

Семинар 1.

Модель множественной регрессии.

1. Предположим, что модель

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

удовлетворяет условиям классической регрессии. Рассматривается следующая оценка коэффициента β :

$$\tilde{\beta} = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \bar{y}}{x_i - \bar{x}}.$$

- (a) Является ли оценка $\tilde{\beta}$ несмещенной? Является ли она линейной по y ?
- (b) Вычислите дисперсию оценки $\tilde{\beta}$.
- (c) Проверьте теорему Гаусса–Маркова, сравнив полученную дисперсию оценки $\tilde{\beta}$ с дисперсией МНК–оценки параметра β .

Решение:

- а) Рассчитаем математическое ожидание оценки $\tilde{\beta}$:

$$\begin{aligned} \mathbb{E}(\tilde{\beta}) &= \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{E}(y_i) - \mathbb{E}(\bar{y})}{x_i - \bar{x}} = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{E}(\alpha + \beta x_i + \varepsilon_i) - \mathbb{E}(\alpha + \beta \bar{x} + \bar{\varepsilon})}{x_i - \bar{x}} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{(\alpha + \beta x_i) - (\alpha + \beta \bar{x})}{x_i - \bar{x}} = \frac{1}{n} \sum_{i=1}^n \frac{\beta (x_i - \bar{x})}{x_i - \bar{x}} = \beta. \end{aligned}$$

Следовательно, оценка $\tilde{\beta}$ является несмещенной. Представим $\tilde{\beta}$ в следующем виде:

$$\begin{aligned} \tilde{\beta} &= \frac{1}{n} \sum_{i=1}^n \frac{y_i - \bar{y}}{x_i - \bar{x}} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i - \bar{x}} - \frac{1}{n} \sum_{j=1}^n \frac{\bar{y}}{x_j - \bar{x}} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i - \bar{x}} y_i - \frac{1}{n^2} \left(\sum_{i=1}^n y_i \right) \left(\sum_{j=1}^n \frac{1}{x_j - \bar{x}} \right) \\ &= \sum_{i=1}^n \frac{1}{n} \left(\frac{1}{x_i - \bar{x}} - \frac{1}{n} \sum_{j=1}^n \frac{1}{x_j - \bar{x}} \right) y_i = \sum_{i=1}^n c_i y_i. \end{aligned}$$

Так как переменные $x_i - \bar{x}$ являются неслучайными, то оценка $\tilde{\beta}$ является линейной.

- б) Вычислим дисперсию оценки $\tilde{\beta}$. При этом воспользуемся тем, что Y_t — некоррелированные случайные величины с дисперсиями $V(y_i) = \sigma^2$.

$$V(\tilde{\beta}) = V\left(\sum_{i=1}^n c_i y_i\right) = \sum_{i=1}^n c_i^2 V(y_i) = \sigma^2 \sum_{i=1}^n c_i^2 = \sigma^2 \sum_{i=1}^n \frac{1}{n^2} \left(\frac{1}{x_i - \bar{x}} - \frac{1}{n} \sum_{s=1}^n \frac{1}{x_s - \bar{x}} \right)^2$$

- в) Для проверки теоремы Гаусса-Маркова воспользуемся неравенством Коши-Буняковского.

$$\begin{aligned}\frac{V(\tilde{\beta})}{V(\hat{\beta}_{OLS})} &= \left(\sigma^2 \frac{1}{n^2} \sum_{i=1}^n \left(\frac{1}{x_i - \bar{x}} - \frac{1}{n} \sum_{j=1}^n \frac{1}{x_j - \bar{x}} \right)^2 \right) \left(\frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right) \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n \left(\frac{1}{x_i - \bar{x}} - \frac{1}{n} \sum_{j=1}^n \frac{1}{x_j - \bar{x}} \right)^2 \right) \left(\sum_{i=1}^n (x_i - \bar{x})^2 \right) \\ &\geq \frac{1}{n^2} \left(\sum_{i=1}^n \left(\frac{1}{x_i - \bar{x}} - \frac{1}{n} \sum_{j=1}^n \frac{1}{x_j - \bar{x}} \right) (x_i - \bar{x}) \right)^2 = \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n \left(1 - \frac{1}{n} \sum_{j=1}^n \frac{x_i - \bar{x}}{x_j - \bar{x}} \right) \right)^2 = \\ &= \frac{1}{n^2} \left(n - \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{x_j - \bar{x}} \sum_{i=1}^n (x_i - \bar{x}) \right) \right)^2 = \frac{1}{n^2} n^2 = 1\end{aligned}$$

Здесь мы также использовали тождество $\sum_{i=1}^n (x_i - \bar{x}) = 0$. Таким образом, мы получили, что $V(\tilde{\beta}) \geq V(\hat{\beta}_{OLS})$.

2. Модель, порождающая данные, имеет вид $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, \dots, n$. Ошибки независимы, и их дисперсии имеют вид $V(\varepsilon_i) = \sigma^2$. Регрессоры детерминированы. Для оценки дисперсии σ^2 используется формула $s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$. Является ли $\hat{\sigma}^2$ несмещенной оценкой σ^2 ? Если оценка смещена то что можно сказать о знаке смещения?

Решение:

$$\begin{aligned}\mathbb{E} \sum (y_i - \bar{y})^2 &= \sum \mathbb{E} (y_i - \bar{y})^2 = \sum \mathbb{E} (\alpha + \beta x_i + \varepsilon_i - \alpha - \beta \bar{x} - \bar{\varepsilon})^2 = \\ &= \sum \mathbb{E} (\beta (x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon}))^2 = \\ &= \sum \mathbb{E} (\beta^2 (x_i - \bar{x})^2 + 2\beta (x_i - \bar{x}) (\varepsilon_i - \bar{\varepsilon}) + (\varepsilon_i - \bar{\varepsilon})^2) = \\ &= \beta^2 \sum (x_i - \bar{x})^2 + 0 + \sum \mathbb{E} (\varepsilon_i - \bar{\varepsilon})^2 = \beta^2 \sum (x_i - \bar{x})^2 + (n-1)\sigma^2.\end{aligned}$$

Поэтому, $\mathbb{E} \hat{\sigma}^2 = \mathbb{E} \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} (\beta^2 \sum (x_i - \bar{x})^2 + (n-1)\sigma^2) = \sigma^2 + \frac{1}{n-1} \beta^2 \sum (x_i - \bar{x})^2 \geq \sigma^2$. Следовательно, оценка, вообще говоря, смещенная (кроме случая $\beta = 0$). Смещение положительно.

3. Рассмотрим оценку вида $\tilde{\beta} = ((X'X)^{-1} + \gamma I)X'y$ для вектора коэффициентов регрессионного уравнения $y = X\beta + \varepsilon$, удовлетворяющего условиям классической регрессионной модели. Найдите $E(\tilde{\beta})$ и $Var(\tilde{\beta})$.

Решение:

Преобразуем выражение для $\tilde{\beta}$:

$$\begin{aligned}\tilde{\beta} &= \left((\mathbf{X}'\mathbf{X})^{-1} + \gamma\mathbf{I} \right) \mathbf{X}'\mathbf{y} = \left((\mathbf{X}'\mathbf{X})^{-1} + \gamma\mathbf{I} \right) \mathbf{X}'(\mathbf{X}\beta + \varepsilon) \\ &= \beta + \gamma\mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\varepsilon + \gamma\mathbf{X}'\varepsilon = \beta + \gamma\mathbf{X}'\mathbf{X}\beta + \mathbf{C}\varepsilon,\end{aligned}$$

где $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' + \gamma\mathbf{X}'$.

Так как $\mathbb{E}(\varepsilon) = \mathbf{0}$, то $\mathbb{E}(\tilde{\beta}) = \beta + \gamma\mathbf{X}'\mathbf{X}\beta$. Далее, поскольку вектор $\beta + \gamma\mathbf{X}'\mathbf{X}\beta$ является неслучайным, то

$$\begin{aligned}V(\tilde{\beta}) &= V(\mathbf{C}\varepsilon) = \sigma^2 \mathbf{C}\mathbf{C}' = \sigma^2 \left((\mathbf{X}'\mathbf{X})^{-1} + \gamma\mathbf{I} \right) \mathbf{X}'\mathbf{X} \left((\mathbf{X}'\mathbf{X})^{-1} + \gamma\mathbf{I} \right) \\ &= \sigma^2 \left((\mathbf{X}'\mathbf{X})^{-1} + 2\gamma\mathbf{I} + \gamma^2\mathbf{X}'\mathbf{X} \right) = V\left(\hat{\beta}_{\text{OLS}}\right) + \sigma^2 (2\gamma\mathbf{I} + \gamma^2\mathbf{X}'\mathbf{X}).\end{aligned}$$

4. (Математическое ожидание квадратичной формы) Пусть y — случайный вектор-столбец размерности $n \times 1$, A — детерминированная матрица размерности $n \times n$. Покажите, что справедливо следующее:

$$\mathbb{E}(y' Ay) = \text{tr}(A \text{Var}(y)) + \mathbb{E}(y') A \mathbb{E}(y).$$

5. Используя матрицы $P = X(X'X)^{-1}X'$ и $\pi = \vec{1}(\vec{1}'\vec{1})^{-1}\vec{1}'$,

- (а) запишите TSS, RSS и ESS в матричной форме;
- (б) вычислите $\mathbb{E}(\text{TSS})$, $\mathbb{E}(\text{ESS})$.

Примечание: $\vec{1}$ — вектор размерности $n \times 1$, состоящий из единиц.

Список использованных источников

1. Катышев П.К., Магнус Я.Р., Пересецкий А.А., Головань С.В. Сборник задач к начальному курсу эконометрики: Учеб. пособие. — 4-е изд., перераб. и доп. — М.: Дело, 2007. — 368 с.