

## MODUL 10

### UJI MULTIKOLINEARITAS

#### **Pengertian *Multikolinearitas***

Sebagaimana dalam daftar asumsi klasik yang ke-10 telah dibahas pengertian *multikolinearitas*. *Multikolinearitas* adalah korelasi linear yang “*perfect*” atau eksak di antara variabel penjelas yang dimasukkan ke dalam model. Misalnya kita melakukan penelitian mengenai perilaku variabel Y (Kinerja), dan dijelaskan oleh beberapa variabel yang kita masukkan ke dalam model katakanlah  $X_1$ ,  $X_2$ ,  $X_3$ , dan  $X_4$ . Persamaan kita tulis:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + e$$

Jika antara  $X_1$ ,  $X_2$ ,  $X_3$ , dan  $X_4$  ada yang memiliki korelasi tinggi maka hal tersebut mengindikasikan adanya problem *multikolinearitas*.

#### **Akibat Adanya *Multikolinearitas***

Jika antara  $X_1$ , dan  $X_2$  terjadi *multikolinear*, misalnya secara sempurna seluruh data menunjukkan bahwa  $X_1 = 2 X_2$ , maka nilai  $b_1$  dan  $b_2$  tidak dapat ditentukan hasilnya karena dari formula OLS sebagaimana dibahas terdahulu,

$$b_1 = \frac{(\sum yx_1)(\sum x_2^2) - (\sum yx_2)(\sum x_1x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2}$$

akan menghasilkan bilangan pembagian,

$$b_1 = \frac{0}{0}$$

Dengan demikian hasilnya tidak menentu. Demikian juga *standar error* ( $S_{b_1}$ ) akan menjadi sangat besar. Jika *multikolinearitas* tidak begitu sempurna tetapi tetap tinggi akibatnya adalah parameter *estimate*  $b_1$  yang diperoleh tetap *valid*, tetapi  $S_{b_1}$  akan bias membesar. Akibatnya uji t yang rumusnya berupa,  $t = b_1/S_{b_1}$  akan cenderung kecil.

### **Konsekuensi Adanya *Multikolinearitas* yang Tidak Sempurna**

Karena yang bias membesar dalam kasus *multikolinearitas* adalah  $S_b$  maka sepanjang uji  $t$  sudah signifikan maka *multikolinearitas* tidak perlu dirisaukan. Hal ini disebabkan arah pembiasan  $S_b$  yang selalu membesar. Logikannya jika  $t = b$  dibagi bilangan yang bias membesar masih signifikan maka  $t$  sebenarnya lebih signifikan lagi. Akan tetapi, jika  $t$  menjadi tidak signifikan maka *multikolinearitas* perlu ditangani.

### **Cara Menangani Adanya *Multikolinearitas***

Pada hakekatnya jika  $X_1$  dan  $X_2$  *multikolinear* maka keduanya bersifat saling mewakili dalam mempengaruhi variabel tergantung  $Y$ . Oleh karena itu penanganannya adalah dibuat persamaan yang terpisah.

Contoh: kita memiliki regresi sbb.

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + e$$

Karena  $X_1$  dan  $X_2$  memiliki *kolinearitas* yang tinggi, maka regresi dapat dibuat menjadi dua model.

$$Y = a + b_1X_1 + b_3X_3 + b_4X_4 + e$$

dan 
$$Y = a + b_2X_2 + b_3X_3 + b_4X_4 + e$$

Pada prinsipnya kita harus *mengestimate* dampak  $X_1$  terhadap  $Y$  dan  $X_2$  terhadap  $Y$  secara terpisah tidak dapat bersama-sama.

### Cara Menguji *Multikolinearitas*

Dengan contoh di atas kita dapat menggunakan korelasi metrik. Aturannya jika korelasi antara X lebih besar dari korelasi X dan Y maka variabel bebas tersebut mengindikasikan *multikolinear*.

Matrik korelasi

|                | Y     | X <sub>1</sub> | X <sub>2</sub> | X <sub>3</sub> | X <sub>4</sub> |
|----------------|-------|----------------|----------------|----------------|----------------|
| Y              | 1     |                |                |                |                |
| X <sub>1</sub> | .490  | 1              |                |                |                |
| X <sub>2</sub> | .564  | .347           | 1              |                |                |
| X <sub>3</sub> | .353  | .403           | .719           | 1              |                |
| X <sub>4</sub> | -.008 | -.005          | -.007          | -.006          | 1              |

\*\* Correlation is significant at the 0.01 level (2-tailed)

Terlihat dari tabel di atas korelasi Y dan Y = 1 (korelasi sempurna). Korelasi Y dan X<sub>1</sub>=0,49. Korelasi Y dan X<sub>2</sub> = 0,564 korelasi X<sub>3</sub> dan X<sub>2</sub> = 0,719. Disini terlihat bahwa antara X<sub>2</sub> dan X<sub>3</sub> terjadi korelasi yang tinggi. Hal tersebut menunjukkan adanya gejala *multikolinearitas*.

Solusinya sebaiknya model dipisah menjadi dua regresi yang mengandung X<sub>2</sub> dan tidak mengandung X<sub>3</sub> dan regresi yang lain mengandung X<sub>3</sub> tetapi tidak mengandung X<sub>2</sub>.