

Семинар 4.

Множественная регрессия.

1. Рассмотрим классическую линейную модель регрессии

$$y = X\beta + \varepsilon.$$

Найдите:

- (a) $\text{Cov}(e, \hat{\beta})$;
- (b) $\text{Cov}(e, y)$;
- (c) $\text{Cov}(e, \hat{y})$.

2. Пусть регрессионная модель $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, $i = 1, \dots, n$, 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 $\varepsilon \sim N(0, \sigma_\varepsilon^2 \cdot I_n)$. Известно также, что:

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

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) Рассчитайте при помощи метода наименьших квадратов оценку для вектора неизвестных коэффициентов.
- (b) Рассчитайте несмещенную оценку для неизвестного параметра σ_ε^2 регрессионной модели.
- (c) Рассчитайте $\widehat{\text{Var}}(\hat{\beta})$, оценку для ковариационной матрицы МНК-оценки $\hat{\beta}$ вектора коэффициентов β .
- (d) Рассчитайте TSS , RSS и ESS .
- (e) Сформулируйте основную и альтернативную гипотезы, которые соответствуют тесту на значимость переменной X_2 в уравнении регрессии.
- (f) Протестируйте на значимость переменную X_2 в уравнении регрессии на уровне значимости 10%:

- i. Приведите формулу для тестовой статистики.
 - ii. Укажите распределение тестовой статистики при верной H_0 .
 - iii. Вычислите наблюдаемое значение тестовой статистики.
 - iv. Укажите границы области, где основная гипотеза не отвергается.
 - v. Сделайте статистический вывод о значимости переменной X_2 .
- (g) Найдите p -value, соответствующее наблюдаемому значению тестовой статистики (t_{obs}) из предыдущего пункта. На основе полученного p -value сделайте вывод о значимости переменной X_2 .
- (h) Постройте 90%-ый доверительный интервал для оценки коэффициента β_2 .
3. По n наблюдениям 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 $\widehat{\text{Var}}(\hat{\beta}_2)$, $\widehat{\text{Var}}(\hat{\beta}'_2)$ и $\widehat{\text{Var}}(\hat{\beta}''_2)$?
- (e) Как выглядит матрица $\widehat{\text{Var}}(\hat{\beta}')$?
- (f) Как связаны между собой t -статистики $t_{\hat{\beta}_2}$, $t_{\hat{\beta}'_2}$ и $t_{\hat{\beta}''_2}$?
- (g) Как связаны между собой R^2 , $R^{2'}$ и $R^{2''}$?
- (h) В нескольких предложениях прокомментируйте последствия перехода к стандартизированным переменным.