Прикладная статистика в машинном обучении

# Квиз #3A

16 ноября 2020 г.

### Критические значения.

- $t_{500-6;0.025} = -1.964778$
- $t_{500-6,0.05} = -1.647944$
- $F_{2,1;0.95} = 199.5$
- $Z_{0.025} = -1.959964$
- $Z_{0.05} = -1.644854$
- $F_{1,1;0.95} = 161.4476.$

### Условие для задач 1-4.

Рассмотрим модель $y_i = \beta_0 + \beta_1\sqrt{X_{1i}} + \beta_2 X_{1i}^2 + \beta_3 X_{2i} + \beta_4 e^{X_{3i}} + \beta_5 X_{4i} + u_i$, которая оценивается по 700 наблюдениям при помощи МНК. Оценённая модель имеет следующий вид:

$$\hat{y}_i = \underset{(0.05)}{20} + \underset{(0.02)}{4}\sqrt{X_{1i}} + \underset{(9)}{184}X_{1i}^2 - \underset{(0.9)}{1.1}X_{2i} - \underset{(0.01)}{0.3}e^{X_{3i}} + \underset{(0.1)}{15}X_{4i}.$$

В скобках указаны стандартные ошибки оценок коэффициентов. Будем считать, что все предпосылки теоремы Гаусса-Маркова выполнены, и $u \sim \mathcal{N}(0, \sigma^2 I)$. Также известно, что

$$\hat{\mathrm{Var}}(\hat{\beta}) = \begin{pmatrix} 0.0025 & -0.04 & 0.001 & -0.001 & 0.47 & -0.1 \\ -0.04 & 0.0004 & 0.09 & 1 & 0.9 & 0.4 \\ 0.001 & 0.09 & 81 & 0.017 & -0.48 & 9 \\ -0.001 & 1 & 0.017 & 0.81 & -0.002 & 4.78 \\ 0.47 & 0.9 & -0.48 & -0.002 & 10^{-4} & 1.17 \\ -0.1 & 0.4 & 9 & -4.78 & 1.17 & 0.01 \end{pmatrix}$$

## Задача 1.

a) Проверьте коэффициенты $\beta_1$, $\beta_2$ и $\beta_5$ на значимость на 5% уровне.

b) Проверьте коэффициенты $\beta_1$, $\beta_3$ и $\beta_4$ на значимость на 5% уровне.

c) Проверьте коэффициенты $\beta_0$, $\beta_1$ и $\beta_3$ на значимость на 5% уровне.

d) Проверьте коэффициенты $\beta_0$, $\beta_3$ и $\beta_5$ на значимость на 5% уровне.

e) Проверьте коэффициенты $\beta_3$, $\beta_4$ и $\beta_5$ на значимость на 5% уровне.

f) Проверьте коэффициенты $\beta_1$, $\beta_4$ и $\beta_5$ на значимость на 5% уровне.

g) Проверьте коэффициенты $\beta_0$, $\beta_3$ и $\beta_4$ на значимость на 5% уровне.

## Задача 2.

a) Проверьте следующие гипотезы на уровне значимости 5%.

$$\begin{cases} H_0: \beta_4 = -0.2, \\ H_1: \beta_4 \neq -0.2 \end{cases} \text{и} \begin{cases} H_0: \beta_2 = 190, \\ H_1: \beta_2 < 190. \end{cases}$$

b) Проверьте следующие гипотезы на уровне значимости 5%.

$$\begin{cases} H_0: \beta_3 = -1, \\ H_1: \beta_3 \neq -1 \end{cases} \text{и} \begin{cases} H_0: \beta_5 = 13, \\ H_1: \beta_5 < 13. \end{cases}$$

c) Проверьте следующие гипотезы на уровне значимости 5%.

$$\begin{cases} H_0: \beta_1 = 5 \\ H_1: \beta_1 \neq 5 \end{cases} \text{и} \begin{cases} H_0: \beta_2 = 180, \\ H_1: \beta_2 > 180. \end{cases}$$

d) Проверьте следующие гипотезы на уровне значимости 5%.

$$\begin{cases} H_0: \beta_1 = 3, \\ H_1: \beta_1 \neq 3 \end{cases} \text{и} \begin{cases} H_0: \beta_4 = -0.2, \\ H_1: \beta_4 < -0.2. \end{cases}$$

## Задача 3.

a) Постройте 95%-ый доверительный интервал для $\beta_4$.

b) Постройте 95%-ый доверительный интервал для $\beta_1$.

c) Постройте 95%-ый доверительный интервал для $\beta_5$.

d) Постройте 95%-ый доверительный интервал для $\beta_3$.

e) Постройте 95%-ый доверительный интервал для $\beta_2$.

## Задача 4.

a) Постройте 95%-ый доверительный интервал для $\mathbb{E}(y_{701}|X_1 = X_2 = 4, X_3 = 0, X_4 = 1)$.

b) Постройте 95%-ый доверительный интервал для $\mathbb{E}(y_{701}|X_1 = X_2 = 0, X_3 = X_4 = 1)$.

c) Постройте 95%-ый доверительный интервал для $\mathbb{E}(y_{701}|X_1 = X_2 = 4, X_3 = X_4 = 0)$.

d) Постройте 95%-ый доверительный интервал для $\mathbb{E}(y_{701}|X_1 = X_3 = 0, X_2 = X_4 = 1)$.

## Задача 5.

Рассмотрим модель $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$, оцениваемую при помощи МНК. Предположим, что все предпосылки теоремы Гаусса-Маркова выполнены, и $u \sim \mathcal{N}(0, \sigma^2 I)$. Модель оценили на следующих данных:

$$X = \begin{bmatrix} 1 & -9 & 2 \\ 1 & -2 & 0 \\ 1 & -5 & 3 \\ 1 & -1 & 2 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}.$$

Оказалось, что $\hat{\beta} = \begin{pmatrix} 1.9 & 0.01 & -0.4 \end{pmatrix}'$.

a) Проверьте регрессию на значимость в целом на уровне значимости 5%.

b) Проверьте гипотезу

$$\begin{cases} H_0: \beta_1 = \beta_2, \\ H_1: \beta_1 \neq \beta_2 \end{cases}$$

на уровне значимости 5%, если

$$(X'X)^{-1}_R = \begin{pmatrix} 0.3 & 0.03 \\ 0.03 & 0.02 \end{pmatrix}, \hat{\beta}_R = \begin{pmatrix} 1.4 \\ 0.07 \end{pmatrix}$$

## Задача 6.

Рассмотрим модель $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$, оцениваемую при помощи МНК. Предположим, что все предпосылки теоремы Гаусса-Маркова выполнены, и $u \sim \mathcal{N}(0, \sigma^2 I)$. Модель оценили на следующих данных:

$$X = \begin{bmatrix} 1 & 4 & 0 \\ 1 & -2 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}, y = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}.$$

Оказалось, что $\hat{\beta} = \begin{pmatrix} 2.3 & 0.2 & -1.3 \end{pmatrix}'$.

a) Проверьте регрессию на значимость в целом на уровне значимости 5%.

b) Проверьте гипотезу

$$\begin{cases} H_0: \beta_1 = \beta_2, \\ H_1: \beta_1 \neq \beta_2 \end{cases}$$

на уровне значимости 5%, если

$$(X'X)^{-1}_R = \begin{pmatrix} 0.28 & -0.04 \\ -0.04 & 0.05 \end{pmatrix}, \hat{\beta}_R = \begin{pmatrix} 1.6 \\ 0.2 \end{pmatrix}$$

## Задача 7.

Рассмотрим линейную модель $y = X\beta + u$, оцениваемую при помощи МНК. Пусть $\mathbb{E}(u) = 0$, $\mathrm{Var}(u) = \sigma^2 I$, число наблюдений равно $n$, число регрессоров, включая константный, равно $k$. Найдите следующие величины и укажите их размеры

a) $\mathbb{E}(\hat{\beta})$, $\mathrm{Var}(\hat{u})$, $\mathrm{Cov}(\hat{\beta}, \hat{u})$, $\mathrm{Cov}(u, \hat{\beta})$.

b) $\mathbb{E}(\hat{y})$, $\mathrm{Var}(\hat{y})$, $\mathrm{Cov}(\hat{\beta}, \hat{u})$, $\mathrm{Cov}(\hat{u}, u)$.

c) $\mathbb{E}(y)$, $\mathrm{Var}(\hat{\beta})$, $\mathrm{Cov}(u, \hat{u})$, $\mathrm{Cov}(\hat{u}, \hat{y})$.

## Задача 8.

Рассмотрим модель парной регрессии

$$y_i = \beta_0 + \beta_1 x_i + u_i,$$

оцениваемую при помощи МНК. Пусть известно, что

| $y_i$ | $x_i$ |
|---|---|
| 5 | 1 |
| 1 | 0 |
| 1 | 4 |
| 2 | 1 |

$$(X'X)^{-1} = \begin{bmatrix} 0.5 & -0.17 \\ -0.17 & 0.1 \end{bmatrix}$$

a) Постройте 95%-ый доверительный интервал для $\hat{\beta}_0$, используя стандартные ошибки $HC_0$.

b) Постройте 95%-ый доверительный интервал для $\hat{\beta}_1$, используя стандартные ошибки $HC_0$.

c) Постройте 95%-ый доверительный интервал для $\hat{\beta}_0$, используя стандартные ошибки $HC_3$.

d) Постройте 95%-ый доверительный интервал для $\hat{\beta}_1$, используя стандартные ошибки $HC_3$.