

**PENERAPAN MODEL *MARKOV SWITCHING AUTOREGRESSIVE*
(MSAR) PADA PERAMALAN JUMLAH PENUMPANG
KERETA API DI SUMATERA**

(Skripsi)

Oleh

SILVIA SINTA SARI



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025**

ABSTRACT

PENERAPAN MODEL *MARKOV SWITCHING AUTOREGRESSIVE* (MSAR) PADA PERAMALAN JUMLAH PENUMPANG KERETA API DI SUMATERA

By

SILVIA SINTA SARI

This study aims to apply the Markov Switching Autoregressive (MSAR) model in modeling and predicting data on the number of train passengers in Sumatera. The MSAR model is a nonlinear autoregressive model that can experience changes in fluctuations that occur in the data, form a dynamic model with respect to data patterns, can determine transition opportunities or known as state switching opportunities, and determine the average duration in each state. The data used is monthly data on the number of train passengers in Sumatera for the period January 2010 to July 2024. MSAR modeling obtained as many as 5 models with 2 states. In addition, the MSAR model has state variables and the probability value of the transition matrix calculated using the Maximum Likelihood estimation (MLE) method. Parameter estimation is performed on all models and evaluated based on the Akaike's Information Criterion corrected (AICc) and Mean Absolute Percentage Error (MAPE) values, with the aim of finding the model for the best forecasting results. The best MSAR model is MS(2)AR(4) which has an AICc value of 1919,0440 and a MAPE value of 3,89%. The model and forecasting results obtained state 1 (increase) and state 2 (decrease). Analysis of forecasting results for the next 12 months that the data on the number of train passengers in Sumatera as a whole shows that the data is stable from August 2024 - April 2025, showing a slow upward trend in May 2025 and June 2025 and a drastic increase in July 2025.

Keywords: Number of Train Passengers, Forecasting, Nonlinear, State, MSAR

ABSTRAK

PENERAPAN MODEL *MARKOV SWITCHING AUTOREGRESSIVE* (MSAR) PADA PERAMALAN JUMLAH PENUMPANG KERETA API DI SUMATERA

Oleh

SILVIA SINTA SARI

Penelitian ini bertujuan untuk menerapkan model *Markov Switching Autoregressive* (MSAR) dalam memodelkan dan memprediksi data jumlah penumpang kereta api di Sumatera. Model MSAR merupakan model *autoregressive* nonlinier yang dapat mengalami perubahan fluktuasi yang terjadi pada data, membentuk model dinamis sehubungan dengan pola data, dapat menentukan peluang transisi atau dikenal sebagai peluang perpindahan *state*, dan menentukan lama durasi rata-rata dalam setiap *state*. Data yang digunakan adalah data bulanan data jumlah penumpang kereta api di Sumatera periode Januari 2010 sampai Juli 2024. Pemodelan MSAR diperoleh sebanyak 5 model dengan 2 *state*. Selain itu, Model MSAR mempunyai variabel *state* dan nilai peluang matriks transisi yang dihitung dengan menggunakan metode *Maximum Likelihood estimation* (MLE). Estimasi parameter dilakukan pada seluruh model dan dievaluasi berdasarkan nilai *Akaike's Information Criterion corrected* (AICc) dan *Mean Absolute Percentage Error* (MAPE), dengan tujuan menemukan model untuk hasil peramalan yang terbaik. Adapun model MSAR terbaik yaitu MS(2)AR(4) yang memiliki nilai AICc sebesar 1919,0440 dan nilai MAPE sebesar 3,89%. Model dan hasil peramalan diperoleh kondisi *state* 1 (peningkatan) dan *state* 2 (penurunan). Analisis hasil peramalan untuk 12 bulan ke depan bahwa data jumlah penumpang kereta api di Sumatera secara keseluruhan menunjukkan bahwa data stabil dari bulan Agustus 2024 - April 2025, menunjukkan adanya tren naik secara perlahan pada bulan Mei 2025 dan Juni 2025 serta naik secara drastis pada bulan Juli 2025.

Kata Kunci : Jumlah Penumpang Kereta Api, Peramalan, Nonlinear, *State*, MSAR

**PENERAPAN MODEL *MARKOV SWITCHING AUTOREGRESSIVE*
(MSAR) PADA PERAMALAN JUMLAH PENUMPANG
KERETA API DI SUMATERA**

Oleh

SILVIA SINTA SARI

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025**

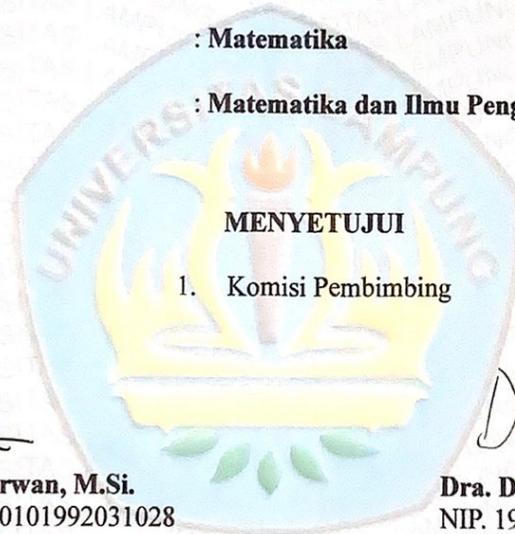
Judul Skripsi : **PENERAPAN MODEL *MARKOV SWITCHING*
AUTOREGRESSIVE (MSAR) PADA
PERAMALAN JUMLAH PENUMPANG
KERETA API DI SUMATERA**

Nama Mahasiswa : **Silvia Sinta Sari**

Nomor Pokok Mahasiswa : **2117031007**

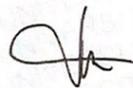
Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

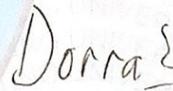


MENYETUJUI

1. **Komisi Pembimbing**

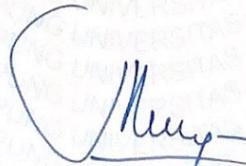


Drs. Nusyirwan, M.Si.
NIP. 196610101992031028



Dra. Dorrah Aziz, M.Si.
NIP. 196101281988112001

2. **Ketua Jurusan Matematika**

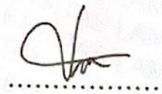


Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

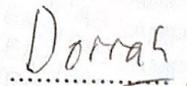
MENGESAHKAN

1. Tim Penguji

Ketua : **Drs. Nusyirwan, M.Si.**



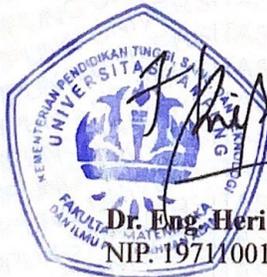
Sekretaris : **Dra. Dorrah Aziz, M.Si.**



Penguji
Bukan Pembimbing : **Dr. Khoirin Nisa, S.Si., M.Si.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.

NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : **31 Januari 2025**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan dibawah ini : **SILVIA SINTA SARI**
Nomor Pokok Mahasiswa : **2117031007**
Jurusan : **Matematika**
Judul Skripsi : **PENERAPAN MODEL *MARKOV SWITCHING AUTOREGRESSIVE (MSAR)* PADA PERAMALAN JUMLAH PENUMPANG KERETA API DI SUMATERA**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku.

Bandar Lampung, 10 Februari 2025
Penulis,



Silvia Sinta Sari
NPM. 2117031007

RIWAYAT HIDUP

Penulis bernama Silvia Sinta Sari lahir di Mandala Sari pada 07 Agustus 2003. Penulis merupakan anak pertama dari 2 bersaudara pasangan Bapak Suhartono dan Ibu Helena Puspa Hayati.

Penulis mengawali pendidikan di Raudhatul Athfal (RA) Nurul Iman Mandala Sari pada tahun 2008 – 2009 dan menempuh pendidikan Sekolah Dasar di SDN Mandala Sari pada tahun 2009 – 2015. Kemudian penulis melanjutkan pendidikan Sekolah Menengah Pertama di SMP Negeri 1 Bandar Sribhawono pada tahun 2015 – 2018 dan Sekolah Menengah Atas di SMAN 1 Bandar Sribhawono pada tahun 2018 – 2021. Setelah itu, penulis diterima sebagai mahasiswa Program Studi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui Jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN).

Selama menjadi mahasiswa, penulis aktif di beberapa kegiatan di antaranya aktif dalam kepengurusan organisasi Himpunan Mahasiswa Jurusan Matematika (HIMATIKA) FMIPA Unila sebagai anggota Bidang Minat dan Bakat serta menjadi Bendahara Pelaksana Himatika Competition pada tahun 2022. Selain itu, penulis juga aktif dalam organisasi Ikatan Mahasiswa Lampung Timur sebagai anggota Departemen Sosial Masyarakat pada tahun 2022. Selain aktif di organisasi, penulis juga pernah menjadi asisten dosen pada mata kuliah Aljabar Matriks Statistika pada tahun 2024. Pada bulan Desember 2023 – Februari 2024, penulis melaksanakan Kerja Praktik (KP) di Dinas Komunikasi, Informatika dan Statistik Provinsi Lampung pada Bidang Persandian dan Statistik yang berlokasi di

Jl. Wolter Monginsidi N0.69, Telukbetung Selatan, Bandar Lampung dan menulis laporan yang berjudul “ Analisis Regresi Data Panel dengan Pendekatan *Common Effect Model* (CEM), *Fixed Effect Model* (FEM) dan *Random Effect Model* (REM) Pada Persentase Penduduk Miskin Menurut Kabupaten/Kota di Provinsi Lampung Tahun 2019 – 2023 ”. Selanjutnya, pada bulan Juni – Agustus 2024, penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Braja Asri, Kecamatan Way Jepara, Kabupaten Lampung Timur, Provinsi Lampung. Kemudian sebagai bentuk aplikasi bidang ilmu di dunia kerja, penulis telah menyelesaikan skripsi yang berjudul “ Penerapan Model *Markov Switching Autoregressive* (MSAR) Pada Peramalan Jumlah Penumpang Kereta Api di Sumatera ”.

MOTTO

“ Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya ”

(QS. Al-Baqarah: 286)

“ Dan bersabarlah kamu, sesungguhnya janji Allah itu benar ”

(QS. Ar-Rum: 60)

“ Hidup bukan saling mendahului, bermimpilah sendiri-sendiri ”

(Daniel Baskara Putra Mahendra aka Hindia)

*“ Mengapa takut pada lara, sementara semua rasa bisa kita cipta, akan selalu
ada tenang disela-sela gelisah yang menunggu reda ”*

(Payung Teduh)



PERSEMBAHAN

Alhamdulillahirabbil'alamiin, puji Syukur kepada Allah SWT, atas nikmat dan hidayah-Nya kepada penulis untuk dapat menyelesaikan skripsi ini dengan baik.

Rasa syukur dan bahagia penulis persembahkan skripsi ini kepada:

Diri Sendiri

Kepada diri saya sendiri. Silvia Sinta Sari. Terima kasih sudah bertahan sejauh ini dan tetap memilih berusaha serta merayakan diri sendiri sampai di titik ini, walau sering kali merasa putus asa. Terima kasih sudah menjadi manusia yang selalu mau berusaha dan tidak lelah mencoba. Terima kasih telah menyelesaikan apa yang sudah dimulai dan tidak menyerah sesulit apapun proses penyusunan skripsi ini. Berbahagialah selalu dimana pun kamu berada, Silvia. Apapun kurang dan lebihmu mari tetap merayakan diri sendiri.

Kedua Orang Tua

Kepada Bapak Suhartono dan Mama Helena Puspa Hayati, terima kasih telah tulus mendoakan tiada henti, selalu memberikan kasih sayang, cinta, dukungan, motivasi dan semangat. Terima kasih sudah menciptakan keluarga yang harmonis, selalu memberi pelukan hangat dan menjadi tempat pulang yang selalu mendengar dan tidak pernah menghakimi. Terima kasih selalu mempercayai penulis diatas keraguan orang lain. Menjadi suatu kebanggaan memiliki orang tua yang

mendukung anaknya untuk mencapai cita-cita. Terima kasih sudah membuktikan kepada dunia bahwa anak petani bisa menjadi sarjana.

Adik Tersayang

Kepada adikku Rhaditia Tata Pradita yang selalu menjadi alasan penulis untuk lebih keras lagi dalam berjuang. Terima kasih selalu mendoakan, menyemangati, mendukung, menghibur, mendengarkan keluh kesah, dan selalu ada. Terima kasih selalu kebersamai dan tetaplah tumbuh menjadi adik yang baik.

Dosen

Kepada bapak dan ibu dosen, terima kasih telah memberikan ilmunya selama penulis berkuliah di Jurusan Matematika FMIPA Unila, khususnya kepada Bapak Drs. Nusyirwan, M.Si., Ibu Dra. Dorrah Aziz, M.Si., dan Ibu Dr. Khoirin Nisa, S.Si., M.Si. yang telah mendukung penulis dalam menyelesaikan skripsi.

Almamater Tercinta

Universitas Lampung

SANWACANA

Dengan mengucapkan puji syukur kepada Allah SWT, yang telah memberikan kekuatan, ketabahan, serta petunjuk dan tuntunan sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam semoga senantiasa kita haturkan kepada junjungan dan suri tauladan seluruh umat manusia, Nabi Muhammad SAW. Semoga kita mendapatkan syafaatnya di hari akhir kelak.

Skripsi dengan judul “ **Penerapan Model *Markov Swiching Autoregressive (MSAR)* dalam Peramalan Jumlah Penumpang Kereta Api di Sumatera** ” dibuat sebagai salah satu syarat untuk meraih gelar sarjana matematika (S.Mat) di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung. Penulis menyadari masih terdapat banyak kekurangan dan kesalahan dalam penyusunan skripsi ini. Dengan terselesaikannya skripsi ini penulis mengucapkan terima kasih kepada:

1. Kedua orang tua tercinta, Bapak Suhartono dan Ibu Helena Puspa Hayati yang selalu menjadi penyemangat penulis dan sandaran terkuat dari kerasnya dunia, yang tiada henti memberikan kasih sayang dengan penuh cinta, tak kenal lelah mendoakan serta memberi perhatian dan dukungan kepada anak-anaknya.
2. Adik laki-lakiku, Rhaditia Tata Pradita yang menjadi teman tumbuh, memberi semangat, dukungan, motivasi, dan doa. Semoga kita tetap tumbuh dan tangguh dalam menghadapi dinamika kehidupan yang tidak menentu.
3. Bapak Drs. Nusyirwan, M.Si., selaku Pembimbing 1 yang telah sabar dan ikhlas dalam memberikan ilmu, bimbingan, motivasi, kritik/saran, dan bantuan selama perkuliahan maupun dalam penyusunan skripsi ini.

4. Ibu Dra. Dorrah Aziz, M.Si., selaku Pembimbing II yang telah sabar dan ikhlas dalam memberikan ilmu, bimbingan, dukungan, motivasi, kritik/saran, dan bantuan selama perkuliahan maupun dalam penyusunan skripsi ini.
5. Ibu Dr. Khoirin Nisa, S.Si., M.Si., selaku Pembahas yang telah sabar dan senantiasa memberikan masukan dan arahan dalam menyelesaikan skripsi ini.
6. Bapak Prof. Drs. Mustofa, M.A., Ph.D., selaku Dosen Pembimbing Akademik yang senantiasa memberikan motivasi dan membimbing selama menjalani perkuliahan.
7. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika FMIPA Universitas Lampung.
8. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu pengetahuan Alam, Universitas Lampung.
9. Ibu Prof. Dr. Ir. Lusmeilia Afriani, DEA, IPM., selaku Rektor Universitas Lampung.
10. Seluruh Dosen, Staf, dan Karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
11. Rosa Amanda Putri dan Syifa Arindia Septiana, yang telah kebersamai penulis dalam keadaan susah dan senang, memberikan doa, dukungan serta motivasi. Terima kasih telah bersedia berjalan bersama, semoga tetap menjadi teman tumbuh di segala kondisi yang kadang tak terduga.
12. Teman-teman Istri Konglomerat, KKN Braja Asri, KP DISKOMINFOTIK, Mikat'22, Sirkel Bimbingan dan teman-teman seperjuangan yang senantiasa memberi semangat, dukungan, bantuan, motivasi dan doa.
13. Teman-teman mahasiswa Jurusan Matematika, terima kasih atas kebersamaan, pengalaman, bantuan, dukungan, dan kisah yang telah kalian berikan selama menempuh pendidikan dan penyelesaian penyusunan skripsi. Semoga kita semua bisa sukses sesuai cita-cita masing-masing dan bermanfaat bagi lingkungan sekitar.
14. Mochi, Olio dan Kembar selaku kucing peliharaan yang selalu menghibur dan menjadi teman penulis di rumah.

15. Lagu Danilla Riyadi, For Revenge, Fourtwnty, Hindia, Hivi, Juicy Luicy, Maliq & D'Essentials, My Chemical Romance, Neckdeep, Payung Teduh, Reality Club, Rumahsakit dan lagu lainnya yang menemani penulis setiap saat.
16. Konser musik yang telah hadir di sela-sela kesibukan yang menjadi penghibur bagi penulis.
17. Seseorang yang nantinya akan bersama penulis dan menjadi pasangan hidup yang belum diketahui orangnya siapa, ini sebagai suatu usaha penulis memproseskan dan menginvestasikan diri dalam hal ilmu, pemikiran dan wawasan agar saat bersama kita bisa menyetarakan dalam segi ilmu, pemikiran, wawasan dan susunan kehidupan lain nantinya.
18. Semua pihak yang tidak bisa disebutkan satu persatu yang telah memberi pengalaman, semangat, motivasi, dan doa kepada penulis.
19. Terakhir, terima kasih untuk diri sendiri, karena telah mampu berusaha keras dan berjuang sejauh ini, mampu mengendalikan diri dari tekanan luar dan mampu berdiri tegak ketika dihantam permasalahan yang ada. Terima kasih pada hati yang tetap tegar dan ikhlas menjalani dinamika kehidupan. Semoga tetap rendah hati, ini baru awal dari permulaan hidup, semangat.

Semoga Allah SWT mamberikan keberkahan dan membalas kebaikan kepada semua pihak yang telah membantu penulis secara langsung maupun tidak langsung. Penulis menyadari masih banyak kekurangan dalam penulisan skripsi ini, kritik dan saran yang membangun sangat diharapkan. Semoga skripsi ini berguna dan bermanfaat bagi kita semua.

Bandar lampung, 10 Februari 2025

Penulis,

Silvia Sinta Sari

DAFTAR ISI

	Halaman
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xix
I. PENDAHULUAN	
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian	3
II. TINJAUAN PUSTAKA	
2.1 Analisis Deret Waktu	4
2.2 <i>Autoregressive</i> (AR)	4
2.3 Stasioneritas	5
2.4 Transformasi Box-Cox	6
2.5 Uji <i>Augmented Dickey-Fuller</i> (ADF)	7
2.6 Pembeda (<i>Differencing</i>)	8
2.7 Fungsi Autokorelasi dan Fungsi Autokorelasi Parsial	
2.7.1 Fungsi Autokorelasi (ACF)	9
2.7.2 Fungsi Parsial Autokorelasi (PACF)	10
2.8 Uji Nonlinieritas	14
2.9 <i>State</i>	15
2.10 <i>Markov Switching</i>	15
2.11 Model <i>Markov Switching Autoregressive</i> (MSAR)	
2.11.1 Peluang Perpindahan <i>State</i>	16
2.11.2 Rata-rata Jangka Waktu pada Masing-masing <i>State</i>	17
2.12 Peluang (<i>Probability</i>)	17
2.13 Matriks (<i>Matrix</i>)	18
2.14 Matriks Transisi (<i>Transition Matriks</i>)	19
2.15 Estimasi Parameter Model <i>Markov Switching Autoregressive</i> (MSAR)	
2.15.1 <i>Filtering</i>	21
2.15.2 <i>Smoothing</i>	24
2.16 Pemilihan Model Terbaik	
2.16.1 <i>Akaike's Information Criterion corrected</i> (AICc)	31

2.16.2	<i>Mean Absolute Percentage Error (MAPE)</i>	31
2.17	Pemeriksaan Diagnostik	
2.17.1	Uji Normalitas	33
2.17.2	Uji Asumsi	33
2.17.3	Uji Heteroskedastisitas	34
2.17.4	Uji Autokorelasi.....	35
2.18	Jumlah Penumpang Kereta Api	35
III. METODOLOGI PENELITIAN		
3.1	Waktu dan Tempat Penelitian	37
3.2	Data Penelitian.....	37
3.3	Metode Penelitian	37
IV. HASIL DAN PEMBAHASAN		
4.1	Analisis Deskriptif.....	40
4.2	Uji Stasioneritas.....	40
4.3	Uji Nonlinieritas	47
4.4	Identifikasi Model	48
4.5	Estimasi Parameter	49
4.6	Pemilihan Model Terbaik	54
4.7	Uji Diagnostik	
4.7.1	Uji Normalitas	57
4.7.2	Uji Asumsi	58
4.7.3	Uji Heteroskedastisitas	59
4.7.4	Uji Autokorelasi.....	60
4.8	Peramalan	61
V.	KESIMPULAN	63
	DAFTAR PUSTAKA	64
	LAMPIRAN	67

DAFTAR TABEL

Tabel	Halaman
1. Transformasi Box-Cox	6
2. Kriteria nilai MAPE	32
3. Deskriptif data	40
4. Uji ADF dari data	43
5. Uji ADF dari data <i>differencing</i>	46
6. Uji Nonlinieritas Teräsvirta	47
7. Estimasi parameter model MS(2)AR(1)	49
8. Probabilitas transisi model MS(2)AR(1)	49
9. Estimasi parameter model MS(2)AR(2)	50
10. Probabilitas transisi model MS(2)AR(2)	50
11. Estimasi parameter model MS(2)AR(3)	51
12. Probabilitas transisi model MS(2)AR(3)	51
13. Estimasi parameter model MS(2)AR(4)	52
14. Probabilitas transisi model MS(2)AR(4)	52
15. Estimasi parameter model MS(2)AR(5)	53
16. Probabilitas transisi model MS(2)AR(5)	54
17. Perbandingan nilai AICc dan MAPE	54
18. Hasil uji Kolmogorov-Smirnov	57

19. Hasil uji Breusch-Pagan.....	59
20. Hasil uji Ljung-Box	60
21. Peramalan model MSAR pada data Jumlah Penumpang Kereta Api di Sumatera	61

DAFTAR GAMBAR

Gambar	Halaman
1. Diagram Alir Pemodelan MSAR	39
2. Plot data Jumlah Penumpang Kereta Api di Sumatera	41
3. (a) Plot ACF dan (b) Plot ACF Jumlah Penumpang	42
4. Plot Transformasi Box-Cox	43
5. Plot differencing data Jumlah Penumpang Kereta Api di Sumatera	44
6. (a) Plot ACF dan (b) Plot PACF Data <i>Differencing</i>	45
7. Plot Transformasi Box-Cox Setelah <i>Differencing</i>	46
8. Plot Matriks Peluang Transisi	56
9. Plot Model MS(2)AR(4)	58
10. Plot Peramalan Model MSAR Pada Data Jumlah Penumpang Kereta	62

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Salah satu faktor yang mendukung mobilitas di suatu wilayah yaitu transportasi. Di Indonesia, banyak perusahaan (Persero) yang mengelola transportasi, baik darat, laut maupun udara. Menurut Badan Pusat Statistik Provinsi Lampung, jumlah penumpang yang menggunakan transportasi umum pada bulan Juni 2024 sebanyak 179.086 orang karena terjadi adanya peningkatan 8,83 % dari jumlah penumpang sebanyak 164.552 orang pada bulan Mei 2024. Transportasi darat terbagi menjadi beberapa jenis, salah satunya yaitu kereta api. Kereta api menjadi transportasi umum yang bersifat massal yang dapat mengangkut penumpang dalam satu perjalanan sehingga lebih efisien dibandingkan dengan alat angkut lainnya (Mirza, 2023). Kereta api menjadi transportasi darat yang memiliki jumlah penumpang tertinggi diantara transportasi darat yang lain. Kereta api menjadi pilihan utama dalam memilih transportasi karena mampu menawarkan harga murah, durasi perjalanan, dan fasilitas yang disediakan cukup kompetitif (Hasanah, *et al.*, 2024).

Data jumlah penumpang kereta api di Sumatera dari tahun 2010 hingga 2024 disajikan dalam format runtun waktu atau *time series*. Tiga model klasik yang sering digunakan dalam pemodelan *time series* yaitu *Autoregressive Integrated Moving Average* (ARIMA), *Autoregressive Conditional Heteroscedastic* (ARCH), dan *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Model klasik tersebut cukup baik, namun tidak memperhitungkan adanya perubahan

struktur atau bentuk (Khoerunnisa, *et al.*, 2022). Alternatif model *time series* yang digunakan dalam menganalisis dan menyelidiki data yang mengalami perubahan struktur dimodelkan dengan Model *Markov Switching Analysis* (MSA). *Switching Markov* merupakan model *time series* nonlinier yang sering digunakan. Model ini sering disebut model *switching rezim* sebagai alternatif pemodelan deret waktu yang digunakan untuk melihat kondisi atau peralihan *regime/state*. Dalam model peralihan *Markov*, rantai *Markov* memiliki sifat menetapkan nilai variabel status yang mengontrol peralihan struktural atau fluktuasi data berdasarkan nilai sebelumnya. Melalui proses pertukaran, struktur lain menggantikan struktur yang telah berubah dari waktu ke waktu (Rizki, *et al.*, 2024).

Data jumlah penumpang kereta api dapat mengalami fluktuasi yang kemungkinan terjadi karena masyarakat belum cukup puas terhadap layanan jasa kereta api (Effendi & Yenita, 2023). Fluktuasi juga dapat disebabkan oleh musim liburan atau hari raya, kondisi ekonomi, kebijakan pemerintah serta faktor lain yang mengubah nilai data. Perubahan kondisi fluktuasi pada data penumpang kereta api umumnya memiliki 2 macam perubahan, yaitu apresiasi (peningkatan jumlah penumpang kereta api) dan depresiasi (penurunan jumlah penumpang kereta api). Dibutuhkan teknik yang tepat untuk menganalisis data *time series* dengan mempertimbangkan fluktuasi yang terjadi karena perubahan kondisi fluktuasi seringkali tidak diperhatikan atau diabaikan.

Salah satu metode alternatif yang dapat digunakan adalah *Markov switching* karena metode tersebut dapat menangkap dinamika pergerakan yang lebih kompleks dan menghitung probabilitas perubahan kondisi dan durasi pada setiap kondisi. Model *Markov Switching* dikombinasikan dengan model *Autoregressive* (AR) yang disebut model *Markov Switching Autoregressive* (MSAR). Model MSAR dapat digunakan untuk membentuk model dinamis sehubungan dengan pola data, dapat menentukan peluang transisi atau dikenal sebagai peluang perpindahan *state*, dan menentukan lama durasi rata-rata dalam setiap *state*.

Sehingga, dengan asumsi *autoregressive* yang beragam, model MSAR mampu memodelkan *time series* nonlinier atau *time series* tidak normal berdasarkan pada keadaan yang sesuai dengan pergantian rantai *Markov* (Rizki, *et al.*, 2024).

Penelitian ini akan membahas model MSAR yang diterapkan dalam data jumlah penumpang kereta api Sumatera dari tahun 2010 hingga 2024 karena dianggap pergerakan jumlah penumpang kereta api memiliki perubahan kondisi fluktuasi yaitu apresiasi dan depresiasi. Pada penelitian ini akan mencari model terbaik untuk data, menghitung peluang transisi (peluang untuk suatu *state* berpindah atau bertahan), dan menghitung lama durasi bertahan rata-rata untuk setiap *state*.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengidentifikasi model terbaik dari data.
2. Mengidentifikasi kemungkinan data akan bertahan dalam kondisi apresiasi atau depresiasi, kondisi depresiasi berpindah ke apresiasi atau apresiasi berpindah ke depresiasi.
3. Mengidentifikasi rata-rata lama durasi pada kondisi apresiasi bertahan maupun depresiasi bertahan dari data.

1.3 Manfaat Penelitian

Penelitian ini bermanfaat sebagai:

1. Mempelajari lebih dalam tentang penggunaan model MSAR sebagai salah satu alternatif dalam *time series*.
2. Menemukan informasi tentang kemungkinan kondisi fluktuasi (naik turunnya jumlah penumpang kereta api di Sumatera) yang terjadi.

II. TINJAUAN PUSTAKA

2.1 Analisis Deret Waktu

Peramalan merupakan cara yang digunakan dalam menentukan sikap atas situasi ke depan dengan lebih baik dan terperinci di masa depan untuk meminimalkan tingkat kesalahan dengan menggunakan kumpulan data historis dari periode sebelumnya hingga saat ini (Ahmad, 2020). Peramalan menjadi dasar utama dalam analisis deret waktu. Menurut Wei (2006), suatu variabel yang diambil secara berurutan dalam jangka waktu tertentu pada serangkaian observasi disebut *time series*. Data pengamatan *time series* direpresentasikan dengan Y_t dimana t adalah indeks waktu urutan pengamatan.

2.2 Autoregressive (AR)

Misalkan terdapat pengamatan *time series* yang bergantung pada p dimana p merupakan waktu terakhir, maka distribusi bersyarat pada Y_t dapat ditulis

$$f(y_t|Y_t) = f(y_t|y_{t-1}, \dots, y_{t-p})$$

Jika ketergantungan diasumsikan hingga rata-rata bersyarat dari Y_t merupakan fungsi linier dari y_{t-1}, \dots, y_{t-p} dan varians bersyarat dari Y_t adalah konstan maka Y_t dapat ditunjukkan oleh model linier orde p yang dikenal dengan model *autoregressive* orde p (AR(p)). Model *autoregressive* adalah representasi acak dari

proses yang bergantung secara linier pada nilai sebelumnya. Model *autoregressive* orde p sebagai berikut:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (2.1)$$

dengan

$\phi_1, \phi_2, \dots, \phi_p$ = parameter *autoregressive*

ε_t = residual yang bersifat *white noise*

White noise harus memenuhi tiga asumsi, yaitu:

1. $E[\varepsilon_t] = 0$ $t = 1, 2, \dots, T$
2. $E[\varepsilon_t^2] = \sigma^2$ $t = 1, 2, \dots, T$
3. $E[\varepsilon_s \varepsilon_t] = 0$ $s, t = 1, 2, \dots, T$ dan $s \neq t$

Berdasarkan model yang diasumsikan untuk Y_t maka ε_t tidak diamati dan harus diperkirakan dari data (Franses, *et all.*, 2014).

2.3 Stasioneritas

Menurut Wei (2006), suatu data dianggap stasioner jika tidak terjadi perubahan yang signifikan atau ekstrim, yaitu ketika rata-rata data setiap periode pada ragamnya tetap stabil. Kestasioneran diperlukan untuk memperkecil kekeliruan model pada suatu data yang berhubungan dengan metode estimasi yang digunakan. Stasioneritas dibagi menjadi 2 jenis sebagai berikut:

1. Stasioneritas dalam varians terjadi ketika struktur data runtun waktu memiliki data yang tidak berubah dan konstan dari waktu ke waktu. Stasioneritas dalam varians dapat dianalisis menggunakan Transformasi Box-Cox.
2. Stasioneritas dalam rata-rata berarti bahwa fluktuasi data tetap pada nilai rata-rata konstan, tidak bergantung terhadap waktu atau variansi fluktuasi tersebut. Stasioneritas dalam rata-rata dapat dianalisis menggunakan uji ADF.

2.4 Transformasi Box-Cox

Menurut Wei (2006), pada tahun 1964 Box dan Cox memperkenalkan metode Transformasi Box-Cox yang digunakan untuk mengubah ragam data ketika ragam data tidak stasioner. Transformasi Box-Cox didefinisikan sebagai berikut.

$$Z_t = \begin{cases} \left(\frac{Z_t^\lambda - 1}{\lambda}\right), \lambda \neq 0 \\ \ln Z_t, \lambda = 0 \end{cases} \quad (2.2)$$

dengan

Z_t = data pada waktu ke- t

λ = indikator Transformasi Box-Cox

Rumus transformasi yang digunakan berbeda-beda dengan mempertimbangkan nilai λ . Box Cox memperhitungkan kelas transformasi dengan parameter tunggal, yaitu λ yang dipangkatkan pada variabel respon Z , kemudian transformasinya menjadi Z^λ dan λ merupakan parameter yang perlu diprediksi (Juniawati & Hajarisman, 2023). Aturan Transformasi Box-Cox tertera pada Tabel 1.

Tabel 1. Transformasi Box-Cox

Nilai λ (lamda)	Transformasi
1	Z_t (tidak transformasi)
0,5	$\sqrt{Z_t}$
0	$\ln Z_t$
-0,5	$1/\sqrt{Z_t}$
-1	$1/Z_t$

2.5 Uji *Augmented Dickey-Fuller* (ADF)

Uji ADF digunakan untuk mengetahui suatu data telah stasioner atau tidak (Gujarati & Porter, 2009). Persaman Uji ADF adalah sebagai berikut.

$$\Delta Y_t = \phi Y_{t-1} \sum_{j=1}^{\rho-1} a_j^* \Delta Y_{t-1} + \varepsilon_t \quad (2.3)$$

dengan

ΔY_t = hasil *difference* data pada periode ke-t

ϕ = koefisien *autoregressive*

a_j^* = konstanta

ρ = panjang *lag*

ε_t = residual waktu ke-t

Berikut ini adalah rumus statistik uji *t*.

$$t = \frac{\bar{X} - \theta}{s / \sqrt{n}} \quad (2.4)$$

dengan

\bar{X} = rata-rata sampel

θ = nilai parameter yang diuji (rata-rata populasi)

s = simpangan baku sampel

n = ukuran sampel

Setelah melihat statistik uji *t* kemudian didapatkan nilai ϕ yang selanjutnya dapat digunakan dalam uji statistik pada ADF. Pada model ini, hipotesis yang diuji adalah:

$H_0 : \phi = 0$ (data *time series* tidak stasioner)

$H_1 : \phi < 0$ (data *time series* stasioner)

Dengan kriteria tolak H_0 jika *p-value* < α maka dapat dikatakan data stasioner. Sebaliknya, terima H_0 jika *p-value* > α dapat dikatakan data tidak stasioner.

2.6 Pembeda (*Differencing*)

Menurut Pankrtaz (1991), penyesuaian bisa dilakukan dengan menambahkan data baru yang mempunyai nilai rata-rata yang konstan jika nilai rata-rata kumpulan data tidak konstan. Untuk mencapai tujuan ini, salah satu metode yang paling umum digunakan adalah proses *differencing* data. Langkah awal yang digunakan dalam proses ini yaitu perumusan *differencing* pertama atau $d = 1$.

Hal ini dicapai dengan menghitung berapa banyak perubahan atau peningkatan yang terjadi dalam data deret waktu dari titik waktu ke titik berikutnya. Proses ini juga memungkinkan untuk menghilangkan atau mengurangi pola atau *tren* yang sudah ada dalam data, sehingga membuatnya lebih stasioner dan memungkinkan untuk melakukan analisis lanjutan seperti pemodelan ARIMA atau identifikasi pola lainnya. Berikut ini adalah rumus *differencing* pertama atau $d = 1$.

$$X'_t = X_t - X_{t-1} \quad (2.5)$$

dengan

X'_t = nilai *time series* dari hasil *differencing* pertama pada waktu t

X_t = nilai *time series* pada waktu t

X_{t-1} = nilai *time series* pada waktu sebelumnya, yaitu $t - 1$

Untuk *differencing* kedua atau $d = 2$, didapatkan dengan cara sebagai berikut.

$$\begin{aligned} X''_t &= X'_t - X'_{t-1} \\ &= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \end{aligned} \quad (2.6)$$

2.7 Fungsi Autokorelasi dan Fungsi Autokorelasi Parsial

Dalam menentukan model *time series* dapat diidentifikasi dengan fungsi autokorelasi / *Autocorrelation Function* (ACF) dan fungsi autokorelasi parsial / *Partial Autocorrelation Function* (PACF).

2.7.1 Fungsi Autokorelasi (ACF)

Fungsi autokorelasi atau *Autocorrelation Function* (ACF) adalah alat statistik yang dapat digunakan untuk mengukur ketergantungan nilai pada rangkaian data *time series* dalam periode waktu yang berbeda dan digunakan dalam perhitungan koefisien korelasi. Proses stasioneritas suatu data *time series* (X_t) didapatkan $E(X_t) = \mu$ dan variansi $Var(X_t) = E(X_t - \mu)^2 = \sigma_t^2$ yang konstan serta kovarian $Cov(X_t, X_{t+k})$ berfungsi ketika perbedaan waktu $|t - (t + k)|$. Selanjutnya, hasil dapat ditulis sebagai kovariansi X_t dan X_{t+k} .

$$\begin{aligned}\gamma_k &= Cov(X_{t+1}, X_{t+2}) = E[(X_t - \mu)(X_{t+k} - \mu)] \\ \gamma_0 &= \sqrt{Var(X_t)Var(X_{t+k})} = \sqrt{E(X_t - \mu)^2} \sqrt{E(X_{t+k} - \mu)^2}\end{aligned}$$

fungsi autokorelasi antara X_t dan X_{t+k} yaitu:

$$\begin{aligned}\rho_k &= \frac{Cov(X_{t+1}, X_{t+2})}{\sqrt{Var(X_t)Var(X_{t+k})}} \\ &= \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\sqrt{E(X_t - \mu)^2} \sqrt{E(X_{t+k} - \mu)^2}} = \frac{\gamma_k}{\gamma_0}\end{aligned}\tag{2.7}$$

dengan $Var(X_t) = Var(X_{t+k}) = \gamma_0$ merupakan fungsi dari k , γ_k adalah fungsi autokovarian dan ρ_k sebagai fungsi autokorelasi ACF, γ_k dan ρ_k menggambarkan kovarian dan hubungan antara X_t dan X_{t+k} hanya terpisah dari proses yang sama yaitu oleh *lag* ke- k (Wei, 2006).

Sifat-sifat pada fungsi autokorelasi γ_k dan fungsi autokorelasi ρ_k adalah sebagai berikut.

1. $\gamma_0 = Var(X_t)$; $\rho_0 = 1$

Bukti :

Akan dibuktikan menggunakan definisi korelasi antara X_t dan X_{t+k} bahwa

$$\gamma_0 = Var(X_t); \rho_0 = 1$$

$$\rho_k = \frac{Cov(X_{t+1}, X_{t+2})}{\sqrt{Var(X_t)Var(X_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

Diberikan $k = 0$, maka

$$\rho_0 = \frac{Cov(X_{t+1}, X_{t+0})}{\sqrt{Var(X_t)Var(X_{t+0})}} = \frac{\gamma_0}{\gamma_0}$$

$$\rho_0 = \frac{Cov(X_t, X_t)}{\sqrt{Var(X_t)Var(X_t)}} = \frac{\gamma_0}{\gamma_0}$$

$$\rho_0 = \frac{Var(X_t)}{Var(X_t)} = \frac{\gamma_0}{\gamma_0}$$

$$\rho_0 = \frac{Var(X_t)}{Var(X_t)} = \frac{\gamma_0}{\gamma_0} \text{ (Terbukti).}$$

$$2. |\gamma_k| \leq \gamma_0 ; |\rho_k| \leq 1$$

Bukti :

Sifat tersebut dihasilkan dari persamaan autokorelasi di mana nilai mutlaknya sama dengan atau kurang dari 1.

$$3. \gamma_k = \gamma_{(-k)} \text{ dan } \rho_k = \rho_{(-k)} \text{ untuk semua } k. \gamma_k \text{ dan } \rho_k \text{ adalah fungsi yang sama dan simetri lag } k = 0.$$

Bukti :

Sifat ketiga didapat dari perbedaan waktu antara X_t dan X_{t+k} . Jadi, fungsi autokorelasi biasanya diplotkan untuk lag positif dan plot fungsi autokorelasi biasanya disebut korelogram (Wei, 2006).

2.7.2 Fungsi Parsial Autokorelasi (PACF)

Fungsi autokorelasi parsial atau *Partial Autocorrelation Function* (PACF) adalah alat statistik yang digunakan untuk mengidentifikasi contoh korelasi parsial antara pengamatan pada waktu ke t (Z_t) dan pengamatan terhadap titik waktu sebelumnya ($Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$). Autokorelasi parsial dilakukan untuk mengukur tingkat keeratan hubungan antara X_t dan X_{t+k} setelah menghilangkan pengaruh ketergantungan linier terhadap variabel $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$. Fungsi autokorelasi parsial dinotasikan dengan :

$$corr(X_t, X_{t+1}, X_{t+2}, \dots, X_{t+k})$$

Misalkan X_t adalah proses stasioner dengan $E(X_t) = 0$, sehingga X_{t+k} dinyatakan sebagai berikut.

$$X_{t+k} = \phi_{k1}X_{t+k-1} + \phi_{k2}X_{t+k-2} + \dots + \phi_{kk}X_t + \varepsilon_{t+k} \quad (2.8)$$

dimana ϕ_{ki} merupakan parameter regresi ke- i dan ε_{t+k} adalah nilai kesalahan yang tidak berkorelasi dengan X_{t+k-j} dengan $j = 1, 2, \dots, k$. Untuk mendapatkan nilai PACF, proses pertama yang dilakukan yaitu mengalikan persamaan (2.8) dengan X_{t+k-j} pada kedua ruas data sehingga diperoleh :

$$X_{t+k-j}X_{t+k} = \phi_{k1}X_{t+k-1}X_{t+k-j} + \phi_{k2}X_{t+k-2}X_{t+k-j} + \dots + \phi_{kk}X_tX_{t+k-j}$$

Kemudian, nilai ekspektasinya adalah :

$$E(X_{t+k-j}X_{t+k}) = E(\phi_{k1}X_{t+k-1}X_{t+k-j} + \phi_{k2}X_{t+k-2}X_{t+k-j} + \dots + \phi_{kk}X_tX_{t+k-j} + \varepsilon_{t+k}X_{t+k-j})$$

dimisalkan nilai

$$E(X_{t+k-j}X_{t+k}) = \gamma_j, j = 0, 1, \dots, k$$

karena $E(X_{t+k-j}X_{t+k}) = 0$. Sehingga didapatkan

$$\gamma_j = \phi_{k1}\gamma_{j-1} + \phi_{k2}\gamma_{j-2} + \dots + \phi_{kk}\gamma_{j-k} \quad (2.9)$$

Persamaan (2.9) dibagi dengan γ_0

$$\frac{\gamma_j}{\gamma_0} = \phi_{k1}\frac{\gamma_{j-1}}{\gamma_0} + \phi_{k2}\frac{\gamma_{j-2}}{\gamma_0} + \dots + \phi_{kk}\frac{\gamma_{j-k}}{\gamma_0}$$

Diperoleh nilai untuk autokorelasi parsial sebagai berikut:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-k}, \quad j = 1, 2, \dots, k$$

Untuk $j = 1, 2, \dots, k$ didapatkan nilai persamaan

$$\rho_1 = \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \dots + \phi_{kk}\rho_{j-1}$$

$$\rho_2 = \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \dots + \phi_{kk}\rho_{j-2}$$

⋮

$$\rho_k = \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \dots + \phi_{kk}\rho_0 \quad (2.10)$$

Sistem persamaan (2.10) untuk $j = 1, 2, \dots, k$ digunakan untuk mencari nilai-nilai fungsi autokorelasi parsial lag k yaitu $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$.

a. Untuk lag pertama ($k = 1$) dan ($j = 1$) didapatkan nilai yaitu:

$\rho_1 = \phi_{11}\rho_0$ karena $\rho_0 = 1$ sehingga $\rho_1 = \phi_{11}$ yang berarti bahwa pada lag pertama fungsi autokorelasi parsial akan sama dengan lag pertama fungsi autokorelasi.

- b. Untuk *lag* ke-2 ($k = 2$) dan ($j = 2$) diperoleh persamaan

$$\begin{aligned}\rho_1 &= \phi_{11}\rho_0 + \phi_{22}\rho_1 \\ \rho_2 &= \phi_{11}\rho_1 + \phi_{22}\rho_0\end{aligned}\quad (2.11)$$

Jika ditulis dalam bentuk matriks, maka persamaan (2.11) akan menjadi

$$\begin{bmatrix} \rho_0 & \rho_1 \\ \rho_1 & \rho_0 \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{22} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}$$

Matriks \mathbf{A} adalah matriks kovarians untuk proses dengan satu parameter autokorelasi parsial (PACF), yaitu ρ_1 .

$$\mathbf{A} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}$$

Sedangkan, \mathbf{A}_2 menggambarkan kovarians yang melibatkan dua parameter PACF, ρ_1 dan ρ_2 .

$$\mathbf{A}_2 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}$$

Dalam model *autoregressive* (AR), matriks \mathbf{A} atau \mathbf{A}_2 biasanya digunakan untuk menghitung koefisien PACF pada *lag* tertentu. Matriks-matriks ini dapat digunakan untuk menghitung hubungan parsial antar-*lag*. Dengan menggunakan aturan Cramer didapat

$$\phi_{22} = \frac{\det(\mathbf{A}_2)}{\det(\mathbf{A})} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

- c. Untuk *lag* ke-3 ($k = 3$) dan ($j = 3$) didapat nilai persamaan

$$\begin{aligned}\rho_1 &= \phi_{11}\rho_0 + \phi_{22}\rho_1 + \phi_{33}\rho_2 \\ \rho_2 &= \phi_{11}\rho_1 + \phi_{22}\rho_0 + \phi_{33}\rho_1 \\ \rho_3 &= \phi_{11}\rho_2 + \phi_{22}\rho_1 + \phi_{33}\rho_0\end{aligned}\quad (2.12)$$

Jika ditulis dalam bentuk matriks, persamaan (2.12) akan menjadi

$$\begin{bmatrix} \rho_0 & \rho_1 & \rho_2 \\ \rho_1 & \rho_0 & \rho_1 \\ \rho_2 & \rho_1 & \rho_0 \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{22} \\ \phi_{33} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}$$

Dalam model *autoregressive*, matriks \mathbf{A} menunjukkan hubungan antar-*lag* dengan elemen diagonal utama bernilai 1 dan elemen lainnya

merepresentasikan nilai autokorelasi antar-*lag*. Selain itu, matriks A_3 merupakan variasi dari matriks A dengan semua elemen berada di luar diagonal utama yaitu bernilai ρ_1 dan merepresentasikan kasus khusus bahwa hubungan antar-*lag* dianggap seragam dan hanya bergantung pada ρ_1 .

d. Untuk *lag* ke- $j = 1, 2, \dots, k$ persamaannya yaitu

$$\begin{aligned}\rho_1 &= \phi_{11}\rho_0 + \phi_{22}\rho_1 + \phi_{33}\rho_2 + \dots + \phi_{kk}\rho_{k-1} \\ \rho_2 &= \phi_{11}\rho_1 + \phi_{22}\rho_0 + \phi_{33}\rho_1 + \dots + \phi_{kk}\rho_{k-2} \\ \rho_3 &= \phi_{11}\rho_2 + \phi_{22}\rho_1 + \phi_{33}\rho_0 + \dots + \phi_{kk}\rho_{k-3} \\ &\vdots \\ \rho_k &= \phi_{11}\rho_1 + \phi_{22}\rho_2 + \phi_{33}\rho_3 + \dots + \phi_{kk}\rho_0\end{aligned}\tag{2.13}$$

Dalam bentuk matriks, persamaan (2.13) dinyatakan sebagai

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_k \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{22} \\ \phi_{33} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{bmatrix}$$

dengan aturan Cramer didapat

$$A_k = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_k \end{bmatrix}$$

Hasil dari nilai autokorelasi *lag* k yaitu

$$\phi_{kk} = \frac{\det(A_k)}{\det(A)} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{vmatrix}}$$

ϕ_{kk} disebut PACF antara X_t dan X_{t+k} atau bisa dituliskan dengan

$$\phi_{kk} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

Sehingga didapatkan autokorelasi parsial dari X_t lag k . Himpunan dari $\phi_{kk}\{\phi_{kk}; k = 1, 2, \dots\}$ disebut *Partial Autocorrelation Function* (PACF). Fungsi ϕ_{kk} menjadi notasi standar dalam analisis *time series* untuk autokorelasi parsial antara observasi X_t dan X_{t+k} . Fungsi ϕ_{kk} akan bernilai nol untuk $k > \rho$. Sifat tersebut digunakan untuk mengidentifikasi model AR, yakni nilai ACF yang berlaku untuk model *autoregressive* akan secara perlahan menurun menuju nol dari nilai PACF model AR yaitu $\phi_{kk} = 0, k > \rho$ (Wei, 2006).

2.8 Uji Nonlinieritas

Uji nonlinieritas pada awalnya digunakan pada data deret waktu saat menggunakan model nonlinier untuk memastikan bahwa metode yang dipilih sesuai dengan kumpulan data yang diberikan. Untuk mengetahui hubungan nonlinier setiap variabel digunakan uji Terasvirta (Teräsvirta, 1994). Hipotesis yang digunakan dalam uji Teräsvirta adalah sebagai berikut:

H_0 : $f(x)$ merupakan model linier

H_1 : $f(x)$ merupakan model nonlinier

Statistik uji:

$$F = \frac{(SSR_0 - SSR)/m}{SSR/(N - p - 1 - m)} \quad (2.14)$$

$$SSR_0 = \sum_{i=1}^n (\hat{Y}_0 - \bar{Y})^2, \text{ pada kondisi } H_0 \text{ benar} \quad (2.15)$$

$$SSR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2, \text{ pada kondisi } H_0 \text{ benar} \quad (2.16)$$

dengan

SSR_0 = jumlah kuadrat regresi di bawah hipotesis nol

SSR = jumlah kuadrat regresi

p = jumlah variabel prediktor awal

m = jumlah variabel prediktor kuadratik dan kubik

N = jumlah data

Kriteria keputusan yakni $F_{hitung} < F_{(m, (N-p-1-m))}$ atau $p - value < \alpha$ sehingga tolak H_0 yang berarti bahwa $f(x)$ merupakan fungsi nonlinier x .

2.9 State

Menurut Cox and Miller (1965), *state* merupakan kondisi peubah acak x_t dimana jika suatu peubah acak berada pada *state* tersebut maka dapat berpindah ke *state* lainnya. Dalam $AR(p)$ stasioner, *state* berubah secara bertahap dan kontinu menurut dinamika regresi. Sedangkan, pada MSAR model bisa berpindah antara beberapa *state* yang berbeda dengan probabilitas *Markov* karena *state* memiliki *regime switching* yang terjadi ketika karakteristik model berubah pada secara berkala, biasanya dikendalikan oleh proses *Markov*.

Untuk menentukan *state* adalah apresiasi atau depresiasi dengan mempertimbangkan nilai μ_{st} dengan ketentuan $\mu_2 < \mu_1$ atau rata-rata pada *state* 1 lebih kecil dari rata-rata *state* 2 (Hamilton, 1994).

2.10 Markov Switching

Model *Markov Switching* adalah model yang sering digunakan untuk deret waktu nonlinier. Model tersebut dapat memodelkan data *time series* yang mengalami perubahan struktur dan cocok untuk menjelaskan data berkorelasi yang menunjukkan pola dinamik yang jelas selama perubahan periode waktu. Perubahan (*switching*) bisa terjadi pada rata-rata dan varians. Model dengan perubahan dalam rata-rata dan varians yaitu:

$$y_t = \mu_{st} + \varepsilon_t \quad (2.17)$$

dimana y_t merupakan variabel dependen pada waktu t , μ_{st} merupakan rata-rata pada waktu t dari data yang diamati yang dipengaruhi oleh perubahan keadaan, ε_t

adalah *error* pada waktu t dengan $\varepsilon_t \sim \text{i. i. d } N(0, \sigma_{s_t}^2)$, dan $\sigma_{s_t}^2$ adalah varians dari ε_t yang dipengaruhi oleh perubahan keadaan (Hamilton, 1994).

2.11 Model *Markov Switching Autoregressive* (MSAR)

Penggabungan sederhana model *Markov Switching* dan model *Autoregressive* disebut dengan model MSAR. Menurut Kim dan Nelson (1999), model MSAR dituliskan sebagai berikut.

$$(y_t - \mu_{st}) = \sum_{i=1}^p \phi_i (y_{t-1} - \mu_{s_{t-i}}) + \varepsilon_t$$

atau dapat ditulis:

$$(y_t - \mu_{st}) = \phi_1 (y_{t-1} - \mu_{s_{t-1}}) + \dots + \phi_p (y_{t-p} - \mu_{s_{t-p}}) + \varepsilon_t \quad (2.18)$$

dengan $\varepsilon_t \sim \text{iid } N(0, \sigma_{s_t}^2)$.

dimana

$y_t, y_{t-1}, \dots, y_{t-p}$ = data pengamatan

$\phi_1, \phi_2, \dots, \phi_p$ = koefisien *autoregressive*

$\mu_{st}, \mu_{s_{t-1}}, \dots, \mu_{s_{t-p}}$ = rata-rata pada saat t yang dipengaruhi perubahan *state*

$\sigma_{s_t}^2$ = varians pada saat t yang dipengaruhi perubahan *state*

ε_t = residual pada t

2.11.1 Peluang Perpindahan *State*

Dalam model *Markov Switching Autoregressive*, *state* tidak teramati. Proses *filtering* dan *smoothing* dilakukan dalam menghitung nilai peluang pada model ini. Matriks transisi membentuk peluang perpindahan *state* karena nilai masa lalu dipengaruhi nilai sekarang oleh rantai *Markov* pada matriks transisi, dengan jumlah *entri* pada baris matriks transisi bernilai 1 (Kim and Nelson, 1999).

2.11.2 Rata-rata Jangka Waktu pada Masing-masing *State*

Menurut Kim and Nelson (1999), lama durasi rata-rata dari setiap *state* juga dapat dihitung dengan model MSAR. Matriks peluang transisi memiliki elemen diagonal yang berisi informasi penting tentang durasi rata-rata yang diharapkan dari suatu *state* akan bertahan. Persamaan perhitungan durasi rata-rata *state* adalah sebagai berikut:

$$E(D) = \frac{1}{1-p_{jj}} \quad (2.19)$$

dengan

$E(D)$ = rata-rata durasi waktu

p_{jj} = probabilitas transisi dari *state j* kembali ke *state j*

Jika p_{jj} mendekati 1 maka *state j* akan bertahan lebih lama dan ekspektasi durasinya akan tinggi. Sedangkan, jika p_{jj} rendah maka *state j* lebih cepat berubah ke *state* lain.

2.12 Peluang (*Probability*)

Menentukan frekuensi relatif adalah salah satu cara untuk mengetahui peluang terjadinya suatu peristiwa. Definisi tersebut bermaksud bahwa dalam menduga percobaan dengan ruang sampel berukuran S yang dilakukan berulang kali dengan kondisi yang sama. Pada kejadian E yang terjadi pada ruang sampel S , $n(E)$ dikalikan sebanyak n pertama pengulangan percobaan bahwa kejadian E terjadi. Berikut ini didefinisikan peluang dari kejadian E .

$$P(E) = \frac{n(E)}{n(S)} \quad (2.20)$$

Berarti bahwa $P(E)$ merupakan batas waktu E yang terjadi. Dengan demikian, E dibatasi oleh frekuensi. Pendekatan terakhir adalah pendekatan aksiomatis kontemporer terhadap teori. Diasumsikan bahwa probabilitas E merupakan nilai

$P(E)$ untuk setiap kejadian E di ruang sampel S . Selanjutnya, akan diasumsikan bahwa semua probabilitas ini memenuhi beberapa aksioma tertentu.

Pertimbangkan percobaan yang menggunakan ruang sampel S .

Diasumsikan bahwa bilangan $P(E)$ pada setiap kejadian E di ruang sampel S didefinisikan dan memenuhi tiga aksioma berikut:

Aksioma 1:

$$0 \leq P(E) \leq 1$$

Aksioma 2:

$$P(S) = 1$$

Aksioma 3:

Untuk barisan dari kejadian saling lepas E_1, E_2, \dots (Kejadian yang $E_i E_j = \emptyset$ ketika $i \neq j$).

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

dengan $P(E)$ adalah peluang kejadian E . Aksioma ini menunjukkan aturan dasar untuk kejadian saling lepas. Untuk peluang bersyarat, kejadian tidak harus saling lepas dan peluang dihitung berdasarkan hubungan antara dua kejadian. (Ross, 2010).

2.13 Matriks (*Matrix*)

Barisan bilangan berbentuk persegi panjang disebut matriks. Bilangan pada barisan tersebut dikenal dengan *entri* matriks. Matriks memiliki jumlah baris dan kolom yang menunjukkan ukurannya. Huruf yang menunjukkan matriks dicetak dengan huruf kapital. Dalam matriks \mathbf{A} , *entri* yang terletak dalam baris i dan kolom j dinyatakan dengan a_{ij} atau dapat ditulis dengan $(\mathbf{A})_{ij}$. Secara umum, matriks $m \times n$ dapat ditulis sebagai berikut.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Dalam notasi singkat, pada matriks di atas jika ukuran matriks sangat penting diketahui dapat ditulis sebagai $[a_{ij}]_{m \times n}$ atau jika ukuran matriks tidak terlalu penting diketahui ditulis dengan $[a_{ij}]$ (Anton dan Rorres, 2004).

2.14 Matriks Transisi (*Transition Matriks*)

Evolusi rantai Markov $(Z_n)_{n \in \mathbb{N}}$ secara acak ditentukan oleh data yang menunjukkan bahwa nilai sebelumnya memengaruhi nilai saat ini

$$P_{i,j} = P(Z_1 = j | Z_0 = i), \quad i, j \in S$$

yang bertepatan dengan peluang $P(Z_{n+1} = j | Z_n = i)$ yang independen $n \in \mathbb{N}$.

Data tersebut dapat membuat matriks yang dikenal sebagai matriks transisi yaitu

$$[P_{i,j}]_{i,j \in S} = P(Z_1 = j | Z_0 = i)$$

untuk situasi yang memiliki 2 kondisi (*state*) dapat ditulis sebagai

$$[P_{i,j}]_{i,j \in S} = \begin{bmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{bmatrix}$$

Dilihat bahwa urutan inversi indeks (i, j) dari $P(Z_{n+1} = j | Z_n = i)$ dan $P_{i,j}$.

Secara khusus, bilangan baris dalam baris merupakan keadaan awal i dan bilangan kolom merupakan keadaan akhir j . Sehingga hubungannya adalah:

$$\sum_{j \in S} P(Z_1 = j | Z_0 = i) = 1, i \in N$$

Kondisi yang dimiliki baris pada matriks transisi

$$\sum_{j \in S} P_{i,j} = 1$$

untuk setiap $i \in S$ (Prevault, 2013).

2.15 Estimasi Parameter Model *Markov Switching Autoregressive* (MSAR)

Metode yang digunakan dalam estimasi parameter adalah metode *Maximum Likelihood Estimation* (MLE). Langkah yang dapat dilakukan yaitu mengidentifikasi fungsi densitasnya yang kemudian dapat diubah menjadi fungsi *log likelihood*. Menurut Hamilton (1989), model MSAR memiliki fungsi densitas:

$$f(y_t | s_t, s_{t-1}, \dots, s_{t-p}, \Omega_{t-1}; \theta) = \frac{1}{\sigma_{s_t} \sqrt{2\pi}} \exp \left[-\frac{((y_t - \mu_{s_t}) - \phi_1(y_{t-1} - \mu_{s_{t-1}}) - \phi_2(y_{t-2} - \mu_{s_{t-2}}) - \dots - \phi_p(y_{t-p} - \mu_{s_{t-p}}))^2}{2\sigma_{s_t}^2} \right] \quad (2.21)$$

dengan

$\Omega_{t-1} = (y_{t-1}, y_{t-2}, \dots, y_{t-p})$: data pengamatan masa lalu

$\theta = (\mu_{s_t}, \sigma_{s_t}^2, \phi_p, p_{ij})$: parameter model MSAR

Mengidentifikasi fungsi densitas dari y_t yang diberikan informasi melalui Ω_{t-1} dan membutuhkan nilai $s_t, s_{t-1}, \dots, s_{t-p}$ yang merupakan variabel tak teramati (*state*). Dalam menyelesaikan masalah ini, yang perlu dilakukan adalah mempertimbangkan fungsi densitas bersama dari y_t dan $s_t, s_{t-1}, \dots, s_{t-p}$, yaitu:

1. Memperoleh fungsi densitas bersama dari y_t dan $s_t, s_{t-1}, \dots, s_{t-p}$ bersyarat informasi masalah.

$$f(y_t, s_t, s_{t-1}, \dots, s_{t-p}, \Omega_{t-1}; \theta) = f(y_t | s_t, s_{t-1}, \dots, s_{t-p}, \Omega_{t-1}; \theta) P(s_t, s_{t-1}, \dots, s_{t-p}, \Omega_{t-1}; \theta)$$

2. Untuk mendapatkan $f(y_t | \Omega_{t-1}; \theta)$, maka menggabungkan densitas bersama dari $s_t, s_{t-1}, \dots, s_{t-p}$ dengan menjumlahkan semua kemungkinan densitas bersama dari $s_t, s_{t-1}, \dots, s_{t-p}$:

$$\begin{aligned} f(y_t | \Omega_{t-1}; \theta) &= \sum_{s_t=0}^1 \sum_{s_{t-1}=0}^1 \dots \sum_{s_{t-p}=0}^1 f(y_t, s_t, s_{t-1}, \dots, s_{t-p} | \Omega_{t-1}; \theta) \\ &= \sum_{s_t=0}^1 \sum_{s_{t-1}=0}^1 \dots \sum_{s_{t-p}=0}^1 f(y_t | s_t, s_{t-1}, \dots, s_{t-p}, \Omega_{t-1}; \theta) P[s_t, s_{t-1}, \dots, s_{t-p} | \Omega_{t-1}] \end{aligned}$$

Nilai peluang $P[s_t, s_{t-1}, \dots, s_{t-p} | \Omega_{t-1}]$ dihitung melalui proses *filtering* dan *smoothing*.

2.15.1 Filtering

Menurut Wahyudi, *et all.* (2015), Proses *filtering* digunakan untuk memperoleh peluang *state* pada saat t yang didasarkan pada pengamatan hingga saat t . Proses *filtering* menghasilkan nilai *filtered state probability* dengan langkah sebagai berikut :

Dimulai

$$P[S_0 = s_0, S_{-1} = s_{-1}, \dots, S_{-p+1} = s_{-p+1} | \Omega_{t-1}]$$

dengan

$$P[S_0 = 1 | \Omega_0] = \frac{1 - p_{22}}{2 - p_{22} - p_{11}}$$

$$P[S_0 = 2 | \Omega_0] = \frac{1 - p_{11}}{2 - p_{22} - p_{11}}$$

Selanjutnya, menghitung :

$$\begin{aligned} P[S_t = s_t = j, S_{t-1} = s_{t-1}, \dots, S_{t-p+1} = s_{t-p+1} = i | \Omega_t] \\ &= \frac{f(S_t = s_t = j, S_{t-1} = s_{t-1}, \dots, S_{t-p+1} = s_{t-p+1} = i | \Omega_{t-1})}{f(y_t | \Omega_{t-1})} \\ &= \frac{f(y_t | S_t = s_t, S_{t-1} = s_{t-1}, \dots, S_{t-p} = s_{t-p}, \Omega_{t-1}) \times P[y_t | S_t = s_t, S_{t-1} = s_{t-1}, \dots, S_{t-p} = s_{t-p} | \Omega_{t-1}]}{\sum_{s_t=0}^1 \sum_{s_{t-1}=0}^1 \dots \sum_{s_{t-p}=0}^1 f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \dots, S_{t-p} = s_{t-p} | \Omega_{t-1})} \end{aligned}$$

Sehingga diperoleh hasil dari proses *filtering* :

$$\begin{aligned} P[S_t = s_t = j, S_{t-1} = s_{t-1}, \dots, S_{t-p+1} = s_{t-p+1} = i | \Omega_t] \\ &= \sum_{s_{t-p}=0}^1 P[S_t = s_t, S_{t-1} = s_{t-1}, \dots, S_{t-p} = s_{t-p} | \Omega_t] \end{aligned} \quad (2.22)$$

Tahap awal yang digunakan untuk menentukan nilai *filtered state probability* yaitu dengan cara menetapkan fungsi densitas y_t bagi setiap kemungkinan *state* saat $t = 1$.

$$\begin{aligned} f(y_t | s_t^* = 1, s_{t-1}^* = 1, \Omega_0; \theta) \\ &= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left(- \frac{[(y_t - \mu_{s_t=1}^*) - \phi(y_{t-1} - \mu_{s_{t-1}=1}^*)]^2}{2\sigma_1^2} \right) \end{aligned} \quad (2.23)$$

$$\begin{aligned}
f(y_t | s_t^* = 1, s_{t-1}^* = 2, \Omega_0; \theta) &= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_{s_t^*=1}) - \phi(y_{t-1} - \mu_{s_{t-1}^*=2})]^2}{2\sigma_1^2}\right) \\
&= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_1) - \phi(y_{t-1} - \mu_2)]^2}{2\sigma_1^2}\right)
\end{aligned} \tag{2.24}$$

$$\begin{aligned}
f(y_t | s_t^* = 2, s_{t-1}^* = 1, \Omega_0; \theta) &= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_{s_t^*=2}) - \phi(y_{t-1} - \mu_{s_{t-1}^*=1})]^2}{2\sigma_2^2}\right) \\
&= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_2^2}\right)
\end{aligned} \tag{2.25}$$

$$\begin{aligned}
f(y_t | s_t^* = 2, s_{t-1}^* = 2, \Omega_0; \theta) &= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_{s_t^*=2}) - \phi(y_{t-1} - \mu_{s_{t-1}^*=2})]^2}{2\sigma_2^2}\right) \\
&= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_2)]^2}{2\sigma_2^2}\right)
\end{aligned} \tag{2.26}$$

Sehingga, nilai peluang suatu *state* bisa dihitung dengan cara berikut.

$$\begin{aligned}
P(s_1^* = 1, s_0^* = 1 | \Omega_0; \theta) &= P_{11} \times P(s_0^* = 1 | \Omega_0; \theta) = p_{11} \times \mu_1, \\
P(s_1^* = 1, s_0^* = 2 | \Omega_0; \theta) &= P_{21} \times P(s_0^* = 2 | \Omega_0; \theta) = p_{21} \times \mu_2, \\
P(s_1^* = 2, s_0^* = 1 | \Omega_0; \theta) &= P_{12} \times P(s_0^* = 1 | \Omega_0; \theta) = p_{12} \times \mu_1, \\
P(s_1^* = 2, s_0^* = 2 | \Omega_0; \theta) &= P_{22} \times P(s_0^* = 2 | \Omega_0; \theta) = p_{22} \times \mu_2,
\end{aligned}$$

Jumlah perkalian fungsi densitas y_t bersyarat s_t^* , s_{t-1}^* dan Ω_{t-1} probabilitas *state* untuk semua kemungkinan s_t^* dan s_{t-1}^* digunakan untuk menghitung fungsi densitas y_t bersyarat Ω_{t-1} . Probabilitas y_1 diperoleh untuk semua kemungkinan s_t^* dan s_{t-1}^* sebagai berikut.

$$\begin{aligned}
P(s_1^* = 1, s_0^* = 1 | \Omega_0, y_1; \theta) &= \frac{f(y_1 | s_1^* = 1, s_0^* = 1 | \Omega_0; \theta) \times P(s_1^* = 1, s_0^* = 1 | \Omega_0; \theta)}{\sum_{j=1}^2 \sum_{i=1}^2 f(y_1 | s_1^* = j, s_0^* = i | \Omega_0; \theta) \times P(s_1^* = j, s_0^* = i | \Omega_0; \theta)} \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_1 - \mu_1) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_1^2}\right) \times (p_{11} \times \mu_1)}{f(y_1 | \Omega_0; \theta)}
\end{aligned} \tag{2.27}$$

$$\begin{aligned}
& P(s_1^* = 1, s_0^* = 2 | \Omega_0, y_1; \theta) \\
&= \frac{f(y_1 | s_1^* = 1, s_0^* = 2 | \Omega_0; \theta) \times P(s_1^* = 1, s_0^* = 2 | \Omega_0; \theta)}{\sum_{j=1}^2 \sum_{i=1}^2 f(y_1 | s_1^* = j, s_0^* = i | \Omega_0; \theta) \times P(s_1^* = j, s_0^* = i | \Omega_0; \theta)} \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_1) - \phi(y_{t-1} - \mu_2)]^2}{2\sigma_1^2}\right) \times (p_{21} \times \pi_2)}{f(y_1 | \Omega_0; \theta)} \tag{2.28}
\end{aligned}$$

$$\begin{aligned}
& P(s_1^* = 2, s_0^* = 1 | \Omega_0, y_1; \theta) \\
&= \frac{f(y_1 | s_1^* = 2, s_0^* = 1 | \Omega_0; \theta) \times P(s_1^* = 2, s_0^* = 1 | \Omega_0; \theta)}{\sum_{j=1}^2 \sum_{i=1}^2 f(y_1 | s_1^* = j, s_0^* = i | \Omega_0; \theta) \times P(s_1^* = j, s_0^* = i | \Omega_0; \theta)} \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_1^2}\right) \times (p_{12} \times \pi_1)}{f(y_1 | \Omega_0; \theta)} \tag{2.29}
\end{aligned}$$

$$\begin{aligned}
& P(s_1^* = 2, s_0^* = 2 | \Omega_0, y_1; \theta) \\
&= \frac{f(y_1 | s_1^* = 2, s_0^* = 2 | \Omega_0; \theta) \times P(s_1^* = 2, s_0^* = 2 | \Omega_0; \theta)}{\sum_{j=1}^2 \sum_{i=1}^2 f(y_1 | s_1^* = j, s_0^* = i | \Omega_0; \theta) \times P(s_1^* = j, s_0^* = i | \Omega_0; \theta)} \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_2)]^2}{2\sigma_1^2}\right) \times (p_{22} \times \pi_2)}{f(y_1 | \Omega_0; \theta)} \tag{2.30}
\end{aligned}$$

Dengan demikian, *filtered state probability* diperoleh sebagai berikut.

Untuk $s_t^* = 1$

$$\begin{aligned}
P(s_1^* = 1 | \Omega_1; \theta) &= P(s_1^* = 1, s_0^* = 1 | \Omega_1; \theta) + P(s_1^* = 1, s_0^* = 2 | \Omega_0; \theta) \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_1) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_1^2}\right) \times (p_{11} \times \pi_1)}{f(y_1 | \Omega_0; \theta)} \\
&\quad + \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_1) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_1^2}\right) \times (p_{11} \times \pi_1)}{f(y_1 | \Omega_0; \theta)} \tag{2.31}
\end{aligned}$$

Untuk $s_t^* = 2$

$$\begin{aligned}
P(s_1^* = 2 | \Omega_1; \theta) &= P(s_1^* = 2, s_0^* = 1 | \Omega_1; \theta) + P(s_1^* = 2, s_0^* = 2 | \Omega_0; \theta) \\
&= \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_1)]^2}{2\sigma_1^2}\right) \times (p_{12} \times \pi_1)}{f(y_1 | \Omega_0; \theta)}
\end{aligned}$$

$$+ \frac{\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{[(y_t - \mu_2) - \phi(y_{t-1} - \mu_2)]^2}{2\sigma_1^2}\right) \times (p_{22} \times \pi_2)}{f(y_1|\Omega_0; \theta)} \quad (2.32)$$

Dengan demikian, *filtered state probability* hingga waktu T dapat dihitung jika $P(s_T = j|\psi_T; \eta)$ merepresentasikan hasil akhir dari iterasi proses *filtering*.

2.15.2 Smoothing

Menurut Ariyani, *et al.* (2014), Proses *smoothing* dilakukan untuk memperoleh estimasi yang lebih baik yaitu peluang nilai *state* dihitung berdasarkan informasi dari semua data pengamatan. Proses ini merupakan lanjutan dari proses *filtering*. Proses *smoothing* menghasilkan nilai *smoothed state probability*. Perhitungan proses *smoothing* adalah sebagai berikut.

$$\begin{aligned} & P(s_t^* = j, s_{t+1}^* = k|\Omega_T; \theta) \\ &= \frac{P(s_{t+1}^* = k|\Omega_T; \theta) \times P(s_t^* = j|\Omega_1; \theta) \times P(s_{t+1}^* = k|s_t^* = j, \Omega_1; \theta)}{P(s_{t+1}^* = k|\Omega_1; \theta)} \end{aligned} \quad (2.33)$$

Dilihat persamaan (2.30), proses *smoothing* untuk setiap kemungkinan $t = 1$ sebagai berikut.

$$\begin{aligned} & P(s_1^* = 1, s_2^* = 1|\Omega_T; \theta) \\ &= \frac{P(s_2^* = 1|\Omega_T; \theta) \times P(s_1^* = 1|\Omega_1; \theta) \times P(s_2^* = 1|s_1^* = 1, \Omega_1; \theta)}{P(s_2^* = 1|\Omega_1; \theta)} \end{aligned} \quad (2.34)$$

$$\begin{aligned} & P(s_1^* = 1, s_2^* = 2|\Omega_T; \theta) \\ &= \frac{P(s_2^* = 2|\Omega_T; \theta) \times P(s_1^* = 1|\Omega_1; \theta) \times P(s_2^* = 2|s_1^* = 1, \Omega_1; \theta)}{P(s_2^* = 2|\Omega_1; \theta)} \end{aligned} \quad (2.35)$$

$$\begin{aligned} & P(s_1^* = 2, s_2^* = 1|\Omega_T; \theta) \\ &= \frac{P(s_2^* = 1|\Omega_T; \theta) \times P(s_1^* = 2|\Omega_1; \theta) \times P(s_2^* = 1|s_1^* = 2, \Omega_1; \theta)}{P(s_2^* = 1|\Omega_1; \theta)} \end{aligned} \quad (2.36)$$

$$\begin{aligned} & P(s_1^* = 2, s_2^* = 2|\Omega_T; \theta) \\ &= \frac{P(s_2^* = 2|\Omega_T; \theta) \times P(s_1^* = 2|\Omega_1; \theta) \times P(s_2^* = 2|s_1^* = 2, \Omega_1; \theta)}{P(s_2^* = 2|\Omega_1; \theta)} \end{aligned} \quad (2.37)$$

Sehingga didapatkan *smoothed state probability* sebagai berikut.

$$\begin{aligned}
P(s_1^* = 1|\Omega_T; \theta) &= P(s_1^* = 1, s_2^* = 1|\Omega_T; \theta) + P(s_1^* = 1 | s_2^* = 2, \Omega_T; \theta) \\
&= \frac{P(s_2^* = 1|\Omega_T; \theta) \times P(s_1^* = 1|\Omega_1; \theta) \times P(s_2^* = 1|s_1^* = 1, \Omega_1; \theta)}{P(s_2^* = 1|\Omega_1; \theta)} \\
&\quad + \frac{P(s_2^* = 2|\Omega_T; \theta) \times P(s_1^* = 1|\Omega_1; \theta) \times P(s_2^* = 2|s_1^* = 1, \Omega_1; \theta)}{P(s_2^* = 2|\Omega_1; \theta)} \quad (2.38)
\end{aligned}$$

$$\begin{aligned}
P(s_1^* = 2|\Omega_T; \theta) &= P(s_1^* = 2, s_2^* = 1|\Omega_T; \theta) + P(s_1^* = 2 | s_2^* = 2, \Omega_T; \theta) \\
&= \frac{P(s_2^* = 1|\Omega_T; \theta) \times P(s_1^* = 2|\Omega_1; \theta) \times P(s_2^* = 1|s_1^* = 2, \Omega_1; \theta)}{P(s_2^* = 1|\Omega_1; \theta)} \\
&\quad + \frac{P(s_2^* = 2|\Omega_T; \theta) \times P(s_1^* = 2|\Omega_1; \theta) \times P(s_2^* = 2|s_1^* = 2, \Omega_1; \theta)}{P(s_2^* = 2|\Omega_1; \theta)} \quad (2.39)
\end{aligned}$$

Fungsi *log likelihood* dinyatakan sebagai berikut.

$$\ln L(\theta) = \sum_{t=1}^T \ln \left[\sum_{j=1}^N f(y_t, s_t^* = j|\Omega_T; \theta) \right] \quad (2.40)$$

Pada persamaan (2.37) yang merupakan tahap akhir, fungsi *log likelihood* diturunkan terhadap setiap parameter untuk menentukan nilai maksimum yang akan digunakan untuk estimasi.

1. Parameter μ_1

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \mu_1} &= \frac{\partial}{\partial \mu_1} \sum_{t=1}^T \ln[f(y_t, s_t^* = 1|\Omega_T; \theta)] \\
0 &= \frac{\partial}{\partial \mu_1} \sum_{t=1}^T \ln[f(y_t|s_t^* = 1|\Omega_T; \theta) \times P(s_t^* = 1|\Omega_T; \theta)] \\
&= \frac{\partial}{\partial \mu_1} \left[\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(\frac{-(y_t - \mu_1)^2}{2\sigma^2} \times P(s_t^* = 1|\Omega_T; \theta)\right) \right] \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \frac{(y_t - \mu_1)}{\sigma^2} \times P(y_t, s_t^* = 1|\Omega_T; \theta) \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \frac{(y_t - \mu_1)}{\sigma^2} \times P(s_t^* = 1|y_t, \Omega_T; \theta) \times f(y_t; \theta) \\
&= \sum_{t=1}^T \frac{(y_t - \mu_1)}{\sigma^2} \times P(s_t^* = 1|y_t, \Omega_T; \theta)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^T (y_t - \mu_1) \times P(s_t^* = 1 | \Omega_T; \theta) \\
&= \sum_{t=1}^T y_t \times P(s_t^* = 1 | \Omega_T; \theta) = T\mu_1 \sum_{t=1}^T P(s_t^* = 1 | \Omega_T; \theta) \\
\hat{\mu}_1 &= \frac{\sum_{t=1}^T y_t \times P(s_t^* = 1 | \Omega_T; \theta)}{T \sum_{t=1}^T P(s_t^* = 1 | \Omega_T; \theta)} \tag{2.41}
\end{aligned}$$

setelah itu diperoleh,

$$\begin{aligned}
\frac{\partial^2}{\partial \mu_1^2} \ln L(\theta) &= \frac{\partial}{\partial \mu_1} \left(\sum_{t=1}^T (y_t - \mu_1) \times P(s_t^* = 1 | \Omega_T; \theta) \right) \\
&= \frac{\partial}{\partial \mu_1} \left(\sum_{t=1}^T y_t \times P(s_t^* = 1 | \Omega_T; \theta) - T\mu_1 \times P(s_t^* = 1 | \Omega_T; \theta) \right) \\
&= -T \times P(s_t^* = 1 | \Omega_T; \theta) < 0 \tag{2.42}
\end{aligned}$$

karena $\frac{\partial^2}{\partial \mu_1^2} \ln L(\theta) < 0$ maka dapat disimpulkan bahwa fungsi *log likelihood* berada pada nilai maksimum.

2. Parameter μ_2

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \mu_2} &= \frac{\partial}{\partial \mu_2} \sum_{t=1}^T \ln[f(y_t, s_t^* = 2 | \Omega_T; \theta)] \\
0 &= \frac{\partial}{\partial \mu_2} \sum_{t=1}^T \ln[f(y_t | s_t^* = 2 | \Omega_T; \theta) \times P(s_t^* = 2 | \Omega_T; \theta)] \\
&= \frac{\partial}{\partial \mu_2} \left[\frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(\frac{-(y_t - \mu_2)^2}{2\sigma^2}\right) \times P(s_t^* = 2 | \Omega_T; \theta) \right] \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \frac{(y_t - \mu_2)}{\sigma^2} \times P(y_t, s_t^* = 2 | \Omega_T; \theta) \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \frac{(y_t - \mu_2)}{\sigma^2} \times P(s_t^* = 2 | y_t, \Omega_T; \theta) \times f(y_t; \theta) \\
&= \sum_{t=1}^T \frac{(y_t - \mu_2)}{\sigma^2} \times P(s_t^* = 2 | y_t, \Omega_T; \theta)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^T (y_t - \mu_2) \times P(s_t^* = 2 | \Omega_T; \theta) \\
&= \sum_{t=1}^T y_t \times P(s_t^* = 2 | \Omega_T; \theta) = T\mu_2 \sum_{t=1}^T P(s_t^* = 2 | \Omega_T; \theta) \\
\hat{\mu}_1 &= \frac{\sum_{t=1}^T y_t \times P(s_t^* = 2 | \Omega_T; \theta)}{T \sum_{t=1}^T P(s_t^* = 2 | \Omega_T; \theta)} \tag{2.43}
\end{aligned}$$

setelah itu diperoleh,

$$\begin{aligned}
\frac{\partial^2}{\partial \mu_1^2} \ln L(\theta) &= \frac{\partial}{\partial \mu_2} \left(\sum_{t=1}^T (y_t - \mu_2) \times P(s_t^* = 2 | \Omega_T; \theta) \right) \\
&= \frac{\partial}{\partial \mu_2} \left(\sum_{t=1}^T y_t \times P(s_t^* = 2 | \Omega_T; \theta) - T\mu_2 \times P(s_t^* = 2 | \Omega_T; \theta) \right) \\
&= -T \times P(s_t^* = 2 | \Omega_T; \theta) < 0 \tag{2.44}
\end{aligned}$$

karena $\frac{\partial^2}{\partial \mu_2^2} \ln L(\theta) < 0$ maka dapat disimpulkan bahwa fungsi *log likelihood* berada pada nilai maksimum.

3. Parameter σ^2

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} \sum_{t=1}^T \ln[f(y_t, s_t^* = j | \Omega_T; \theta)] \\
0 &= \frac{\partial}{\partial \sigma^2} \sum_{t=1}^T \ln[f(y_t | s_t^* = j | \Omega_T; \theta) \times P(s_t^* = j | \Omega_T; \theta)] \\
&= \frac{\partial}{\partial \sigma^2} \left[\frac{1}{\sigma \sqrt{2\pi}} \exp\left(\frac{-(y_t - \mu_j)^2}{2\sigma^2} \right) \times P(s_t^* = j | \Omega_T; \theta) \right] \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \left(-\frac{1}{2\sigma^2} + \frac{(y_t - \mu_j)^2}{2\sigma^4} \right) \times P(y_t, s_t^* = j | \Omega_T; \theta) \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \left(-\frac{1}{2\sigma^2} + \frac{(y_t - \mu_j)^2}{2\sigma^4} \right) \times P(s_t^* = j | y_t, \Omega_T; \theta) \times f(y_t; \theta) \\
&= \sum_{t=1}^T \left(-\frac{1}{2\sigma^2} + \frac{(y_t - \mu_j)^2}{2\sigma^4} \right) \times P(s_t^* = j | y_t, \Omega_T; \theta)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^T (-\sigma^2 + (y_t - \mu_2)^2) \times P(s_t^* = j | \Omega_T; \theta) \\
&= \sum_{t=1}^T (y_t - \mu_2)^2 \times P(s_t^* = j | \Omega_T; \theta) = T\sigma^2 \sum_{t=1}^T P(s_t^* = j | \Omega_T; \theta) \\
\hat{\sigma}_1 &= \frac{\sum_{t=1}^T (y_t - \mu_2)^2 \times P(s_t^* = j | \Omega_T; \theta)}{T \sum_{t=1}^T P(s_t^* = j | \Omega_T; \theta)} \tag{2.45}
\end{aligned}$$

setelah itu diperoleh,

$$\begin{aligned}
\frac{\partial^2}{\partial(\sigma^2)^2} \ln L(\theta) &= \frac{\partial}{\partial\sigma^2} \left(\sum_{t=1}^T (-\sigma^2 + (y_t - \mu_j)^2) \times P(s_t^* = j | \Omega_T; \theta) \right) \\
&= \frac{\partial}{\partial\sigma^2} \left(\sum_{t=1}^T (y_t - \mu_j)^2 \times P(s_t^* = j | \Omega_T; \theta) - T\sigma^2 \times P(s_t^* = j | \Omega_T; \theta) \right) \\
&= -T \times P(s_t^* = j | \Omega_T; \theta) < 0 \tag{2.46}
\end{aligned}$$

karena $\frac{\partial^2}{\partial(\sigma^2)^2} \ln L(\theta) < 0$ maka dapat disimpulkan bahwa fungsi *log likelihood* berada pada nilai maksimum.

4. Parameter \emptyset

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \emptyset} &= \frac{\partial}{\partial \emptyset} \sum_{t=1}^T \ln[f(y_t, s_t^* = j | \Omega_T; \theta)] \\
0 &= \frac{\partial}{\partial \emptyset} \sum_{t=1}^T \ln[f(y_t | s_t^* = j | \Omega_T; \theta) \times P(s_t^* = j | \Omega_T; \theta)] \\
&= \frac{\partial}{\partial \emptyset} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(y_t - \mu_j)^2}{2\sigma^2} \times P(s_t^* = j | \Omega_T; \theta)\right) \right] \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \left(\sum_{j=1}^N \frac{\emptyset(y_t - \mu_j)^2}{2\sigma^2} \right) \times P(y_t, s_t^* = j | \Omega_T; \theta) \\
&= \sum_{t=1}^T \frac{1}{f(y_t; \theta)} \times \left(\sum_{j=1}^N \frac{\emptyset(y_t - \mu_j)^2}{2\sigma^2} \right) \times P(y_t, s_t^* = j | \Omega_T; \theta) \times f(y_t; \theta) \\
&= \sum_{t=1}^T \left(\sum_{j=1}^N \frac{\emptyset(y_t - \mu_j)^2}{2\sigma^2} \right) \times P(s_t^* = j | y_t, \Omega_T; \theta)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^T \sum_{j=1}^N (y_t - \mu_j)^2 \times P(s_t^* = j | \Omega_T; \theta) \\
&= T\phi \sum_{t=1}^T \sum_{j=1}^N P(s_t^* = j | \Omega_T; \theta) \\
\hat{\phi} &= \frac{\sum_{t=1}^T \sum_{j=1}^N (y_t - \mu_j)^2 \times P(s_t^* = j | \Omega_T; \theta)}{T \sum_{t=1}^T \sum_{j=1}^N P(s_t^* = j | \Omega_T; \theta)} \tag{2.47}
\end{aligned}$$

setelah itu diperoleh,

$$\begin{aligned}
\frac{\partial^2}{\partial \phi^2} \ln L(\theta) &= \frac{\partial}{\partial \phi} \sum_{t=1}^T \left(\sum_{j=1}^N -\phi (y_t - \mu_j)^2 \right) \times P(s_t^* = j | \Omega_T; \theta) \\
&= \frac{\partial}{\partial \phi} \left(\sum_{t=1}^T \sum_{j=1}^N (y_t - \mu_j)^2 \times P(s_t^* = j | \Omega_T; \theta) - T\phi \times P(s_t^* = j | \Omega_T; \theta) \right) \\
&= -T \times P(s_t^* = j | \Omega_T; \theta) \leq 0 \tag{2.48}
\end{aligned}$$

karena $\frac{\partial^2}{\partial \phi^2} \ln L(\theta) < 0$ maka dapat disimpulkan bahwa fungsi *log likelihood* berada pada nilai maksimum.

Jika diketahui $X_t = ((y_t - \mu_{s_t^*}), \dots, (y_{-r+1} - \mu_{s_{-r+1}^*}))$ untuk rantai *Markov* yang terdiri dari 2 *state* maka estimasi nilai probabilitas transisi p_{11} dan p_{22} dapat dihitung sebagai berikut.

1. Parameter p_{11}

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial p_{11}} &= \frac{\partial}{\partial p_{11}} \sum_{t=2}^T \ln[f(\Omega_T, s_t^*, s_{t-1}^* | X_T; \theta)] \\
0 &= \sum_{t=2}^T \left\{ \begin{array}{l} P(s_t^* = 1, s_{t-1}^* = 1 | X_T; \theta) \times \frac{1}{p_{11}} \\ P(s_t^* = 2, s_{t-1}^* = 1 | X_T; \theta) \times \frac{1}{1 - p_{11}} \end{array} \right\} \\
&= \frac{1 - p_{11}}{p_{11}} = \frac{\sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 1 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 1 | \Omega_T; \theta)} \\
p_{11} &= \frac{\sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 1 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 1 | \Omega_T; \theta) + \sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 1 | \Omega_T; \theta)}
\end{aligned}$$

$$= q = \frac{\sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 1 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 1 | \Omega_T; \theta)} \quad (2.49)$$

2. Parameter p_{22}

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial p_{22}} &= \frac{\partial}{\partial p_{22}} \sum_{t=2}^T \ln[f(\Omega_T, s_t^*, s_{t-1}^* | X_T; \theta)] \\ 0 &= \sum_{t=2}^T \left\{ \begin{array}{l} P(s_t^* = 2, s_{t-1}^* = 2 | X_T; \theta) \times \frac{1}{p_{22}} \\ P(s_t^* = 1, s_{t-1}^* = 2 | X_T; \theta) \times \frac{1}{1 - p_{22}} \end{array} \right\} \\ &= \frac{1 - p_{22}}{p_{22}} = \frac{\sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 2 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 2 | \Omega_T; \theta)} \\ p_{22} &= \frac{\sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 2 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 2 | \Omega_T; \theta) + \sum_{t=2}^T P(s_t^* = 1, s_{t-1}^* = 2 | \Omega_T; \theta)} \\ &= p = \frac{\sum_{t=2}^T P(s_t^* = 2, s_{t-1}^* = 2 | \Omega_T; \theta)}{\sum_{t=2}^T P(s_t^* = 2 | \Omega_T; \theta)} \quad (2.50) \end{aligned}$$

Sehingga diperoleh sebuah matriks transisi sebagai berikut:

$$P = \begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (2.51)$$

dengan p_{11} merepresentasikan probabilitas transisi dari $s_{t-1}^* = 1$ ke $s_t^* = 1$, p_{12}

merepresentasikan probabilitas transisi dari $s_{t-1}^* = 1$ ke $s_t^* = 2$, p_{21}

merepresentasikan probabilitas transisi dari $s_{t-1}^* = 2$ ke $s_t^* = 1$, dan p_{22}

merepresentasikan probabilitas transisi dari $s_{t-1}^* = 2$ ke $s_t^* = 2$. Jumlah dari

$p_{11} + p_{12} = 1$ dan jumlah $p_{21} + p_{22} = 1$.

2.16 Pemilihan Model Terbaik

Pemilihan model terbaik bisa ditentukan menggunakan kriteria *in-sample* yaitu dapat berdasarkan nilai *Akaike's Information Criterion* (AIC_c) dan menggunakan kriteria *out-sample* yaitu dengan menggunakan *Mean Absolute Percentage Error* (MAPE).

2.16.1 Akaike's Information Criterion corrected (AIC_c)

Menurut Sen & Shitan (2002), *Bayesian Information Criterion* (BIC) merupakan modifikasi dari *Akaike's Information Criterion* (AIC) yang tidak efisien secara asimtotik. Oleh karena itu, disarankan versi koreksi yang bias dari AIC yang disebut sebagai *Akaike's Information Criterion corrected* (AIC_c). AIC_c digunakan secara luas sebagai salah satu kriteria pemilihan model terbaik dalam deret waktu yang ditulis sebagai berikut:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \quad (2.52)$$

dengan

AIC = *Akaike's Information Criterion*

k = jumlah parameter dalam model

n = ukuran sampel

Untuk menentukan nilai AIC dapat ditulis sebagai berikut:

$$AIC = -2 \ln(L) + 2k \quad (2.53)$$

dengan

k = jumlah parameter dalam model

L = nilai maksimum dari fungsi *likelihood* dalam model

2.16.2 Mean Absolute Percentage Error (MAPE)

Menurut Makridakis *et al.* (2008), *Mean Absolute Percentage Error* (MAPE) merupakan kriteria yang sering digunakan sebagai alat pengukur kesalahan pada kesalahan melalui akurasi suatu model. MAPE adalah tolak ukur yang digunakan untuk mengidentifikasi perbedaan antara hasil prediksi dan nilai aktual atau observasi. Mengidentifikasi perbedaan tersebut dilakukan dengan

membandingkan data aktual dan data peramalan. MAPE dirumuskan sebagai berikut.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (2.54)$$

dengan

y_t = nilai pengamatan waktu ke- t

\hat{y}_t = nilai peramalan waktu ke- t

n = banyaknya pengamatan

Selain itu, nilai MAPE memiliki standar yang bisa digunakan dalam mengevaluasi kualitas sebuah model. Keriteria nilai MAPE tertera pada Tabel 2. (Martin & Witt, 1989).

Tabel 2. Kriteria nilai MAPE

Nilai MAPE	Kriteria
< 10%	Sangat Baik
10% - 20%	Baik
20% - 50%	Cukup
> 50%	Buruk

2.17 Pemeriksaan Diagnostik

Pemeriksaan diagnostik bertujuan untuk mengetahui kelayakan model melalui nilai probabilitas parameter pada model. Prosesnya dimulai dengan menentukan model dan estimasi parameternya serta dievaluasi apakah model tersebut sesuai asumsi yang diperlukan atau tidak. Asumsi utama yang mendasari model tersebut adalah bahwa residual α_t berdistribusi normal atau merupakan *white noise*.

Asumsi normalitas diidentifikasi dengan uji Kolmogorov-Smirnov dan asumsi *white noise* diidentifikasi menggunakan plot *time series*, uji Breusch-Pagan dan uji Ljung-Box.

2.17.1 Uji Normalitas

Menurut Difinubun, *et all.* (2023), untuk mengetahui kenormalan residu suatu model dapat menggunakan uji normalitas residual. Tujuan dari uji normalitas yaitu mengetahui apakah residu pada model berdistribusi normal atau tidak. Salah satu uji normalitas dapat dilakukan menggunakan uji Kolmogorov-Smirnov.

Tahap pengujinnya adalah sebagai berikut:

Hipotesis:

$$H_0 : F_{0(t)} = S_{(t)} \text{ untuk semua nilai } t, \text{ dengan } F_{0(t)} \sim N(0, \sigma^2)$$

$$H_1 : F_{0(t)} \neq S_{(t)}, t = 1, 2, \dots, T$$

Statistik uji:

$$D = \text{Sup} |S_{(t)} - F_{0(t)}| \quad (2.55)$$

dengan

$S_{(t)}$ = fungsi distribusi komulatif empiris dari sampel

$F_{0(t)}$ = fungsi distribusi komulatif dari distribusi teoritis yang sedang di uji

Kriteria uji:

Jika $D > K_{(1-\alpha)}$ atau dengan $\alpha = 0,05$ jika $p - \text{value} < \alpha$ sehingga tolak H_0 dan terima H_1 jika sebaliknya. Nilai $K_{(1-\alpha)}$ dapat diketahui melalui tabel

Kolmogorov-Smirnov.

2.17.2 Uji Asumsi

Salah satu cara untuk menguji asumsi *white noise* yaitu dengan plot data. Plot data *time series* memiliki banyak jenis, salah satunya scatter plot. Scatterplot sering kali menjadi grafik pilihan untuk menampilkan hubungan antara dua variabel. Scatterplot berguna untuk analisis eksplorasi, tetapi dapat melakukan lebih dari

sekadar mengidentifikasi korelasi. Scatter plot dapat menggambarkan berbagai pola dan hubungan misalnya hubungan linier atau nonlinier, hubungan positif (langsung) atau negatif (terbalik), serta konsentrasi atau penyebaran titik data. Jika titik-titik pada plot membentuk garis lurus berarti bahwa hubungan antar variabel tersebut bersifat linier. Sebaliknya, jika titik-titik pada plot membentuk garis yang tidak lurus berarti bahwa hubungan antar variabel tersebut bersifat nonlinier (Lukman, *et all.*, 2024).

2.17.3 Uji Heteroskedastisitas

Menurut Alwi, *et all.* (2023), hipotesis yang digunakan untuk mendeteksi heteroskedastisitas dalam model regresi dapat menggunakan uji Breusch-Pagan ini sebagai berikut:

$H_0 : \sigma^2 = 0$ (tidak terjadi heteroskedastisitas)

$H_1 : \sigma^2 \neq 0$ (terjadi heteroskedastisitas)

Adapun statistik uji Breusch-Pagan berdasarkan koefisien determinasi (R^2) yang mengikuti distribusi *chi-square* (χ^2) dengan derajat kebebasan sama dengan jumlah variabel independen (k) sebagai berikut:

$$BP = \frac{n}{2} \times R^2 \quad (2.56)$$

dengan

n = jumlah observasi

R^2 = koefisien determinasi dari regresi residual kuadrat

Jika $p - value \leq \alpha$ maka tolak H_0 yang berarti terjadi heteroskedastisitas. Sedangkan, jika $p - value > \alpha$ maka terima H_0 yang berarti tidak terjadi heteroskedastisitas.

2.17.4 Uji Autokorelasi

Untuk menguji apakah nilai autokorelasi sisaan sama dengan nol atau tidak dapat menggunakan uji Ljung-Box. Jika autokorelasi sisaan bernilai nol maka error yang berarti *white noise*, sehingga model bisa digunakan dalam peramalan (Khoerunnisa, *et all.*, 2022). Berikut ini adalah pengujian Ljung-Box.

Hipotesis:

$H_0 : p_1 = p_2 = \dots = p_k = 0$ (residual memenuhi syarat *white noise*)

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

Statistik uji:

$$Q^* = n(n+2) \sum_{k=1}^K \frac{r_k^2}{n-k} \quad (2.57)$$

dengan

Q^* = statistik Uji Ljung-Box

n = jumlah observasi atau data sampel

k = lag ke- k

K = jumlah lag yang diuji

r_k^2 = autokorelasi pada lag ke- k

p dan q = order dari ARMA (p, q)

Kriteria uji:

Tolak H_0 jika $p\text{-value} < \alpha$ atau $Q^* > \chi_{\alpha; df=K-pq}^2$.

2.18 Jumlah Penumpang Kereta Api

Salah satu elemen yang penting untuk keberlangsungan hidup suatu negara adalah transportasi yang handal. Transportasi yang handal adalah faktor pendukung

utama yang menentukan kemajuan suatu negara. Kegiatan sehari-hari masyarakat Indonesia sangat bergantung pada aksesibilitas. Setiap proses perpindahan, baik orang maupun barang, dapat difasilitasi dengan adanya transportasi. Kebanyakan orang menggunakan kereta api sebagai salah satu model transportasi masal yang tersedia bagi masyarakat (Effendi & Yenita, 2023).

Kereta api adalah cara transportasi darat yang digunakan untuk perjalanan jauh. Selain cepat, kereta api juga membantu menghindari kemacetan di jalan, terutama di kota-kota besar. Kereta api tidak hanya murah dan ramah lingkungan, tetapi juga memiliki banyak fasilitas. Jumlah penumpang kereta api biasanya bertambah pada hari libur Natal dan tahun baru, hari raya Idul Fitri, serta hari besar lainnya. Dalam mengatasi jumlah penumpang yang terus bertambah, peramalan diperlukan untuk menyediakan fasilitas yang diperlukan dalam mengoperasikan transportasi kereta api tersebut (Utomo & Fanani, 2020).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap Tahun Ajaran 2024/2025 yang bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

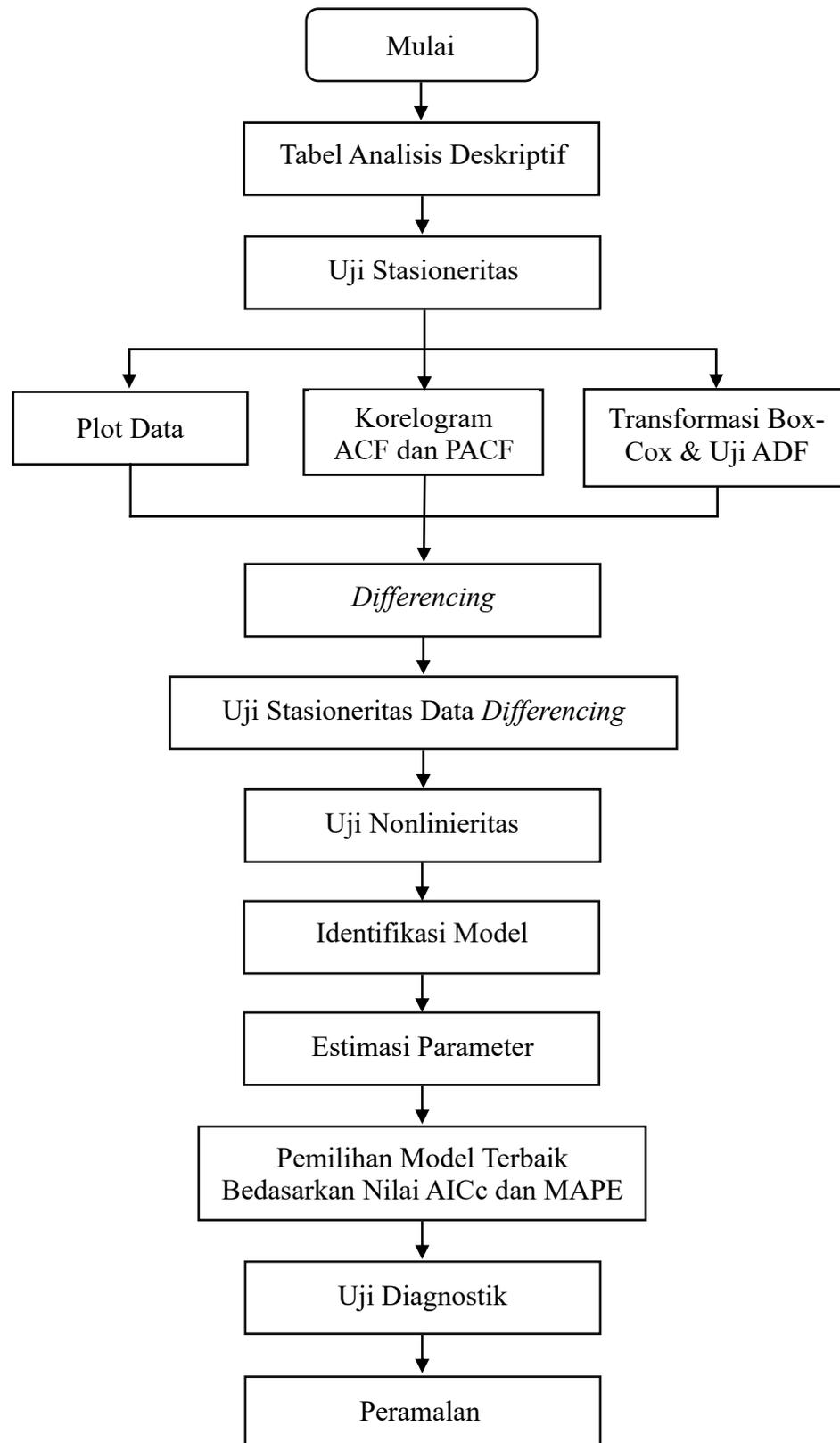
Pada penelitian ini data yang digunakan adalah data *time series* sekunder yang diambil dari <https://bps.go.id> yaitu data bulanan jumlah penumpang kereta api di Sumatera periode Januari 2010 – Juli 2024.

3.3 Metode Penelitian

Berikut langkah-langkah yang digunakan dalam penelitian ini, yaitu:

1. Melakukan analisis deskriptif pada data.
2. Melakukan uji stasioneritas data.
 - a. Membuat plot data.
 - b. Melihat pola korelogram ACF dan PACF.
 - c. Melakukan transformasi Box-Cox.

- d. Melakukan uji *Augmented Dickey-Fuller* (ADF) untuk melihat nilai p – *value*.
3. Jika data tidak stasioner, maka dilakukan *differencing* data.
4. Melakukan kembali uji stasioneritas dengan melihat plot, pola korelogram ACF dan PACF, transformasi Box-Cox dan uji ADF pada data yang telah dilakukan *differencing*.
5. Melakukan uji nonlinieritas menggunakan uji Terasvirta. Jika hasil dari uji nonlinieritas sesuai maka dapat disimpulkan data sesuai dengan model nonlinier sehingga dapat dilanjutkan ke tahap berikutnya.
6. Penentuan *state* yaitu *state* 1 dan *state* 2. Nilai μ_{st} digunakan untuk menentukan *state*. μ_{st} merupakan rata-rata yang dipengaruhi oleh *state*. Pada penelitian ini terdapat dua *state* yaitu apresiasi dan depresiasi sehingga μ_{st} ada dua yakni μ_1 dan μ_2 .
7. Melakukan estimasi parameter, metode yang digunakan dalam estimasi parameter untuk mendapatkan orde yang sesuai pada model MSAR adalah MLE yang dikombinasikan dengan *filtering* dan *smoothing*. Hal ini disebabkan karena pada MSAR terdapat parameter p_{ij} yang merepresentasikan probabilitas dari masing-masing *state*.
8. Melakukan pemilihan model terbaik dengan memeriksa nilai AIC_c dan MAPE.
9. Melakukan uji diagnostik pada model terbaik.
 - a. Memeriksa normalitas residual dengan menggunakan uji Kolmogorov-Smornov.
 - b. Memeriksa apakah residual bersifat *white noise* dengan melihat plot data, uji Breusch-Pagan dan uji Ljung-Box.
10. Melakukan peramalan periode 12 bulan dengan model terbaik.



Gambar 1. Diagram Alir Pemodelan MSAR

V. KESIMPULAN

Berdasarkan hasil dan pembahasan pada penelitian yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Model MSAR terbaik yang dapat digunakan untuk melakukan peramalan jumlah penumpang kereta api di Sumatera adalah model MS(2)AR(4) dengan nilai AICc sebesar 1919,0448 dan MAPE sebesar 3,8878 dalam *state* 1 dan *state* 2. Model yang dihasilkan adalah sebagai berikut:

$$\text{State 1} = -0,3009(z_{t-1} - \mu_{St-1}) - 0,1915(z_{t-2} - \mu_{St-2}) - 0,0717(z_{t-3} - \mu_{St-3}) - 0,0025(z_{t-4} - \mu_{St-4}) + \varepsilon_t$$

$$\text{State 2} = -0,571(z_{t-1} - \mu_{St-1}) + 1,6390(z_{t-2} - \mu_{St-2}) + 0,3481(z_{t-3} - \mu_{St-3}) - 0,4761(z_{t-4} - \mu_{St-4}) + \varepsilon_t$$

2. Hasil dari peramalan jumlah penumpang kereta api di Sumatera untuk bulan Agustus 2024 – Juli 2025 yaitu 1,534956e+02, 7,573674e+02, 3,799271e+03, 1,908896e+04, 9,597587e+04, 4,825546e+05, 2,426231e+06, 1,219882e+07, 6,133435e+07, 3,083824e+08, 1,550513e+09 dan 7,795810e+09.

DAFTAR PUSTAKA

- Ahmad, F. 2020. Penentuan Metode Peramalan pada Produksi *Part New Granada Bowl St* di PT.X. *Jurnal Integrasi Sistem Industri*. 7(1): 31-39.
- Alwi, W., Sauddin, A., & Islamiah, N. 2023. Faktor-faktor Yang Mempengaruhi Angka Harapan Hidup di Sulawesi Selatan Menggunakan Analisis Regresi. *Jurnal Matematika dan Statistika serta Aplikasinya*. 11(1): 72-80.
- Anton, H. & Rorres, C. 2004. *Elementary Linier Algebra: Application Version*. 8th Edition. Alih Bahasa: Revina & Irzam. Erlangga, Jakarta.
- Ariyani, F.D., Warsito, B., & Yasin, H. 2014. Pemodelan *Markov Switching Autoregressive*. *Jurnal Gaussian*. 3(3): 381-390.
- BPS. 2024. Jumlah Keberangkatan Penumpang Transportasi Umum Juni 2024. Jumlah Keberangkatan Penumpang Transportasi Umum Juni 2024 Naik 8,83% - Berita - Badan Pusat Statistik Provinsi Lampung (bps.go.id). Diakses pada Tanggal 05 Oktober 2024.
- Cox, D.R. & Miller, H. D. 1965. *The Theory of Stochastic Process*. Chapman and Hall, London.
- Difinubun, S.H., Nara, O.D., & Abdin, M. 2023. Analisis Pengaruh Sumber Daya Manusia Terhadap Aspek Kinerja Pekerja Pada Proyek Pembangunan Gedung Laboratorium Terpadu Pendukung Blok Masela Universitas Pattimura. *Jurnal Agragate*. 2(1): 76-86.
- Effendi, H & Yenita, Y. 2023. Faktor Penentu Loyalitas Penumpang Kereta Api Antar Kota Lintas Pulau Jawa. *Jurnal Manajemen Maranatha*. 22(2): 169-182.

- Franses, P.H., Dijk, D.V., and Opschoor, A. 2014. *Time Series Model : for Bussines and Economic Forecasting*. 2th Edition. Cambridge, University Press.
- Gujarati, D. N. & Porter, D. C. 2009. *Basic Econometrics*. McGraw-hill, New York.
- Hamilton, J.D. 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*. **57**(2): 357-384.
- Hamilton, J.D. 1994. *Time Series Analysis*. Princeton, University Press.
- Hasanah, I.A.W. & Iriani, S.S. 2024. Pengaruh Harga Tiket Tarif Khusus, Kualitas Layanan Petugas, dan Fasilitas Fisik Stasiun Terhadap Keputusan Penggunaan Transportasi Kereta Api. *Jurnal Ilmu Manajemen*. **12**(2): 211-226.
- Juniawati, D.A & Hajarisman, N. 2023. Pendekatan Baru dari Transformasi Box-Cox untuk Mengatasi Masalah Ketidaknormalan pada Variabel Respon. *Jurnal Statistika*. **3**(2): 556- 565.
- Khoerunnisa, A., Nur, I.M., & Arum, P.R. 2022. Metode *Markov Switching Autoregressive* (MSAR) untuk Peramalan Indeks Saham Syariah Indonesia (ISSI), hlm. 608-623. Prosiding Seminar Nasional UNIMUS, Semarang.
- Kim, C.J. dan Nelson, C.R. 1999. *State Space Models with Regime Switching, Classical and Gibbs Sampling Approaches with Applications*. Cambridge, MA: MIT Press.
- Lukman, Wardhani, R.P., Sarungu, S., & Andrianti, I. 2024. Penggunaan Metode Seven Tool Dalam Pembelajaran Pengendalian Mutu Secara Statistik. *Jurnal Teknosains Kodepena*. **5**(1): 27-33.
- Makridakis, S., Wheelwright, S.C., & Hyndman, R.J. 2008. *Forecasting Method and Applications*. John & Sons, New York.
- Martin, C.A. & Witt, S.F. 1989. Accuracy of Econometric Forecasts of Tourism. *Annals of Tourism Research*. **16**(3): 407-428.
- Mirza, M.Y. 2023. Transportasi Kereta Api SCS (*Semarang Cheribon Stoomstram Maatschappij*) di Kabupaten Kendal Pada Tahun 1897-1930. *Jurnal Sejarah*. **3**(1): 65-71.

- Pankratz, A. 1991. *Forecasting with Dynamic Regression Models*. John Wiley & Sons, New York.
- Privault, N. 2013. *Understanding Markov Chain : Examples and Applications*. Springer, Singapore.
- Rizki, S.W., Martha, S., & Apriliyanti, R. 2024. Model *Markov Switching Autoregressive* pada Data Covid-19 di Indonesia. *Jurnal Probabilitas dan Statistik*. **5**(1): 1-5.
- Ross, S. 2010. *A First Course in Probability*. 8th Edition. Pearson Education, Inc., USA.
- Sen, K.L. & Shitan, M. 2002. The Performance of AICC as an Order Selection Criterion in ARMA Time Series Models. *Pertanika Journal Science & Tecnology*. **10**(1): 25-33.
- Teräsvirta, T. 1994. Specification, Estimation, and Evaluating of Smooth Transition Autoregressive Models. *Jurnal of the American Statistical Assosiation*. **89**(425): 511-512.
- Utomo, P. & Fanani, A. 2020. Peramalan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA). *Jurnal Matematika*. **1**(1): 169-178.
- Wahyudi, O., Warsito. B., & Prahutama A. 2015. Pemodelan Tingkat Inflasi Indonesia Menggunakan *Markov Switching Autoregressive Conditional Heteroskedasticit*. *Jurnal Gaussian*. **4**(1): 103-111.
- Wei, W. W. S. 2006. *Univariate and Multivariate Methods*. 2th Edition. Pearson Education, London.