

**APLIKASI *QUASICOMPLETE SEPARATION* DENGAN METODE
MAXIMUM LIKELIHOOD PADA REGRESI LOGISTIK**

SKRIPSI

**OLEH
EDY SUTRISNO
NIM 13610075**



**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

**APLIKASI *QUASICOMPLETE SEPARATION* DENGAN METODE
MAXIMUM LIKELIHOOD PADA REGRESI LOGISTIK**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Edy Sutrisno
NIM. 13610075**

**JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020**

**APLIKASI *QUASICOMplete SEPARATION* DENGAN METODE
MAXIMUM LIKELIHOOD PADA REGRESI LOGISTIK**

SKRIPSI

Oleh
Edy Sutrisno
NIM. 13610075

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 11 Mei 2020

Pembimbing I,

Dr. Sri Harini, M.Si
NIP. 197310104 200112 2 002

Pembimbing II,

Juhari, S.Pd., M.Si
NIP. 19840201 60801 105

Mengetahui,
Ketua Jurusan Matematika



Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

**APLIKASI *QUASICOMplete SEPARATION* DENGAN METODE
MAXIMUM LIKELIHOOD PADA REGRESI LOGISTIK**

SKRIPSI


Oleh
Edy Sutrisno
NIM. 13610075

Telah Dipertahankan di Depan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 11 Mei 2020

Penguji Utama : Abdul Aziz, M.Si



Ketua Penguji : Ria Dhea Layla Nur Karisma, M.Si



Sekretaris Penguji : Dr. Sri Harini, M.Si



Anggota Penguji : Juhari, S.Pd., M.Si



Mengetahui,
Ketua Jurusan Matematika



Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Edy Sutrisno

NIM : 13610075

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Aplikasi *Quasicomplete Separation* dengan Metode *Maximum Likelihood Estimation (MLE)* pada Regresi Logistik.

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 26 Mei 2020
Yang Membuat Pernyataan,



Edy Sutrisno
NIM. 13610075

MOTO

“Sesungguhnya sesudah kesulitan itu ada kemudahan ,maka apabila kamu telah selesai (dari suatu urusan) kerjakanlah dengan sungguh sungguh (urusan) yang lain.Dan hanya kepada Tuhanlah hendaknya kamu berharap”.

(Q.S Al-Insyirah:6-8)

“Tidak ada yang dapat merubah ketentuan Allah kecuali do’a dan tidak ada yang dapat menambah umat kecuali kebaikan”

(HR.At-Tirmidzi)



PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Ayah Miran dan Ibu Marni

yang selalu sabar tanpa lelah memberikan dukungan, doa, serta limpahan kasih sayangnya.

Dan juga terimakasih kepada kakak Purnomo, Murtinah yang selalu mendukung penulis. Seluruh Keluarga besar ayah ibu.

Teman-teman seperjuangan matematika 2013 yang selalu memberi semangat , Nur fuad, Mubin dan Imam yang selalu mendukung dan mendo'akan serta untuk saya sendiri.



KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah Swt. yang telah melimpahkan nikmat, rahmat serta hidayah-Nya sehingga penulis mampu menyelesaikan penyusunan skripsi ini sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Penulis banyak mendapatkan bimbingan serta arahan dari berbagai pihak selama proses penyusunan skripsi ini. Untuk itu ucapan terima kasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya penulis sampaikan terutama kepada:

1. Prof. Dr. H. Abd. Haris, M.Ag, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Usman Pagalay, M.Si, selaku ketua Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang
4. Dr. Sri Harini, M.Si, selaku dosen pembimbing I yang telah banyak memberikan arahan dan berbagi ilmunya kepada penulis.
5. Juhari, S.Pd., M.Si, selaku pembimbing II yang senantiasa membimbing dengan sabar hingga terselesaikannya penelitian ini.

6. Abdul Aziz, M.Si. selaku penguji utama yang telah banyak memberikan saran dan masukan kepada penulis.
7. Abdul Aziz, M.Si. yang telah sabar dalam mendidik dan memberikan ilmunya kepada penulis hingga terselesaikannya penelitian ini.

Akhirnya penulis berharap semoga skripsi ini bermanfaat bagi penulis dan pembaca.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 1 Mei 2020

Penulis



DAFTAR ISI

HALAMAN PENGAJUAN	
HALAMAN PERSETUJUAN	
HALAMAN PENGESAHAN	
HALAMAN PERNYATAAN KEASLIAN TULISAN	
HALAMAN PERSEMBAHAN	
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR SIMBOL	xi
DAFTAR TABEL	xii
ABSTRAK	xiii
ABSTRACT	xiv
لخص	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
1.6 Sistematika Penulisan	5
BAB II KAJIAN PUSTAKA	6
2.1 Regresi Logistik.....	6
2.2 Model Regresi Logistik Biner.....	8
2.3 Distribusi Normal.....	10
2.4 Uji Multikolinearitas	12
2.5 Estimasi Parameter.....	12
2.6 Pemisahan dalam Metode Maximum Likelihood Estimation	15
2.7 Pengujian Parameter	18
2.7.1 Pengujian Parameter dengan Uji <i>Likelihood</i> Ratio (Uji Simultan atau Uji G)	18
2.7.2 Pengujian Parameter dengan Uji Wald (Uji Parsial).....	20
2.8 Kajian Dalam Al-Qur'an.....	21
BAB III METODE PENELITIAN.....	22
3.1 Variabel Penelitian.....	22

3.2 Sumber Data.....	23
3.3 Tahap Analisis Data.....	23
BAB IV PEMBAHASAN.....	24
4.1 Deskripsi Data.....	24
4.2 Uji Multikolinieritas.....	24
4.3 Tahap Pendugaan Parameter Menggunakan Metode Likelihood	25
4.4 Tahap Pemisahan (Quasi-Complete Separation)	27
4.5 Tahap Pegujian Parameter	29
4.6 Kajian Nilai-nilai Al-Qur'an dengan Teori Estimasi	32
BAB V PENUTUP.....	35
5.1 Kesimpulan	35
5.2 Saran	35
DAFTAR RUJUKAN.....	36
LAMPIRAN	
RIWAYAT HIDUP	

DAFTAR SIMBOL

Simbol-simbol yang digunakan dalam skripsi ini mempunyai makna yaitu sebagai berikut:

Y : Nilai variabel respon

X : Nilai variabel prediktor

$f(y)$: Fungsi probabilitas distribusi *Bernoulli*

$\pi(x)$: Peluang nilai sukses dari variabel x

β : Nilai koefisien regresi

$\hat{\beta}$: Penduga dari parameter β

L : Fungsi likelihood

i : $1, 2, \dots, n$

J : $1, 2, \dots, n$

DAFTAR TABEL

Tabel 4 1	Nilai VIF Variabel Prediktor data Ibu Hamil di Makassar	25
Tabel 4 2	Hasil Estimasi Parameter Regresi Logistik.....	26
Tabel 4 3	Data Penelitian yang Dipisahkan	28
Tabel 4 4	Hasil Estimasi Regresi Logistik yang Setelah Dipisahkan	29
Tabel 4 5	Hasil Uji Ratio Test.....	30
Tabel 4 6	Hasil Uji Wald pada Regresi Logistik	31
Tabel 4 7	Hasil Estimasi Regresi Logistik Biner	32

ABSTRAK

Sutrisno, Edy. 2020. **Aplikasi *Quasicomplete Separation* Dengan Metode *Maximum Likelihood* Pada Regresi Logistik.** Skripsi, Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Pembimbing: (I) Dr. Sri Harini, M.Si (II) Juhari, S.Pd., M.Si

Kata Kunci: Regresi logistik biner, *Quasicomplete Separation*, Estimasi Parameter *Maximum Likelihood*..

Regresi logistik biner merupakan salah satu metode yang digunakan untuk mengetahui pengaruh suatu peubah predikatornya terhadap peubah respon bersifat kategorik. Estimasi parameter model regresi logistik biner menggunakan metode *Maximum Likelihood* agar menghasilkan data yang linier dan signifikan. Metode *Maximum Likelihood* hasilnya tidak dapat digunakan lagi apabila diperoleh penduga yang tidak konvergen atau tidak signifikan. Maka harus menggunakan *Quasi-Complete Separation* supaya bisa hasilnya signifikan atau hasilnya lebih kecil. Tujuan penelitian ini untuk mengaplikasikan kasus *Quasi-Complete Separation* dengan metode *Maximum Likelihood Estimation* pada regresi logistik. Regresi logistik biner tahap pemisahan (*Quasi-Complete Separation*) menghasilkan faktor-faktor yang mempengaruhi (signifikan) ibu hamil yang terkena anemia di Makassar bulan Januari-Agustus pada tahun 2017 adalah variabel (x_1) adalah jarak kelahiran, usia bayi (x_2) dan pendidikan (x_4). Estimasi regresi logistik biner dengan menggunakan metode *Maximum Likelihood* menghasilkan data tidak signifikan maka harus ada pemisahan data dengan menggunakan metode *Quasi-Complete Separation* persamaannya linier sehingga dapat signifikan dan hasilnya salah satunya terdapat yang tidak signifikan tapi hasil yang diperoleh untuk σ dan standar eror lebih kecil dari pada menggunakan *Maximum Likelihood Estimation*.

ABSTRACT

Sutrisno, Edy. 2020. **Application of Quasicomplete Separation by Method**

Maximum Likelihood in Logistic Regression.. Thesis, Department of Mathematics, Faculty of Science and Technology of the State Islamic University of Maulana Malik Ibrahim Malang. Supervisor: (I) Dr. SriHarini, M.Si (II) Juhari, M.Si

Keywords: Binary logistic regression, quasicomplete separation, Maximum Likelihood Parameter Estimation.

Binary logistic regression is a method used to determine the effect of its predictor variables on categorical response variables. Parameter estimation of binary logistic regression models using the Maximum Likelihood method produces linear and significant data. The Maximum Likelihood Method cannot be used as a result which is not convergent or insignificant. To use Quasi-Complete Separation the results are significant or the results are smaller. The research objective is to apply the Quasi-Complete Separation method of Maximum Likelihood Estimation in logistic regression. Binary logistic regression (Quasi-Complete Separation) produces (significant) factors affecting pregnant women affected by anemia in Makassar in January-August in 2017 are variables (x_1) are birth spacing, baby age (x_2) and education (x_4) . Estimation of binary logistic regression using the Maximum Likelihood method results in insignificant data and must separate data using the Quasi-Complete Separation method. The equation is linear so that it is significant and the results are not significant but the results obtained for σ and the standard error are smaller than using the Maximum Likelihood Estimation.

لخص

سوتريسنو ، إيدي، 2020. الفصل شبه الكامل بالطريقة. تطبيق Maximum likelihood
 أطروحة ، قسم الرياضيات ، كلية العلوم والتكنولوجيا بالجامعة الإسلامية الحكومية في .في الانحدار اللوجستي
 مولانا مالك إبراهيم مالانج
 سري هاريني ، ماجستير .د (ط): (1)، ماجستير المشرف .(1) جوهري S.Pd

المفتاحية الكلمات: الثنائي اللوجستي الانحدار, Maximum likelihood, *quasicomplete separation*

حدار اللوجستي الثنائي هو أحد الأساليب المستخدمة لتحديد تأثير المتغير التنبئي على متغير الاستجابة الفئوية. تستخدم المعلمات المقدرة لنموذج الانحدار اللوجستي الثنائي طريقة maximum likelihood لإنتاج بيانات خطية وهامة. لا يمكن استخدام طريقة الحد Maximum likelihood إذا لم يكن المقدر متقاربًا أو غير ذي أهمية. ثم يجب عليك استخدام الفصل شبه الكامل حتى تكون النتائج كبيرة أو تكون النتائج أصغر. الغرض من هذه الدراسة هو تطبيق حالة الفصل شبه الكامل مع طريقة تقدير الاحتمالية القصوى في الانحدار اللوجستي. تنتج مرحلة الانحدار اللوجستي الثنائي للانفصال (الفصل شبه الكامل) العوامل التي تؤثر (بشكل ملحوظ) على النساء الحوامل المصابات بفقر الدم في ماكاسار في الفترة من يناير إلى أغسطس في عام 2017 ، وهي متغيرات (x_1) هي المباشرة بين الولادات وعمر الطفل (x_2) والتعليم (x_4) . ينتج عن الانحدار اللوجستي المقدر باستخدام طريقة أقصى احتمالية بيانات غير مهمة ، ثم يجب أن يكون هناك فصل للبيانات باستخدام طريقة الفصل شبه الكامل ، والمعادلة خطية بحيث يمكن أن تكون مهمة والنتائج غير مهمة ولكن النتائج التي تم الحصول عليها لـ σ والأخطاء المعيارية أصغر من حول استخدام تقدير أقصى احتمالية.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi memiliki peubah dibedakan ke dalam dua jenis peubah yaitu respon (Y) dan prediktor (X). Pada umumnya, sebelum penelitian sering menggunakan data kuantitatif sebagai peubah responnya. Akan tetapi dalam kenyataannya banyak ditemukan kasus dengan peubah responnya berupa data kualitatif yang berbentuk biner, misalnya jenis kelamin (laki-laki atau perempuan) dan pengambilan keputusan (ya atau tidak). Analisis regresi untuk peubah respon bersifat kualitatif sedangkan peubah predikatornya bisa bersifat kuantitatif, kualitatif atau gabungan keduanya yang dapat digambarkan ke dalam suatu model yang dikenal sebagai model respon biner. Salah satu bagian dari analisis yang mempunyai model respon bersifat biner adalah analisis regresi logistik.

Analisis regresi logistik digunakan untuk mengetahui pengaruh suatu peubah predikatornya terhadap peubah respon bersifat kategorik. Parameter model regresi logistik diduga dengan metode *Maximum Likelihood Estimation (MLE)*. Namun, metode *MLE* tidak dapat digunakan lagi apabila diperoleh penduga yang tidak konvergen. Hal ini dapat terjadi karena kombinasi beberapa prediktor menyebabkan peubah prediktor dan peubah respon terpisah secara kurang sempurna. Kasus seperti ini dalam bidang statistika disebut dengan pemisahan kurang sempurna (*quasicomplete separation*) (Albert dan Anderson, 1984).

Parameter dari model regresi logistik biner diduga dengan metode *MLE*. Namun, dalam suatu kondisi tertentu metode *MLE* tidak dapat digunakan atau hasilnya tidak signifikan dan hasilnya terlalu besar serta hasil yang diperoleh

penduga yang tidak konvergen. Kondisi *MLE* yang tidak signifikan atau hasilnya tidak konvergen maka, digunakan pendekatan metode *Penalized Maximum Likelihood Estimation (PMLE)* yang pertama kali diusulkan oleh Firth (1993). Metode *PMLE* merupakan hasil modifikasi fungsi skor *likelihood* menjadi fungsi skor *Penalized likelihood*. Maka dari itu mengatasi kasus pemisahan kurang sempurna (*Quasicomplete Separation*) dengan memisahkan datanya dalam metode *MLE* supaya hasil yang didapatkan terdapat signifikan satu peubah prediktor atau lebih dan hasilnya lebih kecil bila menggunakan data yang penuh atau sempurna dengan menerapkan pada regresi logistik biner.

Estimasi adalah suatu metode yang dapat memperkirakan nilai dari suatu populasi dengan menggunakan nilai dari sampel. Sementara dalam Kamus Besar Bahasa Indonesia estimasi adalah perkiraan, penilaian, atau pendapat. Dalam al-Quran telah dijelaskan mengenai estimasi atau pendugaan sebagai berikut:

وَالَّذِي يَبْسُ مِنَ الْمَجِضِ مَنْ يَسَابِكُمْ إِنْ أَرْتَبْتُمْ فَعِدَّتُهُنَّ ثَلَاثَةُ أَشْهُرٍ وَالَّذِي لَمْ يَجِضْ وَأَوْلَتْ الْأَحْمَالِ أَجَلُهُنَّ أَنْ يَضَعْنَ حَمْلَهُنَّ وَمَنْ يَتَّقِ اللَّهَ يَجْعَلْ لَهُ مِنْ أَمْرِهِ يُسْرًا

Artinya: Dan perempuan-perempuan yang tidak haid lagi di antara perempuan-perempuan jika kamu ragu (tentang masa iddahnya), maka masa iddah mereka adalah tiga bulan, dan begitu (pula) perempuan-perempuan yang tidak haid. Dan perempuan-perempuan yang hamil waktu iddah mereka itu ialah sampai mereka melahirkan kandungannya. Dan barang siapa yang bertakwa kepada Allah, niscaya Allah menjadikannya baginya kemudahan dalam urusan (Qs. At-Talaq/65:4).

Qs. At-Talaq ayat 4 menjelaskan bahwa wanita-wanita tertalak yang menopause karena usia yang sudah tua, jika kalian ragu-ragu bagaimana iddah mereka, mereka iddahnya adalah tiga bulan. Jika diperhatikan lebih rinci, maka terdapat keraguan yang di alami oleh kebanyakan wanita yang menagalami iddahnya maka pada selanjutnya dijelaskan bahwa masa iddahnya kurang lebih tiga

bulan. Dari gambaran di atas diketahui bahwa itulah contoh estimasi (pendugaan) dalam Al-Qur'an.

Berdasarkan penelitian yang terdahulu yaitu Sholihin (2014) dan Atika (2014), telah dibahas tentang PMLE untuk masalah pemisahan kurang sempurna dengan menggunakan data sekunder tentang kredit ayam potong yang mengandung faktor-faktor yang paling mempengaruhi pemberian kredit tersebut dan studi kasus di badan pengelolaan tambak ikan di kabupaten Magetan. Tujuan dari penelitian ini adalah mengatasi kasus pemisahan kurang sempurna pada regresi logistik biner dengan menggunakan data yang berbeda pada penelitian sebelumnya yaitu Sholihin (2014) dan Atika (2014). Kasus ini terjadi karena satu atau kombinasi beberapa prediktor akan menyebabkan peubah prediktor dan peubah respon terpisah secara kurang sempurna, sehingga penduga menjadi tidak konvergen.

1.2 Rumusan Masalah

Berdasarkan latar belakang maka rumusan masalah pada penelitian ini adalah bagaimana aplikasi kasus *Quasi-Complete Separation* dengan metode *Maximum Likelihood Estimation (MLE)* pada regresi logistik.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah maka tujuan penelitian ini untuk mengaplikasikan kasus *Quasi-Complete Separation* dengan metode *Maximum Likelihood Estimation (MLE)* pada regresi logistik.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat bagaimana mengatasi kasus pemisahan kurang sempurna (*Quasycomplete separation*) yang mana terjadi jika vektor koefisien dari fungsi linier variabel independen melewati paling sedikit satu pengamatan pada setiap kategorik dependennya pada regresi logistik biner.

1.5 Batasan Masalah

Untuk mendekati sasaran yang diharapkan, maka perlu diadakan pembatasan permasalahan dalam penelitian itu yaitu sebagai berikut:

1. Jenis model regresi logistik yang dipakai regresi logistik biner.
2. Metode estimasi parameter regresi logistik biner menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan pemisahan sempurna (*Quasi-Complete Separation*).
3. Penelitian ini diterapkan pada data faktor-faktor yang diduga berpengaruh terhadap terjadinya anemia pada ibu hamil di Makassar pada bulan Januari-Agustus pada tahun 2017 dengan variabel yang digunakan adalah anemia pada ibu yang hamil sebagai variabel respon (Y) yang dikategorikan menjadi dua kategori yaitu ibu hamil tidak terkena anemia (0) dan ibu hamil terkena anemia (1), dengan lima variabel prediktor yaitu jarak kehamilan (X_1), usia ibu (X_2), usia kehamilan (X_3), pendidikan (X_4) dikategorikan menjadi dua yaitu tidak sekolah (0) dan sekolah (1), dan pekerjaan (X_5) dikategorikan menjadi dua yaitu tidak bekerja (0) dan bekerja (1).

1.6 Sistematika Penulisan

Penulisan skripsi ini, peneliti menggunakan sistematika penulisan yang terdiri dari lima bab, dan masing-masing bab dibagi dalam subbab dengan sistematika penulisan sebagai berikut:

BAB I Pendahuluan

Meliputi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

BAB II Kajian Pustaka

Meliputi regresi logistik, regresi logistik biner, metode maksimum *likelihood*, uji multikolinieritas, pemisahan dalam *estimasi maximum likelihood* dan pengujian parameter

BAB III Metode Penelitian

Meliputi pendekatan penelitian dan tahap analisis data.

BAB IV Pembahasan

Deskripsi data yang digunakan pada model regresi logistik biner, penduga parameter dengan metode *maximum likelihood*, pemisahan dalam *estimasi maximum likelihood*, pengujian parameter dan kajian tentang estimasi pada surat AT-Talaq..

BAB V Penutup

Meliputi kesimpulan dan saran.

BAB II
KAJIAN PUSTAKA

2.1 Regresi Logistik

Menurut Collett (1991) untuk populasi yang variabel responya (Y) berdistribusi binomial, distribusi peluangnya adalah sebagai berikut:

$$b(y_i; n, \pi(x_i)) = \binom{n}{y} \pi(x_i)^{y_i} [1 - \pi(x_i)]^{n-y_i} \quad (2.1)$$

di mana:

$$\pi(x) = P(y = 1) = \text{peluang kejadian } y = 1.$$

$$1 - \pi(x) = P(y = 0) = \text{peluang } y = 0.$$

Ragamnya merupakan sebagai berikut $Var(y) = \pi(x_i)[1 - (x_i)]$. Misalkan r adalah banyaknya kejadian berhasil maka penduga bagi proporsi kejadian berhasil adalah $\pi(x_i) = \frac{r}{n}$. Sehingga nilai harapan proporsi dari variabel responnya dapat didekati dengan sebaran atau fungsi logistik atau lebih tepat dikatakan regresi logistik yang merupakan suatu pemodelan matematika dapat digunakan untuk menggambarkan hubungan antara beberapa variabel penjelas X dengan variabel respon Y yang berbentuk dikotomi. Secara matematis model regresi logistik di bangun dari suatu fungsi logistik yang dirumuskan sebagai berikut (Collett, 1991):

$$F(z) = \frac{1}{1 + e^{-z}} \quad (2.2)$$

dengan $z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$

Menurut Agung (2011) dengan mengambil dari pengertian regresi linear, yang menggambarkan nilai rata-rata variabel respon Y untuk bersyarat ini variabel penjelas X merupakan $E(y|x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$. Sedangkan

untuk kasus regresi logistik yang memiliki landasan fungsi logistik seperti pada persamaan 2.2 dimana variabel responnya berbentuk dikotomi, maka:

$$\begin{aligned} E(y|x) &= P(y = 1|x_1, x_2, \dots, x_k) \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} \\ &= (x_i) \end{aligned} \quad (2.3)$$

dari persamaan 2.3 maka dapat pula ditulis sebagai berikut:

$$(x_i) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)} \quad (2.4)$$

Persamaan 2.4 kemudian ditransformasikan dengan menggunakan tranformasi logistik (logit), yang secara matematika dapat dituliskan sebagai berikut Agung (2011) :

$$\begin{aligned} \text{logit}(x_i) &= \log \frac{\pi(x_i)}{[1 - \pi(x_i)]} \\ &= \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k) \end{aligned} \quad (2.5)$$

dengan mentransformasikan peluang keberhasilan (x_i) diharapkan bahwa persamaan 2.5 bernilai pada interval $(-\infty, \infty)$, bersifat linear dan kontinu, serta perlu diperhatikan pula ketentuan-ketentuan berikut ini:

1. Variabel respon biner (dikotomi) yang memiliki sebaran Bernoulli atau binomial.
2. Pemakaian persamaan logit 2.5 dengan asumsi bahwa antara $\log \left\{ \frac{x_i}{1-x_i} \right\}$ dan variabel-variabel penjelas X mempunyai hubungan linear, dengan ketentuan sebagai berikut:

- a. Jika untuk setiap nilai variabel penjelas X terdapat cukup banyak observasi, maka nilai $\log \left\{ \frac{x_i}{1-x_i} \right\}$ dapat dihitung dan dapat menunjukkan kebenaran adanya hubungan linear antara $\log \left\{ \frac{x_i}{1-x_i} \right\}$ dan variabel penjelas X.
- b. Jika variabel penjelas X merupakan indikator 1 dan 0, maka asumsi hubungan linear seperti persamaan logit mutlak berlaku.

2.2 Model Regresi Logistik Biner

Regresi logistik merupakan salah satu pendekatan model matematis yang digunakan untuk menganalisis hubungan satu atau beberapa variabel prediktor (*independent*) dengan sebuah variabel respon (*dependent*) kategori yang bersifat dikotomi. Pada regresi logistik jika variabel responnya terdiri dari dua kategori misalnya $Y = 1$ menyatakan hasil yang diperoleh adalah “sukses” dan $Y = 0$ menyatakan hasil yang diperoleh adalah “gagal” maka regresi logistik tersebut menggunakan regresi logistik biner. Menurut Agresti (2002), variabel (y) yang demikian lebih tepat dikatakan sebagai variabel indikator dan memenuhi distribusi Bernoulli. Fungsi probabilitas distribusi Bernoulli, yaitu:

$$f(y_i; \pi_i) = \begin{cases} \pi_i^{y_i} (1-\pi_i)^{1-y_i} \\ 0 \end{cases} \quad (2.6)$$

dimana : untuk $y_i = 0,1$ dan untuk $1 - y_i$ yang lain,

dengan $\pi_i = P(y_i = 1)$. Dari fungsi distribusi tersebut diperoleh rata-rata:

$$E(Y) = 1 \cdot P(Y = 1) + 0 \cdot P(Y = 0)$$

$$= P(Y = 1) \quad (2.7)$$

Misalkan probabilitas ini dinotasikan sebagai $\pi(x)$ yang bergantung dengan variabel independen $x = (X_1, \dots, X_k)$ dengan $E(y) = \pi$ dan $0 \leq \pi \leq 1$, sehingga diperoleh:

$$E(Y^2) = 1^2\pi(x) + 0^2[1 - \pi(x)] = \pi(x) \quad (2.8)$$

dan variansi dari Y adalah

$$V(Y) = E(Y^2) - [E(Y)]^2 = \pi(x)[1 - \pi(x)] \quad (2.9)$$

Secara umum model probabilitas regresi logistik dengan melibatkan beberapa variabel independen x dapat diformulasikan sebagai berikut:

$$\pi(x_i) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}} \quad (2.10)$$

dimana k adalah banyaknya variabel independen dan $\pi(x_i)$ merupakan peluang terjadinya kejadian puas $\pi(x)$ dan tidak puas $(1 - \pi(x))$. Fungsi $\pi(x)$ merupakan fungsi non linear sehingga perlu dilakukan transformasi logit untuk memperoleh fungsi yang linier agar dapat dilihat hubungan antara variabel dependen (y) dengan variabel independen (x). Model logit dari $\pi(x)$ dinyatakan sebagai $g(x)$, yaitu:

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} \quad (2.11)$$

Model umum regresi logistik setelah disubstitusikan dengan model logit dari $\pi(x)$ diperoleh :

$$\text{Logit } \pi(x_i) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (2.12)$$

Regresi logistik terdapat beberapa perbedaan dengan regresi pada umumnya yaitu regresi logistik tidak mengasumsikan suatu hubungan yang linear antara variabel independen dengan variabel dependen, tidak memerlukan asumsi *multivariate normality* pada variabel independennya, tidak ada asumsi homokedastisitas, variabel independen tidak perlu diubah ke dalam bentuk metrik (interval atau skala ratio), variabel dependen harus bersifat dikotomi (2 kategori, contoh: tinggi dan rendah atau puas dan tidak puas), tidak adanya multikolinearitas, kategori dalam variabel independen harus terpisah satu sama lain atau bersifat eksklusif, dan sampel yang diperlukan dalam jumlah relatif besar, minimum dibutuhkan hingga 50 sampel.

2.3 Distribusi Normal

Menurut Sembiring (1995) distribusi yang penting dalam statistik adalah distribusi normal atau sering pula disebut Gauss. Distribusi normal pertama kali diperkenalkan oleh Abraham de Moivre seorang ahli matematika berkebangsaan Perancis yang melarikan diri ke Inggris sekitar tahun 1685. Fungsi dari distribusi normal adalah

$$f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2}\right)\left(\frac{(x-\mu)^2}{\sigma^2}\right)} \quad (2.13)$$

dengan $e = 2,71828 \dots$ dan $\pi = 3,14159$.

Distribusi ini mempunyai rata-rata μ dan variansi σ^2 . Grafiknya mirip lonceng dan tertentu sepenuhnya bila μ dan σ^2 diketahui. Suatu peubah acak X yang

berdistribusi normal dengan rata-ran μ dan variansi σ^2 sering disingkat dengan lambang $X \sim N(\mu, \sigma^2)$ (Sembiring, 1995).

Distribusi normal dengan rata-ran 0 dan variansi 1 atau ditulis $N(0,1)$. Suatu distribusi $N(\mu, \sigma^2)$ berlaku (Sembiring, 1995):

1. $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0,6828 = 0,68$
2. $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,9544 = 0,95$
3. $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0,9774 = 0,98$

Kita inginkan rata-rata yang terletak ditengah seperti $P(-1,96 < X < 1,96)$. Karena distribusi normal adalah simetris, diketahui bahwa $P(-1,96 < X < 0) = 0,475$. Dapat diperoleh,

$$\begin{aligned} P(-1,96 < X < 1,96) &= P(-1,96 < X < 0) + P(0 < X < 1,96) \\ &= 0,475 + 0,475 = 0,95 \end{aligned}$$

Maka interval $-1,96 < X < 1,96$ atau “interval 2σ ” terdiri dari 95% dari daerah dibawah kurva normal (sebenarnya $X = 1,96$ menghasilkan 95% daerah). Karena titik tidak memiliki daerah dalam distribusi kontinu maka peluang $P(-1,96 \leq X \leq 1,96)$ sama seperti $P(-1,96 < X < 1,96)$ untuk kesederhanaan saja (Doane, 2007).

1. $P(-1,00 < X < 1,00) = 2 \times P(0 < X < 1,00) = 2 \times 0,3413 = 0,6826$
= 68,26%
2. $P(-2,00 < X < 2,00) = 2 \times P(0 < X < 2,00) = 2 \times 0,4772 = 0,9544$
= 95,44%
3. $P(-3,00 < X < 3,00) = 2 \times P(0 < X < 3,00) = 2 \times 0,4987 = 0,9974$

2.4 Uji Multikolinearitas

Uji multikolinieritas merupakan suatu uji yang memiliki tujuan untuk menguji apakah model regresi ditemukan adanya korelasi antar variabel bebas. Model regresi yang baik seharusnya bebas multikolinearitas. Multikolinieritas dapat dilihat dari nilai *tolerance* dan lawannya *Variance Inflation Factor* (VIF). *Tolerance* mengukur variabel independen yang terpilih yang tidak dijelaskan oleh variabel independen lainnya. Jadi, nilai *tolerance* yang rendah sama dengan nilai VIF yang tinggi (Ghozali, 2011).

Nilai VIF dapat diperoleh dengan rumus berikut:

$$VIF = \frac{1}{Tolerance} \quad (2.14)$$

Batas *tolerance value* adalah 0,10 atau VIF adalah 10. Jika $VIF > 10$ dan nilai *tolerance* < 0.10 , maka terjadi multikolinearitas tinggi antar variabel bebas dengan variabel bebas lainnya. Jika $VIF < 10$ dan nilai *tolerance* > 0.10 , maka dapat diartikan tidak terdapat multikolinearitas. Regresi yang baik memiliki VIF disekitar angka 1 (satu) dan mempunyai angka *tolerance* mendekati 1. (Santosa,2012).

2.5 Estimasi Parameter

Regresi logistik dapat dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) untuk memperoleh estimasi dari parameter. Menurut Homser dan Lemeshow (1989) metode MLE digunakan untuk mengestimasi parameter-parameter dalam regresi logistik dan pada dasarnya metode maksimum likelihood memberikan nilai estimasi β . Persamaan regresi yang sesuai memiliki, nilai-nilai

parameter yang terdapat pada model harus diestimasi terlebih dahulu. Metode estimasi yang digunakan dalam regresi logistik adalah metode maksimum *likelihood*. Metode ini memiliki prinsip bahwa nilai estimasi yang digunakan adalah nilai estimasi yang memberikan nilai fungsi *Likelihood* yang paling besar.

Misalkan suatu sampel terdiri dari n observasi dari pasangan $(X_1, Y_k), k = 1, 2, \dots, n$. Model regresi logistik $\pi(x_k) = \frac{e^{g(x_k)}}{1+e^{g(x_k)}}$ setiap pasangan (X_1, Y_k) mempunyai fungsi kepadatan peluang:

$$f(x_i) = \pi(x_k)^{y_k} [1 - \pi(x_k)]^{1-y_k} \quad (2.15)$$

Karena setiap pengamatan diasumsikan independen maka fungsinya *likelihood*nya merupakan perkalian antara masing-masing fungsi *likelihood* yaitu:

$$L(\beta) = \prod_{i=1}^n f(x_i) \quad (2.16)$$

Karena berdistribusi normal maka fungsi distribusi peluang gabungan dari (Y) adalah

$$\begin{aligned} P(Y|\beta, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{\sqrt{2\pi\sigma^2}}(Y-X\beta)^T(Y-X\beta)} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{\sqrt{2\pi\sigma^2}}(Y^T Y - 2\beta^T X^T Y + \beta^T X^T X \beta)} \end{aligned} \quad (2.17)$$

Penduga parameter dapat ditentukan dengan menggunakan metode *maximum likelihood estimation*, terlebih dahulu ditentukan fungsi *likelihood* yang diperoleh dari persamaan 2.17:

$$\begin{aligned}
L(\beta, \sigma^2 | Y) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
&= \frac{1}{(2\pi\sigma)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
&= \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \quad (2.18)
\end{aligned}$$

Maka *log likelihoodnya* adalah

$$\begin{aligned}
L(\beta, \sigma^2 | Y) &= \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
L = \ln l &= \ln \frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}}} \ln e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
&= -\ln(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}} + \ln e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
&= -\ln(2\pi)^{\frac{n}{2}} - \ln(\sigma^2)^{\frac{n}{2}} + \ln e^{-\frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta)} \\
&= -\ln(2\pi)^{\frac{n}{2}} - \ln(\sigma^2)^{\frac{n}{2}} - \frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta) \\
&= -\frac{\pi}{2}(2\pi)^{\frac{n}{2}} - \frac{\pi}{2}(\sigma^2)^{\frac{n}{2}} - \frac{1}{2\sigma^2}(Y^T Y - 2\beta X^T Y + \beta^T X^T X \beta) \quad (2.19)
\end{aligned}$$

Nilai β dapat diperoleh dengan memaksimalkan nilai $L(\beta)$ dan mendiferensialkan $L(\beta)$ terhadap β dan menyamakannya dengan nol. Persamaan ini dapat ditulis dalam bentuk sebagai berikut:

$$\frac{\partial L(\beta, \sigma^2 | Y)}{\partial \beta} = \frac{\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} Y^T Y + \frac{1}{2\sigma^2} 2\beta X^T Y - \frac{1}{2\sigma^2} 2\beta^T X^T Y}{\sigma \beta}$$

$$\begin{aligned}
&= -0 - 0 - 0 + \frac{1}{2\sigma^2} X^T Y - \frac{1}{2\sigma^2} 2 X^T X \beta - \frac{1}{2\sigma^2} X^T X \beta - \frac{1}{2\sigma^2} \beta^T X^T X^T \\
&= \frac{1}{\sigma^2} X^T Y - \frac{1}{\sigma^2} X^T X \beta
\end{aligned} \tag{2.20}$$

Kemudian persamaan 2.20 disama dengankan nol.

$$\frac{\partial L(\beta, \sigma^2 | Y)}{\partial \beta} = 0$$

$$\frac{1}{\sigma^2} X^T Y - \frac{1}{\sigma^2} X^T X \beta = 0$$

$$-\frac{1}{\sigma^2} X^T X \beta = -\frac{1}{\sigma^2} X^T Y$$

$$\frac{1}{\sigma^2} X^T X \beta = \frac{1}{\sigma^2} X^T Y$$

$$X^T X \beta = X^T Y$$

$$(X^T X)^{-1} (X^T X) \beta = (X^T X)^{-1} X^T Y$$

$$\beta = (X^T X)^{-1} X^T Y$$

Jadi estimasi dari parameter β adalah

$$\beta = (X^T X)^{-1} X^T Y \tag{2.21}$$

2.6 Pemisahan dalam Metode Maximum Likelihood Estimation

Menurut Albert and Anderson (1984) analisis regresi biner merupakan jenis analisis statistika yang digunakan untuk mengetahui pengaruh satu atau beberapa variabel dependen terhadap satu variabel dependen yang bersifat kualitatif biner. Umumnya parameter model ini diestimasi dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan bantuan metode iteratif *Newton Raphson*.

Namun, penggunaan metode MLE memiliki kelemahan yakni jika dalam data mengandung pemisahan (*separation*) maka akan mengakibatkan hasil estimasi parameter dan *standart error* menjadi tidak konvergen, sehingga hasil pengujian hipotesis cenderung akan menerima H_0 . Pada kondisi seperti ini, maka hasil estimasi parameter dengan metode MLE tidak dapat digunakan.

Menurut Albert and Anderson (1984) bahwa pemisahan disebabkan oleh adanya satu atau kombinasi linier beberapa variabel independen yang mengakibatkan nilai-nilai variabel independen dan dependen terpisah secara sempurna. Kondisi ini sering terjadi pada kasus yang melibatkan ukuran sampel kecil (*small/sparse dataset*). Pemisahan dapat dibedakan menjadi pemisahan sempurna (*complete separation*) dan pemisahan kurang sempurna (*quasi-complete separation*). Pemisahan sempurna terjadi jika vektor koefisien dari fungsi linier variabel independen sama dengan nol yang mengakibatkan pengamatan terbagi ke dalam setiap variabel dependen secara sempurna. Adapun pemisahan kurang sempurna terjadi jika vektor koefisien dari fungsi linier variabel independen melewati paling sedikit satu pengamatan pada setiap kategori dependen.

Pemisahan dalam data pada respon biner perlu mendapatkan perhatian serius karena menghambat proses analisis data. Untuk mengatasi hal ini bisa dengan mengeluarkan variabel yang menjadi penyebab pemisahan atau menambah ukuran sampel. Mengeluarkan variabel independen yang menjadi penyebab pemisahan bisa menjadi kerugian karena ada kemungkinan justru variabel tersebut yang memiliki pengaruh terbesar terhadap variabel dependen. Menambah ukuran sampel juga bukan pilihan yang tepat karena akan menghabiskan sumber daya baik biaya, tenaga maupun waktu (Albert and Anderson ,1984).

Model regresi logistik biner memiliki pola dasar untuk variabel dependen biner $Y_i, i \in 1, 2, \dots, N$ dan vektor k kovariat X_i dengan koefisien $k \times 1$ yang sesuai vektor β :

$$\Pr(Y_i = 1|X_i\beta) = \pi_i = \frac{1}{1 + \exp(-X_i\beta)} \quad (2.22)$$

Kemudian *log-likelihoodnya* untuk model adalah sebagai berikut:

$$\begin{aligned} \ln L(\beta|Y) &= \sum_{i=1}^N \{Y_i \ln \left[\frac{1}{1 + \exp(-X_i\beta)} \right] \\ &+ (1 - Y_i) \ln \left[\frac{1}{1 + \exp(-X_i\beta)} \right] \end{aligned} \quad (2.23)$$

seperti fungsi skor:

$$\frac{\partial \ln L(\beta|Y)}{\partial \beta} = U(\beta) = \sum_{i=1}^N \left[Y_i - \frac{1}{1 + \exp(-X_i\beta)} \right] X_i = 0 \quad (1) \quad (2.24)$$

Menurut McCullagh dan Nelder (1989) persamaan (2.24) dapat diselesaikan untuk $\hat{\beta}$ menggunakan metode iteratif standar kemudian dapat dicapai dengan mempertimbangkan elemen-elemen diagonal dari kebalikan dari matriks informasi standar, dievaluasi pada $\hat{\beta}$ sebagai berikut:

$$\text{var } \hat{\beta} = -[I(\beta)]^{-1} = - \left\{ E \left[\frac{\partial^2 \ln L(\beta|Y)}{\partial \beta \partial \beta'} \right] \right\}^{-1} \quad (2.25)$$

Estimasi regresi logistik biner dalam pemisahan terjadi ketika satu atau lebih kovariat model dengan sempurna memprediksi variabel hasil Y . Menurut Albert dan Anderson (1984), pemisahan yang terakhir menunjukkan kasus di mana prediksi sempurna seperti itu hanya terjadi untuk sebagian pengamatan dalam data.

Pemisahan menyiratkan keberadaan subvektor $X_s \subseteq X$ dimana semua pengamatan N dapat dikategorikan dengan benar sebagai $Y_i = 0$ atau $Y_i = 1$. Kehadiran subvektor ini menyebabkan monotonitas dalam kemungkinan (*log*) dan menghasilkan estimasi kemungkinan maksimum. Dimana $\hat{\beta}_s$ untuk variabel dalam X_s yang sama dengan infinity positif atau negatif dan estimasi kesalahan standar yang terkait juga tidak terbatas. Memahami secara intuitif hubungan antara fenomena pemisahan dan estimasi parameter tak terhingga yang dihasilkan, pertimbangkan kebalikan pemisahan (Albert dan Anderson, 1984).

2.7 Pengujian Parameter

Regresi logistik membandingkan nilai hasil pengamatan dari variabel respon menjadi nilai prediksi dilihat dari model dengan atau tanpa variabel pertanyaannya. Regresi logistik perbandingan dalam penelitian dan nilai prediksi ditentukan dalam fungsi *log-likelihood*). Uji signifikansi masing-masing variabel prediktor yang terdapat dalam model dapat dilakukan menggunakan Uji *Wald*. Uji *Wald* didapat dengan membandingkan *estimasi maximum likelihood* dari parameter $\beta_1, \beta_2, \dots, \beta_i$ dengan estimasi dari *standart error* (Hosmer dan Lemeshow, 1989).

2.7.1 Pengujian Parameter dengan Uji *Likelihood Ratio* (Uji Simultan atau Uji G)

Statistika uji G, yaitu uji rasio kemungkinan (*likelihood ratio test*) digunakan untuk menguji peranan variabel independen didalam model secara bersama-sama. Uji rasio kemungkinan (*likelihood ratio test*) diperoleh dengan cara membandingkan fungsi *log likelihood* dari seluruh variabel bebas dengan fungsi *log*

likelihood tanpa variabel bebas (Raharjanti dan Wildiharih, 2005). Uji G digunakan untuk menguji hipotesis:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (tidak ada variabel independen yang berpengaruh terhadap variabel dependen)

H_1 : Paling sedikit ada satu j dengan $\beta_j \neq 0$ $j = 1, 2, 3, \dots, p$ (paling sedikit ada satu variabel independen yang berpengaruh terhadap variabel dependen)

Menurut *Hosmer-Lemeshow* (1989), statistik uji rasio *likelihood*, G adalah fungsi dari L_0 dan L_k yang berdistribusi X^2 (*chi-square*) dengan derajat bebas p yang didefinisikan sebagai berikut:

$$G = -\ln\left[\frac{L_0}{L_k}\right]$$

$$G = -2 \ln(L_0) - (-2 \ln(L_k))$$

Kemudian L_0 adalah *likelihood* tanpa variabel independen dan L_k adalah *likelihood* dengan variabel independen.

Jika hanya ada satu variabel independen, maka L_0 adalah fungsi *likelihood* yang dihitung sebagai β_0 dan L_1 adalah fungsi kemungkinan *likelihood* yang diestimasi dengan $\hat{\beta}$. Statistika uji G ini mengikuti sebaran *chi-square* bila n mendekati tak terhingga dengan derajat bebas p dimana $p = (r - 1)(c - 1)$, r dan c masing-masing adalah banyaknya kategori pada variabel independen dan variabel dependen. H_0 akan ditolak pada tingkat signifikan α apabila nilai $G > X^2_{(p, \alpha)}$ atau H_0 ditolak ($p - value$) $< \alpha$. Statistik G akan bernilai positif dan semakin kecil nilai $\frac{L_0}{L_1}$ maka akan meningkatkan nilai $2 \ln\left[\frac{L_0}{L_1}\right]$, sehingga peneliti mempunyai bukti

cukup kuat untuk menolak H_0 (Agresti,2007). Kesimpulan bahwa variabel independen, dapat juga dikatakan bahwa paling sedikit ada satu koefisien $\beta_j \neq 0$. untuk mengetahui β_j mana yang berpengaruh signifikan, dapat dilakukan uji parameter β_j secara parsial dengan Uji Wald.

2.7.2 Pengujian Parameter dengan Uji Wald (Uji Parsial)

Pengujian variabel dilakukan satu per satu menggunakan statistika Uji Wald Hosmer dan Lemeshow (1989). Uji ini dilakukan dengan membandingkan model terbaik yang dihasilkan oleh uji simultan terhadap model tanpa variabel bebas di dalam model terbaik.

Hipotesis yang akan diuji sebagai berikut:

$$H_0: \beta_j = 0$$

$$H_1: \beta_j, \text{ dimana } j : 1, 2, 3, 4, \dots, p$$

Statistika uji $W = \left[\frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \right]; j = 1, 2, 3, \dots, p$ dimana $\hat{\beta}_j$ adalah penduga dari β_j

dan $Se(\hat{\beta}_j)$ adalah penduga galat baku dari β_j .

W diasumsikan mengikuti sebaran *chi square* dengan derajat bebas 1. H_0 akan ditolak jika nilai $W > X_2^2$ atau ($p - value$) $< \alpha$. Jika H_0 ditolak maka dapat disimpulkan bahwa β_j signifikan. Variabel independen X secara parsial berpengaruh signifikan terhadap variabel dependen.

2.8 Kajian Dalam Al-Qur'an

Pada surat at-Talaq terdapat ayat yang menyinggung masalah matematika yaitu tentang estimasi. Surat at-Talaq merupakan surat Madaniyyah yang turun sebelum nabi Muhammad Saw hijrah ke Makkiyah. Surat at-Talaq pada tafsir al-Muyassari bahwa wanita hamil pada masa iddahnya adalah melahirkan kandungannya. Barang siapa takut kepada Allah dan menerapkan hukum-hukumNya, niscaya Allah menjadikan urusanNya mudah, di dunia dan di akhirat.

Menurut tafsir al-Mukhtasar bahwa Wanita-wanita tertalak yang menopause karena usia yang sudah tua, jika kalian ragu-ragu bagaimana iddah mereka, maka iddah mereka adalah tiga bulan. Dan wanita wanita yang hamil, batas iddah mereka karena ditalak atau karena ditinggal mati suami adalah jika mereka telah melahirkan. Barang siapa bertakwa kepada Allah dengan menjalankan segala perintah-Nya dan menjauhi segala larangan-Nya niscaya Allah memudahkan berbagai urusannya dan kesulitan. Bagaimana pun wanita-wanita dalam kehidupan di dalam masyarakat kadang ragu-ragu pada masa iddahnya maka hanya Allah swt yang Maha mengetahui. Namun, dalam lanjutan tafsir tersebut maka dijelaskan bahwa masa iddah wanita-wanita itu tiga bulan. Maka dari ayat tersebut untuk estimasi atau perkiraan saat iddahnya untuk setiap wanita tua atau ditinggalkan suaminya mati sudah termasuk sebuah estimasi dalam statistika (Al-Mukhtasar, 1974:286).

BAB III

METODE PENELITIAN

3.1 Variabel Penelitian

Variabel dinyatakan segala sesuatu yang akan menjadi obyek penelitian. Menurut fungsi variabel penelitian dibedakan menjadi variabel bebas dan tergantung. Variabel bebas satu atau lebih dari variabel yang lain itu memungkinkan dipilih sebagai variabel yang sengaja dipelajari pengaruhnya terhadap variabel tergantung. Sedangkan variabel tergantung adalah variabel yang menjadi pusat persoalan dan karena itu tidak mengherankan kalau sering pula disebut kriterium (Suryabrata,2006).

Berdasarkan judul yang penulis teliti dalam penelitian ini, penulis menggunakan dua fungsi variabel bebas dan variabel bergantung variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

Y : Status Anemia (0 = Tidak terkena anemia, 1= Terkena anemia)

X_1 : Jarak kehamilan (Tahun)

X_2 : Usia ibu (Tahun)

X_3 : Usia Kehamilan (Tahun)

X_4 : Pendidikan (0 = Tidak Sekolah , 1= Sekolah)

X_5 : Pekerjaan (0 = Tidak bekerja, 1=Bekerja)

3.2 Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah data dan informasi yang telah dipublikasikan oleh skripsi (Aprilyani, 2017) yang diperoleh dari RSKD ibu dan Anak Siti Fatimah Makassar. Variabel respon pada penelitian ini adalah *headcount index* kehamilan di kota Makassar. *Head count index* adalah prosentase ibu hamil yang terkena anemia. Data yang digunakan merupakan data hamil yang diasumikan mengikuti regresi logistik biner beserta faktor-faktor yang diduga berpengaruh terhadap terjadinya anemia pada ibu hamil di Makassar pada bulan Januari-Agustus (Aprilyani, 2017). Data tersebut diambil melalui situs <http://eprints.unm.ac.id/6321> pada tanggal 11 Maret 2020.

3.3 Tahap Analisis Data

Langkah-langkah untuk menganalisis data dengan regresi logistik yang mengikuti prosedur analisis sebagai berikut:

1. Deskripsi data
2. Tahap uji multikolinieritas
3. Tahap Pemodelan Pendugaan Parameter Menggunakan Metode *Maximum Likelihood*
4. Tahap pemisahan kurang sempurna
5. Tahap pengujian parameter
 - a. Pengujian Uji *Likelihood Ratio*
 - b. Tahap uji parameter dengan Uji Wald (Uji Parsial)

BAB IV

PEMBAHASAN

4.1 Deskripsi Data

Pada penelitian ini model Regresi Logistik Biner diterapkan pada kasus faktor-faktor yang diduga berpengaruh terhadap terjadinya anemia pada ibu hamil di Makassar pada bulan Januari-Agustus pada tahun 2017. Variabel respon yang digunakan dalam penelitian ini adalah anemia pada ibu yang hamil sebagai variabel respon (Y) yang dikategorikan menjadi dua kategori yaitu ibu hamil tidak terkena anemia (0) dan ibu hamil terkena anemia (1), dengan lima variabel prediktor yaitu jarak kehamilan (X_1), usia ibu (X_2), usia kehamilan (X_3), pendidikan (X_4) dikategorikan menjadi dua yaitu tidak sekolah (0) dan sekolah (1), dan pekerjaan (X_5) dikategorikan menjadi dua yaitu tidak bekerja (0) dan bekerja (1). Data yang akan digunakan dalam penelitian ini dapat dilihat pada Lampiran 1.

4.2 Uji Multikolinieritas

Model regresi logistik biner dapat digunakan untuk melihat antara faktor-faktor yang diduga terjadinya kasus anemia pada ibu hamil di Makassar. Model regresi logistik biner maka terlebih dahulu dilakukan pengujian multikolinieritas pada data pada Lampiran 1. Untuk mengetahui adanya multikolinieritas antar variabel prediktor menggunakan Nilai *Variance Inflation Factor* (VIF). Merujuk pada persamaan (2.14) nilai VIF untuk X_1 adalah:

$$VIF = \frac{1}{0,524} = 1,908$$

Nilai VIF yang lebih dari 10 menunjukkan adanya multikolinieritas antar variabel prediktor. VIF untuk X_1 sampai X_5 dapat dilihat pada tabel 4.1

Tabel 4 1 Nilai VIF Variabel Predikator data Ibu Hamil di Makassar

Variabel	Nilai VIF
X_1	1,907
X_2	1,987
X_3	1,066
X_4	1,027
X_5	1,059

Tabel 4.1 Menunjukkan bahwa antara variabel predikatornya yaitu kehamilan (X_1), usia ibu (X_2), usia kehamilan (X_3), pendidikan (X_4) dan pekerjaan (X_5) tidak terdapat multikolinieritas karena nilai untuk setiap VIF lebih dari 1,0 sehingga semua variabel predikator yang mempengaruhi kasus anemia pada ibu hamil di Makassar dapat digunakan dalam pembentukan model regresi logistik.

4.3 Tahap Pendugaan Parameter Menggunakan Metode Likelihood

Metode *Maximum Likelihood Estimation* (MLE) merupakan metode yang digunakan untuk megestimasi parameter-parameter dalam regresi logistik dan pada dasarnya untuk memberikan nilai estimasi β dengan memaksimalkan fungsi *likelihood*-nya. Maka untuk pemeriksaan keberadaan pemisahan dalam model regresi biner MLE menerapkan pada persamaan (2.12) sebagai berikut:

Syarat Y bernilai 1 dan 0, dimana $E(Y) = \pi(x)$, $0 < \pi(x) < 1$ maka dari

rumus $\pi(x) = \frac{e^{\beta x}}{1+e^{\beta x}}$ untuk mencari Y menjadi Z normal maka diperoleh rumus :

$$Z = \text{logit } \pi(x) = g(x) = \ln \frac{\pi(x)}{1-\pi(x)}, \text{ dengan syarat } Z = g(x)$$

$$\begin{aligned}
&= \ln \pi(x) - \ln(1 - \pi(x)) \\
&= \ln\left(\frac{e^{\beta x}}{1 + e^{\beta x}}\right) - \ln\left(1 - \frac{e^{\beta x}}{1 + e^{\beta x}}\right) \\
&= \ln\left(\frac{e^{\beta x}}{1 + e^{\beta x}}\right) - \ln\left(\frac{1 + e^{\beta x} - e^{\beta x}}{1 + e^{\beta x}}\right) \\
&= \ln\left(\frac{e^{\beta x}}{1 + e^{\beta x}}\right) - \ln\left(\frac{1}{1 + e^{\beta x}}\right) \\
&= \beta x - \ln(1 + e^{\beta x}) - \ln(\ln 1 - (\ln(1 + e^{\beta x}))) \\
&= \beta x - \ln(1 + e^{\beta x}) - \ln 1 + \ln(1 + e^{\beta x}) \\
&= \beta x - \ln 1 \\
&= \beta x
\end{aligned}$$

Maka diperoleh hasil yang linear sebagai berikut:

$$Z = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_3 X_3 + \beta_5 X_5$$

dimana Z yakni dengan membandingkan standar error estimasi parameter dari β_0 sampai β_5 yaitu sebagai berikut:

Tabel 4 2 Hasil Estimasi Parameter Regresi Logistik

Parameter	Estimasi	Standar error	P-value
β_0	3.756	1.814	0.055
β_1	-0.021	0.105	0.834
β_2	0.002	0.041	0.96
β_3	-0.033	0.028	0.232
β_4	-0.938	0.899	0.303
β_5	-0.714	0.754	0.350

Tabel 4.2 dapat dilihat bahwa dari 5 parameter yang menunjukkan koefisien regresi logistik dengan data normal yang digunakan tidak ada satu pun variabel yang berpengaruh signifikan terhadap variabel dependen (yang berstatus anemia). Hal ini dikarenakan dari lima parameter jarak kehamilan (X_1), usia ibu (X_2), usia kehamilan (X_3), pendidikan (X_4) dan pekerjaan (X_5) memiliki nilai signifikan lebih dari 0,05. Selain itu, dilihat berdasarkan nilai standar error yang dihasilkan, tampak bahwa standar error untuk β_0 tergolong cukup besar dari X_1 sampai X_5 . Dua keadaan ini, merupakan sinyal awal adanya pemisahan dalam data atau model yang digunakan. Oleh karena itu, perlu dilakukan pemeriksaan terhadap keberadaan pemisahan dalam model.

4.4 Tahap Pemisahan (Quasi-Complete Separation)

Pemisahan dibedakan menjadi dua yaitu pemisahan sempurna dan pemisahan kurang sempurna. Pemisahan sempurna merupakan bila terjadi kombinasi linier dari prediksi menghasilkan prediksi yang sempurna dari variabel respon sedangkan pemisahan kurang sempurna (*quasi-complete separation*) hampir mirip dengan pemisahan lengkap atau sempurna. Predikatornya menghasilkan nilai yang sempurna dari variabel respons untuk sebagian besar nilai prediktor, tetapi tidak sempurna.

Permasalahan pada pemisahan kurang sempurna (*quasi-complete separation*) maka peneliti mengolah data pada Lampiran 1 yang mana data akan dipisahkan (dibagi) menjadi dua data dan diurutkan dari terkecil hingga terbesar untuk mengamati peristiwa dengan probabilitas rendah. Semakin banyak prediktor dalam model maka semakin besar kemungkinan pemisahan terjadi karena masing-

masing kelompok dalam data memiliki ukuran sampel yang lebih kecil. Maka akan diperoleh data sebagai berikut:

Tabel 4 3 Data Penelitian yang Dipisahkan

Data asli	Y	X1	X2	X3	X4	X5
9.8	0	4	24	40	1	0
9.8	0	2	25	39	1	0
9.8	0	1	24	30	1	0
9.9	0	0	20	38	1	0
9.9	0	0	3	39	1	0
9.9	0	6	36	42	1	0
10.1	0	0	28	16	1	0
10.1	0	1	27	40	1	0
10.1	0	5	35	40	1	0
10.1	0	2	37	20	1	0
10.2	0	2	38	37	1	0
10.2	0	3	35	38	1	1
10.3	0	2	25	42	0	0
10.3	0	0	2	24	1	0
10.5	1	0	25	41	1	0
10.5	1	1	28	36	0	0
10.7	1	0	20	40	1	0
10.8	1	0	21	31	1	0
10.8	1	2	27	39	1	0
10.8	1	0	22	30	1	0
10.9	1	0	21	40	1	0
12.6	1	0	20	40	1	0
12.6	1	8	39	12	1	0

Tabel 4.3 merupakan untuk variabel kategorik untuk Y dengan syarat untuk nilai kurang dari 10,5 bernilai 0 dan jika Y lebih dari 10,5 atau sama maka akan bernilai 1. Sedangkan untuk nilai X kita ambil X_1 yang memiliki nilai atau bernilai 2 maka untuk Y nya bernilai 1 atau 0. Ini merupakan suatu data tumpang tindih direntang data membuat pemisahan kurang sempurna (*quasi-complete seperation*).

Maka dari Tabel 4.3 diperoleh hasil sebagai berikut:

Tabel 4.4 Hasil Estimasi Regresi Logistik yang Setelah Dipisahkan

Parameter	Estimasi	Standar error	<i>P-value</i>
β_0	0.164	0.752	0.830
β_1	-0.020	0.066	0.770
β_2	0.001	0.016	0.045
β_3	-0.004	0.013	0.764
β_4	0.334	0.399	0.403
β_5	-0.316	0.546	0.571

Tabel 4.4 dapat dilihat bahwa dari 5 parameter yang menunjukkan koefisien regresi logistik dengan data normal yang dipisahkan, tidak ada dari lima parameter jarak kehamilan (X_1), usia ibu (X_2), usia kehamilan (X_3), pendidikan (X_4) dan pekerjaan (X_5) yang berpengaruh signifikan terhadap variabel dependen (yang berstatus anemia). Hal ini dikarenakan untuk lima parameter tersebut memiliki nilai Z normal lebih dari $\alpha = 0.05$. Suatu variabel independen dikatakan berpengaruh signifikan jika memiliki lebih dari α . Selain itu, dilihat berdasarkan nilai standar error yang dihasilkan, tampak bahwa standar error untuk β_0 sudah semakin kecil. Dibandingkan menggunakan data keseluruhan (sebelum dipisahkan) untuk variabel independen yang berpengaruh signifikan terhadap variabel dependen yang menghasilkan standar error yang cukup besar.

4.5 Tahap Pengujian Parameter

Regresi logistik biner yang didalamnya ada sebuah pengujian supaya hasil yang didapatkan menjadi hasil yang baik maka adanya sebuah pengujian yang mana dilakukan oleh peneliti dari kesimpulan bahwa pada Tabel 4.4 yang menyatakan

bahwa adanya signifikan terhadap X_2 dan standar error semakin mengecil maka akan adanya sebuah pengujian sebagai berikut:

1. Uji *Likelihood* Ratio

Pengujian parameter model regresi logistik biner dilakukan secara serentak (bersama) dan parsial (individu). Pengujian secara serentak dilakukan dengan menggunakan uji nisbah kemungkinan (*Likelihood Ratio Test*) dengan hipotesis sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (tidak ada variabel independen yang berpengaruh terhadap variabel dependen)

H_1 : Paling sedikit ada satu j dengan $\beta_j \neq 0$ $j = 1, 2, 3, \dots, p$ (paling sedikit ada satu variabel independen yang berpengaruh terhadap variabel dependen)

Tabel 4.5 Hasil Uji *Ratio Test*

	Df	Chi-square	P-value
Model	8	88.813	0.000

Berdasarkan Tabel 4.5 Uji *Likelihood Ratio test* nilai *P-value* model sebesar 0,000. Dengan taraf signifikansi 10% maka nilai *P-value* kurang dari $\alpha = 0,1$ dan didapatkan nilai G sebesar 88.813 nilai $X_{0,1,5}^2 = 44,419$. Sehingga $G > X_{0,1,5}^2 = 44,419$ sehingga tolak H_0 . Jadi dapat disimpulkan bahwa variabel prediktor yang digunakan secara bersama-sama (serentak) berpengaruh signifikan terhadap anemia pada ibu hamil atau minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap anemia pada ibu hamil.

2. Tahap Uji Parameter dengan Uji Wald (Uji Parsial)

Pada tahap ini peneliti menggunakan uji Wald. Uji Wald merupakan pengujian variabel dilakukan satu per satu menggunakan statistik Uji Wald dilakukan dengan membandingkan model terbaik yang dihasilkan oleh uji simultan terhadap model tanpa variabel bebas di dalam model terbaik.

$H_0: \beta_j = 0$; (tidak terdapat variabel X yang berpengaruh terhadap variabel Y)

$H_1: \beta_j \neq 0$; $j = 0,1,2, \dots , 5$ (paling sedikit ada satu variabel X yang berpengaruh terhadap variabel Y)

Tabel 4 6 Hasil Uji Wald pada Regresi Logistik

Parameter	Estimasi	Wald	<i>P-value</i>
β_0	-0.086	0.000	0.927
β_1	1.262	1.824	0.036
β_2	0.035	2.2951	0.058
β_3	-0.006	0.013	0.911
β_4	-0.361	3.056	0.013
β_5	-20.751	-4.000	1.000

Berdasarkan Tabel 4.6 maka diperoleh nilai Uji Wald yang akan digunakan oleh peneliti sebagai pengujian parsial yang akan dibandingkan dengan nilai $Z_{0,05} = 1,64$, taraf signifikan 10% sehingga dari lima variabel prediktor yang signifikan atau memiliki nilai uji Wald $> 1,64$ sehingga H_0 ditolak sehingga dapat disimpulkan bahwa β_j signifikan. Variabel independen X secara parsial berpengaruh signifikan terhadap variabel dependen adalah jarak kelahiran (X_1), usia bayi (X_2), pendidikan (X_4). Kemudian dilakukan pengujian kembali untuk mendapatkan model regresi logistik biner. Sehingga diperoleh hasil estimasi yang bisa dilihat pada Tabel 4.7:

Tabel 4 7 Hasil Estimasi Regresi Logistik Biner

Parameter	Estimasi	Wald	P-value
β_0	-0.311	-0.024	0,877
β_1	1.269	1.824	0.0371
β_2	0.027	2.295	0.0581
β_4	-0.426	3.056	0.0776

Sehingga model regresi logistik biner yang dihasilkan adalah

$$\pi(x) = \frac{\exp(1,269X_1 + 0,027X_2 - 0.426X_4)}{1 + \exp(1,269X_1 + 0.027X_2 - 0.426X_4)}$$

Berdasarkan model regresi logistik biner yang terbentuk dapat diinterpretasikan bahwa setiap kenaikan presentase kenaikan jarak kelahiran (X_1) sebesar satu persen. Maka akan meningkatkan rata-rata ibu hamil yang terkena anemia di Makassar sebesar $\exp(1,269) = 3,5254 \approx 4$ kasus. Selanjutnya untuk setiap kenaikan presentase usia bayi (X_2) sebesar satu persen. Maka akan meningkatkan rata-rata ibu hamil yang terkena anemia di Makassar sebesar $\exp(0,027) = 1,0202 \approx 1$ kasus. Setiap kenaikan presentase pendidikan (X_4) sebesar satu persen. Maka akan meningkatkan rata-rata ibu hamil yang terkena anemia di Makassar sebesar $\exp(0,426) = 1,521 \approx 2$ kasus.

4.6 Kajian Nilai-nilai Al-Qur'an dengan Teori Estimasi

Pada surat at-Talaq terdapat ayat yang menyinggung masalah matematika yaitu tentang estimasi. Surat at-Talaq merupakan surat Madaniyyah yang turun sebelum nabi Muhammad Saw hijrah ke Makkiyah. Surat at-Talaq pada tafsir al-Muyassari bahwa wanita hamil pada masa iddahnya adalah melahirkan kandungannya. Barang siapa takut kepada Allah dan menerapkan hukum-hukumNya, niscaya Allah menjadikan urusanya mudah, di dunia dan di akhirat.

Menurut tafsir al-Mukhtasar bahwa Wanita-wanita tertalak yang menopause karena usia yang sudah tua, jika kalain ragu-ragu bagaimana iddah mereka, maka iddah mereka adalah tiga bulan. Dan wanita wanita yang hamil, batas iddah mereka karena ditalak atau karena ditinggal mati suami adalah jika mereka telah melahirkan. Barang siapa bertakwa kepada Allah dengan menjalankan segala perintah-Nya dan manjauhi segala larangan-Nya niscaya Allah memudahkan berbagai urusanya dan kesulitan (Al-Mukhtasar, 1974:286).

Menurut Tafsir al-Wajiz bahwa perempuan-perempuan yang telah sampai pada usia menopause karena sudah tua atau semacamnya sehingga haidnya sudah berhenti. Jika kalian menanyakan tentang masa iddah mereka (makanya kalian tidak tahu sehingga betanya) dan perempuan-perempuan kecil atau sedang sakit sehingga darah (haid) mereka berhenti maka masa, iddahnya adalah 3 bulan selama dalam keadaan talak bukan karena ditinggal mati, sedangkan masa iddah wanita hamil adalah sampai melahirkan. Barang siapa menaati Allah, maka Dia akan memudahkan urusannya di dunia dan akhirat serta menolongnya dalam setiap kebaikan. Ayat ini diturunkan bagi perempuan yang masih muda maupun yang sudah tua yang sudah tidak haid dan mereka sedang hamil (Al- Wajiz, 1974:278).

Bagaimana pun wanita-wanita dalam kehidupan di dalam masyarakat kadang ragu-ragu pada masa iddahnya maka hanya Allah swt yang Maha mengetahui. Namun, dalam lanjutan tafsir tersebut maka dijelaskan bahwa masa iddah wanita-wanita itu tiga bulan. Maka dari ayat tersebut untuk estimasi atau perkiraan saat iddahnya untuk setiap wanita tua atau ditinggalkan suaminya mati sudah termasuk sebuah estimasi dalam statiska.

Adanya pemahaman dan pendalaman teori serta penerapan dalam suatu aplikasi, maka pokok pembahasan ini mengikuti paradigma ulul albab yang mengembangkan pendekatan nasionalis, empiris dan logis (*Bayani dan Burhani*) sekaligus pendekatan intuitif, imajinatif dan metafisis (*Irfani*). Menurut Abdussakir (2007:155-156) konsep tabiyatul ulul albab berlaku dimana dunia akademik dengan adanya kegiatan mendidik dan belajar yang dilakukan oleh dosen dan mahasiswa semata-mata hanya untuk mendekatkan diri kepada Allah SWT. Mahasiswa mencari ilmu pengetahuan melalui berbagai penelitian, karena derajat ulul albab wajib disandang oleh seorang mahasiswa.



BAB V

PENUTUP

5.1 Kesimpulan

Metode Maximum Likelihood Estimation (MLE) memiliki hasil yang tidak signifikan atau hasilnya tidak konvergen maka hal ini digunakan untuk kasus pemisahan. Faktor-faktor yang mempengaruhi ibu hamil dengan tahap pemisahan (Quasi-Complete Separation), maka faktor-faktor yang mempengaruhi ibu hamil yang terkena anemia di Makassar bulan Januari-Agustus pada tahun 2017 adalah variabel X_1 adalah jarak kelahiran, usia bayi (X_2) dan pendidikan (X_4). Model regresi logistik biner dengan parameter maximum likelihood diperoleh sebagai berikut:

Probabilitas kasus ibu hamil yang tidak terkena anemia .

$$\hat{\pi} = P(Y = 1|X)$$

$$\pi(x) = \frac{\exp(1,269X_1 + 0,027X_2 - 0.426X_4)}{1 + \exp(1,269X_1 + 0.027X_2 - 0.426X_4)}$$

Probabilitas kasus ibu hamil yang terkena anemia

$$1 - \hat{\pi} = 1 - P(Y = 0|X) = \frac{1}{1 + \exp(1,269X_1 + 0.027X_2 - 0.426X_4)}$$

5.2 Saran

Penelitian berikutnya adalah Menerapkan Quasi-Complete Separation metode *Maximum Likelihood Estimation* pada model regresi lain seperti regresi logistik ordinal atau multinomial.

DAFTAR RUJUKAN

- Abdussakir. 2007. *Matematika Dalam Al-Qur'an*. Malang: UIN-Maliki Press
- Al- Mukhtasar, Ahmad Musthafa. 1974. *Tafsir Al-Mukhtasar*. Toha Putra:
Semarang.
- Atika,G. R. 2014. *Kajian Metode Maximum Penalized Likelihood Estimation
Untuk Mengatasi Kasus Quaicomplete Separation Pada Regresi Logistik
Biner* :Universitas Negeri Malang.
- Albert, A., and J. A. Anderson. 1984. *On the existence of maximum likelihood
estimates in logistic regression models*.*Biometrika* 71(1):1–10.
- Agresti, A.1990.*Categorical Data Analysis*.John Wiley And Sons:New York.
- Agresti, A.2007.*Categorical Data Analysis*.John Wiley And Sons:New York.
- Agung, G.N. 2011. *Analisis Hubungan Kausal Berdasarkan Data Kategorik*. PT
Raja Grafindo Persada, Jakarta.
- Collet, D. 1991. *Modelling Binary Data*. Chapman and Hall, London..
- Firth, David.1993. *Bias reduction of maximum likelihood estimates*. *Biometrika*
80(1):27–38.
- Hosmer, D.W., dan Lemeshow, S. 1989. *Aplied Logistic Regression*. John Wiley
& Sons Inc., New York.
- Hosmer ,D.W. and Lemeslow.1985. *Applied Logistic Regression Second Edition*.
John Wilwy Dan Sons,Inc: New York.
- Georg Heinze and Michael Schemper.1985. *A solution to the problem of separation
in logistic regression*.
- Ghozali, I.2011. *Aplikasi Multivariate dengan Program IBM SPSS, Edisi Ketiga*.
Semarang: Universitas Diponegoro.

- McCullagh, Peter, and John A. Nelder. 1989. *Generalized linear models*, 2nd ed. Boca Raton, FL: Chapman and Hall.
- Santoso, S. 2012. *Panduan Lengkap SPSS versi.20*. Jakarta: PT Elex Media Komputindo.
- Sembiring, R.K. 2003. *Analisis Regresi*. Bandung: Penerbit ITB.
- Sholihin, Miftahus. 2014. *Implementasi Metode Penalized Maximum Likelihood Estimation Pada Model Regresi Logistik Biner*: Universitas Jember.
- Suryabarata, Sumadi. 2006. *Rancangan-rancangan Berbagai metode dan macam Penelitian dan Peranan Statistik dalam proses penelitian*. Mataram : PT Grafindo Persada.
- Varamita, Aprilyani. 2017. *Analisis Regresi Logistik Dan Aplikasinya*. Di akses melalui internet yaitu. <http://eprints.unm.ac.id/6321> pada tanggal 13 November 2019 pada pukul 09.09 AM.
- Walpole, R. 1995. *Pengantar Statistika*. Ed. Ke-3. Gramedia Pustaka Utama, Jakarta.
- Zorn, Christopher. 2005. *A solution to separation in binary response models*. Political Analysis 13(2):157–70.

LAMPIRAN

LAMPIRAN 1

Data Penelitian Kode	Y	X1	X2	X3	X4	X5
1	8.1	2	33	36	1	0
2	12.6	0	20	40	1	0
3	10.2	2	38	37	1	0
4	9.8	0	16	37	1	0
5	9	2	24	38	1	1
6	7.6	0	20	39	1	0
7	12.6	8	39	12	1	0
8	9.8	9	36	38	1	0
9	8.9	5	29	41	1	0
10	10.1	0	28	16	1	0
11	8.7	2	26	39	1	0
12	10.5	0	25	41	1	0
13	9.2	0	23	33	1	0
14	8.1	4	29	32	1	0
15	9.4	2	22	36	1	0
16	9.5	2	22	40	1	0
17	9.6	7	24	32	1	0
18	9.7	0	21	38	1	0
19	10.5	1	28	36	0	0
20	10.3	2	25	42	0	0
21	8.3	0	20	25	1	0
22	10.8	0	21	31	1	0
23	8.8	6	42	40	1	0
24	7.6	1	22	39	1	0
25	10.1	1	27	40	1	0
26	10.8	2	27	39	1	0
27	9.8	4	24	40	1	0
28	10.1	5	35	40	1	0
29	10.1	2	37	20	1	0
30	8.5	2	25	39	1	0
31	9.9	0	20	38	1	0
32	10.2	3	35	38	1	1
33	9.8	2	25	39	1	0
34	7.8	1	26	37	1	0
35	8.1	1	20	37	1	0
36	7.6	1	32	24	1	1
37	10.8	0	22	30	1	0
38	9.8	1	24	30	1	0
39	9.9	0	3	39	1	0

40	10.3	0	2	24	1	0
41	9.7	4	1	38	1	0
42	8.3	4	38	38	1	0
43	9.9	6	36	42	1	0
44	10.7	0	20	40	1	0
45	10.9	0	21	40	1	0
46	7.1	6	29	38	1	0



LAMPIRAN 2

Coefficients^a

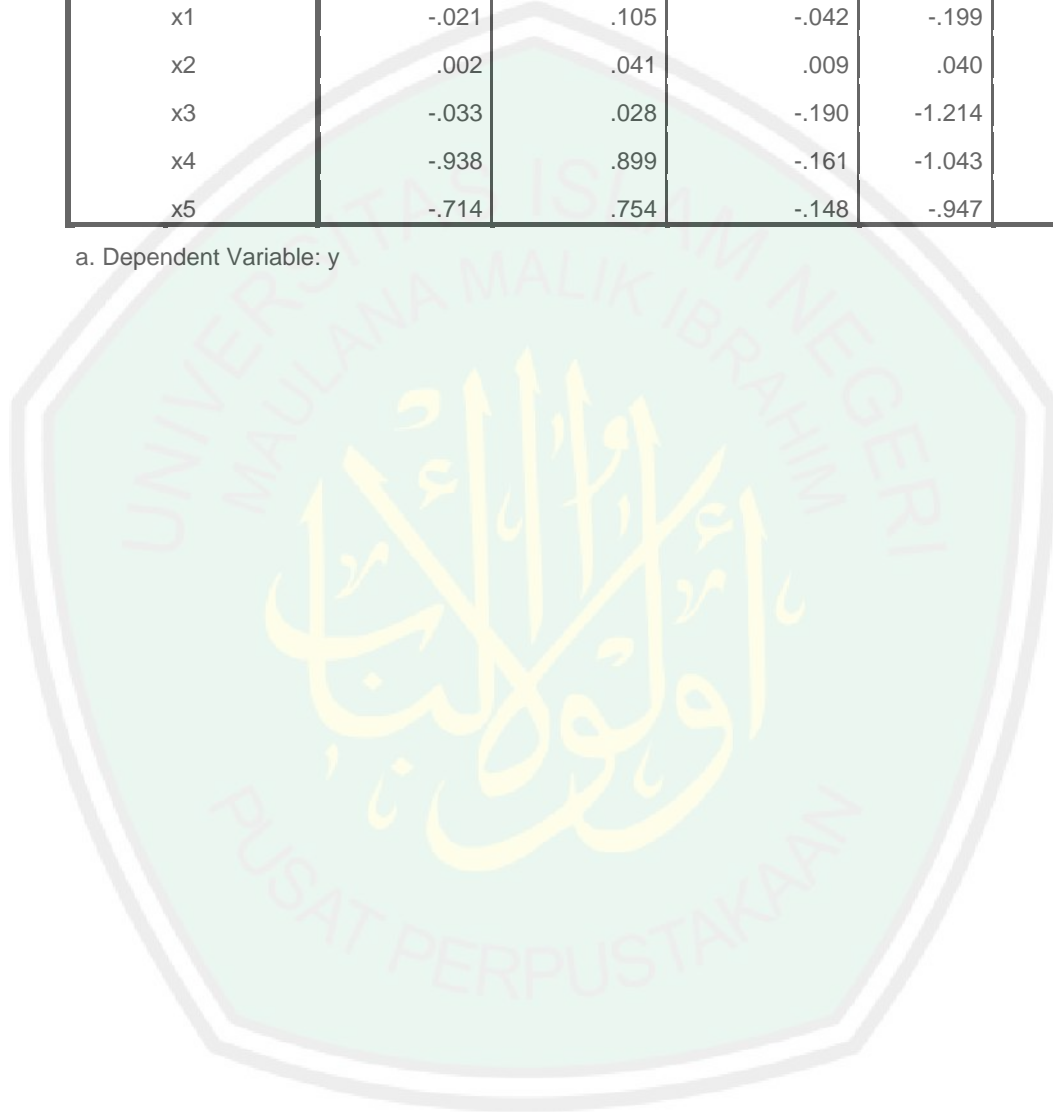
Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.069	.657		.105	.917		
	x1_jarak_kelahiran	.092	.038	.451	2.406	.021	.524	1.907
	x2_usia_bayi	-.030	.015	-.390	-2.038	.048	.503	1.987
	x3_usia_kehamilan	.015	.010	.211	1.505	.140	.938	1.066
	x4_pendidikan	.644	.326	.272	1.977	.055	.974	1.027
	x5_pekerjaan	.177	.273	.090	.647	.521	.944	1.059

LAMPIRAN 3

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.756	1.814		2.070	.045
	x1	-.021	.105	-.042	-.199	.843
	x2	.002	.041	.009	.040	.968
	x3	-.033	.028	-.190	-1.214	.232
	x4	-.938	.899	-.161	-1.043	.303
	x5	-.714	.754	-.148	-.947	.350

a. Dependent Variable: y



LAMPIRAN 4

Data yang di urutkan dan dipisahakan

Y	X1	X2	X3	X4	X5
7.1	6	29	38	1	0
7.6	0	20	39	1	0
7.6	1	22	39	1	0
7.6	1	32	24	1	1
7.8	1	26	37	1	0
8.1	2	33	36	1	0
8.1	4	29	32	1	0
8.1	1	20	37	1	0
8.3	0	20	25	1	0
8.3	4	38	38	1	0
8.5	2	25	39	1	0
8.7	2	26	39	1	0
8.8	6	42	40	1	0
8.9	5	29	41	1	0
9	2	24	38	1	1
9.2	0	23	33	1	0
9.4	2	22	36	1	0
9.5	2	22	40	1	0
9.6	7	24	32	1	0
9.7	0	21	38	1	0
9.7	4	25	38	1	0
9.8	0	16	37	1	0
9.8	9	36	38	1	0
9.8	4	24	40	1	0
9.8	2	25	39	1	0
9.8	1	24	30	1	0
9.9	0	20	38	1	0
9.9	0	3	39	1	0
9.9	6	36	42	1	0

10.1	0	28	16	1	0
10.1	1	27	40	1	0
10.1	5	35	40	1	0
10.1	2	37	20	1	0
10.2	2	38	37	1	0
10.2		35	38	1	1
10.3	2	25	42	0	0
10.3	0	2	24	1	0
10.5	0	25	41	1	0
10.5	1	28	36	0	0
10.7	0	20	40	1	0
10.8	0	21	31	1	0
10.8	2	27	39	1	0
10.8	0	22	30	1	0
10.9	0	21	40	1	0
12.6	0	20	40	1	0
12.6	8	39	12	1	0

LAMPIRAN 5

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	88.813	8	.001



LAMPIRAN 6

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a x1	1.262	.289	1.824	1	.364	.770
x2	.035	.064	2.295	1	.587	1.035
x3	-.006	.055	.013	1	.911	.994
x4	-.361	1.528	3.056	1	.813	.697
x5	-20.751	40192.970	-4.000	1	1.000	.000
Constant	-.286	3.106	.008	1	.927	.752

a. Variable(s) entered on step 1: x1, x2, x3, x4, x5.



LAMPIRAN 7

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a x1	1.262	.289	1.824	1	.364	.770
x2	.035	.064	2.295	1	.587	1.035
x3	-.006	.055	.013	1	.911	.994
x4	-.361	1.528	3.056	1	.813	.697
x5	-20.751	40192.970	-4.000	1	1.000	.000
Constant	-.286	3.106	.008	1	.927	.752

a. Variable(s) entered on step 1: x1, x2, x3, x4, x5.



RIWAYAT HIDUP



Edy Sutrisno, biasa dipanggil Edy, lahir di Magetan pada tanggal 17 Februari 1995. Dia merupakan anak ketiga dari tiga bersaudara dari Bapak Miran dan Marni dan tinggal di Serut RT 01/ RW 01, kelurahan Sidokerto, Kecamatan Sidorejo, Kabupaten Magetan, Jawa Timur.

Pendidikan dasarnya ditempuh di SDN Sidokerto 1 dan lulus pada tahun 2007, kemudian dia melanjutkan Pendidikan menengah pertama di MTs Negeri Sidorejo dan berhasil lulus pada tahun 2010. Setelah itu dia melanjutkan Pendidikan di SMAN 2 Magetan, selama menempuh Pendidikan di SMAN 2 Magetan dia pernah ikut andil dalam kegiatan OSIS dan ikut kegiatan Pramuka BANTARA serta mengambil program Ilmu Pengetahuan Alam dan lulus pada tahun 2013. Selajutnya dia menempuh kuliah di Universitas Islam Negeri Maulana Malik Ibrahim Malang dengan mengambil jurusan Matematika di Fakultas Sains dan Teknologi.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax.(0341)558933


BUKTI KONSULTASI SKRIPSI

Nama : Edy Sutrisno
NIM : 13610075
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Aplikasi *Quasicomplete Separation* dengan Metode *Maximum Likelihood Estimation (MLE)* pada Regresi Logistik.
Pembimbing I : Dr. Sri Harini, M.Si
Pembimbing II : Juhari, S.Pd., M.Si

No	Tanggal	Hal	Tanda Tangan
1.	07 Oktober 2019	Konsultasi Bab I, Bab II dan Bab III	1.
2.	17 Oktober 2019	Konsultasi Kajian Agama Bab I dan Bab II	2.
3.	17 Desember 2019	Revisi Bab I dan Bab III	3.
4.	27 Desember 2019	Revisi Kajian Agama Bab I	4.
5.	07 Februari 2020	ACC Bab I, Bab II dan Bab III	5.
6.	17 Februari 2020	ACC Kajian Agama Bab I, Bab II dan Bab III	6.
7.	19 Februari 2020	Konsultasi Bab III dan Bab IV	7.
8.	26 Februari 2020	Revisi Bab III	8.
9.	04 Maret 2020	Revisi Bab III dan Bab IV	9.
10.	20 Maret 2020	Konsultasi Kajian Agama Bab III dan Bab IV	10.
11.	27 Maret 2020	Revisi Kajian Agama Bab III dan Bab IV	11.
12.	17 Juni 2020	ACC Keseluruhan	12.
13.	17 Juni 2020	ACC Kajian Agama Keseluruhan	13.

Malang, 20 Juni 2020

Mengetahui,
Ketua Jurusan Matematika


Dr. Usman Pagalay, M.Si
NIP. 19650414 200312 1 001