

## **BAB 2**

### **LANDASAN TEORI**

#### **Definisi Pariwisata dan Wisatawan**

Pariwisata adalah istilah yang diberikan apabila seseorang wisatawan melakukan perjalanan itu sendiri, atau dengan kata lain aktivitas dan kejadian yang terjadi ketika seseorang pengunjung melakukan perjalanan (Sutrisno, 1998, hal: 23). Pariwisata secara singkat dapat dirumuskan sebagai kegiatan dalam masyarakat yang berhubungan dengan wisatawan (Soekadijo, 2000, hal: 2).

Pariwisata merupakan bagian yang tidak terpisahkan dari kehidupan manusia terutama menyangkut kegiatan sosial dan ekonomi. Diawali dari kegiatan yang semula hanya dinikmati oleh segelintir orang-orang yang relatif kaya pada awal abad ke-20, kini telah menjadi bagian dari hak azasi manusia. Hal ini terjadi tidak hanya di negara maju tetapi mulai dirasakan pula di negara berkembang. Indonesia sebagai negara yang sedang berkembang dalam tahap pembangunannya, berusaha membangun industri pariwisata sebagai salah satu cara untuk mencapai neraca perdagangan luar negeri yang berimbang. Melalui industri ini diharapkan pemasukan devisa dapat bertambah (Pendit, 2002). Sebagaimana diketahui bahwa sektor pariwisata di Indonesia masih menduduki peranan yang sangat penting dalam menunjang pembangunan nasional sekaligus merupakan salah satu faktor yang sangat strategis untuk meningkatkan pendapatan masyarakat dan devisa negara.

Menurut Pendit (1994), ada beberapa jenis pariwisata yang sudah dikenal, antara lain:

- a. Wisata budaya, yaitu perjalanan yang dilakukan atas dasar keinginan untuk memperluas pandangan hidup seseorang dengan cara mengadakan kunjungan ke tempat lain atau ke luar negeri, mempelajari keadaan rakyat, kebiasaan dan adat istiadat mereka, cara hidup mereka, kebudayaan dan seni mereka.

- b. Wisata kesehatan, yaitu perjalanan seseorang wisatawan dengan tujuan untuk menukar keadaan dan lingkungan tempat sehari-hari di mana ia tinggal demi kepentingan beristirahat baginya dalam arti jasmani dan rohani.
- c. Wisata olahraga, yaitu wisatawan-wisatawan yang melakukan perjalanan dengan tujuan berolahraga atau memang sengaja bermaksud mengambil bagian aktif dalam pesta olahraga di suatu tempat atau Negara.
- d. Wisata komersial, yaitu termasuk perjalanan untuk mengunjungi pameran-pameran dan pekan raya yang bersifat komersial, seperti pameran industri, pameran dagang dan sebagainya.
- e. Wisata industri, yaitu perjalanan yang dilakukan oleh rombongan pelajar atau mahasiswa, atau orang-orang awam ke suatu kompleks atau daerah perindustrian, dengan maksud dan tujuan untuk mengadakan peninjauan atau penelitian.
- f. Wisata Bahari, yaitu wisata yang banyak dikaitkan dengan danau, pantai atau laut.
- g. Wisata Cagar Alam, yaitu jenis wisata yang biasanya diselenggarakan oleh agen atau biro perjalanan yang mengkhususkan usaha-usaha dengan mengatur wisata ke tempat atau daerah cagar alam, taman lindung, hutan daerah pegunungan dan sebagainya yang kelestariannya dilindungi oleh undang-undang.
- h. Wisata bulan madu, yaitu suatu penyelenggaraan perjalanan bagi pasangan-pasangan pengantin baru yang sedang berbulan madu dengan fasilitas-fasilitas khusus dan tersendiri demi kenikmatan perjalan.

Definisi wisatawan menurut Norval (Yoeti, 1995) adalah setiap orang yang datang dari suatu Negara yang alasannya bukan untuk menetap atau bekerja di situ secara teratur, dan yang di Negara dimana ia tinggal untuk sementara itu membalanjakan uang yang didapatkannya di lain tempat.

Sedangkan menurut Soekadijo (2000), wisatawan adalah pengunjung di Negara yang dikunjunginya setidaknya-tidaknya tinggal 24 jam dan yang datang berdasarkan motivasi:

1. Mengisi waktu senggang atau untuk bersenang-senang, berlibur, untuk alasan kesehatan, studi, keluarga, dan sebagainya.
2. Melakukan perjalanan untuk keperluan bisnis.

3. Melakukan perjalanan untuk mengunjungi pertemuan-pertemuan atau sebagai utusan (ilmiah, administrative, diplomatik, keagamaan, olahraga dan sebagainya).
4. Dalam rangka pelayaran pesiar, jika kalau ia tinggal kurang dari 24 jam.

Menurut Undang-Undang Republik Indonesia no 9 tentang kepariwisataan, Bab I Ketentuan Umum Pasal 1 ayat 1 dan 2 dirumuskan.

- a. Wisata adalah kegiatan perjalanan atau sebagian dari kegiatan tersebut yang dilakukan secara sukarela serta bersifat sementara untuk menikmati objek dan daya tarik wisata.
- b. Wisatawan adalah orang yang melakukan kegiatan wisata.

Berdasarkan sifat perjalanan, lokasi di mana perjalanan dilakukan wisatawan dapat diklasifikasikan sebagai berikut (Karyono, 1997).

- a. *Foreign Tourist* (Wisatawan asing)

Orang asing yang melakukan perjalanan wisata, yang datang memasuki suatu negara lain yang bukan merupakan Negara di mana ia biasanya tinggal. Wisatawan asing disebut juga wisatawan mancanegara atau disingkat wisman.

- b. *Domestic Foreign Tourist*

Orang asing yang berdiam atau bertempat tinggal di suatu negara karena tugas, dan melakukan perjalanan wisata di wilayah negara di mana ia tinggal. Misalnya, staf kedutaan Belanda yang mendapat cuti tahunan, tetapi ia tidak pulang ke Belanda, tetapi melakukan perjalanan wisata di Indonesia (tempat ia bertugas).

- c. *Domestic Tourist* (Wisatawan Nusantara)

Seorang warga negara suatu negara yang melakukan perjalanan wisata dalam batas wilayah negaranya sendiri tanpa melewati perbatasan negaranya. Misalnya warga negara Indonesia yang melakukan perjalanan ke Bali atau ke Danau Toba. Wisatawan ini disingkat wisnus.

- d. *Indigenous Foreign Tourist*

Warga negara suatu negara tertentu, yang karena tugasnya atau jabatannya berada di luar negeri, pulang ke negara asalnya dan melakukan perjalanan wisata di wilayah negaranya sendiri. Misalnya, warga negara Perancis yang bertugas sebagai konsultan di perusahaan asing di Indonesia, ketika liburan ia kembali ke Perancis dan melakukan perjalanan wisata di sana. Jenis wisatawan ini merupakan kebalikan dari *Domestic Foreign Tourist*.

e. *Transit Tourist*

Wisatawan yang sedang melakukan perjalanan ke suatu Negara tertentu yang terpaksa singgah pada suatu pelabuhan/airport/stasiun bukan atas kemauannya sendiri.

f. *Business Tourist*

Orang yang melakukan perjalanan untuk tujuan bisnis bukan wisata tetapi perjalanan wisata akan dilakukannya setelah tujuannya yang utama selesai. Jadi perjalanan wisata merupakan tujuan sekunder, setelah tujuan primer yaitu bisnis selesai dilakukan.

### **2.1.1 Pengembangan Pariwisata**

Perencanaan dan pengembangan pariwisata merupakan suatu proses yang dinamis dan berkelanjutan menuju ketataran nilai yang lebih tinggi dengan cara melakukan penyesuaian dan koreksi berdasar pada hasil monitoring dan evaluasi serta umpan balik implementasi rencana sebelumnya yang merupakan dasar kebijaksanaan dan merupakan misi yang harus dikembangkan. Perencanaan dan pengembangan pariwisata bukanlah system yang berdiri sendiri, melainkan terkait erat dengan sistem perencanaan pembangunan yang lain secara inter sektoral dan inter regional. Perencanaan pariwisata haruslah di dasarkan pada kondisi dan daya dukung dengan maksud menciptakan interaksi jangka panjang yang saling menguntungkan diantara pencapaian tujuan pembangunan pariwisata, peningkatan kesejahteraan masyarakat setempat, dan berkelanjutan daya dukung lingkungan di masa mendatang (Fandeli, 1995). Indonesia sebagai negara yang sedang berkembang dalam tahap pembangunannya, berusaha membangun industri pariwisata sebagai salah satu cara

untuk mencapai neraca perdagangan luar negeri yang berimbang. Pengembangan kepariwisataan saat ini tidak hanya untuk menambah devisa negara maupun pendapatan pemerintah daerah. Akan tetapi juga diharapkan dapat memperluas kesempatan berusaha disamping memberikan lapangan pekerjaan baru untuk mengurangi pengangguran. Pariwisata dapat menaikkan taraf hidup masyarakat yang tinggal di kawasan tujuan wisata tersebut melalui keuntungan secara ekonomi. Dengan mengembangkan fasilitas yang mendukung dan menyediakan fasilitas rekreasi, wisatawan dan penduduk setempat saling diuntungkan. Pengembangan daerah wisata hendaknya memperlihatkan tingkatnya budaya, sejarah dan ekonomi dari tujuan wisata.

Perkiraan jumlah wisatawan asing di Pulau Samosir Sumatera Utara sangat diperlukan dalam rangka perencanaan pengembangan pariwisata nasional dan dapat berfungsi sebagai sarana pemerataan pembangunan di daerah yang sekaligus untuk menciptakan kesempatan berusaha atau kesempatan bekerja serta meningkatkan kesejahteraan bagi masyarakat sekitar sehingga mereka tidak tertinggal oleh perkembangan usaha jasa dan sarana pariwisata.

### **2.1.2 Pengaruh Wisata dan Ekonomi**

Pariwisata bukan saja sebagai sumber devisa, tetapi juga merupakan faktor dalam menentukan lokasi industri dalam perkembangan daerah-daerah yang miskin sumber-sumber alam sehingga perkembangan pariwisata adalah salah satu cara untuk memajukan ekonomi di daerah-daerah yang kurang berkembang tersebut sebagai akibat kurangnya sumber-sumber alam (Yoeti, 1997). Gunn (1988), mendefinisikan pariwisata sebagai aktivitas ekonomi yang harus dilihat dari dua sisi yakni sisi permintaan (demand side) dan sisi pasokan (supply side). Lebih lanjut dia mengemukakan bahwa keberhasilan dalam pengembangan pariwisata di suatu daerah sangat tergantung kepada kemampuan perencana dalam mengintegrasikan kedua sisi tersebut secara berimbang ke dalam sebuah rencana pengembangan pariwisata. Menurut Robert (Toety, 1990). Kelincahan dalam berusaha harus dilakukan agar pendapatan selama musim kedatangan wisatawan bisa menjadi penyeimbang bagi

musim sepi wisatawan. Pengaruh yang ditimbulkan oleh pariwisata terhadap ekonomi ada dua ciri, pertama produk pariwisata tidak dapat disimpan, kedua permintaannya sangat tergantung pada musim, berarti pada bulan tertentu ada aktivitas yang tinggi, sementara pada bulan-bulan yang lain hanya ada sedikit kegiatan.

### **2.1.3 Potensi Wisata Pulau Samosir**

Pulau Samosir adalah pulau yang berada di tengah-tengah Danau Toba di Sumatera Utara. Suatu pulau dengan ketinggian 1.000 meter di atas permukaan laut. Samosir menjadi kabupaten pada Januari 2004. Sebelumnya, pulau ini merupakan bagian dari Kabupaten Toba Samosir. Terdiri atas 9 kecamatan, yaitu Pangururan, Harian, Sianjur Mulamula, Nainggolan, Onan Runggu, Palipi, Ronggur Nihuta, Simanindo, dan Sitio-Tio. Masing-masing kecamatan memiliki objek wisata, yang bila dikelola dengan baik akan mendatangkan nilai tambah bagi pulau yang berpenduduk 131.000 jiwa. Namun, sayang potensi wisata itu belum dikelola maksimal. Sehingga wisatawan asing maupun domestik kurang tertarik untuk berkunjung.

Pulau Samosir diyakini sebagai daerah asal orang Batak. Pasalnya, di pulau ini tepatnya di Pusuk Buhit Kecamatan Sianjur Mulamula merupakan asal orang Batak. Pusuk Buhit merupakan perbukitan dengan ketinggian lebih dari 1.800 meter di atas permukaan Danau Toba. Pulau Samosir memiliki banyak objek wisata yang dapat dikunjungi seperti wisata alam, wisata seni budaya dan wisata peninggalan sejarah yang sangat berpeluang untuk memberikan kontribusi terhadap Pendapatan Asli Daerah (PAD).

Adapun objek wisata yang menjadi unggulan Danau Toba di Pulau Samosir saat ini ialah:

1. Kawasan hotel dan restoran di Tuktuk Siadong, Kec. Simanindo
2. Makam Tua Raja Sidabutar di Tomok, Kec. Simanindo
3. Pertunjukan Sigale-gale di Tomok, Kec. Simanindo
4. Batu kursi persidangan di Siallagan, Kec. Simanindo
5. Museum Hutabolon di Simanindo, Kec. Pangururan
6. Pemandian Air Panas, Kec. Pangururan
7. Pemandian Aek Sipitu Dai, Kec. Sianjur Mula-mula

## 8. Perkampungan Si Raja Batak di Sigulatti, Kec. Simanindo

Dan masih banyak obyek wisata Pulau Samosir yang berpotensi layak dikunjungi dan masih dalam tahap pengembangan.

### 2.2 PERAMALAN

Ramalan merupakan dugaan atau pikiran mengenai terjadinya kejadian atau peristiwa dari waktu yang akan datang (Supranto, 1993). Peramalan dapat juga diartikan sebagai suatu usaha memperkirakan perubahan, agar tidak disalahpahami bahwa peramalan tidak memberikan jawaban yang pasti tentang apa yang akan terjadi, melainkan akan mencari yang sedekat mungkin dengan apa yang akan terjadi (Mulyono, 2000). *Forecasting* adalah peramalan yang akan terjadi pada masa yang akan datang, sedangkan rencana merupakan penentuan apa yang akan dilakukan pada waktu yang akan dilakukan pada waktu yang akan datang (Pangestu, 1986).

Suatu keputusan akan lebih baik hasilnya bila memadukan antara hasil peramalan dan intuisi. Arsham (1994) menyatakan bahwa peramalan adalah bagian yang integral dari kegiatan pengambilan keputusan manajemen. Hampir setiap organisasi memerlukan ramalan baik secara eksplisit maupun secara implisit, karena hampir setiap organisasi harus membuat perencanaan agar sesuai dengan kondisi masa depan yang tidak diketahui dengan baik. Selain itu, peramalan dibutuhkan pada semua lini fungsional, begitu pula pada semua jenis organisasi. Peramalan dibutuhkan dalam bidang keuangan, pemasaran, personalia, dan lingkup produksi, dalam pemerintahan dan organisasi pencari laba, dalam klub sosial kecil, dan dalam partai politik nasional. (Hanke *et al*, 2003).

#### 2.2.1 Tujuan Peramalan

Peramalan dan perencanaan adalah dua kegiatan yang saling melengkapi, dan bukan substitusi satu dengan yang lain (Santoso, 2009). Peramalan dan rencana mempunyai

hubungan yang cukup erat, karena rencana itu disusun berdasarkan ramalan yang dimungkinkan terjadi di masa mendatang. Dalam beberapa hal terutama dalam ilmu social ekonomi, sering terkait dengan sesuatu yang serba tidak pasti dan sukar untuk diperkirakan secara tepat, oleh karena itu dalam hal ini kita membutuhkan adanya perkiraan atau peramalan.

Dalam membuat ramalan diupayakan untuk dapat meminimumkan pengaruh ketidakpastian, dengan kata lain *forecast* bertujuan mendapatkan ramalan yang bisa meminimumkan *error* (kesalahan) meramal. *Error* yang digunakan sebagai ukuran akurasi peramalan dapat berupa ME (mean error), MAE (mean absolute error), MSE (mean absolute error), MPE (mean percentage error), dan MAPE (mean absolute percentage error). Semakin besar ukuran itu berarti semakin tidak akurat peramalannya. Sebaliknya, semakin kecil ukuran itu berarti semakin akurat peramalannya. (Aritonang, 2002, hal: 35).

## **2.2.2 Metode Peramalan**

Metode peramalan dapat berdasarkan pengalaman, penilaian, opini dari ahli atau model matematika yang menggambarkan pola data historis. Peramalan tidak dapat dihindarkan karena sifat dari investasi yang mempunyai umur beberapa tahun. Suatu rencana investasi yang didasarkan pada suatu peramalan yang salah, sudah pasti akan membawa kerugian, bukan saja pada saat investasi itu harus dibuat tapi juga selama umur investasi itu. Metode peramalan terbagi menjadi dua yaitu metode peramalan kualitatif dan metode kuantitatif.

### **2.2.2.1 Metode Kualitatif**

Peramalan kualitatif pada hakekatnya didasarkan pada intuisi atau pengalaman empiris dari perencana atau pengambil keputusan, sehingga relative lebih subjektif. Pada situasi manajemen dan industri yang masih sederhana, peramalan kualitatif dapat memberikan akurasi hasil peramalan yang relatif sama dengan peramalan kuantitatif.



Metode kualitatif membutuhkan input yang tergantung pada metode tertentu dan biasanya merupakan hasil dari pemikiran intuitif, pertimbangan dan pengetahuan yang didapat. Pendekatan dengan metode ini seringkali memerlukan input dari sejumlah orang yang telah terlatih secara khusus. (Makridakis *et al*, 1999).

### **2.2.2.2 Metode Kuantitatif**

Peramalan kuantitatif memiliki sifat yang objektif berdasarkan pada keadaan aktual yang diolah dengan menggunakan metode-metode tertentu (Makridakis *et al*, 1999). Penggunaan suatu metode juga harus didasarkan pada fenomena manajemen atau bisnis apa yang akan diramalkan dan tujuan yang ingin dicapai melalui peramalan. Peramalan kuantitatif dapat diterapkan bila terdapat tiga kondisi berikut:

1. Tersedia informasi masa lalu.
2. Informasi tersebut dapat dikuantitatifkan dalam bentuk data numerik.
3. Dapat diasumsikan bahwa pola masa lalu akan terus berlanjut dimasa datang.

Pada dua asumsi pertama merupakan syarat keharusan bagi penerapan metode peramalan kuantitatif. Asumsi ketiga merupakan syarat kecukupan, artinya walaupun asumsi ketiga dilanggar, model yang dirumuskan masih dapat digunakan. Hal tersebut akan memberikan kesalahan peramalan yang relatif besar bila perubahan pola data maupun bentuk hubungan fungsional tersebut terjadi secara sistematis. Metode kuantitatif dapat dibagi dalam dua kelompok, yaitu metode-metode peramalan dengan analisis deret berkala dan metode kausal.

Metode-metode peramalan dengan analisa deret berkala, yaitu:

1. Metode Pemulusan Eksponensial dan Rata-rata Bergerak.
2. Metode Proyeksi Trend dengan Regresi.
3. Metode Box Jenkis.

Metode-metode kausal, yaitu :

1. Metode Regresi dan Korelasi.
2. Metode Ekonometri.
3. Model Input dan Output.

### 2.2.3 Metode Deret Berkala

Runtun waktu adalah serangkaian pengamatan terhadap suatu peristiwa, kejadian, gejala atau variabel yang diambil dari waktu ke waktu, dicatat secara teliti menurut urutan waktu terjadinya dan kemudian disusun sebagai data statistik (Hadi, 1968). Pada umumnya pencatatan ini dilakukan dalam jangka waktu tertentu misalnya satu bulan, satu tahun, sepuluh tahunan dan sebagainya.

Analisis runtun waktu adalah suatu metode kuantitatif untuk menentukan pola data masa lampau yang telah dikumpulkan secara teratur. Secara umum analisis deret waktu mempunyai beberapa tujuan, yaitu peramalan, pemodelan, dan kontrol (Chatfield, 2001). Peramalan berkaitan dengan pembentukan model dan metode yang dapat digunakan untuk menghasilkan suatu ramalan yang akurat. Sebuah model deret berkala adalah suatu fungsi yang menghubungkan nilai deret berkala dengan nilai awal deret berkala, kesalahannya, atau yang berhubungan dengan deret berkala lainnya. Metode deret berkala meramalkan sifatnya untuk masa depan. Jika ada persamaan yang ditentukan maka sifat sistem dapat diketahui dengan menyelesaikan persamaan tersebut yang kondisi awalnya sudah diketahui. Pada peramalan runtun waktu, persamaan dan kondisi awal mungkin diketahui kedua-duanya atau mungkin hanya salah satu saja. Oleh karena itu diperlukan suatu aturan untuk menentukan perkembangan dan keakuratan sistem. Penentuan aturan tersebut mungkin mengacu dari pencocokkan data masa lalu.

Menurut Hanke *et al.* (2003), faktor utama yang mempengaruhi pemilihan teknik peramalan untuk data deret waktu adalah identifikasi dan pemahaman pola historis data. Pola data tersebut terbagi menjadi empat, yaitu:

1. Pola Horisontal

Pola ini terjadi pada saat data observasi berfluktuasi disekitar nilai rata-rata konstan. Pola ini disebut juga pola stasioner.

2. Pola Trend

Pola ini muncul ketika observasi data menaik atau menurun pada periode yang panjang. Contoh dari rangkaian trend adalah pertumbuhan populasi, inflasi harga, perubahan teknologi, preferensi konsumen dan kenaikan produktifitas.

### 3. Pola Siklis (cyclus)

Pola ini muncul pada saat observasi data memperlihatkan kenaikan dan penurunan pada periode yang tidak tetap. Komponen siklik mirip fluktuasi gelombang disekitar *trend* yang sering dipengaruhi oleh kondisi ekonomi.

### 4. Pola Musiman (seasonality)

Pola terjadi pada saat data observasi dipengaruhi oleh faktor musiman. Komponen musiman mengacu pada suatu pola perubahan yang berulang dengan sendirinya dari tahun ketahun. Untuk deret bulanan, komponen musiman mengukur keragaman deret dari setiap Januari, setiap Februari dan seterusnya. Untuk deret triwulanan, ada empat elemen musim, masing-masing satu untuk setiap triwulan.

## 2.2.4 Komponen Deret Berkala (Time Series)

Pada umumnya, time series adalah hasil pekerjaan dari empat gerak, yang disebut juga komponen-komponen dari *time series* itu. Diterangkan secara singkat keempat komponen itu sebagai persiapan kepada uraian yang lebih panjang. Keempat komponen itu antara lain:

#### a. Gerak Jangka Panjang (Longterm Movement atau Seculer Trend)

Yaitu suatu gerak yang menunjukkan ke arah mana tujuan dari time series itu mungkin saja berbentuk garis lurus atau garis lengkung. Cara menentukan *trend*, yaitu:

1. Dengan memakai tangan saja
2. Cara semi average
3. Cara rata – rata bergerak (Moving Average)
4. Cara least square

b. Gerak Berulang (Cyclical movement)

Yaitu gerak naik turun yang terjadi dalam jangka waktu yang lama. Gerak seperti ini terjadi dengan teratur atau hampir teratur, akan tetapi mungkin juga amplitudonya berbeda dari waktu ke dalam jangka waktu yang singkat (bagian dari tahun atau musim). Oleh karena gerak ini hampir teratur atau benar-benar teratur, maka gerak ini disebut dengan gerak periodik.

c. Gerak Tak Teratur (Irregular Movement)

Yaitu gerak yang terjadi hanya sekali kali dan tidak mengikuti aturan tertentu dan karenanya tidak dapat diramalkan lebih dahulu. Gerak ini biasanya tidak berpengaruh untuk jangka waktu yang lama, walaupun kekecualian selalu ada.

Teknik peramalan yang digunakan dalam peramalan *time series* terdiri dari beberapa model. Pembagian model tersebut beragam menurut para ahli, namun pada dasarnya memiliki maksud dan tujuan yang sama. Model-model peramalan *time series* tersebut, adalah sebagai berikut:

1. Model Trend
2. Model *Naive* (naif)
3. Model rata-rata
4. Model *Exponential Smoothing* (pemulusan eksponensial)
5. Model *Box-Jenkins* (ARIMA)
6. Model Dekomposisi

Dalam penelitian ini, metode peramalan yang akan digunakan adalah ARIMA karena dianggap sebagai metode yang dapat digunakan sebagai metode peramalan untuk kasus data kunjungan wisata dan metode ini cukup memungkinkan untuk mengatasi kasus data *time series* yang tidak stasioner. Beberapa penelitian terdahulu memberikan informasi mengenai metode yang dapat diterapkan untuk meramalkan jumlah kunjungan wisata. *Naive Models*, *Exponential Smoothing* dan *Seasonal ARIMA* merupakan metode yang pernah digunakan oleh Wang dan Lim (2005) untuk meramalkan kunjungan wisatawan asal Australia ke Jepang. Sedangkan Athanasopoulos dan Hyndman (2007) memanfaatkan *Time Series Regression*,

*Exponential Smoothing* dan *Innovations state space models* dalam meramalkan jumlah kunjungan wisata ke Australia.

Pada model ARIMA diperlukan penetapan karakteristik data deret berkala seperti stasioner, musiman dan sebagainya, yang memerlukan suatu pendekatan sistematis untuk mendapatkan gambaran yang jelas mengenai model-model dasar yang akan ditangani. Hal utama yang mencirikan dari model ARIMA dalam analisis data deret waktu dibandingkan metode pemulusan adalah perlunya pemeriksaan keacakan data dengan melihat koefisien autokorelasinya. Model ARIMA juga bisa digunakan untuk mengatasi masalah sifat keacakan, trend, musiman bahkan sifat siklis data data deret waktu yang dianalisis.

### **2.3 Metode ARIMA (Autoregressive Integrated Moving average)**

Model *Autoregressive Intrgrated Moving Average* (ARIMA) merupakan metode yang secara intensif dikembangkan oleh George Box dan Gwilym Jenkins. Metode ARIMA berbeda dengan metode peramalan lain karena metode ini tidak mensyaratkan suatu pola data tertentu supaya model dapat bekerja dengan baik. Metode ARIMA akan bekerja dengan baik apabila data deret berkala yang dipergunakan bersifat dependen atau berhubungan satu sama lain secara statistik.

ARIMA merupakan suatu metode yang menghasilkan ramalan-ramalan berdasarkan sintesis dari pola data secara historis (Arsyad, 1995). Metode ARIMA merupakan uji linear yang istimewa. ARIMA sama sekali mengabaikan variabel independen karena model ini menggunakan nilai sekarang dan nilai-nilai lampau dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat.

Arsyad (1995) juga menyebutkan bahwa metodologi Box-Jenkins ini dapat digunakan:

1. Untuk meramal tingkat *employment*.
2. Menganalisis pengaruh promosi terhadap penjualan barang-barang konsumsi.

3. Menganalisis persaingan antara jalur kereta api dengan jalur pesawat terbang.
4. Mengestimasi perubahan struktur harga suatu industri.

Hasil para peneliti terdahulu mengenai ARIMA dapat disimpulkan sebagai berikut:

1. ARIMA merupakan suatu metode yang menghasilkan ramalan-ramalan berdasarkan sintesis dari pola data secara historis (Arsyad, 1995). Variabel yang digunakan adalah nilai-nilai terdahulu bersama nilai kesalahannya.
2. ARIMA memiliki tingkat keakuratan peramalan yang cukup tinggi karena setelah mengalami tingkat pengukuran kesalahan peramalan MAE (*mean absolute error*) nilainya mendekati nol (Francis dan Hare, 1994).
3. ARIMA mempunyai tingkat keakuratan peramalan sebesar 83.33% dibanding model logit 66.37% dan OLS 58.33% (Dunis, 2002).
4. Menurut penelitian Mulyono (2000), teknik Box-Jenkins menunjukkan bahwa metode ini cocok untuk meramal sejumlah variabel dengan cepat, sederhana, dan murah karena hanya membutuhkan data variabel yang akan diramal.

Secara umum model ARIMA dirumuskan dengan notasi sebagai berikut.

ARIMA (p, d, q) dengan,

*p* menunjukkan orde atau derajat *autoregressive* (AR)

*d* menunjukkan orde atau derajat *differencing* (pembedaan) dan

*q* menunjukkan orde atau derajat *moving average* (MA).

### 2.3.1 Model AR (ARIMA (p,0,0))

Kasus yang paling sering dihadapi dalam proses *autoregressive* adalah apabila  $p = 1$  dan  $p = 2$ , yaitu berturut-turut untuk model AR(1) dan AR(2) yang didefinisikan sebagai berikut (Makridakis *et al*, 1999):

ARIMA(1,0,0)

$$X_t = \mu + \phi_1 X_{t-1} + e_t \quad (2.15)$$

atau

$$(1 - \phi_1 B)X_t = \mu' + e_t \quad (2.16)$$

ARIMA(2,0,0)

$$X_t = \mu' + \phi_1 X_{t-1} + \phi_2 X_{t-2} + e_t \quad (2.17)$$

atau

$$(1 - \phi_1 B - \phi_2 B^2)X_t = \mu' + e_t \quad (2.18)$$

### 2.3.2 Model MA (ARIM (0.0.q))

Dalam praktek, kasus yang kemungkinan besar muncul adalah apabila  $q = 1$  dan  $q = 2$  yaitu berturut-turut proses MA(1) dan MA(2). Dua kasus ini dapat ditulis dengan persamaan berikut (Makridakis *et al*, 1999):

ARIMA (0,0,1)

$$X_t = \mu + (1 - \theta_1 B)e_t \quad (2.19)$$

ARIMA (0,0,2)

$$X_t = \mu + (1 - \theta_1 B - \theta_2 B^2)e_t \quad (2.20)$$

### 2.3.3 Model ARIMA (p,d,q)

Secara umum model ARIMA dapat dituliskan sebagai berikut.

$$X_t = \phi_1 \mu + \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (2.21)$$

Dengan operator *shift* mundur dapat ditulis sebagai berikut.

$$\phi(B)X_t = \theta(B)e_t \Leftrightarrow (1 - \phi_1 B + \dots + \phi_p B^p)X_t = (1 - \theta_1 B - \dots - \theta_q B^q)e_t \quad (2.22)$$

Model ARIMA menggunakan nilai sebelumnya dari variabel independen maupun nilai residual periode sebelumnya. Dengan penggabungan ini diharapkan model ARIMA dapat mengakomodasi pola data yang tidak dapat diidentifikasi secara sendiri-sendiri oleh model MA atau AR. Orde dari model ARIMA ditentukan oleh

jumlah periode variabel independen baik dari nilai sebelumnya dari variabel independen maupun nilai residual periode sebelumnya.

### 2.3.4 Model ARIMA dan Musiman

Musiman berarti kecendrungan mengulangi pola tingkah gerak dalam periode musim, biasanya satu tahun untuk data bulanan. Karena itu, deret waktu musiman mempunyai karakteristik yang ditunjukkan oleh adanya korelasi beruntun yang kuat pada jarak semusim (Cryer,1986). Untuk data yang stasioner, faktor musiman dapat ditentukan dengan mengidentifikasi koefisien autokorelasi pada dua atau tiga *time-lag* yang berbeda nyata dari nol. Autokorelasi yang secara signifikan berbeda dari nol menyatakan adanya suatu pola dalam data. Notasi umum yang singkat untuk menangani musiman adalah

ARIMA (p,d,q)(P,D,Q)<sup>S</sup>

dengan,

(p,d,q) = bagian yang tidak musiman dari model

(P,D,Q) = bagian musiman dari model

S = jumlah periode per musim

### 2.3.5 Model AR Musiman

Bentuk umum dari proses *autoregresif* musiman periode  $S$  tingkat  $P$  atau AR(P)<sup>S</sup> didefinisikan sebagai berikut.

$$X_t = \Theta_1 X_{t-s} + \Theta_2 X_{t-2s} + \dots + \Theta_p X_{t-ps} + \alpha'_t \quad (2.23)$$

Persamaan diatas dapat juga ditulis dalam bentuk:

$$\Phi(B)X_t = \alpha_t \quad (2.24)$$



dengan  $\Phi(B) = 1 - \Phi_1 B^S - \Phi_2 B^{2S}$  yang dikenal sebagai operator  $AR(P)^S$  dimana  $\alpha_t$  adalah independen dan berdistribusi normal dengan *mean* 0 dan *varians*  $\sigma_a^2$ . Sebagai contoh model  $AR(p)^S$  akan dijelaskan model  $AR(1)^{12}$  yakni suatu proses  $\{X_t\}$  dikatakan mengikuti model  $AR(1)^{12}$  jika mengikuti model

$$X_t = \Phi_1 X_{t-12} + \alpha_t \quad (2.25)$$

dengan  $E(X_t) = 0$ , untuk semua  $k$  diperoleh

$$E(X_t X_{t-k}) = E[(\Phi_1 X_{t-12} + \alpha_t)(\alpha_1 X_{t-12} + \alpha_{t-k})] \quad (2.26)$$

Dengan membagi persamaan diatas dengan  $\gamma_0$  akan diperoleh

$$\rho_k = \Phi_1 \rho_{k-12}, \quad \text{untuk } k \geq 1 \quad (2.27)$$

Jelas bahwa

$$\rho_{12} = \Phi_1 \rho_0 \text{ dan } \rho_{22} = \Phi_1 \rho_{12} = \Phi_1^2$$

Sehingga secara umum akan diperoleh

$$\rho_{12k} = \Phi_1^k, \text{ untuk } k = 1, 2, \dots$$

Selanjutnya, untuk  $k = 1$  dan  $k = 11$  dengan  $\rho_k = -\rho_k$  akan diperoleh

$$\rho_1 = \Phi_1 \rho_{11} \text{ dan } \rho_{11} = \Phi_1 \rho_1 \quad (2.28)$$

Yang berimplikasi bahwa  $\rho_1 = \rho_{11} = 0$

### 2.3.6 Model MA Musiman

Bentuk umum dari proses *moving average* musiman periode  $S$  tingkat  $Q$  atau  $MA(Q)^S$  didefinisikan sebagai berikut.

$$X_t = \alpha_t - \Theta_1 \alpha_{t-S} - \Theta_2 \alpha_{t-2S} - \dots - \Theta_Q \alpha_t - Q_S \quad (2.29)$$

Persamaan (2.29) dapat juga ditulis dalam bentuk:

$$X_t = \Theta(B) \alpha_t \quad (2.30)$$

dengan,

$$\Theta(B) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \quad (2.31)$$

yang dikenal sebagai operator  $MA(Q)^S$ , dengan  $\alpha_t$  adalah independen dan berdistribusi normal dengan *mean* 0 dan *varians*  $\sigma_a^2$ . Sebagai contoh dari model  $MA(Q)^S$  akan dijelaskan model  $MA(1)^{12}$ . Suatu proses  $X_t$  dikatakan mengikuti model  $MA(1)^{12}$  jika  $\{X_t\}$  mengikuti model:

$$X_t = \alpha_t - \Theta_1 \alpha_{t-12} \quad (2.31)$$

Jelas bahwa  $X_t$ , yaitu  $E(X_t) = 0$  dan untuk semua  $k$

$$E(X_t X_{t-k}) = E[(\alpha_t - \Theta_1 \alpha_{t-12})(\alpha_{t-k} - \Theta_1 \alpha_{t-k-12})] \quad (2.32)$$

Dalam hal ini  $E(X_t X_{t-k}) = 0$  untuk  $k \neq 12$ , yang berarti proses tidak mempunyai korelasi diluar lag 12.

Sebagai ringkasan, untuk suatu series yang mengikuti proses  $MA(1)^{12}$ , maka

$$\gamma_0 = \text{var}(X_t) = \sigma_a^2 (1 + \Theta_1^2) \quad (2.33)$$

$$\gamma_{12} = -\Theta_1 \sigma_a^2$$

$$\rho_{12} = \frac{-\Theta_1}{1 + \Theta_1^2}$$

dan  $\gamma_k = \rho_k = 0$  untuk  $k \neq 12$

## 2.3.7 Analisis Model ARIMA

### 2.3.7.1 Plot Data

Langkah pertama yang baik untuk menganalisis data deret berkala adalah membuat plot data tersebut secara grafis. Hal ini bermanfaat untuk memplot berbagai data *moving average* untuk menetapkan adanya *trend* (penyimpangan nilai tengah) dan adanya pengaruh musiman pada data (*deseasonalize the data*).

### 2.3.7.2 Koefisien Autokorelasi

Kovariansi dan korelasi antara  $X_t$  dan  $X_{t+k}$  berturut-turut dapat dituliskan sebagai berikut.

$$\gamma_k = \text{Cov}(X_t, X_{t+k}) = E(X_t - \mu, X_{t+k} - \mu) \quad (2.1)$$

dan

$$\rho_k = \frac{\text{kov}(X_t, X_{t+k})}{\sqrt{\text{var}(X_t)}\sqrt{\text{var}(X_{t+k})}} = \frac{\gamma_k}{\gamma_0}, \quad (2.2)$$

Dengan:

$X_t$  = pengamatan

$X_{t+k}$  = sejumlah  $k$  pengamatan setelahnya

Dari suatu runtun waktu stasioner  $X_1, X_2, \dots, X_n$  maka dapat diestimasi fungsi autokorelasi dengan menggunakan persamaan:

$$\hat{\rho} = r_k = \frac{\hat{\gamma}_k \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\hat{\gamma}_0 \sum_{t=1}^n (X_t - \bar{X})^2} \quad (2.3)$$

Untuk proses Gaussian, rumus pendekatan kovariannya untuk  $k > 0$  dan  $k+j > 0$  (Wei, 1990)

$$\text{kov}(\hat{\rho}_k, \hat{\rho}_{k+j}) \cong \frac{1}{n} \sum (\rho_i \rho_{i+j} + \rho_{i+k+j} \rho_{i-k} - 2\rho_k \rho_i \rho_{i-k-j} - 2\rho_{k+j} \rho_i \rho_{i-k} + 2\rho_k \rho_{k+j} \rho_i^2) \quad (2.4)$$

Untuk  $n$  besar,  $\hat{\rho}_k$  mendekati distribusi normal dengan rata-rata  $\rho_k$  dan varian:

$$\text{var}(\hat{\rho}_k) \cong \frac{1}{n} \sum_{i=-\infty}^{\infty} (\rho_i^2 + \rho_{i+k} \rho_{i-k} - 4\rho_k \rho_i \rho_{i-k} + 2\rho_k^2 \rho_i^2) \quad (2.5)$$

Untuk proses  $\rho_k = 0$  dengan  $k > m$ , pendekatan terhadap persamaan di atas adalah sebagai berikut:

$$\text{var}(\hat{\rho}_k) \cong \frac{1}{n} \sum_{i=-\infty}^{\infty} (1 + 2\rho_i^2 + 2\rho_2^2 + \dots + 2\rho_m^2) \quad (2.6)$$

Dalam kenyataannya  $\rho_1$  tidak diketahui maka digunakan  $\hat{\rho}_k$  ( $i = 1, 2, \dots, m$ ).

Standar error  $\hat{\rho}_k$  untuk lag besar adalah:

$$SE \hat{\rho}_k = \sqrt{\frac{1}{n}(1 + 2\hat{\rho}_1^2 + \dots + 2\hat{\rho}_m^2)} \quad (2.7)$$

dengan  $k = 0, 1, 2, \dots$

### 2.3.7.3 Koefisien Autokorelasi Parsial

Misalkan terdapat model regresi dimana variabel dependen  $X_{t+k}$  dari proses stasioner diregresikan pada k-lag variabel  $X_{t+k-1}, X_{t+k-2}, \dots$  dan  $X_t$ , yaitu:

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

dengan  $\phi_{k1}$  merupakan parameter regresi dan  $e_{t+k}$  adalah suku sesatan normal yang tidak berkorelasi dengan  $X_{t+k-j}$  untuk  $j \geq 1$  (Wei, 1990). Dengan menggunakan aturan *cramer*, untuk  $k = 1, 2, \dots$ , diperoleh

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

$$\phi_{11} = \frac{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}}$$

$$\phi_{33} = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}}$$

$$\phi_{kk} = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{bmatrix}}$$

Dengan  $\phi_{kk}$  adalah autokorelasi parsial antara  $X_t$  dan  $X_{t+k}$ . Karena merupakan fungsi dari  $k$ , himpunan  $\{\phi_{kk}; k = 0,1\}$  dinamakan fungsi autokorelasi parsial (partial autocorrelation function), disingkat dengan PACF.

### 2.3.7.4 Stasioner dan Nonstasioner

Hal pertama yang harus diperhatikan adalah bahwa kebanyakan deret berkala bersifat nonstasioner. Stasioneritas berarti bahwa tidak terdapat pertumbuhan atau penurunan pada data. Data secara kasarnya harus horizontal sepanjang sumbu waktu. Dengan kata lain, fluktuasi data berada disekitar pada suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan varians dari fluktuasi tersebut tetap konstan setiap waktu. Nilai-nilai autokorelasi dari data stasioner akan turun sampai nol sesudah *time-lag* kedua atau ketiga, sedangkan untuk data yang tidak stasioner, nilai-nilai tersebut berbeda signifikan dari nol untuk beberapa periode waktu. Apabila disajikan secara grafik, autokorelasi data yang tidak stasioner memperlihatkan suatu *trend* searah diagonal dari kanan ke kiri bersama dengan meningkatnya jumlah *time-lag* (selisih waktu). Karena model ARIMA membutuhkan terpenuhinya asumsi stasioneritas, maka ketidakstasioneran dari data perlu dihilangkan terlebih dahulu sebelum melangkah lebih lanjut pada pengestimasian model deret berkala. Untuk

menstasionerkan nilai mean yaitu dengan melakukan perbedaan deret data (differencing).

### 2.3.7.5 Operator Backward Shift (Shift mundur)

Notasi yang sangat bermanfaat dalam metode deret berkala ARIMA *Box-Jenkins* adalah operator shift mundur (backward shift) dinotasikan  $B$ , yang penggunaannya adalah sebagai berikut.

$$BX_t = X_{t-1} \quad (2.9)$$

Notasi  $B$  yang dipasang pada  $X_t$ , mempunyai pengaruh menggeser data 1 periode ke belakang. Dua penerapan  $B$  untuk *shift*  $X$  akan menggeser data tersebut 2 (dua) periode ke belakang.

$$B(BX_t) = B^2X_t = X_{t-2} \quad (2.10)$$

Operator *shift* mundur juga dapat digunakan untuk menggambarkan proses perbedaan (differencing). Sebagai contoh apabila suatu deret berkala tidak stasioner maka data tersebut dapat dibuat lebih mendekati stasioner dengan melakukan perbedaan pertama dari deret data. Perbedaan pertama dirumuskan sebagai berikut.

$$X_t^I = X_t - X_{t-1} \quad (2.11)$$

Dengan menggunakan operator *shift* mundur, perbedaan pertama dapat dituliskan sebagai berikut.

$$X_t^I = X_t - BX_t = (1 - B)X_t \quad (2.12)$$

Pembedaan orde kedua dirumuskan sebagai berikut.

$$\begin{aligned} X_t^{II} &= X_t^I - X_{t-1}^I \\ &= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \\ &= X_t - 2X_{t-1} + X_{t-2} \end{aligned} \quad (2.13)$$

Dengan menggunakan operator *shift* mundur maka perbedaan orde kedua dapat ditulis sebagai berikut.

$$\begin{aligned} X_t^{II} &= X_t - 2X_{t-1} + X_{t-2} \\ &= X_t - 2BX_t + B^2X_t \end{aligned} \quad (2.14)$$

$$\bullet (1 - 2B + B^2)X_t$$

Pembedaan orde kedua dinyatakan oleh  $(1-B)^2$ . Salah satu hal yang penting adalah bahwa pembedaan orde kedua yang dinotasikan  $(1-B)^2$  tidak sama dengan pembedaan kedua yang dinotasikan dengan  $(1-B^2)$ .

## 2.4 Tahapan dalam metode ARIMA

### 2.4.1 Tahap Identifikasi

Pada tahap identifikasi, variabel yang akan diramalkan terlebih dahulu diuji kestasioneran datanya. Kestasioneran data dapat diuji dengan cara plot data dan menghitung *autocorrelation function* (ACF). Melalui plot data, dilihat secara visual apakah data memiliki kecenderungan semakin meningkat, semakin menurun, atau terdapat fluktuasi musiman. Sedangkan dari nilai ACF, jika nilai ACF mendekati nol pada lag kedua atau ketiga, maka data tersebut stasioner. Jika data yang diamati memiliki pola musiman, pada plot ACF akan terlihat nilai ACF yang signifikan pada kelipatan musimnya. Deret data non-stasioner dapat dijadikan stasioner dengan melakukan proses *differencing* (pembedaan). Jumlah berapa kali dilakukan proses *differencing* menunjukkan tingkat differensiasi model. Untuk pola data yang mengandung unsur musiman, secara khusus dapat digunakan model *seasonal* ARIMA. Unsur musiman dapat dihilangkan dengan *seasonal differencing*. Setelah data menjadi stasioner, langkah yang selanjutnya dilakukan adalah menentukan model *tentative*. Untuk menentukan model *tentative*, diperlukan analisis dari ACF dan PACF. Pola ACF dan PACF bisa berpola *cut off* dan *dies down*. Pertama, ACF dan PACF dari data *time series* bisa berpola *cut off*. Pola *cut off* adalah pola ketika garis ACF dan PACF signifikan pada lag pertama atau kedua tetapi kemudian tidak ada garis ACF dan PACF yang signifikan pada lag berikutnya. Untuk pola *cut off*, perbedaan antara ACF dan PACF yang signifikan dengan ACF dan PACF yang tidak signifikan adalah besar sehingga garis ACF dan PACF terlihat terpotong (*cut off*). Kedua, ACF dan PACF dikatakan memiliki perilaku *dies down* jika kedua fungsi

tersebut tidak terpotong, melainkan menurun secara bertahap. Bentuk penurunannya bisa tanpa ataupun dengan berbentuk gelombang sinus. Penentuan apakah suatu data *time series* dimodelkan dengan AR, MA atau ARIMA tergantung pola ACF dan PACF. Model AR digunakan jika plot ACFnya *dies down* sementara PACF-nya *cut off*. Model MA digunakan jika plot ACFnya *cut off* dan plot PACF-nya *dies down*. Sedangkan jika kedua plot ACF dan PACF sama-sama *dies down*, maka model yang digunakan adalah model ARIMA.

**Tabel 2.1 kriteria ACF dan PACF pada model ARIMA**

<b>Proses</b>	<b>ACF</b>	<b>PACF</b>
AR(p)	<i>dies down</i> (menurun) mengikuti bentuk eksponensial atau gelombang sinus.	<i>Cut off</i> (terputus) setelah lag ke- <i>p</i>
MA(q)	<i>Cut off</i> setelah lag ke- <i>q</i>	<i>dies down</i> mengikuti bentuk eksponensial atau gelombang sinus.
ARMA(p, q)	<i>dies down</i> setelah lag (q-p)	<i>dies down</i> setelah lag (p-q)

(Wei, 1990)

## 2.4.2 Tahap Penaksiran Parameter

Setelah model sementara untuk suatu runtun waktu diidentifikasi, langkah selanjutnya adalah mencari estimasi terbaik untuk parameter-parameter dalam model sementara tersebut.

### 2.4.2.1 Proses tidak Musiman AR(1) dan AR(2)



Untuk proses *autoregresif* pada orde  $p$  dengan pendekatan *Yule-Walker* didefinisikan sebagai berikut.

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_2 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p\end{aligned}\tag{2.34}$$

dengan,

$\rho_1, \rho_2, \rho_3$  adalah autokorelasi teoritis berturut-turut untuk *time-lag* 1, 2, 3, ...,  $p$ ,  $\phi_1, \phi_2, \dots, \phi_p$  adalah  $p$  buah koefisien AR dari proses AR( $p$ ). Karena nilai teoritis  $\rho$  tidak diketahui maka digantikan dengan nilai empirisnya dan kemudian digunakan untuk memecahkan nilai-nilai  $\phi$ . Untuk proses AR(1), persamaan (2.34) menjadi sebagai berikut.

$$\rho_1 = \phi_1\tag{2.35}$$

Jika  $\rho_1$  yang tidak diketahui diganti dengan  $r_1$  yang diketahui (autokorelasi empiris) diperoleh nilai taksiran parameter  $\phi_1$  untuk proses AR(1) sebagai berikut.

$$\hat{\phi}_1 = r_1\tag{2.36}$$

Untuk proses AR(2), persamaan (2.34) menjadi sebagai berikut.

$$\begin{aligned}\rho_1 &= \phi_1 + \dots + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \dots + \phi_2\end{aligned}\tag{2.37}$$

Jika  $\rho_1$  dan  $\rho_2$  diganti dengan  $r_1$  dan  $r_2$  diperoleh nilai taksiran parameter  $\phi_1$  dan  $\phi_2$  untuk proses AR(2) sebagai berikut.

$$\begin{aligned}\hat{\phi}_1 &= \frac{r_1(1-r_2)}{1-r_1^2}, \\ \hat{\phi}_2 &= \frac{r_2-r_1^2}{1-r_1^2}.\end{aligned}\tag{2.38}$$

#### 2.4.2.2 Proses tidak Musiman MA(1) dan MA(2)

Estimasi atas koefisien MA(q) tidak dilakukan dengan metode linier karena walaupun modelnya modelnya berbentuk linier tetapi koefisiennys sendiri bersifat nonlinier sehingga digunakan metode estimasi nonlinier (Lerbin, 2002). Autokorelasi teoritis untuk proses MA(q) dapat digunakan dalam bentuk koefisien-koefisien MA sebagai berikut.

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+2} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q. \\ 0, & k > q \end{cases} \quad (2.39)$$

Nilai taksiran dari  $\theta_1, \theta_2, \dots, \theta_q$  dapat diperoleh dengan mensubstitusikan autokorelasi empiris,  $r_k$  pada persamaan (2.39) dan kemudian diselesaikan. Untuk proses MA(1), persamaan (2.39) menjadi sebagai berikut.

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2}, & k = 1, \\ 0, & k \geq 2. \end{cases} \quad (2.40)$$

Dengan mensubstitusikan  $r_1$  untuk  $\rho_1$  pada persamaan (2.40) diperoleh persamaan kuadratik sebagai berikut.

$$\theta_1^2 + \left(\frac{1}{r_1}\right) \theta_1 + 1 = 0. \quad (2.41)$$

Dari persamaan (2.40) diperoleh dua penyelesaian yang harus terletak di antara -1 dan 1.

Untuk proses MA(2) persamaannya adalah sebagai berikut.

$$\begin{aligned} \rho_1 &= \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2} \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \\ \rho_k &= 0, \quad k \geq 3 \end{aligned} \quad (2.42)$$

Dengan mensubstitusikan  $r_1$  dan  $r_2$  untuk  $\rho_1$  dan  $\rho_2$  akan menghasilkan dua persamaan dalam  $\theta_1$  dan  $\theta_2$  yang tidak diketahui.

Ada dua cara yang mendasar untuk mendapatkan parameter-parameter tersebut, yaitu:

1. Dengan cara *trial and error* (mencoba-coba), menguji beberapa nilai yang berbeda dan memilih satu nilai tersebut atau sekumpulan nilai, apabila terdapat lebih dari

satu parameter yang akan ditaksir yang meminimumkan *sum of squared residual* (jumlah kuadrat nilai sisa).

2. Perbaiki secara iteratif, memilih taksiran awal dan kemudian membiarkan program komputer memperhalus penaksiran tersebut secara iteratif.

### 2.4.3 Tahap Pemeriksaan Diagnostik

Setelah menaksir nilai-nilai parameter dari model ARIMA yang ditetapkan sementara, selanjutnya perlu dilakukan pemeriksaan untuk membuktikan bahwa model tersebut cukup memadai. Diagnosis model dilakukan untuk mendeteksi adanya korelasi dan kenormalan antar *residual*. Dalam runtun waktu (*time series*) ada asumsi bahwa *residual* mengikuti proses *white noise* yang berarti *residual* harus independen (tidak berkorelasi) dan berdistribusi normal dengan rata-rata mendekati 0 ( $\mu = 0$ ) dan standar deviasi ( $\sigma$ ) tertentu (Iriawan, 2006).

Untuk mendeteksi adanya proses *white noise*, maka perlu dilakukan diagnosis model. Ada beberapa cara dalam hal ini, antara lain:

1. Uji independensi *residual*

Uji dilakukan untuk mendeteksi independensi *residual* antar lag. Dua lag dikatakan independen (tidak berkorelasi) apabila antar lag tidak ada korelasi cukup berarti (Iriawan, 2006: 362). Dalam penelitian ini, uji dilakukan dengan membandingkan  $\chi^2_{Ljung-Box}$  dan  $\chi^2_{(\alpha, df)}$  pada output proses *Ljung-Box-Pierce*.

Hipotesis:

$$H_0: \rho_{a_t, a_{t+k}} = 0 \text{ (Ada korelasi antar-lag)}$$

$$H_1: \rho_{a_t, a_{t+k}} \neq 0 \text{ (tidak ada korelasi antar lag atau minimal ada 1 lag yang } \rho_{a_t, a_{t+k}} \neq 0 \text{)}$$

Kriteria penolakan  $H_0$  yaitu jika  $\chi^2_{Ljung-Box} < \chi^2_{(\alpha, df)}$ , di mana distribusi  $\chi^2$  yang digunakan mempunyai  $df = K-k$ .

Selain dengan pengujian hipotesis, *independensi* antar lag akan ditunjukkan pula oleh grafik fungsi autokorelasi (ACF) residual. Suatu *residual* model

dikatakan telah independen jika tidak ada satu lag pun pada grafik fungsi autokorelasi (ACF) residual yang keluar batas garis (Iriawan, 2006).

2. Uji kenormalan *residual*

Uji dilakukan untuk mendeteksi kenormalan *residual* model. Dalam penelitian ini, uji dilakukan hanya dengan membandingkan nilai *P-Value* pada output proses *Ljung-Box-Pierce* dengan level toleransi ( $\alpha$ ) yang digunakan dalam uji hipotesis.

Hipotesis:

$H_0$ : *Residual* model berdistribusi normal dengan rata-rata mendekati 0 ( $\mu=0$ )

$H_1$ : *Residual* model tidak berdistribusi normal dengan rata-rata mendekati 0 ( $\mu \neq 0$ )

Kriteria penolakan  $H_0$  yaitu jika *P-Value* < level toleransi ( $\alpha$ ) (Iriawan, 2006).

3. Uji *Lack of fit* (kekurangan kesesuaian) dengan *box-pierce Q* statistik *box-pierce Q* dihitung dengan model sebagai berikut.

$$Q = n \sum_{k=1}^m r_k^2 \quad (2.43)$$

dengan,

$m$  = lag maksimum

$n$  = banyaknya data asli

$r_k$  = nilai koefisien autokorelasi time lag  $k$

Jika nilai  $Q$  lebih kecil dari nilai pada tabel *Chi-Square* dengan derajat kebebasan selisih dari  $m-p-q$ , dimana  $p$  dan  $q$  adalah orde dari AR dan MA. Sebaliknya jika nilai  $Q$  lebih kecil dari *Chi-Square* model belum dianggap memadai dan harus mengulangi langkah sebelumnya.

4. *Overfitting* model ARIMA, yaitu menggunakan beberapa parameter lebih banyak daripada yang diperlukan. Namun, dalam hal ini perlu diperhatikan bahwa dalam metode ARIMA berlaku prinsip *parsimony* artinya model yang

dipilih adalah model yang paling sederhana yaitu yang jenjangnya paling rendah dan komponennya paling sedikit.

#### 2.4.4 Peramalan dengan Model ARIMA

Setelah metode peramalan ditetapkan, maka model ARIMA dapat diterapkan pada data, dan dapat dilakukan perkiraan pada data tersebut untuk beberapa periode ke depan. Dengan berjalannya waktu perlu dilakukan evaluasi ulang terhadap model yang sudah dipilih karena ada kemungkinan model perlu diperbaiki karena pola data mungkin berubah.

##### 2.4.4.1 Pengukuran Kesalahan Peramalan

Ada beberapa metode yang digunakan untuk menunjukkan kesalahan yang disebabkan oleh suatu teknik peramalan tertentu. Hampir semua ukuran tersebut menggunakan beberapa fungsi dari perbedaan antara nilai sebenarnya dengan nilai peramalannya. Perbedaan nilai sebenarnya dengan nilai peramalan ini biasanya disebut sebagai residual.

Menurut Arsyad (1995) ada beberapa teknik untuk mengevaluasi hasil peramalan, diantaranya:

1. *Mean Absolute Deviation* (MAD) atau simpangan absolut rata-rata

$$MAD = \frac{\sum_{t=1}^n (X_t - \hat{X}_t)}{n} \quad (2.44)$$

MAD ini sangat berguna jika seorang analis ingin mengukur kesalahan peramalan dalam unit ukuran yang sama seperti data aslinya.

2. *Mean Squared Error* (MSE) atau Kesalahan rata-rata kuadrat

$$MSE = \frac{\sum_{t=1}^n (X_t - \hat{X}_t)^2}{n} \quad (2.45)$$

Pendekatan ini menghukum suatu kesalahan yang besar karena dikuadratkan. Pendekatan ini penting karena satu teknik yang menghasilkan kesalahan yang moderat yang lebih disukai oleh suatu peramalan yang biasanya menghasilkan kesalahan yang lebih kecil tetapi kadang-kadang menghasilkan kesalahan yang sangat besar.

3. *Mean Absolute Percentage Error* (MAPE) atau persentase kesalahan absolut rata-rata

$$MAPE = \frac{\sum_{t=1}^n \frac{|X_t - \hat{X}_t|}{X_t}}{n} \quad (2.46)$$

Pendekatan ini sangat berguna jika ukuran variabel peramalan merupakan factor penting dalam mengevaluasi akurasi peramalan tersebut. MAPE memberikan petunjuk seberapa besar kesalahan peramalan dibandingkan dengan nilai sebenarnya dari series tersebut.

4. *Mean Percentage Error* (MPE) atau Persentase kesalahan rata-rata

$$MAD = \frac{\sum_{t=1}^n (X_t - \hat{X}_t)}{n} \quad (2.47)$$

MPE diperlukan untuk menentukan apakah suatu metode peramalan bias atau tidak. Jika pendekatan peramalan tersebut tidak bias, maka hasil perhitungan MPE akan menghasilkan persentase mendekati nol.