

- Математическое ожидание и дисперсия распределения параметров $\underline{a_0}$ и $\underline{a_1}$

$$M(a_0) = \hat{a}_0$$

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

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

Оценка дисперсии a_0

$$\hat{\sigma}_{a_0}^2 = \hat{\sigma}_u^2 \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$M(a_1) = \hat{a}_1$$

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

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

Оценка дисперсии a_1

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

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

Для проверки значимости параметров
модели рассчитывают t -статистику
Стьюдента

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

с $\nu = n - k$ степенями свободы

p_i – оценка рассматриваемого i -го параметра

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

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

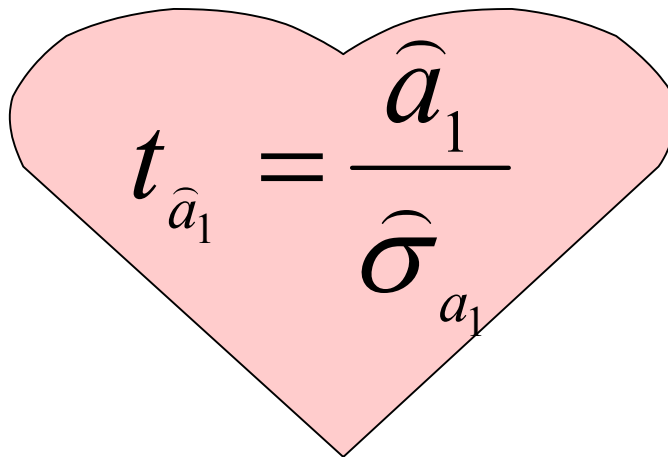
n – размер выборки

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

Оценка параметра $\hat{a}_1$

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

Статистика Стьюдента


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

Находим табличное значение

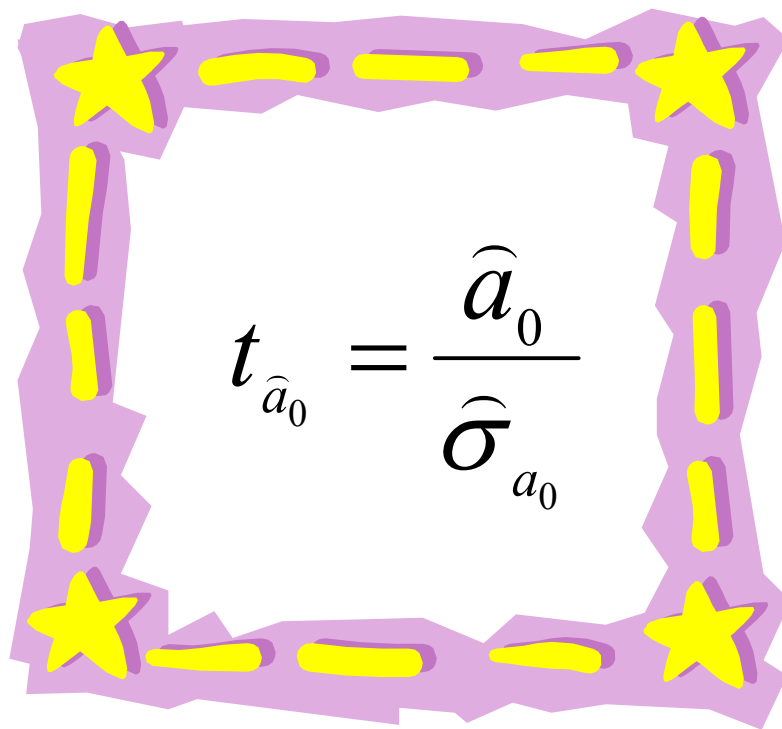
$$t_{(n-k, \alpha)}$$

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 $\hat{a}_1$ и $\hat{a}_0$

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

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

Найдем оценку дисперсии a_1 и
среднеквадратическое отклонение

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

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

Рассчитаем статистику Стьюдента

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

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

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

Так как

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

то значение коэффициента является
значимым

Найдем оценку дисперсии и
среднеквадратическое отклонение a_0

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

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

Рассчитаем статистику Стьюдента

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

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

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

2. Выбираем уровень значимости α

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

Получаем

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

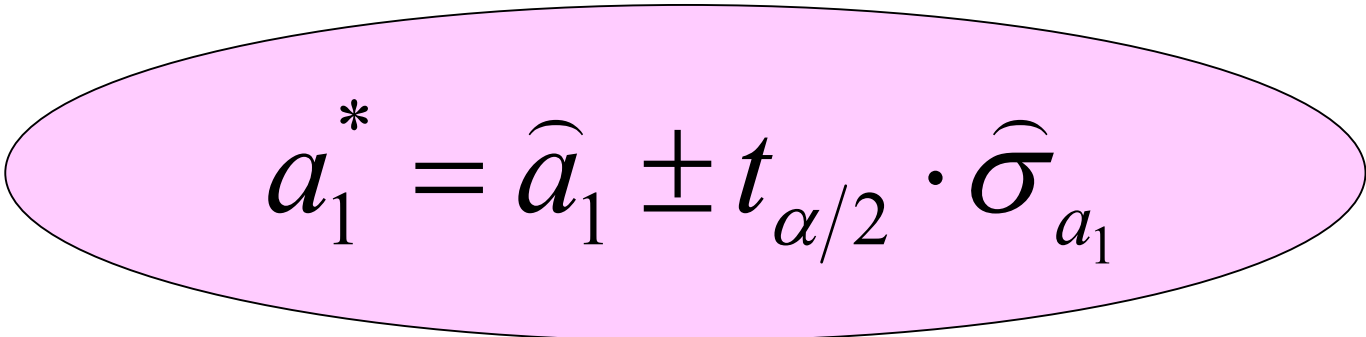
или

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

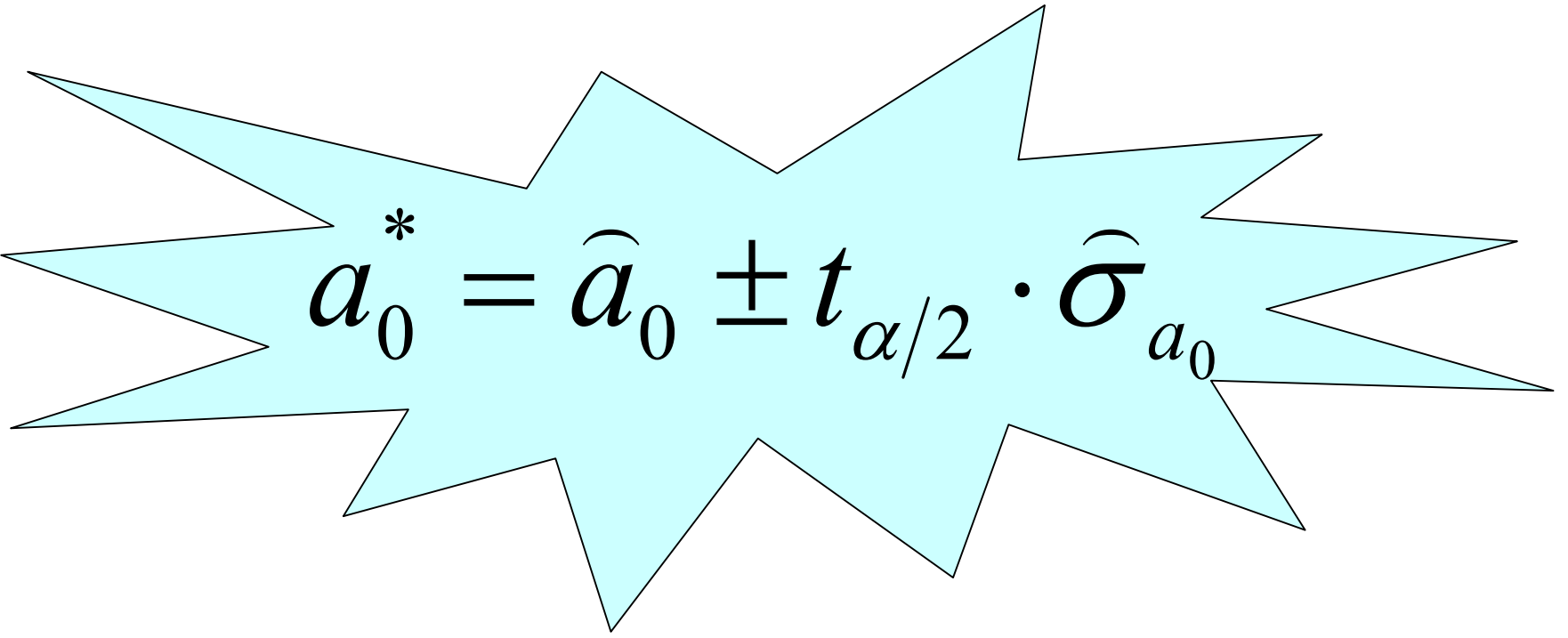
Отсюда имеем

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

Для параметра $\hat{a}_1$ получаем


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

Для параметра $\hat{a}_0$


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

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

$$a_1^* = \hat{a}_1 \pm t_{0.025} \cdot \hat{\sigma}_{a_1} = 0,576 \pm 2,447 \cdot 0,0464 = \\ = 0,576 \pm 0,1136$$



$$0,4625 \leq \hat{a}_1 \leq 0,6897$$

Для $\hat{a}_0$

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

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



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

- Прогнозирование по модели простой линейной регрессии

Виды прогнозов

- интервальный
- точечный

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

Ошибка прогноза u_{n+1}

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

Дисперсия

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

Оценка дисперсии

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

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

$$t_{\hat{y}_{n+1}} = \frac{\hat{y}_{n+1} - y_{n+1}^*}{\hat{\sigma}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}}$$

Доверительный интервал для прогнозного значения переменной y

$$y_{n+1}^* = \hat{y}_{n+1} \pm t_{\alpha/2} \cdot \hat{\sigma}_u \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Построение доверительных интервалов для
математического ожидания (точечный прогноз)

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

Ошибка

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

Дисперсия ошибок

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

Оценка дисперсии ошибок

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

Выражение для расчета доверительного интервала для точечного прогноза

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

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

и доверительные интервалы прогноза для

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

Определим значение $x_9 = 20$

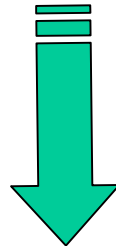
Найдем

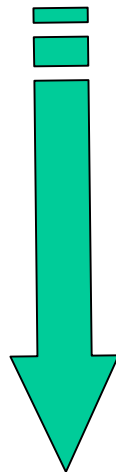
$$\hat{y}_9 = -1,44 + 0,576 \cdot x_9 = -1,44 + 0,576 \cdot 20 = 10,081$$

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

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

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





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

Доверительные интервалы для
математического ожидания y_{n+1}

для $\alpha/2 = 0,025$

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

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



$$8,7678 \leq M(y_9) \leq 11,3953$$