STAT763: Applied Regression Analysis

Multiple linear regression

4.4 Hypothesis testing

Chunsheng Ma

E-mail: cma@math.wichita.edu

4.4.1 Significance of regression

Null hypothesis (Test whether all $\beta_j = 0$, $j = 1, \ldots, k$)

$$\mathcal{H}_0: \quad \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

4.4.1 Significance of regression

Null hypothesis (Test whether all $\beta_j = 0$, $j = 1, \ldots, k$)

$$\mathcal{H}_0: \quad \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

Alternative hypothesis

$$\mathcal{H}_a: \quad \text{Not all } \beta\text{s are equal to zero}$$

4.4.1 Significance of regression

Null hypothesis (Test whether all $\beta_j = 0$, $j = 1, \ldots, k$)

$$\mathcal{H}_0: \quad \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

Alternative hypothesis

$\mathcal{H}_a$: Not all βs are equal to zero

Test statistic

$$F = \frac{MS_R}{MS_{Res}}$$

4.4.1 Significance of regression

Null hypothesis (Test whether all $\beta_j = 0$, $j = 1, \ldots, k$)

$$\mathcal{H}_0: \quad \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

Alternative hypothesis

$\mathcal{H}_a$: Not all βs are equal to zero

Test statistic

$$F = \frac{MS_R}{MS_{Res}}$$

ANOVA table

Source	SS	df	MS	Test statistic
Regression	SS_R	k	MS_R	F
Residuals	SS_{Res}	n-k-1	MS_{Res}	
Total	SS_T	n-1		

Delivery time data (Example 3.1)

```
y = c(16.68, 11.50, 12.03, 14.88, 13.75, 18.11, 8.00, 17.83,
         79.24, 21.50, 40.33, 21.00, 13.50, 19.75, 24.00,29.00,
         15.35, 19.00, 9.50, 35.10, 17.90, 52.32, 18.75, 19.83,
         10.75)

 x1 = c(7, 3, 3, 4, 6, 7, 2, 7, 30, 5, 16, 10, 4, 6, 9, 10,
        6, 7, 3, 17, 10, 26, 9, 8, 4)

 x2 = c(560, 220, 340, 80, 150, 330, 110, 210, 1460, 605, 688,
        215, 255, 462, 448, 776, 200, 132, 36, 770, 140, 810,
        450, 635, 150)

 myfit = lm ( y ~ x1+x2)
```

ANOVA table

```
> anova (myfit)

Analysis of Variance Table

Response: y
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x1	1	5382.4	5382.4	506.619	< 2.2e-16
x2	1	168.4	168.4	15.851	0.0006312
Residuals	22	233.7	10.6		

ANOVA table

```
> anova (myfit)

Analysis of Variance Table

Response: y

            Df   Sum Sq   Mean Sq   F value    Pr(>F)
x1          1    5382.4   5382.4    506.619    < 2.2e-16
x2          1    168.4    168.4     15.851     0.0006312
Residuals   22   233.7     10.6
```

From the above table, we obtain $SS_{Res} = 233.7$, or

```
> anova(myfit)[3,2]

[1] 233.7317
```

SS_R

$SS_R = 5382.4 + 168.4 = 5550.8$, or

```
> SS_R = anova(myfit)[1,2]+anova(myfit)[2,2]
  [1] 5550.811
```

SS_R

$SS_R = 5382.4 + 168.4 = 5550.8$, or

```
> SS_R = anova(myfit)[1,2]+anova(myfit)[2,2]
  [1] 5550.811
```

The observed F value

$$F = \frac{SS_R/2}{MS_{Res}} = \frac{5550.8/2}{10.6} = 261.8302$$

SS_R

$SS_R = 5382.4 + 168.4 = 5550.8$, or

```
> SS_R = anova(myfit)[1,2]+anova(myfit)[2,2]
[1] 5550.811
```

The observed F value

$$F = \frac{SS_R/2}{MS_{Res}} = \frac{5550.8/2}{10.6} = 261.8302$$

```
MS_R = SS_R/(anova(myfit)[1,1]+anova(myfit)[2,1])

F = MS_R/anova(myfit)[3,3]
```

SS_R

$SS_R = 5382.4 + 168.4 = 5550.8$, or

```
> SS_R = anova(myfit)[1,2]+anova(myfit)[2,2]
  [1] 5550.811
```

The observed F value

$$F = \frac{SS_R/2}{MS_{Res}} = \frac{5550.8/2}{10.6} = 261.8302$$

```
MS_R = SS_R/(anova(myfit)[1,1]+anova(myfit)[2,1])

F = MS_R/anova(myfit)[3,3]
```

p-value (tail probability) $P(F_{2,22} > \textit{observed F value})$

```
> 1-pf(F,2,22)
[1] 4.440892e-16
```

Two ways to make decision

(1) Report $p - value$, which is evaluated from the data.

A small p-value leads to reject H_0.

(2) Use the significance level α, which is previously given.

Take $\alpha = 0.05$, say.

From the F table, we find the threshold, $qf(1 - 0.05, 2, 22)$

```
> qf(1-0.05,2,22)
[1] 3.443357
```

Then, compare the observed F value with this threshold.

Finally, make the decision.

◦ **Derive the F test via the** *likelihood-ratio-test method*

$$\mathcal{H}_0: \; \beta_1 = \beta_2 = \ldots = \beta_k = 0 \quad \longleftrightarrow \quad \mathcal{H}_a: \; \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

◦ **Derive the F test via the** *likelihood-ratio-test method*

$$\mathcal{H}_0: \ \beta_1 = \beta_2 = \ldots = \beta_k = 0 \quad \longleftrightarrow \quad \mathcal{H}_a: \ \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

The parameter space associated with $\mathcal{H}_0$ is

$$\Theta_0 = \{(\beta_0, 0, \ldots, 0, \sigma^2) : \ \beta_0 \in \mathbb{R}, \sigma^2 \in (0, \infty)\}$$

◦ **Derive the F test via the** *likelihood-ratio-test method*

$$\mathcal{H}_0: \ \beta_1 = \beta_2 = \ldots = \beta_k = 0 \quad \longleftrightarrow \quad \mathcal{H}_a: \ \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

The parameter space associated with $\mathcal{H}_0$ is

$$\Theta_0 = \{(\beta_0, 0, \ldots, 0, \sigma^2) : \ \beta_0 \in \mathbb{R}, \sigma^2 \in (0, \infty)\}$$

Likelihood function of parameter $(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)$

$$\ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right\}$$

Likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}$$

Likelihood ratio

$$\Lambda = \frac{\sup\limits_{\Theta_0} \ell(\beta_0, \beta_1, \dots, \beta_k, \sigma^2)}{\sup\limits_{\Theta} \ell(\beta_0, \beta_1, \dots, \beta_k, \sigma^2)}$$

$$\begin{aligned}
\sup_{\Theta_0} \ell(\beta_0, \beta_1, \dots, \beta_k, \sigma^2) &= \ell\left(\bar{y}, 0, \ \dots, 0, \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2\right) \\
&= \ell\left(\bar{y}, 0, \ \dots, 0, \frac{1}{n} SS_T\right)
\end{aligned}$$

Likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}$$

$$\begin{aligned}
\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) &= \ell\left(\bar{y}, 0, \ \ldots, 0, \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2\right) \\
&= \ell\left(\bar{y}, 0, \ \ldots, 0, \frac{1}{n} SS_T\right)
\end{aligned}$$

$$\begin{aligned}
\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) &= \ell\left(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k, \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right) \\
&= \ell\left(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k, \frac{1}{n} SS_{Res}\right)
\end{aligned}$$

The likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)} = \left(1 + \frac{k}{n-k-1} F\right)^{-\frac{n}{2}},$$

which is a decreasing function of F.

The likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)} = \left(1 + \frac{k}{n-k-1} F\right)^{-\frac{n}{2}},$$

which is a decreasing function of F.

Thus, a larger F would lead us to conclude $\mathcal{H}_a$.

4.4.2 Test whether a single $\beta_j = 0$

$$\mathcal{H}_0 : \beta_j = 0 \qquad \longleftrightarrow \qquad \mathcal{H}_a : \beta_j \neq 0$$

Two ways for this hypothesis testing question

(1). Student's t test

$$T = \frac{\hat{\beta}_j}{s\{\hat{\beta}_j\}}$$

The corresponding value can be obtained in **R** by summary(lm(y x1+...)), where $s\{\hat{\beta}_j\}$ is the standard error of $\hat{\beta}_j$.

Rational

A Student's t variable is formulated based on the following three facts

$$\hat{\beta}_j \sim N\left(\beta_j, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}\right), \quad \text{or} \quad \frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}} \sim N(0,1)$$

Rational

A Student's t variable is formulated based on the following three facts

$\hat{\beta}_j \sim N\left(\beta_j, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}\right),$ or $\frac{\hat{\beta}_j-\beta_j}{\sqrt{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}} \sim N(0,1)$

$\frac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k-1},$

Rational

A Student's t variable is formulated based on the following three facts

$\hat{\beta}_j \sim N\left(\beta_j, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}\right)$, or $\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}} \sim N(0,1)$

$\frac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k-1}$,

$\hat{\beta}_j$ and SS_{Res} are independent each other.

Rational

A Student's t variable is formulated based on the following three facts

$\hat{\beta}_j \sim N\left(\beta_j, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}\right)$, or $\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}} \sim N(0,1)$

$\frac{SS_{Res}}{\sigma^2} \sim \chi^2_{n-k-1}$,

$\hat{\beta}_j$ and SS_{Res} are independent each other.

Student's t variable

$$\frac{\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}}}{\sqrt{\frac{SS_{Res}}{\sigma^2}/(n-k-1)}} \sim t_{n-k-1}$$

or

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{MS_{Res}\ (\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}} = \frac{\hat{\beta}_j - \beta_j}{s(\hat{\beta}_j)} \sim t_{n-k-1}$$

When $\mathcal{H}_0 : \beta_j = 0$ is true,

$$\frac{\hat{\beta}_j}{s(\hat{\beta}_j)} \sim t_{n-k-1},$$

where

$$s(\hat{\beta}_j) = \sqrt{MS_{Res}\ (\mathbf{X}'\mathbf{X})^{-1}_{j+1,j+1}}$$

Test statistic for $\mathcal{H}_0 : \beta_j = 0$

$$T = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}$$

Two ways could be used to make a decision: p-value or significance level.

(2). Partial F test

F test statistic

$$
\begin{aligned}
F &= \frac{SS_R(X_j|X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}{1} \div \frac{SS_{Res}(X_1,\ldots,X_k)}{n-k-1} \\
&= \frac{SS_R(X_j|X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}{MS_{Res}}
\end{aligned}
$$

$$
\begin{aligned}
& SS_R(X_j|X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k) \\
= \; & SS_R(\text{use all predictor variables in the model}) \\
& -SS_R(\text{use predictor variables } X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)
\end{aligned}
$$

◦ Sum of squares

SS_T: *WITHOUT taking any predictor variables into account.*

$$SS_T = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

◦ Sum of squares

SS_T: WITHOUT taking any predictor variables into account.

$$SS_T = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

(i) Taking X_1 into account, Y is regressed on X_1 alone (simple linear regression).

$$\begin{aligned} SS_R(X_1) &= \text{the regression sum of squares when } X_1 \text{ only is in the model} \\ SS_{Res}(X_1) &= \text{the error sum of squares when } X_1 \text{ only is in the model} \end{aligned}$$

$$SS_T = SS_R(X_1) + SS_{Res}(X_1).$$

◦ Sum of squares

SS_T: WITHOUT taking any predictor variables into account.

$$SS_T = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

(i) Taking X_1 into account, Y is regressed on X_1 alone (simple linear regression).

$$\begin{aligned} SS_R(X_1) &= \text{the regression sum of squares when } X_1 \text{ only is in the model} \\ SS_{Res}(X_1) &= \text{the error sum of squares when } X_1 \text{ only is in the model} \end{aligned}$$

$$SS_T = SS_R(X_1) + SS_{Res}(X_1).$$

(ii) Taking both X_1 and X_2 into account, Y is now regressed on X_1 and X_2 (two predictor variables).

$$SS_R(X_1, X_2) = \text{ the regression sum of squares when } X_1 \text{ and } X_2 \text{ are in the model}$$

$$SS_{Res}(X_1, X_2) = \text{ the error sum of squares when } X_1 \text{ and } X_2 \text{ are in the model}$$

$$SS_T = SS_R(X_1, X_2) + SS_{Res}(X_1, X_2).$$

Questions

Which one is larger, $SS_{Res}(X_1)$ or $SS_{Res}(X_1, X_2)$?
Which one is larger, $SS_R(X_1)$ or $SS_R(X_1, X_2)$?

Questions

Which one is larger, $SS_{Res}(X_1)$ or $SS_{Res}(X_1, X_2)$?
Which one is larger, $SS_R(X_1)$ or $SS_R(X_1, X_2)$?

Answers

$$SS_{Res}(X_1) \geq SS_{Res}(X_1, X_2)$$

$$SS_R(X_1) \leq SS_R(X_1, X_2)$$

Questions

Which one is larger, $SS_{Res}(X_1)$ or $SS_{Res}(X_1, X_2)$?
Which one is larger, $SS_R(X_1)$ or $SS_R(X_1, X_2)$?

Answers

$$SS_{Res}(X_1) \geq SS_{Res}(X_1, X_2)$$

$$SS_R(X_1) \leq SS_R(X_1, X_2)$$

Extra sum of squares

$$SS_{Res}(X_2|X_1) = SS_{Res}(X_1) - SS_{Res}(X_1, X_2)$$

It measures the marginal effect of adding X_2 to the regression model, when X_1 is already in the model (and X_2 is regarded as an extra variable).

$$SS_R(X_2|X_1) = SS_R(X_1, X_2) - SS_R(X_1)$$

it is the marginal increase in the regression sum of squares.

Questions

Which one is larger, $SS_{Res}(X_1)$ or $SS_{Res}(X_1, X_2)$?
Which one is larger, $SS_R(X_1)$ or $SS_R(X_1, X_2)$?

Answers

$$SS_{Res}(X_1) \geq SS_{Res}(X_1, X_2)$$

$$SS_R(X_1) \leq SS_R(X_1, X_2)$$

Extra sum of squares

$$SS_{Res}(X_2|X_1) = SS_{Res}(X_1) - SS_{Res}(X_1, X_2)$$

It measures the marginal effect of adding X_2 to the regression model, when X_1 is already in the model (and X_2 is regarded as an extra variable).

$$SS_R(X_2|X_1) = SS_R(X_1, X_2) - SS_R(X_1)$$

it is the marginal increase in the regression sum of squares.

$$SSTO = SSR(X_1) + SSR(X_2|X_1) + SSE(X_1, X_2)$$

4.4.3 Test several $\beta_j = 0, j = 1, \ldots, r$

$$\mathcal{H}_0 : \beta_1 = \ldots = \beta_r = 0 \qquad \longleftrightarrow \qquad \mathcal{H}_a : \text{not all } \beta_j \neq 0$$

4.4.3 Test several $\beta_j = 0, j = 1, \ldots, r$

$$\mathcal{H}_0 : \beta_1 = \ldots = \beta_r = 0 \qquad \longleftrightarrow \qquad \mathcal{H}_a : \text{not all } \beta_j \neq 0$$

Test statistic

$$\begin{aligned} F &= \frac{SS_R(X_1, \ldots, X_r | X_{r+1}, \ldots, X_k)}{r} \div \frac{SS_{Res}(X_1, \ldots, X_k)}{n-k-1} \\ &= \frac{SS_R(X_1, \ldots, X_r | X_{r+1}, \ldots, X_k)}{r \ MS_{Res}} \end{aligned}$$

$$\begin{aligned} & SS_R(X_1, \ldots, X_r | X_{r+1}, \ldots, X_k) \\ = \ & SS_R(\text{use all predictor variables in the model}) \\ & - SS_R(\text{use predictor variables } X_{r+1}, \ldots, X_k) \\ = \ & SS_{Res}(\text{use predictor variables } X_{r+1}, \ldots, X_k) \\ & - SS_{Res}(\text{use all predictor variables in the model}) \end{aligned}$$

◦ **Derive the F test via the** *likelihood-ratio-test method*

$$\mathcal{H}_0: \ \beta_1 = \beta_2 = \ldots = \beta_r = 0 \quad \longleftrightarrow \quad \mathcal{H}_1: \ \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

◦ Derive the F test via the *likelihood-ratio-test method*

$$\mathcal{H}_0: \ \beta_1 = \beta_2 = \ldots = \beta_r = 0 \quad \longleftrightarrow \quad \mathcal{H}_1: \ \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

The parameter space associated with $\mathcal{H}_0$ is

$$\Theta_0 = \{(\beta_0, 0, \ldots, 0, \beta_{r+1}, \ldots, \beta_k, \sigma^2) : \ \beta_0 \in \mathbb{R}, \beta_{r+1} \in \mathbb{R}, \ldots, \beta_k \in \mathbb{R}, \sigma^2 \in (0, \infty)\}$$

◦ Derive the F test via the *likelihood-ratio-test method*

$$\mathcal{H}_0 : \ \beta_1 = \beta_2 = \ldots = \beta_r = 0 \quad \longleftrightarrow \quad \mathcal{H}_1 : \ \text{Not all } \beta s \text{ are equal to zero}$$

The whole parameter space in this case is

$$\Theta = \{(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) : (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1}, \sigma^2 \in (0, \infty)\}$$

The parameter space associated with $\mathcal{H}_0$ is

$$\Theta_0 = \{(\beta_0, 0, \ldots, 0, \beta_{r+1}, \ldots, \beta_k, \sigma^2) : \ \beta_0 \in \mathbb{R}, \beta_{r+1} \in \mathbb{R}, \ldots, \beta_k \in \mathbb{R}, \sigma^2 \in (0, \infty)\}$$

Likelihood function of parameter $(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)$

$$\ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)\right\}$$

Likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}$$

Likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}$$

$$\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) \quad = \quad \ell(*, 0, \ \ldots, 0, *, \ldots, *, ?)$$

Likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}$$

$$\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) \quad = \quad \ell(*, 0, \ \ldots, 0, *, \ldots, *, ?)$$

$$\begin{aligned} \sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2) \quad &= \quad \ell\left(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k, \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right) \\ &= \quad \ell\left(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k, \frac{1}{n} SS_{Res}\right) \end{aligned}$$

The likelihood ratio

$$\Lambda = \frac{\sup\limits_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup\limits_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)} = (1+?F)^{-\frac{n}{2}},$$

which is a decreasing function of F.

The likelihood ratio

$$\Lambda = \frac{\sup_{\Theta_0} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)}{\sup_{\Theta} \ell(\beta_0, \beta_1, \ldots, \beta_k, \sigma^2)} = (1+?F)^{-\frac{n}{2}},$$

which is a decreasing function of F.

Thus, a larger F would lead us to conclude $\mathcal{H}_a$.