STK333
Pengantar Model Linear

Pengujian Hipotesis

Model Linier

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i$ dimana i=1,2, …,n

$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

X berukuran n×p (dimana p=k+1) dengan r(X)= k+1=p,

$\boldsymbol{\varepsilon}$ vektor acak berukuran nx1 menyebar normal ganda dengan

$E[\boldsymbol{\varepsilon}] = 0$ dan $Var[\boldsymbol{\varepsilon}] = I\sigma^2$ $\quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 I)$

Penduga parameter $\boldsymbol{\beta}$ untuk model berpangkat penuh dan sisaannya yaitu

$\mathbf{b} = (X'X)^{-1}X'\mathbf{y}$ $\qquad \mathbf{e} = \mathbf{y} - X'\mathbf{b}$

$$JK_{Res} = \mathbf{y}'(\mathbf{I} - X(X'X)^{-1}X')\mathbf{y}$$
$$= \mathbf{y}'\mathbf{y} - \mathbf{y}'X(X'X)^{-1}X'\mathbf{y}$$
$$= JKTotal - JKRe_g$$

$$JK_{Total} = JKRe_g + JKRe_s$$

Theorema 4.1.1: $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 I) \Longrightarrow \mathbf{y} \sim N(X\boldsymbol{\beta}, \sigma^2 I)$

$$\frac{JK_{Reg}}{\sigma^2} \sim \chi^2_{r,\lambda} \qquad r(X(X'X)^{-1}X') = k+1 \qquad \lambda = \frac{1}{2\sigma^2}\boldsymbol{\beta}'(X'X)\boldsymbol{\beta}$$

Theorema 4.1.2: $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 I) \Longrightarrow \mathbf{y} \sim N(X\boldsymbol{\beta}, \sigma^2 I)$

$$\frac{JK_{Res}}{\sigma^2} \sim \chi^2_{r,\lambda} \qquad r(X(X'X)^{-1}X') = n - (k+1) \qquad \lambda = 0$$

Theorema 4.1.3: JK_{Reg} dan JK_{Res} saling bebas

Jika $r(X) = k+1 \Rightarrow X'X$ definit positif

Bila $\varepsilon \sim N0,\sigma^2 I$ maka $y \sim N(X\beta,\sigma^2 I)$ sehingga

$$\frac{JK_{Reg}}{\sigma^2} \sim \chi^2_{r,\lambda} \qquad r(X(X'X)^{-1}X') = k+1 \qquad \lambda = \frac{1}{2\sigma^2}\boldsymbol{\beta}'(X'X)\boldsymbol{\beta}$$

Kalau p=k+1, maka

$$\begin{aligned} E\left[\frac{JK_{Reg}}{p}\right] &= E\left[KT_{Reg}\right] \\ &= \sigma^2 + \frac{1}{p}\boldsymbol{\beta}'(X'X)\boldsymbol{\beta} \\ &= \sigma^2 \text{, bila } \boldsymbol{\beta} = \mathbf{0} \end{aligned}$$

Selanjutnya ingin diketahui apakah model linier yang dibangun mampu menjelaskan **keberagaman dari nilai respon y**. Jika tidak maka seluruh koefisien bernilai **nol**. Sebaliknya, minimal ada **satu koefisien yang tidak bernilai nol**.

Untuk menguji model tersebut, dibentuk hipotesis

$$H_0: \beta=0 \text{ vs } H_1: \beta\neq 0$$

Metode yang digunakan untuk menguji H_0 yaitu dengan **Analisis Ragam** (Anova) dimana statistik ujinya adalah **rasio dari JK Reg/db dengan JK Res/db**.

Jika H_0 benar

$$\boxed{\frac{JK_{Reg}/p\sigma^2}{JK_{Res}/(n-p)\sigma^2} = \frac{JK_{Reg}/p}{JK_{Res}/(n-p)} = \frac{KT_{Reg}}{KT_{Res}}}$$

akan menyebar menurut sebaran $F_{(p,n-p)}$

Tabel Anova untuk menguji kecukupan model.
(p = jumlah parameter dalam model).

Sumber Keragaman	Jumlah Kuadrat	Derajat Bebas	Kuadrat Tengah	F
Regresi	$\mathbf{y}'X(X'X)^{-1}X'\mathbf{y}$	p=k+1	JK_{Reg}/p	KT_{Reg}/KT_{Res}
Sisaan (Galat)	$\mathbf{y}'(\mathbf{I} - X(X'X)^{-1}X')\mathbf{y}$	n-p	$JK_{Res}/(n-p)$	
Total	$\mathbf{y}'\mathbf{y}$	n		

Teladan1.

y	x1	x2	x3
22.6	4	44	18
15.0	2	33	15
78.1	20	80	80
28.0	6	24	21
80.5	6	227	50
24.5	3	20	18
20.5	4	41	13
147.6	16	187	137
4.2	4	19	15
48.2	6	50	21
20.5	5	48	17

Teladan1.

Model regresi linier:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i \qquad i = 1,2, \dots 11$$

$$X = \begin{bmatrix} 1 & 4 & 44 & 18 \\ 1 & 2 & 33 & 15 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 5 & 48 & 17 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 22.6 \\ 15.0 \\ \vdots \\ \vdots \\ 20.5 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 0.3197263 & -0.0408268 & -0.00202208 & 0.005305965 \\ -0.0408268 & 0.0140738 & 0.0003717104 & -0.00224159 \\ -0.00202208 & 0.0003717104 & 0.00005188447 & -0.000113861 \\ 0.005305965 & -0.00224159 & -0.000113861 & 0.0004938527 \end{bmatrix}$$

Teladan1.

Untuk mengetahui apakah model tersebut cukup untuk menjelaskan keberagaman nilai y maka dilakukan pengujian terhadap hipotesis berikut

H_0: $\boldsymbol{\beta}=\mathbf{0}$ vs H_1: $\boldsymbol{\beta}\neq\mathbf{0}$

$$\text{JKTotal} = \mathbf{y}'\mathbf{y} = 39667.01$$

$$\text{JKReg} = \mathbf{y}'X(X'X)^{-1}X'\mathbf{y} = 38978.38$$

$$\text{JKRes} = \mathbf{y}'(\mathbf{I} - X(X'X)^{-1}X')\mathbf{y} = \text{JKTotal} - \text{JKReg} = 688.63$$

$$\text{KTReg} = \text{JK}_{\text{Reg}}/p = 38978.38/4 = 9744.595$$

$$\text{KTRes} = \text{JK}_{\text{Res}}/(n-p) = 688.63/(11-4) = 98.375$$

$$F_h = \text{KTReg}/\ \text{KTRes} = 9744.595/\ 98.375 = 99.055$$

Teladan1.

Tabel Anova untuk menguji **β** secara **bersama**

Sumber Keragaman	Jumlah Kuadrat	Derajat Bebas	Kuadrat Tengah	F
Regresi	38978.38	4	9744.595	99.055
Sisaan (Galat)	688.63	7	98.375	
Total	39667.01	11		

Pada pengujian parameter secara bersama, ketika menghasilan keputusan tolak H_0 maka minimal ada satu parameter yang tidak bernilai nol, dan itu bisa parameter apa saja.

Seringkali peneliti ingin mengetahui bagaimana pengaruh peubah bebas tertentu terhadap peubah respon ketika peubah bebas lainnya sudah ada dalam model.

Untuk melakukan pengujian terhadap kasus seperti ini maka perlu dilakukan partisi terhadap parameter yang akan diuji.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ . \\ \beta_{r-1} \\ --- \\ \beta_r \\ . \\ \beta_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ --- \\ \boldsymbol{\gamma}_2 \end{bmatrix} \qquad X = \begin{bmatrix} \underset{n \times r}{X_1} & \underset{n \times (k+1-r)}{X_2} \end{bmatrix}$$

Hipotesisnya:

H_0: $\boldsymbol{\gamma}_1 = 0$ vs H_1: $\boldsymbol{\gamma}_1 \neq 0$

(H_0: apakah kovariat pada partisi pertama berpengaruh terhadap y)

$\mathbf{y} = \mathrm{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

(ketika H_0 **benar** maka model tereduksi menjadi $\mathbf{y} = \mathrm{X}_2\boldsymbol{\gamma}_2 + \boldsymbol{\varepsilon}^*$

$\mathrm{R}(\boldsymbol{\beta}) = \mathbf{y}'\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}'\mathbf{y}$

$\mathrm{R}(\boldsymbol{\gamma}_2) = \mathbf{y}'\mathrm{X}_2(\mathrm{X}_2'\mathrm{X}_2)^{-1}\mathrm{X}_2'\mathbf{y}$

$\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2) = \mathrm{R}(\boldsymbol{\beta}) - \mathrm{R}(\boldsymbol{\gamma}_2)$

$\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)$ adalah bagian keragaman yang tidak acak (random) tetapi tidak dapat diukur oleh model tereduksi.

Perhatikan persamaan identitas berikut:

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'\mathrm{X}_2(\mathrm{X}_2\mathrm{X}_2)^{-1}\,\mathrm{X}_2{}'\mathbf{y} + \mathbf{y}'\mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}' - \mathrm{X}_2(\mathrm{X}_2\mathrm{X}_2)^{-1}\,\mathrm{X}_2{}')\mathbf{y} + \mathbf{y}'(\mathbf{I} - \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}')\mathbf{y}$$

$$\mathbf{y}'\mathrm{y} = \mathrm{R}(\boldsymbol{\gamma}_2) + \mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2) + \mathrm{JKRes}$$

$$\mathbf{y}'\mathrm{y}/\sigma^2 = \mathrm{R}(\boldsymbol{\gamma}_2)/\sigma^2 + \mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/\sigma^2 + \mathrm{JKRes}/\sigma^2$$

Lemma:

$$\mathrm{r}(\mathrm{X}_2(\mathrm{X}_2\mathrm{X}_2)^{-1}\,\mathrm{X}_2{}') = p - r$$

$$A = \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}' - \mathrm{X}_2(\mathrm{X}_2\mathrm{X}_2)^{-1}\,\mathrm{X}_2{}' \text{ adalah matriks idempoten}$$

$$\mathrm{r}(A) = r$$

$$\mathrm{r}(\mathbf{I} - \mathrm{X}(\mathrm{X}'\mathrm{X})^{-1}\mathrm{X}') = n - p$$

Theorema (Cohran-Fisher):

Bila $z \sim N(\mu, I)$ dan $z'z = \sum_{i=1}^{m} y' A_i\, y$ maka

semua $\mathbf{y}' A_i\, \mathbf{y}$ bebas dan menyebar menurut $\chi^2_{r_i, \lambda i}$

dengan $r_i = r(A_i)$ dan $\lambda_i = (1/2)\, \mu' A_i\, \mu$ $\Longleftrightarrow$ $\sum_{i=1}^{m} r_i = ni$ [$A_i\, A_j = 0$]

Berdasarkan persamaan identitas berikut:

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'X_2(X_2X_2)^{-1}X_2'\mathbf{y} + \mathbf{y}'X(X'X)^{-1}X' - X_2(X_2X_2)^{-1}X_2')\mathbf{y} + \mathbf{y}'(\mathbf{I} - X(X'X)^{-1}X')\mathbf{y}$$

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'A_2\mathbf{y} + \mathbf{y}'A_1\mathbf{y} + \mathbf{y}'A_3\mathbf{y}$$

Bila $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, $r(X) = p = k+1$, $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 I)$ dan $z = y/\sigma$

dan $z = y/\sigma$ maka $\mathbf{z} \sim N(X\boldsymbol{\beta}/\sigma, I)$

$$\mathbf{z}'\mathbf{z} = \mathbf{y}'\mathbf{y}/\sigma^2 = \mathrm{R}(\boldsymbol{\gamma}_2)/\sigma^2 + \mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/\sigma^2 + \mathrm{JKRes}/\sigma^2$$

persamaan ini memenuhi theorema Cochran-Fisher

$\frac{\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)}{\sigma^2} \sim \chi^2_{r,\lambda}$ dengan $\lambda = \frac{1}{\sigma^2}(\mathrm{X}\boldsymbol{\beta})'\mathrm{A}_2\mathrm{X}\boldsymbol{\beta}$

Theorema 4.2.2:

Bila H_0: $\boldsymbol{\gamma}_1 = 0$ maka

$$\frac{\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/\mathrm{r}\sigma^2}{\mathrm{JK}_{\mathrm{Res}}/(n-p)\sigma^2} = \frac{\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/\mathrm{r}}{\mathrm{JK}_{\mathrm{Res}}/(n-p)} \sim F_{(r;n-p)}$$

Tabel Anova untuk menguji hipotesis subvektor **β.**

p = jumlah parameter dalam model

r = banyaknya parameter dalam model yang ditanyakan

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F
Regresi				
Model Penuh	p	$R(\boldsymbol{\beta})$		
Model Tereduksi	p-r	$R(\boldsymbol{\gamma}_2)$		
$\boldsymbol{\gamma}_1$ setelah $\boldsymbol{\gamma}_2$	r	$R(\boldsymbol{\beta}) - R(\boldsymbol{\gamma}_2)$	$R(\boldsymbol{\gamma}_1 \vert \boldsymbol{\gamma}_2)/r$	$KT(\boldsymbol{\gamma}_1 \vert \boldsymbol{\gamma}_2)/KT_{Res}$
Sisaan / Galat	n - p	$\mathbf{y}'\mathbf{y} - R(\boldsymbol{\beta})$	$JK_{Res}/(n-p)$	
Total	n	$\mathbf{y}'\mathbf{y}$		

Teladan 2.

Pada Teladan1 model terdiri dari satu intercept dan 3 regressor diperoleh kesimpulan bahwa minimal ada satu parameter **β** yang tidak sama dengan nol. Hal ini mengandung kemungkinan intercept $\beta_0 \neq 0$. Selajutnya pada Teladan 2 ini ingin menguji lebih lanjut apakah regressor x1, x2 dan x3 perlu ditambahkan saat intercept sudah ada dalam model.

Penyelesaian:
Sebelumnya kita akan mengatur ulang dan mempartisi matriks X dan vektor **β** sesuai dengan pengujian yang akan dilakukan

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \vdots & 1 \\ x_{21} & x_{22} & x_{23} & \vdots & 1 \\ x_{31} & x_{32} & x_{33} & \vdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{111} & x_{112} & x_{113} & \vdots & 1 \end{bmatrix} = [X_1 \mid X_2]$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ -- \\ \beta_0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ --- \\ \boldsymbol{\gamma}_2 \end{bmatrix}$$

Kita akan menguji $H_0{:}\boldsymbol{\gamma}_1=0$ (model tereduksi adalah model yang tepat) versus $H_1{:}\boldsymbol{\gamma}_1\neq 0$ (model penuh adalah model yang lebih tepat).

Model penuhnya adalah

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Model tereduksinya adalah

$$\mathbf{y} = \beta_0 + \boldsymbol{\varepsilon}^*$$

$$R(\boldsymbol{\gamma}_2) = \mathbf{y}' X_2 (X_2 X_2)^{-1} X_2' \mathbf{y}$$

$$X_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \Rightarrow$$

$$R(\boldsymbol{\gamma}_2) = \mathbf{y}' X_2 (X_2 X_2)^{-1} X_2' \mathbf{y}$$

$$R(\boldsymbol{\gamma}_2) = \left(\sum_{i=1}^{n} y_i\right)^2 / n$$

$$= \left(\sum_{i=1}^{11} y_i\right)^2 / 11$$

$$= (22.6 + 15.0 + \ldots + 20.5)^2 / 11$$

$$= (489.7)^2 / 11$$

$$= 21800.55$$

Pada Teladan 1 telah diperoleh $\mathrm{R}(\boldsymbol{\beta}) = 38978.38$

sehingga $\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2) = \mathrm{R}(\boldsymbol{\beta}) - \mathrm{R}(\boldsymbol{\gamma}_2)$

$$= 38978.38 - 21800.55$$

$$= 17177.83$$

Pada Teladan 1 juga telah diperoleh $\mathrm{JKRes} = 688.63$

Sehingga F untuk menguji H_0 adalah

$$\frac{\mathrm{R}(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/\mathrm{r}}{\mathrm{JK}_{\mathrm{Res}}/(n-p)} = \frac{17177.83/3}{688.63/7} = 58.2$$

Titik kritis pada taraf nyata α=0.01 adalah F = 8.45. Nilai Fhitung jauh melampaui titik kritisnya, sehingga cukup bukti untuk mengatakan bahwa model yang hanya terdiri dari intercept tidak cukup untuk menjelaskan keberagaman dari nilai respon (y).

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F
Regresi				
Model Penuh	4	38978.38		
Model Tereduksi	1	21800.55		
γ_1 setelah γ_2	3	17177.83	5725.9	58.2
Sisaan / Galat	7	688.63	98.375	
Total	11	39667.01		

Dalam analisis regresi linier peneliti lebih tertarik untuk mengetahui bagaimana pengaruh dari peubah bebasnya/regressor (ketika intercept sudah ada dalam model) terhadap peubah respon. Dapat dilakukan dengan membuat $\boldsymbol{\gamma}_2 = \beta_0$

Bila $\boldsymbol{\gamma}_2 = \beta_0$ maka $\mathbf{x}_2' = [1 \quad 1 \quad \dots \quad 1]$

Sehingga $R(\boldsymbol{\gamma}_2) = \mathbf{y}' X_2 (X_2 X_2)^{-1} X_2' \mathbf{y} = (\sum_{i=1}^{n} y_i)^2 / n$

Tabel Anova untuk menguji hipotesis semua regressor saat intercept sudah ada dalam model berdasarkan total yang belum terkoreksi

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F
Regresi				
Model Penuh	p	$R(\boldsymbol{\beta})$		
Model Tereduksi	1	$R(\boldsymbol{\gamma}_2)$		
γ_1 setelah $\boldsymbol{\gamma}_2$	p – 1	$R(\boldsymbol{\beta}) - R(\boldsymbol{\gamma}_2)$	$R(\boldsymbol{\gamma}_1 \vert \boldsymbol{\gamma}_2)/(p-1)$	$KT(\boldsymbol{\gamma}_1 \vert \boldsymbol{\gamma}_2)/KT_{Res}$
Sisaan / Galat	n – p	$\mathbf{y}'\mathbf{y} - R(\boldsymbol{\beta})$	$JK_{Res}/(n-p)$	
Total	n	$\mathbf{y}'\mathbf{y}$		

Catatan: Teladan untuk pengujian ini sama seperti Teladan 2

Tabel Anova untuk menguji untuk pengujian keseluruhan regressor saat intercept sudah ada dalam model berdasarkan Total terkoreksi.

Sumber Keragaman	Jumlah Kuadrat	Derajat Bebas	Kuadrat Tengah	F
Regresi	$R(\boldsymbol{\gamma}_1 \mid \boldsymbol{\gamma}_2)$	p-1	$R(\boldsymbol{\gamma}_1 \mid \boldsymbol{\gamma}_2)/(p-1)$	$KT_{(\boldsymbol{\gamma}_1 \mid \boldsymbol{\gamma}_2)}/KT_{Res}$
Sisaan (Galat)	$\mathbf{y'y} - R(\boldsymbol{\beta})$	n-p	$JK_{Res}/(n-p)$	
Total	$\mathbf{y'y}$	n-1		

Teladan 3.

Pada Teladan 3 ini dilakukan pengujian untuk semua yaitu x1, x2 dan x3, apakah perlu ditambahkan saat intercept sudah ada dalam model berdasarkan total terkoreksi.

Penyelesaian:
Kita akan menguji H_0:$\boldsymbol{\gamma}_1$=0 (regressor tidak perlu ditambahkan dalam model) versus H_0:$\boldsymbol{\gamma}_1$≠0 (minimal ada satu regressor yang perlu ditambahkan dalam model)

Dari hasil perhitungan pada Teladan 1 dan 2 diperoleh

$$\mathbf{y}'\mathbf{y} = 39667.01 \qquad \mathrm{R}(\boldsymbol{\beta}) = 38978.38$$

$$\mathrm{JKRes} = \mathbf{y}'\mathbf{y} - \mathrm{R}(\boldsymbol{\beta}) = 688.63$$

$$\mathrm{R}(\boldsymbol{\gamma}_2) = \left(\sum_{i=1}^{11} y_i\right)^2/11 = 21800.55$$

$$R(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2) = R(\boldsymbol{\beta}) - R(\boldsymbol{\gamma}_2)$$
$$= 38978.38 - 21800.55$$
$$= 17177.83$$

$$JKT_{\text{Terkoreksi}} = \mathbf{y'y} - R(\boldsymbol{\gamma}_2)$$
$$= 39667.01 - 21800.55$$
$$= 17866.46$$

Sehingga nilai F untuk menguji H_0 adalah

$$\frac{R(\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2)/(p-1)}{JK_{\text{Res}}/(n-p)} = \frac{17177.83/3}{688.63/7} = 58.2$$

Titik kritis pada taraf nyata α=0.01 adalah F = 8.45. Nilai Fhitung jauh melampau titik kritisnya, sehingga cukup bukti untuk mengatakan bahwa model yang hanya terdiri dari intercept tidak cukup untuk menjelaskan keberagaman dari nilai respon (y).

Sumber Keragaman	Jumlah Kuadrat	Derajat Bebas	Kuadrat Tengah	F
Regresi	17177.83	3	5725.9	58.2
Sisaan (Galat)	688.63	7	98.375	
Total	17866.46	10		

Pengujian secara umum untuk $\boldsymbol{\beta}^* \neq \mathbf{0}$ adalah

H_0: $\boldsymbol{\beta} = \boldsymbol{\beta}^*$ vs H_1: $\beta \neq \boldsymbol{\beta}^*$ $\Longrightarrow$ H_0: $\boldsymbol{\beta} - \boldsymbol{\beta}^* = \mathbf{0}$ vs H_1: $\boldsymbol{\beta} - \boldsymbol{\beta}^* \neq \mathbf{0}$

Teorema: $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, $r(X) = p$, $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 I)$

$\mathbf{b} \sim N(\boldsymbol{\beta}, (X'X)^{-1}\sigma^2)$

$\mathbf{z} = \mathbf{b} - \boldsymbol{\beta}^*$

$\mathbf{z} \sim N(\boldsymbol{\beta} - \boldsymbol{\beta}^*, (X'X)^{-1}\sigma^2)$

$$\mathbf{z}'A\mathbf{z} = \frac{1}{\sigma^2}(\mathbf{b} - \boldsymbol{\beta}^*)'(X'X)(\mathbf{b} - \boldsymbol{\beta}^*) \sim \chi^2_{r,p}$$

$r(X'X)^{-1} = p$ $\qquad\qquad$ $\lambda = \frac{1}{2\sigma^2}(\boldsymbol{\beta} - \boldsymbol{\beta}^*)'(X'X)(\boldsymbol{\beta} - \boldsymbol{\beta}^*)$

$$\mathbf{E}[(\mathbf{b} - \boldsymbol{\beta}^*)'(X'X)(\mathbf{b} - \boldsymbol{\beta}^*)] = p\sigma^2 + (\boldsymbol{\beta} - \boldsymbol{\beta}^*)'(X'X)(\boldsymbol{\beta} - \boldsymbol{\beta}^*)$$

$$\frac{(\mathbf{b} - \boldsymbol{\beta}^*)'(X'X)(\mathbf{b} - \boldsymbol{\beta}^*)/\sigma^2 p}{JK_{Res}/\sigma^2(n-p)} = \frac{(\mathbf{b} - \boldsymbol{\beta}^*)'(X'X)(\mathbf{b} - \boldsymbol{\beta}^*)/p}{JK_{Res}/\,(n-p)} \sim \mathbf{F}p;(n-p)$$

Secara umum sama untuk menguji H_0: $\boldsymbol{\gamma_1} = \boldsymbol{\gamma_*}$ vs H_1: $\boldsymbol{\gamma_1} \neq \boldsymbol{\gamma_*}$

$$\frac{(\hat{\gamma}_1 - \boldsymbol{\gamma_*})'(X'X)(\hat{\gamma}_1 - \boldsymbol{\gamma_*})/r}{JK_{Res}/\,(n-p)} \sim \mathbf{F}r;(n-p)$$

Pengujian yang lebih umum adalah mengabungkan beberapa hipotesis menjadi bentuk

H_0: C$\boldsymbol{\beta} = \boldsymbol{\delta}^*$ vs H_1: C$\boldsymbol{\beta} \neq \boldsymbol{\delta}^*$

$\Rightarrow$ H_0: C$\boldsymbol{\beta}$ - $\boldsymbol{\delta}^* = \mathbf{0}$ vs H_1: C$\boldsymbol{\beta}$ - $\boldsymbol{\delta}^* \neq \mathbf{0}$

$\mathbf{z} = C\boldsymbol{\beta} - \boldsymbol{\delta}^*$

$\mathbf{z} \sim N(C\boldsymbol{\beta} - \boldsymbol{\delta}^*, (C(X'X)^{-1}C')\sigma^2)$

Lemma: Jika $r(C_{r\times p}) = r < p$ dan $r(X_{n\times p}) = p$ ➔ $C(X'X)^{-1}C'$ definit positif

$$\mathbf{z}'A\mathbf{z} = \frac{1}{\sigma^2}(C\boldsymbol{\beta} - \boldsymbol{\delta}*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}*) \sim \chi^2_{r,p}$$

$r(C(X'X)^{-1}C')^{-1} = r$ $\qquad \lambda = \frac{1}{2\sigma^2}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)$

$$\mathbf{E}[C\boldsymbol{\beta} - \boldsymbol{\delta}^*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)] = r\sigma^2 + (C\boldsymbol{\beta} - \boldsymbol{\delta}^*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)$$

$$\frac{(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)/\sigma^2 r}{JK_{Res}/\sigma^2(n-p)} = \frac{(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)'(C(X'X)-1C')^{-1}(C\boldsymbol{\beta} - \boldsymbol{\delta}^*)/r}{JK_{Res}/\,(n-p)} \sim \mathbf{F}_{r;(n-p)}$$

Teladan 4.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i, \, i = 1,2,\ldots.., n$$

H_0: $\beta_1 = \beta_2 = \beta_3$

H_0: paling sedikit ada satu $\beta_i \neq \beta_j$ untuk $i \neq j$; I, j = 1,2,3

Ubah H_0: $\beta_1 = \beta_2 = \beta_3$ menjadi

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_4 \end{bmatrix}$$

H_0: $\beta_1 - \beta_2 = 0$

$\beta_2 - \beta_3 = 0$

$\beta_1 - \beta_3 = 0$

$$H_0: \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

H_0: $C\boldsymbol{\beta} = \mathbf{0}$ $\qquad r(C) = 2$

Pengujian Hipotesis