Гипергеометрическое распределение

Пример. Пусть в партии из N изделий имеется M стандартных ($M < N$). Из партии случайно отбираются n изделий.

Гипергеометрическое распределение

Пример. Пусть в партии из N изделий имеется M стандартных ($M < N$). Из партии случайно отбираются n изделий.

Обозначим через X случайную величину – число m стандартных изделий среди n отобранных;

$$X \in \left\{0, 1, 2, ..., \min\{M, n\}\right\}.$$

Гипергеометрическое распределение

Пример. Пусть в партии из N изделий имеется M стандартных ($M < N$). Из партии случайно отбираются n изделий.

Обозначим через X случайную величину – число m стандартных изделий среди n отобранных;

$$X \in \left\{0,1,2,\ldots,\min\{M,n\}\right\}.$$

Найдем вероятность события A: «среди n отобранных изделий m стандартных».

Гипергеометрическое распределение

Пример. Пусть в партии из N изделий имеется M стандартных ($M < N$). Из партии случайно отбираются n изделий.

Обозначим через X случайную величину – число m стандартных изделий среди n отобранных;

$$X \in \{0, 1, 2, \ldots, \min\{M, n\}\}.$$

Найдем вероятность события A: «среди n отобранных изделий m стандартных».

Общее число элементарных исходов - C_N^n.

Гипергеометрическое распределение

Пример. Пусть в партии из N изделий имеется M стандартных ($M < N$). Из партии случайно отбираются n изделий.

Обозначим через X случайную величину – число m стандартных изделий среди n отобранных;

$$X \in \{0, 1, 2, \ldots, \min\{M, n\}\}.$$

Найдем вероятность события A: «среди n отобранных изделий m стандартных».

Общее число элементарных исходов - C_N^n.

Число исходов, благоприятствующих событию $X = m$:
m стандартных изделий можно извлечь из M стандартных изделий C_M^m способами, при этом остальные $n - m$ нестандартных изделий из $N - M$ нестандартных изделий можно выбрать C_{N-M}^{n-m} способами.

Следовательно, число благоприятствующих исходов равно $C_M^m \cdot C_{N-M}^{n-m}$.

Следовательно, число благоприятствующих исходов равно $C_M^m \cdot C_{N-M}^{n-m}$.

Искомая вероятность равна

$$p_m = P\left(X = m\right) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$$

Следовательно, число благоприятствующих исходов равно $C_M^m \cdot C_{N-M}^{n-m}$.

Искомая вероятность равна

$$p_m = P(X = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$$

Распределение вероятностей, определяемой этой формулой при различных значениях $m = 0, 1, 2, \ldots, \min(M, n)$, называют **гипергеометрическим**

$$\boxed{X \sim HG(M, N, n)}.$$

Следовательно, число благоприятствующих исходов равно $C_M^m \cdot C_{N-M}^{n-m}$.

Искомая вероятность равна

$$p_m = P(X = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$$

Распределение вероятностей, определяемой этой формулой при различных значениях $m = 0, 1, 2, \ldots, \min(M, n)$, называют **гипергеометрическим**

$$\boxed{X \sim HG(M, N, n)}.$$

Гипергеометрическое распределение моделирует количество удачных выборок без возвращения из конечной совокупности

В случае выбора из большой генеральной совокупности биномиальное распределение $Bin(n, \frac{M}{N})$ более удобно, чем гипергеометрическое.

В случае выбора из большой генеральной совокупности биномиальное распределение $Bin(n, \frac{M}{N})$ более удобно, чем гипергеометрическое.

Гипергеометрическое распределение более корректно для выборок без возврата (в этом случае биномиальное распределение применять нельзя, так как вероятность успеха p меняется в ходе эксперимента и зависит от предыдущих исходов).

В случае выбора из большой генеральной совокупности биномиальное распределение $Bin(n, \frac{M}{N})$ более удобно, чем гипергеометрическое.

Гипергеометрическое распределение более корректно для выборок без возврата (в этом случае биномиальное распределение применять нельзя, так как вероятность успеха p меняется в ходе эксперимента и зависит от предыдущих исходов).

При достаточно большом значении и малом объеме выборки (когда $\frac{n}{N} \leq 0,1$) гипергеометрическое распределение приблизительно совпадает с биномиальным.

Абсолютно непрерывные функции распределения

Определение. Функция распределение $F(x)$ называется абсолютно непрерывной, если существует функция $f(u)$, такая что для любого x

$$F(x) = \int_{-\infty}^{x} f(u)\,du.$$

Абсолютно непрерывные функции распределения

Определение. Функция распределение $F(x)$ называется абсолютно непрерывной, если существует функция $f(u)$, такая что для любого x

$$F(x) = \int_{-\infty}^{x} f(u)\,du.$$

Функция $f(u)$ называется плотностью распределения случайной величины X.

Абсолютно непрерывные функции распределения

Определение. Функция распределение $F(x)$ называется абсолютно непрерывной, если существует функция $f(u)$, такая что для любого x

$$F(x)=\int_{-\infty}^{x} f(u)\,du.$$

Функция $f(u)$ называется плотностью распределения случайной величины X.

Для краткости будем называть случайную величину с абсолютно непрерывной функцией распределения непрерывной случайной величиной.

Свойства плотности распределения

1. Плотность есть производная от функции распределения: $f(x) = F'(x)$.

Свойства плотности распределения

1.Плотность есть производная от функции распределения: $f(x) = F'(x)$.

2.Плотность распределения неотрицательна: $f(x) \geq 0$.

Справедливость утверждения следует из определения: $F(x)$ – функция неубывающая, а $f(x)$ – ее производная, поэтому она неотрицательна.

Свойства плотности распределения

1.Плотность есть производная от функции распределения: $f(x) = F'(x)$.

2.Плотность распределения неотрицательна: $f(x) \geq 0$.

Справедливость утверждения следует из определения: $F(x)$ – функция неубывающая, а $f(x)$ – ее производная, поэтому она неотрицательна.

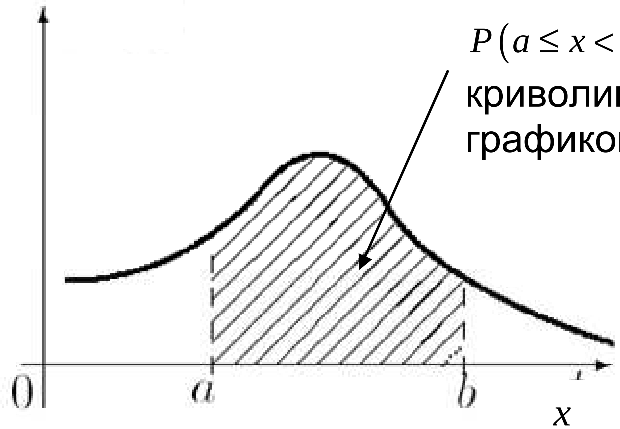
3. Интеграл от плотности распределения случайной величины по области ее определения равен 1:

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

Свойство следует из определения абсолютно непрерывной функции (при $x \to \infty$).

4. Вероятность попадания непрерывной случайной величины в заданный интервал $[a,b)$ равна

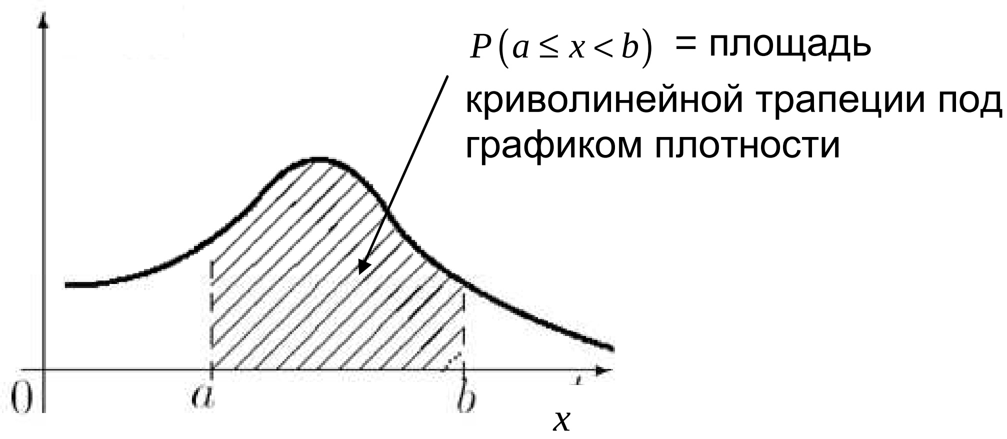
$$P\left(a \le x < b\right) = \int_a^b f\left(x\right)dx.$$



$P\left(a \le x < b\right)$ = площадь криволинейной трапеции под графиком плотности

4. Вероятность попадания непрерывной случайной величины в заданный интервал $[a, b)$ равна

$$P(a \le x < b) = \int_a^b f(x)\,dx.$$



Замечание 1. Для непрерывных случайных величин $P(a \le x < b) = P(a < x < b) = P(a < x \le b) = P(a \le x \le b)$.

Пример. Задана плотность вероятности случайной величины X: $f(x)=\begin{cases} 0 & при \quad x\le 0 \\ 2x & при \ 0<x\le 1 \\ 0 & при \quad x>1 \end{cases}$.

Найти вероятность того, что в результате испытания величина X примет значение, принадлежащее интервалу $(0.5,\ 1)$.

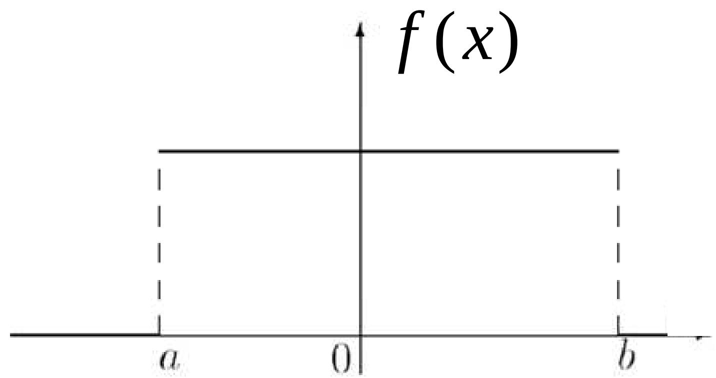
Решение.

$$P(0.5< X <1)=2\int_{0.5}^{1} xdx = x^2\Big|_{0.5}^{1} =1-0.25=0.75.$$

Равномерное распределение

Если плотность вероятности задается формулой

$$f(x)=\begin{cases} 0, & x<a,\ x>b, \\ \dfrac{1}{b-a}, & x\in[a,b] \end{cases},$$

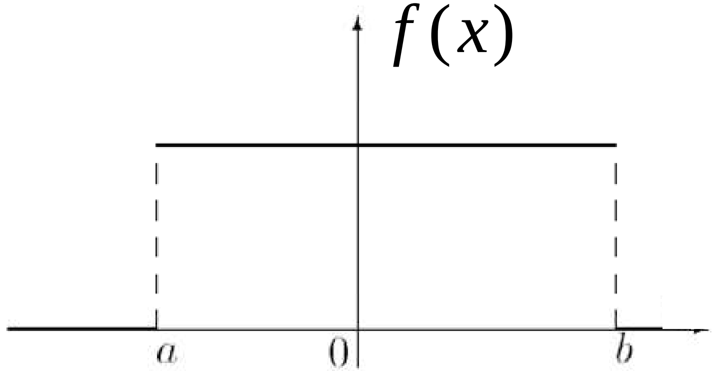


то такое распределение вероятностей называют равномерным на отрезке $[a,b]$.

Равномерное распределение

Если плотность вероятности задается формулой

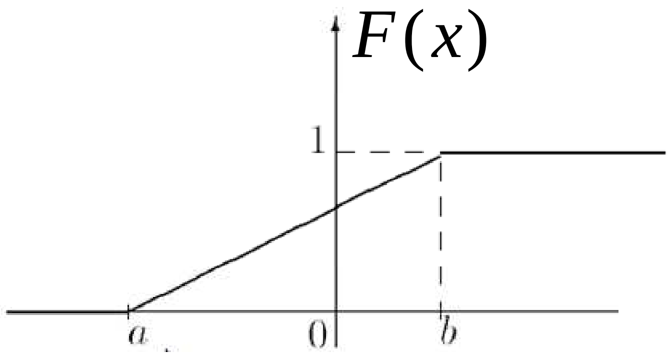
$$f(x)=\begin{cases}0, & x<a,\ x>b,\\ \dfrac{1}{b-a}, & x\in[a,b]\end{cases},$$



то такое распределение вероятностей называют равномерным на отрезке $[a,b]$.

Функция распределения при этом законе имеет вид:

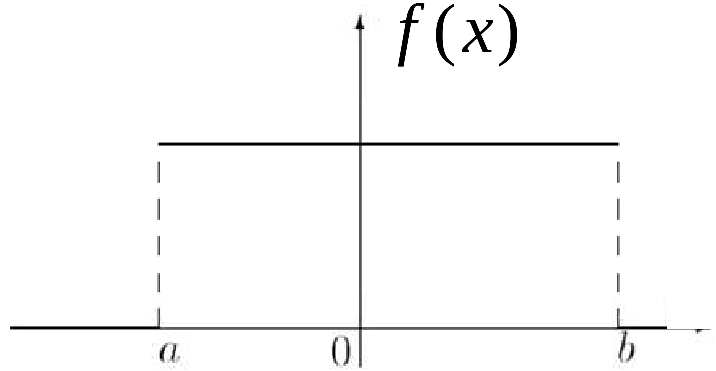
$$F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & x\in[a,b],\\ 1, & x>b.\end{cases}$$



Равномерное распределение

Если плотность вероятности задается формулой

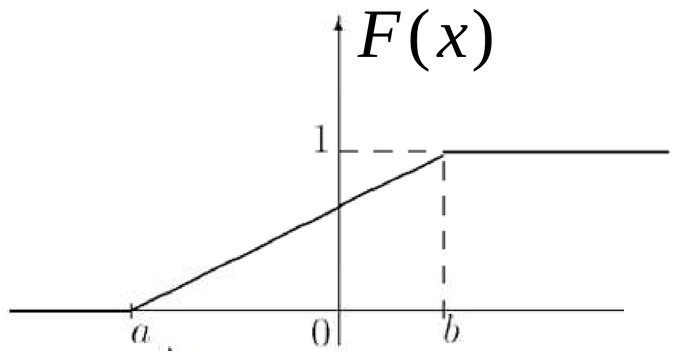
$$f(x)=\begin{cases}0, & x<a,\ x>b,\\ \dfrac{1}{b-a}, & x\in[a,b]\end{cases},$$



то такое распределение вероятностей называют равномерным на отрезке $[a,b]$.

Функция распределения при этом законе имеет вид:

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



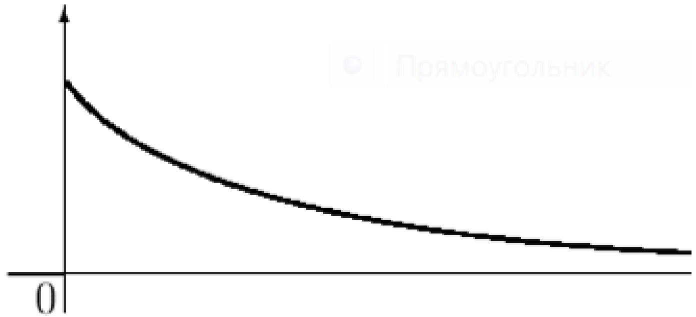
Обозначение: $X\sim U(a,b)$.

Равномерное распределение применяется в случае, когда нет информации о том, что одни значения случайной величины являются более вероятными, чем другие.

Экспоненциальное распределение

Плотность экспоненциального (показательного), распределения задается как

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0,\end{cases}$$

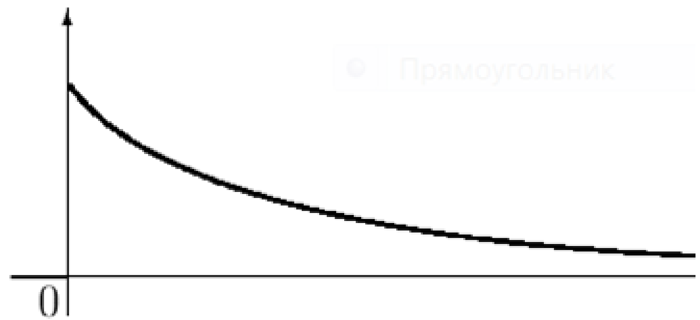


$\lambda > 0$ - параметр распределения.

Экспоненциальное распределение

Плотность экспоненциального (показательного), распределения задается как

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0,\end{cases}$$



$\lambda > 0$ - параметр распределения.

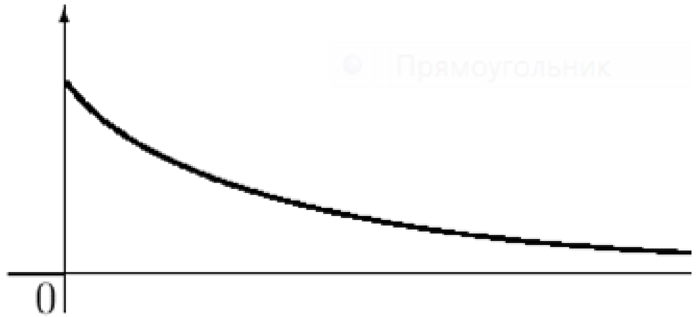
Функция распределения:

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

Экспоненциальное распределение

Плотность экспоненциального (показательного), распределения задается как

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0,\end{cases}$$



$\lambda > 0$ - параметр распределения.

Функция распределения:

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

Обозначение: $X \sim Exp(\lambda)$.

Нормальное распределение

Нормальным (гауссовским) называют распределение с плотностью

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

Нормальное распределение

Нормальным (гауссовским) называют распределение с плотностью

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

Величины a и σ называют **нормального закона распределения**.

Нормальное распределение

Нормальным (гауссовским) называют распределение с плотностью

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

Величины a и σ называют **параметрами нормального закона распределения**.

$$\boxed{X \sim N(a,\sigma)}$$

Нормальное распределение

Нормальным (гауссовским) называют распределение с плотностью

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

Величины a и σ называют **параметрами нормального закона распределения**.

$$\boxed{X\sim N(a,\sigma)}$$

Если $a=0$, $\sigma=1$, то распределение вероятности носит название *стандартного нормального распределения*.

Нормальное распределение

Нормальным (гауссовским) называют распределение с плотностью

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

Величины a и σ называют **параметрами нормального закона распределения**.

$$\boxed{X\sim N(a,\sigma)}$$

Если $a=0,\ \sigma=1$, то распределение вероятности носит название *стандартного нормального распределения*. Плотность:

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

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

так как

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$ (интеграл Пуассона).

График плотности нормального распределения называют *нормальной кривой* *или кривой Гаусса*.

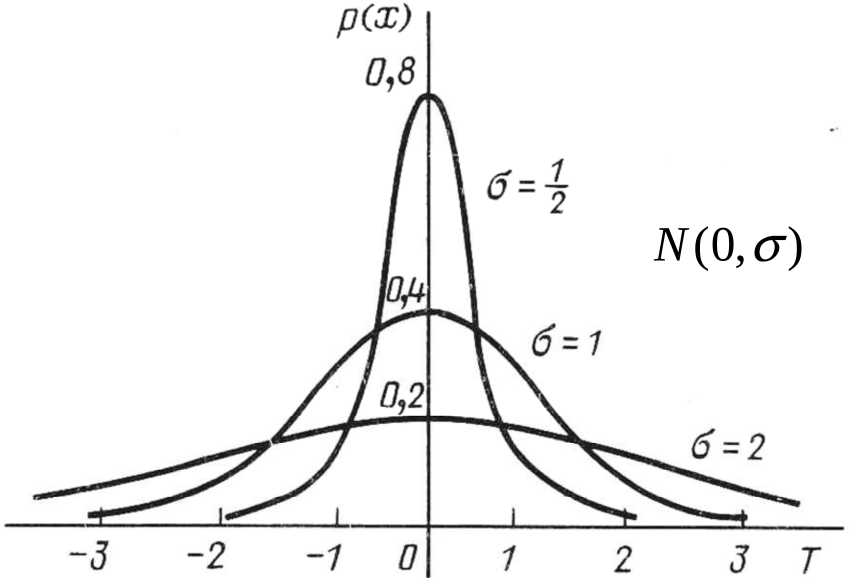
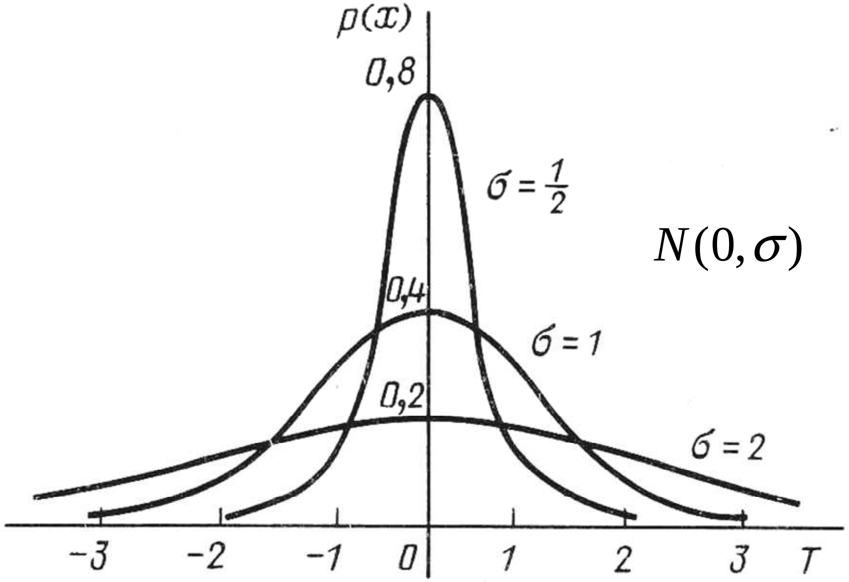


График плотности нормального распределения называют *нормальной кривой* *или кривой Гаусса*.



$N(0, \sigma)$

Нормальное распределение служит моделью распределения случайных величин для большого числа практических задач (распределение ошибок измерений, распределение среднего значения и т.д.)

Функция Лапласа

Пусть известно, что $X \sim N(a, \sigma)$.

Как найти вероятность события $X \in (x_1, x_2)$?

Функция Лапласа

Пусть известно, что $X \sim N(a,\sigma)$.
Как найти вероятность события $X \in (x_1, x_2)$?

Перейдем к стандартному нормальному распределению:

$$Z = \frac{X - a}{\sigma};$$

Функция Лапласа

Пусть известно, что $X \sim N(a,\sigma)$.
Как найти вероятность события $X \in (x_1, x_2)$?

Перейдем к стандартному нормальному распределению:

$$Z = \frac{X-a}{\sigma};$$

тогда $Z \sim N(0,1)$.

Функция Лапласа

Пусть известно, что $X \sim N(a, \sigma)$.
Как найти вероятность события $X \in (x_1, x_2)$?

Перейдем к стандартному нормальному распределению:

$$Z = \frac{X - a}{\sigma};$$

тогда $Z \sim N(0,1)$.

Если $X = x_1$, то

$$Z = \frac{x_1 - a}{\sigma},$$

Функция Лапласа

Пусть известно, что $X \sim N(a, \sigma)$.
Как найти вероятность события $X \in (x_1, x_2)$?

Перейдем к стандартному нормальному распределению:

$$Z = \frac{X - a}{\sigma};$$

тогда $Z \sim N(0,1)$.

Если $X = x_1$, то

$$Z = \frac{x_1 - a}{\sigma},$$

если же $X = x_2$, то

$$Z = \frac{x_2 - a}{\sigma}.$$

Значит

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

Значит

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

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{x_1-a}{\sigma}}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz$$

Значит

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

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{x_1 - a}{\sigma}}^{\frac{x_2 - a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x_2 - a}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x_1 - a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

Значит

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

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{x_1-a}{\sigma}}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\frac{x_1-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

$$= \Phi_0\left(\frac{x_2 - a}{\sigma}\right) - \Phi_0\left(\frac{x_1 - a}{\sigma}\right),$$

где $\Phi_0\left(x\right) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int\limits_0^x e^{-\frac{t^2}{2}} dt = \int\limits_0^x \varphi(t)\, dt,$

Значит

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

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{x_1-a}{\sigma}}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{x_1-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

$$= \Phi_0\left(\frac{x_2 - a}{\sigma}\right) - \Phi_0\left(\frac{x_1 - a}{\sigma}\right),$$

где $\Phi_0\left(x\right) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \int_0^x \varphi(t)\, dt$,

$\varphi(t)$ - плотность стандартного нормального распределения.

Значит

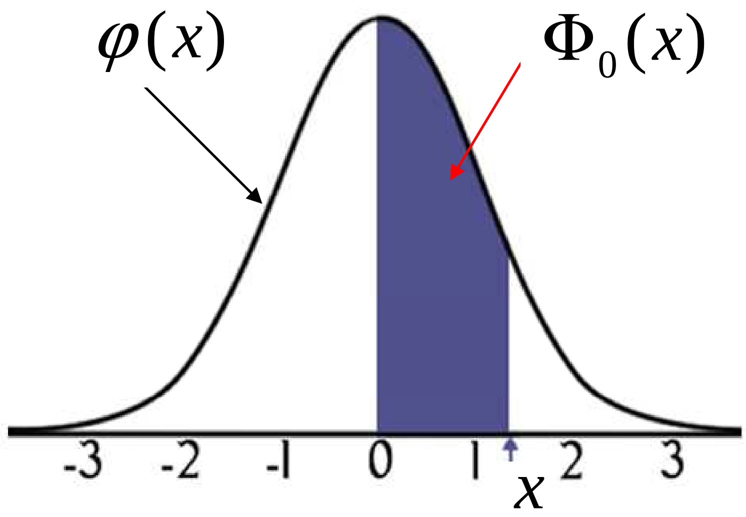
$$P\left(x_1 < X < x_2\right) = P\left(\frac{x_1 - a}{\sigma} < Z < \frac{x_2 - a}{\sigma}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{x_1-a}{\sigma}}^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x_2-a}{\sigma}} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x_1-a}{\sigma}} e^{-\frac{z^2}{2}} dz =$$

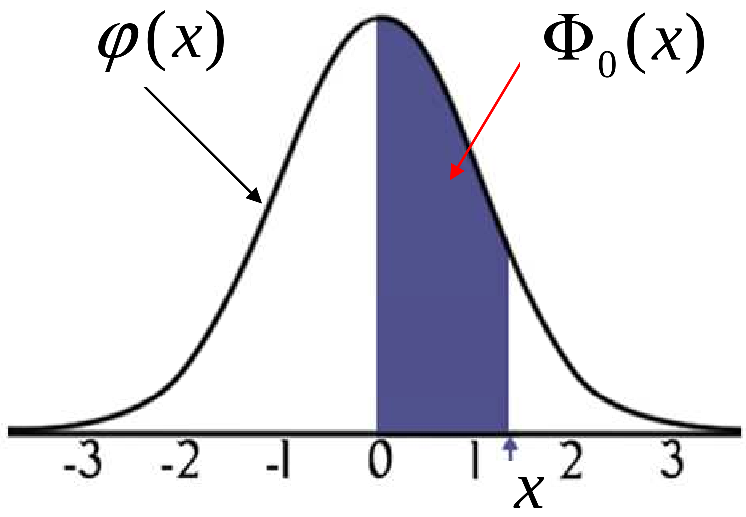
$$= \Phi_0\left(\frac{x_2 - a}{\sigma}\right) - \Phi_0\left(\frac{x_1 - a}{\sigma}\right),$$

где $\Phi_0\left(x\right) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-\frac{t^2}{2}} dt = \int_0^x \varphi(t)\, dt,$

$\varphi(t)$ - плотность стандартного нормального распределения. Функцию $\Phi_0\left(x\right)$ называют интегральной функцией Лапласа.

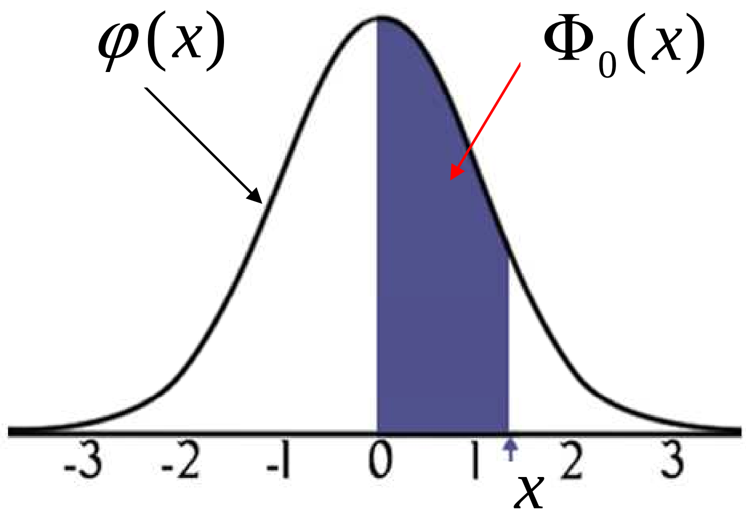


$\varphi(x)$
$\Phi_0(x)$
-3
-2
-1
0
1
2
3
x



Для функции распределения стандартной нормальной величины справедливо:

$$F(x) = \frac{1}{2} + \Phi_0(x), \quad \Phi_0(-x) = -\Phi_0(x).$$



Для функции распределения стандартной нормальной величины справедливо:

$$F(x) = \frac{1}{2} + \Phi_0(x), \quad \Phi_0(-x) = -\Phi_0(x).$$

Интеграл $\Phi_0(x)$ не берется; вычисляется с помощью численных методов. Существуют таблицы распределения функции Лапласа.

Пример. Случайная величина X распределена по нормальному закону с параметрами a=30, σ=10.
Найти вероятность того, что X примет значение, принадлежащее интервалу (10,50).

Пример. Случайная величина X распределена по нормальному закону с параметрами a=30, σ=10. Найти вероятность того, что X примет значение, принадлежащее интервалу (10,50).

Решение. По условию $x_1 = 10$, $x_2 = 50$, a=30, σ=10, следовательно, $\frac{x_2 - a}{\sigma} = 2$, $\frac{x_1 - a}{\sigma} = -2$.

Пример. Случайная величина X распределена по нормальному закону с параметрами a=30, σ=10.
Найти вероятность того, что X примет значение, принадлежащее интервалу (10,50).
Решение. По условию $x_1 = 10$, $x_2 = 50$, a=30, σ=10,

следовательно, $\dfrac{x_2 - a}{\sigma} = 2$, $\dfrac{x_1 - a}{\sigma} = -2$.

Поэтому

$$P\left(10 < X < 50\right) = \Phi_0\left(2\right) - \Phi_0\left(-2\right) = 2 \cdot \Phi_0\left(2\right)$$

$$= 2 \cdot 0.4772 = 0.9544.$$

Правило 3-х сигм

Пусть случайная величина X распределена нормально с параметрами a, σ. Найдем вероятность того, что отклонение случайной величины X от a по абсолютной величине будет меньше 3σ:

Правило 3-х сигм

Пусть случайная величина X распределена нормально с параметрами a, σ. Найдем вероятность того, что отклонение случайной величины X от a по абсолютной величине будет меньше 3σ:

$$P\left(\left|X-a\right|<3\sigma\right)=P\left(\frac{\left|X-a\right|}{\sigma}<3\right)=$$

$$=\Phi_0\left(3\right)-\Phi_0\left(-3\right)=2\Phi_0\left(3\right)=0.9973.$$

Правило: нормально распределенная случайная величина с вероятностью близкой к 1 не выходит за пределы интервала с границами $a \pm 3\sigma$.

Правило: нормально распределенная случайная величина с вероятностью близкой к 1 не выходит за пределы интервала с границами $a \pm 3\sigma$.

Другая формулировка: если для некоторой случайной величины, закон распределения которой неизвестен, "правило 3σ" **не выполняется**, то есть основания сомневаться в том, что ее распределение может быть нормальным.

Гамма распределение

Плотность гамма распределения задается как

$$f(x) = \begin{cases} \dfrac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, & x > 0 \\ 0, & x \le 0, \end{cases}$$

где $a > 0,\ \lambda > 0$ - параметры распределения,

Гамма распределение

Плотность гамма распределения задается как

$$f(x) = \begin{cases} \dfrac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, & x > 0 \\ 0, & x \le 0, \end{cases}$$

где $a > 0, \lambda > 0$ - параметры распределения,

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$$

- гамма функция ($\Gamma(a) = (a-1)!$ при $a = 1, 2, \ldots$).

Гамма распределение

Плотность гамма распределения задается как

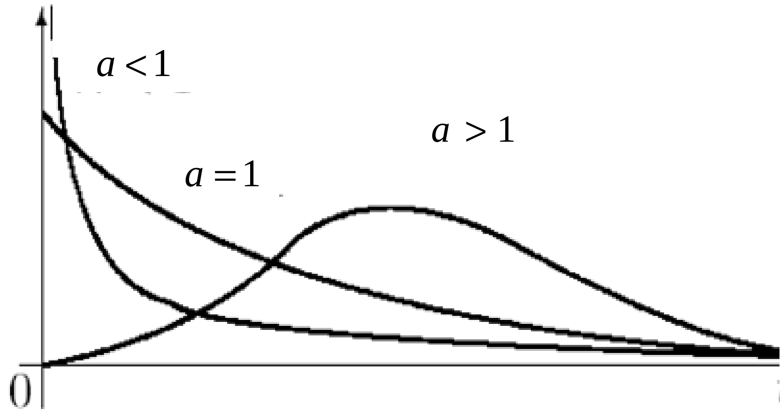
$$f(x) = \begin{cases} \dfrac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}, & x > 0 \\ 0, & x \le 0, \end{cases}$$

где $a > 0,\ \lambda > 0$ - параметры распределения,

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$$

- гамма функция ($\Gamma(a) = (a-1)!$ при $a = 1, 2, \ldots$).

$$X \sim \Gamma(a, \lambda)$$



Распределение Коши

Плотность распределения:

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}.$$

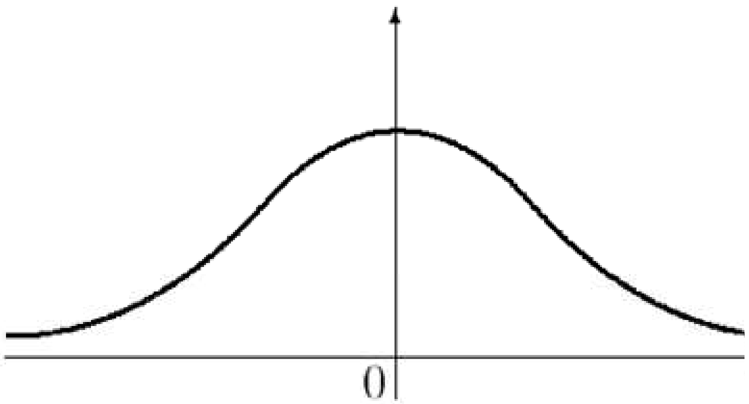
$$X \sim K$$

Распределение Коши

Плотность распределения:

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$

$$X \sim K$$

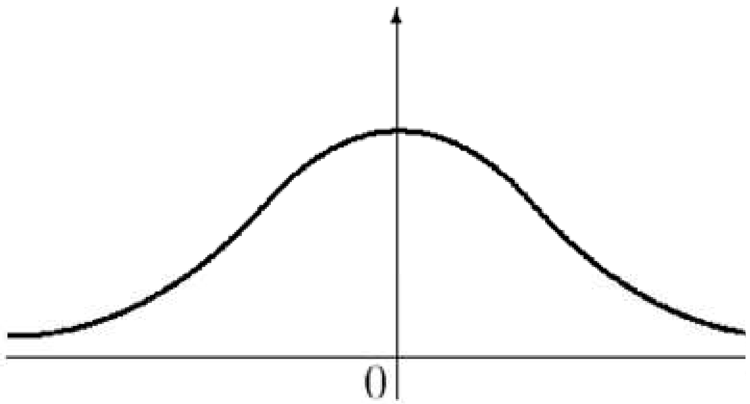


Распределение Коши

Плотность распределения:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

$$X \sim K$$



Функция распределения

$$F(x) = \frac{1}{\pi}\int\limits_{-\infty}^{x}\frac{1}{1+t^2}dt = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x.$$