§ 2.2.Числові характеристики випадкових величин

В деяких задачах можна обмежитися розглядом більш простих, але менш повних характеристик випадкової величини.

Математичним сподіванням функції $\varphi(X)$ випадкової величини X з щільністю ймовірності $p(x)$ називають інтеграл:

$$M\big(\varphi(X)\big) = \int_{-\infty}^{+\infty} \varphi(x)\, p(x)\, dx .$$

Для дискретної випадкової величини дана формула набуває вигляду

$$M\big(\varphi(X)\big) = \sum_i \varphi(x_i) \cdot p_i .$$

Зокрема, при $\varphi(X) = X$ маємо

$$m_X = M\big(X\big) = \int_{-\infty}^{+\infty} x \cdot p(x)\, dx . \tag{2.1}$$

Величину m_X називають математичним сподіванням випадкової величини X .

Для дискретної випадкової величини маємо

$$M\big(X\big) = \sum_i x_i \cdot p_i . \tag{2.2}$$

По суті операція математичного сподівання означає середньозважене даної величини з відповідною щільністю ймовірності. Математичне сподівання є однією з характеристик центру розподілу випадкової величини.

Властивості математичного сподівання $M(X)$:

1. Математичне сподівання сталої величини дорівнює цій сталій
$$M(C) = C .$$
2. Сталий множник можна виносити за знак математичного сподівання
$$M(CX) = CM(X) .$$
3. Математичне сподівання алгебраїчної суми двох випадкових величин дорівнює алгебраїчній сумі їх математичних сподівань
$$M(X \pm Y) = M(X) \pm M(Y) .$$
4. Математичне сподівання добутку двох незалежних випадкових величин дорівнює добутку їх математичних сподівань
$$M(XY) = M(X) \cdot M(Y) .$$

Зауваження. Випадкові величини X і Y називаються незалежними, якщо розподіл однієї з них не залежить від значення, якого набуде інша випадкова величина.

Характеристиками центру розподілу випадкової величини є також мода і медіана.

Модою $M_0(X)$ випадкової величини X називається точка локального максимуму її щільності ймовірності:

$$M_0(X) = x_0,\;\; p(x_0) = \max p(x). \tag{2.3}$$

Якщо щільність ймовірностей має єдиний максимум, то випадкова величина називається унімодальною. Якщо максимумів два або більше, то – бімодальною або мультимодальною відповідно.

Медіаною $M_e(X)$ випадкової величини X називається точка x_e, яка ділить площу під графіком функції щільності навпіл, тобто медіана є будь-який корінь рівняння $F(x_e) = 0{,}5$. Тобто

$$M_e(X) = x_e,\;\; F(x_e) = 0{,}5. \tag{2.4}$$

Розглядають поняття квантілі. Квантіллю порядку p є корінь рівняння $F(x_p) = p$, де p – деяке задане число ($0 < p < 1$). З цієї точки зору медіана є квантіль порядку 1/2.

Типові співвідношення між показниками центру розподілу показано на рис. 1.

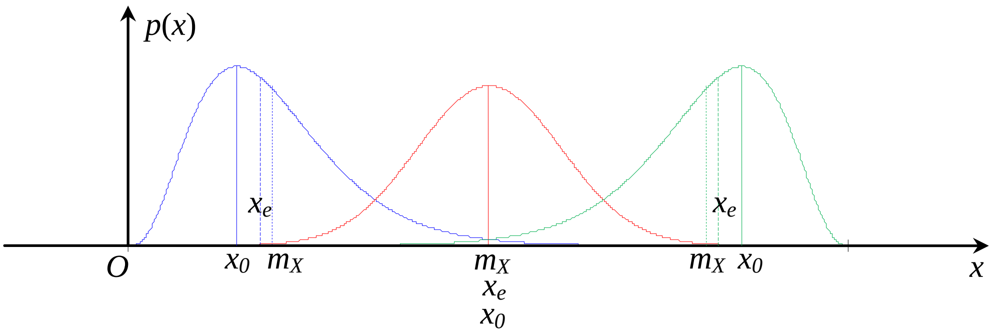


Рис. 1

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

$$D(X) = M\left((X - m_X)^2\right) = \int_{-\infty}^{+\infty} (x - m_X)^2 \cdot p(x)\, dx . \tag{2.5}$$

Для дискретної випадкової величини маємо

$$D(X) = \sum_i (x_i - m_X)^2 \cdot p_i . \tag{2.6}$$

Дисперсія характеризує ступінь розсіювання значень випадкової величини відносно її математичного сподівання.

В практичних розрахунках ефективніше використовувати формулу розрахунку дисперсії, яка подана в наступній теоремі.

Теорема 1. ○

$$D(X) = M\left(X^2\right) - m_x^{\;2} . \bullet \tag{2.7}$$

Для неперервних випадкових величин формула (2.7) запишеться у вигляді

$$D(X) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)\, dx - m_X^{\;2} , \tag{2.8}$$

а для дискретної випадкової величини

$$D(X) = \sum_i x_i^2 \cdot p_i - m_X^{\;2} . \tag{2.9}$$

Властивості дисперсії $D(X)$:

1. Дисперсія будь-якої випадкової величини невід'ємна.
2. Дисперсія сталої дорівнює нулю: $D(C) = 0$.
3. Сталий множник можна винести за знак дисперсії, піднісши його до квадрата:
$$D(CX) = C^2 \cdot D(X) .$$
4. Дисперсія алгебраїчної суми двох незалежних випадкових величин дорівнює сумі їх дисперсій:
$$D(X \pm Y) = D(X) + D(Y) .$$
5. Дисперсія добутку двох незалежних випадкових величин обчислюється за формулою:
$$D(XY) = D(X) \cdot D(Y) + M^2(X) \cdot D(Y) + M^2(Y) \cdot D(X).$$

Одиниці вимірювання дисперсії є квадратними по відношенню до одиниць вимірювання самої випадкової величини. Цього недоліку позбавлений наступний показник розсіювання.

Середнім квадратичним відхиленням випадкової величини X називається квадратний корінь із дисперсії

$$\sigma_X = \sqrt{D(X)} . \tag{2.10}$$

Залежність форми розподілу від середньоквадратичного відхилення σ показано на рис. 2.

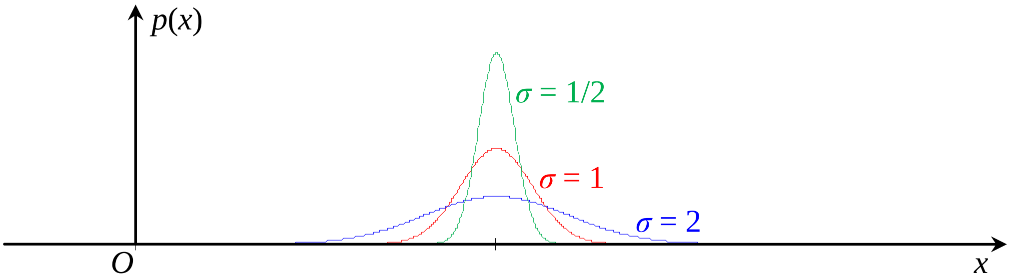


Рис. 2

Наприклад:

П49. Знайдемо математичне сподівання і дисперсію випадкової величини X з таким рядом розподілу

X	1	2	3	4
p_i	$\frac{1}{30}$	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{1}{6}$

Розв'язання. За формулою (2.2) знайдемо математичне сподівання

$$m_x = 1\cdot\frac{1}{30} + 2\cdot\frac{3}{10} + 3\cdot\frac{1}{2} + 4\cdot\frac{1}{6} = \frac{84}{30} = 2{,}8.$$

Мода даної випадкової величини дорівнює 3, $M_0(X)=3,$ тому що значення 3 має найбільшу ймовірність $P(X=3)=0{,}5$.

Дисперсію обчислюємо за формулою (2.9)

$$D(X) = 1^2\cdot\frac{1}{30} + 2^2\cdot\frac{3}{10} + 3^2\cdot\frac{1}{2} + 4^2\cdot\frac{1}{6} - 2{,}8^2 = 8{,}4 - 7{,}84 = 0{,}56\,.$$

Тоді, згідно із формулою (2.8), середнє квадратичне відхилення дорівнює

$$\sigma_x = \sqrt{0{,}56} = 0{,}748\,. \ \square$$

Узагальненням поняття математичного сподівання та дисперсії є початкові та центральні моменти випадкової величини.

Початковим моментом k*-го* порядку випадкової величини X називають математичне сподівання k-го степеня цієї величини:

$$m_k = M\left(X^k\right). \qquad (2.11)$$

В цих термінах математичне сподівання випадкової величини є початковим моментом 1-го порядку цієї величини (при $k=1 \quad m_1 = M(X)$).

Центральним моментом k-го порядку випадкової величини X називають математичне сподівання k-го степеня відхилення $X - M(X)$:

$$\mu_k = M\left[(X - m_x)^k\right]. \qquad (2.12)$$

Користуючись формулою бінома Ньютона та властивостями математичного сподівання можна виразити центральні моменти через початкові і навпаки. Зокрема

$\mu_1 = 0$, $\mu_2 = m_2 - m_1^2 = D(X)$, $\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$, $\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$, …

Отже дисперсія є центральним моментом 2-го порядку даної випадкової величини.

$m_1 = m_X$, $m_2 = D(X) + m_X^2$, $m_3 = 3m_X D(X) + m_X^3$, $m_4 = 3D^2(X) + 6m_X^2 D(X) + m_X^4$, …

Початкові моменти є коефіцієнтами розкладу в ряд Маклорена характеристичної функції випадкової величини:

$$\Phi(\omega) = 1 + \sum_{k=1}^{\infty} \frac{m_k}{k!}(i\omega)^k\,,$$

де $m_k = -i^k \left.\frac{d^k\Phi(\omega)}{d\omega^k}\right|_{\omega=0}$.

Цей розклад дозволяє відновити розподіл випадкової величини за відомою послідовністю моментів. Часто більш результативним є розклад в ряд Маклорена не самої характеристичної функції, а її логарифма:

$$\ln \Phi(\omega) = \sum_{k=1}^{\infty} \frac{\kappa_k}{k!} (i\omega)^k ,$$

де $\kappa_k = -i^k \left. \frac{d^k \ln \Phi(\omega)}{d\omega^k} \right|_{\omega=0}$.

Коефіцієнти κ_k називають <u>кумулянтами</u> k-го порядку.

Кумулянт k-го порядку є многочленом від моментів m_1, m_2, …, m_k і навпаки.

$\kappa_1 = m_1 = m_X$,

$\kappa_2 = m_2 - m_1^2 = \mu_2 = D(X)$,

$\kappa_3 = m_3 - 3m_1m_2 + 2m_1^3 = \mu_3$,

$\kappa_4 = m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m_1^4 = \mu_4 - 3\mu_2^2$, …

Починаючи з 4-го порядку кумулянти відрізняються від центральних моментів.

Безрозмірні відношення

$$\gamma_k = \frac{\kappa_k}{\sigma^k} \tag{2.13}$$

називають <u>кумулянтними коефіцієнтами</u>.

Коефіцієнт $\gamma_3 = \frac{\mu_3}{\sigma^3}$ називають <u>коефіцієнтом асиметрії</u>, а $\gamma_4 = \frac{\mu_4}{\sigma^4} - 3$ – <u>коефіцієнтом ексцесу</u> випадкової величини.

Вплив параметрів γ_3 і γ_4 на форму розподілу проілюстровано на рис. 3 і 4.

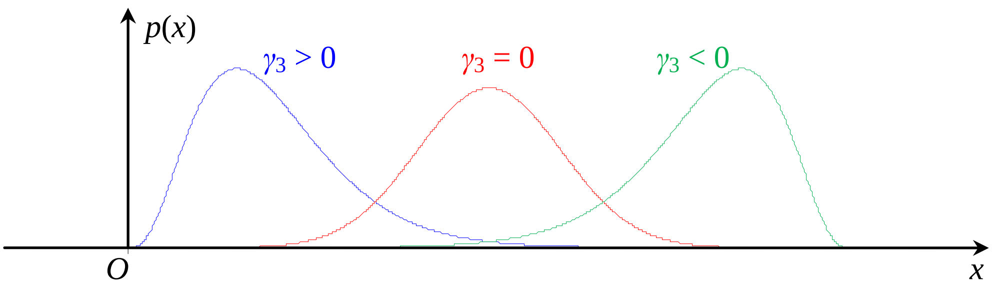


Рис. 3

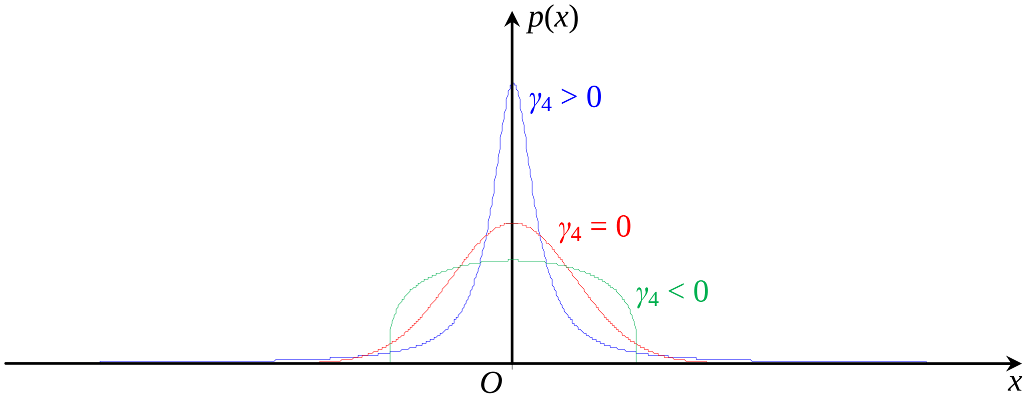


Рис. 4

Зауважимо, що для будь-якого розподілу $\gamma_4 \geq -2$. Причому $\gamma_4 = -2$ для розподілу бінарної альтернативи (див. П47).

Наприклад:

П50. Випадкова величина X задана функцією розподілу

$$F(x)=\begin{cases}0, & x\le 0;\\ \dfrac{1}{2}-\dfrac{1}{2}\cos x, & 0<x\le\pi;\\ 1, & x>\pi.\end{cases}$$

Знайдемо $M(X), D(X), \sigma(X), M_0, M_e$.

Спочатку знайдемо щільність розподілу:

$$p(x)=F'(x)=\begin{cases}0, & x\le 0;\\ \dfrac{1}{2}\sin x, & 0<x\le\pi;\\ 0, & x>\pi.\end{cases}$$

Використовуючи формули 2.1, 2.8, 2.10 знайдемо:

$$M(X)=\frac{1}{2}\int_0^{\pi}x\sin x\,dx=\begin{bmatrix}u=x,\ du=dx\\ dv=\sin xdx,\ v=-\cos x\end{bmatrix}=-\frac{1}{2}x\cos x\Big|_0^{\pi}+\frac{1}{2}\int_0^{\pi}\cos xdx=$$

$$=\frac{1}{2}\pi+\frac{1}{2}\sin x\Big|_0^{\pi}=\frac{\pi}{2};$$

$$D(X)=\frac{1}{2}\int_0^{\pi}x^2\sin x\,dx-\frac{\pi^2}{4}=\frac{\pi^2-4}{4};$$

$$\sigma(X)=\frac{\sqrt{\pi^2-4}}{2}.$$

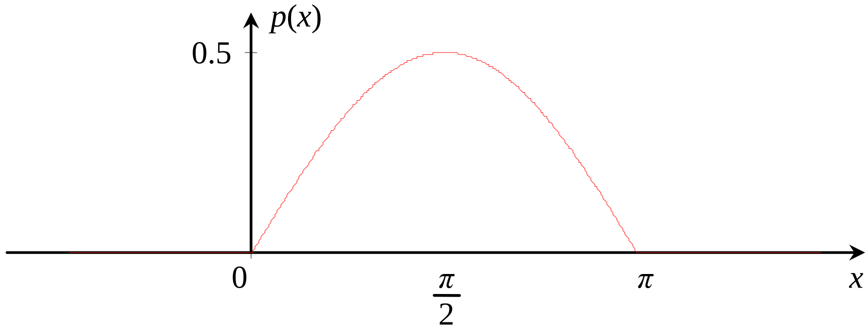
Оскільки $p\left(\dfrac{\pi}{2}\right)=0{,}5$ є максимальним значенням щільності $p(x)$, то $M_0=\dfrac{\pi}{2}$.

Для знаходження медіани скористаємося формулою (2.4):

$$F(Me)=\frac{1}{2}-\frac{1}{2}\cos(M_e)=0{,}5\Rightarrow\cos(M_e)=0\Rightarrow M_e=\frac{\pi}{2}.$$

Таким чином, в даному прикладі значення математичного сподівання, моди і медіани співпадають. Це обумовлено тим, що розподіл випадкової величини X симетричний (графік щільності $p(x)$ симетричний відносно прямої $x=\dfrac{\pi}{2}$).

Побудуйте графік щільності розподілу.



П51. Обчислимо асиметрію і ексцес випадкової величини X за заданою щільністю розподілу. Побудуємо графік щільності розподілу.

$$p(x)=\begin{cases}0, & x\le 0;\\ \dfrac{3}{4}x(2-x), & 0<x\le 2;\\ 0, & x>2.\end{cases}$$

Розв’язання.

$$M(X)=\int_0^2 xp(x)dx=\frac{3}{4}\int_0^3 x^2(2-x)dx=1\ .$$

Оскільки $p(x)=\frac{3}{4}x(2-x)=\frac{3}{4}\left(1-(x-1)^2\right)$ симетрична функція відносно математичного сподівання $M(X)=1$, тоді асиметрія $\gamma_3=0$.

Для обчислення ексцеса γ_4 необхідно знайти μ_4 і σ:

$$\mu_4=\int_0^2 (x-1)^4\frac{3}{4}\left(2x-x^2\right)dx=\frac{3}{35},$$

$$D(X)=\int_0^2 x^2\frac{3}{4}\left(2x-x^2\right)dx-M^2(X)=\frac{1}{5},$$

$$\sigma(X)=\frac{1}{\sqrt{5}},$$

$$\gamma_4=\frac{\mu_4}{\sigma^4}-3=\frac{3\cdot 25}{35}-3=-\frac{6}{7}.$$