BAB II
TINJAUAN PUSTAKA

2.1 Analisis Regresi

Apabila memperhatikan ketergantungan antara suatu peubah respon Y terhadap peubah bebas X yang bervariasi, maka persamaan yang menghubungkan Y dan X disebut persamaan regresi. Suatu persamaan regresi diasumsikan bahwa peubah bebas (X) bersifat tetap atau *fixed*, sedangkan peubah respon (Y) bersifat acak (Drapper dan Smith, 1992).

Berikut merupakan persamaan garis lurus antara dua peubah, yaitu peubah Y dan X :

$$Y = \beta_0 + \beta_1 X \tag{2.1}$$

persamaan regresi linier (2.1) dengan peubah respon Y dan peubah bebas X, menggambarkan hubungan antara nilai peubah Y dan X, di mana besar nilai peubah Y dipengaruhi oleh besar nilai peubah X dengan β_0 adalah *intercept* dan β_1 adalah *slope* (koefisien regresi). Dalam banyak kasus nilai dari koefisien β_0 tidak selalu diartikan sebagai besar nilai Y apabila X bernilai nol, akan tetapi sebagai koefisien penyesuai. β_1 menyatakan besarnya perubahan nilai Y pada setiap satuan perubahan nilai X, sehingga bisa dinyatakan sebagai berikut :

$$\beta_1 = \frac{\Delta Y}{\Delta X} \tag{2.2}$$

Hubungan antara peubah X dan Y tidak selalu tepat mengkuti persamaan linier (2.1), tetapi terdapat keragaman pada setiap nilai peubah bebas. Pada x_i misalnya, dapat dihasilkan bermacam-macam nilai *output* Y yang pusatnya diperkirakan terletak pada titik $Y_i = \beta_0 + \beta_1 X_i$. Adanya perbedaan nilai pengamatan dari garis regresi, maka Persamaaan (2.1) harus dimodifikasi. Dimisalkan perbedaan antara nilai amatan dari Y dengan $\hat{Y}$ adalah galat (ε_i) yaitu peubah acak yang menjelaskan kesalahan dari persamaan dalam mencocokkan data dengan tepat. Oleh karena itu, persamaan untuk menjelaskan sebuah garis regresi adalah :

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{2.3}$$

2.2 Uji Asumsi Klasik

Dalam analisis regresi terdapat beberapa asumsi yang harus dipenuhi agar persamaan regresi yang didapat baik dan mampu menggambarkan data yang sebenarnya sehingga menghasilkan kesimpulan yang sahih dan tidak menyesatkan. Asumsi-asumsi yang harus dipenuhi dalam analisis regresi adalah :

1. Asumsi Linieritas

   Linieritas dapat diartikan bahwa antar peubah mempunyai pola hubungan linier $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$. Jika hubungan antara peubah tidak linier, hasil analisis regresi akan di bawah duga. Untuk mendeteksi linieritas dapat digunakan beberapa metode yaitu (Montgomery and Peck, 1991) :

   a. Metode Grafik

      Melalui metode ini, pemeriksaan dilakukan melalui diagram pencar untuk melihat apakah hubungan antar peubah linier atau tidak. Diagram pencar menunjukkan hubungan antara nilai prediksi dengan galat yang menunjukkan pola acak.

   b. Perbandingan nilai R^2

      Nilai R^2 adalah persentase kemampuan peubah bebas menerangkan peubah respon. Semakin besar nilai R^2, maka semakin baik persamaan yang terbentuk.

   c. Uji F

      Statistik Uji F dapat dijadikan sebagai salah satu pertimbangan untuk menentukan apakah data yang digunakan adalah linier atau tidak. Untuk mengetahui apakah antar peubah mempunyai pola hubungan linier atau tidak, dilakukan pengujian terhadap hipotesis :

      H_0: model linier
      H_1: model tidak linier

      Statistik uji F didapatkan dari rumus :

$$F_{hitung} = \frac{Kuadrat\ \ Tenga\,h\ Linieritas}{Kuadrat\ \ Tenga\,h\ Galat} \qquad (2.4)$$

di mana :

$$KT_L = \frac{JK_L}{db_L} = \frac{\sum(Y_i - \hat{Y})^2 - \sum_x \left\{\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}\right\}}{k-2} \qquad (2.5)$$

$$KT_G = \frac{JK_G}{db_G} = \frac{\sum_x \left\{ \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \right\}}{n-k} \quad (2.6)$$

Jika $F_{hitung} < F_{tabel}$ maka model hubungan yang teridentifikasi adalah linier.

2. Asumsi Kenormalan Galat

Galat atau ε_i merupakan peubah acak normal dengan nilai tengah (*mean*) nol dan ragam sebesar σ^2, atau $E(\varepsilon_i) = 0$, $var(\varepsilon_i) = \sigma^2$. Kenormalan dari galat tersebut berpengaruh terhadap hasil dari pendugaan parameter, yaitu menyebabkan keputusan yang di bawah duga atau kelebihan duga terhadap taraf nyata percobaan atau α. Untuk mengetahui apakah galat menyebar normal atau tidak, dilakukan pengujian terhadap hipotesis :

H_0: galat menyebar normal
H_1: galat tidak menyebar normal

Salah satu metode yang digunakan adalah uji *Anderson-Darling* dengan statistik uji :

$$A^2 = -N - \left(\frac{1}{N}\right) \sum (2i-1)(lnF(Y_i) + \ln(1 - F(Y_{N+1-i}))) \quad (2.7)$$

di mana A menyebar *Anderson Darling* dan F adalah fungsi sebaran kumulatif dari sebaran normal dan Y_i adalah nilai pengamatan ke-i.

Kriteria yang digunakan dalam pengambilan keputusan adalah

$$A^2 \begin{cases} \le A_\alpha;\ H_0\ diterima \\ > A_\alpha;\ H_0\ ditolak \end{cases} \quad (2.8)$$

(Drapper dan Smith, 1992).

3. Asumsi Autokorelasi

Satu dari asumsi penting dari model regresi linier klasik adalah bahwa galat ε_{it} yang masuk ke dalam fungsi regresi populasi adalah random atau tak berkorelasi. Jika asumsi ini dilanggar, maka terdapat masalah autokorelasi (Nachrowi dan Usman, 2006). Untuk mengetahui ada tidaknya autokorelasi pada model, dapat dilihat dari statistik uji *Durbin-Watson*, dengan hipotesis sebagai berikut :

H_0 : tidak ada autokorelasi baik positif atau korelasi negatif

H_1 : Ada autokorelasi baik positif atau korelasi negatif

Pengujian dilakukan dengan membandingkan statistik uji *Durbin-Watson* (d) dengan nilai batas atas (*Upper Limit*/ d_U) dan nilai batas bawah (*Lower Limit*/ d_L) pada tabel dengan banyaknya pengamatan n, banyaknya peubah bebas k dan tingkat signifikansi (taraf nyata) α, dengan kriteria pengujian sebagai berikut :

- $d < d_L$: Tolak H_0 (terdapat autokorelasi positif)
- $d \geq 4\text{-}d_L$: Tolak H_0 (terdapat autokorelasi negatif)
- $d_U < d < 4\text{-} d_U$: Terima H_0 (tidak terdapat autokorelasi)
- $d_L \leq d \leq d_U$ atau $4\text{-} d_U \leq d \leq 4\text{-} d_L$

Di mana pada umumnya statistik uji *Durbin-Watson* dirumuskan sebagai berikut :

$$d = \frac{\sum_{i=2}^{N}(\hat{\varepsilon}_i - \hat{\varepsilon}_{(i-1)})}{\sum_{i=1}^{N}\hat{\varepsilon}_i} \qquad (2.9)$$

4. Asumsi Multikolinieritas

Multikolinieritas adalah adanya hubungan linier antara peubah bebas dalam model regresi. Salah satu cara yang digunakan untuk mendeteksi adanya masalah ini adalah dengan melihat nilai VIF (*Variance Inflation Factor*), apabila nilai VIF $\leq$ 10 disimpulkan tidak terjadi muktikolinieritas, namun jika nikai VIF > 10 maka disimpulkan terjadi multikolinieritas. Di mana $VIF_j = \frac{1}{(1-{R_j}^2)}$, dan R_j^2 adalah koefisien determinasi antara peubah bebas ke-j dengan peubah bebas lainnya (Nachrowi dan Usman, 2006).

5. Asumsi Homoskedastisitas

Menurut Gujarati (1991), homoskedastisitas adalah keadaan di mana masing-masing galat mempunyai ragam yang sama atau konstan, yaitu: $var(\varepsilon_{it}) = \sigma^2$. Konsekuensi adanya heteroskedastisitas adalah nilai penduga parameter regresi yang didapatkan dari OLS tidak lagi efisien. Salah satu cara untuk mengetahui ada tidaknya heteroskedastisitas, adalah melalui uji *White*.

Pendeteksian kehomogenan ragam galat dengan menggunakan uji *White* diperlukan adanya *auxiliary regression*. Fungsi dari *auxiliary regression* tersebut adalah untuk menghasilkan nilai R^2. *Auxiliary regression* dilakukan dengan meregresikan galat kuadrat ($\hat{\varepsilon}_i^2$) dengan peubah bebas, peubah bebas yang dikuadratkan, dan hasil kali silang peubah bebas.

Sebagai ilustrasi, pertimbangkan model regresi pada persamaan berikut :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \qquad (2.10)$$

Auxiliary regression dapat dilakukan setelah persamaan tersebut diduga dengan menggunakan Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Square* (OLS). Galat yang diperoleh dikuadratkan dan diregresikan berdasarkan persamaan regresi berikut :

$$\hat{\varepsilon}_i^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \cdots + \alpha_k X_{ki} + \alpha_{k+1} {X_{1i}}^2 + \alpha_{k+2} {X_{2i}}^2 + \cdots + \alpha_{k+k} {X_{k+1}}^2 + \alpha_{k+k+1} X_{1i} X_{2i} + \alpha_{k+k+2} X_{1i} X_{3i} + \cdots + \alpha_{k+k+{}_kC_2} X_{ki} X_{ki} + v_i \qquad (2.11)$$

di mana :

k = banyak peubah bebas

${}_kC_2$ = banyak kombinasi hasil kali silang peubah bebas

Dari *auxiliary regression* akan diperoleh nilai R^2 (koefisien determinasi) yang akan dijadikan sebagai statistik uji dalam uji *White*.

2.3 Konsep Dasar Homoskedastisitas

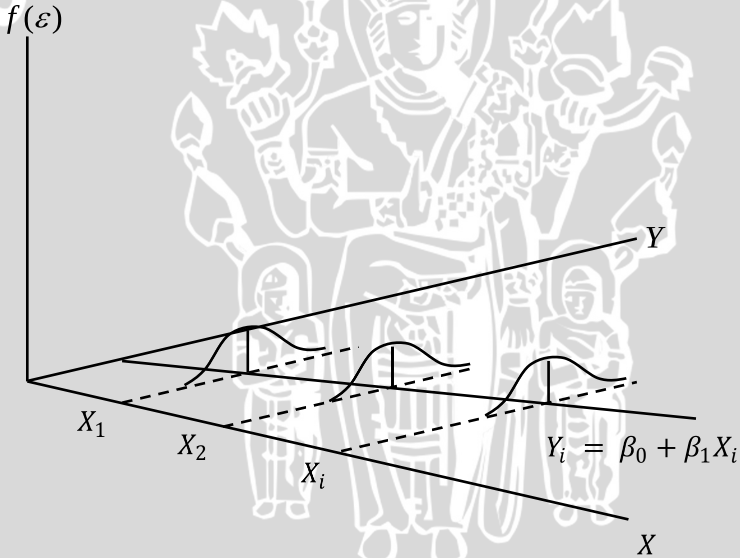
Salah satu asumsi utama dalam model regresi adalah homoskedastisitas yang menyatakan bahwa ragam galat ε_i, adalah sama atau konstan untuk semua nilai X, sebesar σ^2. Secara sistemastis dapat dituliskan :

$$\begin{aligned} var(\varepsilon_i) &= E(\varepsilon_i - E(\varepsilon_i)^2) \\ &= E({\varepsilon_i}^2) - E(\varepsilon_i)^2 \\ &= E({\varepsilon_i}^2) - 0 \\ &= E({\varepsilon_i}^2) \end{aligned}$$

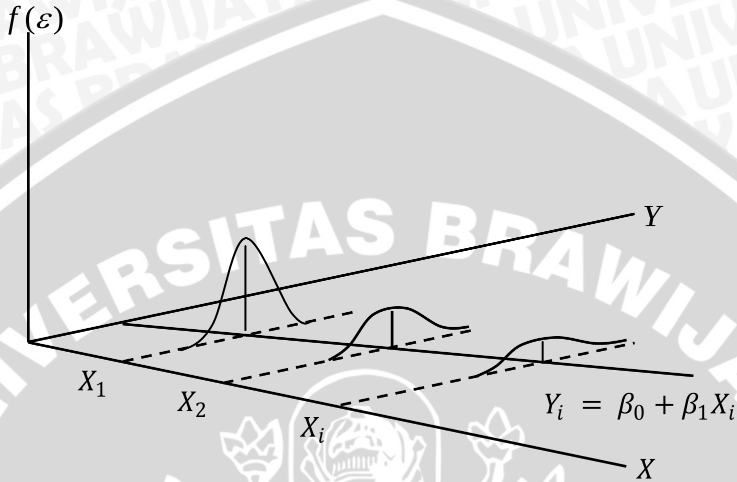
$$= \sigma^2 \tag{2.12}$$

di mana $E(\varepsilon_i) = 0$, berarti ragam dari ε_i konstan sebesar σ^2 dari pengamatan satu ke pengamatan yang lain. Jadi berapapun nilai X, ragamnya tidak berubah. Kondisi sebaliknya yaitu bila ragam galat $var(\varepsilon_i) = E({\varepsilon_i}^2) = \sigma_i^2$ untuk setiap nilai i, dinamakan heteroskedastisitas di mana ragam galat ε_i tidak lagi konstan (Gujarati, 2004).

Dalam bentuk diagram, dengan menggunakan regresi satu peubah $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$. Gambar 2.1 menunjukkan kondisi homoskedastisitas, di mana ragam bersyarat dari Y_i bergantung pada nilai X_i, tetap sama tanpa melihat nilai yang diambil untuk peubah X. Sebaliknya, Gambar 2.2 menunjukkan bahwa ragam bersyarat dari Y_i meningkat dengan meningkatnya X_i, sehingga ragam tidak sama yang menunjukkan kondisi heteroskedastisitas.

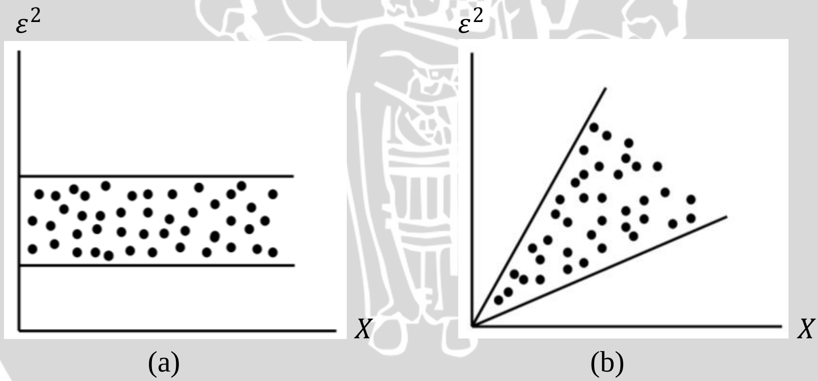


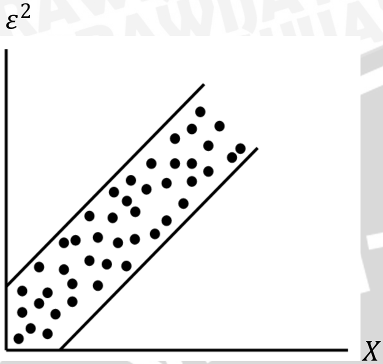
Gambar 2.1. Kondisi Homoskedastisitas



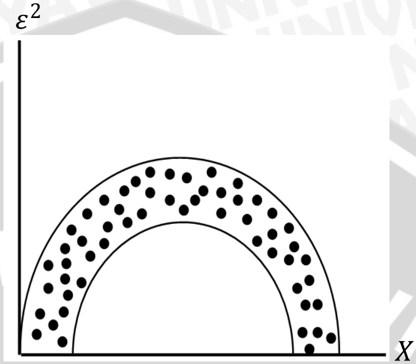
Gambar 2.2. Kondisi Heteroskedastisitas

Beberapa pola heteroskedastisitas dapat dilihat pada Gambar 2.3, yaitu dengan memetakan ε^2 dengan peubah bebas X.

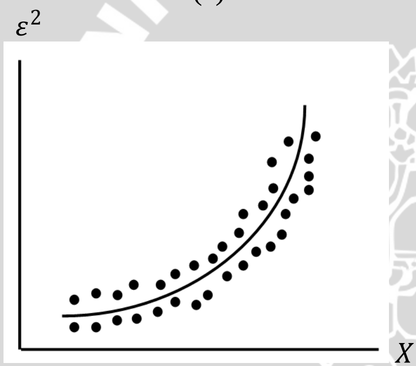




(c)



(d)



(e)

Gambar 2.3 Diagram pencar residual kuadrat yang ditaksir tehadap *X*
Sumber : Gujarati, 2006

Dalam Gambar 2.3(a) tidak terdapat heteroskedastisitas karena tidak terdapat pola dalam plot, sedangkan Gambar 2.3(b) hingga 2.3(e) menunjukkan adanya pola dalam plot sehingga terdapat heteroskedastisitas dalam model regresi.

Menurut Sumodiningrat (1994), jika masalah heteroskedastisitas tidak dikoreksi, penggunaan hasil regresi akan berpengaruh pada ketelitian interval kepercayaan dan pengujian hipotesis. Heteroskedastisitas tidak mempengaruhi sifat tak bias dan konsistensi dari koefisien regresi, namun berpengaruh pada *Standard Error* yang mana nilai *Standard Error* pada situasi heteroskedastisitas lebih besar daripada situasi homoskedastisitas, sehingga tidak efisien.

2.4 *Ordinary Least Square* (OLS)

Pendugaan parameter untuk persamaan regresi linier sederhana biasa menggunakan OLS di mana nilai dari Jumlah Kuadrat Galat (JKG) akan menjadi minimum. Drapper dan Smith (1992) menyatakan bahwa pola hubungan antara peubah Y dan X yang bersifat linier dan sederhana dapat dimodelkan seperti Persamaan (2.3) di mana $i = 1, 2, 3, \dots, n$, sehingga JKG pada pengamatan-pengamatan garis regresi sebenarnya adalah :

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 \qquad (2.13)$$

Sebagai nilai dugaan β_0 dan β_1, dipilih nilai $\hat{\beta}_0$ dan $\hat{\beta}_1$. Untuk menentukan nilai $\hat{\beta}_0$ dan $\hat{\beta}_1$, Persamaan (2.13) dideferensial parsial terhadap β_0 dan β_1 dan kemudian menyamakan hasil diferensial dengan nol, sehingga :

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i) \qquad (2.14)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^{n} X_i (Y_i - \beta_0 - \beta_1 X_i) \qquad (2.15)$$

sehingga nilai dugaan $\hat{\beta}_0$ dan $\hat{\beta}_1$ dapat diperoleh dari :

$$\sum_{i=1}^{n} Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} X_i = 0$$
$$\sum_{i=1}^{n} X_i Y_i - \hat{\beta}_0 \sum_{i=1}^{n} X_i - \hat{\beta}_1 \sum_{i=1}^{n} {X_i}^2 = 0 \qquad (2.16)$$

atau

$$\hat{\beta}_0 n + \hat{\beta}_1 \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i$$
$$\hat{\beta}_0 \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} {X_i}^2 = \sum_{i=1}^{n} X_i Y_i \qquad (2.17)$$

Persamaan (2.17) disebut dengan persamaan normal.

Penyelesaian dari Persamaan (2.17) untuk menentukan nilai $\hat{\beta}_0$ dan $\hat{\beta}_1$ adalah :

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \qquad (2.18)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i - \left[\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right) \right]}{\sum_{i=1}^{n} {X_i}^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \qquad (2.19)$$

Model dugaan untuk regresi linier sederhana adalah :

$$\begin{aligned} \hat{Y} &= \hat{\beta}_0 + \hat{\beta}_1 X \\ &= \bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 X \\ &= \bar{Y} - \hat{\beta}_1 (\bar{X} - X) \end{aligned} \qquad (2.20)$$

Berdasarkan Persamaan (2.20), OLS terkait dengan konsep rata-rata karena garis regresi tepat melalui titik $(\bar{X}, \bar{Y})$.

2.4.1 Pendugaan *Ordinary Least Square* (OLS)

Salah satu metode yang dapat digunakan untuk menduga parameter bagi persamaan regresi adalah Metode Kuadrat Terkecil (MKT) atau *Ordinary Least Square* (OLS). Ada beberapa asumsi-asumsi terhadap ε_i agar hasil pengujian bagi parameter-parameter dalam regresi linier sederhana bersifat sahih, antara lain :

1. $(\varepsilon_i) = 0$, nilai tengah galat adalah nol untuk semua i $(i = 1,2,\ldots,n)$
2. $var(\varepsilon_i) = \sigma^2$, ragam galat konstan sebesar σ^2 untuk semua i $(i = 1,2,\ldots,n)$, dikenal dengan homoskedastisitas.
3. $cov\left(\varepsilon_i, \varepsilon_j\right) = 0$, tidak terdapat korelasi antar galat untuk $i \neq j$, dikenal dengan autokorelasi.
4. $\varepsilon_i \sim N(0,\sigma^2)$, galat menyebar normal dengan rata-rata 0 dan ragam σ^2.

Bila asumsi-asumsi ini terpenuhi, maka OLS akan menghasilkan pendugaan parameter yang memberikan jumlah kuadrat simpangan atau galat yang sekecil-kecilnya (Kmenta, 1990).

Yitnosumarto (1990) menyebutkan bahwa sifat-sifat penduga parameter dengan OLS antara lain :

1. Tak Bias

   Penduga parameter dikatakan tak bias jika penduga mendekati nilai sebenarnya dari parameter yang diduga. Misalkan $\hat{\beta}$ penduga tak bias dari parameter β maka :

   $$E\left(\hat{\beta}\right) = \beta \qquad (2.21)$$

2. Efisien

   Penduga parameter dikatakan efisien jika ragamnya lebih kecil daripada ragam penduga parameter yang lain. Misalkan $\hat{\beta}_0$ dengan ragam $var(\hat{\beta}_0)$ dan $\hat{\beta}_1$ dengan ragam $var(\hat{\beta}_1)$, adalah penduga untuk parameter tak bias untuk β, maka $\hat{\beta}_0$ dikatakan lebih efisien dari $\hat{\beta}_1$ apabila :

   $$var\left(\hat{\beta}_0\right) \leq var(\hat{\beta}_1) \qquad (2.22)$$

3. Konsisten

   Penduga parameter dikatakan konsisten jika semakin mendekati parameter yang diduga dengan semakin besarnya ukuran contoh n. Misal $\hat{\beta}$, merupakan penduga yang konsisten jika $\hat{\beta} \rightarrow \theta$ dengan $n \rightarrow \sim$. Secara matematis dapat dituliskan :

$$\lim_{n\to\infty} P\left(\left|\hat{\beta}_n - \beta\right| < \varepsilon\right) = 1 \qquad (2.23)$$

2.4.2 Pendugaan OLS pada Keberadaan Heteroskedastisitas

Konsekuensi dari keadaan heteroskedastisitas adalah batas-batas keyakinan dan uji hipotesis tidak akan bisa diterapkan. Artinya jika ragam dari penduga model tidak memenuhi asumsi homoskedatisitas, maka pendugaan mengenai koefisien populasi keliru atau tidak valid. Selain itu, peramalan yang didasarkan pada pendugaan tersebut juga akan tidak bias. Tetapi karena penduga yang digunakan tidak efisien, maka peramalan juga akan tidak efisien (Ramanathan, 1992).

Berikut ini akan diperlihatkan apakah OLS masih tetap memberikan penduga yang memenuhi sifat-sifat yang baik apabila terjadi heteroskedastisitas.

a. Tak bias

Menurut Ghosh (1991), penduga $\widehat{\boldsymbol{\beta}}$ dikatakan penduga tak bias dari parameter β kalau nilai harapan $\widehat{\boldsymbol{\beta}}$ sama dengan parameter $\boldsymbol{\beta}$, yaitu $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$.

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{Y}) \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'(\boldsymbol{X\beta} + \boldsymbol{\varepsilon}) \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{X})\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{\varepsilon}) \\
&= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{\varepsilon}) \qquad (2.24)
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{E}(\widehat{\boldsymbol{\beta}}) &= \boldsymbol{E}[\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\boldsymbol{\varepsilon})] \\
&= \boldsymbol{E}(\boldsymbol{\beta}) + (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{E}(\boldsymbol{\varepsilon}) \\
&= \boldsymbol{\beta} + \mathbf{0} \\
&= \boldsymbol{\beta} \qquad (2.25)
\end{aligned}$$

Sifat ketidakbiasan tidak memerlukan bahwa galat, ε_i, adalah homoskedastisitas sehingga meskipun dibawah kondisi heteroskedastisitas, OLS memberikan penduga yang tidak bias.

b. Efisien

Menurut Ghosh (1991), jika asumsi klasik dari model regresi linier terpenuhi, maka,

$$\boldsymbol{var(\varepsilon)} = \boldsymbol{E(\varepsilon\varepsilon')} = \boldsymbol{\sigma^2 I} = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

(2.26)

Sedangkan, jika tidak memenuhi asumsi homoskedastisitas, maka,

$$\boldsymbol{var(\varepsilon)} = \boldsymbol{E(\varepsilon\varepsilon')} = \boldsymbol{\sigma^2 \Omega} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} \omega_1 & 0 & 0 & \dots & 0 \\ 0 & \omega_2 & 0 & \dots & 0 \\ 0 & 0 & \omega_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \omega_n \end{bmatrix} \qquad (2.27)$$

di mana $\sigma_i^2 = \sigma^2\ \omega_i$, i = 1, 2, …, n, sehingga

$$\boldsymbol{var(\hat{\beta})} = \boldsymbol{E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\,]}$$
$$= \boldsymbol{E\{[(X'X)^{-1}X'Y - \beta][(X'X)^{-1}X'Y - \beta]'\}}$$
$$= \boldsymbol{E\{[(X'X)^{-1}X'(X\beta + \varepsilon) - \beta][(X'X)^{-1}X'(X\beta + \varepsilon - \beta'\}}$$
$$= \boldsymbol{E\{[(X'X)^{-1}(X'X)\beta + (X'X)^{-1}X'\varepsilon - \beta X'X-1X'X\beta + X'X-1X'\varepsilon - \beta'\}}$$
$$= \boldsymbol{E\{[(X'X)^{-1}X'\varepsilon][(X'X)^{-1}X'\varepsilon]'\}}$$
$$= \boldsymbol{E\{[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}\}}$$
$$= \boldsymbol{(X'X)^{-1}X'E(\varepsilon'\varepsilon)X(X'X)^{-1}} \qquad (2.28)$$
$$\boldsymbol{var(\hat{\beta})} = \boldsymbol{(X'X)^{-1}X'\sigma^2\Omega X(X'X)^{-1}}$$
$$= \boldsymbol{\sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}}$$
$$= \boldsymbol{var(\hat{\beta}_{homo})X'\Omega X(X'X)^{-1}} \qquad (2.29)$$

di mana $\boldsymbol{var}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{homo}})$ adalah ragam $\widehat{\boldsymbol{\beta}}$ dengan asumsi homoskedastisitas. Jadi ragam $\widehat{\boldsymbol{\beta}}$ dengan heteroskedastisitas akan lebih besar daripada ragam $\widehat{\boldsymbol{\beta}}$ dengan homoskedastisitas. Berarti dibawah kondisi heteroskedastisitas, penduga parameter dengan menggunakan OLS tidak lagi bersifat efisien.

c. Konsisten

OLS dikatakan konsisten apabila $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ dan $\lim_{n\to\infty} var(\widehat{\boldsymbol{\beta}}) = 0$

$\lim_{n\to\infty} \sigma^2 (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} = 0$

$$\lim_{n\to\infty} \frac{\sigma^2}{n}\left(\frac{(\boldsymbol{X}'\boldsymbol{X})}{n}\right)^{-1}\left(\frac{\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}}{n}\right)\left(\frac{(\boldsymbol{X}'\boldsymbol{X})}{n}\right)^{-1} = 0 \qquad (2.30)$$

Jika $\lim_{n\to\infty} \frac{\sigma^2}{n} = 0,$ (2.31)

$\lim_{n\to\infty}\left(\frac{(\boldsymbol{X}'\boldsymbol{X})}{n}\right)^{-1}$ dan $\lim_{n\to\infty}\left(\frac{\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}}{n}\right)$ nilainya terbatas, maka $\widehat{\boldsymbol{\beta}}$ merupakan penduga yang konsisten (Daniel, 2002).

Menurut Gujarati dan Porter (2010), penduga *Ordinary Least Square* (OLS) dan nilai ragamnya jika memiliki kondisi heteroskedastisitas dengan menganggap $E({\varepsilon_i}^2) = \sigma_i^2$, tetapi tetap mempertahankan semua asumsi model klasik lainnya adalah dengan memperhatikan kembali model dua peubah berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (2.32)$$

sehingga didapat penduga OLS dari β_1 adalah

$$\beta_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} \qquad (2.33)$$

akan tetapi, kini ragamnya diekspresikan sebagai berikut :

$$\begin{aligned} var(\hat{\beta}_1) &= E(k_1^2u_1^2 + k_2^2u_2^2 + \cdots + k_n^2u_n^2 + 2\,produk\ silang) \\ &= E(k_1^2u_1^2 + k_2^2u_2^2 + \cdots + k_n^2u_n^2) \qquad (2.34) \end{aligned}$$

oleh karena ekspektasi dari produk silang adalah nol akibat dari asumsi tidak ada korelasi serial atau autokorelasi, maka

$$\begin{aligned} var(\hat{\beta}_1) &= k_1^2E(u_1^2) + k_2^2E(u_2^2) + \cdots + k_n^2E(u_n^2) \\ &= k_1^2\sigma_1^2 + k_2^2\sigma_2^2 + \cdots + k_n^2\sigma_n^2 \qquad (2.35) \end{aligned}$$

$$var(\hat{\beta}_1) = \sum k_i^2\sigma_i^2 = \sum\left[\left(\frac{x_i}{\sum x_i^2}\right)^2 \sigma_i^2\right] = \frac{\sum x_i^2\sigma_i^2}{(\sum x_i^2)^2} \qquad (2.36)$$

di mana jelas berbeda dari formula ragam umum yang didapatkan di bawah asumsi homoskedastisitas, yaitu

$$var(\hat{\beta}_1) = \frac{\sigma^2}{\sum x_i^2} \qquad (2.37)$$

$\hat{\beta}_1$ adalah penduga terbaik, linier, dan tidak bias atau bersifat *Best Linear Unbiased Estimator* (BLUE).

2.5 Uji Parsial Model Regresi

Uji parsial digunakan untuk menguji apakah sebuah peubah bebas X benar-benar memberikan kontribusi terhadap peubah respon Y. Dalam pengujian ini ingin diketahui apakah jika secara terpisah, suatu peubah bebas X masih memberikan kontribusi secara signifikan terhadap peubah respon Y.

Hipotesis untuk uji ini adalah :

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

Uji parsial ini menggunakan uji t, yaitu :

$$t_{hitung} = \left|\frac{\hat{\beta}_1 - \beta_1}{se\ (\hat{\beta}_1)}\right| \sim t_{tabel}\ (n-2) \qquad (2.38)$$

jika $t_{hitung} \leq t_{tabel}\ (n-2)$, maka terima H_0.

jika $t_{hitung} > t_{tabel}\ (n-2)$, maka tolak H_0,

Apabila H_0 ditolak, maka peubah bebas X tersebut memiliki kontribusi yang signifikan terhadap peubah respon Y (Montgomery and Peck, 1991).

2.6 Uji Simultan Model Regresi

Uji simultan (bersama-sama) pada konsep regresi linier adalah pengujian mengenai apakah model regresi yang didapatkan benar-benar dapat diterima. Uji simultan bertujuan untuk menguji apakah antara peubah bebas X dan respon Y, atau setidak-tidaknya antara salah satu peubah X dengan peubah respon Y, benar-benar terdapat hubungan linier (*linear relation*).

Hipotesis yang berlaku untuk pengujian ini adalah :

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

Penjabaran secara hitungan untuk uji simultan ini dapat ditemui pada Tabel 2.1.

Tabel 2.1 Tabel ANOVA (*Analysis Of Variance*) Uji Simultan Regresi Linier

SK	db	JK	KT	F_{hitung}
$Regresi$	1	JK_{Reg}	KT_{Reg}	KT_{Reg}/KT_R
$Residual$	$n-2$	JK_R	KT_R	
$Total$	$n-1$	JK_T		

di mana :

$$JK_{Reg} = \hat{\beta}_1 \left(\sum_{i=1}^{n} x_i\, y_i - \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i)}{n} \right) \quad (2.39)$$

$$JK_R = \sum_{i=1}^{n} y_i^2 - \frac{(\sum_{i=1}^{n} y_i)^2}{n} - \hat{\beta}_1 \left(\sum_{i=1}^{n} x_i\, y_i - \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i)}{n} \right) \quad (2.40)$$

$$JK_T = \sum_{i=1}^{n} y_i^2 - \frac{(\sum_{i=1}^{n} y_i)^2}{n} \quad (2.41)$$

$$KT_{Reg} = \frac{JK_{Reg}}{db_{Reg}} = \frac{JK_{Reg}}{1} \quad (2.42)$$

$$KT_R = \frac{JK_R}{db_R} = \frac{JK_R}{n-1} \quad (2.43)$$

Di dalam tabel ANOVA akan ditemui nilai F_{hitung} , di mana :

jika $F_{hitung} \leq F_{tabel}\ (db_{Reg}, db_R)$, maka terima H_0,

jika $F_{hitung} > F_{tabel}\ (db_{Reg}, db_R)$, maka tolak H_0.

Apabila H_0 ditolak, maka model regresi yang diperoleh dapat digunakan (Montgomery and Peck, 1991).

Menurut Gujarati (1991), terdapat hubungan antara uji F dengan uji t pada analisis regresi linier sederhana, dapat ditunjukkan bahwa kuadrat nilai t dengan derajat kebebasan $N-k$ adalah nilai F dengan derajat kebebasan 1 dan $N-k$. Sehingga pada analisis regresi linier sederhana, nilai t^2 sama dengan nilai F.

2.7 Uji Heteroskedastisitas *White*

Metode *White* merupakan salah satu metode yang digunakan untuk mendeteksi kehomogenan ragam galat. Metode ini mudah diterapkan daripada metode lain. Fomby (2006) menunjukkan bahwa Metode *White* sudah dipelajari sejak 1980 dan metode ini sering disebut dengan uji *White*. Uji ini mengunakan galat dalam pengujiannya. Pendeteksian kehomogenan ragam galat dengan menggunakan uji *White* tidak memerlukan asumsi yang spesifik (Jeeshim, 2003).

Uji *White* sering digunakan pada sampel besar yaitu sampel yang mempunyai banyak pengamatan (n) lebih dari atau sama dengan tiga puluh. Dalam uji *White* menggunakan *auxiliary regression* yaitu dengan meregresikan galat kuadrat (${\hat{\varepsilon}_i}^2$) dengan semua peubah bebas, peubah bebas yang dikuadratkan, dan hasil kali silang peubah bebas. Kemudian dari regresi tersebut didapatkan nilai R^2 yang digunakan sebagai statistik uji (Ramanathan, 1992).

Koefisien determinasi (R^2) adalah ukuran kecukupan dari sebuah model. Koefisien determinasi mengukur keragaman di sekitar nilai tengah yang dapat dijelaskan oleh regresi tersebut. Koefisien determinasi dirumuskan dengan :

$$R^2 = \frac{JKR}{JKT} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \qquad (2.44)$$

di mana JKR adalah jumlah kuadrat regresi, dan JKT adalah jumlah kuadrat total (Montgomery and Peck, 1991).

Pengujian umum yang dikemukakan oleh *White* tidak bergantung asumsi normalitas dan mudah untuk diimplementasikan. Sebagai ilustrasi dari ide dasarnya, kita perhatikan model regresi dengan satu peubah bebas :

$$Y_i = \beta_0 + \beta_1 X_{1i} \qquad (2.45)$$

Prosedur pengujian uji *White* adalah sebagai berikut :

Langkah 1. Dari hasil pendugaan model regresi pada Persamaan (2.45), didapatkan nilai residualnya, $\hat{\varepsilon}_i$.

Langkah 2. Kemudian dilakukan *auxiliary regression* sebagai berikut :

$${\hat{\varepsilon}_i}^2 = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 {X_{1i}}^2 \qquad (2.46)$$

Hal ini berarti bahwa residual yang dikuadratkan dari regresi awal data diregresikan terhadap peubah X awal, nilai-nilai X yang dikuadratkan, dan produk silang dari peubah X. Regresi dengan pangkat yang lebih tinggi juga dapat digunakan. Kemudian dapatkan nilai R^2 dari regresi pada Persamaan (2.46).

Langkah 3. Hipotesis pendeteksian kehomogenan ragam galat dengan menggunakan uji *White* adalah sebagai berikut :

H_0 : $\alpha_1 = 0$ (kondisi homoskedastisitas) vs

H_1 : $\alpha_1 \neq 0$ (kondisi heteroskedastisitas)

Apabila H_0 benar maka statistik uji *White* merupakan perkalian dari banyaknya pengamatan (n) dengan nilai R^2 yang diperoleh dari *auxiliary regression* pada Persamaan (2.46), yaitu :

$$\chi^{2*} = n.R^2 \sim \chi^2_m \tag{2.47}$$

di mana m adalah banyak peubah bebas dalam *auxiliary regression*.
Langkah 4. Kaidah pengambilan keputusan dengan menggunakan tabel kritis χ^2 yaitu:

$$\chi^{2*} \begin{cases} \leq \chi^2_{m(\alpha)}, H_0\ diterima \\ > \chi^2_{m(\alpha)}, H_0\ ditolak \end{cases} \tag{2.48}$$

Jika nilai *chi-square* yang didapatkan dari Persamaan (2.47) lebih besar dari nilai *chi-square* kritis pada tingkat signifikansi yang dipilih, maka terdapat heteroskedastisitas. Sebaliknya, jika nilainya lebih kecil dari *chi-square* kritis, maka tidak terdapat heteroskedastisitas (Gujarati, 2004).

2.8 *White's Robust Standard Error*

Metode *White's Robust Standard Error* digunakan apabila σ_i^2 tidak diketahui dan untuk memperoleh penduga ragam dan peragam dari penduga OLS yang konsisten jika terdapat masalah heteroskedastisitas. *White* telah menunjukkan bahwa pendugaan ini dapat dilakukan sehingga, secara asimtotik, inferensi statistik yang valid mengenai nilai-nilai parameter yang sebenarnya dapat dibuat. Metode ini juga dikenal dengan nama *White's Robust Standard Error*.

Sebagai ilustrasi dari ide dasar metode *White's Robust Standard Error*, perhatikan model regresi dua peubah berikut :

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{2.49}$$

$$var\ (\varepsilon_i) = \sigma_i^2 \tag{2.50}$$

Seperti ditujukkan pada Persamaan (2.36),

$$var(\hat{\beta}_1) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2} \tag{2.51}$$

Oleh karena σ_i^2 tidak dapat diobservasi secara langsung, White menyarankan untuk menggunakan ε_i^2, residual kuadrat untuk setiap i, untuk menggantikan σ_i^2 dan menduga $var(\hat{\beta}_1)$ sebagai berikut:

$$var(\hat{\beta}_1) = \frac{\sum x_i^2 \varepsilon_i^2}{(\sum x_i^2)^2} \tag{2.52}$$

White telah menunjukkan bahwa Persamaan (2.52) merupakan penduga yang konsisten dari Persamaan (2.51), yaitu seiring dengan meningkatnya sampel secara tidak hingga, Persamaan (2.52) mengerucut ke Persamaan (2.51) (Gujarati dan Porter, 2010).

Hubungan antara σ_i^2 dengan satu atau beberapa peubah bebas dapat diperkirakan karena residual kuadrat ε_i^2 diperoleh dari regresi pada metode OLS yang merupakan penduga bagi σ_i^2 apabila fungsi regresi ini telah sesuai. Ragam dari residual ε_i apabila terdapat masalah heteroskedastisitas adalah σ_i^2, dapat dinyatakan sebagai berikut :

$$\sigma_i^2 = E\{\varepsilon_i^2\} - (E\{\varepsilon_i\})^2 \tag{2.53}$$

di mana $E\{\varepsilon_i\} = 0$ sesuai dengan model regresi, dapat diperoleh :

$$\sigma_i^2 = E\{\varepsilon_i^2\} \tag{2.54}$$

Oleh karena itu, residual kuadrat ε_i^2 merupakan penduga bagi σ_i^2.

Pada kondisi heteroskedastisitas, penduga koefisien regresi yang dihasilkan dari metode OLS tetap bersifat tak bias dan konsisten, tetapi tidak lagi mempunyai ragam penduga yang mininum. Sehingga rumus ragam penduganya dapat ditulis sebagai betikut :

$$\boldsymbol{var(\hat{\beta})} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'\sigma^2\Omega X}(\boldsymbol{X'X})^{-1} \tag{2.55}$$

Jika ragam galat tidak sama dan tidak diketahui, penduga yang sesuai dari $\boldsymbol{var(\hat{\beta})}$ dapat diduga menggunakan metode *White's Robust Standard Error*, dengan rumus sebagai berikut :

$$\boldsymbol{var(\hat{\beta})} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'S_0X}(\boldsymbol{X'X})^{-1} \tag{2.56}$$

di mana :

$$\boldsymbol{S_0} = \begin{bmatrix} \varepsilon_1^2 & 0 & \dots & 0 \\ 0 & \varepsilon_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varepsilon_n^2 \end{bmatrix} \tag{2.57}$$

Nilai $\varepsilon_i, \dots, \varepsilon_n$ adalah penduga residual dari metode OLS. Metode *White's Robust Standard Error* dapat digunakan untuk membuat kesimpulan tentang parameter regresi yang sesuai jika didasarkan pada metode OLS, tanpa harus menentukan bentuk ragam kesalahan yang tidak konstan (Kurtner, M.H., C.J. Nachtsheim, J. Neter, and W. Li, 2005).

Menurut Asteriou and Hall (2007), nilai *Standard Error* dari penduga koefisien metode *Ordinary Least Square* (OLS) pada masalah heteroskedastisitas adalah tidak benar, sehingga pada metode *White's Robust Standard Error* dilakukan koreksi terhadap nilai *Standard Error*. Nilai *Standard Error* yang dihasilkan dapat lebih besar atau lebih kecil daripada metode OLS. Untuk menghasilkan penduga yang lebih akurat, nilai *Standard Error* pada metode *White's Robust Standard Error* seharusnya lebih besar daripada *Standard Error* pada metode OLS. Peningkatan nilai *Standard Error* membuat nilai pada uji t atau F menurun serta selang kepercayaan menjadi lebih lebar yang mencerminkan sifat data yang sebenarnya. Oleh karena itu, hasil dari metode *White's Robust Standard Error* akan menghasilkan penduga yang lebih baik dan menghasilkan informasi yang valid.

2.9 Kriteria Keberartian Model

2.9.1 *Standard Error* (SE)

Standard Error adalah galat dari $\hat{Y}$. Ukuran ini dipakai sebagai salah satu kriteria evaluasi model, dengan rumus :

$$SE = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon_i^2}{n-k}} = \sqrt{KTG} \qquad (2.58)$$

di mana :

ε_i : galat pada observasi ke-i ($i = 1, 2, \dots, n$).

n : banyaknya observasi.

k : banyaknya parameter regresi yang diduga.

KTG : Kuadrat Tengah Galat penduga model regresi.

KTG dihitung dengan rumus $KTG = \frac{JKG}{db\ galat}$ (Gujarati, 2004).

2.9.2 *P-Value*

P-value merupakan taraf keberartian terkecil sehingga nilai statistik uji yang diperoleh masih berarti dan dihitung berdasarkan peluang eksak berlandaskan statistik uji yang digunakan (Walpole, 1995). *P-value* menunjukkan tingkat terkecil untuk menolak H_0. Secara umum, jika *P-value* kurang dari atau sama dengan α maka H_0 akan ditolak. Sedangkan jika *P-value* melampui α maka H_0 tidak dapat ditolak.

Menurut Hines dan Montgomery (1990), keuntungan dari *P-value* adalah tidak hanya menentukan hasil dalam suatu pengambilan keputusan tentang H_0, tetapi juga memberikan informasi tentang kuatnya keputusan yang diambil. Selain itu, dengan *P-value* juga dapat diketahui besarnya resiko salah secara eksak dalam pengambilan keputusan. Sehingga dalam berbagai pengujian, *P-*

value sering digunakan sebagai indikator dalam pengambilan keputusan. Semakin kecil *P-value* maka semakin kecil peluang untuk membuat kesalahan dengan menolak H_0.

P-value yang sangat kecil, misal 0.001 menunjukkan bahwa terdapat kemungkinan kecil untuk menolak H_0 yang benar yaitu sebesar 0.001. *P-value* dapat diperoleh dengan menghitung :

$$\begin{aligned} P - value &= P[|t| \geq t_{hitung}] \\ &= P\left[t \leq -t_{hitung}\right] + P\left[t \geq t_{hitung}\right] \\ &= 2 \times P\left[t \geq t_{hitung}\right] \qquad (2.59) \end{aligned}$$

Kemudian *P-value* dibandingkan dengan α. Mengenai besarnya nilai α, yang dapat digunakan sebagai pembanding bagi *P-value,* tergantung pada seberapa besar resiko salah yang masih ditolerir menurut jenis penelitian yang dilakukan (Johnson and Bhattacharyya, 1985).