

[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]

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 $\tilde{\beta}_2$ as follows:

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

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

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

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

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

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

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

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

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

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

When these conditions hold, we can rewrite $\tilde{\beta}_2$ as:

$$\tilde{\beta}_2 = \beta_2 + \sum_{i=1}^n (\omega_i + d_i) u_i.$$

The variance of $\tilde{\beta}_2$ is derived as:

$$\begin{aligned} V(\tilde{\beta}_2) &= V\left(\beta_2 + \sum_{i=1}^n (\omega_i + d_i) u_i\right) = V\left(\sum_{i=1}^n (\omega_i + d_i) u_i\right) = \sum_{i=1}^n V\left((\omega_i + d_i) u_i\right) \\ &= \sum_{i=1}^n (\omega_i + d_i)^2 V(u_i) = \sigma^2 \left(\sum_{i=1}^n \omega_i^2 + 2 \sum_{i=1}^n \omega_i d_i + \sum_{i=1}^n d_i^2 \right) \\ &= \sigma^2 \left(\sum_{i=1}^n \omega_i^2 + \sum_{i=1}^n d_i^2 \right). \end{aligned}$$

From unbiasedness of $\tilde{\beta}_2$, using $\sum_{i=1}^n d_i = 0$ and $\sum_{i=1}^n d_i x_i = 0$, we obtain:

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

which is utilized to obtain the variance of $\tilde{\beta}_2$ in the third line of the above equation.

From (15), the variance of $\hat{\beta}_2$ is given by: $V(\hat{\beta}_2) = \sigma^2 \sum_{i=1}^n \omega_i^2$.

Therefore, we have:

$$V(\tilde{\beta}_2) \geq V(\hat{\beta}_2),$$

because of $\sum_{i=1}^n d_i^2 \geq 0$.

When $\sum_{i=1}^n d_i^2 = 0$, i.e., when $d_1 = d_2 = \cdots = d_n = 0$,

we have the equality: $V(\tilde{\beta}_2) = V(\hat{\beta}_2)$.

Thus, in the case of $d_1 = d_2 = \cdots = d_n = 0$, $\hat{\beta}_2$ is equivalent to $\tilde{\beta}_2$.

As shown above, the least squares estimator $\hat{\beta}_2$ gives us the **minimum variance linear unbiased estimator** (最小分散線形不偏推定量), or equivalently the **best linear unbiased estimator** (最良線形不偏推定量, **BLUE**), which is called the **Gauss-Markov theorem** (ガウス・マルコフ定理).

Asymptotic Properties (漸近的性質) of $\hat{\beta}_2$: We assume that as n goes to infinity we have the following:

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \longrightarrow m < \infty,$$

where m is a constant value. From (12), we obtain:

$$n \sum_{i=1}^n \omega_i^2 = \frac{1}{(1/n) \sum_{i=1}^n (x_i - \bar{x})} \longrightarrow \frac{1}{m}.$$

Note that $f(x_n) \longrightarrow f(m)$ when $x_n \longrightarrow m$, called **Slutsky's theorem (スルツキ一定理)**, where m is a constant value and $f(\cdot)$ is a function.

We show both **consistency (一致性)** of $\hat{\beta}_2$ and **asymptotic normality (漸近正規性)** of $\sqrt{n}(\hat{\beta}_2 - \beta_2)$.

● First, we prove that $\hat{\beta}_2$ is a consistent estimator of β_2 .

[Review] **Chebyshev's inequality** (チェビシェフの不等式) is given by:

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}, \quad \text{where } \mu = E(X), \sigma^2 = V(X) \text{ and any } \epsilon > 0.$$

[End of Review]

Replace X , $E(X)$ and $V(X)$ by:

$$\hat{\beta}_2, \quad E(\hat{\beta}_2) = \beta_2, \quad \text{and} \quad V(\hat{\beta}_2) = \sigma^2 \sum_{i=1}^n \omega_i^2 = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})}.$$

Then, when $n \rightarrow \infty$, we obtain the following result:

$$P(|\hat{\beta}_2 - \beta_2| > \epsilon) \leq \frac{\sigma^2 \sum_{i=1}^n \omega_i^2}{\epsilon^2} = \frac{\sigma^2 n \sum_{i=1}^n \omega_i^2}{n \epsilon^2} \rightarrow 0,$$

where $\sum_{i=1}^n \omega_i^2 \rightarrow 0$ because $n \sum_{i=1}^n \omega_i^2 \rightarrow \frac{1}{m}$ from the assumption.

Thus, we obtain the result that $\hat{\beta}_2 \rightarrow \beta_2$ as $n \rightarrow \infty$.

Therefore, we can conclude that $\hat{\beta}_2$ is a **consistent estimator** (一致推定量) of β_2 .

● Next, we want to show that $\sqrt{n}(\hat{\beta}_2 - \beta_2)$ is asymptotically normal.

[Review] The **Central Limit Theorem** (中心極限定理, CLT) is: for random variables X_1, X_2, \dots, X_n ,

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\sum_{i=1}^n X_i - E(\sum_{i=1}^n X_i)}{\sqrt{V(\sum_{i=1}^n X_i)}} \longrightarrow N(0, 1), \quad \text{as } n \longrightarrow \infty,$$

where $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

X_1, X_2, \dots, X_n are not necessarily iid, if $V(\bar{X})$ is finite as n goes to infinity.

[End of Review]

Note that $\hat{\beta}_2 = \beta_2 + \sum_{i=1}^n \omega_i u_i$ as in (13), and X_i is replaced by $\omega_i u_i$.

From the central limit theorem, asymptotic normality is shown as follows:

$$\frac{\sum_{i=1}^n \omega_i u_i - E(\sum_{i=1}^n \omega_i u_i)}{\sqrt{V(\sum_{i=1}^n \omega_i u_i)}} = \frac{\sum_{i=1}^n \omega_i u_i}{\sigma \sqrt{\sum_{i=1}^n \omega_i^2}} = \frac{\hat{\beta}_2 - \beta_2}{\sigma / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \rightarrow N(0, 1),$$

where

- $E(\sum_{i=1}^n \omega_i u_i) = 0$,
- $V(\sum_{i=1}^n \omega_i u_i) = \sigma^2 \sum_{i=1}^n \omega_i^2$, and
- $\sum_{i=1}^n \omega_i u_i = \hat{\beta}_2 - \beta_2$

are substituted in the first and second equalities.

Moreover, we can rewrite as follows:

$$\frac{\hat{\beta}_2 - \beta_2}{\sigma / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} = \frac{\sqrt{n}(\hat{\beta}_2 - \beta_2)}{\sigma / \sqrt{(1/n) \sum_{i=1}^n (x_i - \bar{x})^2}}.$$

Replacing $(1/n) \sum_{i=1}^n (x_i - \bar{x})^2$ by its converged value m , we have:

$$\frac{\sqrt{n}(\hat{\beta}_2 - \beta_2)}{\sigma / \sqrt{m}} \longrightarrow N(0, 1),$$

which implies

$$\sqrt{n}(\hat{\beta}_2 - \beta_2) \longrightarrow N(0, \frac{\sigma^2}{m}).$$

Thus, the asymptotic normality of $\sqrt{n}(\hat{\beta}_2 - \beta_2)$ is shown.