



TUGAS AKHIR - KM184801

APLIKASI MODEL BINOMIAL NEGATIF *GENERALIZED SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (GSARIMA) PADA PERAMALAN JUMLAH PENDERITA PENYAKIT INFEKSI SALURAN PERNAPASAN AKUT (ISPA)

**WINDYA HARIESKA PRAMUJATI
0611164000030**

Dosen Pembimbing
Dra. Laksmi Prita Wardhani, M.Si
Dra. Nuri Wahyuningsih, M.Kes

Departemen Matematika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember
Surabaya 2020



TUGAS AKHIR – KM184801

APLIKASI MODEL BINOMIAL NEGATIF *GENERALIZED SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE* (GSARIMA) PADA PERAMALAN JUMLAH PENDERITA PENYAKIT INFEKSI SALURAN PERNAPASAN AKUT (ISPA)

**Windya Harieska Pramujati
0611164000030**

**Dosen Pembimbing:
Dra. Laksmi Prita Wardhani, M.Si
Dra. Nuri Wahyuningsih, M.Kes**

**Departemen Matematika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember
Surabaya 2020**



FINAL PROJECT – KM184801

***APPLICATION OF NEGATIVE BINOMIAL GENERALIZED
SEASONAL AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (GSARIMA) MODELS IN FORECASTING THE
NUMBER OF ACUTE RESPIRATORY INFECTIONS (ARI)***

**Windya Harieska Pramujati
0611164000030**

**Supervisors:
Dra. Laksmi Prita Wardhani, M.Si
Dra. Nuri Wahyuningsih, M.Kes**

**Mathematics Department
Faculty of Science and Data Analytics
Institut Teknologi Sepuluh Nopember
Surabaya 2020**

LEMBAR PENGESAHAN

**APLIKASI MODEL BINOMIAL NEGATIF GENERALIZED
SEASONAL AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (GSARIMA) PADA PERAMALAN JUMLAH
PENDERITA PENYAKIT INFEKSI SALURAN PERNAPASAN
AKUT (ISPA)**

**APPLICATION OF NEGATIVE BINOMIAL GENERALIZED
SEASONAL AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE (GSARIMA) MODELS IN FORECASTING THE
NUMBER OF ACUTE RESPIRATORY INFECTIONS (ARI)**

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Matematika
Bidang Studi Matematika Terapan
Program Studi S-1 Departemen Matematika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember Surabaya


Oleh:

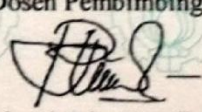
Windya Harieska Pramujati
NRP. 06111640000030

Menyetujui,

Dosen Pembimbing II,

Dosen Pembimbing I,


Dra. Nuri Wahyuningsih, M.Kes
NIP. 19660220 198903 2 002


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

Mengetahui,

Kepala Departemen Matematika FSAD ITS


Subchan, Ph.D

NIP. 19710513 199702 1 001

DEPARTEMEN
MATEMATIKA
Surabaya, Agustus 2020

Aplikasi Model Binomial Negatif *Generalized Seasonal Autoregressive Integrated Moving Average* (GSARIMA) Pada Peramalan Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA)

Nama Mahasiswa : Windya Harieska Pramujati
NRP : 0611164000030
Dosen Pembimbing : Dra. Laksmi Prita Wardhani, M.Si
Dra. Nuri Wahyuningsih, M.Kes

ABSTRAK

Infeksi Saluran Pernapasan Akut (ISPA) merupakan penyakit yang disebabkan oleh virus atau bakteri. Jumlah penderita ISPA biasanya meningkat karena adanya pengaruh musim. Salah satu provinsi di Indonesia dengan prevalensi ISPA tertinggi adalah Jawa Timur yaitu di Kota Surabaya. Berdasarkan fakta tersebut, maka akan dilakukan peramalan jumlah penderita ISPA di Kota Surabaya. Metode peramalan yang sering digunakan pada data *time series* adalah ARIMA. Namun pada peramalan data *count*, model Gaussian tidak selalu tepat digunakan, dan pada penelitian ini ditemukan adanya *overdispersion* (nilai varian variabel respon lebih besar dari pada nilai *mean*). Sehingga diterapkan sebuah model peramalan pada data *count* dengan pendekatan distribusi Binomial Negatif dan melibatkan efek musiman yaitu model Binomial Negatif GSARIMA. Pada penelitian ini didapatkan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ yang didapatkan dari identifikasi model SARIMA. Estimasi parameter model GSARIMA didapatkan dengan metode Bayesian. Terdapat dua model GSARIMA yaitu model GSARIMA transformasi ZQ1 dan transformasi ZQ2. Hasil yang didapatkan, peramalan dengan model Binomial Negatif GSARIMA transformasi ZQ1 lebih baik dibandingkan transformasi ZQ2 dengan nilai MARE sebesar 0,1311 yang diaplikasikan pada data jumlah penderita ISPA.

Kata Kunci: Binomial Negatif GSARIMA, ISPA, SARIMA

***Application Of Negative Binomial Generalized Seasonal
Autoregressive Integrated Moving Average (GSARIMA) Models
In Forecasting The Number Of Acute Respiratory Infections
(ARI)***

Name : Windya Harieska Pramujati
NRP : 06111640000030
Supervisors : Dra. Laksmi Prita Wardhani, M.Si
Dra. Nuri Wahyuningsih, M.Kes

ABSTRACT

Acute Respiratory Infection (ARI) is a disease caused by a virus or bacteria. The number of ARI sufferers usually increases due to seasonal influences. One of the provinces in Indonesia with the highest ARI prevalence is East Java, that is in Surabaya city. Based on these facts, forecasting the number of ARI in Surabaya will be conducted. Forecasting method, that is often used in time series data is ARIMA. However, in forecasting data count, the Gaussian model is not always appropriate, and this research found overdispersion (the variance value of the response variable is greater than the mean value). So the forecasting model is applied to the data count with Negative Binomial distribution and involves seasonal effects, that is Negative Binomial GSARIMA model. In this research, the Negative Binomial GSARIMA model $(2,1,1)(0,1,1)^{\phi}$ was obtained from the identification of the SARIMA model. The estimated parameters of the GSARIMA model are obtained by the Bayesian method. There are two GSARIMA models, namely the GSARIMA ZQ1 transformation model and ZQ2 transformation. The results obtained, forecasting with the Negative Binomial GSARIMA ZQ1 transformation model is better than the ZQ2 transformation with a MARE value is 0.1311 which is applied to data the number of ARI patients.

Keywords: ARI, Negative Binomial GSARIMA, SARIMA

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, atas limpahan rahmat, karunia, petunjuk, dan hidayahNya, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul:

“Aplikasi Model Binomial Negatif *Generalized Seasonal Autoregressive Integrated Moving Average* (GSARIMA) Pada Peramalan Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA)”

sebagai salah satu prasyarat akademis dalam menyelesaikan Program Sarjana Departemen Matematika FSAD Institut Teknologi Sepuluh Nopember Surabaya.

Tugas Akhir ini dapat diselesaikan berkat kerjasama, bantuan, dan dukungan dari banyak pihak. Sehubungan dengan hal tersebut, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua penulis yaitu Bapak Harko Sujiran dan Ibu Eni Susilowati atas segala doa, nasihat, motivasi, kasih sayang, pengorbanan, dukungan, dan segalanya yang selalu diberikan kepada penulis.
2. Bapak Subchan, Ph.D, selaku Kepala Departemen Matematika FSAD ITS yang telah memberikan dukungan fasilitas, sarana, dan prasarana.
3. Ibu Dr. Dwi Ratna Sulistyaningrum, S.Si, MT, selaku Sekretaris Departemen Bidang Akademik S1 Departemen Matematika FSAD ITS yang telah membantu kebutuhan administrasi selama periode pengambilan Tugas Akhir ini.
4. Ibu Dra. Laksmi Prita Wardhani, M.Si dan Ibu Dra. Nuri Wahyuningsih, M.Kes selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi kepada penulis dalam mengerjakan Tugas Akhir ini hingga dapat terselesaikan dengan baik.
5. Bapak Drs. Sentot Didik Surjanto, M.Si dan Ibu Endah Rokhmah M.P., Ph.D selaku dosen penguji yang telah memberikan saran demi kebaikan Tugas Akhir ini.

6. Bapak Drs. Komar Baihaqi, M.Si dan Bapak Prof. DR. Mohammad Isa Irawan, MT selaku dosen wali yang telah memberikan nasihat dan arahan selama penulis menempuh perkuliahan.
7. Seluruh Bapak dan Ibu Dosen dan para Staf Departemen Matematika ITS yang tidak dapat penulis sebutkan satu-persatu.
8. Bapak Faisol, Ibu Dian, dan Ibu Arofah yang telah membantu penulis dalam pengambilan data di Dinas Kesehatan Kota Surabaya.
9. Saudara penulis Ghilar Aushaftha Hanggara, Fazilya Triata Novaizhar, dan semua keluarga besar yang ikut dalam memberikan dukungan kepada penulis.
10. Para sahabat penulis Ika, Dian, Ratna, Nita, Safira, Farida, Trisna yang telah menjadi teman diskusi, serta memberikan doa, dukungan, dan semangat bagi penulis.
11. Teman-teman Matematika ITS 2016 terutama teman-teman “Lemniscate” yang telah memberikan banyak cerita selama perkuliahan dan banyak pihak yang tidak dapat dituliskan satu-persatu yang telah membantu penulis selama penulisan Tugas Akhir ini.

Penulis menyadari bahwa dalam penulisan Tugas Akhir ini terdapat banyak kekurangan, maka penulis mengharapkan kritik dan saran yang bersifat membangun demi penyempurnaan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi banyak pihak. Aamiin.

Madiun, Agustus 2020

Penulis

DAFTAR ISI

| | Halaman |
|--|---------|
| HALAMAN JUDUL..... | i |
| LEMBAR PENGESAHAN..... | v |
| ABSTRAK | vii |
| <i>ABSTRACT</i> | ix |
| KATA PENGANTAR..... | xi |
| DAFTAR ISI..... | xiii |
| DAFTAR GAMBAR | xv |
| DAFTAR TABEL | xvii |
| DAFTAR LAMPIRAN | xix |
| | |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 4 |
| 1.3 Batasan Masalah..... | 4 |
| 1.4 Tujuan | 4 |
| 1.5 Manfaat | 5 |
| 1.6 Sistematika Penulisan..... | 5 |
| | |
| BAB II TINJAUAN PUSTAKA | 7 |
| 2.1 Penelitian Terdahulu | 7 |
| 2.2 Analisis <i>Time Series</i> | 8 |
| 2.3 Stasioneritas | 8 |
| 2.3.1 Stasioneritas dalam Varian..... | 9 |
| 2.3.2 Stasioneritas dalam <i>Mean</i> | 9 |
| 2.4 Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF)..... | 10 |
| 2.5 Model-Model <i>Time Series</i> | 11 |
| 2.5.1 Model ARIMA | 11 |
| 2.5.2 Model SARIMA..... | 15 |
| 2.6 Generalized Linear <i>Model</i> | 15 |
| 2.7 Model Binomial Negatif Data <i>Count</i> | 16 |

| | | |
|----------------|--|------------|
| 2.8 | Model Binomial Negatif GSARIMA | 17 |
| 2.9 | Maximum Likelihood Estimation (MLE) | 20 |
| 2.10 | Metode Bayesian | 20 |
| 2.11 | Kriteria Pemilihan Model Terbaik..... | 22 |
| | 2.11.1 <i>Akaike's Information Criterion</i> (AIC)..... | 23 |
| | 2.11.2 <i>Mean Absolute Relative Error</i> (MARE).... | 23 |
| BAB III | METODOLOGI PENELITIAN..... | 25 |
| 3.1 | Sumber Data | 25 |
| 3.2 | Variabel Penelitian | 25 |
| 3.3 | Metode Penelitian..... | 26 |
| 3.4 | Diagram Alir..... | 27 |
| BAB IV | ANALISIS DAN PEMBAHASAN | 31 |
| 4.1 | Analisis Deskriptif..... | 31 |
| 4.2 | Tahap Identifikasi Model SARIMA..... | 34 |
| | 4.2.1 Tahap Identifikasi..... | 34 |
| | 4.2.2 Tahap Estimasi dan Uji Signifikansi Parameter | 39 |
| | 4.2.3 Uji Diagnostik | 43 |
| | 4.2.4 Pemilihan Model Terbaik..... | 45 |
| 4.3 | Model Binomial Negatif GSARIMA | 46 |
| 4.4 | Estimasi Parameter Model Binomial Negatif GSARIMA | 51 |
| 4.5 | Peramalan Model Binomial Negatif GSARIMA .. | 54 |
| BAB V | PENUTUP..... | 63 |
| 5.1 | Kesimpulan | 63 |
| 5.2 | Saran..... | 64 |
| | DAFTAR PUSTAKA..... | 65 |
| | LAMPIRAN | 69 |
| | BIODATA PENULIS..... | 105 |

DAFTAR GAMBAR

| | Halaman |
|-------------|---|
| Gambar 3.1 | Diagram Alir Pengerjaan Tugas Akhir.....28 |
| Gambar 3.2 | Diagram Alir Tahap Identifikasi Model SARIMA29 |
| Gambar 4.1 | Diagram Rata-rata Jumlah Hari Hujan Kota Surabaya 2015-2019.....32 |
| Gambar 4.2 | Plot <i>Time Series</i> Data Jumlah Penderita Penyakit ISPA33 |
| Gambar 4.3 | <i>Box-Cox Plot</i> Data Jumlah Penderita Penyakit ISPA35 |
| Gambar 4.4 | Transformasi 1 <i>Box-Cox Plot</i> Data Jumlah Penderita Penyakit ISPA35 |
| Gambar 4.5 | Transformasi 2 <i>Box-Cox Plot</i> Data Jumlah Penderita Penyakit ISPA36 |
| Gambar 4.6 | Plot ACF Hasil Transformasi 2 <i>Box-Cox Plot</i>36 |
| Gambar 4.7 | Plot PACF Hasil Transformasi 2 <i>Box-Cox Plot</i> ...37 |
| Gambar 4.8 | Plot ACF Hasil <i>Differencing Lag</i> Nonmusiman ..37 |
| Gambar 4.9 | Plot PACF Hasil <i>Differencing Lag</i> Nonmusiman 38 |
| Gambar 4.10 | Plot ACF Hasil <i>Differencing</i> pada <i>Lag</i> Musiman 38 |
| Gambar 4.11 | Plot PACF Hasil <i>Differencing</i> pada <i>Lag</i> Musiman39 |
| Gambar 4.12 | <i>Trace Plot</i> Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶52 |
| Gambar 4.13 | <i>Density Plot</i> Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶53 |
| Gambar 4.14 | Plot Hasil Peramalan Data <i>In-Sample</i> dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶55 |

| | | |
|-------------|--|----|
| Gambar 4.15 | Plot Hasil Peramalan Data <i>Out-Sample</i> dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ | 59 |
| Gambar 4.16 | Plot CDF Distribusi Binomial Negatif <i>One-Step</i> Ke Depan dengan $\mu t = 629,47$ | 59 |

DAFTAR TABEL

| | Halaman |
|--|---------|
| Tabel 2.1 Transformasi <i>Box-Cox</i> | 9 |
| Tabel 2.2 Kriteria Signifikansi Parameter | 22 |
| Tabel 3.1 Struktur Data Jumlah Penderita Penyakit ISPA | 25 |
| Tabel 4.1 Stastistik Deskriptif Data Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kota Surabaya Januari 2014 - Desember 2019 | 32 |
| Tabel 4.2 Estimasi Parameter Model SARIMA(2,1,2)(1,1,1) ⁶ .. | 40 |
| Tabel 4.3 Hasil Uji <i>White Noise</i> pada Model SARIMA Hasil <i>overfitting</i> yang Telah Memenuhi Uji Signifikansi Paremeter..... | 44 |
| Tabel 4.4 Hasil Uji <i>Kolmogorov-Smirnov</i> pada Model SARIMA Hasil <i>Overfitting</i> yang Telah Memenuhi Asumsi <i>White Noise</i> | 45 |
| Tabel 4.5 Kriteria Pemilihan Model Terbaik SARIMA | 46 |
| Tabel 4.6 Hasil Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ Transformasi ZQ1 | 54 |
| Tabel 4.7 Hasil Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ Transformasi ZQ2 | 54 |
| Tabel 4.8 Hasil Peramalan pada Data <i>Out-Sample</i> dengan Model Binomial Negatif GSARIMA (2,1,1)(0,1,1) ⁶ Transformasi ZQ1 dan ZQ2 | 56 |
| Tabel 4.9 Perbandingan Nilai MARE Hasil Peramalan pada Data <i>Out-Sample</i> dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ Transformasi ZQ1 dan ZQ2 | 57 |
| Tabel 4.10 Hasil Peramalan Jumlah Penderita Penyakit ISPA Kota Surabaya Tahun 2020 | 60 |

DAFTAR LAMPIRAN

| | Halaman |
|------------|---|
| LAMPIRAN A | <i>Mixed Distribution</i> Gamma dan Poisson 69 |
| LAMPIRAN B | Data Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA) Kategori Pneumonia Kota Surabaya Januari 2014-Desember 2019. 73 |
| LAMPIRAN C | Perhitungan Autokorelasi (ACF) dan Autokorelasi Parsial (PACF) 75 |
| LAMPIRAN D | Estimasi dan Uji Signifikansi Model <i>Overfitting</i> 81 |
| LAMPIRAN E | Uji <i>Ljung-box</i> pada Model SARIMA Hasil <i>Overfitting</i> yang Memenuhi Uji Signifikansi Parameter 89 |
| LAMPIRAN F | Uji <i>Kolmogorov-Smirnov</i> pada Model SARIMA Hasil <i>Overfitting</i> yang Memenuhi Uji <i>White Noise</i> 95 |
| LAMPIRAN G | <i>Syntax</i> Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ Transformasi ZQ1 dan ZQ2 di <i>R Software</i> 97 |
| LAMPIRAN H | <i>Output</i> Hasil Estimasi Model Binomial Negatif GSARIMA(2,1,1)(0,1,1) ⁶ di <i>R Software</i> 103 |

BAB I

PENDAHULUAN

Pada bab ini dibahas mengenai hal-hal yang menjadi latar belakang permasalahan dalam Tugas Akhir ini, kemudian permasalahan tersebut disusun kedalam suatu rumusan masalah. Selanjutnya dijabarkan batasan masalah untuk memperoleh tujuan serta manfaat.

1.1 Latar Belakang

Infeksi Saluran Pernapasan Akut (ISPA) merupakan penyakit yang disebabkan oleh virus atau bakteri yang menyerang salah satu bagian atau lebih dari saluran pernapasan mulai dari hidung (saluran atas) hingga alveoli (saluran bawah) [1]. ISPA menjadi salah satu penyebab utama morbiditas dan mortalitas penyakit menular di dunia. Hampir empat juta orang meninggal disetiap tahunnya akibat penyakit ISPA [2]. Di Indonesia ISPA merupakan salah satu penyebab utama kunjungan pasien di puskesmas dan rumah sakit [1]. Menurut Wantania dkk (2012) salah satu faktor yang menyebabkan terjadinya penyakit ISPA adalah keadaan lingkungan. Polutan lingkungan dapat mengiritasi saluran pernapasan sehingga memudahkan terjadinya infeksi. Pada saat musim kemarau kondisi udara sangat kering, sehingga meningkatkan polutan di udara yang menjadi salah satu faktor penyebab penyakit ISPA. Selain itu, curah hujan yang berlebihan menyebabkan kondisi lingkungan menjadi lembab, sehingga dapat juga meningkatkan risiko penyakit ISPA. Enam provinsi di Indonesia dengan prevalensi ISPA tertinggi berdasarkan diagnosis tenaga kerja menurut provinsi 2018 adalah Papua, Bengkulu, Papua Barat, Nusa Tenggara Timur, Kalimantan Tengah, dan Jawa Timur [3]. Sedangkan berdasarkan diagnosis tenaga kesehatan menurut kabupaten atau kota Provinsi Jawa Timur 2018, prevalensi ISPA tertinggi berada di Kota Surabaya [4]. Maka dari itu, pada penelitian ini dilakukan peramalan pada jumlah penderita penyakit ISPA di Kota Surabaya.

Metode peramalan yang paling dikembangkan saat ini adalah *time series*. Analisis *time series* adalah suatu analisis dengan menggunakan data dari masa lalu untuk memprediksi masa depan [5]. Data bentuk *time series* dapat dicatat dalam periode waktu yang sama seperti: harian, mingguan, bulanan, tahunan, ataupun periode waktu tertentu lainnya. Metode yang sering digunakan pada data bentuk *times series* adalah ARIMA. Data yang diproses pada model ARIMA adalah data stasioner yaitu proses tidak berubah dari waktu ke waktu dengan rata-rata dan varian yang konstan [6]. Model ARIMA diasumsikan eror berdistribusi normal, yang berakibat variabel respon hanya berdistribusi normal. Namun pada beberapa kasus, asumsi tersebut tidak terpenuhi. Salah satu contoh kasusnya adalah peramalan model ARIMA pada data jumlahan (*count data*). Data *count* adalah data yang berada pada domain diskrit dan merupakan bilangan bulat positif $\{0,1,2,3,\dots\}$ [7], contoh data *count* antara lain: jumlah kasus pencurian pada suatu wilayah dalam selang waktu tertentu, jumlah serangan hama pada satu hektar tanaman padi selama satu tahun, jumlah kecelakaan di jalan tol dalam selang waktu tertentu, dan lain-lain. Data *count* sering menyebabkan data tidak menyebar secara normal, untuk mengatasi permasalahan tersebut telah dikembangkan sebuah model peramalan non-Gaussian yang diterapkan pada data *count*, dimana distribusi yang digunakan seperti distribusi Poisson dan distribusi Binomial Negatif.

McCullagh dan Nelder mengembangkan model *Generalized Linear Models* (GLM) untuk menganalisis hubungan antara variabel respon dengan variabel prediktor dimana variabel respon tidak harus berdistribusi normal (Gaussian), tetapi termasuk dalam keluarga ekponensial [8]. Distribusi Poisson merupakan distribusi keluarga ekponensial, dimana terdapat asumsi bahwa data harus bersifat *equisdispersion* yaitu nilai varian variabel respon sama dengan nilai rata-rata, namun kondisi *equisdispersion* tidak selalu terpenuhi. Untuk mengatasinya dapat menggunakan distribusi yang juga termasuk keluarga ekponensial yaitu distribusi Binomial Negatif. Dalam penelitian yang dilakukan

Benjamin dkk (2003), dikembangkan model *Generalized Autoregressive Moving Average* (GARMA) untuk data-data yang mengikuti distribusi non-Gaussian seperti distribusi Poisson dan distribusi Binomial Negatif [9]. Model GARMA merupakan pengembangan dari model *Generalized Linear Models* (GLM) yang menghubungkan komponen ARMA dengan variabel prediktor ke transformasi parameter rata-rata dari distribusi data dengan menggunakan fungsi *link*. Pada tahun 2013, Briet dkk mengembangkan suatu model *Generalized Seasonal Autoregressive Integrated Moving Average* (GSARIMA) dengan melibatkan efek musiman dan orde *differencing* yang diterapkan pada data malaria di Sri Lanka dengan mengasumsikan data malaria mengikuti distribusi Binomial Negatif [7]. Briet dkk menggunakan pendekatan Bayesian untuk mengestimasi parameter-parameter model GSARIMA.

Di Indonesia, Asrirawan dkk (2014) menerapkan model GARMA dengan melibatkan efek musiman yaitu simulasi perbandingan tingkat akurasi GSARIMA dan SARIMA pada kasus jumlah penderita demam berdarah di Surabaya, dengan hasil yang diperoleh model GSARIMA relatif lebih baik dibandingkan dengan model SARIMA [10]. Kemudian pada tahun 2019, Diana Indrawati melakukan penelitian mengenai peramalan jumlah kriminalitas pencurian di wilayah Polrestabes Kota Surabaya dengan menggunakan model Poisson GARMA dan model Binomial Negatif GARMA dengan hasil model Binomial Negatif GARMA lebih baik daripada model Poisson GARMA [11]. Kemudian Moch. Fitriawan Eka Saputra dan Muhammad Rizky melakukan peramalan jumlah penderita penyakit ISPA pada laki-laki tahun 2019 yang memiliki pola musiman dengan metode ARIMA [12].

Berdasarkan uraian diatas maka pada penelitian ini akan diaplikasikan model Binomial Negatif GSARIMA pada peramalan suatu data yang merupakan salah satu data *count* dan terdapat unsur musiman yaitu data jumlah penderita penyakit ISPA di Kota Surabaya.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas maka dapat dirumuskan suatu rumusan masalah yaitu:

1. Bagaimana model peramalan jumlah penderita penyakit ISPA di Kota Surabaya menggunakan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$.
2. Bagaimana hasil peramalan jumlah penderita penyakit ISPA di Kota Surabaya menggunakan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$.

1.3 Batasan Masalah

Dalam penulisan Tugas Akhir ini, permasalahan yang dibahas akan dibatasi ruang lingkup pembahasannya. Batasan masalah dalam Tugas Akhir ini antara lain:

1. Data yang digunakan berupa data sekunder yaitu data bulanan jumlah penderita penyakit ISPA di Kota Surabaya pada bulan Januari 2014-Desember 2019.
2. Data penyakit ISPA yang digunakan adalah penyakit ISPA dengan kategori pneumonia.
3. Jenis distribusi keluarga eksponensial yang digunakan adalah Binomial Negatif sehingga dilakukan peramalan menggunakan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$.
4. Pemilihan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$ terbaik berdasarkan perbandingan nilai MARE yang terendah.
5. Hasil peramalan mengabaikan adanya kasus pandemi Covid-19.
6. *Software* yang digunakan dalam membantu menyelesaikan Tugas Akhir ini adalah Minitab16 dan *R software*.

1.4 Tujuan

Tujuan dari penulisan Tugas Akhir ini adalah:

1. Mendapatkan model peramalan jumlah penderita penyakit ISPA di Kota Surabaya dengan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$.

2. Memperoleh nilai peramalan jumlah penderita penyakit ISPA di Kota Surabaya dengan model Binomial Negatif GSARIMA(p, d, q)(P, D, Q)^s.

1.5 Manfaat

Manfaat yang diperoleh dari penulisan Tugas Akhir ini adalah sebagai berikut :

1. Mengembangkan wawasan keilmuan dan pengetahuan tentang Binomial Negatif GSARIMA(p, d, q)(P, D, Q)^s.
2. Memberikan informasi model terbaik yang diterapkan pada studi kasus jumlah penderita penyakit ISPA di Kota Surabaya.
3. Memberikan informasi tentang peramalan pada instansi atau pihak yang membutuhkan dan dapat dijadikan sebagai bahan pertimbangan dalam hal pengambilan keputusan terkait kasus ISPA.

1.6 Sistematika Penulisan

Sistematika penulisan dalam Tugas Akhir ini adalah sebagai berikut:

1. BAB I PENDAHULUAN

Bab I berisi tentang gambaran umum dari penulisan Tugas Akhir yang meliputi latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Bab II berisi tentang penelitian terdahulu beserta teori-teori dan konsep yang terkait dengan Tugas Akhir ini. Teori- teori yang digunakan dalam tugas akhir ini antara lain: model peramalan ARIMA, model SARIMA, model Binomial Negatif GSARIMA, metode estimasi parameter, dan metode pemilihan model terbaik.

3. BAB III METODOLOGI PENELITIAN

Bab III menjelaskan tentang sumber data, variabel penelitian dan langkah-langkah yang dilakukan dalam mengerjakan Tugas Akhir ini.

4. BAB IV ANALISIS DAN PEMBAHASAN

Bab IV berisi tentang aplikasi model Binomial Negatif GSARIMA terhadap data yang sudah ditentukan. Dimulai dari tahap identifikasi model SARIMA, pembentukan model Binomial Negatif GSARIMA, estimasi parameter model Binomial Negatif GSARIMA, kemudian pemilihan model terbaik, dan peramalan.

5. BAB V PENUTUP

Bab V berisi kesimpulan dari hasil pembahasan pada BAB IV serta saran untuk pengembangan penelitian berikutnya.

BAB II

TINJAUAN PUSTAKA

Pada bab ini dijabarkan beberapa teori dan konsep penunjang yang digunakan dalam menyelesaikan permasalahan dalam Tugas Akhir ini, yaitu teori dan konsep yang mendukung proses pengaplikasian model Binomial Negatif GSARIMA pada peramalan jumlah penderita penyakit ISPA di Kota Surabaya.

2.1 Penelitian Terdahulu

Dalam penelitian ini didasarkan pada penelitian terdahulu yaitu pada tahun 2003 Benjamin dkk mengembangkan model *Generalized Autoregressive Moving Average* (GARMA) untuk data-data yang mengikuti distribusi non-Gaussian seperti distribusi Poisson dan distribusi Binomial Negatif [9]. Model GARMA merupakan pengembangan dari model *Generalized Linear Models* (GLM) yang menghubungkan komponen ARMA dengan variabel prediktor ke transformasi parameter rata-rata dari distribusi data dengan menggunakan fungsi *link*. Kemudian Briet, Amerasinghe, dan Vounatsou (2013) mengembangkan model *Generalized Seasonal Autoregressive Integrated Moving Average* (GSARIMA) dengan melibatkan efek musiman dan orde *differencing* yang diterapkan pada data malaria di Sri Lanka [7]. Briet dkk mengasumsikan bahwa data malaria mengikuti distribusi Binomial Negatif karena datanya berada pada domain diskrit dan merupakan bilangan bulat positif serta telah dibuktikan bahwa terjadi kasus *overdispersi*, dimana nilai varian variabel respon lebih besar dari nilai rata-ratanya. Briet dkk menggunakan pendekatan Bayesian untuk mengestimasi parameter-parameter model GSARIMA.

Selanjutnya di Indonesia penerapan model GARMA dengan melibatkan stasioner dan musiman dilakukan oleh Asrirawan dkk (2014) yakni berupa simulasi perbandingan tingkat akurasi GSARIMA dan SARIMA pada kasus jumlah penderita demam berdarah di Surabaya, dengan hasil yang diperoleh model GSARIMA relatif lebih baik dibandingkan dengan model

SARIMA [10]. Penerapan selanjutnya dilakukan pada tahun 2018 oleh Agil Desti Fauzia yakni Estimasi Parameter pada Model Poisson GARMA dengan Algoritma IRLS (studi kasus: Peramalan Jumlah Kecelakaan di Jalan Tol Surabaya-Gempol) [13] dan Mada Aqil H yakni Estimasi Parameter pada Model Binomial Negatif GARMA dengan Algoritma IRLS (studi kasus : Peramalan Jumlah Kecelakaan di Jalan Tol Surabaya-Gempol) [14]. Kemudian pada tahun 2019 oleh Diana Indrawati yakni berupa Model Peramalan Jumlah Kriminalitas Pencurian di Wilayah Polrestabes Kota Surabaya dengan Menggunakan Model Poisson GARMA dan Model Binomial Negatif GARMA, dengan hasil model Binomial Negatif GARMA lebih baik daripada model Poisson GARMA [11]. Kemudian ditahun yang sama Moch. Fitriawan Eka Saputra dan Muhammad Rizky melakukan Peramalan Jumlah Penderita Penyakit ISPA pada Laki-laki Tahun 2019 dengan Metode ARIMA [12].

2.2 Analisis *Time Series*

Data *time series* adalah deretan rangkaian suatu pengamatan yang diambil dari waktu ke waktu dan dicatat menurut urutan waktu kejadian dengan periode waktu yang tetap. Data bentuk *time series* ini dapat dicatat dalam periode waktu seperti: harian, mingguan, bulanan, tahunan, ataupun periode waktu tertentu lainnya. Analisis *time series* pertama kali diperkenalkan dan dikembangkan pada tahun 1970 oleh Box dan Jenkins. Analisis *time series* adalah suatu analisis dengan data dari masa lalu untuk memprediksi masa depan [5]. Secara garis besar, tujuan dari analisis *time series* adalah untuk mendapatkan hubungan dinamis dari Z_t , yaitu pengamatan Z pada waktu ke- t dengan waktu-waktu sebelumnya ($t - 1$, $t - 2$, dan seterusnya) [5].

2.3 Stasioneritas

Dalam memodelkan data *time series* terdapat asumsi yang harus dipenuhi yaitu asumsi stasioner. Asumsi stasioner memiliki arti bahwa tidak terjadi pertumbuhan dan penurunan data. Data

dapat dikatakan stasioner apabila pola data berada disekitar *mean* dan varian yang konstan [6]. Maka dari itu kestasioneran suatu data dapat dilihat dari dua hal yaitu stasioneritas dalam varian dan stasioneritas dalam *mean* (rata-rata).

2.3.1 Stasioneritas dalam Varian

Suatu data *time series* dikatakan stasioner dalam varian jika struktur data dari waktu ke waktu mempunyai fluktuasi data yang tetap atau konstan. Secara teoritis dapat dikatakan bahwa pemodelan data *time series* $Y_1, Y_2, Y_3, \dots, Y_t$ memenuhi asumsi $Var(Y_t) = Var(Y_{t+k}) = \sigma^2$ untuk semua t [6]. Kestasioneran data dalam varian dapat dilihat dari *Box-Cox plot*, dimana data dikatakan stasioner apabila nilai *rounded value* sama dengan 1. Jika suatu deret data belum stasioner dalam varian, salah satu cara dalam mengatasi ketidakstasioneran tersebut adalah dengan adanya transformasi *Box-Cox*.

Transformasi *Box-Cox* adalah transformasi pangkat oleh λ pada variabel tak bebas Y sehingga transformasinya menjadi Y^λ , dimana λ adalah parameter yang perlu diduga. Secara umum aturan pada transformasi *Box-Cox* dan nilai λ dapat dilihat pada Tabel 2.1.

Tabel 2.1 Transformasi *Box-Cox*

| Estimasi λ | Transformasi |
|--------------------|--------------------------------|
| -1 | $1/Y_t$ |
| -0,5 | $1/\sqrt{Y_t}$ |
| 0 | $\ln(Y_t)$ |
| 0,5 | $\sqrt{Y_t}$ |
| 1 | Y_t (tidak ada transformasi) |

2.3.2 Stasioneritas dalam Mean

Suatu data *time series* dikatakan stasioner dalam *mean* apabila fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, dengan kata lain $E(Y_t) = E(Y_{t+k}) = \mu$ untuk semua t [6]. Stasioneritas dalam *mean* secara umum dapat dilihat dari plot *Autocorrelation Function* (ACF). Fungsi autokorelasi suatu data

stasioner, akan menurun cepat mendekati nol sesudah *lag* kedua atau ketiga. Sedangkan untuk data yang tidak stasioner, fungsi autokorelasi cenderung lambat menuju nol untuk beberapa periode waktu.

Apabila suatu data belum stasioner dalam *mean*, maka dapat distasionerkan dengan cara melakukan *differencing* (pembeda). Proses *differencing* pada orde pertama merupakan selisih antara data ke- t dengan data ke- $(t - 1)$, yaitu:

$$\Delta Z_t = Z_t - Z_{t-1} \quad (2.1)$$

Menggunakan operator shift mundur (*backshift*), persamaan (2.1) dapat ditulis kembali menjadi:

$$\Delta Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t$$

dengan ΔZ_t adalah variabel respon pada waktu ke t setelah dilakukan *differencing* dan $(1 - B)$ adalah pembeda pertama. Adapun bentuk *differencing* pada orde ke-2, yaitu:

$$\begin{aligned} \Delta^2 Z_t &= \Delta Z_t - \Delta Z_{t-1} \\ &= Z_t - Z_{t-1} - (Z_{t-1} - Z_{t-2}) \\ &= Z_t - 2Z_{t-1} + Z_{t-2} \\ &= (1 - 2B + B^2)Z_t \\ &= (1 - B)^2 Z_t \end{aligned}$$

dengan $(1 - B)^2$ adalah pembeda orde ke-2. Sehingga bentuk *differencing* pada orde ke- d dapat dinyatakan sebagai:

$$\Delta^d Z_t = (1 - B)^d Z_t$$

2.4 Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF)

Stasioneritas mengasumsikan bahwa distribusi probabilitas gabungan $p(Z_{t_1}, Z_{t_2})$ adalah sama untuk semua waktu t_1, t_2 yang merupakan interval konstan yang terpisah. Suatu data *time series* dipisahkan oleh interval k atau *lag* k . Sehingga dibawah asumsi stasioneritas, nilai korelasi Z_t dan Z_{t+k} yang dipisahkan oleh *lag* k harus sama untuk semua t . Fungsi autokorelasi (ACF) pada *lag* k didefinisikan sebagai berikut [5]:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}$$

dengan:

$\hat{\rho}_k$: autokorelasi pada lag k

$\hat{\gamma}_k$: autokovariansi pada lag k

\bar{Z} : rata-rata

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

n : jumlah data

Fungsi Autokorelasi Parsial (PACF) digunakan untuk mengukur tingkat kerataan antara Z_t dan Z_{t+k} apabila pengaruh lag $t+1, t+2, \dots, t+k-1$ dianggap terpisah. Untuk PACF dapat didekati dengan persamaan sebagai berikut:

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_j}$$

dengan:

$\hat{\rho}_k$: autokorelasi pada lag k

$\hat{\phi}_{k,k}$: autokorelasi parsial pada lag k

$\hat{\phi}_{k+1,j} = \hat{\phi}_{k,j} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}; j = 1, 2, \dots, k$

2.5 Model-Model Time Series

Salah satu model yang paling sering digunakan pada data *time series* adalah model ARIMA *Box-Jenkins*. Apabila suatu model ARIMA melibatkan unsur *seasonal* atau musiman, maka disebut sebagai model ARIMA musiman (SARIMA).

2.5.1 Model ARIMA

Model *time series* yang tidak stasioner dan telah dilakukan *differencing* dinamakan model *Autoregressive Integrated Moving Average* (ARIMA), yang dapat dituliskan sebagai berikut [5]:

$$\phi_p(B)(1-B)^d Z_t = \mu' + \theta_q(B)\alpha_t \quad (2.2)$$

dengan:

$\phi_p(B)$: operator *Autoregressive* (AR)

$\theta_q(B)$: operator *Moving Average* (MA)

Z_t : data observasi ke- t

μ' : suatu konstanta

B : operator *backshift*

$$\begin{aligned}
 (1 - B)^d Z_t &: \text{time series yang stasioner pada differencing ke-}d \\
 p &: \text{orde Autoregressive (AR)} \\
 d &: \text{orde differencing} \\
 q &: \text{orde Moving Average (MA)} \\
 \alpha_t &: \text{nilai error pada waktu ke- }t \\
 \phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\
 \theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q
 \end{aligned}$$

Untuk memilih model ARIMA yang sesuai pada data *time series*, terdapat prosedur *Box-Jenkins* yang digunakan. Prosedur ini terdiri dari empat tahapan yaitu sebagai berikut:

1. Identifikasi Model

Identifikasi model ARIMA dapat dilakukan dengan melihat plot *time series*, plot ACF, dan plot PACF. Plot ACF digunakan untuk menentukan orde q , plot PACF digunakan untuk menentukan orde p , dan banyaknya *differencing* untuk menentukan orde d dari model ARIMA(p, d, q).

2. Estimasi Parameter dan Uji Signifikansi Parameter

Dari model ARIMA(p, d, q) yang telah diidentifikasi pada tahap pertama, selanjutnya dilakukan estimasi parameter dari model. Ada beberapa metode yang dapat digunakan untuk mengestimasi parameter, salah satunya adalah metode *Maximum Likelihood Estimation* (MLE). Parameter yang diestimasi pada model ARIMA(p, d, q) adalah ϕ (parameter *Autoregressive*) dan θ (parameter *Moving Average*).

Setelah dilakukan perhitungan nilai estimasi kemudian setiap parameter tersebut diuji signifikansinya. Jika setiap parameter pada model ARIMA(p, d, q) signifikan, maka model layak digunakan. Uji signifikansi parameter ϕ meliputi [15]:

Hipotesis:

$H_0: \phi_p = 0$ (parameter ϕ_p tidak signifikan)

$H_1: \phi_p \neq 0$ (parameter ϕ_p signifikan)

Statistik uji:

$$t_{hitung} = \frac{\hat{\phi}_p}{SE(\hat{\phi}_p)}$$

dengan:

$\hat{\phi}_p$: parameter hasil estimasi

$SE(\hat{\phi}_p)$: *standart error* estimasi parameter

kriteria pengujian:

Jika $|t_{hitung}| > t_{tabel} = t_{\frac{\alpha}{2}; n-1}$ maka H_0 ditolak, yang artinya parameter ϕ signifikan.

3. Uji Diagnostik (*Diagnostic Check*)

Selanjutnya adalah perlu diperiksa kecukupan model dengan memeriksa apakah asumsi model sudah dipenuhi. Uji Diagnostik terdiri dari *overfitting* dan beberapa uji yang dilakukan untuk memeriksa kesesuaian model antara lain residual pada model ARIMA(p, d, q) harus memenuhi uji asumsi residual *white noise* dan uji asumsi residual berdistribusi normal.

a. Uji Asumsi Residual *White Noise*

Asumsi residual bersifat *white noise* artinya tidak terdapat korelasi antar residual atau residual bersifat identik dan *independent*. Uji asumsi *white noise* dapat dilakukan dengan uji *Ljung-Box*, yang meliputi [6]:

Hipotesis:

$H_0 : \rho_1 = \rho_2 = \dots = \rho_i = 0$ (residual *white noise*)

$H_1 : \text{Minimal ada satu } \rho_i \neq 0, i = 1, 2, \dots, k$ (residual tidak *white noise*)

Statistik uji:

$$Q = n(n+1) \sum_{k=1}^K \frac{(\hat{\rho}_k)^2}{n-k}, n > k$$

dengan:

$\hat{\rho}_k$: ACF residual pada *lag* k

n : banyaknya pengamatan

K : *lag* maksimum

Kriteria pengujian:

Jika $Q > \chi^2(\alpha; K - (p + q))$ atau *p-value* $< \alpha$ maka H_0 ditolak yang berarti bahwa residual tidak *white noise*.

dengan:

p : orde *Autoregressive*

q : orde *Moving Average*

b. Uji Asumsi Residual Berdistribusi Normal

Uji kenormalan pada residual dapat dilakukan dengan menggunakan uji *Kolmogorov Smirnov* yang meliputi [15]:

Hipotesis:

$H_0 : F(x) = F_0(x)$ (residual berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ (residual tidak berdistribusi normal)

Statistik uji:

$$D_{hitung} = \sup |S(x) - F_0(x)|$$

dengan:

$F(x)$: fungsi distribusi yang belum diketahui

$F_0(x)$: fungsi yang dihipotesiskan yaitu berdistribusi normal

$S(x)$: fungsi peluang kumulatif yang dihitung dari data sampel

Kriteria uji:

Jika $D_{hitung} > D_{tabel} = D_{\alpha;n}$ atau $p\text{-value} < \alpha$ maka H_0 ditolak artinya residual tidak berdistribusi normal.

c. *Overfitting*

Salah satu prosedur diagnostik cek yang dikemukakan Box Jenkins adalah *overfitting*. Pada tahap ini dilakukan perluasan terhadap model yang dihasilkan pada tahap identifikasi. *Overfitting* terhadap model dapat dipilih dengan cara memilih model dengan orde lebih rendah atau kombinasi dari orde pada model yang diidentifikasi sebelumnya. Model yang dihasilkan dari *overfitting* dijadikan sebagai model alternatif yang kemudian dicari model terbaik diantara model-model yang signifikan [6].

4. Peramalan (*Forecasting*)

Setelah melalui beberapa tahapan di atas, tahapan terakhir dari analisis *time series* adalah peramalan. Dalam praktiknya, model yang ditemukan dalam peramalan ini bukan merupakan model yang sebenarnya, melainkan hanya pendekatannya saja yang selalu mengandung kesalahan, baik dalam langkah identifikasi maupun estimasi [5].

2.5.2 Model SARIMA

Model ARIMA juga dapat diterapkan pada data yang memiliki pola musiman yang dirumuskan sebagai berikut:

$$\Phi_P(B^S)(1 - B^S)^D Z_t = \mu' + \theta_Q(B^S)\alpha_t \quad (2.3)$$

Kombinasi dari persamaan (2.2) dan (2.3) menghasilkan model multiplikatif *seasonal* ARIMA (SARIMA) yang dapat ditulis sebagai berikut [5]:

$$\phi_p(B)\Phi_P(B^S)(1 - B)^d(1 - B^S)^D Z_t = \mu' + \theta_q(B)\theta_Q(B^S)\alpha_t$$

dengan:

$\Phi_P(B^S)$: operator *Seasonal Autoregressive* (SAR)

$\theta_Q(B^S)$: operator *Seasonal Moving Average* (SMA)

μ' : suatu konstanta

P : orde *seasonal Autoregressive*

D : orde *seasonal differencing*

Q : orde *seasonal Moving Average*

S : periode musiman

$(1 - B)^d(1 - B^S)^D Z_t$: *time series* yang stasioner pada *differencing* nonmusiman ke- d , dan musiman ke- D

$$\Phi_P(B^S) = 1 - \phi_1 B^S - \phi_2 B^{2S} - \dots - \phi_P B^{PS}$$

$$\theta_Q(B^S) = 1 - \theta_1 B^S - \theta_2 B^{2S} - \dots - \theta_Q B^{QS}$$

2.6 Generalized Linear Model

Generalized Linear Models (GLM) merupakan pengembangan dari model linier klasik, dimana pada model ini variabel respon diasumsikan mengikuti distribusi keluarga eksponensial [8]. Menurut Keedem and Fokianos, *time series* yang mengikuti *Generalized Linear Model* terbagi menjadi beberapa komponen yaitu [16]:

1. Komponen Random

Densitas probabilitas respon termasuk ke dalam distribusi keluarga eksponensial pada bentuk natural atau kanonikal seperti berikut:

$$f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) = \exp \left\{ \frac{y_t \theta_t - b(\theta_t)}{\alpha_t(\phi)} + c(y_t; \phi) \right\}$$

dengan:

θ_t : fungsi kanonik

$\alpha_t(\phi) = \phi/w_t$

ϕ : parameter dispersi

w_t : parameter yang diketahui berupa bobot.

2. Komponen Sistemik

Terdapat fungsi monoton $g(\cdot)$ seperti berikut:

$$g(\mu_t) = \eta_t = \sum_{j=1}^p \beta_j Z_{(t-1)j} = \mathbf{Z}'_{t-1} \boldsymbol{\beta}$$

dengan:

$g(\cdot)$: fungsi *link*

η_t : prediktor linear dari model

2.7 Model Binomial Negatif Data Count

Salah satu model regresi yang merupakan terapan dari GLM adalah regresi Binomial Negatif. Sehingga distribusi Binomial Negatif memiliki ketiga komponen yaitu komponen random, komponen sistematis dan fungsi *link*. Pada Regresi Binomial Negatif variabel respon Y_t diasumsikan berdistribusi Binomial Negatif yang dihasilkan dari distribusi *mixture* Poisson-gamma. Untuk membentuk suatu model regresi pada distribusi Binomial Negatif, maka nilai parameter dari distribusi Poisson-gamma *mixture* dinyatakan dalam bentuk $\mu_t = \alpha\beta$ dan $k = 1/\alpha$ sehingga diperoleh *mean* dan varian dalam bentuk $E(y_t) = \mu_t$ dan $Var(y_t) = \mu_t + k\mu_t^2$. Pembentukan distribusi *mixture* Poisson-gamma dapat dilihat pada Lampiran A. Sehingga didapatkan fungsi kepadatan peluang distribusi Binomial Negatif adalah sebagai berikut [10]:

$$f(y_t; \mu_t, k) = \frac{\Gamma(y_t + 1/k)}{y_t! \Gamma(1/k)} \left(\frac{k\mu_t}{k\mu_t + 1} \right)^{y_t} \left(\frac{1}{k\mu_t + 1} \right)^{1/k}$$

dengan:

$y_t : 1, 2, \dots$

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

k : parameter dispersi

Sehingga fungsi distribusi keluarga eksponensial dari distribusi Binomial Negatif adalah sebagai berikut:

$$f(y_t; \mu_t, k) = \exp \left\{ \ln \left(\frac{\Gamma(y_t + 1/k)}{y_t! \Gamma(1/k)} \right) + y_t \ln \left(\frac{k\mu_t}{k\mu_t + 1} \right) + \frac{1}{k} \ln \left(\frac{1}{k\mu_t + 1} \right) \right\}$$

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

$$\eta_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_p x_{pt}$$

dalam matriks dituliskan dalam bentuk $\boldsymbol{\eta} = \mathbf{X}^T \boldsymbol{\beta}$, dengan $\boldsymbol{\eta}$ adalah vektor $n \times 1$ dari observasi, \mathbf{X}^T adalah matriks $n \times c$ dari variabel bebas, $\boldsymbol{\beta}$ adalah matriks $c \times 1$ dari koefisien regresi, dengan $c = p + 1$. Nilai ekspektasi dari variabel respon Y_t adalah diskrit dan bernilai positif. Maka untuk mentransformasikan nilai η_t ke rentang yang sesuai dengan rentang pada respon Y diperlukan suatu fungsi link $g(\cdot)$ yaitu:

$$g(\mu_t) = \ln(\mu_t) = \mathbf{X}^T \boldsymbol{\beta}$$

2.8 Model Binomial Negatif GSARIMA

Model GSARIMA merupakan pengembangan dari model GARMA dengan melibatkan efek musiman dan *differencing*. Diberikan $y^T = (y_t, y_{t+1}, \dots, y_{t+n})$ adalah model data *time series*, dan misalkan $y_t \sim BN(\mu_t, \alpha)$ dengan $E(y_t) = \mu_t$ dan $Var(y_t) = \mu_t + \frac{\mu_t^2}{\alpha}$. Jika $\alpha \rightarrow \infty$ maka y_t akan mengikuti distribusi Poisson. Adapun model GARMA(p, q) adalah sebagai berikut [7]:

$$g(\mu_t) = \phi_p(B) [\mathbf{X}_t^T \boldsymbol{\beta} - g(y_t)] + g(y_t) - \theta_q(B) [g(y_t) - g(\mu_t)] + g(y_t) - g(\mu_t) \quad (2.4)$$

dengan:

$g(\cdot)$: fungsi link

B : operator *backshift*
 β : koefisien vektor \mathbf{X}_t^T
 $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$
 $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$
 $B^d y_t = y_{t-d}$
 $\beta^T = (\beta_0 \beta_1 \dots \beta_p)$
 $\mathbf{X}_t^T = (x_0 x_{1t} \dots x_{pt})$
 x_0 : *intercept*, biasanya digunakan $x_0 = 1$

Pada GARMA, data *count* dimodelkan dengan menggunakan sebuah fungsi *link logarithmic* atau *identity*. Zeger dan Qaqish (1998) mengusulkan sebuah transformasi dari y'_t yaitu $y'_t = \max(y_t, c)$ dengan $0 < c \leq 1$ yang selanjutnya disebut transformasi ZQ1. Zeger dan Qaqish juga mengusulkan metode alternatif yaitu dengan menambahkan nilai parameter c pada masing-masing data, yang kemudian disebut transformasi ZQ2 [7].

Pada model Binomial Negatif GARMA, variabel respon Y_t diasumsikan berdistribusi Binomial Negatif yang dihasilkan dari distribusi *mixture* Poisson-gamma. Sehingga untuk membentuk model Binomial Negatif GARMA diperlukan fungsi *link logarithmic* yaitu $g(\mu_t) = \ln(\mu_t)$, sehingga persamaan (2.4) dengan transformasi ZQ1 menjadi:

$$\begin{aligned}
 \ln(\mu_t) &= \phi_p(B)\{\mathbf{X}_t^T \beta - \ln(y'_t)\} + \ln(y'_t) \\
 &\quad - \theta_q(B)(\ln(y'_t) - \ln(\mu_t)) + \ln(y'_t) - \ln(\mu_t) \\
 \ln(\mu_t) &= \phi_p(B)\{\mathbf{X}_t^T \beta - \ln(y'_t)\} + \ln(y'_t) - \theta_q(B) \ln\left(\frac{y'_t}{\mu_t}\right) \\
 &\quad + \ln\left(\frac{y'_t}{\mu_t}\right) \tag{2.5}
 \end{aligned}$$

dengan $y'_t = \max(y_t, c)$ dan $0 < c \leq 1$. Jika persamaan (2.5) dikembangkan menjadi model Binomial Negatif GSARIMA dengan memasukkan unsur musiman dan komponen *differencing*, maka diperoleh persamaan [7]:

$$\begin{aligned}
\ln(\mu_t) &= \phi_p(B)\Phi_P(B^S) (1 - B)^d(1 - B^S)^D \{\mathbf{X}_t^T \boldsymbol{\beta} - \ln(y'_t)\} \\
&\quad + \ln(y'_t) - \theta_q(B)\theta_Q(B^S)\ln\left(\frac{y'_t}{\mu_t}\right) \\
&\quad + \ln\left(\frac{y'_t}{\mu_t}\right)
\end{aligned} \tag{2.6}$$

dengan $y'_t = \max(y_t, c)$ dan $0 < c \leq 1$, dimana persamaan (2.6) merupakan model Binomial Negatif GSARIMA dengan transformasi ZQ1. Apabila persamaan (2.6) diubah ke dalam transformasi ZQ2 akan menjadi persamaan (2.7), dimana persamaan (2.7) merupakan model Binomial Negatif GSARIMA dengan transformasi ZQ2 dengan formula sebagai berikut [7]:

$$\begin{aligned}
\ln(\mu_t) &= \phi_p(B)\Phi_P(B^S)(1 - B)^d(1 - B^S)^D \{\ln[\exp(\mathbf{X}_t^T \boldsymbol{\beta} + c)] \\
&\quad - \ln(y_t + c)\} + \ln(y_t + c) \\
&\quad - \theta_q(B)\theta_Q(B^S)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\
&\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)
\end{aligned} \tag{2.7}$$

dengan

$\phi_p(B)$: operator *Autoregressive* (AR)

$\theta_q(B)$: operator *Moving Average* (MA)

$\Phi_P(B^S)$: operator musiman *Autoregressive* (AR)

$\theta_Q(B^S)$: operator musiman *Moving Average* (MA)

p : orde *Autoregressive* nonmusiman

d : orde *differencing* nonmusiman

q : orde *Moving Average* nonmusiman

P : orde *Autoregressive* musiman

D : orde *differencing* musiman

Q : orde *Moving Average* musiman

S : periode musiman

B : operator *backshift*

$\mathbf{X}_t^T \boldsymbol{\beta}$: prediktor linier GLM

2.9 Maximum Likelihood Estimation (MLE)

Metode *Maximum Likelihood Estimation* adalah salah satu metode pendugaan parameter dengan memaksimalkan fungsi *Likelihood*. Adapun fungsi *Likelihood* $L(\theta)$ sebagai berikut [17].

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta)$$

Misal diketahui populasi $X \sim f(x|\theta)$, maka langkah-langkah metode MLE adalah 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

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

3. Mendapatkan $\hat{\theta}$ yang memaksimalkan fungsi *Likelihood* $\ln L(\theta)$.

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0$$

dimana

$$\ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta)$$

yang kemudian didapatkan nilai $\hat{\theta}$.

2.10 Metode Bayesian

Metode Bayesian merupakan metode estimasi parameter dengan menggabungkan informasi dari sampel dengan informasi lain yang telah tersedia. Kelebihan dari metode Bayesian terletak pada kemampuannya dalam mengakomodasi informasi awal yang dimiliki oleh peneliti. Sehingga kesimpulan akhir dari nilai dugaan parameter tidak hanya berdasarkan pada informasi dari sampel saja, melainkan juga berdasarkan informasi subjektif dari peneliti (*prior*) yang biasanya dikuantifikasi dalam bentuk distribusi

parameternya (distribusi *prior*). Teorema Bayesien dapat ditulis sebagai berikut [18]:

$$p(\theta|x) = \frac{L(\theta|x)p(\theta)}{p(x)} \quad (2.8)$$

dengan:

x : data yang diobservasi

θ : parameter data

$p(\theta|x)$: informasi *posterior*

$L(\theta|x)$: informasi sampel (fungsi *likelihood*)

$p(\theta)$: informasi *prior*

$p(x)$: fungsi konstanta

Persamaan (2.8) dapat juga ditulis sebagai berikut [18]:

$$p(\theta|x) \propto L(\theta|x)p(\theta)$$

dengan kata lain *posterior* yang merupakan hasil yang digunakan dalam mengambil keputusan didapatkan dengan mengalikan *likelihood* yang berisi informasi dari sampel, dengan *prior* parameter yang berisi informasi dari data sebelumnya.

Di dalam analisis Bayesian, penggunaan metode *Markov Chain Monte Carlo* (MCMC) dapat mempermudah analisis dan komputasi numerik. *Markov Chain Monte Carlo* (MCMC) merupakan metode simulasi untuk membangkitkan peubah-peubah acak menggunakan rantai markov sehingga diperoleh kondisi yang konvergen. Teknik simulasi ini banyak digunakan dalam menyelesaikan permasalahan akibat tidak dapat disimulasikan secara langsung, salah satunya pada inferensia Bayesian. Salah satu prinsip penggunaan *Markov Chain Monte Carlo* (MCMC) yaitu dengan *Gibbs Sampling*. Algoritma dari *Gibbs Sampling* adalah sebagai berikut [19]:

1. Menentukan nilai awal (*initial value*) untuk masing-masing parameter

$$\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)}).$$

2. Membangkitkan komponen $\theta^{(m+1)} = (\theta_1^{(m+1)}, \theta_2^{(m+1)}, \dots, \theta_p^{(m+1)})$, dimana nilai dari

$$\begin{aligned} \theta_1^{(m+1)} &\text{ berasal dari } p\left(\theta_1 \mid \mathbf{x}, \theta_2^{(m)}, \theta_3^{(m)}, \dots, \theta_p^{(m)}\right) \\ \theta_2^{(m+1)} &\text{ berasal dari } p\left(\theta_2 \mid \mathbf{x}, \theta_1^{(m+1)}, \theta_3^{(m)}, \dots, \theta_p^{(m)}\right) \\ &\vdots \\ \theta_p^{(m+1)} &\text{ berasal dari } p\left(\theta_p \mid \mathbf{x}, \theta_1^{(m+1)}, \theta_2^{(m+1)}, \dots, \theta_{p-1}^{(m+1)}\right) \end{aligned}$$

3. Melakukan *monitoring* terhadap kekonvergenan algoritma. Apabila algoritma belum konvergen, maka perlu membangkitkan lebih banyak observasi dan mengulangi langkah (1) dan (2).
4. Membuang b observasi pertama (*burn in period*) untuk menghindari pengaruh *initial value*.
5. Menganggap $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(x)}$ dan sebanyak b observasi pertama sudah dibuang, kemudian diperoleh $\boldsymbol{\theta}^{(b+1)}, \boldsymbol{\theta}^{(b+2)}, \dots, \boldsymbol{\theta}^{(k)}$ sebagai sampel untuk analisis *posterior*.
6. Membuat *plot* distribusi *posterior*.
7. Mendapatkan *mean*, median, simpangan baku dari distribusi *posterior*.

Parameter hasil estimasi pada metode Bayesian dikatakan signifikan pada taraf signifikansi α , apabila memenuhi kriteria pada Tabel 2.2

Tabel 2.2 Kriteria Signifikansi Parameter

| <i>Confident Interval</i> | | | Keterangan |
|---------------------------|--------|------------|---------------------------|
| α | Median | $1-\alpha$ | |
| + | + | + | Estimasi bernilai positif |
| - | - | - | Estimasi bernilai negatif |

2.11 Kriteria Pemilihan Model Terbaik

Terdapat beberapa macam ukuran yang dapat digunakan sebagai kriteria pemilihan model terbaik. Dalam Tugas Akhir ini kriteria pemilihan model yang digunakan adalah *Akaike's Information Criterion* (AIC), dan *Mean Absolute Relative Error* (MARE).

2.11.1 Akaike's Information Criterion (AIC)

Akaike's Information Criterion (AIC) adalah penduga kualitas relatif dari model statistik untuk sekumpulan data yang diberikan. Diberikan beberapa model untuk data, AIC memperkirakan kualitas masing-masing model, maka AIC menyediakan sarana untuk pemilihan model tersebut. Kriteria AIC dapat dirumuskan sebagai berikut [5] :

$$AIC = -2 \ln L(\hat{\theta}) + 2M$$

dengan:

$L(\hat{\theta})$: nilai maksimum dari fungsi *likelihood* untuk model

M : jumlah parameter pada model

Kriteria AIC diperkenalkan pertama kali oleh Akaike pada tahun 1973, dimana kriteria ini mempertimbangkan banyaknya parameter pada model dan dapat digunakan untuk membandingkan model dengan distribusi probabilitas. Kriteria AIC dalam penelitian ini digunakan untuk memilih model SARIMA terbaik yang akan digunakan sebagai orde pada model Binomial Negatif GSARIMA. Pemilihan model terbaik didapat dari model yang memiliki nilai AIC terkecil.

2.11.2 Mean Absolute Relative Error (MARE)

Mean Absolute Relative Error (MARE) merupakan salah satu kriteria pemilihan model peramalan terbaik yang dilakukan dengan menggunakan pendekatan *out-sample*. MARE menyatakan seberapa besar akurasi nilai dugaan pada suatu model peramalan, yang dinyatakan dalam bentuk rata-rata absolut kesalahan. Perhitungan nilai MARE dapat dilakukan dengan formula sebagai berikut [7]:

$$MARE = \frac{\sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t + 1} \right|}{N} \quad (2.9)$$

dengan:

y_t : nilai data ke-t

\hat{y}_t : nilai peramalan ke-t

N : banyak pengamatan

Pada penelitian ini MARE digunakan untuk menentukan model peramalan terbaik GSARIMA yang dapat diaplikasikan pada data jumlah penderita penyakit ISPA di Kota Surabaya. Model terbaik adalah model yang memiliki nilai MARE yang paling kecil.

BAB III

METODOLOGI PENELITIAN

Pada bab ini dijabarkan mengenai hal mendasar beserta langkah-langkah dalam penelitian, diantaranya sumber data, variabel penelitian, metode, dan tahapan penelitian.

3.1 Sumber Data

Pada penelitian ini digunakan data pada Lampiran B yaitu berupa data sekunder jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA) kategori pneumonia di Kota Surabaya pada bulan Januari 2014 sampai bulan Desember 2019 dengan jumlah data sebanyak 72. Data tersebut diperoleh dari Dinas Kesehatan Kota Surabaya.

3.2 Variabel Penelitian

Pada penelitian ini variabel respon Y yang digunakan adalah jumlah penderita penyakit ISPA di Kota Surabaya tahun 2014-2019. Adapun struktur data dalam penelitian ini ditunjukkan pada Tabel 3.1

Tabel 3.1 Struktur Data Jumlah Penderita Penyakit ISPA

| Tahun | Bulan | Jumlah Penderita ISPA (Y) |
|-------|----------|---------------------------|
| 2014 | Januari | Y_1 |
| | Februari | Y_2 |
| | ⋮ | ⋮ |
| | Desember | Y_{12} |
| 2015 | Januari | Y_{13} |
| | Februari | Y_{14} |
| | ⋮ | ⋮ |
| | Desember | Y_{24} |
| ⋮ | ⋮ | ⋮ |
| 2019 | Januari | Y_{61} |
| | Februari | Y_{62} |
| | ⋮ | ⋮ |
| | Desember | Y_{72} |

3.3 Metode Penelitian

Metode dan tahapan penelitian yang dilakukan adalah sebagai berikut:

1. Studi literatur

Studi literatur ini dilakukan untuk memperoleh dan memahami teori-teori yang berhubungan dengan pemecahan masalah pada Tugas Akhir ini. Konsep yang harus dipahami oleh peneliti antara lain mengenai ISPA, *time series*, model ARIMA, model SARIMA, model binomial negatif GSARIMA, metode estimasi, dan metode pemilihan model terbaik.

2. Pengambilan data

Tahap pengumpulan data ini dilakukan untuk memperoleh bahan penelitian sesuai dengan tujuan penelitian yang telah ditetapkan. Data yang dibutuhkan dalam penelitian ini adalah data bulanan jumlah penderita penyakit ISPA di Kota Surabaya.

3. Melakukan analisis deskriptif terhadap variabel penelitian Y berupa:

- a. Menghitung nilai *mean*, varian, standar deviasi, nilai minimum, dan nilai maksimum.
- b. *Time series plot* dengan bantuan *software* minitab, serta mengevaluasi adanya unsur musiman pada data.

4. Identifikasi model SARIMA

Orde dari model GSARIMA(p,d,q)(P,D,Q)^s diperoleh dari tahap identifikasi model SARIMA(p,d,q)(P,D,Q)^s terlebih dahulu. Tahap-tahap yang dilakukan adalah sebagai berikut:

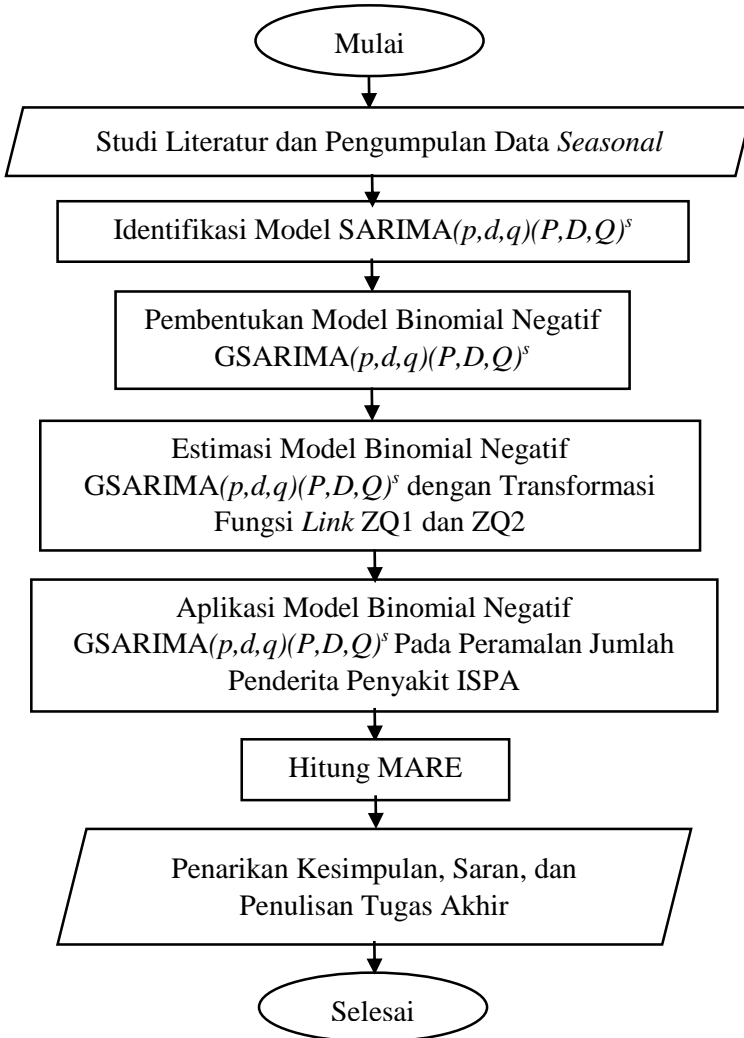
- a. Menguji kestasioneran data *time series* baik stasioneritas dalam varian maupun *mean*.
- b. Mengidentifikasi model SARIMA(p,d,q)(P,D,Q)^s dugaan sementara dengan cara menentukan orde AR, MA, SAR, SMA, dan apakah terdapat unsur *differencing* yang dapat dilihat dari grafik ACF dan PACF.
- c. Estimasi parameter model SARIMA(p,d,q)(P,D,Q)^s dugaan sementara dan dilanjutkan uji signifikansi parameter.
- d. Melakukan pemeriksaan diagnostik yang meliputi uji residual *white noise*, uji normalitas residual, dan *overfitting*.

- e. Menentukan model terbaik $SARIMA(p,d,q)(P,D,Q)^s$.
5. Pembentukan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$, dimana orde-ordennya didapat dari model terbaik $SARIMA(p,d,q)(P,D,Q)^s$ pada tahap identifikasi model sebelumnya.
 6. Mendapatkan penaksir parameter model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$ dengan menggunakan metode Bayesian. Terdapat dua jenis transformasi fungsi *link* yang dapat digunakan dalam proses estimasi $GSARIMA$, yaitu transformasi fungsi *link* ZQ1 dan ZQ2.
 7. Melakukan peramalan model Binomial Negatif $GSARIMA(p,d,q)(P,D,Q)^s$ terhadap data jumlah penderita penyakit ISPA. Model yang diperoleh adalah sebuah model pendekatan, dimana terdapat dua hasil peramalan yaitu peramalan dengan model $GSARIMA$ dengan transformasi ZQ1 dan transformasi ZQ2, kemudian dihitung nilai MARE pada masing-masing hasil peramalan. Model terbaik diperoleh dari model yang memiliki nilai MARE yang paling kecil.
 8. Kesimpulan dan Saran
Setelah semua pengolahan, interpretasi, dan analisis data maka ditarik suatu kesimpulan yang merupakan ringkasan akhir dari hasil yang mampu menjawab tujuan penelitian yang dilakukan. Setelah itu diberikan pula saran-saran berupa perbaikan maupun pengembangan untuk penelitian mendatang.
 9. Penyusunan laporan Tugas Akhir
Penulisan laporan Tugas Akhir ini dilakukan dari awal dilakukannya penelitian hingga waktu yang telah ditentukan.

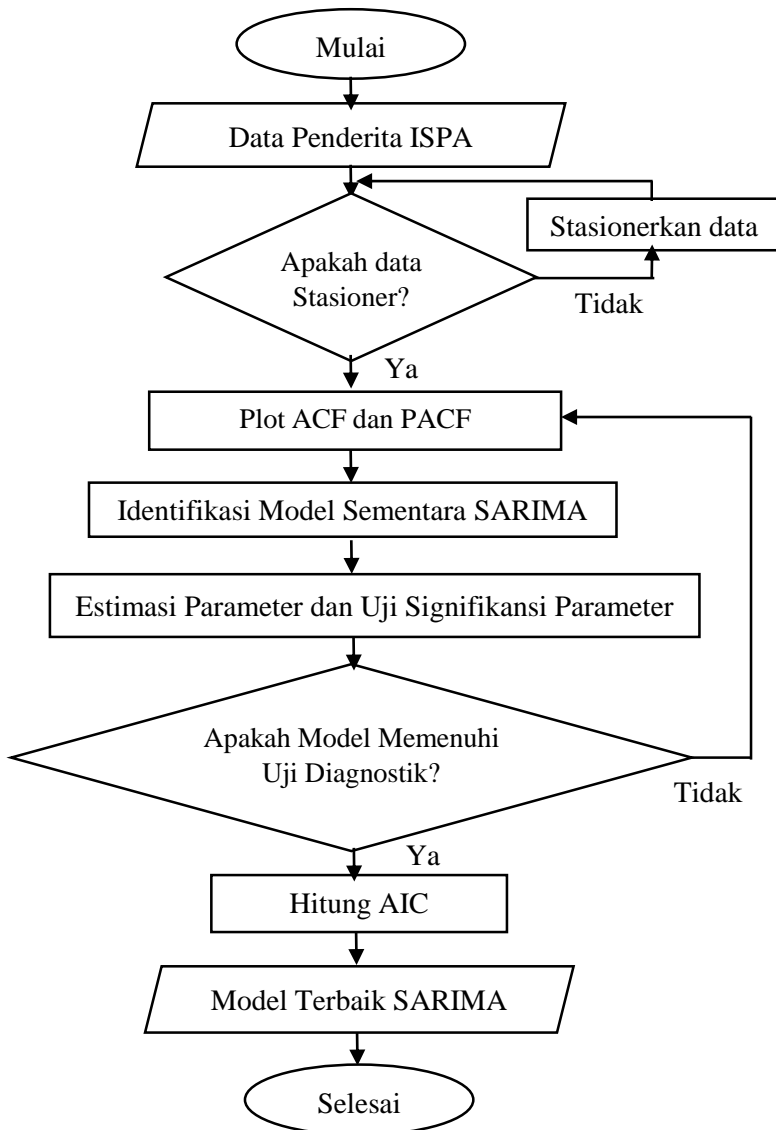
3.4 Diagram Alir

Tahap pengerjaan Tugas Akhir terhadap aplikasi model Binomial Negatif $GSARIMA$ pada peramalan jumlah penderita penyakit ISPA ini disajikan dalam bentuk diagram alir pada Gambar 3.1. Model Binomial Negatif $GSARIMA$ didapatkan dari orde model $SARIMA$ terbaik dengan cara identifikasi model

SARIMA terlebih dahulu. Identifikasi model SARIMA ini disajikan pada Gambar 3.2.



Gambar 3.1 Diagram Alir Pengerjaan Tugas Akhir



Gambar 3.2 Diagram Alir Tahap Identifikasi Model SARIMA

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab ini dibahas mengenai aplikasi model Binomial Negatif GSARIMA pada peramalan jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kota Surabaya. Diawali dari analisis deskriptif data, tahap identifikasi model SARIMA, pembentukan model Binomial Negatif GSARIMA, estimasi parameter, dan kemudian tahap peramalan dengan model Binomial Negatif GSARIMA.

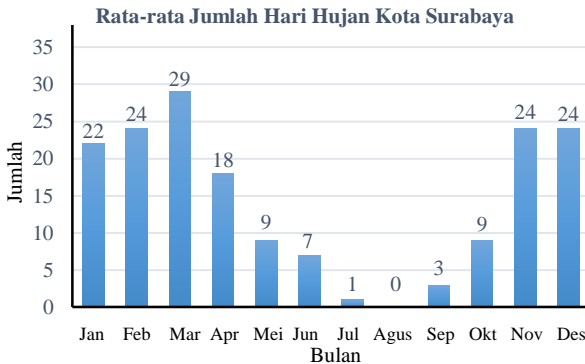
4.1 Analisis Deskriptif

Analisis deskriptif jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kota Surabaya dijelaskan menggunakan statistik deskriptif dan plot *time series*. Statistik deskriptif meliputi rata-rata (*mean*), standar deviasi, varian, nilai minimum, dan nilai maksimum. Hasil statistik deskriptif ditunjukkan pada Tabel 4.1.

Berdasarkan Tabel 4.1 nilai rata-rata jumlah penderita penyakit ISPA di Kota Surabaya selama 72 bulan terakhir adalah 348,76 yang berarti bahwa rata-rata jumlah penderita penyakit ISPA setiap bulan di Kota Surabaya sebanyak 348 orang. Rata-rata jumlah penderita penyakit ISPA tertinggi terjadi pada bulan Agustus yaitu sebesar 401,00 dengan jumlah penderita minimum sebanyak 241 dan jumlah penderita maksimum sebanyak 645. Sedangkan rata-rata jumlah penderita penyakit ISPA terendah terjadi pada bulan Januari yaitu sebesar 253,17 dengan jumlah penderita minimum sebanyak 148 dan jumlah penderita maksimum sebanyak 459. Selain itu, terlihat bahwa rata-rata jumlah penderita penyakit ISPA dari bulan Januari sampai bulan Maret cenderung mengalami kenaikan, kemudian pada bulan Maret sampai bulan Juni cenderung menurun. Pada bulan Juni sampai bulan Agustus rata-rata jumlah penderita penyakit ISPA cenderung mengalami kenaikan kembali, kemudian pada bulan Agustus sampai bulan Desember cenderung mengalami penurunan.

Tabel 4.1 Statistik Deskriptif Data Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kota Surabaya Januari 2014 - Desember 2019

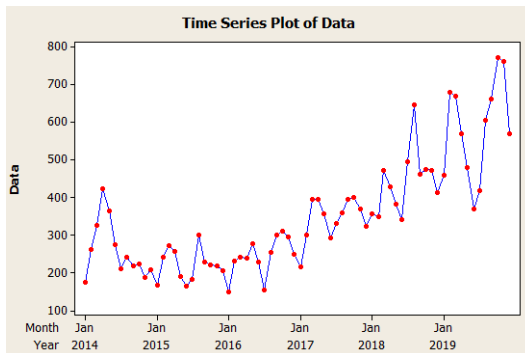
| Bulan | Mean | Standar Deviasi | Varian | Min | Max |
|-----------|--------|-----------------|---------|-----|-----|
| Januari | 253,17 | 125,830 | 15833,4 | 148 | 459 |
| Februari | 343,67 | 169,795 | 28830,3 | 232 | 679 |
| Maret | 396,00 | 158,156 | 25013,2 | 240 | 670 |
| April | 384,50 | 122,525 | 15012,3 | 238 | 569 |
| Mei | 341,50 | 98,393 | 9681,1 | 190 | 479 |
| Juni | 278,00 | 74,305 | 5521,2 | 165 | 369 |
| Juli | 298,50 | 138,201 | 19099,5 | 155 | 494 |
| Agustus | 401,00 | 179,250 | 32130,4 | 241 | 645 |
| September | 377,00 | 168,464 | 28380,0 | 217 | 661 |
| Oktober | 400,00 | 207,065 | 42876,0 | 222 | 771 |
| November | 384,00 | 212,253 | 45051,2 | 188 | 762 |
| Desember | 327,83 | 142,385 | 20273,4 | 205 | 569 |
| Total | 348,76 | 150,448 | 22634,5 | 148 | 771 |



Gambar 4.1 Diagram Rata-rata Jumlah Hari Hujan Kota Surabaya 2015-2019

Berdasarkan Gambar 4.1 rata-rata jumlah hari hujan Kota Surabaya pada tahun 2015-2019 tertinggi sebesar 29 terjadi pada bulan Maret. Sedangkan rata-rata jumlah hari hujan terendah sebesar 0 terjadi pada bulan Agustus. Hal ini menunjukkan bahwa rata-rata jumlah penderita penyakit ISPA memiliki pola yang berkaitan dengan rata-rata jumlah hari hujan di Kota Surabaya. Ketika rata-rata jumlah hari hujan sangat tinggi, rata-rata jumlah penderita penyakit ISPA juga tinggi. Demikian juga ketika rata-rata jumlah hari hujan sangat rendah, rata-rata jumlah penderita penyakit ISPA juga menunjukkan angka yang tinggi.

Hal ini sesuai dengan Wantania, dkk (2012) yang menyatakan bahwa salah satu faktor yang menyebabkan terjadinya penyakit ISPA adalah keadaan lingkungan. Polutan lingkungan dapat mengiritasi saluran pernapasan sehingga memudahkan terjadinya infeksi. Dimana pada saat musim kemarau kondisi udara sangat kering, sehingga meningkatkan polutan di udara yang menjadi salah satu faktor penyebab penyakit ISPA. Selain itu, curah hujan yang berlebihan menyebabkan kondisi lingkungan menjadi lembab, sehingga dapat juga meningkatkan risiko penyakit ISPA.



Gambar 4.2 Plot *Time Series* Data Jumlah Penderita Penyakit ISPA

Peningkatan dan penurunan pada data jumlah penderita penyakit ISPA juga terlihat dari plot *time series* pada Gambar 4.2,

dimana plot *time series* data terlihat naik turun pada waktu tertentu. Sehingga mengindikasikan adanya faktor musiman pada data jumlah penderita penyakit ISPA.

4.2 Tahap Identifikasi Model SARIMA

Tahap pembentukan model SARIMA dilakukan untuk memperoleh orde model GSARIMA pada data jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA) berdasarkan *Box-Jenkins*. Berikut hasil pemodelan pada data jumlah penderita ISPA di Kota Surabaya dengan menggunakan model ARIMA *Box-Jenkins*:

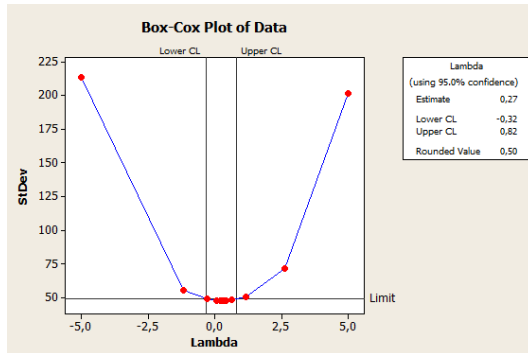
4.2.1 Tahap Identifikasi

Tahapan identifikasi diawali dengan membuat plot *time series*, plot ACF (Autokorelasi), dan plot PACF (Autokorelasi Parsial). Proses identifikasi ini dilakukan untuk memastikan apakah data jumlah penderita penyakit ISPA telah memenuhi asumsi stasioner, baik stasioneritas terhadap varian dan stasioneritas terhadap *mean* serta untuk mengidentifikasi model SARIMA(p,d,q)(P,D,Q)^s yang sesuai dengan plot ACF dan PACF.

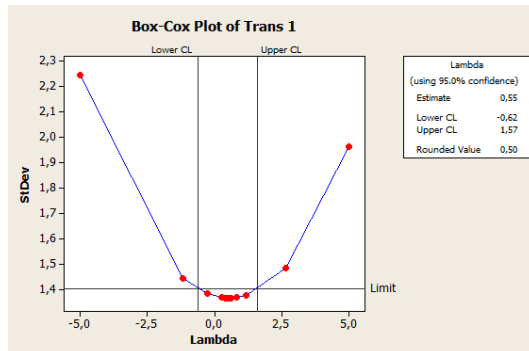
Data dikatakan stasioner apabila plot *time series* berfluktuasi di sekitar garis yang sejajar dengan sumbu waktu (t) atau nilai *mean* dan varian tidak ada perubahan secara signifikan dalam periode waktu tertentu. Berdasarkan Gambar 4.2 terdapat unsur *trend* pada plot *time series*. Selain itu, terlihat bahwa titik satu dengan titik lainnya sangat bervariasi dan tidak sejajar dengan sumbu waktu. Sehingga dapat dikatakan data belum stasioner dalam varian dan *mean*.

Identifikasi pertama adalah untuk mengetahui stasioneritas data dalam varian yang dapat diidentifikasi melalui *Box-Cox plot*. Pada *Box-Cox plot*, data dikatakan telah stasioner terhadap varian jika telah memenuhi syarat bahwa *rounded value* atau nilai taksiran lamda sama dengan satu. Berdasarkan Gambar 4.3 nilai *rounded value* atau nilai pendugaan lamda (λ) pada *Box-Cox Plot* sebesar 0,50, maka dapat dikatakan data belum stasioner dalam varian,

sehingga perlu dilakukan transformasi. Transformasi yang akan dilakukan adalah $Y^{0,5}$, artinya data awal jumlah penderita penyakit ISPA dipangkatkan dengan 0,5.



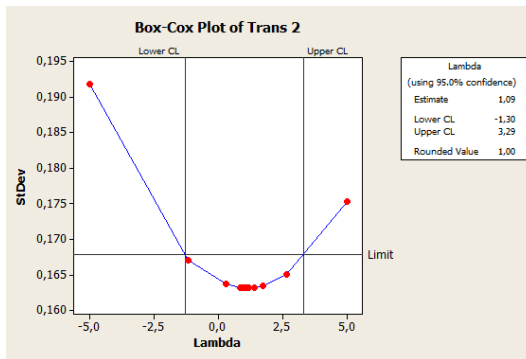
Gambar 4.3 *Box-Cox Plot* Data Jumlah Penderita Penyakit ISPA



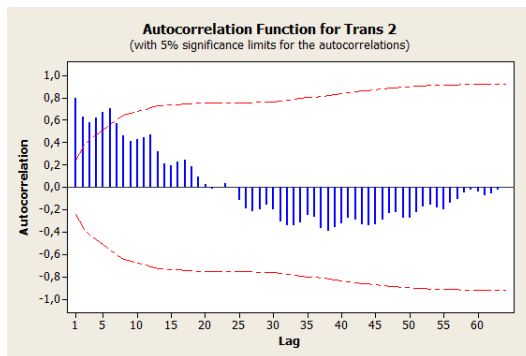
Gambar 4.4 Transformasi 1 *Box-Cox Plot* Data Jumlah Penderita Penyakit ISPA

Hasil *Box-Cox plot* transformasi 1 yaitu pada Gambar 4.4 didapatkan nilai *rounded value* sebesar 0,50, atau dapat dikatakan data belum stasioner dalam varian. Sehingga harus dilakukan transformasi kembali. Sehingga transformasi yang akan dilakukan adalah $(Y^{0,5})^{0,5}$.

Berdasarkan Gambar 4.5 nilai *rounded value* pada hasil *Box-Cox Plot* transformasi 2 sebesar 1,00, atau dapat dikatakan data sudah stasioner dalam varian. Sehingga tidak perlu dilakukan transformasi kembali.



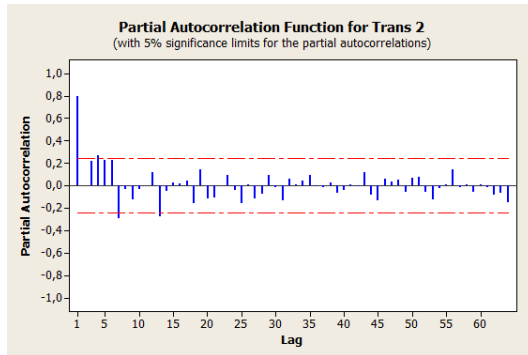
Gambar 4.5 Transformasi 2 *Box-Cox Plot* Data Jumlah Penderita Penyakit ISPA



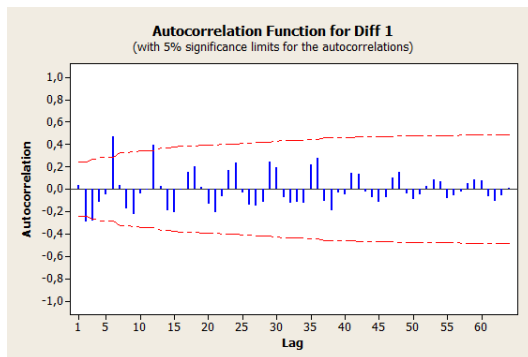
Gambar 4.6 Plot ACF Hasil Transformasi 2 *Box-Cox Plot*

Setelah data dinyatakan stasioner dalam varian, selanjutnya dilakukan pengecekan stasioneritas dalam *mean* dengan melihat plot ACF dari hasil transformasi 2 *Box-Cox* terhadap data jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA). Perhitungan nilai ACF, PACF beserta

selang kepercayaan atau garis signifikan pada plot ACF dan PACF dari data hasil transformasi 2 *Box-Cox* dapat dilihat pada Lampiran C. Sedangkan Plot ACF dan PACF dari data hasil transformasi 2 *Box-Cox* dapat dilihat pada Gambar 4.6 dan Gambar 4.7.



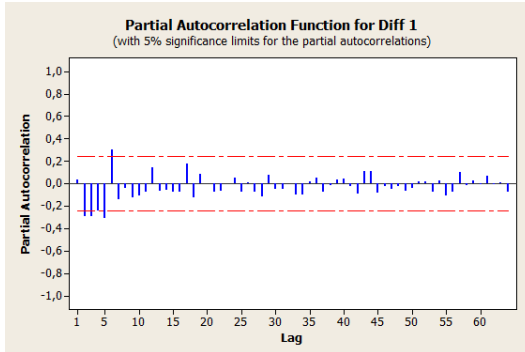
Gambar 4.7 Plot PACF Hasil Transformasi 2 *Box-Cox* Plot



Gambar 4.8 Plot ACF Hasil *Differencing Lag* Nonmusiman

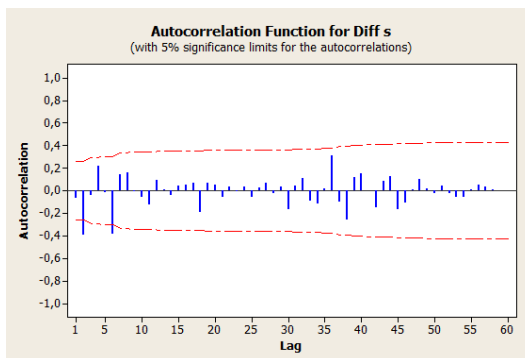
Pada Gambar 4.6 diketahui bahwa plot ACF turun menuju 0 secara perlahan, dan dimana nilai ACF pada tiga *lag* pertama melebihi batas garis signifikan, atau dapat dikatakan data belum stasioner dalam *mean*. Sehingga perlu dilakukan *differencing* pada

data hasil transformasi 2 *Box Cox*. Plot ACF dan PACF hasil *differencing* ditunjukkan pada Gambar 4.8 dan Gambar 4.9.

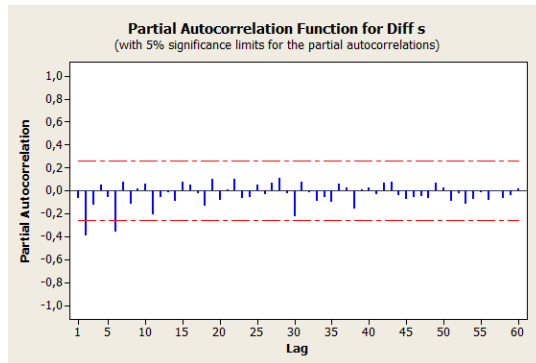


Gambar 4.9 Plot PACF Hasil *Differencing Lag* Nonmusiman

Berdasarkan Gambar 4.8 terlihat bahwa plot ACF pada *lag* 6, 12 melebihi batas garis signifikan, maka dapat dikatakan terdapat orde musiman pada *lag* 6. Untuk menghilangkan unsur musiman dapat dilakukan dengan adanya *differencing* pada *lag* musimannya yaitu *lag* 6. Plot ACF dan plot PACF hasil *differencing* pada *lag* musiman ditunjukkan pada Gambar 4.10 dan Gambar 4.11



Gambar 4.10 Plot ACF Hasil *Differencing* pada *Lag* Musiman



Gambar 4.11 Plot PACF Hasil *Differencing* pada *Lag* Musiman

Plot ACF digunakan untuk menentukan orde q , sedangkan plot PACF digunakan untuk menentukan orde p , dan banyaknya *differencing* digunakan untuk menentukan orde d . Berdasarkan plot ACF pada Gambar 4.10 terlihat bahwa terdapat 2 *lag* yang signifikan, dengan 1 *lag* berada pada *lag* musiman yaitu *lag* 6. *Lag* yang signifikan adalah *lag* yang melebihi garis merah atau batas signifikan. Sedangkan berdasarkan plot PACF pada Gambar 4.11 terdapat 2 *lag* yang signifikan pula, dengan 1 *lag* berada pada *lag* musiman yaitu *lag* 6. Banyaknya *differencing* yang dilakukan pada *lag* nonmusiman dan *lag* musiman masing-masing sebanyak satu kali. Sehingga model dugaan sementara untuk data jumlah penderita penyakit ISPA di Kota Surabaya adalah model SARIMA(2,1,2)(1,1,1)⁶.

4.2.2 Tahap Estimasi dan Uji Signifikansi Parameter

Dari tahap identifikasi model SARIMA, didapatkan model dugaan sementara yaitu model SARIMA(2,1,2)(1,1,1)⁶. Kemudian tahap selanjutnya adalah tahap estimasi dan uji signifikansi parameter. Estimasi parameter model dugaan sementara SARIMA(2,1,2)(1,1,1)⁶ dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan menggunakan bantuan *software* Minitab16. Hasil estimasi parameter terhadap

model dugaan sementara SARIMA(2,1,2)(1,1,1)⁶ dapat dilihat pada Tabel 4.2.

Tabel 4.2 Estimasi Parameter Model SARIMA(2,1,2)(1,1,1)⁶

| Parameter | Koef | SE | T | P-value |
|-----------|---------|--------|-------|---------|
| AR 1 | 0,4354 | 0,4516 | 0,96 | 0,339 |
| AR 2 | -0,2470 | 0,3109 | -0,79 | 0,430 |
| SAR 1 | 0,0109 | 0,2147 | 0,05 | 0,960 |
| MA 1 | 0,7340 | 0,4678 | 1,57 | 0,123 |
| MA 2 | 0,1512 | 0,4690 | 0,32 | 0,748 |
| SMA 1 | 0,7621 | 0,1868 | 4,08 | 0,000 |

Setelah parameter pada model SARIMA(2,1,2)(1,1,1)⁶ diestimasi, kemudian tahap selanjutnya adalah pengujian signifikansi parameter. Uji signifikansi parameter dapat dilakukan dengan menggunakan uji-t dengan $\alpha = 5\%$. Uji signifikansi parameter model SARIMA dugaan sementara yaitu SARIMA(2,1,2)(1,1,1)⁶ adalah sebagai berikut:

1. Uji parameter ϕ_1 (AR 1)

Hipotesis:

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

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

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\phi}_1}{SE(\hat{\phi}_1)} \\ &= \frac{0,4354}{0,4516} \\ &= 0,9641 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| < t_{tabel}$ maka H_0 diterima, yang artinya parameter ϕ_1 tidak signifikan.

2. Uji parameter ϕ_2 (AR 2)

Hipotesis:

 $H_0 : \phi_2 = 0$ (parameter ϕ_2 tidak signifikan) $H_1 : \phi_2 \neq 0$ (parameter ϕ_2 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\phi}_2}{SE(\hat{\phi}_2)} \\ &= \frac{-0,2470}{0,3109} \\ &= -0,7945 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| < t_{tabel}$ maka H_0 diterima, yang artinya parameter ϕ_2 tidak signifikan.

3. Uji parameter θ_1 (MA 1)

Hipotesis:

 $H_0 : \theta_1 = 0$ (parameter θ_1 tidak signifikan) $H_1 : \theta_1 \neq 0$ (parameter θ_1 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\theta}_1}{SE(\hat{\theta}_1)} \\ &= \frac{0,7340}{0,4678} \\ &= 1,5690 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| < t_{tabel}$ maka H_0 diterima, yang artinya parameter θ_1 tidak signifikan.

4. Uji parameter θ_2 (MA 2)

Hipotesis:

 $H_0: \theta_2 = 0$ (parameter θ_2 tidak signifikan) $H_1: \theta_2 \neq 0$ (parameter θ_2 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\theta}_2}{SE(\hat{\theta}_2)} \\ &= \frac{0,1512}{0,4690} \\ &= 0,3224 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| < t_{tabel}$ maka H_0 diterima, yang artinya parameter θ_2 tidak signifikan.

5. Uji parameter Φ_1 (SAR 1)

Hipotesis:

 $H_0: \Phi_1 = 0$ (parameter Φ_1 tidak signifikan) $H_1: \Phi_1 \neq 0$ (parameter Φ_1 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\Phi}_1}{SE(\hat{\Phi}_1)} \\ &= \frac{0,0109}{0,2147} \\ &= 0,0508 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| < t_{tabel}$ maka H_0 diterima, yang artinya parameter Φ_1 tidak signifikan.

6. Uji parameter θ_1 (SMA 1)

Hipotesis:

 $H_0 : \theta_1 = 0$ (parameter θ_1 tidak signifikan) $H_1 : \theta_1 \neq 0$ (parameter θ_1 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\theta}_1}{SE(\hat{\theta}_1)} \\ &= \frac{0,7621}{0,1868} \\ &= 4,079 \end{aligned}$$

dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

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

Signifikansi parameter juga dapat dilihat dari nilai p -value masing-masing parameter. Parameter dikatakan signifikan apabila nilai p -value lebih kecil dari taraf signifikannya yaitu 5%. Karena terdapat parameter pada model SARIMA(2,1,2)(1,1,1)⁶ yang tidak signifikan, maka model SARIMA(2,1,2)(1,1,1)⁶ bukan merupakan model yang signifikan.

4.2.3 Uji Diagnostik

Kemudian langkah selanjutnya adalah melakukan uji diagnostik dari model SARIMA dugaan sementara. Model SARIMA dugaan sementara bukan merupakan model yang signifikan, sehingga diperlukan model lain yang memenuhi asumsi pada SARIMA. Sehingga langkah selanjutnya adalah melakukan *overfitting* pada model dugaan sementara. Pada tahap *overfitting* ini, dihasilkan model-model alternatif yang merupakan perluasan dari model SARIMA dugaan sementara. Model hasil *overfitting* ini kemudian diestimasi dan dilakukan uji signifikansi pada setiap

parameternya. Model hasil *overfitting* dan hasil estimasi serta uji signifikansi parameter dapat dilihat pada Lampiran D.

Model-model alternatif pada tahap *overfitting* yang memenuhi uji signifikansi parameter adalah sebagai berikut:

1. SARIMA(2,1,1)(0,1,1)⁶
2. SARIMA(1,1,1)(1,1,0)⁶
3. SARIMA(1,1,1)(0,1,1)⁶
4. SARIMA(1,1,1)(0,1,0)⁶
5. SARIMA(0,1,2)(1,1,0)⁶
6. SARIMA(0,1,2)(0,1,1)⁶
7. SARIMA(0,1,2)(0,1,0)⁶
8. SARIMA(0,1,1)(0,1,0)⁶
9. SARIMA(0,1,0)(1,1,0)⁶
10. SARIMA(0,1,0)(0,1,1)⁶

Kemudian langkah selanjutnya adalah melakukan beberapa uji untuk mengetahui apakah model SARIMA yang terbentuk pada tahap *overfitting* memenuhi asumsi *white noise* dan berdistribusi normal pada residualnya.

Tabel 4.3 Hasil Uji *White Noise* pada Model SARIMA Hasil *overfitting* yang Telah Memenuhi Uji Signifikansi Parameter

| Model | Hasil Uji White Noise |
|-----------------------------------|--------------------------|
| SARIMA(2,1,1)(0,1,1) ⁶ | <i>White noise</i> |
| SARIMA(1,1,1)(1,1,0) ⁶ | <i>White noise</i> |
| SARIMA(1,1,1)(0,1,1) ⁶ | <i>White noise</i> |
| SARIMA(1,1,1)(0,1,0) ⁶ | Tidak <i>white noise</i> |
| SARIMA(0,1,2)(1,1,0) ⁶ | <i>White noise</i> |
| SARIMA(0,1,2)(0,1,1) ⁶ | <i>White noise</i> |
| SARIMA(0,1,2)(0,1,0) ⁶ | <i>White noise</i> |
| SARIMA(0,1,1)(0,1,0) ⁶ | Tidak <i>white noise</i> |
| SARIMA(0,1,0)(1,1,0) ⁶ | <i>White noise</i> |
| SARIMA(0,1,0)(0,1,1) ⁶ | <i>White noise</i> |

Asumsi residual *white noise* artinya tidak terdapat korelasi antar residualnya, dimana variansi bernilai konstan dan nilai rata-

rata sama dengan nol. Uji residual *white noise* dapat dilakukan dengan uji *Ljung-Box*. Perhitungan uji *Ljung-Box* terhadap model SARIMA yang dihasilkan pada tahap *overfitting* dan telah memenuhi asumsi signifikansi parameter dapat dilihat pada Lampiran E. Hasil uji *white noise* disajikan pada Tabel 4.3.

Berdasarkan Tabel 4.3, didapatkan beberapa model yang memenuhi asumsi residual *white noise*. Kemudian model yang telah memenuhi asumsi residual *white noise*, tahap selanjutnya adalah dilakukan uji residual berdistribusi normal. Uji normalitas pada residual dapat dihitung menggunakan uji *Kolmogorov-Smirnov*. Perhitungan uji *Kolmogorov-Smirnov* terhadap model yang telah memenuhi asumsi *white noise* dapat dilihat pada Lampiran F. Model-model yang memenuhi uji normalitas residual disajikan pada Tabel 4.4.

Tabel 4.4 Hasil Uji *Kolmogorov-Smirnov* pada Model SARIMA Hasil *Overfitting* yang Telah Memenuhi Asumsi *White Noise*

| Model | Hasil Uji <i>Kolmogorov-Smirnov</i> |
|-----------------------------------|-------------------------------------|
| SARIMA(2,1,1)(0,1,1) ⁶ | Normal |
| SARIMA(1,1,1)(1,1,0) ⁶ | Normal |
| SARIMA(1,1,1)(0,1,1) ⁶ | Normal |
| SARIMA(0,1,2)(1,1,0) ⁶ | Normal |
| SARIMA(0,1,2)(0,1,1) ⁶ | Tidak Normal |
| SARIMA(0,1,2)(0,1,0) ⁶ | Normal |
| SARIMA(0,1,0)(1,1,0) ⁶ | Normal |
| SARIMA(0,1,0)(0,1,1) ⁶ | Normal |

4.2.4 Pemilihan Model Terbaik

Model SARIMA hasil *overfitting* yang telah memenuhi uji diagnostik kemudian akan dipilih model terbaik. Model SARIMA terbaik dipilih berdasarkan kriteria nilai AIC, dimana akan dipilih model dengan nilai AIC terkecil. Hasil perhitungan dengan kriteria nilai AIC pada model SARIMA yang telah memenuhi uji diagnostik dapat dilihat pada Tabel 4.5.

Berdasarkan Tabel 4.5 model terbaik SARIMA dengan nilai AIC terkecil untuk data jumlah penderita penyakit ISPA di Kota Surabaya adalah model SARIMA(2,1,1)(0,1,1)⁶ yang kemudian akan dijadikan orde pada model Binomial Negatif GSARIMA dengan nilai AIC sebesar 674,86.

Tabel 4.5 Kriteria Pemilihan Model Tebaik SARIMA

| Model | AIC |
|-----------------------------------|--------|
| SARIMA(2,1,1)(0,1,1) ⁶ | 674,86 |
| SARIMA(1,1,1)(1,1,0) ⁶ | 678,62 |
| SARIMA(1,1,1)(0,1,1) ⁶ | 677,29 |
| SARIMA(0,1,2)(1,1,0) ⁶ | 676,12 |
| SARIMA(0,1,2)(0,1,0) ⁶ | 682,60 |
| SARIMA(0,1,0)(1,1,0) ⁶ | 683,28 |
| SARIMA(0,1,0)(0,1,1) ⁶ | 684,36 |

4.3 Model Binomial Negatif GSARIMA

Orde model Binomial Negatif GSARIMA diperoleh berdasarkan identifikasi dari orde model SARIMA yang didapatkan sebelumnya. Sehingga model yang terbentuk adalah model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶, yaitu dengan $p = 2, P = 0, q = 1, Q = 1, d = 1, D = 1, S = 6$. Sehingga model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dengan transformasi ZQ1 pada persamaan (2.6) menjadi:

$$\begin{aligned}
 \ln(\mu_t) = & (1 - B)^{-1}(1 - B^6)^{-1}\phi_2(B)\Phi_0(B^6)\{X_t^T\beta - \ln(y'_t)\} \\
 & + \ln(y'_t) - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y'_t}{\mu_t}\right) \\
 & + \ln\left(\frac{y'_t}{\mu_t}\right) \tag{4.1}
 \end{aligned}$$

Pada penelitian ini, variabel prediktor X diabaikan, menurut Marinho G. Andrade dan Ricardo(2015) $X_t^T\beta$ dapat diganti suatu konstanta a , sehingga Persamaan (4.1) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= (1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6)\{a - \ln(y'_t)\} + \ln(y'_t) \\
&\quad - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y'_t}{\mu_t}\right) \\
&\quad + \ln\left(\frac{y'_t}{\mu_t}\right)
\end{aligned} \tag{4.2}$$

jika $\beta_0 = ((1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6))a$, maka Persamaan (4.2) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= \beta_0 + (1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6)\{-\ln(y'_t)\} \\
&\quad + \ln(y'_t) - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y'_t}{\mu_t}\right) \\
&\quad + \ln\left(\frac{y'_t}{\mu_t}\right)
\end{aligned} \tag{4.3}$$

dengan:

$$\begin{aligned}
\phi_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\
\Phi_p(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps} \\
\theta_q(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\
\Theta_q(B^s) &= 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_p B^{qs} \\
B^d y_t &= y_{t-d}
\end{aligned}$$

sehingga persamaan (4.3) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= \beta_0 + (1 - B)(1 - B^6)(1 - \phi_1 B \\
&\quad - \phi_2 B^2)(1)\{-\ln(y'_t)\} + \ln(y'_t) \\
&\quad - (1 - \theta_1 B)(1 - \theta_1 B^6)\ln\left(\frac{y'_t}{\mu_t}\right) + \ln\left(\frac{y'_t}{\mu_t}\right) \\
\ln(\mu_t) &= \beta_0 + (1 - B - B^6 + B^7 - \phi_1 B + \phi_1 B^2 + \phi_1 B^7 - \phi_1 B^8 \\
&\quad - \phi_2 B^2 + \phi_2 B^3 + \phi_2 B^8 \\
&\quad - \phi_2 B^9)\{-\ln(y'_t)\} + \ln(y'_t) \\
&\quad - (1 - \theta_1 B - \theta_1 B^6 + \theta_1 \theta_1 B^7)\ln\left(\frac{y'_t}{\mu_t}\right) \\
&\quad + \ln\left(\frac{y'_t}{\mu_t}\right)
\end{aligned}$$

$$\begin{aligned}
\ln(\mu_t) &= \beta_0 - \ln(y'_t) + \ln(y'_{t-1}) + \ln(y'_{t-6}) - \ln(y'_{t-7}) \\
&\quad + \phi_1 \ln(y'_{t-1}) - \phi_1 \ln(y'_{t-2}) - \phi_1 \ln(y'_{t-7}) \\
&\quad + \phi_1 \ln(y'_{t-8}) + \phi_2 \ln(y'_{t-2}) - \phi_2 \ln(y'_{t-3}) \\
&\quad - \phi_2 \ln(y'_{t-8}) + \phi_2 \ln(y'_{t-9}) + \ln(y'_t) - \ln\left(\frac{y'_t}{\mu_t}\right) \\
&\quad + \theta_1 \ln\left(\frac{y'_{t-1}}{\mu_{t-1}}\right) + \theta_1 \ln\left(\frac{y'_{t-6}}{\mu_{t-6}}\right) - \theta_1 \theta_1 \ln\left(\frac{y'_{t-7}}{\mu_{t-7}}\right) \\
&\quad + \ln\left(\frac{y'_t}{\mu_t}\right) \\
\ln(\mu_t) &= \beta_0 + \ln(y'_{t-1}) + \ln(y'_{t-6}) - \ln(y'_{t-7}) + \phi_1 \ln(y'_{t-1}) \\
&\quad - \phi_1 \ln(y'_{t-2}) - \phi_1 \ln(y'_{t-7}) + \phi_1 \ln(y'_{t-8}) \\
&\quad + \phi_2 \ln(y'_{t-2}) - \phi_2 \ln(y'_{t-3}) - \phi_2 \ln(y'_{t-8}) \\
&\quad + \phi_2 \ln(y'_{t-9}) + \theta_1 \ln\left(\frac{y'_{t-1}}{\mu_{t-1}}\right) + \theta_1 \ln\left(\frac{y'_{t-6}}{\mu_{t-6}}\right) \\
&\quad - \theta_1 \theta_1 \ln\left(\frac{y'_{t-7}}{\mu_{t-7}}\right) \tag{4.4}
\end{aligned}$$

Persamaan (4.4) dapat juga ditulis sebagai berikut:

$$\begin{aligned}
\mu_t &= \exp \left\{ \beta_0 + \ln(y'_{t-1}) + \ln(y'_{t-6}) \right. \\
&\quad - \ln(y'_{t-7}) + \phi_1 \ln(y'_{t-1}) - \phi_1 \ln(y'_{t-2}) \\
&\quad - \phi_1 \ln(y'_{t-7}) + \phi_1 \ln(y'_{t-8}) \\
&\quad + \phi_2 \ln(y'_{t-2}) - \phi_2 \ln(y'_{t-3}) - \phi_2 \ln(y'_{t-8}) \\
&\quad + \phi_2 \ln(y'_{t-9}) + \theta_1 \ln\left(\frac{y'_{t-1}}{\mu_{t-1}}\right) + \theta_1 \ln\left(\frac{y'_{t-6}}{\mu_{t-6}}\right) \\
&\quad \left. - \theta_1 \theta_1 \ln\left(\frac{y'_{t-7}}{\mu_{t-7}}\right) \right\} \tag{4.5}
\end{aligned}$$

dengan $y'_{t-j} = \max(y_{t-j}, c)$ dan $0 < c \leq 1$.

Sedangkan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dengan transformasi ZQ2 pada Persamaan (2.7) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= (1 - B)^1(1 - B^6)^1\phi_2(B)\Phi_0(B^6)\{\ln[\exp(\mathbf{X}_t^T\boldsymbol{\beta}) + c] \\
&\quad - \ln(y_t + c)\} + \ln(y_t + c) \\
&\quad - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\
&\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)
\end{aligned} \tag{4.6}$$

Pada penelitian ini, variabel prediktor X diabaikan, menurut Marinho G. Andrade dan Ricardo(2015) $\mathbf{X}_t^T\boldsymbol{\beta}$ dapat diganti dengan suatu konstanta a , sehingga Persamaan (4.6) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= (1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6)\{\ln[\exp(a) + c] - \ln(y_t \\
&\quad + c)\} + \ln(y_t + c) - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\
&\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)
\end{aligned} \tag{4.7}$$

Jika $\beta_0 = ((1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6))\{\ln[\exp(a) + c]\}$ maka Persamaan (4.7) menjadi:

$$\begin{aligned}
\ln(\mu_t) &= \beta_0 + (1 - B)(1 - B^6)\phi_2(B)\Phi_0(B^6)\{-\ln(y_t + c)\} \\
&\quad + \ln(y_t + c) - \theta_1(B)\theta_1(B^6)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\
&\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)
\end{aligned} \tag{4.8}$$

dengan:

$$\begin{aligned}
\phi_p(B) &= 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p \\
\Phi_p(B^S) &= 1 - \Phi_1B^S - \Phi_2B^{2S} - \dots - \Phi_pB^{pS} \\
\theta_q(B) &= 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q \\
\Theta_q(B^S) &= 1 - \Theta_1B^S - \Theta_2B^{2S} - \dots - \Theta_pB^{qS} \\
B^d y_t &= y_{t-d}
\end{aligned}$$

sehingga persamaan (4.8) menjadi:

$$\begin{aligned}\ln(\mu_t) &= \beta_0 + (1 - B)(1 - B^6)(1 - \phi_1 B \\ &\quad - \phi_2 B^2)(1)\{-\ln(y_t + c)\} + \ln(y_t + c) \\ &\quad - (1 - \theta_1 B)(1 - \theta_1 B^6)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\ &\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)\end{aligned}$$

$$\begin{aligned}\ln(\mu_t) &= \beta_0 + (1 - B - B^6 \\ &\quad + B^7)(1 - \phi_1 B - \phi_2 B^2)\{-\ln(y_t + c)\} \\ &\quad + \ln(y_t + c) \\ &\quad - (1 - \theta_1 B)(1 - \theta_1 B^6)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\ &\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)\end{aligned}$$

$$\begin{aligned}\ln(\mu_t) &= \beta_0 + (1 - B - B^6 + B^7 - \phi_1 B + \phi_1 B^2 + \phi_1 B^7 - \phi_1 B^8 \\ &\quad - \phi_2 B^2 + \phi_2 B^3 + \phi_2 B^8 \\ &\quad - \phi_2 B^9)\{-\ln(y_t + c)\} + \ln(y_t + c) \\ &\quad - (1 - \theta_1 B - \theta_1 B^6 + \theta_1 \theta_1 B^7)\ln\left(\frac{y_t + c}{\mu_t + c}\right) \\ &\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)\end{aligned}$$

$$\begin{aligned}\ln(\mu_t) &= \beta_0 - \ln(y_t + c) + \ln(y_{t-1} + c) + \ln(y_{t-6} + c) \\ &\quad - \ln(y_{t-7} + c) + \phi_1 \ln(y_{t-1} + c) \\ &\quad - \phi_1 \ln(y_{t-2} + c) - \phi_1 \ln(y_{t-7} + c) \\ &\quad + \phi_1 \ln(y_{t-8} + c) \\ &\quad + \phi_2 \ln(y_{t-2} + c) - \phi_2 \ln(y_{t-3} + c) \\ &\quad - \phi_2 \ln(y_{t-8} + c) + \phi_2 \ln(y_{t-9} + c) \\ &\quad + \ln(y_t + c) - \ln\left(\frac{y_t + c}{\mu_t + c}\right) + \theta_1 \ln\left(\frac{y_{t-1} + c}{\mu_{t-1} + c}\right) \\ &\quad + \theta_1 \ln\left(\frac{y_{t-6} + c}{\mu_{t-6} + c}\right) - \theta_1 \theta_1 \ln\left(\frac{y_{t-7} + c}{\mu_{t-7} + c}\right) \\ &\quad + \ln\left(\frac{y_t + c}{\mu_t + c}\right)\end{aligned}$$

$$\begin{aligned}
\ln(\mu_t) = & \beta_0 + \ln(y_{t-1} + c) + \ln(y_{t-6} + c) - \ln(y_{t-7} + c) \\
& + \phi_1 \ln(y_{t-1} + c) \\
& - \phi_1 \ln(y_{t-2} + c) - \phi_1 \ln(y_{t-7} + c) \\
& + \phi_1 \ln(y_{t-8} + c) + \phi_2 \ln(y_{t-2} + c) \\
& - \phi_2 \ln(y_{t-3} + c) - \phi_2 \ln(y_{t-8} + c) \\
& + \phi_2 \ln(y_{t-9} + c) + \theta_1 \ln\left(\frac{y_{t-1} + c}{\mu_{t-1} + c}\right) \\
& + \theta_1 \ln\left(\frac{y_{t-6} + c}{\mu_{t-6} + c}\right) - \theta_1 \theta_1 \ln\left(\frac{y_{t-7} + c}{\mu_{t-7} + c}\right) \quad (4.9)
\end{aligned}$$

Persamaan (4.9) dapat juga ditulis sebagai berikut:

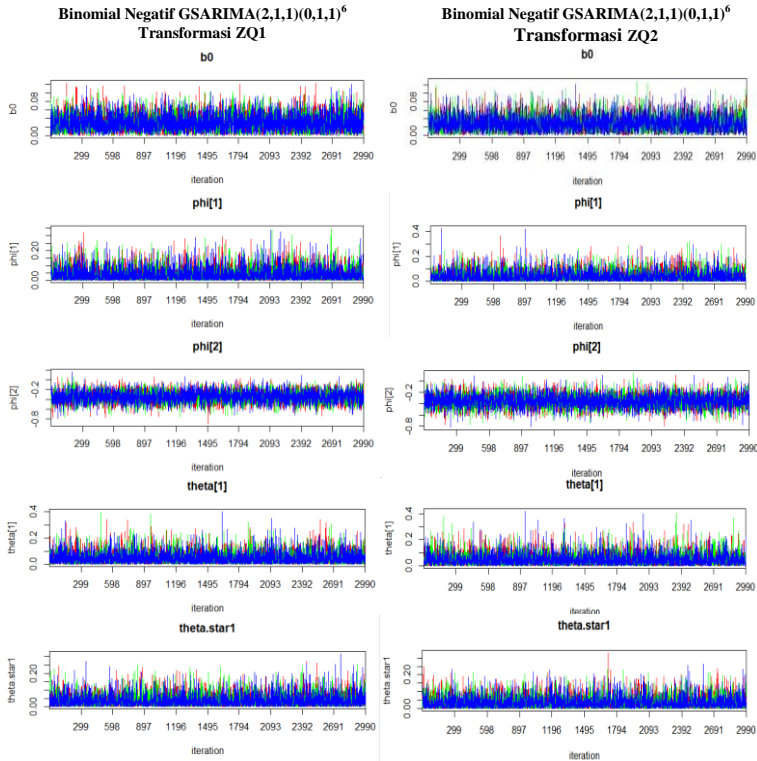
$$\begin{aligned}
\mu_t = \exp \left\{ & \beta_0 + \ln(y_{t-1} + c) + \ln(y_{t-6} + c) - \ln(y_{t-7} + c) \right. \\
& + \phi_1 \ln(y_{t-1} + c) \\
& - \phi_1 \ln(y_{t-2} + c) - \phi_1 \ln(y_{t-7} + c) \\
& + \phi_1 \ln(y_{t-8} + c) + \phi_2 \ln(y_{t-2} + c) \\
& - \phi_2 \ln(y_{t-3} + c) - \phi_2 \ln(y_{t-8} + c) \\
& + \phi_2 \ln(y_{t-9} + c) + \theta_1 \ln\left(\frac{y_{t-1} + c}{\mu_{t-1} + c}\right) \\
& \left. + \theta_1 \ln\left(\frac{y_{t-6} + c}{\mu_{t-6} + c}\right) - \theta_1 \theta_1 \ln\left(\frac{y_{t-7} + c}{\mu_{t-7} + c}\right) \right\} \quad (4.10)
\end{aligned}$$

Dengan demikian model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 adalah Persamaan (4.5). Sedangkan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ2 adalah Persamaan (4.10).

4.4 Estimasi Parameter Model Binomial Negatif GSARIMA

Estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ didapatkan dengan menggunakan inferensia Bayesian dengan metode *Markov Chain Monte Carlo* (MCMC) dengan bantuan *R software*. *Syntax* dan hasil estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶

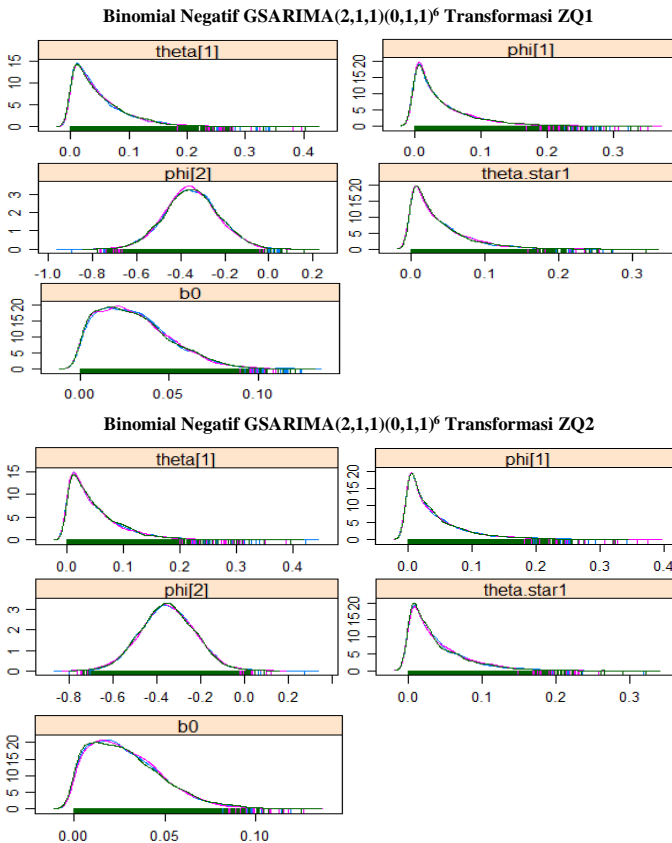
dapat dilihat pada Lampiran G dan Lampiran H. Pada penelitian ini digunakan algoritma MCMC dengan iterasi yang digunakan sebanyak 300.000 sehingga didapatkan nilai estimasi parameter model yang konvergen.



Gambar 4.12 Trace Plot Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶

Konvergenitas pada estimasi parameter dapat dilihat melalui *trace plot*. Berdasarkan *trace plot* pada Gambar 4.12, estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ telah memenuhi syarat konvergen yaitu nilainya memusat ke satu titik. Konvergenitas dapat juga dilihat

dari *density plot* estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ pada Gambar 4.13. Berdasarkan Gambar 4.13 terlihat bahwa kurva terpusat ke satu titik atau ke satu puncak, dengan kata lain tidak ditemukan kurva estimasi parameter yang memiliki dua puncak atau lebih. Sehingga dapat dikatakan nilainya telah konvergen. Titik puncak tersebut merupakan nilai estimasi parameter yang didapatkan dengan metode Bayesian.



Gambar 4.13 *Density Plot* Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶

Terdapat dua transformasi fungsi *link* pada estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶, yaitu transformasi fungsi *link* ZQ1 dan fungsi *link* ZQ2. Sehingga terdapat dua hasil estimasi pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶. Hasil estimasi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dapat dilihat pada Tabel 4.6 dan Tabel 4.7

Tabel 4.6 Hasil Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ Transformasi ZQ1

| Parameter | Estimasi | St.Dev | SE | 2,5% | 97,5% |
|--------------------|----------|--------|--------|--------|--------|
| β_0 | 0,030 | 0,020 | 0,0004 | 0,002 | 0,077 |
| AR1 (ϕ_1) | 0,038 | 0,042 | 0,0014 | 0,000 | 0,147 |
| AR2 (ϕ_2) | -0,351 | 0,124 | 0,0021 | -0,593 | -0,106 |
| MA1(θ_1) | 0,048 | 0,046 | 0,0012 | 0,001 | 0,169 |
| SMA1(θ_1) | 0,038 | 0,040 | 0,0012 | 0,001 | 0,148 |

Tabel 4.7 Hasil Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ Transformasi ZQ2

| Parameter | Estimasi | St.Dev | SE | 2,5% | 97,5% |
|--------------------|----------|--------|--------|--------|--------|
| β_0 | 0,029 | 0,020 | 0,0005 | 0,002 | 0,076 |
| AR1 (ϕ_1) | 0,041 | 0,046 | 0,0017 | 0,000 | 0,163 |
| AR2 (ϕ_2) | -0,349 | 0,123 | 0,0021 | -0,588 | -0,113 |
| MA1(θ_1) | 0,051 | 0,050 | 0,0015 | 0,001 | 0,182 |
| SMA1(θ_1) | 0,037 | 0,038 | 0,0012 | 0,000 | 0,140 |

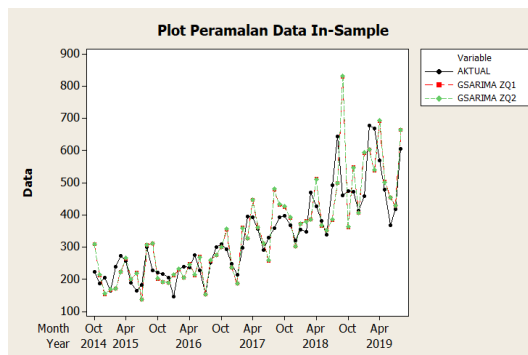
Signifikansi parameter model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dapat dilihat dari *confident interval* pada nilai 2,5% sampai dengan 97,5%. Pada Tabel 4.6 dan Tabel 4.7 menunjukkan bahwa setiap parameter pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ baik transformasi ZQ1 maupun transformasi ZQ2 telah memenuhi signifikansi parameter. Hal ini ditunjukkan dari *confident interval* pada titik 2,5% sampai titik 97,5% bernilai negatif pada parameter ϕ_2 dan bernilai positif pada

parameter β_0 , ϕ_1 , θ_1 , dan Θ_1 . Dengan demikian nilai estimasi setiap parameter pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ tidak melewati angka 0, sehingga setiap parameternya telah signifikan.

4.5 Peramalan Model Binomial Negatif GSARIMA

Setelah dilakukan tahap estimasi parameter dan uji signifikansi parameter pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶, kemudian langkah selanjutnya adalah tahap peramalan. Peramalan dengan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dilakukan dengan menggunakan *R software* yang diaplikasikan pada data jumlah penderita penyakit ISPA di Kota Surabaya. Peramalan dilakukan dengan 68 data sebagai *in-sample* dan 4 data sebagai *out-sample*. Hasil peramalan dari model Binomial Negatif GSARIMA merupakan nilai median dari distribusi bersyarat *k-step* ke depan, sehingga didapatkan hasil peramalan yang merupakan data *count*.

Model dikatakan baik, apabila pada peramalan *in-sample* pola data hasil peramalan akan mengikuti pola data aktual. Plot hasil peramalan data *in-sample* dengan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ dapat dilihat pada Gambar 4.14.



Gambar 4.14 Plot Hasil Peramalan Data *In-Sample* dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶

Berdasarkan Gambar 4.14 hasil peramalan data *in-sample* jumlah penderita penyakit ISPA di Kota Surabaya dengan Model Binomial Negatif GSARIMA $(2,1,1)(0,1,1)^6$ transformasi ZQ1 dan ZQ2 yang masing-masing ditunjukkan dengan garis berwarna merah dan hijau telah menunjukkan pola yang sama dengan data aktual *in-sample* yang ditunjukkan dengan garis berwarna hitam. Sehingga model Binomial Negatif GSARIMA $(2,1,1)(0,1,1)^6$ dapat digunakan untuk meramalkan jumlah penderita penyakit ISPA di Kota Surabaya.

Kemudian langkah selanjutnya adalah melakukan peramalan pada data *out-sample*. Hasil peramalan pada data *out-sample* jumlah penderita penyakit ISPA di Kota Surabaya dengan model Binomial Negatif GSARIMA $(2,1,1)(0,1,1)^6$ disajikan pada Tabel 4.8.

Tabel 4.8 Hasil Peramalan pada Data *Out-Sample* dengan Model Binomial Negatif GSARIMA $(2,1,1)(0,1,1)^6$ Transformasi ZQ1 dan ZQ2

| No | Periode | Ramalan | |
|----|----------------|---------|-----|
| | | ZQ1 | ZQ2 |
| 1 | September 2019 | 605 | 604 |
| 2 | Oktober 2019 | 578 | 575 |
| 3 | November 2019 | 656 | 654 |
| 4 | Desember 2019 | 540 | 539 |

Hasil peramalan pada data *out-sample* yang telah didapatkan dengan model Binomial Negatif GSARIMA $(2,1,1)(0,1,1)^6$ transformasi ZQ1 dan ZQ2, akan dibandingkan tingkat akurasinya dengan data aktual yang ada. Selanjutnya untuk kedua model akan dipilih model terbaik yang digunakan sebagai model peramalan. Pemilihan model terbaik pada penelitian ini menggunakan kriteria *Mean Absolute Relative Error* (MARE). Model dikatakan baik jika memiliki nilai MARE terkecil. Perbandingan hasil peramalan pada data *out-sample*

dengan data aktual serta nilai MARE pada masing-masing model dapat dilihat pada Tabel 4.9.

Tabel 4.9 Perbandingan Nilai MARE Hasil Peramalan pada Data *Out-Sample* dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ Transformasi ZQ1 dan ZQ2.

| No | Y Aktual | \hat{Y} GSARIMA ZQ1 | \hat{Y} GSARIMA ZQ2 |
|----|----------|--------------------------|--------------------------|
| 1 | 661 | 605 | 604 |
| 2 | 771 | 578 | 575 |
| 3 | 762 | 656 | 654 |
| 4 | 569 | 540 | 539 |
| | MARE | 0,13109 | 0,13354 |

Pada Tabel 4.9 menunjukkan bahwa hasil peramalan dengan menggunakan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 lebih mendekati data aktual dari pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ2. Hal ini dapat dilihat juga dari nilai MARE pada masing-masing model, yang menunjukkan bahwa nilai MARE pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 relatif lebih kecil dari pada transformasi ZQ2 yaitu dengan nilai MARE sebesar 0,1311.

Besarnya nilai MARE pada masing-masing model dapat dihitung dengan persamaan (2.9). Perhitungan nilai MARE pada model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 dengan persamaan (2.9) adalah sebagai berikut:

$$\begin{aligned} \text{MARE} &= \frac{\sum_{t=1}^4 \left| \frac{y_t - \hat{y}_t}{y_t + 1} \right|}{4} \\ &= \frac{\left| \frac{661 - 605}{661 + 1} \right| + \left| \frac{771 - 578}{771 + 1} \right| + \dots + \left| \frac{569 - 540}{569 + 1} \right|}{4} \end{aligned}$$

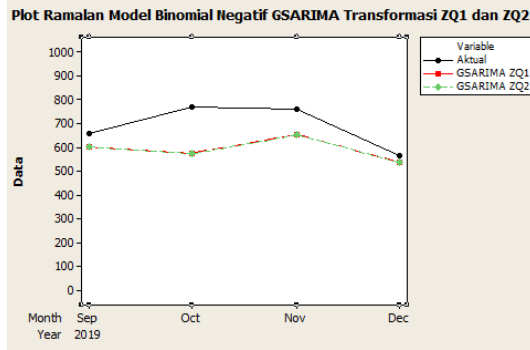
$$\begin{aligned}
 \text{MARE} &= \frac{0,0846 + 0,2500 + 0,1389 + 0,0509}{4} \\
 &= \frac{0,5244}{4} \\
 &= 0,1311
 \end{aligned}$$

Plot perbandingan hasil peramalan pada data *out-sample* dengan data aktual menggunakan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 dan ZQ2 dapat dilihat pada Gambar 4.15. Sehingga model terbaik pada penelitian ini yang selanjutnya dapat digunakan untuk model peramalan pada data jumlah penderita penyakit ISPA di Kota Surabaya adalah model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1, dengan model sebagai berikut:

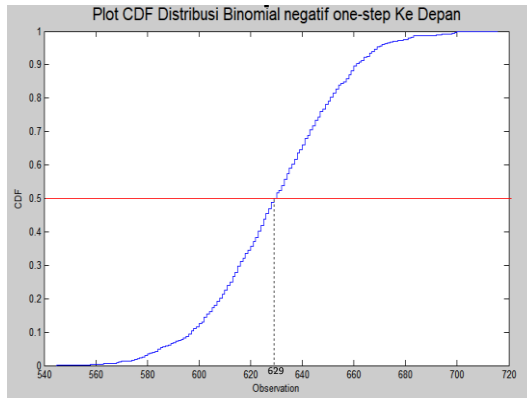
$$\begin{aligned}
 \mu_t = \exp \left\{ &0,03 + \ln(y'_{t-1}) + \ln(y'_{t-6}) \right. \\
 &- \ln(y'_{t-7}) + 0,038 \ln(y'_{t-1}) - 0,038 \ln(y'_{t-2}) \\
 &- 0,038 \ln(y'_{t-7}) + 0,038 \ln(y'_{t-8}) \\
 &- 0,351 \ln(y'_{t-2}) + 0,351 \ln(y'_{t-3}) \\
 &+ 0,351 \ln(y'_{t-8}) - 0,351 \ln(y'_{t-9}) \\
 &+ 0,048 \ln \left(\frac{y'_{t-1}}{\mu_{t-1}} \right) + 0,038 \ln \left(\frac{y'_{t-6}}{\mu_{t-6}} \right) \\
 &\left. - 0,0018 \ln \left(\frac{y'_{t-7}}{\mu_{t-7}} \right) \right\} \quad (4.11)
 \end{aligned}$$

Hasil peramalan dari model Binomial Negatif GSARIMA merupakan nilai median dari distribusi bersyarat *k-step* ke depan. Dari persamaan (4.11), kita dapatkan hasil peramalan untuk *one-step* ke depan yaitu didapatkan nilai $\hat{\mu}_{73} = 629,47$. Kemudian dengan membangkitkan suatu data, dimana data yang di random sebanyak 1000 data dengan $y_t \sim \text{Binomial negatif}(\mu_t, k)$, dengan nilai μ_t untuk *one-step* ke depan yaitu $\hat{\mu}_{73} = 629,47$ dan k adalah parameter dispersi yang diperoleh dari hasil estimasi, diperoleh nilai median sebesar 629. Hal ini dapat juga dilihat dari plot CDF

distribusi binomial negatif *one-step* kedepan pada Gambar 4.16 yang menunjukkan bahwa nilai \hat{y} yang memberikan nilai median untuk fungsi tersebut adalah 629. Sehingga nilai peramalan untuk *one-step* ke depan dengan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 terhadap data jumlah penderita penyakit ISPA di Kota Surabaya adalah 629.



Gambar 4.15 Plot Hasil Peramalan Data *Out-Sample* dengan Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶



Gambar 4.16 Plot CDF Distribusi Binomial Negatif *One-Step* ke Depan dengan $\hat{\mu}_t = 629,47$

Kemudian dengan prosedur yang sama, didapatkan hasil peramalan untuk satu tahun kedepan terhadap data jumlah penderita penyakit ISPA di Kota Surabaya menggunakan model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 yang telah disajikan pada Tabel 4.10.

Tabel 4.10 Hasil Peramalan Jumlah Penderita Penyakit ISPA Kota Surabaya Tahun 2020

| No | Periode | Hasil Ramalan |
|----|----------------|---------------|
| 1 | Januari 2020 | 629 |
| 2 | Februari 2020 | 751 |
| 3 | Maret 2020 | 869 |
| 4 | April 2020 | 947 |
| 5 | Mei 2020 | 780 |
| 6 | Juni 2020 | 598 |
| 7 | Juli 2020 | 708 |
| 8 | Agustus 2020 | 842 |
| 9 | September 2020 | 886 |
| 10 | Oktober 2020 | 913 |
| 11 | November 2020 | 879 |
| 12 | Desember 2020 | 800 |

Berdasarkan hasil peramalan untuk satu tahun ke depan pada Tabel 4.10, rata-rata jumlah penderita penyakit ISPA di Kota Surabaya pada tahun 2020 adalah sebanyak 816 orang, dengan jumlah penderita maksimum sebanyak 947 orang yaitu pada bulan April dan penderita minimum sebanyak 598 orang yaitu pada bulan Juni. Pada bulan Januari hingga April jumlah penderita penyakit ISPA cenderung mengalami kenaikan, dan pada bulan April hingga Juni cenderung mengalami penurunan. Kemudian pada bulan Juni hingga Oktober mengalami kenaikan kembali, dan pada bulan Oktober hingga Desember cenderung mengalami penurunan kembali. Sehingga dari Tabel 4.10 terlihat bahwa jumlah penderita

penyakit ISPA di Kota Surabaya pada tahun 2020 cenderung naik pada periode 6 bulan sekali yaitu pada bulan April dan bulan Oktober yang masing-masing dengan jumlah penderita sebanyak 947 dan 913 orang.

BAB V PENUTUP

Pada bab ini berisi kesimpulan yang merupakan ringkasan akhir dari hasil yang mampu menjawab tujuan penelitian yang dilakukan. Setelah itu diberikan pula saran-saran berupa perbaikan maupun pengembangan untuk penelitian mendatang.

5.1 Kesimpulan

Kesimpulan yang didapatkan pada penelitian ini adalah sebagai berikut:

1. Model terbaik yang dapat digunakan untuk peramalan pada data jumlah penderita penyakit Infeksi Saluran Pernapasan Akut (ISPA) di Kota Surabaya adalah model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ transformasi ZQ1 dengan model sebagai berikut:

$$\begin{aligned} \mu_t = \exp \left\{ & 0,03 + \ln(y'_{t-1}) + \ln(y'_{t-6}) \right. \\ & - \ln(y'_{t-7}) + 0,038 \ln(y'_{t-1}) \\ & - 0,038 \ln(y'_{t-2}) - 0,038 \ln(y'_{t-7}) \\ & + 0,038 \ln(y'_{t-8}) \\ & - 0,351 \ln(y'_{t-2}) + 0,351 \ln(y'_{t-3}) \\ & + 0,351 \ln(y'_{t-8}) - 0,351 \ln(y'_{t-9}) \\ & + 0,048 \ln \left(\frac{y'_{t-1}}{\mu_{t-1}} \right) + 0,038 \ln \left(\frac{y'_{t-6}}{\mu_{t-6}} \right) \\ & \left. - 0,0018 \ln \left(\frac{y'_{t-7}}{\mu_{t-7}} \right) \right\} \end{aligned}$$

2. Rata-rata jumlah penderita penyakit ISPA di Kota Surabaya pada tahun 2020 adalah sebanyak 816 orang, dengan jumlah penderita maksimum sebanyak 947 orang yaitu pada bulan April dan penderita minimum sebanyak 598 orang yaitu pada bulan Juni. Jumlah penderita penyakit ISPA di Kota Surabaya

pada tahun 2020 cenderung naik pada periode 6 bulan sekali yaitu pada bulan April dan bulan Oktober, dengan jumlah penderita masing-masing sebanyak 947 dan 913 orang.

5.2 Saran

Saran yang dapat digunakan untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Pada penelitian ini tidak mempertimbangkan variabel prediktor atau faktor luar yang mempengaruhi jumlah penderita penyakit ISPA. Penelitian selanjutnya diharapkan mempertimbangkan faktor luar yang mungkin berpengaruh terhadap variabel penelitian.
2. Pada penelitian selanjutnya, dapat diaplikasikan pada data musiman menggunakan kasus lain dan wilayah lain selain Kota Surabaya.

DAFTAR PUSTAKA

- [1] Kemenkes RI. 2011. *Pedoman Pengendalian Infeksi Saluran Pernafasan Akut*. Jakarta: Dirjen Pengendalian Penyakit dan Penyehatan Lingkungan.
- [2] World Health Organization, 2007. *Pencegahan dan Pengendalian Infeksi Saluran Pernapasan Akut (ISPA) yang Cenderung Menjadi Epidemik dan Pandemi Di Fasilitas Pelayanan Kesehatan*. Diakses 25 Januari 2020. https://www.who.int/csr/resources/publications/WHO_CD_S_EPR_2007_8bahasa.pdf?ua=1
- [3] Kemenkes RI. 2018. *Hasil Utama Riskesdas 2018*. Jakarta: Balitbang Kemenkes RI.
- [4] Kemenkes RI. 2018. *Hasil Utama Riskesdas 2018 Provinsi Jawa Timur*. Jakarta: Balitbang Humaniora dan Manajemen Kesehatan Kemenkes RI.
- [5] Wei, W., 2006. *Time Series Analysis*. United States of America: Pearson Education, Inc.
- [6] Box, G. E., Jenkins, G. M. & Reinsel, G. C., 2008. *Time Series Analysis Forecasting and Control Fourth Edition*. New Jersey: A John Wiley and Sons, Inc. Publication.
- [7] Briet, J. T. O., Amerasinghe, H. P., dan Vounatsou, P. 2013. *Generalized Seasonal Autoregressive Integrated Moving Average Models for Count Data with Application to Malaria Time Series with Low Case Numbers*. *Malaria Journal*, PLoS ONE, 8(6): e65761.
- [8] McCullagh, P. & Nelder, J. A., 1989. *Generalized Linear Models*. London: Chapman and Hall.
- [9] Benjamin, M. A., Rigby, R. A. & Stasinopoulos, D. M., 2003. Generalized Autoregressive Moving Average Models. *Journal of the American Statistical Association*, Vol 98, No 461, pp. 214-216.
- [10] Asrirawan, 2015. Simulasi Perbandingan Metode Peramalan Model Generalized Seasonal Autoregressive Integrated Moving Average (GSARIMA) dengan Seasonal

- Autregressive Integrated Moving Average (SARIMA). *Jurnal Dinamika*, 06(1), pp. 61-66
- [11] Diana, I., 2019. Model Peramalan Jumlah Kriminalitas Pencurian di Wilayah Polrestabes Kota Surabaya dengan Menggunakan Model Poisson GARMA dan Model Binomial Negatif GARMA. Tugas Akhir Jurusan Matematika, FMKSD, Institut Teknologi Sepuluh Nopember.
- [12] Fitriawan, E. S., 2019. Peramalan Jumlah Kasus Infeksi Saluran Pernafasan Akut (ISPA) pada Laki-laki Tahun 2019 dengan Metode ARIMA. *Jurnal Biometrika dan Kependudukan*, 8(2), pp.134-141
- [13] Fauzia, A. D., 2018. Estimasi Parameter pada Model Poisson Generalized Autoregressive Moving Average (GARMA) dengan Algoritma IRLS (Studi Kasus: Peramalan Jumlah Kecelakaan di Jalan Tol Surabaya-Gempol). Tugas Akhir Jurusan Matematika, FMKSD, Institut Teknologi Sepuluh Nopember.
- [14] Habibi, M. A., 2018. Estimasi Parameter pada Model Binomial Negatif Generalized Autoregressive Moving Average (GARMA) dengan Algoritma IRLS (Studi Kasus: Peramalan Jumlah Kecelakaan di Jalan Tol Surabaya-Gempol). Tugas Akhir Jurusan Matematika, FMKSD, Institut Teknologi Sepuluh Nopember.
- [15] Makridakis, S., Wheelwright, S.C. & McGee, V.E., 1988. *Metode dan Aplikasi Peramalan*. Jakarta: PT. Glora Aksara, Pratama, 1999.
- [16] Keedem, B. & Fokianos, K., 2002. *Regression Models for Time Series Analysis*. Wiley: s.n.
- [17] J. Bain & M. Engelhardt, 1992. *Introduction to Probability and Mathematical Statistics*. United States of America: Duxbury.
- [18] Box, G. E. & Tiao, G. C., 1973. *Bayesian Inference in Statistical Analysis*. London: Addison-Wisley

- [19] Ntzoufras, I. ,2009. *Bayesian Modeling Using WinBUGS*.
New York: John Wiley & Sons, Inc.

LAMPIRAN A

Mixed Distribution Gamma dan Poisson

Binomial Negatif merupakan distribusi Poisson *mixture* yaitu distribusi Poisson-Gamma *mixture*, dimana *mean* distribusi Poisson dinyatakan dalam distribusi Gamma. Misalkan diketahui:

$$y_t \sim POI(\mu)$$

dengan

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

dan

$$\mu \sim GAM(\alpha, \beta)$$

dengan

$$f(\mu; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \mu^{\alpha-1} e^{-\frac{\mu}{\beta}} \text{ untuk } x > 0$$

Maka distribusi Poisson-Gamma *mixture* memiliki pdf

$$\begin{aligned} f(y_t; \alpha, \beta) &= \int_{\tilde{0}}^{\tilde{\infty}} f(y_t|\mu) f(\mu) d\mu \\ &= \int_{\tilde{0}}^{\tilde{\infty}} f(y_t|\mu) f(\mu) d\mu + \int_0^{\tilde{\infty}} f(y_t|\mu) f(\mu) d\mu \\ &= 0 + \int_0^{\tilde{\infty}} \frac{e^{-\mu} \mu^{y_t}}{y_t!} \frac{1}{\beta^\alpha \Gamma(\alpha)} \mu^{\alpha-1} e^{-\frac{\mu}{\beta}} d\mu \\ &= \frac{1}{\beta^\alpha y_t! \Gamma(\alpha)} \int_0^{\tilde{\infty}} e^{-\mu} \mu^{y_t} \mu^{\alpha-1} e^{-\frac{\mu}{\beta}} d\mu \\ &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \int_0^{\tilde{\infty}} e^{-\mu(1+\frac{1}{\beta})} \mu^{y_t+\alpha-1} d\mu \end{aligned}$$

LAMPIRAN A Lanjutan

Misal jika $v = \mu \left(1 + \frac{1}{\beta}\right)$

$$\mu = \frac{v}{1 + \frac{1}{\beta}}$$

maka $dv = \left(1 + \frac{1}{\beta}\right) d\mu$

$$d\mu = \frac{dv}{1 + \frac{1}{\beta}}$$

Jika $\mu = 0$ maka $v = 0$

Jika $\mu = \sim$ maka $v = \sim$

Sehingga *pdf* distribusi Poisson-Gamma *mixture* menjadi:

$$\begin{aligned} f(y_t; \alpha, \beta) &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \int_0^{\sim} e^{-v} \left(\frac{v}{1 + 1/\beta}\right)^{y_t + \alpha - 1} \frac{dv}{1 + 1/\beta} \\ &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \int_0^{\sim} e^{-v} \left(\frac{\beta v}{\beta + 1}\right)^{y_t + \alpha - 1} \left(\frac{\beta}{\beta + 1}\right) dv \\ &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \int_0^{\sim} e^{-v} \left(\frac{\beta}{\beta + 1}\right)^{y_t + \alpha - 1} v^{y_t + \alpha - 1} \\ &\quad \left(\frac{\beta}{\beta + 1}\right) dv \\ &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \int_0^{\sim} e^{-v} \left(\frac{\beta}{\beta + 1}\right)^{y_t + \alpha} v^{y_t + \alpha - 1} dv \\ &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \left(\frac{\beta}{\beta + 1}\right)^{y_t + \alpha} \int_0^{\sim} v^{y_t + \alpha - 1} e^{-v} dv \end{aligned}$$

Karena

$$\Gamma(n) = \int_0^{\sim} x^{n-1} e^{-x} dx$$

LAMPIRAN A Lanjutan

sehingga

$$\begin{aligned}
 f(y_t; \alpha, \beta) &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \left(\frac{\beta}{\beta + 1} \right)^{y_t + \alpha} \Gamma(y_t + \alpha) \\
 &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \left(\frac{\beta}{\beta + 1} \right)^{y_t} \left(\frac{\beta}{\beta + 1} \right)^{\alpha} \Gamma(y_t + \alpha) \\
 &= \frac{\beta^{-\alpha}}{y_t! \Gamma(\alpha)} \left(\frac{\beta}{\beta + 1} \right)^{y_t} \beta^{\alpha} \left(\frac{1}{\beta + 1} \right)^{\alpha} \Gamma(y_t + \alpha) \\
 &= \frac{\Gamma(y_t + \alpha)}{y_t! \Gamma(\alpha)} \left(\frac{\beta}{\beta + 1} \right)^{y_t} \left(\frac{1}{\beta + 1} \right)^{\alpha}
 \end{aligned}$$

misalkan $k = \frac{1}{\alpha}$ atau $\alpha = \frac{1}{k}$ dan $\mu_t = \alpha\beta$, maka $\beta = k\mu_t$

$$f(y_t; \mu_t, k) = \frac{\Gamma(y_t + 1/k)}{y_t! \Gamma(1/k)} \left(\frac{k\mu_t}{k\mu_t + 1} \right)^{y_t} \left(\frac{1}{k\mu_t + 1} \right)^{1/k}$$

LAMPIRAN B

Data Jumlah Penderita Penyakit Infeksi Saluran Pernapasan Akut (ISPA) Kategori Pneumonia Kota Surabaya Januari 2014-Desember 2019

Tabel B. Data Jumlah Penderita Penyakit ISPA Kota Surabaya

| | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|-----------|------|------|------|------|------|------|
| Januari | 175 | 166 | 148 | 215 | 356 | 459 |
| Februari | 262 | 241 | 232 | 299 | 349 | 679 |
| Maret | 326 | 273 | 240 | 396 | 471 | 670 |
| April | 422 | 257 | 238 | 394 | 427 | 569 |
| Mei | 364 | 190 | 277 | 357 | 382 | 479 |
| Juni | 274 | 165 | 228 | 292 | 340 | 369 |
| Juli | 210 | 183 | 155 | 330 | 494 | 419 |
| Agustus | 241 | 300 | 254 | 360 | 645 | 606 |
| September | 217 | 228 | 300 | 394 | 462 | 661 |
| Oktober | 223 | 222 | 310 | 399 | 475 | 771 |
| November | 188 | 217 | 295 | 369 | 473 | 762 |
| Desember | 207 | 205 | 250 | 322 | 414 | 569 |

LAMPIRAN C

Perhitungan Autokorelasi (ACF) dan Autokorelasi Parsial (PACF)

Perhitungan autokorelasi (ACF) dan autokorelasi parsial (PACF) hasil transformasi 2 *Box-Cox Plot* adalah sebagai berikut:

1. Nilai autokorelasi (ACF)

$$\begin{aligned}
 \bar{Z} &= \sum_{t=1}^n Z_t \\
 &= \sum_{t=1}^{68} Z_t \\
 &= \frac{3,63 + 4,02 + \dots + 4,52 + 4,96}{68} \\
 &= \frac{285,76}{68} \\
 &= 4,20
 \end{aligned}$$

Autokorelasi pada lag k

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}$$

Autokorelasi pada lag 1 ($\hat{\rho}_1$)

$$\begin{aligned}
 &= \frac{\sum_{t=1}^{67} (Z_t - \bar{Z})(Z_{t+1} - \bar{Z})}{\sum_{t=1}^{68} (Z_t - \bar{Z})^2} \\
 &= \frac{(Z_1 - \bar{Z})(Z_2 - \bar{Z}) + \dots + (Z_{67} - \bar{Z})(Z_{68} - \bar{Z})}{(Z_1 - \bar{Z})^2 + \dots + (Z_{68} - \bar{Z})^2} \\
 &= \frac{(3,63 - 4,2)(4,02 - 4,2) + \dots}{(3,63 - 4,2)^2 + \dots + (4,96 - 4,2)^2} \\
 &\quad \dots + (4,52 - 4,2)(4,96 - 4,2) \\
 &= \frac{\dots + (4,96 - 4,2)^2}{(3,63 - 4,2)^2 + \dots + (4,96 - 4,2)^2}
 \end{aligned}$$

LAMPIRAN C Lanjutan

$$\begin{aligned}
 &= \frac{8,33}{10,51} \\
 &= 0,79
 \end{aligned}$$

Autokorelasi pada lag 10 ($\hat{\rho}_{10}$)

$$\begin{aligned}
 &= \frac{\sum_{t=1}^{58} (Z_t - \bar{Z})(Z_{t+10} - \bar{Z})}{\sum_{t=1}^{68} (Z_t - \bar{Z})^2} \\
 &= \frac{(Z_1 - \bar{Z})(Z_{11} - \bar{Z}) + \dots + (Z_{58} - \bar{Z})(Z_{68} - \bar{Z})}{(Z_1 - \bar{Z})^2 + \dots + (Z_{68} - \bar{Z})^2} \\
 &= \frac{(3,63 - 4,2)(3,70 - 4,2) + \dots}{(3,63 - 4,2)^2 + \dots + (4,96 - 4,2)^2} \\
 &= \frac{\dots + (4,67 - 4,2)(4,96 - 4,2)}{\dots + (4,67 - 4,2)(4,96 - 4,2)} \\
 &= \frac{(3,63 - 4,2)^2 + \dots + (4,96 - 4,2)^2}{(3,63 - 4,2)^2 + \dots + (4,96 - 4,2)^2} \\
 &= \frac{4,49}{10,51} \\
 &= 0,43
 \end{aligned}$$

2. Selang kepercayaan pada plot ACF

Batas atas selang kepercayaan ACF pada lag k

$$= t_{n-1, 1-\frac{\alpha}{2}} \times SE_{(\hat{\rho}_k)}$$

Batas bawah selang kepercayaan ACF pada lag k

$$= t_{n-1, \frac{\alpha}{2}} \times SE_{(\hat{\rho}_k)}$$

dimana

$$SE_{(\hat{\rho}_k)} = \sqrt{\frac{(1 + 2 \sum_{i=1}^{k-1} \hat{\rho}_i^2)}{n}}$$

dengan

$$k = 1, 2, 3, \dots$$

n = jumlah data pengamatan

$\hat{\rho}_i$ = autokorelasi pada lag i

LAMPIRAN C Lanjutan

Selang kepercayaan plot ACF pada lag 1

$$\begin{aligned} SE_{(\hat{\rho}_1)} &= \sqrt{\frac{1}{68}} (1) \\ &= 0,12 \end{aligned}$$

dengan $\alpha = 5\%$

Batas atas selang kepercayaan ACF pada lag 1

$$\begin{aligned} &= t_{68-1; 1-\frac{0,05}{2}} \times SE_{(\hat{\rho}_1)} \\ &= t_{67; 0,975} \times SE_{(\hat{\rho}_1)} \\ &= 1,996 \times 0,12 \\ &= 0,24 \end{aligned}$$

Batas bawah selang kepercayaan ACF pada lag 1

$$\begin{aligned} &= t_{68-1; \frac{0,05}{2}} \times SE_{(\hat{\rho}_1)} \\ &= t_{67; 0,025} \times SE_{(\hat{\rho}_1)} \\ &= (-1,996) \times 0,12 \\ &= -0,24 \end{aligned}$$

3. Nilai autokorelasi parsial (PACF)

$\hat{\phi}_{k,k}$ adalah nilai PACF pada lag k, dengan

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_j}$$

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{k,j} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}$$

Nilai PACF lag 1

$$\begin{aligned} \hat{\phi}_{1,1} &= \frac{\hat{\rho}_1}{1} \\ &= 0,80 \end{aligned}$$

LAMPIRAN C Lanjutan

Nilai PACF lag 2

$$\begin{aligned}\hat{\phi}_{2,2} &= \frac{\hat{\rho}_2 - \sum_{j=1}^1 \hat{\phi}_{1j} \hat{\rho}_{1+1-j}}{1 - \sum_{j=1}^1 \hat{\phi}_{1j} \hat{\rho}_j} \\ &= \frac{\hat{\rho}_2 - \hat{\phi}_{1,1} \hat{\rho}_1}{1 - \hat{\phi}_{1,1} \hat{\rho}_1} \\ &= \frac{0,59 - (0,80)(0,80)}{1 - (0,80)(0,80)} \\ &= -0,12\end{aligned}$$

dan

$$\begin{aligned}\hat{\phi}_{2,1} &= \hat{\phi}_{1,1} - \hat{\phi}_{2,2} \hat{\phi}_{1,1+1-1} \\ &= \hat{\phi}_{1,1} - \hat{\phi}_{2,2} \hat{\phi}_{1,1} \\ &= 0,80 - (-0,12)(0,80) \\ &= 0,89\end{aligned}$$

dengan $\hat{\phi}_{2,2}$ dan $\hat{\phi}_{2,1}$ dapat digunakan untuk mencari PACF pada lag 3.

4. Selang kepercayaan pada plot PACF

$$\begin{aligned}S_{\hat{\phi}_{k,k}} &= \pm \frac{t_{n-1; 1-\frac{\alpha}{2}}}{\sqrt{n}} \\ &= \pm \frac{t_{67; 0,975}}{\sqrt{67}} \\ &= \pm \frac{1,996}{\sqrt{67}} \\ &= \pm 0,24\end{aligned}$$

Nilai ACF pada lag 1 sebesar 0,79 dengan batas atas selang kepercayaan pada lag 1 sebesar 0,24, dalam kata lain ACF pada lag 1 melebihi selang kepercayaan. Sehingga dapat diartikan terdapat korelasi antara Z_t dengan Z_{t+1} .

LAMPIRAN C Lanjutan

Sedangkan nilai PACF pada *lag* 1 sebesar 0,80 dan nilai PACF pada *lag* 2 sebesar -0,12, dengan selang kepercayaan pada setiap *lag* sebesar $\pm 0,24$. Karena $\hat{\phi}_{1,1} > S_{\hat{\phi}_{1,1}}$ dan $\hat{\phi}_{2,2} < S_{\hat{\phi}_{2,2}}$ dengan kata lain nilai PACF pada *lag* 1 melebihi selang kepercayaan, sedangkan nilai PACF pada *lag* 2 tidak melebihi selang kepercayaan. Sehingga dengan menghilangkan dependensi antar variabel respon, dapat diartikan bahwa terdapat korelasi antara Z_t dengan Z_{t+1} , dan tidak ada korelasi antara Z_t dengan Z_{t+2} .

LAMPIRAN D

Estimasi dan Uji Signifikansi Model *Overfitting*

Tabel D. Tabel Estimasi dan Uji Signifikansi Model *Overfitting*

| Model | Var | Koef | Koef se | T | p | Ket |
|---|---------|---------|---------|-------|-------|-------------------------|
| SARIMA (2,1,2) (1,1,0) ⁶ | AR 1 | 0,3162 | 0,2286 | 1,38 | 0,172 | Tidak Signifi kan |
| | AR 2 | -0,0037 | 0,1807 | -0,02 | 0,984 | |
| | SAR 1 | -0,4456 | 0,1311 | -3,40 | 0,001 | |
| | MA 1 | 0,6538 | 0,1907 | 3,43 | 0,001 | |
| | MA 2 | 0,4134 | 0,1748 | 2,36 | 0,022 | |
| | Konstan | 1,4573 | 0,0236 | 61,81 | 0,000 | |
| SARIMA (2,1,2) (0,1,1) ⁶ | AR 1 | 1,0047 | 0,2690 | 3,73 | 0,000 | Tidak Signifi kan |
| | AR 2 | -0,5841 | 0,1878 | -3,11 | 0,003 | |
| | MA 1 | 1,3257 | 0,2914 | 4,55 | 0,000 | |
| | MA 2 | -0,4480 | 0,2912 | -1,54 | 0,130 | |
| | SMA 1 | 0,7553 | 0,1365 | 5,53 | 0,000 | |
| | Konstan | 0,9743 | 0,2244 | 4,34 | 0,000 | |
| SARIMA (2,1,2) (0,1,0) ⁶ | AR 1 | 0,1524 | 0,4527 | 0,34 | 0,738 | Tidak Signifi kan |
| | AR 2 | -0,0859 | 0,2899 | -0,30 | 0,768 | |
| | MA 1 | 0,5692 | 0,4490 | 1,27 | 0,210 | |
| | MA 2 | 0,3807 | 0,4570 | 0,83 | 0,408 | |
| | Konstan | 1,3934 | 0,9445 | 1,48 | 0,146 | |
| SARIMA (2,1,1) (1,1,1) ⁶ | AR 1 | 0,5857 | 0,1524 | 3,84 | 0,000 | Tidak Signifi kan |
| | AR 2 | -0,3387 | 0,1400 | -2,42 | 0,019 | |
| | SAR 1 | -0,0008 | 0,2118 | -0,00 | 0,997 | |
| | MA 1 | 0,8834 | 0,1119 | 7,90 | 0,000 | |
| | SMA 1 | 0,7575 | 0,1835 | 4,13 | 0,000 | |
| | Konstan | 1,2298 | 0,2313 | 5,32 | 0,000 | |
| SARIMA (2,1,1) (1,1,0) ⁶ | AR 1 | -1,0079 | 0,1499 | -6,72 | 0,000 | Tidak Signifi kan |
| | AR 2 | -0,1173 | 0,1362 | -0,86 | 0,393 | |
| | SAR 1 | -0,5261 | 0,1326 | -3,97 | 0,000 | |
| | MA 1 | -0,9551 | 0,0967 | -9,88 | 0,000 | |
| | Konstan | -2,57 | 16,13 | -0,16 | 0,874 | |
| SARIMA (2,1,1) (0,1,1) ⁶ | AR 1 | 0,6476 | 0,1425 | 4,55 | 0,000 | Signifi kan |
| | AR 2 | -0,3834 | 0,1369 | -2,80 | 0,007 | |
| | MA 1 | 0,9166 | 0,0824 | 11,12 | 0,000 | |

LAMPIRAN D Lanjutan

| Model | Var | Koef | Koef se | T | p | Ket |
|---|---------|---------|---------|-------|-------|-------------------------|
| | SMA 1 | 0,8036 | 0,1161 | 6,92 | 0,000 | |
| | Konstan | 1,1883 | 0,1818 | 6,54 | 0,000 | |
| SARIMA (2,1,1) (0,1,0) ⁶ | AR 1 | 0,5152 | 0,1335 | 3,86 | 0,000 | Tidak Signifi kan |
| | AR 2 | -0,2585 | 0,1349 | -1,92 | 0,60 | |
| | MA 1 | 0,9594 | 0,0625 | 15,35 | 0,000 | |
| | Konstan | 1,1463 | 0,6241 | 1,84 | 0,071 | |
| SARIMA (2,1,0) (1,1,1) ⁶ | AR 1 | -0,0761 | 0,1269 | -0,60 | 0,551 | Tidak Signifi kan |
| | AR 2 | -0,3495 | 0,1259 | -2,78 | 0,007 | |
| | SAR 1 | 0,1475 | 0,2186 | 0,67 | 0,503 | |
| | SMA 1 | 0,8051 | 0,1843 | 4,37 | 0,000 | |
| | Konstan | 0,8050 | 1,6140 | 0,50 | 0,620 | |
| SARIMA (2,1,0) (1,1,0) ⁶ | AR 1 | -0,0936 | 0,1253 | -0,75 | 0,458 | Tidak Signifi kan |
| | AR 2 | -0,3257 | 0,1260 | -2,59 | 0,012 | |
| | SAR 1 | -0,5070 | 0,1264 | -4,01 | 0,000 | |
| | Konstan | -0,8600 | 7,8410 | -0,11 | 0,913 | |
| SARIMA (2,1,0) (0,1,1) ⁶ | AR 1 | -0,0701 | 0,1269 | -0,55 | 0,583 | Tidak Signifi kan |
| | AR 2 | -0,3361 | 0,1259 | -2,67 | 0,010 | |
| | SMA 1 | 0,7615 | 0,1138 | 6,69 | 0,000 | |
| | Konstan | 1,1230 | 1,9730 | 0,57 | 0,572 | |
| SARIMA (2,1,0) (0,1,0) ⁶ | AR 1 | -0,1833 | 0,1193 | -1,54 | 0,130 | Tidak Signifi kan |
| | AR 2 | -0,4158 | 0,1193 | -3,49 | 0,001 | |
| | Konstan | -1,4210 | 8,5950 | -0,17 | 0,869 | |
| SARIMA (1,1,2) (1,1,1) ⁶ | AR 1 | 0,1048 | 0,2677 | 0,39 | 0,697 | Tidak Signifi kan |
| | SAR 1 | 0,0080 | 0,2092 | 0,04 | 0,970 | |
| | MA 1 | 0,4552 | 0,2483 | 1,83 | 0,072 | |
| | MA 2 | 0,4562 | 0,1940 | 2,35 | 0,022 | |
| | SMA 1 | 0,7711 | 0,1763 | 4,37 | 0,000 | |
| | Konstan | 1,4500 | 0,2022 | 7,17 | 0,000 | |
| SARIMA (1,1,2) (1,1,0) ⁶ | AR 1 | 0,2538 | 0,2426 | 1,05 | 0,300 | Tidak Signifi kan |
| | SAR 1 | -0,5136 | 0,1311 | -3,92 | 0,000 | |
| | MA 1 | 0,5478 | 0,2287 | 2,40 | 0,020 | |
| | MA 2 | 0,3937 | 0,1974 | 1,99 | 0,051 | |
| | Konstan | 1,5765 | 0,6183 | 2,55 | 0,014 | |

LAMPIRAN D Lanjutan

| Model | Var | Koef | Koef se | T | p | Ket |
|---|---------|---------|---------|-------|-------|-------------------------|
| SARIMA (1,1,2) (0,1,1) ⁶ | AR 1 | 0,1013 | 0,2468 | 0,38 | 0,703 | Tidak Signifi kan |
| | MA 1 | 0,4552 | 0,2446 | 1,86 | 0,068 | |
| | MA 2 | 0,4592 | 0,1918 | 2,39 | 0,020 | |
| | SMA 1 | 0,7709 | 0,1204 | 6,40 | 0,000 | |
| | Konstan | 1,4737 | 0,2099 | 7,02 | 0,000 | |
| SARIMA (1,1,2) (0,1,0) ⁶ | AR 1 | 0,0427 | 0,2675 | 0,16 | 0,874 | Tidak Signifi kan |
| | MA 1 | 0,4587 | 0,2472 | 1,86 | 0,069 | |
| | MA 2 | 0,4902 | 0,2078 | 2,36 | 0,022 | |
| | Konstan | 1,3900 | 1,0940 | 1,27 | 0,209 | |
| SARIMA (1,1,1) (1,1,1) ⁶ | AR 1 | 0,4769 | 0,1772 | 2,69 | 0,009 | Tidak Signifi kan |
| | SAR 1 | 0,0333 | 0,2423 | 0,14 | 0,891 | |
| | MA 1 | 0,9134 | 0,1180 | 7,74 | 0,000 | |
| | SMA 1 | 0,7115 | 0,2176 | 3,27 | 0,002 | |
| | Konstan | 0,7267 | 0,3348 | 2,17 | 0,034 | |
| SARIMA (1,1,1) (1,1,0) ⁶ | AR 1 | 0,5417 | 0,1353 | 4,00 | 0,000 | Signifi kan |
| | SAR 1 | -0,5260 | 0,1291 | -4,07 | 0,000 | |
| | MA 1 | 0,9567 | 0,0654 | 12,62 | 0,000 | |
| | Konstan | 0,8617 | 0,4690 | 1,84 | 0,071 | |
| SARIMA (1,1,1) (0,1,1) ⁶ | AR 1 | 0,4779 | 0,1735 | 2,75 | 0,008 | Signifi kan |
| | MA 1 | 0,9150 | 0,1144 | 8,00 | 0,000 | |
| | SMA 1 | 0,6983 | 0,1303 | 5,36 | 0,000 | |
| | Konstan | 0,7418 | 0,3349 | 2,22 | 0,031 | |
| SARIMA (1,1,1) (0,1,0) ⁶ | AR 1 | 0,4444 | 0,1295 | 3,43 | 0,001 | Signifi kan |
| | MA 1 | 0,9734 | 0,0582 | 16,73 | 0,000 | |
| | Konstan | 0,6925 | 0,5156 | 1,34 | 0,184 | |
| SARIMA (1,1,0) (1,1,0) ⁶ | AR 1 | -0,0666 | 0,1322 | -0,50 | 0,616 | Tidak Signifi kan |
| | SAR 1 | -0,5676 | 0,1236 | -4,59 | 0,000 | |
| | Konstan | -1,1730 | 8,1750 | -0,14 | 0,886 | |
| SARIMA (1,1,0) (1,1,1) ⁶ | AR 1 | -0,0491 | 0,1352 | -0,36 | 0,718 | Tidak Signifi kan |
| | SAR 1 | 0,0986 | 0,2124 | 0,46 | 0,644 | |
| | SMA 1 | 0,8148 | 0,1856 | 4,39 | 0,000 | |
| | Konstan | 0,7120 | 1,6830 | 0,42 | 0,674 | |

LAMPIRAN D Lanjutan

| Model | Var | Koef | Koef se | T | p | Ket |
|---|---------|---------|---------|-------|-------|-------------------|
| SARIMA (1,1,0) (0,1,1) ⁶ | AR 1 | -0,0471 | 0,1342 | -0,35 | 0,727 | Tidak Signifi kan |
| | SMA 1 | 0,7894 | 0,1204 | 6,56 | 0,000 | |
| | Konstan | 0,9170 | 2,1140 | 0,43 | 0,666 | |
| SARIMA (1,1,0) (0,1,0) ⁶ | AR 1 | -0,0471 | 0,1342 | -0,35 | 0,727 | Tidak Signifi kan |
| | Konstan | -1,8160 | 9,3660 | -0,19 | 0,847 | |
| SARIMA (0,1,2) (1,1,1) ⁶ | SMA 1 | 0,7894 | 0,1204 | 6,56 | 0,000 | Tidak Signifi kan |
| | Konstan | 0,9170 | 2,1140 | 0,43 | 0,666 | |
| | AR 1 | -0,1251 | 0,1294 | -0,97 | 0,337 | |
| | SMA 1 | 0,7057 | 0,1768 | 3,99 | 0,000 | |
| | Konstan | 1,6290 | 0,1650 | 9,87 | 0,000 | |
| SARIMA (0,1,2) (1,1,0) ⁶ | SAR 1 | -0,4717 | 0,1343 | -3,51 | 0,001 | Signifi kan |
| | MA 1 | 0,4008 | 0,1152 | 3,48 | 0,001 | |
| | MA 2 | 0,5415 | 0,1173 | 4,62 | 0,000 | |
| | Konstan | 2,1622 | 0,5692 | 3,80 | 0,000 | |
| SARIMA (0,1,2) (0,1,1) ⁶ | MA 1 | 0,5168 | 0,1274 | 4,06 | 0,000 | Signifi kan |
| | MA 2 | 0,5696 | 0,1364 | 4,18 | 0,000 | |
| | SMA 1 | 0,6855 | 0,1189 | 5,77 | 0,000 | |
| | Konstan | 1,6840 | 0,3303 | 5,10 | 0,000 | |
| SARIMA (0,1,2) (0,1,0) ⁶ | MA 1 | 0,4258 | 0,1251 | 3,40 | 0,001 | Signifi kan |
| | MA 2 | 0,5211 | 0,1258 | 4,14 | 0,000 | |
| | Konstan | 1,4650 | 1,1610 | 1,26 | 1,212 | |
| SARIMA (0,1,1) (1,1,1) ⁶ | SAR 1 | 0,1131 | 0,2247 | 0,50 | 0,616 | Tidak Signifi kan |
| | MA 1 | 0,2023 | 0,1333 | 1,52 | 0,135 | |
| | SMA 1 | 0,8107 | 0,1986 | 4,08 | 0,000 | |
| | Konstan | 0,6030 | 1,3620 | 0,44 | 0,660 | |
| SARIMA (0,1,1) (1,1,0) ⁶ | SAR 1 | -0,5583 | 0,1243 | -4,49 | 0,000 | Tidak Signifi kan |
| | MA 1 | 0,1909 | 0,1302 | 1,47 | 0,148 | |
| | Konstan | -0,9640 | 6,5780 | -0,15 | 0,884 | |
| SARIMA (0,1,1) (0,1,1) ⁶ | MA 1 | 0,1812 | 0,1327 | 1,37 | 0,177 | Tidak Signifi kan |
| | SMA 1 | 0,7807 | 0,1225 | 6,37 | 0,000 | |
| | Konstan | 0,8050 | 1,7950 | 0,45 | 0,656 | |

LAMPIRAN D Lanjutan

| Model | Var | Koef | Koef se | T | p | Ket |
|---|---------|---------|---------|-------|-------|-------------------------|
| SARIMA (0,1,1) (0,1,0) ⁶ | MA 1 | 0,7682 | 0,0952 | 8,07 | 0,000 | Signifi kan |
| | Konstan | 0,9230 | 2,1560 | 0,43 | 0,670 | |
| SARIMA (0,1,0) (1,1,1) ⁶ | SAR 1 | 0,0934 | 0,2053 | 0,45 | 0,651 | Tidak Signifi kan |
| | SMA 1 | 0,8111 | 0,1789 | 4,53 | 0,000 | |
| | Konstan | 0,7100 | 1,6890 | 0,42 | 0,676 | |
| SARIMA (0,1,0) (1,1,0) ⁶ | SAR 1 | -0,5692 | 0,1208 | -4,71 | 0,000 | Signifi kan |
| | Konstan | -1,1230 | 8,131 | -0,14 | 0,891 | |
| SARIMA (0,1,0) (0,1,1) ⁶ | SMA 1 | 0,7927 | 0,1167 | 6,79 | 0,000 | Signifi kan |
| | Konstan | 0,9060 | 2,1240 | 0,43 | 0,671 | |

Uji Signifikansi Parameter Model SARIMA hasil *overfitting*

Uji signifikansi parameter model SARIMA(2,1,1)(0,1,1)⁶ adalah sebagai berikut:

1. Uji parameter ϕ_1 (AR 1)

Hipotesis:

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

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

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\phi}_1}{SE(\hat{\phi}_1)} \\ &= \frac{0,6476}{0,1425} \\ &= 4,5456 \end{aligned}$$

Dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

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

LAMPIRAN D Lanjutan

2. Uji parameter ϕ_2 (AR 2)

Hipotesis:

$H_0: \phi_2 = 0$ (parameter ϕ_2 tidak signifikan)

$H_1: \phi_2 \neq 0$ (parameter ϕ_2 signifikan)

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\phi}_2}{SE(\hat{\phi}_2)} \\ &= \frac{-0,3834}{0,1369} \\ &= -2,8006 \end{aligned}$$

Dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

Jika $|t_{hitung}| > t_{tabel}$ maka H_0 ditolak, yang artinya parameter ϕ_2 signifikan.

3. Uji parameter θ_1 (MA 1)

Hipotesis:

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

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

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\theta}_1}{SE(\hat{\theta}_1)} \\ &= \frac{0,9166}{0,0824} \\ &= 11,124 \end{aligned}$$

Dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

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

LAMPIRAN D Lanjutan

4. Uji parameter θ_1 (SMA 1)

Hipotesis:

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

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

Statistik uji:

$$\begin{aligned} t_{hitung} &= \frac{\hat{\theta}_1}{SE(\hat{\theta}_1)} \\ &= \frac{0,8036}{0,1161} \\ &= 6,9216 \end{aligned}$$

Dengan tabel distribusi t diperoleh:

$$t_{tabel} = t_{\frac{\alpha}{2}; n-1} = t_{0,025; 67} = 1,99601$$

Kriteria pengujian:

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

Karena semua parameter pada model signifikan, maka model SARIMA(2,1,2)(1,1,1)⁶ merupakan model yang signifikan.

LAMPIRAN E

Uji *Ljung-Box* pada Model SARIMA Hasil *Overfitting* yang Memenuhi Uji Signifikansi Parameter

Uji *Ljung-Box* pada model SARIMA(2,1,1)(0,1,1)⁶

Hipotesis:

H₀ : $\rho_1 = \rho_2 = \dots = \rho_k = 0$ (residual *white noise*)

H₁ : Minimal ada satu $\rho_i \neq 0$, $i = 1, 2, \dots, k$ (residual tidak *white noise*)

Statistik uji:

$$Q = n'(n' + 2) \sum_{k=1}^K \frac{(\hat{\rho}_k)^2}{n' - k}$$

$$\begin{aligned} n' &= n - (d + SD) \\ &= 68 - (1 + 6(1)) \\ &= 68 - 7 \\ &= 61 \end{aligned}$$

Untuk $K = 12$ diperoleh

$$\begin{aligned} Q &= 61(61 + 2) \sum_{k=1}^{12} \frac{(\hat{\rho}_k)^2}{61 - k} \\ &= 61(63) \left(\frac{(\hat{\rho}_1)^2}{61 - 1} + \frac{(\hat{\rho}_2)^2}{61 - 2} + \dots + \frac{(\hat{\rho}_{12})^2}{61 - 12} \right) \\ &= 3843 \left(\frac{(-0,0878)^2}{60} + \frac{(0,0166)^2}{59} + \frac{(0,0295)^2}{58} \right. \\ &\quad \left. + \frac{(0,0953)^2}{57} + \frac{(-0,1589)^2}{56} + \frac{(0,0930)^2}{55} \right. \\ &\quad \left. + \frac{(0,2187)^2}{54} + \frac{(-0,0247)^2}{53} + \frac{(-0,0643)^2}{52} \right. \\ &\quad \left. + \frac{(0,0369)^2}{51} + \frac{(-0,1482)^2}{50} + \frac{(-0,0131)^2}{49} \right) \\ &= 9,0766 \end{aligned}$$

LAMPIRAN E Lanjutan

dengan nilai

$$\begin{aligned} & \chi^2(\alpha; K - (p + q + P + Q)) \\ &= \chi^2(0,05; 12 - (2 + 1 + 0 + 1)) \\ &= \chi^2(0,05; 8) \\ &= 15,507 \end{aligned}$$

Untuk $K = 24$ diperoleh

$$\begin{aligned} Q &= 61(61 + 2) \sum_{k=1}^{24} \frac{(\hat{p}_k)^2}{61 - k} \\ &= 61(63) \left(\frac{(\hat{p}_1)^2}{61 - 1} + \dots + \frac{(\hat{p}_{24})^2}{61 - 24} \right) \\ &= 3843 \left(\frac{(-0,0878)^2}{60} + \dots + \frac{(-0,1001)^2}{37} \right) \\ &= 15,0397 \end{aligned}$$

dengan nilai

$$\begin{aligned} & \chi^2(\alpha; K - (p + q + P + Q)) \\ &= \chi^2(0,05; 24 - (2 + 1 + 0 + 1)) \\ &= \chi^2(0,05; 20) \\ &= 31,410 \end{aligned}$$

Untuk $K = 36$ diperoleh

$$\begin{aligned} Q &= 61(61 + 2) \sum_{k=1}^{36} \frac{(\hat{p}_k)^2}{61 - k} \\ &= 61(63) \left(\frac{(\hat{p}_1)^2}{61 - 1} + \dots + \frac{(\hat{p}_{36})^2}{61 - 36} \right) \\ &= 3843 \left(\frac{(-0,0878)^2}{60} + \dots + \frac{(0,1175)^2}{25} \right) \\ &= 26,7435 \end{aligned}$$

LAMPIRAN E Lanjutan

dengan nilai

$$\begin{aligned} & \chi^2(\alpha; K - (p + q + P + Q)) \\ &= \chi^2(0,05; 36 - (2 + 1 + 0 + 1)) \\ &= \chi^2(0,05; 32) \\ &= 46,194 \end{aligned}$$

Untuk $K = 48$ diperoleh

$$\begin{aligned} Q &= 61(61 + 2) \sum_{k=1}^{48} \frac{(\hat{\rho}_k)^2}{61 - k} \\ &= 61(63) \left(\frac{(\hat{\rho}_1)^2}{61 - 1} + \dots + \frac{(\hat{\rho}_{48})^2}{61 - 48} \right) \\ &= 3843 \left(\frac{(-0,0878)^2}{60} + \dots + \frac{(-0,0097)^2}{13} \right) \\ &= 33,9765 \end{aligned}$$

dengan nilai

$$\begin{aligned} & \chi^2(\alpha; K - (p + q + P + Q)) \\ &= \chi^2(0,05; 48 - (2 + 1 + 0 + 1)) \\ &= \chi^2(0,05; 44) \\ &= 60,481 \end{aligned}$$

Kriteria pengujian:

Karena pada *lag* 12, 24, 36, dan 48 nilai $Q < \chi^2(\alpha; K - (p + q + P + Q))$, maka H_0 diterima artinya model memenuhi asumsi residual *white noise*.

LAMPIRAN E Lanjutan

Tabel E. Ringkasan Proses Uji *Ljung-Box*

| Model | Lag | <i>Ljung-Box</i> (<i>Q</i>) | $\chi^2_{(0,05;ab)}$ | <i>p-value</i> | Ket |
|---|-----|----------------------------------|----------------------|----------------|-----------------------------|
| SARIMA (2,1,1) (0,1,1) ⁶ | 12 | 9,1 | 15,507 | 0,247 | <i>White noise</i> |
| | 24 | 15,0 | 31,410 | 0,720 | |
| | 36 | 26,7 | 46,194 | 0,685 | |
| | 48 | 34,0 | 60,481 | 0,836 | |
| SARIMA (1,1,1) (1,1,0) ⁶ | 12 | 11,8 | 16,919 | 0,161 | <i>White noise</i> |
| | 24 | 18,3 | 32,671 | 0,569 | |
| | 36 | 27,7 | 47,400 | 0,684 | |
| | 48 | 47,5 | 61,656 | 0,332 | |
| SARIMA (1,1,1) (0,1,1) ⁶ | 12 | 13,4 | 16,919 | 0,098 | <i>White noise</i> |
| | 24 | 18,7 | 32,671 | 0,544 | |
| | 36 | 28,3 | 47,400 | 0,654 | |
| | 48 | 46,7 | 61,656 | 0,364 | |
| SARIMA (1,1,1) (0,1,0) ⁶ | 12 | 19,4 | 18,307 | 0,022 | Tidak <i>white noise</i> |
| | 24 | 22,8 | 33,924 | 0,356 | |
| | 36 | 36,2 | 48,602 | 0,320 | |
| | 48 | 62,2 | 62,830 | 0,045 | |
| SARIMA (0,1,2) (1,1,0) ⁶ | 12 | 8,6 | 16,919 | 0,377 | <i>White noise</i> |
| | 24 | 13,3 | 32,671 | 0,863 | |
| | 36 | 24,5 | 47,400 | 0,825 | |
| | 48 | 39,9 | 61,656 | 0,647 | |
| SARIMA (0,1,2) (0,1,1) ⁶ | 12 | 8,5 | 16,919 | 0,391 | <i>White noise</i> |
| | 24 | 12,5 | 32,671 | 0,896 | |
| | 36 | 22,2 | 47,400 | 0,903 | |
| | 48 | 34,8 | 61,656 | 0,840 | |
| SARIMA (0,1,2) (0,1,0) ⁶ | 12 | 12,9 | 18,307 | 0,166 | <i>White noise</i> |
| | 24 | 16,4 | 33,924 | 0,748 | |
| | 36 | 29,5 | 48,602 | 0,644 | |
| | 48 | 45,7 | 62,830 | 0,442 | |

LAMPIRAN E Lanjutan

| Model | Lag | Ljung-Box (Q) | $\chi^2_{(0,05;db)}$ | p-value | Ket |
|---|-----|------------------|----------------------|---------|----------------------------------|
| SARIMA (0,1,2) (0,1,1) ⁶ | 12 | 8,5 | 16,919 | 0,391 | <i>White noise</i> |
| | 24 | 12,5 | 32,671 | 0,896 | |
| | 36 | 22,2 | 47,400 | 0,903 | |
| | 48 | 34,8 | 61,656 | 0,840 | |
| SARIMA (0,1,2) (0,1,0) ⁶ | 12 | 12,9 | 18,307 | 0,166 | <i>White noise</i> |
| | 24 | 16,4 | 33,924 | 0,748 | |
| | 36 | 29,5 | 48,602 | 0,644 | |
| | 48 | 45,7 | 62,830 | 0,442 | |
| SARIMA (0,1,1) (0,1,0) ⁶ | 12 | 20,3 | 19,675 | 0,026 | <i>Tidak white noise</i> |
| | 24 | 24,1 | 35,172 | 0,341 | |
| | 36 | 36,5 | 49,802 | 0,352 | |
| | 48 | 67,8 | 64,001 | 0,020 | |
| SARIMA (0,1,0) (1,1,0) ⁶ | 12 | 15,0 | 19,675 | 0,133 | <i>White noise</i> |
| | 24 | 24,3 | 35,172 | 0,334 | |
| | 36 | 33,7 | 49,802 | 0,480 | |
| | 48 | 50,4 | 64,001 | 0,303 | |
| SARIMA (0,1,0) (0,1,1) ⁶ | 12 | 18,0 | 19,675 | 0,054 | <i>White noise</i> |
| | 24 | 27,0 | 35,172 | 0,210 | |
| | 36 | 35,6 | 49,802 | 0,391 | |
| | 48 | 49,5 | 64,001 | 0,336 | |

LAMPIRAN F

Uji *Kolmogorov-Smirnov* pada Model SARIMA Hasil *Overfitting* yang Memenuhi Uji *White Noise*

Uji *Kolmogorov-Smirnov* pada model SARIMA(2,1,1)(0,1,1)⁶

Hipotesis:

$H_0 : F(x) = F_0(x)$ (residual berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ (residual tidak berdistribusi normal)

Statistik uji:

$$D_{hitung} = \sup |S(x) - F_0(x)| \\ = 0,094$$

$$D_{tabel} = D_{\alpha;n} \\ = D_{0,05;59} \\ = 0,1771$$

Kriteria uji:

Karena $D_{hitung} < D_{tabel}$, maka H_0 diterima artinya residual berdistribusi normal.

Uji *Kolmogorov-Smirnov* juga dapat dilihat dari nilai *p-value* pada masing-masing model. Jika nilai *p-value* $> \alpha = 5\%$, maka model dikatakan memenuhi asumsi residual berdistribusi normal. Ringkasan hasil uji *Kolmogorov-Smirnov* dapat dilihat pada Tabel F.

Tabel F. Ringkasan Proses Uji *Kolmogorov-Smirnov*

| Model | D_{hitung} | p -value | Ket |
|---------------------------------------|--------------|------------|--------|
| SARIMA(2,1,1) (0,1,1) ⁶ | 0,096 | >0,150 | Normal |
| SARIMA(1,1,1) (1,1,0) ⁶ | 0,081 | >0,150 | Normal |
| SARIMA(1,1,1) (0,1,1) ⁶ | 0,100 | 0,144 | Normal |

LAMPIRAN F Lanjutan

| Model | D_{hitung} | p -value | Ket |
|---------------------------------------|--------------|------------|-----------------|
| SARIMA(0,1,2) (1,1,0) ⁶ | 0,091 | >0,150 | Normal |
| SARIMA(0,1,2) (0,1,1) ⁶ | 0,131 | 0,017 | Tidak Normal |
| SARIMA(0,1,2) (0,1,0) ⁶ | 0,073 | >0,150 | Normal |
| SARIMA(0,1,0) (1,1,0) ⁶ | 0,101 | 0,134 | Normal |
| SARIMA(0,1,0) (0,1,1) ⁶ | 0,109 | 0,082 | Normal |

LAMPIRAN G

Syntax Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ Transformasi ZQ1 dan ZQ2 di R Software

```

library(rstan)
library(ggplot2)
library(R2jags)
library(coda)
library(lattice)
library(boot)
library(R2jags)
library(rjags)
library(gsarima)

data<-read.csv("D://DataTA.csv", sep=";", header=T)
y<-data$A
N<-length(y)
#N<-68
#M<-72
yt<-as.ts(y)
t=c(1:68)
data2=cbind(t, yt)
data2

s<-6
w<-s+3
c<-1
p<-2
q<-2
inits<- list(list(r=2), list(r=4), list(r=8))
#untuk outsample
#parameters<-c(paste("y.p1[", w+1, ":", M, "]", sep=""),
paste("u[", w+1, ":", N, "]", sep=""))
#data<-list(N=N, w=w, p=p, c=c, y= yt[1:M], q=q, M=M)
parameters<-c("b0", "phi", "theta", "phi.star1",
"theta.star1", "r")
data<-list(N=N, w=w, p=p, c=c, y= y[1:N], q=q)
ptm <- proc.time()
jags.output.model211011x <- jags(data= data, inits,
parameters, model.file= "D://ZQ1Estimasi.txt",
n.iter=300000, n.burnin=1000, n.chains=3, n.thin=100)

```

LAMPIRAN G Lanjutan

jags.output.model211011x

```

model{
  #Prior
  b0~dbeta(1,1)
  r~dgamma(0.000001,0.000001)
  phi.star1<-0 #use this option in a model without a
first order seasonal autoregressive parameter
  theta.star1~dbeta(0.8036,1)
  theta[1]~dbeta(0.9166,1) #use this option if phi[1]
is included
  theta[2]<-0 #use this option in a model without
theta[2] parameter
  phi[1]~dbeta(0.6476,1) #use this option if phi[1]
is included
  phi[2]~dnorm(-0.3834,0.000001) #use this option if
phi[2] is included

  #Deviance
  Dev <- -2*sum(LL[(w+1): N])
  for (t in 1:w){
    u[t] <-0
  }

  #Likelihood
  for (t in (w+1):N){
    y[t]~dnegbin(pr[t],r)
    y.pl[t]~dnegbin(pr[t],r)
    pr[t]<-r/(r+lambda[t])
    lambda[t]<-exp(m[t])
    m[t]<-b0+ log(max(c,y[t-1]))+ log(max(c,y[t-6]))-
log(max(c,y[t-7]))+ sum(AR[ ,t])+ (theta.star1)*u[t-
6]+ theta[1]*u[t-1]- (theta[1]*theta.star1)*u[t-7]
    for (k in 1:p){
      AR[k,t]<-phi[k]*log(max(c,y[t-k]))-
phi[k]*log(max(c,y[t-k-1])) - phi[k]*log(max(c,y[t-k-
6])) +phi[k]*log(max(c,y[t-k-1-6]))
    }
    u[t] <- log(max(c,y[t])/lambda[t])
    LL[t] <-r*log(pr[t])+y[t]*log(1-
pr[t])+loggam(y[t]+r)-loggam(y [t]+1) - loggam(r)
  }
}

```

LAMPIRAN G Lanjutan

```

#peramalan outsample
for (t in (N+1):M){
  y.p1[t]~dnegbin(pr[t],r)
  pr[t]<-r/(r+lambda[t])
  lambda[t]<-exp(m[t])
  m[t]<-b0+ log(max(c,y[t-1]))+ log(max(c,y[t-6]))-
log(max(c,y[t-7]))+ sum(AR[ ,t])+ (theta.star1)*u[t-
6]+theta[1]*u[t-1]-(theta[1]*theta.star1)*u[t-7]
  for (k in 1:p){
    AR[k,t]<-phi[k]*log(max(c,y[t-k]))-
phi[k]*log(max(c,y[t-k-1]))- phi[k]*log(max(c,y[t-k-
6]))+ phi[k]*log(max(c,y[t-k-1-6]))
  }
  u[t] <- log(max(c,y[t])/lambda[t])
  LL[t] <-r*log(pr[t])+y[t]*log(1-
pr[t])+loggam(y[t]+r)-loggam(y [t]+1) - loggam(r)
}
}

model.mcmc<-as.mcmc(jags.output.model211011x)
summary(model.mcmc)

traceplot(jags.output.model211011x)
model.mcmc<-as.mcmc(jags.output.model211011x)
densityplot(model.mcmc)
autocorr.plot(model.mcmc)
summary(model.mcmc)

```


LAMPIRAN G Lanjutan

```

library(rstan)
library(ggplot2)
library(R2jags)
library(coda)
library(lattice)
library(boot)
library(R2jags)
library(rjags)
library(gsarima)

data<-read.csv("D://DataTA.csv", sep=";", header=T)
y<-data$A
N<-length(y)
#N<-68
#M<-72
yt<-as.ts(y)
t=c(1:68)
data2=cbind(t, yt)
data2

s<-6
w<-s+3
c<-1
p<-2
q<-2
inits<- list(list(r=2), list(r=4), list(r=8))

#untuk outsample
#parameters<-c(paste("y.p1[", w+1, ":", M, "]", sep=""),
paste("u[", w+1, ":", N, "]", sep=""))
#data<-list(N=N, w=w, p=p, c=c, y= yt[1:M], q=q, M=M)
parameters<-c("b0", "phi", "theta", "phi.star1",
"theta.star1", "r")
data<-list(N=N, w=w, p=p, c=c, y= y[1:N], q=q)
ptm <- proc.time()
jags.output.model211011x <- jags(data= data, inits,
parameters, model.file= "D://ZQ2Estimasi.txt",
n.iter=300000, n.burnin=1000, n.chains=3, n.thin=100)
jags.output.model211011x

```

LAMPIRAN G Lanjutan

```

model{
  #Prior
  b0~dbeta(1,1)
  r~dgamma(0.000001,0.000001)
  phi.star1<-0 #use this option in a model without a
  first order seasonal autoregressive parameter
  theta.star1~dbeta(0.8036,1)
  theta[1]~dbeta(0.9166,1) #use this option if phi[1]
  is included
  theta[2]<-0 #use this option in a model without
  theta[2] parameter
  phi[1]~dbeta(0.6476,1) #use this option if phi[1]
  is included
  phi[2]~dnorm(-0.3834,0.000001) #use this option if
  phi[2] is included

  #Deviance
  Dev <- -2*sum(LL[(w+1): N])
  for (t in 1:w){
    u[t] <-0
  }

  #Likelihood
  for (t in (w+1):N){
    y[t]~dnegbin(pr[t],r)
    y.pl[t]~dnegbin(pr[t],r)
    pr[t]<-r/(r+lambda[t])
    lambda[t]<-exp(m[t])
    m[t]<-b0+ log(y[t-1]+c)- log(y[t-6]+1)- log(y[t-
  7]+c)+ sum(AR[ ,t])+ theta[1]*u[t-1]+
  (theta.star1)*u[t-6]- theta[1]*(theta.star1)*u[t-7]
    for (k in 1:p){
      AR[k,t]<-phi[k]*log(y[t-k]+c)- phi[k]*log(y[t-
  k-1]+c)- phi[k]*log(y[t-k-6]+c)+ phi[k]*log(y[t-k-1-
  6]+c)
    }
    u[t] <- log((y[t]+c)/(lambda[t]+c))
    LL[t] <-r*log(pr[t])+y[t]*log(1-pr[t])+
  loggam(y[t]+r)- loggam(y[t]+1)- loggam(r)
  }
}

```

LAMPIRAN G Lanjutan

```

#peramalan outsample
for (t in (N+1):M){
  y[t]~dnegbin(pr[t],r)
  y.pl[t]~dnegbin(pr[t],r)
  pr[t]<-r/(r+lambada[t])
  lambda[t]<-exp(m[t])
  m[t]<-b0+ log(y[t-1]+c)- log(y[t-6]+1)- log(y[t-
7]+c)+ sum(AR[ ,t])+ theta[1]*u[t-1]+
(theta.star1)*u[t-6]- theta[1]*(theta.star1)*u[t-7]
  for (k in 1:p){
    AR[k,t]<-phi[k]*log(y[t-k]+c)- phi[k]*log(y[t-
k-1]+c)- phi[k]*log(y[t-k-6]+c)+ phi[k]*log(y[t-k-1-
6]+c)
  }
  u[t] <- log((y[t]+c)/(lambda[t]+c))
  LL[t] <-r*log(pr[t])+y[t]*log(1-pr[t])+
loggam(y[t]+r)- loggam(y[t]+1)- loggam(r)}
}

model.mcmc<-as.mcmc(jags.output.model211011x)
summary(model.mcmc)

traceplot(jags.output.model211011x)
model.mcmc<-as.mcmc(jags.output.model211011x)
densityplot(model.mcmc)
autocorr.plot(model.mcmc)
summary(model.mcmc)

```

LAMPIRAN H

Output Hasil Estimasi Parameter Model Binomial Negatif GSARIMA(2,1,1)(0,1,1)⁶ di R Software

Transformasi ZQ1

1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:

| | Mean | SD | Naive SE | Time-series SE |
|-------------|------------|---------|-----------|----------------|
| b0 | 3.005e-02 | 0.02045 | 0.0002640 | 0.0004151 |
| deviance | 6.510e+02 | 3.52732 | 0.0455375 | 0.0895226 |
| phi.star1 | 0.000e+00 | 0.00000 | 0.0000000 | 0.0000000 |
| phi[1] | 3.777e-02 | 0.04179 | 0.0005395 | 0.0013527 |
| phi[2] | -3.509e-01 | 0.12360 | 0.0015956 | 0.0021009 |
| r | 30.81848 | 6.50417 | 0.0839684 | 0.1163786 |
| theta.star1 | 3.815e-02 | 0.03994 | 0.0005157 | 0.0012019 |
| theta[1] | 4.778e-02 | 0.04603 | 0.0005942 | 0.0012354 |
| theta[2] | 0.000e+00 | 0.00000 | 0.0000000 | 0.0000000 |

2. Quantiles for each variable:

| | 2.5% | 25% | 50% | 75% | 97.5% |
|-------------|------------|------------|------------|------------|------------|
| b0 | 1.586e-03 | 0.013923 | 0.026824 | 4.304e-02 | 7.689e-02 |
| deviance | 6.461e+02 | 648.444166 | 650.397266 | 6.530e+02 | 6.595e+02 |
| phi.star1 | 0.000e+00 | 0.000000 | 0.000000 | 0.000e+00 | 0.000e+00 |
| phi[1] | 1.710e-04 | 0.006572 | 0.023587 | 5.573e-02 | 1.470e-01 |
| phi[2] | -5.925e-01 | -0.434104 | -0.353071 | -2.670e-01 | -1.060e-01 |
| r | 1.946e+01 | 26.21636 | 30.28517 | 34.74161 | 44.91152 |
| theta.star1 | 3.789e-04 | 0.009440 | 0.025296 | 5.325e-02 | 1.483e-01 |
| theta[1] | 9.009e-04 | 0.013195 | 0.033990 | 6.833e-02 | 1.686e-01 |
| theta[2] | 0.000e+00 | 0.000000 | 0.000000 | 0.000e+00 | 0.000e+00 |

Transformasi ZQ2

1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:

| | Mean | SD | Naive SE | Time-series SE |
|-------------|------------|---------|-----------|----------------|
| b0 | 0.029163 | 0.02059 | 0.0002658 | 0.0004569 |
| deviance | 651.034801 | 3.78733 | 0.0488943 | 0.1070259 |
| phi.star1 | 0.000000 | 0.00000 | 0.0000000 | 0.0000000 |
| phi[1] | 0.041991 | 0.04698 | 0.0006065 | 0.0016703 |
| phi[2] | -0.348608 | 0.12370 | 0.0015969 | 0.0021319 |
| r | 30.94755 | 6.50639 | 0.0839971 | 0.118702 |
| theta.star1 | 0.039304 | 0.03875 | 0.0005002 | 0.0011886 |
| theta[1] | 0.048967 | 0.04893 | 0.0006316 | 0.0014520 |
| theta[2] | 0.000000 | 0.00000 | 0.0000000 | 0.0000000 |

LAMPIRAN H Lanjutan

2. Quantiles for each variable:

| | 2.5% | 25% | 50% | 75% | 97.5% |
|-------------|------------|------------|------------|------------|------------|
| b0 | 1.385e-03 | 0.012963 | 0.025771 | 0.041228 | 7.781e-02 |
| deviance | 6.458e+02 | 648.279127 | 650.321074 | 652.999207 | 6.601e+02 |
| phi.star1 | 0.000e+00 | 0.000000 | 0.000000 | 0.000000 | 0.000e+00 |
| phi[1] | 2.280e-04 | 0.008077 | 0.025238 | 0.060161 | 1.702e-01 |
| phi[2] | -5.911e-01 | -0.429342 | -0.347788 | -0.267458 | -1.027e-01 |
| r | 1.984e+01 | 26.256596 | 30.45149 | 35.08597 | 45.05279 |
| theta.star1 | 5.690e-04 | 0.010320 | 0.027213 | 0.056275 | 1.463e-01 |
| theta[1] | 8.869e-04 | 0.013316 | 0.033862 | 0.068266 | 1.812e-01 |
| theta[2] | 0.000e+00 | 0.000000 | 0.000000 | 0.000000 | 0.000e+00 |

BIODATA PENULIS



Windya Harieska Pramujati atau biasa dipanggil Windy lahir di Madiun tanggal 23 Juni 1998. Pendidikan formal yang pernah ditempuh yaitu RA Al-Kautsar Tirtamulya, SD Negeri Banaran, SMP Negeri 1 Balerejo dan SMA Negeri 1 Mejayan. Kemudian penulis melanjutkan pendidikan S1 di Departemen Matematika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember, Surabaya dengan bidang Matematika Terapan khususnya Riset Operasi dan Pengolahan Data (ROPD). Selama kuliah, penulis aktif di organisasi HIMATIKA ITS, Lembaga Dakwah Jurusan IBNU MUQLAH dan Unit Kegiatan Mahasiswa TDC ITC. Pada tahun 2017-2018 penulis menjadi anggota aktif HIMATIKA ITS sebagai Staff *Entrepreneur Development*, Staff Kemuslimahan IBNU MUQLAH, dan Staff *Human Resources Development* TDC ITS. Pada tahun 2018-2019 penulis tetap aktif di organisasi tersebut dengan mengemban amanah sebagai Ketua Departemen *Entrepreneur Development* HIMATIKA ITS dan Staff Dana Usaha IBNU MUQLAH. Demikian biodata tentang penulis. Jika ingin memberikan saran, kritik, dan diskusi mengenai Tugas Akhir ini, dapat dikirimkan melalui email windyaharieska@gmail.com. Terimakasih.