Семинар 4.
Множественная регрессия.

1. Рассмотрим классическую линейную модель $y = X\beta + \varepsilon$ с предпосылками Гаусса — Маркова: $\mathbb{E}(\varepsilon) = 0$ и $\mathbb{V}\mathrm{ar}(\varepsilon) = \sigma^2 I$. Для всех случайных векторов (y, $\hat{\beta}$, $\hat{y}$, ε, $\hat{\varepsilon}$, $\bar{y}$) найдите все возможные условные математические ожидания и ковариационные матрицы. $\mathbb{E}(\cdot)$, $\mathbb{V}\mathrm{ar}(\cdot)$, $\mathbb{C}\mathrm{ov}(\cdot, \cdot)$.

2. Рассмотрим модель $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$, где

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}.$$

   Случайные ошибки ε_i независимы и нормально распределены с $\mathbb{E}(\varepsilon) = 0$ и $\mathbb{V}\mathrm{ar}(\varepsilon) = \sigma^2 I$.

   Для удобства расчётов даны матрицы: $X^T X$, $(X^T X)^{-1}$ и $X^T y$:

$$X^T X = \begin{pmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad (X^T X)^{-1} = \begin{pmatrix} 0.5 & -0.5 & 0 \\ -0.5 & 1 & -0.5 \\ 0 & -0.5 & 1.5 \end{pmatrix}, \quad X^T y = \begin{pmatrix} 15 \\ 12 \\ 5 \end{pmatrix}.$$

   (a) Определите n и k.

   (b) Вычислите МНК оценку вектора β.

   (c) Найдите $\hat{\sigma}^2$, $\mathbb{E}(\hat{\sigma}^2)$.

   (d) Найдите $\mathbb{V}\mathrm{ar}(\varepsilon_1)$, $\mathbb{V}\mathrm{ar}(\beta_1)$, $\mathbb{V}\mathrm{ar}(\hat{\beta}_1)$, $\widehat{\mathbb{V}\mathrm{ar}}(\hat{\beta}_1)$, $\mathbb{E}(\hat{\beta}_1^2) - \beta_1^2$;

   (e) Найдите $\mathbb{C}\mathrm{ov}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\mathbb{C}\mathrm{ov}}(\hat{\beta}_2, \hat{\beta}_3)$, $\mathbb{V}\mathrm{ar}(\hat{\beta}_2 - \hat{\beta}_3)$, $\widehat{\mathbb{V}\mathrm{ar}}(\hat{\beta}_2 - \hat{\beta}_3)$;

   (f) Найдите $\mathbb{V}\mathrm{ar}(\beta_2 - \beta_3)$, $\mathbb{C}\mathrm{orr}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\mathbb{C}\mathrm{orr}}(\hat{\beta}_2, \hat{\beta}_3)$;

3. Вася оценил исходную модель:

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$$

   Для надежности Вася стандартизировал переменные, т.е. перешёл к $y_i^* = (y_i - \bar{Y})/\hat{\sigma}_y$ и $x_i^* = (x_i - \bar{x})/\hat{\sigma}_x$. Затем Вася оценил ещё две модели:

$$y_i^* = \beta_1' + \beta_2' x_i^* + \varepsilon_i'$$

и

$$y_i^* = \beta_2'' x_i^* + \varepsilon_i''$$

В решении можно считать $\hat{\sigma}_x$ и $\hat{\sigma}_y$ известными.

(a) Найдите $\hat{\beta}_1'$.

(b) Как связаны между собой $\hat{\beta}_2$, $\hat{\beta}_2'$ и $\hat{\beta}_2''$?

(c) Как связаны между собой e_i, e_i' и e_i''?

(d) Как связаны между собой $\widehat{Var}\left(\hat{\beta}_2\right)$, $\widehat{Var}\left(\hat{\beta}_2'\right)$ и $\widehat{Var}\left(\hat{\beta}_2''\right)$?

(e) Как выглядит матрица $\widehat{Var}\left(\hat{\beta}'\right)$?

(f) Как связаны между собой R^2, $R^{2\prime}$ и $R^{2\prime\prime}$?

(g) В нескольких предложениях прокомментируйте последствия перехода к стандартизированным переменным.