Regresi & Korelasi Linier Sederhana

1. Pendahuluan

- 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*)

- Diagram Pencar = *Scatter Diagram*

  Diagram yang menggambarkan nilai-nilai observasi peubah takbebas dan peubah bebas.

  Nilai peubah bebas ditulis pada sumbu X (sumbu horizontal)
  Nilai peubah takbebas ditulis pada sumbu Y (sumbu vertikal)

  Nilai peubah takbebas ditentukan oleh nilai peubah bebas

Anda sudah dapat menentukan mana peubah takbebas dan peubah bebas?

Contoh 1:

Umur Vs Tinggi Tanaman	(X : Umur, Y : Tinggi)
Biaya Promosi Vs Volume penjualan	(X : Biaya Promosi, Y : Vol. penjualan)

- Jenis-jenis Persamaan Regresi : a. Regresi Linier :
  - Regresi Linier Sederhana
  - Regresi Linier Berganda

  b. Regresi Nonlinier
  - Regresi Eksponensial

- 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 Non Linier
  - Bentuk umum Regresi Eksponensial

$$Y = ab^x$$

$$\log Y = \log a + (\log b)\, x$$

2. Regresi Linier Sederhana

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

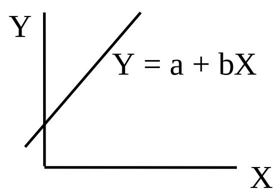
  - Bentuk Umum Regresi Linier Sederhana :

$$Y = a + bX$$

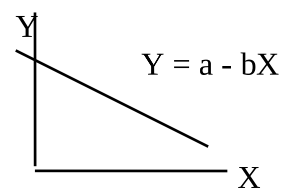
Y	: peubah takbebas	X	: peubah bebas
a	: konstanta	b	: kemiringan

Nilai b dapat positif (+) dapat negartif (-)

b : positif $\rightarrow$



b : negatif $\rightarrow$



- Penetapan Persamaan Regresi Linier Sederhana

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

$$a = \bar{y} - b\bar{x} \qquad \text{sehingga} \qquad a = \frac{\sum_{i=1}^{n} y_i}{n} - b\,\frac{\sum_{i=1}^{n} x_i}{n}$$

n : banyak pasangan data

y_i : nilai peubah takbebas Y ke-i
x_i : nilai peubah bebas X ke-i

Contoh 2 :

Berikut adalah data Biaya Promosi dan Volume Penjualan PT BIMOIL perusahaan Minyak Goreng.

Tahun	x Biaya Promosi (Juta Rupiah)	y Volume Penjualan (Ratusan Juta Liter)	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

bentuk umum persaman regresi linier sederhana : Y = a + b X

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

$$b = \frac{(5\times 232)-(26\times 40)}{(5\times 158)-(26^2)} = \frac{1160-1040}{790-676} = \frac{120}{114} = 1.05263\ldots = 1.053$$

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

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

Y = a + b X → Y = 2.530 + 1.053 X

- Peramalan dengan Persamaan Regresi

Contoh 3 :

Diketahui hubungan Biaya Promosi (X dalam Juta Rupiah) dan Y (Volume penjualan dalam Ratusan Juta liter) dapat dinyatakan dalam persamaan regresi linier berikut
$Y = 2.530 + 1.053 X$

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

Jawab : $Y = 2.530 + 1.053 X$
$X = 10$

$Y = 2.53 + 1.053 (10) = 2.53 + 10.53 = 13.06$ (ratusan juta liter)

Volume penjualan = 13.06 x 100 000 000 liter

3. Korelasi Linier Sederhana

- Koefisien Korelasi (r) : ukuran hubungan linier peubah X dan Y
  Nilai r berkisar antara (+1) sampai (-1)
  Nilai r yang (+) ditandai oleh nilai b yang (+)
  Nilai r yang (-) ditandai oleh nilai b yang (-)

  Jika nilai r mendekati +1 atau r mendekati -1 maka
  X dan Y memiliki korelasi linier yang tinggi

  Jika nilai r = +1 atau r = -1 maka X dan Y memiliki korelasi linier sempurna

  Jika nilai r = 0 maka X dan Y tidak memiliki relasi (hubungan) linier
  (dalam kasus r mendekati 0, anda dapat melanjutkan analisis ke regresi eksponensial)

- Koefisien Determinasi Sampel = $R = r^2$
  Ukuran proporsi keragaman total nilai peubah Y yang dapat dijelaskan oleh nilai peubah X melalui hubungan linier.

- Penetapan & Interpretasi Koefisien Korelasi dan Koefisien Determinasi

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

$$R = r^2$$

Contoh 4 :

Lihat Contoh 2, setelah mendapatkan persamaan Regresi Y = 2.530 + 1.053 X, hitung koef. korelasi (r) dan koef determinasi (R).

Gunakan data berikut (lihat Contoh 2)

$\Sigma x = 26$ $\Sigma y = 40$ $\Sigma xy = 232$ $\Sigma x^2 = 158$ $\Sigma y^2 = 346$

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

$$r = \frac{(5\times 232) - (26\times 40)}{\sqrt{\left[(5\times 158) - (26^2)\right]\times\left[(5\times 346) - (40^2)\right]}} = \frac{1160 - 1040}{\sqrt{[790 - 676]\times[1730 - 1600]}} = \frac{120}{\sqrt{114\times 130}}$$

$$= \frac{120}{\sqrt{14820}} = \frac{120}{121.73...} = 0.9857...$$

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

$R = r^2 = 0.9857...^2 = 0.97165.... = 97\ \%$

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.

4. Regresi Linier Berganda

- Pembahasan akan meliputi regresi linier dengan 2 Variabel Bebas (X_1 dan X_2) dan 1 Variabel Tak Bebas (Y).

- Bentuk Umum : $Y = a + b_1 X_1 + b_2 X_2$

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

- a , b_1 dan b_2 didapatkan dengan menyelesaikan tiga persamaan Normal berikut:

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

n : banyak pasangan data	y_i : nilai peubah takbebas Y ke-i
x_{1i} : nilai peubah bebas X_1 ke-i	x_{2i} : nilai peubah bebas X_2 ke-i

Contoh 4:

Berikut adalah data Volume Penjualan (juta unit) Mobil dihubungkan dengan variabel biaya promosi (X_1 dalam juta rupiah/tahun) dan variabel biaya penambahan aksesoris (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
$\sum x_1 =$ 31	$\sum x_2 =$ 40	$\sum y =$50	$\sum x_1 x_2 =$ 239	$\sum x_1 y =$ 296	$\sum x_2 y =$ 379	$\sum x_1^2 =$ 187	$\sum x_2^2 =$ 306	$\sum y^2 =$ 470

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

n = 6

$\sum x_1 = 31$ $\qquad \sum x_2 = 40$ $\qquad \sum y = 50$

$\sum x_1 x_2 = 239$ $\qquad \sum x_1 y = 296$ $\qquad \sum x_2 y = 379$

$\sum x_1^2 = 187$ $\qquad \sum x_2^2 = 306$ $\qquad \sum y^2 = 470$

Masukkan notasi-notasi ini dalam ketiga persamaan normal,

(i) $\quad n\,\mathrm{a} + \mathrm{b}_1 \sum_{i=1}^{n} x_{1i} + \mathrm{b}_2 \sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i$

(ii) $\quad \mathrm{a} \sum_{i=1}^{n} x_{1i} + \mathrm{b}_1 \sum_{i=1}^{n} x_{1i}^2 + \mathrm{b}_2 \sum_{i=1}^{n} x_{2i} x_{1i} = \sum_{i=1}^{n} x_{1i} y_i$

(iii) $\quad \mathrm{a} \sum_{i=1}^{n} x_{2i} + \mathrm{b}_1 \sum_{i=1}^{n} x_{2i} x_{1i} + \mathrm{b}_2 \sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i} y_i$

Sehingga didapatkan tiga persamaan berikut:

(i) $\quad 6a + 31\, b_1 + 40\, b_2 = 50$

$$(ii)\quad 31a + 187b_1 + 239b_2 = 296$$
$$(iii)\quad 40a + 239b_1 + 306b_2 = 379$$

Lakukan Eliminasi, untuk menghilangkan (a)

$$\begin{array}{llllll}
(ii) & 31a + 187b_1 & + & 239b_2 & = 296 & \times 6 \\
(i) & 6a + 31b_1 & + & 40b_2 & = 50 & \times 31
\end{array}$$

$$\begin{array}{lllll}
(ii) & \cancel{189a} + 1122b_1 & + & 1434b_2 & = 1776 \\
(i) & \cancel{189a} + 961b_1 & + & 1240b_2 & = 1550 \\
\hline
(iv) & 161b_1 & + & 194b_2 & = 226
\end{array}$$

Lalu

$$\begin{array}{llllll}
(iii) & 40a + 239b_1 & + & 306b_2 & = 379 & \times 6 \\
(i) & 6a + 31b_1 & + & 40b_2 & = 50 & \times 40
\end{array}$$

$$\begin{array}{lllll}
(iii) & \cancel{240a} + 1434b_1 & + & 1836b_2 & = 2274 \\
(i) & \cancel{240a} + 1240b_1 & + & 1600b_2 & = 2000 \\
\hline
(v) & 194b_1 & + & 236b_2 & = 274
\end{array}$$

Selanjutnya, eliminasi (b_1) dan dapatkan nilai (b_2)

$$\begin{array}{llllll}
(v) & 194b_1 & + & 236b_2 & = 274 & \times 161 \\
(iv) & 161b_1 & + & 194b_2 & = 226 & \times 194
\end{array}$$

$$\begin{array}{lllll}
(v) & \cancel{31234b_1} & + & 37996b_2 & = 44114 \\
(iv) & \cancel{31234b_1} & + & 37636b_2 & = 43844 \\
\hline
 & & & 360b_2 & = 270 \\
 & & & b_2 & = 0.75
\end{array}$$

Dapatkan Nilai (b_1) dan nilai (a) dengan melakukan substitusi, sehingga:

$$(v)\quad 194b_1 + 236b_2 = 274$$

Perhatikan $b_2 = 0.75$

$$194\ b_1 + 236\ (0.75) = 274$$
$$194\ b_1 + 177 = 274$$
$$194\ b_1 = 97$$
$$b_1 = 0.50$$

(i) $6a + 31\ b_1 + 40\ b_2 = 50$

Perhatikan $b_1 = 0.50$ dan $b_2 = 0.75$

$$6a + 31(0.50) + 40\ (0.75) = 50$$
$$6a + 15.5 + 30 = 50$$
$$6a = 4.5$$
$$a = 0.75$$

Sehingga Persamaan Regresi Berganda

$a + b_1 X_1 + b_2 X_2$ dapat ditulis sebagai $0.75 + 0.50\ X_1 + 0.75\ X_2$

5. 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

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

- Rumus

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

JKG : Jumlah Kuadrat Galat
s_y^2 : Jumlah Kuadrat y (terkoreksi)

di mana

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

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

Contoh 5:

Jika diketahui (dari Contoh 4)
n = 6

$\sum x_1 = 31$	$\sum x_2 = 40$	$\sum y = 50$
$\sum x_1 x_2 = 239$	$\sum x_1 y = 296$	$\sum x_2 y = 379$
$\sum x_1^{\,2} = 187$	$\sum x_2^{\,2} = 306$	$\sum y^2 = 470$

Maka tetapkan $R_{y.12}^2$ dan jelaskan artinya nilai tersebut!

$$s_y^2 = \frac{n\sum y^2 - \left(\sum y\right)^2}{n(n-1)} = \frac{6(470)-(50)^2}{6(6-5)} = \frac{2820-2500}{30} = \frac{320}{30} = 10.667$$

$JKG = \sum y^2 - a\sum y - b_1\sum x_1 y - b_2\sum x_2 y$ = 470 - 0.75(50) - 0.5 (296) - 0.75 (379)
= 470 - 37.5 - 148 - 284.25
= 0.25

$$R_{y.12}^2 = 1 - \frac{JKG}{(n-1)s_y^2} = 1 - \frac{0.25}{5\times 10.667} = 1 - \frac{0.25}{53.333}$$

= 1 - 0.0046875

= 0.9953125

= 99.53%

Nilai $R^2_{y.12}$ = 99.53% menunjukkan bahwa 99.53% proporsi keragaman nilai peubah Y (volume penjualan) dapat dijelaskan oleh nilai peubah X (biaya promosi) dan X_2 (biaya aksesoris) melalui hubungan linier.

Sisanya sebesar 0.47% dijelaskan oleh hal-hal lain.

⌖ *Selesai* ⌖