

KUG STATISTIKA REGRESI

1. Berdasarkan data jamak berikut ini ditabel 8.3

A. Tentukan Model Estimasi regresi jarak tempuh bensin terhadap pendapatan rumah (X_1) dan jumlah banal kendaraan (X_2) dan interpretasikan.

Jawaban Dan Pembahasan :

Untuk memperoleh model estimasi regresi dipergunakan perhitungan regresi linear berganda dengan X_1 dan X_2 sebagai variabel independen dan Y sebagai variabel dependen. Dengan melakukan proses perhitungan estimasi beta. Diperoleh model estimasi regresi :

$$\hat{Y} = 32,9200 + (-0,0530X_1) + 0,9295 X_2$$

Dengan Interpretasi :

Dari model estimasi regresi tersebut, diperoleh interpretasi bahwa variabel X_1 memiliki pengaruh negatif terhadap variabel dependen. Artinya, semakin tinggi nilai variabel X_1 , maka semakin rendah nilai variabel Y . Sementara, variabel X_2 memiliki pengaruh positif terhadap variabel dependen. Artinya, semakin tinggi nilai variabel X_2 , maka semakin tinggi variabel Y .

B. Hitunglah tabel Analysis of Variance (ANOVA) dan simpulkan hasil uji F yang dilakukan.

TABEL ANOVA

Source of Variation	Degrees Of Freedom	Sum Of Squares	Mean Square	Fo
Regression	2	972,8983957	486,4491978	
Residual	29	264,6456918	9,125713511	53,305333
Total	31	1237,544088		

Nilai-nilai Diatas dapat diperoleh dari rumus dan hasil perhitungan yang akan dijelaskan lebih lanjut, seperti dibawah ini :

* Perhitungan Degrees Of Freedom

1) DF Regression

$$DF_{Reg} = K = 2$$

2) DF Residual

$$DF_{Res} = n - K - 1 = 32 - 2 - 1 = 29$$

3) DF Total

$$DF_{Total} = n - 1 = 32 - 1 = 31$$

* Perhitungan Sum Of Squares

1) SS Regression

Pada nilai SS_{Reg} diperoleh dari perhitungan :

$$SS_{Reg} = \hat{\beta}'X'Y - \frac{(\sum Y_i)^2}{n} = 972,8983957$$

2) SS Residual

Pada nilai SS_{Res} diperoleh dari perhitungan :

$$SS_{Res} = Y'Y - \hat{\beta}'X'Y = 264,6456918$$

3) SS Total

Pada nilai SST diperoleh dari perhitungan:

$$SST = y'y - \left(\frac{\left(\sum_{i=1}^n y_i \right)^2}{n} \right) = 1259,54408811$$

* Perhitungan Mean Square

1) MS Regression

Pada nilai MSReg diperoleh dari perhitungan:

$$MS_{Reg} = \frac{SS_{Reg}}{k} = \frac{972,8983957}{2} = 486,4491978$$

2) MS Residual

Pada nilai MS Res diperoleh dari perhitungan:

$$MS_{Res} = \frac{SS_{Res}}{n-p} = \frac{264,6456918}{32-3} = 9,125713511$$

* Perhitungan F0

Pada nilai F0 diperoleh dari perhitungan:

$$F_0 = \frac{MS_{Reg}}{MS_{Res}} = \frac{486,4491978}{9,125713511} = 53,30533363$$

KESIMPULAN DAN INTERPRETASI UJI F

F-Hitung	53,3053363
F-Tabel	3,3276545

* Perhitungan F-Hitung

$$F\text{-Hitung} = \frac{MS_{Reg}}{MS_{Res}} = \frac{486,4491978}{9,125713511} = 53,30533363$$

* Perhitungan F-Tabel

$$F\text{-Tabel} = F_{\alpha, k, n-p} = 0,5, 2, (32-3) = 3,3276545$$

Diperoleh kesimpulan dan interpretasi:

H_0 ditolak.

H_0 ditolak, karena $F\text{-Hitung} > F\text{-Tabel}$. Dimana, $53,3053363 > 3,3276545$. Maka, dapat diambil kesimpulan bahwa terdapat minimal satu variabel independen yang berpengaruh terhadap variabel dependen.

c. Hitung dan interpretasikan hasil dan R^2 dan R^2 Adj untuk model ini. Apakah terdapat perbedaan antara perhitungannya tersebut. Jika berbeda, kenapa?

Jawaban dan Pembahasan:

Diperoleh Hasil Perhitungan R^2 (Adj)

R^2 (Adj)	0,771404412
R^2 (Adj) %	77,14%
Sisa	0,228595588
Sisa %	22,86%

Diperoleh Hasil Perhitungan R^2

R^2	0,7802575
R^2 %	78,02%
Sisa	0,2138475
Sisa %	21,38%

* Perhitungan R^2

$$R^2 = 1 - \frac{SS_{Res}}{SST} = 1 - \frac{264,6456918}{1237,544088} = 0,7861525 = 78,61\%$$

$$Sisa = 1 - R^2 = 1 - 0,7861525 = 0,2138475 = 21,38\%$$

Interpretasi : Diperoleh nilai R^2 sebesar 0,7861525 menunjukkan bahwa 78,61% Varians Y yang telah terjadi dapat dijelaskan oleh model regresi. Sisa 21,38% Varians Y yang telah terjadi masih belum dapat dijelaskan oleh model regresi.

* Perhitungan R^2 (Adj)

$$R^2 (Adj) = 1 - \frac{SS_{Res} (n-p)}{SST (n-1)} = 1 - \frac{264,6456918 (32-3)}{1237,544088 (32-1)} = 0,771404412 = 77,14\%$$

$$Sisa = 1 - R^2 (Adj) = 1 - 0,771404412 = 0,228595588 = 22,86\%$$

Interpretasi : Diperoleh nilai R^2 (Adj) sebesar 0,771404412 menunjukkan bahwa 77,14% Varians Y yang telah terjadi dapat dijelaskan oleh model regresi. Sedangkan 22,86% Varians Y yang telah terjadi masih belum dapat dijelaskan oleh model regresi. Setelah memperhitungkan jumlah variabel bebas

Alasan Perbedaan Perhitungan Antara R^2 dan R^2 (Adj)

Diperoleh perbedaan hasil perhitungan antara R^2 dan R^2 (Adj). Hal ini dikarenakan, pada nilai R^2 menunjukkan seberapa besar Varians Variabel dependen Y yang dapat dijelaskan oleh model regresi. Sedangkan, pada R^2 (Adj) menunjukkan seberapa besar Varians Variabel dependen Y yang dapat dijelaskan oleh model regresi setelah memperhitungkan jumlah variabel independen yang digunakan.

D. Temukan dan interpretasikan 95% confidence interval untuk β_1 .

Jawaban dan Pembahasan :

CONFIDENCE INTERVAL β_1

Batas Bawah	-0,065593538
Batas Atas	-0,040455979

* Perhitungan Batas Bawah

$$\begin{aligned} BB &= \hat{\beta}_1 - (t_{\text{tabel}} * \text{standar Error } \hat{\beta}_1) \\ &= -0,053024758 - (2,04522964 * 0,00614541) \\ &= -0,065593538 \end{aligned}$$

$$\text{Dimana, nilai } t_{\text{tabel}} \text{ diperoleh dari : } t_{\frac{\alpha}{2}} (n-p) = \frac{0,05}{2} (32-3) = 2,04522964$$

* Perhitungan Batas Atas

$$\begin{aligned} BA &= \hat{\beta}_1 + (t_{\text{tabel}} * \text{standar Error } \hat{\beta}_1) \\ &= -0,053024758 + (2,04522964 * 0,00614541) \\ &= -0,040455979 \end{aligned}$$

* Interpretasi

Diyakini sebesar 95% β_1 dapat mewakili parameter β_1 pada populasi. Dengan nilai rentang koefisien β_1 ditentang -0,065593538 dan -0,040455979.



E. Hitunglah uji T statistik untuk $H_0: \beta_1 = 0$ dan $H_0: \beta_6 = 0$. Apa kesimpulan yang dapat diambil dari pengujian tersebut.

Jawaban Dan Pembahasan.

* Uji Beta Topi 1.

$$T\text{-Hitung} \quad 8,628348363$$

$$T\text{-tabel} \quad 2,045229642.$$

Penghitungan T-Hitung

$$T\text{-Hitung} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{-0,053024758}{0,00614541} = 8,628348363 //$$

Penghitungan T-tabel

$$T\text{-tabel} = t_{\frac{\alpha}{2}} (n-p) = \frac{0,105}{2} (32-3) = 2,045229642$$

Kesimpulan Dan Interpretasi

H_0 ditolak.

Variabel X_1 berpengaruh terhadap variabel Y .

* Uji Beta Topi 6

$$T\text{-Hitung} \quad 1,387090097$$

$$T\text{-tabel} \quad 2,045229642$$

Penghitungan T-Hitung

$$T\text{-Hitung} = \frac{\hat{\beta}_6}{SE(\hat{\beta}_6)} = \frac{0,192949867}{0,167040764} = 1,387090097 //$$

Penghitungan T-tabel

$$T\text{-tabel} = t_{\frac{\alpha}{2}} (n-p) = \frac{0,105}{2} (32-3) = 2,045229642$$

Kesimpulan Dan Interpretasi :

H_0 Gagal ditolak.

Variabel X_6 tidak berpengaruh terhadap variabel Y .

KESIMPULAN :

Dapat disimpulkan bahwa Variabel X_1 berpengaruh terhadap Variabel Y . Sedangkan Variabel X_6 tidak berpengaruh terhadap Variabel Y .

F. Tentukan dan interpretasikan 95% confidence interval pada rata-rata jarak tempuh basin ketika $x_1 = 275$ dan $x_6 = 2$.

Jawaban dan Pembahasan:

Dikann pada perhitungan ini, diperlukan nilai dari $\hat{\beta}_1$ dan $\hat{\beta}_6$ dengan perhitungan confidence interval.

CONFIDENCE INTERVAL $\hat{\beta}_6$

Batas Bawah	-0,4410243
Batas Atas	2,3000237

Dikann nilai ini diperoleh dari:

* Perhitungan Batas Bawah

$$BB = \hat{\beta}_6 - (T\text{-tabel} \times \text{standar error } \hat{\beta}_6) \\ = -0,4410243,$$

* Perhitungan Batas Atas

$$BA = \hat{\beta}_6 + (T\text{-tabel} \times \text{standar error } \hat{\beta}_6) \\ = 2,3000237,$$

* Interpretasi

95% kita yakin bahwa $\hat{\beta}_6$ terdapat mewakili parameter β_6 pada populasi.

Untuk nilai $x_1 = 275$ dan $x_6 = 2$. Maka:

* Perhitungan Batas Bawah

$$Y = 32,910041 - (-0,065593538)(275) - (-0,040455979)(2) \\ = 11,7909061,$$

* Perhitungan Batas Atas

$$Y = 32,910041 + (-0,040455979)(275) + 2,3000237(2) \\ = 26,38469412,$$

* Interpretasi

95% kita yakin bahwa nilai tersebut berada diantara pada rata-rata jarak tempuh basin dan mewakili parameter β_1 pada populasi.

G. Tentukan dan interpretasikan 95% prediction interval untuk observasi baru pada jarak tempuh basin ketika $x_1 = 275$ dan $x_6 = 2$.

Jawaban dan Pembahasan:

Batas Bawah	13,98976938
Batas Atas	26,38469412

* Perhitungan Batas Bawah

$$Y = 32,910041 - (-0,065593538)(275) - (-0,4410243)(2) \\ = 13,98976938,$$

* Perhitungan Batas Atas

$$Y = 32,910041 + (-0,040455979)(275) + (2,3000237)(2) \\ = 26,38469412,$$

* Interpretasi

95% kita yakin bahwa nilai tersebut berada diantara pada observasi baru yang
terbaca benar dan mewakili parameter β_0 pada populasi.

2. Berdasarkan Data Jarak tempur bensin dicat 8.3

A. Lakukan analisis residual menggunakan standarised residual dan studentized residual. Apa hasil analisis yang diperoleh.

Jawaban Dan Pembahasan:

* Perhitungan Standarised Residual

Untuk nilai standarised Residual diperoleh pada data "Data 8.3" yang memiliki standarised tertinggi sebesar 2,07024839.

$$\text{Dimana nilai tersebut diperoleh dari } \frac{e}{\sqrt{MSES}} = \frac{6,523971562}{\sqrt{9,125713511}} = 2,07024839$$

Interpretasi:

Nilai residual dan data terbesar pada data Automobile "Data 8.3" dengan nilai standarised residual sebesar 2,07024839.

Dikarenakan, nilai tersebut < 3. Maka tidak terjadi outlier pada setoran observasi data.

* Perhitungan Studentized Residual

Untuk nilai studentized Residual diperoleh pada data "Data 8.3" yang memiliki nilai studentized residual sebesar 2,225047062.

$$\text{Dimana nilai tersebut diperoleh dari } \frac{e}{\sqrt{MSES \cdot (1 - h_{ii})}} = \frac{6,523971562}{\sqrt{9,125713511 \cdot (1 - 0,0860477)}} = 2,225047062$$

Interpretasi:

Nilai residual dan data tersebut terdapat pada data Automobile "Data 8.3" dengan nilai studentized residual sebesar 2,225047062.

Dikarenakan, nilai tersebut < 3. Maka, tidak terjadi outlier pada setoran observasi.

B. Lakukan perhitungan dan interpretasi terhadap R^2 prediction. Apakah ada perbedaan antara R^2 dengan R^2 prediction. Menurutmu kenapa?

Jawaban Dan Pembahasan:

Nilai R^2

$$R^2 = 0,786357514$$

Nilai R^2 Prediction

$$R^2 \text{ Prediction} = 0,733506101$$

Perhitungan Nilai R^2 Prediction

$$R^2 \text{ Prediction} = 1 - \frac{PRESS}{SST} = 1 - \frac{329,7979494}{12371544088} = 0,733506101$$

Kesimpulan Dan Interpretasi:

Dapat diketahui bahwa nilai R^2 prediction memiliki nilai 0,7335 yang berarti bahwa nilai R^2 prediction menunjukkan hubungan model yang berkebalikan dengan nilai sebenarnya dan nilai R^2 sebesar 0,7861 yang berarti bahwa R^2 menunjukkan kinerja dependen yang dapat dijelaskan oleh model regresi.



Diketahui bahwa terdapat data eksperimen ukuran bekunya tanyan dipengaruhi oleh konsentrasi Asid.

A. Hitunglah tabel Analysis of Variance (ANOVA) dan simpulkan hasil uji F yang dilakukan.

Jawaban Dan Pembahasan:

TABEL ANOVA

Source Of Variation	Sum Of Squares	Degrees Of Freedom	Mean Square	Fo
Regression	392,850467	1	392,850467	
Residual	3698,363429	24	154,0984762	2,5493465
Total	4091,213896	25		

Dimana nilai-nilai tersebut diperoleh dengan menggunakan rumus yang sama seperti nomor tetapi menggunakan nilai n sebesar 26, nilai k sebesar 1, dan nilai p sebesar 2.

KESIMPULAN DAN INTERPRETASI UJI F

F-Hitung	2,5493465
F-Tabel	4,5296773

* Perhitungan F-Hitung

$$F\text{-Hitung} = \frac{MS_{reg}}{MS_{res}} = \frac{392,850467}{154,0984762} = 2,5493465$$

* Perhitungan F-Tabel

$$F\text{-Tabel} = F_{\alpha, k, n-p} = 0,05, 1, (26-2) = 4,5296773$$

* Kesimpulan Dan Interpretasi

H_0 gagal ditolak.

H_0 gagal ditolak, karena $F\text{-Hitung} < F\text{-Tabel}$. Dimana $2,5493465 < 4,5296773$.

Maka, dapat diambil kesimpulan bahwa terdapat minimal satu variabel independen yang tidak berpengaruh terhadap variabel dependen.

B. Lakukan analisis uji Lack of Fit dari data tersebut.

Jawaban Dan Pembahasan:

TABEL LACK OF FIT

Level of X	Fakta-Rata Y	$\bar{Y} - Y_i$	Degrees of Freedom
0,3	89,3375	1099,53875	7
0,5	80,4	366	7
0,7	99,5	966,82	7
Total	269,2375	2462,35875	21

Source Of Variation	Sum Of Squares	Degrees Of Freedom	Mean Square	F0
Regression	3921,8504167	1	3921,8504167	
Residual	3698,363429	21	176,0984762	
(Lack Of Fit)	1236,004679	3	412,0015598	3,513717388
(Pure Error)	2462,35875	21	117,2553766	
Total (Reg+Res)	4091,213846	25		

Penjelasan Perhitungan Lack Of Fit

* Level of X

Pada kolom level of X diperoleh dari nilai variabel X yang muncul lebih dari 1 kali dimana terdapat nilai 0,3, 0,5, dan 0,7.

* Rata-rata Y

Pada kolom Rata-rata Y diperoleh dari nilai Y variabel X yang muncul lebih dari 1

kali dengan menggunakan rumus = $\frac{\sum Y_i}{n}$

Dimana, diperoleh nilai-nilai: 89,3375, 80,4, 99,5.

dengan n yang dipergunakan adalah jumlah dari total Y yang dihitung atau banyaknya munculnya level of X.

* $\bar{Y} - Y_i$

Pada kolom $\bar{Y} - Y_i$ diperoleh dari nilai rata-rata Y dikurangi dengan Y_i dan dipangkatkan

2 dengan menggunakan rumus: $(\bar{Y} - Y_i)^2$.

Dimana, diperoleh nilai-nilai 1099,53875, 36, 996,87.

* Degrees Of Freedom

Pada kolom DF ini diperoleh dari nilai jumlah total munculnya X dikurangi 1.

Dimana diperoleh nilai-nilai 7,7,7.

Penjelasan Perhitungan Tabel Of Variance

Pada perhitungan ini diperlukan ingkud nilai SS_{Reg} , MS_{Reg} , DF_{Reg} , SS_{Res} , MS_{Res} , DF_{Res} , SS_{LoF} , MS_{LoF} , DF_{LoF} .

Dimana untuk nilai-nilai dari Pure Error, yakni:

$$SS_{Pe} = (\bar{Y} - Y_i)$$

$$MS_{Pe} = \frac{SS_{Pe}}{n-m}$$

Dimana untuk nilai-nilai dari Lack Of Fit, yakni:

$$SS_{LoF} = SS_{Res} - SS_{Pe}$$

$$MS_{LoF} = \frac{SS_{LoF}}{m-2}$$

Dimana m merupakan jumlah variabel X yang memiliki satu data. Jadi, dipermisakan bahwa $X = 0,3$ muncul sebanyak 8 kali, maka akan dihitung 1 saja. Diperoleh nilai $m=8$.

