BAB II
TINJAUAN PUSTAKA

2.1 Statistika

Menurut Sudjana (2005), Statistik merupakan kata yang digunakan untuk menyatakan sekumpulan data, bilangan maupun non bilangan yang disusun dalam sebuah tabel dan atau diagram yang menggambarkan suatu persoalan. Sedangkan statistika adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisisan dan penarikan kesimpulan berdasarkan data dan penganalisisan yang digunakan. Jenis data yang digunakan pada statistik sendiri terbagi menjadi dua macam yaitu:

1. Data kuantitatif merupakan data yang berbentuk bilangan yang sifatnya berubah-ubah atau bersifat variabel. Dari nilainya sendiri, data kuantitatif terbagi menjadi dua golongan yaitu data dengan variabel diskrit atau data yang nilainya merupakan bilangan caca, sedangkan data dengan variabel kontinu merupakan data yang nilainya terdapat pada suatu interval tertentu.
2. Data kualitatif merupakan data yang bukan kuantitatif atau data yang tidak berbentuk bilangan. Data ini sendiri dikategorikan menurut objek yang dipelajari atau biasa disebut juga dengan nama *atribut*.

Sugiyono (2011) menyatakan bahwa pengumpulan data menurut sumbernya dibagi menjadi dua macam yaitu sumber primer dan sumber sekunder. Sumber primer merupakan kondisi dimana sumber data langsung memberikan datanya kepada pengumpul data, sedangkan sumber sekunder merupakan kondisi dimana sumber data tidak langsung memberikan datanya kepada pengumpul data melainkan melalui orang lain. Selain itu, Sugiyono (2011) juga menyebutkan bahwa teknik analisis kuantitatif dibagi menjadi dua macam yaitu statistik deskriptif dan statistik inferensia sebagai berikut:

1. Statistik deskriptif merupakan statistik yang digunakan untuk menganalisis data dengan cara menggambarkan data yang terkumpul sebagaimana adanya tanpa bermaksud untuk membuat kesimpulan yang berlaku secara umum.

2. Statistik inferensial (*statistik induktif* atau *statistik probabilitas*) merupakan teknik statistik yang digunakan untuk menganalisis data sampel yang hasilnya diberlakukan untuk populasi. Statistik ini disebut sebagai statistik *probability* karena kesimpulan dari hasil data sampelnya bersifat peluang (*probability*). Dimana kesimpulan dari data sampel yang diberlakukan untuk populasinya memiliki peluang kesalahan dan kebenaran (kepercayaan) yang dinyatakan dalam bentuk persentase seperti peluang kesalahan 5% maka taraf nyatanya 95%.

2.2 Matriks

Menurut Kariadinata (2019), matriks merupakan susunan sekelompok bilangan dalam suatu jajaran yang berbentuk persegi panjang yang diatur berdasarkan baris dan kolom, serta diletakkan diantara dua tanda kurung. Dalam suatu ukuran matriks berordo $a \times a$, angka pertama menyatakan jumlah baris (garis horizontal), dan angka kedua menyatakan jumlah kolom (garis vertikal). Sebuah matriks yang hanya mempunyai satu kolom disebut *matriks kolom*. Sedangkan sebuah matriks yang hanya memiliki satu baris disebut *matriks baris*. Sebuah matriks yang hanya terdiri dari satu kolom maupun satu baris juga disebut sebagai vektor.

1. Jenis-jenis matriks

   Matriks terdiri atas berbagai jenis antara lain, matriks nol, matriks baris, matriks kolom, matriks persegi, matriks segitiga atas, matriks segitiga bawah, matriks diagonal, dan matriks identitas. Berikut merupakan bentuk dari jenis-jenis matriks (Gumilar, 2008) :

   a. Matriks Nol

      Matriks nol merupakan matriks yang seluruh elemennya bernilai nol, contohnya

      $$\boldsymbol{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \boldsymbol{C} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

   b. Matriks Baris

      Matriks baris merupakan matriks yang hanya terdiri dari satu baris saja atau berordo $1 \times \mathrm{a}$, contohnya

$$E = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \quad F = \begin{bmatrix} -1 & 3 \end{bmatrix} \quad G = \begin{bmatrix} -2 & 4 & 6 & -8 \end{bmatrix}$$

Matriks E berordo 1×3, matriks F berordo 1×2, dan matriks G berordo 1×4. Matriks E,F, dan G di atas hanya memiliki satu baris saja sehingga disebut sebagai matriks baris.

c. Matriks Kolom

Matriks kolom merupakan matriks yang hanya terdiri dari satu kolom saja atau berordo $a \times 1$, contohnya

$$H = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \quad J = \begin{bmatrix} -4 \\ 3 \\ 10 \end{bmatrix} \quad K = \begin{bmatrix} 5 \\ -7 \\ 2 \\ 1 \end{bmatrix}$$

Matriks H berordo 2×1, matriks J berordo 3×1, dan matriks K berordo 4×1. Matriks H,J, dan K di atas hanya memiliki satu kolom saja sehingga disebut sebagai matriks kolom.

d. Matriks Persegi

Matriks persegi merupakan matriks yang elemen baris dan kolomnya memiliki banyak yang sama, contohnya

$$L = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} \quad N = \begin{bmatrix} 1 & 0 & -12 \\ 6 & -3 & 0 \\ 4 & 3 & 6 \end{bmatrix}$$

Matriks L berordo 2×2 dan matriks N berordo 3×3. Dikarenakan banyaknya elemen pada baris sama dengan banyaknya elemen pada kolom, maka matriks L dan matriks N disebut sebagai matriks persegi.

e. Matriks Segitiga Atas

Matriks segitiga atas merupakan matriks persegi yang elemen di bawah diagonal utamanya bernilai nol, contohnya

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$$

f. Matriks Segitiga Bawah

Matriks segitiga bawah merupakan matriks persegi yang elemen di atas diagonal utamanya bernilai nol, contohnya

$$\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$$

g. Matriks Diagonal

Matriks diagonal merupakan matriks persegi yang elemen-elemennya bernilai nol, kecuali diagonal utamanya tidak selalu nol, contohnya

$$\begin{bmatrix} \mathrm{a} & 0 & 0 \\ 0 & \mathrm{b} & 0 \\ 0 & 0 & \mathrm{c} \end{bmatrix}$$

h. Matriks Identitas

Matriks identitas merupakan matriks skalar yang elemen-elemen pada diagonal utamanya bernilai 1, contohnya

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2. Transpose

Transpose matriks adalah matriks yang dihasilkan dari hasil mengubah baris-baris matriks menjadi kolom-kolom matriks. Jika suatu matriks A berordo $m \times n$, maka transpose dari matriks tersebut adalah matriks berordo $n \times m$. Misalkan sebuah matriks $\boldsymbol{A}$ dimana $\boldsymbol{A} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, maka transpose dari matriks $\boldsymbol{A}$ adalah $\boldsymbol{A^T} \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$.

Adapun beberapa sifat dari matriks transpose yaitu (Gumilar, 2008) :

a. $(\boldsymbol{A}+\boldsymbol{B})^{\boldsymbol{T}} = (\boldsymbol{A})^{\boldsymbol{T}} + (\boldsymbol{B})^{\boldsymbol{T}}$
b. $((\boldsymbol{A})^{\boldsymbol{T}})^{\boldsymbol{T}} = \boldsymbol{A}$
c. $(\boldsymbol{AB})^{\boldsymbol{T}} = (\boldsymbol{B})^{\boldsymbol{T}}.(\boldsymbol{A})^{\boldsymbol{T}}$
d. $(c\boldsymbol{A})^{\boldsymbol{T}} = c.\boldsymbol{A^T}$, dengan $c =$ konstanta

3. Determinan

Determinan didefinisikan sebagai selisih antara hasil kali elemen-elemen pada diagonal utama dengan hasil kali elemen-elemen pada diagonal sekunder (Gumilar, 2008).

a. Determinan matriks ordo 2×2

Misalkan $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, maka determinan dari matriks $\boldsymbol{A}$ atau dapat juga dinotasikan dengan $\det \boldsymbol{A}$ atau $|\boldsymbol{A}|$ adalah sebagai berikut :

$$\det \boldsymbol{A} = |\boldsymbol{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = (a \times d) - (b \times c) = ad \times bc$$

Elemen a dan d merupakan elemen-elemen pada diagonal utama, sedangkan elemen b dan c merupakan elemen-elemen pada diagonal sekunder.

b. Determinan matriks ordo 3×3

Misalkan, $\boldsymbol{A}$ matriks persegi berordo 3×3 berikut ini.

$$\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Untuk mencari determinan dari matriks $\boldsymbol{A}$ yang berordo 3×3 diperlukan metode Sarrus sebagai berikut :

$$\det \boldsymbol{A} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix}$$

$$\det \boldsymbol{A} = (a)(e)(i) + (b)(f)(g) + (c)(d)(h) - (g)(e)(c) - (h)(f)(a) - (i)(d)(b)$$

Berdasarkan nilai diskriminannya, suatu matriks dapat dibedakan menjadi dua jenis yaitu matriks singular dan matriks non singular. Matriks singular adalah matriks yang nilai determinannya nol, sedangkan matriks non singular adalah matriks yang nilai determinannya tidak sama dengan nol.

4. Invers

Invers matriks adalah matriks kebalikan dari matriks persegi. Apabila suatu matriks dikalikan dengan inversnya maka akan diperoleh matriks identitas atau $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$, maka $\boldsymbol{A}$ dikatakan dapat dibalik (*invertible*) dan $\boldsymbol{B}$ dinamakan invers dari matriks $\boldsymbol{A}$.

a. Invers matriks ordo 2×2

Misalkan $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, maka invers dari matriks $\boldsymbol{A}$ adalah sebagai berikut :

$$\boldsymbol{A^{-1}} = \frac{1}{\det A} \, Adjoin\ A$$

$$\boldsymbol{A^{-1}} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\boldsymbol{A^{-1}} = \begin{bmatrix} \frac{d}{\det A} & \frac{-b}{\det A} \\ \frac{-c}{\det A} & \frac{a}{\det A} \end{bmatrix}$$

b. Invers matriks ordo 3×3

Misalkan $\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, maka invers dari matriks $\boldsymbol{A}$ adalah sebagai berikut :

$$\boldsymbol{A^{-1}} = \frac{1}{\det A}\ Adjoin\ A$$

$$\boldsymbol{A^{-1}} = \frac{1}{\det A} \begin{bmatrix} +\begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & +\begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & +\begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ +\begin{vmatrix} b & c \\ e & f \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} & +\begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$$

2.3 Regresi Linear

Menurut Suyono (2015), model regresi linear sederhana adalah model probabilistik yang menyatakan hubungan linear antara dua variabel yang menganggap bahwa salah satu variabel dapat mempengaruhi variabel lainnya. Variabel yang mempengaruhi variabel lainnya disebut dengan variabel bebas, sedangkan variabel yang dipengaruhi oleh variabel lainnya disebut variabel terikat. Model probabilistik regresi linear sederhana adalah sebagai berikut:

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

Regresi linear juga dapat melibatkan lebih dari satu variabel bebas. Regresi linear yang melibatkan lebih dari satu variabel bebas disebut regresi linear berganda. Model probabilistik dari regresi linear berganda adalah sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \tag{2.2}$$

Dimana :

Y : Variabel terikat

X : Variabel bebas

β_0 : Koefisien intersep

$\beta_1, \beta_2, \dots, \beta_k$: Koefisien slope

ε : Galat acak atau residual

Berdasarkan Persamaan (2.2) diketahui bahwa untuk mendapati nilai Y perlu dilakukan estimasi $(\beta_1, \beta_2, \dots, \beta_k)$ dan residualnya. Ada beberapa cara untuk mendapatkan estimator atau penduga untuk $\beta_1, \beta_2, \dots, \beta_k$, salah satunya yaitu dengan menggunakan metode kuadrat biasa (*Ordinary Least Square*), persamaan ini akan dipaparkan pada Persamaan (2.20). Sedangkan untuk mengestimasi residual pada regresi linier dapat dilakukan dengan menggunakan rumus sebagai berikut.

$$S^2 = \frac{JK_{Res}}{n - K - 1} \tag{2.3}$$

dengan

$$JK_{Res} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \tag{2.4}$$

Dimana

S^2	:	Estimator variasi galat acak atau estimasi residual
JK_{Res}	:	Jumlah kuadrat residual
$\hat{Y}$	:	Penduga parameter
n	:	Banyaknya data
K	:	Banyaknya variabel bebas

Estimator S^2 merupakan estimator yang bersifat tak bias untuk variansi galat. Semakin kecil nilai S^2 maka semakin sesuai model regresinya (Suyono, 2015).

A. Koefisien Korelasi

Menurut Suyono (2015), koefisien korelasi merupakan suatu nilai yang digunakan untuk menyatakan keeratan hubungan linear antara dua variabel. Koefisien korelasi antar variabel terikat dan variabel bebas dapat didefinisikan sebagai

$$r = \frac{J_{XY}}{\sqrt{J_{XX}\, JK_{Total}}} \tag{2.5}$$

dimana

$$J_{XX} = \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} {X_i}^2 - \frac{(\sum_{i=1}^{n} X_i)^2}{n} \tag{2.6}$$

$$JK_{Total} = \sum_{i=1}^{n} (Y_i - \overline{Y})^2 = \sum_{i=1}^{n} {Y_i}^2 - \frac{(\sum_{i=1}^{n} Y_i)^2}{n} \tag{2.7}$$

dan

$$J_{XY} = \sum_{i=1}^{n} (X_i - \bar{\mathrm{X}})(Y_i - \bar{\mathrm{Y}}) = \sum_{i=1}^{n} X_i Y_i - \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n} \tag{2.8}$$

Persamaan (2.5) dapat ditulis juga

$$r = \frac{n \sum_{i=1}^{n} X_i Y_i - (\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{\sqrt{n \sum_{i=1}^{n} {X_i}^2 - (\sum_{i=1}^{n} X_i)^2} \sqrt{n \sum_{i=1}^{n} {Y_i}^2 - (\sum_{i=1}^{n} Y_i)^2}} \tag{2.9}$$

Dimana

r	:	Koefisien korelasi
J_{XY}	:	Total variasi variabel X dan Y
J_{XX}	:	Total variasi variabel bebas
JK_{Total}	:	Jumlah kuadrat total
$\bar{\mathrm{X}}$	:	Rata-rata dari variabel bebas
$\bar{\mathrm{Y}}$	:	Rata-rata dari variabel terikat
n	:	Banyak data

Untuk menghitung koefisien korelasi sampel r juga tidak mensyaratkan X dan Y berdistribusi normal. Nilai koefisien korelasi r berkisar antara -1 sampai 1. Jika $r = 0$, maka tidak ada hubungan linear antara variabel X dan Y. Jika r mendekati 0, maka hubungan linear antara X dan Y semakin lemah. Jika nilai r mendekati 1 atau -1 menunjukkan bahwa semakin kuat hubungan linear antara variabel X dan Y. Nilai positif menunjukkan adanya hubungan positif antara variabel X dan Y. Sedangkan nilai negatif menunjukkan adanya hubungan negatif antara variabel X dan Y.

B. Koefisien Determinasi

Dalam mengukur kesesuaian model regresi linear dapat dilakukan dengan berbagai cara. Menurut Suyono (2015), salah satu cara untuk mengukur kesesuaian model regresi linear dapat dilakukan dengan mengukur kontribusi yang diberikan oleh variabel *X* dalam memprediksi *Y*. Dengan menggunakan $\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_k X_k$ untuk memprediksi *Y* diperlukan kontribusi dari *X* menggunakan perhitungan jumlah kuadrat residual pada Persamaan (2.4), maka selanjutnya dilakukan proporsi pengurangan dalam jumlah kuadrat penyimpangan yang diberikan oleh X terhadap JK_{Total} adalah :

$$\frac{JK_{Total} - JK_{Res}}{JK_{Total}} = \frac{Variasi\ dalam\ sampel\ yang\ dijelaskan\ oleh\ X}{Total\ variansi\ dalam\ sampel}$$

$$= Proporsi\ total\ variasi\ dalam\ sampel\ yang\ dijelaskan\ oleh\ hubungan\ linier \tag{2.10}$$

Proporsi pada Persamaan (2.9) dinamakan dengan *koefisien determinasi*. Umumnya koefisien determinasi disimbolkan dengan R^2 dengan rumus yaitu:

$$R^2 = \frac{JK_{Total} - JK_{Res}}{JK_{Total}} = 1 - \frac{JK_{Res}}{JK_{Total}} \tag{2.11}$$

Dimana

R^2	:	Koefisien determinasi
JK_{Res}	:	Jumlah kuadrat residual-
JK_{Total}	:	Total variasi variabel terikat

2.4 Regresi Data Panel

Menurut Caraka (2017), data panel merupakan gabungan antara data runtun waktu (*time series*) dan data silang (*cross section*). Data runtun waktu merupakan data dari suatu objek atau individu yang terdiri dari beberapa periode (harian, bulanan, triwulan, maupun tahunan). Sedangkan data silang merupakan data yang terdiri dari beberapa objek (misalnya; perusahaan, negara, provinsi) dengan berbagai jenis data dalam suatu periode tertentu. Adapun keuntungan yang diperoleh dengan menggunakan data panel ini yaitu mampu menghasilkan data yang lebih banyak dengan derajat bebas yang lebih besar karena menggunakan gabungan dari data *cross section* dan *time series*. Selain itu, dengan menggabungkan informasi data *time series* dan *cross section* juga dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel.

Regresi yang menggunakan data panel disebut dengan model regresi data panel. Persamaan model regresi dengan menggunakan data *cross section* dapat ditulis sebagai berikut (Caraka, 2017) :

$$y_i = \beta_{0i} + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_K X_{Ki} + \varepsilon_i \tag{2.12}$$

dengan $i = 1,2, \dots, N$ dimana i adalah data *cross section*. Sedangkan persamaan model regresi dengan menggunakan data *time series* dapat ditulis sebagai berikut :

$$y_t = \beta_{0t} + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_K X_{Kt} + \varepsilon_t \qquad (2.13)$$

dengan $t = 1,2, \dots, T$ dimana t adalah data *time series*.

Secara umum model regresi data panel ditulis sebagai berikut:

$$y_{it} = \beta_{0it} + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_K X_{Kit} + \varepsilon_{it} \qquad (2.14)$$

$$y_{it} = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{NT} \end{bmatrix};\ X_{it} = \begin{bmatrix} X_{1i1} & X_{2i1} & \dots & X_{Ki1} \\ X_{1i2} & X_{2i2} & \dots & X_{Ki2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1NT} & X_{2NT} & \dots & X_{KNT} \end{bmatrix};\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}; \varepsilon_{it} = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

dengan $i = 1,2, \dots, N$; $t = 1,2, \dots, T$; $k = 1,2, \dots, K$, dimana k adalah vari0abel bebas.

Keterangan :

y_{it}	: Variabel terikat individu ke-i dan waktu ke-t
β_0	: Koefisien intersep
X_{kit}	: Variabel bebas dari individu ke-i dan waktu ke-t
$\beta_1, \beta_2, \dots, \beta_k$	: Koefisien slope variabel bebas
ε_{it}	: *error* untuk individu ke-i dan waktu ke-t

2.4.1 Model Regresi Data Panel

Dalam teknik analisisnya, data panel dapat mendeteksi dan mengukur efek dengan lebih baik yang tidak dapat diamati secara murni oleh data deret waktu atau data *cross section* murni. Namun demikian, Gujarati (2003) menyatakan bahwa dalam penggunaan regresi data panel terdapat beberapa kemungkinan model yang akan muncul, yaitu:

1. Intersep dan koefisien slope adalah konstan sepanjang waktu dan individu
2. Koefisien slope konstan tetapi intersep berbeda antar individu
3. Koefisien slope konstan tetapi intersep berbeda antar individu dan waktu
4. Semua koefisien (baik intersep maupun koefisien slope) berbeda antar individu
5. Intersep dan koefisien slope berbeda antar individu dan waktu

Terdapat beberapa teknik estimasi untuk mengatasi masalah-masalah tersebut diantaranya yaitu *common effect model* (CEM), *fixed effect model* (FEM), dan *random effect model* (REM) atau *error component* model (ECM). Pada estimasi CEM dan FEM menggunakan pendekatan OLS, hanya saja pada FEM intersepnya dinyatakan dengan variabel *dummy* atau *Least Square Dummy Variable* (LSDV), sedangkan pada estimasi REM menggunakan pendekatan *Generalized Least Square* (GLS).

A. Common Effect Model (CEM)

Common effect model sering disebut sebagai model regresi data panel dengan teknik estimasi paling sederhana diantara teknik estimasi model lainnya. Menurut Gujarati (2003), pada pemodelan ini perbedaan dimensi individu dan waktu diabaikan. Dengan kata lain, pada CEM perilaku dari setiap individu sama dalam berbagai periode waktu. Oleh karena itu, pada model ini data *cross section* dan data *time series* dikombinasikan menjadi satu kesatuan tanpa melihat perbedaan waktu dan individu.

Adapun model regresi data panel dengan menggunakan pendekatan OLS maka model dari CEM adalah sebagai berikut (Gujarati,2003):

$$Y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it}$$

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_K X_{Kit} + \varepsilon_{it} \quad (2.15)$$

dengan indeks *i* menyatakan individu ke-*i* dan indeks *t* menyatakan periode ke-*t*.

Dari Persamaan (2.15), selanjutnya akan ditaksir parameternya sebagai berikut:

$$\hat{Y}_{it} = b_0 + b_1 X_{1it} + b_2 X_{2it} + \cdots + b_K X_{Kit} \quad (2.16)$$

Nilai koefisien-koefisien b_0, b_1, b_2,…, b_K dapat diperoleh dengan metode kuadrat terkecil atau OLS, dengan cara meminimumkan jumlah kuadrat sisanya terlebih dahulu :

$$\begin{aligned} S &= \sum_{i=1}^{n} e_{it}^2 \\ &= \sum_{i=1}^{n} \left(Y_{it} - \hat{Y}_{it}\right)^2 \\ &= \sum_{i=1}^{n} (Y_{it} - b_0 - b_1 X_{1it} - b_2 X_{2it} - \cdots - b_K X_{Kit})^2 \quad (2.17) \end{aligned}$$

Kemudian dari Persamaan (2.17), selanjutnya dideferensialkan terhadap masing-masing koefisien b_0, b_1, b_2,…, b_K sehingga diperoleh persamaan-persamaan sebagai berikut:

Diferensial terhadap koefisien b_0,

$$\frac{\partial}{\partial b_0}\sum_{i=1}^{n} e_{it}^2 = 0$$

$$\frac{\partial}{\partial b_0}\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_KX_{Kit}))^2 = 0$$

$$2[\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_KX_{Kit})^2\,(-1)] = 0$$

$$2[-\sum_{i=1}^{n} Y_{it} + b_0\sum_{i=1}^{n} 1 + b_1\sum_{i=1}^{n} X_{1it} + \; b_2\sum_{i=1}^{n} X_{2it} + \cdots +$$

$$b_K\sum_{i=1}^{n} X_{Kit}] = 0$$

$$2[-\sum_{i=1}^{n} Y_{it} + b_0n + b_1\sum_{i=1}^{n} X_{1it} + b_2\sum_{i=1}^{n} X_{2it} + \; \ldots + b_K\sum_{i=1}^{n} X_{Kit}] = 0$$

Sehingga diperoleh :

$$\sum_{i=1}^{n} Y_{it} = b_0n + b_1\sum_{i=1}^{n} X_{1it} + b_2\sum_{i=1}^{n} X_{2it} + \cdots + b_K\sum_{i=1}^{n} X_{Kit}$$

Diferensial terhadap koefisien b_1,

$$\frac{\partial}{\partial b_1}\sum_{i=1}^{n} e_{it}^2 = 0$$

$$\frac{\partial}{\partial b_1}\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_KX_{Kit})^2 = 0$$

$$2[\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_kX_{Kit})^2\,(-X_{1it})] = 0$$

$$2[-\sum_{i=1}^{n} Y_{it}X_{1it} + b_0\sum_{i=1}^{n} X_{1it} + b_1\sum_{i=1}^{n} X_{1it}^2 +$$

$$b_2\sum_{i=1}^{n} X_{2it}X_{0000000000000000000000000000000000001it} + \cdots +$$

$$b_K\sum_{i=1}^{n} X_{Kit}X_{1it}] = 0$$

Sehingga diperoleh :

$$\sum_{i=1}^{n} Y_{it}X_{1it} = b_0\sum_{i=1}^{n} X_{1it} + b_1\sum_{i=1}^{n} X_{1it}^2 + b_2\sum_{i=1}^{n} X_{2it}X_{1it} + \cdots + b_K\sum_{i=1}^{n} X_{Kit}X_{1it}$$

Dengan cara yang sama maka akan diperoleh persamaan untuk b_K adalah sebagai berikut :

$$\frac{\partial}{\partial b_K}\sum_{i=1}^{n} e_{it}^2 = 0$$

$$\frac{\partial}{\partial b_k}\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_KX_{Kit})^2 = 0$$

$$2[\sum_{i=1}^{n}(Y_{it} - b_0 - b_1X_{1it} - b_2X_{2it} - \cdots - b_KX_{Kit})^2\,(-X_{Kit})] = 0$$

$$2[-\sum_{i=1}^{n} Y_{it}X_{Kit} + b_0\sum_{i=1}^{n} X_{Kit} + b_1\sum_{i=1}^{n} X_{1it}X_{Kit} +$$

$$b_2\sum_{i=1}^{n} X_{2it}X_{Kit} + \cdots + b_k\sum_{i=1}^{n} X_{Kit}^2] = 0$$

Sehingga diperoleh :

$$\sum_{i=1}^{n} Y_{it}X_{kit} = b_0\sum_{i=1}^{n} X_{Kit} + b_1\sum_{i=1}^{n} X_{kit}^2 + b_2\sum_{i=1}^{n} X_{2it}X_{Kit} + \cdots + b_k\sum_{i=1}^{n} X_{Kit}^2$$

Sehingga diperoleh persamaan-persamaan sebagai berikut :

$$\begin{aligned}
\sum_{i=1}^{n} Y_{it} &= b_0 n + b_1 \sum_{i=1}^{n} X_{1it} + \cdots + b_K \sum_{i=1}^{n} X_{Kit} \\
\sum_{i=1}^{n} Y_{it} X_{1it} &= b_0 \sum_{i=1}^{n} X_{1it} + b_1 \sum_{i=1}^{n} X_{1it}^2 + \cdots + b_K \sum_{i=1}^{n} X_{Kit} X_{1it} \\
\vdots \qquad & \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\sum_{i=1}^{n} Y_{it} X_{kit} &= b_0 \sum_{i=1}^{n} X_{Kit} + b_1 \sum_{i=1}^{n} X_{kit}^2 + \cdots + b_k \sum_{i=1}^{n} X_{Kit}^2
\end{aligned}$$

Apabila persamaan di atas ditulis dalam bentuk matriks, maka matriks akan menjadi:

$$\begin{bmatrix} \sum_{i=1}^{n} \boldsymbol{Y_{it}} \\ \sum_{i=1}^{n} \boldsymbol{Y_{it} X_{1it}} \\ \vdots \\ \sum_{i=1}^{n} \boldsymbol{Y_{it} X_{Kit}} \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} X_{1it} & \cdots & \sum_{i=1}^{n} X_{Kit} \\ \sum_{i=1}^{n} X_{1it} & \sum_{i=1}^{n} X_{1it}^2 & \cdots & \sum_{i=1}^{n} X_{Kit} X_{1it} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{n} X_{kit} & \sum_{i=1}^{n} X_{1it} X_{kit} & \cdots & \sum_{i=1}^{n} X_{Kit}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_K \end{bmatrix} \tag{2.18}$$

Persamaan (2.18) dapat juga ditulis menjadi

$$\boldsymbol{X^T Y} = (\boldsymbol{X^T X})\boldsymbol{\hat{\beta}} \tag{2.19}$$

kemudian ruas kiri dan kanan pada Persamaan (2.19) dikalikan dengan $(\boldsymbol{X^T X})^{-\mathbf{1}}$. Sehingga didapatkan estimator untuk OLS adalah sebagai berikut:

$$\boldsymbol{\hat{\beta}} = (\boldsymbol{X^T X})^{-\mathbf{1}} \boldsymbol{X^T Y} \tag{2.20}$$

B. Fixed Effect Model (FEM)

Menurut Gujarati (2003), *fixed effect model* merupakan model estimasi regresi data panel yang mengasumsikan bahwa nilai intersep dari unit *cross section* atau *time series* berbeda, namun dengan slope koefisien yang tetap. FEM mengestimasi data panel dengan menggunakan variabel *dummy* untuk menangkap adanya perbedaan intersep, model ini juga mengasumsikan bahwa koefisien slope tetap antar individu maupun antar waktu. *Least Square Dummy Variable* (LSDV) merupakan metode estimasi dari pendekatan model ini, dimana LSDV ini dipakai dalam menduga parameter regresi linear dengan menggunakan OLS pada model variabel *dummy* untuk intersep yang berbeda pada setiap individu maupun waktu.

1. FEM Antar Individu

Pada asumsi ini, variasi hanya terjadi pada individu sehingga faktor waktu diabaikan. Adapun persamaan FEM dengan variasi antar individu sebagai berikut (Gujarati,2003):

$$Y_{it} = \beta_0 + \sum_{j=2}^{N} \beta_{0j} D_{ij} + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \tag{2.21}$$

Persamaan (2.21) menunjukkan bahwa pada koefisien β_0 terdapat indeks *j* yang menunjukkan adanya variasi pada unit individu namun tidak terjadi variasi pada waktu. Variasi pada intersep tersebut dapat dinyatakan sebagai berikut:

$$D_{ij} = \begin{cases} 1, & jika\ i = j \\ 0, & jika\ i \neq j \end{cases}$$

Keterangan :

i, j : $1, 2, \dots, N$

D_{ij} : Variabel *dummy* individu ke-i dan *j*

Persamaan (2.21) dapat dijabarkan sebagai berikut:

$$\begin{aligned} Y_{1t} &= \beta_0 + \beta_{02} D_{12} + \cdots + \beta_{0N} D_{1N} + \beta_1 X_{11t} + \cdots + \beta_K X_{K1t} + \varepsilon_{1t} \\ Y_{2t} &= \beta_0 + \beta_{02} D_{22} + \cdots + \beta_{0N} D_{2N} + \beta_1 X_{12t} + \cdots + \beta_K X_{K2t} + \varepsilon_{2t} \\ Y_{3t} &= \beta_0 + \beta_{02} D_{32} + \cdots + \beta_{0N} D_{3N} + \beta_1 X_{13t} + \cdots + \beta_K X_{K3t} + \varepsilon_{3t} \\ &\vdots \\ Y_{Nt} &= \beta_0 + \beta_{02} D_{N2} + \cdots + \beta_{0N} D_{NN} + \beta_1 X_{1Nt} + \cdots + \beta_K X_{KNt} + \varepsilon_{Nt} \end{aligned} \tag{2.22}$$

Dengan mengikuti asumsi variasi intersep pada FEM antara individu, maka Persamaan (2.22) dapat ditulis menjadi :

$$\begin{aligned} Y_{1t} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0N}(0) + \beta_1 X_{11t} + \cdots + \beta_K X_{K1t} + \varepsilon_{1t} \\ Y_{2t} &= \beta_0 + \beta_{02}(1) + \cdots + \beta_{0N}(0) + \beta_1 X_{12t} + \cdots + \beta_K X_{K2t} + \varepsilon_{2t} \\ Y_{3t} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0N}(0) + \beta_1 X_{13t} + \cdots + \beta_K X_{K3t} + \varepsilon_{3t} \\ &\vdots \\ Y_{Nt} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0N}(1) + \beta_1 X_{1Nt} + \cdots + \beta_K X_{KNt} + \varepsilon_{Nt} \end{aligned} \tag{2.23}$$

Persamaan (2.23) dapat dituliskan dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \\ \vdots \\ Y_{Nt} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_{02} \\ \beta_{03} \\ \vdots \\ \beta_{0N} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{21} & \cdots & X_{K1} \\ X_{12} & X_{22} & \cdots & X_{K2} \\ X_{13} & X_{23} & \cdots & X_{K3} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1N} & X_{2N} & \dots & X_{KN} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}$$

atau dapat ditulis secara ringkas menjadi:

$$\begin{aligned} \mathbf{Y}_{N\times 1} &= \boldsymbol{D}_{N\times N}\, \boldsymbol{\beta_0}_{N\times 1} + \boldsymbol{X}_{N\times K}\, \boldsymbol{\beta}_{K\times 1} + \boldsymbol{\varepsilon} \\ &= [\boldsymbol{D} \quad \boldsymbol{X}] \begin{bmatrix} \boldsymbol{\beta_0} \\ \boldsymbol{\beta} \end{bmatrix} + \boldsymbol{\varepsilon} \end{aligned} \tag{2.24}$$

2. FEM Antar waktu

Pada asumsi ini, variasi hanya terjadi pada waktu sehingga faktor individu diabaikan. Adapun persamaan FEM dengan variasi antar waktu sebagai berikut (Gujarati,2003):

$$Y_{it} = \beta_0 + \sum_{s=2}^{T} \beta_{0s} D_{st} + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \tag{2.25}$$

Persamaan (2.25) menunjukkan bahwa pada koefisien β_0 terdapat indeks s yang menunjukkan adanya variasi pada waktu namun tidak terjadi variasi pada unit individu. Variasi pada intersep tersebut dapat dinyatakan sebagai berikut :

$$D_{st} = \begin{cases} 1, & jika\, s = t \\ 0, & jika\, s \neq t \end{cases}$$

Keterangan :

s, t : $1,2, \ldots, T$

D_{st} : Variabel *dummy* waktu ke-t dan s

Persamaan (2.25) dapat dijabarkan sebagai berikut:

$$\begin{aligned} Y_{i1} &= \beta_0 + \beta_{02} D_{21} + \cdots + \beta_{0T} D_{T1} + \beta_1 X_{1i1} + \cdots + \beta_K X_{Ki1} + \varepsilon_{i1} \\ Y_{i2} &= \beta_0 + \beta_{02} D_{22} + \cdots + \beta_{0T} D_{T2} + \beta_1 X_{1i2} + \cdots + \beta_K X_{Ki2} + \varepsilon_{i2} \\ Y_{i3} &= \beta_0 + \beta_{02} D_{23} + \cdots + \beta_{0T} D_{T3} + \beta_1 X_{1i3} + \cdots + \beta_K X_{Ki3} + \varepsilon_{i3} \\ &\vdots \\ Y_{iN} &= \beta_0 + \beta_{02} D_{T2} + \cdots + \beta_{0T} D_{TT} + \beta_1 X_{1iT} + \cdots + \beta_K X_{KiT} + \varepsilon_{iT} \end{aligned} \tag{2.26}$$

Dengan mengikuti asumsi variasi intersep pada FEM antara waktu, maka Persamaan (2.26) dapat ditulis menjadi :

$$\begin{aligned} Y_{i1} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0T}(0) + \beta_1 X_{1i1} + \cdots + \beta_K X_{Ki1} + \varepsilon_{i1} \\ Y_{i2} &= \beta_0 + \beta_{02}(1) + \cdots + \beta_{0T}(0) + \beta_1 X_{1i2} + \cdots + \beta_K X_{Ki2} + \varepsilon_{i2} \\ Y_{i3} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0T}(0) + \beta_1 X_{1i3} + \cdots + \beta_K X_{Ki3} + \varepsilon_{i3} \\ &\vdots \\ Y_{iN} &= \beta_0 + \beta_{02}(0) + \cdots + \beta_{0T}(1) + \beta_1 X_{1iT} + \cdots + \beta_K X_{KiT} + \varepsilon_{iT} \end{aligned} \tag{2.27}$$

Persamaan (2.27) dapat dituliskan dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ \vdots \\ Y_{iT} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_{02} \\ \beta_{03} \\ \vdots \\ \beta_{0T} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{21} & \ldots & X_{K1} \\ X_{12} & X_{22} & \cdots & X_{K2} \\ X_{13} & X_{23} & \cdots & X_{K3} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1T} & X_{2T} & \ldots & X_{KT} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

atau dapat ditulis secara ringkas menjadi:

$$\mathbf{Y}_{\boldsymbol{T}\times\mathbf{1}} = \boldsymbol{D}_{T\times T}\, \boldsymbol{\beta}_{\mathbf{0}_{T\times 1}} + \boldsymbol{X}_{T\times K}\, \boldsymbol{\beta}_{K\times 1} + \boldsymbol{\varepsilon}$$

$$= [\mathbf{D} \quad \mathbf{X}]\begin{bmatrix}\boldsymbol{\beta_0}\\ \boldsymbol{\beta}\end{bmatrix} + \boldsymbol{\varepsilon} \tag{2.28}$$

Untuk menduga masing-masing parameter diperlukan estimasi LSDV. Pada estimasi LSDV, Persamaan (2.24) dan (2.28) didefinisikan sebagai berikut :

$$[\boldsymbol{D} \quad \boldsymbol{X}] = \boldsymbol{M}\ ;\ \begin{bmatrix}\boldsymbol{\beta_0}\\ \boldsymbol{\beta}\end{bmatrix} = \boldsymbol{\theta} \tag{2.29}$$

maka pada Persamaan (2.24) dan Persamaan (2.28) dapat ditulis menjadi

$$y = \boldsymbol{M\theta} + \boldsymbol{\varepsilon} \tag{2.30}$$

Kemudian dilakukan estimasi parameter $\boldsymbol{\theta}$ pada Persamaan (2.30) dengan menggunakan metode OLS dengan cara meminimumkan fungsi total kuadrat *error*.

$$\begin{aligned} S &= \sum_{i=1}^{n} e_{it}^2 \\ &= \sum_{i=1}^{n} \left(Y_{it} - \hat{Y}_{it}\right)^2 \\ &= (\boldsymbol{y} - \boldsymbol{M\theta})^{\mathbf{2}} \\ &= (\boldsymbol{y} - \boldsymbol{M\theta})^T(\boldsymbol{y} - \boldsymbol{M\theta}) \end{aligned} \tag{2.31}$$

Untuk meminimalkan fungsi pada Persamaan (2.31), maka dilakukan penurunan pertama S terhadap $\boldsymbol{\theta}$, kemudian menyamakannya dengan nol.

$$\begin{aligned} \frac{\partial S}{\partial \boldsymbol{\theta}} &= \frac{\partial((\boldsymbol{y} - \boldsymbol{M\theta})^T(\boldsymbol{y} - \boldsymbol{M\theta}))}{d\boldsymbol{\theta}} \\ &= \frac{\partial((\boldsymbol{y}^T - \boldsymbol{\theta}^T\boldsymbol{M}^T)(\boldsymbol{y} - \boldsymbol{M\theta}))}{\partial\boldsymbol{\theta}} \\ &= \frac{\partial(\boldsymbol{y}^T\boldsymbol{y} - \boldsymbol{y}^T\boldsymbol{M\theta} - \boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{y} + \boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M\theta})}{\partial\boldsymbol{\theta}} \\ &= \frac{\partial(\boldsymbol{y}^T\boldsymbol{y} - \boldsymbol{y}^T\boldsymbol{M\theta} - \boldsymbol{y}^T\boldsymbol{M\theta} + \boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M\theta})}{\partial\boldsymbol{\theta}} \\ &= \frac{\partial(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{y}^T\boldsymbol{M\theta} + \boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M\theta})}{\partial\boldsymbol{\theta}} \\ &= 0 - 2\boldsymbol{y}^T\boldsymbol{M} + \boldsymbol{M}^T\boldsymbol{M\theta} + (\boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M})^T \end{aligned}$$

$$-2\boldsymbol{y}^T\boldsymbol{M} + \boldsymbol{M}^T\boldsymbol{M\theta} + \boldsymbol{M}^T\boldsymbol{M\theta} = \mathbf{0}$$

$$-2\boldsymbol{y}^T\boldsymbol{M} + 2\boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M} = 0$$

$$2\boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M} = 2\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{M}$$

$$(\boldsymbol{\theta}^T\boldsymbol{M}^T\boldsymbol{M})^T = (\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{M})^{\boldsymbol{T}}$$

$$\boldsymbol{M}^T\boldsymbol{M}\widehat{\boldsymbol{\theta}} = \boldsymbol{M}^T\boldsymbol{y}$$

Dimana $[\boldsymbol{D} \quad \boldsymbol{X}] = \boldsymbol{M}$ dan $\begin{bmatrix} \boldsymbol{\beta_0} \\ \boldsymbol{\beta} \end{bmatrix} = \boldsymbol{\theta}$, sehingga ditulis menjadi

$$\begin{bmatrix} \boldsymbol{D}^T \\ \boldsymbol{X}^T \end{bmatrix} [\boldsymbol{D} \quad \boldsymbol{X}] \begin{bmatrix} \widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} \\ \widehat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}^T \\ \boldsymbol{X}^T \end{bmatrix} \boldsymbol{y}$$

$$\begin{bmatrix} \boldsymbol{D}^T\boldsymbol{D} & \boldsymbol{D}^T\boldsymbol{X} \\ \boldsymbol{X}^T\boldsymbol{D} & \boldsymbol{X}^T\boldsymbol{X} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} \\ \widehat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}^T \\ \boldsymbol{X}^T \end{bmatrix} \boldsymbol{y}$$

$$\boldsymbol{D}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} + \boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{D}^T\boldsymbol{y} \tag{2.32}$$

$$\boldsymbol{X}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} + \boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y} \tag{2.33}$$

Dari Persamaan (2.32), maka akan diperoleh parameter dari $\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}}$ sebagai berikut:

$$\boldsymbol{D}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} + \boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{D}^T\boldsymbol{y}$$

$$\boldsymbol{D}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} = \boldsymbol{D}^T\boldsymbol{y} - \boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

$$(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} = (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} - (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} = (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} - (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{2.34}$$

Kemudian Persamaan (2.34) disubtitusikan ke dalam Persamaan (2.33) untuk mendapatkan estimasi dari parameter $\widehat{\boldsymbol{\beta}}$.

$$\boldsymbol{X}^T\boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}} + \boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y}$$

$$\boldsymbol{X}^T\boldsymbol{D}\left[(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} - (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}\right] + \boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y}$$

$$\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y}$$

$$\boldsymbol{X}^T\boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} + \boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T]\,\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y}$$

Misalkan $\boldsymbol{D}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T = \boldsymbol{P}$, maka diperoleh

$$\boldsymbol{X}^T\boldsymbol{P}\boldsymbol{y} + \boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{P}]\,\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y}$$

$$\boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{P}]\,\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{P}\boldsymbol{y}$$

$$\boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{P}]\,\boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$$

$$\widehat{\boldsymbol{\beta}} = [\boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{P}]\,\boldsymbol{X}]^{-1}\boldsymbol{X}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y}$$

Sehingga diperoleh estimator untuk LSDV yaitu:

$$\widehat{\boldsymbol{\beta}} = [\boldsymbol{X}^T[\boldsymbol{I} - \boldsymbol{P}]\,\boldsymbol{X}]^{-1}\boldsymbol{X}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{y} \tag{2.35}$$

Pada pemodelan FEM terdapat dua model yang dihasilkan yaitu FEM antar individu dan FEM antar waktu. Dikarenakan Uji Chow membandingkan model FEM dengan model CEM, maka perlu ditentukan model FEM yang akan digunakan untuk perbandingan tersebut. Menurut Gujarati (2003), pemilihan model FEM terbaik antara FEM antar individu dan FEM antar waktu dapat dilakukan dengan melihat koefisien determinasinya. Model dengan koefisien determinasi paling besar yang layak terpilih menjadi model FEM terbaik untuk dibandingkan pada pengujian model terbaik.

C. Random Effect Model (REM)

Menurut Gujarati (2003), *random effect model* merupakan teknik estimasi yang menambahkan variabel gangguan atau *error terms* yang mungkin saja akan muncul pada hubungan antar waktu dan antar individu. Pada REM data diasumsikan terdapat perbedaan intersep untuk setiap individu, sehingga terdapat dua komponen residual yaitu residual secara menyeluruh dan residual secara individu. Residual secara menyeluruh merupakan kombinasi antara *time-series* dan *cross section*, sedangkan residual secara individual merupakan residual dari masing-masing unit *cross section*. Nilai intersep untuk setiap individu dapat dituliskan sebagai berikut:

$$\beta_{0i} = \beta_0 + \mu_i \tag{2.36}$$

Dimana μ_i adalah karakteristik random dari observasi ke-i yang tetap sepanjang waktu. Sehingga dari Persamaan (2.36), maka diperoleh model umum regresi data panel dengan pendekatan REM sebagai berikut (Gujarati,2003):

$$Y_{it} = \beta_0 + \mu_i + \sum_{k=1}^{K} \beta_k X_{kit} + \varepsilon_{it}$$

$$Y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{kit} + \mu_i + \varepsilon_{it} \tag{2.37}$$

dengan asumsi,

$E[\varepsilon_{it}] \quad = E[\mu_i] = 0;\ E[\varepsilon^2{}_{it}] = \sigma^2{}_\varepsilon, E\left[\mu^2{}_i\right] = \sigma^2{}_\mu$

$E[\varepsilon_{it}\mu_i] \ = 0,$ untuk semua i, t dan j

$E\left[\varepsilon_{it}\varepsilon_{js}\right] = 0,$ jika $t \neq s$ atau $i \neq j$

$E\left[\mu_i\mu_j\right] \ = 0,$ jika $i \neq j$

Keterangan:

Y_{it} : Variabel terikat unit individu ke-i dan waktu ke-t

X_{1it} : Variabel bebas unit individu ke-i dan waktu ke-t

β_{0i} : Intersep untuk unit individu ke-i

β_k : Koefisien slope variabel bebas

μ_i : error untuk individu ke-i

ε_{it} : *error* untuk individu ke-i dan waktu ke-t

Matriks varians komponen dari $(\mu_i + \varepsilon_{it})$ untuk unit individu ke-i adalah:

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma^2{}_\mu + \sigma^2{}_\varepsilon & \sigma^2{}_\mu & \cdots & \sigma^2{}_\mu \\ \sigma^2{}_\mu & \sigma^2{}_\mu + \sigma^2{}_\varepsilon & \cdots & \sigma^2{}_\mu \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2{}_\mu & \sigma^2{}_\mu & \cdots & \sigma^2{}_\mu + \sigma^2{}_\varepsilon \end{bmatrix} \tag{2.38}$$

Matriks varians komponen $\boldsymbol{\Omega}$ identic untuk setiap individu, sehingga varians komponen untuk seluruh observasi dapat dituliskan sebagai berikut:

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{\Omega} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Omega} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\Omega} \end{bmatrix} \tag{2.39}$$

Dengan $\boldsymbol{W}$ merupakan matriks non singular dan definit positif, sehingga terdapat matriks $\boldsymbol{A}$ simetrik non singular berukuran $n \times n$ dengan $\boldsymbol{A^T A} = \boldsymbol{AA} = \boldsymbol{W}.$

Didefinisikan variabel-variabel baru sebagai berikut :

$$\boldsymbol{Y}^* = \boldsymbol{A^{-1}Y}, \qquad \boldsymbol{X}^* = \boldsymbol{A^{-1}X}, \qquad \boldsymbol{\varepsilon}^* = \boldsymbol{A^{-1}\varepsilon} \tag{2.40}$$

Metode GLS digunakan untuk mengestimasi parameter model regresi linear berbentuk :

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.41}$$

Persamaan (2.41) diubah sesuai dengan definisi variabel pada (2.38) menjadi:

$$\boldsymbol{A^{-1}Y} = \boldsymbol{A^{-1}X\beta} + \boldsymbol{A^{-1}\varepsilon}$$

$$\boldsymbol{Y}^* = \boldsymbol{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^* \tag{2.42}$$

kemudian dari Persamaan (2.42) akan dicari estimator dari parameternya dengan meminimumkan bentuk kuadrat :

$$\begin{aligned} S &= \textstyle\sum_{i=1}^{n} e_{it}^2 \\ &= \textstyle\sum_{i=1}^{n} \left(Y_{it} - \hat{Y}_{it}\right)^2 \\ &= (\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})^2 \\ &= (\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})^T(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta}) \end{aligned} \tag{2.43}$$

Analogi dengan metode OLS pada Persamaan (2.19), diperoleh persamaan normal metode GLS dalam lambang matriks sebagai berikut:

$$\left(\boldsymbol{X^{*T}X^*}\right)\widehat{\boldsymbol{\beta}} = \boldsymbol{X^{*T}Y^*}$$

$$\left((\boldsymbol{A^{-1}X})^{\boldsymbol{T}}(\boldsymbol{A^{-1}X})\right)\widehat{\boldsymbol{\beta}} = (\boldsymbol{A^{-1}X})^{\boldsymbol{T}}(\boldsymbol{0A^{-1}Y})$$

$$(\boldsymbol{X}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{X})\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{Y})$$
$$(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{Y})$$
$$(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{Y})$$
$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{Y})$$

Sehingga diperoleh estimator untuk GLS yaitu:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{W}^{-1}\boldsymbol{Y}) \tag{2.44}$$

2.4.2 Pemilihan Model Regresi Data Panel Terbaik

Menurut Greene (2003) terdapat tiga jenis uji khusus yang digunakan untuk memilih model regresi data panel terbaik berdasarkan pendekatan model yang telah dilakukan yaitu uji *Chow*, uji *Hausman*, dan uji *Lagrange Multiplier*.

A. Uji *Chow*

Uji Chow merupakan uji yang digunakan untuk mengetahui model regresi data panel terbaik yang telah diperoleh berdasarkan pendekatan CEM dengan model yang telah diperoleh dengan pendekatan FEM. Adapun langkah-langkahnya adalah sebagai berikut:

1. Formulasi hipotesis

   H_0 : $\beta_1 = \beta_2 = \cdots = \beta_K$ (Model yang sesuai adalah CEM)

   H_1 : $\exists\beta_k \neq 0$ (Model yang sesuai adalah FEM)
2. Taraf signifikansi dengan α sebesar 5%
3. Kriteria pengujian

   $\mathrm{H_0}$ ditolak apabila $\mathrm{F_{hitung}} > \mathrm{F_{(\alpha;df(N-1,NT-N-K))}}$ atau *p-value* < α

   $\mathrm{H_0}$ gagal ditolak apabila $\mathrm{F_{hitung}} < \mathrm{F_{(\alpha;df(N-1,NT-N-K))}}$ atau *p-value* ≥ α
4. Statistik uji

   Rumus F_{hitung} dinyatakan sebagai berikut :

$$F_{hitung} = \frac{\frac{(R^2_{FEM} - R^2_{CEM})}{(N-1)}}{\frac{(1 - R^2_{FEM})}{(NT - N - K)}} \tag{2.45}$$

   Keterangan :

   R^2_{FEM} : Koefisien determinasi dari FEM

R_{CEM}^2 : Koefisien determinasi dari CEM
N : Banyaknya unit *cross section*
T : Banyaknya unit *time series*
K : Banyaknya variabel bebas

B. Uji *Hausman*

Uji *Hausman* merupakan uji yang digunakan untuk mengetahui model data panel terbaik diantara model yang diperoleh dengan pendekatan REM dengan model yang diperoleh dari pendekatan FEM. Adapun langkah-langkah pengerjaannya adalah sebagai berikut:

1. Formulasi hipotesis

   $H_0 :\ corr(X_{it}, \varepsilon_{it}) = 0$ (Model yang sesuai adalah REM)

   $H_1 :\ corr(X_{it}, \varepsilon_{it}) \neq 0$ (Model yang sesuai adalah FEM)

2. Taraf signifikansi dengan α sebesar 5%
3. Kriteria pengujian

   H_0 ditolak apabila W $> \mathcal{X}^2_{(\alpha,k)}$ atau *p-value* < α

   H_0 gagal ditolak apabila W $< \mathcal{X}^2_{(\alpha,k)}$ atau *p-value* ≥ α

4. Statistik uji

$$W = \mathcal{X}^2[K-1] = \left[\hat{\beta}_{REM} - \hat{\beta}_{FEM}\right]^T \psi^{-1}\left[\hat{\beta}_{REM} - \hat{\beta}_{FEM}\right]$$

$$\psi = Var\left[\hat{\beta}_{REM}\right] - Var\left[\hat{\beta}_{FEM}\right]$$

$$W = \left[\hat{\beta}_{REM} - \hat{\beta}_{FEM}\right]^T \left[Var\left[\hat{\beta}_{REM}\right] - Var\left[\hat{\beta}_{FEM}\right]\right]^{-1} \left[\hat{\beta}_{REM} - \hat{\beta}_{FEM}\right] \tag{2.46}$$

   Dengan $\hat{\beta}_{GLS}$ adalah vektor estimasi REM dan $\hat{\beta}_{LSDV}$ adalah vektor estimasi FEM.

C. Uji *Lagrange Multiplier*

Uji ini merupakan uji yang digunakan untuk mengetahui model regresi data panel terbaik dari pendekatan REM dengan model yang diperoleh dengan pendekatan CEM. Adapun langkah-langkah prosedur pengerjaannya adalah sebagai berikut:

1. Formulasi hipotesis

   $H_0 :\ \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = 0$ (CEM lebih baik dari REM)

   $H_1 :\ \exists\sigma_i^2 \neq 0$(REM lebih baik dari CEM)
2. Taraf signifikansi dengan α sebesar 5%
3. Kriteria pengujian

   $\mathrm{H_0}$ ditolak apabila LM $> \mathcal{X}^2_{(\alpha,\mathrm{k})}$atau *p-value* < α

   $\mathrm{H_0}$ gagal ditolak apabila LM $< \mathcal{X}^2_{(\alpha,\mathrm{k})}$atau *p-value* > α
4. Statistik uji

$$LM = \frac{NT}{2(T-1)}\left[\frac{\sum_{i=1}^{N}[\sum_{t=1}^{T}\varepsilon_{it}]^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2} - 1\right]^2 \qquad (2.47)$$

2.5 Pengujian Parameter Model Regresi

Pengujian model regresi digunakan untuk mengetahui apakah ada hubungan antara variabel bebas dan variabel terikat. Pengujian model regresi ini terbagi menjadi dua tahap yaitu uji serentak dan uji parsial.

A. Uji Serentak

Uji serentak merupakan uji yang dilakukan untuk mengetahui apakah ada hubungan linear antara variabel terikat dengan variabel bebas, dimana pengujian ini sering dianggap sebagai uji keseluruhan untuk kecukupan model. Uji serentak ini biasa disebut juga dengan uji ANOVA. Uji ini merupakan generalisasi dari analisis varians yang digunakan pada regresi linear berganda (Suyono, 2015).

1. Formulasi hipotesis

   $H_0 :\ \beta_1 = \beta_2 = \cdots = \beta_K$ (Tidak ada pengaruh antara variabel X secara bersama-sama terhadap variabel Y)

   $H_1 :\ \exists\beta_k \neq 0$ (Minimal terdapat satu variabel X yang berpengaruh terhadap variabel Y)
2. Taraf signifikansi dengan α sebesar 5%
3. Kriteria pengujian

   $\mathrm{H_0}$ gagal ditolak apabila $\mathrm{F_{hitung}} \leq \mathrm{F_{(\alpha,NT-K-1)}}$ atau *p-value* ≥ α

   $\mathrm{H_0}$ ditolak apabila $\mathrm{F_{hitung}} > \mathrm{F_{(\alpha,NT-K-1)}}$ atau *p-value* < α
4. Statistik uji

$$F_{hitung} = \frac{\left(\sum_{t=1}^{T}\sum_{t=1}^{T}\left(\hat{Y}_{it} - \overline{Y}_i\right)^2\right)/K}{\left(\sum_{t=1}^{N}\sum_{t=1}^{T}\left(Y_{it} - \hat{Y}_{it}\right)^2\right)/(n-K-1)} \tag{2.48}$$

B. Uji Parsial

Setelah dilakukan uji secara serentak dan didapati hasil bahwa adalah salah satu variabel bebas yang berpengaruh terhadap variabel terikat, maka selanjutnya dilakukan uji secara parsial terhadap setiap variabel bebas (Suyono, 2015). Pengujian ini dilakukan pada $\beta_1, \beta_2, \dots, \beta_k$

1. Formulasi hipotesis

   $H_0 : \beta_k = 0$ (Tidak ada pengaruh variabel X_k terhadap variabel *Y*)

   $H_1 : \beta_k \neq 0$; k = 1, 2, …, K (Ada pengaruh variabel X_k terhadap variabel *Y*)

2. Taraf signifikansi dengan α sebesar 5%
3. Kriteria pengujian

   $\mathrm{H_0}$ gagal ditolak apabila $-\mathrm{t}_{\left(\frac{\alpha}{2};\mathrm{NT-K-1}\right)} < \mathrm{t_{hitung}} < \mathrm{t}_{\left(\frac{\alpha}{2};\mathrm{NT-K-1}\right)}$ atau *p-value* ≥ α

   $\mathrm{H_0}$ ditolak apabila $\mathrm{t_{hitung}} > \mathrm{t}_{\left(\frac{\alpha}{2};\mathrm{NT-K-1}\right)}$ atau $\mathrm{t_{hitung}} < -\mathrm{t}_{\left(\frac{\alpha}{2};\mathrm{NT-K-1}\right)}$ atau *p-value* < α

4. Statistik uji

$$t_{hitung} = \frac{b_k}{SE(b_k)} \tag{2.49}$$

   Keterangan :

   b_k : estimator untuk β_k

   $SE(b_k)$: standar *error* dari estimator b_k

2.6 Pengujian Asumsi

Uji asumsi klasik merupakan persyaratan yang harus dipenuhi pada analisis regresi linear berganda yang berbasis OLS. Walaupun demikian, Basuki dan Prawoto (2016) menyatakan bahwa tidak semua uji asumsi klasik harus dilakukan pada setiap model yang berbasis OLS. Menurutnya, apabila suatu regresi linear diperoleh dari pendekatan OLS, maka uji asumsi klasik yang perlu dilakukan hanya uji multikolinearitas dan uji heteroskedastisitas. Sedangkan untuk regresi linear dengan pendekatan GLS, Ekananda (2016) menyatakan bahwa uji yang dilakukan hanya uji multikolinearitas dan uji normalitas. Uji heteroskedastisitas tidak perlu dilakukan dikarenakan GLS sendiri sudah mampu menyelesaikan masalah heteroskedastisitas.

A. Uji Normalitas

Distribusi normal menurut Gujarati (2003) merupakan suatu distribusi yang relatif sederhana yang hanya melibatkan dua parameter yaitu rata-rata dan varian yang bertujuan untuk untuk mengetahui apakah nilai residual berdistribusi normal atau tidak. Suatu distribusi dikatakan normal apabila kurvanya membentuk lonceng. Untuk melihat apakah suatu distribusi berdistribusi normal maka dapat di uji dengan uji normalitas. Uji normalitas dapat dilakukan dengan menggunakan Uji *Jarque-Bera*. Uji ini menggunakan perhitungan *skewness* (kecondongan) dan *kurtosis* (keruncingan). Hipotesis yang digunakan yaitu:

H_0 : Sebaran data berdistribusi normal

H_1 : Sebaran data tidak berdistribusi normal

Statistik uji yang digunakan yaitu:

$$JB = \frac{n}{6}\left[S^2 + \frac{(Kt-3)^2}{4}\right] \tag{2.50}$$

dengan,

$$S^2 = \left(\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^3}{(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2)^{\frac{3}{2}}}\right)^2 \tag{2.51}$$

dan

$$Kt = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^4}{(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2)^{\frac{3}{2}}} \tag{2.52}$$

dimana,

n : banyaknya data

Kt : kurtosis

S^2 : *skewness*

Kriteria uji yang digunakan yaitu H_0 ditolak apabila JB > $\mathcal{X}^2_{(\alpha,K)}$, yang berarti bahwa error tidak berdistribusi secara normal

B. Uji Multikolinearitas

Multikolinearitas adalah adanya hubungan linear antara variabel bebas. Salah satu asumsi model regresi klasik adalah tidak terdapat hubungan linear antara variabel bebas dalam model regresi. Menurut Gujarati (2003) multikolinearitas digunakan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik multikolinearitas, yaitu adanya hubungan linear antar variabel bebas dalam model regresi. Ada beberapa metode pengujian yang bisa digunakan salah satunya yaitu dengan menggunakan *Variance Inflation Factor* (VIF). *Variance Inflation Factor* dari variabel bebas X didefinisikan sebagai berikut:

$$VIF = \frac{1}{1 - R^2} \tag{2.53}$$

R^2 adalah korelasi ganda antara variabel X dengan semua variabel bebas lainnya. Pada dasarnya tidak ada kriteria untuk menentukan kapan terjadi multikolinearitas, namun beberapa para ahli menyatakan bahwa multikolinearitas menjadi masalah apabila nilai VIF lebih besar dari 10.

C. Uji Heteroskedastisitas

Heteroskedastisitas adalah variasi residual tidak sama untuk semua pengamatan. Uji ini digunakan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik heteroskedastisitas, yaitu adanya ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi. Ada beberapa metode pengujian yang bisa digunakan salah satunya dengan menggunakan uji *White* (Hill, Griffiths, & Lim, 2011). Hipotesis pada uji ini yaitu sebagai berikut:

H_0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$ (tidak terdapat heteroskedastisitas)

H_1 : terdapat minimal satu $\sigma_k^2 \neq \sigma^2$ (terdapat heteroskedastisitas)

Statistik uji yang digunakan yaitu:

$$\mathcal{X}^2{}_{hitung} = nR^2 \tag{2.54}$$

Dimana n merupakan banyaknya data dan R^2 merupakan koefisien determinasi. Distribusi $\mathcal{X}^2{}_{hitung}$ memiliki derajat kebebasan K-1 Kriteria pengujian untuk uji heteroskedastisitas yaitu H_0 ditolak apabila $\mathcal{X}^2_{hitung} > \mathcal{X}^2_{\alpha,(K-1)}$ atau *p-value* < α yang berarti bahwa terjadi heteroskedastisitas pada data yang digunakan.

2.7 Model *Double Log*

Menurut Gujarati & Porter (2004) model regresi memiliki sifat yaitu linear dan nonlinear yang dapat memungkinkan suatu regresi bersifat nonlinear dalam variabel tetapi bersifat linear pada parameter atau dengan kata lain regresi tersebut dibuat sesuai dengan menggunakan transformasi variabel. Transformasi variabel sendiri memiliki berbagai macam cara, salah satunya yaitu dengan menggunakan *double logarithm*. Model dari *double logarithm* (*exponential regression model*) adalah sebagai berikut.

$$\log Y_i = \beta_0 + \beta_1 \log X_i + \beta_2 \log(X_2) + \cdots + \beta_k \log(X_k) + \varepsilon_i \tag{2.55}$$

Model pada Persamaan (2.55) dapat diestimasikan menggunakan OLS. Karena linearitas inilah maka model ini disebut juga sebagai model *log-log, double-log,* atau *log-linear*.

2.8 Produk Domestik Regional Bruto (PDRB)

Keberhasilan pembangunan pada suatu daerah dapat dilihat dari tingkat pertumbuhan ekonominya. Laju pertumbuhan ekonomi suatu wilayah dapat dilihat dari indikator ekonominya yaitu indikator ekonomi regional atau PDRB (Astuti, Susilawati & Suciptawati, 2021). PDRB merupakan nilai tambah bruto dari seluruh barang dan jasa yang tercipta di wilayah domestik suatu negara, dimana hal ini timbul akibat adanya berbagai aktivitas ekonomi dalam suatu periode tertentu tanpa memperhatikan faktor produksi yang dimiliki residen atau non-residen (Badan Pusat Statistik Provinsi Papua, 2022).

Menurut Badan Pusat Statistik, terdapat tiga pendekatan yang dilakukan dalam penyusunan PDRB. Pendekatan itu adalah pendekatan produksi, pendekatan pengeluaran, dan pendekatan pendapatan (Badan Pusat Statistik Provinsi Papua, 2022).

a. Pendekatan Produksi

PDRB dengan pendekatan produksi merupakan PDRB yang tercipta dari jumlah nilai tambah atas barang dan jasa yang dihasilkan oleh berbagai unit produksi di wilayah suatu daerah dalam jangka waktu tertentu. Distribusi PDRB menurut sektor dapat digunakan untuk menunjukkan struktur perekonomian suatu daerah dimana sektor-sektor ekonomi yang memiliki peran besar dalam menunjukkan basis perekonomian suatu daerah.

b. Pendekatan Pendapatan

PDRB dengan pendekatan pendapatan merupakan PDRB yang tercipta dari jumlah balas jasa yang diterima oleh faktor-faktor produksi yang ikut serta dalam proses produksi di suatu daerah dalam periode tertentu. Balas jasa yang dimaksudkan pada pendekatan ini yaitu hasil dari kegiatan produksi tersebut, seperti upah atau gaji, sewa tanah, bunga modal, dan keuntungan. Hasil dari balas jasa tersebut merupakan hasil bersih tanpa adanya pemotongan pajak penghasilan dan pajak langsung lainnya.

c. Pendekatan Pengeluaran

PDRB dengan pendekatan pengeluaran merupakan PDRB atas semua komponen permintaan akhir yang terdiri dari: Pengeluaran Konsumsi Rumah Tangga; Pengeluaran Konsumsi Lembaga Non Profit; Pengeluaran Konsumsi Pemerintah; Pembentukkan Modal Tetap Bruto; Perubahan Inventori atau Stok; dan Ekspor neto. Pendekatan ini sendiri tidak dijelaskan apakah kepemilikan faktor produksi merupakan milik penduduk di wilayah tersebut atau bukan.

Meskipun dalam penyusunannya dibutuhkan tiga pendekatan, tetapi secara konsep tiga pendekatan tersebut akan menghasilkan angka yang sama. Atau dengan kata lain, jumlah pendapatan untuk faktor-faktor produksi akan sama dengan jumlah barang dan jasa akhir yang dihasilkan dan harus sama pula dengan jumlah pengeluarannya.

Setelah dilakukan penyusunan dengan menggunakan tiga pendekatan, nantinya data PDRB akan disajikan dalam dua macam bentuk penyajian yaitu disajikan atas dasar harga berlaku dan atas dasar harga konstan (Badan Pusat Statistik Provinsi Papua, 2022).

1. PDRB Atas Dasar Harga Berlaku (ADHB)
   PDRB atas dasar harga berlaku sering disebut juga dengan PDRB nominal merupakan PDRB yang disusun berdasarkan harga berlaku pada periode perhitungan. Tujuan dari PDRB ini yaitu untuk melihat struktur perekonomian di suatu wilayah..
2. PDRB Atas Dasar Harga Konstan (ADHK)
   PDRB atas dasar harga konstan atau sering disebut juga dengan PDRB riil merupakan PDRB yang disusun berdasarkan harga pada tahun dasar dan bertujuan untuk mengukur pertumbuhan ekonomi secara keseluruhan atau setiap sektor dari tahun ke tahun.

2.9 Pendapatan Asli Daerah

Dalam UU Nomor 23 Tahun 2014 tentang Pemerintahan Daerah Pasal 285 ayat (1) Peraturan Daerah Pasal 21 ayat (1) dan Peraturan Menteri Dalam Negeri Nomor 13 Tahun 2006 tentang Pedoman Pengelolaan Keuangan Daerah Pasal 26 ayat (1) Pendapatan Asli Daerah diartikan sebagai hak pemerintah daerah yang diakui sebagai penambah nilai kekayaan bersih yang diperoleh dari Pajak Daerah, Retribusi Daerah, Hasil Pengelolaan Kekayaan Daerah yang Dipisahkan serta Lain-lain. Selain itu, PAD menurut Nasution (2010) dapat diartikan sebagai pendapatan yang diperoleh daerah yang dipungut berdasarkan peraturan daerah dengan perundang-undangan, dimana bertujuan untuk memberikan hubungan antara pemerintah pusat dan pemerintah daerah.

Pendapatan Asli Daerah memiliki hubungan yang kuat terhadap pertumbuhan ekonomi, dimana menurut Sabilla & Sumarsono (2022) apabila nilai PAD yang dimiliki suatu daerah meningkat, maka menunjukkan bahwa daerah tersebut mandiri dalam mengambil keputusan dan kebijakan pembangunan. Suatu daerah dianggap mandiri apabila seluruh pengeluaran yang dilakukan oleh pemerintah berasal dari dana pendapatan asli daerah tersebut. Dengan meningkatnya PAD akan mengakibatkan dana yang diterima oleh pemerintah daerah juga meningkat, sehingga berdampak pada meningkatnya kemandirian daerah tersebut. Hal tersebut berpotensi dalam meningkatkan pertumbuhan ekonomi di daerah tersebut.

2.10 Belanja Daerah

Belanja pemerintah menurut Sabilla & Sumarsono (2022) merupakan segala bentuk pengeluaran kas pemerintah selama waktu anggaran tertentu, dimana pengeluaran yang dilakukan berdasarkan pada program dan kegiatan yang dapat mengurangi kekayaan pemerintah. Belanja pemerintah sendiri bertujuan untuk memberikan manfaat bagi masyarakat. Belanja pemerintah tercermin dari belanja daerah yang terdapat dalam anggaran pendapatan dan belanja daerah (APBD) Belanja daerah menurut Megasari (2020) merupakan semua bentuk pengeluaran pemerintah daerah pada suatu periode anggaran. Dari definisi tersebut dapat dikatakan bahwa baik belanja pegawai maupun belanja daerah memiliki arti dan tujuan yang sama yaitu bentuk pengeluaran kas pemerintah dalam kurun waktu anggaran tertentu yang bertujuan untuk memberikan manfaat bagi masyarakat di daerah tersebut.

Belanja daerah atau belanja pemerintah memiliki peranan penting dalam meningkatkan pertumbuhan ekonomi di suatu daerah, dimana menurut Megasari (2020) semakin besar belanja daerah maka semakin besar pula dampaknya bagi pertumbuhan ekonomi di daerah tersebut. Hal ini dibuktikan dengan pendapat dari Sabilla & Sumarsono (2022) yang menyatakan bahwa apabila pengeluaran pemerintah meningkat maka Produk Nasional Bruto juga akan meningkat. Begitupun sebaliknya, apabila pengeluaran pemerintah terlalu sedikit maka akan menghambat pertumbuhan ekonomi. Hal tersebut sejalan dengan pernyataan oleh Nasution (2010) yang menyatakan bahwa pengeluaran pemerintah mencerminkan kebijakan pemerintah. Sehingga apabila pemerintah telah menetapkan suatu kebijakan untuk membeli barang maupun jasa, maka pengeluaran tersebut akan mencerminkan biaya yang dikeluarkan oleh pemerintah untuk melaksanakan kebijakan tersebut.

2.11 Tenaga Kerja

Ketenagakerjaan menurut Undang-Undang Republik Indonesia Nomor 13 Tahun 2003 Bab I Pasal 1 ayat 1 merupakan segala hal yang berhubungan dengan tenaga kerja pada waktu sebelum, selama, dan sesudah masa kerja. Dimana pada ayat 2 menyatakan bahwa tenaga kerja merupakan setiap orang yang mampu melakukan pekerjaan guna menghasilkan barang dan/atau jasa baik untuk memenuhi kebutuhan sendiri maupun untuk masyarakat. Kemudian dalam Bab II Pasal 4 menjelaskan bahwa Pembangunan ketenagakerjaan bertujuan untuk; memberdayakan dan memberdayagunakan tenaga kerja secara optimal dan manusiawi; mewujudkan pemerataan kesempatan kerja dan penyediaan tenaga kerja yang sesuai dengan kebutuhan pembangunan nasional dan daerah; memberikan perlindungan kepada tenaga kerja dalam mewujudkan kesejahteraannya; dan meningkatkan kesejahteraan tenaga kerja dan keluarganya.

Tenaga kerja secara spesifik berpengaruh terhadap pertumbuhan ekonomi. Ahrizal (2022) menyatakan bahwa tenaga kerja dapat berpengaruh terhadap pertumbuhan ekonomi, karena dengan terciptanya lapangan pekerjaan akan berdampak pada kesempatan kerja yang penuh (*full employment*), sehingga dapat meminimalkan pengangguran yang bersifat permanen atau jangan panjang. Dimana para ekonom klasik percaya bahwa pengangguran bersifat sementara, sehingga dengan adanya penurunan pengangguran yang bersifat jangka panjang akan berdampak pada perekonomian ekonomi di daerah tersebut.

2.12 Rata-rata Lama Sekolah

Pendidikan menurut Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 Bab I Pasal 1 ayat 1 merupakan usaha sadar dan terencana untuk mewujudkan belajar dan proses pembelajaran agar peserta didik dapat secara aktif mengembangkan potensi diri untuk meningkatkan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keperluan untuk dirinya, masyarakat, bangsa dan negara. Menurut Rasnino, Nuryadin, dan Suharsih (2022) pendidikan merupakan salah satu investasi manusia yang dapat menghasilkan sumber daya manusia menjadi lebih baik. Sumber daya manusia yang baik menjadi salah satu faktor yang sangat kuat bagi pertumbuhan ekonomi di suatu wilayah.

Keberhasilan pengembangan sumber daya manusia dapat dilihat melalui tingkat keberhasilan pendidikan di suatu daerah. Menurut Badan Pusat Statistik salah satu indikator yang menunjukkan keberhasilan pendidikan di suatu daerah yaitu rata-rata lama sekolah (RLS). RLS sendiri merupakan metode baru dalam perhitungan kualitas pendidikan di Indonesia, dimana RLS menggantikan indikator pendidikan sebelumnya yaitu angka melek huruf. Hal ini dikarenakan RLS dapat mengGambarkan situasi yang lebih relevan terhadap pendidikan dan perubahan yang terjadi. RLS menunjukkan jenjang pendidikan yang pernah atau sedang diduduki oleh seseorang, semakin tinggi angka RLS maka semakin tinggi jenjang pendidikan yang diselesaikan (Badan Pusat Statistik Provinsi Papua, 2022).

2.13 Penelitian Terdahulu

Penelitian ini mengacu pada beberapa penelitian terdahulu baik dalam metode maupun variabel yang digunakan, dengan maksud agar hasil yang dihasilakn valid dan baik. Dalam penelitian ini sendiri mengacu pada 6 penelitian yang sudah pernah dilakukan sebelumnya. Berikut merupakan penelitian terdahulu yang digunakan sebagai acuan pada penelitian ini.

Tabel 2.1 Penelitian Terdahulu

No	Nama	Judul	Hasil Penelitian
1.	Yozi Aulia Rahman, dan Ayunda Lintang Chamelia (2015)	Faktor-Faktor Yang Mempengaruhi PDRB Kabupaten/Kota Jawa Tengah Tahun 2008-2012	Model analisis yang digunakan adalah analisis regresi data panel, dimana model yang terpilih yaitu FEM dengan hasil penelitian yang menunjukkan bahwa Pendapatan Asli Daerah, Realisasi Belanja Daerah, Tabungan, dan Posisi Kredit berhubungan positif dengan PDRB di Kabupaten/Kota di Jawa Tengah.
2.	Machmud Al Amrie, Adi Aspian Nur, dan	Pengaruh Belanja Daerah Serta Tenaga Kerja	Hasil analisis menunjukkan bahwa baik Belanja Daerah maupun Tenaga Kerja

No	Nama	Judul	Hasil Penelitian
	Amelia Ramadhani (2017)	Terhadap Produk Domestik Regional Bruto Provinsi Kalimantan Utara	berpengaruh secara signifikan terhadap Produk Domestik Regional Bruto di Provinsi Kalimantan Utara. Model analisis yang digunakan yaitu analisis regresi linear berganda.
3.	Megasari (2020)	Analisis Pengaruh Pendapatan Asli Daerah dan Belanja Daerah Terhadap Pertumbuhan Ekonomi di Kabupaten Luwu Utara	Hasil penelitian dengan model analisis yaitu regresi linear berganda, didapati hasil bahwa Pendapatan Asli Daerah secara parsial tidak berpengaruh terhadap pertumbuhan ekonomi di Kabupaten Luwu Utara. Sedangkan Belanja Daerah berpengaruh secara parsial terhadap pertumbuhan ekonomi di Kabupaten Luwu Utara.
4.	Ghaly Rizquillah Ahrizal (2022)	Analisis PDRB 34 Provinsi di Indonesia Pada Masa Pemerintahan Presiden Joko Widodo	Model analisis yang digunakan yaitu analisis regresi data panel. Dimana model yang terpilih yaitu FEM dengan hasil analisis yaitu variabel pendidikan dan tenaga kerja memiliki pengaruh yang signifikan terhadap PDRB, sedangkan variabel indeks pembangunan teknologi, informasi, dan komunikasi, dan PMDN tidak memiliki pengaruh yang signifikan terhadap PDRB.
5.	Tara May Sabilla, dan Hadi Sumarsono (2022)	Pengaruh Belanja Pemerintah, Pendapatan Asli Daerah, Penanaman Modal Dalam Negeri, Indeks Pembangunan Manusia Terhadap PDRB	Model analisis yang digunakan yaitu analisis regresi data panel, dimana model yang terpilih yaitu FEM. Hasil analisis menunjukkan bahwa belanja daerah, Pendapatan Asli Daerah dan IPM memiliki hubungan signifikan dan positif terhadap PDRB di Jawa Timur periode 2015-2020. Sedangkan PMDN memiliki pengaruh yang tidak signifikan terhadap PDRB di Jawa Timur periode 2015-2020.

No	Nama	Judul	Hasil Penelitian
6.	Cass A. Rasnino, Didi Nuryadin, dan Sri Suharsih (2022)	Pengaruh Angka Harapan Hidup, Rata-rata Lama Sekolah dan Konsumsi Rumah Tangga Terhadap Pertumbuhan Ekonomi di Kabupaten/Kota Provinsi Lampung tahun 2014-2019	Metode analisis yang digunakan yaitu analisis regresi data panel dengan pendekatan FEM. Hasil analisis menunjukkan bahwa RLS dan konsumsi rumah tangga berpengaruh secara positif dan signifikan terhadap pertumbuhan ekonomi di Kabupaten/Kota Lampung tahun 2014-2019. Sedangkan angka harapan hidup tidak berpengaruh terhadap pertumbuhan ekonomi di Kabupaten/Kota Lampung tahun 2014-2019.