Econometrics II

(Tue., 8:50-10:20)

TA Session (by Mr. Kinoshita):

Thu., 14:40 - 16:10

Room # 605 (法経大学院総合研究棟)

Econometrics (Undergraduate Course)

Wed., 10:30-12:00

Fri., 8:50-10:20

- If you have not taken Econometrics in undergraduate level, attend the class.
- Textbook: 『計量経済学』(山本 拓 著，新世社)
- The prerequisite of this class is to have knowledge of **Econometrics I** (last semistar) and **Econometrics** (undergraduate level).

1 Regression Analysis (回帰分析)

1.1 Setup of the Model

When (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n) are available, suppose that there is a linear relationship between y and x, i.e.,

$$y_i = \beta_1 + \beta_2 x_i + u_i, \tag{1}$$

for $i = 1, 2, \cdots, n$. x_i and y_i denote the ith observations.

$\longrightarrow$ **Single (or simple) regression model (単回帰モデル)**

y_i is called the **dependent variable (従属変数)** or the **explained variable (被説明変数)**, while x_i is known as the **independent variable (独立変数)** or the **explanatory (or explaining) variable (説明変数)**.

β_1 = **Intercept (切片)** β_2 = **Slope (傾き)**

β_1 and β_2 are unknown **parameters (パラメータ，母数)** to be estimated.

β_1 and β_2 are called the **regression coefficients** (回帰係数).

u_i is the unobserved **error term (誤差項)** assumed to be a random variable with mean zero and variance σ^2.
σ^2 is also a parameter to be estimated.

x_i is assumed to be **nonstochastic (非確率的)**, but y_i is **stochastic (確率的)** because y_i depends on the error u_i.

The error terms u_1, u_2, $\cdots$, u_n are assumed to be mutually independently and identically distributed, which is called ***iid***.

It is assumed that u_i has a distribution with mean zero, i.e., $\mathrm{E}(u_i) = 0$ is assumed.

Taking the expectation on both sides of (1), the expectation of y_i is represented as:

$$\begin{aligned}\mathrm{E}(y_i) &= \mathrm{E}(\beta_1 + \beta_2 x_i + u_i) = \beta_1 + \beta_2 x_i + \mathrm{E}(u_i) \\ &= \beta_1 + \beta_2 x_i, \end{aligned} \tag{2}$$

for $i = 1, 2, \cdots, n$.

Using $\mathrm{E}(y_i)$ we can rewrite (1) as $y_i = \mathrm{E}(y_i) + u_i$.

(2) represents the true regression line.

Let $\hat\beta_1$ and $\hat\beta_2$ be estimates of β_1 and β_2.

Replacing β_1 and β_2 by $\hat\beta_1$ and $\hat\beta_2$, (1) turns out to be:

$$y_i = \hat\beta_1 + \hat\beta_2 x_i + e_i, \tag{3}$$

for $i = 1, 2, \cdots, n$, where e_i is called the **residual (残差)**.

The residual e_i is taken as the experimental value (or realization) of u_i.

We define $\hat y_i$ as follows:

$$\hat y_i = \hat\beta_1 + \hat\beta_2 x_i, \tag{4}$$

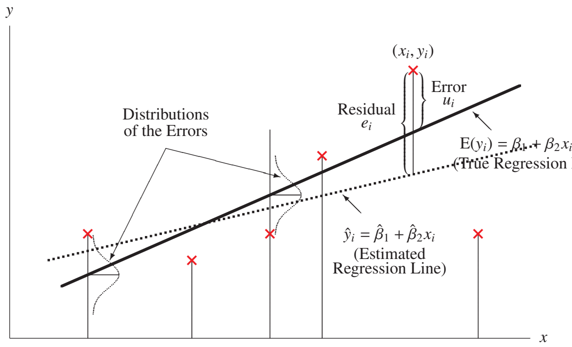
for $i = 1, 2, \cdots, n$, which is interpreted as the **predicted value (予測値)** of y_i.

(4) indicates the estimated regression line, which is different from (2).

Moreover, using $\hat y_i$ we can rewrite (3) as $y_i = \hat y_i + e_i$.

(2) and (4) are displayed in Figure 1.

Figure 1. True and Estimated Regression Lines (回帰直線)



Consider the case of $n = 6$ for simplicity.

$\times$ indicates the observed data series.

The true regression line (2) is represented by the solid line, while the estimated regression line (4) is drawn with the dotted line.

Based on the observed data, β_1 and β_2 are estimated as: $\hat\beta_1$ and $\hat\beta_2$.

In the next section, we consider how to obtain the estimates of β_1 and β_2, i.e., $\hat\beta_1$ and $\hat\beta_2$.

1.2 Ordinary Least Squares Estimation

Suppose that (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n) are available.
For the regression model (1), we consider estimating β_1 and β_2.
Replacing β_1 and β_2 by their estimates $\hat\beta_1$ and $\hat\beta_2$, remember that the residual e_i is given by:

$$e_i = y_i - \hat y_i = y_i - \hat\beta_1 - \hat\beta_2 x_i.$$

The sum of squared residuals is defined as follows:

$$S(\hat\beta_1, \hat\beta_2) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat\beta_1 - \hat\beta_2 x_i)^2.$$

It might be plausible to choose the $\hat\beta_1$ and $\hat\beta_2$ which minimize the sum of squared residuals, i.e., $S(\hat\beta_1, \hat\beta_2)$.
This method is called the **ordinary least squares estimation (最小二乗法，OLS)**.
To minimize $S(\hat\beta_1, \hat\beta_2)$ with respect to $\hat\beta_1$ and $\hat\beta_2$, we set the partial derivatives equal to zero:

$$\frac{\partial S(\hat\beta_1, \hat\beta_2)}{\partial \hat\beta_1} = -2 \sum_{i=1}^n (y_i - \hat\beta_1 - \hat\beta_2 x_i) = 0,$$

$$\frac{\partial S(\hat\beta_1, \hat\beta_2)}{\partial \hat\beta_2} = -2 \sum_{i=1}^n x_i (y_i - \hat\beta_1 - \hat\beta_2 x_i) = 0,$$

which yields the following two equations:

$$\overline{y} = \hat{\beta}_1 + \hat{\beta}_2\overline{x}, \tag{5}$$

$$\sum_{i=1}^{n} x_i y_i = n\overline{x}\hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^{n} x_i^2, \tag{6}$$

where $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

Multiplying (5) by $n\overline{x}$ and subtracting (6), we can derive $\hat{\beta}_2$ as follows:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{7}$$

From (5), $\hat{\beta}_1$ is directly obtained as follows:

$$\hat{\beta}_1 = \overline{y} - \hat{\beta}_2\overline{x}. \tag{8}$$

When the observed values are taken for y_i and x_i for $i = 1, 2, \cdots, n$, we say that $\hat{\beta}_1$ and $\hat{\beta}_2$ are called the **ordinary least squares estimate**s (or simply the **least squares estimate**s, **最小二乗推定値**) of β_1 and β_2.

When y_i for $i = 1, 2, \cdots, n$ are regarded as the random sample, we say that $\hat{\beta}_1$ and $\hat{\beta}_2$ are called the **ordinary least squares estimator**s (or the **least squares estimator**s, **最小二乗推定量**) of β_1 and β_2.

1.3 Properties of Least Squares Estimator

Equation (7) is rewritten as:

$$\begin{aligned}
\hat{\beta}_2 &= \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2} - \frac{\overline{y}\sum_{i=1}^{n}(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \\
&= \sum_{i=1}^{n} \frac{x_i - \overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2} Y_i = \sum_{i=1}^{n} \omega_i y_i.
\end{aligned} \tag{9}$$

In the third equality, $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$ is utilized because of $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$.

In the fourth equality, ω_i is defined as: $\omega_i = \dfrac{x_i - \overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$.

ω_i is nonstochastic because x_i is assumed to be nonstochastic.

ω_i has the following properties:

$$\sum_{i=1}^{n} \omega_i = \sum_{i=1}^{n} \frac{x_i - \overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 0, \tag{10}$$

$$\sum_{i=1}^{n} \omega_i x_i = \sum_{i=1}^{n} \omega_i(x_i - \overline{x}) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = 1, \tag{11}$$

$$\sum_{i=1}^{n} \omega_i^2 = \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{\left(\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^2} = \frac{1}{\sum_{i=1}^{n}(x_i - \overline{x})^2}. \tag{12}$$

The first equality of (11) comes from (10).

From now on, we focus only on $\hat{\beta}_2$, because usually β_2 is more important than β_1 in the regression model (1).

In order to obtain the properties of the least squares estimator $\hat{\beta}_2$, we rewrite (9) as:

$$\begin{aligned}
\hat{\beta}_2 &= \sum_{i=1}^{n} \omega_i y_i = \sum_{i=1}^{n} \omega_i(\beta_1 + \beta_2 x_i + u_i) \\
&= \beta_1 \sum_{i=1}^{n} \omega_i + \beta_2 \sum_{i=1}^{n} \omega_i x_i + \sum_{i=1}^{n} \omega_i u_i = \beta_2 + \sum_{i=1}^{n} \omega_i u_i.
\end{aligned} \tag{13}$$

In the fourth equality of (13), (10) and (11) are utilized.

Mean and Variance of $\hat{\beta}_2$: $u_1, u_2, \cdots, u_n$ are assumed to be mutually independently and identically distributed with mean zero and variance σ^2, but they are not necessarily normal.

Remember that we do not need normality assumption to obtain mean and variance but the normality assumption is required to test a hypothesis.

From (13), the expectation of $\hat{\beta}_2$ is derived as follows:

$$\begin{aligned}
\mathrm{E}(\hat{\beta}_2) &= \mathrm{E}(\beta_2 + \sum_{i=1}^{n} \omega_i u_i) = \beta_2 + \mathrm{E}(\sum_{i=1}^{n} \omega_i u_i) \\
&= \beta_2 + \sum_{i=1}^{n} \omega_i \mathrm{E}(u_i) = \beta_2.
\end{aligned} \tag{14}$$

It is shown from (14) that the ordinary least squares estimator $\hat{\beta}_2$ is an unbiased estimator of β_2.

From (13), the variance of $\hat{\beta}_2$ is computed as:

$$\begin{aligned}\mathrm{V}(\hat{\beta}_2) &= \mathrm{V}(\beta_2 + \sum_{i=1}^n \omega_i u_i) = \mathrm{V}(\sum_{i=1}^n \omega_i u_i) = \sum_{i=1}^n \mathrm{V}(\omega_i u_i) = \sum_{i=1}^n \omega_i^2 \mathrm{V}(u_i) \\ &= \sigma^2 \sum_{i=1}^n \omega_i^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \overline{x})^2}. \end{aligned} \tag{15}$$

The third equality holds because $u_1, u_2, \cdots, u_n$ are mutually independent.

The last equality comes from (12).

Thus, $\mathrm{E}(\hat{\beta}_2)$ and $\mathrm{V}(\hat{\beta}_2)$ are given by (14) and (15).

Gauss-Markov Theorem (ガウス・マルコフ定理): It has been discussed above that $\hat{\beta}_2$ is represented as (9), which implies that $\hat{\beta}_2$ is a linear estimator, i.e., linear in y_i.

In addition, (14) indicates that $\hat{\beta}_2$ is an unbiased estimator.

Therefore, summarizing these two facts, it is shown that $\hat{\beta}_2$ is a **linear unbiased estimator (線形不偏推定量)**.

Furthermore, here we show that $\hat{\beta}_2$ has minimum variance within a class of the linear unbiased estimators.

Consider the alternative linear unbiased estimator $\widetilde{\beta}_2$ as follows:

$$\widetilde{\beta}_2 = \sum_{i=1}^n c_i y_i = \sum_{i=1}^n (\omega_i + d_i) y_i,$$

where $c_i = \omega_i + d_i$ is defined and d_i is nonstochastic.

Then, $\widetilde{\beta}_2$ is transformed into:

$$\begin{aligned}\widetilde{\beta}_2 &= \sum_{i=1}^n c_i y_i = \sum_{i=1}^n (\omega_i + d_i)(\beta_1 + \beta_2 x_i + u_i) \\ &= \beta_1 \sum_{i=1}^n \omega_i + \beta_2 \sum_{i=1}^n \omega_i x_i + \sum_{i=1}^n \omega_i u_i + \beta_1 \sum_{i=1}^n d_i + \beta_2 \sum_{i=1}^n d_i x_i + \sum_{i=1}^n d_i u_i \\ &= \beta_2 + \beta_1 \sum_{i=1}^n d_i + \beta_2 \sum_{i=1}^n d_i x_i + \sum_{i=1}^n \omega_i u_i + \sum_{i=1}^n d_i u_i. \end{aligned}$$

Equations (10) and (11) are used in the forth equality.

Taking the expectation on both sides of the above equation, we obtain:

$$\begin{aligned}\mathrm{E}(\widetilde{\beta}_2) &= \beta_2 + \beta_1 \sum_{i=1}^n d_i + \beta_2 \sum_{i=1}^n d_i x_i + \sum_{i=1}^n \omega_i \mathrm{E}(u_i) + \sum_{i=1}^n d_i \mathrm{E}(u_i) \\ &= \beta_2 + \beta_1 \sum_{i=1}^n d_i + \beta_2 \sum_{i=1}^n d_i x_i. \end{aligned}$$

Note that d_i is not a random variable and that $\mathrm{E}(u_i) = 0$.

Since $\widetilde{\beta}_2$ is assumed to be unbiased, we need the following conditions:

$$\sum_{i=1}^n d_i = 0, \qquad \sum_{i=1}^n d_i x_i = 0.$$