PERBANDINGAN REGRESI *ZERO INFLATED POISSON* DAN *ZERO INFLATED NEGATIVE BINOMIAL* PADA DATA *OVERDISPERSION*

SKRIPSI

OLEH
KAMALIA RIZKI RAHMAWATI
NIM. 13610082

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2018

PERBANDINGAN REGRESI *ZERO INFLATED POISSON* DAN *ZERO INFLATED NEGATIVE BINOMIAL* PADA DATA *OVERDISPERSION*

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
Kamalia Rizki Rahmawati
NIM. 13610082

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2018

PERBANDINGAN REGRESI *ZERO INFLATED POISSON* DAN *ZERO INFLATED NEGATIVE BINOMIAL* PADA DATA *OVERDISPERSION*

SKRIPSI

Oleh
Kamalia Rizki Rahmawati
NIM. 13610082

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal 28 Mei 2018

Pembimbing I,	Pembimbing II,
Dr. Sri Harini, M.Si	Dr. Abdussakir, M.Pd
NIP. 19731014 200112 2 002	NIP. 19751006 200312 1 001

Mengetahui,
Ketua Jurusan Matematika

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

PERBANDINGAN REGRESI *ZERO INFLATED POISSON* DAN *ZERO INFLATED NEGATIVE BINOMIAL* PADA DATA *OVERDISPERSION*

SKRIPSI

Oleh
Kamalia Rizki Rahmawati
NIM. 13610082

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan
Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 17 Juli 2018

Penguji Utama	: Dr. Usman Pagalay, M.Si	....
Ketua Penguji	: Dr. Anwar Fitrianto, Ph.D	....
Sekretaris Penguji	: Dr. Sri Harini, M.Si	....
Anggota Penguji	: Dr. Abdussakir, M.Pd	....

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	: Kamalia Rizki Rahmawati
NIM	: 13610082
Jurusan	: Matematika
Fakultas	: Sains dan Teknologi
Judul Skripsi	: Perbandingan Regresi *Zero Inflated Poisson* dan *Zero Inflated Negative Binomial* pada Data *Overdispersion*

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, 06 Desember 2018
Yang membuat pernyataan,

Kamalia Rizki Rahmawati
NIM. 13610082

MOTO

إِذَا صَدَقَ العَزْمُ وَضَحَ السَّبِيْلُ

“Jika benar tekadnya maka akan jelas perjalanannya”
(Muhafadhah Bahasa Arab)

PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:

Ibunda Raudlatul Jannah dan ayahanda Rukin tercinta yang tiada henti dengan penuh ikhlas dan sabar mendoakan, memberi dukungan, motivasi, dan ridla serta mendengarkan segala keluh kesah penulis, kakak Muhammad Ridho Suaidi serta adik Ahmad Yazid Afthon semoga menjadi anak yang sholeh, dan seluruh keluarga besar yang selalu memberikan doa dan motivasi kepada penulis.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah Swt. atas rahmat, taufik serta hidayah-Nya sehingga penulis mampu menyelesaikan 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. Shalawat serta salam kepada Nabi Muhammad Saw., yang telah menuntun umat manusia menuju jalan yang terang benderang yakni *ad-Diin al-Islam* dan dinantikan syafaatnya kelak.

Proses penyusunan skripsi ini, penulis mendapat banyak bimbingan dan arahan dari berbagai pihak. Untuk itu penulis memberikan ucapan terima kasih 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 sekaligus dosen pembimbing I yang selalu sabar memberikan bimbingan, arahan, dan berbagai ilmunya kepada penulis.
3. Dr. Usman Pagalay, M.Si, selaku ketua Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Abdussakir, M.Pd, selaku dosen pembimbing II yang telah memberikan bimbingan, arahan, dan berbagai ilmunya kepada penulis.

5. Segenap sivitas akademika Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang terutama seluruh dosen, terima kasih atas ilmu dan bimbingannya.
6. Segenap keluarga terutama Ayah dan Ibu yang selalu memberikan doa, semangat, serta motivasi kepada penulis sampai saat ini.
7. Seluruh teman-teman di Jurusan Matematika angkatan 2013 yang tiada hentinya membantu, mendukung, dan mendoakan dalam mewujudkan cita-cita, terima kasih atas kenangan-kenangan indah yang dirajut bersama dalam menggapai cita-cita.
8. Seluruh pihak yang ikut membantu dalam menyelesaikan skripsi ini baik moril maupun materil serta baik secara langsung ataupun tidak langsung.

Penulis berharap agar skripsi ini dapat bermanfaat dan diambil hikmahnya bagi pembaca maupun bagi penulis.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, Mei 2018

Penulis

DAFTAR ISI

DAFTAR SIMBOL

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

i	: banyaknya lokasi pengamatan
k	: banyaknya variabel bebas
y_i	: variabel terikat pada lokasi pengamatan ke-i
λ_i	: rata-rata variabel dependen pada lokasi pengamatan ke-i yang merupakan ekspektasi dari Y dengan $Y \sim poisson\ (\lambda)$
x_i	: vektor variabel bebas pada lokasi pengamatan ke-i
β	: vektor parameter regresi ZIP dan ZINB berukuran $(k+1) \times 1$
γ	: vektor parameter regresi ZIP dan ZINB berukuran $(k+1) \times 1$
$\boldsymbol{Z}$	: matriks berdimensi $n \times (k+1)$ yang berisi variabel bebas yang telah distandarisasikan untuk setiap pengamatan
ε_i	: residual untuk pengamatan ke-i dari model regresi Poisson
β_0	: koefisien intersep model ZIP dan ZINB untuk model diskrit
γ_0	: koefisien intersep model ZIP dan ZINB untuk model *zero inflation*
β_k	: parameter regresi ke-k
γ_k	: parameter regresi ke-k

x_{ik} : nilai variabel bebas ke-*k* pada setiap lokasi pengamatan ke-*i*

$\hat{\lambda}_i$: penduga rata-rata variabel terikat pada lokasi pengamatan ke-*i*

G : nilai *devians*

$L\,(\widehat{\omega})$: nilai *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor

$L\,(\widehat{\Omega})$: nilai *likelihood* untuk model lengkap dengan melibatkan variabel prediktor

db : derajat bebas

p : banyaknya parameter termasuk konstanta

n : banyaknya pengamatan

θ : parameter dispersi

χ^2 : nilai *pearson chi-square*.

$\boldsymbol{X_i^T}$: matriks variabel prediktor

ω : probabilitas observasi benilai nol model ZIP.

ρ : probabilitas observasi benilai nol model ZINB.

DAFTAR TABEL

ABSTRAK

Rahmawati, Kamalia Rizki. 2018. **Perbandingan Regresi *Zero Inflated Poisson* dan *Zero Inflted Negative Binomial* pada Data *Overdispersion*.** Skripsi. Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Sri Harini, M.Si (II) Dr. Abdussakir, M.Pd.

Kata Kunci: Regresi *Zero Inflated Poisson*, Regresi *Zero Inflated Negative Binomial*, *Maximum Likelihood Estimation*, Angka Kematian Ibu.

Regresi Poisson merupakan suatu bentuk analisis regresi yang digunakan untuk memodelkan data yang berbentuk *count* (jumlah). Metode regresi Poisson mensyaratkan adanya equidispersi yaitu kondisi yang nilai *mean* dan *varians* dari variabel respon bernilai sama. Namun adakalanya terjadi fenomena overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi berarti data memiliki *varians* yang lebih besar daripada *mean*. Overdispersi menunjukkan bahwa terdapat heterogenitas populasi. Akibatnya estimasi parameter pada data dengan kondisi yang demikian menjadi tidak tepat. Salah satu metode untuk mengatasi overdispersi adalah dengan model *Zero Inflated Poisson Regression* (ZIP). Kemudian pengembangan dari regresi ZIP juga ada *Zero Infated Negative Binomial Regression* (ZINB). Kedua metode tersebut untuk mengatasi terjadinya banyak pengamatan nol. Penaksiran parameter model regesi ZIP dan ZINB dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) dan diselesaikan menggunakan algoritma Ekspektasi-Maksimalisasi (EM). Penelitian ini menjelaskan tentang faktor-faktor yang mempengaruhi Angka Kematian Ibu di seluruh kecamatan di Kabupaten Malang.

ABSTRACT

Rahmawati, Kamalia Rizki. 2018. **Comparison of Zero Inflated Poisson Regression with Zero Inflated Negative Binomial Regression in Overdispersion Data.** Thesis. Department of Mathematics, Faculty of Sains and Technology, Maulana Malik Ibrahim Islamic University State of Malang. Advisors: (I) Dr. Sri Harini, M.Si (II) Dr. Abdussakir, M.Pd.

Keywords: Zero Inflated Poisson Regression, Zero Inflated Negative Binomial Regression, Maximum Likelihood Estimation, Maternal Mortality Rate.

Poisson regression is a form of regression analysis used to model data in the form of count (number). Poisson regression method requires the existence of equidispersi that is the condition where the mean and variance of the response variable is equal. But sometimes there is an overdispersion phenomenon in the data modeled with the Poisson distribution. Overdispersion means that data has a variance greater than the mean. Overdispersion shows that there is a population heterogeneity. Consequently, the estimation of the parameters on the data under such conditions becomes imprecise. One method to overcome the overdispersion is the Zero Inflated Poisson Regression (ZIP) model. Then the development of ZIP regression is also Zero Inflated Negative Binomial Regression (ZINB). Both methods are to overcome many zero observations. Estimation of ZIP and ZINB regression model parameters is performed using the Maximum Likelihood Estimation (MLE) method and completed using the Expectation-Maximization (EM) algorithm. This research explains about the factors that influence Maternal Mortality Rate in all district in Malang Regency.

ملخص

كامالية.٢٠١٨. **مقارنة الانحدار *Zero Inflated Poisson* و *Zero Inflted Negative Binomial* في بيانات*Overdispersion*** . بحث جامعي .شعبة الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية دولة مالانج .المستشارون: (١)الدكتورة سري هاريني الماجستير (٢) الدكتور عبد الشاكر الماجستير

كلمات البحث: الانحدار*Zero Inflated Poisson* ، الانحدار *Zero Inflated Negative Binomial*، *Maximum Likelihood Estimation* معدل وفيات الأمهات.

انحدار بواسون هو كان يستخدم شكلا من أشكال تحليل الانحدار لنمذجة البيانات في عدد النموذج (عدد). يتطلب بواسون طريقة الانحدار وهو الشرط الذي قيمة المتوسط والتباين للمتغير استجابة من نفس القيمة. ولكن في بعض الأحيان هناك ظاهرة الإفراط في الإفراط في البيانات على غرار نموذج توزيع بواسون. *Overdispersion* يعني أن البيانات لها تباين أكبر من المتوسط. فرط التبخُر يبين أن هناك تباينًا سكانيًا. وبالتالي ، يصبح تقدير المعلمات على البيانات في ظل هذه الظروف غير دقيق. يتم نفخ طريقة واحدة للتغلب على نماذج الانحدار *Zero Inflated Poisson* (ZIP). ثم تطور الانحدارZIP أيضا الانحدار*Zero Infated Negative Binomial* (ZINB). كلتا الطريقتين للحل على العديد من الملاحظات الصفرية. تقدير معالم ZIP و ZINB باستخدام *Maximum Likelihood Estimation* (MLE) وأكمل باستخدام *Expectation-Maximization algorithm* (EM). يشرح هذا البحث العوامل التي تؤثر على معدل وفيات الأمهات في جميع المناطق في مالانج.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Persamaan regresi linier digunakan untuk menganalisis variabel respons yang berupa peubah acak kontinu dan mengikuti distribusi normal, namun banyak ditemukan variabel respons yang tidak berdistribusi normal dan tidak linier dalam parameter. Untuk mengatasi hal tersebut dikembangkan *Generalized Linear Model* (GLM). GLM digunakan sebagai perluasan model regresi umum dengan variabel responnya berdistribusi keluarga eksponensial, antara lain distribusi normal, binomial, Poisson, binomial negatif, eksponensial, gamma, dan invers normal (Myers, dkk, 2010).

Metode utama yang dikembangkan untuk memahami distribusi dari data cacah *(count)* adalah distribusi Poisson, dan kemudian menjadi standar yang digunakan untuk memodelkan data cacah (Hilbe, 2011). Analisis Regresi Poisson adalah suatu model yang digunakan untuk menganalisis hubungan antara variabel respon Y yang berupa data cacah dengan variabel prediktor X. Pada model regresi Poisson dalam analisis data cacah disyaratkan keadaan yang *equidisprsion* yaitu nilai mean dan varians dari variabel respon sama. Ketika nilai varians lebih besar dari nilai meannya disebut *overdispersion*, sedangkan ketika nilai varians lebih kecil dari nilai meannya disebut *underdispersion* (McCullagh & Nelder, 1989). Salah satu metode yang juga dapat digunakan untuk mengatasi *overdispersion* pada distribusi Poisson adalah distribusi *Negative Binomial*, karena distribusi *Negative Binomial* tidak mengharuskan nilai varians sama dengan *mean*nya (Bouk, 2016).

Distribusi *Negative Binomial* dapat mengatasi *overdispersion* pada saat varians lebih dari *mean*. Tidak menutup kemungkinan pada *Negative Binomial* selain data dalam keadaan *overdispersion* juga diperparah dengan kasus *excess zeros.* Dalam beberapa penelitian tertentu, angka nol dapat berarti sangat penting, sehingga tidak dapat diabaikan dalam proses analisisnya. Adanya *overdispersion* dapat menyebabkan model yang terbentuk akan menghasilkan penduga parameter yang bias, menghasilkan nilai penduga bagi kesalahan baku yang lebih kecil *(underestimates)* dan dapat menimbulkan kesalahan dalam pendugaan parameter, selain itu, regresi *Poisson* menjadi tidak tepat lagi untuk memodelkan data.

Regresi *Poisson* yang diterapkan pada data yang mengandung *overdispersion* akan mempengaruhi nilai *standard error* yang menjadi turun (Hilbe, 2011). Masalah *overdispersion* pada regresi Poisson umumnya dapat diatasi dengan menggunakan *Generalized Poisson Regression* atau *Negative Binomial Regression.* Pada umumnya masalah *overdispersion* disebabkan karena kelebihan nol atau terdapat banyak nilai nol pada variabel terikat. Untuk mengatasi hal tersebut salah satu metode yang digunakan adalah metode regresi *Zero Inflated Poisson* (Cameron dan Trivedi, 1998). Hilbe (2011) memperkenalkan juga model distribusi *Zero Inflated Negative Binomial* yang merupakan perluasan dari *Negative Binomial* untuk mengatasi data yang *overdispersion* dan data cacahan nol berlebih atau disebut *excess zeros*.

Pada penelitian sering dijumpai kondisi terlalu banyak nol pada peubah respon (lebih dari 50 persen). Besarnya proporsi data yang bernilai nol dapat berakibat pada ketepatan (presisi) dari inferensia. Dona (2015) merupakan salah satu yang meneliti *Zero Inflated Poisson* (ZIP) untuk memodelkan faktor-faktor yang mempengaruhi terjadinya kebakaran di Kabupaten Sidoarjo. Ariawan (2012) merupakan salah satu yang meneliti *Zero Inflated Negative Binomial* (ZINB) dengan model aditif untuk data respon diskrit dengan *excess zeros*. Pada penelitian sebelumnya yaitu studi kasus angka kematian ibu di Provinsi Bali yang membandingkan antara model regresi ZIP dengan model regresi ZINB menghasilkan bahwa model regresi ZINB lebih tepat digunakan (Dewanti, dkk, 2016:138).

Adapun kajian tentang perbandingan untuk mencari yang terbaik dibahas dalam Islam disinggung dalam al-Quran surat Az-Zumar/39:18 tersebut adalah sebagai berikut:

ٱلَّذِينَ يَسْتَمِعُونَ ٱلْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُۥٓ ۚ أُو۟لَـٰٓئِكَ ٱلَّذِينَ هَدَىٰهُمُ ٱللَّهُ ۖ وَأُو۟لَـٰٓئِكَ هُمْ أُو۟لُوا۟ ٱلْأَلْبَـٰبِ ١٨

"yang mendengarkan perkataan lalu mengikuti apa yang paling baik di antaranya. Mereka itulah orang-orang yang telah diberi Allah petunjuk dan mereka itulah orang-orang yang mempunyai akal" (QS. Az-Zumar/39:18).

Nabi Muhammad diperintahkan untuk memberi kabar gembira kepada umatnya yang selalu menyembah Allah, dan selalu mendengar perkataan yang benar, serta mengerjakan mana yang paling baik dari semua perkataan yang benar itu. Orang yang mendengarkan perkataan yang baik dan mengerjakan yang paling baik dari semua perkataan yang benar adalah orang yang mendapat taufik dari Allah dan selalu menggunakan akal pikirannya. Dalam ayat ini jika diintegrasikan dengan ilmu statistika yang dimaksud perkataan dapat dimaknai secara luas, dapat berupa suatu metode, model, ataupun yang lainnya. Dari model-model yang didapatkan dalam statistika perlu disimpulkan untuk memilih model terbaik yang cocok atau sesuai dengan sebaran suatu data. Sesuai dengan anjuran dalam Islam yang termaktub dalam ayat di atas untuk mengerjakan yang paling baik atau memilih yang terbaik dalam segala hal kebenaran tentunya dengan menggunakan akal pikirannya. Maka yang demikian itu adalah orang yang mendapat taufik dari Allah Swt.

Berdasarkan latar belakang di atas, maka penulis menyusun suatu penelitian dengan judul "Perbandingan Regresi *Zero Inflated Poisson* dan *Zero Inflated Negative Binomial* pada Data *Overdispersion"*.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana model regresi *Zero Inflated Poisson* pada data *overdispersion*?
2. Bagaimana model regresi *Zero Inflated Negative Binomial* pada data *overdispersion*?
3. Bagaimana perbandingan model regresi *Zero Inflated Poisson* dan model regresi *Zero Inflated Negative Binomial* pada data *overdispersion*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui model regresi *Zero Inflated Poisson* pada data *overdispersion.*

2. Untuk mengetahui model regresi *Zero Inflated Negative Binomial* pada data *overdispersion*.
3. Untuk membandingkan dan mendapatkan model terbaik dari regresi *Zero Inflated Poisson* dan regresi *Zero Inflated Negative Binomial* pada data *overdispersion.*

1.4 Manfaat Penelitian

Adapun manfaat penelitian ini secara teoritis adalah mendapatkan wawasan baru mengenai analisis regresi terutama tentang model regresi untuk data cacah yang menunjukkan sifat overdispersi, serta pengetahuan tentang model regresi *Zero Inflated Poisson* (ZIP) dan model regresi *Zero Inflated Negative Binomial* (ZINB). Secara praktis, diharapkan dapat menentukan model yang sesuai atau model terbaik dengan tipe data yang ada dan dapat menganalisis data sehingga menghasilkan suatu kesimpulan yang bermanfaat.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut:

1. Estimasi *Maximum Likelihood* yang digunakan adalah metode Newton-Raphson.
2. Pengujian signifikansi secara parsial menggunakan uji *Wald*.
3. Pemodelan dicontohkan pada kasus Angka Kematian Ibu (AKI) dengan usia ibu 20-34 tahun di Kabupaten Malang pada tahun 2014.

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 teori-teori yang berhubungan dengan pembahasan antara lain konsep *Generalized Linear Model*, overdispersi, distribusi Poisson, uji kesesuaian distribusi Poisson, regresi Poisson, multikolinearitas, distribusi ZIP, model regresi ZIP, regresi Binomial Negatif, model regresi ZINB pengujian signifikansi parameter, pemilihan model terbaik, dan kajian keislaman mengenai perbandingan dan *Overdispersion*.

Bab III Metode Penelitian

Meliputi pendekatan penelitian, jenis dan sumber data, variabel penelitian, dan analisis data.

Bab IV Pembahasan

Meliputi perbandingan model regresi ZIP dan ZINB, aplikasi pada data jumlah kasus Angka Kematian Ibu (AKI) di Kabupaten Malang Tahun 2014, dan kajian tentang perbandingan pada surat Az-Zumar.

Bab V Penutup

Meliputi kesimpulan dan saran.

BAB II

KAJIAN PUSTAKA

2.1 *Generalized Linear Model* (GLM)

Pada dasarnya hampir semua jenis data hasil pengamatan dapat dibuat modelnya. Model ini dapat memberikan informasi tentang karakteristik data yang nantinya informasi tersebut dapat memberikan suatu kesimpulan untuk dapat diambil suatu keputusan tertentu. Keanekaragaman jenis data menghasilkan model yang beragam juga. Jika model yang digunakan tidak sesuai, maka dapat berakibat pengambilan kesimpulan atau keputusan yang salah.

Analisis regresi yang responnya termasuk salah satu keluarga eksponensial disebut Generalisasi Model Linier atau lebih dikenal dengan *Generalized Linear Model* (GLM). GLM merupakan perluasan dari proses pemodelan linier untuk pemodelan data yang mengikuti distribusi.

Tirta (2009) menjelaskan bahwa model linier terampat didefinisikan kedalam tiga komponen penting, yaitu:

a. Komponen distribusi, yaitu y berdistribusi keluarga eksponensial;

b. Komponen prediktor X dan koefisien β membentuk prediktor linier, yaitu $\eta = x^T\beta$;

Fungsi *link* yaitu fungsi monoton dan memiliki diferensial g sehingga $g\,(\mu) = \eta$. Adanya fungsi *link* memungkinkan prediktor linier memiliki daerah rentang seluruh bilangan riil $(-\infty < x < \infty)$ tetapi respon y memiliki rentang tertentu (misalnya $0 < y < 1$ untuk binomial) dan bilangan cacah untuk respon hasil pencacahan (*count data*).

2.2 Overdispersi

Hilbe (2011) menjelaskan bahwa overdispersi pada regresi Poisson terjadi ketika varians dari respon lebih besar dari meannya. Berikut adalah beberapa hal yang dapat menyebabkan overdispersi, antara lain:

a. terdapat korelasi antar pengamatan
b. terdapat pelanggaran asumsi distribusi Poisson, yaitu $Var(Y) > E(Y)$
c. terdapat *excess zeros*
d. terdapat *outlier* dalam data,

Overdispersi menyebabkan nilai devians model menjadi sangat besar dan menyebabkan model yang dihasilkan menjadi kurang tepat. Salah satu cara untuk mengatasiadanya kasus overdispersi dalam regresi Poisson adalah dengan mengganti asumsi distribusi Poisson dengan distribusi lain yang lebih fleksibel.

Kategori lain yang digunakan untuk mendeteksi adanya overdispersi dan underdispersi adalah nilai deviance. Bentuk statistika deviance adalah

$$D = 2\sum_{i=0}^{n} y_i \log\left(\frac{y_i}{\mu_i}\right) \tag{2.1}$$

Jika hasil bagi antara nilai statistik D terhadap derajat bebasnya atau statistik X^2 terhadap derajat bebasnya lebih besar dari 1, maka indikasi bahwa telah terjadi overdispersi pada model regresi Poisson. Sedangkan jika nilai hasil bagi kurang dari atau samadengan 1 maka diidentifikasi telah terjadi underdispersi.

2.3 Distribusi Poisson

Sebaran peluang suatu peubah acak Y dikatakan berdistribusi Poisson karena nilai hanya bergantung pada rata-rata banyak sukses atau terjadinya suatu kejadian dalam interval tertentu yaitu waktu atau daerah tertentu. Distribusi Poisson merupakan peluang suatu peristiwa atau kejadian yang dalam keadaan tertentu diharapkan terjadinya sangat jarang. Dalam kehidupan banyak sekali contoh dari distribusi Poisson adalah masalah kecelakaan, perceraian, kelulusan, kematian, kebakaran, banjir, gempa bumi, dan masih banyak lagi.

Distribusi Poisson memiliki bentuk fungsi kepadatan peluang (pdf) yaitu:

$$f(y) = \frac{e^{-\lambda}\lambda^{y}}{y!}\left(y = 1, 2, 3, \ldots\right) \tag{2.2}$$

dan dinotasikan dengan $Y{\sim}POI(\lambda)$, dengan $\lambda(\lambda > 0)$ adalah parameter dari distribusi Poisson. Menurut Bain dan Engelhardt (1992) untuk distribusi Poisson dapat pula didekati dengan distribusi Binomial. Dinotasikan dengan $Y{\sim}BIN(n,p)$ dengan $y = 1,2,3, \ldots n$ dan $(p \to 0)$ dengan $np = \lambda$ maka

$$\lim_{n\to\infty}\binom{n}{y}p^{y}\left(1-p\right)^{n-y} = \frac{e^{-\lambda}\lambda^{y}}{y!} \tag{2.3}$$

Pembuktian pernyataan tersebut sebagai berikut:

$$\binom{n}{y}p^{y}\left(1-p\right)^{n-y} = \frac{n!}{y!(n-y)!}\left(\frac{\lambda}{n}\right)^{y}\left(1-\frac{\lambda}{n}\right)^{n-y}$$

$$= \frac{\lambda^{y}}{y!}\frac{n(n-1)\ldots(n-y+1)}{n^{y}}\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-y}$$

$$= \frac{\lambda^{y}}{y!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\ldots\left(\frac{n-y+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-y} \tag{2.4}$$

Dengan memberikan limit pada kedua ruas persamaan (2.4) dan menggunakan sifat limit pada kalkulus maka

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda} \tag{2.5}$$

dan

$$\lim_{n\to\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\ldots\left(\frac{n-y+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{-y}=1 \tag{2.6}$$

Sehingga persamaan (2.4) menjadi persamaan (2.3) yaitu:

$$\lim_{n\to\infty}\binom{n}{y}p^{y}(1-p)^{n-y}=\frac{e^{-\lambda}\lambda^{y}}{y!}$$

Bain dan Engelhardt (1992) menjelaskan rata-rata dan varian distribusi Poisson dapat dicari dengan terlebih dahulu menentukan fungsi pembangkit momennya sebagai berikut:

$$M(t)=E\left(\mathrm{e}^{tY}\right)=\sum_{Y=0}^{\infty}\mathrm{e}^{tY}\frac{\mathrm{e}^{-\lambda}\lambda^{y}}{y!}=\mathrm{e}^{-\lambda}\sum_{Y=0}^{\infty}\frac{\left(\lambda\mathrm{e}^{t}\right)^{y}}{y!} \tag{2.7}$$

dengan deret Taylor

$$e^{\lambda}=1+\frac{\lambda^{1}}{1!}+\frac{\lambda^{2}}{2!}+\frac{\lambda^{3}}{3!}+\ldots=\sum_{Y=0}^{\infty}\frac{\lambda^{y}}{y!} \tag{2.8}$$

Sehingga persamaan (2.7) menjadi

$$M(t)=e^{-\lambda}e^{\lambda e^{t}}=e^{\lambda\left(e^{t}-1\right)} \tag{2.9}$$

untuk semua nilai t. Selanjutnya menentukan turunan pertama dan kedua dari persamaan (2.9) sehingga menjadi

$$M'(t)=e^{\lambda\left(e^{t}-1\right)}\left(\lambda e^{t}\right) \tag{2.10}$$

$$M''(t)=e^{\lambda\left(e^{t}-1\right)}\left(\lambda e^{t}\right)+e^{\lambda\left(e^{t}-1\right)}\left(\lambda e^{t}\right)^{2} \tag{2.11}$$

Ketika $t=0$ persamaan (2.10) dan (2.11) menjadi

$$M'(0)=e^{\lambda\left(e^{0}-1\right)}\left(\lambda e^{0}\right)=e^{\lambda(1-1)}(\lambda)=e^{\lambda(0)}(\lambda)=e^{0}(\lambda)=\lambda \tag{2.12}$$

$$M''(0)=e^{\lambda\left(e^{0}-1\right)}\left(\lambda e^{0}\right)+e^{\lambda\left(e^{0}-1\right)}\left(\lambda e^{0}\right)^{2}$$
$$=e^{\lambda(1-1)}(\lambda)+e^{\lambda(1-1)}(\lambda)^{2}=e^{\lambda(0)}(\lambda)+e^{\lambda(0)}(\lambda)^{2}=\lambda+\lambda^{2} \tag{2.13}$$

Sehingga rata-rata dan variannya adalah

$$E(Y)=M'(0)=\lambda \tag{2.14}$$

$$Var(Y)=E\left(Y^{2}\right)-\left[E(Y)\right]^{2}=M''(0)-\left[M'(0)\right]^{2}=\lambda+\lambda^{2}-\lambda^{2}=\lambda \tag{2.15}$$

2.4 Uji Kesesuaian Distribusi Poisson

Uji kesesuaian distribusi Poisson merupakan suatu pengujian yang digunakan untuk menunjukkan apakah suatu data tersebut mengikuti distribusi ini atau tidak. Dalam pengujian Kolmogorov-Smirnov atau Chi Square menggunakan Software SPSS. Pada pengujian Kolmogorov-Smirnov hipotesis pengujiannya sebagai berikut:

H_0 : Data mengikuti pola distribusi Poisson.

H_1 : Data tidak mengikuti pola distribusi Poisson.

Menurut Simarmata dkk (2011) statistik uji Kolmogorov-Smirnov dapat ditulis sebagai berikut:

$$D=Max\left|F(y)-F_{0}(y)\right| \tag{2.16}$$

dengan:

D : deviasi maksimum

$F(y)$: fungsi distribusi frekuensi kumulatif teoritis.

$F_0(y)$: fungsi distribusi frekuensi kumulatif sampel.

Nilai D kemudian dibandingkan $D^*(\alpha)$ yaitu nilai kritis pada tabel *Kolmogorov-Smirnov*. Pengambilan kesimpulan pengujian hipotesis sebagai berikut:

Nilai teramati maksimum $D \geq D^*(\alpha)$ maka tolak H_0

Nilai teramati maksimum $D \leq D^*(\alpha)$ maka terima H_0

Pengambilan kesimpulan pengujian hipotesis juga dapat dilakukan sebagai berikut:

$$P_{value} > a \text{ maka terima } H_0$$

$$P_{value} < a \text{ maka tolak } H_0$$

Untuk nilai D, daerah kritis $D^*(\alpha)$ dan juga P_{value} dapat dilihat pada hasil luaran dari *software SPSS.*

2.5 Regresi Poisson

Regresi Poisson merupakan suatu bentuk analisis regresi yang digunakan untuk memodelkan data diskrit. Model regresi Poisson merupakan model standar untuk data diskrit dan termasuk dalam model regresi linier (Cameron & Trivedi, 1998).

Asumsi Regresi Poisson menurut Hilbe (2014) adalah:

1. Distribusi peluang diskrit dengan parameter tunggal yaitu rata-rata.
2. Peubah respon berupa bilangan cacah.
3. Antar pengamatan saling bebas.
4. Peubah respon tidak boleh mengandung banyak nilai nol.
5. Rata-rata dan ragam peubah identik.

2.5.1 Model Regresi Poisson

Model regresi Poisson adalah model regresi nonlinier yang variabel terikatnya distribusi Poisson dan nilainya berupa integer tidak negatif. Jika Y adalah peubah acak Poisson dengan parameter $\lambda > 0$, maka fungsi kepadatan peluang dapat ditulis dalam bentuk:

$$f(y;\lambda) = \exp\left[y_i \ln(\lambda_i) - \lambda_i - \ln(y_i!)\right]$$

$$= \exp\left[y_i\theta_i - b(\theta_i) - \ln(y_i!)\right] \quad (2.17)$$

Persamaan tersebut merupakan suatu bentuk persamaan fungsi distribusi keluarga eksponensial, dengan $\theta_i = \ln(\lambda_i)\,, b(\theta_i) = \lambda_i = \exp(\theta_i)$.

Model regresi Poisson dapat ditulis sebagai

$$y_i = \lambda_i + \varepsilon_i \left(i = 1, 2, \ldots, n\right) \quad (2.18)$$

dengan

y_i: respon ke-i

λ_i: rata-rata peubah respon yang dipengaruhi oleh nilai pengamatan peubah prediktor pada pengamatan ke-i

ε_i: galat ke-i

Menurut Agresti (2007) model regresi Poisson merupakan *Generalized Linear Model* dengan peubah respon (komponen acak) diasumsikan berdistribusi Poisson dan pada GLM terdapat sebuah fungsi linier yang menghubungkan rata-rata peubah respon dengan sebuah peubah prediktor, yaitu:

$$g(\lambda_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{ip}$$

$$= \beta_0 + \sum_{j=1}^{p} \beta_j X_{i,j} \quad (2.19)$$

Fungsi g disebut dengan fungsi penghubung *(link function).*

Menurut Cahyandari (2014) model regresi Poisson menggunakan konsep *Generalized Linier Model* (GLM) dengan fungsi penghubungnya yaitu log karena rata-rata variabel terikatnya berbentuk fungsi Eksponensial dan taksirannya akan bernialai positif. Konsep GLM untuk distribusi Poisson dijelaskan pada saat $g(\lambda_i)$ sama dengan parameter natural $\theta_i (g(\lambda_i) = \theta_i = \ln(\lambda_i))$, sehingga hubungan λ_i dengan prediktor linier η_i dinyatakan dengan $\ln(\lambda_i) = \eta_i$. Dengan menggunakan fungsi hubung log natural tersebut diperoleh model regresi Poisson dalam bentuk:

$$\begin{aligned} \ln(\lambda_i) &= \eta_i \\ \ln(\lambda_i) &= \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} \\ \lambda_i &= \exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}) \end{aligned} \tag{2.20}$$

Menurut Agresti dkk (2007), *link function* yang digunakan pada model regresi Poisson adalah log natural. Dengan demikian model regresi Poisson dapat ditulis sebagai:

$$\ln(\lambda_i) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{i,j} \tag{2.21}$$

$$\lambda_i = \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j X_{i,j}\right) \tag{2.22}$$

Berdasarkan persamaan (2.18) dan (2.22) maka persamaan regresi Poisson secara umum adalah

$$y_i = \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j X_{i,j}\right) + \varepsilon_i \tag{2.23}$$

2.5.2 Estimasi Parameter Regresi Poisson

Penaksiran estimasi parameter regresi Poisson diproleh dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Berdasarkan persamaan (2.2) maka fungsi *likelihood* untuk model regresi Poisson berikut:

$$L(\beta) = \prod_{i=1}^{n} P(y_i; \beta)$$

$$=\prod_{i=1}^{n}\frac{e^{-\lambda}\lambda_i^{y_i}}{y_i!}$$

$$=\frac{\exp\left(-\sum_{i=1}^{n}\lambda_i\right)\left(\prod_{i=1}^{n}\lambda_i^{y_i}\right)}{\prod_{i=1}^{n}y_i} \tag{2.24}$$

Fungsi *log natural-likelihood* diperoleh dengan mengambil nilai ln dari fungsi *likelihood* seperti dijelaskan pada persamaan (2.25) berikut:

$$\ln \mathrm{L}(\beta)=\ln\left\{\prod_{i=1}^{n}P(y_i;\beta)\right\}$$

$$=\ln\left\{\frac{\exp\left(-\sum_{i=1}^{n}\lambda_i\right)\left(\prod_{i=1}^{n}\lambda_i^{y_i}\right)}{\prod_{i=1}^{n}y_i}\right\}$$

$$=-\sum_{i=1}^{n}\lambda_i+\sum_{i=1}^{n}y_i\ln\lambda_i-\sum_{i=1}^{n}\ln y_i! \tag{2.25}$$

Dari persamaan (2.22) bahwa $\lambda_i=\exp(x_i^T\beta)$, persamaan (2.25) dapat ditulis menjadi persamaan (2.26) berikut:

$$\ln L(\beta)=-\sum_{i=1}^{n}\exp\left(x_i^T\beta\right)+\sum_{i=1}^{n}y_i\ln\left(\exp\left(x_i^T\beta\right)\right)-\sum_{i=1}^{n}\ln y_i! \tag{2.26}$$

Kemudian persamaan (2.26) tersebut diturunkan terhadap β_j dan disamadengankan nol sesuai persamaan (2.27) berikut:

$$\frac{\partial\ln L(\beta)}{\partial\beta_j}=\sum_{i=1}^{n}y_i x_{ij}\frac{\exp\left(x_i^T\beta\right)}{\exp\left(x_i^T\beta\right)}-\sum_{i=1}^{n}\exp\left(x_i^T\beta\right)x_{ij}$$

$$=\sum_{i=1}^{n}y_i x_{ij}-\sum_{i=1}^{n}\lambda_i x_{ij}$$

$$=\sum_{i=1}^{n}(y_i-\lambda_i)x_{ij}=0\ ,\ j=1,2,\dots,p \tag{2.27}$$

p adalah jumlah parameter yang ada. Selanjutnya pengestimasian parameter β dilakukan secara iteratif dengan bantuan komputer yang didasarkan pada algoritma *Iterative Reweighted Least Square* (IWLS).

2.6 Multikolinearitas

Pengujian multikolinearitas dilakukan untuk membuktikan ada tidaknya hubungan yang linier antara variabel bebas satu dengan variabel bebas yang lainnya. Adanya multikolinearitas di antara variable-variabel bebas akan menimbulkan kesulitan dalam memisahkan pengaruh masing-masing variabel bebasnya. Multikolinearitas juga merupakan pendorong terjadinya *overdispersion* (Kartiningrum dan Nursaidah, 2013).

Uji multokolinearitas bertujuan untuk menguji apakah variabel dalam model regresi ditemukan adanya korelasi antara variabel bebas. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel bebas. Jika variabel bebas saling berkorelasi, maka variabel-variabel ini tidak ortogonal. Variabel ortogonal adalah variabel bebas yang nilai korelasi sesama variabel bebas sama dengan nol (Ghozali, 2005).

Untuk mendeteksi ada atau tidaknya multikolinearitas didalam model regresi, digunakan: nilai toleransi dan *Variance Inflation Factor* (VIF). Kedua ukuran tersebut menunjukkan setiap variabel bebas menjadi variabel terikat dan diregresi terhadap variabel bebas lainnya. Nilai toleransi mengukur variabilitas variabel bebas yang terpilih yang tidak dijelaskan oleh variabel bebas lainnya. Dengan kriteria pengambilan keputusan suatu model regresi bebas multikolinearitas adalah sebagai berikut:

a. Mempunyai nilai VIF di bawah 10

b. Mempunyai nilai toleransi di atas 0.10

Jika variabel bebas dapat memenuhi kriteria tersebut, maka variabel bebas tersebut tidak mempunyai persoalan multikolinearitas dengan variabel bebas lainnya.

2.7 Distribusi *Zero Inflated Poisson* (ZIP)

Distribusi ZIP merupakan modifikasi dari perumuman distribusi Poisson. Distribusi Poisson mungkin tidak dapat dipakai ketika terdapat banyak 0 pada data sehingga menggunakan distribusi ZIP. Menurut Hossain dan Howlader, (2015) distribusi ZIP mempunyai keadaan yaitu keadaan nol dengan probabilitas ω yang hanya menghasilkan data bernilai nol dan keadaan Poisson dengan probabilitas $(1-\omega)$ yang berdistribusi Poisson dengan rata-rata λ. Dengan dua proses tersebut memberikan distribusi campuran dengan dua komponen yang disebut distribusi ZIP dengan fungsi kepadatan peluang (pdf) yaitu:

$$P(Y=y)=\begin{cases}\omega+(1-\omega)e^{-\lambda}, y=0\\(1-\omega)e^{-\lambda}\left(\dfrac{\lambda^{y}}{y!}\right), y>0\end{cases} \tag{2.28}$$

dengan $0 \leq \omega \leq 1$ dan $\lambda \geq 0$. Distribusi ZIP dituliskan dengan ZIP (ω, λ) dengan parameter ω dan λ.

Rata-rata varian distribusi ZIP dapat dicari dengan terlebih dahulu menentukan fungsi pembangkit momennya sebagai berikut:

$$M(t) = E\left(e^{tY}\right) = \sum_{y=0}^{\infty} e^{ty} f(y)$$

$$= e^{t(0)}\left(\omega + (1-\omega)e^{-\lambda}\right) + \sum_{y=1}^{\infty} e^{ty}\left((1-\omega)e^{-\lambda}\left(\frac{\lambda^y}{y!}\right)\right)$$

$$= \omega + (1-\omega)e^{-\lambda} + (1-\omega)e^{-\lambda}\sum_{y=1}^{\infty} e^{ty}\left(\frac{\lambda^y}{y!}\right)$$

$$= \omega + (1-\omega)e^{-\lambda} + (1-\omega)e^{-\lambda}\sum_{Y=1}^{\infty}\left(\frac{\left(\lambda e^t\right)^y}{y!}\right)$$

$$= \omega + (1-\omega)e^{-\lambda} + (1-\omega)e^{-\lambda}\left(e^{\lambda e^t} - 1\right)$$

$$= \omega + (1-\omega)e^{-\lambda} + (1-\omega)e^{-\lambda}e^{\lambda e^t} - (1-\omega)e^{-\lambda}$$

$$= \omega + (1-\omega)e^{-\lambda}e^{\lambda e^t}$$

$$= \omega + (1-\omega)e^{\lambda\left(e^t - 1\right)} \tag{2.29}$$

Dengan pendekatan menggunakan deret Taylor sesuai persamaan (2.8) maka persamaan $\sum_{y=1}^{\infty} e^{ty}\left(\frac{\lambda^y}{y!}\right)$ dapat ditulis menjadi $e^{\lambda e^t} - 1$ sebagai bukti pada persamaan (2.29).

$$e^{\lambda e^t} = 1 + \frac{\left(\lambda e^t\right)^1}{1!} + \frac{\left(\lambda e^t\right)^2}{2!} + \frac{\left(\lambda e^t\right)^3}{3!} + \ldots = \sum_{Y=0}^{\infty}\left(\frac{\left(\lambda e^t\right)^y}{y!}\right)$$

$$e^{\lambda e^t} = 1 + \sum_{Y=1}^{\infty}\left(\frac{\left(\lambda e^t\right)^y}{y!}\right)$$

$$e^{\lambda e^t} - 1 = \sum_{Y=1}^{\infty}\left(\frac{\left(\lambda e^t\right)^y}{y!}\right)$$

Selanjutnya dari persamaan (2.29) dari fungsi pembangkit momennya, maka rata-rata dan varian sebagai berikut:

$$M(t) = \omega + (1-\omega)e^{-\lambda}e^{\lambda e^t} = \omega + (1-\omega)e^{\lambda\left(e^t - 1\right)}$$

$$M'(t) = (1-\omega)\left(\lambda e^t\right)e^{\lambda\left(e^t - 1\right)} \tag{2.30}$$

$$M''(t)=(1-\omega)\left(\lambda e^{t}\right)e^{\lambda\left(e^{t}-1\right)}+(1-\omega)\left(\lambda e^{t}\right)^{2}e^{\lambda\left(e^{t}-1\right)} \tag{2.31}$$

Untuk $t = 0$ persamaan (2.30) dan (2.31) dapat ditulis menjadi

$$\begin{aligned}M'(0)&=(1-\omega)\left(\lambda e^{0}\right)e^{\lambda\left(e^{0}-1\right)}=(1-\omega)(\lambda)e^{\lambda(1-1)}\\&=(1-\omega)\left(\lambda e^{0}\right)e^{\lambda\left(e^{0}-1\right)}=(1-\omega)(\lambda)e^{\lambda(1-1)}\\&=(1-\omega)(\lambda)e^{\lambda(0)}=(1-\omega)\lambda\end{aligned} \tag{2.32}$$

$$\begin{aligned}M''(0)&=(1-\omega)\left(\lambda e^{0}\right)e^{\lambda\left(e^{0}-1\right)}+(1-\omega)\left(\lambda e^{0}\right)^{2}e^{\lambda\left(e^{0}-1\right)}\\&=(1-\omega)(\lambda)e^{\lambda(1-1)}+(1-\omega)(\lambda)^{2}e^{\lambda(1-1)}\\&=(1-\omega)(\lambda)e^{\lambda(0)}+(1-\omega)(\lambda)^{2}e^{\lambda(0)}\\&=(1-\omega)\lambda+(1-\omega)\lambda^{2}\end{aligned} \tag{2.33}$$

Sehingga rata-rata dan variannya adalah

$$E(Y)=M'(0)=(1-\omega)\lambda \tag{2.34}$$

$$\begin{aligned}&Var(Y)=E\left(Y^{2}\right)-\left[E(Y)\right]^{2}=M''(0)-\left[M'(0)\right]^{2}\\&=\left((1-\omega)\lambda+(1-\omega)\lambda^{2}\right)-\left((1-\omega)\lambda\right)^{2}\\&=\left((\lambda-\lambda\omega)+\left(\lambda^{2}-\lambda^{2}\omega\right)\right)-(\lambda-\lambda\omega)^{2}\\&=\left(\lambda-\lambda\omega+\lambda^{2}-\lambda^{2}\omega\right)-\left(\lambda^{2}-2\lambda^{2}\omega+(\lambda\omega)^{2}\right)\\&=\lambda-\lambda\omega+\lambda^{2}-\lambda^{2}\omega-\lambda^{2}-2\lambda^{2}\omega+(\lambda\omega)^{2}\\&=\lambda-\lambda\omega+\lambda^{2}\omega-(\lambda\omega)^{2}\\&=\lambda\left(1-\omega+\lambda\omega-\lambda\omega^{2}\right)\\&=\lambda(1-\omega)(1+\lambda\omega)\end{aligned} \tag{2.35}$$

2.8 Model Regresi *Zero Inflated Poisson* (ZIP)

Model regresi *Zero Inflated Poisson* (ZIP) adalah model campuran yang sederhana untuk data diskrit dengan banyak peristiwa nol (Lambert, 1992). Jika Y_i merupakan peubah acak bebas yang berdistribusi ZIP, maka nilai nol yang terdapat pada observasi diduga telah terjadi dengan dua cara yang sesuai dengan keadaan *(state)* yang terpisah. Keadaan pertama disebut dengan *zero state* dengan probabilitas ω_i dan keadaan kedua disebut dengan *Poisson state* dengan probabilitas $1-\omega_i$. Kedua keadaan memberikan distribusi campuran dua komponen.

Model hubungan untuk λ dan ω pada model regresi ZIP (Lambert, 1992) adalah sebagai berikut:

$$\begin{gathered} \ln(\lambda_i) = \sum_{k=0}^{p} x_{ik}\beta_k \\ logit(\omega) = \ln\left(\frac{\omega_i}{1-\omega_i}\right) = \sum_{k=0}^{p} x_{ik}\gamma_k \end{gathered} \tag{2.36}$$

Sehingga persamaan (2.36) dapat ditulis sebagai persamaan berikut:

$$\lambda_i = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) \text{ dan } \omega_i = \frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \tag{2.37}$$

2.9 Regresi Binomial Negatif

Regresi Binomial Negatif merupakan salah satu cara untuk mengatasi masalah *overdispersion* pada data cacah yang didasarkan pada model campuran Poisson-Gamma (Hardin & Hilbe, 2007). Fungsi peluang Binomial Negatif adalah

$$p(y|\lambda,\theta) = \frac{\Gamma\left(y+\frac{1}{\theta}\right)}{y!\,\Gamma\left(\frac{1}{\theta}\right)}\left(\frac{\theta\lambda}{1+\theta\lambda}\right)^{y}\left(\frac{1}{1+\theta\lambda}\right)^{\frac{1}{\theta}} \tag{2.38}$$

dengan y merupakan nilai dari data cacah, λ nilai harapan, dan θ merupakan parameter dispersi, $\theta > 0$ (Yesilova, 2010). Jika $\theta \to 0$ maka distribusi ini mendekati sebaran Poisson (λ). Sebaran Binomial Negatif memiliki nilai tengah $E(Y) = \lambda$ dan ragam $Var(y) = \lambda + \lambda^2\theta$(Ismail & Jemain, 2007).

Peubah respon didefinisikan sebagai peubah acak berdistribusi Binomial Negatif dengan fungsi penghubung ln. Model regresi yang akan dibentuk yaitu:

$$\ln(\lambda_i) = X_i\beta \tag{2.39}$$

Parameter regresi Binomial Negatif (β) diduga menggunakan penduga kemungkinan maksimum.

2.10 Model Regresi *Zero-Inflated Negative Binomial* (ZINB)

Model regresi yang dapat digunakan untuk mengatasi overdispersi yang disebabkan oleh *excess zeros* adalah model regresi ZINB. Model regresi ZINB adalah model yang dibentuk dari campuran distribusi Poisson dan Gamma. Jika Y_i merupakan peubah acak yang diskrit, maka nilai nol yang terdapat pada observasi diduga telah terjadi dengan dua cara yang sesuai dengan keadaan (*state*) yang terpisah. Keadaan yang pertama disebut *zero state* dengan probabilitas p_i dan keadaan kedua disebut *negative binomial state* dengan probabilitas$1 - p_i$. Kedua keadaan tersebut memberikan distribusi campuran dua komponen, maka fungsi peluang dari model regresi ZINB (Garay & Hashimoto, 2011) adalah sebagai berikut:

$$P(Y_i = y_i) = \begin{cases} \rho_i + (1-\rho_i)\left(\dfrac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}} & \text{untuk } y_i = 0 \\ (1-\rho_i)\dfrac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y_i!}\left(\dfrac{1}{1+\theta\lambda_i}\right)^{\frac{1}{k}}\left(\dfrac{\theta\lambda_i}{1+\theta\lambda_i}\right)^{y_i} & \text{untuk } y_i = 1, 2, \ldots, k \end{cases} \tag{2.40}$$

Dari fungsi peluang model regresi ZINB, didapat rata-rata dan varians sebagai berikut:

$$E(Y_i) = (1-\rho_i)\lambda_i \quad (2.41)$$

$$Var(Y_i) = (1-\rho_i)\lambda_i(1+\theta\lambda_i+\rho_i\lambda_i) \quad (2.42)$$

Model regresi ZINB dibagi menjadi dua komponen yaitu model data diskrit untuk λ_i dan model *zero-inflation* untuk ρ_i yaitu:

$$\begin{aligned} \ln(\lambda_i) &= \sum_{k=0}^{p} x_{ik}\beta_k \\ logit(\rho_i) &= \ln\left(\frac{\rho_i}{1-\rho_i}\right) = \sum_{k=0}^{p} z_{ik}\gamma_k \end{aligned} \quad (2.43)$$

2.11 *Maximum Likelihood Estimation* (MLE)

Fungsi densitas bersama dari variabel random $X_1, X_2, \dots, X_n$ yang bernilai $x_1, x_2, \dots, x_n$ adalah $L(\beta) = f(x_1, x_2, \dots, x_n; \beta)$ yang merupakan fungsi *likelihood*. Untuk $x_1, x_2, \dots, x_n$ tetap, fungsi *likelihood* merupakan fungsi dari ϑ dan dilambangkan dengan $L(\beta)$. Jika $X_1, X_2, \dots, X_n$ mewakili sebuah sampel random dari $f(x; \beta)$, maka $L(\vartheta) = f(x_1; \beta) \cdot f(x_2; \beta) \dots f(x_n; \beta)$ dapat dituliskan sebagai berikut:

$$\begin{aligned} L(\beta) &= f(\hat{x}; \beta) \\ &= f(x_1, x_2, \dots, x_n; \beta) \\ &= f(x_1; \beta) \cdot f(x_2; \beta) \dots f(x_n; \beta) \\ &= \prod_{i=1}^{n} (x_i; \beta) \end{aligned} \quad (2.44)$$

$L(\beta) = f(x_1, x_2, \dots, x_n; \beta)$ merupakan fungsi densitas probabilitas dari $x_1, x_2, \dots, x_n$. Untuk hasil pengamatan $x_1, x_2, \dots, x_n$, nilai $\hat{\beta}$ berada dalam $\Omega(\hat{\beta}\epsilon\Omega)$, dimana $L(\beta)$ maksimum yang disebut sebagai *Maximum Likelihood Estimation* (MLE) dari β. Jadi $\hat{\beta}$ merupakan nilai dugaan dari β (Hogg & Craig, 1995).

Jika $(x_1, x_2, \dots, x_n) = \max f(x_1, x_2, \dots, x_n; \beta)\beta\epsilon\Omega$, maka untuk memperoleh nilai $\hat{\beta}$ tersebut yang memaksimumkan harus diderivatifkan dengan langkah-langkah sebagai berikut:

1. Nilai $\hat{\beta}$ diperoleh dari derivatif pertama jika:

$$\frac{\partial}{\partial\beta} L(\beta)|_{\beta=\hat{\beta}} = 0$$

2. Nilai $\hat{\beta}$ dikatakan memaksimumkan $L(\beta)$ jika:

$$\frac{\partial^2}{\partial^2\beta} L(\beta)|_{\beta=\hat{\beta}} < 0$$

Selain dengan memaksimumkan fungsi *likelihood*, nilai $\hat{\beta}$ juga dapat diperoleh dengan memaksimumkan fungsi ln *likelihood*, karena dengan memaksimumkan fungsi ln *likelihood*, juga akan memaksimumkan fungsi *likelihood*. Karena fungsi ln *likelihood* merupakan fungsi yang monoton naik, maka untuk memperoleh $\hat{\beta}$ dengan memaksimumkan fungsi ln *likelihood* dapat dilakukan dengan langkah-langkah yang sama, yaitu:

1. Nilai $\hat{\beta}$ diperoleh dari derivatif pertama jika:

$$\frac{\partial}{\partial\beta} \ln L(\beta)|_{\beta=\hat{\beta}} = 0$$

2. Nilai $\hat{\beta}$ dikatakan memaksimumkan $L(\beta)$ jika:

$$\frac{\partial^2}{\partial^2\beta} \ln L(\beta)|_{\beta=\hat{\beta}} < 0$$

2.12 Metode Iterasi *Newton Raphson*

Apabila dalam proses estimasi parameter didapat persamaan akhir yang non linier, maka tidak mudah memperoleh estimasi parameter tersebut, sehingga diperlukan suatu metode numerik untuk memecahkan persamaan non linier tersebut. Salah satu metode yang sering digunakan adalah metode *Newton Raphson*. Metode *Newton Raphson* adalah metode untuk menyelesaikan persamaan non linier secara *iterative* seperti persamaan *likelihood* yang mencari lokasi yang memaksimalkan suatu fungsi. Dasar dari metode ini adalah dengan pendekatan deret *Taylor* linier.

$$f(x) = f\left(x^0\right) + \sum_{i=1}^{p} \frac{1}{i!} \frac{\partial^p f\left(x^0\right)}{\partial\left(x^0\right)^i}\left(x - x^0\right)^i \tag{2.45}$$

Perluasan dari bentuk orde I:

$$\frac{\partial F(\beta)}{\partial \beta} = 0$$

diperoleh:

$$\frac{\partial f(\beta)}{\partial \beta} = f\left(\beta^0\right) + H^0\left(\beta - \beta^0\right)$$

$$\frac{\partial f(\beta)}{\partial \beta} = G^0 + H^0\left(\beta - \beta^0\right)$$

Jika β^0 merupakan nilai awal (inisialisasi) dari β atau β^0 merupakan nilai ke-1 dari β, maka dapat dimisalkan $\beta^0 = \beta^t$ dan $\beta = \beta^{t+1}$ dengan $t_0 = 0$. Begitu pula dengan G dan H. Maka diperoleh iterasi sebagai berikut:

$$\beta^{t+1} = \beta^t - \left(H^t\right)^{-1} G^t \tag{2.46}$$

dengan indeks t menyatakan ukuran iterasi.

Metode ini dapat diperluas untuk menyelesaikan system persamaan dengan lebih dari satu parameter. Misal $\beta_1, \beta_2, \ldots, \beta_k$ maka iterasinya sebagai berikut:

$$\hat{\beta}^{t+1} = \hat{\beta}^{t} - \left(H^{t}\right)^{-1} G^{t} \tag{2.47}$$

dengan $\hat{\beta}^{t+1}$ dan $\hat{\beta}^{t}$ adalah bentuk vektor, yaitu:

$$\hat{\beta}^{t+1} = \begin{bmatrix} \hat{\beta}_1{}^{t+1} \\ \vdots \\ \hat{\beta}_k{}^{t+1} \end{bmatrix} \text{ dan } \hat{\beta}^{t} = \begin{bmatrix} \hat{\beta}_1{}^{t} \\ \vdots \\ \hat{\beta}_k{}^{t} \end{bmatrix}$$

$$H = \begin{bmatrix} \frac{\partial^2 F(\beta)}{\partial \beta_1^2} & \cdots & \frac{\partial^2 F(\beta)}{\partial \beta_1 \beta_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F(\beta)}{\partial \beta_k^2} & \cdots & \frac{\partial^2 F(\beta)}{\partial \beta_k^2} \end{bmatrix} \text{ dan } G = \begin{bmatrix} \frac{\partial F(\beta)}{\partial \beta_1} \\ \vdots \\ \frac{\partial F(\beta)}{\partial \beta_k} \end{bmatrix}$$

2.13 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter digunakan untuk mengetahui kesesuaian atau kecocokan model regresi *Zero Inflated Poisson* dan model regresi *Zero Inflated Negative Binomial*. Pengujian ini dilakukan secara simultan dan parsial.

2.13.1 Pengujian secara Simultan

Pengujian parameter regresi secara simultan dilakukan untuk menguji signifikansi dari koefisien regresi secara bersama-sama. Pengujian ini digunakan untuk mengetahui apakah semua variabel bebasnya secara bersama-sama berpengaruh terhadap variabel terikatnya. Fittriyah dkk (2014) menjelaskan pengujian ini menggunakan *Likelihood Ratio Test* yang didefinisikan sebagai berikut:

$$G = -2\log\left[\frac{L(y:\hat{w})}{L(y:\Omega)}\right] \tag{2.48}$$

Hipotesis yang digunakan adalah:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = \gamma_1 = \gamma_2 = \cdots = \gamma_p = 0$

H_1 : paling sedikit ada satu $\beta_j \neq 0$ atau $\gamma_j \neq 0$ dengan $j = 1, 2, \ldots, p$

2.13.2 Pengujian secara Parsial

Pengujian koefisien regresi secara parsial dapat menggunakan uji *Wald*. Uji *Wald* merupakan salah satu statistik uji yang sering digunakan untuk menguji signifikasnsi parameter regresi secara individual pada masing-masing variabel prediktor. Pengujian parameter model regresi ZIP dan model regresi ZINB secara parsial dibagi menjadi dua yaitu model $\ln(\lambda) = X\beta$ dan model $\text{logit}(\omega) = X\gamma$. Hipotesis yang digunakan untuk model $\ln(\lambda) = X\beta$ yaitu:

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, dengan $j = 1, 2, \ldots, k$

Sedangkan untuk model $\text{logit}(\omega) = X\gamma$ yaitu:

$H_0 : \gamma_j = 0$

$H_1 : \gamma_j \neq 0$, dengan $j = 1, 2, \ldots, k$

Persamaan statistik unutuk uji *Wald* adalah

$$W_j = \frac{\beta_j}{SE(\beta_j)} \text{ dan } W_j = \frac{\gamma_j}{SE(\gamma_j)} \qquad (2.49)$$

Dengan kriteria pengujian $W_j = t_{hitung}$ yaitu:

$$|t_{hitung}| < t_{\left(\frac{\alpha}{2};n-k\right)} \text{ maka Terima } H_0$$

$$|t_{hitung}| > t_{\left(\frac{\alpha}{2};n-k\right)} \text{ maka Tolak } H_0$$

2.14Pemilihan Model Terbaik

Dari model-model yang didapatkan perlu disimpulkan model yang terbaik dan sesuai. Oleh karena itu perlu ada metode atau cara untuk menentukannya. Pada penelitian ini dalam menentukan metode analisis terbaik ZIP dan ZINB dapat dilihat dari nilai *Akaike Information Criteria (AIC)*. Akaike (1969) dalam Rawlings dkk, (1998) menjelaskan nilai AIC dapat didefinisikan sebagai:

$$AIC = n\ln\left(SSRes\right) + 2k - n\ln(n) \qquad (2.50)$$

Nilai $SSRes$ adalah jumlah kuadrat residual, n banyaknya data dan banyaknya parameter termasuk konstata dinyatakan dengan k. Dalam pengujian ini nilai AIC terkecil adala yang terbaik yaitu missal jika nilai $AIC_A < AIC_B$ maka model A lebih baik.

Selain AIC, Rawlings dkk, (1998) menyatakan beberapa kriteria yang digunakan untuk melihat tepat tidaknya model regresi yang diperoleh, salah satunya yaitu dengan melihat koefisien detrminasi (R^2). Nilai koefisien determinasi (R^2) dapat dinyatakan dengan persen dan dicari dengan jumlah kuadrat variabel terikat regresi ($SS_{regresi}$) dengan jumlah kuadrat variabel terikat total (SS_{total}) dan didefinsikan sebagai berikut:

$$R^2 = \frac{SS_{regresi}}{SS_{total}} \times 100\% \qquad (2.51)$$

Nilai koefisien determinasi R^2 menjelaskan semakin besar nilainya, maka taksiran model regresi yang diperoleh semakin baik dalam menjelaskan pengaruh variabel bebas dengan variabel terikat, begitu pula sebaliknya.

2.15 Angka Kematian Ibu (AKI)

Angka Kematian Ibu (AKI) berguna untuk menggambarkan status gizi dan kesehatan ibu, kondisi lingkungan, tingkat pelayanan kesehatan, terutama untuk ibu hamil, melahirkan dan masa nifas. Menurut Depkes (2009), kematian ibu adalah kematian yang terjadi pada ibu selama masa kehamilan atau dalam 42 hari setelah berakhirnya kehamilan oleh setiap penyebab yang berhubungan dengan atau diperberat oleh kehamilan atau penanganan kehamilan dan persalinan tetapi bukan terjadi akibat dari kecelakaan. Faktor-faktor yang mempengaruhi angka kematian ibu antara lain adalah:

a. Kunjungan Kehamilan K1

Kunjungan kehamilan K1 (akses pelayanan ibu hamil) merupakan gambaran jumlah ibu hamil yang telah melakukan kunjungan pertama ke fasilitas pelayanan kesehatan untuk mendapatkan pelayanan. Tahun 2014 cakupan rata-rata kunjungan ibu hamil (K1) di Kabupaten Malang turun sebesar 98,72% (44.537 kunjungan). Cakupan K1 yang dilaporkan selama ini adalah akses, dengan demikian dapat menggambarkan keterjangkauan pelayanan program (Dinkes, 2015).

b. Kunjungan Kehamilan K4

Kunjungan kehamilan K4 adalah gambaran jumlah ibu hamil yang mendapat pelayanan 4 (empat) kali yaitu sekali pada trimester pertama, sekali pada trimester kedua, dan dua kali pada trimester ketiga. Tahun 2014 cakupan rata-rata kunjungan ibu hamil (K4) di Kabupaten Malang meningkat sebesar 97,07% (43.802

kunjungan dari 45.115 ibu hamil yang ada). Cakupan tersebut lebih tinggi bila dibandingkan dengan target nasional yaitu 90% (Dinkes, 2015).

c. Ibu Nifas Mendapat Vitamin A

Vitamin A adalah salah satu zat gizi dari golongan vitamin yang sangat diperlukan oleh tubuh yang berguna untuk kesehatan mata dan untuk kesehatan tubuh. Pemberian Vitamin A pada ibu nifas sangat penting karena ibu hamil dan menyusui beresiko mengalami KVA sehingga pada masa tersebut ibu membutuhkan vitamin A yang tinggi untuk pertumbuhan janin dan produksi ASI. Penanggulangan KVA saat ini masih bertumpu pada pemberian kapsul vitamin A dosis tinggi. Oleh karena itu, pemberian secara periodik dilakukan kepada ibu nifas 2 kapsul vitamin A warna merah yang diminum, 1 kapsul setelah melahirkan dan 1 kapsul lagi setelah 24 jam (Depkes RI, 2011).

d. Persalinan Ditolong Tenaga Kesehatan

Persalinan ditolong tenaga kesehatan adalah pertolongan persalinan oleh tenaga kesehatan yang memiliki kompetensi kebidanan. Komplikasi dan kematian ibu maternal dan bayi baru lahir sebagian besar terjadi pada masa disekitar persalinan, hal ini disebabkan pertolongan tidak dilakukan oleh tenaga kesehatan yang mempunyai kompetensi kebidanan (profesional). Pada tahun 2014 cakupan Persalinan Ditolong Tenaga Kesehatan di Kabupaten Malang turun sebesar 99,85% (42.999 dari 43.064 sasaran ibu bersalin). Cakupan Kabupaten Malang ini lebih tinggi bila dibandingkan dengan target SPM yaitu 90% (Dinkes, 2015).

2.16 Kajian Perbandingan dan Memilih Model Terbaik dalam Islam

2.16.1 Kajian Mengenai Perbandingan dalam Islam

Perbandingan merupakan salah satu kegiatan yang bisa dilakukan dalam berbagai ilmu termasuk dalam ilmu statistika. Perbandingan biasanya digunakan untuk mencari solusi yang lebih tepat atau mencari yang terbaik dalam mengatasi suatu kasus. Dalam al-Quran terdapat ayat yang memuat konsep perbandingan untuk mencari yang terbaik disebutkan dalam al-Quran al-Baqarah/2:148 tersebut sebagai berikut:

وَلِكُلٍّ وِجْهَةٌ هُوَ مُوَلِّيهَا ۖ فَٱسْتَبِقُوا۟ ٱلْخَيْرَٰتِ ۚ أَيْنَ مَا تَكُونُوا۟ يَأْتِ بِكُمُ ٱللَّهُ جَمِيعًا ۚ إِنَّ ٱللَّهَ عَلَىٰ كُلِّ شَىْءٍ قَدِيرٌ ﴿١٤٨﴾

"Dan bagi tiap-tiap umat ada kiblatnya (sendiri) yang ia menghadap kepadanya. Maka berlomba-lombalah (dalam membuat) kebaikan. Di mana saja kamu berada pasti Allah akan mengumpulkan kamu sekalian (pada hari kiamat). Sesungguhnya Allah Maha Kuasa atas segala sesuatu" (QS. al-Baqarah/2:148).

Surat al-Baqarah yang artinya sapi betina adalah surat Madaniyah yakni turun ketika Nabi hijrah ke Madinah. Setiap umat mempunyai kiblat masing-masing. Nabi Ibrahim dan Nabi Ismail a.s, menghadap Ka'bah. Bani Israil menghadap ke Baitulmakdis dan orang Nasrani menghadap ke timur, yang prinsip ialah beriman kepada Allah dan mematuhi segala perintah-Nya. Karena Allah telah memerintahkan agar kaum Muslimin menghadap ke Ka'bah dalam shalat, maka fitnah dan cemoohan dari orang yang ingkar itu tidak perlu dilayani, tetapi hendaklah kaum Muslimin bekerja dengan giat, beramal, bertobat dan berlomba membuat kebajikan. Allah nanti akan menghimpun umat manusia untuk menghitung serta membalas segala amal perbuatannya, dan Allah Maha Kuasa atas segala sesuatu, tidak ada yang dapat melemahkan-Nya untuk mengumpulkan semua manusia pada hari pembalasan (Depag RI, 2010).

Pada surat al-Baqarah ayat 148 tersebut dijelaskan bahwa pada ayat-ayat sebelum ini diterangkan adanya perubahan kiblat bagi umat Islam yaitu menjadi menghadap ke Masjidilharam, maka pada ayat ini ditegaskan bahwa kebenaran adalah dari Allah. Manusia tidak boleh menetapkan kebenaran hanya didasarkan pada emosi dan perasaan saja, melainkan memikirkan kebenaran tersebut dan melengkapinya dengan mengadakan pengamatan dan observasi, serta membahas dan menganalisis dengan akal yang sehat. Oleh karena itu, dengan adanya studi perbandingan dalam ilmu statistik sangat berperan penting untuk membantu mencari solusi atau metode yang paling tepat dari kasus-kasus statistik yang ada.

2.16.2 Kajian Mengenai Pemilihan Model Terbaik dalam Islam

Pemilihan model terbaik dalam ilmu statistika sangat sering digunakan untuk mendapatkan model yang terbaik dan sesuai mengikuti sebaran data yang ada. Adapun kajian keislaman mengenai pemilihan yang terbaik terdapat dalam al-Quran surat az-Zumar/39:18 sebagai berikut:

ٱلَّذِينَ يَسْتَمِعُونَ ٱلْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُۥٓ ۚ أُو۟لَـٰٓئِكَ ٱلَّذِينَ هَدَىٰهُمُ ٱللَّهُ ۖ وَأُو۟لَـٰٓئِكَ هُمْ أُو۟لُوا۟ ٱلْأَلْبَـٰبِ ﴿١٨﴾

"yang mendengarkan perkataan lalu mengikuti apa yang paling baik di antaranya. Mereka itulah orang-orang yang telah diberi Allah petunjuk dan mereka itulah orang-orang yang mempunyai akal" (QS. Az-Zumar/39:18).

Surat az-Zumar termasuk kelompok surat-surat Makkiyyah diturunkan sesudah surat Saba'. Ayat ini menerangkan orang-orang yang selalu menjaga dirinya dan menghindar diri dari menyembah tagut, berhala, serta tabah dalam menghadapi godaan setan, menghambakan diri dan menyembah kepada Allah semata, tidak menyembah selain Allah. Mereka akan memperoleh kabar gembira dari para rasul bahwa mereka akan terhindar dari azab kubur sesudah mati, kesengsaraan di Padang Mahsyar. Mereka akan mendapatkan kenikmatan yang abadi di dalam surga. Oleh karena itu, Nabi Muhammad diperintahkan untuk memberi kabar gembira kepada umatnya yang selalu menyembah Allah, dan selalu mendengar perkataan yang benar, serta mengerjakan mana yang paling baik dari semua perkataan yang benar itu. Mereka pun akan memperoleh apa yang diperoleh oleh hamba-hamba Allah yang takwa. Mereka adalah orang-orang yang selalu mengikuti petunjuk Allah dan selalu menggunakan akal yang sehat (Depag RI, 2010).

Nabi Muhammad diperintahkan untuk memberi kabar gembira kepada umatnya yang selalu menyembah Allah, dan selalu mendengar perkataan yang benar, serta mengerjakan mana yang paling baik dari semua perkataan yang benar itu. Orang yang mendengarkan perkataan yang baik dan mengerjakan yang paling baik dari semua perkataan yang benar adalah orang yang mendapat taufik dari Allah dan selalu menggunakan akal pikirannya. Dalam ayat ini jika diintegrasikan dengan ilmu statistika yang dimaksud perkataan dapat dimaknai secara luas dapat berupa metode, model, ataupun yang lainnya. Dari model-model yang didapatkan dalam statistik perlu disimpulkan untuk memilih model terbaik yang cocok atau sesuai dengan sebaran suatu data. Sesuai dengan anjuran dalam Islam yang termaktub dalam ayat di atas untuk mengerjakan yang paling baik atau memilih yang terbaik dalam segala hal kebenaran tentunya dengan menggunakan akal pikirannya. Maka yang demikian itu adalah orang yang mendapat taufik dari Allah Swt.

Allah menegaskan bahwa umat Islam adalah memang diciptakan untuk menjadi teladan bagi umat-umat yang lain karena mereka membawa misi dakwah, yaitu mengajak kepada perbuatan–perbuatan yang baik dan benar, serta mencegah segala perbuatan yang keji dan mungkar. Hadits tentang anjuran melakukan kebaikan terdapat dalam hadits arbain hadits ke tigapuluh tujuh:

عَنْ ابْنِ عَبَّاسٍ رَضِيَ اللهُ عَنْهُمَا، عَنْ رَسُوْلِ اللهِ صَلى الله عليه وسلم فِيْمَا يَرْوِيْهِ عَنْ رَبِّهِ تَبَارَكَ وَتَعَالَى : إِنَّ اللهَ كَتَبَ الْحَسَنَاتِ وَالسَّيِّئَاتِ، ثُمَّ بَيَّنَ ذَلِكَ : فَمَنْ هَمَّ بِحَسَنَةٍ فَلَمْ يَعْمَلْهَا كَتَبَهَا عِنْدَهُ حَسَنَةً كَامِلَةً، وَإِنْ هَمَّ بِهَا فَعَمِلَهَا كَتَبَهَا اللهُ عِنْدَهُ عَشْرَةَ حَسَنَاتٍ إِلَى سَبْعِمِائَةِ ضِعْفٍ إِلَى أَضْعَافٍ كَثِيْرَةٍ، وَإِنْ هَمَّ بِسَيِّئَةٍ فَلَمْ يَعْمَلْهَا كَتَبَهَا اللهُ عِنْدَهُ حَسَنَةً كَامِلَةً، وَإِنْ هَمَّ بِهَا فَعَمِلَهَا كَتَبَهَا اللهُ سَيِّئَةً وَاحِدَةً "[رواه البخاري ومسلم في صحيحهما بهذه الحروف]

"Dari Ibnu Abbas radhiallahuanhuma, dari Rasulullah Shallallahu'alaihi wasallam sebagaimana dia riwayatkan dari Rabbnya Yang Maha Suci dan Maha Tinggi: Sesungguhnya Allah telah menetapkan kebaikan dan keburukan, kemudian menjelaskan hal tersebut: Siapa yang ingin melaksanakan kebaikan kemudian dia tidak mengamalkannya, maka dicatat disisi-Nya sebagai satu kebaikan penuh. Dan jika dia berniat melakukannya dan kemudian melaksanakannya maka Allah akan mencatatnya sebagai sepuluh kebaikan hingga tujuh ratus kali lipat bahkan hingga kelipatan yang banyak. Dan jika dia berniat melaksanakan keburukan kemudian dia tidak melaksanakannya maka baginya satu kebaikan penuh, sedangkan jika dia berniat kemudian dia melaksanakannya Allah mencatatnya sebagai satu keburukan" (HR. Bukhari dan Muslim).

Seorang muslim hendaklah meniatkan perbuatan baik selalu dan membuktikannya, diharapkan dengan begitu akan ditulis pahala dan ganjarannya. Allah tidak menghitung keinginan hati dan kehendak perbuatan manusia kecuali jika kemudian dibuktikan dengan amal perbuatan dan praktik. Segala kebaikan dalam Islam adalah dianjurkan sebagaimana dalam hadits tersebut jika manusia berniat melakukan kebaikan maka Allah sudah mencatatnya dan apabila telah berbuat kebaikan maka semakin berlipat pahalanya. Namun tidak dengan keburukan apabila manusia masih berniat tidak dicatat sebagai dosa, jika telah berbuat keburukan maka hanya dicatat satu dosa. Maka sesuai dengan isi kandungan surat az-Zumar ayat 18 apabila mengerjakan kebaikan yang paling baik atau memilih yang terbaik dalam segala hal kebenaran tentunya semakin berlipat-lipat pula pahala di sisi Allah.

BAB III

METODE PENELITIAN

3.1 Pendekatan Penelitian

Pendekatan penelitian yang digunakan pada penelitian ini adalah pendekatan studi literatur dan deskriptif kuantitatif. Pada studi literatur yaitu dengan mengumpulkan bahan-bahan pustaka mengenai model regresi ZIP dan ZINB. Sedangkan pendekatan deskriptif kuantitatif yaitu dengan mengaplikasikan dan menganalisis data jumlah kasus Angka Kematian Ibu di Kabupaten Malang Tahun 2014.

3.2 Sumber Data

Data yang digunakan dalam penelitian adalah data sekunder dari Dinas Kesehatan Kabupaten Malang mengenai jumlah kasus Angka Kematian Ibu di Kabupaten Malang Tahun 2014. Terdapat 33 kecamatan di Kabupaten Malang.

3.3 Variabel Penelitian

Pada penelitian ini variabel penelitian dibagi menjadi dua, yaitu variabel terikat (Y) yang merupakan jumlah kasus Angka Kematian Ibu di Kabupaten Malang Tahun 2014 dan untuk variabel bebas (X) yaitu:

Tabel 3.1 Variabel Bebas

Variabel	Keterangan
X_1	Persentase cakupan kunjungan kehamilan k1
X_2	Persentase cakupan kunjungan kehamilan k4
X_3	Persentase ibu nifas mendapatkan vitamin A
X_4	Persentase persalinan ditolong tenaga kesehatan

3.4 Tahap Analisis Data

3.4.1 Estimasi Parameter Model Regresi ZIP dan ZINB

Langkah-langkah estimasi parameter model Regresi ZIP dan ZINB sebagai berikut:

1. Membentuk fungsi kepadatan peluang untuk y_i pada persamaan (2.28) untuk ZIP dan (2.40) untuk ZINB.
2. Membentuk fungsi *likelihood.*
3. Membentuk fungsi *log-natural likelihood* untuk $y_i = 0$ dan $y_i > 0$.
4. Bila terjadi *incomplete data* maka dilakukan algoritma EM, sebagai berikut:
   a. Menentukan distribusi dari variabel z_i untuk ZIP dan v_i untuk ZINB.
   b. Membentuk distribusi gabungan antara y_i dan z_i untuk ZIP serta y_i dan v_i untuk ZINB.
   c. Tahap ekspektasi dengan mengganti variabel Z_i dengan $Z_i^{(m)}$ untuk ZIP dan V_i dengan $V_i^{(m)}$ untuk ZINB.
   d. Tahap maksimalisasi yaitu memaksimalkan β dan γ dengan menghitung $\beta^{(m)}$ dan $\gamma^{(m)}$ dengan menggunakan metode *Newton Raphson.*

3.4.2 Aplikasi pada Data Jumlah Angka Kematian Ibu Di Kabupaten Malang Tahun 2014

Langkah-langkah pada data jumlah kasus Angka Kematian Ibu (AKI) di Kabupaten Malang Tahun 2014 adalah sebagai berikut:

1. Melakukan pengujian distribusi Poisson.
2. Melakukan pengujian Multikolinearitas.

3. Melakukan pengujian Overdispersi.
4. Penggunaan Regresi ZIP dan ZNB menggunakan *software SAS* (9.3).
5. Pendugaan parameter model Regresi ZIP dan ZINB dengan menggunakan fungsi *Likelihood.*
6. Pengujian signifikansi parameter model regresi menggunakan uji *G* dan uji *Wald.*
7. Pemilihan model terbaik dengan melihat nilai AIC, AICC, dan BIC yang terkecil.

BAB IV

PEMBAHASAN

4.1 Regresi ZIP

Menurut Jansakul dan Hinde (2002), jika Y merupakan peubah respon yang memiliki sebaran *Zero Inflated Poisson*, maka nilai nol pada pengamatan (peubah respon) terjadi dalam dua tahap. Tahap pertama terjadi dengan peluang ω dan hanya menghasilkan data bernilai nol. Sedangkan tahap kedua terjadi dengan peluang $(1-\omega)$ dan memiliki sebaran Poisson dengan rata-rata λ. Regresi ZIP memiliki model dan estimsi parameter sebagai berikut:

4.1.1 Model Regresi *Zero-Inflated Poisson*(ZIP)

Lambert (1992) menyarankan hubungan model untuk λ dan ω adalah sebagaiberikut:

$$\ln(\lambda_i) = \sum_{k=0}^{p} x_{ik}\beta_k$$

$$logit(\omega_i) = \ln\left(\frac{\omega_i}{1-\omega_i}\right) = \sum_{k=0}^{p} x_{ik}\gamma_k \tag{4.1}$$

$$\ln(\lambda_i) = \sum_{k=0}^{p} x_{ik}\beta_k$$

$$\exp[\ln(\lambda_i)] = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)$$

$$\lambda_i = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) \tag{4.2}$$

$$\ln\left(\frac{\omega_i}{1-\omega_i}\right)=\sum_{k=0}^{p}x_{ik}\gamma_k$$

$$\exp\left[\ln\left(\frac{\omega_i}{1-\omega_i}\right)\right]=\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\frac{\omega_i}{1-\omega_i}=\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\omega_i=(1-\omega_i)\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\omega_i=\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)-\omega_i\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\omega_i+\omega_i\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)=\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\omega_i\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)=\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)$$

$$\omega_i=\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)} \tag{4.3}$$

Sehingga persamaan (4.2) dan (4.3) dapat ditulis sebagai persamaan berikut:

$$\lambda_i=\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\text{dan}\,\omega_i=\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)} \tag{4.4}$$

4.1.2 Estimasi Parameter ZIP

Estimasi Parameter ZIP dengan metode *Maximum Likelihood Estimation* (MLE). Metode ini dapat digunakan untuk menaksir parameter suatu model yang diketahui distribusinya. Fungsi *likelihood* persamaan (2.28) sebagai berikut:

$$L(\beta,\gamma)=\prod_{i=1}^{n}P(y_i;\beta,\gamma) \tag{4.5}$$

$$L(\beta,\gamma)=\begin{cases}\prod_{i=1}^{n}\left[\omega_i+(1-\omega_i)\exp(-\lambda_i)\right], y_i=0\\ \prod_{i=1}^{n}\left[(1-\omega_i)\exp(-\lambda_i)\left(\frac{\lambda_i^{y_i}}{y_i!}\right)\right], y_i>0\end{cases}$$

Dari persamaan (4.5) fungsi *log natural-likelihood* diperoleh dengan mengambil nilai ln dari fungsi *likelihood* dapat ditulis menjadi persamaan berikut:

$$\ln L(\beta,\gamma)=\ln\left\{\prod_{i=1}^{n}P(y_i;\beta,\gamma)\right\}$$

$$\ln L(\beta,\gamma)=\ln\begin{cases}\prod_{i=1}^{n}\left[\omega_i+(1-\omega_i)\exp(-\lambda_i)\right] & y_i=0\\ \prod_{i=1}^{n}\left[(1-\omega_i)\exp(-\lambda_i)\left(\frac{\lambda_i^{y_i}}{y_i!}\right)\right] & y_i>0\end{cases} \tag{4.6}$$

Kemudian dari persamaan (4.6) diselesaikan dengan mensubstitusikan persamaan (4.5) sebagai berikut:

Untuk $y_i = 0$

$$\ln L(\beta,\gamma)=\ln\prod_{i=1}^{n}\left[\omega_i+(1-\omega_i)\exp(-\lambda_i)\right]$$

$$=\ln\prod_{i=1}^{n}\left[\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\right)+\left(1-\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\right)\right)\right.$$

$$\left.\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right]$$

$$=\ln\prod_{i=1}^{n}\left[\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\right)+\left(\frac{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\right.\right.$$

$$\left.\left.-\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\right)\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) + \left(\frac{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \right.$$

$$\left. \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) + \left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \right.$$

$$\left. \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} + \frac{\exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right]$$

$$= \ln \prod_{i=1}^{n} \left[\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) + \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right]$$

$$= \sum_{i=1}^{n} \ln \left[\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) + \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right]$$

$$(\beta,\gamma) = \sum_{i=1}^{n} \ln\left(\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right) + \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right) - \sum_{i=1}^{n} \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right) \right) \quad (4.7)$$

Untuk $y_i > 0$

$$\ln L(\beta,\gamma) = \ln \prod_{i=1}^{n} \left[(1-\omega_i)\exp(-\lambda_i)\left(\frac{\lambda_i^{y_i}}{y_i!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(1 - \frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} - \frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \right.$$

$$\left. \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i\,!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \right.$$

$$\left. \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i\,!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i\,!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right) \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i\,!} \right) \right]$$

$$= \ln \prod_{i=1}^{n} \left[\frac{\exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right) \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right) y_i\,!} \right]$$

$$= \sum_{i=1}^{n} \ln \left[\frac{\exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right) \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right) y_i\,!} \right]$$

$$= \sum_{i=1}^{n} \ln \exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) + \sum_{i=1}^{n} \ln \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right)^{y_i} -$$

$$\sum_{i=1}^{n} \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right) \right) - \sum_{i=1}^{n} \ln y_i\,!$$

$$= \sum_{i=1}^{n} -\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) + \sum_{i=1}^{n} y_i \ln \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) -$$

$$\sum_{i=1}^{n} \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right) \right) - \sum_{i=1}^{n} \ln y_i\,!$$

$$= \sum_{i=1}^{n} -\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) + \sum_{i=1}^{n} y_i \left(\sum_{k=0}^{p} x_{ik}\beta_k \right) - \sum_{i=1}^{n} \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right) \right) - \sum_{i=1}^{n} \ln y_i\,!$$

$$= \sum_{i=1}^{n}\left(y_i \left(\sum_{k=0}^{p} x_{ik}\beta_k \right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right)\right) - \sum_{i=1}^{n} \ln\left(1 + \exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k \right)\right) - \sum_{i=1}^{n} \ln y_i ! \quad (4.8)$$

Nilai maksimal dari persamaan (4.7) dan (4.8) diperoleh dengan algoritma *Expectation Maximization* (EM). Misal variabel Y berkaitan dengan variabel indikator Z yaitu:

$$z = \begin{cases} 1, \text{jika } y_i \text{ dari keadaan nol} \\ 0, \text{jika } y_i \text{ dari keadaan Poisson} \end{cases}$$

Langkah-langkah estimasi parameter ZIP adalah yang pertama menentukan distribusi variabel Z

$$P(z_i = 1) = \omega_i \text{ dan } P(z_i = 0) = P(y_i \sim Poisson(\lambda)) = 1 - \omega_i$$

Sehingga $z_i \sim Binomial\ (1, \omega_i), E(z_i) = \omega_i$ dan $var(z_i) = \omega_i\ (1 - \omega_i)$

Setelah menentukan distribusi variabel Z kemudian membentuk distribusi gabungan y_i dan z_i sehingga didapatkan fungsi gabungan berikut:

$$\begin{aligned} f(y_i, z_i \mid \omega_i, \lambda_i) &= f(z_i)\, f(y_i \mid z_i) \\ &= f(z_i \mid 1, \omega_i)\, f(y_i \mid z_i, \lambda_i) \\ &= (1-\omega_i)^{(1-z_i)} (\omega_i)^{z_i} \left(\frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i !} \right)^{(1-z_i)} \\ &= \left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right)^{(1-z_i)} \left(\frac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)} \right)^{z_i} \\ &\quad \left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)^{y_i}}{y_i !} \right)^{(1-z_i)} \end{aligned}$$

$$
=\frac{1}{\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{(1-z_i)}}\frac{\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}{\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}
$$

$$
\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i!}\right)^{(1-z_i)}
$$

$$
=\frac{\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}{\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{(1-z_i)}\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}
$$

$$
\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i!}\right)^{(1-z_i)}
$$

$$
=\frac{\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}{\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{(1-z_i+z_i)}}\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i!}\right)^{(1-z_i)}
$$

$$
=\frac{1}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}
$$

$$
\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i!}\right)^{(1-z_i)}
$$

$$
=\frac{\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i!}\right)^{(1-z_i)} \tag{4.9}
$$

Selanjutnya dari persamaan (4.9) maka diperoleh fungsi *likelihood* berikut:

$$
L(\beta,\gamma|y_i,z_i)=\prod_{i=1}^{n}\frac{1}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}
$$

$$\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i\,!}\right)^{(1-z_i)} \tag{4.10}$$

Kemudian dari persaman (4.10) maka diperoleh ln-*likelihood* berikut:

$$\ln\prod_{i=1}^{n}(\beta,\gamma|y_i,z_i)$$

$$=\ln\prod_{i=1}^{n}\frac{1}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i\,!}\right)^{(1-z_i)}$$

$$=\sum_{i=1}^{n}\ln\frac{1}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i\,!}\right)^{(1-z_i)}$$

$$=\sum_{i=1}^{n}\ln\frac{1}{1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)}+\sum_{i=1}^{n}\ln\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)^{z_i}+$$

$$\sum_{i=1}^{n}\ln\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i\,!}\right)^{(1-z_i)}$$

$$=\sum_{i=1}^{n}\left(\ln 1-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}z_i\ln\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)+$$

$$\sum_{i=1}^{n}(1-z_i)\ln\left(\frac{\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}}{y_i\,!}\right)$$

$$=\sum_{i=1}^{n}\left(0-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}z_i\ln\left(\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)+$$

$$\sum_{i=1}^{n}(1-z_i)\ln\left(\exp\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)^{y_i}\right)-\ln y_i\,!$$

$$= \sum_{i=1}^{n}\left(-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}z_i\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)+$$

$$\sum_{i=1}^{n}(1-z_i)\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)+y_i\ln\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)-\ln y_i!\right)$$

$$= \sum_{i=1}^{n}\left(-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}z_i\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)+$$

$$\sum_{i=1}^{n}(1-z_i)\left(-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)+y_i\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)-\ln y_i!\right)$$

$$= \sum_{i=1}^{n}\left(z_i\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right)+$$

$$\sum_{i=1}^{n}\left((1-z_i)\left(y_i\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)-\sum_{i=1}^{n}(1-z_i)\ln y_i! \qquad (4.11)$$

Persamaan (4.11) dapat disebut *complete data likelihood*. Persamaan ini selanjutnya dimaksimumkan dengan algoritma EM. Bertujuan mendapt estimasi parameter β dan γ dapat diestimasi terpisah dengan menuliskan persamaan (4.11) menjadi:

$$\ln L(\beta,\gamma,y,z)=\ln L(\beta,y,z)+\ln L(\gamma,y,z)-\sum_{i=1}^{n}(1-z_i)\ln y_i!$$

Dengan

$$\ln L(\gamma,y,z)=\sum_{i=1}^{n}\left[z_i\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p}x_{ik}\gamma_k\right)\right)\right] \qquad (4.12)$$

$$\ln L(\beta,y,z)=\sum_{i=1}^{n}\left[(1-z_i)\left(y_i\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)-\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right] \qquad (4.13)$$

Sedangkan persamaan $(1-z_i)\ln y_i!$ dapat diabaikan karena tidak terdapat parameter β dan γ.

Pada tahap ekspektasi ini dengan cara dapat dilakukan dengan mengganti variabel z_i dengan $z_i^{(m)}$ yang merupakan ekspektasi dari z_i

$$z_i^{(m)} = E\left(z_i \mid y_i, \gamma^{(m)}, \beta^{(m)}\right) = P\left(z_i = 1 \mid y_i, \gamma^{(m)}, \beta^{(m)}\right)$$

$$z_i^{(m)} = \begin{cases} \dfrac{1}{1+\exp\left(-\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}, & y_i = 0 \\ 0 & y_i > 0 \end{cases} \tag{4.14}$$

Sehingga persamaan (4.12) dan (4.13) dapat ditulis menjadi

$$\ln L\left(\beta^{(m)}, y, z^{(m)}\right) = \sum_{i=1}^{n}\left[\left(1 - z_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right] \tag{4.15}$$

$$\ln L\left(\gamma^{(m)}, y, z^{(m)}\right) = \sum_{i=1}^{n}\left[z_i^{(m)}\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \ln\left(1 - \exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right)\right] \tag{4.16}$$

Pada tahap maksimalisasi dengan cara memaksimalkan β dan γ pada persamaan (4.15) dan (4.16) dengan menghitung $\beta^{(m+1)}$ dan $\gamma^{(m+1)}$ dengan metode *Newton Raphson*. Misalkan $\beta^{(m)}$ dan $\gamma^{(m)}$ adalah aproksimasi metode maksimum *likelihood* untuk $\hat{\beta}$ dan $\hat{\gamma}$. Dengan menggunakan metode *Newton Raphson* maka:

$$\beta^{(m+1)} = \beta^{(m)} - \left(H^{(m)}\right)^{-1} U^{(m)} \tag{4.17}$$

$$\gamma^{(m+1)} = \gamma^{(m)} - \left(H^{(m)}\right)^{-1} U^{(m)} \tag{4.18}$$

Dengan H adalah turunan kedua dari ln $L\left(\beta^{(m)}, y, z^{(m)}\right)$ dan ln $L\left(\gamma^{(m)}, y, z^{(m)}\right)$, U adalah turunan pertama dari ln $L\left(\beta^{(m)}, y, z^{(m)}\right)$ dan ln $L\left(\gamma^{(m)}, y, z^{(m)}\right)$.

Taksiran maksimum likelihood $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$ dan $\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_p$ sebagai berikut :

a. Untuk ln $L\left(\beta^{(m)}, y, z^{(m)}\right)$

Misalkan :

$\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) = \exp\left(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}\right)$ dan $U_j(\beta)$ turunan pertama dari ln $\mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$ terhadap $\beta_j (j = 0,1,2, \dots, p)$ maka:

$$U_p(\beta) = \frac{\partial}{\partial \beta_p} \ln \mathrm{L}\left(\beta^{(m)}, y_i, z_i^{(m)}\right)$$

$$= \frac{\partial}{\partial \beta_p}\left[\left(1 - z_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right]$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \beta_p}\left[\left(1 - z_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\frac{\partial}{\partial \beta_p}\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right) - \frac{\partial}{\partial \beta_p}\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right) - \frac{\partial}{\partial \beta_p}\left(\exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\beta_0 + \frac{\partial}{\partial \beta_p}\beta_1 x_{1i} + \ldots + \frac{\partial}{\partial \beta_p}\beta_p x_{pi}\right) - \right.$$

$$\left(\frac{\partial}{\partial \beta_p}\beta_0 \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right) + \frac{\partial}{\partial \beta_p}\beta_1 x_{1i} \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right.$$

$$\left.\left. + \ldots + \frac{\partial}{\partial \beta_p}\beta_p x_{pi} \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i x_{pi} - \left(x_{pi} \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right)\left(y_i x_{pi} - x_{pi} \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)$$

$$= \sum_{i=1}^{n}\left(1 - z_i^{(m)}\right) x_{pi}\left(y_i - \exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right) \tag{4.19}$$

b. Untuk ln $L\left(\gamma^{(m)}, y, z^{(m)}\right)$

Misalkan :

$\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) = \exp\left(\gamma_0 + \gamma_1 x_{1i} + \cdots + \gamma_p x_{pi}\right)$ dan $U_j(\gamma)$ turunan pertama dari $\ln \mathrm{L}\left(\gamma^{(m)}, y, z^{(m)}\right)$ terhadap $\gamma_j (j = 0,1,2, \dots, p)$ maka:

$$U_p(\gamma) = \frac{\partial}{\partial \gamma_p} \ln \mathrm{L}\left(\gamma^{(m)}, y, z^{(m)}\right)$$

$$= \frac{\partial}{\partial \gamma_p}\left[\sum_{i=1}^{n} z_i^{(m)}\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right)\right]$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \gamma_p}\left[z_i^{(m)}\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right) - \ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left[\frac{\partial}{\partial \gamma_p}\left(z_i^{(m)}\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right) - \frac{\partial}{\partial \gamma_p}\left(\ln\left(1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left[z_i^{(m)}\left(\frac{\partial}{\partial \gamma_p}\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right) - \frac{\partial}{\partial \gamma_p}\left(\ln\left(1+\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left[z_i^{(m)}\left(\frac{\partial}{\partial \gamma_p}\gamma_0 + \frac{\partial}{\partial \gamma_p}\gamma_1 x_{1i} + \ldots + \frac{\partial}{\partial \gamma_p}\gamma_p x_{pi}\right) - \frac{1}{\left(1+\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)}\right.$$

$$\left.\frac{\partial}{\partial \gamma_p}\left(1+\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left[z_i^{(m)} x_{pi} - \frac{1}{\left(1+\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)}\left(\frac{\partial}{\partial \gamma_p}1 + \frac{\partial}{\partial \gamma_p}\left(\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left[z_i^{(m)} x_{pi} - \frac{1}{\left(1+\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)}\left(0 + \frac{\partial}{\partial \gamma_p}\gamma_0\left(\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right) + \right.\right.$$

$$\left.\left.\frac{\partial}{\partial \gamma_p}\gamma_1 x_{1i}\left(\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right) + \ldots + \frac{\partial}{\partial \gamma_p}\gamma_p x_{pi}\left(\exp\left(\gamma_0 + \gamma_1 x_{1i} + \ldots + \gamma_p x_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[z_i^{(m)}x_{pi}-\frac{\left(x_{pi}\left(\exp\left(\gamma_0+\gamma_1x_{1i}+\ldots+\gamma_px_{pi}\right)\right)\right)}{\left(1+\exp\left(\gamma_0+\gamma_1x_{1i}+\ldots+\gamma_px_{pi}\right)\right)}\right] \tag{4.20}$$

Selanjutnya, turunan parsial kedua dari $\ln \mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$ terhadap $\beta_j(j = 0,1,2,\ldots,p)$ dan $\ln \mathrm{L}\left(\gamma^{(m)}, y, z^{(m)}\right)$ terhadap $\gamma_j(j = 0,1,2,\ldots,p)$ yaitu:

c. Untuk $\ln \mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$

Misalkan $\mathrm{H}_{jk}(\beta)$ adalah turunan parsial kedua dari $\ln \mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$ terhadap $\beta_j(j = 0,1,2,\ldots,p)$ dengan $j,k = 0,1,2,\ldots,p$ maka:

$$H_{pp}(\beta)=\frac{\partial^2}{\partial\beta_p^2}\ln \mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$$

$$=\frac{\partial}{\partial\beta_p}\left[\sum_{i=1}^{n}\left(1-z_i^{(m)}\right)x_{pi}\left(y_i-\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right]$$

$$=\sum_{i=1}^{n}\frac{\partial}{\partial\beta_p}\left[\left(1-z_i^{(m)}\right)x_{pi}\left(y_i-\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-z_i^{(m)}\right)x_{pi}\frac{\partial}{\partial\beta_p}\left(y_i-\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-z_i^{(m)}\right)x_{pi}\left(\frac{\partial}{\partial\beta_p}y_i-\left(\frac{\partial}{\partial\beta_p}\beta_0\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)+\right.\right.\right.$$

$$\left.\left.\left.\frac{\partial}{\partial\beta_p}\beta_1x_{1i}\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)+\ldots+\frac{\partial}{\partial\beta_p}\beta_px_{pi}\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-z_i^{(m)}\right)x_{pi}\left(0-\left(x_{pi}\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-z_i^{(m)}\right)x_{pi}\left(-\left(x_{pi}\exp\left(\beta_0+\beta_1x_{1i}+\ldots+\beta_px_{pi}\right)\right)\right)\right]$$

$$= -\sum_{i=1}^{n}\left[\left(1-z_i^{(m)}\right)x_{pi}^{\ 2}\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right] \tag{4.21}$$

d. Untuk ln L$\left(\gamma^{(m)}, y, z^{(m)}\right)$

Misalkan $H_{jk}(\gamma)$ adalah turunan parsial kedua dari ln L$\left(\gamma^{(m)}, y, z^{(m)}\right)$ terhadap $\gamma_j (j = 0,1,2, \ldots, p)$ dengan, $k = 0,1,2, \ldots, p$ maka:

$$H_{pp}(\gamma) = \frac{\partial^2}{\partial \gamma_p^2}\ln \mathrm{L}\left(\gamma^{(m)}, y, z^{(m)}\right)$$

$$= \frac{\partial}{\partial \gamma_p}\left[\sum_{i=1}^{n}\left[z_i^{(m)}x_{pi} - \frac{\left(x_{pi}\left(\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)}\right]\right]$$

$$= \sum_{i=1}^{n}\left[\frac{\partial}{\partial \gamma_p}\left[z_i^{(m)}x_{pi} - \frac{\left(x_{pi}\left(\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)}\right]\right]$$

$$= \sum_{i=1}^{n}\left[\frac{\partial}{\partial \gamma_p}\left(z_i^{(m)}x_{pi}\right) - \frac{\partial}{\partial \gamma_p}\left(\frac{x_{pi}\left(\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)}{1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)}\right)\right]$$

$$= \sum_{i=1}^{n}\left[0 - \left(\frac{x_{pi}^{\ 2}\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)^2}\right)\right]$$

$$= \sum_{i=1}^{n}\left[-\left(\frac{x_{pi}^{\ 2}\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)^2}\right)\right]$$

$$= -\sum_{i=1}^{n}\left(\frac{x_{pi}^{\ 2}\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 x_{1i}+\ldots+\gamma_p x_{pi}\right)\right)^2}\right) \tag{4.22}$$

Pada tahap selanjutnya mengganti $\beta^{(m)}$ dan $\gamma^{(m)}$ dengan $\beta^{(m+1)}$ dan $\gamma^{(m+1)}$ pada iterasi selanjutnya, kemudian kembali melakukan tahap ekspektasi.

Tahap ekspektasi dan tahap maksimalisasi dilakukan secara berulang-ulang untuk sampai diperoleh penaksir parameter yang konvergen $\left(\left|\beta^{(m)} - \beta^{(m+1)}\right| \leq \varepsilon \; dan \left|\gamma^{(m)} - \gamma^{(m+1)}\right| \leq \varepsilon, biasanya \; \varepsilon = 10^{-5}\right)$ dengan ε merupakan nilai error terkecil atau mendekati nol. Hasil penaksir yang diperoleh $\hat{\beta}^{(m+1)}$ pada iterasi terakhir.

4.2 Regresi ZINB

Regresi ZINB digunakan untuk memodelkan data diskrit dengan banyak nilai nol pada peubah respon (*Zero Inflation*) dan terjadi *overdispersion*. Regresi ZINB memiliki model dan estimasi parameter sebagai berikut:

4.2.1 Model Regresi *Zero-Inflated Negative Binomial* (ZINB)

Garay (2011) menyarankan hubungan model untuk λ dan ρ adalah sebagai berikut:

$$\begin{aligned} \ln(\lambda) &= \sum_{k=0}^{p} x_{ik}\beta_k, \\ logit(\rho) &= \ln\left(\frac{\rho}{1-\rho}\right) = \sum_{k=0}^{p} z_{ik}\gamma_k \end{aligned} \tag{4.23}$$

$$\ln(\lambda) = \sum_{k=0}^{p} x_{ik}\beta_k$$

$$\exp\left[\ln(\lambda)\right] = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)$$

$$\lambda = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) \tag{4.24}$$

$$\ln\left(\frac{\rho}{1-\rho}\right)=\sum_{k=0}^{p} z_{ik}\gamma_k$$

$$\exp\left[\ln\left(\frac{\rho}{1-\rho}\right)\right]=\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\frac{\rho}{1-\rho}=\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\rho=(1-\rho)\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\rho=\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)-\rho\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\rho+\rho\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)=\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\rho\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)=\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)$$

$$\rho=\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \tag{4.25}$$

Sehingga persamaan (4.24) dan (4.25) dapat ditulis sebagai persamaan berikut:

$$\lambda_i=\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) \text{dan } \rho_i=\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \tag{4.26}$$

4.2.2 Estimasi Parameter ZINB

Estimasi Parameter ZINB dengan metode *Maximum Likelihood Estimation* (MLE). Metode ini dapat digunakan untuk menaksir parameter suatu model yang diketahui distribusinya. Fungsi *likelihood* persamaan (2.40) sebagai berikut:

$$L(\beta,\gamma)=\prod_{i=1}^{n} P(y_i;\beta,\gamma)$$

$$L(\beta,\gamma)=\begin{cases}\prod_{i=1}^{n}\left[\rho_i+(1-\rho_i)\left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}}\right], y_i=0\\ \prod_{i=1}^{n}\left[(1-\rho_i)\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y_i!}\left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta\lambda_i}{1+\theta\lambda_i}\right)^{y_i}\right], y_i>0\end{cases} \tag{4.27}$$

Dari persamaan (4.27) fungsi *log natual-likelihood* diperoleh dengan mengambil nilai ln dari fungsi *likelihood* dapat ditulis menjadi persamaan berikut:

$$\ln L(\beta,\gamma)=\ln\left\{\prod_{i=1}^{n}P(y_i;\beta,\gamma)\right\}$$

$$\ln L(\beta,\gamma)=\ln\begin{cases}\prod_{i=1}^{n}\left[\rho_i+(1-\rho_i)\left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}}\right]\\ \prod_{i=1}^{n}\left[(1-\rho_i)\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y_i!}\left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}}\left(\frac{\theta\lambda_i}{1+\theta\lambda_i}\right)^{y_i}\right]\end{cases} \tag{4.28}$$

Kemudian dari persamaan (4.28) diselesaikan dengan mensubstitusikan persamaan (4.26) sebagai berikut:

Untuk $y_i = 0$

$$\ln L(\beta,\gamma)=\ln\prod_{i=1}^{n}\left[\rho_i+(1-\rho_i)\left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}}\right]$$

$$=\ln\prod_{i=1}^{n}\left[\left(\frac{\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\right)+\left(1-\left(\frac{\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\right)\right)\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right) + \left(\frac{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} - \right.\right.$$

$$\left.\left.\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right)\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right) + \left(\frac{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right) - \exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right)\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right) + \left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right)\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right]$$

$$= \sum_{i=1}^{n} \ln \left[\left(\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right) + \left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right)\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right] \tag{4.29}$$

Untuk $y_i > 0$

$$\ln L(\beta,\gamma) = \ln \prod_{i=1}^{n} \left[(1-\rho_i) \frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) y_i!} \left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}} \left(\frac{\theta\lambda_i}{1+\theta\lambda_i}\right)^{y_i} \right]$$

$$\ln L(\beta,\gamma) = \ln \prod_{i=1}^{n} \left[\left(1 - \frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}\right) \frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) y_i!} \right.$$

$$\left. \left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}} \left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{y_i} \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} - \frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \right) \frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) y_i!} \right.$$

$$\left. \left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}} \left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}\right)^{y_i} \right]$$

$$= \ln \prod_{i=1}^{n} \left[\left(\frac{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right) - \exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \right) \frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) y_i!} \right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\right]$$

$$=\ln\prod_{i=1}^{n}\left[\left(\left(\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\right)\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y_i!}\right.\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\right]$$

$$=\sum_{i=1}^{n}\ln\left[\left(\left(\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\right)\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)y_i!}\right.\right.$$

$$\left.\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\right]$$

$$=\sum_{i=1}^{n}\ln 1-\sum_{i=1}^{n}\ln\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)+\sum_{i=1}^{n}\ln\left(\frac{\Gamma\left(y_i+\frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)}\right)-\sum_{i=1}^{n}\ln\left(y_i!\right)+$$

$$\sum_{i=1}^{n}\frac{1}{\theta}\ln 1-\sum_{i=1}^{n}\frac{1}{\theta}\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)+\sum_{i=1}^{n}y_i\ln\left(\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)-$$

$$\sum_{i=1}^{n}y_i\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)$$

$$= 0 - \sum_{i=1}^{n} \ln\left(1 + \exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right) + \sum_{i=1}^{n} \ln\left(\frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)}\right) - \sum_{i=1}^{n} \ln\left(y_i!\right) + 0 -$$

$$\sum_{i=1}^{n} \frac{1}{\theta} \ln\left(1 + \theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right) + \sum_{i=1}^{n} y_i \ln\left(\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right) -$$

$$\sum_{i=1}^{n} y_i \ln\left(1 + \theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)$$

$$= -\sum_{i=1}^{n} \ln\left(1 + \exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right) + \sum_{i=1}^{n} \ln\left(\frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)}\right) - \sum_{i=1}^{n} \ln\left(y_i!\right) -$$

$$\sum_{i=1}^{n} \left(\frac{1}{\theta} + y_i\right) \ln\left(1 + \theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right) + \sum_{i=1}^{n} y_i \ln\left(\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right) \quad (4.30)$$

Nilai maksimal dari persamaan (4.29) dan (4.30) diperoleh dengan algoritma *Expectation Maximization* (EM). Misal variabel Y berkaitan dengan variabel indikator V yaitu:

$$V = \begin{cases} 1, \text{ jika } y_i \text{ dari keadaan nol} \\ 0, \text{ jika } y_i \text{ dari keadaan Poisson} \end{cases}$$

Langkah-langkah estimasi parameter ZINB adalah yang pertama menentukan distribusi variabel V

$$P(v_i = 1) = \rho_i dan\ P(v_i = 0) = P\big(y_i \sim NB(\lambda)\big) = 1 - \rho_i$$

Sehingga $v_i \sim Binomial\ (1, \rho_i), E(v_i) = \rho_i$ dan $var(v_i) = \rho_i\ (1 - \rho_i)$

Setelah menentukan distribusi variabel v kemudian membentuk distribusi gabungan y_i dan v_i sehingga didapatkan fungsi gabungan berikut:

$$f(y_i, v_i \mid \rho_i, \lambda_i) = f(v_i)\, f(y_i \mid v_i)$$

$$= f(v_i \mid 1, \rho_i)\, f(y_i \mid v_i, \lambda_i)$$

$$= (1-\rho_i)^{(1-v_i)} (\rho_i)^{v_i} \left(\frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) y_i!} \left(\frac{\theta\lambda_i}{1+\theta\lambda_i}\right)^{y_i} \left(\frac{1}{1+\theta\lambda_i}\right)^{\frac{1}{\theta}} \right)^{(1-v_i)}$$

$$= \left(\frac{1}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \right)^{(1-v_i)} \left(\frac{\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)} \right)^{v_i} \left(\left(\left(\frac{1}{\theta}\right)^{y_i} \prod_{r=0}^{y_i-1} (1+r\theta) \right) \right.$$

$$\left. \left(\frac{1}{y_i!}\right) \left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)} \right)^{y_i} \left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)} \right)^{\frac{1}{\theta}} \right)^{(1-v_i)}$$

$$= \frac{1}{\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{(1-v_i)}} \frac{\left(\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{v_i}}{\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{v_i}} \left(\left(\prod_{r=0}^{y_i-1} (1+r\theta) \right) \left(\frac{1}{y_i!}\right) \right.$$

$$\left. \left(\frac{1}{\theta}\right)^{y_i} \left(\frac{\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)} \right)^{y_i} \left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)} \right)^{\frac{1}{\theta}} \right)^{(1-v_i)}$$

$$= \frac{\left(\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{v_i}}{\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{(1-v_i)} \left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)^{v_i}} \left(\left(\prod_{r=0}^{y_i-1} (1+r\theta) \right) \left(\frac{1}{y_i!}\right) \right.$$

$$\left(\frac{1}{\theta}\right)^{y_i}(\theta)^{y_i}\left(\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)}$$

$$=\frac{\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}}{\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{(1-v_i+v_i)}}\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\right.$$

$$\left.\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)}$$

$$=\frac{\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\right.$$

$$\left.\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)}$$

$$=\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\right.$$

$$\left.\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)} \quad (4.31)$$

Selanjutnya dari persamaan (4.31) maka diperoleh fungsi *likelihood* berikut:

$$L(\beta,\gamma|y_i,v_i)=\prod_{i=1}^{n}\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\right.$$

$$\left.\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)} \tag{4.32}$$

Kemudian dari persaman (4.32) maka diperoleh *log natural -likelihood* berikut:

$$\ln\prod_{i=1}^{n}(\beta,\gamma|y_i,v)$$

$$=\ln\prod_{i=1}^{n}\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\right.$$

$$\left.\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)^{(1-v_i)}$$

$$=\sum_{i=1}^{n}\left(\ln\frac{1}{1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)}+\ln\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)^{v_i}\right)+\sum_{i=1}^{n}(1-v_i)$$

$$\ln\left(\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)\left(\frac{1}{y_i!}\right)\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)$$

$$=\sum_{i=1}^{n}\left(\ln 1-\ln\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)+v_i\ln\left(\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}(1-v_i)$$

$$\left(\ln\left(\prod_{r=0}^{y_i-1}(1+r\theta)\right)+\ln\left(\frac{1}{y_i\,!}\right)+\ln\left(\frac{\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{y_i}+\ln\left(\frac{1}{1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)}\right)^{\frac{1}{\theta}}\right)$$

$$=\sum_{i=1}^{n}\left(0-\ln\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)+v_i\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)+\sum_{i=1}^{n}(1-v_i)$$

$$\Bigg(\sum_{r=0}^{y_i-1}\ln(1+r\theta)+\ln\left(\frac{1}{y_i\,!}\right)+\ln\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)^{y_i}-\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)^{y_i}+$$

$$\ln(1)^{\frac{1}{\theta}}-\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)^{\frac{1}{\theta}}\Bigg)$$

$$=\sum_{i=1}^{n}\left(v_i\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}(1-v_i)$$

$$\Bigg(\sum_{r=0}^{y_i-1}\ln(1+r\theta)-\ln(y_i\,!)+y_i\ln\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)-y_i\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)+$$

$$\frac{1}{\theta}\ln(1)-\frac{1}{\theta}\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)\Bigg)$$

$$=\sum_{i=1}^{n}\left(v_i\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p}z_{ik}\gamma_k\right)\right)\right)+\sum_{i=1}^{n}(1-v_i)$$

$$\Bigg(\sum_{r=0}^{y_i-1}\ln(1+r\theta)-\ln(y_i\,!)+y_i\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)-y_i\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)$$

$$0-\frac{1}{\theta}\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p}x_{ik}\beta_k\right)\right)\right)\Bigg)$$

$$= \sum_{i=1}^{n} \left(v_i \left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) - \ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) \right) \right) + \sum_{i=1}^{n} (1-v_i)$$

$$\left(\sum_{r=0}^{y_i-1} \ln(1+r\theta) - \ln(y_i!) + y_i \left(\sum_{k=0}^{p} x_{ik}\beta_k \right) - \left(y_i + \frac{1}{\theta} \right) \ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right) \right)$$

$$= \sum_{i=1}^{n} \left(v_i \left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) - \ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) \right) \right) + \sum_{i=1}^{n} (1-v_i) \left(y_i \left(\sum_{k=0}^{p} x_{ik}\beta_k \right) - \right.$$

$$\left. \left(y_i + \frac{1}{\theta} \right) \ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right) \right) + \sum_{i=1}^{n} (1-v_i) \left(\sum_{r=0}^{y_i-1} \ln(1+r\theta) - \ln(y_i!) \right) \quad (4.33)$$

Persamaan (4.33) dapat disebut *complete data likelihood*. Persamaan ini selanjutnya dimaksimumkan dengan algoritma EM. Bertujuan mendapt estimasi parameter β dan γ dapat diestimasi terpisah dengan menuliskan persamaan (4.33) menjadi:

$$\ln L(\beta,\gamma,y,v) = \ln L(\beta,y,v) + \ln L(\gamma,y,v) - \sum_{i=1}^{n} (1-v_i) \left(\sum_{r=0}^{y_i-1} \ln(1+r\theta) - \ln(y_i!) \right)$$

dengan

$$\ln L(\gamma,y,v) = \sum_{i=1}^{n} \left[v_i \left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) - \ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k \right) \right) \right] \quad (4.34)$$

$$\ln L(\beta,y,v) = \sum_{i=1}^{n} \left[(1-v_i) \left(y_i \left(\sum_{k=0}^{p} x_{ik}\beta_k \right) - \left(y_i + \frac{1}{\theta} \right) \ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k \right) \right) \right) \right) \right] \quad (4.35)$$

Sedangkan persamaan $(1-v_i)\left(\sum_{r=0}^{y_i-1} \ln(1+rk) - \ln(y_i!)\right)$ dapat diabaikan karena tidak terdapat parameter β dan γ.

Pada tahap ekspektasi ini dapat dilakukan dengan cara mengganti variabel v_i dengan $v_i^{(m)}$ yang merupakan ekspektasi dari v_i

$$v_i^{(m)} = E\left(v_i \mid y_i, \gamma^{(m)}, \beta^{(m)}\right) = P\left(v_i = 1 \mid y_i, \gamma^{(m)}, \beta^{(m)}\right)$$

$$v_i^{(m)} = \begin{cases} \dfrac{1}{1+\exp\left(-\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)-\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)}, & y_i = 0 \\ 0 & y_i > 0 \end{cases} \tag{4.36}$$

Sehingga persamaan (4.34) dan (4.35) dapat ditulis menjadi

$$\ln L\left(\beta^{(m)}, y, v^{(m)}\right) = \sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)-\left(y_i+\frac{1}{\theta}\right)\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)\right)\right] \tag{4.37}$$

$$\ln L\left(\gamma^{(m)}, y, v^{(m)}\right) = \sum_{i=1}^{n}\left[v_i^{(m)}\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)\right] \tag{4.38}$$

Pada tahap maksimalisasi dengan cara memaksimalkan β dan γ pada persamaan (4.37) dan (4.38) dengan menghitung $\beta^{(m+1)}$ dan $\gamma^{(m+1)}$ dengan metode Newton-Raphson. Misalkan $\beta^{(m)}$ dan $\gamma^{(m)}$ adalah aproksimasi metode maksimum *likelihood* untuk $\hat{\beta}$ dan $\hat{\gamma}$. Dengan menggunakan metode Newton-Raphson maka:

$$\beta^{(m+1)} = \beta^{(m)} - \left(H^{(m)}\right)^{-1} U^{(m)} \tag{4.39}$$

$$\gamma^{(m+1)} = \gamma^{(m)} - \left(H^{(m)}\right)^{-1} U^{(m)} \tag{4.40}$$

Dengan H adalah turunan kedua dari ln L$\left(\beta^{(m)}, y, v^{(m)}\right)$ dan ln L$\left(\gamma^{(m)}, y, v^{(m)}\right)$, U adalah turunan pertama dari ln L$\left(\beta^{(m)}, y, v^{(m)}\right)$ dan ln L$\left(\gamma^{(m)}, y, v^{(m)}\right)$.

Taksiran maksimum *likelihood* $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$ dan $\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_p$ sebagai berikut :

a. Untuk ln L$\left(\beta^{(m)}, y, v^{(m)}\right)$

Misalkan :

$\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) = \exp\left(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}\right)$ dan $U_j(\beta)$ turunan pertama dari ln $\mathrm{L}\left(\beta^{(m)}, y, v^{(m)}\right)$ terhadap $\beta_j (j = 0,1,2,\ldots,p)$ maka:

$$U_p(\beta) = \frac{\partial}{\partial \beta_p} \ln \mathrm{L}\left(\beta^{(m)}, y, v^{(m)}\right)$$

$$= \frac{\partial}{\partial \beta_p}\left[\left(1-v_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \left(y_i + \frac{1}{\theta}\right)\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)\right)\right]$$

$$= \sum_{i=1}^{n}\frac{\partial}{\partial \beta_p}\left[\left(1-v_i^{(m)}\right)\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \left(y_i + \frac{1}{\theta}\right)\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)\right)\right]$$

$$= \sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\frac{\partial}{\partial \beta_p}\left(y_i\left(\sum_{k=0}^{p} x_{ik}\beta_k\right) - \left(y_i + \frac{1}{\theta}\right)\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right) - \frac{\partial}{\partial \beta_p}\left(y_i + \frac{1}{\theta}\right)\ln\left(1+\theta\left(\exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right) - \left(y_i + \frac{1}{\theta}\right)\right.$$

$$\left.\frac{\partial}{\partial \beta_p}\ln\left(1+\theta\left(\exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)\right)$$

$$= \sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i\left(\frac{\partial}{\partial \beta_p}\beta_0 + \frac{\partial}{\partial \beta_p}\beta_1 x_{1i} + \ldots + \frac{\partial}{\partial \beta_p}\beta_p x_{pi}\right) - \left(y_i + \frac{1}{\theta}\right)\right.$$

$$\left.\frac{1}{\left(1+\theta\left(\exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)}\left(\frac{\partial}{\partial \beta_p}\left(1+\theta\left(\exp\left(\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}\right)\right)\right)\right)\right)$$

$$=\sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i x_{pi}-\left(y_i+\frac{1}{\theta}\right)\frac{1}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right.$$

$$\left(\frac{\partial}{\partial\beta_p}1+\frac{\partial}{\partial\beta_p}\theta\beta_0\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)+\frac{\partial}{\partial\beta_p}\theta\beta_1 x_{1i}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots\right.\right.\right.$$

$$\left.\left.\left.+\beta_p x_{pi}\right)\right)+\ldots+\frac{\partial}{\partial\beta_p}\theta\beta_p x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)\right)$$

$$=\sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i x_{pi}-\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right) \tag{4.41}$$

b. Untuk ln L$\left(\gamma^{(m)}, y, v^{(m)}\right)$

Misalkan :

$\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)=\exp\left(\gamma_0+\gamma_1 z_{1i}+\cdots+\gamma_p z_{pi}\right)$ dan $U_j(\gamma)$ turunan pertama dari ln L$\left(\gamma^{(m)}, y, v^{(m)}\right)$ terhadap $\gamma_j (j=0,1,2,\ldots,p)$ maka:

$$U_p(\gamma)=\frac{\partial}{\partial\gamma_p}\ln \mathrm{L}\left(\gamma^{(m)}, y, v^{(m)}\right)$$

$$=\frac{\partial}{\partial\gamma_p}\left[\sum_{i=1}^{n} v_i^{(m)}\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)\right]$$

$$=\sum_{i=1}^{n}\frac{\partial}{\partial\gamma_p}\left[v_i^{(m)}\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)-\ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[\frac{\partial}{\partial\gamma_p}\left(v_i^{(m)}\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)-\frac{\partial}{\partial\gamma_p}\left(\ln\left(1+\exp\left(\sum_{k=0}^{p} z_{ik}\gamma_k\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[v_i^{(m)}\left(\frac{\partial}{\partial\gamma_p}\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)-\frac{\partial}{\partial\gamma_p}\left(\ln\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[v_i^{(m)}\left(\frac{\partial}{\partial\gamma_p}\gamma_0+\frac{\partial}{\partial\gamma_p}\gamma_1 z_{1i}+\ldots+\frac{\partial}{\partial\gamma_p}\gamma_p z_{pi}\right)-\frac{1}{\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)}\right.$$

$$\left.\frac{\partial}{\partial \gamma_p}\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[v_i^{(m)} z_{pi}-\frac{1}{\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)}\left(\frac{\partial}{\partial \gamma_p}1+\frac{\partial}{\partial \gamma_p}\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[v_i^{(m)} z_{pi}-\frac{1}{\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)}\left(0+\frac{\partial}{\partial \gamma_p}\gamma_0\right.\right.$$

$$\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)+\frac{\partial}{\partial \gamma_p}\gamma_1 z_{1i}\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)+\ldots$$

$$\left.\left.+\frac{\partial}{\partial \gamma_p}\gamma_p z_{pi}\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right)\right]$$

$$=\sum_{i=1}^{n}\left[v_i^{(m)} z_{pi}-\frac{\left(z_{pi}\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)}\right] \tag{4.42}$$

Selanjutnya, turunan parsial kedua dari $\ln \mathrm{L}\left(\beta^{(m)}, y, v^{(m)}\right)$ terhadap $\beta_j (j = 0,1,2, \ldots, p)$ dan $\ln \mathrm{L}\left(\gamma^{(m)}, y, v^{(m)}\right)$ terhadap $\gamma_j (j = 0,1,2, \ldots, p)$ yaitu:

c. Untuk $\ln \mathrm{L}\left(\beta^{(m)}, y, v^{(m)}\right)$

Misalkan $H_{jk}(\beta)$ adalah turunan parsial kedua dari $\ln \mathrm{L}\left(\beta^{(m)}, y, z^{(m)}\right)$ terhadap $\beta_j (j = 0,1,2, \ldots, p)$ dengan $j, k = 0,1,2, \ldots, p$ maka:

$$H_{pp}(\beta)=\frac{\partial^2}{\partial \beta_p^2}\ln \mathrm{L}\left(\beta^{(m)}, y, v^{(m)}\right)$$

$$=\frac{\partial}{\partial \beta_p}\sum_{i=1}^{n}\left(1-v_i^{(m)}\right)\left(y_i x_{pi}-\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)$$

$$=\sum_{i=1}^{n}\frac{\partial}{\partial\beta_p}\left[\left(1-v_i^{(m)}\right)\left(y_i x_{pi}-\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(\frac{\partial}{\partial\beta_p}y_i x_{pi}-\frac{\partial}{\partial\beta_p}\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(0-\left(y_i+\frac{1}{\theta}\right)\frac{\partial}{\partial\beta_p}\frac{\theta x_{pi}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(0-\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}^{2}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)\right]$$

$$=\sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(-\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}^{2}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right)\right]$$

$$=-\sum_{i=1}^{n}\left[\left(1-v_i^{(m)}\right)\left(y_i+\frac{1}{\theta}\right)\frac{\theta x_{pi}^{2}\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)}{\left(1+\theta\left(\exp\left(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}\right)\right)\right)}\right] \tag{4.43}$$

d. Untuk $\ln \mathrm{L}\left(\gamma^{(m)}, y, v^{(m)}\right)$

Misalkan $H_{jk}(\gamma)$ adalah turunan parsial kedua dari $\ln \mathrm{L}\left(\gamma^{(m)}, y, v^{(m)}\right)$ terhadap $\gamma_j (j = 0,1,2, \ldots, p)$ dengan $j, k = 0,1,2, \ldots, p$ maka:

$$H_{pp}(\gamma)=\frac{\partial^2}{\partial\gamma_p^2}\ln \mathrm{L}\left(\gamma^{(m)}, y, v^{(m)}\right)$$

$$=\frac{\partial}{\partial\gamma_p}\left[\sum_{i=1}^{n}\left[v_i^{(m)} z_{pi}-\frac{\left(z_{pi}\left(\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)\right)}{\left(1+\exp\left(\gamma_0+\gamma_1 z_{1i}+\ldots+\gamma_p z_{pi}\right)\right)}\right]\right]$$

$$= \sum_{i=1}^{n} \left[\frac{\partial}{\partial \gamma_p} \left[v_i^{(m)} z_{pi} - \frac{\left(z_{pi} \left(\exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right) \right)}{\left(1 + \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right)} \right] \right]$$

$$= \sum_{i=1}^{n} \left[\frac{\partial}{\partial \gamma_p} \left(v_i^{(m)} z_{pi} \right) - \frac{\partial}{\partial \gamma_p} \left(\frac{z_{pi} \left(\exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right)}{1 + \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right)} \right) \right]$$

$$= \sum_{i=1}^{n} \left[0 - \left(\frac{z_{pi}^{\;2} \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right)}{\left(1 + \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right)^2} \right) \right]$$

$$= \sum_{i=1}^{n} \left[- \left(\frac{z_{pi}^{\;2} \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right)}{\left(1 + \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right)^2} \right) \right]$$

$$H_{pp}(\gamma) = -\sum_{i=1}^{n} \left(\frac{z_{pi}^{\;2} \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right)}{\left(1 + \exp\left(\gamma_0 + \gamma_1 z_{1i} + \ldots + \gamma_p z_{pi} \right) \right)^2} \right) \tag{4.44}$$

Pada tahap selanjutnya mengganti $\beta^{(m)}$ dan $\gamma^{(m)}$ dengan $\beta^{(m+1)}$ dan $\gamma^{(m+1)}$ pada iterasi selanjutnya, kemudian kembali melakukan tahap ekspektasi.

Tahap ekspektasi dan tahap maksimalisasi dilakukan secara berulang-ulang untuk sampai diperoleh penaksir parameter yang konvergen $\left(\left| \beta^{(m)} - \beta^{(m+1)} \right| \leq \varepsilon \; dan \; \left| \gamma^{(m)} - \gamma^{(m+1)} \right| \leq \varepsilon, biasanya \; \varepsilon = 10^{-5} \right)$ dimana ε merupakan nilai error terkecil atau mendekati nol. Hasil penaksir yang diperoleh $\hat{\beta}^{(m+1)}$ pada iterasi terakhir.

4.3 Perbandingan ZIP dan ZINB

4.3.1 Uji Kesesuaian Distribusi Poisson

Tabel 4.1 Hasil Uji Kesesuaian Distribusi

		Y
N		33
Poisson Parameter[a,b]	Mean	.6061
Most Extreme Differences	Absolute	.033
	Positive	.033
	Negative	-.027
Kolmogorov-Smirnov Z		.190
Asymp. Sig. (2-tailed)		1.000

Uji kesesuaian distribusi Poisson pada data Angka Kematian Ibu Bersalin di Jawa Timur Tahun 2015 menggunakan Uji Kolmogorov-Smirnov dengan nilai *P value* adalah 0,6061 lebih besar dari α (0,05) sehingga dapat diambil keputusan terima H0 dan dapat ditarik kesimpulan bahwa variabel respon berdistribusi Poisson. Jadi, jumlah kematian ibu usia 20-34 tahun pada data Angka Kematian Ibu di Kabupaten Malang Tahun 2014 adalah berdistribusi Poisson.

4.3.2 Uji Multikolinearitas

Berikut merupakan hasil nilai VIF antara satu variabel prediktor dengan beberapa variabel prediktor lainnya.

Tabel 4.2 Nilai VIF dari Variabel Prediktor

Variabel	VIF
X_1	1.02
X_2	1.05
X_3	1.03
X_4	1.04

Tabel 4.2 menunjukkan bahwa nilai VIF dari masing-masing variabel prediktor memiliki nilai kurang dari 10 sehingga dapat disimpulkan bahwa tidak terdapat korelasi yang tinggi antar variabel prediktor.

4.3.3 Overdispersi

Berikut merupakan hasil nilai *deviance* yang diperoleh dari *software Minitab17* untuk mengetahui apakah data tersebut mengalami overdispersi atau tidak.

Tabel 4.3 Uji overdispersi

Test	Db	Estimate	Mean
Deviance	28	31,60880	1,12889

Nilai *deviance* dibagi dengan derajat bebas menghasilkan nilai lebih dari 1 maka data tersebut mengalami overdispersi.

4.3.4 Uji Signifikansi Parameter

Pengujian signifikansi parameter model regresi secara simultan dan parsial. Pengujian parameter secara simultan dan parsial data Angka Kematian Ibu Kabupaten Malang Tahun 2014 dengan menggunakan uji G dan uji *Wald.*

Hasil pengujian didapat nilai statistik uji G untuk ZIP adalah uji G = 25.0169 $>\chi^2 = 18{,}29$ dan untuk ZINB uji G = 25.0616 $>\chi^2 = 18{,}29$ maka diambil keputusan H_0 ditolak. Dari keputusan yang diambil, untuk untuk model regresi ZIP maupun model regresi ZINB dapat disimpulkan bahwa Persentasi K1, Persentase K4, Persentase Ibu Nifas Mendapat Vitamin A dan Persentase Persalinan Ditolong Tenaga Kesehatan secara bersama-sama dapat memberikan pengaruh yang besar terhadap Angka Kematian Ibu di Kabupaten Malang Tahun 2014.

Uji signifikansi secara parsial pada model ZIP menggunakan uji *Wald* di dalam lampiran 6 dengan dibandingkan dengan nilai t table 2,048407 maka tolak H_0 karena t hitung = 2,7992> t tablel = 2,048407 jadi ada peubah prediktor tersebut yang signifikan terhadap kasus Angka Kematian Ibu di Kabupaten Malang.

4.3.5 Pemodelan ZIP

Model regresi Zero-*Inflated* Poisson yang terbaik untuk menentukan hubungan variabel angka kematian ibu di Jawa Timur dengan faktor-faktornya yaitu:

$$\lambda_i = \exp\left(1,9529 - 0,0362x_1 - 0,0291x_2 + 0,0159x_3 + 0,0222x_4\right)$$

$$\omega_i = \frac{\exp\left(14,8054 - 0,2007x_1 - 0,5753x_2 - 0,4401x_3 + 0,8637x_4\right)}{1 + \exp\left(14,8054 - 0,2007x_1 - 0,5753x_2 - 0,4401x_3 + 0,8637x_4\right)}$$

Jika persentase kunjungan kehamilan K1 meningkat 1% maka akan menurunkan rata-rata jumlah kasus kematian ibu sebesar exp(0,0362)=1,0369 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase kujungan kehamilan K4 meningkat 1% maka akan menurunkan rata-rata jumlah kasus kematian ibu sebesar exp(0,0291)=1,0295 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase ibu nifas mendapat vitamin A meningkat 1% maka akan memberikan rata-rata jumlah kasus kematian ibu sebesar exp(0,0159)=1,0160 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase persalinan ditolong tenaga kesehatan meningkat 1% maka akan memberikan rata-rata jumlah kasus kematian ibu sebesar exp(0,0222)=1,0224 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap.

Jika persentase kunjungan kehamilan K1 meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu sebesar 1/exp(0,2007)=0,8182 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase kujungan kehamilan K4 meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu 1/exp(0,5753)=0,5625 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase ibu nifas mendapat vitamin A meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu 1/exp(0,4401)=0,6440 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase persalinan ditolong tenaga kesehatan meningkat 1% maka akan memberikan peluang jumlah kasus kematian ibu 1/exp(0,8637)=0,4216 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap.

4.3.6 Pemodelan ZINB

Model regresi Zero-*Inflated* Negative Binomial yang terbaik untuk menentukan hubungan variabel angka kematian ibu di Jawa Timur dengan faktor-faktornya yaitu:

$$\lambda_i = \exp\left(2,5845 - 0,0362x_1 - 0,0254x_2 + 0,0089x_3 + 0,0189x_4\right)$$

$$\omega_i = \frac{\exp\left(57,9854 - 0,2847x_1 - 1,1475x_2 - 0,8016x_3 + 1,2713x_4\right)}{1 + \exp\left(57,9854 - 0,2847x_1 - 1,1475x_2 - 0,8016x_3 + 1,2713x_4\right)}$$

Jika persentase kunjungan kehamilan K1 meningkat 1% maka akan menurunkan rata-rata jumlah kasus kematian ibu sebesar exp(0,0362)=1,0368 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase kujungan kehamilan K4 meningkat 1% maka akan menurunkan rata-rata jumlah kasus kematian ibu sebesar exp(0,0254)=1,0257 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase ibu nifas mendapat vitamin A meningkat 1% maka akan memberikan rata-rata jumlah kasus kematian ibu sebesar exp(0,0089)=1,0089 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase persalinan ditolong tenaga kesehatan meningkat 1% maka akan memberikan rata-rata jumlah kasus kematian ibu sebesar exp(0,0189)=1,0191 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap.

Jika persentase kunjungan kehamilan K1 meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu 1/exp(0,2847)=0,7522 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase kujungan kehamilan K4 meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu 1/exp(-1,1475)=0,3174 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase ibu nifas mendapat vitamin A meningkat 1% maka akan menurunkan peluang jumlah kasus kematian ibu 1/exp(0,8016)=0,4486 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap. Jika persentase persalinan ditolong tenaga kesehatan meningkat 1% maka akan memberikan peluang jumlah kasus kematian ibu 1/exp(1,2713)=0,2805 kali lipat dibanding sebelumnya dengan asumsi nilai variabel prediktor lain tetap.

4.3.7 Pemilihan Model Terbaik

Pemelihan model terbaik adalah berdasarkan nilai AIC,AICC, dan BIC dimana nilai AIC, AICC, dan BIC yang paling kecil adalah yang terbaik.

Tabel 4.4 Nilai AIC,AICC, dan BIC dari Distribusi

Model	**AIC**	**AICC**	**BIC**
ZIP	80,6554	90,6554	95,6205
ZINB	82,0880	94,6594	98,5498

Nilai AIC, AICC, dan BIC yang paling kecil adalah dari model ZIP jadi dapat kita simpulkan bahwa kasus Angka Kematian Ibu di Kabupaten Malang dengan dipengaruhi variabel prediktor kunjungan kehamilan K1 (X_1), kunjungan kehamilan K4 (X_2), ibu nifas mendapatkan vitamin A (X_3), dan persalinan ditolong tenaga kesehatan (X_4) menggunakan model ZIP adalah yang terbaik.

Pemilihan model terbaik dapat diintegrasikan dengan surat az-Zumar ayat 18 yang berisi tentang Allah menyeru untuk mengikuti perkataan yang paling baik. Sebagaimana dalam Tafsir Ibnu Katsir sebagai berikut:

وَٱلَّذِينَ ٱجْتَنَبُوا۟ ٱلطَّٰغُوتَ أَن يَعْبُدُوهَا وَأَنَابُوٓا۟ إِلَى ٱللَّهِ لَهُمُ ٱلْبُشْرَىٰ ۚ فَبَشِّرْ عِبَادِ ﴿١٧﴾ ٱلَّذِينَ
يَسْتَمِعُونَ ٱلْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُۥٓ ۚ أُو۟لَٰٓئِكَ ٱلَّذِينَ هَدَىٰهُمُ ٱللَّهُ ۖ وَأُو۟لَٰٓئِكَ هُمْ أُو۟لُوا۟
ٱلْأَلْبَٰبِ ﴿١٨﴾

"Dan orang-orang yang menjauhi tagut (yaitu) tidak menyembahnya dan kembali kepada Allah, bagi mereka berita gembira; sebab itu sampaikanlah berita itu kepada hamba-hamba-Ku, yang mendengarkan perkataan, lalu mengikuti apa yang paling baik di antaranya. Mereka itulah orang-orang yang telah diberi Allah petunjuk dan mereka itulah orang-orang yang mempunyai akal" (QS. Az-Zumar/39:17-18).

Abdur Rahman ibnu Zaid ibnu Aslam telah meriwayatkan dari ayahnya sehubungan dengan makna firman-Nya: Dan orang-orang yang menjauhi tagut (yaitu) tidak menyembahnya. (az-Zumar: 17) Ayat ini diturunkan berkenaan dengan Zaid ibnu Amr ibnu Nufail r.a, Abu Zar r.a, dan Salman Al-Farisi r.a.

Tetapi yang benar ayat ini mencakup mereka dan orang-orang selain mereka dari kalangan orang-orang yang menjauhi penyembahan berhala dan selalu taat menyembah Tuhan Yang Maha Pemurah. Maka merekalah orang-orang yang mendapat berita gembira dalam kehidupan dunia dan akhiratnya. Selanjutnya Allah Swt. berfirman:

{فَبَشِّرْ عِبَادِ الَّذِينَ يَسْتَمِعُونَ الْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُ}

"sebab itu sampaikanlah berita itu kepada hamba-hamba-Ku yang mendengarkan perkataan, lalu mengikuti apa yang paling baik di antaranya" (QS. Az-Zumar/39:17-18).

Yakni mereka memahaminya dan mengamalkan apa yang dipesankan olehnya, semakna dengan apa yang disebutkan di dalam firman-Nya kepada Musa a.s. ketika diberikan kitab Taurat kepadanya:

{فَخُذْهَا بِقُوَّةٍ وَأْمُرْ قَوْمَكَ يَأْخُذُوا بِأَحْسَنِهَا}

"Berpegang teguhlah kepadanya dan suruhlah kaummu berpegang kepada (perintah-perintahnya) dengan sebaik-baiknya" (QS. Al-A'raf /7:145).

Adapun firman Allah Swt.:

{أُولَئِكَ الَّذِينَ هَدَاهُمُ اللَّهُ}

"Mereka itulah orang-orang yang telah diberi Allah petunjuk" (QS. Az-Zumar/39:18).

Maksudnya, orang-orang yang mempunyai sifat ini adalah mereka yang mendapat petunjuk dari Allah di dunia dan di akhirat.

وَأُولَئِكَ هُمْ أُولُو الْأَلْبَابِ

"dan mereka itulah orang-orang yang mempunyai akal" (QS. Az-Zumar/39:18).

Yakni mempunyai akal yang sehat dan fitrah yang lurus.

Berdasarkan tafsiran di atas maka sangat jelas sekali bahwa Allah memerintahkan kepada kita untuk mendengar segala perkataan dalam hal ini bisa diartikan ilmu dan kita memilih yang paling baik. Begitu juga dalam ilmu statistik dalam memodelkan suatu data juga dipilih model yang terbaik. Karena Allah memberi akal kepada kita yakni manusia tentunya sudah selayaknya kita menggunakan akal fikiran untuk kebaikan dan berbuat sebaik-baiknya dan dapat menggunakan akal fikiran kita untuk mengetahui dan memilih mana yang *haq* maupun yang *bathil* dalam beribadah kepadaNya dan sebagai khalifah di muka bumi ini.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan di atas, dapat disimpulkan bahwa model yang terbaik adalah model regresi ZIP:

1. Estimasi parameter Regresi *Zero Inflated Poisson Regression* (ZIP) adalah

   Untuk model log : $\lambda_i = \exp\left(\sum_{k=0}^{p} x_{ik}\beta_k\right)$

   Untuk model logit : $\omega_i = \dfrac{\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}{1+\exp\left(\sum_{k=0}^{p} x_{ik}\gamma_k\right)}$

2. Model Regresi *Zero Inflated Poisson* (ZIP) pada jumlah Angka Kematian Ibu adalah

   $$\lambda_i = \exp\left(1,9529 - 0,0362x_1 - 0,0291x_2 + 0,0159x_3 + 0,0222x_4\right)$$

   dan

   $$\text{logit}\left(\omega_i\right) = 14,8054 - 0.2007X_1 - 0,5753X_2 - 0,4401X_3 + 0,8637X_4$$

5.2 Saran

Penelitian ini masih terdapat beberapa permasalahan yaitu terdapat variabel yang hasilnya kurang sesuai dengan teori. Saran untuk penelitian selanjutnya sebaiknya menambahkan variabel lain yang masih berkaitan dengan jumlah Angka Kematian Ibu.

DAFTAR RUJUKAN

Agresti, A. 2007. *An Introduction to Categorical Data Analysis.* New York: John Wiley & Sons.

Ariawan, B., Suparti & Sudarno. 2012. Pemodelan Regresi Zero Inflated Negative Binomial (ZINB) untuk Data Respon Diskrit dengan Excess Zeros. *Jurnal Gaussian*, 1(1): 55-64.

Bain, L. J. & Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*. Edisi Kedua. Belmont, California: Duxbury Press.

Bouk, M. A. 2016. Pendugaan Model Regresi Binomial Negatif dengan Metode Kemungkinan Maksimum. Skripsi Tidak dipublikasikan. Yogyakarta: Universitas Sanata Dharma Yogyakarta.

Cahyandari, R. 2014. Pengujian Overdispersi pada Model Regresi Poisson (Studi Kasus: Laka Lantas Mobil Penumpang di Provinsi Jawa Barat). *Jurnal Statistika*, 14(2):69-76.

Cameron, A.C., & Trivedi, P. K. 1998. *Regression Analysis of count data.* Cambridge: Cambridge University Press.

Departemen Agama RI. 2010a. *Al-Qur'an dan Tafsirnya Jilid I.* Jakarta: Lentera Abadi.

Departemen Agama RI. 2010b. *Al-Qur'an dan Tafsirnya Jilid II.* Jakarta: Lentera Abadi.

Departemen Kesehatan RI. 2009. *Sistem Kesehatan Nasional.* Jakarta: Departemen Kesehatan RI.

Departemen Kesehatan RI. 2011. *Analisis Kematian Ibu di Indonesia.* Jakarta: Departemen Kesehatan RI.

Dinas Kesehatan Kabupaten Malang. 2015. *Profil Kesehatan Kabupaten Malang Tahun 2015*. Jawa Timur: Dinas Kesehatan Kabupaten Malang.

Dewanti, NPP., Susilawati, M., & Srinadi, IGAM. 2016. Perbandingan Regresi Zero Inflated Poisson (ZIP) dan Regresi Zero Inflated Negative Binomial (ZINB) pada Data Overdispersion. *E-Jurnal Matematika*, 5(4):133-138.

Dona, F.R. 2015. Zero Inflated Poisson Regression untuk Memodelkan Faktor-faktor yang Mempengaruhi Terjadinya Kebakaran di Kabupaten Sidoarjo. Skripsi Tidak dipublikasikan. Malang: Universitas Negeri Malang.

Fittriyah, N., Hadi, A. F., Dewi, Y.S. 2014. Pemodelan Jumlah Kematian Akibat Difteri di Provinsi Jawa Timur dengan Regresi Binomial Negatif dan Zero-Inflated Poisson. *Prosiding Seminar Nasional Mateatika.* Jember. 201-214.

Garay, A.M. & Hashimoto, E.M. 2011.On Estimation and Influence Diagnostics for Zero Inflated Negative Binomial Regression Model. *Computational Statistics & Data Analysis*, 55(3): 1304-1318.

Ghozali, Imam. 2005, *Aplikasi Multivariate dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.

Hardin, J. W. & Hilbe, J.M. 2007. *Generalized Linear Models and Extensions.* Texas: A Stata Press Publication.

Hilbe, J.M. 2011. *Negative Binomial Regression.* New York: Cambridge University Press.

Hilbe, J.M. 2014. *Modeling Count Data.* New York: Cambridge University Press.

Hogg, R. V., & Craig, A. T. 1995. *Introduction to Mathematical Statistics (5th ed).* New Jersey: Prentice-Hall International.

Hossain, S., & Howlader, H. A. 2015. Estimation Techniques for Regression Model with Zero-Inflated Poisson Data. *International Joural of Statistics and Probability*, 4(4): 64-76.

Ismail, N. & Jemain, A. A. 2007. *Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Model.* Malaysia: Causalty Actuarial Society Forum.

Jansakul, N. & J.P. Hinde. 2002. Score Tests for Zero-Inflated Poisson Models.*Computational Statistics & Data Analysis*, 40(1):75-96.

Kartiningrum, D. E. & Nursaidah. 2013. Pemodelan Faktor yang Mempengaruhi Kematian Ibu di Provinsi Jawa Timur Menggunakan Zero Inflated Poisson Regression. *Prosiding Seminar Nasional 2013 Menuju Masyarakat Madani dan Lestari.* Majapahit. 497-508.

Lambert, D. 1992. Zero Inflated Poisson Regression, with an Application to Detect in Manufacturing.*Technometrics*, 34(1):1-14.

McCullagh, P. & Nelder, J. A. 1989. *Generalized Linier Models. Second Edition.* London: Chapman & Hall.

Myers, R.H., Dougals, C. Montgomery, G. Geoffrey Vining, & Timothy J, Robinson. 2010. *Generalized Linear Models Second Edition: with Applications in Engineering and Sciences.* New Jersey: John Wiley and Sons.

Rawlings, J. O., Pantula, S. G. & Dicky, D. A. 1998. *Aplied Regression Analysis: A Research Tool.* Edisi Kedua. Verlag New York: Springer.

Simarmata, R. T. & Ispriyanti, D. 2011. Penanganan Overdispersi pada Model Regresi Poisson Menggunakan Model Regresi Binomial Negatif. *Jurnal Media Statistika*, 4(2):95-104.

Tirta, I.M. 2009. *Analisis Regresi dengan R.* Jember: UNEJ Press.

Yesilova, A., Kaya, Y., Kaki, B., & Kasab, I. 2010. *Analysis of Plant Protection Studies with Excess Zeros Using Zero-Inflated and Negative Binomial Hardle Models.* Turkey: Gazi University.

LAMPIRAN

Lampiran 1. Data

No	Kecematan	Y	X1	X2	X3	X4
1	Tumpang	0	85.12	82.5	74.54	99.4
2	Poncokusumo	1	85.97	94.9	94.37	100
3	Jabung	0	90.41	99.6	98.28	98.9
4	Pakis	1	88.73	72.8	97.92	99.4
5	Lawang	1	89.36	91.4	88.83	97.6
6	Singosari	2	88.02	83.05	91.02	94.1
7	Karangploso	1	82.26	90.2	100	91.5
8	Dau	0	84.72	85.5	73.46	96.4
9	Pujon	0	94.61	92.2	90.58	97.8
10	Ngantang	1	94.3	86.4	83.24	94.8
11	Kasembon	1	79.08	100	90.05	99.4
12	Kepanjen	2	87.72	85.3	91.43	95.2
13	Sumber Pucung	1	84.04	88.8	71.88	98.4
14	Kromengan	0	78.76	96.8	97.81	91.7
15	Pakisaji	0	100	84.2	87.69	93.7
16	Ngajum	0	91.38	87.7	94.18	94.4
17	Wonosari	0	81.99	89.7	93.09	97.2
18	Wagir	1	93.21	93.2	97.61	92.2
19	Pagak	0	82.32	93.35	47.96	94.5
20	Donomulyo	0	94.86	77.6	91.25	99.4
21	Kalipare	1	75.31	81.2	76.78	91.6
22	Bantur	0	89.77	93.1	63.93	100
23	Gedangan	0	83.15	83.7	79.71	98.4
24	Gondanglegi	4	88.61	86.15	90.21	98.75
25	Bululawang	0	90.66	82.7	81.11	92.1
26	Wajak	1	97.03	88.3	87.83	98.3
27	Tajinan	0	63.04	80.8	85.72	97.3
28	Turen	1	93.62	88.1	91.92	75.6
29	Dampit	1	83.95	95.9	97.61	91.15
30	Sumber Manjing Wetan	0	91.19	87.15	94.54	92.05
31	Ampelgading	0	98.83	87.4	78.26	87.2
32	Tirtoyudo	0	81.78	82.9	22.27	93.9
33	Pagelaran	0	89.47	89.7	51.43	90.7

Lanjutan Lampiran 1

Keterangan:

Y : Jumlah kasus Angka Kematian Ibu di Kabpaten Malang Tahun 2014 Usia 20-34 Tahun (orang)

X_1 : Cakupan kunjungan kehamilan K1 (%)

X_2 : Cakupan kunjngan kehamilan K4 (%)

X_3 : Ibu nifas mendapatkan vitamin A (%)

X_4 : Persalinan ditolong tenaga kesehatan (%)

Lampiran 2. Uji Kolmogoov Smirnov untuk pemeriksaan sebaran

One-Sample Kolmogorov-Smirnov Test						
		Y	**X1**	**X2**	**X3**	**X4**
N		**33**	**33[c]**	**33[d]**	**33[e]**	**33[f]**
Poisson Parameter[a,b]	**Mean**	**.6061**	**98.0318**	**97.6485**	**95.7712**	**85.5706**
Most Extreme Differences	**Absolute**	**.033**				
	Positive	**.033**				
	Negative	**-.027**				
Kolmogorov-Smirnov Z		**.190**				
Asymp. Sig. (2-tailed)		**1.000**				
a. Test distribution is Poisson.						
b. Calculated from data.						

Lampiran 3. Pemeriksaan Nonmultikolinieritas dengan VIF

Coefficients

Term	Coef	SE Coef	VIF
Constant	-3.68	7.52	
X1	-0.0022	0.0312	1.02
X2	-0.0195	0.0335	1.05
X3	0.0466	0.0257	1.03
X4	0.0112	0.0484	1.04

Lampiran 4. Pemeriksaan Overdispersi

Goodness-of-Fit Tests

Test	DF	Estimate	Mean	Chi-Square	P-Value
Deviance	28	31.60880	1.12889	31.61	0.291
Pearson	28	32.12669	1.14738	32.13	0.269

Lampiran 5. *Syntax* dengan *Software* SAS Model Regresi ZIP

```
data data2;
input y x1 x2 x3 x4;
cards;
0        85.12    82.50    74.54    99.40
1        85.97    94.90    94.37    100.00
0        90.41    99.60    98.28    98.90
1        88.73    72.80    97.92    99.40
1        89.36    91.40    88.83    97.60
2        88.02    83.05    91.02    94.10
1        82.26    90.20    100      91.50
0        84.72    85.50    73.46    96.40
0        94.61    92.20    90.58    97.80
1        94.30    86.40    83.24    94.80
1        79.08    100      90.05    99.40
2        87.72    85.30    91.43    95.20
1        84.04    88.80    71.88    98.40
0        78.76    96.80    97.81    91.70
0        100      84.20    87.69    93.70
0        91.38    87.70    94.18    94.40
0        81.99    89.70    93.09    97.20
1        93.21    93.20    97.61    92.20
0        82.32    93.35    47.96    94.50
0        94.86    77.60    91.25    99.40
1        75.31    81.20    76.78    91.60
0        89.77    93.10    63.93    100
0        83.15    83.70    79.71    98.40
4        88.61    86.15    90.21    98.75
0        90.66    82.70    81.11    92.10
1        97.03    88.30    87.83    98.30
0        63.04    80.80    85.72    97.30
1        93.62    88.10    91.92    75.60
1        83.95    95.90    97.61    91.15
0        91.19    87.15    94.54    92.05
0        98.83    87.40    78.26    87.20
0        81.78    82.90    22.27    93.90
0        89.47    89.70    51.43    90.70
;
run;
proc genmod data=data2;
model y=x1 x2 x3 x4/dist=zip;
zeromodel x1 x2 x3 x4/link=logit;
run;
```

Lampiran 6. Hasil Pendugaan Regresi ZIP

The SAS System

The GENMOD Procedure

Model Information

Data Set	WORK.DATA2
Distribution	Zero Inflated Poisson
Link Function	Log
Dependent Variable	y

Number of Observations Read	33
Number of Observations Used	33

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance		60.6554	
Scaled Deviance		60.6554	
Pearson Chi-Square	23	25.0169	1.0877
Scaled Pearson X2	23	25.0169	1.0877
Log Likelihood		-25.7634	
Full Log Likelihood		-30.3277	
AIC (smaller is better)		80.6554	
AICC (smaller is better)		90.6554	
BIC (smaller is better)		95.6205	

WARNING: The relative Hessian convergence criterion of 0.0269894684 is greater than the limit of 0.0001. The convergence is questionable.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	1.9529	7.4072	-12.5650	16.4708	0.07	0.7920
x1	1	-0.0362	0.0394	-0.1134	0.0410	0.84	0.3582

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
x2	1	-0.0291	0.0348	-0.0974	0.0392	0.70	0.4042
x3	1	0.0159	0.0350	-0.0528	0.0845	0.21	0.6507
x4	1	0.0222	0.0524	-0.0804	0.1249	0.18	0.6715
Scale	0	1.0000	0.0000	1.0000	1.0000		

Note: The scale parameter was held fixed.

Analysis Of Maximum Likelihood Zero Inflation Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	14.8054	71.1151	-124.578	154.1884	0.04	0.8351
x1	1	-0.2007	0.2658	-0.7216	0.3203	0.57	0.4503
x2	1	-0.5753	1.0076	-2.5502	1.3995	0.33	0.5680
x3	1	-0.4401	0.6239	-1.6629	0.7827	0.50	0.4805
x4	1	0.8637	0.9875	-1.0719	2.7992	0.76	0.3818

Lampiran 7. *Syntax* dengan *Software* SAS Model Regresi ZINB

```
data data2;
input y x1 x2 x3 x4;
cards;
0        85.12    82.50    74.54    99.40
1        85.97    94.90    94.37    100
0        90.41    99.60    98.28    98.90
1        88.73    72.80    97.92    99.40
1        89.36    91.40    88.83    97.60
2        88.02    83.05    91.02    94.10
1        82.26    90.20    100      91.50
0        84.72    85.50    73.46    96.40
0        94.61    92.20    90.58    97.80
1        94.30    86.40    83.24    94.80
1        79.08    100      90.05    99.40
2        87.72    85.30    91.43    95.20
1        84.04    88.80    71.88    98.40
0        78.76    96.80    97.81    91.70
0        100      84.20    87.69    93.70
0        91.38    87.70    94.18    94.40
0        81.99    89.70    93.09    97.20
1        93.21    93.20    97.61    92.20
0        82.32    93.35    47.96    94.50
0        94.86    77.60    91.25    99.40
1        75.31    81.20    76.78    91.60
0        89.77    93.10    63.93    100
0        83.15    83.70    79.71    98.40
4        88.61    86.15    90.21    98.75
0        90.66    82.70    81.11    92.10
1        97.03    88.30    87.83    98.30
0        63.04    80.80    85.72    97.30
1        93.62    88.10    91.92    75.60
1        83.95    95.90    97.61    91.15
0        91.19    87.15    94.54    92.05
0        98.83    87.40    78.26    87.20
0        81.78    82.90    22.27    93.90
0        89.47    89.70    51.43    90.70
;
run;
proc genmod data=data2;
model y=x1 x2 x3 x4/dist=zinb;
zeromodel x1 x2 x3 x4/link=logit;
run;
```

Lampiran 8. Hasil Pendugaan Regresi ZINB

The SAS System

The GENMOD Procedure

Model Information

Data Set	WORK.DATA2
Distribution	Zero Inflated Negative Binomial
Link Function	Log
Dependent Variable	y

Number of Observations Read	33
Number of Observations Used	33

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance		60.0880	
Scaled Deviance		60.0880	
Pearson Chi-Square	23	25.0616	1.0896
Scaled Pearson X2	23	25.0616	1.0896
Log Likelihood		-1.79769E308	
Full Log Likelihood		-30.0440	
AIC (smaller is better)		82.0880	
AICC (smaller is better)		94.6594	
BIC (smaller is better)		98.5496	

ERROR: Error in estimation routine.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	0	2.5845	0.0000	2.5845	2.5845	.	.
x1	0	-0.0362	0.0000	-0.0362	-0.0362	.	.
x2	0	-0.0254	0.0000	-0.0254	-0.0254	.	.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
x3	0	0.0089	0.0000	0.0089	0.0089	.	.
x4	0	0.0189	0.0000	0.0189	0.0189	.	.
Dispersion	0	0.0056	0.0000	0.0056	0.0056		

Note: The negative binomial dispersion parameter was estimated by maximum likelihood.

Analysis Of Maximum Likelihood Zero Inflation Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	0	57.9854	0.0000	57.9854	57.9854	.	.
x1	0	-0.2847	0.0000	-0.2847	-0.2847	.	.
x2	0	-1.1475	0.0000	-1.1475	-1.1475	.	.
x3	0	-0.8016	0.0000	-0.8016	-0.8016	.	.
x4	0	1.2713	0.0000	1.2713	1.2713	.	.

RIWAYAT HIDUP

Kamalia Rizki Rahmawati, lahir di Malang pada tanggal 12 Maret 1995, biasa dipanggil Rahma, tinggal di Desa Belung, Kecamatan Poncokusumo, Kabupaten Malang. Anak kedua dari tiga bersaudara, putri bapak Rukin dan ibu Raudlatul Jannah.

Pendidikan dasarnya ditempuh di MI KH. Romly Tamim Belung, lulus pada tahun 2007. Setelah itu dia melanjutkan ke SMP NEGERI 1 Tumpang Kabupaten Malang dan lulus pada tahun 2010. Kemudian melanjutkan pendidikan menengah atas di MAN Gondanglegi Kabupaten Malang dan lulus pada tahun 2013. Selanjutnya, pada tahun 2013 dia menempuh kuliah di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengambil Jurusan Matematika Fakultas Sains dan Teknologi.

Selain menempuh pendidikan formal dia juga menempuh pendidikan keagamaannya di TPQ al-Ittihad Belung hingga tahun 2007. Kemudian melanjutkan pendidikannya di PP. Subulus Salam Belung hingga tahun 2010. Selama menempuh pendidikan menengah atas juga menempuh madrasah diniyah dan mondok di PP. Baitul Karim Putat Lor Gondanglegi lulus pada tahun 2013. Setelah menjadi mahasiswa, dia mengikuti pendidikan di Ma’had Sunan Ampel Al-‘Aly Universitas Islam Negeri Maulana Malik Ibrahim Malang lulus pada tahun 2014. Pengalaman organisasi tahun 2018 menjadi bendahara Pimpinan Ranting Ikatan Pelajar Putri Nahdlatul Ulama (IPPNU) Belung dan menjadi anggota Pimpinan Anak Cabang Ikatan Pelajar Putri Nahdlatul Ulama (IPPNU) Poncokusumo Kabupaten Malang.

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 : Kamalia Rizki Rahmawati
NIM : 13610082
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Perbandingan Regresi *Zero Inflated Poisson* dan *Zero Inflated Negative Binomial* pada Data *Overdispersion*
Pembimbing I : Dr. Sri Harini, M.Si
Pembimbing II : Dr. Abdussakir, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	15 Maret 2017	Konsultasi BAB I dan II	1.
2.	21 April 2017	Konsultasi BAB I, II, dan III	2.
3.	12 Mei 2017	Konsultasi Agama BAB I	3.
4.	29 Mei 2017	Konsultasi Agama BAB II	4.
5.	6 Juni 2017	Revisi Agama BAB II	5.
6.	9 Juni 2017	Konsultasi BAB III dan IV	6.
7.	4 Desember 2017	ACC untuk Seminar Proposal	7.
8.	6 Desember 2018	ACC untuk Seminar Proposal	8.
9.	22 Maret 2018	Konsultasi BAB III dan IV	9.
10.	20 April 2018	Konsultasi BAB IV dan V	10.
11.	4 Mei 2018	Revisi BAB IV dan V	11.
12.	28 Mei 2018	ACC Keseluruhan	12.
13.	25 Juni 2018	Konsultasi Agama	13.
14.	28 Juni 2018	ACC Agama Keseluruhan	14.

Malang, 28 Juni 2018
Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19770521 200501 2 004

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 : Kamalia Rizki Rahmawati
NIM : 13610082
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Perbandingan Regresi *Zero Inflated Poisson* dan *Zero Inflated Negative Binomial* pada Data *Overdispersion*
Pembimbing I : Dr. Sri Harini, M.Si
Pembimbing II : Dr. Abdussakir, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	15 Maret 2017	Konsultasi BAB I dan II	1.
2.	21 April 2017	Konsultasi BAB I, II, dan III	2.
3.	12 Mei 2017	Konsultasi Agama BAB I	3.
4.	29 Mei 2017	Konsultasi Agama BAB II	4.
5.	6 Juni 2017	Revisi Agama BAB II	5.
6.	9 Juni 2017	Konsultasi BAB III dan IV	6.
7.	4 Desember 2017	ACC untuk Seminar Proposal	7.
8.	6 Desember 2018	ACC untuk Seminar Proposal	8.
9.	22 Maret 2018	Konsultasi BAB III dan IV	9.
10.	20 April 2018	Konsultasi BAB IV dan V	10.
11.	4 Mei 2018	Revisi BAB IV dan V	11.
12.	28 Mei 2018	ACC Keseluruhan	12.
13.	25 Juni 2018	Konsultasi Agama	13.
14.	28 Juni 2018	ACC Agama Keseluruhan	14.

Malang, 28 Juni 2018
Mengetahui,
Ketua Jurusan Matematika

Dr. Usman Pagalay, M.Si
NIP. 19770521 200501 2 004