

**KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL)  
DALAM PENDUGAAN PARAMETER MODEL REGRESI  
LOGISTIK MULTILEVEL**

**SKRIPSI**

oleh:

**AHMAD TANTOWI JAUHARI**

**105090501111014**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
2014**

**KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL)  
DALAM PENDUGAAN PARAMETER MODEL REGRESI  
LOGISTIK MULTILEVEL**

**SKRIPSI**

**Sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains  
dalam bidang Statistika**

oleh:  
**AHMAD TANTOWI JAUHARI**  
**105090501111014**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
2014**

**LEMBAR PENGESAHAN SKRIPSI**

**KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL)  
DALAM PENDUGAAN PARAMETER MODEL REGRESI  
LOGISTIK MULTILEVEL**

oleh:  
**AHMAD TANTOWI JAUHARI**  
**105090501111014**

Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 17 Juli 2014  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika

**Dosen Pembimbing**

**Samingun Handoyo, S. Si., M. Cs.**  
**NIP. 197304151998021002**

**Mengetahui,**  
**Ketua Jurusan Matematika**  
**Fakultas MIPA**  
**Universitas Brawijaya**

**Dr. Abdul Rouf Alghofari, M.Sc**  
**NIP. 1967090719920131001**

## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : AHMAD TANTOWI JAUHARI  
NIM : 105090501111014  
Jurusan : Matematika  
Program Studi : STATISTIKA  
Skripsi berjudul :

### KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL) DALAM PENDUGAAN PARAMETER MODEL REGRESI LOGISTIK MULTILEVEL

Dengan ini menyatakan bahwa:

1. Isi dari skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam skripsi ini.
2. Apabila dikemudian hari ternyata skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 17 juli 2014  
Yang menyatakan,

(AHMAD TANTOWI JAUHARI)  
NIM. 105090501111014

# KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL) DALAM PENDUGAAN PARAMETER MODEL REGRESI LOGISTIK MULTILEVEL

## ABSTRAK

Dalam meneliti hubungan suatu peubah prediktor yang mempengaruhi peubah respon dapat digunakan analisis regresi sederhana. Namun untuk peubah respon bernilai kategorik, analisis regresi yang digunakan yaitu analisis regresi logistik. Untuk data yang memiliki struktur hirarki, analisis regresi logistik dikembangkan menjadi analisis regresi logistik multilevel. Analisis ini berguna untuk mengetahui keragaman kelompok pada level yang lebih tinggi yang nantinya mempengaruhi unit pengamatan pada level yang lebih rendah. Terdapat beberapa metode dalam menduga parameter pada regresi logistik multilevel, seperti metode *Penalized Quasi Likelihood* (PQL) dan metode *Maximum Likelihood Estimation* (MLE). Penelitian ini bertujuan untuk mengetahui metode yang lebih baik antara metode PQL atau MLE berdasarkan nilai *Apparent Error Rate* (APER). Pada tiga data yang dianalisis, didapat nilai APER yang sama untuk dua data dan satu data terdapat perbedaan namun dengan selisih yang sangat kecil. Sehingga dapat diambil kesimpulan hasil dari pendugaan menggunakan metode PQL dan MLE memiliki nilai keakuratan yang sama. Namun untuk hal komputasi, metode PQL memiliki jumlah iterasi yang lebih sedikit serta membutuhkan waktu yang relatif lebih cepat jika dibandingkan dengan metode MLE. Untuk metode PQL dilakukan menggunakan bantuan *software* MLwiN 2.30, sedangkan untuk metode MLE dilakukan menggunakan bantuan *software* SAS 9.3 yaitu PROC NL MIXED.

*Kata Kunci: Penalized Quasi Likelihood, Maximum Likelihood Estimation, Apparent Error Rate*



# ASSESSMENT PENALIZED QUASI LIKELIHOOD (PQL) METHOD FOR ESTIMATION PARAMETERS IN A MULTILEVEL LOGISTIC REGRESSION

## ABSTRACT

In examining the relationship of the predictor variables that affect the response variable can be used simple regression analysis. But for categorical response variable-value, regression analysis is used logistic regression analysis. For data that have a hierarchical structure, logistic regression analysis was developed into a multilevel logistic regression analysis. This analysis is useful to know the diversity of the group at a higher level that will affect the unit of observation at a lower level. There are several methods of predicting the parameters in multilevel logistic regression, such as the method Penalized Quasi Likelihood (PQL) and the method of Maximum Likelihood Estimation (MLE). This study aims to determine which method is better between PQL or MLE method based on the value of Apparent Error Rate (APER). In the three analyzed data, obtained APER same value for two data and one data there is a difference, but the difference is very small. So it can be concluded from the results of the estimation using the MLE method PQL and has the same accuracy rate. However, for this computation, the method PQL has fewer number of iterations and requires a relatively more rapid when compared with the MLE method. For PQL method performed using statistical software MLwiN 2.30 while for the MLE method performed using statistical software SAS 9.3 PROC NLMIXED.

*Keywords: Penalized Quasi Likelihood, Maximum Likelihood Estimation, Apparent Error Rate*

## KATA PENGANTAR

Alhamdulillah, puji syukur kehadiran Allah SWT atas rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi yang berjudul “KAJIAN METODE *PENALIZED QUASI LIKELIHOOD* (PQL) DALAM PENDUGAAN PARAMETER MODEL REGRESI LOGISTIK MULTILEVEL”. Dalam penyusunan skripsi ini, banyak pihak yang telah memberikan bantuan kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Samingun Handoyo, S. Si., M. Cs. selaku dosen pembimbing yang telah memberikan bimbingan dan nasehat hingga skripsi ini terselesaikan dengan baik.
2. Eni Sumarminingsih, S. Si., MM. selaku dosen penguji I yang telah memberikan saran, masukan, dan bimbingan hingga skripsi ini terselesaikan dengan baik.
3. Ir. Heni Kusdarwati, MS. selaku dosen penguji II yang telah memberikan saran dan masukan hingga skripsi ini terselesaikan dengan baik.
4. Dr. Abdul Rouf Alghofari, M.Sc. selaku ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang.
5. Kedua orang tua atas segala usaha dan doa yang diberikan.
6. Seluruh keluarga yang telah memberikan semangat serta doa.
7. Teman-teman seperjuangan Statistika Universitas Brawijaya, khususnya angkatan 2010.
8. Semua pihak yang telah membantu menyelesaikan penyusunan skripsi ini yang tidak dapat disebutkan satu persatu terima kasih atas bantuan yang telah diberikan dan semoga Allah memeberikan balasan kebaikan.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan oleh karena itu saran dan kritik yang membangun demi kesempurnaan penulisan selanjutnya sangat penulis harapkan. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membutuhkan.

Malang, Juli 2014

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	ii
<b>LEMBAR PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	ix
<b>DAFTAR TABEL</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Model Regresi Multilevel .....	5
2.2 Model Regresi Logistik 2 Level.....	6
2.3 Pendugaan Parameter Model Regresi Logistik 2 Level	8
2.3.1 Metode <i>Penalized Quasi Likelihood</i> (PQL) .....	8
2.3.2 Metode <i>Maximum Likelihood Estimation</i> (MLE).....	20
2.4 Pengujian Signifikasi Parameter .....	21
2.5 Pengujian Kesesuaian Model .....	22
2.6 Kriteria Pemilihan Model Terbaik .....	23
2.7 Interpretasi Model Regresi Logistik Multilevel.....	24
<b>BAB III METODE PENELITIAN</b>	
3.1 Data Penelitian .....	27
3.1.1 Data 1 .....	27
3.1.2 Data 2 .....	28
3.1.3 Data 3 .....	31
3.2 Tahap Analisis.....	32



## **BAB IV HASIL DAN PEMBAHASAN**

### **4.1 Hasil Pendugaan Parameter dan Uji Signifikansi**

Parameter .....	39
4.1.1 Data 1 .....	39
4.1.2 Data 2 .....	42
4.1.3 Data 3 .....	46
4.2 Pengujian Kesesuaian Model .....	49
4.3 Pemilihan Model Terbaik .....	49
4.3.1 Data 1 .....	49
4.3.2 Data 2 .....	50
4.3.3 Data 3 .....	52
4.4 Interpretasi Model .....	55
4.4.1 Nilai Peluang .....	55
4.4.2 <i>Odds Ratio</i> .....	57

## **BAB IV KESIMPULAN DAN SARAN**

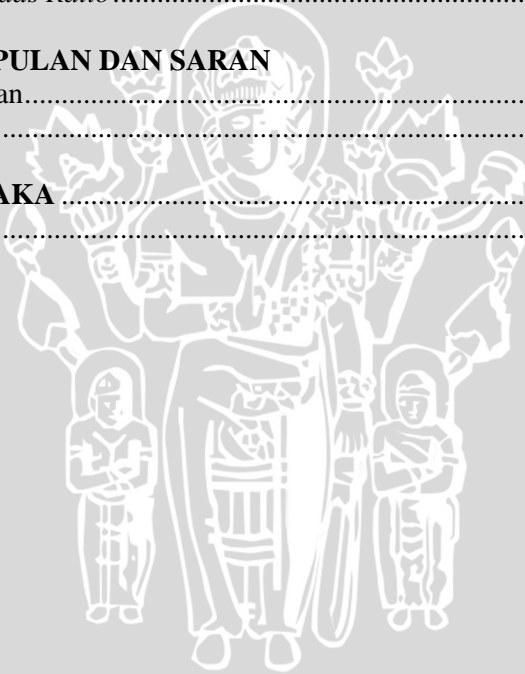
5.1 Kesimpulan .....	61
5.2 Saran .....	61

## **DAFTAR PUSTAKA**

63

## **LAMPIRAN**

65



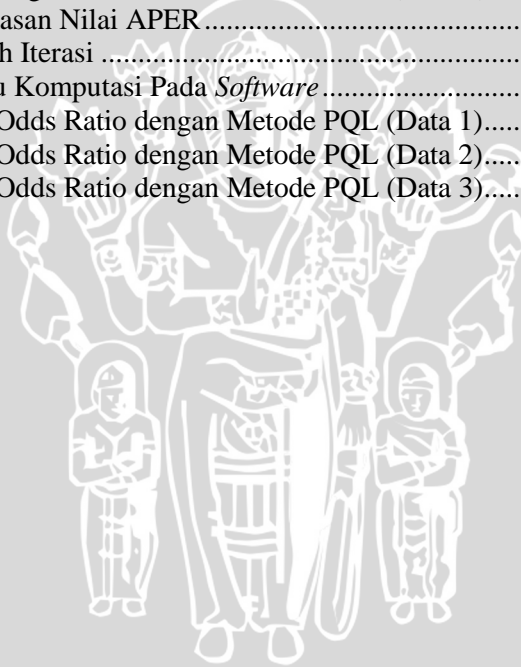
## DAFTAR GAMBAR

	Halaman
Gambar 3.1 Struktur Hirarki Data 1 .....	28
Gambar 3.2 Struktur Hirarki Data 2 .....	30
Gambar 3.3 Struktur Hirarki Data 3 .....	32
Gambar 3.4 Diagram Alir Metode Penelitian.....	35
Gambar 3.5 Diagram Alir Metode <i>Penalized Quasi Likelihood</i> (PQL) .....	36
Gambar 3.6 Diagram Alir Metode <i>Maximum Likelihood</i> <i>Estimation</i> (MLE).....	37



## DAFTAR TABEL

	Halaman
Tabel 2.1 Perhitungan Nilai APER.....	24
Tabel 4.1 Hasil Pendugaan Parameter (Data 1).....	39
Tabel 4.2 Hasil Pendugaan Parameter (Data 2).....	42
Tabel 4.3 Hasil Pendugaan Parameter (Data 3).....	46
Tabel 4.4 Hasil Uji Kesesuaian Model.....	49
Tabel 4.5 Perhitungan Nilai APER Metode PQL (Data 1).....	49
Tabel 4.6 Perhitungan Nilai APER Metode MLE (Data 1).....	50
Tabel 4.7 Perhitungan Nilai APER Metode PQL (Data 2).....	51
Tabel 4.8 Perhitungan Nilai APER Metode MLE (Data 2).....	51
Tabel 4.9 Perhitungan Nilai APER Metode PQL (Data 3).....	52
Tabel 4.10 Perhitungan Nilai APER Metode MLE (Data 3).....	52
Tabel 4.11 Ringkasan Nilai APER.....	53
Tabel 4.12 Jumlah Iterasi .....	54
Tabel 4.13 Waktu Komputasi Pada <i>Software</i> .....	54
Tabel 4.14 Nilai Odds Ratio dengan Metode PQL (Data 1).....	57
Tabel 4.15 Nilai Odds Ratio dengan Metode PQL (Data 2).....	58
Tabel 4.16 Nilai Odds Ratio dengan Metode PQL (Data 3).....	59



## DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Data Bakat Siswa TK/RA di Desa Maron Kec. Srengat Dalam Mewarnai Tahun 2009 (Data 1).... 65
Lampiran 2	Data Nilai Akreditasi SMK di Jawa Timur Tahun 2006 (Data 2)..... 67
Lampiran 3	Data Bangladesh <i>Fertility Survey</i> (Dhaka) Tahun 1989 (Data 3)..... 69
Lampiran 4	Syntax Prosedur NLMIXED Pada <i>Software</i> SAS 9.3 (Data 1) ..... 72
Lampiran 5	Syntax Prosedur NLMIXED Pada <i>Software</i> SAS 9.3 (Data 2) ..... 73
Lampiran 6	Syntax Prosedur NLMIXED Pada <i>Software</i> SAS 9.3 (Data 3) ..... 74
Lampiran 7	<i>Output</i> Metode PQL Dengan MLwiN 2.30 (Data 1)..... 75
Lampiran 8	<i>Output</i> Metode PQL Dengan MLwiN 2.30 (Data 2)..... 76
Lampiran 9	<i>Output</i> Metode PQL Dengan MLwiN 2.30 (Data 3)..... 77
Lampiran 10	<i>Output</i> Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 1) ..... 78
Lampiran 11	<i>Output</i> Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 2) ..... 81
Lampiran 12	<i>Output</i> Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 3) ..... 84
Lampiran 13	Model Logistik Multilevel Menggunakan Metode PQL (Data 1) ..... 87
Lampiran 14	Model Logistik Multilevel Menggunakan Metode PQL (Data 2) ..... 88
Lampiran 15	Model Logistik Multilevel Menggunakan Metode PQL (Data 3) ..... 89
Lampiran 16	Model Logistik Multilevel Menggunakan Metode MLE (Data 1) ..... 91
Lampiran 17	Model Logistik Multilevel Menggunakan Metode MLE (Data 2) ..... 92
Lampiran 18	Model Logistik Multilevel Menggunakan Metode MLE (Data 3) ..... 93

UNIVERSITAS BRAWIJAYA





# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada umumnya, dalam meneliti pengaruh satu peubah (peubah prediktor) terhadap satu peubah yang lain (peubah respon) atau pengaruh beberapa peubah terhadap satu peubah yang lain digunakan suatu analisis yang dikenal sebagai analisis regresi, baik analisis regresi linier sederhana ataupun analisis regresi linier berganda. Namun dalam analisis regresi linier, peubah respon disyaratkan harus memiliki suatu skala pengukuran yang kontinyu, sedangkan untuk skala pengukuran pada peubah prediktor dapat berupa skala pengukuran yang kontinyu, diskrit, ataupun campuran.

Untuk data yang memiliki struktur hirarki atau memiliki unit yang tersarang pada kolompok-kelompok tertentu, analisis regresi biasa tidak dapat langsung digunakan. Dalam perkembangannya, untuk data yang berstruktur hirarki dapat dilakukan menggunakan analisis regresi multilevel (Hox, 2002).

Sedangkan untuk peubah respon yang bernilai diskrit, sudah dikembangkan suatu analisis yaitu analisis regresi logistik. Dalam analisis regresi logistik, peubah respon pun dapat berupa peubah respon dikotomus ataupun peubah respon polikotomus (nominal dan ordinal). Analisis regresi logistik pun dapat dibedakan lagi, yaitu model regresi logistik dan model regresi logistik multilevel. Model regresi logistik digunakan pada analisis dengan data bersifat tanpa dikelompokkan, namun apabila data bersifat tersarang atau data memiliki struktur hirarki dapat digunakan analisis dengan model regresi logistik multilevel.

Untuk data yang memiliki struktur hirarki 2 level dengan respon biner, maka analisis yang dapat digunakan yaitu model regresi logistik 2 level dengan respon biner. Menurut Goldstein (2007), untuk melakukan pendugaan parameter dalam model logistik 2 level, terdapat beberapa cara yang dapat digunakan antara lain

metode *Quasi Likelihood* (*Marginal Quasi Likelihood* (MQL) atau *Penalized Quasi Likelihood* (PQL)) order pertama dan order kedua, metode *Bayesian Markov Chain Monte Carlo*, metode *Iterative Bootstrap*, dan metode *Quadrature*.

Metode PQL ini sebelumnya sudah dibahas oleh Rodriguez dan Goldman (1995) serta Khan dan Shaw (2011). Selain itu, metode PQL juga telah dibahas oleh Dewi (2008) dan didapatkan bahwa metode PQL merupakan metode yang melinierkan model regresi logistik multilevel, sehingga parameternya dapat diduga dengan metode pada model mmultilevel yang linier. Selain itu, model logistik multilevel sudah dibahas oleh Nazhiirah (2010) yang membahas tentang perbandingan metode *Maximum Likelihood Estimation* (MLE) dengan metode *Bayesian Markov Chain Monte Carlo* (MCMC). Dalam skripsi tersebut, didapatkan hasil bahwa metode MLE lebih baik daripada metode Bayesian MCMC berdasarkan nilai *Deviance Information Criterion* (DIC). Sehingga dalam skripsi ini, penulis akan membahas cara pendugaan parameter model logistik 2 level menggunakan metode *Penalized Quasi Likelihood* (PQL) dan sebagai pembanding dari metode PQL yaitu metode MLE yang memang biasa digunakan dalam melakukan analisis regresi logistik multilevel.

## **1.2 Rumusan Masalah**

Rumusan masalah pada skripsi ini yaitu:

1. Apakah metode *Penalized Quasi Likelihood* (PQL) lebih baik dalam pendugaan parameter model regresi logistik multilevel dibandingkan dengan metode *Maximum Likelihood Estimation* (MLE) berdasarkan nilai *Apparent Residual Rate* (APER)?
2. Apa keunggulan metode *Penalized Quasi Likelihood* (PQL) dibandingkan dengan metode *Maximum Likelihood Estimation* (MLE)?

### 1.3 Batasan Masalah

Batasan masalah pada skripsi ini adalah :

1. Peubah respon bersifat biner dengan peubah prediktor berskala diskrit dan kontinyu.
2. Model logistik multilevel yang dibahas adalah model regresi logistik dengan banyaknya level adalah 2 dengan *random intercept*.
3. Pendugaan parameter menggunakan metode *Penalized Quasi Likelihood* order pertama (PQL-1) dan *Maximum Likelihood Estimation* (MLE).
4. Dalam melakukan pemilihan model terbaik, digunakan nilai *Apparent Residual Rate* (APER).

### 1.4 Tujuan Penelitian

Tujuan dalam skripsi ini yaitu:

1. Menentukan metode yang lebih baik berdasarkan nilai *Apparent Residual Rate* (APER).
2. Mengetahui keunggulan metode *Penalized Quasi Likelihood* (PQL) dibandingkan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE).

### 1.5 Manfaat Penelitian

Dengan skripsi ini, diharapkan mampu memberikan informasi dan alternatif cara pendugaan parameter pada regresi logistik multilevel. Serta dapat diketahui metode yang lebih baik antara menggunakan metode PQL atau metode MLE.

UNIVERSITAS BRAWIJAYA



## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Model Regresi Multilevel

Suatu model analisis regresi untuk data berstruktur hirarki disebut dengan model regresi multilevel. Secara umum model regresi multilevel sama seperti model regresi biasa, yaitu terdapat sebuah peubah respon yang diukur pada setiap unit (level 1) serta satu atau lebih peubah prediktor yang diukur pada setiap level. Beda dengan model regresi biasa, pada model regresi multilevel peubah respon pada level 1 akan tersarang pada kelompok-kelompok dengan struktur hirarki yang lebih tinggi nantinya disebut sebagai level 2.

Model regresi multilevel yang paling sederhana adalah model 2 level dengan *random intercept*. Model matematisnya dapat dituliskan sebagai berikut (Goldstein, 1999):

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \quad 2.1$$

$$\beta_{0j} = \gamma_0 + u_{0j} \quad 2.2$$

di mana,

$Y_{ij}$  = peubah respon (hanya diukur pada level 1)

$X_{ij}$  = peubah prediktor pada level 1

$\beta_1$  = parameter koefisien regresi pada tahap 1

$\gamma_0$  = parameter koefisien regresi pada tahap 2

$e_{ij}$  = residual pada level 1

$u_{0j}$  = residual pada level 2

$i$  = individu dalam kelompok ke- $i$  ( $i = 1, 2, \dots, n_j$ )

$j$  = kelompok ( $j = 1, 2, \dots, m$ )

Pada model regresi multilevel, terdapat korelasi antar pengamatan yang disebut sebagai *Intra-Class Correlation* (ICC) yang menunjukkan keragaman yang terjadi pada peubah respon yang diakibatkan oleh kelompok. ICC adalah penduga untuk menjelaskan



keragaman proporsi dalam populasi. Untuk model pada persamaan 2.1, ICC antar dua individu adalah :

$$ICC = \frac{\sigma_2^2}{\sigma^2 + \sigma_2^2} \quad 2.3$$

di mana,

$\sigma^2$  = ragam residual pada level 1 ( $\text{var}(\epsilon_{ij})$ )

$\sigma_2^2$  = ragam residual pada level 2 ( $\text{var}(u_{0j})$ )

## 2.2 Model Regresi Logistik 2 Level

Model regresi logistik multilevel digunakan jika pada data, peubah respon merupakan peubah kategori misalnya saja jika peubah respon bernilai biner (2 kategori). Misalkan, bila  $Y=1$  menyatakan kejadian sukses dan  $Y=0$  menyatakan kejadian gagal, maka peluang akan terjadi sukses pada subyek ke- $i$  dan level ke- $j$  dinyatakan dengan  $P(Y=1) = \pi_{ij}$  dan peluang akan terjadi gagal pada subyek ke- $i$  dan level ke- $j$  yaitu  $P(Y=0) = 1 - \pi_{ij}$ . Sehingga ragam  $Y$  dinyatakan sebagai :

$$\text{Var}(Y_{ij}) = \pi_{ij} (1 - \pi_{ij}) \quad 2.4$$

Seperti pada model regresi logistik biasa, bentuk persamaan model regresi logistik 2 level dengan *random intercept* dapat ditulis sebagai berikut:

$$\pi_{ij} = f(\mathbf{X}'_{ij}\boldsymbol{\beta} + u_{0j}) = \frac{1}{1 + \exp[-(\mathbf{X}'_{ij}\boldsymbol{\beta} + u_{0j})]} \quad 2.5$$

Pada peubah respon biner, peubah respon dapat dituliskan sebagai jumlah dari probabilitas  $\pi_{ij}$  dengan residual level 1  $\epsilon_{ij}$  yang diasumsikan mempunyai distribusi kumulatif logistik:

$$y_{ij} = \pi_{ij} + \epsilon_{ij} \quad 2.6$$

di mana  $\pi_{ij}$  seperti ditunjukkan pada Persamaan 2.5,  $\varepsilon_{ij}$  merupakan residual model level 1 yang mempunyai rata-rata nol dan ragam  $\pi_{ij}$  ( $1 - \pi_{ij}$ ), serta  $y_{ij}$  berdistribusi Bernoulli dengan probabilitas  $\pi_{ij}$  ( $y_{ij} \sim \text{Bernoulli}(1, \pi_{ij})$ ).

Fungsi  $\pi_{ij}$  merupakan fungsi non linier sehingga perlu dilakukan transformasi logit untuk memperoleh fungsi yang linier agar dapat dilihat hubungan antara peubah respon (Y) dengan peubah prediktor (X) (Hosmer dan Lemeshow, 2000). Bentuk logit dari  $\pi_{ij}$  digunakan sebagai  $g(x)$ , yaitu :

$$g(x_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) \quad 2.7$$

Sehingga model regresi logistik 2 level dapat dituliskan sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) &= \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) \\ &= \beta_0 + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + u_{0j} \\ &= \mathbf{X}'_{ij}\boldsymbol{\beta} + u_{0j} \end{aligned} \quad 2.8$$

$\text{logit}(\pi_{ij})$  merupakan prediktor linier yang terdiri dari *fixed effect part* dari model ( $\mathbf{X}'_{ij}\boldsymbol{\beta}$ ) dan *random effect part* ( $u_{0j}$ ).

Dari persamaan 2.8, untuk mengetahui pengaruh level 2 dijelaskan melalui residual level 2 dalam persamaan berikut ini :

$$\hat{u}_{0j} = \frac{n_j \sigma_{u_0}^2}{n_j \sigma_{u_0}^2 + \sigma_\varepsilon^2} \hat{y}_j \quad 2.9$$

di mana

$\sigma_\varepsilon^2$  = ragam residual pada level 1

$\sigma_{u_0}^2$  = ragam residual pada level 2

$$\hat{y}_j = \frac{\sum \hat{y}_{ij}}{n_j}$$

Untuk nilai *Intra-Class Correlation* (ICC) pada model regresi logistik multilevel dapat dihitung dengan rumus sebagai berikut (Goldstein, 2007):

$$ICC = \frac{\sigma_{u_{0j}}^2}{\sigma_{u_{0j}}^2 + \sigma_{\varepsilon}^2} \quad 2.10$$

di mana,

$\sigma_{u_{0j}}^2$  = ragam residual level-2

$\sigma_{\varepsilon}^2$  =  $\pi^2/3$

## 2.3 Pendugaan Parameter Model Regresi Logistik 2 Level

### 2.3.1 Metode *Penalized Quasi Likelihood* (PQL)

Dalam melakukan pendugaan parameter model regresi logistik 2 level terdapat beberapa metode yang dapat digunakan. Pada skripsi ini pendugaan parameter model regresi logistik 2 level yang digunakan ialah metode *Penalized Quasi Likelihood* order pertama (PQL-1). Menurut Khan dan Shaw (2011) dijelaskan bahwa pendugaan parameter dengan metode PQL pada dasarnya adalah melinierkan bagian yang nonlinier dari model sehingga didapat model yang linier terlebih dahulu. Parameter-parameter dari model yang sudah linier diduga menggunakan prosedur pendugaan model linier 2 level. Hasil dari pendugaan parameter-parameter akan menghasilkan model yang telah linier yang nantinya akan digunakan dalam melakukan pendugaan parameter-parameter model regresi logistik 2 level berulang-ulang sampai konvergen karena prosedur yang digunakan merupakan prosedur *iterative*.

Sesuai Rodriguez dan Goldman (1995), langkah pertama dari metode PQL yaitu melinierkan bagian yang nonlinier dari Persamaan

2.6, yaitu  $\pi_{ij}$  atau sama dengan melinierkan fungsi dari  $\mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j}$ . Cara yang digunakan untuk melinierkan  $f(\mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j})$  ialah dengan perluasan deret Taylor .

Misalkan  $\mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} = H$ , sehingga dapat ditulis  $f(H) = \pi_{ij}$  dan turunan pertama dari  $f(H)$  adalah  $f'(H) = \pi_{ij}(1 - \pi_{ij})$ . Maka perluasan deret Taylor order pertama untuk fungsi  $f(H)$  dinyatakan sebagai berikut:

$$f(H) \approx f(H_0) + (H - H_0)f'(H_0)$$

Dengan mensubstitusikan  $H$  dengan  $\mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j}$ , dan  $H_0$  menyatakan suatu nilai, maka persamaan di atas menjadi:

$$\begin{aligned} f(H) &\approx f(H_0) + \left( \mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} - \mathbf{X}_{ij}\boldsymbol{\beta}^{(0)} + u_{0j}^{(0)} \right) f'(H_0) \\ &= f(H_0) + \left( (\mathbf{X}_{ij}\boldsymbol{\beta} - \mathbf{X}_{ij}\boldsymbol{\beta}^{(0)}) + (u_{0j} - u_{0j}^{(0)}) \right) f'(H_0) \\ &= f(H_0) + (\mathbf{X}_{ij}\boldsymbol{\beta} - \mathbf{X}_{ij}\boldsymbol{\beta}^{(0)})f'(H_0) + (u_{0j} - u_{0j}^{(0)})f'(H_0) \\ &= f(H_0) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)})f'(H_0) + (u_{0j} - u_{0j}^{(0)})f'(H_0) \end{aligned} \quad 2.11$$

PQL merupakan metode yang bersifat iteratif, di mana pada setiap iterasi dilakukan pelinierisasian pada nilai  $\boldsymbol{\beta}$  dan  $u_{0j}$ . Nilai keduanya berbeda pada setiap iterasi, karena nilai  $\boldsymbol{\beta}$  dan  $u_{0j}$  pada waktu ke- $t$  merupakan nilai yang diperoleh dari nilai di iterasi sebelumnya. Selanjutnya, pelinierisasian bagian yang nonlinier dari model pada iterasi ke- $t$  mengikuti ketentuan metode PQL order pertama dan dapat dituliskan sebagai berikut:

$$\begin{aligned} \pi_{ij} &= f(\mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j}) \\ &\approx f(H_t) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(t)})f'(H_t) + (u_{0j} - u_{0j}^{(t)})f'(H_t) \end{aligned} \quad 2.12$$

di mana

$t$  = menyatakan langkah iterasi,  $t = 0, 1, 2, \dots$ , (sampai iterasi berhenti)

$t = 0$  menyatakan nilai inisial

$f(H_t)$  = merupakan nilai dari fungsi  $f(H)$  pada titik  $H_t$  saat iterasi ke  $t$ , dimana  $f(H_t) = \pi_{ij}^{(t)}$

$\pi_{ij}^{(t)}$  = nilai  $\pi$  untuk unit level 1 ke- $i$  pada level 2 ke- $j$  saat iterasi ke- $t$

$\mathbf{x}_{ij}$  = vektor ukuran  $1 \times (P+1)$  berisi observasi dari peubah-peubah prediktor untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level 2

$\boldsymbol{\beta}$  = vektor *fixed parameter* dengan ukuran  $(P+1) \times 1$

$\boldsymbol{\beta}^{(t)}$  = vektor *fixed parameter* dengan ukuran  $(P+1) \times 1$  pada iterasi ke- $t$

$f'(H_t)$  = merupakan nilai turunan pertama dari fungsi  $f(H)$  pada titik  $H_t$  saat iterasi ke- $t$

$u_{0j}$  = *random effect* untuk unit ke- $j$  pada level 2

$u_{0j}^{(t)}$  = *random effect* untuk unit ke- $j$  pada level 2 saat iterasi ke- $t$

Setelah  $\pi_{ij}$  mempunyai bentuk linier dengan pendekatan deret Taylor seperti pada persamaan 2.12, substitusikan bentuk linier tersebut ke dalam persamaan 2.6 dan dihasilkan:

$$y_{ij} = f(H_t) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(t)})f'(H_t) + (u_{0j} - u_{0j}^{(t)})f'(H_t) + \varepsilon_{ij} \quad 2.13$$

dengan

$y_{ij}$  = nilai respon untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level 2

$\varepsilon_{ij}$  = residual untuk unit ke- $i$  pada level-1 dalam unit ke- $j$  level-2



Pada langkah awal ( $t=0$ ), persamaan (2.13) ditulis:

$$y_{ij} = f(H_0) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)})f'(H_0) + (u_{0j} - u_{0j}^{(0)})f'(H_0) + \varepsilon_{ij} \quad 2.14$$

dan diberikan suatu nilai awal  $\pi_{ij}^{(0)}$ ,  $\boldsymbol{\beta}^{(0)}$ , dan  $u_{0j}^{(0)}$ .

Selanjutnya untuk memperoleh model dalam bentuk linier, bagi persamaan 2.14 dengan  $f'(H_t) = \pi_{ij}^{(t)} (1 - \pi_{ij}^{(t)}) = w_{ij}^{(t)}$  sehingga diperoleh:

$$\begin{aligned} y_{ij} &= \pi_{ij} + \varepsilon_{ij} \\ y_{ij} &= f(H_t) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(t)})f'(H_t) + (u_{0j} - u_{0j}^{(t)})f'(H_t) + \varepsilon_{ij} \\ \frac{y_{ij}}{f'(H_t)} &= \frac{f(H_t) + \mathbf{X}_{ij}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(t)})f'(H_t) + (u_{0j} - u_{0j}^{(t)})f'(H_t) + \varepsilon_{ij}}{f'(H_t)} \\ \frac{y_{ij}}{f'(H_t)} &= \frac{f(H_t)}{f'(H_t)} + \mathbf{X}_{ij}\boldsymbol{\beta} - \mathbf{X}_{ij}\boldsymbol{\beta}^{(t)} + u_{0j} - u_{0j}^{(t)} + \frac{\varepsilon_{ij}}{f'(H_t)} \\ \mathbf{X}_{ij}\boldsymbol{\beta}^{(t)} + u_{0j}^{(t)} + \frac{y_{ij} - f(H_t)}{f'(H_t)} &= \mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} + \frac{\varepsilon_{ij}}{f'(H_t)} \\ \mathbf{X}_{ij}\boldsymbol{\beta}^{(t)} + u_{0j}^{(t)} + \frac{y_{ij} - \pi_{ij}^{(t)}}{\pi_{ij}^{(t)}(1 - \pi_{ij}^{(t)})} &= \mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} + \frac{\varepsilon_{ij}}{\pi_{ij}^{(t)}(1 - \pi_{ij}^{(t)})} \\ \mathbf{X}_{ij}\boldsymbol{\beta}^{(t)} + u_{0j}^{(t)} + \frac{y_{ij} - \pi_{ij}^{(t)}}{w_{ij}^{(t)}} &= \mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} + \frac{\varepsilon_{ij}}{w_{ij}^{(t)}} \quad 2.15 \end{aligned}$$

Dari persamaa 2.15 dilakukan pengaturan sedemikian rupa sehingga diperoleh:

$$y_{ij}^{**} = \mathbf{X}_{ij}\boldsymbol{\beta} + u_{0j} + \varepsilon_{ij}^{**} \quad 2.16$$

Persamaan 2.16 merupakan model dalam bentuk linier, dimana berdasarkan persamaan 2.15 dan 2.16 menjadi:

$$y_{ij}^{**t) = X_{ij}\beta^{(t)} + u_{0j}^{(t)} + \frac{y_{ij} - \pi_{ij}^{(t)}}{w_{ij}^{(t)}} \quad 2.17$$

dengan

$t$  = menyatakan langkah iterasi,  $t = 0, 1, 2, \dots$ , (sampai iterasi berhenti)

$t = 0$  menyatakan nilai inisial

$y_{ij}^{**t)}$  = merupakan nilai respon yang telah ditransformasi untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level 2 saat iterasi ke- $t$

$x_{ij}$  = vektor ukuran  $1 \times (P+1)$  berisi observasi dari peubah-peubah prediktor untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level 2

$\beta^{(t)}$  = vektor *fixed parameter* dengan ukuran  $(P+1) \times 1$  pada iterasi ke- $t$

$u_{0j}^{(t)}$  = *random effect* untuk unit ke- $j$  pada level 2 saat iterasi ke- $t$

$y_{ij}$  = merupakan nilai observasi (peubah respon) untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level 2

$\pi_{ij}^{(t)}$  = nilai  $\pi$  untuk unit ke- $i$  pada level 1 dalam unit ke- $j$  pada level-2 saat iterasi ke- $t$

$$w_{ij}^{(t)} = f'(H_t) = \pi_{ij}^{(t)}(1 - \pi_{ij}^{(t)})$$

dan

$$\varepsilon_{ij}^{**} = \frac{\varepsilon_{ij}}{w_{ij}^{(t)}} \quad 2.18$$

Model dalam persamaan 2.16 merupakan model yang telah di linierisasi dengan respon transformasi inisial adalah  $y_{ij}^{**t)}$  dan

$$\text{var}(\varepsilon_{ij}^{**}) = \text{var}\left(\frac{\varepsilon_{ij}}{w_{ij}^{(t)}}\right) \approx \frac{1}{w_{ij}^{(t)}}$$

Parameter-parameter dalam model yang telah dilinierisasi pada iterasi ke- $t$  yang dinyatakan dalam persamaan 2.16 dapat diduga

menggunakan prosedur pendugaan dari model multilevel linier, yaitu metode *Iterative Generalized Least Square* (IGLS).

Setelah diperoleh taksiran dari *fixed effect* dan *random effect* yang konvergen dengan IGLS, hitung nilai taksiran residual yang akan digunakan dalam proses linierisasi di iterasi berikutnya dengan:

$$u_{0j}^{(t)} = \frac{n_j \sigma_{u_0}^2}{n_j \sigma_{u_0}^2 + \sigma_\varepsilon^2} \bar{y}_j \quad 2.19$$

Pada persamaan 2.21 merupakan persamaan untuk nilai residual pada level 2, sehingga nilai residual pada level 1 adalah:

$$\varepsilon_{ij}^{(t)} = y_{ij}^{**} - u_{0j}^{(t)} \quad 2.20$$

di mana

$u_{0j}^{(t)}$  = *random effect* (residual level 2) untuk unit level 2 ke-j pada iterasi ke-t (baru bisa dicari pada iterasi pertama dan seterusnya,  $t = 1, 2, 3, \dots$ )

$n_j$  = banyaknya unit level 1 dalam unit ke-j pada level 2

$\sigma_{u_0}^2$  = ragam dari *random effect* (residual level 2) yang diperoleh dari penaksiran IGLS

$\sigma_\varepsilon^2$  = ragam residual level 1

$\bar{y}_j$  = rata-rata dari  $\tilde{y}_{ij}$  untuk setiap unit level 2 ke-j ( $\bar{y}_j = \tilde{y}_{ij} / n_j$ ), dengan  $\tilde{y}_{ij} = y_{ij} - \hat{y}_{ij}$  seperti pada proses IGLS.

Berikut merupakan langkah-langkah dari *Iterative Generalized Least Square* (IGLS):

Model regresi logistik multilevel yang sudah linier dapat dilakukan penaksiran menggunakan IGLS.

$$Y = X\beta + E$$

2.21

$Y$  merupakan vektor ukuran  $nx1$  yang berisi nilai observasi,

$$Y = \{y_{ij}\} = \begin{bmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{n_1 1} \\ y_{12} \\ y_{22} \\ \vdots \\ y_{n_1 1} \\ \vdots \\ y_{1m} \\ \vdots \\ y_{n_m m} \end{bmatrix}$$

$y_{ij}$  = respon untuk unit ke- $i$  dalam unit level-2 ke- $j$

$j = 1, 2, \dots, m$  menyatakan unit-unit level-2

$i = 1, 2, \dots, n_j$  menyatakan unit-unit level-1 yang bersarang dalam tiap unit level-2

total observasi dinyatakan oleh  $n$ , dengan  $n = \sum_{j=1}^m n_j$

$X$  merupakan *design matrix* ukuran  $nx(P+1)$  berisi konstanta satu dan observasi untuk  $P$ -peubah penjelas,  $m$  unit level-2, dan  $n$  jumlah total observasi, ditulis sebagai berikut:

$$X = \begin{bmatrix} 1 & x_{111} & \dots & x_{p11} \\ 1 & x_{121} & \dots & x_{p21} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n_m m} & \dots & x_{pn_m m} \end{bmatrix}$$

$\beta$  adalah vektor ukuran  $(P+1)x1$  yang berisi *fixed* parameter,

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

$E$  menyatakan matriks yang berisi penjumlahan residual level-1 dan level-2, di mana  $E = \{e_i\}$ ,  $e_{ij} = \varepsilon_{ij} + u_{0j}$ ,  $\sigma_e^2$  menyatakan variansi residual level-1 dan  $\sigma_u^2$  menyatakan variansi residual level-2.

Parameter-parameter yang akan ditaksir pada model regresi 2-level seperti yang telah dituliskan ulang dalam persamaan 2.21 adalah *fixed* parameter  $\beta_p$  dengan  $p = 1, 2, \dots, P$  dan *random* parameter,  $\sigma_{u0}^2$ . Langkah pertama dalam menaksir parameter dengan metode IGLS adalah menaksir *fixed* parameter  $\beta$ , untuk suatu matriks varians-kovarians  $V$  yang diketahui, dengan menggunakan *Generalized Least Square* (GLS):

$$\hat{\beta} = (X'V^{-1}X)^{-1}(X'V^{-1}Y) \quad 2.22$$

Sebagai taksiran inisial, matriks varian-kovarians yang digunakan pada persamaan 2.22 adalah  $V = \sigma_{\epsilon}^2 I$  (diasumsikan  $\sigma_{u0}^2 = 0$ ), di mana  $I$  adalah matriks identitas ukuran  $n \times n$ . Dengan diketahui  $V = \sigma_{\epsilon}^2 I$ , artinya pada taksiran inisial nilai taksiran diperoleh seperti pada *Ordinary Least Square*,  $\hat{\beta} = (X'X)^{-1}(X'Y)$ .

Setelah taksiran dari  $\beta$  diketahui, hitung nilai-nilai taksiran untuk  $Y$ , yaitu  $\hat{Y} = X\hat{\beta}$ . Sehingga dapat diperoleh nilai residual yang dinyatakan dalam bentuk:

$$\tilde{Y} = Y - \hat{Y} = Y - X\hat{\beta}$$

Bentuk cross product matriks  $\tilde{Y}\tilde{Y}'$  :

$$\tilde{Y}\tilde{Y}' = \begin{bmatrix} \tilde{y}_{11} \\ \tilde{y}_{21} \\ \vdots \\ \tilde{y}_{n_11} \\ \tilde{y}_{12} \\ \tilde{y}_{22} \\ \vdots \\ \tilde{y}_{n_22} \\ \vdots \\ \tilde{y}_{1m} \\ \vdots \\ \tilde{y}_{n_m m} \end{bmatrix} [\tilde{y}_{11} \quad \tilde{y}_{21} \quad \dots \quad \tilde{y}_{n_11} \quad \tilde{y}_{12} \quad \tilde{y}_{22} \quad \dots \quad \tilde{y}_{n_22} \quad \dots \quad \tilde{y}_{1m} \quad \dots \quad \tilde{y}_{n_m m}]$$



$$= \begin{bmatrix} \tilde{y}_{11}^2 & \tilde{y}_{11}\tilde{y}_{21} & \cdots & \tilde{y}_{11}\tilde{y}_{n_1} & \tilde{y}_{11}\tilde{y}_{12} & \tilde{y}_{11}\tilde{y}_{22} & \cdots & \tilde{y}_{11}\tilde{y}_{n_2} & \cdots & \tilde{y}_{11}\tilde{y}_{1m} & \cdots & \tilde{y}_{11}\tilde{y}_{n_m} \\ \tilde{y}_{21}\tilde{y}_{11} & \tilde{y}_{21}^2 & \cdots & \tilde{y}_{21}\tilde{y}_{n_1} & \tilde{y}_{21}\tilde{y}_{12} & \tilde{y}_{21}\tilde{y}_{22} & \cdots & \tilde{y}_{21}\tilde{y}_{n_2} & \cdots & \tilde{y}_{21}\tilde{y}_{1m} & \cdots & \tilde{y}_{21}\tilde{y}_{n_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{n_1,1}\tilde{y}_{11} & \tilde{y}_{n_1,1}\tilde{y}_{21} & \cdots & \tilde{y}_{n_1,1}^2 & \tilde{y}_{n_1,1}\tilde{y}_{12} & \tilde{y}_{n_1,1}\tilde{y}_{22} & \cdots & \tilde{y}_{n_1,1}\tilde{y}_{n_2} & \cdots & \tilde{y}_{n_1,1}\tilde{y}_{1m} & \cdots & \tilde{y}_{n_1,1}\tilde{y}_{n_m} \\ \tilde{y}_{12}\tilde{y}_{11} & \tilde{y}_{12}\tilde{y}_{21} & \cdots & \tilde{y}_{12}\tilde{y}_{n_1} & \tilde{y}_{12}^2 & \tilde{y}_{12}\tilde{y}_{22} & \cdots & \tilde{y}_{12}\tilde{y}_{n_2} & \cdots & \tilde{y}_{12}\tilde{y}_{1m} & \cdots & \tilde{y}_{12}\tilde{y}_{n_m} \\ \tilde{y}_{22}\tilde{y}_{11} & \tilde{y}_{22}\tilde{y}_{21} & \cdots & \tilde{y}_{22}\tilde{y}_{n_1} & \tilde{y}_{22}\tilde{y}_{12} & \tilde{y}_{22}^2 & \cdots & \tilde{y}_{22}\tilde{y}_{n_2} & \cdots & \tilde{y}_{22}\tilde{y}_{1m} & \cdots & \tilde{y}_{22}\tilde{y}_{n_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{n_2,2}\tilde{y}_{11} & \tilde{y}_{n_2,2}\tilde{y}_{21} & \cdots & \tilde{y}_{n_2,2}\tilde{y}_{n_1} & \tilde{y}_{n_2,2}\tilde{y}_{12} & \tilde{y}_{n_2,2}\tilde{y}_{22} & \cdots & \tilde{y}_{n_2,2}^2 & \cdots & \tilde{y}_{n_2,2}\tilde{y}_{1m} & \cdots & \tilde{y}_{n_2,2}\tilde{y}_{n_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{1m}\tilde{y}_{11} & \tilde{y}_{1m}\tilde{y}_{21} & \cdots & \tilde{y}_{1m}\tilde{y}_{n_1} & \tilde{y}_{1m}\tilde{y}_{12} & \tilde{y}_{1m}\tilde{y}_{22} & \cdots & \tilde{y}_{1m}\tilde{y}_{n_2} & \cdots & \tilde{y}_{1m}^2 & \cdots & \tilde{y}_{1m}\tilde{y}_{n_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{n_m,m}\tilde{y}_{11} & \tilde{y}_{n_m,m}\tilde{y}_{21} & \cdots & \tilde{y}_{n_m,m}\tilde{y}_{n_1} & \tilde{y}_{n_m,m}\tilde{y}_{12} & \tilde{y}_{n_m,m}\tilde{y}_{22} & \cdots & \tilde{y}_{n_m,m}\tilde{y}_{n_2} & \cdots & \tilde{y}_{n_m,m}\tilde{y}_{1m} & \cdots & \tilde{y}_{n_m,m}^2 \end{bmatrix}$$

Lakukan pemvektorisan pada matriks  $\tilde{Y}\tilde{Y}'$ :

$$\mathbf{Y}^* = \text{vec}(\tilde{Y}\tilde{Y}') = \begin{bmatrix} \tilde{y}_{11}^2 \\ \vdots \\ \tilde{y}_{n_m,m}\tilde{y}_{11} \\ \tilde{y}_{11}\tilde{y}_{21} \\ \vdots \\ \tilde{y}_{n_m,m}\tilde{y}_{21} \\ \vdots \\ \tilde{y}_{11}\tilde{y}_{n_m} \\ \vdots \\ \tilde{y}_{n_m,m}^2 \end{bmatrix} \quad 2.23$$

Operator  $\text{vec}$  merupakan operator yang membuat matriks ukuran  $n \times n$  menjadi vektor ukuran  $mn \times 1$  dengan menyusun entri-entri matriks pada kolom  $(s+1)$  di bawah entri terakhir kolom ke- $s$ , dengan  $s=1, 2, \dots, n$ .

Matriks varian-kovarians  $V$  ukuran  $n \times n$  adalah matriks block diagonal yang dinyatakan dalam bentuk sebagai berikut:

$$\mathbf{V} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_m \end{bmatrix} \quad 2.24$$

$m$  = banyaknya unit level-2 yang diobservasi, dan  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$  adalah matriks varian-kovarians untuk masing-masing unit level-2, yang didefinisikan sebagai berikut:

$$\begin{aligned}
 \mathbf{A}_1 &= \sigma_{u0}^2 \mathbf{J}_{(n_1)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_1)} \\
 \mathbf{A}_2 &= \sigma_{u0}^2 \mathbf{J}_{(n_2)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_2)} \\
 &\vdots \\
 \mathbf{A}_m &= \sigma_{u0}^2 \mathbf{J}_{(n_m)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_m)}
 \end{aligned}
 \tag{2.25}$$

Jika dijabarkan, matriks varian-kovarians untuk unit level-2 ke- $j$ ,  $\mathbf{A}_j$ , dijabarkan sebagai berikut:

$$\mathbf{A}_j = \begin{bmatrix} \sigma_{u0}^2 + \sigma_\varepsilon^2 & \sigma_{u0}^2 & \dots & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2 + \sigma_\varepsilon^2 & \dots & \sigma_{u0}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u0}^2 & \sigma_{u0}^2 & \dots & \sigma_{u0}^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

$\mathbf{A}_j$  berukuran  $n_j \times n_j$ , dengan  $\mathbf{I}_{(n_j)}$  adalah matriks identitas berukuran  $n_j \times n_j$  dan  $\mathbf{J}_{(n_j)}$  adalah matriks yang entri-entrinya berisi konstanta 1 ukuran  $n_j \times n_j$ .

Dari persamaan 2.24 dan persamaan 2.25, matriks varian-kovarians untuk  $n$  observasi, di mana  $n = \sum_{j=1}^m n_j$ , dinyatakan dalam bentuk:

$$\mathbf{V} = \begin{bmatrix} \sigma_{u0}^2 \mathbf{J}_{(n_1)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_1)} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \sigma_{u0}^2 \mathbf{J}_{(n_2)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_2)} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \sigma_{u0}^2 \mathbf{J}_{(n_m)} + \sigma_\varepsilon^2 \mathbf{I}_{(n_m)} \end{bmatrix}
 \tag{2.26}$$

Lakukan pemvektorisasian pada matriks varian-kovarians  $\mathbf{V}$  dengan menyusun entri-entri dari kolom ke- $(s+1)$  di bawah entri terakhir dari kolom ke- $s$  dari matriks  $\mathbf{V}$ , dengan  $s=1, 2, \dots, n$ . Vektorisasi matriks  $\mathbf{V}$  dinyatakan dalam notasi  $\mathbf{V}^*$ , dengan  $\mathbf{V}^*$  berukuran  $nm \times 1$ :

$$\mathbf{V}^* = \text{vec}(\mathbf{V}) = \begin{bmatrix} \sigma_{u0}^2 + \sigma_{\varepsilon}^2 \\ \sigma_{u0}^2 \\ \vdots \\ \sigma_{u0}^2 + \sigma_{\varepsilon}^2 \\ \sigma_{u0}^2 \\ \vdots \\ \sigma_{u0}^2 + \sigma_{\varepsilon}^2 \end{bmatrix} \quad 2.27$$

Diketahui nilai ekspekstasi dari  $\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}'$  adalah  $\mathbf{V}$ :

$$\begin{aligned} E\left((\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))'\right) &= E\left((\mathbf{Y} - \hat{\mathbf{Y}})(\mathbf{Y} - \hat{\mathbf{Y}})'\right) \\ &= E\left((\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\right) \\ &= E(\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}') = \mathbf{V} \end{aligned} \quad 2.28$$

Dengan pengaturan sedemikian rupa, bisa dibentuk model linier berdasarkan persamaan 2.28:

$$E(\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}') = \mathbf{V}$$

$$E(\text{vec}(\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}')) = \text{vec}(\mathbf{V})$$

$$E(\mathbf{Y}^*) = \mathbf{V}^*$$

Sehingga diperoleh model  $\mathbf{Y}^* = \mathbf{V}^* + \mathbf{R}$ , dengan  $\mathbf{R}$  menyatakan residual, yang dijabarkan dalam bentuk berikut:

$$\mathbf{Y}^* = \mathbf{V}^* + \mathbf{R}$$

$$\begin{bmatrix} \tilde{y}_{11}^2 \\ \vdots \\ \tilde{y}_{n_m m} \tilde{y}_{11} \\ \tilde{y}_{11} \tilde{y}_{21} \\ \vdots \\ \tilde{y}_{n_m m} \tilde{y}_{21} \\ \vdots \\ \tilde{y}_{11} \tilde{y}_{n_m m} \\ \vdots \\ \tilde{y}_{n_m m}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{u0}^2 + \sigma_{\varepsilon}^2 \\ \vdots \\ 0 \\ \sigma_{u0}^2 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ \sigma_{u0}^2 + \sigma_{\varepsilon}^2 \end{bmatrix} + \mathbf{R} = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix} \sigma_{u0}^2 + \begin{bmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix} \sigma_{\varepsilon}^2 + \mathbf{R} \quad 2.29$$

Pada model linier yang terbentuk dalam persamaan persamaan 2.29,  $Y^*$  dijadikan sebagai respon,  $\sigma_{u0}^2$  dan  $\sigma_\epsilon^2$  menjadi koefisien model, dan vektor-vektor berisi konstanta 0 dan 1 yang bersesuaian dengan  $\sigma_{u0}^2$  dan  $\sigma_\epsilon^2$  menjadi peubah-peubah penjelas. Sehingga pada persamaan 2.29, parameter-parameter yang akan ditaksir adalah  $\sigma_{u0}^2$  dan  $\sigma_\epsilon^2$ .

Jika vektor-vektor yang bersesuaian dengan  $\sigma_{u0}^2$  dan  $\sigma_\epsilon^2$  dalam persamaan 2.29 dinotasikan sebagai  $Z_1^*$  dan  $Z_2^*$ , kemudian dibentuk matriks  $Z^*=[Z_1^* Z_2^*]$ , dan parameter-parameter *random* yang akan ditaksir tergabung dalam vektor  $\theta$ , di mana  $\theta = \begin{bmatrix} \sigma_{u0}^2 \\ \sigma_\epsilon^2 \end{bmatrix}$ , maka persamaan 2.29 dapat dimodelkan dalam persamaan:

$$E(Y^*)=Z^*\theta \tag{2.30}$$

Dengan membentuk model yang dinyatakan dalam persamaan 2.32, parameter-parameter *random* yang ingin diketahui ( $\sigma_{u0}^2$  dan  $\sigma_\epsilon^2$ ) dapat ditaksir. Penaksiran parameter-parameter *random* dilakukan dengan metode yang sama seperti pada penaksiran parameter-parameter tetap  $\beta_p, p=1, 2, \dots, P$ , yaitu dengan menggunakan metode *Generalized Least Square* (GLS):

$$\theta = (Z^{*'} (V^*)^{-1} Z^*)^{-1} (Z^{*'} (V^*)^{-1} Y^*) \tag{2.31}$$

Dengan  $V^* = V \otimes V$ ,  $V^*$  berukuran  $nnxn$ .

Setelah diperoleh taksiran dari parameter-parameter *random*, ulangi langkah pengestimasiian *fixed* parameter dengan nilai matriks varian-kovarians yang baru, kemudian hasil penaksiran *fixed* parameter digunakan untuk menaksir *random* parameter, selanjutnya dilakukan penaksiran berulang-ulang secara bergantian antara *fixed* parameter dan *random* parameter sampai konvergen, yaitu nilai taksiran tidak lagi berfluktuasi pada iterasi-iterasi berikutnya.

### 2.3.2 Metode *Maximum Likelihood Estimation* (MLE)

Salah satu metode yang dapat digunakan dalam menduga parameter model regresi logistik multilevel adalah *Maximum Likelihood Estimation* (MLE). Metode ini memberikan nilai dugaan parameter  $\beta$  dengan cara memaksimumkan fungsi kemungkinan (*likelihood function*). Jika fungsi sebaran peluang untuk  $y_{ij}$  adalah:

$$f(y_{ij}) = \pi^{y_{ij}} (1 - \pi)^{1-y_{ij}} \quad 2.32$$

Maka fungsi *likelihood* untuk  $n_j$  pengamatan bebas adalah:

$$\begin{aligned} L(\beta_0, \beta_1, \beta_2, \dots, \beta_p) &= \prod_{j=1}^m \prod_{i=1}^{n_j} \{ [\pi(x_{ij})]^{y_{ij}} [1 - \pi(x_{ij})]^{1-y_{ij}} \} \\ &= \prod_{j=1}^m \prod_{i=1}^{n_j} \left\{ \left[ \frac{\pi(x_{ij})}{1 - \pi(x_{ij})} \right]^{y_{ij}} [1 - \pi(x_{ij})] \right\} \\ &= \left[ \frac{\pi(x_{ij})}{1 - \pi(x_{ij})} \right]^{\sum_{i=1}^{n_j} \sum_{j=1}^m y_{ij}} [1 - \pi(x_{ij})]^{n_j} \end{aligned} \quad 2.33$$

Pada model regresi logistik multilevel terdapat parameter dari *random effect* yang berdistribusi normal yang harus diduga, sehingga dalam pendugaan digunakan *marginal maximum likelihood*. Pada metode ini, *marginal likelihood* data observasi didapatkan dari integrasi distribusi *random effect*.

Menurut Flom, dkk (2006), fungsi *likelihood* yang digunakan untuk menyelesaikan pendugaan parameter pada model regresi logistik multilevel dinyatakan pada:

$$L(\pi, \sigma_{u_{0j}}^2 | y, x) = \prod_j \int_{-\infty}^{+\infty} \prod_i g(Y_{ij} | X_{ij}, u_{0j}) f(u_{0j}) du_{0j} \quad 2.34$$

dengan,



$$g(Y_{ij} | X_{ij}, u_{0j}) = \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}}$$

Secara umum persamaan 2.34 dapat diselesaikan melalui prosedur iterasi. Untuk menyelesaikan persamaan fungsi kemungkinan harus menggunakan teknik integrasi numerik. Dalam model multilevel prosedur yang digunakan adalah prosedur *Gaussian Quadrature* untuk menghitung integral secara numerik.

*Gaussian Quadrature* merupakan merupakan teknik integrasi secara numerik yang bermanfaat untuk menyelesaikan perhitungan integral dari fungsi seperti pada persamaan 2.34 (Flom, dkk, 2006).

$$\int g(x)f(x)dx = \sum_{i=1}^m w_m g(x_m) \tag{2.35}$$

di mana  $x_m$  merupakan poin gaussian quadrature dan  $w_m$  merupakan bobot.

## 2.4 Pengujian Signifikansi Parameter

Untuk menguji apakah peubah prediktor yang terdapat dalam model tersebut memiliki kontribusi yang nyata terhadap peubah respon, maka perlu dilakukan pengujian terhadap parameter.

Uji koefisien regresi secara parsial digunakan untuk memeriksa peranan koefisien regresi dari masing-masing peubah prediktor secara individu pada model, yaitu dengan membandingkan penduga dengan ragam penduganya (Hosmer dan Lemeshow, 2000).

Uji yang digunakan yaitu uji Wald dengan hipotesis :

$$H_0 : \beta_k = 0$$

Lawan

$$H_1 : \beta_k \neq 0, \text{ untuk } k = 0, 1, 2, \dots, p.$$

Statistik uji yang digunakan adalah :

$$W = \frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)} \tag{2.36}$$

$\widehat{\beta}_k$  adalah penduga  $\beta_k$  dan  $SE(\widehat{\beta}_k)$  merupakan penduga residual baku dari  $\widehat{\beta}_k$  yang didapat dari persamaan 2.11. Nilai  $SE(\widehat{\beta}_k)$  ditentukan dari nilai diagonal utama matriks kovarin yaitu :

$$\mathbf{cov}(\widehat{\beta}_k) = [\mathbf{X}'\mathbf{V}\mathbf{X}]^{-1}$$

$$\mathbf{V} = \begin{bmatrix} \widehat{\pi}_1(1 - \widehat{\pi}_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \widehat{\pi}_n(1 - \widehat{\pi}_n) \end{bmatrix}$$

$\text{var}(\widehat{\beta}_k) =$  diagonal utama matriks  $\mathbf{cov}(\widehat{\beta}_k)$

$$SE(\widehat{\beta}_k) = \sqrt{\text{var}(\widehat{\beta}_k)} \quad 2.37$$

Statistik uji W mengikuti sebaran normal, sehingga  $H_0$  ditolak jika  $W > Z_{\alpha/2}$  atau nilai- $p < \alpha$ .

## 2.5 Pegujian Kesesuaian Model

Untuk menguji apakah model yang terbentuk dari hasil pendugaan parameter sudah sesuai atau belum terhadap data, maka harus dilakukan pengujian kesesuaian model. Menurut Hosmer dan Lemeshow (2000), salah satu cara dalam melakukan pengujian kesesuaian model pada analisis regresi logistik yaitu menggunakan uji *Pearson* dengan hipotesis sebagai berikut:

$H_0$  : model sesuai dengan data

$H_1$  : model tidak sesuai dengan data

Statistik uji yang digunakan adalah :

$$\chi_{pearson}^2 = \sum_{j=1}^m \sum_{i=1}^{n_j} \frac{(Y_{ij} - \widehat{\pi}_{ij})^2}{\widehat{\pi}_{ij}(1 - \widehat{\pi}_{ij})} \quad 2.38$$

dengan

$Y_{ij}$  = nilai observasi peubah respon unit (level-1) ke- $i$  dan kelompok (level-2) ke- $j$

$\hat{\pi}_{ij}$  = peluang peubah respon unit level-1 ke- $i$  dan kelompok (level-2) ke- $j$

Statistik uji  $\chi^2_{pearson}$  menyebar mengikuti sebaran *chi-square*, sehingga  $H_0$  ditolak jika  $\chi^2_{pearson} > \chi^2_{(\alpha, n-p-1)}$  atau nilai- $p < \alpha$ .

## 2.6 Kriteria Pemilihan Model Terbaik

Untuk menentukan model terbaik dari dua model regresi logistik multilevel yang terbentuk, salah satu cara yang dapat digunakan yaitu menggunakan nilai *Apparent Residual Rate* (APER). Nilai APER merupakan suatu nilai yang berfungsi untuk melihat proporsi kesalahan dalam melakukan proses pengklasifikasian objek. Pengklasifikasian ini dilakukan dengan mencocokkan nilai peubah respon pada observasi yang nantinya dicocokkan dengan peubah respon prediksi sesuai model yang telah terbentuk.

Karena peubah respon memiliki nilai dengan dua kategori (biner), maka diberikan *cut value* yaitu sebesar 0,5 untuk nilai peluang. Jika peluang suatu model untuk unit ke- $i$  kelompok ke- $j$  bernilai lebih dari atau sama dengan 0,5 ( $\pi_{ij} \geq 0,5$ ) maka dapat dikatakan nilai duga adalah 1. Apabila peluang bernilai kurang dari 0,5 ( $\pi_{ij} < 0,5$ ) maka nilai duga adalah 0. Peluang untuk unit ke- $i$  pada kelompok ke- $j$  ( $\pi_{ij}$ ) dihitung sesuai model logistik pada persamaan 2.10 (Hosmer dan Lemeshow, 2000).

Perhitungan nilai APER dilakukan pada model regresi logistik multilevel yang terbentuk dari pendugaan menggunakan metode PQL maupun menggunakan metode MLE. Nilai APER akan menyatakan seberapa besar proporsi sampel yang salah diklasifikasikan. Model terbaik yaitu model dengan nilai APER yang lebih kecil dari model yang lain. Berikut merupakan cara menghitung nilai APER (Johnson dan Wichern, 2007):

Tabel 2.1 Perhitungan Nilai APER

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	$n_{11}$	$n_{12}$	$n_{11} + n_{12}$
	1	$n_{21}$	$n_{22}$	$n_{21} + n_{22}$
Total		$n_{11} + n_{21}$	$n_{12} + n_{22}$	$n_{11} + n_{12} + n_{21} + n_{22}$

di mana

$n_{11}$  = banyaknya unit dengan nilai respon observasi = 0 dan prediksi = 0

$n_{21}$  = banyaknya unit dengan nilai respon observasi = 0 dan prediksi = 1

$n_{12}$  = banyaknya unit dengan nilai respon observasi = 1 dan prediksi = 0

$n_{22}$  = banyaknya unit dengan nilai respon observasi = 1 dan prediksi = 1

$$APER = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \quad 2.39$$

## 2.7 Interpretasi Model Regresi Logistik Multilevel

Dalam melakukan interpretasi model regresi logistik multilevel, dapat digunakan nilai peluang seperti pada persamaan 2.10. Selain itu, dapat pula diinterpretasi menggunakan *Odds ratio* (OR). *Odds ratio* merupakan perbandingan peluang suatu kejadian sukses dengan peluang kejadian gagal. Dalam regresi logistik, *odds ratio* dapat digunakan untuk mempermudah interpretasi model yang terbentuk.

$$OR = \frac{\pi_{ij}(Y_{ij} | X_{ij} = 1) / [1 - \pi_{ij}(Y_{ij} | X_{ij} = 1)]}{\pi_{ij}(Y_{ij} | X_{ij} = 0) / [1 - \pi_{ij}(Y_{ij} | X_{ij} = 0)]} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{\beta_1} \quad 2.40$$

Nilai *odds ratio* digunakan untuk menunjukkan kecenderungan hubungan suatu peubah prediktor terhadap peubah respon. Misalkan  $OR=1$  berarti bahwa  $X=1$  mempunyai resiko yang sama dengan  $X=0$  untuk menghasilkan  $Y=1$ . Bila  $1 < OR < \infty$  berarti  $X=1$  memiliki resiko lebih tinggi sebanyak  $OR$  kali dibanding  $X=0$  untuk menghasilkan  $Y=1$ , sedangkan bila  $0 < OR < 1$  maka  $X=1$  memiliki resiko lebih tinggi  $1/OR$  kali dibanding  $X=0$  untuk menghasilkan  $Y=1$  (Hosmer dan Lemeshow, 2000).





UNIVERSITAS BRAWIJAYA



## BAB III METODE PENELITIAN

### 3.1 Data Penelitian

Dalam penelitian ini, data yang digunakan sebanyak tiga data dari berbagai sumber. Setiap data berstruktur hirarki dengan peubah respon bernilai biner dan peubah prediktor berskala diskrit ataupun kontinyu.

#### 3.1.1 Data 1

Data pertama merupakan data sekunder dari Nazhiirah (2010) mengenai bakat siswa TK/RA di Desa Maron Kecamatan Srengat dalam mewarnai tahun 2009. Data yang digunakan berjumlah 79 siswa TK/RA yang tersarang atau berkelompok pada 3 sekolah TK/RA. Peubah yang digunakan terdiri dari satu peubah respon dan empat peubah prediktor. Berikut keterangan dari peubah yang digunakan dalam penelitian:

1. Peubah respon: Bakat siswa ( $Y$ )

Keterangan: 0 = Kurang berbakat

1 = Berbakat

2. Peubah prediktor

- a. Skor pemilihan warna ( $X_1$ )

- b. Skor kerapian ( $X_2$ )

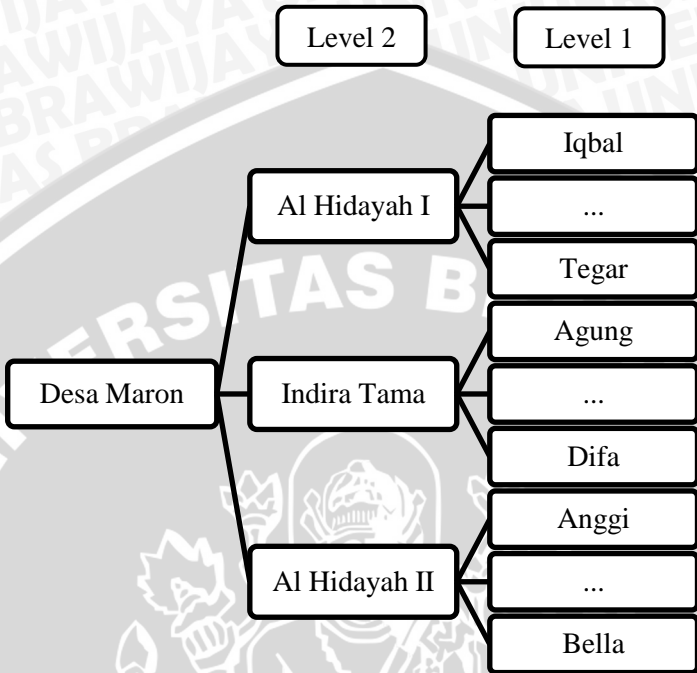
- c. Skor kebersihan ( $X_3$ )

- d. Jenis kelamin ( $X_4$ )

Keterangan: 0 = Perempuan

1 = Laki-laki

Pada data 1 ini, struktur hirarki terdapat pada setiap siswa TK/RA yang tersarang atau terkolompok pada tiga sekolah di Desa Maron Kec. Srengat di mana siswa tersebut bersekolah. Untuk lebih memperjelas struktur hirarki pada data bakat siswa TK/RA di Desa Maron Kec. Srengat dalam mewarnai, maka dapat digambarkan sebagai berikut:



Gambar 3.1 Struktur Hirarki Data 1

Pada gambar 4.1 dapat dijelaskan bahwa terdapat beberapa siswa yang tersarang pada sekolah sesuai di mana mereka sekolah. Sebagai contoh yaitu Iqbal dan Tia yang tersarang pada RA Al Hidayah I, Agung dan Yona pada TK Indira Tama, serta Anggi dan Moh. Riko Bima pada RA Al Hidayah II. Dapat dijelaskan pula bahwa siswa merupakan level pertama sedangkan sekolah merupakan level kedua pada model multilevel yang dibahas dalam skripsi ini.

### 3.1.2 Data 2

Untuk data kedua merupakan data nilai akreditasi SMK di Jawa Timur tahun 2006 yang diambil dari Nazhiirah (2010). Terdapat 99 SMK di Jawa Timur yang tersarang atau berkelompok

pada 21 kota/kabupaten. Peubah yang digunakan yaitu satu peubah respon dan tujuh peubah prediksi sebagai berikut:

1. Peubah respon: Nilai Akreditasi (Y)

Keterangan: 0 = nilai selain A

1 = nilai A

2. Peubah prediktor

a. Status sekolah ( $X_1$ )

Keterangan: 0 = swasta

1 = negeri

b. Lama sekolah berdiri (tahun) ( $X_2$ )

c. Banyak siswa aktif ( $X_3$ )

d. Banyak guru aktif ( $X_4$ )

e. Banyak alumni yang diterima dunia usaha dan industri setahun terakhir ( $X_5$ )

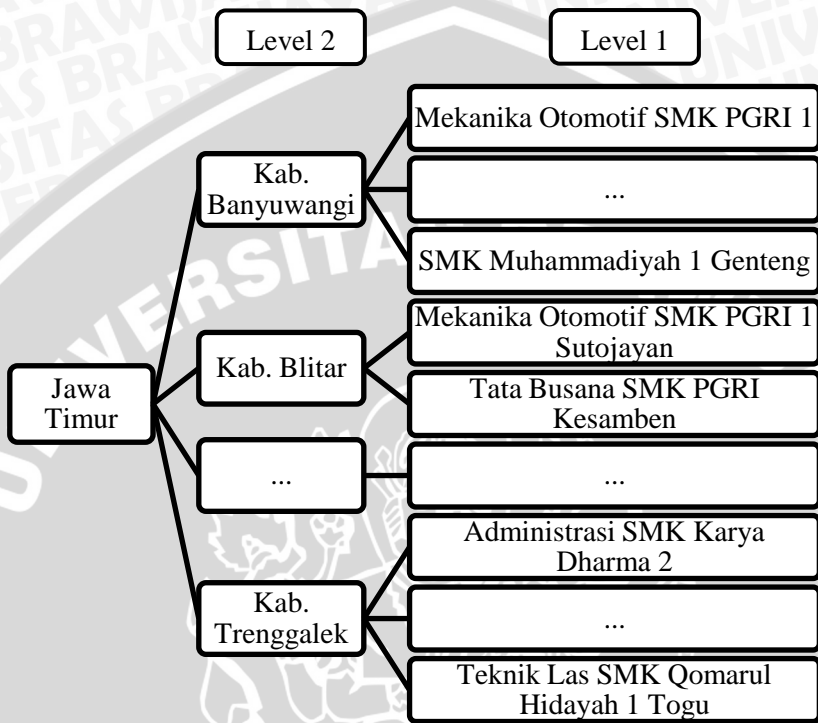
f. Status tanah bangunan ( $X_6$ )

Keterangan: 0 = bukan milik sendiri

1 = milik sendiri

g. Nilai rata-rata ujian nasional ( $X_7$ )

Struktur hirarki terdapat pada setiap Sekolah Menengah Kejuruan (SMK) di Jawa Timur yang tersarang pada kota/kabupaten di mana sekolah tersebut berada. Untuk lebih memperjelas struktur hirarki pada data nilai akreditasi SMK di Jawa Timur tahun 2006, maka dapat digambarkan sebagai berikut:



Gambar 3.2 Struktur Hirarki Data 2

Pada gambar 4.2 dapat dijelaskan bahwa terdapat beberapa SMK yang tersarang pada kota/kabupaten tertentu sesuai di mana SMK tersebut berada. Sebagai contoh yaitu Mekanika Otomotif SMK PGRI 1 yang tersarang pada Kabupaten Banyuwangi sesuai SMK tersebut berada, serta Jurusan Tata Busana SMK PGRI Kesamben yang tersarang pada Kabupaten Blitar. Dapat dijelaskan pula bahwa SMK merupakan level pertama sedangkan kota merupakan level kedua pada model multilevel yang dibahas dalam skripsi ini.



### 3.1.3 Data 3

Data ketiga merupakan data hasil survei yaitu *Bangladesh Fertility Survey* tahun 1989 tentang penggunaan alat kontrasepsi bagi wanita (Huq dan Cleland, 1990). Data yang digunakan merupakan data hirarki dengan jumlah observasi sebanyak 2711 yang tersarang pada 49 distrik yang ada di Bangladesh. Peubah yang digunakan terdiri dari satu peubah respon dan lima peubah prediktor sebagai berikut:

1. Peubah respon: status penggunaan alat kontrasepsi (Y)

Keterangan: 0 = tidak menggunakan alat kontrasepsi

1 = menggunakan alat kontrasepsi

2. Peubah prediktor:

- a. Status memiliki anak ( $X_1$ )

Keterangan: 0 = belum memiliki anak

1 = sudah memiliki anak

- b. Umur (tahun) ( $X_2$ )

- c. Daerah tinggal ( $X_3$ )

Keterangan: 0 = pedesaan

1 = perkotaan

- d. Status pendidikan ( $X_4$ )

Keterangan: 0 = tidak pernah bersekolah

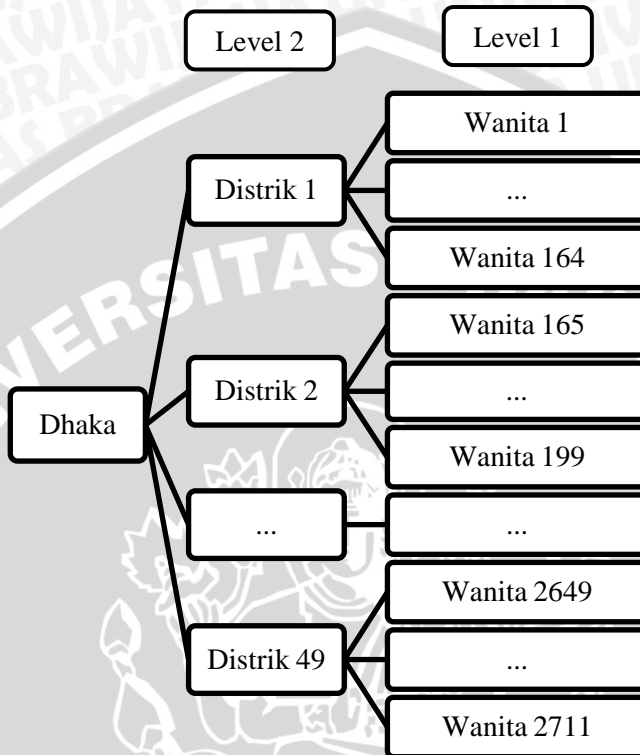
1 = pernah bersekolah

- e. Agama ( $X_5$ )

Keterangan: 0 = Muslim

1 = Hindu

Struktur hirarki terdapat pada setiap individu wanita yang tersarang pada distrik di mana mereka tinggal. Untuk lebih memperjelas struktur hirarki pada data hasil survei penggunaan alat kontrasepsi di Bangladesh tahun 1989, maka dapat digambarkan sebagai berikut:



Gambar 3.3 Struktur Hirarki Data 3

Pada Gambar 4.3 dapat dijelaskan bahwa terdapat beberapa wanita yang tersarang pada distrik tertentu sesuai di mana wanita tersebut tinggal. Terdapat 2711 wanita sebagai obyek yang diteliti yang tersarang pada 49 distrik. Untuk memudahkan, maka dilakukan penomoran pada setiap individu dan distrik sesuai pada Gambar 4.3.

### 3.2 Tahap Analisis

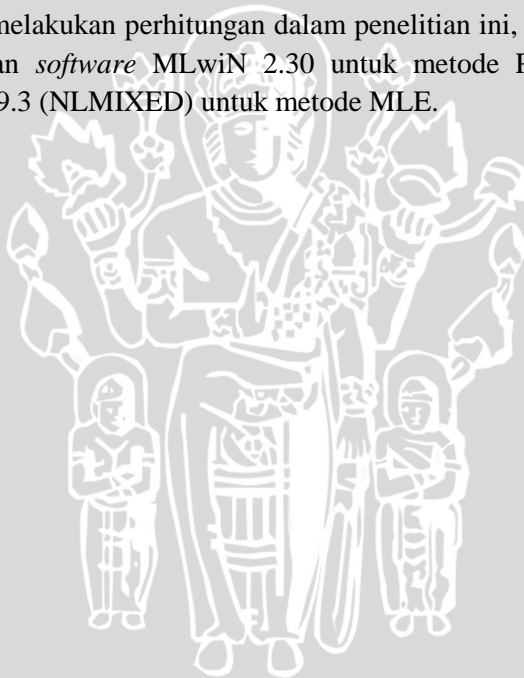
Tahapan analisis pada skripsi ini adalah sebagai berikut :

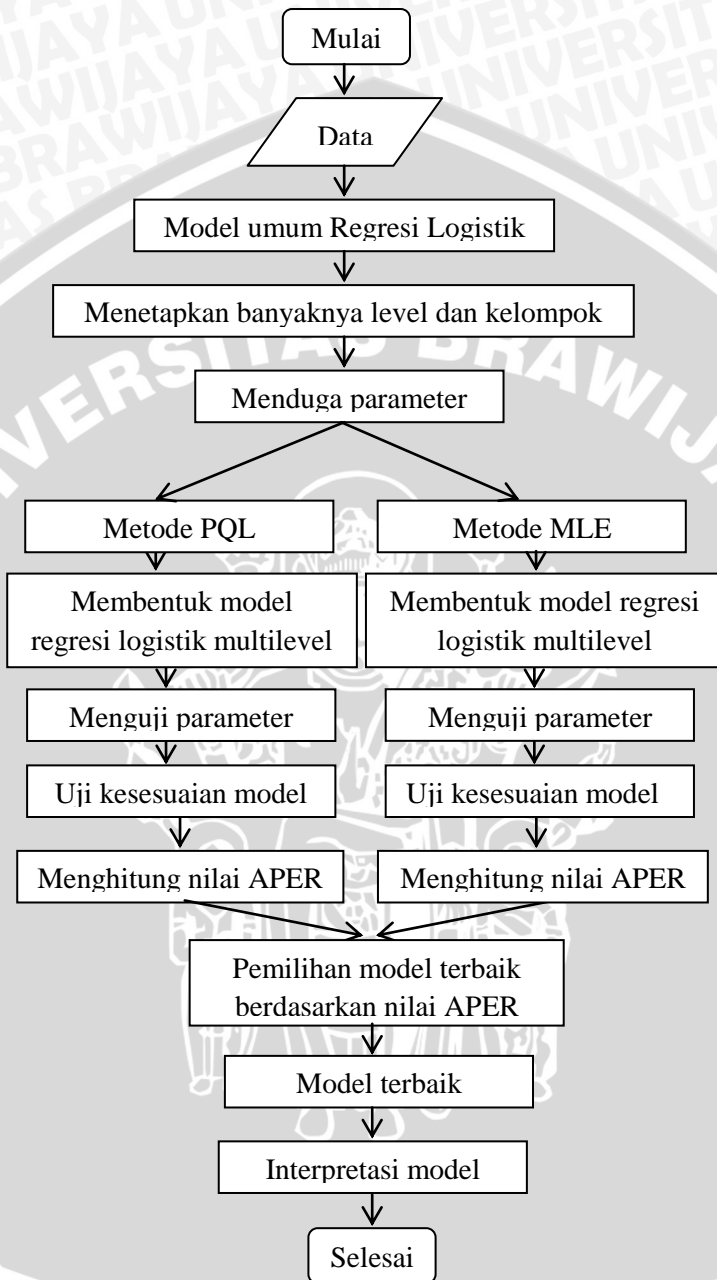
1. Menentukan model umum regresi logistik multilevel sesuai data.

2. Menetapkan banyaknya level dan kelompok yang akan digunakan dalam analisis. Dalam hal ini, level yang digunakan adalah 2 seperti pada batasan masalah.
3. Menduga parameter-parameter regresi logistik menggunakan dua metode, yaitu metode *Penalized Quasi Likelihood* (PQL) dengan langkah-langkah sebagai berikut:
  - a. Linierisasi model logistik pada persamaan 2.5 menggunakan perluasan deret Taylor order pertama sehingga terbentuk rumus sesuai pada persamaan 2.15.
  - b. Menduga *fixed* parameter dan *random* parameter pada persamaan 2.17 menggunakan *Iterative Generalized Least Square* (IGLS) hingga konvergen.
  - c. Mencatat waktu pada saat menjalankan program MLwiN 2.30.
  - d. Menghitung nilai  $u_{0j}$  untuk setiap kelompok pada level 2 seperti pada persamaan 2.9.
  - e. Membentuk model regresi logistik multilevel sesuai hasil pendugaan parameter 2.8.Sedangkan untuk langkah-langkah metode *Maximum Likelihood Estimation* (MLE) sebagai berikut:
  - a. Menentukan fungsi *Likelihood* untuk model regresi logistik multilevel seperti pada persamaan 2.34.
  - b. Karena fungsi *Likelihood* dalam menduga parameter model regresi logistik multilevel pada persamaan 2.34 tidak linier, maka proses pendugaan menggunakan prosedur *Gaussian Quadrature* dengan bantuan program NLMIXED SAS 9.3.
  - c. Mencatat waktu pada saat menjalankan program SAS 9.3.
  - d. Menghitung nilai  $u_{0j}$  untuk setiap kelompok pada level 2 seperti pada persamaan 2.9.
  - e. Membentuk model regresi logistik multilevel sesuai hasil pendugaan parameter 2.8.
4. Menguji signifikansi parameter secara parsial pada dua model yang telah didapat dari dua metode pendugaan menggunakan statistik uji Wald sesuai pada Persamaan 2.28.

5. Menguji kesesuaian model pada enam model yang telah terbentuk menggunakan uji *Pearson* sesuai pada persamaan 2.38.
6. Membandingkan hasil pendugaan parameter untuk dua metode menggunakan nilai APER menggunakan rumus pada Persamaan 2.39.
7. Menentukan model akhir yang akan diinterpretasi berdasarkan nilai APER pada langkah no. 5. Model yang lebih baik yang akan dipilih sebagai model akhir.
8. Melakukan interpretasi model akhir sesuai data yang digunakan dengan menggunakan nilai *Odds Ratio* seperti pada persamaan 2.40.

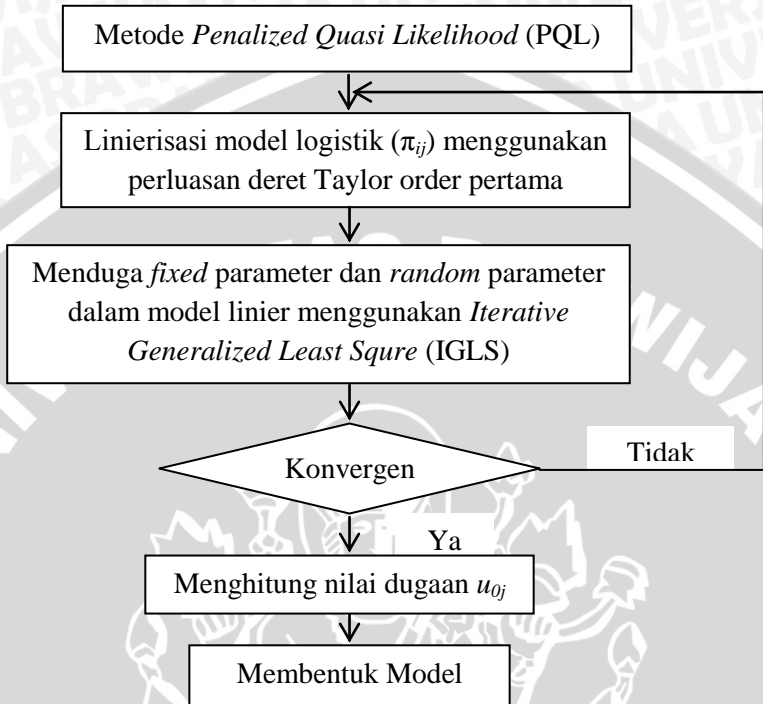
Dalam melakukan perhitungan dalam penelitian ini, dilakukan dengan bantuan *software* MLwiN 2.30 untuk metode PQL serta *software* SAS 9.3 (NLMIXED) untuk metode MLE.



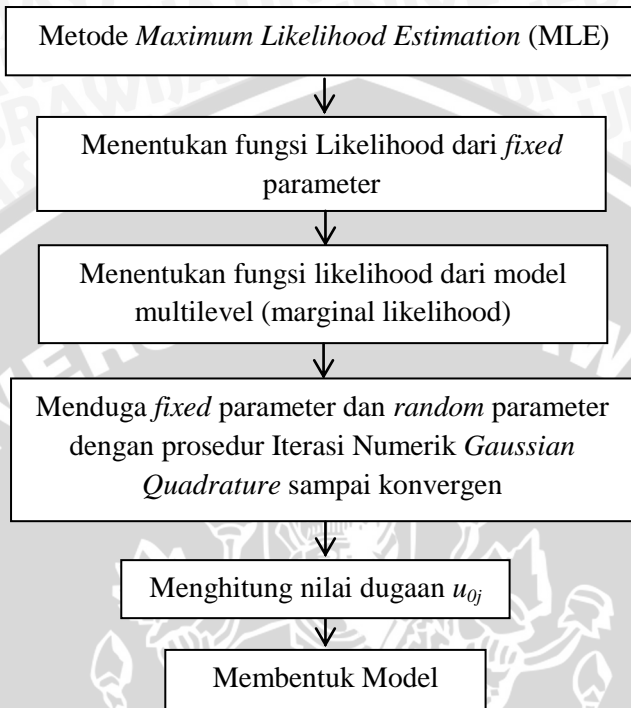


Gambar 3.4 Diagram Alir Metode Penelitian





Gambar 3.5 Diagram Alir Metode *Penalized Quasi Likelihood* (PQL)



Gambar 3.6 Diagram Alir Metode *Maximum Likelihood Estimation* (MLE)

UNIVERSITAS BRAWIJAYA



## BAB IV HASIL DAN PEMBAHASAN

### 4.1 Hasil Pendugaan Parameter dan Uji Signifikansi Parameter

Pada subbab ini akan dijelaskan mengenai hasil pendugaan parameter-parameter serta hasil uji signifikansi parameter dari model regresi logistik multilevel menggunakan metode PQL dan MLE. Untuk mengetahui peubah prediktor yang mempengaruhi peubah respon, maka perlu dilakukan pengujian signifikansi parameter yang telah didapat dari pendugaan parameter menggunakan metode PQL ataupun metode MLE. Dalam pengujian koefisien regresi ini dilakukan secara parsial (uji Wald).

Berikut hipotesis yang digunakan dalam melakukan pengujian secara parsial:

$$H_0 : \hat{\beta}_k = 0$$

Lawan

$$H_1 : \hat{\beta}_k \neq 0, \text{ untuk } k = 0, 1, 2, \dots, p.$$

#### 4.1.1 Data 1

Hasil pendugaan parameter-parameter serta hasil uji signifikansi parameter dari model regresi logistik multilevel menggunakan metode PQL dan MLE untuk data bakat siswa TK/RA yang disajikan dalam Tabel 4.1 sebagai berikut:

Tabel 4.1 Hasil Pendugaan Parameter (Data1)

Metode PQL				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-74,423	19,488	-3,819	<0,001
$\hat{\beta}_1$	0,169	0,058	2,914	0,0036
$\hat{\beta}_2$	0,053	0,025	2,12	0,034
$\hat{\beta}_3$	0,125	0,044	2,841	0,0045
$\hat{\beta}_4$	0,807	0,974	0,828	0,407
$\hat{\sigma}_{u_{0j}}^2$	9,886	8,572		

Tabel 4.1. Lanjutan

Metode MLE				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-82,7082	27,5875	-3,009	0,0026
$\hat{\beta}_1$	0,1859	0,07123	2,61	0,0091
$\hat{\beta}_2$	0,05972	0,02816	2,121	0,0339
$\hat{\beta}_3$	0,1407	0,06212	2,265	0,0235
$\hat{\beta}_4$	0,8901	1,0518	0,846	0,397
$\hat{\sigma}_{u_{0j}}^2$	12,4007	14,2819		

Berdasarkan hasil pada Tabel 4.1, maka didapat model umum regresi logistik multilevel menggunakan metode PQL untuk data bakat siswa TK/RA sebagai berikut:

Model Level 1:

$$\text{logit}(\pi_{ij}) = \beta_{0j} + 0,169X_{1ij} + 0,053X_{2ij} + 0,125X_{3ij} + 0,807X_{4ij} \quad 4.1$$

Model Level 2:

$$\beta_{0j} = -74,423 + u_{0j} \quad 4.2$$

Sehingga jika persamaan 4.2 disubstitusikan pada persamaan 4.1 akan menjadi model logistik dua level seperti pada persamaan 2.8 sebagai berikut:

$$\text{logit}(\pi_{ij}) = -74,423 + 0,169X_{1ij} + 0,053X_{2ij} + 0,125X_{3ij} + 0,807X_{4ij} + u_{0j} \quad 4.3$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap distrik akan memiliki nilai  $u_{0j}$  yang berbeda-beda.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:



$$ICC = \frac{9,886}{9,886 + 3,29} = 0,7503$$

Nilai tersebut merupakan korelasi antar wanita dalam satu distrik yang sama dan merupakan proporsi dari variansi antar distrik terhadap total variansi, yaitu sebesar 75,05%. Itu berarti bahwa 75,03% dari total variansi dapat diperhitungkan sebagai variansi antar sekolah.

Sedangkan model umum regresi logistik multilevel menggunakan metode MLE untuk bakat siswa TK/RA sebagai berikut:

Model Level 1:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 0,1859X_{1ij} + 0,05972X_{2ij} + 0,1407X_{3ij} \\ & + 0,8901X_{4ij} \end{aligned} \quad 4.4$$

Model Level 2:

$$\beta_{0j} = -82,7082 + u_{0j} \quad 4.5$$

Sehingga jika persamaan 4.5 disubstitusikan pada persamaan 4.4 akan menjadi model logistik dua level seperti pada persamaan 2.8 sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -82,7082 + 0,1859X_{1ij} + 0,05972X_{2ij} \\ & + 0,1407X_{3ij} + 0,8901X_{4ij} + u_{0j} \end{aligned} \quad 4.6$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap distrik akan memiliki nilai  $u_{0j}$  yang berbeda-beda.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:

$$ICC = \frac{12,4007}{12,4007 + 3,29} = 0,7903$$

Nilai tersebut merupakan korelasi antar siswa dalam satu sekolah yang sama dan merupakan proporsi dari variansi antar sekolah

terhadap total variansi, yaitu sebesar 79,03%. Itu berarti bahwa 79,03% dari total variansi dapat diperhitungkan sebagai variansi antar sekolah. Sehingga dengan nilai ICC sebesar 79,03% untuk metode MLE dan 75,03% untuk metode PQL, maka analisis multilevel perlu untuk digunakan.

Untuk uji Wald ini,  $H_0$  ditolak jika nilai statistik uji Wald ( $|W|$ ) lebih besar daripada nilai  $Z_{0,025}$  ( $|W| > 1,96$ ) atau jika nilai-p kurang dari  $\alpha$  (nilai-p  $< 0,05$ ). Berdasarkan Tabel 4.1 nilai  $|W|$  yang lebih besar dari 1,96 yaitu pada peubah pemilihan warna, peubah kerapian, serta peubah kebersihan dalam mewarnai. Hasil pengujian signifikansi parameter antara model dengan metode PQL dan menggunakan metode MLE didapat hasil yang sama. Hal tersebut juga dapat dibuktikan dengan membandingkan nilai-p yang lebih kecil dari 0,05, sehingga dapat diambil kesimpulan bahwa peubah yang mempengaruhi bakat seorang siswa TK/RA di Desa Maron adalah peubah pemilihan warna, peubah kerapian, serta peubah kebersihan dalam mewarnai pada taraf nyata 95%.

#### 4.1.2 Data 2

Hasil pendugaan parameter-parameter serta hasil uji signifikansi parameter dari model regresi logistik multilevel menggunakan metode PQL dan MLE untuk data nilai akreditasi SMK di Jawa Timur tahun 2006 disajikan dalam Tabel 4.2 sebagai berikut:

Tabel 4.2 Hasil Pendugaan Parameter (Data 2)

Metode PQL				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-7,448	2,673	-2,786	0,0053
$\hat{\beta}_1$	2,145	1,463	1,467	0,1426
$\hat{\beta}_2$	0,058	0,022	2,636	0,0084
$\hat{\beta}_3$	-0,000	0,001	0,000	1,0000

Tabel 4.2. Lanjutan

Metode PQL				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_4$	0,047	0,024	1,9583	0,0502
$\hat{\beta}_5$	-0,005	0,008	-0,625	0,532
$\hat{\beta}_6$	1,326	1,225	1,082	0,279
$\hat{\beta}_7$	0,138	0,11	1,255	0,2096
$\hat{\sigma}_{u_{oj}}^2$	0,466	0,538		
Metode MLE				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-7,446	2,6722	-2,786	0,0053
$\hat{\beta}_1$	2,126	1,5194	1,399	0,1617
$\hat{\beta}_2$	0,05804	0,02501	2,321	0,0203
$\hat{\beta}_3$	-0,00005	0,001292	-0,0387	0,969
$\hat{\beta}_4$	0,04659	0,02652	1,757	0,0789
$\hat{\beta}_5$	-0,00467	0,008228	-0,568	0,570
$\hat{\beta}_6$	1,3356	1,1968	1,116	0,264
$\hat{\beta}_7$	0,1382	0,114	1,212	0,225
$\hat{\sigma}_{u_{oj}}^2$	0,4679	0,7013		

Berdasarkan hasil pada Tabel 4.2, maka didapat model umum regresi logistik multilevel menggunakan metode PQL untuk data nilai akreditasi SMK di Jawa Timur tahun 2006 sebagai berikut:

Model Level 1:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 2,145X_{1ij} + 0,058X_{2ij} + (-0,000)X_{3ij} \\ & + 0,047X_{4ij} + (-0,005)X_{5ij} + 1,326X_{6ij} \\ & + 0,138X_{7ij} \end{aligned}$$

4.7

Model Level 2:

$$\beta_{0j} = -7,448 + u_{0j}$$

4.8

43

Sehingga jika persamaan 4.8 disubstitusikan pada persamaan 4.7 akan menjadi model logistik dua level seperti pada persamaan 2.8 sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -7,448 + 2,145X_{1ij} + 0,058X_{2ij} \\ & + (-0,000)X_{3ij} + 0,047X_{4ij} + (-0,005)X_{5ij} \\ & + 1,326X_{6ij} + 0,138X_{7ij} + u_{0j} \end{aligned} \quad 4.9$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap kota akan memiliki nilai  $u_{0j}$  yang berbeda-beda. Model regresi logistik multilevel tiap-tiap kota/kabupaten secara lengkap dapat dilihat pada Lampiran 5.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:

$$ICC = \frac{0,466}{0,466 + 3,29} = 0,1241$$

Nilai tersebut merupakan korelasi antar SMK dalam satu kota yang sama dan merupakan proporsi dari variansi antar kota terhadap total variansi, yaitu sebesar 12,41%. Itu berarti bahwa 12,41% dari total variansi dapat diperhitungkan sebagai variansi antar kota.

Untuk model umum regresi logistik multilevel menggunakan metode MLE untuk data nilai akreditasi SMK di Jawa Timur tahun 2006, berdasarkan hasil pada Tabel 4.2 maka didapat sebagai berikut:

Model Level 1:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 2,126X_{1ij} + 0,05804X_{2ij} \\ & + (-0,00005)X_{3ij} + 0,04659X_{4ij} \\ & + (-0,00467)X_{5ij} + 1,3356X_{6ij} + 0,1382X_{7ij} \end{aligned} \quad 4.10$$

Model Level 2:

$$\beta_{0j} = -7,446 + u_{0j} \quad 4.11$$

Sehingga jika persamaan 4.11 disubstitusikan pada persamaan 4.10 akan menjadi model logistik dua level seperti pada persamaan 2.8 sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -7,446 + 2,126X_{1ij} + 0,05804X_{2ij} \\ & + (-0,00005)X_{3ij} + 0,04659X_{4ij} \\ & + (-0,00467)X_{5ij} + 1,3356X_{6ij} + 0,1382X_{7ij} \\ & + u_{0j} \end{aligned} \quad 4.12$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap kota akan memiliki nilai  $u_{0j}$  yang berbeda-beda. Model regresi logistik multilevel tiap-tiap kota/kabupaten secara lengkap dapat dilihat pada Lampiran 6.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:

$$ICC = \frac{0,4679}{0,4679 + 3,29} = 0,1245$$

Nilai tersebut merupakan korelasi antar SMK dalam satu kota yang sama dan merupakan proporsi dari variansi antar kota terhadap total variansi, yaitu sebesar 12,45%. Itu berarti bahwa 12,45% dari total variansi dapat diperhitungkan sebagai variansi antar kota. Sehingga dengan nilai ICC sebesar 12,45% untuk metode MLE serta 12,41 untuk metode PQL, maka analisis multilevel perlu untuk digunakan.

Untuk uji Wald ini,  $H_0$  ditolak jika nilai statistik uji Wald ( $W$ ) lebih besar daripada nilai  $Z_{0,025}$  ( $|W| > 1,96$ ) atau jika nilai-p kurang dari  $\alpha$  (nilai-p  $< 0,05$ ). Berdasarkan Tabel 4.2 nilai  $|W|$  untuk pendugaan menggunakan metode PQL dan metode MLE yang lebih besar dari 1,96 yaitu pada parameter  $\beta_2$ . Hal tersebut juga dapat dibuktikan dengan membandingkan nilai-p dari  $\beta_2$  lebih kecil dari 0,05, sehingga dapat diambil kesimpulan bahwa peubah yang mempengaruhi nilai akreditasi SMK di Jawa Timur yaitu lama berdirinya sekolah (tahun) pada taraf nyata 95%.



### 4.1.3 Data 3

Hasil pendugaan parameter-parameter serta hasil uji signifikansi parameter dari model regresi logistik multilevel menggunakan metode PQL dan MLE untuk data 3 disajikan dalam Tabel 4.3 sebagai berikut:

Tabel 4.3 Hasil Pendugaan Parameter (Data 3)

Metode PQL				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-1,626	0,179	-9,0838	<0,001
$\hat{\beta}_1$	1,294	0,119	10,874	<0,001
$\hat{\beta}_2$	-0,011	0,006	-1,833	0,067
$\hat{\beta}_3$	0,65	0,105	6,19	<0,001
$\hat{\beta}_4$	0,705	0,091	7,747	<0,001
$\hat{\beta}_5$	0,409	0,13	3,146	0,0016
$\hat{\sigma}_{u_{oj}}^2$	0,208	0,062		
Metode MLE				
Parameter	Koefisien	SE	W	Nilai-p
$\hat{\beta}_0$	-1,6251	0,1798	-9,038	<0,001
$\hat{\beta}_1$	1,2933	0,1202	10,76	<0,001
$\hat{\beta}_2$	-0,01113	0,005589	-1,991	0,046
$\hat{\beta}_3$	0,6492	0,1054	6,159	<0,001
$\hat{\beta}_4$	0,7041	0,09163	7,684	<0,001
$\hat{\beta}_5$	0,4086	0,1301	3,141	0,0017
$\hat{\sigma}_{u_{oj}}^2$	0,2077	0,06372		

Berdasarkan hasil pada Tabel 4.3, maka didapat model umum regresi logistik multilevel untuk data hasil Bangladesh Fertility Survey tentang penggunaan alat kontrasepsi tahun 1989 sebagai berikut:

Model Level 1:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 1,294X_{1ij} + (-0,011)X_{2ij} + 0,65X_{3ij} + \\ & 0,705X_{4ij} + 0,409X_{5ij} \end{aligned} \quad 4.13$$

Model Level 2:

$$\beta_{0j} = -1,626 + u_{0j} \quad 4.14$$

Sehingga jika persamaan 4.14 disubstitusikan pada persamaan 4.13 akan menjadi model logistik dua level seperti pada persamaan 2.8 sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -1,626 + 1,294X_{1ij} + (-0,011)X_{2ij} + 0,65X_{3ij} \\ & + 0,705X_{4ij} + 0,409X_{5ij} + u_{0j} \end{aligned} \quad 4.5$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap distrik akan memiliki nilai  $u_{0j}$  yang berbeda-beda.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:

$$ICC = \frac{0,208}{0,208 + 3,29} = 0,0595$$

Nilai tersebut merupakan korelasi antar wanita dalam satu distrik yang sama dan merupakan proporsi dari variansi antar distrik terhadap total variansi, yaitu sebesar 5,95%. Itu berarti bahwa 5,95% dari total variansi dapat diperhitungkan sebagai variansi antar distrik.

Untuk model umum regresi logistik multilevel menggunakan metode MLE untuk data hasil survei penggunaan alat kontrasepsi di Bangladesh tahun 1989 berdasarkan hasil pada Tabel 4.3 maka didapat sebagai berikut:

Model Level 1:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 1,2933X_{1ij} + (-0,01113)X_{2ij} + \\ & 0,6492X_{3ij} + 0,7041X_{4ij} + 0,4086X_{5ij} \end{aligned} \quad 4.16$$

Model Level 2:

$$\beta_{0j} = -1,6251 + u_{0j} \quad 4.17$$

Sehingga jika persamaan 4.17 disubstitusikan pada persamaan 4.16 akan menjadi model logistik dua level sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -1,6251 + 1,2933X_{1ij} + (-0,01113)X_{2ij} \\ & + 0,6492X_{3ij} + 0,7041X_{4ij} + 0,4086X_{5ij} + u_{0j} \end{aligned} \quad 4.18$$

Dengan nilai  $u_{0j}$  merupakan *random effect* yang nantinya akan dihitung seperti pada persamaan 2.9 dan setiap distrik akan memiliki nilai  $u_{0j}$  yang berbeda-beda.

Untuk nilai *Intra-Class Correlation* (ICC) sesuai persamaan 2.10, maka didapat nilai sebagai berikut:

$$ICC = \frac{0,2077}{0,2077 + 3,29} = 0,0594$$

Nilai tersebut merupakan korelasi antar wanita dalam satu distrik yang sama dan merupakan proporsi dari variansi antar distrik terhadap total variansi, yaitu sebesar 5,94%. Itu berarti bahwa 5,94% dari total variansi dapat diperhitungkan sebagai variansi antar distrik. Sehingga dengan nilai ICC sebesar 5,94% untuk metode MLE dan 5,94% untuk metode PQL, maka analisis multilevel perlu untuk digunakan.

Pada uji Wald ini,  $H_0$  ditolak jika nilai statistik uji Wald ( $|W|$ ) lebih besar daripada nilai  $Z_{0,025}$  ( $|W| > 1,96$ ) atau jika nilai-p kurang dari  $\alpha$  (nilai-p < 0,05). Berdasarkan Tabel 4.3 nilai  $|W|$  yang lebih besar dari 1,96 untuk metode PQL dan metode MLE terdapat perbedaan, yaitu pada parameter  $\beta_2$  di mana pada metode PQL tidak berpengaruh nyata sedangkan untuk metode MLE memiliki pengaruh yang nyata. Hal tersebut juga dapat dibuktikan dengan membandingkan nilai-p yang lebih kecil dari 0,05, sehingga dapat diambil kesimpulan bahwa peubah yang mempengaruhi seorang wanita dalam penggunaan alat kontrasepsi di Bangladesh adalah

peubah status memiliki anak, peubah umur (tahun), peubah daerah tinggal, peubah tingkat pendidikan, serta peubah agama pada taraf nyata 95%.

#### 4.2 Pegujian Kesesuaian Model

Digunakan uji Pearson untuk uji kesesuaian model kali ini, dengan hipotesis sebagai berikut:

$H_0$  : model sesuai dengan data

$H_1$  : model tidak sesuai dengan data

Berikut hasil uji Pearson untuk tiga data yang dianalisis:

Tabel 4.4 Hasil Uji Kesesuaian Model

	Metode PQL			Metode MLE		
	$\chi^2_{pearson}$	db	<i>p-value</i>	$\chi^2_{pearson}$	db	<i>p-value</i>
Data 1	63,08	74	0,8134	78,14	74	0,3487
Data 2	87,76	91	0,5768	87,59	91	0,5818
Data 3	2771,45	2705	0,1827	2769,6	2705	0,1893

Berdasarkan uji Pearson, bahwa nilai-p untuk ketiga data dan kedua metode bernilai lebih besar dari  $\alpha$ , sehingga  $H_0$  diterima. Dapat disimpulkan bahwa model yang terbentuk dari hasil pendugaan parameter untuk kedua metode sesuai dengan data.

#### 4.3 Pemilihan Model Terbaik

##### 4.3.1 Data 1

Dalam pemilihan model terbaik, digunakan nilai APER seperti pada Persamaan 2.39 sebagai tolok ukur. Nilai APER akan menyatakan besar proporsi sampel yang salah diklasifikasikan sesuai Tabel 2.1. Berikut nilai APER untuk data 1:

Tabel 4.5 Perhitungan Nilai APER Metode PQL (Data 1)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	33	8	41
	1	0	38	38

Total	33	46	79
-------	----	----	----

Berdasarkan Tabel 4.5, nilai APER untuk model pada Persamaan 4.3 yang diduga menggunakan metode PQL dapat dihitung yaitu sebesar:

$$APER_{PQL} = \frac{8 + 0}{79} \times 100\% = 10,13\%$$

Untuk nilai APER model menggunakan metode MLE dijelaskan pada Tabel 4.6 berikut:

Tabel 4.6 Perhitungan Nilai APER Metode MLE (Data 1)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	33	8	41
	1	0	38	38
Total		33	46	79

Berdasarkan Tabel 4.6, nilai APER untuk model pada Persamaan 4.6 yang diduga menggunakan metode MLE dapat dihitung yaitu sebesar:

$$APER_{MLE} = \frac{8 + 0}{79} \times 100\% = 10,13\%$$

Dari dua nilai APER yang telah dihitung, maka selanjutnya nilai tersebut dibandingkan satu sama lain untuk mengetahui model mana yang lebih baik. Nilai APER yang lebih kecil mengindikasikan bahwa model tersebut yang lebih baik. Nilai APER dari dua model memiliki nilai yang sama yaitu 10,13% ( $APER_{PQL} = APER_{MLE} = 10,13\%$ ), sehingga disimpulkan bahwa tidak bisa dikatakan model yang dihasilkan dari pendugaan menggunakan metode PQL lebih baik dari model yang dihasilkan dari pendugaan menggunakan metode MLE dan sebaliknya.

#### 4.3.2 Data 2



Nilai APER untuk model data 2 hasil pendugaan menggunakan metode PQL dan MLE akan dijelaskan pada Tabel 4.7 dan Tabel 4.8 berikut:

Tabel 4.7 Perhitungan Nilai APER Metode PQL (Data 2)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	48	18	66
	1	13	20	33
Total		61	38	99

Berdasarkan Tabel 4.7, nilai APER untuk model pada Persamaan 4.9 yang diduga menggunakan metode PQL dapat dihitung yaitu sebesar:

$$APER_{PQL} = \frac{18 + 13}{99} \times 100\% = 31,31\%$$

Untuk nilai APER model pada Persamaan 4.12 dijelaskan pada Tabel 4.8 berikut:

Tabel 4.8 Perhitungan Nilai APER Metode MLE (Data 2)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	48	18	66
	1	13	20	33
Total		61	38	99

Berdasarkan Tabel 4.8, nilai APER untuk model pada Persamaan 4.12 yang diduga menggunakan metode MLE dapat dihitung yaitu sebesar:

$$APER_{MLE} = \frac{18 + 13}{99} \times 100\% = 31,31\%$$

Dari dua nilai APER yang telah dihitung, maka selanjutnya nilai tersebut dibandingkan satu sama lain untuk mengetahui model mana yang lebih baik. Nilai APER yang lebih kecil mengindikasikan

bahwa model tersebut yang lebih baik. Nilai APER dari dua model memiliki nilai yang sama yaitu 31,31% ( $APER_{PQL} = APER_{MLE} = 31,31\%$ ), sehingga disimpulkan bahwa tidak bisa dikatakan model yang dihasilkan dari pendugaan menggunakan metode PQL lebih baik dari model yang dihasilkan dari pendugaan menggunakan metode MLE dan sebaliknya.

### 4.3.3 Data 3

Berikut nilai APER untuk model untuk data 3 hasil pendugaan menggunakan metode PQL dan MLE akan dijelaskan pada Tabel 4.9 dan Tabel 4.10.

Tabel 4.9 Perhitungan Nilai APER Metode PQL (Data 3)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	741	886	1627
	1	189	895	1084
Total		930	1781	2711

Berdasarkan Tabel 4.9, nilai APER untuk model pada Persamaan 4.15 yang diduga menggunakan metode PQL dapat dihitung yaitu sebesar:

$$APER_{PQL} = \frac{886 + 189}{2711} \times 100\% = 39,65\%$$

Untuk nilai APER model pada Persamaan 4.18 dijelaskan pada Tabel 4.10 berikut:

Tabel 4.10 Perhitungan Nilai APER Metode MLE (Data 3)

Observasi		Prediksi		Total
		Y		
		0	1	
Y	0	741	886	1627
	1	190	894	1084
Total		931	1780	2711

Berdasarkan Tabel 4.10, nilai APER untuk model pada Persamaan 4.18 yang diduga menggunakan metode MLE dapat dihitung yaitu sebesar:

$$APER_{MLE} = \frac{886 + 190}{2711} \times 100\% = 39,69\%$$

Dari dua nilai APER yang telah dihitung, maka selanjutnya nilai tersebut dibandingkan satu sama lain untuk mengetahui model mana yang lebih baik. Nilai APER yang lebih kecil mengindikasikan bahwa model tersebut yang lebih baik. Nilai APER dari model 4.9 sebesar 39,65% ( $APER_{PQL} = 39,65\%$ ), sedangkan untuk model 4.18 memiliki nilai APER sebesar 39,69% ( $APER_{MLE} = 39,69\%$ ). Terdapat perbedaan nilai APER untuk kedua model, namun selisih nilai APER sangatlah sedikit sehingga disimpulkan bahwa tidak bisa dikatakan model yang dihasilkan dari pendugaan menggunakan metode PQL lebih baik dari model yang dihasilkan dari pendugaan menggunakan metode MLE dan sebaliknya.

Tabel 4.11 Ringkasan Nilai APER

	Nilai APER	
	Metode PQL	Metode MLE
Data 1	10,13%	10,13%
Data 2	31,31%	31,31%
Data 3	39,65%	39,69%

Berdasarkan hasil nilai APER untuk tiga data yang telah dianalisis, didapatkan bahwa pendugaan parameter menggunakan metode PQL dan metode MLE memiliki nilai APER yang sama. Adapun selisih nilai APER pada data 3 tidaklah berpengaruh terhadap keakuratan karena memang selishnya sangatlah kecil. Sehingga dapat diambil kesimpulan bahwa pendugaan parameter menggunakan metode PQL dan metode MLE memiliki nilai keakuratan yang sama.

Tetapi pada dasarnya, metode PQL lebih mudah dilakukan dari pada metode MLE karena pada proses pendugaan parameter

menggunakan fungsi iterative, metode PQL dilakukan pada fungsi yang linier sedangkan pada metode MLE fungsi yang digunakan bersifat nonlinier. Selain itu, secara komputasi metode PQL lebih cepat dijalankan jika dibandingkan dengan metode MLE pada saat dijalankan menggunakan program. Dapat dilihat dari jumlah iterasi dan lama proses analisis pada *software* pada lampiran 7, 8, dan 9 untuk output dari metode PQL dan lampiran 10, 11, dan 12 untuk output metode MLE. Untuk lebih lengkap akan ditampilkan dalam bentuk Tabel berikut:

Tabel 4.12 Jumlah Iterasi

	Jumlah Iterasi	
	Metode PQL	Metode MLE
Data 1	14	27
Data 2	10	17
Data 3	5	12

Berdasarkan Tabel 4.12, dapat dijelaskan metode PQL memiliki jumlah iterasi lebih sedikit untuk ketiga data yang dianalisis yaitu 14 kali iterasi untuk data 1, 15 kali iterasi untuk data 2, dan untuk data 3 sebanyak 8 kali iterasi. Sedangkan untuk metode MLE terjadi 24 kali iterasi untuk data 1, 17 kali iterasi untuk data 2, serta 12 kali iterasi untuk data 3.

Selain dari jumlah iterasi, waktu komputasi pada setiap metode juga dicatat pada Tabel berikut:

Tabel 4.13 Waktu Komputasi Pada *Software*

	Waktu Proses Analisis (detik)	
	Metode PQL	Metode MLE
Data 1	1,4	13
Data 2	1,3	4,1
Data 3	0,9	17,5

Untuk waktu komputasi sesuai Tabel 4.13, *software* yang digunakan yaitu *software* MLwiN 2.30 untuk metode PQL dan *software* SAS 9.3 (PROC NL MIXED) untuk metode MLE. Dapat

dijelaskan bahwa metode PQL membutuhkan waktu 1,4 detik untuk proses analisis data 1, 1,3 detik untuk data 2, serta 0,9 detik untuk data 3. Sedangkan untuk metode MLE membutuhkan waktu yang lebih lama yaitu, 13 detik untuk analisis data 1, 4,1 detik untuk data 2, serta 17,5 detik untuk data 3. Meskipun berdasarkan Tabel 4.16 metode PQL memiliki waktu yang lebih cepat daripada metode MLE, namun sesuai *software* yang digunakan memang *software* MLwiN 2.30 merupakan *software* yang hanya mengakomodir analisis data multilevel, sedangkan *software* SAS 9.3 merupakan *software* yang tersedia untuk berbagai macam analisis statistika.

#### 4.4 Interpretasi Model

Dalam melakukan interpretasi model, seharusnya dipilih model yang lebih baik. Karena model hasil pendugaan menggunakan metode PQL dan menggunakan metode MLE memiliki nilai APER yang sama, maka kedua model dapat digunakan dalam melakukan interpretasi model. Namun dalam interpretasi model ini, digunakan model hasil pendugaan menggunakan metode PQL. Dalam interpretasi model, terdapat dua cara yang digunakan, yaitu menggunakan nilai peluang dan nilai *Odds ratio*.

##### 4.4.1 Nilai Peluang

- **Data 1**

Model hasil pendugaan menggunakan metode PQL untuk data 1 sudah dituliskan pada persamaan 4.3 yaitu:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & -74,423 + 0,169X_{1ij} + 0,053X_{2ij} + 0,125X_{3ij} \\ & + 0,807X_{4ij} + u_{0j} \end{aligned}$$

Dari model yang terbentuk masih terdapat *random effect* ( $u_{0j}$ ) yang nilainya berbeda pada tiap sekolah. Untuk memudahkan dalam interpretasi maka diambil contoh dua sekolah yaitu RA Al Hidayah I ( $u_{01} = 6,342$ ) serta TK Indira Tama ( $u_{02} = 3,45$ ). Peluang (seperti



persamaan 2.10) seorang siswa di RA Al Hidayah I dikatakan berbakat dalam mewarnai dengan skor pemilihan warna ( $X_1=200$ ), skor kerapian ( $X_2=200$ ), serta skor kebersihan ( $X_3=200$ ) adalah 0,789 ( $\pi_{i1} = 0,789$ ). Sedangkan peluang seorang siswa dengan kondisi yang sama dari TK Indira Tama adalah lebih kecil yaitu 0,172 ( $\pi_{i1} = 0,172$ ).

- **Data 2**

Model hasil pendugaan menggunakan metode PQL untuk data 2 sudah dituliskan pada persamaan 4.9 yaitu:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 2,145X_{1ij} + 0,058X_{2ij} + (-0,000)X_{3ij} \\ & + 0,047X_{4ij} + (-0,005)X_{5ij} + 1,326X_{6ij} \\ & + 0,138X_{7ij} \end{aligned}$$

Model yang terbentuk masih terdapat *random effect* ( $u_{0j}$ ) yang nilainya berbeda pada tiap kota/kab. Diambil contoh dua kota/kab. dalam melakukan interpretasi, Kota Surabaya ( $u_{01} = 0,833$ ) serta Kab. Pacitan ( $u_{02} = 0,199$ ). Sehingga dapat dihitung nilai peluang suatu sekolah di Kota Surabaya mendapat nilai Akreditasi A dengan status sekolah adalah swasta ( $X_1=0$ ), lama berdiri 25 tahun ( $X_2=25$ ), banyak siswa 260 siswa ( $X_3=260$ ), banyak guru 30 orang ( $X_4=30$ ), banyak alumni yang diterima di dunia usaha dan industri sebanyak 40 orang ( $X_5=40$ ), status bangunan milik sendiri ( $X_6=1$ ), serta nilai rata-rata UN sebesar 20 ( $X_7=20$ ) yaitu sebesar 0,533 ( $\pi_{i20} = 0,533$ ). Sedangkan peluang suatu sekolah mendapatkan nilai akreditasi A di Kab. Pacitan dengan kondisi yang sama dengan Kota Surabaya didapat lebih kecil yaitu sebesar 0,377 ( $\pi_{i16} = 0,377$ ).

- **Data 3**

Model hasil pendugaan menggunakan metode PQL untuk data 3 sudah dituliskan pada persamaan 4.15 yaitu:

$$\text{logit}(\pi_{ij}) = -1,626 + 1,294X_{1ij} + (-0,011)X_{2ij} + 0,65X_{3ij} \\ + 0,705X_{4ij} + 0,409X_{5ij} + u_{0j}$$

Karena model yang terbentuk masih terdapat *random effect* ( $u_{0j}$ ) yang nilainya berbeda pada tiap distrik, maka diambil contoh dua distrik dalam melakukan interpretasi model yaitu Distrik 4 ( $u_{04} = 0,821$ ) dan Distrik 44 ( $u_{044} = 0,256$ ). Sehingga dapat dihitung nilai peluang seorang wanita di Distrik 4 menggunakan alat kontrasepsi dengan status memiliki anak adalah sudah memiliki anak ( $X_1=1$ ), umur 35 tahun ( $X_2=35$ ), daerah tinggal adalah pedesaan ( $X_3=0$ ), status pendidikan yaitu tidak pernah bersekolah ( $X_4=0$ ), serta beragama Muslim ( $X_5=0$ ) sebesar  $0,525$  ( $\pi_{i4} = 0,525$ ) dan lebih besar dari nilai peluang suatu wanita menggunakan alat kontrasepsi di Distrik 44 yang bernilai  $0,368$  ( $\pi_{i44} = 0,368$ ) dengan kondisi yang sama.

#### 4.5.2 Odds ratio

- **Data 1**

Berdasarkan persamaan regresi logistik multilevel pada persamaan 4.3, maka terbentuk nilai *Odds ratio* untuk masing-masing peubah prediktor sebagai berikut:

Tabel 4.14 Nilai *Odds ratio* dengan Metode PQL (Data 1)

Peubah Prediktor	<i>Odds ratio</i>
Skor pemilihan warna ( $X_1$ )	1,184
Skor kerapian ( $X_2$ )	1,054
Skor kebersihan ( $X_3$ )	1,133
Jenis kelamin ( $X_4$ )	2,241

Interpretasi dari nilai *odds ratio* pada Tabel 4.14 adalah:

1. Nilai *odds ratio* untuk peubah skor pemilihan warna didapat sebesar 1,184 yang berarti bahwa setiap kenaikan satu satuan skor pemilihan warna, maka peluang seorang siswa TK/RA dikatakan berbakat dalam mewarnai meningkat 1,184 kali.

2. *Odds ratio* pada peubah skor kerapian bernilai 1,054 dapat diartikan bahwa setiap kenaikan satu satuan skor kerapian suatu siswa TK/RA dapat meningkatkan peluang suatu siswa TK/RA dikatakan berbakat dalam mewarnai sebesar 1,054 kali.
3. Dengan nilai *odds ratio* peubah skor kebersihan bernilai 1,133, dapat diartikan bahwa peluang suatu siswa TK/RA meningkat sebesar 1,133 kali dengan meningkatnya skor kebersihan sebesar satu satuan.
4. Pada peubah jenis kelamin, nilai *odds ratio* sebesar 2,241 yang berarti bahwa peluang siswa TK/RA laki-laki berpeluang dikatakan berbakat dalam mewarnai 2,241 kali daripada siswa TK/RA dengan jenis kelamin perempuan.

- **Data 2**

Berdasarkan persamaan regresi logistik multilevel pada persamaan 4.9, maka terbentuk nilai *odds ratio* untuk masing-masing peubah prediktor sebagai berikut:

Tabel 4.15 Nilai *Odds ratio* dengan Metode PQL (Data 2)

Peubah Prediktor	<i>Odds ratio</i>
Status Sekolah ( $X_1$ )	8,542
Lama sekolah berdiri ( $X_2$ )	1,06
Banyak siswa aktif ( $X_3$ )	1
Banyak guru aktif ( $X_4$ )	1,048
Banyak alumni bekerja ( $X_5$ )	0,995
Status tanah bangunan ( $X_6$ )	3,766
Nilai rata-rata UN ( $X_7$ )	1,148

Interpretasi dari nilai *odds ratio* pada Tabel 4.15 adalah:

1. Nilai *odds ratio* untuk peubah status sekolah bernilai 8,542, sehingga dapat diartikan bahwa peluang suatu sekolah dengan status negeri mendapat nilai akreditasi A 8,542 kali lebih besar daripada sekolah dengan status sekolah swasta.
2. Untuk peubah lama berdiri sekolah, nilai *odds ratio* yaitu sebesar 1,06 dan dapat diartikan bahwa setiap kenaikan satu tahun

berdirinya sekolah, peluang mendapatkan nilai akreditasi A meningkat 1,06 kali.

3. Dengan nilai *odds ratio* untuk peubah banyak siswa aktif sebesar 1, maka dengan jumlah siswa aktif yang berbeda peluang suatu sekolah mendapat nilai akreditasi A adalah sama.
4. Nilai *odds ratio* untuk peubah banyak guru aktif sebesar 1,048 yang berarti bahwa setiap bertambahnya satu orang guru di suatu sekolah dapat meningkatkan peluang mendapat nilai akreditasi A sebesar 1,048 kali.
5. Nilai *odds ratio* untuk peubah banyaknya alumni yang diserap dunia usaha bernilai 0,995 yang berarti bahwa setiap bertambah satu orang alumni yang diserap dunia usaha maka peluang sekolah mendapat nilai akreditasi A menurun 0,995 kali.
6. Peubah status tanah bangunan memiliki nilai *odds ratio* sebesar 3,776 yang berarti bahwa sekolah dengan status bangunan milik sendiri memiliki peluang mendapat nilai akreditasi A lebih besar 3,776 kali dari sekolah dengan status tanah bangunan bukan milik sendiri.
7. Peubah rata-rata nilai UN memiliki nilai *odds ratio* sebesar 1,148 yang berarti setiap kenaikan satu satuan nilai rata-rata UN dapat meningkatkan peluang mendapatkan nilai akreditasi A sebesar 1,148 kali.

### • Data 3

Berdasarkan persamaan regresi logistik multilevel pada persamaan 4.15, maka terbentuk nilai *odds ratio* untuk masing-masing peubah prediktor sebagai berikut:

Tabel 4.16 Nilai *Odds ratio* dengan Metode PQL (Data 3)

Peubah Prediktor	<i>Odds ratio</i>
Status memiliki anak ( $X_1$ )	3,647
Umur ( $X_2$ )	0,989
Daerah tinggal ( $X_3$ )	1,916
Status pendidikan ( $X_4$ )	2,024

Agama ( $X_5$ )	1,505
-----------------	-------

Interpretasi dari nilai *odds ratio* pada Tabel 4.17 adalah:

1. Nilai *odds ratio* untuk peubah status memiliki anak sebesar 3,647 yang berarti bahwa seorang yang memiliki anak berpeluang menggunakan alat kontrasepsi 3,647 kali daripada wanita yang belum memiliki anak.
2. Nilai *odds* untuk peubah umur bernilai 0,989, berarti bahwa setiap bertambahnya satu tahun umur seorang wanita, maka peluang menggunakan alat kontrasepsi menurun 0,989 kali.
3. Pada peubah daerah tinggal nilai *odds ratio* sebesar 1,916. Berarti bahwa seorang wanita yang tinggal di perkotaan berpeluang menggunakan alat kontrasepsi 1,916 kali daripada seorang wanita yang tinggal di pedesaan.
4. Nilai *odds ratio* untuk peubah status pendidikan yaitu sebesar 2,024. Berarti bahwa seorang wanita yang pernah bersekolah memiliki peluang untuk menggunakan alat kontrasepsi 2,024 kali daripada seorang wanita yang tidak pernah bersekolah.
5. Untuk peubah agama didapatkan nilai *odds ratio* sebesar 1,505 yang berarti bahwa seorang wanita yang beragama Hindu memiliki peluang menggunakan alat kontrasepsi 1,505 kali daripada wanita yang beragama Muslim.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil penelitian, dapat diambil kesimpulan:

1. Dari tiga data yang dianalisis, didapatkan nilai APER untuk model hasil pendugaan menggunakan metode PQL dan metode MLE memiliki nilai yang sama yaitu pada data 1 dan data 2. Pada data 3 terdapat perbedaan nilai APER yang sangat kecil, sehingga dapat disimpulkan bahwa berdasarkan nilai APER pendugaan parameter menggunakan metode PQL dan metode MLE memiliki keakuratan yang sama.
2. Secara komputasi metode PQL lebih cepat dibandingkan dengan metode MLE pada saat dijalankan menggunakan program untuk ketiga data, meskipun pada proses komputasi *software* yang digunakan merupakan *software* yang berbeda. Dapat dilihat juga dari jumlah iterasi dari proses komputasi melalui program bahwa metode PQL memiliki jumlah iterasi lebih sedikit daripada metode MLE untuk ketiga data yang dianalisis.

#### 5.2 Saran

Saran yang dapat diberikan pada skripsi ini yaitu:

1. Disarankan untuk menggunakan data dengan jumlah observasi dan jumlah kelompok yang bervariasi lagi untuk melihat kestabilan metode PQL dan metode MLE.
2. Untuk penelitian selanjutnya, dalam menduga parameter model regresi logistik multilevel dapat digunakan metode PQL order kedua (PQL-2) atau metode *Marginal Quasi Likelihood* baik order pertama (MQL-1) ataupun order kedua (MQL-2). Model regresi logistik multilevel juga dapat dikembangkan dengan peubah respon polikotomus (ordinal dan nominal).

UNIVERSITAS BRAWIJAYA



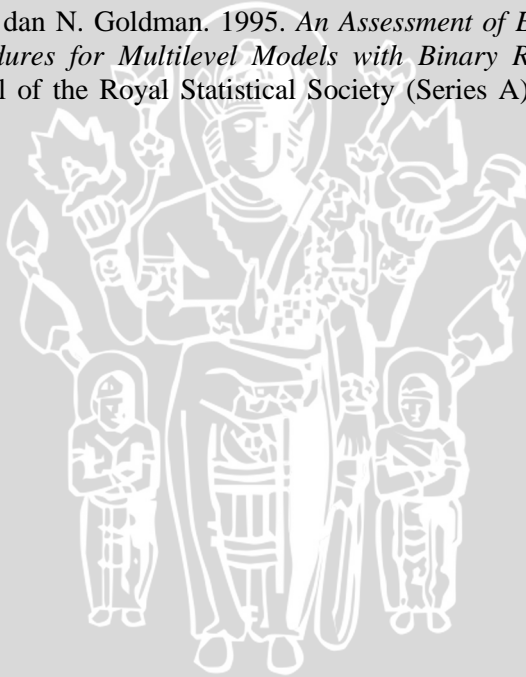
## DAFTAR PUSTAKA

- Dewi, A. 2008. *Estimasi Parameter Pada Model Regresi Logistik 2-Level*. Skripsi S1 Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia. Tidak dipublikasikan.
- Flom, P. L., J. M. Mc Mahon, dan E. R. Pouget. 2006. *Using PROC NLMIXED and PROC GLMMIX to Analyze Dyadic Data With Binary Outcomes*.  
<http://www.resug.org/proceedings/nesug06/an/da08.pdf>.  
Diakses tanggal 27 Desember 2013.
- Goldstein, H. 1999. *Multilevel Statistical Models*. Arnold Publisher. London.
- Goldstein, H. 2007. *Multilevel Modelling*. [www.cmm.bris.ac.uk](http://www.cmm.bris.ac.uk).  
Diakses tanggal 24 Desember 2013.
- Hosmer, D. W. dan S. Lemeshow. 2000. *Applied Logistic Regression*. John Wiley and Sons, Inc. New York.
- Hox, J. 2002. *Multilevel Analysis Techniques and Applications*. Lawrence Erlbaum Associates, Inc., Publisher. New Jersey.
- Huq, N.M. dan J. Cleland. 1990. *Bangladesh Fertility Survey 1989*. National Institute of Population Research and Training (NIPORT). Dhaka.
- Johnson, R. A. dan D. W. Wichern. 2007. *Applied Multivariate Statistical Analysis. Sixth Edition*. Prentice Hall International Inc. New Jersey.

Khan, M. H. R. dan J. E. H. Shaw. 2011. *Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data*. Journal of Data Science 9: 93-110.

Nazhiirah, A. 2010. *Perbandingan Pendugaan Parameter Maximum Likelihood dan Bayesian Markov Chain Monte Carlo (MCMC) pada Model Regresi Logistik Multilevel dengan Peubah Respon Biner*. Skripsi S1 Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya. Tidak dipublikasikan.

Rodriguez, G. dan N. Goldman. 1995. *An Assessment of Estimation Procedures for Multilevel Models with Binary Responses*. Journal of the Royal Statistical Society (Series A) 158: 73-89.



Lampiran 1. Data Bakat Siswa TK/RA di Desa Maron Kec. Srengat Dalam Mewarnai Tahun 2009 (Data 1)

NAMA	TK/RA	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	Y
Iqbal	Al Hidayah I	255	230	200	1	1
Tia	Al Hidayah I	250	215	195	0	1
Bimo	Al Hidayah I	245	215	195	1	1
Debi	Al Hidayah I	255	195	190	0	1
Roziq	Al Hidayah I	245	215	190	1	1
Linda	Al Hidayah I	245	195	195	0	1
Sita	Al Hidayah I	240	195	180	0	1
Yoga	Al Hidayah I	230	210	165	1	1
Ima	Al Hidayah I	245	205	170	0	1
Anggi	Al Hidayah I	250	195	165	0	1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
Dewi	Al Hidayah I	215	120	140	0	0
Dewi	Al Hidayah I	170	120	130	0	0
Angel	Al Hidayah I	190	60	140	0	0
Tegar	Al Hidayah I	120	90	140	1	0
Agung	Indria Tama	255	210	170	1	1
Yona	Indria Tama	240	210	180	0	1
Vanaesa	Indria Tama	240	220	190	0	1
Willy	Indria Tama	250	180	170	1	1
Yossy	Indria Tama	250	150	180	1	1
Tegar	Indria Tama	250	170	140	1	0
Edo	Indria Tama	225	170	170	1	0
Naswa	Indria Tama	210	120	170	0	0
Akbar	Indria Tama	210	120	165	1	0
Samuel	Indria Tama	220	100	130	1	0
Inez	Indria Tama	205	110	120	0	0



Lampiran 1. Lanjutan

NAMA	TK/RA	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	Y
Cindy	Indria Tama	250	245	195	0	1
Wulan	Indria Tama	255	230	200	0	1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
Angel	Indria Tama	245	215	195	0	1
Angga	Indria Tama	250	200	180	1	0
Difa	Indria Tama	250	200	180	0	0
Anggi	Al Hidayah II	270	245	235	0	1
Moh. Riko Bima	Al Hidayah II	245	215	235	1	1
Yuda	Al Hidayah II	270	205	225	1	1
Dhona	Al Hidayah II	245	230	225	0	1
Weyke	Al Hidayah II	255	205	215	0	1
Nova	Al Hidayah II	245	210	220	0	1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
Mamat	Al Hidayah II	260	180	170	1	0
Ratna	Al Hidayah II	240	180	180	0	0
Yusuf	Al Hidayah II	220	180	180	1	0
Nizar	Al Hidayah II	225	170	140	1	0
Bella	Al Hidayah II	200	120	130	0	0

Keterangan:

Y = Bakat siswa; 0 = kurang berbakat, 1 = berbakat

X<sub>1</sub> = Skor Pemilihan warna

X<sub>2</sub> = Skor Kerapian

X<sub>3</sub> = Sko Kebersihan

X<sub>4</sub> = Jenis kelamin; 0 = Perempuan, 1 = Laki-laki

Lampiran 2. Data Nilai Akreditasi SMK di Jawa Timur Tahun 2006  
(Data 2)

<b>Nama Jurusan SMK</b>	<b>Kota/ Kab.</b>	<b>X<sub>1</sub></b>	<b>X<sub>2</sub></b>	<b>X<sub>3</sub></b>	<b>X<sub>4</sub></b>	<b>X<sub>5</sub></b>	<b>X<sub>6</sub></b>	<b>X<sub>7</sub></b>	<b>Y</b>
Mekanika Otomotif SMK PGRI 1 Banyuwangi	1	0	30	114	30	20	1	21,34	0
Akuntansi SMK 17 Agustus Banyuwangi	1	0	37	488	25	35	1	20,54	1
Teknik Otomotif SMK Muh. 2 Genteng	1	0	31	734	47	97	1	18,67	1
Mekanik Otomotif SMK Muh. 6 Rogojampi Banyuwangi	1	0	14	990	39	80	1	20,01	0
Akuntansi SMK Muh. 1 Genteng Banyuwangi	1	0	39	464	43	131	1	21,61	0
Mekanik Otomotif SK PGRI 1 Sutojayan Kab. Blitar	2	0	23	213	15	60	1	21,53	0
Tata Busana SMK PGRI Kesamben Kab. Blitar	2	0	17	40	12	8	0	20,1	0
Teknik Informatika SMK Bima Bojonegoro	3	0	5	20	14	20	1	19,74	0
Akuntansi SMK Madinatul Ulum Bojonegoro	3	0	11	82	16	11	1	15,47	0
Admin PKT SMK Muh. 2 Gresik	4	0	15	66	18	5	1	19,66	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.

Lampiran 2. Lanjutan

<b>Nama Jurusan SMK</b>	<b>Kota/ Kab.</b>	<b>X<sub>1</sub></b>	<b>X<sub>2</sub></b>	<b>X<sub>3</sub></b>	<b>X<sub>4</sub></b>	<b>X<sub>5</sub></b>	<b>X<sub>6</sub></b>	<b>X<sub>7</sub></b>	<b>Y</b>
Nautika Niaga SMK Wira Maritim Surabaya	20	0	6	60	18	33	1	21,89	0
Mekanik Otomotif SMKN 1 Trenggalek	21	1	7	386	68	35	1	21,56	1
Akuntansi SMK Karya Dharma 2 Kab. Trenggalek	21	0	24	184	29	40	1	21,67	0
Admin PKT SMK Muh. 1 Trenggalek	21	0	18	127	19	61	1	20,55	0
Teknik Las SMK Qomarul Hidayah 1 Togu Kab. Trenggalek	21	0	22	271	24	0	0	17,55	0
Admin PKT SMK Karya Dharma 2 Kab. Trenggalek	21	0	24	76	29	40	1	18,09	1

Keterangan :

Y : Nilai akreditasi; 1 = Nilai A, 0 = Nilai selain A

X<sub>1</sub> : Status sekolah; 1 = Sekolah negeri, 0 = Sekolah swasta

X<sub>2</sub> : Lama berdiri sekolah (tahun)

X<sub>3</sub> : Banyaknya siswa aktif

X<sub>4</sub> : Banyaknya guru aktif

X<sub>5</sub> : Banyaknya alumni yang diterima di dunia usaha setahun terakhir

X<sub>6</sub> : Status tanah bangunan; 1 = Milik sendiri, 0=Bukan milik sendiri

X<sub>7</sub> : Nilai rata-rata Ujian Nasional Sekolah

Lampiran 3. Data Bangladesh *Fertility Survey* (Dhaka) Tahun 1989  
(Data 3)

Individu	Distrik	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	1	0	3	48	1	1	0
2	1	1	0	41	1	3	1
3	1	0	0	24	1	1	1
4	1	0	2	31	1	2	0
5	1	0	3	38	1	1	0
6	1	0	0	16	1	1	0
7	1	0	0	18	1	1	0
8	1	0	3	48	1	1	0
9	1	1	0	17	1	3	0
10	1	0	3	26	1	1	0
11	1	0	1	24	1	1	0
12	1	0	0	26	1	4	0
13	1	1	3	41	1	1	0
14	1	0	3	43	1	1	0
15	1	0	3	31	1	1	0
16	1	1	0	18	1	1	0
17	1	0	0	27	1	1	0
18	1	0	1	25	1	1	0
19	1	0	3	35	1	1	0
20	1	0	3	47	1	1	0
21	1	0	3	29	1	1	0
22	1	1	3	34	1	1	0
23	1	1	0	18	1	1	0
24	1	0	0	24	1	1	0
25	1	1	1	22	1	4	0
26	1	1	3	29	1	2	0
27	1	1	1	23	1	4	0
28	1	0	2	26	1	1	0

Lampiran 3. Lanjutan

Individu	Distrik	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
29	1	0	0	25	1	3	0
30	1	0	3	45	1	1	0
31	1	0	0	20	1	1	0
32	1	0	1	20	1	4	0
33	1	0	3	32	1	2	0
34	1	0	3	30	1	3	0
35	1	1	2	32	1	4	0
36	1	1	1	25	1	4	0
37	1	0	0	25	1	4	0
38	1	0	3	44	1	4	0
39	1	1	0	23	1	4	0
40	1	0	3	42	1	1	0
41	1	1	1	26	1	1	0
42	1	0	2	31	1	4	0
43	1	1	1	24	1	3	0
44	1	1	1	28	1	4	0
45	1	0	3	40	1	4	0
46	1	0	2	23	1	3	0
47	1	1	2	27	1	4	0
48	1	0	2	22	1	1	0
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
2690	49	0	3	31	0	1	0
2691	49	0	0	16	0	1	0
2692	49	0	2	24	0	1	0
2693	49	0	3	44	0	1	0
2694	49	0	3	49	0	1	0
2695	49	0	0	18	0	1	0



Lampiran 3. Lanjutan

Individu	Distrik	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
2696	49	1	2	20	0	1	0
2697	49	0	2	27	0	1	0
2698	49	1	0	24	0	4	0
2699	49	0	3	43	0	1	0
2700	49	0	3	44	0	1	0
2701	49	0	2	25	0	1	0
2702	49	1	2	32	0	3	0
2703	49	0	1	24	0	1	0
2704	49	0	3	44	0	1	0
2705	49	0	3	34	0	1	0
2706	49	0	0	16	0	1	0
2707	49	0	3	37	0	1	0
2708	49	0	3	40	0	1	0
2709	49	1	3	41	0	1	1
2710	49	1	3	41	0	1	1
2711	49	0	3	39	0	1	1

Keterangan:

Y : Penggunaan Alat Kontrasepsi; 1 = Menggunakan, 0 = Tidak Menggunakan

X<sub>1</sub> : Status memiliki anak; 0 = Tidak Memiliki, 1 = Memiliki

X<sub>2</sub> : Umur (tahun)

X<sub>3</sub> : Daerah tinggal, 1 = Perkotaan, 0 = Pedesaan

X<sub>4</sub> : Tingkat pendidikan, 0 = Tidak pernah bersekolah, 1 = Pernah bersekolah

X<sub>5</sub> : Agama, 1 = Hindu, 0 = Muslim

#### Lampiran 4. Syntax Prosedur NLMIXED Pada Software SAS 9.3 (Data 1)

Syntax	Keterangan
proc import datafile="D:\data\Bakat Siswa TK.xlsx" out=bakat dbms=excel2007 replace;	Prosedur untuk melakukan import data (Data berupa file Microsoft Excel)
run;	Akhir syntax
proc nlmixed data=bakat;	Prosedur yang digunakan dalam analisis
parms beta0=-34.561 beta1=0.091 beta2=0.052 beta3=0.016 beta4=0.292 s2u=0.05;	Parameter yang akan diduga dan nilai awal untuk setiap parameter
eta= beta0+beta1*x1+beta2*x2+beta3*x3+beta4*x4+u;	Model regresi logistik multilevel
expeta=exp(eta);	Nilai exponential dari model
p = expeta/(1+expeta);	Peluang terjadi kejadian sukses
model y~ binomial(n,p);	Distribusi dari peubah respon
random u~ normal (0,s2u);	Distribusi dari <i>random</i> parameter
subject=sekolah;	Kelompok dalam analisis multilevel (level-2)
run;	Akhir syntax

Lampiran 5. Syntax Prosedur NLMIXED Pada Software SAS 9.3 (Data 2)

Syntax	Keterangan
proc import datafile="D:\data\Nilai Akreditasi SMK.xlsx" out=nilai dbms=excel2007 replace;	Prosedur untuk melakukan import data (Data berupa file Microsoft Excel)
run;	Akhir syntax
proc nlmixed data=nilai;	Prosedur yang digunakan dalam analisis
parms beta0=0 beta1=0 beta2=0 beta3=0 beta4=0 beta5=0 beta6=0 beta7=0 s2u=1;	Parameter yang akan diduga dan nilai awal untuk setiap parameter
eta= beta0+beta1*x1+beta2*x2+beta3*x3+beta4*x4+ beta5*x5+beta6*x6+beta7*x7+u;	Model regresi logistik multilevel
expeta=exp(eta);	Nilai exponential dari model
p = expeta/(1+expeta);	Peluang terjadi kejadian sukses
model y~ binomial(n,p);	Distribusi dari peubah respon
random u~ normal (0,s2u),	Distribusi dari <i>random</i> parameter
subject=kota;	Kelompok dalam analisis multilevel (level-2)
run;	Akhir syntax

Lampiran 6. Syntax Prosedur NLMIXED Pada Software SAS 9.3 (Data 3)

Syntax	Keterangan
proc import datafile="D:\data\Survey Bangladesh 1989.xlsx" out=SurveyBangladesh dbms=excel2000 replace; run;	Prosedur untuk melakukan import data (Data berupa file Microsoft Excel)
proc nlmixed data=SurveyBangladesh; parms beta0=-1.427 beta1=0.424 beta2=-0.02 beta3=0.619 beta4=0.363 beta5=0.439 s2u=0.05; eta= beta0+beta1*x1+beta2*x2+beta3*x3+ beta4*x4+beta5*x5+u; expeta=exp(eta); p = expeta/(1+expeta); model y~ binomial(n,p); random u~ normal (0,s2u); subject=distrik; run;	Akhir syntax Prosedur yang digunakan dalam analisis Parameter yang akan diduga dan nilai awal untuk setiap parameter Model regresi logistik multilevel Nilai exponential dari model Peluang terjadi kejadian sukses Distribusi dari peubah respon Distribusi dari <i>random</i> parameter Kelompok dalam analisis multilevel (level-2) Akhir syntax

## Lampiran 7. *Output* Metode PQL Dengan MLwiN 2.30 (Data 1)

MLwiN - J:\Ahmad Tantowi Jauhar\105090501111014\HASIL\DATA\TK\dataTK.wsz - [Equations]

Start More Stop IGLS Estimation control

$y_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$   
 $\text{logit}(\pi_{ij}) = \beta_{0j} + 0.169(0.058)x_{1ij} + 0.053(0.025)x_{2ij} + 0.125(0.044)x_{3ij} + 0.807(0.974)x_{4ij}$   
 $\beta_{0j} = -74.423(19.488) + u_{0j}$   
 $[u_{0j}] \sim N(0, \Omega_j) : \Omega_j = [9.886(8.572)]$   
 $\text{var}(y_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij}) / \text{denom}_{ij}$

(79 of 79 cases in use)

Name	+	-	Add Term	Estimates	Nonlinear	Clear	Iteration	Responses	Store	Help	Zoom	100
random				fixed			iteration 14					

Langkah-langkah metode PQL menggunakan software MLwiN 2.30:

- Input data yang akan dianalisis
- Klik **Model**, kemudian pilih **Equations**
- Klik pada tombol **Name**
- Klik pada y pada jendela peubah Y, isi sebagai berikut:
  - y : y
  - N levels : 2-ij
  - Level 2(j) : sekolah
  - Level 1(i) : siswa
- Klik **Done**
- Klik pada N di jendela **Equations**
- Pada jendela **Respons type**, pilih **Binomial** sebagai distribusi dan **logit** sebagai link function, Klik **Done**
- Klik pada X<sub>0</sub> pilih **cons** dari daftar yang disajikan dan klik **Done**
- Klik pada tombol **Add term**, dan pilih peubah yang akan dimasukkan dalam model, Klik **Done**
- Klik **cons** (koefisien  $\beta_0$ ) pada jendela **Equations**, Centang **j** (**sekolah**) pada jendela peubah X, Klik **Done**
- Klik tombol **Nonlinier** pada sisi bawah jendela **Equations**, Pilih **Ist Order** pada **Linieritation**, Pilih **PQL** pada **Estimation type**, Klik **Done**
- Klik **Estimation control**, Pilih **IGLS** pada **Method**, Klik **Done**
- Klik **Estimates**
- Klik **Start**



## Lampiran 8. *Output* Metode PQL Dengan MLwiN 2.30 (Data 2)

```

MLwiN - J:\Ahmad Tantowi Jauhari\105090501111014\HASIL\DATA99\Nilai SMK.wsz - [Equations]
File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help
Start More Stop IGLS Estimation control..

yij ~ Binomial(nij, πij)
logit(πij) = β0jx0 + 2.145(1.463)x1ij + 0.058(0.022)x2ij + -0.000(0.001)x3ij +
0.047(0.024)x4ij + -0.005(0.008)x5ij + 1.326(1.225)x6ij +
0.138(0.110)x7ij
β0j = -7.448(2.673) + u0j

[u0j] ~ N(0, Ωj) : Ωj = [0.466(0.538)]

var(yij | πij) = πij(1 - πij)/nij

(99 of 99 cases in use)

Name * - Add Term Estimates Nonlinear Iteration Notation Responses Store Help Zoom 100
random fixed iteration 10 Names Equations
    
```

Langkah-langkah metode PQL menggunakan software MLwiN 2.30:

- Input data yang akan dianalisis
- Klik **Model**, kemudian pilih **Equations**
- Klik pada tombol **Name**
- Klik pada y pada jendela peubah Y, isi sebagai berikut:
  - y
  - N levels : 2-ij
  - Level 2(j) : kota/kab
  - Level 1(i) : smk
- Klik **Done**
- Klik pada N di jendela **Equations**
- Pada jendela **Respons type**, pilih **Binomial** sebagai distribusi dan **logit** sebagai link function, Klik **Done**
- Klik pada X<sub>0</sub> pilih **cons** dari daftar yang disajikan dan klik **Done**
- Klik pada tombol **Add term**, dan pilih peubah yang akan dimasukkan dalam model, Klik **Done**
- Klik **cons** (koefisien β<sub>0</sub>) pada jendela **Equations**, Centang **j (kota/kab)** pada jendela peubah X, Klik **Done**
- Klik tombol **Nonlinier** pada sisi bawah jendela **Equations**, Pilih **Ist Order** pada **Linieritation**, Pilih **PQL** pada **Estimation type**, Klik **Done**
- Klik **Estimation control**, Pilih **IGLS** pada **Method**, Klik **Done**
- Klik **Estimates**
- Klik **Start**

## Lampiran 9. *Output* Metode PQL Dengan MLwiN 2.30 (Data 3)

MLwiN - J:\Ahmad Tantowi Jauhar\105090501111014\HASIL\BANGLADESH (2).wsz - [Equations]

File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help

Start More Stop IGLS Estimation control.

$$y_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_0 \text{cons} + 1.294(0.119)x1_{ij} + -0.011(0.006)x2_{ij} + 0.650(0.105)x3_{ij} + 0.705(0.091)x4_{ij} + 0.409(0.130)x5_{ij}$$

$$\beta_{0j} = -1.626(0.179) + u_{0j}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0.208(0.062)]$$

$$\text{var}(y_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}$$

(2711 of 2711 cases in use)

Name	+	-	Add Term	Estimates	Nonlinear	Clear	Notation	Responses	Store	Help	Zoom	100
random				fixed			Iteration 5					

Langkah-langkah metode PQL menggunakan software MLwiN 2.30:

- Input data yang akan dianalisis
- Klik **Model**, kemudian pilih **Equations**
- Klik pada tombol **Name**
- Klik pada y pada jendela peubah Y, isi sebagai berikut:
 

y	:	y
N levels	:	2-ij
Level 2(j)	:	wilayah
Level 1(i)	:	individu
- Klik **Done**
- Klik pada N di jendela **Equations**
- Pada jendela **Respons type**, pilih **Binomial** sebagai distribusi dan **logit** sebagai link function, Klik **Done**
- Klik pada X<sub>0</sub> pilih **cons** dari daftar yang disajikan dan klik **Done**
- Klik pada tombol **Add term**, dan pilih peubah yang akan dimasukkan dalam model, Klik **Done**
- Klik **cons** (koefisien β<sub>0</sub>) pada jendela **Equations**, Centang **j** (**district**) pada jendela peubah X, Klik **Done**
- Klik tombol **Nonlinier** pada sisi bawah jendela **Equations**, Pilih **Ist Order** pada **Linieritation**, Pilih **PQL** pada **Estimation type**, Klik **Done**
- Klik **Estimation control**, Pilih **IGLS** pada **Method**, Klik **Done**
- Klik **Estimates**
- Klik **Start**

Lampiran 10. *Output* Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 1)

The SAS System

The NLMIXED Procedure

<b>Specifications</b>	
Data Set	WORK.BAKAT
Dependent Variable	y
Distribution for Dependent Variable	Binomial
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	sekolah
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

<b>Dimensions</b>	
Observations Used	79
Observations Not Used	0
Total Observations	79
Subjects	3
Max Obs Per Subject	42
Parameters	6
Quadrature Points	1

<b>Parameters</b>						
beta0	beta1	beta2	beta3	beta4	s2u	NegLogLike
-34.561	0.091	0.052	0.016	0.292	0.05	24.6093817

Lampiran 10. Lanjutan

<b>Iteration History</b>					
<b>Iter</b>	<b>Calls</b>	<b>NegLogLike</b>	<b>Diff</b>	<b>MaxGrad</b>	<b>Slope</b>
1	6	24.5983493	0.011032	8.000897	-975.243
2	8	24.5115872	0.086762	9.129415	-6.67422
3	9	23.9762247	0.535362	24.30145	-8.82981
4	11	23.1439088	0.832316	115.4757	-41.8575
5	13	22.6006187	0.54329	83.01222	-0.48339
6	15	22.2834586	0.31716	45.57188	-0.21762
7	17	22.105947	0.177512	21.57015	-0.15909
8	19	22.0410383	0.064909	6.879782	-0.0568
9	21	22.0207071	0.020331	3.948979	-0.02746
10	23	22.0122827	0.008424	7.433726	-0.00646
11	24	22.0016932	0.01059	6.79741	-0.00699
12	26	21.8683162	0.133377	20.56866	-0.01018
13	27	21.7416337	0.126683	47.88555	-0.0961
14	28	21.5744489	0.167185	41.32269	-0.08481
15	30	19.55884	2.015609	34.89934	-0.22019
16	32	18.8315721	0.727268	8.974945	-0.73498
17	34	18.7744444	0.057128	1.175762	-0.08317
18	36	18.7714865	0.002958	0.434493	-0.0031
19	38	18.7697287	0.001758	2.180819	-0.00132
20	39	18.7669668	0.002762	3.05616	-0.001
21	41	18.7656333	0.001333	5.116563	-0.00148
22	43	18.7651872	0.000446	4.810931	-0.00041

### Lampiran 10. Lanjutan

23	44	18.764774	0.000413	2.525196	-0.00011
24	45	18.7642944	0.00048	1.129012	-0.00045
25	47	18.7641898	0.000105	0.084472	-0.00014
26	49	18.7641836	6.186E-6	0.009532	-8.79E-6
27	51	18.7641835	6.997E-8	0.000727	-1.27E-7

NOTE: GCONV cnvergence criterion satisfied.

Fit Statistics	
-2 Log Likelihood	37.5
AIC (smaller is better)	49.5
AICC (smaller is better)	50.7
BIC (smaller is better)	44.1

Parameter Estimates									
Parameter	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper	Gradient
beta0	-82.7082	27.4875	2	-3.01	0.0950	0.05	-200.98	35.5609	2.484E-6
beta1	0.1859	0.07123	2	2.61	0.1208	0.05	-0.1206	0.4924	0.000727
beta2	0.05972	0.02816	2	2.12	0.1680	0.05	-0.06144	0.1809	0.000531
beta3	0.1407	0.06212	2	2.27	0.1518	0.05	-0.1266	0.4080	0.000226
beta4	0.8901	1.0518	2	0.85	0.4865	0.05	-3.6354	5.4157	3.783E-6
s2u	12.4007	14.2891	2	0.87	0.4770	0.05	-49.0805	73.8819	1.335E-6



Lampiran 11. *Output* Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 2)

The SAS System

The NLMIXED Procedure

Specifications	
Data Set	WORK.NILAI
Dependent Variable	Y
Distribution for Dependent Variable	Binomial
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	kota
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	99
Observations Not Used	0
Total Observations	99
Subjects	21
Max Obs Per Subject	14
Parameters	9
Quadrature Points	1

Parameters									
beta0	beta1	beta2	beta3	beta4	beta5	beta6	beta7	s2u	NegLogLike
-7.472	1.928	0.05	0	0.038	-0.006	1.305	0.166	0.05	51.1299939

Lampiran 11. Lanjutan

Iteration History					
Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	6	51.1180471	0.011947	3.639619	-1232.39
2	9	51.1169196	0.001127	4.392105	-0.52513
3	11	51.1063388	0.010581	4.009683	-0.77292
4	13	51.1037094	0.002629	3.046304	-0.74314
5	15	51.0721651	0.031544	7.108669	-0.21913
6	17	50.8249813	0.247184	8.438592	-1.79799
7	19	50.7936378	0.031344	10.26103	-0.0356
8	21	50.7880709	0.005567	3.041886	-0.00919
9	23	50.7874719	0.000599	5.04652	-0.00029
10	25	50.7844288	0.003043	11.57062	-0.00057
11	26	50.7798815	0.004547	6.592002	-0.003
12	28	50.7796225	0.000259	0.738726	-0.00049
13	29	50.7794991	0.000123	7.410757	-0.00002
14	31	50.7789533	0.000546	7.891226	-0.00016
15	33	50.7788943	0.000059	0.817867	-0.00009
16	35	50.7788923	2.006E-6	0.112246	-3E-6
17	37	50.7788815	0.000011	0.943338	-3.52E-7

NOTE: GCONV convergence criterion satisfied.

## Lampiran 11. Lanjutan

Fit Statistics	
-2 Log Likelihood	101.6
AIC (smaller is better)	119.6
AICC (smaller is better)	121.6
BIC (smaller is better)	129.0

Parameter Estimates									
Parameter	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper	Gradient
beta0	-7.4460	2.6722	20	-2.79	0.0114	0.05	-13.0202	-1.8719	-0.00525
beta1	2.1260	1.5194	20	1.40	0.1771	0.05	-1.0434	5.2954	-0.00166
beta2	0.05804	0.02501	20	2.32	0.0310	0.05	0.005865	0.1102	-0.0753
beta3	-0.00005	0.001292	20	-0.04	0.9686	0.05	-0.00275	0.002643	-0.94334
beta4	0.04659	0.02652	20	1.76	0.0942	0.05	-0.00872	0.1019	-0.06708
beta5	-0.00467	0.008228	20	-0.57	0.5765	0.05	-0.02183	0.01249	0.280423
beta6	1.3356	1.1968	20	1.12	0.2777	0.05	-1.1609	3.8320	-0.00152
beta7	0.1382	0.1140	20	1.21	0.2396	0.05	-0.09962	0.3760	-0.03886
s2u	0.4679	0.7013	20	0.67	0.5122	0.05	-0.9949	1.9307	0.001452

Lampiran 12. *Output* Metode MLE Dengan SAS 9.3 PROC NLMIXED (Data 3)

The NLMIXED Procedure

Specifications	
Data Set	WORK.SURVEYBANGLADESH
Dependent Variable	y
Distribution for Dependent Variable	Binomial
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	wilayah
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions	
Observations Used	2711
Observations Not Used	0
Total Observations	2711
Subjects	49
Max Obs Per Subject	173
Parameters	7
Quadrature Points	1

Parameters							
beta0	beta1	beta2	beta3	beta4	beta5	s2u	NegLogLike
-1.427	0.424	-0.02	0.619	0.363	0.439	0.05	1841.28605

Lampiran 12. Lanjutan

<b>Iteration History</b>					
<b>Iter</b>	<b>Calls</b>	<b>NegLogLike</b>	<b>Diff</b>	<b>MaxGrad</b>	<b>Slope</b>
1	5	1690.62113	150.6649	299.5394	-1004082
2	8	1687.69466	2.926478	1578.451	-475.748
3	10	1656.28493	31.40972	688.7273	-68.5249
4	11	1652.28934	3.995593	289.8409	-9.78238
5	13	1650.9781	1.311239	209.3194	-1.1881
6	15	1650.18535	0.792748	11.55568	-0.65546
7	17	1649.97396	0.211396	113.8338	-0.136
8	19	1649.885	0.088956	24.55568	-0.08894
9	21	1649.87706	0.007939	7.712531	-0.01082
10	23	1649.87674	0.000318	0.726648	-0.00046
11	25	1649.87673	0.000012	0.043797	-0.00002
12	27	1649.87673	8.867E-8	0.011033	-1.51E-7

NOTE: GCONV convergence criterion satisfied.

<b>Fit Statistics</b>	
-2 Log Likelihood	3299.8
AIC (smaller is better)	3313.8
AICC (smaller is better)	3313.8
BIC (smaller is better)	3327.0



## Lampiran 12. Lanjutan

Parameter Estimates									
Parameter	Estimate	Standard Error	DF	t Value	Pr >  t	Alpha	Lower	Upper	Gradient
<b>beta0</b>	-1.6251	0.1798	48	-9.04	<.0001	0.05	-1.9866	-1.2637	-0.0003
<b>beta1</b>	1.2933	0.1202	48	10.76	<.0001	0.05	1.0515	1.5350	-0.00025
<b>beta2</b>	-0.01113	0.005589	48	-1.99	0.0522	0.05	-0.02237	0.000108	-0.01103
<b>beta3</b>	0.6492	0.1054	48	6.16	<.0001	0.05	0.4374	0.8611	-0.00009
<b>beta4</b>	0.7041	0.09163	48	7.68	<.0001	0.05	0.5199	0.8883	-0.00022
<b>beta5</b>	0.4086	0.1301	48	3.14	0.0029	0.05	0.1470	0.6702	-0.00048
<b>s2u</b>	0.2077	0.06372	48	3.26	0.0021	0.05	0.07959	0.3358	-0.00014



Lampiran 13. Model Logistik Multilevel Menggunakan Metode PQL  
(Data 1)

Model Level 1

$$\text{logit}(\pi_{ij}) = \beta_{0j} + 0,169X_{1ij} + 0,053X_{2ij} + 0,125X_{3ij} + 0,807X_{4ij}$$

Model Level 2

$$\beta_{0j} = -74,423 + u_{0j}$$

TK/RA	$u_{0j}$	$\beta_{0j}$
Al Hidayah I	6,342	-68,081
Indira Tama	3,45	-70,973
Al Hidayah II	-2,707	-77,13

Lampiran 14. Model Logistik Multilevel Menggunakan Metode PQL  
(Data 2)

Model Level 1

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 2,146X_{1ij} + 0,058X_{2ij} + 0,000X_{3ij} \\ & + 0,047X_{4ij} + (-0,005)X_{5ij} + 1,326X_{6ij} \\ & + 0,138X_{7ij} \end{aligned}$$

Model Level 2

$$\beta_{0j} = -7,448 + u_{0j}$$

Kab./Kota	$u_{0j}$	$\beta_{0j}$
Kab. Banyuwangi	0,242	-7,206
Kab. Blitar	0,506	-6,942
Kab. Bojonegoro	0,575	-6,873
Kab. Gresik	0,303	-7,145
Kab. Jember	0,689	-6,759
Kab. Jombang	0,394	-7,054
Kota Kediri	0,204	-7,244
Kab. Kediri	0,150	-7,298
Kab. Lamongan	0,433	-7,015
Kab. Lumajang	0,005	-7,443
Kab. Magetan	0,414	-7,034
Kab. Malang	0,672	-6,776
Kab. Mojokerto	0,926	-6,522
Kab. Nganjuk	0,353	-7,095
Kab. Ngawi	0,824	-6,624
Kab. Pacitan	0,199	-7,249
Kab. Pasuruan	0,846	-6,602
Kota Probolinggo	0,241	-7,207
Kab. Sidoarjo	0,836	-6,612
Kota Surabaya	0,833	-6,615
Kab. Trenggalek	0,456	-6,992

Lampiran 15. Model Logistik Multilevel Menggunakan Metode PQL  
(Data 3)

Model Level 1

$$\text{logit}(\pi_{ij}) = \beta_{0j} + 1,294X_{1ij} + (-0,011)X_{2ij} + 0,65X_{3ij} \\ + 0,705X_{4ij} + 0,409X_{5ij}$$

Model Level 2

$$\beta_{0j} = -1,626 + u_{0j}$$

Distrik	$u_{0j}$	$\beta_{0j}$
1	0,483	-1,143
2	0,644	-0,982
3	0,476	-1,150
4	0,818	-0,808
5	0,759	-0,867
6	0,807	-0,819
7	0,653	-0,973
8	0,455	-1,171
9	0,651	-0,975
10	0,652	-0,974
11	0,727	-0,899
12	0,608	-1,018
13	0,631	-0,995
14	0,888	-0,738
15	0,780	-0,846
16	0,696	-0,930
17	0,546	-1,080
18	0,739	-0,887
19	0,524	-1,102
20	0,490	-1,136
21	0,619	-1,007

Lampiran 15. Lanjutan

<b>Distrik</b>	<b><math>u_{0j}</math></b>	<b><math>\beta_{0j}</math></b>
22	0,885	-0,741
23	0,710	-0,916
24	0,920	-0,706
25	0,827	-0,799
26	0,911	-0,715
27	0,667	-0,959
28	0,698	-0,928
29	0,307	-1,319
30	0,629	-0,997
31	0,774	-0,852
32	0,505	-1,121
33	0,714	-0,912
34	0,508	-1,118
35	0,688	-0,938
36	0,823	-0,803
37	0,726	-0,900
38	0,498	-1,128
39	0,920	-0,706
40	0,733	-0,893
41	0,790	-0,836
42	0,569	-1,057
43	0,903	-0,723
44	0,253	-1,373
45	0,956	-0,670
46	0,762	-0,864
47	0,866	-0,760
48	0,559	-1,067
49	0,754	-0,872



Lampiran 16. Model Logistik Multilevel Menggunakan Metode MLE  
(Data 1)

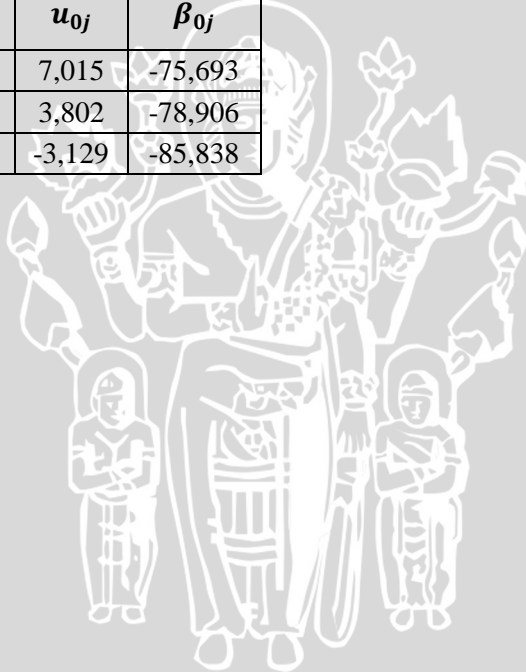
Model Level 1

$$\text{logit}(\pi_{ij}) = \beta_{0j} + 0,169X_{1ij} + 0,053X_{2ij} + 0,125X_{3ij} \\ + 0,807X_{4ij}$$

Model Level 2

$$\beta_{0j} = -82,7082 + u_{0j}$$

<b>TK/RA</b>	<b><math>u_{0j}</math></b>	<b><math>\beta_{0j}</math></b>
Al Hidayah I	7,015	-75,693
Indira Tama	3,802	-78,906
Al Hidayah II	-3,129	-85,838



Lampiran 17. Model Logistik Multilevel Menggunakan Metode MLE  
(Data 2)

Model Level 1

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{0j} + 2,126X_{1ij} + 0,05804X_{2ij} + (-0,00005)X_{3ij} \\ & + 0,04659X_{4ij} + (-0,00467)X_{5ij} + 1,3356X_{6ij} \\ & + 0,1382X_{7ij} \end{aligned}$$

Model Level 2

$$\beta_{0j} = -7,446 + u_{0j}$$

<b>Kota/Kab.</b>	$u_{0j}$	$\beta_{0j}$
Kab. Banyuwangi	0,244	-7,202
Kab. Blitar	0,504	-6,942
Kab. Bojonegoro	0,573	-6,873
Kab. Gresik	0,303	-7,143
Kab. Jember	0,688	-6,758
Kab. Jombang	0,390	-7,056
Kota Kediri	0,202	-7,244
Kab. Kediri	0,150	-7,296
Kab. Lamongan	0,431	-7,015
Kab. Lumajang	0,006	-7,440
Kab. Magetan	0,414	-7,032
Kab. Malang	0,671	-6,775
Kab. Mojokerto	0,925	-6,521
Kab. Nganjuk	0,357	-7,089
Kab. Ngawi	0,827	-6,619
Kab. Pacitan	0,200	-7,246
Kab. Pasuruan	0,846	-6,600
Kota Probolinggo	0,242	-7,204
Kab. Sidoarjo	0,833	-6,613
Kota Surabaya	0,831	-6,615
Kab. Trenggalek	0,458	-6,988

Lampiran 18. Model Logistik Multilevel Menggunakan Metode MLE  
(Data 3)

Model Level 1

$$\text{logit}(\pi_{ij}) = \beta_{0j} + 1,2933X_{1ij} + (-0,01113)X_{2ij} + 0,6492X_{3ij} \\ + 0,7041X_{4ij} + 0,4086X_{5ij}$$

Model Level 2

$$\beta_{0j} = -1,6251 + u_{0j}$$

Distrik	$u_{0j}$	$\beta_{0j}$
1	0,487	-1,138
2	0,646	-0,979
3	0,479	-1,146
4	0,821	-0,804
5	0,762	-0,863
6	0,809	-0,816
7	0,656	-0,969
8	0,457	-1,168
9	0,653	-0,972
10	0,655	-0,970
11	0,730	-0,895
12	0,612	-1,013
13	0,633	-0,992
14	0,890	-0,735
15	0,782	-0,843
16	0,699	-0,926
17	0,549	-1,076
18	0,740	-0,885
19	0,527	-1,099
20	0,492	-1,133
21	0,621	-1,004

Lampiran 18. Lanjutan

<b>Distrik</b>	<b><math>u_{0j}</math></b>	<b><math>\beta_{0j}</math></b>
22	0,888	-0,737
23	0,714	-0,912
24	0,923	-0,702
25	0,830	-0,795
26	0,914	-0,711
27	0,670	-0,955
28	0,701	-0,925
29	0,310	-1,315
30	0,633	-0,993
31	0,777	-0,848
32	0,508	-1,117
33	0,716	-0,909
34	0,512	-1,113
35	0,690	-0,935
36	0,826	-0,799
37	0,729	-0,897
38	0,501	-1,124
39	0,924	-0,702
40	0,736	-0,889
41	0,792	-0,833
42	0,572	-1,054
43	0,907	-0,719
44	0,256	-1,369
45	0,959	-0,666
46	0,765	-0,860
47	0,868	-0,757
48	0,562	-1,063
49	0,757	-0,868