

#### 4.4 Функція розподілу та її властивості

Функція розподілу випадкової величини  $X$  – це функція  $F(x) = P\{X < x\}$ , яка визначає для кожного значення  $x$  ймовірність того, що випадкова величина  $X$  прийме значення менше  $x$  в результаті експерименту. Функцію  $F(x)$  також називають *інтегральною функцією розподілу*.

Геометрично визначення функції розподілу можна інтерпретувати наступним чином:  $F(x)$  існує ймовірність того, що в. в.  $X$  прийме значення, яке представлено на числовій осі точкою, що лежить зліва від точки  $x$ , тобто випадкова точка  $X$  потрапить в інтервал  $(-\infty, x)$ , див. рис. 4.1.

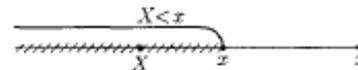


Рисунок 4.1

Властивості функції розподілу:

1.  $0 \leq F(x) \leq 1$ . Це випливає з того, що  $F(x)$  дорівнює ймовірності, а ймовірність будь-якої події знаходиться між нулем і одиницею.

2. Відзначимо також, що  $F(-\infty) = P\{X < -\infty\} = 0$  і  $F(+\infty) = P\{X < +\infty\} = 1$ , оскільки події  $X < -\infty$  і  $X < +\infty$  відповідно неможлива і достовірна.

3. Функція розподілу неспадна, тобто  $F(x_1) \leq F(x_2)$  при  $x_1 < x_2$ . Дійсно, коли  $x_1 < x_2$  поява події  $X < x_2$  еквівалентно виникненню однієї з несумісних подій  $X < x_1$  і  $x_1 \leq X < x_2$ . Тому

$$P\{X < x_2\} = P\{X < x_1\} + P\{x_1 \leq X < x_2\}$$

або

$$F(x_2) - F(x_1) = P\{x_1 \leq X < x_2\}.$$

У правій частині останньої рівності знаходиться невід'ємне значення, тому  $F(x_1) \leq F(x_2)$ . Ця рівність означає, що ймовірність попадання випадкової величини  $X$  в напівінтервал  $[x_1; x_2)$  дорівнює приросту функції розподілу на цьому напівінтервалі.

Графік функції розподілу однієї з неперервних випадкових величин наведено на рис. 4.2.

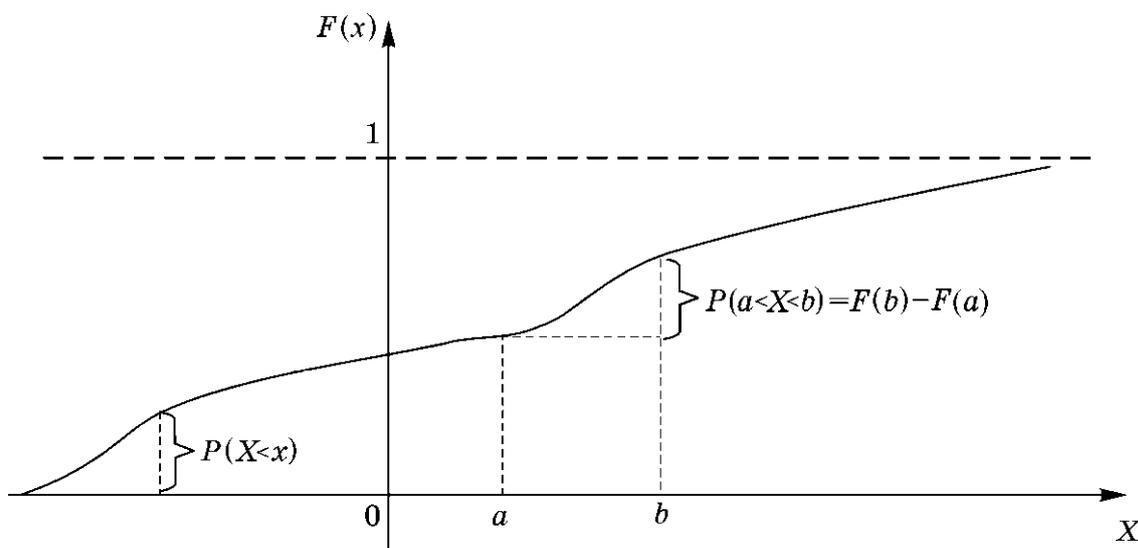


Рисунок 4.2

Функція розподілу може бути задана як для неперервної, так і для дискретної випадкової величини. Для дискретної випадкової величини функція розподілу є функцією накопичених ймовірностей:

$$F(x) = \sum P\{X = x_i\},$$

де підсумовування поширюється на всі значення індексу  $i$ , для яких  $x_i < x$ .

Якщо дискретна випадкова величина  $X$  має закон розподілу:

$X$	$x_1$	$x_2$	...	$x_n$	...
$P$	$p_1$	$p_2$	...	$p_n$	...

тоді її функція розподілу має вигляд ступінчастої функції, причому стрибки функції дорівнюють ймовірності відповідних значень  $X$  (рис. 4.3).

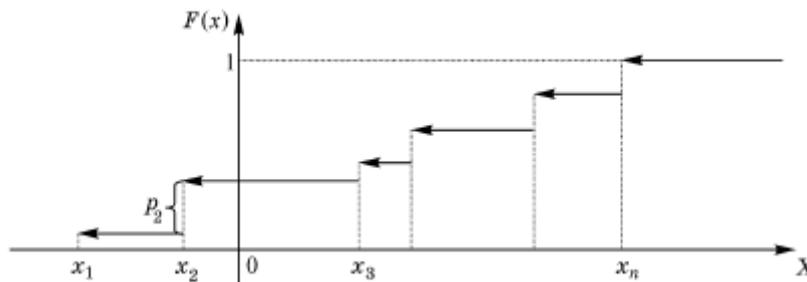


Рисунок 4.3

Функція розподілу неперервної випадкової величини є неперервною, для дискретної випадкової величини це ступінчаста функція.

**Приклад 10** В урні 8 кульок, з яких 5 – білі, інші – чорні. З неї навмання виймають 3 кульки. Знайдіть функцію розподілу кількості білих кульок у виборці і побудуйте її графік.

*Розв'язання.* Даваймо встановимо різні значення  $x$  і знайдемо для них  $F(x) = P\{X < x\}$ :

якщо  $x \leq 0$ , то  $F(x) = P\{X < 0\} = 0$ ;

якщо  $0 < x \leq 1$ , то  $F(x) = P\{X < x\} = P\{X = 0\} = \frac{1}{56}$ ;

якщо  $1 < x \leq 2$ , то  $F(x) = P\{X = 0\} + P\{X = 1\} = \frac{1}{56} + \frac{15}{56} = \frac{16}{56}$ ;

якщо  $2 < x \leq 3$ , то  $F(x) = P\{X = 0\} + P\{X = 1\} + P\{X = 2\} =$   
 $= \frac{1}{56} + \frac{15}{56} + \frac{30}{56} = \frac{46}{56}$ ;

якщо  $3 < x$ , то  $F(x) = P\{X = 0\} + P\{X = 1\} + P\{X = 2\} + P\{X = 3\} =$   
 $= \frac{1}{56} + \frac{15}{56} + \frac{30}{56} + \frac{10}{56} = 1$ .

Таким чином,

$$F(x) = \begin{cases} 0, & \text{якщо } x \leq 0; \\ \frac{1}{56}, & \text{якщо } 0 < x \leq 1; \\ \frac{16}{56}, & \text{якщо } 1 < x \leq 2; \\ \frac{46}{56}, & \text{якщо } 2 < x \leq 3; \\ 1, & \text{якщо } 3 < x. \end{cases}$$

Графік побудованої функції наведено на рис. 4.4.

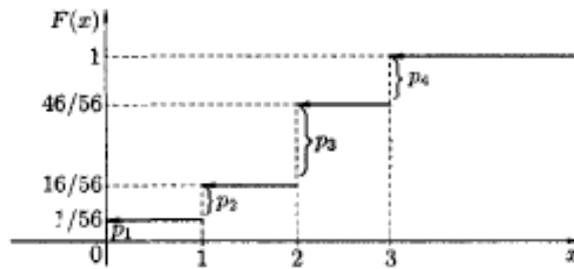


Рисунок 4.4

#### 4.5 Найважливіші дискретні випадкові величини та їх числові характеристики

##### *Вироджена випадкова величина.*

Будь-яку константу  $C$  можна розглядати як випадкову величину, що приймає єдине значення:  $X = X(\omega) = C$  для будь-якої  $\omega \in \Omega$ .

Закон розподілу виродженої випадкової величини має вигляд:

$X$	$C$
$P$	1

Вираз для функції розподілу виродженої випадкової величини та її графік також мають вироджений вигляд:

$$F(x) = \begin{cases} 0, & \text{якщо } x \leq C; \\ 1, & \text{якщо } x > C. \end{cases}$$

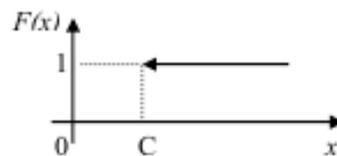


Рисунок 4.5

Математичне сподівання  $MC = C$ , дисперсія  $DC = 0$ .

##### *Індикаторна випадкова величина.*

Будь-яка випадкова подія  $A$  може бути пов'язана з випадковою величиною виду:

$$X = I_A(\omega) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases}$$

Випадкова величина  $X = I_A$  називається індикатором випадкової події  $A$  або індикатором випадкової величини. Вона приймає тільки два значення  $x_1 = 0$  і  $x_2 = 1$ , при цьому

$$P(X = 1) = P(A) = p, \quad P(X = 0) = P(\bar{A}) = 1 - p = q.$$

Закон розподілу індикаторної випадкової величини має вигляд:

$X$	0	1
$P$	$p$	$q$

Аналітичним виразом і графіком функції розподілу є:

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

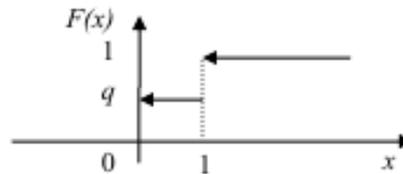


Рисунок 4.6

Математичне сподівання  $MX = MI_A = \sum_i x_i p_i = 0 \cdot q + 1 \cdot p = p$ , дисперсія  $DX = DI_A = MX - (MX)^2 = \sum_i x_i^2 p_i - p^2 = 0^2 \cdot q + 1^2 \cdot p - p^2 = p(1 - p) = pq$ .

**Біноміальна випадкова величина.**

Біноміальною називається дискретна випадкова величина  $X$ , яка є числом успіхів в  $n$  незалежних випробуваннях, проведених за схемою Бернуллі, з ймовірністю успіху в одному випробуванні, що дорівнює  $p$ .

Множиною можливих значень біноміальної випадкової величини є:

$$X = \{0, 1, \dots, n\} = \{x_k = k, k = \overline{0, n}\}.$$

Ймовірності, з якими приймаються значення, визначаються за формулою Бернуллі:

$$p_k = P(X = k) = P_n(k) = C_n^k p^k q^{n-k}, k = \overline{0, n}.$$

Закон розподілу такий:

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

і називається біноміальним законом розподілу.

Скорочене позначення для біноміальної випадкової величини:  $X \sim Bi(n; p)$ .

Математичне сподівання  $MX = np$ , дисперсія  $DX = MX^2 - (MX)^2 = n^2 p^2 - np^2 + np - n^2 p^2 = np(1 - p) = npq$ .

**Геометрична випадкова величина.**

Геометричною називається дискретна випадкова величина  $X$ , яка являє собою кількість випробувань, виконаних за схемою Бернуллі до появи першого успіху з ймовірністю успіху в одному випробуванні, рівному  $p$ .

Геометрична випадкова величина має злічений набір можливих значень:

$$X = \{1, 2, \dots, n, \dots\} = \{x_k = k, k = 1, 2, \dots\}.$$

Ймовірності значень визначаються за формулою:

$$p_k = P(X = k) = q^{k-1} p, k = 1, 2, \dots$$

Закон розподілу такий:

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

і називається геометричним законом розподілу.

Скорочене позначення для геометричної випадкової величини:  $X \sim G(p)$ .

Математичне сподівання 
$$MX = \sum_i x_i p_i = \sum_{i=1}^{\infty} i q^{i-1} p = p \sum_{i=1}^{\infty} i q^{i-1} = p \frac{d}{dq} \left( \sum_{i=1}^{\infty} q^i \right) = p \frac{d}{dq} \left( \frac{q}{1-q} \right) = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p},$$
 дисперсія 
$$DX = MX^2 - (MX)^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

**Випадкова величина Пуассона.**

Пуассонівською називається цілочисельна випадкова величина, набір можливих значень якої

$$X = \{0, 1, \dots, n, \dots\} = \{x_k = k, k = 0, 1, 2, \dots\},$$

а ймовірності, з якими приймаються значення, задаються формулою:

$$p_k = P(X = k) = \frac{a^k}{k!} e^{-a}, k = 0, 1, 2, \dots$$

Число  $a > 0$  називається параметром випадкової величини Пуассона.

Закон розподілу такий:

$X$	0	1	...	$n$	...
$P$	$e^{-a}$	$ae^{-a}$	...	$\frac{a^n e^{-a}}{n!}$	...

і називається законом розподілу Пуассона.

Скорочене позначення для пуассонівської випадкової величини:  $X \sim P(a)$ .

Математичне сподівання 
$$MX = \sum_i x_i p_i = \sum_{i=0}^{\infty} i \frac{a^i}{i!} e^{-a} = a e^{-a} \sum_{i=1}^{\infty} \frac{a^{i-1}}{(i-1)!} = a e^{-a} e^a = a,$$
 дисперсія 
$$DX = MX^2 - (MX)^2 = a^2 + a - a^2 = a.$$

**4.6 Найважливіші неперервні випадкові величини та їх числові характеристики**

**Рівномірна випадкова величина.**

Кажуть, що неперервна випадкова величина  $X$  має рівномірний закон розподілу (рівномірний розподіл) на відрізку  $[a; b]$ , якщо множина її можливих значень  $X = [a; b]$ , а щільність ймовірності є постійною на цьому відрізку:

$$f(x) = \begin{cases} C, & x \in [a; b]; \\ 0, & x \notin [a; b]. \end{cases}$$

Константа  $C$  в цьому випадку однозначно визначається з умови нормування:

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_a^b C dx = C(b - a), \text{ тобто } C = \frac{1}{b-a}.$$

Таким чином, рівномірно розподілена випадкова величина має щільність ймовірності:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a; b]; \\ 0, & x \notin [a; b] \end{cases}$$

і для неї використовується позначення:  $X \sim R[a; b]$ .

Знайдемо функцію розподілу  $F(x)$  випадкової величини  $X \sim R[a; b]$ . Для цього розглянемо три випадки:

- 1) якщо  $x < a$ , то  $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x 0dt = 0$ ;
- 2) якщо  $x \in [a; b]$ , то  $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^a 0dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$ ;
- 3) якщо  $x > b$ , то  $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^a 0dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0dt = 1$ .

Нарешті маємо:

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

Графіки щільності ймовірностей (в точках  $x = a$  і  $x = b$  функція  $f(x)$  має розрив) і функцій розподілу (в точках  $x = a$  і  $x = b$  функція  $F(x)$  не існує) випадкової величини  $X \sim R[a; b]$  представлені на рис. 4.7.

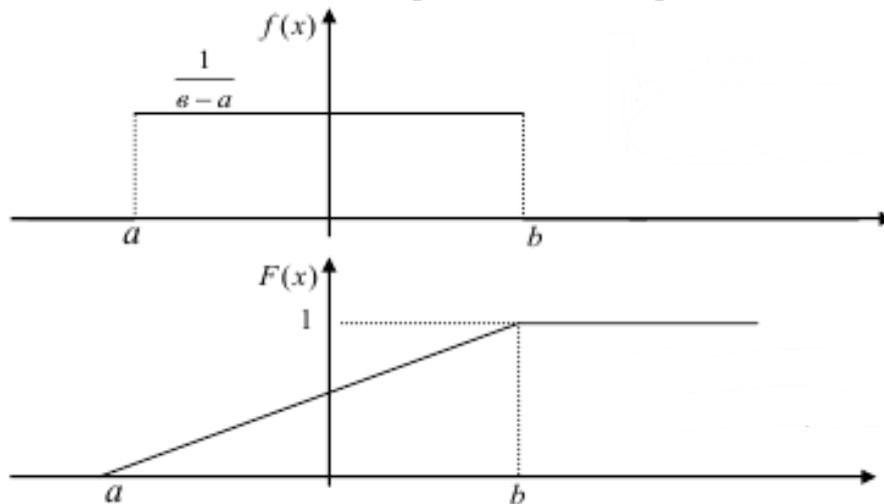


Рисунок 4.7 – Графіки розподілу ймовірностей і функції щільності рівномірної випадкової величини

### **Показникова (експоненціальна) випадкова величина.**

Кажуть, що неперервна випадкова величина  $X$  має показниковий закон розподілу (показниковий, експоненціальний розподіл), якщо множина її можливих значень  $X = [0; +\infty)$ , а щільність ймовірності має вигляд:

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

Число називається  $\lambda > 0$  параметром показникового розподілу, а для показникової випадкової величини використовується скорочене позначення:  $X \sim E(\lambda)$ .

Знайдемо функцію розподілу випадкової величин  $X \sim E(\lambda)$ . Для цього розглянемо два випадки:

- 1) якщо  $x < 0$ , то  $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x 0dt = 0$ ;
- 2) якщо  $x \geq 0$ , то  $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 0dt + \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$ .

Нарешті маємо:

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

Графіки щільності ймовірностей (в точці  $x = 0$  функція  $f(x)$  має розрив) і функцій розподілу (в точці  $x = 0$  функція  $F(x)$  не існує) випадкової величини  $X \sim E(\lambda)$  представлені на рис. 4.8.

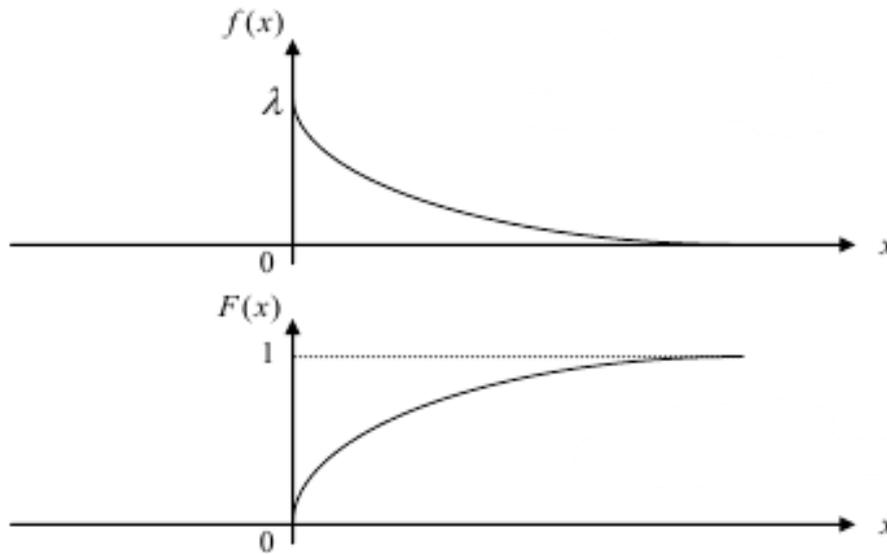


Рисунок 4.8 – Графіки щільності ймовірностей та функції розподілу показникової випадкової величини

**Нормальна (гауссова) випадкова величина.**

Кажуть, що неперервна випадкова величина  $X$  має нормальний закон розподілу (нормальний, розподіл Гауса) з параметрами  $(a, \sigma^2)$ , якщо набір її можливих значень  $X = (-\infty; +\infty)$ , а щільність ймовірності має вигляд:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, -\infty < a < +\infty, \sigma > 0.$$

Скорочене позначення нормальної випадкової величини:  $X \sim N(a, \sigma^2)$ .

Крива щільності ймовірності (рис. 4.9) має симетричну форму по відношенню до прямої  $x = a$  і має максимум в точці  $x = a$ .

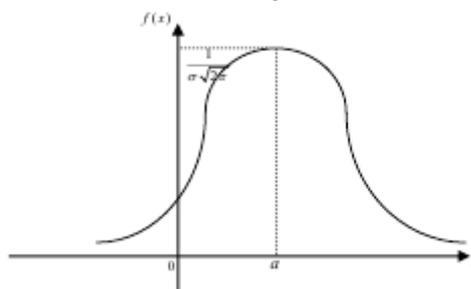


Рисунок 4.9 – Графік щільності ймовірності нормальної випадкової величини

Залежно від зміни параметрів щільність ймовірності нормального закону розподілу змінюється наступним чином: якщо параметр фіксований  $\sigma^2$ , то при зміні  $a$  крива  $f(x)$ , не змінюючи своєї форми, просто зміщується уздовж осі абсцис (див. рис. 4.10). Таким чином, параметр  $a$  – це параметр зсуву (положення). Також параметр  $a$  характеризує середнє значення випадкової величини.

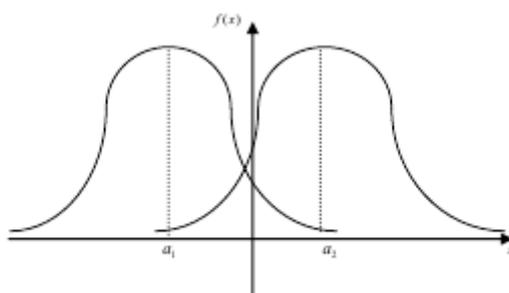


Рисунок 4.10 – Зміна в графіку нормальної щільності ймовірності в залежності від параметра  $a$

Зміна  $\sigma^2$  при фіксованому  $a$  еквівалентно зміні масштабу кривої  $f(x)$  по обох осях: при збільшенні  $\sigma^2$  щільність ймовірності стає більш плоскою, розтягуючись уздовж осі абсцис; при зменшенні  $\sigma^2$  – вона витягується вгору, одночасно стискаючись зі сторін (ефект дії умови нормування). Таким чином, параметр  $\sigma^2$  є параметром масштабу (див. рис. 4.11).

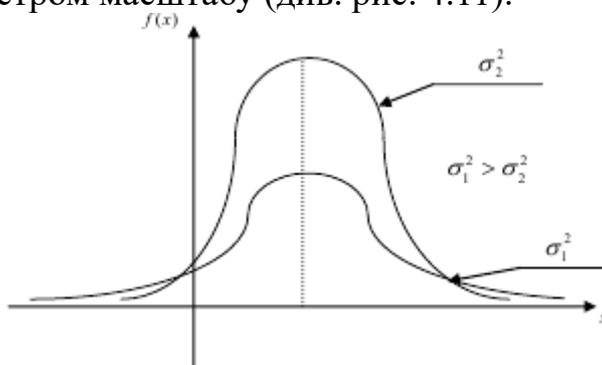


Рисунок 4.11 – Зміна графіка нормальної щільності ймовірності в залежності від параметра  $\sigma^2$

Також параметр  $\sigma^2$  характеризує ступінь розбросу значень випадкової величини навколо середнього значення  $a$  в наступному сенсі. Чим менше  $\sigma^2$ , тим більша при фіксованому  $l$  ймовірність виду  $P(|X - a| < l)$ , як площа під щільністю ймовірності, або, іншими словами, при меншому  $l$  можна отримати задану ймовірність виду  $P(|X - a| < l)$ . Це означає, що при зменшенні  $\sigma^2$  значення випадкової величини  $X \sim N(a, \sigma^2)$  більш щільно групуються навколо  $a$ , тобто ступінь розбросу значень випадкової величини становить приблизно середнє значення  $a$  або менше.

Якщо  $a = 0$  і  $\sigma^2 = 1$ , то нормальний закон розподілу називається стандартним, його щільність ймовірностей має вигляд:

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

і називається *функцією Гауса*.

Функція розподілу випадкової величини  $X \sim N(0,1)$  має такий вигляд:

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

і не виражається в елементарних функціях. Функція називається  $\Phi(x)$  *функцією Лапласа* (або інтегралом ймовірностей).

**Властивості функції Лапласа  $\Phi(x)$ :**

1.  $\Phi(0) = 0,5$ ;
2.  $\Phi(-x) = 1 - \Phi(x)$  для  $x > 0$ .

Значення функції Лапласа  $\Phi(x)$  для  $x > 0$  зведені у таблицю.

Імовірність потрапляння випадкової величини  $X \sim N(a, \sigma^2)$  в заданий інтервал  $[x_1, x_2)$  визначається за формулою:

$$P\{x_1 \leq X < x_2\} = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right).$$

Найбільш простою виражається у функції Лапласа ймовірність потрапляння випадкової величини  $X \sim N(a, \sigma^2)$  в інтервал довжини  $2l$ , симетричний відносно точки  $x = a$ :

$$\begin{aligned} P\{a - l \leq X < a + l\} &= P\{|X - a| < l\} = \Phi\left(\frac{a + l - a}{\sigma}\right) - \Phi\left(\frac{a - l - a}{\sigma}\right) = \\ &= \Phi\left(\frac{l}{\sigma}\right) - \Phi\left(-\frac{l}{\sigma}\right) = 2\Phi\left(\frac{l}{\sigma}\right) - 1. \end{aligned}$$

Далі, якщо покласти  $l = 3\sigma$  і врахувати, що  $\Phi(3) = 0,9987$ , ми отримаємо:

$$P\{a - 3\sigma \leq X < a + 3\sigma\} = P\{|X - a| < 3\sigma\} = 2\Phi(3) - 1 = 0,9974.$$

Результат називається «Правило трьох сигм» (рис. 4.12), що означає, що «практично всі» значення випадкової величини  $X \sim N(a, \sigma^2)$  знаходяться в межах інтервалу  $(a - 3\sigma, a + 3\sigma)$  в тому сенсі, що ймовірність випадкової величини  $X \sim N(a, \sigma^2)$  прийняти значення, яке не належить цьому інтервалу, мізерно мала ( $\approx 0,0026$ ).

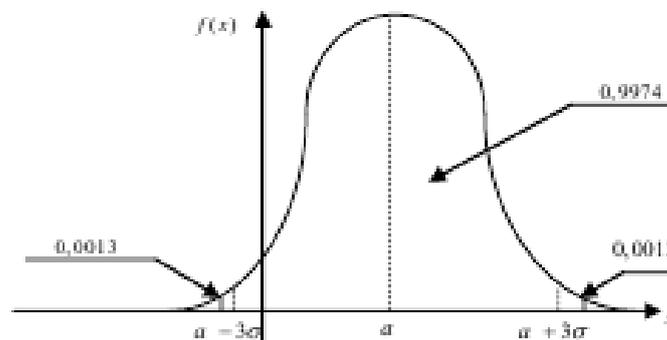


Рисунок 4.12 – Геометрична ілюстрація «Правила трьох сигм»