BAB II
REGRESI LINIER SEDERHANA

Model regresi linier sederhana adalah model dengan satu variabel bebas dan satu variabel tak bebas, serta hubungannya antara kedua variabel tersebut dalam sebuah garis lurus.

Dalam regresi linier sederhana diasumsikan bahwa tiap kesalahan (ε_i) dalam fungsi regresi populasi;

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Didistribusikan secara normal dengan,

$$\text{Rata - rata} \quad : E(\varepsilon_i) = 0$$

$$\text{Varians} \quad : E(\varepsilon_i) = \sigma^2$$

$$\text{Cov}(\varepsilon_i, \varepsilon_j) : E(\varepsilon_i\, \varepsilon_j) = 0 \qquad i \neq j$$

Asumsi ini secara ringkas dinyatakan sebagai :

$$\varepsilon_i \sim N(0, \sigma^2)$$

2.1. PENAKSIRAN UNTUK PARAMETER DAN VARIANS

Metode kuadrat terkecil biasa (OLS = Ordinary Least Squares) digunakan untuk menaksir α dan β , yaitu menaksir α dan β sehingga jumlah kuadrat antara jarak hasil observasi (Y_i) dengan garis lurus yang dicari harus sekecil mungkin (minimum).

$$Y_i = \hat{\alpha} + \hat{\beta} X_i + e_i \qquad , i = 1,2,\ldots,n$$

$$S(e_i^2) = \Sigma(Y_i - \hat{\alpha} - \hat{\beta} X_i)^2$$

Taksiran kuadrat terkecil dari α dan β , katakanlah $\hat{\alpha}$ dan $\hat{\beta}$ harus memenuhi :

$$\frac{\partial S(e_i^2)}{\partial \hat{\alpha}} = -2\Sigma(Y_i - \hat{\alpha} - \hat{\beta} X_i) = 0$$

Dan

$$\frac{\partial S(e_i^2)}{\partial \hat{\beta}} = -2\Sigma(Y_i - \hat{\alpha} - \hat{\beta} X_i)(X_i) = 0$$

Dari dua persamaan diatas akan didapat :

$$\left.\begin{aligned} n\hat{\alpha} + \hat{\beta}\,\Sigma X_i &= \Sigma Y_i \\ \hat{\alpha}\,\Sigma X_i + \hat{\beta}\,\Sigma X_i^2 &= \Sigma Y_i X_i \end{aligned}\right\} \quad \ldots\ldots\ldots (2.1)$$

Persamaan (2.1) dinamakan persamaan persamaan normal kuadrat terkecil, yang mempunyai penyelesaian yaitu :

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\,\overline{X} \quad \ldots\ldots(2.2)$$

dan

$$\hat{\beta} = \frac{\Sigma Y_i X_i - \dfrac{(\Sigma Y_i)(\Sigma X_i)}{n}}{\Sigma X_i^2 - \dfrac{(\Sigma X_i)^2}{n}} \quad \ldots\ldots(2.3)$$

dimana $\overline{Y} = \frac{1}{n}\Sigma Y_i$ dan $\overline{X} = \frac{1}{n}\Sigma X_i$

n = besarnya sampel

bukti :

1. $n\hat{\alpha} + \hat{\beta}\,\Sigma X_i = \Sigma Y_i$

   $\hat{\alpha} + \hat{\beta}\frac{1}{n}\Sigma X_i = \frac{1}{n}\Sigma Y_i$

   $\hat{\alpha} = \overline{Y} - \hat{\beta}\,\overline{X}$

2. $\hat{\alpha}\,\Sigma X_i + \hat{\beta}\,\Sigma X_i^{2} = \Sigma Y_i X_i$

$$(\bar{Y} - \hat{\beta}\,\bar{X})\,\Sigma X_i + \hat{\beta}\,\Sigma X_i^2 = \Sigma Y_i X_i$$

$$\hat{\beta}\left[\Sigma X_i^{\,2} - \frac{(\Sigma X_i)^2}{n}\right] = \Sigma Y_i X_i - \frac{(\Sigma Y_i)(\Sigma X_i)}{n}$$

$$\hat{\beta} = \frac{\Sigma Y_i X_i - \dfrac{(\Sigma X_i)\ (\Sigma Y_i)}{n}}{\Sigma X_i^2 - \dfrac{(\Sigma X_i)^2}{n}}$$

Jadi pengepasan untuk model regresi linier sederhana adalah : $\hat{Y} = \hat{\alpha} + \hat{\beta}\,X$

Penyebut pada persamaan (2.3) adalah jumlah kuadrat kesalahan dari pada X_i, sedangkan pembilangnya disebut jumlah kesalahan hasil kali X_i dan Y_i yang ditulis dalam notasi :

$$S_{xx} = \Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n} = \Sigma\ (X_i - \bar{X})^2$$

misalkan :

$$X_i - \bar{X} = x_i \quad \text{maka}$$

$$S_{xx} = \Sigma X_i^2 - (\Sigma X_i)^2/n = \Sigma(X_i - \bar{X}) = \Sigma x_i^2$$

$$S_{xy} = \Sigma Y_i X_i - \frac{(\Sigma Y_i)(\Sigma X_i)}{n} = \Sigma Y_i\ (X_i - \bar{X})$$

$$= \Sigma(Y_i - \bar{Y})(X_i - \bar{X})$$

misalkan :

$$Y_i - \bar{Y} = y_i \quad \text{maka}$$

$$S_{xy} = \Sigma\ y_i x_i$$

Persamaan (2.3) dapat ditulis menjadi :

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} \quad \text{atau} \quad \hat{\beta} = \frac{\Sigma y_i x_i}{\Sigma\ x_i^{\,2}} \qquad \ldots\ldots\ldots\ldots (2.4)$$

Definisi 1 :

Perbedaan antara hasil observasi Y_i dengan kecocokan nilai yang diharapkan $\hat{Y}_i$ disebut Residual

$e_i = Y_i - \hat{Y}_i \quad , \; i = 1,2,\ldots,n$

Sedangkan untuk penaksiran σ^2 didapat dari jumlah kuadrat kesalahan (residual) :

$$SSE = \Sigma\, e_i^2 = \Sigma\, (Y_i - \hat{Y}_i)^2 \qquad \ldots\ldots\ldots\ldots\ldots(2.5)$$

Substitusikan $\hat{Y}_i = \hat{\alpha} + \hat{\beta}\, X_i$ kedalam persamaan (2.5), sehingga :

$$\begin{aligned} SSE &= \Sigma\, (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 \\ &= \Sigma\, [\, Y_i - (\bar{Y} - \hat{\beta}\,\bar{X}) - \hat{\beta}X_i\,]^2 \\ &= \Sigma\, \{(Y_i - \bar{Y}) - \hat{\beta}\,(X_i - \bar{X})\}^2 \\ &= \Sigma\, \{(Y_i - \bar{Y})^2 - 2\,\hat{\beta}\,(Y_i - \bar{Y})(X_i - \bar{X}) + \hat{\beta}^2\,(X_i - \bar{X})^2\} \\ &= \Sigma\, (Y_i - \bar{Y})^2 - 2\,\hat{\beta}\,\Sigma(Y_i - \bar{Y})(X_i - \bar{X}) + \hat{\beta}^2\,\Sigma(X_i - \bar{X})^2 \\ &= \Sigma\, (Y_i - Y)^2 - 2\,\hat{\beta}\,S_{xy} + \hat{\beta}\,\frac{S_{xy}}{S_{xx}}\,.\,S_{xx} \end{aligned}$$

$$SSE = S_{yy} - \hat{\beta}\,S_{xy} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \;(2.6)$$

Jumlah kuadrat kesalahan (SEE) mempunyai n - 2 derajat kebebasan, sehingga penduga dari σ^2 adalah :

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = MSE$$

2.2. SIFAT - SIFAT PENAKSIR KUADRAT TERKECIL

<u>THEOREMA</u> : GAUSS MARKOV

Jika semua asumsi model regresi liner sederhana dipenuhi, penaksir kuadrat terkecil ($\hat{\alpha}$ dan $\hat{\beta}$) adalah tak bias linier terbaik (Best Linier Unbiased Estimator) yaitu dalam kelas semua penaksir tak bias linier mereka mempunyai varians yang minimum.

Bukti :

Definisi 2 :

Ketidakbiasan (unbiasedness) suatu penaksir $\hat{\theta}$ dikatakan penaksir tak bias θ, jika nilai yang diharapkan dari $\hat{\theta}$ adalah sama dengan θ sebenarnya, yaitu :

$$E (\hat{\theta}) = \theta$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$$

Subtitusikan β dari persamaan (2.4), didapat

$$\hat{\alpha} = \bar{Y} - \frac{\Sigma x_i y_i}{\Sigma x_i} \bar{X}$$

$$\hat{\alpha} = \frac{\Sigma Y_i}{n} - \bar{X} \frac{\Sigma x_i}{\Sigma x_i^2} (Y_i - \bar{Y})$$

$$\hat{\alpha} = \Sigma \left(\frac{1}{n} - \bar{X} \frac{x_i}{\Sigma x_i^2} \right) Y_i \qquad \ldots\ldots\ldots\ldots\ (2.7)$$

karena

$$\bar{X}\ \bar{Y} \frac{\Sigma x_i}{\Sigma x_i^2} = 0$$

sebab $\Sigma x_i = \Sigma (X_i - \bar{X})$

$= \Sigma X_i - \Sigma X_i$

$= 0$

dengan demikian :

$$\hat{\alpha} = \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)(\alpha + \beta X_i + \varepsilon_i)$$

$$= \alpha\,\Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right) + \beta\,\Sigma\left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right) X_i$$

$$+ \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)\varepsilon_i$$

$$\hat{\alpha} = \alpha + 0 + \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)\varepsilon_i \quad \ldots\ldots\ldots\ (2.8)$$

maka akan didapat :

$$\hat{\alpha} = \alpha + \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)\varepsilon_i$$

$$E(\hat{\alpha}) = E\left[\alpha + \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)\varepsilon_i\right]$$

$$= \alpha + \Sigma \left(\frac{1}{n} - \bar{X}\frac{x_i}{\Sigma x_i^2}\right)E(\varepsilon_i)$$

$$= \alpha + 0 = \alpha$$

Jadi $E(\hat{\alpha}) = \alpha$

Terbukti bahwa $\hat{\alpha}$ adalah penaksir tak bias dari α.

Sekarang untuk $\hat{\beta}$:

$$\hat{\beta} = \frac{\Sigma x_i y_i}{\Sigma x_i^2}$$

dengan model regresi sampel :

$$Y_i = \hat{\alpha} + \hat{\beta} X_i + e_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ (2.9A)$$

$$\Sigma Y_i = n\hat{\alpha} + \hat{\beta}\,\Sigma X_i + \Sigma e_i$$

$$\Sigma Y_i = n\hat{\alpha} + \hat{\beta}\,\Sigma X_i \quad \ldots\ldots\ldots\ldots\ldots\ (2.9)$$

Karena :

$$\Sigma e_i = \Sigma (Y_i - \hat{Y}_i) = \Sigma (Y_i - \hat{\alpha} - \hat{\beta}X_i)$$
$$= \Sigma Y_i - n \hat{\alpha} - \hat{\beta} \Sigma X_i$$
$$= \Sigma Y_i - n \overline{Y} + n \hat{\beta}\overline{X} - \hat{\beta} \Sigma X_i$$
$$= \Sigma Y_i - \Sigma Y_i + \hat{\beta} \Sigma X_i - \hat{\beta} \Sigma X_i = 0$$

dengan membagi persamaan (2.9) seluruhnya dengan n, diperoleh :

$$\overline{Y} = \hat{\alpha} + \hat{\beta}\overline{X} \qquad \text{............. (2.10)}$$

kurangkan (2.10) dari (2.9A) diperoleh :

$$Y_i - \overline{Y} = \hat{\beta} (X_i - \overline{X}) + e_i$$
$$y_i = \hat{\beta} x_i + e_i \qquad \text{............. (2.11)}$$

Model regresi populasi :

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad \text{............. (2.12)}$$
$$\Sigma Y_i = n \alpha + \beta \Sigma X_i + \Sigma \varepsilon_i$$
$$\frac{1}{n} \Sigma Y = \alpha + \beta \frac{1}{n} \Sigma X_i + \frac{1}{n} \Sigma \varepsilon_i$$
$$\overline{Y} = \alpha + \beta \overline{X} + \overline{\varepsilon} \qquad \text{............. (2.13)}$$

dengan mengurangkan (2.13) dari (2.12) diperoleh :

$$Y_i - \overline{Y} = \beta (X_i - \overline{X}) + (\varepsilon_i - \overline{\varepsilon})$$
$$y_i = \beta x_i + (\varepsilon_i - \overline{\varepsilon}) \qquad \text{............. (2.14)}$$
$$\beta = \frac{\Sigma x_i y_i}{\Sigma x_i^2}$$

Subtitusikan persamaan (2.14) ke persamaan diatas :

$$\hat{\beta} = \frac{\Sigma x_i (\beta x_i + (\varepsilon_i - \overline{\varepsilon}))}{\Sigma x_i^2}$$

$$\hat{\beta} = \frac{\beta \Sigma x_i^{\,2} + \Sigma x_i \varepsilon_i - \overline{\varepsilon} \Sigma x_i}{\Sigma x_i^2}$$

$$= \beta + \frac{\Sigma\, x_i \varepsilon_i}{\Sigma\, x_i^2} \qquad \text{.......... (2.15)}$$

$$E(\hat{\beta}) = E\left(\beta + \frac{\Sigma\, x_i \varepsilon_i}{\Sigma\, x_i^2}\right)$$

$$= \beta + \frac{\Sigma\, x_i}{\Sigma\, x_i^2}\, E(\varepsilon_i)$$

$$= \beta + 0$$

jadi $E(\hat{\beta}) = \beta$

terlihat juga bahwa $\hat{\beta}$ penaksir tak bias dari β

<u>Definisi 3</u> :

Kelinieran (LINEARITY) suatu penaksir $\hat{\theta}$ dikatakan merupakan penaksir linier dari θ, jika merupakan fungsi linier dari observasi sampel.

Dari persamaan (2.7) kita mempunyai :

$$\hat{\alpha} = \Sigma \left(\frac{1}{n} - \bar{X}\, \frac{x_i}{\Sigma\, x_i^2}\right) Y_i$$

misalkan $k_i = \dfrac{x_i}{\Sigma\, x_i^2}$, maka

$$\hat{\alpha} = \Sigma \left(\frac{1}{n} - \bar{X}\, k_i\right) Y_i \qquad \text{...... (2.16)}$$

Ini menunjukkan bahwa $\hat{\alpha}$ adalah penaksir linier karena merupakan fungsi linier dari Y.

Demikian juga untuk $\hat{\beta}$:

$$\hat{\beta} = \frac{\Sigma\, x_i y_i}{\Sigma\, x_i^2}$$

$$\hat{\beta} = \frac{\Sigma x_i (Y_i - \overline{Y})}{\Sigma x_i^2}$$

$$= \frac{\Sigma x_i Y_i}{\Sigma x_i^2} - \overline{Y} \frac{\Sigma x_i}{\Sigma x_i^2}$$

$$\hat{\beta} = \frac{\Sigma x_i Y_i}{\Sigma x_i^2} \text{, misalkan } k_i = \frac{x_i}{\Sigma x_i^2} \quad \ldots . \ (2.17)$$

$$\hat{\beta} = \Sigma k_i Y_i$$

juga menunjukan bahwa $\hat{\beta}$ adalah penaksir linier karena merupakan fungsi linier dari Y.

Definisi 4 :

Varians minimum $\hat{\theta}_1$ dikatakan sebagai penaksir dengan varians minimum dari θ jika varians dari $\hat{\theta}_1$, kecil dari atau paling besar sama dengan varians dari $\hat{\theta}_2$, yang merupakan penaksir lain manapun dari θ.

Dari persamaan (2.8) kita mempunyai :

$$\hat{\alpha} - \alpha = \Sigma \left(\frac{1}{n} - \overline{X} \frac{x_i}{\Sigma x_i^2} \right) \varepsilon_i$$

Varians $\hat{\alpha}$ adalah :

$$E\,[(\hat{\alpha} - E(\hat{\alpha})]^2 = E\,[\hat{\alpha} - \alpha]^2$$

$$E\,(\hat{\alpha} - \alpha)^2 = E\left[\Sigma\left(\frac{1}{n} - \overline{X}\frac{x_i}{\Sigma x_i^2}\right)\varepsilon_i\right]^2$$

$$= \Sigma \left(\frac{1}{n^2} - \frac{2}{n} \frac{\overline{X} \quad x_i}{\Sigma x_i^2} + \frac{\overline{X}^2 \ x_i^2}{(\Sigma x_i^2)^2} \right) E\,(\varepsilon_i^2)$$

$$= \left[\frac{n}{n^2} - \frac{2}{n}\bar{X}\frac{\Sigma x_i}{\Sigma x_i^2} + \bar{X}^2\frac{\Sigma x_i^2}{(\Sigma x_i^2)^2}\right] E(\varepsilon_i^2)$$

$$= \left[\frac{1}{n} - 0 + \bar{X}^2\frac{\Sigma x_i^2}{(\Sigma x_i^2)^2}\right]\sigma^2$$

$$\mathrm{Var}\ (\hat{\alpha}) = \left[\frac{1}{n} + \frac{\bar{X}^2}{\Sigma x_i^2}\right]\sigma^2 \qquad \text{........ (2.18)}$$

Misalkan α^* penaksir linier yang lain dari α , yaitu

$$\alpha^* = \Sigma\ Y_i\ (\frac{1}{n} - \bar{X}\ w_i)$$

Dimana w_i tidak perlu sama dengan k_i pada persamaan (2.17)

$$E\ (\alpha^*) = E\ [\Sigma\ Y_i (\ \frac{1}{n} - \bar{X}\ w_i\ )]$$

$$= \Sigma\ ((\frac{1}{n} - \bar{X}\ w_i)\ E(Y_i))$$

$$= \Sigma\ (\frac{1}{n} - \bar{X}\ w_i)(\alpha + \beta X_i)$$

$$= \alpha + \beta\ \bar{X} - \alpha\ \bar{X}\ \Sigma w_i + \beta\ \bar{X}\ \Sigma w_i X_i$$

$$= \alpha - \alpha\ \bar{X}\ \Sigma w_i + \beta\ \bar{X}\ (1 - \Sigma w_i X_i)$$

agar α^* tak bias, maka

$$\Sigma\ w_i = 0 \quad \text{dan} \quad \Sigma\ w_i X_i = 1$$

$$\mathrm{var}\ (\alpha^*) = \mathrm{var}\ [\Sigma Y_i (\frac{1}{n} - \bar{X}\ w_i)]$$

$$= \mathrm{var}\ (\frac{\Sigma Y_i}{n}) + \bar{X}^2\ \Sigma w_i^2\ \mathrm{var}\ (Y_i)$$

$$= \frac{\sigma^2}{n} + \bar{X}^2\ \sigma^2\ \Sigma w_i^2$$

$$= \sigma^2\ \left[\frac{1}{n} + \bar{X}^2\ \Sigma w_i^2\right]$$

$$= \sigma^2\ \left[\frac{1}{n} + \bar{X}^2\ \Sigma\left(w_i - \frac{x_i}{\Sigma x_i^2} + \frac{x_i}{\Sigma x_i^2}\right)^2\right]$$

$$= \sigma^2 \left[\frac{1}{n} + \bar{X}^2 \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^2}\right)^2 + \bar{X}^2 \frac{\Sigma x_i^2}{(\Sigma x_i^2)^2} + \right.$$

$$\left. 2 \bar{X}^2 \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^2}\right)\left(\frac{x_i}{\Sigma x_i{}^2}\right) \right]$$

sedangkan suku ke-3 ruas kanan pada persamaan diatas :

$$2 \bar{X}^2 \Sigma \left(w_i - \frac{x_i}{\Sigma x_i{}^2} \right) \left(\frac{x_i}{\Sigma x_i{}^2} \right) = 2 \bar{X}^2 \left(\frac{\Sigma w_i x_i}{\Sigma x_i^2} - \frac{\Sigma x_i{}^2}{(\Sigma x_i^2)^2} \right)$$

$$= 2 \bar{X}^2 \left(\frac{\Sigma w_i x_i}{\Sigma x_i^2} - \frac{1}{(\Sigma x_i^2)} \right) = 2 \bar{X}^2 \left(\frac{\Sigma w_i (X_i - \bar{X})}{\Sigma x_i^2} - \frac{1}{\Sigma x_i^2} \right)$$

$$= 2 \bar{X}^2 \left(\frac{\Sigma w_i X_i - \bar{X} \Sigma w_i - 1}{\Sigma x_i{}^2} \right)$$

$$= 2 \bar{X}^2 \left(\left(\frac{\Sigma w_i X_i - 1}{\Sigma x_i{}^2} \right) - \left(\frac{\bar{X} \Sigma w_i}{\Sigma x_i{}^2} \right) \right)$$

karena $\Sigma w_i X_i = 1$ dan $\Sigma w_i = 0$, maka

$$2 \bar{X}^2 \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^2}\right)\left(\frac{x_i}{\Sigma x_i{}^2}\right) = 0$$

jadi $\text{var}(\alpha^*) = \sigma^2 \left[\frac{1}{n} + \bar{X}^2 \Sigma \left(w_i - \frac{x_i}{\Sigma x_i{}^2} \right)^2 \right.$

$$\left. + \bar{X}^2 \frac{\Sigma x_i{}^2}{(\Sigma x_i{}^2)^2} \right]$$

var (α^*) dapat minimum hanya dengan memanipulasikannya, yaitu dengan mengambil :

$$w_i = \frac{x_i}{\Sigma x_i^2}$$

sehingga var $(\alpha^*) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\Sigma x_i^2} \right)$

$$= \text{var } (\hat{\alpha})$$

Ini menunjukkan bahwa dengan $w_i = k_i$, varians dari α^* adalah sama dengan varians taksiran kuadrat terkecil α^*, dengan kata lain var $(\alpha^*) \geq$ var $(\hat{\alpha})$ adalah varians yang minimum untuk α

Sedangkan untuk β, dengan persamaan (2.15) :

$$\hat{\beta} - \beta = \frac{\Sigma x_i \varepsilon_i}{\Sigma x_i^2}$$

varians dari $\hat{\beta}$ adalah

$$E\ (\hat{\beta} - E(\hat{\beta}))^2 = E\ (\hat{\beta} - \beta)^2 \text{ , maka}$$

$$\hat{E}(\beta - \beta) = E\left(\frac{\Sigma x_i \varepsilon_i}{\Sigma x_i^2} \right)^2$$

$$E\ (\hat{\beta} - \beta) = E \left[\left(\frac{x_1 \varepsilon_1}{\Sigma x_i^2} \right)^2 + \left(\frac{x_2 \varepsilon_2}{\Sigma x_i^2} \right)^2 + \ldots + \left(\frac{x_n \varepsilon_n}{\Sigma x_i^2} \right)^2 \right.$$

$$+ 2 \left(\frac{x_1 x_2}{\Sigma x_i^2} \varepsilon_1 \varepsilon_2 \right) + 2 \left(\frac{x_1 x_3}{\Sigma x_i^2} \varepsilon_1 \varepsilon_3 \right) + \ldots +$$

$$\left. 2 \left(\frac{x_1 x_n}{\Sigma x_i^2} \varepsilon_1 \varepsilon_n \right) + \ldots + 2 \left(\frac{x_{n-1} x_n}{\Sigma x_i^2} \varepsilon_{n-1} \varepsilon_n \right) \right]$$

$$= E \left(\frac{\Sigma x_i^2 \ \varepsilon_i^2}{(\Sigma x_i^2)^2} + 2 \ \frac{\sum_{i>j} x_i x_j}{\Sigma x_i^2} \varepsilon_i \varepsilon_j \right)$$

$$= \frac{1}{\Sigma x_i^2} E(\varepsilon_i^2) \text{ , karena } E(\varepsilon_i \varepsilon_j) = 0 \text{ untuk } i \neq j$$

$$\text{var } (\hat{\beta}) = \frac{\sigma^2}{\Sigma x_i^{\ 2}} \qquad (2.19)$$

misalkan β^* penaksir linier dari β, yaitu

$$\beta^* = \Sigma\, w_i Y_i$$

dimana w_i tidak perlu sama dengan k_i pada persamaan (2.17)

$$E(\beta^*) = \Sigma\, w_i E(Y_i)$$

$$= \Sigma\, w_i (\alpha + \beta X_i)$$

$$= \alpha\, \Sigma\, w_i + \beta\, \Sigma\, w_i X_i$$

agar β^* tak bias maka

$$\Sigma\, w_i = 0 \qquad \text{dan} \qquad \Sigma\, w_i X_i = 1$$

$$\text{var } (\beta^*) = \text{var } (\Sigma\, w_i Y_i)$$

$$= \Sigma\, w_i^{\ 2}.\text{var } (Y_i)$$

$$= \sigma^2\, \Sigma\, w_i^{\ 2}$$

$$= \sigma^2\, \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^{\ 2}} + \frac{x_i}{\Sigma x_i^{\ 2}} \right)^2$$

$$= \sigma^2\, \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^{\ 2}} \right)^2 + \sigma^2 \frac{\Sigma x_i^{\ 2}}{(\Sigma x_i^{\ 2})^2} +$$

$$2\, \sigma^2\, \Sigma \left(w_i - \frac{x_i}{\Sigma x_i^{\ 2}} \right) \left(\frac{x_i}{\Sigma x_i^{\ 2}} \right)$$

varians β^* dapat menjadi minimum jika dimisalkan

$$w_i = \frac{x_i}{\Sigma x_i^2}$$

sehingga $\text{var } (\beta^*) = \dfrac{\sigma^2}{\Sigma x_i^{\ 2}}$

terlihat juga dengan $w_i = k_i$, varians β^* sama dengan varians kuadrat terkecil $\hat{\beta}$, sehingga var $(\hat{\beta})$ adalah varians minimum untuk β.

Dengan pembuktian-pembuktian diatas, $\hat{\alpha}$ dan $\hat{\beta}$ adalah linier, tak bias dan mempunyai varians minimum, maka $\hat{\alpha}$ dan $\hat{\beta}$ disebut penaksir tak bias linier terbaik (Best Linear Unbiased Estimator).

Sedangkan $\hat{\sigma}^2$ adalah estimator tak bias dari σ^2 sebab dengan mensubtitusikan (2.11) kedalam (2.14) dihasilkan :

$$e_i = \beta x_i + (\varepsilon_i - \bar{\varepsilon}) - \hat{\beta} x_i$$

$$e_i^2 = (\hat{\beta} - \beta)^2 . x_i^2 + (\varepsilon_i - \bar{\varepsilon})^2 - 2 (\hat{\beta} - \beta) x_i (\varepsilon_i - \bar{\varepsilon})$$

$$\Sigma e_i^2 = (\hat{\beta} - \beta)^2 \Sigma x_i^2 + \Sigma (\varepsilon - \bar{\varepsilon})^2$$
$$- 2 (\hat{\beta} - \beta) \Sigma x_i (\varepsilon_i - \bar{\varepsilon})$$

$$E(\Sigma e_i^2) = \Sigma x_i^2 \, E(\hat{\beta} - \beta)^2 + E[\Sigma (\varepsilon_i - \bar{\varepsilon})^2] -$$
$$2 E(\hat{\beta} - \beta) \Sigma x_i (\varepsilon_i - \bar{\varepsilon})$$
$$= \sigma^2 + (n-1) \sigma^2 - 2 \sigma^2$$
$$= (n-2) \sigma^2$$

sedangkan $\hat{\sigma}^2 = \dfrac{\Sigma e_i^2}{n - 2}$

Nilai yang diharapakannya adalah :

$$E(\hat{\sigma}^2) = \frac{1}{n - 2} E (\Sigma e_i^2)$$
$$= \sigma^2$$

Terbukti bahwa $\hat{\sigma}^2$ merupakan penaksir tak bias dari σ^2.

2.3. Test Hipotesa Parameter

Teori pengujian hipotesa berkenaan dengan pengembangan aturan atau prosedur untuk memutuskan apakah menerima atau menolak hipotesa. Secara garis besarnya adalah suatu prosedur dimana hasil sampel digunakan untuk menguji kebenaran atau kepalsuan suatu hipotesa nol. Keputusan untuk menerima atau menolak H_o dibuat atas dasar nilai statistik uji yang diperoleh dari data yang dimiliki.

2.3.1 Uji Parameter β

$H_o : \beta = \beta_o$

$H_1 : \beta \neq \beta_o$

Karena $\hat{\beta}$ kombinasi linier dari observasi Y_i, sehingga $\hat{\beta}$ berdistribusi normal dengan rata-rata β dan varians σ^2/S_{xx}, maka

$$Z = \frac{\hat{\beta} - \beta_o}{\sqrt{\sigma^2 / S_{xx}}}$$

adalah berdistribusi N(0,1) jika H_o benar.

$\hat{\sigma}^2$ atau MSE adalah estimator tak bias dari σ^2 dan distribusi dari (n-2)MSE/σ^2 adalah χ^2_{n-2} . Selain itu MSE dan $\hat{\beta}$ adalah variabel random independent.

Ini menyarankan bahwa jika kita mengganti σ^2 pada Z dengan $\hat{\sigma}^2$ = MSE, statistik

$$t_o = \frac{\hat{\beta} - \beta_o}{\sqrt{MSE/S_{xx}}}$$

adalah berdistribusi t (*student*) dengan derajat kebebasan n-2 jika H_o benar. Statistik t_o digunakan untuk test H_o: $\beta = \beta_o$ dengan membandingkan nilai observasi pada t_o dengan a/2 persentasi dari distribusi $t_{n-2}(t_{\alpha/2,n-2})$ dan tolak H_o jika $|t_o| > t_{\alpha/2,n-2}$

2.3. 2 Uji Parameter α

H_o: $\alpha = \alpha_o$

H_1: $\alpha \neq \alpha_o$

$$t_o = \frac{\hat{\alpha} - \alpha_o}{\sqrt{MSE\left[\frac{1}{n} + \bar{X}^2/S_{xx}\right]}}$$

Tolak H_o jika $|t_o| > t_{\alpha/2,n-2}$

2.3.3 Uji Independensi

H_o: $\beta = 0$

H_1: $\beta \neq 0$

Prosedur test untuk uji independensi dapat dikembangkan dari dua pendekatan.

Pendekatan pertama dimulai dengan kesamaan :

$$Y_i - \bar{Y} = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \quad \text{.................. (2.20)}$$

kuadratkan kedua sisi dari persamaan (2.20) dan jumlahkan lagi seluruh n observasi, dihasilkan :

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma(\hat{Y}_i - \bar{Y})^2 + \Sigma(Y_i - \hat{Y}_i)^2 +$$

$$2 \sum (\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i)$$

sedangkan $2 \sum (\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i) = 2 \sum \hat{Y}_i (Y_i - \hat{Y}_i) - 2\bar{Y} \sum (Y_i - \hat{Y}_i)$

$$= 2 \sum \hat{Y}_i e_i - 2 \bar{Y} \sum e_i$$

$$= 0$$

sehingga

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2 \ldots\ldots\ldots(2.21)$$

Ruas kiri dari persamaan (2.21) adalah jumlah kuadrat yang menunjukkan kesalahan observasi yaitu S_{yy}.

Dua komponen dari ukuran S_{yy} adalah jumlah perubahan pada observasi Y_i yang dijelaskan oleh garis regresi dan perbedaan residual sisa yang tidak dijelaskan oleh garis regresi, $S_{yy} = SSR + SSE$ ....(2.22)

Dengan membandingkan (2.6) dengan (2.22) kita melihat bahwa jumlah kuadrat regresi dapat dihitung dengan :

$$SSR = \hat{\beta}\, S_{xy}$$

Sehingga untuk test hipotesa $H_o : \beta = 0$ menggunakan test statistik

$$F_o = \frac{SSR/1}{SSE/n-2} = \frac{MSR}{MSE} \qquad \ldots\ldots\ldots\ldots (2.23)$$

tolak H_o jika $F_o > F_{\alpha,1,n-2}$

Pendekatan kedua dengan menggunakan t-test,

$$t_o = \frac{\hat{\beta}}{\sqrt{MSE/S_{xx}}} \qquad \ldots\ldots\ldots\ldots (2.24)$$

Jika pada (2.24) kedua sisinya dikuadratkan akan didapat

$$t_o^2 = \frac{\hat{\beta}^2 \; S_{xx}}{MSE} = \frac{MSR}{MSE}$$

Terlihat bahwa hasilnya sama dengan F_o pada (2.23).

Tolak H_o jika $|t_o| > t_{\alpha/2,n-2}$

2.4. Interval Konfidensi

Jika error (kesalahan) adalah normal dan berdistribusi secara bebas, maka distribusi sampling keduanya :

$$\frac{\hat{\beta} - \beta}{\sqrt{MSE/S_{xx}}} \quad \text{dan} \quad \frac{\hat{\alpha} - \alpha}{\sqrt{MSE\,[\frac{1}{n} + \bar{X}^2/S_{xx}]}}$$

adalah t (*student*) dengan n-2 derajat bebas.

Maka dari itu 100(1-a) persen interval konfidensi pada β diberikan oleh

$$\hat{\beta} - t_{\alpha/2,n-2}\sqrt{MSE/S_{xx}} \le \beta \le \hat{\beta} + t_{\alpha/2,n-2}\sqrt{MSE/S_{xx}}$$

dan 100(1-a) persen interval konfidensi pada α adalah

$$\hat{\alpha}-t_{\alpha/2,n-2}\sqrt{MSE\,[\frac{1}{n} + \bar{X}^2/S_{xx}} \le \alpha \le \hat{\alpha}+t_{\alpha/2,n-2}\sqrt{MSE\,[\frac{1}{n} + \bar{X}^2/S_{xx}}$$

Sedangkan distribusi sampling dari $\frac{(n-2)\;MSE}{\sigma^2}$ adalah

chi-kuadrat dengan n-2 derajat bebas, jadi

$$P\{\ \chi^2_{1-\alpha/2,n-2} \le \frac{(n-2)MSE}{\sigma^2} \le \chi^2_{\alpha/2,n-2}\} = 1 - \alpha$$

dan akibatnya 100(1-a) persen interval konfidensi pada σ^2

adalah :

$$\frac{(n-2)\ MSE}{\chi^2_{\alpha/2,\,n-2}} \le \sigma^2 \le \frac{(n-2)\ MSE}{\chi^2_{1-\alpha/2,\,n-2}}$$