

Семинар 4.

Модель множественной регрессии.

1. 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.

- Найдите $\hat{\beta}'_1$.
- Как связаны между собой $\hat{\beta}_2$, $\hat{\beta}'_2$ и $\hat{\beta}''_2$?
- Как связаны между собой e_i , e'_i и e''_i ?
- Как связаны между собой $\widehat{Var}(\hat{\beta}_2)$, $\widehat{Var}(\hat{\beta}'_2)$ и $\widehat{Var}(\hat{\beta}''_2)$?
- Как выглядит матрица $\widehat{Var}(\hat{\beta}')$?
- Как связаны между собой t -статистики $t_{\hat{\beta}_2}$, $t_{\hat{\beta}'_2}$ и $t_{\hat{\beta}''_2}$?
- Как связаны между собой R^2 , R'^2 и R''^2 ?
- В нескольких предложениях прокомментируйте последствия перехода к стандартизированным переменным.

2. Рассмотрим следующую модель регрессии

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + \varepsilon_i,$$

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

Ошибки ε_i независимы и нормально распределены с $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2 I$.

Для удобства расчётов даны 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) Укажите число наблюдений.
- (b) Укажите число регрессоров в модели, учитывая свободный член.
- (c) С помощью МНК найдите оценку для вектора неизвестных коэффициентов β .
- (d) Найдите $TSS = \sum_{i=1}^n (Y_i - \bar{Y})^2$.
- (e) Найдите $RSS = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$.
- (f) Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оценённого уравнения регрессии.
- (g) Сформулируйте основную и альтернативную гипотезы, которые соответствуют тесту на значимость переменной X_2 в уравнении регрессии.
- (h) 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 .
- (i) 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 .
- (j) На уровне значимости 10% проверьте гипотезу $H_0 : \beta_2 = 1$ против альтернативной $H_a : \beta_2 \neq 1$:
 - 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.
- (k) На уровне значимости 10% проверьте гипотезу $H_0 : \beta_2 = 1$ против альтернативной $H_a : \beta_2 > 1$:
 - 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.
- (l) На уровне значимости 10% проверьте гипотезу $H_0 : \beta_2 = 1$ против альтернативной $H_a : \beta_2 < 1$:
- 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.
- (m) Сформулируйте основную гипотезу, которая соответствует тесту на значимость регрессии «в целом».
- (n) На уровне значимости 5% проверьте гипотезу о значимости регрессии «в целом»:
- 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.
- (o) Найдите p -value, соответствующее наблюдаемому значению тестовой статистики (T_{obs}) из предыдущего пункта. На основе полученного p -value сделайте вывод о значимости регрессии «в целом».
- (p) На уровне значимости 5% проверьте гипотезу $H_0 : \beta_2 + \beta_3 = 2$ против альтернативной $H_a : \beta_2 + \beta_3 \neq 2$:
- 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.
- (q) На уровне значимости 5% проверьте гипотезу $H_0 : \beta_2 + \beta_3 = 2$ против альтернативной $H_a : \beta_2 + \beta_3 > 2$:
- 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.
- (г) На уровне значимости 5% проверьте гипотезу $H_0 : \beta_2 + \beta_3 = 2$ против альтернативной $H_a : \beta_2 + \beta_3 < 2$:
- 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.