## BAB III

## IDENTIFIKASI VARIABEL MODERATOR KATEGORIK

### 3.1 Identifikasi Variabel Moderator

Misalkan $X_1, X_2, \ldots, X_k$ merupakan variabel prediktor dan Y merupakan variabel respon, serta terdapat n observasi. Model regresi linear berganda antara $X_1, X_2, \ldots, X_k$ dengan Y dapat ditulis sebagai berikut :

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon\,, \quad \varepsilon \sim NIID\ (0, \sigma^2)$$

Berdasarkan model tersebut, besarnya kekuatan hubungan antara X dan Y ditentukan oleh besarnya $\beta$, sehingga besarnya $\alpha$ tidak diperhatikan dalam model.

Berdasarkan Sharma *al*. (1981:297), langkah awal untuk mengidentifikasi apakah variabel Z merupakan variabel moderator atau bukan yaitu dengan membuat model regresi antara variabel Y dengan variabel X dan Z.

Perhatikan model regresi dimana variabel Z dilibatkan dalam model :

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \beta_{k+1} Z + \beta_{k+2} X_1 Z + \beta_{k+3} X_2 Z + \cdots + \beta_{2k+1} X_i Z + \varepsilon$$

Jika variabel Z dan atau interaksi antara X dan Z signifikan, maka Z merupakan variabel prediktor dan bukan merupakan variabel moderator Dalam hal ini, analisis dilakukan seperti biasa. Disisi lain jika Z dan interaksinya dengan X tidak signifikan, maka model regresinya yaitu:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Secara sepintas terlihat bahwa besarnya kekuatan hubungan antara X dan Y tidak dipengaruhi oleh Z. Akan tetapi, bisa jadi variabel Z tercakup dalam variabel galat. Oleh karena Z tercakup dalam variabel galat, maka secara tidak langsung Z akan mempengaruhi besarnya kekuatan hubungan antara X dan Y. Dalam arti, jika sampel dipartisi menjadi subgrup berdasarkan variabel Z, maka akan didapat taksiran dari $\beta_1, \beta_2, \ldots, \beta_k$ yang berbeda setiap subgrup. Jika taksiran dari $\beta_1, \beta_2, \ldots, \beta_k$ dalam setiap subgrup berbeda maka Z merupakan variabel moderator.

Pengujian untuk mengetahui taksiran dari $\beta_1, \beta_2, \dots, \beta_k$ sama atau tidak untuk setiap subgrup dilakukan dengan menggunakan uji Chow.

## 3.2 Uji Chow untuk Menguji Kesamaan Koefisien Regresi dari Dua Model Regresi

Misalkan dari suatu populasi diambil sampel acak berukuran n. Selanjutnya akan dilihat bentuk hubungan antara variabel $X_1, X_2, \dots, X_k$ dan $Y$ yang dinyatakan dalam sebuah persamaan. Untuk melihat bentuk hubungan antara variabel $X_1, X_2, \dots, X_k$ dan $Y$ dapat digunakan analisis regrsi linear berganda.

Model regresi untuk pengamatan di atas yaitu :

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \tag{3.1}$$

Misalkan responden dalam populasi tersebut dapat dikelompokkan menjadi dua subgrup yang saling lepas berdasarkan variabel $Z$. Misalkan pada subgrup pertama terdapat $n_1$ pengamatan, sedangkan pada subgrup kedua terdapat $n_2$ pengamatan sehingga $n = n_1 + n_2$

Tabel 3.1 Pengamatan pada Model Regresi Linear Berganda

| Subgrup | Pengamatan ke- | $X_1$ | $X_2$ | ... | $X_k$ | $Y$ |
|---|---|---|---|---|---|---|
| 1 | 1 | $X_{111}$ | $X_{112}$ | | $X_{11k}$ | $Y_{11}$ |
| | 2 | $X_{121}$ | $X_{122}$ | | $X_{12k}$ | $Y_{12}$ |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | $n_1$ | $X_{1n_11}$ | $X_{1n_12}$ | | $X_{1n_1k}$ | $Y_{1n_1}$ |
| 2 | 1 | $X_{211}$ | $X_{212}$ | | $X_{21k}$ | $Y_{21}$ |
| | 2 | $X_{221}$ | $X_{222}$ | | $X_{22k}$ | $Y_{22}$ |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

| | | | | | | |
|---|---|---|---|---|---|---|
| | $n_2$ | $X_{2n_2 1}$ | $X_{2n_2 2}$ | | $X_{2n_2 k}$ | $Y_{2n_2}$ |

Untuk subgrup pertama, misalkan didapatkan model regresi :

$$Y^{(1)} = \alpha + \beta_1 X_1^{(1)} + \beta_2 X_2^{(1)} + \cdots + \beta_k X_k^{(1)} + \varepsilon^{(1)} \quad (3.2)$$

Untuk subgrup kedua, misalkan didapatkan model regresi :

$$Y^{(2)} = \alpha + \beta_1^* X_1^{(2)} + \beta_2^* X_2^{(2)} + \cdots + \beta_k^* X_k^{(2)} + \varepsilon^{(2)} \quad (3.3)$$

Penaksiran dari $\beta_1, \beta_2, \ldots, \beta_k$ dan $\beta_1^*, \beta_2^*, \ldots, \beta_k^*$ dilakukan dengan metode OLS. Berdasarkan model (3.2) dan (3.3) akan diuji apakah koefisien regresi dari kedua model regresi tersebut sama atau tidak.

Dengan hipotesis sebagai berikut :

$$H_0: \beta_1, \beta_2, \ldots, \beta_k = \beta_1^*, \beta_2^*, \ldots, \beta_k^*$$

$$H_1: \beta_1, \beta_2, \ldots, \beta_k \neq \beta_1^*, \beta_2^*, \ldots, \beta_k^* \quad (3.4)$$

Dari model persamaan (3.1) dapat dicari *Sum of Square Error (SSE)* dimana

$$SSE = \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2$$

Tabel ANAVA untuk model persamaan (3.1) adalah sebagai berikut :

Tabel 3.2 Anava untuk Data Gabungan Subgrup 1 dan Subgrup 2

| Sumber Variasi | *Sum of Square* | Derajat Bebas | *Mean Square* | $F_0$ |
|---|---|---|---|---|
| Regresi | $SS_{Reg}$ | 1 | $MS_{Reg} = \frac{SS_{Reg}}{1}$ | $\frac{MS_{Reg}}{MSE}$ |
| Residual(Galat) | $SS_E$ | $n-2$ | $MSE = \frac{SS_E}{n-2}$ | |
| Total | $SS_T$ | $n-1$ | | |

Dari model regresi linear untuk subgrup 1 pada persamaan (3.2) dapat dicari $SSE_1$, yaitu :

$$SSE_1 = \sum_{i=1}^{n} (Y_{1i} - \hat{Y}_{1i})^2$$

Tabel ANAVA untuk model persamaan (3.2) adalah sebagai berikut :

Tabel 3.3 Anava untuk Data Subgrup 1 dengan $n_1$ Pengamatan

| Sumber Variasi | *Sum of Square* | Derajat Bebas | *Mean Square* | $F_0$ |
|---|---|---|---|---|
| Regresi | $SS_{Reg1}$ | 1 | $MS_{Reg1} = \frac{SS_{Reg1}}{1}$ | $\frac{MS_{Reg1}}{MSE_1}$ |
| Residual(Galat) | $SS_{E1}$ | $n_1 - 2$ | $MSE_1 = \frac{SS_{E_1}}{n_1 - 2}$ | |
| Total | $SS_{T1}$ | $n_1 - 1$ | | |

Dari model regresi linear untuk subgrup 2 pada persamaan (3.3) dapat dicari $SSE_2$, yaitu :

$$SSE_2 = \sum_{i=1}^{n_2} (Y_{2i} - \hat{Y}_{2i})^{2}$$

Tabel ANAVA untuk model persamaan (3.3) adalah sebagai berikut :

Tabel 3.4 Anava untuk Data Subgrup 2 dengan $n_2$ Pengamatan

| Sumber Variasi | *Sum of Square* | Derajat Bebas | *Mean Square* | $F_0$ |
|---|---|---|---|---|
| Regresi | $SS_{Reg2}$ | 1 | $MS_{Reg2} = \frac{SS_{Reg2}}{1}$ | $\frac{MS_{Reg2}}{MSE_2}$ |
| Residual(Galat) | $SS_{E2}$ | $n_2 - 2$ | $MSE_2 = \frac{SS_{E_2}}{n_2 - 2}$ | |
| Total | $SS_{T2}$ | $n_2 - 1$ | | |

$\frac{SSE}{\sigma^2} \sim \chi^2_{n-p}$ , $p =$ banyaknya parameter dalam model.

Misalkan parameter pada regresi linear berganda ada 4 buah, maka diperoleh hasil:

1. Untuk data gabungan subgrup 1 dan subgrup 2

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n-4} \tag{3.5}$$

Karena $n = n_1 + n_2$ maka $\frac{SSE}{\sigma^2} \sim \chi^2_{(n_{1_1}+n_2)-4}$ (3.6)

2. Untuk data subgrup 1

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n_1-4} \tag{3.7}$$

3. Untuk data subgrup 2

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n_2-4} \tag{3.8}$$

Berdasarkan sifat 1 pada subbab 2.5 maka,

$$\frac{SSE_1}{\sigma^2} + \frac{SSE_2}{\sigma^2} = \frac{SSE_1 + SSE_2}{\sigma^2} \sim \chi^2_{(n_1-4)+(n_2-4)}$$

$$\frac{(SSE_1 + SSE_2)}{\sigma^2} \sim \chi^2_{n_1+n_2-8} \tag{3.9}$$

Berdasarkan sifat 2 pada subbab 2.5 maka,

$$\frac{SSE}{\sigma^2} - \frac{SSE_1+SSE_2}{\sigma^2} = \frac{SSE-SSE_1+SSE_2}{\sigma^2} \sim \chi^2_{(n_1+n_2-4)-(n_1+n_2-8)}$$

$$\frac{SSE - (SSE_1 + SSE_2)}{\sigma^2} \sim \chi^2_4 \tag{3.10}$$

Dengan demikian, diperoleh statistik uji untuk pengujian hipotesis pada persamaan (3.4) adalah :

$$F = \frac{(SSE - (SSE_1 + SSE_2))\big/4}{(SSE_1 + SSE_2)\big/n_1 + n_2 - 8} \sim F_{4,n_1+n_2-8} \tag{3.11}$$

Dari persamaan (3.4) jika $\beta_1, \beta_2, \ldots, \beta_k = \beta_1^*, \beta_2^*, \ldots, \beta_k^* = \beta$ diperoleh :

$$SSE = \sum_{j=1}^{n_i} \sum_{i=1}^{2} (Y_{ij} - \hat{Y}_{ij})^2$$

$$= \sum_{j=1}^{n_1} \left(Y_{1j} - \hat{Y}_{1j}\right)^2 + \sum_{j=1}^{n_2} \left(Y_{2j} - \hat{Y}_{2j}\right)^2$$

$$= \sum_{j=1}^{n_1} \left(Y_{1j} - X_{1j}\hat{\beta}\right)^2 + \sum_{j=1}^{n_2} \left(Y_{2j} - X_{2j}\hat{\beta}\right)^2$$

$$= SSE_1 + SSE_2$$

Di sisi lain, jika $\beta_1, \beta_2, \dots, \beta_k \neq \beta_1^*, \beta_2^*, \dots, \beta_k^*$ maka $SSE > SSE_1 + SSE_2$ semakin besar selisih antara $\beta_1, \beta_2, \dots, \beta_k$ dan $\beta_1^*, \beta_2^*, \dots, \beta_k^*$ maka $SSE - (SSE_1 + SSE_2)$ akan semakin besar, begitu juga dengan nilai $F$ akan semakin besar. Maka untuk menguji $\beta_1, \beta_2, \dots, \beta_k = \beta_1^*, \beta_2^*, \dots, \beta_k^*$ dapat digunakan statistik uji di atas.

Penolakan $H_0$ mengimplikasikan bahwa $\beta_1, \beta_2, \dots, \beta_k \neq \beta_1^*, \beta_2^*, \dots, \beta_k^*$, artinya model regresi linear berganda untuk kedua subgrup dari $z$ tidak berbeda, maka $z$ bukan merupakan variabel moderator.

Jika $z$ bukan merupakan variabel moderator, analisis regresi linear berganda dapat dilakukan seperti biasa. Tetapi jika $z$ merupakan variabel moderator maka analisis regresi linear berganda dilakukan secara terpisah untuk setiap kategori.