**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] Content of Special Lectures in Economics (Statistical Analysis)**

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.

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

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 $n$.

**[Review 1]** $t$ **Distribution:**

$Z \sim N(0,1)$, $V \sim \chi^2(k)$, and $Z$ is independent of $V$. Then, $\dfrac{Z}{\sqrt{V/k}} \sim t(k)$.

**[End of Review 1]**

**[Review 2]** $t$ **Distribution:**

Suppose that $X_1, X_2. \cdots, X_n$ are mutually independently, identically and normally distributed with mean $\mu$ and variance $\sigma^2$.

$\overline{X} \sim N(\mu, \sigma^2/n)$, i.e., $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

Define $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$, which is an unbiased estimator of $\sigma^2$.

It is known that $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ and $\overline{X}$ is independent of $S^2$. (The proof is skipped.)

Then, we obtain $\dfrac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$.

As a result, replacing $\sigma^2$ by $S^2$, $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$.

**[End of Review 2]**

Back to OLS:

Replacing $\sigma^2$ 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,

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

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 $i$th row and $j$th 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 $i$th row of $A'$ is the $i$th 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. If $A$ is symmetric, $A = A'$.

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

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

   Especially, when $A$ is symmetric,

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

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

8. 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.:** $A : k \times k, \qquad B : n \times k, \qquad C : k \times n.$

1. The **trace (トレース)** of $A$ is: $\mathrm{tr}(A) = \sum_{i=1}^{k} a_{ii}$, 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 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, rank($A$) =tr($A$) .

7. tr($BC$) =tr($CB$)

**Distributions in Matrix Form:**

1. Let $X$, $\mu$ and $\Sigma$ 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:

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

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

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

(*) In the univariate case, when $X \sim N(\mu, \sigma^2)$, the density function of $X$ is:

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big).$$

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

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

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