7

Analisis Korelasi dan Regresi

Analisis Korelasi dan Regresi

Pengantar

- Gagasan perhitungan ditetapkan oleh Sir Francis Galton (1822-1911).
- Persamaan regresi :Persamaan matematik yang memungkinkan peramalan nilai suatu peubah takbebas (*dependent variable*) dari nilai peubah bebas (*independent variable*).
- Regresi dan korelasi digunakan untuk mempelajari pola dan mengukur hubungan statistik antara dua atau lebih variabel.
- Jika digunakan hanya dua variabel disebut regresi dan korelasi sederhana.
- Jika digunakan lebih dari dua variabel disebut regresi dan korelasi berganda.

PENGERTIAN KORELASI

Analisis Korelasi adalah Analisis statistika yang memanfaatkan hubungan antara dua atau lebih variable sehingga dapat diukur keeratannya.

Teknik Statistika → Hubungan → Antar Variabel

Pengertian Korelasi

- Variabel yang akan diduga disebut variabel terikat (tidak bebas) atau **dependent variable**, biasa dinyatakan dengan **variabel Y**.
- Variabel yang menerangkan perubahan variabel terikat disebut variabel bebas atau **independent variable**, biasa dinyatakan dengan **variabel X**.
- Analisa korelasi digunakan untuk mengukur keeratan hubungan antara variabel-variabel tersebut.

Jenis-jenis Persamaan Regresi

- Regresi Linier :
  - Regresi Linier Sederhana
  - Regresi Linier Berganda

- Regresi Nonlinier
  - Regresi Eksponensial

Jenis-jenis Persamaan Regresi

- Regresi Linier :
- Bentuk Umum **Regresi Linier Sederhana**

$$Y = a + bX$$

- Y : peubah takbebas
- X : peubah bebas
- a : konstanta
- b : kemiringan

JENIS-JENIS PERSAMAAN REGRESI

- Regresi Linier :
- Bentuk Umum **Regresi Linier Sederhana**

$$Y = a + bX$$

Y : peubah takbebas
X : peubah bebas
a : konstanta
b : kemiringan

- Bentuk Umum **Regresi Linier Berganda**

$$Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n$$

Y : peubah takbebas	a : konstanta
X_1 : peubah bebas ke-1	b_1 : kemiringan ke-1
X_2 : peubah bebas ke-2	b_2 : kemiringan ke-2
X_n : peubah bebas ke-n	b_n : kemiringan ke-n

REGRESI LINIER SEDERHANA

- Metode Kuadrat terkecil (*least square method*): metode paling populer untuk menetapkan persamaan regresi linier sederhana
- Bentuk Umum **Regresi Linier Sederhana**

$$\mathbf{Y = a + bX}$$

Y : peubah takbebas

X : peubah bebas

a : konstanta

b : kemiringan

- Penetapan Persamaan Regresi Linier Sederhana

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$a = \bar{y} - b\bar{x}$$

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\frac{\sum_{i=1}^{n} x_i}{n}$$

n : banyak pasangan data

y_i : nilai peubah takbebas Y ke-i

x_i : nilai peubah bebas X ke-i

Regresi Linier Sederhana (Contoh)

Tahun	Biaya Promosi (Juta Rupiah) (x)	Volume Penjualan (Ratusan Juta Liter) (y)	xy	x^2	y^2
1992	2	5	10	4	25
1993	4	6	24	16	36
1994	5	8	40	25	64
1995	7	10	70	49	100
1996	8	11	88	64	121
Σ	**Σx = 26**	**Σy = 40**	**Σxy = 232**	**Σx^2 =158**	**Σy^2 = 346**

$$n = 5$$

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad b = \frac{(5\times 232)-(26\times 40)}{(5\times 158)-(26^2)} = \frac{1160-1040}{790-676} = \frac{120}{114} = 1.0526 = 1.053$$

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\frac{\sum_{i=1}^{n} x_i}{n}$$

$$a = \frac{40}{5} - \left(1.05263\ldots\times\frac{26}{5}\right) = 8-(1.05263\ldots\times 5.2) = 8-5.4736\ldots = 2.5263\ldots = 2.530$$

$$\boxed{Y = a + b\,X \quad \rightarrow \quad Y = 2.530 + 1.053\,X}$$

Regresi Linier Sederhana (Contoh)

Diketahui hubungan Biaya Promosi (X dalam Juta Rupiah) dan Y (Volume penjualan dalam Ratusan Juta liter) dapat dinyatakan dalam persamaan regresi linier pada contoh sebelumnya yaitu:

$$\mathbf{Y = 2{,}530 + 1{,}053\,X}$$

Perkirakan Volume penjualan jika dikeluarkan biaya promosi Rp. 10 juta?

Jawab:

$Y = 2{,}530 + 1{,}053\,X$

Jika $X = 10$, maka

$Y = 2{,}530 + 1{,}053\,(10) = 2{,}530 + 10{,}53$

Y = 13,06 (ratusan juta liter)

Analisis Korelasi Linier Sederhana

- **ANALISA KORELASI** digunakan untuk mengukur kekuatan keeratan hubungan antara dua variabel melalui sebuah bilangan yang disebut **koefisien korelasi**.
- **Koefisien korelasi linier (r)** adalah ukuran hubungan linier antara dua variabel/peubah acak X dan Y, dengan nilai antara 0 – 1.
- Bila dua peubah tidak berhubungan; korelasinya 0
- Bila sempurna korelasinya 1 (korelasinya linier)
- **Koefisien Determinasi** Sampel = R = r^2

  Ukuran proporsi keragaman total nilai peubah Y yang dapat dijelaskan oleh nilai peubah X melalui hubungan linier.

Model persamaan Korelasi Linier Sederhana

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}}$$

$$R = r^2$$

MODEL PERSAMAAN KORELASI LINIER SEDERHANA (CONTOH)

Tahun	Biaya Promosi (Juta Rupiah) (x)	Volume Penjualan (Ratusan Juta Liter) (y)	xy	x^2	y^2
1992	2	5	10	4	25
1993	4	6	24	16	36
1994	5	8	40	25	64
1995	7	10	70	49	100
1996	8	11	88	64	121
Σ	$\Sigma x = 26$	$\Sigma y = 40$	$\Sigma xy = 232$	$\Sigma x^2 = 158$	$\Sigma y^2 = 346$

Dari contoh sbelumnya, setelah mendapatkan persamaan Regresi Y = 2,530 + 1,053 X, hitung koef. korelasi (r) dan koef determinasi (R).

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}}$$

$$R = r^2$$

MODEL PERSAMAAN KORELASI LINIER SEDERHANA (CONTOH)

$$r = \frac{(5\times 232)-(26\times 40)}{\sqrt{\left[(5\times 158)-(26^2)\right]\times\left[(5\times 346)-(40^2)\right]}} = \frac{1160-1040}{\sqrt{[790-676]\times[1730-1600]}}$$

$$= \frac{120}{\sqrt{114\times 130}} = \frac{120}{\sqrt{14820}} = \frac{120}{121.73\ldots} = 0.9857\ldots$$

Nilai r = 0.9857 menunjukkan bahwa peubah X (biaya promosi) dan Y (volume penjualan) berkorelasi **linier yang positif dan tinggi**

$R = r^2 = 09857^2 = 0.97165 = 97\%$

Nilai R = 97% menunjukkan bahwa 97% proporsi keragaman nilai peubah Y (volume penjualan) dapat dijelaskan oleh nilai peubah X (biaya promosi) melalui hubungan linier. Sisanya, yaitu 3% dijelaskan oleh hal-hal lain.

REGRESI LINIER BERGANDA

- Pembahasan akan meliputi regresi linier dengan 2 atau lebih Variabel Bebas (X_1 , X_2 , dan X_n) dan 1 Variabel Tak Bebas (Y)

- Bentuk Umum **Regresi Linier Berganda**

$$Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n$$

Y : peubah takbebas	a : konstanta
X_1 : peubah bebas ke-1	b_1 : kemiringan ke-1
X_2 : peubah bebas ke-2	b_2 : kemiringan ke-2
X_n : peubah bebas ke-n	b_n : kemiringan ke-n

- Untuk regresi linier berganda dengan 2 Variabel Bebas (X_1 dan X_2) dan 1 Variabel Tak Bebas (Y).

$$Y = a + b_1X_1 + b_2X_2$$

REGRESI LINIER BERGANDA

$$Y = a + b_1X_1 + b_2X_2$$

a , b1 dan b2 didapatkan dengan menyelesaikan tiga persamaan Normal berikut:

$$\text{(i)} \quad n\,a + b_1\sum_{i=1}^{n} x_{1i} + b_2\sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i$$

$$\text{(ii)} \quad a\sum_{i=1}^{n} x_{1i} + b_1\sum_{i=1}^{n} x_{1i}^{\;2} + b_2\sum_{i=1}^{n} x_{2i}x_{1i} = \sum_{i=1}^{n} x_{1i}y_i$$

$$\text{(iii)} \quad a\sum_{i=1}^{n} x_{2i} + b_1\sum_{i=1}^{n} x_{2i}x_{1i} + b_2\sum_{i=1}^{n} x_{2i}^{\;2} = \sum_{i=1}^{n} x_{2i}y_i$$

n : banyak pasangan data

x_{1i} : nilai peubah bebas X1 ke-i

y_i : nilai peubah takbebas Y ke-i

x_{2i} : nilai peubah bebas X2 ke-i

REGRESI LINIER BERGANDA

Berikut adalah data Volume Penjualan (juta unit) Mobil dihubungkan dengan variabel biaya promosi (X_1 dalam juta rupiah/tahun) dan variabel biaya penambahan asesoris (X_2 dalam ratusan ribu rupiah/unit).

x_1	x_2	y	$x_1\ x_2$	x_1y	x_2y	x_1^2	x_2^2	y^2
2	3	4	6	8	12	4	9	16
3	4	5	12	15	20	9	16	25
5	6	8	30	40	48	25	36	64
6	8	10	48	60	80	36	64	100
7	9	11	63	77	99	49	81	121
8	10	12	80	96	120	64	100	144
$\sum x = 31$	$\sum x = 40$	$\sum y = 50$	$\sum x_1x_2 = 239$	$\sum x_1\ y = 296$	$\sum x_2y = 379$	$\sum x_1^2 = 187$	$\sum x_2^2 = 306$	$\sum y = 470$

Tetapkan Persamaan Regresi Linier Berganda $Y = a + b_1 X_1 + b_2 X_2$

REGRESI LINIER BERGANDA

Masukkan notasi-notasi ini dalam ketiga persamaan normal,

(i) $$n\,a + b_1\sum_{i=1}^{n} x_{1i} + b_2\sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i$$

(ii) $$a\sum_{i=1}^{n} x_{1i} + b_1\sum_{i=1}^{n} x_{1i}^{\,2} + b_2\sum_{i=1}^{n} x_{2i}x_{1i} = \sum_{i=1}^{n} x_{1i}y_i$$

(iii) $$a\sum_{i=1}^{n} x_{2i} + b_1\sum_{i=1}^{n} x_{2i}x_{1i} + b_2\sum_{i=1}^{n} x_{2i}^{\,2} = \sum_{i=1}^{n} x_{2i}y_i$$

Sehingga didapatkan tiga persamaan berikut:

(i)	6a	+	31 b_1	+	40 b_2	=	50
(ii)	31 a	+	187 b_1	+	239 b_2	=	296
(iii)	40 a	+	239 b_1	+	306 b_2	=	379

Selesaikan persamaan diatas dengan cara subtitusi atau eliminasi :
Sehingga didapat nilai a, b1 dan b2

a = 0,95 b1 = 0,5 b2 = 0,75,

Sehingga Persamaan Regresi Berganda

Y = a + b1 X1 + b2 X2

Y =0.75 + 0.50 X1 + 0.75 X2

KORELASI LINIER BERGANDA

Koefisien Determinasi Sampel untuk Regresi Linier Berganda diberi notasi sebagai berikut:

$$R^2_{y.12}$$

Sedangkan Koefisien Korelasi adalah akar positif Koefisien Determinasi atau

$$r_{y.12} = \sqrt{R^2_{y.12}}$$

Model

$$R^2_{y.12} = 1 - \frac{JKG}{(n-1)s_y^2}$$

JKG : Jumlah Kuadrat Galat

s_y^2 : Jumlah Kuadrat y (terkoreksi)

di mana

$$s_y^2 = \frac{n\sum y^2 - \left(\sum y\right)^2}{n(n-1)}$$

$$JKG = \sum y^2 - a\sum y - b_1\sum x_1 y - b_2 \sum x_2 y$$