

**ANALISIS REGRESI DAN KORELASI LINIER**

1

---

---

---

---

---

---

---

---

**OUTLINE**

**Bagian I Statistik Induktif**

- Metode dan Distribusi Sampling
- Teori Pendugaan Statistik
- Pengujian Hipotesa Sampel Besar
- Pengujian Hipotesa Sampel Kecil
- Analisis Regresi dan Korelasi Linier**
  - Pengertian Korelasi Sederhana
  - Uji Signifikansi Koefisien Korelasi
  - Analisis Regresi: Metode Kuadrat Terkecil
  - Kesalahan Baku Pendugaan
  - Asumsi-asumsi Metode Kuadrat Terkecil
  - Perkiraan Interval dan Pengujian Hipotesa
  - Hubungan Koefisien Korelasi, Koefisien Determinasi dan Kesalahan Baku Pendugaan
- Analisis Regresi dan Korelasi Berganda
- Fungsi, Variabel, dan Masalah dalam Analisis Regresi

2

---

---

---

---

---

---

---

---

**PENGERTIAN ANALISIS KORELASI**

**Analisis Korelasi**

Suatu teknik statistika yang digunakan untuk mengukur keeratan hubungan atau korelasi antara dua variabel.

3

---

---

---

---

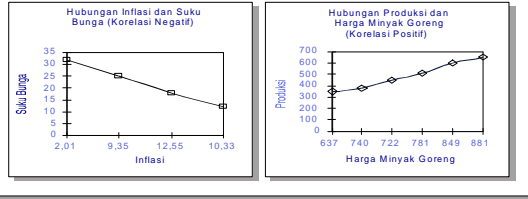
---

---

---

---

**HUBUNGAN POSITIF DAN NEGATIF**



Gambar pertama menunjukkan hubungan antara variabel inflasi dan suku bunga. Apabila dilihat pada gambar saat inflasi rendah, maka suku bunga tinggi dan pada saat inflasi tinggi, suku bunga rendah. Gambar tersebut menunjukkan adanya hubungan antara inflasi dan suku bunga yang bersifat negatif.

Gambar kedua memperlihatkan hubungan yang positif antara variabel produksi dan harga minyak goreng yaitu apabila harga meningkat, maka produksi juga meningkat.

---

---

---

---

---

---

---

---

---

---

---

---

**RUMUS KOEFISIEN KORELASI**

Rumus koefisien korelasi tersebut dinyatakan sebagai berikut:

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

Di mana:

- r : Nilai koefisien korelasi
- $\sum X$  : Jumlah pengamatan variabel X
- $\sum Y$  : Jumlah pengamatan variabel Y
- $\sum XY$  : Jumlah hasil perkalian variabel X dan Y
- $(\sum X^2)$  : Jumlah kuadrat dari pengamatan variabel X
- $(\sum X)^2$  : Jumlah kuadrat dari jumlah pengamatan variabel X
- $(\sum Y^2)$  : Jumlah kuadrat dari pengamatan variabel Y
- $(\sum Y)^2$  : Jumlah kuadrat dari jumlah pengamatan variabel Y
- n : Jumlah pasangan pengamatan Y dan X

---

---

---

---

---

---

---

---

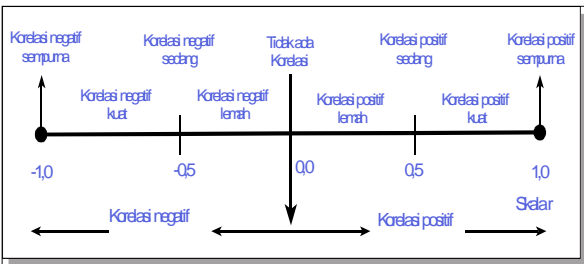
---

---

---

---

**HUBUNGAN KUAT DAN LEMAHNYA SUATU KORELASI**




---

---

---

---

---

---

---

---

---

---

---

---

**CONTOH: PERMINTAAN DIPENGARUHI HARGA DAN PENDAPATAN**

Tahun	Investasi (milliar)	Suku bunga (%/th)
1994	34.285	19,25
1995	43.141	17,75
1996	50.825	18,88
1997	57.399	19,21
1998	74.873	21,98
1999	31.180	32,27
2000	28.897	28,89
2001	38.056	18,43
2002	45.962	19,19

7

---

---

---

---

---

---

---

---

---

---

**CONTOH: PERMINTAAN DIPENGARUHI HARGA DAN PENDAPATAN**

Rumus koefisien korelasi

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

n	Y	X	X <sup>2</sup>	XY	Y <sup>2</sup>
1	34.285	19,25	371	659.986	1.175.461.225
2	43.141	17,75	315	765.753	1.861.145.881
3	50.825	18,88	356	959.576	2.583.180.625
4	57.399	19,21	369	1.102.635	3.294.645.201
5	74.873	21,98	483	1.645.709	5.605.966.129
6	31.180	32,27	1041	1.006.179	972.192.400
7	28.897	28,89	835	834.834	835.036.609
8	38.056	18,43	340	701.372	1.448.259.136
9	45.962	19,19	368	882.011	2.112.505.444
Jumlah	404.618	196	4478	8.558.054	19.888.392.650

$$r = \frac{9 \times 8558054 - 196 \times 404618}{\sqrt{[9(4478) - (196)^2][9(19888392650) - (404618)^2]}} = -0,412$$

8

---

---

---

---

---

---

---

---

---

---

**PENGERTIAN KOEFISIEN DETERMINASI**

Koefisien determinasi

Bagian dari keragaman total variabel tak bebas Y (variabel yang dipengaruhi atau dependent) yang dapat diterangkan atau diperhitungkan oleh keragaman variabel bebas X (variabel yang mempengaruhi atau independent).

Koefisien determinasi = r<sup>2</sup>

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

9

---

---

---

---

---

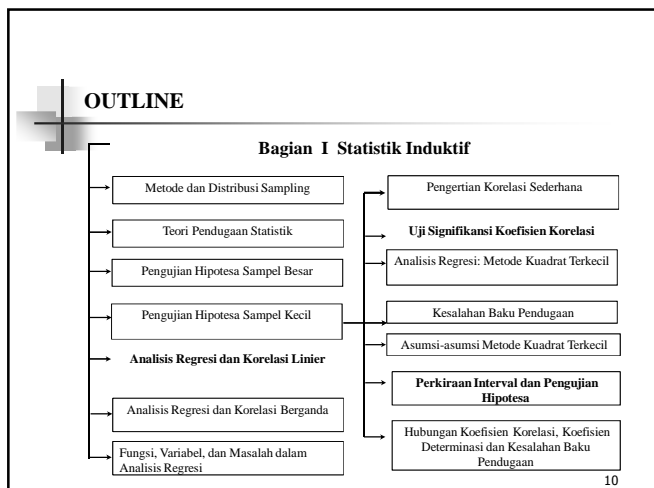
---

---

---

---

---




---

---

---

---

---

---

---

---

### RUMUS UJI t UNTUK UJI KORELASI

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{atau} \quad t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

**Di mana:**

- t : Nilai t-hitung
- r : Nilai koefisien korelasi
- n : Jumlah data pengamatan

---

---

---

---

---

---

---

---

### CONTOH UJI t UNTUK UJI KORELASI SOAL A

Ujilah apakah (a) nilai  $r = -0,412$  pada hubungan antara suku bunga dan investasi dan (b)  $r = 0,86$  pada hubungan antara harga minyak dan produksi kelapa sawit sama dengan nol pada taraf nyata 5%?

- Perumusan hipotesa:  
 Hipotesa yang diuji adalah koefisien korelasi sama dengan nol. Korelasi dalam populasi dilambangkan dengan  $\rho$  sedang pada sampel  $r$ .  
 $H_0 : \rho = 0$   
 $H_1 : \rho \neq 0$
- Taraf nyata 5% untuk uji dua arah ( $\alpha/2 = 0,025$ ) dengan derajat bebas (df) =  $n - k = 9 - 2 = 7$ . Nilai taraf nyata  $\alpha/2 = 0,025$  dan df = 7 adalah = 2,36. Ingat bahwa n adalah jumlah data pengamatan yaitu = 9, sedangkan k adalah jumlah variabel yaitu Y dan X, jadi k=2.
- Menentukan nilai uji t

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{-0,41}{\sqrt{\frac{1-(,041)^2}{9-2}}} = -1,21$$

---

---

---

---

---

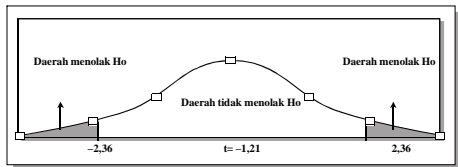
---

---

---

**CONTOH UJI t UNTUK UJI KORELASI**

4. Menentukan daerah keputusan dengan nilai kritis 2,36



5. Menentukan keputusan. Nilai t-hitung ternyata terletak pada daerah tidak menolak  $H_0$ . Ini menunjukkan bahwa tidak terdapat cukup bukti untuk menolak  $H_0$ , sehingga dapat disimpulkan bahwa korelasi dalam populasi sama dengan nol, hubungan antara tingkat suku bunga dengan investasi lemah dan tidak nyata.

---

---

---

---

---

---

---

---

**CONTOH UJI t UNTUK UJI KORELASI SOAL B**

1. Perumusan hipotesa:

Hipotesa yang diuji adalah koefisien korelasi sama dengan nol. Korelasi dalam populasi dilambangkan dengan  $\rho$  dan sampel  $r$ .

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

2. Taraf nyata 5% untuk uji dua arah ( $\alpha/2=0,05/2=0,025$ ) dengan derajat bebas (df) =  $n-k - 2 = 10$ . Nilai taraf nyata  $\alpha/2=0,025$  dan  $df=10$  adalah = 2,23.

3. Menentukan nilai uji t

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0,86}{\sqrt{\frac{1-(0,86)^2}{12-2}}} = 5,33$$

---

---

---

---

---

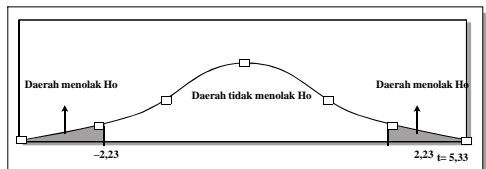
---

---

---

**RUMUS KOEFISIEN DETERMINASI**

4. Menentukan daerah keputusan dengan nilai kritis 2,23



5. Menentukan keputusan. Nilai t-hitung berada di daerah menolak  $H_0$ , yang berarti bahwa  $H_0$  ditolak dan menerima  $H_1$ . Ini menunjukkan bahwa koefisien korelasi pada populasi tidak sama dengan nol, dan ini membuktikan bahwa terdapat hubungan yang kuat dan nyata antara harga minyak dan produksi kelapa sawit.

---

---

---

---

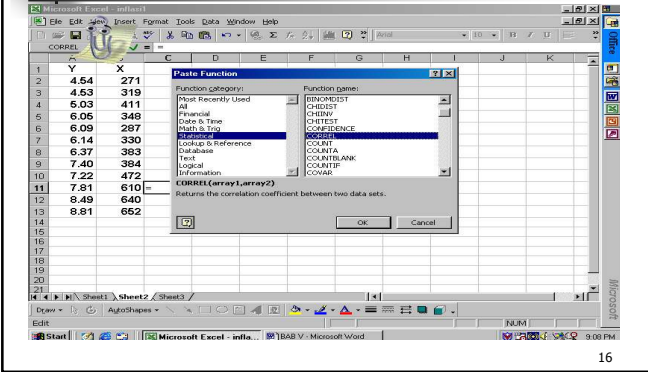
---

---

---

---

## MENGGUNAKAN MS EXCEL UNTUK Mencari KORELASI



---

---

---

---

---

---

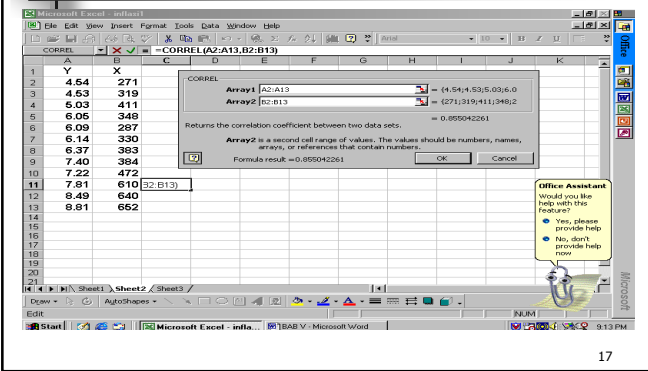
---

---

---

---

## MENGGUNAKAN MS EXCEL UNTUK Mencari KORELASI



---

---

---

---

---

---

---

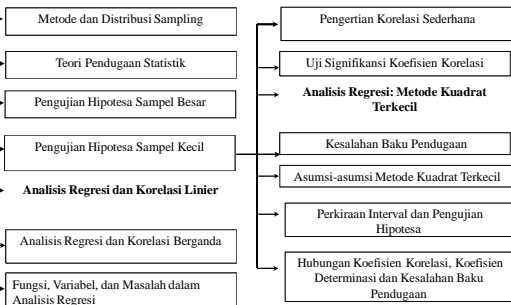
---

---

---

## OUTLINE

### Bagian I Statistik Induktif



---

---

---

---

---

---

---

---

---

---

**RUMUS PERSAMAAN REGRESI**

**Persamaan regresi**

Suatu persamaan matematika yang mendefinisikan hubungan antara dua variabel.

---

---

---

---

---

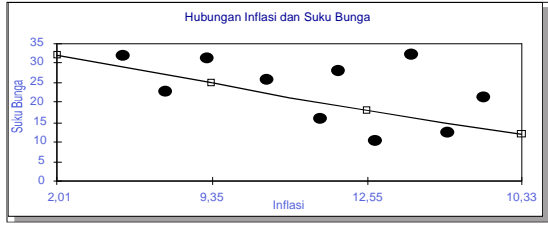
---

---

---

**SCATTER DIAGRAM UNTUK MEMBANTU MENARIK GARIS REGRESI**

Scatter diagram untuk hubungan antara inflasi dan suku bunga dapat digambarkan sebagai berikut:



Gambar A

---

---

---

---

---

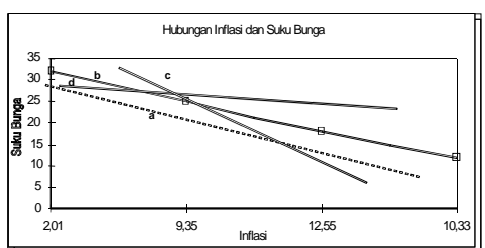
---

---

---

**SCATTER DIAGRAM UNTUK MEMBANTU MENARIK GARIS REGRESI**

Scatter diagram untuk hubungan antara inflasi dan suku bunga dapat digambarkan sebagai berikut:



Gambar B

---

---

---

---

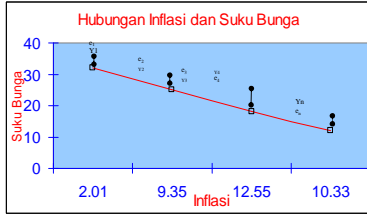
---

---

---

---

**CONTOH SELISIH ANTARA DUGAAN DAN AKTUAL LEBIH KECIL**



Gambar A: selisih antara dugaan dan aktual lebih kecil

---

---

---

---

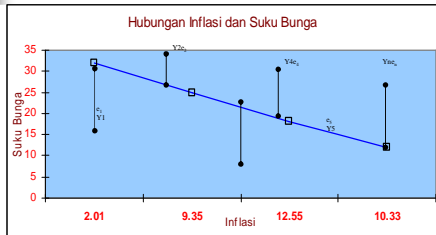
---

---

---

---

**CONTOH SELISIH ANTARA DUGAAN DAN AKTUAL LEBIH BESAR**




---

---

---

---

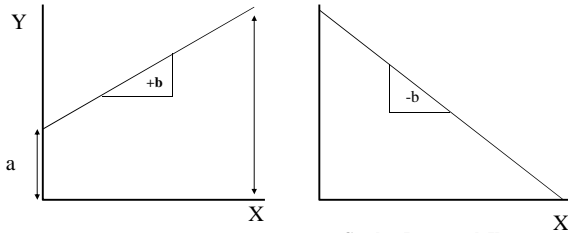
---

---

---

---

**GAMBAR PERSAMAAN REGRESI**



Gambar A:  $\hat{y} = a + b X$

Gambar B:  $\hat{y} = a - b X$

---

---

---

---

---

---

---

---



**RUMUS Mencari Koefisien a dan b**

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

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

- Y : Nilai variabel bebas Y
- a : Intersep yaitu titik potong garis dengan sumbu Y
- b : Slope atau kemiringan garis yaitu perubahan rata-rata pada  $\hat{Y}$  untuk setiap unit perubahan pada variabel X
- X : Nilai variabel bebas X
- n : Jumlah sampel

---

---

---

---

---

---

---

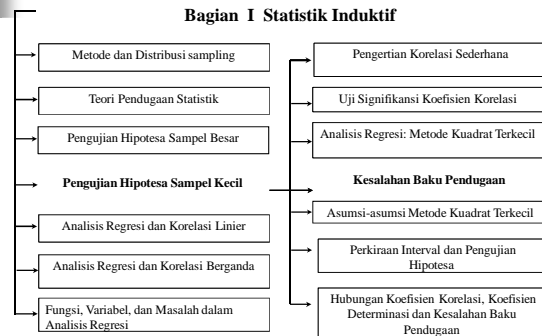
---

---

---

**OUTLINE**

**Bagian I Statistik Induktif**




---

---

---

---

---

---

---

---

---

---

**CONTOH HUBUNGAN ANTARA PRODUKSI DENGAN HARGA MINYAK KELAPA SAWIT**

$$= a + b X$$

Yi	X	= 2,8631 + 0,0086 X		e=Y-
4,54	271	= 2,8631 + 0,0086 x 271	5,1853	-0,6453
4,53	319	= 2,8631 + 0,0086 x 319	5,5966	-1,0666
5,03	411	= 2,8631 + 0,0086 x 411	6,3850	-1,3550
6,05	348	= 2,8631 + 0,0086 x 348	5,8451	0,2049
6,09	287	= 2,8631 + 0,0086 x 287	5,3224	0,7676
6,14	330	= 2,8631 + 0,0086 x 330	5,6909	0,4491
6,37	383	= 2,8631 + 0,0086 x 383	6,1450	0,2250
7,40	384	= 2,8631 + 0,0086 x 384	6,1536	1,2464
7,22	472	= 2,8631 + 0,0086 x 472	6,9077	0,3123
7,81	610	= 2,8631 + 0,0086 x 610	8,0902	-0,2802
8,49	640	= 2,8631 + 0,0086 x 640	8,3473	0,1427

---

---

---

---

---

---

---

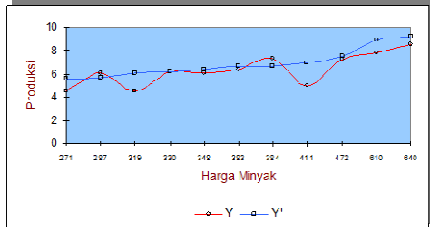
---

---

---

**CONTOH HUBUNGAN ANTARA PRODUKSI DENGAN HARGA MINYAK KELAPA SAWIT**

Persamaan  $\hat{y} = 2,8631 + 0,0086 X$ .



Gambar A: Koordinat antara Y dan  $\hat{y}$

---

---

---

---

---

---

---

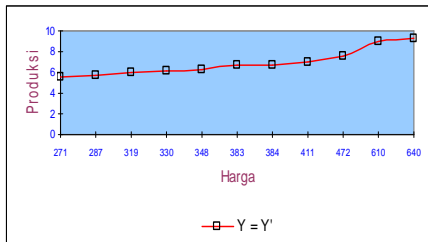
---

---

---

**CONTOH HUBUNGAN ANTARA PRODUKSI DENGAN HARGA MINYAK KELAPA SAWIT**

Persamaan  $\hat{Y} = 2,8631 + 0,0086 X$ .



Gambar B: Koordinat antara Y dan  $\hat{y}$ , dimana  $Y = \hat{y}$

---

---

---

---

---

---

---

---

---

---

**DEFINISI**

**Standar error atau kesalahan baku Pendugaan**

Suatu ukuran yang mengukur ketidakakuratan pencaran atau persebaran nilai-nilai pengamatan (Y) terhadap garis regresinya ( $\hat{y}$ ).

---

---

---

---

---

---

---

---

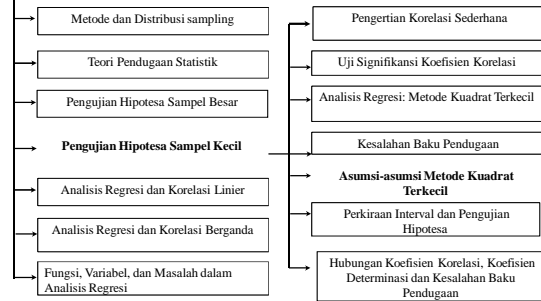
---

---



OUTLINE

Bagian I Statistik Induktif




---

---

---

---

---

---

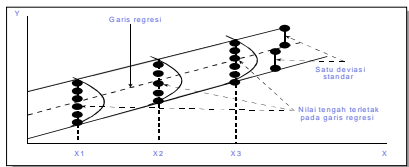
---

---

ASUMSI METODE KUADRAT TERKECIL

Beberapa asumsi penting metode kuadrat terkecil adalah sebagai berikut:

1. Nilai rata-rata dari *error term* atau *expected value* untuk setiap nilai X sama dengan nol. Asumsi ini dinyatakan  $E(e_i/X_i) = 0$ .
2. Nilai error dari  $E_i$  dan  $E_j$  atau biasa disebut dengan kovarian saling tidak berhubungan atau berkorelasi. Asumsi ini biasa dilambangkan sebagai berikut,  $Cov(E_i, E_j) = 0$ , di mana  $i \neq j$ . Berdasarkan pada asumsi nomor 1, pada setiap nilai  $X_i$  akan terdapat  $E_i$ , dan untuk  $X_j$  akan ada  $E_j$ , yang dimaksud dengan nilai kovarian = 0 adalah nilai  $E_i$  dari  $X_i$  tidak ada hubungan dengan nilai  $E_j$  dari  $X_j$ .




---

---

---

---

---

---

---

---

ASUMSI METODE KUADRAT TERKECIL

3. Varian dari error bersifat konstan. Ingat bahwa varian dilambangkan dengan  $s^2$ , sehingga asumsi ini dilambangkan dengan  $Var(E_i/E_j) = E(e_i - e_j)^2 = s^2$ . Anda perhatikan pada gambar di atas bahwa nilai  $E_i$  (yang dilambangkan dengan tanda titik) untuk setiap X yaitu  $X_1$ ,  $X_2$  dan  $X_3$  tersebar secara konstan sebesar variannya yaitu  $s^2$ . Pada gambar tersebut nilai E tersebar 1 standar deviasi di bawah garis regresi dan 1 standar deviasi di atas garis regresi. Seluruh sebaran nilai  $E_i$  untuk  $X_i$  dan  $E_j$  untuk  $X_j$ , di mana  $i \neq j$  terlihat sama dengan ditunjukkan kurva yang berbentuk simetris dengan ukuran yang sama, hal inilah yang dikenal dengan varians dari error bersifat konstan.
4. Variabel bebas X tidak berkorelasi dengan error term E, ini biasa dilambangkan dengan  $Cov(E_i, X_i) = 0$ . Pada garis regresi  $Y = a + bX_i + e_i$  maka nilai  $X_i$  dan  $E_i$  tidak saling mempengaruhi, sebab apabila saling mempengaruhi maka pengaruh masing-masing yaitu X dan E tidak saling dapat dipisahkan. Ingat bahwa yang mempengaruhi Y selain X adalah pasti E yaitu faktor diluar X. Oleh sebab itu varians dari E dan X saling terpisah atau tidak berkorelasi.

---

---

---

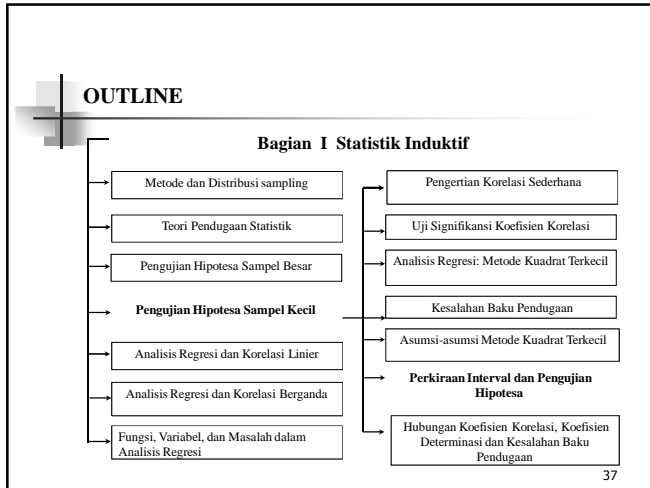
---

---

---

---

---




---

---

---

---

---

---

---

---

### RUMUS

$$\hat{Y} \pm t(S_{yx}) \sqrt{\frac{1}{n} + \frac{(X - \bar{X})^2}{\sum X^2 - (\sum X)^2 / n}}$$

$\hat{Y}$  : Nilai dugaan dari Y untuk nilai X tertentu  
 $t$  : Nilai t-tabel untuk taraf nyata tertentu  
 $S_{yx}$  : Standar error variabel Y berdasarkan variabel X yang diketahui  
 $X$  : Nilai data pengamatan variabel bebas  
 $\bar{X}$  : Nilai rata-rata data pengamatan variabel bebas  
 $n$  : Jumlah sampel

38

---

---

---

---

---

---

---

---

### PENDUGAAN INTERVAL NILAI KOEFISIEN REGRESI A DAN B

Dengan menggunakan asumsi bahwa nilai  $E_i$  bersifat normal, maka hasil dugaan a dan b juga mengikuti distribusi normal. Sehingga nilai  $t = (b - B)/\sigma_b$ , juga merupakan variabel normal. Dalam praktiknya nilai standar deviasi populasi  $\sigma_b$  sulit diketahui, maka standar deviasi populasi biasa diduga dengan standar deviasi sampel yaitu  $S_b$ , sehingga nilai  $t$  menjadi  $t = (b - B)/S_b$ . Selanjutnya probabilitasnya dinyatakan sebagai berikut:

$$P(-t_{\alpha/2} \leq (b - B)/S_b \leq t_{\alpha/2}) = 1 - \alpha$$

$$P(-t_{\alpha/2} \cdot S_b \leq (b - B) \leq t_{\alpha/2} \cdot S_b) = 1 - \alpha$$

Sehingga interval B adalah:  
 $(b - t_{\alpha/2} \cdot S_b \leq B \leq b + t_{\alpha/2} \cdot S_b)$   
 sedangkan dengan cara yang sama interval A adalah:  
 $(a - t_{\alpha/2} \cdot S_a \leq A \leq a + t_{\alpha/2} \cdot S_a)$

di mana  $S_a$  dan  $S_b$  adalah sebagai berikut:  
 $S_b = S_{y.x} / \sqrt{\sum X^2 - (\sum X)^2 / n}$   
 $S_a = \sqrt{(\sum X^2 \cdot S_{y.x}) / (n \sum X^2 - (\sum X)^2)}$

39

---

---

---

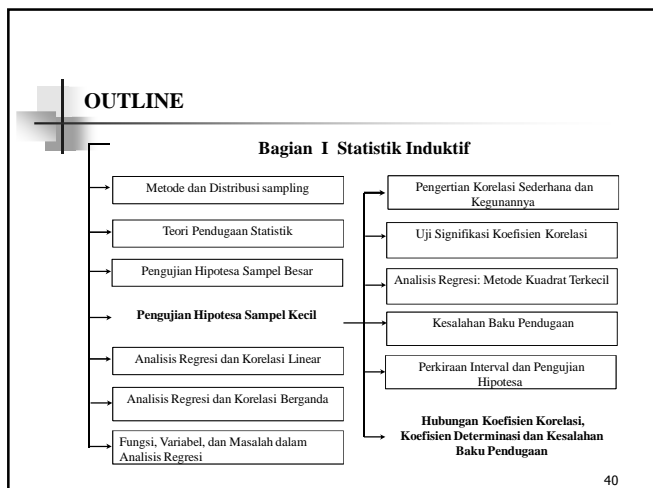
---

---

---

---

---




---

---

---

---

---

---

---

---

### ANALISIS VARIANS ATAU ANOVA

#### Analisis Varians atau ANOVA

Merupakan alat atau peranti yang dapat menggambarkan hubungan antara koefisien korelasi, koefisien determinasi dan kesalahan baku pendugaan. Untuk mengukur kesalahan baku kita menghitung error yaitu selisih Y dengan  $\hat{Y}$  atau dapat dinyatakan dalam bentuk persamaan:

$$e = Y - \hat{Y}$$

atau dalam bentuk lain yaitu

$$Y = \hat{Y} + e$$

Di mana

- $Y$  adalah nilai sebenarnya,
- $\hat{Y}$  adalah nilai regresi
- $e$  adalah error atau kesalahan

41

---

---

---

---

---

---

---

---

### TABEL ANOVA

Sumber Keragaman (Source)	Derajat bebas (df)	Sum Square (SS)	Mean Square (MS)
Regresi (Regression)	1 (jumlah var bebas, X)	$SSR = \sum(\hat{Y} - Y)^2$	$MSR = SSR/1$
Kesalahan (error)	n-2	$SSE = \sum(Y - \hat{Y})^2$	$MSE = SSE/(n-2)$
Total	n-1	$SST = \sum(Y - \bar{Y})^2$	

42

---

---

---

---

---

---

---

---