

PENGUJIAN ASUMSI REGRESI

1. NORMALITAS

Uji ini bertujuan untuk menguji apakah alam model regresi, variable pengganggu atau residual memiliki distribusi normal? Seperti diketahui bahwa Uji t dan Uji F mengasumsikan bahwa nilai residual mengikuti distribusi normal.

1.1. Uji One sample Kolmogorov Smirnov

Cara:

1. Analyze...nonparametric test...1-sample KS
2. Masukkan nilai yg akan dites dan biarkan default normal. OK.
3. Pada output perhatikan nilai signifikansi bila $>0,05$ berarti H_0 tidak dapat ditolak dimana data berdistribusi normal.

1.2. Nilai Standar

Cara:

1. Analyze...descriptive statistics...descriptive
2. masukkan variabel yang akan diuji
3. Klik save standardized values as variables. OK.
4. pada spss data editor perhatikan muncul nilai Z
5. Terjadi outlier bila terdapat nilai $Z < -1,96$ dan $Z > 1,96$
6. pada output spss : Bila nilai descriptive statistics tidak ada yang menunjukkan nilai Z yang lebih tinggi dari 3.0 maka berarti tidak ada outlier.

1.3. Grafik

Cara:

1. Graph...histogram.
2. masukkan variabel yg akan diuji
3. klik display normal curve. OK.
4. Perhatikan output: apakah menceng kiri/kanan atau normal?

2. MULTIKOLINEARITAS

Uji ini bertujuan untuk menguji apakah model regresi ditemukan adanya korelasi antar variabel bebasnya.

2.1. Perhatikan nilai R^2 yang sangat tinggi, tetapi secara individual variabel independennya banyak yang tidak signifikan

2.2. Matriks korelasi antar variabel independent. Bila antar variabel independent ada korelasi yang cukup tinggi (umumnya diatas 0,9) berarti ada gejala multikol.

Cara:

1. Analyze...correlate...bivariate.
2. masukkan variabel yang akan diuji
3. lihat pengujian yg sesuai apakah pearson, Kendal atau spearmen. OK.

2.3. Tolerance dan VIF

Nilai Cutoff tolerance adalah $< 0,10$ atau $VIF > 10$ menunjukkan adanya gejala multikol.

Cara:

1. Analyze...regresi...linear
2. masukkan variabel dependen dan independent
3. Klik statistic: aktifkan collinearity diagnostic. OK

2.4. Regresi Parsial

Cara:

1. lakukan estimasi pada model regresi awal $y = f(x_1, x_2, x_3, x_n)$ dan dapatkan nilai R^2 .
2. lakukan auxiliary regression antar var independen :
 $x_1 = f(x_2, x_3, x_n)$
 $x_2 = f(x_1, x_3, x_n)$
 $x_3 = f(x_1, x_2, x_n)$
 $x_n = f(x_1, x_2, x_3)$
3. Nilai R^2 masing-masing regresi pada point 2 diatas kemudian dibandingkan dengan nilai R^2 regresi utama point 1. bila nilai R^2 point 2 lebih tinggi dari model utama, maka didalam regresi parsial tersebut terdapat multikol.

3. AUTOKORELASI

Uji ini bertujuan untuk menguji apakah dalam model regresi linear terdapat korelasi antara kesalahan pengganggu pada periode t dengan kesalahan pengganggu pada periode $t-1$. autokorelasi timbul karena observasi yang berurutan sepanjang waktu berkaitan satu sama lainnya. Masalah ini timbul karena residual tidak bebas dari satu observasi ke observasi lainnya. Hal ini sering muncul pada data runtut waktu.

3.1. Durbin-Watson (DW Test)

Cara:

1. Analyze...regresi...linear
2. masukan variabel dependen dan independent
3. klik statistic: aktifkan residual durbin watson.
4. Bandingkan nilai DW hitung dengan DW table

3.2. Lagrange Multiplier (LM Test)

Untuk sample besar diatas 100 observasi. Uji LM akan menghasilkan statistic Breusch-Godfrey (BG Test) yg dilakukan dengan meregres residual ut menggunakan autoregresif model dengan order p :

$$U_t = p_1.U_{t-1} + p_2.U_{t-2} + \dots + p_n.U_{t-n} + e_t$$

Dg hipotesis $H_0 : p_1 = p_2 = \dots = p_n$ (koefisien autoregresif secara simultan sama dengan nol, yg menunjukkan tidak terdapat autokorelasi pada setiap orde).

Cara:

1. Analyze...regresi...linear
2. masukan dependen Y dan independent X1, X2, Xn
3. klik save: aktifkan unstandardized residual
4. pada spss data editor muncul nilai res_1
5. buat variabel lag residual (Ut-1, Ut-2, dst) dg perintah transform, compute.
6. pada kotak target variabel isikan: res_2 yg merupakan var.residual lag 2 (Ut-2) dan pada kotak numeric expression pilih fungsi lag (variabel) dan isikan menjadi lag(res_1). OK.
7. uji breusch-godfrey dg meregres pers. Sbb:
 $Res_1 = b_0 + b_1.X_1 + b_2.X_2 + \dots + b_n.X_n + b_{n+1}.Res_2$
8. perhatikan hasil regresi pada koefisien parameter untuk residual lag 2 (res_2) bila sig <0,05 hal ini berarti terdapat masalah autokorelasi.

3.3. Statistic Q: Box Pierce dan Ljung Box

Uji ini untuk melihat autokorelasi dg lag >2 (by default spss menguji sampai 16 lag)

Cara:

1. Dengan menggunakan regresi biasa, buatlah variabel unstandardized residual.
2. Graph...time series...autocorrelation
3. Masukkan variabel unstandardized residual (res_1)
4. pilih display autocorellation. OK.
5. pada output perhatikan signifikansi mulai lag 1 s/d lag 16. criteria adanya autokorelasi jika jumlah lag yang signifikan >2, bila lag yg signifikan 2 atau kurang dari 2 maka dikatakan tidak ada autokorelasi.

3.4. Run Test

Run test merupakan nonparamterik yg dapat digunakan untuk melihat apakah residual terjadi secara random atau tidak.

H0: residual random

H1: residual tidak random

Cara:

1. Dengan menggunakan regresi biasa, buatlah variabel unstandardized residual.
2. Analyze...nonparametric...runs
3. masukkan variabel unstandardized residual (res_1). OK.
4. perhatikan output, bila sig < 0,05, berarti H0 ditolah yg berarti data residual tidak random atau terjadi autokorelasi.

4. HETEROSKEDASTISITAS

Uji ini dipergunakan untuk menguji apakah dalam model regresi terjadi ketidaksamaan varians dari residual satu pengamatan ke pengamatan lainnya. Jika varians dari residual satu pengamatan ke pengamatan lainnya tetap, hal ini disebut sebagai homoskedastisitas. Kebanyakan data *cross section* mengandung situasi heteroskedastisitas karena data ini menghimpun data yang mewakili berbagai ukuran (kecil, sedang dan besar).

4.1. Uji Park

Dasarnya bahwa variance merupakan fungsi dari variabel independent, namun karena umumnya variance tidak diketahui maka ditaksir dengan menggunakan residual U_t , dg persamaan:

$$\ln U_t^2 = a + b.X_i$$

Cara:

1. Dengan menggunakan regresi biasa, buatlah variabel unstandardized residual.
2. ubah nilai `res_1` dg transform, compute.
3. Target var: `lnres`, numeric expresion: `ln(res_1**2)`. OK
4. Analyze...regresi...linear
5. dependen: `lnres`, independennya tetap. OK
6. Lihat ouput, bila `sig > 5%`, H_0 tidak dapat ditolak berarti homoskedastisitas.

4.2. Uji Glesjer

Cara: sama dg uji park namun pada variabel dependen bukan $\ln U_t^2$ namun absolute dari residual (`|res_1|`)

4.3. Scatterplot

Dasar: Jika ada pola tertentu yang teratur (seperti bergelombang, melebar kemudian menyempit) maka terdapat heteroskedastisitas. Bila tidak ada pola yang jelas, serta titik-titik menyebar diatas dan dibawah angka 0 pada sumbu y, maka tidak terjadi heteroskedastisitas.

Cara:

1. Analyze...regresi...linear
2. masukan var independent dan dependen
3. klik plot: Y:Sresid dan X:Zpred. OK.

5. LINEARITAS

Uji ini digunakan untuk melihat apakah spesifikasi model yang digunakan sudah benar atau tidak. Apakah fungsi yang digunakan dalam suatu studi empiris sebaiknya berbentuk linear, kuadrat atau kubik?

5.1. Uji Durbin Watson

Uji ini biasanya dilakukan untuk melihat adanya autokorelasi atau tidak. Cara menguji apakah model sebaiknya linear atau kuadrat dapat dilakukan dg cara sbb:

1. Lakukan regresi dengan 2 persamaan yaitu linear atau kuadrat, contoh:
(1) $Y = a + b.X_1 + c.X_2 + d.X_3$
(2) $Y = a + b.X_1 + c.X_2 + d.X_3 + e.X_1^2 + f.X_2^2 + g.X_3^2$
2. Dapatkan nilai DW untuk masing-masing model.
3. Berdasarkan pada table DW, bandingkan nilai statistiknya. Bila keduanya signifikan atau berada pada daerah autokorelasi +/-, maka spesifikasi persamaan utama adalah salah.

5.2. Ramsey Test

Dikembangkan oleh Ramsey (1969) yang menyarankan suatu uji yang disebut *general test of specification* atau RESET. Uji ini bertujuan untuk menghasilkan F hitung, caranya:

1. Dapatkan *fitted value* dari variable dependen dengan cara dari linear regresi, pilih save dan aktifkan Dfit pada *influence statistics*.
2. kemudian variable fitted tersebut diregres bersama dengan model semula sebagai variable independent. Dapatkan nilai R^2 untuk menghitung F statistic dengan rumus:

$$F = \frac{(R^2_{\text{new}} - R^2_{\text{old}})/m}{(1 - R^2_{\text{new}})/(n-k)}$$

Dimana:

- m = jumlah variable independent yang baru masuk
- n = jumlah data observasi
- k = banyaknya parameter dalam persamaan yang baru.
- R^2_{new} = nilai R^2 dari persamaan regresi baru
- R^2_{old} = nilai R^2 dari persamaan regresi lama.

Dari hasil perhitungan nilai F hitung, kemudian dibandingkan dengan F table, jika F hitung > F Tabel, maka hipotesis nol yang menyatakan bahwa spesifikasi model dalam bentuk fungsi linear adalah ditolak.

5.3. Uji Lagrange Multiplier

Dikembangkan oleh Engle (1982) yang merupakan pengembangan Ramsey test. Uji ini bertujuan untuk mendapatkan nilai chisquare/ χ^2 hitung atau ($n \times R^2$).

Caraya:

1. Lakukan regresi dengan persamaan utama $Y = f(X_1, X_2, X_3)$
2. Jika dianggap persamaan utama tersebut benar spesifikasinya, maka nilai residual harus dihubungkan dengan nilai kuadrat variable independent dengan persamaan regresi:
$$U_t = a + b.X_1^2 + c. X_2^2 + d. X_3^2$$
3. Dapatkan nilai R^2 untuk menghitung nilai chisquare/ c^2 hitung.
4. Jika Chisquare/ C^2 hitung $>$ Chisquare table , maka hipotesis yang model linear adalah ditolak.