ВИПАДКОВІ ВЕЛИЧИНИ

- Способи задання випадкових величин
- Числові характеристики випадкових величин
- Основні дискретні розподіли
- Основні неперервні розподіли
- Граничні теореми

Означення. ***Випадковою величиною*** називають величину, яка в результаті випробування набуває одне і тільки одне можливе значення, наперед не відоме і не залежні від випадкових причин, які не можуть бути враховані наперед.

Приклади.

1. Кількість хлопчиків , що народилися серед 100 новонароджених.

2. Відстань, яку пролетить снаряд при пострілі.

Дискретні і ***неперервні*** випадкові величини.

Випадкові величини: $X, Y, Z, \ldots$, їх значення: $x, y, z, \ldots$

Означення. ***Законом розподілу*** дискретної випадкової величини називають відповідність між можливими значеннями і їх імовірностями.

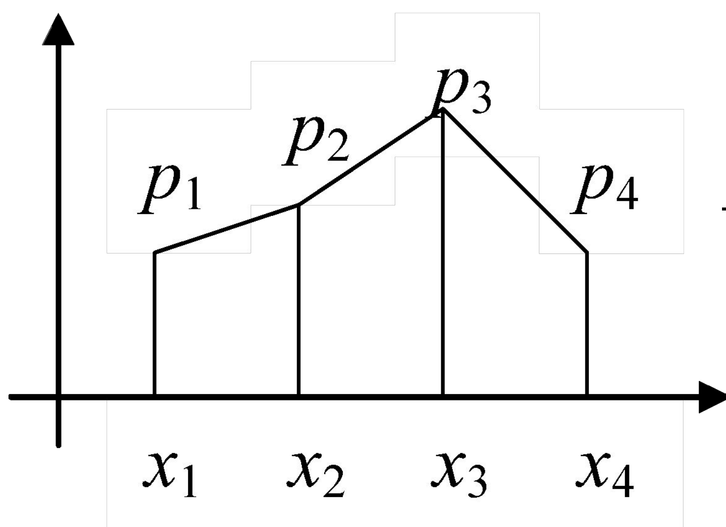
Таблично:

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

$$p_1 + p_2 + \ldots + p_n = 1$$

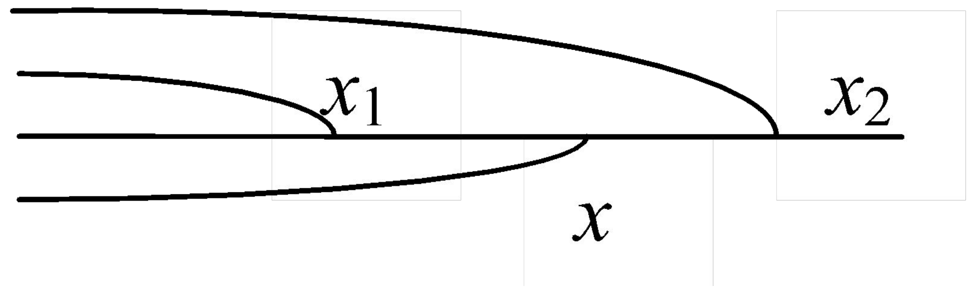
Аналітично: $p(x_i) = f(i), \quad i = 1, 2, \ldots, n$

Графічно:



– ***багатокутник розподілу***

$p(X < x)$



Означення. ***Функцією розподілу*** випадкової величини X називається функція *F(x)*, що визначає імовірність того, що випадкова величина X приймає значення, менше x, тобто

$$\boxed{F(x) = p(X < x)}.$$

Функцію $F(x)$ називають ***інтегральною*** функцією розподілу.

Означення. Випадкова величина X називається ***неперервною***, якщо її функція розподілу неперервна на всій числовій осі.

Властивості функції розподілу:

1) $0 \le F(x) \le 1$

2) Якщо $x_1 < x_2$, то $F(x_1) \le F(x_2)$

3) $p(a \le x < b) = F(b) - F(a)$

4) Якщо можливі значення випадкової величини належать інтервалу (a, b), то а) $F(x)=0$ при $x \le a$ і б) $F(x)=1$ при $x \ge b$

5) Якщо X – непрерывна випадкова величина, то імовірність того, що вона прийме одне визначене значення дорівнює нулю: $p(X=x) = 0$.

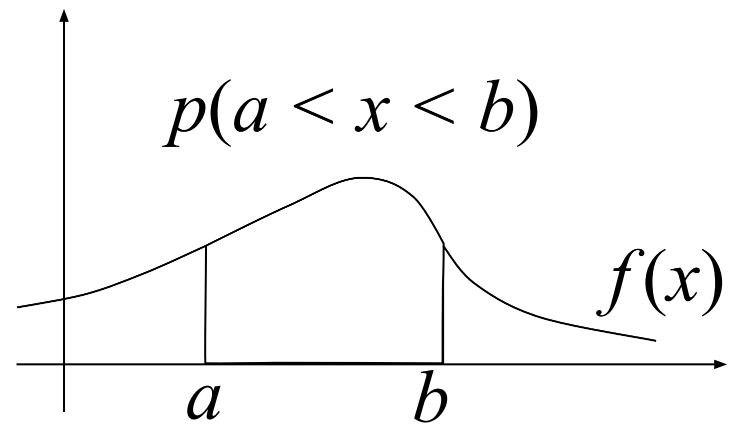
Означення. ***Щільністю розподілу*** *імовірностей* неперервної випадкової величини X називають функцію, яка є похідною від функції розподілу:

$$\boxed{f(x) = F'(x).}$$

Функцію $f(x)$ називають ***щільністю імовірності*** або ***диференціальною*** функцією розподілу.

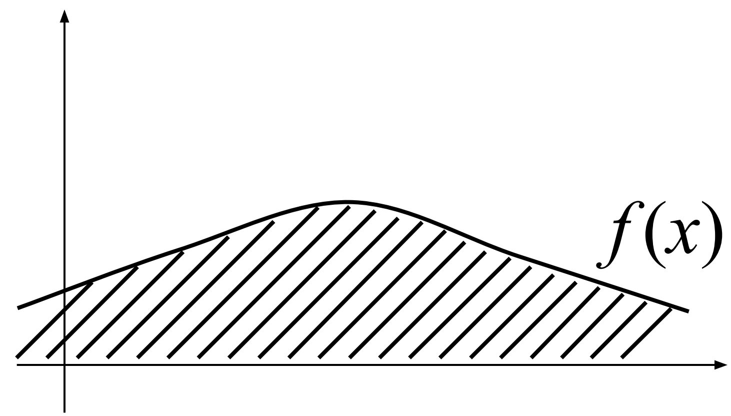
Властивості щільності ймовірності:

1) $$p(a < x < b) = \int_a^b f(x)dx$$



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

3) $$f(x) \geq 0$$



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

Дискретні і неперервні випадкові величини.

Дискретна випадкова величина:
приймає окремі, ізольовані значення.

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Неперервна випадкова величина:
можливі значення цілком заповнюють деякій проміжок.

$F(x) = p(X < x)$
функція розподілу

$f(x) = F'(x)$
щільність розподілу

Означення. Математичне сподівання дискретної випадкової величини X називається сума добутків всіх можливих значень випадкової величина на відповідні їм імовірності. Позначається *M (X)*.

Нехай

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n$$

Якщо випадкова величина *X* приймає нескінченну множину значень, то

$$M(X) = \sum_{i=1}^{\infty} x_i p_i$$

Імовірнісний зміст математичного сподівання: математичне сподівання наближено дорівнює середньому арифметичному значень випадкової величини.

Доведення

Нехай *n* – кількість випробувань (достатньо велика).

Нехай випадкова величина *X* приймає значення $x_1, x_2, \ldots, x_k$ відповідно $m_1, m_2, \ldots, m_k$ раз.

$$m_1 + m_2 + \ldots + m_k = n$$

Знайдем середнє арифметичне всіх значень:

$$\overline{X} = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_k m_k}{n} = x_1 \cdot \frac{m_1}{n} + x_2 \cdot \frac{m_2}{n} + \ldots + x_k \cdot \frac{m_k}{n} \approx$$

$$x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots \; x_k \cdot p_k = M(X)$$

Означення. *Математичне сподівання неперервної випадкової величини x,* можливі значення, якої належать відрізку *[a, b],* називається визначений інтеграл

$$M(X) = \int_{a}^{b} xf(x)dx$$

Якщо можливі значення випадкової величини розподілені по всій осі *OX*, то

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

Властивості математичного сподівання

- 1. $M(C)=C,$ де $C=const.$
- 2. $M(CX)=CM(X)$
- 3. $M\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}M\left(X_i\right)$
- 4. $M(aX+b)=aM(X)+b$
- 5. $M(XY)=M(X)M(Y)$ - <u>тільки для незалежних ВВ !</u>
- 6. $M(X)=0,$ якщо $f(x)=f(-x)$

Приклад.

X	-0.1	0.1
p	$1/2$	$1/2$

$$M(X) = -0.1 \cdot \frac{1}{2} + 0.1 \cdot \frac{1}{2} = 0$$

Y	-100	100
p	$1/2$	$1/2$

$$M(Y) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} = 0$$

$M(X) = M(Y)$, але X і Y сильно відрізняються

Питання: чи можна оцінити дисперсії можливих значень випадкової величини розрахувавши відхилення кожного з цих значень від математичного сподівання і потім знайти їх середнє?

$X-M(X)$	$x_1-M(X)$	$x_2-M(X)$	…	$x_n-M(X)$
p	p_1	p_2	…	p_n

$$M(X-M(X)) =$$

$$(x_1 - M(X))\cdot p_1 + (x_2 - M(X))\cdot p_2 + \ldots + (x_n - M(X))\cdot p_n =$$

$$\underbrace{x_1\cdot p_1 + x_2\cdot p_2 + \ldots + x_n\cdot p_n}_{M(X)} - M(X)\cdot\underbrace{(p_1 + p_2 + \ldots + p_n)}_{1} =$$

$$M(X) - M(X) = 0$$

$(X-M(X))^2$	$(x_1-M(X))^2$	$(x_2-M(X))^2$	…	$(x_n-M(X))^2$
p	p_1	p_2	…	p_n

Означення. Дисперсією випадкової величини X називається математичне сподівання квадрата відхилення випадкової величини X від математичного сподівання:

$$D(X) = M[X - M(X)]^2$$

1. Дискретна випадкова величина

$$D(X) = M[X - M(X)]^2$$

$(X - M(X))^2$	$(x_1 - M(X))^2$	$(x_2 - M(X))^2$	$\ldots$	$(x_n - M(X))^2$
p	p_1	p_2	$\ldots$	p_n

$$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 +$$

$$\ldots + (x_n - M(X))^2 p_n$$

або $$D(X) = \sum_i (x_i - M(X))^2 p_i$$

$$D(X) = M(X^2) - [M(X)]^2$$

2. Неперервна випадкова величина

За означенням $\qquad D(X) = M[X - M(X)]^2$

Але $M(X) = \int_a^b xf(x)dx \quad \Rightarrow$

$$M[X - M(X)]^2 = \int_a^b (x - M(X))^2 f(x)dx \quad \Rightarrow$$

$$D(X) = \int_a^b (x - M(X))^2 f(x)dx$$

$$D(X) = \int_a^b x^2 f(x)dx - [M(X)]^2$$

Середнє квадратичне відхилення

Означення. Середнім квадратичним відхиленням випадкової величини X називається корінь з її дисперсії:

$$\sigma(X) = \sqrt{D(X)}$$

Властивості

Дисперсія	*Середнє квадратичне відхилення*
1. $D(C)=0$, де $C=const.$	1. $\sigma(C)=0$, де $C=const.$
2. $D(C+X)=D(X)$	2. $\sigma(C+X)=\sigma(X)$
3. $D(CX)=C^2D(X)$	3. $\sigma(CX)=C\sigma(X)$
4. $D\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}D\left(X_i\right)$	

Приклад

- Дискретна ВВ задана рядом розподілу:

x_i	0	1	2
p_i	0.2	0.5	0.3

- Обчислить математичне сподівання. дисперсію и середнє квадратичне відхилення.

- 1. $M(X)=\sum_{i=1}^{n} x_i p_i = 0\cdot 0.2+1\cdot 0.5+2\cdot 0.3=1.1$.
- 2. $D(X)=M(X^2)-\left[M(X)\right]^2$.
- $M(X^2)=0^2\cdot 0.2+1^2\cdot 0.5+2^2\cdot 0.3=1.7$.
- $D(X)=1.7-(1.1)^2=0.49$.
- 3. $\sigma(X)=+\sqrt{D(X)}=\sqrt{0.49}=0.7$.

- Неперервна ВВ задана щільністю імовірності

$$f(x)=\begin{cases} 2x & при\ 0 \le x < 1 \\ 0 & при\ x < 0\ і\ x \ge 1 \end{cases}$$

- Обчислить математичне сподівання. дисперсію и середнє квадратичне відхилення.

1. $$M(x)=\int_{-\infty}^{+\infty} x\cdot f(x)dx=\int_0^1 x\cdot 2xdx=2\int_0^1 x^2dx=2\frac{x^3}{3}\Big|_0^1=\frac{2}{3}$$

2. $$D(X)=M(X^2)-[M(X)]^2$$

$$M(x^2)=\int_{-\infty}^{+\infty} x^2 f(x)dx=\int_0^1 x^2\,2xdx=2\int_0^1 x^3dx=2\frac{x^4}{4}\Big|_0^1=\frac{1}{2}$$

$$D(X)=\frac{1}{2}-\left(\frac{2}{3}\right)^2\approx 0.28$$

3. $$\sigma(X)=+\sqrt{D(X)}=\sqrt{0.28}\approx 0.53$$

4. *Початковий момент порядку k*

$$\nu_k = M(X^k)$$

$$\nu_k = \sum_i x_i^k \cdot p_i$$ – дискретна

$$\nu_k = \int_{-\infty}^{+\infty} x^k \cdot f(x)dx$$ – неперервна

Початковий момент першого порядку:

$$\nu_1 = M(X)$$ – математичне сподівання

- Якщо $k = 0$, то $\nu_0 = M(X^0) = 1$.
- Якщо $k = 1$, то $\nu_1 = M(X^1) = M(X)$.
- Якщо $k = 2$, то $\nu_2 = M(X^2)$.
- математичне сподівання ВВ це початковий <u>момент 1-го порядку</u>, а дисперсія може бути виражена через початкові моменти 1-го и 2-го порядків:
- $$D(X) = M(X^2) - [M(X)]^2 = \nu_2 - \nu_1^2$$.

5. Центральний момент порядку k

$$\mu_k = M(X - M(X))^k$$

$$\mu_k = \sum_i (x_i - M(X))^k \cdot p_i$$ – дискретна

$$\mu_k = \int_{-\infty}^{+\infty} (x - M(X))^k \cdot f(x)dx$$ – неперервна

Центральний момент другого порядку:

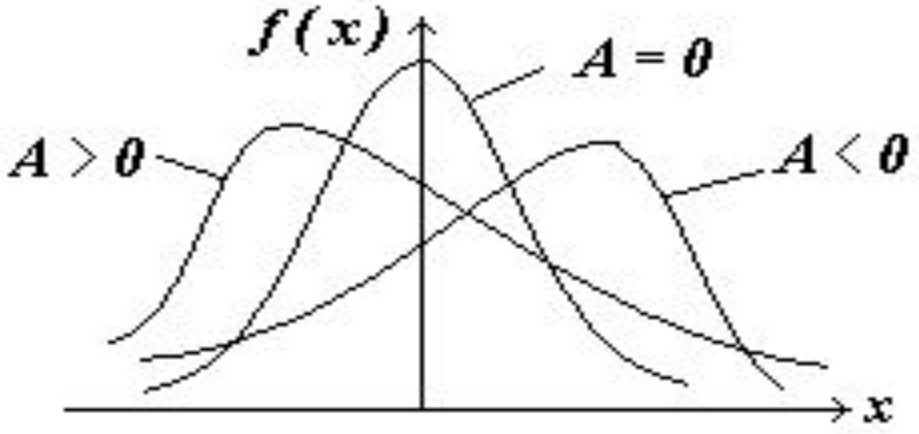
$$\mu_2 = M(X - M(X))^2$$ – дисперсія

- Якщо $k = 0$, то $\mu_0 = M(\delta^0) = 1$;
- Якщо k = 1,то
$$\mu_1 = M(\delta^1) = M\left\{\left[X - M(X)\right]^1\right\} =$$
- $$= M(X) - M(X) = 0$$;
- Якщо $k = 2$, то
- $$\mu_2 = M\left\{\left[X - M(X)\right]^2\right\} = D(X)$$,
- *дисперсія ВВ це центральний момент другого порядку :* $\mathbf{D(X) = \mu_2}$

- Вводиться <u>коефіцієнт асиметрії</u> *(характеристика скошеності)*:

$$A = \frac{\mu_3}{\sigma_x^3},$$

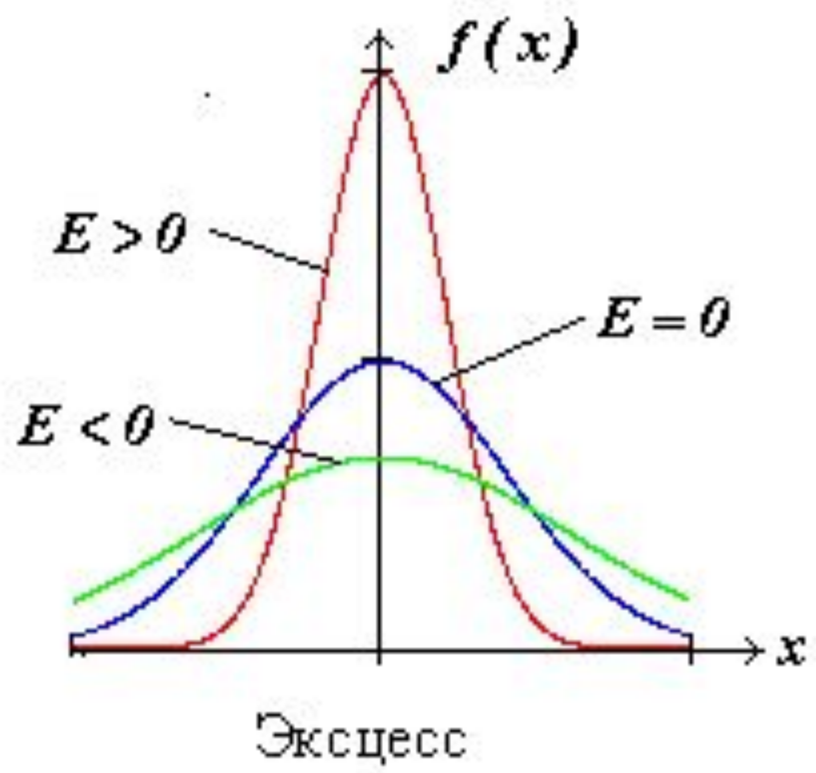
- де – $\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$
- центральний момент 3-го порядку, а σ_x – середнє квадратичне відхилення ВВ



Асимметрия

- Центральний $\mu_4 = \nu_3 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4$ момент 4-го порядку використовується для характеристики положення вершини кривої (крутості кривої) розподілу відносно еталона – *нормального розподілу*, для якого відношення
- $$\frac{\mu_4}{\sigma_x^4} = 3 \cdot$$

- Вводиться числова характеристика, яка називається ***ексцесом*** *кривої розподілу* і обчислюється як



$$E = \frac{\mu_4}{\sigma_x^4} - 3$$

6. *Мода*

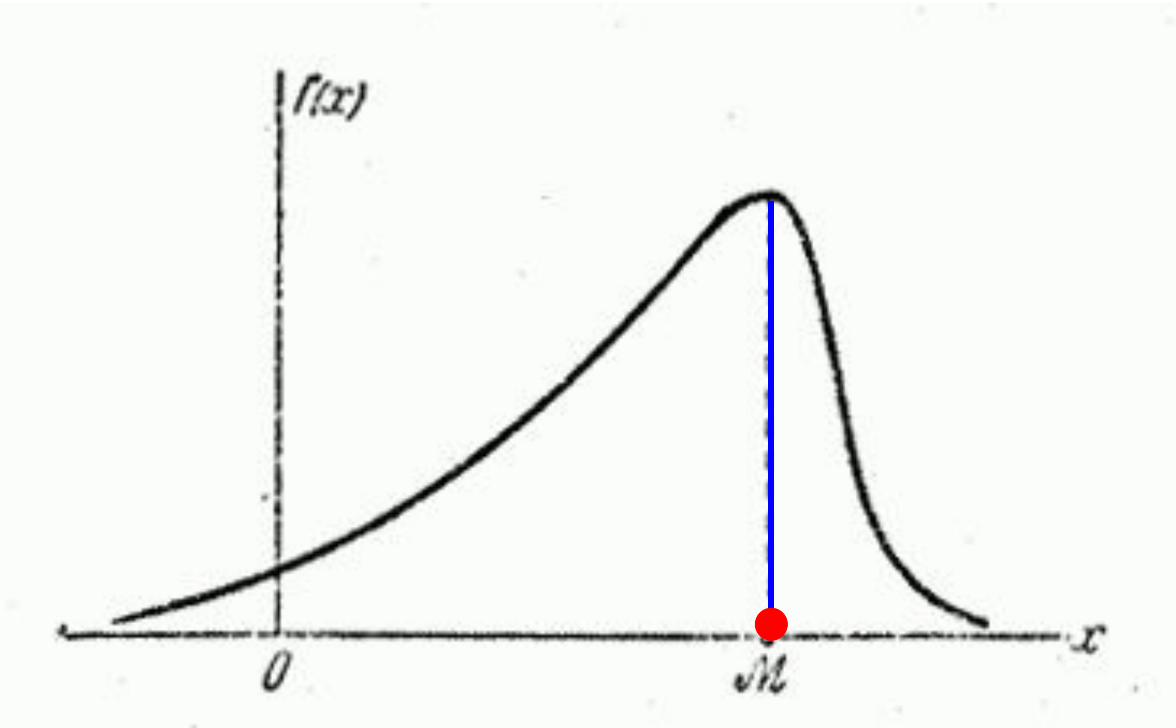
Для дискретної випадкової величини ***мода*** – це найбільш імовірне значення ВВ.

x_i	10	20	30	40	50	60
p_i	0,24	0,36	0,20	0,15	0,03	0,02

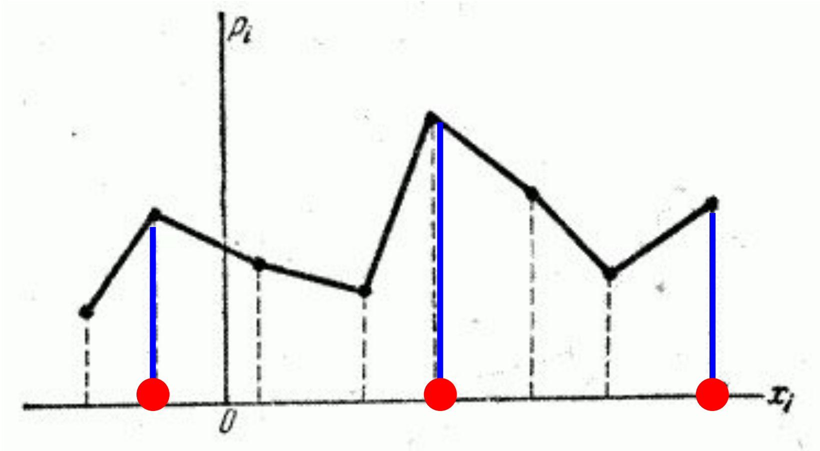
$$0{,}24 < 0{,}36 > 0{,}20$$

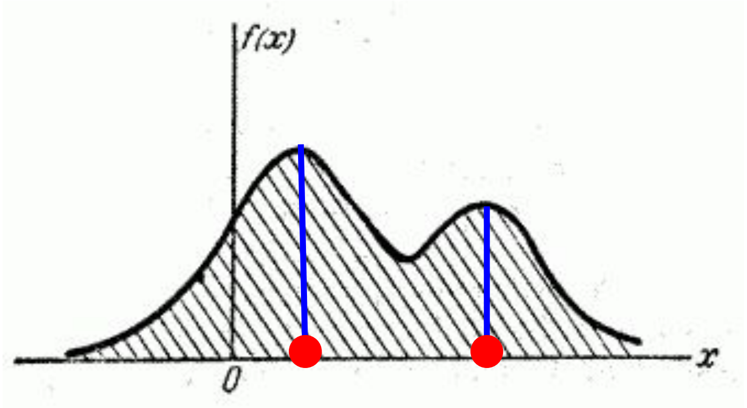
Мода: 20

Для неперервної випадкової величини ***мода*** – значення, при якому щільність розподілу $f(x)$ досягає максимуму.



У випадкової величини може бути декілька мод.





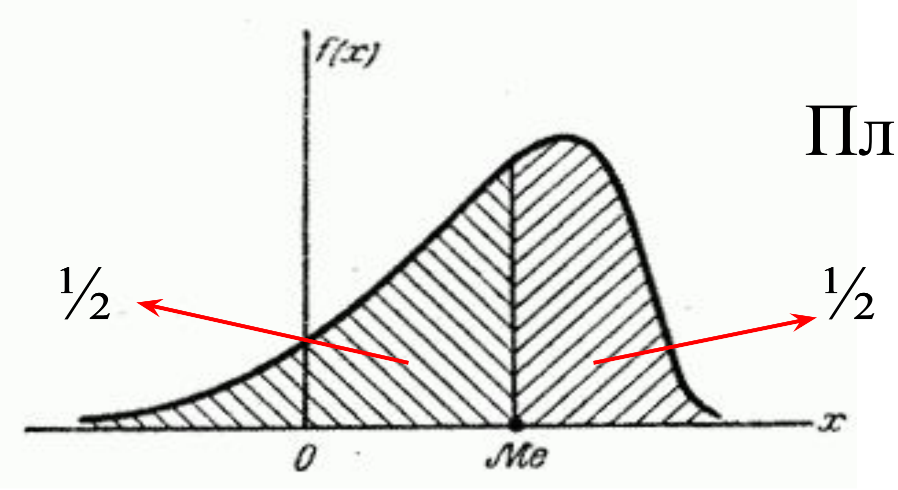
Розподіли з однією, двома або більшим числом мод називаються відповідно ***унімодальними***, ***бімодальними*** або ***мультимодальними***.

7. *Медіана*

таке число m, для якого однаково імовірно, що випадкова величина менше m або більше m, то б то

$$p(X < m) = p(X > m) = 0.5$$

Геометрично медіана – це абсциса точки, в якій площа, обмежена кривою щільності розподілу, ділиться навпіл.



Площа всієї фігури: = 1

8. *Квантіль рівня p*

таке число x_p, що $p(X < x_p) = p \quad \Leftrightarrow \quad F(x_p) = p$

$F(x)$ – функція розподілу

$$x_p = F^{-1}(p)$$

$F^{-1}(x)$ – функція, обернена до функції розподілу

Квантіль рівня 0.5 – це ***медіана***.

Квантілі рівня ¼, ½, ¾ називають відповідно першим, другим і третім ***квартілями***.

Квантілі рівня 0.1, 0.2, 0.3, …, 0.9 називають ***децілями***.

Квантілі рівня 0.01, 0.02, 0.03, …, 0.99 називають ***процентілями***.

Основні дискретні розподіли

- ***Біноміальний розподіл***
- ***Розподіл Пуассона***
- ***Геометричний розподіл***
- ***Гіпергеометричний розподіл***
- ***Рівномірний розподіл***

1. Біноміальний розподіл

p – імовірність події А

X – число появ події А в n незалежних випробуваннях

Можливі значення: $k = 0, 1, 2, \ldots, n$

$q = 1 - p.$

$$p(k) = p^k q^{n-k} C_n^k$$

p – параметр розподілу

Біном Ньютона: $(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k}$

$M(X) = np$ і $D(X) = npq$

2. Розподіл Пуассона

n – дуже велике, p – дуже мале, $np = \lambda$

X – число появ події А в n незалежних випробуваннях

Можливі значення: $k = 0, 1, 2, \ldots, n$

$$p(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

λ – параметр розподілу

$$M(X) = D(X) = \lambda$$

3. Геометричний розподіл

p – імовірність події А

X – число випробувань, яке необхідно провести ***до першої появи події А***

Можливі значення: всі натуральні числа $k = 1, 2, 3, \ldots$

$q = 1 - p.$

$$P(k) = q^{k-1} p$$

p – параметр розподілу

$p, qp, q^2p, q^3p, q^4p, \ldots$ – геометрична прогресія

$$M(X) = \frac{q}{p} \quad \text{і} \quad D(X) = \frac{q}{p^2}$$

4. Гіпергеометричний розподіл

Партія з N виробів містить M стандартних.
З партії випадковим чином вибирають n виробів.

X – число стандартних виробів серед відібраних

Можливі значення: $k = 0, 1, 2, \ldots, \min(M,n)$

$$P(k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

N, M, n – параметри розподілу

$$M(X) = \frac{n \cdot M}{N} \quad \text{і} \quad D(X) = \frac{M \cdot (N-M) \cdot (N-n) \cdot n}{N^2 \cdot (N-1)}$$

5. Рівномірний розподіл

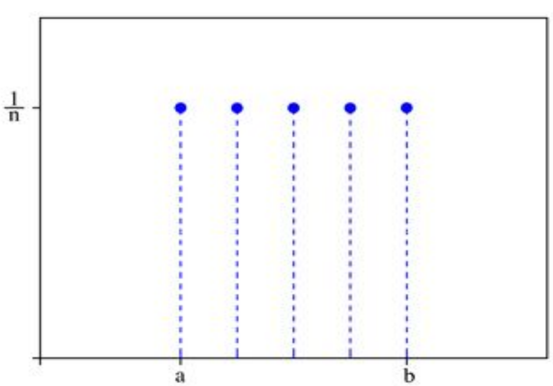
Дискретна випадкова величина має рівномірний розподіл, якщо її функція імовірності на всій області визначення (a,b) має вид

$$P(x)=1/n,$$

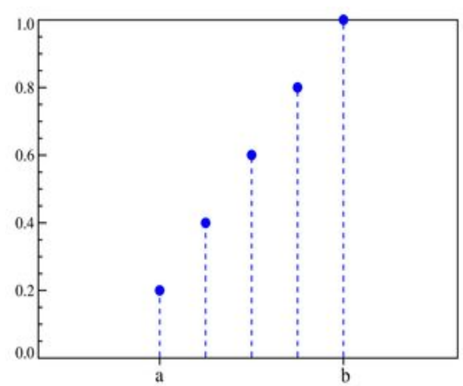
де n — число випробувань

$M[x]=(a+b)/2$ - математичне сподівання

$D[x]=(n^2-1)/12$ - дисперсія



Графік характеристичної функції



Графік кумулятивної функції

Основні неперервні розподіли

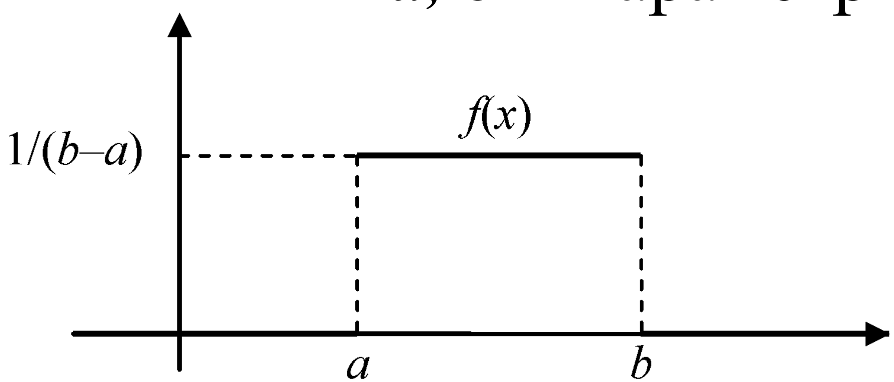
- ***Рівномірний розподіл***
- ***Показниковий (експоненціальний) розподіл***
- ***Нормальний розподіл***
- ***Розподіл Пірсона***
- ***Розподіл Стьюдента***
- ***Розподіл Фішера***

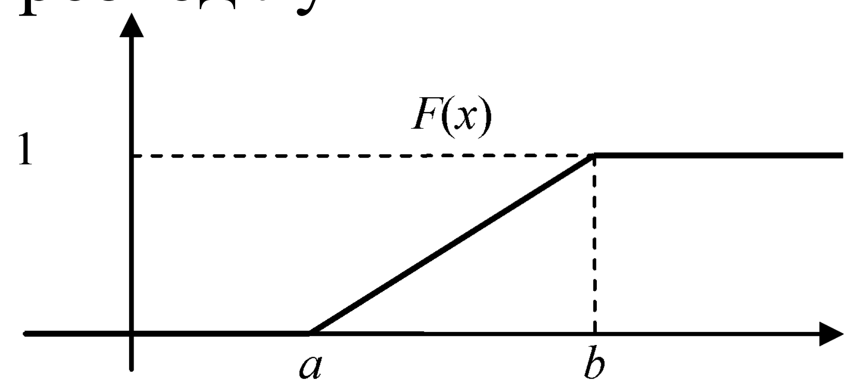
1. Рівномірний розподіл

В інтервалі (a, b) стала щільність розподілу

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

a, b – параметри розподілу



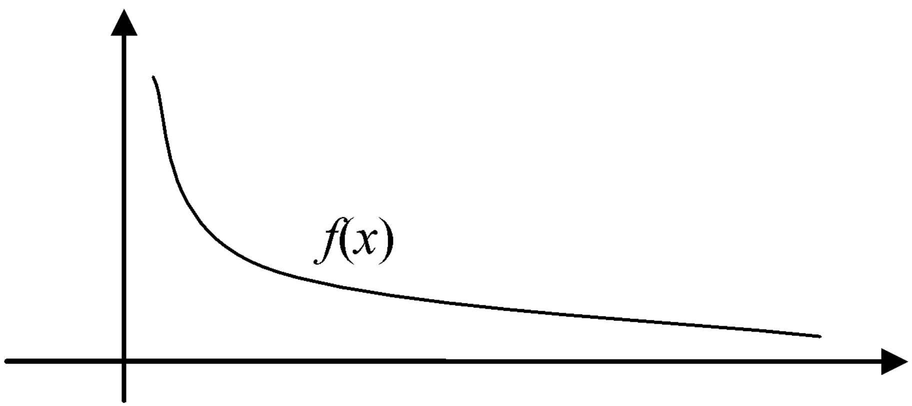


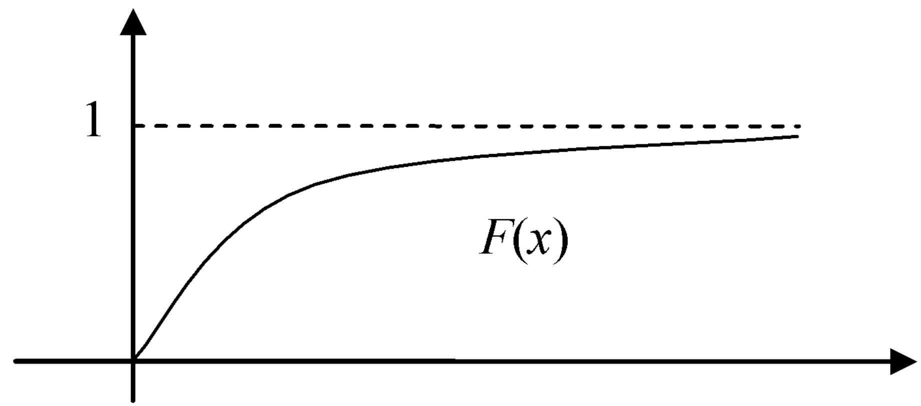
$$M(X)=\frac{a+b}{2}\qquad \text{и} \qquad D(X)=\frac{(b-a)^2}{12}$$

2. Показниковий (експоненціальний) розподіл

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

λ – параметр розподіл



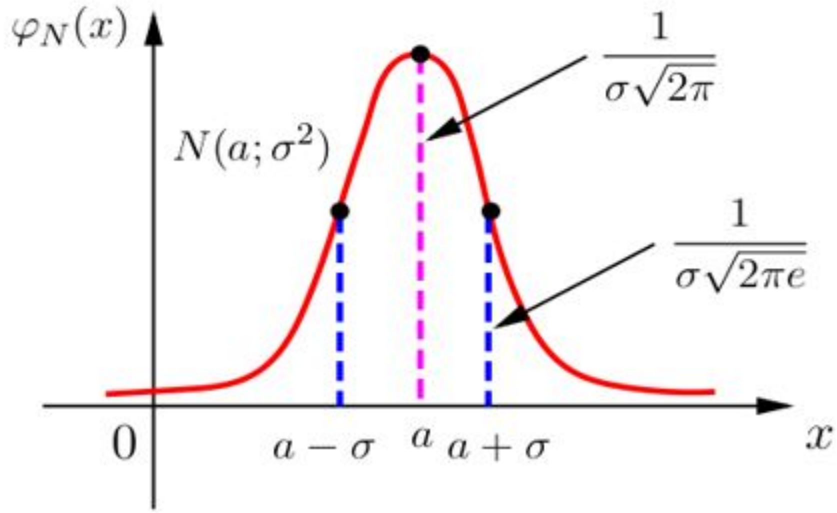


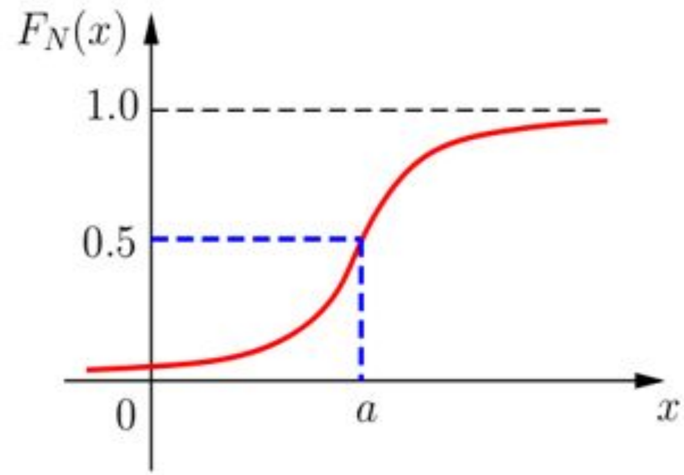
$$M(X) = \frac{1}{\lambda} \quad \text{і} \quad D(X) = \frac{1}{\lambda^2}$$

3. Нормальний розподіл

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-a)^2}{2\sigma^2}} \qquad F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{\frac{-(x-a)^2}{2\sigma^2}}dx$$

a, *σ* – параметри розподілу





$$M(X)=a \quad \text{і} \quad D(X)=\sigma^2$$

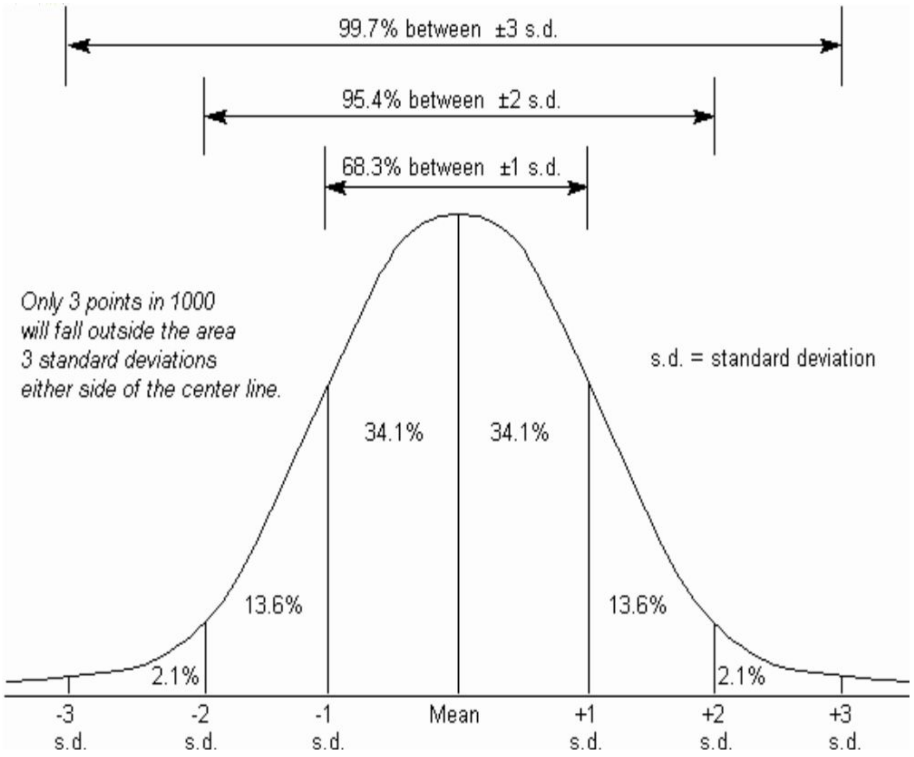

GS9397398N8
DEUTSCHE BUNDESBANK
Banknote
10
ZEHN DEUTSCHE MARK
Deutsche Bundesbank
Frankfurt am Main
1.September 1999
GS9397398N8

Правило 3 сігм

При нормальному розподілі:

- $M(+/-)\sigma=68,26\%$
- $M(+/-)2\sigma=95,44\%$
- $M(+/-)3\sigma=99,72\%$,

$M(+/-)3\sigma$ - інтервал всіх можливих значень



Властивості нормального розподілу

- Правило 3 сігм (99,72% значень лежать в межах M+/-3σ)
- Розподіл симетричний (A=0), ексцес E = 0
- Мода, медіана і математичне сподівання співпадають
- Значення, що лежать на однаковій відстані від M(X), будуть мати однакову імовірність

4. Розподіл Пірсона (χ^2-розподіл)

Нехай незалежні випадкові величини $X_1, X_2, \ldots, X_k$ мають нормальний розподіл, до того ж матаматичне сподівання кожної з них дорівнює 0, а середнє квадратичне відхилення дорівнює 1.

Тоді сума квадратів цих величин:

$$\chi^2 = X_1^2 + X_2^2 + \ldots + X_k^2$$

k – параметр розподілу

має χ^2 – розподіл з k ступенями свободи.

1) випадкова величина $\chi^2 \geq 0$.

2) при збільшення числа ступенів свободи розподіл Пірсона повільно наближається до нормального.

5. Розподіл Стьюдента (t-розподіл)

Нехай X і Y – незалежні випадкові величини.

X має нормальний розподіл з математичним сподіванням, що дорівнює 0, і середнім квадратичним відхиленням, що дорівнює 1.

Y має χ^2 – розподіл з k ступенями свободи.

Тоді величина

$$T = \frac{X}{\sqrt{Y/k}}$$

має розподіл Стьюдента з k ступенями свободи

k – параметр розподілу

При збільшенні числа ступенів свободи розподіл Стьюдента швидко наближається до нормального.

6. Розподіл Фішера (F-розподіл)

Нехай X і Y – незалежні випадкові величини.

X має χ^2– розподіл з k_1 ступенями свободи.

Y має χ^2– розподіл з k_2 ступенями свободи.

Тоді величина

$$F = \frac{X / k_1}{Y / k_2}$$

має розподіл Фішера з k_1 і k_2 ступенями свободи.

k_1 і k_2 – параметри розподілу

Оскільки $X \geq 0$ і $Y \geq 0$, то $F \geq 0$.

Граничні теореми

1. Закон великих чисел.

2. Центральна гранична теорема.

Теорема Чебишева. Якщо дисперсії незалежних випадкових величин $X_1, X_2, \ldots, X_n, \ldots$ не перевищують сталого числа C, то для довільного скільки завгодно малого числа $\varepsilon > 0$ справедлива рівність

$$\lim_{n\to\infty} p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| \le \varepsilon\right) = 1$$

Наслідок з теореми Чебишева. Якщо випадкові величини $X_1, X_2, \ldots, X_n, \ldots$ незалежні і однаково розподілені, з математичним сподіванням a і дисперсією σ^2, то для довільного скольки завгодно малого числа $\varepsilon > 0$ справедлива рівність

$$\lim_{n\to\infty} p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| \le \varepsilon\right) = 1$$

Теорема Бернуллі. Якщо в кожному з n незалежних випробуваннях імовірність p того, що відбудеться подія A стала, то для довільного, скільки завгодно малого числа $\varepsilon > 0$ справедлива рівність

$$\lim_{n\to\infty} p\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) = 1,$$

де m – число появ події A в n випробуваннях.

Теорема Ляпунова. Якщо випадкова величина X дорівнює сумі великої кількість взаємнонезалежних випадкових величин, усі значення якої мають скінченні математичні сподівання і дисперсії, і жодне із значень різко не відрізняється від всіх інших, тобто має незначний вплив на їх суму, то X має розподіл, який близький до нормального.

Теорема. Нехай незалежні випадкові величины X_1, $X_2,\ldots,$ X_n однаково розподілені з математичним сподіванням a і середнім квадратичним відхиленням σ. Тоді випадкова величина

$$\frac{1}{n}\sum_{i=1}^{n} X_i$$

має розподіл, близький до нормального з математичним сподіванням, що дорівнює a і середнім квадратичним відхиленням $\sigma/\sqrt{n}$

Наслідок. Нехай незалежні випадкові величины X_1, $X_2,\ldots,$ X_n однаково розподілені з математичним сподіванням a і середнім квадратичним відхиленням σ . Тоді випадкова величина

$$\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - a}{\sigma / \sqrt{n}}$$

має розподіл, близький до нормального з математичним сподіванням, що дорівнює 0, і середнім квадратичним відхиленням, рівним 1.