Вычисление стандартного отклонения и размаха для логарифмического нормального распределения

На практике аналитик часто встречается с ситуацией, когда в результате проверки на корреляцию оказывается, что экспериментальные данные не могут быть описаны линейной зависимостью. В этом случае пытаются преобразовать результаты в удобную для статистической обработки форму. Часто целесообразным является использование полулогарифмического или логарифмического преобразования.

Этот приём обычно применяется:

- при анализе очень малых концентраций (анализ следов);
- при проведении анализа в очень широкой области концентраций (несколько десятков процентов);
- при большой случайной ошибке (например, при полуколичественном спектральном анализе) и т.д.

Если имеет место зависимость вида $y=bx^a$, то после логарифмирования будем иметь линейную зависимость

$$Y=aX+B$$

где $Y=\lg y$, $X=\lg x$, $B=\lg b$.

Далее обработка результатов проводится по методике, обработки прямых линий. После определения $X_к \pm \Delta X_к$ для пробы К с неизвестной концентраций необходимо перейти от логарифмов к реальным значениям параметра x. Поскольку $X_к \pm \Delta X_к = \lg x \pm \lg \Delta x$, то это соответствует интервалу от $x \cdot \Delta x$ до $x/\Delta x$.

Следует обратить внимание на то, что в этом случае доверительный интервал задаётся относительной ошибкой.

Пример. При определении зависимости разности электрического потенциала на конденсаторе (u, В) от концентрации микропримесей кобальта (c, мас. %) на приборе МФС-4 были получены следующие экспериментальные данные:

$u_{фон}$ = 40 В, 35 В, 42 В;

при $c_1 = 10^{-3}$ мас. % u_1 = 265 В, 332 В;

при $c_2 = 10^{-2}$ мас. % u_2 = 675 В;

при $c_3 = 10^{-1}$ мас. % u_3 = 1771 В, 2139 В, 1811 В.

Среднее значение фона составляет 39 В. Обозначив $y=u=u_i - u_{фон}$, а концентрацию кобальта в пробе x, проводим логарифмическое преобразование данных ($Y=\lg y$, $X=\lg x$) и вносим полученные результаты в таблицу.

№	x_i	y_i	X_i	Y_i	X_iY_i	X_i^2	Y_i^2
1	10^{-3}	226	-3	2,354108	$-7{,}062324$	9	5,541827
2	10^{-3}	293	-3	2,466868	$-7{,}400603$	9	6,085436
3	10^{-2}	636	-2	2,803457	$-5{,}606914$	4	7,859372
4	10^{-1}	1732	-1	3,238548	$-3{,}238548$	1	10,488192
5	10^{-1}	2100	-1	3,322219	$-3{,}322219$	1	11,037141
6	10^{-1}	1772	-1	3,248464	$-3{,}248464$	1	10,552517
n=6			$\Sigma X_i = -11$ $(\Sigma X_i)^2=121$	$\Sigma Y_I=17{,}433664$ $(\Sigma Y_I)^2=303{,}932643$	$\Sigma X_iY_i=$ $-29{,}879073$	$\Sigma X_i^2=$ $=25$	$\Sigma Y_i^2=$ $=51{,}564484$

Перед началом обработки данных рассчитываем коэффициент корреляции:

$$r = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{\sqrt{[\,n\sum X_i^2 - \left(\sum X_i\right)^2\,] \cdot [\,n\sum Y_i^2 - \left(\sum Y_i\right)^2\,]}}$$

$$r = \frac{6 \cdot (-29{,}879073) - (-11) \cdot 17{,}433664}{\sqrt{\left(6 \cdot 25 - 121\right) \cdot \left(6 \cdot 51{,}564484 - 303{,}93264\right)}} = 0{,}993572$$

В таблице находим *r*(*P*=0,95; *f*=4) = 0,81. Поскольку |*r*| > 0,81, то между значениями *X* и *Y* существует линейная зависимость.

Определяем коэффициенты регрессии:

$$a = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^{\,2} - \left(\sum_{i=1}^{n} X_i\right)^2}$$

$$a = \frac{6 \cdot (-29{,}879073) - (-11) \cdot 17{,}433664}{6 \cdot 25 - 121} = 0{,}4308919$$

$$B = \frac{\sum_{i=1}^{n} Y_i - a\sum_{i=1}^{n} X_i}{n}$$

$$B = \frac{17{,}433664 - 0{,}4308919 \cdot (-11)}{6} = 3{,}69557915$$

Далее рассчитываем дисперсии:

$$s_0^2 = \frac{\sum_{i=1}^{n} Y_i^2 - B\sum_{i=1}^{n} Y_i - a\sum_{i=1}^{n} X_i Y_i}{n-2}$$

$$s_0{}^2 = \frac{51{,}564484 - 3{,}69557915 \cdot 17{,}433664 - 0{,}4308919 \cdot (-29{,}879073)}{6-2} =$$

$$= 2{,}912296 \cdot 10^{-3}$$

$$s_a^2 = \frac{n \cdot s_0^2}{n\sum_{i=1}^{n} \mathrm{X}_i{}^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

$$s_a^2 = \frac{6 \cdot 2{,}912296 \cdot 10^{-3}}{6 \cdot 25 - 121} = 6{,}025439 \cdot 10^{-4}$$

$$s_B^2 = \frac{s_a^2 \cdot \sum_{i=1}^{n} X_i^2}{n}$$

$$s_B^2 = \frac{6{,}025439 \cdot 10^{-4} \cdot 25}{6} = 2{,}510600 \cdot 10^{-3}$$

При P = 0,95 и f = 6 − 2 = 4 табличное значение $t(P,f)$ = 2,78, откуда доверительные интервалы для a и B равны:

$$\Delta a = 2{,}78 \cdot \sqrt{6{,}025439 \cdot 10^{-4}} = 0{,}068240$$

$$\Delta B = 2{,}78 \cdot \sqrt{2{,}510600 \cdot 10^{-3}} = 0{,}139294$$

Окончательный вид уравнения регрессии с учётом точности определения коэффициентов a и B имеет вид:

$$\lg u = (0{,}43_1 \pm 0{,}06_8) \cdot \lg c + (3{,}7_0 \pm 0{,}1_4)$$

Поскольку коэффициент B существенно больше нуля, то проверку возможности преобразования полученного уравнения к виду $Y=aX$ можно не проводить.

При анализе контрольной пробы (К) на содержание кобальта было получено значение разности потенциалов $u_{К+ф}$ = 489, 462, 474 В. Используя уравнение, обратное градуировочной функции, определяем десятичный логарифм концентрации компонента:

$$lg\, c_{К} = \frac{lg\, u - 3{,}70}{0{,}431} = \frac{lg\, 436 - 3{,}70}{0{,}431} = -2{,}4512686$$

Далее находим стандартное отклонение и ошибку определения концентрации в пробе К:

$$s_{X_K} = \frac{s_0}{a}\sqrt{\frac{1}{n} + \frac{1}{m} + \left(\frac{s_a}{a}\right)^2 \cdot \left(\frac{\overline{Y}_K - \overline{Y}}{s_0}\right)^2}$$

$$s_{lg\, c_{К}} = \frac{\sqrt{2{,}912296 \cdot 10^{-3}}}{0{,}4308919} \cdot \sqrt{\frac{1}{6} + \frac{1}{3} + \left(\frac{\sqrt{6{,}025439 \cdot 10^{-4}}}{0{,}4308919}\right)^2 \cdot \left(\frac{lg\, 436 - \frac{17{,}433664}{6}}{\sqrt{2{,}912296 \cdot 10^{-3}}}\right)^2} =$$

$$= 0{,}09529925$$

$$\Delta X_{К} = s_{X_{К}} \cdot t(P, f)$$

$$\Delta\, lg\, c_{К} = s_{lg\, c_{К}} \cdot t(\,P = 0{,}95;\, f = 6 - 2 = 4\,) = 0{,}09529925 \cdot 2{,}78 = 0{,}2649319$$

Отсюда

$$\lg c_К = -2{,}4_5 \pm 0{,}2_6 .$$

Потенцированием получаем, что

$$c_К = 10^{-2,45} = 3{,}53778 \cdot 10^{-3} \text{ мас. \%},$$

Доверительный интервал:

$$\Delta c_К = 10^{0,26} = 1{,}81970$$

Следовательно, искомое значение концентрации кобальта в пробе лежит в диапазоне

$$3{,}5 \cdot 10^{-3} \cdot 1{,}8 < c_К < \frac{3{,}5 \cdot 10^{-3}}{1{,}8}$$

Таким образом, окончательно получаем:

$$1{,}95 \cdot 10^{-3} < c_К < 6{,}46 \cdot 10^{-3} \text{ (мас. \%)}.$$