- **Математическое ожидание и дисперсия распределения параметров a_0 и a_1**

$$M\left(a_0\right)=\widehat{a}_0$$

$$D\left(a_0\right)=M\left[a_0-M\left(a_0\right)\right]^2=M\left(a_0-\widehat{a}_0\right)^2=$$

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

Оценка дисперсии a_0

$$\widehat{\sigma}^2_{a_0} = \widehat{\sigma}^2_u \frac{\sum\limits_{i=1}^{n} x_i^2}{n\sum\limits_{i=1}^{n}\left(x_i - \overline{x}\right)^2}$$

$$M\left(a_1\right) = \widehat{a}_1$$

$$D\left(a_1\right) = M\left[a_1 - M\left(a_1\right)\right]^2 = M\left(a_1 - \widehat{a}_1\right)^2 =$$

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

Оценка дисперсии a_1

$$\hat{\sigma}^2_{a_1} = \hat{\sigma}^2_u \frac{1}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}$$

Проверка значимости параметров a_0 и a_1 по критерию Стьюдента и построение доверительных интервалов

Для проверки значимости параметров модели рассчитывают t-статистику Стьюдента

$$t_{p_i} = \frac{p_i - p_i^*}{\sqrt{\hat{\sigma}^2_{p_i}}}$$

с $\upsilon = n - k$ степенями свободы

p_i – оценка рассматриваемого i-го параметра

p_i^* – гипотетическое значение, которое должен принять параметр $\left(H_0 : p_i = p_i^*\right)$

i

$\widehat{\sigma}_{p_i}^2$ – оценка дисперсии параметра p_i

n – размер выборки

k – число оцениваемых параметров модели (для парной регрессии $k = 2$)

Оценка параметра $\widehat{a}_1$

Выдвигается гипотеза $H_0 : \widehat{a}_1 = 0$

Статистика Стьюдента

$$t_{\widehat{a}_1} = \frac{\widehat{a}_1}{\widehat{\sigma}_{a_1}}$$

Находим табличное значение

$$t_{(n-k,\alpha)}$$

Если $\left|t_{\hat{a}_1}\right| > t_{(n-k,\alpha)}$, то гипотеза $H_0 : \hat{a}_1 = 0$ отвергается,

если $\left|t_{\hat{a}_1}\right| \leq t_{(n-k,\alpha)}$ – то принимается

Оценка значимости параметра $\hat{a}_0$ осуществляется аналогично

$$t_{\hat{a}_0} = \frac{\hat{a}_0}{\hat{\sigma}_{a_0}}$$

Оценим значимость параметров модели $\widehat{a}_1$ и $\widehat{a}_0$

Рассчитаем оценку дисперсии ошибок $\widehat{\sigma}_u^2$

$$\widehat{\sigma}_u^2 = \frac{\sum_{i=1}^{n} u_i^2}{n-2} = \frac{10,71}{8-2} = 1,78$$

Найдем оценку дисперсии a_1 и среднеквадратическое отклонение

$$\hat{\sigma}^2_{a_1} = \hat{\sigma}^2_u \cdot \frac{1}{\sum\limits_{i=1}^{n}\left(x_i - \bar{x}\right)^2} = 1{,}78 \cdot \frac{1}{828} = 0{,}002155$$

$$\hat{\sigma}_{a_1} = \sqrt{\hat{\sigma}^2_{a_1}} = 0{,}046$$

Рассчитаем статистику Стьюдента

$$t_{\widehat{a}_1} = \frac{\widehat{a}_1}{\widehat{\sigma}_{a_1}} = \frac{0,576}{0,046} = 12,41$$

Для $\alpha = 0,05$ и $v = n - 2 = 8 - 2 = 6$

$$t_{(6,0.05)} = 1,943$$

Так как

$$t_{\widehat{a}_1} = 12{,}4 > t_{(6,\,0.05)} = 1{,}943\ ,$$

то значение коэффициента является значимым

Найдем оценку дисперсии и среднеквадратическое отклонение a_0

$$\hat{\sigma}^2_{a_0} = \hat{\sigma}^2_u \cdot \frac{\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} = 1{,}78 \cdot \frac{6030}{8 \cdot 828} = 1{,}62$$

$$\hat{\sigma}_{a_0} = \sqrt{\hat{\sigma}^2_{a_0}} = 1{,}27$$

Рассчитаем статистику Стьюдента

$$t_{\widehat{a}_0} = \frac{\widehat{a}_0}{\widehat{\sigma}_{a_0}} = \frac{-1,44}{1,27} = -1,13$$

Сравним рассчитанное значение статистики по модулю с табличным

$$\left|t_{\hat{a}_0}\right| = \left|-1{,}13\right| < t_{(6,\,0.05)} = 1{,}943$$

Вывод: для выбранного уровня доверия значение коэффициента не является значимым

Расчет доверительных интервалов для параметров

Этапы расчета

1. Рассчитываем статистику Стьюдента

$$t_{p_i} = \frac{p_i - p_i^*}{\sqrt{\hat{\sigma}^2_{p_i}}}$$

2. Выбираем уровень значимости α

По таблице критерия Стьюдента находим значения $\pm t_{\alpha/2}$ с $n-k$ степенями свободы (для парной зависимости).

Получаем

$$P\left(-t_{\alpha/2} \le t_{p_i} \le t_{\alpha/2}\right) = 1-\alpha$$

или

$$P\left(-t_{\alpha/2} \le \frac{p_i - p_i^*}{\sqrt{\hat{\sigma}_{p_i}^2}} \le t_{\alpha/2}\right) = 1 - \alpha$$

Отсюда имеем

$$p_i^* = p_i \pm t_{\alpha/2} \cdot \widehat{\sigma}_{p_i}$$

Для параметра $\widehat{a}_1$ получаем

$$a_1^* = \widehat{a}_1 \pm t_{\alpha/2} \cdot \widehat{\sigma}_{a_1}$$

Для параметра $\widehat{a}_0$

$$a_0^* = \widehat{a}_0 \pm t_{\alpha/2} \cdot \widehat{\sigma}_{a_0}$$

Расчет доверительных интервалов для параметров a_1 и a_0.

Пусть $\alpha = 0{,}05$

Для $n - 2 = 8 - 2 = 6$ степеней свободы находим табличное значение

$$t_{\alpha/2} = t_{0{,}025} = 2{,}447$$

С учетом этого

$$a_1^* = \widehat{a}_1 \pm t_{0.025} \cdot \widehat{\sigma}_{a_1} = 0{,}576 \pm 2{,}447 \cdot 0{,}0464 =$$

$$= 0{,}576 \pm 0{,}1136$$

$$\downarrow$$

$$0{,}4625 \le \widehat{a}_1 \le 0{,}6897$$

Для $\hat{a}_0$

$$a_0^* = \hat{a}_0 \pm t_{0.025} \cdot \hat{\sigma}_{a_0} =$$

$$= -1,44 \pm 2,447 \cdot 1,2745 = -1,44 \pm 3,1188$$

$$\Downarrow$$

$$-4,55897 \leq \hat{a}_0 \leq 1,6785$$

•Прогнозирование по модели простой линейной регрессии

Виды прогнозов

- интервальный
- точечный

$$\widehat{y}_{n+1} = \widehat{a}_0 + \widehat{a}_1 x_{n+1}$$

Ошибка прогноза u_{n+1}

$$u_{n+1} = y_{n+1} - \widehat{y}_{n+1}$$

Дисперсия

$$\sigma^2_{u_{n+1}} = \sigma^2_u \left[1 + \frac{1}{n} + \frac{\left(x_{n+1} - \bar{x}\right)^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2} \right]$$

Оценка дисперсии

$$\hat{\sigma}^2_{u_{n+1}} = \hat{\sigma}^2_u \left[1 + \frac{1}{n} + \frac{\left(x_{n+1} - \bar{x}\right)^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2} \right]$$

Статистика Стьюдента для переменной $\widehat{y}_{n+1}$

$$t_{\widehat{y}_{n+1}} = \frac{\widehat{y}_{n+1} - y^{*}_{n+1}}{\widehat{\sigma}_u \sqrt{1 + \dfrac{1}{n} + \dfrac{\left(x_{n+1} - \bar{x}\right)^2}{\sum\limits_{i=1}^{n} \left(x_i - \bar{x}\right)^2}}}$$

Доверительный интервал для прогнозного значения переменной y

$$y^*_{n+1} = \widehat{y}_{n+1} \pm t_{\alpha/2} \cdot \widehat{\sigma}_u \sqrt{1 + \frac{1}{n} + \frac{\left(x_{n+1} - \overline{x}\right)^2}{\sum\limits_{i=1}^{n} \left(x_i - \overline{x}\right)^2}}$$

Построение доверительных интервалов для математического ожидания (точечный прогноз)

$$M\left(y_{n+1}\right) = a_0 + a_1 x_{n+1}$$

Ошибка

$$u_{n+1} = M\left(y_{n+1}\right) - \widehat{y}_{n+1}$$

Дисперсия ошибок

$$\sigma^2_{u_{n+1}} = \sigma^2_u \left[\frac{1}{n} + \frac{\left(x_{n+1} - \bar{x}\right)^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2} \right]$$

Оценка дисперсии ошибок

$$\hat{\sigma}^2_{u_{n+1}} = \hat{\sigma}^2_u \left[\frac{1}{n} + \frac{\left(x_{n+1} - \bar{x}\right)^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2} \right]$$

Выражение для расчета доверительного интервала для точечного прогноза

$$M\left(y^{*}_{n+1}\right) = \widehat{y}_{n+1} \pm t_{\alpha/2} \cdot \widehat{\sigma}_{u} \sqrt{\frac{1}{n} + \frac{\left(x_{n+1} - \bar{x}\right)^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}}$$

Рассчитаем прогноз величины $\widehat{y}_{n+1}$

и доверительные интервалы прогноза для

$$n+1=8+1=9$$

Определим значение $x_9 = 20$

Найдем

$$\widehat{y}_9 = -1{,}44 + 0{,}576 \cdot x_9 = -1{,}44 + 0{,}576 \cdot 20 = 10{,}081$$

Доверительные интервалы для уровня значимости $\alpha/2 = 0{,}025$

$$y_9^* = \widehat{y}_9 \pm t_{\alpha/2} \cdot \widehat{\sigma}_u \cdot \sqrt{1 + \frac{1}{n} + \frac{\left(x_9 - \overline{x}\right)^2}{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}} =$$

$$= 10{,}082 \pm 2{,}447 \cdot 1{,}336 \cdot \sqrt{1 + \frac{1}{8} + \frac{\left(20 - 25{,}5\right)^2}{828}} = 10{,}082 \pm 3{,}523$$

$$6,559 \leq y_{n+1} \leq 13,60$$

Доверительные интервалы для математического ожидания y_{n+1}

для $\alpha/2 = 0{,}025$

$$M\left(y_9^*\right) = \hat{y}_9 \pm t_{\alpha/2} \cdot \hat{\sigma}_u \cdot \sqrt{\frac{1}{n} + \frac{\left(x_9 - \bar{x}\right)^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}} =$$

$$= 10{,}082 \pm 2{,}447 \cdot 1{,}336 \cdot \sqrt{\frac{1}{8} + \frac{\left(20 - 25{,}5\right)^2}{828}} = 10{,}082 \pm 1{,}3138$$

$$8{,}7678 \le M\left(y_9\right) \le 11{,}3953$$