

**PEMODELAN REGRESI 2-LEVEL DENGAN METODE
ITERATIVE GENERALIZED LEAST SQUARE (IGLS)
(Studi Kasus: Tingkat pendidikan Anak di Kabupaten Semarang)**



SKRIPSI

Oleh:

DYAN ANGGUN KRISMALA

NIM: J2E 009 040

**JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2013

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

(Studi Kasus: Tingkat Pendidikan Anak di Kabupaten Semarang)

Disusun Oleh :

DYAN ANGGUN KRISMALA

J2E 009 040

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains pada Jurusan Statistika Universitas Diponegoro

JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG

2013

HALAMAN PENGESAHAN I

Judul : Permodelan Regresi 2-Level Dengan Metode *Iterative Generalized Least Square* (IGLS)
(Studi Kasus: Tingkat Pendidikan Anak di Kabupaten Semarang)

Nama : Dyan Anggun Krismala

NIM : J2E 009 040

Jurusan : Statistika

Telah diujikan pada sidang Tugas Akhir tanggal 20 Desember 2013 dan dinyatakan lulus pada tanggal 27 Desember 2013

Semarang, 30 Desember 2013

Mengetahui,

Ketua Jurusan Statistika
Universitas Diponegoro



Lilik Dwi Ispryanti, M.Si
NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir
Ketua,

A handwritten signature in blue ink, appearing to be "H. Mustafid", written over a faint grid background.

Prof. Drs. H. Mustafid, M.Eng, Ph.D
NIP. 1955 05 28 1980 03 1 002

HALAMAN PENGESAHAN II

Judul : Pemodelan Regresi 2-Level Dengan Metode *Iterative Generalized Least Square* (IGLS).
(Studi Kasus: Tingkat Pendidikan Anak di Kabupaten Semarang).
Nama : Dyan Anggun Krismala
NIM : J2E 009 040
Jurusan : Statistika

telah diujikan pada sidang Tugas Akhir tanggal 20 Desember 2013

Semarang, 30 Desember 2013

Pembimbing I



Dri. Dwi Isprivanti, M.Si.

NIP. 195709141986032002

Pembimbing II



Moch. Abdul Mukid, S.Si, M.Si.

NIP. 197808172005011001

KATA PENGANTAR

Puji syukur kehadirat Allah SWT, karena atas rahmat, hidayah, kemudahan, dan segala limpahan nikmat-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “**Pemodelan Regresi 2-Level Dengan Metode *Iterative Generalized Least square (IGLS)* (Studi Kasus: Tingkat Pendidikan Anak di Kabupaten Semarang)**”.

Penulis menyadari tanpa adanya dukungan dan bantuan dari berbagai pihak maka penulisan tugas akhir ini tidak akan berjalan dengan baik. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Dra. Hj. Dwi Ispriyanti, M.Si selaku Ketua Jurusan Statistika FSM Universitas Diponegoro dan Dosen pembimbing I yang telah memberikan bimbingan serta pengarahan dalam penulisan Tugas Akhir ini.
2. Moch. Abdul Mukid, S.Si, M.Si selaku dosen pembimbing II yang telah memberikan bimbingan dan pengarahan dalam penulisan Tugas Akhir ini.
3. Bapak dan Ibu Dosen Jurusan Statistika, FSM Universitas Diponegoro yang telah memberikan ilmu yang sangat berguna.
4. Semua pihak yang telah memberikan bantuan, semangat dan doa sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini.

Penulis sadar bahwa Tugas Akhir ini masih jauh dari sempurna, untuk itu penulis mengharapkan kritik dan saran untuk perbaikan lebih lanjut.

Semarang, Desember 2013

Penulis

ABSTRAK

Dalam suatu penelitian terkadang diperoleh bentuk data yang berstruktur hirarki. Data hirarki adalah data yang diperoleh melalui *multistage sampling* dari populasi berjenjang dengan variabel bebas didefinisikan dalam setiap level dan variabel tak bebas didefinisikan dalam level terendah. Salah satu analisis yang dapat digunakan untuk data dengan struktur hirarki adalah analisis regresi multilevel. Analisis regresi multilevel yang paling sederhana adalah analisis regresi 2-level. Analisis regresi 2-level akan digunakan untuk membentuk model regresi tentang tingkat pendidikan anak di Kabupaten Semarang dimana anak (level-1) bersarang pada kecamatan (level-2) dengan faktor-faktor yang mempengaruhinya. Penaksiran parameter dalam model regresi 2-level dapat menggunakan beberapa metode, salah satunya adalah metode *Iterative Generalized Least Square* (IGLS). Dari hasil pembahasan menunjukkan bahwa faktor-faktor yang berpengaruh terhadap tingkat pendidikan anak di Kabupaten Semarang adalah pendidikan ibu, pendidikan ayah, dan prosentase keluarga petani. Keragaman tingkat pendidikan anak di Kabupaten Semarang lebih banyak disebabkan variasi antar anak dibandingkan variasi antar kecamatan.

Kata Kunci : Struktur hirarki, *multistage sampling*, Regresi multilevel, regresi 2-level

ABSTRACT

In a research, data was used often hierarchical structure. Hierarchical data is data obtained through multistage sampling from a population with independent variables can be defined within each level and dependent variable can be defined at the lowest level. One analysis that can be used for data with a hierarchical structure is a multilevel regression analysis. Multilevel regression analysis is the most simple regression analysis 2-levels. 2-level regression analysis will be used to construct a regression model the education level of children in Semarang where children (level-1) nested on the districts (level-2) with the factors that influence. Estimation of parameter in 2-level regression model can use some methods, one of them is *Iterative Generalized Least Square* (IGLS). From the results of the discussion indicates that the factors which affect the level of education of children in Semarang is the mother's education, father's education, and percentage of farm families. The diversity level of the education of children in Semarang caused more variation among children than the variation between districts.

Keywords : Hierarchical structure, multistage sampling, multilevel regression, 2-level regression

DAFTAR ISI

	halaman
HALAMAN JUDUL	i
PENGESAHAN I	ii
PENGESAHAN II.....	iii
KATA PENGANTAR.....	iv
ABSTRAK.....	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	x
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Tujuan Penulisan	3
BAB II TINJAUAN PUSTAKA	
2.1 Gambaran Umum	4
2.1.1 Definisi Pendidikan	4
2.1.2 Faktor Yang Mempengaruhi Pendidikan	5
2.2 Profil Kabupaten Semarang.....	6
2.3 Model Regresi Linier Berganda	7
2.4 Data Hirarki	11
2.5 Model Regresi 2-Level	13
2.6 Model Umum Regresi 2-Level dan Sub Modelnya	15
2.6.1 Model Intersep (<i>Intersept Only-Model</i>)	16
2.6.2 Model Intersep Acak (<i>Random Intercept Model</i>)	17
2.6.3 Model Kemiringan Acak (<i>Random Slope Model</i>)	20
2.7 Korelasi Intraklas (<i>Intra-class Corelation</i>)	21
2.8 <i>Generalized Least Square</i> (GLS)	22
2.9 Pemilihan dan Perbandingan Model	31
2.10 Eksplorasi Data	32
2.11 Uji Signifikansi Parameter	33

2.12 Asumsi Model Regresi 2-Level	34
2.12.1 Normalitas	34
2.12.2 Non Autokorelasi	35
2.12.3 Multikolinieritas	36
BAB III DATA DAN METODE	
3.1 Data.....	37
3.2 Metode Penelitian	38
3.3 Diagram Alir Analisis Data	39
BAB IV HASIL DAN PEMBAHASAN	
4.1 Aplikasi	40
4.2 Eksplorasi Data	42
4.3 Pemodelan Regresi 2-Level	43
4.3.1 Pemilihan Struktur Intersep Acak	44
4.3.2 Pemilihan Struktur Kemiringan Acak	47
4.3.3 Pemilihan Struktur Efek Tetap	53
4.3.4 Penyusunan Model	55
4.4 Uji Signifikansi Parameter	56
4.5 Model Akhir	58
4.6 Korelasi Intraklas (ICC).....	60
4.7 Uji Asumsi	61
BAB V KESIMPULAN	
Kesimpulan	63
DAFTAR PUSTAKA	64
LAMPIRAN	

DAFTAR TABEL

	halaman
Tabel 1 Statistik Deskriptif Tingkat Pendidikan di Seluruh Kecamatan	40
Tabel 2 Hasil Pendugaan Parameter Model M.1.1	44
Tabel 3 Hasil Pendugaan Parameter Model M.1.2	45
Tabel 4 Hasil Perbandingan Model M.1.1 dan M.1.2.....	46
Tabel 5 Hasil Pendugaan Parameter Model M.2.2	47
Tabel 6 Hasil Pendugaan Parameter Model M.2.3	48
Tabel 7 Hasil Pendugaan Parameter Model M.2.4	49
Tabel 8 Hasil Pendugaan Parameter Model M.2.5	50
Tabel 9 Hasil Perbandingan Model Kemiringan Acak	50
Tabel 10 Hasil Pendugaan Parameter Model M.3.1	51
Tabel 11 Hasil Pendugaan Parameter Model M.3.2	53
Tabel 12 Hasil Perbandingan Model M.3.1 dan M.3.2.....	53
Tabel 13 Nilai Pendugaan Parameter Model Regresi 2-Level (M.4.1).....	55
Tabel 14 Nilai Uji Signifikansi Parameter	56
Tabel 15 Hasil Pendugaan Parameter pada Model Akhir	57
Tabel 16 Hasil Pendugaan Komponen Acak (Varian dan Kovarian) Model Akhir.....	59

DAFTAR LAMPIRAN

	halaman
Lampiran 1 Data Pendidikan Anak	68
Lampiran 2 Statistik Deskriptif	78
Lampiran 3 Plot Interaksi Variabel Bebas Antar Level	79
Lampiran 4 Model M.1.1	82
Lampiran 5 Model M.1.2	85
Lampiran 6 Model M.2.2	87
Lampiran 7 Model M.2.3	89
Lampiran 8 Model M.2.4	91
Lampiran 9 Model M.2.5	93
Lampiran 10 Model M.3.1	95
Lampiran 11 Model M.3.2	97
Lampiran 12 Model M.4.1	99
Lampiran 13 Model Akhir	101
Lampiran 14 Uji Asumsi Normalitas	103
Lampiran 15 Uji Asumsi Non Autokorelasi	104
Lampiran 16 Uji Asumsi Multikolinieritas	105
Lampiran 17 Tabel Distribusi Nilai Chi-Square	106

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam suatu penelitian terkadang diperoleh struktur data yang merupakan data hirarki (*hierarchical*) atau data berjenjang. Data yang terstruktur hirarki merupakan data yang timbul karena individu-individu terkumpul dalam kelompok-kelompoknya, dimana individu-individu dalam kelompok yang sama memiliki karakteristik yang cenderung sama. Struktur hirarki mengindikasikan bahwa data yang dianalisis berasal dari beberapa level, dimana level yang lebih rendah tersarang pada level yang lebih tinggi. Analisis dari data seperti ini disebut sebagai analisis data multilevel.

Menurut Hox (1995), penelitian multilevel menyangkut tentang populasi dengan struktur hirarki atau berjenjang, dimana sampel dari populasi dapat digambarkan sebagai sampel multi-tahap (*multistage sample*). Variabel-variabel dapat didefinisikan dari setiap tingkat dari jenjang. Sebagian variabel ini dapat diukur secara langsung dari tingkat aslinya. Bentuk data hirarki dapat dijumpai dalam kehidupan sehari-hari di berbagai bidang penelitian. Salah satu contoh tentang data hirarki yaitu dalam bidang pendidikan.

Perkembangan dunia pendidikan yang sangat maju tidak dipungkiri menjadi kebutuhan yang sangat primer dalam kehidupan. Selain itu pendidikan merupakan salah satu unsur yang mempunyai peranan penting dalam meningkatkan mutu sumber daya manusia yang baik dan berkualitas, karena pendidikan merupakan dasar untuk mengasah ketrampilan seseorang dalam

berfikir. Banyak orang beranggapan bahwa semakin baik pendidikan maka semakin baik pula kualitas sumber daya manusianya

Terbentuknya sumber daya manusia yang berkualitas menunjukkan keberhasilan proses pendidikan. Keberhasilan pendidikan seseorang dipengaruhi oleh banyak hal diantaranya faktor orang itu sendiri (faktor diri sendiri), faktor keluarga dan lingkungan masyarakat yang saling berkorelasi. Misalnya kecamatan (faktor lingkungan), dalam arti bahwa pendidikan seseorang dipengaruhi kecamatan tempat mereka tinggal, dan sifat-sifat dari kecamatan tersebut terbentuk oleh individu-individu yang berdomisili di kecamatan tersebut. Secara umum individu dan kecamatan merupakan suatu sistem hirarki. Bentuk data seperti contoh tersebut dapat dianalisis dengan menggunakan analisis regresi multilevel dua tingkatan atau analisis regresi 2-level dimana faktor individu (diri sendiri) dan faktor keluarga diukur pada level terendah atau level-1, sedangkan untuk faktor ketiga yaitu faktor lingkungan (kecamatan) diukur pada level yang lebih tinggi yaitu level kelompok (level-2).

Dalam beberapa sumber penelitian model regresi multilevel dikenal dengan berbagai nama, antara lain model koefisien acak, model komponen variansi, dan *Hierarchical Linear Model* (HLM), dengan beberapa macam metode penaksiran yang digunakan dalam analisis, diantaranya metode *Iterative Generalized Least Square* (IGLS), *Two Stage Least Square* (2SLS), *Generalized Estimating Equation* (GEE) (Liang and Zeger, 1986), dan *Restricted Maximum Likelihood* (REML) (Goldstein, 1995). Metode IGLS banyak digunakan oleh peneliti karena dalam metode ini model yang digunakan merupakan model yang telah disubstitusikan.

Dalam penulisan tugas akhir ini, akan dibahas mengenai pemodelan regresi multilevel pada data yang berstruktur hirarki dan dibatasi hanya pada data hirarki dua tingkatan (2-level) serta aplikasinya pada data tingkat pendidikan anak di Kabupaten Semarang. Metode yang digunakan dalam model regresi 2-level yaitu dengan metode *Iterative Generalized Least Square (IGLS)* serta aplikasinya untuk mengetahui hubungan fungsional antar tingkat pendidikan anak dengan mempertimbangkan jenis kelamin, pendidikan ibu, pendidikan ayah dan tempat tinggal sebagai variabel bebas level-1, serta banyaknya SD, banyak SMP, banyak SMA dan prosentase keluarga petani di kecamatan sebagai variabel bebas level-2.

1.2 Tujuan Penulisan

Berdasarkan dari uraian di atas maka penulisan tugas akhir ini bertujuan untuk :

1. Mengestimasi parameter model regresi 2-level dengan menggunakan metode IGLS.
2. Menentukan model terbaik pada model regresi 2-level.
3. Menerapkan model regresi 2-level untuk mengetahui faktor-faktor yang mempengaruhi pendidikan anak di Kabupaten Semarang.