PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* PADA FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEMISKINAN DI PULAU SUMATERA MENGGUNAKAN METODE *GENERALIZED LEAST SQUARE*

(Skripsi)

Oleh

SISILIA AMARA
2017031010

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

ABSTRAK

PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* PADA FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEMISKINAN DI PULAU SUMATERA MENGGUNAKAN METODE *GENERALIZED LEAST SQUARE*

Oleh

SISILIA AMARA

Model regresi data panel adalah model yang digunakan untuk mengetahui pengaruh satu atau beberapa peubah independen terhadap suatu peubah dependen dengan struktur data berupa data panel. Data panel merupakan kombinasi dari data *cross section* dan *time series.* Tujuan penelitian ini adalah untuk menganalisis pengaruh dari Pengeluaran Perkapita disesuaikan, Tingkat Pengangguran Terbuka, dan Rata Rata Lama Sekolah terhadap Tingkat kemiskinan di Pulau Sumatera. Setiap variabel independen pada masing-masing provinsi dilakukan peramalan menggunakan analisis *trend*. Hasil peramalan tersebut disubstitusikan ke dalam model regresi data panel yang telah diperoleh untuk meramalkan tingkat kemiskinan pada tahun 2023-2025. Hasil dari penelitian ini didapatkan model regresi data panel untuk memodelkan Tingkat Kemiskinan di Pulau Sumatera tahun 2015-2022 adalah *Random Effect Model* (REM) dengan estimasi GLS dimana variabel Pengeluaran Perkapita disesuaikan, variabel Tingkat Tengangguran Terbuka, dan variabel Rata-Rata Lama Sekolah mampu menjelaskan variabel tingkat kemiskinan di Pulau Sumatera sebesar 73,38% sedangkan sisanya sebesar 26,62% dijelaskan oleh variabel lain diluar model. Dengan model persamaan hasil estimasi yaitu $\hat{Y}_{it} = 25{,}7186 - 0{,}0008X_{1it} + 0{,}1125X_{2it} - 0{,}7977X_{3it}$. Hasil peramalan menunjukan bahwa tingkat kemiskinan pada tahun 2023-2025 mengalami penurunan pada beberapa provinsi dan juga mengalami kenaikan di provinsi lainnya.

Kata kunci: data panel, tingkat kemiskinan, *random effect model,* analisis *trend*

ABSTRACT

REGRESSION MODELING OF RANDOM EFFECT PANEL DATA ON FACTORS AFFECTING POVERTY LEVELS ON THE ISLAND OF SUMATRA USING THE GENERALIZED LEAST SQUARE METHOD

By

SISILIA AMARA

The panel data regression model is a model used to determine the influence of one or several independent variables on a dependent variable with a data structure in the form of panel data. Panel data is a combination of cross section and time series data. The aim of this research is to analyze the influence of adjusted per capita expenditure, open unemployment rate, and average length of schooling on the poverty level on the island of Sumatra. Each independent variable in each province is forecasted using *trend* analysis. The forecasting results are substituted into the panel data regression model that has been obtained to predict the poverty level in 2023-2025. The results of this research obtained a panel data regression model to model the level of poverty on the island of Sumatra in 2015-2022, namely the Random Effect Model (REM) with GLS estimation where the Per Capita Expenditure variable is adjusted, the Open Unemployment Rate variable, and the Average Years of Schooling variable are able to explain The poverty level variable on Sumatra Island is 73,38% while the remaining 26.62% is explained by other variables outside the model. With the equation model, the estimation results are: $\hat{Y}_{it} = 25{,}7186 - 0{,}0008X_{1it} + 0{,}1125X_{2it} - 0{,}7977X_{3it}$. The forecasting results show that the poverty rate in 2023-2025 has decreased in some provinces and also increased in others.

Kata kunci : panel data, poverty level, random effect model, trend analysis

PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* PADA FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEMISKINAN DI PULAU SUMATERA MENGGUNAKAN METODE *GENERALIZED LEAST SQUARE*

Oleh

SISILIA AMARA

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024

Judul Skripsi	: **PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* PADA FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEMISKINAN DI PULAU SUMATERA MENGGUNAKAN METODE *GENERALIZED LEAST SQUARE***
Nama Mahasiswa	: Sisilia Amara
Nomor Pokok Mahasiswa	: **2017031010**
Jurusan	: **Matematika**
Fakultas	: **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Prof. Drs. Mustofa Usman, MA., Ph.D	**Siti Laelatul Chasanah, S.Pd., M.Si.**
NIP. 19570101 198403 1 020	**NIP. 19930601 201903 2 021**

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Prof. Drs. Mustofa Usman, M.A., Ph.D.** ....

Sekretaris : **Siti Laelatul Chasanah, S.Pd., M.Si.** ....

Penguji
Bukan Pembimbing : **Dr. Khoirin Nisa, S.Si., M.Si.** ....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : 9 Juli 2024

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini :

Nama	: **Sisilia Amara**
Nomor Pokok Mahasiswa	: **2017031010**
Jurusan	: **Matematika**
Judul Skripsi	: **PEMODELAN REGRESI DATA PANEL *RANDOM EFFECT* PADA FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEMISKINAN DI PULAU SUMATERA MENGGUNAKAN METODE *GENERALIZED LEAST SQUARE***

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila di kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 9 Juli 2024
Penulis,

METERAI TEMPEL
95F46ALX236914349

Sisilia Amara
NPM. 2017031010

RIWAYAT HIDUP

Penulis bernama lengkap Sisilia Amara, lahir di Bukit Kemuning pada tanggal 25 Agustus 2002. Penulis merupakan anak pertama dari empat bersaudara yang merupakan anak dari pasangan Bapak Odo Waker Amru, S.Pd. dan Ibu Emawaty S.Pd.

Penulis menempuh pendidikan di Taman Kanak-kanak (TK) Muslimin Bukit Kemuning pada tahun 2007-2008, melanjutkan ke Sekolah Dasar Negeri (SDN) 3 Bukit Kemuning pada tahun 2008-2014, Sekolah Menengah Pertama (SMP) Negeri 1 Bukit Kemuning pada tahun 2014-2017, dan Sekolah Menengah Atas (SMA) YP UNILA Bandar Lampung pada tahun 2017-2020.

Pada tahun 2020 penulis melanjutkan pendidikan di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswi, penulis aktif di organisasi sebagai Anggota Bidang Minat dan Bakat Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) periode 2021. Kemudian pada tahun 2022 aktif di organisasi Badan Eksekutif Mahasiswa (BEM) FMIPA Universitas Lampung sebagai Anggota Dinas Kominfo (Komunikasi dan Informasi).

Pada tahun 2023 penulis melaksanakan Kerja Praktik (KP) di Dinas Kominfo dan Statistik (DISKOMINFOTIK) Provinsi Lampung sebagai bentuk penerapan ilmu yang telah penulis peroleh selama kuliah. Sebagai bentuk pengabdian dan pelaksanaan Tri Dharma Perguruan Tinggi, penulis pada bulan Juni hingga Agustus 2023 melaksanakan Kuliah Kerja Nyata (KKN) yang dilaksanakan di Kampung Sidomulyo, Kecamatan Bangun Rejo, Kabupaten Lampung Tengah.

KATA INSPIRASI

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya."
(QS. Al-Baqarah: 286)

"Sesungguhnya bersama kesulitan itu ada kemudahan. Apabila engkau telah selesai (dari suatu urusan), teruslah bekerja keras (untuk urusan yang lain). Dan hanya kepada Tuhanmulah engkau berharap."
(QS. Al-Insyirah: 6-8)

"Bukan kesulitan yang membuat kita takut, tapi sering ketakutanlah yang membuat jadi sulit. Jadi jangan mudah menyerah.
-Joko Widodo-

"Sekali lagi kuatkan bahunya, minta restu orang tua mu karena masih banyak hal-hal baik dan indah yang menunggu mu diluar sana."
Penulis-

PERSEMBAHAN

Puji Syukur kepada Allah SWT atas limpahan rahmat, taufiq, hidayah serta karunia-Nya. Dengan ketulusan hati dan ucapan terimakasih kupersembahkan karya sederhana dengan penuh perjuangan ini kepada:

Ayah dan Ibu tercinta

Terima kasih Ayah dan Ibu yang tiada hentinya selalu memberikan kasih sayang, doa, dan motivasi dengan penuh ke ikhlasan yang tak terhingga. Terimakasih atas segala hal yang telah diberikan untukku.

Adik-adikku tersayang

Yang telah membantu dan memberikan semangat.
Terimakasih sudah menjadi tempat untuk berbagi cerita.

Dosen Pembimbing dan Pembahas

Terima kasih kepada dosen pembimbing dan pembahas yang sangat berjasa, selalu membimbing, memberikan arahan dan juga ilmu yang berharga kepada penulis.

Almamater Tercinta, Universitas Lampung

SANWACANA

Puji Syukur kehadirat Allah SWT, atas segala rahmat dan karunianya sehingga penulis dapat menyelesaikan skripsi dengan judul "Pemodelan Regresi Data Panel *Random Effect* Pada Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan di Pulau Sumatera Menggunakan Metode *Generalized Least Square*". Penulis menyadari bahwa skripsi ini tidak akan selesai dengan baik tanpa adanya arahan, bimbingan, serta kritik dan saran dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis iningin mengucapkan terimakasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, MA., Ph.D., selaku Pembimbing I yang selalu bersedia memberikan waktunya untuk memberikan arahan, bimbingan dan dukungan yang membangun sehingga penulis selalu dipermudah dalam proses penyelesaian skripsi ini.
2. Ibu Siti Laelatul Chasanah, S.Pd., M.Si., selaku Pembimbing II yang telah memberikan waktunya untuk memberi bimbingan, saran serta dukungan kepada penulis selama penyusunan skripsi ini.
3. Ibu Dr. Khoirin Nisa, S.Si, M.Si., selaku dosen Pembahas yang telah bersedia memberikan kritik dan saran serta evaluasi kepada penulis.
4. Ibu Dina Eka Nurvazly, M.Si., dan Alm. Bapak Amanto S.Si., M.Si selaku dosen Pembimbing Akademik atas bimbingannya selama ini.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen, staf dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang telah memberikan

wawasan, ilmu, dan pengetahuan yang sangat berharga bagi penulis selama proses perkuliahan berlangsung.

8. Kedua orang tua ku ayahanda Odo Waker Amru, S.Pd., dan Ibu Emawaty, S.Pd., terimakasih sebesar-besarnya karena telah mendidik penulis, memberikan kasih sayang, motivasi, dan semangat yang tiada hentinya ketika penulis merasa putus asa dan tidak mampu. Ibu dan Ayah menjadi penguat dan pengingat yang paling hebat. Terimakasih sudah menjadi tempatku untuk pulang. Gelar ini ku persembahkan untuk kalian.
9. Adik-adikku tersayang Meylin Sabrina, Zaskia Mauliya Azhar dan Citra Khazanah Amru terimakasih telah menghibur dan memberikan semangat serta doa kepada penulis. Tumbuhlah kalian menjadi versi yang paling hebat
10. Untuk rumah kedua ku Bapak, Mamah, Hilal, mba Inggit dan Adek terimakasih sudah menerimaku dengan baik, selalu peduli senantiasa mendukung dan memotivasi semoga kalian selalu diberikan rezeki berlimpah dan sehat selalu.
11. Sahabat terbaik penulis Hilal, Yulian, Fegy, Nispril, Winda, Hanafi, terimakasih telah bersedia menjadi tempat keluh kesah selalu memberikan motivasi dan dukungan tanpa henti sehingga penulis bisa menyelesaikan skripsi ini. Kalian adalah orang orang pilihan yang berada di balik layar, membersamai dalam perjuangan dan selalu mau penulis repotkan. Terimakasih untuk banyak momen sekaligus pelajaran hidup selama perkuliahan.
12. *Last but not least*, terimakasih kepada diri saya sendiri atas kerja keras dan semangat nya sehingga tidak pernah menyerah dalam mengerjakan tugas akhir skripsi ini. Tetap rendah hati karena ini baru awal dari semuanya.

Bandar Lampung, 9 Juli 2024
Penulis

Sisilia Amara

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi adalah analisis yang digunakan untuk menganalisis pengaruh variabel independen terhadap variabel dependen (Gujarati & Porter, 2012). Secara umum, analisis regresi terbagi menjadi dua yaitu regresi linear sederhana dan regresi linear berganda. Menurut Alamsyah, dkk. (2022), regresi linear sederhana digunakan untuk mengetahui hubungan antar satu variabel independen dan variabel dependen. Sebaliknya, regresi linear berganda digunakan untuk mendapatkan hubungan antar dua atau lebih variabel independen dan variabel dependen. Terdapat 3 jenis data yang digunakan dalam analisis regresi yaitu data *cross section*, *time series*, dan data panel (Basuki, 2016).

Data panel merupakan gabungan antara data *cross section* dan data *time series.* Data *cross section* adalah data yang diperoleh dengan mengamati banyak subjek dalam waktu yang sama, sedangkan data *time series* adalah data yang diperoleh dari pengamatan satu objek pada beberapa periode waktu tertentu (Alamsyah, dkk., 2022). Analisis regresi data panel merupakan suatu metode yang digunakan untuk memodelkan pengaruh variabel independen terhadap variabel dependen selain itu, analisis regresi data panel juga digunakan untuk melakukan peramalan variabel dependen pada setiap *cross section* yang ada. Namun, untuk meramalkannya perlu dilakukan peramalan terlebih dahulu untuk variabel independen pada masing-masing *cross section*. Terdapat tiga pendekatan yang biasanya dilakukan dalam mengestimasi parameter model regresi data panel yaitu pendekatan *common effect model* (CEM), *fixed effect model* (FEM), dan *random effect model* (REM).

Metode REM dapat diestimasi menggunakan metode *Generalized Least Square* (GLS). Metode GLS menghasilkan estimasi yang takbias, konsisten, dan varians minimum. Selain itu, metode ini mampu mengatasi pelanggaran asumsi heterokedastisitas dan autokorelasi yang sering terjadi pada model yang menggunakan metode *Ordinary Least Square* (OLS). Dengan demikian, metode GLS dapat menghasilkan model yang baik untuk mengatasi gejala heterokedastisitas dan autokorelasi antar variabel pengamatan. Salah satu masalah yang dapat diselesaikan dengan menggunakan metode *random effect model* adalah memodelkan faktor-faktor yang berpengaruh pada tingkat kemiskinan.

Menurut Badan Pusat Statistik (2022), kemiskinan merupakan kondisi ketika ekonomi tidak dapat memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran, sedangkan tingkat kemiskinan adalah persentase jumlah penduduk yang berada dibawah garis kemiskinan. Tingkat kemiskinan merupakan masalah yang dihadapi di berbagai negara, termasuk Indonesia. Indonesia berada di peringkat ke enam dengan tingkat kemiskinan tertinggi di Asia Tenggara dengan persentase sebesar 9,5% per Maret tahun 2023 (GoodStats, 2023).

Dari beberapa pulau di Indonesia, Pulau Sumatera menduduki peringkat ke dua dengan jumlah penduduk miskin terbanyak. Jumlah penduduk miskin di Sumatera tercatat sebanyak 5,67 juta jiwa atau 21,89% dari total keseluruhan penduduk miskin per akhir Maret 2023 (GoodStats, 2023). Salah satu cara yang dapat digunakan untuk mengidentifikasi faktor-faktor yang mempengaruhi tingkat kemiskinan adalah pendekatan regresi data panel model *random effect* dengan metode *generalized least square* serta meramalkan tingkat kemiskinan dengan faktor faktor yang diduga berpengaruh seperti pengeluaran perkapita disesuaikan, tingkat pengangguran terbuka, dan rata rata lama sekolah.

Beberapa penelitian telah dilakukan dengan menggunakan data panel. Nandita (2019) menggunakan regresi data panel untuk mengetahui faktor-faktor yang mempengaruhi PDRB di Provinsi DIY tahun 2011-2015. Rizal dan Yantieka (2022) menganalisis kemiskinan Pulau Sumatera dengan variabel pendidikan tahun 2016-2021 menggunakan regresi data panel. Pada tahun 2022 Septianingsih melakukan pemodelan data panel menggunakan *random effect model* untuk mengetahui faktor-faktor yang mempengaruhi umur harapan hidup di Indonesia. Saputri, dkk. (2020) melakukan analisis regresi data panel pada frekuensi penerbangan dan volume *cargo* terhadap jumlah penumpang pesawat tahun 2018. Kemudian Hijrawati (2022) menganalisis faktor-faktor yang mempengaruhi tingkat kemiskinan Provinsi Sulawesi Tenggara tahun 2017-2020 menggunakan REM. Oleh karena itu, pada penelitian ini penulis akan membahas tentang pemodelan dan peramalan regresi data panel *random effect* pada faktor-faktor yang mempengaruhi tingkat kemiskinan di Pulau Sumatera menggunakan metode *generalized least square.*

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menentukan model tingkat kemiskinan di Pulau Sumatera menggunakan model *random effect* dengan metode *generalized least square*
2. Melakukan peramalan pada data tingkat kemiskinan di Pulau Sumatera untuk periode yang akan datang.

1.3 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini yaitu:

1. Menambah wawasan dan pengetahuan bagi penulis dan pembaca tentang analisis regresi data panel.
2. Sebagai referensi bagi penelitian selanjutnya dalam pengolahan data menggunakan data panel.
3. Dapat mengaplikasikan estimasi model regresi data panel *random effect* dengan metode *generalized least square* yang dapat digunakan untuk meramalkan data tingkat kemiskinan di Pulau Sumatera untuk periode mendatang.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Salah satu teknik analisis yang dapat digunakan untuk memodelkan dan menganalisis hubungan antar variabel adalah analisis regresi (Montgomery & Peck, 1992). Dalam analisis regresi, terdapat dua jenis variabel yaitu variabel terikat (dependen) dan variabel bebas (independen). Variabel independen merupakan variabel yang mempengaruhi perubahan variabel dependen.

Terdapat dua jenis analisis regresi yaitu analisis regresi linear dan analisis regresi non-linear. Analisis regresi linear terdiri dari analisis regresi linear sederhana dan analisis regresi linear berganda. Regresi linear sederhana berfungsi untuk menguji secara linear hubungan antara variabel penyebab (variabel independen) dengan variabel akibatnya (variabel dependen). Model regresi linear sederhana dapat ditulis pada persamaan (2.1) (Pangestika, 2015).

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \;;\; i = 1,2,\dots,n \tag{2.1}$$

Sebaliknya, regresi linear berganda adalah hubungan antara dua atau lebih variabel independen dan variabel dependen. Model regresi linear berganda dengan k variabel independen dapat ditulis pada persamaan (2.2) (Montgomery & Peck, 1992).

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \tag{2.2}$$

dengan,

$$i = 1,2,\dots,n$$

keterangan:

Y_i = nilai variabel dependen pada pengamatan ke- i

X_{ik}	= nilai variabel independen ke- k dengan $k = 1,2, \ldots, n$
β_0	= intersep model regresi
$\beta_{1,}\beta_2, \ldots, \beta_k$	= koefisien *slope*
ε_i	= nilai *error* ; $\varepsilon_i \sim N(0, \sigma^2)$
k	= jumlah variabel independen
n	= jumlah data pengamatan

Parameter i menunjukan pengamatan, maka n memiliki persamaan yaitu:

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_k X_{1k} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_k X_{2k} + \varepsilon_2 \\
&\vdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_k X_{nk} + \varepsilon_n
\end{aligned}
\tag{2.3}
$$

Persamaan (2.3) dalam bentuk matriks ditulis sebagai berikut:

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.4}
$$

Persamaan (2.4) dapat ditulis dalam bentuk matriks menjadi :

$$
\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.5}
$$

dengan,

$$
\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.6}
$$

keterangan:

$\boldsymbol{Y}$ = vektor kolom dari variabel dependen yang berukuran $n \times 1$

$\boldsymbol{X}$ = matriks dari variabel independen yang berukuran $n \times k$

$\boldsymbol{\beta}$ = vektor kolom dari parameter yang berukuran $k \times 1$

$\boldsymbol{\varepsilon}$ = vektor kolom dari *error* yang berukuran $n \times 1$

Menurut Widarjono (2005), model linear regresi berganda memiliki beberapa asumsi, antara lain:

1. Hubungan antara variabel dependen (Y) dan variabel independen (X) adalah linear dalam parameter.
2. Tidak adanya hubungan linear antar variabel independen atau tidak terdapat multikolinearitas antar variabel independen.
3. Nilai harapan atau rata-rata dari *error* ε_i adalah nol (0).

$$E(\varepsilon_i) = 0 \tag{2.7}$$

dengan $i = 1,2,\dots,n$. Persamaan (2.7) dapat dituliskan dalam bentuk matriks pada persamaan (2.8).

$$E(\boldsymbol{\varepsilon}) = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ E(\varepsilon_I) \\ \vdots \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0} \tag{2.8}$$

4. Tidak terdapat korelasi antar *error* (ε_i) dan *error* lainnya (ε_j).

$$E\left(\varepsilon_i \varepsilon_j\right) = 0\,;\, i \neq j \tag{2.9}$$

5. Varian dari setiap *error* adalah sama (homokedastisitas).

$$\begin{aligned} var(\boldsymbol{\varepsilon}) &= E[\varepsilon_i - E({\varepsilon_i}^2)];\, i = 1,2,\dots,n \\ &= E({\varepsilon_i}^2) \\ &= \sigma^2 \end{aligned} \tag{2.10}$$

Persamaan (2.10) jika dituliskan dalam bentuk dapat ditulis pada persamaan (2.11).

$$\begin{aligned} Var(\boldsymbol{\varepsilon}) &= E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') \\ &= E\left(\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \right) \\ &= \begin{bmatrix} E({\varepsilon_1}^2) & E(\varepsilon_1\varepsilon_2) & \dots & E(\varepsilon_1\varepsilon_n) \\ E(\varepsilon_2\varepsilon_1) & E({\varepsilon_2}^2) & \vdots & E(\varepsilon_2\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_i\varepsilon_1) & E(\varepsilon_i\varepsilon_2) & \vdots & E(\varepsilon_i\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_n\varepsilon_1) & E(\varepsilon_n\varepsilon_2) & \cdots & E({\varepsilon_n}^2) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \sigma_\varepsilon{}^2 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon{}^2 & \vdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \square & \vdots \\ 0 & 0 & \vdots & \sigma_\varepsilon{}^2 & \square & 0 \\ \vdots & \vdots & \square & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 0) & \cdots & \sigma_\varepsilon{}^2 \end{bmatrix}$$

$$= \sigma_\varepsilon{}^2 \boldsymbol{I} \tag{2.11}$$

6. *Error* berdistribusi normal.

$$\varepsilon_i \sim N(0, \sigma^2); i = 1,2, \dots, n. \tag{2.12}$$

2.2 Analisis Regresi Data Panel

Data panel merupakan kombinasi antara data *time series* dan *cross section* di mana observasi dilakukan pada berbagai entitas seperti perusahaan atau negara, yang diamati sepanjang periode waktu tertentu. Pendekatan ini memungkinkan analisis terhadap variabilitas antar individu (*cross-sectional*) dan dalam waktu (*time series*), yang dapat memberikan wawasan mendalam mengenai dinamika dan hubungan antar variabel dalam konteks yang lebih luas (Widarjono, 2005).

Analisis regresi data panel merupakan suatu teknik yang dapat digunakan untuk memodelkan pengaruh variabel independen terhadap variabel dependen di beberapa wilayah yang diamati dari suatu populasi penelitian dalam jangka *time series* tertentu. Variabel dependen adalah variabel akibat atau variabel yang dipengaruhi sedangkan variabel independen adalah variabel yang mempengaruhi (Kusumawati, dkk., 2017).

Model regresi data panel secara umum dapat dinyatakan pada persamaan (2.13) (Baltagi, 2005).

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it} \tag{2.13}$$

keterangan:

Y_{it} = variabel dependen untuk unit *cross section* ke-*i* dan *time series* ke-*t*

X_{it} = variabel independen untuk unit *cross section* ke-*i* dan *time series* ke-*t*

β_0 = intersep

β_1 = *slope* untuk semua unit

ε_{it} = nilai *error* untuk *cross section* ke-*i* dan *time series* ke-*t*

Menurut Hsiao (2003), terdapat beberapa keuntungan dalam penggunaan data panel di dalam penelitian, yaitu sebagai berikut:

1. Data panel merupakan kombinasi antara data *cross section* dan data *time series*, memungkinkan untuk meningkatkan jumlah data yang dapat meningkatkan derajat bebas dan dengan demikian dapat mengurangi adanya multikolinearitas antar variabel independen.
2. Dengan menggabungkan informasi dari data *time series* dan *cross section* dapat mengurangi masalah yang muncul ketika variabel dihilangkan.

Menurut Jaya & Sunengsih (2009), analisis regresi data panel adalah regresi yang didasarkan pada data panel untuk mengamati hubungan antara satu variabel tak bebas (dependen variabel) dengan satu atau lebih variabel bebas (independent variabel). Beberapa alternatif model yang dapat diselesaikan dengan data panel yaitu :

1. Semua koefisien baik intersep maupun *slope* koefisien konstan.

$$Y_{it} = \beta_0 + \sum_{K=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \tag{2.14}$$

2. *Slope* koefisien konstan tetapi intersep berbeda akibat perbedaan unit *cross section*.

$$Y_{it} = \beta_{0i} + \sum_{K=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \tag{2.15}$$

3. *Slope* koefisien konstan tetapi intersep berbeda akibat perbedaan unit *cross section* dan berubahnya *time series*.

$$Y_{it} = \beta_{0it} + \sum_{K=1}^{K} \beta_k X_{kit} + \varepsilon_{it} \tag{2.16}$$

4. Intersep dan *slope* koefisien berbeda akibat perbedaan unit *cross section.*

$$Y_{it} = \beta_{0i} + \sum_{K=1}^{K} \beta_{ki} X_{kit} + \varepsilon_{it} \tag{2.17}$$

5. Intersep dan *slope* koefisien berbeda akibat perbedaan unit *cross section* dan berubahnya *time series.*

$$Y_{it} = \beta_{0it} + \sum_{K=1}^{K} \beta_{kit} X_{kit} + \varepsilon_{it} \tag{2.18}$$

Keterangan:

i = 1, 2, ..., N

t = 1, 2, ..., T

N = banyaknya unit *cross section*

T = banyaknya unit *time series*

Y_{it} = nilai variabel dependen untuk unit *cross section* ke-i dan *time series* ke-t

X_{kit} = nilai variabel independen ke-k untuk unit *cross section* ke-i dan *time series* ke-t

β_{kit} = parameter yang diduga atau ditaksir

ε_{it} = nilai error untuk unit *cross section* ke-i dan *time series* ke-t

k = banyaknya parameter dalam regresi yang akan diduga atau ditaksir

2.3 Estimasi parameter Model Regresi Data Panel

Untuk mengestimasi parameter model regresi data panel terdapat 3 teknik yang dapat dilakukan yaitu, *common effect model*, *fixed effect model*, dan *random effect model*.

2.3.1 *Common Effect Model* (CEM)

Teknik yang paling sederhana untuk mengestimasi model regresi data panel adalah *common effect model* yaitu dengan menggabungkan data *time series* dan data *cross section* tanpa memperhatikan dimensi *cross section* maupun *time series*. Dalam metode ini, diasumsikan bahwa intersep setiap variabel adalah sama, dan setiap unit *cross section* dan *time series* memiliki *slope* yang konstan (Gujarati & Porter, 2012). *Common effect model* dapat ditulis pada persamaan (2.19).

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \varepsilon_{it} \tag{2.19}$$

keterangan:

Y_{it} = nilai variabel dependen untuk unit *cross section* ke-i dan *time series* ke-t

X_{it} = nilai variabel independen untuk unit *cross section* ke-i dan *time series* ke-t

β_0 = intersep model regresi

β_i = koefisien *slope*; $i = 1,2, \dots, k$

ε_{it} = nilai *error* untuk unit *cross section* ke-i dan *time series* ke-t

Apabila *common effect model* apabila memiliki *n* persamaan dan *t* observasi maka model dapat ditulis sebagai berikut.

- $i = 1$ maka

  $Y_{1t} = \beta_0 + \beta_1 X_{11t} + \beta_2 X_{21t} + \cdots + \beta_k X_{k1t} + \varepsilon_{1t}$
- $i = 2$ maka

  $Y_{2t} = \beta_0 + \beta_1 X_{12t} + \beta_2 X_{22t} + \cdots + \beta_k X_{k2t} + \varepsilon_{2t}$
- $i = 3$ maka

  $Y_{3t} = \beta_0 + \beta_1 X_{13t} + \beta_2 X_{23t} + \cdots + \beta_k X_{k3t} + \varepsilon_{3t}$

$$\vdots$$

dst.

- $t = 1$ maka

  $Y_{11} = \beta_0 + \beta_1 X_{111} + \beta_2 X_{211} + \cdots + \beta_k X_{k11} + \varepsilon_{11}$
- $t = 2$ maka

  $Y_{12} = \beta_0 + \beta_1 X_{112} + \beta_2 X_{212} + \cdots + \beta_k X_{k12} + \varepsilon_{12}$
- $t = 3$ maka

  $Y_{13} = \beta_0 + \beta_1 X_{113} + \beta_2 X_{213} + \cdots + \beta_k X_{k13} + \varepsilon_{13}$

$$\vdots$$

dst.

Maka diperoleh persamaan sebagai berikut:

$$Y_{1t} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{j1t} + \varepsilon_{1t}$$

$$Y_{2t} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{j2t} + \varepsilon_{2t}$$

$$Y_{3t} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{j3t} + \varepsilon_{3t}$$

$$\vdots$$

$$Y_{nt} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{jnt} + \varepsilon_{nt} \qquad (2.20)$$

Pada persamaan (2.20) dapat dituliskan dalam bentuk matriks sebagai berikut:

- Untuk $i = 1$ dan $t = 1, 2, 3, \dots, t$.

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1t} \end{bmatrix} = \begin{bmatrix} 1 & X_{111} & \cdots & X_{k11} \\ 1 & X_{112} & \cdots & X_{k12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{11t} & \cdots & X_{k1t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1t} \end{bmatrix}$$

- Untuk $i = 2$ dan $t = 1, 2, 3, \dots, t$.

$$\begin{bmatrix} Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & X_{121} & \cdots & X_{k21} \\ 1 & X_{122} & \cdots & X_{k22} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{12t} & \cdots & X_{k2t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2t} \end{bmatrix}$$

- Untuk $i = 3$ dan $t = 1, 2, 3, \dots, t$.

$$\begin{bmatrix} Y_{31} \\ Y_{32} \\ \vdots \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} 1 & X_{131} & \cdots & X_{k31} \\ 1 & X_{132} & \cdots & X_{k32} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{13t} & \cdots & X_{k3t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{31} \\ \varepsilon_{32} \\ \vdots \\ \varepsilon_{3t} \end{bmatrix}$$

- Untuk $i = n$ dan $t = 1, 2, 3, \dots, t$.

$$\begin{bmatrix} Y_{n1} \\ Y_{n2} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} 1 & X_{1n1} & \cdots & X_{kn1} \\ 1 & X_{1n2} & \cdots & X_{kn2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1nt} & \cdots & X_{knt} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{n1} \\ \varepsilon_{n2} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Secara keseluruhan matriks pada observasi *nt* dapat ditulis sebagai berikut:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} 1 & X_{111} & \cdots & X_{k11} \\ 1 & X_{112} & \cdots & X_{k12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1nt} & \cdots & X_{knt} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Metode estimasi parameter yang digunakan pada model *common effect* pada dasarnya sama dengan metode regresi linear biasa yaitu metode *ordinary least square* atau yang dikenal sebagai metode kuadrat terkecil. Prinsip dasar dari metode kuadrat terkecil yaitu:

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X\beta} \qquad (2.21)$$

Dalam mengestimasi parameter menggunakan metode *ordinary least square* adalah dengan meminumkan jumlah kuadrat residual (Gujarati & Porter, 2012). Maka didapat jumlah kuadrat *error* sebagai berikut:

$$\begin{aligned} SSE &= \sum_{i=1}^{n} \boldsymbol{\varepsilon_i}^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \\ &= (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta}) \\ &= (\boldsymbol{Y}' - \boldsymbol{\beta}'\boldsymbol{X}')(\boldsymbol{Y} - \boldsymbol{X\beta}) \\ &= \boldsymbol{Y'Y} - \boldsymbol{\beta'X'Y} - \boldsymbol{Y'X\beta} + \boldsymbol{\beta'X'X\beta} \\ &= \boldsymbol{Y'Y} - \boldsymbol{2\beta'X'Y} + \boldsymbol{X'\beta'X\beta} \end{aligned} \tag{2.22}$$

Apabila matriks $(\boldsymbol{X\beta})' = \boldsymbol{\beta'X'}$, maka skalar $\boldsymbol{\beta'X'Y} = \boldsymbol{Y'X\beta}$. Apabila ingin mendapatkan penduga untuk parameter $\boldsymbol{\beta}$ maka hasil turunan harus disamakan dengan 0 sehingga didapat persamaan (2.23) sebagai berikut:

$$\begin{aligned} \frac{\partial(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon})}{\partial\boldsymbol{\beta}} &= \boldsymbol{0} \\ \frac{\partial(\boldsymbol{Y'Y} - \boldsymbol{2\beta'X'Y} + \boldsymbol{\beta'X'X\beta})}{\partial\boldsymbol{\beta}} &= \boldsymbol{0} \\ -\boldsymbol{2X'Y} + \boldsymbol{2X'X\hat{\beta}} &= \boldsymbol{0} \\ \boldsymbol{2X'X\hat{\beta}} &= \boldsymbol{2X'Y} \\ \boldsymbol{X'X\hat{\beta}} &= \boldsymbol{X'Y} \\ (\boldsymbol{X'X})^{-1}(\boldsymbol{X'X})\boldsymbol{\hat{\beta}} &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \\ \boldsymbol{I\hat{\beta}} &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \\ \boldsymbol{\hat{\beta}} &= (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \end{aligned} \tag{2.23}$$

Apabila diasumsikan β_0 dan β_1 sama atau konstan untuk setiap data *time series* dan *cross section*, maka β_0 dan $\beta_1$dapat diestimasi dengan metode *ordinary least square*. Sehingga model *common effect* dengan $n \times t$ dapat ditulis pada persamaan (2.24).

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \varepsilon_{it}; i = 1,2,\dots,n;\ \ t = 1,2,\dots,T \tag{2.24}$$

2.3.2. *Fixed Effect Model* (FEM)

Fixed effect model merupakan model regresi data panel dengan asumsi adanya perbedaan *cross section* sehingga setiap unit *cross section* akan mempunyai intersep yang berbeda-beda sedangkan koefisien *slope* (kemiringan) konstan (Gujarati & Porter, 2012). *Fixed effect model* dapat ditulis pada persamaan (2.25) yaitu:

$$Y_{it} = \beta_{0i} + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \varepsilon_{it}; \tag{2.25}$$

Keterangan:

Y_{it} = nilai variabel terikat (*dependent* variabel) *cross section* ke-i dan *time series* ke-t

X_{it} = nilai variabel bebas (*independent variabel*) untuk unit *cross section* ke-i dan *time series* ke-t

β_0 = intersep model regresi

β_i = koefisien *slope* ; $i = 1,2,\ldots,k$

ε_{it} = nilai *error* untuk unit *cross section* ke-i dan *time series* ke-t

Dalam mengestimasi parameter pada *fixed effect model* dapat dilakukan dengan teknik menambahkan variabel *dummy* yang biasa dikenal dengan *least square dummy variabel* atau sering disebut dengan LSDV.

Least square dummy variabel adalah metode yang digunakan untuk mengestimasi parameter regresi linear dengan menerapkan metode kuadrat terkecil pada model ini, variabel *dummy* digunakan sebagai salah satu variabel bebasnya atau variabel independen. Variabel *dummy* adalah variabel yang digunakan untuk mengkuantitatifkan variabel yang bersifat kualitatif. Variabel *dummy* hanya mempunyai 2 (dua) nilai yaitu 1 dan nilai 0, serta diberi simbol D. *Dummy* memiliki nilai 1 ($D = 1$) untuk salah satu kategori dan nol ($D = 0$) untuk kategori yang lain. *Fixed effect model* menggunakan variabel *dummy* guna menjelaskan adanya perbedaan intersep antar unit *cross section*. *Fixed effect model* dengan variabel *dummy* dapat ditulis pada persamaan (2.26).

$$Y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{i=2}^{N} \beta_{0i} D_{ki} + \varepsilon_{it} \tag{2.26}$$

Subskrip $0i$ yang terdapat pada konstanta β_{0i} menunjukan bahwa i adalah objeknya yang diartikan bahwa setiap objek memiliki konstanta yang berbeda. Variabel *dummy* D_{it} mengartikan bahwa 1 untuk objek pertama, sedangkan 0 untuk objek lainnya. Variabel *dummy* yang dibentuk sebanyak *N-1,* dengan β_0 sebagai intersep untuk unit *cross section* yang pertama (Singh & Sachdeva, 2020).

Apabila terdapat banyak *n* persamaan dan *t* observasi maka *fixed effect model* dapat ditulis sebagai berikut:

$$Y_{1t} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{k1t} + \beta_{01} D_{1t} + \varepsilon_{1t}$$

$$Y_{2t} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{k2t} + \beta_{02} D_{2t} + \varepsilon_{2t}$$

$$Y_{3t} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{k3t} + \beta_{03} D_{3t} + \varepsilon_{3t}$$

$$\vdots$$

$$Y_{nt} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{knt} + \beta_{0n} D_{nt} + \varepsilon_{nt}$$

Untuk $i = 1$ dan dan $t = 1,2,3, \dots, t$ sehingga *fixed effect model* pada persamaan (2.26) dapat ditulis sebagai berikut:

- Jika $t = 1$ maka

$$\begin{aligned} Y_{11} &= \beta_{01} D_{11} + \beta_1 X_{111} + \beta_2 X_{211} + \beta_3 X_{311} + \cdots + \beta_k X_{k11} + \varepsilon_{11} \\ &= \beta_{01}.1 + \beta_1 X_{111} + \beta_2 X_{211} + \beta_3 X_{311} + \cdots + \beta_k X_{k11} + \varepsilon_{11} \end{aligned}$$

- Jika $t = 2$ maka

$$\begin{aligned} Y_{12} &= \beta_{01} D_{12} + \beta_1 X_{112} + \beta_2 X_{212} + \beta_3 X_{312} + \cdots + \beta_k X_{k12} + \varepsilon_{12} \\ &= \beta_{01}.0 + \beta_1 X_{112} + \beta_2 X_{212} + \beta_3 X_{312} + \cdots + \beta_k X_{k12} + \varepsilon_{12} \end{aligned}$$

- Jika $t = 3$ maka

$$\begin{aligned} Y_{13} &= \beta_{01} D_{13} + \beta_1 X_{113} + \beta_2 X_{213} + \beta_3 X_{313} + \cdots + \beta_k X_{k13} + \varepsilon_{13} \\ &= \beta_{01}.0 + \beta_1 X_{113} + \beta_2 X_{213} + \beta_3 X_{313} + \cdots + \beta_k X_{k13} + \varepsilon_{13} \end{aligned}$$

$$\vdots$$

- Jika $t = t$ maka

$$\begin{aligned} Y_{1t} &= \beta_{01}D_{1t} + \beta_1 X_{11t} + \beta_2 X_{21t} + \beta_3 X_{31t} + \cdots + \beta_k X_{k1t} + \varepsilon_{1t} \\ &= \beta_{01}.0 + \beta_1 X_{11t} + \beta_2 X_{21t} + \beta_3 X_{31t} + \cdots + \beta_k X_{k1t} + \varepsilon_{1t} \end{aligned} \quad (2.27)$$

Jika persamaan diatas ditulis dalam bentuk matriks maka dapat dilihat pada persamaan (2.28).

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1t} \end{bmatrix} = \begin{bmatrix} 1 & X_{111} & \cdots & X_{k11} \\ 1 & X_{112} & \cdots & X_{k12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{11t} & \cdots & X_{k1t} \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1t} \end{bmatrix} \quad (2.28)$$

$i = 2$ dan dan $t = 1,2,3, \ldots, t$ sehingga *fixed effect model* pada persamaan (2.26) menggunakan langkah yang sama seperti persamaan (2.27) dapat ditulis dalam bentuk matriks pada persamaan (2.29).

$$\begin{bmatrix} Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & X_{121} & \cdots & X_{k21} \\ 1 & X_{122} & \cdots & X_{k22} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{12t} & \cdots & X_{k2t} \end{bmatrix} \begin{bmatrix} \beta_{02} \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2t} \end{bmatrix} \quad (2.29)$$

$i = 3$ dan $t = 1,2,3, \ldots, t$ sehingga *fixed effect model* pada persamaan (2.26) menggunakan langkah yang sama seperti persamaan (2.27) dapat ditulis dalam bentuk matriks pada persamaan (2.30).

$$\begin{bmatrix} Y_{31} \\ Y_{32} \\ \vdots \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} 1 & X_{131} & \cdots & X_{k31} \\ 1 & X_{132} & \cdots & X_{k32} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{13t} & \cdots & X_{k3t} \end{bmatrix} \begin{bmatrix} \beta_{03} \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{31} \\ \varepsilon_{32} \\ \vdots \\ \varepsilon_{3t} \end{bmatrix} \quad (2.30)$$

$i = n$ dan dan $t = 1,2,3, \ldots, t$ sehingga *fixed effect model* pada persamaan (2.27) menggunakan langkah yang sama seperti persamaan (2.27) dapat ditulis dalam bentuk matriks pada persamaan (2.31):

$$\begin{bmatrix} Y_{n1} \\ Y_{n2} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} 1 & X_{1n1} & \cdots & X_{kn1} \\ 1 & X_{1n2} & \cdots & X_{kn2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1nt} & \cdots & X_{knt} \end{bmatrix} \begin{bmatrix} \beta_{0n} \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_{n1} \\ \varepsilon_{n2} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix} \quad (2.31)$$

2.3.3 *Random Effect Model* (REM)

Random effect model merupakan model regresi data panel dengan asumsi bahwa setiap unit *cross section* memiliki intersep yang berbeda. Tujuan dari model *random effect* adalah untuk mengestimasi data panel dengan menggunakan variabel *error* yang memiliki hubungan antara variabel individu dan waktu (Winarno, 2007). Terdapat dua komponen yang berkontribusi pada pembentukan *error*, yaitu individu dan waktu, maka *error random effect* juga diuraikan menjadi *error* komponen waktu dan gabungan antara *error* komponen waktu dan individu (Nachrowi & Usman, 2006). Model ini digunakan untuk mengatasi masalah yang ditimbulkan oleh model *fixed effect* yang menggunakan variabel *dummy.* Model *random effect* dapat ditulis pada persamaan (2.32):

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + W_{it} \tag{2.32}$$

keterangan:

Y_{it} = nilai variabel dependen untuk unit *cross section* ke-i dan *time series* ke-t

X_{kit} = nilai variabel independen ke-k untuk unit *cross section* ke-i dan *time series* ke-t

β_0 = intersep unit *cross section*

β_k = koefisien *slope*; $i = 1,2, \dots, k$

W_{it} = komponen *error* untuk unit *cross section* ke-i dan *time series* ke-t

Terdapat beberapa asumsi yang berlaku pada REM antara lain:

$$\mu_i \sim N\left(0, \sigma_\mu^2\right)$$
$$\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$$
$$E(\mu_i, \varepsilon_{it}) = 0$$
$$\mathrm{E}\left(\mu_i, \mu_j\right) = 0;\ i \neq j$$
$$E(\varepsilon_{it}, \varepsilon_{is}) = E\left(\varepsilon_{it}, \varepsilon_{jt}\right) = E\left(\varepsilon_{it}, \varepsilon_{js}\right) = 0; i \neq j; t \neq s \tag{2.33}$$

dengan $W_{it} = \mu_i + \varepsilon_{it}$ dimana μ_i merupakan komponen *error* untuk *cross section* dan ε_{it} merupakan kombinasi antara komponen *error cross section* dan *time series.* Terdapat beberapa hal terkait output estimasi *random effect* yaitu, penjumlahan dari nilai *random effect* adalah 0, karena komponen error W_{it} merupakan kombinasi antara *time series error* dan *cross section error*. Model yang sesuai dalam mengestimasi *random effect model* adalah *generalized least square.*

Jika $i = 1$ maka $Y_{1t} = \beta_0 + \beta_1 X_{11t} + \beta_2 X_{21t} + \cdots + \beta_k X_{k1t} + \mu_1 + \varepsilon_{1t}$

$i = 2$ maka $Y_{2t} = \beta_0 + \beta_1 X_{12t} + \beta_2 X_{22t} + \cdots + \beta_k X_{k2t} + \mu_2 + \varepsilon_{2t}$

$i = 3$ maka $Y_{3t} = \beta_0 + \beta_1 X_{13t} + \beta_2 X_{23t} + \cdots + \beta_k X_{k3t} + \mu_3 + \varepsilon_{3t}$

$\vdots$

dst

Jika $t = 1$ maka $Y_{11} = \beta_0 + \beta_1 X_{111} + \beta_2 X_{211} + \cdots + \beta_k X_{k11} + \mu_1 + \varepsilon_{11}$

$t = 2$ maka $Y_{12} = \beta_0 + \beta_1 X_{112} + \beta_2 X_{212} + \cdots + \beta_k X_{k12} + \mu_1 + \varepsilon_{12}$

$t = 3$ maka $Y_{13} = \beta_0 + \beta_1 X_{113} + \beta_2 X_{213} + \cdots + \beta_k X_{k13} + \mu_1 + \varepsilon_{13}$

$\vdots$

dst

Kemudian diperoleh persamaan dibawah ini:

$$Y_{1t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k1t} + \mu_1 + \varepsilon_{1t}$$

$$Y_{2t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k2t} + \mu_2 + \varepsilon_{2t}$$

$$Y_{3t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k3t} + \mu_3 + \varepsilon_{3t}$$

$$\vdots$$

$$Y_{nt} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{knt} + \mu_n + \varepsilon_{nt}$$

Jika $i = 1$ dan $t = 1,2,3, \ldots, t$, maka model *random effect* pada persamaan (2.33) dapat ditulis sebagai berikut:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1t} \end{bmatrix} = \begin{bmatrix} 1 & X_{111} & \cdots & X_{k11} \\ 1 & X_{112} & \cdots & X_{k12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{11t} & \cdots & X_{k1t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1t} \end{bmatrix} \tag{2.34}$$

Sehingga untuk *cross section* ke-1 dapat dituliskan pada persamaan berikut:

$$Y_{1t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k1t} + \mu_1 + \varepsilon_{1t} \tag{2.35}$$

Jika $i = 2$ dan $t = 1,2,3, \dots, t$, maka model *random effect* pada persamaan (2.33) dapat ditulis sebagai berikut:

$$\begin{bmatrix} Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & X_{121} & \cdots & X_{k21} \\ 1 & X_{122} & \cdots & X_{k22} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{12t} & \cdots & X_{k2t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2t} \end{bmatrix} \tag{2.36}$$

Sehingga untuk *cross section* ke-2 dapat dituliskan pada persamaan berikut:

$$Y_{2t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k2t} + \mu_2 + \varepsilon_{2t} \tag{2.37}$$

Jika $i = 3$ dan $t = 1,2,3, \dots, t$, maka model *random effect* pada persamaan (2.33) dapat ditulis pada persamaan 2.38.

$$\begin{bmatrix} Y_{31} \\ Y_{32} \\ \vdots \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} 1 & X_{131} & \cdots & X_{k31} \\ 1 & X_{132} & \cdots & X_{k32} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{13t} & \cdots & X_{k3t} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_{31} \\ \mu_{32} \\ \vdots \\ \mu_{3t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{31} \\ \varepsilon_{32} \\ \vdots \\ \varepsilon_{3t} \end{bmatrix} \tag{2.38}$$

Sehingga untuk *cross section* ke-3 dapat dituliskan pada persamaan berikut:

$$Y_{3t} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{k3t} + \mu_3 + \varepsilon_{3t} \tag{2.39}$$

Jika $i = n$ dan $t = 1,2,3, \dots, t$, maka model *random effect* pada persamaan (2.33) dapat ditulis sebagai berikut:

$$\begin{bmatrix} Y_{n1} \\ Y_{n2} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} 1 & X_{1n1} & \cdots & X_{kn1} \\ 1 & X_{1n2} & \cdots & X_{kn2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1nt} & \cdots & X_{knt} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_{n1} \\ \mu_{n2} \\ \vdots \\ \mu_{nt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{n1} \\ \varepsilon_{n2} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix} \tag{2.40}$$

Sehingga untuk cross section ke- n dapat dituliskan pada persamaan berikut:

$$Y_{nt} = \beta_0 + \sum_{k=1}^{k} \beta_k X_{knt} + \mu_n + \varepsilon_{nt} \tag{2.41}$$

Secara keseluruhan persamaan diatas dapat ditulis dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} 1 & X_{111} & \cdots & X_{k11} \\ 1 & X_{112} & \cdots & X_{k12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1nt} & \cdots & X_{knt} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{nt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix} \tag{2.42}$$

Estimasi parameter untuk model *random effect* dilakukan dengan metode *generalized least square*. Menurut Greene (1997), *generalized least square* dapat digunakan untuk mengatasi heterokedastisitas dan autokorelasi pada data *cross-section.* Menurut Gujarati & Porter (2012), pada data panel metode GLS lebih baik dan konsisten dari pada OLS. Penentuan estimasi parameter $\boldsymbol{\beta}$ pada model *random effect* model dilakukan dengan mencari jumlah kuadrat error. GLS menggunakan informasi secara eksplisit dan mampu menghasilkan *best linier unbiased estimator* (BLUE).

Diketahui model regresi linear :

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.43}$$

dalam GLS diasumsikan bahwa *error* memiliki varians-kovarians yang tidak identik dan heterokedastisitas atau ada autokorelasi.

$$E(\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}) = \boldsymbol{S} \tag{2.44}$$

Dengan $\boldsymbol{S}$ merupakan matriks varians kovarians yang mencakup heteroskedastisitas dan autokorelasi. $\boldsymbol{S}$ adalah matriks simetris dan *positive definite* maka ada matriks $\boldsymbol{p}$ simetrik non singular berukuran $n \times n$ dengan $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{PP} = \boldsymbol{S}$.

Untuk mengatasi heteroskedastisitas dan autokorelasi di lakukan transformasi pada model regresi linear, maka model di transformasi menjadi

$$\boldsymbol{P}^{-1}\boldsymbol{Y} = \boldsymbol{P}^{-1}\boldsymbol{X\beta} + \boldsymbol{P}^{-1}\boldsymbol{\varepsilon} \tag{2.45}$$

dengan $\boldsymbol{Y}^* = \boldsymbol{P}^{-1}\boldsymbol{Y}$, $\boldsymbol{X}^* = \boldsymbol{P}^{-1}\boldsymbol{X}$, $\boldsymbol{\varepsilon}^* = \boldsymbol{P}^{-1}\boldsymbol{\varepsilon}$.

Model yang telah di transformasi kemudian di estimasi menggunakan metode OLS dengan cara meminimumkan jumlah kuadrat *error.*

$$\begin{aligned}\boldsymbol{SSE} = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} &= (\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})'(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta}) \\ &= (\boldsymbol{Y}^{*\prime} - \boldsymbol{\beta}'\boldsymbol{X}^{*\prime})(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta}) \\ &= \boldsymbol{Y}^{*\prime}\boldsymbol{Y}^* - \boldsymbol{Y}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta} - \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{Y}^* + \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta}\end{aligned} \tag{2.46}$$

Karena $\boldsymbol{Y}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta}$ merupakan skalar, maka matriks *transpose* nya adalah $(\boldsymbol{Y}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta})' = \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{Y}^*$. Sehingga

$$\boldsymbol{SSE} = \boldsymbol{Y}^{*\prime}\boldsymbol{Y}^* - \boldsymbol{2\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{Y}^* + \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta} \tag{2.47}$$

SSE minimum diperoleh jika turunannya bernilai 0. Oleh karena itu, persamaan (2.47) diturunkan terhadap $\boldsymbol{\beta}$ menjadi persamaan (2.48).

$$\frac{\partial \boldsymbol{SSE}}{\partial \boldsymbol{\beta}} = \mathbf{0}$$

$$\frac{\partial(\boldsymbol{Y}^{*\prime}\boldsymbol{Y}^{*} - \mathbf{2}\boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*} + \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \mathbf{0}$$

$$-\mathbf{2}\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*} + \boldsymbol{X}^{*\prime}\boldsymbol{X}^{*}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*\prime} = \mathbf{0}$$

$$-\mathbf{2}\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*} + \mathbf{2}\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*}\boldsymbol{\beta} = \mathbf{0}$$

$$\mathbf{2}\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*}\boldsymbol{\beta} = \mathbf{2}\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*}$$

$$\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*}\boldsymbol{\beta} = \boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*}$$

$$(\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})^{-\mathbf{1}}(\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})\boldsymbol{\beta} = (\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})^{-\mathbf{1}}(\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*})$$

$$\boldsymbol{\beta} = (\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})^{-\mathbf{1}}(\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*}) \qquad (2.48)$$

sehingga untuk estimasi parameter dengan GLS dinyatakan sebagai berikut

$$\begin{aligned}\boldsymbol{\beta}_{\boldsymbol{GLS}} &= (\boldsymbol{X}^{*\prime}\boldsymbol{X}^{*})^{-\mathbf{1}}(\boldsymbol{X}^{*\prime}\boldsymbol{Y}^{*})\\ &= [(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{X})'(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{X})]^{-\mathbf{1}}(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{X})'(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{Y})\\ &= [\boldsymbol{X}'(\boldsymbol{P}^{-\mathbf{1}})'(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{X})]^{-\mathbf{1}}[\boldsymbol{X}'(\boldsymbol{P}^{-\mathbf{1}})'(\boldsymbol{P}^{-\mathbf{1}}\boldsymbol{Y})]\\ &= (\boldsymbol{X}'(\boldsymbol{P}'\boldsymbol{P})^{-\mathbf{1}}\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'(\boldsymbol{P}'\boldsymbol{P})^{-\mathbf{1}}\boldsymbol{Y}\\ &= (\boldsymbol{X}'\mathbf{S}^{-\mathbf{1}}\boldsymbol{X})^{-\mathbf{1}}(\boldsymbol{X}'\mathbf{S}^{-\mathbf{1}}\boldsymbol{Y}) \qquad (2.49)\end{aligned}$$

2.4 Pemilihan Model Regresi Data Panel

Tiga model dalam analisis regresi data panel yaitu *common effect model, fixed effect model*, dan *random effect model*. Untuk menentukan model terbaik dari analisis regresi data panel terdapat beberapa uji yakni menggunakan uji Chow, uji Hausman dan uji *Lagrange multiplier*.

2.4.1 Uji Chow

Uji Chow pertama kali diperkenalkan oleh Gregory Chow, seorang ekonom dan statistikawan asal amerika yang dikembangkan pada tahun 1960-an. Uji Chow digunakan untuk memilih model terbaik dalam mengestimasi model regresi data panel antara *common effect model* dengan *fixed effect model* (Wibawa, dkk.*,* 2022). Adapun langkah pengujiannya adalah sebagai berikut:

1. Hipotesis

   H_0 : model yang sesuai digunakan dalam regresi data panel adalah *common effect model*

   H_1 : model yang sesuai digunakan dalam regresi data panel adalah *fixed effect model*

2. Statistik uji yang digunakan adalah uji F, yaitu :

   $$F_{hitung} = \frac{(RSS_1 - RSS_2)/(n-1)}{RSS_2/(nt-n-k)} \qquad (2.50)$$

   Sedangkan F nya diperoleh dari:

   $$F_{tabel}; \{\alpha: df(n-1, nt-n-k)\} \qquad (2.51)$$

   keterangan:

   RSS_1 = *residual sum of square* model *common effect*

   RSS_2 = *residual sum of square* model *fixed effect*

   t = jumlah unit *time series*

   n = jumlah unit *cross section*

   k = jumlah variabel independen (variabel bebas)

   α = tingkat signifikansi yang digunakan

3. Taraf signifikansi yang digunakan adalah $\alpha = 0.05$.

4. Daerah kritis

   Terima H_0, jika $F_{hitung} < F_{tabel}$ atau jika $p - value > \alpha$

   Tolak H_0, jika $F_{hitung} > F_{tabel}$ atau jika $p - value < \alpha$

5. Kesimpulan

   Jika H_0 ditolak maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *fixed effect model*. jika H_0 diterima maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *common effect model.*

2.4.2 Uji Hausman

Uji Hausman pertama kali diperkenalkan oleh ekonometris Jerry A. Hausman pada tahun 1978. Uji Hausman digunakan untuk memilih model terbaik dalam mengestimasi model regresi data panel antara *fixed effect model* dengan *random effect model* (Gujarati & Porter, 2012). Adapun langkah pengujiannya sebagai berikut:

1. Hipotesis

   H_0 : model yang sesuai digunakan dalam regresi data panel adalah *random effect model*

   H_1 : model yang sesuai digunakan dalam regresi data panel adalah *fixed effect model*

2. Statistik uji yang digunakan adalah uji *chi-squared* berdasarkan kriteria wald,

$$\boldsymbol{W} = (\widehat{\boldsymbol{\beta}}_{MFE} - \widehat{\boldsymbol{\beta}}_{MRE})'[\boldsymbol{var}(\widehat{\boldsymbol{\beta}}_{MFE} - \widehat{\boldsymbol{\beta}}_{MRE})]^{-1}(\widehat{\boldsymbol{\beta}}_{MFE} - \widehat{\boldsymbol{\beta}}_{MRE}) \quad (2.52)$$

   keterangan:

   $\widehat{\boldsymbol{\beta}}_{MFE}$ = vektor estimasi *slope fixed effect model*

   $\widehat{\boldsymbol{\beta}}_{MRE}$ = vektor estimasi *slope random effect model*

3. Taraf signifikansi yang digunakan adalah $\alpha = 0.05$.

4. Daerah kritis

   Terima H_0, jika nilai $W < X^2{}_{(\alpha,k)}$ atau jika $p - value > \alpha$

   Tolak H_0, jika nilai $W > X^2{}_{(\alpha,k)}$ atau jika $p - value < \alpha$

5. Kesimpulan

   Jika H_0 ditolak maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *fixed effect model*, tetapi jika

H_0 diterima maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *random effect model.*

2.4.3 Uji *Lagrange Multiplier*

Uji *lagrange multiplier* merupakan uji yang digunakan untuk memilih model terbaik dalam mengestimasi model regresi data panel antara *common effect model* dengan *random effect model* (Widarjono, 2005). Adapun langkah pengujiannya sebagai berikut:

1. Hipotesis

   H_0 : model yang sesuai digunakan dalam regresi data panel adalah *common effect model*

   H_1 : model yang sesuai digunakan dalam regresi data panel adalah *random effect model*

2. Statistik uji yang digunakan adalah uji *lagrange multiplier* :

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\hat{\varepsilon}_{it}\right)^2}{\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2} - 1\right]^2 \quad (2.53)$$

   keterangan:

   n = jumlah unit *cross section*

   T = jumlah unit *time series*

   $\hat{\varepsilon}_{it}$= estimasi residual model *common effect model* dengan *cross section* ke-i dan periode ke-t

3. Taraf signifikansi yang digunakan adalah $\alpha = 0.05$

4. Daerah kritis

   Terima H_0, jika nilai $LM < X^2{}_{(\alpha,k)}$ atau jika $p - value > \alpha$

   Tolak H_0, jika nilai $LM > X^2{}_{(\alpha,k)}$atau jika $p - value < \alpha$

5. Kesimpulan

   Jika H_0 ditolak maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *random effect model*, tetapi jika H_0 diterima maka sudah cukup bukti untuk menyatakan bahwa model yang sesuai digunakan dalam regresi data panel adalah *common effect model*.

2.5 Uji Asumsi Klasik

Model analisis regresi data panel dikatakan baik apabila dapat memenuhi beberapa uji asumsi klasik yang terdiri dari uji normalitas, uji multikolinearitas, uji heterokesdastisitas dan uji autokorelasi. Akan tetapi pada *random effect model* hanya perlu dilakukan uji normalitas dan multikolineritas karena model tersebut menggunakan estimasi GLS, estimasi GLS tersebut sudah dapat mengatasi heterokedastisitas dan autokorelasi (Halim, dkk., 2020)

2.5.1 Uji Normalitas

Uji normalitas digunakan untuk menguji apakah nilai residual berdistribusi normal atau tidak berdistribusi normal. Model yang residualnya berdistribusi normal merupakan model yang baik (Draper & Smith, 1992). Pengujian normalitas data dapat dilakukan baik secara visual ataupun menggunakan analisis statistik. Secara visual, pengujian dapat dilakukan dengan menampilkan histogram dan *boxplot*. Sedangkan untuk analisis statistik dapat digunakan dengan menggunakan uji *Jarque Bera* (JB). Uji *Jarqeu Bera* ini menggunakan perhitungan *skewness* dan kurtosis. Adapun langkah pengujiannya sebagai berikut:

1. Hipotesis

   H_0: residual berdistribusi normal

   H_1: residual tidak berdistribusi normal
2. Statistik uji yang digunakan adalah uji *Jarque Bera* yaitu :

$$JB = N\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \tag{2.54}$$

$$K = \frac{\hat{\mu}_4}{\hat{\mu}_2^2} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}^2)\right]^2} \tag{2.55}$$

$$S = \frac{\hat{\mu}_3}{\hat{\mu}_2^{3/2}} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^3}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}^2)\right]^{3/2}} \tag{2.56}$$

   keterangan:

   N = banyaknya data

S = *skewness*

K = kurtosis

3. Taraf signifikansi yang digunakan adalah $\alpha = 0{,}05$
4. Daerah kritis

   Terima H_0, jika nilai $JB <$ dari $X^2{}_{(\alpha,2)}$ atau $p - value > \alpha$

   tolak H_0, jika nilai $JB >$ dari $X^2{}_{(\alpha,2)}$ atau $p - value < \alpha$
5. Kesimpulan

   Jika H_0 ditolak maka sudah cukup bukti untuk menyatakan bahwa residual tidak berdistribusi normal, tetapi jika H_0 diterima maka sudah cukup bukti untuk menyatakan bahwa residual berdistribusi normal.

2.5.2 Uji Multikolinearitas

Uji multikolinearitas bertujuan untuk menguji apakah ditemukan adanya korelasi (hubungan yang kuat) antar variabel independen pada suatu model regresi. Jika pada model regresi memiliki lebih dari satu variabel independen maka perlu dilakukan uji multikolinearitas. Model regresi yang baik adalah model yang tidak mempunyai korelasi yang kuat antar variabel independennya. Adanya korelasi yang kuat antar variabel independen sangatlah tidak disarankan, karena akan mempengaruhi keakuratan estimasi parameter dalam estimasi nilai yang sebenarnya. Terdapat cara untuk mendeteksi ada atau tidaknya multikolinearitas yaitu:

1. Menggunakan nilai *Variance Inflation Factor* (VIF). Apabila nilai $VIF \geq 10$ maka terdapat multikolinearitas. Sebaliknya apabila nilai $VIF < 10$ maka tidak terdapat multikolinearitas (Montgomery & Peck, 1992). Rumus VIF yang digunakan adalah uji multikolinearitas pada persamaan (2.57).

$$VIF = \frac{1}{1-R_j{}^2} \tag{2.57}$$

dengan $R_j{}^2$ merupakan koefisien determinasi yang diperoleh dari regresi variabel bebas X_j terhadap variabel bebas lainnya. Apabila variabel X_j tidak berkorelasi dengan variabel bebas lainnya maka nilai ${R_j}^2$ akan lebih kecil dan nilai VIF mendekati 1. Sebaliknya apabila variabel X_j berkorelasi dengan variabel bebas lainnya maka nilai ${R_j}^2$ akan mendekati 1 dan nilai VIF akan besar (Montgomery & Peck, 1992).

2. Multikolinearitas dapat terdeteksi dengan menggunakan matriks korelasi variabel bebas. Jika terdapat korelasi antar variabel bebasnya dengan nilai lebih dari 0,8 hal ini mengindikasikan bahwa adanya multikolinearitas (Ghozali, 2016).

2.5.3 Uji Heterokesdastisitas

Uji heterokedastisitas bertujuan untuk melihat apakah residual dari model yang terbentuk memiliki varians yang konstan atau tidak (Kuncoro, 2011). Model regresi dikatakan baik apabila model tersebut memiliki varians residual yang konstan. Adanya heterokedastisitas dapat membuat estimasi parameter menjadi tidak efisien (Nachrowi & Usman, 2006). Pengujian adanya heterokedastisitas dalam suatu model regresi dilakukan dengan menggunakan uji *Breusch Pagan* (BP).

Adapun Hipotesis yang digunakan pada uji *Breusch Pagan* sebagai berikut:
H_0 : tidak terdapat heterokedastisitas pada data
H_1 : terdapat heterokedastisitas pada data

Kriteria untuk menolak hipotesis ini adalah jika nilai uji BP < $x^2_{(\alpha,k)}$ atau apabila nilai $p - value$ > α maka H_0 tidak ditolak yang berarti tidak terdapat heterokedastisitas pada data. Begitupun sebaliknya jika nilai uji BP > $x^2_{(\alpha,k)}$ atau apabila nilai $p - value$ < α maka H_0 ditolak. Maka terdapat heterokedastisitas pada data.

2.5.4 Uji Autokorelasi

Uji autokorelasi bertujuan untuk mendeteksi apakah dalam model regresi linear terdapat korelasi antara residual pada suatu pengamatan dengan pengamatan yang lain. Jika ada autokorelasi, asumsi bahwa residual dalam model regresi linear bersifat independen satu sama lain dilanggar. Hal ini dapat menyebabkan estimasi parameter yang tidak akurat dan kurang efisien (Singgih, 2002). Pengujian adanya autokorelasi dalam suatu model regresi dilakukan dengan menggunakan uji *Durbin Watson.*

Adapun hipotesis yang digunakan pada uji *Durbin Watson* sebagai berikut:
H_0 : data tidak terdapat autokorelasi
H_1 : data terdapat autokorelasi

Statistik uji yang digunakan adalah:

$$d = \frac{\sum_{t=1}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2} \quad (2.58)$$

keterangan:
d = nilai *Durbin Watson Statistic*
ε_t = nilai residual periode ke-t
ε_{t-1} = nilai residual periode ke $t-1$

Tabel 1. Kriteria pengujian autokorelasi

Kriteria	Keputusan	Keterangan
$0 < d < d_L$	H_0 ditolak	Terdapat autokorelasi positif
$d_L \leq d \leq d_U$	Tidak ada	Tidak terdapat autokorelasi positif
$4 - d_L \leq d \leq 4$	H_0 ditolak	Terdapat autokorelasi negatif
$4 - d_U \leq d \leq 4 - d_L$	Tidak ada	Tidak terdapat autokorelasi negatif
$d_U < d < 4 - d_U$	H_0 diterima	Tidak terdapat autokorelasi

keterangan:
d_L= nilai Durbin *Lower*
d_U = nilai Durbin *Upper*

Kriteria untuk menolak hipotesis ini adalah sebagai berikut apabila nilai $p - value$ < α maka H_0 ditolak. Maka sudah cukup bukti untuk menyatakan bahwa data terdapat autokorelasi. Begitupun sebaliknya apabila nilai $p - value$ > α maka H_0 diterima. Maka sudah cukup bukti untuk menyatakan bahwa data tidak terdapat autokorelasi.

2.6 Uji Kelayakan Model Regresi Data Panel

Terdapat dua statistik uji yang perlu dilakukan untuk menguji signifikansi koefisien regresi yaitu statistik uji simultan (uji F) dan uji parsial (uji T).

2.6.1 Uji Simultan (Uji F)

Uji F dilakukan untuk menunjukan apakah seluruh variabel independen yang dimasukan dalam model regresi mempunyai pengaruh terhadap variabel dependen (Ghozali, 2016). Uji F digunakan untuk menguji secara simultan *slope* dalam model regresi.

Adapun hipotesis yang digunakan pada uji F sebagai berikut:

H_0 : $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ (secara simultan variabel independen tidak memberikan pengaruh yang signifikan terhadap variabel dependen)

H_1: Minimal terdapat satu $\beta_j \neq 0$; $j = 1,2,3, \ldots, k$ (secara simultan variabel independen memberikan pengaruh signifikan terhadap variabel dependen)

Statistik Uji F yang digunakan adalah:

$$F_{hitung} = \frac{R^2/(N+k-1)}{(1-R^2)/(NT-N-k)} \tag{2.59}$$

keterangan:

N = banyaknya unit *cross section*

T = banyaknya unit *time series*

k = banyaknya variabel independen

Kriteria untuk menolak hipotesis ini adalah apabila nilai F_{hitung} > F_{tabel} atau apabila nilai $p - value$ < α maka H_0 ditolak. Maka sudah cukup bukti untuk menyatakan bahwa secara simultan variabel independen memberikan pengaruh signifikan terhadap variabel dependen. Begitupun sebaliknya apabila nilai F_{hitung} < F_{tabel} atau apabila nilai $p - value$ > α maka H_0 diterima. Maka sudah cukup bukti untuk menyatakan bahwa secara simultan variabel independen tidak memberikan pengaruh signifikan terhadap variabel dependen.

2.6.2 Uji Parsial (Uji T)

Uji T (uji parsial) biasanya digunakan untuk mengetahui apakah variabel independen mempunyai pengaruh signifikan secara individual atau parsial terhadap variabel dependen (Gujarati & Porter, 2012).

Adapun hipotesis yang digunakan pada uji t sebagai berikut:

H_0 : $\beta_j = 0$ (secara individual, variabel independen tidak memberikan pengaruh signifikan terhadap variabel dependen)

H_1 : $\beta_j \neq 0$, $j = 1,2,3, \dots, k$ (secara individual, variabel independen memberikan pengaruh signifikan terhadap variabel dependen)

Uji t didefinisikan sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \tag{2.60}$$

keterangan:

$\hat{\beta}_j$ = penduga parameter ke-j yang dohipotesiskan ($j = 1,2,3, \dots, k$)

$SE(\hat{\beta}_j)$= *standard error* $\hat{\beta}_j$

Kriteria untuk menolak hipotesis ini adalah sebagai berikut apabila nilai $|t_{hitung}| > t_{tabel}$ atau apabila nilai $p - value$ < α maka H_0 ditolak. Maka sudah cukup bukti untuk menyatakan bahwa secara individual variabel independen memberikan pengaruh signifikan terhadap variabel dependen. Begitupun sebaliknya apabila nilai $|t_{hitung}| < t_{tabel}$ atau apabila nilai $p - value$ > α maka H_0 diterima. Maka, secara individual variabel independen tidak memberikan pengaruh signifikan terhadap variabel dependen,

2.7 Koefisien Determinasi

Nilai koefisien determinasi digunakan untuk mengetahui besarnya pengaruh variabel independen (X) terhadap variabel dependen (Y) atau derajat variasi variabel dependen yang dapat dijelaskan oleh variabel independen. Nilai koefisien determinasi mendekati 1 (satu), berarti kemampuan variabel independen dalam menjelaskan variabel dependen cukup baik. Sebaliknya, jika koefisien determinasi mendekati 0 (nol), maka daya penjelas variabel independen cukup terbatas (Kuncoro, 2011).

$$R^2 = 1 - \frac{RSS}{TSS} \tag{2.61}$$

keterangan:

RSS = *residual sum of square*

TSS = *total sum of square*

2.8 Peramalan

Peramalan merupakan teknik yang digunakan untuk memprediksi nilai di masa depan dengan mempertimbangan data di masa lalu dan masa kini (Rizal, dkk., 2021). Model regresi data panel juga dapat digunakan untuk melakukan peramalan dependen. Akan tetapi, meramalkan variabel dependen untuk setiap wilayah dalam beberapa tahun mendatang hanya dapat dilakukan apabila nilai dari variabel independen untuk setiap wilayah selama tahun tersebut telah diketahui. Oleh karena itu, akan dilakukan peramalan varaiabel independen terlebih dahulu untuk beberapa tahun kedepan di setiap wilayahnya menggunakan analisis *trend* linear, *trend* kudratik dan *trend* eksponensial. Hasil peramalan yang terbaik adalah peramalan dengan memiliki nilai MAPE yang paling kecil. Setelah peramalan dilakukan maka hasil peramalan variabel prediktor di subsitusikan ke dalam persamaan regresi data panel sehingga diperoleh ramalan untuk variabel respon.

2.9 Analisis *Trend*

Analisis *trend* adalah jenis analisis yang digunakan untuk melakukan peramalan atau perkiraan pada masa yang akan datang. Untuk melakukan peramalan dibutuhkan berbagai macam informasi pada periode waktu yang relatif panjang (Rahmawati, 2015). Apabila dalam suatu deret terdapat gerakan naik ataupun turun dalam jangka panjang, maka deret tersebut dikatakan deret yang mengandung unsur *trend.* Pemilihan model *trend* yang paling sesuai dilakukan karena metode *trend* yang paling sesuai akan memberikan nilai dugaan yang lebih dekat dengan nilai aktualnya. Pemilihan model yang paling sesuai dapat dilakukan dengan membandingan nilai *error* dari masing-masing metode *trend* yang digunakan untuk menentukan teknik yang terbaik dalam peramalan (Murti, 2019). *Trend* adalah istilah yang digunakan untuk meramalkan variabel yang variabel bebasnya adalah waktu. Analisis *trend* dibagi menjadi 3 yaitu *trend* linear, *trend* kuadratik, dan *trend* eksponensial (Apriyanti, dkk*.,* 2021).

2.9.1 *Trend* Linear

Trend linear merujuk pada kecenderungan data di mana perubahan seiring waktu tetap stabil (konstan). Untuk mengamati *trend* linear jangka panjang, diperlukan periode yang mencakup setidaknya satu siklus.

Menurut Rachmawati, dkk. (2022), model untuk mencari persamaan *trend* linear mirip dengan model regresi, namun dalam hal ini menggunakan waktu (t) sebagai variabel independen. Sebagai contoh, jika kita ingin melihat pola *trend* jangka panjang dari variabel X_{it}, maka model untuk mengestimasikan persamaannya dapat dirumuskan pada persamaan (2.62).

$$X_{it} = a_i + b_i t \tag{2.62}$$

Keterangan:

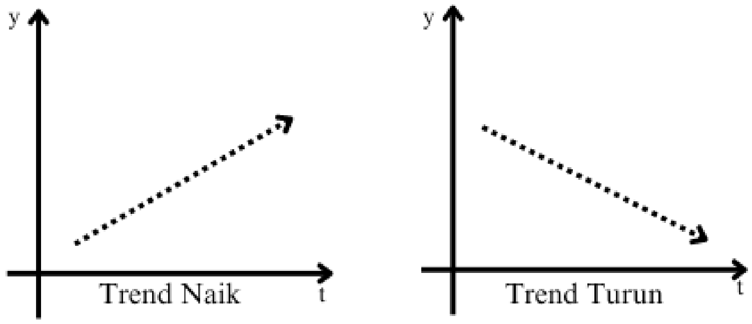
X_{it} = nilai *trend* pada periode tertentu, $t = 1, 2, \dots, n$

α, b = koefisien *trend*

t = waktu/periode

Nilai t untuk waktu diberi nilai 1, waktu berikutnya diberi nilai 2, dan seterusnya waktu terakhir diberi nilai n (Atmaja, 1997).

Trend linear akan berbentuk pola garis lurus seperti yang ditunjukan pada Gambar 1.



Gambar 1. Grafik *trend* linear.

2.9.2 *Trend* Kuadratik

Trend kuadratik merupakan ke kecenderungan data yang kurvanya berpola lengkungan atau parabola. Penggunaan *trend* kuadratik terjadi karena sering kali perkembangan nilai suatu peubah yang dalam jangka pendek atau menengahnya berpola linear menjadi tidak linear dalam jangka panjang. Konsekuensinya harus dibuat persamaan *trend* yang tidak linear (Rachmawati, dkk., 2022). Bentuk persamaan yang digunakan disajikan pada persamaan (2.63).

$$X_{it} = a_i + b_i t + c_i t^2 \qquad (2.63)$$

keterangan:

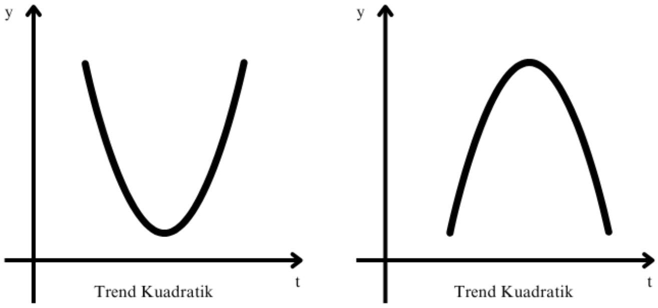
X_{it} = nilai *trend* pada periode tertentu

a, b = nilai *trend* periode dasar

c = nilai perubahan *trend* setiap periode

t = unit periode yang dihitung dari periode dasar

Secara grafik, *trend* kuadratik ditunjukan pada Gambar 2.



Gambar 2. Grafik *trend* kuadratik.

2.9.3 *Trend* Eksponensial

Trend kuadratik atau yang biasa di sebut dengan *trend* parabolik biasanya digunakan untuk menggambarkan tingkat pertambahan yang bertambah secara konstan. Secara matematis selisih kedua dari *trend* kuadratik menjadi positif atau konstan. Apabila *trend* tersebut digambarkan pada kertas berskala hitung maka rasio perubahan konstan tersebut sulit diketahui oleh karena itu, digunakan *trend* eksponensial guna mengetahui rasio perubahan konstan tersebut. *Trend* eksponensial adalah *trend* yang variabel bebasnya dapat naik berlipat ganda dan tidak linier (Rachmawati, dkk., 2022). Model untuk *trend* eksponensial yaitu:

$$X_{it} = a_i + {b_i}^t \tag{2.64}$$

dimana t adalah pangkat eksponen pada b. Logaritma natural pada persamaan diatas digunakan untuk mencari nilai a dan b maka menghasilkan:

$$\log X_{it} = \log a_i + t \log b_i \tag{2.65}$$

Perlu digunakan metode kuadrat terkecil untuk menentukan nilai a dan b. Nilai Y yang telah di transformasikan kedalam $\log Y$, $log\ a$ dan $log\ b$ dapat menghasilkan rumus persamaan sebagai berikut:

$$\log a_i = \frac{\sum \log X_{it}}{n} \tag{2.66}$$

$$\log b_i = \frac{\sum \mathrm{t} \log X_{it}}{\sum t^2} \tag{2.67}$$

dengan demikian nilai koefisien *trend* a dan b dapat diperoleh.

keterangan:

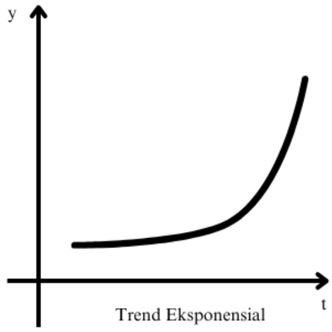
X_{it} = nilai *trend* pada periode tertentu

$a, b =$ nilai konstanta

t = waktu

n = banyaknya data

Secara grafis, *trend* eksponensial ditunjukan pada Gambar 3.



Gambar 3. Grafik *trend* eksponensial.

2.10 Kriteria Memilih *Trend*

Dalam memilih *trend* yang paling baik digunakan untuk analisis ada 2 cara yaitu:

1. Menganalisis grafik data atau *scatter-plot*. Jika data observasi cenderung menunjukkan gejala linear, kita sebaiknya menggunakan *trend* linear. Jika data observasi cenderung menunjukkan ciri-ciri bentuk kuadratik, gunakan *trend* kuadratik. Jika data observasi cenderung menunjukkan tidak linear dan tidak kuadratik, gunakan *trend* eksponensial.
2. Menghitung nilai *Mean Absolute Percentage error* (MAPE)
   MAPE (*Mean Absolute Percentage Error*) yaitu mengukur ketepatan nilai dugaan model yang dinyatakan dalam bentuk rata-rata persentase absolut kesalahan. Menghitung nilai MAPE untuk setiap jenis *trend*, pilih garis *trend* yang memberikan nilai MAPE terkecil. Semakin kecil nilai MAPE menyebabkan semakin baik/tepat model yang didapatkan. Nilai MAPE dapat dihitung menggunakan persamaan (2.68).

$$\text{MAPE}=\frac{\sum_{t=1}^{n}|Y_t-\hat{Y}_t|}{n}\times 100\% \qquad (2.68)$$

keterangan:

Y_t= data aktual pada periode ke-t

$\hat{Y}_t$= data ramalan pada periode ke-t

n = banyaknya periode waktu

Tabel 2. Kriteria nilai MAPE

No	MAPE	Intrepretasi
1	$< 10\%$	Sangat baik
2	$10\% - 20\%$	Baik
3	$20\% - 50\%$	Cukup baik
4	$> 50\%$	Buruk

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun akademik 2023/2024 di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang didapatkan dari *website* resmi Badan Pusat Statistika (BPS) pada link https://www.bps.go.id/. Data pada penelitian ini terdiri dari banyaknya tingkat kemiskinan (Y), pengeluaran perkapita disesuaikan (X_1), tingkat pengangguran terbuka (X_2), dan rata-rata lama sekolah (X_3) di provinsi yang berada di pulau Sumatera pada tahun 2015-2022 yang disajikan pada Lampiran 1.

3.3 Metode Penelitian

Pengolahan data pada penelitian ini dibantu dengan menggunakan *software R Studio* dan *Excel*. Langkah-langkah yang digunakan pada penelitian ini adalah sebagai berikut:

1. Melakukan analisis deskriptif pada data.

2. Menganalisis model regresi data panel dengan pendekatan *common effect model*, *fixed effect model* dan *random effect model*.
3. Melakukan uji pemilihan model terbaik yang terdiri dari
   a. Uji Chow dilakukan untuk pemilihan model antara *common effect model* dan *fixed effect model.*
   b. Uji Hausman dilakukan untuk pemilihan model antara *fixed effect model* dan *random effect model.* Apabila uji Hausman terpenuhi maka selanjutnya dilakukan uji asumsi klasik, tetapi apabila tidak terpenuhi maka dilanjutkan dengan uji *Lagrange Multiplier.*
4. Melakukan uji asumsi klasik yang terdiri:
   a. Uji normalitas menggunakan statistik uji Jarque Bera.
   b. Uji multikolinearitas menggunakan nilai VIF
5. Melakukan uji statistik model yang terdiri dari uji simultan (uji F) dan uji parsial (uji T), dan melakukan uji koefisien determinasi.
6. Menginterpretasikan hasil estimasi pada model regresi data panel.
7. Melakukan peramalan untuk variabel dependen berdasarkan pada model regresi data panel yang telah diperoleh untuk waktu mendatang.
   a. Melakukan peramalan untuk masing-masing variabel independen menggunakan analisis *trend* linear, kuadratik, dan eksponensial.
   b. Menentukan hasil peramalan terbaik pada masing-masing variabel dari ketiga metode tersebut berdasarkan nilai MAPE terkecil.
   c. Mensubstistusikan hasil peramalan variabel independen untuk masing-masing wilayah pada masing-masing model yang telah diperoleh untuk mendapatkan hasil peramalan dari variabel dependen di setiap wilayah.
8. Membuat kesimpulan.

V. KESIMPULAN

Berdasarkan dari hasil dan pembahasan pada penelitian ini, maka diperoleh kesimpulan sebagai berikut:

1. Model regresi data panel yang paling tepat atau yang paling baik digunakan untuk memodelkan tingkat kemiskinan di Pulau Sumatera dari tahun 2015-2022 adalah *random effect model* (REM) dengan metode *generalized least square*. Adapun model persamaan estimasinya adalah $\hat{Y}_{it} = 25.7186 - 0.0008X_{1it} + 0.1125X_{2it} - 0.7977X_{3it} \quad +W_{it}$ dengan diperolah nilai koefisien determinasinya nya sebesar 0.7338 yang artinya variabel pengeluaran perkapita disesuaikan, tingkat pengangguran terbuka, rata-rata lama sekolah hanya mampu menjelaskan variabel tingkat kemiskinan sebesar 73,38%, sedangkan sisanya sebesar 26,62% dijelaskan oleh variabel lain yang tidak digunakan pada model tersebut.
2. Metode terbaik yang digunakan untuk melakukan peramalan yaitu *trend* kuadratik untuk X_1, X_2 dan X_3 sehingga dapat diketahui peramalan tingkat kemiskinan mengalami naik turun pada masing-masing provinsi.

DAFTAR PUSTAKA

Alamsyah, F.A., Esra, R., Awalia, S., & Nohe, D.A. 2022. Analisis Regresi Data Panel untuk Mengetahui Faktor yang Memengaruhi Jumlah Penduduk Miskin di Kalimantan Timur, hlm. 254-266. *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*. Samarinda.

Alwi,W., Rayyan, I., & Nurfadilah. 2018. Analisis Regresi Data Panel pada Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan Provinsi Sulawesi Selatan Tahun 2011-2015. *Jurnal Matematika dan Statistik serta Aplikasinya*. **6**(2): 1-15.

Apriyanti, L., Setiadi, A., Santoso, S.I. 2021. Analisis Peramalan Volume Ekspor Melon di PT. Bumi Sari Lestari Temanggung Jawa Tengah. *Jurnal Ekonomi Pertanian dan Agrbisnis*. **5**(4): 1051-1058.

Atmaja, L. S.1997. *Memahami Statistika Bisnis.* Buku Kedua. Ed. 1. Andi Offset, Yogyakarta.

Badan Pusat Statistik. 2022. Data dan Informasi Kemiskinan Kabupaten/Kota di Indonesia.https://www.bps.go.id/id/publication/2022/11/30/3b084878f782dfa44e0025e0/data-dan-informasi-kemiskinan-kabupaten-kota-tahun-2022.html. Diakses pada 20 September 2023.

Baltagi, B.H. 2005. *Econometrics Analysis of Panel data. 3rd Edition*. Pearson Education, Inc., New Jersey.

Basuki, A.T. 2016. *Pengantar Ekonometrika (Dilengkapi Penggunaan Eviews)*. Edisi ke-1. Danisa Media, Sleman.

Draper, N.R & Smith, H. 1992. *Analisis Regresi Terapan*. Terjemahan Bambang Sumantri. Gramedia Pustaka Utama, Jakarta.

Ghozali, I. 2016. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 23*. Ed. ke-8. Badan Penerbit Universitas Diponegoro. Semarang.

GoodStats. 2023. Tingkat Kemiskinan di Asia Tenggara 2023. https://goodstats.id/infographic/tingkat-kemiskinan-di-asia-tenggara-2023-HBHy1#:~:text=Kementerian%20Keuangan%20mengungkapkan%2C%20secara%20akumulatif%2C%20sejak%20Maret%202021,negara%20dengan%20angka%20kemiskinan%20tertinggi%20di%20Asia%20Tenggara. Diakses pada 9 Desember 2023.

GoodStats. 2023. Jumlah Penduduk Miskin di Pulau Besar Indonesia per Maret 2023. https://data.goodstats.id/statistic/sarahjauhari/jumlah-penduduk-miskin-di-pulau-besar-indonesia-per-maret-2023-XHweA. Diakses pada 9 Desember 2023.

Greene, W.H. 1997. *Economic Analysis*. Prentice-Hall International, Lnc. USA.

Gujarati, D. & Porter, D. C. 2012. *Basic Econometrics*. Tata McGraw-Hill Education. New York.

Halim, I. T., Ramadhanty, A. & Oscarini, D.R. 2020. Faktor-Faktor yang Mempengaruhi Presentase Penduduk Miskin di Indonesia Tahun 2015-2018 Menggunakan RegresI Data Panel. *Jurnal EMACS.* **2**(2): 55-63.

Hijrawati. 2022. Analisis Regresi Data Panel pada Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan Sulawesi Tenggara Tahun 2017-2018. *Jurnal Matematika Komputasi & Statistika*. **2**(3): 1-9.

Hsiao, C. 2003. *Analysis of Panel Data. 2nd Edition*. Cambridge University Press. Cambridge.

Jaya, I.G.N.M. & Sunengsih, N. 2009. Kajian Analisis Regresi dengan Data Panel, hlm. 51-58. *Prosiding Seminar Nasional penelitian, Pendidikan, dan penerapan*. Fakultas MIPA, Universitas Yogyakarta. Yogyakarta.

Kuncoro, M. 2011. *Metode Kuantitatif Teori dan Aplikasi untuk Bisnis & Ekonomi*. Ed. ke-4. UPP STIM YKPN, Yogyakarta.

Kusumawati, N., F. Marisa, I. D. Wijaya, & U. W. Malang. 2017. Prediksi Kurs Rupiah Terhadap Dolar Amerika dengan Menggunakan Metode Regresi Linear. *J. Inform. Merdeka Pasuruan*. **2**(3): 45–56.

Montgomery, D. C. & Peck, E. A. 1992. *Introduction to Linear Regression Analysis.* 2 nd Edition. John Wiley & Sons, Toronto.

Murti, D. 2019. Analisis Trend Pada Harga Garam yang dipengaruhi oleh Curah Hujan di Kabupaten Jeneponto (Skripsi). Jurusan Matematika FST UIN Alauddin Makassar, Makassar.

Nachrowi, N. D. & Usman, H. 2006. *Pendekatan Populer dan Praktis Ekonometrika Untuk Analisis Ekonomi dan Keuangan*. Fakultas Ekonomi Universitas Indonesia, Jakarta.

Nandita, D. A., Alamsyah, L. B., Jati, E. P., & Widodo, E. 2019. Regresi Data Panel untuk Mengetahui Faktor-Faktor yang Mempengaruhi PDRB di Provinsi DIY Tahun 2011-2015. *Indonesian Journal of Applied Statistics*. **2**(1): 42-52.

Pangestika, S. 2015. Analisis Estimasi Model Regresi Data Panel dengan Pendekatan Common Effect Model, Fixed Effect Model, dan Random Effect Model (Skripsi). Jurusan Matematika Universitas Negeri Semarang, Semarang.

Rachmawati, F., Sigalingging, N.S., & Kiftiah, M. 2022. Proyeksi Indeks Pembangunan Gender (IPG) di Kalimantan Barat dengan Metode Trend Parabolik. *Jurnal Formasi.* **2**(2): 83-91.

Rahmawati. 2015. Model Trend untuk Peramalan Jumlah Penduduk Studi Kasus pada Pertumbuhan Penduduk Kabupaten Gowa. *Journal of Technology Research in Information System and Enginering.* **2**(2): 46-52.

Rizal, M., Indah, D.R., Meutia, R. 2021. Analisis Peramaan Menggunakan Trend Moment pada Kilang Padi Doa Ibu Perlak Kacamatan Peureulak. *Jurnal Samudra Ekonometrika*. **5**(2): 161-168.

Rizal, A., Yantieka, N. 2022. Analisis Regresi Data Panel untuk Pemodelan Tingkat Kemiskinan Pulau Sumatera dengan Variabel Pendidikan Tahun 2016-2022. *Jurnal Ekonomi-Teknik*. **1**(7): 518-529.

Saputri, A. N. E., Azka, M., & Dinata, S. A. W. 2020. Analisis Regresi Data Panel Frekuensi Penerbangan dan Volume Cargo Terhadap Jumlah Penumpang Pesawat Tahun 2018. *SPECTA Jurnal of teknologi.* **10**(10); 1-11.

Septianingsih, A. 2022. Pemodelan Data Panel Menggunakan Random Effect Model untuk Mengetahui Faktor yang Mempegaruhi Umur Harapan Hidup di Indonesia. *Jurnal Ilmiah Pendidikan Matematika*. **3**(3): 525-536.

Singgih, S. 2002. *Mengolah Data Statistik Secara Professional*. Elex Media Komputindo. Jakarta.

Singh, S. & Sachdeva, T. K. 2020. Financial Performance of Selected IT Companies in India: A Panel Data Approach. *International Journal of Arts, Science and Humanities.* **7**(3): 7-14.

Taboga, Marco. 2021. *Generalized Least Squares*. Lectures on probability theory and mathematical statistics. Kindle Direct Publishing.

Wibawa, G. N. A., Yahya, I., Baharudin, Rahman, G. A., & Agusrawati. 2022. Analisis Regresi Data Panel pada Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan Sulawesi Tenggara Tahun 2017-2020. *Jurnal Matematika, Komputasi, dan Statistika*. **2**(3): 1-9.

Widarjono, A. 2005. *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Fakultas Ekonomi Universitas Islam Indonesia. Yogyakarta.

Winarno. W. 2007. *Analisis Ekonometrika dan Statistika dengan Eviews Edisi Ke-1.* UPP STIM YKPN. Yogyakarta.