

ANALISI REGRESI GANDA

Regresi linear ganda adalah persamaan regresi yang menggambarkan hubungan antara lebih dari satu peubah bebas (X) dan satu peubah tak bebas (Y) Hubungan peubah-peubah tersebut dapat dituliskan dalam bentuk persamaan:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

Y = Peubah tak bebas, X = Peubah bebas, β_0 = intersep/perpotongan dengan sumbu tegak, $\beta_1, \beta_2, \dots, \beta_{p-1}$ = parameter model regresi, ε_i saling bebas dan menyebar normal $N(0, \sigma^2)$, $i = 1, 2, \dots, n$

Persamaan regresi dugaannya adalah

$$\hat{Y}_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_{p-1} X_{i,p-1}$$

Hipotesis yang harus diuji dalam analisis regresi ganda adalah

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$$

$$H_1 : \text{Tidak semua } \beta_k \text{ (} k=1, 2, \dots, p-1 \text{) sama dengan nol}$$

Untuk melakukan pendugaan parameter model regresi ganda dan menguji signifikansinya dapat dilakukan dengan program SPSS 16.

Asumsi yang harus dipenuhi dalam analisis regresi ganda adalah :

- Tidak ada multikolinearitas (korelasi antara variabel independen)
- Heteroskedastisitas (variansi *error* konstan)
- Normalitas (*error* berdistribusi normal)
- Autokorelasi (*error* bersifat acak)

Multikolinearitas

- Multikolinearitas atau kekolinearan ganda adalah terjadinya korelasi antar peubah bebas.
- Model regresi yang baik seharusnya tidak terjadi korelasi antar peubah bebas.
- Metode yang banyak digunakan untuk mendeteksi adanya multikolinearitas adalah faktor inflasi ragam (*variance inflation factor/VIF*)

- Multikolinearitas terjadi jika nilai VIF > 10

Heteroskedastisitas

- Ragam galat diasumsikan konstan dari satu pengamatan ke pengamatan lain, hal ini disebut **homoskedastisitas**.
- Jika ragam galat berbeda disebut **heteroskedastisitas**.
- Model regresi yang baik adalah tidak terjadi heteroskedastisitas.
- Untuk mendeteksi heteroskedastisitas adalah dengan membuat plot nilai dugaan yang dibakukan (*standardized predicted value*) dengan sisaan yang dibakukan (*studentized residual*).
- Jika ada pola tertentu (bergelombang, melebar kemudian menyempit) maka **terjadi heteroskedastisitas**.
- Jika tidak ada pola jelas, serta titik-titik (sisaan) menyebar di atas dan di bawah angka 0 pada sumbu Y, maka **tidak terjadi heteroskedastisitas**.

Normalitas (*error* berdistribusi normal)

- Untuk mendeteksi normalitas digunakan *normal p-p plot*.
- Jika titik-titik (sisaan) menyebar di sekitar garis diagonal dan mengikuti arah garis diagonal, maka model regresi memenuhi asumsi normalitas.
- Jika titik-titik (sisaan) menyebar jauh dari garis diagonal dan atau tidak mengikuti arah garis diagonal, maka model regresi tidak memenuhi asumsi normalitas

Autokorelasi.

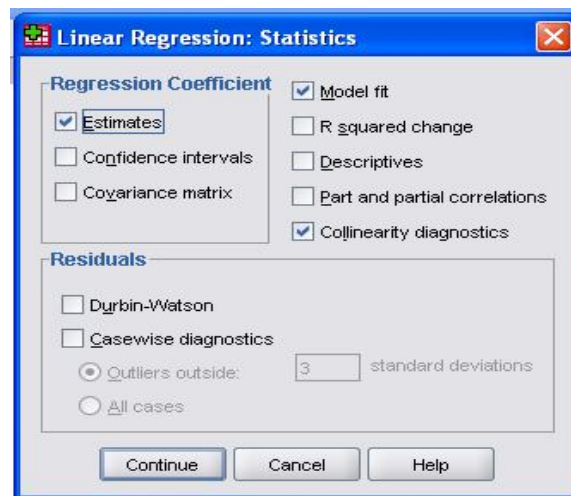
- Bila dalam model regresi linear ganda ada korelasi antara galat pada periode t dengan galat pada periode t-1, maka dinamakan ada masalah autokorelasi.
- Model regresi yang baik adalah model regresi yang bebas dari autokorelasi.

Contoh. Misalkan dipunyai data

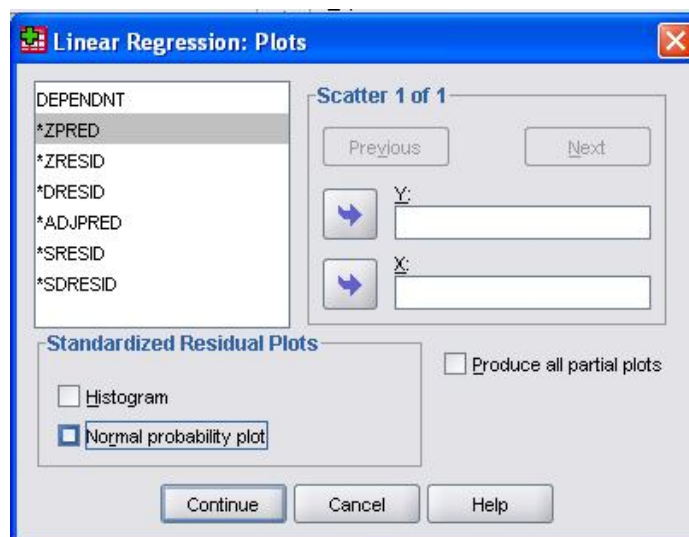
Y	10	6	5	12	10	15	5	12	17	20
X ₁	1.3	2.0	1.7	1.5	1.6	1.2	1.6	1.4	1.0	1.1
X ₂	9	7	5	14	15	12	6	10	15	21

Akan dilakukan pendugaan dan pengujian parameter model regresi, serta uji asumsi dengan menggunakan SPSS 16.

1. Cara memasukkan data dan melakukan analisis sama dengan pada regresi sederhana.
2. Untuk memunculkan hasil uji asumsi pada kotak dialog statistics klik juga collinearity diagnostics baru continue, sebagaimana terlihat pada gambar berikut:



3. Untuk melakukan uji asumsi pada residual klik plots, sehingga akan muncul kotak dialog :



4. Masukkan ZPRED pada kotak X dan ZRESID pada kotak Y, dan beri tanda centang (✓) pada Normal probability plot, kemudian klik continue. Kembali ke kotak dialog awal, dan klik OK.

Hasil analisis dengan ANAVA adalah sebagai berikut:

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	217.699	2	108.849	47.917	.000 ^a
	Residual	15.901	7	2.272		
	Total	233.600	9			

a. Predictors: (Constant), VAR00003, VAR00002

b. Dependent Variable: VAR00001

Terlihat bahwa nilai signifikansi 0,000, sehingga H_0 ditolak, yang berarti ada hubungan linear antara variabel independen X_1 dan X_2 dengan variabel dependen Y.

Hasil uji parsial adalah sebagai berikut :

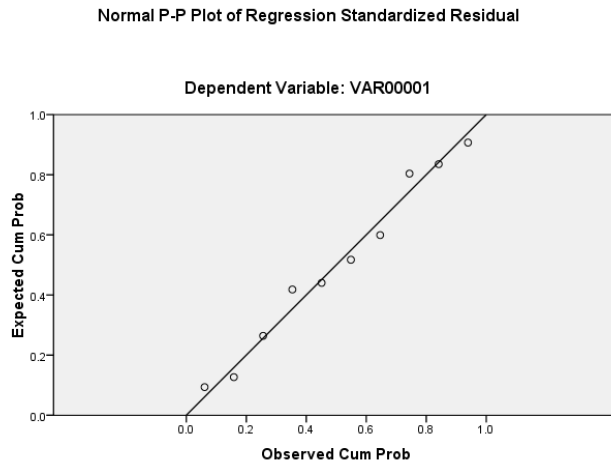
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	16.406	4.343		3.778	.007		
	VAR00002	-8.248	2.196	-.490	-3.756	.007	.572	1.749
	VAR00003	.585	.134	.571	4.377	.003	.572	1.749

a. Dependent Variable: VAR00001

Karena nilai signifikansi 0,007 untuk konstanta dan VAR00002 dan 0,003 untuk VAR00003, sehingga H_0 ditolak untuk semua uji. Jadi konstanta β_0 semua dan koefisien regresi β_1 , dan β_2 signifikan. Persamaan regresi dugaannya adalah :

$$\hat{Y}_i = 16.4068 - 8,248 X_{i1} + 0,585 X_{i2}$$

Hasil uji asumsi multikolinearitas dapat dilihat pada nilai VIF, yaitu $1,749 < 10$, sehingga dapat disimpulkan tidak ada multikolinearitas antara variabel X1 dan X2. Hasil uji normalitas dari error dapat dilihat pada output berikut



Karena plot mendekati garis diagonal, maka dapat disimpulkan *error* memenuhi asumsi normalitas. Uji normalitas error juga dapat dilakukan dengan uji Kolmogorov-Smirnov.

Hasil plot berikut menunjukkan tidak ada pola yang jelas, sehingga dapat disimpulkan tidak terjadi **heteroskedastisitas** atau ragam galat konstan.

