
REGRESI MULTIPLE

March 26, 2024

Sains Data IKOPIN
Pengantar Model Linear
github.com/sevinurafni/SD2201

Model Regresi Multiple

Kasus pada kehidupan nyata hubungan antara variabel bukan hanya terdiri dari dua variabel. Tetapi juga dapat terjadi hubungan antara tiga variabel atau bahkan lebih. Misalnya, apakah kecenderungan berprestasi ini hanya cukup ditinjau dari motivasi saja? mengapa tidak dari ciri kepribadian, kecerdasan dan barangkali juga dari tingkat pendidikan. Apabila ingin dicari hubungan antar variabel-variabel tersebut maka, harus digunakan suatu model yang disebut persamaan regresi multiple dengan persamaan matematikanya adalah

$$Y = \alpha_0 + \alpha_1 X_{1i} + \dots + \alpha_k X_{ki} + \varepsilon_n \quad (1)$$

apabila ditulis dalam bentuk matriks diperoleh

$$Y = \alpha X + \varepsilon \quad (2)$$

dengan,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} & \vdots & X_{nk} \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (3)$$

dengan metode kuadra terkecil dapat dicari taksiran dari α sebagai berikut:

$$\alpha = (X^T X)^{-1} (X^T Y) \quad (4)$$

Untuk mengetahui apakah koefisien regresi mempunyai arti atau tidak maka, perlu diperlakukan pengujian dengan menggunakan daftar Anova, sedangkan bentuk hipotesisnya adalah:

$$H_0 : \alpha^* \rightarrow \text{model yang digunakan} \quad \hat{Y} = \alpha_0 + \varepsilon$$

$$H_1 : \alpha^* \rightarrow \text{model yang digunakan} \quad \hat{Y} = \alpha X + \varepsilon$$

Sumber Variasi	Dk	Jk
Regresi pada X_1, \dots, X_k	k	$\alpha^T (X^T Y) \alpha$
Residu	n-k-1	$(Y^T Y) - \alpha^T (X^T Y) \alpha$
Total	n-1	$(X^T Y)$

Contoh:

Lima rumah tangga petani dari suatu daerah pertanian dipilih sebagai sampel acak untuk diteliti tentang pengaruh pendapatan (X_1) dan kekayaan (X_2) terhadap tingkat konsumsinya (Y). Dari hasil penelitian diperoleh data sebagai berikut:

X_1	80	110	90	60	60
X_2	120	60	60	30	180
Y	74	98	80	53	57

Tentukanlah:

- Persamaan regresi multiple
- Uji keberartian koefisien regresi

Penyelesaian:

Apabila data di atas ditulis dalam bentuk matrik maka diperoleh:

$$Y = \begin{bmatrix} 74 \\ 98 \\ 80 \\ 53 \\ 57 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 80 & 120 \\ 1 & 110 & 60 \\ 1 & 90 & 60 \\ 1 & 60 & 30 \\ 1 & 60 & 180 \end{bmatrix}$$

$$(Y^T Y) = \begin{bmatrix} 74 & 98 & 80 & 53 & 57 \end{bmatrix} \begin{bmatrix} 74 \\ 98 \\ 80 \\ 53 \\ 57 \end{bmatrix} = 27538$$

$$\begin{aligned} (X^T X) &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 80 & 110 & 90 & 60 & 60 \\ 120 & 60 & 60 & 30 & 180 \end{bmatrix} \begin{bmatrix} 1 & 80 & 120 \\ 1 & 110 & 60 \\ 1 & 90 & 60 \\ 1 & 60 & 30 \\ 1 & 60 & 180 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 400 & 450 \\ 400 & 33800 & 34200 \\ 450 & 34200 & 54900 \end{bmatrix} \end{aligned}$$

$$(X^T Y) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 80 & 110 & 90 & 60 & 60 \\ 120 & 60 & 60 & 30 & 180 \end{bmatrix} \begin{bmatrix} 74 \\ 98 \\ 80 \\ 53 \\ 57 \end{bmatrix}$$

$$= \begin{bmatrix} 362 \\ 30500 \\ 31410 \end{bmatrix}$$

a.

$$\alpha = (X^T X)^{-1} (X^T Y)$$

$$= \begin{bmatrix} 68598.10^4 & -657.10^4 & -153.10^4 \\ -657.10^4 & 72.10^3 & 9.10^3 \\ -153.10^4 & 9.10^3 & 9.10^3 \end{bmatrix} \begin{bmatrix} 362 \\ 30500 \\ 31410 \end{bmatrix} = \begin{bmatrix} -1,0365 \\ 0,8849 \\ 0,0294 \end{bmatrix}$$

sehingga persamaan regresinya adalah $\hat{Y} = -1,0365 + 0,8849X_1 + 0,0294X_2$

b.

Sumber Variasi	Dk	Jk	RJk
Regresi pada X_1, X_2	2	27537,691	13768,8455
Residu	2	0,309	0,1545
Total	4	27538	

$$F_{\text{perhitungan}} = \frac{13768,8455}{0,1545} = 89118,7411$$

$$F_{\alpha; 2:2} = 19.0$$

H_0 : ditolak maka H_1 diterima atau model $\hat{Y} = \alpha X + \varepsilon$ diterima

Kesimpulan Dalam Regresi Multiple

Dengan statistika kita berusaha untuk menyimpulkan populasi. Untuk itu sifat populasi dipelajari berdasarkan data yang diambil baik secara sampling maupun sensus. Sifat populasi yang akan ditinjau hanya mengenai parameter populasi dan sample yang digunakan secara acak.

Cara pengambilan kesimpulan yang dibicarakan adalah interval taksiran dan pengujian hipotesis. Sehubungan dengan masalah regresi multiple maka yang dibahas adalah mengenai pengujian hipotesis dan interval taksiran tentang koefisien regresi multiple rata-rata respon dan nilai respon yang tunggal.

A. Pengujian Hipotesis

Menguji keberartian koefisien regresi multiple dari persamaan regresi:

$$Y = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_n \quad (5)$$

maka hipotesisnya berbentuk:

$$H_0 : \beta_j = 0 \quad (\text{tidak ada hubungan/pengaruh } X \text{ terhadap } Y)$$

$$H_1 : \beta_j \neq 0 \quad (\text{ada hubungan/pengaruh } X \text{ terhadap } Y)$$

Statistik yang digunakan adalah

$$\frac{b_j - \beta_{jo}}{\sigma b_j} \quad (6)$$

dengan:

$b_j - \beta_{jo}$: adalah nilai koefisien β_j

b_j : adalah nilai dari $Var(b_j)$ yang ada di matriks.

terima H_0 jika:

$$-t_{\frac{\alpha}{2}} < t < t_{\frac{\alpha}{2}}$$

dengan

$$t_{\frac{\alpha}{2}} = t_{\frac{\alpha}{2}}; n - k - 1$$

Contoh:

Dari hasil penelitian diperoleh data sebagai berikut:

Y	X1	X2	X3
25.5	1.74	5.3	10.8
31.2	6.32	5.42	9.4
25.9	6.22	8.41	7.2
38.4	10.52	4.63	8.5
18.4	1.19	11.6	9.4
26.7	1.22	5.85	9.9
26.4	4.1	6.62	8
25.9	6.32	8.72	9.1
32	4.08	4.42	8.7
25.2	4.15	7.6	9.2
39.7	10.15	4.83	9.4
35.7	1.72	3.12	7.6
26.5	1.7	5.3	8.2

Apakah koefisien regresi mempunyai arti atau tidak?

Penyelesaian:

Untuk menentukan persamaan regresi multiple, maka nilai β diperoleh dari:

$$\beta = (X^T X)^{-1} (X^T Y)$$

$$= \begin{bmatrix} 13 & 59.43 & 81.82 & 115.4 \\ 394.7255 & 360.6621 & 522.078 & 1877.567 \\ 576.7264 & 728.31 & 2246.661 & 3337.780 \\ 1035.96 & & & \end{bmatrix} \begin{bmatrix} 377.5 \\ 1877.567 \\ 2246.661 \\ 3337.780 \end{bmatrix}$$

$$\div [$$