Интервальные оценки

Проблема № 1: **получить численное значение определенной физической величины А экспериментальными методами**.

Исходной информацией является серия измерений неизвестной величины **А**:

$$x_1, x_2, \ldots x_n \qquad (1)$$

В математической статистике серия измерений рассматривается как **выборка** из генеральной совокупности.

Согласно **постулату 2**, измеряемая величина **А** совпадает с генеральным средним $\boldsymbol{M_G}$

Генеральная совокупность является бесконечным множеством значений, характеризуемое ***распределением вероятностей***.

Существуют дискретные и непрерывные генеральные совокупности.

Генеральная совокупность обладает числовыми характеристиками, важнейшие из которых:

генеральное среднее M_G

генеральная дисперсия D_G

Точечные оценки измеряемой величины

Среднее статистическое

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

Выборочная медиана m^*

Состоятельной, несмещенной и эффективной оценкой генерального среднего $\boldsymbol{M_G}$ нормальной генеральной совокупности является **среднее статистическое**.

Так как $\bar{x}$ − случайная величина, равенство

$$\bar{x} \approx M_G$$

является приближенным.

Дисперсия среднего статистического равна $\boldsymbol{D_G / n}$

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

несмещенная оценка генеральной дисперсии

$$\frac{s}{\sqrt{n}}$$

наилучшая оценка **среднеквадратичного отклонения среднего статистического**

Можно ли записать результат измерений в виде :

$$A = \bar{x} \pm \frac{s}{\sqrt{n}}$$

???

Доверительный интервал для измеряемой величины А

(для генерального среднего)

$$P\left(\bar{x} - t\frac{s}{\sqrt{n}} < M_G < \bar{x} + t\frac{s}{\sqrt{n}}\right) = \eta \qquad (3)$$

где **η** – доверительная вероятность.

t – пока неизвестное число

Величина доверительной вероятности выбирается на основе **принципа практической достоверности**

Уравнение (3) означает, что две случайные величины

$$\bar{x} - t\frac{s}{\sqrt{n}} \qquad \text{и} \qquad \bar{x} + t\frac{s}{\sqrt{n}}$$

с вероятностью **η** ограничивают постоянную, но неизвестную величину $\boldsymbol{M_G = A}$.

безразмерное число $\boldsymbol{t}$ –

коэффициент Стьюдента.

Случайное события

$$\bar{x} - t\frac{s}{\sqrt{n}} < A < \bar{x} + t\frac{s}{\sqrt{n}}$$

и

$$-t < \frac{\bar{x} - A}{s/\sqrt{n}} < t$$

эквивалентны

Это означает, что случайная величина

$$\frac{\bar{x} - A}{s / \sqrt{n}}$$

с вероятностью **η** попадает в интервал, ограниченный числами $-t$ и t .

В теории вероятностей доказано, что если генеральная совокупность **нормальная** (гауссова), то случайная величина

$$T_{n-1} = \frac{\bar{x} - M_G}{s/\sqrt{n}}$$

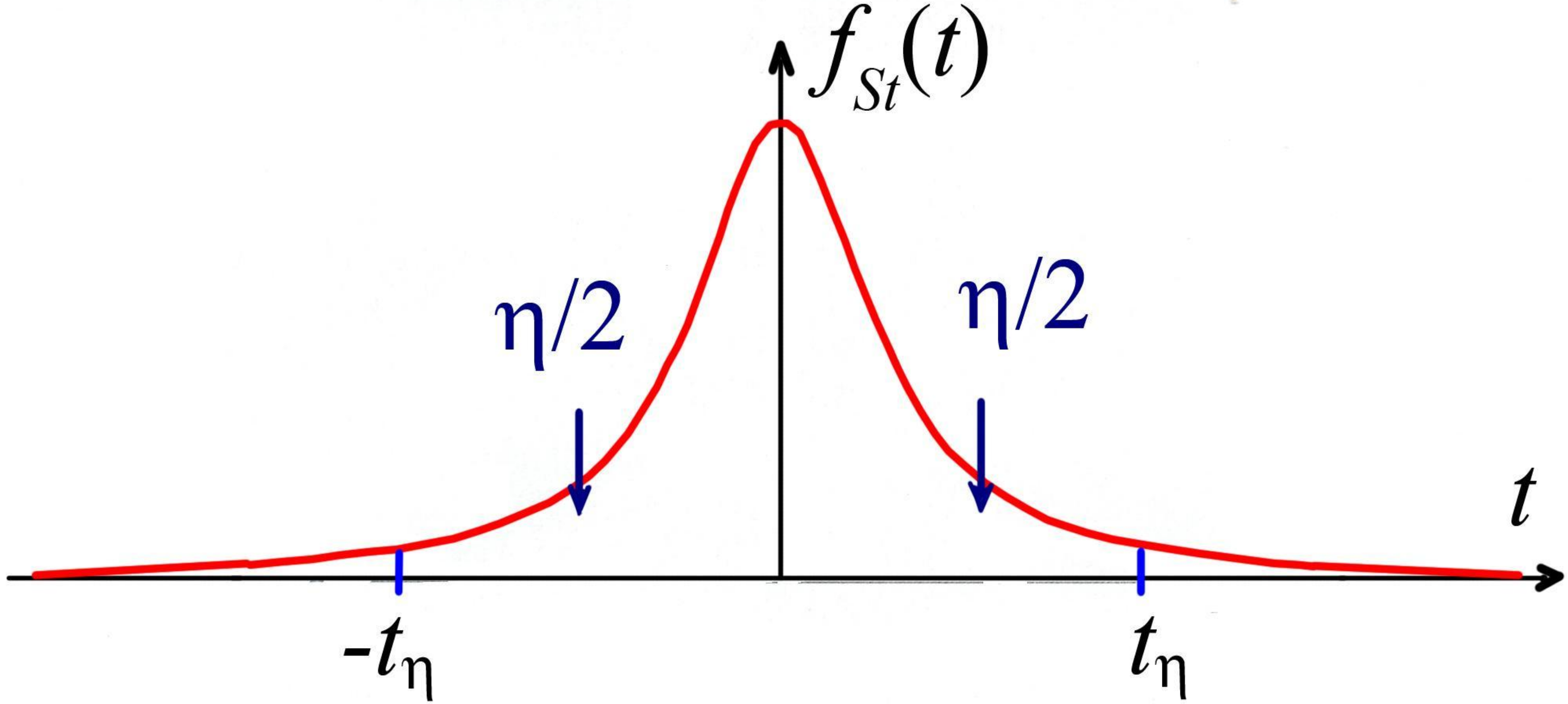
имеет **распределение Стьюдента** с числом степеней свободы $\boldsymbol{\nu = n - 1}$.

Следовательно, числа $\pm t$ представляют собой симметричные пределы интегрирования плотности распределения Стьюдента, соответствующие вероятности η.

$$\int_{-t}^{t} f_{St}(x)\ dx = \eta$$

где f_{St} – функция плотности случайной величины с **распределением Стьюдента**.

η – выбранная доверительная вероятность.



t_η – квантиль для вероятности **(1 + η) / 2**.

Процедура построения доверительного интервала для измеряемой величины.

По данной выборке вычисляются **среднее статистическое** $\bar{x}$ и **несмещенная оценка дисперсии** s^2.

По заданной **доверительной вероятности** η и **числу степеней свободы** $\nu = n - 1$ извлекается из таблицы значение **коэффициента Стьюдента** $t_{\nu,\eta}$.

Рассчитываются границы **доверительного интервала**:

$$\bar{x} - \frac{s}{\sqrt{n}} t_{\nu,\eta} \; ; \; \bar{x} + \frac{s}{\sqrt{n}} t_{\nu,\eta}$$

Доверительный интервал для генеральной дисперсии

Если генеральная совокупность имеет **нормальное** (гауссово) распределение, то случайная величина

$$\chi_{\nu}^{2} = \frac{(n-1)s^{2}}{D_{G}}$$

имеет распределение вероятностей «**хи-квадрат**» с $\boldsymbol{\nu = n - 1}$ степенями свободы

$\boldsymbol{D_G}$ – генеральная дисперсия, $\boldsymbol{n}$ - объем выборки

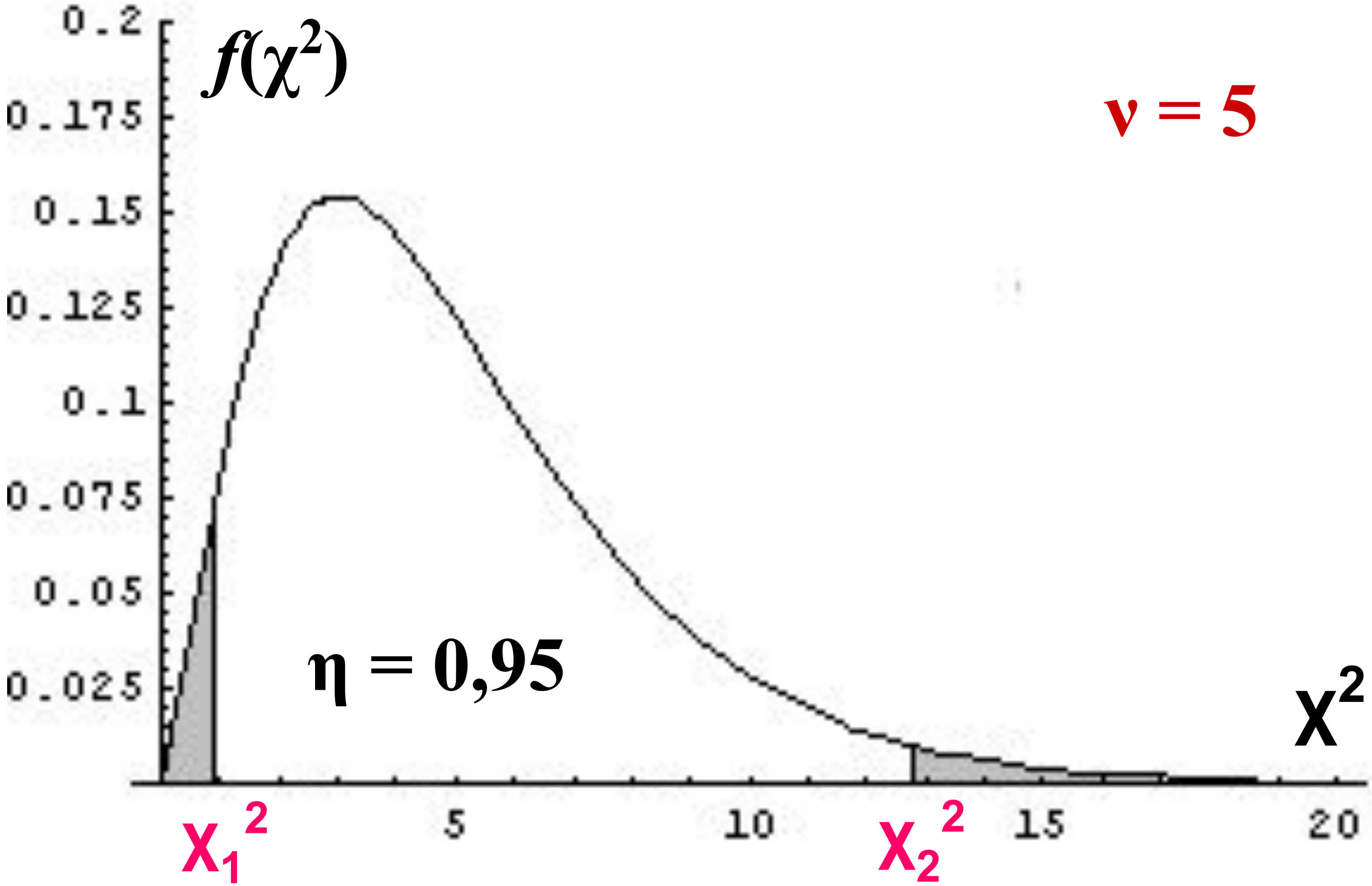
Определим числа χ_1^2 и χ_2^2 следующими уравнениями для вероятностей событий

$$P\left[\frac{(n-1)s^2}{D_G} < \chi_1^2\right] = \frac{1-\eta}{2} \qquad (12)$$

$$P\left[\frac{(n-1)s^2}{D_G} > \chi_2^2\right] = \frac{1-\eta}{2} \qquad (13)$$

Согласно уравнению (12), вероятность того, что значение случайной величины χ_ν^2 ***не превысит*** числа χ_1^2, равна **(1 – η) / 2**.

Согласно уравнению (13), вероятность того, что значение той же случайной величины χ_ν^2 ***превышает*** число χ_2^2 , равна также **(1 – η) / 2**.



Заштрихованы области под кривой, площадь которых равна **(1 – η) / 2**

Вероятностное уравнение, эквивалентное уравнениям (12) и (13) имеет вид :

$$P\left[\frac{(n-1)s^2}{\chi_2^2} < D_G < \frac{(n-1)s^2}{\chi_1^2}\right] = \eta$$

χ_1^2 – квантиль для вероятности **(1 - η) / 2**.,

χ_2^2 – квантиль для вероятности **(1 + η) / 2**.

Иначе говоря, доверительной вероятности **η** соответствует следующий доверительный интервал для генеральной дисперсии

$$\frac{(n-1)s^2}{\chi_2^2} < D_G < \frac{(n-1)s^2}{\chi_1^2} \qquad (14)$$

Величины χ_1^2 и χ_2^2 извлекаются из таблиц квантилей случайной величины «**хи-квадрат**», составленными для различных доверительных вероятностей **η** и определённых чисел степеней свободы **ν**.

Наилучшей оценкой генеральной дисперсии является

$$D_G \approx s^2$$

где

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Извлечение квадратного корня из всех частей неравенств (4) дает доверительный интервал для **генерального среднеквадратичного**

$$\sigma_G = \sqrt{D_G}$$

$$\frac{s\sqrt{n-1}}{\chi_2} < \sigma_G < \frac{s\sqrt{n-1}}{\chi_1}$$