Семинар 5.

Множественная регрессия.

Стандартизированные данные. Тестирование гипотез.

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.

(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 t-статистики $t_{\hat{\beta}_2}$, $t_{\hat{\beta}_2'}$ и $t_{\hat{\beta}_2''}$?

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

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

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

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

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 $k \times k$, $H = D - CA^{-1}B$.

2. Рассмотрим модель множественной регрессии:

$$y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i.$$

Мы хотим проверить гипотезу следующего вида:

$$H_0 : \beta_{k-q+1} = \beta_{k-q+2} = \ldots = \beta_k = 0.$$

Покажите, что для проверки гипотезы H_0 может быть использована статстистика

$$F = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n-k)} \sim F(q, n-k).$$

3. 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

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