

ESTIMASI FUNGSI REGRESI

Untuk menentukan β_0 dan β_1 digunakan Least Squares Method (metode kuadrat terkecil). Misalkan $P = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$

$$\frac{\partial P}{\partial \beta_0} = 2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i) (-1)$$

$$\frac{\partial P}{\partial \beta_1} = 2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i) (-X_i)$$

P akan minimum jika turunan parsial P terhadap β_0 dan β_1 keduanya sama dengan nol.

$$\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i) = 0 \quad (1)$$

$$\sum_{i=1}^n X_i (Y_i - \beta_0 - \beta_1 X_i) = 0 \quad (2)$$

Sistem persamaan ini dinamakan persamaan normal. Dengan menyelesaikan persamaan normal tersebut diperoleh :

$$b_1 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b_0 = \bar{Y} - b_1 \bar{X}$$

Untuk lebih memudahkan penghitungan pada data yang cukup besar, dapat digunakan fasilitas pada kalkulator yang telah menyediakan menu statistik regresi, yang umumnya dikenal dengan a untuk b_0 dan b untuk b_1 .

Contoh : Menghitung b_0 dan b_1 untuk data iklan dan penjualan

$$b_1 = \frac{9(134425) - 295(3970)}{9(10925) - 295^2} = 3.42$$

dan $b_0 = 441.11 - (3.42)(32.78) = 329$

dengan demikian garis regresi dugaannya adalah $\hat{Y} = 329 + 3.42 X$

Persamaan regresi ini berarti saat kita naikkan iklan sebesar satu satuan maka penjualan akan meningkat sebesar 3.42 satuan. Sedangkan jika tidak ada iklan maka penjualan hanya sebesar 329.

Practice

Practice your capability to entry data and find the estimation of regression model using calculator

1. Data

Curah hujan / hari (0.01 inch)	Polusi yang terhilangkan (mikogram per meter kubik)
4.3	128
4.5	121
5.9	116
5.6	120
6.1	114
5.2	118
3.8	140
2.1	141
7.5	108

- Which one is an independent variable ?
 - And which one is a dependent variable ?
 - Estimate the linear regression function after drawing its scatter plot!
2. Tes kebugaran fisik adalah salah satu aspek yang penting untuk seorang atletis. Salah satu indikatornya adalah volume maksimum pengambilan oksigen selama suatu latihan. Penelitian dilakukan pada 24 pria paruh baya untuk mengetahui pengaruh dari waktu yang diperlukan untuk menyelesaikan 2 mil berlari dengan menggunakan treadmill. Data sebagai berikut :

no	max vol O ₂	waktu (detik)	no	max vol O ₂	waktu (detik)
1	42,33	918	13	48,18	858
2	53,1	805	14	43,21	860
3	42,08	892	15	51,81	760

4	50,06	962	16	53,28	747
5	42,45	968	17	53,29	743
6	42,46	907	18	47,18	803
7	47,82	770	19	56,91	683
8	49,92	743	20	47,8	844
9	36,23	1045	21	48,65	755
10	49,66	810	22	53,69	700
11	41,49	927	23	60,62	748
12	46,17	813	24	56,73	775

- a) Tentukan persamaan regresi dugaannya
- b) Berapakah estimasi maksimum O_2 untuk waktu berlari masing-masing : 750; 775; 800; 825 dan 850

Untuk memperkirakan bentuk suatu persamaan regresi terlebih dahulu dilakukan langkah-langkah berikut :

- Buat scatter plotnya
- Cari persamaan regresi yang paling mendekati.

Contoh :

$$Y = f(x) + \varepsilon \quad , \quad \varepsilon \sim (0, \sigma^2)$$

