

**ESTIMASI PARAMETER
PADA MODEL REGRESI LINIER MULTILEVEL
DENGAN METODE *RESTRICTED MAXIMUM LIKELIHOOD* (REML)**



SKRIPSI

Disusun Oleh :

ISNI RAKHMI DIANTI

J2E 006 018

**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS DIPONEGORO
SEMARANG**

2011

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas segala limpahan rahmat, karunia serta hidayah-Nya sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul “**Estimasi Parameter pada Model Regresi Linier Multilevel dengan Metode *Restricted Maximum Likelihood (REML)***”. Skripsi ini dapat tersusun atas bantuan berbagai pihak, untuk itu penulis mengucapkan terima kasih kepada :

1. Ibu Dr. Widowati, M.Si selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Diponegoro.
2. Ibu Dra. Suparti, M.Si selaku Ketua Program Studi Statistika Fakultas MIPA Universitas Diponegoro.
3. Ibu Yuciana Wilandari, S.Si., M.Si dan Bapak Moch. Abdul Mukid, S.Si., M.Si selaku Dosen Pembimbing I dan II yang telah berkenan meluangkan waktu untuk memberikan masukan dan bimbingan kepada penulis sehingga skripsi ini bisa terselesaikan.
4. Ibu Diah Safitri, M.Si selaku dosen wali yang telah memberikan pengarahan selama masa perkuliahan.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, untuk itu penulis mengharapkan kritik dan saran yang membangun guna perbaikan lebih lanjut. Akhir kata semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, Juni 2011

Penulis

ABSTRAK

Individu dan lingkungan dapat membentuk suatu sistem yang berstruktur hierarki. Permasalahan khusus pada data hierarki adalah menganalisis data pada level yang berbeda. Model multilevel dapat menjadi solusi untuk permasalahan ini, dimana model multilevel merupakan bagian dari model linier campuran. Estimasi parameter pada model multilevel menggunakan metode *Restricted Maximum Likelihood* (REML) untuk mengatasi masalah bias pada parameternya. Untuk mendapatkan model multilevel terbaik digunakan uji rasio likelihood, sedangkan untuk menguji koefisien masing-masing parameternya menggunakan uji Wald.

Kata kunci : data hierarki, model linier campuran, model multilevel, *Restricted Maximum Likelihood*

ABSTRACT

Individual and its environment form a system called hierarchy system. Special issue of hierarchy data is about analysis of data at different levels of observation. Multilevel models can be used to solve this problem, where multilevel model is subset of linear mixed model. Estimating parameters used Restricted Maximum Likelihood method to solve bias in its parameters. In order to obtain the best multilevel model, its analysis used likelihood ratio tests, whereas to tests each parameter it used Wald test.

Keywords : hierarchy data, linear mixed model, multilevel model, Restricted Maximum Likelihood

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN I.....	ii
HALAMAN PENGESAHAN II.....	iii
KATA PENGANTAR.....	iv
ABSTRAK.....	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR SIMBOL.....	xi
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Pembatasan Masalah.....	3
1.4 Tujuan Penulisan	3
1.5 Sistematika Penulisan	4
BAB II KONSEP DASAR	5
2.1 Matriks	5
2.1.1 Notasi Matriks	5
2.1.2 Transpose Suatu Matriks.....	6
2.1.3 Matriks Khusus	6
2.1.4 Determinan Matriks	8
2.1.5 Invers Matriks.....	11

2.2	Metode-metode Estimasi Parameter	13
2.2.1	<i>Ordinary Least Square</i> (OLS)	13
2.2.2	<i>Weighted Least Square</i> (WLS)	14
2.2.3	<i>Generalized Least Square</i> (GLS).....	15
2.2.4	Estimasi Maximum Likelihood	17
2.2.5	<i>Restricted Maximum Likelihood</i> (REML).....	18
2.3	Analisis Regresi Linier	21
2.3.1	Model Regresi Linier Sederhana	21
2.3.2	Model Regresi Linier Berganda.....	24
BAB III MODEL REGRESI LINIER MULTILEVEL		27
3.1	Model Linier Campuran (<i>Linier Mixed Model</i>)	27
3.1.1	Tipe dan Struktur Data	28
3.1.2	Level pada Data	28
3.1.3	Efek dalam Model Linier Campuran	29
3.1.4	Spesifikasi Matriks Umum.....	30
3.1.5	Struktur Kovarian pada Matriks D	33
3.1.6	Struktur Kovarian pada Matriks R_j	34
3.2	Model Linier Marjinal dari Sebuah Model Linier Campuran.....	34
3.3	Estimasi Parameter Model Linier Campuran	35
3.4	Model Dasar Regresi Multilevel	39
3.5	Model Umum Regresi Multilevel dan Sub Modelnya.....	42
3.6	Estimasi Parameter Model Multilevel	43
3.7	Galat Baku (<i>Standar Errors</i>) Estimator Model Multilevel.....	48
3.8	Pengujian Hipotesis	48

3.8.1 Uji Rasio Likelihood (<i>Likelihood Ratio Test</i> (LRTs)).....	48
3.8.2 Uji Keberartian Masing-masing Parameter	50
3.9 Asumsi Model Regresi Linier Multilevel	51
3.10 Diagram Alur Penyelesaian Masalah.....	56
3.11 Contoh Penerapan.....	57
3.11.1 Spesifikasi Model.....	58
3.11.2 Uji Hipotesis	62
3.11.3 Model Akhir	67
3.11.4 Uji Asumsi Model Akhir Regresi Linier Multilevel.....	70
BAB IV PENUTUP	72
DAFTAR PUSTAKA	74
LAMPIRAN	75

DAFTAR SIMBOL

- i : Individu dalam taraf level 2 ke- j , dimana $i = 1, 2, \dots, n_j$
- j : Taraf level 2, dimana $j = 1, 2, \dots, J$
- Y_j : Vektor variabel tak bebas pada subjek ke- j
- X_j : Matriks variabel bebas untuk parameter tetap dari p kovariat
- β : Vektor dari koefisien regresi yang belum diketahui nilainya
(parameter efek tetap)
- $\hat{\beta}$: Nilai estimasi dari parameter β
- Z_j : Matriks variabel bebas untuk parameter tetap dari q kovariat
- δ_j : Vektor efek acak dalam matriks Z_j
- e_j : Vektor residual pengamatan untuk subjek ke- j
- D : Matriks varian-kovarian untuk δ_j
- R_i : Matriks varian-kovarian untuk e_j
- θ_D : Vektor yang berisi parameter varian-kovarian pada D
- θ_R : Vektor yang berisi parameter varian-kovarian pada R_j
- \hat{Y}_j : Nilai estimasi dari Y_j
- W_i : Matriks tertimbang pada WLS dan GLS
- V_j : Matriks varian-kovarian pada model linier marjinal
- θ : Vektor dari koefisien regresi yang belum diketahui nilainya
(parameter kovariansi)
- Y_{ij} : Variabel tak bebas pada level 1
- X_{ij} : Variabel bebas pada level 1

- e_{ij} : Residual untuk level 1
- Z_j : Variabel bebas pada level 2
- γ : Efek tetap pada level 2
- δ : Residual pada level 2
- u_{ij} : Residual total model multilevel
- λ : Rasio likelihood

DAFTAR TABEL

Tabel 3.1 Hasil Analisis dengan Metode REML untuk Data *Rat Pup*..... 60

Tabel 3.2 Hasil Analisis Model Akhir dengan Metode REML untuk Data *Rat Pup* 68

DAFTAR LAMPIRAN

Lampiran 1	Data <i>Rat Pup</i>	75
Lampiran 2	Sintaks Statistik Deskriptif Data <i>Rat Pup</i> dengan SAS 9.1.3	84
Lampiran 3	Output Statistik Deskriptif Data <i>Rat Pup</i> dengan SAS 9.1.3	85
Lampiran 4	Sintaks Model 3.1 Data <i>Rat Pup</i> dengan SAS 9.1.3.....	86
Lampiran 5	Output Model 3.1 Data <i>Rat Pup</i> dengan SAS 9.1.3	87
Lampiran 6	Sintaks Model 3.1 A Data <i>Rat Pup</i> dengan SAS 9.1.3.....	89
Lampiran 7	Output Model 3.1 A Data <i>Rat Pup</i> dengan SAS 9.1.3	90
Lampiran 8	Sintaks Model 3.2 A Data <i>Rat Pup</i> dengan SAS 9.1.3.....	92
Lampiran 9	Output Model 3.2 A Data <i>Rat Pup</i> dengan SAS 9.1.3	93
Lampiran 10	Sintaks Model 3.2 B Data <i>Rat Pup</i> dengan SAS 9.1.3.....	95
Lampiran 11	Output Model 3.2 B Data <i>Rat Pup</i> dengan SAS 9.1.3.....	96
Lampiran 12	Sintaks Model 3.3 Menggunakan Metode <i>Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3.....	98
Lampiran 13	Output Model 3.3 Menggunakan Metode <i>Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3.....	99
Lampiran 14	Sintaks Model 3.3 A Menggunakan Metode <i>Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3.....	101
Lampiran 15	Output Model 3.3 A Menggunakan Metode <i>Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3.....	102
Lampiran 16	Sintaks Model 3.3 Menggunakan Metode <i>Restricted Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3	104
Lampiran 17	Output Model 3.3 Menggunakan Metode <i>Restricted Maximum Likelihood</i> Data <i>Rat Pup</i> dengan SAS 9.1.3	105

Lampiran 18	Sintaks Uji Hipotesis 1 Data <i>Rat Pup</i> dengan SAS 9.1.3	107
Lampiran 19	Output Uji Hipotesis 1 Data <i>Rat Pup</i> dengan SAS 9.1.3.....	108
Lampiran 20	Sintaks Uji Hipotesis 2 Data <i>Rat Pup</i> dengan SAS 9.1.3	109
Lampiran 21	Sintaks Uji Hipotesis 3 Data <i>Rat Pup</i> dengan SAS 9.1.3	110
Lampiran 22	Sintaks Uji Hipotesis 4 Data <i>Rat Pup</i> dengan SAS 9.1.3	111
Lampiran 23	Output Uji Asumsi Normalitas Model Akhir Regresi Linier Multilevel dengan SPSS 15.0	112
Lampiran 24	Output Uji Asumsi Autokorelasi Model Akhir Regresi Linier Multilevel dengan SPSS 15.0	113
Lampiran 25	Output Uji Asumsi Multikolinieritas Model Akhir Regresi Linier Multilevel dengan SPSS 15.0	114

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penelitian sosial seringkali terkonsentrasi pada masalah bagaimana menelusuri hubungan antara individu dengan lingkungannya. Konsep umum bahwa individu berkorelasi dengan lingkungannya adalah bahwa suatu individu dipengaruhi oleh lingkungan dimana mereka berada, dan sifat-sifat dari lingkungan tersebut terbentuk dari individu-individu yang ada di dalamnya. Secara umum, individu dan lingkungannya merupakan sistem hierarki (bertingkat). Secara alamiah, suatu sistem dapat diamati pada tingkat yang berbeda-beda, dan sebagai hasilnya akan terbentuk variabel-variabel yang menggambarkan individu dengan variabel yang menggambarkan lingkungannya. Penelitian semacam ini disebut dengan penelitian multilevel. Model multilevel dapat digunakan untuk menganalisis data berstruktur hierarki.

Salah satu contoh penelitian regresi multilevel adalah penelitian tentang anak tikus (*rat pup*) dimana induknya memperoleh satu dari tiga perlakuan (*treatment*) dengan dosis yang berbeda, yaitu tinggi (*high*), rendah (*low*) dan dosis tertentu (*control*). Penelitian ini bertujuan untuk membandingkan berat lahir anak tikus yang lahir dari induk yang mendapatkan perlakuan dosis tinggi (*high treatment*) atau rendah (*low treatment*) dengan yang mendapat perlakuan dengan dosis tertentu (*control*). Masing-masing anak tikus dari induk yang sama (*litter*) tersebar secara acak ke dalam level yang lebih spesifik dari perlakuan (*treatment*), dan anak tikus (*rat pup*) tersarang di dalam *litter* (sejumlah anak tikus dari induk yang sama). Berat lahir dari anak-anak tikus (*rat pup*) dari satu induk yang sama (*litter*) mungkin memiliki korelasi karena anak-anak tikus itu berada dalam lingkungan induk yang sama. Secara umum anak tikus (*rat pup*) dan sejumlah anak tikus dari induk yang sama

(*litter*) merupakan suatu sistem hierarki atau dapat dikatakan sebagai suatu struktur tersarang (*nested*). Secara alamiah suatu sistem dapat diamati pada tingkat yang berbeda-beda, dan sebagai hasilnya akan terbentuk variabel-variabel dalam setiap tingkat. Hal ini menuntun penelitian tersebut kepada interaksi antar variabel-variabel yang menggambarkan anak tikus (*rat pup*) dengan variabel-variabel yang menggambarkan sejumlah anak tikus dari induk yang sama (*litter*). Penelitian seperti ini menghasilkan data yang berstruktur hierarki.

(West *et al*, 2007)

Model multilevel diperkenalkan oleh Goldstein (1995) yang menyebutkan bahwa model multilevel dapat mengatasi masalah yang muncul dari data yang diperoleh dari survey yang dilakukan menggunakan random sampling bertahap (*multistage random sampling*) atau data dengan struktur hierarki, seperti heterogenitas dalam residual dan adanya hubungan antar variabel pada tingkat yang berbeda. Dalam model multilevel, tingkatan dalam struktur hierarki didefinisikan sebagai level. Tingkat yang paling rendah disebut level 1 dan tingkat yang lebih tinggi disebut level 2. Untuk mengestimasi parameter-parameter dalam model multilevel terdapat beberapa metode estimasi parameter, namun para peneliti menyarankan untuk menggunakan metode *Restricted Maximum Likelihood* (REML) untuk mengatasi masalah ketakbiasan parameternya.

1.2 Perumusan Masalah

Berdasarkan latar belakang tersebut maka permasalahannya adalah bagaimana membentuk model regresi linier multilevel dan melakukan estimasi terhadap parameternya pada data kluster yang berstruktur hierarki menggunakan metode *Restricted Maximum Likelihood* (REML).

1.3 Pembatasan Masalah

Model regresi linier multilevel cukup luas penerapannya. Dalam tugas akhir ini pembahasan hanya dilakukan pada pembentukan model regresi dua level serta estimasi parameter-parameternya.

1.4 Tujuan Penulisan

Tujuan penulisan tugas akhir ini adalah :

1. Membentuk model regresi linier multilevel.
2. Melakukan estimasi parameter regresi linier multilevel.
3. Menentukan model regresi linier multilevel terbaik.
4. Menerapkan model regresi linier multilevel untuk faktor-faktor yang mempengaruhi berat lahir anak tikus.

1.5 Sistematika Penulisan

Sistematika penulisan tugas akhir ini adalah sebagai berikut : Bab I merupakan pendahuluan yang berisi garis besar permasalahan yang akan dibahas dan diselesaikan sesuai dengan tujuan yang telah dirumuskan. Bab II berisi teori-teori yang mendukung dan mendasari penulisan ini yaitu tentang matriks, metode-metode estimasi parameter serta analisis regresi. Bab III merupakan bagian utama dari penulisan tugas akhir ini, mengenai estimasi parameter model regresi linier multilevel. Pembahasan difokuskan pada pembentukan model regresi linier multilevel serta estimasi parameter pada model regresi linier multilevel. Beberapa bagian disertakan dengan contoh penerapan terhadap data *Rat Pup* beserta analisisnya. Bab IV berisi kesimpulan secara umum dari keseluruhan pembahasan.

