

PREDIKSI NILAI VARIABEL TERIKAT DALAM REGRESI LINIER SEDERHANA

Catatan pendahuluan:

1. Tulisan ini terkait dengan *posting* berjudul PENGARUH SPP TERHADAP PENJUALAN CILOK yang terdapat di www.edscyclopedia.com.
2. Lambang-lambang, notasi yang digunakan dalam tulisan ini mengacu pada tautan berjudul "Catatan kuliah mengenai regresi linier sederhana" yang terletak di bagian akhir *posting* tersebut di atas.

Prediksi terhadap nilai variabel terikat baru dapat dilakukan apabila kita telah mendapatkan fungsi regresi sampel $Y_i = \hat{\beta}_1 + \hat{\beta}_2 \bar{X} + \hat{u}_i$. [Cara mendapatkan fungsi ini dapat dilihat pada tautan berjudul CARA MENDAPATKAN FUNGSI REGRESI SAMPEL yang terdapat pada bagian bawah *posting* yang dimaksud pada catatan pendahuluan di atas.] Pada dasarnya terdapat dua macam prediksi yang dapat dilakukan dengan model regresi linier sederhana. Pertama, dapat dilakukan prediksi terhadap $E(Y_0|X_0)$, yaitu prediksi terhadap rata-rata nilai variabel terikat apabila diketahui nilai variabel penjelasnya adalah $X = X_0$. Kedua, dapat juga dilakukan prediksi terhadap nilai Y_0 apabila diketahui variabel penjelas bernilai $X = X_0$. Kedua jenis prediksi ini memberikan jawaban berupa suatu selang kepercayaan (*confidence interval*) dengan suatu derajat kepercayaan (*degree of confidence*) yang kita tentukan. Prediksi jenis kedua ini dinamakan prediksi individual (*individual prediction*).

Bagian 1: Prediksi terhadap $E(Y_0|X_0)$:

Selang kepercayaan 100(1- α)% bagi $E(Y_0|X_0)$ adalah:

$$\hat{Y}_0 - t_{\alpha/2} se(\hat{Y}_0) \leq E(Y_0|X_0) \leq \hat{Y}_0 + t_{\alpha/2} se(\hat{Y}_0) \quad [df = n - 2]$$

$$\text{dengan } \hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0 \text{ dan } se(\hat{Y}_0) = \sigma \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n x_i^2}}$$

Nilai σ tidak diketahui sehingga ditaksir menggunakan $\hat{\sigma}$. Nilai \hat{Y}_0 adalah estimasi titik terhadap rata-rata nilai variabel terikat apabila diketahui nilai variabel penjelasnya adalah $X = X_0$.

Bagian 2: Prediksi terhadap Y_0 : (Individual Prediction)

Selang kepercayaan 100(1- α)% bagi Y_0 adalah:

$$\hat{Y}_0 - t_{\alpha/2} \cdot se(Y_0 - \hat{Y}_0) \leq Y_0|X_0 \leq \hat{Y}_0 + t_{\alpha/2} \cdot se(Y_0 - \hat{Y}_0) \quad [df = n - 2]$$

$$\text{dengan } \hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0 \text{ dan } se(Y_0 - \hat{Y}_0) = \sigma \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n x_i^2}}$$

Nilai σ tidak diketahui sehingga ditaksir menggunakan $\hat{\sigma}$.

Sebagai contoh penerapan rumus-rumus tersebut, akan digunakan contoh kasus Cilok Mr. Omon yang tersaji dalam *posting* saya berjudul PENGARUH SPP TERHADAP PENJUALAN CILOK yang terdapat di www.edscyclopedia.com. Misalnya kita ingin mengetahui berapa rata-rata pendapatan dari penjualan cilok terhadap tiap 50 orang mahasiswa di kampus yang SPP-nya Rp 3,6 juta? Untuk menjawab ini, substitusikan $X_0 = 3,6$ ke $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0 = -14,112 + 45,545 \cdot 3,6 = 149,85$.

Kemudian hitunglah $se(\hat{Y}_0) = \sigma \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n x_i^2}} = 8,76716 \sqrt{\frac{1}{20} + \frac{(3,6-4,0)^2}{9,8057}} \approx 2,25565$

Dari Tabel Nilai Kritis t dengan $\alpha = 0,05$ dan $df = 20-2 = 18$ diperoleh $t_{0,025;18} = 2,101$ sehingga diperoleh selang kepercayaan 95% bagi rata-rata pendapatan dari penjualan cilok terhadap tiap 50 orang mahasiswa di kampus yang SPP-nya Rp 3,6 juta sebagai berikut.

$$\begin{aligned} \hat{Y}_0 - t_{\alpha/2} se(\hat{Y}_0) &\leq E(Y_0|X_0) \leq \hat{Y}_0 + t_{\alpha/2} se(\hat{Y}_0) \\ 149,85 - 2,101 \cdot 2,25565 &\leq E(Y_0|3,6) \leq 149,85 + 2,101 \cdot 2,25565 \\ 145,11 &\leq E(Y_0|3,6) \leq 154,59 \end{aligned}$$

Jadi, kita 95% yakin bahwa rata-rata pendapatan dari penjualan cilok terhadap tiap 50 orang mahasiswa di kampus yang SPP-nya Rp 3,6 juta adalah antara Rp 145.110 hingga Rp 154.590.

Sekarang, apabila SPP di suatu kampus adalah Rp 3,6 juta, berapa estimasi bagi pendapatan dari penjualan cilok di sana? Ini suatu contoh pertanyaan mengenai prediksi individual.

$$se(Y_0 - \hat{Y}_0) = \sigma \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n x_i^2}} = 8,76716 \sqrt{1 + \frac{1}{20} + \frac{(3,6 - 4,0)^2}{9,8057}} \approx 9,05269$$

Dari Tabel Nilai Kritis t dengan $\alpha = 0,05$ dan $df = 20-2 = 18$ diperoleh $t_{0,025;18} = 2,101$ sehingga diperoleh selang kepercayaan 95% untuk pendapatan dari penjualan cilok terhadap tiap 50 orang mahasiswa di kampus yang SPP-nya Rp 3,6 juta sebagai berikut.

$$\begin{aligned} \hat{Y}_0 - t_{\alpha/2} \cdot se(Y_0 - \hat{Y}_0) &\leq Y_0|X_0 \leq \hat{Y}_0 + t_{\alpha/2} \cdot se(Y_0 - \hat{Y}_0) \\ 149,85 - 2,101 \cdot 9,05269 &\leq Y_0|X_0 = 3,6 \leq 149,85 + 2,101 \cdot 9,05269 \\ 130,831 &\leq Y_0|X_0 = 3,6 \leq 168,870 \end{aligned}$$

Jadi, apabila SPP di suatu kampus Rp 3,6 juta, kita 95% yakin bahwa pendapatan dari penjualan cilok terhadap tiap 50 orang mahasiswa di kampus tersebut adalah antara Rp 130.831 hingga Rp 168.870.