TUGAS AKHIR – KM184801

PERAMALAN JUMLAH KECELAKAAN DI JALAN TOL DENGAN MODEL GENERALIZED AUTOREGRESSIVE MOVING AVERAGE (GARMA)

THERESIA ANINDYA PUSPANINGRUM

NRP 06111840000035

Dosen Pembimbing

Dra. Laksmi Prita Wardhani, M.Si

NIP 19611208 198803 2 001

Program Sarjana

Departemen Matematika

Fakultas Sains dan Analitika Data

Institut Teknologi Sepuluh Nopember

Surabaya

2022

TUGAS AKHIR – KM184801

PERAMALAN JUMLAH KECELAKAAN DI JALAN TOL DENGAN MODEL GENERALIZED AUTOREGRESSIVE MOVING AVERAGE (GARMA)

THERESIA ANINDYA PUSPANINGRUM

NRP 06111840000035

Dosen Pembimbing

Dra. Laksmi Prita Wardhani

NIP 19611208 198803 2 001

Program Sarjana

Departemen Matematika

Fakultas Sains dan Analitika Data

Institut Teknologi Sepuluh Nopember

Surabaya

2022

FINAL PROJECT – KM184801

FORECASTING THE NUMBER OF ACCIDENTS IN TOLL ROADS USING THE GENERALIZED AUTOREGRESSIVE MOVING AVERAGE (GARMA) MODEL

THERESIA ANINDYA PUSPANINGRUM

NRP 06111840000035

Advisor

Dra. Laksmi Prita Wardhani

NIP 19611208 198803 2 001

Bachelor Program

Department of Mathematics

Faculty of Science and Data Analytics

Sepuluh Nopember Institute of Technology

Surabaya

2022

LEMBAR PENGESAHAN

PERAMALAN JUMLAH KECELAKAAN DI JALAN TOL DENGAN MODEL *GENERALIZED AUTOREGRESSIVE MOVING AVERAGE* (GARMA)

TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Matematika pada
Program Studi S-1 Matematika
Departemen Matematika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh : **THERESIA ANINDYA PUSPANINGRUM**
NRP. 06111840000035

Surabaya, Juli 2022
Disetujui oleh Tim Penguji Tugas Akhir:

Pembimbing

1. Dra. Laksmi Prita Wardhani, M.Si (..................)
   NIP. 19611208 198803 2 001

Penguji

2. Endah Rokhmati M.P., Ph.D (..................)
   NIP. 19761213 200212 2 001

3. Sunarsini S.Si, M.Si (..................)
   NIP. 19691004 199402 2 001

4. Amirul Hakam S.Si, M.Si (..................)
   NIP. 1993202011012

Mengetahui
Kepala Departemen Matematika
Fakultas Sains dan Analitika Data

Subchan, Ph.D
NIP. 19710513 199702 1 001

PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

Nama mahasiswa / NRP	:	Theresia Anindya Puspaningrum / 0611180000035
Program studi	:	Matematika
Dosen Pembimbing / NIP	:	Dra. Laksmi Prita Wardhani M.Si / 19611208 198803 2 001

dengan ini menyatakan bahwa Tugas Akhir dengan judul "Peramalan Jumlah Kecelakaan Di Jalan Tol Dengan Model Generalized Autoregressive Moving Average (GARMA)" adalah hasil karya sendiri, bersifat orisinal, dan ditulis dengan mengikuti kaidah penulisan ilmiah.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku di Institut Teknologi Sepuluh Nopember.

Surabaya, 27 Juli 2022

Mengetahui
Dosen Pembimbing

Dra. Laksmi Prita Wardhani M.Si
NIP. 19611208 198803 2 001

Mahasiswa

Theresia Anindya Puspaningrum
NRP. 06111840000035

ABSTRAK

PERAMALAN JUMLAH KECELAKAAN DI JALAN TOL DENGAN MODEL GENERALIZED AUTOREGRESSIVE MOVING AVERAGE (GARMA)

Nama Mahasiswa / NRP	**:**	**Theresia Anindya Puspaningrum / 06111840000035**
Departemen	**:**	**Matematika FSAD - ITS**
Dosen Pembimbing	**:**	**Dra. Laksmi Prita Wardhani M.Si**

Abstrak

Model peramalan yang sering digunakan adalah model *time series* regresi dan ARIMA. Model ini mempertimbangkan asumsi data yang stasioner dan menyebar normal, namun pada tipe data *time series* yang merupakan data *count*, seringkali ditemukan bahwa data tidak menyebar normal dan stasioner sehingga model klasik Gaussian tidak selalu akurat dalam perhitungannya**.** *Generalized Linear Model* (GLM) merupakan solusi untuk menganalisis data *count*. Setelah dilakukan penelitian secara terus – menerus, dikembangkanlah sebuah model peramalan yaitu *Generalized Autoregressive Moving Average* (GARMA) untuk data yang mengikuti distribusi non – Gaussian seperti distribusi Poisson dan Binomial Negatif. Untuk mengestimasi parameter model Poisson GARMA dan Binomial Negatif GARMA digunakan metode *Maximum Likelihood Estimation* (MLE) dengan pendekatan optimasi *Iteratively Reweighted Least Square* (IRLS). Data yang digunakan dalam penelitian ini adalah data jumlah kecelakaan yang terjadi di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong yang terdistribusi Poisson dan Binomial Negatif. Pada penelitian model ARIMA digunakan dalam mencari orde p dan q, sehingga diperoleh model yang memenuhi untuk dibentuk ke dalam model GARMA yaitu Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1). Kriteria pemilihan model terbaik menggunakan RMSE dan MAPE, dengan hasil RMSE model Poisson GARMA (1,1) sebesar 0.18546485 serta MAPE sebesar 11.57 dan RMSE model Binomial Negatif GARMA (1,1) sebesar 0.150043187 dan MAPE sebesar 10.08 , sehingga dapat dikatakan bahwa model terbaik dalam meramalkan jumlah kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong adalah model Binomial Negatif GARMA (1,1).

Kata kunci: Data *count*, Binomial Negatif GARMA, Poisson GARMA, IRLS.

ABSTRACT

FORECASTING THE NUMBER OF ACCIDENTS IN TOLL ROAD USING GENERALIZED AUTOREGRESSIVE MOVING AVERAGE (GARMA) MODEL

Student Name / NRP	**:**	**Theresia Anindya Puspaningrum / 06111840000035**
Department	**:**	**Matematika FSAD – ITS**
Advisor	**:**	**Dra. Laksmi Prita Wardhani M.Si**

Abstract

Forecasting models that often used are time series regression models and ARIMA models which consider each other assumption of stationary and normally distributed data, but in the time series data type count data, it is often found that data is not normally distributed and stationary, so the classical Gaussian model is not always accurate in calculations. Generalized Linear Model (GLM) is a solution for analyzing count data. After continuous research, a forecasting model is developed, namely the *Generalized Autoregressive Moving Average* (GARMA) for data that follows non-Gaussian distributions such as the Poisson distribution and Negative Binomial. To estimate the parameters of the Poisson GARMA model and the Negative Binomial GARMA model, the Maximum Likelihood Estimation (MLE) method is used with the Iteratively Reweighted Least Square (IRLS) optimization approach. By using the ARIMA model to find the order of p and q, we get models that meet the requirements to be formed into the GARMA model, namely Poisson GARMA (1,1) and Negative Binomial GARMA (1,1). The criteria for selecting the best model using RMSE and MAPE, with the results of the Poisson GARMA (1,1) model RMSE of 0.18546485 with MAPE 11.57 and Negative Binomial (1,1) model RMSE (1,1) of 0.150043187 with MAPE 10.08, so it can be said that the best model in predicting the number of accidents on the Surabaya – Gempol Toll Road Banyu Urip – Porong section is the Negative Binomial GARMA (1,1) model.

Keywords: Data *count*, Binomial Negative GARMA, Poisson GARMA, IRLS.

KATA PENGANTAR

Segala puji syukur kepada Tuhan Yesus Kristus atas segala berkat dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul:

“PERAMALAN JUMLAH KECELAKAAN DI JALAN TOL DENGAN MODEL *GENERALIZED AUTOREGRESSIVE MOVING AVERAGE* (GARMA)”

sebagai salah satu persyaratan akademis dalam menyelesaikan Program Sarjana Departemen Matematika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Tugas Akhir ini dapat diselesaikan dengan baik berkat kerja sama, bantuan, dan dukungan dari banyak pihak. Sehubungan dengan hal itu, penulis ingin mengucapkan terima kasih kepada :

1. Mama yang telah memberikan doa, dukungan dan nasehat sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Bapak Subchan, Ph.D selaku Kepala Departemen Matematika FSAD ITS
3. Ibu Dra. Laksmi Prita Wardhani M.Si selaku dosen pembimbing atas segala masukkan dan motivasi dalam mengerjakan Tugas Akhir ini
4. Ibu Endah Rokhmati M.P., Ph.D, Ibu Sunarsini S.Si, M.Si, dan Bapak Amirul Hakam S.Si, M.Si selaku dosen penguji atas segala kritik dan masukannya sehingga tugas akhir ini dapat terselesaikan dengan baik
5. Bapak Dr. Mahmud Yunus M.Si selaku dosen wali yang memberikan arahan akademik selama penulis menempuh pendidikan di Departemen Matematika FSAD ITS.
6. Bapak dan Ibu dosen serta para staf Departemen Matematika FSAD ITS yang tidak dapat penulis sebutkan satu-persatu, yang telah memberikan ilmu dan pengalaman yang bermanfaat kepada penulis, serta segenap Karyawan, Tendik dan keluarga besar Departemen Matematika Institut Teknologi Sepuluh Nopember atas dukungan dan bantuannya.
7. Teman-teman seperjuangan, Matematika Institut Teknologi Sepuluh Nopember Angkatan 2018 yaitu angkatan MODuLO yang telah mengisi hari - hari penulis dengan penuh keceriaan, motivasi dan pengalaman.

Penulis juga menyadari bahwa dalam tugas akhir ini tentu masih terdapat beberapa kekurangan. Oleh sebab itu, kritik dan saran yang bersifat membangun sangat penulis harapkan demi kesempurnaan tugas akhir ini. Akhirnya, penulis berharap semoga tugas akhir ini dapat bermanfaat dalam pengembangan ilmu bagi banyak pihak.

Surabaya, Juli 2022

Penulis

DAFTAR ISI

Hal

Hal

Hal

DAFTAR GAMBAR

DAFTAR TABEL

Hal

DAFTAR SIMBOL

$\|x\|$	: Harga mutlak dari x
max	: Maksimum
$\exp(x)$	: Eksponen dari x
n	: Banyaknya data
cov	: Kovarian
Var	: Varian
Y	: Variabel respon yang menyatakan jumlah kecelakaan di jalan tol Gempol – Surabaya
$\boldsymbol{X^T}$	: matriks $n \times c$ dengan $X^T = (x_t, x_{t-1}, \dots, x_{t-k})$
$g(\mu_t)$	: Fungsi *link* Model GARMA
μ_t	: Hasil Peramalan ke$-t$
τ_t	: Komponen *Autoregrresive* dan *Moving Average*
ϕ	: Parameter *Autoregressive*
θ	: Parameter *Moving Average*
$L(\theta)$	: Fungsi *Likelihood*
$PL(\theta)$	: Fungsi *partial likelihood*
$l(\theta)$	: Fugsi *log-likelihood*
β_0	: Parameter konstanta model GARMA
c	: Parameter *threshold* yang bernilai $0 < c < 1$

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Peramalan merupakan suatu seni dan ilmu pengetahuan dalam memprediksi sesuatu yang akan terjadi di masa yang akan datang. Peramalan melibatkan pengambilan data historis dan memproyeksikan data tersebut ke masa yang akan datang dengan bantuan sebuah model matematika. Pada suatu peramalan atau prediksi dalam segala bidang, model *time series* merupakan model yang sering digunakan. Metode *time series* merupakan metode peramalan yang menggunakan analisa hubungan antara variabel yang diperkirakan dengan variabel waktu. Adapun model Analisis Regresi dan model *Autoregressive Integrated Moving Average* (ARIMA) adalah dua model yang paling sering digunakan dalam melakukan sebuah peramalan pada data *time series* (Wei 2016).

Ketika membuat peramalan, model ARIMA mengabaikan adanya variabel bebas dan menggunakan data saat ini serta masa lalu dari variabel terikat (dependen). Sama halnya dengan model Regresi, dalam pengerjaan model ARIMA seringkali ditemukan data yang masih bersifat non – stasioner, sedangkan komponen AR (*Autoregressive)* dan MA (*Moving Average)* pada model ARIMA hanya dapat menerima data yang stasioner. Penerapan model ARIMA juga mempertimbangkan asumsi data yang terdistribusi secara normal, sedangkan dalam beberapa kasus terdapat data yang tidak berdistribusi secara normal (Chen, Yuan, and Shu 2008). Hal ini umumnya dapat diselesaikan dengan cara melakukan transformasi data. Setelah melakukan transformasi data masih belum menjamin bahwa data secara otomatis selalu menyebar normal, namun data yang dihasilkan juga dapat memungkinkan untuk menyebar secara tidak normal.

Pada tipe data *time series* yang merupakan data jumlahan (*count*), seringkali ditemukan pada sebuah kasus bahwa data tidak menyebar secara normal, sehingga dilakukan pengembangan secara berkala untuk dapat menerapkan sebuah model pada data yang bertipe data *count*. Data *count* sendiri merupakan data diskrit yang tidak negatif atau nilainya merupakan bilangan bulat tak negatif (Cameron and Trivedi 1998). Beberapa contoh dari data *count* seperti jumlah kejadian kecelakaan yang terjadi dalam selang waktu tertentu, jumlah telepon yang masuk ke dalam perusahaan dalam interval waktu tertentu, jumlah suatu pohon pada suatu area tertentu, jumlah kriminalitas dalam suatu daerah tertentu dan lain sebagainya. Setelah pengembangan dilakukan dari waktu ke waktu, *Generalized Linear Models* (GLM) merupakan solusi untuk menganalisis data *count* dan tipe data diskrit. *Generalized Linier mode*l (GLM) yang dikembangkan oleh McCullagh dan Nelder merupakan perluasan dari model regresi linier dengan asumsi prediktor memiliki efek linier dan dapat digunakan ketika variabel respon masih termasuk dalam anggota keluarga eksponensial . Hubungan antara variabel respon dan prediktor dinyatakan dengan suatu fungsi yang disebut sebagai fungsi link (*link function*) (Cheek, McCullagh, and Nelder 1990).

Model yang umum digunakan untuk data bertipe jumlahan atau data *count* adalah model Regresi Poisson. Model Regresi Poisson dapat digunakan dengan asumsi peluang variabel responnya berdistribusi Poisson terhadap variabel prediktornya. Pada distribusi Poisson juga terdapat asumsi bahwa data diharuskan bersifat *equisdispersion* dimana nilai rata – rata sama dengan nilai varian variabel respon. Namun pada kenyataannya terdapat dua kemungkinan yang mungkin terjadi yaitu bahwa data bersifat *overdispersion* (nilai rata – rata lebih kecil dibandingkan nilai varian) atau bersifat *underdispersion* (nilai rata – rata lebih besar dibandingkan nilai varian), atau juga dapat memungkinkan bahwa di dalam data memiliki banyak nilai nol (Kedem, B & Fokianos 2003).

Pada permasalahan diatas, Benjamin, dkk mengembangkan sebuah model stokastik *Generalized Autoregressive Moving Average* (GARMA) untuk data – data yang mengikuti distribusi non – Gaussian seperti distribusi Poisson dan Binomial Negatif. Model GARMA merupakan gabungan dari komponen *Autoregressive Moving Average* (ARMA) dengan variabel prediktor ke transformasi parameter rata – rata dari distribusi data dengan menggunakan fungsi link (*link function*). Fungsi link (*link function*) digunakan agar distribusi data tetap dalam domain bilangan riil positif sehingga prediksi yang dihasilkan lebih akurat (Benjamin et al. 2003). Model GARMA merupakan perluasan dari *Generalized Linear Models* (GLM) yang dikembangkan oleh McCullagh dan Nelder. Untuk mengestimasi parameter – parameter model GARMA digunakan metode *Maximum Likelihood Estimation* (MLE) serta pendekatan algoritma *Iteratively Reweighted Least Square* (IRLS) untuk optimasi model (Green 1984). Pada penelitian ini dilakukan analisis mengenai model GARMA dengan pendekatan distribusi Poisson atau disebut model Poisson GARMA (p, q) dan pendekatan distribusi Binomial Negatif atau disebut model Binomial Negatif GARMA (p, q).

Sebelumnya telah terdapat penelitian mengenai peramalan dengan menggunakan model Poisson GARMA dan model Binomial Negatif GARMA yang dilakukan oleh Wardhani, dkk (2020). Pada penelitian tersebut dilakukan perbandingan kedua model yang dilihat dari kriteria *Akaike's Information Criterion* (AIC) dengan data yang digunakan yaitu data kriminalitas pencurian di wilayah Polres Surabaya. Penelitian tentang perluasan model GARMA juga pernah dilakukan yang melibatkan efek stasioner dan musiman atau disebut *Generalized Seasonal Autoregressive Integrated Moving Average* (GSARIMA). Pada penelitian tersebut model GSARIMA diterapkan pada sebuah data simulasi dan dibandingkan dengan model *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Selanjutnya dipilih model terbaik dalam peramalan dengan membandingkan tingkat akurasi kedua model berdasarkan *Mean Absolute Percentage Error* (MAPE) (Asrirawan , Heri Kuswanto 2014).

Salah satu contoh dari data *count* adalah jumlah kejadian kecelakaan yang terjadi di jalan tol. Menurut Peraturan Pemerintah Nomor 15 Tahun 2005, jalan tol (*tax on location*) merupakan bagian dari sistem jaringan jalan alternatif dan sebagai jalan nasional yang penggunanya diwajibkan membayar. Pembangunan jalan tol merupakan salah satu bentuk usaha pemerintah dalam memudahkan masyarakat Indonesia karena dapat mengurangi inefisiensi akibat kemacetan pada ruas utama jalan raya, serta untuk meningkatkan pelayanan sistem distribusi barang dan jasa di wilayah yang dapat menjadi sentra perekonomian (Fahza and Widyastuti 2019). Pada kenyataannya tingkat kejadian kecelakaan di jalan tol relatif cukup tinggi. Sebagai pengelola jalan tol di Indonesia, PT. Jasa Marga mengungkapkan bahwa 82% penyebab utama kecelakaan di jalan tol terjadi akibat faktor pengemudi yang lalai dalam berkendara, 17% disebabkan oleh faktor kendaraan dan 1% disebabkan oleh faktor jalan dan lingkungan sekitar. Jumlah kasus kecelakaan yang terjadi di jalan tol ini cukup tinggi, sehingga penelitian mengenai faktor penyebab dan peramalan kejadian kecelakaan yang terjadi beberapa bulan ke depan perlu dilakukan. Hal ini dilakukan agar pihak yang terkait yaitu PT. Jasa Marga dan pengendara jalan tol dapat mengantisipasi dan melakukan upaya dalam mencegah faktor – faktor terjadinya kecelakaan, seperti faktor pengemudi, faktor lingkungan dan cuaca, faktor kendaraan dan lain sebagainya.

Pada penelitian ini, data yang digunakan yaitu data jumlah kecelakaan di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong pada periode Januari 2016 – Desember 2021 yang memiliki tingkat kecelakaan paling tinggi dibandingkan ruas lainnya. Jumlah kecelakaan yang terjadi dalam kurun waktu lima tahun terakhir sebanyak 237 kasus dengan persentase kejadian 94,04% dari semua ruas. Data jumlah kecelakaan ini juga merupakan data *count* yang berdistribusi non – Gaussian yaitu berdistribusi Poisson dan Binomial Negatif. Untuk membandingkan tingkat akurasi kedua model yang dianalisis (Poisson GARMA dan Binomial Negatif GARMA) dan untuk menemukan model terbaik dalam melakukan peramalan jumlah kejadian kecelakaan di jalan tol dapat dilihat berdasarkan nilai *Root Mean Square Error* (RMSE) atau merupakan nilai rata – rata dari jumla kuadrat kesalahan hasil analisis tersebut serta berdasarkan nilai *Mean Absolute Percentage Error* (MAPE).

1.2 Rumusan Masalah

Rumusan masalah dalam Tugas Akhir ini adalah :

1. Bagaimana estimasi parameter pada model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) dengan metode *Maximum Likelihood Estimation* (MLE) dan dengan algoritma IRLS?
2. Bagaimana mengaplikasian model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) pada data jumlah kecelakaan di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong?
3. Bagaimana perbandingan tingkat akurasi peramalan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) pada data jumlah kecelakaan di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong?

1.3 Batasan Masalah

Batasan Masalah dalam Tugas Akhir ini adalah :

1. Data yang digunakan adalah data sekunder yaitu data jumlah kecelakaan yang terjadi di jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong periode bulan Januari 2016 – Desember 2021.
2. Model yang akan diestimasi adalah Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) dengan distribusi yang digunakan adalah distribusi Poisson dan distribusi Binomial Negatif.
3. Software yang digunakan untuk menghitung peramalan data jumlah kecelakaan adalah Minitab, Matlab 2018 dan Microsoft Excel.
4. Pemilihan kriteria model terbaik adalah berdasarkan RMSE (*Root Mean Square Error*).

1.4 Tujuan

Tujuan Penelitian dalam Tugas Akhir ini adalah :

1. Memperoleh estimasi parameter pada model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) dengan metode *Maximum Likelihood Estimation* (MLE) dan dengan menggunakan algoritma IRLS.
2. Mengaplikasikan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) pada data jumlah kejadian kecelakaan di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong.
3. Membandingkan tingkat akurasi dan memilih model terbaik peramalan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) pada data jumlah kecelakaan di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong.

1.5 Manfaat

Manfaat Penelitian dalam Tugas Akhir ini adalah :

1. Mengembangkan wawasan keilmuan serta pengetahuan tentang model non – Gaussian Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q).
2. Mengembangkan wawasan keilmuan dan pengetahuan tentang cara menaksir parameter Poisson GARMA dan Binomial Negatif GARMA dengan menggunakan algoritma IRLS.
3. Memberi informasi dan membantu tentang tata cara peramalan jumlah kecelakaan pada intansi atau pihak yang terkait.

BAB 2

TINJAUAN PUSTAKA

2.1 PT. Jasa Marga

PT. Jasa Marga merupakan sebuah perusahaan milik Badan Usaha Milik Negara (BUMN) Indonesia yang bergerak di bidang penyedia layanan jalan tol dan bisnis terkait lainnya. Perusahaan ini mulai beroperasi secara komersial pada tahun 1978. PT. Jasa Marga didukung oleh beberapa anak perusahaan, salah satunya adalah PT. Jasa Marga Surabaya – Gempol (Lestiningsih and Agustini 2016). Memiliki panjang 43 kilometer, jalan tol ini sudah beroperasi sejak tahun 1986 hingga saat ini atau sudah bergerak selama 36 tahun. Namun dengan bencana lumpur panas Lapindo Brantas Inc, jalan ini terpotong sekitar enam kilometer yang menghubungkan Porong dan Gempol. Pemerintah memutuskan menutup ruas tersebut sebagai jalan tol, sehingga panjangnya saat ini berkurang menjadi 37 kilometer. Jalan yang dikelola oleh cabang Surabaya – Gempol ini memiliki 2×3 lajur (Waru – Dupak) dan 2×2 lajur (Waru – Gempol), tujuh *interchange*, 27 jembatan perlintasan kendaraan, dan dua jembatan penyeberangan orang. Pada ruas ini terdapat 10 gerbang tol yang terdiri dari 6 gerbang tol dengan sistem transaksi terbuka dan 4 gerbang tol dengan sistem transaksi tertutup.

Mengingat pentingnya ruas ini, Jasa Marga memutuskan untuk memindahkan ruas yang terendam lumpur dan menggesernya sekitar 3 kilometer ke arah barat. Hal ini juga merupakan bagian dari jalan tol Trans Jawa yang akan menyambung ke Pasuruan, Probolinggo hingga ke Banyuwangi. Pada tahun 2005 rata – rata harian volume lalu lintas mencapai 173.300 kendaraan, dan merosot pada tahun 2006 setelah tergenang lumpur, hingga pada tahun 2007 rata-rata volume harian lalu lintas hanya sekitar 147.200 kendaraan. Namun pada tahun 2008 menunjukkan kenaikan lagi menjadi 156.000 kendaraan per hari. Pendapatan dari tol ini masih mengalami peningkatan sekalipun volume lalu lintas menurun pada tahun 2006. Pada tahun 2006/2007 pendapatan tol masih meningkat 10,32 persen menjadi sebesar Rp 312 juta per hari, dan tetap meningkat di tahun 2008 menjadi Rp 402,325 juta per hari.

2.2 Penelitian Terdahulu

Pada penelitian tugas akhir ini dilampirkan beberapa penelitian terdahulu yang berkoresponden dengan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) untuk data *count*. Pada penelitian ini studi kasus yang diambil adalah jumlah kecelakaan di jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong. Penelitian mengenai data *count* sebelumnya telah dilakukan oleh Rini Cahyandari yang membahas mengenai pemodelan regresi Poisson studi kasus kecelakaan lalu lintas mobil penumpang di Provinsi Jawa Barat. Berdasarkan hasil analisa data, faktor manusia, faktor jalan, dan faktor alam merupakan faktor pada model regresi Poisson mengalami overdispersi (varian peubah respon lebih besar dari rata-rata) sehingga mendapatkan kesimpulan bahwa model regresi Poisson tidak cocok digunakan untuk memodelkan studi kasus tersebut (Cahyandari 2014).

Selanjutnya penelitian mengenai estimasi model GARMA menggunakan algoritma *Iteratively Reweighted Least Square* (IRLS) juga telah dilakukan oleh Benjamin dkk. Pada penelitian tersebut, Benjamin dkk mengembangkan model GARMA tanpa melibatkan efek stasioner serta musiman dan dilakukan estimasi parameter model GARMA dengan menggunakan *Maximum Likelihood Estimation* (MLE) kemudian dioptimasi dengan algoritma IRLS (Benjamin et al. 2003). Selain itu, terdapat pula penelitian yang dilakukan oleh Asriawan dengan topik perbandingan peramalan *Generalized Seasonal Autoregressive Integrated Moving*

Average (GSARIMA) dan *Seasonal Autoregressive Integrated Moving Average* (SARIMA) pada jumlah penderita Demam Berdarah Dengue (DBD) di Kota Surabaya. Data yang digunakan pada penelitian memiliki tipe data *count* yang memiliki efek musiman sehingga model yang terbentuk adalah GSARIMA $(p,d,q)(P,D,Q)^s$. Hasil yang diperoleh dari penelitian tersebut menunjukkan bahwa model GSARIMA relatif lebih baik dibandingkan model SARIMA berdasarkan kriteria nilai *Mean Absolute Percentage Error* (MAPE) yang terkecil (Asrirawan , Heri Kuswanto 2014).

Sebelumnya telah terdapat penelitian mengenai peramalan dengan menggunakan model Poisson GARMA dan Binomial Negatif GARMA yang dilakukan oleh Wardhani, dkk (2020). Data yang digunakan yaitu data kriminalitas pencurian di wilayah Polres Surabaya. Pada penelitian tersebut Wardhani, dkk memperoleh hasil bahwa pengaplikasian peramalan dengan menggunakan model Binomial Negatif GARMA terhadap data kriminalitas pencurian di Surabaya Pusat, Surabaya Timur, Surabaya Barat dan Surabaya Selatan memiliki hasil yang lebih baik dibandingkan peramalan dengan menggunakan model Poisson GARMA. Pemilihan model terbaik dilihat dari perolehan nilai terkecil berdasarkan kriteria *Akaike's Information Criterion* (AIC).

2.3 Model – model Deret Waktu

Data deret waktu atau data *time series* merupakan rangkaian nilai dari suatu variabel tertentu atau pengamatan yang berurutan dalam interval waktu tertentu secara konstan. Contohnya seperti data yang diambil per jam, per hari, per minggu, per bulan, dan sebagainya. Adanya data *time series* ini dapat digunakan sebagai dasar dalam melakukan perencanaan atau memperkirakan kegiatan di masa yang akan datang atau dengan kata lain dapat disebut dengan peramalan. Setiap pengamatan yang dilakukan berdasarkan waktu tertentu t dimana $t = 1, 2, 3, \ldots$ dapat dinyatakan dalam bentuk variabel Y_t sehingga penulisan dari data *time series* yaitu $Y_1, Y_2, Y_3, \ldots, Y_t$ (Halim 2006).

Beberapa model untuk data *time series* yang bersifat stasioner antara lain model *Autoregressive* (AR), model *Moving Average* (MA), dan model gabungan dari dua model sebelumnya yaitu *Autoregressive Moving Average* (ARMA). Sedangkan model yang digunakan untuk data time series yang bersifat nonstasioner menggunakan model *Autoregressive Integrated Moving Average* (ARIMA) dengan melakukan *differencing* dan/atau transformasi pada data tersebut terlebih dahulu. Pada bagian ini dijelaskan mengenai beberapa model – model deret waktu yang sering digunakan.

1. Model Autoregressive (AR)

   Bentuk umum model *Autoregressive* (AR) dengan orde p (AR(p)) atau model ARIMA (p, 0,0) dinyatakan sebagai berikut:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

dengan,

δ : suatu konstanta
ϕ_p : parameter autoregressive ke$-p$, $p = 1,2, \ldots, n$
e_t : nilai galat pada data ke t, $t = 0,1,2, \ldots, n$; $e_t \sim N(0, \sigma^2)$.

2. Model Moving Average (MA)

   Bentuk umum model *moving average* (MA) orde q atau MA(q) dinyatakan sebagai berikut:

$$Y_t = \delta + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

dengan,

δ : suatu konstanta
θ_q : parameter *moving average* ke$-q$, $q = 1,2, \dots, n$
e_{t-q} : nilai galat pada saat data $t - q$, $t = 0,1,2, \dots, n$; $e_{t-q} \sim N(0, \sigma^2)$.

3. Model Campuran

Terdapat dua model yang dapat dibentuk dari gabungan model *Autoregressive* (AR) dan *Moving Average* (MA), yaitu model *Autoregressive Moving Average* (ARMA) dan *Autoregressive Integrated Moving Average* (ARIMA).

a. Model ARMA

Model umum untuk campuran proses AR(p) dan MA(q) murni untuk data yang stasioner, misalnya ARMA(p, q) atau dapat ditulis dengan ARIMA $(p, 0, q)$ dinyatakan dalam bentuk persamaan sebagai berikut:

$$Y_t = \delta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \quad (2.1)$$

dengan,

δ : suatu konstanta
ϕ_p : parameter *autoregressive*
θ_q : parameter *moving average*
e_t : nilai galat pada data ke t, $t = 0,1,2, \dots, n$; $e_t \sim N(0, \sigma^2)$
e_{t-q} : nilai galat pada saat data $t - q$, $t = 0,1,2, \dots, n$; $e_{t-q} \sim N(0, \sigma^2)$.

b. Model ARIMA

Model ARIMA merupakan model gabungan atau campuran dari model *Autoregressive* AR(p) dan *Moving Average* MA(q) untuk data yang bersifat tidak stasioner dan sudah di *differencing* sehingga menjadi model yang stasioner, dengan banyaknya *differencing* dinotasikan dengan d. Proses penstasioneran dengan *differencing* untuk data tidak stasioner dilakukan terhadap rata – rata (*mean*) dan proses transformasi dilakukan untuk data yang tidak stasioner terhadap varians. Bentuk umum ARIMA (p, d, q) dapat ditulis sebagai berikut:

$$\phi_p(B)(1 - B)^d Y_t = \theta_0 + \theta_q(B) e_t \quad (2.2)$$

dengan,

$\phi_p(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$: operator *autoregressive*
$\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$: operator *moving average*
$(1 - B)^d$: *differencing* orde d
e_t : nilai galat pada saat t; $e_t \sim N(0, \sigma^2)$.

2.4 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF)

Autocorrelation Function (ACF) atau biasa disebut fungsi autokorelasi merupakan suatu hubungan linear pada data *time series* antara Y_t dengan Y_{t+k} yang dipisahkan oleh waktu lag k. ACF sering digunakan sebagai langkah awal dalam mengidentifikasi model sementara data yang diramalkan dan digunakan untuk melihat kestasioneran dalam *mean* atau rata – rata. Secara umum fungsi autokorelasi dirumuskan sebagai berikut:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{cov(Y_t, Y_{t+k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t+k})}} \tag{2.3}$$

dan kovarians antara Y_t dengan Y_{t+k} dapat dilihat pada persamaan (2.4) sebagai berikut:

$$\gamma_k = cov(Y_t, Y_{t+k}) = E(Y_t - \mu)(Y_{t+k} - \mu) \tag{2.4}$$

dengan,

ρ_k : autokorelasi pada *lag* k

γ_k : kovarians pada *lag* k

μ : rata – rata

t : waktu pengamatan; $t = 1,2,3,\ldots$

$Var(Y_t) = Var(Y_{t+k}) = \gamma_0$.

Sedangkan fungsi autokorelasi dalam pengambilan sampel data *time series* dapat dituliskan sebagai berikut:

$$\rho_k = \frac{\sum_{t=1}^{n-k}(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}, k = 0,1,2,\ldots \tag{2.5}$$

dimana,

$$\bar{Y} = \frac{\sum_{t=1}^{n} Y_t}{n}$$

dengan,

ρ_k : autokorelasi pada *lag* k

Y_t : data ke t

$\bar{Y}$: rata-rata

t : waktu pengamatan ; $t = 1,2,3,\ldots$

n : jumlah data.

Sebuah data *time series* dikatakan stasioner dalam *mean* apabila *lag* pada autokorelasinya turun secara cepat menuju nol. Stasioner digunakan untuk memperkecil kekeliruan model yang dilihat dari tidak adanya perubahan drastis pada data. Kestasioneran diidentifikasi dengan bentuk sebaran data yang berfluktuatif di sekitar nilai rata – rata yang konstan dan tidak bergantung pada waktu dan variansi dari fluktuasi tersebut. Sedangkan jika *lag* autokorelasinya turun secara lambat menuju nol maka data tersebut tidak stasioner terhadap *mean*. Sedangkan *Partial Autocorrelation Function* (PACF) atau dapat disebut sebagai Fungsi Autokorelasi Parsial digunakan sebagai alat untuk mengukur tingkat keratan antara Y_t dan Y_{t+k} apabila pengaruh lag $t+1, t+2, \dots, t+k-1$ dianggap terpisah. Untuk PACF dapat didekati dengan persamaan sebagai berikut (Wei 2016) :

$$\phi_{kk} = \frac{cov[(Y_t - \overline{Y}_t), (Y_{t+k} - \overline{Y}_{t+k})]}{\sqrt{Var(Y_t - \overline{Y}_t)}\sqrt{Var(Y_{t+k} - \overline{Y}_{t+k})}} \tag{2.6}$$

dengan,

ϕ_{kk} : autokorelasi parsial

Y_t : nilai variabel *time series* pada waktu ke t

Y_{t+k} : data yang dipisahkan oleh waktu ke-k; $k = 1,2,3, \dots$

$\overline{Y}_t$: rata – rata.

Fungsi autokorelasi parsial dalam pengambilan sampel data *time series* dapat dituliskan sebagai berikut:

$$\phi_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^{k} \phi_{kj}\rho_{k+1-j}}{1 - \sum_{j=1}^{k} \phi_{kj}\rho_j} \tag{2.7}$$

dengan,

ρ_k : autokorelasi pada lag ke k

$\phi_{k+1,k+1}$: autokorelasi parsial pada lag ke k.

Tabel 2.1 Pola ACF dan PACF

Model	**ACF**	**PACF**
AR(p)	Menurun secara eksponensial	Terpotong setelah lag ke$-p$
MA(q)	Terpotong setelah lag ke$-q$	Menurun secara eksponensial
ARMA(p, q)	Menurun secara eksponensial setelah lag ke$-(q-p)$	Menurun secara eksponensial setelah lag ke$-(p-q)$

Tabel 2.1 menunjukkan cara menentukan model AR, MA, dan ARMA. Untuk menentukan orde tertinggi q dapat dilihat dari banyaknya lag yang keluar pada *plot* ACF. Untuk menentukan orde tertinggi p dapat dilihat dari banyaknya lag yang keluar pada *plot* PACF.

2.5 *Generalized Linear Models* (GLM)

Generalized Linear Models (GLM) merupakan salah satu kelompok model statistika yang menghubungkan kombinasi linier antara variabel respon dengan variabel prediktor. Variabel respon atau yang sering disebut variabel dependen adalah variabel yang biasanya dipengaruhi oleh variabel prediktor (independen) dan mendefinisikan tentang semua hal yang terjadi saat percobaan. Variabel prediktor (independen) merupakan variabel yang mempengaruhi variabel lain dan menjadi sebab atau berubanya variabel lain atau dapat dikatakan sebagai faktor yang mempengaruhi sebuah variabel. GLM merupakan pengembangan dari model linier "klasik" untuk mengatasi masalah variabel respon yang tidak berdistribusi normal. Variabel respon dalam GLM diasumsikan berdistribusi yang termasuk ke dalam distribusi eksponensial.

Pendekatan GLM sangat baik karena memberikan struktur teori yang umum untuk kebanyakan model statistika, selain itu implementasinya sederhana untuk dituangkan dalam software statistika dan biasanya algoritma yang sama dapat digunakan untuk estimasi, inferensi, dan kecukupan model untuk semua GLM (Cheek et al. 1990)**.** Terdapat tiga komponen utama dalam GLM, yaitu:

1. Komponen acak, distribusi bersyarat dari variabel respon mengikuti distribusi keluarga eksponensial pada bentuk kanonik yaitu untuk $t = 1,2,\dots,n$.

$$f(y;\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right)$$

2. Komponen sistematik, yaitu hubungan dari sebuah vektor $(\eta_1,\eta_2,\dots,\eta_n)$ yang bertujuan untuk menjabarkan variabel – variabel yang berhubungan dengan model linier, dengan $\boldsymbol{X^T} = (x_t, x_{t-1},\dots,x_{t-k})$ dan $\boldsymbol{\beta}$ merupakan matriks dari koefisien regresi.

$$\eta_t = \sum_{j=1}^{p} \boldsymbol{\beta}_j \boldsymbol{X}_j = \boldsymbol{X^T\beta} \tag{2.8}$$

3. Fungsi link (*link function*) yang menghubungkan suatu fungsi dari nilai tengah komponen acak dengan komponen sistematik dengan persamaan sebagai berikut:

$$g(\mu_t) = \eta_t = \sum_{j=1}^{p} \boldsymbol{\beta_j X_j} = \boldsymbol{X^T\beta} \tag{2.9}$$

2.6 Distribusi Poisson

Distribusi Poisson adalah distribusi probabilitas diskrit yang menyatakan peluang jumlah peristiwa atau kejadian yang terjadi pada periode waktu tertentu apabila nilai rata – rata (*mean*) kejadian tersebut diketahui dan dalam waktu yang saling bebas (independen) sejak kejadian terakhir (Green 1984). Kejadian saling bebas atau dapat disebut kejadian independen memiliki ciri – ciri sebagai berikut:

1. Banyaknya percobaan yang terjadi dalam suatu selang waktu tidak tergantung pada banyaknya hasil percobaan yang terjadi pada selang waktu atau daerah yang terpisah.
2. Peluang terjadinya satu hasil percobaan selama suatu selang waktu yang singkat sebanding dengan panjang selang waktu tertentu atau besarnya daerah tersebut.
3. Peluang bahwa lebih dari satu hasil percobaan akan terjadi dalam selang waktu yang singkat tersebut diabaikan.

Misalkan y_i dengan $i = 0,1,2,3,\dots$ merupakan suatu jumlah kejadian yang terjadi dalam interval waktu tertentu dengan nilai parameter dari distribusi Poisson (μ_t). Variabel respon y merupakan populasi yang berdistribusi Poisson, maka fungsi kepadatan peluang untuk distribusi Poisson dapat ditulis seperti yang tertera pada persamaan (2.10) sebagai berikut:

$$f(y;\mu_t) = \frac{e^{-\mu_t}\mu_t^{y_t}}{y!}; y = 0,1,2,\dots,n \tag{2.10}$$

dengan,

t = $1,2,3,\dots,n$
$E(y)$ = μ_t
$Var(y)$ = μ_t.
Pada distribusi Poisson rata – rata (*mean*) dan variansnya bernilai sama yakni $E(y) = Var(y) = \mu_t$, untuk pembuktiannya dapat dilihat pada Lampiran 5.

2.7 Model Poisson untuk Data Count

Distribusi bersyarat dari setiap hasil observasi y_t, untuk $t = 1,2,3,\dots,n$ diberikan pada himpunan $H_t = \{x_t,\dots,x_1,y_{t-1},\dots,y_1,\mu_{t-1},\dots,\mu_1\}$, yang merupakan keluarga eksponensial. Model Poisson untuk data count diperoleh dari bentuk eksponensial dari persamaan (2.10) yaitu sebagai berikut (Kedem, B & Fokianos 2003):

$$f(y_t;\mu_t|H_{t-1}) = \exp\{(y_t \ln \mu_t - \mu_t) - (\ln(y_t))!\} \quad (2.11)$$

untuk $t = 1,2,\dots,n$

$E(Y_t|H_{t-1}) = Var(\mu_t) = \mu_t$

$b(\theta_t) = \mu_t = \exp\,(\theta_t)$

Parameter Dispersi $(\phi) = 1$.

Fungsi *likelihood* parsial untuk model Poisson adalah sebagai berikut (Kedem & Fokianos, 2005):

$$PL(\mu_t) = \prod_{t=1}^{n} f(y_t;\mu_t|H_{t-1})$$

$$PL(\mu_t) = \prod_{t=1}^{n} \frac{\exp(-\mu_t)\,\mu_t^{y_t})}{y_t!}. \quad (2.12)$$

2.8 Model Binomial Negatif untuk Data Count

Regresi binomial negatif merupakan salah satu model regresi terapan dari *Generalized Linear Models* (GLM). Sebagai penerapan dari GLM maka distribusi binomial negatif memiliki tiga komponen yaitu komponen random, komponen sistematik, dan fungsi link seperti yang sudah dijelaskan sebelumnya. Pada regresi binomial negatif variabel respon y diasumsikan berdistribusi binomial negatif yang dihasilkan dari distribusi *mixture* Poisson Gamma (Startz 2008) yang dapat dilihat pada Lampiran 5. Untuk membentuk suatu model regresi pada distribusi Binomial Negatif, maka nilai parameter dari distribusi Poisson Gamma *mixture* dinyatakan dalam bentuk $\mu_t = \alpha\beta$ sehingga diperoleh *mean* dan variansi dalam bentuk:

$$E(y_t) = \mu_t \text{ dan } Var(y_t) = \mu_t + \alpha\mu_t^2$$

kemudian fungsi kepadatan peluang untuk distribusi Binomial Negatif menjadi:

$$f(y_t;\mu_t;\alpha) = \frac{\Gamma(y_t + \frac{1}{\alpha})}{\Gamma(\mathrm{y_t}+1)\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{1}{1+\alpha\mu_t}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_t}{1+\alpha\mu_t}\right)^{y_t} \quad (2.13)$$

dengan $y_t = 0,1,2, ...$

Saat $\alpha \to 0$ maka distribusi Binomial Negatif memiliki varians $Var(y_t) = \mu_t$, distribusi binomial negatif akan mendekati suatu distribusi Poisson yang mengasumsikan *mean* dan varians sama yaitu $E(y_t) = Var(y_t) = \mu_t$. Fungsi distribusi keluarga eksponensial dari distribusi dapat dilihat pada persamaan (2.14) yang sudah tertera sebagai berikut:

$$f(y_t, \mu, \alpha) = exp\left\{y_t \ln\left(\frac{\alpha\mu_t}{1+\alpha\mu_t}\right) + \frac{1}{\alpha}\ln\left(\frac{1}{1+\alpha\mu_t}\right) + \ln\left(\frac{\Gamma(y+\frac{1}{\alpha})}{\Gamma\left(\frac{1}{\alpha}y_t!\right)}\right)\right\}. \tag{2.14}$$

Kontribusi variabel prediktor dalam model regresi Binomial Negatif dinyatakan dalam bentuk kombinasi linier antara parameter (η) dengan parameter regresi yang akan diestimasi yaitu:

$$\eta_t = \beta_0 + \beta_1 x_{t-1} + \cdots + \beta_k x_{t-k} \tag{2.15}$$

atau dalam bentuk matriks dapat dituliskan seperti pada persamaan (2.16) sebagai berikut:

$$\boldsymbol{\eta} = \boldsymbol{X^T\beta} \tag{2.16}$$

dengan $\boldsymbol{\eta}$ adalah vektor $(n \times 1)$ dari observasi, $\boldsymbol{X}$ adalah matriks $(n \times c)$ dari variabel bebas, $\boldsymbol{\beta}$ adalah matriks $(c \times 1)$ dari koefisien regresi. Nilai ekspektasi dari variabel respon Y adalah diskrit dan bernilai positif. Maka, untuk mentransformasikan nilai η_t (bilangan riil) ke rentang yang sesuai dengan rentang pada variabel respon Y diperlukan suatu fungsi link $g(.)$ seperti pada persamaan (2.17) sebagai berikut:

$$g(\mu_t) = \ln \mu_t = \boldsymbol{X^T\beta}. \tag{2.17}$$

Fungsi *likelihood* parsial untuk model Binomial Negatif adalah sebagai berikut (Kedem & Fokianos, 2005):

$$PL(\beta) = \prod_{t=1}^{n} f(y_t; \beta | H_{t-1})$$

$$PL(\beta) = \prod_{t=1}^{n} exp\left\{y_t ln\left(\frac{\alpha\mu_t}{1+\alpha\mu_t}\right) + \frac{1}{\alpha} ln\left(\frac{1}{1+\alpha\mu_t}\right) + ln\Gamma\left(y_t + \frac{1}{\alpha}\right)\right\} - ln\Gamma(y_t+1) - ln\Gamma\left(\frac{1}{\alpha}\right). \tag{2.18}$$

2.9 Model *Generalized Autoregressive Moving Average* (GARMA)

Model GARMA pertama kali diperkenalkan oleh Benjamin, dkk pada tahun 2003. Bentuk dari model GARMA (p, q) dapat dilihat pada persamaan (2.19) yang tertera sebagai berikut (Benjamin et al. 2003):

$$g(\mu_t) = \boldsymbol{Z_{t-1}^T}\beta = \boldsymbol{X_{t-1}^T\beta} + \tau_t \tag{2.19}$$

dengan,

$$\tau_t = \sum_{j=1}^{p} \phi_j A\big(y_{t-j}, X_{t-j}, \beta\big) + \sum_{j=1}^{q} \theta_j M(y_{t-j}, \mu_{t-j}) \tag{2.20}$$

τ_t : komponen AR dan MA
H : fungsi yang merepresentasikan bentuk *autoregressive*
A : fungsi yang merepresentasikan bentuk *moving average*
ϕ^T : parameter *autoregressive* ; $\phi^T = \big(\phi_1, \phi_2, \dots, \phi_p\big)$
θ^T : parameter *moving average* ; $\theta^T = \big(\theta_1, \theta_2, \dots, \theta_q\big)$.

Bentuk dari model GARMA (p, q) yang didefinisikan oleh Benjamin, dkk (2003) adalah sebagai berikut:

$$g(\mu_t) = \boldsymbol{X_{t-1}^T\beta} + \sum_{j=1}^{p} \phi_j\{g\big(y_{t-j}\big) - \boldsymbol{X_{t-j}^T\beta}\} + \sum_{j=1}^{q} \theta_j\{g\big(y_{t-j}\big) - \boldsymbol{\eta_{t-j}\beta}\}. \tag{2.21}$$

a. Diberikan data deret waktu $\{Y_t\}, t = 1,2, \dots, n$ dengan distribusi bersyarat Poisson pada persamaan (2.11) dan berdistribusi Binomial Negatif pada persamaan (2.14), maka bentuk persamaan *partial likelihood* dapat dilihat pada persamaan (2.22) sebagai berikut:

$$PL(\boldsymbol{\beta}) = \prod_{t=1}^{n} f(y_t; \theta_t, \kappa \,|H_{t-1}) \tag{2.22}$$

dimana $\boldsymbol{\beta}$ merupakan vektor koefisien regresi, sehingga jika persamaan (2.22) diturunkan ke dalam bentuk *log-partial likelihood*, maka akan diperoleh persamaan (2.23) sebagai berikut:

$$l(\boldsymbol{\beta}) = \prod_{t=1}^{n} \ln f(y_t; \theta_t, \kappa \,|H_{t-1})$$

$$= \prod_{t=1}^{n} \left\{\frac{y_t\theta_t - b(\theta_t)}{\alpha_t(\kappa)} + c(y_t, \kappa)\right\} \tag{2.23}$$

$$= \sum_{t=1}^{n} l_t$$

dan misalkan $\nabla = \left(\frac{\partial}{\partial \beta_1}, \frac{\partial}{\partial \beta_2}, \dots, \frac{\partial}{\partial \beta_i}\right)$, maka skor parsial didefinisikan sebagai $\nabla l(\boldsymbol{\beta})$ merupakan turunan pertama dari fungsi *likelihood* parsial logaritma dengan vektor regresi $\boldsymbol{\beta}$ yang tidak diketahui. Adapun peritungannya dibangun dengan aturan rantai sebagai berikut:

$$\frac{\partial l_t}{\partial \beta_j} = \frac{\partial l_t}{\partial \theta_t}\frac{\partial \theta_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_t}\frac{\partial \eta_t}{\partial \beta_j} \tag{2.24}$$

dimana $j = 1,2, \dots, i$ dan

$$\frac{\partial l_t}{\partial \theta_t} = \frac{(y_t - b'(\theta_t))}{\alpha_t(\kappa)} = \frac{(y_t - \mu_t)}{\alpha_t(\kappa)} \tag{2.25}$$

serta

$$\frac{\partial \theta_t}{\partial \mu_t} = \frac{1}{b''(\theta_t)} = \frac{\alpha_t(\kappa)}{Var\,[y_t|\mathcal{F}_{t-1}]} \tag{2.26}$$

karena $\eta_t = \sum_{j=1}^{k} z_{(t-1)j}\beta_j$ maka diperoleh

$$\frac{\partial \eta_t}{\partial \beta_j} = z_{(t-1)j} \tag{2.27}$$

ehingga persamaan (2.24) akan menjadi persamaan (2.28) sebagai berikut:

$$\frac{\partial l_t}{\partial \beta_j} = \frac{(y_t - \mu_t)}{Var[y_t|\mathcal{F}_{t-1}]}\frac{\partial \mu_t}{\partial \eta_t} z_{(t-1)j} \tag{2.28}$$

dan secara umum persamaan fungsi skor parsial untuk vektor koefisien regresi dapat ditulis sebagai berikut:

$$U(\boldsymbol{\beta}) = \nabla l(\boldsymbol{\beta}) = \sum_{t=1}^{n} Z_{t-1}\frac{\partial \mu_t}{\partial \eta_t}\frac{(y_t - \mu_t(\boldsymbol{\beta}))}{\sigma^2(\boldsymbol{\beta})} \tag{2.29}$$

dimana

$$\sigma^2(\boldsymbol{\beta}) = Var\,[y_t|\mathcal{F}_{t-1}] \tag{2.30}$$

$$E\left(\sum_{t=1}^{n} Z_{t-1}\frac{\partial \mu_t}{\partial \eta_t}\frac{(y_t - \mu_t(\boldsymbol{\beta}))}{\sigma^2(\boldsymbol{\beta})}\middle|\mathcal{F}_{t-1}\right) = 0. \tag{2.31}$$

Untuk menyelesaikan persamaan (2.29) maka digunakan algoritma *Fisher Scoring* yang merupakan modifikasi prosedur *Newton – Raphson.* Namun, untuk menggunakan algoritma tersebut maka diperlukan sebuah matriks informasi $I(\boldsymbol{\beta})$ yang rumusnya diberikan sebagai berikut:

$$\begin{aligned} I(\boldsymbol{\beta}) &= \sum_{t=1}^{n} Cov\,\left[\sum_{t=1}^{n} Z_{t-1}\frac{\partial \mu_t}{\partial \eta_t}\frac{(y_t - \mu_t(\boldsymbol{\beta}))}{\sigma^2(\boldsymbol{\beta})}\middle|\mathcal{F}_{t-1}\right] \\ &= \sum_{t=1}^{n} Z_{t-1}\left(\frac{\partial \mu_t}{\partial \eta_t}\right)^2 \frac{1}{\sigma^2(\boldsymbol{\beta})}\boldsymbol{Z}_{t-1}^T \end{aligned}$$

$$= \boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{W}(\boldsymbol{\beta})\boldsymbol{Z} \tag{2.32}$$

dengan

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_{\boldsymbol{0}}^{\boldsymbol{T}} \\ \boldsymbol{Z}_{\boldsymbol{1}}^{\boldsymbol{T}} \\ \vdots \\ \boldsymbol{Z}_{\boldsymbol{n-1}}^{\boldsymbol{T}} \end{bmatrix}$$

adalah matriks yang berukuran $N \times p$ dan $\boldsymbol{W}(\boldsymbol{\beta}) = diag(w_1, w_2, \dots, w_n)$ dan w_t didefinisikan sebagai berikut:

$$w_t = \left(\frac{\partial \mu_t}{\partial \eta_t}\right)^2 \frac{1}{\sigma^2(\boldsymbol{\beta})} \tag{2.33}$$

dengan $t = 1,2, \dots, n$.

Berdasarkan matriks informasi pada persamaan (2.32) tersebut maka diperoleh suatu iterasi sebagai berikut:

$$\widehat{\boldsymbol{\beta}}^{(m+1)} = \widehat{\boldsymbol{\beta}}^{(m)} + I^{-1}\left(\widehat{\boldsymbol{\beta}}^{(m)} U(\widehat{\boldsymbol{\beta}}^{(m)})\right) \tag{2.34}$$

Jika matriks informasi tersebut diasumsikan memiliki sebuah invers maka persamaan (2.34) dapat diubah menjadi persamaan (2.35) sebagai berikut:

$$I\left(\widehat{\boldsymbol{\beta}}^{(m)}\right)\widehat{\boldsymbol{\beta}}^{(m+1)} = I\left(\left(\widehat{\boldsymbol{\beta}}^{(m)}\right)\widehat{\boldsymbol{\beta}}^{(m)} + U\left(\widehat{\boldsymbol{\beta}}^{(m)}\right)\right) \tag{2.35}$$

Perhatikan sisi kanan dari persamaan (2.32) diatas akan diperoleh:

$$\begin{aligned} &\sum_{j=1}^{p}\left[\sum_{t=1}^{n} \frac{Z_{(t-1)j} Z_{(t-1)l}}{\sigma_t^2}\left(\frac{\partial \mu_t}{\partial \eta_t}\right)^2 \widehat{\boldsymbol{\beta}}_j^{\ (k)}\right] + \sum_{t=1}^{n} \frac{\left(y_t - Z_{(t-1)l}\right)}{\sigma_t^2}\left(\frac{\partial \mu_t}{\partial \eta_t}\right) \\ &= \sum_{t=1}^{n} Z_{(t-1)l} w_t \left\{\eta_t + (y_t - \mu_t)\frac{\partial \mu_t}{\partial \eta_t}\right\} \end{aligned} \tag{2.36}$$

untuk $l = 1,2, \dots, p$ dan $\mu_t, \eta_t, \frac{\partial \mu_t}{\partial \eta_t}$ mengevaluasi nilai $\widehat{\boldsymbol{\beta}}^{(k)}$ dan w_t , sehingga diperoleh persamaan (2.37) sebagai berikut:

$$\begin{aligned} z_t^{(m)} &= \sum_{j=1}^{p} Z_{(t-1)j} \widehat{\boldsymbol{\beta}}_j^{\ (m)} + (y_t - \mu_t)\frac{\partial \mu_t}{\partial \eta_t} \\ &= \eta_t\left(\widehat{\boldsymbol{\beta}}^{(m)}\right) + (y_t - \mu_t)\frac{\partial \mu_t}{\partial \eta_t} \end{aligned} \tag{2.37}$$

sehingga persamaan (2.32) menjadi:

$$\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{W}\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)\boldsymbol{Z}\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m+1})} = \boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{W}\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)\boldsymbol{z}^{(\boldsymbol{m})} \tag{2.38}$$

Sehingga diperoleh persamaan *fisher scoring* seperti yang tertera pada persamaan (2.39):

$$\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m}+\mathbf{1})} = \left(\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{W}\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)\boldsymbol{Z}\right)^{-\mathbf{1}}\boldsymbol{W}\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)z^{(m)} \tag{2.39}$$

dimana $\boldsymbol{W}\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)$ dan $z^{(m)}$ dievaluasi pada $\left(\widehat{\boldsymbol{\beta}}^{(\boldsymbol{m})}\right)$**.** Persamaan (2.39) merupakan prosedur algoritma IRLS yang didefiniskan oleh Kedem dan Fokianos (2002).

2.10 Model Poisson GARMA $(\boldsymbol{p}, \boldsymbol{q})$

Pada penelitian ini salah satu model yang digunakan adalah Model Poisson GARMA. Model Poisson GARMA (p, q) merupakan pengembangan dari model *Generalized Autoregressive Moving Average* (GARMA) (p, q) oleh Benjamin, dkk dengan pendekatan variabel respon yang berdistribusi Poisson. Pada sub bab sebelumnya telah dijelaskan bahwa dalam membuat suatu model GARMA diperlukan suatu fungsi link (*link function*) yang bertujuan untuk menyatakan hubungan antara variabel respon dan variabel prediktor. Fungsi link kanonik untuk model Poisson GARMA (p, q) merupakan bentuk fungsi logaritma.

Persamaan (2.21) dapat berbentuk seperti yang tertera pada persamaan (2.40) sebagai berikut (Wardhani et al. 2020):

$$\ln(\mu_t) = \boldsymbol{X}_{\boldsymbol{t-1}}^{\boldsymbol{T}}\boldsymbol{\beta} + \sum_{j=1}^{p} \phi_j\{lny_{t-j}^* - \boldsymbol{X}_{\boldsymbol{t-j}}^{\boldsymbol{T}}\boldsymbol{\beta}\} + \sum_{j=1}^{q} \theta_j \left\{\frac{lny_{t-j}^*}{\mu_{t-j}}\right\}. \tag{2.40}$$

dimana $y_{t-j}^* = \max\,(y_{t-j}, c)$ dan $0 < c < 1$. Jika terdapat nilai 0 pada nilai y_{t-j} maka akan diganti dengan c yang merupakan koefisien *threshold*.

2.11 Model Binomial Negatif GARMA $(\boldsymbol{p}, \boldsymbol{q})$

Selain model Poisson GARMA, dalam penelitian ini juga menggunakan model Binomial Negatif GARMA. Model Binomial Negatif GARMA merupakan perluasan dari model *Generalized Autoregressive Moving Average* (GARMA) oleh Benjamin, dkk. Model Binomial Negatif GARMA (p, q) mengasumsikan variabel responnya berdistribusi Binomial Negatif. Pada sub bab sebelumnya telah dijelaskan bahwa dalam membuat suatu model GARMA diperlukan suatu fungsi link (*link function*) yang bertujuan untuk menyatakan hubungan antara variabel respon dan variabel prediktor. Fungsi link kanonik untuk model Binomial Negatif GARMA (p, q) merupakan bentuk fungsi logaritma.

Persamaan (2.21) dapat berbentuk seperti yang tertera pada persamaan (2.41) sebagai berikut (Wardhani et al. 2020):

$$\ln(\mu_t) = \boldsymbol{X}_{\boldsymbol{t-1}}^{\boldsymbol{T}}\beta + \sum_{j=1}^{p} \phi_j\{lny_{t-j}^* - \boldsymbol{X}_{\boldsymbol{t-j}}^{\boldsymbol{T}}\beta\} + \sum_{j=1}^{q} \theta_j \left\{\frac{lny_{t-j}^*}{\mu_{i-j}}\right\} \tag{2.41}$$

dimana $y_{t-j}^* = \max(y_{t-j}, c)$ dan $0 < c < 1$. Jika terdapat nilai y_{t-j} sama dengan nol maka akan diganti dengan c yang merupakan parameter *threshold*.

2.12 *Maximum Likelihood Estimation* (MLE)

Metode *maximum likelihood estimation* (MLE) adalah metode pendugaan atau estimasi dengan memaksimumkan fungsi *likelihood*. Dalam penelitian ini metode MLE digunakan untuk menduga parameter distribusi Poisson dan distribusi Binomial Negatif yang digunakan dalam model GARMA.

Apabila diketahui populasi $X \sim f(x_t; \theta)$, maka langkah – langkah MLE sebagai berikut:

1. Ambil n sampel random $x_1, x_2, \dots, x_n$ yang berdistribusi sama dengan X.
2. Buat fungsi *likelihood* yaitu fungsi distribusi peluang bersama dari $x_1, x_2, \dots, x_n$ sebagai berikut:

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{t=1}^{n} f(x_t; \theta)$$

3. Menentukan fungsi ln dari $L(\theta)$ atau dapat ditulis sebagai $\ln L(\theta)$
4. Memaksimumkan fungsi $\ln L(\theta)$

Memaksimumkan fungsi $\ln L(\theta)$ dapat diperoleh dengan cara mencari turunan parsial pertama dari parameter yang diestimasi, kemudian disamadengankan dengan nilai nol. Untuk mempermudah perhitungan secara matematis, umumnya digunakan fungsi *log – likelihood* seperti yang dapat dilihat pada persamaan (2.17) sebagai berikut:

$$l(\theta) = lnL(\theta) = \sum_{i=1}^{n} lnf(x_i|\theta). \tag{2.42}$$

Jika pada penggunaan metode MLE menghasilkan bentuk yang tidak *closed form*, maka untuk memaksimumkan persamaan dilanjutkan dengan iterasi numerik yaitu dengan pendekatan optimasi algoritma *Iteratively Reweighted Least Squares* (IRLS).

2.13 Algoritma *Iteratively Reweighted Least Square* (IRLS)

Berdasarkan fungsi kepadatan peluang pada (2.11) dan (2.14), maka bentuk *log-likelihood* dari fungsi tersebut diberikan sebagai berikut:

$$l(\theta, \alpha) = \sum_{t=1}^{n} \frac{w_t(y_t\theta_t + c(\theta))}{\alpha} + \sum_{t=1}^{n} d\left(y_t, \frac{\alpha}{w_t}\right)$$

dimana θ_t merupakan parameter kanonik dengan *mean* dan *varians* sebagai berikut:

$$E(Y_t) = -\frac{\partial c(\theta)}{\partial \theta_t} = -c'(\theta_t) = \mu_t$$

$$Var(Y_t) = -\frac{\partial^2 c(\theta)}{\partial \theta_t^2 w_t} = -\frac{\phi}{w_t} c'(\theta_t) = \frac{\phi}{w_t} V(\mu_t)$$

dengan fungsi link g sebagai berikut:

$$g(\mu_t) = g(-c^T(\theta_t)) = \boldsymbol{X_t^T \beta} = \boldsymbol{\eta_t}$$

dimana $\boldsymbol{\beta^T} = (\beta_0, \beta_1, \beta_2, \dots, \beta_v)$ sebagai parameter dalam model. Contoh algoritma IRLS untuk menentukan estimasi $\boldsymbol{\beta}$ adalah sebagai berikut:

a. Inisialisasi memilih nilai awal untuk $\hat{\mu}_t^{(0)}$ dan $\hat{\eta}_t^{(0)}$ untuk $t = 1,2, \dots, n$
b. Untuk $t = 1,2, \dots, n$

$$\hat{z}_t^{(1)} = \hat{\eta}_t^{(0)} + (y_t - \hat{\mu}_t^{(0)}) g'(\hat{\mu}_t^{(0)})$$

dan

$$\widehat{w}_t^{(1)} = \frac{w_t}{\left\{g'\left(\hat{\mu}_t^{(0)}\right)\right\}^2 v\left(\hat{\mu}_t^{(0)}\right)}$$

dengan $\boldsymbol{W}^{(\mathbf{1})} = diag\left(\widehat{w}_1^{(1)}, \dots, \widehat{w}_n^{(1)}\right)$.

c. Menghitung estimasi

$$\widehat{\boldsymbol{\beta}}^{(\mathbf{1})} = \left(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}^{(\mathbf{1})}\boldsymbol{X}\right)^{-\mathbf{1}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}^{(\mathbf{1})}z^{(1)}$$

d. Mendefinisikan vektor

$$\widehat{\boldsymbol{\eta}}^{(\mathbf{1})} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}^{(\mathbf{1})}$$

$$\widehat{\boldsymbol{\mu}}^{(\mathbf{1})} = \boldsymbol{g}\left(\widehat{\boldsymbol{\eta}}^{(\mathbf{1})}\right)$$

e. Ulangi langkah sampai dengan e menggunakan nilai $\widehat{\boldsymbol{\mu}}^{(\mathbf{1})}$ dan $\widehat{\boldsymbol{\eta}}^{(\mathbf{1})}$ sehingga akan diperoleh nilai baru $\widehat{\boldsymbol{\mu}}^{(\mathbf{2})}$ dan $\widehat{\boldsymbol{\eta}}^{(\mathbf{2})}$

f. Ulangi langkah e sampai diperoleh toleransi sebagai berikut:

$$\left|\hat{\mu}^{(t)} - \hat{\mu}^{(t-1)}\right| < \varepsilon_\mu$$

$$\left|\hat{\eta}^{(t)} - \hat{\eta}^{(t-1)}\right| < \varepsilon_\eta$$

Estimasi model GARMA menggunakan IRLS telah dilakukan oleh Benjamin, dkk., berdasarkan proses yang dilakukan oleh (Green 1984). Misalkan parameter yang akan diestimasi yaitu $\boldsymbol{\gamma}^{\boldsymbol{T}} = (\boldsymbol{\beta}^{\boldsymbol{T}}, \boldsymbol{\phi}^{\boldsymbol{T}}, \boldsymbol{\theta}^{\boldsymbol{T}})$. Ketiga parameter tersebut diestimasi dengan menggunakan *Maximum Likelihood Estimation* (MLE), namun karena bentuk yang dihasilkan tidak *closed form* maka optimasi parameter dilanjutkan dengan menggunakan algoritma IRLS (Garvey, Book, and Covert 2015). Adapun fungsi *log – likelihood* model GARMA dapat dilihat pada persamaan (2.21) sebagai berikut:

$$l(\theta) = lnL(\theta) = \sum_{i=1}^{n} lnf(x_i|\theta).$$

Berdasarkan fungsi *score* pada persamaan $U(\gamma) = \left(\frac{\partial l}{\partial \beta}, \frac{\partial l}{\partial \phi}, \frac{\partial l}{\partial \theta}\right)$ yang dapat diberikan sebagai berikut:

$$U(\gamma) = \sum_{t=m+1}^{n} \frac{\partial \mu_t}{\partial \eta_t}\frac{\partial \eta_t}{\partial \gamma}\frac{1}{v_t}(y_t - \mu_t) \tag{2.43}$$

dimana, $v_t = Var(y_t|\boldsymbol{D_t}) = \varphi Var(\mu_t) = \varphi b''(\theta_t)$. Langkah yang dilakukan selanjutnya adalah memaksimumkan fungsi *log-likelihood* l dengan menggunakan algoritma *Fisher Scoring* dengan persamaan sebagai berikut:

$$\gamma^{(k+1)} = \gamma^{(k)} + \alpha I^{-1}(\gamma^{(k)})U(\gamma^{(k)}) \tag{2.44}$$

dengan, $0 < \alpha \leq 1$ dan $I(\gamma) = -E\left(\frac{\partial^2 l}{\partial\gamma\partial\gamma}\right) = \frac{1}{\varphi}\frac{\partial\eta}{\partial\gamma}\boldsymbol{W}\frac{\partial\eta}{\partial\gamma}$ yang disebut dengan matriks informasi *Fisher*, $\boldsymbol{w} = diag(w_t)$ dengan

$$w_t^{-1} = \left(\frac{\partial\eta}{\partial\mu}\right)^2 V(\mu_t).$$

Adapun langkah – langkah algoritma IRLS secara umum adalah sebagai berikut:

a. Diberikan $\gamma^{(k)}$ kemudian hitung nilai $\boldsymbol{\eta}^{(\boldsymbol{k})}, \boldsymbol{\mu}^{(\boldsymbol{k})}, \boldsymbol{V}(\boldsymbol{\mu})^{(\boldsymbol{k})} \left(\frac{\partial\eta}{\partial\mu}\right)^{(k)}, w^{(k)}, z^{(k)}$, dimana nilai $z^{(k)}$ dan $w^{(k)}$ dikonstruksi berdasarkan variabel dependen yang disesuaikan dengan persamaan berikut:

$$z_t = \frac{\partial\eta}{\partial\gamma}\gamma + \alpha(y_t - \mu_t)\frac{\partial\eta}{\partial\mu}$$

$$w_t^{-1} = \left(\frac{\partial\eta}{\partial\mu}\right)^2 V(\mu_t)$$

untuk $t = 1, \ldots, n$.

b. Estimasi $\gamma^{(k+1)}$ dengan meregresikan $z^{(k)}$ pada $\left(\frac{\partial\eta_t}{\partial\gamma}\right)^{(k)}$ dengan bobot $w^{(k)}$

$$\gamma^{(k+1)} = I^{-1}\left(\gamma^{(k)}\right)\left(\frac{\partial\eta_t}{\partial\gamma}\right)^{(k)} w^{(k)} z^{(k)}$$

c. *Update* k ke $k+1$ dan ulangi langkah a dan b sampai estimasi parameter konvergen.

Untuk $\eta^{(m)}$ dihitung dengan proses rekursif sementara turunan dari $\left(\frac{\partial\eta_t}{\partial\gamma}\right)^{(m)}$ dihitung dari proses rekursif parameter regresi, *autoregressive* dan *moving average*.

2.14 Uji Signifikansi Parameter

Setelah memperoleh estimasi parameter dari tiap model selanjutnya dilakukan uji signifikansi parameter. Adapun tujuan dari uji signifikansi parameter adalah untuk mengetahui apakah parameter dari model tersebut signifikan sehingga layak digunakan atau tidak. Pada penelitian ini digunakan uji-t untuk menguji signifikansi parameter. Jika β, ϕ, dan θ merupakan suatu parameter dari model Poisson GARMA dan Binomial Negatif GARMA, maka $\hat{\beta}, \hat{\phi}$, dan $\hat{\theta}$ merupakan estimasi parameter β, ϕ, dan θ. Dalam melakukan pengujian signifikansi parameter dapat dinyatakan sebagai berikut (Benjamin et al. 2003):

Hipotesis :

$H_0: \beta = 0$ (Parameter model tidak signifikan)
$H_1: \beta \neq 0$ (Parameter model signifikan).

Statistik uji :

$$t_{hitung} = \frac{\hat{\beta}}{SE(\hat{\beta})}\ ,untuk\ SE(\hat{\beta} \neq 0) \qquad (2.45)$$

dengan,

β : parameter hasil estimasi
$SE(\hat{\beta})$: *Standar error* parameter hasil estimasi.

Kriteria Pengujian :
H_0 akan ditolak apabila nilai statistika uji $|t_{hitung}| > t_{\alpha/2,(n-1)}$, yang memiliki arti parameter model signifikan, dimana n adalah jumlah data dan α adalah taraf signifikan.

2.15 Kriteria Pemilihan Model Terbaik

Pada penelitian Tugas Akhir ini digunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) sebagai cara untuk menemukan kriteria model terbaik berdasarkan tingkat akurasi yang terbaik.

a. *Root Mean Square Error* (RMSE)

RMSE adalah metode pengukuran dengan menghitung perbedaan nilai dari prediksi sebuah model sebagai estimasi atas nilai yang diobservasi. *Root Mean Square Error* merupakan hasil dari akar kuadrat *Mean Square Error*. Keakuratan metode estimasi kesalahan pengukuran ditandai dengan adanya nilai RMSE yang kecil. Metode estimasi yang mempunyai nilai RMSE lebih kecil dikatakan lebih akurat daripada metode estimasi yang mempunyai nilai RMSE lebih besar (Chai and Draxler 2014).

Cara menghitung *Root Mean Square Error* (RMSE) adalah dengan mengurangi nilai aktual dengan nilai peramalan kemudian dikuadratkan dan dijumlahkan keseluruhan hasilnya kemudian dibagi dengan banyaknya data. Hasil perhitungan tersebut selanjutnya dihitung kembali untuk mencari nilai dari akar kuadrat, atau dapat dirumuskan sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}. \qquad (2.46)$$

b. *Mean Absolut Percentage error* (MAPE)

MAPE adalah persentase kesalahan rata-rata secara mutlak (absolut). Pengertian *Mean Absolute Percentage Error* adalah pengukuran statistik tentang akurasi perkiraan (prediksi) pada metode peramalan. Pengukuran dengan menggunakan *Mean Absolute Percentage Error* (MAPE) dapat digunakan oleh masyarakat luas karena MAPE mudah dipahami dan diterapkan dalam memprediksi akurasi peramalan. MAPE memberikan informasi seberapa besar kesalahan peramalan dibandingkan dengan nilai sebenarnya dari series tersebut. Semakin kecil nilai presentasi kesalahan (*percentage error*) pada MAPE maka semakin akurat hasil peramalan tersebut.

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| 100}{n} \qquad (2.47)$$

BAB 3

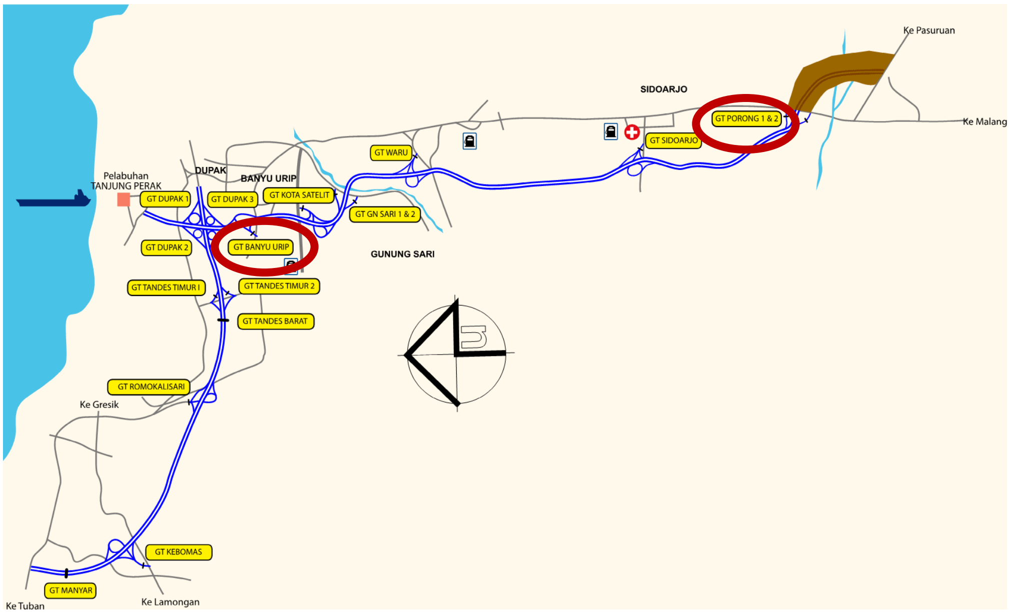
METODOLOGI

Pada bab ini dijabarkan mengenai tahapan – tahapan dalam penelitian yang meliputi penjelasan mengenai sumber data, pembahasan mengenai variabel penelitian, dilanjutkan dengan metode dan tahapan penelitian, dan menyertakan diagram alir penelitian.

3.1 Sumber Data

Pada penelitian ini, data yang digunakan adalah data jumlah kecelakaan yang terjadi di Jalan Tol Surabaya – Gempol ruas Banu Urip – Porong dengan kurun waktu lima tahun terakhir yakni dari bulan Januari 2016 – Desember 2021, data yang didapatkan sebanyak 72 data. Data jumlah kecelakaan ini merupakan data sekunder yang diperoleh dari PT. Jasa Marga Surabaya yang berlokasi di Plaza Tol Kota Satelit Jl. Mayjen Sungkono Surabaya. Pada Jalan Tol Gempol – Surabaya terdapat beberapa ruas, yang meliputi Perak, Dupak, Banyu Urip, KS, Gunungsari, Waru, Sidoarjo, Porong, Kejapanan, dan Gempol. Peta Jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong dapat dilihat pada Gambar 3.1.

Data yang digunakan untuk penelitian ini termasuk ke dalam tipe data *count* (jumlahan) yang merupakan data diskrit non – negatif atau nilainya berupa bilangan bulat non – negatif $\{0,1,2,\ldots\}$. Pada data jumlah kecelakaan ini telah memenuhi syarat agar dapat dibentuk ke dalam model Poisson GARMA (p,q) dan model Binomial Negatif GARMA (p,q), yaitu data sudah berdistribusi Poisson dan juga berdistribusi Binomial Negatif.



Gambar 3. 1 Peta Jalan Tol Gempol – Surabaya

3.2 Variabel Penelitian

Variabel yang digunakan dalam mengindikasikan jumlah kecelakaan yang terjadi di jalan tol Gempol – Surabaya ruas Banyu Urip – Porong yaitu variabel Y, yang ditunjukkan pada tabel 3.1 seperti yang sudah tertera. Untuk data kecelakaan di Jalan Tol Gempol Surabaya ruas Banyu Urip – Porong dapat dilihat pada Lampiran 1.

Tabel 3.1 Struktur Data Jumlah Kecelakaan

Tahun	Bulan	Jumlah Kecelakaan (Y)
2016	Januari Februari ⋮ Desember	Y_1 Y_2 ⋮ Y_{12}
2017	Januari Februari ⋮ Desember	Y_{13} Y_{14} ⋮ Y_{24}
⋮	⋮	⋮
2021	Januari Februari ⋮ Desember	Y_{61} Y_{62} ⋮ Y_{72}

Pada Tabel 3.1 dapat dilihat bahwa variabel Y_1 mengindikasikan data jumlah kecelakaan yang terjadi pada bulan Januari tahun 2016, Y_2 mengindikasikan data jumlah kecelakaan yang terjadi pada bulan Februari tahun 2016 dan seterusnya hingga Y_{72} mengindikasikan data jumlah kejadian kecelakaan yang terjadi pada bulan Desember tahun 2021.

3.3 Metode dan Tahapan Penelitian

1. Studi Literatur

Pada tahap ini dilakukan pencarian serta pengumpulan referensi yang menunjang pengerjaan Tugas Akhir dan pemahaman lebih mendalam tentang metode yang akan digunakan. Literatur yang dipakai adalah buku – buku, jurnal ilmiah, tugas akhir atau tesis yang berkaitan dengan permasalahan, serta artikel yang berasal dari media elektronik.

2. Pengumpulan dan Analisis Data

Pada tahap ini dilakukan pengumpulan data jumlah kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong yang diperoleh dari PT. Jasa Marga Surabaya. Periode yang digunakan adalah Januari 2016 – Desember 2021.

3. Analisis Deskriptif Terhadap Data

Pada tahap ini dilakukan perhitungan nilai rata – rata, varian, standar deviasi, nilai maksimum, nilai minimum, serta *plot time series* dari data yang bertujuan untuk mengeahui kestasioneran dan ada tidaknya unsur musiman dalam data yang digunakan.

4. Identifikasi Model Sementara

Pada tahap ini dilakukan analisis *plot time series* data, identifikasi orde (p, q), seta pembentukan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) dari orde yang sudah didapatkan sebelumnya.

5. Estimasi Parameter dan Uji Signifikansi

Pada tahap ini dilakukan estimasi parameter model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) dan pendekatan algoritma *Iteratively Reweighted Least Square* (IRLS). Setelah diperoleh nilai parameter, selanjutnya dilakukan uji signifikansi terhadap parameter – parameter yang sudah didapatkan, sehingga diperoleh model dengan parameter yang signifikan dan dilanjutkan ke dalam tahap peramalan.

6. Peramalan

Pada tahap ini dilakukan peramalan terhadap jumlah kejadian kecelakaan yang akan terjadi pada bulan Januari 2022 – Desember 2022 dengan model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q). Setelah diperoleh hasil peramalan, maka ditentukan model terbaik berdasarkan tingkat akurasi RMSE.

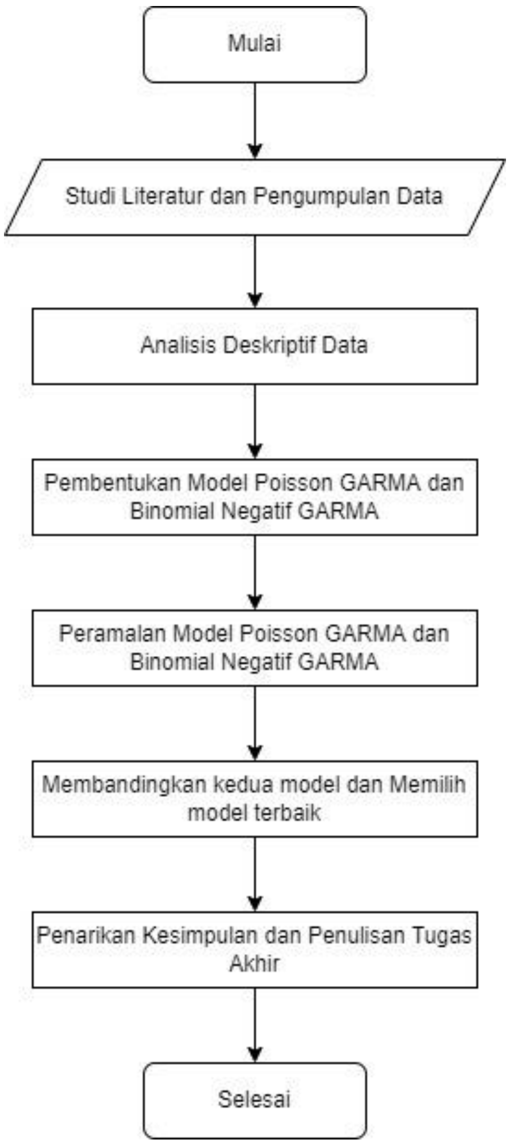
7. Penarikan Kesimpulan

Setelah langkah 1 hingga 5 dilakukan, maka dilakukan penarikan kesimpulan dari pembahasan yang sudah dilakukan sebelumnya.

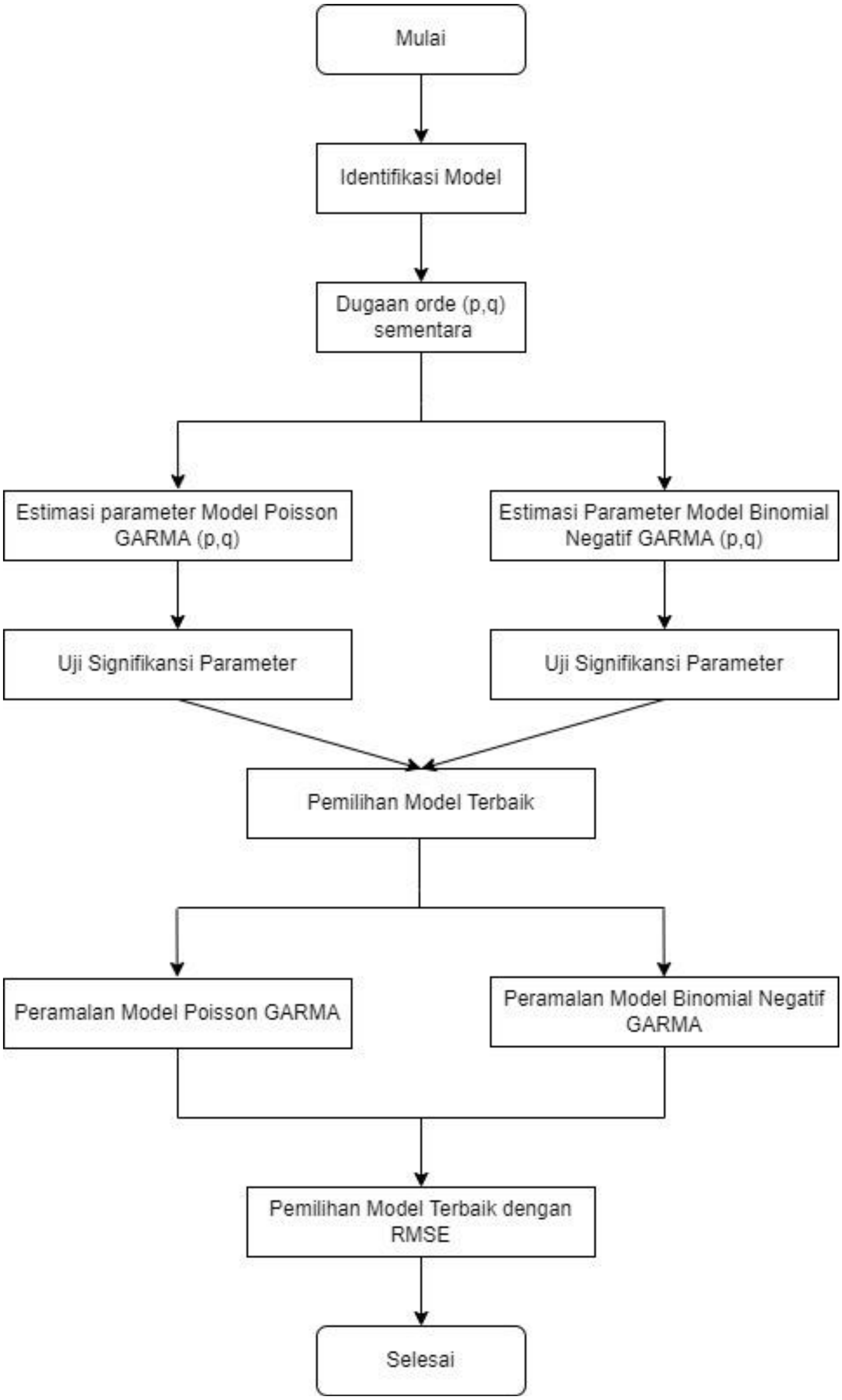
8. Penulisan Laporan Tugas Akhir

Penulisan laporan tugas akhir dilakukan dari awal penelitian hingga waktu yang ditentukan.

3.4 Diagram Alir Penelitian



Gambar 3. 2 Diagram Alir Pengerjaan Tugas Akhir



Gambar 3. 3 Diagram Alir Peramalan Model GARMA

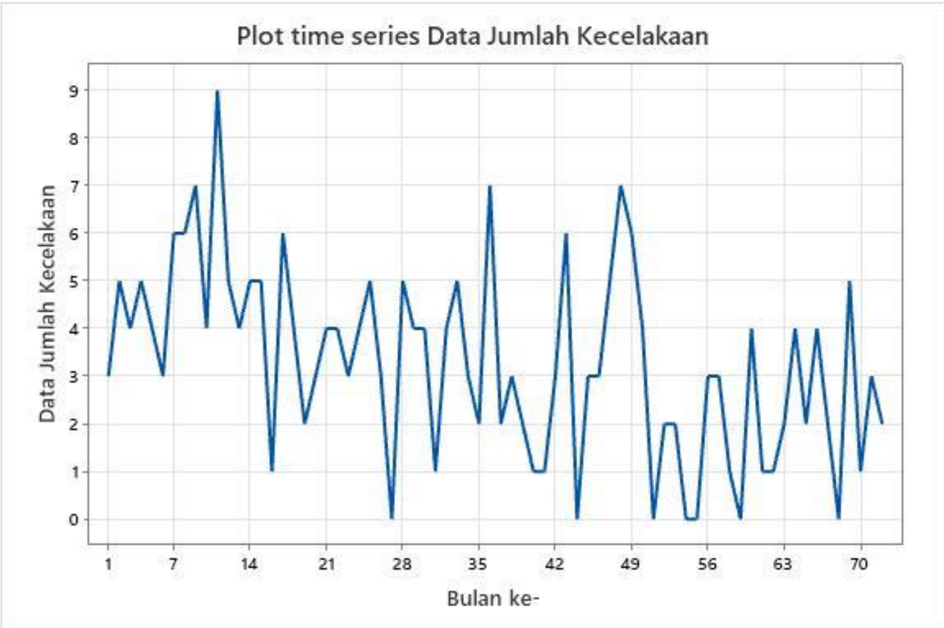
BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini dibahas mengenai tahapan dalam melakukan peramalan jumlah kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong. Tahapan peramalan dimulai dari analisis deskriptif, dilanjutkan dengan identifikasi model, estimasi parameter dengan menggunakan *Maximum Likelihood Estimation* (MLE) dengan optimasi parameter menggunakan *Iteratively Reweighted Least Square* (IRLS) serta dilanjutkan dengan melakukan peramalan selama satu tahun kedepan yaitu untuk bulan Januari 2022 – Desember 2022.

4.1 Analisis Deskriptif

Pada sub bab ini dibahas mengenai estimasi parameter dan peramalan model Poisson , GARMA (p,q) serta model Binomial Negatif GARMA (p,q) yang diaplikasikan pada data jumlah kejadian kecelakaan di jalan tol Surabaya – Gempol ruas Banyu Urip – Porong. Langkah awal yang harus dilakukan yaitu menentukan model terbaik dari data yang digunakan, kemudian dilanjutkan dengan mengestimasi parameter dan mendapatkan hasil peramalan. Model terbaik Poisson GARMA (p,q) dan Binomial Negatif GARMA (p,q) diperoleh berdasarkan hasil identifikasi awal dari model *Autoregressive Integrated Moving Average* (ARIMA). Untuk memperoleh identifikasi model Poisson GARMA (p,q) dan Binomial Negatif GARMA (p,q), tahap pertama yang dilakukan adalah dengan melihat *plot time series* dari data yang diamati. Data yang digunakan adalah data jumlah kecelakaan di jalan tol periode Januari 2016 – Desember 2021 yaitu sebanyak 72 data dengan pembagian 60 data sebagai data *in sample* yang dimulai dari bulan Januari 2016 – Desember 2020 dan 12 data sebagai data *out sample* mulai dari bulan Januari 2021 – Desember 2021 dan dapat dilihat pada Lampiran 1. *Plot* data *time series* untuk data jumlah kecelakaan di jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong yang selanjutnya disebut dengan data observasi Y dapat dilihat pada Gambar 4.1. Pada Gambar 4.1 terlihat bahwa *plot* grafik data fluktuatif atau berubah sepanjang waktu periode dan tidak terlalu menyebar pada sekitaran nilai *mean* (rata – rata).



Gambar 4. 1 Time Series Data Jumlah Kecelakaan

Setelah melihat *plot time series* data jumlah kecelakaan, selanjutnya dilihat tabel statistik deskriptif yang tertera pada Tabel 4.1. Tabel statistik deskriptif bertujuan untuk memberikan gambaran mengenai variabel yang akan diteliti. Pengolahan statistik deskriptif menunjukkan tentang ukuran sampel data yang akan diteliti seperti rata – rata (mean), simpangan baku (standar deviation), nilai maksimum, nilai minimum, dan varians. Rata – rata atau *mean* merupakan sebuah nilai yang mewakili suatu data yang diperoleh dari jumlah keseluruhan data dibagi dengan banyak data. *Standar Deviation* atau simpangan baku adalah nilai ukur penyebaran data statistik yang digunakan untuk menentukan seberapa dekat data sampel dengan nilai *mean* pada penelitian. Varians adalah ukuran dari seberapa jauh penyebaran data dari nilai rata – ratanya. Nilai maksimum merupakan nilai yang paling tinggi dari sebuah data, sedangkan nilai minimum merupakan nilai yang paling kecil yang dihasilkan dari sebuah data.

Pada Tabel 4.1 nilai *mean* atau rata – rata kecelakaan yang terjadi di jalan tol Surabaya – Gempol ruas Banyu Urip – Porong selama 72 bulan adalah 3,29166667. Jumlah kecelakaan selama kurun waktu enam tahun terakhir tidak pernah lebih dari 9 kecelakaan, dengan jumlah kecelakaan tertinggi terjadi pada bulan Juli 2016. Meningkatnya jumlah kecelakaan pada bulan Juli 2016 ini disebabkan oleh adanya faktor arus mudik Hari Raya Idul Fitri. Jumlah kecelakaan tertinggi juga terjadi pada bulan Desember 2018 dan Desember 2019 yang disebabkan oleh libur Hari Raya Natal dan Tahun Baru sehingga mengakibatkan arus mudik padat.

Pada bulan Desember tahun 2020 jumlah kecelakaan yang terjadi di jalan tol sebanyak 2 kecelakaan, hal ini dipengaruhi faktor Pembatasan Sosial Berskala Besar (PSBB) oleh pemerintah akibat adanya wabah *Covid-19* sehingga tidak banyak orang yang melakukan perjalanan atau aktifitas diluar dan lebih memilih untuk tinggal di rumah. Nilai varians dan standar deviasi dari data jumlah kecelakaan sebesar 3,984 dan 1,996. Nilai varians yang tidak terlalu tinggi ini mengindikasikan bahwa data tersebut tidak memiliki banyak keragaman atau bersifat homogen dan *plot* data tidak terlalu fluktuatif atau berubah – ubah secara signifikan.

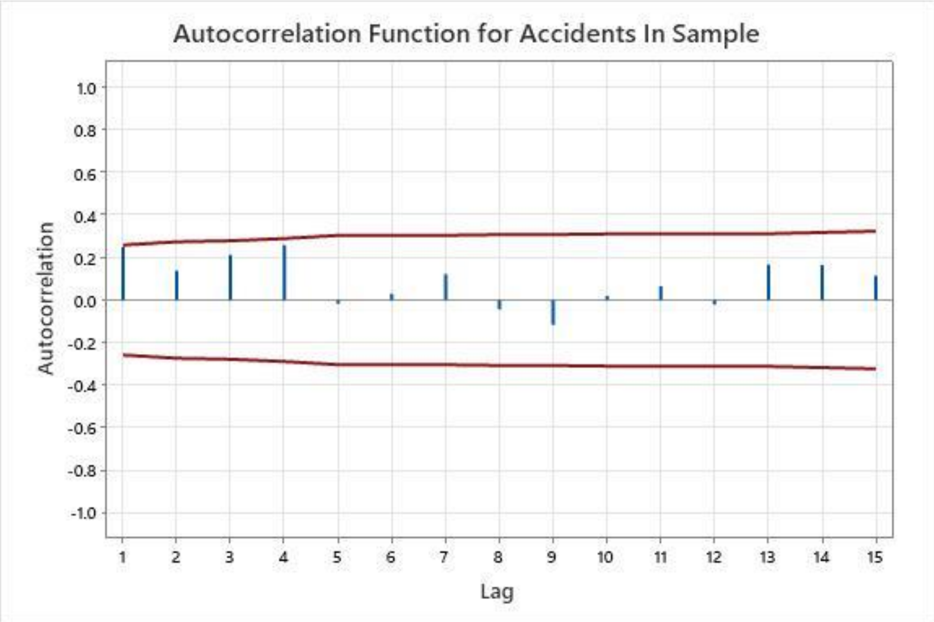
Tabel 4.1 Statistika Deskriptif Data Jumlah Kecelakaan

Variabel	Rata – rata	Simpangan Baku	Varians	Minimum	Maksimum
Jumlah Kecelakaan	3.29166667	1.996	3.984	0.000	9.000

Pada sub bab 4.2 dijelaskan mengenai identifikasi model yang bertujuan untuk mengetahui orde p dan q dengan menggunakan model ARIMA sebagai dugaan orde yang dakan dipakai dalam pembentukan model GARMA. Orde p dan q yang sudah diketahui nantinya dibentuk ke dalam model Poisson GARMA (p, q) dan model Binomial Negatif GARMA (p, q) untuk dilakukan peramalan terhadap data jumlah kecelakaan yang digunakan. Untuk mengetahui nilai orde p dan q ARIMA dapat dilihat dari *lag* yang keluar dalam *plot* ACF dan PACF nya seperti yang dibahas pada sub bab 4.2 sebagai berikut :

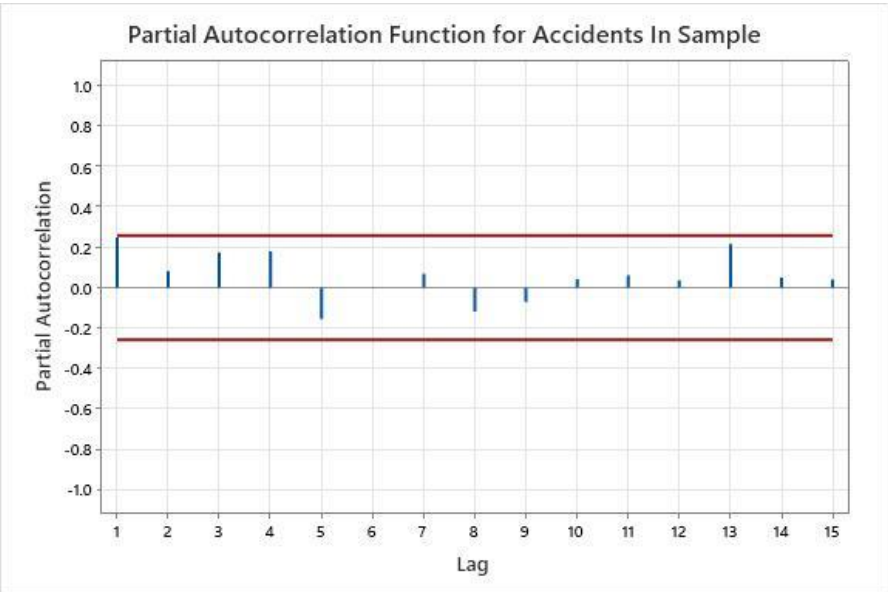
4.2 Identifikasi Model

Dalam memperoleh identifikasi model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) maka hal yang harus dilakukan adalah dengan melihat *plot Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) data. Pada Gambar 4.2 terlihat bahwa *plot* ACF pada *lag* pertama hampir melewati atau keluar dari *significant limit*. Namun, dikarenakan tidak ada *lag* yang keluar dari *significant limit* maka *plot* ACF dapat dikatakan memiliki pola *tentative.* Dari *plot* ACF mengindikasikan bahwa data tidak stasioner terhadap nilai *mean* atau rata – ratanya, hal ini dilihat dari penurunan *lag* yang lambat untuk menuju nol. Selanjutnya dilihat *plot* PACF seperti yang ada pada Gambar 4.3.



Gambar 4.2 Plot ACF Data Jumlah Kecelakaan

Pada *plot* PACF seperti pada Gambar 4.3 terlihat bahwa tidak ada *lag* yang keluar atau melebihi dari *significant limit*, serta *lag* pada *plot* PACF menurun secara perlahan menuju nol yang mengindikasikan bahwa data tidak stasioner dalam varians. Dengan melihat pola ACF dan PACF maka dapat disimpulkan bahwa tidak ada proses *differencing* sehingga dapat dinotasikan nilai $d = 0$. Sehingga diperoleh model awal ARIMA yang memenuhi adalah $(1,0,1)$ atau dapat ditulis ARMA $(1,1)$. Dari model awal tersebut tidak menutup kemungkinan bahwa ada model lain yang dapat terbentuk, seperti model ARMA $(0,1)$ dan ARMA $(1,0)$.



Gambar 4.3 Plot PACF Data Jumlah Kecelakaan

Distribusi data Jumlah Kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong dapat dilihat dengan menggunakan bantuan aplikasi *Easyfit.* Dengan melakukan uji *Goodness of Fit* Kolmogrov Smirnov pada Gambar 4.4 dan 4.5 terlihat bahwa distribusi yang menempati posisi pertama adalah distribusi *Uniform* dengan nilai $a = 0$ dan $b = 6$, dilanjutkan dengan distribusi Geometrik pada posisi kedua, distribusi Binomial Negatif pada posisi ketiga dengan nilai $n = 17$ dan $p = 0{,}83783$, dan pada posisi keempat terdapat distribusi Poisson dengan nilai rata – rata atau $\lambda = 3{,}2917$. Oleh karena model yang digunakan dalam penelitian ini adalah model GARMA dengan syarat data berdistribusi keluarga eksponensial, maka model GARMA yang kemungkinan dapat dibentuk ke dalam sebuah peramalan adalah model Poisson GARMA dan Binomial Negatif GARMA.

#	Distribution	Parameters
1	D. Uniform	a=0 b=6
2	Geometric	p=0,23301
3	Neg. Binomial	n=17 p=0,83783
4	Poisson	λ=3,2917
5	Bernoulli	No fit (data max > 1)
6	Binomial	No fit
7	Hypergeometric	No fit
8	Logarithmic	No fit (data min < 1)

Gambar 4. 4 Hasil Uji Distribusi Data

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	D. Uniform	0,22421	1	26,096	4
2	Geometric	0,34047	4	12,0	3
3	Neg. Binomial	0,24117	3	3,9959	1
4	Poisson	0,23633	2	4,2371	2
5	Bernoulli	No fit (data max > 1)			
6	Binomial	No fit			
7	Hypergeometric	No fit			
8	Logarithmic	No fit (data min < 1)			

Gambar 4. 5 Hasil Uji Distribusi Data dengan Software Easyfit

Pada Gambar 4.4 telah ditunjukkan bahwa data berdistribusi Poisson dan Binomial Negatif, sehingga data dapat dibentuk ke dalam model Poisson GARMA dan Binomial Negatif GARMA. Setelah dilakukan analisis deskriptif dan menentukan orde p dan q pada identifikasi model pada sub bab sebelumnya, maka selanjutnya dapat dijabarkan mengenai Model Poisson GARMA (p, q) dan Model Binomial Negatif GARMA (p, q).

4.3 Model Poisson GARMA (p, q)

Model Poisson GARMA (p, q) diusulkan oleh (Benjamin, dkk, 2003) dibangun dari distribusi bersyarat Poisson y_t yang ditunjukkan pada persamaan (2.11). Pada persamaan model GARMA (2.40) tersebut melibatkan variabel prediktor X_t^T, tetapi dalam penelitian ini tidak melibatkan variabel prediktor X_t^T sehingga menurut Marinho dan Ricardo (2015) variabel prediktor tersebut dapat digantikan oleh sebuah konstanta β_0. Sehingga diperoleh persamaan seperti yang tertera pada persamaan (4.1) sebagai berikut:

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{p} \phi_j \left(\ln y_{t-j}^* - \beta_0\right) + \sum_{j=1}^{q} \theta_j \left\{\ln \frac{y_{t-j}^*}{\mu_{t-j}}\right\} \tag{4.1}$$

dengan $y_{t-j}^* = \max\,(y_{t-j}, c)$ dan $0 < c < 1$. Jika persamaan (4.1) dibentuk kedalam sebuah matriks maka dapat ditulis sebagai berikut:

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \cdots & \ln y_{2-p}^* - \beta_0 & \ln \frac{y_1^*}{\mu_1} & \cdots & \ln \frac{y_{2-q}^*}{\mu_{2-q}} \\ 1 & \ln y_2^* - \beta_0 & \cdots & \ln y_{3-p}^* - \beta_0 & \ln \frac{y_2^*}{\mu_2} & \cdots & \ln \frac{y_{3-q}^*}{\mu_{3-q}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \cdots & \ln y_{(T-1)-p}^* - \beta_0 & \ln \frac{y_{T-1}^*}{\mu_{T-1}} & \cdots & \ln \frac{y_{(T-1)-q}^*}{\mu_{(T-1)-q}} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \\ \vdots \\ \phi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix}$$

atau dapat ditulis menjadi persamaan

$$\boldsymbol{v_t} = \boldsymbol{X\gamma}$$

dengan,

$$\boldsymbol{v_t} = \begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \cdots & \ln y_{2-p}^* - \beta_0 & \ln \frac{y_1^*}{\mu_1} & \cdots & \ln \frac{y_{2-q}^*}{\mu_{2-q}} \\ 1 & \ln y_2^* - \beta_0 & \cdots & \ln y_{3-p}^* - \beta_0 & \ln \frac{y_2^*}{\mu_2} & \cdots & \ln \frac{y_{3-q}^*}{\mu_{3-q}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \cdots & \ln y_{(T-1)-p}^* - \beta_0 & \ln \frac{y_{T-1}^*}{\mu_{T-1}} & \cdots & \ln \frac{y_{(T-1)-q}^*}{\mu_{(T-1)-q}} \end{bmatrix},$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \beta_0 \\ \phi_1 \\ \vdots \\ \phi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix},$$

$t = 2,3,4, \dots T$. Jika nilai μ_t yang dicari maka persamaan (4.1) dapat berubah menjadi persamaan (4.2) seperti yang tertera sebagai berikut:

$$\mu_t = \exp\left[\beta_0 + \sum_{j=1}^{p} \phi_j \left\{\ln y_{t-j}^* - \beta_0\right\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right]. \tag{4.2}$$

Pada sub bab 4.2 telah diperoleh nilai orde p dan q yang memungkinkan untuk dibentuk menjadi model Poisson GARMA yaitu orde (1,0), (0,1), dan (1,1) atau dapat ditulis model Poisson GARMA (1,0), Poisson GARMA (0,1) dan Poisson GARMA (1,1). Pada sub bab selanjutnya dijabarkan beberapa model khusus Poisson GARMA (p, q) dengan orde p dan q yang sudah ditentukan sebelumnya.

4.3.1 Model Poisson GARMA $(\mathbf{1}, \mathbf{0})$

Pada persamaan (4.1) jika disubtitusikan nilai orde $p = 1$ dan orde $q = 0$ maka menghasilkan persamaan model Poisson GARMA (1,0) seperti yang dapat dituliskan pada persamaan (4.3) sebagai berikut:

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{1} \phi_j \left(\ln y_{t-j}^* - \beta_0\right) + \sum_{j=1}^{0} \theta_j \left\{\ln \frac{y_{t-j}^*}{\mu_{t-j}}\right\}$$

$$\ln(\mu_t) = \beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\}. \tag{4.3}$$

Berdasarkan persamaan (4.3) jika dijabarkan dengan memasukkan nilai $t = 2,3 \dots, T$ maka terbentuk sebuah persamaan yang dapat ditulis sebagai berikut:

$$\begin{gathered} \ln(\mu_2) = \beta_0 + \phi_1\{\ln y_1^* - \beta_0\} \\ \ln(\mu_3) = \beta_0 + \phi_1\{\ln y_2^* - \beta_0\} \\ \vdots \\ \ln(\mu_t) = \beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\}. \end{gathered}$$

Persamaan diatas dapat ditulis menjadi persamaan (4.4) sebagai berikut:

$$\mu_t = \exp(\beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\}). \tag{4.4}$$

Selanjutnya dari persamaan (4.4) jika dibentuk ke dalam sebuah matriks maka bentuknya menjadi seperti matriks pada persamaan (4.5) sebagai berikut:

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 \\ 1 & \ln y_2^* - \beta_0 \\ 1 & \ln y_3^* - \beta_0 \\ \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \end{bmatrix}. \tag{4.5}$$

Berdasarkan bentuk matriks pada persamaan (4.5) dapat dilihat bahwa parameter β_0 dan ϕ_1 adalah parameter yang akan dicari nilainya atau diestimasikan.

4.3.2 Model Poisson GARMA $(\mathbf{0}, \mathbf{1})$

Jika mensubtitusikan nilai orde $p = 0$ dan orde $q = 1$ pada persamaan (4.1) maka persamaan tersebut akan berubah menjadi persamaan model Poisson GARMA (0,1) atau dapat dituliskan seperti pada persamaan (4.6) sebagai berikut:

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{0} \phi_j \, (\ln y_{t-j}^* - \beta_0) + \sum_{j=1}^{1} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}$$

$$\ln(\mu_t) = \beta_0 + \theta_1 \left\{\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right\}. \tag{4.6}$$

Jika persamaan (4.6) dijabarkan dengan memasukkan nilai $t = 2,3,\dots,T$ maka akan membentuk sebuah persamaan yang dapat dituliskan sebagai berikut:

$$\begin{gathered}
\ln(\mu_2) = \beta_0 + \theta_1 \left\{\frac{\ln y_1^*}{\mu_1}\right\} \\
\ln(\mu_3) = \beta_0 + \theta_1 \left\{\frac{\ln y_2^*}{\mu_2}\right\} \\
\vdots \\
\ln(\mu_t) = \beta_0 + \theta_1 \left\{\frac{\ln y_{T-1}^*}{\mu_{T-1}}\right\}
\end{gathered}$$

dari penjabaran persamaan diatas dapat dicari nilai μ_t seperti yang dapat ditulis pada persamaan (4.7) sebagai berikut:

$$\mu_t = \exp\left(\beta_0 + \theta_1 \left\{\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right\}\right). \tag{4.7}$$

Selanjutnya dari persamaan (4.7) jika dibentuk ke dalam sebuah matriks maka bentuknya menjadi seperti persamaan (4.8) sebagai berikut:

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_t) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* / \mu_0 \\ 1 & \ln y_2^* / \mu_1 \\ 1 & \ln y_3^* / \mu_2 \\ \vdots & \vdots \\ 1 & \ln y_{t-1}^* / \mu_{t-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \theta_1 \end{bmatrix}. \tag{4.8}$$

Berdasarkan matriks yang terbentuk pada persamaan (4.8) dapat dilihat bahwa parameter β_0 dan θ_1 adalah parameter yang akan dicari nilainya.

4.3.3 Model Poisson GARMA $(\mathbf{1},\mathbf{1})$

Dengan cara yang sama seperti model Poisson GARMA $(1,0)$ dan $(0,1)$, jika mensubtitusikan nilai orde $p = 1$ dan orde $q = 1$ pada persamaan (4.3) maka persamaan tersebut akan berubah menjadi persamaan model Poisson GARMA $(1,1)$, atau dapat dituliskan seperti pada persamaan (4.9).

$$\begin{gathered}
\ln(\mu_t) = \beta_0 + \sum_{j=1}^{1} \phi_j \, (\ln y_{t-j}^* - \beta_0) + \sum_{j=1}^{1} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\} \\
\ln(\mu_t) = \beta_0 + \phi_1 \{\ln y_{t-1}^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right\}.
\end{gathered} \tag{4.9}$$

Jika persamaan (4.9) dijabarkan dengan memasukkan nilai $t = 2,3,\dots,T$ maka terbentuk sebuah persamaan sebagai berikut:

$$\ln(\mu_2) = \beta_0 + \phi_1\{\ln y_1^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_1^*}{\mu_1}\right\}$$
$$\ln(\mu_3) = \beta_0 + \phi_1\{\ln y_2^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_2^*}{\mu_2}\right\}$$
$$\vdots$$
$$\ln(\mu_T) = \beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_{T-1}^*}{\mu_{T-1}}\right\}$$

dari penjabaran persamaan diatas dapat dicari nilai μ_t atau dapat ditulis menjadi persamaan (4.10) sebagai berikut:

$$\mu_t = \exp\left(\beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right\}\right). \tag{4.10}$$

Menggunakan cara yang sama seperti sebelumnya, persamaan (4.10) dapat diubah ke dalam bentuk sebuah matriks seperti yang ditunjukkan pada persamaan (4.11).

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \ln y_1^* / \mu_0 \\ 1 & \ln y_2^* - \beta_0 & \ln y_2^* / \mu_1 \\ 1 & \ln y_3^* - \beta_0 & \ln y_3^* / \mu_2 \\ \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \ln y_{T-1}^* / \mu_{T-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \\ \theta_1 \end{bmatrix}. \tag{4.11}$$

Sehingga dapat dilihat bahwa parameter yang nilainya dicari untuk model Poisson GARMA (1,1) adalah $\beta_0, \phi_1, \theta_1$.

Pada sub bab 4.4 dijelaskan mengenai model umum dari Binomial Negatif GARMA (p, q) yang dilanjutkan dengan menjabarkan beberapa model khusus dari Binomial Negatif GARMA (p, q) dengan orde p dan q tertentu yang sudah ditentukan pada sub bab pembahasan sebelumnya.

4.4 Model Binomial Negatif GARMA $(\boldsymbol{p}, \boldsymbol{q})$

Model Binomial Negatif GARMA (p, q) diusulkan oleh Benjamin pada tahun 2003 dibangun dari distribusi bersyarat Binomial Negatif y_t pada persamaan (2.14). Pada persamaan model (2.41) tersebut melibatkan variabel prediktor X_t^T, tetapi dalam penelitian ini tidak melibatkan variabel prediktor X_t^T sehingga menurut Marinho dan Ricardo (2015) variabel prediktor tersebut dapat digantikan oleh sebuah konstanta β_0. Sehingga diperoleh persamaan seperti yang tertera pada persamaan (4.12) sebagai berikut:

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{p} \phi_j \left(\ln y_{t-j}^* - \beta_0\right) + \sum_{j=1}^{q} \theta_j \left\{\ln\left(y_{t-j}^*\right) - \mu_{t-j}\right\} \tag{4.12}$$

dengan $y_{t-j}^* = \max\left(y_{t-j}, c\right)$ dan $0 < c < 1$. Jika persamaan (4.12) dibentuk ke dalam sebuah matriks maka persamaan tersebut memiliki bentuk seperti yang dapat dituliskan sebagai berikut :

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} =$$

$$\begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \cdots & \ln y_{2-p}^* - \beta_0 & \ln y_1^* - \ln \mu_1 & \cdots & \ln y_{2-q}^* - \ln \mu_{2-q} \\ 1 & \ln y_2^* - \beta_0 & \cdots & \ln y_{3-p}^* - \beta_0 & \ln y_2^* - \ln \mu_2 & \cdots & \ln y_{3-q}^* - \ln \mu_{3-q} \\ 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \cdots & \ln y_{T-1-p}^* - \beta_0 & \ln y_{T-1}^* - \ln \mu_{T-1} & \cdots & \ln y_{T-1-q}^* - \ln \mu_{T-1-q} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \\ \vdots \\ \phi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix}$$

atau dapat ditulis menjadi persamaan

$$\boldsymbol{v_t} = \boldsymbol{X\gamma}$$

dengan,

$$\boldsymbol{X} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \cdots & \ln y_1^* - \beta_0 & \ln \frac{y_1^*}{\mu_1} & \cdots & \ln y_1^* - \mu_1 \\ 1 & \ln y_2^* - \beta_0 & \cdots & \ln y_2^* - \beta_0 & \ln \frac{y_2^*}{\mu_2} & \cdots & \ln y_2^* - \mu_1 \\ 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \cdots & \ln y_{T-1}^* - \beta_0 & \ln \frac{y_{T-1}^*}{\mu_{T-1}} & \cdots & \ln y_{T-1}^* - \mu_{T-1} \end{bmatrix}, \ \boldsymbol{v_t} = \begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix},$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \beta_0 \\ \phi_1 \\ \vdots \\ \phi_p \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix}$$

$t = 2,3,4, \ldots T$. Jika nilai μ_t yang dicari pada persamaan (4.12) maka nilai tersebut dapat dicari seperti yang ditunjukkan pada persamaan (4.13) sebagai berikut:

$$\mu_t = \exp\left[\beta_0 + \sum_{j=1}^{p} \phi_j \left\{\ln y_{t-j}^* - \beta_0\right\} + \sum_{j=1}^{q} \theta_j \left\{\ln\left(y_{t-j}^*\right) - \mu_{t-j}\right\}\right] \tag{4.13}$$

Pada *plot* ACF dan PACF yang sudah dibahas sebelumnya di sub bab 4.2 diperoleh beberapa orde yang dapat dibentuk ke dalam model Binomial Negatif GARMA (p, q), yaitu Binomial Negatif GARMA $(1,0)$, Binomial Negatif GARMA $(0,1)$, dan Binomial Negatif GARMA $(1,1)$ yang dijabarkan sebagai berikut:

4.4.2 Model Binomial Negatif GARMA $(\mathbf{1}, \mathbf{0})$

Pada persamaan (4.12) jika disubtitusikan nilai orde $p = 1$ dan orde $q = 0$ maka menghasilkan persamaan model Binomial Negatif GARMA $(1,0)$ seperti yang dapat dituliskan pada persamaan (4.14) sebagai berikut:

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{1} \phi_j \,(\ln y^*_{t-j} - \beta_0) + \sum_{j=1}^{0} \theta_j \{\ln(y^*_{t-j}) - \mu_{t-j}\}$$
$$\ln(\mu_t) = \beta_0 + \phi_1\{\ln y^*_{t-1}\}. \tag{4.14}$$

Jika persamaan (4.14) dijabarkan dengan memasukkan nilai $t = 2,3,\dots,T$ maka akan membentuk sebuah persamaan yang dapat dituliskan sebagai berikut:

$$\ln(\mu_2) = \beta_0 + \phi_1\{\ln y^*_1 - \beta_0\}$$
$$\ln(\mu_3) = \beta_0 + \phi_1\{\ln y^*_2 - \beta_0\}$$
$$\vdots$$
$$\ln(\mu_t) = \beta_0 + \phi_1\{\ln y^*_{t-1} - \beta_0\}$$

dari penjabaran persamaan diatas dapat dicari nilai μ_t atau dapat ditulis menjadi seperti persamaan (4.15) sebagai berikut:

$$\mu_t = \exp(\beta_0 + \phi_1\{\ln y^*_{t-1} - \beta_0\}). \tag{4.15}$$

Berdasarkan persamaan (4.15) dapat dibentuk ke dalam sebuah matriks yang dapat ditunjukkan pada persamaan (4.16)

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y^*_1 - \beta_0 \\ 1 & \ln y^*_2 - \beta_0 \\ 1 & \ln y^*_3 - \beta_0 \\ \vdots & \vdots \\ 1 & \ln y^*_{T-1} - \beta_0 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \end{bmatrix}. \tag{4.16}$$

Berdasarkan persamaan (4.16) dapat dilihat bahwa parameter β_0 dan ϕ_1 adalah parameter yang akan dicari nilainya.

4.4.3 Model Binomial Negatif GARMA $(\mathbf{0},\mathbf{1})$

Pada persamaan (4.12) jika disubtitusikan nilai orde $p = 0$ dan orde $q = 1$ maka menghasilkan persamaan model Binomial Negatif GARMA $(0,1)$ seperti yang dapat dituliskan pada persamaan (4.17)

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{0} \phi_j \,(\ln y^*_{t-j} - \beta_0) + \sum_{j=1}^{1} \theta_j \{\ln(y^*_{t-j}) - \ln(\mu_{t-j})\}$$
$$\ln(\mu_t) = \beta_0 + \theta_1\{\ln y^*_{t-1} - \ln(\mu_{t-1})\}. \tag{4.17}$$

Jika persamaan (4.17) dijabarkan dengan memasukkan nilai $t = 2,3,\dots,T$ maka akan membentuk sebuah persamaan yang dapat dituliskan sebagai berikut:

$$\ln(\mu_2) = \beta_0 + \theta_1\{\ln y^*_1 - \ln(\mu_1)\}$$
$$\ln(\mu_3) = \beta_0 + \theta_1\{\ln y^*_2 - \ln(\mu_2)\}$$

$$\vdots$$

$$\ln(\mu_t) = \beta_0 + \theta_1\{\ln y_{t-1}^* - \ln(\mu_{t-1})\}$$

dari penjabaran persamaan diatas dapat dicari nilai μ_t atau dapat ditulis menjadi seperti persamaan (4.18)

$$\mu_t = \exp(\beta_0 + \theta_1\{\ln y_{T-1}^* - \ln(\mu_{T-1})\}). \tag{4.18}$$

Berdasarkan persamaan (4.18) dapat dibentuk ke dalam sebuah matriks yang dapat ditunjukkan pada persamaan (4.19)

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* - \ln(\mu_0) \\ 1 & \ln y_2^* - \ln(\mu_1) \\ 1 & \ln y_3^* - \ln(\mu_2) \\ \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \ln(\mu_{T-1}) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \theta_1 \end{bmatrix} \tag{4.19}$$

Dilihat dari persamaan (4.19) dapat diketahui bahwa parameter β_0 dan θ_1 adalah parameter yang akan dicari.

4.4.4 Model Binomial Negatif GARMA $(\mathbf{1}, \mathbf{1})$

Dengan cara yang sama seperti model Binomial Negatif GARMA $(1,0)$ dan $(0,1)$, jika mensubtitusikan nilai orde $p = 1$ dan orde $q = 1$ pada persamaan (4.12) maka persamaan tersebut akan berubah menjadi persamaan model Binomial Negatif GARMA $(1,1)$, atau dapat dituliskan seperti pada persamaan (4.20).

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{1} \phi_j \left(\ln y_{t-j}^* - \beta_0\right) + \sum_{j=1}^{1} \theta_j\{\ln(y_{t-j}^*) - \ln(\mu_{t-j})\}$$

$$\ln(\mu_t) = \beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\} + \theta_1\{\ln y_{t-1}^* - \ln(\mu_{t-1})\}. \tag{4.20}$$

Jika persamaan (4.20) dijabarkan dengan memasukkan nilai $t = 1,2,\dots,T$ maka terbentuk sebuah persamaan sebagai berikut:

$$\ln(\mu_2) = \beta_0 + \phi_1\{\ln y_1^* - \beta_0\} + \theta_1\{\ln y_1^* - \ln(\mu_1)\}$$

$$\ln(\mu_3) = \beta_0 + \phi_1\{\ln y_2^* - \beta_0\} + \theta_1\{\ln y_2^* - \ln(\mu_2)\}$$

$$\vdots$$

$$\ln(\mu_T) = \beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\} + \theta_1\{\ln y_{T-1}^* - \ln(\mu_{T-1})\}$$

dari penjabaran persamaan diatas dapat dicari nilai μ_t atau dapat ditulis menjadi persamaan (4.21).

$$\mu_t = \exp(\beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\} + \theta_1\{\ln y_{T-1}^* - \ln(\mu_{T-1})\}). \tag{4.21}$$

Menggunakan cara yang sama seperti sebelumnya, persamaan (4.21) dapat diubah ke dalam bentuk sebuah matriks seperti yang ditunjukkan pada persamaan (4.22).

$$\begin{bmatrix} \ln(\mu_2) \\ \ln(\mu_3) \\ \ln(\mu_4) \\ \vdots \\ \ln(\mu_T) \end{bmatrix} = \begin{bmatrix} 1 & \ln y_1^* - \beta_0 & \ln y_1^* - \mu_1 \\ 1 & \ln y_2^* - \beta_0 & \ln y_2^* - \mu_2 \\ 1 & \ln y_3^* - \beta_0 & \ln y_3^* - \mu_3 \\ \vdots & \vdots & \vdots \\ 1 & \ln y_{T-1}^* - \beta_0 & \ln y_{T-1}^* - \mu_{T-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \phi_1 \\ \theta_1 \end{bmatrix}. \tag{4.22}$$

Berdasarkan persamaan (4.22) dapat dilihat bahwa parameter $\beta_0, \phi_1, \theta_1$ adalah parameter yang akan dicari atau diestimasi.
Model – model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) yang telah terbentuk sebelumnya dapat dijadikan ke dalam satu tabel seperti yang dapat dilihat pada Tabel 4.2 sebagai berikut:

Tabel 4. 2 Persamaan Model Poisson dan Binomial Negatif GARMA

Model GARMA (p, q)	**Persamaan Model**
Poisson GARMA $(1,0)$	$\ln(\mu_T) = \beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\}$
Poisson GARMA $(0,1)$	$\ln(\mu_T) = \beta_0 + \theta_1 \left\{\frac{\ln y_{T-1}^*}{\mu_{T-1}}\right\}$
Poisson GARMA $(1,1)$	$\ln(\mu_T) = \beta_0 + \phi_1\{\ln y_{T-1}^* - \beta_0\} + \theta_1 \left\{\frac{\ln y_{T-1}^*}{\mu_{T-1}}\right\}$
Binomial Negatif GARMA $(1,0)$	$\mu_t = \exp(\beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\})$
Binomial Negatif GARMA $(0,1)$	$\ln(\mu_t) = \beta_0 + \theta_1\{\ln y_{t-1}^* - \ln(\mu_{t-1})\}$
Binomial Negatif GARMA $(1,1)$	$\ln(\mu_t) = \beta_0 + \phi_1\{\ln y_{t-1}^* - \beta_0\}$ $+ \theta_1\{\ln y_{t-1}^* - \ln(\mu_{t-1})\}$

Setelah membahas tentang model khusus Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q) pada sub bab (4.3) dan (4.4), maka selanjutnya dibahas mengenai estimasi dan uji signifikansi parameter yang dibahas pada sub bab 4.5.

4.5 Estimasi dan Uji Signifikansi Parameter Model

Pada sub bab ini dibahas tentang estimasi parameter model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q). Tujuan dari estimasi parameter ini yaitu untuk memperoleh nilai parameter yang akan digunakan pada model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q). Adapun nilai parameter yang dicari adalah (β_0, ϕ, θ), dimana β_0 adalah suatu konstanta, ϕ merupakan parameter *autoregressive*, dan θ merupakan parameter *moving average*. Untuk mencari nilai estimasi parameter digunakan metode yang memaksimumkan fungsi *likelihood* atau dikenal sebagai metode *Maximum Likelihood Estimation* (MLE).

4.5.1 Estimasi Parameter Model Poisson GARMA $(\boldsymbol{p}, \boldsymbol{q})$

Terdapat beberapa langkah yang harus dilakukan dalam melakukan estimasi model Poisson GARMA (p, q). Langkah pertama yaitu mengasumsikan data $y_1, y_2, \dots, y_t \sim Poisson(\mu_t)$, dimana persamaan μ_t memenuhi model seperti pada persamaan (4.23).

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{\boldsymbol{p}} \phi_j \{\ln\left(y^*_{t-j}\right) - \beta_0) + \sum_{j=1}^{\boldsymbol{q}} \theta_j \left\{\ln \frac{\left(y^*_{t-j}\right)}{\mu_{t-j}}\right\}$$

$$\mu_t = \exp(\beta_0 + \sum_{j=1}^{\boldsymbol{p}} \phi_j \{\ln\left(y^*_{t-j}\right) - \beta_0) + \sum_{j=1}^{\boldsymbol{q}} \theta_j \left\{\ln \frac{\left(y^*_{t-j}\right)}{\mu_{t-j}}\right\}. \qquad (4.23)$$

Langkah selanjutnya yaitu membentuk fungsi *partial likelihood* dari distribusi Poisson yang ditunjukkan pada persamaan (2.10). Persamaan fungsi *partial likelihood* distribusi Poisson dapat dilihat pada persamaan (4.24) sebagai berikut:

$$PL\left(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \prod_{t=1}^{n} f\left(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right)$$

$$PL\left(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \prod_{t=1}^{n} \frac{\exp(-\mu_t)\, \mu_t^{y_t}}{y_t!}. \qquad (4.24)$$

Fungsi *partial likelihood* pada persamaan (4.24) jika diubah menjadi fungsi *partial log-likelihood* akan berbentuk seperti pada persamaan (4.25) sebagai berikut:

$$l\left(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \ln PL\left(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right)$$

$$l\left(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \ln\left[\prod_{t=1}^{n} \frac{\exp(-\mu_t)\, \mu_t^{y_t}}{y_t!}\right]$$

$$l\left(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \sum_{t=1}^{n} y_t \ln \mu_t - \sum_{t=1}^{n} \mu_t - \sum_{t=1}^{n} \ln(y_t!). \qquad (4.25)$$

Jika persamaan (4.1) disubtitusikan ke dalam fungsi *partial log-likelihood* pada persamaan (4.25) maka bentuk *partial log-likelihood* dapat ditulis seperti pada persamaan (4.26)

$$l\left(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}\right) = \sum_{t=1}^{n} y_t \ln \left\{\exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j \{y^*_{t-j}\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y^*_{t-j}}{\mu_{t-j}}\right\}\right)\right\} - \qquad (4.26)$$

$$\sum_{t=1}^{n}\left\{\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)\right\}-\sum_{t=1}^{n}\ln(y_t!).$$

Langkah yang dilakukan selanjutnya adalah mendapatkan parameter $\beta_0, \phi_j, \theta_j$ dengan cara memaksimumkan fungsi *partial log-likelihood* yang sudah didapat sebelumnya pada persamaan (4.26). Untuk memperoleh nilai parameter dapat dilakukan dengan cara mencari turunan parsial pertama terhadap masing – masing parameter $\beta_0, \phi_j, \theta_j$. Turunan pertama dari *partial log-likelihood* terhadap parameter β_0 ditunjukkan pada persamaan (4.27)

$$\frac{\partial l(\mu_t)}{\partial \beta_0}=\frac{\partial l}{\partial \beta_0}\left[\sum_{t=1}^{n} y_t \ln\left\{\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)\right\}\right.$$
$$\left.-\sum_{t=1}^{n}\left\{\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)\right\}-\sum_{t=1}^{n}\ln(y_t!)\right]$$

$$\frac{\partial l(\mu_t)}{\partial \beta_0}=\sum_{t=1}^{n} y_t\left[-\sum_{j=1}^{p}\phi_j\right]-\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)$$
$$-\sum_{t=1}^{n}\left(\sum_{j=1}^{p}-\phi_j\right). \tag{4.27}$$

Menggunakan cara yang sama seperti turunan *partial log-likelihood* terhadap paremeter β_0, diperoleh turunan pertama dari *partial log-likelihood* terhadap parameter ϕ seperti yang dapat ditulis pada persamaan (4.28)

$$\frac{\partial l(\mu_t)}{\partial \phi_1}=\sum_{t=1}^{n} y_t(\ln(y_{t-1}^{*})-\sum_{t=1}^{n}(\ln(y_{t-1}^{*})\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)$$

$$\frac{\partial l(\mu_t)}{\partial \phi_2}=\sum_{t=1}^{n} y_t(\ln(y_{t-2}^{*})-\sum_{t=1}^{n}(\ln(y_{t-2}^{*})\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right)$$

$$\vdots$$

$$\frac{\partial l(\mu_t)}{\partial \phi_j}=\sum_{t=1}^{n} y_t(\ln(y_{t-j}^{*})-\sum_{t=1}^{n}(\ln(y_{t-j}^{*})\exp\left(\beta_0+\sum_{j=1}^{p}\phi_j\{\ln y_{t-j}^{*}\}+\sum_{j=1}^{q}\theta_j\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right). \tag{4.28}$$

Selanjutnya untuk turunan pertama dari *partial log-likeliood* terhadap parameter θ dapat ditulis seperti pada persamaan (4.29)

$$\frac{\partial l(\mu_t)}{\partial \theta_1} = \sum_{t=1}^{n} y_t \left(\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right) - \sum_{t=1}^{n} \left(\frac{\ln y_{t-1}^*}{\mu_{t-1}}\right) \exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j \{\ln y_{t-j}^*\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right)$$

$$\frac{\partial l(\mu_t)}{\partial \theta_2} = \sum_{t=1}^{n} y_t \left(\frac{\ln y_{t-2}^*}{\mu_{t-2}}\right) - \sum_{t=1}^{n} \left(\frac{\ln y_{t-2}^*}{\mu_{t-2}}\right) \exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j \{\ln y_{t-j}^*\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right)$$

$$\vdots$$

$$\frac{\partial l(\mu_t)}{\partial \theta_j} = \sum_{t=1}^{n} y_t \left(\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right) - \sum_{t=1}^{n} \left(\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right) \exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j \{\ln y_{t-j}^*\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right). \quad (4.29)$$

Agar fungsi dari *partial log-likelihood* dapat maksimum maka turunan pertama dari masing masing parameter disamadengankan nol seperti yang tertera pada persamaan (4.30)

$$\frac{\partial l(\mu_t)}{\partial \beta_0} = 0\ ;\ \frac{\partial l(\mu_t)}{\partial \phi_j} = 0\ ;\ \frac{\partial l(\mu_t)}{\partial \theta_j} = 0. \quad (4.30)$$

Pada persamaan (4.27), (4.28) dan (4.29) terdapat suku $\exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j \{\ln y_{t-j}^*\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right)$ yang mengakibatkan bentuk *closed form* dari β_0, ϕ_j, dan θ_j sulit untuk ditemukan sehingga jika menggunakan metode *Maximum Likelihood Estimation* nilai parameter tidak dapat diselesaikan. Oleh karena itu, untuk mengestimasi parameter $\beta_0, \phi_j, \theta_j$ digunakan optimasi dengan menggunakan algoritma *Iteratively Reweighted Least Square* (IRLS) dengan langkah – langkah melakukan iterasi IRLS dapat dilihat pada sub bab 2.12. Iterasi dilakukan hingga memperoleh parameter yang konvergen atau hingga diperoleh hasil $|\hat{\gamma}^{(t)} - \hat{\gamma}^{(t-1)}| < \varepsilon$.

4.5.2 Estimasi Parameter Model Binomial Negatif GARMA $(\boldsymbol{p}, \boldsymbol{q})$

Terdapat beberapa langkah yang harus dilakukan dalam melakukan estimasi model Binomial Negatif GARMA (p, q). Langkah pertama yaitu mengasumsikan data $y_1, y_2, \dots, y_t \sim Binomial\ Negatif(\mu_t)$, dimana persamaan μ_t memenuhi model seperti pada persamaan (4.31).

$$\ln(\mu_t) = \beta_0 + \sum_{j=1}^{\boldsymbol{p}} \phi_j \{\ln(y_{t-j}^*) - \beta_0) + \sum_{j=1}^{\boldsymbol{q}} \theta_j \left\{\ln \frac{(y_{t-j}^*)}{\mu_{t-j}}\right\}$$

$$\mu_t = \exp(\beta_0 + \sum_{j=1}^{\boldsymbol{p}} \phi_j \{\ln(y_{t-j}^*) - \beta_0) + \sum_{j=1}^{\boldsymbol{q}} \theta_j \left\{\ln \frac{(y_{t-j}^*)}{\mu_{t-j}}\right\}. \quad (4.31)$$

Langkah selanjutnya yaitu membentuk fungsi *partial likelihood* dari distribusi Binomial Negatif seperti yang ditunjukkan pada persamaan (4.32)

$$PL(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \prod_{t=1}^{n} f(y_t; \beta_0, \phi_j, \theta_j | H_{t-1})$$

$$PL(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \prod_{t=1}^{N} \left(\frac{1}{1+\alpha\mu_t}\right)^{1/\alpha} \left(\frac{\alpha\mu_t}{\mu_t\alpha+1}\right)^{y_t} \left(\frac{\Gamma(\frac{1}{\alpha}+y_t}{y_t!\,\Gamma(\frac{1}{\alpha})}\right). \qquad (4.32)$$

Fungsi *partial likelihood* pada persamaan (4.32) jika diubah menjadi fungsi *partial log-likelihood* akan berbentuk seperti pada persamaan (4.33)

$$l(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \ln PL(y_1, y_2, \dots y_t; \beta_0, \phi_j, \theta_j | H_{t-1})$$

$$l(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \ln \left[\prod_{t=1}^{n} \left(\frac{1}{1+\alpha\mu_t}\right)^{1/\alpha} \left(\frac{\alpha\mu_t}{\mu_t\alpha+1}\right)^{y_t} \left(\frac{\Gamma(\frac{1}{\alpha}+y_t}{y_t!\,\Gamma(\frac{1}{\alpha})}\right)\right]$$

$$l(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \ln \prod_{t=1}^{N} \Gamma\left(\frac{1}{\alpha}+y_t\right) - \ln \prod_{t=1}^{N} y_t!\,\Gamma\left(\frac{1}{\alpha}\right) + \qquad (4.33)$$

$$\ln \prod_{t=1}^{n} \left(\frac{1}{1+\alpha\mu_t}\right)^{\frac{1}{\alpha}} + \ln\left(\frac{\alpha\mu_t}{\mu_t\alpha+1}\right)^{y_t}$$

$$l(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = \sum_{t=1}^{N} \ln\Gamma\left(\frac{1}{\alpha}+y_t\right) - \sum_{t=1}^{N} \ln(y_t!)\Gamma\left(\frac{1}{\alpha}\right) - \frac{1}{\alpha}\sum_{t=1}^{N} \ln(1+\alpha\mu_t) \qquad (4.34)$$

$$+\sum_{t=1}^{N} y_t \ln \alpha\mu_t - \sum_{t=1}^{N} y_t \ln(1+\alpha\mu_t).$$

Jika persamaan (4.1) disubtitusikan ke dalam fungsi *partial log-likelihood* pada persamaan (4.34) maka bentuk *partial log-likelihood* dapat ditulis seperti pada persamaan (4.35)

$$l(y_t; \beta_0, \phi_j, \theta_j | H_{t-1}) = -\frac{1}{\alpha}\sum_{t=1}^{n} \ln(\exp \beta_0 + \phi_1 \{\ln y_{t-1}^*\}) + \qquad (4.35)$$

$$\sum_{t=1}^{N} y_t \ln \alpha(\exp \beta_0 + \phi_1\{\ln y_{t-1}^*\}) - \sum_{t=1}^{N} y_t \ln(1 + \alpha \exp \beta_0 + \phi_1\{\ln y_{T-1}^*\})).$$

Langkah yang dilakukan selanjutnya adalah mendapatkan nilai parameter $\beta_0, \phi_j, \theta_j$ dengan cara memaksimumkan fungsi *partial log-likelihood* yang sudah didapat sebelumnya pada persamaan (4.35). Untuk memperoleh nilai parameter dapat dilakukan dengan cara mencari turunan parsial pertama terhadap masing – masing parameter $\beta_0, \phi_j, \theta_j$. Turunan pertama dari *partial log-likelihood* terhadap parameter β_0 ditunjukkan pada persamaan (4.36) seperti yang sudah tertera sebagai berikut:

$$\frac{\partial l(\mu_t)}{\partial \beta_0} = \frac{\partial l}{\partial \beta_0}\left[-\frac{1}{\alpha}\sum_{t=1}^{n} \ln(\exp(\beta_0 + \phi_1\{\ln y_{t-1}^*\})) + \sum_{t=1}^{N} y_t \ln \alpha(\exp \beta_0 + \phi_1\{\ln y_{t-1}^*\})) - \sum_{t=1}^{N} y_t \ln(1 + \alpha \exp \beta_0 + \phi_1\{\ln y_{T-1}^*\}))\right]$$

$$\frac{\partial l(\mu_t)}{\partial \beta_0} = -\frac{1}{\alpha}\sum_{t=1}^{N} \frac{\alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}{1 + \alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})} + \sum_{t=1}^{N} y_t - \sum_{t=1}^{N} y_t \frac{\alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}{1 + \alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}. \tag{4.36}$$

Menggunakan cara yang sama seperti turunan *partial log-likelihood* terhadap paremeter β_0, diperoleh turunan pertama dari *partial log-likelihood* terhadap parameter ϕ seperti yang dapat ditulis pada persamaan (4.37)

$$\frac{\partial l(\mu_t)}{\partial \phi_j} = -\frac{1}{\alpha}\sum_{t=1}^{N} \frac{\alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}{1 + \alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})} + \sum_{t=1}^{N} y_t(\ln y_{T-1}^*) - \sum_{t=1}^{N} y_t \frac{\alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}{1 + \alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^*\})}. \tag{4.37}$$

Selanjutnya untuk turunan pertama dari *partial log-likeliood* terhadap parameter θ dapat ditulis seperti pada persamaan (4.38)

$$\frac{\partial l(\mu_t)}{\partial \theta_j} = -\frac{1}{\alpha}\sum_{t=1}^{N} \frac{\alpha \exp(\beta_0 + \theta_1\{\ln y_{T-1}^* - \ln \mu_{T-1}\})}{1 + \alpha \exp(\beta_0 + \theta_1\{\ln y_{T-1}^* - \ln \mu_{T-1}\})} + \sum_{t=1}^{N} y_t(\ln y_{T-1}^* - \ln \mu_{T-1}) - \sum_{t=1}^{N} y_t \frac{\alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^* - \ln \mu_{T-1}\})}{1 + \alpha \exp(\beta_0 + \phi_1\{\ln y_{T-1}^* - \ln \mu_{T-1}\})} \tag{4.38}$$

Agar fungsi dari *partial log-likelihood* dapat maksimum maka turunan pertama dari masing masing parameter disamadengankan nol seperti yang tertera pada persamaan (4.39)

$$\frac{\partial l(\mu_t)}{\partial \beta_0} = 0\,; \frac{\partial l(\mu_t)}{\partial \phi_j} = 0\,; \frac{\partial l(\mu_t)}{\partial \theta_j} = 0. \tag{4.39}$$

Dalam persamaan (4.36), (4.37) dan (4.38) terdapat suku $\exp\left(\beta_0 + \sum_{j=1}^{p} \phi_j\{\ln y_{t-j}^*\} + \sum_{j=1}^{q} \theta_j \left\{\frac{\ln y_{t-j}^*}{\mu_{t-j}}\right\}\right)$ yang mengakibatkan bentuk *closed form* dari β_0, ϕ_j, dan θ_j sulit untuk ditemukan sehingga jika menggunakan metode *Maximum Likelihood Estimation* nilai parameter tidak dapat diselesaikan. Oleh karena itu, untuk mengestimasi parameter $\beta_0, \phi_j, \theta_j$ digunakan optimasi dengan menggunakan algoritma *Iteratively Reweighted Least Square* (IRLS) dengan langkah – langkah melakukan iterasi IRLS dapat dilihat pada sub bab 2.12. Iterasi dilakukan hingga memperoleh parameter yang konvergen atau hingga diperoleh hasil $|\hat{\gamma}^{(t)} - \hat{\gamma}^{(t-1)}| < \varepsilon$.

4.5.3 Hasil Estimasi dan Uji Sigifikansi Parameter

Pada sub bab sebelumnya telah dibahas bahwa bentuk dari persamaan *partial log-likelihood* pada model Poisson GARMA dan Binomial Negatif GARMA tidak *closed form*. Hal ini mengakibatkan estimasi parameter dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) tidak dapat dilakukan, dan digunakan optimasi algoritma *Iteratively Reweighted Least Square* (IRLS) dalam menyelesaikan estimasi parameter tersebut. Penaksiran parameter dengan menggunakan optimasi algoritma IRLS dan dengan bantuan *software* MATLAB2013 diperoleh hasil seperti yang dapat dilihat pada Tabel 4.3.

Tabel 4.3 Estimasi Parameter Model Poisson GARMA dan Binomial Negatif GARMA

Model	**Parameter**	**SE**
Poisson GARMA (0,1)	$\beta_0 = 1{,}214$	0,074
	$\theta_1 = 0{,}064$	0,061
Poisson GARMA (1,0)	$\beta_0 = -1{,}219$	0,075
	$\phi_1 = 0{,}074$	0,056
Poisson GARMA (1,1)	$\beta_0 = 1{,}572$	0,189
	$\phi_1 = 0{,}914$	0,054
	$\theta_1 = -0{,}187$	0,077
Binomial Negatif GARMA (0,1)	$\beta_0 = 1{,}208$	0,082
	$\phi_1 = 0{,}068$	0,067
Binomial Negatif GARMA (1,0)	$\beta_0 = 1{,}217$	0,087
	$\phi_1 = 0{,}095$	0,063
Binomial Negatif GARMA (1,1)	$\beta_0 = 1{,}497$	0,149
	$\phi_1 = 0{,}899$	0,143
	$\theta_1 = -0{,}106$	0,048

Setelah memperoleh nilai parameter $\beta_0, \phi_1, \theta_1$ dari masing – masing model Poisson GARMA (p, q) dan Binomial Negatif GARMA (p, q), selanjutnya akan dilakukan uji signifikansi yang bertujuan untuk mengetahui kelayakan dari setiap parameter yang diperolah agar menghasilkan peramalan yang optimal atau akurat seperti yang sudah dijelaskan dalam sub bab sebelumnya.

1. Uji Signifikansi Model Poisson GARMA $(\mathbf{0}, \mathbf{1})$

Uji parameter $\boldsymbol{\theta_1}$

Hipotesa :

$H_0: \theta_1 = 0$ (parameter θ_1 tidak signifikan)

$H_1: \theta_1 \neq 0$ (parameter θ_1 signifikan)

Statistik Uji :

$$t_{hitung} = \frac{\hat{\theta}_1}{SE(\theta_1)} = \frac{0{,}064}{0{,}061} = 1{,}0492$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0{,}025;60} = 2{,}000298$$

Kriteria pengujian:

Karena $|t_{hitung}| < t_{tabel}$, maka H_0 gagal ditolak yang artinya parameter model θ_1 tidak signifikan.

2. Uji Signifikansi Model Poisson GARMA $(\mathbf{1}, \mathbf{0})$

Uji parameter $\boldsymbol{\phi_1}$

Hipotesa :

$H_0: \phi_1 = 0$ (parameter ϕ_1 tidak signifikan)

$H_1: \phi_1 \neq 0$ (parameter ϕ_1 signifikan)

Statistik Uji :

$$t_{hitung} = \frac{\hat{\phi}_1}{SE(\phi_1)} = \frac{0,074}{0,056} = 1,32143$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0,025;60} = 2,000298$$

Kriteria pengujian:

Karena $|t_{hitung}| < t_{tabel}$, maka H_0 gagal ditolak yang artinya parameter model ϕ_1 tidak signifikan.

3. Uji Signifikansi Model Poisson GARMA $(\mathbf{1}, \mathbf{1})$

Uji parameter $\boldsymbol{\theta_1}$

Hipotesa :

$H_0: \theta_1 = 0$ (parameter θ_1 tidak signifikan)

$H_1: \theta_1 \neq 0$ (parameter θ_1 signifikan)

Statistik Uji :

$$t_{hitung} = \frac{\hat{\theta}_1}{SE(\theta_1)} = \frac{-0,187}{0,077} = 2,429$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0,025;60} = 2,000298$$

Kriteria pengujian:

Karena $|t_{hitung}| > t_{tabel}$, maka H_0 ditolak yang artinya parameter model θ_1 signifikan.

Uji parameter $\boldsymbol{\phi_1}$

Hipotesa :

$H_0: \phi_1 = 0$ (parameter ϕ_1 tidak signifikan)

$H_1: \phi_1 \neq 0$ (parameter ϕ_1 signifikan)

Statistik Uji :

$$t_{hitung} = \frac{\hat{\phi}_1}{SE(\phi_1)} = \frac{0,914}{0,054} = 16,925$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0,025;60} = 2,000298$$

Kriteria pengujian:

Karena $|t_{hitung}| > t_{tabel}$, maka H_0 ditolak yang artinya parameter model ϕ_1 signifikan.

4. **Uji Signifikansi Model Binomial Negatif GARMA $(\mathbf{0},\mathbf{1})$**
Uji parameter $\boldsymbol{\theta_1}$
Hipotesa :
$H_0: \theta_1 = 0$ (parameter θ_1 tidak signifikan)
$H_1: \theta_1 \neq 0$ (parameter θ_1 signifikan)
Statistik Uji :

$$t_{hitung} = \frac{\hat{\theta}_1}{SE(\theta_1)} = \frac{0{,}068}{0{,}067} = 1{,}015$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0{,}025;60} = 2{,}000298$$

Kriteria pengujian:
Karena $|t_{hitung}| < t_{tabel}$, maka H_0 gagal ditolak yang artinya parameter model θ_1 tidak signifikan.

5. **Uji Signifikansi Model Binomial Negatif GARMA $(\mathbf{1},\mathbf{0})$**
Uji parameter $\boldsymbol{\phi_1}$
Hipotesa :
$H_0: \phi_1 = 0$ (parameter ϕ_1 tidak signifikan)
$H_1: \phi_1 \neq 0$ (parameter ϕ_1 signifikan)
Statistik Uji :

$$t_{hitung} = \frac{\hat{\phi}_1}{SE(\phi_1)} = \frac{0{,}095}{0{,}063} = 1{,}508$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0{,}025;60} = 2{,}000298$$

Kriteria pengujian:
Karena $|t_{hitung}| < t_{tabel}$, maka H_0 gagal ditolak yang artinya parameter model ϕ_1 tidak signifikan.

6. **Uji Signifikansi Model Binomial Negatif GARMA $(\mathbf{1},\mathbf{1})$**
Uji parameter $\boldsymbol{\theta_1}$
Hipotesa :
$H_0: \theta_1 = 0$ (parameter θ_1 tidak signifikan)
$H_1: \theta_1 \neq 0$ (parameter θ_1 signifikan)
Statistik Uji :

$$t_{hitung} = \frac{\hat{\theta}_1}{SE(\theta_1)} = \frac{-0{,}106}{0{,}048} = 2{,}2083$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0{,}025;60} = 2{,}000298$$

Kriteria pengujian:
Karena $|t_{hitung}| > t_{tabel}$, maka H_0 ditolak yang artinya parameter model θ_1 signifikan.

Uji parameter ϕ_1

Hipotesa :

$H_0: \phi_1 = 0$ (parameter ϕ_1 tidak signifikan)

$H_1: \phi_1 \neq 0$ (parameter ϕ_1 signifikan)

Statistik Uji :

$$t_{hitung} = \frac{\hat{\phi}_1}{SE(\phi_1)} = \frac{0{,}899}{0{,}143} = 6{,}287$$

Dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{0{,}025;60} = 2{,}000298$$

Kriteria pengujian:

Karena $|t_{hitung}| > t_{tabel}$, maka H_0 ditolak yang artinya parameter model ϕ_1 signifikan.

Dari Uji Signifikansi yang telah dilakukan diperoleh hasil signifikansi parameter yang dicari seperti yang tertera pada tabel 4.3

Tabel 4.4 Hasil Uji Signifikansi Parameter Model Poisson GARMA dan Binomial Negatif GARMA

Model	Parameter	SE	Uji Signifikan
Poisson GARMA (0,1)	$\beta_0 = 1{,}214$	0,074	Signifikan
	$\theta_1 = 0{,}064$	0,061	Tidak Signifikan
Poisson GARMA (1,0)	$\beta_0 = -1{,}219$	0,075	Signifikan
	$\phi_1 = 0{,}074$	0,056	Tidak Signifikan
Poisson GARMA (1,1)	$\beta_0 = 1{,}572$	0,189	Signifikan
	$\phi_1 = 0{,}914$	0,054	Signifikan
	$\theta_1 = -0{,}187$	0,077	Signifikan
Binomial Negatif GARMA (0,1)	$\beta_0 = 1{,}208$	0,082	Signifikan
	$\phi_1 = 0{,}068$	0,067	Tidak Signifikan
Binomial Negatif GARMA (1,0)	$\beta_0 = 1{,}217$	0,087	Signifikan
	$\phi_1 = 0{,}095$	0,063	Tidak Signifikan
Binomial Negatif GARMA (1,1)	$\beta_0 = 1{,}497$	0,149	Signifikan
	$\phi_1 = 0{,}899$	0,143	Signifikan
	$\theta_1 = -0{,}106$	0,048	Signifikan

Dari Tabel 4.4 dapat dilihat bahwa model yang layak digunakan ke dalam tahap peramalan yaitu model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1) dengan parameter $\beta_0, \phi_1, \theta_1$ telah memenuhi uji signifikan.

4.6 Peramalan Model Poisson GARMA dan Binomial Negatif GARMA

Setelah diperoleh parameter $\beta_0, \phi_1, \theta_1$ untuk masing – masing model Poisson GARMA dan Binomial Negatif dan sudah melewati uji signifikansi, maka tahap selanjutnya adalah melakukan peramalan Jumlah Kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong dengan menggunakan model yang sudah terpilih yaitu model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1). Peramalan model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1) dilakukan dengan bantuan *Microsoft Excel*, dengan pembagian data yang digunakan yaitu 60 data sebagai data *in sample* dan 12 data sebagai data *out sample*.

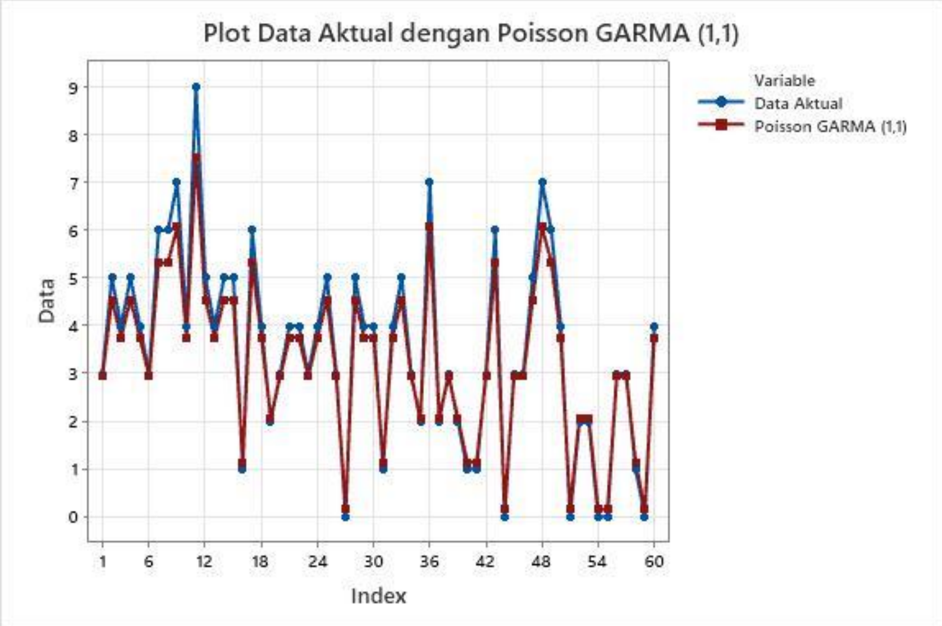
Data yang digunakan pada penelitian ini adalah data jumlah kecelakaan yang termasuk dalam data *count* yang nilainya merupakan bilangan diskrit non negatif 0,1,2,3, ..., maka untuk hasil peramalan ini data yang dihasilkan juga harus mengikuti sifat dari data *count*. Jika hasil peramalan yang dihasilkan merupakan bilangan pecahan maka dapat dilakukan sebuah pembulatan. Pembulatan merupakan menyajikan bentuk bilangan ke dalam digit yang lebih sedikit. Terdapat tiga aturan pembulatan dalam statistika, yaitu jika angka terkiri dari angka yang dihilangkan kurang dari lima, maka angka terkanan dari angka yang mendahuluinya tetap atau tidak berubah. Aturan yang kedua adalah jika angka terkiri dari angka yang dihilangkan lebih dari atau sama dengan lima, maka angka terkanan dari angka yang mendahuluinya bertambah dengan satu. Aturan ketiga adalah jika angka terkiri dari angka yang dihilangkan sama dengan lima, maka angka terkanan dari angka yang mendahuluinya tetap jika angka tersebut genap dan bertambah satu jika angka tersebut ganjil.

4.6.1 Peramalan Model Poisson GARMA $(\mathbf{1},\mathbf{1})$

Pada tahap ini dilakukan peramalan Jumlah Kecelakaan di Jalan Tol dengan menggunakan Model Poisson GARMA (1,1) yang sudah dicari terlebih dahulu. Sebelumnya sudah dilakukan estimasi parameter untuk model Poisson GARMA (1,1) dengan hasil estimasi yang diperoleh adalah nilai $\beta_0 = 1{,}572$; $\phi_1 = 0{,}914$; $\theta_1 = -0{,}187$. Jika parameter yang sudah diperoleh dimasukkan ke dalam persamaan (4.1) maka diperoleh persamaan model Poisson GARMA (1,1) seperti yang tertera pada persamaan (4.37) sebagai berikut :

$$\mu_t = \exp\left[1{,}572 + \sum_{j=1}^{1} 0{,}914\left\{\ln y_{t-j}^{*} - 1{,}572\right\} - \sum_{j=1}^{q} 0{,}187\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right] \tag{4.40}$$

Persamaan (4.40) merupakan persamaan model Poisson GARMA (1,1) yang nantinya persamaan tersebut digunakan untuk menghitung hasil peramalan data *in sample* dan data *out sample* data jumlah kecelakaan pada bulan Januari 2016 – Desember 2021 terlebih dahulu. Setelah memperoleh hasil peramalan data *in sample* dan *out sample*, maka dibandingkan dengan data aktual yang sudah diperoleh sebelumnya dari PT. Jasa Marga. Kemudian perbandingan antara data aktual dengan data *in sample* dan *out sample* di *plot* agar mengetahui pola data dan melihat tingkat akurasi dengan menggunakan *Root Mean Square Error* (RMSE) yang dihasilkan dari perhitungan tersebut. Setelah menghitung tingkat akurasi maka dilanjutkan untuk melakukan peramalan selama satu tahun ke depan yaitu untuk bulan Januari 2022 hingga Desember 2022. Dengan menggunakan bantuan *software* Microsoft Excel dan *Minitab* maka diperoleh grafik peramalan data *in sample* seperti yang tertera pada Gambar 4.6 sebagai berikut :



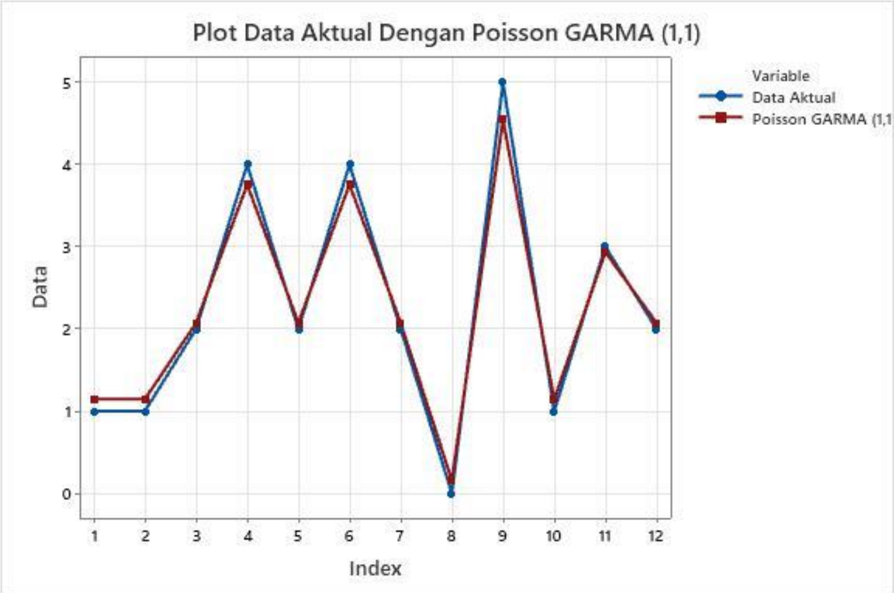
Gambar 4.6 Grafik Model Poisson GARMA (1,1) Data *In Sample*

Pada Gambar 4.6 terlihat bahwa hasil peramalan dengan menggunakan model pada persamaan (4.37) mengikuti pola data aktual dengan perbedaan yang tidak terlalu menonjol. Oleh karena itu, dilanjutkan untuk melakukan peramalan dalam satu tahun ke depan dilihat dari *plot* grafik perbandingan perhitungan data aktual dengan hasil peramalan data *out sample*. Dari peramalan data *out sample* dihitung tingkat akurasi peramalan dengan menggunakan *Root Mean Square Error* (RMSE) agar dapat diketahui model terbaik untuk melakukan peramalan Jumlah Kecelakaan di Jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong. *Plot* grafik perbandingan data aktual dengan data *out sample* data dapat dilihat pada Tabel 4.6 seperti yang sudah tertera sebagai berikut:

Tabel 4. 5 Hasil Perbandingan Data Aktual dengan Hasil Peramalan Out Sample

Data Aktual	**Hasil Peramalan Data *Out Sample***	**Pembulatan**
1	1.144756557	1
1	1.144756557	1
2	2.076484848	2
4	3.767334226	4
2	2.076193776	2
4	3.766276137	4
2	2.075918156	2
0	0.158783325	0
5	4.556666136	5
1	1.144756557	1
3	2.937147032	3
2	2.073974277	2
RMSE	**0.189015023**	
MAPE	**11.57**	

Setelah memperoleh perbandingan data aktual dengan hasil peramalan data *out sample* serta sudah melakukan pembulatan terhadap data ramalan yang dihasilkan, maka dapat di *plot* sebuah grafik dari Tabel 4.6 yang dapat dilihat pada Gambar 4.7 seperti yang sudah tertera sebagai berikut:



Gambar 4.7 Grafik Model Poisson GARMA (1,1) Data *Out Sample*

Pada Gambar 4.7 dapat dilihat bahwa hasil peramalan data *out sample* sudah mengikuti data aktual dengan perbedaan yang tidak jauh berbeda, dan pada tabel 4.6 diperoleh nilai RMSE sebesar 0.18546485. Dengan mengikuti pola peramalan seperti data *out sample* dan dengan menggunakan persamaan model (4.37), maka akan diperoleh peramalan selama satu tahun ke depan yaitu mulai dari bulan Januari 2022 – Desember 2022 seperti yang tertera pada Tabel 4.7 sebagai berikut:

Tabel 4.6 Peramalan Jumlah Kecelakaan Model Poisson GARMA $(1,1)$

Periode (Bulan)	Hasil Peramalan	Pembulatan
Januari 2022	2.073974277	2
Februari 2022	2.139118866	2
Maret 2022	2.196127943	2
April 2022	2.245735966	2
Mei 2022	2.288681785	2
Juni 2022	2.32568316	2
Juli 2022	2.357419736	2
Agustus 2022	2.38452265	2
September 2022	2.407569134	2
Oktober 2022	2.427080734	2
November 2022	2.443524077	2
Desember 2022	2.457313298	2

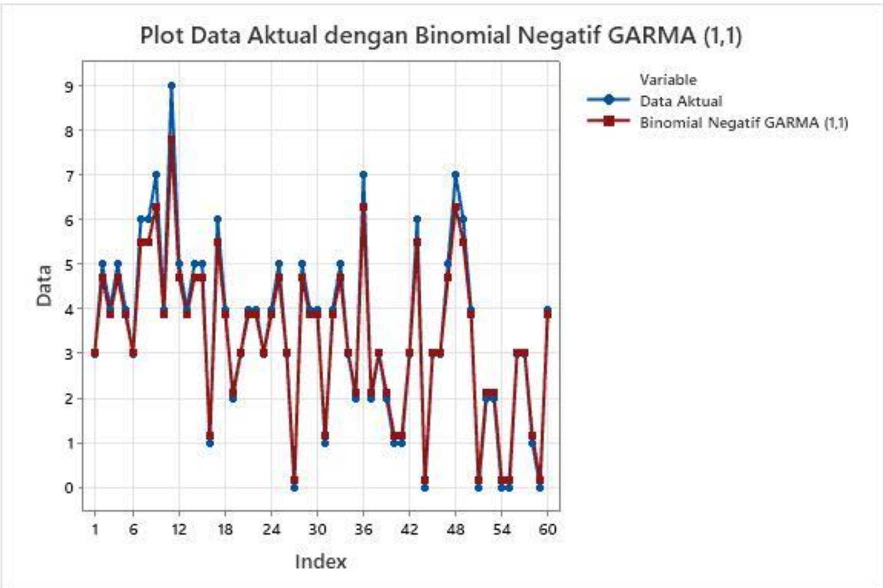
4.6.2 Peramalan Model Binomial Negatif GARMA $(1,1)$

Pada tahap ini dilakukan peramalan Jumlah Kecelakaan dengan menggunakan Model Binomial Negatif GARMA $(1,1)$. Sebelumnya sudah dilakukan estimasi parameter untuk model Binomial Negatif GARMA $(1,1)$ dengan hasil estimasi yang diperoleh nilai $\beta_0 =$

1,497 ; $\phi_1 = 0{,}899$; $\theta_1 = -0{,}106$. Jika parameter yang sudah diperoleh dimasukkan ke dalam persamaan (4.12) akan diperoleh persamaan model Binomial Negatif GARMA (1,1) sebagai berikut :

$$\mu_t = \exp\left[1{,}497 + \sum_{j=1}^{1} 0{,}899\left\{\ln y_{t-j}^{*} - 1{,}497\right\} - \sum_{j=1}^{1} 0{,}106\left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right] \tag{4.41}$$

Persamaan (4.41) merupakan persamaan model Binomial Negatif GARMA (1,1) yang nantinya persamaan tersebut digunakan untuk menghitung hasil peramalan data *in sample* dan data *out sample* data jumlah kecelakaan pada bulan Januari 2016 – Desember 2021 terlebih dahulu. Setelah memperoleh hasil peramalan data *in sample* dan *out sample*, maka dibandingkan dengan data aktual yang sudah diperoleh sebelumnya dari PT. Jasa Marga. Kemudian perbandingan antara data aktual dengan data *in sample* dan *out sample* di *plot* agar mengetahui pola data dan melihat tingkat akurasi dengan menggunakan *Root Mean Square Error* (RMSE) yang dihasilkan dari perhitungan tersebut. Setelah menghitung tingkat akurasi maka dilanjutkan untuk melakukan peramalan selama satu tahun ke depan yaitu untuk bulan Januari 2022 hingga Desember 2022. Dengan menggunakan bantuan *software* Microsoft Excel dan *Minitab* maka diperoleh grafik peramalan data *in sample* seperti yang tertera pada Gambar 4.8 sebagai berikut :



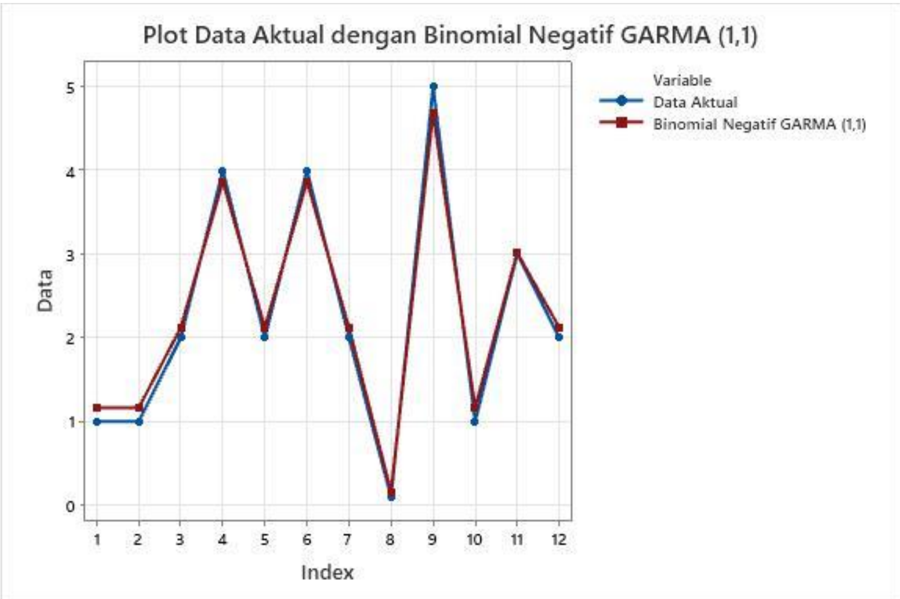
Gambar 4.8 Grafik Model Binomial Negatif GARMA (1,1) Data *In Sample*

Pada Gambar 4.8 terlihat bahwa hasil peramalan dengan menggunakan model pada persamaan (4.38) mengikuti pola data aktual dengan perbedaan yang tidak terlalu menonjol. Oleh karena itu, untuk melakukan peramalan dalam satu tahun ke depan dilihat dari *plot* grafik perbandingan perhitungan data aktual dengan hasil peramalan data *out sample*. Dari peramalan data *out sample* dihitung tingkat akurasi peramalan dengan menggunakan *Root Mean Square Error* (RMSE) agar dapat diketahui model terbaik untuk melakukan peramalan Jumlah Kecelakaan di Jalan Tol. *Plot* grafik perbandingan data aktual dan data *out sample* dapat dilihat pada Tabel 4.8 dan Gambar 4.9 sebagai berikut :

Tabel 4.7 Hasil Perbandingan Data Aktual dengan Hasil Peramalan *Out Sample*

Data Aktual	**Hasil Peramalan Data *Out Sample***	**Pembulatan**
1	1.163225791	1
1	1.163225791	1
2	2.122865256	2
4	3.874643184	4
2	2.122696573	2
4	3.87402629	4
2	2.122536835	2
0	0.157928077	0
5	4.69865863	5
1	1.163225791	1
3	3.015513997	3
2	2.121409984	2
RMSE	**0.151081028**	
MAPE	**10.08**	

Setelah memperoleh perbandingan data aktual dengan hasil peramalan data *out sample* serta sudah melakukan pembulatan terhadap data ramalan yang dihasilkan, maka dapat di *plot* sebuah grafik dari Tabel 4.8 yang dapat dilihat pada Gambar 4.9 seperti yang sudah tertera sebagai berikut:



Gambar 4.9 Grafik Model Binomial Negatif GARMA (1,1) Data Out Sample

Pada Gambar 4.9 dapat dilihat bahwa hasil peramalan data *out sample* sudah mengikuti data aktual dengan perbedaan yang tidak jauh berbeda, dan pada tabel 4.5 diperoleh nailai RMSE sebesar 0.150043187. Dengan mengikuti pola peramalan seperti data *out sample* dan dengan menggunakan persamaan model (4.38), maka dapat diperoleh peramalan selama satu tahun ke depan yaitu mulai dari bulan Januari 2022 – Desember 2022 sebagai berikut :

Tabel 4.8 Peramalan Jumlah Kecelakaan Tol Surabaya – Gempol Model Binomial Negatif GARMA (1,1)

Periode (Bulan)	Hasil Peramalan	Pembulatan
Januari 2022	2.121409984	2
Februari 2022	2.232342265	2
Maret 2022	2.332910163	2
April 2022	2.423463385	2
Mei 2022	2.504518323	3
Juni 2022	2.576699409	3
Juli 2022	2.640691802	3
Agustus 2022	2.697204713	3
September 2022	2.746944069	3
Oktober 2022	2.790593056	3
November 2022	2.828799036	3
Desember 2022	2.862165469	3

4.6.3 Perbandingan Hasil Peramalan

Setelah mendapatkan hasil peramalan *in* sample, *out* sample, dan peramalan dari bulan Januari 2022 – Desember 2022 dari masing – masing model, tahap selanjutnya adalah membandingkan hasil dari kedua model tersebut agar dapat menentukan model terbaik dalam melakukan peramalan Jumlah Kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong selama tahun 2022. Perbandingan Hasil *Out Sample* model Poisson GARMA (1,1) dengan model Binomial Negatif GARMA (1,1) dapat dilihat pada Tabel 4.9 dan Perbandingan Hasil Peramalan Model Poisson GARMA (1,1) dengan model Binomial Negatif GARMA (1,1) dapat dilihat pada Tabel 4.10 sebagai berikut :

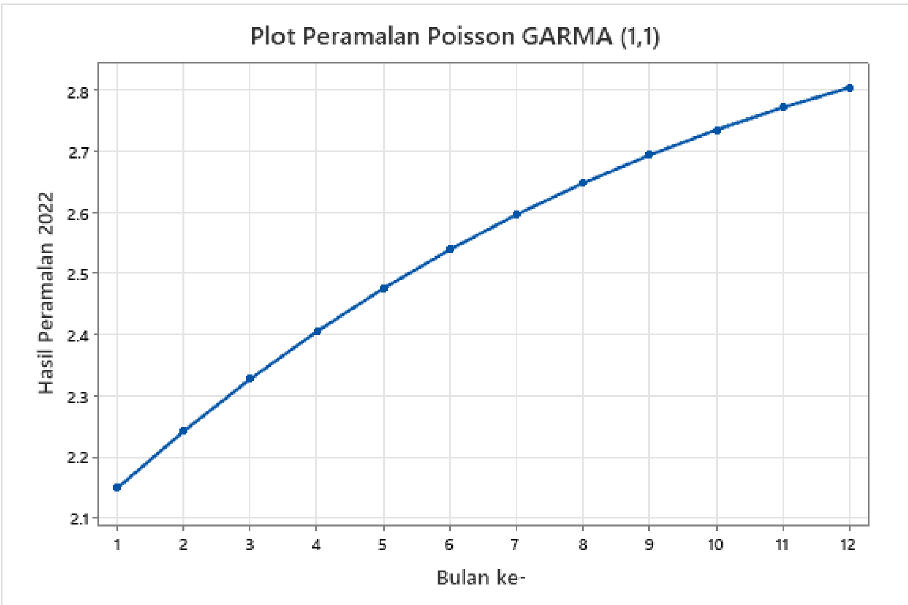
Tabel 4. 9 Perbandingan Hasil Out Sample Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1)

t	Data Aktual	Poisson GARMA $(1,1)$	Binomial Negatif GARMA $(1,1)$
61	1	1.144756557	1.163225791
62	1	1.144756557	1.163225791
63	2	2.076484848	2.122865256
64	4	3.767334226	3.874643184
65	2	2.076193776	2.122696573
66	4	3.766276137	3.87402629
67	2	2.075918156	2.122536835
68	0	0.158783325	0.157928077
69	5	4.556666136	4.69865863
70	1	1.144756557	1.163225791
71	3	2.937147032	3.015513997
72	2	2.073974277	2.121409984

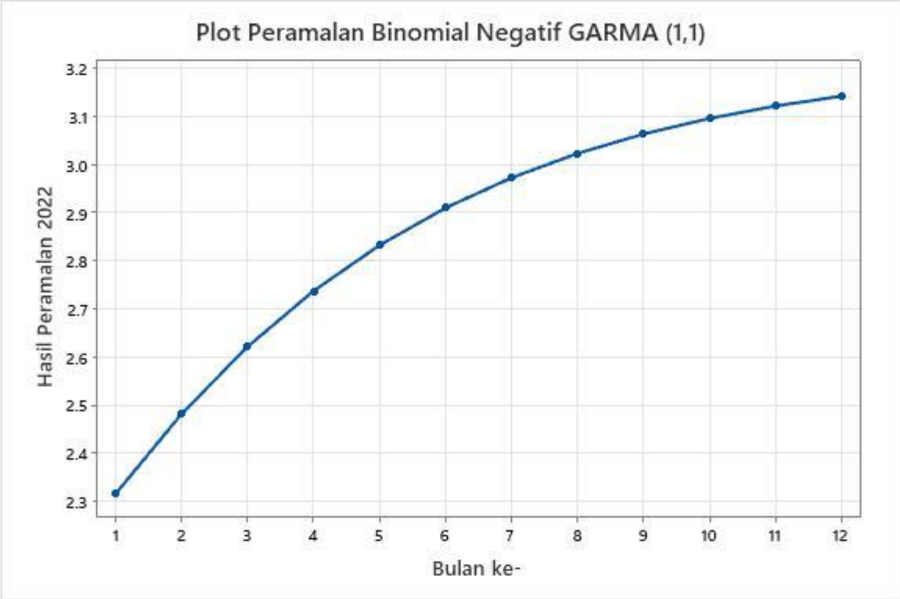
Tabel 4.10 Perbandingan Hasil Peramalan Jumlah Kecelakaan Model Poisson GARMA (1,1) dan Model Binomial Negatif GARMA (1,1)

Periode (Bulan)	Hasil Peramalan Poisson GARMA (1,1)	Hasil Peramalan Binomial Negatif GARMA (1,1)
Januari 2022	2.073974277	2.121409984
Februari 2022	2.139118866	2.232342265
Maret 2022	2.196127943	2.332910163
April 2022	2.245735966	2.423463385
Mei 2022	2.288681785	2.504518323
Juni 2022	2.32568316	2.576699409
Juli 2022	2.357419736	2.640691802
Agustus 2022	2.38452265	2.697204713
September 2022	2.407569134	2.746944069
Oktober 2022	2.427080734	2.790593056
November 2022	2.443524077	2.828799036
Desember 2022	2.457313298	2.862165469
RMSE	**0.189015023**	**0.151081028**
MAPE	**11.57**	**10.08**

Pada hasil peramalan yang sudah dilakukan dapat dilihat bahwa hasilnya berada pada sekitaran nilai rata – rata data. Dilihat dari Tabel 4.10 menunjukkan bahwa peramalan menggunakan model Poisson GARMA (1,1) menghasilkan RMSE sebesar 0.189015023 dan MAPE sebesar 11.57, sedangkan peramalan menggunakan model Binomial Negatif GARMA (1,1) menghasilkan RMSE sebesar 0.151081028 dan MAPE sebesar 10.08. Dari tabel diatas dapat disimpulkan bahwa nilai RMSE dan MAPE model Binomial Negatif GARMA (1,1) lebih kecil dibandingkan nilai RMSE dan MAPE model Poisson GARMA (1,1), yang mengindikasikan bahwa peramalan Jumlah Kecelakaan dengan menggunakan model Binomial Negatif GARMA (1,1) memiliki hasil yang lebih akurat. *Plot* Grafik Peramalan dengan model Poisson GARMA dan Binomial Negatif GARMA dapat dilihat pada Gambar 4.10 dan 4.11 sebagai berikut :



Gambar 4.10 Grafik Peramalan Jumlah Kecelakaan Model Poisson GARMA (1,1)



Gambar 4.11 Grafik Peramalan Jumlah Kecelakaan Model BN GARMA (1,1)

Dari hasil yang sudah diperoleh dari sub bab sebelumnya, selanjutnya dapat dilakukan analisis hasil peramalan Data Jumlah Kecelakaan di Jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong. Dari *plot time series* sebelumnya terlihat bahwa data kejadian kecelakaan ini bukan merupakan data *seasonal* atau musiman, dengan data tidak stasioner terhadap *mean* dan belum stasioner terhadap nilai variansnya. Model yang digunakan dalam penelitian ini adalah model GARMA yang dikembangkan oleh Benjamin dkk, model ini tidak melibatkan efek stasioneritas dan musiman sehingga tidak perlu dilakukan uji stasioner terhadap data yang digunakan.

Model GARMA merupakan pengembangan dari model *Generalized Linear Models* (GLM) dengan melibatkan unsur ARMA di dalamnya. Pada model ini juga tidak perlu dilakukan uji diagnostik yaitu model yang terbentuk harus memenuhi asumsi *white noise* dan berdistribusi normal pada residualnya, karena variable respon dari GLM tidak harus mengikuti distribusi normal namun termasuk ke dalam distribusi keluarga eksponensial. Dalam penelitian ini model yang digunakan ada dua yakni model Poisson GARMA dan Binomial Negatif GARMA. Pada tahap identifikasi model telah diperoleh dugaan model sementara yang akan digunakan dalam melakukan peramalan, yaitu model Poisson GARMA (1,0), Poisson GARMA (0,1), Poisson GARMA (1,1), Binomial Negatif GARMA (1,0), Binomial Negatif GARMA (0,1), dan Binomial Negatif GARMA (1,1). Dari keenam model yang diperoleh, selanjutnya dilakukan estimasi parameter dengan menggunakan *Maximum Likelihood Estimation* (MLE), namun diperoleh hasil yang tidak *closed form* sehingga estimasi parameter dilanjutkan dengan menggunakan algoritma *Iteratively Reweighted Least Square* (IRLS). Selanjutnya hasil parameter yang sudah diperoleh tersebut dilanjutkan dengan melakukan uji signifikansi dan diperoleh model yang signifikan adalah model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1).

Kedua model yang sudah signifikan tersebut dilakukan untuk melakukan peramalan pada data *out sample* kemudian dihitung tingkat akurasinya dengan menggunakan *Root Mean Square Error* (RMSE). Dari grafik hasil peramalan *out sample* dengan data aktual terlihat bahwa hasil peramalan telah mengikuti pola data aktual yang didukung dengan nilai RMSE dan MAPE yang sudah kecil, dimana nilai RMSE dan MAPE terkecil terdapat pada model Binomial Negatif GARMA (1,1) dengan nilai RMSE sebesar 0.151081028 dan MAPE sebesar 10.08. Selanjutnya dilakukan analisis terhadap hasil tersebut. Hasil Peramalan Jumlah Kecelakaan di Jalan Tol Gempol – Surabaya ruas Banyu Urip – Porong dengan menggunakan model Poisson

GARMA (1,1) dan model Binomial Negatif GARMA (1,1) memiliki *plot* grafik yang meningkat setiap bulannya. Pada bulan Januari 2022 hasil peramalan jumlah kecelakaan dengan menggunakan model Poisson GARMA (1,1) adalah sebesar 2.073974277 atau jika dilakukan pembulatan maka banyak kejadian kecelakaan sebesar 2. Sama halnya dengan menggunakan model Binomial Negatif GARMA (1,1), pada bulan Januari 2022 hasil peramalan jumlah kecelakaan adalah sebesar 2.121409984 atau jika dibulatkan ke angka terdekat adalah 2 kejadian kecelakaan. Dari hasil peramalan dengan menggunakan model Poisson GARMA (1,1) maupun Binomial Negatif GARMA (1,1), diperkirakan jumlah kecelakaan terbanyak sepanjang tahun 2022 dari mulai bulan Januari sampai bulan Desember adalah sebanyak 3 kejadian kecelakaan.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang sudah dilakukan pada Bab IV, maka dapat ditarik beberapa kesimpulan antara lain sebagai berikut :

1. Estimasi parameter untuk model Poisson GARMA (p, q) dan model Binomial Negatif GARMA (p, q) dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Setelah melakukan turunan pertama pada masing – masing parameter terlihat bahwa bentuk persamaan tidak *closed form* sehingga untuk melakukan estimasi parameter harus menggunakan optimasi algoritma *Iteratively Reweighted Least Square* (IRLS). Hasil dari estimasi parameter dengan orde p dan q tertentu diperoleh sebagai berikut:

Model	Parameter
Poisson GARMA (0,1)	$\beta_0 = 1{,}214$
	$\theta_1 = 0{,}064$
Poisson GARMA (1,0)	$\beta_0 = -1{,}219$
	$\phi_1 = 0{,}074$
Poisson GARMA (1,1)	$\beta_0 = 1{,}572$
	$\phi_1 = 0{,}914$
	$\theta_1 = -0{,}187$
Binomial Negatif GARMA (0,1)	$\beta_0 = 1{,}208$
	$\phi_1 = 0{,}068$
Binomial Negatif GARMA (1,0)	$\beta_0 = 1{,}217$
	$\phi_1 = 0{,}095$
Binomial Negatif GARMA (1,1)	$\beta_0 = 1{,}497$
	$\phi_1 = 0{,}899$
	$\theta_1 = -0{,}106$

2. Hasil dari mengaplikasikan model Poisson GARMA dan Binomial Negatif GARMA ke dalam data Jumlah Kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong diperoleh dari orde awal p dan q model ARMA yang kemudian dibentuk menjadi model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1) setelah melakukan uji signifikansi. Persamaan model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1) dengan memasukkan parameter yang sudah dicari adalah sebagai berikut :

$$\mu_t = \exp\left[1{,}572 + \sum_{j=1}^{1} 0{,}914 \left\{\ln y_{t-j}^{*} - 1{,}572\right\} - \sum_{j=1}^{1} 0{,}187 \left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right]$$

dan

$$\mu_t = \exp\left[1{,}497 + \sum_{j=1}^{1} 0{,}899 \left\{\ln y_{t-j}^{*} - 1{,}497\right\} - \sum_{j=1}^{1} 0{,}106 \left\{\frac{\ln y_{t-j}^{*}}{\mu_{t-j}}\right\}\right]$$

3. Perbandingan hasil peramalan dengan menggunakan model Poisson GARMA (1,1) dan Binomial Negatif GARMA (1,1) menunjukkan bahwa nilai RMSE Poisson GARMA yakni sebesar 0.189015023 dengan nilai MAPE sebesar 11.57 dan nilai RMSE Binomial Negatif

GARMA sebesar 0.151081028 dengan nilai MAPE sebesar 10.08. Dari nilai RMSE dan MAPE yang sudah diperoleh dapat disimpulkan bahwa Peramalan Jumlah Kecelakaan dengan menggunakan model Binomial Negatif GARMA (1,1) memiliki hasil yang lebih akurat dibanding peramalan dengan menggunakan model Poisson GARMA (1,1) . Sehingga model terbaik yang dapat digunakan untuk meramalkan Jumlah Kecelakaan di Jalan Tol Surabaya – Gempol ruas Banyu Urip – Porong adalah model Binomial Negatif GARMA (1,1).

5.2 Saran

1. Data yang digunakan dalam penelitian ini merupakan data kejadian yang jarang terjadi yaitu bersifat data jumlahan (*count*) yang nilainya merupakan data diskrit non negatif. Namun dri hasil peramalan yang sudah dilakukan data yang diperoleh memiliki sifat decimal sehingga perlu dilakukan pembulatan dalam hasilnya. Maka dari itu perlu dilakukan suatu pembalikan data agar menghasilkan data yang berupa data *count*.
2. Pada penelitian ini tidak melibatkan variabel prediktor sehingga diharpkan pada penelitian selanjutnya dapat melibatkan variabel prediktor agar dapat menghasilkan akurasi peramalan yang lebih baik.

DAFTAR PUSTAKA

Asrirawan , Heri Kuswanto, Suhartono. 2014. “Generalized Seasonal Autoregressive Integrated Moving Average (Gsarima) Models for Forecasting the Number of Dengue Hemorrhagic Fever (Dhf) Patients in Surabaya Average (Gsarima) Models for Forecasting the Number Of.” *Conference Paper* (February 2016).

Benjamin, Michael A., Robert A. Rigby, and D. Mikis Stasinopoulos. 2003. “Generalized Autoregressive Moving Average Models.” *Journal of the American Statistical Association* 98(461):214–23. doi: 10.1198/016214503388619238.

Cahyandari, Rini. 2014. “Pengujian Overdispersi Pada Model Regresi Poisson.” *Statistika* 14(2):69–76.

Cameron, Colin, and Pravin Trivedi. 1998. “Regression Analysis of Count Data (Econometric Society Monographs).” (30):96–137. doi: 10.1017/CCOL0521632013.

Chai, T., and R. R. Draxler. 2014. “Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? -Arguments against Avoiding RMSE in the Literature.” *Geoscientific Model Development* 7(3):1247–50. doi: 10.5194/gmd-7-1247-2014.

Cheek, P. J., P. McCullagh, and J. A. Nelder. 1990. “Generalized Linear Models, 2nd Edn.” *Applied Statistics* 39(3):385.

Chen, Peng, Hongyong Yuan, and Xueming Shu. 2008. “Forecasting Crime Using the ARIMA Model.” *Proceedings - 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008* 5(November):627–30. doi: 10.1109/FSKD.2008.222.

Fahza, Asep, and Hera Widyastuti. 2019. “Analisis Daerah Rawan Kecelakaan Lalu Lintas Pada Ruas Jalan Tol Surabaya-Gempol.” *Jurnal Teknik ITS* 8(1):54–59. doi: 10.12962/j23373539.v8i1.42123.

Garvey, Paul, Stephen Book, and Raymond Covert. 2015. “Iteratively Reweighted Least Squares.” *Probability Methods for Cost Uncertainty Analysis* 9(5):477–83. doi: 10.1201/b19143-23.

Green, P. J. 1984. “Iteratively Reweighted Least Squares for Maximum Likelihood Estimation, and Some Robust and Resistant Alternatives.” *Journal of the Royal Statistical Society: Series B (Methodological)* 46(2):149–70. doi: 10.1111/j.2517-6161.1984.tb01288.x.

Halim, Siana. 2006. “Diktat - Time Series Analysis Prakata.” (January).

Kedem, B & Fokianos, K. 2003. “Regression Models for Time Series Analysis.” *Technometrics* 45(4):364–364. doi: 10.1198/tech.2003.s166.

Lestiningsih, A., and E. Agustini. 2016. “Hubungan Penggunaan E-Toll Card Terhadap Pendapatan Tol Pada PT. Jasa Marga Tbk Cabang Ctc: Studi Kasus Gerbang Tol Tebet 1.” *Jurnal Online Insan Akuntan* 1(2):234091.

Putu, Ni, Ratindia Apriyanti, I. Ketut Gede, Darma Putra, and I. Made Suwija Putra. 2020. “Peramalan Jumlah Kecelakaan Lalu Lintas Menggunakan Metode Support Vector Regression.” *Jurnal Ilmiah Merpati* 8(2):72–80.

Startz, Richard. 2008. “Binomial Autoregressive Moving Average Models with an Application to U.S. Recessions.” *Journal of Business and Economic Statistics* 26(1):1–8. doi: 10.1198/073500107000000151.

Wardhani, L. P., D. Indrawati, N. Wahyuningsih, Setiawan, Suhartono, and H. Kuswanto. 2020. “The Theft Criminality Forecasting in the Surabaya District Police Region Using Poisson GARMA Model and Negative Binomial GARMA Model.” *Journal of Physics: Conference Series* 1490(1). doi: 10.1088/1742-6596/1490/1/012015.

Wei, William W. S. 2016. *Time Series Analysis Univariate and Multivariate Methods Second Edition*. Second Edi. edited by LynchDeirdre. Greg Tobin.

LAMPIRAN 1

Data Jumlah Kecelakaan Di Jalan Tol Surabaya – Gempol Ruas Banyu Urip – Porong

BULAN	TAHUN					
	2016	2017	2018	2019	2020	2021
Januari	3	5	5	2	6	1
Februari	5	4	3	3	4	1
Maret	4	5	0	2	0	2
April	5	5	5	1	2	4
Mei	4	1	4	1	2	2
Juni	3	6	4	3	0	4
Juli	6	4	1	6	0	2
Agustus	6	2	4	0	3	0
September	7	3	5	3	3	5
Oktober	4	4	3	3	1	1
November	9	4	2	5	0	3
Desember	3	3	7	7	4	2

LAMPIRAN 2

Source Code IRLS Estimasi Parameter

```
clear;close all
p       = 0;            % Autoregressive order
q       = 1;            % Moving average order
dist    = 2;            % Distribution, 1: Poisson, 2: Negative Binomial,
3: Gaussian
eStep   = 0;            % exogenous variables lag back w.r.t response
varable
h       = 2;            % h-steps-ahead forecasting
labels = {'intercept  ';};
x=xlsread('DATA TA','Sheet10','C2:C73');
y=xlsread('DATA TA','Sheet4','A1:BT1');
y=y';

% load jumlah kecelakaan               % load the Jumlah Kecelakaan data
set
[b1,Se1] = garmafit(x,y,p,q,dist,h,eStep); % Fit a GARMA model, aligned
exogenous vars
%----------------------------------------------------------------------------
% Print model parameters in a Table
%----------------------------------------------------------------------------
disp('Table 1  GARMA model parameters')
disp( sprintf('                MODEL Est.        Std Error.'))
m = size(x,2);      % number of exogenous variables
for i=1:m
disp(sprintf('%s:       %2.3f           %2.3f',[labels{i}],[b1(i)
Se1(i)]) )
end
for j=1:p
    disp(sprintf('Autoregressive %i:  %2.3f           %2.3f',[j b1(m+j)
Se1(m+j)]))
end
for k=1:q
    disp(sprintf('Moving-average %i:  %2.3f           %2.3f',[k b1(i+p+k)
Se1(i+p+k)]))
end
%----------------------------------------------------------------------------
--------------------
% Produce Figure 1 (Note: the residuals are randomized quantile)
%----------------------------------------------------------------------------
--------------------
%[yHat1,times1,dev1,RMSE,SS,res] = garmaval(b1,x,y,p,q,dist,h,eStep);
%----------------------------------------------------------------------------
--------------------
% Out-of-sample forecasting
%----------------------------------------------------------------------------
--------------------
nTrain  = 60;                  % Number of training data
tTrain  = 1:nTrain;             % Acquire the training time indices
xPre    = x(tTrain,:);          % Training set exogenous vars
yPre    = y(tTrain);            % Training set responses
tTest   = nTrain+1-(max(p,q)+h-1):72;           % Test time indices
xTest   = x(tTest,:);           % Test set exogenous vars
yPost   = y(tTest);             % Test set responses
[b2,Se2] = garmafit(xPre,yPre,p,q,dist,h,eStep); % Fit a garma model
```

Lanjutan Lampiran 2

```
clear;close all
p        = 1;            % Autoregressive order
q        = 1;            % Moving average order
dist     = 2;            % Distribution, 1: Poisson, 2: Negative Binomial,
3: Gaussian
eStep    = 0;            % exogenous variables lag back w.r.t response
varable
h        = 2;            % h-steps-ahead forecasting
labels = {'intercept  ';};
x=xlsread('DATA TA','Sheet10','C2:C73');
y=xlsread('DATA TA','Sheet4','A1:BT1');
y=y';

% load jumlah kecelakaan              % load the Jumlah Kecelakaan data
set
[b1,Se1] = garmafit(x,y,p,q,dist,h,eStep); % Fit a GARMA model, aligned
exogenous vars
%--------------------------------------------------------------------------
% Print model parameters in a Table
%--------------------------------------------------------------------------
disp('Table 1  GARMA model parameters')
disp( sprintf('                 MODEL Est.        Std Error.'))
m = size(x,2);      % number of exogenous variables
for i=1:m
disp(sprintf('%s:       %2.3f           %2.3f',[labels{i}],[b1(i)
Se1(i)]) )
end
for j=1:p
    disp(sprintf('Autoregressive %i:  %2.3f           %2.3f',[j b1(m+j)
Se1(m+j)]))
end
for k=1:q
    disp(sprintf('Moving-average %i:  %2.3f           %2.3f',[k b1(i+p+k)
Se1(i+p+k)]))
end
%--------------------------------------------------------------------------
--------------------
% Produce Figure 1 (Note: the residuals are randomized quantile)
%--------------------------------------------------------------------------
--------------------
%[yHat1,times1,dev1,RMSE,SS,res] = garmaval(b1,x,y,p,q,dist,h,eStep);
%--------------------------------------------------------------------------
--------------------
% Out-of-sample forecasting
%--------------------------------------------------------------------------
--------------------
nTrain   = 60;                  % Number of training data
tTrain   = 1:nTrain;             % Acquire the training time indices
xPre     = x(tTrain,:);          % Training set exogenous vars
yPre     = y(tTrain);            % Training set responses
tTest    = nTrain+1-(max(p,q)+h-1):72;        % Test time indices
xTest    = x(tTest,:);           % Test set exogenous vars
yPost    = y(tTest);             % Test set responses
[b2,Se2] = garmafit(xPre,yPre,p,q,dist,h,eStep); % Fit a garma model
```

LAMPIRAN 3

Source Code Garmafit

```
function [b,sE] = garmafit(x,y,p,q,distr,h,eStep,b0,bValue)
%   [ b,SEs ] = garmafit( x,y,p,q,distr )
%   Input :
%           x       : Matrix dari variabel eksogen (kovariat)
%           y       : Variabel Respon
%           p       : Autoregressive order
%           q       : Moving average order
%           distr   : Asumsi Distribusi :
%                   | 1: Poisson,           link: log
%                   | 2: Negative Binomial, link: log
%                   | 3: Gaussian,          link: identity
%           h       : Langkah melakukan peramalan
%           eStep   : Alignment of exogenous vars w.r.t response var
%                   | 0: Aligned
%                   | 1: lagged back h-step
%           b0      : Initial value untuk model parameters (optional)
%           bvalue  : Nilai pengganti angka 0 (optional)
%   Output :
%            b :    Parameter model
%            sE:    Standard errors untuk parameter model (optional)
%---------------------------------------------------------------------------

%---------------------------------------------------------------------------
if nargin < 6
    warning('one-step ahead forecasting');
    h = 1;
end

if nargin < 7
    warning('Exogenous varibles are aligned with responses');
    eStep = 0;
end
if ~isempty(x)                      % jika terdapat variabel eksogen
    m = size(x,2);                  % angka dari kovariat
    if length(y) ~=size(x,1)        % cek data untuk konsistensi
        error('Error : The response and covariate matrix must have same
number of rows')
    end
else
    m=0;
end
if nargin < 8                         % Jika pengguna tidak memasukkan nilai
inisial yang dicari
    if distr==1 %distribusi poisson
        b0 = 0.5+0.1*randn(m+p+q,1);          % initial point for search
    else
        b0 = 0.5+0.1*randn(m+p+q+1,1);        % initial point for search
(dispersion/variance is the last entry)
    end
else
    if size(b0,1) < 2                 % need a column vector
        b0 = b0';
    end

%---------------------------------------------------------------------------
% Ubah nilai 0 dengan boundary value (for logarithmic link)
```

```
%-------------------------------------------------------------------------
if nargin < 10
    bValue  = 0.1;                  % boundary value for replacing the zeros
end
ystar   = y;                        % work with y*
if distr < 3                        % if non-Gaussian
    ystar = max(y,bValue);          % replace zeros with bValue   parameter
threshold 0<c<1
end
n = length(y);                      % number of data
%-------------------------------------------------------------------------
% Optimization procedure
%-------------------------------------------------------------------------
% Masuk ke dalam prosedur optimasi likelihood
options = optimset('display','none','TolX',1e-11,'TolFun',1e-11,...
    'MaxFunEvals',1e6,'MaxIter',1e5);                          % Optimization
options

if eStep
    b = fminsearch( @(b) -logLik(x,ystar,y,exp(b),p,q,m,distr,h), ...
        b0,options);                                           % log
likelihood maximization
    [logL,mY] = logLik(x,ystar,y,exp(b),p,q,m,distr,h);
else                                                           % aligned
exogenous variables
    b = fminsearch( @(b) -logLik2(x,ystar,y,exp(b),p,q,m,distr,h), ...
        b0, options );
    [logL,mY] = logLik2(x,ystar,y,exp(b),p,q,m,distr,h);      % likelihood
function value, and predicted means
end
%-------------------------------------------------------------------------
% Standard Errors of model parameters
% note: standard error for the scale parameter is not evaluated.
%-------------------------------------------------------------------------
if nargout > 1                                   % Standard errors
calculation requested
    if distr<3
        gMu = ln(mY);
        gY  = ln(ystar);                             %IRLS
    else
        gMu = mY;
        gY  = ystar;
    end
    times    = length(gMu);                      % number of data
    if distr==1
        grad     = zeros(times,length(b));       % for poisson all sE's are
estimated
    else
        grad     = zeros(times,length(b)-1);     % for NB or Gaussian the
scale is excluded
    end
    if ~isempty(x)                               % if exogenous variables
are provided
        xB    = x*b(1:m);                        % sum over j of x_ij * b_j
    else
        xB = zeros(n,1);
    end
    for t = m+p+q+h:times                         % starting time for
gradient matrix
        for k = 1:m                               % calculations for
exogenous variables
```

```
            grad(t,k) = x(t,k);
            for j = 1:p
                grad(t,k) = grad(t,k) - b(m+j)*x(t-j,k);
            end
            for j = 1:q
                grad(t,k) = grad(t,k) - b(m+p+j)* grad(t-j-h+1,k);
            end
        end
        for k = m+1:m+p                          % calculations for
autoregressive terms
            grad(t,k) = grad(t,k)+ gY(t-k-h+1) - xB(t-k);
            for j = 1:q
                grad(t,k) = grad(t,k) - b(m+p+j)*grad(t-q-h+1,k);
            end
        end
        for k = m+p+1:m+p+q                      %                   moving
average terms
            grad(t,k) = grad(t,k) + gY(t-k) - gMu(t-k-h+1);
            for j =1:q
                grad(t,k) = grad(t,k) - b(m+p+j)*grad(t-q-h+1,k);
            end
        end
    end
    if distr==1                              % Poisson distribution
        scale   = 1;                         % exponential family scale
parameter dispersi
        w       = mY;                        % weights
bobot
    elseif distr==2                          % Negative Binomial
distribution
        scale   = 1;                         % parameter dispersi
        dispers = b(end);                    % bobot
        w       = dispers*mY./(dispers+mY);  % bobot
    else                                     % Gaussian distribution
        scale   = b(end);                    % scale is the conditional
variance
        w       = ones(times,1)/scale;
    end
    fI   = grad'*diag(w)*grad;               % Fisher information matrix
(IRLS)
    sE   = scale*sqrt(diag(inv(fI)));        % Standard error of model
parameters
end

end
function [logL,mY,it] = logLik(x,yStar,y,b,p,q,m,distr,h)
%---------------------------------------------------------------------------
--------------------
% Compute the log likelihood and E[Y] data with lagged exogenous variables
%---------------------------------------------------------------------------
--------------------
% xB(t)          = b(1)*x(t,1) + b(2)*x(t,2) + ... + b(m)*x(t,m)    linear
combination
% log E[y(t+1)] = xB(t) + b(m+1)*[ log y(t)     - xB(t-1) ]         auto-
regressive terms
%                        + b(m+2)*[ log y(t-1)   - xB(t-2) ]
%                        + b(m+p)*[ log y(t-p+1) - xB(t-p) ]
%                        + b(m+p+1)*[log (y(t)/E[y(t))]    ]        moving
average terms
%---------------------------------------------------------------------------
--------------------
```

```
b    = log(b);                    % because we input the exponenet of b
n    = length(yStar);             % number of data
i    = max(p+h,q);                % need "i" initial values for mean
if ~isempty(x)                    % if exogenous variables are provided
    xB    = x*b(1:m);             % sum over j of x_ij * b_j
else
    xB = zeros(n,1);
end
if distr < 3
    gY = log(yStar);              % apply link to the reponse
    gMu(1:i) = log(yStar(1:i)); % use y as initial conditions
else
    gY = yStar;
    gMu(1:i) = yStar(1:i);
end

for t = i+1:n                     % t is h-step into future
    gMu(t) = xB(t-h);             % begin with exogenous terms
    for j =1:p
        gMu(t)  = gMu(t) + b(m+j)*( gY(t-j-h+1) - xB(t-j-h) );      % add AR
terms
    end
    for k =1:q
        gMu(t) =gMu(t) + b(m+p+k)*(gY(t-k-h+1) - gMu(t-k-h+1) );       %
add MA terms
    end
end
it   = i+1:n;                             % indices at time t
if distr < 3
    mY   = exp(gMu(it))';                 % result for the mean (inverse of
the link)
else
    mY = gMu(it)';
end
if distr==1                               % Poisson
    lHs = (poisspdf(y(it),mY));          % likelihoods
elseif distr==2                           % Negative Binomial
    dispers = b(end);                     % dispersion parameter
    pNB     = dispers./(mY+dispers);     % convert to MATLAB parameters
    lHs   = nbinpdf(y(it),dispers,pNB);
else                                      % Gaussian
    phi     = b(end);                     % conditional variance
    lHs = normpdf(y(it),mY,sqrt(phi));
end
function [logL,mY,it] = logLik2(x,yStar,y,b,p,q,m,distr,h)
%-----------------------------------------------------------------------
--------------------
% Fits GARMA(p,q) to the data, with  alignment below.
%-----------------------------------------------------------------------
--------------------
% xB(t)          = b(1)*x(t,1) + b(2)*x(t,2) + ... + b(m)*x(t,m)    linear
combination
% log E[y(t)] = xB(t) + b(m+1)*[ log y(t-1)     - xB(t-1) ]         AR terms
%                         + b(m+2)*[ log y(t-2)    - xB(t-2) ]
%
%                         + b(m+p)*[ log y(t-p) - xB(t-p) ]         MA terms
%                         + b(m+p+1)*[log (y(t-1)/E[y(t-1))]
%-----------------------------------------------------------------------
--------------------
b = log(b);              % because we input the exponenet of b
n = length(yStar);       % number of data
```

```
i = max(p,q)+h-1;          % need "i" initial values for mu (mean of process)
if ~isempty(x)             % if exogenous variables are provided
    xB   = x*b(1:m);       % sum over j of x_ij * b_j
else
    xB = zeros(n,1);
end
if distr < 3
    gY = log(yStar);           % apply link to the reponse
else
    gY = yStar;
end
gMu(1:i) = gY(1:i);            % use y as initial conditions

for t = i+1:n                  % t is h-step into future
    gMu(t) = xB(t);            % begin with exogenous terms
    for j =1:p
        gMu(t)  = gMu(t) + b(m+j)*( gY(t-j-h+1) - xB(t-j-h+1) );        %
add AR terms
    end
    for k =1:q
        gMu(t)  = gMu(t) + b(m+p+k)*( gY(t-k-h+1) - gMu(t-k-h+1) );     %
add MA terms
    end
end
it   = i+1:n;                                          % indices at time t
if distr < 3
    mY   = exp(gMu(it))';                              % result for the mean
(inverse of the link)
else
    mY   = gMu(it)';
end
if distr==1                                            % Poisson
    lHs = poisspdf(y(it),mY);                          % likelihoods
elseif distr==2                                        % Negaitve Binomial
    dispers = b(end);                                  % dispersion parameter
    pNB     = dispers./(mY+dispers);                   % convert to MATLAB
parameters
    lHs     = nbinpdf(y(it),dispers,pNB);
else                                                   % Gaussian
    phi     = b(end);                                  % conditional variance
    lHs     = normpdf(y(it),mY,sqrt(phi));
end
lHs     = max(lHs,eps);                                % aviod numerical
issues with log(0)
logL    = sum(log(lHs));                               % log-likelihood (Eq. 3
of paper)
end
%-----------------------------------------------------------------------------
% Please send comments to: m.jalalpour@csuohio.edu
% Disclaimer:
% The authors reserve all rights but do not guaranty that the code is
% free from errors. Furthermore, we shall not be liable in any event
% caused by the use of the program.
%-----------------------------------------------------------------------------
```

LAMPIRAN 4

Hasil Peramalan *In Sample* Poisson GARMA dan Binomial Negatif GARMA

Data Aktual	**Hasil Peramalan *In Sample* Poisson GARMA $(1,1)$**	**Data Aktual**	**Hasil Peramalan *In Sample* Binomial Negatif GARMA $(1,1)$**
3	2.935598435	3	3.014612659
5	4.549669592	5	4.694567728
4	3.7574516	4	3.868878425
5	4.549669592	5	4.694567728
4	3.7574516	4	3.868878425
3	2.936136987	3	3.014926138
6	5.319446455	6	5.498411091
6	5.319446455	6	5.498411091
7	6.071048889	7	6.284530685
4	3.7574516	4	3.868878425
9	7.53077181	9	7.814286521
5	4.549669592	5	4.694567728
4	3.7574516	4	3.868878425
5	4.549669592	5	4.694567728
5	4.549669592	5	4.694567728
1	1.144756557	1	1.163225791
6	5.319446455	6	5.498411091
4	3.7574516	4	3.868878425
2	2.073973808	2	2.121409712
3	2.936136987	3	3.014926138
4	3.7574516	4	3.868878425
4	3.7574516	4	3.868878425
3	2.936136987	3	3.014926138
4	3.7574516	4	3.868878425
5	4.549669592	5	4.694567728
3	2.936136987	3	3.014926138
0	0.158985052	0	0.158041778
5	4.549669592	5	4.694567728
4	3.7574516	4	3.868878425
4	3.7574516	4	3.868878425
1	1.144756557	1	1.163225791
4	3.7574516	4	3.868878425
5	4.549669592	5	4.694567728
3	2.936136987	3	3.014926138
2	2.073973808	2	2.121409712
7	6.071048889	7	6.284530685
2	2.073973808	2	2.121409712
3	2.936136987	3	3.014926138

2	2.073973808	2	2.121409712
1	1.144756557	1	1.163225791
1	1.144756557	1	1.163225791
3	2.936136987	3	3.014926138
6	5.319446455	6	5.498411091
0	0.158985052	0	0.158041778
3	2.936136987	3	3.014926138
3	2.936136987	3	3.014926138
5	4.549669592	5	4.694567728
7	6.071048889	7	6.284530685
6	5.319446455	6	5.498411091
4	3.7574516	4	3.868878425
0	0.158985052	0	0.158041778
2	2.073973808	2	2.121409712
2	2.073973808	2	2.121409712
0	0.158985052	0	0.158041778
0	0.158985052	0	0.158041778
3	2.936136987	3	3.014926138
3	2.936136987	3	3.014926138
1	1.144756557	1	1.163225791
0	0.158985052	0	0.158041778
4	3.7574516	4	3.868878425

LAMPIRAN 5

Tabel Distribusi t

dk	α = 0.1	0.05	0.025	0.01	0.005	0.0025	0.001
40	1.303077	1.683851	2.021075	2.423257	2.704459	2.971171	3.306878
41	1.302543	1.682878	2.019541	2.420803	2.701181	2.966961	3.301273
42	1.302035	1.681952	2.018082	2.418470	2.698066	2.962962	3.295951
43	1.301552	1.681071	2.016692	2.416250	2.695102	2.959157	3.290890
44	1.301090	1.680230	2.015368	2.414134	2.692278	2.955534	3.286072
45	1.300649	1.679427	2.014103	2.412116	2.689585	2.952079	3.281480
46	1.300228	1.678660	2.012896	2.410188	2.687013	2.948781	3.277098
47	1.299825	1.677927	2.011741	2.408345	2.684556	2.945630	3.272912
48	1.299439	1.677224	2.010635	2.406581	2.682204	2.942616	3.268910
49	1.299069	1.676551	2.009575	2.404892	2.679952	2.939730	3.265079
50	1.298714	1.675905	2.008559	2.403272	2.677793	2.936964	3.261409
51	1.298373	1.675285	2.007584	2.401718	2.675722	2.934311	3.257890
52	1.298045	1.674689	2.006647	2.400225	2.673734	2.931765	3.254512
53	1.297730	1.674116	2.005746	2.398790	2.671823	2.929318	3.251268
54	1.297426	1.673565	2.004879	2.397410	2.669985	2.926965	3.248149
55	1.297134	1.673034	2.004045	2.396081	2.668216	2.924701	3.245149
56	1.296853	1.672522	2.003241	2.394801	2.666512	2.922521	3.242261
57	1.296581	1.672029	2.002465	2.393568	2.664870	2.920420	3.239478
58	1.296319	1.671553	2.001717	2.392377	2.663287	2.918394	3.236795
59	1.296066	1.671093	2.000995	2.391229	2.661759	2.916440	3.234207
60	1.295821	1.670649	2.000298	2.390119	2.660283	2.914553	3.231709

LAMPIRAN 6

Distribusi Poisson dan Binomial Negatif

1. Distribusi peluang peubah acak Poisson diberikan oleh persamaan berikut:

$$f(y;\mu_t) = \frac{e^{-\mu_t}\mu_t^{y_t}}{y!}$$

Rataan dari distribusi Poisson$(y;\mu_t)$ diberikan sebagai berikut:

$$E(Y) = \sum_{y=0}^{n} y\, f(y) = \sum_{y=0}^{n} y \frac{e^{-\mu_t}\mu_t^{y}}{y!}$$

$$= \sum_{y=0}^{n} y \frac{e^{-\mu_t}\mu_t\mu_t^{y-1}}{y(y-1)!}$$

$$= \mu_t \sum_{y=0}^{n} \frac{e^{-\mu_t}\mu_t^{y-1}}{(y-1)!}$$

$$= \mu_t$$

Varians dari distribusi Poisson $(y;\mu_t)$ diberikan sebagai berikut:

$$E(Y^2) = E(Y^2) - E(Y) + E(Y)$$

$$= E(Y^2 - Y) + E(Y)$$

$$= E\big(Y(Y-1)\big) + E(Y)$$

dimana,

$$E\big(Y(Y-1)\big) = \sum_{y=0}^{n} y(y-1) f(y)$$

$$= \sum_{y=0}^{n} y(y-1) \frac{e^{-\mu_t}\mu_t^{y}}{y!}$$

$$= \sum_{y=0}^{n} y(y-1) \frac{e^{-\mu_t}\mu_t^{2}\mu_t^{y-2}}{y(y-1)(y-2)!}$$

$$= \mu_t^2 \sum_{y=0}^{n} \frac{e^{-\mu_t}\mu_t^{y-2}}{(y-2)!}$$

$$= \mu_t^2$$

maka nilai harapan Y^2 adalah sebagai berikut:

$$E(Y^2) = E\big(Y(Y-1)\big) + E(Y)$$
$$= \mu_t^2 + \mu_t$$

Sehingga diperoleh nilai varians dari distribusi Poisson adalah sebagai berikut:

$$Var(Y) = E\left[\big(Y - E(Y)\big)^2\right]$$
$$= E\left[Y^2 - 2YE(Y) + \big(E(Y)\big)^2\right]$$
$$= E(Y^2) - 2E(Y)E(Y) + \big(E(Y)\big)^2$$
$$= E(Y^2) - \big(E(Y)\big)^2$$
$$= \mu_t^2 + \mu_t - (\mu_t)^2$$
$$= \mu_t$$

2. Pembentukan distribusi Binomial Negatif diperoleh dari gabungan distribusi Poisson dan distribusi Gamma, misalkan diketahui $y{\sim}Poi(\lambda u)$ dan $\lambda u{\sim}Gamma(v)$ maka gabungan distribusi Poisson dan Gamma adalah sebagai berikut:

$$f(y;u) = \int_0^\infty \frac{e^{-(\lambda u)}(\lambda u)^y}{y!}\frac{v^v}{\Gamma(v)}u^{v-1}e^{-vu}du$$
$$= \frac{\lambda^y}{y!}\frac{v^v}{\Gamma(v)}\int_0^\infty e^{-(\lambda+v)u}u^{y+v-1}du$$
$$= \frac{\lambda^y}{\Gamma(y+1)}\frac{v^v}{\Gamma(v)}\int_0^\infty e^{-(\lambda+v)u}u^{y+v-1}du$$

Misalkan:

$$(\lambda+v)u = a$$
$$(\lambda+v)du = da$$
$$du = \frac{da}{(\lambda+v)}$$

$$= \frac{\lambda^y}{\Gamma(y+1)}\frac{v^v}{\Gamma(v)}\int_0^\infty e^{-a}\left(\frac{a}{\lambda+v}\right)^{y+v-1}\frac{da}{(\lambda+v)}$$
$$= \frac{\lambda^y}{\Gamma(y+1)}\frac{v^v}{\Gamma(v)}\frac{1}{(\lambda+v)^{y+v}}\int_0^\infty e^{-a}a^{y+v-1}da$$
$$= \frac{\lambda^y}{\Gamma(y+1)}\frac{v^v}{\Gamma(v)}\frac{\Gamma(y+v)}{(\lambda+v)^{y+v}}$$
$$= \frac{\lambda^y}{\Gamma(y+1)}\frac{v^v}{\Gamma(v)}\Gamma(y+v)\left(\frac{\lambda}{\lambda+v}\right)^y\frac{1}{\lambda^y}$$

$$= \frac{\Gamma(y+v)}{\Gamma(y+1)\Gamma(v)}\left(\frac{v}{\lambda+v}\right)^{v}\left(\frac{\lambda}{\lambda+v}\right)^{y}$$

$$= \frac{\Gamma(y+v)}{\Gamma(y+1)\Gamma(v)}\left(\frac{1}{1+\frac{\lambda}{v}}\right)^{v}\left(1-\frac{1}{1+\frac{\lambda}{v}}\right)^{y}$$

dengan $v = \frac{1}{k}$ dan $\lambda = \mu$, maka

$$= \frac{\Gamma(y+\frac{1}{k})}{\Gamma(y+1)\Gamma(\frac{1}{k})}\left(\frac{1}{1+\frac{\mu}{\frac{1}{k}}}\right)^{\frac{1}{k}}\left(1-\frac{1}{1+\frac{\mu}{\frac{1}{k}}}\right)^{y}$$

$$= \frac{\Gamma\left(y+\frac{1}{k}\right)}{\Gamma(y+1)\Gamma\left(\frac{1}{k}\right)}\left(\frac{1}{1+k\mu}\right)^{\frac{1}{k}}\left(1-\frac{1}{1+k\mu}\right)^{y}$$

$$= \frac{\Gamma\left(y+\frac{1}{k}\right)}{\Gamma(y+1)\Gamma\left(\frac{1}{k}\right)}\left(\frac{1}{1+k\mu}\right)^{\frac{1}{k}}\left(\frac{1+k\mu-1}{1+k\mu}\right)^{y}$$

$$f(y;\mu) = \frac{\Gamma\left(y+\frac{1}{k}\right)}{\Gamma(y+1)\Gamma\left(\frac{1}{k}\right)}\left(\frac{1}{1+k\mu}\right)^{\frac{1}{k}}\left(\frac{k\mu}{1+k\mu}\right)^{y}$$

UCAPAN TERIMA KASIH

Penyelesaian penulisan tugas akhir ini tidak lepas dari orang-orang terdekat penulis yang telah mendukung dan memotivasi penulis. Oleh sebab itu, penulis mengucapkan terima kasih kepada:

1. Keluarga besar penulis yang selalu mendukung serta memotivasi penulis dalam pembuatan Tugas Akhir ini agar cepat terselesaikan.
2. Wira selaku teman terdekat penulis yang telah memberi semangat, masukkan, serta memberikan banyak cerita dalam kehidupan penulis.
3. Emil, Erlyn, dan Dinda yang telah memberikan cerita dan hal baik kepada penulis selama masa perkuliahan.
4. Teman-teman seperbimbingan yaitu Lola, Ulfa, Mado yang telah berbagi motivasi dan semangat selama pengerjaan Tugas Akhir.
5. Teman-teman seperjuangan, Matematika Institut Teknologi Sepuluh Nopember Angkatan 2018 yaitu angkatan MOD$\bar{u}$LO yang telah mengisi hari - hari penulis dengan penuh keceriaan, motivasi dan pengalaman.
6. Semua pihak yang tidak bisa penulis sebutkan satu persatu yang telah membantu dalam pengerjaan Tugas Akhir ini

Penulis juga mengharapkan kritik dan saran yang membangun dari berbagai pihak untuk penyempurnaan isi tugas akhir ini. Akhir kata, semoga tugas akhir ini bermanfaat bagi semua pihak yang bersangkutan.

Surabaya, Juli 2022

Theresia Anindya Puspaningrum

BIODATA PENULIS

Theresia Anindya Puspaningrum, anak kedua dari tiga bersaudara lahir di Blitar, 22 Maret 2000. Pendidikan formal yang telah ditempuh penulis yakni SDN Dukuh Menanggal II/425, SMPN 22 Surabaya, dan SMAN 15 Surabaya. Setelah lulus dari SMAN 15 Surabaya di tahun 2018, penulis berhasil diterima di Departemen Matematika Institut Teknologi Sepuluh Nopember melalui jalur SNMPTN dengan bidang minat Matematika Terapan. Penulis beberapa kali mengikuti lomba bola voli putri baik semasa SMA maupun kuliah. Selama menjalani masa perkuliahan penulis aktif di beberapa organisasi seperti UKM Voli dan diberi kepercayaan untuk menjadi Staff di Bidang Eksternal. Selain itu, penulis juga aktif dalam Himpunan Matematika ITS dengan menjadi Staff Sport and Art Department. Penulis juga pernah diberi kepercayaan untuk menjadi pemateri Mathematics Learning Video (MLV) yang diadakan oleh HIMATIKA ITS. Penulis juga pernah kerja praktik di Badan Pusat Statistik Sidoarjo dan bekerja di King Sejong Institute Surabaya sebagai Finance Officer. Demikian biodata singkat yang dapat penulis tuliskan, jika terdapat saran, kritik dan diskusi mengenai penelitian ini dapat dikirimkan melalui email nindyoow22@gmail.com. Terima kasih.