівновіддалені). Введемо умовні варіанти

$$u_i=\frac{x_i-C_1}{h_1};\ \ \nu_j=\frac{y_j-C_2}{h_2},$$

де C_1, C_2 − умовні нулі.

Тоді $x_i=h_1u_i+C_1;\ y_j=h_2\nu_j+C_2$ і формула набуває вигляду

$$r=\frac{\sum\limits_{i,j}n_{ij}u_i\nu_j-n\overline{u}\overline{\nu}}{n\overline{s}_u\overline{s}_\nu}. \qquad (8.14)$$

З огляду на можливості сучасної обчислювальної техніки формулою (8.14) користуються набагато рідше, ніж формулою (8.13).

8.2.3. Метод найменших квадратів для прямих регресії

Якщо $y = \alpha_1 x + \alpha_0$ – рівняння вибіркової прямої регресії Y на X, то згідно з означенням

$$\Phi(\alpha_1,\ \alpha_0) = \sum_{i=1}^{n}\left(y_i - \alpha_1 x_i - \alpha_0\right)^2 \to \min.$$

Диференціюючи за α_1 і α_0 функцію $\Phi(\alpha_1,\ \alpha_0)$, отримуємо систему рівнянь для визначення α_1 і α_0, з якої

$$\begin{cases} \alpha_0 = \overline{y} - \alpha_1 \overline{x}; \\ \alpha_1 = \dfrac{\sum\limits_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum\limits_{i=}^{n} x_i^2 - n\overline{x}^2} = \dfrac{\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}. \end{cases}$$

Якщо значення x_i відомі без похибок, а значення y_i незалежні й точні, оцінку дисперсії (похибку вимірювань) величин y_i визначають за формулою $\sigma^2 = \dfrac{\Phi_{\min}}{n-2}$, де $\Phi_{\min} = \sum\limits_{i=1}^{n}\left(y_i - \alpha_1 x_i - \alpha_0\right)^2$.

Оцінки дисперсій коефіцієнтів α_1 і α_0 визначають за формулами

$$\sigma_{\alpha_0}^2 = \frac{\sum\limits_{i=1}^{n} x_i^2}{n\sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2} \cdot \frac{\Phi_{\min}}{n-2},\quad \sigma_{\alpha_1}^2 = \frac{n}{n\sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2} \cdot \frac{\Phi_{\min}}{n-2}.$$

Якщо величини y_i мають нормальний розподіл, то для коефіцієнтів α_1 і α_0 справджуються такі надійні інтервали:

$$\left(\alpha_k - t(n-2,\gamma)\sigma_{\alpha_k},\ \alpha_k + t(n-2,\gamma)\sigma_{\alpha_k}\right),\ k = 0,\ 1,$$

де α_k – оцінки, отримані методом найменших квадратів, а число $t(n-2,\gamma)$ знаходять відповідно до таблиці Стьюдента за кількості степенів вільності $k = n-2$ і $\gamma = P\left(|t| < t(n-2,\gamma)\right)$.

Приклад 6. Для торгових агентів компанії зареєстровано дві ознаки: X –

витрати на представництво, Y – обсяг продажів товарів за певний час (в тис. грн). Дані наведено в таблиці:

x_i	6,5	6,5	6,2	6,7	6,9	6,5	6,1	6,7
y_i	105	125	110	120	140	135	95	130

Знайти вибірковий коефіцієнт кореляції, рівняння прямої регресії Y на X, похибку вимірювань, надійні інтервали для коефіцієнтів прямої регресії за $\gamma = 0{,}95$. Об'єм вибірки n = 8.

Розв'язання. Проміжні результати обчислень заносимо у таку таблицю:

№	x_i	y_i	$x_i y_i$	x_i^2	y_i^2	$\widehat{y}_i$	$\left(y_i - \widehat{y}_i\right)^2$
1	6,5	105	682,5	42,25	11025	119,4120	207,7057
2	6,5	125	812,5	42,25	15625	119,4120	31,2257
3	6,2	110	682,0	38,44	12100	105,2981	22,1079
4	6,7	120	804,0	44,89	14400	128,8212	77,8136
5	6,9	140	966,0	47,61	19600	138,2305	3,1311
6	6,5	135	877,5	42,25	18225	119,4110	242,9857
7	6,1	95	579,5	37,21	9025	100,5934	31,2861
8	6,7	130	871,0	44,89	16900	128,8212	1,3896
$\sum$	52,1	960	6275	339,79	116900	–	617,6454

$$\overline{x} = \frac{1}{8} \cdot 52{,}1 = 6{,}5125; \quad \overline{y} = \frac{1}{8} \cdot 960 = 120;$$

$$\overline{s}_x^2 = \frac{1}{8} \cdot 339{,}79 - 6{,}5125^2 = 0{,}0611; \quad \overline{s}_x = 0{,}2472;$$

$$\overline{s}_y^2 = \frac{1}{8} \cdot 116900 - 120^2 = 212{,}5; \quad \overline{s}_y = 14{,}5774;$$

$$r = \frac{6275 - 8 \cdot 6{,}5125 \cdot 120}{8 \cdot 0{,}2472 \cdot 14{,}5774} = 0{,}7978.$$

Записуємо пряму регресії Y на X за формулою $y - \overline{y} = r\dfrac{\overline{s}_y}{\overline{s}_x}(x - \overline{x})$, тому

$$\alpha_1 = r\frac{\overline{s}_y}{\overline{s}_x} = 0{,}7978 \cdot \frac{14{,}5774}{0{,}2472} = 47{,}0463;$$

$$\alpha_0 = \overline{y} - \alpha_1\overline{x} = 120 - 47,0463 \cdot 6,5125 = -186,3890;$$

$$\widehat{y} = 47,0463x - 186,3890.$$

Обчислимо:

$$\widehat{y}_1 = 47,0463 \cdot 6,5 - 186,389 = 119,4120;$$

$$\widehat{y}_2 = 47,0463 \cdot 6,5 - 186,389 = 119,4120;$$

$$\widehat{y}_3 = 47,0463 \cdot 6,2 - 186,389 = 105,2981;$$

$$\widehat{y}_4 = 47,0463 \cdot 6,7 - 186,389 = 128,8212;$$

$$\widehat{y}_5 = 47,0463 \cdot 6,9 - 186,389 = 138,2305;$$

$$\widehat{y}_6 = 47,0463 \cdot 6,5 - 186,389 = 119,4120;$$

$$\widehat{y}_7 = 47,0463 \cdot 6,1 - 186,389 = 100,5934;$$

$$\widehat{y}_8 = 47,0463 \cdot 6,7 - 186,389 = 128,8212;$$

$$\Phi_{\min} = 617,6454.$$

Знаходимо

$$\Delta = n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 = 8 \cdot 339,79 - 52,1^2 = 3,91.$$

Тоді $\sigma^2 = \dfrac{617,6454}{8-2} = 102,94044;$

$$\sigma_{\alpha_0}^2 = \frac{339,79}{3,91} \cdot 102,9409 = 8945,8537;\ \sigma_{\alpha_0} = 94,5825;$$

$$\sigma_{\alpha_1}^2 = \frac{8}{3,91} \cdot 102,9409 = 210,6208;\ \sigma_{\alpha_1} = 14,1278.$$

За таблицею розподілу Стьюдента знаходимо $t(6;0,95) = 2,45$. Тоді з імовірністю 0,95

$$-186,389 - 2,45 \cdot 94,5825 < \alpha_0 < -186,389 + 2,45 \cdot 94,5825,$$

тобто $-418,1161 < \alpha_0 < 45,3381;$

$$47,0463 - 2,45 \cdot 14,1278 < \alpha_1 < 47,0463 + 2,45 \cdot 14,1278,$$

тобто $12,4332 < \alpha_1 < 81,6594.$

8.3. Коефіцієнт кореляції рангів

Нехай елементи генеральної сукупності мають якісні ознаки A і B, тобто ознаки, які не піддаються кількісним оцінкам, але дають змогу порівнювати елементи між собою і, отже, розміщувати їх у порядку погіршення якості.

Нехай вибірка обсягу n містить незалежні елементи, що мають якісні ознаки A і B. Розмістимо елементи в порядку погіршення якості за ознакою A. Поставимо у відповідність елементу, що стоїть на i-му місці, число – ранг x_i, рівний порядковому номеру елемента $x_i = i$. Потім розмістимо елементи в порядку спадання якості за ознакою B і припишемо кожному з них ранг (порядковий номер) y_i, причому для зручності порівняння рангів індекс i дорівнює порядковому номеру елемента за ознакою A. Наприклад, $y_3 = 4$ означає, що за ознакою A елемент стоїть на третьому місці, а за ознакою B – на четвертому.

В результаті отримаємо дві послідовності рангів:

– за ознакою A: $x_1,\ x_2,\ \ldots,\ x_n$;

– за ознакою B: $y_1,\ y_2,\ \ldots,\ y_n$.

Позначають $d_i = x_i - y_i$ – різниця рангів.

Означення. Коефіцієнт щільності зв'язку між A і B, який визначається формулою

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

називають **коефіцієнтом кореляції рангів Спірмена.**

Якщо не можна визначити рангову відмінність деяких елементів, то беруть середній ранг. У цьому разі використовують **коефіцієнт кореляції рангів Кендела:**

$$\tau = \frac{\frac{n^3 - n}{6} - \left(T_x + T_y\right) - \sum_{i=1}^{n} d_i^2}{\sqrt{\left(\frac{n^3 - n}{6} - 2T_x\right)\left(\frac{n^3 - n}{6} - 2T_y\right)}},$$

де $T_x = T_y = \dfrac{\sum_{i=1}^{l}(t_i^3 - t_i)}{12}$, t_i – число об’єднаних рангів для X та Y.

Абсолютна величина коефіцієнта кореляції рангів не більше одиниці:

$$|\rho| \le 1.$$

Що ближче до нуля $|\rho|$, то залежність між якісними ознаками *A* і *B* менша.

Приклад 7. Два інспектори *A* і *B* перевірили 12 водіїв на швидкість реакції і розмістили їх у порядку погіршення реакції (в дужках вказано порядкові номери водіїв з однаковою реакцією).

A	1	(2, 3, 4)	5	(6, 7, 8)	9	10	11	12
B	3	1 2 6	4	5 7 8	11	10	9	12

Знайти коефіцієнти кореляції рангів Спірмена і Кендела.

Розв’язання. Ранги водіїв з номерами 2, 3, 4 дорівнюють $\dfrac{2+3+4}{3} = 3$, водіїв з номерами 6, 7, 8 відповідно $\dfrac{6+7+8}{3} = 7$. Запишемо послідовність рангів

A	1	3	3	3	5	7	7	7	9	10	11	12
B	3	1	2	6	4	5	7	8	11	10	9	12

$$\sum d_i^2 = 4+4+1+9+1+4+1+4+4 = 32.$$

Коефіцієнт кореляції рангів Спірмена

$$\rho = 1 - \frac{6 \cdot 32}{12^3 - 12} = 0{,}8881.$$

У першій послідовності маємо трьох водіїв з однаковими рангами 3 і трьох водіїв з однаковими рангами 7. Тоді

$$T_x = \frac{(3^3 - 3) + (3^3 - 3)}{12} = 4,\ T_y = 0.$$

Коефіцієнт кореляції рангів Кендела

$$\tau = \frac{\dfrac{12^3 - 12}{6} - (4+0) - 32}{\sqrt{\left(\dfrac{12^3 - 12}{6} - 8\right)\left(\dfrac{12^3 - 12}{6} - 0\right)}} = \frac{250}{281{,}9716} = 0{,}8866.$$

Розділ 9. Статистична перевірка гіпотез

На практиці часто необхідно знати закон розподілу генеральної сукупності. Якщо закон розподілу невідомий, але є припущення його певного вигляду, тоді висувають гіпотезу, що випадкова величина має цей розподіл і виникає потреба у перевірці гіпотези. Розглянемо одну з основних задач математичної статистики – перевірку правдоподібності гіпотез. Для цього потрібно використовувати так звані критерії узгодження, з яких зупинимось на критеріях Пірсона та Романовського.

9.1. Перевірка статистичних гіпотез

Означення. **Статистичною гіпотезою** називають будь-яке твердження про вигляд або властивості розподілу випадкових величин, що спостерігаються в експерименті.

Означення. **Нульовою** (основною) називають висунуту гіпотезу H_0. **Конкуруючою** (альтернативною) називають гіпотезу H_1, яка суперечить нульовій.

Означення. **Простою** називають гіпотезу, що містить лише одне твердження. **Складною** називають гіпотезу, яка складається зі скінченної або нескінченної кількості простих.

Висунута статистична гіпотеза підлягає перевірці, яка здійснюється за даними вибірки. Помилка першого роду полягає в прийнятті альтернативи (відхилення нульової гіпотези) в умовах, коли справджується нульова гіпотеза.

Помилка другого роду полягає в прийнятті нульової гіпотези (відхилення альтернативи) в умовах, коли справджується альтернативна гіпотеза.

Для перевірки нульової гіпотези вибирається деяка випадкова величина K, розподіл якої відомий, і вона називається **статистичним критерієм** перевірки нульової гіпотези.

Означення. **Критичною областю** називають множину значень критерію, за яких нульову гіпотезу відхиляють. ***Областю прийняття гіпотези*** називають множину значень критерію, за яких нульову гіпотезу приймають.

Означення. **Критичними точками** $k_{\hat{e}\eth}$ називають точки, які відокремлюють критичну область від області прийняття гіпотези.

Означення. **Правосторонньою** називають критичну область, що визначається нерівністю $K > k_{\hat{e}\eth}$.

Для визначення правосторонньої критичної області достатньо знайти критичну точку $k_{\hat{e}\eth}$. Для цього задають достатньо малу ймовірність α; число α називається рівнем значущості критерію. Тоді знаходять критичну точку, виходячи з того, що за умови справедливості нульової гіпотези виконується рівність

$$P\left(K > k_{\hat{e}\eth}\right) = \alpha. \qquad (9.1)$$

Порядок дій у разі перевірки статистичних гіпотез.

Для перевірки правильності основної статистичної гіпотези H_0 потрібно:

1) визначити гіпотезу H_1, альтернативну до гіпотези H_0;

2) обрати статистичну характеристику перевірки;

3) визначити допустиму ймовірність похибки першого роду, тобто рівень значущості α;

4) знайти за відповідною таблицею критичну область (критичну точку) для обраної статистичної характеристики.

До критичної області належать такі значення статистичної характеристики, за яких гіпотеза H_0 відхиляється на користь альтернативної гіпотези H_1.

Зауваження. Якщо гіпотеза H_0 правильна, то з імовірністю α значення вибіркової функції будуть належати критичній області.

9.2. Критерії згоди

Критерії, в яких формулюється лише одна гіпотеза H_0, і потрібно перевірити, чи узгоджуються статистичні дані з цією гіпотезою, називають ***критеріями згоди***.

9.2.1. Критерій згоди χ^2 про вигляд розподілу

Критерій згоди χ^2 можна використовувати для будь-яких розподілів. Розглянемо його для перевірки гіпотези H_0: генеральна сукупність має нормальний закон розподілу $N\left(a,\ \sigma^2\right)$, де $a=\overline{x},\ \sigma^2=s^2$.

Розбиваємо числову вісь на r інтервалів, що не перетинаються: $h_1,\ h_2,\ \ldots,\ h_r$; n_i – кількість елементів вибірки, що потрапили в інтервал h_i $(i=1,\ 2,\ \ldots,\ r)$. Позначимо через $p_i=P\left(X\in h_i\right)$, тоді np_i – теоретичні частоти. За формулою обчислення ймовірності $P(\alpha<X<\beta)=\Phi\left(\dfrac{\beta-a}{\sigma}\right)-\Phi\left(\dfrac{\alpha-a}{\sigma}\right)$ для нормального розподілу, якщо $a=\overline{x},\ \sigma=s,\ h_i=(x_i,\ x_{i+1})$, отримаємо:

$$p_i=P\left(x_i<X<x_{i+1}\right)=\Phi\left(\frac{x_{i+1}-\overline{x}}{s}\right)-\Phi\left(\frac{x_i-\overline{x}}{s}\right).$$

За критерій перевірки гіпотези H_0 беруть величину

$$\chi^2=\sum_{i=1}^{r}\frac{\left(n_i-np_i\right)^2}{np_i}. \tag{9.2}$$

Число k степенів вільності дорівнює $k=r-l-1$, де l – кількість параметрів розподілу. Для нормального розподілу $k=r-3$.

Якщо нульова гіпотеза справедлива за даним рівнем значущості α і кількістю степенів вільності k відповідно до таблиці розподілу χ^2 знаходимо критичну точку $\chi^2_{k;\,\alpha}$ з рівності $P\left(\chi^2>\chi^2_{k;\,\alpha}\right)=\alpha$. Якщо $\chi^2<\chi^2_{k;\,\alpha}$, то гіпотезу H_0 приймають, якщо $\chi^2\geq\chi^2_{k;\,\alpha}$, то гіпотезу H_0 відхиляють.

Приклад. Згідно з результатами спостережень, наведених у таблиці, за допомогою критерію згоди χ^2 перевірити гіпотезу про те, що випадкова величина має нормальний розподіл $(\alpha=0{,}05;\ n=200)$.

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
n_i	6	9	26	25	30	26	21	24	20	8	5

Розв'язання. Обчислимо за вибіркою оцінки параметрів нормального розподілу ($\tilde{N} = 1{,}1;\ h = 0{,}2$).

№	x_i	n_i	u_i	$n_i u_i$	$n_i u_i^2$
1	0,3	6	–4	–24	96
2	0,5	9	–3	–27	81
3	0,7	26	–2	–52	104
4	0,9	25	–1	–25	25
5	1,1	30	0	$A_1 = -128$	0
6	1,3	26	1	26	26
7	1,5	21	2	42	84
8	1,7	24	3	72	216
9	1,9	20	4	80	320
10	2,1	8	5	45	225
11	2,3	5	6	30	180
				$A_2 = 295$	
$\sum$		200		167	1357

$$\overline{u} = \frac{1}{200} \cdot 167 = 0{,}835; \quad \overline{x} = h\overline{u} + C = 0{,}2 \cdot 0{,}835 + 1{,}1 = 1{,}267;$$

$$\overline{s}_u^2 = \frac{1}{200} \cdot 1357 - 0{,}835^2 = 6{,}0878;$$

$$\overline{s}^2 = h^2 \overline{s}_u^2 = 0{,}04 \cdot 6{,}0878 = 0{,}2435;$$

$$s^2 = \frac{200}{199} \cdot \overline{s}^2 = 0{,}2447;\ s = 0{,}4947.$$

Обчислимо теоретичні частоти за формулами $n_i = \dfrac{nh}{s}\varphi(z_i)$, де $z_i = \dfrac{x_i - \overline{x}}{s}$, $\varphi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$. Для зручності визначення теоретичних частот складемо розрахункову таблицю:

№	x_i	$z_i = \dfrac{x_i - \bar{x}}{s}$	$\varphi(z_i)$	$np_i = \dfrac{nh}{s}\varphi(z_i)$
1	0,3	–1,95	0,0596	4,82
2	0,5	–1,55	0,1200	9,70
3	0,7	–1,15	0,2059	16,65
4	0,9	–0,74	0,3034	24,53
5	1,1	–0,34	0,3765	30,44
6	1,3	0,07	0,3980	32,18
7	1,5	0,47	0,3572	28,88
8	1,7	0,88	0,2709	21,90
9	1,9	1,28	0,1758	14,21
10	2,1	1,68	0,0973	7,87
11	2,3	2,09	0,0449	3,63
$\sum$	–	–	–	194,81

Порівнюємо теоретичні та статистичні частоти. Для знаходження χ^2 складаємо розрахункову таблицю:

№	n_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$\dfrac{(n_i - np_i)^2}{np_i}$
1	6	4,82	1,18	1,3924	0,289
2	9	9,70	–0,7	0,4900	0,051
3	26	16,65	9,35	87,423	5,251
4	25	24,53	0,47	0,221	0,009
5	30	30,44	–0,44	0,1936	0,006
6	26	32,18	6,18	38,1924	1,187
7	21	28,88	–7,88	62,0944	2,150
8	24	21,90	2,1	4,41	0,201
9	20	14,21	5,79	33,5241	2,359
10	8	7,87	0,13	0,0169	0,002
11	5	3,63	1,37	1,8769	0,517
$\sum$	200	194,81	–	–	12,022

Отже, $\chi^2 = 12,022$. Кількість степенів вільності $k = 11 - 3 = 8$. За таблицею розподілу χ^2 знаходимо критичну точку $\chi^2_{8;\,0,05} = 15,51$. Оскільки $\chi^2 = 12,022 < \chi^2_{8;\,0,05} = 15,51$, гіпотезу приймають.

9.2.2. Критерій незалежності χ^2

Нехай дані, одержані в результаті спостережень над парою випадкових величин X і Y, подано у вигляді кореляційної таблиці:

$Y \backslash X$	x_1	x_2	…	x_l	$\sum$
y_1	n_{11}	n_{21}	…	n_{l1}	n_1
y_2	n_{12}	n_{22}	…	n_{l2}	n_2
…	…	…	…	…	…
y_k	n_{1k}	n_{2k}	…	n_{lk}	n_k
$\sum$	m_1	m_2	…	m_l	n

Потрібно перевірити гіпотезу H_0 про незалежність випадкових величин X і Y.

За критерій перевірки гіпотези H_0 беруть величину

$$\chi_n^2 = \sum_{i=1}^{l}\sum_{j=1}^{k}\frac{\left(n_{ij} - \frac{m_i n_j}{n}\right)^2}{\frac{m_i n_j}{n}}. \tag{9.3}$$

Відповідно до таблиці розподілу χ^2 за даним числом α і кількістю степенів вільності $m = (l-1)(k-1)$ знаходимо число $\chi^2_{m;\alpha} : P\left\{\chi^2(m) \geq \chi^2_{m;\alpha}\right\} = \alpha$. Якщо $\chi_n^2 \geq \chi^2_{m;\alpha}$, то гіпотезу H_0 відхиляють, якщо $\chi_n^2 < \chi^2_{m;\alpha}$, то гіпотезу H_0 приймають.

Приклад. Проведено 200 спостережень над випадковими величинами X і Y:

$Y \backslash X$	1	2	$\sum$
1	25	52	77
2	50	41	91
3	25	7	32
$\sum$	100	100	200

Перевірити за допомогою критерію χ^2 незалежність випадкових величин X і Y за $\alpha = 0,05.$

Розв'язання. Обчислюємо за формулою

$$\chi_n^2 = \frac{(25-38,5)^2}{38,5} + \frac{(52-38,5)^2}{38,5} + \frac{(50-45,5)^2}{45,5} + \frac{(41-45,5)^2}{45,5} + \frac{(25-16)^2}{16} +$$

$$+\frac{(7-16)^2}{16} = 20,48.$$

Кількість степенів вільності $m = (2-1)(3-1) = 2,$ $\chi^2_{2;\,0,05} = 5,99.$ Оскільки 20,4 > 5,99, то гіпотезу незалежності відхиляють.

9.3. Перевірка гіпотез про рівність математичних сподівань та дисперсій для нормальних сукупностей

9.3.1. Гіпотеза про рівність математичних сподівань за відомих дисперсій

Нехай X і Y – незалежні випадкові величини, кожна з яких має нормальний розподіл $N(a_1,\ \sigma_x^2)$ і $N(a_2,\ \sigma_y^2)$. В результаті спостережень цих величин отримано дві вибірки $x_1,\ x_2,\ \ldots,\ x_n$ та $y_1,\ y_2,\ \ldots,\ y_m$ обсягів n і m. Потрібно перевірити гіпотезу $H_0 : a_1 = a_2$ проти альтернативної гіпотези H_1: $a_1 \neq a_2;$ $\sigma_x,$ σ_y – відомі.

За критерій перевірки нульової гіпотези беруть випадкову величину

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_x^2}{n} + \dfrac{\sigma_y^2}{m}}}. \tag{9.4}$$

Критерій Z – нормована випадкова величина, яка має нормальний розподіл, тому двостороння критична область симетрична відносно нуля. Критичну точку z_α знаходиться з рівності $\Phi\left(z_\alpha\right) = \dfrac{1-\alpha}{2}$. За реалізації вибірок обчислюють

$$Z = \frac{\overline{x} - \overline{y}}{\sqrt{\dfrac{\sigma_x^2}{n} + \dfrac{\sigma_y^2}{m}}}.$$

Якщо $|Z| < z_\alpha$, то приймають гіпотезу H_0, якщо $|Z| \geq z_\alpha$, то приймають гіпотезу H_1.

9.3.2. Гіпотеза про рівність математичних сподівань за невідомих дисперсій

Нехай потрібно перевірити гіпотезу H_0: $a_1 = a_2$ проти альтернативної гіпотези H_1: $a_1 \neq a_2$, але $\sigma_x^2 = \sigma_y^2 = \sigma^2$ і величина σ^2 – невідома.

За критерій перевірки нульової гіпотези беруть випадкову величину

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}}. \tag{9.5}$$

Величина T за умови правильності H_0 має розподіл Стьюдента з $k = n + m - 2$ степенями вільності.

Двостороння критична область симетрична відносно нуля, її знаходять з умови, що ймовірність прийняття гіпотези H_1 дорівнює взятому рівню значущості α, коли правильна H_0. Критичну точку $t_{n+m-2,\,\alpha}$ знаходять відповідно до таблиці розподілу Стьюдента за кількості степенів вільності $k = n + m - 2$ і $P\{|t| < t_{k,\,\alpha}\} = 1 - \alpha$. За реалізації вибірок обчислюють

$$T = \frac{\overline{x} - \overline{y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}}},$$

де $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$, $s_y^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \overline{y})^2$.

Якщо $|T| < t_{k,\,\alpha}$, то приймають гіпотезу H_0; якщо $|T| \geq t_{k,\,\alpha}$, то приймають гіпотезу H_1.

Приклад. Випадкові величини X і Y мають нормальний розподіл з рівними дисперсіями. В результаті спостережень над X та Y отримано вибірки:

X: 45, 48, 53, 44, 59, 60, 41, 43, 57;

Y: 51, 50, 42, 44, 39, 40, 48, 38, 59, 55, 55, 51.

Чи можна вважати, що випадкові величини мають однакові математичні сподівання? Вважати похибку першого роду $\alpha = 0,05$.

Розв'язання. За умовою $n=9$, $m=11$.

$$\overline{x} = \frac{1}{9}(45+48+53+44+59+60+41+43+57) = 50,\ s_x^2 = 54,25;$$

$$\overline{y} = 47,\ s_y^2 = 47,8.$$

$$|T| = \frac{|50-47|}{\sqrt{\left(\frac{1}{9}+\frac{1}{11}\right)\frac{8\cdot 54,25+10\cdot 47,8}{2+11-2}}} = 0,938.$$

Кількість степенів вільності $k = 9+11-2 = 18$; $\gamma = 1-\alpha = 0,95$, тоді $t_{18;\,0,05} = t(18;\ 0,95) = 2,10$. Оскільки $0,938 < 2,10$, то можна вважати, що випадкові величини X і Y мають однакові математичні сподівання.

9.3.3. Гіпотеза про рівність дисперсій за невідомих математичних сподівань

Нехай потрібно перевірити гіпотезу $H_0:\ \sigma_x^2 = \sigma_y^2$ проти альтернативної гіпотези $H_1: \sigma_x^2 > \sigma_y^2$.

За критерій перевірки гіпотези беруть відношення більшої виправленої дисперсії s_v^2 до меншої виправленої дисперсії s_u^2, тобто величину

$$F = \frac{s_v^2}{s_u^2}. \tag{9.6}$$

У разі справедливості гіпотези H_0 випадкова величина має розподіл Фішера–Снедекора з (n_1-1, n_2-1) степенями вільності, де n_1 – обсяг вибірки, за якою обчислюють s_v^2, n_2 – обсяг вибірки, за якою обчислюють s_u^2.

Зауваження. Якщо U і V – незалежні випадкові величини, що мають розподіл χ^2 з k_1 та k_2 степенями вільності, то величина

$$F = \frac{U/k_1}{V/k_2}$$

має розподіл Фішера–Снедекора з $(k_1,\ k_2)$ степенями вільності.

Правосторонню критичну область знаходять з умови, що ймовірність прийняття гіпотези H_1 дорівнює взятому рівню значущості α, коли правильна H_0:

$$P\left(F > F_{(\alpha,\, k_1,\, k_2)}\right) = \alpha,$$

де $k_1 = n_1 - 1,\ k_2 = n_2 - 1$. Критичну точку $F_{(\alpha,\, k_1,\, k_2)}$ знаходять за таблицею розподілу Фішера–Снедекора.

За вибірками обчислюємо $F = \dfrac{s_\nu^2}{s_u^2}$, тоді, якщо $F < F_{(\alpha,\, k_1,\, k_2)}$, то гіпотезу H_0 приймають, якщо $F \geq F_{(\alpha,\, k_1,\, k_2)}$, то гіпотезу H_0 відхиляють.

9.3.4. Гіпотеза про рівність дисперсій за відомих математичних сподівань

Цю гіпотезу перевіряють аналогічно до попередньої, але в цьому разі

$$F = \frac{s_\nu^2}{s_u^2}, \tag{9.7}$$

де $s_\nu^2 = \dfrac{1}{n_1}\sum\limits_{i=1}^{n_1}(x_i - a_1)^2,\ s_u^2 = \dfrac{1}{n_2}\sum\limits_{i=1}^{n_2}(y_i - a_2)^2$.

Якщо правильна гіпотеза H_0: $\sigma_x^2 = \sigma_y^2$, то випадкова величина має розподіл Фішера–Снедекора з (k_1, k_2) степенями вільності. Правосторонню критичну область знаходять з умови

$$P\left(F > F_{(\alpha,\, k_1,\, k_2)}\right) = \alpha,$$

коли правильна гіпотеза H_0. Критичну точку $F_{(\alpha,\, k_1,\, k_2)}$ знаходять за таблицею розподілу Фішера–Снедекора.

За вибірками обчислюємо значення $F = \dfrac{s_\nu^2}{s_u^2}$, тоді, якщо $F < F_{(\alpha,\, k_1,\, k_2)}$, то гіпотезу H_0 приймають, якщо $F \geq F_{(\alpha,\, k_1,\, k_2)}$, то гіпотезу H_0 відхиляють.

Приклад. За двома вибірками з нормальних сукупностей обсягу $n_1 = 11,\ n_2 = 15$ підраховано $s_1^2 = 0,76$ і $s_2^2 = 0,36.$ За $\alpha = 0,05$ перевірити гіпотезу про рівність дисперсій цих двох нормальних сукупностей.

Розв'язання. Враховуючи $s^2 = \dfrac{n}{n-1}\overline{s}^2,$ обчислюємо:

$$s_1^2 = \frac{11 \cdot 0,76}{10} = 0,836;\ s_2^2 = \frac{15 \cdot 0,36}{14} = 0,386.$$

Тоді $F = \dfrac{s_1^2}{s_2^2} = \dfrac{0,836}{0,386} = 2,17;\ F_{(0,05;10;14)} = 2,6.$

Оскільки $2,17 < 2,6$, гіпотезу про рівність дисперсій приймають.

Приклади завдань для самостійної роботи

Завдання 1. Класичне означення ймовірності

1. Серед 25 фірм, з яких 10 українських, а інші російські, розігрується 5 урядових контрактів. Вважається, що кожна фірма має рівні шанси на отримання контракту. Знайти ймовірність того, що принаймні дві українські фірми виграють контракт.

2. У фірмі 10 співробітників (6 чоловіків і 4 жінки) претендують на заміщення трьох вакансій. Вважають, що всі кандидатури мають рівні шанси на зайняття цих вакансій. Знайти ймовірність того, що жінки не займуть жодної вакансії.

3. У групі з 12 бізнесменів тільки 8 мають досвід роботи у запропонованій новій галузі. Для проекту потрібно відібрати 4 особи. За припущення, що відбір претендентів проводять навмання, знайти ймовірність того, що в команду з чотирьох чоловік потраплять всі, хто має досвід роботи.

4. Комплект містить 7 виробів першого сорту, 6 – другого сорту і 2 вироби – третього сорту. Навмання обирають 5 виробів. Знайти ймовірність того, що серед них не виявиться виробів третього сорту.

5. Із 15 рейсів, що виконуються з аеропорту протягом доби, 60 % рейсів виконуються на власному літаковому парку. Знайти ймовірність того, що з вибраних навмання 5 рейсів рівно 3 виконуються на власному парку.

6. 12 виробів, серед яких 4 нестандартних, випадковим способом розбиваються на дві рівні партії. Знайти ймовірності того, що: *а*) у кожній партії буде рівна кількість нестандартних виробів; *б*) усі нестандартні вироби будуть в одній партії.

7. У конкурсі газети бере участь 12 чоловіків та 8 жінок. Є два призових місця. За припущення, що відбір претендентів ведуть навмання, яка ймовірність того, що обидва місця займуть жінки?

8. З 10 літаків, що прибувають в аеропорт протягом доби, 80 % мають повне комерційне завантаження. Знайти ймовірність того, що серед п'яти випадковим способом узятих літаків тільки 4 мають повне завантаження.

9. В авіакасі було 15 квитків, серед яких 6 квитків до пункту *А*. До кінця зміни продано 8 квитків. Знайти ймовірність того, що в касі не залишилося квитків до пункту *А*, якщо ймовірність продажу кожного квитка однакова.

10. Партія з 30 виробів містить 10 % браку. Знайти ймовірність того, що серед 7 виробів, узятих випадково: *а*) тільки 2 бракованих; *б*) жодного бракованого.

Відповіді: **1.** 0,9435. **2.** $\frac{1}{6}$. **3.** 0,141. **4.** $\frac{3}{7}$. **5.** $\frac{60}{143}$. **6.** *а*) 0,455; *б*) 0,065. **7.** $\frac{14}{95}$. **8.** $\frac{5}{9}$. **9.** $\frac{4}{715}$. **10.** *а*) 0,119; *б*) 0,436.

Завдання 2. Статистичні та геометричні ймовірності

1. До авіакаси у випадковий час у межах 10 хв звернулось 2 пасажири. Обслуговування одного пасажира триває 2 хв. Знайти ймовірність того, що пасажир, який звернувся другим, буде вимушений зачекати.

2. Відділ технічного контролю видавництва виявив 5 бракованих книжок у партії з випадково відібраних 250 книжок. Знайти відносну частоту виявлення бракованих книжок.

3. Два літаки прибувають у зону аеропорту у випадковий час між 12:00 і 12:30. Знайти ймовірність того, що літак, який прибув другим, не буде вимушений чекати дозволу на посадку, якщо чергову посадку можна здійснювати не раніше, ніж за 10 хв після попередньої.

4. У рівнобедреному трикутнику з бічною стороною 5 см та кутом при вершині 120 ° довільно розміщено квадрат зі стороною 1,5 см. Знайти ймовірність того, що навмання вибрана точка трикутника лежатиме в квадраті.

5. На відрізку довжиною 15 см випадково поставлено дві точки. Знайти ймовірність того, що відстань між цими точками не перевищує 7 см.

6. Відстань між пунктами *M* і *N* літак долає за 1 год, а потяг – за 18 год. Потяг у випадковий час протягом доби вирушає з пункту *M* до *N*. Знайти ймовірність того, що черговий літак прибуде до пункту *N* раніше від потяга, якщо між *M* і *N* виконується за розкладом один рейс літака щодоби.

7. У ромбі зі стороною 5 см та гострим кутом 60 ° лежить прямокутник зі сторонами 2 та 3 см. Знайти ймовірність того, що навмання вибрана у ромбі точка лежатиме і в прямокутнику.

8. З проміжку $[0,1]$ випадковим способом вибирають два дійсних числа. Знайти ймовірність того, що їх сума не більша одиниці, а добуток не перевищує $\frac{2}{9}$.

9. Кожне з двох дійсних додатних чисел не більше 4. Знайти ймовірність того, що їх добуток також буде не більше 4.

10. Усередині кола радіусом 5 мм розташовано прямокутник зі сторонами 4 та 6 мм. Знайти ймовірність того, що навмання вибрана у колі точка лежатиме і в прямокутнику.

Відповіді: **1.** 0,36. **2.** 0,02. **3.** $\frac{4}{9}$. **4.** 0,208. **5.** $\frac{161}{225}$. **6.** $\frac{1103}{1152}$. **7.** $\frac{\sqrt{3}}{25}$. **8.** $\frac{1+2\ln 2}{9}$. **9.** $\frac{1+\ln 4}{4}$. **10.** 0,305.

Завдання 3. Ймовірності суми та добутку подій

1. Фірма має можливість отримати два контракти. Ймовірність отримання першого контракту дорівнює 0,9, а другого – 0,8. Вважаючи ці події незалежними, знайти ймовірності подій: *а*) фірма отримає обидва контракти; *б*) фірма отримає принаймні один контракт.

2. Надійність лінії зв’язку між об’єктами (ймовірність безвідмовної роботи протягом певного часу) дорівнює 0,75. Для підвищення якості зв’язку встановлено резервну лінію надійністю 0,65. Визначити надійність зв’язку з резервною лінією.

3. Через метеорологічні умови літак було відправлено на запасний аеродром, під час наближення до якого у баках літака залишалося палива на 3 заходи на посадку. Ймовірність посадки літака за першого заходу дорівнює 0,8, за другого – 0,95, за третього – 0,995. Знайти ймовірність вдалої посадки літака.

4. Є 8 кандидатів на отримання роботи. Серед них є люди з відповідною кваліфікацією (подія A) і люди, що закінчили Видавничо-поліграфічний інститут (подія B), та інші. Їх кількості подано в таблиці.

	A	$\overline{A}$	Усього
B	1	3	4
$\overline{B}$	2	2	4
Усього	3	5	8

Усі кандидати мають рівні шанси на отримання роботи. Знайти ймовірність того, що роботу отримає людина без кваліфікації, або яка закінчила Видавничо-поліграфічний інститут.

5. Імовірність виготовлення виробу вищого сорту на першому верстаті становить 0,7, на другому – 0,8. На першому верстаті виготовлено 2 вироби, на другому – 3. Знайти ймовірність того, що всі вироби належать до вищого сорту.

6. Відомо, що в деякому регіоні 40 % компаній мають у штаті юриста і 80 % компаній мають у штаті економіста. Вважаємо, що ці дві події незалежні. Знайти ймовірність того, що фірма має в штаті економіста і юриста.

7. Є два ринки цінних паперів. Інвестиційна фірма направила на обидва ринки водночас 260 акцій різноманітної якості. У визначений день на першому ринку зросли ціни 197 акцій (подія A), а на другому ринку зросли ціни 191 акції (подія B). Одночасно на обох ринках зросли ціни 165 акцій. Інші акції не піднялися в ціні. Результати торгів наведено в таблиці.

	A	$\overline{A}$	Разом
B	165		191
$\overline{B}$			
Разом	197		260

Заповнити порожні місця в таблиці. Знайти ймовірність того, що зросла ціна акцій на першому ринку, якщо відомо, що зросла ціна акцій на другому ринку.

8. На фірмі опитано 100 службовців із метою вивчення стану транспортного обслуговування. Виявилося, що 70 осіб користуються метро (подія A).

Іншими видами транспорту користуються 40 осіб (подія *B*), а 20 осіб одночасно користуються метро та іншими видами транспорту. Опитування занесено в таблицю.

	A	$\overline{A}$	Разом
B	20		40
$\overline{B}$			
Разом	70		100

Заповнити порожні місця в таблиці. Знайти ймовірність того, що службовець користується метро або іншими видами транспорту.

9. Радіостанція аеропорту надсилає 3 повідомлення для літака. Імовірність підслуховування розвідувальним агентством першого повідомлення дорівнює 0,6; другого – 0,65; третього – 0,7. Знайти ймовірність того, що агентство підслухало: *а*) тільки два повідомлення; *б*) усі три повідомлення.

10. З опитаних бізнесменів 80 % віддають перевагу зберіганню грошей у банку (подія *A*), 60 % вкладає гроші в цінні папери (подія *B*), 50 % одночасно тримають гроші в банку та вкладають у цінні папери. Результати опитування подано в таблиці.

	A	$\overline{A}$	Разом
B	0,5		0,6
$\overline{B}$			
Разом	0,8		1

Заповнити таблицю до кінця. Знайти ймовірність того, що навмання обраний бізнесмен тримає гроші в банку або у вигляді цінних паперів.

Відповіді: **1.** *а*) 0,72; *б*) 0,98. **2.** 0,9125. **3.** 0,99995. **4.** 0,75. **5.** 0,251. **6.** 0,32. **7.** 0,864. **8.** 0,9. **9.** *а*) 0,446; *б*) 0,273. **10.** 0,9.

Завдання 4. Формула повної ймовірності. Формула Байєса

1. Серед виробів, які випускаються заводом, 96 % відповідають стандарту. Спрощена схема контролю визнає доброякісною стандартну продукцію з імовірністю 0,98 і нестандартну – з імовірністю 0,05.

Знайти:

а) ймовірність того, що навмання взятий виріб пройде спрощений контроль;

б) за умови проходження спрощеного контролю ймовірність того, що він відповідає стандарту.

2. Інвестор вкладає свої гроші в акції без ризику, в акції допустимого ризику та акції великого ризику в пропорціях: 10 %, 30 %, 60 %. Імовірність отримання прибутку від цих акцій становить 1, 0,75 та 0,6 відповідно. Знайти:

а) імовірність того, що інвестор отримає прибуток;

б) якщо інвестор отримав прибуток, то яка ймовірність того, що він отриманий від акцій без ризику?

3. Економічний факультет провів обстеження працевлаштування своїх випускників. Імовірність того, що людина, яка працює в сфері бізнесу та має прибуток вище N гривень становить 0,9, а поза сферою бізнесу – 0,3. З'ясовано, що 80 % випускників працюють у сфері бізнесу. Знайти:

а) імовірність того, що навмання обраний випускник має прибуток вище N гривень;

б) за умови отримання прибутку випускником більше N гривень яка ймовірність того, що він працює в сфері бізнесу?

4. Задачу розв'язують самостійно 2 відмінники, 3 посередні студенти та 5 студентів, що навчаються добре. Ймовірність розв'язання задачі відмінником дорівнює 0,9; студентом, який навчається добре, – 0,8 та посереднім – 0,5. До дошки навмання викликають одного зі студентів.

1) Знайти ймовірність того, що студент розв'язав задачу.

2) Викликаний студент розв'язав задачу. Знайти ймовірність того, що він: *а*) відмінник; *б*) студент із середнім рівнем успішності.

5. Податкові інспектори роблять перевірку діяльності підприємств. Перший інспектор обслуговує 40 підприємств, серед яких 25 % не мають заборгованостей, другий – 60 підприємств, із них 40 % без заборгованостей. Знайти ймовірність того, що навмання обране підприємство:

а) не має заборгованості;

б) підприємство, що не має заборгованості, перевіряв перший інспектор.

6. Фабрика виготовляє однотипну продукцію на трьох конвеєрних лініях, які мають однакову продуктивність. На першій лінії виробляється продукція

тільки першого сорту, на другій лінії продукція першого сорту становить 90 %, а на третій – 85 %. Знайти ймовірність того, що:

а) випадковим способом узятий виріб буде першосортним;

б) випадково взятий виріб виявився першосортним і його виготовлено на третій лінії.

7. Менеджер з інвестицій передбачає три варіанти розвитку економічної ситуації на наступний рік: високе зростання, відсутність зростання та спад. Імовірності цих подій становить: 0,6; 0,3 та 0,1 відповідно. Очікується отримання прибутку з наявного активу. Імовірність отримання прибутку становить: випадок високого зростання – 0,8; випадок відсутності зростання – 0,6; випадок спаду – 0,1, знайти ймовірність того, що:

а) буде отримано прибуток з наявного активу;

б) цей прибуток з наявного активу отримано в умовах високого зростання економіки.

8. Авіатехнічний склад одержує від першого заводу в 4 рази більше агрегатів, ніж від другого. Брак у продукції першого заводу становить 4 %, а другого – 8 %.

1) Знайти ймовірність того, що випадковим способом узятий агрегат виявиться бракованим;

2) Випадковим способом узятий агрегат виявився бракованим. На якому заводі він більш імовірно виготовлений?

9. Два робітники виготовили по однаковій кількості деталей. Брак продукції, виготовленої першим робітником, становить 5 %, а другим – 1 %, знайти ймовірність того, що:

а) узята навмання деталь буде бракованою;

б) бракована деталь виготовлена першим робітником.

10. На фабриці перша машина виробляє 40 %, а друга – 60 % усієї продукції. У середньому 9 з 1000 одиниць продукції, виготовленої першою машиною, виявляється браком, а для другої машини брак становить 2 одиниці на 500 одиниць продукції:

а) знайти ймовірність браку продукції, виготовленої на фабриці;

б) на якій з машин виготовлена більш імовірно певна одиниця продукції, обрана випадково із даної продукції фабрики, що виявилася браком.

Відповіді. **1.** *а*) 0,9428; *б*) 0,9979. **2.** *а*) 0,885; *б*) 0,678. **3.** *а*) 0,78; *б*) 0,923. **4.** 1) 0,73; 2) *а*) 0,247; *б*) 0,205. **5.** *а*) 0,34; *б*) 0,294. **6.** *а*) 0,917; *б*) 0,309. **7.** *а*) 0,67; *б*) 0,716. **8.** 1) 0,048; 2) першим заводом. **9.** *а*) 0,03; *б*) 0,83. **10.** *а*) 0,006; *б*) на першій машині.

Завдання 5. Повторення незалежних випробувань. Формула Бернуллі. Формула Муавра–Лапласа. Формула Пуассона

1. Серед великої кількості виробів, що знаходяться в комплекті, 30 % – нестандартні. Знайти ймовірності того, що серед 5 виробів, навмання взятих із комплекту, буде: *а*) тільки один нестандартний; *б*) принаймні один нестандартний.

2. Імовірність того, що кожен клієнт, який звернувся в авіакасу, замовить квиток до аеропорту *N*, дорівнює 0,1. Знайти ймовірності того, що із 100 клієнтів, що звернулись в касу, замовлять квиток до аеропорту *N*: *а*) менше 15 осіб; *б*) 5 – 12 осіб; *в*) більше 20 осіб.

3. На біржі виставлено 10 цінних паперів. Імовірність того, що вони подорожчають протягом одного дня, дорівнює 0,6. Знайти ймовірності того, що подорожчає: *а*) рівно 5 паперів; *б*) не більше, ніж 4 папери; *в*) 3 – 5 паперів.

4. Авіакомпанія виконує протягом місяця 400 рейсів. Імовірність повного комерційного завантаження кожного рейсу дорівнює 0,8. Знайти ймовірності того, що протягом місяця з повним комерційним завантаженням буде виконано: *а*) не менше 300 рейсів; *б*) більша частина рейсів.

5. За статистичними даними у середньому 1 % пасажирів відмовляється від рейсу. Знайти ймовірності того, що з 300 пасажирів, які мають квитки на рейс, відмовляться від польоту: *а*) не більше 5 пасажирів; *б*) не менше 3 пасажирів.

6. Інвестор укладає договір на фондовій біржі. Ймовірність укладання однієї угоди за день дорівнює 0,7. Виходячи із припущення, що протягом 10 робочих днів укладається не більше однієї угоди в день, знайти ймовірності таких подій: *а*) буде укладено 7 угод; *б*) буде укладено не менше 8 угод; *в*) жодної угоди не буде укладено.

7. Кількість помилок у рахунках торгових підприємств становить 5 %. Аудитор перевіряє 10 навмання вибраних рахунків. Якщо не виявиться жодної помилки, то рахунки підприємства далі не перевірятимуть. Яка ймовірність того, що в 10 рахунках підприємства: *а*) не буде жодної помилки; *б*) буде 3 помилки; *в*) буде 3 – 5 помилок.

8. Телефонна станція обслуговує 2000 абонентів. Імовірність того, що будь-який абонент зателефонує на станцію протягом години, дорівнює 0,001. Знайти ймовірності того, що протягом години на телефонну станцію зателефонують: *а*) 5 абонентів; *б*) не більше 3 абонентів.

9. Імовірність того, що інвестиційний проект принесе через рік прибуток, дорівнює 0,8. Знайти ймовірності того, що із 15 інвестиційних проектів: *а*) 10 проектів виявляться прибутковими; *б*) не менше 8 проектів виявляться прибутковими; *в*) 5 – 9 проектів будуть прибутковими.

10. Фабрика випускає 75 % продукції першого сорту. Знайти ймовірність того, що із 300 виробів, виготовлених фабрикою, кількість першосортних виробів буде: *а*) 250 виробів; *б*) 220 – 235; *в*) не більше 200.

Відповіді. **1.** *а*) 0,36; *б*) 0,832. **2.** *а*) 0,9526; *б*) 0,7; *в*) 0,00048. **3.** *а*) 0,2; *б*) 0,166; *в*) 0,3548. **4.** *а*) 0,99; *б*) 1. **5.** *а*) 0, 916; *б*) 0,084. **6.** *а*) 0,267; *б*) 0,38; *в*) $(0,3)^{10}$. **7.** *а*) 0,01; *б*) 0,0011; *в*) 001146. **8.** *а*) 0,0361; *б*) 0,86. **9.** *а*) 0,103; *б*) 0,9768; *в*) 0,018. **10.** *а*) 0,00023; *б*) 0,648; *в*) 0,00048.

Завдання 6. Дискретні випадкові величини

1. Продавець морозива дійшов висновку, що рівень продажу залежить від погоди: сонячної, похмурої та холодної. Сонячні дні становлять 50 %, холодні – 10 %. Виторг від продажу морозива становить 290, 260 і 225 грн за день відповідно до стану погоди. Неповернені витрати на морозиво становлять 100 грн за день. Побудувати ряд розподілу випадкової величини X – прибутку від продажу морозива та визначити її числові характеристики.

2. Цінні папери на біржі кожного дня можуть із ймовірністю 0,5 подорожчати на 10 %. Спостереження ведеться три дні. Початкова вартість цінних паперів 5 000 грн. Побудувати ряд розподілу випадкової величини X –

вартості цінних паперів, вважаючи її розподіленою за біномним законом та знайти її числові характеристики.

3. Огляд рахунків 500 інвесторів на фондовій біржі дав таку інформацію про кількість угод протягом 10 робочих днів:

Кількість угод	0	1	2	3	4	5	6	7	8	9	10
Кількість власників фінансових інструментів	136	98	84	76	38	28	14	10	10	3	3

Побудувати ряд розподілу випадкової величини X – кількості угод та визначити її числові характеристики.

4. Корпорація вкладає гроші у купівлю земельної ділянки вартістю 100 000 у. о., сподіваючись, що за два роки ціна на земельну ділянку зросте на $q\,\%$. Є чотири прогнози зміни $q\,\%$ за два роки:

q, %	50	45	40	30
Імовірність	0,2	0,3	0,3	0,2

Побудувати ряд розподілу випадкової величини X – вартості земельної ділянки за два роки та визначити її числові характеристики.

5. Продавець кафе встановив такий закон розподілу рівня продажів прохолоджувальних напоїв залежно від погоди:

	Сонячна	Похмура	Холодна
Сума продажів	150	100	40
Ймовірність	0,3	0,5	0,2

Загальна собівартість напоїв становить 80 грн. Побудувати ряд розподілу випадкової величини X – величини прибутку за один день та визначити її числові характеристики.

6. У рекламних цілях торгова фірма вкладає в кожну десяту одиницю товару приз вартістю 100 грн. Побудувати ряд розподілу випадкової величини X – розміру виграшу за 5 покупок та знайти її числові характеристики.

7. Клієнти банку, які не зв'язані між собою, не повертають кредити в зазначений термін з імовірністю 0,1. Побудувати ряд розподілу випадкової

величини X – кількості повернених кредитів із 5 виданих та знайти її числові характеристики.

8. Контрольна робота складається із трьох запитань. На кожне запитання запропоновано 4 відповіді, серед яких одна правильна. Побудувати ряд розподілу випадкової величини X – кількості правильних відповідей за простого вгадування та знайти її числові характеристики.

9. У білеті є три задачі. Ймовірність правильного розв'язання студентом першої задачі дорівнює 0,9; другої – 0,8; третьої – 0,7. Побудувати ряд розподілу випадкової величини X – кількості правильно розв'язаних задач у білеті та знайти її числові характеристики.

10. Скласти закон розподілу випадкової величини X – кількості пакетів з трьох акцій, за якими власником буде отримано прибуток, якщо ймовірність отримання прибутку за кожним пакетом дорівнює 0,5; 0,6 та 0,7. Визначити числові характеристики випадкової величини X.

Відповіді:

1. $M(X)=171{,}5;\quad D(X)=440{,}25.$

2. $M(X)=5788{,}125;\ D(X)=228424{,}61;\ \sigma(X)=477{,}94.$

X	5000	5500	6050	6655
P	$0{,}5^3$	$3\cdot 0{,}5^3$	$3\cdot 0{,}5^3$	$0{,}5^3$

3. $M(X)=2{,}154;\quad D(X)=4{,}566.$

4. $M(X)=41500;\ \ D(X)=45239076.$

5. $M(X)=2060;\quad D(X)=584400\,.$

X	800	2000	3000
P	0,2	0,5	0,3

6.

X	0	100	200	300	400	500
P	$\frac{9^5}{10^5}$	$5\cdot\frac{9^4}{10^5}$	$10\cdot\frac{9^3}{10^5}$	$10\cdot\frac{9^2}{10^5}$	$5\cdot\frac{9^2}{10^5}$	$\frac{1}{10^5}$

7. $M(X)=4{,}5;\quad D(X)=0{,}45.$

X	0	1	2	3	4	5
P	0,00001	0,00045	0,0081	0,0729	0,32805	0,59049

8. $M(X)=0{,}75;\quad D(X)=0{,}5625.$

X	0	1	2	3
P	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

9. $M(X)=2{,}4;\quad D(X)=0{,}46.$

X	0	1	2	3
P	0,006	0,092	0,398	0,504

10. $M(X)=1{,}8;\quad D(X)=0{,}7.$

X	0	1	2	3
P	0,06	0,29	0,44	0,21

Завдання 7. Неперервні випадкові величини

7.1. Задана функція розподілу неперервної випадкової величини X. Знайти коефіцієнт A; записати щільність розподілу $f(x)$; визначити числові характеристики $M(X), D(X)$, а також $P(\alpha \le X \le \beta)$. Накреслити графіки функції розподілу та щільності розподілу.

1. $F(x)=\begin{cases} 0, \ x \le 0; \\ Ax^5, \ 0 < x \le 3, \ \alpha = 0{,}25; \ \beta = 0{,}75; \\ 1, \ x > 3. \end{cases}$

2. $F(x)=\begin{cases} 0, \ x \le 0; \\ A\left(2x - \dfrac{x^2}{3}\right), \ 0 < x \le 3, \ \alpha = 1; \ \beta = 2; \\ 1, \ x > 3. \end{cases}$

3. $F(x)=\begin{cases} 0, \ x \le 0; \\ A\left(-x^2 + 4x\right), \ 0 < x \le 2, \ \alpha = 0{,}5; \ \beta = 1{,}5; \\ 1, \ x > 2. \end{cases}$

4. $F(x)=\begin{cases}0, \ x\le 0;\\ Ax^2, \ 0<x\le 1, \ \alpha=0{,}3; \ \beta=0{,}8;\\ 1, \ x>1.\end{cases}$

5. $F(x)=\begin{cases}0, \ x\le 0;\\ Ax^3, \ 0<x\le 2, \ \alpha=0{,}2; \ \beta=1{,}7.\\ 1, \ x>2.\end{cases}$

7.2. Неперервна випадкова величина задана щільністю розподілу $y=f(x)$. Записати інтегральну функцію розподілу $y=F(x)$; визначити числові характеристики $M(X), D(X)$, а також $P(\alpha\le X\le\beta)$. Накреслити графіки диференціальної та інтегральної функцій розподілу.

6. $f(x)=\begin{cases}0, \ x\le 0, \ x>\pi/2;\\ \cos x, \ 0<x\le \pi/2;\end{cases} \ \alpha=-\pi/4; \ \beta=\pi/4.$

7. $f(x)=\begin{cases}0, \ x\le 0, \ x>\pi/2;\\ \sin x, \ 0<x\le \pi/2;\end{cases} \ \alpha=-\pi; \ \beta=\pi/4.$

8. $f(x)=\begin{cases}0, \ x\le \pi/6, \ x>\pi/3;\\ 3\sin 3x, \ \pi/6<x\le \pi/3;\end{cases} \ \alpha=0; \ \beta=\pi/4.$

9. $f(x)=\begin{cases}\frac{2}{\pi}\cos^2 x, \ |x|\le \pi/2;\\ 0, \ |x|>\pi/2;\end{cases} \ \alpha=0; \ \beta=\pi/4.$

10. $f(x)=\begin{cases}ae^{-ax} \ (a>0), \ x>0;\\ 0, \ x\le 0;\end{cases} \ \alpha=1; \ \beta=2.$

Відповіді:

1. $A=\frac{1}{243}; \ M(X)=2{,}5; \ D(X)=\frac{5}{28}; \ p=0{,}00097.$

2. $A=\frac{1}{3}; \ M(X)=1; \ D(X)=0{,}5; \ p=\frac{1}{3}.$

3. $A=\frac{1}{4};\ M(X)=\frac{2}{3};\ D(X)=\frac{2}{9};\ p=2.$

4. $A=1;\ M(X)=\frac{2}{3};\ D(X)=\frac{1}{18};\ p=0,55.$

5. $A=\frac{1}{8};\ M(X)=\frac{3}{2};\ D(X)=\frac{3}{20};\ p=0,613.$

6. $M(X)=\frac{\pi}{2}-1;\ D(X)=\pi-3;\ p=\frac{\sqrt{2}}{2}.$

7. $M(X)=1;\ D(X)=\pi-3;\ p=\frac{\sqrt{2}}{2}.$

8. $M(X)=\frac{\pi-1}{3};\ D(X)=\frac{\pi-3}{9};\ p=\frac{\sqrt{2}}{2}.$

9. $M(X)=0;\ D(X)=\frac{\pi^2-6}{12};\ p=\frac{\pi+2}{4\pi}.$

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

Завдання 8. Деякі закони розподілу неперервних випадкових величин

1. Магазин отримав 1000 пляшок мінеральної води в скляній тарі. Ймовірність того, що після перевезення пляшка виявиться розбитою, дорівнює 0,003. Знайти ймовірності того, що магазин отримає розбитих пляшок: *а*) тільки дві; *б*) менше двох; *в*) більше двох; *г*) принаймні одну.

2. Випадкова величина X задає час безвідмовної роботи системи і має щільність розподілу $f(x)=Ae^{-0,1t}$, $t\geq 0$. Знайти коефіцієнт A та надійність (імовірність безвідмовної роботи системи) протягом часу τ. Яка ймовірність того, що час безвідмовної роботи системи буде меншим від математичного сподівання?

3. Час відновлення каналу зв'язку має показниковий розподіл. Середній час відновлення дорівнює 10 хв. Яка ймовірність того, що час відновлення буде в межах 5 – 25 хв?

4. Час приймання та оброблення одного повідомлення – випадкова величина X, яка розподілена за показниковим законом. У середньому за хвилину приймають 6 повідомлень. Яка ймовірність того, що повідомлення буде прийняте та оброблене протягом 4 хв?

5. Припускаючи, що річний дохід деяких фізичних осіб розподілений нормально із середнім значенням 15 000 грн і стандартним відхиленням 3 000 грн, визначити, який відсоток родин отримає: *а*) принаймні 15 000 грн; *б*) більше 20 000 грн.

6. Випадкова величина X – кількість продажів ділера, розподілена нормально з середнім значенням 29 і середнім відхиленням 7. Знайти ймовірність того, що кількість продажів перебуватиме в межах: *а*) 16 – 18 одиниць; *б*) 28 – 35 одиниць.

7. Компанія встановила, що випадкова величина X – кількість робочих днів, пропущених співробітниками протягом року через хворобу, розподілена за нормальним законом з середнім значенням 78 днів і стандартним відхиленням – 14 днів. Знайти ймовірність того, що кількість робочих днів, пропущених через хворобу, буде: *а*) не більше 50; *б*) 50 – 70.

8. Відомо, що прибутковість пакета акцій фірми розподілена за законом Гаусса із середнім значенням 12 % і стандартним відхиленням 1,2 %. Знайти ймовірність того, що: *а*) акції матимуть прибутковість не більше 10 %; *б*) акції матимуть прибутковість 10 – 14 %.

9. Час, який співробітники фірми витрачають на дорогу до місця роботи, розподілений за нормальним законом із середнім значенням 40 хв і стандартним відхиленням 6 хв. Знайти ймовірність того, що навмання вибраний співробітник витратить на дорогу: *а*) більше 50 хв; *б*) 30 – 45 хв.

10. Прибуток фірми має нормальний розподіл із середнім значенням 12 000 грн на місяць і стандартним відхиленням 1 000. Визначити інтервал прибутку із надійністю 95 %.

Відповіді: **1.** *а*) 0,224; *б*) 0,199; *в*) 0,577; *г*) 0,95. **2.** $A = 0{,}1$; $p = 1 - e^{-1}$. **3.** 0,524. **4.** ≈ 1. **5.** *а*) 0,5; *б*) 0,047. **6.** *а*) 0,02677; *б*) 0,36078. **7.** *а*) 0,023; *б*) 0,26. **8.** *а*) 0,04746; *б*) 0,90508. **9.** *а*) 0,04746; *б*) 0,74927. **10.** (10,04; 13,96).

Завдання 9. Система двох дискретних випадкових величин. Функція одного випадкового аргумента

1. Випадкова величина X розподілена за нормальним законом із $M(X)=0$. Знайти закон розподілу функції $Y = X^3$.

2. Неперервна випадкова величина X задана диференціальною функцією розподілу

$$f(x)=\begin{cases} 0, \; x \le 0, \; x > \dfrac{\pi}{2}; \\ \sin x, \; 0 < x \le \dfrac{\pi}{2}. \end{cases}$$

Знайти математичне сподівання функції $Y = X^2$.

3. Задані закони розподілів незалежних випадкових величин X та Y

X	–1	0	1
P	0,3	0,5	0,2

та Y

Y	0	1
P	0,4	0,6

Знайти математичне сподівання та дисперсію випадкової величини $Z = X^2 \cdot Y^3$, використовуючи властивість математичного сподівання

$$M(Z) = M(X^2 \cdot Y^3) = M(X^2) \cdot M(Y^3).$$

4. Випадкова величина X рівномірно розподілена в інтервалі (–1, 1). Знайти щільність розподілу імовірності випадкової величини $Y = X^2$.

5. Випадкова величина X рівномірно розподілена в інтервалі (1, 2). Знайти $M(Z) = M(3X^2 - 2X + 1)$ та $D(Z) = D(X^2 + 1)$.

6. Фірма планує відкрити ще одне кафе. Для оптимального планування можливої кількості відвідувачів кафе протягом деякого часу аналітики фірми провели статистичне дослідження кількості відвідувачів й отриманого прибутку. У таблиці подано закон розподілу системи двох дискретних випадкових величин: X – кількість відвідувачів кафе за деякий час і Y – прибуток, який отримала фірма в умовних одиницях. Вважати, що k – номер варіанта.

$Y \backslash X$	k	$k+2$	$k+4$	$k+6$
k	$0{,}002(35-k)$	$0{,}002(k+40)$	$0{,}002(70-k)$	$0{,}002(k+20)$
$k+5$	$0{,}002(k+30)$	$0{,}002(55-k)$	$0{,}002(k+25)$	$0{,}002(40-k)$
$k+10$	$0{,}002(50-k)$	$0{,}002(k+5)$	$0{,}002(80-k)$	$0{,}002(k+50)$

Виконати такі завдання:

1) скласти закони розподілу одновимірних випадкових величин X та Y;

2) знайти математичне сподівання, дисперсію та середнє квадратичне відхилення випадкових величин X та Y;

3) обчислити кореляційний момент $\hat{E}\ (X, Y)$ і коефіцієнт кореляції $\rho\,(X, Y)$;

4) побудувати умовний закон розподілу випадкової величини X за умови, що випадкова величина Y набуває значення $k+10$ та умовний закон розподілу випадкової величини Y за умови, що випадкова величина X набуває значення $k+2$;

5) знайти умовні математичні сподівання $M\left(\dfrac{Y}{X}=k+2\right)$, $M\left(\dfrac{X}{Y}=k+10\right)$.

Відповіді:

1. $g(y)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{\frac{-y^{\frac{2}{3}}}{2\sigma^2}}\cdot\left|(y^{\frac{1}{3}})'\right|.$

2. $\pi-2.$

3. 0,3; 0,2.

4. $g(y)=\begin{cases}\dfrac{1}{4\sqrt{y}},\ y\in(0{,}1);\\ 0,\ y\notin(0{,}1).\end{cases}$

5. 5; 0,76.

6. Для $k=3$: 2) $M(X)=6{,}144$; $D(X)=4{,}579$; $\sigma(X)=2{,}14$ $M(Y)=8{,}2$; $D(Y)=17{,}46$; $\sigma(Y)=4{,}17$; 3) $\hat{E}\ (X,Y)=0{,}8712$; $\rho\,(X,Y)=0{,}0976$;

5) $M\left(Y/_{X=5}\right)=5{,}525$; $M\left(X/_{Y=13}\right)=6{,}463.$

Завдання 10. Вибірка та її основні характеристики

За результатами спостережень над випадковою величиною X, які подано у таблиці, знайти емпіричну функцію розподілу, вибіркове середнє та незміщену оцінку дисперсії.

1.

x_i	–1	1	2	3	4	5
n_i	3	8	20	16	7	2

2.

x_i	1	2	3	4	5	6
n_i	5	8	15	12	7	3

3.

x_i	–2	–1	0	1	2	3
n_i	2	8	20	9	8	3

4.

x_i	–2	–1	0	1	2	3
n_i	3	12	20	8	5	2

5.

x_i	–3	–2	–1	0	1	2
n_i	2	6	25	10	5	2

6.

x_i	–4	–3	–1	0	2	3
n_i	1	7	23	10	7	2

7.

x_i	–1	0	1	2	4	5
n_i	3	9	18	10	8	2

8.

x_i	–2	–1	2	3	4	5
n_i	5	10	15	8	7	5

9.

x_i	–3	–2	–1	1	2	3
n_i	1	7	25	10	5	2

10.

x_i	0	1	2	3	5	6
n_i	2	7	20	10	8	3

Відповіді: **1.** $\overline{x} = 2,34$; $s^2 = 0,915$. **2.** $\overline{x} = 3,22$; $s^2 = 6,298$. **3.** $\overline{x} = 0,44$; $s^2 = 1,517$. **4.** $\overline{x} = 0,12$; $s^2 = 1,414$. **5.** $\overline{x} = -0,68$; $s^2 = 1,161$.

Завдання 11. Надійні інтервали

Знайти надійний інтервал для оцінювання математичного сподівання a нормального розподілу з надійністю 0,95, знаючи вибіркову середню $\overline{x}$, обсяг вибірки n і середнє квадратичне відхилення σ.

1. $\overline{x} = 75,17$; $n = 36$; $\sigma = 6$.
2. $\overline{x} = 75,16$; $n = 49$; $\sigma = 7$.
3. $\overline{x} = 75,15$; $n = 64$; $\sigma = 8$.
4. $\overline{x} = 75,14$; $n = 81$; $\sigma = 9$.
5. $\overline{x} = 75,13$; $n = 100$; $\sigma = 10$.
6. $\overline{x} = 75,12$; $n = 121$; $\sigma = 11$.
7. $\overline{x} = 75,11$; $n = 144$; $\sigma = 12$.
8. $\overline{x} = 75,10$; $n = 169$; $\sigma = 13$.
9. $\overline{x} = 75,09$; $n = 196$; $\sigma = 14$.
10. $\overline{x} = 75,08$; $n = 225$; $\sigma = 15$.

Відповіді: **1.** (73,21; 77,13). **2.** (73,2; 77,12). **3.** (73,19; 77,12). **4.** (73,18; 77,1). **5.** (73,17; 77,09).

Завдання 12. Вибіркові характеристики зв’язку

За вибірковими даними пари випадкових величин (X, Y) знайти вибірковий коефіцієнт кореляції пари і рівняння лінійної регресії Y на X та X на Y.

1.

x_i	0,1	1,1	2.0	3,0	4,0	5,0
y_i	0,5	0,9	1,4	2,4	3,5	3,9

2.

x_i	6,1	7,0	8.0	9,1	10	12
y_i	4,5	5,1	7,1	7,8	8,2	10,4

3.

x_i	1	2	4	8	9	10
y_i	12	10	9	6	4	3

4.

x_i	–1,1	0	2	3	4	6
y_i	0,1	1	1,5	2	3	4

5.

x_i	2	4	6	8	9,5	11
y_i	4,5	7	8	7,5	9	10

6.

x_i	0	2	3,4	4	5	6
y_i	5	4	3,5	2	1	0

7.

x_i	2	4	6	8	10	12
y_i	4,5	7	8	8,5	9	10

8.

x_i	1	2	3	4	5	6
y_i	0,1	0,3	0,5	0,8	0,9	1

9.

x_i	1	2	3	4	5	6
y_i	1	0,9	0,8	0,5	0,3	0,1

10.

x_i	0	2	3	4	5	6
y_i	6	4,5	3,5	2	1	0

Відповіді:

1. $r = 0{,}45;\ y = -0{,}25x + 1{,}475;\ x = 0{,}81y + 0{,}79.$

2. $r = 0{,}255;\ y = 0{,}26x + 4{,}89;\ x = 0{,}25y + 6{,}93.$

3. $r = -0{,}82;\ y = -0{,}76x + 11{,}64;\ x = -0{,}88y + 12{,}12.$

4. $r = 0{,}82;\ y = 0{,}44x + 0{,}91;\ x = 1{,}53y - 0{,}63.$

Завдання 13. Перевірка статистичних гіпотез

13.1. За таблицею статистичного розподілу побудувати полігон розподілу, оцінити правдоподібність гіпотези, що випадкова величина розподілена за законом Пуассона. Перевірити узгодженість між емпіричними та теоретичними даними за допомогою критерію Пірсона. Рівень значущості $\alpha = 0{,}05$.

№	X	0	1	2	3	4	5	6	7	8	9
1	n	37	37	18	6	2	0	0	0	0	0
2		29	36	22	9	3	1	0	0	0	0
3		22	33	25	13	5	2	0	0	0	0
4		17	30	27	16	7	2	1	0	0	0
5		14	27	27	18	9	4	1	0	0	0
6		11	24	27	20	11	5	2	0	0	0
7		8	21	26	21	13	7	3	1	0	0
8		6	18	24	22	15	8	4	2	1	0
9		5	15	22	22	17	10	5	2	1	0
10		4	13	20	22	18	12	6	3	1	1

13.2. Проведено $n = 500$ випробувань випадкової величини X. Результати випробувань зведені в групований статистичний ряд. Користуючись критерієм згоди χ^2, визначити, чи не суперечить вибірковим даним гіпотеза про те, що випадкова величина розподілена за нормальним законом з математичним сподіванням, рівним вибірковому середньому, і дисперсією, рівною виправленій вибірковій дисперсії. Рівень значущості $\alpha = 0{,}1$.

№	інтервали	$[-4,-3)$	$[-3,-2)$	$[-2,-1)$	$[-1,0)$	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$
1	n_i	5	25	70	134	120	90	46	10
2		4	26	68	135	121	92	45	9
3		4	25	70	135	122	90	46	8
4		7	25	69	135	123	89	45	7
5		8	26	72	135	122	91	43	3
6		2	24	72	136	121	90	47	8
7		6	24	69	135	120	92	45	9
8		7	20	73	132	120	92	45	11
9		4	24	69	133	121	91	47	11
10		6	26	71	135	119	89	45	9

Завдання для модульної контрольної роботи

Варіант 1

1.1. Обчислити: $A_6^2 : P_5 \cdot C_6^3$.

1.2. Підприємство виготовляє 96 % стандартних виробів, причому 80 % із них вищого сорту. Знайти ймовірність того, що взятий навмання виріб буде стандартним, а не вищого сорту.

1.3. На склад автомашинами завозять 30 % скляних банок, а решту – залізницею. Ймовірність розбивання скляних банок під час транспортування автомашинами дорівнює 0,01, а під час транспортування залізницею – удвічі менша. Визначити ймовірність того, що узята навмання на складі банка виявиться розбитою.

1.4. За цим законом розподілу дискретної випадкової величини *X*

X	–2	–1	0	4	5	7
p	0,12	0,18	0,2	0,3	0,17	0,03

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина *X* задана функцією розподілу ймовірностей $F(x)$:

$$F(x)=\begin{cases}0, \; x \le 0;\\ \dfrac{x^3}{27}, \; 0 < x \le 3;\\ 1, \; x > 3.\end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини *X*;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (1, 2)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

$Y \backslash X$	2	6	10	14
1	0,25	0,25	0	0,05
4	0,2	0,1	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$, $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right), M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right), D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

$$\begin{matrix} 0 & 2 & 4 & 5 & 6 \\ 2 & 1 & 2 & 0 & 4 \\ 3 & 4 & 3 & 4 & 2 \\ 1 & 2 & 0 & 4 & 4 \\ 4 & 4 & 2 & 5 & 1 \end{matrix}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 2

1.1. Обчислити: $C_7^2 : A_7^2 \cdot P_5$.

1.2. Студент прийшов складати іспит, знаючи з 50 питань програми лише 36. Яка ймовірність того, що студент відповість принаймні на одне із заданих йому трьох запитань?

1.3. Ймовірність присутності студента на занятті дорівнює 0,9. Чому дорівнює найімовірніша кількість студентів, присутніх на занятті, якщо всього у групі 35 студентів?

1.4. За даним законом розподілу дискретної випадкової величини X

X	–2	–1	0	1	3	4
p	0,15	0,2	0,25	0,2	0,15	0,05

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, \ x \le 0; \\ \dfrac{x^2}{4}, \ 0 < x \le 2; \\ 1, \ x > 2. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (1, 2)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

Y \ X	4	6	8	10
1	0,1	0,05	0,15	0,1
4	0,15	0,3	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$.

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$, $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right), M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right), D\left(Y \mid X = x_3\right)$.

1.7. У результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

6	4	6	6	2
2	1	4	0	5
2	0	3	2	1
1	2	1	4	4
4	4	2	2	6

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 3

1.1. Обчислити: $(C_9^7 - A_4^3) : P_4$.

1.2. Підприємство виготовляє 98 % стандартних виробів, причому 85 % із них вищого сорту. Знайти ймовірність того, що взятий навмання виріб буде вищого сорту.

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

1.4. За даним законом розподілу дискретної випадкової величини X

X	–3	–1	0	1	3	5
p	0,13	0,17	0,2	0,3	0,18	0,02

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x)=\begin{cases}0, \; x\le 3;\\ \dfrac{x^2-9}{27}, \; 3<x\le 6;\\ 1, \; x>6.\end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta)=(4, 5)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

Y \ X	1	3	5	7
3	0,05	0,05	0,1	0,1
4	0,2	0,35	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$.

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$.

в) побудувати умовні закони розподілу $X \mid Y=y_2$; $Y \mid X=x_3$.

г) обчислити умовні математичні сподівання $M\left(X \mid Y=y_2\right), M\left(Y \mid X=x_3\right)$ та дисперсії $D\left(X \mid Y=y_2\right), D\left(Y \mid X=x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

1	4	5	0	6
5	1	4	6	3
3	2	1	1	5
2	1	5	4	0
4	5	2	5	6

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 4

1.1. Обчислити: $(A_9^2 - C_8^4) \cdot P_3$.

1.2. Під час стріляння з автоматичної зброї відносна частота влучень у ціль дорівнює 0,95. Знайти кількість влучень, якщо зроблено 160 пострілів.

1.3. В ящику 7 стандартних і 3 нестандартні деталі. Навмання послідовно беруть 3 із поверненням. Знайти ймовірність того, що серед узятих деталей немає стандартних.

1.4. За даним законом розподілу дискретної випадкової величини X

X	2	4	7	9	12	15
p	0,05	0,15	0,35	0,2	0,15	0,1

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, x \le 2; \\ \dfrac{x^2 - 4}{21}, 2 < x \le 5; \\ 1, x > 5. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (3, 4)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

$Y \backslash X$	2	5	6	8
1	0,15	0,05	0,1	0,1
4	0,25	0,2	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2,\ Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right), M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right), D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

$$\begin{array}{ccccc} 6 & 2 & 3 & 3 & 4 \\ 1 & 2 & 2 & 1 & 6 \\ 4 & 0 & 3 & 4 & 3 \\ 2 & 3 & 1 & 2 & 4 \\ 5 & 3 & 5 & 3 & 6 \end{array}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 5

1.1. Обчислити: $A_{10}^4 : P_7 \cdot C_8^4$.

1.2. Для руйнування моста достатньо влучення однієї авіаційної бомби. Знайти ймовірність того, що міст буде зруйновано, якщо на нього скинули 3 бомби та ймовірності влучення яких відповідно дорівнюють 0,6; 0,7 та 0,9.

1.3. Деталь може надійти для обробки на перший верстат з імовірністю 0,8, а на другий верстат – з імовірністю 0,2. Під час оброблення деталі на першому й другому верстатах імовірність допустити брак відповідно дорівнює 0,01 і 0,02. Оброблені деталі складають в одному приміщенні. Навмання взята деталь виявилась бракованою. Знайти ймовірність того, що її обробляли на першому верстаті.

1.4. За даним законом розподілу дискретної випадкової величини *X*

X	3	4	7	9	12	14
p	0,1	0,3	0,2	0,05	0,15	0,2

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина *X* задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, \ x \le 2; \\ \dfrac{x^3 - 4x}{15}, \ 2 < x \le 3; \\ 1, \ x > 3. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини *X*;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (2,5; 4)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

$Y \backslash X$	4	5	7	8
2	0,2	0,05	0	0,1
5	0,2	0,3	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$; $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right)$; $M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right)$; $D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

$$\begin{matrix} 1 & 0 & 2 & 4 & 0 \\ 0 & 2 & 3 & 0 & 5 \\ 2 & 6 & 1 & 5 & 2 \\ 1 & 0 & 2 & 3 & 1 \\ 3 & 2 & 4 & 2 & 5 \end{matrix}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 6

1.1. Обчислити: $A_9^4 : P_6 \cdot C_6^3$.

1.2. Кожен із трьох стрільців виконує один постріл по мішені. Ймовірності їхнього влучення в мішень відповідно дорівнюють 0,6; 0,7 та 0,8. Знайти ймовірність того, що результатом цих трьох пострілів буде тільки:

а) одне влучення;

б) принаймні одне влучення.

1.3. Схожість насіння становить 60%. Яка ймовірність того, що з чотирьох посіяних насінин зійде дві?

1.4. За даним законом розподілу дискретної випадкової величини X

X	–5	–4	0	1	2	4
p	0,15	0,2	0,25	0,2	0,15	0,05

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0,\ x \le -1; \\ \dfrac{(x+1)^3}{8}, -1 < x \le 1; \\ 1,\ x > 1. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (0; 1)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

Y \ X	1	4	7	12
3	0,15	0,05	0,1	0,1
5	0,25	0,2	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$; $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right)$; $M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right)$; $D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

2	4	6	4	5
3	2	5	3	2
0	4	3	4	4
4	1	6	2	6
0	2	2	4	4

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 7

1.1. Обчислити: $P_7 : C_8^5 - A_9^2$.

1.2. Після бурі на ділянці між 20-м та 45-м кілометрами лінії електропередачі відбувся обрив провода. Яка ймовірність того, що обрив відбувся між 30-м і 40-м кілометрами лінії?

1.3. Два економісти заповнюють документи і складають їх у спільну папку. Ймовірність зробити помилку для першого економіста дорівнює 0,1, для другого – 0,2. Перший економіст заповнив 40 документів, другий – 60. Знайти ймовірність того, що навмання взятий з папки документ буде з помилкою.

1.4. За даним законом розподілу дискретної випадкової величини X

X	1	4	8	9	12	13
p	0,13	0,17	0,2	0,3	0,18	0,02

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x)=\begin{cases}0,\ x\le 1;\\ \dfrac{x^2-1}{3},1<x\le 2;\\ 1,\ x>2.\end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha,\beta)=(1{,}5;\ 2)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X,Y) задано відповідною таблицею:

Y \ X	4	6	7	9
5	0,15	0,05	0,2	0,15
6	0,2	0,1	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) Обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) Побудувати умовні закони розподілу $X\mid Y=y_2$; $Y\mid X=x_3$;

г) Обчислити умовні математичні сподівання $M\left(X\mid Y=y_2\right)$; $M\left(Y\mid X=x_3\right)$ та дисперсії $D\left(X\mid Y=y_2\right)$; $D\left(Y\mid X=x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водо провідної системи на території району отримали вибірку:

$$\begin{array}{ccccc} 4 & 5 & 1 & 5 & 6 \\ 3 & 4 & 3 & 0 & 1 \\ 5 & 1 & 4 & 4 & 0 \\ 1 & 4 & 1 & 2 & 4 \\ 4 & 2 & 4 & 3 & 5 \end{array}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 8

1.1. Обчислити: $A_8^2 : C_8^5 + P_2$.

1.2. Підприємство виготовляє 99 % стандартних виробів, причому 70 % із них – вищого сорту. Знайти ймовірність того, що взятий навмання виріб буде вищого сорту.

1.3. Ймовірність присутності студента на занятті дорівнює 0,7. Чому дорівнює найімовірніша кількість студентів, присутніх на занятті, якщо в групі 32 студенти.

1.4. За даним законом розподілу дискретної випадкової величини *X*

X	–2	–1	0	1	4	6
p	0,12	0,28	0,22	0,18	0,12	0,08

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина *X* задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, \ x \le 2; \\ (x-2)^2, \ 2 < x \le 3; \\ 1, \ x > 3. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини *X*;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (1; 2,5)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

$Y \backslash X$	5	8	10	12
7	0,1	0,05	0,15	0,1
12	0,15	0,3	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$.

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$; $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right)$; $M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right)$; $D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

$$\begin{matrix} 2 & 0 & 5 & 4 & 3 \\ 5 & 3 & 4 & 3 & 4 \\ 2 & 1 & 3 & 1 & 2 \\ 4 & 0 & 4 & 3 & 6 \\ 5 & 2 & 3 & 0 & 3 \end{matrix}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 9

1.1. Обчислити: $\left(C_9^4 + A_8^3\right): P_3$.

1.2. Студент прийшов складати іспит, знаючи з 50 питань програми лише 36. Яка ймовірність того, що студент не відповість на всі три запропоновані йому запитання?

1.3. Схожість насіння становить 60 %. Яка ймовірність того, що з чотирьох посіяних насінин зійдуть усі чотири ?

1.4. За даним законом розподілу дискретної випадкової величини X

X	–7	–5	–2	1	5	9
p	0,13	0,17	0,2	0,3	0,18	0,02

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, \ x \le 0; \\ \dfrac{x^3 + x^2}{2}, \ 0 < x \le 1; \\ 1, \ x > 1. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (0{,}5; 2)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

$Y \backslash X$	5	7	10	14
2	0,25	0,25	0	0,05
8	0,2	0,1	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$.

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$.

в) побудувати умовні закони розподілу $X \mid Y = y_2$; $Y \mid X = x_3$.

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right)$; $M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right)$; $D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

$$\begin{array}{ccccc} 2 & 6 & 5 & 1 & 6 \\ 1 & 3 & 1 & 3 & 2 \\ 3 & 4 & 2 & 3 & 5 \\ 2 & 5 & 6 & 0 & 2 \\ 5 & 1 & 3 & 2 & 4 \end{array}$$

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Варіант 10

1.1. Обчислити: $A_6^4 : \left(P_5 - C_9^6\right)$.

1.2. В урні 7 білих і 3 чорних кулі. Навмання одну за одною виймають дві кулі з поверненням. Знайти ймовірність того, що вийняті кулі одного кольору.

1.3. На склад автомашинами завозиться 75 % скляних банок, а решта – залізницею. Ймовірність розбивання скляних банок при транспортуванні на автомашинах дорівнює 0,03, а при транспортуванні залізницею – утричі менше. Узята навмання на складі банка виявилась розбитою. Визначити ймовірність того, що вона привезена залізницею.

1.4. За даним законом розподілу дискретної випадкової величини X

X	–3	–2	0	1	2	4
p	0,13	0,17	0,2	0,3	0,18	0,02

побудувати багатокутник розподілу. Знайти $M(X)$, $D(X)$ і $\sigma(X)$.

1.5. Випадкова величина X задана функцією розподілу ймовірностей $F(x)$:

$$F(x) = \begin{cases} 0, \; x \le -1; \\ \left(x^2 - 1\right)^2, \; -1 < x \le 0; \\ 1, \; x > 0. \end{cases}$$

Потрібно:

а) знайти щільність розподілу ймовірностей $f(x)$;

б) побудувати графіки функцій $F(x)$ і $f(x)$;

в) знайти $M(X)$ і дисперсію випадкової величини X;

г) знайти ймовірність того, що в результаті проведення випробування випадкова величина набуде значення з інтервалу $(\alpha, \beta) = (-1; -0,5)$.

1.6. Закон розподілу двовимірної дискретної випадкової величини (X, Y) задано відповідною таблицею:

Y \ X	6	8	9	10
7	0,15	0,05	0,1	0,1
10	0,25	0,2	0,05	0,1

Потрібно:

а) знайти числові характеристики складових системи: $M(X), D(X), \sigma(X)$ та $M(Y), D(Y), \sigma(Y)$;

б) обчислити кореляційний момент $K(X,Y)$ і коефіцієнт кореляції $\rho(X,Y)$;

в) побудувати умовні закони розподілу $X \mid Y = y_2$; $Y \mid X = x_3$;

г) обчислити умовні математичні сподівання $M\left(X \mid Y = y_2\right)$; $M\left(Y \mid X = x_3\right)$ та дисперсії $D\left(X \mid Y = y_2\right)$; $D\left(Y \mid X = x_3\right)$.

1.7. В результаті вибіркового аналізу добової кількості аварій водопровідної системи на території району отримали вибірку:

0	5	4	5	1
6	3	0	1	4
4	4	2	3	4
0	5	3	0	3
2	3	4	4	5

Для цієї вибірки потрібно:

а) побудувати дискретний розподіл частот;

б) побудувати полігон частот;

в) знайти вибіркову середню $\overline{x}$;

г) знайти виправлену вибіркову дисперсію s^2;

д) знайти емпіричну функцію розподілу $F^*(x)$ і побудувати її графік.

Відповіді

Варіант 1. **1.1.** 5. **1.2.** 0,192. **1.3.** 0,0085. **1.4.** $M\left(X\right)=1{,}84$; $D\left(X\right)=7{,}7944$. **1.5.** $M\left(X\right)=2{,}25$; $D\left(X\right)=0{,}3375$; $p=0{,}259$. **1.6.** X : 5,6; 17,44; 4,176; Y : 2,35; 2,2275; 1,49; $K_{xy}=1{,}14$; $\rho=0{,}1832$. **1.7.** $\overline{x}=2{,}76$; $\overline{s}^2=2{,}86$.

Варіант 2. **1.1.** 60. **1.2.** 0,9814. **1.3.** 32. **1.4.** $M\left(X\right)=0{,}35$; $D\left(X\right)=3{,}0275\,1$. **1.5.** $M\left(X\right)=1{,}333$; $D\left(X\right)=0{,}222$; $p=0{,}75$. **1.6.** X : 6,71; 4,51; 2,12; Y : 6,8; 2,16; 1,5; $K_{xy}=-0{,}66$; $\rho=-0{,}211$. **1.7.** $\overline{x}=2{,}96$; $\overline{s}^2=3{,}62$.

Варіант 3. **1.1.** 0,5. **1.2.** 0,833. **1.3.** 0,25. **1.4.** $M(X)=0,38$; $D(X)=3,62$. **1.5.** $M(X)=4,667$; $D(X)=0,722$; $p=0,333$. **1.6.** X : 3,6; 4,44; 2,11; Y : 3,7; 0,21; 0,46; $K_{xy}=-0,32$; $\rho=-0,33$. **1.7.** $\overline{x}=3,24$; $\overline{s}^2=3,94$.

Варіант 4. **1.1.** 12. **1.2.** 152. **1.3.** 0,027. **1.4.** $M(X)=8,25$; $D(X)=11,99$. **1.5.** $M(X)=3,71$; $D(X)=14,5$; $p=0,33$. **1.6.** X : 4,55; 5,35; 2,31; Y : 2,8; 2,16; 1,47; $K_{xy}=-0,39$; $\rho=-0,115$. **1.7.** $\overline{x}=3,12$; $\overline{s}^2=2,69$.

Варіант 5. **1.1.** 70. **1.2.** 0,988. **1.3.** 0,89. **1.4.** $M(X)=7,95$; $D(X)=17,1475$. **1.5.** $M(X)=1,35$; $D(X)=3,33$; $p=0,625$. **1.6.** X : 5,3; 2,31; 1,52; Y : 3,95; 2,05; 1,42; $K_{xy}=-1,48$; $\rho=-0,68$. **1.7.** $\overline{x}=2,24$; $\overline{s}^2=3,19$.

Варіант 6. **1.1.** 84. **1.2.** *а*) 0,188; *б*) 0,976. **1.3.** 0,3456. **1.4.** $M(X)=-0,85$; $D(X)=7,8275$. **1.5.** $M(X)=0,5$; $D(X)=0,15$; $p=0,875$. **1.6.** X : 4,85; 17,028; 4,13; Y : 4,2; 0,96; 0,98; $K_{xy}=-0,62$; $\rho=-0,153$. **1.7.** $\overline{x}=3,28$; $\overline{s}^2=2,88$.

Варіант 7. **1.1.** 18. **1.2.** 0,4. **1.3.** 0,16. **1.4.** $M(X)=7,57$; $D(X)=13,0651$. **1.5.** $M(X)=1,556$; $D(X)=0,080$; $p=0,58$. **1.6.** X : 6,3; 3,81; 1,952; Y : 4,35; 2,29; 1,49; $K_{xy}=-0,555$; $\rho=-0,19$. **1.7.** $\overline{x}=3,04$; $\overline{s}^2=1,87$.

Варіант 8. **1.1.** 3. **1.2.** 0,693. **1.3.** 23. **1.4.** $M(X)=0,62$; $D(X)=5,3556$. **1.5.** $M(X)=2,667$; $D(X)=18,06$; $p=0,25$. **1.6.** X : 8,45; 6,048; 2,46; Y : 10; 6; 2,45; $K_{xy}=-1,1$; $\rho=-0,1825$. **1.7.** $\overline{x}=2,88$; $\overline{s}^2=2,69$.

Варіант 9. **1.1.** 77. **1.2.** 0,009. **1.3.** 0,1296. **1.4.** $M(X)=-0,78$; $D(X)=17,23$. **1.5.** $M(X)=0,708$; $D(X)=0,05$; $p=0,81$. **1.6.** X : 7,3; 9,51; 3,084; Y : 4,7; 8,91; 2,99; $K_{xy}=1,89$; $\rho=0,205$. **1.7.** $\overline{x}=3,08$; $\overline{s}^2=3,16$.

Варіант 10. **1.1.** 10. **1.2.** 0,58. **1.3.** 0,01. **1.4.** $M(X)=0,01$; $D(X)=3,19$. **1.5.** $M(X)=0,533$; $D(X)=0,093$; $p=0,5625$. **1.6.** X : 7,75; 2,4875; 1,577; Y : 8,8; 2,16; 1,47; $K_{xy}=-0,3$; $\rho=-0,13$. **1.7.** $\overline{x}=3$; $\overline{s}^2=3,25$.

Індивідуальні завдання для контрольної роботи

Завдання 1. Підкидають два гральні кубики. Визначити ймовірність того, що:

а) сума очок не перевищує N;

б) добуток очок не перевищує N;

в) добуток очок ділиться на N.

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
N	3	4	5	6	7	8	9	10	3	4
Варіант	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
N	5	6	7	8	9	10	11	12	13	14
Варіант	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
N	15	16	17	18	19	20	3	4	5	6

Завдання 2. Серед n лотерейних білетів k виграшних. Навмання взяли m білетів. Визначити ймовірність того, що серед них l виграшних.

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
n	10	10	10	10	11	11	11	12	12	12	9	9	9	8	8
l	2	2	3	3	2	3	3	3	2	2	2	3	2	2	2
m	4	3	5	5	5	4	5	8	8	5	4	5	3	4	5
k	6	6	7	6	7	8	7	5	3	4	6	6	7	5	4
Варіант	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
n	8	10	10	10	12	8	8	8	8	8	9	9	9	9	9
l	3	4	5	4	4	2	2	2	3	1	2	3	2	4	3
m	4	6	7	6	8	3	3	4	5	4	3	4	6	5	5
k	5	5	7	7	6	4	5	3	4	2	5	4	3	5	4

Завдання 3. У ліфт k-поверхового будинку сіло n пасажирів ($n<k$). Кожен незалежно від інших із однаковою ймовірністю може вийти на довільному (починаючи з другого) поверсі. Визначити ймовірність того, що:

а) усі вийшли на різних поверхах;

б) принаймні двоє вийшли на одному поверсі.

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
k	6	**7**	8	9	10	11	12	13	14	13	12	11	10	9	8
n	4	4	5	5	6	4	4	3	3	4	3	3	4	3	4
Варіант	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
k	7	6	7	8	9	10	11	12	13	14	12	11	10	9	8
n	3	4	4	5	5	6	4	4	3	3	3	3	4	4	3

Завдання 4. У крузі радіусом R навмання обирають точку. Визначити ймовірність того, що вона потрапить в одну із двох фігур, які не перетинаються і площі яких дорівнюють S_1 та S_2.

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
R	11	12	13	14	11	12	13	14	11	12
S_1	2,25	2,37	2,49	2,55	2,27	2,39	2,51	2,57	2,29	2,41
S_2	3,52	3,5	3,54	1,57	0,57	5,57	2,57	3,56	3,6	3,58
Варіант	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
R	13	14	15	16	18	12	13	14	15	16
S_1	2,53	2,59	2,5	2,6	3,2	2,4	2,5	2,6	2,7	3,1
S_2	2,34	5,57	4,57	3,57	2,57	1,57	0,57	1,8	7,9	9,24
Варіант	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
R	11	12	13	14	15	11	12	13	13	14
S_1	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3,1	3,2	3,2
S_2	4,12	4,23	4,24	5,65	6,78	6,53	1,45	5,78	4,98	7,35

Завдання 5. У двох партіях k_1 та k_2% якісних виробів відповідно. Навмання вибирають по одному виробу з кожної партії. Яка ймовірність виявити серед них:

а) принаймні один бракований виріб;

б) два браковані вироби;

в) один якісний та один бракований виріб?

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
k_1	71	78	87	72	79	86	73	81	85	74	82	84	75	83	75
k_2	47	39	31	46	38	32	45	37	33	44	36	34	43	35	42
Варіант	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
k_1	77	47	39	31	72	38	32	73	81	33	44	36	84	75	83
k_2	41	71	78	87	46	79	86	45	37	85	74	82	74	82	43

Завдання 6. Ймовірність того, що в ціль влучає з одного пострілу перший снайпер дорівнює p_1, другий – p_2. Перший зробив n_1, другий – n_2 пострілів. Визначити ймовірність того, що ціль не була уражена (в неї не влучив жоден із снайперів).

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
p_1	0,61	0,62	0,63	0,64	0,65	0,66	0,67	0,68	0,69	0,71
p_2	0,55	0,54	0,53	0,52	0,51	0,49	0,48	0,47	0,46	0,45
n_1	2	3	2	3	2	3	2	3	2	3
n_2	3	2	3	2	3	2	3	2	3	2
Варіант	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
p_1	0,72	0,73	0,74	0,75	0,76	0,77	0,78	0,39	0,38	0,37
p_2	0,44	0,43	0,41	0,39	0,42	0,38	0,37	0,45	0,46	0,47
n_1	2	3	2	3	2	3	2	3	2	3
n_2	3	2	3	2	3	2	3	2	3	2
Варіант	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
p_1	0,36	0,35	0,34	0,33	0,32	0,31	0,29	0,28	0,27	0,26
p_2	0,48	0,49	0,51	0,52	0,53	0,54	0,55	0,56	0,57	0,58
n_1	2	3	2	3	2	3	2	3	2	3
n_2	3	2	3	2	3	2	3	2	3	2

Завдання 7. Із 1000 ламп n_i належить i-й партії, i = 1, 2, 3, $\sum_{i=1}^{3} n_i = 1000$. У першій партії – 6 %, у другій – 5 %, у третій – 4 % бракованих ламп. Навмання вибирають одну лампу. Визначити ймовірність того, що вибрана лампа – бракована.

Варіант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
n_1	100	430	170	520	360	700	240	80	630	500
n_2	250	180	540	390	600	90	610	710	230	320
Варіант	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
n_1	810	450	270	380	640	160	590	620	730	350
n_2	70	280	640	470	80	570	210	190	100	200
Варіант	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
n_1	90	220	290	350	470	680	710	180	260	650
n_2	690	550	700	440	360	230	160	270	620	140

Завдання 8. До крамниці надходять однотипні вироби з трьох заводів, причому i-й завод постачає m_i % виробів (i = 1, 2, 3). Серед виробів i-го заводу n_i % першосортних. Куплено один виріб. Він виявився першосортним. Визначити ймовірність того, що куплений виріб випущено j-м заводом.

Варіант	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
m_1	50	50	50	60	60	60	40	40	40	40	40	40	70	70	70
m_2	30	30	30	20	20	20	30	30	30	20	20	20	20	20	20
m_3	20	20	20	20	20	20	30	30	30	40	40	40	10	10	10
n_1	70	70	70	70	70	70	80	80	80	90	90	90	70	70	70
n_2	80	80	80	80	80	80	80	80	80	90	90	90	80	80	80
n_3	90	90	90	90	90	90	90	90	90	80	80	80	90	90	90
j	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Варіант	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
m_1	60	60	60	50	50	50	30	30	30	20	20	20	10	10	10
m_2	10	10	10	20	20	20	30	30	30	40	40	40	50	50	50
m_3	30	30	30	30	30	30	40	40	40	40	40	40	40	40	40
n_1	80	80	80	90	90	90	70	70	70	90	90	90	70	70	70
n_2	90	90	90	80	80	80	70	70	70	70	70	70	90	90	90
n_3	80	80	80	90	90	90	80	80	80	80	80	80	80	80	80
j	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3

Завдання 9. Імовірність виграшу в лотерею на один білет дорівнює p. Куплено n білетів. Знайти найбільш імовірне число виграшних білетів і відповідну ймовірність.

Варіант	1	2	3	4	5	6	7	8	9	10
p	0,3	0,3	0,3	0,3	0,3	0,3	0,4	0,4	0,4	0,4
n	10	14	13	12	11	15	11	13	14	10
Варіант	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
p	0,4	0,4	0,5	0,4	0,5	0,5	0,5	0,5	0,6	0,6
n	12	15	12	12	11	13	14	15	13	11
Варіант	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
p	0,6	0,6	0,6	0,6	0,7	0,7	0,8	0,7	0,7	0,7
n	12	10	15	14	14	10	15	11	12	13

Завдання 10. Для дискретної випадкової величини відомий ряд розподілу. Побудувати багатокутник розподілу та графік функції розподілу цієї випадкової величини. Знайти числові характеристики: $M(2X), D(X), \sigma(X)$.

10.1.

x	–1	0	1	2
p_i	0,2	0,1	0,3	0,4

10.2.

x	–2	–1	0	2	4
p_i	1/3	2/15	1/5	2/15	3/15

10.3.

x	–1	2	3	5
p_i	0,2	0,3	0,4	0,1

10.4.

x	–1	0	1	2	3
p_i	1/4	1/4	1/8	1/8	1/4

10.5.

x	–2	–1	0	1
p_i	0,2	0,4	0,1	0,3

10.6.

x	–1	0	1	2	3
p_i	0,1	0,4	0,2	0,1	0,2

10.7.

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

10.8.

x	–1	0	2	4	5
p_i	0,1	0,1	0,3	0,2	0,3

10.9.

x	–3	–2	–1	1
p_i	0,3	0,2	0,1	0,4

10.10.

x	–3	–2	1	3	4
p_i	0,2	0,1	0,3	0,3	0,1

10.11.

x	–3	–2	–1	0	1
p_i	0,3	0,1	0,1	0,1	0,4

10.12.

x	–1	0	2	3	5
p_i	0,1	0,3	0,1	0,2	0,3

10.13.

x	–3	–1	0	1
p_i	0,15	0,25	0,2	0,4

10.14.

x	–3	–2	–1	0	1
p_i	0,25	0,3	0,15	0,2	0,1

10.15.

x	–3	–2	0	–1
p_i	0,1	0,15	0,35	0,4

10.16.

x	–1	0	1	2
p_i	0,1	0,4	0,2	0,3

10.17.

x	0	1	2	3
p_i	0,1	0,2	0,4	0,3

10.18.

x	1	3	5	7
p_i	0,4	0,1	0,2	0,3

10.19.

x	–2	0	2	4
p_i	0,3	0,1	0,2	0,4

10.20.

x	–3	0	3
p_i	0,1	0,2	0,7

10.21.

x	1	2	3
p_i	0,3	0,2	0,5

10.22.

x	–2	1	4
p_i	0,4	0,5	0,1

10.23.

x	–2	–1	1	2
p_i	0,1	0,3	0,2	0,4

10.24.

x	– 4	0	4	8
p_i	0,1	0,2	0,3	0,4

10.25.

x	–1	0	1
p_i	0,5	0,1	0,4

10.26.

x	–1	1	2	4
p_i	0,2	0,1	0,3	0,4

10.27.

x	–2	0	3	5
p_i	0,1	0,4	0,4	0,1

10.28.

x	–3	–2	0	1
p_i	1/3/	1/6	1/6	1/3

10.29.

x	–1	0	1	2
p_i	0,8	0,15	0,025	0,025

10.30.

x	1	2	3	4
p_i	0,6	0,1	0,2	0,1

Завдання 11. Неперервна випадкова величина задана щільністю розподілу $f(x)$. Зобразити диференціальну $f(x)$ та інтегральну $F(x)$ функції розподілу випадкової величини. Обчислити $M(X), D(4X), \sigma(X)$.

11.1. $f(x)=\begin{cases}0, x\le 1;\\ \frac{2}{3}x, 1<x\le 2;\\ 0, x>2.\end{cases}$

11.2. $f(x)=\begin{cases}0, x\le 0;\\ \frac{1}{4}, 0<x\le 4;\\ 0, x>4.\end{cases}$

11.3. $f(x)=\begin{cases}0, x\le 0;\\ 3x^2, 0<x\le 1;\\ 0, x>1.\end{cases}$

11.4. $f(x)=\begin{cases}0, x\le 0;\\ \cos x, 0<x\le \frac{\pi}{2};\\ 0, x>\frac{\pi}{2}.\end{cases}$

11.5. $f(x)=\begin{cases}0, x\le 1;\\ \frac{2x-1}{2}, 1<x\le 2;\\ 0, x>2.\end{cases}$

11.6. $f(x)=\begin{cases}0, x\le 0;\\ 6x+2, 0<x\le \frac{1}{3};\\ 0, x>\frac{1}{3}.\end{cases}$

11.7. $f(x)=\begin{cases}0, x\le 0;\\ \frac{x}{2}, 0<x\le 2;\\ 0, x>2.\end{cases}$

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

11.9. $f(x)=\begin{cases}0, x\le 0;\\ \frac{3}{8}x^2, 0<x\le 2;\\ 0, x>2.\end{cases}$

11.10. $f(x)=\begin{cases}0, x\le 0;\\ 5x^4, 0<x\le 1;\\ 0, x>1.\end{cases}$

11.11. $f(x)=\begin{cases}0, x\le 2;\\ \frac{1}{2}, 2<x\le 4;\\ 0, x>4.\end{cases}$

11.12. $f(x)=\begin{cases}0, x<1;\\ 2x-2, 1\le x\le 2;\\ 0, x>2.\end{cases}$

11.13. $f(x)=\begin{cases}0, x<0;\\ 1, 0\le x\le 1;\\ 0, x>1.\end{cases}$

11.14. $f(x)=\begin{cases}0, x<0;\\ \frac{x}{2}, 0\le x\le 2;\\ 0, x>2.\end{cases}$

11.15. $f(x)=\begin{cases}0, x<0;\\ 2x, 0\le x\le 1;\\ 0, x>1.\end{cases}$

11.16. $f(x)=\begin{cases}0, x\le -1;\\ \frac{1}{3}, -1<x\le 2;\\ 0, x>2.\end{cases}$

11.17. $f(x)=\begin{cases}0, x\le 0;\\ 2x+\frac{8}{3}, 0<x\le \frac{1}{3};\\ 0, x>\frac{1}{3}.\end{cases}$

11.18. $f(x)=\begin{cases}0, x<0;\\ \frac{1}{2}\sin x, 0\le x\le \pi;\\ 0, x>\pi.\end{cases}$

11.19. $f(x)=\begin{cases}0, x\le 0;\\ 4x^3, 0<x\le 1;\\ 0, x>1.\end{cases}$

11.20. $f(x)=\begin{cases}0, x\le 0;\\ \frac{2x+1}{2}, 0<x\le 1;\\ 0, x>1.\end{cases}$

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

11.22. $f(x)=\begin{cases}0, x\le 0;\\ 24x^2, 0<x\le \frac{1}{2};\\ 0, x>\frac{1}{2}.\end{cases}$

11.23. $f(x)=\begin{cases}0, x\le 0;\\ 2\cos x, 0<x\le \frac{\pi}{6};\\ 0, x>\frac{\pi}{6}.\end{cases}$

11.24. $f(x)=\begin{cases}0, x\le 0;\\ 0,5; 0<x\le 2;\\ 0, x>2.\end{cases}$

11.25. $f(x)=\begin{cases}0, x\le 0;\\ \frac{x}{8}, 0<x\le 4;\\ 0, x>4.\end{cases}$

11.26. $f(x)=\begin{cases}0, x\le 0;\\ \frac{2x}{25}, 0<x\le 5;\\ 0, x>5.\end{cases}$

11.27. $f(x)=\begin{cases}0,\ x\le 0;\\ 6x+2,\ 0<x\le\dfrac{1}{3};\\ 0,\ x>\dfrac{1}{3}.\end{cases}$

11.28. $f(x)=\begin{cases}0,\ x\le\dfrac{3\pi}{2};\\ -\sin x,\ \dfrac{3\pi}{2}<x\le 2\pi;\\ 0,\ x>2\pi.\end{cases}$

11.29. $f(x)=\begin{cases}0,\ x\le\dfrac{\pi}{2};\\ -\cos x,\ \dfrac{\pi}{2}<x\le\pi;\\ 0,\ x>\pi.\end{cases}$

11.30. $f(x)=\begin{cases}0,\ x\le 0;\\ \dfrac{1}{7},\ 0<x\le 7;\\ 0,\ x>7.\end{cases}$

Відповіді до варіанта 30. **1.** 1) 0,417; 2) 0,583; 3) 0,417. **2.** 0,3. **3.** 1) 0,06; 2) 0,94. **4.** 0,012. **5.** 1) 0,6431; 2) 0,0969; 3) 0,5462. **6.** 0,071. **7.** 0,0544. **8.** 0,381. **9.** 9. **10.** 3,6; 1,16; 1,078. **11.** 3,5; 65,33.

Завдання контрольного тесту

1. Кількість комбінацій з n елементів по k $(k \le n)$ обчислюють за формулою:

$$а)\ \tilde{N}_n^k = \frac{n!}{k!(n-k)!};\quad б)\ \tilde{N}_n^k = \frac{n}{k(n-k)!};\quad в)\ \tilde{N}_n^k = \frac{n!}{k!(n-k)}.$$

2. Для несумісних подій A та B виконується рівність:

$$а)\ p(A \cup B) = p(A) + p(B);\quad б)\ p(A \cup B) = p(A)p(B);$$

$$в)\ p(A \cup B) = p(A) + p(B) - p(AB).$$

3. Для незалежних подій A та B виконується рівність:

$$а)\ p(A \cap B) = p(A) + p(B);\quad б)\ p(A \cap B) = p(A)p(B) - p(A);$$

$$в)\ p(A \cap B) = p(A)p(B).$$

4. Умовну ймовірність події A за умови, що відбулась подія B, обчислюють за формулою:

$$а)\ p(A \mid B) = \frac{p(A \cap B)}{p(A)};\quad б)\ p(A \mid B) = \frac{p(A \cap B)}{p(B)};\quad в)\ p(B \mid A) = \frac{p(A \cap B)}{p(A)}.$$

5. Ймовірність настання принаймні однієї з подій $A_1, A_2, \dots A_n$, незалежних в сукупності, обчислюють за формулою:

$$а)\ p(A) = 1 - p_1 p_2 \dots p_n;\quad б)\ p(A) = 1 - q_1 q_2 \dots q_n;$$

$$в)\ p(A) = 1 - p_1 p_2 \dots p_n + q_1 q_2 \dots q_n.$$

6. Формула повної ймовірності має вигляд:

$$а)\ p(A) = \sum_{s=1}^{n} p(H_s)p(A \mid H_s);\quad б)\ p(A) = \sum_{s=1}^{n} p(A_s)p(A \mid H_s);$$

$$в)\ p(A) = \sum_{s=1}^{n} p(H_s) + p(A \mid H_s).$$

7. Ймовірність настання події A k разів у n випробуваннях $(0 \le k \le n)$ (ймовірність настання події A за кожного випробування дорівнює p) знаходять за формулою Бернуллі:

$$а)\ p_n(k) = C_k^n p^k (1-p)^{n-k};\quad б)\ p_n(k) = C_k^n p^{n-k} (1-p)^{n-k};$$

$$в)\ p_n(k) = C_n^k p^k (1-p)^{n-k}.$$

8. Неперервна випадкова величина X має нормальний закон розподілу з параметрами a, σ^2, якщо щільність розподілу має вигляд:

$$а)\ f(x)=\frac{1}{2\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}};\quad б)\ f(x)=\frac{1}{2\sigma\sqrt{\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}};$$

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

9. Записати числові характеристики дискретних випадкових величин.

10. Який зв’язок між інтегральною та диференціальною функціями розподілу ймовірностей? Записати формули визначення числових характеристик неперервних випадкових величин.

11. Як позначають та визначають емпіричну функцію розподілу? Які основні властивості цієї функції?

12. Вказати числові характеристики вибірки та формули їх обчислення.

13. Які бувають статистичні розподіли вибірки?

Додатки

Таблиця Д.1

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

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2035	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0941	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0271	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0159	0155	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
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	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблиця Д.2

Значення функції Лапласа $\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,00000	00399	00798	01197	01595	01994	02392	02790	03188	03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43872	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48860	48899
2,3	48928	48966	48983	49010	49036	49061	49086	49110	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3,0	0,49865		3,1	49903	3,2	49931	3,3	49952	3,4	49966
3,5	49977		3,6	49984	3,7	49989	3,8	49993	3,9	49995
4,0	499968									
4,5	499997									
5,0	49999997									

Таблиця Д.3

Значення χ^2 залежно від кількості степенів вільності k

і рівня значущості α

k	α 0,95	α 0,90	α 0,50	α 0,30	α 0,20	α 0,10	α 0,05	α 0,025	α 0,01
1	0,004	0,016	0,455	1,074	1,642	2,71	3,84	5,0	6,6
2	0,103	0,211	1,386	2,41	3,22	4,60	5,99	7,4	9,2
3	0,352	0,584	2,37	3,67	4,64	6,25	7,82	9,4	11,3
4	0,711	1,064	3,36	4,88	5,99	7,78	9,49	11,1	13,3
5	1,145	1,61	4,35	6,06	7,29	9,24	11,07	12,8	15,1
6	1,635	2,20	5,35	7,23	8,56	10,64	12,59	14,4	16,8
7	2,17	2,83	6,35	8,38	9,80	12,02	14,07	16,0	18,5
8	2,73	3,49	7,34	9,52	11,03	13,36	15,51	17,5	20,1
9	3,32	4,17	8,34	10,66	12,24	14,68	16,92	19,0	21,7
10	3,94	4,86	9,34	11,78	13,44	15,99	18,31	20,5	23,2
11	4,58	5,58	10,34	12,90	14,63	17,28	19,68	21,9	24,7
12	5,23	6,30	11,34	14,01	15,84	18,55	21,0	23,3	26,2
13	5,89	7,04	12,34	15,12	16,98	19,81	22,4	24,7	27,7
14	6,57	7,79	13,34	16,22	18,15	21,1	23,7	26,1	29,1
15	7,26	8,55	14,34	17,32	19,31	22,3	25,0	27,5	30,6
16	7,96	9,31	15,34	18,42	20,5	23,5	26,3	28,8	32,0
17	8,67	10,08	16,34	19,51	21,6	24,8	27,6	30,2	33,4
18	9,39	10,86	17,34	20,6	22,8	26,0	28,9	31,5	34,8
19	10,11	11,65	18,34	21,7	23,9	27,2	30,1	32,9	36,2
20	10,85	12,44	19,34	22,8	25,0	28,4	31,4	34,2	37,6
21	11,59	13,24	20,3	23,9	26,2	29,6	32,7	35,5	38,9
22	12,34	14,04	21,3	24,9	27,3	30,8	33,9	36,8	40,3
23	13,09	14,85	22,3	26,0	28,4	32,0	35,2	38,1	41,6
24	13,85	15,66	23,3	27,0	29,6	33,2	36,4	39,4	43,0
25	14,61	16,47	24,3	28,2	30,7	34,4	37,7	40,6	44,3
26	15,38	17,29	25,3	29,2	31,8	35,6	38,9	41,9	45,6
27	16,15	18,11	26,3	30,3	32,9	36,7	40,1	43,2	47,0
28	16,93	18,94	27,3	31,4	34,0	37,9	41,3	44,5	48,3
29	17,71	19,77	28,3	32,5	35,1	39,1	42,6	45,7	49,6
30	18,49	20,60	29,3	33,5	36,2	40,3	43,8	47,0	50,9

Таблиця Д.4

Значення $t_\gamma = t(k, \gamma)$ (розподіл Стьюдента)

Кількість степенів вільності	$\gamma = 1 - \alpha$					
k	0,9	0,95	0,98	0,99	0,998	0,999
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,21	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,37	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,06	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,74	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,47	2,76	3,41	3,67
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	20,2	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Таблиця Д.5.1

Критичні точки розподілу *F* Фішера–Снедекора
за рівня значущості $\alpha = 0,01$

k_2	k_1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	4052	4999	5403	5625	5764	5889	5928	5981	6022	6056	6082	6106
2	98,49	99,01	90,17	99,25	99,33	99,30	99,34	99,36	99,36	99,40	99,41	99,42
3	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	15,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,86	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45

Таблиця Д.5.2

Критичні точки розподілу *F* Фішера–Снедекора
за рівня значущості $\alpha = 0,05$

k_2	k_1											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,37	19,40	19,41
3	10,12	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,20	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38

Таблиця Д.6

Значення $p_k = \dfrac{\lambda^k}{k!} e^{-\lambda}$

k	λ								
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494
4		0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111
5				0,0001	0,0002	0,0004	0,0007	0,0012	0,0020
6							0,0001	0,0002	0,0003

k	λ								
	1,0	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0
0	0,3679	0,1353	0,0498	0,0183	0,0067	0,0025	0,0009	0,0003	0,0001
1	0,3679	0,2707	0,1494	0,0733	0,0337	0,0149	0,0064	0,0027	0,0011
2	0,1839	0,2707	0,2240	0,1465	0,0842	0,0446	0,0223	0,0107	0,0050
3	0,0613	0,1804	0,2240	0,1954	0,1404	0,0892	0,0521	0,0286	0,0150
4	0,0153	0,0902	0,1680	0,1954	0,1755	0,1339	0,0912	0,0573	0,0337
5	0,0031	0,0361	0,1008	0,1563	0,1755	0,1606	0,1277	0,0916	0,0607
6	0,0005	0,0120	0,0504	0,1042	0,1462	0,1606	0,1490	0,1221	0,0911
7	0,0001	0,0034	0,0216	0,0595	0,1044	0,1377	0,1490	0,1396	0,1171
8		0,0009	0,0081	0,0298	0,0653	0,1033	0,1304	0,1396	0,1318
9		0,0002	0,0027	0,0132	0,0363	0,0688	0,1014	0,1241	0,1318
10			0,0008	0,0053	0,0181	0,0413	0,0710	0,0993	0,1186
11			0,0002	0,0019	0,0082	0,0225	0,0452	0,0722	0,0970
12			0,0001	0,0006	0,0034	0,0113	0,0264	0,0481	0,0728
13				0,0002	0,0013	0,0052	0,0142	0,0296	0,0504
14				0,0001	0,0005	0,0022	0,0071	0,0169	0,0324
15					0,0002	0,0009	0,0033	0,0090	0,0194
16						0,0003	0,0014	0,0045	0,0109
17						0,0001	0,0006	0,0021	0,0058
18							0,0002	0,0009	0,0029
19							0,0001	0,0004	0,0014
20								0,0002	0,0006
21								0,0001	0,0003
22									0,0001

ТРИКУТНИК ПАСКАЛЯ

1

1 1

1 2 1

1 3 3 1

1 4 6 4 1

1 5 10 10 5 1

1 6 15 20 15 6 1

1 7 21 35 35 21 7 1

1 8 28 56 70 56 28 8 1

1 9 36 84 126 126 84 36 9 1

1 10 45 120 210 252 210 120 45 10 1

Список використаної літератури

1. Барковський В. В. Теорія ймовірностей та математична статистика / В. В. Барковський, Н. В. Барковська, О. К. Лопатін. – К. : ЦУЛ, 2002. – 448 с.

2. Вентцель Е. С. Теория вероятностей и математическая статистика / Е. С. Вентцель. – М. : Высш. шк., 1999. – 576 с.

3. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М. : Высш. шк., 1997. – 400 с.

4. Гмурман В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М. : Высш. шк., 2004. – 479 с.

5. Гнеденко Б. В. Курс теории вероятностей. – М. : Наука, 1988. – 446 с.

6. Дорош А. К. Теорія ймовірностей та математична статистика / А. К. Дорош, О. П. Коханівський. – К. : НТУУ «КПІ», 2006. – 268 с.

7. Дорош А. К. Теорія ймовірностей та математична статистика. Збірник задач та індивідуальних завдань / А. К. Дорош, О. П. Коханівський. – К. : «Київський політехнік», 2000. – 125 с.

8. Іванюта І. Д. Елементи теорії ймовірностей та математичної статистики / І. Д. Іванюта, В. І. Рибалка, І. А. Рудоміно-Дусятська. – К. : Слово, 2003. – 272 с.

9. Каніовська І. Ю. Теорія ймовірностей у прикладах і задачах / І. Ю. Каніовська. – К. : ІВЦ "Видавництво «Політехніка»", ТОВ "Фірма «Періодика»", 2004. – 156 с.

10. Королюк В. С. Справочник по теории вероятностей и математической статистике / В. С. Королюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин. – М. : Наука. Гл. ред. физ.-мат. л-ры, 1985. – 640 с.

11. Коханівський О. П. Збірник задач з теорії ймовірностей / О. П. Коханівський, О. І. Кушлик, Б. П. Орел та ін. – К. : НТУУ «КПІ», 1999. – 40 с.

12. Розанов Ю. А. Лекции по теории вероятностей / Ю. А. Розанов. – М. : «Наука», 1986. – 120 с.

ЗМІСТ