



PERBANDINGAN MODEL *NEGATIVE BINOMIAL* DAN *ZERO INFLATED NEGATIVE BINOMIAL* UNTUK MENGATASI *OUTLIER* DAN *EXCESS ZEROS* PADA *ROW COLUMN INTERACTION MODELS*

SKRIPSI

Oleh

**Rian Rizky Fauzy
NIM 1118101013**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015**



PERBANDINGAN MODEL *NEGATIVE BINOMIAL* DAN *ZERO INFLATED NEGATIVE BINOMIAL* UNTUK MENGATASI *OUTLIER* DAN *EXCESS ZEROS* PADA *ROW COLUMN INTERACTION MODELS*

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat untuk menyelesaikan Program Studi Matematika (S1) dan mencapai gelar Sarjana Sains

Oleh

Rian Rizky Fauzy
NIM 1118101013

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

1. Allah SWT yang telah memberikan kehidupan sempurna ini;
2. Bapak Sutrisno dan Ibunda Pertiwi, yang telah memberikan doa, cinta dan kasih sayangnya dalam perjalanan hidupku;
3. guru-guru sejak taman kanak-kanak sampai perguruan tinggi, yang telah memberikan ilmu dan membimbing dengan penuh kesabaran;
4. Almamater Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

MOTTO

Rasulullah SAW. Bersabda, “janganlah kalian menuntut ilmu untuk membanggakan terhadap para ulama’ dan untuk diperdebatkan di kalangan orang-orang bodoh dan buruk perangainya. Jangan pula menuntut ilmu untuk penampilan dalam majelis (pertemuan atau rapat) dan untuk menarik perhatian orang-orang kepadamu. Barangsiapa seperti itu maka baginya neraka...neraka”.

(HR. Attirmidzi dan Ibnu Majah)¹⁾

Manusia, seribu orang di antara mereka seperti satu orang
Seseorang seperti seribu orang apabila memiliki keahlian.

(‘Aidh Al-Qarni)²⁾

¹⁾ Gus AA. 2009. *Matematika Al-Quran (mengungkap M’jizat dengan Bahasa Angka)*.

Surakarta: Rahma Medika Pustaka

²⁾ Alqarni, ‘Aidh. 2010. *Detik-detik usia yang Paling Mahal*. Yogyakarta: Bening

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

nama : Rian Rizky Fauzy

NIM : 111810101013

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Perbandingan Model *Negative Binomial* dan *Zero Inflated Negative Binomial* Untuk Megatasi *Outlier* dan *Excess Zeros* pada *Row Column Interaction Models*” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, Juni 2015

Yang menyatakan,

Rian Rizky Fauzy

NIM 111810101013

SKRIPSI

PERBANDINGAN MODEL *NEGATIVE BINOMIAL* DAN *ZERO INFLATED NEGATIVE BINOMIAL* UNTUK MENGATASI *OUTLIER* DAN *EXCESS ZEROS* PADA *ROW COLUMN INTERACTION MODELS*

Oleh

Rian Rizky Fauzy
NIM 111810101013

Pembimbing:

Dosen Pembimbing Utama : Dr. Alfian Futuhul Hadi, S.Si., M.Si

Dosen Pembimbing Anggota : Dian Anggraeni, S.Si., M.Si

PENGESAHAN

Skripsi berjudul “Perbandingan Model *Negative Binomial* dan *Zero Inflated Negative Binomial* Untuk Mengatasi *Outlier* dan *Excess Zeros* Pada *Row Column Interaction Models*” telah diuji dan disahkan pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Jember.

Tim Penguji:

Ketua,

Sekretaris,

Dr. Alfian Futuhul Hadi, S.Si., M.Si
NIP. 19740719 200012 1 001

Dian Anggraeni, S.Si., M.Si
NIP. 19820216 200604 2 002

Penguji I,

Penguji II,

Prof. Drs. I Made Tirta, M.Sc., Ph.D
NIP. 19591220 198503 1 002

Ika Hesti Agustin, S.Si., M.Si
NIP. 19840801 200801 2 006

Mengesahkan

Dekan,

Prof. Drs. Kusno, DEA., Ph.D.
NIP 19610108 198602 1 001

RINGKASAN

Perbandingan Model *Negative Binomial* dan *Zero Inflated Negative Binomial* Untuk Megatasi *Outlier* dan *Excess Zeros* Pada *Row Column Interaction Models*;
Rian Rizky Fauzy; 111810101013; 2015; 54 Halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Distribusi *Zero Inflated Negative Binomial* (ZINB) merupakan metode yang dibentuk dari distribusi *Negative Binomial*. Metode ini mengasumsikan bahwa data cacah yang memiliki pengamatan nol di dalamnya terbagi menjadi 2 bagian. Bagian pertama disebut *zero state*, terjadi dengan probabilitas ω dan bagian kedua adalah pengamatan selain nol yang bersebar secara *Negative Binomial* dengan peluang $(1-\omega)$.

Pendugaan data pada penelitian model aditif dengan menggunakan distribusi ZINB cukup sering dijumpai, namun untuk penelitian model multiplikatif masih jarang digunakan. Salah satu model yang digunakan untuk model multiplikatif adalah model AMMI. AMMI merupakan suatu teknik analisis data percobaan dua faktor perlakuan dengan pengaruh utama perlakuan bersifat aditif sedangkan untuk pengaruh interaksi dimodelkan dengan model multiplikatif (bilinier). Untuk mengatasi data yang mempunyai respon tidak normal, telah dikembangkan model multiplikatif *Generalized AMMI* (GAMMI) sebagai perluasan dari AMMI. Seiring dengan berkembangnya penggunaan model GAMMI ini, muncul suatu metode baru yang dapat memberikan pemodelan layaknya model GAMMI. Model tersebut adalah RCIM (*Row-Column Interaction Model*). RCIM memberikan pendugaan model interaksi dengan menggunakan konsep RR-VGLM (*Reduced Rank Vector Generalized Linear Model*) dimana pendugaan dilakukan dengan menggunakan

metode *reduced rank regression*. RCIM dapat memberikan pendekatan pada model GAMMI disebabkan oleh kesamaan dalam proses pendugaannya yaitu dengan menggunakan metode pendugaan bolak balik.

Penelitian ini bertujuan untuk membandingkan model *Negative Binomial* dengan model *Zero Inflated Negative Binomial* pada *Row-Column Interaction Model*. Data yang digunakan dalam penelitian ini adalah data cacah yang mengalami kondisi *outlier* dan *excess zeros* dengan menggunakan 12 galur dan 5 lokasi dalam percobaannya, data ini cukup menarik karena terdapat angka nol penuh pada satu baris serta data cacah yang juga mengalami kondisi *outlier* dan *excess zeros* yang menggunakan 15 galur dan 8 lokasi dalam percobaannya. Hasilnya adalah pada data cacah yang mengalami *excess zeros* dan terdapat *outlier* serta terdapat angka nol penuh pada satu baris, MSE RCIM *Negative Binomial* dapat memberikan pendugaan yang lebih baik daripada MSE RCIM ZINB, namun RCIM1 ZINB dengan 1 pengaruh interaksi sudah cukup untuk menduga data sedangkan RCIM *Negative Binomial* tidak signifikan pada semua RCIM. Sedangkan pada data yang mengandung *outlier* dan *excess zeros*, RCIM ZINB dapat memberikan pendugaan yang lebih baik dan lebih kecil daripada RCIM *Negative Binomial*. Hal ini dapat dilihat dari besar MSE RCIM ZINB lebih kecil dan RCIM1 ZINB dengan 1 pengaruh interaksi sudah cukup untuk menduga data sedangkan RCIM *Negative Binomial* harus menggunakan RCIM3 *Negative Binomial* untuk menduga data.

Selanjutnya peneliti menggunakan data cacah berdistribusi *Negative Binomial* menggunakan 15 galur dan 8 lokasi dengan pengamatan nol dalam percobaannya. Prosesnya secara bertahap dengan mengisikan setiap kolom yang memiliki nilai pengamatan terendah dengan pengamatan nol hingga setiap kolom terdapat pengamatan nol. Pengamatan nol ini dimulai dari prosentase 8,33% hingga 20%. Data ke-1 memiliki 10 pengamatan nol, data ke-2 memiliki 11 pengamatan nol, dan begitu seterusnya hingga data ke-15. Kemudian dilakukan pendugaan pada masing-masing data dengan pendugaan RCIM ZINB dan RCIM *Negative Binomial*. Hasilnya adalah RCIM ZINB dapat memberikan MSE yang cukup stabil dan memiliki tren menurun.

Selain itu nilai MSE pada RCIM *Negative Binomial* cenderung tidak stabil pada setiap modelnya. Hal ini memunculkan keraguan dalam pendugaan dengan menggunakan RCIM *Negative Binomial* pada data cacah dalam keadaan *excess zeros*.



PRAKATA

Alhamdulillah, puji syukur kehadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga tugas akhir yang berjudul “Perbandingan Model *Negative Binomial* dan *Zero Inflated Negative Binomial* Untuk Megatasi *Outlier* dan *Excess Zeros* pada *Row Column Interaction Models*” dapat terselesaikan dengan baik. Tugas akhir ini disusun untuk memenuhi syarat menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan skripsi ini tidak terlepas dari bantuan berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Dr. Alfian Futuhul Hadi, S. Si, M. Si selaku Dosen Pembimbing Utama dan Dian Anggraeni, S. Si, M. Si selaku Dosen Pembimbing Anggota yang telah memberikan bimbingan secara intensif dan bantuan untuk penyempurnaan skripsi ini;
2. Prof. Drs. I Made Tirta, M. Sc., Ph. D dan Ika Hesti Agustin, S. Si, M. Si. Selaku Dosen Penguji yang telah memberikan kritik dan saran yang membangun untuk penyempurnaan skripsi ini;
3. seluruh dosen dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah memberikan ilmu serta fasilitas yang membantu selama proses perkuliahan berlangsung;
4. sahabati tercinta Dewi Evianti yang telah memberikan dukungan serta motivasi dalam pembuatan skripsi ini;
5. sahabat Feri, Sahabat Ilyas, Sahabat Sholeh, Sahabati Retna, Sahabati Sari, dan seluruh Sahabat-sahabati PMII Rayon FMIPA yang tidak bisa saya sebut satu per

satu yang selalu senantiasa menemani, memberi dukungan, semangat perjuangan, serta saran dalam proses menyelesaikan tugas akhir;

6. para akademisi (Yuli, Haki, Darul, Rafi, Ivan, Hendri, Emil, Feri, Jefri) yang senantiasa memberi dukungan dalam proses menyelesaikan tugas akhir;
7. Miranda Diyah Paramudhita, Mochamad Andi Saputra, Ganda Arif Sujendro, Yuwan Setiawan Wibowo, Rahmawati Fitri dan Indah Arum yang senantiasa memberi dukungan dalam hal apapun;
8. Keluarga besar KRAMAT '11 yang selalu memberikan dukungan dalam hal apapun.

Penulis menyadari bahwa dalam menyusun skripsi ini masih terdapat kekurangan baik isi maupun susunannya. Oleh karena itu, penulis mengharapkan saran dan kritik demi penyempurnaan skripsi ini. Akhirnya penulis berharap semoga skripsi ini dapat memberi manfaat dan sumbangan bagi pembaca.

Jember, Juni 2015

Rian Rizky Fauzy

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN	iv
HALAMAN PEMBIMBINGAN	v
HALAMAN PENGESAHAN	vi
RINGKASAN	vii
PRAKATA	x
DAFTAR ISI	xii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	4
BAB 2. TINJAUAN PUSTAKA	5
2.1 Generalized Linear Model	5
2.2 Overdispersi	6
2.3 Outlier	6
2.4 Regresi <i>Negative Binomial</i>	7
2.5 Excess Zeros	8
2.5.1 <i>Distribusi Zero-Inflated Negative Binomial (ZINB)</i>	8

2.5.2 Regresi <i>Zero-Inflated Negative Binomial</i> (ZINB).....	9
2.6 RR-VGLM dan <i>Row Column Interaction Model</i> (RCIM)	10
2.7 Uji Kelayakan Model	11
2.7.1 Devians.....	11
2.7.2 <i>Mean Square Error</i> (MSE)	12
2.8 Analisis Biplot.....	13
BAB 3. METODOLOGI PENELITIAN	16
3.1 Desain Penelitian	16
3.2 Data Penelitian dan Skenario Simulasi	18
3.2.1 Data Penelitian.....	19
3.2.2 Skenario Simulasi	19
3.3 Struktur Fungsi Program R.....	20
BAB 4. HASIL DAN PEMBAHASAN	22
4.1 Hubungan <i>Negative Binomial</i> dan ZINB	22
4.2 Karakteristik Data Skenario.....	23
4.3 Pengaruh Outlier pada Model <i>Negative Binomial</i> dan ZINB	24
4.4 Pengaruh <i>Excess Zeros</i> pada Model <i>Negative Binomial</i> dan ZINB .	25
4.5 Pengaruh Outlier dan <i>Excess Zeros</i> pada <i>Negative Binomial</i> dan ZINB	35
4.6 Pengaruh Pengamatan Nol Pada Model <i>Negative Binomial</i>	46
4.7 Analisis Biplot RCIM	49
BAB 5. PENUTUP	51
5.1 Kesimpulan.....	51
5.2 Saran.....	52
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR TABEL

	Halaman
4.1 Nilai <i>Log-Likelihood</i> skenario 2 dengan pendekatan rcim.....	25
4.2 Nilai <i>Log-Likelihood</i> Data 2 dengan pendekatan rcim.....	26
4.3 Nilai Model Devians dan Analisis Residu Devians Pada Skenario 2 .	27
4.4 Nilai Model Devians dan Analisis Residu Devians Pada Data 2.....	27
4.5 Analisis Residu Devians RCIM6 (Skenario 2, <i>Negative Binomial</i>) ...	28
4.6 Analisis Residu Devians RCIM4 (Skenario 2, <i>Negative Binomial</i>) ...	28
4.7 Analisis Residu Devians RCIM3 (Skenario 2, <i>Negative Binomial</i>) ...	29
4.8 Analisis Residu Devians RCIM2 (Skenario 2, <i>Negative Binomial</i>) ...	29
4.9 Analisis Residu Devians RCIM1 (Skenario 2, <i>Negative Binomial</i>) ...	29
4.10 Analisis Residu Devians RCIM6 (Skenario 2, ZINB)	30
4.11 Analisis Residu Devians RCIM5 (Skenario 2, ZINB)	30
4.12 Analisis Residu Devians RCIM4 (Skenario 2, ZINB)	31
4.13 Analisis Residu Devians RCIM3 (Skenario 2, ZINB)	31
4.14 Analisis Residu Devians RCIM2 (Skenario 2, ZINB)	31
4.15 Analisis Residu Devians RCIM1 (Skenario 2, ZINB)	32
4.16 Analisis Residu Devians RCIM3 (Data 2, <i>Negative Binomial</i>).....	32
4.17 Analisis Residu Devians RCIM2 (Data 2, <i>Negative Binomial</i>).....	33
4.18 Analisis Residu Devians RCIM1 (Data 2, <i>Negative Binomial</i>).....	33
4.19 Analisis Residu Devians RCIM3 (Data 2, ZINB).....	34
4.20 Analisis Residu Devians RCIM2 (Data 2, ZINB).....	34
4.21 Nilai <i>Mean Square Error</i> Skenario 2 dengan pendekatan rcim.....	35
4.22 Nilai <i>Mean Square Error</i> Data 2 dengan pendekatan rcim.....	35
4.23 Nilai <i>Log-Likelihood</i> skenario 3 dengan pendekatan rcim.....	36

4.24	Nilai <i>Log-Likelihood</i> skenario 4 dengan pendekatan <i>rcim</i>	36
4.25	Nilai Model Devians dan Analisis Residu Devians Pada Skenario 3 .	37
4.26	Nilai Model Devians dan Analisis Residu Devians Pada Skenario 3 .	38
4.27	Analisis Residu Devians RCIM6 (Skenario 3, <i>Negative Binomial</i>)....	39
4.28	Analisis Residu Devians RCIM3 (Skenario 3, <i>Negative Binomial</i>)....	39
4.29	Analisis Residu Devians RCIM2 (Skenario 3, <i>Negative Binomial</i>)....	39
4.30	Analisis Residu Devians RCIM1 (Skenario 3, <i>Negative Binomial</i>)....	40
4.31	Analisis Residu Devians RCIM6 (Skenario 3, <i>ZINB</i>)	41
4.32	Analisis Residu Devians RCIM4 (Skenario 3, <i>ZINB</i>)	41
4.33	Analisis Residu Devians RCIM3 (Skenario 3, <i>ZINB</i>)	41
4.34	Analisis Residu Devians RCIM2 (Skenario 3, <i>ZINB</i>)	42
4.35	Analisis Residu Devians RCIM1 (Skenario 3, <i>ZINB</i>)	42
4.36	Analisis Residu Devians RCIM3 (Skenario 4, <i>Negative Binomial</i>)....	43
4.37	Analisis Residu Devians RCIM2 (Skenario 4, <i>Negative Binomial</i>)....	43
4.38	Analisis Residu Devians RCIM1 (Skenario 4, <i>Negative Binomial</i>)....	43
4.39	Analisis Residu Devians RCIM2 (Skenario 4, <i>ZINB</i>)	44
4.40	Analisis Residu Devians RCIM1 (Skenario 4, <i>ZINB</i>)	44
4.41	Nilai <i>Mean Square Error</i> Skenario 3 dengan pendekatan <i>rcim</i>	45
4.42	Nilai <i>Mean Square Error</i> Skenario 4 dengan pendekatan <i>rcim</i>	45

DAFTAR GAMBAR

	Halaman
2.1 Plot Pemetaan Hasil Analisis Biplot	14
3.1 Diagram Metode Penelitian	18
4.1 Grafik Tren Penurunan/Kenaikan MSE RCIM.....	47
4.2 MSE ZINB dan <i>Negative Binomial</i>	48
4.3 Biplot RCIM2 Data Simulasi Polong Isi Kedelai (73,94%).....	50

DAFTAR LAMPIRAN

	Halaman
A. DATA PENELITIAN	55
B. Script Data Serangan Hama Terhadap Tanaman Kacang Hijau (<i>Negative Binomial</i>).....	58
C. Script Data Serangan Hama Terhadap Tanaman Kacang Hijau (<i>ZINB</i>).....	60
D. Script Data Jumlah Polong Isi Kedelai dengan <i>Excess Zeros</i> (<i>Negative Binomial</i>).....	62
E. Script Data Jumlah Polong Isi Kedelai dengan <i>Excess Zeros</i> (<i>ZINB</i>).....	65
F. Script Biplot Data Jumlah Polong Isi Kedelai dengan <i>Excess Zeros</i> (<i>ZINB</i>).....	67

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Metode utama yang dikembangkan untuk memahami distribusi dari data cacah (*count*) adalah distribusi Poisson, dan kemudian menjadi metode 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 equidispersi yaitu nilai mean dan varians dari variabel respon sama. Ketika nilai varians lebih besar dari mean disebut overdispersi, sedangkan ketika nilai varians lebih kecil dari nilai meannya disebut underdispersi (McCullagh & Nelder, 1989). Salah satu metode yang dapat digunakan untuk mengatasi overdispersi pada distribusi Poisson adalah distribusi *Negative Binomial*, karena distribusi *Negative Binomial* tidak mengharuskan nilai varians sama dengan meannya (Pradawati, 2013).

Distribusi *Negative Binomial* dapat mengatasi overdispersi pada *outlier* maupun pada saat varians lebih besar dari mean. Tidak menutup kemungkinan pada *Negative Binomial* selain data dalam keadaan overdispersi karena terdapat *outlier* juga diperparah dengan kasus *excess zeros*. Dalam beberapa penelitian tertentu, angka nol dapat berarti sangat penting, sehingga tidak dapat diabaikan dalam proses analisisnya. Hilbe (2011) memperkenalkan model distribusi *Zero-Inflated Negative Binomial* yang merupakan perluasan dari *Negative Binomial* untuk mengatasi data dengan overdispersi dan data cacahan nol berlebih atau disebut *excess zeros*. Ariawan (2012) merupakan salah satu yang meneliti *Zero-Inflated Negative Binomial* dengan model aditif untuk data respon diskrit dengan *excess zeros*.

Pendugaan data pada penelitian model aditif cukup sering dijumpai, namun untuk penelitian model multiplikatif masih jarang digunakan. Salah satu model yang

digunakan untuk model multiplikatif adalah model AMMI. AMMI merupakan suatu teknik analisis data percobaan dua faktor perlakuan dengan pengaruh utama perlakuan bersifat aditif sedangkan untuk pengaruh interaksi dimodelkan dengan model multiplikatif (bilinier) (Mattjik *et al*, 2011). Untuk mengatasi data yang mempunyai respon tidak normal, Van Eeuwijk (1995) memperkenalkan model multiplikatif *Generalized AMMI* (GAMMI) sebagai perluasan dari AMMI.

Yee dan Hadi (2014) memperkenalkan sebuah model yang disebut *Row Column Interaction Model* (RCIM) yang dapat menganalisis tabel dua arah dengan matriks responnya bisa dalam bentuk cacahan (*count*), variabel respon Y pada RCIM berupa angka dan variabel prediktor X pada RCIM berupa baris ke-1 hingga ke-*i* dan dari kolom ke-1 hingga kolom ke-*j*. Model ini merupakan bagian dari *Reduced-Rank Vector Generalized Linear Model* (RR-VGLM) yang mengaplikasikan konsep *Reduced Rank Regression* kedalam jenis kelas VGLM.

Ahadiyah (2015) telah meneliti bahwa model GAMMI terlihat identik dengan *Row Column Interaction Model* (RCIM). Hasil pengepasan model GAMMI dengan menggunakan pendekatan *Row Column Interaction Model* sama dengan pengepasan model GAMMI dengan menggunakan regresi bolak-balik berdasarkan Eeuwijk (1995) terlihat dari hasil penelitian pada data Hadi (2010). Pendekatan RCIM juga mampu menyediakan analisis devian dari nilai *log-Likelihood*. Wicaksono (2015) telah meneliti model GAMMI dengan data pada kasus *excess zeros* dengan menggunakan metode *Zero-Inflated Poisson* (ZIP). Hasil nilai MSE dari masing-masing model GAMMI menunjukkan bahwa fungsi distribusi *Zero-Inflated Poisson* lebih baik dalam mengepaskan data yang mengalami kasus *zero-inflated* dari pada dengan menggunakan fungsi distribusi Poisson. MSE adalah *Mean Square Error* atau dapat diartikan sebagai *error* dari masing-masing model. Sehingga untuk model yang terbaik akan memiliki nilai MSE yang lebih kecil. Dalam penelitiannya, Wicaksono (2015) menyarankan data dengan kasus *excess zeros* dengan menggunakan metode *Zero Inflated Negative Binomial*.

Oleh karena itu, berdasarkan penjelasan diatas penelitian ini menggunakan data cacahan yang mengalami kondisi overdispersi yang disebabkan *outlier* dan *excess zeros*, sehingga digunakan model RCIM dengan metode *Zero-Inflated Negative Binomial*. Untuk memudahkan proses perhitungannya, maka digunakan alat bantu yaitu Program Rstudio dan R i386 3.1.0.

1.2 Rumusan Masalah

Sesuai dengan konsep *Reduced Rank Regression* penelitian ini membandingkan Regresi *Zero Inflated Negative Binomial* dan Regresi *Negative Binomial* pada *Row Column Interaction Model* (RCIM) sebagai berikut:

- a. Mengevaluasi model terbaik antara Regresi ZINB dengan *Negative Binomial* melalui tingkat kompleksitas model yang lebih sederhana dan memiliki nilai *Mean Square Error* yang mendekati nol.
- b. Seberapa kecil persentase angka nol pada data sehingga analisis data dengan menggunakan Regresi *Negative Binomial* menjadi tidak baik lagi dan harus beralih menggunakan Regresi *Zero Inflated Negative Binomial*.
- c. Seberapa besar pengaruh data cacahan yang terdapat *outlier* setelah dimodifikasi dengan pengamatan nol, dan seberapa besar pengaruh data cacahan yang mengandung *excess zeros* setelah dimodifikasi dengan *outlier*.

1.3 Tujuan Penelitian

Adapun tujuan dalam penelitian ini adalah:

- a. mendapatkan model regresi terbaik antara Regresi *Zero-Inflated Negative Binomial* dan Regresi *Negative Binomial* pada *Row Column Interaction Model* (RCIM) melalui tingkat kompleksitas model yang lebih sederhana serta memiliki nilai *Mean Square Error* yang mendekati nol
- b. Mengetahui seberapa kecil persentase angka nol pada data sehingga analisis data dengan menggunakan Regresi *Negative Binomial* menjadi tidak baik lagi dan harus beralih menggunakan Regresi *Zero-Inflated Negative Binomial*.

- c. Mengetahui seberapa besar pengaruh data cacahan yang terdapat outlier setelah dimodifikasi dengan pengamatan nol, dan seberapa besar pengaruh data cacahan yang mengandung *excess zeros* setelah dimodifikasi dengan *outlier*.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini ialah menambah pengetahuan mengenai analisis regresi terutama tentang Regresi *Zero-Inflated Negatif Binomial* dengan pendekatan *Row Column Interaction Model* (RCIM) sebagai salah satu alternatif dalam menduga data yang mengalami data dalam keadaan overdispersi karena *outlier* dan *excess zeros*.

BAB 2. TINJAUAN PUSTAKA

2.1 Generalized Linear Model

Pada dasarnya kita dapat memodelkan hampir semua jenis data hasil pengamatan. 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 GLM (*Generalized Linear Models*). *Generalized Linear Model* (GLM) merupakan perluasan dari proses pemodelan linier untuk pemodelan data yang mengikuti distribusi.

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

- a. Komponen distribusi, yaitu y berdistribusi keluarga eksponensial;
- b. Komponen prediktor X dan koefisien β membentuk predictor linier, yaitu $\eta = x^T \beta$;
- c. 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 $E(Y) > Var(Y)$
- 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 mengatasi adanya kasus overdispersi dalam regresi Poisson adalah dengan mengganti asumsi distribusi Poisson dengan distribusi lain yang lebih fleksibel.

2.3 Outlier

Menurut Hawkins (1980), *outlier* adalah pengamatan yang menyimpang jauh dari pengamatan lain yang menimbulkan kecurigaan bahwa pengamatana itu dihasilkan oleh penelitian yang berbeda. Contoh penggunaan *outlier* dalam kehidupan sehari-hari adalah pembelian dari sebuah kartu kredit biasanya berubah saat kartu dicuri. Pengambilan uang dalam skala besar dianggap sebagai *outlier* dalam pengambilan uang pada kartu kredit. Suatu pengamatan dianggap *outlier* jika pengamatan tersebut bernilai lebih besar dari rata-rata ditambahkan dengan tiga atau empat kali standar deviasi (Kriegel,2010). Secara matematis pengamatan dianggap *outlier* jika :

$$y_i \geq \bar{y} + 3.stdev(y_i)$$

$$y_i \leq \bar{y} - 3.stdev(y_i)$$

2.4 Regresi *Negative Binomial*

Percobaan *Negative Binomial* terdiri atas beberapa usaha dan tiap usaha dengan dua kemungkinan hasil yang dapat diberi nama sukses atau gagal dan dilakukan sampai tercapai sejumlah sukses tertentu. Model Regresi Binomial Negatif memiliki kegunaan yang sama dengan model Regresi Poisson yaitu untuk menganalisis hubungan antara suatu variabel respon data *count* dengan satu atau lebih variabel acak penjelas, tetapi model Regresi Binomial Negatif lebih fleksibel dibandingkan dengan model Poisson karena mean dan variansi dari model Binomial Negatif tidak harus sama. Model regresi *Negative Binomial* yang dibangun memiliki sebaran *Negative Binomial* dengan parameter μ dan α . Sehingga sebaran Y_i menjadi:

$$f(y; \mu, \alpha) = \frac{\Gamma(y_i + 1/\alpha)}{y_i! \Gamma(1/\alpha)} \left(\frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha} \left(\frac{\mu_i}{1/\alpha + \mu_i} \right)^{y_i}$$

Distribusi *Negative Binomial* $f(y; \mu, \alpha)$ mempunyai rataan dan variansi

$$E(Y_i) = \mu_i$$

$$Var(Y_i) = \mu_i + \alpha \mu_i^2$$

Y_i dalam model *Negative Binomial* adalah variabel yang berupa data *count* sehingga Y_i merupakan bilangan non-negatif, maka ekspektasi dari Y_i juga tidak mungkin negatif. Karena ruang nilai untuk $\mu_i = \mathbf{x}_i' \boldsymbol{\beta}$ adalah bilangan riil pada interval $(-\infty, \infty)$, membuat model regresi tidak dapat digunakan untuk menganalisis data *count*. Untuk mengatasinya digunakan sebuah fungsi penghubung yang menghubungkan antara *fitted value* (μ_i) dengan prediktor linier ($\mathbf{x}_i' \boldsymbol{\beta}$). Hilbe (2011) menyatakan bahwa model *Negative Binomial* pada umumnya menggunakan fungsi peghubung logaritma atau log *link* sebagai berikut.

$$\ln(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta}$$

$$\ln\{E[Y_i | \mathbf{x}_i]\} = \ln(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta} \quad \text{untuk } i = 1, 2, \dots, n$$

$$\mu_i = \exp(\mathbf{x}_i' \boldsymbol{\beta})$$

Sehingga model regresi *Negative Binomial* untuk memodelkan data *count* dapat dituliskan sebagai berikut.

$$f(y|\alpha, \beta) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i! \Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha} + \exp(\mathbf{x}'_i \boldsymbol{\beta})}\right)^{y_i} \left(\frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{\frac{1}{\alpha} + \exp(\mathbf{x}'_i \boldsymbol{\beta})}\right)^{1/\alpha}$$

2.5 Excess Zeros

Salah satu permasalahan pada regresi poisson yaitu nilai nol yang berlebih (*excess zeros*). Variabel respon pada data diskrit mungkin ditemukan data kosong/tak terisi (bernilai nol). Akan tetapi, dalam banyak kasus, data kosong memiliki arti penting pada penelitian yang bersangkutan. Jika data kosong memiliki arti penting dalam data diskrit maka data tersebut harus dimasukkan dalam analisis. *Excess zeros* dapat dilihat pada proporsi variabel respon yang bernilai nol lebih besar dari data diskrit lainnya. *Excess zeros* merupakan salah satu penyebab terjadinya overdispersi (Ariawan, 2012).

2.5.1 Distribusi Zero-Inflated Negative Binomial (ZINB)

Zero-Inflated Negative Binomial merupakan model yang dibentuk dari distribusi *Negative Binomial*. Ketika memodelkan sejumlah data cacahan, ada dua kondisi yang biasanya digunakan distribusi campuran untuk keadaan yang kurang baik. Jika untuk kondisi yang kurang baik adalah poisson, distribusi campurannya adalah *Zero-Inflated Poisson*. Jika untuk kondisi yang kurang baik adalah *Negative Binomial*, yang mana didefinisikan:

$$Y_i \sim \begin{cases} 0, & \text{dengan peluang } \omega \\ \text{Negatif Binomial}(\mu_i), & \text{dengan peluang } (1 - \omega) \end{cases}$$

Jika Y_i adalah variabel random independen yang diskrit, nilai nol pada observasi diduga muncul pada salah satu dari dua kondisi. Kondisi pertama disebut *zero state*, terjadi dengan probabilitas ω dan hanya menghasilkan observasi bernilai nol. Kondisi yang kedua disebut *Negative Binomial state*, terjadi dengan probabilitas $(1 - \omega)$ dan berdistribusi *Negative Binomial* dengan mean μ_i , dengan $0 \leq \omega \leq 1$.

Ariawan (2012) mendefinisikan model regresi *Zero-Inflated Negative Binomial* adalah:

$$P(Y_i = y_i) = \begin{cases} \omega + (1 - \omega) \left(\frac{1}{1 + \alpha\mu_i}\right)^{1/\alpha} & , y_i = 0 \\ (1 - \omega) \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \alpha\mu_i}\right)^{1/\alpha} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i} & , y_i = 1, 2, \dots, \end{cases}$$

dimana μ_i dan α adalah ukuran parameter yang masing-masing distribusi campurannya adalah *Zero-Inflated Negatif Binomial*. Mean dan varians masing-masing didefinisikan sebagai berikut:

$$E(Y_i) = (1 - \omega)\mu_i$$

dan

$$Var(Y_i) = E(Y_i)(1 + \alpha\mu_i + \omega\mu_i)$$

2.5.2 Regresi *Zero-Inflated Negative Binomial* (ZINB)

Model Regresi *Zero-Inflated Negative Binomial* memiliki kegunaan yang sama dengan model Regresi *Negative Binomial* yaitu untuk menganalisis hubungan antara suatu variabel respon data *count* dengan satu atau lebih variabel acak penjelas. Model *Zero-Inflated Negative Binomial* dinotasikan dengan $Y_i \sim ZINB(\mu_i, \alpha, \omega)$ dan untuk memodelkan ω umumnya digunakan model logit, yaitu:

$$\omega = \frac{\exp(\mathbf{Z}_i^T \lambda)}{1 + \exp(\mathbf{Z}_i^T \lambda)}$$

Diasumsikan bahwa parameter μ_i dan ω masing-masing bergantung pada variabel x_i , sehingga model dari regresi ZINB dibagi menjadi dua komponen model yaitu:

- a. Model data diskrit untuk μ_i adalah

$$\ln(\mu_i) = x_i^T \beta \quad , \mu_i \geq 0, i = 1, \dots, n.$$

- b. Model *excess zeros* untuk ω adalah

$$\text{logit}(\omega) = \ln\left(\frac{\omega}{1 - \omega}\right) = x_i^T \gamma \quad , 0 \leq \omega \leq 1, i = 1, \dots, n.$$

x_i adalah matriks variabel yang memuat himpunan-himpunan yang berbeda dari faktor eksperimen yang berhubungan dengan peluang pada *mean Negative Binomial* pada *Negative Binomial state*, sedangkan β dan γ adalah parameter regresi yang akan ditaksir.

2.6 RR-VGLM dan Row Column Interaction Model (RCIM)

Yee dan Hastie (2003) memperkenalkan kelas *Reduced-Rank Vector Generalized Linear Model* (RR-VGLM) dimana mengaplikasikan konsep regresi *reduced rank* ke dalam kelas *vektor generalized linear model* (VGLM) yang berkaitan dengan analisis data kategori. Misal data (x_i, y_i) untuk $i = 1, \dots, n$, dan $x_i = (x_{i1}, \dots, x_{ip})^T$ dengan $x_{i1} = 1$ jika ada intersepsi. Variable eksplanatori x_i dan y_i dan respon masing-masing. VGLMs didefinisikan sebagai model di mana

$$f(y|x; \mathbf{B}) = f(y, \eta_1, \dots, \eta_M)$$

Umumnya VGLM mirip dengan GLM, namun VGLM menyediakan prediktor linier yang banyak. VGLM dapat menangani sebanyak M prediktor linier (dimensi) tergantung dari model yang dipaskan dimana persamaan salah satu prediktor linier ke adalah sebagai berikut :

$$\eta_j(x) = \beta_j^T = \sum_{k=1}^p \beta_{(j)k} x_k, \quad j = 1, \dots, M$$

η_j pada VGLM dapat diterapkan langsung pada parameter distribusi. Pada banyak kasus, koefisien regresi dihubungkan satu sama lain. seperti contoh, beberapa $\beta_j^T x$ mungkin sama, nol, atau penambahan ke jumlah tertentu. Pada situasi ini mungkin dilakukan dengan menggunakan matriks *constraint*. VGLMs secara umum diberikan

$$\eta_j(x) = \sum_{k=1}^p H_k \beta_{(k)}^*, \quad j = 1, \dots, M$$

dimana H_1, H_2, \dots, H_p merupakan *constraint* matriks *full column-rank* dan $\beta_{(k)}^*$ merupakan vektor dari parameter yang diketahui. Dengan tidak ada *constraint*, $H_1 = H_2 = \dots = H_p = I_2$. Sehingga RR-VGLM diberikan:

$$\eta = B_{(1)}^T x_1 + B_{(2)}^T x_2$$

dimana B_2 dengan *reduced-rank regression*

$$B_2 = AC^T$$

C dan A merupakan matriks $p_2 \times R$ dan $M \times R$.

RCIM merupakan bagian dari RR-VGLM dimana prediktor linier yang pertama dimodelkan dengan penjumlahan dari pengaruh baris pengaruh kolom, dan pengaruh interaksi yang mana pada pengaruh interaksi ditunjukkan seperti regresi *reduced-rank*. Sehingga sebuah model RCIM secara umum didefinisikan seperti RR-VLGM yang diaplikasikan pada variabel respon Y sebagai berikut :

$$g_1(\theta_1) \equiv \eta_{1ij} = \mu + \alpha_i + \gamma_j + \sum_{r=1}^R c_{ir} a_{jr}$$

dimana $R \leq \min(M, p_2)$. Hal ini berarti parameter pertama pada model statistik yang berhubungan dengan variabel respon pada matriks. Setelah ditransformasikan, hasilnya sama dengan penjumlahan dari *intercept*, pengaruh baris dan kolom, serta interaksi yang berbentuk $A^T C$. Variabel respon Y pada RCIM berupa angka dan variabel prediktor X pada RCIM berupa baris ke-1 hingga baris ke- i dan dari kolom ke-1 hingga kolom ke- j .

2.7 Uji Kelayakan Model

Terdapat beberapa metode untuk dapat membandingkan suatu model dengan model yang lain. Namun tidak selalu hasil pengujian tersebut memberikan kesimpulan yang sama. Hal ini tidak menjadi persoalan yang rumit karena umumnya setiap peneliti memiliki keyakinan masing-masing dalam pengujian kelayakan model.

2.7.1 Devians

Salah satu aplikasi yang sangat penting pada statistik rasio likelihood ialah konsep devians pada GLM yang digunakan untuk analisis devians. Secara formal, kita dapat menganggap devian mirip dengan jarak antara 2 model probabilistik. Pada kasus khusus, devian juga dapat digunakan sebagai ukuran model yang kurang baik.

Nilai devians dapat didapatkan dengan membandingkan nilai residu devians dari model yang kecil dengan model yang besar. Namun terkadang nilai residu tidak muncul pada hasil estimasi dengan bantuan komputasi. Sehingga perlu ada metode lain untuk dapat memperoleh nilai residu devian tersebut. Pawitan (2001) menjelaskan bahwa nilai residu devian untuk μ didefinisikan sebagai rasio likelihood dari *saturated model* dengan *particular model* :

$$L(y, \mu) = 2 \log \frac{L(y; y)}{L(\mu; y)}$$

dimana $L(y, \mu)$ merupakan nilai Likelihood dari *particular model* dan $L(y, y)$ merupakan nilai likelihood dari *saturated model*.

Jika terdapat 2 GLM yang berbeda dengan D_1 merupakan devians dari model 1 dan D_2 merupakan devians dari model 2, devians dari kedua model tersebut didefinisikan dengan selisih dari model yang kecil dengan model yang besar :

$$D = D_2 - D_1$$

Ini mirip dengan pengujian ketika model yang terekduksi cukup baik dan model kompleks juga cukup baik pada regresi linier. Devians yang lebih besar dari derajat bebas mengindikasikan bahwa kedua model memiliki perbedaan yang signifikan yang selanjutnya dengan menambahkan *p-value* dapat membantu untuk proses analisis pada tabel analisis devians.

2.7.2 Mean Square Error (MSE)

Terdapat beberapa metode yang biasa digunakan untuk menguji model manakah yang lebih baik. Namun terkadang dua metode yang berbeda dapat memberikan hasil yang berbeda, sehingga tidak ada ketetapan khusus metode manakah yang lebih baik dalam mengukur kecocokan model. Peneliti menggunakan metode *Mean Square Error* (MSE) untuk menentukan model manakah yang lebih baik. Nilai MSE diperoleh dengan membagi *Sum Square Error* (SSE) dengan jumlah data atau dalam bentuk matematis dapat ditulis sebagai berikut :

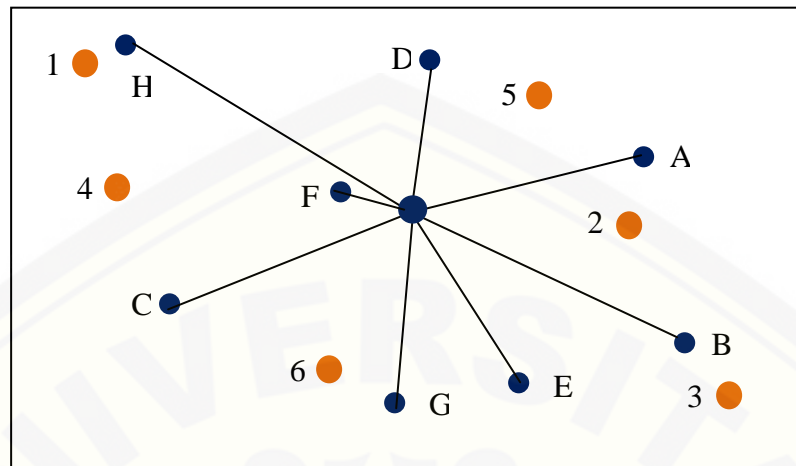
$$MSE = \frac{\sum(y - \hat{y})^2}{n}$$

dimana y merupakan variabel respon, \hat{y} merupakan prediktor dan n adalah banyaknya data. Model dengan nilai MSE yang mendekati angka nol mengindikasikan bahwa model tersebut merupakan model yang lebih baik.

2.8 Analisis Biplot

Biplot merupakan teknik statistika deskriptif yang digunakan untuk menyajikan secara simultan n objek pengamatan dan p atribut dalam ruang bidang datar sehingga ciri-ciri objek pengamatan dan atribut serta posisi antara keduanya dapat dianalisis. Biplot pertama kali diperkenalkan oleh Gabriel pada tahun 1971. Menurut Gabriel (1971), biplot merupakan metode yang sering digunakan dalam analisis multivariat untuk menggambarkan elemen baris dan kolom ke dalam suatu grafik. Dalam Analisis Mutivariat terdapat banyak metode yang dapat digunakan untuk menyelesaikan masalah atau mengolah data yang melibatkan banyak variabel. Misalnya pada saat melakukan suatu penelitian, data yang diperoleh adalah rekapan data yang berupa tabel nilai rata-rata dari beberapa peubah/variabel pada beberapa objek. Metode Multivariat yang dapat digunakan untuk mengatasi masalah tersebut salah satunya adalah biplot.

Dasar dari analisis biplot adalah SVD (*Singular Value Decomposition*) dari suatu matriks. SVD (*Singular Value Decomposition*) dari suatu matrik merupakan teorema dasar yang banyak digunakan dalam perhitungan matriks. SVD (*Singular Value Decomposition*) banyak digunakan karena secara komputerisasi lebih efisien dalam menghasilkan skor komponen utama dan kedua.



Gambar 2.1 Plot Pemetaan Hasil Analisis Biplot

Anak panah yang berupa vektor merupakan gambaran dari variabel penelitian sedangkan titik-titik merupakan gambaran dari objek penelitian. Biplot adalah salah satu upaya menggambarkan data-data yang ada pada tabel ringkasan dalam grafik berdimensi dua. Empat hal penting yang bisa didapatkan dari tampilan biplot adalah:

a. Kedekatan antar objek yang diamati

Informasi ini dapat dijadikan panduan untuk mengetahui objek yang memiliki kemiripan karakteristik dengan objek lain. Penafsiran ini mungkin akan berbeda untuk setiap bidang terapan, namun inti dari penafsiran ini adalah bahwa dua objek yang memiliki karakteristik sama akan digambarkan sebagai dua titik dengan posisi yang berdekatan.

b. Keragaman peubah

Informasi ini digunakan untuk melihat apakah ada variabel yang mempunyai nilai keragaman yang hampir sama untuk setiap objek. Dengan informasi ini, bisa diperkirakan pada variabel mana strategi tertentu harus ditingkatkan, dan juga sebaliknya.

c. Korelasi antar peubah

Dari informasi ini bisa diketahui bagaimana suatu variabel mempengaruhi ataupun dipengaruhi variabel yang lain. Pada biplot, variabel akan digambarkan

sebagai garis berarah. Dua variabel yang memiliki nilai korelasi positif akan digambarkan sebagai dua buah garis dengan arah yang sama atau membentuk sudut sempit. Sementara itu, dua variabel yang memiliki nilai korelasi negatif akan digambarkan dalam bentuk dua garis dengan arah yang berlawanan atau membentuk sudut lebar (tumpul). Sedangkan dua variabel yang tidak berkorelasi akan digambarkan dalam bentuk dua garis dengan sudut yang mendekati 90^0 (siku-siku).

d. Nilai peubah pada suatu objek

Dalam informasi ini digunakan untuk melihat keunggulan dari setiap objek. Objek yang terletak searah dengan arah vektor variabel dikatakan bahwa objek tersebut mempunyai nilai di atas rata-rata. Namun jika objek terletak berlawanan dengan arah dari vektor variabel tersebut, maka objek tersebut memiliki nilai di bawah rata-rata. Sedangkan objek yang hampir berada ditengah-tengah berarti objek tersebut memiliki nilai dekat dengan rata-rata.

Gabriel (1971) mengemukakan bahwa ukuran pendekatan matriks X dengan biplot dalam bentuk:

$$\rho^2 = \frac{(\lambda_1 + \lambda_2)}{\sum_{k=1}^r \lambda_k}$$

Dengan λ_1 adalah nilai eigen terbesar ke-1, λ_2 adalah nilai eigen terbesar ke-2 dan $\lambda_k, k = 1, 2, \dots, r$ adalah nilai eigen ke- k . Apabila ρ^2 mendekati nilai satu, maka biplot memberikan penyajian yang semakin baik mengenai informasi data yang sebenarnya.

Perlu dipahami sebelumnya bahwa biplot adalah upaya membuat gambar di ruang berdimensi banyak menjadi gambar di ruang berdimensi dua. Pereduksian dimensi ini mengakibatkan menurunnya informasi yang terkandung dalam biplot. Biplot yang mampu memberikan informasi sebesar 70% dari seluruh informasi dianggap cukup (Mattjik & Sumertajaya, 2011).

BAB 3 METODOLOGI PENELITIAN

3.1 Desain Penelitian

Data pada penelitian ini adalah data yang dibangkitkan kemudian proses perhitungannya dibantu dengan salah satu aplikasi statistik yaitu program Rstudio dan R i386 3.1.0. Penjelasan dari masing-masing langkah penelitian pada skema desain penelitian adalah sebagai berikut :

a. Data Penelitian

Pada tahap ini peneliti menyiapkan 2 data dalam bentuk tabel dua arah. Kedua data tersebut memiliki karakteristik yang berbeda. Data yang akan digunakan ialah skenario 3 dan skenario 4. Skenario 3 adalah data jumlah isi polong kedelai yang terdapat *outlier* dan memiliki kondisi *excess zeros*. Skenario 4 adalah data yang memiliki kondisi *excess zeros* dan *outlier* serta terdapat pengamatan nol satu baris. Sedangkan untuk menguji seberapa kecil prosentase angka nol yang mengganggu model *Negative Binomial*, peneliti menggunakan data 1 yang dimodifikasi dengan *outlier* saja yang selanjutnya data ini menjadi skenario 1.

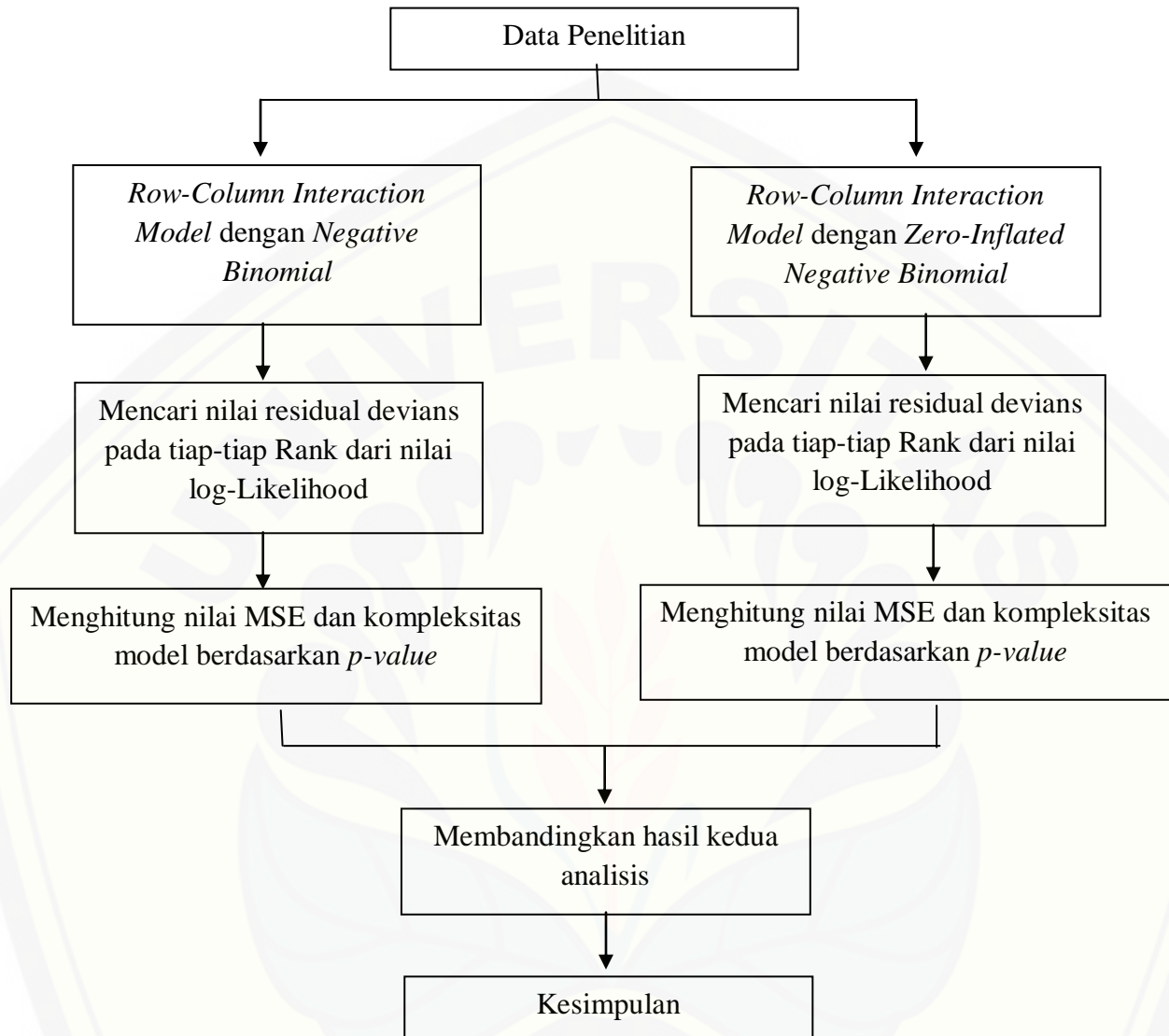
b. *Row-Column Interaction Model* dengan *Negative Binomial*

Memodelkan data menggunakan *Row-Column Interaction Model* dengan menetapkan family distribusinya yaitu *negbinomial* untuk fungsi distribusi *Negative Binomial*. Estimasi parameter model menggunakan bantuan fungsi *rcim* pada program Rstudio.

c. *Row-Column Interaction Model* dengan *Zero-Inflated Negative Binomial*

Memodelkan data menggunakan *Row-Column Interaction Model* dengan menetapkan family distribusinya yaitu *zinegbinomial* untuk fungsi distribusi *Zero-Inflated Negative Binomial*. Estimasi parameter model menggunakan bantuan fungsi *rcim* pada program R.

- d. Mencari nilai residual devians pada tiap-tiap Rank dari nilai log-Likelihood
Mencari nilai log likelihood dan nilai residual devians dari *Null* model hingga saturated model. Nilai log likelihood tersebut akan digunakan untuk memperoleh nilai devian dari masing-masing model.
- e. Menghitung nilai MSE dan kompleksitas model berdasarkan *p-value*
Nilai *p-value* didapatkan dari hasil analisis devians dan digunakan untuk menentukan kompleksitas model. Model terbaik yang terpilih kemudian akan dicari nilai MSE-nya.
- f. Kesimpulan
Setelah semua tahapan penelitian dilakukan, maka ditarik kesimpulan yang merupakan jawaban dari rumusan masalah pada bab 1.
Secara sistematis akan dijelaskan melalui gambar 3.1.



Gambar 3.1 Diagram Metode Penelitian

3.2 Data Penelitian dan Skenario Simulasi

Data penelitian ini menggunakan 2 jenis data yang berbeda. Kedua data tersebut merupakan data sekunder yang diperoleh dari Balitkabi Malang. Data secara lengkap dapat dilihat pada Lampiran A.

3.2.1 Data Penelitian

Data jumlah polong isi kedelai yang selanjutnya disebut data 1 merupakan data yang digunakan oleh Ahadiyah (2015), data ini meliputi 15 jenis galur yang ditanam pada 8 lokasi yang berbeda. Data ini adalah data cacahan yang memiliki kondisi tanpa *outlier* dan pengamatan nol. Data ini disajikan kedalam bentuk matriks dengan ukuran 15×8 .

Data serangan hama penyakit pada tanaman kacang hijau yang selanjutnya disebut data 2 merupakan data riil yang digunakan oleh Wicaksono (2015). Percobaan ini melibatkan 12 jenis galur yang ditanam pada 5 lokasi yang berbeda. Data riil ini adalah data cacahan yang memiliki kondisi *excess zeros*. Dapat dilihat pada L1 (lokasi 1) bahwa pengamatan nol terdapat pada semua genotip yang mengakibatkan pengamatan nol terstruktur pada baris tersebut. Selain itu terdapat pula pengamatan nol pada G11 (galur 11) di L2 (lokasi 2). Data riil ini disajikan kedalam bentuk matriks dengan ukuran 5×12 . Sehingga data ini berjenis data cacahan dalam keadaan overdispersi yang memiliki kondisi *excess zeros*.

3.2.2 Skenario Simulasi

- a. Skenario 1 merupakan data simulasi jumlah isi polong kedelai. Skenario 1 dimodifikasi dengan cara menghitung nilai standar deviasi dari masing-masing baris dan kolom, kemudian ditambahkan dengan nilai maksimal pada tiap baris dan kolom. Secara matematis dituliskan $max + 3\sqrt{stdev}$. Hal ini dilakukan untuk member nilai outlier pada data yaitu dengan menambahkan nilai maksimal dengan 3 kali standar deviasi atau lebih (Kriegel, 2010). Sehingga data ini berjenis data cacahan yang terdapat *outlier*.
- b. Skenario 2 merupakan data simulasi jumlah isi polong kedelai. Skenario 2 ini adalah data yang digunakan Ahadiyah (2015) yang dimodifikasi dengan cara memasukkan 3 nilai nol yang paling minimal pada tiap kolom. Sehingga data ini berjenis data cacahan yang memiliki kondisi *excess zeros*.

- c. Skenario 3 merupakan data simulasi jumlah isi polong kedelai. Skenario 3 ini adalah skenario 1 yang dimodifikasi dengan cara memasukkan 3 nilai nol yang paling minimal pada tiap kolom. Sehingga data ini berjenis data cacahan yang terdapat *outlier* dan memiliki kondisi *excess zeros*.
- d. Skenario 4 merupakan data serangan hama penyakit pada tanaman kacang hijau yang dimodifikasi dengan cara menghitung nilai standar deviasi dari masing-masing baris dan kolom, kemudian ditambahkan dengan nilai maksimal pada tiap baris dan kolom. Secara matematis dituliskan $max + 3\sqrt{stdev}$. Sehingga data ini berjenis data cacahan yang terdapat *outlier* dan memiliki kondisi *excess zeros*.

3.3 Struktur Fungsi Program R

Metode yang digunakan untuk menganalisis data pada penelitian ini adalah pendekatan RCIM. Untuk mempermudah proses perhitungan digunakan bantuan salah satu *software* dalam statistika yaitu Rstudio. Cara penggunaan Rstudio mirip dengan program R. Rstudio adalah *Integrated Development Environment (IDE)* utama bagi pengembang yang menggunakan bahasa R. Uniform Resource Locator (URL) Rstudio dapat diunduh pada link <http://www.rstudio.com/ide/download/desktop>. Menurut Tirta (2009), *source* kode dari program R dapat diakses, dimodifikasi, dan dikembangkan secara bebas sesuai dengan keperluan.

Untuk menganalisis data menggunakan RCIM dibantu dengan paket yang tersedia pada Rstudio yaitu paket VGAM. Paket ini menyediakan fungsi untuk vektor GLM dan model aditif lainnya seperti VGLM VGAM. Fungsi yang digunakan dalam penelitian ini adalah:

- a. `vglm`, digunakan untuk memodelkan suatu data menjadi bentuk VGLM.

Struktur fungsi `vglm` adalah sebagai berikut :

```
vglm(formula, family, data, ...)
```

- b. `array`, digunakan untuk membuat vektor. Struktur fungsi `array` adalah sebagai berikut :
- ```
array(data, dim, ...)
```
- c. `rcim`, digunakan untuk memodelkan data dengan *Row-Column Interaction Models*. Struktur fungsi `rcim` adalah sebagai berikut :
- ```
rcim(data, family, constrain, rank)
```
- d. `logLik`, digunakan untuk memunculkan nilai log likelihood hasil iterasi. Struktur fungsi `logLik` adalah sebagai berikut :
- ```
logLik(object, ...)
```
- e. `deviance`, digunakan untuk memperoleh nilai residu devian pada model. Struktur fungsi `deviance` adalah sebagai berikut :
- ```
deviance(object, ...)
```
- f. `residuals`, digunakan untuk memperoleh nilai residu dari hasil pengepasan model pada data. Struktur fungsi `residuals` adalah sebagai berikut :
- ```
residuals(object)
```
- g. `plot`, digunakan untuk memplotkan suatu data. Struktur fungsi `plot` adalah sebagai berikut :
- ```
plot(x, y, ...)
```
- h. `pf`, digunakan untuk memperoleh nilai P-Value dari model. Struktur fungsi `pf` adalah sebagai berikut :
- ```
pf(q, df1, df2, ncp, lower.tail = FALSE)
```
- i. `mean`, digunakan untuk mencari nilai rata-rata dari suatu data. Struktur fungsi `mean` adalah sebagai berikut :
- ```
mean(object)
```
- j. `svd`, digunakan untuk menghitung nilai SVD (*singular value decomposition*). Struktur fungsi `svd` adalah sebagai berikut :
- ```
svd(x, ...)
```

## BAB 4 HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai hasil dari penelitian tentang penggunaan metode *Row Colum Interaction Model* (RCIM) pada data berdistribusi *Zero Inflated Negative Binomial* (ZINB). Regresi *Negative Binomial* merupakan salah satu metode statistik yang digunakan untuk mengatasi permasalahan *outlier* pada data, sedangkan Regresi *Zero Inflated Negative Binomial* (ZINB) merupakan salah satu metode statistik yang digunakan untuk mengatasi permasalahan *outlier* dan *excess zeros* pada data berjenis cacahan. Dalam penelitian ini, pengolahan data dibantu dengan menggunakan *software* R studio dan R i386 3.1.0 serta paket yang digunakan adalah paket VGAM.

### 4.1 Hubungan *Negative Binomial* dan ZINB

Model regresi *Negative Binomial* yang dibangun memiliki sebaran *Negative Binomial* dengan parameter  $\mu$  dan  $\alpha$ . Sehingga sebaran  $Y$  menjadi:

$$f(y; \mu, \alpha) = \frac{\Gamma(y_i + 1/\alpha)}{y_i! \Gamma(1/\alpha)} \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha} \left( \frac{\mu_i}{1/\alpha + \mu_i} \right)^{y_i}$$

*Zero-Inflated Negative Binomial* merupakan model yang dibentuk dari distribusi *Negative Binomial* dan campuran poisson gamma. Jika nilai  $\omega = 0$  maka mean dan varians ZINB akan menjadi bentuk mean dan varians dari *Negative Binomial*.

Mean ZINB saat  $\omega = 0$ .

$$E(Y_i) = (1 - \omega)\mu_i$$

$$E(Y_i) = (1 - 0)\mu_i$$

$$E(Y_i) = \mu_i$$

Varians ZINB saat  $\omega = 0$ .

$$Var(Y_i) = E(Y_i)(1 + \alpha\mu_i + \omega\mu_i)$$

$$Var(Y_i) = (1 - 0) \mu_i(1 + \alpha\mu_i + 0\mu_i)$$

$$Var(Y_i) = \mu_i(1 + \alpha\mu_i)$$

Secara garis besar model *Zero-Inflated Negative Binomial* didefinisikan sebagai berikut:

$$Y_i \sim \begin{cases} 0, & \text{dengan peluang } \omega \\ \text{Negatif Binomial}(\mu_i), & \text{dengan peluang } (1 - \omega) \end{cases}$$

Pada keadaan peluang  $\omega$ , Jika  $Y_i = 0$ , maka persamaannya ialah:

$$= \omega + (1 - \omega) \text{Negative Binomial}$$

$$= \omega + (1 - \omega) \frac{\Gamma(y_i + 1/\alpha)}{y_i! \Gamma(1/\alpha)} \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha} \left( \frac{\mu_i}{1/\alpha + \mu_i} \right)^{y_i}$$

$$= \omega + (1 - \omega) \frac{\Gamma(0 + 1/\alpha)}{0! \Gamma(1/\alpha)} \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha} \left( \frac{\mu_i}{1/\alpha + \mu_i} \right)^0$$

$$= \omega + (1 - \omega) \frac{\Gamma(1/\alpha)}{\Gamma(1/\alpha)} \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha}$$

$$= \omega + (1 - \omega) \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha}$$

Pada keadaan peluang  $(1 - \omega)$ , Jika  $Y_i = 1, 2, \dots$ , maka persamaannya ialah:

$$= (1 - \omega) \text{Negative Binomial}$$

$$= (1 - \omega) \frac{\Gamma(y_i + 1/\alpha)}{y_i! \Gamma(1/\alpha)} \left( \frac{1/\alpha}{1/\alpha + \mu_i} \right)^{1/\alpha} \left( \frac{\mu_i}{1/\alpha + \mu_i} \right)^{y_i}$$

#### 4.2 Karakteristik Data Skenario

Skenario 1 merupakan data simulasi jumlah isi polong kedelai yang dimodifikasi dengan cara menghitung nilai standar deviasi dari masing-masing baris



dan kolom, kemudian ditambahkan dengan nilai maksimal pada tiap baris dan kolom. Sehingga data ini berjenis data cacahan yang terdapat *outlier*.

Skenario 2 merupakan data simulasi jumlah isi polong kedelai. Skenario 2 ini adalah data yang dimodifikasi dengan cara memasukkan 3 nilai nol yang paling minimal pada tiap kolom. Sehingga data ini berjenis data cacahan yang memiliki kondisi *excess zeros*.

Skenario 3 merupakan data simulasi jumlah isi polong kedelai. Skenario 3 ini adalah skenario 1 yang dimodifikasi dengan cara memasukkan 3 nilai nol yang paling minimal pada tiap kolom. Sehingga data ini berjenis data cacahan yang terdapat *outlier* dan memiliki kondisi *excess zeros*.

Skenario 4 merupakan data simulasi serangan hama penyakit pada tanaman kacang hijau yang dimodifikasi dengan cara menghitung nilai standar deviasi dari masing-masing kolom, kemudian ditambahkan dengan nilai maksimal pada tiap kolom. Sehingga data ini berjenis data cacahan yang terdapat *outlier* dan memiliki kondisi *excess zeros*.

#### **4.3 Pengaruh *Outlier* pada Model *Negative Binomial* dan *ZINB***

Skenario 1 merupakan data simulasi jumlah isi polong kedelai yang dimodifikasi dengan cara menghitung nilai standar deviasi dari masing-masing baris dan kolom, kemudian ditambahkan dengan nilai maksimal pada tiap baris dan kolom. Pada data ini *Zero Inflated Negative Binomial* tidak dapat menganalisanya dikarenakan tidak terdapat pengamatan nol. *Zero Inflated Negative Binomial* dapat menganalisa pengamatan nol pada data cacahan dengan persentase sebesar 8,33%. Susilowati (2015) telah meneliti bahwa regresi *Negative Binomial* dapat mengatasi dengan baik data cacahan yang terdapat *outlier* didalamnya. Model *Negative Binomial* dapat memberikan analisis devian yang kompleks dan nilai MSE yang lebih baik dari model Poisson.

#### 4.4 Pengaruh *Excess Zeros* pada Model *Negative Binomial* dan ZINB

Skenario 2 merupakan data simulasi polong kacang hijau yang memiliki kondisi *excess zeros*. Data 2 merupakan data serangan hama pada tanaman kacang hijau yang terdapat pengamatan nol pada data. Pengamatan nol pada data ini cukup menarik, karena terdapat angka nol penuh pada 1 baris data. Tabel 4.1 dan Tabel 4.2 akan menyajikan nilai *Log Likelihood* pada skenario 2 dan data 2.

Tabel 4.1 Nilai *Log-Likelihood* skenario 2 dengan pendekatan *rcim*

| Model    | Regresi ZINB | Regresi <i>Negative Binomial</i> |
|----------|--------------|----------------------------------|
| Null     | -556,5205    | -620,6408                        |
| Baris    | -536,1783    | -612,6517                        |
| Kolom    | -462,3490    | -617,5125                        |
| Rank = 0 | -539,2798    | -607,8516                        |
| Rank = 1 | -434,0176    | -503,4749                        |
| Rank = 2 | -366,1121    | -421,6781                        |
| Rank = 3 | -336,6064    | -353,7470                        |
| Rank = 4 | -317,6875    | -325,6028                        |
| Rank = 5 | -308,0453    | -316,8316                        |
| Rank = 6 | -303,1202    | -315,5627                        |
| Rank = 7 | -302,2272    | -311,2199                        |

Skenario 2 merupakan data simulasi polong kacang hijau yang telah dimodifikasi dengan mengganti 3 nilai nol paling minimal pada tiap kolom sehingga mengalami kondisi *excess zeros*.

Tabel 4.2 Nilai *Log-Likelihood* data 2 dengan pendekatan rcim

| Model    | Regresi ZINB | Regresi <i>Negative Binomial</i> |
|----------|--------------|----------------------------------|
| Null     | -294,3771    | -330,1365                        |
| Baris    | -266,1469    | -274,2027                        |
| Kolom    | -245,3563    | -325,7052                        |
| Rank = 0 | -264,3893    | -227,6392                        |
| Rank = 1 | -196,3969    | -186,4723                        |
| Rank = 2 | -174,1760    | -166,9237                        |
| Rank = 3 | -162,9126    | -155,6934                        |
| Rank = 4 | -155,8735    | -155,6856                        |

Data 2 merupakan data serangan hama pada tanaman kacang hijau yang terdapat pengamatan nol satu baris pada data.

Pada skenario 2 menunjukkan Nilai *Log-Likelihood* RCIM ZINB lebih besar daripada RCIM *Negative Binomial* sedangkan pada data 2 menunjukkan Nilai *Log-Likelihood* RCIM ZINB lebih kecil daripada RCIM *Negative Binomial*. Nilai *Log-Likelihood* yang didapatkan dari masing-masing model selanjutnya digunakan untuk mencari nilai model devians. Cara mengitung nilai model devians untuk mengetahui analisis residu devians pada masing-masing data. Cara menghitung nilai model devians telah dijelaskan sebelumnya pada subbab 2.6.1 yaitu dua kali perbandingan nilai *Log-Likelihood* dari *saturated* model dengan *null* model. Model devians pada Regresi *Negative Binomial* dan Regresi *Zero Inflated Negative Binomial* hanya didapatkan nilai *Log-Likelihood* oleh karena itu dibutuhkan penghitungan secara manual untuk mendapatkan nilai Model devians. Pada Tabel 4.3 dan Tabel 4.4, berikut menunjukkan hasil dari perhitungan model devians dan juga analisis residu deviansnya.

Tabel 4.3 Nilai Model Devians dan Analisis Residu Devians Pada Skenario 2

| Model    | Regresi ZINB  |                | Regresi <i>Negative Binomial</i> |                |
|----------|---------------|----------------|----------------------------------|----------------|
|          | Model Devians | Residu Devians | Model Devians                    | Residu Devians |
| Null     | 508,5866      | -              | 618,8418                         | -              |
| Baris    | 467,9022      | 40,6844        | 602,8636                         | 15.9782        |
| Kolom    | 320,2436      | 188,3430       | 612,5852                         | 6.2566         |
| Rank = 0 | 474,1052      | -              | 593,2634                         | -              |
| Rank = 1 | 263,5808      | 210,5244       | 384,5100                         | 208.7534       |
| Rank = 2 | 127,7698      | 135,8110       | 220,9164                         | 163.5936       |
| Rank = 3 | 68,7584       | 59,0114        | 85,0542                          | 135.8622       |
| Rank = 4 | 30,9206       | 37,8378        | 28,7658                          | 56.2884        |
| Rank = 5 | 11,6362       | 19,2844        | 11,2234                          | 17.5424        |
| Rank = 6 | 1,7860        | 9,8502         | 8,6856                           | 2.5378         |
| Rank = 7 | 0             | 1,7860         | 0                                | 8.6856         |

Tabel 4.4 Nilai Model Devians dan Analisis Residu Devians Pada Data 2

| Model    | Regresi ZINB  |                | Regresi <i>Negative Binomial</i> |                |
|----------|---------------|----------------|----------------------------------|----------------|
|          | Model Devians | Residu Devians | Model Devians                    | Residu Devians |
| Null     | 277,0072      | -              | 348,9018                         | -              |
| Baris    | 220,5468      | 56,4604        | 237,0342                         | 111,8676       |
| Kolom    | 178,9656      | 98,0416        | 340,0392                         | 8,8626         |
| Rank = 0 | 217,0316      | -              | 143,9072                         | -              |
| Rank = 1 | 70,0346       | 146,9970       | 66,8188                          | 77,0884        |
| Rank = 2 | 36,6050       | 33,4296        | 22,4762                          | 44,3426        |
| Rank = 3 | 14,0782       | 22,5268        | 0,0156                           | 22,4606        |
| Rank = 4 | 0             | 14,0782        | 0                                | 0,0156         |

Nilai Residu devians baris dan kolom diperoleh dari selisih antara nilai model devians *Null* model dengan model devians dari model baris dan kolom. Sedangkan untuk nilai residu devians pada model rank = 1 diperoleh dari selisih antara nilai model devians dari model rank = 0 dengan model rank = 1. Nilai residu devians untuk model rank = 2 diperoleh dari selisih antara model devians dari model rank = 1 dengan model rank = 2. Begitu seterusnya hingga model rank maksimum. Sehingga dari nilai residu devians model RCIM diatas diperoleh analisis residu devians sebagai berikut:.

Tabel 4.5 Analisis Residu Devians RCIM6 (Skenario 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 208,7534          | 10,43767          | 9,61377              | 0,001361 | **           |
| RCIM 2   | 18 | 163,5936          | 9,08853           | 8,37113              | 0,002305 | **           |
| RCIM 3   | 16 | 135,8622          | 8,49139           | 7,82112              | 0,003054 | **           |
| RCIM 4   | 14 | 56,2884           | 4,02060           | 7,70323              | 0,034336 | *            |
| RCIM 5   | 12 | 17,5424           | 1,46187           | 1,34647              | 0,344353 | Ns           |
| RCIM 6   | 10 | 2,5378            | 0,25378           | 0,23375              | 0,982223 | Ns           |
| Residual | 8  | 8,6856            | 1,08570           |                      |          |              |

Tabel 4.6 Analisis Residu Devians RCIM4 (Skenario 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 208,7534          | 10,43767          | 10,88550             | 7,84E-09 | ***          |
| RCIM 2   | 18 | 163,5936          | 9,08853           | 9,47848              | 6,14E-08 | ***          |
| RCIM 3   | 16 | 135,8622          | 8,49139           | 8,85571              | 2,11E-07 | ***          |
| RCIM 4   | 14 | 56,2884           | 4,02060           | 4,19310              | 0,000483 | ***          |
| Residual | 30 | 8,6856            | 0,95886           |                      |          |              |



Tabel 4.7 Analisis Residu Devians RCIM3 (Skenario 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 208,7534          | 10,43767          | 5,39959              | 1,58E-06 | ***          |
| RCIM 2   | 18 | 163,5936          | 9,08853           | 4,70166              | 0,000014 | ***          |
| RCIM 3   | 16 | 135,8622          | 8,49139           | 4,39274              | 0,000048 | ***          |
| Residual | 44 | 85,0542           | 1,93305           |                      |          |              |

Tabel 4.8 Analisis Residu Devians RCIM2 (Skenario 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 208,7534          | 10,43767          | 2,83483              | 0,000970 | ***          |
| RCIM 2   | 18 | 163,5936          | 9,08853           | 2,46841              | 0,004686 | **           |
| Residual | 60 | 220,9164          | 3,68194           |                      |          |              |

Tabel 4.9 Analisis Residu Devians RCIM1 (Skenario 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 208,7534          | 10,43767          | 2,11734              | 0,010170 | *            |
| Residual | 78 | 384,5100          | 4,92962           |                      |          |              |

Kode signifikan: 0 – 0,001 “\*\*\*”  
 0,001 – 0,01 “\*\*”  
 0,01 – 0,05 “\*”

Untuk mencari model terbaik dilakukan analisis nilai Rasio Rataan Devians. Model RCIM dikatakan signifikan jika  $p\text{-value} < 0,05$ . Hasil analisis residu devians *Negative Binomial* pada hasil pendugaan skenario 2 tabel 4.5 menunjukkan model RCIM6 dengan 6 komponen utama dan RCIM5 dengan 5 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.6 dan tabel 4.7 menunjukkan model

RCIM4 dengan 4 komponen utama dan model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$  begitu juga dengan tabel 4.8 menunjukkan model RCIM2 dengan 2 komponen utama signifikan dimana  $p\text{-value} < 0,05$  dan tabel 4.9 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 4 tabel yang signifikan menunjukkan bahwa RCIM4 dengan 4 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM *Negative Binomial*. Untuk pendugaan analisis residu devians RCIM ZINB pada skenario 2 akan ditunjukkan pada tabel sebagai berikut:

Tabel 4.10 Analisis Residu Devians RCIM6 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | $p\text{-value}$ | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|------------------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 47,14992             | 3,37E-06         | ***          |
| RCIM 2   | 18 | 135,8110          | 7,54506           | 33,79644             | 1,29E-05         | ***          |
| RCIM 3   | 16 | 59,0114           | 3,68821           | 16,52055             | 0,000208         | ***          |
| RCIM 4   | 14 | 37,8378           | 2,70270           | 12,10616             | 0,000696         | ***          |
| RCIM 5   | 12 | 19,2844           | 1,60703           | 7,19836              | 0,004589         | **           |
| RCIM 6   | 10 | 9,8502            | 0,98502           | 4,41218              | 0,023119         | *            |
| Residual | 8  | 1,7860            | 0,22325           |                      |                  |              |

Tabel 4.11 Analisis Residu Devians RCIM5 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | $p\text{-value}$ | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|------------------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 16,28298             | 9,35E-08         | ***          |
| RCIM 2   | 18 | 135,8110          | 7,54506           | 11,67142             | 1,60E-06         | ***          |
| RCIM 3   | 16 | 59,0114           | 3,68821           | 5,70528              | 0,000334         | ***          |
| RCIM 4   | 14 | 37,8378           | 2,70270           | 4,18080              | 0,002676         | **           |
| RCIM 5   | 12 | 19,2844           | 1,60703           | 2,48592              | 0,039431         | *            |
| Residual | 18 | 11,6362           | 0,64646           |                      |                  |              |

Tabel 4.12 Analisis Residu Devians RCIM4 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 10,21282             | 1,69E-08       | ***          |
| RCIM 2   | 18 | 135,8110          | 7,54506           | 7,32042              | 1,13E-06       | ***          |
| RCIM 3   | 16 | 59,0114           | 3,68821           | 3,57840              | 0,001267       | ***          |
| RCIM 4   | 14 | 37,8378           | 2,70270           | 2,62223              | 0,013081       | ***          |
| Residual | 30 | 30,9206           | 1,03069           |                      |                |              |

Tabel 4.13 Analisis Residu Devians RCIM3 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 6,73596              | 7,35E-08       | ***          |
| RCIM 2   | 18 | 135,8110          | 7,54506           | 4,82825              | 1,02E-05       | ***          |
| RCIM 3   | 16 | 59,0114           | 3,68821           | 2,36017              | 0,012501       | *            |
| Residual | 44 | 68,7584           | 1,56269           |                      |                |              |

Tabel 4.14 Analisis Residu Devians RCIM2 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 4,94306              | 7,38E-07       | ***          |
| RCIM 2   | 18 | 135,8110          | 7,54506           | 3,54312              | 0,000164       | ***          |
| Residual | 60 | 127,7698          | 2,12950           |                      |                |              |

Tabel 4.15 Analisis Residu Devians RCIM1 (Skenario 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 210,5244          | 10,52622          | 3,11497              | 0,000175 | ***          |
| Residual | 78 | 263,5808          | 3,37924           |                      |          |              |

Hasil analisis residu devians ZINB pada hasil pendugaan skenario 2 tabel 4.10 menunjukkan model RCIM6 dengan 6 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.11 menunjukkan model RCIM5 dengan 5 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.12 menunjukkan model RCIM4 dengan 4 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.13 menunjukkan model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.14 menunjukkan model RCIM2 dengan 2 komponen utama signifikan dimana  $p\text{-value} < 0,05$ , begitu pula pada tabel 4.15 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 6 tabel yang signifikan menunjukkan bahwa RCIM6 dengan 6 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM ZINB. Untuk pendugaan analisis residu devians RCIM *Negative Binomial* pada data 2 akan ditunjukkan pada tabel sebagai berikut:

Tabel 4.16 Analisis Residu Devians RCIM3 (Data 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 14 | 77,0884           | 5,50631           | 2823,751             | 3,52E-13 | ***          |
| RCIM 2   | 12 | 44,3426           | 3,69522           | 1894,983             | 1,92E-12 | ***          |
| RCIM 3   | 10 | 22,4606           | 2,24606           | 1151,826             | 1,62E-11 | ***          |
| Residual | 8  | 0,0156            | 0,00195           |                      |          |              |

Tabel 4.17 Analisis Residu Devians RCIM2 (Data 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 14 | 77,0884           | 5,50631           | 4,40972              | 0,001971       | **           |
| RCIM 2   | 12 | 44,3426           | 3,69522           | 2,95930              | 0,018562       | *            |
| Residual | 18 | 22,4762           | 1,24868           |                      |                |              |

Tabel 4.18 Analisis Residu Devians RCIM1 (Data 2, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 14 | 77,0884           | 5,50631           | 2,47220              | 0,018382       | *            |
| Residual | 30 | 66,8188           | 2,22729           |                      |                |              |

Hasil analisis residu devians *Negative Binomial* pada hasil pendugaan data 2 tabel 4.16 menunjukkan bahwa model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$ , begitu juga dengan tabel 4.17 menunjukkan model RCIM2 dengan 2 komponen utama signifikan dimana  $p\text{-value} < 0,05$  dan tabel 4.18 menunjukkan model RCIM1 dengan riil komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 3 tabel yang signifikan menunjukkan bahwa RCIM3 dengan 3 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM *Negative Binomial*. Untuk pendugaan analisis residu devians RCIM ZINB pada data 2 akan ditunjukkan pada sebagai berikut:



Tabel 4.19 Analisis Residu Devians RCIM3 (Data 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 14 | 135,9848          | 9,71320           | 5,51957              | 0,010229 | *            |
| RCIM 2   | 12 | 44,4418           | 3,70348           | 2,10452              | 0,148555 | ns           |
| RCIM 3   | 10 | 22,5268           | 2,25268           | 1,28010              | 0,370414 | ns           |
| Residual | 8  | 14,0782           | 1,75978           |                      |          |              |

Tabel 4.20 Analisis Residu Devians RCIM1 (Data 2, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 14 | 135,9848          | 9,71320           | 3,59540              | 0,001595 | **           |
| Residual | 30 | 81,0468           | 2,70156           |                      |          |              |

Hasil analisis residu devians ZINB pada hasil pendugaan data 2 tabel 4.19 menunjukkan bahwa model RCIM3 dengan 3 komponen utama dan model RCIM2 dengan 2 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . Tabel 4.20 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 1 tabel yang signifikan menunjukkan bahwa RCIM1 dengan 1 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM ZINB. Hasil pendugaan skenario 2 dengan RCIM *Negative Binomial* dan RCIM ZINB lebih baik dengan menggunakan RCIM ZINB. Dapat dilihat pada Tabel 4.2 yang menunjukkan besarnya nilai *log-Likelihood* dari kedua metode. Begitu juga dengan Tabel 4.4 yang menunjukkan analisis deviansnya. Tabel 4.21 berikut merupakan tabel besar MSE dari kedua metode. Tabel tersebut menunjukkan bahwa MSE ZINB lebih baik dari MSE *Negative Binomial*. Sehingga dapat disimpulkan bahwa RCIM ZINB dapat menduga lebih baik daripada RCIM *Negative Binomial* pada data *excess zeros*.

Tabel 4.21 Nilai *Mean Square Error* Skenario 2 dengan pendekatan *rcim*

| Model  | MSE ZINB   | MSE <i>Negative Binomial</i> |
|--------|------------|------------------------------|
| RCIM 1 | 0,07589846 | 0,20876010                   |
| RCIM 2 | 0,07504630 | 0,09743042                   |
| RCIM 3 | 0,02441078 | 0,10276670                   |
| RCIM 4 | 0,01666468 | 0,05146207                   |
| RCIM 5 | 0,00899084 | 0,02056970                   |
| RCIM 6 | 0,00535431 | 0,02681639                   |

Perhitungan besar nilai MSE pada Tabel 4.22 menunjukkan bahwa meskipun model RCIM dengan ZINB dapat memberikan nilai MSE yang lebih kecil pada RCIM1 dan RCIM2 daripada model RCIM dengan *Negative Binomial*, namun pada RCIM3 MSE *Negative Binomial* lebih kecil dari model RCIM ZINB.

Tabel 4.22 Nilai *Mean Square Error* Data 2 dengan Pendekatan *rcim*

| Model  | MSE ZINB   | MSE <i>Negative Binomial</i> |
|--------|------------|------------------------------|
| RCIM 1 | 0,02143623 | 0,02549026                   |
| RCIM 2 | 0,01260851 | 0,01268638                   |
| RCIM 3 | 0,01209246 | 9,86e-06                     |

#### 4.5 Pengaruh *Outlier* dan *Excess Zeros* pada Model *Negative Binomial* dan ZINB

Skenario 3 merupakan data simulasi polong kacang hijau yang terdapat *outlier* dan memiliki kondisi *excess zeros*. Skenario 4 merupakan data serangan hama pada tanaman kacang hijau yang terdapat pengamatan nol satu baris pada data dan telah dimodifikasi dengan *outlier*. Tabel 4.23 dan Tabel 4.24 akan menyajikan nilai *Log Likelihood* pada skenario 3 dan skenario 4.

Tabel 4.23 Nilai *Log-Likelihood* skenario 3 dengan pendekatan *rcim*

| Model    | Regresi ZINB | Regresi <i>Negative Binomial</i> |
|----------|--------------|----------------------------------|
| Null     | -562,5072    | -623,3174                        |
| Baris    | -541,8295    | -614,8219                        |
| Kolom    | -472,5752    | -620,2632                        |
| Rank = 0 | -545,2684    | -610,0339                        |
| Rank = 1 | -430,7079    | -498,9305                        |
| Rank = 2 | -370,9601    | -430,1561                        |
| Rank = 3 | -343,3580    | -357,0737                        |
| Rank = 4 | -321,2976    | -331,2666                        |
| Rank = 5 | -308,4064    | -317,5134                        |
| Rank = 6 | -305,2641    | -313,1086                        |
| Rank = 7 | -302,8724    | -311,6426                        |

Skenario 3 merupakan data simulasi polong kacang hijau yang telah dimodifikasi dengan *outlier* dan memiliki kondisi *excess zeros*.

Tabel 4.24 Nilai *Log-Likelihood* skenario 4 dengan pendekatan *rcim*

| Model    | Regresi ZINB | Regresi <i>Negative Binomial</i> |
|----------|--------------|----------------------------------|
| Null     | -296,9352    | -332,3046                        |
| Baris    | -268,5822    | -276,5002                        |
| Kolom    | -258,3841    | -328,1949                        |
| Rank = 0 | -266,9144    | -238,8735                        |
| Rank = 1 | -204,4886    | -200,5827                        |
| Rank = 2 | -181,8155    | -166,2092                        |
| Rank = 3 | -163,6287    | -156,6098                        |
| Rank = 4 | -156,7656    | -156,5977                        |

Skenario 4 merupakan data serangan hama pada tanaman kacang hijau yang terdapat pengamatan nol satu baris pada data dan dimodifikasi dengan *outlier*.

Pada skenario 3 menunjukkan Nilai *Log-Likelihood* RCIM ZINB lebih besar daripada RCIM *Negative Binomial* sedangkan pada skenario 4 menunjukkan Nilai *Log-Likelihood* RCIM ZINB pada tiap rank lebih kecil daripada RCIM *Negative Binomial*. Nilai *Log-Likelihood* yang didapatkan dari masing-masing model selanjutnya digunakan untuk mencari nilai model devians. Cara mengitung nilai model devians untuk mengetahui analisis residu devians pada masing-masing data. Cara menghitung nilai model devians telah dijelaskan sebelumnya pada subbab 2.6.1 yaitu dua kali perbandingan nilai *Log-Likelihood* dari *saturated* model dengan *null* model. Model devians pada Regresi *Negative Binomial* dan Regresi *Zero Inflated Negative Binomial* hanya didapatkan nilai *Log-Likelihood* oleh karena itu dibutuhkan penghitungan secara manual untuk mendapatkan nilai Model devians. Pada Tabel 4.25 dan Tabel 4.26, berikut menunjukkan hasil dari perhitungan model devians dan juga analisis residu deviansnya.

Tabel 4.25 Nilai Model Devians dan Analisis Residu Devians Pada Skenario 3

| Model    | Regresi ZINB  |                | Regresi <i>Negative Binomial</i> |                |
|----------|---------------|----------------|----------------------------------|----------------|
|          | Model Devians | Residu Devians | Model Devians                    | Residu Devians |
| Null     | 519,2696      | -              | 623,3496                         | -              |
| Baris    | 477,9142      | 41,3554        | 606,3586                         | 16,9910        |
| Kolom    | 339,4056      | 179,8640       | 617,2412                         | 6,1084         |
| Rank = 0 | 484,7920      | -              | 596,7826                         | -              |
| Rank = 1 | 255,6710      | 229,1210       | 374,5758                         | 222,2068       |
| Rank = 2 | 136,1754      | 119,4956       | 237,0270                         | 137,5488       |
| Rank = 3 | 80,9712       | 55,2042        | 90,8622                          | 146,1648       |
| Rank = 4 | 36,8504       | 44,1208        | 39,2480                          | 51,6142        |
| Rank = 5 | 11,0680       | 25,7824        | 11,7416                          | 27,5064        |
| Rank = 6 | 4,7834        | 6,2846         | 2,9320                           | 8,8096         |
| Rank = 7 | 0             | 4,7834         | 0                                | 2,9320         |

Tabel 4.26 Nilai Model Devians dan Analisis Residu Devians Pada Skenario 4

| Model    | Regresi ZINB  |                | Regresi <i>Negative Binomial</i> |                |
|----------|---------------|----------------|----------------------------------|----------------|
|          | Model Devians | Residu Devians | Model Devians                    | Residu Devians |
| Null     | 280,3392      | -              | 351,4138                         | -              |
| Baris    | 223,6332      | 56,7060        | 239,8050                         | 111,6088       |
| Kolom    | 203,2370      | 77,1022        | 343,1944                         | 8,2194         |
| Rank = 0 | 220,2976      | -              | 164,5516                         | -              |
| Rank = 1 | 95,4460       | 124,8516       | 87,9700                          | 76,5816        |
| Rank = 2 | 50,0998       | 45,3462        | 19,2230                          | 68,7470        |
| Rank = 3 | 13,7262       | 36,3736        | 0,0242                           | 19,1988        |
| Rank = 4 | 0             | 13,7262        | 0                                | 0,0242         |

Nilai Residu devians baris dan kolom diperoleh dari selisih antara nilai model devians *Null* model dengan model devians dari model baris dan kolom. Sedangkan untuk nilai residu devians pada model rank = 1 diperoleh dari selisih antara nilai model devians dari model rank = 0 dengan model rank = 1. Nilai residu devians untuk model rank = 2 diperoleh dari selisih antara model devians dari model rank = 1 dengan model rank = 2. Begitu seterusnya hingga model rank maksimum. Sehingga dari nilai residu devians model RCIM diatas diperoleh analisis residu devians sebagai berikut:



Tabel 4.27 Analisis Residu Devians RCIM6 (Skenario 3, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 222,2068          | 11,11034          | 6,99148              | 0,004130 | **           |
| RCIM 2   | 18 | 137,5488          | 7,64160           | 4,80868              | 0,014570 | *            |
| RCIM 3   | 16 | 146,1648          | 9,13530           | 5,74864              | 0,008500 | **           |
| RCIM 4   | 14 | 51,6142           | 3,68673           | 2,31997              | 0,116737 | ns           |
| RCIM 5   | 12 | 27,5064           | 2,29220           | 1,44243              | 0,307932 | ns           |
| RCIM 6   | 10 | -0,9714           | -0,09714          | -0,06113             |          | 1 ns         |
| Residual | 8  | 12,7130           | 1,58913           |                      |          |              |

Tabel 4.28 Analisis Residu Devians RCIM3 (Skenario 3, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 222,2068          | 11,11034          | 5,38018              | 1,65E-06 | ***          |
| RCIM 2   | 18 | 137,5488          | 7,64160           | 3,70044              | 0,000202 | ***          |
| RCIM 3   | 16 | 146,1648          | 9,13530           | 4,42377              | 4,47E-05 | ***          |
| Residual | 44 | 90,8622           | 2,06505           |                      |          |              |

Tabel 4.29 Analisis Residu Devians RCIM2 (Skenario 3, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 20 | 222,2068          | 11,11034          | 2,81242              | 0,001053 | **           |
| RCIM 2   | 18 | 137,5488          | 7,64160           | 1,93436              | 0,029651 | *            |
| Residual | 60 | 237,0270          | 3,95045           |                      |          |              |

Tabel 4.30 Analisis Residu Devians RCIM1 (Skenario 3, *Negative Binomial*)

|                   | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|-------------------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1            | 20 | 222,2068          | 11,11034          | 2,31357              | 0,004617       | **           |
| Residual          | 78 | 374,5758          | 4,80225           |                      |                |              |
| Kode signifikan : |    | 0 – 0,001         | “***”             |                      |                |              |
|                   |    | 0,001 – 0,01      | “**”              |                      |                |              |
|                   |    | 0,01 – 0,05       | “*”               |                      |                |              |

Untuk mencari model terbaik dilakukan analisis nilai Rasio Rataan Devians. Model RCIM dikatakan signifikan jika  $p\text{-value} < 0,05$ . Hasil analisis residu devians *Negative Binomial* pada hasil pendugaan skenario 3 tabel 4.27 menunjukkan model RCIM6 dengan 6 komponen utama, RCIM5 dengan 5 komponen utama, dan RCIM4 dengan 4 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . Pada tabel 4.28 menunjukkan model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$  begitu juga dengan dan tabel 4.29 menunjukkan model RCIM2 dengan 2 komponen utama signifikan dimana  $p\text{-value} < 0,05$  dan tabel 4.30 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 3 tabel yang signifikan menunjukkan bahwa RCIM3 dengan 3 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM *Negative Binomial*. Untuk pendugaan analisis residu devians RCIM ZINB pada skenario 3 akan ditunjukkan pada sebagai berikut:

Tabel 4.31 Analisis Residu Devians RCIM6 (Skenario 3, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 134,4232          | 6,72116           | 23,47389             | 5,00E-05       | ***          |
| RCIM 2   | 18 | 59,4354           | 3,30197           | 11,53223             | 0,000738       | ***          |
| RCIM 3   | 16 | 69,2432           | 4,32770           | 15,11464             | 0,000290       | ***          |
| RCIM 4   | 14 | 32,7006           | 2,33576           | 8,15771              | 0,002798       | **           |
| RCIM 5   | 12 | 9,2524            | 0,77103           | 2,69286              | 0,083596       | ns           |
| RCIM 6   | 10 | 4,9652            | 0,49652           | 1,73411              | 0,223518       | ns           |
| Residual | 8  | 2,2906            | 0,28633           |                      |                |              |

Tabel 4.32 Analisis Residu Devians RCIM4 (Skenario 3, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 134,4232          | 6,72116           | 12,21422             | 1,90E-09       | ***          |
| RCIM 2   | 18 | 59,4354           | 3,30197           | 6,00059              | 9,12E-06       | ***          |
| RCIM 3   | 16 | 69,2432           | 4,32770           | 7,86464              | 7,86E-07       | ***          |
| RCIM 4   | 14 | 32,7006           | 2,33576           | 4,24472              | 0,000438       | ***          |
| Residual | 30 | 16,5082           | 0,55027           |                      |                |              |

Tabel 4.33 Analisis Residu Devians RCIM3 (Skenario 3, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 134,4232          | 6,72116           | 6,00972              | 3,70E-07       | ***          |
| RCIM 2   | 18 | 59,4354           | 3,30197           | 2,95245              | 0,001748       | **           |
| RCIM 3   | 16 | 69,2432           | 4,32770           | 3,86961              | 0,000186       | ***          |
| Residual | 44 | 49,2088           | 1,11838           |                      |                |              |

Tabel 4.34 Analisis Residu Devians RCIM2 (Skenario 3, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 134,4232          | 6,72116           | 3,40450              | 0,000125       | ***          |
| RCIM 2   | 18 | 59,4354           | 3,30197           | 1,67256              | 0,070735       | ns           |
| Residual | 60 | 118,4520          | 1,97420           |                      |                |              |

Tabel 4.35 Analisis Residu Devians RCIM1 (Skenario 3, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 20 | 134,4232          | 6,72116           | 2,94709              | 0,000348       | ***          |
| Residual | 78 | 177,8874          | 2,28061           |                      |                |              |

Hasil analisis residu devians ZINB pada hasil pendugaan skenario 3 tabel 4.31 menunjukkan bahwa model RCIM6 dengan 6 komponen utama dan model RCIM5 dengan 5 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . Tabel 4.32 menunjukkan model RCIM4 dengan 4 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Tabel 4.33 menunjukkan model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Pada Tabel 4.34 menunjukkan model RCIM2 dengan 2 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ , dan Tabel 4.35 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 1 tabel yang signifikan menunjukkan bahwa RCIM1 dengan 1 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM ZINB. Untuk pendugaan analisis residu devians RCIM *Negative Binomial* pada skenario 4 akan ditunjukkan pada sebagai berikut:

Tabel 4.36 Analisis Residu Devians RCIM3 (Skenario 4, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 14 | 76,5816           | 5,47011           | 1808,3022            | 2,09E-12       | ***          |
| RCIM 2   | 12 | 68,7470           | 5,72892           | 1893,8567            | 1,93E-12       | ***          |
| RCIM 3   | 10 | 19,1988           | 1,91988           | 634,6711             | 1,75E-10       | ***          |
| Residual | 8  | 0,0242            | 0,00303           |                      |                |              |

Tabel 4.37 Analisis Residu Devians RCIM2 (Skenario 4, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 14 | 76,5816           | 5,47011           | 5,12210              | 0,000805       | ***          |
| RCIM 2   | 12 | 68,7470           | 5,72892           | 5,36443              | 0,000766       | ***          |
| Residual | 18 | 19,2230           | 1,06794           |                      |                |              |

Tabel 4.38 Analisis Residu Devians RCIM1 (Skenario 4, *Negative Binomial*)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | <i>p-value</i> | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------------|--------------|
| RCIM 1   | 14 | 76,5816           | 5,47011           | 1,86545              | 0,074425       | ns           |
| Residual | 30 | 87,9700           | 2,93233           |                      |                |              |

Hasil analisis residu devians *Negative Binomial* pada hasil pendugaan skenario 4 tabel 4.36 menunjukkan bahwa model RCIM3 dengan 3 komponen utama signifikan dimana  $p\text{-value} < 0,05$ , begitu juga dengan tabel 4.37 menunjukkan model RCIM2 dengan 2 komponen utama signifikan dimana  $p\text{-value} < 0,05$ , namun pada tabel 4.38 model RCIM1 dengan 1 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . karena RCIM1 dengan 1 komponen utama tidak signifikan menunjukan bahwa model RCIM *Negative Binomial* tidak signifikan. Untuk pendugaan analisis residu devians RCIM ZINB pada skenario 4 akan ditunjukkan pada sebagai berikut:



Tabel 4.39 Analisis Residu Devians RCIM3 (Skenario 4, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 14 | 124,8516          | 8,91797           | 5,19764              | 0,012376 | *            |
| RCIM 2   | 12 | 45,3462           | 3,77885           | 2,20242              | 0,134361 | ns           |
| RCIM 3   | 10 | 36,3736           | 3,63736           | 2,11995              | 0,149617 | ns           |
| Residual | 8  | 13,7262           | 1,71578           |                      |          |              |

Tabel 4.40 Analisis Residu Devians RCIM1 (Skenario 4, ZINB)

|          | Db | Residu<br>Devians | Rataan<br>Devians | Ras. Rat.<br>Devians | p-value  | taraf<br>sig |
|----------|----|-------------------|-------------------|----------------------|----------|--------------|
| RCIM 1   | 14 | 124,8516          | 8,91797           | 2,80304              | 0,008724 | **           |
| Residual | 30 | 95,4460           | 3,18153           |                      |          |              |

Hasil analisis residu devians ZINB pada hasil pendugaan skenario 4 tabel 4.39 menunjukkan bahwa model RCIM3 dengan 3 komponen utama dan model RCIM2 dengan 2 komponen utama tidak signifikan dimana  $p\text{-value} < 0,05$ . Tabel 4.40 menunjukkan model RCIM1 dengan 1 komponen utama signifikan dimana  $p\text{-value} < 0,05$ . Dengan adanya 1 tabel yang signifikan menunjukkan bahwa RCIM1 dengan 1 pengaruh interaksi menjadi model yang paling layak dalam menduga data dengan menggunakan RCIM ZINB. Hasil pendugaan skenario 3 dengan RCIM *Negative Binomial* dan RCIM ZINB lebih baik dengan menggunakan RCIM ZINB. Dapat dilihat pada Tabel 4.23 yang menunjukkan besarnya nilai *log-Likelihood* dari kedua metode. Begitu juga dengan Tabel 4.25 yang menunjukkan analisis deviansnya. Tabel 4.41 berikut merupakan tabel besar MSE dari kedua metode. Tabel tersebut menunjukkan bahwa MSE ZINB lebih baik dari MSE *Negative Binomial*. Sehingga dapat disimpulkan bahwa RCIM ZINB dapat menduga lebih baik daripada RCIM *Negative Binomial* pada data yang mengandung *outlier* dan *excess zeros*. Pada

tabel 4.21 skenario 2 dan skenario 3 menunjukkan ada kesamaan nilai MSE pada RCIM2, RCIM5, dan RCIM6. Untuk nilai MSE *Negative Binomial* antara skenario 2 dan skenario 3 lebih baik skenario 3.

Tabel 4.41 Nilai *Mean Square Error* Skenario 3 dengan pendekatan *rcim*

| Model  | MSE ZINB          | MSE <i>Negative Binomial</i> |
|--------|-------------------|------------------------------|
| RCIM 1 | 0,08142671        | 0,25555470                   |
| RCIM 2 | <b>0,07504630</b> | 0,14247590                   |
| RCIM 3 | 0,05165929        | 0,07846524                   |
| RCIM 4 | 0,05012614        | 0,05101145                   |
| RCIM 5 | <b>0,00899084</b> | 0,02153536                   |
| RCIM 6 | <b>0,00535431</b> | 0,01880404                   |

Hasil pendugaan skenario 4 dengan RCIM *Negative Binomial* dan RCIM ZINB lebih baik dengan menggunakan RCIM ZINB dikarenakan RCIM *Negative Binomial* tidak signifikan. Tabel 4.42 berikut merupakan tabel besar MSE dari kedua metode. Tabel tersebut menunjukkan bahwa MSE *Negative Binomial* lebih baik dari MSE ZINB. Untuk MSE ZINB dan *Negative Binomial* pada data 2 dan skenario 4 lebih baik nilai MSE data 2.

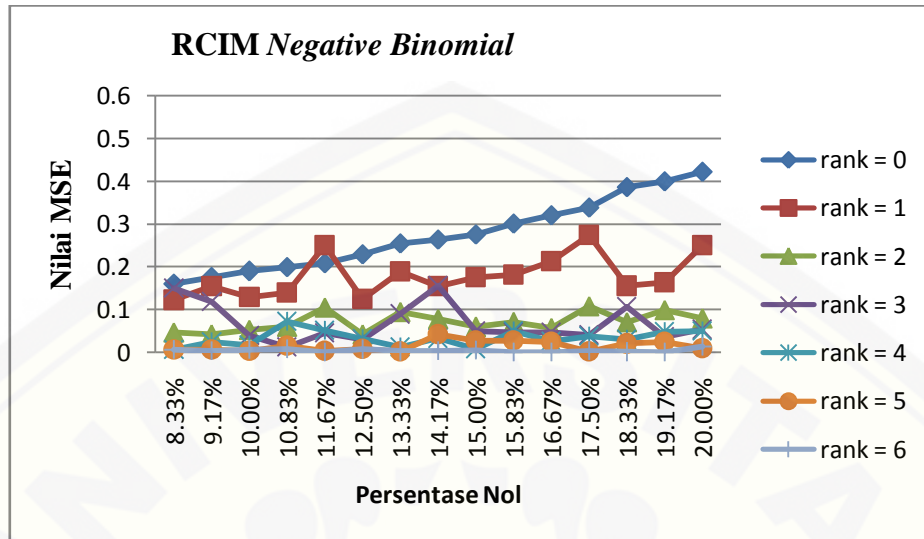
Tabel 4.42 Nilai *Mean Square Error* Skenario 4 dengan Pendekatan *rcim*

| Model  | MSE ZINB   | MSE <i>Negative Binomial</i> |
|--------|------------|------------------------------|
| RCIM 1 | 0,03839160 | 0,02187914                   |
| RCIM 2 | 0,05245484 | 0,02145327                   |
| RCIM 3 | 0,01474227 | 1,06e-05                     |

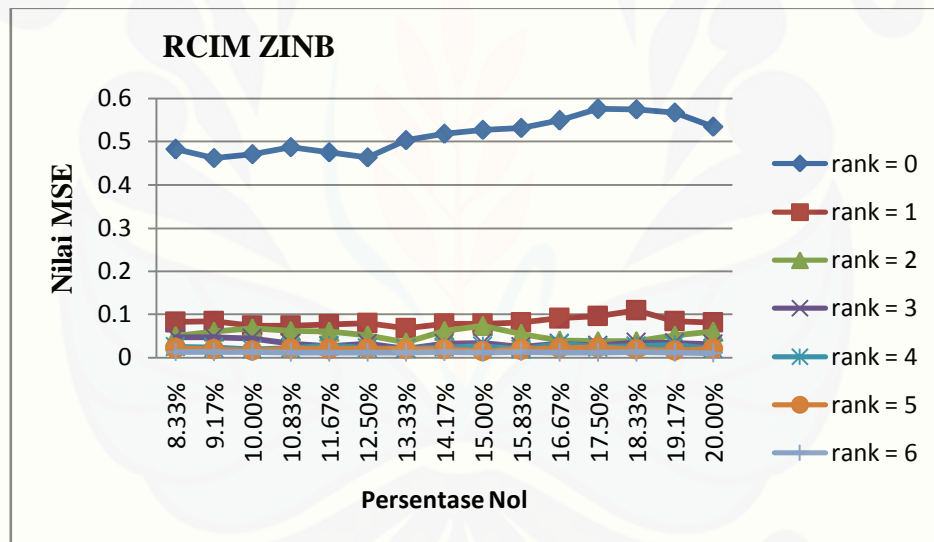
#### 4.6 Pengaruh Pengamatan Nol pada Model *Negative Binomial*

Pembahasan ini dilakukan guna mendapatkan informasi tentang seberapa kecil prosentase pengamatan nol yang terdapat pada data sehingga pendugaan dengan Regresi *Negative Binomial* tidak baik lagi. Analisis dilakukan sebatas pada besar nilai MSE yang dihasilkan pada masing-masing metode. Data yang digunakan adalah data dengan *outlier* yang berdistribusi *Negative Binomial* (skenario 1) yang telah dikontaminasi oleh pengamatan nol.

Proses pemberian pengamatan nol bertahap sehingga terdapat pada setiap kolom yang memiliki nilai pengamatan terendah. Pengamatan nol dimulai dari 10 nol atau prosentasenya 8,33% dikarenakan program ZINB tidak dapat menganalisanya. Proses pemilihan kolom dilakukan secara acak. Pertama-tama memilih kolom hingga setiap kolom memiliki pengamatan nol. Apabila terdapat 2 data yang sama-sama memiliki nilai pengamatan terendah, maka akan di pilih salah satu secara acak, sampai terdapat 24 data yang akan diduga. Sedangkan besar prosentase pengamatan nol pada data ialah 8,33% hingga 20%. Gambar 4.1 dan Gambar 4.2 berikut merupakan grafik tren kenaikan/penurunan nilai MSE pada model RCIM *Negative Binomial* dan RCIM ZINB.

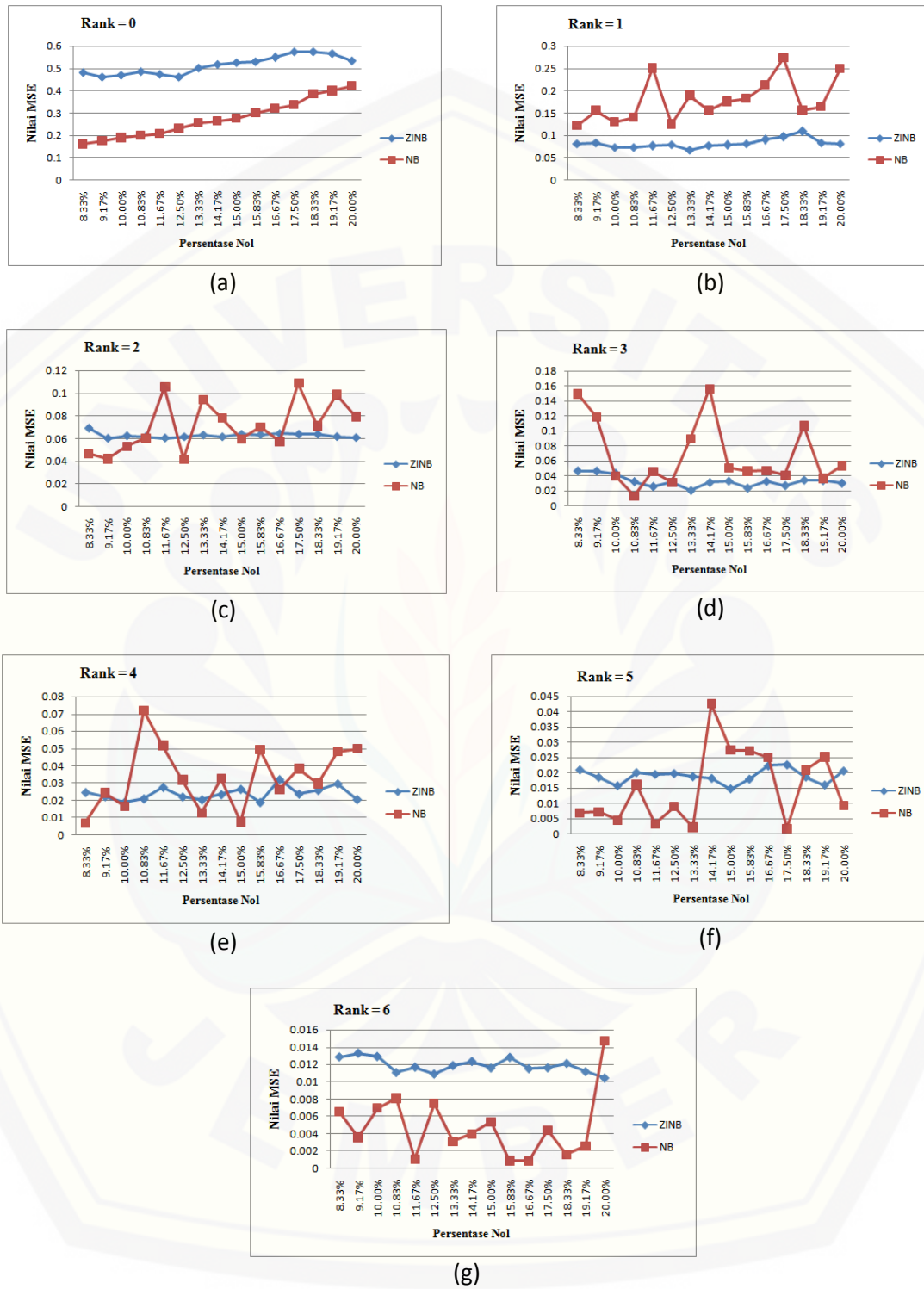


(a)



(b)

Gambar 4.1 Grafik Tren Penurunan/Kenaikan MSE RCIM



Gambar 4.2 MSE ZINB dan *Negative Binomial*



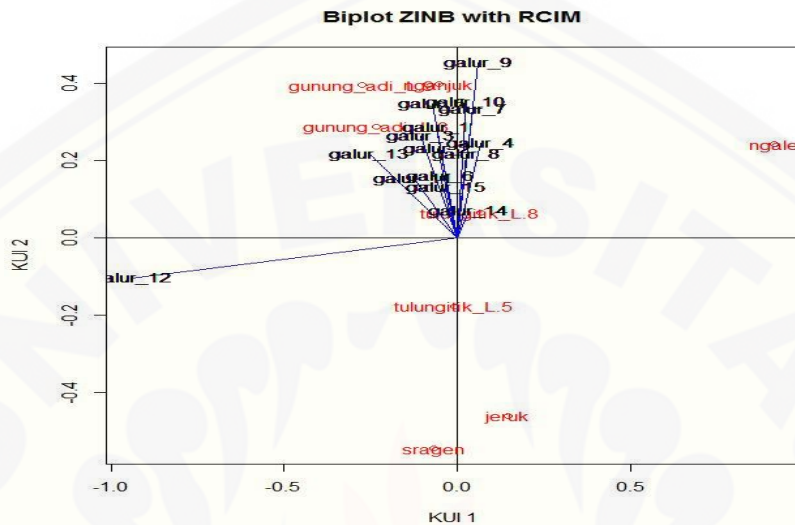
Gambar 4.1 (a) menunjukkan bahwa penambahan pengamatan nol pada data mengakibatkan nilai MSE model RCIM *Negative Binomial* meningkat. Pada rank = 0 MSE nya cenderung meningkat stabil. Sedangkan pada rank = 1 sampai rank = 6 terjadi peningkatan dan penurunan MSE namun semakin besar persentase pengamatan nol, maka grafik MSE cenderung meningkat.

Gambar 4.2 menunjukkan grafik MSE kedua model pada masing-masing rank. Pada gambar 4.2 (a) MSE *Negative Binomial* meningkat stabil, pada gambar 4.2 (b) MSE *Negative Binomial* naik turun agak tajam dan tidak stabil namun masih lebih baik MSE ZINB. Pada gambar 4.2 (c) hingga (e) terlihat beberapa MSE *Negative Binomial* lebih baik dari MSE ZINB namun grafiknya naik turun agak tajam dan tidak stabil. Pada gambar 4.2 (f) dan (g) MSE *Negative Binomial* lebih baik dari MSE ZINB namun grafiknya naik turun dengan tajam dan tidak stabil. Dapat dilihat bahwa RCIM ZINB dapat memberikan nilai MSE yang lebih baik daripada RCIM *Negative Binomial* pada semua jenis model dikarenakan RCIM ZINB dapat memberikan nilai MSE yang stabil pada setiap data sedangkan RCIM *Negative Binomial* memberikan nilai MSE yang tidak stabil. Informasi lain yang dapat diambil ialah RCIM *Negative Binomial* dapat memberikan MSE yang cukup kecil apabila model yang digunakan semakin sederhana. Oleh karena itu ketidakstabilan nilai MSE dari RCIM *Negative Binomial* menimbulkan kekhawatiran dalam penggunaannya pada data dengan kasus *excess zeros*.

#### 4.7 Analisis Biplot RCIM

Langkah selanjutnya adalah menginterpretasikan model RCIM pada biplot. Pada data kedua akan diinterpretasikan ke dalam biplot untuk menunjukkan kestabilan Galur pada Lingkungan yang beragam. Berdasarkan pendekatan *Row Column Interaction Model* (RCIM) pada data 2 ditunjukkan bahwa model dibagi menjadi 7 model yang diperlihatkan pada pembagian Rank yang sebanyak 7 pada model multiplikatif. Namun untuk menginterpretasikan ke dalam biplot digunakan nilai residual model pada saat Rank = 0 (model aditif) karena pada model RCIM yang

diperlihatkan dalam analisis adalah interaksi multiplikatifnya. Hasil Interpretasi biplot RCIM ditunjukkan pada Gambar 4.3 sebagai berikut.



Gambar 4.3 Biplot RCIM2 Data Simulasi Polong Isi Kedelai (73,94%)

Biplot RCIM2 menunjukkan plot 2 dimensi dari model RCIM. Berdasarkan Gambar 4.3 terlihat bahwa Komponen Utama 1 menunjukkan informasi sebesar 45,76%, sedangkan komponen utama 2 menunjukkan informasi sebesar 28,18%. Sehingga dapat dari visualisasi biplot pada data polong isi kedelai secara keseluruhan menunjukkan informasi sebesar 73,94% yang berarti Gambar 4.3 menerangkan informasi sebesar 73,94 % dari total keragaman data yang sebenarnya.

Mattjik dan Sumertadjaya (2011) menyatakan bahwa objek yang hampir berada di titik pusat berarti objek tersebut memiliki nilai dekat dengan rata-rata. Galur 14 dapat beradaptasi dengan baik pada semua lokasi dengan ditunjukkan pada gambar 4.3 bahwa galur 14 berada dekat dengan titik pusat.

## BAB 5 PENUTUP

### 5.1 Kesimpulan

Hasil analisis yang dilakukan pada kedua data menghasilkan beberapa kesimpulan sebagai berikut :

- a. Hasil analisis devian model RCIM pada data cacah yang mengalami kondisi *excess zeros* dan terdapat *outlier* serta terdapat pengamatan nol satu baris menunjukkan bahwa model RCIM *Negative Binomial* tidak signifikan pada semua RCIM, sedangkan ZINB memberikan kelayakan model RCIM1. Pada data cacah yang terdapat *outlier* dan mengalami *excess zeros*, pendugaan dengan ZINB memberikan model RCIM yang lebih sederhana daripada *Negative Binomial*. Hal ini ditunjukkan oleh kelayakan model pada data cacah yang terdapat *outlier* dan mengalami *excess zeros* adalah model RCIM3 jika diduga dengan *Negative Binomial* sedangkan ZINB memberikan kelayakan model RCIM1.
- b. Hasil perbandingan nilai MSE dari masing–masing model RCIM menunjukkan pada data cacah yang mengalami kondisi *excess zeros* dan terdapat *outlier* serta pengamatan nol satu baris fungsi distribusi *Negative Binomial* lebih baik dalam menduga data yang mengalami kasus *excess zeros* dari pada dengan menggunakan fungsi distribusi *Zero Inflated Negative Binomial*. Sedangkan pada data cacah yang terdapat *outlier* dan mengalami *excess zeros* fungsi distribusi *Zero Inflated Negative Binomial* lebih baik dalam menduga data cacahan dari pada dengan menggunakan fungsi distribusi *Negative Binomial*.
- c. Pada pengujian pengaruh pengamatan nol menghasilkan bahwa pendugaan dengan menggunakan fungsi distribusi *Zero Inflated Negative Binomial* memberikan besar MSE yang lebih kecil daripada dengan menggunakan fungsi distribusi *Negative Binomial*. Hal lain yang dapat dilihat pada Gambar 4.1 (a)

bahwa *Negative Binomial* tidak stabil dalam menduga data yang mengandung pengamatan nol. Tidak seperti *Negative Binomial*, *Zero Inflated Negative Binomial* dapat menduga data dengan stabil pada data yang mengandung pengamatan nol.

- d. Pengaruh data cacahan yang mengalami kondisi *outlier* kemudian ditambah dengan *excess zeros* yaitu pada model *Negative Binomial* nilai MSE-nya lebih baik dan analisis deviansnya lebih sederhana dengan *outlier* dan *excess zeros*, pada model *Zero Inflated Negative Binomial* juga memberikan analisis devians yang lebih sederhana dan memberikan nilai MSE yang sama dengan data cacahan yang mengandung *outlier* dan *excess zeros*.
- e. Pengaruh data cacahan yang mengalami kondisi *excess zeros* kemudian ditambah dengan *outlier* yaitu pada model *Negative Binomial* nilai MSE-nya lebih baik dan analisis deviansnya lebih sederhana dengan *excess zeros* saja tanpa *outlier*, pada model *Zero Inflated Negative Binomial* memberikan analisis devians yang sama yaitu RCIM1 dan memberikan nilai MSE yang lebih baik dengan data cacahan yang mengandung *excess zeros*.

## 5.2 Saran

Setiap hasil penelitian selalu memberikan saran untuk dapat mengembangkan penelitian – penelitian sebelumnya. Adapun saran dari peneliti ialah :

1. Penelitian selanjutnya dapat membandingkan kasus *excess zeros* untuk distribusi yang berbeda seperti ZINB (*Zero-Inflated Negative Binomial*) dan ZIP (*Zero-Inflated Poisson*) ataupun yang lainnya.
2. Peneliti juga menyarankan untuk meningkatkan dimensi plot menjadi 3 (3 dimensi) agar pengaruh interaksi dapat dijelaskan lebih detail lagi.



**DAFTAR PUSTAKA**

- Ahadiyah, K. 2015. "Pendekatan *Row Column Interaction Model* untuk Model *Generalized AMMI (GAMMI)* Pada Data Berdistribusi Poisson dengan *Log-Likelihood* dan Analisis Devian". Tidak diterbitkan. Skripsi. Jember : FMIPA Universitas Jember.
- Ariawan, B., Suparti, dan Sudarno. 2012. Pemodelan Regresi *Zero-Inflated Negative Binomial (ZINB)* Untuk Data Repon Diskrit dengan *Exceeds Zeros*. *Jurnal GAUSSIAN* vol. **1 (1)** : 55-64.
- Hadi A. F., Mattjik, A. A., dan Sumertajaya, I. M. 2010. Generalized AMMI Models for Assessing The Endurance of Soybean to Leaf Pest. *Jurnal ILMU DASAR 2* : 151-159.
- Hawkins, D. 1980. *Identification of Outliers*. Chapman and Hall.
- Hilbe, J.M. 2011. *Negative Binomial Regression Second Edition*. New York: Cambridge University Press.
- Kriegel, H. P. 2010. *Outlier Detection Techniques*. Munich: University Munchen.
- Mattjik, A. A & Sumertajaya, I. M. 2011. *Sidik Peubah Ganda dengan Menggunakan SAS*. Departemen Statistika IPB: IPB Press.
- McCullagh, P. & Nelder, J.A. 1989. *Generalized Linier Models, second edition*. London: Chapman & Hall.
- Pawitan, Y. 2011. *In All Likelihood : Statistical Modelling and Inference Using Likelihood*. Ireland : Clarendon Press. Oxford



- Susilowati, U. 2015. “Penangan Overdispersi pada Pemodelan Data Cacahan Menggunakan Regresi Binomial Negatif pada RCIM (*Row Column Interaction Model*)”. Tidak diterbitkan. Skripsi. Jember : FMIPA Universitas Jember.
- Tirta, I.M. 2009. *Analisis Regresi dengan R*. Jember : UNEJ Press.
- Van Eeuwijk, F.A. 1995. Multiplicative Interaction in Generalized Linear Models. *Biometrics*, **51**, 1017-1032
- Wicaksono, D. B. C. 2015. “Perbandingan Model GAMMI Poisson dan GAMMI *Zero-Inflated Poisson* dengan Pendekatan RCIM (*Row Column Interaction Models*)”. Tidak diterbitkan. Skripsi. Jember : FMIPA Universitas Jember.
- Yee, T.W., Hastie, T.J. 2003. Reduced-rank Vector Generalized Linear Models. *Statistical Modelling* **3**, 15-41.
- Yee, T.W., dan Hadi, A.F. 2014. Row-Column Interaction Models, with an R Implementation. *Comput Stat* DOI 10.1007/s00180-014-0499-9.

**LAMPIRAN**

**A. DATA PENELITIAN**

**A.1 Data Jumlah Polong Isi Kedelai Ahadiyah (2015)**

Tabel A.1 Jumlah Polong Isi Kedelai (Data 1)

|          | L1 | L2  | L3 | L4  | L5  | L6  | L7 | L8 |
|----------|----|-----|----|-----|-----|-----|----|----|
| galur_1  | 84 | 118 | 85 | 155 | 115 | 44  | 63 | 49 |
| galur_2  | 57 | 118 | 85 | 176 | 115 | 49  | 80 | 35 |
| galur_3  | 61 | 101 | 82 | 197 | 110 | 54  | 59 | 38 |
| galur_4  | 94 | 122 | 53 | 189 | 130 | 43  | 62 | 41 |
| galur_5  | 72 | 113 | 48 | 195 | 119 | 44  | 53 | 37 |
| galur_6  | 83 | 159 | 72 | 201 | 146 | 115 | 50 | 73 |
| galur_7  | 75 | 137 | 70 | 156 | 125 | 30  | 43 | 41 |
| galur_8  | 56 | 119 | 75 | 160 | 139 | 38  | 51 | 49 |
| galur_9  | 58 | 95  | 87 | 178 | 143 | 26  | 35 | 36 |
| galur_10 | 71 | 113 | 63 | 212 | 155 | 35  | 31 | 37 |
| galur_11 | 66 | 146 | 72 | 214 | 184 | 67  | 45 | 59 |
| galur_12 | 53 | 111 | 79 | 137 | 242 | 15  | 46 | 27 |
| galur_13 | 79 | 195 | 96 | 236 | 217 | 59  | 50 | 54 |
| galur_14 | 42 | 99  | 57 | 166 | 148 | 25  | 31 | 46 |
| galur_15 | 42 | 96  | 68 | 103 | 148 | 16  | 40 | 37 |

Tabel A.2 Jumlah Polong Isi Kedelai dengan *Outlier* (Skenario 1)

|          | L1  | L2  | L3  | L4  | L5  | L6  | L7 | L8 |
|----------|-----|-----|-----|-----|-----|-----|----|----|
| galur_1  | 84  | 118 | 85  | 167 | 115 | 44  | 63 | 49 |
| galur_2  | 57  | 118 | 85  | 189 | 115 | 49  | 87 | 35 |
| galur_3  | 61  | 101 | 82  | 210 | 110 | 54  | 59 | 38 |
| galur_4  | 101 | 122 | 53  | 203 | 130 | 43  | 62 | 41 |
| galur_5  | 72  | 113 | 48  | 209 | 119 | 44  | 53 | 37 |
| galur_6  | 83  | 159 | 72  | 215 | 146 | 124 | 50 | 80 |
| galur_7  | 75  | 137 | 70  | 169 | 125 | 30  | 43 | 41 |
| galur_8  | 56  | 119 | 75  | 174 | 139 | 38  | 51 | 49 |
| galur_9  | 58  | 95  | 87  | 193 | 143 | 26  | 35 | 36 |
| galur_10 | 71  | 113 | 63  | 227 | 155 | 35  | 31 | 37 |
| galur_11 | 66  | 146 | 72  | 229 | 184 | 67  | 45 | 59 |
| galur_12 | 53  | 111 | 79  | 137 | 284 | 15  | 46 | 27 |
| galur_13 | 79  | 205 | 103 | 279 | 217 | 59  | 50 | 54 |
| galur_14 | 42  | 99  | 57  | 181 | 148 | 25  | 31 | 46 |
| galur_15 | 42  | 96  | 68  | 103 | 161 | 16  | 40 | 37 |

Tabel A.3 Jumlah Polong Isi Kedelai dengan *Excess Zeros* (Skenario 2)

|          | L1 | L2  | L3 | L4  | L5  | L6  | L7 | L8 |
|----------|----|-----|----|-----|-----|-----|----|----|
| galur_1  | 84 | 118 | 85 | 0   | 115 | 44  | 63 | 49 |
| galur_2  | 57 | 118 | 85 | 176 | 115 | 49  | 80 | 0  |
| galur_3  | 61 | 101 | 82 | 197 | 110 | 54  | 59 | 38 |
| galur_4  | 94 | 122 | 0  | 189 | 130 | 43  | 62 | 41 |
| galur_5  | 72 | 113 | 0  | 195 | 119 | 44  | 53 | 37 |
| galur_6  | 83 | 159 | 72 | 201 | 146 | 115 | 50 | 73 |
| galur_7  | 75 | 137 | 70 | 156 | 125 | 30  | 43 | 41 |
| galur_8  | 0  | 119 | 75 | 160 | 139 | 38  | 51 | 49 |
| galur_9  | 58 | 0   | 87 | 178 | 143 | 26  | 0  | 0  |
| galur_10 | 71 | 113 | 63 | 212 | 155 | 35  | 0  | 37 |
| galur_11 | 66 | 146 | 72 | 214 | 184 | 67  | 45 | 59 |
| galur_12 | 0  | 111 | 79 | 0   | 242 | 0   | 46 | 0  |
| galur_13 | 79 | 195 | 96 | 236 | 217 | 59  | 50 | 54 |
| galur_14 | 0  | 0   | 0  | 166 | 148 | 0   | 0  | 46 |
| galur_15 | 0  | 0   | 68 | 0   | 148 | 0   | 0  | 37 |

Tabel A.4 Jumlah Polong Isi Kedelai dengan *Outlier* dan *Excess Zeros* (Skenario 3)

|          | L1  | L2  | L3  | L4  | L5  | L6  | L7 | L8 |
|----------|-----|-----|-----|-----|-----|-----|----|----|
| galur_1  | 84  | 118 | 85  | 0   | 0   | 44  | 63 | 49 |
| galur_2  | 57  | 118 | 85  | 189 | 0   | 49  | 87 | 0  |
| galur_3  | 61  | 101 | 82  | 210 | 0   | 54  | 59 | 38 |
| galur_4  | 101 | 122 | 0   | 203 | 130 | 43  | 62 | 41 |
| galur_5  | 72  | 113 | 0   | 209 | 0   | 44  | 53 | 37 |
| galur_6  | 83  | 159 | 72  | 215 | 146 | 124 | 50 | 80 |
| galur_7  | 75  | 137 | 70  | 169 | 125 | 30  | 43 | 41 |
| galur_8  | 0   | 119 | 75  | 174 | 139 | 38  | 51 | 49 |
| galur_9  | 58  | 0   | 87  | 193 | 143 | 26  | 0  | 0  |
| galur_10 | 71  | 113 | 63  | 227 | 155 | 35  | 0  | 37 |
| galur_11 | 66  | 146 | 72  | 229 | 184 | 67  | 45 | 59 |
| galur_12 | 0   | 111 | 79  | 0   | 284 | 0   | 46 | 0  |
| galur_13 | 79  | 205 | 103 | 279 | 217 | 59  | 50 | 54 |
| galur_14 | 0   | 0   | 0   | 181 | 148 | 0   | 0  | 46 |
| galur_15 | 0   | 0   | 68  | 0   | 161 | 0   | 0  | 37 |

**A.2 Data Serangan Hama Terhadap Tanaman Kacang Hijau**

Tabel A.5 Serangan Hama Terhadap Tanaman Kacang Hijau (Data 2)

|    | G1  | G2  | G3  | G4  | G5  | G6  | G7  | G8  | G9 | G10 | G11 | G12 |
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|-----|-----|-----|
| L1 | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0  | 0   | 0   | 0   |
| L2 | 167 | 217 | 200 | 133 | 200 | 133 | 50  | 50  | 50 | 83  | 0   | 67  |
| L3 | 100 | 250 | 217 | 200 | 200 | 150 | 100 | 67  | 83 | 117 | 50  | 133 |
| L4 | 150 | 233 | 183 | 183 | 233 | 167 | 67  | 100 | 83 | 133 | 33  | 117 |
| L5 | 150 | 250 | 217 | 133 | 367 | 150 | 83  | 67  | 83 | 83  | 33  | 117 |

Tabel A.6 Serangan Hama Terhadap Tanaman Kacang Hijau (Skenario 4)

|    | G1  | G2  | G3  | G4  | G5  | G6  | G7  | G8  | G9  | G10 | G11 | G12 |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| L1 | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| L2 | 192 | 217 | 200 | 133 | 200 | 133 | 50  | 50  | 50  | 83  | 0   | 67  |
| L3 | 100 | 281 | 246 | 227 | 200 | 150 | 119 | 67  | 83  | 117 | 64  | 155 |
| L4 | 150 | 233 | 183 | 183 | 233 | 192 | 67  | 118 | 101 | 155 | 33  | 117 |
| L5 | 150 | 250 | 217 | 133 | 401 | 150 | 83  | 67  | 83  | 83  | 33  | 117 |

## B. Script Data Serangan Hama Terhadap Tanaman Kacang Hijau (Binomial Negatif)

### ###Memodelkan kedalam model RCIM###

```
>library(VGAM)
>modelfull<-
vglm(JUMLAH_POLONG_ISI~1,negbinomial,data=v_data_asli)
>modelfull
>modellokasi<-
vglm(JUMLAH_POLONG_ISI~LOKASI,negbinomial,data=v_data_asli)
>modellokasi
>modelgen<-
vglm(JUMLAH_POLONG_ISI~GENOTYPE,negbinomial,data=v_data_asli)
>modelgen
>a=rcim(data_asli, negbinomial, Svd.arg=TRUE, Rank=0,
trace=TRUE, wzeption = 0.217, maxiter = 1000)
>logLik(a)
>b=rcim(data_asli, negbinomial, Svd.arg=TRUE, Rank=1,
trace=TRUE, wzeption = 0.217, maxiter = 1000)
>logLik(b)
>c=rcim(data_asli, negbinomial, Svd.arg=TRUE, Rank=2,
trace=TRUE, wzeption = 0.217, maxiter = 1000)
>logLik(c)
>d=rcim(data_asli, negbinomial, Svd.arg=TRUE, Rank=3,
trace=TRUE, wzeption = 0.217, maxiter = 1000)
>logLik(d)
>e=rcim(data_asli, negbinomial, Svd.arg=TRUE, Rank=4,
trace=TRUE, wzeption = 0.217, maxiter = 1000)
>logLik(e)
```

### ###Analisis Devian model RCIM###

```
>devh<-348.9018-237.0342
>devg<-348.9018-340.0392
>dbham<-ncol(data_asli)-1
>dbgen<-nrow(data_asli)-1
>t<-min(dbham,dbgen)-1
>dev<-c(1:(t+2))
>dev[1]<-143.9072
>dev[2]<-66.8188
>dev[3]<-22.4762
>dev[4]<-0.0156
>dev[5]<-0
>tabel<-array(0,c((4+t),5,t))
>gammi<-paste(c("RCIM"),1:t)
>gam<-c(1:t)
>for (l in 1:t){
 gam[l]<-paste(c("RCIM"),t-l+1)
```



```

}
>dimnames(tabel)<-
 list(c("Hama", "Genotipe", gammi, "Residual", "Total"),
 c("db", "Devians", "Rataan Devians", "Ras. Rat.
Devians", "Nilai-p"), gam)
>for(k in 1:t){
 tabel[1:2,1,k]<-c(dbham, dbgen)
 tabel[1:2,2,k]<-c(devh, devg)
 tabel[1:2,3,k]<-c(devh/dbham, devg/dbgen)
 t1<-(4+t)-1
 t2<-(4+t)
 n<-(t-k)+1
 for (l in 1:n){
 tabel[2+l,1,k]<-dbham+1+dbgen+1-1-2*1
 tabel[2+l,2,k]<-dev[l]-dev[l+1]
 tabel[2+l,3,k]<-tabel[2+l,2,k]/tabel[2+l,1,k]
 }
 tabel[t1,1,k]<-dbham*dbgen-sum(tabel[3:(2+n),1,k])
 tabel[t1,2,k]<-dev[n+1]
 tabel[t2,1:2,k]<-
c(sum(tabel[1:t1,1,k]), sum(tabel[1:t1,2,k]))
 tabel[t1:t2,3,k]<-
c(tabel[t1,2,k]/tabel[t1,1,k], sum(tabel[1:t1,2,k])/sum(tabel
[1:t1,1,k]))
 for (p in 1:(n+2)){
 tabel[p,4,k]<-tabel[p,3,k]/tabel[t1,3,k]
 tabel[p,5,k]<-
 pf(tabel[p,4,k], tabel[p,1,k], tabel[t1,1,k], lower.tail
 =FALSE)
 }
}
}
>Table

```

### ###Mean Square Error model RCIM###

```

>resia <- residuals(a)
>resiaa <- resia[,c(1,3,5,7,9,11,13,15)]
>msea <- mean(resiaa^2)
>resib <- residuals(b)
>resiba <- resib[,c(1,3,5,7,9,11,13,15)]
>mseb <- mean(resiba^2)
>resic <- residuals(c)
>resica <- resic[,c(1,3,5,7,9,11,13,15)]
>msec <- mean(resica^2)
>resid <- residuals(d)

```

```

>resida <- resid[,c(1,3,5,7,9,11,13,15)]
>msed <- mean(resida^2)
>resie <- residuals(e)
>resiea <- resie[,c(1,3,5,7,9,11,13,15)]
>msee <- mean(resiea^2)
>msea;mseb;msec;mseed;msee

```

### C. Script Data Script Data Serangan Hama Terhadap Tanaman Kacang Hijau (ZINB)

#### ###Memodelkan kedalam model RCIM###

```

>library(VGAM)
>modelfull<-
vglm(JUMLAH_POLONG_ISI~1,zinegbinomial,data=v_data_asli)
>modelfull
>modellokasi<-
vglm(JUMLAH_POLONG_ISI~LOKASI,zinegbinomial,data=v_data_asli)
>modellokasi
>modelgen<-
vglm(JUMLAH_POLONG_ISI~GENOTYPE,zinegbinomial,data=v_data_asli)
)
>modelgen
>a=rcim(data_asli, zinegbinomial, Svd.arg=TRUE, Rank=0,
trace=TRUE, wzeppilon = 0.217, maxiter=1000)
>logLik(a)
>b=rcim(data_asli, zinegbinomial, Svd.arg=TRUE, Rank=1,
trace=TRUE, wzeppilon = 0.217, maxiter=1000)
>logLik(b)
>c=rcim(data_asli, zinegbinomial, Svd.arg=TRUE, Rank=2,
trace=TRUE, wzeppilon = 0.217, maxiter=1000)
>logLik(c)
>d=rcim(data_asli, zinegbinomial, Svd.arg=TRUE, Rank=3,
trace=TRUE, wzeppilon = 0.217, maxiter=1000)
>logLik(d)
>e=rcim(data_asli, zinegbinomial, Svd.arg=TRUE, Rank=4,
trace=TRUE, wzeppilon = 0.217, maxiter=1000)
>logLik(e)

```

#### ###Analisis Devian model RCIM###

```

>devh<-277.0072-220.5468
>devg<-277.0072-178.9656
>dbham<-ncol(data_asli)-1
>dbgen<-nrow(data_asli)-1
>t<-min(dbham,dbgen)-1
>dev<-c(1:(t+2))
>dev[1]<-217.0316
>dev[2]<-81.0468

```

```

>dev[3]<-36.605
>dev[4]<-14.0782
>dev[5]<-0
>tabel<-array(0,c((4+t),5,t))
>gammi<-paste(c("GAMMI"),1:t)
>gam<-c(1:t)
>for (l in 1:t){
 gam[l]<-paste(c("GAMMI"),t-l+1)
}
>dimnames(tabel)<-
 list(c("Hama","Genotipe",gammi,"Residual","Total"),
 c("db","Devians","Rataan Devians","Ras. Rat.
Devians","Nilai-p"),gam)
>for(k in 1:t){
 tabel[1:2,1,k]<-c(dbham,dbgen)
 tabel[1:2,2,k]<-c(devh,devg)
 tabel[1:2,3,k]<-c(devh/dbham,devg/dbgen)
 t1<-(4+t)-1
 t2<-(4+t)
 n<-(t-k)+1
 for (l in 1:n){
 tabel[2+l,1,k]<-dbham+1+dbgen+1-1-2*1
 tabel[2+l,2,k]<-dev[l]-dev[l+1]
 tabel[2+l,3,k]<-tabel[2+l,2,k]/tabel[2+l,1,k]
 }
 tabel[t1,1,k]<-dbham*dbgen-sum(tabel[3:(2+n),1,k])
 tabel[t1,2,k]<-dev[n+1]
 tabel[t2,1:2,k]<-
c(sum(tabel[1:t1,1,k]),sum(tabel[1:t1,2,k]))
 tabel[t1:t2,3,k]<-
c(tabel[t1,2,k]/tabel[t1,1,k],sum(tabel[1:t1,2,k])/sum(tabel
[1:t1,1,k]))
 for (p in 1:(n+2)){
 tabel[p,4,k]<-tabel[p,3,k]/tabel[t1,3,k]
 tabel[p,5,k]<-
 pf(tabel[p,4,k],tabel[p,1,k],tabel[t1,1,k],lower.tail
 =FALSE)
 }
}
>Table

```

**###Script Mean Square Error model RCIM###**

```

>resia <- residuals(a)
>resiaa <- resia[,c(2,5,8,11,14,17,20,23)]
>msea <- mean(resiaa^2)
>resib <- residuals(b)
>resiba <- resib[,c(2,5,8,11,14,17,20,23)]
>mseb <- mean(resiba^2)
>resic <- residuals(c)
>resica <- resic[,c(2,5,8,11,14,17,20,23)]
>msec <- mean(resica^2)
>resid <- residuals(d)
>resida <- resid[,c(2,5,8,11,14,17,20,23)]
>msted <- mean(resida^2)
>resie <- residuals(e)
>resiea <- resie[,c(2,5,8,11,14,17,20,23)]
>msee <- mean(resiea^2)
>msea;mseb;msec;msted;msee

```

**D. Script Data Jumlah Polong Isi Kedelai dengan *excess zeros* (Binomial Negatif)****###Memodelkan kedalam model RCIM###**

```

>library(VGAM)
>modelfull<-
vglm(JUMLAH_POLONG_ISI~1,negbinomial,data=v_polong_isi_over_ze
ros)
>modelfull
>modellokasi<-
vglm(JUMLAH_POLONG_ISI~LOKASI,negbinomial,data=v_polong_isi_ov
er_zeros)
>modellokasi
>modelgen<-
vglm(JUMLAH_POLONG_ISI~GENOTYPE,negbinomial,data=v_polong_isi_
over_zeros)
>modelgen
>a=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=0, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(a)
>b=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=1, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(b)
>c=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=2, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(c)
>d=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=3, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(d)

```

```

>e=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=4, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(e)
>f=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=5, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(f)
>g=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=6, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(g)
>h=rcim(polong_isi_over_zeros, negbinomial, Svd.arg=TRUE,
Rank=7, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(h)

```

### ### Analisis Devian model RCIM###

```

>devh<-623.3496-606.3586
>devg<-623.3496-617.2412
>dbham<-ncol(polong_isi_over_zeros)-1
>dbgen<-nrow(polong_isi_over_zeros)-1
>t<-min(dbham,dbgen)-1
>dev<-c(1:(t+2))
>dev[1]<-596.7826
>dev[2]<-374.5758
>dev[3]<-237.027
>dev[4]<-90.8622
>dev[5]<-39.248
>dev[6]<-11.7416
>dev[7]<-12.713
>dev[8]<-0
>tabel<-array(0,c((4+t),5,t))
>gammi<-paste(c("RCIM"),1:t)
>gam<-c(1:t)
>for (l in 1:t){
 gam[l]<-paste(c("RCIM"),t-l+1)
}
>dimnames(tabel)<-
 list(c("Hama","Genotipe",gammi,"Residual","Total"),
 c("db","Devians","Rataan Devians","Ras. Rat.
Devians","Nilai-p"),gam)
>for(k in 1:t){
 tabel[1:2,1,k]<-c(dbham,dbgen)
 tabel[1:2,2,k]<-c(devh,devg)
 tabel[1:2,3,k]<-c(devh/dbham,devg/dbgen)
 t1<-(4+t)-1
 t2<-(4+t)
 n<-(t-k)+1
 for (l in 1:n){
 tabel[2+l,1,k]<-dbham+1+dbgen+1-1-2*1

```



```

 tabel[2+1,2,k]<-dev[1]-dev[1+1]
 tabel[2+1,3,k]<-tabel[2+1,2,k]/tabel[2+1,1,k]
 }
 tabel[t1,1,k]<-dbham*dbgen-sum(tabel[3:(2+n),1,k])
 tabel[t1,2,k]<-dev[n+1]
 tabel[t2,1:2,k]<-
c(sum(tabel[1:t1,1,k]),sum(tabel[1:t1,2,k]))
 tabel[t1:t2,3,k]<-

c(tabel[t1,2,k]/tabel[t1,1,k],sum(tabel[1:t1,2,k])/sum(tabel
[1:t1,1,k]))
 for (p in 1:(n+2)){
 tabel[p,4,k]<-tabel[p,3,k]/tabel[t1,3,k]
 tabel[p,5,k]<-
 pf(tabel[p,4,k],tabel[p,1,k],tabel[t1,1,k],lower.tail
 =FALSE)
 }
}
>Table

```

### ###Script Mean Square Error model RCIM###

```

>resia <- residuals(a)
>resiaa <- resia[,c(1,3,5,7,9,11,13,15)]
>msea <- mean(resiaa^2)
>resib <- residuals(b)
>resiba <- resib[,c(1,3,5,7,9,11,13,15)]
>mseb <- mean(resiba^2)
>resic <- residuals(c)
>resica <- resic[,c(1,3,5,7,9,11,13,15)]
>msec <- mean(resica^2)
>resid <- residuals(d)
>resida <- resid[,c(1,3,5,7,9,11,13,15)]
>mсед <- mean(resida^2)
>resie <- residuals(e)
>resiea <- resie[,c(1,3,5,7,9,11,13,15)]
>msee <- mean(resiea^2)
>resif <- residuals(f)
>resifa <- resif[,c(1,3,5,7,9,11,13,15)]
>msef <- mean(resifa^2)
>resig <- residuals(g)
>resiga <- resig[,c(1,3,5,7,9,11,13,15)]
>mseg <- mean(resiga^2)
>resih <- residuals(h)
>resiha <- resih[,c(1,3,5,7,9,11,13,15)]
>mseh <- mean(resiha^2)

```

```
>msea;mseb;msec;msed;msee;msef;mseg;mesh
```

### E. Script Data Jumlah Polong Isi Kedelai dengan *Excess Zeros* (ZINB)

#### ###Memodelkan kedalam model RCIM

```
>library(VGAM)
>modelfull<-
vglm(JUMLAH_POLONG_ISI~1,zinegbinomial,data=v_polong_isi_over_
zeros)
>modelfull
>modelgen<-
vglm(JUMLAH_POLONG_ISI~GENOTYPE,zinegbinomial,data=v_polong_isi_
i_over_zeros)
>modelgen
>modellokasi<-
vglm(JUMLAH_POLONG_ISI~LOKASI,zinegbinomial,data=v_polong_isi_
over_zeros)
>modellokasi
>a=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=0, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(a)
>b=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=1, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(b)
>c=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=2, trace=TRUE, wzepsilon = 0.225, maxiter=1000)
>logLik(c)
>d=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=3, trace=TRUE, wzepsilon = 0.225, maxiter=1000)
>logLik(d)
>e=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=4, trace=TRUE, wzepsilon = 0.225, maxiter=1000)
>logLik(e)
>f=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=5, trace=TRUE, wzepsilon = 0.225, maxiter=1000)
>logLik(f)
>g=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=6, trace=TRUE, wzepsilon = 0.225, maxiter=1000)
>logLik(g)
>h=rcim(polong_isi_over_zeros, zinegbinomial, Svd.arg=TRUE,
Rank=7, trace=TRUE, wzepsilon = 0.225, maxiter = 1000)
>logLik(h)
```

#### ###Analisis Devian model RCIM###

```
>devh<-519.2696-477.9142
>devg<-519.2696-339.4056
```

```

>dbham<-ncol(polong_isi_over_zeros)-1
>dbgen<-nrow(polong_isi_over_zeros)-1
>t<-min(dbham,dbgen)-1
>dev<-c(1:(t+2))
>dev[1]<-484.792
>dev[2]<-255.671
>dev[3]<-136.1754
>dev[4]<-80.9712
>dev[5]<-36.8504
>dev[6]<-11.068
>dev[7]<-4.7834
>dev[8]<-0
>tabel<-array(0,c((4+t),5,t))
>gammi<-paste(c("RCIM"),1:t)
>gam<-c(1:t)
>for (l in 1:t){
 gam[l]<-paste(c("RCIM"),t-l+1)
}
>dimnames(tabel)<-
 list(c("Lokasi","Genotipe",gammi,"Residual","Total"),
 c("db","Devians","Rataan Devians","Ras. Rat.
Devians","Nilai-p"),gam)
>for(k in 1:t){
 tabel[1:2,1,k]<-c(dbham,dbgen)
 tabel[1:2,2,k]<-c(devh,devg)
 tabel[1:2,3,k]<-c(devh/dbham,devg/dbgen)
 t1<-(4+t)-1
 t2<-(4+t)
 n<-(t-k)+1
 for (l in 1:n){
 tabel[2+l,1,k]<-dbham+1+dbgen+1-1-2*l
 tabel[2+l,2,k]<-dev[1]-dev[1+l]
 tabel[2+l,3,k]<-tabel[2+l,2,k]/tabel[2+l,1,k]
 }
 tabel[t1,1,k]<-dbham*dbgen-sum(tabel[3:(2+n),1,k])
 tabel[t1,2,k]<-dev[n+1]
 tabel[t2,1:2,k]<-
c(sum(tabel[1:t1,1,k]),sum(tabel[1:t1,2,k]))
 tabel[t1:t2,3,k]<-
c(tabel[t1,2,k]/tabel[t1,1,k],sum(tabel[1:t1,2,k])/sum(tabel
[1:t1,1,k]))
 for (p in 1:(n+2)){
 tabel[p,4,k]<-tabel[p,3,k]/tabel[t1,3,k]
 tabel[p,5,k]<-
 pf(tabel[p,4,k],tabel[p,1,k],tabel[t1,1,k],lower.tail

```

```

 =FALSE)
 }
}
>tabel

```

### ###Script Mean Square Error model RCIM###

```

>resia <- residuals(a)
>resiaa <- resia[,c(2,5,8,11,14,17,20,23)]
>msea <- mean(resiaa^2)
>resib <- residuals(b)
>resiba <- resib[,c(2,5,8,11,14,17,20,23)]
>mseb <- mean(resiba^2)
>resic <- residuals(c)
>resica <- resic[,c(2,5,8,11,14,17,20,23)]
>msec <- mean(resica^2)
>resid <- residuals(d)
>resida <- resid[,c(2,5,8,11,14,17,20,23)]
>msted <- mean(resida^2)
>resie <- residuals(e)
>resiea <- resie[,c(2,5,8,11,14,17,20,23)]
>msee <- mean(resiea^2)
>resif <- residuals(f)
>resifa <- resif[,c(2,5,8,11,14,17,20,23)]
>msef <- mean(resifa^2)
>resig <- residuals(g)
>resiga <- resig[,c(2,5,8,11,14,17,20,23)]
>mseg <- mean(resiga^2)
>resih <- residuals(h)
>resiha <- resih[,c(2,5,8,11,14,17,20,23)]
>mseh <- mean(resiha^2)
>msea;mseb;msec;msted;msee;msef;mseg;mesh

```

### F. Script Biplot Data Jumlah Polong Isi Kedelai dengan *Excess Zeros* (ZINB)

```

>resia<-residuals(a)
>resiaa<-resia[,c(3,6,9,12,15,18,21,24)]
>svd.a<-svd(resiaa)
>svd.a
>aa=svd.a$V[,c(1,2)]## Col
>aa
>bb=svd.a$U[,c(1,2)] ## Row
>bb
>cc=rbind(aa,bb)
>cc
>row.names(aa)=row.names(t(data_polong_isi_over_zeros))

```

```
>row.names(aa)
>row.names(bb)=row.names(data_polong_isi_over_zeros)
>row.names(bb)
>plot(1, type = "n", xlim = range(cc[,1]),
 ylim = range(cc[,2]), xlab = "KUI 1", ylab = "KUI 2",
 main = "Biplot ZINB with RCIM")
abline(h = 0)
abline(v = 0)
points(aa[, 1], aa[, 2], type = "p", col = "red")
text(aa[, 1], aa[, 2], labels = row.names(aa),
 adj = c(0.5, 0.5), col = "red")
for (i in 1:15) {
 x2 <- bb[, 1]
 y2 <- bb[, 2]
 x21 <- c(0, x2[i])
 y21 <- c(0, y2[i])
 points(x21, y21, type = "l", col = "blue")
 text(bb[, 1], bb[, 2], labels = row.names(bb))
}
```