14.5 : Testing for significance

→ Model : $y = \beta_0 + \beta_1 X + \varepsilon$.

→

$H_0 : \beta_1 = 0$ means • X is not a significant variable
• $y = \beta_0 + \beta_1 X + \varepsilon$ is not a significant Model .

$H_1 : \beta_1 \neq 0$ means • X is a significant variable .
• $y = \beta_0 + \beta_1 X + \varepsilon$ is a significant variable .

Assuming : 1. $E(\varepsilon) = 0$

2. $Var(\varepsilon) = \sigma^2$

3. ε indep

4. ε Normal .

Two-tail test

→ test - statistic :

• t - test : $t = \dfrac{b_1}{s_{b_1}}$ where $s_{b_1} = \sqrt{\dfrac{MSE}{(n-1)\, S_x^2}}$

with $df = n - 2$.

→ Rejection Rule :

Reject H_0 if $|t| \geq t_{\frac{\alpha}{2}}$.

→ Mean square Error (estimate of σ^2).

$$S^2 = MSE = \frac{SSE}{n-2}.$$

→ Standard Error of the estimate:

$$S = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}}.$$

→ Model:

$$Y = \beta_0 + \beta_1 X + \varepsilon.$$

→ sampling distribution of b_1:

- $E(b_1) = \beta_1$

- $\sigma_{b_1} = \dfrac{\sigma}{\sqrt{\Sigma(X_i - \bar{X})^2}} = \dfrac{\sigma}{\sqrt{(n-1)\,S_x^2}}$ standard deviation.

- Distribution b_1 is Normal.

→ Estimated standard distribution of b_1:

$$S_{b_1} = \frac{S}{\sqrt{(n-1)\,S_x^2}}.$$

→ $(1-\alpha)$ CI for b_1:

$$b_1 \mp t_{\frac{\alpha}{2}}\, S_{b_1}.$$

→ F-test :

ANOVA Table.

Source of Variation	df	SS	MS	F
Regression	1	SSR	MSR	$F = \frac{MSR}{MSE}$.
Error	n-2	SSE	MSE	
Total	n-1	SST	–	

- Mean square Regression (MSR) = SSR
- Mean square Error (MSE) = SSE\n-2 .
- Test statistic F = MSR\MSE.
- Reject H_0 if • $F \geq F_\alpha$ with $df_1 = 1$ and $df_2 = n-2$.
  - p-value $\leq \alpha$

Recall the example in section 14.3 :

$\hat{y} = 0.2 + 2.6\,X$, $SSR = 67.6$, $SSE = 12.4$, $SST = 80$, $n = 5$.

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$.

1. use t-test :

$$t = \frac{b_1}{S_{b_1}} = \frac{2.6}{0.64} = 4.06$$

But $$S_{b_1} = \frac{S}{\sqrt{(n-1)S_x^2}} = \frac{2.03}{\sqrt{4(2.5)}} = \frac{2.03}{\sqrt{10}} = 0.64$$

But $$S = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{12.4}{5-2}} = 2.03 .$$

Conti. $\alpha = 0.05$

→ Reject ??

$\alpha = 0.05 \rightsquigarrow \frac{\alpha}{2} = 0.025$

$df = n-2 = 3$

t dup $t_{\frac{\alpha}{2}}$ with $df = 3$ → $t_{0.025}$ with $df = 3$

$\boxed{t_{0.025} = 3.182}$

$|t| = 4.06 \geq 3.182 = t_{\frac{\alpha}{2}}$ so we Reject H_0 $(\alpha = 0.05)$

so $\beta_1 \neq 0$ $(\alpha = 0.05)$.

→ So The Model $y = \beta_0 + \beta_1 X + \varepsilon$ is significant. $(\alpha = 0.05)$

→ The variable X is significant $(\alpha = 0.05)$

question : Try to make The same conclusion using a 95% CI for β_1 :

$0.95\ CI = b_1 \mp t_{\frac{\alpha}{2}} S_{b_1}$

$= 2.6 \mp (3.182)(0.64)$

$= 2.6 \mp 2.04$

$= [0.56, 4.64]$

② By F-test :

Source of Variation	df	SS	MS	F
Regression	1	67.6	67.6	$\frac{67.6}{4.13} = 16.37$.
Error	3	12.4	4.13	
Total	4	80	–	

→ Find F_α → $F_{0.05}$ with $df_1 = 1$, $df_2 = 3$

$F_{0.05} = 10.13$.

$\Longrightarrow$ $F > F_\alpha$ so we Reject H_0 $(\alpha = 0.05)$

$\beta_1 \neq 0$ $(\alpha = 0.05)$.

→ The variable X is significant variable $(\alpha = 0.05)$

→ The Model $Y = \beta_0 + \beta_1 X + \varepsilon$ is significant $(\alpha = 0.05)$.