

BAB II

KAJIAN TEORI

Pada Bab II ini dijelaskan mengenai pengertian-pengertian dasar yang digunakan sebagai landasan dalam pembahasan. Pengertian-pengertian dasar tersebut harus dikuasi untuk memudahkan memahami pembahasan. Pengertian-pengertian dasar itu diantaranya adalah pengertian tentang konsep dasar analisis runtun waktu, model regresi linear sederhana, model regresi linear ganda, analisis korelasi sederhana, derajat integrasi, *likelihood ratio test*, distribusi normal, stasioneritas, nilai eigen dan eigen vektor, model *vector autoregressive (VAR)*, *sembako*, *EViews* dan klasifikasi data.

A. Konsep Dasar Analisis Runtun Waktu

Dalam kehidupan sehari-hari banyak data yang diperoleh dari observasi suatu fenomena di berbagai bidang berdasarkan waktu. Misalnya di bidang pertanian, produksi dan harga hasil panen per tahun. Di bidang ekonomi dan bisnis, harga saham per hari, suku bunga per minggu, indeks harga per bulan, jumlah penjualan per kuartal dan pendapatan per tahun. Di bidang meteorologi, kecepatan angin per jam, temperatur, udara per hari dan curah hujan tiap tahun. Di bidang ilmu sosial dipelajari angka kelahiran, angka kematian, angka kecelakaan dan angka pelbagai tindakan kejahatan dan lain sebagainya. Data demikian disebut data runtun waktu yang bisa diobservasi tidak pernah berakhir (Rusdi, 2002: 6).

Data runtun waktu dapat dituliskan sebagai $y_t = y_1, y_2, y_3, \dots, y_n ; t = 1, 2, 3, \dots, n$ dengan y_t adalah data ke t dari runtun waktu tersebut. Jika $t \in A$

dengan A himpunan bilangan asli maka runtun waktu ini dinamakan runtun waktu diskret. Sedangkan jika $t \in R$ dengan R himpunan bilangan riil maka runtun waktu ini dinamakan runtun waktu kontinu.

Menurut (Dedi, 2011:1), untuk dapat memahami pemodelan runtun waktu, perlu diketahui beberapa jenis data menurut waktu, yang dapat dibedakan sebagai berikut :

1. Data *Time Series* (Runtun Waktu) yakni jenis data yang terdiri atas variabel-variabel yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Jika waktu dipandang bersifat diskrit (waktu dapat dimodelkan bersifat kontinu), frekuensi pengumpulan selalu sama (*equidistant*). Dalam kasus diskrit, frekuensi dapat berupa misalnya detik, menit, jam, hari, minggu, bulan atau tahun dan lain-lain.
2. Data *Cross-Section* yakni jenis data yang terdiri atas variabel-variabel yang dikumpulkan pada sejumlah individu atau kategori pada suatu titik waktu tertentu. Model yang digunakan untuk memodelkan data tipe ini seperti model regresi (*Cross-Section*).
3. Data Panel atau *Pooled* yakni tipe data yang terdiri atas variabel-variabel yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu pada sejumlah individu atau kategori. Model yang digunakan untuk pemodelan data ini seperti model data panel dan model runtun waktu multivariat. Secara ekuivalen, dikenal juga tipe data longitudinal dengan frekuensi data tidak harus *equidistant*, namun analisa statistika yang

dilakukan dalam model longitudinal berbeda tujuannya dengan analisa untuk model panel.

B. Model Regresi Linear Sederhana

Secara kualitatif hubungan antara variabel independen dengan variabel dependen dapat dimodelkan dalam suatu persamaan matematik, sehingga dapat diduga nilai variabel dependen jika variabel independennya diketahui. Persamaan matematik yang menghubungkan antara variabel dependen dengan variabel independen disebut persamaan regresi. Model regresi linear klasik yang melibatkan satu variabel independen disebut model regresi linear sederhana, fungsi regresi dapat ditulis sebagai berikut (Draper and Smith, 1992:8):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (2.1)$$

dengan

Y_i : nilai variabel dependen dalam pengamatan ke-i

X_i : nilai variabel independen dalam pengamatan ke-i

ε_i : error yang bersifat acak dalam pengamatan ke-i

β_0 : parameter suku konstan dari model regresi, menunjukkan nilai Y_i jika $X_i = 0$

β_1 : parameter koefisien dari X_i , menunjukkan kemiringan dari garis regresi

Persamaan 2.1 diatas dapat dikatakan linear dalam parameter dan linear dalam variabel independen. Dikatakan linear dalam parameter karena tidak ada parameter yang muncul sebagai eksponen atau hasil kali atau hasil bagi dengan parameter lain. Persamaan 2.1 dikatakan linear dalam variabel independen karena variabel independen dalam model berpangkat satu (Gujarati, 2006:124-125). Beberapa hal yang perlu diperhatikan dalam persamaan 2.1 adalah:

1. $E(Y_i) = \beta_0 + \beta_1 X_i$ karena $E(\varepsilon_i) = 0$

2. $\text{Var}(Y_i) = \sigma^2$ karena $\text{Var}(\varepsilon_i) = \sigma^2$
3. Y_i dan Y_j tidak berkorelasi, karena ε_i dan ε_j tidak berkorelasi
4. ε_i berdistribusi normal

C. Model Regresi Linear Ganda

Model regresi linear ganda dengan k variabel independen adalah sebagai berikut (Draper and Smith, 1992:186):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad (2.2)$$

dengan

- Y_i : nilai variabel dependen dalam pengamatan ke- i
- X_{ik} : nilai variabel independen ke- k dalam pengamatan ke- i
- ε_i : *error* yang bersifat acak dalam pengamatan ke- i
- β_0 : parameter suku konstan dari model regresi, menunjukkan nilai Y_i apabila $X_{ik} = 0$
- β_k : parameter koefisien dari X_{ik} , menunjukkan kemiringan dari garis regresi.

Koefisien-koefisien dalam regresi linear sebenarnya adalah nilai dugaan dari parameter model regresi. Parameter regresi diduga melalui teknik perhitungan yang disebut *Ordinary Least Square (OLS)* atau metode kuadrat terkecil. Model regresi yang diperoleh dari metode kuadrat terkecil merupakan model regresi yang menghasilkan penduga parameter terbaik jika memenuhi asumsi-asumsi sebagai berikut:

1. *Error* menyebar normal dengan rata-rata nol dan suatu ragam (variansi) tertentu, $\varepsilon \sim N(0, \sigma^2)$.
2. Ragam dari *error* bersifat homogen (homokedastisitas)
Maksud dari ragam bersifat homogen adalah bahwa *error* memiliki nilai ragam yang sama antara *error* ke- i dan *error* ke- j .
3. *Error* tidak mengalami autokorelasi

Autokorelasi yaitu terdapat pengaruh dari variabel dalam model melalui tenggang waktu (*timeline*). Misalnya nilai suatu variabel saat ini akan berpengaruh terhadap nilai variabel lain pada saat masa mendatang.

4. Tidak terjadi multikolinearitas antar variabel independen X.

Asumsi ini digunakan dalam kasus regresi linear berganda. Multikolinearitas berarti bahwa terjadi korelasi linear yang erat antara variabel independen.

D. Analisis Korelasi Sederhana

Analisis korelasi pertama kali dikembangkan oleh Karl Pearson pada tahun 1990. Tujuan dari analisis korelasi adalah untuk menentukan seberapa erat hubungan antara dua variabel. Menurut (Subaryadi & Purwanto, 2004: 158), analisis korelasi adalah suatu teknik statistika yang digunakan untuk mengukur keeratan hubungan atau korelasi antara dua variabel.

Analisis korelasi mencoba mengukur keeratan hubungan antara dua variabel X dan Y. Keeratan hubungan antara dua variabel tersebut dinyatakan dalam bentuk *koefisien korelasi* yang dilambangkan dengan huruf r . Koefisien korelasi (r) menunjukkan seberapa dekat titik kombinasi antara variabel X dan Y pada garis lurus sebagai garis dugaannya. Semakin dekat titik kombinasinya dengan garis dugaannya maka nilai korelasinya semakin membesar. Sebaliknya, semakin menyebar dari garis dugaannya, maka nilai korelasi semakin kecil.

Ukuran statistik yang dapat menggambarkan hubungan antara suatu variabel dengan variabel lain adalah koefisien korelasi dan koefisien determinasi.

1. Koefisien Korelasi

Ukuran korelasi antara dua buah variabel yang paling banyak digunakan adalah koefisien korelasi momen yang dikembangkan oleh Pearson. Rumus koefisien korelasi tersebut dinyatakan sebagai berikut (Subaryadi & Purwanto, 2004: 159) :

$$r = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{\sqrt{[n(\Sigma X^2) - (\Sigma X)^2][n(\Sigma Y^2) - (\Sigma Y)^2]}} \quad (2.3)$$

dengan

- r : nilai koefisien korelasi
- ΣX : jumlah pengamatan variabel X
- ΣY : jumlah pengamatan variabel Y
- ΣXY : jumlah hasil perkalian variabel X dan Y
- (ΣX^2) : jumlah kuadrat dari pengamatan variabel X
- $(\Sigma Y)^2$: jumlah kuadrat dari jumlah pengamatan variabel X
- (ΣY^2) : jumlah kuadrat dari pengamatan variabel Y
- $(\Sigma Y)^2$: jumlah kuadrat dari jumlah pengamatan variabel Y
- n : banyaknya pasang pengamatan Y dan X

Koefisien korelasi mempunyai nilai antara -1 sampai 1. Nilai $r = -1$ yang disebut dengan linear sempurna negatif, terjadi apabila titik sampel atau kombinasi terletak tepat pada suatu garis lurus yang mempunyai kemiringan negatif. Nilai $r = 1$ disebut dengan linear sempurna positif, dan hal ini terjadi apabila semua titik contoh terletak tepat pada satu garis lurus dengan kemiringan positif. Nilai koefisien korelasi yang mendekati -1 atau 1 menyatakan bahwa hubungan kedua variabel adalah kuat atau korelasi kedua variabel tinggi. Akan tetapi apabila nilai r mendekati 0, hubungan

antara kedua variabel sangat lemah atau mungkin tidak ada sama sekali (Subaryadi & Purwanto, 2004: 159).

2. Koefisien Determinasi

Menurut Subaryadi & Purwanto (2004:162-163), koefisien determinasi merupakan ukuran untuk mengetahui kesesuaian atau ketepatan antara nilai dugaan atau garis regresi dengan data sampel. Jika semua data observasi terletak pada garis regresi akan diperoleh garis regresi yang sesuai atau sempurna, namun apabila data observasi tersebar jauh dari nilai dugaan atau garis regresinya, maka nilai dugaannya menjadi kurang sesuai.

Koefisien determinasi adalah bagian dari keragaman total variabel terikat Y (variabel yang dipengaruhi atau *dependent*) yang dapat diterangkan atau diperhitungkan oleh keragaman variabel bebas X (variabel yang mempengaruhi atau *independent*). Jadi koefisien determinasi adalah kemampuan variabel X (variabel independen) mempengaruhi variabel Y (variabel terikat). Semakin besar koefisien determinasi menunjukkan semakin baik kemampuan X menerangkan Y. Besarnya koefisien determinasi adalah kuadrat dari koefisien korelasi dan dirumuskan sebagai berikut:

$$r^2 = \frac{[n(\Sigma XY) - (\Sigma X)(\Sigma Y)]^2}{\sqrt{[n(\Sigma X^2) - (\Sigma X)^2][n(\Sigma Y^2) - (\Sigma Y)^2]}} \quad (2.4)$$

Apabila nilai koefisien korelasi sudah diketahui, maka untuk mendapatkan koefisien determinasi dapat diperoleh dengan mengkuadratkannya.

E. Derajat Integrasi

Data runtun waktu dikatakan terintegrasi pada orde d atau dinotasikan $I(d)$ jika runtun waktu tersebut dapat di-*differencing* sebanyak d kali dan hasil *differencing* adalah stasioner (Gujarati, 2003). Dengan kata lain derajat integrasi adalah dimana pada derajat tersebut data runtun waktu bersifat stasioner. Tujuan derajat integrasi adalah untuk melihat apakah runtun waktu terintegrasi atau tidak.

F. Likelihood Ratio Test

Likelihood Ratio Test (LRT) adalah uji signifikansi yang digunakan untuk membandingkan data *likelihood* hipotesis null terhadap data *likelihood* hipotesis alternatif.

Definisi 2.1 : *Likelihood Ratio Test* (Bain and Engelhardt, 1992: 418)

Misalkan X variabel random yang mempunyai fungsi peluang $f(y; \mu)$ dan mempertimbangkan hipotesis $H_0: \mu \in \Omega_0$ melawan $H_a: \mu \in \Omega - \Omega_0$, dengan Ω_0 (parameter dibatasi) himpunan bagian dari Ω (semua parameter).

Likelihood ratio didefinisikan:

$$\lambda(x) = \frac{f(y; \hat{\mu}_0)}{f(y; \hat{\mu})} \quad (2.5)$$

$\hat{\mu}$ dan $\hat{\mu}_0$ diperoleh dengan memaksimalkan $f(y; \mu)$ berturut-turut dalam Ω dan Ω_0 atau $f(y; \hat{\mu}_0)$ adalah nilai maksimum dari data *likelihood* dengan semua parameter dan $f(y; \hat{\mu})$ adalah nilai maksimum dari data *likelihood* ketika parameternya dibatasi atau dikurangi jumlahnya berdasarkan asumsi.

Statistik ujinya $G = -2 \ln \lambda \sim \chi_{\alpha, v}^2$

H_0 ditolak jika $G_{hitung} > \chi_{\alpha, v}^2$ dengan $v = n - k - 1$

dengan

n : banyaknya pengamatan

k : banyaknya variabel prediktor dari data *likelihood* dengan semua parameter

G. Stasioneritas

Sebuah data dikatakan stasioner jika tidak terdapat pertumbuhan atau penurunan data. Dengan kata lain, fluktuasi data berada disekitar suatu nilai rata-rata yang konstan. Stasioner atau tidaknya suatu data dapat dilihat dengan grafik. Jika grafik cenderung naik maka dapat disimpulkan sementara bahwa data tersebut tidak stasioner.

Suatu data dikatakan stasioner jika memenuhi keadaan sebagai berikut (Thomas, 1997:374):

1. $E(X_t) = \mu$ konstan untuk semua t
2. $Var(X_t) = \sigma^2$ konstan untuk semua t
3. $Cov(X_t, X_{t+k}) = \gamma_k$ konstan untuk semua t dan $k \neq 0$

Jenis-jenis data yang tidak stasioner adalah data dengan adanya pola *trend*, data dengan adanya pengaruh seasonal dan data dengan adanya pengaruh siklis. Data dengan adanya pola *trend* ditandai dengan adanya kecenderungan arah data bergerak menaik atau menurun pada jangka panjang. Data dengan adanya pengaruh seasonal ditandai dengan adanya pola perubahan yang berulang secara otomatis dari tahun ke tahun. Data dengan adanya pengaruh siklis maksudnya adalah data yang terjadi di sekitar garis *trend*. Pola siklis agak sulit diprediksi karena pola yang ada cenderung tidak stabil.

Kestasioneran data terbagi menjadi dua yaitu kestasioneran dalam *mean* dan varians. Data runtun waktu yang tidak stasioner dapat dikatakan distasionerkan dengan cara sebagai berikut:

1. Ketidakstasioneran dalam *mean*

Data runtun waktu yang tidak stasioner dalam *mean* berarti memiliki *mean* yang tidak konstan atau *meannya* terpengaruh oleh waktu. Cara yang digunakan untuk mengatasi ketidakstasioneran dalam *mean* ini dengan melakukan pembedaan (*differencing*) terhadap data runtun waktu.

2. Ketidakstasioneran dalam varians

Data runtun waktu yang tidak stasioner dalam varians berarti mempunyai varians yang tidak konstan atau variansnya terpengaruh oleh waktu. Cara yang digunakan untuk mengatasi ketidakstasioneran dalam varians yaitu dengan melakukan transformasi.

Proses stasioneritas rata-rata pada data dapat menggunakan metode *differencing*, sedangkan untuk proses stasioneritas varians pada data dapat menggunakan transformasi *box-cox*.

1. *Differencing*

Salah satu cara umum yang digunakan dalam stasioneritas data adalah metode *differencing*. Tujuan menghitung pembedaan adalah untuk mencapai stasioneritas dalam rata-rata. Metode ini dilakukan dengan cara mengurangi nilai data pada suatu periode dengan data periode sebelumnya. Pembedaan dapat dilakukan untuk beberapa order sampai diperoleh data yang stasioner.

Notasi B (operator *shift* mundur / *backward shift*) digunakan untuk melakukan pembedaan. Operator B didefinisikan sebagai berikut:

$$BY_t = Y_{t-1} \quad (2.6)$$

dan secara umum untuk d periode dapat ditulis:

$$B^d Y_t = Y_{t-d} \quad (2.7)$$

Pembedaan periode pertama adalah sebagai berikut:

$$Y'_t = Y_t - Y_{t-1} \quad (2.8)$$

yang dapat ditulis dengan operator B:

$$Y_t - BY_t = (1 - B)Y_t \quad (2.9)$$

Pembedaan pada periode kedua adalah sebagai berikut:

$$Y''_t = Y'_t - Y'_{t-1}$$

$$Y''_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$$

$$Y''_t = Y_t - 2Y_{t-1} + Y_{t-2}$$

$$Y''_t = (1 - 2B + B^2)Y_t$$

$$Y''_t = (1 - B)^2 Y_t \quad (2.10)$$

Tujuan menghitung pembedaan adalah untuk mencapai stasioneritas. Secara umum, apabila terdapat pembedaan untuk periode ke- d untuk mencapai stasioneritas dapat dinyatakan sebagai berikut:

$$Y_t^d = (1 - B)^d Y_t \quad (2.11)$$

2. Transformasi *Box-Cox*

Transformasi *Box-Cox* adalah salah satu metode untuk proses stasioneritas data dalam varians yang dikenalkan oleh Box dan Tiao Cox.

Transformasi ini didefinisikan sebagai:

$$T(Y_t) = \frac{Y_t^\lambda}{\lambda}, \lambda \neq 0 \text{ dan } \ln Y_t, \lambda = 0 \quad (2.12)$$

λ disebut parameter transformasi. Jenis transformasi yang digunakan tergantung dari nilai estimasi λ . Estimasi nilai λ dapat diperoleh dengan menggunakan minitab. Tabel 2.1 menunjukkan transformasi yang harus dilakukan pada data runtun waktu berdasarkan estimasi λ yang diperoleh (Wei, 2006: 84)

Tabel 2.1 Transformasi Box-Cox

Nilai λ (lamda)	Transformasi
-1	$-1/Y_t$
-0,5	$-2/\sqrt{Y_t}$
0	$\ln Y_t$
0,5	$2\sqrt{Y_t}$
1	Y_t (tidak ada transformasi)

Pengujian stasioneritas dapat juga dilakukan dengan menguji *Augmented Dickey Fuller Test (ADF Test)*. Menurut (Gujarati, 2004: 814) langkah-langkah stasioneritas dengan menggunakan *ADF Test* adalah sebagai berikut:

1. Langkah pertama

Misalkan terdapat persamaan sebagai berikut:

$$Y_t = \beta Y_{t-1} + \varepsilon_t \quad (2.13)$$

Dimana β adalah koefisien *autoregressive*, ε_t adalah *error* yang bersifat random dimana *errornya* memiliki rata-rata sama dengan nol dan varians konstan serta tidak mengandung autokorelasi. Jika $\beta = 1$ maka dapat dinyatakan bahwa variabel Y_t tidak stasioner.

Hipotesisnya adalah:

$H_0: \beta = 1$ (runtun waktu tidak stasioner)

$H_1: \beta < 1$ (runtun waktu stasioner)

Apabila H_0 ditolak maka Y_t stasioner, sedangkan jika H_0 diterima maka dilanjutkan langkah berikutnya.

2. Langkah kedua

Persamaan (2.13) diatas dapat dinyatakan dalam bentuk lain (*first difference*) yaitu:

$$Y_t - Y_{t-1} = \beta Y_{t-1} - Y_{t-1} + \varepsilon_t \quad (2.14)$$

$$\Delta Y_t = (\beta - 1)Y_{t-1} + \varepsilon_t \quad (2.15)$$

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad (2.16)$$

$\delta = (\beta - 1)$ dan Δ adalah *first difference* atau dengan mudah dinyatakan dalam bentuk $\Delta Y_t = (Y_t - Y_{t-1})$, sehingga hipotesisnya menjadi:

$H_0 : \delta = 1$ (runtun waktu tidak stasioner)

$H_1 : \delta < 1$ (runtun waktu stasioner)

Jika $\delta = 0$ maka persamaan (2.16) dapat ditulis:

$$\Delta Y_t = \varepsilon_t \quad (2.17)$$

Persamaan ini menunjukkan bahwa turunan persamaan dari runtun waktu yang *random walk* adalah sebuah runtun waktu stasioner dengan asumsi bahwa ε_t adalah benar-benar *random*.

3. Langkah ketiga

Setelah didapat persamaannya, statistik uji yang digunakan adalah statistik *ADF Test* berikut:

$$t_{hitung} = \frac{\hat{\beta}}{SE(\hat{\beta})} \quad (2.18)$$

Dengan melihat nilai dari statistik *ADF Test* yang merupakan koefisien *autoregressivenya* dapat diketahui bahwa runtun waktu stasioner atau tidak. Jika nilai *ADF Test* (t hitung) absolut kurang dari nilai kritis absolut tabel MacKinnon dengan $db = n - k, n$ adalah banyaknya pengamatan dan k adalah banyaknya parameter yang digunakan, maka H_0 diterima atau dapat dikatakan bahwa runtun waktu tidak stasioner. Sebaliknya, jika nilai *ADF Test* lebih dari nilai kritis absolut tabel MacKinnon, maka H_0 ditolak atau dapat dikatakan bahwa runtun waktu stasioner.

H. Nilai Eigen dan Vektor Eigen

Untuk setiap matriks persegi A , terdapat suatu skalar λ dan vektor kolom tak nol x yang memenuhi (Roman, 1992: 137)

$$Ax = \lambda x \quad (2.19)$$

dengan λ merupakan nilai eigen (nilai karakteristik) dan x merupakan vektor eigen (vektor karakteristik) dari matriks A .

Persamaan (2.19) dapat ditulis juga menjadi persamaan berikut:

$$Ax - \lambda x = 0$$

atau

$$(A - \lambda I)x = 0 \quad (2.20)$$

Nilai λ pada persamaan (2.20) dapat diselesaikan menggunakan (Rencher and Schaalje, 2008 : 47)

$$\det(A - \lambda I) = 0 \quad (2.21)$$

yang disebut sebagai persamaan karakteristik.

Apabila A berukuran $n \times n$, maka persamaan karakteristik (2.21) akan memiliki n akar karakteristik atau memiliki n nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_n$. Nilai λ tidak harus semua berbeda, tidak harus semuanya tidak nol, ataupun harus merupakan bilangan real (Rencher and Schaalje, 2008 : 47). Setelah ditemukan nilai eigen, dapat ditentukan nilai vektor eigen x_1, x_2, \dots, x_n menggunakan persamaan (2.20).

I. Model Vector Autoregressive (VAR)

Model *Vector Autoregressive (VAR)* merupakan gabungan dari beberapa model *Autoregressive (AR)*. Menurut Makridakis (1995:385), secara umum model runtun waktu *autoregressive* berorde p atau $AR(p)$ dirumuskan sebagai berikut:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (2.22)$$

dengan

$$\varepsilon_t \sim N(0, \sigma_t^2)$$

X_t : data observasi ke- t

$\phi_1, \phi_2, \dots, \phi_p$: parameter model AR ke- $i, i = 1, 2, \dots, p$

$X_{t-1}, X_{t-2}, \dots, X_{t-p}$: nilai masa lalu deret runtun waktu yang bersangkutan pada saat $t - 1, t - 2, \dots, t - p$

Model *VAR* merupakan model yang berguna dalam memahami adanya hubungan timbal balik antara variabel ekonomi. Dalam pendekatan ini, semua variabel dianggap sebagai endogen (*dependent*) dan estimasi dilakukan secara serentak. Konsep *VAR* diperkenalkan oleh Christopher A. Sims pada pembentukan model persamaan simultan. Sims berpendapat, dalam persamaan simultan, jika terdapat hubungan yang simultan antar variabel yang diamati, maka variabel-variabel tersebut diperlakukan sama, akibatnya tidak ada lagi variabel endogen dan eksogen (*independent*). Berdasarkan pemikiran Sims

memperkenalkan konsep yang sering disebut *Vector Autoregressive* (Heni, 2012).

Misalkan runtun waktu $Y_t = (Y_{1t}, Y_{2t}, Y_{3t}, \dots, Y_{mt})$, model vector autoregressive dari orde p , ditulis $VAR(p)$ didefinisikan sebagai (Hamilton, 1994):

$$\begin{aligned}
Y_{1t} &= \alpha_1 + \beta_{11}^{(1)}Y_{1,t-1} + \beta_{12}^{(1)}Y_{2,t-1} + \dots + \beta_{1m}^{(1)}Y_{m,t-1} + \beta_{11}^{(2)}Y_{1,t-2} + \\
&\beta_{12}^{(2)}Y_{2,t-2} + \dots + \beta_{1m}^{(2)}Y_{m,t-2} + \dots + \beta_{11}^{(p)}Y_{1,t-p} + \beta_{12}^{(p)}Y_{2,t-p} + \dots + \\
&\beta_{1m}^{(p)}Y_{m,t-p} + \varepsilon_{1t} \\
Y_{2t} &= \alpha_2 + \beta_{21}^{(1)}Y_{1,t-1} + \beta_{22}^{(1)}Y_{2,t-1} + \dots + \beta_{2m}^{(1)}Y_{m,t-1} + \beta_{21}^{(2)}Y_{1,t-2} + \\
&\beta_{22}^{(2)}Y_{2,t-2} + \dots + \beta_{2m}^{(2)}Y_{m,t-2} + \dots + \beta_{21}^{(p)}Y_{1,t-p} + \beta_{22}^{(p)}Y_{2,t-p} + \dots + \\
&\beta_{2m}^{(p)}Y_{m,t-p} + \varepsilon_{2t} \\
&\vdots \\
&\vdots \\
&\vdots \\
Y_{mt} &= \alpha_m + \beta_{m1}^{(1)}Y_{1,t-1} + \beta_{m2}^{(1)}Y_{2,t-1} + \dots + \beta_{mm}^{(1)}Y_{m,t-1} + \beta_{m1}^{(2)}Y_{1,t-2} + \\
&\beta_{m2}^{(2)}Y_{2,t-2} + \dots + \beta_{mm}^{(2)}Y_{m,t-2} + \dots + \beta_{m1}^{(p)}Y_{1,t-p} + \beta_{m2}^{(p)}Y_{2,t-p} + \dots + \\
&\beta_{mm}^{(p)}Y_{m,t-p} + \varepsilon_{mt}
\end{aligned} \tag{2.23}$$

Model $VAR(p)$ pada persamaan (2.23) dapat ditulis secara sederhana sebagai berikut:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \varepsilon_t \tag{2.24}$$

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \varepsilon_t \tag{2.25}$$

dengan

Y_t : vektor berukuran $m \times 1$ dengan entri-entri m variabel dalam VAR

α : vektor berukuran $m \times 1$ dengan entri-entri intersep

β_i : matriks berukuran $m \times m$ dengan entri-entri koefisien-koefisien dalam VAR

ε_t : vektor berukuran $m \times 1$ dengan entri-entri galat dari model VAR

Selain itu, persamaan (2.23) dapat ditulis dalam matriks menjadi:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{mt} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} + \begin{bmatrix} \beta_{11}^1 & \beta_{12}^1 & \cdots & \beta_{1m}^1 \\ \beta_{21}^1 & \beta_{22}^1 & \cdots & \beta_{2m}^1 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1}^1 & \beta_{m2}^1 & \cdots & \beta_{mm}^1 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ \vdots \\ Y_{m,t-1} \end{bmatrix} + \cdots +$$

$$\begin{bmatrix} \beta_{11}^p & \beta_{12}^p & \cdots & \beta_{1m}^p \\ \beta_{21}^p & \beta_{22}^p & \cdots & \beta_{2m}^p \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1}^p & \beta_{m2}^p & \cdots & \beta_{mm}^p \end{bmatrix} \begin{bmatrix} Y_{1,t-p} \\ Y_{2,t-p} \\ \vdots \\ Y_{m,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{mt} \end{bmatrix} \quad (2.26)$$

Menurut Gujarati (2003), model VAR memiliki beberapa keunggulan, antara lain:

1. Model VAR terbebas dari beberapa batasan teori ekonomi yang sering muncul.
2. VAR membangun model secara bersamaan didalam suatu sistem yang kompleks, sehingga dapat dilihat hubungan keseluruhan variabel didalam persamaan.
3. Peneliti tidak perlu membedakan mana variabel endogen maupun eksogen karena semua variabel VAR adalah endogen.
4. Model VAR dapat melihat hubungan antar variabel didalam suatu sistem persamaan.

5. Estimasi parameter dapat menggunakan metode sederhana yaitu dengan metode kuadrat terkecil atau *Least Square (LS)* dan dapat dibuat model terpisah untuk masing-masing variabel endogen.
6. Hasil estimasi peramalan dapat diperoleh melalui model *VAR* dalam banyak kasus lebih baik dari pada hasil estimasi dari model-model persamaan simultan yang lebih kompleks.

Sekalipun banyak keunggulan, model *VAR* juga memiliki beberapa kelemahan.

Menurut Gujarati (2003), beberapa kelemahan *VAR* antara lain:

1. Model *VAR* lebih bersifat a-teoritik (tidak memiliki landasan teori) karena tidak memanfaatkan informasi atau teori terdahulu. Oleh karenanya, model ini sering disebut sebagai model yang tidak struktural.
2. Model *VAR* lebih menitikberatkan peramalan (*forecasting*), sehingga model *VAR* dianggap kurang sesuai untuk analisis kebijakan.
3. Pemilihan banyaknya lag yang digunakan dalam persamaan menimbulkan permasalahan tersendiri.
4. Semua variabel dalam model *VAR* harus memenuhi kestasioneran.
5. Koefisien dalam model *VAR* sulit untuk diinterpretasikan.

J. Sembako

Sembilan bahan pokok atau sering disingkat sembako adalah sembilan jenis kebutuhan pokok masyarakat yang terdiri atas berbagai bahan-bahan makanan dan minuman. Menurut keputusan Menteri Industri dan Perdagangan no 115/mpp/kep/2/1998 tanggal 27 Februari 1998, kesembilan bahan pokok itu adalah beras, sagu dan jagung, gula pasir, sayur-sayuran dan buah-buahan,

daging sapi dan ayam, minyak goreng dan margarin, susu, telur, minyak tanah atau gas elpiji, garam beryodium dan bernatrium.

K. EViews

EViews adalah program komputer yang digunakan untuk mengolah data statistika dan data ekonometrika. Program ini dapat dijalankan pada sistem operasi *Microsoft Windows*, sejak versi XP atau sesudahnya, baik versi 32 maupun 64 bit. *EViews* merupakan kelanjutan dari program *MicroTSP* yang dikeluarkan pada tahun 1981. Program *EViews* dibuat oleh QMS (*Quantitative Micro Software*) yang berkedudukan di Irvine, California, Amerika Serikat. Alamat situsnya ada di www.eviews.com (Wing, 2011: 1).

EViews dapat dijalankan dengan dua cara, yaitu dengan memilih menu atau dengan menuliskan perintah. Cara pertama dengan memilih menu dapat menggunakan *mouse* atau *keyboard* sering juga disebut dengan cara interaktif. Cara ini lebih mudah dilakukan kebanyakan pemakai komputer karena pemakai tinggal memilih menu yang sesuai. Sedangkan cara kedua dengan menuliskan perintah sering disebut dengan metode *batch*. Cara ini sedikit lebih sulit karena pemakai komputer harus menghafal perintah-perintah yang disediakan. Keunggulan dengan menuliskan perintah adalah perintah-perintah yang ditulis dapat disimpan dan dijalankan pada waktu lain atau dikirim ke pemakai lainnya tanpa perlu memilih-milih menu lagi (Wing, 2011: 2).

EViews bersifat *window based* sehingga memakan banyak memori. Namun demikian sisi positifnya adalah sifat *user friendly*. Meskipun *EViews* mampu mengelola berbagai tipe data, namun *software* ini dianggap memiliki

kemampuan lebih dalam hal *prosesing* data runtun waktu karena banyaknya tipe analisis yang dapat dilakukan. Dalam menggunakan *EViews* pengguna akan berinteraksi dengan beberapa komponen penting, yakni *workfile*, *object*, *series* dan *database* (Doddy, 2012: 195-196).

Keunggulan *EViews* terletak pada kemampuannya untuk mengolah data yang bersifat *time series*, meskipun tetap dapat mengolah data *cross section* maupun data panel. Keunggulan lainnya, *EViews* tidak memerlukan langkah yang panjang seperti pada program sejenis untuk mengolah data. Pada program *EViews*, cukup dengan beberapa mengklik *mouse*, hasil akan tampak di layar. Hasil analisis *EViews* selalu ditampilkan dalam satu layar, sehingga mudah dan praktis untuk dianalisis (Wing, 2011: 3).

Kelemahan *EViews* yang utama adalah cara penggunaannya yang tidak biasa seperti pada program lainnya. Contoh kasus adalah untuk menjalankan regresi, menunya menggunakan istilah *Quick Estimate*, dan bukan *Analysis* atau *Regression* seperti pada program statistika lainnya. Kelemahan lain, menurut beberapa peneliti yang sudah banyak menggunakan program *EViews* adalah kelemahannya mengolah data grafik. Oleh karenanya, bila banyak mengolah data dan perlu membuat grafik, dianjurkan tetap mengolah datanya dengan program *EViews* ini, namun grafiknya diselesaikan dengan program *spredsheets* seperti *Microsoft Excel* dan *Lotus 1-2-3*. Ini dikarenakan kemampuan pembuatan grafiknya jauh lebih baik bila dibanding dengan kemampuan *EViews* (Wing, 2011: 3).

Untuk dapat menggunakan program *EViews*, dibutuhkan pengetahuan mengenai statistika, baik statistika deskriptif, statistika parametrik maupun statistika nonparametrik. Selain itu, agar dapat memanfaatkan *EViews* dengan baik juga dibutuhkan pengetahuan mengenai cara pengoperasian komputer di dalam sistem grafik seperti *Microsoft Windows*. Program *EViews* sangat membantu dalam proses penyelesaian uji kointegrasi khususnya dengan menggunakan metode Johansen.

L. Klasifikasi Data

Menurut (Syofian, 2010: 6) data adalah bahan mentah yang perlu diolah sehingga menghasilkan informasi atau keterangan, baik kualitatif maupun kuantitatif yang menunjukkan fakta. Data juga dapat didefinisikan sebagai kumpulan fakta atau angka atau segala sesuatu yang dapat dipercaya kebenarannya sehingga dapat digunakan sebagai dasar menarik suatu kesimpulan. Menurut (Hasan 2004: 19) suatu data dapat diklasifikasikan menjadi empat macam yaitu berdasarkan sumber pengambilan, waktu pengambilan, sifat data dan tingkat pengukuran. Keempat klasifikasi data tersebut adalah sebagai berikut:

1. Berdasarkan Sumber Pengambilannya

Berdasarkan sumber pengambilannya, data dibedakan menjadi dua yaitu data primer dan data sekunder.

a. Data Primer

Data primer adalah data yang diperoleh atau dikumpulkan langsung dilapangan oleh orang yang melakukan penelitian atau yang

bersangkutan yang memerlukannya. Data primer disebut juga data asli atau data baru.

Contoh : data kuesioner, data survei dan sebagainya.

b. Data Sekunder

Data sekunder adalah data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian dari sumber-sumber yang telah ada. Data ini biasanya diperoleh dari perpustakaan atau dari laporan-laporan penelitian terdahulu.

Contoh: data yang sudah tersedia ditempat-tempat tertentu seperti perpustakaan, BPS (Badan Pusat Statistika), kantor-kantor dan sebagainya.

2. Berdasarkan Waktu Pengumpulannya

Berdasarkan waktu pengumpulannya, data dibedakan menjadi dua yaitu data runtun waktu (*Time Series*) dan data *cross section*.

a. Data runtun waktu (*Time Series*)

Data runtun waktu (*time series*) adalah data yang terkumpul dari waktu ke waktu untuk memberikan gambaran perkembangan suatu kegiatan atau keadaan.

Contoh: data perkembangan harga sembilan macam bahan pokok selama 10 bulan terakhir yang dikumpulkan setiap bulan.

b. *Data Cross Section*

Data cross section adalah data yang terkumpul pada suatu waktu tertentu untuk memberikan gambaran perkembangan suatu kegiatan atau keadaan pada waktu itu.

Contoh: data sensus penduduk tahun 2013.

3. Berdasarkan Sifat Data

Berdasarkan sifatnya, data dibedakan menjadi dua yaitu data kualitatif dan data kuantitatif.

a. Data Kualitatif

Data kualitatif adalah data yang tidak berbentuk bilangan.

Contoh: jenis kelamin, agama, warna.

b. Data Kuantitatif

Data kuantitatif adalah data yang berbentuk bilangan.

Contoh: tinggi, panjang, umur, dan sebagainya.

4. Berdasarkan Tingkat Pengukurannya

Berdasarkan tingkat pengukurannya (skala), data dibedakan menjadi empat yaitu data nominal, data ordinal, data interval rasio.

a. Data Nominal

Data nominal adalah data yang berasal dari pelabelan berdasarkan kategori atau klasifikasi tertentu yang perbedaannya hanya menunjukkan perbedaan kualitatif. Ciri-ciri data nominal adalah posisi data setara dan tidak bisa dilakukan operasi matematika (\times , $+$, $-$ atau \div).

Contoh: jenis kelamin, pria diberi label 1 dan wanita diberi label 0.

b. Data Ordinal

Data ordinal adalah data yang berdasarkan urutan yang disusun menurut besarnya, dari tingkat terendah ke tingkat tertinggi atau sebaliknya, dengan jarak atau rentang yang tidak harus sama. Data ordinal disebut juga sebagai data berurutan, data berjenjang, data berpangkat, data tata jenjang atau data bertingkat. Ciri-ciri data ordinal adalah mempunyai urutan, posisi data tidak setara dan tidak dapat dilakukan operasi matematika.

Contoh: kepuasan pelanggan sebuah cafe, misalnya sangat puas diberi label 1, puas diberi label 2, cukup puas diberi label 3, tidak puas diberi label 4, sangat tidak puas diberi label 5.

c. Data Interval

Data interval adalah data yang diperoleh dengan cara pengukuran, dimana jarak antar dua titik pada skala sudah diketahui. Pada data interval ini tidak terdapat titik nol absolut atau mutlak. Ciri-ciri data interval adalah tidak ada pelabelan seperti pada data nominal dan data ordinal serta dapat dilakukan operasi matematika.

Contoh: suhu dapat diukur dengan Celcius atau Fahrenheit dengan masing-masing mempunyai skala tersendiri. Misal Celcius pada 0°C sampai 100°C . Fahrenheit pada 323°F sampai 212°F .

d. Data Rasio

Data rasio adalah data yang diperoleh dengan cara pengukuran, dimana jarak dua titik pada skala sudah diketahui dan mempunyai titik

nol absolut. Ini berbeda dengan data interval dimana tak ada titik nol absolut atau mutlak. Seperti titik 0°C tentu berbeda dengan titik 0°F . Ciri-ciri data rasio adalah tidak ada pemberian label atau kode dan dapat dilakukan operasi matematika.

Contoh: panjang sebuah benda adalah 100cm.