

Econometrics I

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

Room # 1 (法経講義棟)

- This class is based on **Statistics** (統計, 『コア・テキスト 統計学』 大屋 幸輔 著, 新世社) and **Econometrics** (計量経済, 『計量経済学』 山本 拓 著, 新世社), which are provided by Department of Economics, and **Basic Statistics** (統計基礎, 『コア・テキスト 統計学』 大屋 幸輔 著, 新世社), provided by Graduate School of Economics.

Thus, Statistics and Econometrics of undergraduate level are prerequisites.

- Furthermore, **Special Lectures in Economics (Statistical Analysis)** or **Statistical Analysis** (統計解析), provided by Graduate School of Economics, should be studied with this class.

Or, do self-study using the lecture notes of

<https://www2.econ.osaka-u.ac.jp/~tanizaki/class/2012/econome1/index.htm>

(The notes are written in English with Japanese translation for econometrics terms).

TA Session:

TAs: **Shintaro Minamoto (源 真太郎)**
 u681900g [at] ecs. osaka-u. ac. jp

Jukina Hatakeyama (畠山 樹輝凪):
 u868710a [at] ecs. osaka-u. ac. jp

Date: **Thur. 13:30-15:00**

Place: **Room #1 (1 番講義室)**

Contents: **Basic Statistics, Matrix Algebra, and etc.**

Ask TAs directly in the TA session
 if you have questions about class, homework and etc.

See

<https://www2.econ.osaka-u.ac.jp/~tanizaki/class/2025/econome1/index.htm>

1 Regression Analysis (回帰分析) — Review

1.1 Setup of the Model

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

$$y_i = \beta_1 + \beta_2 x_i + u_i, \quad (1)$$

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

→ **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)**.

$\beta_1 = \text{Intercept (切片)}, \quad \beta_2 = \text{Slope (傾き)}$

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, \dots, u_n are assumed to be mutually independently and identically distributed, which is called **iid**. \longrightarrow discussed later.

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

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

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

for $i = 1, 2, \dots, n$. Using $E(y_i)$ we can rewrite (1) as $y_i = E(y_i) + u_i$.

(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 $i = 1, 2, \dots, n$, 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:

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i, \quad (4)$$

for $i = 1, 2, \dots, n$, 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 $y_i = \hat{y}_i + e_i$.

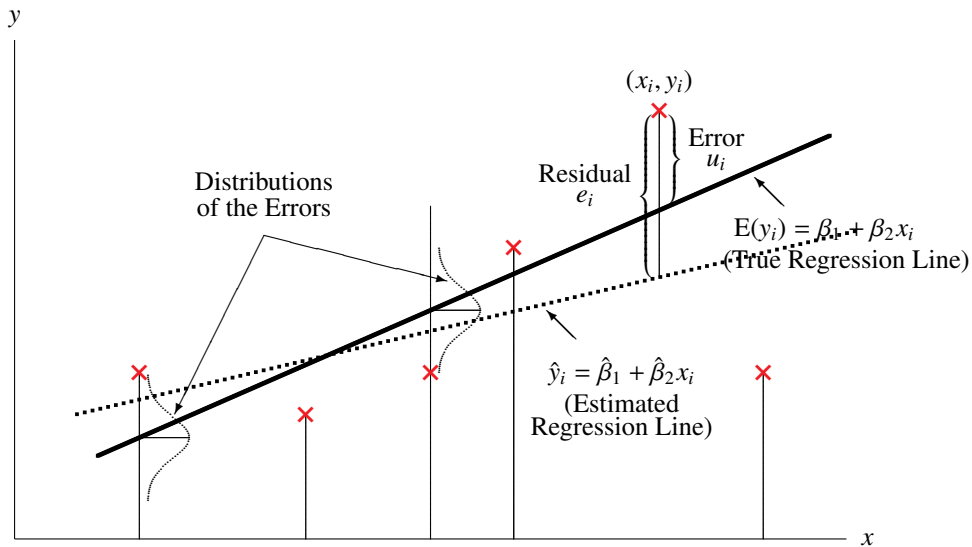
(2) and (4) are displayed in Figure 1.

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 $\hat{\beta}_2$.

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



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 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 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** (最小二乘法, 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:

$$\begin{aligned}\frac{\partial S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1} &= -2 \sum_{i=1}^n (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0, \\ \frac{\partial S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2} &= -2 \sum_{i=1}^n x_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0.\end{aligned}$$

The second order condition for minimization is:

$$\begin{pmatrix} \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1^2} & \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_1 \partial \hat{\beta}_2} \\ \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2 \partial \hat{\beta}_1} & \frac{\partial^2 S(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2^2} \end{pmatrix} = \begin{pmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{pmatrix}$$

should be a positive definite matrix.

The diagonal elements $2n$ and $2 \sum_{i=1}^n x_i^2$ are positive.

The determinant:

$$\begin{vmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{vmatrix} = 4n \sum_{i=1}^n x_i^2 - 4 \left(\sum_{i=1}^n x_i \right)^2 = 4n \sum_{i=1}^n (x_i - \bar{x})^2$$

is positive. \implies The second-order condition is satisfied.

The first two equations yield the following two equations:

$$\bar{y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{x}, \quad (5)$$

$$\sum_{i=1}^n x_i y_i = n \bar{x} \hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n x_i^2, \quad (6)$$

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

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

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

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

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}. \quad (8)$$

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

When y_i for $i = 1, 2, \dots, n$ are regarded as the random sample, we say that $\hat{\beta}_1$ and $\hat{\beta}_2$ are called the **ordinary least squares estimators** (or the **least squares estimators**, 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:

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

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

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

ω_i is nonstochastic because x_i is assumed to be nonstochastic.

ω_i has the following properties:

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

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

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

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.

[Review] Random Variables:

Let X_1, X_2, \dots, X_n be n random variables, which are mutually independently and identically distributed.

mutually independent $\implies f(x_i, x_j) = f_i(x_i)f_j(x_j)$ for $i \neq j$.

$f(x_i, x_j)$ denotes a joint distribution of X_i and X_j .

$f_i(x)$ indicates a marginal distribution of X_i .

identical $\implies f_i(x) = f_j(x)$ for $i \neq j$.

[End of Review]

[Review] Mean and Variance:

Let X and Y be random variables (continuous type), which are independently distributed.

Definition and Formulas:

- $E(g(X)) = \int g(x)f(x)dx$ for a function $g(\cdot)$ and a density function $f(\cdot)$.
- $V(X) = E((X - \mu)^2) = \int (x - \mu)^2 f(x)dx$ for $\mu = E(X)$.
- $E(aX + b) = aE(X) + b$ and $V(aX + b) = V(aX) = a^2V(X)$ for constant a and b .
- $E(X \pm Y) = E(X) \pm E(Y)$ and $V(X \pm Y) = V(X) + V(Y)$.

[End of Review]

Mean and Variance of $\hat{\beta}_2$: u_1, u_2, \dots, 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:

$$E(\hat{\beta}_2) = E(\beta_2 + \sum_{i=1}^n \omega_i u_i) = \beta_2 + E(\sum_{i=1}^n \omega_i u_i) = \beta_2 + \sum_{i=1}^n \omega_i E(u_i) = \beta_2. \quad (14)$$

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:

$$V(\hat{\beta}_2) = V(\beta_2 + \sum_{i=1}^n \omega_i u_i) = V(\sum_{i=1}^n \omega_i u_i) = \sum_{i=1}^n V(\omega_i u_i) = \sum_{i=1}^n \omega_i^2 V(u_i)$$

$$= \sigma^2 \sum_{i=1}^n \omega_i^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (15)$$

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

The last equality comes from (12).

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

[Review] Three Good Properties on Estimator:

θ : Parameter

$\hat{\theta}$: Estimator of θ , i.e., $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$,

where X_1, X_2, \dots, X_n are mutually independent random variables.

(*) Estimate of θ : $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$, where x_i denotes the observed data of X_i .

- Unbiasedness (不偏性): $E(\hat{\theta}) = \theta$.

- Efficiency (有効性):

The minimum variance estimator within all the unbiased estimators.

(*) It is not easy to check efficiency in general. Instead, consider the **best linear unbiased estimator (BLUE, 最良線型不偏推定量)**.

- Consistency (一致性): $\hat{\theta} \rightarrow \theta$ as $n \rightarrow \infty$. Note that $\hat{\theta}$ depends on # of obs.

[End of Review]