Случайные величины. Распределения случайных величин

Тишков Артем Валерьевич, к.ф.-м.н., доцент
Микрюкова Надежда Николаевна

Случайная величина

Случайная величина – это

- числовая переменная, которая принимает свои значения в зависимости от случайных обстоятельств.
- функция, действующая из вероятностного пространства (множество событий) в множество вещественных чисел.

.***Дискретная*** (точечная) ***СВ*** принимает отдельные числовые значения (число студентов в аудитории, игральная кость: 1,2,3,4,5,6)

Непрерывная случайная величина принимает любые значения из некоторого интервала (масса тела, рост студентов), возможно бесконечного.

Случайная величина

Случайные величины будем обозначать заглавными последними буквами латинского алфавита:**X,Y,Z…**,а их возможные значения прописными буквами: $X\ \{x_1, x_2, \ldots, x_n\}$, $Y\ \{y_1, y_2, \ldots, y_m\}$

Любое правило, которое устанавливает связь между возможными значениями случайной величины и вероятностями, с которыми она эти значения принимает, называется ***законом распределения случайной величины***.

Закон распределения СВ можно задавать ***в виде***: 1) таблицы, 2) графика, 3) Функции распределения.

Закон распределения случайной величины

Любое правило, которое устанавливает связь между возможными значениями случайной величины и вероятностями, с которыми она эти значения принимает, называется ***законом распределения случайной величины***.

Закон распределения случайной величины можно задавать ***в виде***:

1) Таблицы
2) Графика
3) Функции распределения.

Дискретная СВ. Таблица распределения

Ряд распределения(может быть конечным или бесконечным)

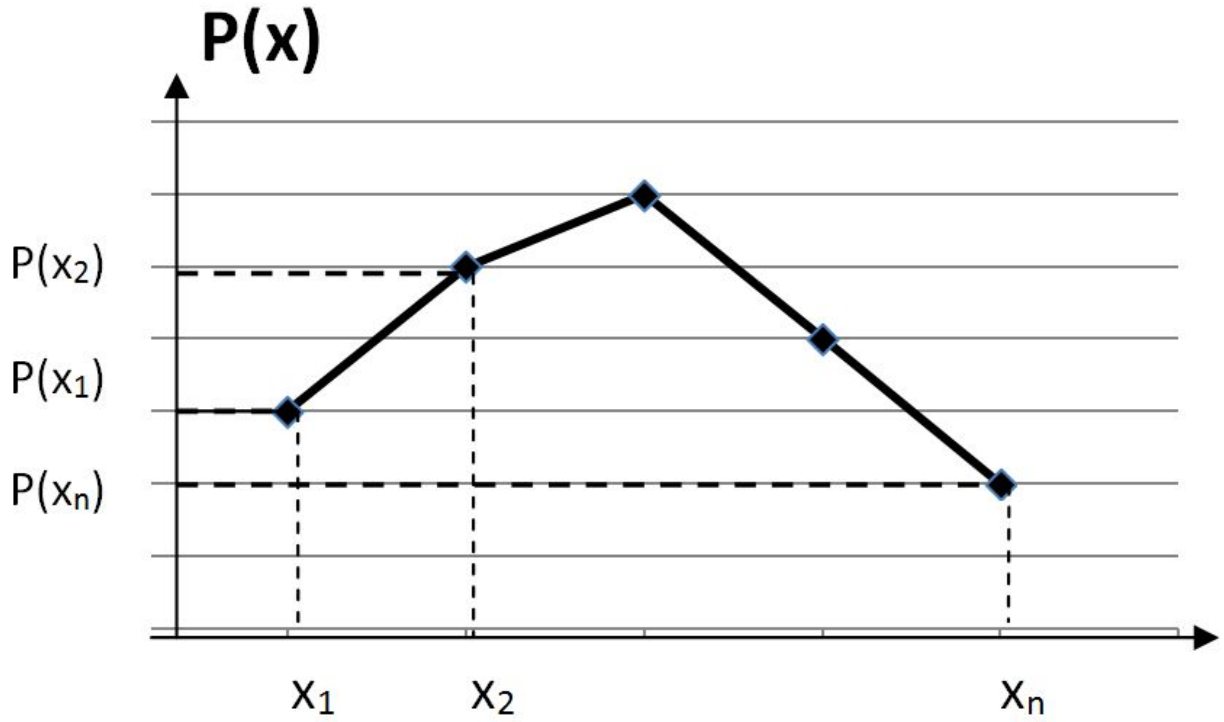
X	x_1	x_2	…	…	…	x_n
P(x)	$P(x_1)$	$P(x_2)$				$P(x_n)$

Так как события $X=x_1$, $X=x_2$…. попарно несовместны и составляют полную группу событий, следовательно

$$P(x_1)+P(x_2)+\cdots+P(x_n)=\sum_{i=1}^{n}P(x_i)=1: \quad \text{Условие нормировки}$$

Дискретная СВ. График распределения

***График**: многоугольник распределения.*

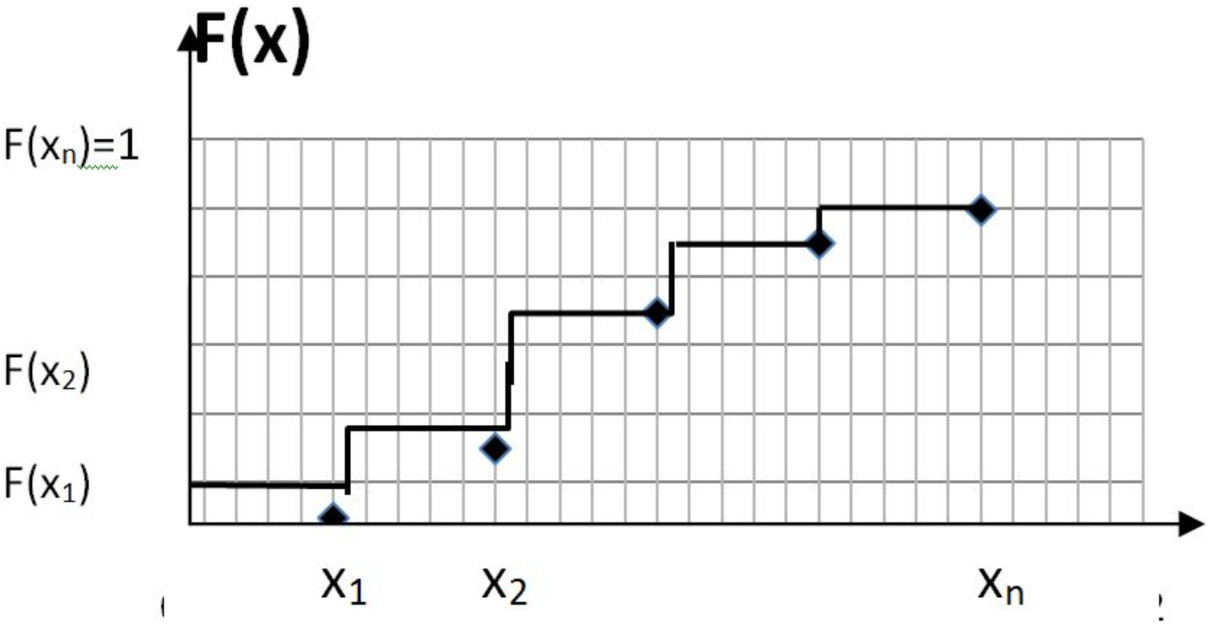


Дискретная СВ. Функция распределения

Функция распределения **$F(x_0)$–** это вероятность того, что случайная величина X принимает значения меньшие или равные x_0.

$$F(x_0) = P(X \le x_0)$$

$$F(x_0) = \sum_{i/x_i \le x_0} P(x_i)$$



1). F(x) неубывающая: $F(x_2) \geq F(x_1)$ если $x_2 \geq x_1$

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

$$\boxed{P(a \leq X \leq b) = F(b) - F(a)}$$

Пример

X	2	4	6	8	10
P(x)	0,1	0,2	0,4	0,2	0,1
F(x)	0,1	0,3	0,7	0,9	1

$F(4)=P(X\leq 4)=P(2)+P(4)=0{,}1+0{,}2=0{,}3$

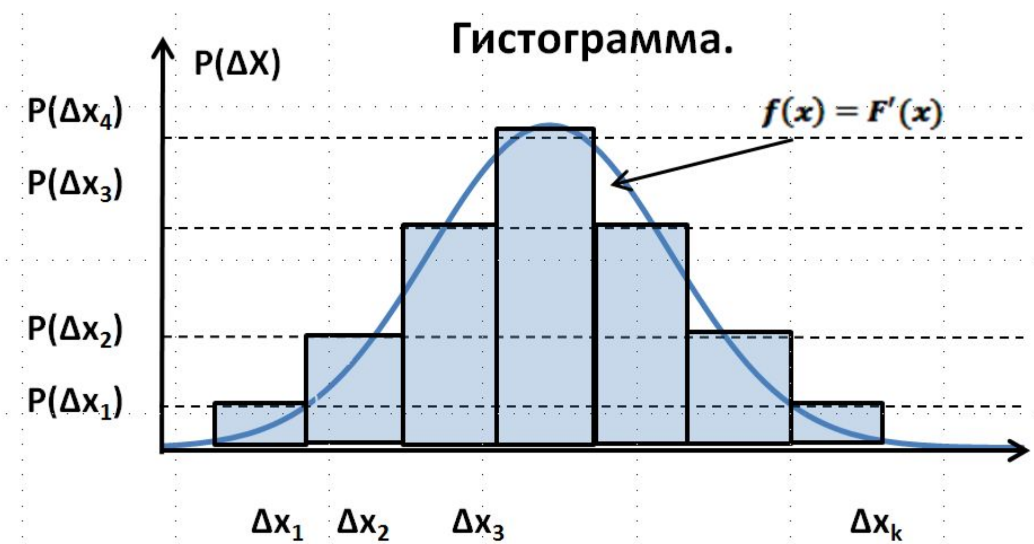
$F(8)=P(X\leq 8)=P(2)+P(4)+P(6)+P(8)=0{,}1+0{,}2+0{,}4+0{,}2=0{,}9$

$P(4<X\leq 8)=F(8)-F(4)=0{,}9-0{,}3=0{,}6$

Непрерывная случайная величина

Таблица: *Интервальный ряд распределения.*

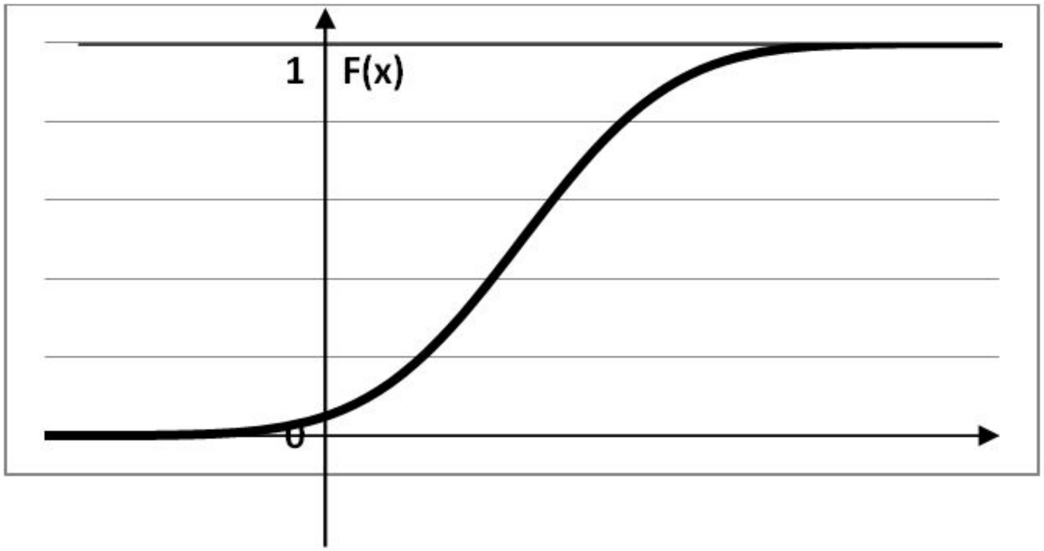
X	Δx_1	Δx_2				Δx_k
$P(\Delta x)$	$P(\Delta x_1)$	$P(\Delta x_2)$				$P(\Delta x_k)$



Непрерывная случайная величина

Функция распределения

$$\boxed{F(x_0) = P(X \le x_0)}$$



1) F(x) неубывающая:
$F(x_2) \geq F(x_1)$ если $x_2 \geq x_1$

2) $F(-\infty)=0$; $F(+\infty)=1$

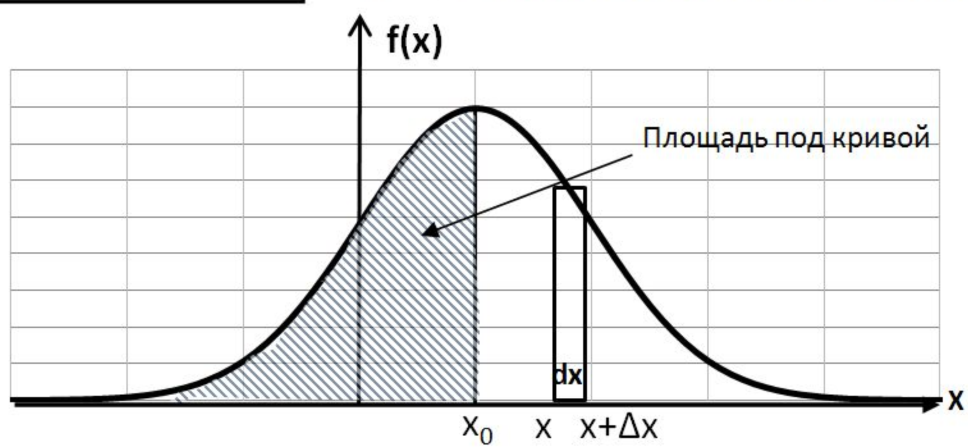
Непрерывная случайная величина

Функция плотности распределения ***f(x):*** (только для непрерывной случайной величины).

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

Найдём предел: $\boxed{\lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = F'(x)}$

Обозначим: $\boxed{f(x) = F^{'}(x)}$ это *функция плотности распределения*.



Функция плотности распределения

f(x) неотрицательная функция (f(x)≥0)

Вероятность попадания в элементарный интервал <u>*dx=(x+Δx)-x равна f(x)dx=dP.*</u>

$$P(x < X \leq x + \Delta x) = F(x + \Delta x) - F(x) = \Delta F(x) = F'(x) \cdot dx = f(x)dx = \boldsymbol{dP} -$$

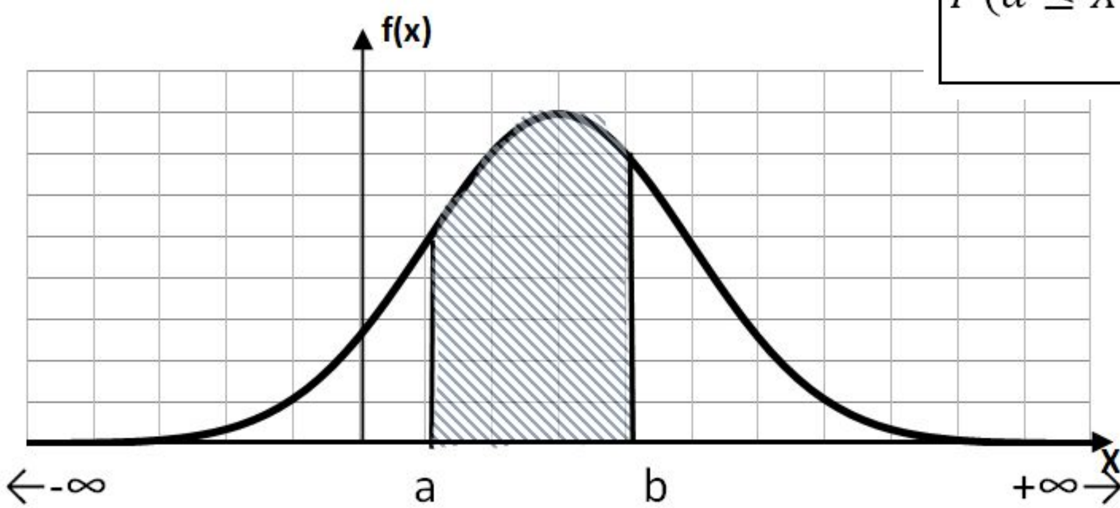
−элемент вероятности.

$$\boxed{P(X \leq x_0) = F(x_0) = \int_{-\infty}^{x_0} f(x)dx}$$

Функция плотности распределения

Вероятность попадания случайной величины в интервал [a,b]:

$$P(a \le X \le b) = \int_a^b f(x)\,dx = F(b) - F(a)$$



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

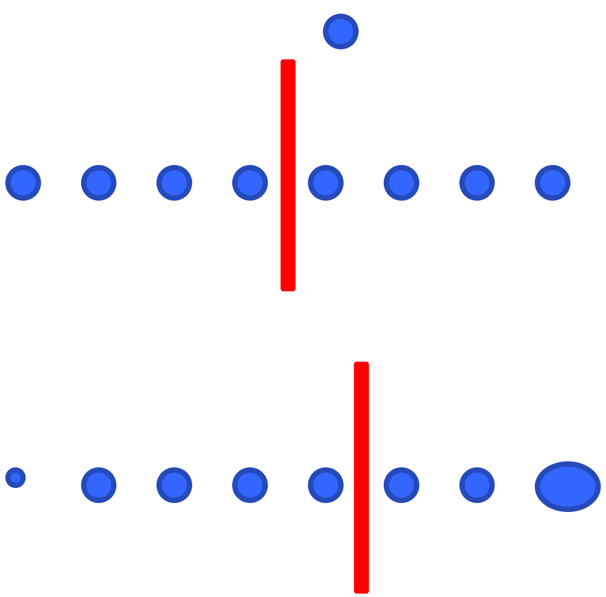
Числовые характеристики (параметры) случайной величины

1) *Математическое ожидание*
2) *Дисперсия (рассеивание)*
3) *Средне-квадратическое или стандартное отклонение*

Математическое ожидание

<u>*Дискретная случайная величина*</u>

$$M[X] = \sum_{i=1}^{k} x_i \cdot P(x_i)$$



<u>*Непрерывная случайная величина*</u>

$$M[X] = \int_{-\infty}^{+\infty} x \cdot f(x)dx$$

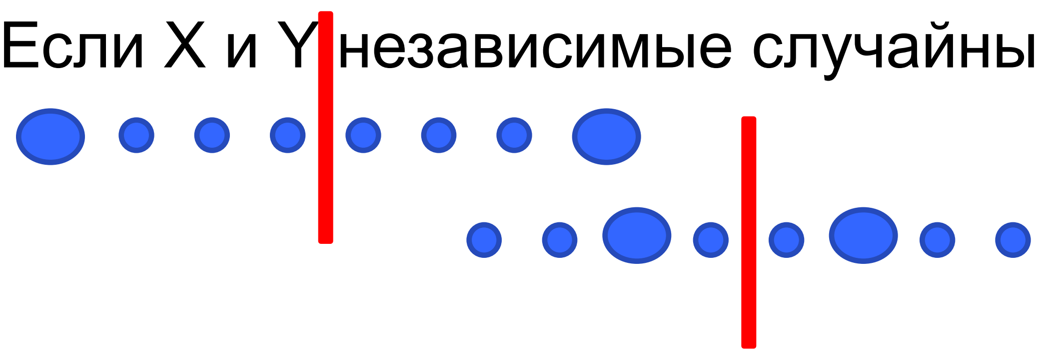
Дисперсия (рассеивание)

это математическое ожидание (среднее значение) квадрата отклонения случайной величины X от её математического ожидания.

$$D[X] = \sum_{i=1}^{k} (x_i - M[X])^2 \cdot P(x_i)$$

$$\boxed{D[X] = M[X^2] - (M[X])^2}$$

Если X и Y независимые случайные величины, то



$$D[X+Y] = D[X] + D[Y]$$

Непрерывная случайная величина:

$$D[X] = \int_{-\infty}^{+\infty} (x - M[X])^2 \cdot f(x)dx$$

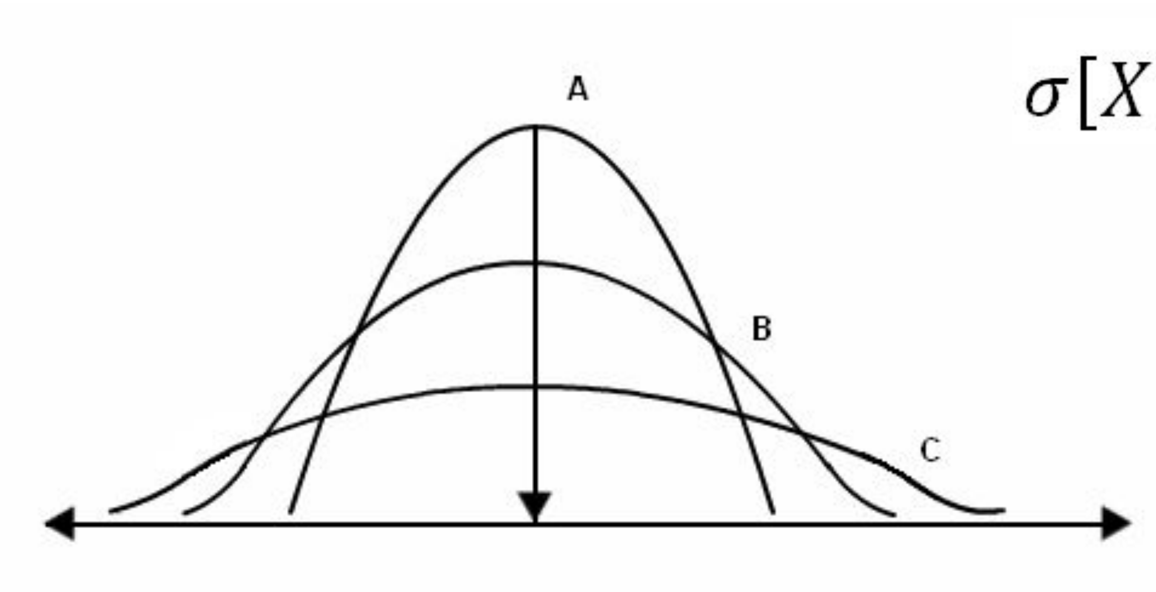
Равномерное или прямоугольное распределение

Случайная величина называется равномерно распределённой на интервале [c,d], если функция плотности распределения её на этом интервале постоянна, а вне него равна нулю

$$f(x)=\begin{cases} const, & \text{если } c \le x \le d \\ 0 & \text{если } x<c,\; x>d \end{cases}$$

Стандартное отклонение

Средне-квадратическое или стандартное отклонение:



$$\sigma[X] = \sqrt{D[X]}$$

Равномерное распределение. Чему равна константа

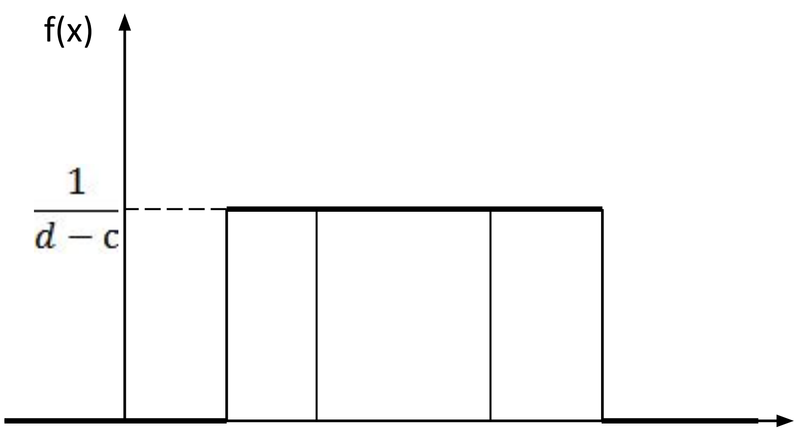
Из условия нормировки $\int_{-\infty}^{+\infty} f(x)dx = 1$

$$\int_c^d f(x)dx = \int_c^d const\,dx = const\int_c^d dx = const\cdot x\Big|_c^d = const\cdot(d-c) = 1$$

$$\Rightarrow \quad const = \frac{1}{d-c}$$

Равномерное распределение. Вероятность попадания в интервал

$$P(a \le x \le b) = \int_a^b \frac{1}{d-c}dx = \frac{1}{d-c} \cdot x\Big|_a^b = \frac{b-a}{d-c}$$



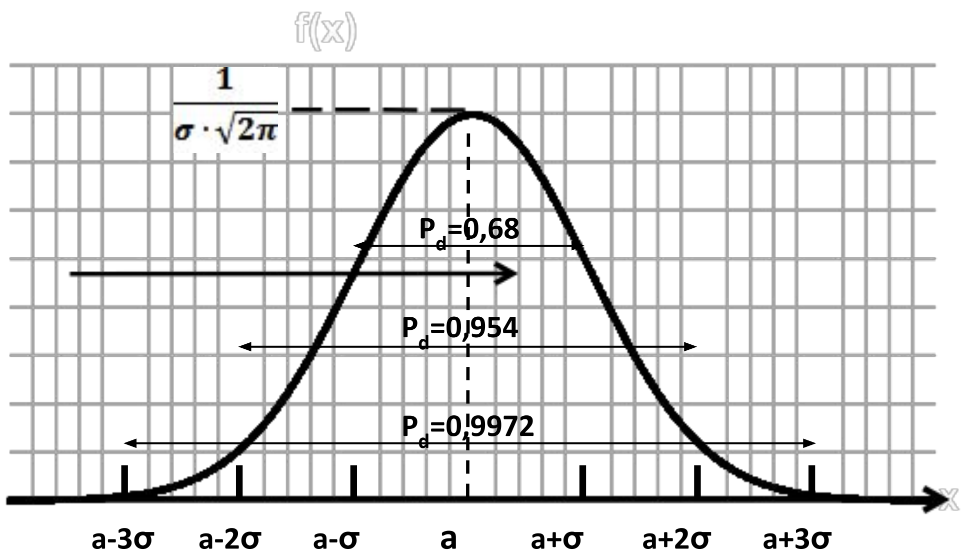
Нормальное распределение или распределение Гаусса

Случайная величина распределена по *нормальному закону*, если функция плотности её распределения имеет вид:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{\frac{-(x-a)^2}{2\cdot\sigma^2}}$$

где a,σ – параметры распределения.

Нормальное распределение. График плотности распределения



Кривая симметрична относительно прямой x=**a**

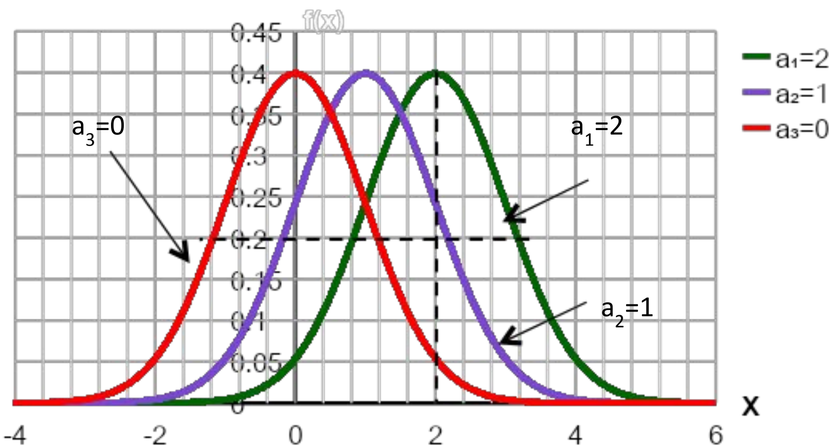
$$f_{\max}(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}}$$

достигается в этой же точке x=**a**

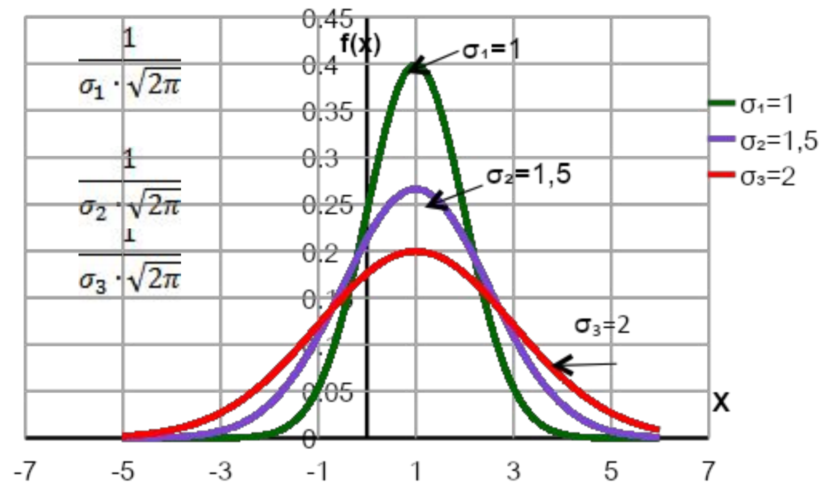
На графике представлены вероятности попадания в интервалы *среднее значение* плюс-минус *одна, две и три сигмы*

Нормальное распределение. Примеры графиков плотности распределения

Графики плотности распределения с разными значениями параметра a. (σ=1)



Графики плотности распределения с разными значениями параметра σ . ($\sigma_1<\sigma_2<\sigma_3$, a=1)



Нормальное распределение. Математическое ожидание и дисперсия

Математическое ожидание н.р. равно *a*:

$$M[X]=\int_{-\infty}^{+\infty} x\cdot f(x)dx=\frac{1}{\sigma\cdot\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\cdot e^{\frac{-(x-a)^2}{2\cdot\sigma^2}}dx=a$$

Дисперсия н.р. равна σ^2

$$D[x]=\int_{-\infty}^{+\infty}(x-M[x])^2\cdot f(x)dx=\frac{1}{\sigma\cdot\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2\cdot e^{\frac{-(x-a)^2}{2\cdot\sigma^2}}dx=\sigma^2$$

Величину σ называют среднеквадратичным отклонением

$$\boldsymbol{\sigma}=\sqrt{\boldsymbol{D}[\boldsymbol{X}]}=\boldsymbol{\sigma}[\mathrm{X}]$$

Нормальное распределение. Нормированная случайная величина

Введем замену переменной $t = \dfrac{x-a}{\sigma} = \dfrac{x-M[X]}{\sigma[X]}$

t – безразмерная случайная величина. Важные свойства: M[t]=0 D[t]=1 σ[t] =1

Так как 99,7% всех значений случайной величины X отличаются от M[X] не больше, чем на **3·σ[X],** следовательно для любого значения *x* получим:

$$-3 \le t = \frac{x-M[x]}{\sigma[x]} \le 3$$

с вероятностью P=0,997.

Нормальное распределение. Нормальная функция распределения

Функция распределения н.р.

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

Введем замену переменной $t = \frac{x-a}{\sigma} = \frac{x - M[X]}{\sigma[X]}$ $\quad dt = \frac{1}{\sigma} dx$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-t^2}{2}} dt = \Phi(t) = \Phi\left(\frac{x - M[X]}{\sigma[X]}\right)$$

Φ(t) называется функцией Гаусса нормальной функцией распределения

Значения функции Ф(t) для $0 \le t \le 3$

t	Ф(t)	t	Ф(t)	t	Ф(t)
0	0,5	1,1	0,864334	2,2	0,986097
0,1	0,539828	1,2	0,88493	2,3	0,989276
0,2	0,57926	1,3	0,9032	2,4	0,991802
0,3	0,617911	1,4	0,919243	2,5	0,99379
0,4	0,655422	1,5	0,933193	2,6	0,995339
0,5	0,691462	1,6	0,945201	2,7	0,996533
0,6	0,725747	1,7	0,955435	2,8	0,997445
0,7	0,758036	1,8	0,96407	2,9	0,998134
0,8	0,788145	1,9	0,971283	3	0,99865
0,9	0,81594	2	0,97725	3,1	0,999032
1	0,841345	2,1	0,982136	3,2	0,999313

$$\boxed{\Phi(0)=0{,}5; \quad \Phi(-\infty)=0; \quad \Phi(+\infty)=1; \quad \Phi(-t)=1-\Phi(t)}$$

Вероятность попадания значений н.р. случайной величины в интервал

Интервал [a;b]

$$P(a \le x \le b) = F(b) - F(a) = \Phi\left(\frac{b - M[X]}{\sigma[X]}\right) - \Phi\left(\frac{a - M[X]}{\sigma[X]}\right)$$

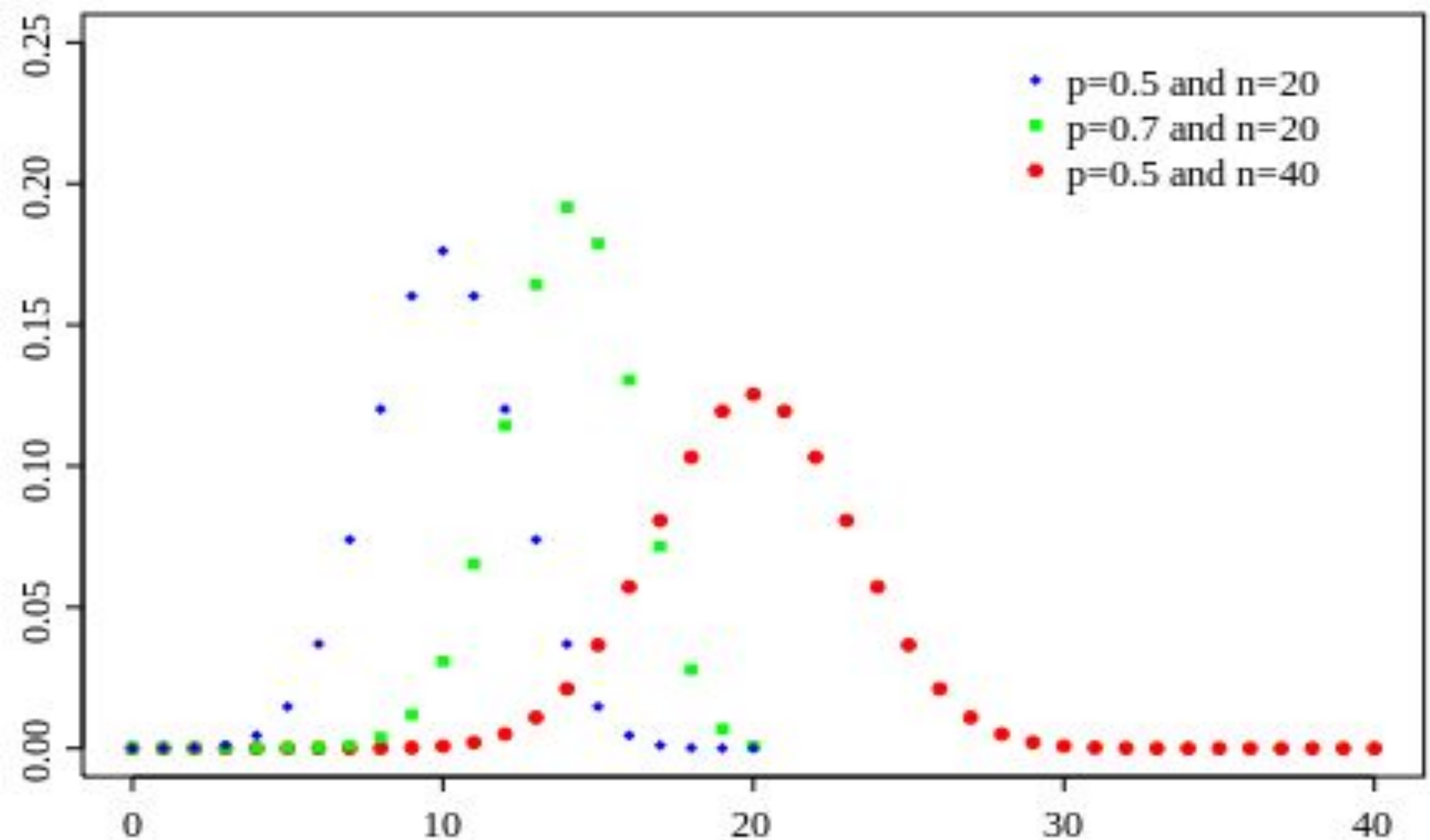
$$P(\mathrm{M[X]} - \sigma[\mathrm{X}] \le x \le \mathrm{M[X]} + \sigma[\mathrm{X}]) = \Phi\left(\frac{\mathrm{M[X]} + \sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) - \Phi\left(\frac{\mathrm{M[X]} - \sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) =$$
$$= \Phi(+1) - \Phi(-1) = \Phi(+1) - \big(1 - \Phi(+1)\big) = 2 \cdot \Phi(+1) - 1 = 2 \cdot 0{,}84 - 1 = \mathbf{0,68}$$

$$P(\mathrm{M[X]} - 2\sigma[\mathrm{X}] \le x \le \mathrm{M[X]} + 2\sigma[\mathrm{X}]) = \Phi\left(\frac{\mathrm{M[X]} + 2\sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) - \Phi\left(\frac{\mathrm{M[X]} - 2\sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) =$$
$$= \Phi(+2) - \Phi(-2) = \Phi(+2) - \big(1 - \Phi(+2)\big) = 2 \cdot \Phi(+2) - 1 = 2 \cdot 0{,}977 - 1 = \mathbf{0{,}954}$$

Правило трёх сигм:

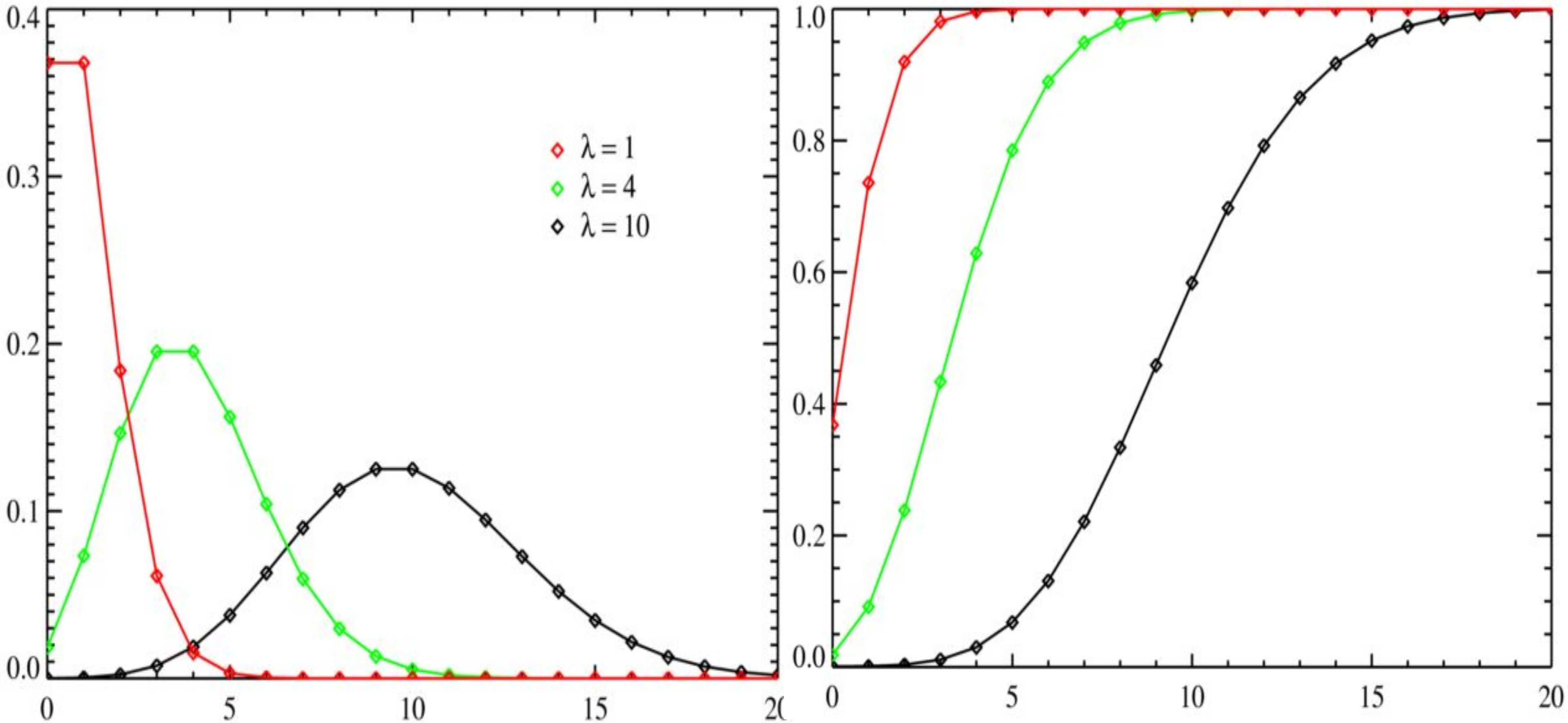
$$P(\mathrm{M[X]} - 3\sigma[\mathrm{X}] \le x \le \mathrm{M[X]} + 3\sigma[\mathrm{X}]) = \Phi\left(\frac{\mathrm{M[X]} + 3\sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) - \Phi\left(\frac{\mathrm{M[X]} - 3\sigma[\mathrm{X}] - M[X]}{\sigma[X]}\right) =$$
$$= \Phi 3 - \Phi(-3) = \Phi(+3) - \big(1 - \Phi(+3)\big) = 2 \cdot \Phi(+3) - 1 = 2 \cdot 0{,}9986 - 1 = \mathbf{0{,}9972}$$

Биномиальное распределение



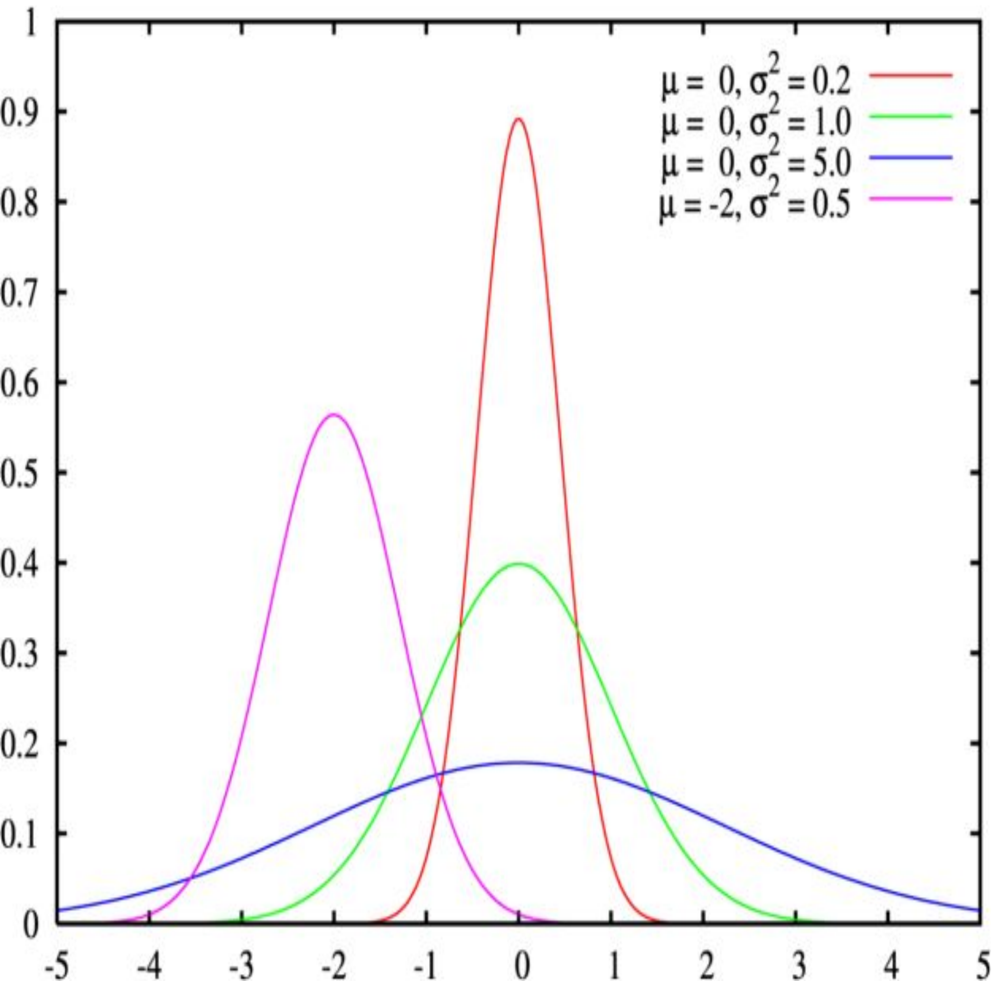
Обозначение	B (n, p)
Параметры	$n \geq 0$ — число «испытаний»
	$0 \leq p \leq 1$ - вероятность «успеха»
Носитель	$k \in \{0, \ldots, n\}$
Функция плотности распределения	$\binom{n}{k} p^k q^{n-k}$
Функция распределения	I_{1-p}(n - [k], 1+ [k])
Математическое ожидание	np
Дисперсия	npq

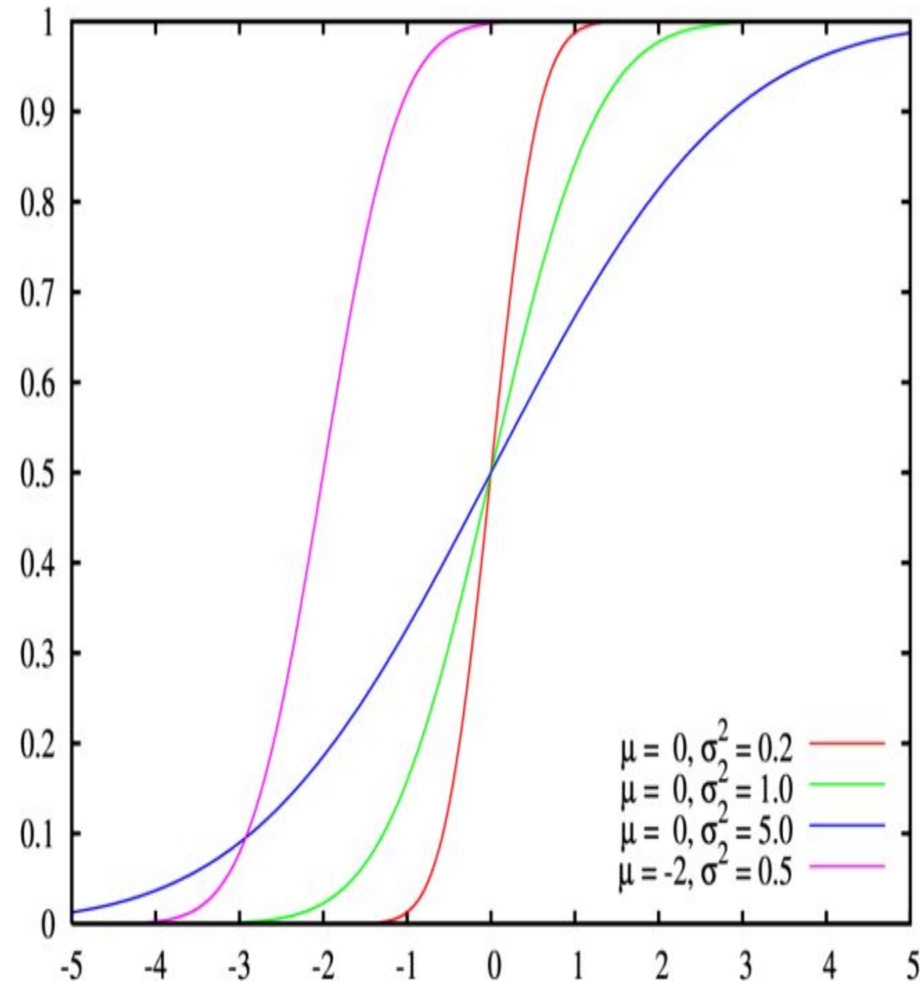
Распределение Пуассона



Обозначение	$P(\lambda)$
Параметры	$\lambda \in (0, \infty)$
Носитель	$\in$
Функция вероятности	$\dfrac{e^{-\lambda}\lambda^{k}}{k!}$
Функция распределения	$\dfrac{\Gamma(k+1,\lambda)}{k!}$
Математическое ожидание	λ
Дисперсия	λ

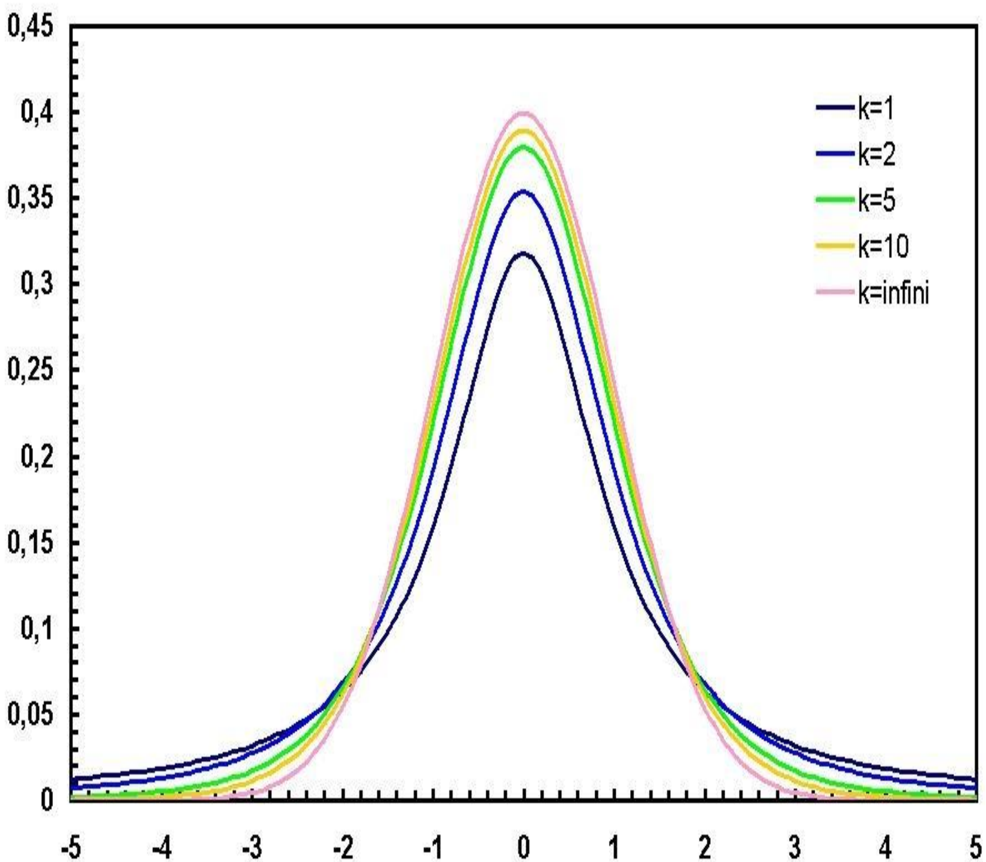
Распределение Гаусса

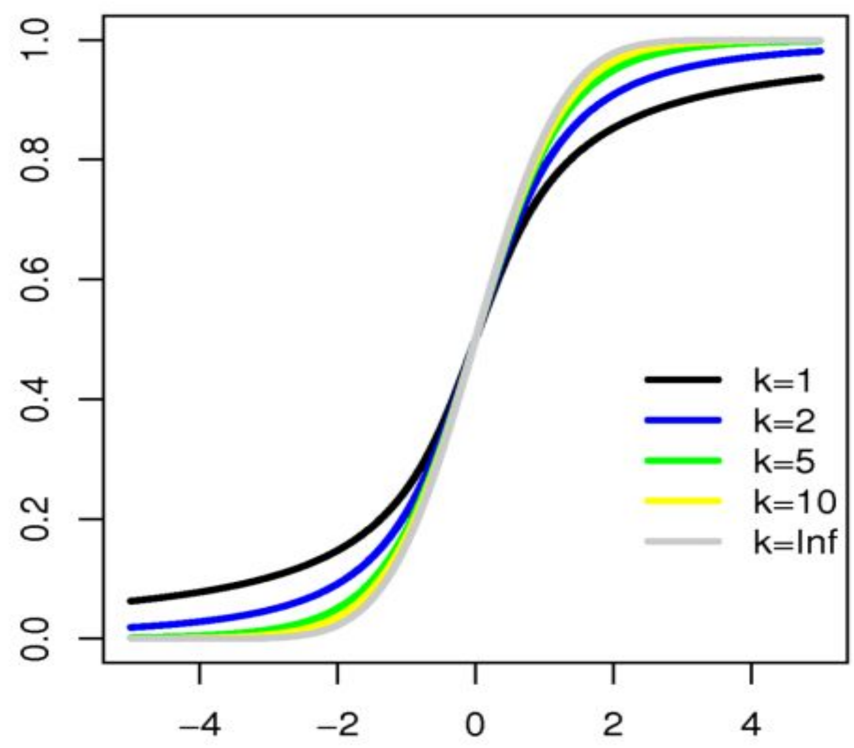




обозначение	N (μ,σ^2)
Параметры	μ - коэффициент сдвига(вещественное число) σ >0 - коэффициент масштаба(вещественный, строго положительный)
Носитель	$x \in (-\infty;+\infty)$
Плотность вероятности	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Функция распределения	$\Phi\left(\frac{x-\mu}{\sigma}\right)$
Математическое ожидание	μ
Дисперсия	σ^2

Распределение Стьюдента





Обозначение	t (n)
Параметры	n>0 — число степеней свободы
Носитель	$x \in \left(-\infty; +\infty\right)$
Плотность вероятности	$\dfrac{\Gamma((n+1)/2)}{\sqrt{n\pi}\,\Gamma(n/2)(1+x^2/n)^{(n+1)/2}}$
Функция распределения	$\dfrac{1}{2} + x\,\dfrac{\Gamma((n+1)/2)}{\sqrt{\pi n}\,\Gamma(n/2)} \cdot \dfrac{{}_2F_1(\frac{1}{2},(n+1)/2;\frac{3}{2};-\frac{x^2}{n})}{\sqrt{\pi n}\,\Gamma(n/2)}$

где ${}_2F_1$ — гипергеометрическая функция

Математическое ожидание — 0, если n>1

Дисперсия — $\dfrac{n}{n-2}$, если n>2