● 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 $X_1, X_2, \cdots, X_n$,

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

where $\overline{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$.

$X_1, X_2, \cdots, X_n$ are not necesarily iid, if $\mathrm{V}(\overline{X})$ 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:

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

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.

33

Moreover, we can rewrite as follows:

$$\frac{\hat{\beta}_2 - \beta_2}{\sigma / \sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}} = \frac{\sqrt{n}(\hat{\beta}_2 - \beta_2)}{\sigma / \sqrt{(1/n)\sum_{i=1}^{n}(x_i - \overline{x})^2}} \longrightarrow \frac{\sqrt{n}(\hat{\beta}_2 - \beta_2)}{\sigma / \sqrt{m}} \longrightarrow N(0, 1).$$

Or equivalently,

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

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

Finally, replacing σ^2 by its consistent estimator s^2, it is known as follows:

$$\frac{\hat{\beta}_2 - \beta_2}{s / \sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}} \longrightarrow N(0, 1), \tag{16}$$

where s^2 is defined as:

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} e_i^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)^2, \tag{17}$$

which is a consistent and unbiased estimator of σ^2. $\longrightarrow$ Proved later.

Thus, using (16), in large sample we can construct the confidence interval and test the hypothesis.

Exact Distribution of $\hat{\beta}_2$: We have shown asymptotic normality of $\sqrt{n}(\hat{\beta}_2 - \beta_2)$, which is one of the large sample properties.

Now, we discuss the small sample properties of $\hat{\beta}_2$.

In order to obtain the distribution of $\hat{\beta}_2$ in small sample, the distribution of the error term has to be assumed.

Therefore, the extra assumption is that $u_i \sim N(0, \sigma^2)$.

Writing (13), again, $\hat{\beta}_2$ is represented as:

$$\hat{\beta}_2 = \beta_2 + \sum_{i=1}^{n} \omega_i u_i.$$

First, we obtain the distribution of the second term in the above equation.

[Review] Note that the **moment-generating function** (積率母関数, MGF) is given by $M(\theta) \equiv \mathrm{E}(\exp(\theta X)) = \exp(\mu\theta + \frac{1}{2}\sigma^2\theta^2)$ when $X \sim N(\mu, \sigma^2)$.

$X_1, X_2, \cdots, X_n$ are mutually independently distributed as $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, \cdots, n$.

MGF of X_i is $M_i(\theta) \equiv \mathrm{E}(\exp(\theta X_i)) = \exp(\mu_i\theta + \frac{1}{2}\sigma_i^2\theta^2)$.

Consider the distribution of $Y = \sum_{i=1}^{n}(a_i + b_i X_i)$, where a_i and b_i are constant.

$$
\begin{aligned}
M_y(\theta) &\equiv \mathrm{E}(\exp(\theta Y)) = \mathrm{E}(\exp(\theta \sum_{i=1}^{n}(a_i + b_i X_i))) \\
&= \prod_{i=1}^{n} \exp(\theta a_i)\mathrm{E}(\exp(\theta b_i X_i)) = \prod_{i=1}^{n} \exp(\theta a_i)M_i(\theta b_i) \\
&= \prod_{i=1}^{n} \exp(\theta a_i) \exp(\mu_i\theta b_i + \frac{1}{2}\sigma_i^2(\theta b_i)^2) = \exp(\theta \sum_{i=1}^{n}(a_i+b_i\mu_i) + \frac{1}{2}\theta^2 \sum_{i=1}^{n} b_i^2\sigma_i^2),
\end{aligned}
$$

which implies that $Y \sim N(\sum_{i=1}^{n}(a_i + b_i\mu_i), \sum_{i=1}^{n} b_i^2\sigma_i^2)$.

[End of Review]

Substitute $a_i = 0$, $\mu_i = 0$, $b_i = \omega_i$ and $\sigma_i^2 = \sigma^2$.

Then, using the moment-generating function, $\sum_{i=1}^{n} \omega_i u_i$ is distributed as:

$$\sum_{i=1}^{n} \omega_i u_i \sim N(0, \ \sigma^2 \sum_{i=1}^{n} \omega_i^2).$$

Therefore, $\hat{\beta}_2$ is distributed as:

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

or equivalently,

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

for any n.

Moreover, replacing σ^2 by its estimator s^2 defined in (17), it is known that we have:

$$\frac{\hat{\beta}_2 - \beta_2}{s / \sqrt{\sum_{i=1}^n (x_i - \overline{x})^2}} \sim t(n - 2),$$

where $t(n - 2)$ denotes t distribution with $n - 2$ degrees of freedom.

Thus, under normality assumption on the error term u_i, the $t(n - 2)$ distribution is used for the confidence interval and the testing hypothesis in small sample.

Or, taking the square on both sides,

$$\Big(\frac{\hat{\beta}_2 - \beta_2}{s/\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}}\Big)^2 \sim F(1, n - 2),$$

which will be proved later.

Before going to **multiple regression model** (重回帰モデル),

2 Some Formulas of Matrix Algebra

1. Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lk} \end{pmatrix} = [a_{ij}],$

which is a $l \times k$ matrix, where a_{ij} denotes ith row and jth column of A.

The **transposed matrix** (転置行列) of A, denoted by A', is defined as:

$$A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{l1} \\ a_{12} & a_{22} & \cdots & a_{l2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{lk} \end{pmatrix} = [a_{ji}],$$

where the ith row of A' is the ith column of A.

2. $(Ax)' = x'A'$,

   where A and x are a $l \times k$ matrix and a $k \times 1$ vector, respectively.

3. $a' = a,$

   where a denotes a scalar.

4. $\dfrac{\partial a' x}{\partial x} = a,$

   where a and x are $k \times 1$ vectors.

5. $\dfrac{\partial x' A x}{\partial x} = (A + A')x,$

   where A and x are a $k \times k$ matrix and a $k \times 1$ vector, respectively.

Especially, when A is symmetric,

$$\frac{\partial x'Ax}{\partial x} = 2Ax.$$

6. Let A and B be $k \times k$ matrices, and I_k be a $k \times k$ **identity matrix** (単位行列) (one in the diagonal elements and zero in the other elements).

   When $AB = I_k$, B is called the **inverse matrix** (逆行列) of A, denoted by $B = A^{-1}$.

   That is, $AA^{-1} = A^{-1}A = I_k$.

7. Let A be a $k \times k$ matrix and x be a $k \times 1$ vector.

   If A is a **positive definite matrix** (正値定符号行列), for any x except for $x = 0$ we have:

   $$x'Ax > 0.$$

   If A is a **positive semidefinite matrix** (非負値定符号行列), for any x except for $x = 0$ we have:

   $$x'Ax \geq 0.$$

If *A* is a **negative definite matrix** (負値定符号行列), for any *x* except for $x = 0$ we have:

$$x'Ax < 0.$$

If *A* is a **negative semidefinite matrix** (非正値定符号行列), for any *x* except for $x = 0$ we have:

$$x'Ax \leq 0.$$

Trace, Rank and etc.: $\quad A : k \times k, \qquad B : n \times k, \qquad C : k \times n.$

1. The **trace (トレース)** of A is: $\text{tr}(A) = \sum_{i=1}^{k} a_{ii}$, where $A = [a_{ij}]$.

2. The **rank (ランク，階数)** of A is the maximum number of linearly independent column (or row) vectors of A, which is denoted by $\text{rank}(A)$.

3. If A is an **idempotent matrix (べき等行列)**, $A = A^2$.

4. If A is an idempotent and symmetric matrix, $A = A^2 = A'A$.

5. A is idempotent if and only if the eigen values of A consist of 1 and 0.

6. If A is idempotent, $\text{rank}(A) = \text{tr}(A)$.

7. $\text{tr}(BC) = \text{tr}(CB)$

Distributions in Matrix Form:

1. Let X, μ and Σ be $k \times 1$, $k \times 1$ and $k \times k$ matrices.

   When $X \sim N(\mu, \Sigma)$, the density function of X is given by:

   $$f(x) = \frac{1}{(2\pi)^{k/2}|\Sigma|} \exp\left(-\frac{1}{2}(x - \mu)'\Sigma^{-1}(x - \mu)\right).$$

   $\mathrm{E}(X) = \mu$ and $\mathrm{V}(X) = \mathrm{E}\left((X - \mu)(X - \mu)'\right) = \Sigma$

   The moment-generating function: $\phi(\theta) = \mathrm{E}\left(\exp(\theta'X)\right) = \exp(\theta'\mu + \frac{1}{2}\theta'\Sigma\theta)$

2. If $X \sim N(\mu, \Sigma)$, then $(X - \mu)'\Sigma^{-1}(X - \mu) \sim \chi^2(k)$.

   Note that $\quad X'X \sim \chi^2(k)$ when $X \sim N(0, I_k)$.

3. $X: n \times 1, \qquad Y: m \times 1, \qquad X \sim N(\mu_x, \Sigma_x), \qquad Y \sim N(\mu_y, \Sigma_y)$

   X is independent of Y, i.e., $E\big((X - \mu_x)(Y - \mu_y)'\big) = 0$ in the case of normal random variables.

   $$\frac{(X - \mu_x)'\Sigma_x^{-1}(X - \mu_x)/n}{(Y - \mu_y)'\Sigma_y^{-1}(Y - \mu_y)/m} \sim F(n, m)$$

4. If $X \sim N(0, \sigma^2 I_n)$ and A is a symmetric idempotent $n \times n$ matrix of rank G, then $X'AX/\sigma^2 \sim \chi^2(G)$.

Note that $X'AX = (AX)'(AX)$ and $\text{rank}(A) = \text{tr}(A)$ because A is idempotent.

5. If $X \sim N(0, \sigma^2 I_n)$, A and B are symmetric idempotent $n \times n$ matrices of rank G and K, and $AB = 0$, then

$$\frac{X'AX}{G\sigma^2} \Big/ \frac{X'BX}{K\sigma^2} = \frac{X'AX/G}{X'BX/K} \sim F(G, K).$$