

Обработка результатов измерения одной величины.

Измерения, проводимые в опытах эксперимента, сопровождаются ошибками, ввиду конечной точности приборов и не идеальности условий эксперимента. Ошибки делятся на три типа.

- Систематические
- Грубые
- Случайные

Из-за наличия ошибок, точное значение измеряемой величины y^* установить не удастся. Поэтому при n повторных измерений одной и той же величины y^* получают серию различных результатов $y_1, y_2, y_3, \dots, y_n$ и наиболее вероятной оценкой измеряемой величины y^* будет являться среднее значение результатов серии.

$$\bar{y} = \frac{y_1 + y_2 + y_3 + \dots + y_n}{n} = \frac{\sum_{i=1}^n y_i}{n} \quad \text{xsr} = \text{mean}(y);$$

Замена точного значения измеряемой величины y^* значением \bar{y} влечёт ошибку, значение которой точно указать нельзя, а можно определить приближенно с необходимой доверительной вероятностью β . Нам надо определить величину ε_β в неравенстве

$$\left| y^* - \bar{y} \right| \leq \varepsilon_\beta$$

Очевидно, ε_β будет тем больше, чем с больше вероятностью β мы будем её определять, чем грубее был проведен эксперимент и чем меньше n (количество опытов в серии измерений).

Для оценки качества измерений, вводят понятие дисперсии, которая вычисляется по формуле:

$$S_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{f}, \quad \text{где } f - \text{число степеней свободы, } f = n - 1 \quad s2x=\text{var}(y);$$

Среднеквадратичным отклонением или стандартом называют величину:

$$S_y = \sqrt{S_y^2} \quad \text{standx}=\text{std}(y);$$

Для определения, является ли измеренное значение грубой ошибкой, можно воспользоваться U критерием:

$$U_{\text{расч}} = \frac{|y^{\text{под}} - \bar{y}|}{\sqrt{S_y^2 \cdot \frac{n-1}{n}}}$$

Если $U_{\text{расч}} > U_{p,f}$, то подозреваемое значение с вероятностью β является грубой ошибкой. Грубая ошибка исключается из серии. Критерий $U_{p,f}$ определяется из табл. 1 при уровне значимости $p = 1 - \beta$ и числе степеней свободы $f = n - 2$.

Таблица 1

f/p	0.05	0.01
1	1.412	1.414
2	1.689	1.723
3	1.869	1.955
4	1.996	2.130
5	2.093	2.265
6	2.172	2.374

function $x=U(p,f)$
 $tr=t(2*p/(f+2),f);$
 $x=tr*sqrt((f+1)/(f+tr^2));$

В статистике доверительную ошибку вычисляют по формуле:

$$\varepsilon_{\beta} = t_{p,f} \cdot \sqrt{\frac{S_y^2}{n}}$$

где $t_{p,f}$ – критерий Стьюдента, который определяется из табл. 2 при $p = 1 - \beta$ и $f = n - 1$.

Таблица 2

f/p	0.10	0.05	0.01
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.01	2.57	4.03
6	1.94	2.45	3.78

function $x=t(p,f)$
 $x=тинв(1-p/2, f);$

Интервал, который с доверительной вероятностью β покрывает точное значение y^* определяется, значением ε_{β} и называется доверительным и определяется как:

$$\bar{y} - \varepsilon_{\beta} \leq y^* \leq \bar{y} + \varepsilon_{\beta}$$

Пример: $p = 0.05$ $\beta = 0.95$ $n = 6$

i	1	2	3	4	5	6
y_i	6.28	6.47	6.54	7.02	6.45	6.40

$$\bar{y} = 39.16/6 = 6.527 \quad S_y^2 = 0.0659 \quad U^{\text{таб}} \text{ для } f = 6-2 = 4 \quad p = 0.05 \text{ имеет значение } 1.996$$

Подозреваемое значение $y_4 = 7.02$ т.к. $|7.02 - 6.527| = 0.493$ максимальна

$$U_{\text{расч.}} = \frac{|7.02 - 6.527|}{\sqrt{0.0659 * \frac{6-1}{6}}} = 2.105$$

$2.105 > 1.996$ поэтому $x_4 = 7.02$ является грубой ошибкой и удаляется из серии $n = 5$

i	1	2	3	4	5
y_i	6.28	6.47	6.54	6.45	6.40

$$\bar{y} = 32.14 / 5 = 6.428 \quad S_y^2 = 0.0094 \quad U^{\text{таб}} \text{ для } f = 5-2 = 3 \quad p = 0.05 \text{ имеет значение } 1.869$$

Подозреваемое значение $x_1 = 6.28$ т.к. $|6.28 - 6.428| = 0.148$ максимальна

$$U_{\text{расч.}} = \frac{|6.28 - 6.428|}{\sqrt{0.0659 * \frac{5-1}{5}}} = 1.709 < 1.869 \text{ поэтому } y_1 = 6.28 \text{ не является грубой ошибкой.}$$

Для последней серии строим доверительный интервал

$$t_{0.05, 4}^{\text{таб}} = 2.78 \quad \varepsilon_{\beta} = 2.78 * \sqrt{\frac{0.0094}{5}} = 0.12 \quad 6.308 < y^* < 6.548$$

