

Семинар 3.

Спецификация модели и гетероскедастичность.

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

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, i = 1, \dots, n$$

представлены в таблице ниже. Постройте 95%-ый доверительный интервал для величины стажа $x = x_0$, при котором заработная плата 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

2. Рассмотрим следующую регрессионную модель, в которой $2n$ наблюдений разбиты на две равные группы по n наблюдений в каждой:

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

$$\mathbb{E}(\varepsilon) = 0; \text{Cov}(\varepsilon_t, \varepsilon_s) = 0, t \neq s$$

$$\text{Var}(\varepsilon_t) = \sigma_1^2, t = 1, \dots, n; \text{Var}(\varepsilon_t) = \sigma_2^2, t = n + 1, \dots, 2n.$$

Введём естественное разбиение матриц на блоки:

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

- (а) Выведите следующие формулы для GLS-оценок:

$$\hat{\beta}_{GLS} = \left(\frac{X_1' X_1}{\sigma_1^2} + \frac{X_2' X_2}{\sigma_2^2} \right)^{-1} \left(\frac{X_1' y_1}{\sigma_1^2} + \frac{X_2' y_2}{\sigma_2^2} \right),$$

$$\text{Var}(\hat{\beta}_{GLS}) = \left(\frac{X_1' X_1}{\sigma_1^2} + \frac{X_2' X_2}{\sigma_2^2} \right)^{-1}.$$

- (б) Опишите процедуру получения FGLS-оценок для данной модели.

3. Рассмотрим модель

$$y_i = \beta x_i + \varepsilon_i, i = 1, \dots, n.,$$

где $E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \alpha x_i^2$, $E(\varepsilon_i \varepsilon_j) = 0$ при $i \neq j$ и $\sum_{i=1}^n x_i^2 = n$.

- (а) Покажите, что МНК-оценка $\hat{\beta}$ параметра β является несмещенной, но неэффективной.
- (б) Покажите, что стандартная оценка дисперсии $\hat{\beta}$ смещена вниз по отношению к истинной дисперсии β .

4. Дана стандартная модель парной регрессии

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i, i = 1, \dots, n.$$

- (а) Чему равна МНК-оценка коэффициента β_2 при ограничении $\beta_1 = 0$.
- (б) Чему равна дисперсия оценки в пункте (а)? Покажите, что она меньше, чем $\sigma^2 / \sum_{i=1}^n (x_i - \bar{x})^2$ — дисперсия МНК-оценки β_2 в регрессии без ограничения. Противоречит ли это теореме Гаусса–Маркова?

5. Найдите наиболее эффективную оценку коэффициента β_1 для модели

$$y_i = \beta_1 + \varepsilon_i,$$

$$\mathbb{E}(\varepsilon_i) = 0, \mathbb{E}(\varepsilon_i \varepsilon_j) = 0, \text{Var}(\varepsilon_i) = \sigma_\varepsilon^2 / x_i, x_i > 0$$

в классе линейных несмещённых оценок. Рассчитайте дисперсию этой оценки и сравните её с дисперсией МНК-оценки.

Список используемой литературы.

Магнус Я.Р., Катышев П.К., Пересецкий А.А. (2007). Эконометрика. Начальный курс : учебник для вузов.