

РОЗДІЛ 6

РОЗПОДІЛИ ВИПАДКОВИХ ВЕЛИЧИН

6.1. Рівномірний розподіл

Рівномірний розподіл можуть мати як дискретні, так і неперервні випадкові величини.

Дискретна випадкова величина X має *рівномірний* розподіл, якщо ймовірності всіх її значень однакові. Якщо така величина може набувати n значень, то ймовірність кожного значення дорівнює $\frac{1}{n}$. Відповідний ряд розподілу має вигляд:

X	x_1	x_2	...	x_n
p	$\frac{1}{n}$	$\frac{1}{n}$...	$\frac{1}{n}$

Многокутник розподілу в цьому разі – це відрізок прямої, що паралельний осі абсцис із кінцями в точках $(1; \frac{1}{n})$ і $(n; \frac{1}{n})$.

Прикладом дискретної випадкової величини, яка має рівномірний розподіл, є кількість очок, які випадають при киданні гральної кості. Випадковою величиною, що має неперервний рівномірний розподіл, є, наприклад, похибка вимірювання, якщо результат заокруглюється до найближчого цілого значення.

Неперервна випадкова величина рівномірно розподілена на відрізку $[a; b]$, якщо її густина ймовірності на цьому відрізку дорівнює сталій c , а поза його межами – нулеві, тобто

$$f(x) = \begin{cases} 0, & \text{якщо } x < a, \ x > b, \\ c, & \text{якщо } a \leq x \leq b, \end{cases} \quad (1)$$

де стала c задовольняє умову нормування: $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Оскільки для рівномірного розподілу $f(x) = 0$, якщо $x \notin [a; b]$, то умова нормування набирас вигляду:

$$\int_a^b c dx = 1, \text{ тобто } c(b-a) = 1,$$

звідки $c = \frac{1}{b-a}$.

Отже, диференціальна функція рівномірного розподілу

$$f(x) = \begin{cases} 0, & \text{якщо } x < a, x > b, \\ \frac{1}{b-a}, & \text{якщо } a \leq x \leq b. \end{cases} \quad (2)$$

Графік функції $f(x)$ показано на рис. 6.1.

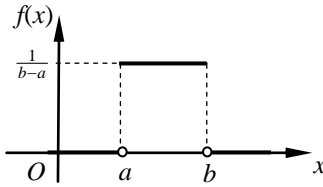


Рис. 6.1

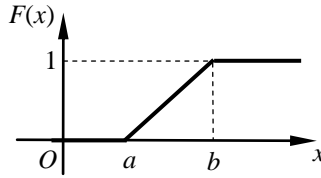


Рис. 6.2

Побудуємо інтегральну функцію $F(x)$ рівномірного розподілу:

$$F(x) = \int_{-\infty}^x f(x) dx.$$

Позаяк $f(x)$ залежить від проміжку зміни аргументу, то необхідно розглянути кожен проміжок зокрема.

Якщо $x \leq a$, то $F(x) = \int_{-\infty}^x 0 \cdot dx = 0$.

Якщо $a < x \leq b$, то $F(x) = F(a) + \int_a^x \frac{dx}{b-a} = \frac{x-a}{b-a}$.

Якщо $x > b$, то $F(x) = F(b) + \int_b^x 0 \cdot dx = 1 + 0 = 1$.

Графік функції $F(x)$ зображено на рис. 6.2.

Як бачимо, в точках a і b функція $F(x)$ неперервна, але не гладка, тому похідні $F'(a)$ і $F'(b)$ не існують.

Ймовірність того, що рівномірно розподілена випадкова величина потрапляє в інтервал $[c; d] \subset [a; b]$

$$P(c < X < d) = \int_c^d \frac{dx}{b-a} = \frac{d-c}{b-a}.$$

Відшукаймо основні числові характеристики випадкової величини, що має рівномірний розподіл. Математичне сподівання

$$M(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{a+b}{2}.$$

Позаяк розподіл симетричний, то медіана величини X збігається з її середнім значенням (з математичним сподіванням):

$$\text{Me}X = \frac{a+b}{2}.$$

Моди рівномірний розподіл не має.

Дисперсія і середнє квадратичне відхилення:

$$D(X) = \int_a^b \left(x - \frac{a+b}{2} \right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}, \quad \sigma_X = \frac{\sqrt{3}(b-a)}{6}.$$

Нескладно відшукати характеристики форми графіка розподілу – скошеність та ексцес. Зробіть це самотужки.

6.2. Біномний розподіл

Біномним називають розподіл імовірностей кількості m появ події в n незалежних випробуваннях, у кожному з яких імовірність появи події стала і дорівнює p . Значеннями випадкової величини X є кількість появ події в n випробуваннях, тобто $X = m$, де $m = \overline{0, n}$. Ймовірність можливих значень випадкової величини визначають за формулою Бернуллі. Отже, *біномний закон розподілу* має вигляд

$$P(X = m) = C_n^m p^m (1-p)^{n-m},$$

де сталі p й n називають *параметрами розподілу*, $1-p = q$.

Біномний розподіл описує лише дискретну випадкову величину. Його можна задати рядом розподілу:

X	0	1	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p^1 q^{n-1}$...	$C_n^k p^k q^{n-k}$...	$C_n^n p^n q^0$

Інтегральна функція біномного розподілу

$$F(m) = \begin{cases} 0, & m \leq 0, \\ \sum_{n < m} P_n(m), & 0 < m \leq n, \\ 1, & m > n. \end{cases}$$

Відшукаймо числові характеристики біномного закону розподілу.

Випадкова величина X – кількість появ події A в n випробуваннях – є сумою n незалежних випадкових величин X_i :

$$X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i,$$

де кожен із доданків

$$X_i = \begin{cases} 1, & \text{якщо подія } A \text{ наступила в } i\text{-му випробуванні,} \\ 0, & \text{якщо подія } A \text{ не наступила в } i\text{-му випробуванні.} \end{cases}$$

Ряд розподілу випадкової величини X_i має вигляд

X	1	0
P	p	q

Математичне сподівання кількості появ події A в одному випробуванні

$$M(X_i) = \sum_{k=1}^2 (x_i)_k p_k = 1 \cdot p + 0 \cdot q = p.$$

За властивістю математичного сподівання суми незалежних випадкових величин

$$M(X) = M\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n M(X_i) = \sum_{i=1}^n p = np.$$

Відшукаймо дисперсію випадкової величини X_i . За означенням

$$\begin{aligned} D(X_i) &= \sum_{k=1}^2 ((x_i)_k - p)^2 p_k = (1-p)^2 p + (0-p)^2 (1-p) = \\ &= p(1-p)(1-p+p) = pq. \end{aligned}$$

За теоремою про дисперсію суми незалежних випадкових величин

$$D(X) = D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i) = \sum_{i=1}^n pq = npq.$$

Якщо $n \rightarrow \infty$, то біномний розподіл наближається до нормального.

На рис. 6.3 подано многокутники розподілу для $n = 20$ і $p = 0,3; 0,5$.
Якщо $p = q = 0,5$, то многокутник розподілу симетричний.

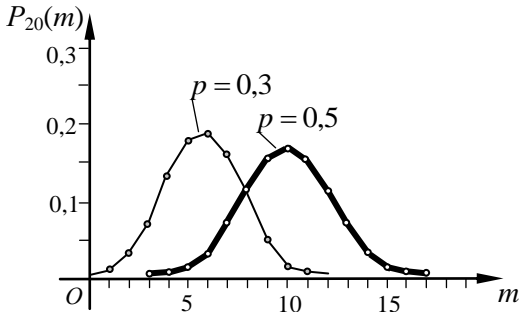


Рис. 6.3

Найбільшого значення ймовірність досягає тоді, коли $m = m_0$, де m_0 задовольняє умову:

$$np - q \leq m_0 \leq np + p.$$

Пригадаймо: m_0 – це найімовірніша кількість появ події в n випробуваннях, тобто мода біномного розподілу.

6.3. Геометричний розподіл

Відшукаймо ймовірність того, що подія A в серії випробувань перший раз з'явиться в m -у випробуванні, тобто після $m-1$ появ події \bar{A} . Випадковою величиною X є кількість випробувань, внаслідок яких вперше появилася A . Очевидно, X може набувати значень $1, 2, \dots, n, \dots$. Побудуємо ряд розподілу цієї величини.

За теоремою множення ймовірностей ймовірність того, що випадкова величина набула значення m ,

$$P(X = m) = P(m, p) = (1 - p)^{m-1} p.$$

Отже, ряд розподілу залежить лише від одного параметра p і має вигляд ($q = 1 - p$):

X	1	2	3	...	m	...
P	p	qp	$q^2 p$...	$q^{m-1} p$...

Цей ряд розподілу називають *геометричним розподілом* випадкової величини. Прикладом геометрично розподіленої величини є кількість пострілів у ціль до першого влучення (за умови, що ймовірність влучення в кожному пострілі та сама).

Скориставшись формулою суми членів нескінченної спадної геометричної прогресії, просто пересвідчитися, що сума ймовірностей дорівнює 1.

Можна довести, що математичне сподівання й дисперсія випадкової величини, яка має геометричний розподіл, визначаються формулами:

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

Кількість випробувань вважається необмеженою. Однак якщо вона достатньо велика, то ці формули дають добрі наближення.

6.4. Розподіл Пуассона

Розподіл Пуассона характеризує дискретну випадкову величину. Він трапляється в задачах, пов'язаних із *найпростішим потоком подій*. Потік подій – це деяка послідовність подій, що настають одна за одною у випадкові моменти часу зі сталою інтенсивністю. Під *інтенсивністю* ν розуміють середню кількість появ події за одиницю часу.

Вважатимемо, що потік подій має такі властивості:

1) ймовірність появи деякої кількості подій упродовж заданого проміжку часу залежить лише від довжини проміжку і не залежить від його положення на часовій осі (властивість *стаціонарності*);

2) ймовірність появи більше ніж однієї події упродовж досить малого проміжку часу практично дорівнює нулеві (властивість *ординарності*);

3) ймовірність появи деякої кількості подій упродовж фіксованого проміжку часу не залежить від кількості подій, що появилися в інші проміжки часу (властивість *відсутності післядії*).

Найпростіший потік подій – це потік, який задовольняє властивості стаціонарності, ординарності і відсутності післядії. Найпростіший потік подій називають також *пуассонівським потоком*.

Можна довести: ймовірність того, що за проміжок часу t за сталої інтенсивності ν подія появиться m разів, визначається формулою:

$$P_t(m) = \frac{(\nu t)^m}{m!} e^{-\nu t} = \frac{\lambda^m}{m!} e^{-\lambda}. \quad (1)$$

Величина $\lambda = \nu t$ – це середня кількість появ події за час t .

Формула (1) справедлива й тоді, коли події розподілені зі сталою густиною в деякій просторовій області (тоді ν – математичне сподівання кількості подій). Тому, незалежно від конкретних міркувань, що приводять до формули (1), можна розглядати розподіл імовірностей величини X , яка набуває значень $x = 0, 1, 2, \dots, m, \dots$ з імовірностями

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}. \quad (2)$$

Ця формула – закон розподілу Пуассона.

Ряд розподілу випадкової величини, що має розподіл Пуассона, виглядає так:

X	0	1	2	...	m	...
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2} e^{-\lambda}$...	$\frac{\lambda^m}{m!} e^{-\lambda}$...

Позаяк $e^\lambda = \sum_{m=0}^{\infty} \frac{\lambda^m}{m!}$, то зрозуміло, що ряд імовірностей збігається:

$$p_0 + p_1 + p_2 + \dots + p_m + \dots = e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} \cdot e^\lambda = 1.$$

Математичне сподівання і дисперсія випадкової величини з розподілом Пуассона однакові і дорівнюють параметрові розподілу λ :

$$M(X) = D(X) = \lambda. \quad (3)$$

Випадкові величини, які підлягають закономірному розподілу Пуассона, часто трапляються на практиці. Це, зокрема, кількість альфа-частинок, що виникають при розпаді радіоактивного елемента, кількість електронів, що вилітають із поверхні розжареного металу, кількість виробів, що зазнають пошкоджень при транспортуванні, кількість n_k клітин із k зміненими внаслідок рентгенівського опромінення хромосомами, прихід покупців у магазин, помилки на сторінці тексту, затримки в русі потягів на деякій ділянці шляху, дефекти виробів у партії тощо.

6.5. Нормальний закон розподілу

Досі ми розглядали розподіли дискретних випадкових величин (за винятком рівномірного неперервного розподілу). Нормальний розподіл характеризує лише неперервну випадкову величину. Попри те що на практиці здебільшого маємо справу з дискретними випадковими вели-

чинами (дані експерименту, результати спостереження тощо), розподіл більшості з них добре апроксимується нормальним розподілом. Широке розповсюдження нормального закону розподілу пояснює центральна гранична теорема, про яку мова далі (п. 6.10). Згідно з цією теоремою інші розподіли за певних умов наближаються до нормального.

6.5.1. Означення нормального закону розподілу. *Нормальний розподіл* – це розподіл неперервної випадкової величини X , який описується густиною ймовірності

$$f_N(x, a, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \equiv \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad (1)$$

де a й σ – параметри розподілу, до того ж $\sigma > 0$, а $-\infty < x < \infty$ (зміст параметрів розподілу буде з'ясовано далі).

Часто нормальний розподіл із параметрами a і σ позначають $N(a, \sigma)$, а його густина ймовірності – $f_N(x)$.

Нескладно пересвідчитися, що

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

Справді,

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (3)$$

Для відшукування інтеграла зробимо заміну змінної

$$t = \frac{x-a}{\sigma}, \text{ звідки } x = a + \sigma t \text{ і } dx = \sigma dt. \quad (4)$$

Скориставшись добре відомим інтегралом Ойлера – Пуассона

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}, \quad (5)$$

дістанемо:

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{\sigma\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1.$$

Особливість нормального закону розподілу полягає в тому, що, як уже мовилося, він є граничним законом, до якого наближаються інші закони розподілу.

Аналіз свідчить, що найважливішою умовою виникнення нормального розподілу ознаки є формування цієї ознаки як суми великої кількості взаємно незалежних доданків (факторів). При цьому жоден із доданків не є визначальним порівняно з іншими, тобто його дисперсія не є значно більшою від дисперсій решти доданків. Строго це питання розв'язує так зв. центральна гранична теорема Ляпунова. Докладніше про цю теорему мова далі.

6.5.2. Ймовірнісний зміст параметрів нормального розподілу. Щоб з'ясувати зміст параметрів a і σ нормального розподілу, відшукаймо математичне сподівання і дисперсію випадкової величини X .

За означенням математичного сподівання з огляду на заміну змінної (4) дістанемо:

$$\begin{aligned} M(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (a + \sigma t) e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} a \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \sigma \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt. \end{aligned}$$

Перший інтеграл у правій частині цієї рівності – це інтеграл Ойлера – Пуассона, а другий, як нескладно пересвідчитися, дорівнює нулеві. Отже,

$$M(X) = \frac{1}{\sqrt{2\pi}} \cdot a \cdot \sqrt{2\pi} = a, \quad (6)$$

тобто параметр a має зміст математичного сподівання.

Аналогічно можна довести, що дисперсія нормально розподіленої випадкової величини дорівнює σ^2 :

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2,$$

тобто параметр σ нормального розподілу є середнім квадратичним відхиленням.

Оскільки густина ймовірностей (1) є парна функція, то крива нормального розподілу симетрична щодо прямої $x = a$, тобто точка $x = a$ є водночас медіаною. У точці $x = a$ функція $f(x)$ набуває найбільшого значення, тобто значення $x = a$ – це мода випадкової величини. Отже, математичне сподівання, мода й медіана нормального розподілу збігаються.

Точку $x = a$ називають *центром розподілу ймовірностей* або *центром розсіяння* випадкової величини.

Крива асимптотично наближається до осі абсцис, якщо $x \rightarrow \pm\infty$.

Графік розподілу щільності ймовірностей (1) – гладка дзвоноподібна крива. На рис. 6.7 подано криві нормального розподілу для $a = 0$ і різних значень дисперсії. Найбільше значення щільності ймовірності дорівнює

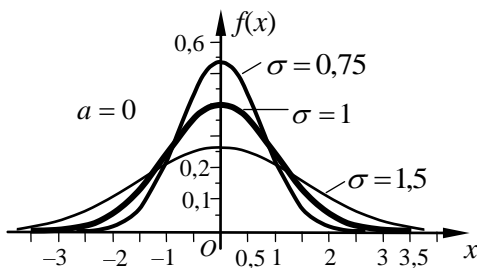


Рис. 6.7

$\frac{1}{\sigma\sqrt{2\pi}}$. Це значення змен-

шується зі збільшенням параметра σ . Але при цьому, відповідно до рівності (2), площа під кривою розподілу залишається рівною одиниці. Тому зі збільшенням дисперсії крива розподілу опускається донизу і розтягується.

Якщо в формулі (1) $a = 0$ і $\sigma = 1$, то розподіл називають *стандартним нормальним розподілом*. Значення функції щільності ймовірності стандартного нормального розподілу табульовано (додаток 1). Докладніше цей розподіл розглянемо в наступному параграфі.

6.5.3. Асиметрія та ексцес. Як ми знаємо, мірою асиметрії (скошеності) форми кривої розподілу є коефіцієнт асиметрії (п. 5.7):

$$A_S = \frac{\mu_3}{\sigma^3}, \quad (7)$$

де третій центральний момент $\mu_3 = M((x - m_X)^3)$, $m_X = a$. Крива нормального розподілу симетрична щодо прямої $x = a$, тому $\mu_3 = 0$ і $A_S = 0$.

Ексцес кривої розподілу визначають за формулою (п. 5.7):

$$E_S = \frac{\mu_4}{\sigma^4} - 3, \quad (8)$$

де $\mu_4 = M((x - a)^4)$ – четвертий центральний момент. Можна довести, що для нормального розподілу $\mu_4 = 3\sigma^4$, тому $E_S = 0$.

6.5.4. Ймовірність потрапляння нормально розподіленої випадкової величини в заданий інтервал. Як ми знаємо, ймовірність того, що неперервна випадкова величина зі щільністю ймовірності $f(x)$ потрапить у проміжок $[\alpha; \beta]$ визначають за формулою (п. 4.4):

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx.$$

Для нормального розподілу зі щільністю ймовірності (1) ця формула набирає вигляду:

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (9)$$

Після заміни змінної (4) з огляду на те що $t = \frac{\alpha-a}{\sigma}$, якщо $x = \alpha$, і $t = \frac{\beta-a}{\sigma}$, якщо $x = \beta$, останнє співвідношення набирає вигляду:

$$P(\alpha < X < \beta) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt.$$

Нескладно зрозуміти: ймовірність того, що нормально розподілена випадкова величина $N(a, \sigma)$ потрапить у проміжок $[\alpha; \beta)$,

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right), \quad (10)$$

де $\Phi(x)$ – функція Лапласа (інтеграл імовірностей), що визначається формулою (9). Значення функції Лапласа табульовано (додаток 2).

Приклад 1. Діаметр d кульки підшипника можна розглядати як випадкову величину, що має нормальний закон розподілу. Кульку вважають стандартною, якщо вона проходить через отвір із діаметром $d_2 = 3,63$ мм і не проходить через отвір із діаметром $d_1 = 3,57$ мм. Якщо не справджується одна з цих вимог, то кульку вважають бракованою. Відшукаймо дисперсію діаметра кульки, якщо брак складає 5 %.

Розв’язання. Середнє значення (математичне сподівання) діаметра

$$m_d = \frac{d_1 + d_2}{2} = \frac{3,57 + 3,63}{2} = 3,60 \text{ (мм)}. \quad (a)$$

Ймовірність того, що кульку не буде забраковано, знаходимо за формулою (10):

$$P = P(d_1 < d \leq d_2) = \Phi\left(\frac{d_2 - m_d}{\sigma_d}\right) - \Phi\left(\frac{d_1 - m_d}{\sigma_d}\right).$$

З огляду на співвідношення (a) остання формула набирає вигляду:

$$P(d_1 < d \leq d_2) = \Phi\left(\frac{d_2 - d_1}{2\sigma_d}\right) - \Phi\left(\frac{d_1 - d_2}{2\sigma_d}\right).$$

Беручи до уваги непарність функції Лапласа, дістаємо

$$P(d_1 < d \leq d_2) = 2\Phi\left(\frac{d_2 - d_1}{2\sigma_d}\right).$$

За умовою задачі ця ймовірність дорівнює 0,95. Тому

$$2\Phi\left(\frac{d_2 - d_1}{2\sigma_d}\right) = 0,95, \quad \Phi\left(\frac{d_2 - d_1}{2\sigma_d}\right) = 0,475.$$

За таблицею значень функції Лапласа (додаток 2) знаходимо:

$$\frac{d_2 - d_1}{2\sigma_d} = 1,96, \quad \text{звідки} \quad \sigma_d = \frac{d_2 - d_1}{2 \cdot 1,96} = 0,0153 \text{ (мм)}.$$

Отже, дисперсія діаметра кульки $\sigma_d^2 = 0,00023 \text{ мм}^2$. ►

6.5.5. Ймовірність заданого відхилення нормально розподіленої випадкової величини. Правило трьох сигм. Нехай випадкова величина X характеризується нормальним законом розподілу. Відшукаймо ймовірність того, що її відхилення од математичного сподівання менше від заданого додатного числа ε , тобто $P(|X - a| < \varepsilon)$.

Запишімо нерівність $|X - a| < \varepsilon$ у вигляді $-\varepsilon < X - a < \varepsilon$, або

$$a - \varepsilon < X < a + \varepsilon.$$

Тоді за формулою (10)

$$\begin{aligned} P(|X - a| < \varepsilon) &= P(a - \varepsilon < X < a + \varepsilon) = \\ &= \Phi\left(\frac{a + \varepsilon - a}{\sigma}\right) - \Phi\left(\frac{a - \varepsilon - a}{\sigma}\right) = \Phi\left(\frac{\varepsilon}{\sigma}\right) - \Phi\left(\frac{-\varepsilon}{\sigma}\right). \end{aligned}$$

Позаяк функція Лапласа непарна, то

$$P(|X - a| < \varepsilon) \equiv P(a - \varepsilon < X < a + \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right). \quad (11)$$

Як видно з формули (11), ймовірність потрапляння в заданий інтервал більша для тої величини, дисперсія котрої менша. Геометричний зміст цього висновку ілюструє рис. 6.7. Справді, як ми знаємо, ймовірність потрапляння випадкової величини в заданий інтервал чисельно дорівнює площі відповідної криволінійної трапеції. Якщо розглянути, для прикладу, інтервал $(-1; 1)$, то, як видно з рисунка, більшу площу має та трапеція, яка обмежена кривою, що відповідає розподілові з меншою дисперсією.

Приклад 2. Випадкова величина X має нормальний розподіл. Математичне сподівання і середнє квадратичне відхилення дорівнюють відповідно 20 і 10. Відшукаймо ймовірність того, що відхилення випадкової величини від математичного сподівання за абсолютною величиною буде менше 3.

Розв'язання. За формулою (11)

$$P(20 - 3 < X < 20 + 3) = 2\Phi\left(\frac{3}{10}\right).$$

З огляду на те що $\Phi(0,3) = 0,1179$ (додаток 2), знаходимо:

$$P(17 < X < 23) = 2 \cdot 0,1179 \approx 0,24. \blacktriangleright$$

На практиці за одиницю вимірювання відхилення часто вибирають середнє квадратичне відхилення σ . Зміна параметра a не змінює форми кривої (не змінює характеру розсіювання випадкової величини). Ця зміна лише визначає величину зсуву кривої розподілу ліворуч або праворуч. Тому без втрати загальності можна вважати, що $a = 0$ (крива розподілу симетрична щодо осі Oy). Тоді за формулою (11) дістанемо:

$$P(-\sigma < X < \sigma) = 2\Phi(\sigma/\sigma) = 2\Phi(1) = 2 \cdot 0,3413 \approx 0,683,$$

$$P(-2\sigma < X < 2\sigma) = 2\Phi(2) = 2 \cdot 0,4772 \approx 0,954,$$

$$P(-3\sigma < X < 3\sigma) = 2\Phi(3) = 2 \cdot 0,49865 \approx 0,997.$$

Як бачимо, майже неможливо, що нормально розподілена випадкова величина відхилиться од математичного сподівання більше ніж на 3σ . Це припущення називають *правилом трьох сигм*. На підставі цього правила можна оцінити середнє квадратичне відхилення. Для цього з ряду спостережень треба вибрати найбільше та найменше значення і поділити на шість їхню різницю. Отримане число – це груба оцінка σ , за умови, що ознака підлягає нормальному законові розподілу.

6.6. Стандартний нормальний розподіл

Диференціальну функцію нормально розподіленої випадкової величини X табулювати складно, позаяк вона залежить від двох параметрів – від a й σ . Тому розглядають нормовану (стандартизовану) випадкову величину

$$Z = \frac{X - a}{\sigma}. \quad (1)$$

Можна довести, що щільність розподілу ймовірностей величини (1)

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

Формула (2) – це щільність нормального розподілу з параметрами $a = 0$ і $\sigma = 1$, які мають зміст математичного сподівання і дисперсії нормованої випадкової величини Z . Розподіл, заданий формулою (2), називають *стандартним нормальним законом розподілу*.

Графік функції $\varphi(z)$ називають *кривою Гаусса* (рис. 6.7).

Значення функції $\varphi(z)$ табульовано (додаток 1). За значеннями функції $\varphi(z)$ можна знайти значення функції $f_N(x, a, \sigma)$:

$$f_N(x, a, \sigma) = \varphi\left(\frac{x-a}{\sigma}\right) \cdot \frac{1}{\sigma}.$$

6.7. Розподіли, що пов'язані з нормальним розподілом

Для побудови деяких статистичних моделей використовуються закони розподілу, що пов'язані з нормальним і по суті є розподілами деяких функцій від нормально розподілених випадкових величин.

6.7.1. Розподіл Пірсона (розподіл хі-квадрат). Нехай X_1, X_2, \dots, X_n – незалежні нормально розподілені випадкові величини, математичне сподівання кожної з яких дорівнює нулеві, а дисперсія – одиниці. Розгляньмо суму квадратів цих величин:

$$\chi^2 = \sum_{i=1}^n X_i^2. \quad (1)$$

Доведено, що величина, визначена формулою (1), має χ^2 -розподіл із k ступенями вільності. Функція щільності ймовірності $f_{\chi^2}(x, k)$ величини

χ^2 доволі складна, тому не розглядаємо її. Але важливо, що розподіл χ^2 залежить лише від одного параметра – кількості ступенів вільності k , яка дорівнює кількості незалежних доданків у сумі (1). У разі наявності l зв'язків між величинами X_i , які не ведуть до порушення нормального закону розподілу, кількість незалежних величин зменшується, а отже, зменшується й кількість ступенів вільності суми квадратів: вона дорівнює

$k = n - l$. Графіки розподілу Пірсона залежно від кількості ступенів вільності показано на рис. 6.12.

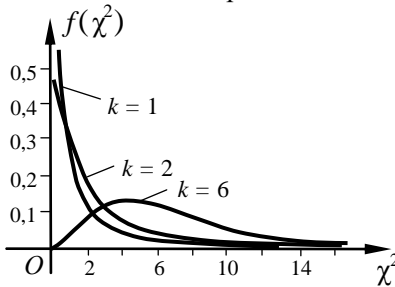


Рис. 6.12

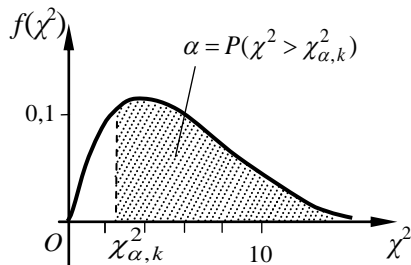


Рис. 6.13

Можна довести, що математичне сподівання і дисперсія випадкової величини $\chi^2(k)$ дорівнюють відповідно

$$M(\chi^2(k)) = k, \quad D(\chi^2(k)) = 2k. \quad (3)$$

Диференціальну функцію розподілу хі-квадрат інтегрувати складно. Тому складено таблиці, де вказано критичні точки розподілу випадкової величини χ^2 (додаток 3). Критичні точки хі-квадрат розподілу – це такі граничні (найменші) значення $\chi^2_{k,\alpha}$, які випадкова величина, що має χ^2 -розподіл із відомою кількістю ступенів вільності k , перевищить з імовірністю α (рис. 6.13). Отже, $\alpha = P(\chi^2 > \chi^2_{\alpha,k})$.

Ймовірність α в цьому разі іноді називають *рівнем значущості*.

Як видно з рис. 6.12, вигляд кривої щільності ймовірностей χ^2 -розподілу істотно залежить від кількості ступенів вільності k . На рис. 6.13 схематично зображено криву розподілу для $k = 10$.

Якщо $k \rightarrow \infty$, то розподіл хі-квадрат прямує до нормального розподілу з параметрами $a = k$, $\sigma = \sqrt{2k}$. Якщо $k \geq 30$, то цей розподіл уже можна апроксимувати нормальним розподілом.

6.7.2. Розподіл Стьюдента (t -розподіл). Нехай $k + 1$ незалежних випадкових величин X, X_1, X_2, \dots, X_k розподілені нормально з параметрами

$a = 0, \sigma^2 = 1$. Тоді, як відомо, величина $\chi^2 = \frac{1}{k} \sum_{i=1}^k X_i^2$ має розподіл хі-квадрат із k ступенями вільності. Доведено, що випадкова величина

$$t = \frac{X}{\sqrt{\frac{1}{k} \sum_{i=1}^k X_i^2}} = \frac{X}{\sqrt{\frac{\chi^2}{k}}} \quad (4)$$

має розподіл Стюдента з k ступенями вільності. Вигляд функції щільності ймовірності $f_S(x)$ величини t складний. Розподіл Стюдента, як і розподіл Пірсона, залежить від єдиного параметра – кількості ступенів вільності k .

Для розподілу випадкової величини t укладено таблиці (додаток 4).

Математичне сподівання і дисперсія випадкової величини t дорівнюють відповідно

$$M(t(k)) = 0, \quad D(t(k)) = \frac{k}{k-2}, \quad k > 2. \quad (6)$$

Очевидно, що

$$\lim_{n \rightarrow \infty} D(t(k)) = \lim_{n \rightarrow \infty} \frac{k}{k-2} = 1.$$

Крива t -розподілу симетрична щодо осі ординат. На рис. 6.14 схематично показані крива t -розподілу (для $k = 2$) та крива стандартного нормального розподілу. Якщо $k \rightarrow \infty$, то розподіл Стюдента прямує до стандартного нормального розподілу, і вже для $k \geq 25$ його заміняють цим розподілом.

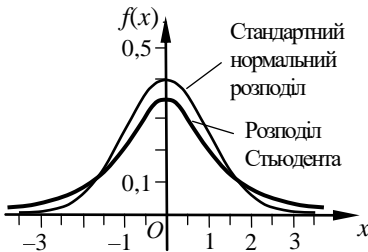


Рис. 6.14

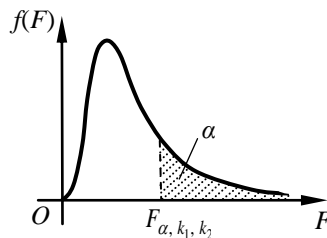


Рис. 6.15

6.7.3. Розподіл Фішера-Снедекора (F -розподіл). Нехай задано $k_1 + k_2$ незалежних нормально розподілених випадкових величин $X_1, X_2, \dots, X_{k_1}, Y_1, Y_2, \dots, Y_{k_2}$ з нульовими математичними сподіваннями й одиничними дисперсіями. Тоді, як ми знаємо, величини

$$\chi_1^2 = \sum_{i=1}^{k_1} X_i^2 \quad \text{та} \quad \chi_2^2 = \sum_{i=1}^{k_2} Y_i^2$$

мають χ^2 -розподіли з k_1 та k_2 ступенями вільності відповідно.

Статистика (випадкова величина)

$$F = \frac{\chi_1^2}{k_1} / \frac{\chi_2^2}{k_2} \quad (7)$$

має розподіл Фішера – Снедекора з k_1 і k_2 ступенями вільності.

Функція розподілу щільності ймовірностей $f_F(k_1, k_2, x)$ випадкової величини F залежить лише від кількостей ступенів вільності k_1 та k_2 . На рис. 6.15 схематично зображена крива розподілу.

Для випадкової величини F складено таблиці (додаток 6), в яких різним поєднанням кількостей ступенів вільності чисельника k_1 і знаменника k_2 відповідають такі значення F_{α, k_1, k_2} , для яких справджується рівність $P(F > F_{\alpha, k_1, k_2}) = \alpha$ (рис. 6.15).

6.8. Показниковий розподіл

Розглянемо найпростіший (пуасонівський) потік подій з інтенсивністю λ . Введемо неперервну невід’ємну випадкову величину T – проміжок часу між двома появами подій. Побудуємо функцію розподілу цієї величини

$$F(t) = P(T < t).$$

Ймовірність протилежної події ($T \geq t$) дорівнює ймовірності того, що в проміжку часу $(0, t)$ не наступить жодна з подій потоку, тобто

$$P(T \geq t) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}. \quad (1)$$

Отже, функція розподілу

$$F(t) = 1 - e^{-\lambda t}, \quad \lambda > 0, t \geq 0. \quad (2)$$

Закон розподілу випадкової величини T , що описується функцією (2), називають *показниковим*, або *експоненційним*.

Диференціальна функція показникового розподілу

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{якщо } t \geq 0, \\ 0, & \text{якщо } t < 0. \end{cases} \quad (3)$$

Показниковий розподіл залежить від одного параметра λ .

На рис. 6.16 зображені графіки інтегральної та диференціальної функцій показникового розподілу для значення параметра $\lambda = 0,8$.

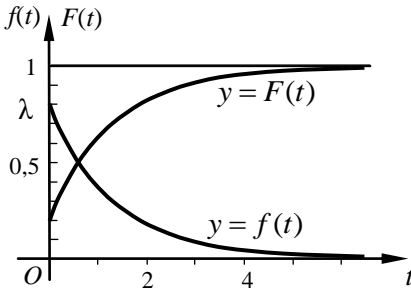


Рис. 6.16

Відшукаймо числові характеристики випадкової величини T , що має показниковий розподіл.

За означенням математичного сподівання

$$M(T) = \lambda \int_0^{+\infty} t e^{-\lambda t} dt.$$

Інтегруючи частинами і враховуючи, що $\lim_{t \rightarrow +\infty} t e^{-\lambda t} = 0$, матимемо:

$$\int_0^{+\infty} t e^{-\lambda t} dt = -\frac{1}{\lambda} t e^{-\lambda t} \Big|_0^{+\infty} + \frac{1}{\lambda} \int_0^{+\infty} e^{-\lambda t} dt = -\frac{1}{\lambda^2} e^{-\lambda t} \Big|_0^{+\infty} = \frac{1}{\lambda^2}.$$

Отже,

$$M(T) = 1/\lambda,$$

тобто математичне сподівання показникового розподілу обернене параметрові λ . λ — це кількість подій за одиницю часу, $1/\lambda$ — це середній проміжок часу між двома послідовними подіями.

Дисперсія

$$D(T) = \int_{-\infty}^{+\infty} t^2 f(t) dt - M^2(T) = \lambda \int_0^{+\infty} t^2 e^{-\lambda t} dt - \frac{1}{\lambda^2}.$$

Двічі інтегруючи частинами, отримаємо:

$$D(T) = \frac{1}{\lambda^2}, \quad \text{звідки} \quad \sigma(T) = \frac{1}{\lambda}. \quad (4)$$

Як бачимо, середнє квадратичне відхилення випадкової величини, розподіленої за показниковим законом, дорівнює її математичному сподіванню.

Формулу (4) можна використати для оцінювання параметра λ за дослідними даними: якщо \bar{T} — середній проміжок часу між двома подіями, то за наближене значення λ природно взяти величину $1/\bar{T}$. Якщо потік спостерігають упродовж часу τ і подія настала n разів, то

$$\lambda \approx \frac{n}{\tau} \quad (5)$$

(тобто оцінкою параметра λ буде число $\lambda^* = n/\tau$).

Грунтуючись на означенні показникового розподілу, можна довести, що ймовірність проміжку часу T між появами подій найпростішого потоку не залежить від тривалості попередніх проміжків часу. Ця властивість – це властивість відсутності післядії, про яку вже мовилося.

Розгляньмо цю властивість докладніше. Нехай подія

$$A = \{\text{безвідмовна робота пристрою в інтервалі } (0; t_1)\},$$

а подія

$$B = \{\text{безвідмовна робота пристрою в інтервалі } (t_1; t_1 + t)\}.$$

Тоді подія

$$AB = \{\text{безвідмовна робота пристрою в інтервалі } (0; t_1 + t)\}.$$

Нехай випадкова величина T – час безвідмовної роботи – має показниковий розподіл і описується функцією (2). Тоді

$$P(A) = P(T > t_1) = 1 - (1 - e^{-\lambda t_1}) = e^{-\lambda t_1},$$

$$P(AB) = P(T > t_1 + t) = e^{-\lambda(t_1 + t)}.$$

Умовна ймовірність того, що пристрій працюватиме безвідмовно в інтервалі часу $(t_1; t_1 + t)$ за умови, що він працював безвідмовно в інтервалі $(0; t_1)$:

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{e^{-\lambda(t_1 + t)}}{e^{-\lambda t_1}} = e^{-\lambda t}.$$

Як бачимо, умовна ймовірність $P(B|A)$ не залежить від t_1 , а лише від t .

Це означає, що час безвідмовної роботи пристрою в інтервалі $(0; t_1)$ не впливає на величину ймовірності його безвідмовної роботи впродовж інтервалу часу $(t_1; t_1 + t)$, і залежить лише від довжини t цього інтервалу, тобто володіє властивістю відсутності післядії.

Експоненційний розподіл широко застосовують у теорії надійності, яка вивчає умови безвідмовної роботи систем, якщо відмови в роботі цих систем утворюють найпростіший потік. А саме: як щойно доведено, формула (1) може описувати ймовірність того, що впродовж часу t система працює безвідмовно. Тому функцію

$$R(t) = P(T \geq t) = 1 - F(t) = e^{-\lambda t} \quad (6)$$

називають *функцією надійності*, а параметр λ – *інтенсивністю відмов*.

Використовуючи її, нескладно знайти ймовірність того, що час безвідмовної роботи системи знаходиться в інтервалі $(t_1; t_2)$. Справді,

$$P(t_1 \leq T \leq t_2) = F(t_2) - F(t_1),$$

звідки, беручи до уваги формулу (2), отримуємо:

$$P(t_1 \leq T \leq t_2) = e^{-\lambda t_1} - e^{-\lambda t_2}. \quad (7)$$

Якщо задати ймовірність $P(T \geq t) = p$, то можна знайти час, упродовж якого з імовірністю p система працюватиме безвідмовно:

$$t = -\ln p / \lambda. \quad (8)$$

Приклад. Упродовж 100 годин роботи системи зареєстровано 5 відмов. Побудувати функцію надійності системи. Знайти ймовірність того, що система працюватиме безвідмовно упродовж 50 годин, а також час безвідмовної роботи, гарантований з імовірністю 0,99.

Розв'язання. За формулою (5) знаходимо інтенсивність потоку подій

$$\lambda \approx \frac{n}{\tau} = \frac{5}{100} = 0,05.$$

Функція надійності

$$R(t) = e^{-0,05t}.$$

Ймовірність безвідмовної роботи впродовж 50 годин знаходимо за формулою (1) (або за формулою (7)):

$$P(T \geq 50) = e^{-\lambda t} = e^{-0,05 \cdot 50} \approx 0,61.$$

Нарешті, за формулою (8) отримуємо:

$$t = -\frac{\ln 0,99}{0,05} \approx 0,2 \text{ (год)}. \blacktriangleright$$

6.9. Поняття про закон великих чисел.

Центральна гранична теорема

Статистичні закономірності і можливість їх теоретико-ймовірнісного моделювання ґрунтуються на підтверженому досвідом факті існування цих закономірностей. Дуже важливими є теореми, які встановлюють умови цих закономірностей. Оскільки статистичні закономірності проявляються лише за великої кількості дослідів або, в математичній формі, як граничні властивості за необмеженого збільшення кількості випробувань, то ці твердження називають *теоремами закону великих чисел*.

Одна група теорем закону великих чисел стосується стійкості відносних частот і стійкості середніх. Друга група теорем присвячена закономірностям, яким підлягає сума великої кількості випадкових доданків: виявля-

ється, що за деяких припущень і необмеженого збільшення кількості доданків їхня сума підлягає в границі нормальному законові розподілу. Цю форму закону великих чисел називають *центральною граничною теоремою*.

Історично першою формою закону великих чисел була теорема Бернуллі про стійкість відносних частот у моделі повторних незалежних випробувань.

Теорема 1 (теорема Бернуллі). Нехай m – кількість появ події A в n випробуваннях Бернуллі з імовірністю p появи в одному випробуванні. Тоді для будь-якого числа $\varepsilon > 0$ справджується співвідношення

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

Іноді теорему Бернуллі інтерпретують неправильно, стверджучи, ніби з неї випливає, що зі збільшенням кількості випробувань частота $w(A) = m/n$ події A прямує до її ймовірності: $\lim_{n \rightarrow \infty} w(A) = p$. За

означенням границі, ця рівність означає, що нерівність $|w(A) - p| < \varepsilon$ справджується для всіх $n > N$, де N – деяке натуральне число, а $\varepsilon > 0$ – як завгодно мале число. Теорема ж Бернуллі стверджує інше, а саме: за $n \rightarrow \infty$ близькою до 1 є ймовірність $P(|w(A) - p| < \varepsilon)$, але сама нерівність $|w(A) - p| < \varepsilon$ для деяких навіть великих n може не справджуватися.

Є різні версії центральної граничної теореми. Одна з найпростіших – така.

Теорема 2 (центральна гранична теорема). Якщо X_1, X_2, \dots, X_n – послідовність взаємно незалежних однаково розподілених випадкових величин із математичними сподіваннями $a = M(X_i)$ і дисперсіями σ^2 , то за $n \rightarrow \infty$ закон розподілу суми цих величин

$$Y_n = X_1 + X_2 + \dots + X_n \quad (1)$$

прямує до нормального закону розподілу. \blacktriangle

Загальніше формулювання центральної граничної теореми дав, зокрема, О. М. Ляпунов (теорема Ляпунова).

Приклад. Кожна зі 100 незалежних випадкових величин X_i ($i = \overline{1, 100}$) рівномірно розподілена на відріжку $[0; 0,1]$. Побудуємо диференціальну й інтегральну функції розподілу суми Y цих величин.

Розв'язання. Щільність ймовірності окремої випадкової величини

$$f(x) = \begin{cases} \frac{1}{0,1}, & \text{якщо } 0 \leq x \leq 0,1, \\ 0, & \text{якщо } x < 0, x > 0,1. \end{cases}$$

Відшукаймо її математичне сподівання й дисперсію.

$$m_X = \int_0^{0,1} x \cdot \frac{1}{0,1} dx = \frac{1}{0,1} \cdot \frac{x^2}{2} \Big|_0^{0,1} = 0,05,$$
$$\sigma_X^2 = m_{X^2} - (m_X)^2 = \int_0^{0,1} x^2 \cdot \frac{1}{0,1} dx - 0,05^2 =$$
$$= \frac{1}{0,1} \cdot \frac{x^3}{3} \Big|_0^{0,1} - 0,05^2 \approx 0,00333 - 0,0025 = 0,00083.$$

Оскільки випадкові величини X_i незалежні, то

$$m_Y = 100 m_X = 5, \quad \sigma_Y^2 = 100 \sigma_X^2 \approx 0,083, \quad \sigma_Y \approx 0,2881.$$

Згідно з центральною граничною теоремою випадкова величина Y , як сума достатньо великої кількості незалежних випадкових величин, має розподіл, близький до нормального розподілу з параметрами $m_Y = 5$, $\sigma_Y \approx 0,2881$. Тому

$$f(y) \approx \frac{1}{\sqrt{0,166\pi}} e^{-\frac{(y-5)^2}{2 \cdot 0,083}},$$
$$F(y) \approx \frac{1}{\sqrt{0,166\pi}} \int_{-\infty}^y e^{-\frac{(y-5)^2}{2 \cdot 0,083}} dy. \blacktriangleright$$

Задачі до розділу 6

1. Рух автомобілів на перехресті регулюється світлофором, який дозволяє проїзд перехрестя впродовж 1 хв і забороняє впродовж 0,5 хв. Автомобіль підїжджає до перехрестя у випадковий момент часу. Яка ймовірність того, що він проїде перехрестя, не зупиняючись? Чому дорівнює середній час зупинки автомобіля перед світлофором?
2. Інтервал між відправленнями автобусів із зупинки дорівнює 8 хв. Яка ймовірність того, що пасажир, який прийшов на зупинку у випадковий момент часу чекатиме на автобус не більше 5 хв?
3. Записати біномний закон розподілу випадкової величини X для $n = 4$, $p = \frac{1}{4}$ і знайти її математичне сподівання та дисперсію.
4. Випадкова величина X розподілена за нормальним законом.
 - а) $a = 30$, $\sigma = 10$. Знайти $P(10 < X < 30)$.
 - б) $a = 15$, $\varepsilon = 3$, $P(|X - a| < \varepsilon) = 0,8664$. Знайти σ .
 - в) $a = 10$, $\sigma = 5$, $P(|X - a| < \varepsilon) = 0,9544$. Знайти ε .
 - г) $a = 8$, $\sigma = 4$, $\alpha = 8$, $P(\alpha < X < \beta) = 0,4772$. Знайти β .
 - д) $a = 6$, $\sigma = 2$, $\beta = 5$, $P(\alpha < X < \beta) = 0,2880$. Знайти α .
5. Випадкова величина X розподілена нормально з центром розподілу $a = 0,5$ і дисперсією $\sigma^2 = \frac{1}{8}$. Яка ймовірність того, що випадкова величина потрапляє в інтервал $(0,4; 0,6)$?
6. Верстат-автомат виготовляє деталі, довжину яких можна вважати випадковою величиною з нормальним законом розподілу. Параметри розподілу $m_X = 10$, $\sigma^2 = \frac{1}{200}$. Знайти ймовірність браку, якщо допустимі розміри деталей мають бути в межах $10 \pm 0,05$.
7. Знайти ймовірність влучення у смугу шириною 3,5 м, якщо помилки стрільби розподілені за нормальним законом з параметрами $a = 0$ і $\sigma = 1,9$.
8. Випадкова величина X характеризується нормальним законом розподілу. Математичне сподівання і середнє квадратичне відхилення дорівнюють відповідно 0 і 4. Відшукати ймовірність того, що відхилення випадкової величини від її середнього значення не перевищує 2.
9. Дохід на душу населення можна вважати випадковою величиною, що має нормальний розподіл із математичним сподіванням 12000 гр. од. середнім квадратичним відхиленням $\sigma = 300$ гр. од. В яких межах можна практично гарантувати дохід на душу населення?

10. Вважаючи, що результати вимірювань мають нормальний розподіл, знайти, який відсоток результатів відрізняється від середнього: а) більше ніж на чверть середнього квадратичного відхилення; б) менше ніж на половину середнього квадратичного відхилення.

11. У нормально розподіленій сукупності 15 % значень ознаки X менші від 12 і 40 % значень більші від 16,2. Знайти середнє значення ознаки і стандартне (середнє квадратичне) відхилення розподілу.

14. Щорічний дохід від продажу товару кожним продавцем має нормальний розподіл, середнє значення якого 720 тис. гр. од., а середнє квадратичне відхилення 80 тис. грн. Визначити ймовірність того, що річний дохід навання вибраного продавця: а) перевищує 900 тис. гр. од.; б) знаходиться в межах від 600 до 800 тис. гр. од.; в) менший ніж 500 тис. гр. од.; г) знаходиться в межах від 540 до 660 тис. гр. од.

15. Банк провів дослідження про наявність річних заощаджень в осіб, вік яких є не більший, ніж 21 рік. Дослідження засвідчили, що річні заощадження на одну особу нормально розподіляються з середнім значенням 1850 гр. од. і середнім квадратичним відхиленням 350 гр. од. Визначити ймовірність того, що навання вибрана особа має такі заощадження: а) більше ніж 2200 гр. од.; б) менше ніж 1500 гр. од.; в) у межах від 1080 гр. од. до 2375 гр. од.; г) менше ніж 800 гр. од.

16. Статистичні дослідження, які впродовж 10 років проводилися в регіоні, засвідчили, що врожайність пшениці з одного гектара є нормально розподіленою випадковою величиною з середньою врожайністю 45 ц і середнім квадратичним відхиленням 10 ц. Визначити ймовірність того, що в наступному після досліджень році врожайність пшениці з гектара: а) буде меншою, ніж 50 ц; б) в межах від 45 ц до 50 ц.

17. Верстат-автомат виготовляє вироби, які вважаються придатними, якщо відхилення розміру виробу (випадкова величина X) від проєктного розміру за абсолютним значенням не перевищує 0,8 мм. Яка ймовірність такого відхилення? Яка найімовірніша кількість придатних виробів із 200, якщо X має розподіл з параметром $\sigma = 0,4$ мм?

20. Для випадкової величини X , що розподілена за показниковим законом, знайти ймовірність того, що $X < M(X)$.

21. Практика свідчить, що 95 % приладів виходить з ладу після 1000 годин експлуатації. Знайти ймовірність того, що прилад вийде з ладу в проміжку від 800 до 900 годин експлуатації, вважаючи, що час виходу з ладу підлягає експоненційному законові розподілу.

22. Час безвідмовної роботи пристрою розподілений за законом $f(x) = 0,005e^{-0,005x}$. Знайти ймовірність того, що а) пристрій працюватиме безвідмовно 50 год; б) пристрій вийде з ладу упродовж 50 год.

РОЗДІЛ 8

БАГАТОВИМІРНІ ВИПАДКОВІ ВЕЛИЧИНИ

8.1. Закон розподілу дискретної двовимірної випадкової величини

Нехай (X, Y) – дискретна двовимірна випадкова величина, а $(x_i; y_j)$, $i = \overline{1, n}$, $j = \overline{1, m}$ – її можливі значення, ймовірності яких

$$p_{ij} = P(X = x_i, Y = y_j),$$

де подію $(X = x_i) \cap (Y = y_j)$ позначено $(X = x_i, Y = y_j)$.

Законом розподілу дискретної двовимірної випадкової величини (X, Y) називають перелік усіх можливих її значень та відповідних їм імовірностей. Зазвичай закон розподілу дискретної двовимірної випадкової величини задають у вигляді таблиці.

Табл. 8.1

$X \backslash Y$	x_1	x_2	\dots	x_i	\dots	x_n
y_1	p_{11}	p_{21}	\dots	p_{i1}	\dots	p_{n1}
y_2	p_{12}	p_{22}	\dots	p_{i2}	\dots	p_{n2}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
y_j	p_{1j}	p_{2j}	\dots	p_{ij}	\dots	p_{nj}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
y_m	p_{1m}	p_{2m}	\dots	p_{im}	\dots	p_{nm}

Події $(X = x_i, Y = y_j)$ утворюють повну групу, набуття випадковою величиною (X, Y) одного зі своїх значень є подія вірогідна, тому

$$\sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1.$$

У лівій частині цієї рівності стоїть знак подвійної суми. Розкривши цю суму за першим індексом підсумовування, матимемо:

$$\sum_{i=1}^n \sum_{j=1}^m p_{ij} = \sum_{j=1}^m p_{1j} + \sum_{j=1}^m p_{2j} + \dots + \sum_{j=1}^m p_{nj}.$$

Кожна із сум у правій частині дорівнює сумі елементів відповідного стовпця таблиці.

Табл. 8.1 дає повну характеристику двовимірної величини. Зокрема, на підставі таблиці можна побудувати закони розподілу складових X та Y , кожна з яких є одновимірною випадковою величиною. Справді, якщо, наприклад, відомо, що складова X набула значення x_1 , то це означає, що відбулася одна з таких несумісних подій:

$$(X = x_1, Y = y_1), (X = x_1, Y = y_2), \dots, (X = x_1, Y = y_m).$$

Оскільки події несумісні, то

$$p(x_1) = P(X = x_1) = p_{11} + p_{12} + \dots + p_{1m} = \sum_{j=1}^m p_{1j}.$$

Отже, ймовірність того, що $X = x_i$, дорівнює сумі ймовірностей стовпця x_i табл. 8.1:

$$p(x_i) = P(X = x_i) = \sum_{j=1}^m p_{ij}, \quad i = \overline{1, n}.$$

Аналогічно, ймовірність того, що складова Y набула значення y_j , дорівнює сумі ймовірностей рядка y_j :

$$p(y_j) = P(Y = y_j) = \sum_{i=1}^n p_{ij}, \quad j = \overline{1, m}.$$

Отже, в разі дискретної випадкової величини ймовірності значень її складових X та Y визначаються підсумковим рядком та підсумковим стовпцем таблиці розподілу.

Приклад. Закон розподілу випадкової величини (X, Y) має вигляд

$X \backslash Y$	-1	2	3
1	0,1	0,2	0,12
3	0,25	0,13	0,2

Побудуємо ряди розподілу випадкових величин X та Y .

Розв'язання. Випадкова величина X може набувати значень $-1, 2, 3$ з такими ймовірностями:

$$P(X = -1) = 0,1 + 0,25 = 0,35 ,$$

$$P(X = 2) = 0,2 + 0,13 = 0,33 ,$$

$$P(X = 3) = 0,12 + 0,2 = 0,32 .$$

Отже, ряд розподілу складової X має вигляд:

X	-1	2	3
P	$0,35$	$0,33$	$0,32$

Ймовірності можливих значень складової Y дорівнюють сумі ймовірностей відповідних рядків таблиці розподілу. Ряд розподілу складової Y має вигляд:

Y	1	3
P	$0,42$	$0,58$

Суми ймовірностей дорівнюють одиниці. ►

8.2. Функція розподілу двовимірної випадкової величини

Функцією розподілу двовимірної випадкової величини (X, Y) називають функцію двох змінних $F(x, y)$, яка для кожної пари значень x та y визначає ймовірність одночасного справдження нерівностей $X < x, Y < y$:

$$F(x, y) = P(X < x, Y < y) . \quad (1)$$

Функцію $F(x, y)$ називають ще *двовимірною функцією розподілу*, або *спільним розподілом величин X та Y* , або *розподілом вектора (X, Y)* .

Геометричне тлумачення двовимірної функції розподілу зрозуміле з рис. 8.2: $F(x, y)$ – це ймовірність того, що випадкова точка (X, Y) потрапить у нескінченну область, яка лежить лівіше і нижче від точки $A(x; y)$.

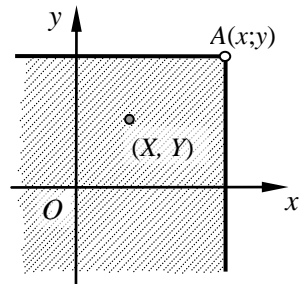
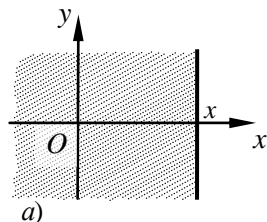
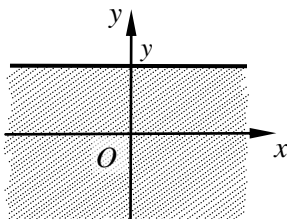


Рис. 8.2

Нескладно дати також геометричну інтерпретацію окремих складових двовимірної випадкової величини.



а)



б)

Рис. 8.3

Функція розподілу $F_1(x)$ випадкової складової X дорівнює ймовірності того, що випадкова точка (X, Y) потрапить у півплощину $X < x$ (рис. 8.3, а)); функція розподілу $F_2(y)$ випадкової складової Y дорівнює ймовірності того, що випадкова точка (X, Y) потрапить у півплощину $Y < y$ (рис. 8.3, б)).

Властивості функції $F(x, y)$ подібні до властивостей одновимірної функції розподілу.

Властивість 1.

$$0 \leq F(x, y) \leq 1. \quad (2)$$

Властивість очевидна й випливає з того, що функція $F(x, y)$ має зміст імовірності. ▲

Властивість 2. Функція $F(x, y)$ є неспадна функція за кожним з аргументів:

$$F(x_2, y) \geq F(x_1, y), \text{ якщо } x_2 > x_1, \quad (3)$$

$$F(x, y_2) \geq F(x, y_1), \text{ якщо } y_2 > y_1. \quad \blacktriangle (4)$$

Властивість 3. Справджуються співвідношення

$$F(+\infty, y) = F_2(y), \quad (6)$$

$$F(x, +\infty) = F_1(x). \quad (7)$$

Ці рівності мають зрозумілий геометричний зміст. Справді, нехай, наприклад $x \rightarrow +\infty$. Тоді, як видно з рис. 8.2, права межа безмежного квадранта зсувається в $+\infty$, і квадрант обертається на півплощину, показану на рис. 8.3, б). Ймовірність потрапляння випадкової точки (X, Y) у цю півплощину і є функція розподілу $F_2(y)$.

Аналогічно можна проінтерпретувати рівність (7). ▲

Властивість 4.

$$F(+\infty, +\infty) = 1. \quad (8)$$

Властивість 5.

$$F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0. \quad (9)$$

Геометричне тлумачення властивостей 4 і 5 очевидне. ▲

8.3. Ймовірність потрапляння випадкової точки в нескінченну напівсмугу і в прямокутник

Відшукаймо ймовірність того, що випадкова точка потрапить у нескінченну напівсмугу, яка визначена системою нерівностей $x_1 \leq X < x_2$, $Y < y$ і зображена на рис. 8.4, а).

Формально потрібно знайти ймовірність $P(x_1 \leq X < x_2, Y < y)$. Ця ймовірність визначена рівністю (5) із попереднього параграфу і дорівнює частинному приростові функції $F(x, y)$ за аргументом x :

$$\begin{aligned} P(x_1 \leq X < x_2, Y < y) &= \\ &= F(x_2, y) - F(x_1, y). \end{aligned} \quad (1)$$

Аналогічно доходимо висновку, що ймовірність потрапляння випадкової точки в нескінченну напівсмугу $X < x$, $y_1 \leq Y < y_2$ (рис. 8.4, б)) дорівнює відповідному частинному приростові функції $F(x, y)$ за аргументом y :

$$\begin{aligned} P(X < x; y_1 \leq Y < y_2) &= \\ &= F(x, y_2) - F(x, y_1). \end{aligned} \quad (2)$$

Відшукаймо ймовірність того, що випадкова величина потрапить у прямокутник, сторони якого паралельні осям координат (рис. 8.5).

З геометричних міркувань зрозуміло, що шукана ймовірність дорівнює різниці ймовірностей потрапляння випадкової точки у нескінченні вертикальні напівсмуги, перша з яких обмежена згори відрізком із кінцями в точках $A(x_1; y_2)$ та $B(x_2; y_2)$, а друга – відрізком із кінцями в точках $C(x_1; y_1)$ та $D(x_2; y_1)$. Ймовірності потрапляння

випадкової точки в ті напівсмуги дорівнюють відповідно:

$$\begin{aligned} P(x_1 \leq X < x_2, Y < y_2) &= F(x_2, y_2) - F(x_1, y_2), \\ P(x_1 \leq X < x_2, Y < y_1) &= F(x_2, y_1) - F(x_1, y_1). \end{aligned}$$

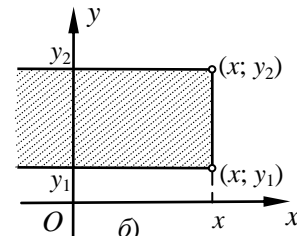
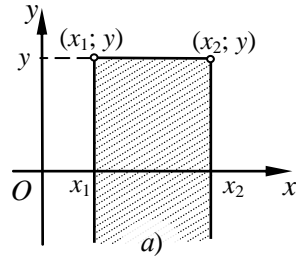


Рис. 8.4

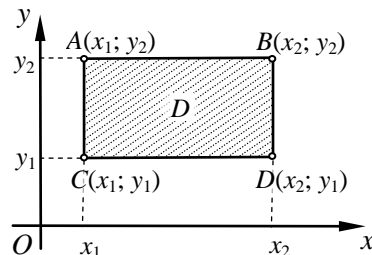


Рис. 8.5

Отже, ймовірність потрапляння випадкової точки у прямокутну область, зображену на рис. 8.5,

$$P(x_1 \leq X < x_2, y_1 \leq Y < y_2) = \\ = (F(x_2, y_2) - F(x_1, y_2)) - (F(x_2, y_1) - F(x_1, y_1)). \quad (3)$$

Приклад. Випадкова величина описується функцією розподілу

$$F(x, y) = \begin{cases} (1 - e^{-x})(1 - e^{-2y}), & \text{якщо } x > 0, y > 0, \\ 0, & \text{якщо } x \leq 0 \text{ або } y \leq 0. \end{cases}$$

Відшукаймо ймовірність того, що випадкова величина потрапить у прямокутник із вершинами в точках $O(0;0)$, $A(0;1)$, $B(2;1)$, $C(2;0)$.

Розв'язання. За формулою (3)

$$P(0 \leq X < 2, 0 \leq Y < 1) = (F(2;1) - F(0;1)) - (F(2;0) - F(0;0)) = \\ = (1 - e^{-2})(1 - e^{-2}) \approx (1 - 0,1353)^2 \approx 0,78. \blacktriangleright$$

8.4. Щільність імовірності двовимірної неперервної випадкової величини

Вважатимемо, що функція розподілу двовимірної випадкової величини $F(x, y)$ неперервна і має неперервну мішану похідну другого порядку $F''_{xy}(x, y)$. Тоді за аналогією з одновимірною випадковою величиною можна ввести двовимірну щільність імовірності:

$$f(x, y) = F''_{xy}(x, y). \quad (1)$$

Геометрично двовимірною щільністю імовірності $f(x, y)$ – це деяка поверхня, яку називають *поверхнею розподілу*.

З рівності (1) отримуємо:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv. \quad (2)$$

Приклад 1. Двовимірною функцією розподілу має вигляд:

$$F(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y}, & \text{якщо } x \geq 0, y \geq 0, \\ 0, & \text{якщо } x < 0 \text{ або } y < 0. \end{cases} \quad (a)$$

Відшукаймо двовимірну щільність імовірностей.

Розв'язання. Скористаймося формулою (1). Відшукаймо мішану похідну:

$$(1 - e^{-x} - e^{-y} + e^{-x-y})'_x = e^x - e^{-x-y}, \\ (1 - e^{-x} - e^{-y} + e^{-x-y})''_{xy} = (e^x - e^{-x-y})'_y = e^{-x-y}.$$

Отже, двовимірна щільність ймовірності

$$f(x, y) = \begin{cases} e^{-x-y}, & \text{якщо } x \geq 0, y \geq 0, \\ 0, & \text{якщо } x < 0, y < 0. \end{cases} \blacktriangleright \quad (b)$$

Вираз $f(x, y)dxdy$ називають *елементом ймовірності двовимірної випадкової величини*. Як видно з формули (2), за малих значень $\Delta x = dx$ та $\Delta y = dy$ елемент ймовірності наближено дорівнює ймовірності того, що випадкова точка потрапить у прямокутник зі сторонами dx та dy , який розташований поблизу точки $A(x; y)$ (рис. 8.6).

Маючи елемент ймовірності, можна знайти ймовірність потрапляння випадкової величини в довільну плоску область D . Справді, ймовірність потрапляння випадкової величини в прямокутний елемент області поблизу точки $A(x; y)$ (рис. 8.6) наближено дорівнює елементові ймовірності $f(x, y)dxdy$, а ймовірність того, що випадкова величина потрапить в область D , дорівнює, очевидно, відповідному подвійному інтегралові:

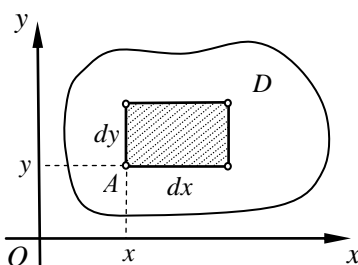


Рис. 8.6

$$P((X, Y) \in D) = \iint_D f(x, y)dxdy.$$

З огляду на геометричний зміст подвійного інтеграла зрозуміло, що ймовірність $P((X, Y) \in D)$ чисельно дорівнює об'єму вертикального циліндроїда, що обмежений зверху поверхнею $z = f(x, y)$, а його основою є область D .

Властивості двовимірної щільності розподілу ймовірностей $f(x, y)$.

Властивість 1. Щільність ймовірності є невід'ємна функція:

$$f(x, y) \geq 0. \quad (3)$$

Це випливає з означення (3) щільності ймовірності як границі відношення двох невід'ємних величин: ймовірності P та приросту площі $\Delta x \cdot \Delta y$. \blacktriangle

Властивість 2. Ймовірність потрапляння двовимірної випадкової величини у прямокутну область $D = \{a \leq X \leq b, c \leq Y \leq d\}$:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y)dxdy. \blacktriangle \quad (4)$$

Властивість 3. Зв'язок між $f(x, y)$ і $F(x, y)$:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy. \blacktriangle \quad (5)$$

Властивість 4. Інтеграл за нескінченними межами від щільності розподілу ймовірностей дорівнює одиниці:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1. \quad (6)$$

Це співвідношення випливає з рівності (2) і властивості 3 функції розподілу, згідно з якою $F(+\infty; -\infty) = 1$. \blacktriangle

Якщо всі можливі значення випадкової величини (X, Y) належать деякій скінченній області D , то остання рівність набирає вигляду:

$$\iint_D f(x, y) dx dy = 1. \quad (7)$$

Геометрично властивість 4 означає, що об'єм тіла (циліндроїда), обмеженого поверхнею розподілу $f(x, y)$, областю D площини Oxy та відповідними твірними, паралельними осі Oz , дорівнює одиниці. Якщо поверхня розподілу нескінченна, то й тіло, обмежене нею, буде нескінченним, але об'єм того тіла буде скінченним і дорівнюватиме 1.

Приклад 2. Надходження товару до споживача і поступлення вимоги на нього з однаковою ймовірністю можуть відбутися в будь-який із семи днів тижня. Відшукаймо ймовірність того, що: 1) товар надійде раніше, ніж поступить вимога на нього (подія A); 2) товар буде відправлено після отримання вимоги (подія B).

Розв'язання. Нехай X та Y – часи надходження товару та вимоги. За умовою $0 \leq X \leq 7$, $0 \leq Y \leq 7$, тому $f(x, y) = \frac{1}{7 \cdot 7} = \frac{1}{49}$. Отже,

$$\begin{aligned} P(A) &= P(0 \leq X \leq y, 0 \leq Y \leq 7) = \\ &= \frac{1}{49} \int_0^7 \int_0^y dx dy = \frac{1}{49} \int_0^7 y dy = \frac{1}{49} \cdot \frac{7^2}{2} = \frac{1}{2}; \\ P(B) &= P(y \leq X \leq 7, 0 \leq Y \leq 7) = \frac{1}{49} \int_0^7 \int_y^7 dx dy = \\ &= \frac{1}{49} \int_0^7 (7 - y) dy = -\frac{1}{49} \cdot \frac{(7 - y)^2}{2} \Big|_0^7 = \frac{1}{2}. \end{aligned}$$

Отже, ймовірності подій однакові і дорівнюють 0,5. \blacktriangleright

Як ми знаємо, для дискретних випадкових величин розподіли складових задаються підсумковими рядком і стовпцем таблиці розподілу (п. 8.3). Не вдаючись у деталі, зауважимо, що для двовимірної неперервної випадкової величини щільності ймовірностей складових мають вигляд:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \quad f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

8.5. Умовні закони розподілу складових двовимірної випадкової величини

Закони розподілу складових не дають повної характеристики двовимірної величини, зокрема, не відображають залежності складових. Цю властивість величини $Z = (X, Y)$ відображають умовні розподіли складових X і Y .

Як ми знаємо, якщо випадкові події A і B залежні, то умовна ймовірність визначається за формулою:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1)$$

Умовну ймовірність того, що складова X набуде значення x_i за умови, що складова $Y = y_j$, позначають $p(x_i|y_j)$.

За рівністю (1) умовні ймовірності визначають за формулою:

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}. \quad (2)$$

Умовним законом розподілу дискретної випадкової величини X за фіксованого значення $Y = y_j$ називають перелік можливих значень x_1, x_2, \dots, x_n величини X та відповідних їм умовних імовірностей $p(x_1|y_j), p(x_2|y_j), \dots, p(x_n|y_j)$.

Умовний розподіл складової X можна подати у вигляді таблиці.

$X = x_i$	x_1	x_2	...	x_i	...	x_n
$p(x_i y_j)$	$p(x_1 y_j)$	$p(x_2 y_j)$...	$p(x_i y_j)$...	$p(x_n y_j)$

Поняття умовного розподілу складової Y вводиться аналогічно.

Умовним законом розподілу дискретної випадкової величини Y за фіксованого значення $X = x_i$ ($i = \overline{1, n}$) називають перелік можливих

значень y_1, y_2, \dots, y_m величини Y та відповідних їм умовних імовірностей $p(y_1|x_i), p(y_2|x_i), \dots, p(y_m|x_i)$.

Таблиця умовного розподілу складової Y має вигляд:

$Y = y_j$	y_1	y_2	...	y_j	...	y_m
$p(y_j x_i)$	$p(y_1 x_i)$	$p(y_2 x_i)$...	$p(y_j x_i)$...	$p(y_m x_i)$

Відповідні умовні ймовірності визначають за формулою:

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}. \quad (3)$$

Зрозуміло, що

$$\sum_{i=1}^n p(x_i|y_j) = 1, \quad \sum_{j=1}^m p(y_j|x_i) = 1. \quad (4)$$

Як видно з формул (2) і (3), у разі дискретної випадкової величини умовні розподіли складових характеризуються відповідними *нормованими* стовпцем і рядком таблиці розподілу. Під нормуванням розуміється ділення елементів стовпця чи рядка на суму цих елементів.

Приклад. Закон розподілу випадкової величини (X, Y) має вигляд:

	Табл. 8.6.1			
$X \backslash Y$	-1	2	5	$p(y_j)$
3	0,2	0,1	a	0,4
4	0,1	0,3	0,2	0,6
$p(x_i)$	0,3	0,4	0,3	$\sum p = 1$

Побудувати умовні розподіли складових X та Y .

Розв'язання. Насамперед зауважмо, що розподіли (безумовні) складових X і Y задаються відповідно підсумковими рядком і стовпцем закону розподілу (у таблиці вони виділені). Зокрема,

$$p(x_1 = -1) = 0,3, \quad p(y_2 = 4) = 0,6.$$

За формулою (2) знаходимо умовні розподіли складової X .

Розподіл складової X за умови $Y = y_1$:

$$p(x_1|y_1) = \frac{p(x_1, y_1)}{p(y_1)} = \frac{0,2}{0,4} = \frac{1}{2}, \quad p(x_2|y_1) = \frac{0,1}{0,4} = \frac{1}{4}, \quad p(x_3|y_1) = \frac{1}{4}.$$

Аналогічно знаходимо розподіл складової X за умови $Y = y_2$:

$$p(x_1|y_2) = \frac{p(x_1, y_2)}{p(y_2)} = \frac{0,1}{0,6} = \frac{1}{6}, \quad p(x_2|y_2) = \frac{0,3}{0,6} = \frac{1}{2}, \quad p(x_3|y_2) = \frac{1}{3}.$$

За формулою (3) знаходимо умовні розподіли складової Y :

$$p(y_1|x_1) = \frac{p(x_1, y_1)}{p(x_1)} = \frac{0,2}{0,3} = \frac{2}{3}, \quad p(y_2|x_1) = \frac{p(x_1, y_2)}{p(x_1)} = \frac{1}{3};$$

$$p(y_1|x_2) = \frac{p(x_2, y_1)}{p(x_2)} = \frac{0,1}{0,4} = \frac{1}{4}, \quad p(y_2|x_2) = \frac{p(x_2, y_2)}{p(x_2)} = \frac{3}{4};$$

$$p(y_1|x_3) = \frac{p(x_3, y_1)}{p(x_3)} = \frac{0,1}{0,3} = \frac{1}{3}, \quad p(y_2|x_3) = \frac{p(x_3, y_2)}{p(x_3)} = \frac{2}{3}.$$

Умовні розподіли можна подати в табличному вигляді. Зокрема,

$X = x_i$	-1	2	5
$p(x_i y_2)$	1/6	1/2	1/3

$Y = y_j$	3	4
$p(y_j x_3)$	1/3	2/3

Аналогічно можна побудувати решту рядів умовних розподілів заданої випадкової величини (X, Y) . ►

Подібно вводять поняття умовних розподілів неперервної двовимірної випадкової величини.

Умовною щільністю $f(x|y)$ розподілу складової X за фіксованого значення складової $Y = y$ називають відношення двовимірної щільності ймовірності $f(x, y)$ випадкової величини (X, Y) до щільності ймовірності $f_2(y)$ складової Y :

$$f(x|y) = \frac{f(x, y)}{f_2(y)} = f(x, y) \Big/ \int_{-\infty}^{+\infty} f(x, y) dx. \quad (5)$$

Аналогічно визначають умовну щільність $f(y|x)$ розподілу складової Y за фіксованого значення складової $X = x$.

Властивості умовної щільності ймовірності аналогічні властивостям звичайної щільності ймовірності. Зокрема,

$$f(x|y) \geq 0, \quad \int_{-\infty}^{+\infty} f(x|y) dx = 1. \quad (6)$$

З означень (2) і (3) та (5) і (6) випливає така теорема.

Теорема. Щільність розподілу двовимірної випадкової величини дорівнює добуткові щільності ймовірності однієї зі складових на щільність умовного розподілу іншої:

$$f(x, y) = f_1(x) \cdot f(y | x) = f_2(y) \cdot f(x | y). \blacktriangle \quad (7)$$

Ця рівність аналогічна виразу для ймовірності добутку двох подій.

Залежність випадкових величин означає аналітичну залежність умовного розподілу однієї з них від значень, яких набуває інша.

8.6. Умовні числові характеристики двовимірної випадкової величини

Одним з аспектів дослідження багатовимірного розподілу є опис його властивостей за допомогою числових характеристик, які вводяться аналогічно до характеристик одновимірної випадкової величини. Такі характеристики можна ввести і для умовних законів розподілу багатовимірної випадкової величини.

Числові характеристики умовного закону розподілу випадкової величини називають *умовними*. Основними характеристиками умовних розподілів є умовне математичне сподівання та умовна дисперсія.

Формули для відшукування умовних числових характеристик аналогічні відповідним формулам для безумовних характеристик. Зокрема, умовні математичні сподівання дискретної двовимірної випадкової величини (X, Y) визначають за формулами:

$$M(X|Y = y_j) = \sum_{i=1}^n x_i p(x_i | y_j) = \frac{1}{p(y_j)} \sum_{i=1}^n x_i p(x_i, y_j), \quad (1)$$

$$M(Y|X = x_i) = \sum_{j=1}^m y_j p(y_j | x_i) = \frac{1}{p(x_i)} \sum_{j=1}^m y_j p(x_i, y_j). \quad (2)$$

Умовну дисперсію величини X визначають за формулою:

$$D(X|Y = y) = M\left(\left((X|Y = y) - M(X|Y = y)\right)^2\right)$$

або

$$D(X|Y = y) = M(X^2|Y = y) - M^2(X|Y = y).$$

Аналогічно визначають умовну дисперсію складової Y :

$$D(Y|X = x) = M\left(\left((Y|X = x) - M(Y|X = x)\right)^2\right)$$

або

$$D(Y|X = x) = M(Y^2|X = x) - M^2(Y|X = x).$$

Зокрема, для умовних дисперсій дискретної випадкової величини:

$$D(X|Y = y_j) = \sum_{i=1}^n x_i^2 p(x_i|y_j) - M^2(X|Y = y_j), \quad (3)$$

$$D(Y|X = x_i) = \sum_{j=1}^m y_j^2 p(y_j|x_i) - M^2(Y|X = x_i). \quad (4)$$

Умовним математичним сподіванням $M(X|y)$ неперервної випадкової величини називають математичне сподівання відповідної умовної щільності розподілу ймовірностей $f(x|y)$:

$$M(X|y) \equiv M(X|Y = y) = \int_{-\infty}^{+\infty} xf(x|y)dx. \quad (5)$$

Аналогічно,

$$M(Y|x) \equiv M(Y|X = x) = \int_{-\infty}^{+\infty} yf(y|x)dy. \quad (6)$$

Умовні дисперсії неперервної випадкової величини визначають за формулами:

$$D(X|y) \equiv D(X|Y = y) = \int_{-\infty}^{+\infty} x^2 f(x|y)dx - M^2(X|y),$$

$$D(Y|x) \equiv D(Y|X = x) = \int_{-\infty}^{+\infty} y^2 f(y|x)dy - M^2(Y|x).$$

Умовні щільності визначають за формулами п. 8.6.

Як бачимо, умовна дисперсія визначається так само як безумовна, але замість безумовного розподілу слід узяти умовний розподіл.

З формули (1) видно, що умовне математичне сподівання $M(X|y)$ кожному можливому значенню y ставить у відповідність єдине значення математичного сподівання випадкової величини X . Отже, умовне математичне сподівання $M(X|y)$ є функцією від y :

$$M(X|y) = g(y). \quad (7)$$

Цю функцію називають *регресією випадкової величини X на Y* (або *регресією X щодо Y*). Функція $M(X|Y = y)$ показує, яким буде в середньому значення випадкової змінної X , якщо змінна Y набуває значення y .

Аналогічно, умовне математичне сподівання $M(Y|x)$ випадкової змінної Y , яке розглядається як функція x , називають *функцією регресії випадкової змінної Y на X* (або *регресією Y щодо X*):

$$M(Y|x) = \varphi(x).$$

Ця функція регресії виражає математичне сподівання змінної Y за умови, що друга змінна набуває певного числового значення. Функція $M(Y|x)$ показує, яким буде в середньому значення випадкової змінної Y , якщо змінна X набуває значення x .

Зі сказаного зрозуміло, що функція регресії має важливе значення для статистичного аналізу залежностей між змінними і може використовуватися для прогнозування значень однієї з випадкових змінних, якщо відоме значення другої. Точність такого прогнозування визначається дисперсією умовного розподілу.

8.7. Залежні та незалежні випадкові величини.

Стохастична залежність

Дві випадкові величини X і Y називають *незалежними*, якщо закон розподілу однієї з них не залежить від того, яких значень набуває інша. Поняття незалежності випадкових величин є взаємним: якщо X не залежить від Y , то й Y не залежить від X .

Поняття залежності системи випадкових величин (X, Y) є узагальненням поняття залежності подій. Під подіями A і B розумітимемо справдження нерівностей $X < x$ і $Y < y$. За такої інтерпретації випадкові величини X і Y є незалежними, якщо ймовірність спільного настання двох подій $(X < x)$ і $(Y < y)$ для будь-яких x та y дорівнює добуткові ймовірності цих подій:

$$P((X < x)(Y < y)) = P(X < x)P(Y < y). \quad (1)$$

З огляду на означення функції розподілу, останнє співвідношення можна записати у вигляді:

$$F(x, y) = F_1(x) \cdot F_2(y). \quad (2)$$

Рівність (2) є означенням незалежності випадкових величин X і Y . Якщо функцію розподілу $F(x, y)$ не можна подати як добуток функцій $F_1(x)$ і $F_2(y)$, то випадкові величини X і Y називають *залежними*.

Диференціюючи обидві частини останньої рівності за змінними x та y , дістанемо умову незалежності випадкових величин X і Y у термінах диференціальних функцій розподілу:

$$f(x, y) = f_1(x) \cdot f_2(y). \quad (3)$$

Якщо $f(x, y) \neq f_1(x)f_2(y)$, то величини X і Y залежні.

Приклад. Випадкові величини X та Y , спільна щільність розподілу ймовірностей яких має вигляд

$$f(x, y) = \begin{cases} abe^{-ax-by}, & \text{якщо } x \geq 0, y \geq 0, \\ 0, & \text{якщо } x < 0 \text{ або } y < 0, \end{cases} \quad (a)$$

є незалежними, оскільки $f(x, y) = f_1(x) \cdot f_2(y)$, де

$$f_1(x) = \begin{cases} ae^{-ax}, & \text{якщо } x \geq 0, \\ 0, & \text{якщо } x < 0, \end{cases} \quad f_2(y) = \begin{cases} be^{-by}, & \text{якщо } y \geq 0, \\ 0, & \text{якщо } y < 0. \end{cases} \blacktriangleright$$

Умову незалежності можна записати також в іншій формі. А саме: за теоремою множення законів розподілу (п. 8.6),

$$f(x, y) = f_1(x)f(y|x) = f_2(y)f(x|y). \quad (4)$$

Якщо величина X та Y незалежні, то праву частину формули (4) за означенням (3) можна подати у вигляді $f_1(x)f_2(y)$. Після скорочення на $f_1(x)$, або на $f_2(y)$ дістанемо дві рівносильні умови незалежності:

$$f_2(y) = f(y|x) \quad \text{або} \quad f_1(x) = f(x|y). \quad (5)$$

Отже, незалежність двох випадкових величин означає, що щільність умовного розподілу кожної з них збігається з відповідною щільністю безумовного розподілу.

Залежність випадкових величин є узагальненням поняття функціональної залежності $y = f(x)$, коли кожному значенню x відповідає єдине значення y . У разі залежності між випадковими величинами окремому $X = x$ відповідає деяка сукупність значень (умовний розподіл) змінної Y . Залежність випадкових величин означає аналітичну залежність умовного розподілу однієї з них від значень, яких набуває інша. Таку залежність називають *ймовірнісною* або *стохастичною*.

Залежність умовного математичного сподівання однієї випадкової величини від значень іншої називають *кореляційною залежністю*.

8.8. Числові характеристики двовимірної випадкової величини. Коефіцієнт кореляції

Як ми знаємо, найважливішою характеристикою випадкової величини є її закон розподілу. Однак на практиці часто використовують хоч і не такі повні, зате зручніші числові характеристики. Для двовимірної випадкової величини (X, Y) такими характеристиками є, зокрема, здобуті раніше математичні сподівання та дисперсії її складових.

Математичним сподіванням двовимірної випадкової величини (X, Y) називають сукупність математичних сподівань складових X та Y , тобто не випадковий вектор із координатами m_X і m_Y .

Аналогічно, під *дисперсією двовимірної випадкової величини (X, Y)* розуміють сукупність дисперсій складових X і Y , тобто не випадковий вектор із компонентами σ_X^2 і σ_Y^2 . Дисперсії σ_X^2 і σ_Y^2 описують розсіяння випадкової точки навколо математичних сподівань m_X і m_Y у напрямках осей Ox і Oy відповідно.

Але, як уже згадувалося, для характеристики системи двох випадкових величин не досить вивчити кожен складову окремо. Необхідно вивчити також взаємний зв'язок (залежність) між складовими X та Y , який проявляється в тому, що умовний розподіл однієї зі складових змінюється залежно від того, яких значень набуває інша.

Далі розглядаємо дві числові характеристики ймовірнісного зв'язку: кореляційний момент (коваріація) випадкових величин і коефіцієнт кореляції.

Кореляційним моментом (коваріацією) випадкової величини (X, Y) називають математичне сподівання добутку відхилень її складових X і Y від своїх математичних сподівань:

$$\mu_{XY} \equiv \text{cov}(X, Y) = M((X - m_X)(Y - m_Y)), \quad (1)$$

або $\overset{\circ}{X} = X - m_X$ $\overset{\circ}{Y} = Y - m_Y$ – центровані випадкові величини),

$$\mu_{XY} = M \begin{pmatrix} \overset{\circ}{X} & \overset{\circ}{Y} \end{pmatrix}. \quad (2)$$

Коваріація – це другий мішаний центральний момент випадкових величин X і Y .

Використовуючи властивості математичного сподівання, формулу (1) можна записати у зручнішому вигляді. А саме:

$$\begin{aligned}\mu_{XY} &= M((X - m_X)(Y - m_Y)) = \\ &= M(XY - X m_Y - m_X Y + m_X m_Y) = \\ &= M(XY) - m_Y M(X) - m_X M(Y) + m_X m_Y,\end{aligned}$$

тобто, кореляційний момент випадкових величин X і Y

$$\mu_{XY} = M(XY) - m_X m_Y. \quad (3)$$

Як видно з формул (1)–(3), якщо $X = Y$, то кореляційний момент дорівнює дисперсії випадкової величини X :

$$\mu_{XX} = M(X^2) - m_X^2 = D(X) \equiv \sigma_X^2. \quad (4)$$

З огляду на це кореляційний момент μ_{XY} іноді позначають σ_{XY} ($\sigma_{XX} = \sigma_X^2$) і називають *мішаною дисперсією*.

Якщо випадкові величини X та Y дискретні, то кореляційний момент розраховують за формулою

$$\mu_{XY} = \sum_{j=1}^m \sum_{i=1}^n (x_i - m_X)(y_j - m_Y) p_{ij}, \quad (5)$$

або за формулою

$$\mu_{XY} = \sum_{j=1}^m \sum_{i=1}^n x_i y_j p_{ij} - m_X m_Y, \quad (6)$$

де p_{ij} – це ймовірність події, яка полягає в тому, що випадкова величина X набула значення x_i , а випадкова величина Y – значення y_j :

$$p_{ij} = P((X = x_i) \cap (Y = y_j)).$$

Якщо випадкові величини X і Y неперервні, то кореляційний момент вираховують за формулою

$$\mu_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y) f(x, y) dx dy,$$

або за формулою

$$\mu_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - m_X m_Y, \quad (7)$$

де $f(x, y)$ – густина спільного розподілу величин X і Y .

Теорема 1. Кореляційний момент двох незалежних випадкових величин дорівнює нулеві. ▲

Доведення. Ця властивість випливає з означення (3), бо для незалежних величин математичне сподівання добутку дорівнює добуткові математичних сподівань: $M(XY) = M(X)M(Y)$. ▲

Оскільки для незалежних величин $\text{cov}(X, Y) = 0$, то нерівність $\text{cov}(X, Y) \neq 0$ є умовою залежності випадкових величин X і Y . Але слід розуміти, що з рівності $\text{cov}(X, Y) = 0$ не випливає незалежність випадкових величин X і Y .

Використовуючи властивості математичного сподівання, нескладно довести такі властивості коваріації (C_1 і C_2 – стали):

$$\begin{aligned} \text{cov}(X, Y) &= \text{cov}(Y, X), \\ \text{cov}(X + C_1, Y + C_2) &= \text{cov}(X, Y), \\ \text{cov}(C_1 X_1 + C_2 X_2, Y) &= C_1 \text{cov}(X_1, Y) + C_2 \text{cov}(X_2, Y). \end{aligned} \quad (8)$$

З означення випливає, що коваріація μ_{XY} має розмірність, яка дорівнює добуткові розмірностей випадкових величин X і Y . Це означає, що значення коваріації залежать від одиниць вимірювання X і Y . Природно вимагати, щоби величина, яка є мірою стохастичного зв'язку між величинами, не залежала від вибраного масштабу вимірювання вихідних параметрів. Для цього досить від вихідних випадкових величин X і Y перейти до нормованих випадкових величин X_σ і Y_σ за формулами

$$X_\sigma = \frac{X - m_X}{\sigma_X}, \quad Y_\sigma = \frac{Y - m_Y}{\sigma_Y}, \quad (10)$$

які, очевидно, є безрозмірними й, окрім того, $m_{X_\sigma} = m_{Y_\sigma} = 0$, $\sigma_{X_\sigma}^2 = \sigma_{Y_\sigma}^2 = 1$. З огляду на властивості коваріації дістанемо:

$$\begin{aligned} \text{cov}(X_\sigma, Y_\sigma) &= \text{cov}\left(\frac{X - m_X}{\sigma_X}, \frac{Y - m_Y}{\sigma_Y}\right) = \\ &= \frac{\text{cov}(X - m_X, Y - m_Y)}{\sigma_X \sigma_Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}. \end{aligned} \quad (11)$$

Ця величина, на відміну од коваріації випадкових величин X та Y , є безрозмірною, і, очевидно, також може слугувати мірою їх залежності.

Коефіцієнтом кореляції ρ_{XY} випадкових величин X і Y називають відношення коваріації цих величин до добутку їхніх середніх квадратичних відхилень:

$$\rho_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}. \quad (12)$$

Приклад. Знайти коваріацію і коефіцієнт кореляції випадкової величини (X, Y) , закон розподілу якої задано табл. 8.6.1.

Розв'язання. Скористаємося формулою (6). Знаючи розподіли складових (це підсумкові рядок і стовпець таблиці розподілу), знаходимо їхні математичні сподівання, дисперсії і середні квадратичні відхилення:

$$m_X = \sum_{i=1}^3 x_i \cdot p(x_i) = (-1) \cdot 0,3 + 2 \cdot 0,4 + 5 \cdot 0,3 = 2,$$

$$m_Y = \sum_{j=1}^2 y_j \cdot p(y_j) = 3 \cdot 0,4 + 4 \cdot 0,6 = 3,6;$$

$$\sigma_X^2 = m_{X^2} - m_X^2 = \left((-1)^2 \cdot 0,3 + 2^2 \cdot 0,4 + 5^2 \cdot 0,3 \right) - 2^2 = 4,4.$$

$$\sigma_Y^2 = m_{Y^2} - m_Y^2 = \left(3^2 \cdot 0,4 + 4^2 \cdot 0,6 \right) - 3,6^2 = 0,24;$$

$$\sigma_X = \sqrt{4,4} \approx 2,10, \quad \sigma_Y = \sqrt{0,24} \approx 0,49.$$

Математичне сподівання добутку випадкових величин X та Y :

$$m_{XY} = \sum_{j=1}^2 \sum_{i=1}^3 x_i y_j p_{ij} = (-1)(3 \cdot 0,2 + 4 \cdot 0,1) + \\ + 2(3 \cdot 0,1 + 4 \cdot 0,3) + 5(3 \cdot 0,1 + 4 \cdot 0,2) = 7,5.$$

Отже, кореляційний момент

$$\mu_{XY} = 7,5 - 2 \cdot 3,6 = 0,3.$$

Коефіцієнт кореляції

$$\rho_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y} \approx \frac{0,3}{2,1 \cdot 0,49} \approx 0,3.$$

Оскільки $\mu_{XY} \neq 0$, то випадкові величини X і Y залежні. ►

Зі співвідношень (9)–(10) випливає, що коефіцієнт кореляції випадкових величин X та Y дорівнює коваріації відповідних нормованих величин X_σ і Y_σ :

$$\rho_{XY} = \rho_{X_\sigma Y_\sigma} = \text{cov}(X_\sigma, Y_\sigma). \quad (13)$$

Для незалежних випадкових величин $\rho_{XY} = 0$, тому що тоді $\text{cov}(X, Y) = 0$. Але обернене твердження неправильне: з умови $\rho_{XY} = 0$ не випливає незалежність величин X і Y . Рівність $\rho_{XY} = 0$ (або $\text{cov}(X, Y) = 0$) є лише необхідною, але не достатньою умовою незалежності випадкових величин X і Y . З огляду на це доцільно розглянути вужче поняття, аніж поняття залежності двох випадкових величин, а саме поняття корельованості.

Випадкові величини X і Y називають *корельованими*, якщо їхній коефіцієнт кореляції не дорівнює нулеві, і *некорельованими*, якщо коефіцієнт кореляції дорівнює нулеві.

Отже, якщо величини незалежні, то вони некорельовані, але з некорельованих величин можуть бути залежними.

Якщо величини X і Y мають нормальний закон розподілу, то з некорельованості цих величин випливає їх незалежність, тобто для нормально розподілених випадкових величин поняття “некорельованість” і “незалежність” рівносильні.

Властивості коефіцієнта кореляції описують такі теореми.

Теорема 2. Абсолютне значення коефіцієнта кореляції не перевищує одиниці: $|\rho_{XY}| \leq 1$. ▲

Теорема 3. $|\rho_{XY}| = 1$ тоді і лише тоді, коли X і Y пов’язані лінійною функціональною залежністю, тобто $Y = aX + b$, де a і b – сталі, до того ж, якщо $\rho_{XY} = 1$, то $a > 0$, а якщо $\rho_{XY} = -1$, то $a < 0$. ▲

Як впливає з теореми 2, у разі, коли $-1 < \rho_{XY} < 1$, точки (X, Y) розкидані навколо прямої лінії тим щільніше, чим ближча до одиниці величина $|\rho_{XY}|$.

Коефіцієнт кореляції показує, наскільки зв’язок між X і Y є близьким до лінійного функціонального зв’язку. Навіть якщо $\rho_{XY} = 0$, тобто за повної відсутності лінійної залежності між X та Y , між ними може бути сильна стохастична і навіть нелінійна функціональна залежність.

Якщо $\rho_{XY} > 0$, то кажуть, що між X і Y є *додатна кореляція* (*додатний зв’язок*), яка проявляється в тому, що зі збільшенням однієї величини друга також має тенденцію до зростання (у середньому зростає). Якщо $\rho_{XY} < 0$, то між X і Y є *від’ємна кореляція* (*від’ємний зв’язок*), що означає наявність протилежних тенденцій у зміні величин X і Y .

8.9. Лінійна регресія

Здебільшого зв’язок між випадковими величинами X і Y стохастичний, і тоді одну змінну не можна точно подати як функцію іншої.

Розгляньмо випадок, коли змінну Y можна наближено подати у вигляді лінійної функції від X :

$$Y \approx g(X) = aX + b,$$

де a і b – невідомі параметри, які потрібно визначити. Найчастіше для визначення параметрів використовують метод найменших квадратів (МНК).

Критерієм для відшукування a і b тоді є мінімум дисперсії відхилення Y від функції $g(X) = aX + b$:

$$D(Y - aX - b) = M(Y - aX - b)^2 = Q(a, b) \rightarrow \min. \quad (1)$$

Функцію $g(X)$ називають *середньою квадратичною регресією Y на X* .

На практиці часто можна обмежитися лінійною формою залежності між змінними попри простоту цієї залежності. Мірою ступеня лінійності зв'язку між випадковими величинами є, як ми знаємо, коефіцієнт кореляції.

У математичній статистиці здебільшого розглядають функції, диференційовні за своїми коефіцієнтами (параметрами). Тоді необхідною умовою мінімуму функції $Q(a, b)$ є рівність нулеві її перших частинних похідних:

$$\frac{\partial Q(a, b)}{\partial a} = 0, \quad \frac{\partial Q(a, b)}{\partial b} = 0.$$

Ці рівності називають *нормальними рівняннями*.

З цієї системи, з огляду на означення коваріації, дисперсії і коефіцієнта кореляції, отримуємо значення параметрів a і b :

$$a = \rho_{XY} \frac{\sigma_Y}{\sigma_X}, \quad b = m_Y - \rho_{XY} \frac{\sigma_Y}{\sigma_X} m_X. \quad (2)$$

Тепер можна записати дисперсію відхилення $Y - g(X)$. Беручи до уваги вирази (2), матимемо:

$$\begin{aligned} D(Y - aX - b) &= M(Y - aX - b)^2 = \\ &= M\left(Y - \rho_{XY} \frac{\sigma_Y}{\sigma_X} (X - m_X) - m_Y\right)^2. \end{aligned}$$

Після піднесення до квадрата отримаємо:

$$D(Y - aX - b) = \sigma_Y^2 (1 - \rho_{XY}^2). \quad (3)$$

З цього співвідношення видно: що ближче ρ_{XY}^2 до одиниці, то менша дисперсія відхилення Y від найкращого лінійного наближення. У цьому сенсі коефіцієнт кореляції можна вважати показником ступеня лінійності зв'язку двох випадкових величин.

Отже, найкраще (в сенсі МНК) лінійне наближення функції регресії Y на X має вигляд:

$$y - m_Y = \rho_{XY} \frac{\sigma_Y}{\sigma_X} (x - m_X). \quad (4)$$

Цю пряму називають *прямою середньоквадратичної регресії Y на X* .

Величину $\sigma_Y^2(1 - \rho_{XY}^2)$ називають *залишковою дисперсією*. Вона описує коливання випадкової величини Y навколо регресійної прямої і характеризує величину помилки, що спричинена заміною Y на функцію $g(X) = aX + b$. Якщо $\rho_{XY} = \pm 1$, то залишкова дисперсія обертається на нуль: у цьому разі зв'язок між X та Y функціональний.

Рівняння регресії X на Y аналогічне рівнянню (4):

$$x - m_X = \rho_{XY} \frac{\sigma_X}{\sigma_Y} (y - m_Y). \quad (5)$$

Регресійні прямі (4)–(5) проходять через центр спільного розподілу величин X та Y , тобто через *середню точку* $(m_X; m_Y)$.

Кутовий коефіцієнт прямої (4) $k_1 = \operatorname{tg} \alpha = \rho_{XY} \frac{\sigma_Y}{\sigma_X}$, а кутовий коефі-

цієнт прямої (5) $k_2 = \operatorname{tg} \beta = \frac{1}{\rho_{XY}} \cdot \frac{\sigma_Y}{\sigma_X}$. Позаяк $0 < |\rho_{XY}| < 1$, то

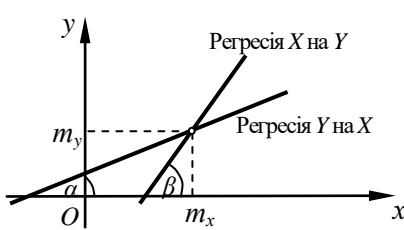


Рис. 8.10

$\frac{1}{|\rho_{XY}|} > |\rho_{XY}|$, тобто пряма (5) нахилена до осі Ox під більшим кутом, аніж пряма (4) (рис. 8.10). Що ближче $|\rho_{XY}|$ до одиниці, то менший кут між прямими (4) та (5).

Якщо $|\rho_{XY}| = 1$, то $k_1 = k_2$, тобто прямі зливаються, як і має бути у разі функціонального зв'язку. Якщо $\rho_{XY} = 0$ (X і Y – некорельовані), то рівняння прямих (4) і (5) набувають вигляду $y = m_Y$ і $x = m_X$, тобто прямі паралельні відповідним осям координат.

8.10. Кореляційна матриця

Нехай задано n -вимірну випадкову величину (випадковий вектор)

$$Z = (X_1, X_2, \dots, X_n). \quad (1)$$

Математичним сподіванням випадкової величини (1) називають випадковий вектор, компонентами якого є математичні сподівання складових цієї випадкової величини:

$$M(Z) = (m_{X_1}; m_{X_2}; \dots; m_{X_n}). \quad (2)$$

Коваріаційною (дисперсійно-коваріаційною) матрицею випадкової величини $Z = (X_1, X_2, \dots, X_n)$ називають матрицю n -го порядку, елементами якої є коваріації $\sigma_{ij} = \text{cov}(X_i, X_j)$ кожної пари складових X_i і X_j :

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{pmatrix}. \quad (3)$$

Діагональні елементи коваріаційної матриці дорівнюють дисперсіям випадкових величин X_1, X_2, \dots, X_n :

$$\sigma_{ii} \equiv \sigma_i^2 = \text{cov}(X_i, X_i) = \sigma_{X_i}^2, \quad i = \overline{1, n}. \quad (4)$$

Визначник коваріаційної матриці називають *узагальненою дисперсією*. Узагальнена дисперсія $\det \Sigma \equiv |\Sigma|$ може слугувати мірою розсіяння n -вимірної випадкової величини.

З огляду на симетричність коваріаційної матриці, для її задання потрібно $n(n+1)/2$ елементів (а не n^2). Тому її записують також у вигляді:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ & \sigma_{22} & \dots & \sigma_{2n} \\ & & \dots & \dots \\ & & & \sigma_{nn} \end{pmatrix}. \quad (5)$$

Математичне сподівання і коваріаційна матриця є основними числовими характеристиками випадкового вектора Z .

Дисперсія суми n випадкових величин X_1, X_2, \dots, X_n дорівнює сумі елементів коваріаційної матриці цих величин:

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}.$$

Оскільки для некорельованих, зокрема для незалежних, випадкових величин коваріаційна матриця діагональна, тобто

$$\Sigma = \begin{pmatrix} \sigma_{X_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{X_2}^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_{X_n}^2 \end{pmatrix},$$

то дисперсія суми некорельованих випадкових величин дорівнює сумі їхніх дисперсій:

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \sigma_{X_i}^2.$$

Зрозуміло, що взаємозв'язок випадкових величин (1) можна описувати за допомогою коефіцієнта кореляції

$$\rho_{ij} \equiv \rho_{X_i X_j} = \frac{\text{cov}(X_i, X_j)}{\sigma_{ij}}.$$

Тоді отримаємо *кореляційну матрицю* (нормовану коваріаційну матрицю), елементами якої є коефіцієнти кореляції:

$$R^* = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1m} \\ \rho_{21} & 1 & \dots & \rho_{2m} \\ \dots & \dots & \dots & \dots \\ \rho_{m1} & \rho_{m2} & \dots & 1 \end{pmatrix}. \quad (10)$$

Приклад. Розподіл двовимірної випадкової величини має вигляд:

$Y \backslash X$	2	3	5	6	p_j
0	0,1	0	0,2	0	0,3
1	0	0,1	0,1	0,05	0,25
4	0,2	0,1	0,1	0,05	0,45
p_i	0,3	0,2	0,4	0,1	1

Побудуємо коваріаційну і кореляційну матриці.

Розв'язання. Відповідно до формул (3) і (10), для розв'язання задачі треба знайти коваріацію і дисперсії її складових X і Y . За формулою (6) п. 8.11

$$\mu_{XY} = \sigma_{XY} = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_{ij} - M(X)M(Y).$$

Для даних, наведених у таблиці, отримуємо:

$$M(X) = \sum_{i=1}^4 x_i p_i = 2 \cdot 0,3 + 3 \cdot 0,2 + 5 \cdot 0,4 + 6 \cdot 0,1 = 3,8,$$

$$M(Y) = \sum_{j=1}^3 y_j p_j = 0 \cdot 0,3 + 1 \cdot 0,25 + 4 \cdot 0,45 = 2,05,$$

$$M(XY) = \sum_{i=1}^4 \sum_{j=1}^3 x_i y_j p_{ij} = 2(1 \cdot 0 + 4 \cdot 0,2) + 3(1 \cdot 0,1 + 4 \cdot 0,1) + 5(1 \cdot 0,1 + 4 \cdot 0,1) + 6(1 \cdot 0,05 + 4 \cdot 0,05) = 7,1.$$

Позаяк $M(X) \cdot M(Y) \neq M(XY)$, то величини X і Y залежні. Отже,

$$\text{cov}(X, Y) = \text{cov}(Y, X) = 7,1 - 3,8 \cdot 2,05 = -0,69.$$

Відшукаймо дисперсії складових.

$$M(X^2) = \sum_{i=1}^4 x_i^2 p_i = 4 \cdot 0,3 + 9 \cdot 0,2 + 25 \cdot 0,4 + 36 \cdot 0,1 = 16,6,$$

$$D(X) = M(X^2) - M^2(X) = 16,6 - 3,8^2 = 2,16.$$

Аналогічно знаходимо дисперсію складової Y :

$$M(Y^2) = \sum_{j=1}^3 y_j^2 p_j = 0 + 1 \cdot 0,25 + 16 \cdot 0,45 = 7,45.$$

$$D(Y) = 7,45 - 2,05^2 = 3,2475.$$

Отже, коваріаційна матриця $\Sigma = \begin{pmatrix} 2,16 & -0,69 \\ -0,69 & 3,2475 \end{pmatrix}$.

Середні квадратичні відхилення складових $\sigma_X \approx 1,4697$, $\sigma_Y \approx 1,8021$; коефіцієнт кореляції

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} = -\frac{0,69}{2,6485} = -0,26.$$

Отже, кореляційна матриця має вигляд:

$$R^* = \begin{pmatrix} 1 & -0,26 \\ -0,26 & 1 \end{pmatrix}. \blacktriangleright$$

8.11. Двовимірний нормальний розподіл. Нормальна кореляція

Двовимірним нормальним розподілом називають розподіл двовимірної випадкової величини (X, Y) , щільність імовірності якого

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2)\right), \quad (1)$$

де

$$u = (x - m_X)/\sigma_X, \quad v = (y - m_Y)/\sigma_Y.$$

Отже, нормальний закон на площині визначається п'ятьма параметрами: $m_X, m_Y, \sigma_X, \sigma_Y, \rho_{XY}$. Ці параметри мають такий зміст: m_X, m_Y – математичні сподівання складових X і Y , σ_X, σ_Y – середні квадратичні відхилення складових X і Y , $\rho = \rho_{XY}$ – коефіцієнт кореляції між X і Y .

Якщо складові X і Y некорельовані, тобто $\rho = 0$, то функцію розподілу (1) можна подати у вигляді:

$$f(x, y) = f_1(x) \cdot f_2(y), \quad (2)$$

де

$$f_1(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \cdot e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}, \quad f_2(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \cdot e^{-\frac{(y-m_Y)^2}{2\sigma_Y^2}}.$$

Рівність (2) є означенням незалежності випадкових величин X і Y .

Отже, якщо складові двовимірної нормально розподіленої випадкової величини некорельовані, то вони водночас є незалежними. Для випадкових величин, коливання яких підлягають нормальному закону, некорельованість ($\rho = 0$) є достатньою умовою незалежності.

Якщо функції регресії Y на X і X на Y лінійні, то кажуть, що X і Y пов'язані лінійною кореляційною залежністю. Графіки лінійних функцій регресії збігаються з прямими середньоквадратичної регресії.

Якщо спільним розподілом ймовірностей величин X і Y є двовимірний нормальний розподіл, то кажуть, що між величинами X і Y існує нормальна кореляція. Нормальна кореляція завжди лінійна.

Задачі до розділу 8

1. Розподіл системи двох випадкових величин (X, Y) має вигляд:

1)

$Y \setminus X$	1	5	7
3	0,25	0,15	0,32
8	0,1	0,05	0,13

2)

$Y \setminus X$	-3	-1	2
1	0,2	0,1	0,1
9	0,1	0,3	0,2

3)

$Y \setminus X$	-1	4
-1	0,2	0,25
0	0,05	0,15
1	0,25	0,1

4)

$Y \setminus X$	0	1
-3	$\frac{1}{9}$	$\frac{1}{2}$
-1	$\frac{1}{18}$	$\frac{1}{6}$
0	$\frac{1}{18}$	$\frac{1}{9}$

Потрібно визначити:

- 1) $P(x_1 \leq X < x_3, y_1 \leq Y < y_2)$;
- 2) розподіли складових X і Y ;
- 3) математичні сподівання складових;
- 4) дисперсії складових;
- 5) умовні розподіли кожної складової, якщо друга складова набуває свого першого значення;
- 6) умовні математичні сподівання;
- 7) коваріацію і коефіцієнт кореляції.

2. Задано розподіл двовимірної випадкової величини (X, Y) :

а)

$Y \setminus X$	-1	1
3	0,1	0,3
8	0,1	0,5

б)

$Y \setminus X$	0	1	3
1	0,06	0,09	0,15
2	0,14	0,21	0,35

Побудувати коваріаційну та кореляційну матриці цієї випадкової величини. Зробити висновки.

3. Компонентами випадкової величини (X, Y) є відповідно кількість влучень у мішень і кількість промахів. Роблять два постріли з імовірністю влучення $p = 0,8$ у кожному пострілі. Побудувати закон розподілу випадкової величини (X, Y) і знайти її числові характеристики. Скласти коваріаційну та кореляційну матриці випадкової величини (X, Y) .

ЧАСТИНА II. МАТЕМАТИЧНА СТАТИСТИКА

РОЗДІЛ 10

ВИБІРКА

10.1. Основні поняття теорії вибіркового методу

Основне завдання математичної статистики – вивчення масових явищ і процесів на підставі результатів *спостережень* або *експериментів*. Математична статистика вивчає методiku систематизації, обробки й використання статистичних даних та виявляє статистичні закономірності.

Статистичне спостереження може бути суцільним і вибірковим.

Усю множину об'єктів, яка вивчається, називають генеральною сукупністю. Частину об'єктів, яку відібрано для вивчення, називають *вбірковою сукупністю* (*вбіркою*). Вибіркове спостереження є одним із основних методів отримання статистичної інформації, а сам метод називають *вбірковим*. Якщо відбір об'єктів і спосіб їх дослідження науково обґрунтовані, то висновки, отримані на підставі вибірки, цілком задовольняють потреби практики.

Математична статистика базується на теорії ймовірностей. Завданням математичної статистики є оцінювання характеристик генеральної сукупності за вибірковими даними. Відомості, отримані за вибіркою, неминуче містять похибку, бо ґрунтуються на вивченні не всієї сукупності, а лише деякі її частини. Звідси випливають дві задачі математичної теорії вибірки: 1) організація вибіркового спостереження в такий спосіб, щоб отримати найбільш правдиву й повну інформацію (проблема репрезентативності вибірки); 2) використання вибіркових результатів так, щоб висновки про властивості й параметри генеральної сукупності були якомога надійнішими (проблема оцінювання).

Для підвищення вірогідності результатів дослідження треба правильно організувати статистичне спостереження. Для цього насамперед має бути чітко сформульована його ціль і точно визначена статистична сукупність. Однією з найважливіших вимог до процесу організації статистичного спостереження є *однорідність* сукупності. Це означає, що окремі елементи сукупності мають бути *нерозрізними*, точніше, *зіставні* за всіма ознаками, окрім тієї, що вивчається.

Є два істотно різні способи відбору елементів генеральної сукупності у вибірку: випадковий і невипадковий.

Випадковий відбір – це такий відбір, за якого кожен елемент генеральної сукупності може потрапити у вибірку випадково. Такий відбір є послідовністю повторних випробувань і тому дає змогу застосувати для вивчення вибірки ймовірнісну модель, закон великих чисел, методи дослідження закономірностей випадкових явищ.

Випадкова вибірка може бути повторною і безповторною. У разі повторної вибірки взятий елемент після обстеження знову повертають у набір. У разі формування безповторної вибірки відібрані елементи не повертають у набір.

Для скінченної генеральної сукупності випадковий безповторний відбір веде до залежності окремих випробувань, а за повторного вибору випробування незалежні. З математичного погляду останній випадок простіший, його моделлю є схема повторних незалежних випробувань Бернуллі, але на практиці здебільшого трапляються безповторні вибірки. Відмінність між цими вибірками незначна, якщо їхні обсяги становлять не більше 5–10% обсягу генеральної сукупності. Тоді відмінностями можна знехтувати і для безповторної вибірки використовувати простіші співвідношення для моделі повторних незалежних випробувань.

За *механічного* способу утворення вибірки для обстеження з генеральної сукупності вибираються об'єкти через певний інтервал, наприклад, кожен двадцятий (п'ятивідсоткова вибірка), кожен десятий (десятивідсоткова вибірка) тощо. Цим способом можна скористатися за умови, що елементи генеральної сукупності мають певний порядок (наприклад, поява однотипних виробів із конвеєра).

Типовий спосіб утворення вибірки полягає в тому, що генеральну сукупність розбивають на групи за певними ознаками. Ці групи не перекриваються між собою. Так, для вивчення попиту на товар сукупність можна розбити на групи, що різняться за рівнем доходу, соціальною приналежністю, географічним розташуванням, а потім з кожної групи за схемою випадкової вибірки беруть певну (загалом різну) кількість елементів. Усі вибрані елементи утворюють *типову вибірку*.

За *серійного* способу формування вибіркової сукупності генеральну сукупність поділяють на серії (групи), що не перекриваються, і за схемою випадкової вибірки відбирають певну кількість серій, усі члени яких і утворюють серійну вибірку.

На практиці для утворення вибірки водночас можуть застосовуватися різні методи відбору.

10.2. Найпростіші характеристики вибірки

10.2.1. Статистичний розподіл вибірки. Нехай із генеральної сукупності X зроблено вибірку, і значення ознаки (випадкової величини X), яке дорівнює x_1 , спостерігаються n_1 разів, значення ознаки, яке дорівнює $x_2 - n_2$ разів, ..., значення ознаки, яке дорівнює $x_m - n_m$ разів. Обсяг вибірки $\sum_{i=1}^m n_i = n$. Спостережувані значення x_i називають *варіантами*, числа $n_i -$ *статистичними частотами варіант*, а відношення $w_i = n_i/n -$ *відносними статистичними частотами варіант*. Зрозуміло, що $\sum_{i=1}^m w_i = 1$.

Зміну значення ознаки називають її *варіюванням*.

Послідовність варіант, що записані в порядку зростання, називають *статистичним*, або *варіаційним*, *рядом* (*дискретним варіаційним рядом*).

Статистичним розподілом вибірки називають перелік варіант і відповідних їм частот або відносних частот. Статистичні розподіли задають у вигляді таблиць.

x_i	x_1	x_2	...	x_i	...	x_m
n_i	n_1	n_2	...	n_i	...	n_m

x_i	x_1	x_2	...	x_i	...	x_m
w_i	w_1	w_2	...	w_i	...	w_m

Якщо ознака може набувати не лише дискретних, а й неперервних значень, або кількість різних варіант велика, то для опису вибірки використовують *інтервальний статистичний розподіл* – відповідність між інтервалами значень варіаційного ряду і статистичними частотами або відносними статистичними частотами. Для цього область реалізації розбивають на m інтервалів і для кожного інтервалу визначають частоти. Зазвичай частинні інтервали вибирають так, щоб вони мали однакову довжину. Довжина інтервалу має бути оптимальною. Це означає, що варіаційний ряд не має бути надто громіздким, і водночас він має відображати характерні особливості явища. Часто цю довжину визначають за формулою

$$\Delta x = \frac{x_{\max} - x_{\min}}{1 + 3,2 \lg n}$$

Інколи визначають не довжину, а кількість інтервалів, вважаючи, що вона має задовольняти умову $m \leq 5 \lg n$. Слід зазначити також, що на практиці нерідко спосіб розбиття на інтервали є очевидним.

Інтервальний розподіл вибірки має вигляд:

$(z_{i-1}; z_i]$	$(z_0; z_1]$	$(z_1; z_2]$...	$(z_{m-1}; z_m]$
n_i	n_1	n_2	...	n_m

Аналогічно записується інтервальний розподіл відносних частот.

Іноді неперервний інтервальний статистичний розподіл замінюють дискретним. Для цього частинні інтервали $(z_{i-1}; z_i]$ замінюють числами, які є серединами інтервалів, тобто приймають

$$x_i = \frac{1}{2}(z_{i-1} + z_i).$$

10.2.2. Числові характеристики варіаційного ряду. *Генеральною середньою* називають середнє значення ознаки генеральної сукупності. Якщо N – обсяг генеральної сукупності і всі значення ознаки x_1, x_2, \dots, x_N різні, то генеральна середня

$$\bar{x}_G = \frac{1}{N}(x_1 + x_2 + \dots + x_N).$$

Якщо значення x_1, x_2, \dots, x_k мають відповідно частоти N_1, N_2, \dots, N_k , то

$$\bar{x}_G = \frac{1}{N}(x_1 N_1 + x_2 N_2 + \dots + x_m N_m) = \frac{1}{N} \sum_{i=1}^m x_i N_i = \sum_{i=1}^m x_i w_i,$$

де $N_1 + N_2 + \dots + N_m = N$.

Як бачимо, генеральна середня – це середня зважена з вагами, що дорівнюють відповідним частотам.

Аналогічно розраховується *вибіркова середня* – середнє арифметичне значення ознаки у вибірці. Якщо n – обсяг вибірки і всі $x_i, i = \overline{1, n}$ різні, то

$$\bar{x} = \bar{x}_B = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Якщо значення x_1, \dots, x_m мають відповідно частоти n_1, \dots, n_m , то

$$\bar{x} = \bar{x}_B = \frac{1}{n} \sum_{i=1}^m x_i n_i = \sum_{i=1}^m x_i w_i, \text{ де } n_1 + n_2 + \dots + n_m = n.$$

ЗАУВАГА. Якщо тлумачити ознаку X генеральної сукупності як випадкову величину, то математичне сподівання $m_X = a$ ознаки дорівнює генеральній середній: $m_X = \bar{x}_G$. Тому далі поняття генеральної середньої і математичного сподівання вважаються еквівалентними, тобто $\bar{x}_G = a$. ☉

Відхиленням називають різницю між значенням ознаки та її середнім:

$$x_i - \bar{x}.$$

Вибіркова дисперсія

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2 \text{ або } D_B = \frac{1}{n} \sum_{i=1}^m n_i (x_i - \bar{x}_B)^2.$$

Вибіркове середнє квадратичне відхилення (вибіркове стандартне відхилення) визначають за формулою: $\sigma_B = \sqrt{D_B}$. Часто стандартне відхилення вибіркового розподілу називають *стандартною похибкою*.

Аналогічно визначають генеральну дисперсію D_G і генеральне середнє квадратичне відхилення $\sigma_G = \sqrt{D_G}$.

Дисперсія і середнє квадратичне відхилення описують розсіяння спостережуваних значень навколо вибіркового середнього.

Зазвичай на практиці для відшукання дисперсії використовують зручнішу формулу, а саме:

$$D = \overline{x^2} - \bar{x}^2, \text{ де } \overline{x^2} = \frac{1}{n} \sum_{i=1}^m n_i x_i^2.$$

У правильності цієї формули нескладно пересвідчитися. Справді,

$$\begin{aligned} D &= \frac{1}{n} \sum_{i=1}^n n_i (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^m n_i (x_i^2 - 2x_i \bar{x} + \bar{x}^2) = \\ &= \frac{1}{n} \left(\sum_{i=1}^m n_i x_i^2 - 2\bar{x} \sum_{i=1}^m n_i x_i + \bar{x}^2 \sum_{i=1}^m n_i \right) = \overline{x^2} - 2\bar{x}^2 + \bar{x}^2 = \overline{x^2} - \bar{x}^2. \end{aligned}$$

Окрім дисперсії і середнього, є й інші характеристики варіаційного ряду: мода, медіана, розмах варіації, коефіцієнт варіації.

Модою називають варіанту, яка має найбільшу частоту.

Медіаною називають варіанту, яка поділяє варіаційний ряд на дві рівні (за кількістю варіант) частини. Якщо n – непарне, тобто $n = 2k + 1$, то

$$\text{Мех} = x_{k+1}; \text{ якщо } n \text{ – парне, тобто } n = 2k, \text{ то } \text{Мех} = \frac{1}{2} (x_k + x_{k+1}).$$

Розмахом варіації R називають різницю між найбільшою та найменшою варіантами: $R = x_{\max} - x_{\min}$.

Коефіцієнт варіації V – це відношення середнього квадратичного відхилення до середнього:

$$V = \frac{\sigma_B}{\bar{x}} = \frac{\sqrt{\sum (x_i - \bar{x})^2} / \sqrt{n}}{\sum x_i / n}.$$

Часто коефіцієнт варіації виражають у відсотках, помноживши праву частину формули на 100 %.

Центральним емпіричним моментом k -го порядку μ_k статистичного розподілу вибірки називають середнє арифметичне k -го степеня відхилень варіант від їхнього середнього значення, тобто

$$\mu_k = \frac{1}{n} \sum_{i=1}^m n_i (x_i - \bar{x}_B)^k.$$

Зокрема, $\mu_0 = 1$, $\mu_1 = 0$, $\mu_2 = D_B$.

Асиметрію й ексцес статистичного розподілу вибірки визначають так:

$$A = \frac{\mu_3}{\sigma_B^3}, \quad E = \frac{\mu_4}{\sigma_B^4} - 3.$$

Як відомо, асиметрія й ексцес нормально розподіленої випадкової величини дорівнюють нулю (п. 5.7). Отже, що більше віддаляються від нуля асиметрія й ексцес статистичного розподілу вибірки, то менше підстав вважати, що вибірка, з якої утворено варіаційний ряд, зроблена з нормально розподіленої генеральної сукупності.

10.3. Емпірична функція розподілу

Нехай задано статистичний розподіл кількісної ознаки X . Нехай n – загальна кількість спостережень (обсяг вибірки), а n_x – кількість спостережень, в яких значення ознаки менші від x . Тоді $w_x = n_x/n$ – це відносна частота події $X < x$. Відносна частота є функцією від x .

Емпіричною функцією розподілу (функцією розподілу вибірки) називають функцію $F_n(x)$, яка для кожного значення x дорівнює відносній частоті події $X < x$:

$$F_n(x) = \frac{n_x}{n}. \quad (1)$$

Функцію розподілу генеральної сукупності

$$F(x) = P(X < x) \quad (2)$$

називають *теоретичною інтегральною функцією розподілу*. З означень зрозуміла відмінність між функціями $F(x)$ і $F_n(x)$: $F(x)$ – це ймовірність події $X < x$, а $F_n(x)$ – відносна частота цієї події. За теоремою Бернуллі, відносна частота події $X < x$ прямує за ймовірністю до ймовірності цієї події, тобто

$$F(x) = \text{plim}_{n \rightarrow \infty} F_n(x).$$

Тобто за достатньо великих значень n для кожного фіксованого x числа $F_n(x)$ і $F(x)$ мало відрізняються одне від одного в тому сенсі, що

$$\lim_{n \rightarrow \infty} P(|F(x) - F_n(x)| < \varepsilon) = 1.$$

Звідси зрозуміло, що емпіричну функцію розподілу можна використовувати як наближення (*оцінку*) теоретичної функції розподілу генеральної сукупності. За міру точності цього наближення можна взяти модуль максимальної різниці значень теоретичної і статистичної функцій розподілу:

$$\max_x |F(x) - F_n(x)|.$$

Зауважмо, що є критерії, на підставі яких можна оцінити істотність знайденого відхилення, зокрема, критерій Колмогорова.

Нескладно пересвідчитися, що функція $F_n(x)$ має такі самі властивості, як і функція $F(x)$:

- 1) значення $F_n(x)$ належать проміжковій $[0; 1]$: $0 \leq F_n(x) \leq 1$;
- 2) $F_n(x)$ є неспадною функцією;
- 3) $F_n(x) = 0$, якщо $x \leq x_{\min}$, і $F_n(x) = 1$, якщо $x > x_{\max}$, де x_{\min} та x_{\max} – найменша та найбільша варіанти відповідно.

Емпірична функція дискретного розподілу виглядає так:

$$F_n(x) = \begin{cases} 0, & x \leq x_1, \\ n_1/n, & x_1 < x \leq x_2, \\ (n_1 + n_2)/n, & x_2 < x \leq x_3, \\ \dots & \dots \\ (n_1 + n_2 + \dots + n_{k-1})/n, & x_{k-1} < x \leq x_k, \\ 1, & x > x_k. \end{cases} \quad (3)$$

Графік функції – розривну східчасту криву – зображено на рис. 10.1.

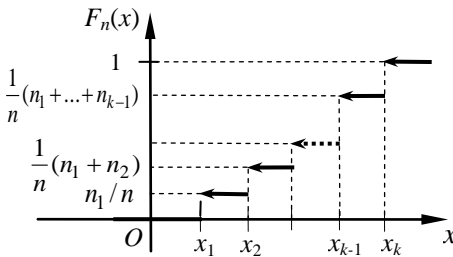


Рис. 10.1

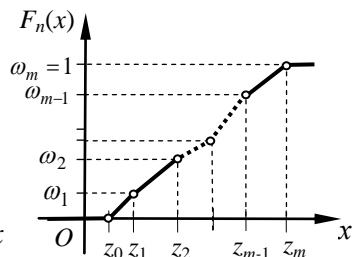


Рис. 10.2

Емпірична функція інтервального статистичного розподілу:

$$F_n(x) = \begin{cases} 0, & x \leq z_0, \\ F_1^*(x), & z_0 < x \leq z_1, \\ F_2^*(x), & z_1 < x \leq z_2, \\ \dots & \dots \\ F_k^*(x), & z_{k-1} < x \leq z_k, \\ 1, & x > z_k, \end{cases} \quad (4)$$

де

$$F_i^*(x) = \frac{w_i}{z_i - z_{i-1}}(x - z_{i-1}) + \omega_{i-1}, \quad (5)$$

$$\omega_{i-1} = w_1 + w_2 + \dots + w_{i-1}, \quad i = \overline{1, k}, \quad (6)$$

Графіком емпіричної функції інтервального статистичного розподілу є неперервна ламана лінія, складена з відрізків прямих, які з'єднують точки $(z_i; \omega_i)$, де $\omega_i = w_1 + w_2 + \dots + w_i$ (рис. 10.2).

10.4. Полігон і гістограма

Окрім функції розподілу важливими графічними характеристиками вибірки є, зокрема, полігон та гістограма.

Полігон розподілу – це графічне зображення статистичного ряду. Відкладемо на осі Ox варіанти x_i , а на осі Oy – відповідні їм частоти n_i , $i = \overline{1, k}$. З'єднавши точки $(x_1; n_1)$, $(x_2; n_2)$, ..., $(x_k; n_k)$, відрізками прямих, дістанемо ламану, яку називають *полігоном частот*.

Аналогічно можна побудувати *полігон відносних частот* – ламану, відрізки якої з'єднують точки $(x_1; w_1)$, $(x_2; w_2)$, ..., $(x_k; w_k)$. Для конкретних вибірок полігон частот і полігон відносних частот зображено на рис. 10.3.

Для геометричного зображення інтервального статистичного розподілу використовують *гістограму частот* (*гістограму відносних частот*).

Розбиймо весь інтервал спостережуваних значень ознаки на кілька частинних інтервалів $[z_0; z_1]$, $(z_1; z_2]$, ..., $(z_{k-1}; z_k]$ довжиною ℓ . *Гістограмою частот* називають східчасту фігуру, яка складена з прямокутників, основами яких є частинні інтервали довжиною ℓ , а висоти прямокутників $h_i = n_i / \ell$ (n_i / ℓ – щільність (густина) частот).

Щоб побудувати гістограму частот, на осі абсцис відкладають частинні інтервали $(z_{i-1}; z_i)$ і на них як на основах будують прямокутники з висотами h_i . Гістограму конкретного статистичного розподілу зображено на рис. 10.5.

Площа i -го прямокутника дорівнює сумі частот варіант i -го інтервалу, тобто $\ell \cdot \frac{n_i}{\ell} = n_i$. Отже, площа гістограми частот дорівнює сумі всіх частот, тобто обсягові вибірки.

Аналогічно можна побудувати *гістограму відносних частот* – східчасту фігуру, що складається з прямокутників з основами ℓ і висотами $h_i = \frac{w_i}{\ell}$ ($\frac{w_i}{\ell}$ – щільність відносних частот).

Площа i -го частинного прямокутника дорівнює відносній частоті відповідної варіанти: $\ell \cdot h_i = \ell \cdot \frac{w_i}{\ell} = w_i$. Отже, площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто, як було вже доведено, одиниці.

Полігон і гістограма відносних частот дають наближене уявлення про криву розподілу генеральної сукупності.

Приклад. Для вивчення ефективності діяльності підприємств регіону побудовано вибірку (відношення прибутку до обсягів виробництва у відсотках) 9, 8, 6, 7, 9, 10, 11, 8, 9, 9, 8, 6, 9, 8, 10, 10, 12, 7, 7, 10. На підставі цих даних потрібно:

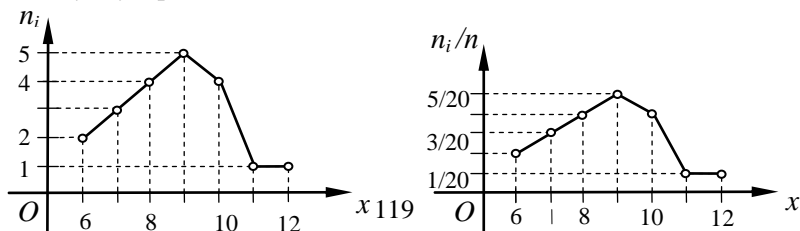
- 1) побудувати статистичний розподіл вибірки, полігони частот і відносних частот та емпіричну функцію розподілу;
- 2) вирахувати вибіркові середню та середнє квадратичне відхилення і зробити аналіз результатів;
- 3) відшукати моду, медіану, розмах та коефіцієнт варіації;
- 4) побудувати інтервальний статистичний розподіл відносних частот, розбивши діапазон даних на 4 рівні проміжки; побудувати гістограму й емпіричну функцію інтервального розподілу;

Розв'язання. Випадкова величина X (відсоткове відношення прибутку підприємства до обсягів виробництва) набуває значень x_i , $i = \overline{1, 7}$, вказаних у вибірці. Обсяг вибірки $n = 20$.

1) Статистичні розподіли частот і відносних частот виглядають так:

x_i	6	7	8	9	10	11	12
n_i	2	3	4	5	4	1	1
w_i	2/20	3/20	4/20	5/20	4/20	1/20	1/20

Полігон частот – це ламана, відрізки якої з'єднують точки $(x_i; n_i)$ (рис. 10.3, а)), а полігон відносних частот – це ламана, відрізки якої з'єднують точки $(x_i; w_i)$ (рис. 10.3, б)).



а)

б)

Рис. 10.3

За статистичним розподілом будемо емпіричну функцію розподілу.

Найменша варіанта $x_1 = 6$, тому $F_n(x) = 0$, якщо $x \leq 6$.

Значення $x < 7$ (тобто $x = 6$) спостерігається 2 рази, тому

$$F_n(x) = 2/20, \text{ якщо } 6 < x \leq 7.$$

Значення $x < 8$ (тобто $x = 6, x = 7$) спостерігається $2 + 3 = 5$ разів, тому

$$F_n(x) = 5/20, \text{ якщо } 7 < x \leq 8.$$

Міркуючи аналогічно, отримаємо емпіричну функцію розподілу:

$$F_n(x) = \begin{cases} 0, & x \leq 6, \\ 2/20, & 6 < x \leq 7, \\ 5/20, & 7 < x \leq 8, \\ 9/20, & 8 < x \leq 9, \\ 14/20, & 9 < x \leq 10, \\ 18/20, & 10 < x \leq 11, \\ 19/20, & 11 < x \leq 12, \\ 1, & x > 12. \end{cases}$$

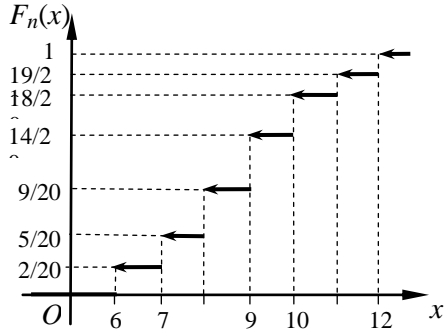


Рис. 10.4

Графік функції зображено на рис. 10.4.

2) Вибіркова середня:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^7 x_i n_i = \frac{1}{20} (6 \cdot 2 + \dots + 12 \cdot 1) = \frac{173}{20} = 8,65.$$

Вибіркову дисперсію обчислюємо за однією з формул:

$$D_B = \frac{1}{n} \sum n_i (x_i - \bar{x})^2 \text{ або } D_B = \overline{x^2} - \bar{x}^2 = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2.$$

Скориставшись, наприклад, другою формулою, матимемо:

$$D_B = \frac{1}{20} \cdot (6^2 \cdot 2 + 7^2 \cdot 3 + 8^2 \cdot 4 + 9^2 \cdot 5 + 10^2 \cdot 4 +$$

$$+11^2 \cdot 1 + 12^2 \cdot 1) - 8,65^2 = \frac{1}{20} \cdot 1545 - 74,8225 = 2,4275 .$$

Вибіркове середнє квадратичне відхилення:

$$\sigma_B = \sqrt{D_B} = \sqrt{2,4275} \approx 1,56 .$$

З отриманих результатів випливає, що середній прибуток одного підприємства становить 8,65 % від обсягу виробництва. Приблизно можна вважати, що відсотковий прибуток підприємств коливається в межах $\bar{x} \pm \sigma_B$, тобто від 7,09 % до 10,21 %.

3) Мода (варіанта з найбільшою частотою $n_4 = 5$) дорівнює 9 ; медіана також дорівнює 9 :

$$MoX = MeX = 9 .$$

Розмах варіації:

$$R = x_{\max} - x_{\min} = 12 - 6 = 6 .$$

Коефіцієнт варіації:

$$V = \frac{\sigma_B}{\bar{x}} \cdot 100\% = \frac{1,56}{8,65} \cdot 100\% \approx 18\% .$$

4) Довжина частинного інтервалу дорівнює $1/4$ довжини проміжку, тобто $R/4 = 1,5$. Інтервальний статистичний розподіл частот і відносних частот можна подати у вигляді таблиці:

$(z_{i-1}; z_i]$	[6; 7,5]	(7,5; 9]	(9; 10,5]	(10,5; 12]
n_i	5	9	4	2
w_i	5/20	9/20	4/20	2/20

Гістограма – східчаста фігура (рис. 10.5), що складається з прямокутників з основами $[z_{i-1}; z_i] = 1,5$ та висотами

$$h_1 = \frac{5}{1,5} \approx 3,3, \quad h_2 = \frac{9}{1,5} = 6, \quad h_3 = \frac{4}{1,5} \approx 2,7, \quad h_4 = \frac{2}{1,5} \approx 1,3 .$$

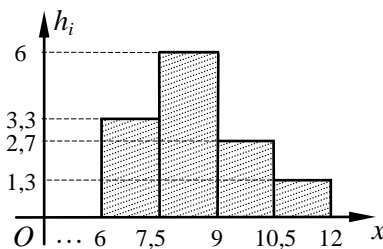


Рис. 10.5

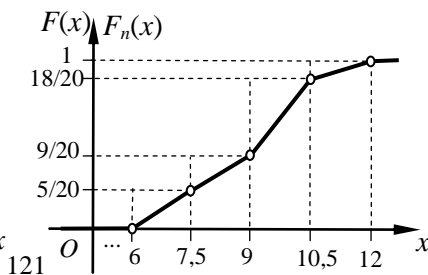


Рис. 10.6

Емпірична функція інтервального статистичного розподілу відповідно до формули (4) п. 10.3 має вигляд:

$$F_n(x) = \begin{cases} 0, & x \leq 6, \\ \frac{1}{6}(x-6), & 6 < x \leq 7,5, \\ \frac{3}{10}(x-7,5) + \frac{5}{20}, & 7,5 < x \leq 9, \\ \frac{2}{15}(x-9) + \frac{14}{20}, & 9 < x \leq 10,5, \\ \frac{1}{15}(x-10,5) + \frac{18}{20}, & 10,5 < x \leq 12, \\ 1, & x > 12. \end{cases} \quad (\text{a})$$

Графік функції зображено ламаною лінією на рис. 10.6.

Задачі до розділу 10

1. Кількість X комп'ютерів, проданих у восьми магазинах упродовж тижня, подано в таблиці:

№ магазину	1	2	3	4	5	6	7	8
X	3	6	5	6	8	8	5	8

Побудувати:

- 1) варіаційний ряд;
- 2) статистичний розподіл частот та відносних частот;
- 3) числові характеристики вибірки: обсяг, найменше та найбільше значення варіанти, розмах, моду, медіану, середню та дисперсію, коефіцієнт варіації;
- 4) полігон частот та відносних частот;
- 5) емпіричну функцію розподілу та її графік.

2. Для вивчення генеральної сукупності зроблено вибірку, яку подано у вигляді такого статистичного розподілу частот:

a)	<table border="1" style="display: inline-table;"><tr><td>x_i</td><td>4</td><td>7</td><td>9</td></tr><tr><td>n_i</td><td>2</td><td>3</td><td>5</td></tr></table>	x_i	4	7	9	n_i	2	3	5
x_i	4	7	9						
n_i	2	3	5						

b)	<table border="1" style="display: inline-table;"><tr><td>x_i</td><td>5</td><td>7</td><td>12</td><td>16</td></tr><tr><td>n_i</td><td>3</td><td>6</td><td>7</td><td>4</td></tr></table>	x_i	5	7	12	16	n_i	3	6	7	4
x_i	5	7	12	16							
n_i	3	6	7	4							

Побудувати розподіл відносних частот. Знайти вибіркву середню, вибіркву дисперсію і коефіцієнт варіації. Побудувати емпіричну функцію розподілу і її графік.

3. Для вивчення якості тієї самої продукції, що випускається десятима різними виробниками, обстежено по 100 одиниць продукції кожного підприємства. Внаслідок обстеження серед кожних 100 виробів виявлено таку кількість неякісних:

3, 1, 0, 2, 1, 4, 0, 3, 1, 1.

Побудувати варіаційний ряд і статистичні розподіли частот та відносних частот. Знайти коефіцієнт варіації, моду і медіану. Побудувати емпіричну функцію розподілу і її графік. Побудувати полігон частот і полігон відносних частот.

4. Внаслідок вибіркового обстеження неперервної ознаки генеральної сукупності отримано такий інтервальный статистичний розподіл вибірки:

a)	$(z_{i-1}; z_i]$	(2;5]	(5;8]	(8;11]	(11;14]
	n_i	5	14	7	4

b)	$(z_{i-1}; z_i]$	(1; 5]	(5; 9]	(9; 13]	(13; 17]	(17; 21]
	n_i	9	22	31	28	10

Побудувати гістограму, емпіричну функцію розподілу та її графік.

5. Побудувати гістограму і полігон частот для такого інтервального статистичного розподілу відносних частот:

$(z_{i-1}; z_i]$	(0; 1,5]	(1,5; 3]	(3; 4,5]	(4,5; 6]
w_i	?	0,2	0,4	0,1