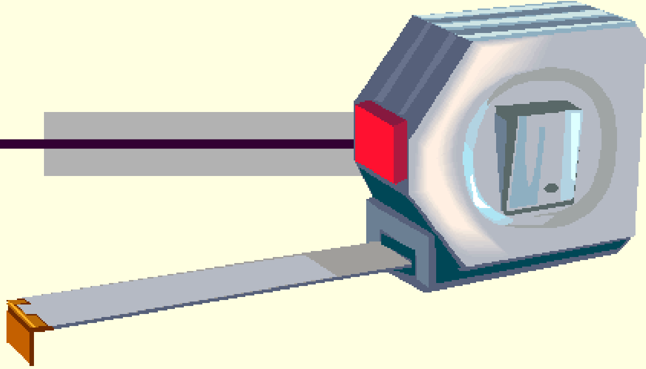


СЛУЧАЙНЫЕ ВЕЛИЧИНЫ.

СЛУЧАЙНЫЕ И ГРУБЫЕ ПОГРЕШНОСТИ

Погрешность

В реальных условиях даже очень точные измерения будут содержать **погрешность Δ**, которая является отклонением результата измерения x от истинного значения $x_{\text{ист}}$ измеряемой величины.

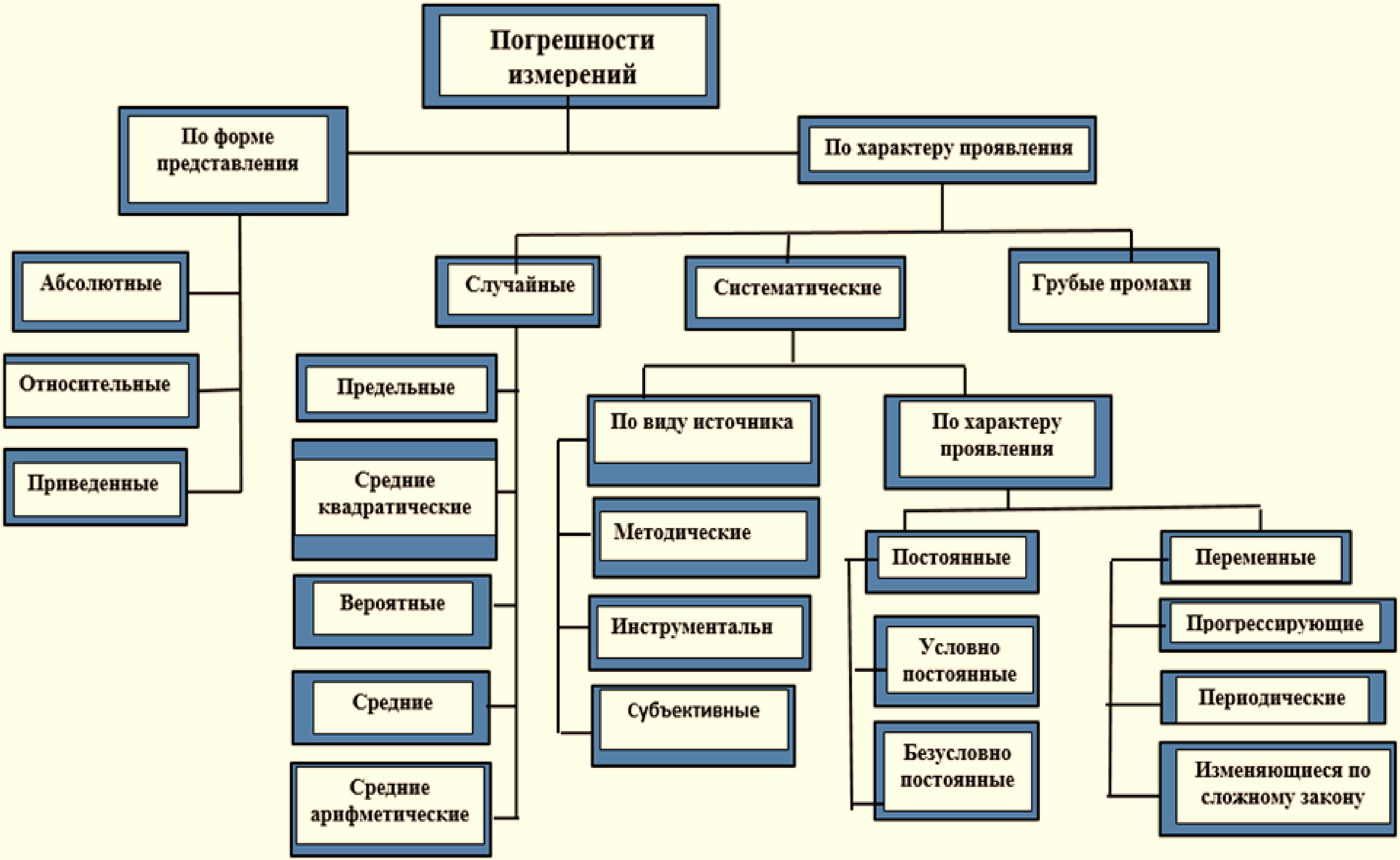
Погрешность измерения $\Delta x_{изм}$ – это отклонение результата измерения x от истинного (действительного) $x_и$ $(x_д)$ значения измеряемой величины:

$$\Delta x_{изм} = x - x_и \,, \qquad \textit{или} \qquad \Delta x_{изм} = x - x_д \,.$$

Погрешность средства измерения - отклонение показания средства измерения от истинного (действительного) значения измеряемой величины. Оно характеризует точность результатов измерений, проводимых данным средством.

Эти два понятия во многом близки друг к другу и классифицируются по одинаковым признакам.

Классификация погрешностей



Классификация погрешностей

По характеру проявления, способам обнаружения и учета погрешности измерений подразделяются на **систематические** и **случайные и грубые погрешности (промахи)**.

Систематическая погрешность измерения – составляющая погрешности результата измерения, остающаяся постоянной или закономерно изменяющаяся при повторных измерениях одной и той же физической величины.

- Примечание – в зависимости от характера измерения систематические погрешности подразделяют на **постоянные, прогрессивные**, **периодические** и погрешности, **изменяющиеся по сложному закону**.

Случайная погрешность измерения – составляющая погрешности результата измерения, изменяющаяся случайным образом (по знаку и значению) при повторных измерениях, проведенных с одинаковой тщательностью, одной и той же физической величины.

Грубая погрешность – погрешность результата отдельного измерения, входящего в ряд измерений, которая для данных условий резко отличается от остальных результатов этого ряда.

СЛУЧАЙНЫЕ погрешности

Случайные погрешности

Случайные погрешности вызываются большой совокупностью причин, остающихся при проведении измерений неизвестными.

Случайные погрешности неизбежны и неустранимы.

При многократных измерениях в одних и тех же условиях, одинаковым средством измерения и тем самым оператором какого-то параметра, мы получаем результаты, которые отличаются друг от друга и от действительного значения.

Погрешность, возникающая при этом носит случайный характер и может быть определена по формулам, которые базируются на ***законах теории вероятности*** *и* ***математической статистики.***

Случайная погрешность, как и всякая случайная величина, наиболее полно характеризуется **законом распределения.**

Случайной событие. Вероятность

Событие — некоторый факт, который в результате опыта может произойти или не произойти.

Событие называется случайным, если в результате опыта оно может произойти или не произойти.

Событие называется **достоверным,** если оно обязательно появляется в результате данного опыта, и невозможным, если оно не может появиться в этом опыте.

Пусть в N испытаниях событие A произошло $n(A)$ раз. Отношение $n(A)/N$ называется относительной частотой или просто частотой появления события А. Если провести несколько серий опытов по N испытаний в каждой, то отношение $n(A)/N$ будет различным для разных серий, но при увеличении N это отношение будет стремиться к некоторому постоянному числу, называемому **вероятностью появления события *A*:**

$$n(A)/N \to P(A) \text{ при } N \to \infty$$

Случайная величина

Пусть некоторая величина X в ряде испытаний может принимать различные числовые значения.

Если значение величины X в каждом данном испытании не может быть указано заранее (непредсказуемо), то величина X называется **случайной величиной**.

Другими словами *случайной* называют величину, которая в результате опыта (наблюдения, измерения) принимает одно возможное, но заранее неизвестное значение. Случайная величина может быть **дискретной** или **непрерывной**.

Случайные величины

Если случайная величина может принимать лишь дискретные значения, то она называется **дискретной случайной величиной.**

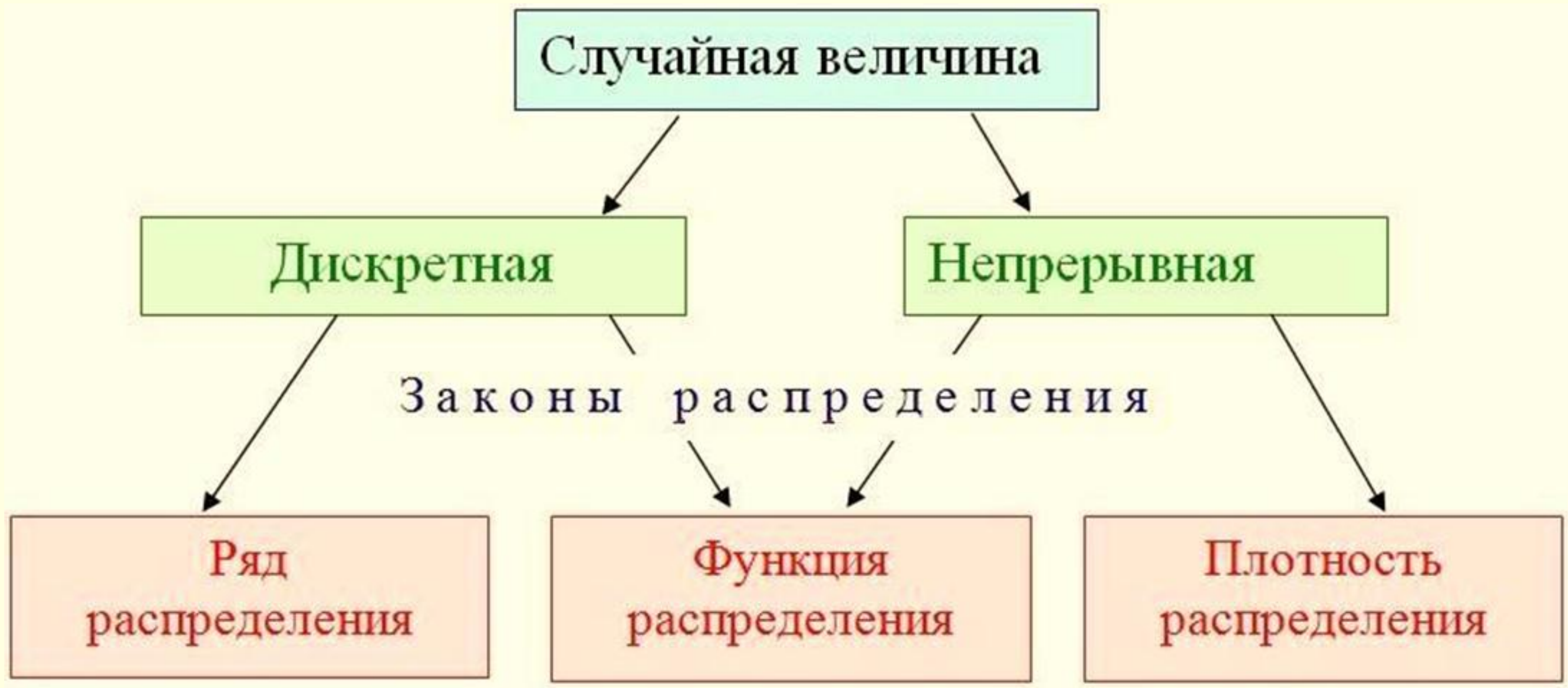
Примеры дискретных случайных величин: число отказов электронного устройства за рассматриваемый календарный период времени, например два года (возможные значения 0, 1, 2, 3, 4, ...); частота попадания сопротивления резистора, взятого из партии с сопротивлением R = 1 кОм 10%, в диапазон (950…1000 Ом).

Если случайная величина может принимать бесконечное множество значений, причем эти значения могут быть сколь угодно близки друг к другу, то такая величина называется ***непрерывной случайной величиной*.**

Примеры непрерывной случайной величины: зависимость сопротивления резистора от температуры, температура воздуха в определённый день

Охарактеризовать случайную величину можно при помощи закона распределения.

Случайная величина

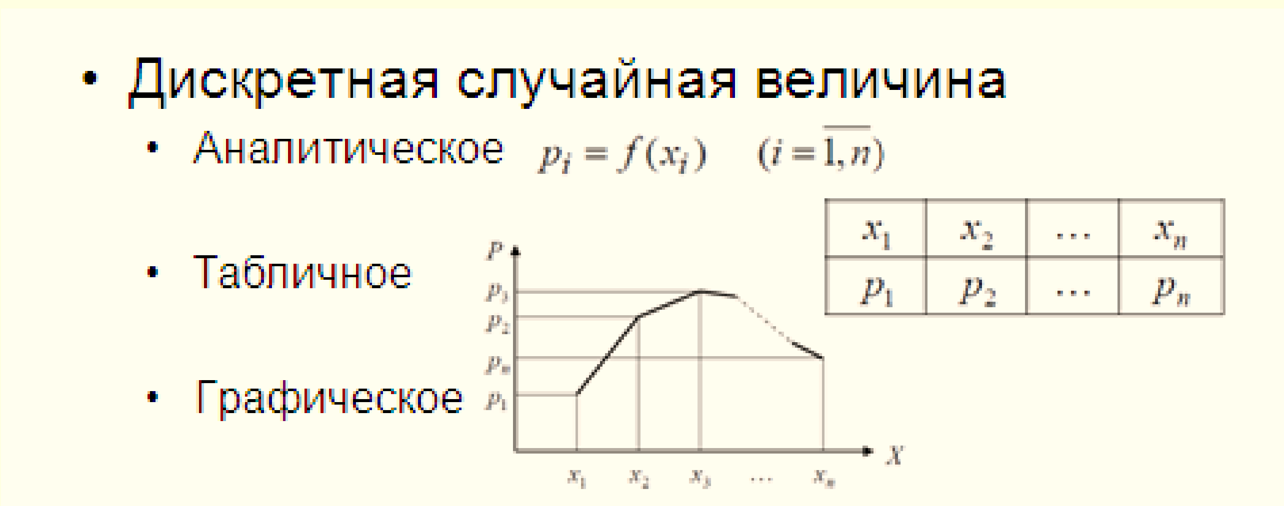


Закон распределения случайной величины

Под **законом распределения случайной величины** понимается соответствие, устанавливающее связь между возможными значениями случайной величины и вероятностями принятия этих значений.

Закон распределения дискретной случайной величины представляет собой перечень всех её возможных значений и соответствующих вероятностей. **Сумма всех вероятностей $\Sigma p_i = 1$**.

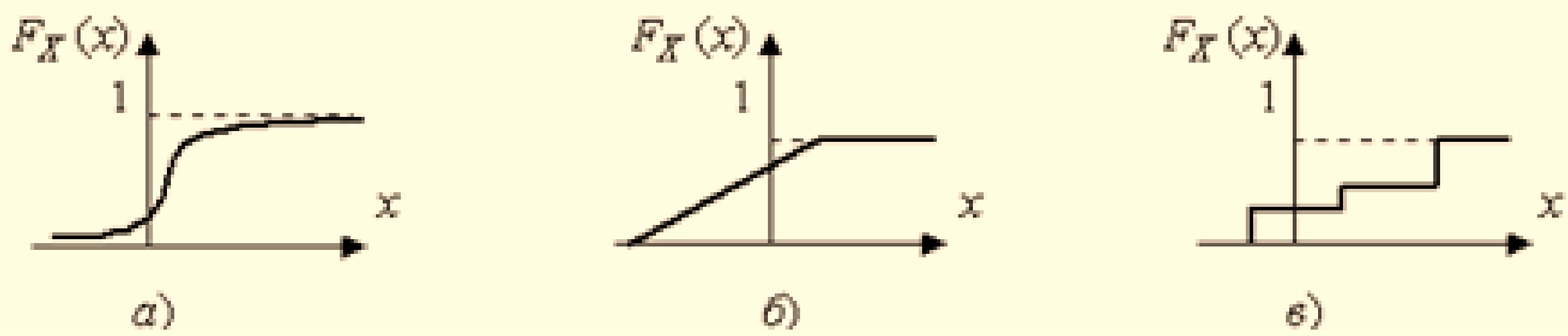
Закон распределения также может быть задан аналитически (формулой) и графически (многоугольником распределения, соединяющим точки (x_i; p_i)



Интегральная функция распределения

Наиболее универсальный способ описания случайных величин заключается в отыскании их интегральных или дифференциальных функций распределения.

Под **интегральной функцией распределения** результатов наблюдений понимается зависимость вероятности того, что результат наблюдения X в i-м опыте окажется меньшим некоторого текущего значения x_i, от самой величины x. Другими словами под функцией распределения случайной величины X для текущего значения x понимают вероятность не события $X = x$, а вероятность события $X < x$. Обозначают это как **$F(x) = P(X < x)$.** На рис. показаны примеры функций распределения вероятности.



Свойства интегральной функции распределения

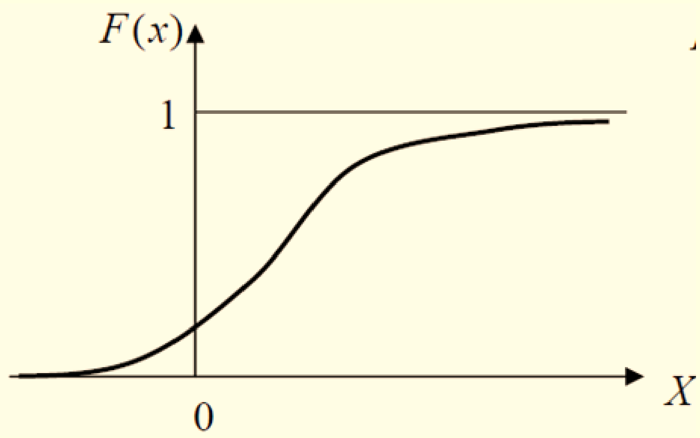
Свойства функции $F(x)$:

1. $F(x)$ – неубывающая функция,

т.е. $F(x_2) \geq F(x_1)$ при $x_2 > x_1$.

2. $F(x = -\infty) = 0$.

3. $F(x = +\infty) = 1$

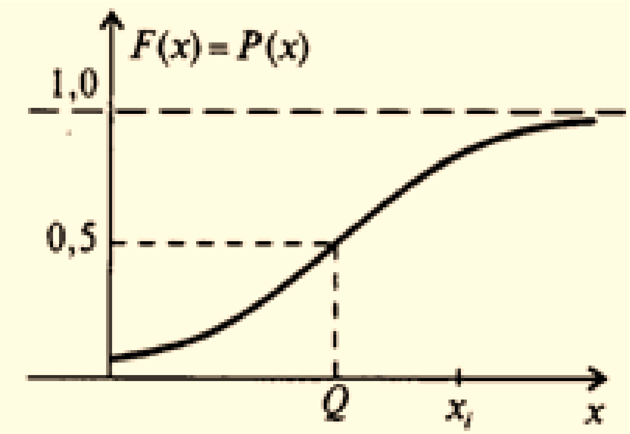


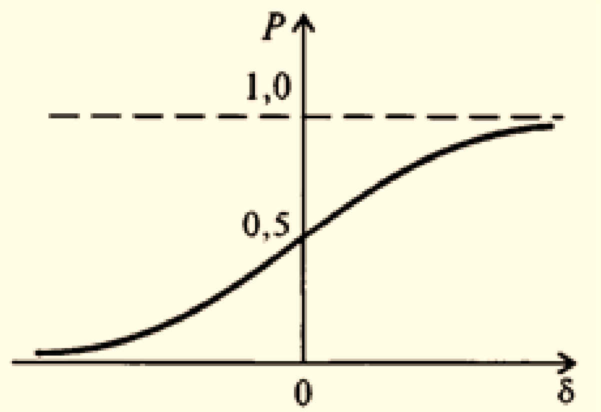
Интегральная функция распределения F(x) — неубывающая непрерывная функция, асимптотически стремящаяся к нулю при x → -∞ и к единице при x → ∞. Непрерывность функции F(x) понимается таким образом, что результат измерения может принять любое значение, но с вероятностью, близкой к нулю.

Интегральная функция распределения погрешности

Случайную погрешность δ можно рассматривать как случайную величину и ее поведение также описывать интегральной функцией распределения.

Интегральной функцией распределения погрешности δ называется зависимость вероятности того, что погрешность в i-м измерении x_i - Q будет меньше, чем погрешность x - Q. Получается эта функция путем переноса начала координат в точку $x = Q$.

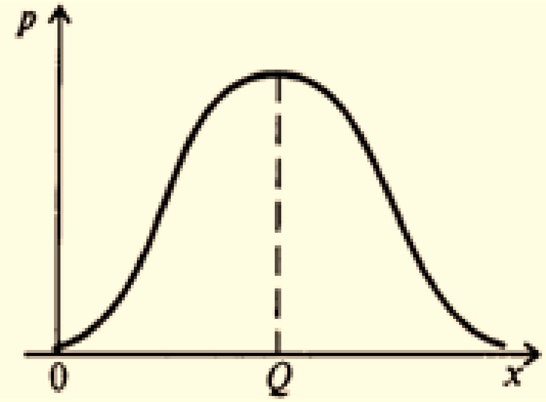


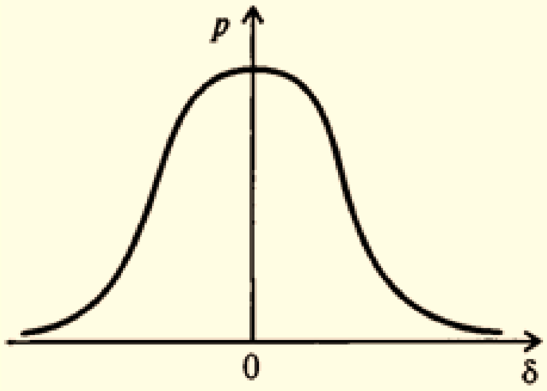


Дифференциальная функция распределения

Наибольшее распространение получили как более наглядные **дифференциальные функции распределения** результатов измерений p(x) и погрешностей p{δ), являющиеся производными от интегральных функций:

$$p(x) = \frac{dF(x)}{dx}, \qquad p(\delta) = \frac{dF(\delta)}{d\delta}.$$





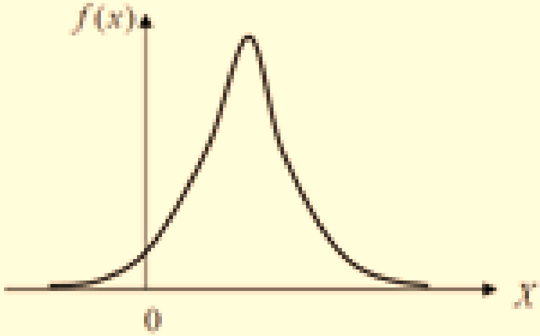
Дифференциальная функция распределения называемой также **плотностью распределения вероятностей**

Свойства дифференциальной функции распределения

Плотность распределения вероятностей $f(x)$ определяется как *производная* от функции распределения $F(x)$ по x:

$$f(x) = F'(x) = \frac{dF(x)}{dx}$$

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



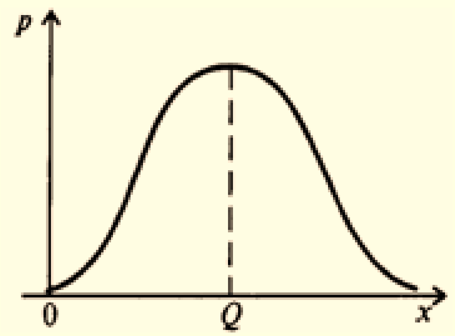
Свойства плотности распределения:

- плотность распределения есть функция *неотрицательная*: $f(x) \geq 0$;
- интеграл в бесконечных пределах от плотности распределения равен *единице*:

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

Свойства дифференциальной функции распределения

Обычно максимум дифференциальная функция распределения (или плотности распределения) совпадает (при отсутствии систематической погрешности) с истинным значением ФВ, поэтому дифференциальная функция распределения является более наглядной .



Дифференциальные функции позволяют определить вероятность попадания погрешности или результата измерения в заданный интервал:

$$P(\delta_1 \prec \delta \le \delta_2) = \int_{\delta_1}^{\delta_2} p(\delta)d\delta = P(-\infty \prec \delta \le \delta_2) - P(-\infty \prec \delta \le \delta_1) = \int_{-\infty}^{\delta_2} p(\delta)d\delta - \int_{-\infty}^{\delta_1} p(\delta)d\delta;$$

Очевидны также следующие равенства **(условия нормировки):**

$$\int_{-\infty}^{\infty} p(x)dx = P(-\infty \prec x \prec \infty) = 1; \int_{-\infty}^{\infty} p(\delta)d\delta = P(-\infty \prec \delta \prec \infty) = 1.$$

Генеральная совокупность и выборка

Каждая отдельная измерительная операция (отсчет, замер) называется **наблюдением,** а получаемое при этом значение физической величины – **результатом наблюдения**. Всё множество значений, которые измеряемая величина может принимать в эксперименте, называется **генеральной совокупностью.**

На практике ограничиваются конечным числом наблюдений (от единиц до нескольких десятков). Полученный при этом ряд значений физической величины: x_1, x_2, ..., x_N называют ***выборкой из генеральной совокупности*** или просто ***выборкой***. Число N результатов наблюдений в выборке называют ***объёмом выборки.***

Результаты наблюдений, входящие в выборку, можно упорядочить, т. е. расположить их в порядке возрастания или убывания: $x_1 \le x_2 \le ... \le x_N$. Полученную выборку называют упорядоченной или ***ранжированной.***

Величина $\boldsymbol{R = x_{\max} - x_{\min}}$ называется ***размахом выборки***.

Эмпирическое распределение результатов наблюдений

Чтобы получить представление о законе распределения измеряемой величины, экспериментальные данные группируют.

Для этого область полученных значений делят на некоторое количество интервалов одинаковой ширины Δx и подсчитаем количество измерений, попавших в каждый из этих интервалов.

Обозначают $m_1, m_2, m_3, \ldots m_k$ – количество измерений, попавших, соответственно в первый, второй, и т.д. интервал длиной Δx.

Относительная частота попадания результатов измерений в какой-либо интервал $(x_i, x_i + \Delta x)$ равна

$$\frac{m_i}{n}$$

Ось абсцисс разбивают на конечное число промежутков Δx.

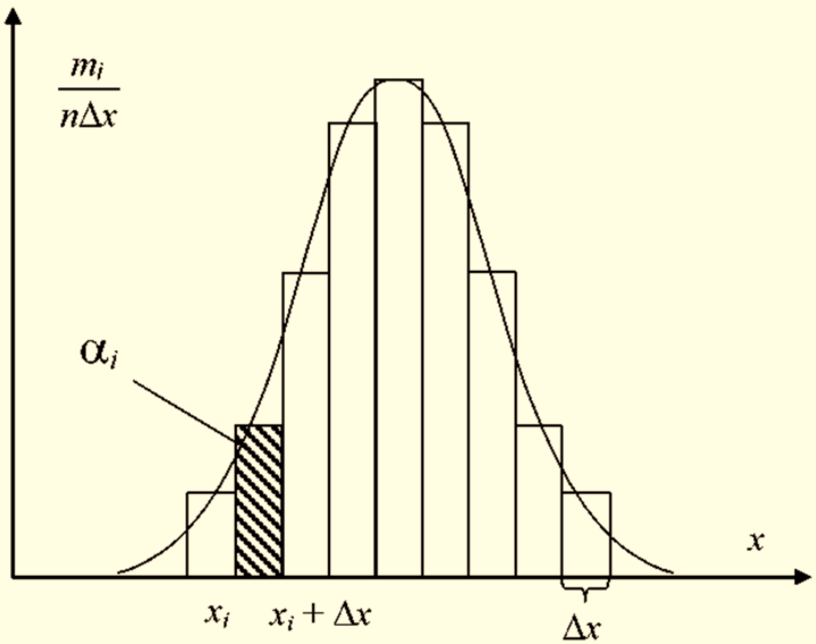
Над каждым промежутком $(x_i, x_i + \Delta x)$ рисуем прямоугольник высотой, равной частоте попадания результатов измерений в данный интервал, или величине

$$\frac{m_i}{n\Delta x}$$

Гистограмма

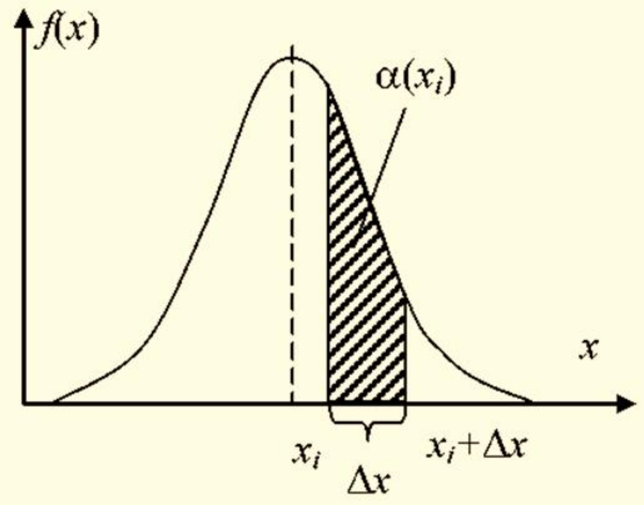
Построенный таким образом ступенчатый график называется **гистограммой** выборки (рис.). Такое частотное распределение позволяет наглядно показать исход серии измерений. Хотя результат каждого измерения определяется случайными причинами, из рисунка хорошо видно, что эта случайность подчиняется определенному закону.

При большом числе измерений n относительную частоту того, что величина x может принимать значения в интервале от x_i до $x_i + \Delta x$, называют **вероятностью.**



Гистограмма

Следует отметить, что при увеличении числа интервалов до ∞, длина
интервала Δx стремится к нулю.



Тогда гистограмма в предельном переходе заменится гладкой кривой $f(x)$, которая называется кривой распределения или **плотностью вероятности величины *x***. На рисунке – это кривая, симметричная относительно максимума.

Вероятность $\alpha(x_i)$ попадания результата измерения величины x в интервал от x_i до $x_i + \Delta x$ численно равна площади под кривой функции плотности вероятности на этом интервале (заштрихованный участок с основанием Δx на рис., которая вычисляется путем интегрирования функции плотности вероятности $f(x)$

Числовые характеристики (параметры) распределений случайных величин

Параметры распределения

Функции *F(x)* и *p*(*x*) несут о случайной величине одну и ту же информацию, но в разной форме. Для их определения необходимо проведение весьма длительных и кропотливых исследований и вычислений.

В большинстве случаев бывает достаточно охарактеризовать случайные величины с помощью ограниченного числа специальных параметров, основными из которых являются:

- **центр распределения;**
- **начальные и центральные моменты** и производные от них коэффициенты – **математическое ожидание (МО), дисперсия, среднее квадратическое отклонение (СКО), эксцесс, контрэксцесс и коэффициент асимметрии.**

Числовые характеристики дискретной СВ

Математическим ожиданием дискретной случайной величины называют сумму произведений всех ее возможных значений на их вероятности:

$$(\quad)$$

Математическое ожидание приближенно равно среднему арифметическому значений случайной величины:

$$(\quad) \quad -$$

Однако математическое ожидание не может в достаточной степени характеризовать случайную величину. На практике часто требуется оценить **рассеяние возможных значений случайной величины вокруг ее среднего значения.**

Дисперсия дискретной СВ

Дисперсией случайной величины X называется математическое ожидание квадрата ее отклонения от ее математического ожидания:

( ) [ ]

Дисперсия — это мера рассеяния случайной величины около ее математического ожидания. Если X — дискретная случайная величина, то дисперсию вычисляют по следующим формулам:

( ) ( )

или

( ) ( ) ( )

( )

Моменты распределения непрерывной СВ

Все моменты (параметры) распределения представляют собой некоторые средние значения, причем если усредняются величины, отсчитываемые от начала координат, то моменты называют **начальными**, а если от центра распределения, то **центральными.**

Начальные и центральные моменты *r-го* порядка определяются соответственно по формулам:

$$\alpha_r[x] = \int_{-\infty}^{+\infty} x^r p(x) dx \qquad \mu_r[x] = \int_{-\infty}^{+\infty} (x - m_x)^r p(x) dx$$

Нулевой начальный момент равен единице. Он используется для задания **условия нормирования** плотности распределения:

$$\int_{-\infty}^{+\infty} x^0 p(x) dx = 1$$

Моменты распределения непрерывной СВ

Математическое ожидание (первый начальный момент) непрерывной случайной величины определяется по формуле:

$$m_x = M[x] = \int_{-\infty}^{+\infty} xp(x)dx$$

Дисперсия (второй центральный момент) характеризует степень рассеяния значений случайной величины относительно среднего значения.

Дисперсия непрерывной случайной величины определяется как

$$D[x] = \int_{-\infty}^{+\infty} (x - m_x)^2 p(x)dx$$

Средним квадратичным отклонением называется квадратный корень из дисперсии

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

Моменты распределения непрерывной СВ

Третий центральный момент $\mu_3[x] = \int_{-\infty}^{+\infty} (x - m_x)^3 \, p(x) dx$

служит характеристикой асимметрии, или скошенности распределения.

С его использованием вводится **коэффициент асимметрии**:

$$v = \frac{\mu_3[x]}{\sigma^3}$$

Моменты распределения непрерывной СВ

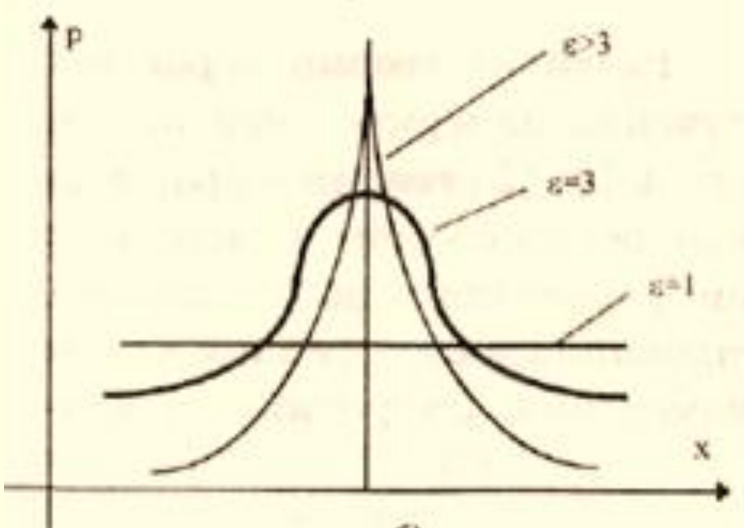
Четвертый центральный момент

$$\mu_4[x] = \int_{-\infty}^{+\infty} (x - m_x)^4 \, p(x) dx$$

служит для характеристики плоско– или островершинности распределения.

Эти свойства описываются с помощью **эксцесса**

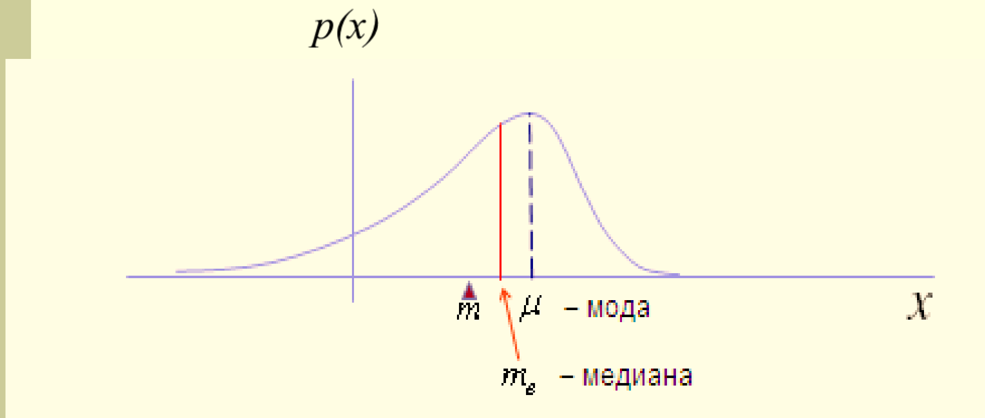
$$\varepsilon = \frac{\mu_4[x]}{\sigma^4}$$



б)

Числовые характеристики непрерывной СВ

Модой непрерывной случайной величины X называется ее наиболее вероятное значение (для которого плотность вероятности *p(x)* достигает максимума).



Медианой $M_e(x)$ непрерывной случайной величины X называется такое ее значение, для которого

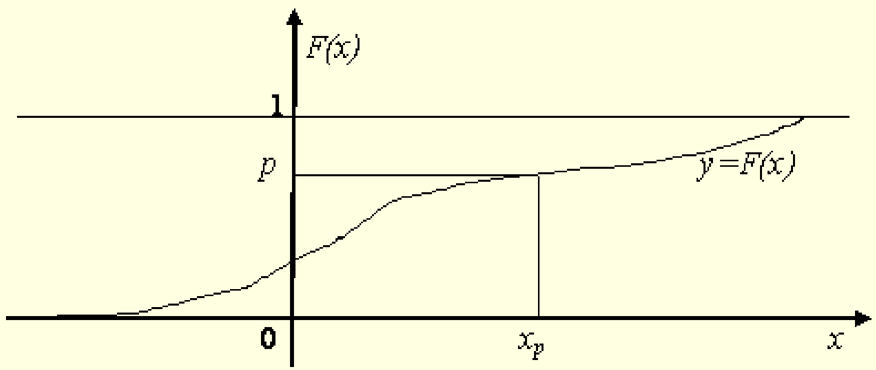
[ ] [ ] –

Квантили распределения

При нахождении доверительных границ для параметров распределений применяется такое понятие, как «квантиль порядка *p*»,

где $0 < p < 1$ (обозначается x_p).

Квантиль порядка *p* – значение случайной величины, для которого функция распределения принимает значение *p*.



Для непрерывных функций распределения, как правило, существует единственный квантиль x_p порядка *p* , причем

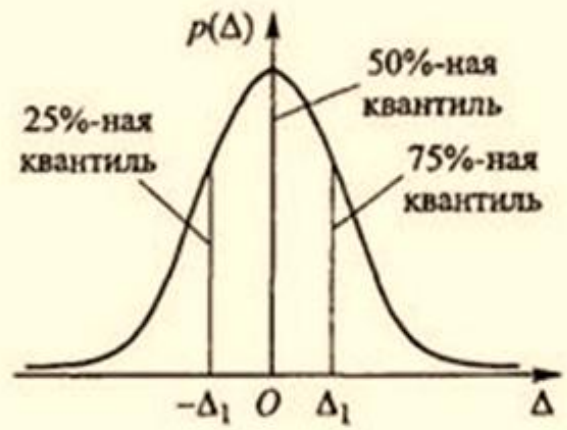
$$F(x_p) = p.$$

Квантили распределения

Квантилем уравнения α (или **α - квантилем**) называется такое значение x_a случайной величины X, при котором ее функция распределения принимает значение, равное a

$$F(x_\alpha) = P(X \le x_\alpha) = \alpha$$

Квантили, наиболее часто встречающиеся в практических задачах, имеют свои названия:



Медиана - квантиль уровня 0.5;
Нижняя квартиль - квантиль уровня 0.25;
Верхняя квартиль - квантиль уровня 0.75;
Децили - квантили уровней 0.1, 0.2, …, 0.9;
Процентили - квантили уровней 0.01, 0.02, …, 0.99.

Точечные оценки распределения

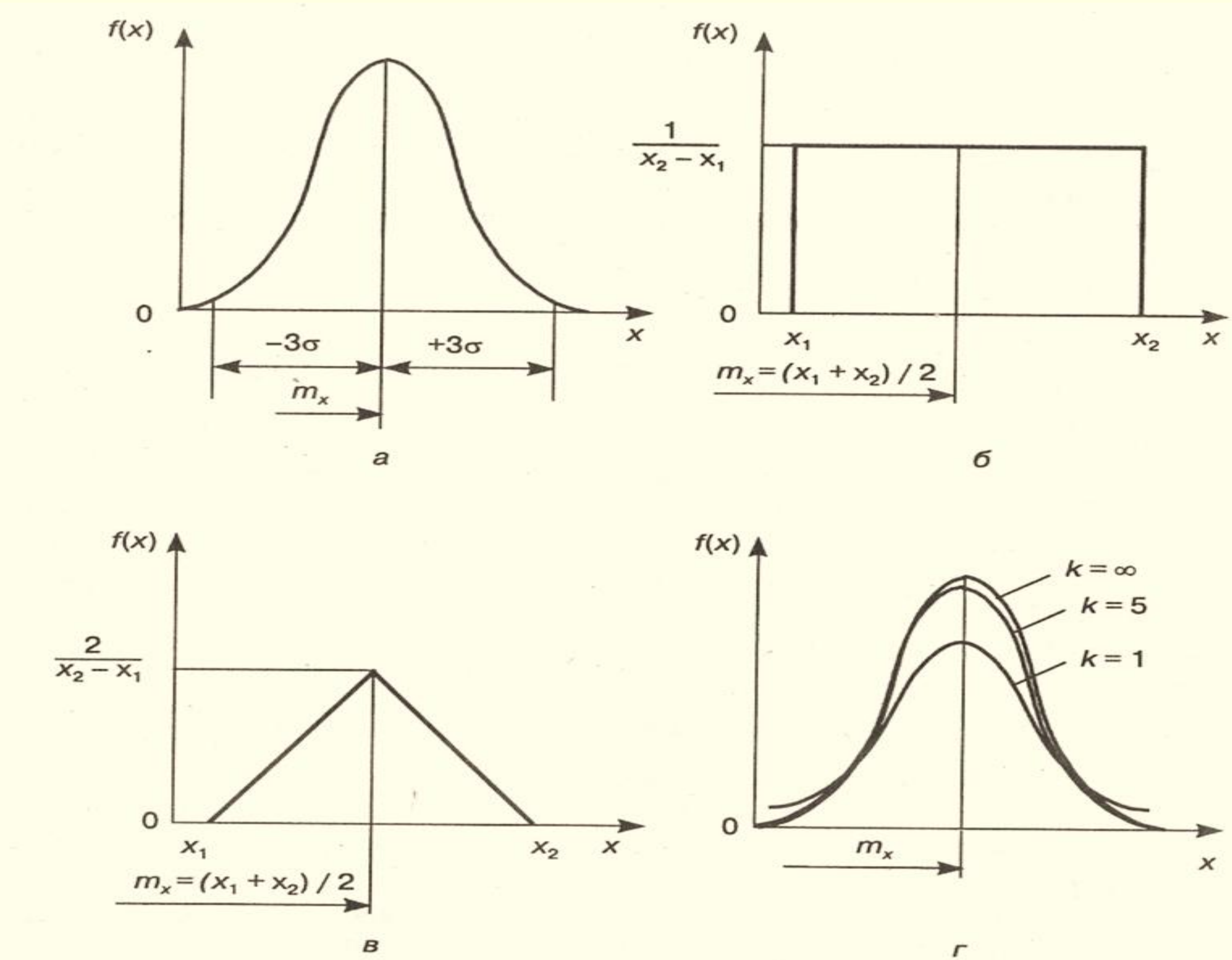
Точечные оценки могут быть состоятельными, несмещенными и эффективными.

Состоятельной называется оценка, которая при увеличении объема выборки стремится по вероятности к истинному значению числовой характеристики.

Несмещенной называется оценка, математическое ожидание которой равно оцениваемой числовой характеристике (параметру).

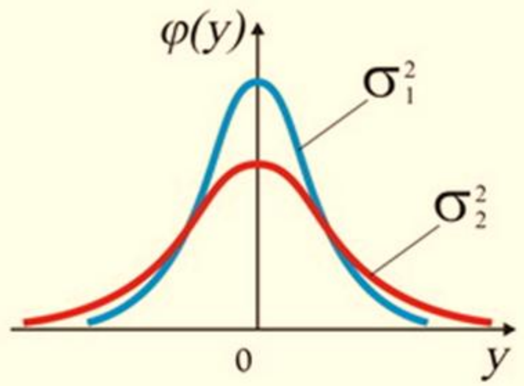
Оценка называется **эффективной**, если ее дисперсия меньше дисперсии любой другой оценки данного параметра, т.е. наиболее эффективной считают ту из нескольких возможных несмещенных оценок, которая имеет наименьшую дисперсию.

Законы распределения СВ



Закон нормального распределения (Гаусса)

Нормальный закон распределения (часто называемый законом Гаусса) играет исключительно важную роль в теории вероятностей и занимает среди других законов распределения особое положение.

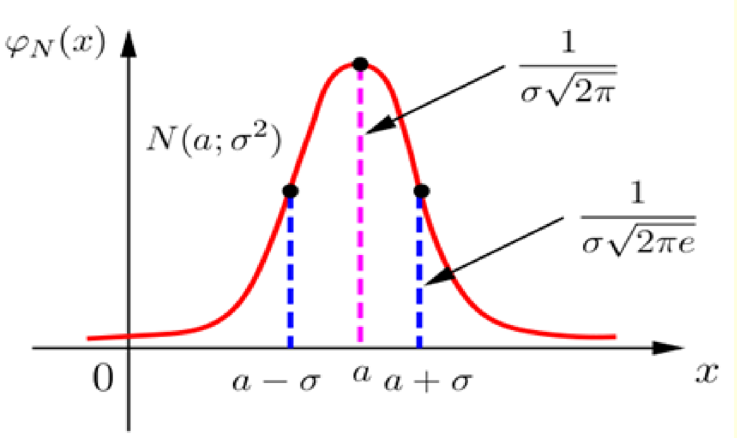


Это — наиболее часто встречающийся на практике закон распределения. Главная особенность, выделяющая нормальный закон среди других законов, состоит в том, что он является предельным законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.

Закон нормального распределения (Гаусса)

Случайную величину называют нормально распределенной, если ее плотность имеет вид

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



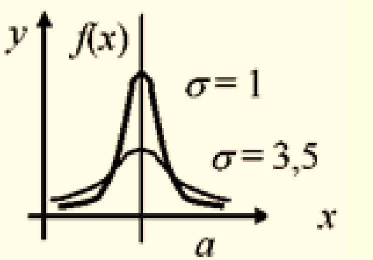
где a – математическое ожидание;

σ – среднее квадратическое отклонение

$$\sigma=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

С увеличением параметра нормальная кривая становится ниже, положе и шире.

С изменением ***а*** форма нормальной кривой не изменяется, только кривая смещается вправо, если значение увеличивается, или влево, если ***а*** уменьшается. Площадь под кривой плотности равна 1.

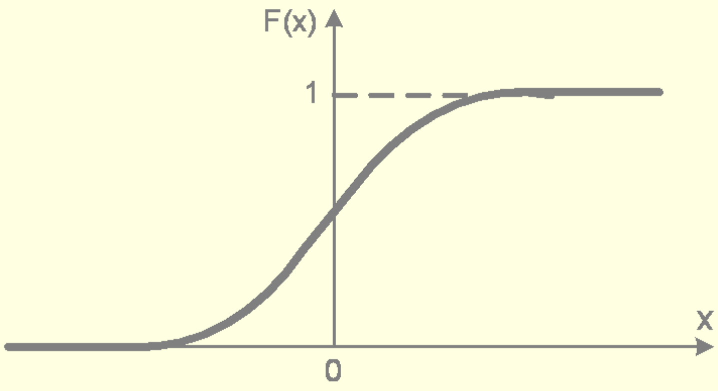


Закон нормального распределения (Гаусса)

Интегральная функция нормального распределения

$$F(x) = \int_0^x p(x)dx = \frac{1}{2} + \int_0^x f(x)dx$$

$$= \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right),$$



Следовательно, вероятность принятия значения из интервала **[x$_1$; x$_2$]** для нормального распределения случайной величины с параметрами ***a*** и **σ** может быть определена по формуле:

$$P(x_1 \le x \le x_2) = \Phi\left(\frac{x_2-a}{\sigma}\right) - \Phi\left(\frac{x_1-a}{\sigma}\right)$$

при помощи **таблицы значений функции Лапласа Ф(х)**.

Таблица значений функции Лапласа

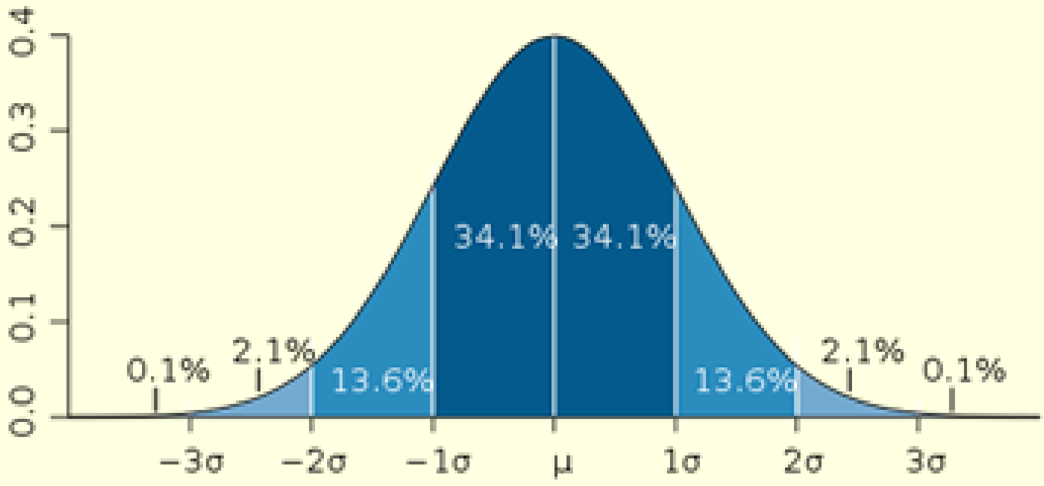
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \exp\left[\frac{(x-m)^2}{2\sigma^2}\right]$$

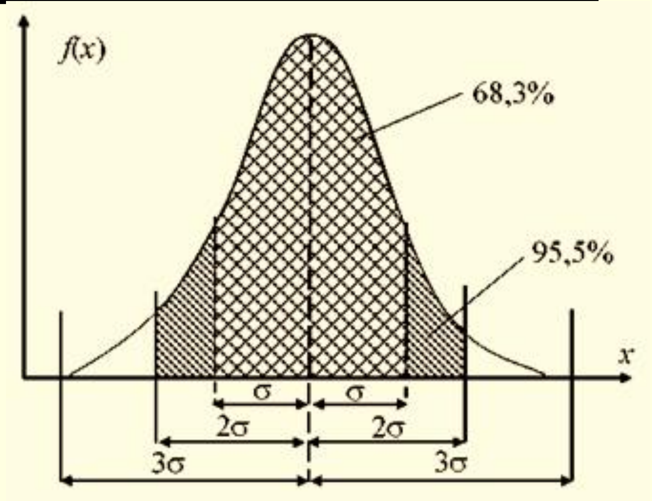
x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,00000	0,50	0,19146	1,00	0,34134	1,50	0,43319	2,00	0,47725	3,00	0,49865
0,01	0,00399	0,51	0,19497	1,01	0,34375	1,51	0,43448	2,02	0,47831	3,05	0,49886
0,02	0,00798	0,52	0,19847	1,02	0,34614	1,52	0,43574	2,04	0,47932	3,10	0,49903
0,03	0,01197	0,53	0,20194	1,03	0,34849	1,53	0,43699	2,06	0,48030	3,15	0,49918
0,04	0,01595	0,54	0,20540	1,04	0,35083	1,54	0,43822	2,08	0,48124	3,20	0,49931
0,05	0,01994	0,55	0,20884	1,05	0,35314	1,55	0,43943	2,10	0,48214	3,25	0,49942
0,06	0,02392	0,56	0,21226	1,06	0,35543	1,56	0,44062	2,12	0,48300	3,30	0,49952
0,07	0,02790	0,57	0,21566	1,07	0,35769	1,57	0,44179	2,14	0,48382	3,35	0,49960
0,08	0,03188	0,58	0,21904	1,08	0,35993	1,58	0,44295	2,16	0,48461	3,40	0,49966
0,09	0,03586	0,59	0,22240	1,09	0,36214	1,59	0,44408	2,18	0,48537	3,45	0,49972
0,10	0,03983	0,60	0,22575	1,10	0,36433	1,60	0,44520	2,20	0,48610	3,50	0,49977

Правило трех сигм

Вероятности попадания значений нормально распределённой случайной величины в некоторые интервалы приведены в таблице

Границы промежутка	Площадь под кривой плотности
Односигмовые [*a - σ; a + σ*]	**0,6827**
Двухсигмовые [a - 2σ; a + 2σ]	**0,9545**
Трёхсигмовые [a - 3σ; a + 3σ]	**0,9973**





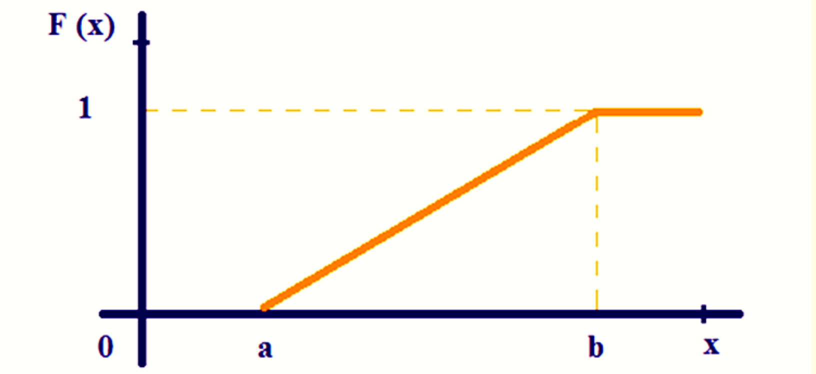
Правило трех σ:* площадь, лежащая за пределами трехсигмового интервала, равна 1 - 0,9973 = 0,0027. *Вероятность того, что случайная величина выйдет за его пределы, весьма близка к нулю.

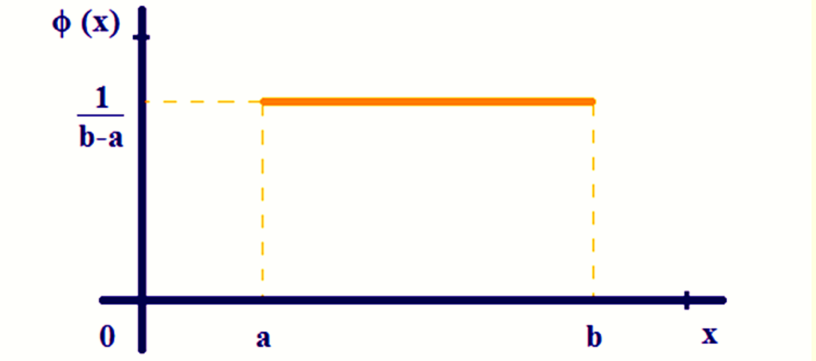
Равномерный закон распределения

Если плотность вероятности ϕ(x) есть величина постоянная на определенном промежутке [a,b], то закон распределения называется **равномерным.**

$$\phi(x) = \begin{cases} \dfrac{1}{b-a} & \text{при } a \le x \le b \\ 0 & \text{при } x < a \quad x > b \end{cases}$$

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



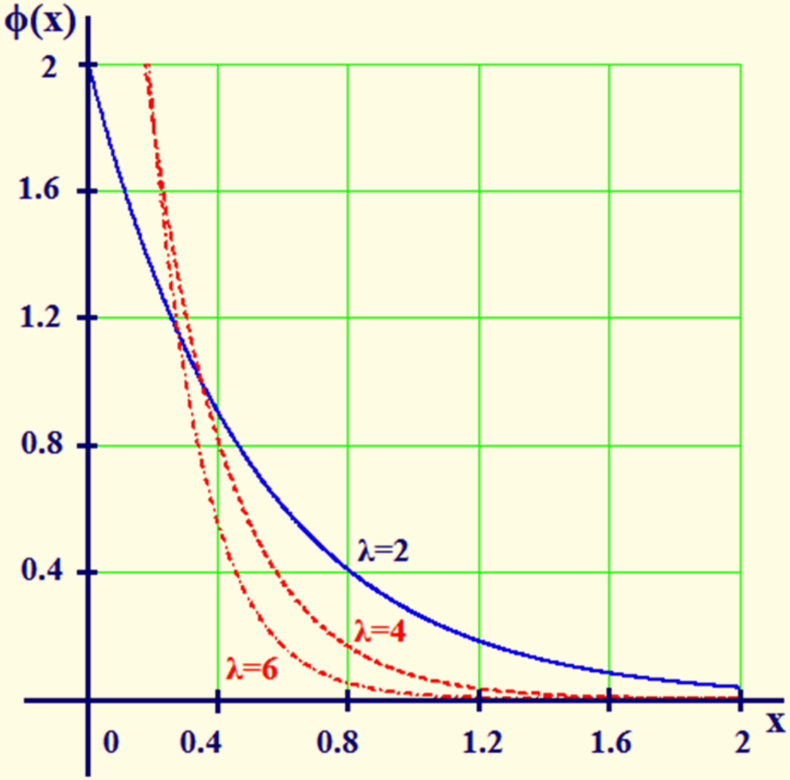


Показательный (экспоненциальный) закон

Закон распределения случайной величины X называется показательным (или экспоненциальным), если плотность вероятности имеет вид:

$$\phi(x) = \begin{cases} \lambda e^{-\lambda x} & \text{при } x \geq 0 \\ 0 & \text{при } x < 0 \end{cases}$$

где λ - параметр обратно-пропорциональный математическому ожиданию.

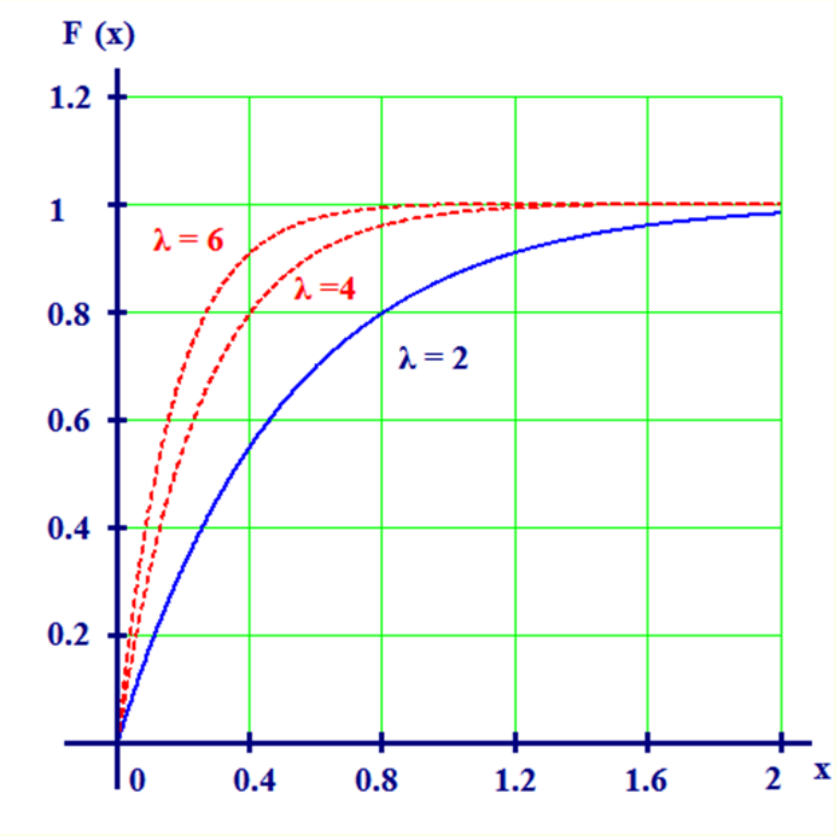


На рисунке изображен график плотности вероятности с параметрами $\lambda = 2$, $\lambda = 4$, $\lambda = 6$ и 8.

Показательный (экспоненциальный) закон

Функция распределения случайной величины X, которая имеет показательное распределение, имеет вид:

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



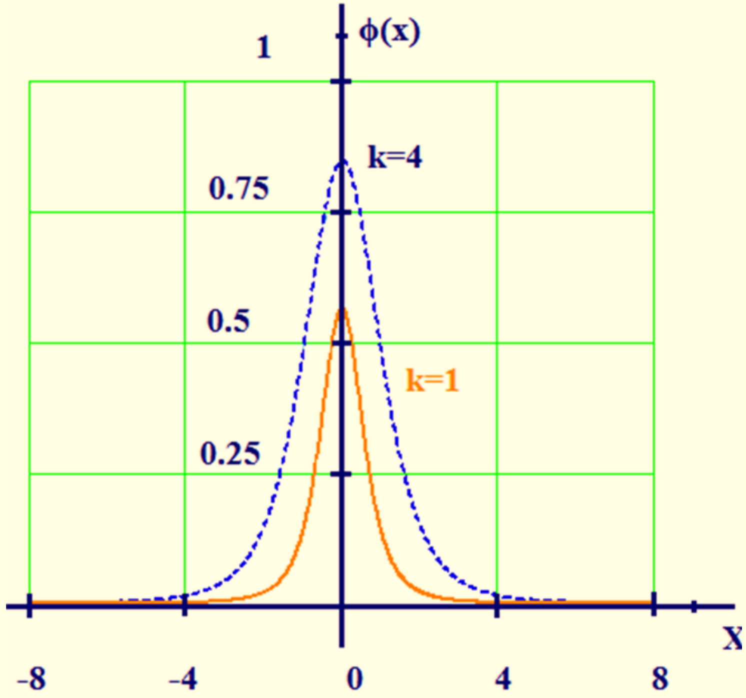
Если функцию распределения случайной величины выразить через плотность вероятности при x ≥ a, то она примет вид:

$$F(x) = \int_0^x \lambda e^{-\lambda x}\, dx$$

Распределение Стьюдента (t - распределение)

Распределение непрерывной случайной величины называется **распределением Стьюдента**, если плотность вероятности имеет вид:

$$\phi(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}}$$



где

$\Gamma(y)$ - гамма-функция

k - число степеней свободы

На рисунке изображена плотность вероятности распределения Стьюдента. Из графика можно увидеть, что чем больше k, тем больше кривая приближается к нормальному распределению.

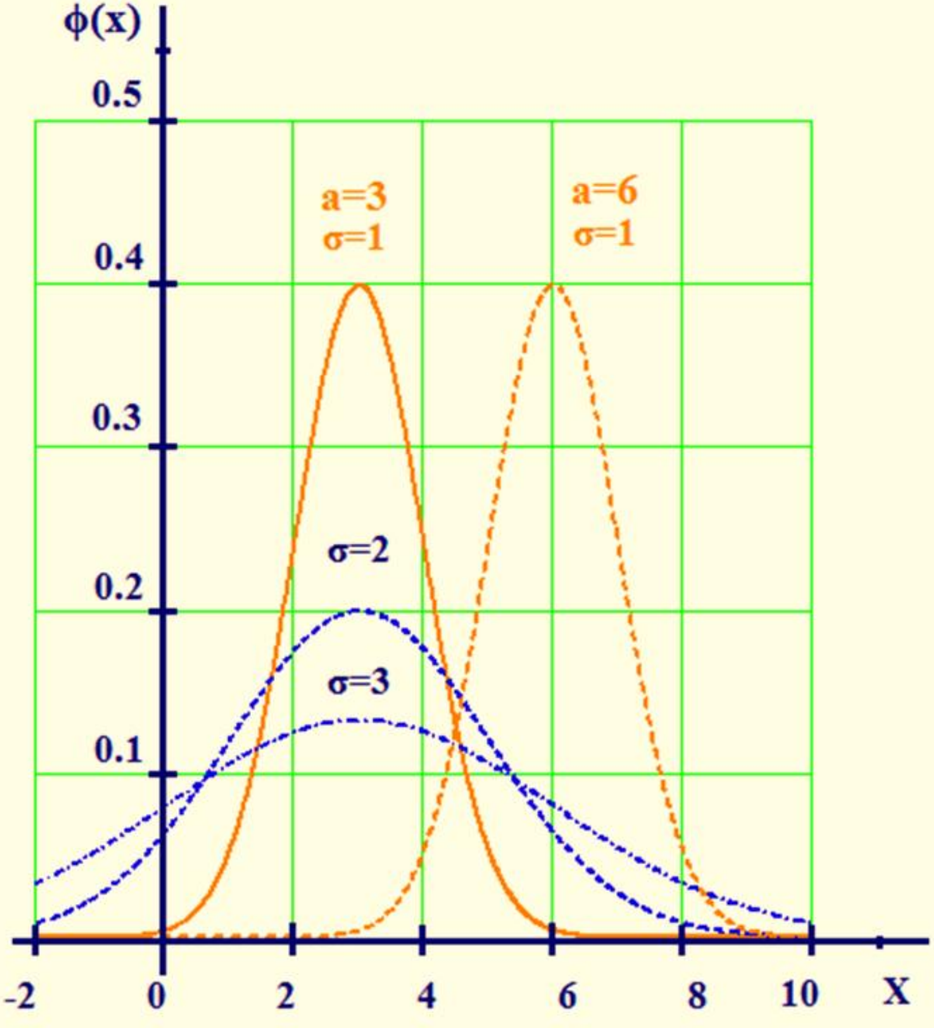
Параметры нормального распределения

Показатель	Нормальный закон распределения	Примечание
Плотность вероятности	$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{\frac{-(x-m_x)^2}{2\sigma_x^2}}$	где m_x - математическое ожидание , σ_x – среднее квадратическое отклонение.
Функция распределения	$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{t^2}{2}} dt$	$u = \frac{x-m_x}{\sigma_x}$
Вероятность попадания в интервал (a;b)	$P(a \leq X \leq b) = \Phi\left(\frac{b-m_x}{\sigma_x}\right) - \Phi\left(\frac{a-m_x}{\sigma_x}\right)$	$\Phi(u) = F(u) - \frac{1}{2} =$ $= \frac{1}{\sqrt{2\pi}} \int_{0}^{u} e^{-\frac{t^2}{2}} dt$ **интегральная функция Лапласа**
Математическое ожидание	$m_x = \int\limits_{-\infty}^{\infty} xf(x)dx.$	
Дисперсия	$\sigma^2 = \int\limits_{-\infty}^{\infty} (\bar{x} - x)^2 f(x)dx,$	

Параметры распределений

Показатель	Равномерный закон распределения	Показательный закон распределения
Вероятность попадания в интервал	$P(x_1 < X < x_2 = \frac{x_2 - x_1}{b - a}$	$P(a < X < b) = \int_a^b \lambda e^{-\lambda x} dx == e^{-\lambda e P} - e^{-\lambda b}$
Математическое ожидание	$M(X) = \frac{a + b}{2}$	$M(X) = \frac{1}{\lambda}$
Дисперсия	$D(X) = \frac{(b - a)^2}{12}$	$D(X) = \frac{1}{\lambda^2}$
Среднее квадратическое отклонение	$\sigma(X) = \frac{b - a}{2\sqrt{3}}$	$\sigma(X) = \frac{1}{\lambda}$

Какой закон распределения?



Пример решения задачи

Случайная величина имеет нормальное распределение с математическим ожиданием - 15 и средним квадратическим отклонением - 5 . Найти вероятность того, что случайная величина примет значение, принадлежащее интервалу [10;30].

Для нормально распределённой случайной величины вероятность попадания в интервал $[x_1;x_2]$

$$P(\mathrm{X1} < X < \mathrm{X2}) = \Phi\left(\frac{\mathrm{X2} - a}{\sigma}\right) - \Phi\left(\frac{\mathrm{X1} - a}{\sigma}\right)$$

где a – математическое ожидание, σ – среднее квадратическое отклонение.

По условию $a = 15$, $\sigma=5$, $\mathrm{X1} = 10, \mathrm{X2} = 30$. Следовательно,

$$P(10 < X < 30) = \Phi\left(\frac{30 - 15}{5}\right) - \Phi\left(\frac{10 - 15}{5}\right) = \Phi(3) - \Phi(-1) = 0{,}83995$$

Ответ: $P(10 < X < 30) = 0{,}83995$.

Грубые погрешности (промахи)

Грубые погрешности (промахи)

Грубые погрешности (промахи)

Под промахом понимается значение погрешности, отклонение которого от центра распределения существенно превышает значение, оправданное объективными условиями измерения.

Поэтому с точки зрения теории вероятности появление промаха маловероятно.

Причинами грубых погрешностей могут быть неконтролируемые изменения условий измерений, неисправность, ошибки оператора и др.

Для исключения грубых погрешностей применяют аппарат проверки статистических гипотез (статистические критерии).

*Рассмотрим следующие критерии проверки подозрительных (с точки зрения погрешностей) результатов наблюдений: **Ирвина, Романовского, вариационного размаха, Диксона, Смирнова, Шовене**.*

Критерий Ирвина

Для полученного ряда экспериментальных данных (выборки) определяют коэффициент **λ** по формуле:

$$\lambda = \frac{(x_{n+1} - x_n)}{S},$$

где x_{n+1}, x_n – наибольшие значения случайной величины; S – среднее квадратическое отклонение, вычисленное по всем значениям выборки.

Затем этот коэффициент сравнивается с табличным значением $\boldsymbol{\lambda_q}$,

Если $\lambda > \lambda_q$, то нулевая гипотеза не подтверждается, т. е. результат ошибочный, и он должен быть исключен при дальнейшей обработке результатов наблюдений.

Критерий Романовского

Гипотеза о наличии грубых погрешностей в подозрительных результатах подтверждается, если выполняется неравенство:

$$\left|\overline{X}_{ц.р.} - x_{iпод}\right| \geq t_p S,$$

где t_p- квантиль распределения Стьюдента при заданной доверительной вероятности с числом степеней свободы $k = n - k_n$ (k_n - число подозрительных результатов наблюдений).

Квантиль t_p распределения Стьюдента берется из таблиц.

Точечные оценки распределения и СКО S результатов наблюдений вычисляется без учета k_n подозрительных результатов наблюдений.

Критерий вариационного размаха

Критерий вариационного размаха является одним из простых методов исключения грубой погрешности измерений (промаха). Для его использования определяют размах вариационного ряда упорядоченной совокупности наблюдений *($x_1 \le x_2 \le \ldots \le x_k \le \ldots \le x_n$):*

$$\overline{X} - z \cdot R_n \prec x_k \prec \overline{X} + z \cdot R_n$$

где X - выборочное среднее арифметическое значение, вычисленное после исключения предполагаемого промаха; z - критериальное значение, взятое из таблицы.

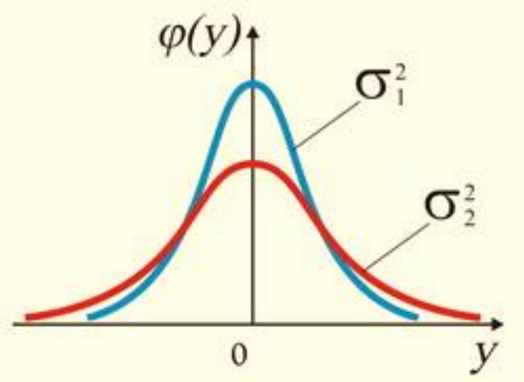
n	5	6	7	8-9	10-11	12-15	16-22	23-25	26-63	64-150
z	1,7	1,6	1,5	1,4	1,3	1,2	1,1	1,0	0,9	0,8

Гипотезу (об отсутствии грубой погрешности) принимают, если указанное неравенство выполняется. Если x_k не удовлетворяет этому условию, то этот результат исключают из вариационного ряда.

Критерии “Трех сигм”или “3S”

Критерий “трех сигм” является одним из простейших для проверки результатов, подчиняющихся нормальному закону распределения.

Сущность правила трех сигм: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.



$$\left| x_{iпод} - \overline{X}_{ц.р.} \right| \geq 3S$$

Этому критерию аналогичен **критерий Райта**, основанный на том, что если остаточная погрешность больше четырех сигм, то этот результат измерения является грубой погрешностью и должен быть исключен при дальнейшей обработке. Оба критерия надежны при числе измерений больше 20…50.

Критерий Смирнова

Критерий Смирнова используется при объемах выборки $n \geq 25$, при известных значениях генеральных среднего и *СКО*. Он устанавливает менее жесткие границы грубой погрешности. Для реализации этого критерия вычисляются действительные значения квантилей распределения (наблюдаемое значение критерия) по формуле:

$$\beta = \frac{\max\left|x_{inod} - \overline{X}\right|}{S}.$$

Найденное значение сравнивается с предельным критериальным βk , приведенным в таблице.

Объем выборки n	Предельное значение β_k при уровне значимости q				
	0,100	0,050	0,0010	0,005	0,001
1	2	3	4	5	6
1	1,282	1,645	2,326	2,576	3,090
2	1,632	1,955	2,575	2,807	3,290
3	1,818	2,121	2,712	2,935	3,403
4	1,943	2,234	2,806	3,023	3,481
5	2,036	2,319	2,877	3,090	3,540
6	2,111	2,386	2,934	3,143	3,588
7	2,172	2,442	2,981	3,188	3,628
8	2,224	2,490	3,022	3,227	3,662
9	2,269	2,531	3,057	3,260	3,692
10	2,309	2,568	3,089	3,290	3,719
15	2,457	2,705	3,207	3,402	3,820
20	2,559	2,799	3,289	3,480	3,890
25	2,635	2,870	3,351	3,539	3,944
30	2,696	2,928	3,402	3,587	3,988
40	2,792	3,015	3,480	3,662	4,054
50	2,860	3,082	3,541	3,716	4,108
100	3,076	3,285	3,723	3,892	4,263
250	3,339	3,534	3,946	4,108	4,465
500	3,528	3,703	4,108	4,263	4,607

Критерий Шовене

Критерий Шовене строится на определении числа ожидаемых результатов наблюдений $n_{ож}$, которые имеют столь же большие погрешности, как и подозрительный. Гипотеза о наличии грубой погрешности принимается, если выполняется условие:

$$n_{ож} \le 0{,}5.$$

Сначала рассчитывают

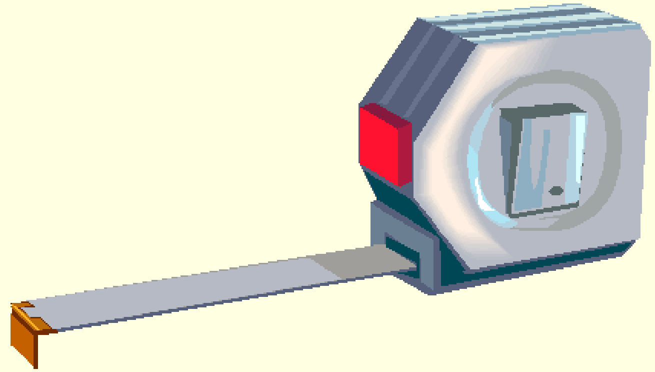
$$z = \frac{\left| x_{iпод} - \overline{X}_{ц.р.} \right|}{S}$$

Из таблицы нормированного нормального распределения определяется вероятность появления подозрительного результата в генеральной совокупности чисел n:

$$P\left(z \cdot S \prec \left| x_{iпод} - \overline{X}_{ц.р.} \right|\right);$$

Число ожидаемых результатов $n_{ож}$ определяется по формуле:

$$n_{ож} = n \cdot P.$$



БЛАГОДАРЮ
ЗА ВНИМАНИЕ