

## Семинар 5.

## Проверка гипотез и блочные матрицы.

**Задание 1.** Всегда ли доверительный интервал для коэффициентов регрессии  $\beta_1 + \beta_2$  шире каждого доверительного интервала для  $\beta_1$  и  $\beta_2$ ? Если да, то почему?

**Задание 2.** Рассмотрим модель  $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i$ . При оценке модели по 24 наблюдениям оказалось, что  $RSS = 15$ ,  $\sum (Y_i - \bar{Y} - X_{i3} + \bar{X}_3)^2 = 20$ . На уровне значимости 1% протестируйте гипотезу

$$H_0 : \begin{cases} \beta_2 + \beta_3 + \beta_4 = 1 \\ \beta_2 = 0 \\ \beta_3 = 1 \\ \beta_4 = 0 \end{cases}$$

**Задание 3.** В файле *dataflats.xlsx* хранятся данные о стоимости квартир в Москве (тыс.долл.).

- (а) Оцените следующие модели регрессии для стоимости одного квадратного метра жилья:

$$price\_sq_i = \beta_1 + \beta_2 livesp_i + \beta_3 dist_i + \varepsilon_i,$$

$$price\_sq_i = \beta'_1 + \beta'_2 livesp_i + \beta'_3 dist_i + \beta'_4 metrdist_i + \varepsilon_i.$$

- (b) Для оцененных моделей проверьте гипотезу о незначимости модели в целом.
- (c) Используя  $p$  – value коэффициентов, укажите для каждой из моделей, какие из переменных являются значимыми, а какие — незначимыми на 5% уровне значимости?
- (d) Проинтерпретируйте оценки коэффициентов при значимых переменных. Согласуются ли знаки данных оценок с интуицией?
- (e) Постройте 90%-ые доверительные интервалы для коэффициентов обеих моделей.
- (f) Для каждой из моделей проверьте гипотезу (с помощью t-теста и с помощью F-теста) о том, что коэффициент при расстоянии до центра ( $dist$ ) равен -0.1. Содержательно проинтерпретируйте результаты тестирования.
- (g) Для второй модели проверьте гипотезу  $H_0 : \beta'_2 + \beta'_4 = 0$ . Содержательно проинтерпретируйте результаты тестирования.
- (h) Переоцените вторую модель регрессии на шакалированных данных. Какой из факторов оказывает наибольшее влияние на стоимость квартиры? Вспомните выводы о

соотношении оценок регрессии для исходных переменных и регрессии для стандартизированных показателей, полученные на предыдущем семинаре. Сравните теоретические выводы с полученными практическими результатами.

**Умножение блочных матриц.** Если размеры блоков допускают операцию умножения, то:

$$\left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \cdot \left[ \begin{array}{c|c} E & F \\ \hline G & H \end{array} \right] = \left[ \begin{array}{c|c} AE + BG & AF + BH \\ \hline CE + DG & CF + DH \end{array} \right].$$

**Транспонирование блочных матриц.** Пусть

$$M = \left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right].$$

Тогда

$$M' = \left[ \begin{array}{c|c} A' & C' \\ \hline B' & D' \end{array} \right].$$

**Формула Фробениуса (блочное обращение).**

$$\left[ \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]^{-1} = \left[ \begin{array}{c|c} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ \hline -H^{-1}CA^{-1} & H^{-1} \end{array} \right],$$

где  $A$  — невырожденная квадратная матрица размерности  $n \times n$ ,  $D$  — квадратная матрица размерности  $k \times k$ ,  $H = D - CA^{-1}B$ .

**Задание 4.** Вместо того чтобы оценивать параметры  $\beta_1, \beta_2$  в модели

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon$$

( $X_1, X_2$  —  $n \times k_1, n \times k_2$  матрицы соответственно,  $\beta_1, \beta_2$  — векторы размерности  $k_1, k_2$  соответственно), строятся МНК-оценки этих параметров исходя из модели

$$y = X_1^*\beta_1 + X_2\beta_2 + \varepsilon^*,$$

где  $X_1^*$  — матрица остатков, полученных в результате регрессии каждого столбца матрицы  $X_1$  на  $X_2$ .

- (а) Покажите, что полученная таким образом оценка вектора  $\beta_2$  совпадает с оценкой, полученной в результате регрессии  $y$  только на  $X_2$ .

- (b) Найдите смещение оценки вектора  $\beta_2$ .

**Задание 5.** Для блочной матрицы

$$\left[ \begin{array}{c|c} A & B \\ \hline 0 & D \end{array} \right]$$

найдите обратную матрицу.

*Предположения:*  $\det A_{n \times n} \neq 0$ ;  $\det D_{k \times k} \neq 0$ .

**Задание 6.**

- (a) Рассмотрим модель множественной регрессии:

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon.$$

Что произойдет с МНК–оценками коэффициентов модели, если добавить константу  $c_1$  к каждому наблюдению признака  $X_2$  и другую константу  $c_2$  к каждому наблюдению признака  $X_3$ ?

- (b) Что произойдет с МНК–оценками коэффициентов множественной регрессии, если умножить зависимую переменную  $y$  на константу  $c$ ? А если на константу умножить какой-нибудь регрессор?