ЛЕКЦІЯ 6

ВИПАДКОВІ ВЕЛИЧИНИ

6.1. Випадкові величини та закони їх розподілу

Означення та види випадкових величин

В практичній діяльності часто зустрічаються експерименти, випробування, досліди, результатами яких є чисельні значення. Наприклад, кількість замовлень на авіаквитки, що надходить до системи бронювання та продажу квитків впродовж часу t, не є сталою величиною і може приймати різні значення $0,1,2, \ldots, n, \ldots$ в залежності від впливу факторів випадкового характеру. Величина «кількість замовлень» відноситься до величин, які називаються випадковими. Вони дають кількісну оцінку результату випробування на відміну від випадкових подій, розглянутих попередньо, які характеризують результат випробування якісно.

Випадковою називається величина, яка в результаті випробування приймає те чи інше можливе значення, заздалегідь невідоме, яке змінюється від випробування до випробування і залежить від ряду випадкових факторів.

Випадкові величини позначаються великими літерами $X, Y, Z, \ldots$, а їх можливі значення відповідними малими літерами з індексами. Наприклад, випадкова величина X, її можливі значення $x_1, x_2, \ldots, x_n, \ldots$

Застосовується також інше означення випадкової величини.

Випадковою величиною називається функція X, означена на множині наслідків $x_1, x_2, \ldots, x_n, \ldots$ даного випробування.

Наведемо приклади випадкових величин.

1. Кількість електричних ламп, що виходять з ладу в системі освітлення та сигналізації аеропорту на протязі доби, не є сталою і змінюється в залежності від якості ламп, умов експлуатації, рівня напруги в електромережі тощо. Ця випадкова величина має множину можливих значень $\{0,1,2, \ldots, n, \ldots\}$, яка теоретично може бути нескінченною.
2. Рівень напруги в електромережі аеропорту також не є сталою величиною і змінюється в залежності від режиму роботи електростанції, кількості споживачів, системи стабілізації тощо. Ця випадкова величина має множину можливих значень, які суцільно заповнюють деякий інтервал.

Випадкові величини бувають двох видів: дискретні і неперервні.

1. *Дискретні випадкові величини* – величини, які в результаті випробувань приймають окремі, ізольовані можливі значення, множина яких може бути скінченною або нескінченною. Можливі значення дискретної величини зображуються точками числової осі. Прикладами дискретних випадкових величин є кількість літаків в зоні диспетчера по керуванню повітряним рухом, кількість пасажирів на рейсі, кількість квитків, виданих на протязі зміни по запитам пасажирів системою продажу авіаквитків, число вузлів системи, які вийшли з ладу впродовж певного часу тощо.
2. *Неперервні випадкові величини* – величини, які в результаті випробувань приймають можливі значення, які суцільно заповнюють деякий інтервал числової осі, скінченний або нескінченний. Множина можливих значень неперервної випадкової величини нескінченна і незліченна. Прикладами неперервних величин є похибки вимірювань фізичних величин з допомогою

приладів, час безвідмовної роботи окремих вузлів системи і всієї системи в цілому, відхилення геометричних розмірів виготовленої деталі від стандартних тощо.

Закон розподілу випадкової величини. Ряд розподілу

Для задання випадкової величини недостатньо перелічити всі її можливі значення, необхідно також вказати ймовірності, з якими ця величина приймає те чи інше можливе значення (для дискретної випадкової величини), або ймовірності, з якими випадкова величина попадає в деякий інтервал (для неперервної випадкової величини). Такі повні дані про випадкову величину дають так звані закони розподілу випадкової величини.

Законом розподілу випадкової величини називається залежність (таблиця, графік, функція тощо) між її можливими значеннями і відповідними їм імовірностями.

Найпростішою формою закону розподілу дискретної випадкової величини є *ряд розподілу*. Ряд розподілу являє собою таблицю, в першому рядку якої наведені всі можливі значення дискретної випадкової величини, а в другому – ймовірності, з якими випадкова величина приймає ці значення (табл. 6.1):

Таблиця 6.1. *Ряд розподілу дискретної випадкової величини*

x_i	x_1	x_2	$\dots$	x_{n-1}	x_n
p_i	p_1	p_2	$\dots$	p_{n-1}	p_n

Та обставина, що в результаті випробування випадкова величина X приймає певне можливе значення x_i, може розглядатися, як випадкова подія $\{X = x_i\}$. Оскільки в результаті випробування величина X приймає лише можливе значення x_i, то події $\{X = x_i\}$ $(0,1,2,\dots,n)$ утворюють повну групу несумісних подій. Тому справедлива умова (6.1):

$$\sum_{i=1}^{n} p_i = 1. \qquad (6.1)$$

де $p_i = P\{X = x_i\} = P(x_i)$ – ймовірність настання події $\{X = x_i\}$.

Приклад 6.1. *Проводиться випробування надійності системи, яка складається з трьох працюючих незалежно приладів. Надійність* (*ймовірність безвідмовної роботи*) *першого приладу дорівнює* 0,9, *другого* – 0,8, *третього* – 0,7. *Побудувати ряд розподілу випадкової величини* X *– числа надійних приладів в системі.*

Розв'язання. Випадкова величина X приймає можливі значення 0,1,2,3. Позначимо через g_1, g_2, g_3 ймовірності безвідмовної роботи відповідно першого, другого, третього приладів, тоді за умовою задачі $g_1 = 0{,}9$; $g_2 = 0{,}8$; $g_3 = 0{,}7$, отже, ймовірності виходу з ладу приладів відповідно дорівнюють $\overline{g_1} = 0{,}1$; $\overline{g_2} = 0{,}2$; $\overline{g_3} = 0{,}3$. Застосувавши теореми додавання і множення ймовірностей, обчислимо ймовірності того, що випадкова величина X приймає можливі значення 0,1,2,3:

$$p_0 = P\{X = 0\} = \overline{g_1} \cdot \overline{g_2} \cdot \overline{g_3} = 0{,}006;$$
$$p_1 = P\{X = 1\} = g_1 \cdot \overline{g_2} \cdot \overline{g_3} + g_2 \cdot \overline{g_1} \cdot \overline{g_3} + g_3 \cdot \overline{g_1} \cdot \overline{g_2} = 0{,}092;$$
$$p_2 = P\{X = 2\} = g_1 \cdot g_2 \cdot \overline{g_3} + g_1 \cdot g_3 \cdot \overline{g_2} + g_2 \cdot g_3 \cdot \overline{g_1} = 0{,}398;$$
$$p_3 = P\{X = 3\} = g_1 \cdot g_2 \cdot g_3 = 0{,}504.$$

Ряд розподілу випадкової величини X запишеться у вигляді (табл.6.2):

Таблиця 6.2. Ряд розподілу випадкової величини X

x_i	0	1	2	3
p_i	0,006	0,092	0,398	0,504

Для контролю обчислень перевіримо виконання умови (6.1):

$$\sum_{i=1}^{n} p_i = 0{,}006 + 0{,}092 + 0{,}398 + 0{,}504 = 1.$$

Геометричне зображення ряду розподілу називають *багатокутником розподілу*, для побудови якого на осі абсцис відкладають можливі значення випадкової величини X, а на осі ординат – відповідні їм імовірності, після чого одержані точки з'єднують прямолінійними відрізками.

Багатокутник розподілу для ряду, одержаного в прикладі 6.1, наведено на рис. 6.1.

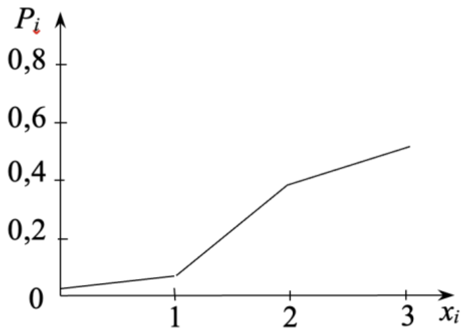

Рис. 6.1. Багатокутник розподілу

Відповідь. Ряд розподілу наведено в табл. 6.2.

6.2. Функція розподілу випадкової величини та її властивості

Ряд розподілу досить повно характеризує випадкову величину, проте побудувати його можна лише для дискретної випадкової величини, оскільки множина можливих значень неперервної випадкової величини – незліченна і, отже, їх не можна перелічити в ряді розподілу.

Загальною формою задання закону розподілу, яка застосовується як для дискретних, так і для неперервних випадкових величин, є функція розподілу, яку інколи також називають інтегральною функцією розподілу.

Функцією розподілу випадкової величини X називається функція $F(x)$, яка для кожного значення x дорівнює ймовірності того, що випадкова величина X прийме значення, менше за x, тобто (6.2):

$$F(x) = P\{X < x\}. \tag{6.2}$$

Геометрично функція розподілу $F(x)$ для кожного фіксованого x подає ймовірність попадання випадкової величини в напівінтервал $(-\infty, x)$, який знаходиться на числовій осі лівіше точки x.

Побудуємо графік функції розподілу $F(x)$ для дискретної випадкової величини X, заданої рядом розподілу (табл. 6.1).

1. Нехай $x \le x_1$. Оскільки випадкова величина X не приймає можливих значень, менших за x, то подія $\{X < x\}$ в цьому випадку неможлива і, отже, її ймовірність дорівнює нулю:

$$F(x) = P\{X < x\} = 0.$$

2. Нехай тепер $x_1 < x \le x_2$. При цьому випадкова величина X приймає єдине можливе значення x_1, менше за x, з ймовірністю p_1. Тому

$$F(x) = P\{X < x\} = P\{X = x_1\} = p_1.$$

3. Нехай далі $x_2 < x \le x_3$. При цьому випадкова величина X може прийняти або значення x_1 з ймовірністю p_1, або значення x_2 з ймовірністю p_2.

Тому, застосовуючи теорему додавання ймовірностей несумісних подій, одержимо:

$$F(x) = P\{X < x\} = P\{X = x_1\} + P\{X = x_2\} = p_1 + p_2.$$

4. Для випадку $x_{n-1} < x \leq x_n$ аналогічно одержимо:

$$F(x) = P\{X < x\} = P\{X = x_1\} + P\{X = x_2\} + \cdots + P\{X = x_{n-1}\} = p_1 + p_2 + \\ + \cdots + p_{n-1}.$$

5. Нехай, нарешті, $x > x_n$. Тоді випадкова величина X приймає одне з усіх можливих значень $x_1, x_2, \ldots, x_n$. Ця подія достовірна і, отже, її ймовірність дорівнює одиниці, тобто $F(x) = 1$.

Таким чином, *функція розподілу $F(x)$ для дискретної випадкової величини X*, заданої рядом розподілу (табл. 6.1), має такий аналітичний вираз (6.3):

$$F(x) = \begin{cases} 0, & \text{при } x \leq x_1; \\ p_1, & \text{при } x_1 < x \leq x_2; \\ p_1 + p_2, & \text{при } x_2 < x \leq x_3; \\ \ldots & \\ p_1 + p_2 + + \cdots + p_{n-1}, & \text{при } x_{n-1} < x \leq x_n; \\ 1, & \text{при } x > x_n. \end{cases} \qquad (6.3)$$

Побудуємо графік функції $F(x)$ (рис. 6.2).

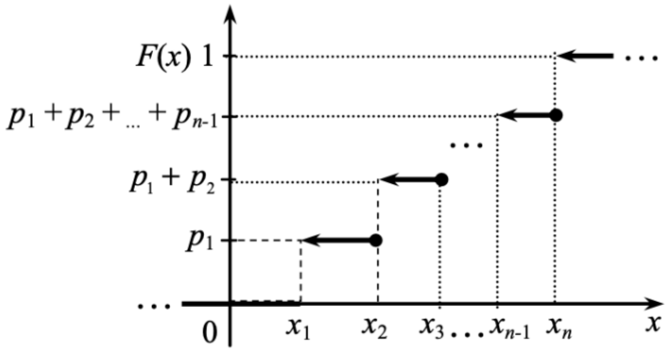

Рис. 6.2. Графік функції розподілу $F(x)$

Як видно з рис. 6.2, графік функції розподілу $F(x)$ дискретної випадкової величини X є розривна східчаста лінія, стала в інтервалах між можливими значеннями випадкової величини, причому розмір стрибка функції $F(x)$ в точках x_i дорівнює ймовірності p_i, з якою випадкова величина приймає відповідне можливе значення x_i.

При збільшенні числа n можливих значень, які приймає випадкова величина X, довжини східців і розміри стрибків в точках розриву зменшуються, і графік функції $F(x)$ наближається до певної плавної неперервної кривої.

У випадку *неперервної випадкової величини*, у якої множина можливих значень на інтервалі $(a, b]$ незліченна, графік функції $F(x)$ є неперервною лінією, яка схематично зображена суцільною кривою на рис. 6.3.

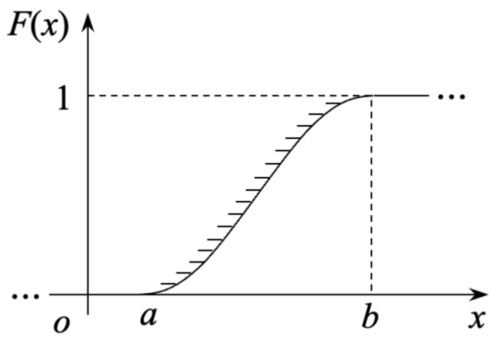

Рис. 6.3. Графік функції розподілу неперервної випадкової величини

Що ж стосується конкретної неперервної випадкової величини X, то її функція розподілу $F(x)$ повинна бути заданою аналітично або графічно. Наприклад, якщо неперервна випадкова величина X задана функцією розподілу $F(x)$:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 0; \\ \dfrac{x \cdot \sqrt{x}}{8}, & \text{при } 0 < x \leq 4 \\ 1, & \text{при } x > 4, \end{cases}$$

то графік функції $F(x)$ має вигляд, представлений на рис. 6.4.

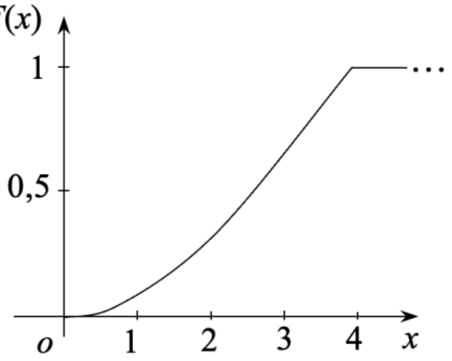

Рис. 6.4. Графік функції розподілу $F(x)$

Властивості функції розподілу випадкової величини

1. *Функція розподілу приймає значення з відрізка* $[0; 1]$, *тобто справедлива нерівність* (6.4):

$$0 \leq F(x) \leq 1. \tag{6.4}$$

2. $F(x)$ *— неспадна функція, тобто при* $x_2 \geq x_1$ *справедлива нерівність* (6.5):

$$F(x_2) \geq F(x_1). \tag{6.5}$$

3. *Ймовірність того, що випадкова величина* X *в результаті випробування прийме значення з інтервалу* $[\alpha; \beta)$, *дорівнює приросту функції розподілу* $F(x)$ *на цьому інтервалі* (6.6):

$$P\{\alpha \leq X < \beta\} = F(\beta) - F(\alpha). \tag{6.6}$$

Подібний канонічний запис інтервалу $\alpha \leq X < \beta$ має істотне значення у випадку дискретної випадкової величини.

У випадку неперервної випадкової величини факт включення чи не включення границь інтервалу не має значення.

З властивості 3 одержуємо такий важливий висновок:

4. *Ймовірність того, що випадкова величина* X *в результаті випробування прийме одне конкретне можливе значення* x_i, *обчислюється за формулою* (6.7):

$$P\{X = x_i\} = F(x_i + 0) - F(x_i). \tag{6.7}$$

Зокрема, якщо в точці x_i *функція* $F(x)$ *неперервна, то виконується рівність* (6.8):

$$P\{X = x_i\} = 0, \tag{6.8}$$

оскільки за означенням неперервної функції в точці x_i: $F(x_i + 0) = F(x_i)$.

Таким чином, не має сенсу розглядати ймовірність того, що неперервна випадкова величина прийме одне конкретне можливе значення, доцільно розглядати ймовірність її подання в деякий інтервал, нехай навіть досить малий.

5. *Якщо випадкова величина X приймає всі можливі значення на інтервалі $[\alpha; \beta)$, то $F(x) = 0$ при $x \le a$ і $F(x) = 1$ при $x > b$.*

З властивості 5 одержуємо такий важливий висновок:

6. *Якщо випадкова величина X приймає можливі значення на всій числовій осі, то справедливі рівності* (6.9):

$$F(-\infty) = 0; \; F(+\infty) = 1. \qquad (6.9)$$

Приклад 6.2. *Випадкова величина X має ряд розподілу* (*табл.* 6.3)

Таблиця 6.3. *Ряд розподілу випадкової величини X*

x_i	1	2	3	4
p_i	0,006	0,092	0,398	0,504

Знайти функцію розподілу $F(x)$ та обчислити ймовірності подій: ***а)*** $\{X < 3\}$; ***б)*** $\{2 \le X \le 4\}$; ***в)*** $\{2 < X < 4\}$; *г)* $\{X = 3\}$; ***д)*** $\{X = 3{,}5\}$.

Розв'язання. Функція розподілу $F(x)$ будується за схемою (6.3).

$$F(x) = \begin{cases} 0, & \text{при } x \le 1; \\ 0{,}006, & \text{при } 1 < x \le 2; \\ 0{,}098, & \text{при } 2 < x \le 3; \\ 0{,}496, & \text{при } 3 < x \le 4; \\ 1, & \text{при } x > 4. \end{cases}$$

а) Ймовірність події $\{X < 3\}$ обчислюється за формулою (6.6):

$$P\{X < 3\} = F(3) - F(-\infty) = 0{,}098 - 0 = 0{,}098.$$

б) Для обчислення ймовірності події $\{2 \le X \le 4\}$ застосовуємо формули (6.6) і (6.7):

$$P\{2 \le X \le 4\} = P\{2 \le X < 4\} + P\{X = 4\} = F(4) - F(2) + F(4+0) - F(4) =$$
$$= F(4+0) - F(2) = 1 - 0{,}006 = 0{,}994.$$

в) $P\{2 < X < 4\} = P\{2 \le X < 4\} - P\{X = 2\} = F(4) - F(2) - F(2+0) + F(2) = F(4) - F(2+0) = 0{,}496 - 0{,}098 = 0{,}398.$

г) $P\{X = 3\} = F\{3+0\} - F\{3\} = 0{,}498 - 0{,}098 = 0{,}398.$

д) $P\{X = 3{,}5\} = 0$, оскільки в точці $x = 3{,}5$ функція розподілу $F(x)$ неперервна.

Відповідь. а) $P\{X < 3\} = 0{,}098$; б) $P\{2 \le X \le 4\} = 0{,}994$; в) $P\{2 < X < 4\} = 0{,}398$; г) $P\{X = 3\} = 0{,}398$; д) $P\{X = 3{,}5\} = 0$.

Приклад 6.3. *Неперервна випадкова величина X задана функцією розподілу*

$$F(x) = \begin{cases} 0, & \text{при } x \le -\dfrac{\pi}{2}; \\ A \cdot \big(1 + sin(x)\big), & \text{при } -\dfrac{\pi}{2} < x \le \dfrac{\pi}{2}; \\ 1, & \text{при } x > \dfrac{\pi}{2}. \end{cases}$$

1. *Знайти коефіцієнт A*; **2.** *Побудувати графік функції $F(x)$*; **3.** *Обчислити ймовірність того, що в результаті випробування X прийме значення з інтервалу* $\left(-\frac{\pi}{4}; \frac{\pi}{4}\right)$.

Розв'язання. 1. Випадкова величина X приймає можливі значення на інтервалі $\left(-\frac{\pi}{2};\frac{\pi}{2}\right)$. Коефіцієнт A знайдемо за властивістю 4, згідно якої $F\left(\frac{\pi}{2}\right)=1$, отже, $A=\frac{1}{2}$.

2. Графік функції $F(x)$ на інтервалі $\left(-\frac{\pi}{2};\frac{\pi}{2}\right)$ одержуємо з графіка функції $y=sin\,(x)$ зсувом на одиницю в додатному напрямі осі ординат і стисканням уздовж цієї осі вдвічі.

Графік функції $F(x)$ поданий на рис. 6.5.

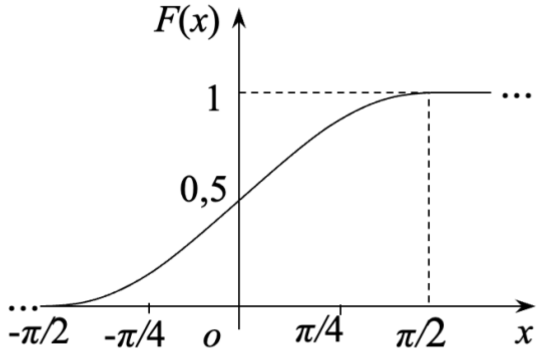

Рис. 6.5. Графік функції розподілу $F(x)$

3. Ймовірність того, що випадкова величина прийме можливе значення з інтервалу$\left(-\frac{\pi}{4};\frac{\pi}{4}\right)$, обчислимо за формулою (6.6):

$$P\left\{-\frac{\pi}{4}\leq X<\frac{\pi}{4}\right\}=F\left(\frac{\pi}{4}\right)-F\left(-\frac{\pi}{4}\right)=\frac{1}{2}\cdot\left(1+\frac{\sqrt{2}}{2}\right)-\frac{1}{2}\cdot\left(1-\frac{\sqrt{2}}{2}\right)=\frac{\sqrt{2}}{2}\approx 0{,}707.$$

Відповідь. 1. $A=\frac{1}{2}$; 2. Рис. 6.5; 3. $P\left\{-\frac{\pi}{4}\leq X<\frac{\pi}{4}\right\}\approx 0{,}707$.

6.3. Щільність розподілу неперервної випадкової величини і її властивості

Щільність розподілу або щільність ймовірності $f(x)$ є однією з форм закону розподілу випадкової величини, який застосовується лише для неперервних величин. Термін «щільність ймовірності» походить із задач механіки: якщо в відношенні (6.10)

$$\frac{P\{x\leq X<x+\Delta x\}}{\Delta x} \tag{6.10}$$

відношення ймовірності $P\{x\leq X<x+\Delta x\}$ того, що в результаті випробування випадкова величина X потрапить в інтервал $[x;x+\Delta x)$, до довжини інтервалу Δx ймовірність $P\{x\leq X<x+\Delta x\}$ інтерпретувати як масу речовини, то дане відношення буде не що інше, як середня лінійна щільність (густина) речовини.

Враховуючи, що за формулою (6.6)

$$P\{x\leq X<x+\Delta x\}=F(x+\Delta x)-F(x),$$

представимо відношення (6.10) у вигляді:

$$\frac{P\{x\leq X<x+\Delta x\}}{\Delta x}=\frac{F(x+\Delta x)-F(x)}{\Delta x}$$

або, переходячи до границі при $\Delta x\to 0$, одержимо в правій частині похідну $F'(x)$ функції розподілу, тобто рівність (6.11):

$$\lim_{\Delta x\to 0}\left(\frac{P\{x\leq X<x+\Delta x\}}{\Delta x}\right)=\lim_{\Delta x\to 0}\left(\frac{F(x+\Delta x)-F(x)}{\Delta x}\right)=F'(x). \tag{6.11}$$

В лівій частині формули (6.11) знаходиться щільність ймовірності в точці x, яка позначається $f(x)$, тому отримуємо рівність (6.12):

$$f(x) = \lim_{\Delta x \to 0} \left(\frac{P\{x \leq X < x + \Delta x\}}{\Delta x} \right) = F'(x). \quad (6.12)$$

Щільністю розподілу або *щільністю ймовірності* $f(x)$ неперервної випадкової величини X називається перша похідна від її функції розподілу $F(x)$. Функцію $f(x)$ інколи називають також *диференціальною функцією* розподілу випадкової величини.

Властивості щільності розподілу неперервної випадкової величини $f(x)$

1. *Щільність ймовірності – невід'ємна функція:* $f(x) \geq 0$.
2. *Ймовірність того, що в результаті випробування неперервна випадкова величина прийме можливе значення з інтервалу* $[\alpha; \beta)$, *дорівнює визначеному інтегралу від щільності ймовірності на цьому інтервалі* (6.13):

$$P\{\alpha \leq X < \beta\} = \int_{\alpha}^{\beta} f(x)dx. \quad (6.13)$$

Оскільки випадкова величина є неперервною факт включення чи не включення границь інтервалу не має значення.

Якщо крива щільності розподілу $f(x)$ має схематичний вигляд, представлений на рис. 6.6, то геометрично ймовірність $P\{\alpha \leq X < \beta\}$ дорівнює площі криволінійної трапеції, що спирається на відрізок $[\alpha; \beta)$ і обмежена зверху кривою $f(x)$.

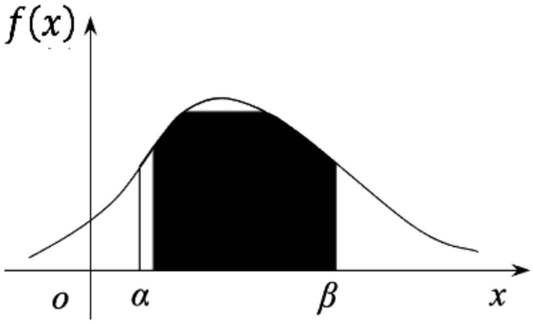

Рис.6.6. Графік функції щільності розподілу $f(x)$

3. *Функція розподілу* $F(x)$ *виражається через щільність розподілу* $f(x)$ *формулою* (6.14):

$$F(x) = \int_{-\infty}^{x} f(x)dx. \quad (6.14)$$

Геометрично значення функції розподілу $F(x)$ в точці x дорівнює площі, яка розташована лівіше точки x і обмежена кривою щільності розподілу і віссю OX (рис. 6.7).

Рис. 6.7. Геометрична інтерпретація властивості 3

4. *Якщо неперервна випадкова величина* X *приймає можливі значення на всій числовій осі, то виконується умова* (6.15):

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \qquad (6.15)$$

Якщо випадкова величина приймає всі можливі значення на обмеженому інтервалі$(a; b]$, *то умова* (6.15) *набуває вигляду* (6.16):

$$\int_a^b f(x)dx = 1. \qquad (6.16)$$

Із щільністю ймовірності зв'язане важливе поняття елемента ймовірності, яке використовується для знаходження числових характеристик неперервної випадкової величини.

Із формули (6.16) з точністю до нескінченно малих вищих порядків випливає рівність:

$$P\{x \le X < x + \Delta x\} = f(x) \cdot \Delta x = f(x)dx. \qquad (6.17)$$

Величина $f(x)dx$ (або $f(x) \cdot \Delta x$) називається *елементом ймовірності* для точки x неперервної випадкової величини, оскільки вона дорівнює ймовірності того, що випадкова величина в результаті випробування потрапить в досить малий інтервал Δx.

Геометрично елемент ймовірності наближено дорівнює площі елементарної криволінійної трапеції з основою dx, прилеглою до точки x (рис. 6.8).

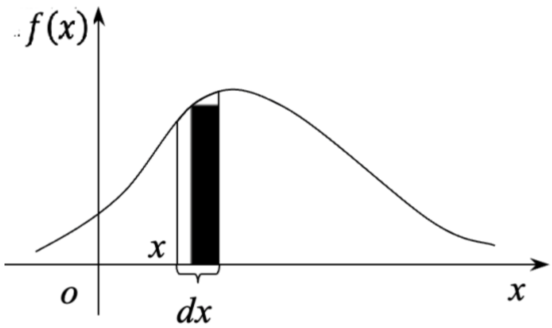

Рис. 6.8. Геометрична інтерпретація властивості 4

Приклад 6.4. *Неперервна випадкова величина має щільність ймовірності*:

$$f(x) = \begin{cases} 0, & \text{при } x \le 1; \\ A \cdot ln(x), & \text{при } 1 < x \le e^2; \\ 0, & \text{при } x > e^2. \end{cases}$$

Знайти: **1.** *Сталий параметр* A; **2.** *Функцію розподілу* $F(x)$ (*побудувати її графік*); **3.** *Ймовірність того, що в результаті випробування випадкова величина потрапить в інтервал* $(e; e^2)$.

Розв'язання. 1. Параметр A знайдемо з умови (6.16):

$$\int_a^b f(x)dx = A \cdot \int_1^{e^2} ln(x)dx = \left| \begin{matrix} u = ln(x); dv = dx \\ du = \frac{1}{x}dx; v = x \end{matrix} \right| =$$

$$= A \cdot \left(\left(x \cdot ln(x)\right) \Big|_1^{e^2} - x \Big|_1^{e^2} \right) = A \cdot (e^2 + 1) = 1.$$

Звідки

$$A = \frac{1}{e^2+1}, f(x) = \frac{ln(x)}{e^2+1}, \text{ при } 1 < x \le e^2, f(x) = 0 \text{ поза цим інтервалом.}$$

2. Функція розподілу знаходиться за формулою (6.14):

$$F(x) = \int_{-\infty}^{x} f(x)dx = \frac{1}{e^2+1} \cdot \int_{1}^{x} ln(x)dx = \frac{1}{e^2+1} \cdot (x \cdot ln(x) - x)\Big|_1^x =$$
$$= \frac{1}{e^2+1} \cdot (x \cdot ln(x) - x + 1).$$

Отже, функція розподілу $F(x)$ має вигляд:

$$F(x) = \begin{cases} 0, & \text{при } x \le 1; \\ \dfrac{x \cdot ln(x) - x + 1}{e^2+1}, & \text{При } 1 < x \le e^2; \\ 1, & \text{при } x > e^2. \end{cases}$$

Графік $F(x)$ представлений на рис. 6.9.

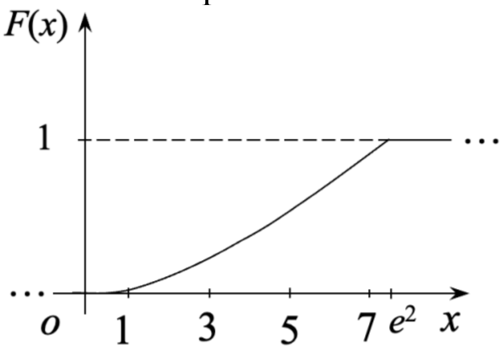

Рис. 6.9. Графік функції розподілу $F(x)$

3. Ймовірність події $\{e < X < e^2\}$ обчислимо за формулою (6.13):

$$P\{e < X < e^2\} = \int_{e}^{e^2} f(x)dx = \frac{1}{e^2+1} \cdot \int_{e}^{e^2} ln(x)dx = \frac{1}{e^2+1} \cdot (x \cdot ln(x) - x)\Big|_e^{e^2}$$
$$= \frac{e^2}{e^2+1} \approx 0{,}881.$$

Відповідь. 1. $A = \frac{1}{e^2+1}$; 2. $F(x) = \begin{cases} 0, \text{ при } x \le 1; \\ \frac{x \cdot ln(x) - x + 1}{e^2+1}, \text{ при } 1 < x \le e^2; \\ 1, \text{при } x > e^2, \end{cases}$ 3. $P\{e < X < e^2\} \approx 0{,}881$.

6.4. Мода та медіана випадкових величин

Ймовірнісне трактування

Модою $Mo(X)$ випадкової величини X називається її найбільш ймовірне можливе значення.

Для дискретної випадкової величини X, яка має ряд розподілу (табл. 6.1), $Mo(X)$ дорівнює тому можливому значенню, якого величина набуває з найбільшою ймовірністю. Так, для випадкової величини X, ряд розподілу якої побудовано у прикладі 6.1, мода $Mo(X) = 3$.

Для неперервної випадкової величини X, заданої щільністю ймовірності $f(x)$ на проміжку $(a; b)$, мода дорівнює тому значенню x, при якому $f(x)$ досягає максимуму або (у разі відсутності максимуму) набуває найбільшого значення на відрізку.

Якщо $f(x)$ набуває найбільшого значення в одній точці x з проміжку $(a; b)$, то розподіл випадкової величини X називається *унімодальним*, а якщо в кількох точках — *полімодальним*.

Медіаною $Me(X)$ випадкової величини X називається те значення x, для якого виконується рівність (6.18):

$$P(X < Me(X)) = P(X > Me(X)) = \frac{1}{2}. \qquad (6.18)$$

Враховуючи формулу (6.6), з формули (6.18) отримуємо:

$$P(X < Me(X)) = F(Me(X)) - F(-\infty) = F(Me(X)) = \frac{1}{2}.$$

Таким чином, якщо випадкова величина задана своєю функцією розподілу, то для обчислення медіани варто використовувати формулу (6.19):

$$F(Me(X)) = \frac{1}{2}. \qquad (6.19)$$

Медіана, як правило, застосовується для неперервної випадкової величини і геометрично подає абсцису точки, в якій пряма $x = Me(X)$ поділяє навпіл площу, обмежену кривою розподілу $f(x)$ і віссю OX.

У випадку дискретної випадкової величини медіана обчислюється у тих випадках, коли її функція розподілу на певному інтервалі $[a; b)$ (*медіанному відрізку*) дорівнює $\frac{1}{2}$, згідно формули (6.19). Тоді справедлива рівність (6.20):

$$Me(X) = \frac{a+b}{2}. \qquad (6.20)$$

Рис. 6.10. Геометрична візуалізація моди та медіани функції густини ймовірностей

Приклад 6.5. *Неперервна випадкова величина X має щільність ймовірності*

$$f(x) = \begin{cases} -6 \cdot x^2 + 18 \cdot x - 12, & \text{при } x \in [1; 2); \\ 0, & \text{при } x \notin [1; 2). \end{cases}$$

Знайти моду та медіану цієї величини.

Розв'язання. Для визначення моди перевіримо функцію $f(x)$ на наявність екстремумів, для чого визначимо, при якому значенні x її похідна дорівнює нулю:

$$f'(x) = -12 \cdot x + 18 = 0; \; x = 1{,}5.$$

При переході через точку $x = 1{,}5$ похідна змінює знак з «+» на «–», тому в цій точці $f(x)$ має максимум, а отже, мода $Mo(X) = 1{,}5$.

Оскільки крива розподілу $f(x)$ симетрична відносно прямої $x = 1{,}5$ (рис. 6.10), то ця пряма поділяє навпіл площу, обмежену кривою $f(x)$ і віссю OX, а отже, $Me(X) = 1{,}5$.

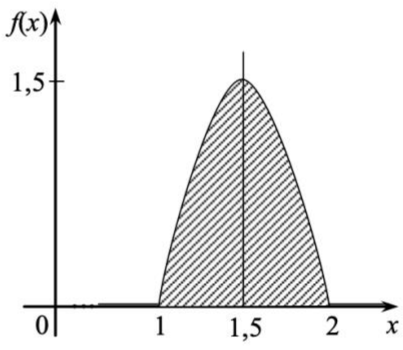

Рис.6.11. Графік функції щільності розподілу $f(x)$

Взагалі, якщо крива розподілу симетрична відносно прямої $x = a$ і має в точці a максимум, то $Mo(X) = Me(X) = a$.

Відповідь. $Mo(X) = Me(X) = 1{,}5$.

Приклад 6.6. *Неперервна випадкова величина X має щільність ймовірності*

$$f(x) = \begin{cases} A \cdot x^2, \text{при } x \in [0;2); \\ 0, \quad \text{при } x \notin [0;2). \end{cases}$$

Знайти: **1.** *Коефіцієнт A*; **2.** *Моду і медіану величини X.*

Розв'язання. 1. Коефіцієнт A знайдемо за властивістю 4 щільності ймовірності:

$$\int_a^b f(x)dx = A \cdot \int_0^2 x^2\, dx = \frac{8 \cdot A}{3} = 1;\ A = \frac{3}{8}.$$

Тому щільність ймовірності має вигляд

$$f(x) = \frac{3}{8} \cdot x^2, \text{при } x \in [0;2).$$

Графік кривої розподілу зображено на рис. 6.11.

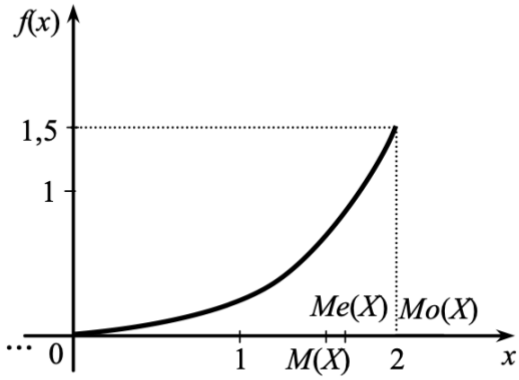

Рис.6.12. Графік функції щільності розподілу $f(x)$

2. На інтервалі $(0;2)$ крива розподілу не має максимуму, тому мода $Mo(X) = 2$, оскільки при $x = 2$ функція $f(x)$ набуває найбільшого значення.

Медіану знайдемо за означенням (формула (6.18)):

$$P\big(X < Me(X)\big) = \frac{3}{8} \cdot \int_0^{Me(X)} x^2\, dx = \frac{x^3}{8}\Big|_0^{Me(X)} = \frac{\big(Me(X)\big)^3}{8} = \frac{1}{2},$$

звідки $Me(X) = \sqrt[3]{4} \approx 1{,}59$.

Відповідь. 1. $A = \frac{3}{8}$; 2. $Mo(X) = 2$; $Me(X) \approx 1{,}59$.

Приклад 6.7. *Знайти моду і медіану дискретної випадкової величини X (табл.* 6.4).

Таблиця 6.4. *Ряд розподілу випадкової величини X*

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

Розв'язання. Складемо функцію розподілу випадкової величини X:

$$F(x) = \begin{cases} 0, \quad \text{при } x \le 1; \\ 0{,}1, \text{при } 1 < x \le 2; \\ 0{,}5, \text{при } 2 < x \le 3; \\ 0{,}8, \text{при } 3 < x \le 4; \\ 1, \quad \text{при } x > 4. \end{cases}$$

$Mo(X) = 2$, оскільки 2-ці відповідає найбільше значення ймовірності.

$Me(X) = 2{,}5$ – за формулою (6.19), оскільки $F(x) = 0{,}5$, при $2 < x \le 3$.

Відповідь. $Mo(X) = 2$; $Me(X) = 2{,}5$.

Приклад 6.8. *Знайти моду і медіану дискретної випадкової величини X (табл.* 6.5).

Таблиця 6.5. *Ряд розподілу випадкової величини X*

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

Розв'язання. Складемо функцію розподілу випадкової величини X:

$$F(x) = \begin{cases} 0, & \text{при } x \leq 1; \\ 0{,}2, & \text{при } 1 < x \leq 2; \\ 0{,}6, & \text{при } 2 < x \leq 3; \\ 0{,}7, & \text{при } 3 < x \leq 4; \\ 1, & \text{при } x > 4. \end{cases}$$

$Mo(X) = 2$, оскільки 2-ці відповідає найбільше значення ймовірності.

$Me(X)$ немає – за формулою (6.19), оскільки $F(x) \neq 0{,}5$ на кожному з інтервалів.

Відповідь. $Mo(X) = 2$; $Me(X)$ немає.

Приклад 6.9. *Знайти моду і медіану неперервної випадкової величини, заданої функцією розподілу*

$$F(x) = \begin{cases} 0, & \text{при } x \leq -1; \\ \dfrac{x+1}{2}, & \text{при } -1 < x \leq 1; \\ 1, & \text{при } x > 1. \end{cases}$$

Розв'язання. Щільність розподілу отримуємо з формули (6.12).

$$f(x) = \begin{cases} 0, \text{ при } x \leq -1; \\ \dfrac{1}{2}, \text{ при } -1 < x \leq 1; \\ 0, \text{ при } x > 1\,. \end{cases}$$

Функція щільності набуває максимального значення на всьому інтервалі $(-1; 1]$. Таким чином, в даному розподілі модами є всі значення $-1 < x \leq 1$. Тобто він є полімодальним.

З формули (6.19)

$$\frac{Me(X) + 1}{2} = \frac{1}{2},$$

тобто $Me(X) = 0$.

Відповідь. $Mo(X)$ – весь інтервал $(-1; 1]$; $Me(X) = 0$.

Статистичне трактування

Для визначення моди та медіани у математичній статистиці введено декілька означень.

Статистична вибірка — це підмножина об'єктів, елементів або спостережень, відібрана з більшої сукупності (генеральної сукупності) для аналізу і вивчення. Вибірка використовується для того, щоб зробити висновки про всю генеральну сукупність на основі меншої кількості даних.

Варіаційний ряд — це впорядкований за зростанням або спаданням набір значень варіанти (спостережень) випадкової величини або статистичної вибірки (табл. 6.6). Він відображає розподіл частот значень досліджуваної величини, дозволяючи оцінювати розподіл даних та проводити різноманітні статистичні аналізи.

Таблиця 6.6. *Варіаційний ряд*

x_i	x_1	x_2	…	x_k
n_i	n_1	p_2	…	n_k

x_i – елементи вибірки ($i = \overline{1, k}$);

n_i – частоти елементів вибірки;

$n = \sum_{i=1}^{k} n_i$ – об'єм вибірки.

Інтервальний варіаційний ряд — це спосіб подання статистичних даних, де всі спостереження розподіляються за інтервалами (табл. 6.7). Кожен інтервал включає певний діапазон значень змінної, а також частоти або частки, які показують, скільки спостережень належать до кожного інтервалу. Інтервальний варіаційний ряд використовується для групування даних і надання загальної картини їх розподілу.

Таблиця 6.7. *Інтервальний варіаційний ряд*

$[x_i; x_{i+1})$	$[x_1; x_2)$	$[x_2; x_3)$	…	$[x_m; x_{m+1})$
n_i	n_1	n_2	…	n_m

$[x_i; x_{i+1})$ – інтервали $(i = \overline{1, m})$;
n_i – частоти інтервалів вибірки;
$n = \sum_{i=1}^{m} n_i$ – об'єм вибірки.

В статистиці *медіана* — це величина ознаки, що розташована посередині ранжованого ряду вибірки, тобто — це величина, що розташована в середині ряду величин, розташованих у зростальному або спадному порядку; в теорії ймовірності — характеристика розподілення випадкової величини.

Знаходження медіани:
1. Необхідно ранжувати статистичну вибірку;
2. Якщо n – непарне число, то $Me = x_{\frac{1+k}{2}}$ – середина вибірки;
3. Якщо n – парне число, то $Me = \frac{x_{\frac{k}{2}} + x_{\frac{2+k}{2}}}{2}$ – середнє арифметичне двох середніх елементів вибірки.

Модою називають значення ознаки, що має найбільшу частоту в статистичному ряду розподілу.

Спосіб обчислення моди залежить від того, в якому вигляді дано значення ознаки: дискретного чи інтервального ряду розподілу. В дискретних варіаційних рядах моду обчислюють без додаткових розрахунків за значенням варіанти з найбільшою частотою.

Тип	Опис	Приклад	Результат
Середнє арифметичне	Сума всіх значень вибірки поділена на кількість цих елементів вибірки: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	(1+2+2+3+4+7+9) / 7	**4**
Медіана	Середнє значення, що відокремлює більшу половину і меншу половину вибірки	1, 2, 2, **3**, 4, 7, 9	**3**
Мода	Значення, що зустрічається у вибірці найчастіше	1, **2**, **2**, 3, 4, 7, 9	**2**

Рис. 6.13. Порівняння різних загальних середніх значень для вибірки {1,2,2,3,4,7,9}

При розрахунку моди в інтервальному варіаційному ряду розподілу спочатку потрібно визначити модальний інтервал, в межах якого знаходяться мода, а потім значення модальної величини ознаки.

В інтервальному варіаційному ряду розподілу модою наближенно вважають центральний вараінт так званого модального інтервалу, тобто того інтервалу, який має найбільшу частоту. В межах інтервалу необхідно знайти те значення ознаки, яке є модою.

В інтервальних варіаційних рядах розподілів моду визначають за формулою (6.21):

$$Mo = x_{Mo} + i_{Mo} \cdot \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}, \quad (6.21)$$

де Mo – мода;
x_{Mo} – нижня межа модального інтервалу;

i_{Mo} – довжина модального інтервалу;
f_{Mo} – частота модального інтервалу;
f_{Mo-1} – частота інтервалу, що передує модальному;
f_{Mo+1} – частота інтервалу, що є наступним за модальним.

Формула ґрунтується на припущенні, що відстані від нижньої межі модального інтервалу до моди і від моди до верхньої межі модального інтервалу прямо пропорційні різницям між чисельностями (частотами) модального інтервалу і інтервалів, що прилягають до нього.

В інтервальних варіаційних рядах розподілів медіану визначають за формулою. (6.22):

$$Me = x_{Me} + i_{Me} \cdot \frac{\frac{n}{2} - n^{\text{н}}_{Me-1}}{n_{Me}}, \qquad (6.22)$$

де Me – медіана;
x_{Me} – нижня межа медіанного інтервалу;
i_{Me} – довжина медіанного інтервалу;
n – об'єм вибірки;
$n^{\text{н}}_{Me-1}$ – накопичувана частота інтервалу, що передує медіанному;
n_{Me} – частота медіанного інтервалу.

Приклад 6.10. *Знайти моду і медіану вибірки, що задана інтервальним рядом* (*табл.* 6.8).

Таблиця 6.8. *Інтервальний варіаційний ряд*

Інтервал добової потужності $[x_i; x_{i+1})$	Кількість заводів (частота) n_i	Кумулятивна (накопичувана) частота $n_i^{\text{н}}$
[20,25)	7	7
[25,35)	14	21
[35,40)	6	27
[40,45)	5	32
[45,50)	8	40
Разом n	40	–

Розв'язання. Обчислимо моду.
[25,35) – модальний інтервал, оскільки має найбільшу частоту.

$$Mo = 25 + 10 \cdot \frac{14-7}{(14-7)+(14-6)} \approx 30.$$

Обчислимо медіану.
[25,35) – медіанний інтервал, оскільки на цей інтервал припадає середина вибірки.

$$Me = 25 + 10 \cdot \frac{\frac{40}{2} - 7}{14} \approx 34.$$