II. TINJAUAN PUSTAKA

2.1. Model Linear

Model Linier dengan n pengamatan dan p variable penjelas biasa ditulis sebagai berikut :

$$Y_i = \beta_0 + X_{i1}\beta_1 + X_{i2}\beta_2 + \cdots + X_{ip}\beta_p + \varepsilon_i; \quad i = 1,2,\ldots,n$$

bila dirinci untuk setiap pengamatan :

$$\begin{array}{ccccccc}
Y_1 = \beta_0 & + X_{11}\beta_1 & + X_{12}\beta_2 & + \cdots & + X_{1p}\beta_p & + \varepsilon_1 \\
Y_2 = \beta_0 & + X_{21}\beta_1 & + X_{22}\beta_2 & + \cdots & + X_{2p}\beta_p & + \varepsilon_2 \\
\vdots \quad \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
Y_n = \beta_0 & + X_{n1}\beta_1 & + X_{n2}\beta_2 & + \cdots & + X_{np}\beta_p & + \varepsilon_n
\end{array}$$

dengan pendekatan matriks dapat ditulis sebagai berikut :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \ldots & X_{1p} \\ 1 & X_{21} & X_{22} & \ldots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \ldots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Persamaan diatas juga dapat ditulis dalam notasi matriks :

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

dimana

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \;;\; \Sigma(\mathbf{Y}) = \sigma^2\mathbf{I};\;\; \mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$$

Y = n x 1 vektor pengamatan.

X = matriks nxp dengan elemen-elemennya diketahui $(n > p)$, kolom pertama terdiri dari angka 1 yang menyatakan unsur intersep.

$\boldsymbol{\beta}_{\mathrm{p\ x\ 1}}$ = vektor parameter yang harus diduga.

$\boldsymbol{\varepsilon}_{\mathrm{n\ x\ 1}}$ = vektor galat, dengan $\boldsymbol{\varepsilon} \sim \mathrm{N_n}\left(\underline{\mathbf{0}}, \sigma^2\mathbf{I}\right)$.

Pada model linear umum terdapat beberapa asumsi dasar, yaitu :

1. Nilai rata-rata galat nol, yaitu: $\mathrm{E}(\varepsilon_\mathrm{i}) = 0$, untuk $\mathrm{i} = 1,2, \dots, \mathrm{n}$.
2. $\mathrm{var}(\varepsilon_\mathrm{i}) = \sigma^2$ adalah konstan (asumsi homoskedastisitas).
3. Tidak ada korelasi serial (*autocorrelation*) antara galat, berarti $\mathrm{kovarian}\left(\varepsilon_\mathrm{i}\varepsilon_\mathrm{j}\right) = 0$.
4. Tidak ada korelasi antar variabel bebas **X** (Gujarati, 1997).

2.2. Pendugaan Parameter

Pendugaan parameter adalah proses yang menggunakan sampel statistik untuk menduga atau menakasir parameter populasi yang tidak diketahui berdasarkan informasi dari sampel. Penduga adalah suatu statistik yang digunakan untuk menduga parameter. Menurut Hoog dan Craig (1995), Kriteria penduga yang baik adalah :

1. Takbias

   Suatu statistik dikatakan penduga tidak bias dari parameter θ apabila nilai harapan penduga sama dengan parameter θ, sebaliknya jika nilai harapan statistik tersebut tidak sama dengan parameter θ maka disebut penduga θ yang berbias.

Misal :

penduga $U(\underline{X})$ merupakan penduga tak bias bagi $g(\underline{\theta})$ bila $E\left(U(\underline{X})\right) = g(\underline{\theta})$ dan jika $E\left(U(\underline{X})\right) - g(\underline{\theta}) = b(\underline{\theta})$, maka $b(\underline{\theta})$ adalah bias bagi penduga $U(\underline{X})$ terhadap $g(\underline{\theta})$.

2. Varians Minimum

Suatu penduga $U(\underline{X})$ dikatakan mempunyai varians minimum apabila penduga tersebut memiliki varians yang kecil. Apabila terdapat lebih dari satu penduga, penduga yang efisien adalah penduga yang memiliki varians terkecil.

3. Konsisten

Suatu statistik dikatakan penduga yang konsisten jika peluangnya konvergen ke parameter θ.

$$\lim_{n\to\infty} P\left(\left|U(\underline{X}) - \underline{\theta}\right| < \varepsilon\right) = 1$$

atau

$$\lim_{n\to\infty} P\left(\left|U(\underline{X}) - \underline{\theta}\right| > \varepsilon\right) = 0$$

Semakin besarnya ukuran dari sampel maka ragam penduga semakin kecil.

4. Statistik Cukup

Definisi : $X \sim f(x, \theta)$; $\theta \in \Omega$

$x_1, x_2, \dots, x_n$ sampel acak, $Y = U(x_1, x_2, \dots, x_n)$ dikatakan statistik cukup bagi θ, jika :

$$f\left((x_1, x_2, \dots, x_n) \middle| U(x_1, x_2, \dots, x_n)\right) = \frac{f(x_1, \theta) \dots f(x_n, \theta)}{g\left(U(x_1, x_2, \dots, x_n)\right)}$$

$$= H(x_1, x_2, \dots, x_n)$$

$H(x_1, x_2, \dots, x_n)$ merupakan fungsi yang bebas dari θ untuk setiap nilai **Y**

5. Kelengkapan

   Misalkan peubah acak Z baik kontinu maupun diskrit mempunyai fungsi kepekatan peluang (fkp) bagi peubah acak kontinu dan fungsi peluang (pms) bagi peubah acak diskrit merupakan keluarga eksponensial dari $\{h(z;\theta) : \theta \in \Omega\}$. Jika kondisi $E[u(Z)] = 0$ untuk setiap $\theta \in \Omega$, memenuhi $u(z) = 0$ kecuali pada titik dimana probabilitasnya nol untuk setiap pdf $h(z;\theta) : \theta \in \Omega$. Maka keluarga eksponensial $\{h(z;\theta) : \theta \in \Omega\}$ disebut keluarga eksponensial lengkap dari fungsi kepekatan peluangnya.

2.3. Metode Pendugaan Parameter pada Model Linear

Untuk memperoleh penduga parameter pada model linear dapat dilakukan dengan beberapa metode, diantaranya :

1. Metode kuadrat terkecil/*Ordinary least square method* (OLS)

   Prinsip dasar dari metode kuadrat terkecil adalah mencari $\hat{\beta} = \{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p\}$ dengan meminimumkan bentuk kuadratik

$$\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

   Penduga parameter β pada metode metode kuadrat terkecil adalah

$$\boldsymbol{\beta}_{\mathrm{OLS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

   sedangkan penduga untuk parameter dispersi adalah

$$\hat{\sigma}^2 = \frac{\left((\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right)}{n - p}$$

   Selama asumsi dasar dipenuhi oleh data, maka dugaan metode kuadrat terkecil bersifat tak bias dengan varians minimum (Mustofa, U. dan Warsono, 2009).

2. Metode kemungkinan maksimum/*maximum likelihood method* (MLE)

Untuk menduga parameter model linear pada metode likelihood maksimum hal pertama yang dilakukan adalah menentukan fungsi likelihood; menyatakan fungsi likelihood L sebagai fungsi $\boldsymbol{\beta}$ dan σ^2; membangun ln L; memaksimumkan ln L terhadap $\boldsymbol{\beta}$ dan σ^2 untuk mencari penduga maksimum likelihoodnya. Karena $\mathbf{Y} \sim \mathrm{N_n}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, maka fungsi likelihoodnya adalah

$$\mathrm{L}(\beta, \sigma^2) = \left[\frac{1}{2\pi\sigma^2}\right]^{n/2} \exp\left[-\frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right]$$

Melogaritma asli kedua ruas, maka diperoleh

$$\ln \mathrm{L}(\beta, \sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}(\mathbf{Y}'\mathbf{Y} - \mathbf{2Y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta})$$

Untuk mencari nilai $\boldsymbol{\beta}$ dan σ^2 yang memaksimumkan fungsi likelihood, yaitu dengan cara mencari nilai turunan parsialnya yang disamakan dengan nol untuk masing-masing $\boldsymbol{\beta}$ dan σ^2.

$$\frac{\partial}{\partial\beta}\ln \mathrm{L}(\beta, \sigma^2) = \frac{2}{2\sigma^2}(\mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{X}\boldsymbol{\beta})$$

$$\frac{\partial}{\partial\sigma^2}\ln \mathrm{L}(\beta, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\big((\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\big)$$

sehingga, $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y}$

$$\sigma^2 = ((\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\ )\ /\ n$$

Karena $\mathrm{R(X)} = \mathrm{p}$, maka X'X juga berperingkat p dan nonsingular, sehingga penduga maksimum likelihoodnya adalah

$$\boldsymbol{\beta}_{\mathbf{MLE}} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}; \quad \sigma^2 = \frac{\mathbf{Y}'[\mathbf{I} - \mathbf{XX}^-]\mathbf{Y}}{n}$$

(Mustofa, U. dan Warsono, 2009).

3. Penggunaan *Singular Value Decompotition* (SVD) pada matriks $\mathbf{X}$

Menurut Rao, C.R (1971), jika model linear $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, dengan $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ dan $\Sigma(\mathbf{Y}) = \sigma^2\mathbf{I}$, akan tetapi $\mathrm{Rank}\,(\mathbf{X}) = \mathrm{r} < \mathrm{p}$ atau terjadi multikolinearitas, maka dapat diperoleh penduga yang baik dengan *Singular Value Decompotition* (SVD) pada matriks $\mathbf{X}$.

$$\mathbf{X}_{\mathrm{nm}} = \mathbf{P}_{\mathrm{nn}}\boldsymbol{\Lambda}_{\mathrm{nm}}\mathbf{Q}'_{\mathrm{mm}}$$

Dimana

$\mathbf{P}$ = kolom $\mathbf{P}$ adalah eigen vektor dari $\mathbf{XX}'$ yang ortonormal sedemikian sehingga $\mathbf{P}'\mathbf{P} = \mathbf{I}$.

$\mathbf{Q}$ = kolom $\mathbf{Q}$ adalah eigen vektor dari $\mathbf{X}'\mathbf{X}$ yang ortonormal sedemikian sehingga $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$.

$\boldsymbol{\Lambda}$ = matriks diagonal dengan unsur diagonalnya akar kuadrat dari nilai eigen tak nol dari matriks $\mathbf{X}'\mathbf{X}$ (Baker, 2005).

dengan $\boldsymbol{\gamma} = \boldsymbol{\Lambda}\mathbf{Q}'\boldsymbol{\beta}$ dan $\boldsymbol{\gamma}$ menjadi parameter baru, maka model direduksi menjadi :

$$\mathbf{Y} = \mathbf{P}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

sehingga model diatas telah memenuhi asumsi, sehingga teorema-teorema pada pendugaan OLS/MLE dapat dipergunakan.

$\hat{\boldsymbol{\gamma}} = \mathbf{P}'\mathbf{Y}$ adalah penduga takbias bagi γ, sehingga penduga dari fungsi parameter $\boldsymbol{\beta} = \mathbf{Q}\boldsymbol{\Lambda}^{-1}\mathbf{P}'\mathbf{Y} = \mathbf{Q}\boldsymbol{\Lambda}^{-1}\hat{\boldsymbol{\gamma}}$.

4. Metode *Generalized Least Square* (GLS)

Metode GLS adalah metode pendugaan parameter model linear $\mathbf{Y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, dengan asumsi $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$, $\Sigma(\mathbf{Y}) = \sigma^2\mathbf{V}$, V matriks definit positif, dan Rank $(\mathbf{X}) = \mathrm{p}$. Jika dibandingkan dengan asumsi pada metode MLE atau OLS pada metode GLS asumsi heteroskedastisitas dan galat yang saling berkorelasi dapat diizinkan. Pada GLS dilakukan pendekatan dengan transformasi matriks pengamatan $[\mathbf{Y}\ \ \mathbf{X}]$, ide transformasi ini menggunakan konsep matriks sehingga matriks kovarian $\sigma^2\mathbf{V}$ menjadi $\sigma^2\mathbf{I}$.

Karena matriks V adalah matriks simetriks definit positif sehingga matriks $\mathbf{V}$ nonsingular dan terdapat matriks n x n nonsingular $\mathbf{G}$, sehingga:

$$\mathbf{G}'\mathbf{G} = \mathbf{V}^{-1}$$

Pada model linier umum dikalikan dengan matriks **G** sehingga $\mathbf{GY} = \mathbf{GX}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\varepsilon}$ Sehingga matriks pengamatannya $[\mathbf{GY}\ \ \mathbf{GX}]$ dengan vektor galat $\mathbf{G}\boldsymbol{\varepsilon}$, sehingga matriks kovarian dari galatnya :

$$\begin{aligned}
\mathrm{cov}\,(\mathbf{G}\boldsymbol{\varepsilon}) &= \mathbf{G}\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')(\mathbf{G})' \\
&= \sigma^2\mathbf{GVG}' \\
&= \sigma^2\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' \\
&= \sigma^2\mathbf{I}
\end{aligned}$$

(Theil , 1971).

Karena matriks kovariannya $\mathrm{cov}(\mathbf{G}\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$, sehingga model transformasi memenuhi asumsi MLE ataupun OLS, sehingga teorema-teorema pada MLE / OLS dapat dipergunakan dan diperoleh **b** penduga parameter $\boldsymbol{\beta}$:

$\mathbf{b} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ dan $\mathrm{cov}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$

(Rao, 1973).

5. Invers Partisi Matriks

Menurut Rao, C.R (1971), jika model linier $\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$, mempunyai $\text{rank}(\mathbf{X}) = \text{r} < p$ dan $\boldsymbol{\varepsilon} \sim \text{N}_\text{n}\left(\underline{\mathbf{0}}, \sigma^2\mathbf{V}\right)$. Parameter $\boldsymbol{\beta}$ dapat diduga dengan menggunakan invers partisi matriks $\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{bmatrix}^{-} = \begin{bmatrix} \mathbf{C_1} & \mathbf{C_2} \\ \mathbf{C_3} & -\mathbf{C_4} \end{bmatrix}$ dapat dicari penduga dari $\boldsymbol{\beta}$, matriks dispersi dari $\widehat{\boldsymbol{\beta}}$, dan penduga takbias untuk σ^2 :

i. [menggunakan $\mathbf{C_2}$ atau $\mathbf{C_3}$]

   $\mathbf{p}'\widehat{\boldsymbol{\beta}}$ adalah penduga BLUE dari $\mathbf{p}'\boldsymbol{\beta}$ dimana $\widehat{\boldsymbol{\beta}} = \mathbf{C_2'Y}$ atau $\widehat{\boldsymbol{\beta}} = \mathbf{C_3Y}$.

ii. [menggunakan $\mathbf{C_4}$]

   Matriks dispersion dari $\widehat{\boldsymbol{\beta}}$ adalah $\sigma^2\mathbf{C_4}$

iii. [menggunakan $\mathbf{C_1}$]

   Penduga takbias untuk $\sigma^2 = \text{f}^{-1}\mathbf{Y}'\mathbf{C_1}\mathbf{Y}$, dimana $\text{f} = \text{R}(\mathbf{V} \quad \mathbf{X}) - \text{R}(\mathbf{X})$ dan f adalah deerajat bebasnya.

2.4. Pengujian Hipotesis

Dalam pengujian hipotesis yang digunakan adalah Uji F

Uji F digunakan untuk menguji secara bersamaan apakah parameter dalam model menerangkan respon secara signifikan atau tidak.

Hipotesis:

$\text{Ho} : \mathbf{H\beta} = \mathbf{h}$

$\text{H}_\text{a} : \mathbf{H\beta} \neq \mathbf{h}$

Statistik uji:

$$F = \frac{(SSR(\mathbf{H}) - SSR)/q}{SSR/(n-p)}$$

dimana :

SSR : Jumlah Kuadrat Galat

SSR($\mathbf{H}$) : Jumlah Kuadrat Galat dengan kendala $\mathbf{H\beta} = \mathbf{h}$

Hipotesis nol ditolak jika:

$$F_{hitung} > F_{\alpha(q,n-p)}$$

P-value juga dapat digunakan untuk menolak atau tidak tolak hipotesis. Semakin kecil p-value, semain kecil peluang membuat kesalahan yang diakibatkan menolak Ho. Artinya berdasarkan nilai peluang yang ada, p-value < α maka Ho ditolak pada tingkatan α tertentu.

Pada pengujian hipotesis dua macam kesalahan :

Kesimpulan \ H_0	**Hipotesis Benar**	**Hipotesis Salah**
Tidak tolak H_0	Benar	Kesalahan Tipe II (Beta) ⇨ Kuasa Uji = 1 - beta
Tolak H_0	Kesalahan Tipe I Taraf nyata (α)	Benar

α = P (menolak H_0 | H_0 benar) atau dengan kata lain α adalah P (nilai yang diamati dari statistik uji akan jatuh di wilayah penolakan ketika H_0 benar), sedangkan beta (kesalahan jenis II) adalah peluang menerima H_0 dimana H_0 salah, atau P (nilai yang diamati dari statistik uji tidak akan jatuh dalam penolakan wilayah ketika hipotesis nol adalah salah). Sehingga *power* / kuasa uji adalah peluang menolak H_0 dimana H_0 tidak benar) atau sama saja dengan peluang (statistik uji akan jatuh dalam penolakan wilayah ketika hipotesis nol salah). Kuasa uji = 1 - beta.

2.5. Invers Partisi Matriks / IPM

Menurut Rao, C.R (1972) Didefinisikan matriks

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix}^{-} = \begin{bmatrix} \mathbf{C_1} & \mathbf{C_2} \\ \mathbf{C_3} & -\mathbf{C_4} \end{bmatrix}$$

X' adalah transpose matriks **X**.

Berdasarkan definisi g-invers $\mathbf{AA^-A} = \mathbf{A}$, maka

$$\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{C_1} & \mathbf{C_2} \\ \mathbf{C_3} & -\mathbf{C_4} \end{bmatrix}\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{VC_1V} + \mathbf{XC_3V} + \mathbf{VC_2X'} - \mathbf{XC_4X'} & \mathbf{VC_1X} + \mathbf{XC_3X} \\ \mathbf{X'C_1V} + \mathbf{X'C_2X'} & \mathbf{X'C_1X} \end{bmatrix} = \begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix} \qquad (1)$$

$\begin{bmatrix} \mathbf{C_1} & \mathbf{C_2} \\ \mathbf{C_3} & -\mathbf{C_4} \end{bmatrix}$ disebut juga invers partisi matriks dari $\begin{bmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X'} & \mathbf{0} \end{bmatrix}$.