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 $\widetilde{\beta}$:

$$\mathbb{E}(\widetilde{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}\left(y_i\right) - \mathbb{E}(\bar{y})}{x_i - \bar{x}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{E}\left(\alpha + \beta x_i + \varepsilon_i\right) - \mathbb{E}(\alpha + \beta\bar{x} + \bar{\varepsilon})}{x_i - \bar{x}}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\frac{\left(\alpha + \beta x_i\right) - \left(\alpha + \beta\bar{x}\right)}{x_i - \bar{x}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\beta\left(x_i - \bar{x}\right)}{x_i - \bar{x}} = \beta.$$

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

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

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

б) Вычислим дисперсию оценки $\widetilde{\beta}$. При этом воспользуемся тем, что Y_t– некоррелированные случайные величины с дисперсиями $\mathrm{V}\left(y_i\right) = \sigma^2$.

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

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

$$\frac{\mathrm{V}(\widetilde{\beta})}{\mathrm{V}\left(\hat{\beta}_{\mathrm{OLS}}\right)} = \left(\sigma^2 \frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{1}{x_i-\bar{x}} - \frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j-\bar{x}}\right)^2\right)\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\left(\frac{1}{x_i-\bar{x}} - \frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j-\bar{x}}\right)^2\right)\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$$

$$\geqslant \frac{1}{n^2}\left(\sum_{i=1}^{n}\left(\frac{1}{x_i-\bar{x}} - \frac{1}{n}\sum_{j=1}^{n}\frac{1}{x_j-\bar{x}}\right)(x_i-\bar{x})\right)^2 =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\left(1 - \frac{1}{n}\sum_{j=1}^{n}\frac{x_i-\bar{x}}{x_j-\bar{x}}\right)\right)^2 =$$

$$= \frac{1}{n^2}\left(n - \frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{x_j-\bar{x}}\sum_{i=1}^{n}(x_i-\bar{x})\right)\right)^2 = \frac{1}{n^2}n^2 = 1$$

Здесь мы также использовали тождество $\sum_{i=1}^{n}(x_i-\bar{x}) = 0$. Таким образом, мы получили, что $\mathrm{V}(\widetilde{\beta}) \geqslant \mathrm{V}\left(\widehat{\beta}_{\mathrm{OLS}}\right)$.

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

Решение:

$$\mathbb{E}\sum(y_i-\bar{y})^2 = \sum\mathbb{E}(y_i-\bar{y})^2 = \sum\mathbb{E}(\alpha+\beta x_i+\varepsilon_i-\alpha-\beta\bar{x}-\bar{\varepsilon})^2 =$$

$$= \sum\mathbb{E}(\beta(x_i-\bar{x})+(\varepsilon_i-\bar{\varepsilon}))^2 =$$

$$= \sum\mathbb{E}\left(\beta^2(x_i-\bar{x})^2+2\beta(x_i-\bar{x})(\varepsilon_i-\bar{\varepsilon})+(\varepsilon_i-\bar{\varepsilon})^2\right)^2 =$$

$$= \beta^2\sum(x_i-\bar{x})^2+0+\sum\mathbb{E}(\varepsilon_i-\bar{\varepsilon})^2 = \beta^2\sum(x_i-\bar{x})^2+(n-1)\sigma^2.$$

Поэтому, $\mathbb{E}s^2 = E\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2 = \frac{1}{n-1}\left(\beta^2\sum(x_i-\bar{x})^2+(n-1)\sigma^2\right) = \sigma^2 + \frac{1}{n-1}\beta^2\sum(x_i-\bar{x})^2 \geq \sigma^2$. 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 $\widetilde{\boldsymbol{\beta}}$:

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

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

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

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

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:

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

5. Используя матрицы $\mathrm{P} = X(X'X)^{-1}X'$ и $\pi = \vec{1}(\vec{1}'\vec{1})^{-1}\vec{1}'$,
   (a) запишите TSS, RSS и ESS в матричной форме;
   (b) вычислите $\mathbb{E}(\mathrm{TSS})$, $\mathbb{E}(\mathrm{ESS})$.

   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.