

ANALISIS INTERVENSI DATA DERET WAKTU
UNTUK PERAMALAN
PENDAPATAN DOMESTIK BRUTO INDONESIA

Skripsi

Disusun untuk melengkapi syarat-syarat
guna memperoleh gelar Sarjana Sains



KURNIA WANTO
3125090115

PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA
2016

LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI

ANALISIS INTERVENSI DATA DERET WAKTU

UNTUK PERAMALAN

PENDAPATAN DOMESTIK BRUTO INDONESIA

Nama : Kurnia Wanto

No. Registrasi : 3125090115

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan	: Prof. Dr. Suyono, M.Si. NIP. 19671218 199303 1 005
Wakil Penanggung Jawab			
Pembantu Dekan I	: Dr. Muktiningsih, M.Si. NIP. 19640511 198903 2 001
Ketua	: Drs. Mulyono, M.Kom. NIP. 19660517 199403 1 003
Sekretaris	: Med Irzal, M.Kom. NIP. 19770615 200312 1 001
Penguji	: Eti Dwi Wiraningsih, S.Pd, M.Si. NIP. 19810203 200604 2 001
Pembimbing I	: Prof. Dr. Suyono, M.Si. NIP. 19671218 199303 1 005
Pembimbing II	: Dian Handayani, M.Si. NIP. 19740415 199803 2 001

Dinyatakan lulus ujian skripsi tanggal: 28 Juli 2016

PERSEMBAHANKU...

"Sungguh, Kami telah memberikan kepadamu kemenangan yang nyata, agar Allah memberikan ampunan kepadamu (Muhammad) atas dosamu yang lalu dan yang akan datang, serta menyempurnakan nikmat-Nya atasmu dan menunjukimu ke jalan yang lurus..."

(QS. Al Fath : 1-2)

Skripsi ini kupersembahkan untuk....
kedua orangtuaku Bapak Surono dan Ibu Sunarni,
saudaraku Widhi Arti, Andi Setiawan, Devi Rahmawati, serta keluarga besarku.
"Terima kasih atas dukungan, do'a, serta kasih sayang kalian..."

ABSTRACT

Kurnia Wanto, 3125090115. Intervention Analysis on Time Series Data for Forecasting GDP Indonesia. Thesis. Faculty of Mathematics and Natural Sciences, State University of Jakarta. 2016.

This thesis discusses how to model and forecast the time series data containing an intervention when the time of the intervention has known. Intervention analysis method used to measure the magnitude and duration of effect of an intervention. A time series data containing interventions Z_t , grouped into the data prior to the intervention and the data after the intervention. Data prior to the intervention N_t modeled using Autoregressive Integrated Moving Average (ARIMA) models. The magnitude and duration of the effect of the intervention were analyzed by comparing the results of the forecast data prior to the intervention using ARIMA models with observational data after the intervention. Forecasting time series data containing the intervention, conducted by calculating the value of the forecast data before the intervention and the future value of the function of intervention $f(I_t)$.

Keywords : Time Series Data, ARIMA, Forecasting, Intervention, Intervention Analysis Method, Wold Decomposition.

ABSTRAK

KURNIA WANTO, 3125090115. Analisis Intervensi pada Data Deret Waktu untuk Peramalan Pendapatan Domestik Bruto Indonesia. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. 2016.

Skripsi ini membahas bagaimana memodelkan dan meramalkan data deret waktu yang mengandung suatu intervensi saat waktu terjadinya intervensi diketahui. Metode analisis intervensi digunakan untuk mengukur besar dan lamanya dampak dari suatu intervensi. Suatu data deret waktu yang mengandung intervensi Z_t , dikelompokkan menjadi data sebelum terjadinya intervensi dan data setelah terjadinya intervensi. Data sebelum terjadinya intervensi N_t dimodelkan dengan menggunakan model *Auto Regressive Integrated Moving Average* (ARIMA). Besar dan lamanya dampak dari intervensi, dianalisis dengan membandingkan antara hasil peramalan dari data sebelum intervensi menggunakan model ARIMA dengan data observasi setelah intervensi. Peramalan data deret waktu yang mengandung intervensi, dilakukan dengan menghitung nilai peramalan data sebelum intervensi dan future value dari fungsi intervensi $f(I_t)$.

Kata kunci : Data Deret Waktu, ARIMA, Peramalan, Intervensi, Metode Analisis Intervensi, Dekomposisi Wold.

KATA PENGANTAR

Puji syukur kepada Allah SWT atas pemberian pengetahuan dan kemampuan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Intervensi Pada Data Deret Waktu Untuk Peramalan Pendapatan Domestik Bruto Indonesia" yang merupakan salah satu syarat dalam memperoleh gelar Sarjana Jurusan Matematika Universitas Negeri Jakarta.

Skripsi ini berhasil diselesaikan tak lepas dari bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Bapak Prof. Dr. Suyono, M.Si. selaku Dosen Pembimbing I dan Ibu Dian Handayani, M.Si. selaku Dosen Pembimbing II yang telah meluangkan waktu, saran, nasihat, serta arahan dalam pengerjaan skripsi ini sehingga menjadi lebih baik. Terima kasih banyak dan maaf atas segala kekurangan, semoga kesehatan selalu tercurah kepada Bapak dan Ibu.
2. Ibu Lukita Ambarwati S.Pd, M.Si. selaku Ketua Prodi Matematika FMIPA UNJ. Terima kasih atas segala bantuan dan kerja sama Ibu selama pengerjaan skripsi ini.
3. Bapak Drs. Sudarwanto, M.Si, DEA selaku Pembimbing Akademik atas segala bimbingan dan kerja sama Bapak selama perkuliahan penulis dan seluruh Bapak/Ibu dosen atas pengajaran yang telah diberikan, serta Ibu Lastri dan karyawan/karyawati FMIPA UNJ lainnya atas informasi dan bantuan yang diberikan dalam menyelesaikan skripsi ini.
4. Orang tua tercinta, Bapak Surono dan Ibu Sunarni yang senantiasa memberikan doa, perhatian, kesabaran, nasihat, serta bantuan secara moral maupun material.

5. Teman-teman satu prodi Matematika 2009, Desti, Vickri, Sulis, Hardi, Vlan, Eka, Hendi, Silvi, Novi, Yepni, Ria, Suci, Syifa, Jurin, Gilang, Benni, Dena, Veli. Terima kasih atas segala dukungan, kerjasama, perhatian, dan kebersamaan yang diberikan, semoga pertemanan ini terus terjalin.

Penulis menyadari skripsi ini masih jauh dari sempurna. Masukan dan kritik akan sangat berarti. Semoga skripsi ini dapat bermanfaat bagi pembaca.

Jakarta, 7 Agustus 2016



Kurnia Wanto

DAFTAR ISI

ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Pembatasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penulisan	3
1.6 Metode Penelitian	3
II LANDASAN TEORI	4
2.1 Data Deret Waktu	4
2.2 Kestasioneran Data Deret Waktu	6
2.2.1 Pembedaan (<i>differencing</i>)	7
2.2.2 Transformasi Box-Cox	8
2.3 Fungsi Autokorelasi dan Autokorelasi Parsial	9
2.3.1 Fungsi Autokorelasi (ACF)	9
2.3.2 Fungsi Autokorelasi Parsial (PACF)	11
2.4 Model ARIMA	13
2.4.1 Identifikasi Model ARIMA	16
2.4.2 Estimasi Parameter Model	17
2.4.3 Diagnosis Model	20
2.4.4 Kriteria Pemilihan Model	22
2.4.5 Peramalan Model ARIMA	24

III PEMBAHASAN	26
3.1 Dekomposisi Wold	26
3.2 Analisis Intervensi	27
3.2.1 Fungsi Intervensi	29
3.2.2 Variabel Intervensi	32
3.2.3 Orde Intervensi	33
3.3 Pemodelan Intervensi	35
3.3.1 Pengelompokan Data Deret Waktu	35
3.3.2 Pemodelan ARIMA Untuk Data Sebelum Intervensi	35
3.3.3 Identifikasi Orde Intervensi	36
3.3.4 Estimasi Parameter Intervensi	36
3.3.5 Diagnosis Model Intervensi	38
3.4 Peramalan Menggunakan Model Intervensi	38
3.5 Aplikasi Penggunaan Analisis Intervensi	40
3.5.1 Data Pendapatan Domestik Bruto	40
3.5.2 Pemodelan Data PDB Sebelum Intervensi (N_t)	41
3.5.3 Identifikasi Orde Intervensi	45
3.5.4 Estimasi Parameter Intervensi	47
3.5.5 Diagnosis Model Intervensi	48
3.5.6 Peramalan PDB Menggunakan Model Intervensi	48
IV PENUTUP	50
4.1 Kesimpulan	50
4.2 Saran	51
DAFTAR PUSTAKA	51
LAMPIRAN	55
A Flowchart Prosedur Pemodelan ARIMA	56
B Flowchart Prosedur Pemodelan Intervensi	58
C Ilustrasi Pola Pembobot Impuls untuk $r = 0, 1, 2$	61
D R Script	62

DAFTAR GAMBAR

2.1	Pola Gerak Jangka Panjang	4
2.2	Pola Siklis	5
2.3	Pola Musiman	6
2.4	Pola Gerak Tak Teratur	6
3.1	Pola Intervensi Langsung Permanen	30
3.2	Pola Intervensi Langsung Sementara	31
3.3	Pola Intervensi Gradual Permanen	31
3.4	Pola Intervensi Gradual Sementara	32
3.5	Tipe Dari Komponen Intervensi. (McDowall, 1980)	34
3.6	Plot Data PDB dari tahun 1930 sampai 2013	41
3.7	Plot Data PDB dari tahun 1930 sampai 1997 dengan nilai ACF dan PACF yang bersesuaian.	42
3.8	Selang kepercayaan 95% untuk λ	42
3.9	Plot W_t dengan nilai ACF dan PACF yang bersesuaian	43
3.10	Plot perbandingan antara Z_t dengan hasil peramalan N_t	46
3.11	Plot pola intervensi	46
3.12	Plot Hasil Permalan PDB	49

DAFTAR TABEL

2.1	Tabel Identifikasi Model	17
3.1	Nilai AIC, AICc, dan BIC untuk berbagai model ARIMA	44
3.2	Estimasi Parameter Intervensi	47

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Data deret waktu sering sekali dipengaruhi oleh beberapa kejadian eksternal seperti bencana alam, kebijakan pemerintah, hari libur nasional, promosi penjualan, dan lain-lain. Kejadian-kejadian eksternal ini dalam analisis data deret waktu disebut sebagai intervensi. Pada analisis data deret waktu, pemodelan data yang dilakukan tanpa memperhatikan dampak dari intervensi akan menghasilkan nilai kesalahan model (*error*) yang besar. Nilai *error* yang semakin besar mengakibatkan model yang diperoleh menjadi tidak akurat dan tentunya model tersebut kurang sesuai untuk menggambarkan data yang diamati.

George Box dan Gwilyn Jenkins (1970) mengembangkan suatu metode untuk memodelkan data deret waktu atau yang biasa dikenal dengan metode *Autoregressive Integrated Moving Average* (ARIMA). Metode ARIMA menduga suatu nilai deret waktu dengan menggunakan kombinasi linear dari nilai *error* dan nilai amatan sebelumnya. Metode ini mempunyai beberapa pendekatan untuk mengurangi nilai kesalahan dari model seperti asumsi kestasioneran, pembedaan (*differencing*), serta transformasi data. Meskipun begitu, pendekatan ini masih tidak mampu untuk mengurangi nilai kesalahan model yang disebabkan oleh suatu intervensi.

Dampak dari intervensi dalam suatu data deret waktu dapat dianalisis dengan menggunakan suatu metode yang disebut analisis intervensi. Analisis intervensi telah digunakan dalam mempelajari dampak dari kontrol polusi udara terhadap kebijakan ekonomi pada kota Los Angeles dari tahun 1955 sampai 1972 (Box

dan Tiao, 1975), dampak embargo minyak Arab terhadap meningkatnya harga produk perminyakan Amerika Serikat pada tahun 1973 (Montgomery dan Weatherby, 1980), dampak *Blackout* di kota New York pada 9 November 1965 terhadap kenaikan angka kelahiran kota New York sembilan bulan setelah kejadian *Blackout* tersebut (Izenman dan Zabell, 1981), dan masih banyak kejadian lainnya.

Secara umum, analisis intervensi dibedakan menjadi dua macam, yaitu analisis intervensi fungsi *step* dan analisis intervensi fungsi *pulse*. Kebijakan pemerintah, pergantian pemerintah, dan perang adalah contoh intervensi yang bersifat jangka panjang sehingga dalam kasus seperti ini digunakan analisis intervensi fungsi *step*. Untuk intervensi yang bersifat sementara atau jangka pendek seperti hari libur keagamaan atau potongan harga dari pusat perbelanjaan digunakan analisis intervensi fungsi *pulse*. Berdasarkan waktu intervensi, analisis intervensi dibagi menjadi dua yaitu, saat waktu intervensi tidak diketahui yang mengarah kepada analisis pencilaan data deret waktu dan saat waktu intervensi diketahui yang berfokus untuk mengukur pengaruh dari intervensi (William Wei, 2006).

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas, masalah yang akan dibahas dalam skripsi ini adalah:

1. Bagaimana memodelkan data deret waktu yang mengandung intervensi?
2. Bagaimana meramalkan data deret waktu yang mengandung intervensi?

1.3 Pembatasan Masalah

Pada penelitian ini pemodelan dan peramalan deret waktu hanya dilakukan pada data yang waktu terjadinya intervensi diketahui dan berbasis pada metode

ARIMA tanpa adanya pengaruh musiman (*seasonal*).

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penulisan ini adalah untuk mendapatkan model dan hasil peramalan dari suatu data deret waktu yang mengandung intervensi saat waktu terjadinya intervensi diketahui.

1.5 Manfaat Penulisan

Manfaat yang diharapkan dari penulisan ini adalah dapat membantu para pengguna dalam melakukan pemodelan dan peramalan data deret waktu yang lebih akurat dengan menggunakan analisis intervensi saat waktu intervensi diketahui.

1.6 Metode Penelitian

Skripsi ini merupakan kajian teori dalam bidang pemodelan menggunakan analisis intervensi data deret waktu yang didasarkan pada buku-buku dan jurnal-jurnal tentang pemodelan data deret waktu.

BAB II

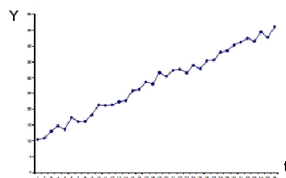
LANDASAN TEORI

2.1 Data Deret Waktu

Data deret waktu adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan perkembangan suatu kegiatan atau sekumpulan hasil observasi yang diatur dan didapat menurut urutan kronologis waktu. Ada empat komponen gerak/variasi data deret waktu, yaitu:

1. Gerak Jangka Panjang dan Trend

Trend melukiskan gerak data deret waktu selama jangka waktu yang cukup lama. Gerak ini mencerminkan sifat kontinuitas atau keadaan yang terus-menerus dari waktu ke waktu selama jangka waktu tertentu. Dengan sifat kontinuitas ini, maka trend dianggap sebagai gerak stabil dan menunjukkan arah perkembangan secara umum (kecenderungan naik/turun). Trend sangat berguna untuk membuat peramalan (*forecasting*) yang merupakan perkiraan untuk masa depan yang diperlukan bagi perencanaan.



Gambar 2.1: Pola Gerak Jangka Panjang

Trend dibedakan menjadi dua jenis, yakni :

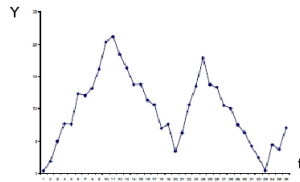
- a. Trend Linear: mengikuti pola garis lurus ($Y = a + bt$)

- b. Trend Non Linear: mengikuti pola lengkung (parabola, eksponensial, logaritma, dan lain-lain).

Pada Gambar 2.1, plot data menunjukkan pola data deret waktu yang mengikuti pola trend linear naik.

2. Gerak Siklis

Gerak siklis adalah gerak/variasi jangka panjang di sekitar garis trend (tempunya lebih pendek). Gerak siklis terjadi berulang-ulang namun tidak perlu periodik, artinya bisa berulang setelah jangka waktu tertentu atau bisa juga tidak berulang dalam jangka waktu yang sama. Gerak siklis melukiskan terjadinya empat fase kejadian dalam jangka waktu tertentu, yakni kemajuan, kemunduran, depresi dan pemulihan. Contoh plot data deret waktu yang mengikuti pola siklis dapat dilihat pada Gambar 2.2.



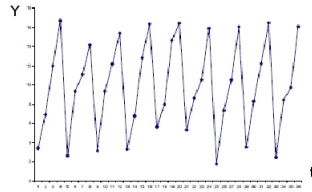
Gambar 2.2: Pola Siklis

3. Gerak Musiman

Gerak musiman terjadi lebih teratur dibandingkan gerak siklis dan bersifat lengkap, biasanya selama satu tahun kalender. Gerak ini berpola tetap dari waktu ke waktu. Faktor utama yang menyebabkan gerak ini adalah iklim dan kebiasaan. Contoh plot data deret waktu yang mengikuti pola musiman dapat dilihat pada Gambar 2.3.

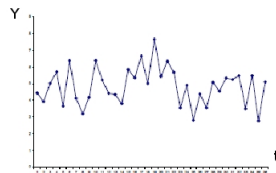
4. Gerak Ireguler atau Faktor Residu (Gerak Tak Teratur)

Gerak ini bersifat tidak teratur. Perang, bencana alam, dan kekacauan adalah beberapa faktor yang dikenal dapat menyebabkan gerak ini terjadi. Dengan



Gambar 2.3: Pola Musiman

adanya pengaruh tersebut, maka gerak ireguler sulit untuk dilukiskan dalam suatu model. Contoh plot data deret waktu yang mengikuti pola ireguler dapat dilihat pada Gambar 2.4.



Gambar 2.4: Pola Gerak Tak Teratur

2.2 Kestasioneran Data Deret Waktu

Kestasioneran merupakan asumsi yang diperlukan dalam analisis data deret waktu, karena dengan asumsi kestasioneran ini kesalahan pemodelan dapat diminimalisir. Data deret waktu yang memenuhi asumsi kestasioneran mempunyai rata-rata dan ragam yang konstan dengan kovarians dan korelasi yang tergantung hanya pada selisih waktu (Wei, 2006).

Definisi 2.2.1. Suatu data deret waktu Z_t dikatakan (*weakly*) stasioner jika untuk setiap $t \in \mathbb{Z}$:

- (i) $E(Z_t) = \mu(t) = \mu$,
- (ii) $Var(Z_t) = \sigma^2 = \gamma_0$,
- (iii) $\gamma_k = E[(Z_t - \mu)(Z_{t+k} - \mu)]$,

dengan γ_k adalah autokovarians pada lag- k , nilai μ dan γ_k untuk setiap k adalah konstan.

Bentuk ketakstasioneran yang mungkin terjadi pada data deret waktu meliputi ketakstasioneran nilai rata-rata dan ketakstasioneran ragam. Apabila suatu data deret waktu mempunyai ragam yang tidak stasioner, maka perlu dilakukan transformasi terhadap data tersebut untuk membuat ragam data menjadi stasioner. Transformasi yang banyak digunakan adalah transformasi Box-Cox. Sementara itu untuk data deret waktu yang mempunyai nilai rata-rata tidak stasioner, maka dapat digunakan proses pembedaan (*differencing*).

2.2.1 Pembedaan (*differencing*)

Suatu data deret waktu yang stasioner dalam ragam, namun tak stasioner dalam rata-rata, dapat diupayakan menjadi suatu deret yang stasioner dengan memilih suatu pembeda (*difference*) yang bersesuaian dengan data deret waktu tersebut (Wei, 2006).

Definisi 2.2.2. Pembedaan atau *differencing* terhadap suatu data deret waktu Z_t dilakukan dengan menggunakan operator ∇ yang didefinisikan

$$\begin{aligned}\nabla Z_t &= Z_t - Z_{t-1} = (1 - B)Z_t \\ \nabla^2 Z_t &= \nabla Z_t - \nabla Z_{t-1} = (1 - B)^2 Z_t \\ &\vdots \\ \nabla^d Z_t &= \nabla^{d-1} Z_t - \nabla^{d-1} Z_{t-1} = (1 - B)^d Z_t,\end{aligned}$$

dengan B adalah *Back-shift operator*. Dengan kata lain, jika suatu data deret waktu Z_t tidak stasioner, maka pembedaan ke- d , $(1 - B)^d Z_t$, untuk suatu $1 \leq d < n$ adalah stasioner.

2.2.2 Transformasi Box-Cox

Pada deret waktu yang tidak stasioner dalam ragam, dapat diupayakan menjadi stasioner melalui transformasi (Wei, 2006). Box dan Cox pada tahun 1964 memperkenalkan suatu transformasi bentuk pangkat:

$$T(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & , \lambda \neq 0 \\ \ln Z_t & , \lambda = 0 \end{cases} \quad (2.1)$$

dengan λ adalah parameter transformasi. Transformasi Box-Cox hanya dapat digunakan jika observasi Z_t bernilai lebih dari 0. Jika terdapat $Z_t \leq 0$, maka suatu konstanta c dapat ditambahkan sedemikian sehingga $Z_t + c > 0, \forall t$ dan transformasi Box-Cox dapat digunakan terhadap $Z_t + c$. Hal ini dimungkinkan karena sebuah konstanta selalu dapat ditambahkan pada data tanpa mempengaruhi nilai korelasi data tersebut.

Teorema 2.2.1. *Jika X dan Y adalah sebuah peubah acak dengan nilai ekspektasi μ_X dan μ_Y , simpangan baku σ_X dan σ_Y , dan c adalah konstan, maka*

$$\text{Corr}(X + c, Y + c) = \text{Corr}(X, Y). \quad (2.2)$$

Bukti.

$$\begin{aligned} \text{Corr}(X + c, Y + c) &= \frac{\text{Cov}(X + c, Y + c)}{\sqrt{\text{Var}(X + c) \cdot \text{Var}(Y + c)}} \\ &= \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} \\ &= \text{Corr}(X, Y). \end{aligned}$$

□

Bentuk transformasi Box-Cox yang sudah dimodifikasi jika terdapat nilai

$Z_t \leq 0$:

$$T(Z_t) = \begin{cases} \frac{(Z_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1}, & \lambda_1 \neq 0 \\ \ln(Z_t + \lambda_2) & \lambda_1 = 0 \end{cases} \quad (2.3)$$

dengan λ_1 adalah parameter transformasi dan λ_2 adalah nilai konstan yang dipilih sedemikian sehingga $Z_t + \lambda_2 > 0$.

2.3 Fungsi Autokorelasi dan Autokorelasi Parsial

2.3.1 Fungsi Autokorelasi (ACF)

Untuk Z_t yang stasioner, $E(Z_t) = \mu$ dan $Var(Z_t) = \sigma^2$ adalah konstan dan $Cov(Z_t, Z_s)$ adalah fungsi dari selisih waktu $|t - s|$. Kovarians dan korelasi antara Z_t dan Z_{t+k} berturut-turut adalah

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu, Z_{t+k} - \mu)$$

dan

$$\rho_k = Corr(Z_t, Z_{t+k}) = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}.$$

Sebagai fungsi dari k , γ_k dinamakan fungsi autokovarians dan ρ_k dinamakan fungsi autokorelasi (*autocorrelation function*), disingkat dengan ACF. Beberapa sifat dari autokovarians γ_k dan autokorelasi ρ_k :

1. $\gamma_0 = Var(Z_t)$ dan $\rho_0 = 1$
2. $|\gamma_k| \leq \gamma_0$ dan $|\rho_k| \leq 1$
3. $\gamma_k = \gamma_{-k}$ dan $\rho_k = \rho_{-k}$.

Untuk menghitung ACF pada sampel digunakan formula berikut:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}, \quad (2.4)$$

dengan ρ_k adalah nilai korelasi sampel pada lag- k dan n adalah banyaknya pengamatan.

Untuk mempermudah dalam melihat apakah suatu ACF signifikan atau tidak, dapat dibuat plot dari ACF dan disandingkan dengan galat bakunya (*standard error*). Plot dari ACF ini dinamakan *correlogram*. Untuk dapat membuat suatu *correlogram*, perlu dihitung terlebih dahulu nilai galat baku dari ACF tersebut. Galat baku ini digunakan untuk melihat apakah ACF berbeda secara nyata dengan nol. Formula untuk menghitung galat baku sebagai berikut:

$$SE_{(\hat{\rho}_k)} = \sqrt{\frac{1 + 2\sum_{i=1}^{k-1} \hat{\rho}_i^2}{n}}, \quad (2.5)$$

dengan $SE_{(\hat{\rho}_k)}$ adalah nilai galat baku dari $\hat{\rho}_k$ dan $\hat{\rho}_i^2$ adalah nilai autokorelasi sampel pada lag i , dan k adalah selisih waktu.

Pada uji korelasi, H_0 didefinisikan dengan $\rho_k = 0$ yaitu tidak ada autokorelasi, sedangkan H_1 adalah $\rho_k \neq 0$ yaitu ada autokorelasi antar observasi. Statistik uji yang digunakan dalam uji autokorelasi adalah statistik t yang dirumuskan sebagai berikut:

$$t_{hit} = \frac{\hat{\rho}_k}{SE_{(\hat{\rho}_k)}} \quad (2.6)$$

dengan $df = n - k$. Daerah penolakan yang digunakan adalah H_0 ditolak jika $|t_{hit}| > t_{\frac{\alpha}{2}, df}$. Selain menggunakan statistik t , plot ACF dapat digunakan untuk melihat ada atau tidaknya autokorelasi antar observasi. Apabila tidak terdapat lag yang keluar dari batas signifikansi, maka dapat disimpulkan tidak ada korelasi antar lag.

2.3.2 Fungsi Autokorelasi Parsial (PACF)

Autokorelasi parsial atau *partial autocorrelation* (PACF) antara Z_t dan Z_{t+k} adalah korelasi antara Z_t dan Z_{t+k} setelah ketergantungan linearnya dengan $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$ dihilangkan. Autokorelasi parsial antara Z_t dan Z_{t+k} , dinotasikan dengan ϕ_{kk} , dirumuskan sebagai berikut:

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\det(P_2^*)}{\det(P_2)} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\det(P_3^*)}{\det(P_3)} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

⋮

$$\phi_{kk} = \frac{\det(P_k^*)}{\det(P_k)} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-4} & \rho_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \dots & 1 & \rho_{k-1} \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{vmatrix}}, \quad (2.7)$$

dengan P_k adalah matrik autokorelasi $k \times k$ dan $\det(P_k) \neq 0$ dan ρ_k adalah autokorelasi pada lag k . Matriks P_k^* adalah matriks P_k dengan kolom terakhir disubstitusi dengan transpose dari $(\rho_1 \ \rho_2 \ \dots \ \rho_k)$ dan $\det(P_k^*) \neq 0$.

Durbin (1960) mengemukakan sebuah alternatif cara untuk menghitung autokorelasi parsial tanpa perlu sebelumnya menentukan nilai $\det(P_k)$ dan $\det(P_k^*)$, yaitu dengan metode rekursif sebagai berikut:

$$\phi_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^k \phi_{kj} \rho_{k+1-j}}{1 - \sum_{j=1}^k \phi_{kj} \rho_j} \quad (2.8)$$

dan

$$\phi_{k+1,j} = \phi_{kj} - \phi_{k+1,k+1} \phi_{k,k+1-j}, \quad j = 1, 2, \dots, k. \quad (2.9)$$

Quenouille (1949) telah menunjukkan bahwa untuk suatu proses *white noise*, nilai ragam dari ϕ_{kk} dapat diaproksimasi dengan:

$$\text{Var}(\phi_{kk}) \approx \frac{1}{n} \quad (2.10)$$

Dengan demikian, $\pm \frac{2}{\sqrt{n}}$ dapat digunakan sebagai batas kritis ϕ_{kk} untuk menguji hipotesis dari suatu proses *white noise*.

2.4 Model ARIMA

Metode *Autoregressive Integrated Moving Average* (ARIMA) atau biasa disebut juga sebagai metode Box-Jenkins merupakan metode yang secara intensif dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970. Metode Box Jenkins merupakan model yang menggambarkan data deret waktu yang stasioner. Sementara itu identifikasi kestasioneran pada rata-rata dan ragam data yang digunakan harus dilakukan sebelum menentukan model ARIMA yang bersesuaian. Untuk data deret waktu yang tidak stasioner dalam rata-rata maka perlu dilakukan *differencing* terlebih dahulu untuk menstasionerkan rata-rata tersebut. Untuk data deret waktu yang tidak stasioner dalam ragam, maka data tersebut perlu ditransformasi agar memperoleh ragam yang stasioner.

Model Autoregresif (*Autoregressive*, AR)

Menurut Halim (2006), model Autoregresif (AR) pertama kali diperkenalkan oleh Yule pada tahun 1926 dan dikembangkan oleh Walker pada tahun 1931, model ini memiliki asumsi bahwa data periode sekarang dipengaruhi oleh data pada periode sebelumnya.

Definisi 2.4.1. Suatu deret waktu Z_t dikatakan *invertible* jika terdapat

$$\phi_p(B) = \phi_0 + \phi_1 B + \phi_2 B^2 + \dots \quad (2.11)$$

dengan $\sum_{j=0}^{\infty} |\phi_j| < \infty$ dan $\phi_p(B)Z_t = a_t$.

Model Autoregresif dengan ordo p disingkat AR(p) atau ARIMA ($p,0,0$)

dan diformulasikan sebagai berikut:

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \quad (2.12)$$

dengan μ adalah konstanta, $\phi_1, \phi_2, \dots, \phi_p$ adalah koefisien parameter autoregresif, dan a_t adalah residual pada saat ke- t dan $a_t \sim WN(0, \sigma^2)$.

Model Rataan Bergerak (*Moving Average, MA*)

Proses Moving Average berorde q menyatakan hubungan ketergantungan antara nilai pengamatan dengan nilai-nilai kesalahan yang berurutan dari periode t sampai $t - q$.

Definisi 2.4.2. Suatu deret waktu Z_t dikatakan *causal* jika terdapat

$$\theta_q(B) = \theta_0 + \theta_1 B + \theta_2 B^2 + \dots \quad (2.13)$$

dengan $\sum_{j=0}^{\infty} |\theta_j| < \infty$ dan $Z_t = \theta_q(B) a_t$.

Model Moving Average (MA) pertama kali diperkenalkan oleh Slutsky pada tahun 1973, dengan orde q ditulis MA (q) atau ARIMA ($0, 0, q$). Dikembangkan oleh Wadsworth pada tahun 1989 yang memiliki formulasi sebagai berikut (Halim, 2006):

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (2.14)$$

dengan $\theta_1, \theta_2, \dots, \theta_q$ adalah koefisien parameter *moving average*.

Model ARMA (*Autoregressive Moving Average*)

Model AR (p) dan MA (q) dapat disatukan menjadi model yang dikenal dengan Autoregressive Moving Average (ARMA), sehingga memiliki asumsi bahwa data periode sekarang dipengaruhi oleh data pada periode sebelumnya dan nilai

residual pada periode sebelumnya (Assauri, 1984). Model ARMA dengan berorde p dan q ditulis ARMA (p, q) atau ARIMA $(p, 0, q)$ yang memiliki formulasi sebagai berikut (Halim, 2006) :

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.15)$$

atau

$$\phi_p(B)Z_t = \mu + \theta_q(B)a_t, \quad (2.16)$$

dengan ϕ_p adalah parameter autoregresif dan θ_q adalah parameter rata-rata bergerak (*moving average*).

Teorema 2.4.1. *Jika ϕ dan θ tidak memiliki faktor bersama, maka terdapat suatu solusi stasioner yang unik untuk $\phi(B)Z_t = \theta_q(B)a_t$ jika dan hanya jika akar dari $\phi(z)$ tidak berada pada unit circle:*

$$|z| = 1 \rightarrow \phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0. \quad (2.17)$$

ARMA(p,q) dikatakan causal jika dan hanya jika akar dari $\phi(z)$ berada di luar unit circle:

$$|z| \leq 1 \rightarrow \phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0 \quad (2.18)$$

dan suatu ARMA(p,q) dikatakan invertible jika dan hanya jika akar dari $\theta(z)$ berada di luar unit circle:

$$|z| \leq 1 \rightarrow \theta(z) = 1 + \theta_1 z + \dots + \theta_p z^p \neq 0. \quad (2.19)$$

Model ARIMA (*Autoregressive Integrated Moving Average*)

Model AR, MA dan ARMA menggunakan asumsi bahwa data deret waktu yang dihasilkan sudah bersifat stasioner. Pada kenyataannya, data deret waktu lebih banyak bersifat tidak stasioner (Sadeq, 2008). Jika data tidak stasioner maka metode yang digunakan untuk membuat data stasioner dilakukan adalah *differencing* untuk data yang tidak stasioner dalam rata-rata dan proses transformasi untuk data yang tidak stasioner dalam ragam (Mulyana, 2004). Bentuk umum model ARIMA dapat dinyatakan dalam persamaan berikut (Sartono, 2006):

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t, \quad (2.20)$$

dengan d adalah parameter pembeda (*difference*) dan θ_0 adalah parameter konstan.

2.4.1 Identifikasi Model ARIMA

Langkah pertama dalam pembentukan model ARIMA adalah membuat plot data deret waktu. Pada plot tersebut dapat dilihat apakah pola suatu data deret waktu mengikuti pola horizontal, trend, siklis, ataupun pola musiman. Pembuatan plot data deret waktu bertujuan pula untuk menyelidiki stasioneritas data deret waktu. Stasioneritas data deret waktu adalah hal pertama yang harus diperhatikan karena aspek-aspek AR dan MA dari model ARIMA hanya berkenaan dengan data deret waktu yang stasioner dalam ragam dan rata-rata.

Data yang belum stasioner dalam ragam harus terlebih dahulu dilakukan transformasi Box-Cox agar data tersebut menjadi stasioner dalam ragam. Apabila data sudah stasioner dalam ragam namun belum stasioner dalam rata-rata maka dapat dilakukan pembedaan pada lag 1, lag 2, dan seterusnya sampai data stasioner. Secara umum model dari data deret waktu dapat dinyatakan sebagai model *Autoregressive Integrated Moving Average* atau ARIMA (p, d, q) yang stasioner dengan:

Proses	ACF	PACF
AR(p)	Menurun secara eksponensial atau mengikuti gelombang sinus	Terpotong setelah lag p
MA(q)	Terpotong setelah lag q	Menurun secara eksponensial atau mengikuti gelombang sinus
ARMA(p, q)	Terpotong setelah lag $(q - p)$	Terpotong setelah lag $(p - q)$

Tabel 2.1: Tabel Identifikasi Model

2.4.2 Estimasi Parameter Model

Hukum Bilangan Besar (Law of Large Number)

Jika distribusi dari sampel yang i.i.d. X_1, X_2, \dots, X_n sedemikian sehingga X_1 mempunyai ekspektasi yang terhingga yaitu $|E(X_1)| < \infty$, maka rata-rata sampel

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow E(X_1) \quad (2.21)$$

konvergen menuju ekspektasinya, yang berarti bahwa untuk sembarang $\varepsilon > 0$,

$$P[|\bar{X}_n - E(X_1)|] \rightarrow 0 \quad \text{untuk} \quad n \rightarrow \infty. \quad (2.22)$$

Teorema Nilai Tengah (Central Limit Theorem)

Teorema 2.4.2. *Jika distribusi dari sampel yang i.i.d. X_1, X_2, \dots, X_n sedemikian sehingga X_1 mempunyai ekspektasi dan ragam yang terhingga yaitu $|E(X_1)| < \infty$ dan $\sigma^2 = \text{Var}(X_1) < \infty$, maka*

$$\sqrt{n}(\bar{X}_n - E(X_1)) \rightarrow^d N(0, \sigma^2) \quad (2.23)$$

konvergen dalam distribusi menuju distribusi normal dengan rata-rata 0 dan ragam σ^2 , yang berarti untuk sembarang interval $[a, b]$,

$$P(\sqrt{n}[\bar{X}_n - E(X_1)] \in [a, b]) \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx. \quad (2.24)$$

Dengan kata lain, teorema nilai tengah menyatakan bahwa suatu peubah acak $\sqrt{n}[\bar{X}_n - E(X_1)]$ akan berperilaku seperti peubah acak normal saat n semakin besar.

Maximum Likelihood Estimation (MLE)

Fungsi kepadatan peluang atau *probability density function* (pdf) untuk peubah acak x yang dikondisikan untuk suatu parameter θ dinotasikan $f(x|\theta)$. Fungsi kepadatan peluang bersama dari suatu n observasi yang i.i.d. adalah hasil perkalian dari setiap kepadatan individual:

$$f(x_1, x_2, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta) = L(\theta | x) \quad (2.25)$$

atau $L(\theta | x)$ disebut juga sebagai fungsi likelihood, didefinisikan sebagai suatu fungsi dari vektor parameter yang tidak diketahui, θ , dimana x mengindikasikan koleksi dari data sampel. Kemudian, MLE $\hat{\theta}$ didefinisikan sebagai *maximizer* dari

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^n \log f(x_i | \theta). \quad (2.26)$$

Beberapa sifat dari MLE:

1. Penduga $\hat{\theta}$ dari θ adalah penduga yang *tak bias* jika $E(\hat{\theta}) = \theta$.
2. Penduga $\hat{\theta}$ dari θ adalah penduga yang *konsisten* jika $P(|\hat{\theta} - \theta| > \varepsilon) \rightarrow 0$ untuk $n \rightarrow \infty$.
3. Penduga $\hat{\theta}$ dari θ adalah penduga yang *asimtotik normal* jika $\sqrt{n}(\hat{\theta} - \theta) \rightarrow^d$

$N(0, \sigma_{\hat{\theta}}^2)$, dengan $\sigma_{\hat{\theta}}^2$ disebut sebagai ragam asimtotik dari dugaan $\hat{\theta}$.

4. Misal $\hat{\theta}$ dan θ^* adalah penduga tak bias dari θ maka $\hat{\theta}$ dikatakan sebagai penduga yang *lebih efektif* daripada θ^* jika $Var(\hat{\theta}) \leq Var(\theta^*)$ untuk semua θ .

Conditional MLE Pada Model ARIMA

Untuk model umum ARMA(p, q) yang stasioner, fungsi kepadatan peluang bersama dari $a = (a_1, a_2, \dots, a_n)'$ dinyatakan

$$P(a|\phi, \theta, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} \exp \left[-\frac{1}{2\sigma_a^2} \sum_{t=1}^n a_t^2 \right]. \quad (2.27)$$

Persamaan (2.15) dapat dituliskan menjadi

$$a_t = \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p}, \quad (2.28)$$

kemudian misal $Z = (Z_1, Z_2, \dots, Z_n)'$ selanjutnya asumsikan kondisi atau syarat $Z_* = (Z_{1-p}, \dots, Z_{-1}, Z_0)'$ dan $a_* = (a_{1-p}, \dots, a_{-1}, a_0)'$ sudah diketahui. Fungsi log-likelihood bersyarat dituliskan

$$\log L_*(\phi, \theta, \sigma_a^2) = -\frac{n}{2} \log 2\pi\sigma_a^2 - \frac{S_*(\phi, \theta)}{2\sigma_a^2}, \quad (2.29)$$

dengan

$$S_*(\phi, \theta) = \sum_{t=1}^n a_t^2(\phi, \theta | Z, a_*, Z_*) \quad (2.30)$$

adalah fungsi jumlah kuadrat bersyarat. Nilai dari $\hat{\phi}$ dan $\hat{\theta}$ yang memaksimalkan fungsi pada Persamaan (2.29) disebut sebagai penduga likelihood maksimum bersyarat (*conditional maximum likelihood estimator*). Karena $\log L_*(\phi, \theta, \sigma_a^2)$ hanya melibatkan data melalui $S_*(\phi, \theta)$, penduga ini dapat dianggap sebagai penduga ku-

adrat terkecil bersyarat (*conditional least square estimator*) yang didapat dengan meminimalkan jumlah kuadrat bersyarat $S_*(\phi, \theta)$. Hal yang perlu diperhatikan adalah penduga ini tidak mengandung parameter σ_a^2 . Berdasarkan bahwa Z_t adalah proses yang stasioner dan a_t menyebar i.i.d. $N(0, \sigma_a^2)$ maka untuk nilai Z_t yang tidak diketahui dapat diganti dengan rataan sampel \bar{Z} , dengan \bar{Z} adalah penduga tak bias untuk μ dan mengganti nilai a_t dengan nilai ekpektasinya yaitu 0. Fungsi jumlah kuadrat bersyarat pada Persamaan (2.30) menjadi

$$S_*(\phi, \theta) = \sum_{t=1}^n a_t^2(\phi, \theta|Z). \quad (2.31)$$

Setelah mendapatkan estimasi dari parameter $\hat{\phi}$ dan $\hat{\theta}$, estimasi untuk $\hat{\sigma}_a^2$ dari σ_a^2 dapat dihitung dengan

$$\hat{\sigma}_a^2 = \frac{S(\hat{\phi}, \hat{\theta})}{d.f.} = \frac{\sum_{t=1}^n (a_t)^2}{n - (2p + q + 1)}. \quad (2.32)$$

2.4.3 Diagnosis Model

Setelah melakukan estimasi parameter, langkah selanjutnya adalah mengecek model yang telah didapat apakah sudah cukup memadai untuk data yang ada. Untuk melihat apakah model sudah memadai atau belum, dapat diperiksa melalui residual a_t yang dihasilkan oleh model. Agar model ini dapat dikatakan memadai, residual yang dihasilkan harus memenuhi asumsi *white noise* dan normalitas.

White Noise

Definisi 2.4.3. Suatu model dikatakan memadai jika residual a_t yang dihasilkan oleh model tersebut memenuhi asumsi *white noise*, yaitu:

- (i) a_t adalah residual yang *independent*.
- (ii) Mempunyai distribusi tertentu dengan rataan dan ragam yang konstan, $E(a_t) =$

$\mu_{a_t}, \text{Var}(a_t) = \sigma_{a_t}^2$ dan $\gamma_k = \text{Cov}(a_t, a_{t+k}) = 0$ untuk $k \neq 0$ (Box dan Jenkins, 1976).

Rumusan hipotesis yang digunakan:

$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0$ (nilai residual independen)

$H_1 : \text{ada satu } \rho_i \neq 0 \text{ untuk } i = 1, 2, 3, \dots, k$ (nilai residual dependen).

Untuk melihat independensi a_t , perlu dilakukan pengujian terhadap nilai koefisien autokorelasi dari a_t dengan menggunakan statistik Uji Ljung-Box (Ljung dan Box, 1978):

$$Q = n'(n' + 2) \sum_{k=1}^m \frac{\gamma_k^2}{(n' - k)}. \quad (2.33)$$

Statistik Q menyebar Khi Kuadrat (χ^2) dengan derajat bebas (df)= $(k - p - q)$, dengan $n' = n - d$, $m = \min(10, n/5)$, d adalah orde perbedaan bukan faktor musiman, dan $\gamma_k =$ autokorelasi untuk time lag 1, 2, 3, 4, ..., k . Kriteria pengujian:

- Jika $Q \leq \chi_{(\alpha, db)}^2$, maka a_t independen (model diterima).
- Jika $Q > \chi_{(\alpha, db)}^2$, maka a_t dependen (model tidak diterima).

Normalitas

Uji kenormalan residual digunakan untuk memeriksa apakah suatu residual a_t mempunyai distribusi normal atau tidak. Rumusan hipotesis yang digunakan adalah:

$H_0 : \text{residual } a_t \text{ berdistribusi normal}$

$H_1 : \text{residual } a_t \text{ tidak berdistribusi normal.}$

Uji Kolmogorov-Smirnov digunakan untuk membandingkan kesesuaian dari distribusi sampel dengan suatu distribusi pembanding. Uji Kolmogorov-Smirnov dapat digunakan sebagai uji kenormalan jika distribusi pembanding yang diambil ada-

lah distribusi normal. Statistik uji yang digunakan adalah statistik uji Kolmogorov-Smirnov dengan formula sebagai berikut:

$$D = KS = \max|F_0(X) - S_n(X)| \quad (2.34)$$

dengan $F_0(X)$ adalah fungsi distribusi kumulatif pembanding dan $S_n(X)$ adalah fungsi distribusi kumulatif observasi. Kriteria pengujian:

- Jika $D_{hit} < D_{tabel}$, maka distribusi sampel sesuai dengan distribusi pembanding.
- Jika $D_{hit} \geq D_{tabel}$, maka distribusi sampel tidak sesuai dengan distribusi pembanding.

2.4.4 Kriteria Pemilihan Model

Prinsip Parsimony

Prinsip *parsimony* mempunyai arti bahwa model yang dipilih adalah model tersederhana. Begitu pula dalam pemodelan data deret waktu, dapat dimungkinkan bahwa ada lebih dari satu model yang dapat menggambarkan data tersebut. Jika terdapat lebih dari satu kemungkinan model untuk memodelkan suatu data deret waktu, prinsip *parsimony* dapat dilakukan dengan memilih nilai orde $AR(p)$ dan $MA(q)$ sesederhana mungkin. Misal, setelah identifikasi model didapatkan model $ARIMA(2,1,0)$, $ARIMA(3,1,1)$ dan $ARIMA(0,1,2)$, dengan prinsip *parsimony* maka sebaiknya model yang dipilih adalah $ARIMA(2,1,0)$.

Akaike's Information Criterion (AIC) dan Bayesian Information Criterion (BIC)

Salah satu kriteria untuk memilih model ARIMA terbaik dari beberapa model yang mungkin adalah dengan memilih model yang mempunyai nilai *Mean*

Square Error (MSE) terendah. Ada suatu masalah vital dalam menggunakan kriteria pemilihan tersebut, karena nilai MSE dapat selalu berkurang apabila menambahkan parameter AR dan MA. Berarti model dapat mempunyai nilai MSE sangat rendah namun tingkat kerumitan dari model tersebut menjadi sangat tinggi. Untuk itu diperlukan suatu kriteria yang dapat menyeimbangkan antara pemilihan nilai MSE terendah dan banyaknya parameter dalam suatu model yang akan dipilih.

Kriteria lain yang dapat digunakan untuk memilih model ARIMA dari beberapa model yang mungkin adalah dengan melihat pada kriteria informasi masing-masing model. Ada dua kriteria yang dapat digunakan untuk pemilihan ini yaitu, *Akaike's Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC). Formula untuk AIC dan BIC adalah sebagai berikut:

$$AIC = -2\text{Log}(L) + 2m \quad (2.35)$$

$$BIC = -2\text{Log}(L) + m\text{Log}(n), \quad (2.36)$$

dengan L adalah likelihood dari data dengan model tertentu, m adalah banyaknya parameter dalam model atau $m = p + q$ untuk sebuah model ARIMA(p, d, q), dan n adalah banyaknya observasi.

Model yang akan dipilih adalah model yang mempunyai nilai AIC dan BIC terendah di antara semua model yang mungkin. Meskipun begitu terkadang masih terlalu sulit untuk menghitung likelihood dari model tertentu, sehingga digunakan aproksimasi:

$$-2\text{Log}(L) \approx n(1 + 2\text{Log}(2\pi)) + n\text{Log}(s^2), \quad (2.37)$$

dengan n adalah banyaknya observasi dan s^2 adalah estimasi variasi dari residual model. Kemudian berdasarkan pada prosedur pemodelan ARIMA di atas, dapat

dibuat *flowchart* prosedur pemodelan ARIMA (Lampiran A).

2.4.5 Peramalan Model ARIMA

Salah satu unsur terpenting dalam analisis data deret waktu adalah meramalkan nilai yang akan datang. Peramalan perlu dilakukan untuk memperbesar peluang memperoleh keuntungan dan memperkecil terjadinya kerugian. Untuk peramalan, misal pada $t = n$ terdapat $Z_n, Z_{n-1}, Z_{n-2}, \dots$ observasi dan akan meramalkan l langkah ke depan Z_{n+l} sebagai kombinasi linier dari $Z_n, Z_{n-1}, Z_{n-2}, \dots$ dengan penduga untuk Z_{n+l} adalah $\hat{Z}_n(l)$ yaitu:

$$\hat{Z}_n(l) = E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots). \quad (2.38)$$

Model ARIMA pada Persamaan (2.20) dapat dituliskan sebagai proses AR yaitu:

$$\pi(B)Z_{t+l} = \frac{\phi(B)(1-B)^d}{\theta(B)}Z_{t+l} = a_{t+l}$$

atau

$$Z_{t+l} = \sum_{j=1}^{\infty} \pi_j Z_{t+l-j} + a_{t+l}, \quad (2.39)$$

jadi penduga $\hat{Z}_n(l)$ adalah

$$\hat{Z}_n(l) = \sum_{j=1}^{\infty} \pi_j \hat{Z}_n(l-j). \quad (2.40)$$

Kemudian dengan dasar bahwa suatu proses ARIMA yang tak stationer dapat dinotasikan sebagai proses stationer ARMA, maka Persamaan (2.16) dapat dituliskan sebagai proses MA yaitu:

$$W_{n+l} = \sum_{j=0}^{\infty} \psi_j a_{n+l-j}, \quad (2.41)$$

dengan $W_{n+l} = (1 - B)^d Z_{n+l}$ untuk suatu $d \geq 1$. Dengan demikian penduga $\hat{W}_n(l)$ adalah

$$\begin{aligned}\hat{W}_n(l) &= E(W_{n+l} | W_n, W_{n-1}, W_{n-2}, \dots) \\ &= \psi_l a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + \dots\end{aligned}\quad (2.42)$$

Dengan hasil dari Persamaan (2.41) dan (2.42) didapat nilai *error* untuk peramalan sebagai berikut:

$$e_n(l) = W_{n+l} - \hat{W}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j}, \quad (2.43)$$

karena nilai $E(e_n(l) | Z_t, t \leq n) = 0$, maka hasil peramalan tak bias dengan ragam dari *error* adalah

$$\text{Var}(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2. \quad (2.44)$$

Sehingga untuk suatu proses yang normal, selang kepercayaan $(1 - \alpha)100\%$ untuk batas peramalan didapat

$$\hat{Z}_n(l) \pm N_{\alpha/2} \left[1 + \sum_{j=0}^{l-1} \psi_j^2 \right]^{1/2} \sigma_a, \quad (2.45)$$

dengan $N_{\alpha/2}$ adalah sebaran normal baku dengan $P(N > N_{\alpha/2}) = \alpha/2$.

BAB III

PEMBAHASAN

3.1 Dekomposisi Wold

Dekomposisi Wold atau Teorema Representasi Wold menyatakan bahwa setiap deret waktu x_t yang stasioner dalam kovarian dapat dituliskan sebagai penjumlahan dari deret waktu deterministik dan deret waktu stokastik (Cochrane, 1997).

Teorema 3.1.1. *Jika x_t adalah deret waktu yang stasioner dalam kovarian dengan $E(x_t) = 0$ dan fungsi kovarian, $\gamma(j) = E(x_t x_{t-j}) \forall j$, maka*

$$x_t = \sum_{j=0}^{\infty} d_j \varepsilon_{t-j} + \eta_t, \quad (3.1)$$

dengan $\varepsilon_t = x_t - P(x_t | x_{t-1}, x_{t-2}, \dots)$. Operator $P(x_t | x_{t-1}, x_{t-2}, \dots)$ adalah *projection operator* yaitu nilai proyeksi x_t berdasarkan regresi linear dari $x_{t-1}, x_{t-2}, x_{t-3}, \dots$. Nilai $E(\varepsilon_t) = 0, E(\varepsilon_t^2) = \sigma_\varepsilon^2$ sama untuk setiap t dan $E(\varepsilon_t \varepsilon_s) = 0$ untuk $t \neq s$. Semua akar dari $d(B)$ berada di luar *unit circle* yaitu polinom dari MA adalah *invertible*. Nilai $d_0 = 1$, dan $\sum_{j=0}^{\infty} |d_j|^2 < \infty$. Kemudian η_t adalah suatu proses *linearly deterministic* yaitu $\eta_t = P(\eta_t | x_{t-1}, x_{t-2}, \dots)$.

Pada Persamaan 3.1, $\sum_{j=0}^{\infty} d_j \varepsilon_{t-j}$ disebut sebagai *purely indeterministic*. Suatu proses dikatakan *purely indeterministic* apabila regresi dari x_t pada $x_{t-1}, x_{t-2}, x_{t-3}, \dots$ memiliki kekuatan penjelas yang menuju 0 saat $q \rightarrow \infty$. Dalam Persamaan 3.1, η_t disebut bagian deterministik karena η_t dapat diprediksi berdasarkan pengamatan masa lampau dari x_t . Berdasarkan teorema di atas, dapat dinyatakan bahwa

η_t bukan merupakan suatu trend waktu, karena asumsi kestasioneran dalam kovarian dari x_t mencegah adanya suatu trend waktu. Keberadaan dari trend waktu itu sendiri akan mengimplikasikan bahwa rata-rata dari x_t adalah sebuah fungsi dari t .

3.2 Analisis Intervensi

Model intervensi adalah suatu model yang digunakan saat kejadian eksternal di luar perkiraan (waktu intervensi tidak diketahui) maupun kejadian-kejadian eksternal yang diperkirakan (waktu intervensi diketahui) mempengaruhi variabel yang diobservasi. Tujuan utama dari analisis ini adalah mengukur besar dan lamanya efek intervensi pada suatu data deret waktu (William Wei, 2006). Berdasarkan Teorema 3.1.1, maka deret waktu Z_t dapat dilakukan suatu dekomposisi sedemikian sehingga deret waktu Z_t adalah penjumlahan dari suatu deret waktu deterministik yang berpengaruh pada saat $t \geq T$ dan suatu deret waktu stokastik N_t .

Definisi 3.2.1. Bentuk umum dari model intervensi adalah:

$$W_t = f(I_t) + N_t, \quad (3.2)$$

dengan W_t adalah suatu transformasi yang bersesuaian dari Z_t , misal $\log Z_t, (Z_t)^{1/2}$, atau Z_t itu sendiri. Fungsi $f(I_t)$ adalah fungsi dari variabel intervensi pada waktu t , variabel intervensi I_t sendiri bernilai 0 pada saat tidak ada pengaruh intervensi dan bernilai 1 pada saat ada pengaruh intervensi. Kemudian N_t adalah model ARIMA tanpa adanya pengaruh intervensi.

Jika T adalah waktu terjadinya intervensi, maka nilai $W_t = N_t$ untuk $t < T$. Pada saat intervensi terjadi $t = T$, nilai W_t sama dengan nilai N_t ditambah dengan pengaruh intervensi $f(I_t)$. Dikarenakan N_t adalah model ARIMA yang dibentuk dengan menggunakan data amatan pada saat $t < T$, maka untuk $t \geq T$ nilai N_t yang

digunakan pada Persamaan (3.2) adalah hasil peramalan N_t untuk $t \geq T$.

Contoh: Misalkan $W_t = Z_t$ untuk $t = 1, 2, \dots, n$ dan T adalah waktu terjadinya intervensi. Anggap N_t memenuhi model ARMA(1,1) yaitu

$$\begin{aligned}(1 - \phi B) N_t &= (1 - \theta B) a_t \\ N_t &= \phi N_{t-1} - \theta a_{t-1} + a_t.\end{aligned}$$

Untuk meramalkan l langkah ke depan dari N_t , menggunakan Persamaan (2.40) didapat penduga untuk $\hat{N}_{T-1}(l)$ yaitu

$$\hat{N}_{T-1}(l) = \sum_{j=1}^{\infty} \phi_j \hat{N}_{T-1}(l - j + 1). \quad (3.3)$$

Kemudian misalkan nilai pengaruh intervensi $f(I_t) = 0$ untuk $t < T$ dan bernilai ω untuk $t \geq T$, yaitu

$$f(I_t) = \begin{cases} 0 & , t < T \\ \omega & , t \geq T, \end{cases} \quad (3.4)$$

maka model intervensi seperti pada Persamaan (3.2) saat $t < T$ dapat dituliskan

$$\begin{aligned}W_t &= f(I_t) + N_t \\ Z_t &= 0 + \phi N_{t-1} - \theta a_{t-1} + a_t \\ &= \phi N_{t-1} - \theta a_{t-1} + a_t,\end{aligned}$$

dan pada saat terjadinya intervensi $t \geq T$, menggunakan penduga untuk $\hat{N}_{T-1}(l)$

pada Persamaan (3.3) didapat model intervensi

$$\begin{aligned} W_t &= f(I_t) + N_t \\ Z_t &= \omega + \hat{N}_{T-1}(l) \\ &= \omega + \sum_{j=1}^{\infty} \phi_j \hat{N}_{T-1}(l - j + 1), \end{aligned}$$

Sehingga model intervensi dari Z_t secara singkat dapat ditulis sebagai berikut

$$Z_t = \begin{cases} \phi N_{t-1} - \theta a_{t-1} + a_t & , t < T \\ \omega + \sum_{j=1}^{\infty} \phi_j \hat{N}_{T-1}(l - j + 1) & , t \geq T. \end{cases} \quad (3.5)$$

3.2.1 Fungsi Intervensi

Fungsi dari dampak intervensi $f(I_t)$ dapat meliputi suatu dampak deterministik dari waktu t dan perkembangan (*growth*) dari variabel intervensi I_t . Besarnya dampak dari variabel intervensi dapat direpresentasikan sebagai respon dinamis dari intervensi $\omega_s(B)/\delta_r(B)$.

Definisi 3.2.2. Fungsi intervensi untuk satu variabel intervensi didefinisikan

$$f(I_t) = y_t(s, r, b) = \frac{\omega_s(B)}{\delta_r(B)} B^b I_t, \quad (3.6)$$

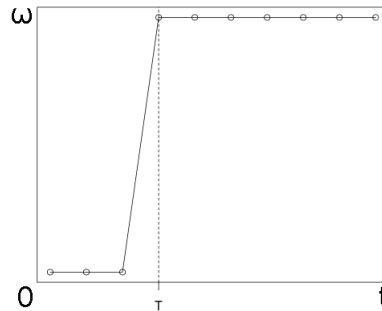
dengan b adalah jumlah periode sebelum input I_t mulai mempengaruhi Z_t . Besarnya kenaikan/penurunan level (*mean*) ditunjukkan dengan ω_s , dengan s menunjukkan jumlah dari nilai I_t yang mempengaruhi Z_t , $\omega_s(B) = \omega_0 - \omega_1 B - \dots - \omega_s B^s$. Kemudian δ_r adalah bentuk *decay* sebagai akibat pengaruh intervensi, r menunjukkan jumlah dari masa lalu yang berhubungan dengan Z_t , $\delta_r(B) = 1 - \delta_1 B - \dots - \delta_r B^r$.

Akar-akar dari $\omega_s(z)$ dan $\delta_r(z)$ diasumsikan berada di luar *unit circle*. Asumsi ini diperlukan karena, apabila akar-akar dari $\omega_s(B)$ dan $\delta_r(B)$ berada di da-

lam *unit circle* maka $\omega_s(z)$ dan $\delta_r(z)$ tidak konvergen atau akan membesar seiring berjalannya waktu. Beberapa contoh pola pengaruh fungsi intervensi, misalkan I_t adalah variabel intervensi yang bernilai 1 untuk $t < T$ dan bernilai 0 untuk $t \geq T$:

Fungsi Intervensi Langsung Permanen

Intervensi terjadi secara langsung pada waktu $t = T$ dan pengaruh dari intervensi berlangsung secara permanen atau tetap pada waktu $t = T + 1, T + 2, \dots, n$ atau secara ilustratif dapat dilihat pada Gambar 3.1.



Gambar 3.1: Pola Intervensi Langsung Permanen

Fungsi intervensi langsung permanen dinotasikan

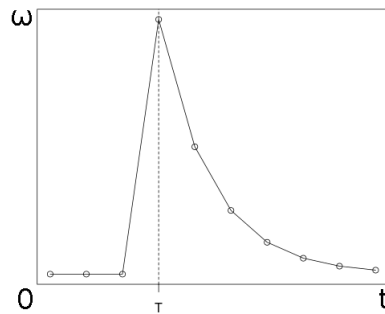
$$f(I_t) = \omega_0 I_t. \quad (3.7)$$

Fungsi Intervensi Langsung Sementara

Intervensi terjadi secara langsung pada waktu $t = T$ namun pengaruh dari intervensi berkurang seiring dengan berjalannya waktu $t = T + 1, T + 2, \dots, n$ atau secara ilustratif dapat dilihat pada Gambar 3.2.

Fungsi intervensi langsung sementara dinotasikan

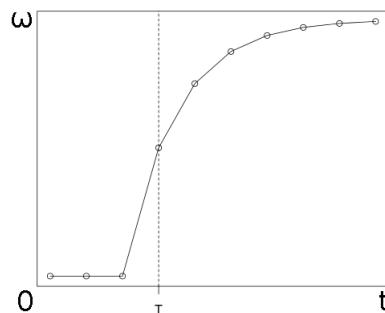
$$f(I_t) = (\omega_0 - \omega_1 B) I_t. \quad (3.8)$$



Gambar 3.2: Pola Intervensi Langsung Sementara

Fungsi Intervensi Gradual Permanen

Intervensi terjadi secara bertahap dimulai dari waktu $t = T$ dan terus bertambah secara perlahan seiring berjalannya waktu, pengaruh intervensi ini berlangsung secara permanen seperti yang dapat dilihat pada Gambar 3.3.



Gambar 3.3: Pola Intervensi Gradual Permanen

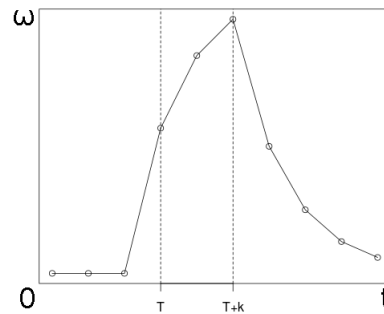
Fungsi intervensi gradual permanen dinotasikan

$$f(I_t) = \frac{\omega_0}{1 - \delta B} I_t. \quad (3.9)$$

Fungsi Intervensi Gradual Sementara

Intervensi terjadi secara bertahap dimulai dari waktu $t = T$ dan terus bertambah sampai waktu $t = T + k$ kemudian setelah itu pengaruh intervensi perlahan-

lahan berkurang untuk $t > T + k$, ilustrasi pola pengaruh intervensi ini dapat dilihat pada Gambar 3.4.



Gambar 3.4: Pola Intervensi Gradual Sementara

Fungsi intervensi gradual sementara dinotasikan

$$f(I_t) = \frac{\omega_0 - \omega_1 B}{1 - \delta B} I_t. \quad (3.10)$$

3.2.2 Variabel Intervensi

Secara umum ada dua jenis variabel intervensi (I_t), yaitu fungsi *step* dan fungsi *pulse*. Kejadian intervensi yang terjadi sejak waktu T dan berdampak terus-menerus dalam waktu yang cukup panjang disebut fungsi *step* dan kejadian intervensi terjadi hanya pada waktu T saja dan tidak berlanjut pada waktu selanjutnya disebut fungsi *pulse*.

Definisi 3.2.3. Jika T adalah waktu terjadinya intervensi dan $t \in \mathbb{Z}$, maka fungsi *Step* (S_t) dan *Pulse* (P_t) didefinisikan

$$S_t = \begin{cases} 0 & , t < T \\ 1 & , t \geq T \end{cases} \quad (3.11)$$

dan

$$P_t = \begin{cases} 0 & ,t \neq T \\ 1 & ,t = T. \end{cases} \quad (3.12)$$

Teorema 3.2.1. *Jika S_t adalah fungsi step, P_t adalah fungsi Pulse dan T adalah waktu terjadinya intervensi, maka*

$$(1 - B) S_t = P_t \quad ,t \in \mathbb{Z} \quad (3.13)$$

Bukti. Misal T adalah waktu terjadinya intervensi, maka

$$(1 - B) S_t = S_t - S_{t-1} \quad (3.14)$$

$$= \begin{cases} 0 & ,t < T \\ 1 & ,t = T \\ 0 & ,t > T \end{cases} \quad (3.15)$$

sehingga berdasarkan Definisi 3.2.3 fungsi $(1 - B)S_t$ dapat tuliskan kembali menjadi

$$(1 - B) S_t = \begin{cases} 0 & ,t \neq T \\ 1 & ,t = T \end{cases} = P_t \quad ,t \in \mathbb{Z}. \quad (3.16)$$

□

Dengan demikian, dapat dinyatakan bahwa fungsi *step* untuk suatu deret waktu yang stasioner akan menghasilkan respon yang identik dengan fungsi *pulse* untuk deret waktu yang terintegrasi ARIMA($p, 1, q$).

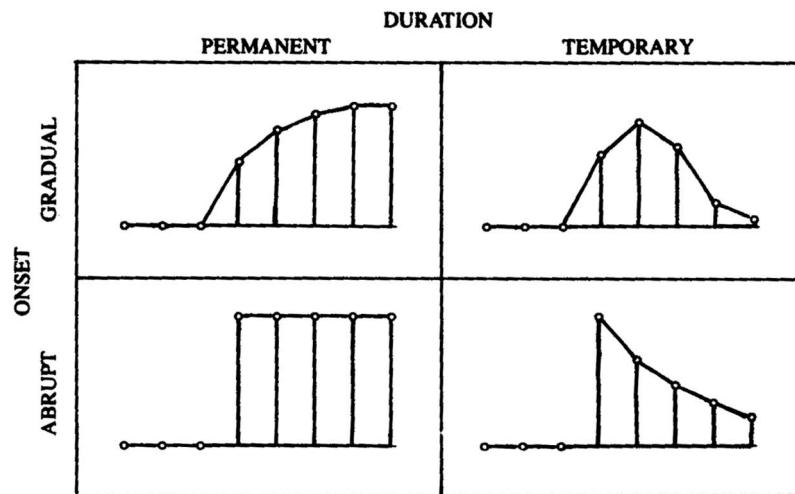
3.2.3 Orde Intervensi

Dalam Persamaan (3.6), b , r , dan s disebut sebagai orde untuk model intervensi. Orde intervensi ini dapat diidentifikasi dengan melihat plot residual yang

diperoleh dari selisih antara data hasil pengamatan dengan nilai peramalan menggunakan *noise model*. Misal residual dinotasikan dengan Z_t^* , maka

$$Z_t^* = Z_t - N_t = f(I_t). \quad (3.17)$$

Nilai b ditentukan dengan melihat kapan intervensi mulai terjadi, nilai s menunjukkan kapan gerak bobot respon mulai mengalami kenaikan/penurunan (*delay*), dan r menunjukkan pola dari residual (*decay*).



Gambar 3.5: Tipe Dari Komponen Intervensi. (McDowall, 1980)

Berdasarkan orde intervensi, seperti yang terlihat pada Gambar 3.5, $f(I_t)$ dapat dibedakan menjadi empat tipe dengan dua karakteristik yaitu, waktu permulaan (*onset*) dan durasi terjadinya intervensi (*duration*). Waktu permulaan (*onset*) intervensi dimungkinkan terjadi secara mendadak (*abrupt*) ataupun secara berting-
kat (*gradual*). Dalam durasi terjadinya, dampak intervensi (*duration*) dibedakan menjadi dampak permanen dan dampak sementara (*temporary*). Beberapa contoh pola pembobot bersesuaian dengan orde intervensi dapat dilihat pada Lampiran C, $v(B)$ disebut sebagai pembobot respon impuls, $|v_i| \leq 1$ untuk $i = 1, 2, 3, \dots$

3.3 Pemodelan Intervensi

3.3.1 Pengelompokan Data Deret Waktu

Berdasarkan waktu dari intervensi, analisis intervensi juga dibagi menjadi dua yaitu, analisis intervensi saat waktu intervensi diketahui dan analisis intervensi saat waktu intervensi tidak diketahui. Analisis intervensi saat waktu intervensi tidak diketahui mengarah kepada analisis pencilaan data deret waktu yang lebih umum, sedangkan saat waktu intervensi diketahui lebih berfokus kepada mengukur pengaruh dari intervensi (William Wei, 2006).

Saat waktu intervensi T telah diketahui, langkah pertama dalam analisis intervensi adalah mengelompokan data deret waktu menjadi dua bagian yaitu, data deret waktu sebelum intervensi $t < T$ dan data deret waktu setelah intervensi $t \geq T$. Data deret waktu sebelum intervensi akan digunakan untuk memodelkan N_t yang mengikuti prosedur dari pemodelan ARIMA. Kemudian, hasil dari peramalan N_t pada saat $t \geq T$ akan digunakan untuk mengidentifikasi orde intervensi.

3.3.2 Pemodelan ARIMA Untuk Data Sebelum Intervensi

Pemodelan ARIMA data deret waktu sebelum intervensi N_t dilakukan dalam beberapa tahapan seperti yang ada dalam Sub-Bab 2.4, yaitu:

- Identifikasi Model
- Estimasi Parameter Model
- Diagnosis Model
- Kriteria Pemilihan Model

Secara ilustratif prosedur pemodelan ARIMA dapat dilihat dalam *flowchart* prosedur pemodelan ARIMA (Lampiran A).

3.3.3 Identifikasi Orde Intervensi

Setelah mendapatkan model N_t yang cukup memadai, selanjutnya model N_t akan digunakan untuk mengidentifikasi orde intervensi dengan cara mengamati pola residual yang terjadi antara hasil peramalan N_t dengan Z_t pada saat $t \geq T$ seperti yang terlihat dalam Persamaan (3.17). Orde (b, s, r) pada Persamaan (3.17) dapat diketahui dengan melihat grafik residual dengan batas 3 kali nilai akar MSE (RMSE) dari N_t . Nilai b ditentukan dengan melihat kapan efek intervensi mulai terjadi, nilai s ditentukan dari waktu *delay* yang diperlukan agar data menjadi kembali stabil dihitung dari waktu terjadinya intervensi, dan r merupakan *r time lag* berikutnya (setelah b dan s) saat data menunjukkan *decay* dari residual.

3.3.4 Estimasi Parameter Intervensi

Untuk model ARMA yang stasioner dan mengandung intervensi yang terjadi pada waktu $t = T$, Persamaan (3.2) dapat dituliskan menjadi

$$\phi_p(B)Z_t = \frac{\omega_s(B)}{\delta_r(B)} B^b I_t + \theta_q(B)a_t, \quad (3.18)$$

dengan menyamakan penyebut maka persamaan di atas dapat ditulis menjadi:

$$\delta_r(B) \phi_p(B) Z_t = \omega_s(B) I_{t-b} + \delta_r(B) \theta_q(B) a_t,$$

atau sama dengan

$$k(B)Z_t = \omega_s(B) I_{t-b} + m(B)a_t, \quad (3.19)$$

dengan:

$$k(B) = 1 - k_1B - k_2B - \dots - k_{p+r}B^{p+r}$$

$$m(B) = 1 - m_1B - m_2B - \dots - m_{q+r}B^{q+r}$$

sehingga diperoleh nilai untuk a_t yaitu:

$$a_t = \sum_{i=1}^{q+r} m_i a_{t-i} + Z_t - \sum_{j=1}^{p+r} k_j Z_{t-j} - \omega_s(B)I_{t-b}. \quad (3.20)$$

Dengan menggunakan metode *conditional MLE* seperti pada Persamaan (2.26) dapat diperoleh estimasi parameter model intervensi ω, δ diberikan $Z, \hat{\phi}, \hat{\theta}$ dengan meminimumkan $S_*(\omega, \delta)$ jumlah kuadrat bersyarat

$$S_*(\delta, \omega) = \sum_{t=T}^n a_t^2(\omega, \delta | Z, \hat{\phi}, \hat{\theta}). \quad (3.21)$$

Contoh: misalkan $W_t = Z_t$ untuk $t = 1, 2, \dots, n$ dan $t = T$ adalah waktu terjadinya intervensi. Anggap N_t memenuhi model ARMA(1,1) dan dituliskan

$$(1 - 0.43B) N_t = (1 - 0.37B) a_t, \quad (3.22)$$

lalu misalkan fungsi intervensi $f(I_t)$ memiliki orde (0,0,1) dan dituliskan

$$f(I_t) = \frac{\omega_0}{1 - \delta B} P_t, \quad (3.23)$$

maka didapat fungsi untuk Z_t

$$(1 - 0.43B) Z_t = \frac{\omega_0}{1 - \delta B} P_t + (1 - 0.37B) a_t \quad (3.24)$$

atau

$$a_t = (\delta + 0.37) a_{t-1} - 0.37\delta a_{t-2} + Z_t - (\delta + 0.43)Z_{t-1} + 0.43\delta Z_{t-2} - \omega_0 P_t, \quad (3.25)$$

sehingga estimasi untuk parameter ω_0 dan δ dapat diperoleh dengan meminimumkan jumlah kuadrat bersyarat $S_*(\omega_0, \delta)$ seperti pada Persamaan (3.21).

3.3.5 Diagnosis Model Intervensi

Setelah melakukan estimasi parameter intervensi, langkah selanjutnya adalah memeriksa model yang telah didapat apakah sudah cukup memadai untuk data yang ada. Untuk melihat apakah model sudah memadai atau belum, dapat diperiksa melalui residual a_t yang dihasilkan oleh model. Model ini dikatakan memadai, jika residual memenuhi asumsi *white noise* dan normalitas. *Flowchart* prosedur pemo-
delan intervensi dapat dilihat pada Lampiran B.

3.4 Peramalan Menggunakan Model Intervensi

Misal pada $t = n$ terdapat $Z_n, Z_{n-1}, Z_{n-2}, \dots$ observasi dan akan meramal-
kan l langkah ke depan Z_{n+l} dari Persamaan (3.17) diperoleh

$$Z_{n+l} = X_{n+l} + N_{n+l}, \quad (3.26)$$

dengan $X_t = f(I_t)$ adalah pengaruh intervensi pada waktu t dan N_t adalah model ARIMA tanpa adanya pengaruh intervensi. Jika Z_t mengandung suatu intervensi, maka untuk peramalan l langkah ke depan selain dibutuhkan peramalan dari N_t dibutuhkan pula *future value* dari pengaruh intervensi X_t , sehingga penduga $\hat{Z}_n(l)$ adalah

$$\begin{aligned} \hat{Z}_n(l) &= E(Z_{n+l} | Z_n, Z_{n-1}, Z_{n-2}, \dots) \\ &= E(X_{n+l} | X_n, X_{n-1}, X_{n-2}, \dots) \\ &\quad + E(N_{n+l} | N_n, N_{n-1}, N_{n-2}, \dots). \end{aligned} \quad (3.27)$$

Berdasarkan Persamaan (3.2), didapat *future value* dari pengaruh intervensi X_{n+l} sebagai berikut,

$$\begin{aligned} X_{n+l} &= \frac{\omega_s(B) B^b}{\delta_r(B)} I_{n+l} \\ X_{n+l} &= \sum_{i=0}^{\infty} h_i I_{n+l-i}, \end{aligned} \quad (3.28)$$

sehingga penduga $\hat{Z}_n(l)$ yang mengandung intervensi adalah

$$\hat{Z}_n(l) = \sum_{i=0}^{\infty} h_i \hat{I}_n(l-i) + \sum_{j=1}^{\infty} \pi_j \hat{N}_n(l-j). \quad (3.29)$$

Dengan selang kepercayaan $(1 - \alpha)100\%$ untuk batas peramalan sama seperti pada Persamaan (2.45).

Contoh: Misalkan parameter intervensi pada Persamaan (3.24) adalah $\omega_0 = 50.34$ dan $\delta = 0.61$, maka berdasarkan pada Persamaan (3.29) didapat penduga $\hat{Z}_n(l)$ yaitu

$$\hat{Z}_n(l) = 50.34 \left[\sum_{i=0}^{\infty} (0.61)^i \hat{P}_n(l-i) \right] - \sum_{j=1}^{\infty} (-0.06)^j (0.37)^{j-1} \hat{N}_n(l-j). \quad (3.30)$$

3.5 Aplikasi Penggunaan Analisis Intervensi

3.5.1 Data Pendapatan Domestik Bruto

Produk Domestik Bruto (PDB) merupakan salah satu indikator penting untuk mengetahui kondisi ekonomi di suatu negara dalam suatu periode tertentu. PDB pada dasarnya merupakan jumlah nilai tambah yang dihasilkan oleh seluruh unit usaha dalam suatu negara tertentu, atau merupakan jumlah nilai barang dan jasa akhir yang dihasilkan oleh seluruh unit ekonomi. Berikut ini adalah data Pendapatan Domestik Bruto Indonesia dari tahun 1970 sampai dengan tahun 2013 dalam US Dollar (sumber: <http://data.worldbank.org/indicator/NY.GDP.MKTP.CD>):

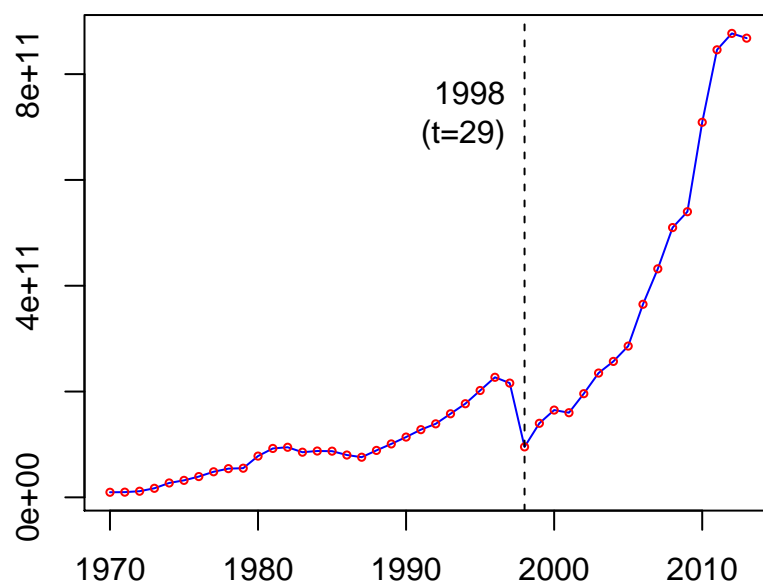
```
Time Series:
Start = 1970
End = 2013
Frequency = 1
[1] 9656740014 9849117953 11605084560 17171181163
[5] 27227710999 32147953008 39328674730 48396143465
[9] 54298158340 55122620334 78013206038 92473878832
[13] 94715163814 85369201879 87612439197 87338874330
[17] 80060657612 75929617715 88787623310 101000000000
[21] 114000000000 128000000000 139000000000 158000000000
[25] 177000000000 202000000000 227000000000 216000000000
[29] 95445548017 140000000000 165000000000 160000000000
[33] 196000000000 235000000000 257000000000 286000000000
[37] 365000000000 432000000000 510000000000 540000000000
[41] 709000000000 846000000000 877000000000 868000000000
```

dengan waktu intervensi $T = 29$ atau intervensi terjadi pada tahun 1998. Tahun 1998 dipilih menjadi waktu terjadinya intervensi karena pada tahun tersebut Indonesia mengalami berbagai kejadian bersejarah seperti demonstrasi besar-besaran mahasiswa, lengsernya Presiden Soeharto, dan kerusuhan terhadap etnis Tionghoa yang mengakibatkan krisis moneter sampai sekitar tahun 2000. Data PDB ini ke-

mudian dikelompokkan menjadi dua yaitu data sebelum intervensi (N_t) untuk waktu $t = 1, 2, \dots, 28$ dan data setelah intervensi untuk waktu $t \geq 29$.

3.5.2 Pemodelan Data PDB Sebelum Intervensi (N_t)

Pola untuk data PDB Indonesia dari tahun 1970 sampai dengan 2013 (Z_t) diperlihatkan pada Gambar 3.6:

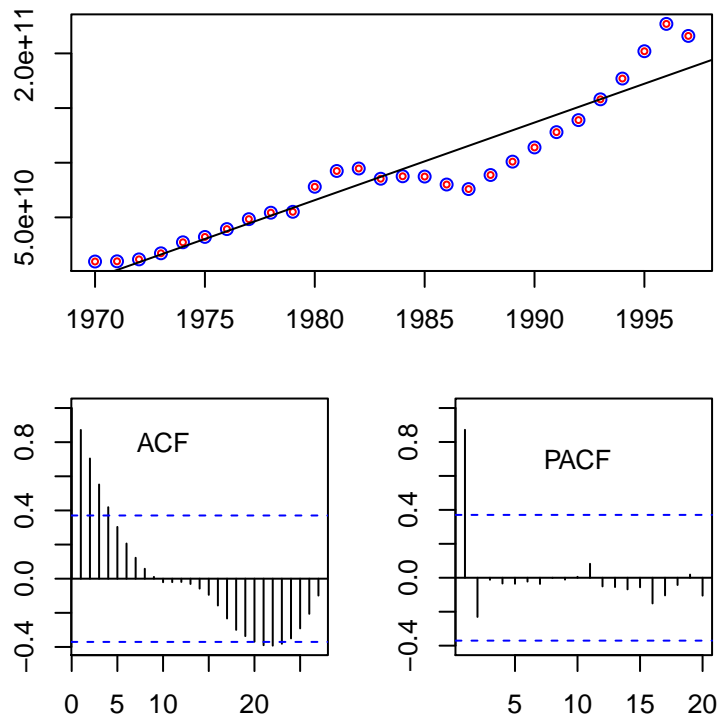


Gambar 3.6: Plot Data PDB dari tahun 1930 sampai 2013

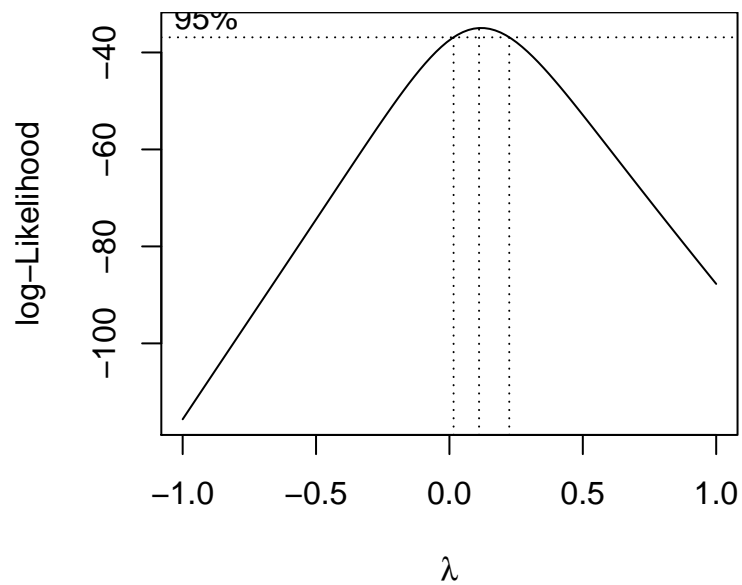
dari Gambar 3.6 terlihat bahwa data PDB dari tahun 1970 mempunyai trend naik, namun pada tahun 1998 data PDB mengalami penurunan yang kemudian pada tahun-tahun berikutnya kembali mengalami trend naik yang lebih cepat.

Identifikasi Model ARIMA N_t

Data PDB sebelum intervensi terjadi pada waktu $t < 29$ atau terjadi dari tahun 1970 sampai tahun 1997. Dari Gambar 3.7, terlihat bahwa data PDB sebelum intervensi mengikuti suatu pola trend naik.



Gambar 3.7: Plot Data PDB dari tahun 1930 sampai 1997 dengan nilai ACF dan PACF yang bersesuaian.

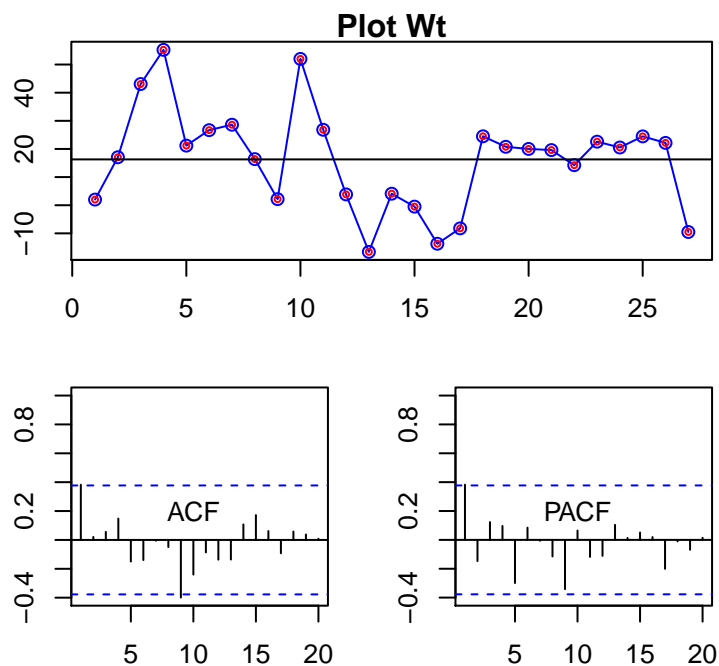


Gambar 3.8: Selang kepercayaan 95% untuk λ .

Sebelum melakukan analisis lebih lanjut, harus diperiksa terlebih dahulu apakah data sudah stationer dalam ragam dan rata-rata. Data dikatakan sudah stasio-

ner dalam ragam jika nilai λ untuk data tersebut sama dengan 1. Data PDB dari tahun 1930 sampai 1997 bersifat tidak stasioner karena memiliki nilai $\lambda = 0.2$. Hal ini diperlihatkan oleh plot antara log likelihood dengan beberapa nilai Lambda yang disajikan pada Gambar 3.8, dimana fungsi log likelihood maksimum pada 0.2. Untuk membuat data menjadi stasioner dalam ragam dilakukan transformasi. Fungsi transformasi Box-Cox yang bersesuaian untuk $\lambda = 0.2$ berdasarkan Persamaan (2.1) adalah $Y_t = \frac{Z_t^{0.2} - 1}{0.2}$.

Pada Gambar 3.7 terlihat bahwa N_t mengikuti suatu trend naik dan nilai ACF menurun secara perlahan yang berarti N_t belum stasioner dalam rata-rata karena nilai rata-rata masih bergantung pada waktu. Agar N_t menjadi stasioner dalam rata-rata maka perlu dilakukan pembedaan terhadap Y_t yaitu $W_t = (1 - B)Y_t = Y_t - Y_{t-1}$.



Gambar 3.9: Plot W_t dengan nilai ACF dan PACF yang bersesuaian

Plot W_t beserta nilai ACF dan PACF pada Gambar 3.9 mengindikasikan bahwa W_t sudah stasioner dalam rata-rata setelah dilakukan pembedaan. ACF W_t terpotong pada lag-9 namun tidak terpotong pada lag sebelumnya ataupun berulang

secara periodik, ini mengindikasikan bahwa lag-9 dapat diabaikan dan model yang mungkin untuk N_t berdasarkan Tabel 2.1 adalah ARIMA(0,1,0), ARIMA(1,1,0), ARIMA(0,1,1), dan ARIMA(1,1,1).

Estimasi Parameter ARIMA

Tahap selanjutnya setelah identifikasi model adalah melakukan estimasi terhadap parameter model. Dengan bantuan *software R*, didapat parameter untuk tiap-tiap model (lampiran D) dan nilai AIC serta BIC yang bersesuaian seperti pada Tabel 3.1.

MODEL	AIC	AICc	BIC
ARIMA(0,1,0)	250.28	250.44	251.57
with drift	235.65	236.15	238.24
ARIMA(1,1,0)	235.21	235.71	237.8
with drift	233	234.04	236.89
ARIMA(0,1,1)	239.95	240.45	242.54
with drift	232.27	233.31	236.16
ARIMA(1,1,1)	237.14	238.18	241.02
with drift	234.05	235.87	239.24

Tabel 3.1: Nilai AIC, AICc, dan BIC untuk berbagai model ARIMA

Berdasarkan Tabel 3.1, nilai AIC, AICc, dan BIC terendah dimiliki oleh model ARIMA(0,1,1) *with drift* yang dinotasikan sebagai berikut

$$Y_t = 15.65 + Y_{t-1} + a_t - 0.44a_{t-1}, \quad (3.31)$$

artinya nilai pengamatan pada waktu ke- t (Y_t) dipengaruhi oleh *drift* (sebesar 15.65), nilai pengamatan pada waktu sebelumnya (Y_{t-1}) serta *error* pada waktu ke- t (a_t) dan *error* pada waktu sebelumnya (a_{t-1}) dimana $N_t = a_t = 0$ untuk $t \leq 0$.

Diagnosis Model ARIMA

Suatu model ARIMA dikatakan telah memadai jika model tersebut menghasilkan residual yang memenuhi asumsi *white noise* dan normalitas. Untuk menguji apakah suatu model telah memenuhi asumsi *white noise* dan normalitas, uji yang digunakan berturut-turut adalah Uji Ljung-Box dan Uji Kolmogorov-Smirnov. Jika model tidak memenuhi salah satu asumsi di atas, maka model dikatakan belum cukup memadai sehingga perlu dicari model lain yang dapat memenuhi asumsi tersebut.

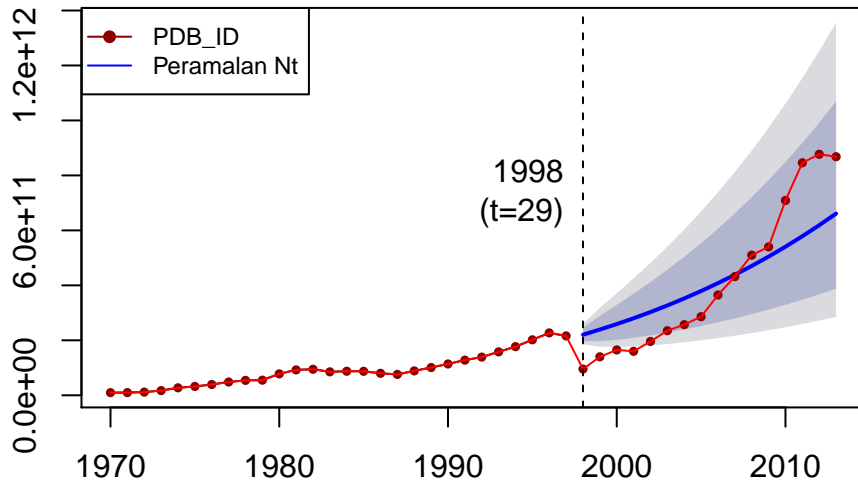
<pre>Box-Ljung test data: Nt_arima\$residuals X-squared = 4.6519, df = 5, p-value = 0.4598</pre>
<pre>One-sample Kolmogorov-Smirnov test data: Nt_arima\$residuals D = 0.12779, p-value = 0.7032 alternative hypothesis: two-sided</pre>

Dari hasil perhitungan statistik uji untuk uji Ljung-Box dan uji Kolmogorov-Smirnov di atas, terlihat bahwa pada $\alpha = 0.05$ hipotesis nol tidak dapat ditolak yang berarti bahwa nilai residual untuk model ARIMA(0,1,1) *with drift* memenuhi asumsi *white noise* dan normalitas. Dengan dipenuhinya asumsi ini, maka model ARIMA(0,1,1) *with drift* dikatakan memadai dan dapat digunakan untuk peramalan.

3.5.3 Identifikasi Orde Intervensi

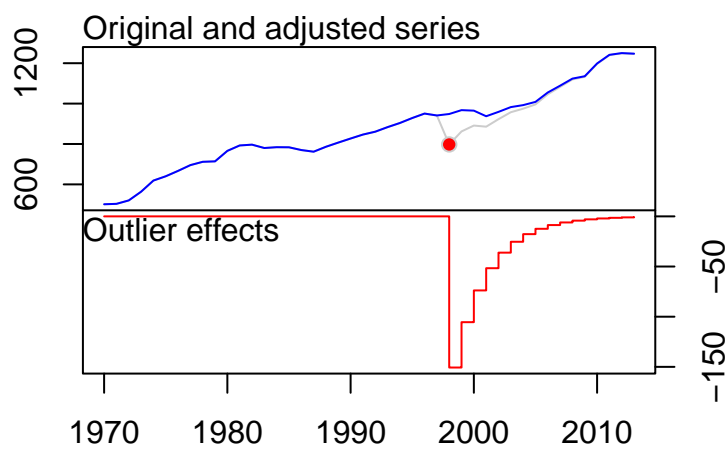
Identifikasi orde intervensi dilakukan dengan mengamati selisih antara hasil peramalan dari N_t untuk $t \geq T$ dan nilai amatan Z_t . Secara ilustratif, perbandingan antara Z_t dan hasil peramalan N_t dapat dilihat pada Gambar 3.10. Terlihat bahwa

nilai Z_t terletak di dalam selang kepercayaan 80% dari hasil peramalan N_t pada saat $t \geq 34$. Hal ini mengindikasikan dampak intervensi sudah mulai menghilang dan Z_t kembali mengikuti pola dari N_t pada saat $t \geq 34$.



Gambar 3.10: Plot perbandingan antara Z_t dengan hasil peramalan N_t

Untuk mengidentifikasi nilai orde b, s , dan r , perlu diamati apakah pola dari dampak intervensi mengikuti salah satu dari empat pola yang ada pada Gambar 3.5. Dengan menggunakan *tsoulier package* yang ada dalam *software R* didapat pola dari dampak intervensi seperti pada Gambar 3.11. Terlihat bahwa intervensi terjadi



Gambar 3.11: Plot pola intervensi

secara mendadak (*abrupt*) dan hanya berdampak sementara (*temporary*) sehingga pola dari dampak intervensi mengikuti pola *abrupt temporary* yang mengindikasikan bahwa intervensi menggunakan fungsi *pulse* dengan nilai $b = 0, s = 0$, dan $r = 1$. Fungsi dari dampak intervensi untuk data PDB Indonesia dapat dituliskan sebagai berikut:

$$f(I_t) = \frac{\omega_0}{1 - \delta_1 B} P_{29}. \quad (3.32)$$

3.5.4 Estimasi Parameter Intervensi

Tahap selanjutnya setelah identifikasi orde intervensi adalah melakukan estimasi terhadap parameter intervensi. Dengan bantuan *software R*, didapat parameter intervensi ω_0 dan δ_0 seperti yang terlihat pada Tabel 3.2:

Tabel 3.2: Estimasi Parameter Intervensi

<i>Dependent variable:</i>	
ma1	0.439*** (0.118)
T29-AR1	0.205** (0.080)
T29-MA0	-112.491*** (14.570)

Note: *p<0.1; **p<0.05; ***p<0.01

didapat nilai $\omega_0 = -112.491$ dan $\delta_0 = 0.205$. Kedua parameter intervensi tersebut signifikan pada taraf nyata $\alpha = 0.05$. Model intervensi untuk data PDB Indonesia dari tahun 1970 sampai 2013 berdasarkan Persamaan (3.31) dan (3.32) adalah

$$Y_t = \frac{-112.491}{1 - 0.205B} P_{29} + 15.65 + Y_{t-1} + a_t - 0.44a_{t-1}. \quad (3.33)$$

3.5.5 Diagnosis Model Intervensi

Suatu model intervensi dikatakan telah memadai jika model tersebut menghasilkan residual yang memenuhi asumsi *white noise* dan normalitas. Untuk menguji apakah suatu model telah memenuhi asumsi *white noise* dan normalitas, uji yang digunakan berturut-turut adalah Uji Ljung-Box dan Uji Kolmogorov-Smirnov. Jika model tidak memenuhi salah satu asumsi di atas, maka model dikatakan belum cukup memadai sehingga perlu dicari model lain yang dapat memenuhi asumsi tersebut.

```
Box-Ljung test

data: PDB_arima$residuals
X-squared = 14.784, df = 8, p-value = 0.06348

One-sample Kolmogorov-Smirnov test

data: PDB_arima$residuals
D = 0.14849, p-value = 0.2597
alternative hypothesis: two-sided
```

Dari hasil perhitungan uji Ljung-Box dan uji Kolmogorov-Smirnov, terlihat bahwa pada $\alpha = 0.05$ hipotesis nol tidak dapat ditolak yang berarti bahwa nilai residual untuk model intervensi sudah memenuhi asumsi *white noise* dan normalitas. Dengan dipenuhinya asumsi ini, maka model intervensi dikatakan memadai dan dapat digunakan untuk peramalan.

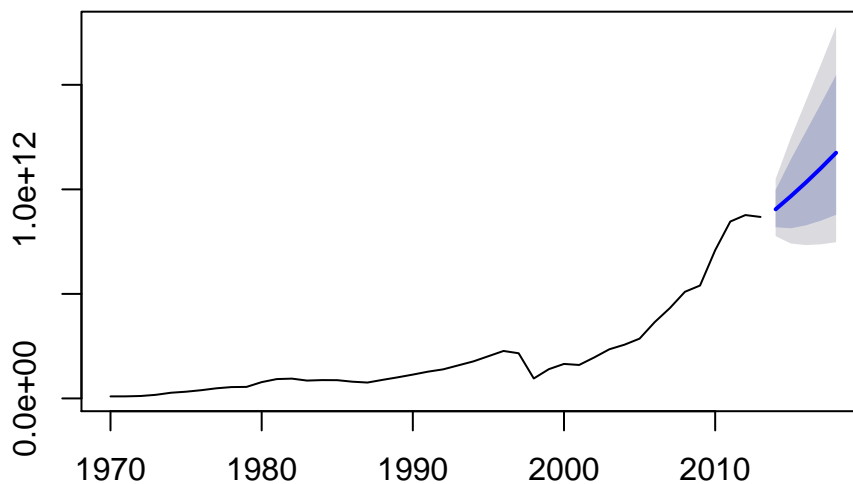
3.5.6 Peramalan PDB Menggunakan Model Intervensi

Model intervensi yang sudah memadai berdasarkan diagnosis model selanjutnya dapat digunakan untuk peramalan. Peramalan dilakukan pada model Y_t seperti yang terlihat pada Persamaan (3.33), dimana Y_t adalah hasil transformasi Box-Cox dari Z_t dengan nilai $\lambda = 0.201$. Hasil peramalan menggunakan model Y_t

tersebut kemudian dilakukan invers transformasi Box-Cox dengan nilai $\lambda = 0.201$ agar hasil yang didapat dari peramalan kembali menggunakan satuan aslinya. Perhitungan peramalan dilakukan dengan bantuan *software R* dengan hasil sebagai berikut:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2014	9.045913e+11	818883650577	9.973193e+11	776215514470	1.049392e+12
2015	9.669871e+11	813820112972	1.142359e+12	740978902963	1.244971e+12
2016	1.032775e+12	827954842710	1.276198e+12	733510015384	1.422443e+12
2017	1.102090e+12	850383812017	1.410060e+12	737126698089	1.598875e+12
2018	1.175072e+12	878177165628	1.547203e+12	747311925442	1.779178e+12

atau secara ilustratif dapat diperlihatkan dalam grafik pada gambar 3.12.



Gambar 3.12: Plot Hasil Permalan PDB

Hasil peramalan menunjukkan bahwa nilai PDB Indonesia dari tahun 2014 sampai 2018 akan mengalami trend naik dengan ragam yang semakin besar setiap tahunnya. Berdasarkan selang kepercayaan 95% dari hasil peramalan, kemungkinan nilai PDB terendah selama lima tahun tersebut berada pada tahun 2016 dengan nilai PDB sebesar 739 milyar US Dollar dan kemungkinan nilai PDB tertinggi berada pada tahun 2018 sebesar 1762 milyar US Dollar.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan uraian pada bab-bab sebelumnya mengenai metode pemodelan intervensi maka diperoleh kesimpulan sebagai berikut:

1. Pemodelan data deret waktu yang mengandung perubahan pola dengan menggunakan analisis intervensi dilakukan dengan terlebih dahulu mengelompokan antara data sebelum terjadinya intervensi (N_t) dan data sesudah terjadinya intervensi ($f(I_t)$). Data sesudah terjadinya intervensi atau dapat disebut sebagai fungsi dari dampak intervensi ($f(I_t)$) didapat dengan mengestimasi ω_s (parameter perubahan level) dan δ_r (parameter *decay* dari dampak intervensi). Parameter tersebut didapat dengan membandingkan antara nilai dari peramalan model ARIMA N_t dengan data amatan pada saat $t \geq T$. Berdasarkan hasil perbandingan ini, dapat ditentukan juga apakah I_t merupakan suatu fungsi *step* atau fungsi *pulse*.
2. Peramalan data deret waktu yang mengandung perubahan pola dilakukan dengan terlebih dahulu menghitung nilai peramalan dari N_t dan *future value* dari $f(I_t)$ untuk $t = n + l$, dimana n adalah banyaknya data amatan dan l adalah banyaknya step yang ingin diramalkan.
3. Model intervensi untuk data Pendapatan Domestik Bruto Indonesia dari tahun

1970 sampai dengan 2013 pada contoh kasus dalam skripsi ini adalah

$$Y_t = \frac{-112.491}{1 - 0.205B} P_{29} + 15.65 + Y_{t-1} + a_t - 0.44a_{t-1}. \quad (4.1)$$

4. Hasil peramalan PDB Indonesia untuk tahun 2014, 2015, 2016, 2017, dan 2018 dengan menggunakan model intervensi untuk data PDB Indonesia dari tahun 1970 sampai 2013 pada contoh kasus dalam skripsi adalah sebagai berikut:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2014	9.045913e+11	818883650577	9.973193e+11	776215514470	1.049392e+12
2015	9.669871e+11	813820112972	1.142359e+12	740978902963	1.244971e+12
2016	1.032775e+12	827954842710	1.276198e+12	733510015384	1.422443e+12
2017	1.102090e+12	850383812017	1.410060e+12	737126698089	1.598875e+12
2018	1.175072e+12	878177165628	1.547203e+12	747311925442	1.779178e+12

dengan nilai real dari PDB Indonesia untuk tahun 2014 dan 2015 adalah 890 dan 861 milyar US Dollar.

4.2 Saran

Metode yang dibahas dalam skripsi ini berbasis pada ARIMA Box-Jenkins tanpa adanya pengaruh musiman (*seasonal*) dan waktu terjadinya intervensi sudah terlebih dahulu diketahui. Diharapkan untuk penulisan selanjutnya, dapat dibahas metode ARIMA Box-Jenkins dengan adanya pengaruh musiman (*seasonal*) atau waktu terjadinya intervensi tidak diketahui (*Outlier Analysis*).

DAFTAR PUSTAKA

- [1] Akbar, M. Sjahid. Jerry Dwi Trijoyo Purnomo dan Nurul Azizah. 2009. Analisis Peramalan Index Harga Saham KOSPI dengan Menggunakan Metode Intervensi. *Seminar Nasional Statistika IX*. Institut Teknologi Sepuluh Nopember.
- [2] Assauri, S. 1984. *Teknik dan Metode Peramalan. Penerapannya Dalam Ekonomi dan Dunia Usaha Edisi Satu*. LP Fakultas Ekonomi UI. Jakarta.
- [3] Budiarti, Lenny. Tarno dan Budi Warsito. 2013. Analisis Intervensi dan Deteksi Outlier Pada Wisatawan Domestik. *Jurnal Gaussian*. Vol. 2, No. 1, 2013, Hal: 39-48.
- [4] Box G.E.P. dan G.C. Tiao. 1975. Intervention Analysis with Applications to Economic and Environmental. *Journal of the American Statistical Association*, Vol 70, No. 349 (Mar., 1975), 70-79.
- [5] Box G.E.P. dan G.M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control, Revised Edition*. California: Holden-Day.
- [6] Brockwell, Peter J and Richard A. Davis. 2002. *Introduction to Time Series and Forecasting Second Edition*. New York: Springer.
- [7] Cryer, Jonathan D. and Kung-Sik Chan. 2008. *Time Series Analysis With Applications in R, Second Edition*. New York: Springer.
- [8] Cochrane, John H. 1997. *Time Series for Macroeconomics and Finance*. University of Chicago.

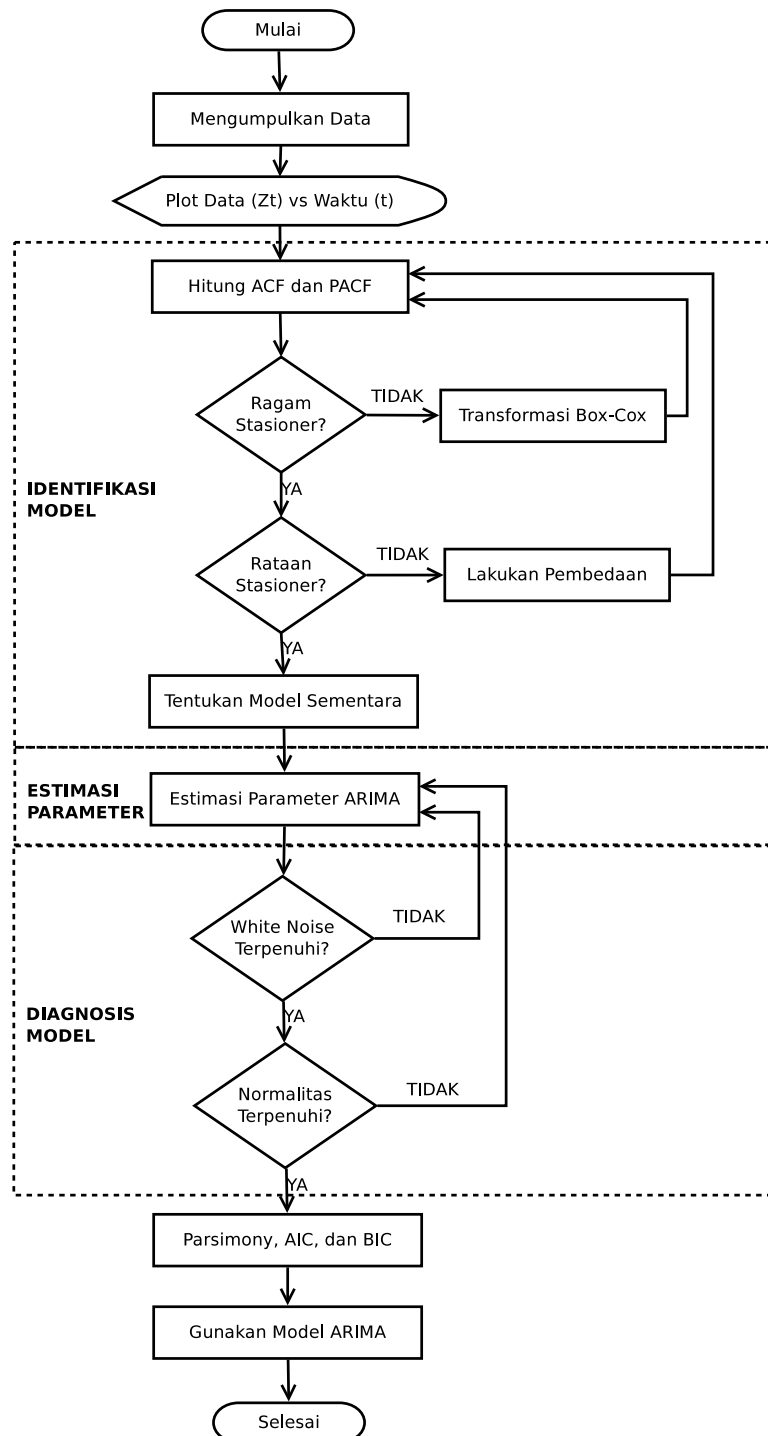
- [9] Durbin, J. 1960. The Fitting of Time-Series Models. *Review of the International Statistical Institute*, Vol. 28, No. 3, pp. 233-244.
- [10] Halim. 2006. *Diktat Time Series*. Universitas Kristen Petra. Surabaya.
- [11] Harris, Richard and Robert Sollis. 2003. *Applied Time Series Modelling and Forecasting*. England: John Willey & Sons Ltd.
- [12] Hyndman, Rob J. and Yeasmin Khandakar. 2008. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, July 2008, Vol 27, Issue 3.
- [13] Izenman, A. J. dan Zabell, S. L. 1981. Babies and the blackout : The genesis of a misconception. *Social Science Research*, 10, 282-299.
- [14] Ljung G.M. dan G.E.P. Box. 1978. On a Measure of a Lack of Fit in Time Series Models. *Biometrika*, 65, 297-303.
- [15] Maity, Bipasha dan Bani Chatterjee. 2012. FORECASTING GDP GROWTH RATES OF INDIA: An Empirical Study. *International Journal of Economics and Management Sciences* Vol. 1, No. 9, 2012, pp. 52-58.
- [16] McDowall, David, Richard McCleary and Errol Meidinger. 1980. Interrupted Time Series Analysis. *Sage University Paper series on Quantitative Applications in the Social Sciences*, series no.07-021. Beverly Hills and London: Sage Publications.
- [17] Montgomery, D.C. and Weatherby, G. 1980. Modeling and forecast time series using transfer function and intervention methods, *AIIE Transactions*, 289-307.
- [18] Mulyana. 2004. *Buku Ajar Analisis Deret Waktu*. Universitas Padjajaran FMI-PA Jurusan Statistika. Bandung.

- [19] Quenouille, M.H. 1949. Approximate Tests of Correlation in Time-Series. *Journal of Royal Statistical Society, Series B (Methodological)*, Vol. 11, No. 1, pp. 68-84.
- [20] Rosadi, Dedi. 2006. *Diktat Kuliah Pengantar Analisis Runtun Waktu*. Universitas Gadjah Mada. Yogyakarta
- [21] Sadeq, A. 2008. *Analisis Prediksi Indeks Harga Saham Gabungan dengan Metode ARIMA*. Tesis. Pasca Sarjana UNDIP. Semarang.
- [22] Sridharan, Sanjeev. Suncica Vujic dan Siem Jam Koopman. 2003. Intervention Time Series Analysis of Crime Rates. *Tinbergen Institute Discussion Paper*. Amsterdam.
- [23] Wei, William.W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods Second Edition*. New York: Pearson Addison Wesley.

LAMPIRAN

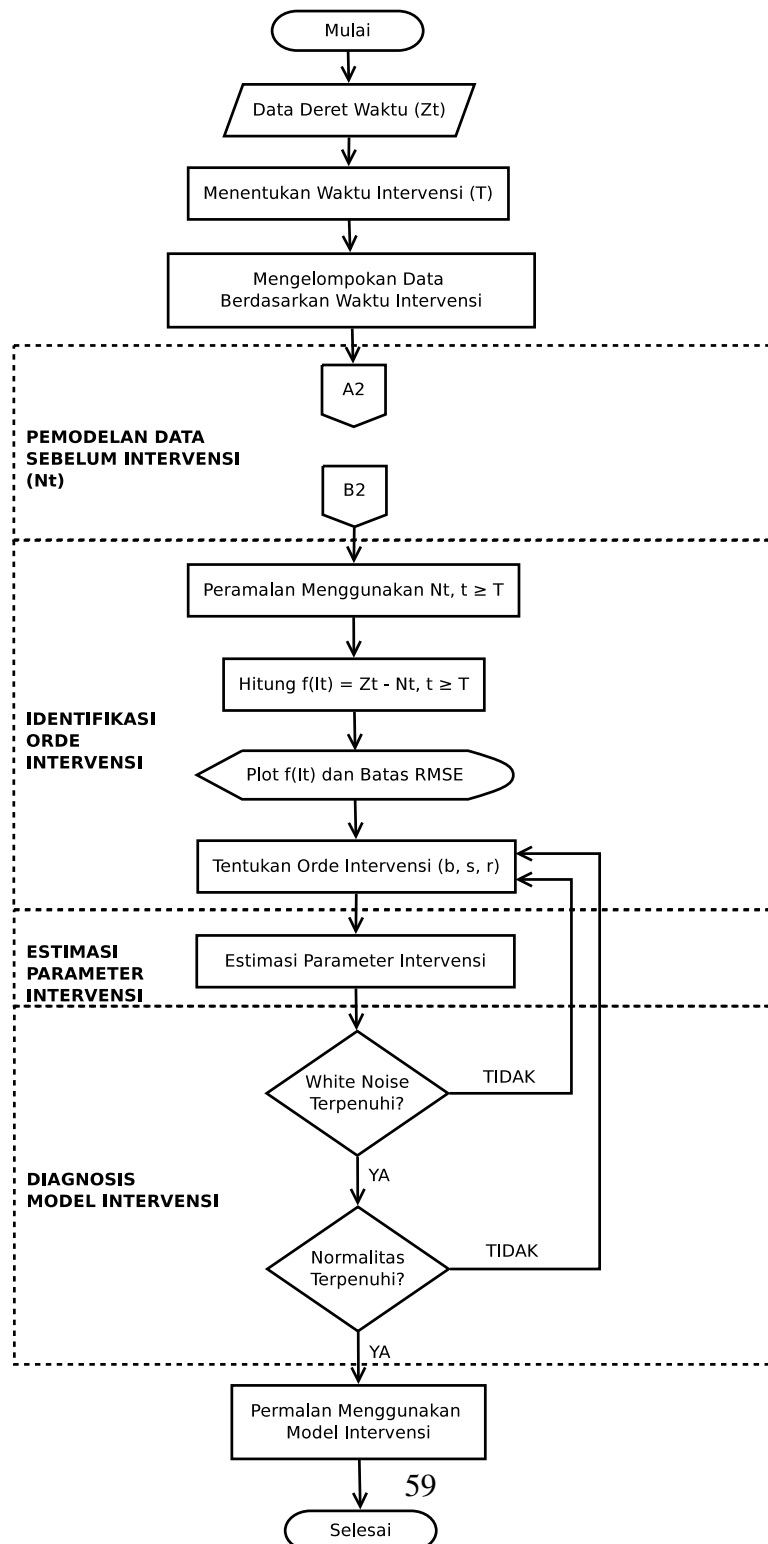
LAMPIRAN A

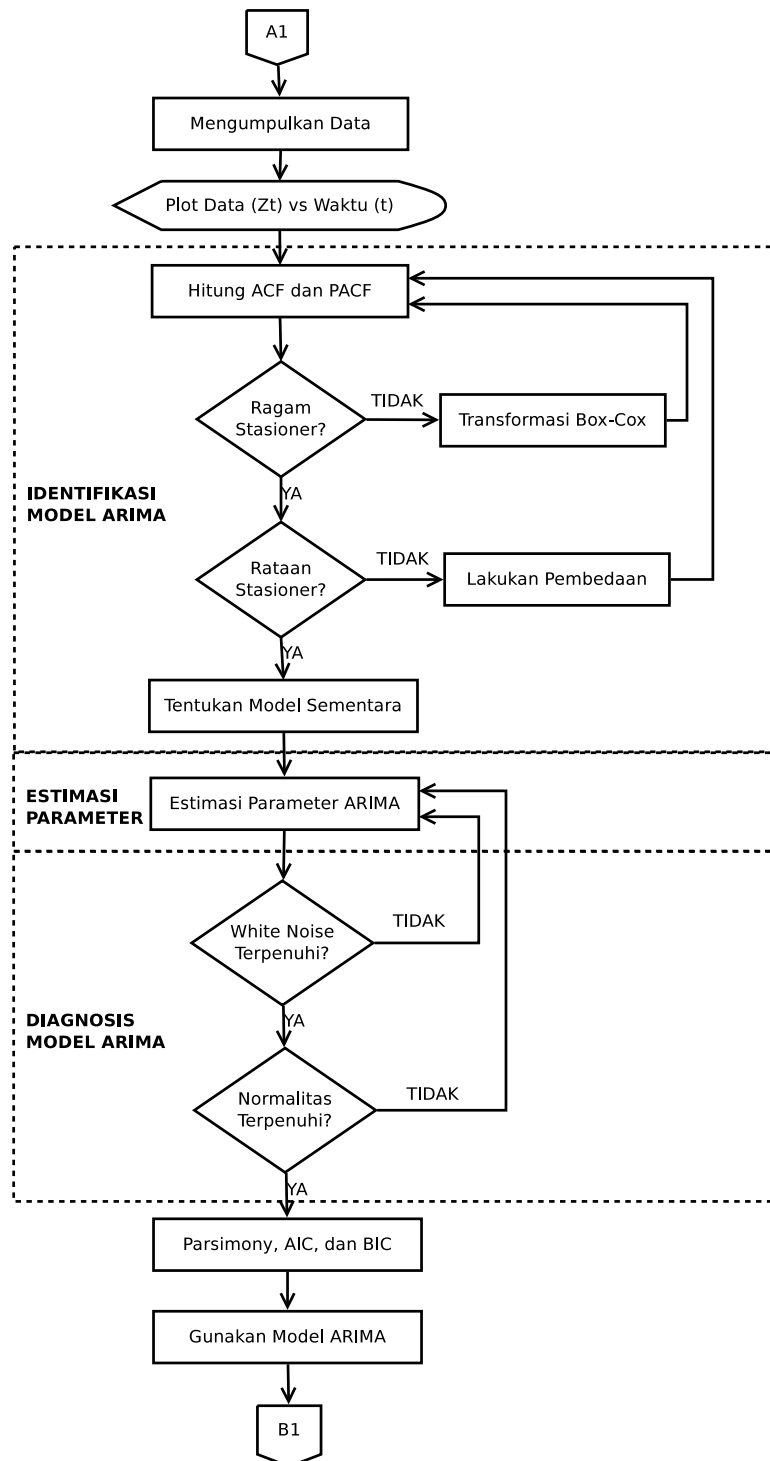
Flowchart Prosedur Pemodelan ARIMA



LAMPIRAN B



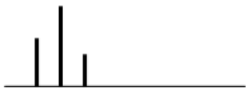
Flowchart Prosedur Pemodelan Intervensi

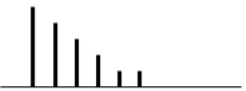

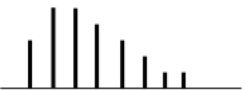







LAMPIRAN C

Ilustrasi Pola Pembobot Impuls untuk $r = 0, 1, 2$

(b, r, s)	Model fungsi transfer	Pola pembobot impuls
(2, 0, 0)	$v(B)X_t = \omega_0 X_{t-2}$	
(2, 0, 1)	$v(B)X_t = (\omega_0 - \omega_1 B) X_{t-2}$	
(2, 0, 2)	$v(B)X_t = (\omega_0 - \omega_1 B - \omega_2 B^2) X_{t-2}$	

(2, 1, 0)	$v(B)X_t = \frac{\omega_0}{(1 - \sigma_1 B)} X_{t-2}$	
(2, 1, 1)	$v(B)X_t = \frac{(\omega_0 - \omega_1 B)}{(1 - \sigma_1 B)} X_{t-2}$	
(2, 1, 2)	$v(B)X_t = \frac{(\omega_0 - \omega_1 B - \omega_2 B^2)}{(1 - \sigma_1 B)} X_{t-2}$	

(2, 2, 0)	$v(B)X_t = \frac{\omega_0}{(1 - \sigma_1 B - \sigma_2 B^2)} X_{t-2}$	
(2, 2, 1)	$v(B)X_t = \frac{(\omega_0 - \omega_1 B)}{(1 - \sigma_1 B - \sigma_2 B^2)} X_{t-2}$	
(2, 2, 2)	$v(B)X_t = \frac{(\omega_0 - \omega_1 B - \omega_2 B^2)}{(1 - \sigma_1 B - \sigma_2 B^2)} X_{t-2}$	

LAMPIRAN D

R Script

```
---
title: "Skripsi-R_LaTeX"
author: "Kurnia Wanto"
date: "July 12, 2016"
geometry: left=3cm, right=2.5cm, top=2.5cm, bottom=2.5cm
output:
  pdf_document:
    fig_height: 3.5
    fig_width: 5
    number_sections: yes
    toc: yes
---

This is an R Markdown document. Markdown is a simple formatting syntax for
authoring HTML, PDF, and MS Word documents. For more details on using R
Markdown see <http://rmarkdown.rstudio.com>.

# Libraries and Import Data

```{r Libraries}
----- Libraries -----

library(forecast) # --- BoxCox Arima auto.arima function is in forecast
package
library(MASS) # --- boxcox function is in MASS package
library(FitAR) # --- LjungBoxTest function is in FitAR package
library(tsoutliers) # --- tso function is in tsoutliers package
library(lmtest) # --- coeftest function is in lmtest package
library(stargazer) # --- stargazer function is in stargazer package
library(TSA) # --- arimax function is in TSA package
```

```{r Import_Data}
```

```

----- 0_Import_Data.R -----

options(width=80)
raw_data <- read.csv("/media/kurnia/Brain/Skripsi/DATA/World\ Bank/inflation.
 csv")
PDB_ID <- ts(raw_data$GDP, start=1970, end=2013)
#PDB_ID <- PDB_ID/1000000000
PDB_ID

stat.desc(rwpdb)

plot(PDB_ID, xlab="Waktu", ylab="US Dollar", col="blue", type="l",
 main="Plot PDB Indonesia 1970-2013")
points(PDB_ID, cex = .5, col = "red")
abline(v=1998, col=1, lty=2)
text(1998, PDB_ID[41], "1998\n(t=29)", pos=2)
'''

Pemodelan ARIMA

Plot N_t
'''{r Plot_Nt}

----- 1-Plot_Nt.R -----

Creating Nt
Nt <- ts(PDB_ID[1:28], start=1970, end=1997)

Creating Layout 1,1;2,3
m <- rbind(c(1, 1), c(2, 3))
layout(m)
par(mar = c(3, 3, 1, 1))

plot(Nt, xlab="Waktu", ylab="US Dollar", col="blue", type="p",
 main="Plot PDB Indonesia 1970-1997 (Nt)")

points(Nt, cex = .5, col = "red")

Trend Linear
t <- 1970:1997
trend_PDB <- glm(Nt~t)
abline(trend_PDB, col="dark blue", lwd=2)

```

```

ACF
acf(Nt, 20, xlim=c(1,20))
text(10,0.8, "ACF")

PACF
pacf(Nt, 20, ylim=c(-.4,1))
text(10,0.8, "PACF")

Show the ACF & PACF value
acf(Nt, 27, plot=FALSE)
pacf(Nt, 27, plot=FALSE)
'''

Box--Cox Transformation
'''{r BoxCox_Transformation}
----- 2-BoxCox_Trans.R -----

t1 <- 1:length(Nt)
t1 <- 1:length(PDB_ID)

--- Search for optimal lambda
par(mar = c(2, 4, 0, 1))
MASS::boxcox(lm(PDB_ID~t1), lambda= seq(-1,1,1/10),
 ylab = "log-Likelihood")

lambda.model <- forecast::BoxCox.lambda(PDB_ID)

--- Box-Cox Transformation
Nt_box <- forecast::BoxCox(Nt, lambda = lambda.model)

--- It is ambiguous, either use lambda for Nt (0.5) or PDB_ID (0.2)
'''

Plot W_t
'''{r Plot_Wt}
----- 3-Plot_Wt.R -----

--- Generating Wt
Wt <- as.numeric(diff(Nt_box))

```

```

--- Creating Layout 1,1;2,3
m <- rbind(c(1, 1), c(2, 3))
layout(m)
par(mar = c(3, 3, 1, 1))

plot(Wt, xlab="Waktu", ylab="US Dollar",
 col="blue", type="o",
 main="Plot Wt",
)
points(Wt, cex = .5, col = "red")

--- Trend Linear
t <- 1:27
trend_PDB <- glm(Wt~t)
abline(h=mean(Wt))

--- ACF
acf(Wt, 20, xlim=c(1,20))
text(10,0.2, "ACF")

--- PACF
pacf(Wt, 20)
text(10,0.2, "PACF")

--- Reset the par
par(mfrow=c(1,1))
'''

Identifikasi Model ARIMA
'''{r Identifikasi_Model_ARIMA}
----- 4-Id_Model.R -----

Save ARIMA model
Nt_arima <- auto.arima(Nt, max.d=2, seasonal=FALSE, lambda=lambda.model,
 stepwise=TRUE, trace=TRUE, max.p=5, max.q=5)
'''

Estimasi Parameter ARIMA
'''{r Estimasi_Parameter_ARIMA}
----- 5-Est_Par.R -----

```

```

--- Creating All Possible Models
model_010 <- Arima(Nt, order=c(0,1,0), lambda = lambda.model, include.drift=
 FALSE)
model_110 <- Arima(Nt, order=c(1,1,0), lambda = lambda.model, include.drift=
 FALSE)
model_011 <- Arima(Nt, order=c(0,1,1), lambda = lambda.model, include.drift=
 FALSE)
model_111 <- Arima(Nt, order=c(1,1,1), lambda = lambda.model, include.drift=
 FALSE)
model_010d <- Arima(Nt, order=c(0,1,0), lambda = lambda.model, include.drift=
 TRUE)
model_110d <- Arima(Nt, order=c(1,1,0), lambda = lambda.model, include.drift=
 TRUE)
model_011d <- Arima(Nt, order=c(0,1,1), lambda = lambda.model, include.drift=
 TRUE)
model_111d <- Arima(Nt, order=c(1,1,1), lambda = lambda.model, include.drift=
 TRUE)

--- Function for rounded value
AIC
raic <- function (model) {
 round(model$aic,2)
}
AICc
raicc <- function (model) {
 round(model$aicc,2)
}
BIC
rbic <- function (model) {
 round(model$bic,2)
}

--- Calculate parameters P-Value
PValue <- function (model) {
 (1-pnorm(abs(model$coef)/sqrt(diag(model$var.coef))))*2
}

```

```

--- Estimating parameters for all possible models:
#
Generating Function
EstimasiParamater <- function (model_010) {
 cat("Parameter:\n")
 print(model_010$coef)
 cat("\nP-Values:\n")
 print(PValue(model_010))
}

--- ARIMA(0,1,0)
EstimasiParamater(model_010)

--- ARIMA(0,1,0) with drift
EstimasiParamater(model_010d)

--- ARIMA(1,1,0)
EstimasiParamater(model_110)

--- ARIMA(1,1,0) with drift
EstimasiParamater(model_110d)

--- ARIMA(0,1,1)
EstimasiParamater(model_011)

--- ARIMA(0,1,1) with drift
EstimasiParamater(model_011d)

--- ARIMA(1,1,1)
EstimasiParamater(model_111)

--- ARIMA(1,1,1) with drift
EstimasiParamater(model_111d)
'''

Diagnosis Model ARIMA
''',{r Diagnosis_Model_ARIMA}
----- 6-Diag_Mod.R -----

--- Ljung-Box Test for Nt
--- OR can use stats::Box.test, lag = min(10,n/5) <-- Rob J. Hyndman

```

```

Box.test(Nt_arima$residuals, lag = round(length(Nt)/5,0),
 type = "Ljung-Box", fitdf = 1)

--- Kolmogorov-Smirnov Test
ks.test(Nt_arima$residuals, "pnorm",
 mean(Nt_arima$residuals),
 sd(Nt_arima$residuals))
'''

Analisis Intervensi
Identifikasi Orde Intervensi
```{r Identifikasi_Model_Intervensi}
# ----- 7-Id_Intv.R -----

# --- Box-Cox Transformation on PDB_ID, not sure using lambda for Nt or PDB
(?)
PDB_box <- forecast::BoxCox(PDB_ID, lambda.model)
Nt_forecast <- forecast(Nt_arima, h = 16)

# --- h-step forecast for Nt
Nt_forecast

# --- Identification intervention order with plot of model residuals
error_idintv <- rep(0,44)
error_idintv[1:28] <- model_011d$residuals
error_idintv[29:44] <- PDB_ID[29:44] - Nt_forecast$mean
plot(error_idintv, type="h", xlab="Waktu (T)", ylab = "Residual", xaxt = "n")
abline(h=c(-3*sd(model_011d$residuals), 3*sd(model_011d$residuals)), col="blue",
        lty=2)
abline(v = 29, col = "red", lty = 3, lwd = 1.5)
text(29, 200, "T=29", cex = .8, pos = 2)
axis(1, at = c(0,10,20,30,40), labels = c("T-29", "T-19", "T-9", "T+1", "T+11"))

error_idintv

# --- Plot PDB_ID vs forecasting of Nt
plot(forecast(model_011d, h=16), main =NA, ylab="US Dollar" )
points(PDB_ID, cex=.5, col="dark red", pch=19)
lines(PDB_ID, col="red")

```

```

abline(v=1998, lty=2)
text(1998, PDB_ID[41], "1998\n(t=29)", pos=2)
legend("topleft", legend = c("PDB_ID", "Peramalan Nt"), cex=0.75, lty=1,
      col=c("dark red", "blue"), pch=c(19,NA))

# --- Detecting Outliers with tsoutliers::tso
PDB_outlier <- tsoutliers::tso(PDB_box, types = c("A0","LS","TC"),
                              maxit.iloop=10, tsmethod = "auto.arima")

plot(PDB_outlier)
'''

## Estimasi Parameter Intervensi
```{r Estimasi_Parameter_Intervensi}
----- 8-Est_Intv.R -----

--- Pulse: Abrupt Temporary
cobalah2 <- TSA::arimax(PDB_box, order = c(0,1,1), xtransf = data.frame(
 T29 = 1*(seq(PDB_ID)==29)), transfer = list(c(1,0)))

cobalah2

--- Test significance of model coefficients
coefptest(cobalah2)
stargazer(coefptest(cobalah2))
'''

Diagnosis Model Intervensi
```{r Diagnosis_Model_Intervensi}
# ----- 9-Diag_Intv.R -----

# --- Generating f(It), based on cobalah2
pulse29 <- filter(1*(seq(PDB_ID)==29), filter = 0.2055,
  method = "rec", sides = 1) * -112.4912

# --- Compute Model for PDB_ID
PDB_arima <- Arima(PDB_ID, lambda = lambda.model, order = c(0,1,1),
  include.constant = TRUE, xreg = pulse29)

# --- Ljung-Box Test for Independence of residuals

```



```

# , lag = min(10,n/5) <-- Rob J. Hyndman
Box.test(PDB_arima$residuals,
         lag=round(length(PDB_ID)/5,0), type = "Ljung-Box", fitdf = 1)

# --- Kolmogorov-Smirnov Test for Normality of residuals
ks.test(PDB_arima$residuals, "pnorm",
        mean(PDB_arima$residuals),
        sd(PDB_arima$residuals))
'''

## Peramalan Model Intervensi
'''{r Peramalan_Model_Intervensi}
# --- Generating f(It)
pulse29 <- filter(1*(seq(PDB_ID)==29),
                 filter = 0.2055, method = "rec",
                 sides = 1)*-112.4912

# --- Compute Model for PDB_ID
PDB_arima <- Arima(PDB_ID, lambda = lambda.model,
                  order = c(0,1,1),
                  include.constant = TRUE, xreg = pulse29)

# --- Future Value of xreg
xreg.rob = forecast(auto.arima(pulse29), h=5)$mean

# --- Forecasting
forecast(PDB_arima, xreg = xreg.rob)
plot(forecast(PDB_arima, xreg = xreg.rob), main=NA)
'''

```

SURAT PERNYATAAN KEASLIAN SKRIPSI

Dengan ini Saya yang bertanda tangan di bawah ini, mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

Nama : Kurnia Wanto
No. Registrasi : 3125090115
Jurusan : Matematika
Program Studi : Matematika

Menyatakan bahwa skripsi ini yang Saya buat dengan judul "**Analisis Intervensi Data Deret Waktu Untuk Peramalan Pendapatan Domestik Bruto Indonesia**" adalah :

1. Dibuat dan diselesaikan oleh Saya sendiri.
2. Bukan merupakan duplikat skripsi yang pernah dibuat oleh orang lain atau jiplakan karya tulis orang lain.

Pernyataan ini dibuat dengan sesungguhnya dan Saya bersedia menanggung segala akibat yang timbul jika pernyataan Saya tidak benar.

Jakarta, 7 Agustus 2016

Yang membuat pernyataan



Kurnia Wanto

DAFTAR RIWAYAT HIDUP

Kurnia Wanto. Lahir di Jakarta, 18 Juni 1991. Anak kedua dari pasangan Bapak Surono dan Ibu Sunarni. Saat ini bertempat tinggal di Jalan M. Saidi Rt. 007 Rw. 006 Petukangan Selatan, Pesanggrahan, Jakarta Selatan 12270.



No. Ponsel : 0856 7724 847

Email : wanto47@gmail.com

Riwayat Pendidikan : Penulis mengawali pendidikan di TK Dewi Sartika dan kemudian melanjutkan pendidikan di SD Negeri 011 Cipulir selama 6 tahun. Tahun 2003, setelah itu melanjutkan ke SMP Negeri 161 Jakarta hingga tahun 2006. Di tahun 2006 penulis melanjutkan ke SMA Negeri 47 Jakarta dan lulus tahun 2009. Di tahun itu pula penulis mengikuti Penelusuran Minat dan Kemampuan (PMDK) dan mendapat Jurusan Matematika, Universitas Negeri Jakarta (UNJ). Pada tahun 2016, penulis telah memperoleh gelar Sarjana Sains untuk Jurusan Matematika, Program Studi Matematika, FMIPA, UNJ.