

Семинар 2.

Модель множественной регрессии.

Тестирование гипотез.

1. Рассмотрим нормальную классическую линейную модель множественной регрессии $y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i$ с неслучайными регрессорами. Дополнительно известно, что на самом деле $\beta_2 = \dots = \beta_k = 0$.

(a) Найдите $\mathbb{E}(R^2)$.

(b) Найдите $\mathbb{E}(R_{adj}^2)$.

(c) Покажите, что $nR^2 \sim \chi^2(k-1)$.

2. Ниже представлены результаты МНК-оценивания двух регрессий, часть из которых не сохранилась. Утерянные оценки коэффициентов регрессий заменены символами.

$$\text{Модель 1: } \hat{y}_i = \hat{\beta}_1 + \underset{(2)}{\hat{\beta}_2} x_i$$

$$\text{Модель 2: } \hat{y}_i = \hat{\alpha}_1 + \hat{\alpha}_2 x_i + \underset{(2)}{10w_i}, R^2 = 0.8$$

Оценивание проводилось по 103 наблюдениям. В скобках под оценками коэффициентов указаны их стандартные ошибки. Восстановите значение оценки коэффициента $\hat{\beta}_2$ первой регрессии?

3. Рассмотрим модель $Y_i = \beta_1 + \beta_2 X_{i1} + \beta_3 X_{i2} + \beta_4 X_{i3} + \varepsilon_i$. При оценке модели по 24 наблюдениям оказалось, что $RSS = 15$, $\sum (Y_i - \bar{Y} - X_{i2} + \bar{X}_2)^2 = 20$. На уровне значимости 1% протестируйте гипотезу

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

4. Пусть x — стаж сотрудника (в годах), а y — ежемесячная заработная плата (в рублях). В результате оценивания регрессии вида

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$$

были получены следующие результаты:

Постройте 95%-ый доверительный интервал для значения цены $x = x_0$, при котором выручка максимальна.

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 50				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	100.2079	1.967171	50.94010	0.0000
X	10.03677	0.903745	11.10576	0.0000
X2	-0.817382	0.084591	-9.662803	0.0000
R-squared	0.765563	Mean dependent var		123.4150
Adjusted R-squared	0.755587	S.D. dependent var		8.089480
S.E. of regression	3.999287	Akaike info criterion		5.668234
Sum squared resid	751.7320	Schwarz criterion		5.782955
Log likelihood	-138.7058	Hannan-Quinn criter.		5.711920
F-statistic	76.74024	Durbin-Watson stat		1.738403
Prob(F-statistic)	0.000000			

Coefficients covariance matrix			
	C	X	X2
C	3.869764	-1.598561	0.134292
X	-1.598561	0.816755	-0.074654
X2	0.134292	-0.074654	0.007156