МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

Методичні рекомендації і завдання до виконання контрольних робіт з навчальної дисципліни "ТЕОРІЯ ЙМОВІРНОСТЕЙ, ІМОВІРНІСНІ ПРОЦЕСИ ТА МАТЕМАТИЧНА СТАТИСТИКА"

для студентів напряму підготовки
6.050101 "Комп'ютерні науки"
заочної форми навчання

Харків
ХНЕУ ім. С. Кузнеця
2016

Затверджено на засіданні кафедри вищої математики й економіко-математичних методів.

Протокол № 8 від 04.03.2016 р.

Укладачі: В. Ф. Сенчуков
Т. В. Денисова

М 54 **Методичні** рекомендації і завдання до виконання контрольних робіт з навчальної дисципліни "Теорія ймовірностей, імовірнісні процеси та математична статистика" для студентів напряму підготовки 6.050101 "Комп'ютерні науки" заочної форми навчання / уклад. В. Ф. Сенчуков, Т. В. Денисова. – Харків : ХНЕУ ім. С. Кузнеця, 2016. – 112 с.

Викладено теоретичний матеріал у стислому вигляді з ілюстративними прикладами, контрольні запитання для самодіагностики, зміст завдань і зразок виконання типового варіанта контрольної роботи.

Рекомендовано для студентів напряму підготовки 6.050101 "Комп'ютерні науки" заочної форми навчання.

Вступ

Запропоновані методичні рекомендації і завдання для контрольних робіт укладені відповідно до робочої програми навчальної дисципліни "Теорія ймовірностей, імовірнісні процеси та математична статистика" для студентів напряму підготовки "Комп'ютерні науки" й охоплюють увесь теоретичний матеріал із відповідних розділів дисципліни.

Теоретичний матеріал викладено у стислому вигляді з ілюстративними прикладами. Завдання для самостійного розв'язання містять набір задач, які відображають основні теоретичні положення та їх застосування. Наведено зразок виконання типового варіанта контрольної роботи.

У процесі заочного навчання контрольні роботи мають виключно важливу роль. Вони не слугують матеріалом для остаточного оцінювання знань студента, але дозволяють викладачеві судити про його самостійну роботу, про те, як студент готується до складання заліку чи екзамену. В останні роки практикується захист контрольних робіт – співбесіда студента з викладачем стосовно виконаних робіт, яка дає змогу з'ясувати ступінь підготовки студента. Треба чітко усвідомити, що виконання контрольної роботи можна розпочинати тільки тоді, коли є впевненість у тому, що засвоєно весь теоретичний матеріал того чи іншого завдання. Роботи, виконані не за належним варіантом, не зараховуються.

Контрольні роботи студент повинен виконувати самостійно, дотримуючись правил:

перед розв'язанням кожної задачі необхідно записати її умову, причому розв'язки та відповіді розташовувати в тій послідовності, в якій наведені задачі контрольного завдання. Робота може містити й неповні (часткові) розв'язки задач, якщо студент не може самостійно довести розв'язання задачі до логічного кінця (у таких випадках слід звернутися за консультацією до викладача та допрацювати незрозумілі моменти). Контрольна робота вважається захищеною, якщо вона виконана повністю і студент може висвітлити хід розв'язання кожної задачі;

розв'язання задачі слід подавати разом з усіма проміжними перетвореннями. Необхідно керуватися зразками розв'язання задач, які вміщені в підручниках або методичних посібниках. Хід розв'язання задачі треба супроводжувати стислим поясненням (що до чого, куди, як, звідки);

писати контрольну роботу потрібно чітким, розбірливим почерком, чорнилами (пастою) синього або чорного кольору (але не червоного);

не допускати перекреслювань, будь-яких позначень, які не є загальноприйнятими в літературі з відповідної дисципліни; рисунки виконувати за допомогою креслярських інструментів, а ескізи – олівцем від руки;

наприкінці роботи слід зазначити, якою літературою (підручниками, посібниками) користувався студент під час вивчення дисципліни та виконання контрольної роботи;

на обкладинці зошита, в якому подано виконану контрольну роботу, необхідно вказати назву дисципліни, курс навчання, спеціальність, номер групи, прізвище, ініціали.

Отже, контрольна робота повинна поєднувати ретельно продуманий зміст із бездоганним літературним і технічним її оформленням.

Запорукою успішного розв'язання завдань контрольної роботи є:

систематичність самостійних занять, що сприяє розвиткові творчої думки (безсистемні заняття з тривалими перервами не можуть дати міцних знань; ніякі короткочасні, епізодичні, навіть дуже інтенсивні заняття не дають таких результатів, які забезпечуються систематичним вивченням матеріалу);

самоконтроль як спосіб перевірки ступеня засвоєння матеріалу (відповіді на запитання для самоперевірки допомагають більш глибоко усвідомити матеріал дисципліни та закріпити його в пам'яті; треба намагатись відповідати на запитання, не підглядаючи в підручник).

Варіант контрольної роботи визначається числом, яке відповідає останній цифрі номера залікової книжки або студентського квитка. Номери задач для кожного варіанта наведені в табл. 1.

Таблиця 1

Номери задач контрольної роботи за варіантами

Номер варіанта	Номери задач контрольної роботи
0	1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0
1	1.1, 2.1, 3.1, 4.1, 5.1, 6.1, 7.1, 8.1, 9.1, 10.1
2	1.2, 2.2, 3.2, 4.2, 5.2, 6.2, 7.2, 8.2, 9.2, 10.2
3	1.3, 2.3, 3.3, 4.3, 5.3, 6.3, 7.3, 8.3, 9.3, 10.3
4	1.4, 2.4, 3.4, 4.4, 5.4, 6.4, 7.4, 8.4, 9.4, 10.4
5	1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5
6	1.6, 2.6, 3.6, 4.6, 5.6, 6.6, 7.6, 8.6, 9.6, 10.6
7	1.7, 2.7, 3.7, 4.7, 5.7, 6.7, 7.7, 8.7, 9.7, 10.7
8	1.8, 2.8, 3.8, 4.8, 5.8, 6.8, 7.8, 8.8, 9.8, 10.8
9	1.9, 2.9, 3.9, 4.9, 5.9, 6.9, 7.9, 8.9, 9.9, 10.9

1. Теоретичні відомості

1.1. Об'єкт і предмет теорії ймовірностей (ТЙ), основні поняття. Класичне означення ймовірності випадкової події

Витоком основних понять ТЙ є поняття експерименту. **Експеримент** – спосіб дослідження явища (процесу) за строго визначених умов, які дозволяють відтворити явище достатню кількість разів. Розрізняють експерименти:

детерміновані (передбачувані) – такі, що результат експерименту визначається певним числом однозначно. Таким експериментам відповідають різні закони природничих наук;

недетерміновані (*непередбачувані, стохастичні, ймовірнісні, випадкові*) – це такі експерименти E, результати проведення яких за одних і тих самих умов можуть бути різними; кажуть, що "результат залежить від випадку".

Приклади: E_1 – підкидання монети над плоскою горизонтальною поверхнею; E_2 – гра в "Спортлото".

Теорія ймовірностей – математична наука, *об'єктом* вивчення якої є стохастичні експерименти, а *предметом* – установлення загальних закономірностей, яким вони підкоряються.

У ТЙ одноразове проведення стохастичного експерименту називають **випробуванням**, а результат випробування ω – **наслідком** або **елементарною подією** (**е/п**). Сукупність (множина) Ω усіх можливих наслідків, які можуть мати місце в результаті проведення E, називають **простором елементарних подій**.

Наприклад, для E_1 – підкидання монети, маємо: $\Omega_1 = \{\omega_1, \omega_2\} = \{Г, Ц\} = \{\text{випадає герб, випадає цифра}\}$; $\omega_1 = Г$, $\omega_2 = Ц$ – е/п (наслідки).

Простір е/п Ω може бути скінченною або нескінченною множиною, дискретною або неперервною:

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ – скінченна множина;

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$ – нескінченна дискретна множина;

$\Omega = \{\omega \mid \omega \in [a, b]\}$ – нескінченна неперервна множина.

Подумайте, який простір е/п має E_3 =<підкидаються дві монети>.

Будь-яку підмножину S простору е/п Ω: $S \subset \Omega$, називають **випадковою подією** (далі – просто **подією**).

Наприклад, $(\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\},\ S = \{\omega_1, \omega_3\}) \Rightarrow S \subset \Omega$ – подія у просторі Ω.

Якщо подія S має місце (відбулася, настала, з'явилася, здійснилася), то це означає, що настав один із наслідків E, які належать S, або декілька з них, або всі. І навпаки, якщо відбулася хоча б одна з е/п, то це означатиме, що відбулася і подія S.

Наприклад, для E_3 – підкидання двох монет – простір е/п має вигляд: $\Omega = \{ЦЦ, ЦГ, ГЦ, ГГ\}$. Якщо подія $S_1 = \{ЦЦ, ГГ\}$ настала, то це означає, що випадає або дві цифри, або два герби. Якщо в результаті проведення E_3 з'явиться наслідок $ЦЦ$ або $ГГ$, то це означатиме, що подія $S_1 = \{ЦЦ, ГГ\}$ відбулася. *Опишіть* самостійно, у чому полягають події: $S_2 = \{ГГ, ЦГ, ГЦ\}$, $S_3 = \{ГГ\}$, $S_4 = \{ЦГ, ГЦ\}$, $S_5 = \{ГГ, ЦЦ, ЦГ, ГЦ\}$, $S_6 = \{ГГГ\}$.

Е/п, які належать події S (як елементи простору Ω), називають **сприятливими для S** (**сприяють появі S**). На основі цього означення проводять "класифікацію подій".

Якщо подія S така, що їй не сприяє жодна з е/п (сприяють усі е/п) із простору Ω, то її називають **неможливою (вірогідною)**.

У нашому розгляді: S_6 – неможлива подія; S_5 – вірогідна подія.

Події S_1, S_2 називають **сумісними**, якщо існують е/п, які сприяють обом подіям, інакше (мовою теорії множин):

$$S_1,\ S_2 \text{ – сумісні} \Leftrightarrow S_1 \cap S_2 \neq \varnothing;$$

у протилежному випадку події S_1 і S_2 називають **несумісними**:

$$S_1,\ S_2 \text{ – несумісні} \Leftrightarrow S_1 \cap S_2 = \varnothing.$$

Якщо S_1 – підмножина S_2: $S_1 \subset S_2$, то кажуть, що настання події S_1 **тягне за собою** настання S_2.

Кажуть, що множина подій S_1, S_2, …, S_k складає (утворює) **повну групу подій**, якщо з настанням якоїсь е/п із Ω відбувається хоча б одна із подій S_i, $i = \overline{1,k}$.

Зазвичай розглядають повні групи несумісних подій, тобто коли $S_i \bigcap S_j = \varnothing \ \forall i \neq j$.

У нас: $\{S_1, S_2\}$ – повна група подій сумісних, оскільки наслідок $ГГ$ належить обом подіям; $\{S_1, S_4\}$ – повна група подій несумісних.

Установіть: $S_5 = \{ГГ, ЦЦ, ЦГ, ГЦ\} = \Omega$ – повна група *яких* подій?

Операції над подіями

1. $S = S_1 \bigcup S_2$ – **сума подій** – це подія S, яка відбувається тоді і тільки тоді, коли настає хоча б одна із двох подій S_1 і S_2 (рис. 1.1).

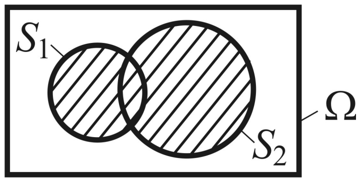

Рис. 1.1. **Сума подій**

2. $S = S_1 \bigcap S_2$ – **перетин подій** – це подія S, яка відбувається тоді і тільки тоді, коли настають обидві події S_1 і S_2 (рис. 1.2).

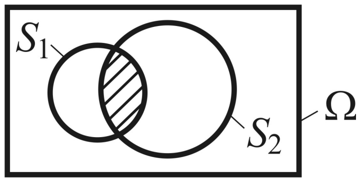

Рис. 1.2. **Перетин подій**

3. $S = S_1 \setminus S_2$ – **різниця подій** – це подія S, яка відбувається тоді і тільки тоді, коли здійснюється подія S_1 і не настає подія S_2 (рис. 1.3).

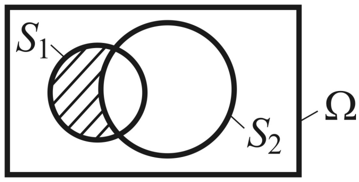

Рис. 1.3. **Різниця подій**

4. $\overline{S} = \Omega \setminus S$ – **протилежна подія** відносно S – це подія, яка настає тоді і тільки тоді, коли не настає S (різниця між простором Ω і подією S) (рис. 1.4).

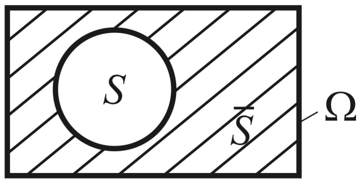

Рис. 1.4. **Протилежна подія**

Класичне означення ймовірності

В описовому плані в ТЙ під *імовірністю події* S розуміють число, яке характеризує ступінь можливості здійснення цієї події; позначення: $P(S)$.

Існує декілька підходів до встановлення вказаного числа як ступеня можливості настання (появи) S: *класичний*, *геометричний*, *статистичний*, *аксіоматичний.*

Нехай $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ – скінченна множина елементарних подій – простір е/п (наслідків), які:

1) **попарно несумісні**: $\{\omega_i\} \cap \{\omega_j\} = \varnothing \ \ \forall i \neq j$, тобто поява ω_i чи ω_j виключає появу іншої е/п;

2) **єдино можливі**, тобто з проведенням E якесь із ω_i, $i = \overline{1, n}$, обов'язково повинно настати;

3) **рівноможливі**, тобто немає об'єктивних причин стверджувати, що ступінь можливості настання е/п ω_i буде іншим, ніж для ω_j.

Саме для такої повної групи подій Ω формувався класичний підхід до означення ймовірності.

Нехай події S сприяє m е/п ($0 \le m \le n$).

Імовірністю події S у класичному розумінні називають відношення числа m е/п, які сприяють S, до числа n усіх можливих е/п:

$$P(S) = \frac{m}{n}. \tag{1.1}$$

Основні властивості:

1) $0 \le P(S) \le 1$, бо $0 \le m \le n \ \Leftrightarrow \ \frac{0}{n} = 0 \le \frac{m}{n} \le \frac{n}{n} = 1$;

2) $P(S) = 0 \ \Leftrightarrow \ S = \varnothing$ – неможлива подія;

3) $P(S) = 1 \ \Leftrightarrow \ S = \Omega$ – вірогідна подія.

Переваги класичного означення над іншими в тому, що: $P(S)$ описується дуже простим співвідношенням; Ω – скінченна множина.

Недоліки: іноді важко встановити наявність чи відсутність групи несумісних рівноможливих е/п; означення втрачає сенс, коли Ω – нескінченна множина.

Приклад 1.1. З урни, в якій чотири білих і п'ять чорних куль, навмання одна за одною виймають усі кулі і, не дивлячись, відкладають убік. Знайти ймовірність того, що остання куля буде білою.

Розв'язання. Позначимо через A подію, яка полягає в тому, що остання куля біла. Тоді склад попередньо витягнутих куль відомий: три білі та п'ять чорних куль. Отже, загальне число рівноможливих наслідків

$n = C_9^8$ – число способів вибрати вісім куль із дев'яти наявних, а число сприятливих наслідків (за правилом добутку) $m = m_1 \cdot m_2 = C_4^3 \cdot C_5^5$ – число способів вибрати три білі кулі з чотирьох білих ($m_1 = C_4^3$) і п'ять чорних куль з п'яти чорних ($m_2 = C_5^5$). За формулою (1.1) отримаємо:

$$P(A) = \frac{m}{n} = \frac{C_4^3 \cdot C_5^5}{C_9^8} = \frac{4 \cdot 1}{9} = \frac{4}{9}.$$

Відповідь: $P(A) = 4/9 \approx 0{,}44$.

1.2. Геометричне означення ймовірності

У геометрії кожній фігурі як множині точок ставиться у відповідність певне невід'ємне дійсне число μ: $\mu \geq 0$, яке називають **мірою** фігури. Мірою відрізка є його *довжина* l (рис. 1.5а), мірою плоскої фігури – *площа* F (рис. 1.5б), а просторового тіла – *об'єм* V (рис. 1.5в).

а б в

Рис. 1.5. **Фігури: а) на прямій; б) на площині; в) у просторі**

Розглянемо детальніше плоский випадок. Нехай Ω – простір е/п – нескінченна множина точок (рис. 1.6), а S – деяка подія – його підмножина. **Ймовірністю події S у геометричному розумінні** називають відношення міри підмножини $\mu(S)$ до міри простору е/п $\mu(\Omega)$:

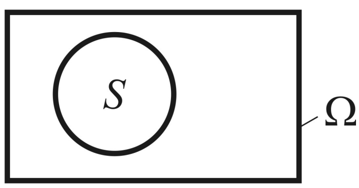

Рис. 1.6. **Простір Ω і подія S**

$$P(S) = \frac{\mu(S)}{\mu(\Omega)}. \qquad (1.2)$$

Аналогічно для одновимірного та тривимірного просторів.

Приклад 1.2. На ділянці площею 30 м^2 розташовано бак із пальним, який займає площу 10 м^2. Яка ймовірність влучання снаряда в бак із пальним під час обстрілу ділянки?

Розв'язання. Позначимо через S подію, яка полягає в тому, що снаряд влучає в бак із пальним, тоді за формулою (1.2) маємо:

$$P(S)=\frac{\mu(S)}{\mu(\Omega)} \Rightarrow \left| \begin{matrix} \mu(S)=10 \\ \mu(\Omega)=30 \end{matrix} \right| \Rightarrow P(S)=\frac{10}{30}=\frac{1}{3}.$$

Зауваження. За допомогою класичного підходу до означення ймовірності цю задачу розв'язати неможливо (*чому?*).

1.3. Теореми додавання та множення ймовірностей

Теорема додавання (про ймовірність суми двох подій)

Нехай подія S є сумою двох подій S_1, S_2: $S=S_1\bigcup S_2$.

Теорема 1.1 (*про ймовірність суми двох подій*). Імовірність суми двох подій дорівнює сумі ймовірностей цих подій без імовірності їх сумісного настання:

$$P(S_1\bigcup S_2)=P(S_1)+P(S_2)-P(S_1\bigcap S_2). \tag{1.3}$$

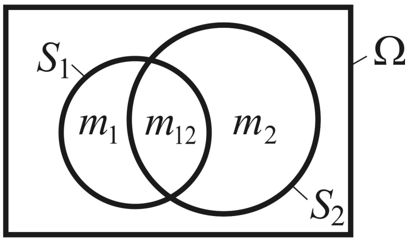

Рис. 1.7. **Ілюстрація до підрахунку числа сприятливих е/п**

Дійсно, якщо m_1 (m_2) – число е/п, що сприяють тільки події S_1 (S_2); m_{12} – число е/п, що сприяють і S_1, і S_2, то отримуємо (рис. 1.7):

$$\underbrace{\frac{m_1+m_{12}+m_2}{n}}_{P(S_1\bigcup S_2)}=\underbrace{\frac{m_1+m_{12}}{n}}_{P(S_1)}+\underbrace{\frac{m_{12}+m_2}{n}}_{P(S_2)}-\underbrace{\frac{m_{12}}{n}}_{P(S_1\bigcap S_2)}.$$

Наслідки з теореми 1.1:

1. Ймовірність суми двох несумісних подій дорівнює сумі ймовірностей цих подій:

$$S_1\bigcap S_2=\varnothing \Rightarrow P(S_1\bigcup S_2)=P(S_1)+P(S_2). \tag{1.4}$$

2. Ймовірність суми взаємно протилежних подій дорівнює одиниці:

$$(S \bigcap \overline{S} = \varnothing) \wedge (S \bigcup \overline{S} = \Omega) \;\Rightarrow\; P(S) + P(\overline{S}) = 1. \tag{1.5}$$

Під час розв'язання задач часто буває простіше підрахувати ймовірність протилежної події $\overline{S}$, а потім – події S, яку треба знайти за умовою задачі: $P(S) = 1 - P(\overline{S})$.

3. Узагальнення (1.4) на довільне скінченне число подій:

$$P(S_1 \bigcup S_2 \bigcup ... \bigcup S_k) = P(S_1) + P(S_2) + ... + P(S_k). \tag{1.6}$$

Формула (1.3) також узагальнюється на довільне скінченне число доданків.

4. Узагальнення (1.3) для трьох ($k = 3$) подій (рис. 1.8):

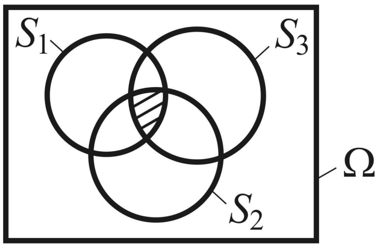

Рис. 1.8. **Ілюстрація до формули ймовірності суми трьох подій**

$$P(S_1 \bigcup S_2 \bigcup S_3) = P(S_1) + P(S_2) + P(S_3) -$$
$$- P(S_1 \bigcap S_2) - P(S_1 \bigcap S_3) - P(S_2 \bigcap S_3) +$$
$$+ P(S_1 \bigcap S_2 \bigcap S_3).$$

Слід зауважити, що спільні е/п, сприятливі для будь-яких двох подій і для всіх трьох подій, ураховуються тільки один раз!

Приклад 1.3. Знайти ймовірність того, що навмання взяте із множини $A = \{1, 2, 3, ..., 30, 31\}$ число виявиться кратним двом або п'яти.

Розв'язання. 1. Вводимо позначення подій, про які йдеться в умові задачі: K_2 $(K_5, K_{2\vee 5})$ =<навмання вибране число, кратне двом (п'яти, двом або п'яти)>.

2. Аналізуємо умову задачі з метою встановлення: простору е/п Ω; сумісності (несумісності) чи залежності (незалежності) розглядуваних подій; зв'язку між подіями K_2, K_5, $K_{2\vee 5}$; теорем (формул), якими слід скористатися.

Простором е/п Ω є множина A як сукупність попарно несумісних (оскільки вибір одного з чисел виключає появу іншої е/п), єдино можливих (оскільки якась із е/п обов'язково повинна настати); рівноможливих наслідків (оскільки числа вибираються навмання). Події K_2 і K_5 сумісні, адже є числа, кратні і двом, і п'яти (наприклад, 10); $K_{2\vee 5}=K_2\bigcup K_5$, отже, треба скористатися теоремою додавання для сумісних подій (див. формулу (1.3)).

3. Записуємо формулу та виконуємо відповідні розрахунки:

$$P(K_2\bigcup K_5)=P(K_2)+P(K_5)-P(K_2\bigcap K_5).$$

Число сприятливих е/п: $m_2=\left[\frac{31}{2}\right]=15$, $m_5=\left[\frac{31}{5}\right]=6$, $m_{2\wedge 5}=\left[\frac{31}{2\cdot 5}\right]=3$, де квадратні дужки – [ ] – символ цілої частини числа (функція антьє).

Отже,

$$P(K_2)=\frac{15}{31};\ P(K_5)=\frac{6}{31};\ P(K_2\bigcap K_5)=\frac{3}{31}\Rightarrow P(K_{2\vee 5})=\frac{18}{31}\approx 0{,}58.$$

Залежні та незалежні події. Теорема множення (про ймовірність добутку двох подій)

Нехай події S_1, S_2 під час проведення E настають з імовірностями $P(S_1)$, $P(S_2)$ відповідно. Часто приходять до задач, де ймовірності обчислюють з урахуванням того, настала чи ні інша подія, і це приводить до розгляду "умовних ймовірностей".

Імовірність події S_1, яка обчислюється за умови, що відбулася подія S_2, називають **умовною ймовірністю** S_1 **відносно** S_2, і пишуть: $P_{S_2}(S_1)$, або $P(S_1|S_2)$.

Імовірності $P(S_1)$, $P(S_2)$, знайдені без урахування того, відбулася чи не відбулася, відповідно, подія S_2, S_1, слід називати **безумовними** або **абсолютними**.

Події S_1, S_2 називають взаємно **залежними (незалежними)**, якщо виконується співвідношення: умовні й абсолютні ймовірності не рівні (рівні) між собою:

$$S_1,\ S_2 \text{ – залежні } \Leftrightarrow \left[\begin{array}{l} P(S_1\mid S_2)\neq P(S_1), \\ P(S_2\mid S_1)\neq P(S_2). \end{array}\right. \tag{1.7}$$

$$S_1,\ S_2 \text{ – незалежні } \Leftrightarrow \left[\begin{array}{l} P(S_1\mid S_2)= P(S_1), \\ P(S_2\mid S_1)= P(S_2). \end{array}\right. \tag{1.8}$$

Теорема 1.2 (*про ймовірність добутку* (*суміщення*) *двох подій*). Ймовірність добутку двох подій дорівнює добутку ймовірності однієї з подій і умовної ймовірності іншої події у припущенні, що перша подія відбулася:

$$P(S_1 \cap S_2) = P(S_1) \cdot P(S_2 \mid S_1) \quad \text{або} \quad P(S_1 \cap S_2) = P(S_2) \cdot P(S_1 \mid S_2). \qquad (1.9)$$

Справедливість (1.9) випливає з арифметичних тотожностей:

$$\underbrace{\frac{m_{12}}{n}}_{P(S_1 \cap S_2)} = \underbrace{\frac{m_1 + m_{12}}{n}}_{P(S_1)} \cdot \underbrace{\frac{m_{12}}{m_1 + m_{12}}}_{P(S_2|S_1)}; \quad \underbrace{\frac{m_{12}}{n}}_{P(S_1 \cap S_2)} = \underbrace{\frac{m_{12} + m_2}{n}}_{P(S_2)} \cdot \underbrace{\frac{m_{12}}{m_{12} + m_2}}_{P(S_1|S_2)} \quad \text{(див. рис. 1.7).}$$

Наслідки з теореми 1.2:

1. Теорема множення для незалежних подій:

$$\left[\begin{array}{l} P(S_1 \mid S_2) = P(S_1) \\ P(S_2 \mid S_1) = P(S_2) \end{array} \right. \Leftrightarrow \quad P(S_1 \cap S_2) = P(S_1) \cdot P(S_2). \qquad (1.10)$$

2. Узагальнення (1.10) на випадок довільного скінченного числа незалежних подій:

$$P(S_1 \cap S_2 \cap ... \cap S_k) = P(S_1) \cdot P(S_2) \cdot ... \cdot P(S_k). \qquad (1.11)$$

3. Узагальнення (1.9) на випадок довільного скінченного числа залежних подій:

$$P(\bigcap_{i=1}^{k} S_i) = P(S_1) \cdot P(S_2 \mid S_1) \cdot P(S_3 \mid S_1 S_2) \cdot ... \cdot P(S_k \mid S_1 S_2 ... S_{k-1}), \qquad (1.12)$$

зокрема ($k = 3$):

$$P(S_1 \cap S_2 \cap S_3) = P(S_1) \cdot P(S_2 \mid S_1) \cdot P(S_3 \mid S_1 S_2).$$

Приклад 1.4. Серед десяти кавунів два зелених. Яка ймовірність того, що вибрані навмання два кавуни стиглі?

Введемо позначення подій: C_i $(i = 1, 2)$ – подія, яка полягає в тому, що i-й кавун стиглий; C =<обидва кавуни стиглі>: $C = C_1 \cap C_2$.

За умовою: $n=10$, $m_c=8$. Тоді за формулою (1.9) маємо:

$$\left(P(C_1)=\frac{m_c}{n}=\frac{8}{10},\ P(C_2\,|\,C_1)=\frac{m_c-1}{n-1}=\frac{7}{9}\right) \quad \Rightarrow$$

$$\Rightarrow \quad P(C)=P(C_1\cap C_2)=\frac{m_c}{n}\cdot\frac{m_c-1}{n-1}=\frac{28}{45}\approx 0{,}62=62\ \%.$$

1.4. Формула повної ймовірності. Формула Байєса

Задача 1.1. Нехай H_1, H_2, H_3, …, H_k – повна група несумісних подій, тобто $H_1\cup H_2\cup H_3\cup\ldots\cup H_k=\Omega$ і $H_i\cap H_j=\varnothing\ \ \forall i\neq j$ (рис. 1.9), з відомими ймовірностями $P(H_i)=p_i$, $i=\overline{1,k}$. Знайти абсолютну ймовірність події S: $P(S)$, якщо відомі її умовні ймовірності відносно подій H_i: $P(S\,|\,H_i)$, $i=\overline{1,k}$.

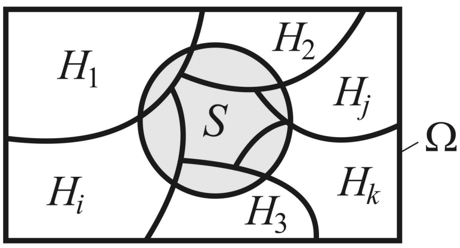

Рис. 1.9. **Повна група несумісних подій**

Розв'язання. Оскільки H_i – несумісні, то S можна подати у вигляді суми перетинів події $S\cdot$з подіями H_i:

$$S=\sum_{i=1}^{k}SH_i.$$

За теоремою додавання для несумісних подій (1.4) і теоремою множення (1.9) отримуємо:

$$P(S)=P(\sum_{i=1}^{k}SH_i)=P(SH_1)+P(SH_2)+\ldots+P(SH_k)=$$

$$=p_1\cdot P(S\,|\,H_1)+p_2\cdot P(S\,|\,H_2)+\ldots+p_k\cdot P(S\,|\,H_k) \quad \Rightarrow$$

$\Rightarrow \quad P(S)=\sum\limits_{i=1}^{k} p_i\cdot P(S|H_i)$ – **формула повної ймовірності**. (1.13)

Зауваження. Події H_i називають *гіпотезами* відносно настання події S тому, що наперед невідомо, які з них відбудуться, а які не відбудуться з настанням S.

Приклад 1.5. Закуплено шкіряні сумки: $60\ \%$ – у Польщі, решту – в Туреччині. Польські сумки мають $5\ \%$ браку, турецькі – $10\ \%$. Знайти ймовірність того, що придбана покупцем сумка буде якісною.

Розв'язання. Введемо позначення подій:

$П\,(Т)$ $=<$ сумку куплено в Польщі (Туреччині) $>$ – гіпотези;

$Б\,(Я)$ $=<$ придбана покупцем сумка бракована (якісна) $>$;

$Я\,|\,П\,(Я\,|\,Т)$ $=<$ якісна сумка закуплена в Польщі (Туреччині) $>$.

Обчислюємо повну ймовірність події $Я$ з урахуванням того, що $P(Я\,|\,П)=1-P(Б\,|\,П)=0{,}95,\ P(Я\,|\,Т)=1-P(Б\,|\,Т)=0{,}90$:

$$P(Я)=P((Я\cap П)\cup(Я\cap Т))=P(Я\cap П)+P(Я\cap Т)=$$
$$=P(П)\cdot P(Я\,|\,П)+P(Т)\cdot P(Я\,|\,Т)=\frac{60}{100}\cdot\frac{95}{100}+\frac{40}{100}\cdot\frac{90}{100}=0{,}93=93\ \%.$$

Слід сподіватися, що зі 100 придбаних покупцями сумок 93 будуть якісними.

Формула Байєса (формула перерахунку ймовірностей гіпотез)

Задача 1.2. Знайти ймовірність гіпотези H_i $(i=\overline{1,k})$ за умови появи події S: $P(H_i\,|S)$, тобто ступінь можливості настання подій H_i у припущенні здійснення події S.

Розв'язання. Як наслідок із теореми множення (1.9) маємо:

$$P(S_1)\cdot P(S_2|S_1)=P(S_2)\cdot P(S_1|S_2)\ – \qquad (1.14)$$

властивість взаємозамінності (у теоремі множення).

У нашому розгляді властивість виглядає так:

$$P(H_i)\cdot P(S|H_i)=P(S)\cdot P(H_i|S),$$

звідки

$$P(H_i\,|\,S)=\frac{P(H_i)\cdot P(S\,|\,H_i)}{P(S)},\ i=\overline{1,k}, \qquad (1.15)$$

– **формула Байєса**,

де $P(S)=\sum\limits_{i=1}^{k}P(H_i)\cdot P(S\,|\,H_i)$ – повна ймовірність події S.

Наслідок. Якщо гіпотези H_i до випробування мають однакові ймовірності, то їхні ймовірності після досліду пропорційні ймовірностям події S за умови настання відповідної гіпотези:

$$P(H_i)=h_i=h-const,\ i=\overline{1,k}\ \Rightarrow\ P(H_i\,|\,S)=C\cdot P(S|H_i),\ C-const.$$

У світлі прикладу 1.5 обчислимо ймовірність того, що придбана якісна сумка виготовлена в Польщі (Туреччині):

$$P(\text{П})\cdot P(\text{Я}\,|\,\text{П})=P(\text{Я})\cdot P(\text{П}\,|\,\text{Я})\ \Rightarrow\ P(\text{П}\,|\,\text{Я})=\frac{P(\text{П})\cdot P(\text{Я}\,|\,\text{П})}{P(\text{Я})}\ \Rightarrow$$

$$\Rightarrow\ P(\text{П}\,|\,\text{Я})=\frac{0{,}6\cdot 0{,}95}{0{,}93}=\frac{19}{31}\approx 0{,}61.$$

$$P(\text{Т})\cdot P(\text{Я}\,|\,\text{Т})=P(\text{Я})\cdot P(\text{Т}\,|\,\text{Я})\ \Rightarrow\ P(\text{Т}\,|\,\text{Я})=\frac{P(\text{Т})\cdot P(\text{Я}\,|\,\text{Т})}{P(\text{Я})}\ \Rightarrow$$

$$\Rightarrow\ P(\text{Т}\,|\,\text{Я})=\frac{0{,}4\cdot 0{,}90}{0{,}93}=\frac{12}{31}\approx 0{,}39.$$

Нескладно помітити, що $P(\text{П}|\text{Я})+P(\text{Т}|\text{Я})=1$ як сума ймовірностей двох взаємно протилежних подій; умовні ймовірності гіпотез мало відрізняються від вихідних: $P(\text{П})=0{,}6$; $P(\text{Т})=0{,}4$, але так буває не завжди.

1.5. Повторні незалежні випробування. Формула Бернуллі, теореми Муавра – Лапласа, формула Пуассона

Незалежні випробування. Схема Бернуллі (СБ). Імовірності $P_n(k)$, $P_n(k_1 \le k \le k_2)$

Розглянемо множину дослідів (випробувань, експериментів), під час проведення яких подія S може настати, а може не настати, тобто розглядається стохастичний експеримент E з простором е/п $\Omega=\{\omega_1,\omega_2\}=$ $=\{S,\overline{S}\}$. Події S і $\overline{S}$ називають, відповідно, "успіхом" і "невдачею": $\Omega=\{$Успіх, Невдача$\}$.

Класичний приклад: E – "підкидання монети": $\Omega=\{S,\overline{S}\}=\{\text{Г},\text{Ц}\}$.

Якщо під час проведення E виконуються умови:

1) усі наслідки випробувань взаємно незалежні, тобто результат одного випробування не впливає на результат іншого;

2) успіх настає з однією і тією ж імовірністю: $P(S)=p$ $(0<p<1)$, то кажуть, що **випробування здійснюється за схемою незалежних випробувань, або схемою Бернуллі**.

Зрозуміло, що "невдача" $\overline{S}$ настає з імовірністю $P(\overline{S})=1-p$, яку позначимо через q: $P(\overline{S})=q$; отже $p+q=1$ як сума ймовірностей взаємно протилежних подій.

Задача 1.3. Проводиться серія із n випробувань за СБ. Знайти ймовірність $P_n(k)$ події, яка полягає в тому, що успіх S у n випробуваннях настає k разів ($0 \le k \le n$), тоді невдача настане ($n-k$) разів.

Розв'язання.

1. Окремий випадок: $n=2$:

Наслідок	SS	$S\overline{S}$	$\overline{S}S$	$\overline{S}\overline{S}$	– множина єдино можливих подій Ω_2
Ймовірність	$p \cdot p$	$p \cdot q$	$q \cdot p$	$q \cdot q$	$\Rightarrow$ $p^2+2pq+q^2=(p+q)^2=1.$

2. Загальний випадок: результат проведення n експериментів можна описати подією $\tilde{S}_1 \cap \tilde{S}_2 \cap \tilde{S}_3 \cap \ldots \cap \tilde{S}_n$ або $\tilde{S}_1\tilde{S}_2\tilde{S}_3 \ldots \tilde{S}_n$, де $\tilde{S}_i$ – це або S, або $\overline{S}$ $(i=\overline{1,k})$, і S повторюється k разів. Для крайніх випадків маємо:

$$P(SSS\ldots S)=P_n(n)=ppp\ldots p=p^n; \quad P(\overline{S}\,\overline{S}\,\overline{S}\ldots\overline{S})=P_n(0)=qqq\ldots q=q^n.$$

В інших випадках результат n випробувань $\tilde{S}_1\tilde{S}_2\tilde{S}_3 \ldots \tilde{S}_n$ буде містити успіх S k разів:

$$\underbrace{SSS\ldots S}_{k}\underbrace{\overline{S}\,\overline{S}\,\overline{S}\ldots\overline{S}}_{n-k} \quad \text{або} \quad \underbrace{SSS\ldots S}_{k-1}\underbrace{\overline{S}\,\overline{S}\,\overline{S}\ldots\overline{S}}_{n-k-2}\,S\,\overline{S}\,\overline{S} \quad \text{і таке інше.}$$

Отже, для всіх таких наслідків успіх наставатиме з імовірністю p^k, а $\tilde{S}_1\tilde{S}_2\tilde{S}_3 \ldots \tilde{S}_n$ – з імовірністю $p^k q^{n-k}$. Число ж усіх подій, у яких успіх зустрічатиметься k разів, визначається числом k-елементних підмножин n-елементної множини, тобто числом комбінацій із n елементів по k:

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

Таким чином,

$$P_n(k) = C_n^k \cdot p^k q^{n-k}$$ – **формула Бернуллі**. (1.16)

Як і для $n=2$, маємо: $\sum\limits_{i=0}^{n} P_n(i) = P_n(0) + P_n(1) + \ldots + P_n(n) = 1$.

Приклад 1.6. E – монета правильної геометричної форми підкидається над горизонтальною плоскою поверхнею п'ять разів. Підрахувати ймовірність того, що "успіх" (випадання цифри) настане 0, 1, 2, 3, 4, 5 разів.

Розв'язання. Оскільки випробування незалежні між собою й успіх настає з однією і тією ж імовірністю $p = 1/2$ $(q = 1/2)$, то E відбувається за СБ і відповідні ймовірності можна підрахувати за формулою (1.16).

Для підрахунку C_n^k скористаємось властивістю: $C_n^k = C_n^{n-k}$. Дійсно: $\dfrac{n!}{k!(n-k)!} = \dfrac{n!}{(n-k)!\,k!}$. *Наприклад*, $C_5^2 = \dfrac{5\cdot 4}{1\cdot 2} = \dfrac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3} = C_5^3$.

$$P_5(0) = P_5(5) = C_5^0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32};$$

$$P_5(1) = P_5(4) = C_5^1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{32};$$

$$P_5(2) = P_5(3) = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{32}.$$

На рис. 1.10 зображено так званий **полігон** (**многокутник**) розподілу ймовірностей – ламана з вершинами в точках $(k,\ P_n(k))$.

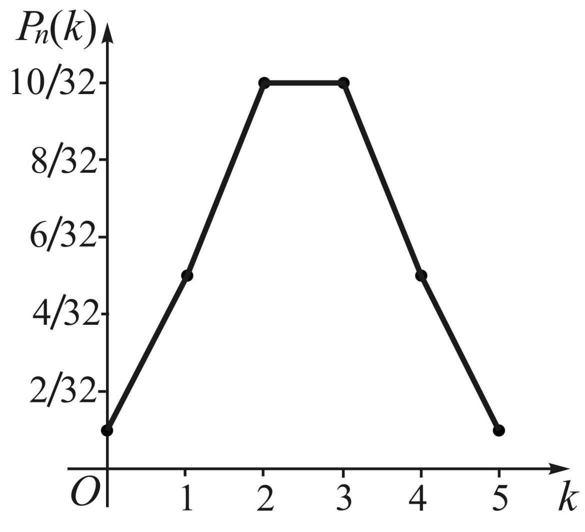

Рис. 1.10. **Полігон розподілу**

Задача 1.4. Здійснено n випробувань за СБ. Обчислити ймовірність того, що успіх настане не менше k_1 разів і не більше k_2 разів: $P_n(k_1 \le k \le k_2)$, якщо $P(S) = p$, $P(\overline{S}) = q$.

Розв'язання. Подія $(k_1 \le k \le k_2)$, або пишуть (k_1, k_2), є сумою подій, які полягають у тому, що успіх настає або k_1, або $k_1 + 1$, або $k_1 + 2$, ..., або k_2 рази. Ці події несумісні, а тому за формулою (1.6) отримаємо:

$$P_n(k_1,k_2) = P_n(k_1) + P_n(k_1+1) + \ldots + P_n(k_2) \;\Rightarrow$$

$$\Rightarrow\; P_n(k_1,k_2) = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i}. \qquad (1.17)$$

Зауваження: у випадку великих n використання формул (1.16), (1.17) призводить до "важкої" арифметики навіть для ЕОМ, а тому знайдені наближені формули для обчислення ймовірностей.

Найімовірніше число успіхів у СБ

Число успіхів k_0, яке настає з найбільшою ймовірністю, тобто $P_n(k_0) = \max\limits_{0 \le k \le n} \{P_n(k)\}$, називають **найімовірнішим числом успіхів**.

Вказане число визначається подвійною нерівністю:

$$np - q \le k_0 \le np + p. \qquad (1.18)$$

Цю нерівність задовольняє тільки одне число (ціла частина числа $np + p$), за винятком випадку, коли $np + p$ – ціле число.

У прикладі 1.6 маємо два найімовірніших числа успіхів: 2 і 3, оскільки $np + p = 5 \cdot 0{,}5 + 0{,}5 = 3$.

Локальна теорема (формула) Муавра – Лапласа

Наближені формули $\widetilde{P}_n(k)$, $\widetilde{P}_n(k_1,k_2)$ для обчислення ймовірностей $P_n(k)$, $P_n(k_1,k_2)$ виведені за умови, що границя відношення наближеного та точного значень ймовірності дорівнює одиниці, якщо число випробувань необмежено зростає ($n \to \infty$):

$$\lim_{n\to\infty}\frac{\widetilde{P}_n(k)}{P_n(k_2)}=1, \quad \lim_{n\to\infty}\frac{\widetilde{P}_n(k_1,k_2)}{P_n(k_1,k_2)}=1.$$

Збільшуючи n, похибку можна зробити як завгодно малою.

Наближені формули називають **асимптотичними формулами** для схеми Бернуллі.

Теорема 1.3 (*формула Муавра – Лапласа локальна, тобто для одного* k). Асимптотичною формулою для ймовірності $P_n(k)$ є формула:

$$\widetilde{P}_n(k)=\frac{1}{\sqrt{npq}}\cdot\varphi(x), \text{ де } \varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-x^2/2}, \; x=\frac{k-np}{\sqrt{npq}}. \tag{1.19}$$

Для функції $\varphi(x)$ – **функції Гаусса** – складено спеціальні таблиці (табл. А.1 додатка А) з урахуванням її властивостей (рис. 1.11):

1) функція $\varphi(x)$ парна: $\varphi(-x)=\varphi(x)$, а тому в таблицях значення її для від'ємних x не наводяться;

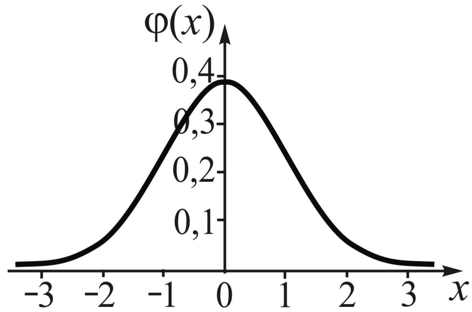

Рис. 1.11. **Функція Гаусса**

2) функція $\varphi(x)$ є спадною для $x>0$ і $\varphi(x)\to 0$, якщо $x\to\infty$. Оскільки $\varphi(5)\approx 0{,}15\cdot 10^{-5}$, то для значень $x>5$ приймають $\varphi(x)=0$.

Інтегральна (тотальна) формула Муавра – Лапласа

Теорема 1.4 (*формула Муавра – Лапласа тотальна, тобто для багатьох* k). Асимптотичною формулою для ймовірності $P_n(k_1,k_2)$ є формула:

$$\widetilde{P}_n(k_1,k_2)=\int_{t_1}^{t_2}\varphi(t)dt=\frac{1}{\sqrt{2\pi}}\int_{t_1}^{t_2}e^{-\frac{t^2}{2}}dt, \text{ де } t_i=\frac{k_i-np}{\sqrt{npq}}, \; i=1{,}2. \tag{1.20}$$

Зауваження. Локальна й інтегральна формули є досить точними, якщо величина npq порядку 10^5, а в задачах, які не потребують високої точності, – і за менших значень (але не менших за 20).

Обчислення визначених інтегралів в (1.20) здійснюється за допомогою так званої **функції Лапласа** (рис. 1.12):

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

для якої складені таблиці (табл. Б.1 додатка Б) з урахуванням її властивостей:

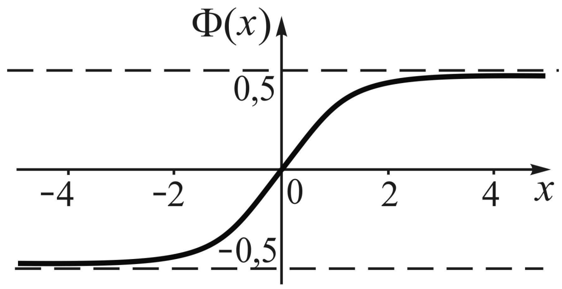

Рис. 1.12. **Функція Лапласа**

1) $\Phi(x)$ є непарною функцією: $\Phi(-x)=-\Phi(x)$, тому її значення для від'ємних x у таблицях не наводяться;

2) якщо $x\to+\infty$, то $\Phi(x)\to 1/2$ і $\Phi(5)\approx 0{,}49999997\approx 1/2$, тому для значень $x>5$ приймають $\Phi(x)=1/2$.

Інтегральна формула Муавра – Лапласа із залученням функції Лапласа $\Phi(x)$ і властивостей визначеного інтеграла набуває вигляду:

$$\tilde{P}_n(k_1,k_2)=\Phi(t_2)-\Phi(t_1). \tag{1.21}$$

Дійсно, $\tilde{P}_n(k_1,k_2)=\frac{1}{\sqrt{2\pi}}\int_0^{t_2} e^{-\frac{t^2}{2}}dt-\frac{1}{\sqrt{2\pi}}\int_0^{t_1} e^{-\frac{t^2}{2}}dt$, оскільки $\int_{t_1}^{t_2}=\int_0^{t_2}-\int_0^{t_1}$.

Приклад 1.7. Імовірність того, що деталь виготовлена з порушенням стандартів, дорівнює $0{,}2$. Знайти ймовірність того, що серед 400 випадково відібраних деталей виявиться: 1) 70 нестандартних; 2) нестандартних деталей від 70 до 100.

Розв'язання: 1. За умовою $p=0{,}2$, тоді $q=0{,}8$; $n=400$, $k=70$; $npq=64>20$. Використовуємо формулу (1.19), для чого підраховуємо x, знаходимо за таблицею значення $\varphi(x)$ і обчислюємо $\tilde{P}_n(k)$:

$$x=\frac{k-np}{\sqrt{npq}}=\frac{70-400\cdot 0{,}2}{\sqrt{400\cdot 0{,}2\cdot 0{,}8}}=-1{,}25 \;\Rightarrow\; \tilde{P}_{400}(70)=1/8\cdot 0{,}1826\approx 0{,}0228.$$

2. Використовуємо формулу (1.20), для чого: підраховуємо t_1 і t_2; за таблицею знаходимо значення $\Phi(t_1)$, $\Phi(t_2)$ і підставляємо у (1.21):

$$t_1 = -1{,}25 \quad \Rightarrow \quad \Phi(1{,}25) = 0{,}3944;$$

$$t_2 = \frac{100 - 400 \cdot 0{,}2}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} = 2{,}5 \quad \Rightarrow \quad \Phi(2{,}5) = 0{,}4938;$$

$$\tilde{P}_{400}(70 \le k \le 100) = \Phi(2{,}5) - \Phi(-1{,}25) = \Phi(2{,}5) + \Phi(1{,}25) = 0{,}8882.$$

Наслідки з теореми 1.4:

1. Формула ймовірності події $<$модуль різниці між числом успіхів k і величиною np не перевищує заданого $\varepsilon>$:

$$\tilde{P}_n(|k - np| \le \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sqrt{npq}}\right).$$

2. Формула ймовірності події $<$модуль різниці між відносним числом успіхів k/n і ймовірністю p не перевищує заданого $\varepsilon>$:

$$\tilde{P}_n\left(\left|\frac{k}{n} - p\right| \le \varepsilon\right) = 2\Phi\left(\frac{\varepsilon n}{\sqrt{npq}}\right) = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Формула Пуассона

Події в ТЙ, імовірність настання яких мала: $p < 0{,}1$, називають **рідкісними,** або **малоймовірними**. Виявляється, що формули Муавра – Лапласа у випадку великих n ($n>100$), але невеликих $\lambda = np$ ($\lambda<10$) і малих p дають значні похибки.

Теорема 1.5 (*формула Пуассона для малоймовірних подій*). Асимптотичною формулою для ймовірності $P_n(k)$ рідкісних подій є формула:

$$\tilde{P}_n(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \text{ де } \lambda = np - const, \tag{1.22}$$

при цьому похибка описується нерівністю:

$$\left|\tilde{P}_n(n) - P_n(k)\right| \le 2\lambda^2/n.$$

Для обчислення $\tilde{P}_n(k)$ за (1.22) складені спеціальні таблиці з двома входами для λ і k (табл. В.1 додатка В).

Приклад 1.8. Завод відправив на базу 500 виробів. Імовірність пошкодження виробу в дорозі $0{,}004$. Знайти ймовірність того, що в дорозі пошкоджено: а) один виріб; б) менше трьох виробів.

Розв'язання. За умовою дано: $n=500$, $p=0{,}004$, $\lambda=np=2<10$. Отже, можна використати формулу Пуассона (1.22):

а) $\tilde{P}_{500}(1)=\dfrac{2^1}{1!}\cdot e^{-2}=\dfrac{2}{e^2}\approx 0{,}2707$ (співпадає зі значенням у таблиці);

б) $\tilde{P}_{500}(0\le k\le 2)=\tilde{P}_{500}(0)+\tilde{P}_{500}(1)+\tilde{P}_{500}(2)=e^{-2}+\dfrac{2}{1!}e^{-2}+\dfrac{4}{2!}e^{-2}=$

$$=e^{-2}(1+2+2)=5/e^2\approx 0{,}6767.$$

Цей приклад показує, що в деяких випадках можна обійтись без таблиць.

Зауважимо, що на практиці часто зустрічаються явища, у яких імовірність $P_1(\Delta t)$ настання одного успіху прямо пропорційна відрізку часу Δt з точністю до нескінченно малої більш високого порядку, ніж Δt, тобто $P_1(\Delta t)=\lambda\,\Delta t+o\,(\Delta t)$. Тоді ймовірність настання k успіхів за проміжок часу t підраховується за формулою, аналогічною до формули Пуассона:

$$P_k(t)=\frac{(\lambda t)^k}{k!}\cdot e^{-\lambda t}, \qquad (1.23)$$

де λ – середнє число успіхів за одиницю часу;

λt – середнє число успіхів за час t.

Приклад 1.9. Під час роботи приладу виникають випадкові неполадки (збої), середнє число яких за годину дорівнює 2. Знайти ймовірність того, що за три години настане п'ять збоїв.

Розв'язання. За умовою маємо: $\lambda=2$, $t=3$, $k=5$. Отже, згідно з (1.23) отримуємо:

$$P_5(3)=\frac{(2\cdot 3)^5}{5!}\cdot e^{-2\cdot 3}=\frac{6^5}{120!}\cdot e^{-6}\approx 0{,}1606.$$

Таке саме значення ймовірності одержуємо за таблицею значень розподілу Пуассона (табл. В.1 додатка В) для $\lambda=6$ і $k=5$.

1.6. Дискретні випадкові величини (ДВВ): закон і функція розподілу, числові характеристики

Означення ДВВ, закон розподілу, способи його задання

Числову множину $A=\{a_1,a_2,..,a_n\}$, скінченну або нескінченну, називають **дискретною,** якщо для кожного елемента a_i, $i=\overline{1,n}$ (∞), існує такий інтервал (a,b), який не містить жодної точки із A, крім a_i.

Числову функцію $\xi=\xi(\omega)$, де $\omega\in\Omega$, називають **дискретною ВВ,** якщо простір е/п Ω є скінченною або нескінченною дискретною множиною (не обов'язково числовою). Інакше, $\xi=\xi(\omega)$ – це функція, яка кожній е/п ω із Ω ставить у відповідність деяке число.

Приклад 1.10. E_1 – спостереження за світлофором із простором наслідків $\Omega=\{\omega_1,\omega_2,\omega_3\}=\{ч,ж,з\}$. Поставимо у відповідність кожній е/п *ч*, *ж*, *з* якесь число, наприклад, 1, 2, 3. Тоді одержимо ДВВ ξ з областю визначення $D(\xi)=\{ч,ж,з\}$ і областю значень $E(\xi)=\{1,2,3\}$. Замість того щоб казати, що світлофор горить червоним (жовтим, зеленим) кольором, кажуть, що ВВ набуває значення 1 (2, 3).

Відповідність між значеннями, яких набуває ДВВ: $\xi=x_1,x_2,...,x_n(...)$, і ймовірностями p_i, $i=\overline{1,n}$ (∞), називають **законом розподілу ймовірностей,** або **розподілом** ДВВ ξ.

Способи задання ДВВ ξ як числових функцій $\xi=\xi(\omega)$ нічим не відрізняються від способів задання функцій у математичному аналізі.

1. *Табличний спосіб* – подання закону розподілу парами (x_i,p_i), $i=\overline{1,n}$ (∞), у вигляді таблиці з двох рядків (або стовпців):

ξ	x_1	x_2	x_3	$\dots$	x_n	– значення дискретної випадкової величини,
p	p_1	p_2	p_3	$\cdots$	p_n	– ймовірності, з якими вони приймаються.

Найпростішою ДВВ є так званий **індикатор події** S, який набуває тільки двох значень: $\chi_s(\omega)=\begin{cases}0, \omega\in S\\ 1, \omega\notin S\end{cases}$, із законом розподілу:

χ_s	0	1
P	q	p

.

Зауважимо, що $\sum_{i=1}^{n(\infty)} p_i = p_1 + p_2 + \ldots + p_n + \ldots = 1$ (*чому?*).

2. *Графічний спосіб* – зображення закону розподілу у вигляді ламаної з вершинами в точках (x_i, p_i), $i = \overline{1, n}$ (∞), яку називають **многокутником,** або **полігоном** розподілу ймовірностей ДВВ (рис. 1.13).

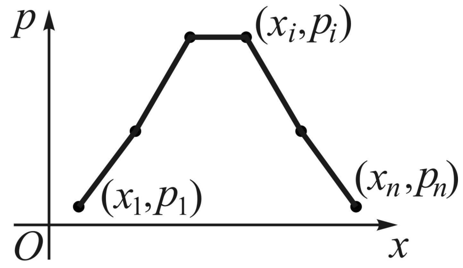

Рис. 1.13. **Полігон розподілу**

3. *Аналітичний спосіб* – подання закону розподілу формулою: $P(\xi = x_i) = p_i$, яка описує ймовірність події $(\xi = x_i)$, тобто ймовірність того, що ДВВ ξ набуде значення x_i, $i = \overline{1, n}$ (∞).

Приклади аналітичного задання важливих розподілів ДВВ.

1. Біноміальний розподіл – розподіл ВВ, значеннями якої є число успіхів k у n випробуваннях за схемою Бернуллі:

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

Такий розподіл називають також *розподілом Бернуллі*.

2. Розподіл Пуассона – розподіл ВВ у випробуваннях за схемою Бернуллі, значеннями якої є число успіхів k у рідкісних подіях:

$$P(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \text{ де } k = 0, 1, 2, \ldots, n; \ \lambda = np.$$

3. Геометричний розподіл – розподіл ВВ, значеннями якої є число випробувань до настання першого успіху:

$$P(\xi = k) = q^{k-1} p, \text{ де } p + q = 1, \ \ k = 1, 2, \ldots.$$

4. Гіпергеометричний розподіл – розподіл ВВ, значеннями якої є число об'єктів m, що мають певну якість серед n вибраних із загального числа об'єктів N, серед яких M мають цю якість:

$$P(\xi = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}, \quad m = 0, 1, 2, \ldots, \min\{n, M\}.$$

Закон розподілу ВВ дозволяє обчислювати ймовірність не тільки події $(\xi = x_i)$, а й більш складних подій, а саме:

$(\xi < x) = (-\infty < \xi < x)$, де $(\xi < x)$ – подія, яка полягає в тому, що ВВ ξ набуває значень, менших за деяке $x \in R$;

$(a \le \xi \le b)$ – подія, яка полягає в тому, що ВВ ξ набуває значень, не менших за a, і не більших за b, і варіації цієї події зі строгими знаками нерівності.

Якщо відомий закон розподілу ВВ, то ймовірність зазначених подій обчислюють за теоремою про ймовірність суми несумісних подій, оскільки одне й те ж значення x_i не може набуватися одночасно з якимось значенням x_j:

$P(\xi < x) = \sum\limits_{\forall i:(x_i < x)} p_i = p_1 + p_2 + ... + p_m$, m – число значень ВВ ξ $x_i < x$;

$$P(a \le \xi \le b) = \sum_{\forall i:(a \le x_i \le b)} p_i . \tag{1.24}$$

Виявляється, що ймовірність будь-якої події можна знайти, відштовхуючись від знання ймовірностей подій вигляду $(\xi < x)$.

Функція розподілу ДВВ: означення, властивості, застосування

Крім закону розподілу ймовірностей: $P(\xi = x_i) = p_i$, для опису ДВВ $\xi = \xi(\omega)$ вводять у розгляд так звану функцію розподілу $F_\xi(x)$, або $F(x)$.

Функцією розподілу ДВВ ξ називають числову функцію $F_\xi(x)$, областю існування якої є вся числова вісь: $(-\infty < x < +\infty)$, а значеннями – ймовірність події $(\xi < x)$, або $(-\infty < \xi < x)$:

$$F_\xi(x) = P(\xi < x) \quad \forall x \in \mathbf{R}. \tag{1.25}$$

Геометрично це означає, що $F_\xi(x)$ визначає ймовірність того, що значення ДВВ ξ розташовані на числовій осі лівіше від точки x (*зобразіть*).

Із (1.25) випливають такі *властивості* $F(x)$:

1) $0 \le F(x) \le 1$, оскільки значеннями $F(x)$ є ймовірності;

2) $F(-\infty)=0$, оскільки подія $(\xi<-\infty)$ неможлива; $F(+\infty)=1$, оскільки подія $(\xi < +\infty)$ вірогідна;

3) $F(x)$ є неспадною функцією: $x_2 > x_1 \Rightarrow F(x_2) \ge F(x_1)$, адже подія $(\xi < x_2)$ тягне за собою подію $(\xi < x_1)$: $(\xi < x_1) \subset (\xi < x_2)$.

Задача 1.5. Обчислити ймовірність події $(x_1 \le \xi < x_2)$, якщо відома функція розподілу $F_\xi(x)$.

Розв'язання. Подію $(\xi < x_2)$ можна подати у вигляді суми двох несумісних подій: $(\xi < x_2) = (\xi < x_1) \bigcup (x_1 \le \xi < x_2)$. За теоремою додавання: $P(\xi < x_2) = P(\xi < x_1) + P(x_1 \le \xi < x_2)$. Звідки

$$P(x_1 \le \xi < x_2) = F(x_2) - F(x_1). \qquad (1.26)$$

Ймовірність того, що ξ набуває значень, не менших за x_1 і менших за x_2, дорівнює різниці значень функції розподілу в точках x_2 та x_1.

Наслідки з (1.26) (*наведіть коментар*):

1) $P(x_1 < \xi < x_2) = F(x_2) - F(x_1) - P(\xi = x_1)$;

2) $P(x_1 \le \xi \le x_2) = F(x_2) - F(x_1) + P(\xi = x_2)$;

3) $P(x_1 < \xi \le x_2) = F(x_2) - F(x_1) + P(\xi = x_2) - P(\xi = x_1)$;

4) $P(\xi \ge x) = 1 - F(x)$, бо $(\xi < x) \bigcup (\xi \ge x) = R = (-\infty, +\infty)$.

Приклад 1.11. Описати в символах функцію розподілу індикатора $\chi_s(\omega)$ події S і побудувати її графік.

Розв'язання. Закон розподілу індикатора має вигляд:

χ_s	0	1
P	q	p

Тоді:

$$F(x) = \begin{cases} 0, & x \le 0, \\ q, & 0 < x \le 1, \\ q + p = 1, & x > 1. \end{cases}$$

За формулою для $F(x)$ будуємо її графік (рис. 1.14).

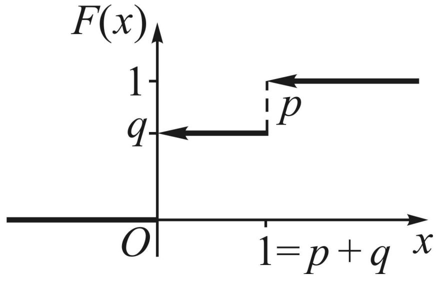

Рис. 1.14. **Графік $F(x)$**

Для $x=0$ функція $F(0)=P(\chi_s<0)=0$, бо χ_s не має значень, менших за 0; тоді і $\forall x\in(-\infty,0)$: $F(x)=P(\chi_s<x)=0$. У точці $x=0$ функція $F(x)$ робить стрибок величини q і залишається такою $\forall x\in(0,1]$. Аналогічно: у точці $x=1$ функція $F(x)$ здійснює стрибок величини p і $\forall x>1$: $F(x)=P(\chi_s<x)=1$, оскільки всі значення індикатора не перевищують одиниці. Стрілочки на графіку вказують на неперервність функції $F(x)$ у точках $x=0$ і $x=1$ зліва: $F(-0)=F(0)$ і $F(1-0)=F(1)$.

Висновок. Графіком функції розподілу ДВВ є східчаста лінія, висота кожної сходинки якої дорівнює ймовірності, з якою випадкова величина набуває відповідного значення; якщо x – точка неперервності, то $P(\xi=x)=0$.

Приклад 1.12. За заданим законом розподілу ДВВ ξ:

ξ	–2	0	2	4
p	0,2	0,1	0,3	0,4

,

описати в символах функцію розподілу $F(x)$ і побудувати її графік.

Розв'язання. Визначаємо проміжки, на які значення ДВВ поділяють числову вісь і на яких $F(x)$ набуває сталих значень (див. стовпець нерівностей у (1.27)). Накопичуємо ймовірності для отримання значень $F(x)$:

$$F(x)=\begin{cases}0, & x\le -2,\\ 0+0{,}2=0{,}2, & -2<x\le 0,\\ 0{,}2+0{,}1=0{,}3, & 0<x\le 2,\\ 0{,}3+0{,}3=0{,}6, & 2<x\le 4,\\ 0{,}6+0{,}4=1, & x>4.\end{cases} \tag{1.27}$$

Будуємо східчасту лінію, висота сходинок якої чисельно дорівнює ймовірності $P(\xi=x)$ набуття випадковою величиною її значень за заданим законом розподілу: $x=-2,\ 0,\ 2,\ 4$ (рис. 1.15).

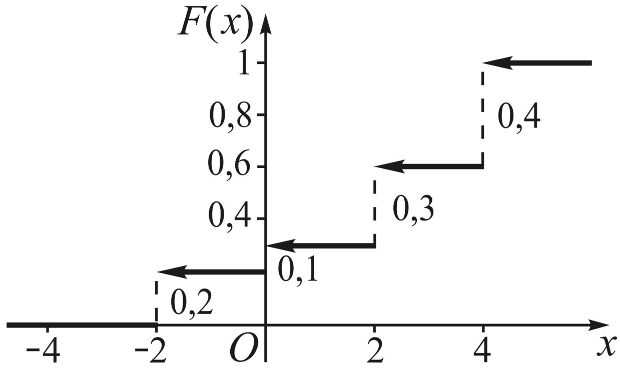

Рис. 1.15. **Графік функції розподілу**

Числові характеристики ДВВ

1. Математичне сподівання (середнє значення, центр розподілу). Дискретна ВВ повністю описується законом розподілу $P(\xi = x) = p$ і функцією розподілу $F(x) = P(\xi < x)$. Але в багатьох випадках виявляється достатнім вказати не зазначені функціональні характеристики ВВ, а числові характеристики (ч/х), які є сталими.

Нехай ВВ ξ є дискретною ВВ зі скінченною або нескінченною множиною значень: $\xi = x_1,\ x_2,\ \ldots,\ x_n$, (…), які набуваються, відповідно, з імовірностями: $p = p_1,\ p_2,\ \ldots,\ p_n$ (…), тобто задано закон розподілу ВВ.

Математичним сподіванням ДВВ ξ називають суму добутків значень ВВ і ймовірностей, з якими вони набуваються:

$$M\xi = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n + \ldots = \sum_{i=1}^{n\,(\infty)} x_i p_i\,, \qquad (1.28)$$

де за умови нескінченної множини значень ВВ числовий ряд збігається.

Приклад 1.13. Задано закон розподілу ДВВ ξ:

ξ	1	2	3
p	0,2	0,3	0,5

Знайти середнє значення ξ і порівняти його із середнім арифметичним значень ВВ.

Розв'язання. За формулою (1.28) маємо:

$$M\xi = \sum_{i=1}^{3} x_i p_i = x_1 p_1 + x_2 p_2 + x_3 p_3 = 1\cdot 0,2 + 2\cdot 0,3 + 3\cdot 0,5 = 2,3.$$

Підраховуємо середнє арифметичне: $(1+2+3)/3 = 2$.

Висновки:

1) $M\xi$ не обов'язково повинно дорівнювати якомусь значенню ВВ;

2) середнє арифметичне – це частинний випадок $M\xi$, коли всі значення ВВ набуваються з однією і тією ж імовірністю $p_i = 1/n$ $(i = \overline{1,3})$: $1\cdot 1/3 + 2\cdot 1/3 + 3\cdot 1/3 = 2$;

3) якщо $D(\xi)$ – нескінченна множина, то приходимо до підрахунку суми нескінченного числа доданків, тобто суми числового ряду.

Різницю між ВВ ξ і її математичним сподіванням називають **центрованою ВВ** і позначають через $\overset{\circ}{\xi}$: $\overset{\circ}{\xi} = \xi - M\xi$.

В умовах прикладу 1.13 знайдемо закон і центр розподілу ВВ $\overset{\circ}{\xi}$:

$$\begin{array}{c|c|c|c} \overset{\circ}{\xi} & -1,3 & -0,3 & 0,7 \\ \hline p & 0,2 & 0,3 & 0,5 \end{array} \Rightarrow \quad M\overset{\circ}{\xi} = -1,3\cdot 0,2 - 0,3\cdot 0,3 + 0,7\cdot 0,5 = 0.$$

Висновок. Середнє значення центрованої ВВ дорівнює нулеві.

Основні властивості $M\xi$:

1. Математичне сподівання виродженої ВВ ξ ($\xi = C - const$) дорівнює C:

$$\xi = C \quad \Rightarrow \begin{array}{c|c} \xi & C \\ \hline p & 1 \end{array} \quad \Rightarrow \; M\xi = C.$$

2. Сталий множник можна виносити за знак (символ) математичного сподівання:

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

3. Математичне сподівання суми (різниці) двох ВВ дорівнює сумі (різниці) математичних сподівань операндів:

$$M(\xi \pm \eta) = M\xi \pm M\eta.$$

Зокрема,

$$M(\xi \pm C) = M\xi \pm C; \quad M(\xi - M\xi) = 0.$$

4. Якщо ВВ ξ, η – незалежні, то середнє значення добутку ВВ дорівнює добутку середніх значень співмножників:

$$M(\xi\cdot\eta) = M\xi\cdot M\eta.$$

Зокрема,

$$M(\xi\cdot C) = C\cdot M\xi; \quad M(\xi^2) = (M\xi)^2.$$

Зауваження. Математичні сподівання важливих розподілів наведені в додатку Г.

2. Дисперсія (розсіяння). Дисперсією ДВВ ξ називають математичне сподівання квадрата відповідної центрованої ВВ:

$$D\xi = M(\overset{\circ}{\xi})^2 = M(\xi - M\xi)^2. \tag{1.29}$$

Інакше, $D\xi$ – математичне сподівання квадрата відхилення ВВ ξ від її математичного сподівання. За відомим законом розподілу як множини пар (x_i, p_i) розрахункова формула має вигляд:

$$D\xi = \sum_{i=1}^{n(\infty)} (x_i - M\xi)^2 \cdot p_i. \tag{1.30}$$

Основні властивості $D\xi$:

1. Дисперсія ВВ дорівнює математичному сподіванню квадрата ВВ без квадрата її математичного сподівання:

$$D\xi = M\xi^2 - M^2\xi \text{ – друга розрахункова формула.} \tag{1.31}$$

2. $\xi = C \quad \Rightarrow \quad D\xi = 0$.

3. $D(C\cdot\xi) = C^2 D\xi$.

4. ξ, η – незалежні $\quad \Rightarrow \quad D(\xi+\eta) = D\xi + D\eta$.

Зокрема, $D(\xi + C) = D\xi$; $\quad D(\xi + C\cdot\xi) = (C+1)^2 D\xi$.

5. ξ, η – незалежні $\quad \Rightarrow \quad D(\xi\cdot\eta) \ge D\xi \cdot D\eta$.

(*Сформулюйте словесно* властивості 2 – 5.)

Арифметичний корінь із дисперсії ВВ називають **середнім квадратичним відхиленням** σ_ξ:

$$\sigma_\xi = \sqrt{D\xi}.$$

Спираючись на σ_ξ, часто пишуть: $D\xi = \sigma_\xi^2$, і кажуть, що дисперсія є квадратом середнього квадратичного відхилення.

Підсумок:

1) дисперсія є числом, яке характеризує ступінь відхилення значень ВВ від її середнього значення;

2) σ_ξ, як і дисперсія, є характеристикою розсіяння значень ВВ навколо центра розподілу, або відхилення значень ВВ від $M\xi$.

Приклад 1.14. В умовах прикладу 1.13 знайти $D\xi$ двома способами.

Розв'язання. Застосовуємо (1.30), ураховуючи, що $M\xi = 2{,}3$:

$$D\xi = \sum_{i=1}^{3}(x_i - M\xi)^2 \cdot p_i = (1-2{,}3)^2 \cdot 0{,}2 + (2-2{,}3)^2 \cdot 0{,}3 + (3-2{,}3)^2 \cdot 0{,}5 = 0{,}61.$$

Для використання (1.31) підраховуємо спочатку $M\xi^2$:

$$M\xi^2 = \sum_{i=1}^{3} x_i^2 p_i = 1^2 \cdot 0{,}2 + 2^2 \cdot 0{,}3 + 3^2 \cdot 0{,}5 = 5{,}9,$$

а потім і дисперсію:

$$D\xi = M\xi^2 - M^2\xi = 5{,}9 - (2,3)^2 = 0{,}61, \text{ звідки } \sigma_\xi \approx 0{,}78.$$

Якій із двох формул, на вашу думку, слід віддати *перевагу*?

1.7. Неперервні випадкові величини (НВВ): функція та щільність розподілу, числові характеристики

НВВ: основні означення, функція та щільність розподілу

НВВ вводяться в розгляд у випадках, коли множина значень ВВ не є дискретною, тобто суцільно займає своїми значеннями деякий проміжок числової осі. Це відповідає тому, що простір е/п Ω є нескінченною множиною і до того ж неперервною. Поняття функції розподілу $F_\xi(x)$ як ймовірності настання події ($\xi < x$) залишається в силі і для неперервних ВВ.

Випадкову величину ξ називають **неперервною**, якщо її функція розподілу визначається формулою:

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(t)dt, \tag{1.32}$$

де функція $f_\xi(x)$ – невід'ємна та визначена на всій числовій осі, а невластивий інтеграл від неї на всій осі дорівнює одиниці.

Функцію $f_\xi(x)$, за допомогою якої визначається $F_\xi(x)$, називають **щільністю розподілу (ймовірностей) НВВ** ξ:

$$f_\xi(x) \text{ – щільність розподілу НВВ } \xi \rightleftharpoons \left[\begin{array}{l} f_\xi(x) \ge 0,\ D(f_\xi) = \mathbf{R}; \\ \int\limits_{-\infty}^{+\infty} f_\xi(x)dx = 1. \end{array}\right. \qquad (1.33)$$

З геометричної точки зору невластивий інтеграл $\int\limits_{-\infty}^{+\infty} f_\xi(x)dx$ чисельно дорівнює площі криволінійної трапеції з основою $(-\infty,+\infty)$ (рис. 1.16).

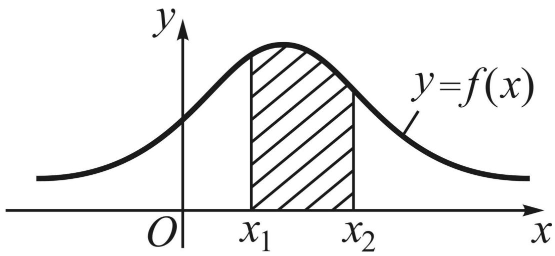

Рис. 1.16. **Щільність розподілу**

Рівність $\int\limits_{-\infty}^{+\infty} f_\xi(x)dx = 1$ з теоретико-ймовірнісної точки зору означає ймовірність вірогідної події $<$всі значення ξ належать множині $\mathbf{R}>$.

Формула (1.32) відповідає ймовірності події $(\xi < x) =<$ВВ ξ набуває значень, менших за $x>$, що узгоджується з призначенням функції $F_\xi(x)$:

$$F_\xi(x) = P(\xi < x) = \int\limits_{-\infty}^{x} f_\xi(t)dt. \qquad (1.34)$$

За допомогою (1.34), як і у випадку ДВВ, легко обчислити ймовірність настання події $(x_1 \le \xi < x_2)$. Дійсно (див. рис. 1.16):

$$(\xi < x_2) = (\xi < x_1) \cup (x_1 \le \xi < x_2) \Rightarrow P(\xi < x_2) = P(\xi < x_1) + P(x_1 \le \xi < x_2) \Rightarrow$$

$$\Rightarrow P(x_1 \le \xi < x_2) = \underbrace{P(\xi < x_2)}_{F(x_2)} - \underbrace{P(\xi < x_1)}_{F(x_1)} = \int\limits_{-\infty}^{x_2} f_\xi(t)dt - \int\limits_{-\infty}^{x_1} f_\xi(t)dt.$$

За властивістю адитивності визначеного інтеграла:

$$P(x_1 \le \xi < x_2) = F(x_2) - F(x_1) = \int\limits_{x_1}^{x_2} f_\xi(t)dt. \qquad (1.35)$$

Якщо в (1.35) покласти $x_2 = x_1 + \Delta x$ і спрямувати Δx до нуля: $\Delta x \to 0$, то отримаємо:

$$\lim_{\Delta x \to 0} \int_{x_1}^{x_1+\Delta x} f_\xi(t)dt = \int_{x_1}^{x_1} f_\xi(t)dt = 0 .$$

Цей результат означає, що ймовірність "влучання в дану точку" дорівнює нулю, тому для НВВ події $(x_1 < \xi < x_2)$, $(x_1 \le \xi < x_2)$, $(x_1 < \xi \le x_2)$, $(x_1 \le \xi \le x_2)$ є рівноможливими. За геометричним змістом інтеграла як площі можна стверджувати, що ймовірність усіх наведених подій чисельно дорівнює площі криволінійної трапеції з основою $[x_1, x_2]$, обмеженою зверху кривою $y = f_\xi(x)$ (див. рис. 1.16).

Співвідношення (1.35) указує на тісний зв'язок між функціями $F_\xi(x)$ і $f_\xi(x)$: $F_\xi(x)$ є первісною функцією для $f_\xi(x)$, отже: $F'_\xi(x) = f_\xi(x)$.

На цій підставі $F_\xi(x)$ називають також **інтегральною функцією розподілу**, а $f_\xi(x)$ – **диференціальною функцією розподілу** НВВ ξ.

Деякі важливі розподіли НВВ

1. Рівномірний розподіл. НВВ ξ має рівномірний розподіл на відрізку $[a, b]$, якщо всі її значення знаходяться на цьому відрізку, причому:

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

Імовірність того, що рівномірно розподілена на $[a, b]$ ВВ ξ набуде значень, що належать проміжку $[x_1, x_2]$, визначається формулами:

$$[x_1, x_2] \subset [a,b] \;\Rightarrow\; P(x_1 < \xi < x_2) = \frac{x_2 - x_1}{b-a},$$

$$[x_1, x_2] \not\subset [a,b] \;\Rightarrow\; P(x_1 < \xi < x_2) = \frac{\beta - \alpha}{b-a}, \quad \text{де } [\alpha, \beta] = [x_1, x_2] \cap [a,b]. \tag{1.36}$$

2. Показниковий розподіл. НВВ ξ має показниковий розподіл із параметром $\lambda > 0$, якщо:

$$f(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0, \end{cases} \qquad F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$$

Для ВВ ξ, розподіленої за показниковим законом, імовірність події $(x_1 < \xi < x_2)$ обчислюють за формулою:

$$P(x_1 < \xi < x_2) = e^{-\lambda x_1} - e^{-\lambda x_2}, \; [x_1, x_2] \subset [0, +\infty). \tag{1.37}$$

3. Нормальний розподіл. НВВ ξ має нормальний розподіл із параметрами a і σ, якщо:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} = \frac{1}{\sigma}\varphi\left(\frac{x-a}{\sigma}\right), \quad F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right),$$

де $\varphi(\cdot)$ і $\Phi(\cdot)$ – функції Гаусса та Лапласа.

У цьому випадку розрахункові формули для ймовірностей подій на проміжку виглядають так:

$$P(x_1 < \xi < x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right) = \frac{1}{\sigma} \int_{x_1}^{x_2} \varphi\left(\frac{x-a}{\sigma}\right) dx. \tag{1.38}$$

Приклад 1.15. Випадкова величина ξ має функцію розподілу:

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

Знайти: а) щільність розподілу $f(x)$; б) ймовірність попадання значень ξ в інтервал $(0, 3/2)$. Побудувати графіки функцій $F(x)$, $f(x)$.

Розв'язання: а) щільність розподілу $f(x)$ дорівнює похідній від функції розподілу:

$$f(x) = F'(x) = \begin{cases} 0, & x \leq 1, \\ x - 1/2, & 1 < x \leq 2, \\ 0, & x > 2; \end{cases}$$

б) ймовірність попадання НВВ ξ в інтервал $(0, 3/2)$ обчислимо двома способами (*вирішіть*, якому зі способів слід віддати перевагу):

$$P(0<\xi<3/2)=\int_0^{3/2} f(x)dx=\int_0^1 0dx+\int_1^{3/2}(x-1/2)dx=\frac{1}{2}(x^2-x)\Big|_1^{3/2}=\frac{3}{8},$$

$$P(0<\xi<3/2)=F(3/2)-F(0)=\frac{3}{8}-0=\frac{3}{8}.$$

Геометричне зображення функції розподілу (рис. 1.17) і щільності розподілу (рис. 1.18) не викликає труднощів, оскільки зводиться до побудови графіків основних елементарних функцій.

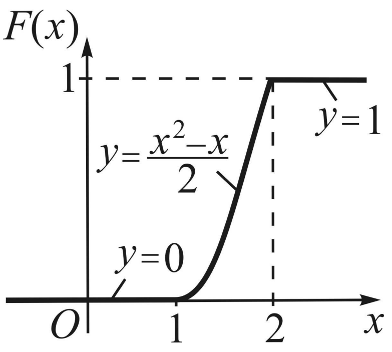

Рис. 1.17. **Функція розподілу**

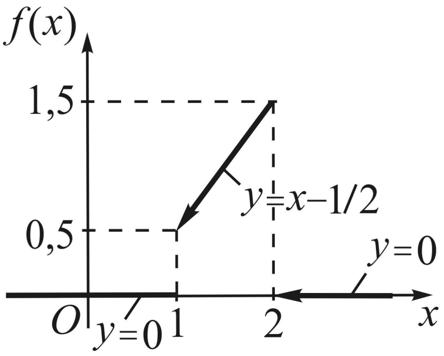

Рис. 1.18. **Щільність розподілу**

Числові характеристики НВВ

1. Математичне сподівання (середнє значення, центр розподілу). Неперервна ВВ повністю описується функцією розподілу $F_\xi(x)=P(\xi<x)$ (1.32) і щільністю розподілу $f_\xi(x)$ (1.33). Але, як і для ДВВ, у багатьох випадках виявляється достатнім вказати не зазначені функціональні характеристики ВВ, а числові характеристики, які є сталими.

Математичним сподіванням НВВ ξ називають число, яке визначається формулою:

$$M\xi=\int_{-\infty}^{+\infty} x\cdot f(x)dx=\int_{-\infty}^{+\infty} x\cdot dF(x). \quad (1.39)$$

Тобто $M\xi$ визначається невластивим інтегралом на всій числовій осі від щільності розподілу ВВ у добутку зі змінною x.

Властивості математичного сподівання НВВ такі ж, як і у ДВВ (див. п. 1.6), тільки тепер треба інтегрувати. Доведемо, наприклад, що математичне сподівання центрованої ВВ $\overset{\circ}{\xi}$ дорівнює нулю. Дійсно,

$$M\overset{\circ}{\xi} = \int_{-\infty}^{+\infty} (x - M\xi) \cdot f(x)dx = \int_{-\infty}^{+\infty} x \cdot f(x)dx - \int_{-\infty}^{+\infty} M\xi \cdot f(x)dx = 0,$$

оскільки $\int_{-\infty}^{+\infty} f(x)dx = 1.$

Приклад 1.16. Знайдемо математичне сподівання показникового розподілу:

$$f_\xi(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases} \Rightarrow \left| M\xi = \int_{-\infty}^{\infty} x \cdot f(x)dx \right| \Rightarrow$$

$$\Rightarrow M\xi = \int_{-\infty}^{0} x \cdot 0\, dx + \lambda \int_{0}^{\infty} x \cdot e^{-\lambda x}\, dx \Rightarrow \left| \begin{matrix} u = x & \Rightarrow & du = dx \\ dv = e^{-\lambda x} dx & \Rightarrow & v = -\lambda^{-1} e^{-\lambda x} \end{matrix} \right| \Rightarrow$$

$$\Rightarrow M\xi = \lambda(u \cdot v\big|_0^\infty - \int_0^\infty v \cdot du) = \lambda\left(- x\lambda^{-1} e^{-\lambda x}\Big|_0^\infty - \lambda^{-2} e^{-\lambda x}\Big|_0^\infty \right) = \frac{1}{\lambda}.$$

2. Дисперсія (розсіяння) НВВ. Для неперервних ВВ залишається без змін означення дисперсії для ДВВ ξ: $D\xi = M(\overset{\circ}{\xi})^2 = M(\xi - M\xi)^2$, але розрахункові формули виглядають інакше, оскільки зводяться до обчислення не сум, а невластивих інтегралів:

$$D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2 \cdot f(x)dx. \tag{1.40}$$

Якщо виходити з другої формули (див. (1.31)), то отримаємо:

$$D\xi = M\xi^2 - M^2\xi \Rightarrow D\xi = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - \left(\int_{-\infty}^{+\infty} x \cdot f(x)dx \right)^2. \tag{1.41}$$

В умовах прикладу 1.16 для показникового розподілу знаходимо, що $D\xi = 1/\lambda^2$ (*переконайтеся самостійно*).

Зауважимо, що всі *властивості* дисперсії, розглянуті для ДВВ, залишаються в силі.

Приклад 1.17. В умовах прикладу 1.15 знайти $M\xi$, $D\xi$, σ_ξ.

Розв'язання. Згідно з формулою $M\xi = \int\limits_{-\infty}^{+\infty} x \cdot f(x)dx$, маємо:

$$M\xi = \int\limits_1^2 x(x-1/2)dx = \int\limits_1^2 \left(x^2 - \frac{x}{2}\right)dx = \left(\frac{x^3}{3} - \frac{x^2}{4}\right)\Bigg|_1^2 = \frac{19}{12}.$$

Дисперсію $D\xi$ обчислимо за формулою $D\xi = M\xi^2 - M^2\xi$:

$$M\xi^2 = \int\limits_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int\limits_1^2 x^2(x-1/2)dx = \int\limits_1^2 \left(x^3 - \frac{x^2}{2}\right)dx = \left(\frac{x^4}{4} - \frac{x^3}{6}\right)\Bigg|_1^2 = \frac{31}{12},$$

$$D\xi = \frac{31}{12} - \left(\frac{19}{12}\right)^2 = \frac{11}{144}, \text{ звідки } \sigma_\xi = \sqrt{D\xi} = \frac{\sqrt{11}}{12} \approx 0{,}28.$$

1.8. Основи математичної статистики (МС). Вибірковий метод побудови математичної моделі масового явища

МС: основні поняття, об'єкт, предмет, основна задача

Статистика – наука про закономірності **масових явищ** (**м/я**), тобто явищ, які зустрічаються у великій кількості, достатній для встановлення закономірностей, яким вони підкоряються. Усі можливі за чисельністю результати прояву певного м/я називають **генеральною сукупністю** (**г/с**).

У вузькому розумінні "статистика", або **статистичний матеріал**, – число об'єктів, які мають ту чи іншу ознаку в досліджуваній сукупності явищ (процесів).

Статистичні методи – методи дослідження м/я на основі статистичного матеріалу.

Залежно від природи досліджуваних явищ, розрізняють такі різновиди статистики, як: соціальна, економічна, політична, виробнича та ін.

Математична статистика – наука про встановлення закономірностей м/я математичним (формальним) методом за допомогою збирання, оброблення і використання статистичного матеріалу (для досягнення поставленої мети).

Об'єктом вивчення МС є м/я, а **предметом** – установлення загальних закономірностей м/я, які *не залежать* від природи досліджуваних явищ. **Основна задача МС** – побудова за статистичними даними математичної моделі м/я, яка описує математичною мовою закономірності перебігу явища і дає змогу прогнозувати характер його подальших проявів.

Зауважимо, що між ТЙ і МС існує тісний зв'язок: МС не тільки використовує моделі, розроблені в ТЙ (ТЙ є теоретичною базою МС), а й впливає на її розвиток, оскільки різноманітність явищ (процесів) потребує розроблення нових імовірнісних моделей.

Повторна вибірка, два її трактування, вибірковий метод МС

Нехай стохастичний експеримент E описується деякою ВВ ξ (дискретною чи неперервною) з областю значень Ω_ξ і невідомою функцією розподілу $F_\xi(x)$.

Для побудови моделі за статистичними (емпіричними) даними експеримент дублюється в однакових умовах незалежним чином n разів. У фізичній інтерпретації це означає, що або проводять *послідовно* n незалежних дослідів на одній і тій самій установці, або здійснюють *одночасно* n дослідів на n однотипних установках.

Набір значень ВВ ξ із генеральної сукупності (г/с) Ω_ξ, скінченної або нескінченної: $W = \{x_1, x_2, x_3, ..., x_n (...)\}$, отриманих у результаті n дослідів (спостережень) називають випадковою **вибіркою обсягу *n***.

Якщо вибірка допускає (не допускає) наявність однакових значень ВВ ξ, то її називають **повторною** (**безповторною**). Повторну (безповторну) вибірку називають також **вибіркою з поверненням (вибіркою без повернення)**.

Смисл термінів "повторна", "безповторна" легко зрозуміти за допомогою моделі "куль і урн", або "предметів і ящиків".

Під час побудови математичної моделі E виходять із того, що вибірка є **репрезентативною** (типовою, характерною), тобто вона достатньо добре представляє (відображує) властивості г/с. Для репрезентативності вибірки її здійснюють так, щоб усі елементи г/с мали однакову можливість бути включеними у вибірку. Для забезпечення цієї умови використовують (за відомої г/с) таблиці або датчик випадкових чисел.

Типовий шлях математико-статистичної побудови E такий.

Задача I (*висунення гіпотези*). Використовуючи графічні методи зображення (подання) вибірки, з урахуванням природи явища зробити висновок щодо можливого закону розподілу ВВ ξ, що еквівалентно висуненню гіпотези про вигляд функції розподілу $F_{\xi}(x)$.

Задача II (*оцінка параметрів*). Знайти наближені значення параметрів гіпотетичної функції розподілу $\widetilde{F}_{\xi}(x)$ за даними вибірки. *Наприклад*, для показникового розподілу – це параметр λ, для нормального – параметри a, σ.

Задача III (*перевірка адекватності*). Перевірити, чи узгоджуються наслідки експерименту із запропонованою моделлю. У випадку позитивного результату розпочинають прогнозування, яке передбачає відповіді на запитання: що відбудеться за заданих умов; яких наслідків слід очікувати, якщо діяти тим чи іншим способом; як треба діяти, щоб сталася певна подія і таке інше. У протилежному випадку уточнюють модель, для чого збільшують обсяг вибірки і (або) точність оцінки параметрів або висувають інші (альтернативні) гіпотези.

Зазначений шлях побудови моделі стохастичного експерименту прийнято називати **вибірковим методом**.

Первинна обробка результатів спостережень. Дискретні та неперервні статистичні (емпіричні) ряди

Нехай під час проведення E із г/с Ω_{ξ} здійснено повторну вибірку обсягу n: $W=\{x_1, x_2, x_3, ..., x_n (...)\}$.

Вибірку, елементи якої розташовані у неспадному порядку: $x_1 \le x_2 \le x_3 \le ... \le x_n$, називають **варіаційним рядом**. *Наприклад:*

$W=\{1,2,5,1,3,2\}$ – вибірка $\Rightarrow$ варіаційний ряд: $(1,1,2,2,3,5)$.

Якщо вибірка обсягу n містить k різних елементів ($k \le n$), при цьому x_i (тут $i=\overline{1,k}$) зустрічається m_i разів, то число m_i називають **частотою** елемента x_i, а відношення $w_i = m_i / n$ – **відносною частотою** (**частістю**) елемента x_i. Різні за значенням члени варіаційного ряду називають **варіантами** досліджуваної ознаки (властивості).

Так, у наведеній вибірці W: $(1,2,3,5)$ – варіанти; $m_1=m_2=2$, $m_3=m_4=1$ – частоти; $w_1=2/6$, $w_2=1/6$ – частості. Помічаємо, що:

$$\sum_{i=1}^{k} m_i = n, \quad \sum_{i=1}^{k} w_i = \sum_{i=1}^{k} \frac{m_i}{n} = 1. \qquad (1.42)$$

Наведіть словесні формулювання.

Частоти та частості варіант об'єднуються загальною назвою – **вага варіант** (абсолютна чи відносна). Підрахунок частот і частостей називають **групуванням даних** за варіантами.

Нехай г/с Ω_ξ – дискретна множина, тобто ξ – ВВ дискретного типу.

Послідовність пар (x_i, m_i) і (x_i, w_i), $i=\overline{1,k}$, називають, відповідно, **дискретним статистичним рядом частот** і **частостей**:

x_i	x_1	x_2	…	x_k
m_i	m_1	m_2	…	m_k

– статряд частот;

x_i	x_1	x_2	…	x_k
w_i	w_1	w_2	…	w_k

– статряд частостей.

Для наведеної вибірки:

x_i	1	2	3	5
m_i	2	2	1	1

– статряд частот;

x_i	1	2	3	5
w_i	$2/6$	$2/6$	$1/6$	$1/6$

– статряд частостей.

Якщо ξ – ВВ неперервного типу, то весь проміжок спостережуваних значень: $[x_{\min}, x_{\max}]=[x_0, x_k]$ поділяють точками x_1, x_2, …, x_{k-1} на k часткових проміжків $(5 \le k \le 15)$ з однаковими довжинами: $[x_0, x_1)$, $[x_1, x_2)$, …, $[x_{k-1}, x_{\max}]$, останній (k-й) проміжок є сегментом. Потім підраховують частоти та частості попадання вибіркових значень ξ у часткові проміжки.

Послідовність пар $\left([x_{i-1}, x_i), m_i\right)$ і $\left([x_{i-1}, x_i), w_i\right)$, $i=\overline{1,k}$, називають, відповідно, **інтервальним**, або **неперервним**, **статистичним рядом частот** і **частостей**:

$[x_{i-1}, x_i)$	$[x_0, x_1)$	$[x_2, x_3)$	…	$[x_{k-1}, x_k]$	Інтервальні статряди:
m_i	m_1	m_2	…	m_k	– частот;
$w_i = m_i/n$	m_1/n	m_2/n	…	m_k/n	– частостей.

Довжину h відрізка $[x_{i-1}, x_i]$ називають **кроком розбиття** всього проміжку $[x_{\min}, x_{\max}]$: $h = x_i - x_{i-1}$, а різницю $R = x_{\max} - x_{\min}$ – **розмахом вибірки**.

Для подальшої обробки неперервний статистичний ряд перетворюють на дискретний, замінюючи кожний із часткових інтервалів його *представником* x_i^* $(i = \overline{1,k})$ – середнім арифметичним кінців інтервалу:

$$x_i^* = (x_{i-1} + x_i)/2 \text{, або } x_i^* = x_{i-1} + h/2 = x_i - h/2. \tag{1.43}$$

Представників x_i^* називають також **центрами усереднення** часткових інтервалів.

Приклад 1.18. За заданим інтервальним статистичним рядом частостей побудувати дискретний ряд за центрами усереднення:

$[x_{i-1}, x_i)$	$[1,5)$	$[5,9)$	$[9,13)$	$[13,17)$	$[17,21]$	– часткові інтервали;
$w_i = \dfrac{m_i}{n}$	$\dfrac{2}{20}$	$\dfrac{5}{20}$	$\dfrac{7}{20}$	$\dfrac{3}{20}$	$\dfrac{3}{20}$	– відносні частоти.

Розв'язання. Візуальний аналіз показує: кількість часткових інтервалів $k=5$, крок розбиття $h = x_i - x_{i-1} = 4$, обсяг вибірки $n = 20$. Згідно з (1.43) знаходимо представників кожного часткового інтервалу:

x_i^*	3	7	11	15	19	– центри усереднення;
$w_i = \dfrac{m_i}{n}$	$\dfrac{2}{20}$	$\dfrac{5}{20}$	$\dfrac{7}{20}$	$\dfrac{3}{20}$	$\dfrac{3}{20}$	– відносні частоти.

Дискретний (неперервний) статистичний ряд частостей називають **статистичним**, або **емпіричним**, законом розподілу ДВВ ξ (НВВ ξ). Їх об'єднують загальною назвою – **статистичні розподіли вибірки**.

Статистичні (емпіричні) розподіли вибірки описуються в термінах ТЙ, перед якими додають слово "статистичний" або "емпіричний": статистичний закон розподілу, статистична функція розподілу і таке інше.

Зауважимо, що в МС статистичний розподіл установлює відповідність між емпіричними даними – елементами вибірки (або інтервалами емпіричних даних) із г/с – і відповідними їм частостями.

А в ТЙ під розподілом ВВ розуміють відповідність між можливими значеннями або інтервалами значень ВВ і ймовірністю їх появи.

З огляду на це, приходимо до **висновку**, що відносні частоти в статистичних рядах дають наближене значення (оцінку) відповідної ймовірності, яка тим краща, чим більший обсяг вибірки.

Дискретний статистичний розподіл (ДСР) вибірки, геометричне зображення, числові характеристики

Емпіричним полігоном частот, **частостей** ДСР, відповідно, називають ламану з вершинами в точках (x_i, m_i), $(x_i, m_i/n)$. Форми полігонів частот і відносних частот схожі: ординати вершин полігону частостей будуть в n разів меншими від ординат вершин полігону частот.

Приклад 1.19. Побудувати полігон частот ДСР із прикладу 1.18.

Розв'язання. Згідно з означенням будуємо вершини ламаної (x_i, m_i), $i=\overline{1,5}$, і з'єднуємо їх відрізками прямих (рис. 1.19). Він співпаде з полігоном частостей, для якого масштаб слід вибрати у двадцять разів більшим.

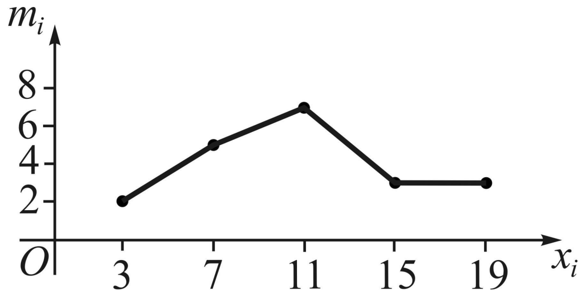

Рис. 1.19. **Полігон частот**

Емпіричною функцією розподілу $F^*(x)$ називають функцію, яка для кожного x визначає відносну частоту події $(\xi < x)$:

$$F^*(x) = w(\xi < x) = \frac{m_x}{n}, \text{ або } F^*(x) = \frac{1}{n} \cdot \sum_{i: x_i < x} m_i , \qquad (1.44)$$

де $m_x = \sum\limits_{i: x_i < x} m_i$ – число значень ВВ ξ у вибірці із г/с Ω_ξ, менших за x;

n – обсяг вибірки;

$w(\xi < x)$ – частість події $<$ВВ ξ набуває значень, менших за $x>$.

$F^*(x)$ називають ще **функцією нагромадження відносних частот**, а її графік – кумулятою.

Кумулята – це східчаста лінія зі скінченними розривами в точках $x = x_i \in \Omega_\xi$. Величина стрибка (сходинки) вказує на частість, з якою зустрічається відповідний елемент у вибірці. У точках неперервності функція розподілу набуває сталих значень (*чому?*).

Властивості $F^*(x)$:

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

2) $F^*(x)\Big|_{x \le x_{\min}} = 0$, де $x_{\min}$ є найменшою варіантою вибірки;

3) $F^*(x)\Big|_{x > x_{\max}} = 1$, де $x_{\max}$ є найбільшою варіантою вибірки;

4) $F^*(x)$ є неспадною функцією: $x_2 > x_1 \Rightarrow F^*(x_2) \ge F^*(x_1)$.

Приклад 1.20. За заданим дискретним статистичним розподілом вибірки побудувати $F^*(x)$ і зобразити її графічно:

$\xi = x_i$	–3	–2	–1	1	2	3	– елементи вибірки;
m_i	5	10	15	20	40	10	– частоти;
w_i	0,05	0,1	0,15	0,2	0,4	0,1	– відносні частоти.

Розв'язання. Згідно з означенням і властивостями функція $F^*(x)$, описана в символах, має такий вигляд:

$$F^*(x) = w(\xi < x) = \frac{m_x}{n} = \begin{cases} 0, & x \le -3, \\ 0{,}05, & -3 < x \le -2, \\ 0{,}15, & -2 < x \le -1, \\ 0{,}3, & -1 < x \le 1, \\ 0{,}5, & 1 < x \le 2, \\ 0{,}9, & 2 < x \le 3, \\ 1, & x > 3. \end{cases}$$

Графічне зображення $F^*(x)$ подано на рис. 1.20.

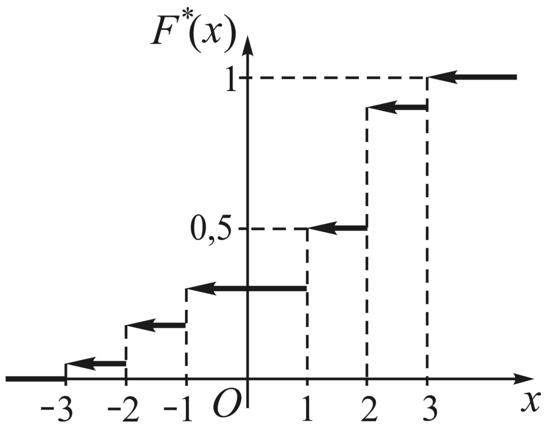

Рис. 1.20. **Емпірична функція розподілу**

Підкреслимо, що емпірична функція розподілу є статистичним аналогом функції розподілу в ТЙ, тільки $F^*(x)$ визначає не ймовірності, а частості як наближені значення ймовірностей.

Вибіркові числові характеристики ДСР: означення, різновиди

Статистичними (емпіричними, вибірковими) числовими характеристиками (ч/х) досліджуваної ВВ ξ називають **наближені значення** (**оцінки**) відповідних ч/х у ТЙ, отримані на основі вибірки. (Інакше кажучи, вибіркові ч/х – це статистичні аналоги відповідних ч/х, які розглядалися в ТЙ і які отримують заміною *ймовірності* події її *відносною частотою*).

1. Вибірковим математичним сподіванням (**середнім**) називають середнє арифметичне елементів вибірки та позначають через $M^*\xi$, $\bar{x}$ або $\bar{x}_{\text{в}}$ – "ікс середнє вибіркове":

$$\bar{x}_{\text{в}} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \,. \tag{1.45}$$

Якщо відомі варіанти $x_j \ (j = \overline{1,k})$ і їхні частоти m_j, то:

$$\bar{x}_{\text{в}} = \frac{1}{n} \cdot \sum_{j=1}^{k} m_j \, x_j \,. \tag{1.46}$$

Приклад 1.21. Здійснено вибірку $W = \{1, 2, 3, 1, 4, 1, 2, 3, 4, 3\}$. Знайти емпіричне середнє.

Розв'язання. За формулами (1.45) і (1.46) маємо:

$$\bar{x}_{\text{в}} = (1 + 2 + 3 + 1 + 4 + 1 + 2 + 3 + 4 + 3)/10 = 2{,}4;$$
$$\bar{x}_{\text{в}} = (3 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 + 2 \cdot 4)/10 = 2{,}4 \,.$$

Величину $\bar{x}_{\text{в}}$ можна розглядати як математичне сподівання ДВВ, яка набуває значень x_1, x_2, …, x_k з імовірностями m_1/n, m_2/n, …, m_k/n, тому що $M\xi = \sum_{i=1}^{n} x_i \, p_i$, а $\bar{x}_{\text{в}}$ – його статистичний аналог.

2. Вибіркова дисперсія – це середнє арифметичне квадратів відхилень елементів вибірки від емпіричного (вибіркового) середнього:

$$D_{\text{в}} = s^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x}_{\text{в}})^2 \quad \text{– за незгрупованими даними;} \tag{1.47}$$

$$s^2 = \frac{1}{n} \cdot \sum_{j=1}^{k} (x_j - \bar{x}_{\text{в}})^2 m_j \quad \text{– за згрупованими даними.} \tag{1.48}$$

Статистичним аналогом формули в ТЙ: $D\xi = \sigma^2 = M\xi^2 - M^2\xi$, є співвідношення:

$$s^2 = \overline{x^2} - \bar{x}^2 \quad \text{– середнє квадратів без квадрата середнього.} \tag{1.49}$$

3. Середнє квадратичне відхилення вибірки – це арифметичний корінь із вибіркової дисперсії: $s=\sqrt{s^2}$. Середнє квадратичне відхилення вимірює розсіяння варіант вибірки відносно $\overline{x}_{\text{в}}$ у тих самих одиницях, в яких вимірюється досліджувана ознака.

Приклад 1.22. В умовах прикладу 1.21 знайти вибіркову дисперсію та середнє квадратичне відхилення.

Розв'язання. Обчислюємо середнє квадратів варіант:

$$\overline{x^2}=(3\cdot 1^2+2\cdot 2^2+3\cdot 3^2+2\cdot 4^2)/10=7,$$

і за формулою (1.49), з урахуванням, що $\overline{x}=2{,}4$, отримуємо:

$$s^2=\overline{x^2}-\overline{x}^2=7-(2{,}4)^2=1{,}24 \;\Rightarrow\; s=\sqrt{1{,}24}\approx 1{,}11.$$

4. Модою (Mo^*) ДСР вибірки називають варіанту, що має найбільшу частоту появи. Мод може бути декілька. Якщо ДСР має одну моду, то його називають **одномодальним**, якщо дві моди – **двомодальним** і так далі. *Наприклад,* вибірка $W=\{1,\,2,\,3,1,\,4,1,\,2,\,3,\,4,\,3\}$ – двомодальна.

5. Медіаною *(*Me^*) ДСР вибірки називають варіанту, яка поділяє варіаційний ряд на дві частини, рівні за кількістю варіант. Якщо k – число варіант – парне число, то Me^* знаходять за формулою:

$$Me^*=(x_{k/2}+x_{k/2+1})/2.$$

Неперервний статистичний розподіл (НСР) вибірки, геометричне зображення, числові характеристики

Емпіричним полігоном частот, **частостей** НСР, відповідно, називають ламану з вершинами у точках $(x_i^*, m_i/h)$, $(x_i^*, w_i/h)$, де h – крок розбиття всього проміжку $[x_{\min}, x_{\max}]$. Форми полігонів частот і відносних частот схожі: ординати вершин полігону частостей будуть в n разів меншими від ординат вершин полігону частот.

Гістограмою частот (частостей) НСР називають фігуру, складену із k прямокутників з основами h і висотами m_i/h (w_i/h), $i=\overline{1,k}$. (*Пропонуємо* обчислити площу одного прямокутника та всієї гістограми, *зробити* відповідний висновок).

Приклад 1.23. За заданою вибіркою скласти інтервальний статистичний ряд із шістьма частковими інтервалами ($k=6$) і побудувати гістограму частостей:

$$W=\{14, 6, 9, 8, 5, 10, 9, 7, 11, 4, 8, 10, 6, 7, 15, 8, 11, 5, 12, 4, 7, 9, 12, 10, 13\}.$$

Розв'язання. Розташовуємо елементи вибірки (за означенням інтервального статряду) у неспадному порядку:

$$(4, 4, 5, 5, 6, 6, 7, 7, 7, 8, 8, 8, 9, 9, 9, 10, 10, 10, 11, 11, 12, 12, 13, 14, 15),$$

звідки: ($x_{\min}=4$, $x_{\max}=15$) $\Rightarrow$ $R=15-4=11$ – розмах вибірки.

Знаходимо орієнтовний крок розбиття: $h=R/k=11/6\approx 2$. Щоб крок виразити цілим числом (для зручності обчислень), на практиці збільшують розмах на величину, яка не перевищує $h/2$. Для цього зменшують $x_{\min}$ і (або) збільшують $x_{\max}$. У нашому випадку діємо так: беремо $[x_{\min}-h/2, x_{\max}]=[3,15]$ або $[x_{\min}, x_{\max}+h/2]=[4,16]$, тобто забезпечуємо $h=12/6=2$. Зупинимось на сегменті $[3,15]$. Розбиваємо його на шість часткових проміжків, будуємо інтервальний статистичний ряд частостей і знаходимо величини w_i/h $(i=\overline{1,6})$ – висоти прямокутників:

$[x_{i-1}, x_i)$	$[3,5)$	$[5,7)$	$[7,9)$	$[9,11)$	$[11,13)$	$[13,15]$
$w_i=\frac{m_i}{n}$	$\frac{2}{25}$	$\frac{4}{25}$	$\frac{6}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{3}{25}$
$\frac{w_i}{h}=\frac{m_i}{nh}$	$\frac{2}{50}$	$\frac{4}{50}$	$\frac{6}{50}$	$\frac{6}{50}$	$\frac{4}{50}$	$\frac{3}{50}$

Далі будуємо гістограму – стовпчасту діаграму (рис. 1.21).

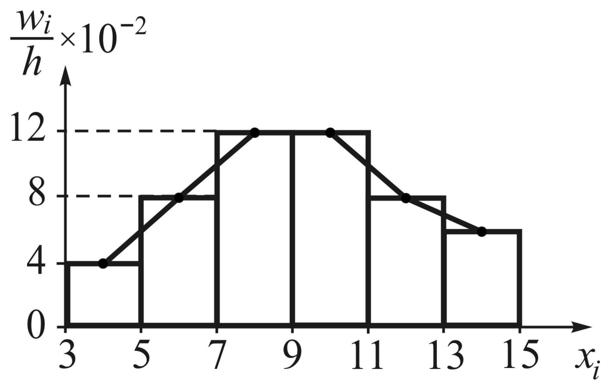

Рис. 1.21. **Гістограма і полігон**

Відлік по горизонтальній осі здійснювався від $x_{\min}-1$, а по вертикалі відкладалися величини, в 100 разів більші від w_i/h.

Якщо з'єднати ламаною середини верхніх основ прямокутників, то отримаємо полігон частостей розподілу.

Емпіричною функцією розподілу НСР (за аналогією з (1.44) для ДСР) називають функцію $F^*(x)$, яка для кожного x визначає відносну частоту події $(\xi < x)$. Тобто $F^*(x)$ – **функція нагромадження відносних частот**: $F^*(x) = w_x^{\text{нагр}}$, а її графік – **кумулятивна крива**.

Графік функції $F^*(x)$ має вигляд ламаної лінії, яка зростає на кожному частковому інтервалі, проходить через точки $(x_i, F^*(x_i))$, $i = \overline{1,k}$, де x_i – права межа часткового інтервалу, та наближається до одиниці. Масштаб на кожній осі вибирають довільно.

Приклад 1.24. В умовах прикладу 1.23 побудувати кумулятивну криву.

Розв'язання. Нагромаджуємо відносні частоти:

x_i	3	7	9	11	13	15
$w_i = m_i/n$	$2/25$	$4/25$	$6/25$	$6/25$	$4/25$	$3/25$
$w_x^{\text{нагр}}$	0,08	0,24	0,48	0,72	0,88	1

.

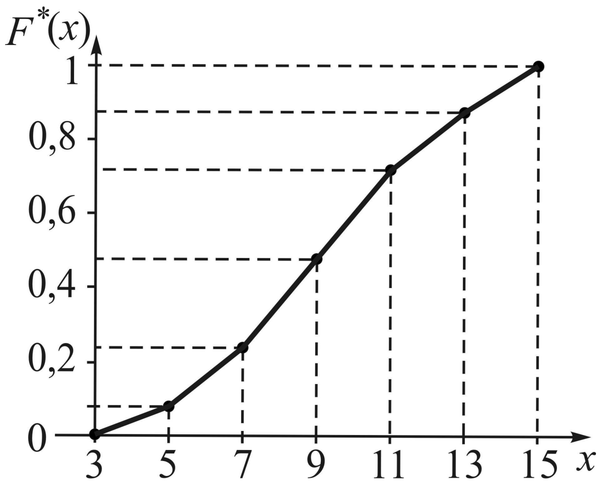

Рис. 1.22. **Кумулятивна крива**

Будуємо ламану згідно з означенням (рис. 1.22).

Зауважимо, що область існування функції розподілу г/с $F_\xi(x)$ розбивається вибіркою на три частини: $(-\infty, x_{\min})$, $[x_{\min}, x_{\max}]$, $(x_{\max}, +\infty)$, де $x_{\min}$ є правою межею першого проміжку, від якої починається ламана. Графік самої $F_\xi(x)$ буде загалом кусково-неперервною лінією.

Емпіричною щільністю НСР називають функцію, яка в центрах усереднення часткових інтервалів x_i^* $(i = \overline{1,k})$ набуває значень, що дорівнюють відношенню частості w_i до величини кроку розбиття h усього проміжку $[x_{\min}, x_{\max}]$:

$$f^*(x_i^*) = \frac{w_i}{h}, \text{ або } f^*(x_i^*) = \frac{m_i}{nh}, \; i = \overline{1,k}. \tag{1.50}$$

Побудова графіка щільності зводиться до нанесення на координатну площину точок $(x_i^*, f^*(x_i^*))$ і послідовного з'єднання їх відрізками прямої. Ураховуючи означення полігону та рис. 1.21, приходимо до **висновку:** графіком емпіричної щільності є полігон відносних частот; за побудованим графіком щільності $f^*(x_i^*)$ можна відразу побудувати гістограму, і навпаки. Звичайно, графіком щільності розподілу г/с $f_\xi(x)$ не є загалом ламана: лінія може бути неперервною або кусково-неперервною.

Наприклад, графіком щільності нормального розподілу з $a=0$ і $\sigma=1$ є "дзвіночок" (рис. 1.23).

На противагу функції розподілу щільність не є універсальною характеристикою, вона існує тільки для неперервних випадкових величин.

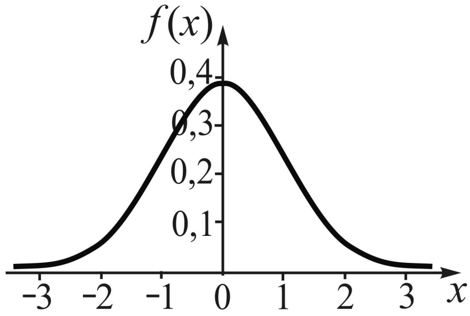

Рис. 1.23. **Дзвіночок**

Вибіркові числові характеристики НСР: означення, різновиди

Для визначення вибіркового середнього $\bar{x}_в$, вибіркової дисперсії $D_в = s^2$, вибіркового середньоквадратичного відхилення $\sigma_в = s$ переходять від інтервального розподілу до дискретного, варіантами якого є середини часткових інтервалів (див. (1.43)):

$$x_i^* = (x_{i-1} + x_i)/2, \text{ або } x_i^* = x_{i-1} + h/2 = x_i - h/2.$$

Названі числові характеристики обчислюють за такими формулами.

1. Вибіркове математичне сподівання (**середнє**):

$$\bar{x}_в = \frac{1}{n} \cdot \sum_{i=1}^{k} x_i^* m_i . \qquad (1.51)$$

2. Вибіркова дисперсія:

$$D_в = s^2 = \frac{1}{n} \cdot \sum_{i=1}^{k} (x_i^*)^2 m_i - (\bar{x}_в)^2. \qquad (1.52)$$

Якщо n лежить у межах від 30 до 50 або $n < 30$ (**мала вибірка**), то користуються формулою так званої **виправленої дисперсії**:

$$\hat{D}_{в} = \frac{n}{n-1} D_{в}, \text{ або } \hat{D}_{в} = \frac{1}{n-1} \cdot \left(\sum_{i=1}^{k} (x_i^*)^2 m_i - n(\bar{x}_{в})^2 \right). \quad (1.53)$$

3. Середнє квадратичне відхилення:

$$\sigma_{в} = s = \sqrt{D_{в}} \text{ або } \hat{\sigma}_{в} = \hat{s} = \sqrt{\hat{D}_{в}}. \quad (1.54)$$

4. Мода (Mo^*) **НСР** – варіанта, що має найбільшу частоту появи.

5. Медіана *(*Me^*) **НСР** – варіанта, яка поділяє варіаційний ряд на дві частини, рівні за кількістю варіант.

Якщо k – число варіант (центрів групування x_i^*, $i = \overline{1,k}$) – парне число, то Me^* знаходять за формулою:

$$Me^* = (x_{k/2}^* + x_{k/2+1}^*)/2.$$

Приклад 1.25. В умовах прикладу 1.23 знайти основні числові характеристики вибірки.

Розв'язання. Розрахунки будемо проводити для збільшеного на одиницю розмаху вибірки. Знаходимо центри групування часткових проміжків:

$[x_{i-1}, x_i)$	$[3,5)$	$[5,7)$	$[7,9)$	$[9,11)$	$[11,13)$	$[13,15]$
x_i^*	4	6	8	10	12	14
m_i	2	4	6	6	4	3

За формулами (1.51), (1.53), (1.54) послідовно обчислюємо:

$$\bar{x}_{в} = \frac{1}{n} \cdot \sum_{i=1}^{6} x_i^* m_i = \frac{1}{25} \cdot (4 \cdot 2 + 6 \cdot 4 + 8 \cdot 6 + 10 \cdot 6 + 12 \cdot 4 + 14 \cdot 3) = 9{,}2;$$

$$\hat{D}_{в} = \frac{1}{24} \cdot \left(\sum_{i=1}^{6} (x_i^*)^2 m_i - 25\,(\bar{x}_{в})^2 \right) = \frac{1}{24} \cdot \left(2324 - 25 \cdot (9{,}2)^2 \right) = \frac{26}{3} \approx 8,(6);$$

$$\hat{\sigma}_{в} = \hat{s} = \sqrt{\hat{D}_{в}} = \sqrt{8,(6)} \approx 2{,}94.$$

Підсумок. Проведена первинна обробка вибірки для ДСР і НСР дає змогу розв'язати **задачу I** *(висунення гіпотези)* і **задачу II** (*оцінки параметрів* гіпотетичного розподілу) вибіркового методу. Так, для НСР у прикладах 1.23 – 1.25 за виглядом гістограми чи полігону можна зробити припущення (висунути гіпотезу), що г/с розподілена за нормальним законом зі щільністю розподілу $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$, де числові параметри a і σ визначаються співвідношеннями: $M\xi=a$, $D\xi=\sigma^2$. Оцінками (наближеними значеннями) параметрів є вибіркові числові характеристики: $a\approx\overline{x}_в=9{,}2$, $\sigma\approx\hat{s}=\sqrt{8,(6)}=2,94$. Таким чином, гіпотетична (теоретична) щільність розподілу г/с $\tilde{f}(x)$ така:

$$\tilde{f}(x)=\frac{1}{2{,}94\cdot\sqrt{2\pi}}\cdot e^{-\frac{3\cdot(x-9{,}2)^2}{2\cdot 26}}. \tag{1.55}$$

Постановка задачі статистичної перевірки гіпотез. Критерії згоди

Нехай **задача I** (*висунення гіпотези)* та **задача II** (*оцінка параметрів)* побудови математичної моделі E вибірковим методом розв'язані. Прийняту гіпотезу називають **початковою** (**основною**, **нульовою**), і пишуть: $H_0:F(x)=\tilde{F}(x,\theta)$, тобто приймають, що функція розподілу г/с $F(x)$ дорівнює теоретичній функції $\tilde{F}(x,\theta)$, побудованій за вибіркою, де θ – параметр (або параметри) функції розподілу. Припущення, яке приймається у разі відхилення H_0, називають **альтернативною гіпотезою**, і пишуть: $H_a:F(x)\neq\tilde{F}(x,\theta)$.

Наступний крок побудови моделі – перевірка адекватності, тобто узгодженості між емпіричною та теоретичною функціями розподілу. Інакше кажучи, постає задача перевірки гіпотези H_0. Як правило, перевірку адекватності проводять, порівнюючи:

1) емпіричні частості $w_i=m_i/n$ з теоретичними ймовірностями $\tilde{p}_i$, які обчислюють за гіпотетичним законом розподілу;

2) емпіричні частоти m_i і теоретичні частоти $\tilde{m}_i$, де $\tilde{m}_i$ виражається через $\tilde{p}_i$: $\tilde{p}_i=\tilde{m}_i/n \;\Rightarrow\; \tilde{m}_i=n\tilde{p}_i$;

3) емпіричну та гіпотетичну функції розподілу: $F^*(x)$, $\tilde{F}(x)$.

Візуально про близькість теоретичного й емпіричного законів розподілу можна судити за геометричним зображенням $F^*(x)$ і $\widetilde{F}(x)$ або $P^*(x)$ і $\widetilde{P}(x)$ чи $f^*(x)$ і $\widetilde{f}(x)$, але це дає тільки якісну оцінку того, чи узгоджуються теоретичний і емпіричний розподіли.

Правила, за допомогою яких вирішується питання про прийняття чи відхилення гіпотези H_0, називають **критеріями згоди** між теоретичним і емпіричним розподілами. Одним із них є **критерій Пірсона** (**хі-квадрат**), який визначається однойменною теоремою.

Нехай ξ – нормально розподілена ВВ з $M\xi = a$ і $D\xi = \sigma^2$, тоді ВВ $u = (\xi - a)/\sigma$ називають **стандартизованою** нормально розподіленою ВВ, і пишуть: $u \in N(0,1)$. Квадрат стандартизованої ВВ u називають **ВВ χ^2** (**хі-квадрат**) **з одним ступенем свободи** ν, тобто $\nu = 1$.

Суму квадратів n стандартизованих ВВ u_i $(i = \overline{1,n})$ називають **ВВ χ^2 з n ступенями свободи**, тобто $\nu = n$: $\chi^2 = u_1^2 + u_2^2 + ... + u_n^2$, або **розподілом Пірсона χ^2**. Величину

$$\chi^2_{\text{в}} = \sum_{i=1}^{k} \frac{(m_i - \widetilde{m}_i)^2}{\widetilde{m}_i}, \text{ де } \widetilde{m}_i = n\widetilde{p}_i, \tag{1.56}$$

називають (вибірковою) **статистикою Пірсона**.

Теорема 1.6 (*Пірсона*). Якщо гіпотеза H_0 є правильною, то незалежно від закону розподілу досліджуваної ВВ статистика $\chi^2_{\text{в}}$ за умови, що $k \to \infty$, має розподіл Пірсона з $\nu = k - r - 1$ ступенями свободи, де r – число параметрів гіпотетичного розподілу.

Для χ^2 (рис. 1.24) складені таблиці (табл. Д.1 додатка Д) так званих **критичних точок** $\chi^2_{\nu,\alpha}$ з двома входами: ν, α, де ν – число ступенів свободи, α – рівень значущості критерію – ймовірність того, що вибіркова статистика перевищує критичну точку $\chi^2_{\text{крит}} = \chi^2_{\nu,\alpha}$.

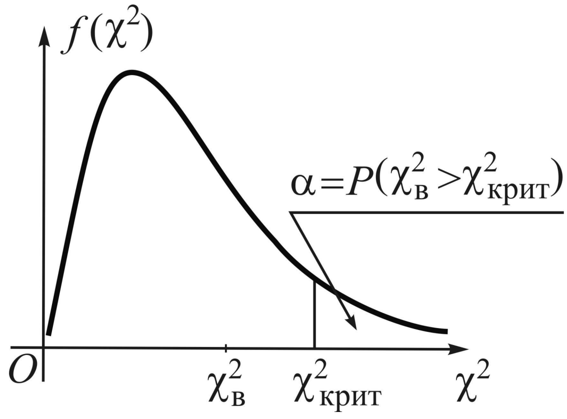

Рис. 1.24. **Розподіл χ^2**

Критерієм χ^2 користуються так:

1) *підраховують* $\chi^2_{\text{в}}$ за деякою вибіркою та відомим гіпотетичним розподілом;

2) *вибирають* рівень довіри до висунутої гіпотези (ймовірність p), близький до 1: $p = 0{,}9;\ 0{,}95;\ 0{,}99$;

3) *знаходять* за таблицями розподілу χ^2 відповідну критичну точку $\chi^2_{\nu,\alpha}$: $\alpha = 1 - p = 0{,}1; 0{,}05; 0{,}01$.

Якщо виявиться, що $\chi^2_{\text{в}} < \chi^2_{\nu,\alpha}$, тобто попадаємо в зону допустимих (докритичних) значень, то нема підстав для відхилення гіпотези.

Якщо $\chi^2_{\text{в}} \ge \chi^2_{\nu,\alpha}$, то гіпотеза відхиляється.

Приклад 1.26. Перевірити на адекватність гіпотезу (1.55), висунуту на підставі результатів обробки вибірки W (див. приклади 1.23 – 1.25) з рівнем значущості $\alpha = 0{,}05$.

Розв'язання. Підраховуємо статистику Пірсона $\chi^2_{\text{в}}$ (1.56), для чого обчислюємо теоретичні частоти $\tilde{m}_i$ за щільністю гіпотетичного розподілу:

$$\tilde{f}(x) = \frac{1}{2{,}94\cdot\sqrt{2\pi}}\cdot e^{-\frac{3\cdot(x-9{,}2)^2}{2\cdot 26}},$$

здійснюючи перехід до стандартизованої змінної $u = (x-a)/\sigma$, з подальшим використанням функції Гаусса $\varphi(x)$ (див. формулу (1.19)).

Результати розрахунків уміщено в табл. 1.1.

Таблиця 1.1

Перевірка адекватності гіпотези за критерієм Пірсона

N п/п	x_i^*	$u_i = \frac{x_i^* - a}{\sigma}$	$\varphi(u_i)$	$\tilde{m}_i = \frac{nh}{\sigma}\cdot\varphi(u_i)$		m_i		$\frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$
1	4	– 1,766	0,084	1,424	5,177	2	6	0,131
2	6	– 1,087	0,221	3,753		4		
3	8	– 0,408	0,367		6,236		6	0,009
4	10	0,272	0,384		6,530		6	0,043
5	12	0,951	0,254	4,310	6,103	4	7	0,132
6	14	1,630	0,106	1,793		3		
Σ					24,046		25	$\chi^2_{\text{в}} = 0{,}315$

Рекомендації щодо використання критерію Пірсона: якщо число спостережень (частота m_i) в інтервалі менше від п'яти, то цей інтервал об'єднують із сусідніми, а їхні частоти додають; відповідні їм теоретичні частоти $\widetilde{m}_i$ також треба додати. У табл. 1.1 було об'єднано перший інтервал з другим і п'ятий із шостим. Далі для визначення числа ступенів свободи ν за k слід брати кількість інтервалів після об'єднання.

У підсумку (див. табл. 1.1) маємо: $\chi^2_{\text{в}} = 0{,}315$. Для числа ступенів свободи $\nu = 4-2-1=1$ і рівня значущості $\alpha=0{,}05$ за таблицею критичних точок визначаємо: $\chi^2_{\text{крит}} = \chi^2_{1;0,05} = 3{,}8$.

Як бачимо, $\chi^2_{\text{в}} = 0{,}315 < \chi^2_{\text{крит}} = 3{,}8$, тому підстав для відхилення гіпотези H_0 про нормальний розподіл г/с немає. (*Подумайте*, як треба було б діяти у протилежному випадку.)

Якщо вважати, що задана вибірка відображає місячний заробіток працівників деякого підприємства (в ум. од.), то виходить, що він розподілений за нормальним законом.

1.9. Кореляційний і регресійний аналіз

Поняття про математичні моделі багатовимірних ВВ. Основні поняття кореляційного та регресійного аналізу

Масові явища навколишнього світу здебільше описуються не однією, а кількома ВВ: ξ_1, ξ_2, ..., ξ_n, які називають **системою ВВ** (або **n-вимірним випадковим вектором** (**ВВ-n**): $\xi = (\xi_1, \xi_2, ..., \xi_n)$.

Вибірковий метод побудови математичної моделі системи ВВ складається з тих самих етапів, що і для одновимірних ВВ: *висунення гіпотези, оцінка параметрів, перевірка адекватності.* У разі випадкових векторів ця задача стає більш складною, оскільки треба встановити закони розподілу для кожної складової випадкового вектора, тобто окремої ВВ, і закон їх сумісного розподілу. Важливим у цьому випадку є питання про залежність (чи незалежність) складових частин. Якщо ВВ незалежні між собою, то закон їх сумісного розподілу знайти дуже легко.

Наприклад, якщо $\mu=(\xi,\eta)$ містить незалежні ВВ ξ, η і відомі функції розподілу $F_\xi(x)$, $F_\eta(y)$, то $F_{\xi,\eta}(x,y)=F_\xi(x)\cdot F_\eta(y)$. Аналогічно для щільностей розподілу $f_{\xi,\eta}(x,y)$, якщо ξ, η – неперервні ВВ.

У випадку залежних ξ і η обмежуються дослідженням так званих **умовних математичних** сподівань – математичних сподівань однієї ВВ залежно від фіксованих значень x і y, відповідно, іншої:

$$\varphi(x)=M(\eta\,|\xi=x)=M(\eta/x),\ \ \psi(y)=M(\xi|\eta=y)=M(\xi/y). \qquad (1.57)$$

Розділ МС, у якому вивчають методи побудови моделей багатовимірних ВВ, називають **багатовимірним** (або **множинним**) **статистичним аналізом** (МСА).

Якщо розподіл однієї ВВ залежить від значень іншої (інших) ВВ, то кажуть, що такі ВВ пов'язані **кореляційною залежністю**. Розділ МСА, в якому вивчають методи дослідження кореляційних залежностей між ВВ, називають **кореляційним аналізом** (**КА**). У разі системи двох ВВ ξ і η КА називають **парним**. КА відповідає на запитання: чи існує зв'язок між досліджуваними ВВ і наскільки він тісний (у тому чи іншому розумінні). КА не передбачає дослідження форми самої залежності. (*Кореляція* – взаємозв'язок.)

Залежність математичного сподівання однієї ВВ від значень іншої (інших) називають **регресійною залежністю**. У випадку (ξ,η) – це функції $\varphi(x)$ і $\psi(y)$. Розділ МСА, в якому вивчають методи дослідження регресійних залежностей, називають **регресійним аналізом** (**РА**). Зокрема, для двох ВВ (ξ,η) – **парним** РА, для ВВ-n $(\xi_1,\xi_2,...,\xi_n)$ – **багатовимірним** (або **множинним**) РА.

Зауваження:

1) усі дослідження в МСА проводять, як і в разі побудови математичної моделі однієї ВВ, вибірковим методом;

2) термін "регресійний" пояснюється тим, що вивчення багатовимірної ВВ зводять, по суті, до вивчення однієї (одновимірної) ВВ (*регресія* – рух назад);

3) теоретичні основи МСА більш-менш розроблені для нормально розподілених ВВ.

Статистичне оцінювання парної кореляції (нормальний випадок)

Нехай ξ і η – нормально розподілені ВВ з параметрами a_ξ, σ_ξ і a_η, σ_η відповідно.

Згідно з положеннями ТЙ мірою ступеня зв'язку (тісноти) між двома ВВ є коефіцієнт кореляції:

$$\rho = \rho_{\xi\eta} = \frac{M(\xi \cdot \eta) - M(\xi) \cdot M(\eta)}{\sigma_\xi \cdot \sigma_\eta}, \tag{1.58}$$

який володіє *властивостями*:

1) $|\rho| \le 1$;

2) $|\rho| = 1 \Leftrightarrow \eta = a\xi + b$ $(\xi = c\eta + d)$; $a, b, c, d - const$;

3) (ξ, η – незалежні ВВ) $\Rightarrow \rho = 0$ (ВВ ξ, η – **некорельовані**);

4) $\rho = 0 \not\Rightarrow$ (ξ, η – незалежні ВВ). Для нормально розподілених ВВ: $\rho = 0 \Leftrightarrow$ (ξ, η – незалежні ВВ).

Це означає, що $\rho = \rho_{\xi\eta}$ цілком може слугувати мірою ступеня зв'язку між ξ і η: чим більше $|\rho|$ (ближче до 1), тим "тісніше" зв'язані між собою ВВ, і навпаки (табл. 1.2 – таблиця Чеддока).

Таблиця 1.2

Модуль коефіцієнта кореляції та тіснота зв'язку

Модуль коефіцієнта кореляції	Тіснота зв'язку
1,00	зв'язок функціональний
0,90 – 0,99	дуже сильний
0,70 – 0,89	сильний
0,50 – 0,69	значний
0,30 – 0,49	помірний
0,10 – 0,29	слабкий
0,00	зв'язок відсутній

Нехай із нормально розподіленої г/с добувається репрезентативна вибірка обсягу n, тобто маємо набір пар (x_i, y_i), $i = \overline{1, n}$.

1. *Первинна обробка* вибірки для визначення емпіричного коефіцієнта кореляції $r = r_{xy}$ здійснюється за відомими оцінками для математичного сподівання та дисперсії, а саме:

$$\bar{x} = \frac{1}{n}\sum x_i, \quad \bar{y} = \frac{1}{n}\sum y_i, \quad \overline{x \cdot y} = \frac{1}{n}\sum x_i y_i, \text{ де } \sum = \sum_{i=1}^{n}; \tag{1.59}$$

$$s_x^2 = \frac{1}{n}\sum(x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2, \quad s_y^2 = \frac{1}{n}\sum(y_i - \bar{y})^2 = \overline{y^2} - \bar{y}^2. \tag{1.60}$$

2. *Оцінка* $\rho = \rho_{\xi\eta}$ – вибірковий (емпіричний) коефіцієнт кореляції r – згідно з (1.58) визначається за формулою:

$$r = r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{s_x \cdot s_y}, \tag{1.61}$$

3. *Перевірка гіпотези* H_0 про відсутність кореляційного зв'язку. Беруть H_0: $\rho = 0$, тобто припускають, що кореляційний зв'язок між ξ і η відсутній. У якості альтернативної буде гіпотеза H_a: $\rho \neq 0$.

Для реалізації перевірки H_0 знадобляться: α – рівень значущості, $p = 1 - \alpha$ – надійна ймовірність (див. п. 1.8, критерій χ^2), $\Phi(x)$ – функція Лапласа.

Спочатку визначається так звана **інтервальна оцінка**, яка будується за допомогою знайденого r і має вигляд:

$$(r - \varepsilon, r + \varepsilon), \text{ де } \varepsilon = \frac{t_p(1 - r^2)}{\sqrt{n}}, \tag{1.62}$$

а величина t_p визначається співвідношеннями:

$$p = 2\Phi(t_p) \;\Rightarrow\; t_p = \Phi^{-1}(p/2). \tag{1.63}$$

У ролі критерію виступає статистичний коефіцієнт кореляції r.

Якщо інтервал $(r - \varepsilon, r + \varepsilon)$ містить у собі 0, то початкова гіпотеза H_0 про відсутність кореляційного зв'язку приймається, в протилежному випадку H_0 відхиляється:

$$\begin{aligned} &0 \in (r - \varepsilon, r + \varepsilon) \;\Rightarrow\; H_0 \text{ приймається;} \\ &0 \notin (r - \varepsilon, r + \varepsilon) \;\Rightarrow\; H_0 \text{ відхиляється.} \end{aligned} \tag{1.64}$$

Звичайно, це здійснюється за певним рівнем значущості $\alpha = 1 - p$.

Приклад 1.27. За заданою вибіркою провести кореляційний аналіз відповідної генеральної сукупності:

ξ \ η	5	10	15
2	5	4	
3	7	15	6
4		9	4

Розв'язання. Здійснюємо первинну обробку вибірки згідно з формулами (1.59) – (1.61), для чого підраховуємо відповідні суми. Для полегшення арифметичних обчислень перейдемо від змінних x і y до умовних змінних-варіант u і v за співвідношеннями:

$$u = \frac{x - C_x}{h_x}, \quad v = \frac{y - C_y}{h_y}, \tag{1.65}$$

де C_x (C_y) – варіанта x (y) з найбільшою частотою для ВВ ξ (ВВ η);

h_x, h_y – кроки, з якими наведені значення ξ, η.

Для заданої вибірки: $C_x = 3$, $C_y = 10$; $h_x = 1$, $h_y = 5$. Тоді $u = x - 3$, $v = (y - 10)/5$. Результати обробки вибірки поміщені в табл. 1.3.

Таблиця 1.3

Обчислення сум для підрахунку емпіричних числових характеристик

ξ \ η		5	10	15	m_i	$m_i u_i$	$m_i u_i^2$
	u \ v	–1	0	1			
2	–1	5	4		9	–9	9
3	0	7	15	6	28	0	0
4	1		9	4	13	13	13
m_j		12	28	10	$\sum_m = 50$	$\sum_1 = 4$	$\sum_2 = 22$
$m_j v_j$		–12	0	10	$\sum_3 = -2$		
$m_j v_j^2$		12	0	10	$\sum_4 = 22$		
$\sum_{i,j} m_{ij} u_i v_j$		5	0	4	$\sum_5 = 9$		

За табл. 1.3 знаходимо:

$$\bar{u} = \frac{1}{50}\sum m_i u_i = \frac{1}{50}\sum\nolimits_1 = \frac{1}{50}4 = 0{,}08;$$

$$\bar{v} = \frac{1}{50}\sum m_j v_j = \frac{1}{50}\sum\nolimits_3 = \frac{1}{50}(-2) = -0{,}04;$$

$$\overline{u^2} = \frac{1}{50}\sum m_i u_i^2 = \frac{1}{50}\sum\nolimits_2 = \frac{1}{50}22 = 0{,}44;$$

$$\overline{v^2} = \frac{1}{50}\sum m_j v_j^2 = \frac{1}{50}\sum\nolimits_4 = \frac{1}{50}22 = 0{,}44;$$

$$\overline{u \cdot v} = \frac{1}{50}\sum_{i,j} m_{ij} u_i v_j = \frac{1}{50}\sum\nolimits_5 = \frac{1}{50}9 = 0{,}18.$$

Тепер можна підрахувати емпіричні дисперсії, середні квадратичні відхилення та коефіцієнт кореляції (s_u^2, s_v^2, s_u, s_v, r_{uv}):

$$s_u^2 = \overline{u^2} - \bar{u}^2 = 0{,}4336, \; s_u = \sqrt{s_u^2} = \sqrt{0{,}4336} \approx 0{,}6585;$$

$$s_v^2 = \overline{v^2} - \bar{v}^2 = 0{,}4384, \; s_v = \sqrt{s_v^2} = \sqrt{0{,}4384} \approx 0{,}6621;$$

$$r_{uv} = \frac{\overline{u \cdot v} - \bar{u} \cdot \bar{v}}{s_u \cdot s_v} = \frac{0{,}18 - 0{,}08 \cdot (-0{,}04)}{\sqrt{0{,}4336} \cdot \sqrt{0{,}4384}} = \frac{0{,}1832}{0{,}4360} = 0{,}4202.$$

Зважаючи на те, що $r_{uv} = r_{xy}$, немає потреби обчислювати r_{xy}.

За табл. 1.2 встановлюємо, що $r = r_{uv} = r_{xy} = 0{,}420$ входить у проміжок $0{,}30 - 0{,}49$, тому за вибіркою зв'язок між ξ і η *помірний.*

Далі перевіримо, чи приймається гіпотеза H_0: $\rho = 0$, беручи рівень значущості $\alpha = 0{,}05$; тоді рівень довіри $p = 1 - \alpha = 0{,}95$. За таблицею значень функції Лапласа знаходимо t_p, потім ε та інтервал довіри, використовуючи формули (1.62) – (1.64):

$$t_p = \Phi^{-1}(p/2) = \Phi^{-1}(0{,}475) = 1{,}96;$$

$$\varepsilon = t_p \cdot (1 - r^2)/\sqrt{n} = 1{,}96 \cdot (1 - 0{,}42^2)/\sqrt{50} \approx 0{,}228;$$

$$(r - \varepsilon, r + \varepsilon) = (0{,}420 - 0{,}228; 0{,}420 + 0{,}228) = (0{,}192; 0{,}648).$$

Як бачимо, $0 \notin (r - \varepsilon, r + \varepsilon)$, тому гіпотеза про відсутність кореляційного зв'язку відхиляється; зв'язок існує, але *помірний.*

Далі здійснюємо перехід від умовних характеристик до фактичних емпіричних центрів розподілу та дисперсій за формулами:

$$\bar{x} = \bar{u} \cdot h_x + C_x, \quad \bar{y} = \bar{v} \cdot h_y + C_y. \tag{1.66}$$

$$s_x^2 = s_u^2 \cdot h_x^2, \quad s_y^2 = s_v^2 \cdot h_y^2. \tag{1.67}$$

За знайденими $\bar{u}$, $\bar{v}$, s_u^2, s_v^2 обчислюємо $\bar{x}$, $\bar{y}$, s_x^2, s_y^2, s_x, s_y:

$$\begin{aligned} &\bar{x} = 0{,}08 \cdot 1 + 3 = 3{,}08; && \bar{y} = -0{,}04 \cdot 5 + 10 = 9{,}80; \\ &s_x^2 = 0{,}4336 \cdot 1^2 = 0{,}4336; && s_y^2 = 0{,}4384 \cdot 5^2 = 10{,}96; \\ &s_x = \sqrt{0{,}4336} \approx 0{,}6585; && s_y = \sqrt{10{,}96} \approx 3{,}3105. \end{aligned} \tag{1.68}$$

Отримані результати будуть використані для проведення статистичного регресійного аналізу.

Статистичне оцінювання парної регресії (***нормальний випадок***)

У ТЙ встановлюється, що коли ВВ-2 (ξ, η) має розподіл Гаусса (нормальний розподіл), то його умовні математичні сподівання описуються співвідношеннями:

$$M(\xi \mid \eta = y) = M(\xi / y) = \rho \frac{\sigma_\xi}{\sigma_\eta}(y - a_\eta) + a_\xi, \tag{1.69}$$

$$M(\eta \mid \xi = x) = M(\eta / x) = \rho \frac{\sigma_\eta}{\sigma_\xi}(x - a_\xi) + a_\eta. \tag{1.70}$$

Аналізуючи наведені формули, помічаємо, що умовне математичне сподівання кожної ВВ є лінійною функцією від фіксованих значень другої, і відповідні прямі проходять через точку (a_ξ, a_η) – **центр умовного розподілу**. Рівняння (1.69) і (1.70) називають, відповідно, **рівнянням регресії ВВ ξ на ВВ η** і **рівнянням регресії ВВ η на ВВ ξ**.

Рівняння регресії, як і рівняння прямої, можна подати у вигляді:

$$M(\xi / y) = \alpha_0 + \alpha_1 y, \text{ де } \alpha_0 = a_\xi - \rho \frac{\sigma_\xi}{\sigma_\eta} a_\eta, \ \alpha_1 = \rho \frac{\sigma_\xi}{\sigma_\eta}; \tag{1.71}$$

$$M(\eta / x) = \beta_0 + \beta_1 x, \text{ де } \beta_0 = a_\eta - \rho \frac{\sigma_\eta}{\sigma_\xi} a_\xi, \ \beta_1 = \rho \frac{\sigma_\eta}{\sigma_\xi}. \tag{1.72}$$

Коефіцієнти при y і x – α_1 і β_1 – називають **коефіцієнтами регресії**, відповідно, ξ на η і η на ξ і позначають так: $\rho_{\xi/\eta}$ і $\rho_{\eta/\xi}$, тобто:

$$\rho_{\xi/\eta}=\rho\frac{\sigma_\xi}{\sigma_\eta},\quad \rho_{\eta/\xi}=\rho\frac{\sigma_\eta}{\sigma_\xi}. \tag{1.73}$$

Для побудови рівнянь регресії числові параметри, що входять у рівняння прямої (a_ξ, a_η, σ_ξ, σ_η, ρ), замінюють їх вибірковими значеннями $(\bar{x},\bar{y},s_x,s_y,r)$, а вибіркові математичні сподівання $M(\xi/y)$ і $M(\eta/x)$ позначають, відповідно, через $\bar{x}(y)$ або $\bar{x}_y$ і $\bar{y}(x)$ або $\bar{y}_x$. Отже, рівняння регресій ξ на η і η на ξ, побудованих на основі вибірки, мають вигляд:

$$\bar{x}_y=r\frac{s_x}{s_y}(y-\bar{y})+\bar{x}\text{, або } \bar{x}_y=a_0+a_1y, \tag{1.74}$$

де $a_0=\bar{x}-r\frac{s_x}{s_y}\bar{y}$, $a_1=r\frac{s_x}{s_y}$;

$$\bar{y}_x=r\frac{s_y}{s_x}(x-\bar{x})+\bar{y},\quad \bar{y}_x=b_0+b_1x, \tag{1.75}$$

де $b_0=\bar{y}-r\frac{s_y}{s_x}\bar{x}$, $b_1=r\frac{s_y}{s_x}$.

Рівняння (1.74) і (1.75) дають наближені центри розподілу ВВ ξ (η) для фіксованих значень y (x).

Приклад 1.28. В умовах прикладу 1.27 за вибіркою: а) побудувати емпіричні лінії регресії ξ на η і η на ξ; б) знайти теоретичні рівняння лінії регресії ξ на η і η на ξ, і зобразити їх; в) оцінити за емпіричним коефіцієнтом кореляції тісноту кореляційного зв'язку.

Розв'язання: а) за табл. 1.3 підраховуємо умовні середні значення $\bar{x}_y$, $\bar{y}_x$:

$\bar{x}_{y=5}=(5\cdot 2+7\cdot 3)/12\approx 2{,}6$; $\quad\bar{y}_{x=2}=(5\cdot 5+4\cdot 10)/9\approx 6{,}1$;

$\bar{x}_{y=10}=(4\cdot 2+15\cdot 3+9\cdot 4)/28\approx 3{,}2$; $\quad\bar{y}_{x=3}=(7\cdot 5+15\cdot 10+6\cdot 4)/28\approx 9{,}8$;

$\bar{x}_{y=15}=(6\cdot 3+4\cdot 4)/10=3{,}4$; $\quad\bar{y}_{x=4}=(9\cdot 10+4\cdot 15)/13\approx 11{,}5$.

Наносимо на координатну площину точки $(\overline{x}_y, y_j)$ і $(x_i, \overline{y}_x)$, $i, j = \overline{1,3}$. Ламані (рис. 1.25) з вершинами у цих точках є **емпіричними лініями регресії** ξ на η і η на ξ відповідно;

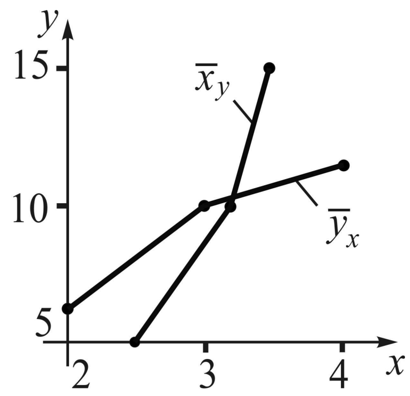

Рис. 1.25. **Емпіричні лінії регресії**

б) теоретичні рівняння ліній регресії отримуємо за формулами (1.74), (1.75) з урахуванням (1.68).

Для регресії $\overline{x}_y = a_0 + a_1 y$:

$$a_0 = \overline{x} - r\frac{s_x}{s_y}\overline{y} = 3{,}08 - 0{,}42 \cdot \frac{0{,}66}{3{,}31} \cdot 9{,}80 = 2{,}26;$$

$$a_1 = r\frac{s_x}{s_y} = 0{,}42 \cdot \frac{0{,}66}{3{,}31} = 0{,}08.$$

Отже, $\overline{x}_y = 0{,}08 \cdot y + 2{,}26$ – *теоретичне рівняння регресії* ξ *на* η.

Для регресії $\overline{y}_x = b_0 + b_1 x$:

$$b_0 = \overline{y} - r\frac{s_y}{s_x}\overline{x} = 9{,}80 - 0{,}42 \cdot \frac{3{,}31}{0{,}66} \cdot 3{,}08 = 3{,}31;$$

$$b_1 = r\frac{s_y}{s_x} = 0{,}42 \cdot \frac{3{,}31}{0{,}66} = 2{,}11.$$

Отже, $\overline{y}_x = 2{,}11 \cdot x + 3{,}31$ – *теоретичне рівняння регресії* η *на* ξ.

Будуємо прямі в одній системі координат. Зважаючи на те, що вони обидві проходять через точку $(\overline{x}, \overline{y})$ – **центр умовного теоретичного розподілу**, – для їх нанесення достатньо вказати ще по одній точці (рис. 1.26).

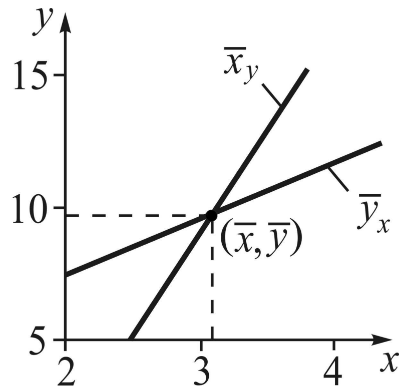

Рис. 1.26. **Теоретичні лінії регресії**

Прямі регресії дають можливість знаходити приблизно вибіркові середні для значень ВВ ξ і η, які не увійшли у вибірку. (*Подумайте*, який висновок слід зробити, якби лінії регресії співпали);

в) ураховуючи результати розв'язання прикладу 1.27, маємо: $r = r_{xy} = 0{,}42$; отже, кореляційний зв'язок між ξ і η *помірний.*

Величина $r^2 \approx 0{,}18$ показує, що близько $18\ \%$ мінливості значень ВВ ξ або η обумовлені лінійною залежністю, а $82\ \%$ пояснюються впливом чинників, зв'язок між якими не є лінійним і які не враховані.

Вибірковий коефіцієнт регресії ξ на η: $a_1 = 0{,}08$, показує, що зі збільшенням значень ВВ ξ на 1 ум. од. значення η зростає на $0{,}08$ ум. од. Вибірковий коефіцієнт регресії η на ξ: $b_1 = 2{,}11$ – показує, що зі збільшенням значень ВВ η на 1 ум. од. значення ξ зростає на $2{,}11$ ум. од.

Висновок: порівнюючи емпіричні та теоретичні лінії регресії, бачимо, що в якісному сенсі теоретичні лінії *досить добре* "вирівнюють" ламані емпіричних регресій, проте, для більш точного опису ступеня зв'язку між ξ і η треба побудувати іншу, нелінійну, регресійну модель.

Зауваження. Для більш глибокого аналізу відшукують інтервальні оцінки параметрів регресії, а також самих регресій. Для проведення КА та РА багатовимірних випадкових векторів зазвичай застосовують пакети прикладних програм.

1.10. Випадкові процеси. Ланцюги Маркова. Системи масового обслуговування

Випадкові (*ймовірнісні*) ***процеси: основні поняття***

Витоком основних понять теорії випадкових процесів є поняття **системи** – множини взаємозв'язаних об'єктів (установок, механізмів, явищ тощо), які розглядаються як єдине ціле. Розрізняють системи: біологічні, виробничі, соціальні, економічні, інформаційні тощо. Математичний підхід до вивчення систем передбачає їх опис сукупністю тих чи інших числових характеристик. Кажуть, що кожний набір числових характеристик визначає **стан системи**. Якщо змінюється хоча б одна характеристика, то це розцінюють як **перехід** в інший стан. Послідовну зміну станів системи називають **процесом** (зміни станів), а перехід із одного стану в інший – **кроком процесу**. За аналогією з різновидами експерименту – витоку основних понять ТЙ – процеси бувають передбачувані та непередбачувані (стохастичні, ймовірнісні).

Нехай $\Omega = \{\omega\}$ – простір елементарних подій (наслідків) деякого стохастичного експерименту E, а $\xi = \xi(\omega)$ – ВВ, задана на цьому просторі, тобто кожному $\omega \in \Omega$ поставлено у відповідність певне число $x \in \mathbf{R} = (-\infty, +\infty)$.

Якщо кожному $\omega \in \Omega$ поставлена у відповідність числова функція $x = x(t)$ від деякого детермінованого (невипадкового) числового параметра $t \in T$ (T – множина значень параметра), то кажуть, що задана **випадкова функція** $\boldsymbol{\xi = \xi(t)}$.

За умови, що в ролі параметра t виступає час, випадкову функцію називають **випадковим процесом** (**ВП**). Залежно від того, якою буде множина T – дискретною чи неперервною, за аналогією з ВВ розрізняють ВП *дискретного типу* (ДВП) і *неперервного типу* (НВП). Множину T називають **областю визначення** (**існування**) ВП, а множину ВВ, яка описує процес, – **фазовим простором**.

Зауваження. Фактично ВП є функцією двох змінних – невипадкового параметра t і випадкової події ω: $\xi = \xi(\omega, t)$, але, як і в позначенні ВВ, символ ω часто пропускають, тобто пишуть ξ замість $\xi(\omega)$, а для ВП замість $\xi = \xi(\omega, t)$ пишуть $\xi(t)$.

ВВ $\xi = \xi(\omega, t_0)$ як значення ВП для фіксованого моменту часу $t = t_0$ називають **перерізом ВП у точці** $\boldsymbol{t_0}$:

$$\xi(\omega, t)\big|_{t=t_0} \text{ – ВП, якщо } t = t_0 \quad \rightleftharpoons \quad \xi_{t_0}(\omega) \text{ – переріз ВП.}$$

Невипадкову функцію $\xi = \xi(\omega_0, t) = \xi_{\omega_0}(t)$, яка отримується із ВП для фіксованого наслідку ω_0, називають **траєкторією** (**реалізацією, вибірковою функцією**) **ВП**:

$$\xi(\omega, t)\big|_{\omega=\omega_0} \text{ – ВП, якщо } \omega = \omega_0 \ \rightleftharpoons \ \xi(\omega, t)\big|_{\omega=\omega_0} = \xi_{\omega_0}(t) \text{ – траєкторія ВП.}$$

Наприклад, якщо вивчати авіаційні польоти (в певному "коридорі"), то можна розглядати ВП, кожний переріз якого є ВВ зі значеннями – висота польоту, а траєкторії ВП співпадають із фізичними траєкторіями польоту (рис. 1.27).

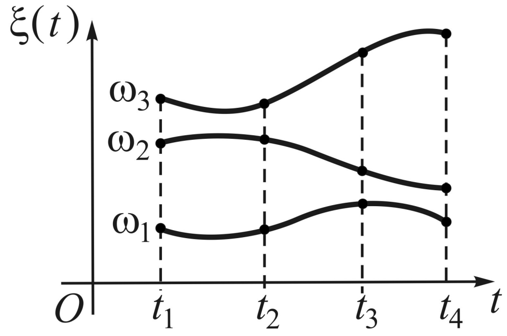

Рис. 1.27. **Перерізи та траєкторії ВП**

Кожне значення t_i $(i=\overline{1,4})$ визначає переріз ВП (рис. 1.27); жирними точками зображені значення відповідних ВВ; траєкторіями ВП є лінії, які проходять через точки всіх перерізів для фіксованого ω і описуються функціями від часу t: $\xi_{\omega_j}(t)$ $(j=\overline{1,3})$.

Необмежену кількість ВП без змістового наповнення (абстрактних) можна отримати з різних типів функції однієї змінної $f(x)$, якщо $x=t$, а числові параметри (константи) тлумачити як ВВ ($const$ – вироджена ВВ):

$$y=ax+b \;\Rightarrow\; \xi=At+B, \text{ де } A,\ B \text{ – ВВ, } t\in T;$$
$$y=\sin kx \;\Rightarrow\; \xi=\sin Kt, \text{ де } K \text{ – ВВ, } t\in T.$$

Зауваження. Для ВП, як і для ВВ, вводяться поняття: функція розподілу, математичне сподівання, дисперсія тощо, але всі вони, на відміну від ВВ, є функціями часу.

Найпростішими є ВП, область визначення й область значень яких – дискретні множини (скінченні або зліченні).

Ланцюги Маркова: означення, матриця перехідних імовірностей, граничні ймовірності

Ланцюг Маркова є прикладом ВП з дискретними множинами станів Ω і області визначення T. Нехай $\{s_1, s_2, ..., s_m\}$ – множина всіх можливих станів деякої системи. З часом система переходить послідовно з одного стану в інший. У будь-який момент часу вона може знаходитись тільки в одному стані. Для опису еволюції системи розглядають послідовність ВВ ξ_0, ξ_1, …, ξ_n, …. Індекс n відіграє роль часу. Якщо в момент часу n система знаходиться у стані s_j, то пишуть $\xi_n=j$. Послідовність ξ_0, ξ_1, …, ξ_n, … називають **ланцюгом Маркова** (**ЛМ**), якщо для будь-якого n і будь-якого i_0, i_1, …, i_{n-2}, і $i\in\{1,2,...,m\}$, маємо:

$$P(\xi_n=j\,|\,\xi_0=i_0,\xi_1=i_1,...,\xi_{n-2}=i_{n-2},\xi_{n-1}=i)=P(\xi_n=j\,|\,\xi_{n-1}=i), \qquad (1.76)$$

тобто ймовірність в момент часу n потрапити в стан j, якщо відома вся передісторія процесу, залежить лише від того, в якому стані система знаходилася в момент $(n-1)$.

Інакше кажучи, "майбутнє" ($\xi_n = j$) з фіксованим "теперішнім" ($\xi_{n-1}=i$) не залежить від "минулого" $(\xi_0 = i_0,...,\xi_{n-2} = i_{n-2})$ (рис. 1.28).

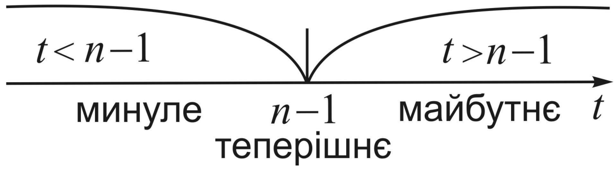

Рис. 1.28. **Марківська властивість**

Імовірності

$$p_{ij}^{(n)} = P(\xi_n = j \mid \xi_{n-1} = i) \tag{1.77}$$

називають **імовірностями переходу** (**перехідними ймовірностями**) системи зі стану i в стан j за один крок (на n-му кроці).

ЛМ називають **однорідною,** якщо ймовірності переходу $p_{ij}^{(n)}$ не залежать від n (номера кроку), а залежать лише від того, з якого стану в який здійснюється перехід:

$$p_{ij}^{(n)} = P(\xi_n = j \mid \xi_{n-1} = i) = p_{ij} - const \;\; \forall n \in \mathbf{N}. \tag{1.78}$$

Однорідні ЛМ повністю описують за допомогою:

1) **вектора** ймовірностей початкового стану: $a = (a_1, a_2, ..., a_m)$;

2) **матриці** ймовірностей переходу за один крок:

$$P = (p_{ij})_{m\times m} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}, \;\; p_{ij} \geq 0, \;\; \sum_{j=1}^{m} p_{ij} = 1 \;\; \forall i \in \{1, 2, \; ..., \; m\}. \tag{1.79}$$

Матриці, які володіють властивостями: $p_{ij} \geq 0$, $\sum\limits_{j=1}^{m} p_{ij} = 1$, називають **стохастичними**.

Для наочності замість матричного опису ЛМ застосовують геометричне – за допомогою графів, вершини яких – стани ланцюга, а орієнтовані ребра (стрілки) від s_i до s_j з числом p_{ij} над ними показують, що перехід із стану i в стан j можливий з імовірністю p_{ij}; якщо $p_{ij} = 0$, то стрілку не проводять.

Імовірності

$$p_{ij}(n) = P(\xi_n = j \mid \xi_0 = i) \qquad (1.80)$$

називають **імовірностями переходу** зі стану i в стан j **за n кроків**.

Матриця ймовірностей переходу за n кроків – $P(n)$ – є n-м степенем матриці переходу за один крок (1.79):

$$P(n) = P^n = \begin{pmatrix} p_{11}(n) & p_{12}(n) & \dots & p_{1m}(n) \\ \dots & \dots & \dots & \dots \\ p_{m1}(n) & p_{m2}(n) & \dots & p_{mm}(n) \end{pmatrix}. \qquad (1.81)$$

Приклад 1.29. Процес міграції населення "село – місто" описується ЛМ з двома станами: $s_1 =$ (селянин), $s_2 =$ (городянин), з вектором імовірностей початкового стану a та матрицею перехідних імовірностей P:

$$a = (1/2,\ 1/2), \quad P = \begin{array}{c} \\ s_1 \\ s_2 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{bmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{bmatrix} \end{array}.$$

Побудувати граф ЛМ і знайти матрицю ймовірностей переходу за два кроки.

Розв'язання. Вибираємо на площині довільним чином дві точки (кружечка) – вершини графа, з'єднуємо їх орієнтованими ребрами та наносимо відповідні написи (рис. 1.29).

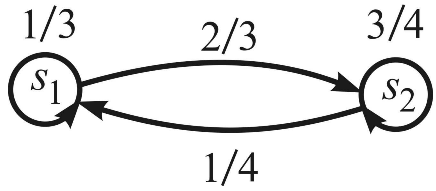

Рис. 1.29. **Граф процесу міграції**

Знаходимо квадрат матриці P:

$$P(2) = \begin{bmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{bmatrix} \cdot \begin{bmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{bmatrix} = \begin{bmatrix} 15/54 & 39/54 \\ 13/48 & 35/48 \end{bmatrix} = \begin{bmatrix} 0{,}2(7) & 0{,}7(2) \\ 0{,}2708(3) & 0{,}7291(6) \end{bmatrix}.$$

Порівнюючи матрицю $P(2)$ з вихідною матрицею P, можна простежити зміни у процесі міграції. *Наприклад,* оскільки $p_{12} = 2/3 = 0{,}6(6)$, а $p_{12}(2) = 0{,}7(2)$, то це означає, що слід чекати збільшення відтоку селян у міста. *Проаналізуйте* самостійно "поведінку" городян.

Теорема 1.7. (*про граничні ймовірності*). Якщо для деякого n_0 усі елементи матриці $P(n_0)$ додатні: $p_{ij}(n_0) > 0$, то для кожного $j = 0,\ 1,\ ...,\ m$ (незалежно від значення $i = 0,\ 1,\ ...,\ m$) існують границі:

$$\forall i = 0,\ 1,\ ...,\ m: \lim_{n\to\infty} p_{ij}(n) = b_j,\ \ j = 0,\ 1,\ ...,\ m, \qquad (1.82)$$

тобто ймовірність перебування системи у стані j практично не залежить від того, в якому стані вона знаходилась у "далекому минулому".

Граничні ймовірності $\boldsymbol{b_j}$ – єдиний розв'язок системи лінійних алгебраїчних рівнянь (як наслідок з (1.81), (1.82)):

$$\begin{cases} \sum\limits_{j=1}^{m} b_j = 1, \\ \sum\limits_{i=1}^{m} b_i p_{ij} = b_j, \end{cases} \quad j = 0,\ 1,\ ...,m;\ \ p_{ij} \in P. \qquad (1.83)$$

Вектор $b = (b_1, b_2, ..., b_m)$ граничних імовірностей називають **стаціонарним розподілом** імовірностей для ЛМ з $P = (p_{ij})_{m\times m}$.

Зауваження. Оскільки в системі (1.83) кількість рівнянь більша від кількості невідомих, то в процесі її розв'язання одне з рівнянь, крім першого (*чому?*), можна відкинути (краще те, що з меншими p_{ij}).

Приклад 1.30. В умовах прикладу 1.29 знайти граничні (фінальні) ймовірності – стаціонарний розподіл для процесу міграції.

Розв'язання. Складаємо систему (1.83) і розв'язуємо її:

$$m=2: \begin{cases} b_1 + b_2 = 1, \\ b_1 p_{11} + b_2 p_{21} = b_1, \\ b_1 p_{12} + b_2 p_{22} = b_2 \end{cases} \Rightarrow \begin{cases} b_1 = \dfrac{p_{21}}{p_{21} + p_{12}}, \\ b_2 = \dfrac{p_{12}}{p_{12} + p_{21}} \end{cases} \Rightarrow$$

$$\Rightarrow \left| P = \begin{bmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{bmatrix} \right| \Rightarrow (b_1,\ b_2) = \left(\frac{3}{11},\ \frac{8}{11} \right).$$

Висновок. За умови тривалого процесу міграції слід очікувати, що кількість селян і городян розподіляться у відношенні 3/8.

Поняття про системи масового обслуговування

Більш складними для вивчення, порівняно з ЛМ, є ВП з дискретною множиною станів і неперервною множиною визначення, оскільки їх закони описуються, як правило, системами диференціальних рівнянь. Одним із різновидів таких ВП є **процес народження (розмноження) і загибелі (ПНЗ)**, граф станів якого подано на рис. 1.30.

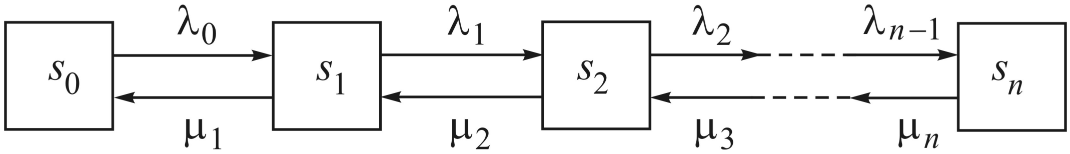

Рис. 1.30. **Граф станів процесу "розмноження і загибелі"**

Теорія ПНЗ веде початок від біологічних задач (від цього походить назва), але в застосуваннях виходить за межі біології: такими ВП моделюють технічні, економічні, фізичні системи, системи **черг**, або **системи масового обслуговування** (**СМО**).

На рисунку 1.30 λ_i $(i=\overline{0,n-1})$ – інтенсивність потоку заявок (величина, обернена до середнього проміжку часу $\overline{(t_з)}_i$ між надходженнями заявок); μ_j, $(j=\overline{1,n})$ – інтенсивність потоку обслуговувань (величина, обернена до середнього проміжку часу $(\overline{t_{об}})_j$ обслуговування заявок).

Прикладами СМО є: каси, аеропорти, вокзали, доки, перукарні, ринки тощо. У більшості випадків основні **вхідні** (**задані**) **параметри** СМО – це кількість каналів (наприклад, крісел у перукарні), інтенсивність потоку заявок (кількість клієнтів), інтенсивність потоку обслуговування (кількість "покрасивілих").

Характеристиками ефективності роботи СМО, як правило, є так звані усереднені характеристики: пропускна спроможність, середня кількість зайнятих каналів, відмов і таке інше. Звичайно, теоретичною основою СМО є теорія ймовірностей, у якій розроблені відповідні закони розподілів.

2. Контрольні запитання для самодіагностики

1. Що є об'єктом, предметом та основними завданнями ТЙ?

2. Які події називають випадковими, на які класи вони поділяються?

3. Які існують операції над випадковими подіями?

4. У чому полягає класичний підхід до означення ймовірності події, які її основні властивості?

5. Що таке міра множини? У чому полягає геометричний підхід до означення ймовірності?

6. Що називають відносною частотою настання події (частістю)? У чому полягає статистичне означення ймовірності?

7. Як формулюється теорема про ймовірність суми сумісних (несумісних) подій?

8. Як формулюється теорема про ймовірність добутку залежних (незалежних) подій?

9. Який вигляд має формула повної ймовірності та формула Байєса?

10. Що розуміють під схемою незалежних випробувань (схемою Бернуллі)? Як виглядають формули обчислення ймовірності певного числа успіхів: $P_n(k)$, $P_n(k_1, k_2)$?

11. За яким співвідношенням визначається найімовірніше число успіхів у схемі Бернуллі?

12. Який вигляд має локальна (інтегральна) формула Муавра – Лапласа?

13. Яку назву мають функції $\varphi(x)$, $\Phi(x)$ та які їхні властивості?

14. Який вигляд має формула Пуассона?

15. Які випадкові величини називають дискретними (ДВВ)? Що таке: закон розподілу ДВВ і способи його задання, функція розподілу та її властивості? Як обчислюються ймовірності на інтервалі, півсегменті та сегменті?

16. Які випадкові величини називають неперервними (НВВ)? Що таке: щільність розподілу (її властивості), функція розподілу та її властивості? Який існує зв'язок між функцією та щільністю розподілу?

17. Як обчислюють імовірність $P\left(x_1 < \xi < x_2\right)$ за допомогою інтегральної (диференціальної) функції розподілу?

18. Як виглядають закони розподілу дискретних ВВ: біноміальний, Пуассона, геометричний, гіпергеометричний?

19. Як виглядають закони розподілу неперервних ВВ: рівномірний, нормальний, показниковий?

20. Що таке математичне сподівання, дисперсія та середнє квадратичне відхилення дискретної (неперервної) ВВ? Які властивості вони мають?

21. Як визначаються мода та медіана ДВВ (НВВ)?

22. За якими формулами обчислюють числові характеристики основних розподілів ДВВ: біноміального, геометричного, гіпергеометричного, пуассонівського?

23. За якими формулами обчислюють числові характеристики основних розподілів НВВ: рівномірного, показникового, нормального?

24. Яке правило називають "правилом трьох сигм"?

25. Що називають масовим явищем, статистичним матеріалом?

26. Що є метою, об'єктом, предметом та основними задачами МС?

27. Що таке "генеральна сукупність", "вибіркова сукупність", "випадкова вибірка", "повторна" та "безповторна" випадкові вибірки, "репрезентативна вибірка"?

28. У чому полягає вибірковий метод побудови моделі масового явища?

29. Як означають дискретні й інтервальні статистичні ряди за частотами та частостями?

30. Що розуміють під полігоном, кумулятою, гістограмою емпіричного розподілу?

31. Що таке "щільність" емпіричного розподілу?

32. Як формулюється постановка задачі статистичного оцінювання невідомих параметрів розподілу генеральної сукупності?

33. Як формулюється постановка задачі перевірки гіпотез про закон розподілу генеральної сукупності?

34. Що таке "нульова гіпотеза", "альтернативна гіпотеза"?

35. Що називають статистичним критерієм згоди між емпіричним і теоретичним розподілами?

36. Що розуміють під рівнем значущості критерію?

37. Як здійснюється перевірка гіпотез про шуканий закон розподілу за критерієм Пірсона?

38. Що називають системою випадкових величин (випадковим вектором, багатовимірною ВВ)?

39. У чому полягає суть кореляційного аналізу?

40. За допомогою чого визначають вплив факторів (незалежних змінних) на вислідний показник (залежну змінну)?

41. У чому полягає суть регресійного аналізу?

42. Який порядок реалізації парного регресійного аналізу?

43. Що називають системою, станом системи, процесом, кроком процесу?

44. Які процеси називають випадковими процесами (ВП)?

45. Що розуміють під областю визначення ВП?

46. У чому полягає принципова відмінність між ВВ і ВП?

47. Що називають фазовим простором, перерізом, траєкторією ВП?

48. Яку послідовність випробувань називають ланцюгом Маркова?

49. Що розуміють під вектором початкових імовірностей ЛМ?

50. Що називають матрицею переходу ЛМ?

51. Що характеризує n-й степінь матриці переходу ЛМ?

52. Які ймовірності ЛМ називають граничними (фінальними)?

53. До розв'язання якої системи лінійних алгебраїчних рівнянь зводиться знаходження граничних імовірностей?

54. Що є метою, об'єктом, предметом та основними завданнями теорії масового обслуговування?

3. Завдання контрольної роботи

Задача 1. Розв'язати задачу, використовуючи класичне означення ймовірності події.

1.0. Серед сорока підручників, що повернули до бібліотеки, п'ять мають пошкодження. Знайти ймовірність того, що серед шести відібраних навмання підручників чотири будуть без пошкоджень.

1.1. У книжковій лотереї сто квитків, причому на двадцять п'ять із них припадають виграші. Учасник купує чотири квитки. Знайти ймовірність виграшу за двома квитками лотереї.

1.2. Для розробки нової моделі виробу із групи з дев'яти інженерів і шести маркетологів навмання вибирають п'ять осіб. З якою ймовірністю можна створити команду у складі двох інженерів і трьох маркетологів?

1.3. Упаковка містить двадцять керамічних плиток, з яких три мають дефекти. Контролер навмання витягає чотири плитки. Знайти ймовірність того, що упаковка буде прийнята контролером, якщо для цього потрібно, щоб він не знайшов жодної бракованої плитки.

1.4. Чотири білети в театр розігруються випадковим чином серед п'яти юнаків і семи дівчат. Знайти ймовірність того, що білети отримають два юнака і дві дівчини.

1.5. Експерт з управління цінними паперами розглядає двадцять об'єктів для інвестування. Знайти ймовірність того, що серед чотирьох навмання вибраних для інвестування об'єктів опиниться об'єкт під номером 8.

1.6. До магазину надійшло п'ятнадцять телевізорів, з яких десять виготовлені на харківському заводі, а решта – на львівському. Знайти ймовірність того, що серед п'яти навмання вибраних телевізорів виявлять три телевізори, виготовлені на харківському заводі.

1.7. За підсумком року акції десяти фірм стали прибутковими, чотирьох – знецінились, а шести – зберегли свою номінальну вартість. Яка ймовірність того, що серед семи навмання придбаних акцій різних фірм три матимуть прибуток?

1.8. Протягом місяця банк мав видати в кредит позику дванадцяти приватним підприємцям і восьми держслужбовцям. Ця операція здійснюється поетапно. Знайти ймовірність того, що за перший тиждень кредити отримають два приватні підприємця та три держслужбовця, якщо всі клієнти мають однакові можливості отримати позику.

1.9. Вісім сторінок газети містять рекламні оголошення, сім сторінок присвячені соціально-політичним проблемам, три – спортивним новинам. Чотири сторінки цієї газети були втрачені. Знайти ймовірність того, що серед них не було сторінок зі спортивними новинами.

Задача 2. Розв'язати задачу, використовуючи геометричне означення ймовірності події.

2.0. У прямокутнику зі сторонами 6 см і 20 см розташовані два круги діаметром 3 см кожний, які не перетинаються. Знайти ймовірність того, що навмання вибрана точка прямокутника не належить жодному з цих кругів.

2.1. У коло радіусом R вписано правильний трикутник. Усередину кола кидають точку. Знайти ймовірність того, що вона потрапить усередину трикутника.

2.2. У квадраті зі стороною a навмання вибирають точку. Знайти ймовірність того, що ця точка належить вписаному в цей квадрат кругу.

2.3. Усередині круга радіусом R навмання вибирають точку. Знайти ймовірність того, що ця точка належить вписаному в цей круг квадрату.

2.4. У коло радіусом 10 см кидають точку. Знайти ймовірність того, що відстань від цієї точки до центра кола не перевищує 4 см.

2.5. Навмання взято два додатні числа x і y, кожне з яких не перевищує 2. Знайти ймовірність того, що сума цих чисел буде не більшою за 2, а частка y/x – не меншою за 1.

2.6. Точку навмання кидають у квадрат, сторона якого дорівнює 1 см. Знайти ймовірність того, що відстань від цієї точки до найближчої сторони квадрата не більша за $0{,}25$ см.

2.7. У квадраті $ABCD$ навмання вибирають точку. Знайти ймовірність того, що ця точка належить трапеції $AMCD$, де точка M – середина сторони BC.

2.8. Середини сторін прямокутника $ABCD$ є вершинами чотирикутника $MNPK$. Периметр прямокутника дорівнює 40 см, а одна сторона в три рази більша за іншу. З прямокутника навмання вибирають точку. Знайти ймовірність того, що вона належить чотирикутнику $MNPK$.

2.9. На відрізку $[-1;1]$ навмання беруть два числа. Знайти ймовірність того, що сума квадратів цих чисел буде не більшою за 1.

Задача 3. Розв'язати задачу, використовуючи теореми про ймовірність суми та ймовірність добутку подій.

3.0. Для перевірки партії з п'ятнадцяти виробів, серед яких п'ять бракованих, навмання вибирають три вироби. Партія вважається забракованою, якщо в ній виявиться хоча б один бракований виріб. Знайти ймовірність того, що вся партія буде забракована.

3.1. Ймовірність своєчасної сплати податків для першого підприємства дорівнює $0{,}8$, для другого – $0{,}7$, для третього – $0{,}5$. Знайти ймовірність того, що тільки одне підприємство своєчасно сплатить податки.

3.2. У туристів залишилося десять однакових консервних банок, серед яких шість – із тушкованим м'ясом, решта – з рибою. Під час зливи етикетки відклеїлися. Знайти ймовірність того, що вміст двох навмання взятих банок виявиться різним.

3.3. Токар поклав у коробку чотири нові різці та, розпочинаючи роботу, щоразу бере навмання один із них, а після роботи знову кладе його в коробку. Знайти ймовірність того, що через чотири зміни всі різці в коробці будуть використані, якщо протягом зміни токар користується лише одним різцем.

3.4. Імовірність банкрутства для першої фірми є розв'язком рівняння $5p^2 - 11p + 2 = 0$, а для другої фірми ця ймовірність на $20\ \%$ більша. Знайти ймовірність того, що хоча б одна з цих фірм збанкрутує.

3.5. Підприємець проводить переговори з трьома постачальниками. Імовірність укладання вигідної угоди з першим постачальником становить $0{,}6$, з другим – $0{,}7$, а з третім – $0{,}9$. Знайти ймовірність того, що підприємець укладе угоду тільки з двома постачальниками.

3.6. Три програмісти тестують нову програму, в якій є помилка. Імовірність знайти помилку для кожного з програмістів дорівнює, відповідно, $0{,}8$; $0{,}9$; $0{,}7$. Знайти ймовірність того, що помилка буде знайдена.

3.7. В аудиторії двадцять комп'ютерів, на п'ятнадцяти з яких встановлене нове програмне забезпечення. Знайти ймовірність того, що з трьох навмання вибраних комп'ютерів нове програмне забезпечення буде встановлене тільки на двох із них.

3.8. Підприємство отримує сировину від трьох постачальників і не виконує контракт на виготовлення продукції, якщо хоча б один із постачальників зриває постачання сировини. Імовірність вчасного постачання сировини для першого постачальника дорівнює $0{,}97$, для другого – $0{,}95$, для третього – $0{,}99$. Знайти ймовірність виконання контракту підприємством-виробником.

3.9. Імовірність повного розрахунку за енергоносії для першого підприємства є коренем рівняння $4 - p = 5p^2$, а для другого – на $10\ \%$ менша. Знайти ймовірність того, що повністю за енергоносії розрахується тільки одне підприємство.

Задача 4. Розв'язати задачу, використовуючи формулу повної ймовірності та формулу Байєса.

4.0. Виробник комп'ютерів отримує комплектуючі деталі від трьох постачальників, частки яких становлять відповідно 20, 45 і $35\ \%$. Деталі першого постачальника мають $2\ \%$ браку, другого – $1{,}5\ \%$, третього – $1{,}7\ \%$. Знайти ймовірність того, що: 1) навмання вибрана деталь буде з браком; 2) бракована деталь отримана від другого постачальника.

4.1. Податкові інспектори перевіряють діяльність підприємств: перший обслуговує 40 підприємств, серед яких $25\,\%$ не мають заборгованостей, другий – 60 підприємств, серед яких $40\,\%$ – без заборгованостей. Знайти ймовірність того, що: 1) навмання вибране підприємство не має заборгованості; 2) підприємство, що не має заборгованості, перевіряв перший інспектор.

4.2. Виробництво певної продукції може проводитись у трьох температурних режимах із ймовірностями $0{,}45$; $0{,}25$; $0{,}3$ відповідно. Залежно від температурного режиму ймовірність отримання продукції вищої якості становить $0{,}8$; $0{,}85$; $0{,}9$. Знайти ймовірність того, що: 1) виготовлена продукція вищої якості; 2) продукцію вищої якості виготовлено при третьому температурному режимі.

4.3. До каси підприємства надійшли банкноти в пачках від двох банків: 50 пачок від першого і 70 – від другого. Імовірність помилки касирів першого банку становить $0{,}15\,\%$, другого – $0{,}2\,\%$. Знайти ймовірність того, що: 1) навмання вибрану пачку сформовано без помилок; 2) пачку без помилок було сформовано касирами другого банку.

4.4. Два цехи заводу випускають кухонні набори білого та синього кольорів. Перший цех виробляє $35\,\%$ продукції, серед яких $40\,\%$ наборів синього кольору. У продукції другого цеху $55\,\%$ синіх наборів. Знайти ймовірність того, що: 1) навмання вибраний набір синього кольору; 2) набір синього кольору зроблено другим цехом.

4.5. Партію виготовлених деталей перевіряли два контролери. Перший перевірив 45 %, а другий – 55 % деталей. Імовірність припуститися помилки під час перевірки для першого контролера становить $0{,}15$, для другого – $0{,}1$. Знайти ймовірність того, що: 1) навмання вибрану деталь під час перевірки помилково визнано стандартною; 2) деталь, яку помилково визнано стандартною, перевіряв другий контролер.

4.6. Магазин отримує продукцію від двох виробників: перший постачає $2/5$ усіх виробів, другий – $3/5$. Імовірність продажу виробів першого постачальника становить $0{,}95$, другого – $0{,}8$. Знайти ймовірність того, що: 1) навмання вибраний виріб не буде реалізовано; 2) нереалізований виріб отримано від першого виробника.

4.7. У рекламному агентстві працює три групи дизайнерів: перша обслуговує 25 фірм, друга – 45, третя – 40. Протягом одного місяця кошти, витрачені на рекламу дизайнерами першої групи, повертаються до $40\,\%$ фірм, другої – до $45\,\%$, третьої – до $35\,\%$. Знайти ймовірність того, що: 1) навмання вибрана фірма окупила витрачені на рекламу кошти протягом місяця; 2) фірма, що окупила протягом місяця витрачені на рекламу кошти, обслуговувалась третьою групою дизайнерів.

4.8. Тираж популярної газети друкується в двох типографіях. Потужності двох типографій відносяться як $3{:}4$, причому перша дає $3{,}5\,\%$ браку, а друга – $2{,}5\,\%$. Знайти ймовірність того, що: 1) навмання вибраний примірник газети буде бракованим; 2) бракований примірник газети надруковано в першій типографії.

4.9. Справи клієнтів банку зберігаються у восьми сейфах: у трьох по 150 справ, у п'яти – по 250. Ймовірність вчасного повернення кредиту клієнтами, справи яких лежать у перших трьох сейфах, становить $0{,}96$, в останніх п'яти – $0{,}95$. Знайти ймовірність того, що: 1) навмання вибрано справу клієнта, який вчасно поверне кредит; 2) справа клієнта, який своєчасно повернув кредит, знаходилась в одному з перших трьох сейфів.

Задача 5. Розв'язати задачу, використовуючи схему незалежних випробувань – схему Бернуллі – або її граничні теореми.

5.0. Відділ доставки піцерії отримує $80\,\%$ замовлень на фірмову піцу. Знайти ймовірність того, що серед 100 отриманих замовлень на фірмову піцу буде: 1) рівно половина; 2) не менше 70.

5.1. Абоненти мобільного зв'язку не отримують відправлені повідомлення з імовірністю $0{,}002$. Знайти ймовірність того, що серед 400 відправлених повідомлень неотриманих буде: 1) рівно три; 2) не більше двох.

5.2. Під час транспортування виробів зі скла пошкоджуються $0{,}3\,\%$. Знайти ймовірність того, що серед 300 відправлених до магазину виробів буде: 1) лише один пошкоджений; 2) хоча б один пошкоджений.

5.3. Серед 500 коробок взуття нової колекції в 400 лежить взуття чорного кольору. Знайти ймовірність того, що серед чотирьох навмання вибраних коробок зі взуттям чорного кольору буде: 1) лише одна; 2) не менше ніж дві.

5.4. Енергетична компанія обслуговує 800 споживачів електроенергії. Перебої в подачі енергії протягом доби виникають з імовірністю $0{,}005$. Знайти ймовірність того, що протягом доби повідомлень про перебої надійде: 1) рівно два; 2) не більше чотирьох, але не менше дев'яти.

5.5. Фірма виконує поліграфічні роботи, причому $20\ \%$ замовлень припадає на виготовлення візитних карток. Знайти ймовірність того, що серед 850 клієнтів візитні картки замовлять: 1) 150 клієнтів; 2) не більше 100 клієнтів.

5.6. Відомо, що три чверті населення міста користується послугами кабельного телебачення. Знайти ймовірність того, що серед 400 мешканців такими послугами користуються: 1) рівно 310 осіб; 2) не менше 280 осіб.

5.7. До агентства нерухомості звертаються з приводу оренди та продажу квартир у співвідношенні $7{:}5$. Знайти ймовірність того, що серед шести довільно вибраних заявок буде: 1) чотири щодо продажу квартир; 2) не менше чотирьох щодо оренди.

5.8. Деяка компанія володіє мережею дилерів на біржі. Імовірність того, що дилер буде грати вдало, становить $0{,}8$. Знайти ймовірність того, що із шести дилерів збитків зазнають: 1) два; 2) хоча б два.

5.9. Локальна мережа складається зі 100 комп'ютерів. Ймовірність виникнення збоїв у роботі протягом доби для кожного з них дорівнює $0{,}002$. Знайти ймовірність того, що протягом доби збої виникнуть у роботі: 1) лише одного комп'ютера; 2) не більше ніж у трьох комп'ютерів.

Задача 6. Для вказаної в умові задачі дискретної випадкової величини ξ: 1) скласти закон розподілу; 2) знайти її математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 3) побудувати полігон розподілу; 4) знайти функцію розподілу $F_\xi(x)$ і побудувати її графік.

6.0. Імовірність укладення угоди між партнерами за результатами ділових переговорів дорівнює $0{,}7$. Розглянути ДВВ ξ, значеннями якої є число укладених угод після трьох ділових зустрічей.

6.1. Імовірність прийняття на роботу для кожного з трьох претендентів дорівнює $0{,}6$. Розглянути ДВВ ξ, значеннями якої є число претендентів, прийнятих на роботу.

6.2. Імовірність потрапити до другого туру виборів для кожного з трьох кандидатів дорівнює $0{,}8$. Розглянути ДВВ ξ, значеннями якої є число кандидатів, які пройшли до другого туру виборів.

6.3. Імовірність підвищення курсу акцій для кожного з трьох підприємств дорівнює $0{,}4$. Розглянути ДВВ ξ, значеннями якої є число підприємств, курс акцій яких підвищився.

6.4. Імовірність того, що виріб не відповідає стандарту, дорівнює $0{,}1$. Перевіряють три вироби. Розглянути ДВВ ξ, значеннями якої є число виробів, що не відповідають стандарту.

6.5. Імовірність того, що витрати електроенергії на підприємстві за один місяць не перевищать норму, дорівнює $0{,}75$. Розглянути ДВВ ξ, значеннями якої є число зимових місяців, протягом яких витрати електроенергії не перевищать норму.

6.6. Прилад містить три елементи. Імовірність виникнення технічних неполадок у кожному з них дорівнює $0{,}2$. Розглянути ДВВ ξ, значеннями якої є число елементів приладу, в яких можуть виникнути технічні неполадки.

6.7. За даними маркетингових досліджень імовірність підвищення попиту на продукцію кожного з трьох видів дорівнює $0{,}3$. Розглянути ДВВ ξ, значеннями якої є число видів продукції, для яких прогнозується підвищення попиту.

6.8. Під час виробництва мікросхем 5% продукції становить брак. Розглянути ДВВ ξ, значеннями якої є число стандартних мікросхем серед трьох, відібраних навмання.

6.9. Встановлено, що 20% приладів виходять з ладу через перепади напруги електромережі. Розглянути ДВВ ξ, значеннями якої є число приладів серед трьох, вибраних навмання, які не вийшли з ладу через перепади напруги.

Задача 7. Треба: а) за заданою функцією розподілу неперервної випадкової величини ξ знайти: 1) щільність розподілу ймовірностей $f(x)$; 2) математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 3) побудувати графіки функцій $f(x)$ і $F(x)$;

б) за заданою щільністю розподілу ймовірностей неперервної випадкової величини ξ визначити, за яким законом вона розподілена, і знайти: 1) математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 2) функцію розподілу $F(x)$; 3) $P(\alpha < \xi < \beta)$.

7.0. а) $F(x) = \begin{cases} 0, & x \le 1, \\ \frac{1}{8}(x^2 - 1), & 1 < x \le 3, \\ 1, & x > 3; \end{cases}$ б) $f(x) = \begin{cases} 0, & x < 1, \\ 1/4, & 1 \le x \le 5, \\ 0, & x > 5; \end{cases}$ $\alpha = 2$, $\beta = 4$.

7.1. а) $F(x) = \begin{cases} 0, & x \le 3, \\ \frac{1}{4}(x-3)^2, & 3 < x \le 5, \\ 1, & x > 5; \end{cases}$ б) $f(x) = \frac{4}{\sqrt{2\pi}} e^{-8(x+3)^2}$, $x \in \mathbf{R}$; $\alpha = -4$, $\beta = -3{,}2$.

7.2. а) $F(x) = \begin{cases} 0, & x \le 1, \\ \frac{1}{9}(x-1)^2, & 1 < x \le 4, \\ 1, & x > 4; \end{cases}$ б) $f(x) = \begin{cases} 0, & x < 0, \\ 0{,}5e^{-0{,}5x}, & x \ge 0; \end{cases}$ $\alpha = 2$, $\beta = 4$.

7.3. а) $F(x) = \begin{cases} 0, & x \le 1, \\ \frac{1}{7}(x^3 - 1), & 1 < x \le 2, \\ 1, & x > 2. \end{cases}$ б) $f(x) = \begin{cases} 0, & x < -2, \\ 1/5, & -2 \le x \le 3, \\ 0, & x > 3; \end{cases}$ $\alpha = -1$, $\beta = 2{,}4$.

7.4. а) $F(x) = \begin{cases} 0, & x \le 3, \\ \frac{1}{16}(x^2 - 9), & 3 < x \le 5, \\ 1, & x > 5; \end{cases}$ б) $f(x) = \frac{3}{\sqrt{2\pi}} e^{-\frac{9(x+1)^2}{2}}$, $x \in \mathbf{R}$; $\alpha = -2$, $\beta = 1$;

7.5. а) $F(x) = \begin{cases} 0, & x \le 0, \\ \frac{1}{8}x^3, & 0 < x \le 2, \\ 1, & x > 2; \end{cases}$ б) $f(x) = \begin{cases} 0, & x < 0, \\ 0{,}2e^{-0{,}2x}, & x \ge 0; \end{cases}$ $\alpha = 1$, $\beta = 5$;

7.6. а) $F(x) = \begin{cases} 0, & x \le 8, \\ (x-8)^2, & 8 < x \le 9, \\ 1, & x > 9; \end{cases}$ б) $f(x) = \begin{cases} 0, & x < -1, \\ 1/3, & -1 \le x \le 2, \\ 0, & x > 2; \end{cases}$ $\alpha = 0$, $\beta = 1{,}5$;

7.7. а) $F(x)=\begin{cases}0, & x\le 2,\\ \dfrac{1}{12}(x^2-4), & 2<x\le 4,\\ 1, & x>4.\end{cases}$ б) $f(x)=\dfrac{5}{\sqrt{2\pi}}e^{-\frac{25(x+4)^2}{2}}$, $x\in\mathbf{R}$;
$\alpha=-5$, $\beta=2$;

7.8. а) $F(x)=\begin{cases}0, & x\le 1,\\ \dfrac{1}{6}(x^3-x), & 1<x\le 2,\\ 1, & x>2.\end{cases}$ б) $f(x)=\begin{cases}0, & x<0,\\ 0{,}4e^{-0{,}4x}, & x\ge 0;\end{cases}$
$\alpha=0{,}25$, $\beta=5$;

7.9. а) $F(x)=\begin{cases}0, & x\le 2,\\ \dfrac{1}{16}(x-2)^2, & 2<x\le 6,\\ 1, & x>6;\end{cases}$ б) $f(x)=\dfrac{2}{\sqrt{2\pi}}e^{-2(x+5)^2}$, $x\in\mathbf{R}$;
$\alpha=-6$, $\beta=1$.

Задача 8. Задано вибірку W, яка характеризує місячний прибуток підприємців (у тис. грн). Побудувати статистичну модель генеральної сукупності вибірковим методом, для чого: скласти інтервальний статистичний ряд; побудувати гістограму та полігон частостей; знайти основні числові характеристики вибірки (вибіркове середнє, вибіркову дисперсію, вибіркове середнє квадратичне відхилення); висунути гіпотезу щодо можливого закону розподілу досліджуваної випадкової величини та перевірити її на адекватність за критерієм Пірсона з рівнем значущості $\alpha=0{,}05$.

8.0. $W = \{5, 9, 4, 11, 6, 5, 7, 1, 9, 7, 8, 5, 1, 6, 2, 4, 6, 4, 7, 2, 12, 3, 8, 3, 7, 3, 9, 5, 7, 7\}$.

8.1. $W = \{7, 8, 5, 8, 3, 8, 5, 8, 3, 8, 5, 8, 3, 8, 4, 9, 5, 9, 6, 9, 2, 10, 6, 10, 7, 10, 2, 12, 7, 13\}$.

8.2. $W = \{10, 8, 6, 11, 10, 9, 4, 12, 3, 4, 6, 11, 4, 8, 9, 14, 7, 11, 7, 5, 9, 5, 8, 6, 5, 6, 7, 8, 9, 8\}$.

8.3. $W = \{11, 10, 10, 6, 4, 12, 7, 4, 11, 10, 7, 5, 12, 8, 10, 9, 8, 5, 12, 8, 5, 9, 8, 5, 13, 9, 8, 9, 5, 15\}$.

8.4. $W = \{9, 12, 9, 14, 12, 9, 16, 8, 11, 8, 11, 8, 13, 7, 11, 7, 13, 7, 6, 10, 6, 10, 5, 13, 5, 10, 10, 5, 13, 10\}$.

8.5. $W = \{12, 14, 13, 11, 11, 13, 11, 12, 9, 6, 9, 6, 7, 6, 17, 9, 6, 7, 7, 9, 8, 10, 8, 10, 12, 10, 12, 14, 14, 16\}$.

8.6. $W = \{3, 5, 1, 8, 12, 8, 3, 11, 2, 5, 1, 4, 7, 9, 7, 9, 4, 7, 9, 2, 5, 1, 3, 7, 6, 4, 6, 6, 5, 4\}$.

8.7. $W = \{8, 13, 7, 5, 4, 8, 6, 2, 7, 4, 6, 8, 5, 3, 7, 6, 3, 9, 5, 3, 10, 4, 5, 9, 5, 6, 10, 7, 9, 11\}$.

8.8. $W = \{11, 7, 7, 11, 6, 6, 5, 12, 14, 8, 10, 11, 5, 3, 6, 8, 11, 3, 4, 9, 7, 4, 9, 10, 7, 6, 9, 6, 9, 9\}$.

8.9. $W = \{6, 4, 4, 4, 8, 8, 6, 4, 8, 5, 1, 8, 9, 5, 6, 1, 7, 11, 5, 1, 12, 2, 7, 5, 2, 6, 3, 7, 3, 7\}$.

Задача 9. Проведено дослідження обчислювального процесу на мережі персональних комп'ютерів (ПК) за показниками: час виконання тієї чи іншої процедури (програми) x (ум. од.) і ресурси ПК y (ум. од.). Необхідно: а) знайти теоретичні рівняння лінії регресії ξ на η і η на ξ і зобразити їх; б) побудувати емпіричні лінії регресії ξ на η і η на ξ; в) оцінити тісноту кореляційного зв'язку між ξ на η (ξ, η – ВВ, значеннями яких є, відповідно, час і ресурси). (Під ресурсами ПК розуміють обсяг оперативної та зовнішньої пам'яті, швидкодію процесора, час звертання до зовнішніх пристроїв тощо.)

9.0.

$\xi \backslash \eta$	15	20	25
6	3	8	
12	5	13	8
18		9	4

9.1.

$\xi \backslash \eta$	5	10	15
4	5	7	
8	4	15	7
12		10	2

9.2.

$\xi \backslash \eta$	10	20	30
5	5	7	
9	4	15	7
13		10	2

9.3.

$\xi \backslash \eta$	5	15	25
3	5	7	
5	4	15	7
7		10	2

9.4.

$\xi \backslash \eta$	10	15	20
7	3	8	
13	6	15	6
19		7	5

9.5.

$\xi \backslash \eta$	5	10	15
10	2	5	
20	9	15	6
30		8	5

9.6.

$\xi \backslash \eta$	4	9	14
6	2	7	
10	8	15	5
14		8	5

9.7.

$\xi \backslash \eta$	11	16	21
5	5	6	
9	5	15	7
13		8	4

9.8.

$\xi \backslash \eta$	6	11	16
10	5	7	
20	4	15	7
30		10	2

9.9.

$\xi \backslash \eta$	10	15	20
3	2	11	
6	6	15	4
9		8	4

Задача 10. Робота персонального комп'ютера (ПК) описується ланцюгом Маркова з трьома станами: S_1 – ПК у робочому стані; S_2 – ПК у режимі очікування; S_3 – ПК у сплячому режимі. За заданою матрицею P перехідних ймовірностей за один крок знайти: а) матрицю ймовірностей переходу за два кроки; б) граничні ймовірності. Побудувати граф ланцюга.

10.0. $P=\begin{pmatrix} 0{,}1 & 0{,}6 & 0{,}3 \\ 0{,}2 & 0{,}4 & 0{,}4 \\ 0{,}5 & 0{,}3 & 0{,}2 \end{pmatrix}.$

10.1. $P=\begin{pmatrix} 0{,}4 & 0{,}2 & 0{,}4 \\ 0{,}2 & 0{,}5 & 0{,}3 \\ 0{,}3 & 0{,}1 & 0{,}6 \end{pmatrix}.$

10.2. $P=\begin{pmatrix} 0{,}3 & 0{,}2 & 0{,}5 \\ 0{,}3 & 0{,}6 & 0{,}1 \\ 0{,}4 & 0{,}4 & 0{,}2 \end{pmatrix}.$

10.3. $P=\begin{pmatrix} 0{,}1 & 0{,}2 & 0{,}7 \\ 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}6 & 0{,}2 & 0{,}2 \end{pmatrix}.$

10.4. $P=\begin{pmatrix} 0{,}2 & 0{,}6 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}3 \\ 0{,}2 & 0{,}7 & 0{,}1 \end{pmatrix}.$

10.5. $P=\begin{pmatrix} 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}7 & 0{,}1 & 0{,}2 \\ 0{,}2 & 0{,}2 & 0{,}6 \end{pmatrix}.$

10.6. $P=\begin{pmatrix} 0{,}3 & 0{,}5 & 0{,}2 \\ 0{,}6 & 0{,}3 & 0{,}1 \\ 0{,}8 & 0{,}1 & 0{,}1 \end{pmatrix}.$

10.7. $P=\begin{pmatrix} 0{,}1 & 0{,}1 & 0{,}8 \\ 0{,}3 & 0{,}6 & 0{,}1 \\ 0{,}3 & 0{,}5 & 0{,}2 \end{pmatrix}.$

10.8. $P=\begin{pmatrix} 0{,}2 & 0{,}3 & 0{,}5 \\ 0{,}1 & 0{,}8 & 0{,}1 \\ 0{,}1 & 0{,}3 & 0{,}6 \end{pmatrix}.$

10.9. $P=\begin{pmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}3 & 0{,}1 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 \end{pmatrix}.$

4. Зразок виконання варіанта контрольної роботи

Задача 1. У партії з десяти деталей сім стандартних. Знайти ймовірність того, що серед шести взятих навмання деталей чотири будуть стандартними.

Розв'язання. Позначимо через A подію, яка полягає в тому, що серед шести взятих навмання деталей чотири будуть стандартними.

Число всіх можливих способів вибрати шість деталей з десяти наявних у партії (загальне число рівноможливих наслідків) дорівнює: $n = C_{10}^{6}$. Сприятливими будуть ті наслідки, коли із шести взятих деталей чотири будуть стандартними, а дві – ні. На кожний із $C_7^4 = 35$ варіантів вибору чотирьох стандартних деталей припадає $C_3^2 = 3$ варіанти вибору двох нестандартних, тому число сприятливих наслідків дорівнює: $m = C_7^4 \cdot C_3^2$. За формулою (1.1) обчислимо шукану ймовірність:

$$P(A) = \frac{m}{n} = \frac{C_7^4 \cdot C_3^2}{C_{10}^6} = \frac{1}{2}.$$

Задача 2. З відрізка $[0,1]$ навмання вибрані два числа x і y. Знайти ймовірність того, що їх сума не більша за 1, а добуток не більший за $2/9$.

Розв'язання. Експеримент E полягає у виборі навмання двох чисел x і y, причому $0 \le x \le 1$, $0 \le y \le 1$. Будемо вважати, що x і y – прямокутні декартові координати точок площини. За простір елементарних подій природно взяти всі точки квадрата $\Omega = \{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1\}$, площа якого $S_\Omega = 1$ (рис. 4.1).

Нехай A – подія, яка полягає у тому, що $x + y \le 1$, а $xy \le 2/9$. Можливі наслідки експерименту, які сприяють події A, задовольняють систему нерівностей:

$$\begin{cases} x + y \le 1, \\ xy \le 2/9, \end{cases} \quad (x, y) \in \Omega.$$

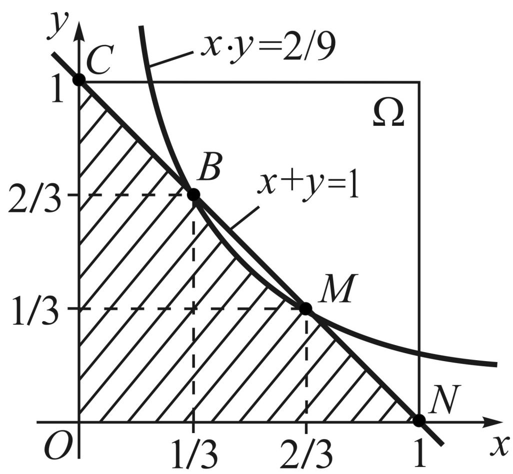

Рис. 4.1. **Простір Ω і подія A**

Побудуємо межі області, координати точок якої є розв'язками цієї системи: сторони квадрата Ω розташовані на прямих $x = 0$, $x = 1$, $y = 0$, $y = 1$; пряма $x + y = 1$ проходить через вершини $(1,0)$ і $(0,1)$ цього квадрата; лінія $xy \le 2/9$ є гіперболою (див. рис. 4.1).

Точки площини, які належать заштрихованій області, сприяють появі події A. Площа S_A цієї області є різницею площі прямокутного трикутника

OCN і площі S_1 фігури, що обмежена гіперболою $xy = 2/9$ і прямою $x + y = 1$:

$$S_{OCN} = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2},$$

$$S_1 = \int_{1/3}^{2/3} \left(1 - x - \frac{2}{9x}\right) dx = \left(x - \frac{x^2}{2} - \frac{2}{9} \ln|x| \right) \Bigg|_{1/3}^{2/3} = \left(\frac{2}{3} - \frac{2}{9} - \frac{2}{9} \ln \frac{2}{3}\right) -$$

$$- \left(\frac{1}{3} - \frac{1}{18} - \frac{2}{9} \ln \frac{1}{3}\right) = \frac{1}{6} - \frac{2}{9} \ln 2 \approx 0{,}0126\,.$$

Таким чином:

$$S_A = S_{OCN} - S_1 = \frac{1}{2} - \left(\frac{1}{6} - \frac{2}{9} \ln 2\right) \approx 0{,}4874.$$

За геометричним означенням (див. формулу (1.2)) знаходимо ймовірність події A:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{S_A}{S_\Omega} = \frac{0{,}4874}{1} = 0{,}4874.$$

Задача 3. Серед 50 електричних ламп є шість бракованих. Знайти ймовірність того, що серед трьох навмання вибраних ламп: 1) усі три стандартні; 2) тільки дві стандартні; 3) тільки одна стандартна; 4) жодної стандартної; 5) принаймні одна стандартна.

Розв'язання. Позначимо через L_i $(i = 1, 2, 3)$ подію, яка полягає в тому, що i-а лампа виявиться стандартною.

1) розглянемо подію $S_1 =$<серед узятих трьох ламп усі стандартні>. Тоді $S_1 = L_1 \cdot L_2 \cdot L_3$. За теоремою про ймовірність добутку залежних подій (див. (1.9)) маємо:

$$P(S_1) = P(L_1 \cdot L_2 \cdot L_3) = P(L_1) P(L_2 \mid L_1) P(L_3 \mid L_1 L_2) = \frac{44}{50} \cdot \frac{43}{49} \cdot \frac{42}{48} \approx 0{,}68;$$

2) $S_2 =$<серед узятих трьох ламп тільки дві стандартні>. Тоді $S_2 = L_1 \cdot L_2 \cdot \overline{L_3} \cup L_1 \cdot \overline{L_2} \cdot L_3 \cup \overline{L_1} \cdot L_2 \cdot L_3$. Події $L_1 \cdot L_2 \cdot \overline{L_3}$, $L_1 \cdot \overline{L_2} \cdot L_3$, $\overline{L_1} \cdot L_2 \cdot L_3$ несумісні, тому за теоремою про ймовірність суми несумісних подій (див. формулу (1.4)) маємо:

$$P(S_2)=P(L_1\cdot L_2\cdot \overline{L_3}\cup L_1\cdot \overline{L_2}\cdot L_3\cup \overline{L_1}\cdot L_2\cdot L_3)=$$

$=P(L_1\cdot L_2\cdot \overline{L_3})+P(L_1\cdot \overline{L_2}\cdot L_3)+P(\overline{L_1}\cdot L_2\cdot L_3)=\Big|$ за теоремою про ймовірність добутку залежних подій $\Big| = P(L_1)P(L_2\,|\,L_1)P(\overline{L_3}\,|\,L_1L_2)+$

$$+P(L_1)P(\overline{L_2}\,|\,L_1)P(L_3\,|\,L_1\overline{L_2})+P(\overline{L_1})P(L_2\,|\,\overline{L_1})P(L_3\,|\,\overline{L_1}L_2)=$$

$$=\frac{44}{50}\cdot\frac{43}{49}\cdot\frac{6}{48}+\frac{44}{50}\cdot\frac{6}{49}\cdot\frac{43}{48}+\frac{6}{50}\cdot\frac{44}{49}\cdot\frac{43}{48}\approx 0{,}29;$$

3) $S_3=$< серед узятих трьох ламп тільки одна стандартна >. Тоді $S_3=L_1\cdot \overline{L_2}\cdot \overline{L_3}\cup \overline{L_1}\cdot L_2\cdot \overline{L_3}\cup \overline{L_1}\cdot \overline{L_2}\cdot L_3$. За формулами (1.4), (1.9) маємо:

$$P(S_3)=P(L_1\cdot \overline{L_2}\cdot \overline{L_3}\cup \overline{L_1}\cdot L_2\cdot \overline{L_3}\cup \overline{L_1}\cdot \overline{L_2}\cdot L_3)=$$

$$=P(L_1\cdot \overline{L_2}\cdot \overline{L_3})+P(\overline{L_1}\cdot L_2\cdot \overline{L_3})+P(\overline{L_1}\cdot \overline{L_2}\cdot L_3)=$$

$$=P(L_1)P(\overline{L_2}\,|\,L_1)P(\overline{L_3}\,|\,L_1\overline{L_2})+P(\overline{L_1})P(L_2\,|\,\overline{L_1})P(\overline{L_3}\,|\,\overline{L_1}L_2)+$$

$$+P(\overline{L_1})P(\overline{L_2}\,|\,\overline{L_1})P(L_3\,|\,\overline{L_1}\overline{L_2})=\frac{44}{50}\cdot\frac{6}{49}\cdot\frac{5}{48}+\frac{6}{50}\cdot\frac{44}{49}\cdot\frac{5}{48}+\frac{6}{50}\cdot\frac{5}{49}\cdot\frac{4}{48}\approx 0{,}03;$$

4) $S_4=$< серед узятих трьох ламп жодної стандартної >. Тоді $S_4=\overline{L_1}\cdot \overline{L_2}\cdot \overline{L_3}$. За теоремою множення для залежних подій (1.9):

$$P(S_4)=P(\overline{L_1}\cdot \overline{L_2}\cdot \overline{L_3})=P(\overline{L_1})P(\overline{L_2}\,|\,\overline{L_1})P(\overline{L_3}\,|\,\overline{L_1}\overline{L_2})=$$

$$=\frac{6}{50}\cdot\frac{5}{49}\cdot\frac{4}{48}\approx 0{,}001;$$

5) $S_5=$< серед узятих трьох ламп принаймні одна стандартна >. Розглянемо протилежну подію $\overline{S_5}=$< серед трьох ламп жодної стандартної >$=S_4$. Тоді:

$$P(S_5)=1-P(\overline{S_5})=1-P(S_4)=1-0{,}001=0{,}999.$$

Задача 4. На трьох верстатах виготовляють однакові деталі, які надходять на конвеєр. З першого верстата надходить $20\,\%$, з другого – $30\,\%$, з третього – $50\,\%$ деталей. Перший верстат дає в середньому $0{,}2\,\%$ браку, другий – $0{,}3\,\%$, третій – $0{,}1\,\%$. Знайти ймовірність того, що: 1) навмання взята з конвеєра деталь буде бракованою; 2) виявлена бракована деталь виготовлена на першому верстаті.

Розв'язання. Позначимо через H_1, H_2, H_3 події, які полягають у тому, що навмання взята деталь виготовлена, відповідно, на першому, другому та третьому верстатах, а подія S – у тому, що навмання взята з конвеєра деталь бракована. Події H_1, H_2, H_3 – гіпотези – несумісні й утворюють повну групу подій. Що ж до події S, то вона може відбутися одночасно з кожною із цих подій.

За умовою задачі: $P(H_1)=0{,}2$, $P(H_2)=0{,}3$, $P(H_3)=0{,}5$. *Переконуємося*, що $\sum\limits_{i=1}^{3} P(H_i)=1$.

Також відомі умовні ймовірності настання події S – ймовірності того, що навмання взята деталь бракована за умови, що вона виготовлена, відповідно, на першому, другому та третьому верстатах: $P(S\,|\,H_1)=0{,}002$; $P(S\,|\,H_2)=0{,}003$; $P(S\,|\,H_3)=0{,}001$. Тоді:

1) за формулою повної ймовірності (1.13) підраховуємо ймовірність події S:

$$P(S)=P(H_1)\cdot P(S|H_1)+P(H_2)\cdot P(S|H_2)+P(H_3)\cdot P(S|H_3)=$$
$$=0{,}2\cdot 0{,}002+0{,}3\cdot 0{,}003+0{,}5\cdot 0{,}001=0{,}0018;$$

2) ймовірність того, що виявлена бракована деталь виготовлена на першому верстаті, обчислимо за формулою Байєса (1.15):

$$P(H_1\,|\,S)=\frac{P(H_1)\cdot P(S\,|\,H_1)}{P(H_1)\cdot P(S\,|\,H_1)+P(H_2)\cdot P(S\,|\,H_2)+P(H_3)\cdot P(S\,|\,H_3)}=$$
$$=\frac{0{,}2\cdot 0{,}002}{0{,}2\cdot 0{,}002+0{,}3\cdot 0{,}003+0{,}5\cdot 0.001}=\frac{2}{9}\approx 0{,}2222.$$

Задача 5. За результатами податкових перевірок кожне друге мале підприємство порушує фінансову дисципліну. Знайти ймовірність того, що серед 1000 зареєстрованих малих підприємств порушення мають: 1) 495 підприємств; 2) від 480 до 540 підприємств.

Розв'язання. Подія $S=<$мале підприємство має порушення фінансової дисципліни$>$. За умовою $P(S)=p=0{,}5$, тоді $P(\overline{S})=q=0{,}5$. Проведено $n=1000$ незалежних випробувань.

1. Оскільки $k = 495$, $npq = 250 > 20$, то для обчислення ймовірності $\widetilde{P}_{1000}(495)$ застосуємо локальну теорему Муавра – Лапласа (1.19):

$$x = \frac{k - np}{\sqrt{npq}} = \frac{495 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} \approx -0{,}32.$$

За таблицею значень функції Гаусса $\varphi(x)$ (див. табл. А.1 додатка А) з урахуванням її парності $(\varphi(-x) = \varphi(x))$ дістанемо: $\varphi(x) = \varphi(-0{,}32) = \varphi(0{,}32) = 0{,}3790$. Таким чином,

$$\widetilde{P}_{1000}(495) = \frac{\varphi(-0{,}32)}{\sqrt{250}} = \frac{0{,}3790}{\sqrt{250}} = 0{,}024.$$

2. Оскільки $k_1 = 480$, $k_2 = 540$, $npq = 250 > 20$, то для обчислення ймовірності $\widetilde{P}_{1000}(480{,}540)$ застосуємо інтегральну теорему Муавра – Лапласа (1.21). Знайдемо значення t_1, t_2:

$$t_1 = \frac{k_1 - np}{\sqrt{npq}} = \frac{480 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} \approx -1{,}26, \; t_2 = \frac{k_2 - np}{\sqrt{npq}} = \frac{540 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} \approx 2{,}53.$$

За таблицею значень функції Лапласа $\Phi(x)$ (див. табл. Б.1 додатка Б) з урахуванням її непарності ($\Phi(-x) = -\Phi(x)$) дістанемо:

$$\Phi(-1{,}26) = -\Phi(1{,}26) = -0{,}3962; \quad \Phi(2{,}53) = 0{,}4943.$$

Отже,

$$\widetilde{P}_{1000}(480{,}540) = \Phi(2{,}53) - \Phi(-1{,}26) = 0{,}4943 + 0{,}3962 = 0{,}8905.$$

Задача 6. Баскетболіст кидає м'яч у корзину з імовірністю влучення після кожного кидка 0,4. Для дискретної випадкової величини ξ, значеннями якої є число влучень після трьох кидків треба: 1) скласти закон розподілу; 2) знайти її математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 3) побудувати полігон розподілу; 4) знайти функцію розподілу $F_\xi(x)$ і побудувати її графік.

Розв'язання.

1. Випадкова величина ξ, значеннями якої є число влучень після трьох кидків, може набувати значень: $0,\ 1,\ 2,\ 3$. Імовірності, з якими ВВ ξ набуває цих значень, визначимо за формулою Бернуллі (1.16), де $n=3$, $p=0{,}4$, $q=0{,}6$, $k=0$, $1,\ 2,\ 3$:

$$P(\xi=0)=C_3^0 p^0 q^3=0{,}6^3=0{,}216;$$

$$P(\xi=1)=C_3^1 p^1 q^2=3\cdot 0{,}4\cdot 0{,}6^2=0{,}432;$$

$$P(\xi=2)=C_3^2 p^2 q^1=3\cdot 0{,}4^2\cdot 0{,}6=0{,}288;$$

$$P(\xi=3)=C_3^3 p^3 q^0=0{,}4^3=0{,}064.$$

Тоді закон розподілу матиме вигляд:

ξ	0	1	2	3
p	0,216	0,432	0,288	0,064

Переконуємося, що $\sum\limits_{i=1}^{4} p_i=1$: $0{,}216+0{,}432+0{,}288+0{,}064=1$.

2. Знайдемо числові характеристики ВВ ξ: $M\xi$, $D\xi$, σ_ξ.

За формулою (1.28) обчислюємо математичне сподівання:

$$M\xi=\sum_{i=1}^{4} x_i p_i=0\cdot 0{,}216+1\cdot 0{,}432+2\cdot 0{,}288+3\cdot 0{,}064=1{,}2.$$

Для знаходження дисперсії спочатку обчислимо $M\xi^2$, а потім скористаємось формулою (1.31):

$$M\xi^2=0^2\cdot 0{,}216+1^2\cdot 0{,}432+2^2\cdot 0{,}288+3\cdot 0{,}064=2{,}16.$$

$D\xi=M\xi^2-M^2\xi=2{,}16-(1{,}2)^2=0{,}72$, звідки $\sigma_\xi=\sqrt{D\xi}\approx 0{,}85$.

Оскільки можливими значеннями ДВВ ξ є число успіхів у схемі незалежних випробувань, а відповідні ймовірності знайдені за формулою Бернуллі, то ВВ ξ розподілена за біноміальним законом. Отже, можна перевірити отримані значення $M\xi$ і $D\xi$ за відповідними формулами для числових характеристик біноміального розподілу (див. додаток Г):

$$M\xi=np=3\cdot 0{,}4=1{,}2;\quad D\xi=npq=3\cdot 0{,}4\cdot 0{,}6=0{,}72.$$

3. Побудуємо полігон розподілу – ламану з вершинами в точках (x_i, p_i), $i=\overline{1,4}$: $(0;0{,}216)$, $(1;0{,}432)$, $(2;0{,}288)$, $(3;0{,}064)$ (рис. 4.2).

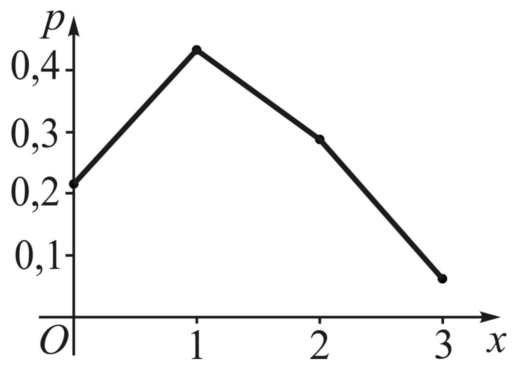

Рис. 4.2. **Полігон розподілу**

4. Для знаходження функції розподілу $F_\xi(x)$ розіб'ємо числову вісь можливими значенням ВВ ξ: 0, 1, 2, 3, на проміжки, на яких $F_\xi(x)$ зберігає стале значення, і скористаємось означенням (1.25): $F_\xi(x)=P(\xi<x)\ \ \forall x\in\mathbf{R}$:

$$x\le 0 \ \Rightarrow\ F_\xi(x)=0;$$

$$0<x\le 1 \ \Rightarrow\ F_\xi(x)=P(\xi=0)=0{,}216;$$

$$1<x\le 2 \ \Rightarrow\ F_\xi(x)=P(\xi=0)+P(\xi=1)=0{,}216+0{,}432=0{,}648;$$

$$2<x\le 3 \ \Rightarrow\ F_\xi(x)=P(\xi=0)+P(\xi=1)+P(\xi=2)=0{,}648+0{,}288=0{,}936;$$

$$x>3 \ \Rightarrow\ F_\xi(x)=P(\xi=0)+P(\xi=1)+P(\xi=2)+P(\xi=3)=1.$$

Отже,

$$F_\xi(x)=\begin{cases}0, & x\le 0,\\ 0{,}216, & 0<x\le 1,\\ 0{,}648, & 1<x\le 2,\\ 0{,}936, & 2<x\le 3,\\ 1, & x>3.\end{cases}$$

Графік функції $F_\xi(x)$ зображено на рис. 4.3.

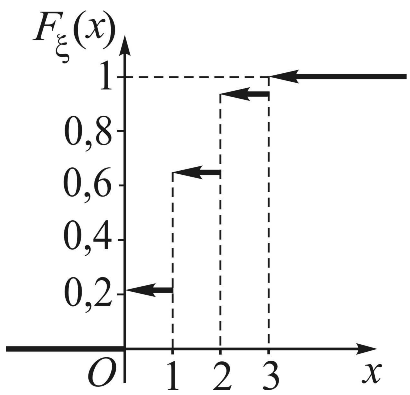

Рис. 4.3. **Графік функції розподілу**

Задача 7. Треба: а) за заданою функцією розподілу неперервної випадкової величини ξ знайти: 1) щільність розподілу ймовірностей $f(x)$; 2) математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 3) побудувати графіки функцій $F(x)$ і $f(x)$;

б) за заданою щільністю розподілу ймовірностей неперервної випадкової величини ξ визначити, за яким законом вона розподілена, і знайти: 1) математичне сподівання $M\xi$, дисперсію $D\xi$, середнє квадратичне відхилення σ_ξ; 2) функцію розподілу $F(x)$; 3) $P(\alpha<\xi<\beta)$.

а) $F(x)=\begin{cases} 0, & x\le -1,\\ \frac{1}{9}(x+1)^2, & -1<x\le 2,\\ 1, & x>2; \end{cases}$ б) $f(x)=\frac{4}{3\sqrt{2\pi}}e^{-\frac{8(x+2)^2}{9}}$, $x\in\mathbf{R}$;

$\alpha=-2{,}5$, $\beta=0$.

Розв'язання:

А. 1. Щільність розподілу ймовірностей $f(x)$ знаходимо як похідну від функції розподілу $F(x)$ з урахуванням, що:

$$(0)'=0,\quad (1)'=0,\quad \left(\frac{1}{9}(x+1)^2\right)'=\frac{1}{9}\cdot 2(x+1)\cdot(x+1)'=\frac{2}{9}(x+1).$$

Отже,

$$f(x)=F'(x)=\begin{cases} 0, & x\le -1,\\ \frac{2}{9}(x+1), & -1<x\le 2,\\ 0, & x>2. \end{cases}$$

2. Математичне сподівання $M\xi$ і дисперсію $D\xi$ знайдемо, відповідно, за формулами (1.39) і (1.41):

$$M\xi=\int_{-\infty}^{+\infty}x\cdot f(x)dx=\int_{-\infty}^{-1}x\cdot 0dx+\int_{-1}^{2}x\cdot\frac{2}{9}(x+1)dx+\int_{2}^{+\infty}x\cdot 0dx=$$

$$=\frac{2}{9}\int_{-1}^{2}(x^2+x)dx=\frac{2}{9}\left(\frac{x^3}{3}+\frac{x^2}{2}\right)\Bigg|_{-1}^{2}=\frac{2}{9}\left(\left(\frac{8}{3}+2\right)-\left(-\frac{1}{3}+\frac{1}{2}\right)\right)=\frac{2}{9}\cdot\frac{9}{2}=1;$$

$$M\xi^2=\int_{-\infty}^{+\infty}x^2\cdot f(x)dx=\int_{-\infty}^{-1}x^2\cdot 0dx+\int_{-1}^{2}x^2\cdot\frac{2}{9}(x+1)dx+\int_{2}^{+\infty}x^2\cdot 0dx=$$

$$=\frac{2}{9}\int_{-1}^{2}(x^3+x^2)dx=\frac{2}{9}\left(\frac{x^4}{4}+\frac{x^3}{3}\right)\Bigg|_{-1}^{2}=\frac{2}{9}\left(\left(4+\frac{8}{3}\right)-\left(\frac{1}{4}-\frac{1}{3}\right)\right)=\frac{2}{9}\cdot\frac{27}{4}=1{,}5;$$

$D\xi=M\xi^2-M^2\xi=1{,}5-1^2=0{,}5$, звідки $\sigma_\xi=\sqrt{0{,}5}\approx 0{,}71$.

3. Побудуємо графіки функцій $F(x)$ і $f(x)$ за їх аналітичними виразами (рис. 4.4, 4.5):

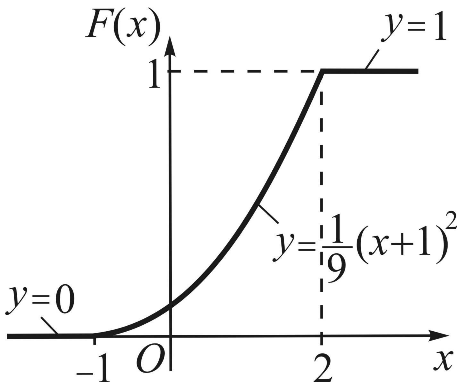

Рис. 4.4. **Графік $\boldsymbol{F(x)}$**

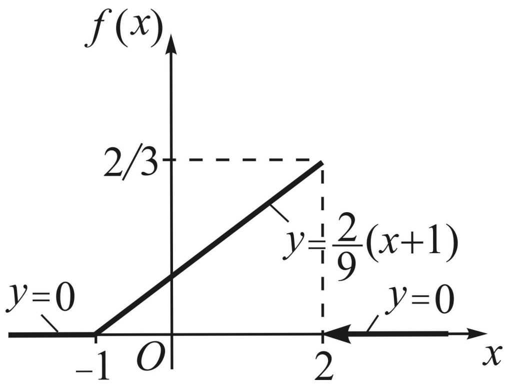

Рис. 4.5. **Графік $\boldsymbol{f(x)}$**

Б. Зіставляємо задану в умові задачі $f(x)$ з виразами для щільностей розподілу ймовірностей основних законів розподілу неперервної ВВ (див. додаток Г) і робимо висновок, що ВВ ξ розподілена за *нормальним законом* з параметрами $a=-2$ і $\sigma=3/4$:

$$f(x)=\frac{4}{3\sqrt{2\pi}}e^{-\frac{8(x+2)^2}{9}}\Rightarrow \left| f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}} \right| \Rightarrow \begin{bmatrix} x-a=x+2\Rightarrow a=-2, \\ \frac{1}{\sigma}=\frac{4}{3}\Rightarrow\sigma=\frac{3}{4}. \end{bmatrix}$$

1. Математичне сподівання $M\xi$, дисперсію $D\xi$ і середнє квадратичне відхилення σ_ξ визначаються параметрами нормального розподілу (див. додаток Г):

$$M\xi=a=-2,\;\; D\xi=\sigma^2=9/16,\;\; \sigma_\xi=\sigma=3/4.$$

2. Функція розподілу $F(x)$ нормально розподіленої ВВ має вигляд (див. додаток Г): $F(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right)$, де $\Phi(x)$ – функція Лапласа. Звідки для $a=-2$ і $\sigma=3/4$ маємо:

$$F(x)=\frac{1}{2}+\Phi\left(\frac{x+2}{3/4}\right)=\frac{1}{2}+\Phi\left(\frac{4(x+2)}{3}\right).$$

3. Ймовірність $P(-2{,}5<\xi<0)$ знайдемо за формулою (1.38):

$$P(-2{,}5<\xi<0)=\Phi\left(\frac{0+2}{3/4}\right)-\Phi\left(\frac{-2{,}5+2}{3/4}\right)=\Phi\left(\frac{8}{3}\right)-\Phi\left(\frac{-2}{3}\right)=$$

$$=\left|\ \Phi(-x)=-\Phi(x)\right|=\Phi\left(\frac{8}{3}\right)+\Phi\left(\frac{2}{3}\right)=\left|\text{ див. додаток Б }\right|=$$

$$=0{,}4962+0{,}2486=0{,}7448.$$

Задача 8. Задано вибірку, яка характеризує місячний прибуток підприємців (у тис. грн.): $W=\{7, 9, 12, 15, 16, 16, 19, 8, 10, 12, 15, 15, 19, 13, 10, 13, 15, 17, 19, 13, 13, 16, 17, 16, 18, 20, 19, 22, 14, 24\}$. Побудувати статистичну модель генеральної сукупності вибірковим методом, для чого: скласти інтервальний статистичний ряд; побудувати гістограму та полігон частостей; знайти основні числові характеристики вибірки (вибіркове середнє, вибіркову дисперсію, вибіркове середнє квадратичне відхилення); висунути гіпотезу щодо можливого закону розподілу досліджуваної випадкової величини та перевірити її на адекватність за критерієм Пірсона з рівнем значущості $\alpha=0{,}05$.

Розв'язання.

І. Первинна обробка вибірки.

Побудуємо варіаційний ряд вибірки, для чого розташуємо її елементи у неспадному порядку: (7, 8, 9, 10, 10, 12, 12, 13, 13, 13, 13, 14, 15, 15, 15, 15, 16, 16, 16, 16, 17, 17, 18, 19, 19, 19, 19, 20, 22, 24). Звідки: $x_{\min}=7$, $x_{\max}=24$, $R=x_{\max}-x_{\min}=24-7=17$ – розмах вибірки; $n=30$ – обсяг вибірки.

Крок розбиття h визначимо за формулою Стерджеса:

$$h=\frac{R}{1+3{,}2\cdot\lg n}=\frac{17}{1+3{,}2\cdot\lg 30}=\frac{17}{5{,}73}\approx 3.$$

Розіб'ємо весь проміжок вибіркових значень: $[x_{\min}, x_{\max}]=[7, 24]$ на часткові проміжки однакової довжини $h=3$, але попередньо збільшимо розмах вибірки на $0{,}5$: $R=[x_{\min}-0{,}5; x_{\max}]=[6{,}5; 24]$, щоб середини проміжків зображалися "зручними" числами. У результаті утворилося шість проміжків ($k=6$).

Знайдемо частості w_i попадань елементів вибірки у кожний із них і складемо інтервальний статистичний ряд:

$[x_{i-1},x_i)$	[6,5;9,5)	[9,5;12,5)	[12,5;15,5)	[15,5;18,5)	[18,5;21,5)	[21,5;24,5]
$w_i=\dfrac{m_i}{n}$	$\dfrac{3}{30}$	$\dfrac{4}{30}$	$\dfrac{9}{30}$	$\dfrac{7}{30}$	$\dfrac{5}{30}$	$\dfrac{2}{30}$

Перевіряємо виконання умови $\sum_{i=1}^{k=6} w_i = 1$: $\frac{3}{30}+\frac{4}{30}+\frac{9}{30}+\frac{7}{30}+\frac{5}{30}+\frac{2}{30}=1.$

Знаходимо величини w_i/h, $i=\overline{1,6}$, де $h=3$ – крок розбиття:

$$\frac{w_i}{h}=\frac{m_i}{nh}:\ \frac{3}{90},\ \frac{4}{90},\ \frac{9}{90},\ \frac{7}{90},\ \frac{5}{90},\ \frac{2}{90}.$$

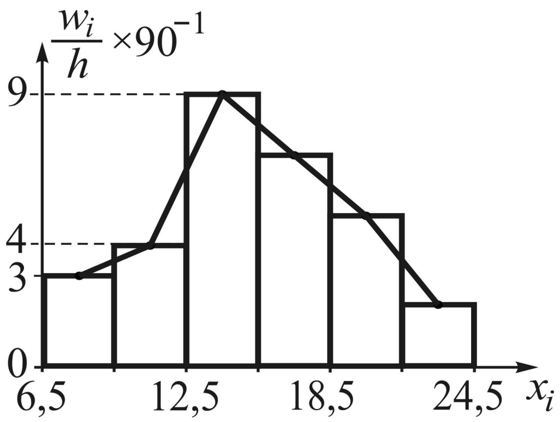

Рис. 4.6. **Гістограма та полігон частостей**

Будуємо власне гістограму, для чого на кожному i-му частковому проміжку ($i=\overline{1,6}$) зводимо прямокутник з основою h і висотою w_i/h (рис. 4.6).

Для побудови полігона частостей розподілу з'єднаємо ламаною середини верхніх основ усіх прямокутників.

Перетворимо інтервальний статистичний ряд на дискретний, для чого знайдемо середини часткових проміжків $x_i^*=(x_{i-1}+x_i)/2$ ($i=\overline{1,6}$):

$[x_{i-1},x_i)$	[6,5;9,5)	[9,5;12,5)	[12,5;15,5)	[15,5;18,5)	[18,5;21,5)	[21,5;24,5]
x_i^*	8	11	14	17	20	23
m_i	3	4	9	7	5	2

Основні числові характеристики вибірки знаходимо за формулами (1.51), (1.53), (1.54):

$$\bar{x}_\text{в}=\frac{1}{n}\cdot\sum_{i=1}^{6}x_i^* m_i=\frac{1}{30}\cdot(8\cdot 3+11\cdot 4+14\cdot 9+17\cdot 7+20\cdot 5+23\cdot 2)=15{,}3;$$

$$\hat{D}_{\text{в}} = \frac{1}{29} \cdot \left(\sum_{i=1}^{6} (x_i^*)^2 m_i - 30\,(\bar{x}_{\text{в}})^2 \right) = \frac{1}{29} \cdot \left(7521 - 30 \cdot (15{,}3)^2 \right) \approx 17{,}18;$$

$$\hat{\sigma}_{\text{в}} = \hat{s} = \sqrt{\hat{D}_{\text{в}}} = \sqrt{17{,}18} \approx 4{,}15.$$

II. Висунення гіпотези щодо можливого закону розподілу й оцінка параметрів розподілу.

За виглядом гістограми частостей (чи полігону) висуваємо *гіпотезу* H_0: генеральна сукупність розподілена за нормальним законом з параметрами a і σ, точковими оцінками яких є, відповідно, $\bar{x}_{\text{в}}$ і $\hat{s}$, тобто $a \approx \bar{x}_{\text{в}} = 15{,}3$, $\sigma \approx \hat{s} = 4{,}15$. Тоді теоретична щільність розподілу набуває вигляду (див. додаток Г):

$$\tilde{f}(x) = \frac{1}{4{,}15 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-15{,}3)^2}{2 \cdot 17{,}18}} = \frac{1}{4{,}35}\,\varphi\!\left(\frac{x - 15{,}3}{4{,}15} \right),$$

де $\varphi(x)$ – функція Гаусса.

III. Перевірка адекватності висуненої гіпотези за критерієм Пірсона з рівнем значущості $\alpha = 0{,}05$.

Гіпотеза H_0: генеральна сукупність розподілена за нормальним законом. Перевіримо узгодженість теоретичного й емпіричного розподілів, для чого складемо таблицю (табл. 4.1) розрахунків статистики Пірсона $\chi^2_{\text{в}}$ (1.56):

Таблиця 4.1

Перевірка адекватності гіпотези за критерієм Пірсона

<table>
<tr><th>N п/п</th><th>x_i^*</th><th>$u_i = \frac{x_i^* - 15{,}3}{4{,}15}$</th><th>$\varphi(u_i)$</th><th colspan="2">$\tilde{m}_i = \frac{nh}{4{,}15} \cdot \varphi(u_i)$</th><th colspan="2">$m_i$</th><th>$\frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$</th></tr>
<tr><td>1</td><td>8</td><td>–1,761</td><td>0,0846</td><td>1,84</td><td rowspan="2">6,9</td><td>3</td><td rowspan="2">7</td><td rowspan="2">0,001</td></tr>
<tr><td>2</td><td>11</td><td>–1,037</td><td>0,2329</td><td>5,06</td><td>4</td></tr>
<tr><td>3</td><td>14</td><td>–0,314</td><td>0,3798</td><td colspan="2">8,25</td><td colspan="2">9</td><td>0,068</td></tr>
<tr><td>4</td><td>17</td><td>0,410</td><td>0,3668</td><td colspan="2">7,96</td><td colspan="2">7</td><td>0,116</td></tr>
<tr><td>5</td><td>20</td><td>1,134</td><td>0,2098</td><td>4,55</td><td rowspan="2">6,09</td><td>5</td><td rowspan="2">7</td><td rowspan="2">0,136</td></tr>
<tr><td>6</td><td>23</td><td>1,858</td><td>0,0711</td><td>1,54</td><td>2</td></tr>
<tr><td>$\sum$</td><td></td><td></td><td></td><td colspan="2">29,2</td><td></td><td>30</td><td>0,321</td></tr>
</table>

Таким чином, за результатами обчислень $\chi^2_{\text{в}} = 0{,}321$. За таблицею критичних точок розподілу χ^2 (див. табл. Д.1 додатка Д) для рівня значущості $\alpha = 0{,}05$ і числа ступенів свободи $\nu = k - r - 1 = 4 - 2 - 1 = 1$, де $k = 4$ – кількість проміжків після об'єднання, $r = 2$ – число параметрів нормального розподілу, знаходимо: $\chi^2_{\text{крит}} = \chi^2_{1;0,05} = 3{,}8$. Оскільки $\chi^2_{\text{в}} = 0{,}321 < \chi^2_{\text{крит}} = 3{,}8$, то немає підстав для відхилення гіпотези H_0 про нормальний розподіл генеральної сукупності.

Висновок. Місячний прибуток підприємців розподілений за нормальним законом зі щільністю розподілу:

$$\tilde{f}(x) = \frac{1}{4{,}15 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-15{,}3)^2}{34{,}36}}.$$

Задача 9. Проведено дослідження обчислювального процесу на мережі персональних комп'ютерів (ПК) за показниками: час виконання тієї чи іншої процедури (програми) x (ум. од.) і ресурси ПК y (ум. од.). Необхідно: а) знайти теоретичні рівняння лінії регресії ξ на η і η на ξ і зобразити їх; б) побудувати емпіричні лінії регресії ξ на η і η на ξ; в) оцінити тісноту кореляційного зв'язку між ξ на η (ξ, η – ВВ, значеннями яких є, відповідно, час і ресурси). (Під ресурсами ПК розуміють обсяг оперативної та зовнішньої пам'яті, швидкодію процесору, час звертання до зовнішніх пристроїв тощо.)

ξ \ η	3	7	11
5	8	6	
8	4	13	5
11		8	6

Розв'язання: а) здійснюємо первинну обробку вибірки згідно з формулами (1.59) – (1.61), для чого підраховуємо відповідні суми. Для полегшення арифметичних обчислень перейдемо від змінних x і y до умовних змінних u і v за співвідношеннями (1.65), в яких C_x (C_y) – варіанта x (y) з найбільшою частотою для ВВ ξ (ВВ η); h_x, h_y – кроки, з якими наведені значення ξ, η:

$$u = \frac{x - C_x}{h_x} = \left| \begin{matrix} C_x = 8, \\ h_x = 3 \end{matrix} \right| = \frac{x-8}{3}, \quad v = \frac{y - C_y}{h_y} = \left| \begin{matrix} C_y = 7, \\ h_y = 4 \end{matrix} \right| = \frac{y-7}{4}.$$

За результатами обробки вибірки заповнюємо табл. 4.2.

Таблиця 4.2

Обчислення сум для підрахунку емпіричних числових характеристик

ξ \ η	u \ v	3	7	11	m_i	$m_i u_i$	$m_i u_i^2$
		-1	0	1			
5	-1	8	6		14	-14	14
8	0	4	13	5	22	0	0
11	1		8	6	14	14	14
m_j		12	27	11	$\sum_m = 50$	$\sum_1 = 0$	$\sum_2 = 28$
$m_j v_j$		-12	0	11	$\sum_3 = -1$		
$m_j v_j^2$		12	0	11	$\sum_4 = 23$		
$\sum\limits_{i,j} m_{ij} u_i v_j$		8	0	6	$\sum_5 = 14$		

За табл. 4.2 знаходимо:

$$\bar{u} = \frac{1}{50} \sum m_i u_i = \frac{1}{50} \sum\nolimits_1 = \frac{1}{50} 0 = 0;$$

$$\bar{v} = \frac{1}{50} \sum m_j v_j = \frac{1}{50} \sum\nolimits_3 = \frac{1}{50} (-1) = -0{,}02;$$

$$\overline{u^2} = \frac{1}{50} \sum m_i u_i^2 = \frac{1}{50} \sum\nolimits_2 = \frac{1}{50} 28 = 0{,}56;$$

$$\overline{v^2} = \frac{1}{50} \sum m_j v_j^2 = \frac{1}{50} \sum\nolimits_4 = \frac{1}{50} 23 = 0{,}46;$$

$$\overline{u \cdot v} = \frac{1}{50} \sum_{i,j} m_{ij} u_i v_j = \frac{1}{50} \sum\nolimits_5 = \frac{1}{50} 14 = 0{,}28.$$

Підраховуємо емпіричні дисперсії, середні квадратичні відхилення та коефіцієнт кореляції (s_u^2, s_v^2, s_u, s_v, r_{uv}):

$$s_u^2 = \overline{u^2} - \overline{u}^2 = 0{,}56, \quad s_u = \sqrt{s_u^2} = \sqrt{0{,}56} \approx 0{,}7483;$$

$$s_v^2 = \overline{v^2} - \overline{v}^2 = 0{,}4596, \quad s_v = \sqrt{s_v^2} = \sqrt{0{,}4596} \approx 0{,}6779;$$

$$r_{uv} = \frac{\overline{u \cdot v} - \overline{u} \cdot \overline{v}}{s_u \cdot s_v} = \frac{0{,}28 - 0 \cdot (-0{,}02)}{\sqrt{0{,}56} \cdot \sqrt{0{,}4596}} = \frac{0{,}28}{0{,}5073} \approx 0{,}552.$$

Оскільки $r_{uv} = r_{xy} = r$, то остаточно маємо: $r = 0{,}552$.

Далі перевіряємо, чи приймається гіпотеза H_0: $\rho = 0$, беручи рівень значущості $\alpha = 0{,}05$; тоді рівень довіри $p = 1 - \alpha = 0{,}95$. За таблицею значень функції Лапласа знаходимо t_p, потім ε й інтервал довіри, використовуючи формули (1.62) – (1.64):

$$t_p = \Phi^{-1}(p/2) = \Phi^{-1}(0{,}475) = 1{,}96;$$

$$\varepsilon = t_p \cdot (1 - r^2)/\sqrt{n} = 1{,}96 \cdot (1 - 0{,}552^2)/\sqrt{50} \approx 0{,}193;$$

$$(r - \varepsilon, r + \varepsilon) = (0{,}552 - 0{,}193; 0{,}552 + 0{,}193) = (0{,}359; 0{,}745).$$

Оскільки $0 \notin (r - \varepsilon, r + \varepsilon)$, то гіпотеза про відсутність кореляційного зв'язку відхиляється.

Далі здійснюємо перехід від умовних характеристик: $\overline{u}$, $\overline{v}$, s_u^2, s_v^2, до фактичних емпіричних центрів розподілу і дисперсій: $\overline{x}$, $\overline{y}$, s_x^2, s_y^2, s_x, s_y, за формулами (1.66), (1.67):

$$\overline{x} = \overline{u} \cdot h_x + C_x = 0 \cdot 3 + 5 = 8; \quad \overline{y} = \overline{v} \cdot h_y + C_y = -0{,}02 \cdot 4 + 7 = 6{,}92;$$

$$s_x^2 = s_u^2 \cdot h_x^2 = 0{,}56 \cdot 3^2 = 5{,}04; \quad s_y^2 = s_v^2 \cdot h_y^2 = 0{,}4596 \cdot 4^2 = 7{,}3536;$$

$$s_x = \sqrt{5{,}04} \approx 2{,}2450; \quad s_y = \sqrt{7{,}3536} \approx 2{,}7118.$$

Теоретичні рівняння ліній регресії отримуємо за формулами (1.74), (1.75) з урахуванням (1.68).

Для регресії ξ на η: $\overline{x}_y = a_0 + a_1 y$,

$$a_0 = \overline{x} - r \frac{s_x}{s_y} \overline{y} = 8 - 0{,}552 \cdot \frac{2{,}2450}{2{,}7118} \cdot 6{,}92 \approx 4{,}84;$$

$$a_1 = r\frac{s_x}{s_y} = 0{,}552 \cdot \frac{2{,}2450}{2{,}7118} \approx 0{,}46.$$

Отже, $\bar{x}_y = 4{,}84 + 0{,}46 \cdot y$ – *теоретичне рівняння регресії* ξ *на* η.

Для регресії η на ξ: $\bar{y}_x = b_0 + b_1 x$,

$$b_0 = \bar{y} - r\frac{s_y}{s_x}\bar{x} = 6{,}92 - 0{,}552 \cdot \frac{2{,}7118}{2{,}2450} \cdot 8 \approx 1{,}59;$$

$$b_1 = r\frac{s_y}{s_x} = 0{,}552 \cdot \frac{2{,}7118}{2{,}2450} \approx 0{,}67.$$

Отже, $\bar{y}_x = 1{,}59 + 0{,}67 \cdot x$ – *теоретичне рівняння регресії* η *на* ξ.

Будуємо прямі в одній системі координат. Обидві вони проходять через точку $(\bar{x}, \bar{y}) = (8; 6{,}92)$, тому для їх нанесення достатньо вказати ще по одній точці (рис. 4.7):

$$y = 3 \Rightarrow \bar{x}_y = 6{,}22 \Rightarrow (6{,}22; 3);$$

$$x = 5 \Rightarrow \bar{y}_x = 4{,}94 \Rightarrow (5; 4{,}94);$$

б) для побудови емпіричних ліній регресії ξ на η і η на ξ за табл. 4.2 підраховуємо умовні середні значення $\bar{x}_y$, $\bar{y}_x$:

$$\bar{x}_{y=3} = (5 \cdot 8 + 8 \cdot 4)/12 = 6;$$

$$\bar{x}_{y=7} = (5 \cdot 6 + 8 \cdot 13 + 11 \cdot 8)/27 \approx 8{,}2;$$

$$\bar{x}_{y=11} = (8 \cdot 5 + 11 \cdot 6)/11 \approx 9{,}6;$$

$$\bar{y}_{x=5} = (3 \cdot 8 + 7 \cdot 6)/14 \approx 4{,}7;$$

$$\bar{y}_{x=8} = (3 \cdot 4 + 7 \cdot 13 + 11 \cdot 5)/28 \approx 7{,}2;$$

$$\bar{y}_{x=11} = (7 \cdot 8 + 11 \cdot 6)/14 \approx 8{,}7.$$

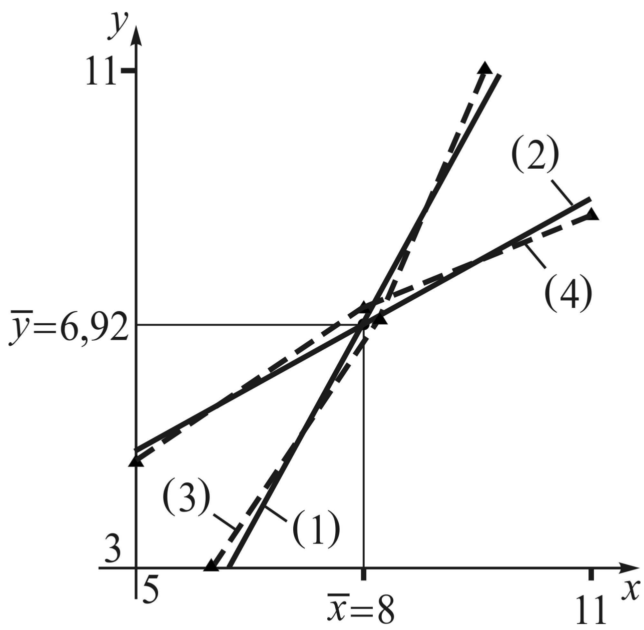

Рис. 4.7. **Емпіричні та теоретичні лінії регресії**

Умовні позначення:

(1) – теоретична лінія регресії ξ на η;

(2) – теоретична лінія регресії η на ξ;

(3) – емпірична лінія регресії ξ на η;

(4) – емпірична лінія регресії η на ξ

Зображуємо на координатній площині (рис. 4.7) точки $(\bar{x}_y, y_j)$, $j = \overline{1,3}$, $((x_i, \bar{y}_x),\ i = \overline{1,3})$, і з'єднуємо їх відрізками прямої, в результаті чого отримуємо **емпіричну лінію регресії** ξ на η (η на ξ);

в) за табл. 1.2 знаходимо, що $r = r_{xy} = 0{,}552$ входить у проміжок $0{,}5 - 0{,}69$, тому за вибіркою зв'язок між ξ і η *значний.*

Величина $r^2 = 0{,}3047 \approx 0{,}3$ показує, що близько $30\,\%$ мінливості часу або ресурсів обумовлені лінійною залежністю, а $70\,\%$ пояснюються впливом чинників, зв'язок між якими не є лінійним і які не враховані. Вибірковий коефіцієнт регресії ξ на η: $a_1 = 0{,}46$, показує, що зі збільшенням ресурсів на 1 ум. од. час зростає на $0{,}46$ ум. од. Вибірковий коефіцієнт регресії η на ξ: $b_1 = 0{,}67$, показує, що зі збільшенням часу обробки програм на 1 ум. од. використання ресурсів зростає на $0{,}67$ ум. од.

Висновок. Для більш точного опису ступеня зв'язку між ξ і η треба побудувати іншу, нелінійну, регресійну модель.

Задача 10. Робота персонального комп'ютера описується ланцюгом Маркова з трьома станами: $s_1 =$ (робочий стан), $s_2 =$ (режим очікування), $s_3 =$ (сплячий режим). Задано матрицю перехідних імовірностей за один крок P:

$$P = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}3 \end{bmatrix} \end{array}.$$

Знайти: а) матрицю ймовірностей переходу за два кроки $P(2)$; б) граничні ймовірності (стаціонарний розподіл). Побудувати граф ЛМ.

Розв'язання: а) для отримання матриці $P(2)$ множимо матрицю P саму на себе:

$$P(2) = \begin{bmatrix} 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}3 \end{bmatrix} \cdot \begin{bmatrix} 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}3 \end{bmatrix} = \begin{bmatrix} 0{,}52 & 0{,}32 & 0{,}16 \\ 0{,}48 & 0{,}36 & 0{,}16 \\ 0{,}48 & 0{,}32 & 0{,}20 \end{bmatrix}.$$

Для контролю правильності підрахунків обов'язково перевірте, чи буде сума елементів кожного рядка матриці $P(2)$ дорівнювати 1 (*чому?*);

б) складаємо систему (1.83) для $m = 3$ і розв'язуємо її:

$$\begin{cases} b_1 + b_2 + b_3 = 1, \\ b_1 p_{11} + b_2 p_{21} + b_3 p_{31} = b_1, \\ b_1 p_{12} + b_2 p_{22} + b_3 p_{32} = b_2, \\ b_1 p_{13} + b_2 p_{23} + b_3 p_{33} = b_3 \end{cases} \Rightarrow \left| \begin{array}{c} \text{відкидаємо друге рівняння} \\ \text{і підставляємо } p_{ij} \text{ із матриці } P \end{array} \right| \Rightarrow$$

$$\Rightarrow \begin{cases} b_1 + b_2 + b_3 = 1, \\ 0{,}4b_1 + 0{,}2b_2 + 0{,}4b_3 = b_2, \\ 0{,}1b_1 + 0{,}2b_2 + 0{,}3b_3 = b_3 \end{cases} \Rightarrow \left| \begin{array}{c} \text{переносимо } b_2,\ b_3 \text{ із правих} \\ \text{частин рівнянь у ліві} \end{array} \right| \Rightarrow$$

$$\Rightarrow \begin{cases} b_1 + b_2 + b_3 = 1, \\ 0{,}4b_1 - 0{,}8b_2 + 0{,}4b_3 = 0, \\ 0{,}1b_1 + 0{,}2b_2 - 0{,}7b_3 = 0 \end{cases} \Rightarrow \left| \begin{array}{c} \text{спрощуємо систему і розв'язуємо} \\ \text{одним із відомих способів} \end{array} \right| \Rightarrow$$

$$\Rightarrow \begin{cases} b_1 + b_2 + b_3 = 1, \\ b_1 - 2b_2 + b_3 = 0, \\ b_1 + 2b_2 - 7b_3 = 0 \end{cases} \Rightarrow \left| \text{за правилом Крамера} \right| \Rightarrow$$

$$\Rightarrow b = (b_1, b_2, b_3) = \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6} \right).$$

Висновок. За тривалого часу роботи ПК (без зміни умов роботи) слід очікувати, що ймовірності станів s_1, s_2, s_3 розподіляться у відношенні $3:2:1$.

Для зображення ЛМ графом зробимо ліворуч і зверху матриці P написи, що відповідають станам ланцюга, які полегшують побудову графа (рис. 4.8):

$$P = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \!\!\begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} 0{,}5 & 0{,}4 & 0{,}1 \\ 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}3 & 0{,}4 & 0{,}3 \end{bmatrix} \end{array} \Rightarrow$$

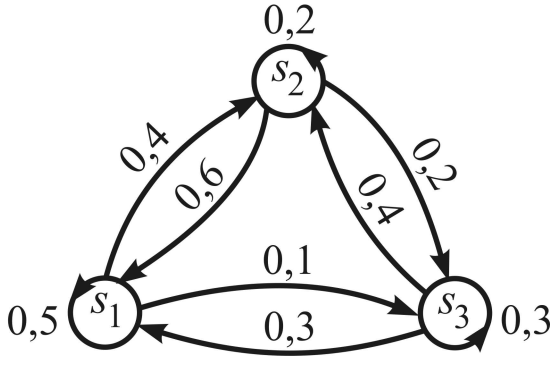

Рис. 4.8. **Граф ланцюга Маркова**

За допомогою графа легко простежити виконання властивості рядків стохастичної матриці: сума написів над орієнтованими ребрами, які виходять із вершин, дорівнює одиниці.

5. Рекомендована література

5.1. Основна

1. Вентцель Е. С. Теория вероятностей : учебник для вузов / Е. С. Вентцель. – 7-е изд., стер. – Москва : Высшая школа, 2001. – 576 с.

2. Вправи з курсу "Математика для економістів" для студентів усіх спеціальностей усіх форм навчання : у 3 ч. / укл. Е. Ю. Железнякова, А. В. Ігначкова, З. Г. Попова. – Ч. 3. – Харків : РВВ ХДЕУ, 2003. – 108 с.

3. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – Москва : Высшая школа, 2001. – 575 с.

4. Гмурман В. Е. Теория вероятностей и математическая статистика : учеб. пособ. для вузов / В. Е. Гмурман. – 6-е изд. – Москва : Высшая школа, 1998. – 480 с.

5. Денисова Т. В. Теорія ймовірностей та математична статистика : конспект лекцій / Т. В. Денисова, Г. К. Снурнікова, О. К. Шевченко. – Харків : Вид. ХДЕУ, 2003. – 132 с.

6. Егоршин А. А. Корреляционно-регрессионный анализ : пособ. для вузов / А. А. Егоршин, Л. М. Малярец. – Харьков : Основа, 1998. – 208 с.

7. Кремер Н. Ш. Теория вероятностей и математическая статистика / Н. Ш. Кремер. – Москва : ЮНИТИ-ДАНА, 2000. – 544 с.

8. Малярець Л. М. Математика для економістів. Теорія ймовірностей та математична статистика : у 3 ч. / Л. М. Малярець, І. Л. Лебедєва. – Ч. 3. – Харків : Вид. ХНЕУ, 2011. – 568 с.

9. Малярець Л. М. Теорія ймовірностей і математична статистика у вправах, прикладах та задачах : навч.-практ. посіб. / Л. М. Малярець, А. В. Ігначкова, Л. Д. Широкорад. – Харків : Вид. ХНЕУ, 2010. – 548 с.

10. Малярець Л. М. Теорія ймовірностей і математична статистика : навч. посіб. / Л. М. Малярець, І. Л. Лебедєва, Е. Ю. Железнякова та ін. – Харків : Вид. ХНЕУ, 2010. – 404 с.

5.2. Додаткова

11. Айвазян С. А. Прикладная статистика. Основы моделирования и первичная обработка данных : справ. изд. / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. – Москва : Финансы и статистика, 1983. – 472 с.

12. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посіб. : у 2 ч. Теорія ймовірностей / В. І. Жлуктенко, С. І. Наконечний. – Ч. 1. – Київ : КНЕУ, 2000. – 304 с.

13. Жлуктенко В. І. Теорія ймовірностей і математична статистика : навч.-метод. посіб. : у 2 ч. Математична статистика / В. І. Жлуктенко, С. І. Наконечний, С. С. Савіна. – Ч. 2. – Київ : КНЕУ, 2000. – 336 с.

14. Іващенко П. О. Багатовимірний статистичний аналіз / П. О. Іващенко, І. В. Семняк, В. В. Іванов. – Харків : Основа, 1992. – 144 с.

15. Ивченко Г. И. Математическая статистика : учеб. пособ. для втузов / Г. И. Ивченко, Ю. И. Медведев. – Москва : Высшая школа, 1984. – 248 с.

16. Карасев А. И. Курс высшей математики для экономических вузов / А. И. Карасев, З. И. Аксютина, Т. И. Савельева. – Москва : Высшая школа, 1982. – Ч. 2. – 320 с.

17. Минюк С. А. Математические методы и модели в экономике : учеб. пособ. / С. А. Минюк, Е. А. Ровба, К. К. Кузьмич. – Минск : ТетраСистемс, 2002. – 432 с.

18. Чернов В. П. Математика для экономистов. Теория массового обслуживания / В. П. Чернов, В. Б. Ивановский. – Москва : ИНФРА-М, 2000. – 158 с.

5.3. Інформаційні ресурси

19. Иглин С. П. Математические расчеты на базе MatLab. Математическая статистика [Электронный ресурс] / С. П. Иглин. – Режим доступа : iglin.exponenta.ru/matst.html.

20. Жлутков В. І. Теорія ймовірностей і математична статистика [Електронний ресурс] : навч.-метод. посіб. : у 2-х ч. Ч. II. Математична статистика / В. І. Жлутков, С. І. Наконечний, С. С. Савіна. – Київ : КНЕУ, 2001. – 336 с. – Режим доступу : dozkontrol.ucoz.ua/index/0-40.

21. Муштари Д. Х. Вероятность, математическая статистика, случайные процессы [Электронный ресурс] : учеб. пособ. / Д. Х. Муштари. – Режим доступа : old.kpfu.ru/infres/00-INT.pdf.

22. Глеч С. Г. Теорія ймовірностей та математична статистика [Електронний ресурс] / С. Г. Глеч, С. Ф. Ледяєв, І. В. Ольшанська. – Режим доступу : sevntu.com.ua/jspui/bitstream/123456789.

Покажчик позначень

$F_\xi(x)$ ($f_\xi(x)$) – функція (щільність) розподілу ВВ ξ – 26, 32

$\overset{\circ}{\xi}$ – центрована ВВ – 29

$M\xi$ – математичне сподівання ВВ ξ – 29

$D\xi$ – дисперсія ВВ ξ – 30

σ_ξ – середнє квадратичне відхилення ВВ ξ – 31

W – вибірка з генеральної сукупності – 39

m_i (w_i) – частота (частість) елемента x_i вибірки – 40

x_i^* – центр усереднення i-го інтервалу – 42

$F^*(x)$ ($f^*(x)$) – емпірична функція (щільність) розподілу – 43, 48

$\overline{x}_{\text{в}}$ – вибіркове середнє – 45

$D_{\text{в}}$, s^2 – вибіркова дисперсія – 45

s ($\sigma_{\text{в}}$) – вибіркове середнє квадратичне відхилення – 46

Mo^* (Me^*) – вибіркова мода (медіана) – 46

$\hat{D}_{\text{в}}$ – виправлена вибіркова дисперсія – 50

H_0 (H_a) – основна (альтернативна) гіпотеза – 51

$\chi^2_{\text{в}}$ – вибіркова статистика Пірсона – 52

$\chi^2_{\nu,\alpha}$ – критична точка розподілу χ^2 з числом ступенів свободи ν і рівнем значущості α – 52

$\xi=(\xi_1,\xi_2,\ldots,\xi_n)$ – випадковий вектор – 54

$M(\eta|\xi=x)$ – умовне математичне сподівання – 55

$\rho_{\xi\eta}$ (r_{xy}) – коефіцієнт (вибірковий коефіцієнт) кореляції – 56, 57

$\rho_{\xi/\eta}$ ($\rho_{\eta/\xi}$) – коефіцієнт регресії ξ на η (η на ξ) – 61

$\xi=\xi(t)$ – випадкова функція – 64

T – область визначення випадкового процесу – 64

$\rightleftharpoons$ – еквівалентність за означенням – 64

$\xi_{t_0}(\omega)$ ($\xi_{\omega_0}(t)$) – переріз (траєкторія) випадкового процесу в точці $t=t_0$ ($\omega=\omega_0$) – 64

$p_{ij}^{(n)}$ ($p_{ij}(n)$) – ймовірність переходу системи із стану i в стан j за один крок (за n кроків) – 65, 67

b_j ($j=\overline{0,m}$) – граничні ймовірності – 68

Додатки

Додаток А

Таблиця А.1

Таблиця значень функції Гаусса $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0395	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002
3,9	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001
4,0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001

Додаток Б

Таблиця Б.1

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

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2258	0,2291	0,2324	0,2356	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3906	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,1049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4274	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4430	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4648	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4700	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4762	0,4757
2,0	0,4772	0,4778	0,4783	0,4788	0,4796	0,4798	0,4803	0,4808	0,48Ї2	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4874	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4903	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4924	0,4927	0,4929	0,4930	0,1932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4958	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4973
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4980	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4986	0,4986	0,4987	0,4987	0,4988	0,4988	0,4988	0,4989	0,4989	0,4990
3,1	0,4990	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993
3,2	0,4993	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998	0,4998	0,4998	0,4998	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,5000	0,5000	0,5000	0,5000	0,5000
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Додаток В

Таблиця В.1

Таблиця значень функції Пуассона $\widetilde{P}_n(k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$

$k \backslash \lambda$	**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	0,1066
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

$k \backslash \lambda$	**1,0**	**2,0**	**3,0**	**4,0**	**5,0**	**6,0**	**7,0**	**8,0**	**9,0**
0	0,3679	0,1358	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	0,0158	0,0902	0,1680	0,1954	0,1755	0,1338	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	0,0588	0,1014	0,1241	0,1318
10			0,0008	0,0053	0,0181	0,0413	0,0710	0,0998	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	0,0263	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								0,0001	0,0003
21									0,0001
22									

Деякі найважливіші розподіли та їх числові характеристики

I. Дискретні розподіли

1. Біноміальний: розподіл імовірностей ВВ, значеннями якої є число k успіхів у n випробуваннях за схемою Бернуллі:

$$P_n(\xi = k) = C_n^k p^k q^{n-k},\ k = 0,1,2,...,n;\quad M\xi = np,\ D\xi = npq.$$

2. Пуассонівський: розподіл імовірностей ВВ, значеннями якої є число k успіхів у рідкісних подіях в n випробуваннях за схемою Бернуллі:

$$P_n(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda},\ \lambda = np > 0,\ k = 0, 1,...;\quad M\xi = \lambda,\ D\xi = \lambda.$$

3. Геометричний: розподіл імовірностей ВВ, значеннями якої є число випробувань k до першого успіху за схемою Бернуллі:

$$P(\xi = k) = q^k p,\ q = 1 - p,\ k = 0,1, ...;\ M\xi = q/p,\ D\xi = q/p^2.$$

4. Гіпергеометричний: розподіл імовірностей ВВ, значеннями якої є число m об'єктів, що мають певну властивість, серед n, вибраних із загального числа N об'єктів, серед яких M мають цю властивість:

$$P(\xi = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n},\ m \le \min(n,M);\quad M\xi = \frac{nM}{N},\ D\xi = \frac{n(M/N)(1-M/N)(N-n)}{N-1}.$$

II. Неперервні розподіли

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

$$M\xi = (a+b)/2,\ D\xi = (b-a)^2/12.$$

2. Показниковий
:
$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-\lambda x}, & x > 0 \end{cases};\ f(x) = \begin{cases} 0, & x \le 0 \\ \lambda e^{-\lambda x}, & x > 0 \end{cases}\ (\lambda > 0);$$

$$M\xi = 1/\lambda,\ D\xi = 1/\lambda^2.$$

3. Нормальний:
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}}dt;\quad f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}};$$

$$M\xi = a,\ D\xi = \sigma^2.$$

Додаток Д

Таблиця Д.1

Критичні точки розподілу χ^2 залежно від рівня значущості α та числа ступенів свободи ν

Число ступенів свободи ν	**Рівень значущості α**					
	0,01	**0,025**	**0,05**	**0,95**	**0,975**	**0,89**
1	6,6	5,0	3,8	0,0039	0,0009	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Зміст

НАВЧАЛЬНЕ ВИДАННЯ

Методичні рекомендації і завдання до виконання контрольних робіт з навчальної дисципліни "ТЕОРІЯ ЙМОВІРНОСТЕЙ, ІМОВІРНІСНІ ПРОЦЕСИ ТА МАТЕМАТИЧНА СТАТИСТИКА"

для студентів напряму підготовки
6.050101 "Комп'ютерні науки"
заочної форми навчання

Укладачі: **Сенчуков** Віктор Федорович
Денисова Тетяна Володимирівна

Відповідальний за видання *Л. М. Малярець*

Редактор *Н. І. Ганцевич*

Коректор *О. С. Новицька*

План 2016 р. Поз. № 34.
Підп. до друку 08.12.2016 р. Формат 60×90 1/16. Папір офсетний. Друк цифровий.
Ум. друк. арк. 7,0. Обл. -вид. арк. 8,75. Тираж 40 прим. Зам. № 285.

Видавець і виготовлювач – ХНЕУ ім. С. Кузнеця, 61166, м. Харків, просп. Науки, 9-А

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру
ДК № 4853 від 20.02.2015 р.