

REGRESI DAN KORELASI

Pada bab sebelumnya kita telah mempelajari analisi data mengenai satu variabel yang menggambarkan satu kejadian. Dalam kehidupan, kita sering menemui kejadian – kejadian yang saling berhubungan satu sama lain. Oleh karena itu kita memerlukan analisis antara kejadian – kejadian tersebut. Berikut adalah contoh – contoh kejadian yang saling berhubungan.

1. Permintaan suatu produk berhubungan dengan harga produk tersebut.
2. Peningkatan suatu produk dipengaruhi oleh tingkat pendapatan masyarakat
3. Berat badan seseorang berkaitan dengan tinggi badan, dsb.

Kali ini kita akan mengukur dua kejadian yang saling berhubungan secara matematis. Ada dua hal yang perlu diukur dan dianalisis:

1. Bagaimana hubungan fungsional dua kejadian tersebut (persamaan matematis yang menghubungkan/ menggambarkan hubungan diantara dua kejadian) -> **analisis regresi**
2. Bagaimana kekuatan atau keeratan hubungan tersebut -> **analisis korelasi**

REGRESI LINEAR SEDERHANA

Garis regresi (regresi) adalah garis lurus atau garis linear yang merupakan garis taksiran atau perkiraan untuk mewakili pola hubungan antara variabel X dan Y. Persamaan umum regresi adalah:

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

Keterangan: Y = nilai taksiran untuk variabel tak bebas Y
X = nilai – nilai variabel bebas
a = intersep
b = koefisien arah garis regresi

Variabel X sering disebut sebagai prediktor, yaitu variabel yang dipakai untuk memprediksi nilai Y, sedangkan variabel Y disebut variabel yang diprediksi atau variabel terikat. Mengingat kita dapat membuat banyak garis regresi, maka ada kriteria, garis regresi yang paling akurat, yaitu garis regresi yang mempunyai total kuadrat kesalahan atau total kuadrat error yang paling minimum.

Beberapa rumus yang digunakan untuk menyelesaikan persamaan regresi:

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

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

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

KESALAHAN BAKU DARI PENAKSIRAN $Y = a + bX$

$$S_{\hat{y}_x} = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n}}$$

KOEFISIEN KORELASI

Koefisien korelasi adalah akar dari koefisien determinasi dan ditulis:

$$r = \sqrt{\frac{\sum (\hat{Y} - \bar{Y})^2}{\sum (Y - \bar{Y})^2}}$$

Koefisien korelasi terletak antara -1 dan 1, yaitu $-1 \leq r \leq 1$:

1. Nilai $r = -1$ disebut korelasi linear negatif (berlawanan arah); artinya terdapat hubungan negatif yang sempurna antara X dan Y
2. Nilai $r = 1$ disebut korelasi linear positif (searah), artinya terdapat hubungan positif yang sempurna antara variabel X dan Y
3. Dan $r = 0$, disebut tidak berkorelasi secara linear, artinya tidak ada hubungan antara variabel X dan Y

Atau dapat juga dicari dengan rumus:

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

Arti dari koefisien relasi r adalah:

1. Bila $0,90 < r < 1,00$ atau $-1,00 < r < -0,90$: hubungan yang sangat kuat
2. Bila $0,70 < r < 0,90$ atau $-0,90 < r < -0,70$: hubungan yang kuat
3. Bila $0,50 < r < 0,70$ atau $-0,70 < r < -0,50$: hubungan yang moderat
4. Bila $0,30 < r < 0,50$ atau $-0,50 < r < -0,30$: hubungan yang lemah
5. Bila $0,0 < r < 0,30$ atau $-0,30 < r < 0,0$: hubungan yang sangat lemah

Contoh:

Data pada suatu pabrik kertas menunjukkan bahwa banyaknya kertas rusak ada hubungannya dengan kecepatan beroperasi mesin cetak.

Kecepatan mesin/ menit	8	10	11	12	13	14	15	16	17	18
Jml kerusakan kertas (lembar)	6	7	8	7	10	9	12	9	11	12

Tentukan:

1. Persamaan regresi linear
2. Berapa perkiraan kertas yang rusak, bila kecepatan mesin permenit adalah 20
3. Tentukan koefisien korelasi data serta berikan artinya
4. Kesalahan baku yang diberikan oleh persamaan regresi tersebut.