EKONOMETRIKA

MODEL REGRESI LINIER

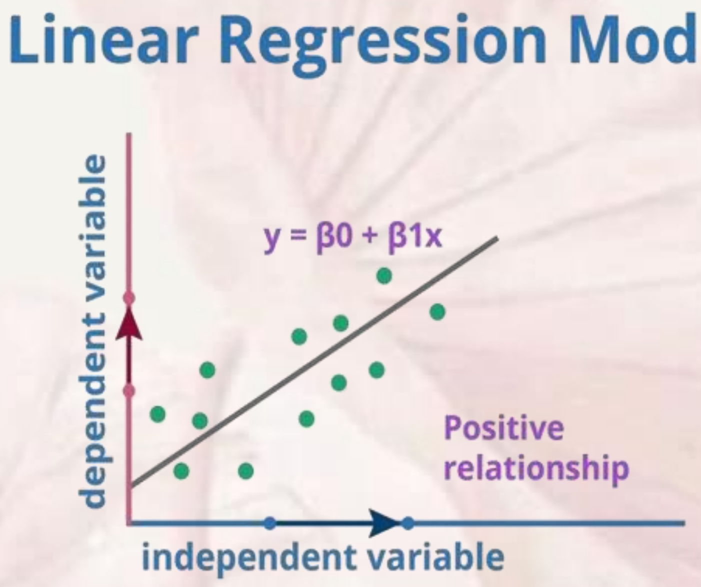
DOSEN PENGASUH:

Prof. Dr. Ir. ZULKIFLI ALAMSYAH, M.Sc.

PROGRAM STUDI AGRIBISNIS
FAKULTAS PERTANIAN UNIVERSITAS JAMBI

Model Regresi Linear

Mengukur ketergantungan suatu variabel tidak bebas (dependent variable) terhadap 1 atau lebih variabel bebas (explanatory/independent variables)



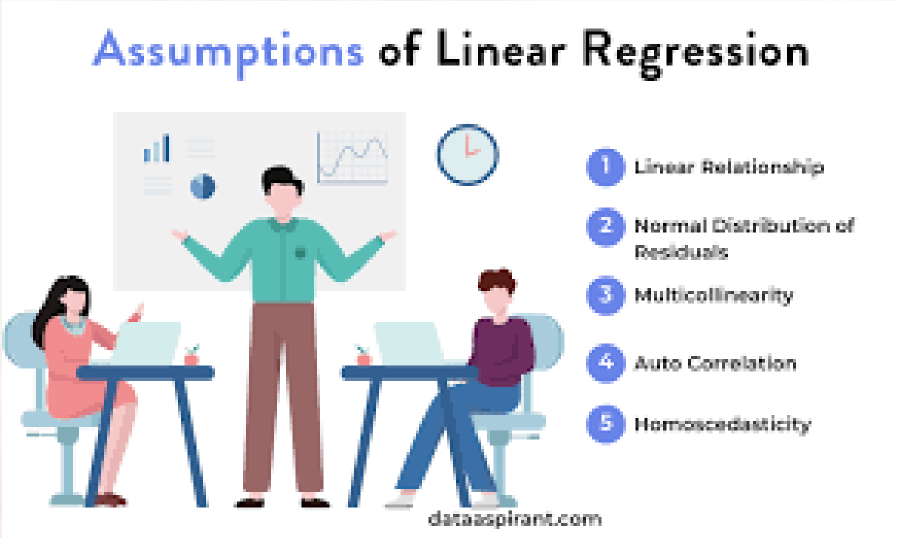
Tujuan

- Mengestimasi nilai tengah Variabel tdk bebas dari nilai rata-rata variabel bebas.
- Untuk menguji hipotesis mengenai sifat ketergantungan sesuai dengan teori ekonomi.

Beda dengan Korelasi

- Asimetri dalam memperlakukan Variabel:
  - Variabel bebas bersifat deterministik
  - Variabel tdk bebas bersifat stochastic (acak)
- ⇨ Untuk setiap nilai X tertentu, dapat memberikan beberapa nilai Y dengan probabilitas tertentu.

Asumsi-asumsi Mengenai μ_I:

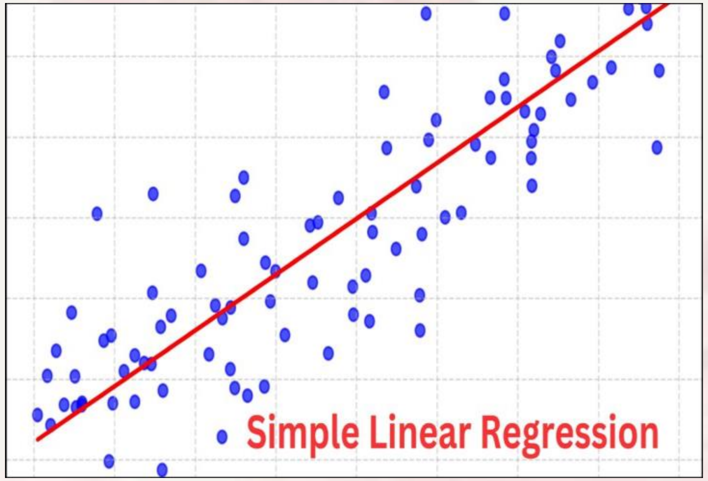


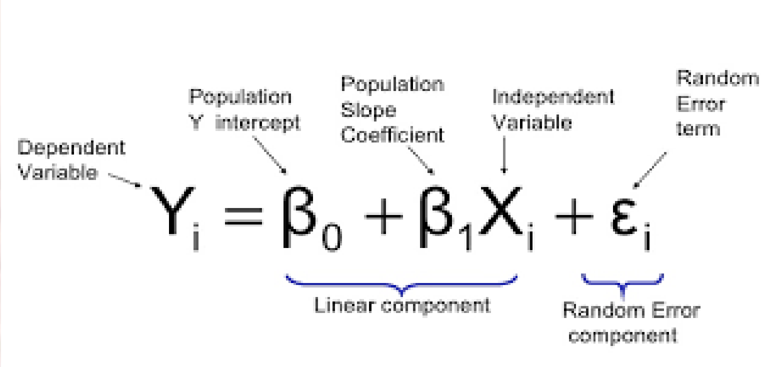
1. μ_i adalah variabel random yg menyebar normal
2. Nilai rata-rata $\mu_i = 0$, $e(\mu_i) = 0$.
3. Tidak tdpt serial korelasi antar μ_i $cov(\mu_i,\mu_j) = 0$
4. Sifat homoskedastistas, $var(\mu_i) = \sigma^2$
5. $Cov(\mu_i,X_i) = 0$
6. Tidak terdapat bias dalam spesifikasi model
7. Tidak terdapat multi-collinearity antar variebel penjelas

Model Regresi Sederhana

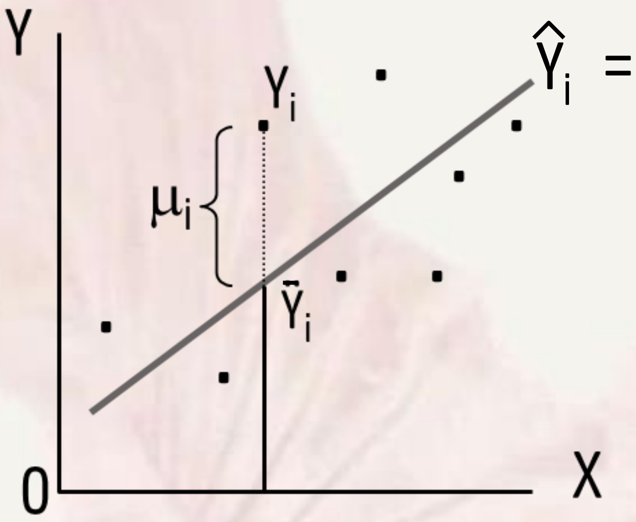
$$Y_i = \beta_0 + \beta_1 X_i + \mu_i$$

- β_0 dan β_1 : parameter dari fungsi yg nilainya akan diestimasi.
- Bersifat stochastik ⇨ untuk setiap nilai X terdapat suatu distribusi probabilitas seluruh nilai Y ***atau*** Nilai Y tidak dapat diprediksi secara pasti karena ada faktor stochastik μ_i yang memberikan sifat acak pada Y.



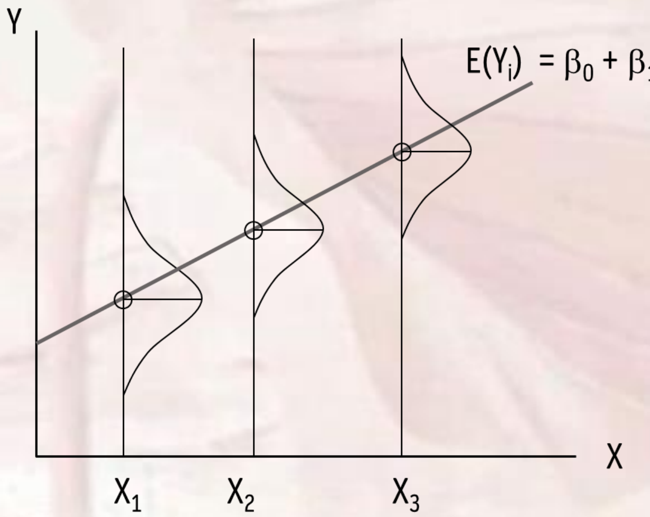


- Adanya variabel μ_i disebabkan karena:
  ① Ketidak-lengkapan teori
  ② Perilaku manusia yang bersifat random
  ③ Ketidak-sempurnaan spesifikasi model
  ④ Kesalahan dalam agregasi
  ⑤ Kesalahan dalam pengukuran



$\hat{Y}_i = b_0 + b_1 X_i$

$$\underbrace{Y_i}_{\text{Variation in Y}} = \underbrace{\beta_0 + \beta_1 X_i}_{\text{Systematic Variation}} + \underbrace{\mu_i}_{\text{Random Variation}}$$



$Y_i = \beta_0 + \beta_1 X_i + \mu_i$

Nilai rata2 Y_i ($\bar{Y}$) :

$E(Y_i) = \beta_0 + \beta_1 X_i$

$\mu_I = Y_i - E(Y_i)$

Zulkifli Alamsyah

Estimasi Parameter
Model Regresi Sederhana

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i$$

β_0 dan β_1 = parameter populasi
β_0 = intercept / konstanta
β_1 = Slope dari garis regresi

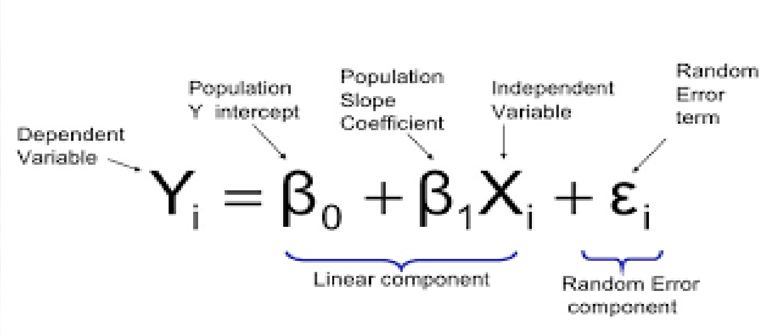
Method of Ordinary Least Squares

$$Y_i = \widehat{Y_i} + \widehat{u_i}$$

$$\widehat{u_i} = Y_i - \widehat{Y_i}$$

error or residual term $\widehat{u_i}$ is the difference between the actual Y_i and estimated $\widehat{Y_i}$

$$\widehat{u_i} = Y_i - (\widehat{\beta_1} + \widehat{\beta_2} X_i) = Y_i - \widehat{\beta_1} - \widehat{\beta_2} X_i$$



Metode Kuadrat Terkecil (Ordinary Least Square – OLS):

Prinsip: Meminimumkan nilai error – mencari jumlah penyimpangan kuadrat ($\Sigma\mu_i^2$) terkecil.

$$\mu_i = Y_i - \beta_0 - \beta_1 X_i$$
$$\mu_i^2 = (Y_i - \beta_0 - \beta_1 X_i)^2$$
$$\Sigma\mu_i^2 = \Sigma (Y_i - \beta_0 - \beta_1 X_i)^2$$

$\Sigma\mu_i^2$ minimum jika:

$$\partial\Sigma\mu_i^2 / \partial\beta_0 = 0 \Rightarrow 2\,\Sigma(Y_i - \beta_0 - \beta_1 X_i) = 0$$
$$\partial\Sigma\mu_i^2 / \partial\beta_1 = 0 \Rightarrow 2\,\Sigma X_i (Y_i - \beta_0 - \beta_1 X_i) = 0$$

Metode Ordinary Least Squares (OLS)

$Y_i = \beta_0 + \beta_1 X_i + \mu_i$ (1) Persamaan umum Regresi sederhana

$Y_i = \beta_0 + \beta_1 X_i + \mu_i$ (2) β_0 dan β_1 adalah nilai estimasi untuk parameter

$\hat{Y}_i = b_0 + b_1 X_i$ (3) $\hat{Y}i$ = nilai estimasi model

$Y_i = \hat{Y}_i + \mu_i$ (4) Y_i = nilai aktual

$\mu_i = Y_i - \hat{Y}_i$ (5) μ_i = nilai residual (error)

$$b_1 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{n \sum x_i y_i}{\sum x_i^2}$$

$$b_0 = \frac{\sum (X_i)^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \bar{Y} - b_1 \bar{X}$$

- b_0 dan b_1 adalah parameter stimasi untuk β_0 dan β_1
- $x = (X_i - \bar{X})$
- $y = (Y_i - \bar{Y})$

Zulkifli Alamsyah

Cara lain untuk memprediksi Nilai Parameter

$$b_1 = \frac{\Sigma (X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma (X_i - \bar{X})^2}$$

$$b_0 = \bar{Y} - b_1\bar{X}$$

Dimana:

b_0 dan b_1 nilai penduga untuk β_0 dan β_1.

$\bar{X}$ dan $\bar{Y}$ adlh nilai rata2 pengamatan X dan Y

Hasil estimasi: $\hat{Y} = b_0 + b_1 X$

Contoh: Perhatikan data pengamatan X dan Y sebagai berikut:

Obs.	1	2	3	4	5	6	7	8	9	10
X_i	2	4	6	8	10	12	14	16	18	20
Y_i	6	9	15	16	18	24	26	30	32	34

Dari data diatas (1) rumuskanlah model estimasi regresi linear pengaruh X terhadap Y, dan (2) apakah pengaruhnya signifikan?

Untuk merumuskan model estimasi regresi linear pengaruh X terhadap Y, hitunglah nilai parameter b_0 dan b_1 sehingga dapat dirumuskan model estimasi $\widehat{Y} = b_0 + b_1 X$

Gunakan Tabel bantu sebagai berikut untuk mendapatkan nilai setiap komponen dari formula untuk menghitung b_0 dan b_1 :

Obs.	X_i	Y_i	$x_i = (X_i - \bar{X})$	$y_i = (Y_i - \bar{Y})$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$
1	2	6	-9	-15	135	81
2	4	9	-7	-12	84	49
3	6	15	-5	-6	30	25
4	8	16	-3	-5	15	9
5	10	18	-1	-3	3	1
6	12	24	1	3	3	1
7	14	26	3	5	15	9
8	16	30	5	9	45	25
9	18	32	7	11	77	49
10	20	34	9	13	117	81
Jumlah	110	210	0	0	**524**	**330**
Rata-rata	**11**	**21**				

$\bar{X} = 11$

$\bar{Y} = 21$

$\Sigma(X_i - \bar{X})(Y_i - \bar{Y}) = 524$

$\Sigma(X_i - \bar{X})^2 = 330$

Gunakan formula sebagai berikut:

$$b_1 = \frac{\Sigma (X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma (X_i - \bar{X})^2}$$

$$b_0 = \bar{Y} - b_1\bar{X}$$

$$b_1 = \frac{524}{330} = 1{,}59$$

$$b_0 = 21 - 1{,}58(11) = 3{,}53$$

$$\boxed{\widehat{Y} = 3{,}53 + 1{,}59\,X}$$

Interpretasi:

Setiap kenaikan 1 unit X, maka Y akan naik sebesar 1,59 unit.

Apakah X berpengaruh signifikan terhadap Y?

Untuk menguji apakah hasil estimasi regresi linear sudah baik dan apakah pengaruh X terhadap Y signifikan, dilakukan langkah-langkah sebagai berikut:

1. Evaluasi Model

- Memastikan bahwa model yang kita bangun valid secara statistik, akurat secara prediksi, dan masuk akal secara ekonomi/ilmiah.
- Komponen penting dalam evaluasi model regresi linear sederhana Adalah:
  - **Uji Asumsi Klasik** (Sebagian)
  - **Koefisien Determinasi** (R^2)
  - **Uji Pengaruh Simultan** (Uji-F)

Uji Asumsi Klasik yang wajib untuk model regresi linear sederhana:

- Uji normalitas, jika jumlah sampel kecil
- Uji hetereskedastisitasi
- Uji autokorelasi (khusus untuk data timeseries)

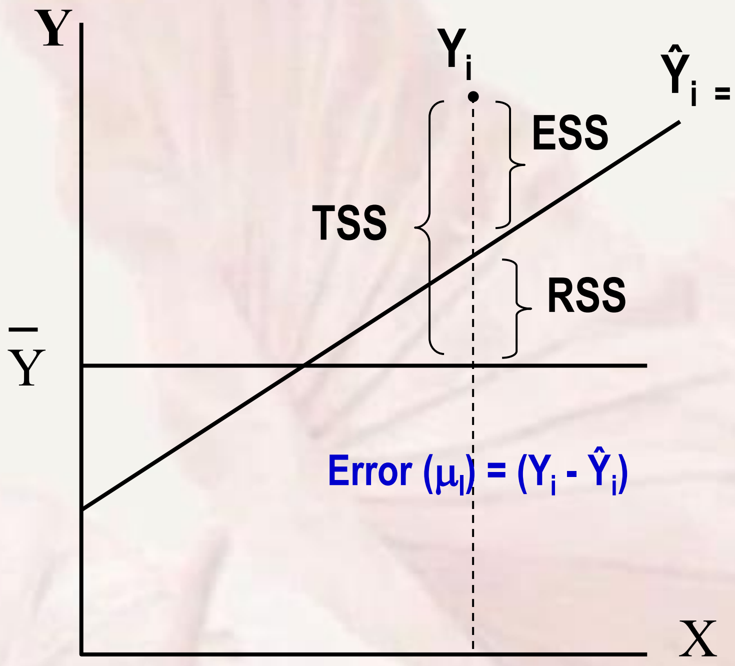
(Uji asumsi klasik akan dibahas secara khusus)

Koefisien Determinasi (R^2)

$$R^2 = 1 - \frac{ESS}{TSS}$$

- R^2 merupakan ukuran statistik dalam model regresi yang menunjukkan seberapa besar proporsi **variasi dalam variabel dependen yang dapat dijelaskan oleh variabel independen**
- R^2 berkisar antara 0 hingga 1, di mana semakin tinggi nilainya, menunjukkan bahwa model dapat menjelaskan variabilitas data dengan baik,
- R^2 tidak bisa langsung menunjukkan tingkat kesalahan prediksi.

Menghitung Koefisien Determinasi (R^2)



$$1 = \frac{RSS}{TSS} + \frac{ESS}{TSS} \quad \Rightarrow \quad \textbf{TSS = RSS + ESS}$$

$RSS = \sum(\hat{Y}_i - \bar{Y})^2$

$ESS = \sum \mu_i^2 = \sum(Y_i - \hat{Y}_i)^2$

$TSS = \sum(Y_i - \bar{Y})^2$

TSS = Total sum of squares
RSS = Residual sum of squares
ESS = Estimated/Explained sum of squares

$$1 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} + \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

$$\boxed{R^2 = \frac{RSS}{TSS}} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \Rightarrow R^2 = 1 - \frac{ESS}{TSS} = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

atau

$$R^2 = \beta_1^2 \left(\frac{\sum x_i^2}{\sum y_i^2}\right)$$

$$R^2 = \frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2}$$

Dimana:
$x_i = (X_i - \bar{X})$
$y_i = (Y_i - \bar{Y})$

Zulkifli Alamsyah

Contoh menghitung Koefisien Determinasi (R^2)

Obs.	X_i	Y_i	$x_i = (X_i - \bar{X})$	$y_i = (Y_i - \bar{Y})$	$x_i y_i = (X_i - \bar{X})(Y_i - \bar{Y})$	x_i^2	y_i^2
1	2	6	-9	-15	135	81	225
2	4	9	-7	-12	84	49	144
3	6	15	-5	-6	30	25	36
4	8	16	-3	-5	15	9	25
5	10	18	-1	-3	3	1	9
6	12	24	1	3	3	1	9
7	14	26	3	5	15	9	25
8	16	30	5	9	45	25	81
9	18	32	7	11	77	49	121
10	20	34	9	13	117	81	169
Jumlah (∑)	110	210	0	0	524	330	844
Rata-rata	11	21					

$$R^2 = \frac{(\sum x_i y_i)^2}{\sum x_i^2 \sum y_i^2}$$

$$R^2 = \frac{524^2}{(330)(884)} = 0{,}9858$$

atau:

$$R^2 = \beta_1^2 \left(\frac{\sum x_i^2}{\sum y_i^2} \right)$$

$$\beta_1 = 1{,}5879$$

$$R^2 = 1{,}5879^2 \left(\frac{330}{844} \right) = 0{,}9858$$

Contoh menghitung Koefisien Determinasi (R^2)

Obs.	X_i	Y_i	$\hat{Y}i = 3{,}53 + 1{,}59\,X$	$y_i = (Y_i - \bar{Y})$	TSS $= (Yi - \bar{Y})^2$	ESS $\mu^2 = (Y_i - \hat{Y}_i)^2$
1	2	6	6,71	-15	225,00	0,50
2	4	9	9,88	-12	144,00	0,78
3	6	15	13,06	-6	36,00	3,76
4	8	16	16,24	-5	25,00	0,06
5	10	18	19,41	-3	9,00	1,99
6	12	24	22,59	3	9,00	1,99
7	14	26	25,76	5	25,00	0,06
8	16	30	28,94	9	81,00	1,12
9	18	32	32,12	11	121,00	0,01
10	20	34	35,29	13	169,00	1,67
Jumlah (∑)	110	210	210,00	0	**844,00**	**11,95**
Rata-rata	**11**	**21**				

$$\boxed{R^2 = \frac{RSS}{TSS}} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

$$R^2 = 1 - \frac{ESS}{TSS} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2}$$

$$R^2 = 1 - \frac{11{,}95}{844} = 0{,}9858$$

Interpretasi:

98.58% variasi Y dapat dijelaskan oleh X, hanya 1.42% dijelaskan oleh faktor lain di luar model.

Zulkifli Alamsyah

Uji Simultan (Uji-F)

- Uji F pada regresi linear sederhana digunakan untuk menentukan apakah model regresi layak digunakan (signifikan) atau tidak
- Menguji apakah variabel independen (X) secara signifikan mempengaruhi variabel dependen (Y)
- Uji ini membandingkan nilai F_{hitung} dengan nilai F_{tabel} atau membandingkan dengan nilai signifikansi (Sig.) α yang digunakan, biasanya α = 5% atau 0,05.

Menghitung Nilai Statistik F

- SSR dan SSE digunakan untuk menghitung nilai statistik F.
- Dalam perhitungannya, SSR dan SSE disesuaikan dengan derajat bebas
- SSR dibagi jumlah variabel independen (k) diperoleh **Mean Square Regression (MSR)**
- SSE dibagi deraja bebas (n-k-1) diperoleh **Mean Square Error (MSE)**

$$MSR = \frac{RSS}{k} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{k}$$

$$MSE = \frac{ESS}{n - k - 1} = \frac{\sum(Y_i - \hat{Y}_i)^2}{n - 2}$$

Karena pada model regresi sederhana k=1, maka

MSR = RSS

MSE = ESS / (n-2)

$$F = \frac{MSR}{MSE} = \frac{RSS / k}{ESS / (n - k - 1)} = \frac{RSS}{ESS / (n-2)} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \hat{Y}_i)^2 / (n-2)}$$

TSS = RSS + ESS
TSS = 844
ESS = 11,95
RSS = 844 – 11,95 = 832,05
n = 10

$$F = \frac{RSS}{ESS / (n-2)} = \frac{832{,}05}{11{,}95 / (10-2)} = \mathbf{556{,}95}$$

Nilai statistik F yang besar membuktikan bahwa model regresi secara efektif menjelaskan variasi pada variabel dependen

Zulkifli Alamsyah

Uji Parsial (Uji-t)

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i$$

- Menguji apakah parameter (β_1) berbeda secara signifikan dari nol.
- Uji-t bertujuan untuk menjawab pertanyaan: "Apakah variabel X berpengaruh secara signifikan terhadap Y"?
- Hipotesis yang akan dibuktikan adalah sebagai berikut:

  Ho : $\beta_1 = 0$

  Ha : $\beta_1 \neq 0$ atau $\beta_1 > 0$ atau $\beta_1 < 0$ (Tergantung kesimpulan sementara)
- Alat uji yang digunakan:

$$t = \frac{\beta_1}{SE(\beta_1)} \qquad SE(\beta_1) = \sqrt{\frac{\sigma^2}{\sum(X_i - \bar{X})^2}} \qquad \sigma^2 = \frac{ESS}{(n-2)} \approx MSE$$

- Kaedah Keputusan:

  Tolak Ho jika t-hitung > t-tabel → ada pengaruh yang signifikan X terhadap Y

Zulkifli Alamsyah

Contoh:

ESS = 11,95

n = 10

$\sum(X_i - \bar{X})^2 = 330$

$\beta_1 = 1{,}5879$

$$\sigma^2 = \frac{ESS}{(n-2)} = \frac{11{,}95}{(10-2)} = 1{,}4939$$

$$SE(\beta_1) = \sqrt{\frac{\sigma^2}{\sum(X_i - \bar{X})^2}} = \sqrt{\frac{1{,}4939}{330}} = \mathbf{0.0673}$$

$$t = \frac{\beta_1}{SE(\beta_1)} = \frac{1{,}5879}{0{,}0673} = \mathbf{23.60}$$

t-tabel (α=0.05; df (n-2) = **1,860**

Kesimpulan: karena t-hitung > t-tabel maka disimpulkan bahwa X berpengaruh signifikan thdp Y

Cara lain untuk menyimpulkan apakah X berpengaruh signifikan terhadap Y adalah:

- menghitung nilai probabilitas (ρ-value) atau tingkat signifikansi secara numerik menggunakan EXCEL dengan formula: **=TDIST(NILAI t-hitung;DF;TAIL)**

  → **df = n–2**, **Tail = 1** jika uji satu arah dan **Tail = 2** jika uji dua arah
- Dengan cara tersebut diperoleh **ρ-value = 5,53E–9 atau 0,0000000553**

Tabel : **Distribusi Student's– t**

υ	α 0,25	0,20	0,15	0,10	0,05	0,025	0,01	0,005	0,0005
1	1,000	1,376	1,963	3,078	6,314	12,706	31,820	63,660	636,600
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,600
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,920
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,032	8,869
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,500	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,073	0,883	1,100	1,387	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,698	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,696	0,873	1,083	1,356	1,782	2,179	2,681	3,054	4,318

LATIHAN-1

Diasumsikan ada 2 variabel yang diduga variabel yang satu berpengaruh terhadap variabel lainnya, dengan bentuk persamaan umum sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \mu_i$$

Gunakan data Var-1 dan Var-2 (nama variabel dapat diubah) **sebanyak 12 pengamatan**, sajikan dalam Tabel, dan **gunakan formula** untuk:

(1) Mengestimasi nilai bo dan b1, lalu tulis dalam bentuk persamaan
(2) Menghitung koefisien determinasi
(3) Menghitung nilai statistik F (uji simultan)
(4) Menghitung nilai satistik t (uji parsial)

Buatlah interpretasi dan kesimpulan dari hasil-hasil peritungan diatas

Terima kasih

zalamsyah@unja.ac.id

Jika ada pertanyaan terkait materi yang disampaikan, dapat diajukan melalui kanal yang tersedia atau melalui forum diskusi pada setiap jadwal kuliah