BAB II

MATERI PENUNJANG

2.1 Heteroskedastisitas

Salah satu asumsi utama dalam model regresi adalah homoskedastisitas, yaitu :

$$E(u_i^2) = \sigma^2 \qquad i = 1, 2, \ldots, n$$

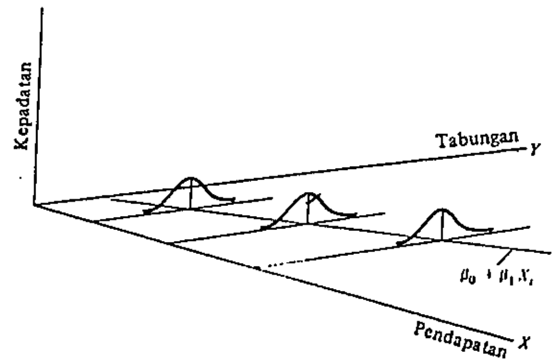
asumsi ini menyatakan bahwa varians dari faktor pengganggu pada model regresi adalah sama untuk semua nilai X (atau semua percobaan).

Jika asumsi ini dilanggar, akan terjadi kasus heteroskedastisitas atau varians tidak sama untuk semua nilai X,

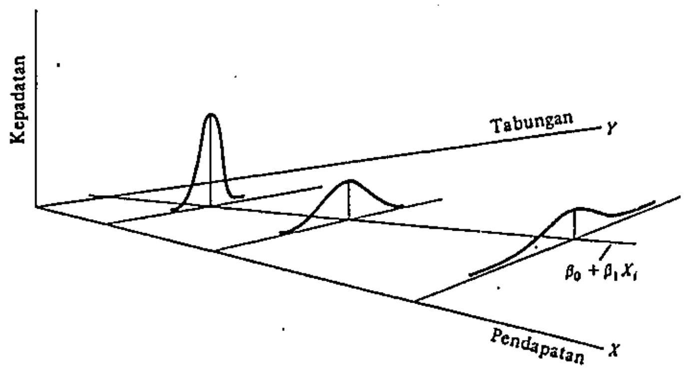
$$E(u_i^2) = \sigma_i^2 \qquad i = 1, 2, \ldots, n$$

Dalam bentuk diagram, dengan menggunakan regresi satu-variabel $Y_i = \beta_0 + \beta_1 X_i + u_i$, Y_i menyatakan tabungan ke-i dan X_i menyatakan pendapatan ke-i. Gambar 1 menunjukkan kasus homoskedastisitas. Dimana varians bersyarat dari Y_i bergantung pada nilai X_i, tetap sama tanpa melihat nilai yang diambil untuk variabel X.

Sebaliknya, Gambar 2 menunjukkan bahwa varians bersyarat dari Y_i meningkat dengan meningkatnya X_i. Disini varians Y_i tidak sama, terjadi heteroskedastisitas. Dengan kata lain, varian tabungan meningkat bersama-sama dengan peningkatan pendapatan.

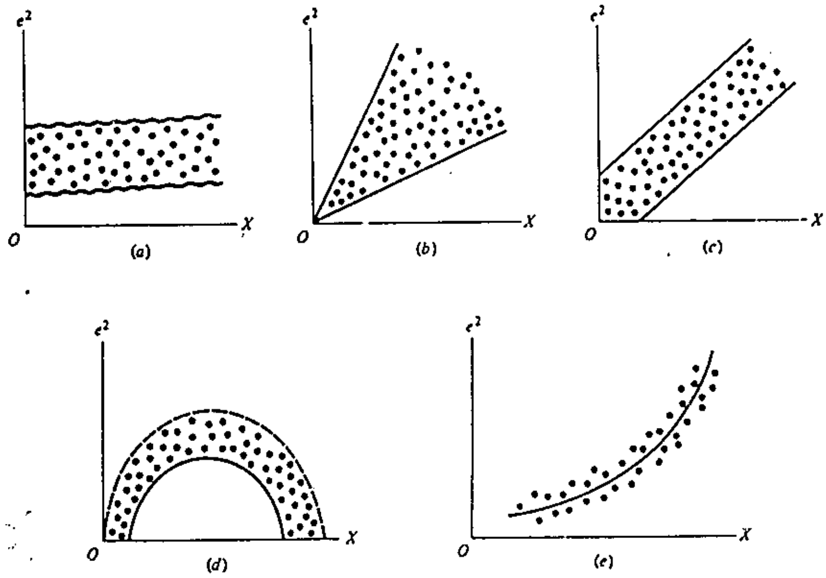


Gambar 1 Gangguan homoskedastik



Gambar 2 Gangguan heteroskedastik

Beberapa pola heteroskedastisitas dapat dilihat pada gambar 3, yaitu dengan memetakan e_i^2 terhadap variabel independent X.



Gambar 3 Diagram pencar residual kuadrat yang ditaksir terhadap X

Dalam gambar 3a tidak terdapat heteroskedastisitas. Sedangkan gambar 3b hingga 3e menunjukkan adanya heteroskedastisitas dalam model regresi.

Jika masalah heteroskedastisitas tidak dikoreksi, penggunaan hasil regresi akan berpengaruh pada ketelitian dari interval kepercayaan dan pengujian hipotesis. Heteroskedastisitas tidak mempengaruhi sifat tak bias dan konsistensi dari koefisien regresi, namun berpengaruh pada standard error. Yang mana lebih besar dari standard error pada situasi homoskedastisitas, dan karenanya tidak effisien.

2.2. Efek Heteroskedastisitas Pada Penaksir Kuadrat Terkecil

Persamaan (1) adalah bentuk ringkas untuk sekumpulan n persamaan simultan berikut :

$$Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + ... + \beta_{k-1} X_{k-1,1} + u_1$$

$$Y_2 = \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + ... + \beta_{k-1} X_{k-1,2} + u_2$$

$$Y_3 = \beta_0 + \beta_1 X_{13} + \beta_2 X_{23} + ... + \beta_{k-1} X_{k-1,3} + u_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots \qquad (2.1)$$

$$Y_n = \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + ... + \beta_{k-1} X_{k-1,n} + u_n$$

dalam bentuk matriks :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k-1,1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k-1,2} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 1 & X_{1n} & X_{2n} & \cdots & X_{k-1,n} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_{k-1} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix} \quad ....(2.2)$$

$$\underset{n \times 1}{Y} = \underset{n \times k}{X} \; \underset{k \times 1}{\beta} + \underset{n \times 1}{u} \quad ....(2.3)$$

dimana : Y = vektor pengamatan nx1

X = matrik pengamatan nxk

β = vektor kolom dari kx1 parameter yang tak diketahui $\beta_0, \beta_1, ..., \beta_{k-1}$

u = error berukuran nx1

Asumsi : 1. $E(u) = 0$

yang berarti:

$$E(u) = E\begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix} = \begin{bmatrix} E(u_1) \\ E(u_2) \\ . \\ . \\ . \\ E(u_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

2. $E(uu') = \sigma^2 I$

bukti :

$$E(uu') = E\begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix}\begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}$$

$$= E\begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & \cdots & u_2u_n \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= \sigma^2 I$$

di mana I adalah matriks identitas nxn

Bila ada heteroskedastisitas maka asumsi diatas menjadi :

$$E(uu') = E \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \ldots & u_n \end{bmatrix}$$

$$= E \begin{bmatrix} u_1^2 & u_1 u_2 & \cdots & u_1 u_n \\ u_2 u_1 & u_2^2 & \cdots & u_2 u_n \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ u_n u_1 & u_n u_2 & \cdots & u_n^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2). \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

$$E(uu') = \Omega \qquad \text{...}(2.4)$$

dimana Ω adalah matrik definit positip yaitu untuk setiap $\mathbf{x} \neq 0$ berlaku $\mathbf{x}'\Omega\mathbf{x} > 0$, dan $\mathbf{x}'\Omega\mathbf{x} = 0$ untuk $x = 0$.

Secara rinci dapat dijelaskan sebagai berikut:

Diberikan sebarang $\mathbf{x} \neq 0$, maka:

$$\mathbf{x}'\Omega\mathbf{x} = [x_1\ x_2\ \ldots\ x_n] \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

$$= [x_1\sigma_1^2\ \ x_2\sigma_2^2\ \ \ldots\ \ x_n\sigma_n^2] \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} \sigma_i^2 x_i^2$$

akan dihasilkan $\mathbf{x}'\Omega\mathbf{x} > 0$.

Karena matrik simetrik Ω adalah matrik definit positip, maka ada matriks nonsingular nxn sedemikian sehingga,

$$H\Omega H' = I$$

$$H^{-1}(H\Omega H') = H^{-1}I$$

$$\Omega H' = H^{-1}$$

$$\Omega H'(H')^{-1} = H^{-1}(H')^{-1}$$

$$\Omega = H^{-1}(H')^{-1} = (H'H)^{-1}$$

sehingga

$$\Omega^{-1} = H'H \qquad \text{...(2.5)}$$

Dalam kasus regresi linier dengan k-1 variabel bebas, penaksir Kuadrat Terkecil diperoleh dengan meminimumkan :

$$\Sigma u_i^2 = \Sigma(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i} - \ldots - \beta_{k-1} X_{k-1,i})^2$$

dimana Σu_i^2 adalah jumlah kudrat sesatan.

Dalam notasi matriks,

$$u'u = \begin{bmatrix} u_1 & u_2 & \ldots & u_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ . \\ . \\ u_n \end{bmatrix} = u_1^2 + u_2^2 + \ldots + u_n^2 = \Sigma u_i^2$$

dari persamaan (2.3) kita peroleh

$$u = Y - X\beta \quad \text{...}(2.6)$$

sehingga

$$u'u = (Y - X\beta)'(Y - X\beta)$$

$$= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$

$$= Y'Y - 2\beta'X'Y + \beta'X'X\beta \quad \text{.................................}(2.7)$$

dimana digunakan sifat-sifat transpose suatu matriks, yaitu $(X\beta)' = \beta'X'$ dan karena $\beta'X'Y$ adalah suatu skalar (suatu angka real), bentuk itu sama dengan transposenya $Y'X\beta$.

Turunkan persamaan (2.7) secara parsial terhadap β dan samakan hasil yang diperoleh dengan nol.

$$\frac{\partial(uu')}{\partial\beta} = -2X'Y + 2X'X\beta$$

$$\frac{\partial^2(uu')}{\partial\beta^2} = 2X'X \quad ====> \quad \text{minimum}$$

maka

$$\frac{\partial(u'u)}{\partial\beta} = 0 \quad ====> \quad -2X'Y + 2X'X\hat{\beta} = 0$$

$$(X'X)\hat{\beta} = (X'Y)$$

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{.............................}(2.8)$$

Untuk matriks varians-kovarians dari $\hat{\beta}$,

$$\hat{\beta} = (X'X)^{-1}X'Y$$

substitusikan $Y = X\beta + u$, didapat :

$$\hat{\beta} = (X'X)^{-1} X'(X\beta + u)$$

$$= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'u$$

$$= \beta + (X'X)^{-1} X'u \quad \text{.....................................(2.9)}$$

Oleh karena itu,

$$\hat{\beta} - \beta = (X'X)^{-1} X'u \quad \text{...(2.10)}$$

Sehingga,

$$\text{var}-\text{cov}(\hat{\beta}) = \mathrm{E}\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right]$$

$$= \mathrm{E}\left[\left((X'X)^{-1} X'u\right)\left((X'X)^{-1} X'u\right)'\right]$$

$$= \mathrm{E}\left[(X'X)^{-1} X'uu'X(X'X)^{-1}\right]$$

$$= (X'X)^{-1} X'\mathrm{E}(uu')X(X'X)^{-1}$$

$$= (X'X)^{-1} X'\sigma^2 X(X'X)^{-1}$$

$$= \sigma^2 (X'X)^{-1} \quad \text{...(2.11)}$$

Bila terdapat heteroskedastisitas maka matriks varians-kovarians dari $\hat{\beta}$ adalah:

$$\text{var}-\text{cov}(\hat{\beta}) = (X'X)^{-1} X'\Omega X(X'X)^{-1} \quad \text{.................................(2.12)}$$

Penerapan metode Kuadrat Terkecil pada model regresi linier yang mengandung heteroskedastisitas akan menghasilkan penaksir yang mempunyai sifat :

1. Tak Bias

   yaitu $E(\hat{\beta}) = \beta$.

   Dari persamaan (2.9) dapat dibuktikan,

$$\begin{aligned} E(\hat{\beta}) &= E\left[\beta + (X'X)^{-1} X'u\right] \\ &= \beta + (X'X)^{-1} X' E(u) \qquad \text{karena } E(u) = 0 \text{ maka} \\ &= \beta + 0 \\ &= \beta \end{aligned}$$

2. Konsisten

   yaitu dengan meningkatnya ukuran sampel secara tak terbatas, penaksir mengarah ke nilai populasi yang sebenarnya. Jika $\hat{\beta}_n$ adalah sebuah estimator β berdasarkan sampel random yang besarnya n, maka $\hat{\beta}_n$ adalah konsisten untuk β jika :

$$\lim_{n\to\infty} P\left(\left|\hat{\beta} - \beta\right| < \varepsilon\right) = 1.$$

Biasanya sukar untuk membuktikan suatu estimator adalah konsisten dengan menggunakan persamaan diatas, tetapi estimator yang mempunyai varian cenderung menuju nol dengan sampel yang biasanya mendekati tak terhingga adalah konsisten.

$$\begin{aligned} \lim_{n\to\infty} \text{var-cov}(\hat{\beta}) &= \lim_{n\to\infty} (X'X)^{-1} X'\Omega X (X'X)^{-1} \\ &= \lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'\Omega X\right)\left(\frac{1}{n}X'X\right)^{-1} \\ &= 0 \end{aligned}$$

3. Variansi tidak lagi minimum

Model regresi mengasumsikan bahwa varian dari faktor pengganggu konstan, karenanya prosedur metode Kuadrat Terkecil yang biasa akan menghasilkan estimator dari asumsi kekonstanan ini. Jika terdapat heteroskedastisitas, varian dari faktor pengganggu tidak sama, bahkan sebaliknya varian ini juga merupakan sebuah variabel yang karenanya tidak lagi minimum.

2.3. Efek heteroskedastisitas pada pengujian hipotesis

Sebuah bagian penting dalam perkiraan pada model regresi multiple linier adalah pengujian hipotesis secara statistik mengenai parameter model dan membentuk interval keyakinan tertentu. Pengujian hipotesis secara statistik dapat dinyatakan sebagai menguji kecocokan suatu pengamatan atau penemuan terhadap suatu hipotesis. Dalam bahasa statistik, hipotesis tersebut dikenal sebagai hipotesis nol dan dinyatakan dengan lambang H_0. Hipotesis nol biasanya diuji terhadap hipotesis alternatif, dinyatakan dengan H_1. Statistik ujinya :

$$t = \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \qquad \text{...(2.13)}$$

dimana $se(\hat{\beta}_i)$ adalah kesalahan standar dari $\hat{\beta}_i$.

Hipotesis H_0 akan ditolak jika $|t_{hit}| > t_{tabel}$.

Dalam bahasa pengujian tingkat arti, suatu statistik dikatakan penting secara statistik (statistically significant) kalau nilai statistik uji terletak dalam daerah kritis, hipotesis H_0 ditolak. Sebaliknya suatu pengujian dikatakan secara statistik tidak penting (statistically insignificant) kalau nilai statistik uji terletak dalam daerah penerimaan. Hipotesis H_0 diterima.

Karena matriks varians-kovarians $\hat{\beta}$ dengan heteroskedastisitas akan lebih besar daripada matriks varians-kovarians dengan homoskedastisitas, maka formula kuadrat Terkecil yang biasa (2.11) akan menaksir terlalu rendah (underestimate) varians dan kesalahan standar dari $\hat{\beta}$ yang diberikan oleh (2.12). Sehingga persamaan (2.11) akan menaksir nilai t yang terlalu tinggi. Akibatnya kita sangat mungkin untuk membesar-besarkan signifikan statistik dari parameter yang ditaksir, artinya taksiran mungkin bisa dinyatakan penting secara statistik, padahal pada kenyataannya, taksiran tidak signifikan.

Selanjutnya dapat ditentukan interval kepercayaan dua arah $100(1-\alpha)$ persen untuk β_i, yaitu:

$$\Pr\left(-t_{\alpha/2,n-k-1} \le t \le t_{\alpha/2,n-k-1}\right) = 1-\alpha \quad \text{..........................(2.14)}$$

substitusikan persamaan (2.13) pada persamaan (2.14),

$$\Pr\left(-t_{\alpha/2,n-k-1} \le \frac{\hat{\beta}_i - \beta_i}{se(\hat{\beta}_i)} \le t_{\alpha/2,n-k-1}\right) = 1-\alpha$$

$$\Pr\left(-t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i) \le \hat{\beta}_i - \beta_i \le t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i)\right) = 1-\alpha$$

$$\Pr\left(-t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i) - \hat{\beta}_i \le -\beta_i \le t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i) - \hat{\beta}_i\right) = 1-\alpha$$

$$\Pr\left(\hat{\beta}_i - t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i) \le \beta_i \le \hat{\beta}_i + t_{\alpha/2,n-k-1}\, se(\hat{\beta}_i)\right) = 1-\alpha \quad \text{.........(2.15)}$$

yang akan menjadi sempit (memberikan kesan ketepatan yang tinggi), pada kenyataannya, selang tadi jauh lebih lebar.

Dalam bahasa pengujian hipotesis, selang keyakinan $100(1-\alpha)$ persen yang ditetapkan dalam (2.15) dikenal sebagai daerah penerimaan (dari hipotesis nol) dan daerah selang keyakinan disebut daerah penolakan (dari H_0) atau daerah kritis. Titik ujung selang keyakinan disebut nilai-nilai kritis.