PEMODELAN ANALISIS REGRESI 2-LEVEL DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD* PADA DATA ANGKA PERTISIPASI SEKOLAH DI INDONESIA

(Skripsi)

Oleh

Reva Noviana Putri
1717031066

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

ABSTRAK

PEMODELAN ANALISIS REGRESI 2-LEVEL DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD* PADA DATA ANGKA PERTISIPASI SEKOLAH DI INDONESIA

Oleh

Reva Noviana Putri

Data berstruktur hirarki adalah data yang terdiri dari beberapa unit yang diamati terkelompok dalam unit level yang lebih tinggi. Melakukan analisis tanpa mempertimbangkan data terhadap level yang lebih tinggi dapat mengakibatkan heteroskedasitas terhadap kesalahan dan ketidakpuasan terhadap hasil pada analisis. Untuk mengetahui faktor-faktor yang mempengaruhi angka partisipasi sekolah di Indonesia pada level provinsi dan kabupaten. Penelitian ini menggunakan regresi multilevel dengan metode *restricted maximum likelihood*. Selain itu, analisis regresi multilevel akan digunakan untuk menentukan nilai keragaman yang dijelaskan di masing-masing level. Menurut hasil analisis diketahui model yang terbaik yaitu model regresi multilevel tanpa menyertakan variabel Z. Sehingga faktor-faktor yang mempengaruhi angka partisipasi sekolah adalah jumlah SD, jumlah SMP, jumlah SMA dan jumlah SMK.

Kata Kunci : Angka Partisipasi Sekolah, Regresi Multilevel, Data Hirarki, *Restricted Maximum Likelihood* (REML).

ABSTRACT

MODELING ANALYSIS REGRESSION 2-LEVEL USING THE *RESTRICTED MAXIMUM LIKELIHOOD* METHOD ON SCHOOL PARTICIPATION NUMBERS IN INDONESIA

By

Reva Noviana Putri

Hierarchically structured data are data consisting of several observed units grouped into higher-level units. Conducting analysis without considering data at a higher level can result in heterosexuality of error and dissatisfaction with the results of the analysis. In order to identify the factors that influence school participation rates in Indonesia at the provincial and district levels, the study used multilevel regression with a method of limiting the maximum probability. In addition, a multi-level regression analysis will be used to determine the value of diversity described at each level. According to the analysis, the best model is the multilevel regression model without the Z variable. Then the factors that influence the number of school participants are the number of elementary schools, the number of junior high schools, the number of senior high schools, and the number of vocational high schools.

Keywords: School Attendance, Multilevel Regression, Hierarchical Data, Restricted Maximum Likelihood (REML).

PEMODELAN ANALISIS REGRESI 2-LEVEL DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD* PADA ANGKA PARTISIPASI SEKOLAH DI INDONESIA

Oleh

Reva Noviana Putri

Skripsi

Sebagai Salah Satu Syarat Untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung

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

Judul skripsi	: **PEMODELAN ANALISIS REGRESI 2-LEVEL DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD* PADA ANGKA PARTISIPASI SEKOLAH DI INDONESIA**
Nama mahasiswa	: **Reva Noviana Putri**
Nomor pokok mahasiswa	: **1717031066**
Jurusan	: **Matematika**
Fakultas	: **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Dr. Khoirin Nisa, S.Si., M.Si.
NIP. 19740726 200003 2 001

Drs. Tiryono Ruby, M.Sc., Ph.D
NIP. 19620704 198803 1 002

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si
NIP. 19750316200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Dr. Khoirin Nisa, S.Si., M.Si.**	....
Sekertaris	: **Drs. Tiryono Ruby, M.Sc., Ph.D.**	....
Penguji Bukan Pembimbing	: **Drs. Nusyirwan, M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi **: 4 April 2024**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama Mahasiswa	: **Reva Noviana Putri**
Nomor Pokok Mahasiswa	: **1717031066**
Jurusan	: **Matematika**
Judul Skripsi	: **Pemodelan Analisis Regresi 2-Level dengan Metode *Restricted Maximum Likelihood* Pada Angka Partisipasi Sekolah Di Indonesia**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan Salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 4 April 2024

Penulis

METERAI TEMPEL

SEPULUH RIBU RUPIAH 10000

05D06ALX134215787

Reva Noviana Putri

NPM. 1717031066

RIWAYAT HIDUP

Penulis bernama Reva Noviana Putri, dilahirkan di Lampung Timur pada tanggal 06 November 1999, sebagai anak pertama dari dua bersaudara dari pasangan Bapak Soldi Hartanto dan Ibu Lulus Setyorini.

Pada tahun 2005, penulis menyelesaikan pendidikan awal di TK Pertiwi 2 Banarjoyo. Kemudian melanjutkan pendidikan sekolah dasar di SD Negeri 2 Banarjoyo dan lulus pada tahun 2011. Kemudian penulis melanjutkan pendidikan sekolah menengah pertama di SMP Negeri 1 Batanghari dan lulus pada tahun 2014. Penulis melanjutkan pendidikan sekolah menengah atas di SMA Kartikatama Metro dan lulus pada tahun 2017. Kemudian penulis diterima sebagai Mahasisw a S1 pada Jurusan Matematika FMIPA Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) pada tahun 2017.

Pada tahun 2020, Penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Ciptamulya, Kecamatan Kebun Tebu, Kabupaten Lampung Barat sebagai bentuk pengabdian kepada masyarakat dan bentuk pelaksanaan tri darma perguruan tinggi negeri. Pada tahun yang sama Penulis juga melaksanakan Kuliah Praktik (KP) di Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Kantor Cabang Metro, sebagai bentuk aplikasi bidang ilmu di dunia kerja.

KATA INSPIRASI

“Laa Tahzan, Innallaha Ma’ana”
(Jangan bersedih, sesungguhnya Allah bersama kita)

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

“When my little steps pile up, They will dorm a firm path one day”
(Woozi)

“Setiap orang memiliki alasan saat merasa kesulitan. Tidak semua orang tahu kesulitan masing-masing, tetapi kamu bisa selalu ada untuk orang itu dan menghirup udara segar bersama dengan mereka ”
(S.coups)

PERSEMBAHAN

Alhamdulillahirobbil'alamin, puji dan Syukur atas kehadiran Allah SWT yang maha pengasih lagi maha penyayang, untuk segala rahmat dan hidayah-nya yang telah memberikan petunjuk sehingga dapat menyelesaikan skripsi ini. Dengan segala kerendahan hati penulis persembahkan skripsi ini kepada:

Ayah dan Ibuku Tercinta

Orangtua yang selalu mendukung dan mendoakan keberhasilan putra- putrinya. Terima kasih untuk semua yang telah kalian berikan.

Adik Tersayang dan Seluruh Keluarga Besar

Adik dan keluarga besar yang selalu memberikan semangat, serta selalu memotivasi untuk bersikap dan bertindak yang baik.

Dosen Pembimbing dan Pembahas

Atas bimbingan dan motivasi dalam menyelesaikan skripsi.

Teman-Teman Terbaik

Sosok yang selalu menjadi tempat berbagi suka duka selama perkuliahan. Terima kasih untuk kehangatan yang telah kalian berikan

Rekan Seperjuangan Matematika 2017

Terima kasih telah menjadi keluarga terbaik selama perkuliahan

Almamater tercinta

Universitas Lampung

SANWACANA

Puji dan Syukur penulis panjatkan kehadirat Allah SWT, yang telah memberikan Rahmat dan hidayah-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pemodelan Analisis Regresi 2-Level Dengan Metode Restricted Maximum Likelihood Pada Data Angka Pertisipasi Sekolah Di Indonesia”.

Penulis menyedari bahwa skripsi ini tidak akan terwujud tanpa adanya bantuan, bimbingan, dan doa dari berbagai pihak sehingga skripsi ini dapat terselesaikan. Dengan segala keredahan hati penulis mengucapkan terima kasih kepada:

1. Ibu Dr. Khoirin Nisa, S.Si., M.Si, selaku Dosen Pembimbing I yang telah bersedia membimbing, memberi saran dan arahan, serta memberikan waktu, motivasi dalam menyelesaikan skripsi ini.
2. Bapak Drs. Tiryono Ruby, M.Sc., Ph.D, selaku Dosen Pembimbing II yang telah memberikan bimbingan, saran dan pengarahan serta motivasi kepada penulis.
3. Bapak Drs. Nusyirwan, M.Si, selaku Dosen Penguji yang telah memberikan kritik dan saran yang membangun serta evaluasi selama proses penyusunan skripsi hingga dapat menjadi lebih baik.
4. Bapak Prof. Drs. Mustofa Usman, M.A, Ph.D., selaku Dosen Pembimbing Akademik yang telah memberi bimbingan dan arahan selama masa perkuliahan.
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, Staff, Karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Ayah, ibu, Sutan, serta seluruh keluarga besar yang selalu memberikan doa, dukungan, serta memberikan semangat dan motivasi yang begitu besar kepada penulis.
9. Teman-temanku Kelvin, Diah, Kurnia, Ilma, Mevita, Nupus, dan Ilham yang telah menemani, memberi dukungan, memberikan bantuan kepada penulis.
10. Teman-teman Matematika 2017 yang telah bersama selama masa perkuliahan.
11. Semua pihak yang telah membantu dalam menyelesaikan skripsi ini.

Akhir kata, semoga Allah SWT senantiasa melimpahkan karunia-Nya dan membalas segala kebaikan pihak-pihak yang telah membantu penulis dalam menyusun skripsi ini. Penulis menyadari bahwa dalam skripsi ini masih banyak kekurangan. Oleh karena itu, kritik dan saran sangat penulis harapkan untuk menjadi bahan perbaikan kedepannya. Semoga skripsi ini dapat memberikan manfaaat bagi pihak yang membutuhkan.

Bandar Lampung, 4 April 2024
Penulis,

Reva Noviana Putri

DAFTAR ISI

Halaman

DAFTAR TABEL

Tabel Halaman

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Penelitian sosial umumnya berkosentrasi pada bagaimana menyelidiki hubungan antara individu dengan komunitas sosialnya. Secara umum, ketika individu berkorelasi terhadap komunitas sosialnya, lingkungan sosial mereka mempengaruhinya. Data yang memiliki struktur hirarki dianalisis menggunakan model multilevel, yang disebut juga sebagai penelitian multilevel. Model multilevel digunakan untuk menganalisis data dari berbagai level, dimana level yang lebih rendah tersarang dalam level yang lebih tinggi (Hox, 2002).

Analisis ini disebut analisis data multilevel atau yang dikenal juga dengan *herarchi linier model* (Harlan, 2016). Melakukan analisis tanpa mempertimbangkan informasi pada level yang lebih tinggi menyebabkan heteroskedastisitas pada galat dan hasil dari analisisnya menyebabkan ketidakpuasan (Tantular, 2009). Digunakan model regresi multilevel untuk mengatasi masalah pada model regresi tersebut.

Dalam menduga parameter model regresi multilevel, terdapat beberapa metode diantaranya adalah metode *maximum likelihood*, metode *restricted maximum likelihod* (REML) dan metode *iterative generalized least square* (IGLS). Metode *maximum likelihood* menghasilkan penduga ragam yang bias, sedangkan *restricted maximum likelihod* (REML) untuk menghasilkan penduga ragam yang tak bias (West, dkk., 2007).

Maka dari itu, dalam penelitian ini dikaji penggunaan metode *restricted maximum likelihood* pada analisis regresi 2-level untuk menentukan faktor apa saja yang mempengaruhi angka partisipasi sekolah (APS) yang ada di Indonesia pada tahun 2022.

1.2 Tujuan Penelitian

Tujuan penelitian ini adalah menerapkan model regresi 2-level menggunakan estimator *restricted maximum likelihood* pada data multilevel untuk mengetahui faktor yang mempengaruhi angka partisipasi sekolah di indonesi tahun 2022.

1.3 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Dapat menerapkan model regresi 2-level pada data multilevel untuk mengetahui faktor-faktor yang mempengaruhi angka partisipasi sekolah anak di Indonesia tahun 2022.
2. Dapat mengetahui estimasi parameter pada model regresi 2-level dengan menggunakan metode *restricted maximum likelihood* pada data angka partisipasi sekolah anak di Indonesia tahun 2022.

II. TINJAUAN PUSTAKA

2.1 Model Regresi

Pada umumnya regresi merupakan suatu metode yang digunakan untuk memprediksi nilai harapan yang bersyarat (Kutner, dkk., 2004). Pada abad ke-19 analisis regresi dikembangkan oleh Sir Francis Galton. Analisis regresi bertujuan untuk mempelajari suatu hubungan antara variabel bebas (X) dengan variabel terikat (Y).

Gujarati (2006), menguraikan tiga tujuan utama model regresi. Tujuan pertama adalah untuk menjelaskan hubungan sebab akibat antara variabel independen dan variabel dependen. Tujuan kedua adalah untuk mengetahui dampak individual masing-masing variabel independen terhadap variabel dependen. Terakhir, tujuan ketiga adalah membuat prediksi nilai variabel terikat berdasarkan nilai tertentu dari variabel bebas. Model regresi dapat menyatakan hubungan antara variabel terikat dengan variabel bebas.

Analisis regresi yang melibatkan satu variabel bebas dan satu variabel terikat merupakan analisis regresi linier sederhana. Persamaan umum dapat ditulis sebagai berikut:

$$Y = \alpha + \beta X + \varepsilon \tag{2.1}$$

dengan

$Y =$ variabel terikat

$\alpha =$ intersep

$\beta =$ *slope*

X = variabel bebas

ε = galat

Sedangkan analisis regresi yang melibatkan satu variabel terikat dan lebih dari satu variabel bebas adalah analisis regresi linier berganda. Persamaan umumnya ditulis sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_j + \varepsilon \qquad (2.2)$$

Model regresi pada persamaan (2.2) dapat disajikan dalam bentuk matriks sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \text{ dengan } \varepsilon \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \qquad (2.3)$$

dengan

$\boldsymbol{Y}$ = vektor variabel terikat berukuran $(nx1)$

$\boldsymbol{X}$ = vektor variabel bebas berukuran $\left(nx(p+1)\right)$

$\boldsymbol{\beta}$ = vektor koefisien regresi berukuran $(p+1x1)$

$\boldsymbol{\varepsilon}$ = vektor galat berukuran $(nx1)$

Nilai duga dari parameter atau kondisi yang sebenarnya disebut koefisien dalam suatu persamaan regresi. Besar persentase keragaman total variabel terikat (Y) yang dapat dijelaskan oleh model regresi dapat dihitung dengan koefisien determinasi. Sementara itu, kekuatan hubungan linier antara variabel terikat dan variabel bebas dapat diukur menggunakan koefisien korelasi.

Asumsi-asumsi pada analisis regresi adalah sebagai berikut:

1. Galat berdistribusi normal $\varepsilon_i \sim N(0, \sigma^2)$;
2. Ragam galat homogen $Var(\varepsilon_1) = \sigma^2; i = 1,2, \dots, n$;
3. Nilai ε_i adalah bebas satu dengan yang lainnya $E(\varepsilon_i) = 0$ dan $E(\varepsilon_i^2) = \sigma^2$;
4. X dan Y terkait secara linier, garis lurus akan terbentuk antara setiap nilai X yang dihubungkan dengan nilai Y.

Asumsi-asumsi tersebut harus dipenuhi saat melakukan analisis regresi (Myers, 1990).

2.2 Model Linier Campuran (Linier Mixed Model)

Model linier campuran (*linier mixed models*) adalah persamaan yang didalamnya menggabungkan model efek acak (*random effect*) dan efek tetap (*fixed effect*). Jika nilai koefisien regresinya sama untuk semua anggota sampel, variabel bebas dianggap memiliki efek tetap. Sebaliknya, jika nilai koefisien regresinya berbeda di antara dua atau lebih subjek sampel, variabel bebas dianggap memiliki efek acak (Harlan, 2016), Menurut Bryck dan Raudenbush (1987), model linier campuran umumnya terdiri dari tiga komponen yaitu:

1. Efek acak, merupakan suatu efek yang diperoleh dari sampel acak yang nilainya memiliki pengaruh suatu variabel.
2. Efek hirarki, merupakan efek yang dihasilkan dari pengaruh berbagai peubah pada berbagai level.
3. Pengukuran berulang, dimana pengamatan sebelumnya dikaitkan pada pengamaan baru.

Menurut Rencher dan Schaalje (2007), persamaan sederhana model linier campuran adalah

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Z_1 u_1} + \boldsymbol{Z_2 u_2} + \cdots + \boldsymbol{Z_m u_m} + \boldsymbol{\varepsilon} \qquad i = 1,2,3,\dots,m \tag{2.4}$$

dengan

$\boldsymbol{Y} =$ vektor respon $(nx1)$

$\boldsymbol{X} =$ matriks prediktor efek tetap (nxp)

$\boldsymbol{\beta} =$ vektor parameter efek tetap $(px1)$

$\boldsymbol{Z_i} =$ matriks prediktor efek acak (nxr_i)

$\boldsymbol{u_i} =$ vektor parameter efek acak (r_ix1)

$\boldsymbol{\varepsilon} =$ vektor galat $(nx1)$

$E(\boldsymbol{\varepsilon}) = \mathbf{0}$ dan $Cov\ (\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I_n}$ untuk $i = 1, \dots\ , m$

$E(\boldsymbol{u_i}) = \mathbf{0}$ dan $Cov\ (\boldsymbol{u_i}) = \sigma^2 \boldsymbol{I_{n_i}}$ untuk $i = 1, \dots\ , m$

$Cov\left(\boldsymbol{u_i}, \boldsymbol{u_j}\right) = \boldsymbol{O}$ untuk $i \neq j$, dimana $\boldsymbol{O}$ adalah $\boldsymbol{r_i x\ r_j}$

$Cov(\boldsymbol{u_i}, \varepsilon) = \boldsymbol{O}$ untuk i, dimana $\boldsymbol{O}$ adalah $\boldsymbol{r_i x\ n}$

Maka $E(\boldsymbol{Y}) = \boldsymbol{X\beta}$ dan $Cov(\boldsymbol{Y}) = \sum_{i=1}^{m} \boldsymbol{\sigma_i^2}\ \boldsymbol{Z_i Z_i'} + \boldsymbol{\sigma^2 I_n}$

2.3 Struktur Data Model 2 Level

Data berstruktur hirarki adalah data yang terdiri dari beberapa unit yang diamati terkelompok dalam unit level yang lebih tinggi. Data multilevel adalah istilah lain untuk data hirarki (Dew, 2008). Dalam pemodelan multilevel dengen 2-level, sebanyak n individu yang berasal dari m kelompok. $Y_{ij}, Y_{2j}, \dots, Y_{nj}$ adalah variabel terikat masing-masing n_j individu pada kelompok ke-j, $j = 1,2,\dots,m$ dan jika $X_{1j}, X_{2j}, \dots, X_{kj}$ adalah variabel bebas pada level 1 untuk kelompok ke- j,serta $Z_1, Z_2, \dots, Z_l$ adalah variabel bebas pada level 2, maka struktur pemodelan multilevel dengan 2-level dapat dilihat dalam tabel berikut.

Tabel 1. Struktur Data Multilevel dengan 2-Level.

Kelompok	Observasi	Variabel terikat Y	Variabel bebas Level 1				Variabel bebas Level 2			
			X_1	X_2	...	X_k	Z_1	Z_2	...	Z_l
1	1	y_{11}	x_{111}	x_{211}	...	x_{k11}	z_{11}	z_{21}	...	z_{l1}
	2	y_{21}	x_{121}	x_{221}		x_{k21}				
	⋮	⋮	⋮	⋮		⋮				
	i	y_{i1}	x_{1i1}	x_{2i1}		x_{ki1}				
	⋮	⋮	⋮	⋮		⋮				
	n_j	y_{nj1}	x_{1nj1}	x_{2nj1}		x_{knj1}				
2	1	y_{12}	x_{112}	x_{212}	...	x_{k12}	z_{12}	z_{22}	...	z_{l2}
	2	y_{22}	x_{122}	x_{222}		x_{k22}				
	⋮	⋮	⋮	⋮		⋮				
	i	y_{i2}	x_{1i2}	x_{2i2}		x_{ki2}				
	⋮	⋮	⋮	⋮		⋮				
	n_j	y_{nj2}	x_{1nj2}	x_{2nj2}		x_{knj2}				
⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	...	⋮
M	1	y_{11}	x_{11m}	x_{21m}	...	x_{k1m}	z_{1m}	z_{2m}	...	z_{lm}
	2	y_{21}	x_{12m}	x_{22m}		x_{k2m}				
	⋮	⋮	⋮	⋮		⋮				
	i	y_{i1}	x_{1im}	x_{2im}		x_{kim}				
	⋮	⋮	⋮	⋮		⋮				
	n_j	y_{nj1}	x_{1njm}	x_{2njm}		x_{knjm}				

2.4 Model Regresi Multilevel

Data berstruktur hirarki, yang juga dikenal sebagai data multilevel atau data bersarang, terdiri dari unit-unit yang diamati yang terkumpul dalam unit level yang lebih tinggi atau bersarang (Tantular, 2009).

Dalam model regresi multilevel, nilai variabel terikat dipengaruhi oleh variabel dengan efek acak dan efek tetap. Model regresi multilevel mencakup dua hal. Hal pertama adalah data tersebut sesuai untuk model dengan berstruktur hirarki, dimana unit level 2 tersarang dalam unit level 2, unit 2 tersarang dalam unit level 1, dan seterusnya. Yang kedua adalah parameter model terlihat memiliki struktur hirarki (Raudenbush dan Bryk, 2002).

Pada regresi multilevel variabel terikat diukur pada level 2, sedangkan variabel bebas dapat ditentukan pada setiap level. Analisis regresi 2-level adalah bentuk paling sederhana dari analisis regresi multilevel. Model regresi 2-level merupakan model 2-level dimana data individu adalah level 1 dan data kelompok adalah level 2 (West *et,al.*, 2007). Ditunjukkan persamaan regresi 2-level sebagai berikut:

1. Model level-1

   Model level 1 merupakan model yang disusun tanpa memperhitungkan level kelompok. Untuk setiap kelompok, pemodelan multilevel dapat ditulis sebagai berikut:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{kj}X_{kij} + \varepsilon_{ij}$$

$$y_{ij} = \beta_{0j} + \sum_{q=1}^{k} \beta_{qj}X_{qij} + \varepsilon_{ij} \tag{2.5}$$

   Dengan $q = 1,2, \dots, k$, individu- $i = 1,2, \dots, n_j$ dan kelompok - $j = 1,2, \dots, m$ atau dinyatakan dalam bentuk matriks berikut

$$\boldsymbol{y_j} = \boldsymbol{X_j}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j \qquad \boldsymbol{\varepsilon}_j \sim N(0, \sigma^2 I) \tag{2.6}$$

   Dengan

$$\boldsymbol{y}_j = [y_{1j} \quad y_{2j} \quad \cdots \quad y_{n_j j}]',$$

$$\boldsymbol{X}_j = \begin{bmatrix} 1 & x_{11j} & x_{21j} & x_{k1j} \\ 1 & x_{12j} & x_{22j} & x_{k2j} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{k_{n_j}j} & x_{kn_jj} & x_{k_{n_j}j} \end{bmatrix}$$

$$\boldsymbol{\beta}_j = [\boldsymbol{\beta}_{0j} \quad \boldsymbol{\beta}_{1j} \quad \cdots \quad \boldsymbol{\beta}_{kj}]',$$

$$\boldsymbol{\varepsilon}_j = [\boldsymbol{\varepsilon}_{0j} \quad \boldsymbol{\varepsilon}_{1j} \quad \cdots \quad \boldsymbol{\varepsilon}_{n_jj}]'$$

Keterangan:

$\boldsymbol{y}_j$: vektor variabel terikat

$\boldsymbol{X}_j$:matriks variabel bebas untuk parameter tetap

$\boldsymbol{\beta}_j$: vektor parameter efek tetap

$\boldsymbol{\varepsilon}_j$: vektor galat

2. Model level-2

Nilai koefisien regresi pada level-1, dinotasikan sebagai β_{pj} dengan $p = 0,1,2, \dots, k$ pada model level-1, menunjukkan variasi antar kelompok yang berbeda. Untuk menjelaskan variasi ini, model level-2 dibangun. Pembentukan model level-2 dilakukan secara individual untuk setiap koefisien regresi sebagai respon ke-p dengan memanfaatkan variabel independen pada level-2. Pemodelan pada level-2 dapat digambarkan sebagai berikut:

$$\beta_{pj} = \gamma_{p0} + \gamma_{1p}Z_{1j} + \cdots + \gamma_{lp}Z_{lj} + u_{pj}$$

$$\beta_{pj} = \gamma_{p0} + \sum_{p=1}^{k} \gamma_{pl}Z_{lj} + u_{pj} \qquad (2.7)$$

Dengan $p = 1,2, \dots, k$, dan $l = 1,2, \dots, n_j$ atau dinyatakan dalam bentuk matriks berikut

$$\boldsymbol{\beta_p} = \boldsymbol{Z\gamma_p} + \boldsymbol{u_p} \qquad \boldsymbol{u_p} \sim N(0, \sigma_\tau^2 I) \qquad (2.8)$$

Dengan

$$\boldsymbol{\beta}_p = [\boldsymbol{\beta_{p1}} \quad \boldsymbol{\beta_{p2}} \quad \cdots \quad \boldsymbol{\beta_{pm}}]',$$

$$\boldsymbol{X}_j = \begin{bmatrix} 1 & Z_{11} & Z_{21} & \cdots & Z_{l1} \\ 1 & Z_{12} & Z_{22} & \cdots & Z_{l2} \\ \vdots & \vdots & \vdots & \vdots & \dots \\ 1 & Z_{1m} & Z_{2m} & \cdots & Z_{lm} \end{bmatrix}$$

$$\boldsymbol{\gamma}_p = [\boldsymbol{\gamma}_{0p} \quad \boldsymbol{\gamma}_{1p} \quad \cdots \quad \boldsymbol{\gamma}_{lp}]',$$

$$\boldsymbol{u}_p = [\boldsymbol{u}_{p1} \quad \boldsymbol{u}_{p2} \quad \cdots \quad \boldsymbol{u}_{pm}]'$$

Jika persamaan 2.6 disubtitusikan dalam persamaan 2.8 maka didapatkan model sebagai berikut:

$$\gamma_{ij} = \beta_{0j} + \beta_{1j}X_{1j} + \beta_{2j}X_{2ij} + \cdots + \beta_{kj}X_{kij} + \varepsilon_{ij}$$

$$\begin{aligned}\gamma_{ij} = \gamma_{oo} + \gamma_{10}Z_{1j} + \cdots + \gamma_{l0}Z_{lj} + u_{0j} \\ + \left(\gamma_{01} + \gamma_{11}Z_{1j} + \gamma_{21}Z_{2j} + \cdots + \gamma_{l1}Z_{lj} + u_{1j}\right)X_{1ij} + \cdots \\ + \left(\gamma_{02} + \gamma_{12}Z_{1j} + \gamma_{22}Z_{2j} + \cdots + \gamma_{l2}Z_{lj} + u_{2j}\right)X_{2ij} + \cdots \\ + \left(\gamma_{0k} + \gamma_{1k}Z_{1j} + \gamma_{2k}Z_{2j} + \cdots + \gamma_{lk}Z_{lj} + u_{kj}\right)X_{kij} + \varepsilon_{ij}\end{aligned}$$

$$y_{ij} = y_{00} + \sum_{q=1}^{l} y_{0q}Z_{qj} + \sum_{p=1}^{k} y_{0p}Z_{pij} + \sum_{p=1}^{k}\sum_{q=1}^{l} y_{qp}Z_{qj}X_{pij} + u_{0j} + \sum_{p=1}^{k} u_{pj}X_{pij} + \varepsilon_{ij} \tag{2.9}$$

Atau dinyatakan dalam bentuk matriks sebagai berikut:

$$\boldsymbol{y}_j = \boldsymbol{X}_j\boldsymbol{Z}_j\boldsymbol{y} + \boldsymbol{X_j}\boldsymbol{u}_j + \boldsymbol{\varepsilon}_j \tag{2.10}$$

Dengan $\boldsymbol{X}_j\boldsymbol{Z}_j\boldsymbol{y}$ adalah efek tetap dan $\boldsymbol{X_j}\boldsymbol{u}_j + \boldsymbol{\varepsilon}_j$ adalah efek acak.

2.5 Uji Asumsi

2.5.1 Uji Normalitas

Suatu uji yang digunakan dalam mengetahui apakah suatu data berdistribusi normal atau tidak berdistribusi normal adalah uji normalitas. Uji Kolmogorov-Smirnov adalah salah satu uji statistik untuk uji normalitas. Uji Kolmogorov-Smirnov membandingkan dua distribusi data yaitu distribusi yang diamati dan distribusi yang dihipotesiskan. Jika distribusi yang diamati sama dengan distribusi yang dihipotesiskan, maka dapat ditarik kesimpulan bahwa data yang diamati memiliki distribusi normal atau sebaran normal (Kurniawan, 2008).

Pada uji Kolmogorov-Smirnov, hipotesis nol menyatakan data yang diamati berasal dari sebuah populasi yang memiliki ditribusi normal. Sebaliknya, hipotesis alternatif menyatakan bahwa data yang diteliti tidak berasal dari populasi yang memiliki distribusi normal. Misalkan $x_1, x_2, x_3, \cdots, x_k$ merupakan nilai pada sampel acak (random sampel). Misalkan $f(x_i)$ menyatakan probabilitas dari nilai x_i, sedangkan $F(x_i) = f(x \leq x_i)$ menyatakan probabilitas kumulatif dari nilai x_i, dimana $i = 1,2,3, \dots, k$. Selanjutnya andaikan Z_i merupakan nilai normal terstandarisasi dari hasil transformasi nilai x_i dan $F(Z_i) = f(Z \leq Z_i)$ menyatakan probabilitas kumulatif dari nilai normal Z_i terstandarisasi. Nilai normal Z_i terstandarisasi merupakan hasil transformasi dari nilai X_i yang dihitung dengan rumus sebagai berikut:

$$Z_i = \frac{X_i - \bar{X}}{s}; \text{ dimana } i = 1,2,3, \dots, k \tag{2.11}$$

Dimana rata-rata sampel yaitu $\bar{X}$ digunakan sebagai estimasi dari rata-rata populasi μ, sedangkan s adalah standar deviasi sampel yang digunakan sebagai estimasi dari standar deviasi populasi σ. Misalkan D_i menyatakan nilai mutlak dari selisih antara $F(Z_i)$ dan $F(X_i)$, yakni

$$D_i = |F(Z_i) - F(X_i)|, i = 1,2,3, \dots, k. \tag{2.12}$$

Nilai D_i maximum (D_{max}) adalah nilai statistik dari uji Kolmogorov-Smirnov. Selanjutnya Nilai D_{max} dibandingkan terhadap nilai kritis berdasarkan tabel distribusi Kolmogorov-Smirnov untuk pengambilan keputusan terhadap hipotesis. Keputusan hipotesis dalam uji Kolmogrov-Smirnov adalah sebagai berikut:

$H_0 =$ galat berdistribusi normal.
$H_1 =$ galat tidak berdistribusi normal.

Dengan aturan dalam pengambilan keputusan terhadap hipotesis berdasarkan uji Kolmogorov-Smirnov.

Uji hipotesis

- Jika $D_{max} \leq$ nilai kritis, maka H_0 diterima.
- Jika $D_{max} >$ nilai kritis, maka H_0 ditolak.

Keputusan terhadap hipotesis dibuat dengan membandingkan nilai probabilitas (p-value) dari uji Kolmogorov-Smirnov dengan tingkat signifikansi α (*significance level*). Pengambilan keputusan berdasarkan pendekatan nilai probabilitas adalah sebagai berikut:

- Jika $p - value \geq$ tingkat signifikansi, maka H_0 diterima.
- Jika $p - value <$ tingkat signifikansi, maka H_0 ditolak.

Berdasarkan pengambilan keputusan jika H_0 diterima maka dapat disimpulkan bahwa data berdistribusi normal, sedangkan jika H_0 ditolak maka dapat disimpulkan bahwa data tidak berdistribusi normal.

2.5.2 Uji Multikolineritas

Uji multikolineritas adalah metode statistik yang sering digunakan dalam mengevaluasi apakah terdapat korelasi yang tinggi antara variabel-variabel bebas dalam sebuah model regresi linier berganda. *Variance Inflation Factor* (VIF) adalah uji statistik yang sering digunakan untuk menguji gangguan multikolineritas. VIF dapat diinterpretasikan sebagai hasil dari antar variabel prediktor ke-*i* pada ragam penduga koefisien regresi.
Perhitungan VIF adalah sebagai berikut:

$$VIF_i = \frac{1}{1-R_I^2}\text{; dimana } I = 1,2,\dots,p-1 \quad (2.13)$$

dimana :

VIF < 10 : dapat diindikasikan bahwa tidak adanya korelasi antara variabel bebasyang signifikan.

VIF ≥ 10 : dapat diindikasikan bahwa terdapat salah satu variabel bebas yang merupakan fungsi dari variabel prediktor yang lain.

Dalam uji statistik *Variance Inflation Factor* (VIF), jika nilai VIF lebih besar dari sepuluh dapat mengindikasikan adanya multkolinieritas yang serius (Zulmi,2011).

2.6 Metode Penduga Parameter

Metode *Maximum Likelihood* atau *Restricted Maximum Likelihood* adalah penduga parameter yang dapat digunakan dalam pemodelean regresi multilevel (Goldstein, 1995). Metode *Maximum Likelihood* mengestimasi parameter dengan cara memaksimumkan fungsi *likelihood* (Hox, 2010). Penduga parameter yang digunakan metode *Maximum Likelihood* ditunjukkan dibawah ini.

Berikut adalah persamaan model linier campuran :

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon} \tag{2.14}$$

dimana $\boldsymbol{E(\varepsilon)} = \boldsymbol{0}, \boldsymbol{cov(\varepsilon)} = \boldsymbol{\sigma^2 I_n} = \boldsymbol{\Sigma_i}$ dan $\boldsymbol{E(u_i)} = \boldsymbol{0}, \boldsymbol{cov}\ \boldsymbol{(u_i)} = \ \boldsymbol{\sigma^2 I_{r_i}} = \boldsymbol{D}$

atau $\begin{bmatrix} \boldsymbol{u_i} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim \boldsymbol{N_{mq+n}} \left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G} & \boldsymbol{0_{mqxn}} \\ \boldsymbol{0_{nxmq}} & \boldsymbol{R} \end{bmatrix} \right)$.

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{D} \end{bmatrix} \text{ dan } \boldsymbol{R} = \begin{bmatrix} \boldsymbol{\Sigma_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma_2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{\Sigma_n} \end{bmatrix}$$

dengan

$\boldsymbol{Y}$ = vektor respon $(nx1)$

$\boldsymbol{X}$ = matriks desain efek tetap (nxp)

$\boldsymbol{\beta}$ = vektor parameter efek tetap $(px1)$

$\boldsymbol{Z}$ = matriks desain efek acak (nxq)

$\boldsymbol{u}$ = vektor parameter efek acak $(qx1)$

$\boldsymbol{\varepsilon}$ = vektor galat $(nx1)$

$\boldsymbol{D}$ = matriks kovarian dari efek acak u

$\boldsymbol{R}$ = matriks kovarian dari vektor galat ε

misalkan

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

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}^* \tag{2.15}$$

dengan

$$\boldsymbol{\varepsilon}^* = \boldsymbol{Zu} + \boldsymbol{\varepsilon}$$

$$= (\boldsymbol{Z} \quad \boldsymbol{I_{nxn}}) \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{\varepsilon} \end{pmatrix}$$

$$= \boldsymbol{A}\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{\varepsilon}\end{pmatrix}$$

$$\boldsymbol{\varepsilon}^{*} \sim \boldsymbol{N_n}(\boldsymbol{0}, \boldsymbol{V})$$

dimana

$$\boldsymbol{V} = \boldsymbol{A}\begin{pmatrix}\boldsymbol{G} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}\end{pmatrix}\boldsymbol{A}^{t}$$

$$= (\boldsymbol{Z} \quad \boldsymbol{I_{nxn}})\begin{pmatrix}\boldsymbol{G} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}\end{pmatrix}\begin{pmatrix}\boldsymbol{Z}^{t}\\ \boldsymbol{I_{nxn}}\end{pmatrix}$$

$$= (\boldsymbol{ZG} \quad \boldsymbol{R})\begin{pmatrix}\boldsymbol{Z}^{t}\\ \boldsymbol{I_{nxn}}\end{pmatrix}$$

$$= \boldsymbol{ZGZ}^{t} + \boldsymbol{R}$$

$$\boldsymbol{V} = \boldsymbol{ZGZ}^{t} + \boldsymbol{R} \tag{2.16}$$

Vektor parameter kovarian θ diduga oleh $\hat{\theta}_{ML}$ dengan memaksimumkan fungsi likelihood.

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon} \tag{2.17}$$

dengan $\begin{bmatrix}\boldsymbol{u_i}\\ \boldsymbol{\varepsilon}\end{bmatrix} \sim \boldsymbol{N_{mq+n}}\left(\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{bmatrix}, \begin{bmatrix}\boldsymbol{G} & \boldsymbol{0_{mqxn}}\\ \boldsymbol{0_{nxmq}} & \boldsymbol{R}\end{bmatrix}\right)$.

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

dengan $\boldsymbol{\varepsilon}^{*} \sim \boldsymbol{N_n}\big(\boldsymbol{0}, \boldsymbol{v}(\boldsymbol{\theta})\big)$dengan$\boldsymbol{V}(\boldsymbol{\theta}) = \boldsymbol{ZG}(\boldsymbol{\theta})\boldsymbol{Z}^{t} + \boldsymbol{R}(\boldsymbol{\theta})$.

Diberikan fungsi *likelihood* sebagai berikut:

$$\ln L(\boldsymbol{\beta}, \boldsymbol{\sigma^2}) = -\frac{1}{2}\{\ln|V(\theta)| + (\boldsymbol{Y} - \boldsymbol{X\beta})^{\boldsymbol{t}}\boldsymbol{V}(\boldsymbol{\theta})^{-\boldsymbol{1}}(\boldsymbol{Y} - \boldsymbol{X\beta})\} \tag{2.18}$$

dengan memaksimumkan log likelihoood untuk θ tetap terhadap β maka diperoleh

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) \coloneqq (\boldsymbol{X}^{t}\boldsymbol{V}(\boldsymbol{\theta})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^{t}\boldsymbol{V}(\boldsymbol{\theta})^{-1}\boldsymbol{Y} \tag{2.19}$$

$$\boldsymbol{Y} \sim \boldsymbol{N_n}(\boldsymbol{X\beta}, \boldsymbol{V}) \text{ dan } \boldsymbol{u} \sim \boldsymbol{N_{mq}}(\boldsymbol{0}, \boldsymbol{G})$$

$$Cov(\boldsymbol{Y}, \boldsymbol{u}) = Cov\ (\boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon}, \boldsymbol{u})$$

$$= Cov\ (\boldsymbol{X\beta}, \boldsymbol{u}) + Cov(\boldsymbol{Zu}, \boldsymbol{u}) + Cov(\boldsymbol{\varepsilon}, \boldsymbol{u})$$

$$= \boldsymbol{0} + \boldsymbol{ZG} + \boldsymbol{0}$$

$$= \boldsymbol{ZG} \tag{2.20}$$

$$\boldsymbol{E}(\boldsymbol{u}|\boldsymbol{Y}) = \boldsymbol{E}(\boldsymbol{u}) + \boldsymbol{cov}(\boldsymbol{u},\boldsymbol{Y})[\boldsymbol{cov}(\boldsymbol{Y})^{-\mathbf{1}}][\boldsymbol{Y} - \boldsymbol{E}(\boldsymbol{Y})]$$
$$= \mathbf{0} + \boldsymbol{G}\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{V}^{-\mathbf{1}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})$$
$$= \boldsymbol{G}\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{V}^{-\mathbf{1}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})$$
$$\boldsymbol{E}(\boldsymbol{u}|\boldsymbol{Y}) = \boldsymbol{G}\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{V}^{-\mathbf{1}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) \tag{2.21}$$

$\boldsymbol{G}\boldsymbol{Z}^{\boldsymbol{t}}\boldsymbol{V}^{-\mathbf{1}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})$ adalah *Best Linier Unbiased Prediktor* (BLUP) dari $\boldsymbol{u}$.
Menurut West dkk (2007) efek tetap $\boldsymbol{\beta}$ dan efek acak $\boldsymbol{u}$ diduga oleh

$$\widehat{\boldsymbol{\beta}} \coloneqq \left(\boldsymbol{X}^{\boldsymbol{t}}\widehat{\boldsymbol{V}}^{-\mathbf{1}}\boldsymbol{X}\right)^{-\mathbf{1}}\boldsymbol{X}^{\boldsymbol{t}}\widehat{\boldsymbol{V}}^{-\mathbf{1}}\boldsymbol{Y}$$
$$\widehat{\boldsymbol{u}} \coloneqq \boldsymbol{G}\boldsymbol{Z}^{\boldsymbol{t}}\widehat{\boldsymbol{V}}^{-\mathbf{1}}\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)$$

dimana $\widehat{\boldsymbol{V}} = \boldsymbol{V}\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{ML}}\right)$.

Jika penduga yang dihasilkan dengan *Maximum Likelihood* adalah penduga bias, maka prosedur pendugaanya harus diubah untuk menghasilkan penduga yang tak bias. Modifikasi ini akan dilakukan pada fungsi profile log-likelihood, yang disebut fungsi *Restricted log-likelihood* berikut ini

$$\ln L(\boldsymbol{\beta},\boldsymbol{\sigma}^{\mathbf{2}}) = -\frac{1}{2}\left\{\ln|\boldsymbol{V}(\boldsymbol{\theta})| + \ln\left|\boldsymbol{X}^{\boldsymbol{t}}\boldsymbol{V}(\boldsymbol{\theta})^{-\mathbf{1}}\boldsymbol{X}\right|\right\} \tag{2.22}$$

Prosedur ini disebut sebagai *Restricted Maximum Likelihood* atau Residual *Maximum Likelihood* (West, dkk., 2007)

2.7 Pengujian Hipotesis

Penduga parameter yang selanjutnya digunakan untuk menguji signifikasi parameter pada model regresi 2-level secara individual.
Hipotesis yang digunakan adalah sebagai berikut:

a. Hipotesis untuk parameter level 1:

   $H_0 :\ \gamma_{pj} = 0$

   $H_1 : \gamma_{pj} \neq 0$

Dengan indeks q=1,2,…,s dan s merupakan banyaknya variabel bebas level 1.

Statistik uji yang digunakan adalah statistik Wald sebagai berikut:

$$t = \frac{\hat{\gamma}_{pj}}{\sqrt{V(\hat{\gamma}_{pj})}} \tag{2.23}$$

b. Hipotesis untuk parameter level 2:

H_0 : $\beta_{qj} = 0$

H_1: $\beta_{qj} \neq 0$

Dengan indeks p =1,2,…,r dan r merupakan banyaknya variabel bebas level 2. Statistik uji yang digunakan adalah statistik Wald sebagai berikut:

$$t = \frac{\hat{\beta}_{qj}}{\sqrt{v(\hat{\beta}_{qj})}} \tag{2.24}$$

Dalam hal ini t mengikuti sebaran t student dengan derajat bebas untuk penduga parameter level 2 adalah $i - r - 1$ sedangkan derajat bebas untuk penduga parameter level 1 adalah $j - s - 1$ (Jones & Steenbergen, 2002).

2.8 Pemilihan Model Terbaik

Nilai deviasi digunakan untuk memilih model terbaik dalam regresi multilevel. Menurut Tantular (2009), uji rasio *likelihood* juga dikenal sebagai sebaran deviasi dapat digunakan untuk memilih model regresi multilevel terbaik. Ukuran deviasi ini digunakan untuk mengukur kesesuaian model. Perhitungan yang digunakan untuk uji coba ini adalah perbedaan nilai deviasi dari dua model, yang terdiri dari:

Hipotesis :

H_0: model tidak signifikan

H_1: model signifikan

Statistik uji :

$$Diff = -2 \log_e \frac{\lambda_0}{\lambda_1} \tag{2.25}$$

dengan

λ_0 = nilai deviasi untuk null model

λ_1 = nillai deviasi untuk full model

Menurut Hox (2010), semakin kecil nilai deviasi model maka semakin cocok. Selain itu, apabila nilai $diff > x^2_{(a,db)}$ maka tolak H_0, dimana db adalah selisih jumlah parameter dari kedua model. Karena itu, dapat disimpulkan bahwa efek acak signifikan, yang artinya terdapat keragaman atau variasi-variasi dependen yang signifikan antar kelompok.

2.9 Korelasi Intraklas (*Intraclass Corelation*)

Korelasi interklas menunjukan proporsi keragaman yang dapat dijelaskan oleh struktur kelompok dalam populasi, yang dapat juga diinterpretasikan sebagai korelasi harapan antara dua unit yang dipilih secara acak yang berada dalam kelompok yang sama (Goldsten,1999). Dengan menggunakan model intersep atau pun model intersep acak korelasi interklas (ρ) dapat dituliskan sebagai berikut:

$$\rho = \frac{\sigma^2_{\delta 0}}{\sigma^2_{\delta 0} + \sigma^2_{\varepsilon}} \qquad 0 \leq \rho \leq 1 \qquad (2.26)$$

dengan $\sigma^2_{\delta 0}$ adalah ragam dari galat pada level-2 dan σ^2_{ε} adalah ragam dari galat pada level-1.

2.10 Keragaman model

Keragaman variabel terikat yang dijelaskan oleh variabel bebas dalam suatu model merupakan koefisien determinasi. Koefisien determinasi digunakan untuk mengukur sebereapa besar keragaman variabel terikat yang dijelaskab oleh model yang telah ditetapkan. Dalam model multilevel akan diperoleh koefisien determinasi lebih dari satu. Koefisien determinasi akan diperoleh pada masing-masing level (Hox, 2002).

Koefisien determinasi untuk level 1 dirumuskan sebagai berikut:

$$R_2^2 = 1 - \frac{\hat{\sigma}_{u0q}^2}{\hat{\sigma}_{u0}^2} \qquad (2.27)$$

Dengan

$\hat{\sigma}_{u0q}^2$ = penduga ragam dari galat pada level 1 dengan q variabel bebas.

$\hat{\sigma}_{u0}^2$ = penduga ragam dari galat pada level 1 tanpa variabel bebas.

Koefisien determinasi untuk level 1 dirumuskan sebagai berikut:

$$R_1^2 = 1 - \frac{\hat{\sigma}_{ep}^2}{\hat{\sigma}_{e0}^2} \qquad (2.28)$$

Dengan

$\hat{\sigma}_{ep}^2$ = penduga ragam dari galat pada level 2 dengan p variabel bebas.

$\hat{\sigma}_{e0}^2$ = penduga ragam dari galat pada level 1 tanpa variabel bebas.

Kisaran nilai koefisien determinasi anatara 0 hingga 1. Apabila nilai koefisien determinasi dikalikan 100% menunjukkan persentase informasi yang diberikan oleh model regresi yang dihasilkan. Semakin besar nilai koefisien determinasi yang didapatkan maka perolehan model regresi akan semakin baik (Kurniawan, 2010).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

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

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari Badan Pusat Statistik yang diunggah pada website https://www.bps.go.id/. Pada penelitian data yang digunakan sebanyak 6 variabel , yang diambil secara acak dari kabupaten/kota di Indonesia. Adapun variabel-variabel penelitian yang digunakan adalah sebagai berikut:

1. Variabel terikat:
   a. Angka partisipasi sekolah (Y)
2. Variabel bebas:
   a. Level 1 (Kabupaten/Kota) : Banyaknya jumlah SD di kabupaten (X_1), banyaknya jumlah SMP di kabupaten (X_2), banyaknya jumlah SMA di kabupaten (X_3), dan banyaknya jumlah SMK di kebupaten (X_4)
   b. Level 2 (Provinsi) : Kepadatan penduduk (Z_1) dan persentase kemiskinan (Z_2)

c.

3.3 Metode Penelitian

Penelitian ini dilaksanakan dengan studi pustaka yaitu dengan pengkajian secara teoritis dan praktik komputasi. Dalam melakukan penilitian ini, proses perhitungan dilakukan dengan menggunkan software R4.0.3 agar lebih efisien dan diperoleh hasil yang akurat.

Adapun langkahh-langkah yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Melakukan analisis deskriptif pada masing-masing variabel pada setiap level.
2. Melakukan uji asumsi pada data.
   a. Uji multikolinieritas dengan melihat nilai *Variance Inflation Factor* (VIF).
   b. Uji normalitas menggunakan Uji Kolmogorov-Smirnov.
3. Melakukan analisis regresi linier berganda pada data dengan memasukkan semua variabel bebas terhadap variabel terikat.
4. Melakukan uji simultan dan koefisien determinasi secara simultan.
5. Mengestimasi parameter model regresi 2-level dengan menggunakan metode *Restrited Maximum likelihood* dengan cara sebagai berikut :
   a. Memilih model tanpa variabel bebas.
   b. Memilih model intersep acak.
   c. Memilih model koefisien acak.
   d. Menghitung nilai koefisien korelasi interklas.
6. Melakukan uji signifikan parameter dalam menentukan model terbaik untuk mengetahui faktor yang berpengaruh pada angka partisipasi sekolah.
7. Menghitung nilai koefisien determinasi.
8. Menguraikan keragaman yang dapat dijelaskan oleh level 1 dan level 2 terhadap angka partisipasi sekolah di provinsi Indonesia.

V. KESIMPULAN

Berdasarkan hasil penelitian ini, diperoleh kesimpulan sebagai berikut:

1. Faktor-faktor yang mempengaruhi angka partisipasi sekolah di Indonesia tahun 2022 adalah banyaknya jumlah SD di kabupaten, banyaknya jumlah SMP di kabupaten, banyaknya jumlah SMA di kabupaten, dan banyaknya jumlah SMK di kebupaten.
2. Berdasarkan hasil analisis regresi multilevel diperoleh model regresi yaitu
   Angka partisipasi sekolah = $4{,}5715 - 0{,}0195$ jumlah SD $- 0{,}0069$ jumlah SMP $+ 0{,}0263$ jumlah SMA $+ 0{,}0099$ jumlah SMK.
   Artinya, jika variabel jumlah SD, jumlah SMP, jumlah SMA, jumlah SMK bernilai nol, maka angka partisipasi sekolah bernilai 4,5715, untuk setiap pertambahan jumlah SD sebesar satu-satuan akan menyebabkan menurunnya angka partisipasi sekolah sebesar -0,0195, variabel jumlah SMP akan menyebabkan menurunnya angka partisipasi sekolah sebesar -0,0069, untuk setiap pertambahan jumlah SMA sebesar satu-satuan akan menyebabkan kenaikan sebesar 0,0263, dan untuk setiap pertambahan jumlah SMk sebesar satu-satuan akan menyebabkan kenaikan sebesar 0,0099.
3. Keragaman yang dapat dijelaskan oleh struktur provinsi yaitu perbedaan provinsi adalah sebesar 18,2% artinya besarnya pengaruh perbedaan provinsi adalah sebesar 18,2%. Sedangkan keragaman yang dapat dijelaskan oleh struktur kabupaten adalah sebesar 88,9% artinya besarnya pengaruh variabel jumlah SD,jumlah SMP, jumlah SMA, jumlah SMK adalah sebesar 88,9%.

DAFTAR PUSTAKA

Bryck, A.S. & Raudenbush, S.W. 1987. *Applying the Hierarchical Linier Models to Measurement of Change Problems*. Psychological Bulletin. 101: 147-158.

Gujarati, D. 2006. *Dasar-Dasar Ekonometrika*. Erlangga, jakarta.

Goldstein, H. 1999. *Multilevel Statistikal Model*. Ed. ke-2. E-Book of Arnold, London.

Harlan, J. 2016. *Analisis Multilevel*. Gunadarma. Depok.

Hox, J.J. 2010. *Multilevel Analysis: Techniques and Applications*. Ed ke-2. Routledge, New York.

Jones, B.S. & Steenbergen, M.R. 1997. *Modeling Multilevel Data Structures.* American Journal of Political Science. 46(1): 218-237.

Kurniawan, D. 2008. *Regresi Linier*. R Development Core Team, Australia.

Kutner, M.H., Nachtsheim, C.J., & Neter, J. 2004. *Applied Linier Regression Models*. Ed ke-4. McGraw-Hill Companies, New York.

Myers, R.H. 1990. *Classical and Modern Regression with Application*. Ed ke-2. PWS Kent, Boston.

Nohe, D. A. 2013. *Biostatistika 1*. Halaman Moeka, Jakarta.

Rencher, A.C. & Schaalje, G.B. 2007. *Linier Model in Statistiks*. Ed ke-2. John Wiley & Sons, Inc., New Jersey.

Santoso, S. 2012. *Panduan Lengkap SPSS Versi 20*. PT Elex Media Komputindo, Jakarta.

Tantular, B. 2009. *Penerapan Model Regresi Linier Multilevel Pada Data Pendidikan dan Data Nilai Ujian*. Tesis. Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor, Bogor.

West, B.T., Welch, K.B. & Galechi, A.T. 2007. *Linier Mixed Models: A Practical Guide Using Stastical Software.* Chapman & Hall, Boca Raton.

Zulmi, R. 2011. *Pengaruh Luas Lahan, Tenaga Kerja, Penggunaan Benih, dan Pupuk Terhadap Produksi Padi di Jawa Tengah Tahun 1994-2008.* Skripsi. Jurusan Ilmu Ekonomi dan Studi Pembangunan Fakultas Ekonomi Universitas Diponegoro, Semarang.