PEMODELAN REGRESI 2-LEVEL DENGAN METODE *ITERATIVE GENERALIZED LEAST SQUARE*

(Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)

SKRIPSI

LA ODE MUHAMMAD IKLIL ANNAUFAL

H051171307

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

NOVEMBER 2022

PEMODELAN REGRESI 2-LEVEL DENGAN METODE *ITERATIVE GENERALIZED LEAST SQUARE*

(Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

LA ODE MUHAMMAD IKLIL ANNAUFAL

H051171307

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

NOVEMBER 2022

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

Pemodelan Regresi 2-Level Dengan Metode Iterative Generalized Least Square (Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 14 November 2022

METERAI TEMPEL 10000 SEPULUH RIBU RUPIAH 95ECDAKX169034302

La Ode Muhammad Iklil Annaufal

NIM H051171307

PEMODELAN REGRESI 2-LEVEL DENGAN METODE ITERATIVE GENERALIZED LEAST SQUARE

(Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)

Disetujui Oleh:

Pembimbing Utama	Pembimbing Pendamping
Sitti Sahriman, S.Si., M.Si.	**Drs. raupong, M.Si.**
NIP. 19881018 201504 2 002	**NIP. 19621015 198810 1 001**

Ketua Program Studi

Dr. Nurtiti Sunusi, S.Si., M.Si.

NIP. 19720117 199703 2 002

Pada 14 November 2022

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama	:	La Ode Muhammad Iklil Annaufal
NIM	:	H051171307
Program Studi	:	Statistika
Judul Skripsi	:	Pemodelan Regresi 2-Level Dengan Metode *Iterative Generalized Least Square* (Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Sitti Sahriman, S.Si., M.Si. (.............................)
2. Sekretaris : Drs. Raupong, M.Si. (.............................)
3. Anggota : Dra. Nasrah Sirajang, M.Si. (.............................)
4. Anggota : Siswanto, S.Si., M.Si. (.............................)

Ditetapkan di : Makassar

Tanggal : 14 November 2022

KATA PENGANTAR

Segala puji bagi Allah *Subhanahu Wata'ala*, sholawat dan salam semoga selalu dilimpahkan kepada baginda Rasulullah *Shallallahu'alaihi Wasallam* nabi yang paling mulia, pemimpin orang-orang takwa, dan kepada para keluarga dan sahabat beliau. Alhamdulillahirobbil'alamiin, berkat rahmat dan hidayah-NYA sehingga penulis dapat menyelesaikan skripsi dengan judul **"Pemodelan Regresi 2-Level Dengan Metode *Iterative Generalized Least Square* (Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)"** yang disusun sebagai salah satu syarat akademik untuk meraih gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin. Penulis berharap skripsi ini dapat memberikan tambahan pengetahuan bagi pembelajar statistika dan bagi pembaca secara umum.

Penulis menyadari dalam penulisan skripsi ini tidak lepas dari hambatan dan masalah namun dapat terselesaikan berkat dukungan dan bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan rasa terima kasih yang setulus-tulusnya serta penghargaan yang setinggi-tingginya untuk orang tua penulis, Ayahanda dan Ibunda tercinta **La Ode Dirham, S.Sos.** dan **Wa Ode Maria, S.E.**, yang telah membesarkan dan mendidik dengan penuh kesabaran, memberikan cinta dan limpahan kasih sayang, dukungan dan doa yang tulus tanpa henti kepada penulis. Rasa terima kasih juga kepada saudari tersayang **Wa Ode Shafira Salsabila Dirham**, serta **Keluarga Besar** atas doa, dukungan, semangat, dan bantuannya kepada penulis.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga penulis ucapkan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.,** selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin, M.Si.** selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si**. selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf Departemen Statistika yang telah membekali

ilmu dan kemudahan-kemudahan kepada Penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.

4. **Bapak Dr. La Podje Talangko, M.Si. (Alm.)** dan **Ibu Dra. Nasrah Sirajang, M.Si.** selaku Penasihat Akademik atas saran, nasehat, motivasi, dan dukungan yang telah diberikan kepada penulis selama menjadi mahasiswa dan meluangkan waktu untuk membimbing dan memberikan masukan dalam penulisan skripsi ini.
5. **Ibu Sitti Sahriman, S.Si., M.Si**. selaku Pembimbing Utama dan **Bapak Drs. Raupong., M.Si.** selaku Pembimbing Pertama yang telah ikhlas meluangkan waktu dan pemikirannya untuk memberikan arahan, pengetahuan, motivasi dan bimbingan ditengah kesibukan beliau serta menjadi tempat berkeluh kesah untuk penulis.
6. **Ibu Dra. Nasrah Sirajang, M.Si.** dan **Bapak Siswanto, S.Si., M.Si.** selaku Tim Penguji atas saran dan kritikan yang membangun dalam penyempurnaan penyusunan skripsi ini serta waktu yang telah diberikan kepada Penulis.
7. **Lalut**, **Fathoni, Islah, Shafwan, Aii**, **Eva, Haura, Putri, Fajri, Adam, Lia, Salsa, Fitri, Hana, Kety, Akil, Fadly, Mirna** yang telah menjadi teman terbaik sejak awal penulisan skripsi dan senantiasa mendengarkan curhatan, memberikan bantuan, dorongan, semangat, dan motivasi dalam setiap keadaan dalam penyusunan skripsi ini.
8. Sahabat terbaik **The Ganglang**,yaitu **Fajar, Radin, Jody, Bagas,** dan **Irfan** yang sampai saat ini masih menjadi sahabat dan setia mendengarkan keluh kesah penulis.
9. Teman-teman **Statistika 2017**, terima kasih atas kebersamaan, suka, dan duka selama menjalani pendidikan di Departemen Statistika. Penulis senang mengenal kalian semua.
10. Keluarga besar **DISKRIT 2017**, terima kasih telah memberikan pelajaran yang berharga dan arti kebersamaan selama ini kepada penulis. Pengalaman yang berharga telah penulis dapatkan dari teman-teman selama berproses.
11. **Keluarga Mahasiswa FMIPA Unhas** terkhusus anggota keluarga **Himatika FMIPA Unhas** dan **Himastat FMIPA Unhas**, terima kasih atas ilmu yang

mungkin tidak bisa didapatkan di proses perkuliahan dan telah menjadi keluarga selama penulis kuliah di Universitas Hasanuddin.

12. Teman-teman **KKN Perisai Anuang Sultra 4 Gelombang 104**. Terima kasih telah menjadi teman sekaligus keluarga selama sebulan lebih, semoga silaturahmi tetap terjalin
13. Semua pihak yang telah banyak berpartisipasi, baik secara langsung maupun tidak langsung yang tak sempat penulis sebutkan satu per satu. Terima kasih untuk segala bantuan dan dukungannya.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga skripsi ini dapat memberikan manfaat bagi pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Makassar, 14 November 2022

La Ode Muhammad Iklil Annaufal

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: La Ode Muhammad Iklil Annaufal
NIM	: H051171307
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-ekslusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

"Pemodelan Regresi 2-Level Dengan Metode Iterative Generalized Least Square (Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.
Dibuat di Makassar pada tanggal 14 November 2022.

Yang menyatakan

(La Ode Muhammad Iklil Annaufal)

ABSTRAK

Analisis regresi 2-level merupakan salah satu analisis yang menduga hubungan antara variabel yang diamati pada level-level yang berbeda dalam struktur data hirarki. Selain itu, regresi 2-level juga mengatasi pelanggaran asumsi homoskedastisitas dan non autokorelasi pada regresi klasik. Salah satu metode pendugaan parameter dalam regresi 2-level yang efisien adalah *Iterative Generalized Least Square* (IGLS). Metode IGLS menduga matriks varian kovarian (V) yang tidak diketahui pada metode *generalized least square*. Matriks varian kovarian diduga secara *iterative* dengan nilai toleransi kekonvergenan sebesar 10^{-1}. Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang mempengaruhi rata-rata nilai Ujian Nasional (UN) SMP di Provinsi Sulawesi Tenggara Tahun 2019 sebagai variabel respon. Variabel prediktor pada level 1 yaitu jumlah peserta UN (X_1), jumlah peserta didik (X_2), dan jumlah rombongan belajar (X_3). Sedangkan variabel prediktor level 2 yaitu angka partisipasi kasar (Z_1), rata-rata lama sekolah (Z_2), dan angka harapan lama sekolah (Z_3). Hasilnya adalah faktor-faktor yang berpengaruh signifikan pada taraf 5% yaitu jumlah peserta didik, interaksi antara jumlah peserta UN dengan angka harapan lama sekolah dan interaksi antara jumlah rombongan belajar dengan rata-rata lama sekolah. Kedua interaksi tersebut memberikan pengaruh positif masing-masing sebesar 0.004 dan 0.123. Sedangkan, jumlah peserta didik memberikan pengaruh negatif sebesar 0.009.

Kata Kunci: *Rata-Rata Nilai UN*, *Regresi 2-Level*, *Generalized Least Square*, *Iterative Generalized Least Square*

ABSTRACT

Two-level regression analysis is a type of analysis that predicts relationships between variables observed at different levels of a hierarchical data structure. Furthermore, two-level regression overcomes the violation of homoscedasticity and non-autocorrelation assumptions in traditional regression. One efficient method for parameter estimation in two-level regression is the iterative generalized least squares method (IGLS). The IGLS method predicts the unknown variance-covariance matrix (V) with generalized least squares. Covariance variance matrices were estimated iteratively with a convergence tolerance of 10^{-1}. The aim of this study was to determine the factors influencing the mean results of the National Examination (UN) of secondary schools in South East Sulawesi in 2019 as response variables. Level 1 predictors are the number of UN participants (X_1), the number of students (X_2), and the number of study groups (X_3). Level 2 predictors are gross enrollment (Z_1), average enrollment (Z_2), and expected enrollment (Z_3). Results are the factors that have a significant impact at the 5% level, namely the number of students, the interaction between the number of UN participants and expected schooling duration, and the interaction between the number of study groups and average schooling duration. school education. The two interactions have positive effects of 0.004 and 0.123 respectively. On the other hand, the number of students has a negative effect of 0.009.

Keywords: *Iterative Generalized Least Square, Generalized Least Square, 2-Level Regression, Average National Examination Score*

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Ujian Nasional (UN) merupakan salah satu sistem evaluasi standar pendidikan di Indonesia. Pada peraturan pemerintah Nomor 13 Tahun 2015 (Puspendik, 2015) hasil UN dapat digunakan sebagai salah satu bahan pertimbangan untuk pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upaya peningkatan mutu pendidikan. Adanya hasil UN diharapkan dapat memberikan gambaran peta mutu pendidikan pada tingkat nasional, provinsi, kabupaten/kota, dan sekolah. Rata-rata nilai UN yang diperoleh pada setiap sekolah dapat dipengaruhi oleh faktor internal dan eksternal. Faktor internal seperti kualitas siswa, kompetensi guru, sarana dan prasarana sekolah (Kemenristekdikti, 2019). Sedangkan faktor eksternal dapat dilihat dari kondisi daerah seperti, Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM), Angka Kelulusan (AL), Angka Melek Huruf (AMH), Rata-Rata Lama Sekolah (RLS) dan Angka Harapan Lama Lulus (AHLS) (Kemendikbud, 2013). Oleh karena itu, maka perlu dilakukan pemodelan hubungan rata-rata nilai UN dengan faktor-faktor yang diduga berpengaruh menggunakan analisis regresi.

Analisis regresi merupakan suatu model yang digunakan untuk menganalisis hubungan antara variabel respon (Y) dengan variabel prediktor (X). Hubungan yang dimaksud dalam hal ini, berupa hubungan sebab-akibat antar kedua variabel tersebut (Drapper & Smith, 1998). Perkembangan metode analisis regresi telah mengalami kemajuan dalam metode estimasi maupun variasi data yang digunakan. Analisis regresi dibedakan atas analisis regresi linier dan analisis regresi nonlinier. Salah satu pengembangan dari analisis regresi linier adalah model regresi multilevel. Pengembangan itu didasari karena data yang digunakan dalam analisis memiliki struktur bertingkat, berjenjang (*hierarchy*) atau berklaster, serta adanya hubungan antara variabel pada tingkat yang berbeda (Goldstein, 1995).

Data berstruktur hirarki adalah data yang terdiri dari kelompok-kelompok yang diobservasi bersarang atau terkelompokkan dalam kelompok level yang lebih tinggi

(Tantular, 2009). Jones dan Steenbergen (1997) menyebutkan bahwa masalah yang muncul akibat mengabaikan informasi kelompok adalah munculnya heteroskedastisitas dalam eror. Heteroskedastisitas mengindikasikan bahwa ada varian variabel pada model regresi yang tidak sama. Selain itu, masalah yang muncul pada data dengan struktur hirarki adalah adanya autokorelasi. Goldstein (1995) memperkenalkan model regresi multilevel yang bertujuan untuk mengatasi masalah pada data yang berstruktur hirarki. Oleh karena itu, untuk mengatasi masalah tersebut maka digunakan analisis multilevel (Tantular, 2016).

Model multilevel merupakan suatu pemodelan untuk menduga hubungan antara variabel yang diamati pada level-level yang berbeda dalam stuktur data berjenjang atau hirarki. Model yang paling sederhana dalam model multilevel adalah model dua level yang mana level 1 merupakan data individu atau sekolah dan level 2 adalah data kelompok atau kabupaten/kota dari individu-individu pada level 1 (West *et al*., 2007). Menurut Goldstein (1995) dalam Tantular (2011), salah satu penaksir parameter pada model multilevel adalah metode *Generalized Least Square* (GLS). Namun menurut Tantular (2011), penaksir pada metode GLS masih mengandung unsur parameter yang nilainya tidak diketahui yaitu pada matriks varian-kovarian. *Iterative Generalized Least Square* (IGLS) merupakan metode pendugaan dalam regresi multilevel yang mampu mengatasi masalah yang ada pada metode GLS. IGLS adalah metode yang digunakan untuk mengestimasi parameter dengan nilai matriks varian-kovarian yang baru, kemudian hasil estimasi parameter tetap tersebut digunakan untuk mengestimasi parameter acak. Selanjutnya, dilakukan estimasi berulang-ulang secara bergantian antara parameter tetap dan parameter acak sampai konvergen (Paramitha *et al*., 2016).

Dianti (2011) menggunakan *Restricted Maximum Likelihood* dalam mengestimasi parameter pada model regresi linier multilevel tidak mempertimbangkan masalah autokorelasi pada data. Pada penelitian yang dilakukan oleh Tantular (2011) membandingkan metode *Two Stage Least Square* (TSLS) dengan metode *Iterative Generalized Least Square* (IGLS) dalam menaksir parameter model multilevel pada beragam kondisi data. Metode IGLS memberikan hasil penaksiran yang lebih baik dibandingkan metode TSLS untuk berbagai ukuran sampel dan berbagai ukuran kelompok. Berdasarkan uraian diatas, maka penulis menyusunnya dalam sebuah penelitian dengan judul **“Pemodelan Regresi 2-Level dengan Metode *Iterative Generalized Least Square* (Studi Kasus: Rata-Rata Nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019)”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibahas, maka rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana estimasi parameter model regresi 2-level menggunakan metode *Iterative Generalized Least Square* pada data rata-rata nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019?
2. Bagaimana faktor-faktor yang mempengaruhi data rata-rata nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019 menggunakan model regresi 2-level dengan metode *Iterative Generalized Least Square*?

1.3 Batasan Masalah

Dalam penelitian ini, penulis memberikan batasan masalah sebagai berikut:

1. Data yang digunakan adalah data UN SMP di Provinsi Sulawesi Tenggara Tahun 2019. Adapun variabel-variabel yang digunakan diantaranya rata-rata nilai UN sebagai variabel respon level 1, serta jumlah peserta UN, jumlah peserta didik, dan jumlah rombongan belajar sebagai variabel prediktor level 1. Variabel angka partisipasi kasar, rata-rata lama sekolah, angka harapan lama sekolah sebagai variabel prediktor level 2.
2. Inisiasi awal yang digunakan adalah hasil estimasi dari metode *Ordinary Least Square* (OLS).

1.4 Tujuan Penelitian

Berdasarkan dari rumusan masalah tersebut, tujuan yang akan dicapai dalam penelitian ini adalah sebagai berikut:

1. Mendapatkan estimasi parameter model regresi 2-level menggunakan metode *Iterative Generalized Least Square* pada data rata-rata nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019.
2. Mendapatkan faktor-faktor yang mempengaruhi data rata-rata nilai UN SMP di Provinsi Sulawesi Tenggara Tahun 2019 menggunakan model regresi 2-level dengan metode *Iterative Generalized Least Square*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan wawasan keilmuan yang lebih khusus kepada penulis tentang estimasi parameter model regresi 2-level menggunakan *Iterative Generalized Least Square*.
2. Menambahan kepustakaan bagi pengguna ilmu statistika tentang cara memodelkan data menggunakan regresi 2-level dan pemahaman dalam analisis persentase rata-rata nilai UN SMP di Provinsi Sulawesi Tenggara dengan faktor-faktor yang mempengaruhinya.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier

Analisis regresi merupakan salah satu analisis statistik yang sering digunakan untuk menganalisis hubungan antara dua variabel atau lebih. Analisis regresi merupakan metode analisis yang dapat digunakan untuk menganalisis data dan mengambil kesimpulan yang bermakna tentang hubungan ketergantungan variabel terhadap variabel lainnya. Hubungan yang didapat pada umumnya dinyatakan dalam bentuk persamaan matematika yang menyatakan hubungan antara variabel prediktor dan variabel respon dalam bentuk persamaan sebagai berikut (Drapper & Smith, 1992):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \cdots + \beta_r X_{ri} + \varepsilon_i$$
$$= \beta_0 + \sum_{p=1}^{r} \beta_p X_{pi} + \varepsilon_i$$

dengan $i = 1,2, \dots, n$; $p = 1,2, \dots, r$

Y_i = variabel respon untuk pengamatan ke-i

β_0 = konstanta/intersep

β_p = koefisien dari variabel prediktor ke-p

X_{pi} = variabel prediktor ke-p pada pengamatan ke-i,

ε_i = eror pada pengamatan ke-i

2.2 Data Hirarki

Suatu penelitian terkadang dijumpai data populasi yang berstruktur hirarki atau data berjenjang. Data berstruktur hirarki merupakan data yang terjadi ketika individu-individu terkumpul dalam kelompok-kelompoknya. Data hirarki yang diperoleh dalam suatu penelitian merupakan data yang diperoleh melalui *multistage sampling* dari populasi berjenjang yang variabel-variabelnya dapat didefinisikan dari setiap level, dengan level yang lebih rendah tersarang pada level yang lebih tinggi. *Multistage sampling* merupakan pengambilan sampel yang dilakukan secara bertahap dengan menggunakan unit sampel yang lebih kecil pada setiap tahap. Data hirarki disebut juga sebagai data multilevel (Hox, 2002). Jika terdapat data dengan kondisi struktur hirarki atau data berjenjang yang memiliki 2 level maka:

a. Terdapat k buah kelompok dengan banyaknya pengamatan pada masing-masing kelompok sebesar $n_1, n_2, \dots, n_k$.

b. Variabel respon (Y) diamati hanya pada level 1.

c. Variabel prediktor yang memiliki 2 level dengan level 1 adalah variabel prediktor X_1 sampai X_r dan level 2 adalah variabel prediktor Z_1 sampai Z_l.

Adapun struktur dari data hirarki 2 level ditunjukkan pada Tabel 1 sebagai berikut:

Tabel 2.1 Gambaran Dari Struktur Data Hirarki 2 Level

j	i	Y_{ij}	X				Z			
			X_{1i}	X_{2i}	$\cdots$	X_{ri}	Z_1	Z_2	$\cdots$	Z_l
	1	y_{11}	$x_{1;1;1}$	$x_{2;1;1}$		$x_{r;1;1}$				
1	2	y_{21}	$x_{1;2;1}$	$x_{2;2;1}$		$x_{r;2;1}$	z_{11}	z_{21}		z_{l1}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$				
	n_1	$y_{n_1 1}$	$x_{1;n_1;1}$	$x_{2;n_1;1}$		$x_{r;n_1;1}$				
	1	y_{12}	$x_{1;1;2}$	$x_{2;1;2}$		$x_{r;1;2}$				
2	2	y_{22}	$x_{1;2;2}$	$x_{2;2;2}$		$x_{r;2;2}$	Z_{12}	Z_{22}		Z_{l2}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$				
	n_2	$y_{n_2 2}$	$x_{1;n_2;2}$	$x_{2;n_2;2}$		$x_{r;n_2;2}$				
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
	1	y_{1k}	$x_{1;1;k}$	$x_{2;1;k}$		$x_{r;1;k}$				
k	2	y_{2k}	$x_{1;2;k}$	$x_{2;2;k}$		$x_{r;2;k}$	z_{1k}	z_{2k}		z_{lk}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$				
	n_k	$y_{n_k k}$	$x_{1;n_k;k}$	$x_{2;n_k;k}$		$x_{r;n_k;k}$				

dengan $i = 1,2, \dots, n_j; p = 1,2, \dots, r; j = 1,2, \dots, k; q = 1,2, \dots, l$

Y_{ij} = variabel respon untuk pengamatan ke-i pada level 1 dalam kelompok ke-j pada level 2

X_{pi} = variabel prediktor ke-p di level 1 untuk pengamatan ke-i pada level 1

Z_{qj} = variabel prediktor ke-q di level 2 untuk kelompok ke-j pada level 2

2.3 Regresi 2-Level

Model multilevel merupakan suatu pemodelan untuk menduga hubungan antar variabel yang diamati pada level-level yang berbeda dalam stuktur data berjenjang atau hirarki. Model yang paling sederhana adalah model dua level yang mana level 1 adalah data individu dan level 2 adalah data kelompok (West *et al*., 2007). Berdasarkan struktur data pada Tabel 1 pemodelan multilevel dengan 2 level adalah sebagai berikut.

2.3.1 Model Level 1

Level 1 adalah model yang disusun tanpa memperhatikan pengaruh dari level 2. Pemodelan multilevel untuk tiap kelompok dapat ditulis sebagai berikut:

$$Y_{ij} = \beta_{0j}X_{0ij} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{pj}X_{pij} + \cdots + \beta_{rj}X_{rij} + \varepsilon_{ij} \tag{2.1}$$

$$= \beta_{0j} + \sum_{p=1}^{r} \beta_{pj}X_{pij} + \varepsilon_{ij} \tag{2.2}$$

dengan $i = 1,2, \dots n_j; j = 1,2, \dots, k; p = 1,2, \dots, r$

Y_{ij} = variabel respon untuk pengamatan ke-i pada level 1 dalam kelompok ke-j pada level 2

β_{0j} = intersep kelompok ke-j pada level 2

β_{pj} = intersep untuk kelompok ke-j pada level 2 yang berasal dari koefisien regresi untuk variabel prediktor ke-p pada level 1

X_{pij} = variabel prediktor ke-p di level 1 untuk pengamatan ke-i pada level 1 dalam kelompok ke-j pada level 2

ε_{ij} = eror untuk pengamatan ke-i pada level 1 dalam kelompok ke-j pada level 2

X_{0ij} = $[1 \quad 1 \quad 1 \quad \cdots \quad 1]'$

atau dapat dinotasikan dalam bentuk matriks adalah

$$\boldsymbol{Y} = \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ Y_{3j} \\ \vdots \\ Y_{n_k j} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & x_{1;1;j} & x_{2;1;j} & & x_{r;1;j} \\ 1 & x_{1;2;j} & x_{2;2;j} & \cdots & x_{r;2;j} \\ 1 & x_{1;3;j} & x_{2;3;j} & & x_{r;3;j} \\ & \vdots & & \ddots & \vdots \\ 1 & x_{1;n_1;j} & x_{2;n_2;j} & \cdots & x_{r;n_k;j} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \\ \beta_{2j} \\ \vdots \\ \beta_{rj} \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{3j} \\ \vdots \\ \varepsilon_{n_k j} \end{bmatrix}$$

dengan bentuk umum sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.3}$$

2.3.2 Model Level 2

Koefisien regresi pada model 1, β_{pj} memiliki nilai yang berbeda antar kelompok, variasi nilai β_{pj} akan dijelaskan dengan membentuk model level 2. Pembentukan model level 2 dilakukan untuk setiap parameter regresi sebagai respon ke-q dengan menggunakan variabel prediktor pada level 2. Bentuk pemodelan pada level 2 dapat ditulis sebagai berikut:

$$\begin{aligned}
\beta_{0j} &= \gamma_{00}Z_{0j} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \cdots + \gamma_{0l}Z_{lj} + u_{0j} \\
\beta_{1j} &= \gamma_{10}Z_{0j} + \gamma_{11}Z_{1j} + \gamma_{12}Z_{2j} + \cdots + \gamma_{1l}Z_{lj} + u_{1j} \\
\beta_{2j} &= \gamma_{20}Z_{0j} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \cdots + \gamma_{2l}Z_{lj} + u_{2j} \\
&\vdots \\
\beta_{rj} &= \gamma_{r0}\,Z_{0j} + \gamma_{r1}\,Z_{1j} + \gamma_{r2}\,Z_{2j} + \cdots + \gamma_{rl}\,Z_{lj} + u_{rj}
\end{aligned}$$

$$\beta_{pj} = \sum_{q=0}^{l} \gamma_{pq}Z_{qj} + u_{pj} \tag{2.4}$$

dengan $q = 1,2,\dots,l$

β_{pj} = intersep untuk kelompok ke-j pada level 2 yang berasal dari koefisien regresi untuk variabel prediktor ke-p pada level 1

γ_{0q} = koefisien regresi untuk variabel prediktor ke-q pada level 2 yang memprediksi koefisien regresi pada intersep persamaan regresi level 1

γ_{pq} = koefisien regresi untuk variabel prediktor ke-q pada level 2 yang memprediksi keragaman pada koefisien regresi variabel ke-p persamaan regresi level 1

Z_{qj} = variabel prediktor ke-q untuk kelompok ke-j pada level 2

u_{pj} = eror untuk kelompok ke-j pada variabel prediktor ke-p, diasumsikan menyebar $N(0, \sigma_\varepsilon^2)$

Z_{0j} = $[1 \quad 1 \quad 1 \quad \cdots \quad 1]'$

atau dinyatakan dalam bentuk matriks adalah

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{p1} \\ \beta_{p2} \\ \beta_{p3} \\ \vdots \\ \beta_{pk} \end{bmatrix}, \boldsymbol{Z} = \begin{bmatrix} 1 & z_{11} & Z_{21} & & Z_{l1} \\ 1 & z_{12} & Z_{22} & \cdots & Z_{l2} \\ 1 & z_{13} & Z_{23} & & Z_{l3} \\ & \vdots & & \ddots & \vdots \\ 1 & Z_{1r} & Z_{2r} & \cdots & Z_{lr} \end{bmatrix}, \boldsymbol{\gamma} = \begin{bmatrix} \gamma_{p1} \\ \gamma_{p2} \\ \gamma_{p3} \\ \vdots \\ \gamma_{pl} \end{bmatrix}, \boldsymbol{u} = \begin{bmatrix} u_{p1} \\ u_{p2} \\ u_{p3} \\ \vdots \\ u_{pr} \end{bmatrix}$$

dan bentuk umum sebagai berikut:

$$\boldsymbol{\beta} = \boldsymbol{Z\gamma} + \boldsymbol{u} \tag{2.5}$$

Jika persamaan (2.4) disubtitusikan pada persamaan (2.2) maka diperoleh model sebagai berikut:

$$\begin{aligned} Y_{ij} = & \left(\gamma_{00}Z_{0j} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \cdots + \gamma_{0l}Z_{lj} + u_{0j}\right)X_{0ij} \\ & +\left(\gamma_{10}Z_{0j} + \gamma_{11}Z_{1j} + \gamma_{12}Z_{2j} + \cdots + \gamma_{1l}Z_{lj} + u_{1j}\right)X_{1ij} \\ & +\left(\gamma_{20}Z_{0j} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \cdots + \gamma_{2l}Z_{lj} + u_{2j}\right)X_{2ij} \\ & + \cdots \\ & +\left(\gamma_{r0}Z_{0j} + \gamma_{r1}Z_{1j} + \gamma_{r2}Z_{2j} + \cdots + \gamma_{rl}Z_{lj} + u_{rj}\right)X_{rij} \\ & +\varepsilon_{ij} \end{aligned}$$

$$Y_{ij} = \sum_{p=0}^{r}\sum_{q=0}^{l}\gamma_{pq}Z_{qj}X_{pij} + \sum_{p=0}^{r}u_{pj}X_{pij} + \varepsilon_{ij} \tag{2.6}$$

$\sum_{p=0}^{r}\sum_{q=0}^{l}\gamma_{pq}Z_{qj}X_{pij}$ = efek tetap

$\sum_{p=0}^{r}u_{pj}X_{pij} + \varepsilon_{ij}$ = efek acak

2.4 *Intraclass Correlation*

Model regresi multilevel dapat menghasilkan nilai dugaan bagi *intraclass correlation*. Nilai dugaan tersebut dapat diperoleh menggunakan model tanpa variabel prediktor atau *null model*. *Null model* pada level 1 dapat dituliskan sebagai berikut:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \tag{2.7}$$

Sedangkan pada level 2 persamaannya sebagai berikut:

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{2.8}$$

Pada Persamaan (2.7) dan (2.8), $u_{0j} \sim N(0, \sigma_{\mathrm{u_0}}^2)$ adalah varian eror pada level 2 (kelompok) dan $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2)$ adalah varian eror pada level 1 (individu) sehingga dari *null model* inilah diperoleh *intraclass correlation* (ρ) dapat dirumuskan sebagai berikut:

$$\rho = \frac{\sigma_{u_o}^2}{\sigma_{u_o}^2 + \sigma_{\varepsilon}^2} \tag{2.9}$$

$\sigma^2_{u_o}$ merupakan varian eror level 2 dan σ^2_ε merupakan varian eror level 1. *Intraclass Correlation* (ICC) merupakan korelasi antara dua individu yang dipilih secara acak dalam kelompok yang sama. (Hox, 2010).

2.5 Metode *Generalized Least Square*

Metode *Generalized Least Square* (GLS) merupakan metode estimasi parameter yang digunakan untuk mengatasi adanya autokorelasi (Alaba *et al*., 2010). Asumsi-asumsi dalam model regresi linear $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ yaitu tidak adanya autokorelasi ($E(\boldsymbol{\varepsilon}) = 0$) dan homoskedasitas ($Var(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{I}$). Apabila asumsi-asumsi tersebut tidak terpenuhi maka metode OLS tidak tepat digunakan untuk mengestimasi parameter pada model regresi linear. Penduga parameter regresi dengan metode GLS adalah sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_{GLS} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y} \tag{2.10}$$

$\boldsymbol{V}$ merupakan matriks varian-kovarian berukuran $n_j \times n_j$.

Penaksiran yang diperoleh pada Persamaan (2.10) mengandung unsur parameter yang nilainya tidak diketahui yaitu pada matriks V. Oleh karena itu, untuk mendapatkan nilai taksirannya harus melalui proses iterasi sehingga metode penaksirnya disebut sebagai metode *Iterative Generalized Least Square* (IGLS) (Tantular, 2011).

2.6 Metode *Iterative Generalized Least Square*

Menurut Paramitha (2015), *Iterative Generalized Least Square* adalah langkah mengestimasi parameter dengan nilai matriks varian-kovarian yang baru, kemudian hasil estimasi parameter tetap tersebut digunakan untuk mengestimasi parameter acak. Langkah pertama dalam menaksir parameter dengan menggunakan metode IGLS adalah menaksir parameter tetap (level 1) β untuk suatu matriks varian-kovarian V yang diketahui dengan menggunakan metode GLS sehingga diperoleh persamaan sebagai berikut:

$$\widehat{\boldsymbol{\beta}}^{(t)} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y} \tag{2.11}$$

Prosedur iterasi memerlukan dugaan awal untuk melakukan estimasi terhadap β. Matriks varian-kovarian pada Persamaan (2.11) adalah $\boldsymbol{V} = \sigma^2_\varepsilon\boldsymbol{I}$ (diasumsikan $\sigma^2_{u_0}$= 0).

Artinya, pada dugaan awal nilai taksiran yang diperoleh sama seperti pada nilai taksiran menggunakan metode OLS.

$$\widehat{\boldsymbol{\beta}}^{(0)} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \tag{2.12}$$

Selanjutnya setelah hasil taksiran dari $\widehat{\boldsymbol{\beta}}^{(0)}$ diketahui, menghitung nilai-nilai taksiran untuk Y, yaitu $\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}^{(t)}$ sehingga dapat diketahui nilai eror yang dinyatakan dalam bentuk:

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}^{(t)} \tag{2.13}$$

Matriks varian-kovarian $\boldsymbol{V}$ ukuran $n \times n$ adalah *block diagonal* yang dinyatakan dalam bentuk sebagai berikut:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{V}_k \end{bmatrix} \tag{2.14}$$

k adalah banyaknya kelompok level 2 yang diobservasi, dan $\boldsymbol{V}_1, \boldsymbol{V}_2, \dots, \boldsymbol{V}_k$ adalah matriks varian-kovarian untuk masing-masing kelompok level 2, yang didefinisikan sebagai berikut:

$$\begin{gathered} \boldsymbol{V_1} = \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_1})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_1})} \\ \boldsymbol{V_2} = \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_2})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_2})} \\ \vdots \\ \boldsymbol{V_j} = \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_j})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_j})} \\ \vdots \\ \boldsymbol{V_k} = \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_k})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_k})} \end{gathered} \tag{2.15}$$

jika dijabarkan, matriks varian-kovarian untuk kelompok ke-j, $\boldsymbol{A_j}$, dijabarkan sebagai berikut:

$$\boldsymbol{V_j} = \begin{bmatrix} \sigma_{u_j}^2 + \sigma_{\varepsilon_j}^2 & \sigma_{u_j}^2 & \dots & \sigma_{u_j}^2 \\ \sigma_{u_j}^2 & \sigma_{u_j}^2 + \sigma_{\varepsilon_j}^2 & \dots & \sigma_{u_j}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u_j}^2 & \sigma_{u_j}^2 & \dots & \sigma_{u_j}^2 + \sigma_{\varepsilon_j}^2 \end{bmatrix} \tag{2.16}$$

$\boldsymbol{V}_j$ berukuran $n_j \times n_j$, dengan $\boldsymbol{I}_{(\boldsymbol{n_j})}$ adalah matriks identitas berukuran $n_j \times n_j$, $j = 1,2, \dots, 17$ dan $\boldsymbol{J}_{(\boldsymbol{n_j})}$ adalah matriks yang entri-entrinya berisi konstanta 1 ukuran

$n_j \times n_j$ dari (2.14) dan (2.15), matriks varian-kovarian untuk n observasi, yang mana $n = \sum_{j=1}^{k} n_j$, dinyatakan dalam bentuk:

$$\boldsymbol{V} = \begin{bmatrix} \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_1})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_1})} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_2})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_2})} & \dots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \sigma_u^2 \boldsymbol{J}_{(\boldsymbol{n_k})} + \sigma_\varepsilon^2 \boldsymbol{I}_{(\boldsymbol{n_k})} \end{bmatrix} \tag{2.17}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\boldsymbol{\varepsilon}}' \hat{\boldsymbol{\varepsilon}}}{\left(n_j - k - p + \lambda\right)} \tag{2.18}$$

$$\tilde{\sigma}_u^2 = \frac{\hat{\boldsymbol{u}}' \hat{\boldsymbol{u}} - (N - p)\hat{\sigma}_\varepsilon^2}{\left(N - trace\left[(\boldsymbol{X}'\boldsymbol{X})^{-1} \sum_{j=1}^{k} n_j^2 \overline{\boldsymbol{x}}'_{.j} \overline{\boldsymbol{x}}_{.j}\right]\right)} \tag{2.19}$$

$$\hat{\sigma}_u^2 = \begin{cases} \tilde{\sigma}_u^2, & \tilde{\sigma}_u^2 > 0 \\ 0, & lainnya \end{cases}$$

dengan

k = banyaknya kelompok level 2

p = banyaknya variabel prediktor level 1

$\hat{\boldsymbol{u}}' \hat{\boldsymbol{u}}$ = jumlah kuadrat eror dari regresi Y terhadap X

λ = jumlah kombinasi linier dari variabel X

$\overline{\boldsymbol{x}}_{.j}$ = vektor $(1 \times p)$ yang memiliki elemen ke-p $\bar{x}_{p.j}$, $p = 1,2, \dots, r$

N = $\sum_{j=1}^{k} n_j$

Penaksiran parameter acak pada level 2 dilakukan dengan cara yang sama seperti pada penaksiran parameter level 1. Pada level 2, β_p $(p = 1,2, \dots, r)$ digunakan sebagai variabel respon dan Z_q $(q = 1,2, \dots, l)$ sebagai variabel prediktor. Selanjutnya, matriks varian-kovarian yang digunakan pada level 2 berbeda dengan matriks varian-kovarian level 1. Matriks varian-kovarian pada level 2 disimbolkan sebagai $(\boldsymbol{V}^*)$ yang ditentukan melalui persamaan sebagai berikut:

$$\hat{\boldsymbol{\gamma}}^{(t)} = (\boldsymbol{Z}'(\boldsymbol{V}^*)^{-\mathbf{1}}\boldsymbol{Z})^{-\mathbf{1}}\boldsymbol{Z}'(\boldsymbol{V}^*)^{-\mathbf{1}}\hat{\boldsymbol{\beta}}^{(t)} \tag{2.20}$$

$$\boldsymbol{V}^* = \begin{bmatrix} \boldsymbol{V}_1^* & \boldsymbol{0} & \cdots & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_2^* & \cdots & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{V}_p^* & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \cdots & \boldsymbol{V}_r^* \end{bmatrix} \qquad (2.21)$$

$\boldsymbol{V}_p^*$ adalah matriks varian-kovarian pada level 2 dengan p menunjukkan banyaknya variabel respon pada level 2.

Setelah diperoleh taksiran dari parameter-parameter acak, ulangi langkah pengestimasian parameter tetap dengan nilai matriks varian-kovarian yang baru, kemudian hasil penaksiran parameter tetap digunakan untuk menaksir parameter acak. Selanjutnya dilakukan penaksiran berulang-ulang secara bergantian antara parameter tetap dan parameter acak sampai konvergen, yaitu nilai taksiran tidak lagi berfluktuasi pada iterasi-iterasi berikutnya.

2.7 Uji Signifikansi Parameter

Pengujian signifikansi parameter ini bertujuan untuk mengetahui ada tidaknya pengaruh dari variabel prediktor terhadap variabel respon. Pengujian signifikansi parameter yang digunakan adalah secara parsial menggunakan uji t, seperti dibahas berikut ini.

Pengujian secara parsial dilakukan apabila pengujian variabel prediktor secara simultan berpengaruh signifikan. Pengujian signifikansi parameter γ secara parsial menggunakan statistik uji t (Neter *et al*., 1997).

Hipotesis uji:

H_0 : $\gamma_{pq} = 0$ (tidak ada pengaruh signifikan antara variabel prediktor dengan variabel respon)

H_1 : $\gamma_{pq} \neq 0;\ p = 0,1,2,\dots,r\ ;\ q = 0,1,2,..,l,$ (ada pengaruh signifikan antara variabel prediktor dengan variabel respon)

Statistik uji:

$$t_{hit} = t_0 = \frac{\hat{\gamma}_{pq}}{se(\hat{\gamma}_{pq})}$$

dengan

$\hat{\gamma}_{pq}$ = nilai taksiran parameter γ_{pq}

$se(\hat{\gamma}_{pq})$ = standar eror dari γ_{pq}

Kriteria pengujian:

Tolak H_0 jika $|t_{hitung}| > t_{tabel}$. Hal ini menunjukkan bahwa ada pengaruh variabel prediktor ke-i terhadap variabel respon (Sembiring, 2003).

2.8 Pemilihan Model Terbaik

Menurut Tantular (2009), untuk memilih model regresi multilevel yang terbentuk, digunakan uji rasio *likelihood* dapat juga disebut sebaran *deviance* yaitu ukuran untuk menentukan cocok tidaknya suatu model. Perhitungan untuk pengujian ini adalah selisih nilai *deviance* antara dua model (*diff*), dapat dituliskan sebagai berikut:

Hipotesis uji:

H_0 : Model tidak signifikan

H_1 : Model signifikan

Taraf signifikansi: $\alpha = 5\%$

Statistik uji:

$$diff \quad = -2[\log(l_0) - \log(l_1)]$$

dengan

l_0 = nilai fungsi *likelihood* pada model sebelumnya

l_1 = nilai fungsi *likelihood* pada model yang diuji

Kriteria pengujian: Tolak H_0 apabila terdapat nilai $diff > \chi^2_{\alpha(db)}$.

2.9 Homoskedastisitas

Uji heteroskedastisitas digunakan untuk menguji apakah dalam model regresi terjadi ketidaksamaan varians dari eror satu pengamatan ke pengamatan yang lain. Jika varians dari eror satu pengamatan ke pengamatan yang lain sama, maka telah terjadi homoskedastisitas. Jika berbeda, maka telah terjadi heteroskedastisitas. Untuk mendeteksi adanya masalah heteroskedastisitas dapat menggunakan uji *Breuch-Pagan* (*BP*).

Hipotesis uji:

H_0 : $\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_2}^2 = \cdots = \sigma_{\varepsilon_n}^2 = \sigma^2$ (homoskedastisitas)

H_1 : Ada $\sigma_{\varepsilon_i}^2 \neq \sigma^2$; $i = 1,2,\ldots,n$ (tidak terjadi homoskedastisitas)

Statistik Uji:

$$BP = \left(\frac{1}{2}\right)\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{f} \sim \chi_{\alpha(db)}^2$$

dengan $\boldsymbol{f} = (\mathrm{f}_1, \mathrm{f}_2, \ldots, \mathrm{f}_\mathrm{n})^\mathrm{T}$ merupakan vektor berukuran, dengan elemen vektor

$$f_i = \frac{e_i^2}{\sigma^2} - 1$$

ε_i = eror untuk pengamatan ke-i dengan vektor berukuran $(n \times 1)$

$\boldsymbol{f}$ = vektor berukuran $(n \times 1)$

σ^2 = varian dari eror ε_i

$\boldsymbol{X}$ = matriks berukuran $n \times (p+1)$

p = banyaknya variabel prediktor

n = banyaknya wilayah pengamatan

Kriteria pengujian: Jika $BP > \chi_{\alpha(db)}^2$ maka tolak H_0

2.10 Autokorelasi

Menurut (Silalahi *et al*., 2014). Uji autokorelasi bertujuan untuk menguji apakah terdapat korelasi antara anggota serangkaian observasi yang diurutkan menurut waktu (seperti data *time series*) atau ruang (seperti data *cross-section*). Dalam penelitian ini digunakan uji *Durbin-Watson* (Uji *DW*). Dalam buku (Gujarati, 2006) nilai *DW* didapatkan dengan rumus:

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

Kriteria uji *DW* dapat dilihat dengan membandingkan nilai *DW* dengan nilai d_l dan nilai d_u yang diperoleh dari tabel *DW*. Pedoman pengambilan keputusan dalam uji autokorelasi sebagai berikut (Ghozali, 2009).

1. Bila nilai *Durbin-Watson* terletak lebih kecil dari pada batas bawah (d_l), maka koefisien autokorelasi lebih besar dari pada nol, berarti ada autokorelasi positif.

2. Bila nilai *Durbin-Watson* terletak antara batas atas (d_u) dan $(4 - d_u)$, maka koefisien autokorelasi sama dengan nol, berarti tidak ada autokorelasi.
3. Bila nilai *Durbin-Watson* lebih besar dari pada $(4 - d_l)$, maka koefisien autokorelasi lebih kecil dari pada nol, berarti ada autokorelasi negatif.
4. Bila nilai *Durbin-Watson* terletak antara $(4 - d_u)$ dan $(4 - d_l)$, maka hasilnya tidak dapat disimpulkan.

2.11 Ujian Nasional

Ujian Nasional (UN) merupakan salah satu sistem evaluasi standar pendidikan di Indonesia. Menurut kemendikbud (2010), Ujian Nasional merupakan salah satu upaya pemerintah dalam rangka memacu peningkatan mutu pendidikan. Ujian Nasional selain berfungsi untuk mengukur dan menilai pencapaian kompetensi lulusan dalam mata pelajaran tertentu, serta pemetaan mutu pendidikan pada tingkat pendidikan dasar dan menengah, juga berfungsi sebagai motivator bagi pihak-pihak terkait untuk bekerja lebih baik guna mencapai hasil ujian yang baik. Berbagai hasil penelitian menunjukkan bahwa dengan adanya Ujian Nasional, siswa terdorong untuk belajar lebih baik dan guru terdorong untuk mengajar lebih baik pula.

Dari perspektif yang lain, informasi tentang peta hasil Ujian Nasional dapat digunakan sebagai umpan balik bagi semua pihak terkait dalam rangka memperbaiki kinerjanya masing-masing. Oleh karena itu, peta hasil Ujian Nasional merupakan bahan informasi yang perlu dikaji secara mendalam oleh semua pihak dalam rangka memperbaiki pembelajaran dan mutu pendidikan secara berkelanjutan.

Rata-rata nilai UN yang diperoleh pada setiap sekolah dapat dipengaruhi oleh faktor internal dan eksternal. Faktor internal seperti kualitas siswa, kompetensi guru, sarana dan prasarana sekolah. Sedangkan faktor eksternal dapat dilihat dari kondisi daerah seperti, Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM), Angka Kelulusan (AL), Angka Melek Huruf (AMH), Rata-Rata Lama Sekolah (RLS) dan Angka Harapan Lama Lulus (AHLS) (Kemendikbud, 2013).

a. Angka Partisipasi Kasar

Angka Partisipasi Kasar (APK) adalah proporsi anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jenjang pendidikan tersebut (Kemendikbud, 2020). Secara umum, APK digunakan untuk mengukur keberhasilan program pembangunan pendidikan yang diselenggarankan dalam rangka memperluas kesempatan bagi penduduk untuk mengenyam pendidikan (Rahmatika, 2016). Pada penelitian ini APK yang digunakan adalah APK pada jenjang Sekolah Menengah Pertama (SMP). Rumus Perhitungan APK menurut Badan Pusat Statistik adalah sebagai berikut:

$$APK = \frac{jumlah\ penduduk\ sekolah}{jumlah\ penduduk\ usia\ tertentu} \times 100\%$$

Keterangan:

Tingkat SD : Kelompok usia 7-12 tahun

Tingkat SMP : Kelompok usia 13-15 tahun

Tingkat SMA : Kelompok usia 16-18 tahun

b. Rata-rata lama sekolah

Angka rata-rata lama sekolah (RLS) adalah rata-rata jumlah tahun yang dihabiskan oleh penduduk berusia 15 tahun ke atas untuk menempuh semua jenis pendidikan formal yang pernah dijalani. Populasi yang digunakan adalah penduduk berumur 15 tahun ke atas karena pada kenyataannya penduduk usia tersebut sudah ada yang berhenti sekolah. Batasan ini diperlukan agar angkanya lebih mencerminkan kondisi sebenarnya, mengingat penduduk yang berusia kurang dari 15 tahun masih dalam proses sekolah atau akan sekolah sehingga belum pantas untuk diperhitungkan rata-rata lama sekolahnya (Yamin *et al*., 2015). Rumus Perhitungan RLS menurut Badan Pusat Statistik adalah sebagai berikut:

$$RLS = \frac{1}{n} \times \sum_{i=1}^{n} x_i$$

dengan

RLS = rata-rata lama sekolah

x_i = lama sekolah penduduk ke-i yang berusia tertentu

N = jumlah penduduk usia tertentu

c. Angka harapan lama sekolah

Harapan Lama Sekolah (HLS) merupakan salah satu *output* yang dapat digunakan untuk memotret pemerataan pembangunan pendidikan di Indonesia. HLS merupakan indikator yang menggambarkan lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang. Angka ini diperoleh dengan cara membagi banyaknya partisipasi sekolah penduduk pada usia a pada tahun t dengan jumlah penduduk yang bersekolah pada usia a pada tahun t. Sebagai catatan indikator ini dianggap peka dalam menggambarkan variasi antar provinsi. Menurut *United Nation Development Programme* (*UNDP*), HLS dihitung dengan cara sebagai berikut (BPS Kabupaten Halmahera Selatan, 2016):

$$HLS_a^t = \sum_{i=a}^{t} \frac{E_i^t}{P_i^t}$$

dengan

HLS_a^t = harapan lama sekolah pada usia a dan pada tahun t

E_i^t = partisipasi sekolah penduduk usia i pada tahun t

P_i^t = populasi penduduk usia i yang bersekolah pada tahun t

i = usia $(a, a+1, \dots, n)$

Keterangan:

Tingkat SD : Kelompok usia 7-12 tahun

Tingkat SMP : Kelompok usia 13-15 tahun

TingkatSMA : Kelompok usia 16-18 tahun