

Analisa Regresi dan Korelasi

Tujuan Pembelajaran :

1. Dapat mengetahui bentuk hubungan antara dua variabel.
2. Dapat menghitung persamaan linear regresion
3. Dapat mengetahui ada tidaknya pengaruh
4. Dapat melakukan peramalan pada masa akan datang



Analisa Regresi dan Korelasi

- Beberapa contoh kasus :
 1. Kenaikan harga BBM akan mempengaruhi kenaikan harga angkutan kota.
 2. Pendapatan akan mempengaruhi konsumsi.
 3. Harga penjualan akan mempengaruhi jumlah permintaan.
- Sebutkan 5 contoh lagi kasus lainnya !

Analisa Regresi :

- Melihat bentuk (pola/model) hubungan antara dua variabel. Misal : var. kenaikan harga BBM dan var. kenaikan harga angkot
- Melihat ada tidaknya hubungan antar kedua variabel.
- Menduga atau membuat peramalan (*forecasting*) ke depan.



Analisa Regresi

```
graph TD; A[Analisa Regresi] --> B[Analisa Regresi Sederhana]; A --> C[Analisa Regresi Berganda]; B --> D[Analisa Regresi Linear Sederhana]; B --> E[Analisa Regresi Nonlinear Sederhana];
```

Analisa Regresi
Sederhana

Analisa Regresi
Berganda

Analisa Regresi
Linear Sederhana

Analisa Regresi
Nonlinear
Sederhana

Analisa Regresi Linear Sederhana

- Analisa Regesi terhadap dua variabel yaitu X dan Y
- Model (persamaan) Linearnya adalah :

$$Y_i = A + BX_i + e_i$$

- Model (persamaan) Linear Duganya adalah :

$$\hat{Y}_i = a + bX_i$$

Keterangan :

- Y_i : variabel terikat (*Dependent variable*)
- X_i : variabel bebas (*Independent variable*)
- A : Perpotongan pada sumbu X (*Intercept*)
- B : Kemiringan (*slope*) = Koefisien regresi
- \hat{Y}_i : penduga bagi Y_i
- a : penduga bagi A
- b : penduga bagi B
- e_i : error

$$e_i = Y_i - \hat{Y}_i$$

Metode Kuadrat Terkecil

- Suatu metode untuk menentukan persamaan regresi dengan meminimumkan jumlah kuadrat beda antara nilai Y dan nilai dugaanya \hat{Y} .

$$\sum ei^2 = \sum (Y_i - \hat{Y}_i)^2 = \text{minimum}$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$a = \bar{Y} - b\bar{X}$$

Sehingga persamaan linear dugaanya adalah :

$$\hat{Y}_i = a + bX_i$$

Contoh soal


Suatu sampel random sebanyak 10 keluarga bertujuan untuk melihat hubungan antara pengeluaran konsumsi (ribuan) dengan pendapatan keluarga (ribuan).

Keluarga	1	2	3	4	5	6	7	8	9	10
Pendapatan	80	100	120	140	160	180	200	220	240	260
Konsumsi	70	65	90	95	110	115	120	140	155	150

- Dugalah bentuk (persamaan) linear hubungan tersebut !
- Hitunglah *Standard error of estimate*-nya !
- Ujilah apakah pendapatan berpengaruh dalam peningkatan konsumsi !
- Interpretasikan koefisien regresi yang diperoleh !
- Hitunglah korelasi antara kedua variabel tersebut !

Jawaban a

- Pertama tentukan dahulu variabel mana yang menjadi variabel bebas dan variabel mana yang menjadi variabel terikat.
- Menurut logika ekonomi :
Variabel Pendapatan sebagai variabel bebas (X)
dan
Variabel Konsumsi sebagai variabel terikat (Y).

Pendapatan (X)  Konsumsi (Y)

Tabel data sampel Pendapatan dan Konsumsi

Pendapatan (X)	Konsumsi (Y)	X ²	XY	Y ²
80	70	6400	5600	4900
100	65	10000	6500	4225
120	90	14400	10800	8100
140	95	19600	13300	9025
160	110	25600	17600	12100
180	115	32400	20400	13225
200	120	40000	24000	14400
220	140	48400	30800	19600
240	155	57600	37200	24025
260	150	67600	39000	22500
1700	1110	322000	205500	132100

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$b = \frac{(10 \times 205500) - (1110 \times 1700)}{10(322000) - (1700)^2} = \mathbf{0,5091}$$

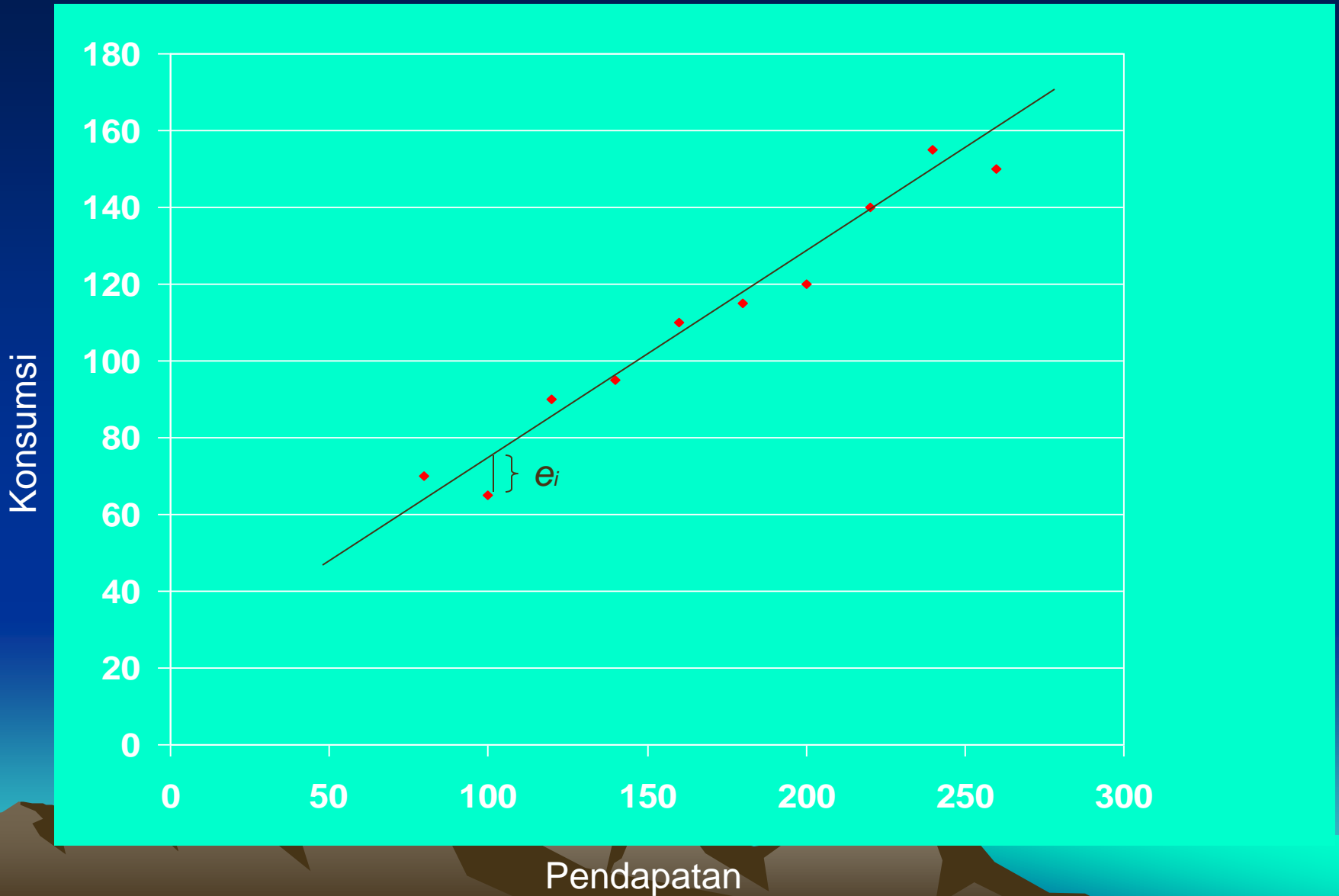
$$a = \bar{Y} - b\bar{X}$$

$$a = 111 - 0,5091(170) = \mathbf{24,454}$$

Sehingga persamaan regresi duganya adalah

$$\hat{Y} = 24,454 + 0,5091 X$$

PLOT DATA DAN GARIS REGRESI DUGA



Kesalahan Baku Pendugaan

- Kesalahan baku pendugaan (*Standar Error of Estimate*) merupakan ukuran tentang variasi titik-titik di sekitar garis regresi, sehingga dapat digunakan sebagai ukuran ketepatan pendugaan.

$$Se = \sqrt{\frac{\sum e^2}{n-2}} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n-2}}$$

$$Se = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n-2}}$$

Jawaban b

Pendapatan (X)	Konsumsi (Y)	$\hat{Y} = 24,45 + 0,51 X$	e (Y - \hat{Y})	e ²
80	70	65.18	4.82	23.2324
100	65	75.36	-10.36	107.3296
120	90	85.55	4.45	19.8025
140	95	95.73	-0.73	0.5329
160	110	105.9	4.1	16.81
180	115	116.1	-1.1	1.21
200	120	126.2	-6.27	39.3129
220	140	136.4	3.54	12.5316
240	155	146.6	8.36	69.8896
260	150	156.82	-6.82	46.5124
				337.1639

Jawaban b

$$Se = \sqrt{\frac{\sum e^2}{n-2}} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n-2}} = \sqrt{337.1639 / (10-2)} = 6,49$$

atau

$$Se = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n-2}}$$

$$= \sqrt{\frac{132100 - 24,454 (1110) - 0,51 (205500)}{10-2}} = \sqrt{\frac{337,273}{8}} = 6,49$$

$$\hat{Y}_0 = 24,45 + 0,51(100) = 75,36$$

Meramal \hat{Y}_0 pada $X = X_0$

- Pendugaan Titik bagi : $\hat{y}_0 = a + bx_0$
- Contoh :
- Dugalah konsumsi yang terjadi \hat{Y}_0 jika pendapatannya Rp 100.000,- ($X_0 = 100$) ?

$$\hat{Y}_0 = 24,45 + 0,51(100) = 75,36$$

- *Maka konsumsi yang terjadi adalah Rp 75.360,-*

$$\mu_{Y|X_0} = \alpha + \beta X_0$$

Meramal $E(\hat{Y}_0) = \mu \hat{Y}_0 = \mu_{Y|X_0}$

Pendugaan Interval

bagi : $\mu_{Y|X_0} = \alpha + \beta X_0$

adalah

$$\hat{y}_0 - t_{\alpha/2} \times Se \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{(n-1)s_x^2}} < \mu_{Y|X_0} < \hat{y}_0 + t_{\alpha/2} \times Se \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{(n-1)s_x^2}}$$

$$t_{\alpha/2 ; (n-2)}$$

Meramal

$$E(\hat{Y}_0) = \mu\hat{Y}_0 = \mu_{Y|X_0}$$

Contoh :

Buatlah Pendugaan Interval bagi rata-rata populasi konsumsi jika pendapatannya Rp 100.000,- pada tingkat keyakinan 95% !

$$t_{0,05/2} (10-2) = 2,306$$

$$75,36 \pm 2,306 \times 6,49 \sqrt{\frac{1}{10} + \frac{(100 - 170)^2}{(10 - 1)(60,553)^2}} = 75,36 \pm 7,46$$

$$67,9 \leq \mu_{Y/100} \leq 82,82 \quad (\text{dalam ribuan})$$

$$\text{Rp } 67.900 \leq \mu_{Y/100.000} \leq \text{Rp } 82.820$$

Uji Hipotesis Koefisien Regresi

> Uji 2 arah :

Ho : $B = 0$ (var. X tidak berpengaruh thd var Y)

Hi : $B \neq 0$ (var. X berpengaruh thd var. Y)

> Uji 1 arah kanan :

Ho : $B \leq 0$ (var. X tidak berpengaruh thd var Y)

Hi : $B > 0$ (var. X berpengaruh positif thd var. Y)

> Uji 1 arah kiri :

Ho : $B \geq 0$ (var. X tidak berpengaruh thd var Y)

Hi : $B < 0$ (var. X berpengaruh negatif thd var. Y)

Statistik Uji

$$t_h = \frac{b - 0}{S_b} = \frac{b}{S_b}$$

Dimana :

$$S_b = \frac{S_e}{\sqrt{n(\sum X^2) - (\sum X)^2}}$$

Jawaban c

1. Type Hipotesis :

$$H_0 : B \leq 0$$

$$H_1 : B > 0$$

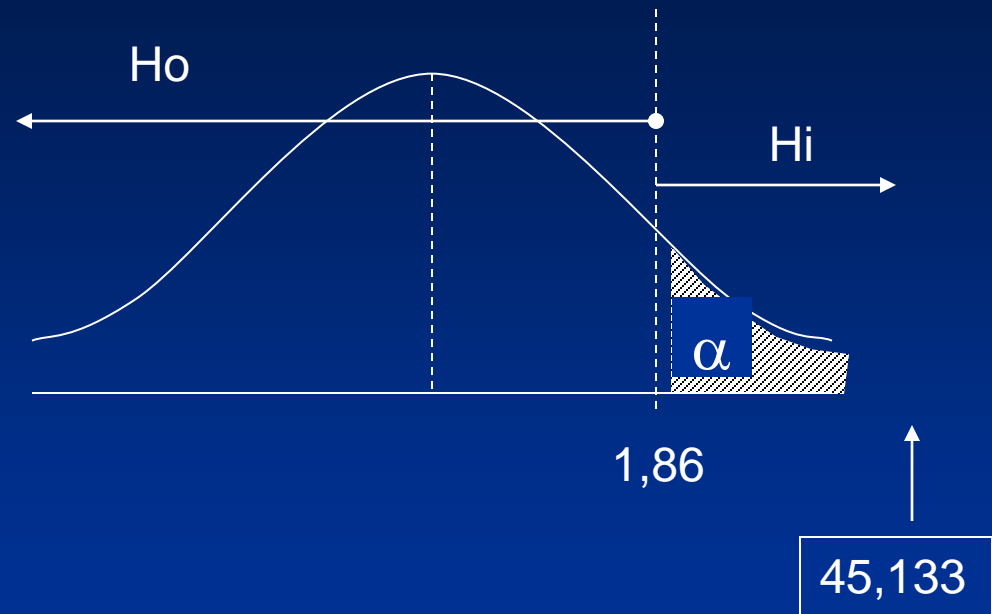
2. Batas kritis : $\alpha = 0,05$

$$t_{0,05, db = 10 - 2} = \mathbf{1,86}$$

3. Statistik Uji t_h :

$$S_b = \frac{S_e}{\sqrt{n(\sum X^2) - (\sum X)^2}} = \frac{6,49}{\sqrt{10(322000) - (1700)^2}} = 0,0113$$

$$t_h = b / S_b = 0,51 / 0,0113 = \mathbf{45,133}$$



4. $t_h = 45,133 > t_{0,05(8)} = 1,86$ terima $H_1 : B > 0$

5. Pendapatan berpengaruh positif terhadap konsumsi

Analisa Korelasi

- Melihat ada tidaknya hubungan antara dua variabel.
- Melihat nilai keeratan hubungan itu.

$$r^2 = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

Korelasi : $r = \sqrt{r^2}$ Dimana $-1 < r < 1$

$r = 0$: tidak ada korelasi

$|r| < 0,5$: korelasi lemah

$0,5 < |r| < 0,75$: korelasi sedang

$0,75 < |r| < 0,9$: korelasi kuat

$0,9 < |r| < 1$: korelasi sangat kuat

$|r| = 1$: berkorelasi sempurna

Jawaban d dan e

d. Interpretasi dari $b = 0,51$ adalah

Setiap kenaikan Rp 1000 pendapatan akan meningkatkan konsumsi sebesar 0,51 (ribuan) atau Rp 510.

e. Korelasi antara Pendapatan dan Konsumsi adalah :

$$r^2 = \frac{10(205500) - (1700)(1110)}{\sqrt{[10(322000) - (1700)^2][10(132100) - (1110)^2]}} = 0,98$$

$r = \sqrt{r^2} = \sqrt{0,98} = 0,99$: berkorelasi sangat kuat

Soal latihan 1

- Tn. Ilham seorang manager pemasaran produsen sepeda motor mengamati bahwa ada suatu bentuk hubungan antara jumlah penduduk atau populasi disuatu wilayah dengan nilai penjualan sepeda motor. Dengan data sample random berikut ini bantulah Pak Ilham dalam menemukan model hubungan tersebut.

<u>Wilayah</u>	<u>Populasi (Ribuan)</u>	<u>Nilai Penjualan (Ratusan juta Rp)</u>
• 1	36	5,4
• 2	26	3,0
• 3	12	2,8
• 4	40	4,8
• 5	24	3,6

- Pertanyaan:
 - a. Carilah persamaan regresi duga untuk melihat bentuk hubungan tersebut. Gunakan variable nilai penjualan sebagai variable terikat dan variable populasi sebagai variable bebas.
 - b. Interpretasi koefisien regresi populasi yang anda peroleh !
 - c. Hitunglah *Standard Error of Estimate* dari model yang anda peroleh !
 - d. Berapa perkiraan nilai penjualan sepeda motor pada wilayah yang berpopulasi 50 000 penduduk ?

Soal latihan 2

- Tn. Fadil manager pemasaran sebuah perusahaan asuransi sedang meneliti hubungan antara Penjualan polis yang dihasilkan seorang sales agen dengan pengalamannya (lamanya) menjadi seorang sales agen. Bila diperoleh data sample sbb :

Pengalaman (tahun)	0	1	2	3	4	5
Penjualan Polis (unit/bl)	5	10	12	15	18	20

- a. Dugalah model hubungan kedua variable di atas dengan menggunakan model regresi linear sederhana.
- b. Hitunglah kasalahan baku (*standard error of estimate*)
- c. Ujilah koefisien regresinya, apakah berpengaruh positif ?
- d. Berapa penjualan polis yang diharapkan, jika pengalaman sbg sales agen asuransi selama 6 tahun ?

-