Diajukan kepada Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Yogyakarta

Untuk Memenuhi Sebagian Persyaratan

Guna Mamperoleh Gelar Sarjana Sains

Oleh:

Anis Septiana Sari

NIM. 08305141039

PROGRAM STUDI MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI YOGYAKARTA

2012

ANALISIS REGRESI LOGISTIK ORDINAL DUA LEVEL

DAN APLIKASINYA

Oleh
Anis Septiana Sari
NIM. 08305141039

ABSTRAK

Penelitian ini bertujuan untuk mendeskripsikan analisis regresi logistik dua level pada data ordinal dan aplikasinya. Analisis regresi logistik ordinal dua level digunakan pada data berhirarki untuk mengetahui hubungan antara variabel tak bebas berskala ordinal dengan variabel-variabel bebasnya. Pada saat variabel tak bebas berskala ordinal, maka model multilevel linier tidak dapat digunakan. Oleh karena itu digunakan pendekatan model logit.

Penelitian ini menggunakan metode Maksimum Likelihood untuk mengestimasi parameter pada analisis regresi logistik ordinal dua level. Parameter diestimasi dengan memaksimumkan fungsi turunan pertama. Analisis ini menggunakan *software* HLM (*Hierarchical Linear and Nonlinear Modeling)* dalam mengestimasi parameter.

Analisis regresi logistik ordinal dua level dapat diaplikasikan untuk mengetahui pengaruh faktor-faktor sekolah (level satu) dan faktor-faktor kecamatan (level dua) terhadap nilai rata-rata ujian IPA tiap Sekolah Dasar di Kota Yogyakarta. Variabel bebas dalam penelitian ini yaitu variabel kepemilikan laboratorium dan variabel rasio kepemilikan buku pada level satu serta variabel MSES (*Mean Socioeconomic Status*) pada level dua. Variabel tak bebas pada penelitian ini berupa data ordinal yaitu nilai rata-rata ujian IPA tiap Sekolah Dasar Kota Yogyakarta yang diklasifikasikan menjadi empat kategori. Hasil penelitian ini menunjukan fungsi *logit* dari regresi logistik ordinal dua level. Variabel-variabel bebas pada fungi logit semuanya tidak signifikan. Variabel-variabel tersebut yaitu variabel kepemilikan laboratorium, variabel rasio kepemilikan buku dan variabel MSES (*Mean Socioeconomic Status*). Hal tersebut diduga karena kurangnya pemanfaatan yang optimal pada laboratorium. Fungsi *logit* pada penelitian ini dapat digunakan untuk mencari peluang komulatif maupun peluang dari masing-masing kategori variabel tak bebas.

Kata kunci :Regresi Logistik Ordinal, Analisis Multilevel, SES (*Socioeconomic Status*)

BAB I

PENDAHULUAN

A. Latar Belakang

Suatu penelitian seringkali meneliti hubungan antara beberapa variabel bebas terhadap variabel tak bebas. Salah satu analisis yang digunakan untuk mengetahui hubungan ini dengan analisis regresi. Analisis regresi terbagi menjadi dua yaitu analisis regresi linier dan analisis regresi non linier. Analisis regresi linier digunakan apabila terdapat hubungan linier antara variabel bebas dengan variabel tak bebasnya. Hubungan linier dapat dijelaskan melalui plot antara variabel bebas dengan variabel tak bebas. Apabila letak titik-titik objek dalam diagram pencar XY (X sebagai variabel bebas dan Y sebagai variabel tak bebas) berada di sekitar garis lurus, maka bisa diduga merupakan regresi linier. Jika letak titik-titik itu tidak di sekitar garis lurus, bisa lengkung, menyebar atau lainnya maka dapat diduga merupakan regresi nonlinier.

Di kehidupan nyata, seringkali dalam suatu penelitian tidak terdapat hubungan linier antara variabel bebas dengan variabel tak bebas. Sehingga analisis regresi linier tidak dapat digunakan. Ada beberapa alternatif yang dapat digunakan untuk menganalisis hubungan antara variabel bebas dengan variabel tak bebas yaitu menggunakan analisi regresi logistik, analisis regresi eksponensial, analisis regresi poisson, analisis regresi kuadratik dan lainnya.

Salah satu alternatif yang mendekati dengan kehidupan nyata yaitu analisis regresi logistik. Analisis regresi logistik dapat digunakan apabila variabel bebas dan variabel tak bebas tidak memiliki hubungan linier. Variabel tak bebas pada

analisis regresi logistik dapat berupa data kategori atau data interval. Data kategorik pada analisis regresi logistik terbagi atas data dikotomi atau data ordinal. Data dikotomi yaitu apabila terdiri dari dua kemungkinan yaitu ya dan tidak, sukses atau gagal, 1 atau 0, dan lain sebagainya. Data ordinal yaitu data berskala yang memiliki beberapa tingkatan. Oleh karena itu analisis regresi logistik dapat dikelompokkan menjadi analisis regresi biner dan analisis logistik ordinal. Analisis biner ditujukan untuk data dikotomi sedangkan analisis regresi ordinal ditujukan untuk data yang berskala ordinal.

Model regresi logistik dapat ditransformasikan menjadi model khusus yaitu model regresi linier tergeneralisasi. Model ini mempunyai tiga komponen yaitu *a random component, a systematic component* dan *link component. Random component* ditujukan pada variabel tak bebas. Variabel tak bebas pada regresi linier merupakan variabel kontinu dan diasumsikan berdistribusi normal sedangkan pada regresi logistik berupa data ordinal atau dikotomi dan diasumsikan berdistribusi binomial. *The systematic component* ditujukan untuk variabel tak bebas dan komponen yang membangun variabel tersebut.

Penelitian seringkali terkonsentrasi tentang bagaimana mengetahui hubungan antara individu dengan lingkungan ataupun kelompok. Individu mempunyai korelasi dengan lingkungan ataupun kelompoknya. Begitu pula sebaliknya, lingkungan ataupun kelompok mempunyai hubungan dengan individu yang berada di dalamnya. Secara umum, individu dan lingkungan ataupun kelompok merupakan sistem hirarki, dimana individu berada pada level satu sedangkan

Misalnya pada ilmu kesehatan, dilakukan penelitian mengenai faktor-faktor yang mempengaruhi kesembuhan pasien setelah menjalani operasi. Pada penelitian tersebut, pasien-pasien (sebagai unit-unit yang diteliti) dipilih secara acak dari tiap rumah sakit yang dipilih secara acak. Variabel-variabel pasien berada pada level satu sedangkan variabel-variabel rumah sakit berada pada level dua. Contoh pada bidang pendidikan, penelitian dilakukan untuk mengetahui faktor-faktor yang mempengaruhi nilai UAN murid Sekolah Dasar di Yogyakarta. Sekolah dipilih secara acak dari masing-masing sekolah di Yogyakarta, kemudian dari masing-masing sekolah dipilih siswa-siswa secara acak. Variabel-variabel murid berada pada level satu sedangkan variabel-variabel sekolah berada pada level dua.

Pada data hirarki, sebenarnya terdapat efek dari setiap kelompok terhadap unit-unit yang terdapat dibawahnya. Oleh karena itu analisis regresi biasa kurang tepat jika digunakan pada data berhirarki. Dalam kasus ini analisis regresi multilevel dapat digunakan untuk mengetahui efek dari setiap kelompok terhadap unit-unit yang terdapat dibawahnya.

Salah satu kejadian sederhana dari analisis regresi multilevel yaitu analisis regresi dua level. Sebagai contoh, penelitian tentang pengaruh variabel-variabel siswa dan variabel-variabel sekolah. Variabel-variabel siswa misalnya tingkat kecerdasan siswa, jenis kelamin dan suku. Variabel-variabel sekolah misalnya status sekolah, golongan sekolah dan lainnya. Kelompok sekolah terdiri dari siswa-siswa. Hal ini menunjukkan variabel-variabel siswa tersarang pada

variabel-variabel sekolah. Variabel siswa sebagai level satu sedangkan variabel sekolah sebagai level dua. Pada kasus ini terdapat dua tingkatan level yang berbeda yaitu level satu dan level dua. Berdasarkan hal tersebut analisis regresi yang digunakan dalam kasus ini yaitu analisis regresi dua level.

Salah satu kejadian khusus dari analisis regresi dua level yaitu analisis regresi logistik dua level. Analisis regresi logistik dua level digunakan apabila data berstruktur hirarki dua level dan variabel tak bebas berupa data kategorik. Jika data kategorik berupa data ordinal maka regresi logistik yang digunakan yaitu regresi logistik ordinal dua level.

Terdapat beberapa penelitian tentang pengaruh SES, kepemilikan laboratorium, kepemilikan buku terhadap prestasi. Penelitian yang dilakukan oleh Sheldon Rothman, yang meneliti tentang pengaruh perubahan SES terhadap prestasi siswa. Penelitan lainnya dilakukan oleh Jennifer Barry tahun 2005 yang menjelaskan tentang pengaruh Kondisi SES terhadap prestasi akademik. Penelitian juga dilakukan oleh Mehmet A. Ozturk tentang pengaruh status sosial ekonomi, Ras, Jenis Kelamin, terhadap Prestasi Siswa.

Dalam skripsi ini dikaji analisis regresi logistik ordinal dua level dan aplikasinya. Aplikasi dari analisis regresi logistik ordinal dua level membahas tentang pengaruh variabel-variabel Sekolah Dasar dan variabel-variabel tingkat kecamatan terhadap nilai rata-rata Ilmu Pengetahuan Alam tiap sekolah. Populasi dalam kajian ini adalah sekolah-sekolah yang ada di Kota Yogyakarta pada tahun ajaran 2010/2011. Variabel-variabel sekolah berada pada level yang paling rendah sehingga dapat dikategorikan variabel-variabel sekolah berada pada level satu.

Variabel ini terdiri dari kepemilikan laboratorium dan rasio kepemilikan buku pegangan siswa mata pelajaran Ilmu Pengetahuan Alam. Variabel-variabel kecamatan berada pada level dua. Variabel ini terdiri dari variabel Mean SES (*Sosioeconomic Status*). Mean SES (*Sosioeconomic Status*) merupakan rata-rata dari SES tiap sekolah pada satu kecamatan yang sama. SES (*Sosioeconomic Status*) dipengaruhi oleh beberapa faktor yaitu kondisi ruang kelas, kepemilikan fasilita-fasilitas sekolah, pendapatan sekolah dan lain sebagainya. Variabel-variabel Sekolah Dasar tersarang pada variabel kecamatan.

Berdasarkan pemaparan di atas perlu adanya analisis tentang pengaruh variabel-variabel Sekolah Dasar dan variabel-variabel tingkat kecamatan terhadap nilai rata-rata Ilmu Pengetahuan Alam tiap sekolah. Analisis yang dapat digunakan yaitu analisis regresi logistik ordinal dua level karena variabel tak bebas berupa data ordinal.

Analisis regresi logistik ordinal dua level mempunyai beberapa metode yang dapat digunakan untuk mengestimasi parameter. Metode-metode tersebut yaitu *Maximum Likelihood Estimation* (MLE), *Restricted Maximum Likelihood* (REML), *Marginal Quasi Likelihood* (MQE), dan *Peralized Quasi Likelihood* (PQL). Dalam skripsi ini dikaji analisis regresi logistik ordinal 2-level menggunakan metode Maksimum Likelihood. Metode Maksimum Likelihood digunakan karena merupakan metode yang paling sederhana.

B. Rumusan Masalah

1. Bagaimana analisis regresi logistik dua level pada data ordinal menggunakan metode Maksimum Likelihood?

2. Bagaimana pengaruh faktor-faktor sekolah (level satu) dan faktor-faktor kecamatan (level dua) terhadap nilai rata-rata ujian IPA tiap sekolah di Kota Yogyakarta pada tahun ajaran 2010/2011?

C. Tujuan Penelitian

1. Mendeskripsikan analisis regresi logistik 2-level pada data ordinal menggunakan metode Maksimum Likelihood.
2. Mengetahui pengaruh pengaruh faktor-faktor sekolah (level satu) dan faktor-faktor kecamatan (level dua) terhadap nilai rata-rata ujian IPA tiap Sekolah Dasar di Kota Yogyakarta pada tahun ajaran 2010/2011.

D. Manfaat Penelitian

1. Bagi Penulis

Manfaat yang diharapkan dalam penelitian ini untuk memperdalam pengetahuan penulis tentang analisis regresi logistik ordinal dua level dan aplikasinya.

2. Bagi Pembaca

Manfaat yang diharapkan dalam penelitian ini untuk mempermudah pembaca memahami analisis regresi logistik ordinal dua level dan aplikasinya

3. Bagi Perpustakaan Universitas Negeri Yogyakarta

Menambah koleksi buku perpustakaan Universitas Negeri Yogyakarta bidang statistika.

Pengukuran merupakan aturan-aturan pemberian angka untuk berbagai objek sedemikian rupa sehingga angka ini mewakili kualitas atribut. Terdapat empat jenis skala yang dapat digunakan untuk mengukur atribut, yaitu: skala nominal, skala ordinal, skala interval, dan skala ratio (John Hendri, 2009:1).

1. Skala nominal

   Merupakan salah satu jenis pengukuran dimana angka dikenakan untuk objek atau kelas objek untuk tujuan identifikasi. Nomor absen siswa, nomor punggung pemain sepakbola, loker, dan lain-lain adalah suatu skala nominal. Demikian juga, jika dalam suatu penelitian tertentu pria diberikan kode 1 dan wanita mendapat kode 2, untuk mengetahui jenis kelamin seseorang adalah melihat apakah orang ini berkode 1 atau 2. Angka-angka tersebut tidak mewakili hal lain kecuali jenis kelamin seseorang. Wanita, meskipun mendapat angka yang lebih tinggi, tidak berarti "lebih baik" dibanding pria, atau "lebih banyak" dari pria (John Hendri, 2009:1).

2. Skala ordinal

   Skala ordinal adalah skala yang bertujuan untuk membedakan antara kategori-kategori dalam satu variabel dengan asumsi bahwa ada urutan atau tingkatan skala. Angka-angka ordinal lebih menunjukkan urutan peringkat. Semakin besar angkanya maka semakin besar peringkat atau tingkatannya. Angka-angka tersebut tidak menunjukkan kuantitas absolut, tidak pula

memberikan petunjuk bahwa jarak antara setiap dua angka itu sama. Contoh, angka 1 untuk mewakili mahasiswa tahun pertama, 2 untuk tahun kedua, 3 untuk tahun ketiga, dan 4 untuk mahasiswa senior (S.S Steven, 1976).

3. Skala interval

   Skala interval adalah skala suatu variabel yang selain membedakan dan mempunyai tingkatan, juga diasumsikan mempunyai jarak yang pasti antara satu kategori dengan kategori yang lain dalam satu variabel. Contoh, nilai prestasi yang telah ditransfer dalam bentuk huruf A, B, C, D, E. Selanjutnya diberi bobot masing-masing 4, 3, 2, 1, dan 0. Sehingga interval A dan B sama dengan interval D dan E, juga interval A dan C sama dengan antara C dan E (S.S Steven, 1976).

4. Skala ratio

   Skala rasio adalah skala suatu variabel yang selain membedakan dan mempunyai tingkatan serta jarak antara suatu nilai dengan nilai yang lainnya, juga diasumsikan bahwa setiap nilai variabel diukur dari suatu keadaan atau titik yang sama (mempunyai titik nol mutlak). Angka-angka pada skala menunjukkan besaran sesungguhnya dari sifat yang diukur. Contoh, berat benda A adalah 30 kg dan berat benda B adalah 60 kg (S.S Steven, 1976).

B. Matriks

1. Definisi Matriks

Sebuah matriks adalah susunan segi empat siku-siku dari bilangan. Bilangan-bilangan dari susunan tersebut dinamakan entri dalam matriks (Howard Anton,

1987). Ukuran matriks dijelaskan dengan menyatakan banyaknya baris (garis horizontal) dan banyaknya kolom (garis vertikal) yang terdapat pada matriks tersebut. Jika ***A*** adalah sebuah matriks, maka kita menggunakan a_{ij} untuk menyatakan entri yang terdapat didalam baris i dan kolom j dari ***A***. Matriks dengan banyaknya baris m dan banyaknya kolom n umumnya dituliskan sebagai berikut:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} atau \left[a_{ij}\right]_{m\,x\,n} \quad (2.1)$$

2. Matriks Persegi

Matriks ***A*** disebut matriks persegi apabila mempunyai jumlah baris dan jumlah kolom yang sama ($m = n$) (Howard Anton, 1987).

3. Transpos Matriks

Jika **A** adalah sebarang matriks m x n, maka transpos **A** dinyatakan oleh $\boldsymbol{A'}$ dan didefinisikan dengan matriks n x m yang kolom pertamanya adalah baris pertama dari **A**, kolom keduanya adalah baris kedua dari **A**, demikian juga dengan kolom ketiga dan seterusnya (Howard Anton, 1987). Sifat-sifat dari tranpos matriks sebagai berikut :

a. $(\boldsymbol{A'})' = \boldsymbol{A}$ (2.2)

b. $(\boldsymbol{A} \pm \boldsymbol{B})' = \boldsymbol{A'} \pm \boldsymbol{B'}$ (2.3)

c. $(k\boldsymbol{A})' = k\boldsymbol{A'}$ dimana k adalah sebarang skalar (2.4)

d. $(\boldsymbol{AB})' = \boldsymbol{B'A'}$ (2.5)

4. Determinan Matriks

Menurut Johnson R.A dan Wichren D.W (2007: 93) determinan matriks persegi $A_{nxn} = \{a_{ij}\}$ dinotasikan |**A**|, dengan skalar

a. $|\mathbf{A}| = a_{11}$, jika $n = 1$ (2.6)

b. $|\boldsymbol{A}| = \sum_{j=1}^{n} a_{1j} \, |\boldsymbol{A}_{1j}|(-1)^{1+j}$, jika $n \geq 1$ (2.7)

dengan $\boldsymbol{A}_{1j}$ matriks (n-1) x (n-1) dengan menghapus baris pertama dan kolom ke-j.

5. Matriks kofaktor

Diberikan matriks A_{nxn} , kofaktor dari matriks A dinotasikan C_{ij} didefinisikan sebagai $C_{ij} = (-1)^{i+j}|\boldsymbol{A}_{ij}|$(Howard Anton, 1987).

6. Adjoint Matriks

Matriks adjoint dari A dinotasikan Adj (A), didefinisikan sebagai tranpos dari matriks kofaktor dari A (Howard Anton, 1987).

7. Invers Matriks

Jika A_{nxn} adalah matriks persegi, dan jika kita dapat mencari matriks $\mathbf{B}_{nxn}$ sehingga **AB** = **BA** = **I**, maka A_{nxn} dikatakan dapat dibalik (*invertible*) dan **B** dinamakan invers dari ***A*** dinotasikan $\mathbf{A}^{-1}$ (Howard Anton, 1987).

Sifat-sifat invers matriks adalah sebagai berikut:

a. $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$ (2.8)

b. $(\boldsymbol{A'})^{-1} = (\boldsymbol{A}^{-1})'$ (2.9)

Persamaan regresi logistik mempunyai fungsi untuk menentukan pengaruh beberapa variabel bebas yang secara bersamaan memprediksikan variabel tak bebas. Regresi logistik dapat digunakan untuk memodelkan hubungan antara dua kategori atau lebih dari variabel tak bebas dan dua atau lebih variabel bebas. Estimasi model regresi logistik untuk masing-masing variabel bebas memberikan hasil estimasi variabel untuk setiap variabel bebas terhadap variabel tak bebas dengan mengikut sertakan variabel bebas lainnya pada permodelan tersebut (Cath Roberts, Rachel Dolman, Anne Kingdon, 2007).

Perbedaan antara analisis regresi berganda dengan analisis logistik yaitu:

1. Analisis regresi berganda terdapat F yaitu uji pengaruh bersama-sama variabel independen terhadap variabel dependen dan uji t yaitu uji untuk mengetahui pengaruh tiap variabel/masing-masing variabel bebas terhadap variabel tak bebas. Sedangkan analisis analisis regresi logistik hanya terdapat uji t.
2. Pada analisis logistik variabel tak bebasnya berupa data kategorik (ordinal dan nominal) maupun data dikotomi. Data dikotomi terdiri dari dua pilihan misalnya ya atau tidak, berkualitas atau tidak berkualitas, lulus atau gagal, sering atau jarang.

Asumsi-asumsi dari regresi logistik yaitu asumsi multikolinieritas. Pada model regresi logistik antara peubah penjelas harus bebas multikolinieritas. Regresi logistik akan baik jika tidak terjadi multikolinearitas. Asumsi ini dapat dilihat

melalui nilai VIF. Jika nilai VIF<10 maka tidak terjadi multikolineritas. (Adriyan, 2010)

D. Metode Maksimum Likelihood

Definisi :

Diberikan L adalah fungsi densitas peluang bersama dari (X_1, X_2 ,X_3 , ..., X_n), yang tergantung pada parameter θ, $\theta \in \Omega$ yaitu L(θ) = f (x_1, x_2, ... ,x_n ; θ). Nilai dari θ yang menghasilkan nilai maksimum untuk L(θ) disebut estimasi maksimum likelihood untuk θ, dan dinyatakan dengan simbol $\hat{\theta}$. Jadi $f\left(x_1, x_2, \dots, x_n; \hat{\theta}\right) = \max_{\theta \in \Omega} f(x_1, x_2, \dots, x_n; \theta)$ (Bain & Engelhard, 1991:294).

Apabila Ω adalah interval terbuka, jika L(θ) diferensiabel dan mencapai nilai maksimum di suatu nilai dalam Ω, maka MLE merupakan solusi dari persamaan $\frac{d}{d\theta} L(\theta) = 0$ (Bain & Engelhard, 1991:294).

Definisi dari maksimum likelihood dapat diterapkan pada kasus dengan lebih dari satu parameter yang tidak diketahui jika $\boldsymbol{\theta} = \theta_1, \theta_2, \dots, \theta_{\mathrm{k}}$ adalah vektor parameter yang tidak diketahui maka estimasi parameter maksimum likelihood dapat diperoleh dengan mendeferensialkan fungsi log likelihood:

$$\frac{d}{d\theta_k} lnL(\theta_1, \theta_2, \dots, \theta_{\mathrm{k}}) = 0 \text{ , untuk } j = 1,2, \dots, k \qquad (2.11)$$

Solusi dari estimasi maksimum likelihood dinotasikan dengan $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{\mathrm{k}}$.

E. Regresi Logistik Ordinal

Regresi logistik ordinal digunakan untuk menganalisis data dengan variabel tak bebas berskala ordinal. Sebagaimana dalam model linier lainnya, variabel

bebasnya dapat berupa data kontinu atau variabel berskala kategorik (Agresi,1990). Jika diasumsikan terdapat variabel dependen Y yang berskala ordinal dengan J kategori $\boldsymbol{x}' = (x_1, x_2, \dots, x_J)$ adalah vektor variabel bebas, $\beta' = (\beta_1, \beta_2, \dots, \beta_j)$ adalah vektor paremeter intersep (slope) maka peluang dari variabel tak bebas kategori ke-*j* pada *p* variabel bebas yang dinyatakan dalam vektor **X** tertentu dapat dinyatakan dengan $P[Y = j|x] = p_j(x)$ dimana $p_j(x) = F\left(\alpha_j + \boldsymbol{\beta}'\mathbf{x}\right) - F\left(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}\right)$ dan peluang komulatifnya adalah (Hosmer&Lemeshow, 2000:290):

$$P[Y \le j|x] = p_1(x) + p_2(x) + \dots + p_j(x) \qquad (2.12)$$

$$= F\left(\alpha_j + \boldsymbol{\beta}'\mathbf{x}\right) \qquad (2.13)$$

$$= \frac{e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}} \qquad (2.14)$$

Model logit komulatif didefinisikan dengan:

$$L_j(x) = \operatorname{logit}\left(P[Y \le j|x]\right) \qquad (2.15)$$

$$= \ln\left(\frac{P[Y \le j|x]}{1 - P[Y \le j|x]}\right) \qquad (2.16)$$

$$= \ln\left(\frac{\frac{e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}}{1 - \left(\frac{e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}\right)}\right) \qquad (2.17)$$

$$= \ln\left(e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}\right) \qquad (2.18)$$

$$= \alpha_j + \boldsymbol{\beta}'\mathbf{x} \qquad (2.19)$$

dimana $j = 1, \dots, \mathrm{J} - 1$ dan $\alpha_1, \dots, \alpha_{\mathrm{J}-1}$ adalah *threshold model* serta $\boldsymbol{\beta}$ merupakan vektor koefisien regresi.

Metode pendugaan parameter yang dapat digunakan pada regresi logistik ordinal diantaranya adalah dengan metode maksimum likelihood. Metode ini dapat dilakukan jika antara amatan yang satu dengan yang lain diasumsikan saling bebas. Metode *maximum likelihood estimation (MLE)* dipilih karena mempunyai beberapa kelebihan dibandingkan dengan metode lain, diantaranya dapat digunakan untuk model tidak linier seperti regresi logistik, serta hasil penaksirannya *unbiased.*

Peluang kumulatif digunakan untuk menduga parameter, maka likelihood dapat ditulis sebagai perkalian J-1 kategori, sehingga Fungsi kumulatif peluang bersama dari $\boldsymbol{Y} = (\mathrm{Y}_1, \mathrm{Y}_2, \dots, \mathrm{Y}_n)$ adalah sama dengan perkalian *n* fungsi multinomial. Fungsi likelihoodnya dapat dinyatakan sebagai (Hosmer& Lemeshow 2000:8).

$$\mathbf{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^{\mathrm{n}} \left[p_1(x_{\mathrm{i}})^{z0i} p_2(x_{\mathrm{i}})^{z1i} \dots p_{\mathrm{j}}(x_{\mathrm{i}})^{zji}\right] \qquad (2.20)$$

$$\text{Dengan } z_{ij} = \begin{cases} 1 \text{ untuk } y = j \\ 0 \text{ untuk } y \ne j \end{cases}$$

Selanjutnya dicari ln dari fungsi likelihood yaitu sebagai berikut:

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^{\mathrm{n}} \mathrm{z_{0i}} \ln[\mathrm{p}_1(\mathrm{x_i})] + \mathrm{z_{1i}} \ln[\mathrm{p}_2(\mathrm{x_i})] + \dots + \mathrm{z_{Ji}} \ln\left[\mathrm{p_J}(\mathrm{x_i})\right] \qquad (2.21)$$

Dimana $p_j(x) = F\left(\alpha_j + \boldsymbol{\beta}'\mathbf{x}\right) - F\left(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}\right) = \frac{e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_j + \boldsymbol{\beta}'\mathbf{x})}} - \frac{e^{(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x})}}$

Maka fungsi ln likelihoodnya menjadi:

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^{\mathrm{n}} \mathrm{z_{0i}} \ln\left[\frac{e^{(\alpha_1 + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_1 + \boldsymbol{\beta}'\mathbf{x})}}\right] + \mathrm{z_{1i}} \ln\left[\frac{e^{(\alpha_2 + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_2 + \boldsymbol{\beta}'\mathbf{x})}} - \frac{e^{(\alpha_1 + \boldsymbol{\beta}'\mathbf{x})}}{1 + e^{(\alpha_1 + \boldsymbol{\beta}'\mathbf{x})}}\right] + \dots +$$

$$\mathrm{z_{Ji}} \ln\left[\frac{e^{\left(\alpha_J + \boldsymbol{\beta}'\mathbf{x}\right)}}{1 + e^{\left(\alpha_J + \boldsymbol{\beta}'\mathbf{x}\right)}} - \frac{e^{\left(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}\right)}}{1 + e^{\left(\alpha_{J-1} + \boldsymbol{\beta}'\mathbf{x}\right)}}\right] \qquad (2.21)$$

$$= \frac{\left(e^{(\alpha_2+\boldsymbol{\beta}'\mathbf{x})}\right)-\left(e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}\right)}{\left(1+e^{(\alpha_2+\boldsymbol{\beta}'\mathbf{x})}\right)\left(1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}\right)} \quad (2.23)$$

$$= \frac{e^{\boldsymbol{\beta}'\mathbf{x}}e^{(\alpha_2-\alpha_1)}}{\left(1+e^{(\alpha_2+\boldsymbol{\beta}'\mathbf{x})}\right)\left(1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}\right)} \quad (2.24)$$

Maka fungsi log likelihood menjadi:

$$ln\mathrm{L}(\boldsymbol{\alpha},\boldsymbol{\beta}) =$$

$$\sum_{i=1}^{n} z_{0i}\left((\alpha_1+\boldsymbol{\beta}'\mathbf{x}) - \left(\mathbf{1}+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}\right)\right) + z_{1i}\left(\boldsymbol{\beta}'\mathbf{x} + \ln(e^{\alpha_2}-e^{\alpha_1}) - ln\left(1+e^{(\alpha_2+\boldsymbol{\beta}'\mathbf{x})}\right) - \ln\left(1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}\right)\right) + \dots +$$

$$-z_{Ji}\ln\left(e^{(\alpha_{j-1}+\boldsymbol{\beta}'\mathbf{x})}\right) \quad (2.25)$$

Dari fungsi log likelihood ini didapatkan turunan log $\mathbf{L}(\boldsymbol{\alpha},\boldsymbol{\beta})$ terhadap $\boldsymbol{\alpha}$ dan $\boldsymbol{\beta}$ sebagai berikut:

$$\frac{\partial \ln L}{\partial \alpha_1} = \sum_{i=1}^{n}\left\{z_{0i}\left(1-\frac{e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}\right) + z_{1i}\left(-\frac{e^{\alpha_1}}{e^{\alpha_2}-e^{\alpha_1}} - \frac{e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}\right)\right\} \quad (2.26)$$

$$\frac{\partial \ln L}{\partial \alpha_1} = \sum_{i=1}^{n}\left\{z_{0i}\left(1-\frac{e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}\right) - z_{1i}\left(\frac{e^{\alpha_1}}{e^{\alpha_2}-e^{\alpha_1}}\right) - z_{1i}\left(\frac{e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}\right)\right\} \quad (2.27)$$

$$\vdots$$

$$\frac{\partial \ln L}{\partial \alpha_j} = \sum_{i=1}^{n}\left\{z_{Ji}\left(\frac{e^{\alpha_j}}{e^{\alpha_j}-e^{\alpha_{j-1}}} - \frac{e^{(\alpha_J+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_J+\boldsymbol{\beta}'\mathbf{x})}}\right) + z_{(J+1)i}\left(\frac{e^{\alpha_j}}{e^{\alpha_{(J+1)i}}-e^{\alpha_j}} - \frac{e^{(\alpha_J+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_J+\boldsymbol{\beta}'\mathbf{x})}}\right)\right\} \quad (2.28)$$

$$\frac{\partial \ln L}{\partial \beta_p} = \sum_{i=1}^{n}\left\{z_{0i}\left(\boldsymbol{x}-\frac{xe^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}\right) + z_{1i}\left(\boldsymbol{x}-\frac{xe^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_1+\boldsymbol{\beta}'\mathbf{x})}} - \frac{xe^{(\alpha_2+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(2+\boldsymbol{\beta}'\mathbf{x})}}\right) + \dots + z_{Ji}\frac{xe^{(\alpha_{J-1}+\boldsymbol{\beta}'\mathbf{x})}}{1+e^{(\alpha_{J-1}+\boldsymbol{\beta}'\mathbf{x})}}\right\} \quad (2.29)$$

Turunan pertama dari log $\mathbf{L}(\boldsymbol{\alpha},\boldsymbol{\beta})$ terhadap $\boldsymbol{\alpha}$ dan $\boldsymbol{\beta}$ adalah non linier parameter. Maka, untuk mendapatkan taksiran parameternya digunakan metode Newton-Raphson dengan iterasi Weigth Least Square.

Metode Newton-Raphson dengan iterasi Weigth Least Square digunakan untuk mendapatkan taksiran parameter yaitu sebagai berikut:

$$\beta^{(t+1)} = \beta^{(t)} - \left(\mathbf{H}(\boldsymbol{\beta}^{(t)})\right)r^{(t)}, t = 0,1,2,\dots \quad (2.30)$$

Dimana $r^{(t)} = \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_k}$ dengan $\mathrm{k} = 1,2,\dots,p$

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1\mathrm{p}} \\ h_{21} & h_{22} & \dots & h_{2\mathrm{p}} \\ \vdots & \vdots & & \vdots \\ h_{\mathrm{p}1} & h_{\mathrm{p}2} & \dots & h_{\mathrm{pp}} \end{bmatrix}, h_{ab} = \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_a \partial \beta_b}$$

Konvergensi $\beta^{(t)}$ menuju $\hat{\beta}$ untuk mudah dilakukan menggunakan metode Newton-Rapson.

Pada analisis regresi ordinal terdapat lima pilihan transformasi seperti tercantum pada tabel 1. Penggunaannya berdasarkan sebaran data yang dianalisis. Logit, probit dan chauchit digunakan untuk sebagian besar sebaran data, *complementary log-log* digunakan untuk data yang mempunyai kecenderungan bernilai tinggi, *negative log-log* digunakan untuk data yang mempunyai nilai yang ekstrim. Perbedaan transformasi logit dan probit terletak pada fungsi kumulatifnya. Model logit menggunakan fungsi logistik kumulatif (*cumulative logistic function*) sedangkan model probit menggunakan fungsi normal kumulatif (normal CDF) disebut juga dengan model normit

Analisis regresi ordinal yang telah dijelaskan sebelumnya adalah analisis regresi ordinal dengan fungsi transformasi atau sering disebut regresi logistik ordinal (Norusis 2010).

Tabel 1. Tabel Transformasi

Fungsi Hubung	Bentuk Fungsi
Logit	$\mathrm{Log}\left(\frac{p}{1-p}\right)$
Complementary log-log	$\mathrm{Log}(-\mathrm{Log}(1-\mathrm{p}))$
Negative log-log	$-\mathrm{Log}(-\mathrm{Log}(\mathrm{p}))$
Probit	$\Phi^{-1}(\mathrm{p})$
Chauchit	$\mathrm{Tan}(\mathrm{phi}(\mathrm{p}-0.5))$

F. Regresi Multilevel

Pemodelan multilevel merupakan suatu teknik statistika yang digunakan untuk menganalisis data dengan struktur hirarki. Pada model multilevel variabel dependen diukur pada level kesatu, sedangkan variabel independen dapat didefinisikan pada setiap level. Bentuk sederhana dari model regresi multilevel adalah regresi dua level. Secara matematis regresi dua level dengan satu peubah bebas pada level kesatu dapat ditulis sebagai berikut (Hox, 2002):

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij} \quad (2.31)$$

$e_{ij} \sim N(0,\sigma_e^2)$; dengan j = 1,2,...,m dan i =1,2,...,n_j

Dengan

j : indeks kelompok pada level dua

i : indeks individu pada level satu

Y_{ij} : variabel respon dari individu ke- *i* dalam kelompok ke-*j*

β_{oj} : intersep pada kelompok ke-*j*

β_{1j} : koefisien regresi (*slope*) pada kelompok ke-*j*

X_{ij} : variabel prediktor dari individu ke-*i* dalam kelompok ke-*j*

ε_{ij} : error individu ke-*i* dalam kelompok ke-*j* dan berdistribusi $N(0,\sigma_\varepsilon^2)$

Pada persamaan (2.31) koefisien regresi β_0 dan β_1 memiliki indeks *j* untuk kelompok pada level dua, yang mengidentifikasikan bahwa koefisien regresi pada level dua dapat memiliki dua nilai yang berbeda. Jika terdapat satu peubah penjelas pada level dua, maka keragaman koefisien regresi tersebut dapat dimodelkan melalui persamaan (2.32) dan (2.33) yaitu:

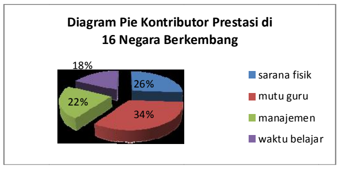
$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j,} \quad (2.32)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (2.33)$$

Dengan Z_j adalah peubah penjelas pada level kedua, u_{0j} dan u_{1j} adalah galat pada level kedua. Diasumsikan $u_{0j} \sim N(0,\sigma_{u0}^2)$ dan $u_{1j} \sim N(0,\sigma_{u1}^2)$ serta u_{0j}, u_{1j} dan ε_{ij} saling bebas (Hox 2002)

Terdapat dua model logistik multilevel yang dapat digunakan untuk data kategori. Model tersebut yaitu model regresi logistik multilevel biner dan model regresi logistik ordinal. Model regresi logistik multilevel biner untuk respon biner (berupa data dikotomi) dan model regresi logistik ordinal multilevel untuk respon ordinal.

belajar siswa. Faktor peranan guru merupakan faktor yang paling signifikan ditengah-tengah keterbatasan sarana dan prasarana. Sebuah studi terhadap peranan pendidikan yang dilakukan di 16 negara berkembang menunjukkan mutu pendidikan bergantung pada faktor-faktor yang berkontribusi terhadap prestasi belajar siswa (Dedi S, Widoyoko, 2008:1).



Gambar 1. Diagram Pie Kontributor Prestasi di 16 Negara Berkembang

Gambar 1 menunjukan bahwa kontributor yang terbesar dalam prestasi yaitu faktor mutu guru atau pengajar sebesar 34%. Faktor-faktor yang lainnya memberikan kontribusi yang lebih rendah terhadap prestasi. Faktor sarana fisik memberikan kontribusi sebesar 26% terhadap prestasi. Faktor manajemen dan waktu belajar hanya memberikan kontribusi masing-masing 22% dan 18% terhadap prestasi.

Menurut Windham (1990) masukan (input) pendidikan adalah sumber-sumber yang digunakan dalam aktifitas produksi pendidikan. Windham membedakan input yang disediakan oleh sekolah dengan input yang ditentukan dari luar

sekolah. Input sekolah mencakup input fisik yaitu kondisi bangunan, kondisi bahan pembelajaran, ketersedian fasilitas dan input manusia yakni mencakup ketersedian dan kualitas guru dan tenaga administrasi.

Input dari bukan sekolah sangat beragam, namun pada umumnya terdiri dari karakteristik murid, karakteristik keluarga dan karakteristik lingkungan masyarakat. karakteristik murid biasanya diukur dengan jenis kelamin, ras, dan kemampuan mula-mula murid. Karakteristik keluarga cenderung diwakili oleh tingkat status sosial ekonomi orang tua yang diukur dengan penghasilan keluarga, tingkat pendidikan ayah dan ibu, jumlah anggota keluarga, jumlah buku di rumah, kepemilikan fasilitas dan lain-lain.

Beberapa hal yang berkenaan dengan apakah pendidikan bermutu atau tidak bermutu, tentu banyak faktor yang yang terlibat, mulai input pendidikan yang dikelompokkan menjadi siswa, Instrumental input (guru, kurikulum, sarana pendidikan/media pembelajaran, buku teks dan buku pelajaran, dsb). Hal yang tidak dapat diabaikan adalah sarana/prasarana, buku teks, buku pelajaran, dan media pembelajaran, karena hal ini akan memberikan sumbangan yang sangat signifikan terhadap peningkatan mutu pendidikan.

Dalam skripsi ini akan membahas tentang pengaruh beberapa fasilitas yang dimiliki sekolah yaitu pengaruh kepemilikan laboratorium dan rasio kepemilikan buku IPA terhadap nilai rata-rata ujian IPA tiap Sekolah dasar di Kota Yogyakarta.

Keadaan sosial ekonomi setiap orang itu berbeda-beda dan bertingkat, ada yang keadaan sosial ekonominya tinggi, sedang, dan rendah. Sosial ekonomi menurut Abdulsyani (1994) adalah kedudukan atau posisi sesorang dalam

kelompok manusia yang ditentukan oleh jenis aktivitas ekonomi, pendapatan, tingkat pendidikan, jenis rumah tinggal, dan jabatan dalam organisasi. Menurut Soerjono Soekanto (2001) sosial ekonomi adalah posisi seseorang dalam masyarakat berkaitan dengan orang lain dalam arti lingkungan pergaulan, prestasinya, dan hak-hak serta kewajibannya dalam hubungannya dengan sumber daya. Menurut Smith (2011) SES (*Socioeconomic Status*) adalah ukuran setiap individu terhadap dalam suatu lingkungan (kelompok) berdasarkan sumber daya yang dimiliki dan kedudukan di kehidupan sosial . Lebih dari 70 tahun yang lalu sudah ditemukan hubungan antara SES dengan intelektual/ kompetensi akademik. Menurut Mc Call (1981) membuktikan bahwa terdapat hubungan yang erat antara SES dengan intelektual/ kompetensi akademik.

Pengukuran SES dapat dilakukan dengan member skor/nilai dari beberapa faktor. Skor gabungan dari SES dapat diperoleh dengan menstandardisasi pendapatan dan pendidikan dan standar penilaian yang lainnya (Oakes, 2006). Ukuran univariat yaitu pendapatan, kekayaan, pencapaian akademik/pendidikan, kemiskinan, ukuran tingkatan level.

Berdasarkan beberapa pendapat di atas, dapat disimpulkan pengertian keadaan sosial ekonomi dalam penelitian ini adalah ukuran Sekolah Dasar terhadap lingkungan di tiap kecamatan berkaitan dengan tingkat pendidikan, akreditasi sekolah dan tingkat pendapatan pemilikan kekayaan atau fasilitas.

Daftar Pustaka

Anton, Howard. (1987). *Aljabar Linier Elementer*. (Alih bahasa: Pantur Silaban). Jakarta: Erlangga.

Bain, Lee J & Engelhardt. (1991). *Introduction To Probability and Mathematical Statistics*. Australia: Duxbury.

Chaimongkol, Saengla. (2005). Modeling Differential Item Functioning (DIF) Using Multilevel Logistic Regression Models: A Bayesian Perspective. University of Florida. *Electronic Theses, Treatises and Dissertations (ETDs)*. Paper 3939.

Croux C. & Haesbroeck G. (2011). Robust estimation for ordinal regression. *Journal of Statistical Planning and Inference*. Hlm. 1-34.

Darnah. (2011). Regresi Logistik Ordinal untuk Menganalisis Faktor-Faktor yang Mempengaruhi Perilaku Sexual Remaja. *Jurnal Eksponensial* Vol. 2 No 1. Hlm 48-52.

Gibbons , Robert D & Hedeker, Donald. (1997). Random Effects Probit and Logistic Regression Models for Three-Level Data. *Journal of Biometrics Vol 53 No 4*. Hlm. 1527-1537.

Hedeker, Donald. (2007). *Handbook of Multilevel Analysis*. New York: Springer.

Hendri, Jhon. Skala Pengukuran dan Teknik Penskalaan. *Riset Pemasaran Universitas Gunadarma*. Jakarta: Universitas Gunadarma.

Herman. (2011). Hubungan Kompetensi dengan Kinerja Guru Ekonomi SMA. *Jurnal Ekonomi Bisnis Th 16 No 1*. Hlm. 17-24.

Hosmer, David W. & Lemeshow, Stanley. (2000). *Applied Logistic Regression*. New York: Wiley.

Hox, Joop J. (2010). *Multilevel Analysis Techniques and Applications*. 2nd. ed. New York: Routledge.

Keating, Kim.A. & Cherry, Steve. (2004). Use and Interpretation of Logistic Regression in Habitat-Selection Studies. *Jurnal of Wildlife Management*. Hlm. 774-789.

Kleinbaum, David G. (2010). *Logistic Regression*. 3rd . ed. USA: Springer.

Kinerja Guru di SMK Negeri Ketapang. *Journal of Research and Educational Research Evaluation*. Hlm. 1-7.

Skrondal, Anders. (2002). Multilevel Logistic Regression for Polytomous Data and Rankings. *Proceeding*. Efron Seminar. Norway: Norwegian Institute of Public Health.

Suprihatmi. (2007). Pengaruh Kompetensi Manajerial Kepala Sekolah terhadap Kinerja Guru SMP Negeri. *Jurnal Manajemen dan Sumber Daya Manusia* Vol. 2 No. 1. Hlm 78-84.

Tantular, Bertho. (2008). Penentuan Ukurean Sampel dalam Analisis Data Multilevel. *Prosiding*, Seminar Nasional Sains dan Teknologi II. Lampung: Universitas Lampung.

Widiarta, Ida B.P. & Wardana, I G. N. (2011). Analisis Pemilihan Moda dengan Regresi Logistik pada Rencana Koridor Trayek Trans Sarbagita. *Jurnal Ilmiah Teknik Sipil* (Vol. 15, No. 2, Juli 2011). Hlm 131-142.

Yang, Min. *Review of HLM 5.04 for Windows*. London: Institute of Education, University of London.