4 Testing Hypotheses about a Single Linear Combination of the Parameter

Consider

$$\log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 \exp er + u$$

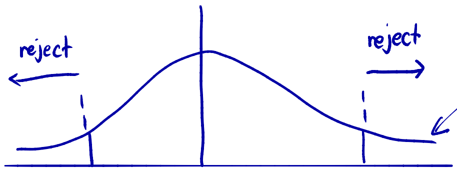
where jc = number of years attending a two-year college
$univ$ = number of years at a four-year college
$\exp er$ = months in the workforce.
We want to test whether $\beta_1 = \beta_2$.

if the returns from 1 more year of education at a junior college is the same as that of univ.

$H_0 : \beta_1 = \beta_2 \rightarrow H_0 : \beta_1 - \beta_2 = 0$

$H_a : \beta_1 \neq \beta_2 \quad H_a : \beta_1 - \beta_2 \neq 0$

2 tailed test



area = $\frac{\text{significant level}}{2}$

compute this t-statistic and compare it with the critical value

$$t = \frac{(\hat{\beta}_1 - \hat{\beta}_2) - 0}{se(\hat{\beta}_1 - \hat{\beta}_2)}$$

where $se(\hat{\beta}_1 - \hat{\beta}_2) = \sqrt{var(\hat{\beta}_1 - \hat{\beta}_2)}$

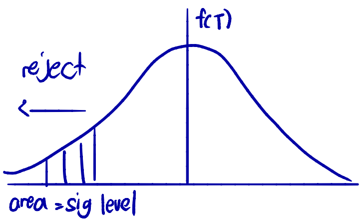
$$= \sqrt{\widehat{var}(\hat{\beta}_1) + \widehat{var}(\hat{\beta}_2) - 2\widehat{cov}(\hat{\beta}_1, \hat{\beta}_2)}$$

another possible hypothesis test (one-tailed alternative)

$H_0 : \beta_1 = \beta_2$ $\qquad$ $H_0 : \beta_1 - \beta_2 = 0$

$H_a : \beta_1 < \beta_2$ $\qquad$ $H_a : \beta_1 - \beta_2 < 0$

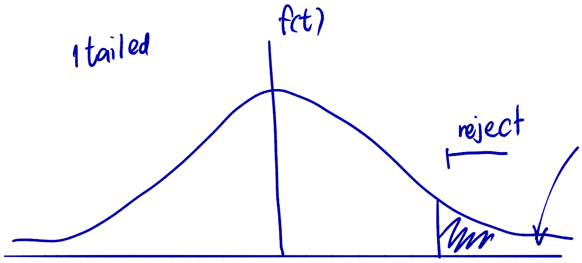
- Assume that β_1 would not be more than β_2 (returns to a 2-year college would never be more than returns to university education



$$t = \frac{(\hat{\beta}_1 - \hat{\beta}_2) - 0}{s.e.\ (\hat{\beta}_1 - \hat{\beta}_2)}$$

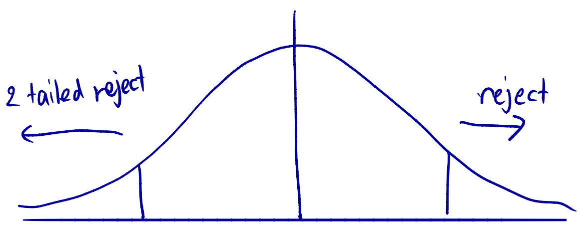
5 Computing p-Values for t-Tests

- What is the significance level given the computed t-statistics?



This shaded area in the rejection is the p-value

$$t = \frac{\hat{\beta}_j - \beta_j}{s.e.\ \beta_j}$$



$$t = \frac{\hat{\beta}_j - \beta_j}{s.e.\ \hat{\beta}_j}$$

- p-value : $P(|T| > |t|)$

T = t- distributed random variable with d.f. $= n-k-1$

t = Computed t - statistic

p- value = probability that a random T value will be greater (in the absolute term than our t - statistic in the H_0 testing

Consider the multiple regression model, assume MLR 1-6 are satisfied

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u.$$

you would like to test $H_0 : \beta_1 - 3\beta_2 = 1$ H_a: otherwise is true

1) Write the t-statistic for testing H_0

$$t = \frac{(\hat{\beta}_1 - 3\hat{\beta}_2) - 1}{s.e.(\hat{\beta}_1 - 3\hat{\beta}_2)}$$

2) Define $\theta_1 = \hat{\beta}_1 - 3\hat{\beta}_2$ $\rightarrow$ $H_0 : \theta = 1$
$H_a : \theta \neq 1$

$t = \frac{\hat{\theta} - 1}{s.e.(\hat{\theta}_1)}$ $\rightarrow$ we need our regression to have θ in it. So stata or OLS estimation will automatically give $\hat{\theta}_1$ & s.e. $\hat{\theta}_1$

Now $\hat{\beta}_1 = \theta_1 + 3\hat{\beta}_2$
or $\beta_1 = \theta_1 + 3\beta_2$

substitute in the main regression and get

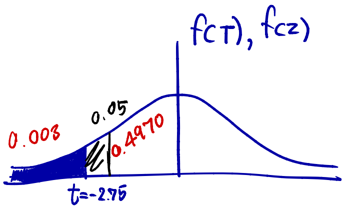
$$\begin{aligned} y &= \beta_0 + (\theta_1 + 3\beta_2) X_1 + \beta_2 X_2 + \beta_3 X_3 + u \\ &= \beta_0 + \theta_1 X_1 + 3\beta_2 X_1 + \beta_2 X_2 + \beta_3 X_3 + u \\ &= \beta_0 + \theta_1 X_1 + \beta_2 (X_2 + 3X_1) + \beta_3 X_3 + u \end{aligned}$$

* Now the explainatory variables are going to be X_1, $X_2 + 3X_1$ and X_3

- We can calculate $t = \frac{\hat{\theta}_1 - 1}{s.e.\ \hat{\theta}_1}$

Example 1: $H_0 : \beta_j \geq 0, \quad H_a : \beta_j < 0, \quad$ d.f.= 140.

P-value = what should be the significant level given the critical value of -2.75 → find the shaded area



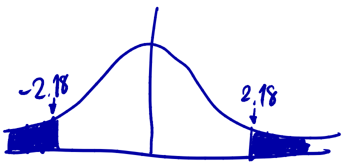
$t_{\hat{\beta}_j} = \frac{(-\hat{\beta}_j - \beta_j)}{se.\ \hat{\beta}_j}$

suppose the calculated $t_{\hat{\beta}_j} = -2.75$

- From the z-table, the value -2.75 corresponds to area = 0.003.
- Thus, p-value = 0.003.
- Would we reject H_0 if we use the significance level = 5%? Yes

we reject H_0 if p value < sig level

Example 2: $H_0 : \beta_j = a_j, \quad H_a : \beta_j \neq a_j, \quad$ d.f.= 18. ← d.f. < 30 use t-table



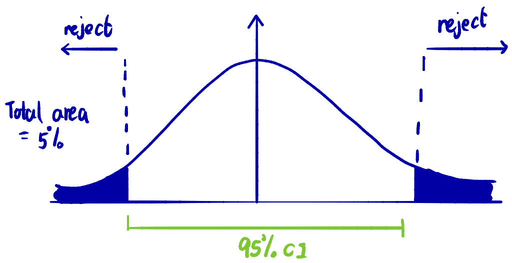
suppose the calculated $t_{\hat{\beta}_j} = -2.18$

- From the t-table, the value -2.18 corresponds to area = 0.02 to 0.05.
- Thus, p-value = between 0.02 to 0.05.
- Would we reject H_0 if we use the significance level = 5%?

Yes, reject H_0 because the area is less than 0.05

6 Confidence Intervals (CI)

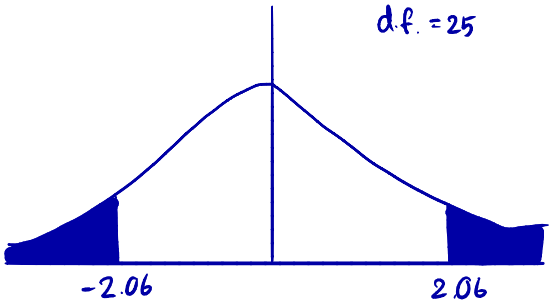
- Confidence Intervales for the POPULATION PARAMETER (β_j) → The range of values that would capture the true β_j at 95% chance
- A 95% CI of β_j is given by



$C_I \longrightarrow \hat{\beta}_j \pm c \times se.(\hat{\beta}_j)$

c is the 97.5 percentile in the t-distribution with n-k-1 d.f.

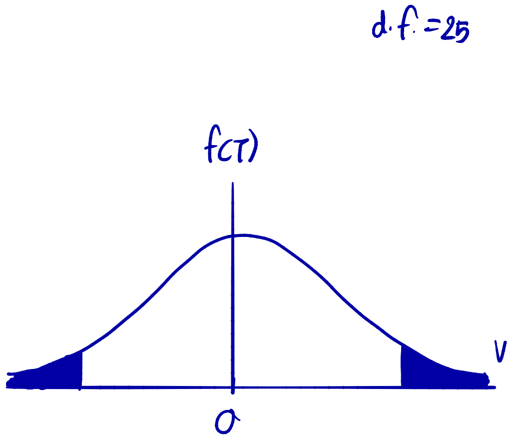
Example 1: **95% CI**



value of the 97.5^{th} percentile in the t_{25} distribution

The 95% C.I. for $\hat{\beta}_j$ = $[\hat{\beta}_j - 2.06 \cdot s.e.(\hat{\beta}_j), \hat{\beta}_j + 2.06 \cdot s.e.(\hat{\beta}_j))$

Example 2: **99% CI**



value of the 99.5% percentile in the t-distribution

The 99% CI for $\hat{\beta}_j$ = $[\hat{\beta}_j - 2.787 \cdot s.e. \hat{\beta}_j, \hat{\beta}_j + 2.787 \cdot s.e. (\hat{\beta}_j))$

F-test

we want to test the significance of a group of hypotheses (multiple hypotheses)

$$Grade_{325} = \beta_0 + \beta_1 \text{ \#times\_front} + \beta_2 \text{ \#times\_back} + \beta_3 \text{ hr\_study} + \beta_4 \text{ past\_GPA} + \beta_5 \text{ gender} + u$$

H_0 : seat position doesn't have impact on GPA

$\beta_1 = 0$ and $\beta_2 = 0$ $\rightarrow$ $\beta_1 = \beta_2 = 0$

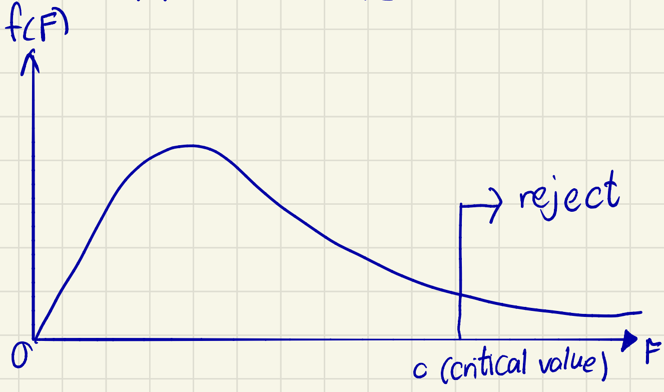
H_a : seat position matters

$\beta_1 \neq 0$ and $\beta_2 \neq 0$

or $\beta_1 \neq 0$ and $\beta_2 = 0$

or $\beta_1 = 0$ and $\beta_2 \neq 0$

at least one of the $\beta_1, \beta_2 \neq 0$



H_0 : $\beta_2 = \beta_3 = 0$

H_a : H_0 not true

$F \sim F_{q, n-k-1}$

# of hypotheses being tested

d.f. of ur model

7 Testing Multiple Linear Restrictions: The F-test

Suppose the model is specified by

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u \\ H_0 &: \beta_2 = 0 \text{ and } \beta_3 = 0 \\ H_1 &: H_0 \text{ is not true} \end{aligned}$$

→ want to test if x_1 and x_2 both have no impact on y

We can use the F-test to test this type of "multiple hypotheses".

1. Our full model is called the "unrestricted" model (ur). Suppose it can be expressed as:
big model (contain all variables)

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$ is true → reject H_0

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_k x_k + u$$

2. The model which takes out x (which we think its associated $\beta = 0$) is called the restricted model (r). ← small model

$y = \beta_0 + \beta_1 x_1 + u$ is true → do not reject H_0

- Suppose there are "q" numbers of β that we would like to perform a joint-test of $= 0$

→ e.g. in this model, $q = 2$

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_{k-q}$

$H_0 : \beta_{k-q+1} = \beta_{k-q+2} = \beta_k = 0$

(the last q $\beta_s = 0$

H_a : H_0 is not true

SSR_{ur} smaller than SSR_r

everytime you add 1 more x, the model will be better explained

*

$$F = \frac{\dfrac{(SSR_r - SSR_{ur})}{q}}{\dfrac{SSR_{ur}}{n-k-1}}$$

q ← number of explainatory variables we want to test

$n-k-1$ ← d.f. of "ur" model

So if everytime you add 1 more X variable, SSR ↓ and R^2 ↑, why don't we keep additional X in the model

→ everytime we add 1 more x, $var(\hat{\beta}_s)$ ↑ making the prediction of β less precise.

→ so we want to keep the additional X_3 if it can improve the model enough, SSR ↓ R^2 ↑ (significant)