

Regresi Linier Berganda

Regresi berganda merupakan alat statistik yang dapat digunakan untuk membangun model untuk memrediksi suatu keluaran (*outcome*) (Higgins, 2005). Menurut Simon (2003), regresi untuk tujuan prediksi adalah yang paling bermanfaat, namun regresi untuk tujuan ini sedikit. Kebanyakan regresi digunakan untuk memeriksa apakah terdapat hubungan signifikan antara Y dengan satu atau lebih prediktor.

REGRESI linier berganda digunakan apabila variabel independen terdiri dari dua atau lebih. Menurut Simon (2003), persamaan regresi linier untuk k variabel adalah:

$$Y_i = \beta_0 + \beta_1 (X_1)_i + \beta_2 (X_2)_i + \beta_3 (X_3)_i + \dots + \beta_k (X_k)_i + \varepsilon_i \quad \dots (3.1)$$

Label 'i' menyatakan tentang apa regresi dimaksud, β_k =koefisien ke-k, ε adalah gangguan atau *error* yang bersifat acak dan tidak teramati atau bukan hasil pengamatan. Umumnya label 'i' tidak disertakan, sehingga model umum regresi linier berganda paling sering ditemukan adalah:

$$Y = \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3)_i + \dots + \beta_k (X_k) + \varepsilon \quad \dots (3.2)$$

XK adalah variabel independen ke-k, β_k =koefisien ke-k, ε adalah gangguan atau error yang bersifat acak dan tidak teramati atau bukan hasil pengamatan. Dalam pelaksanaannya, komponen ' ε ' tidak disertakan pada persamaan. Persamaan demikian disebut *fitted model*, yaitu:

$$\hat{Y} = \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3)_i + \dots + \beta_k (X_k) \quad \dots (3.3)$$

Apabila *fitted model* dipakai untuk memrediksi nilai Y pada kasus ke-i, maka penulisan modelnya adalah:

$$\hat{Y}_i = \beta_0 + \beta_1 (X_1)_i + \beta_2 (X_2)_i + \beta_3 (X_3)_i + \dots + \beta_k (X_k)_i$$

... (3.4)

Penulisan persamaan regresi di atas sebenarnya tidak baku. Ada cara lain penulisan notasi persamaan selain notasi yang kita gunakan, seperti ditampilkan pada Tabel 3.1.

Tabel 3.1. Cara Penulisan Notasi Model Regresi	
NOTASI YANG KITA PAKAI	CARA LAIN PENULISAN NOTASI
Y_i	y_i
X_i	x_i
$\beta_0 + \beta_1 X_1$	$\alpha + \beta X_1$
ε_i	e_i
$(X_1)_i, (X_2)_i, (X_3)_i, \dots, (X_K)_i$	$X_{i1}, X_{i2}, X_{i3}, \dots, X_{iK}$

Kesuksesan regresi dinilai berdasarkan tujuan regresi dilakukan. Apabila dimaksudkan untuk memprediksi \hat{Y}_i , maka keberhasilan regresi tergantung pada seberapa seberapa dekat nilai prediksi Y (yaitu $\hat{Y}_1, \hat{Y}_2, \hat{Y}_3 \dots \hat{Y}_n$) dengan nilai observasi Y ($Y_1, Y_2, Y_3 \dots Y_n$) atau seberapa kecil *error* kuadrat. Namun, menurut Simon (2003), prediksi bukan satu-satunya tujuan regresi linier berganda. Menurutnya, pertanyaan lain adalah apakah persamaan regresi dapat dipercaya? Untuk itu, regresi linier berganda perlu memenuhi berbagai asumsi atau persyaratan.