Семинар 3.

2. Рассмотрим оценку вида $\tilde{\beta} = ((X'X)^{-1} + \gamma I)X'y$ для вектора коэффициентов регрессионного уравнения $y = X\beta + \varepsilon$, удовлетворяющего условиям классической регрессионной модели. Найдите $E(\tilde{\beta})$ и $Var(\tilde{\beta})$.

1.

$$\mathbb{E}(\tilde{\beta}) = ((X'X)^{-1} + \gamma I)X'\mathbb{E}(y) = \\ ((X'X)^{-1} + \gamma I)X'X\beta = \beta + \gamma X'X\beta$$

2.

$$\begin{aligned}\text{Var}(\tilde{\beta}) &= \text{Var}(((X'X)^{-1} + \gamma I)X'y) = \\ &\text{Var}(((X'X)^{-1} + \gamma I)X'\varepsilon) = \\ &(((X'X)^{-1} + \gamma I)X')\,\text{Var}(\varepsilon)(((X'X)^{-1} + \gamma I)X')' = \\ &(((X'X)^{-1} + \gamma I)X')\sigma_\varepsilon^2 I(((X'X)^{-1} + \gamma I)X')' = \\ &\sigma_\varepsilon^2((X'X)^{-1} + \gamma I)X'X((X'X)^{-1} + \gamma I) = \\ &\sigma_\varepsilon^2((X'X)^{-1} + \gamma I)(I + \gamma X'X) = \\ &\sigma_\varepsilon^2((X'X)^{-1} + 2\gamma I + \gamma^2 X'X)\end{aligned}$$

3. Пусть регрессионная модель $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$, $i = 1, \ldots, n$, задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 \end{pmatrix}'$. Известно, что $E(\varepsilon) = 0$ и $Var(\varepsilon) = \sigma^2 \cdot I$. Известно также, что:

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \bar{y} = 3$$

Для удобства расчётов ниже приведены матрицы:

$$X'X = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ и } (X'X)^{-1} = \begin{pmatrix} 1/3 & -1/3 & 0 \\ -1/3 & 4/3 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

a) $n = 5$

б) $k = 3$

c) $\hat{\beta} = (X'X)^{-1} X'y = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$

d) $TSS = \sum_{i=1}^{5} (y_i - \bar{y})^2 = \sum_{i=1}^{5} (y_i - 3)^2 = (1-3)^2 + (2-3)^2 + 0^2 + (4-3)^2 + (5-3)^2 = 10$

$RSS = \sum_{i=1}^{5} (y_i - \hat{y}_i)^2 = (-1)^2 + 1^2 = 2$

$\hat{y} = X\hat{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 4 \\ 5 \end{pmatrix}, \quad y - \hat{y} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

$ESS = TSS - RSS = 10 - 2 = 8$

e) $e_4 = y_4 - \hat{y}_4 = 4 - 4 = 0$

f) $R^2 = \frac{ESS}{TSS} = \frac{8}{10} = 0.8$

g) $\hat{\sigma}_\varepsilon^2 = \frac{RSS}{n-k} = \frac{2}{5-3} = 1$

h) $\hat{Var}(\hat{\beta}) = \hat{\sigma}_\varepsilon^2 (X'X)^{-1} =$

$$= 1 \cdot \begin{pmatrix} 1/3 & -1/3 & 0 \\ -1/3 & 4/3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

i) $\hat{Var}(\hat{\beta}_1) = \hat{\sigma}_\varepsilon^2 \cdot (X'X)^{-1}_{11} = 1 \cdot 1/3 = \frac{1}{3}$

j) $\hat{cov}(\hat{\beta}_1, \hat{\beta}_2) = \hat{\sigma}_\varepsilon^2 \cdot (X'X)^{-1}_{12} = 1 \cdot \left(-\frac{1}{3}\right) = -\frac{1}{3}$

k) $\hat{Var}(\hat{\beta}_1 + \hat{\beta}_2) = \hat{Var}(\hat{\beta}_1) + \hat{Var}(\hat{\beta}_2) +$

$+ 2\hat{cov}(\hat{\beta}_1, \hat{\beta}_2) = 1 \cdot \frac{1}{3} + 1 \cdot \frac{4}{3} + 2 \cdot \left(-\frac{1}{3}\right) =$

$= \frac{3}{3} = 1$

l) $$\widehat{corr}(\hat{\beta}_1, \hat{\beta}_2) = \frac{\widehat{cov}(\hat{\beta}_1, \hat{\beta}_2)}{\sqrt{\widehat{Var}(\hat{\beta}_1)} \cdot \sqrt{Var(\hat{\beta}_2)}} =$$

$$= \frac{\overset{=1}{\hat{\sigma}_\varepsilon^2} \cdot (X'X)^{-1}_{12}}{\sqrt{\underset{1}{\underset{\shortparallel}{\hat{\sigma}_\varepsilon^2}} \cdot (X'X)^{-1}_{11}} \cdot \sqrt{\underset{1}{\underset{\shortparallel}{\hat{\sigma}_\varepsilon^2}} \cdot (X'X)^{-1}_{22}}} = \frac{-1/3}{\sqrt{\frac{1}{3}} \cdot \sqrt{\frac{4}{3}}} =$$

$$= \frac{-1/3}{\sqrt{\frac{4}{9}}} = -\frac{1}{2}$$

m) $$\hat{\sigma}(\hat{\beta}_1) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}} = \sqrt{\widehat{Var}(\hat{\beta}_1)} =$$

$$= \sqrt{\hat{\sigma}_\varepsilon^2 \cdot (X'X)^{-1}_{11}} = \sqrt{1 \cdot \frac{1}{3}} = \frac{1}{\sqrt{3}}$$