

Глава 3. Случайные величины (продолжение).....	2
4. Основные распределения непрерывных случайных величин. Нормальное распределение.	2
Интеграл Пуассона.....	2
Определение нормального распределения.....	3
Свойства плотности нормального распределения.....	4
График плотности нормального распределения.....	4
Доска Гальтона.....	6
Влияние параметров нормального распределения на вид графика:.....	7
Функция распределения. Функция Лапласа.	8
Вероятность попадания в интервал . Правило трёх сигм.	9
Нормальное распределение как предельное для биномиального распределения.	10
Характеристическая функция нормального распределения:.....	10
Коэффициент асимметрии.....	12
Коэффициент эксцесса.	12

Глава 3. Случайные величины (продолжение).

4. Основные распределения непрерывных случайных величин. Нормальное распределение.

Стандартная случайная величина

Пусть случайная величина X имеет конечные мат. ожидание MX и дисперсию DX .

Определение. С.в. $Y = X - MX$ называется центрированной.

Математическое ожидание центрированной случайной величины равно нулю:

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

Определение. С.в. $Y = \frac{X - MX}{\sqrt{DX}}$ называется нормированной.

Дисперсия нормированной случайной величины равна единице:

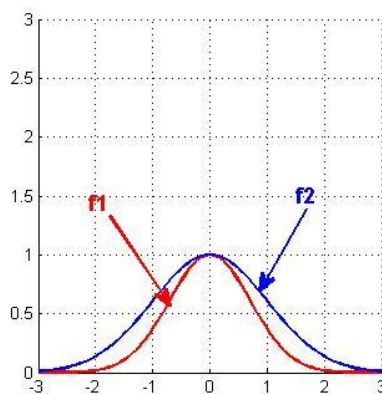
$$D\left(\frac{X - MX}{\sqrt{DX}}\right) = 1.$$

Определение. Центрированная и нормированная случайная величина $\tilde{X} = \frac{X - MX}{\sqrt{DX}}$ называется стандартной.

Для стандартной случайной величины: $MX = 0$, $DX = 1$.

Интеграл Пуассона.

Рассмотрим графики функций $f_1(t) = e^{-t^2}$ и $f_2(t) = e^{-t^2/2}$:



Площади, ограниченные графиками этих функций и осью x , представляют собой интегралы Пуассона:

$$\int_{-\infty}^{\infty} f_1(t) dt = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi} \quad , \quad \int_{-\infty}^{\infty} f_2(t) dt = \int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi}$$

В силу того, что подинтегральные функции являются четными:

$$\int_{-\infty}^0 e^{-t^2} dt = \int_0^{\infty} e^{-t^2} dt = \sqrt{\pi} / 2,$$

$$\int_{-\infty}^0 e^{-t^2/2} dt = \int_0^{\infty} e^{-t^2/2} dt = \sqrt{2\pi} / 2 = \sqrt{\frac{\pi}{2}}$$

Определение нормального распределения.

Определение. Случайная величина имеет нормальное распределение, если ее плотность распределения:

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

Нормальное распределение определяется двумя параметрами: a и σ . Такое распределение еще называют **гауссовским**.

Обозначается: $X \sim N(a, \sigma)$.

Можно показать, что в $N(a, \sigma)$ a является математическим ожиданием, σ - с.к.о. (или стандартным отклонением). Запишем определение математического ожидания:

$$MX = \int_{-\infty}^{\infty} x \cdot f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} dx =$$

$$\frac{x-a}{\sigma} = u, \Rightarrow x = \sigma u + a \Rightarrow dx = \sigma du$$

$$= \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (u\sigma + a) e^{-\frac{1}{2}u^2} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u\sigma e^{-\frac{1}{2}u^2} du + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a e^{-\frac{1}{2}u^2} du =$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u e^{-\frac{1}{2}u^2} du + a \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du =$$

интеграл по симметричному относительно $x=0$ интервалу от нечётной функции равен нулю.

Теперь рассмотрим определение дисперсии, подставив найденное мат.ожидание:

$$DX = \int_{-\infty}^{\infty} (x-a)^2 \cdot f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 \cdot e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} dx =$$

$$\frac{x-a}{\sigma} = u, \Rightarrow dx/\sigma = du \Rightarrow$$

$$= \frac{1 \cdot \sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{x-a}{\sigma}\right)^2 \cdot e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} \frac{dx}{\sigma} = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \cdot e^{-\frac{1}{2}u^2} du = - \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u \cdot d(e^{-\frac{1}{2}u^2}) =$$

$$(интегрируем по частям) = - \frac{\sigma^2}{\sqrt{2\pi}} \left[u \cdot e^{-\frac{1}{2}u^2} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du \right] = - \frac{\sigma^2}{\sqrt{2\pi}} (-\sqrt{2\pi}) = \sigma^2$$

=0 (правило Лопиталя)

$$\int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi}$$

Далее мы найдем математическое ожидание и дисперсию с помощью характеристической функции нормального распределения.

Свойства плотности нормального распределения.

1. $f(x) > 0$;
2. $\lim_{x \rightarrow \pm\infty} f(x) = 0$;
3. $\max f(x) = f(a)$;
4. $(a \pm \sigma, \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2})$ - точка перегиба;
5. условие нормировки: $\int_{-\infty}^{\infty} f(x) dx = 1$

Проверка условия нормировки:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2}[\frac{x-a}{\sigma}]^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}[\frac{x-a}{\sigma}]^2} d(\frac{x-a}{\sigma}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = 1$$

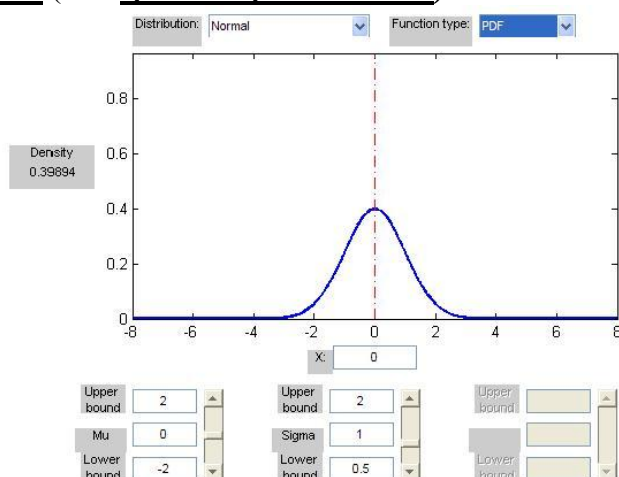
$$(\int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi} \text{ - интеграл Пуассона})$$

График плотности нормального распределения.

Распределение $N(0,1)$ называется стандартным. В этом случае плотность имеет вид:

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

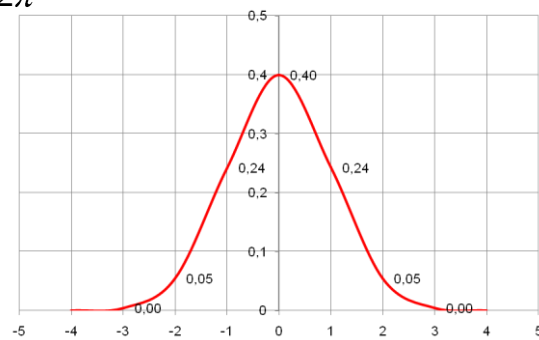
График нормального стандартного распределения представляет собой кривую, которая называется кривой Гаусса (или кривой вероятностей):



Комментарии к рисунку.

Function Type – PDF (probability density function);
показано значение функции $f(0)$ (максимум): $Density=0.018316$;
значение мат.ожидания "Mu" = $m_x=0$, с.к.о. $\sigma = Sigma=1$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \varphi(x) \quad - \text{функция Гаусса}$$



1. Чётная.
2. При $x \geq 4$ можно считать равной 0.



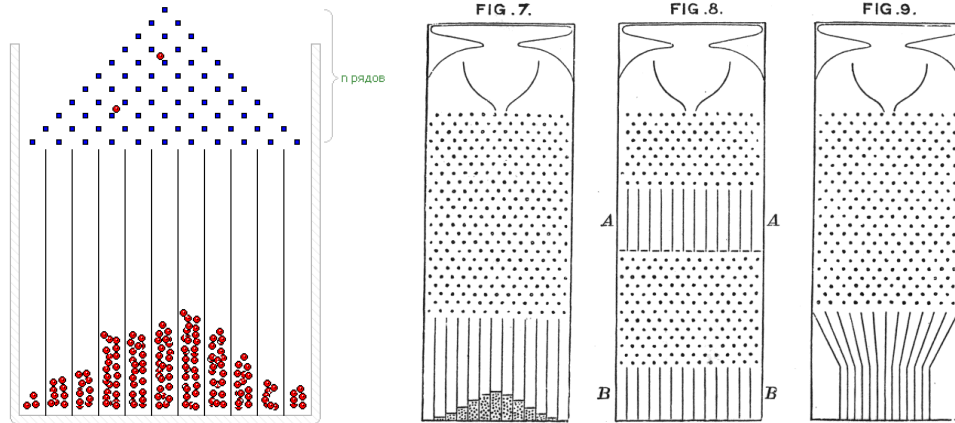
Карл Фридрих Гаусс (нем. *Johann Carl Friedrich Gauß*; [30 апреля 1777](#), Брауншвейг — [23 февраля 1855](#), Гёттинген)

Такую кривую можно получить экспериментально – с помощью доски Гальтона.

Доска Гальтона.

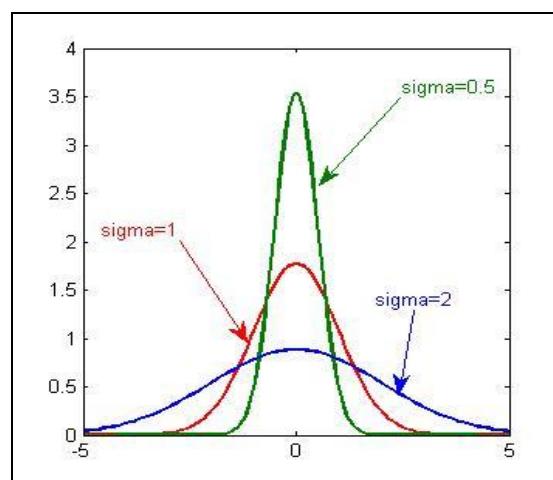
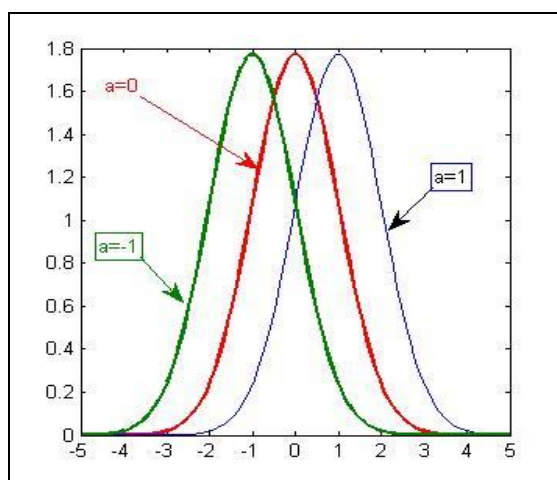
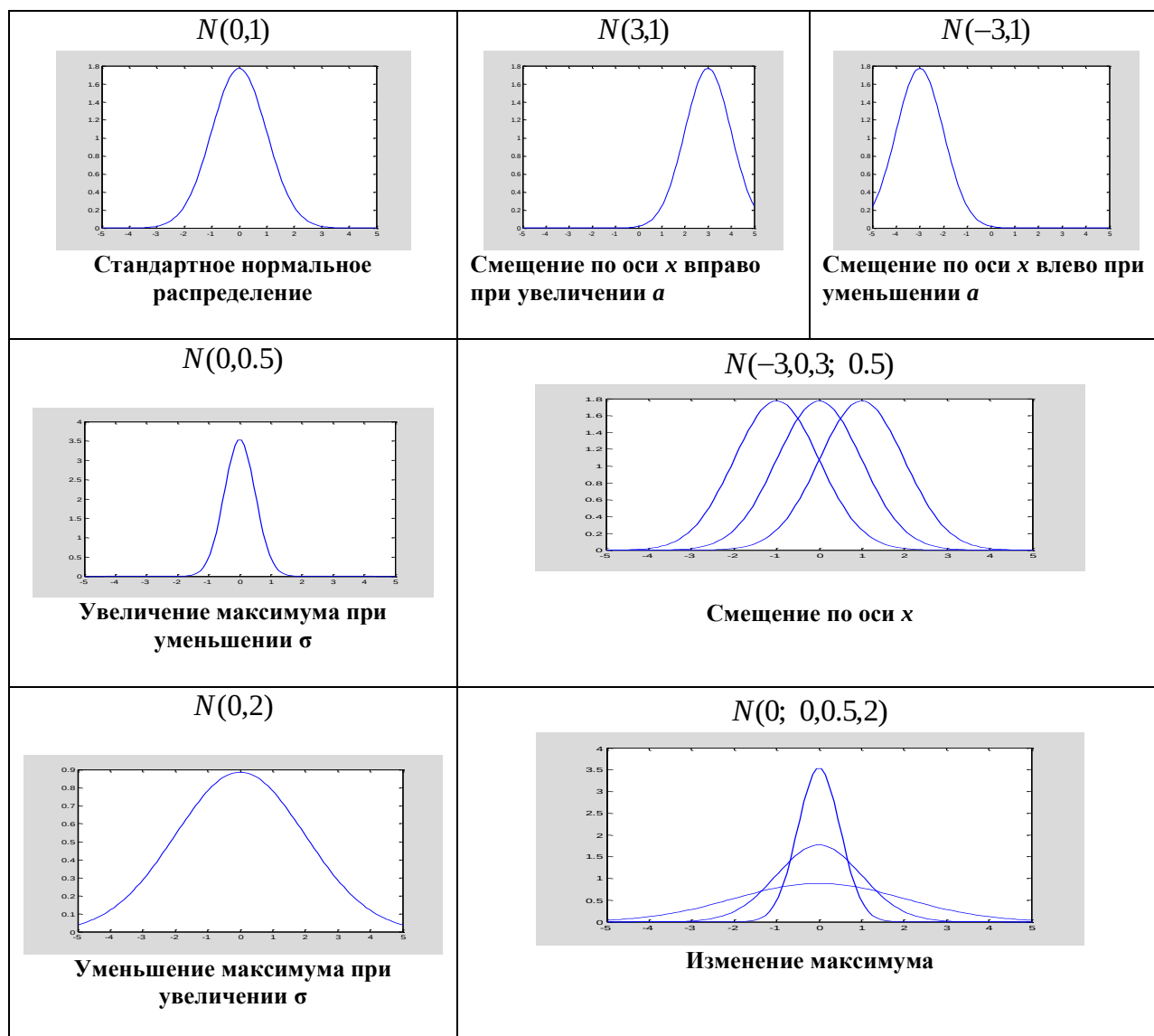
Доска Гальтона представляет собой наклонную доску на которую набиты гвозди – рядами на одинаковом расстоянии друг от друга. Главный принцип – очередной ряд гвоздей смещен по отношению к предыдущему ровно на половину расстояния между гвоздями.

Различные виды таких досок приведены на рисунках:



Далее из воронки высыпаятся шарики одинакового размера, которые сталкиваются на пути вниз как с гвоздями, так и друг с другом. Таким образом, мы моделируем схему, когда конечное положение – попадание шарика в одну из лунок – зависит от суммы случайных величин. (Можно посмотреть ролик [Galton.avi](#)). Если обвести плавной кривой столбики из шариков, то получим кривую Гаусса.

Влияние параметров нормального распределения на вид графика плотности:

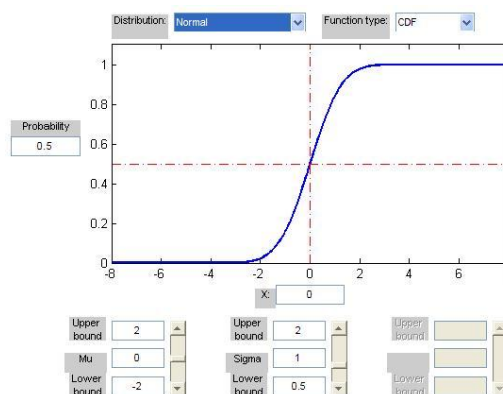


Функция распределения. Функция Лапласа.

Функция распределения (нормального распределения):

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

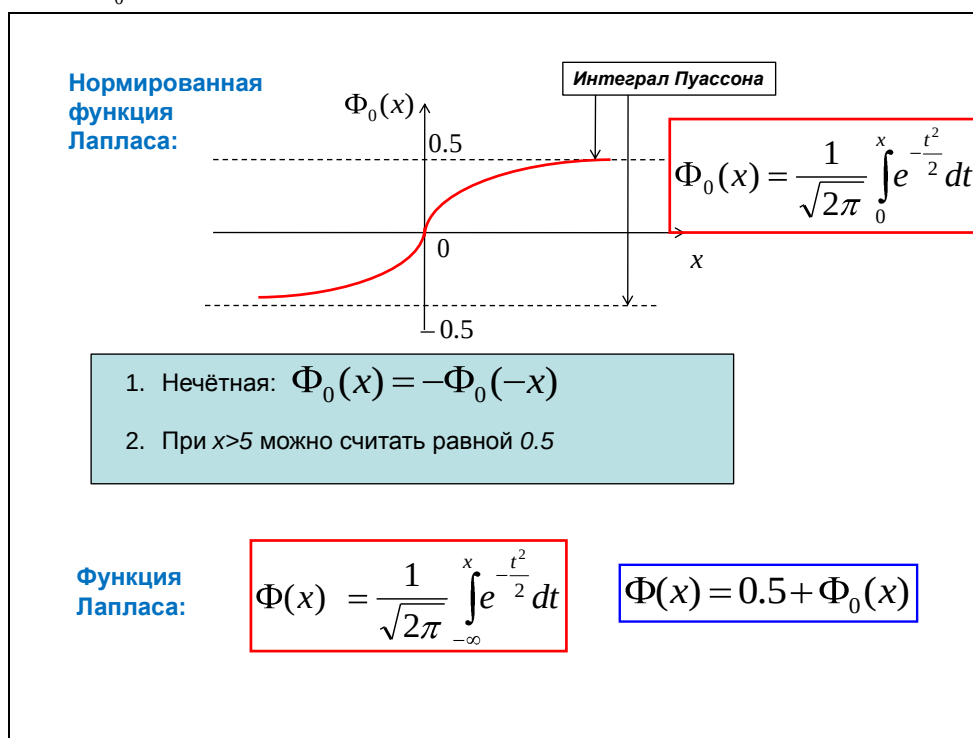
График функции распределения $N(0,1)$:



Функция распределения для $N(0,1)$ есть функция Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \frac{1}{2} + \Phi_0(x),$$

где $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ - нормированная функция Лапласа:



Заметим, что $\Phi_0(\pm\infty) = \pm \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} = \pm \frac{1}{2}$.

Значения функций Лапласа берутся из таблиц:

Приложение 2. Значение функции $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0040	0080	0112	0160	0199	0239	0279	0319	0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3079	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3553	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1,6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4700	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4762	4767

x	Десятые доли x									
	0	1	2	3	4	5	6	7	8	9
2,	4773	4821	4861	4893	4918	4938	4953	4965	4974	4981
3,	4987	4990	4993	4995	4997	4998	4998	4999	4999	5000 ¹

$x_2 = 0.72$

$x_1 = 2.9$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-a}{\sigma}\right)^2} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{u^2}{2}} du = \Phi\left(\frac{x-a}{\sigma}\right) = 0.5 + \Phi_0\left(\frac{x-a}{\sigma}\right)$$

Следовательно, функция распределения для $N(a, \sigma)$ связана с нормированной функцией Лапласа таким соотношением:

$$F(x) = \frac{1}{2} + \Phi_0\left(\frac{x-a}{\sigma}\right).$$

Вероятность попадания в интервал . Правило трёх сигм.

Таким образом, используя свойство функции распределения, можем определить вероятность попадания в интервал $[\alpha, \beta]$ случайной величины $X \sim N(a, \sigma)$:

$$P(X \in [\alpha, \beta]) = \Phi_0\left(\frac{\beta-a}{\sigma}\right) - \Phi_0\left(\frac{\alpha-a}{\sigma}\right) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right)$$

Вероятность попадания в интервал, симметричный относительно математического ожидания (центра рассеяния):

$$P(a-l < X < a+l) = \Phi_0\left(\frac{a+l-a}{\sigma}\right) - \Phi_0\left(\frac{a-l-a}{\sigma}\right) = \Phi_0\left(\frac{l}{\sigma}\right) - \Phi_0\left(\frac{-l}{\sigma}\right) = 2\Phi_0\left(\frac{l}{\sigma}\right)$$

(воспользовались свойством нечетности функции Лапласа).

В частности, для $l = 3\sigma$: $P(a-3\sigma < X < a+3\sigma) = 2\Phi_0\left(\frac{3\sigma}{\sigma}\right) = 2\Phi_0(3) = 2 \cdot 0.49865 \approx 0.9973$

(см. таблицу значений Φ_0).

Отсюда вытекает, так называемое, «правило трех сигм»:

практически достоверно, что с.в. $X \sim N(a, \sigma)$ принимает свои значения в промежутке $(a-3\sigma < X < a+3\sigma)$.

Нормальное распределение как предельное для биномиального распределения.

Нормальное распределение является предельным для биномиального распределения. Вспомним теорему Муавра-Лапласа:

Интегральная теорема Муавра-Лапласа

Теорема 3. Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность

$$P_n(k_1 \leq m \leq k_2) \quad (\text{число успехов из указанного интервала})$$

может быть найдена по приближённой формуле:

$$P_n(k_1 \leq m \leq k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx,$$

где

$$x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}$$

С учетом полученных знаний о нормальном распределении, посмотрим на формулировку изложенной ранее теоремы Муавра-Лапласа:

1. С.в. X распределена по биномиальному закону с $MX = np$, $DX = npq$. Нужно найти вероятность попадания значений случайной величины в заданный интервал $[k_1, k_2]$.
2. Преобразуем к стандартной случайной величине. Границы интервала преобразуются в $[x_1, x_2]$.
3. Далее приближенно вычисляем искомую вероятность как вероятность попадания в интервал $[x_1, x_2]$ значений случайной величины, распределенной по закону $N(0,1)$.

Или: **С.в. X можно считать распределенной по нормальному закону $X \sim N(np, \sqrt{npq})$.**

Тогда $P(X \in [k_1, k_2]) = \Phi_0\left(\frac{k_2 - a}{\sigma}\right) - \Phi_0\left(\frac{k_1 - a}{\sigma}\right)$

Характеристическая функция нормального распределения:

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-(a+it\sigma^2))^2}{2\sigma^2}} \cdot e^{\frac{2ait\sigma^2 + (it\sigma^2)^2}{2\sigma^2}} dx =$$

$$= \frac{1}{\sqrt{\pi}} e^{ita - \frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-u^2} du = e^{ita - \frac{t^2\sigma^2}{2}},$$

(при выводе сначала выделяем квадрат степени экспоненты, затем делаем замену переменных и приводим

интеграл к интегралу Пуассона $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$)

Таким образом, для случайной величины $X \sim N(a, \sigma)$ характеристическая функция имеет вид:

$$\varphi(t) = e^{ita - \frac{t^2 \sigma^2}{2}}$$

Найдем с помощью характеристической функции математическое ожидание и дисперсию:

$$\begin{aligned} \varphi'(0) &= ia, \quad \varphi''(0) = -a^2 - \sigma^2, \\ \alpha_1 &= -i\varphi'(0) = a, \quad \alpha_2 = -\varphi''(0) = a^2 + \sigma^2 \end{aligned}$$

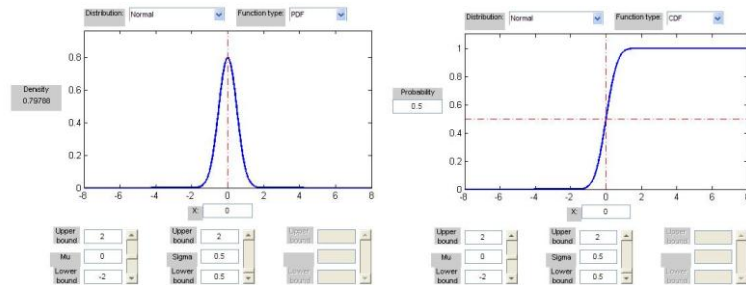
Следовательно, $MX = \alpha_1 = a$, $DX = \alpha_2 - \alpha_1^2 = \sigma^2$.

Замечание. В случае $X \sim N(0,1)$ характеристическая функция имеет вид:

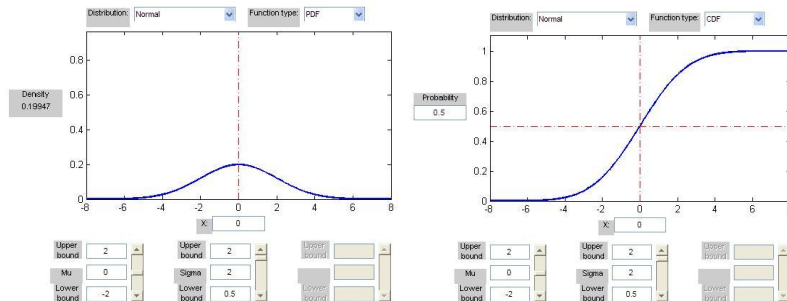
$$\varphi(t) = e^{-\frac{t^2}{2}}$$

Вид графиков плотности и функции распределения для различных значений параметров нормального распределения $N(a; \sigma)$.

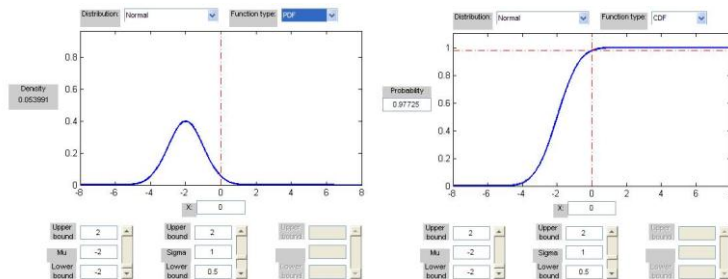
$N(0;0.5)$:



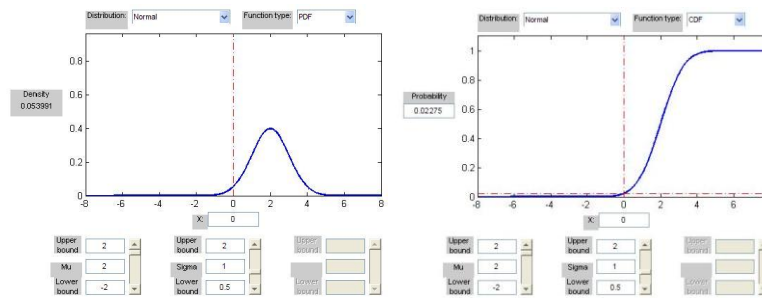
$N(0;2)$:



$N(-2;1)$:



$N(2;1)$:



Коэффициент асимметрии.

Для нормального распределения коэффициент асимметрии («скошенности»)

$$A = \frac{\mu_3}{\sigma^3}$$

равен нулю: $A_{\text{норм.распр.}} = \frac{\mu_3}{\sigma^3} = 0$.

Коэффициенты асимметрии других распределений сравниваются с коэффициентом асимметрии нормального распределения. Получаются либо отрицательными (*negative skew*), либо положительными (*positive skew*):

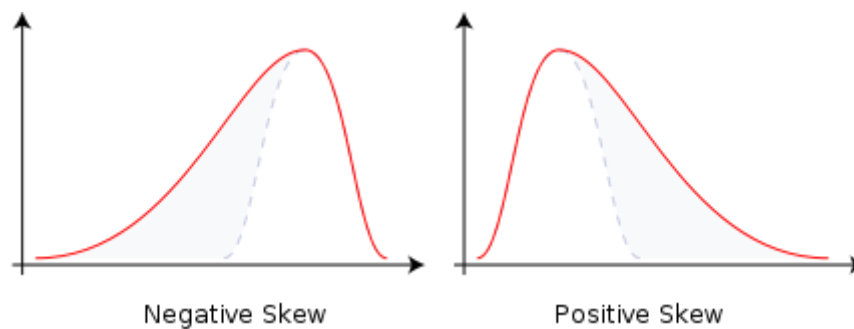


Рисунок. Отрицательная и положительная асимметрия.

Коэффициент эксцесса.

Для нормального распределения значение $\frac{\mu_4}{\sigma^4} = 3$.

Коэффициент эксцесса $E = \frac{\mu_4}{\sigma^4} - 3$ («островершинности») для нормального распределения:

$$E_{\text{норм.распр.}} = \frac{\mu_4}{\sigma^4} - 3 = 0.$$

Коэффициенты эксцесса (*kurtosis*) остальных распределений сравниваются с коэффициентом эксцесса нормального распределения:

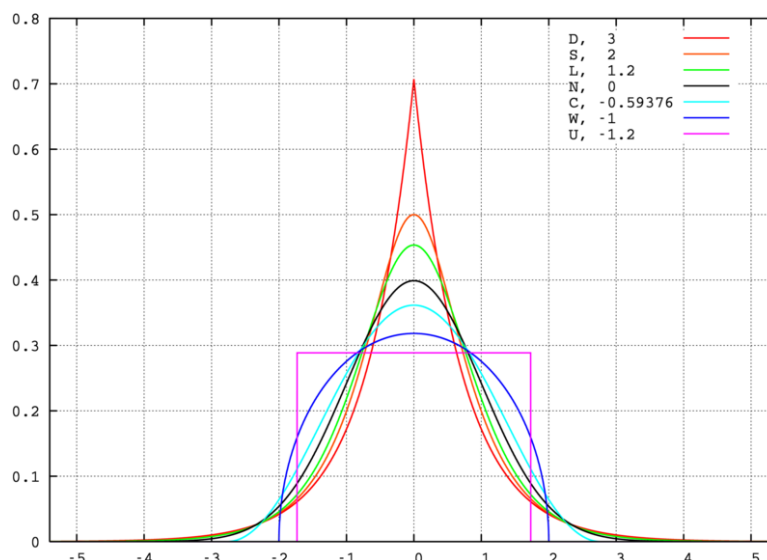


Рисунок. Черным цветом – плотность нормального распределения.
D: Laplace distribution, S: hyperbolic secant distribution, L: logistic distribution,
N: normal distribution, C: raised cosine distribution, W: Wigner semicircle distribution,
U: uniform distribution.

Замечание.

Для расчета коэффициентов асимметрии и эксцесса в Mathcad имеются две встроенные функции:

$kurt(x)$ — коэффициент эксцесса (kurtosis) выборки случайных данных x ;

$skew(x)$ — коэффициент асимметрии (skewness) выборки случайных данных X .

Примеры.

1. При измерении детали получаются случайные ошибки, подчиненные нормальному закону с параметром $\sigma=10$ мм. Производится 3 независимых измерения детали. Найти вероятность того, что ошибка хотя бы одного измерения не превосходит по модулю 2 мм. **Без систематических (одного знака!) ошибок: мат. ожидание ошибки $\alpha=0$.**

Для одного опыта вероятность того, что ошибка не превосходит 2:

$$P\{|X| < 1\} = 2\Phi_0\left(\frac{1}{\sigma}\right) \Rightarrow P\{|X| < 2\} = 2\Phi_0\left(\frac{2}{10}\right) \approx 2 \cdot 0.07926 = 0.15852.$$

Тогда вероятность того, ошибка превысит 2 мм в одном опыте:

$$P\{|X| > 2\} = 1 - P\{|X| < 2\} = 0.84148.$$

Вероятность того, что ошибка не превосходит 2 во всех 3 опытах:

$$0.15852^3 \approx 0.00398.$$

Следовательно, вероятность того, что ошибка хотя бы одного измерения не превосходит по модулю 2 мм: $1 - 0.00398 = 0.99602$.

2.