

REGRESI LINIER GANDA

Pengertian Regresi Linier Ganda

- Merupakan metode yang digunakan untuk memodelkan hubungan linear antara variabel terikat dengan dua/lebih variabel bebas.
- Regresi linier untuk memprediksi variabel terikat dari dua/lebih variabel bebas.

Model Regresi Linier Ganda

- Model regresi linier ganda yang melibatkan p buah variabel bebas adalah:

$$\hat{Y} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \varepsilon_i$$
$$(i = 1, 2, 3, \dots, N)$$

dalam notasi matriks menjadi: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

- Nilai-nilai parameter di atas akan ditaksir sehingga modelnya menjadi:

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \cdots + b_p X_{pi} + e_i$$

Model Regresi Linier Ganda

- Pendugaan parameter regresi β dengan menggunakan metode OLS berdasarkan model $Y = X\beta + \varepsilon$ adalah dengan meminimumkan *sum square error* (SSE) dimana SSE dirumuskan:

$$\mathbf{SSE} = \varepsilon^t \varepsilon = (\mathbf{Y} - \mathbf{X}\beta)^t (\mathbf{Y} - \mathbf{X}\beta)$$

maka penaksir kuadrat terkecil untuk β adalah

$$\hat{\beta} = b = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y}$$

Asumsi-Asumsi Regresi Linier Ganda

1. *Asumsi Univariate*
2. *Asumsi Bivariate*
3. *Asumsi Model Multivariate*

Asumsi *Univariate*

- Variabel-variabel harus berdistribusi NORMAL (terutama variabel terikat), agar analisa dengan uji statistik parametrik
- Salah satu cara untuk menguji kenormalan data yaitu dengan uji *Kolmogorov-Smirnov* (**NPar Test**).
- Hipotesis: H_0 : Variabel berdistribusi normal
 H_1 : Variabel tidak berdistribusi normal
- Dengan menggunakan SPSS, jika nilai sig lebih dari α (5%) maka H_0 diterima.

Asumsi *Bivariate*

- Korelasi antara variabel terikat dengan variabel bebas dapat dideteksi dengan melakukan uji **korelasi Pearson**.

Variabel dengan nilai-p $< 0,15$ (0,25 atau 0,10) merupakan kandidat model.

- Korelasi antar variabel bebas perlu juga diketahui untuk mewaspadaai adanya gejala kolinearitas. Jika nilai $r > 0,8$ maka ada kolinearitas.

Asumsi Model *Multivariate*

- Kelinieran (*Linearity*)
- *Multivariate Normality*
- *Existency*
- *Homocedasticity*
- *Independency*
- *Colinearity*

Kelinieran

- Untuk menguji apakah variabel Y merupakan fungsi linier dari gabungan variabel-variabel X .
- Uji kelinieran ini menggunakan **ANOVA**
- Hipotesis: H_0 : Persamaan regresi tidak linier
 H_1 : Persamaan regresi linier
- Dengan menggunakan SPSS, jika nilai sig lebih dari α (5%) maka H_0 diterima.

Multivariate Normality

- Untuk menguji apakah variabel Y berdistribusi normal untuk gabungan variabel-variabel X .
- Uji *multivariate normality* ini menggunakan plot residual.
- **PLOTS: histogram & Normal Probability plot)**

Existency

- Untuk menguji apakah variabel Y merupakan random variable yang mempunyai mean & SD.
- Uji *existency* ini dilakukan dengan melihat nilai mean Residual pada Residuals Statistics
- Apabila mean Residualnya = 0, maka asumsi *existency* terpenuhi.

Homocedasticity

- Untuk menguji apakah varian nilai Y sama (homogen) untuk setiap variabel X .
- Uji *homocedasticity* dilakukan dengan melihat diagram residual terhadap variabel-variabel bebas pada output **Scatterplot** (antara Nilai prediksi standar (ZPRED) dengan nilai residu standar (ZRESID) atau residu student (SRESID)).
- Apabila Scatterplot tidak membentuk pola tertentu berarti asumsi homocedasticity dipenuhi.

Heterocedasticity

- Heteroskedastisitas terjadi apabila error atau residual dari model yang diamati tidak memiliki varian yang konstan dari satu observasi ke observasi lainnya.
- Uji heterocedasticity dilakukan dengan melihat diagram residual terhadap variabel bebas pada output Scatterplot.
- Jika nilai error membentuk pola tertentu tidak bersifat acak terhadap nol maka dikatakan terjadi heteroskedastisitas

Independency

- Untuk menguji apakah variabel Y mempunyai nilai yang saling bebas (no autocorr) satu dengan lainnya.
- Untuk menguji apakah dalam model regresi linier ada korelasi antar error satu dengan error yang lainnya.
- Uji *independency* ini menggunakan uji Durbin-Watson (DW) pada output Model Summary.
- Jika $-2 < DW < 2$ maka tidak terjadi autokorelasi (asumsi independensi terpenuhi)

Colinearity

- Untuk menguji apakah tidak terdapat korelasi antar variabel-variabel bebas.
- Untuk menguji apakah dalam model regresi linier ditemukan adanya korelasi antar variabel bebas.
- Uji *colinearity* dapat dilakukan dengan melihat nilai Variance Inflasi Factor (VIF) dan tolerance pada output Coefficients.
- VIF dirumuskan dengan

$$VIF = \frac{1}{1 - R_j^2}$$

Colinearity

- $VIF < 1$ mengindikasikan tidak ada korelasi yang signifikan antar variabel bebas
- $VIF > 1$ mengindikasikan bahwa ada korelasi antar variabel bebas namun kecil efeknya
- $VIF > 5 - 10$ mengindikasikan bahwa ada salah satu variabel bebas merupakan fungsi dari variabel bebas lainnya