

22 Sept

$$\Rightarrow \hat{\sigma}^2 = \frac{\text{Res S.S.}}{n-p'}$$

e.g. β_0 is known

$$\text{Res S.S.} = \sum_{i=1}^n (y_i - \beta_0)^2 - \hat{\beta}_1 \sum_{i=1}^n x_{i1} (y_i - \beta_0) - \dots - \hat{\beta}_p \sum_{i=1}^n x_{ip} (y_i - \beta_0)$$

$$\hat{\sigma}^2 = \frac{\text{Res S.S.}}{n-p}$$

Distribution of Res S.S.

$p=1$ dist. of $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2$ \Leftarrow joint dist. of y_i $i=1, \dots, n$

$$\text{let } \hat{\theta}_1 = \frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}}, \quad \hat{\theta}_2 = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}}, \quad \hat{\theta}_3 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\hat{\sigma}^2}$$

$$\text{m.g.f.} = m(t_1, t_2, t_3) = E(e^{t_1 \hat{\theta}_1 + t_2 \hat{\theta}_2 + t_3 \hat{\theta}_3})$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{t_1 \hat{\theta}_1 + t_2 \hat{\theta}_2 + t_3 \hat{\theta}_3} \times$$

$$\frac{1}{(2\pi)^{n/2}} \frac{1}{(\hat{\sigma}^2)^{n/2}} \exp \left\{ -\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1})^2 \right\} dy_1 \dots dy_n$$

= \dots

$$= \exp \left\{ \frac{1}{2} \left(t_1^2 \frac{\sum_{i=1}^n x_{i1}^2 / n}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2} + 2 t_1 t_2 \left(-\frac{\bar{x}_1}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2} \right) + t_2^2 \frac{1}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2} \right) - (n-2) t_3^2 \right\} (1 - 2 t_3)^{-(n-2)/2}$$

$$\Downarrow$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} \sim MN$$

$$\Downarrow$$

$$\frac{\text{Res S.S.}}{\hat{\sigma}^2} \sim \chi^2_{(n-2)}$$

\nwarrow independent \nearrow

For any p

$$\text{Model } \underline{Y} \sim MN(\underline{\mu}, \underline{\Sigma})$$

$$\underline{\mu} = \underline{X} \underline{\beta} \quad \underline{\Sigma} = \sigma^2 \underline{I}$$

\uparrow non-zero mean

Theorem 3.4 $\underline{Y} \sim MN(\underline{\mu}, \underline{\Sigma})$

$\underline{Y}^T \underline{A} \underline{Y}$ has a non-central chi-square with

k d.f. and the non-centrality constant

$$\lambda = \underline{\mu}^T \underline{A} \underline{\mu}$$

iff \underline{A} is a symmetric idempotent matrix of rank k

MATH 3423 $Z \sim N(0, 1)$ $Z^2 \sim \chi^2_{(1)}$

\uparrow

(central) chi-square with d.f. = 1

Theorem 3.3

$$E(\underline{Y}^T \underline{A} \underline{Y}) = \text{trace}(\underline{A} \underline{\Sigma}) + \underline{\mu}^T \underline{A} \underline{\mu}$$

\uparrow \underline{I}

$$= \text{trace}(\underline{A}) + \underline{\mu}^T \underline{A} \underline{\mu}$$

$$= k + \lambda$$

when $\lambda = 0$

\Rightarrow (central) chi-square

with d.f. = k

$$\text{Var}(\underline{Y}^T \underline{A} \underline{Y}) = 2(k + 2\lambda)$$

$$- \text{p.d.f.} = \sum_{i=0}^{\infty} \underbrace{e^{-(\lambda/2)} \frac{(\lambda/2)^i}{i!}}_{P_0(\lambda/2)} \underbrace{f_{Y_{k+2i}}(xy)}_{Y \sim \chi^2_{(k+2i)}}$$

- Poisson weighted mixture of (central) chi-squares v.v. \S

$$- J \sim P_0\left(\frac{\lambda}{2}\right)$$

conditional dist. of Y given $J =$ $\sim \chi^2_{(k+2j)}$
 (unconditional)
 \Rightarrow distribution of $Y \sim \chi^2_{(k)}(k, \lambda)$

$$Y \sim MN(\underline{X}\beta, \sigma^2 \underline{I}) \Rightarrow Y^* = \frac{Y}{\sigma} \sim MN\left(\frac{1}{\sigma} \underline{X}\beta, \underline{I}\right)$$

$$\text{Res. S.S.} = Y^T (\underline{I} - \underbrace{\underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T}_{\underline{H}}) Y \quad \mu^*$$

\underline{A} - symmetric ~~idempotent~~ idempotent

$$\frac{\text{Res S.S.}}{\sigma^2} = \frac{Y^T (\underline{I} - \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T) Y}{\sigma^2}$$

$\underbrace{\quad}_{Y^*} \quad \sigma^2 \quad \underbrace{\quad}_{Y^*}$

$$= Y^{*T} (\underline{I} - \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T) Y^*$$

By Theorem 3.4

$$\text{d.f.} = \text{trace}(\underline{I} - \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T)$$

$$= n - p'$$

$$\lambda = \mu^{*T} (\underline{I} - \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T) \mu^* \quad \mu^* = \frac{\underline{X}\beta}{\sigma}$$

$$= \frac{1}{\sigma^2} \beta^T \underline{X}^T (\underline{I} - \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T) \underline{X} \beta$$

$$= \cancel{\frac{1}{\sigma^2} \beta^T \underline{X}^T \underline{X} \beta}$$

$$= \frac{1}{\sigma^2} \beta^T (\underline{X}^T \underline{X} - \underline{X}^T \underline{X} (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{X}) \beta$$

$$= 0$$

$$\Rightarrow \frac{\text{Res S.S.}}{\sigma^2} \sim \chi^2_{(n-p')} \quad - \text{ (central) chi-square dist. with df.} = n-p'$$

③ Independence of $\hat{\beta}$ & Res S.S. = $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \hat{e}_i^2$

MATH 3423 $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$

sample variance S_{n-1}^2 & sample mean \bar{X} are indep.
 \uparrow \uparrow
 est. of σ^2 est. of μ

MATH 3424 $\frac{\text{Res S.S.}}{n-p'} \rightarrow \sigma^2$ $X \hat{\beta} \rightarrow \mu = X\beta$
 indep.

$p=1$ $\hat{e}_i = \sum_{j=1}^n (\delta_{ij} - (c_j + d_j X_{ci})) y_j$

$\delta_{ij} = 1$ if $i=j$
 $= 0$ otherwise

$y_i - \hat{y}_i$

$\hat{\beta}_0 = \sum_{j=1}^n \left(\frac{1}{n} - \frac{(X_{j1} - \bar{X}_1) \bar{X}_1}{S_{X_1 X_1}} \right) y_j$

$\hat{\beta}_1 = \sum_{j=1}^n \left(\frac{X_{j1} - \bar{X}_1}{S_{X_1 X_1}} \right) y_j$

By Theorem 3.1 $\text{Cov} \left(\sum_{i=1}^n c_i y_i, \sum_{i=1}^n d_i y_i \right) = \sigma^2 \left(\sum_{i=1}^n c_i d_i \right)$

$\text{Cov}(\hat{e}_i, \hat{\beta}_0) = \sigma^2 \sum_{j=1}^n (\delta_{ij} - c_j + d_j X_{ci}) c_j$

$= \sigma^2 \left\{ c_i - \sum_{j=1}^n (c_j + d_j X_{ci}) c_j \right\}$

$= \sigma^2 \left\{ c_i - \left(\sum_{j=1}^n c_j^2 \right) + \left(\sum_{j=1}^n c_j d_j X_{ci} \right) \right\}$

\uparrow from the proof of $\text{Var}(\hat{\beta}_0)$ \uparrow from the proof of $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1)$ (4)

$$= \sigma^2 \left\{ c_i - \left(\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1 x_1}} \right) + \frac{\bar{x}_1}{S_{x_1 x_1}} x_{i1} \right\}$$

$$\uparrow$$

$$c_i = \frac{1}{n} - \frac{(x_{i1} - \bar{x}_1) \bar{x}_1}{S_{x_1 x_1}}$$

$$= 0$$

$\Rightarrow \hat{e}_i, \hat{\beta}_0$ are indep.

$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = 0 \leftarrow \text{proof improve it!}$$

\leftarrow See p.18 &
p.19 of
chapter 1

$p > 1$ Theorem 3.2 (p.14)

$$\text{cov}(\underline{z}, \underline{z}) = \underline{z} \text{Var}(\underline{z}) \underline{z}^T$$

$$\text{cov}(\hat{\underline{e}}, \hat{\underline{\beta}}) = \underbrace{(\underline{I} - \underline{H})}_{\sim} \underbrace{\text{Var}(\underline{Y})}_{\sigma^2 \underline{I}} \underbrace{(\underline{X}^T \underline{X})^{-1} \underline{X}^T}_{\sim}$$

$$\underline{Y} - \underline{\hat{Y}} \quad \underline{(\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}}$$

$$= \sigma^2 (\underline{I} - \underline{H}) (\underline{X} (\underline{X}^T \underline{X})^{-1})$$

$$= \underline{Y} - \underline{X} \hat{\underline{\beta}}$$

$$= \sigma^2 (\underline{X} (\underline{X}^T \underline{X})^{-1} - \underbrace{\underline{X} \underline{(\underline{X}^T \underline{X})^{-1} \underline{X}^T}}_{\underline{H}} (\underline{X}^T \underline{X})^{-1})$$

$$= \underline{Y} - \underline{X} (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}$$

$$= \underline{0}$$

$$= (\underline{I} - \underline{X} (\underline{X}^T \underline{X})^{-1} \underline{X}^T) \underline{Y}$$

$$= \underbrace{(\underline{I} - \underline{H})}_{\sim} \underline{Y}$$

$$\text{cov}(\hat{\underline{e}}, \hat{\underline{\beta}})$$

$$\begin{matrix} \hat{\underline{e}} \\ \sim \\ n \times 1 \end{matrix} \quad \begin{matrix} \hat{\underline{\beta}} \\ \sim \\ p' \times 1 \end{matrix}$$

$$= \begin{pmatrix} \text{cov}(\hat{\beta}_1, \hat{\beta}_0) & \dots & \text{cov}(\hat{e}_1, \hat{\beta}_p) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_0) & \dots & \text{cov}(\hat{e}_2, \hat{\beta}_p) \\ \vdots & & \vdots \\ \text{cov}(\hat{e}_n, \hat{\beta}_0) & \dots & \text{cov}(\hat{e}_n, \hat{\beta}_p) \end{pmatrix}_{n \times p'} = \underline{0}$$

$\Rightarrow \hat{\underline{e}} \text{ d } \hat{\underline{\beta}}$ are indep.

$$E(y_i) = \underline{x}_i^T \underline{\beta} = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \leftarrow \text{diff. mean}$$

Section 4 Confidence interval & hypothesis testing of β

$$\hat{\beta} \sim MN(\beta, (X^T X)^{-1} \sigma^2)$$

$$\frac{\text{Res S.S.}}{n-p'} \sim \chi^2(n-p')$$

4.1 T-test

$$p=1 \quad \hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{S_{x_1 x_1}})$$

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\sigma / \sqrt{S_{x_1 x_1}}} \sim N(0, 1)$$

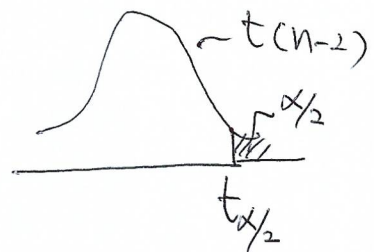
$$\Rightarrow \frac{\frac{\hat{\beta}_1 - \beta_1}{\sigma / \sqrt{S_{x_1 x_1}}} \sim N(0, 1)}{\sqrt{\frac{\text{Res S.S.}}{\sigma^2} / (n-p')} \sim \chi^2} \sim t(n-p')$$

indep

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma} / \sqrt{S_{x_1 x_1}}} \sim t(n-p') \leftarrow t(n-2)$$

$(1-\alpha)100\%$
C.I.

$$Pr(-t_{\alpha/2, n-2} \leq \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma} / \sqrt{S_{x_1 x_1}}} \leq t_{\alpha/2, n-2}) = 1-\alpha$$



$$\hat{\beta}_1 - t_{\alpha/2, n-2} \frac{\hat{\sigma}}{\sqrt{S_{x_1 x_1}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{S_{x_1 x_1}}}$$

random interval

$$\begin{aligned} \sum_{i=1}^n (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1})) &= 0 \\ \sum_{i=1}^n x_{i1} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1})) &= 0 \\ \Rightarrow \left. \begin{aligned} \sum_{i=1}^n \hat{e}_i &= 0 \\ \sum_{i=1}^n x_{i1} \hat{e}_i &= 0 \end{aligned} \right\} &\text{Properties of } \hat{e}_i \end{aligned}$$

Example 1: Example in Simple Linear Regression

i	1	2	3	4	5	6	7	8	9
x_i	1.5	1.8	2.4	3.0	3.5	3.9	4.4	4.8	5.0
y_i	4.8	5.7	7.0	8.3	10.9	12.4	13.1	13.6	15.3

Summary statistics:

$$\begin{aligned} \sum_{i=1}^9 x_i &= 30.3 & \sum_{i=1}^9 y_i &= 91.1 & \sum_{i=1}^9 x_i y_i &= 345.09 & S_{x_1 x_1} &= \sum_{i=1}^n x_{i1}^2 - n \bar{x}_1^2 = 13.1 \\ \sum_{i=1}^9 x_i^2 &= 115.11 & \bar{x} &= 3.3667 & \bar{y} &= 10.1222 & \text{Res S.S.} &= S_{yy} - \hat{\beta}_1 S_{x_1 y} \\ \Rightarrow \hat{\beta}_1 &= 2.9303 & \text{and } \hat{\beta}_0 &= 0.2568. & \text{Res S.S.} &= S_{yy} - \hat{\beta}_1^2 S_{x_1 x_1} & & \uparrow \\ & & & & & & & \frac{S_{x_1 y}}{S_{x_1 x_1}} = \hat{\beta}_1 \end{aligned}$$

Thus, the estimated regression line (or the fitted line) is given by

$$\hat{y} = 0.2568 + 2.9303x$$

For any p

$$\text{Min } \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\Rightarrow \text{Min } \sum_{i=1}^n (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_p x_{ip}))^2$$

$$= \sum_{i=1}^n \hat{e}_i^2$$

$$= (\hat{e}_1 \quad \dots \quad \hat{e}_n) \begin{pmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{pmatrix}$$

$$= (y_1 - (\hat{\beta}_0 + \hat{\beta}_1 x_{11} + \dots + \hat{\beta}_p x_{1p}) \quad \dots \quad y_n - (\hat{\beta}_0 + \hat{\beta}_1 x_{n1} + \dots + \hat{\beta}_p x_{np})) \begin{pmatrix} y_1 - (\hat{\beta}_0 + \hat{\beta}_1 x_{11} + \dots + \hat{\beta}_p x_{1p}) \\ \vdots \\ y_n - (\hat{\beta}_0 + \hat{\beta}_1 x_{n1} + \dots + \hat{\beta}_p x_{np}) \end{pmatrix}$$

$$= (Y^T - \hat{\beta}^T X^T)(Y - X\hat{\beta})$$

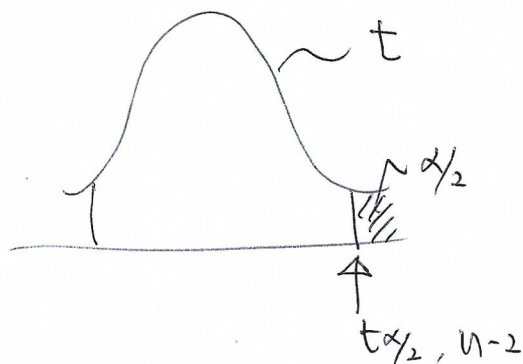
$$= (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

$$H_0 = \beta_1 = \beta_{10} \quad \text{e.g. } H_0 = \beta_1 = 2.5$$

$$\frac{\hat{\beta}_1 - \beta_{10}}{\hat{\sigma} / \sqrt{S_{xx1}}} \sim t(n-2) \quad \leftarrow \text{dist. under } H_0$$

$$\Rightarrow t_{\text{obs}} = \frac{\hat{\beta}_1 - \boxed{\beta_{10}}}{\hat{\sigma} / \sqrt{S_{xx1}}} =$$

\uparrow
 test stat.



$$H_1 = \beta \neq \beta_{10}$$

Reject H_0 if $|t_{\text{obs}}| > t_{\alpha/2, n-2}$

$$H_1 = \beta > \beta_{10} \quad \text{or} \quad H_1 = \beta < \beta_{10}$$

Reject H_0 if $|t_{\text{obs}}| > t_{\alpha, n-2}$

Example 1: Example in Simple Linear Regression (cont.)

Summary statistics:

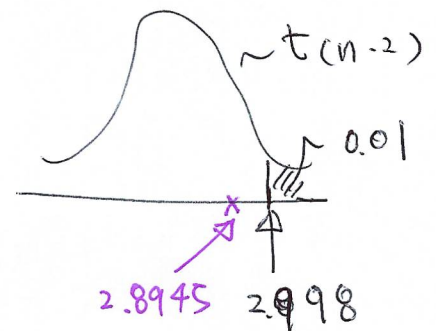
$$\begin{aligned}\sum_{i=1}^9 y_i^2 &= 1036.65 \\ S_{xx} &= 115.11 - \frac{(30.3)^2}{9} = 13.10 \\ S_{yy} &= 1036.65 - \frac{(91.1)^2}{9} = 114.52 \\ S_{xy} &= 345.09 - \frac{(30.3)(91.9)}{9} = 38.39 \\ \hat{\beta}_1 &= 2.9303 \\ \hat{\sigma}^2 &= \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{n-2} \\ &= \frac{114.52 - (2.9303)(38.39)}{7} = 0.2894 \\ \Rightarrow \hat{\sigma} &= 0.538 \\ t_{0.05/2, 7} &= 2.365\end{aligned}$$

$$\begin{aligned}H_0 &: \beta_1 = 2.5 \\ \text{vs } H_1 &: \beta_1 > 2.5 \\ t &= \frac{\hat{\beta}_1 - 2.5}{0.538/\sqrt{13.1}}\end{aligned}$$

= 2.8945
↑
value of test stat.

critical value

$$= t_{0.01, 7} = 2.998$$



$$H_0 : \beta_1 = 2.5,$$

$$H_1 : \beta_1 > 2.5$$

$$\begin{aligned}t &= \frac{2.9303 - 2.5}{0.538/\sqrt{13.10}} \\ &= 2.8945 \\ &< 2.998 = t_{0.01, 7}\end{aligned}$$

Can't reject $H_0 \Rightarrow \beta_1$ does not significantly differ from 2.5

OR There is no enough evidence to support that β_1 is greater than 2.5

95% C.I. of β_1 :

$$\begin{aligned}&(2.9305 - \frac{2.365 * 0.538}{\sqrt{13.10}}, 2.9305 + \frac{2.365 * 0.538}{\sqrt{13.10}}) \\ &\Rightarrow 2.579 < \beta_1 < 3.282\end{aligned}$$

$$H_0 : \beta_0 = 0$$

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma} \sqrt{\frac{1}{n} + \frac{\bar{x}_1^2}{S_{x_1 x_1}}}} = \frac{0.2568 - 0}{0.538 \sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}} = 0.4831 < 1 \text{ (can't reject } H_0)$$

95% C.I. of β_0 :

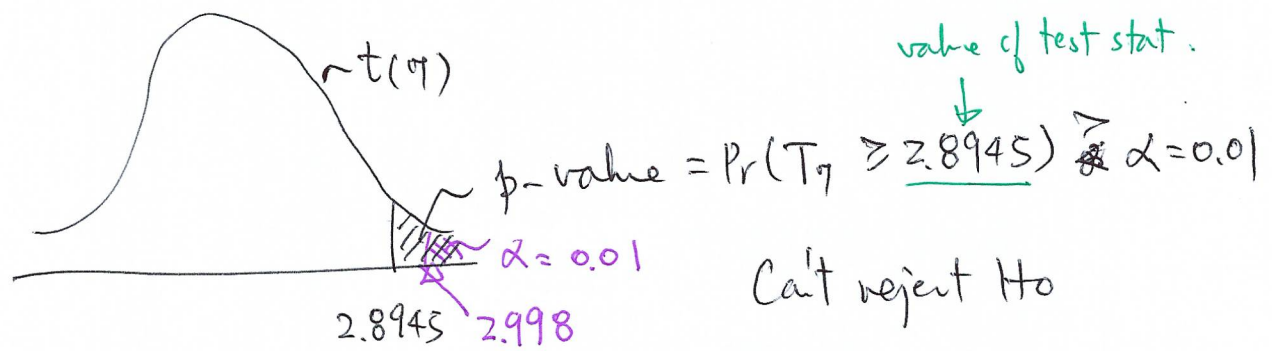
$$\begin{aligned}&(0.2568 - (2.365)(0.538) \sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}, 0.2568 + (2.365)(0.538) \sqrt{\frac{1}{9} + \frac{3.3667^2}{13.10}}) \\ &\Rightarrow -1.0005 < \beta_0 < 1.514\end{aligned}$$

Example 4: Intercept is known (cont.)

$$H_0 : \beta_1 = \beta_{10}$$

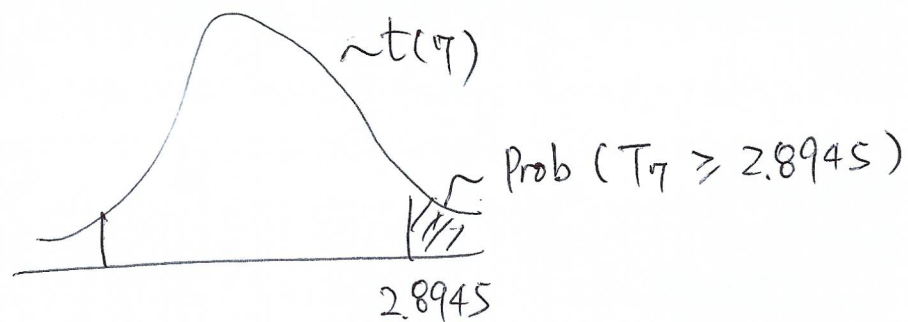
$$\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{\sum_{i=1}^n x_i^2})$$

① p-value



if $p\text{-value} < \alpha$, then reject H_0 .

② $H_1: \beta_1 \neq 2.5$



$$p\text{-value} = \Pr(T_7 \geq 2.8945) \times 2$$

③ p-value in computer output is normally for two-sided alternative

For one-sided alternative, $p\text{-value} = (\text{p-value from computer-output}) / 2$

$$H_0: \beta_1 \geq 2.5$$

④ $H_1: \beta_1 < 2.5$ ← No.

← meaningless

$\hat{\beta}_1 = 2.9303 \Rightarrow \text{data}_{\text{set}} \text{ supports that } \beta_1 > 2.5$

\Rightarrow Can't reject H_0