

**GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA
DATA LONGITUDINAL DISKRIT DENGAN VARIABEL
RESPON ORDINAL**

**(Pada Data Penggunaan Mariyuana dan Data Penyakit
Schizophrenia)**

SKRIPSI

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

oleh:

RIAN RIZKI SUBADINA

0610950048-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2011**

UNIVERSITAS BRAWIJAYA



**GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA
DATA LONGITUDINAL DISKRIT DENGAN VARIABEL
RESPON ORDINAL**

**(Pada Data Penggunaan Mariyuana dan Data Penyakit
Schizophrenia)**

SKRIPSI

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

oleh:

RIAN RIZKI SUBADINA

0610950048-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2011**

LEMBAR PENGESAHAN SKRIPSI

**GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA
DATA LONGITUDINAL DISKRIT DENGAN VARIABEL
RESPON ORDINAL**

**(Pada Data Penggunaan Mariyuana dan Data Penyakit
Schizophrenia)**

oleh:

RIAN RIZKI SUBADINA

0610950048-95

**Setelah dipertahankan di depan Majelis Penguji
pada tanggal Februari 2011
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika**

Pembimbing I

Pembimbing II

**Prof. Dr. Ir. Loekito A. S, M.Agr
NIP. 19470327 197412 1 001**

**Adji Achmad Rinaldo F, S.Si, M.Sc
NIP. 19810908 200501 1 002**

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

**Dr. Abdul Rouf Alghofari, M.Sc
NIP. 19670907 199203 1 001
LEMBAR PERNYATAAN**

Saya yang bertanda tangan di bawah ini :

Nama : RIAN RIZKI SUBADINA
NIM : 0610950048-95
Program Studi : STATISTIKA
Penulis Skripsi Berjudul :

***GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA
DATA LONGITUDINAL DISKRIT DENGAN VARIABEL
RESPON ORDINAL***

**(Pada Data Penggunaan Mariyuana dan Data Penyakit
Schizophrenia)**

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 di kemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko yang akan saya terima.**

Demikian pernyataan ini dibuat dengan segala kesadaran.

**Malang, Februari 2011
Yang menyatakan,**

**RIAN RIZKI SUBADINA
NIM. 0610950048-95**

GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA DATA LONGITUDINAL DISKRIT DENGAN VARIABEL RESPON ORDINAL

**(Pada Data Penggunaan Mariyuana dan Data Penyakit
Schizophrenia)**

ABSTRAK

Dengan menggunakan regresi logistik ordinal (OLR) yang merupakan metode analisis untuk mengetahui bentuk hubungan antara satu variabel respon yang berskala ordinal (berdasarkan pada rangking yang diurutkan dari jenjang yang lebih tinggi sampai jenjang terendah atau sebaliknya) dengan satu atau lebih variabel penjelas, memudahkan dalam menaksir dan peramalan model. Dalam data longitudinal, pengamatan terhadap suatu subyek (*cross sectional*) diamati selama beberapa kali (unit waktu) yang menimbulkan korelasi antar pengamatan. Salah satu metode penyelesaian untuk data longitudinal adalah *Generalized Linear Mixed Model* (GLMMs) yang merupakan perluasan GLMs dengan menyertakan efek acak di samping efek tetap. Tujuan penambahan efek acak adalah untuk mengatasi korelasi antar pengamatan. Pendugaan parameter dalam GLMMs menggunakan *Maximum Likelihood* dan *Adaptive Gaussian Quadrature*. Pemilihan model terbaik dalam GLMMs menggunakan AIC dan *intraclass correlation* untuk mengetahui besarnya korelasi antar pengamatan. Data mariyuana diperoleh model terbaik dengan variabel *intercept*, *time* dan *gender* dengan efek acak intersep, dengan peluang laki-laki menggunakan mariyuana lebih besar dibandingkan perempuan. Data keparahan penyakit *Schizophrenia* diperoleh model terbaik dengan variabel *intercept*, *time* dan *treatment* dengan efek acak intersep, dengan penggunaan *drug* lebih efektif menurunkan keparahan penyakit *Schizophrenia*.

Kata Kunci: GLMMs, Data Longitudinal, Ordinal

**GENERALIZED LINEAR MIXED MODEL (GLMMs) IN
DISCRETE LONGITUDINAL DATA WITH RESPONSE
VARIABLE ORDINAL
(In Data Used Mariyuana And Schizophrenia Illness Data)**

ABSTRACT

By using ordinal logistic regression (OLR), which is a method of analysis to determine the form of the relationship between an ordinal scale response variables (based on the sorted ranking of the higher level until the lowest level or opposite) with one or more explanatory variables, facilitate inestimating and forecasting models. In longitudinal data, the observation on a subject (cross sectional) were observed for several times (unit time) that causes the correlation between observations. One method of settlement for longitudinal data is the Generalized Linear Mixed Model (GLMMs) which is the extension GLMs by including random effects in addition to fixed effects. The purpose of the addition of random effects is to address the correlation between observations. Estimation of parameters in GLMMs using Maximum Likelihood and Adaptive Gaussian Quadrature. Selection of the best model in GLMMs using AIC and the intraclass correlation to determine the magnitude of correlation between observations. Data obtained marijuana the best model with variable intercept, time and gender with random effects intercept, with opportunities to use marijuana male larger than female. Schizophrenia disease severity data obtained the best model with variable intercept, time and treatment with random effects intercept, with the use of the drug is more effective to reduce the severity of the disease Schizophrenia.

Keywords: GLMMs, Longitudinal Data, Ordinal

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT yang senantiasa memberikan rahmat, hidayah dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul **“GENERALIZED LINEAR MIXED MODEL (GLMMs) PADA DATA LONGITUDINAL DISKRIT DENGAN VARIABEL RESPON ORDINAL (Pada Data Penggunaan Mariyuana dan Data Penyakit Shhizophrenia)”** sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika. Pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Loekito A. S, M.Agr selaku Dosen Pembimbing I dan Bapak Adji Ahmad Rinaldo F, S.Si, M.Sc. selaku Dosen Pembimbing II atas bimbingan, masukan, serta bantuan yang telah diberikan.
2. Bapak Prof. Dr. Ir. Henny Pramodyo, MS, Ibu Ir. Heni Kusdarwati, MS, dan Ibu Eni Sumarminingsih, S.Si., MM., selaku dosen penguji atas saran dan nasihat yang telah diberikan.
3. Bapak Dr. Abdul Rouf Alghofari, M.Sc, selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya.
4. Keluargaku tercinta (Subadi, Muhayati, Pipit, Yola, Devi, Handy) yang telah memberikan kasih sayang, doa, dan dukungan yang tulus selama ini.
5. Teman-teman Statistika 2005, 2006, 2007 khususnya Lembuters dan D'pompong atas dukungan, bantuan dan kebersamaannya selama ini.
6. Staf TU Jurusan Matematika yang telah membantu keperluan administrasi.
7. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu dan memberikan dukungan selama penulisan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan selanjutnya. Semoga skripsi ini dapat bermanfaat bagi pembaca.

Malang, Februari 2011

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
 BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
 BAB II TINJAUAN PUSTAKA	
2.1 Regresi Logistik	5
2.2 Regresi Logistik Ordinal (<i>Ordinal Logistic Regression</i>)	8
2.2.1 Proportional Odds	10
2.2.2 Model Regresi Logistik Ordinal	12
2.2.3 Pendugaan Parameter Model Regresi Logistik Ordinal	14
2.2.4 Pengujian Signifikansi Parameter Model	17
2.2.5 Uji kesesuaian Model (<i>Goodness of fit</i>)	20
2.3 Data Longitudinal.....	21
2.4 Eksplorasi Data.....	23
2.5 <i>Generalized Linier Mixed Model (GLMMs)</i>	25
2.5.1 Model GLMMs.....	26
2.5.2 Pendugaan Parameter	29
2.5.3 Pengujian Hipotesis Parameter.....	31
2.5.4 Pemilihan Model Terbaik.....	31
2.5.4 <i>Intraclass Correlation</i>	32

BAB III METODE PENELITIAN

3.1 Sumber Data	33
3.2 Metode	34

BAB IV HASIL DAN PEMBAHASAN

4.1 Penelitian Penggunaan Mariyuana	39
4.1.1 Model Logistik Ordinal.....	39
4.1.1.1 Uji Kesesuaian Model	40
4.1.1.2 Model Akhir Regresi Logistik Ordinal	41
4.1.2 Pembentukan GLMMs pada data mariyuana	43
4.1.2.1 Eksplorasi data	43
4.1.2.2 Pemilihan Efek Acak.....	48
4.1.2.3 Pemilihan Efek Tetap.....	49
4.1.2.4 Pendugaan Parameter.....	49
4.1.2.5 Pengujian Korelasi dalam Kelas (Intraclass Corelation).....	50
4.1.2.6 Model Akhir <i>Generalized Linier Mixed Model</i>	50
4.1.2.7 Evolusi Variabel Respon	53
4.2 Penelitian Tingkat Keparahan Penyakit.....	54
4.2.1 Model Logistik Ordinal.....	54
4.2.1.1 Uji Kesesuaian Model	55
4.2.1.2 Model Akhir Regresi Logistik Ordinal	55
4.2.2 Pembentukan GLMMs pada data mariyuana	59
4.2.2.1 Eksplorasi data	59
4.2.2.2 Pemilihan Efek Acak.....	64
4.2.2.3 Pemilihan Efek Tetap.....	65
4.2.2.4 Pendugaan Parameter.....	66
4.2.2.5 Pengujian Korelasi dalam Kelas (Intraclass Corelation)	67
4.2.2.6 Model Akhir <i>Generalized Linier Mixed Models</i>	67
4.2.2.7 Evolusi Variabel Respon	71

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan.....	73
5.2 Saran	74

DAFTAR PUSTAKA.....	75
LAMPIRAN.....	79

UNIVERSITAS BRAWIJAYA



DAFTAR GAMBAR

	Halaman
Gambar 3.1. Diagram alir Proses pembentukan model Regresi Logistik Ordinal	36
Gambar 3.2 Diagram alir Proses Pembentukan GLMMs	37
Gambar 4.1. Peluang penggunaan mariyuana tiap kategori	42
Gambar 4.2. Grafik <i>mean profile</i> data penggunaan mariyuana..	44
Gambar 4.3. Grafik <i>variance profile</i> data penggunaan mariyuana	45
Gambar 4.4 Grafik eksplorasi jenis kelamin	46
Gambar 4.5 Grafik penggunaan mariyuana tiap kategori.....	51
Gambar 4.6 Pola penggunaan mariyuana	53
Gambar 4.7 Grafik peluang tingkat keparahan penyakit	57
Gambar 4.8 Grafik <i>mean profile</i> data keparahan penyakit.....	60
Gambar 4.9 Grafik <i>variance profile</i> data keparahan penyakit...	61
Gambar 4.10 Grafik keparahan penyakit tiap kategori.....	70
Gambar 4.11 Pola keparahan penyakit tiap kategori	72

DAFTAR TABEL

	Halaman
Tabel 2.1 Kerangka data longitudinal seimbang.....	22
Tabel 3.1 Data Penelitian	33
Tabel 4.1 Hasil pengujian signifikansi model logistik.....	39
Tabel 4.2 Hasil pengujian signifikansi parameter logistik.....	40
Tabel 4.3 Struktur korelasi data mariyuana	45
Tabel 4.4 Hasil pengujian serentak variabel bebas	47
Tabel 4.5 Hasil pengujian parsial variabel bebas.....	47
Tabel 4.6 Pemilihan efek acak data mariyuana	48
Tabel 4.7 Pemilihan efek tetap dengan efek acak intersep	49
Tabel 4.8 Pendugaan parameter efek tetap model	49
Tabel 4.9 Hasil pengujian signifikansi model logistik.....	54
Tabel 4.10 Hasil pengujian signifikansi parameter logistik.....	55
Tabel 4.11 Struktur korelasi data keparahan penyakit.....	62
Tabel 4.12 Hasil pengujian serentak variabel bebas	63
Tabel 4.13 Hasil pengujian parsial variabel bebas.....	63
Tabel 4.14 Pemilihan efek acak data keparahan penyakit	65
Tabel 4.15 Pemilihan efek tetap dengan efek acak intersep	65
Tabel 4.16 Pendugaan parameter efek tetap model	66

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Mariyuana	79
Lampiran 2. Data keparahan penyakit.....	80
Lampiran 3. <i>Syntax SAS 9.1.3 untuk Pembentukan Generalized Linier Mixed Models</i> pada data mariyuana.....	81
Lampiran 4. <i>Syntax SAS 9.1.3 untuk Pembentukan Generalized Linier Mixed Models</i> pada data keparahan penyakit	86
Lampiran 5. <i>Output SAS 9.1.3 untuk Pembentukan Generalized Linier Mixed Models</i> pada data mariyuana.....	92
Lampiran 6. <i>Output SAS 9.1.3 untuk Pembentukan Generalized Linier Mixed Models</i> pada data keparahan penyakit	96



BAB I

PENDAHULUAN

1.1 Latar Belakang

Regresi logistik merupakan metode analisis untuk mengetahui bentuk hubungan antara satu variabel respon dengan satu atau lebih variabel penjelas, di mana variabel respon berupa data kategori (kualitatif) dan variabel penjelas berupa data kategori (kualitatif) atau numerik. Jika variabel respon mempunyai lebih dari dua kategori disebut regresi logistik poltomus dan didasarkan pada distribusi multinomial (Hosmer dan Lemeshow, 2000). Ada kalanya permasalahan yang dihadapi adalah berupa variabel respon kategori yang lebih dari dua kategori dan bersifat ordinal. Sebagai contoh adalah pendapat responden (tidak setuju, setuju, sangat setuju) dan tingkat keparahan penyakit (tidak parah, cukup parah, sangat parah). Menurut Fahmeir dan Tutz (1994) apabila terdapat variabel respon dengan k kategori berskala ordinal yang memiliki tiga kategori atau lebih sedangkan variabel penjelas bersifat kontinyu atau kategori, maka digunakan model regresi logistik ordinal (*Ordinal Logistic Regression*). Kejadian berdistribusi multinomial dengan variabel respon ordinal seringkali ditemukan pada variabel respon untuk data longitudinal.

Data longitudinal disebut juga *repeated measurement data* (data pengukuran berulang) merupakan data yang didapatkan dari hasil pengukuran berulang pada unit *cross-sectional* (data pengamatan terhadap individu yang berbeda) dalam waktu berturut-turut (unit waktu). Penggunaan data *cross-sectional* dan data deret waktu, juga mampu menaksir dan mengukur pengaruh-pengaruh dengan lebih baik dibandingkan kedua data tersebut (Gujarati, 2003). Pengamatan terhadap unit *cross-sectional* yang sama cenderung untuk berkorelasi sehingga hal tersebut menyebabkan data longitudinal bersifat dependen, namun pengamatan antar unit *cross-sectional* yang berbeda dapat diasumsikan saling bebas. Suatu metode analisa statistika diperlukan untuk menangani adanya korelasi pada data longitudinal tersebut.

Gunawan (2009) telah menerapkan metode regresi untuk mengatasi korelasi antar unit waktu dengan menyertakan efek acak (*random effect*) dalam setiap unit *cross sectional* yang disebut *Generalized Linear Mixed Models* (GLMMs). Menurut Carriere dan

Bouyer (2006) studi longitudinal sangat menarik digunakan dalam ilmu epidemiologi dengan variabel respon ordinal sehingga dapat menjelaskan suatu proses evolusi suatu penyakit.

Seringkali penelitian di bidang pertumbuhan, epidemiologi dan kesehatan menggunakan data longitudinal, di mana benda hidup atau subyek manusia atau hewan sebagai individu yang diamati pada unit *cross-sectional* sedangkan waktu sebagai bahan pertimbangan untuk mendiagnosis evolusi dari subyek yang diteliti tiap unit waktu. Data penelitian yang pertama adalah data penggunaan mariyuana atau yang biasa di kenal dengan ganja pada kalangan remaja dengan variabel respon (Y) berupa skala ordinal (tidak pernah, tidak lebih dari satu bulan, lebih dari satu bulan), subyek pengamatan yang diteliti adalah kalangan remaja yang berjenis kelamin perempuan dan laki-laki selama lima tahun secara berturut turut. Variabel penjelas (X) adalah waktu (*time*) dan jenis kelamin (*gender*), dengan adanya perulangan berupa waktu maka dapat diketahui bagaimana perkembangan penggunaan mariyuana pada kalangan remaja dari tahun ke tahun yang semakin meningkat atau menurun dan bagaimana respon dari remaja yang berjenis kelamin perempuan dan laki-laki dalam ketertarikannya menggunakan mariyuana dengan mempertimbangkan adanya korelasi antar waktu.

Data penelitian yang kedua adalah data penyakit schizophrenia yaitu penyakit yang menyerang otak yang timbul akibat ketidakseimbangan salah satu sel kimia dalam otak (dopamine) dengan variabel respon (Y) berupa skala ordinal (normal, tidak parah, cukup parah, sangat parah) pada pasien penderita schizophrenia dengan pemberian *treatment* berupa *placebo* dan *drug* selama 0, 1, 3, 6 minggu. Variabel penjelas (X) adalah waktu (*time*) dan perlakuan (*treatment*), dengan adanya perulangan berupa waktu maka dapat diketahui bagaimana perkembangan penyakit schizophrenia pada penderita dari tahun ke tahun dengan efek pemberian *treatment* sehingga dapat diketahui peluang kesembuhan pasien yang semakin membaik ataupun memburuk.

Oleh karena itu dalam penelitian ini diangkat mengenai analisis data longitudinal dengan respon ordinal. GLMMs dapat digunakan untuk menganalisis data longitudinal dengan respon berupa data diskrit yang berdistribusi multinomial, selain itu GLMMs juga dapat mengatasi korelasi antar unit waktu dengan menyertakan efek acak dalam setiap unit *cross-sectional*.

1.2 Rumusan Masalah

Permasalahan dalam penelitian ini adalah bagaimana pembentukan model *Generalized Linier Mixed Models* pada variabel respon ordinal pada data penggunaan mariyuana dan data penyakit schizophrenia.

1.3 Batasan Masalah

Permasalahan yang dibahas di dalam penelitian ini dibatasi pada:

1. Model yang digunakan pada penelitian ini adalah model longitudinal seimbang.
2. Variabel pengamatan diukur hanya satu kali pada setiap waktu untuk setiap subyek atau individu.
3. Spesifikasi peragam tidak terstruktur (*unstructured covariance*) akan digunakan tanpa memperhatikan ada tidaknya pola korelasi pada data.
4. Fungsi link yang digunakan adalah fungsi link logit.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk membentuk model *Generalized Linier Mixed Models* pada data longitudinal dengan variabel respon ordinal pada data penggunaan mariyuana dan data penyakit schizophrenia.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah:

1. Sebagai penyelesaian masalah pada analisis data longitudinal ordinal dan memberikan masukan bagi penanganan masalah data longitudinal dengan variabel respon ordinal.
2. Untuk membentuk model pada variabel respon ordinal dengan korelasi antar waktu pada data longitudinal.

UNIVERSITAS BRAWIJAYA



BAB II TINJAUAN PUSTAKA

2.1 Regresi Logistik

Hosmer dan Lemeshow (2000) menyatakan bahwa regresi logistik merupakan salah satu metode regresi yang digunakan untuk mencari hubungan antara variabel respon kategori dengan salah satu atau lebih variabel penjelas (variabel eksplanatori) yang kontinu atau kategori. Banyaknya kategori variabel respon dalam model logistik dapat berbentuk dikotomis (biner) atau politomis.

Regresi logistik yang digunakan untuk data yang menyebar mengikuti keluarga eksponensial adalah *Generalized Linear Models* (GLMs). Agresti (2002) *Generalized Linear Models* merupakan metode untuk mengatasi kendala variabel respon yang tidak berdistribusi normal dan memodelkan nilai tengah sebagai fungsi dari variabel-variabel penjelas. Dalam GLMs antar pengamatan diasumsikan tidak terdapat hubungan atau korelasi. Variabel respon dalam GLMs diasumsikan memiliki distribusi yang termasuk dalam keluarga eksponensial.

Fahrmeir dan Tutz (1994) menyatakan bahwa model *Generalized Linear Models* pada regresi logistik terdiri atas tiga komponen, yaitu:

1. *Random Component* atau fungsi sebaran variabel acak $f(y)$ yang termasuk dalam keluarga eksponensial untuk suatu variabel acak yang tergantung pada parameter nilai tengah μ atau parameter lainnya.
2. *Systematic Component* atau fungsi dari variabel penjelas η yang mencakup p variabel penjelas $x_1, x_2, \dots, x_p$ dan β adalah parameter variabel penjelas dengan bentuk :

$$\eta(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2.1)$$

Kombinasi linier dari variabel penjelas pada persamaan (2.1) disebut *linier predictor*.

3. *Link Function* atau fungsi penghubung menggambarkan hubungan antara *random component* dan *systematic component* yang dinyatakan dalam bentuk $\eta = g(\mu_i)$. Misalkan terdapat $\mu_i = E(Y_i)$ di mana μ_i adalah nilai harapan dari sebaran acak Y_i , $i=1, \dots, n$ dan n adalah banyak pengamatan, μ_i dihubungkan dengan $\eta = g(\mu_i)$, dengan model *link function* antara μ_i dan p variabel bebas adalah

$$g(\mu_i) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2.2)$$

Variabel respon (Y) yang berskala ordinal menghasilkan k kategori, dengan mempertimbangkan $Y \in \{1, 2, \dots, k\}$ maka variabel respon multikategori (Y) menyebar menurut sebaran multinomial dengan *link function* atau fungsi hubung yang secara umum digunakan adalah fungsi logistik atau fungsi logit.

Menurut Agresti (2002), data biner adalah data yang hanya memiliki dua kategori, di mana seringkali berupa respon sukses (1) atau gagal (0). Pada suatu penelitian dengan variabel respon biner di mana jumlah percobaan sebanyak satu kali ($n=1$), maka peluang sukses Y_i mengikuti distribusi bernoulli. Sedangkan jika banyaknya pengamatan lebih dari 1 kali ($n \geq 2$) maka Y_i mengikuti distribusi binomial (n_i, π_i), yaitu :

$$Y_i \sim \binom{n_i}{y_i} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i}; \quad i=0, 1, \dots, n_i \quad (2.3)$$

Model logit dengan beberapa variabel penjelas adalah sebagai berikut:

$$\log[P(Y_i = 1 | x_i)] = \log\left[\frac{\pi_i}{1 - \pi_i}\right] = \alpha + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (2.4)$$

Model regresi logistik merupakan invers dari fungsi *logit*

$$P(Y_i = 1 | x_i) = \pi_i = \frac{\exp(\alpha + \beta_1 x_{i1} + \dots + \beta_p x_{ip})}{1 + \exp(\alpha + \beta_1 x_{i1} + \dots + \beta_p x_{ip})} \quad (2.5)$$

Di mana:

- α = intersep model logit
- β = vektor koefisien variabel penjelas
- $x_{i1}, \dots, x_{ip}$ = variabel penjelas
- p = banyaknya variabel penjelas ($p = 1, 2, \dots, p$)

Model logistik politomus adalah model logistik dengan variabel respon politomus (lebih dari dua kategori) yang merupakan perluasan dari model regresi logit biner atau model logistik dikotomus, model ini dikenal dengan istilah multinomial logit. Sebaran multinomial merupakan dasar yang digunakan oleh McCullagh dan Nelder (2003) untuk menjelaskan beberapa respon multikategori baik yang terukur dalam skala nominal, ordinal maupun interval.

Terkadang permasalahan yang dihadapi adalah variabel respon merupakan kategori yang bersifat ordinal bukan nominal. Sebagai contoh adalah pendapat responden (tidak setuju, setuju, sangat setuju), kinerja karyawan (tidak memuaskan, memuaskan, sangat memuaskan) dan tingkat keparahan (tidak parah, cukup parah, sangat parah). Pada permasalahan seperti ini bisa saja diterapkan metode regresi logistik multinomial, namun bagaimanapun juga metode ini tidak memperhitungkan sifat ordinalitas variabel respon.

Model regresi logistik politomus dengan variabel respon yang hanya mampu membedakan dikenal dengan nama regresi logistik politomus nominal, sedangkan pada variabel respon dengan kategori yang berfungsi penamaan dan juga dapat diurutkan dikenal dengan nama regresi politomus ordinal (Hosmer dan Lemeshow, 2000).

2.2 Regresi Logistik Ordinal (Ordinal Logistic Regression)

Fahrmeir dan Tutz (1994) berpendapat bahwa apabila terdapat variabel respon (Y) dengan k kategori berskala ordinal yang memiliki tiga kategori atau lebih sedangkan variabel bebas (x) dapat berupa kontinyu atau kategori, maka digunakan model regresi logistik skala ordinal (Ordinal Logistic Regression). Skala ordinal yaitu skala yang didasarkan pada rangking, diurutkan dari jenjang yang lebih tinggi sampai jenjang terendah atau sebaliknya. Sifat ordinal pada model logistik ordinal didasarkan pada peluang kumulatif.

Apabila terdapat sampel acak dari distribusi bersama (Y,X), di mana Y merupakan variabel respon berskala ordinal dan $\mathbf{X}' = (X_1, X_2, \dots, X_p)$ adalah vektor variabel penjelas. Variabel respon ordinal Y dengan k kategori dapat dipandang sebagai k titik pada variabel acak kontinyu tidak teramati (U). Sehingga bila $Y=k$ dan α_k berada diantara $\alpha_0 = -\infty$ dan $\alpha_r = +\infty$, maka dapat dinyatakan bahwa $\alpha_{k-1} < U \leq \alpha_k$ untuk $k = 1, 2, \dots, k-1$, sedangkan $-\infty = \alpha_0 < \alpha_1 < \dots < \alpha_k = +\infty$ dengan mengasumsikan bahwa variabel U ditentukan oleh variabel penjelas dengan bentuk linier :

$$U = -\mathbf{X}'\boldsymbol{\beta} + \varepsilon \quad (2.6)$$

di mana $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$ adalah vektor koefisien dan ε adalah variabel acak. Dari asumsi persamaan (2.6) diperoleh model untuk variabel respon Y sebagai berikut:

$$\begin{aligned} P(Y \leq j|\mathbf{X}) &= F(\alpha_k + \mathbf{X}'\boldsymbol{\beta}) \\ &= P(Y = 1|\mathbf{X}) + P(Y = 2|\mathbf{X}) + \dots + P(Y = k|\mathbf{X}) \end{aligned} \quad (2.7)$$

Model persamaan (2.7) disebut model peluang kumulatif atau *cumulative probabilities model*. Secara umum fungsi distribusi logistik sebagai berikut:

$$F(x) = \frac{1}{1 + \exp(-x)} \quad (2.8)$$

Maka model peluang kumulatif, yaitu:

$$P(Y \leq k|X) = F(\alpha_k + X'\beta)$$

$$= \frac{1}{1 + \exp(-(\alpha_k + X'\beta))}$$

$$= \frac{1}{1 + \exp(-(\alpha_k + X'\beta))} \cdot \frac{\exp(\alpha_k + X'\beta)}{\exp(\alpha_k + X'\beta)}$$

$$= \frac{\exp(\alpha_k + X'\beta)}{\exp(\alpha_k + X'\beta) + 1}$$

$$= \frac{\exp(\alpha_k + X'\beta)}{1 + \exp(\alpha_k + X'\beta)}$$

(2.9)

Di mana:

$k = 1, 1, \dots, k-1$

α_k = intersep

$\beta = (\beta_1, \beta_2, \dots, \beta_p)$ adalah vektor koefisien variabel penjelas

Dari persamaan (2.9) selanjutnya diperoleh persamaan sebagai berikut:

$$P(Y > k|X) = 1 - P(Y \leq k|X)$$

$$= 1 - \frac{\exp(\alpha_k + X'\beta)}{1 + \exp(\alpha_k + X'\beta)}$$

$$= \frac{1}{1 + \exp(\alpha_k + X'\beta)}$$

(2.10)

$$\begin{aligned}
\frac{P(Y \leq k | \mathbf{X})}{P(Y > k | \mathbf{X})} &= \frac{P(Y \leq k | \mathbf{X})}{1 - P(Y \leq k | \mathbf{X})} \\
&= \frac{\left(\frac{\exp(\alpha_k + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_k + \mathbf{X}'\boldsymbol{\beta})} \right)}{\left(\frac{1}{1 + \exp(\alpha_k + \mathbf{X}'\boldsymbol{\beta})} \right)} \\
&= \exp(\alpha_k + \mathbf{X}'\boldsymbol{\beta})
\end{aligned} \tag{2.11}$$

Untuk menghasilkan model yang linier maka:

$$\begin{aligned}
\log it[P(Y \leq k | \mathbf{X})] &= \ln \left(\frac{P(Y \leq k | \mathbf{X})}{1 - P(Y \leq k | \mathbf{X})} \right) \\
&= \ln(\exp(\alpha_k + \mathbf{X}'\boldsymbol{\beta})) \\
g(\mu) &= \alpha_k + \mathbf{X}'\boldsymbol{\beta}
\end{aligned} \tag{2.12}$$

Persamaan (2.12) disebut dengan *logit link function* atau fungsi hubung logit. Jika $\alpha_{k-1} < \alpha_k$ maka model ini adalah model kumulatif dengan slope yang sama, yaitu model regresi yang berdasarkan pada peluang kumulatif kategori respon.

2.2.1 Proportional Odds Model

Menurut Agresti (2002), sebuah model secara bersamaan menggunakan *cumulative logit* berdasarkan persamaan (2.12), yaitu

$$\log it[P(Y \leq k | x_i)] = \alpha_k + \beta' x_i, j = 1, \dots, j-1 \tag{2.13}$$

Tiap *cumulative logit* memiliki intersep sendiri, α_j akan meningkat di k karena $(Y \leq k | x)$ meningkat pada k untuk x tetap dan logit dapat meningkatkan fungsi dari peluang. Model persamaan (2.13) memiliki efek β yang sama untuk tiap logit karena β tidak dipengaruhi oleh i . Menurut persamaan (2.13) persamaan untuk *cumulative odds*

$$\log \left\{ \frac{P(Y \leq k | \mathbf{X})}{P(Y > k | \mathbf{X})} \right\} \quad k = 1, 2, \dots, q \quad (2.14)$$

Persamaan (2.14) ditentukan oleh bentuk linier dari variabel penjelas. Model kumulatif logistik juga disebut *proportional odds model*. Jika ada dua populasi yang dimisalkan dengan x_1 dan x_2 maka *ratio* dari *cumulative odds* untuk dua populasi berdasarkan persamaan (2.11) sebagai berikut:

$$\frac{P(Y \leq k | x_1) / P(Y > k | x_1)}{P(Y \leq k | x_2) / P(Y > k | x_2)} = \exp\{(x_1 - x_2)' \beta\} \quad (2.15)$$

Persamaan (2.15) disebut *cumulative odds ratio*. Karena *cumulative odds ratio* tidak bergantung pada kategori, ini menunjukkan bahwa pada *cumulative odds ratio* untuk dua populasi adalah sama untuk semua *cumulative odds* (Fahrmeir dan Tutz, 1994).

Model regresi logistik yang digunakan dalam penelitian ini adalah model *proportional odd*. Menurut Hosmer dan Lemeshow (2000), pada model ini perbandingan dilakukan antara peluang mendapatkan nilai respon sama dengan atau kurang dari nilai tertentu, dengan peluang mendapatkan nilai respon yang lebih besar. Misalkan nilai respon Y sama dengan j dengan syarat x dapat dinyatakan sebagai $P(y = j | x) = \pi_j(x)$, sehingga model logitnya adalah sebagai berikut:

$$\begin{aligned} \pi_k(x) &= \log \left[\frac{P(Y \leq k | x)}{P(Y > k | x)} \right] = \log \left[\frac{\pi_0(x) + \pi_1(x) + \dots + \pi_k(x)}{\pi_{k+1}(x) + \pi_{k+2}(x) + \dots + \pi_q(x)} \right] \\ &= \alpha_k + \mathbf{X}'\beta \end{aligned}$$

untuk $k = 1, 2, \dots, k - 1$

(2.16)

2.2.2 Model Regresi Logistik Ordinal

Agresti (2002) berpendapat bahwa model regresi logistik ordinal merupakan invers dari fungsi hubung logit persamaan (2.12) adalah sebagai berikut:

$$[P(Y_i \leq k|X)] = \frac{\exp(\alpha_k + X'\beta)}{1 + \exp(\alpha_k + X'\beta)} \quad (2.17)$$

Di mana :

$g(\mu) = \text{logit } P(y \leq j)$

$Y_{(n \times 1)}$ = vektor variabel respon

$X_{(n \times p)}$ = $(x_{11}, \dots, x_{ip})$ adalah matriks p variabel penjelas

$\beta_{(p \times 1)}$ = $(\beta_p, \dots, \beta_p)$ adalah vektor koefisien model logit

α_k = intersep model logit ke-k ($j=1,2,\dots,q$)

Model logit untuk variabel respon ordinal juga biasa di sebut dengan *Cumulative Logit Model* karena didasarkan pada peluang kumulatif dari variabel respon untuk setiap kategori. Regresi logistik ordinal biasa juga disebut dengan *Proportional Odds Model* karena koefisien variabel penjelas untuk j model logit tidak bergantung pada kategori variabel respon atau sama untuk j model logit (McCullagh dan Nelder, 2003). Berdasarkan persamaan (2.17), untuk variabel respon dengan kategori $k=1,2,\dots,q$ dengan peluang kumulatif $P(Y_i \leq k) = \pi_1(X) + \dots + \pi_k(X)$, maka diperoleh $\pi_k(X)$ yang merupakan peluang terjadinya j kategori variabel respon y dengan persamaan:

$$\begin{aligned} P(Y = k) &= \pi_k(X) \\ &= P(Y \leq k) - P(Y \leq k-1) \\ &= \frac{\exp(\alpha_k + X'\beta)}{1 + \exp(\alpha_k + X'\beta)} - \frac{\exp(\alpha_{k-1} + X'\beta)}{1 + \exp(\alpha_{k-1} + X'\beta)} \end{aligned} \quad (2.18)$$

Model logistik ordinal dengan variabel respon $y \sim \text{Multinomial}(y_1, y_2, \dots, y_k; \pi_1, \pi_2, \dots, \pi_k)$ yang terbentuk dengan tiga kategori, $j=1,2,3$ adalah sebagai berikut:

$$\begin{aligned}
\pi_1(X) &= P(Y \leq 1) = \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})} \\
\pi_2(X) &= P(Y \leq 2) - P(Y \leq 1) \\
&= \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})} - \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})} \\
\pi_3(X) &= 1 - P(Y \leq 2) \\
&= 1 - \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})} \\
&= \frac{1}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})}
\end{aligned} \tag{2.19}$$

Sedangkan untuk empat kategori ordinal adalah sebagai berikut:

$$\begin{aligned}
\pi_1(X) &= P(Y \leq 1) = \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})} \\
\pi_2(X) &= P(Y \leq 2) - P(Y \leq 1) \\
&= \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})} - \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta})} \\
\pi_3(X) &= P(Y \leq 3) - P(Y \leq 2) \\
&= \frac{\exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta})} - \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta})} \\
\pi_4(X) &= 1 - P(Y \leq 3) \\
&= 1 - \frac{\exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta})}{1 + \exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta})} \\
&= \frac{1}{1 + \exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta})}
\end{aligned} \tag{2.20}$$

2.2.3 Pendugaan Parameter Model Regresi Logistik Ordinal

Salah satu cara pendugaan parameter yang dapat dipergunakan pada regresi logistik ordinal adalah dengan *Maximum Likelihood Estimator* (MLE). Konsepnya adalah memaksimumkan fungsi likelihood dari sampel random untuk menduga parameter (Hosmer dan Lemeshow, 2000). Ketika lebih dari satu observasi Y muncul pada nilai X_i , adalah cukup menyatakan jumlah observasi untuk setiap munculnya nilai j pada Y dengan n_{ji} , dimana j adalah kategori tertentu dari variabel respon ($j=1,2,..., k$) (Kim, 2004). Jika variabel respon $y \sim \text{Multinomial}(y_1, y_2, ..., y_k; \pi_1, \pi_2, ..., \pi_k)$ maka fungsi *likelihood* untuk pengamatan y adalah:

$$\ell(\alpha, \beta) = \prod_{i=1}^n (\pi_1(X))^{y_{1i}} (\pi_2(X))^{y_{2i}} \dots (\pi_k(X))^{y_{ki}} \quad (2.21)$$

di mana:

y_k = vektor variabel untuk j kategori ($k=1, ..., q$)

n = total pengamatan ($y_{1i}, y_{2i}, ..., y_{ki}$)

$\pi_k(X)$ = vektor peluang untuk y_k

Untuk mencari nilai maksimum dari fungsi *likelihood* maka kedua ruas harus dilogaritmakan. Fungsi *log-likelihood* untuk persamaan :

$$\begin{aligned} L(\alpha, \beta) &= \log \left[\prod_{i=1}^n (\pi_1(X))^{y_{1i}} (\pi_2(X))^{y_{2i}} \dots (\pi_k(X))^{y_{ki}} \right] \\ &= \sum_{i=1}^n [y_{1i} \log(\pi_1(X)) + y_{2i} \log(\pi_2(X)) + \dots + y_{ki} \log(\pi_{ki}(X))] \\ &= \sum_{i=1}^n \left[y_{1i} \log \left(\frac{e^{\alpha_1 + X' \beta}}{1 + e^{\alpha_1 + X' \beta}} \right) + y_{2i} \log \left(\frac{e^{\alpha_2 + X' \beta}}{1 + e^{\alpha_2 + X' \beta}} \right) + \right. \\ &\quad \left. \dots + y_{ki} \log \left(\frac{e^{\alpha_{k-1} + X' \beta}}{1 + e^{\alpha_{k-1} + X' \beta}} \right) \right] \end{aligned} \quad (2.22)$$

Nilai $\alpha = [\alpha_1, \dots, \alpha_{k-1}]$ dan $\beta = [\beta_1, \beta_2, \dots, \beta_p]$ diperoleh dari turunan pertama masing-masing fungsi *log-likelihood*. Seperti halnya

dengan model dikotomis, pada model berskala ordinal ini juga memaksimumkan peluangnya dengan membuat turunan pertamanya terhadap parameter yang diduga sama dengan nol (Hyun, 2004). Hasil dari turunan pertama adalah sebagai berikut:

Hasil turunan pertama terhadap α_1 adalah sebagai berikut:

$$\frac{\partial L}{\partial \alpha_1} = \sum_{i=1}^n \left\{ y_{1i} \left(1 - \frac{e^{\alpha_1 + X' \beta}}{1 + e^{\alpha_1 + X' \beta}} \right) + y_{2i} \left(-\frac{e^{\alpha_1}}{e^{\alpha_2 - \alpha_1}} - \frac{e^{\alpha_1 + X' \beta}}{1 + e^{\alpha_1 + X' \beta}} \right) \right\}$$

Disamadengankan nol :

$$\frac{\partial L}{\partial \alpha_1} = \sum_{i=1}^n \left\{ y_{1i} \left(1 - \frac{e^{\alpha_1 + X' \beta}}{1 + e^{\alpha_1 + X' \beta}} \right) + y_{2i} \left(-\frac{e^{\alpha_1}}{e^{\alpha_2 - \alpha_1}} - \frac{e^{\alpha_1 + X' \beta}}{1 + e^{\alpha_1 + X' \beta}} \right) \right\} = 0 \quad (2.23)$$

Hasil turunan pertama terhadap α_{k-1} adalah sebagai berikut:

$$\begin{aligned} \frac{\partial L}{\partial \alpha_{k-1}} = \sum_{i=1}^n \left\{ (y_{k-1} - y_{k-2}) \left(\frac{e^{\alpha_{k-1}}}{e^{\alpha_{k-1}} - e^{\alpha_{k-2}}} - \frac{e^{\alpha_{k-1} + X' \beta}}{1 + e^{\alpha_{k-1} + X' \beta}} \right) \right. \\ \left. - (n - y_{k-1}) \left(\frac{e^{\alpha_{k-1} + \beta' X}}{1 + e^{\alpha_{k-1} + \beta' X}} \right) \right\} \end{aligned}$$

Disamadengankan nol :

$$\begin{aligned} \frac{\partial L}{\partial \alpha_{k-1}} = \sum_{i=1}^n \left\{ (y_{k-1} - y_{k-2}) \left(\frac{e^{\alpha_{k-1}}}{e^{\alpha_{k-1}} - e^{\alpha_{k-2}}} - \frac{e^{\alpha_{k-1} + X' \beta}}{1 + e^{\alpha_{k-1} + X' \beta}} \right) \right. \\ \left. - (n - y_{k-1}) \left(\frac{e^{\alpha_{k-1} + \beta' X}}{1 + e^{\alpha_{k-1} + \beta' X}} \right) \right\} = 0 \end{aligned}$$

(2.24)

Hasil turunan pertama terhadap β adalah sebagai berikut:

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^n \left\{ y_1 \left(\mathbf{X}' - \frac{\mathbf{X}' e^{\alpha_1 + \mathbf{X}' \beta}}{1 + e^{\alpha_1 + \mathbf{X}' \beta}} \right) + (y_2 - y_1) \left(\mathbf{X}' - \frac{\mathbf{X}' e^{\alpha_1 + \mathbf{X}' \beta}}{1 + e^{\alpha_1 + \mathbf{X}' \beta}} \right) + \dots + (1 - y_{k-1}) \left(- \frac{\mathbf{X}' e^{\alpha_{k-1} + \mathbf{X}' \beta}}{1 + e^{\alpha_{k-1} + \mathbf{X}' \beta}} \right) \right\}$$

Disamadengankan nol :

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^n \left\{ y_1 \left(\mathbf{X}' - \frac{\mathbf{X}' e^{\alpha_1 + \mathbf{X}' \beta}}{1 + e^{\alpha_1 + \mathbf{X}' \beta}} \right) + (y_2 - y_1) \left(\mathbf{X}' - \frac{\mathbf{X}' e^{\alpha_1 + \mathbf{X}' \beta}}{1 + e^{\alpha_1 + \mathbf{X}' \beta}} \right) + \dots + (1 - y_{k-1}) \left(- \frac{\mathbf{X}' e^{\alpha_{k-1} + \mathbf{X}' \beta}}{1 + e^{\alpha_{k-1} + \mathbf{X}' \beta}} \right) \right\} = 0 \quad (2.25)$$

Semua bentuk persamaan di atas adalah persamaan tidak linier terhadap parameternya. Untuk menduga dari persamaan tidak linier tidaklah semudah pada persamaan linier. Untuk mendapatkan nilai α dan β diperlukan suatu metode numerik untuk mempermudah perhitungan karena fungsi *loglikelihood* merupakan fungsi nonlinier (Hosmer dan Lemeshow, 2000).

Agresti (2002) mengatakan bahwa metode iterasi *Newton-Raphson* adalah metode yang menentukan nilai dari penduga parameter secara berulang-ulang sampai konvergen pada suatu nilai tertentu. Metode iterasi *Newton-Raphson* dapat menyelesaikan persamaan seperti fungsi *likelihood*, untuk menentukan nilai $\hat{\alpha}$ dan $\hat{\beta}$ dari $\theta_i = (\alpha_1, \alpha_2, \beta)$. Metode ini menggunakan rumus iterasi sebagai berikut:

$$\theta^{(t+1)} = \theta^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{g}^{(t)} \quad (2.26)$$

di mana

$\theta^{(t)}$: taksiran dari θ pada iterasi ke- t

$\mathbf{H}^{(t)}$: matriks turunan parsial kedua dari $L(\theta)$ di hitung pada

$\theta = \theta^{(t)}$

$\mathbf{g}^{(t)}$: vektor turunan pertama dari $L(\theta)$ di hitung pada $\theta = \theta^{(t)}$

sehingga $\mathbf{H}^{(t)}$ dan $\mathbf{g}^{(t)}$ berupa matriks, yaitu:

$$\mathbf{H}^{(t)} = \frac{\partial^2 \ell(\theta)}{\partial \theta_i} = \mathbf{X}' \mathbf{V}^{-1} \mathbf{X} \quad (2.27)$$

$$\mathbf{g}^{(t)} = \frac{\partial \ell(\theta)}{\partial \theta_i} = \mathbf{X}'(\mathbf{Y} - \boldsymbol{\pi}) \quad (2.28)$$

$(\mathbf{H}^{(t)})^{-1}$ adalah matriks varian kovarian. Matriks $\mathbf{X}$ berukuran $(n \times p+1)$ dengan p variabel bebas, $\mathbf{V}^{-1}$ adalah matriks dengan diagonal utama $\hat{\pi}_{ji}(1 - \hat{\pi}_{ji})$, dengan $j=1,2$ dan $i=1, \dots, n$. Sehingga didapatkan iterasi Newton Raphson sebagai berikut:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} (\mathbf{X}'(\mathbf{Y} - \boldsymbol{\pi})) \quad (2.29)$$

Melalui teknik iterasi *Newton Raphson* $\mathbf{g}^{(t)}$ dan $\mathbf{H}^{(t)}$ digunakan untuk menduga $\boldsymbol{\theta}^{(t)}$ pada iterasi ke t ($t=0,1,2,\dots$). Iterasi untuk pendugaan nilai θ_i terus dilakukan sampai mencapai kondisi konvergen untuk setiap i

$$|\hat{\theta}_i^{(t)} - \hat{\theta}_i^{(t-1)}| < \varepsilon$$

Di mana $j=1,2,\dots, n$, θ_i dianggap konvergen jika nilai ε sebesar 10^{-6} (Agresti, 2002).

2.2.4 Pengujian Signifikansi Parameter Model

Menurut Hosmer dan Lemeshow (2000), pengujian signifikansi digunakan untuk memeriksa apakah parameter yang diduga dengan variabel tertentu di dalam model lebih baik dibandingkan dengan model tanpa variabel tersebut. Pengujian signifikansi parameter ini digunakan untuk menentukan apakah variabel-variabel penjelas dalam model mempunyai hubungan yang nyata dengan variabel respon.

Pengujian terhadap parameter dilakukan secara parsial maupun serentak.

a. Pengujian signifikansi parameter secara parsial

Pengujian signifikansi parameter secara parsial digunakan untuk mengetahui pengaruh variabel penjelas terhadap variabel respon secara individu menggunakan uji *Wald* dengan hipotesis

$$H_0 : \beta_j = 0 \text{ lawan}$$

$$H_0 : \beta_j \neq 0$$

Dengan statistik uji:

$$W = \frac{(\hat{\beta}_j)^2}{[SE(\hat{\beta}_j)]^2} \sim \chi^2_{(v)} \quad (2.30)$$

v = derajat bebas dari variabel penjelas

$\hat{\beta}_j$ adalah penduga bagi β_j dan $SE(\hat{\beta}_j)$ merupakan standard error dan merupakan akar dari elemen diagonal $Cov(\beta_j) = [\mathbf{X}'\mathbf{V}\mathbf{X}]^{-1}$ seperti pada persamaan (2.27), di mana:

$$\mathbf{V}^{-1} = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1(k-1)} \\ V_{21} & V_{22} & \cdots & V_{2(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ V_{(k-1)} & V_{(k-1)2} & \cdots & V_{(k-1)(k-1)} \end{bmatrix}$$

$$= \begin{bmatrix} \hat{\pi}_1(1-\hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2(1-\hat{\pi}_2) & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & \cdots & \cdots & \hat{\pi}_q(1-\hat{\pi}_q) \end{bmatrix}$$

$$Var(\hat{\beta}_p) = \text{diagonal utama matriks } [\mathbf{X}'\mathbf{V}\mathbf{X}]^{-1}$$

$$SE(\hat{\beta}_q) = \sqrt{Var(\hat{\beta}_q)}$$

Jika $W \geq \chi^2_{(v)}$ atau $p\text{-value} < \alpha$ maka keputusannya tolak H_0 dengan kesimpulan bahwa variabel penjelas berpengaruh nyata terhadap variabel respon.

b. Pengujian signifikansi parameter secara serentak

Menurut Hosmer dan Lemeshow (2000), pengujian parameter secara serentak dilakukan untuk menguji keberartian dari parameter secara serentak dengan hipotesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p$$

$$H_1 : \text{Paling tidak ada satu } \beta_q \neq 0, q=1, 2, \dots, p$$

Pengujian signifikansi parameter secara serentak ini bertujuan untuk membandingkan nilai pengamatan respon dengan nilai dugaan responnya untuk kedua model. Menurut Agresti (2002), pendekatan uji statistik diperoleh melalui pendekatan *likelihood ratio test* antara model tanpa menyertakan β_i dan model yang menyertakan β_i , di mana $i=1,2,3,\dots,p$. Menurut Hosmer dan Lemeshow (2000), model yang menyertakan β_i disebut *full model* sedangkan untuk model yang tanpa menyertakan β_i disebut *reduce model*.

$$G^2 = -2 \ln \left(\frac{\text{likelihood without } \beta_i}{\text{likelihood with } \beta_i} \right), i = 1, 2, 3, \dots, p$$

$$G^2 = -2(L_0 - L_1) \quad (2.31)$$

Di mana :

L_0 = *loglikelihood* tanpa β_i , di mana $i=1,2,3,\dots,p$

L_1 = *loglikelihood* dengan β_i , di mana $i=1,2,3,\dots,p$

Nilai G^2 dibandingkan dengan statistik $\chi^2_{(v)}$ dengan derajat bebas v sesuai dengan jumlah parameter yang diduga. Hipotesis nol ditolak jika nilai probabilitas (*P-value*) lebih kecil dari α , dan diterima jika *P-value* lebih besar dari α .

2.2.5 Uji Kesesuaian Model (*Goodness of fit*)

Pengujian ini dilakukan untuk mengetahui apakah model yang dihasilkan sudah sesuai atau belum, dengan hipotesis:

H_0 = Model sesuai versus

H_1 = Model tidak sesuai

Menurut Fahmeir dan Tutz (1994), statistik uji yang digunakan untuk menguji kesesuaian model logistik ordinal adalah uji *Pearson* dan uji *Deviance*. Secara umum kedua uji statistik tersebut dirumuskan sebagai berikut:

1. Statistik *Pearson*

$$\chi^2 = \sum_{i=1}^g \chi_p^2(y_i, \hat{\pi}_i) \quad (2.32)$$

Di mana: $\chi_p^2(y_i, \hat{\pi}_i) = n_i \sum_{j=1}^k \frac{(y_{ij} - \hat{\pi}_{ij})^2}{\hat{\pi}_{ij}}$

2. Statistik *Deviance*

$$D = 2 \sum_{i=1}^g \chi_D^2(y_i, \pi_i) \quad (2.33)$$

Di mana: $\chi_D^2(y_i, \pi_i) = n_i \sum_{j=1}^k y_{ij} \log \left(\frac{y_{ij}}{\hat{\pi}_{ij}} \right)$

Dengan :

$\hat{\pi}_{ij}$ = peluang Y pada kategori ke-j dan x dengan kategori ke-i

y_{ij} = pengamatan Y pada kategori ke-j dan x dengan kategori ke-i

p = jumlah parameter pada model

g = jumlah kategori variabel x

k = jumlah kategori variabel respon

Statistik *Pearson* dan *Deviance* menyebar menurut distribusi khi-kuadrat dengan derajat bebas $g(k-1)-p$. keputusan tolak H_0 , jika $\chi_p^2, D > \chi_{(g(k-1)-p)}^2$ atau $P[\chi_v^2 > \chi_{pearson}^2]$ dan $P[\chi_v^2 > D]$ lebih

besar dari peluang yang diinginkan. Dengan ditolakny hipotesis nol maka kesimpulanny model yang diperoleh adalah tidak sesuai, dan sebaliknya jika peluang dihasilkan lebih besar dari peluang yang diinginkan maka H_0 diterima yang berarti model sudah sesuai.

2.3 Data Longitudinal

Data longitudinal merupakan data yang diamati dan di ukur berulang kali pada suatu interval waktu tertentu. Dibandingkan dengan data yang diperoleh dari studi *cross-sectional* yang umumnya dilakukan pada bidang sosial dan ekonomi, dimana pengukuran terhadap obyek hanya dilakukan sekali saja, maka data longitudinal memiliki keunggulan yaitu kemampuan dalam mengenali pengaruh waktu pengukuran terhadap respon. Data longitudinal diasumsikan obyek-obyek saling independen satu sama lainnya tetapi antara pengamatan di dalam obyek yang sama dependen, karena cenderung berkorelasi (Wu dan Zhang, 2006).

Data deret waktu diperoleh dari hasil mengukur suatu variabel selama waktu berturut-turut (runtun). Periode waktu dapat berupa tahun, bulan, minggu atau hari. Sedangkan data *cross-sectional* merupakan hasil pengamatan yang dilakukan pada individu berbeda pada waktu tertentu. Data yang didapatkan dari hasil pengukuran berulang pada beberapa individu (unit *cross-sectional*) dalam waktu berturut-turut (unit waktu) dikenal dengan data longitudinal yang disebut juga *repeated measurement data* (data pengukuran berulang). Menurut Weiss (2005), tujuan dari setiap analisis longitudinal adalah untuk mempelajari bagaimana subyek berubah dari waktu ke waktu, dan pengaruh dari karakteristik perubahan tersebut. Dengan adanya perulangan pengamatan yang dilakukan lebih dari satu sehingga variabel respon saling berkorelasi untuk setiap subyek (unit *cross-sectional*).

Danardono (2003), mengemukakan beberapa keuntungan dari data longitudinal, yaitu:

1. Dapat digunakan untuk mengetahui pola perubahan
2. Setiap individu dapat menjadi kontrol bagi dirinya sendiri
3. Dapat membedakan efek dari umur dengan efek dari cohort maupun efek dari periode

4. Memungkinkan untuk meneliti kausalitas

Data penelitian di mana banyaknya unit waktu untuk setiap unit *cross-sectional* sama disebut data longitudinal seimbang (Greene, 1997). Kerangka data longitudinal seimbang ditunjukkan pada Tabel 2.1 memperlihatkan unit pengukuran yang sama (tetap) dan jumlah pengamatan untuk unit *cross-sectional* (subyek) sama pada semua subyek yang diamati. Hun (2005) menyajikan data longitudinal seimbang (*balanced longitudinal data*) dalam bentuk tabel dengan kerangka sebagai berikut:

Tabel 2.1 Kerangka Data Longitudinal Seimbang

i	t	Y_{it}	X_{it}
1	1	Y_{11}	X_{11}
	2	Y_{12}	X_{12}
	$\vdots$	$\vdots$	$\vdots$
	T	Y_{1T}	X_{1T}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
S	1	Y_{S1}	X_{S1}
	2	Y_{S2}	X_{S2}
	$\vdots$	$\vdots$	$\vdots$
	T	Y_{ST}	X_{ST}

Di mana:

i = unit *cross-sectional* ke- i ($i = 1, 2, \dots, S$)

t = unit waktu ke- t ($t = 1, 2, \dots, T$)

Y_{it} = nilai variabel respon unit *cross-sectional* ke- i dan unit waktu ke- t

X_{it} = nilai variabel penjelas unit *cross-sectional* ke- i dan unit waktu ke- t

2.4 Eksplorasi Data

Sebuah eksplorasi analisis data digunakan untuk menentukan struktur rata-rata (efek tetap) dan ragam (efek acak) yang sangat penting untuk menentukan struktur awal model. Menurut Verbeke dan Molenberghs (2000) pembentukan model dan aspek formal kesimpulan awal dilakukan melalui eksplorasi data. Eksplorasi ini dapat dilakukan dengan menggunakan *mean profile*, *variance profile* dan *correlation structure*. Eksplorasi terhadap *concomitant variable* juga dapat dilakukan untuk mengetahui pengaruh *concomitant variable* terhadap subyek yang diteliti (Gunawan, 2009).

a. Mean Profile

Mean Profile menunjukkan bagaimana profil untuk sejumlah subpopulasi yang relevan (atau populasi secara keseluruhan) berubah seiring perubahan waktu. Hasil dari eksplorasi ini akan berguna untuk memilih struktur efek-tetap pada GLMM. Dapat berupa proporsi kejadian sukses yang pada setiap unit waktu untuk masing-masing treatment. *Mean profile* menggambarkan pengaruh efek tetap berupa waktu untuk subyek yang diteliti. Sehingga model sementara untuk efek tetap dapat diperoleh dari tingkat perubahan proporsi kejadian terhadap waktu. Jika pola dari *mean profile* menunjukkan ada fluktuasi atau perubahan di tiap waktu sehingga terlihat ada korelasi antar waktu maka efek tetap berupa waktu akan diikutsertakan dalam model tentatif. Jika pola dari *mean profile* konstan tiap unit waktu sehingga menyimpulkan tidak ada korelasi antar waktu maka efek tetap tidak diikutsertakan dalam model tentatif (Verbeke dan Molenbergh, 2000).

b. Variance Profile

Variance profile merupakan keragaman yang timbul pada tiap unit waktu. Perubahan terhadap variasi penting diamati untuk pembentukan model longitudinal yang tepat, hal ini diperoleh melalui pembentukan plot kuadrat residual General Linear Model (GLM) menggunakan metode Ordinary Least Square (OLS) terhadap waktu. Dengan mengamati perubahan keragaman respon pada setiap subyek seiring dengan perubahan waktu, akan diperoleh kesimpulan awal mengenai perlu atau tidaknya mengikutsertakan efek acak selain efek tetap ke dalam model data longitudinal. Seperti pada *mean profile* jika pola

variance profile menunjukkan adanya fluktuasi atau perubahan tiap unit waktu yang mengindikasikan adanya korelasi maka efek acak diikutsertakan dalam model tentatif. Jika pola *variance profile* konstan tiap unit waktu yang mengindikasikan tidak adanya korelasi maka efek acak tidak diikutsertakan dalam model tentatif (Verbeke dan Molenbergh, 2000).

c. *Correlation structure* (Struktur korelasi)

Struktur korelasi menggambarkan bagaimana pengukuran di dalam subyek berkorelasi. Korelasi menjelaskan hubungan antar waktu pengamatan yang terjadi. Pengukuran berulang terhadap subyek yang sama akan menyebabkan adanya ketidakbebasan dan ragam biasanya tidak konstan selama waktu pengamatan. Fungsi korelasi bergantung pada pasangan waktu dan di bawah asumsi stasioneritas. Struktur korelasi dapat terlihat dari plot residual dari waktu yang distandarisasi (standardized residuals) berguna dalam hal ini (Verbeke dan Molenbergh, 2000).

Fitzmaurice, et al (2004) struktur korelasi menunjukkan korelasi antar waktu pengamatan pada individu yang sama selain itu menunjukkan pola-pola tertentu antar waktu pengamatan. Oleh karena itu berdasarkan struktur korelasi tersebut dapat diprediksi bahwa struktur peragam yang sesuai dalam pembentukan model pada penelitian ini adalah *unstructured covariance* (peragam tidak terstruktur).

d. *Concomitant Variable*

Concomitant variable atau variabel pengiring merupakan variabel yang menjelaskan respon akan tetapi tidak dapat diubah. Eksplorasi terhadap *concomitant variable* dilakukan dengan membandingkan proporsi kejadian sukses untuk masing-masing nilai *concomitant variable* (Everitt, 2002).

2.5 Generalized Linier Mixed Models (GLMMs)

Generalized Linear Mixed Models (GLMMs) merupakan perluasan dari GLMs untuk data berkorelasi pada tiap pengamatan seperti pada data longitudinal dengan menambahkan efek acak pada persamaannya. Variabel respon dalam GLMMs diasumsikan saling bebas dengan penambahan efek acak pada masing-masing subyek. Secara umum model dalam GLMMs adalah :

$$g(\mu_{ij}) = \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'\mathbf{b}_i \quad (2.34)$$

di mana,

i : subyek pengamatan, $i = 1, 2, 3, \dots, S$

j : pengamatan masing-masing subyek, $j = 1, 2, 3, \dots, n$

$g(\cdot)$: fungsi link

μ_{ij} : nilai tengah untuk subyek ke- i , pengamatan ke- j

$\mathbf{X}$: vektor kovariat untuk efek tetap subyek ke- i , pengamatan ke- j

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

$\mathbf{Z}$: vektor kovariat efek acak untuk subyek ke- i , pengamatan ke- j

$\mathbf{b}_i$: efek acak untuk subyek ke- i ,

Efek acak b_0 diasumsikan menyebar normal dengan rata-rata 0 dan ragam σ_b^2 untuk efek acak intersep, atau $b_0 \sim N(0, \sigma_b^2)$.

Menurut Hedeker dan Gibbons (2006), σ_b^2 dapat digolongkan kecil jika korelasi yang terbentuk dalam subyek kecil. Menurut Agresti (2007), metode interpretasi dengan menyertakan efek acak intersep ke dalam model menggunakan selang kepercayaan untuk b_0 seperti pada persamaan (2.34)

$$\log(P(Y = 1|\mathbf{X})) = \mathbf{X}'\boldsymbol{\beta} + \left(\overline{b_{0i}} \pm Z_{\alpha/2} \sigma_b \right) \quad (2.35)$$

Di mana : $\overline{b_{0i}}$ = rata-rata efek acak intersep, yaitu bernilai 0

$Z_{\alpha/2}$ = titik kritis dengan tingkat kesalahan sebesar α

σ_b = standar deviasi efek acak

Proportional odds mengasumsikan parameter efek tetap (β) tidak tergantung atau bebas terhadap respon k kategori. Diketahui bahwa Y_{ij} adalah respon ordinal dengan k kategori, sehingga model yang digunakan untuk respon ordinal adalah model *cumulative logit* dengan

$$\log it(P(Y_{ij} \geq 1)) \geq \dots \log it(P(Y_{ij} \geq c)) \geq \dots \log it(P(Y_{ij} \geq k-1)).$$

Untuk model GLMMs dengan respon ordinal akan dijelaskan pada uraian 2.5.1.

2.5.1 Model GLMMs

Molenberghs dan Verbeke (2005) berpendapat bahwa GLMMs dapat digunakan untuk menganalisis data longitudinal diskrit, termasuk untuk respon ordinal. Hubungan antara variabel respon dengan parameter efek tetap dan efek acak dengan fungsi link menggunakan model logit, sehingga didapatkan hubungan antara variabel respon yang berdistribusi multinomial. Dengan *cumulative logits* biasa juga disebut *proportional odds model*, menunjukkan hasil yang mungkin untuk y_{ij} , observasi t pada kelompok i, dengan $1, 2, \dots, S$. GLMMs untuk *cumulative logits* mempunyai bentuk:

$$\log it[P(Y_{ij} \leq k | \mathbf{X}, \mathbf{Z})] = \alpha_k + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'\mathbf{b}_i \quad (2.36)$$

Di mana:

α_k = vektor intersep regresi logistik ($k=1, \dots, k-1$)

$\mathbf{X}$ = vektor kovariat untuk efek tetap subyek ke-i, pengamatan ke-j

$\boldsymbol{\beta}$ = parameter efek tetap

$\mathbf{Z}$ = vektor kovariat efek acak untuk subyek ke-i, pengamatan ke-j

$\mathbf{b}_i$ = vektor efek acak untuk subyek ke-i

dimana $\mathbf{X}$ dan $\mathbf{Z}$ adalah matriks dari efek tetap dan efek acak, serta $\boldsymbol{\beta}$ dan $\mathbf{b}_i$ adalah vektor dari parameter efek tetap dan efek acak diasumsikan menyebar normal multivariate dengan rata-rata 0 dan kovarians Σ atau $b_i \sim N(0, \Sigma)$.

Seperti pada *logit link function* atau fungsi hubung logit pada persamaan (2.12), model logistik digunakan untuk mengetahui hubungan antara variabel penjelas dengan variabel respon maka berdasarkan bentuk persamaan 2.36 diperoleh model GLMMs untuk tiga kategori :

$$\begin{aligned}
 \pi_{ij1} &= P(Y_{ij} \leq 1 | \mathbf{X}_i, \mathbf{Z}_i) \\
 &= \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} \\
 \pi_{ij2} &= P(Y_{ij} \leq 2 | \mathbf{X}, \mathbf{Z}) - P(Y_{ij} \leq 1 | \mathbf{X}, \mathbf{Z}) \\
 &= \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} - \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} \\
 \pi_{ij3} &= 1 - P(Y_{ij} \leq 2 | \mathbf{X}, \mathbf{Z}) \\
 &= \frac{1}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}
 \end{aligned} \tag{2.37}$$

Sedangkan untuk empat kategori adalah sebagai berikut:

$$\begin{aligned}
 \pi_{ij1} &= P(Y_{ij} \leq 1 | \mathbf{X}_i, \mathbf{Z}_i) \\
 &= \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} \\
 \pi_{ij2} &= P(Y_{ij} \leq 2 | \mathbf{X}, \mathbf{Z}) - P(Y_{ij} \leq 1 | \mathbf{X}, \mathbf{Z}) \\
 &= \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} - \frac{\exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_1 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} \\
 \pi_{ij3} &= P(Y_{ij} \leq 3 | \mathbf{X}, \mathbf{Z}) - P(Y_{ij} \leq 2 | \mathbf{X}, \mathbf{Z}) \\
 &= \frac{\exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_3 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} - \frac{\exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)} \\
 \pi_{ij4} &= 1 - P(Y_{ij} \leq 3 | \mathbf{X}, \mathbf{Z}) \\
 &= \frac{1}{1 + \exp(\alpha_2 + \mathbf{X}'\boldsymbol{\beta} + \mathbf{Z}'b_i)}
 \end{aligned} \tag{2.38}$$

Terdapat satu perbedaan bahwa telah dipilih intersep tetap dengan kategori yang tidak saling bebas, ini menunjukkan bahwa tidak ada parameter intersep pada vektor β . Ini bukanlah masalah untuk mengintegrasikan intersep tetap ke dalam matriks desain efek tetap $\mathbf{X}$, sehingga tidak ada perbedaan antara model efek acak untuk data ordinal dan model umum. Cara paling mudah untuk menjelaskan model ordinal adalah dengan menyusun semua indikator kumulatif Z_{ijk} ke dalam vektor $\mathbf{Z}$ sehingga terbentuk:

$$\log it[E(\mathbf{Z}|\mathbf{X}, \mathbf{Z}, \beta, b_i)] = \mathbf{X}\beta + \mathbf{Z}b_i \quad (2.39)$$

Model *proporsional odds* mengikuti persamaan (2.12), dengan memastikan semua efek kovariat yang umum untuk pengukuran j , bebas untuk titik potong k (Verbeke dan Molenberghs, 2005).

Untuk parameter efek tetap β , pendugaan diperoleh dari kondisi efek acak. Efek acak merupakan komponen variasi yang tidak dijelaskan dalam variabel prediktor. Sedangkan efek tetap merupakan pengaruh yang ditimbulkan oleh variabel respon maupun *concomitant variable*. Menurut Everitt (2002) *concomitant variable* atau variabel pengiring dapat digolongkan menjadi variabel penjelas akan tetapi bukan variabel utama dalam pengamatan dan tetap diamati karena tetap mempengaruhi variabel respon sehingga perlu dipertimbangkan untuk disertakan ke dalam model.

2.5.2 Pendugaan Parameter

Menurut Molenberghs dan Verbeke (2005), efek tetap dan efek acak dapat diduga dengan pemaksimalan *marginal likelihood*, yaitu didapatkan dengan mengintegralkan efek acak. Menurut Jiang (2007), persamaan *likelihood* untuk masing-masing subyek adalah :

$$\begin{aligned} L(\beta, D) &= \prod_{i=1}^N f_i(y_i | \beta, D) \\ &= \prod_{i=1}^N \int \prod_{j=1}^{n_i} f_{ij}(y_{ij} | b_i, \beta) f(b_i | D) db_i \end{aligned} \quad (2.40)$$

Di mana

$$\begin{aligned} f_{ij}(y_i | b_i, \beta) &= \pi_1^{y_{1i}} \pi_2^{y_{2i}} \dots \pi_k^{y_{ki}} \\ &= \left(\frac{e^{\alpha_1 + X' \beta + Z' b_i}}{1 + e^{\alpha_1 + X' \beta + Z' b_i}} \right)^{y_{1i}} \left(\frac{e^{\alpha_2 + X' \beta + Z' b_i} - e^{\alpha_1 + X' \beta + Z' b_i}}{\left(1 + e^{\alpha_2 + X' \beta + Z' b_i} \right) \left(1 + e^{\alpha_1 + X' \beta + Z' b_i} \right)} \right)^{y_{2i}} \dots \\ &\quad \left(\frac{e^{\alpha_k + X' \beta}}{1 + e^{\alpha_k + X' \beta}} \right)^{y_{ki}} \end{aligned} \quad (2.41)$$

$f(b_i | D)$ merupakan fungsi kepekatan peluang normal multivariat di mana $b_i \sim N(0, D)$.

Integral tersebut tidak didapatkan hasil secara langsung, oleh karena itu diperlukan sebuah metode pendekatan untuk menyelesaikannya. Salah satu metode yang dapat digunakan untuk menyelesaikan integral tersebut adalah dengan menggunakan pendekatan *Gaussian Quadrature* (Molenberghs dan Verbeke, 2005).

Dalam *Gaussian Quadrature*, secara umum bentuk integral terhadap suatu fungsi pada suatu batas tertentu yang dapat di tulis

$$\int f(z) \phi(z) dz \approx \sum_{q=1}^Q w_q f(z_q) \quad (2.42)$$

di mana w_q merupakan bobot atau fungsi kernel, sedangkan Q merupakan titik *quadrature*, semakin besar jumlah titik *quadrature* maka akan semakin kecil kesalahan yang dihasilkan. Fungsi $f(z)$

mengambil dari bentuk kepekatan keluarga eksponensial dengan fungsi eksponensial. Jika dalam *Gaussian Quadrature* titik-titik *quadrature*-nya tidak memberikan kontribusi terhadap w_q , oleh karena itu perlu dilakukan pemusatan titik-titik tersebut agar semua titik memberikan kontribusi terhadap w_q . Metode pemusatan titik-titik ini disebut *Adaptive Gaussian Quadrature* (Gunawan, 2009).

Dalam *Adaptive Gaussian Quadrature* titik-titik *quadrature* dipusatkan dan persamaan $f(z)\phi(z)$ diasumsikan menyebar normal. Rata-rata dari distribusi normal menjadi $\hat{z}$ dari $\ln[f(z)\phi(z)]$ dan varian dapat dihitung dengan rumus

$$\left[-\frac{\partial^2}{\partial z^2} \ln[f(z)\phi(z)] \right]_{z=\hat{z}}$$

Titik *quadrature* dan bobot diperbaharui dengan :

$$z_q^+ = \hat{z} + \left[-\frac{\partial^2}{\partial x^2} \ln(f(z)\phi(z)) \right]_{z=\hat{z}}^{-1/2} z_q \quad (2.43)$$

$$w_q^+ = \left[-\frac{\partial^2}{\partial x^2} \ln(f(z)\phi(z)) \right]_{z=\hat{z}}^{-1/2} \frac{\phi(z_q^+)}{\phi(z_q)} w_q \quad (2.44)$$

Menurut Molenberghs dan Verbeke (2005), bentuk integral pada *adaptive gaussian quadrature* untuk menduga efek acak adalah

$$\int_{-\infty}^{\infty} f(z)\phi(z)dx \approx \sum_{q=1}^Q w_q^+ f(z_q^+) \quad (2.45)$$

2.5.3 Pengujian Hipotesis Parameter

Pengujian pendugaan terhadap parameter efek tetap digunakan untuk memilih efek tetap yang sesuai pada model. Pendekatan yang dapat dilakukan adalah dengan menggunakan uji t. pengujian hipotesis setiap parameter β_j pada vector β , $j=1,2,...,p$ menggunakan uji t. secara umum untuk sembarang matriks L yang diketahui, pengujian hipotesis:

$$H_0: L\beta = 0 \text{ versus}$$

$$H_1: L\beta \neq 0$$

Pendekatan uji t (t_{hitung}) dan interval kepercayaan diperoleh melalui pendekatan distribusi

$$t_{hitung} = (\hat{\beta} - \beta_j) / s.\hat{e}(\hat{\beta}) \quad (2.46)$$

$$s.\hat{e}(\hat{\beta}) = \sqrt{\text{Var}(\hat{\beta})}$$

Dengan distribusi t yang bersesuaian, di mana derajat bebas uji t diduga dari data pengamatan. Hipotesis nol akan diterima apabila diperoleh nilai probabilitas (p-value) lebih besar dibandingkan peluang berbuat salah sebesar 0,05 dan sebaliknya apabila p-value lebih kecil dari 0,05 maka hipotesis nol ditolak.

2.5.4 Pemilihan Model Terbaik

Menurut Weiss (2005) pemilihan model terbaik menggunakan AIC (*Akaike Information Criteria*) dengan rumus sebagai berikut:

$$AIC = n \ln(\sigma_a^2) + 2m \quad (2.47)$$

Di mana:

σ_a^2 = penduga ragam sisaan

m = banyaknya parameter yang diduga dalam model

n = banyaknya pengamatan

Menurut Agresti (2002) model terbaik adalah model yang memiliki AIC terkecil, karena memiliki ragam sisaan yang lebih kecil.

2.5.5 Intraclass Correlation

Intraclass correlation atau korelasi dalam kelas menunjukkan proporsi keragaman yang tidak dapat diterangkan dalam subyek dengan kata lain mencerminkan besarnya keragaman antar subyek. Menurut Hedeker dan Gibbons (2006) dalam bentuk regresi logistik ragam sisa diasumsikan mengikuti *standard logistic distribution*

dengan ragam residual $\frac{\pi^2}{3}$ atau 3,29.

Nilai sisa selalu tetap karena y tidak diamati. Anggap bahwa σ_b^2 merupakan ragam efek acak, maka persamaan korelasi dalam kelas dapat digambarkan dengan :

$$\rho = \frac{\sigma_b^2}{\sigma_{res}^2 + \sigma_b^2}$$

Karena $\sigma_{res}^2 = \frac{\pi^2}{3}$ sehingga korelasi dalam kelas dapat juga

ditulis :

$$\rho = \frac{\sigma_b^2}{\frac{\pi^2}{3} + \sigma_b^2} \quad (2.48)$$

Menurut Foy (2001), nilai korelasi dalam kelas yang timbul karena penelitian berulang dinyatakan kecil jika nilai mutlaknya kurang dari atau sama dengan 0,4.

BAB III METODE PENELITIAN

3.1 Sumber Data

Data longitudinal yang akan dimodelkan dalam penelitian ini merupakan data longitudinal seimbang atau memiliki jumlah pengamatan yang sama untuk tiap subyek dengan respon ordinal. Data penelitian tersaji pada Tabel 3.1.

Tabel 3.1 Data Penelitian

No	Judul Penelitian	Variabel Prediktor	Variabel Respon
1.	Penggunaan mariyuana pada National Youth Survey (http://spitswww.uvt.nl/~vermunt/hauspie2004.pdf)	- Time : waktu pengamatan pada tahun 1976-1980 - Gender : ➤ Perempuan=0 ➤ Laki-laki=1	Penggunaan mariyuana tiap tahunnya ➤ Tidak pernah (1) ➤ Tidak lebih dari satu bulan (2) ➤ Lebih dari satu bulan (3)
2.	Penyakit <i>Schizophrenia</i> pada National Institute of Mental Health <i>Schizophrenia Collaborative Study</i> (http://tigger.uic.edu/~hedeker/SCHIZX1.DAT.txt)	- Time : waktu pengamatan ➤ Minggu ke-0 =0 ➤ minggu ke-1=1 ➤ Minggu ke-3 =3 ➤ Minggu ke-6 =6 - Group : ➤ Placebo (0) ➤ Drug (1)	Informasi mengenai tingkat keparahan penyakit <i>Schizophrenia</i> tiap minggunya ➤ normal atau tidak sakit (1), ➤ ringan atau tidak parah (2), ➤ cukup parah (3), ➤ sangat parah (4).

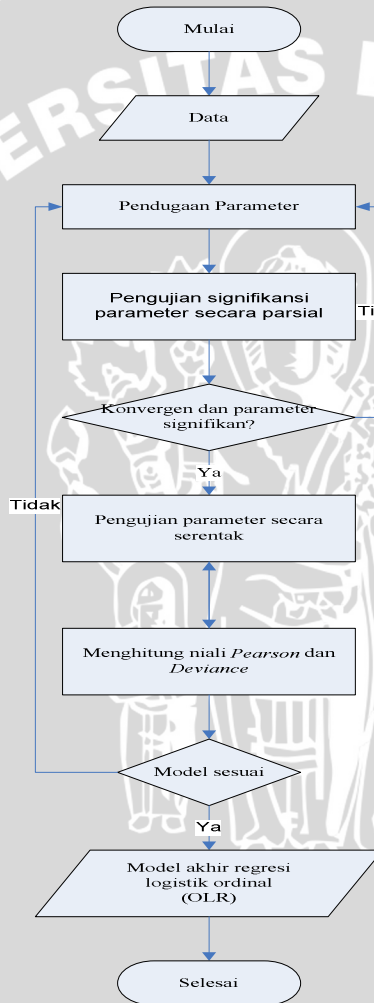
3.2 Metode

Tahapan pembentukan *Generalized Linear Mixed Model* (GLMMs) pada data longitudinal dengan respon ordinal adalah sebagai berikut:

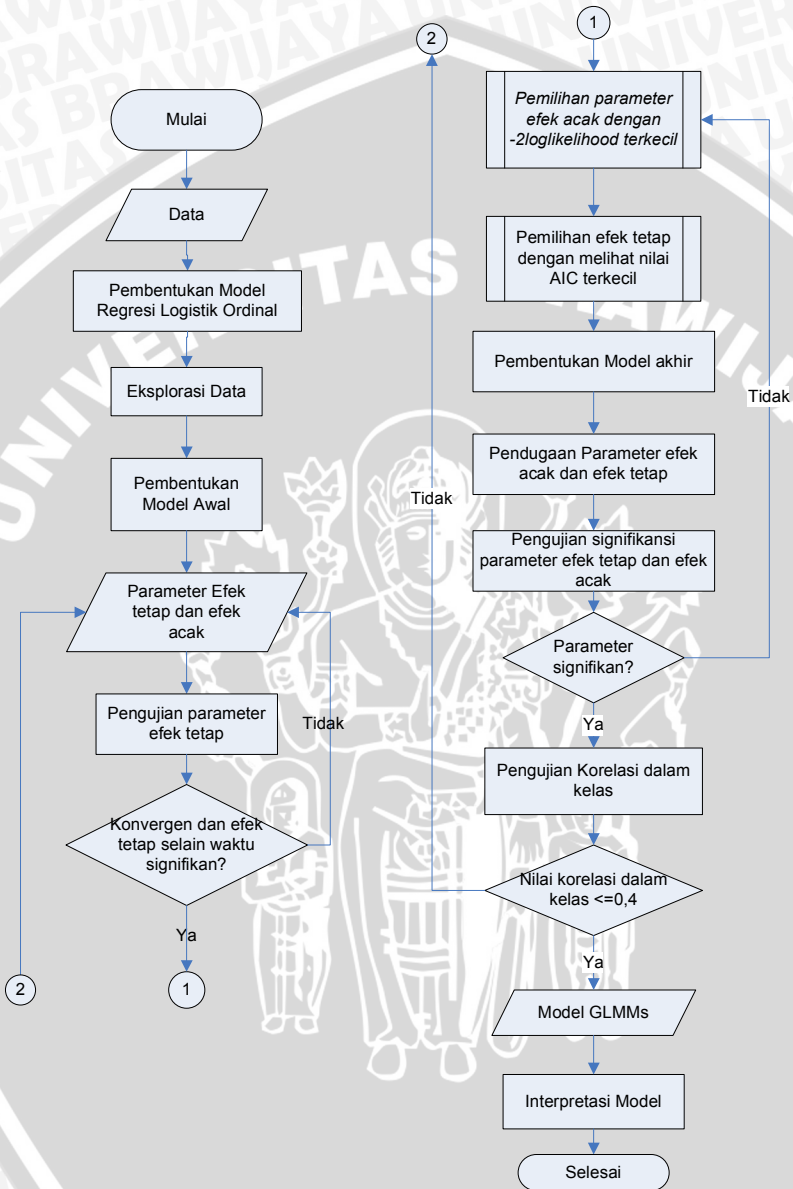
1. Pembentukan model regresi logistik ordinal sebagai inisialisasi efek tetap dalam GLMMs
 - a. Pendugaan parameter model regresi logistik ordinal dengan *Maximum Likelihood Estimator* (MLE) dengan memaksimalkan fungsi likelihood . Didapatkan persamaan yang tidak linier sehingga untuk mendapatkan nilai α dan β menggunakan iterasi *Newton-Raphson* pada persamaan (2.29), Iterasi dilakukan sampai diperoleh batas yang konvergen, yaitu :
$$\left| \hat{\theta}_j^{(t)} - \hat{\theta}_j^{(t-1)} \right| < 10^{-6}$$
 - b. Pengujian signifikansi parameter secara parsial dengan uji *Wald* pada persamaan (2.30). Jika ada parameter yang tidak signifikan maka kembali ke tahap 1a dan dilakukan pengujian parameter tanpa mengikutsertakan parameter yang tidak signifikan
 - c. Pengujian signifikansi parameter secara serentak seperti pada persamaan (2.31).
 - d. Pengujian kesesuaian model menggunakan uji Pearson dan uji Deviance pada persamaan (2.32) dan (2.33). Jika model telah sesuai maka beranjak ke tahap 1e namun jika model tidak sesuai maka kembali ke tahap 1b.
 - e. Model akhir regresi logistik ordinal.
2. Eksplorasi Data pada data longitudinal dengan
 - a. *Mean profile*, jika pola dari proporsi kumulatif menunjukkan pola linier naik atau turun maka efek tetap akan dipertimbangkan dalam pembentukan model tentatif.
 - b. *Variance profile*, jika struktur ragam menunjukkan adanya perubahan di tiap waktu maka perlu penyertaan efek acak selain efek tetap ke dalam model tentatif.
 - c. *Correlation structure*, jika struktur korelasi yang terbentuk untuk setiap waktu berbeda-beda maka struktur korelasi yang digunakan adalah *unstructured correlation* (korelasi tidak terstruktur)

- d. *Concomitant variable*, jika terdapat perbedaan proporsi kumulatif variabel pengiring tiap waktu maka variabel pengiring dipertimbangkan dalam model tentatif.
3. Pembentukan model tentatif dengan penentuan jumlah efek tetap awal dan efek acak sementara. Efek tetap dan efek acak sementara diperoleh dari eksplorasi data, jika *mean profile* dan *variance profile* berpengaruh maka perlu adanya penyertaan efek acak dan efek tetap dalam model.
4. Pendugaan parameter efek tetap pada model awal menggunakan inisialisasi dari hasil pendugaan parameter pada regresi logistik ordinal pada tahap 1b sehingga di peroleh nilai pendugaan awal dari variabel penjelas.
5. Pemeriksaan signifikansi penduga parameter efek tetap pada model awal menggunakan uji Wald pada persamaan (2.30), bila terdapat efek tetap (selain efek tetap waktu) yang tidak signifikan maka kembali ke tahap 3 dan membentuk kembali model awal tanpa mengikutsertakan efek tetap selain waktu yang tidak signifikan tersebut ke dalam model.
6. Pemilihan efek acak dengan menghitung nilai *loglikelihood* semua model efek acak. Nilai *loglikelihood* dibandingkan dengan menggunakan *likelihood ratio test*, model efek acak didapat dari nilai *likelihood ratio test* yang signifikan dengan melihat nilai $-2\loglikelihood$ terkecil.
7. Pemilihan efek tetap berdasarkan efek acak yang telah ditentukan pada tahap 6. Model efek tetap diperoleh dari nilai AIC model secara keseluruhan dengan menyertakan efek acak yang telah diketahui.
8. Pembentukan model dan dilakukan pemeriksaan signifikansi kembali terhadap efek tetap setelah penyertaan efek acak yang telah dipilih melalui tahap 7 ke dalam model.
9. Pemeriksaan korelasi dalam kelas (*intraclass correlation*) dengan menggunakan persamaan (2.48), jika nilai korelasi masih tinggi (lebih besar dari 0,4) maka diperlukan pembentukan model kembali dengan penambahan variabel penjelas yang belum diikutsertakan. Jika korelasi kecil maka dilanjutkan ke tahap 10.
10. Model akhir
11. Interpretasi model

Pembentukan model longitudinal *Generalized Linier Mixed Models* (GLMMs) menggunakan bantuan *software* SAS 9.1.3. Diagram pembentukan regresi logistik ordinal (*Ordinal Logistic Regression*) disajikan pada Gambar 3.1. Diagram pembentukan *Generalized Linier Mixed Models* disajikan pada Gambar 3.2.



Gambar 3.1 Diagram alir Pembentukan model Regresi Logistik Ordinal



Gambar 3.2. Diagram alir Proses pembentukan GLMMs

UNIVERSITAS BRAWIJAYA



BAB IV

HASIL DAN PEMBAHASAN

4.1 Penelitian Penggunaan Mariyuana

Penelitian penggunaan mariyuana adalah informasi penggunaan mariyuana pada 237 individu yang berbeda dalam kurun waktu selama 5 tahun dengan variabel respon ordinal (Y) tiga kategori yaitu tidak pernah (1), tidak lebih dari satu bulan (2), lebih dari satu bulan (3). Data ini merupakan data longitudinal seimbang dengan jumlah pengamatan setiap subyek sama, yaitu selama 5 tahun yang diteliti pada subyek individu berjenis kelamin laki-laki dan perempuan. Variabel penjelas (X) dalam data ini adalah *time* dan *gender*.

4.1.1 Model Regresi Logistik Ordinal

Dalam data longitudinal antar pengamatan terdapat korelasi karena pengaruh pengamatan secara berulang terhadap waktu. Model logistik dibentuk dengan mengabaikan korelasi antar pengamatan tersebut. Pengujian signifikansi untuk data mariyuana dilakukan dengan menggunakan uji serentak *likelihood ratio* dengan statistik khi kuadrat pada uji *Wald* sesuai dengan persamaan (2.32). Hasil pengujian signifikansi parameter model logistik secara parsial dapat dilihat pada Tabel 4.1.

Tabel 4.1 Hasil Pengujian signifikansi model logistik

Variabel bebas	Db	Koefisien	Wald χ^2	P-value	Keterangan
<i>Intercept 1</i>	1	2,8781	192,3960	<0,0001	Signifikan
<i>Intercept 2</i>	1	3,9616	311,0168	<0,0001	Signifikan
<i>Gender</i>	1	-0,6167	20,3382	<0,0001	Signifikan
<i>Time</i>	1	-0,4752	86,6206	<0,0001	Signifikan

Dari Tabel 4.1 dapat dikatakan bahwa variabel *intercept 1* dan *intercept 2* memberikan pengaruh yang signifikan terhadap data mariyuana sedangkan pengaruh dari *gender* dan *time* signifikan terhadap data mariyuana. Setelah dilakukan pengujian signifikansi parameter model logistik secara parsial, langkah berikutnya adalah

pengujian signifikansi parameter model logistik secara serentak yang dapat disajikan pada Tabel 4.2.

Tabel 4.2. Hasil Pengujian signifikansi parameter logistik

Variabel bebas	db	χ^2	P-value	Keterangan
Likelihood Ratio	2	112,9701	<0,0001	Signifikan
Score	2	107,0788	<0,0001	Signifikan
Wald	2	102,0429	<0,0001	Signifikan

Dari tabel 4.2 dapat diketahui bahwa model logistik dengan variabel bebas *intercept*, *time* dan *gender* adalah signifikan untuk α sebesar 0,05. Hal ini dapat dilihat dari *P-value* yang lebih kecil dari α . Dengan demikian, persamaan untuk model logistik ordinal menyertakan *intercept* 1, *intercept* 2, *time* dan *gender* sehingga variabel *gender* berpengaruh terhadap penggunaan mariyuana.

Model logistik ordinal untuk data mariyuana disajikan pada persamaan (4.1) adalah

$$P(y_i \leq 1) = \frac{\exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})}{1 + \exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})}$$

$$P(y_i \leq 2) = \frac{\exp(3,9616 - 0,6167 * \text{gender} - 0,4752 * \text{time})}{1 + \exp(3,9616 - 0,6167 * \text{gender} - 0,4752 * \text{time})} \quad (4.1)$$

4.1.1.1 Uji Kesesuaian Model

Setelah model data mariyuana selanjutnya adalah menguji kesesuaian model. uji kesesuaian model menggunakan uji *Pearson* dan *Deviance*. Dari hasil kesesuaian model uji *Pearson* dan *Deviance* pada persamaan (2.34) dan (2.35) didapatkan χ^2 secara berturut-turut sebesar 19,6154 dan 19,6154. Nilai *P-value* lebih besar daripada α yaitu <0,0001, sehingga diputuskan bahwa H_0 diterima dan modelnya sesuai.

4.1.1.2 Model Akhir Regresi Logistik Ordinal

Dari ketiga model pada persamaan (4.1), didapatkan model peluang untuk regresi logistik ordinal tiap kategori, yaitu:

- a. Untuk $k=1$, Tidak pernah menggunakan mariyuana

$$P(y_i = 1) = \frac{\exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})}{1 + \exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})}$$

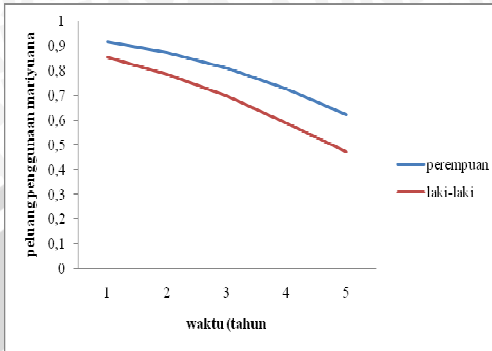
- b. Untuk $k=2$, tidak lebih dari satu bulan menggunakan mariyuana

$$\begin{aligned} P(y_i = 2) &= P(y_i \leq 2) - P(y_i \leq 1) \\ &= \frac{\exp(3,9616 - 0,6167 * \text{gender} - 0,4752 * \text{time})}{1 + \exp(3,9616 - 0,6167 * \text{gender} - 0,4752 * \text{time})} \\ &\quad - \frac{\exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})}{1 + \exp(2,8781 - 0,6167 * \text{gender} - 0,4752 * \text{time})} \end{aligned}$$

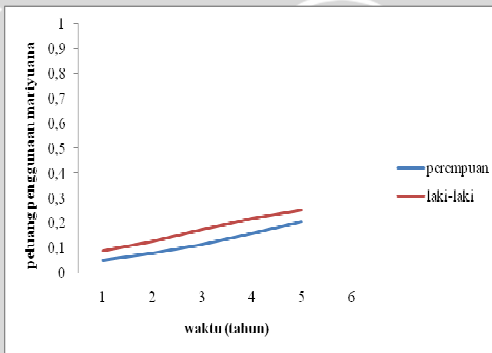
- c. Untuk $k=3$, lebih dari satu bulan menggunakan mariyuana

$$\begin{aligned} P(y_i = 3) &= 1 - P(y_i \leq 2) \\ &= \frac{1}{1 + \exp(3,9616 - 0,6167 * \text{gender} - 0,4752 * \text{time})} \end{aligned} \quad (4.2)$$

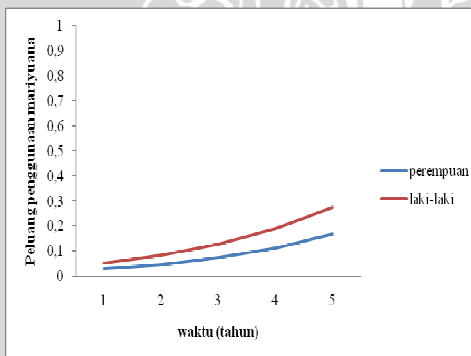
Gambar 4.1 menunjukkan peluang penggunaan mariyuana untuk tiap kategori setiap tahun jika nilai-nilai variabel *time* dan *gender* disubsitusikan ke dalam model logistik ordinal pada persamaan (4.2).



(a)



(b)



(c)

Gambar 4.1 Peluang Penggunaan Mariyuana tiap kategori

(a) $P(Y_i = 1)$, (b) $P(Y_i = 2)$, (c) $P(Y_i = 2)$

Berdasarkan Gambar 4.1 diperoleh kesimpulan bahwa pasien yang tidak menggunakan mariyuana sampai tahun ke-5 mengalami penurunan sedangkan pasien yang menggunakan mariyuana yang kurang lebih selama sebulan sampai tahun ke-5 mengalami peningkatan. Hal ini menyimpulkan bahwa variabel waktu berpengaruh terhadap penggunaan mariyuana. Terlihat dari Gambar 4.1(b) dan (c) yang menggambarkan peluang pasien menggunakan mariyuana, pasien yang berjenis kelamin laki-laki lebih banyak menggunakan mariyuana dibandingkan pasien yang berjenis kelamin perempuan.

4.1.2 Pembentukan GLMMs pada data mariyuana

4.1.2.1 Eksplorasi data

Eksplorasi data merupakan tahap awal pembentukan model data longitudinal menggunakan Generalized Linier Mixed Models (GLMMs). Tahap ini berguna untuk membentuk model awal berdasarkan kesimpulan awal mengenai struktur *mean profile*, *variance profile* dan *concomitant variable*.

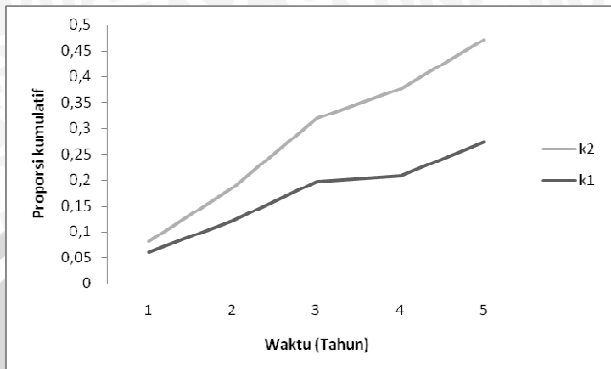
a. Mean Profile

Eksplorasi *mean profile* menggambarkan perubahan proporsi kejadian sukses untuk keseluruhan subyek setiap waktu. *Mean profile* yang terbentuk untuk data mariyuana disajikan pada Gambar 4.2. Garis menunjukkan perubahan proporsi kumulatif penggunaan mariyuana setiap waktu. Garis k1 menunjukkan proporsi kumulatif dari kategori

pertama $\frac{n_2}{n_1 + n_2}$ dan garis k2 menunjukkan proporsi kumulatif dari

kategori kedua $\frac{n_3}{n_1 + n_2 + n_3}$. Dari Gambar 4.2 dapat dilihat fluktuasi

proporsi kumulatif jumlah penggunaan tiap kategori untuk setiap unit waktu. Hal ini menunjukkan bahwa efek waktu mempengaruhi tingkat penggunaan mariyuana.

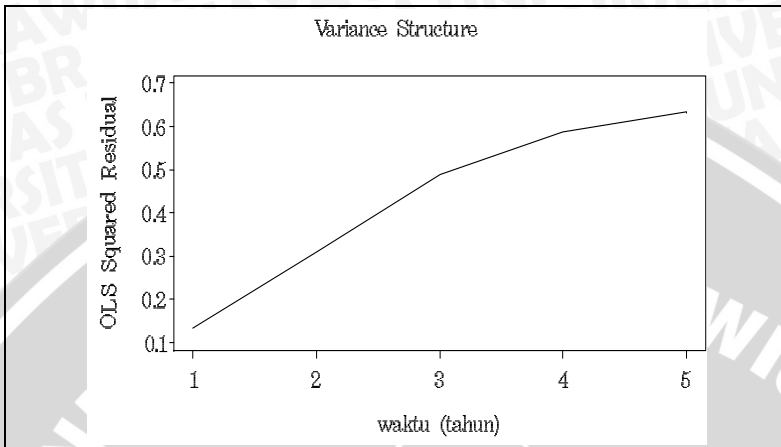


Gambar 4.2 Grafik *mean profile* data mariyuana

Grafik *mean profile* menunjukkan adanya kecenderungan pola naiknya proporsi kumulatif penggunaan mariyuana setiap unit waktu yang menunjukkan adanya korelasi antar pengamatan, hal ini menunjukkan bahwa waktu mempengaruhi tingkat penggunaan mariyuana. Dengan demikian, perlu menambahkan efek tetap dalam model tentatif yang akan dibentuk.

b. *Variance Profile*

Eksplorasi *variance profile* bertujuan untuk mengetahui pengaruh efek acak. Pengaruh efek acak didapatkan dari variasi, yaitu berasal dari nilai residual fungsi link logit yang dikuadratkan. Efek acak perlu ditambahkan dalam model jika dalam grafik *variance profile* menunjukkan perubahan nilai residual tiap unit waktu. *Variance profile* untuk data mariyuana dapat dilihat pada Gambar 4.3



Gambar 4.3 Grafik *variance profile* data mariyuana

Berdasarkan *variance profile* pada Gambar 4.3 menunjukkan adanya perubahan keragaman respon (berupa positif) pada pasien pengguna mariyuana seiring perubahan waktu, hal ini perlu menyertakan efek acak selain efek tetap ke dalam model. Terdapatnya efek acak mengindikasikan bahwa setiap pasien memiliki perilaku yang berbeda, sehingga efek masing-masing individu layak dipertimbangkan dalam model.

c. *Correlation Structure*

Struktur korelasi menggambarkan bagaimana hubungan antar waktu pengamatan pada data longitudinal. Nilai dari struktur korelasi untuk data mariyuana terdapat di Lampiran 5.

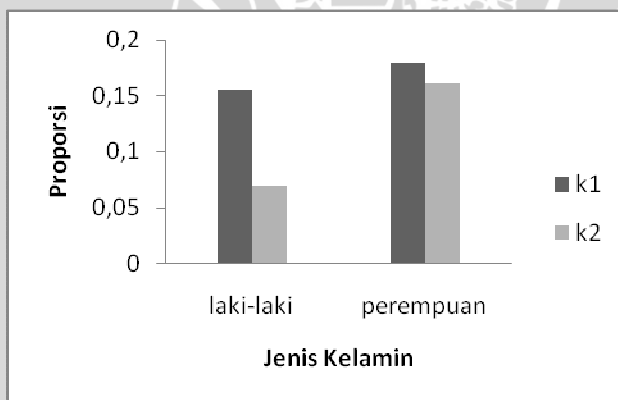
Tabel 4.3. Struktur korelasi data mariyuana

	Time 1	Time 2	Time 3	Time 4	Time 5
Time 1	1	0,46210	0,36368	0,29786	0,20653
Time 2	0,46210	1	0,65378	0,48408	0,39940
Time 3	0,36368	0,65378	1	0,68135	0,57702
Time 4	0,29786	0,48408	0,68135	1	0,67428
Time 5	0,20653	0,39940	0,57702	0,67428	1

Berdasarkan Tabel 4.3 didapatkan *correlation structure* untuk data penelitian mariyuana. Struktur korelasi menunjukkan korelasi antar waktu pengamatan pada individu yang sama selain itu menunjukkan pola-pola tertentu antar waktu pengamatan. Dapat terlihat pada Tabel 4.3 bahwa nilai korelasi antar pengamatan pada selang waktu yang sama menunjukkan adanya pola-pola tertentu yang terbentuk antar selang pengamatan yang berbeda. korelasi sebagai salah satu penentu derajat hubungan antar dua peubah memiliki sifat yang ekuivalen dengan peragam. oleh karena itu berdasarkan struktur korelasi tersebut dapat diprediksi bahwa struktur peragam yang sesuai dalam pembentukan model adalah *unstructured covariance* (peragam tidak terstruktur)

d. *Concomitant Profile*

Eksplorasi terhadap *concomitant profile* atau variabel pengiring perlu juga untuk dilakukan. Pengaruh variabel pengiring tidak dapat diabaikan karena variabel pengiring dapat mempengaruhi respon. Pada data mariyuana terdapat satu variabel pengiring, yaitu jenis kelamin. Eksplorasi terhadap variabel pengiring disajikan pada Gambar 4.4



Gambar 4.4 Grafik eksplorasi jenis kelamin

Dari grafik eksplorasi terhadap variabel pengiring pada Gambar 4.4 dapat diketahui bahwa variabel *gender* memiliki pengaruh yang relatif sama antara laki-laki dan perempuan sehingga perlu dipertimbangkan penyertaan variabel *gender*.

Berdasarkan hal tersebut maka didapatkan model dengan efek *gender* dan tanpa efek *gender* secara berturut-turut ditunjukkan pada persamaan(4.3) dan (4.4) :

$$\log it[P(y_i \leq k)] = \alpha_k + \beta_1 time + \beta_2 gender \quad (4.3)$$

$$\log it[P(y_i \leq k)] = \alpha_k + \beta_1 time \quad (4.4)$$

Selanjutnya dilakukan pengujian masing-masing parameter untuk mengetahui signifikansi. Pengujian signifikansi dilakukan secara serentak maupun secara parsial. Uji serentak dilakukan dengan menggunakan uji *likelihood ratio test*. Hasil uji serentak untuk data mariyuana disajikan pada Tabel 4.4.

Tabel 4.4. Hasil pengujian serentak variabel penjelas

Variabel penjelas	db	χ^2	P-value	Keterangan
<i>Intercept, time, Gender</i>	2	112,9701	<0,0001	Signifikan
<i>Intercept, time</i>	1	92,3203	<0,0001	Signifikan

Uji signifikansi secara parsial terhadap masing-masing variabel respon menggunakan statistik uji Wald. Hasil pengujian variabel penjelas untuk data mariyuana disajikan pada Tabel 4.5.

Table 4.5 Hasil pengujian parsial variabel penjelas

Variabel bebas	db	Wald χ^2	P-value	Keterangan
<i>Intercept1</i>	1	192,3960	<0,0001	Signifikan
<i>Intercept2</i>	1	311,0168	<0,0001	Signifikan
<i>Time</i>	1	20,3382	<0,0001	Signifikan
<i>Gender</i>	1	86,6206	<0,0001	Signifikan

Dari Tabel 4.5 dapat diketahui bahwa variabel penjelas *time* dan *gender* merupakan variabel penjelas yang akan disertakan ke dalam model serta *intercept1* dan *intercept2* juga akan disertakan dalam model. Dengan demikian model tentatif untuk data mariyuana merupakan kombinasi variabel-variabel di atas.

Model tentatif untuk data penelitian mariyuana tampak pada persamaan (4.5)

$$\begin{aligned}\log it(P \leq y_1) &= (\alpha_1 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} + \beta_2 Gender_{ij} \\ \log it(P \leq y_2) &= (\alpha_2 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} + \beta_2 Gender_{ij}\end{aligned}\quad (4.5)$$

Di mana $\alpha_1, \alpha_2, \beta_1, \beta_2$ berturut-turut merupakan efek tetap *intercept1*, *intercept2*, *slop time* dan *gender* untuk subyek ke-*i* dan waktu ke-*j*. Sedangkan b_{0i} dan b_{1i} merupakan efek acak intersep dan *slop time* untuk individu ke-*i*.

4.1.2.2 Pemilihan Efek Acak

Setelah model tentatif diperoleh, selanjutnya akan dilakukan pemilihan terhadap efek acak yang akan disertakan dalam model. pemilihan efek acak meliputi efek acak intersep dan *Time* untuk model pertama, efek acak berupa intersep untuk model kedua dan tanpa efek acak untuk model ketiga. Pemilihan efek acak disajikan pada Tabel 4.6.

Tabel 4.6 Pemilihan efek acak data mariyuana

Efek Acak	-2loglikelihood	$-2\ln\lambda_N$	Db	P-value	Keterangan
<i>Intercept, time</i>	1688,8	14,4	2	<0,000	Signifikan
<i>Intercept</i>	1674,4	3	1	<0,000	Signifikan
Tanpa efek acak	1677,4		1		

berdasarkan Tabel 4.6 diketahui bahwa model efek acak intersep merupakan model efek acak yang akan disertakan dalam model karena model *intercept* memiliki nilai *-2loglikelihood* terkecil sehingga model yang akan dibentuk adalah model efek acak intersep.

4.1.2.3 Pemilihan Efek Tetap

Hasil pemilihan efek acak pada tahap sebelumnya telah didapatkan bahwa efek acak yang akan disertakan adalah efek acak intersep

Table 4.7 Pemilihan Efek Tetap dengan Efek Acak Intersep

Efek Tetap	AIC
<i>Time</i>	1704,9
<i>Time, Gender</i>	1684,4

Berdasarkan pembentukan efek tetap dan efek acak tersebut maka struktur model untuk data mariyuana dapat ditulis seperti persamaan (4.6)

$$\begin{aligned} \log it(P \leq y_1) &= (\alpha_1 + b_{0i}) + \beta_1 Time_{ij} + \beta_2 Gender_{ij} \\ \log it(P \leq y_2) &= (\alpha_2 + b_{1i}) + \beta_1 Time_{ij} + \beta_2 Gender_{ij} \end{aligned} \quad (4.6)$$

di mana β_1 dan β_2 merupakan efek tetap *time* dan *gender* untuk subyek ke-*i* pada waktu ke-*j*.

4.1.2.4 Pendugaan Parameter

Hasil pendugaan efek tetap dan komponen ragam pada model bagi data mariyuana menggunakan metode *Maximum Likelihood* (ML) disajikan pada Tabel 4.8.

Tabel 4.8 Pendugaan Parameter Efek Tetap Model

Parameter	Penduga	Standard Error	χ^2 hitung	P-value	Keterangan
<i>Intercept 1</i>	4,0345	0,8002	25,4016	<0,0001	Signifikan
<i>Intercept 2</i>	5,4995	1,0479	27,5625	<0,0001	Signifikan
<i>Time</i>	-0,6594	0,1398	22,2784	<0,0001	Signifikan
<i>Gender</i>	-0,8804	0,4257	12,3201	0,0074	Signifikan

Dari Tabel 4.8 dapat kita lihat bahwa variabel *time* dan *gender* signifikan pada α sebesar 5%. Nilai ragam untuk data penelitian mariyuana adalah sebesar 2,4774 dengan *standard error* sebesar 1,9176.

4.1.2.5 Pengujian Korelasi Dalam Kelas (*Intraclass Correlation*)

Tahapan terakhir pembentukan model adalah pengujian korelasi dalam subyek. Tujuan dari penyertaan efek acak adalah untuk menghilangkan korelasi dalam subyek. Hasil perhitungan korelasi dalam kelas sesuai pada persamaan (2.48) untuk data mariyuana adalah 0,4294. Nilai korelasi 0,4294 yang menurut Foy (2001) dapat dikategorikan kecil. Hal ini menunjukkan bahwa dengan penambahan efek acak telah mengurangi korelasi dalam subyek. Sehingga dapat disimpulkan bahwa model untuk data mariyuana telah selesai.

4.1.2.6 Model Akhir *Generalized Linier Mixed Models*

Dengan mensubsitusikan hasil pendugaan parameter sebelumnya ke dalam model pada persamaan (4.6) , maka *Generalized Linier Mixed Models* (GLMMs) bagi data mariyuana adalah:

1. Kategori tidak pernah menggunakan mariyuana

$$\log it(P(y_{ij} \leq 1)) = (4,0345 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij}$$

2. Kategori tidak pernah menggunakan mariyuana atau menggunakan mariyuana tidak lebih dari satu bulan

$$\log it(P(y_{ij} \leq 2)) = (5,4995 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij} \quad (4.7)$$

Persamaan (4.7) merupakan model untuk tingkat penggunaan mariyuana dipengaruhi perubahan waktu linier. Keragaman penyimpangan subyek pada intersep (b_0) yang rendah (2,4774) menyimpulkan peluang penggunaan mariyuana masing-masing subyek tidaklah besar. Berdasarkan nilai σ_b^2 sebesar 2,4774 sehingga didapatkan nilai σ_b sebesar 1,5740 dapat disimpulkan bahwa dengan tingkat kesalahan sebesar 5%, maka nilai efek acak intersep pada persamaan (2.37) berkisar antara -3,0850 hingga 3,0850.

Berdasarkan persamaan (4.7), model peluang GLMMs untuk tiap kategori adalah sebagai berikut:

1. Kategori tidak pernah menggunakan mariyuana

$$P(y_{ij} = 1) = \frac{\exp((4,0345 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij})}{1 + \exp((4,0345 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij})}$$

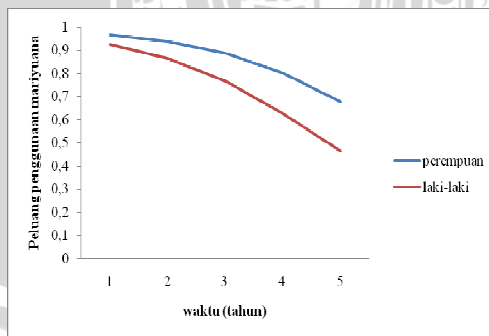
2. Kategori tidak lebih dari satu bulan menggunakan mariyuana

$$P(y_{ij} = 2) = \frac{\exp(5,4995 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij}}{1 + (\exp(5,4995 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij}) + \frac{\exp((4,0345 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij})}{1 + \exp((4,0345 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij})}}$$

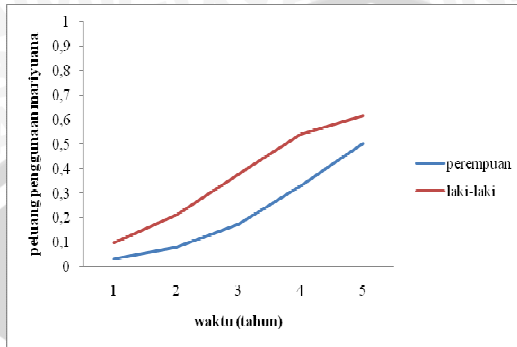
3. Kategori lebih dari satu bulan menggunakan mariyuana

$$P(y_{ij} = 3) = \frac{1}{1 + (\exp(5,4995 + b_{0i}) - 0,6594Time_{ij} - 0,8804Gender_{ij})} \quad (4.8)$$

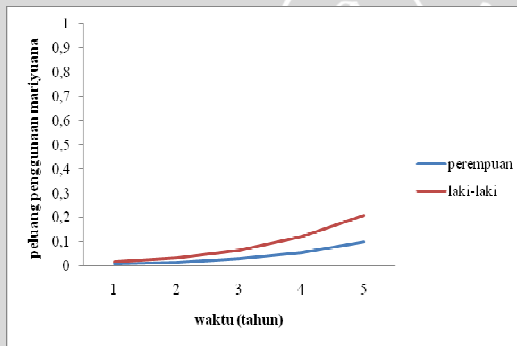
Dengan memasukkan nilai variabel *Gender* bernilai nol untuk jenis kelamin perempuan dan bernilai satu untuk jenis kelamin laki-laki serta menjaga variabel lainnya konstan. Penurunan penggunaan mariyuana yang dipengaruhi oleh waktu (*Time*) pada 237 responden tidak jauh berbeda. Gambar 4.5 akan menunjukkan peluang penggunaan mariyuana setiap unit waktu jika nilai-nilai variabel bebas disubsitusikan ke dalam model persamaan (4.8).



(a)



(b)



(c)

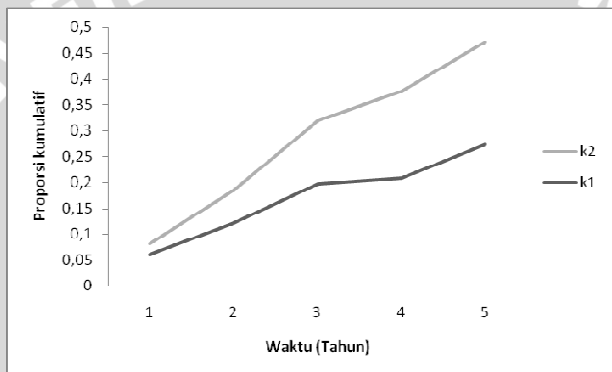
Gambar 4.5 Grafik penggunaan mariyuana tiap kategori
(a) $P(y_{ij} = 1)$, (b) $P(y_{ij} = 2)$, (c) $P(y_{ij} = 3)$

Pada Gambar 4.5(a) dapat kita lihat pada kategori pertama yaitu tidak pernah menggunakan mariyuana, cenderung perempuan lebih banyak tidak menggunakan mariyuana daripada laki-laki. Hal ini mengindikasikan bahwa setiap tahunnya pasien yang tidak menggunakan mariyuana semakin menurun. Gambar 4.5(b) menunjukkan pasien yang menggunakan mariyuana tidak lebih dari satu bulan mengalami peningkatan di tiap tahunnya serta Gambar 4.5(c) menunjukkan pasien yang menggunakan mariyuana lebih dari satu bulan mengalami peningkatan tiap tahun. Pasien berjenis kelamin laki-laki lebih banyak menggunakan mariyuana daripada pasien

berjenis kelamin perempuan. Setiap kelompok *gender* memiliki peluang yang berbeda, hal ini berarti bahwa *gender* sebagai variabel pengiring juga mempengaruhi penggunaan mariyuana.

4.1.2.7 Evolusi Variabel Respon

Pola perubahan proporsi penggunaan mariyuana berdasarkan model persamaan (4.6) disajikan pada Gambar 4.6. Gambar tersebut menunjukkan rata-rata peningkatan proporsi penggunaan mariyuana pada tahun 1976 sampai tahun 1980.



Gambar 4.6 Pola Penggunaan Mariyuana

Berdasarkan Gambar 4.6 pada k1 merupakan proporsi kumulatif penggunaan mariyuana tidak lebih dari satu bulan mengalami peningkatan dari tahun 1976 sampai tahun 1980, begitu juga dengan k2 yang merupakan proporsi kumulatif penggunaan mariyuana lebih dari satu bulan mengalami peningkatan dari tahun 1976 sampai tahun 1980. Setiap tahun terdapat peningkatan mariyuana di kalangan remaja.

4.2 Penelitian Tingkat Keparahan Penyakit

Data keparahan penyakit adalah informasi mengenai keparahan penyakit *Schizophrenia* dari *National Institut of Mental Health (NIMH) Schizophrenia Collaborat Study* pada 437 pasien rawat inap. Variabel respon di hitung berdasarkan 79 item pada IMPS (Inpatient Multidimensional Psychiatric Scale) yang variabel respon (Y) dikategorikan menjadi empat, yaitu: normal atau tidak sakit (1), ringan atau tidak parah (2), cukup parah (3), sangat parah (4). Pada penelitian ini pasien secara acak diberikan empat obat, yaitu *placebo*, *chloropromazine*, *fluphenazine*, atau *thioridazine*. Pada penelitian sebelumnya terjadi efek yang sama pada kelompok obat *antipsychotic* (*chloropromazine*, *fluphenazine*, atau *thioridazine*) dan *placebo* adalah tidak mengkonsumsi obat tetapi berasumsi bahwa dirinya akan sembuh dengan sendirinya sehingga terdapat dua perlakuan kelompok obat yaitu *placebo* (0) dan *drug* (1). Data ini merupakan data longitudinal seimbang dengan jumlah pengamatan setiap subyek sama, yaitu selama 0, 1, 3, dan 6 minggu (X).

4.2.1 Model Logistik Ordinal

Dalam data longitudinal antar pengamatan terdapat korelasi. Model logistik dibentuk dengan mengabaikan korelasi antar pengamatan. Pengujian signifikansi untuk data mawiyuana dilakukan dengan menggunakan uji serentak *likelihood ratio* dengan statistik khi kuadrat pada uji wald sesuai dengan persamaan (2.32). Hasil pengujian signifikansi parameter model logistik secara parsial dapat dilihat pada Tabel 4.9.

Tabel 4.9 Hasil Pengujian signifikansi model logistik

Variabel bebas	Db	Koefisien	Wald χ^2	P-value	Keterangan
<i>Intercept 1</i>	1	-4,0402	678,4713	<0,0001	Signifikan
<i>Intercept 2</i>	1	-1,9975	266,8244	<0,0001	Signifikan
<i>Intercept 3</i>	1	-0,7024	38,8431	<0,0001	Signifikan
<i>Treatment</i>	1	0,8374	53,3853	<0,0001	Signifikan
<i>Time</i>	1	0,4245	330,0199	<0,0001	Signifikan

Dari Tabel 4.9 dapat dikatakan bahwa variabel *intercept 1*, *intercept 2* dan *intercept 3* memberikan pengaruh yang signifikan terhadap data

tingkat keparahan penyakit sedangkan pengaruh dari *treatment* dan *time* signifikan juga terhadap data tingkat keparahan penyakit. Setelah dilakukan pengujian signifikansi parameter model logistik secara parsial, langkah berikutnya adalah pengujian signifikansi parameter model logistik secara serentak yang dapat disajikan pada Tabel 4.10.

Tabel 4.10. Hasil Pengujian signifikansi parameter logistik

Variabel bebas	Db	χ^2	P-value	Keterangan
Likelihood Ratio	2	417,1100	<0,0001	Signifikan
Score	2	371,9683	<0,0001	Signifikan
Wald	2	368,9284	<0,0001	Signifikan

Dari tabel 4.10 dapat diketahui bahwa model logistik dengan variabel bebas *intercept*, *time* dan *treatment* adalah signifikan untuk α sebesar 0,05. Hal ini dapat dilihat dari *P-value* yang lebih kecil dari α . Dengan demikian, persamaan untuk model logistik ordinal menyertakan *intercept 1*, *intercept 2*, *intercept 3*, *time* dan *treatment*. variabel *time* dan *treatment* berpengaruh terhadap keparahan penyakit.

4.2.2.1 Uji Kesesuaian Model

Setelah model data mariyuana selanjutnya adalah menguji kesesuaian model. uji kesesuaian model menggunakan uji *Pearson* dan *Deviance*. Dari hasil kesesuaian model uji *Pearson* dan *Deviance* pada persamaan (2.34) dan (2.35) didapatkan χ^2 secara berturut-turut sebesar 101,1762 dan 98,0597. Nilai *P-value* lebih besar daripada α yaitu <0,0001, sehingga diputuskan bahwa H_0 diterima dan modelnya sesuai.

4.2.2.2 Model Akhir Regresi Logistik Ordinal

Model logistik ordinal untuk data tingkat keparahan penyakit disajikan pada persamaan (4.8) adalah

$$P(y_i \leq 1) = \frac{\exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}$$

$$P(y_i \leq 2) = \frac{\exp(-1,9975 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-1,9975 + 0,9179 * treatment + 1,1079 * time)}$$

$$P(y_i \leq 3) = \frac{\exp(-0,8374 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-0,4245 + 0,8374 * treatment + 0,4245 * time)} \quad (4.9)$$

Dari ketiga model pada persamaan (4.8), didapatkan model peluang untuk regresi logistik ordinal, yaitu:

1. Untuk k=1, normal

$$P(y_i = 1) = \frac{\exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}$$

2. Untuk k=2, parah

$$P(y_i = 2) = P(y_i \leq 2) - P(y_i \leq 1)$$

$$= \frac{\exp(-1,9975 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-1,9975 + 0,8374 * treatment + 0,4245 * time)} - \frac{\exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-4,0402 + 0,8374 * treatment + 0,4245 * time)}$$

3. Untuk k=3, cukup parah

$$P(y_i = 3) = P(y_i \leq 3) - P(y_i \leq 2)$$

$$= \frac{\exp(-0,7024 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-0,7024 + 0,8374 * treatment + 0,4245 * time)} - \frac{\exp(-1,9975 + 0,8374 * treatment + 0,4245 * time)}{1 + \exp(-1,9975 + 0,8374 * treatment + 0,4245 * time)}$$

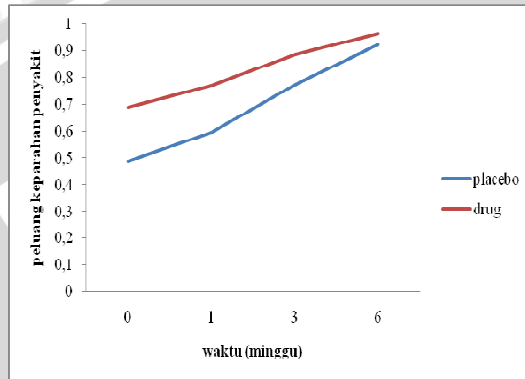
4. Untuk k=4, sangat parah

$$P(y_i = 4) = 1 - P(y_i \leq 4)$$

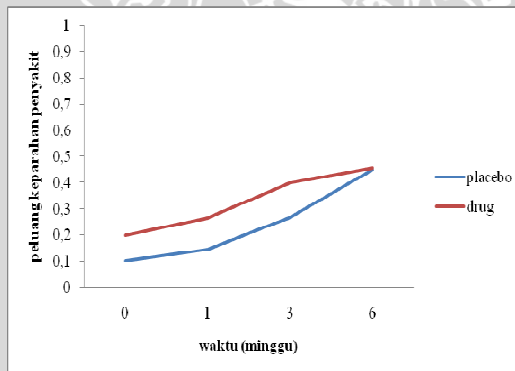
$$= \frac{1}{1 + \exp(-0,7024 + 0,8374 * treatment + 0,4245 * time)}$$

(4.10)

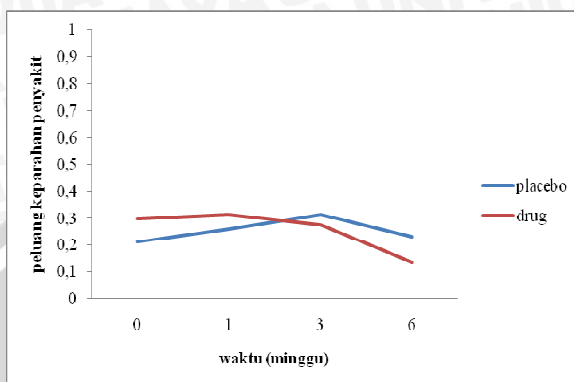
Gambar 4.7 menunjukkan peluang tingkat keparahan penyakit untuk tiap kategori setiap tahun jika nilai-nilai variabel *time* dan *treatment* di subsitusikan ke dalam model logistik ordinal pada persamaan (4.10).



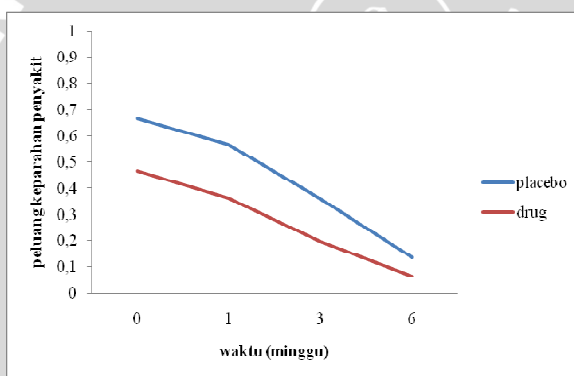
(a)



(b)



(c)



(d)

Gambar 4.7. Grafik Peluang Tingkat keparahan penyakit
(a) $P(y_{ij} = 1)$, (b) $P(y_{ij} = 2)$, (c) $P(y_{ij} = 3)$, (d) $P(y_{ij} = 4)$

Berdasarkan Gambar 4.7 (a), (b), (c) dan (d) diperoleh kesimpulan bahwa pasien yang tidak sakit atau normal sampai minggu ke-6 mengalami peningkatan sedangkan pasien yang tidak parah penyakitnya sampai minggu ke-6 mengalami peningkatan. Begitu juga dengan pasien yang cukup parah penyakitnya sampai minggu ke-6 mengalami peningkatan serta pasien yang sangat parah penyakitnya menurun. Pada Gambar 67 (c) dan (d) dapat kita lihat bahwa untuk pemberian *drug* pada pasien menyebabkan peluang keparahan penyakit menurun, ini menandakan bahwa setiap kelompok *treatment* memiliki

peluang yang berbeda. Hal ini berarti bahwa *treatment* mempengaruhi tingkat keparahan penyakit pasien.

4.2.2 Pembentukan GLMMS pada data tingkat keparahan penyakit

4.2.2.1 Eksplorasi data

Eksplorasi data merupakan tahap awal pembentukan model data longitudinal menggunakan Generalized Linier Mixed Models (GLMMs). Tahap ini berguna untuk membentuk model awal berdasarkan kesimpulan awal mengenai struktur *mean profile*, *variance profile*, *structure correlation* dan *concomitant variable*.

a. Mean Profile

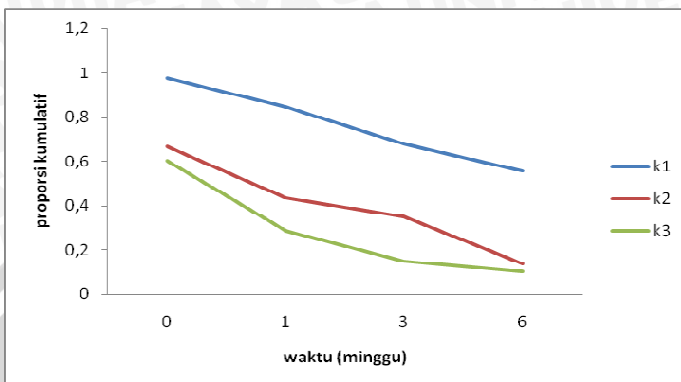
Eksplorasi *mean profile* menggambarkan perubahan proporsi kejadian sukses untuk keseluruhan subyek setiap waktu. *Mean profile* yang terbentuk untuk data keparahan penyakit disajikan pada Gambar 4.2. Garis menunjukkan perubahan proporsi kumulatif penggunaan mariyuana setiap waktu. Garis k1 menunjukkan proporsi kumulatif dari

kategori pertama $\frac{n_2}{n_1 + n_2}$, garis k2 menunjukkan proporsi kumulatif

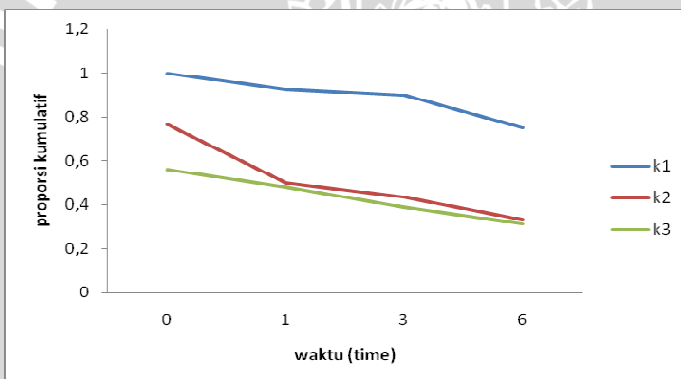
dari kategori kedua $\frac{n_3}{n_1 + n_2 + n_3}$ dan garis k3 menunjukkan proporsi

kumulatif dari kategori ketiga $\frac{n_4}{n_1 + n_2 + n_3 + n_4}$. Dari Gambar 4.2

dapat dilihat fluktuasi proporsi kumulatif jumlah penggunaan tiap kategori untuk setiap unit waktu. Hal ini menunjukkan bahwa efek waktu mempengaruhi tingkat penggunaan mariyuana.



(a)



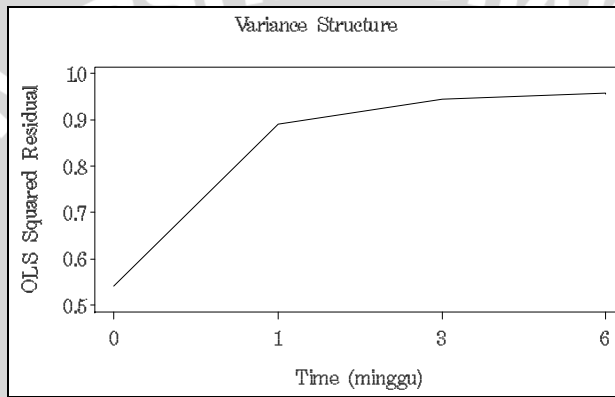
(b)

Gambar 4.8 Grafik *Mean Profile* data keparahan penyakit
(a) *Treatment Placebo*, (b) *treatment Drug*

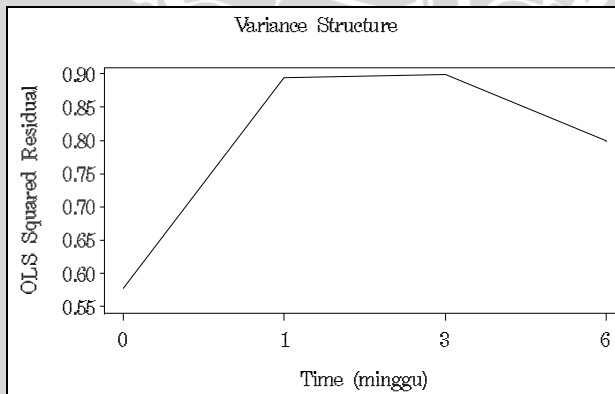
Grafik *mean profile* menunjukkan adanya kecenderungan pola turunnya proporsi keparahan penyakit setiap unit waktu yang menunjukkan adanya korelasi antar pengamatan, hal ini menunjukkan bahwa waktu mempengaruhi tingkat keparahan penyakit. Dengan demikian, perlu menambahkan efek tetap dalam model tentatif yang akan dibentuk.

b. *Variance Profile*

Eksplorasi *variance profile* bertujuan untuk mengetahui pengaruh efek acak. Pengaruh efek acak didapatkan dari variasi, yaitu berasal dari nilai residual fungsi link logit yang dikuadratkan. Efek acak perlu ditambahkan dalam model jika dalam grafik *variance profile* terdapat fluktuasi nilai residual tiap unit waktu. *Variance profile* untuk data mariyuana dapat dilihat pada Gambar 4.9.



(a)



(b)

Gambar 4.9. Grafik *Variance Profile* data keparahan penyakit
(a) *Treatment Placebo* dan (b) *Treatment Drug*

Berdasarkan *variance profile* pada Gambar 4.9 (a) dan (b) menunjukkan adanya perubahan keragaman respon naik turun (pola acak) pada data keparahan penyakit seiring perubahan waktu, untuk setiap *treatment* memiliki proporsi kumulatif yang berbeda. Hal ini menunjukkan perlu menyertakan efek acak selain efek tetap ke dalam model. Terdapatnya efek acak mengindikasikan bahwa setiap pasien memiliki perilaku yang berbeda, sehingga efek masing-masing individu layak dipertimbangkan dalam model.

c. *Correlation Structure*

Struktur korelasi menggambarkan bagaimana hubungan antar waktu pengamatan pada data longitudinal. Nilai dari struktur korelasi untuk data keparahan penyakit terdapat pada Lampiran 6.

Tabel 4.11. Struktur korelasi data keparahan penyakit

	Time 0	Time 1	Time 3	Time 6
Time 0	1	0,4405	0,2894	0,1669
Time 1	0,4405	1	0,5882	0,4610
Time 3	0,2894	0,5882	1	0,6163
Time 6	0,1669	0,4610	0,6163	1

Berdasarkan Tabel 4.11 didapatkan *correlation structure* untuk data penelitian tingkat keparahan penyakit. Struktur korelasi menunjukkan korelasi antar waktu pengamatan pada individu yang sama selain itu menunjukkan pola-pola tertentu antar waktu pengamatan. dapat terlihat pada tabel bahwa nilai korelasi antar pengamatan pada selang waktu yang sama menunjukkan adanya pola-pola tertentu yang terbentuk antar selang pengamatan yang berbeda. korelasi sebagai salah satu penentu derajat hubungan antar dua peubah memiliki sifat yang ekuivalen dengan peragam. oleh karena itu berdasarkan struktur korelasi tersebut dapat diprediksi bahwa struktur peragam yang sesuai dalam pembentukan model adalah *unstructured covariance* (peragam tidak terstruktur)

d. *Concomitant Profile*

Eksplorasi terhadap *concomitant profile* atau variabel pengiring perlu juga untuk dilakukan. Pengaruh variabel pengiring tidak dapat diabaikan karena variabel pengiring dapat mempengaruhi respon. Pada

data keparahan penyakit tidak terdapat variabel pengiring, sehingga tidak dilakukan eksplorasi *concomitant profile*.

Berdasarkan hasil eksplorasi data yang mempertimbangkan struktur efek tetap waktu (*time*) dan efek acak tersebut maka didapatkan model dengan *treatment* dan tanpa *treatment* secara berturut-turut ditunjukkan pada persamaan(4.11) dan (4.12) :

$$\log it[P(y_i \leq k)] = \alpha_k + \beta_1 time + \beta_2 treatment \quad (4.11)$$

$$\log it[P(y_i \leq k)] = \alpha_k + \beta_1 time \quad (4.12)$$

Selanjutnya dilakukan pengujian masing-masing parameter untuk mengetahui signifikansi dari variabel bebas. Pengujian signifikansi dilakukan secara serentak maupun secara parsial. Uji serentak dilakukan dengan menggunakan uji *likelihood ratio test*. Hasil uji serentak untuk data keparahan penyakit disajikan pada Tabel 4.12.

Tabel 4.12. Hasil pengujian serentak variabel bebas

Variabel bebas	db	χ^2	P-value	Keterangan
<i>Intercept, time, treatment</i>	2	417,1100	<0,0001	Signifikan
<i>Intercept, time</i>	1	362,5693	<0,0001	Signifikan

Uji signifikansi secara parsial terhadap masing-masing variabel respon menggunakan statistik uji Wald. Hasil pengujian variabel penjelas untuk data tingkat keparahan penyakit disajikan pada Tabel 4.13.

Table 4.13 Hasil pengujian parsial variabel bebas

Variabel bebas	Db	Koefisien	Wald χ^2	P-value	Keterangan
<i>Intercept 1</i>	1	-4,0402	678,4713	<0,0001	Signifikan
<i>Intercept 2</i>	1	-1,9975	266,8244	<0,0001	Signifikan
<i>Intercept 3</i>	1	-0,7024	38,8431	<0,0001	Signifikan
<i>treatment</i>	1	0,8374	53,3853	<0,0001	Signifikan
<i>time</i>	1	0,4245	330,0199	<0,0001	Signifikan

Dari Tabel 4.13 dapat diketahui bahwa variabel penjelas *time* dan *treatment* merupakan variabel penjelas yang akan disertakan ke dalam

model serta *intercept1*, *intercept2*, dan *intercept3* juga akan disertakan dalam model. Dengan demikian model tentatif untuk data tingkat keparahan penyakit merupakan kombinasi variabel-variabel di atas. Model tentatif untuk data penelitian tingkat keparahan penyakit tampak pada persamaan (4.13)

$$\log it(P \leq y_1) = \begin{cases} (\alpha_1 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_2 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_3 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \end{cases} \quad \text{Placebo}$$

$$\log it(P \leq y_1) = \begin{cases} (\alpha_4 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_5 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_6 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \end{cases} \quad \text{Drug}$$

(4.13)

Di mana $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$ berturut-turut merupakan efek tetap intersep1, intersep2, intersep3, slop *time* dan untuk subyek ke-i dan waktu ke-j. Sedangkan b_{0i} dan b_{1i} merupakan efek acak intersep dan slop *Time* untuk individu ke-i.

4.2.2.2 Pemilihan Efek Acak

Setelah model tentatif diperoleh, selanjutnya akan dilakukan pemilihan terhadap efek acak yang akan disertakan dalam model. pemilihan efek acak meliputi efek acak intersep dan slop *time* untuk model pertama, efek acak berupa intersep untuk model kedua dan tanpa efek acak untuk model ketiga. Pemilihan efek acak disajikan pada Tabel 4.14.

Tabel 4.14 Pemilihan efek acak data keparahan penyakit

Efek Acak	-2loglikelihood	$-2\ln \lambda_N$	Db	P-value	Keterangan
<i>Intercept, time</i>	-	-	-	-	Tidak konvergen
<i>Intercept</i>	3733,6	4	1	<0,000	Signifikan
Tanpa efek acak	3737,6				

Dari Tabel 4.14 diketahui bahwa model efek acak intersep merupakan model efek acak yang akan disertakan dalam model karena memiliki nilai *-2loglikelihood* terkecil sehingga model yang akan dibentuk adalah model efek acak intersep.

4.2.2.3 Pemilihan Efek Tetap

Hasil pemilihan efek acak pada tahap sebelumnya telah didapatkan bahwa efek acak yang akan disertakan adalah efek acak intersep.

Table 4.15 Pemilihan Efek Tetap dengan Efek Acak Intersep

Efek Tetap	AIC
<i>Time</i>	3801,0
<i>Time, treatment</i>	3745,6

Berdasarkan pembentukan efek tetap dan efek acak tersebut maka struktur model untuk data keparahan penyakit dapat ditulis seperti persamaan (4.14).

$$\log it(P \leq y_1) = \begin{cases} (\alpha_1 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_2 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_3 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \end{cases} \quad placebo$$

$$\log it(P \leq y_1) = \begin{cases} (\alpha_4 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_5 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \\ (\alpha_6 + b_{0i}) + (\beta_1 + b_{1i})Time_{ij} \end{cases} \quad drug$$

(4.14)

di mana β_1 dan β_2 merupakan efek tetap *time* untuk subyek ke-*i* pada waktu ke-*j* dan b_{0i} adalah efek acak intersep.

4.2.2.4 Pendugaan Parameter

Hasil pendugaan efek tetap dan komponen ragam pada model bagi data keparahan penyakit menggunakan metode *Maximum Likelihood* (ML) disajikan pada Tabel 4.16.

Tabel 4.16 Pendugaan Parameter Efek Tetap Model

Parameter	Penduga	Standard Error	χ^2 hitung	P-value	Ket
<i>Intercept 1 (Placebo)</i>	-5,9888	0,5199	132,7104	<0,0001	Signifikan
<i>Intercept 2 (Placebo)</i>	-2,7016	0,3411	62,7264	<0,0001	Signifikan
<i>Intercept 3 (Placebo)</i>	-0,4765	0,2855	2,7889	0,0981	Tidak Signifikan
<i>Intercept 1 (Drug)</i>	-4,7093	0,2194	460,9609	<0,0001	Signifikan
<i>Intercept 2 (Drug)</i>	-1,8363	0,1372	155,7504	<0,0001	Signifikan
<i>Intercept 3 (Drug)</i>	0,04952	0,1320	0,1444	0,7077	Tidak Signifikan
<i>Time (Placebo)</i>	0,3179	0,05687	31,2481	<0,0001	Signifikan
<i>Time (Drug)</i>	0,6835	0,03592	361,3801	<0,0001	Signifikan

Dari Tabel 4.16 dapat kita lihat bahwa variabel *intercept* dan *time* signifikan pada α sebesar 5%. Nilai ragam untuk data penelitian tingkat keparahan penyakit adalah sebesar 1,8069 dengan *standard error* sebesar 1,1039.

4.2.2.5 Pengujian Korelasi Dalam Kelas (Intraclass Correlation)

Tahapan terakhir pembentukan model adalah pengujian korelasi dalam subyek. Tujuan dari penyertaan efek acak adalah untuk menghilangkan korelasi dalam subyek. Hasil perhitungan korelasi dalam kelas sesuai pada persamaan (2.50) untuk data keparahan penyakit adalah 0,2512. menurut Foy (2001) nilai korelasi sebesar 0,2512 dapat dikategorikan rendah. Hal ini menunjukkan bahwa dengan penambahan efek acak telah mengurangi korelasi dalam subyek.

4.2.2.6 Model Akhir *Generalized Linier Mixed Models*

Dengan mensubstitusikan hasil pendugaan parameter sebelumnya ke dalam model pada persamaan (4.14) , maka *Generalized Linier Mixed Models* (GLMMs) bagi data tingkat keparahan penyakit adalah:

1. Kategori normal

$$\log it(P(y_{ij} \leq 1)) = \begin{cases} (-5,9888 + b_{0i}) + 0,3179Time_{ij} & Placebo \\ (-4,7093 + b_{0i}) + 0,6835Time_{ij} & Drug \end{cases}$$

2. Kategori normal atau tidak parah

$$\log it(P(y_{ij} \leq 2)) = \begin{cases} (-2,7016 + b_{0i}) + 0,3179Time_{ij} & Placebo \\ (-1,8363 + b_{0i}) + 0,6835Time_{ij} & Drug \end{cases}$$

3. Kategori normal atau tidak parah atau cukup parah

$$\log it(P(y_{ij} \leq 3)) = \begin{cases} (-0,4765 + b_{0i}) + 0,3179Time_{ij} & Placebo \\ (0,04952 + b_{0i}) + 0,6835Time_{ij} & Drug \end{cases}$$

(4.15)

Persamaan 4.15 merupakan model untuk tingkat keparahan penyakit dipengaruhi perubahan waktu linier. Keragaman penyimpangan subyek pada intersep (b_0) yang rendah (1,8069) menyimpulkan peluang keparahan penyakit masing-masing subyek tidaklah besar. Berdasarkan

nilai σ_b^2 sebesar 1,8069 sehingga didapatkan nilai σ_b sebesar 1,3442 dapat disimpulkan bahwa dengan tingkat kesalahan sebesar 5%, maka nilai efek acak intersep pada persamaan (2.37) berkisar antara -2,6347 hingga 2,6347.

Berdasarkan persamaan 4.15, model peluang GLMMs untuk tiap kategori adalah sebagai berikut:

1. Kategori normal (1)

$$P(y_{ij} = 1) = \begin{cases} \frac{\exp((-5,9888 + b_{0i}) + 0,3179Time_{ij})}{1 + \exp((-5,9888 + b_{0i}) + 0,3179Time_{ij})} & \text{Placebo} \\ \frac{\exp((-4,7093 + b_{0i}) + 0,6835Time_{ij})}{1 + \exp((-4,7093 + b_{0i}) + 0,6835Time_{ij})} & \text{Drug} \end{cases}$$

2. Kategori cukup parah (2)

$$P(y_{ij} = 2) = \begin{cases} \frac{\exp((-2,7016 + b_{0i}) + 0,3179Time_{ij})}{1 + \exp((-2,7016 + b_{0i}) + 0,3179Time_{ij})} - \frac{\exp((-5,9888 + b_{0i}) + 0,3179Time_{ij})}{1 + \exp((-5,9888 + b_{0i}) + 0,3179Time_{ij})} & \text{Placebo} \\ \frac{\exp((-1,8363 + b_{0i}) + 0,6835Time_{ij})}{1 + \exp((-1,8363 + b_{0i}) + 0,6835Time_{ij})} - \frac{\exp((-4,7093 + b_{0i}) + 0,6835Time_{ij})}{1 + \exp((-4,7093 + b_{0i}) + 0,6835Time_{ij})} & \text{Drug} \end{cases}$$

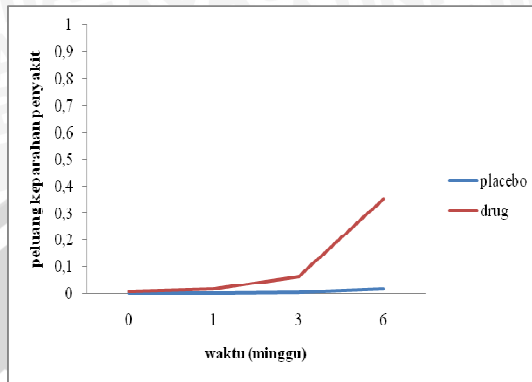
3. Kategori parah (3)

$$P(y_{ij} = 3) = \begin{cases} \frac{\exp((-0,4765 + b_{0i}) + 0,3179Time_{ij})}{1 + \exp((-0,4765 + b_{0i}) + 0,3179Time_{ij})} & \text{Placebo} \\ - \frac{\exp((-2,7016 + b_{0i}) + 0,3179Time_{ij})}{1 + \exp((-2,7016 + b_{0i}) + 0,3179Time_{ij})} \\ \frac{\exp((0,04952 + b_{0i}) + 0,6835Time_{ij})}{1 + \exp((0,04952 + b_{0i}) + 0,6835Time_{ij})} \\ - \frac{\exp((-1,8363 + b_{0i}) + 0,6835Time_{ij})}{1 + \exp((-1,8363 + b_{0i}) + 0,6835Time_{ij})} & \text{Drug} \end{cases}$$

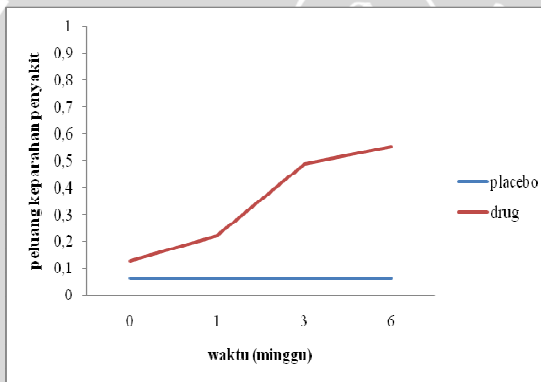
4. Kategori sangat parah (4)

$$P(y_{ij} = 4) = \begin{cases} \frac{1}{1 + \exp((-0,4765 + b_{0i}) + 0,3179Time_{ij})} & \text{Placebo} \\ \frac{1}{1 + \exp((0,04952 + b_{0i}) + 0,6835Time_{ij})} & \text{Drug} \end{cases} \quad (4.16)$$

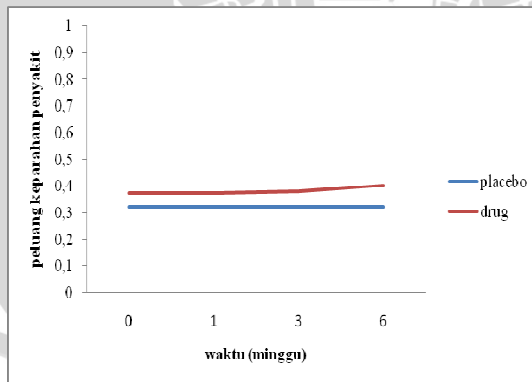
Dengan memasukkan nilai variabel *treatment* untuk perlakuan *placebo* dan *drug* tidak berbeda secara signifikan. Tingkat keparahan suatu penyakit yang dipengaruhi oleh waktu (*Time*) pada 437 pasien tidak jauh berbeda. Gambar 4.10 akan menunjukkan peluang keparahan penyakit setiap unit waktu jika nilai-nilai variabel bebas disubsitusikan ke dalam model persamaan (4.16).



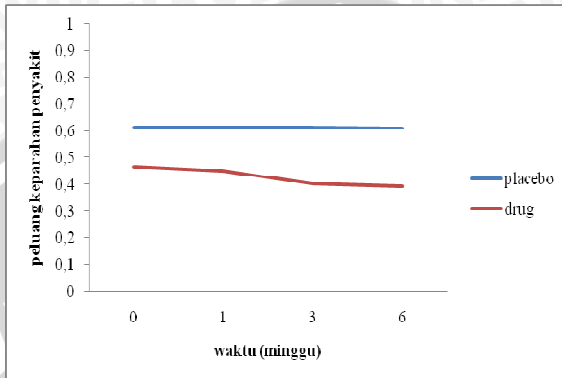
(a)



(b)



(c)



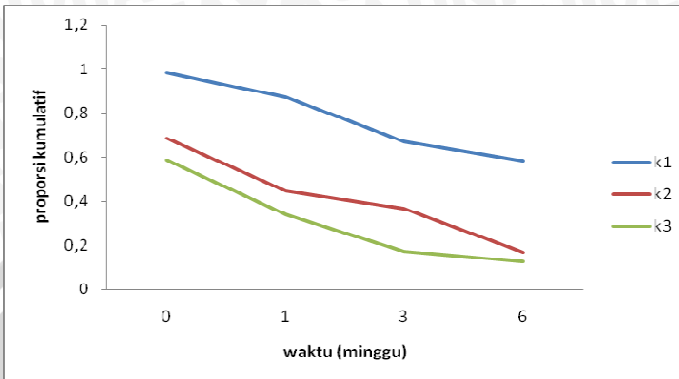
(d)

Gambar 4.10 Grafik keparahan penyakit tiap kategori $P(y_{ij} = 1)$, (b) $P(y_{ij} = 2)$, (c) $P(y_{ij} = 3)$, dan $P(y_{ij} = 4)$

Pada Gambar 4.10 (a), (b), (c) dan (d) dapat kita lihat peluang atau tingkat keparahan penyakit untuk kategori normal dengan pemberian *treatment drug* semakin meningkatkan. Untuk kategori tidak parah, cukup parah dan sangat parah dengan pemberian *treatment drug* akan menurunkan peluang keparahan penyakit sehingga sampai minggu ke-6 banyak pasien yang kondisinya mulai membaik. Setiap kelompok *treatment* memiliki peluang yang berbeda, hal ini berarti bahwa *treatment* mempengaruhi tingkat keparahan pasien.

4.2.2.7 Evolusi Variabel Respon

Pola perubahan proporsi keparahan penyakit berdasarkan model persamaan (4.15) disajikan pada Gambar 4.11. Gambar tersebut menunjukkan rata-rata penurunan proporsi keparahan penyakit pada minggu ke-0 sampai minggu ke-6.



Gambar 4.11 Pola keparahan penyakit

Pada k1 merupakan proporsi kumulatif keparahan penyakit yang cukup parah mengalami penurunan dari minggu ke-0 sampai minggu ke-6, begitu juga dengan k2 yang merupakan proporsi kumulatif proporsi keparahan penyakit yang parah mengalami penurunan dari minggu ke-0 sampai minggu ke-6. Pada k1 merupakan proporsi kumulatif keparahan penyakit yang sangat parah mengalami penurunan dari minggu ke-0 sampai minggu ke-6. Hal ini berarti pada minggu ke-6 pasien yang di amati secara keseluruhan telah mengalami kesembuhan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil perhitungan yang dilakukan dapat diambil kesimpulan penggunaan nilai awal dari Regresi Logistik Ordinal memudahkan dalam inisialisasi model tentatif untuk memperoleh pendugaan dari parameter model GLMMs. Hasil model GLMMs yang terbentuk untuk kedua data pada penelitian ini antara lain:

1. Penelitian terhadap data penggunaan mariyuana, di mana perlakuan untuk setiap subyek sama. Model yang terbentuk untuk data ini adalah :

- a. Kategori tidak pernah menggunakan mariyuana

$$P(y_{ij} = 1) = \frac{\exp((4,0345 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij})}{1 + \exp((4,0345 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij})}$$

- b. Kategori tidak lebih dari satu bulan menggunakan mariyuana

$$P(y_{ij} = 2) = \frac{\exp(5,4995 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij}}{1 + (5,4995 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij}} - \frac{\exp((4,0345 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij})}{1 + \exp((4,0345 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij})}$$

- c. Kategori lebih dari satu bulan menggunakan mariyuana

$$P(y_{ij} = 3) = \frac{1}{1 + (5,4995 + b_{0i}) - 0,6594 \text{ Time}_{ij} - 0,8804 \text{ Gender}_{ij}}$$

Di mana b_{0i} adalah efek acak intersep. Peluang remaja berjenis kelamin laki-laki menggunakan mariyuana lebih besar daripada remaja berjenis kelamin perempuan. Hal ini berarti bahwa *gender* dan *time* mempengaruhi penggunaan mariyuana.

2. Penelitian terhadap data penyakit *schizophrenia*, di mana perlakuan untuk setiap subyek tidak sama. Model yang terbentuk untuk data ini adalah :

a. Kategori normal (1)

$$P(y_{ij} = 1) = \begin{cases} \frac{\exp((-5,9888 + b_{0i}) + 0,3179 \text{ Time}_{ij})}{1 + \exp((-5,9888 + b_{0i}) + 0,3179 \text{ Time}_{ij})} & \text{Placebo} \\ \frac{\exp((-4,7093 + b_{0i}) + 0,6835 \text{ Time}_{ij})}{1 + \exp((-4,7093 + b_{0i}) + 0,6835 \text{ Time}_{ij})} & \text{Drug} \end{cases}$$

b. Kategori cukup parah (2)

$$P(y_{ij} = 2) = \begin{cases} \frac{\exp((-2,7016 + b_{0i}) + 0,3179 \text{ Time}_{ij})}{1 + \exp((-2,7016 + b_{0i}) + 0,3179 \text{ Time}_{ij})} - \frac{\exp((-5,9888 + b_{0i}) + 0,3179 \text{ Time}_{ij})}{1 + \exp((-5,9888 + b_{0i}) + 0,3179 \text{ Time}_{ij})} & \text{Placebo} \\ \frac{\exp((-1,8363 + b_{0i}) + 0,6835 \text{ Time}_{ij})}{1 + \exp((-1,8363 + b_{0i}) + 0,6835 \text{ Time}_{ij})} - \frac{\exp((-4,7093 + b_{0i}) + 0,6835 \text{ Time}_{ij})}{1 + \exp((-4,7093 + b_{0i}) + 0,6835 \text{ Time}_{ij})} & \text{Drug} \end{cases}$$

c. Kategori parah (3)

$$P(y_{ij} = 3) = \begin{cases} \frac{\exp((-0,4765 + b_{0i}) + 0,3179 \text{ Time}_{ij})}{1 + \exp((-0,4765 + b_{0i}) + 0,3179 \text{ Time}_{ij})} - \frac{\exp((-2,7016 + b_{0i}) + 0,3179 \text{ Time}_{ij})}{1 + \exp((-2,7016 + b_{0i}) + 0,3179 \text{ Time}_{ij})} & \text{Placebo} \\ \frac{\exp((0,04952 + b_{0i}) + 0,6835 \text{ Time}_{ij})}{1 + \exp((0,04952 + b_{0i}) + 0,6835 \text{ Time}_{ij})} - \frac{\exp((-1,8363 + b_{0i}) + 0,6835 \text{ Time}_{ij})}{1 + \exp((-1,8363 + b_{0i}) + 0,6835 \text{ Time}_{ij})} & \text{Drug} \end{cases}$$

d. Kategori sangat parah (4)

$$P(y_{ij} = 4) = \begin{cases} \frac{1}{1 + \exp((-0,4765 + b_{0i}) + 0,3179 \text{ Time}_{ij})} & \text{Placebo} \\ \frac{1}{1 + \exp((0,04952 + b_{0i}) + 0,6835 \text{ Time}_{ij})} & \text{Drug} \end{cases}$$

Di mana b_{0i} adalah efek acak intersep. Peluang pemberian *drug* pada pasien *schizophrenia* akan menurunkan tingkat keparahan penyakit *schizophrenia* sehingga banyak pasien yang kondisinya mulai membaik. Setiap kelompok *treatment* memiliki peluang yang berbeda, hal ini berarti bahwa *treatment* dan *time* mempengaruhi tingkat keparahan penyakit *schizophrenia*.

5.2 Saran

Pada penelitian ini dilakukan analisis data longitudinal dengan variabel respon ordinal menggunakan *Generalized Linear Mixed Model* (GLMMs) untuk mengatasi variabel respon yang berkorelasi sehingga untuk penelitian berikutnya disarankan untuk membentuk model selain GLMMs, karena tidak semua data dengan variabel respon berdistribusi multinomial dapat di bentuk dalam model ini.

UNIVERSITAS BRAWIJAYA



DAFTAR PUSTAKA

- Agresti, A. 2002. **An Introduction to Categorical Data Analysis**. John Wiley and Sons. New York
- Carriere, I and J. Bouyer. 2006. **Random-effect models for ordinal responses: Application to self-reported disability among older persons**. http://hal.archives-ouvertes.fr/docs/00/07/76/15/PDF/Article2-RESP_revison3c.pdf. Tanggal akses: 20 Agustus 2010.
- Danardono. 2003. **Biostatistika dan Epidemiologi**. Universitas Gajah Mada. Yogyakarta.
- Everitt, B.S. 2002. **The Cambridge Dictionary of Statistics**. Cambridge University Press. Cambridge.
- Fahrmeir, L. and G. Tutz. 1994. **Multivariate Statistical Modelling Based on Generalized Linier Mixed Model**. John Willey and Sons. New York
- Fitmaurice, G.M, N.M. Laird and J.H. Ware. 2004. **Applied Longitudinal Analysis**. <Http://biosun1.harvard.edu/~fitzmaur/ala/>. Tanggal akses: 20 Agustus 2010.
- Foy, P. 2001. **Intraclass Correlation and Variance Components as Population Attributes and Measures of Sampling Efficiency in Pirls 2001**. http://www.iea.nl/fileadmin/user_upload/IRC2004/Foy.pdf. Tanggal akses : 20 September 2010.
- Greene, W.H. 1997. **Econometric Analysis**. McMillan Publishing Company. New York
- Gujarati, D. 2003. **Ekonometrika Dasar**. Alih bahasa oleh Zain, S. Erlangga. Jakarta

Gunawan, B. 2009. **Analisis Longitudinal Diskrit Dengan Respon Binom Menggunakan Generalized Linier Mixed Models (GLMMS)**. Fakultas MIPA Universitas Brawijaya, Malang. Skripsi. Tidak dipublikasikan

Hardin, J. W. and J. Hilbe. 2007. **Generalized Linear Models and Extensions**. Stata Press. Texas

Hedeker, D. and R.D. Gibbons. 2006. **Longitudinal Data Analysis**. John Wiley & Sons. New York.

Hosmer, D.W. and S. Lemeshow. 2000. **Applied Logistic Regression, Second Edition**. John Wiley and Sons Inc. New York

Hyun, S.K. 2004. **Topics In Ordinal Logistic Regression and Its Application**. <http://etd-tamu-2004B-STAT-Kim-2.pdf>. Tanggal akses: 20 Agustus 2010

McCullagh, P and J.A. Nelder. 2003. **Regression Models for Ordinal Data**. Journal of the Royal Statistical Society, Series B (Methodological), Volume 42, Issue 2 (1980), 109-142. Tanggal akses : 20 Agustus 2010

Molenberghs, G and G. Verbeke. 2005. **Models for Discrete Longitudinal Data**. Springer Series in Statistics. New York.

Weiss, R.E. 2005. **Modeling Longitudinal Data**. Springer, New York.

Wu, H and Z. T. Zhang. 2006. **Nonparametric Regression Method for Longitudinal Data Analysis**. John Wiley and Sons Inc. New Jersey

Verbeke, G and G. Molenberghs. 2000. **Linier Mixed Model for Longitudinal Data**. Springer Series in Statistics. New York.

Lampiran 1. Data Mariyuana

Subyek	Time	Gender	Respon
1	1	0	1
1	2	0	1
1	3	0	1
1	4	0	1
1	5	0	1
2	1	0	1
2	2	0	1
2	3	0	1
2	4	0	1
2	5	0	1
3	1	0	1
3	2	0	1
.	.	.	.
.	.	.	.
.	.	.	.
237	1	1	3
237	2	1	3
237	3	1	3
237	4	1	3
237	5	1	3

Keterangan :

Subyek : pasien pengguna mariyuana

Time : waktu pengamatan (tahun)

Respon : Tidak pernah (1), Tidak lebih dari satu bulan (2), Lebih dari satu bulan (3)

Gender : perempuan (0) dan laki-laki (1)

Lampiran 2. Data keparahan penyakit

id	time	respon	Treatment
1103	0	4	1
1103	1	2	1
1103	3	2	1
1103	6	2	1
1104	0	4	1
1104	1	2	1
1104	3	1	1
1104	6	2	1
1105	0	2	1
1105	1	2	1
1105	3	1	1
1105	6	-9	1
1106	0	2	1
1106	1	1	1
1106	3	1	1
1106	6	1	1
1107	0	3	0
.	.	.	.
.	.	.	.
.	.	.	.
9316	0	4	0
9316	1	4	0
9316	3	4	0
9316	6	4	0

Keterangan :

id : pasien rawat inap

Time : waktu pengamatan 0, 1, 3, 6 minggu

Respon : normal (1) , cukup parah (2) , parah (3), sangat parah(4)

treatment : Placebo (0), Drug (1)

Lampiran 3. *Syntax SAS 9.1.3 untuk Pembentukan Generalized Linear Mixed Models* pada data mariyuana

```
data mariyuana;
input subyek time gender respon;
cards;
```

```
1      1      0      1
1      2      0      1
1      3      0      1
1      4      0      1
1      5      0      1
2      1      0      1
2      2      0      1
2      3      0      1
```

```
:      :      :      :
:      :      :      :
:      :      :      :
```

```
236    2      1      3
236    3      1      3
236    4      1      3
236    5      1      1
237    1      1      3
237    2      1      3
237    3      1      3
237    4      1      3
237    5      1      3
```

```
;;
run;
proc print data=mariyuana;
run;
proc means data=mariyuana;
by subyek;
var respon;
run;
/*--Variance Structure ---*/
```

```
proc glm data=mariyuana;
```

Lampiran 3 (lanjutan Syntax data mariyuana)

```
model respon= time / p;  
output out=out1 residual=res;  
run;  
proc sort data=out1;  
by subyek;  
run;  
proc print data=out1;  
run;  
data out1;  
set out1;  
res2=res*res;  
run;  
proc transpose data=out1 out=out2 prefix=res;  
by subyek; id time;  
var res; run;  
proc print data=out2; run;  
ods select pearsoncorr;  
ods output pearsoncorr=pearsoncorr;  
title ' Correlation Matrix for residual';  
proc corr data=out2;  
var res1--res5;  
run;  
proc sort data=out1;  
by time;  
run;  
proc means data=out1 nway noprint;  
var res2;  
by time;  
output out=means2 mean=mean var=var;  
run;  
goptions reset=all ftext=centx htext=1.5;  
title 'Variance Structure';  
axis1 label=(' waktu (tahun) ') offset=(0.5cm)  
order=(1 to 5 by 1) minor=none;  
axis2 label=(angle=90 ' OLS Squared Residual ')  
minor=none ;  
proc gplot data=means2;  
plot mean*time /haxis=axis1 vaxis=axis2;  
symbol1 i=j c=black r=1 line=1;
```

Lampiran 3 (lanjutan Syntax data mariyuana)

```
run;
quit;
data mariyuana;
set mariyuana;
time2=time*time;
time3=time2*time;
time4=time3*time;
run;

proc logistic data=work.data1;
model respon =gender time;
run;

proc logistic data=work.data1;
model respon =time;
run;

proc nlmixed data=work.data1 tech=newrap;
parms i1=2.8781 i2=3.9616 beta1=-0.4752
      beta2=-0.6167; ;
eta1 = i1 + beta1*time + beta2*gender;
eta2 = i2 + beta1*time + beta2*gender;
if (respon=1) then z=1/(1+exp(-eta1));
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-
eta1));
else z=1-(1/(1+exp(-eta2)));
ll=log(z);
model respon ~ general(ll); subject=subyek;
run;

proc nlmixed data=work.data1 tech=newrap;
parms i1=2.8781 i2=3.9616 d=1
      beta1=-0.4752 beta2=-0.6167 ;
eta1 = i1 + beta1*time + beta2*gender + b;
eta2 = i2 + beta1*time + beta2*gender + b;
if (respon=1) then z=1/(1+exp(-eta1));
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-
eta1));
else z=1-(1/(1+exp(-eta2)));
ll=log(z);
model respon ~ general(ll);
random b ~ normal(0,d**2) subject=subyek;
```

Lampiran 3 (lanjutan Syntax data mariyuana)

```
estimate 'var(b)' d**2;  
run;
```

```
proc nlmixed data=work.data1 tech=newrap;  
parms i1=2.8781 i2=3.9616 d11=1 d12=0 d22=0  
      beta1=-0.4752 beta2=-0.6167 ;  
eta1 = i1 + b0 + beta1*time + b1*time + beta2*gender ;  
eta2 = i2 + b0 + beta1*time + b1*time + beta2*gender ;  
if (respon=1) then z=1/(1+exp(-eta1));  
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-  
eta1));  
else z=1-(1/(1+exp(-eta2)));  
ll=log(z);  
model respon ~ general(ll);  
random b0 b1 ~ normal([0,0], [d11, d12, d22])  
subject=subyek;  
run;
```

```
proc nlmixed data=work.data1 tech=newrap;  
parms i1=2.8781 i2=3.9616 d=1  
      beta1=-0.4752 beta2=-0.6167 ;  
eta1 = i1 + beta1*time + beta2*gender + b;  
eta2 = i2 + beta1*time + beta2*gender + b;  
if (respon=1) then z=1/(1+exp(-eta1));  
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-  
eta1));  
else z=1-(1/(1+exp(-eta2)));  
ll=log(z);  
model respon ~ general(ll);  
random b ~ normal(0,d**2) subject=subyek;  
estimate 'var(b)' d**2;  
run;
```

```
proc nlmixed data=work.data1 tech=newrap;  
parms i1=2.8781 i2=3.9616 d=1  
      beta1=-0.4752 ;  
eta1 = i1 + beta1*time + b;  
eta2 = i2 + beta1*time + b;  
if (respon=1) then z=1/(1+exp(-eta1));  
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-  
eta1));  
else z=1-(1/(1+exp(-eta2)));  
ll=log(z);
```

Lampiran 3 (lanjutan Syntax data mariyuana)

```
model respon ~ general(11);  
random b ~ normal(0,d**2) subject=subyek;  
estimate 'var(b)' d**2;  
run;
```

```
proc nlmixed data=work.data1 tech=newwrap;  
parms i1=2.8781 i2=3.9616 d=1  
      beta1=-0.4752 beta2=-0.6167 ;  
eta1 = i1 + beta1*time + beta2*gender + b;  
eta2 = i2 + beta1*time + beta2*gender + b;  
if (respon=1) then z=1/(1+exp(-eta1));  
else if (respon=2) then z=1/(1+exp(-eta2))-1/(1+exp(-  
eta1));  
else z=1-(1/(1+exp(-eta2)));  
ll=log(z);  
model respon ~ general(11);  
random b ~ normal(0,d**2) subject=subyek;  
estimate 'var(b)' d**2;  
run;
```



Lampiran 4. Syntax SAS 9.1.3 untuk Pembentukan *Generalized Linear Mixed Models* pada data keparahan penyakit

```

data dua;
input id      time      respon treatment;
cards;
1103      0      4      1
1103      1      2      1
1103      3      2      1
1103      6      2      1
1104      0      4      1
1104      1      2      1
1104      3      1      1
1104      6      2      1
1105      0      2      1
1105      1      2      1
1105      3      1      1
:      :      :      :
:      :      :      :
:      :      :      :
9314      1      2      1
9314      3      2      1
9314      6      2      1
9315      0      4      1
9315      1      4      1
9315      3      3      1
9315      6      4      1
9316      0      4      0
9316      1      4      0
9316      3      4      0
9316      6      4      0
; ;
run;

/* deleted missing value*/
data dua;
set dua;
if respon=-9 then delete;
run;
proc means data=dua;

```

Lampiran 4 (lanjutan *Syntax* data keparahan penyakit)

```
by id;
var respon;
run;
/*--Variance Structure ---*/

proc glm data=dua;
model respon= time / p;
output out=out1 residual=res;
run;
proc sort data=out1;
by id;
run;
proc print data=out1;
run;
data out1;
set out1;
res2=res*res;
run;
proc transpose data=out1 out=out2 prefix=res;
by id; id time;
var res; run;
proc print data=out2; run;
ods select pearsoncorr;
ods output pearsoncorr=pearsoncorr;
title ' Correlation Matrix for residual';
proc corr data=out2;
var res0--res6;
run;
proc sort data=out1;
by time;
run;
proc means data=out1 nway noprint;
var res2;
by time;
output out=means2 mean=mean var=var;
run;
options reset=all ftext=centx htext=1.5;
title 'Variance Structure';
```

Lampiran 4 (lanjutan Syntax data keparahan penyakit)

```
axis1 label =(' Time (minggu) ') offset=(0.5cm)
order=(0,1,3,6) minor=none;
axis2 label=(angle=90 ' OLS Squared Residual ')
minor=none ;
proc gplot data=means2;
plot mean*time /haxis=axis1 vaxis=axis2;
symbol1 i=j c=black r=1 line=1;
run;
quit;
data dua;
set dua;
time2=time*time;
time3=time2*time;
time4=time3*time;
run;

PROC LOGISTIC data=dua;
MODEL respon = treatment time;
RUN;
proc genmod data=dua;
model respon = time treatment time*treatment /
dist=multinomial link=cumlogit;
run;
proc genmod data=dua;
model respon = treatment time*treatment /
dist=multinomial link=cumlogit;
run;
PROC NL MIXED DATA=dua tech=newrap;
PARMS B1=0.2265 d=1 i1=-3,8996 i2=-1.5376 i3=-
0.2655;
eta1 = i1 + B1*time + b;
eta2 = i2 + B1*time + b;
eta3 = i3 + B1*time + b;
IF (respon=1) THEN
  z = 1 / (1 + EXP(-eta1));
ELSE IF (respon=2) THEN
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));
ELSE IF (respon=3) THEN
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));
```

Lampiran 4 (lanjutan Syntax data keparahan penyakit)

```
ELSE IF (respon=4) THEN
  z = 1 - (1 / (1 + EXP(-eta3)));
  ll = log(z);
MODEL respon ~ GENERAL(ll);
run;
PROC NLMIXED DATA=dua tech=newrap;
PARMS B1=0.2265 d=1 i1=-3,8996 i2=-1.5376 i3=-
0.2655;
eta1 = i1 + B1*time + b;
eta2 = i2 + B1*time + b;
eta3 = i3 + B1*time + b;
IF (respon=1) THEN
  z = 1 / (1 + EXP(-eta1));
ELSE IF (respon=2) THEN
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));
ELSE IF (respon=3) THEN
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));
ELSE IF (respon=4) THEN
  z = 1 - (1 / (1 + EXP(-eta3)));
  ll = log(z);
MODEL respon ~ GENERAL(ll);
random b ~ normal(0,d**2) subject=id;
estimate 'var(b)' d**2;
run;
PROC NLMIXED DATA=dua tech=newrap;
PARMS B1=0.2265 d=1 i1=-3,8996 i2=-1.5376 i3=-
0.2655;;
eta1 = i1 + b0 + B1*time + b1*time;
eta2 = i2 + b0 + B1*time + b1*time;
eta3 = i3 + b0 + B1*time + b1*time;
IF (respon=1) THEN
  z = 1 / (1 + EXP(-eta1));
ELSE IF (respon=2) THEN
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));
ELSE IF (respon=3) THEN
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));
ELSE IF (respon=4) THEN
  z = 1 - (1 / (1 + EXP(-eta3)));
  ll = log(z);
```

Lampiran 4 (lanjutan Syntax data keparahan penyakit)

```
MODEL respon ~ GENERAL(l1);
RANDOM b0 b1 ~ NORMAL([0,0],[d11,d12,d22])
SUBJECT=id ;
RUN;
PROC NLMIXED DATA=dua tech=newrap;
PARMS B1=0.4845 d=1 i1=-3,3625 i2=-1.2993
i3=0.0136;
eta1 = i1 + B1*time + b;
eta2 = i2 + B1*time + b;
eta3 = i3 + B1*time + b;
IF (respon=1) THEN
  z = 1 / (1 + EXP(-eta1));
ELSE IF (respon=2) THEN
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));
ELSE IF (respon=3) THEN
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));
ELSE IF (respon=4) THEN
  z = 1 - (1 / (1 + EXP(-eta3)));
l1 = log(z);
MODEL respon ~ GENERAL(l1);
run;
PROC NLMIXED DATA=dua tech=newrap;
PARMS B1=0.4845 d=1 i1=-3,3625 i2=-1.2993
i3=0.0136;
eta1 = i1 + B1*time + b;
eta2 = i2 + B1*time + b;
eta3 = i3 + B1*time + b;
IF (respon=1) THEN
  z = 1 / (1 + EXP(-eta1));
ELSE IF (respon=2) THEN
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));
ELSE IF (respon=3) THEN
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));
ELSE IF (respon=4) THEN
  z = 1 - (1 / (1 + EXP(-eta3)));
l1 = log(z);
MODEL respon ~ GENERAL(l1);
random b ~ normal(0,d**2) subject=id;
estimate 'var(b)' d**2;
```

Lampiran 4 (lanjutan Syntax data keparahan penyakit)

```
run;  
PROC NLMIXED DATA=dua tech=newrap;  
PARMS B1=0.4845 d=1 i1=-3,3625 i2=-1.2993  
i3=0.0136;  
eta1 = i1 + b0 + B1*time + b1*time;  
eta2 = i2 + b0 + B1*time + b1*time;  
eta3 = i3 + b0 + B1*time + b1*time;  
IF (respon=1) THEN  
  z = 1 / (1 + EXP(-eta1));  
ELSE IF (respon=2) THEN  
  z = (1/(1 + EXP(-eta2))) - (1/(1 + EXP(-eta1)));  
ELSE IF (respon=3) THEN  
  z = (1/(1 + EXP(-eta3))) - (1/(1 + EXP(-eta2)));  
ELSE IF (respon=4) THEN  
  z = 1 - (1 / (1 + EXP(-eta3)));  
l1 = log(z);  
MODEL respon ~ GENERAL(l1);  
RANDOM b0 b1 ~ NORMAL([0,0],[d11,d12,d22])  
SUBJECT=id ;  
RUN;
```

Lampiran 5. Output SAS 9.1.3 untuk Pembentukan *Generalized Linear Mixed Models* data mariyuana

The CORR Procedure

Pearson Correlation Coefficients, N = 237
Prob > |r| under H0: Rho=0

	res1	res2	res3	res4	res5
res1	1.00000 <.0001	0.46210 <.0001	0.36368 <.0001	0.29786 <.0001	0.20653 0.0014
res2	0.46210 <.0001	1.00000	0.65378 <.0001	0.48408 <.0001	0.39940 <.0001
res3	0.36368 <.0001	0.65378 <.0001	1.00000	0.68135 <.0001	0.57702 <.0001
res4	0.29786 <.0001	0.48408 <.0001	0.68135 <.0001	1.00000	0.67428 <.0001
res5	0.20653 0.0014	0.39940 <.0001	0.57702 <.0001	0.67428 <.0001	1.00000

The SAS System 10:56 Sunday, May 30, 2010 23

The LOGISTIC Procedure

Model Information

Data Set	WORK.DATA1	
Response Variable	respon	respon
Number of Response Levels	3	
Model	cumulative logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	1185
Number of Observations Used	1185

Response Profile

Ordered Value	respon	Total Frequency
1	1	874
2	2	175
3	3	136

Probabilities modeled are cumulated over the lower Ordered Values.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Lampiran 5 (lanjutan *Output* SAS 9.1.3 data mariyuana)

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
7.9178	2	0.0191

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	1794.409	1685.439
SC	1804.564	1705.749
-2 Log L	1790.409	1677.439

The SAS System

10:56 Sunday, May 30, 2010 24

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	112.9701	2	<.0001
Score	107.0788	2	<.0001
Wald	102.0429	2	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept 1	1	2.8781	0.2075	192.3960	<.0001
Intercept 2	1	3.9616	0.2246	311.0168	<.0001
sex	1	-0.6167	0.1367	20.3382	<.0001
time	1	-0.4752	0.0511	86.6206	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
sex	0.540	0.413 0.706
time	0.622	0.563 0.687

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.7	Somers' D	0.367
Percent Discordant	27.0	Gamma	0.405
Percent Tied	9.3	Tau-a	0.155
Pairs	295614	c	0.684

The NLMIXED Procedure

Specifications

Data Set	WORK.MARIYUANA
Dependent Variable	respon
Distribution for Dependent Variable	General
Optimization Technique	Newton-Raphson
Integration Method	None

Dimensions

Observations Used	1185
Observations Not Used	0
Total Observations	1185
Parameters	4

Lampiran 5 (lanjutan *Output* SAS 9.1.3 data mariyuana)

```

Parameters
      i1      i2      beta1      beta2      NegLogLike
      2.8781      3.9616      -0.4752      -0.6167      838.71928

Iteration History
      Iter      Calls      NegLogLike      Diff      MaxGrad      Slope
      1          12      838.719278      1.957E-6      3.078E-6      -3.91E-6

NOTE: GCONV convergence criterion satisfied.

Fit Statistics
      -2 Log Likelihood      1677.4
      AIC (smaller is better)      1685.4
      AICC (smaller is better)      1685.5
      BIC (smaller is better)      1705.7

Parameter Estimates
      Parameter      Estimate      Standard      DF      t Value      Pr > |t|      Alpha      Lower      Upper      Gradient
      i1      2.8779      0.2075      1185      13.87      <.0001      0.05      2.4708      3.2850      -6.9E-7
      i2      3.9614      0.2241      1185      17.67      <.0001      0.05      3.5217      4.4012      -3.11E-7
      beta1      -0.4752      0.05105      1185      -9.31      <.0001      0.05      -0.5753      -0.3750      -3.08E-6
      beta2      -0.6165      0.1368      1185      -4.51      <.0001      0.05      -0.8850      -0.3481      -2.94E-8
      Variance Structure      14:57 Sunday, May 30, 2010 465

The NL MIXED Procedure

Specifications
Data Set      WORK.MARIYUANA
Dependent Variable      respon
Distribution for Dependent Variable      General
Random Effects      b
Distribution for Random Effects      Normal
Subject Variable      subyek
Optimization Technique      Newton-Raphson
Integration Method      Adaptive Gaussian Quadrature

Dimensions
Observations Used      1185
Observations Not Used      0
Total Observations      1185
Subjects      1185
Max Obs Per Subject      1
Parameters      5
Quadrature Points      5

Parameters
      i1      i2      d      beta1      beta2      NegLogLike
      2.8781      3.9616      1      -0.4752      -0.6167      844.418724

Iteration History
      Iter      Calls      NegLogLike      Diff      MaxGrad      Slope
      1          14      837.751627      6.667097      2.262314      -13.0895
      2          21      837.539609      0.212018      16.33917      -0.93943
      3          28      837.209054      0.330555      0.111158      -0.65835
      4          35      837.20583      0.003224      0.05817      -0.00652
      5          42      837.205824      6.105E-6      6.668E-6      -0.00001

```

Lampiran 5 (lanjutan *Output* SAS 9.1.3 data mariyuana)

NOTE: ABGCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	1674.4
AIC (smaller is better)	1684.4
AICC (smaller is better)	1684.5
BIC (smaller is better)	1709.8
Variance Structure	14:57 Sunday, May 30, 2010 466

The NL MIXED Procedure

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper	Gradient
i1	4.0345	0.8002	1184	5.04	<.0001	0.05	2.4645	5.6045	9.869E-7
i2	5.4995	1.0479	1184	5.25	<.0001	0.05	3.4435	7.5555	1.4E-6
d	1.5740	0.6091	1184	2.58	0.0099	0.05	0.3789	2.7691	2.913E-6
beta1	-0.6591	0.1398	1184	-4.72	<.0001	0.05	-0.9333	-0.3848	6.668E-6
beta2	-0.8804	0.2507	1184	-3.51	0.0005	0.05	-1.3723	-0.3885	1.14E-6

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
var(b)	2.4774	1.9176	1184	1.29	0.1966	0.05	-1.2848	6.2396

Lampiran 6. (lanjutan *Output* SAS 9.1.3 data keparahan penyakit)

Correlation Matrix for residual

377

01:35 Monday,

May 31, 2010

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	res0	res1	res3	res6
res0	1.00000 434	0.44048 <.0001 423	0.28937 <.0001 371	0.16695 0.0023 332
res1	0.44048 <.0001 423	1.00000 426	0.58820 <.0001 368	0.46100 <.0001 328
res3	0.28937 <.0001 371	0.58820 <.0001 368	1.00000 374	0.61063 <.0001 320
res6	0.16695 0.0023 332	0.46100 <.0001 328	0.61063 <.0001 320	1.00000 335

The SAS System

11:32 Wednesday, May

19, 2010 26

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	417.1100	2	<.0001
Score	371.9683	2	<.0001
Wald	368.9284	2	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept 1	1	-4.0402	0.1551	678.4713	<.0001
Intercept 2	1	-1.9975	0.1223	266.8244	<.0001
Intercept 3	1	-0.7024	0.1127	38.8431	<.0001
treatment	1	0.8374	0.1146	53.3853	<.0001
time	1	0.4245	0.0234	330.0199	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
treatment	2.310	1.846	2.892
time	1.529	1.460	1.601

Association of Predicted Probabilities and Observed Responses

Percent Concordant	64.3	Somers' D	0.432
Percent Discordant	21.1	Gamma	0.506

Percent Tied Pairs	14.5 889763	Tau-a c	0.313 0.716
--------------------	----------------	------------	----------------

The NLMIXED Procedure

Specifications

Data Set	WORK.DUA
Dependent Variable	respon
Distribution for Dependent Variable	General
Optimization Technique	Newton-Raphson
Integration Method	None

Dimensions

Observations Used	1569
Observations Not Used	0
Total Observations	1569
Parameters	7

Parameters

B1	B2	d	i1	i2	i3	b	NegLogLike
0.8374	0.4245	1	-4.0402	-1.9975	-0.7024	1	2095.41707

Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1*	18	1874.6052	220.8119	128.1596	-479.081
2*	27	1868.83556	5.769635	4.775779	-11.083
3*	36	1868.80622	0.029345	0.018966	-0.05839
4*	45	1868.80621	1.837E-6	2.124E-6	-3.67E-6

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	3737.6
AIC (smaller is better)	3751.6
AICC (smaller is better)	3751.7
BIC (smaller is better)	3789.1

Variance Structure

01:35 Monday, May 31, 2010 388

The NLMIXED Procedure

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper	Gradient
B1	0.8374	0.1151	1569	7.28	<.0001	0.05	0.6116	1.0632	-6.2E-8
B2	0.4245	0.02349	1569	18.07	<.0001	0.05	0.3784	0.4706	-8.3E-7
d	1.0000	0	1569	Infty	<.0001	0.05	-Infty	Infty	0
i1	-4.1539	0.07790	1569	-53.32	<.0001	0.05	-4.3067	-4.0011	7.414E-7
i2	-2.1112	0.04398	1569	-48.00	<.0001	0.05	-2.1974	-2.0249	-2.12E-6
i3	-0.8162	0.04555	1569	-17.92	<.0001	0.05	-0.9055	-0.7268	1.308E-6
b	0.1137	0.09107	1569	1.25	0.2120	0.05	-0.06493	0.2923	-7.41E-8

Lampiran 6. (lanjutan *Output* SAS 9.1.3 data keparahan penyakit)

The NL MIXED Procedure

Specifications

Data Set	WORK.DUA
Dependent Variable	respon
Distribution for Dependent Variable	General
Random Effects	b
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Newton-Raphson
Integration Method	Adaptive Gaussian Quadrature

Dimensions

Observations Used	1569
Observations Not Used	0
Total Observations	1569
Subjects	1569
Max Obs Per Subject	1
Parameters	6
Quadrature Points	3

Parameters

B1	B2	d	i1	i2	i3	NegLogLike
0.8374	0.4245	1	-4.0402	-1.9975	-0.7024	1882.31341

Iteration History

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	16	1867.48846	14.82495	4.896792	-28.3452
2	24	1866.96313	0.525332	3.96523	-0.92155
3	32	1866.81645	0.146676	0.389294	-0.26762
4	40	1866.81111	0.005341	0.03862	-0.01049
5	48	1866.8111	9.101E-6	0.00002	-0.00002

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	3733.6
AIC (smaller is better)	3745.6
AICC (smaller is better)	3745.7
BIC (smaller is better)	3777.8

Variance Structure 01:35 Monday, May 31, 2010 390

The NL MIXED Procedure

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper	Gradient
B1	1.1143	0.2037	1568	5.47	<.0001	0.05	0.7147	1.5139	2.222E-6
B2	0.5474	0.07073	1568	7.74	<.0001	0.05	0.4086	0.6861	-4.64E-8
d	1.3442	0.4106	1568	3.27	0.0011	0.05	0.5388	2.1496	0.000011
i1	-5.2376	0.6549	1568	-8.00	<.0001	0.05	-6.5223	-3.9530	0.000015
i2	-2.6242	0.3569	1568	-7.35	<.0001	0.05	-3.3242	-1.9242	8.427E-6
i3	-0.9256	0.1832	1568	-5.05	<.0001	0.05	-1.2850	-0.5662	-0.00002

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > t	Alpha	Lower	Upper
var(b)	1.8069	1.1039	1568	1.64	0.1019	0.05	-0.3584	3.9722