

## BAB III

### AUTOKORELASI

#### 3.1 TIMBULNYA PERSOALAN AUTOKORELASI

Istilah Autokorelasi didefinisikan sebagai korelasi antara anggota serangkaian observasi yang disusun menurut urutan waktu (data time series) atau data silang sesaat (data cross-section), atau korelasi pada dirinya sendiri. Dalam analisa regresi, model regresi linier mengasumsikan bahwa autokorelasi seperti itu tidak terdapat dalam kesalahan  $\varepsilon_i$ .

Atau dengan lambang

$$E(\varepsilon_i \varepsilon_j) = 0, \quad i \neq j$$

Model sederhana ini menganggap bahwa kesalahan  $\varepsilon_i$  yang berhubungan dengan data observasi ke-i tidak akan dipengaruhi oleh kesalahan  $\varepsilon_j$  yang berhubungan dengan data observasi ke-j ( $i, j = 1, 2, \dots, n$ ).

Contoh :

1. Pada data time-series.

Terdapat data kuartalan yang berkenaan dengan regresi output (Y) / hasil atas masukan tenaga kerja ( $X_1$ ) dan modal ( $X_2$ ). Dan misalnya ada pemogokan para karyawan yang mempengaruhi hasil dalam satu kuartal, maka tidak

ada alasan untuk percaya bahwa pemogokan ini akan berlangsung terus sampai ke kuartal berikutnya.

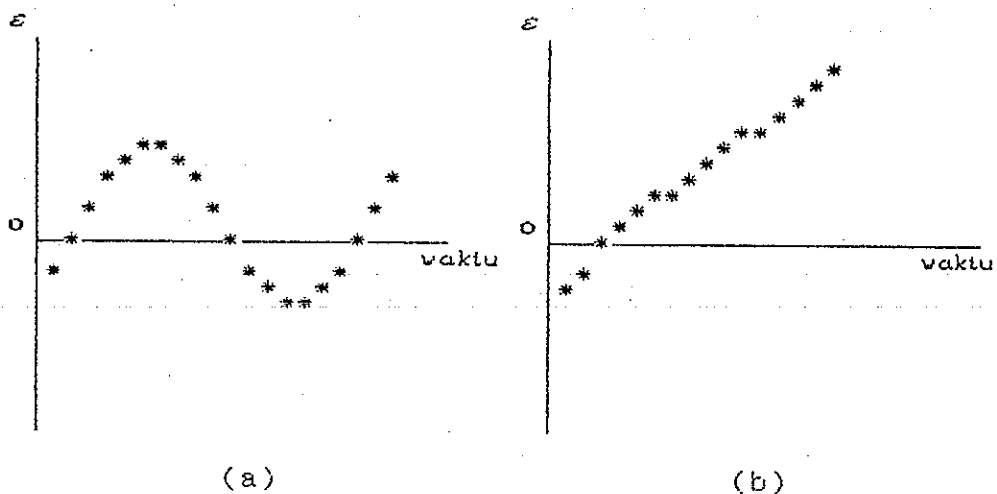
2. Pada data cross section.

Jika terdapat data yang berkenaan dengan regresi mengenai konsumsi rumah tangga (Y) terhadap pendapatan rumah tangga (X), maka pengaruh kenaikan pendapatan dari suatu rumah tangga akan mempengaruhi konsumsi rumah tangga yang bersangkutan, tetapi tidak diharapkan mempengaruhi konsumsi rumah tangga lain.

Tetapi jika terdapat ketergantungan antara  $\varepsilon_i$  dan  $\varepsilon_j$ , maka dikatakan ada autokorelasi, secara matematis dinyatakan dengan

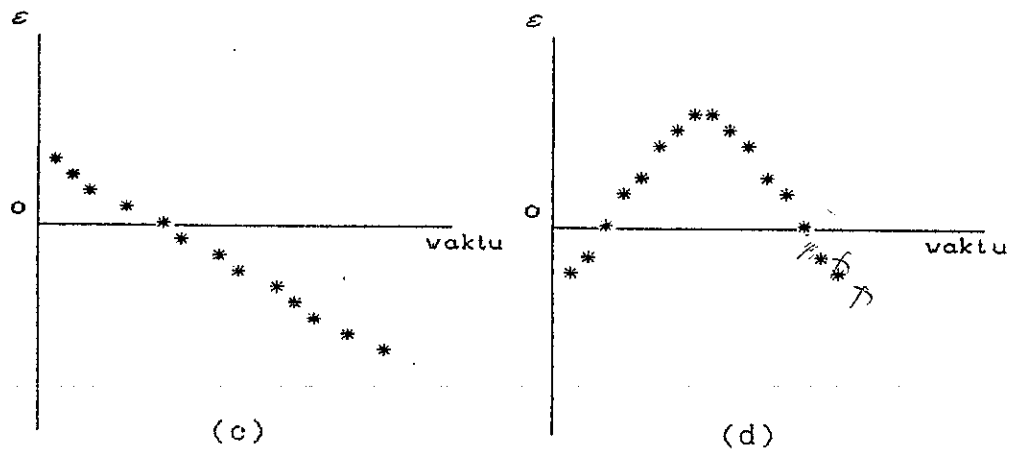
$$E(\varepsilon_i \varepsilon_j) \neq 0, i \neq j$$

Secara visual, pola yang bisa diterima daripada autokorelasi dan tidak adanya autokorelasi, ditunjukkan oleh gambar 3.1 dan gambar 3.2

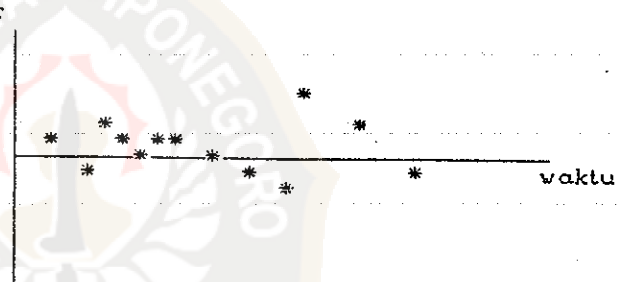


(a)

(b)



Gambar 3.1 Pola Autokorelasi



Gambar 3.2 Pola Non-autokorelasi

Terlihat bahwa gambar 3.1 mengikuti pola yang sistematis, sedangkan pada gambar 3.2 tidak menunjukkan pola yang sistematis untuk  $\epsilon$ .

Beberapa alasan terjadinya autokorelasi :

#### 1. Kelembaman (Inertia).

Sifat atau tanda yang menonjol daripada data time-series ekonomi ialah kelembaman. Data time-series seperti GNP, indeks harga, produksi, kesempatan kerja, dan pengangguran menunjukkan pola siklus (siklus bisnis). Mulai dari bawah pada waktu terjadi resesi,

ketika ekonomi mulai membaik, variabel-variabel pada suatu waktu lebih besar / lebih kecil dari waktu sebelumnya. Jadi dalam regresi yang mencakup time-series, data observasi yang berturutan saling tergantung.

## 2. Kasus variabel penting tidak dimasukkan.

Dalam analisa data empiris, seringkali peneliti memulai dengan model regresi yang dapat diterima, walaupun disadari model tersebut belum sempurna. Setelah membuat analisa regresi, peneliti kemudian mengadakan pengujian untuk mengetahui apakah hasilnya cocok dengan harapan sebelumnya. Jika tidak, suatu tindakan koreksi segera dilakukan, misalnya peneliti tersebut menggambarkan residual  $e_i$  yang diperoleh dari analisa regresi dan mencocokkan dengan pola seperti pada gambar 3.1 dan gambar 3.2. Residual  $e_i$  ini yang merupakan pendekatan untuk  $\epsilon_i$ , mungkin menunjukkan bahwa variabel yang semula merupakan calon untuk dimasukkan dalam model regresi karena berbagai alasan ternyata kemudian tidak dimasukkan. Ini merupakan kasus bias dalam spesifikasi, karena variabel yang tidak dimasukkan. Sering terjadi bahwa dengan memasukkan variabel yang dimaksud dapat menghilangkan pola korelasi yang terjadi diantara residual.

Contoh :

Terdapat model berikut :

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \varepsilon_t \quad (3.1)$$

dimana :

$Y$  = Jumlah permintaan daging sapi

$X_2$  = Harga daging sapi

$X_3$  = Pendapatan konsumen

$X_4$  = Harga daging kambing

$t$  = Waktu

Tetapi untuk beberapa alasan regresi yang digunakan sebagai berikut :

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + V_t \quad (3.2)$$

Jika (3.1) dianggap model yang benar, maka menggunakan model (3.2) adalah memisalkan:

$$V_t = \beta_4 X_{4t} + \varepsilon_t$$

Kalau harga daging kambing mempengaruhi tingkat konsumsi daging sapi, maka unsur kesalahan  $V_t$  akan menunjukkan suatu pola sistematis. Jadi menimbulkan autokorelasi.

Pengujian sederhana mengenai hal ini adalah melakukan regresi (3.1) dan (3.2) kemudian mengecek apakah ada autokorelasi. Jika ada, yang terlibat dalam model (3.2) akan hilang apabila model (3.1) yang digunakan.

3. Bentuk fungsi yang digunakan tidak tepat.

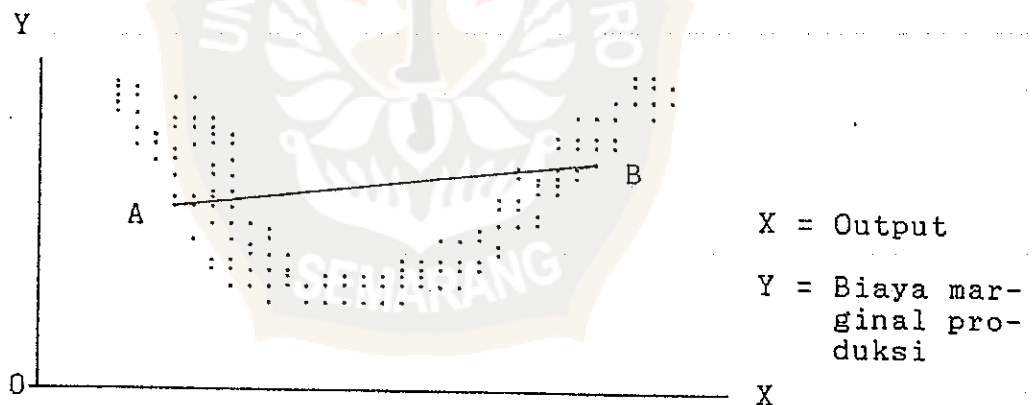
Misalkan model yang benar untuk studi biaya dan output adalah sebagai berikut :

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \varepsilon_i \quad (3.3)$$

Tetapi model yang digunakan :

$$Y_i = \alpha_1 + \alpha_2 X_i + V_i \quad (3.4)$$

Kurva dari model (3.3) yang nonlinier dan model (3.4) yang linier ditunjukkan oleh gambar 3.3



Gambar 3.3 Bias spesifikasi: Bentuk fungsional tidak tepat

Antara titik A dan B kurva biaya marginal linier secara konsisten akan menaksir terlalu tinggi (overestimate) biaya marginal sebenarnya. Sedangkan di luar batas titik A dan B kurva tadi secara konsisten akan menaksir terlalu rendah (underestimate) biaya marginal sebenarnya.

Karena unsur kesalahan :

$$V_i = X_i^2 + \varepsilon_i$$

Maka mencakup pengaruh yang terjadi secara sistematis dari output terhadap biaya marginal. Dalam hal ini unsur kesalahan  $V_i$  akan menunjukkan autokorelasi yang disebabkan oleh penggunaan fungsi yang tidak tepat.

#### 4. Keterlambatan (Lag).

Dalam regresi konsumsi terhadap pendapatan, biasanya terjadi bahwa konsumsi pada waktu  $t$  tergantung pada konsumsi waktu sebelumnya ( $t-1$ ).

Secara matematis dinyatakan :

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 Y_{t-1} + \varepsilon_t \quad (3.5)$$

dimana :  $Y$  = Konsumsi.

$X$  = Pendapatan.

Regresi (3.5) disebut sebagai autoregresi, karena salah satu variabel bebasnya merupakan variabel tak bebas dengan keterlambatan satu periode waktu.

Dasar pemikiran model (3.5) adalah sebagai berikut :

Konsumen tidak sering merubah kebiasaan berkonsumsi karena adanya faktor-faktor psikologis, teknis atau kelembagaan. Kalau unsur keterlambatan  $Y_{t-1}$  dalam model (3.5) diabaikan, maka unsur kesalahan yang dihasilkan adalah

$$V_t = \beta_3 Y_{t-1} + \varepsilon_t$$

Unsur kesalahan ini akan menunjukkan suatu pola sistematis (pola autokorelasi), disebabkan adanya pengaruh konsumsi waktu lalu terhadap konsumsi saat ini.

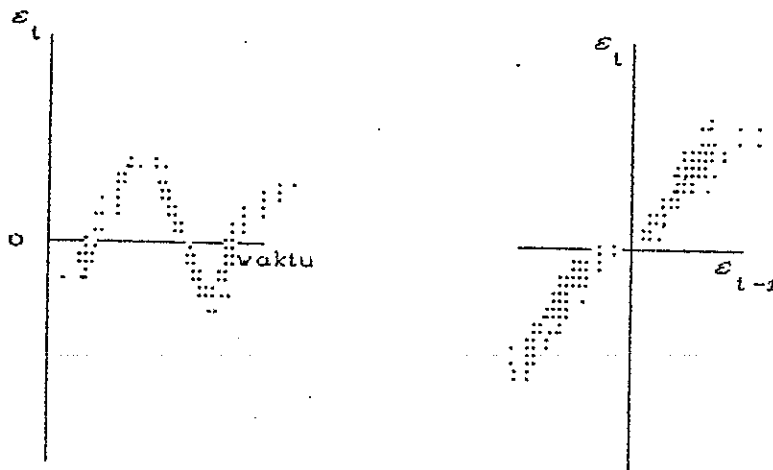
#### 5. Manipulasi Data.

Dalam analisa empiris, data mentah sering dimanipulir. Misalkan dalam regresi yang menggunakan data time series meliputi data triwulan, data tersebut biasanya merupakan penjumlahan dari 3 data bulanan, kemudian dibagi 3. Proses rata-rata ini akan menghasilkan penghalusan (smoothness) data asli, sebab fluktuasi data bulanan tidak akan terlihat lagi. Sehingga kurva data triwulan akan terlihat lebih halus daripada data bulanan. Penghalusan semacam ini akan menimbulkan pola yang sistematis dalam kesalahan yang berarti terdapat autokorelasi.

Sumber manipulasi lainnya adalah interpolasi data. Data hasil interpolasi tersebut akan menyebabkan pola sistematis (autokorelasi) yang tidak ada pada data aslinya.

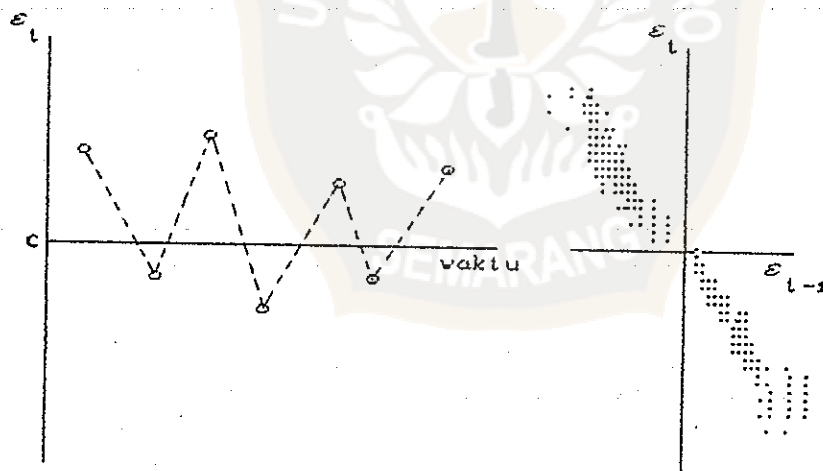
Autokorelasi dapat positif ataupun negatif, seperti terlihat pada gambar 3.4 dan gambar 3.5.





Gambar 3.4 Autokorelasi positif

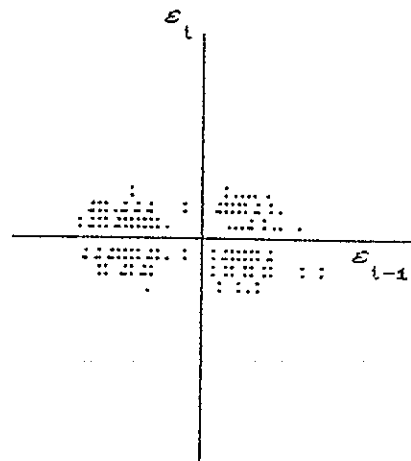
Pada autokorelasi positif beberapa unsur kesalahan secara berurutan mempunyai tanda yang sama.



Gambar 3.5 Autokorelasi negatif

Pada autokorelasi negatif, unsur kesalahan yang berurutan tandanya berubah secara bergantian.

Sedangkan plot  $\varepsilon_t$  terhadap  $\varepsilon_{t-1}$  untuk kasus non autokorelasi adalah sebagai berikut



Gambar 3.6 Pola Non-autokorelasi

### 3.2 PENGARUH AUTOKORELASI

Timbulnya masalah kesalahan yang berkorelasi atau autokorelasi didalam kesalahan, menyebabkan penaksiran dengan metode kuadrat terkecil tidak lagi menghasilkan penaksir parameter-parameter seperti yang diinginkan, yaitu tak bias linier terbaik (BLUE).

Seperti telah dikemukakan dalam pembatasan masalah pada BAB I, bahwa kesalahan  $\varepsilon$  diasumsikan memenuhi hubungan

$$\varepsilon_t = \rho \varepsilon_{t-1} + \mu_t \quad (3.2.1)$$

dimana  $\rho$  adalah koefisien autokorelasi dengan nilai  $-1 < \rho < 1$ .

Koefisien tersebut mengukur derajat hubungan antara dua variabel random  $\varepsilon_t$  dan  $\varepsilon_{t-1}$ . Untuk koefisien yang bernilai mendekati 1 atau -1 menunjukkan derajat hubungan yang tinggi, sedangkan koefisien yang bernilai

mendekati 0 menunjukkan derajat hubungan yang rendah.  
Diasumsikan pula  $\mu_t$  memenuhi syarat-syarat sebagai berikut

$$\left. \begin{aligned} E(\mu_t) &= 0 \\ \text{Var}(\mu_t) &= \sigma_\mu^2 \\ \text{Cov}(\mu_t, \mu_{t-s}) &= 0, \quad s \neq 0 \end{aligned} \right\} \quad (3.2.2)$$

untuk semua t.

Persamaan (3.2.2) disebut skema autoregresif derajat satu Markov atau autoregresif derajat pertama (AR(1)), yang menunjukkan bahwa kesalahan pada periode t ( $\varepsilon_t$ ) bergantung pada kesalahan periode sebelumnya t-1 ( $\varepsilon_{t-1}$ ).

Bentuk lengkap dari AR(1) adalah :

$$\begin{aligned} \varepsilon_t &= \rho \varepsilon_{t-1} + \mu_t \\ \varepsilon_{t-1} &= \rho \varepsilon_{t-2} + \mu_{t-1} \\ &\vdots \\ \varepsilon_{t-r} &= \rho \varepsilon_{t-(r+1)} + \mu_{t-r} \end{aligned}$$

Maka

$$\begin{aligned} \varepsilon_t &= \rho(\rho \varepsilon_{t-2} + \mu_{t-1}) + \mu_t \\ &= \rho^2 \varepsilon_{t-2} + \rho \mu_{t-1} + \mu_t \\ &= \rho^2(\rho \varepsilon_{t-3} + \mu_{t-2}) + \rho \mu_{t-1} + \mu_t \\ &= \rho^3 \varepsilon_{t-3} + \rho^2 \varepsilon_{t-2} + \rho \mu_{t-1} + \mu_t \\ &\vdots \\ &= \rho^t \varepsilon_0 + \rho^{t-1} \mu_1 + \rho^{t-2} \mu_2 + \dots + \rho^2 \mu_{t-2} \\ &\quad + \rho \mu_{t-1} + \mu_t \\ &= \mu_t + \rho \mu_{t-1} + \rho^2 \mu_{t-2} + \rho^3 \mu_{t-3} + \dots \end{aligned}$$

Jadi

$$\varepsilon_t = \sum_{r=0}^{\infty} \rho^r \mu_{t-r} \quad (3.2.3)$$

Sekarang akan dicari *mean*, *varians* dan *kovarians* dari kesalahan yang berkorelasi.

### 1. Mean

$$E(\varepsilon_t) = E \left[ \sum_{r=0}^{\infty} \rho^r \mu_{t-r} \right] = \sum_{r=0}^{\infty} \rho^r E(\mu_{t-r})$$

Dari asumsi  $E(\mu_t) = 0$  untuk semua  $t$ , maka  $E(\mu_{t-r}) = 0$

Jadi

$$E(\varepsilon_t) = 0 \quad (3.2.4)$$

### 2. Varians dan Kovarians

$$\begin{aligned} E(\varepsilon_t^2) &= E \left[ \sum_{r=0}^{\infty} \rho^r \mu_{t-r} \right]^2 \\ &= \sum (\rho^r)^2 E(\mu_{t-r})^2 \\ &= \sum (\rho^r)^2 \text{Var}(\mu_{t-r}) \\ &= \sum \rho^{2r} \sigma_{\mu}^2 \\ &= \sigma_{\mu}^2 (1 + \rho^2 + \rho^4 + \rho^6 + \dots) \end{aligned}$$

Jadi varians-nya adalah :

$$\sigma^2 = \frac{\sigma_{\mu}^2}{1 - \rho^2} \quad (3.2.5)$$

$$E(\varepsilon_t \varepsilon_{t-1})$$

$$\begin{aligned} &= E \left[ (\mu_t + \rho\mu_{t-1} + \rho^2\mu_{t-2} + \dots)(\mu_{t-1} + \rho\mu_{t-2} + \rho^2\mu_{t-3} + \dots) \right] \\ &= E \left\{ \left[ \mu_t + \rho(\mu_{t-1} + \rho\mu_{t-2} + \dots) \right] \left[ \mu_{t-1} + \rho\mu_{t-2} + \dots \right] \right\} \end{aligned}$$

$$\begin{aligned}
&= E \left\{ \mu_t \left[ \mu_{t-1} + \rho \mu_{t-2} + \dots \right] + \rho \left[ \mu_{t-1} + \rho \mu_{t-2} + \dots \right]^2 \right\} \\
&= \rho \cdot E \left[ \left( \mu_{t-1} + \rho \mu_{t-2} + \dots \right)^2 \right] \\
&= \rho E \left( \mu_{t-1}^2 + \rho^2 \mu_{t-2}^2 + \dots \right) \\
&= \rho \left[ E \left( \mu_{t-1}^2 \right) + \rho^2 E \left( \mu_{t-2}^2 \right) + \dots \right] \\
&= \rho \left[ \sigma_\mu^2 \left( 1 + \rho^2 + \rho^4 + \dots \right) \right] \\
&= \rho \frac{\sigma_\mu^2}{1 - \rho^2}
\end{aligned}$$

$$E(\varepsilon_t \varepsilon_{t-1}) = \rho \sigma^2$$

Dengan cara yang sama, maka

$$E(\varepsilon_t \varepsilon_{t-2}) = \rho^2 \sigma^2$$

$$E(\varepsilon_t \varepsilon_{t-3}) = \rho^3 \sigma^2$$

⋮

$$E(\varepsilon_t \varepsilon_{t-s}) = \rho^s \sigma^2 \quad (\text{untuk } s \neq t) \quad (3.2.6)$$

Yang merupakan kovarians dari kesalahan yang berkorelasi.

Jika (3.2.5) dan (3.2.6) disusun dalam suatu matriks,

maka :

$$E(\tilde{\varepsilon} \tilde{\varepsilon}') = V = \begin{bmatrix} \sigma^2 & \rho \sigma^2 & \rho^2 \sigma^2 & \dots & \rho^{n-1} \sigma^2 \\ \rho \sigma^2 & \sigma^2 & \rho \sigma^2 & \dots & \rho^{n-2} \sigma^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} \sigma^2 & \rho^{n-2} \sigma^2 & \rho^{n-3} \sigma^2 & \dots & \sigma^2 \end{bmatrix}$$

$$E(\underline{\varepsilon}\underline{\varepsilon}') = V = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix}$$

Merupakan matriks varians dan kovarians dari kesalahan  $\varepsilon$  yang berkorelasi.

Selanjutnya akan ditunjukkan pengaruh-pengaruh terjadinya autokorelasi terhadap sifat, interval kepercayaan, dan uji hipotesa dari taksiran kuadrat terkecil.

### 3.2.1 Pengaruh Autokorelasi Terhadap Sifat Taksiran Kuadrat Terkecil

Apabila kesalahan  $\varepsilon$  dalam model (2.1.2) berkorelasi, dan dari (2.2.1) didapat bentuk penaksir kuadrat terkecil  $\hat{\beta}$ .

$$\hat{\beta} = \beta + (X'X)^{-1} X'\varepsilon$$

maka

$$\begin{aligned} E(\hat{\beta}) &= E[\beta + (X'X)^{-1} X'\varepsilon] \\ &= E(\beta) + (X'X)^{-1} X' E(\varepsilon) \\ &= \beta \end{aligned}$$

Ini menunjukkan bahwa  $\hat{\beta}$  masih tetap merupakan penaksir tak bias dari  $\beta$  meskipun kesalahan  $\varepsilon$  berkorelasi.

Untuk varians dari  $\hat{\beta}$  yang diberikan oleh (2.2.4),

apabila kesalahan  $\varepsilon$  berkorelasi, varians tersebut menjadi

$$\begin{aligned}\text{Var}(\hat{\beta}) &= E \{ (\hat{\beta} - \beta) (\hat{\beta} - \beta)' \} \\ &= E \{ (X'X)^{-1} X' \varepsilon \varepsilon' X (X'X)^{-1} \} \\ &= (X'X)^{-1} X' E(\varepsilon \varepsilon') X (X'X)^{-1}\end{aligned}$$

jadi

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X' V X (X'X)^{-1} \quad (3.2.7)$$

Sekarang akan diselidiki apakah varians ini merupakan varians yang minimum

Pandang model sederhana sebagai berikut :

$$y_t = \beta x_t + \varepsilon_t^* \quad (3.2.8)$$

dimana :

$$\begin{aligned}y_t &= Y_t - \bar{Y} \\ x_t &= X_t - \bar{X} \\ \varepsilon_t^* &= \varepsilon_t - \bar{\varepsilon}\end{aligned}$$

Persamaan (3.2.8) disebut bentuk simpangan (deviation form).

Varians pada (2.2.4) untuk model sederhana diatas adalah

$$\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

dimana :

$$(X'X) = [ x_1 \quad x_2 \quad \dots \quad x_n ] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1^2 + x_2^2 + \dots + x_n^2 = \sum_{t=1}^n x_t^2$$

$$(X'X)^{-1} = \frac{1}{\sum_{t=1}^n x_t^2}$$

Jadi

$$\text{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{t=1}^n x_t^2} \quad (3.2.9)$$

Tetapi varians pada (3.2.7) untuk model sederhana di atas adalah :

$$\text{Var}(\hat{\beta}) = (X'X)^{-1} X'V X(X'X)^{-1}$$

$$\text{Var}(\hat{\beta}) = \left[ \frac{1}{\sum_{t=1}^n x_t^2} \right] \sigma^2 [x_1 \quad x_2 \quad \dots \quad x_n]$$

$$\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \left[ \frac{1}{\sum_{t=1}^n x_t^2} \right]$$

$$= \frac{\sigma^2}{\left[ \sum_{t=1}^n x_t^2 \right]^2} \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$$

$$\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$



$$\begin{aligned}
&= \frac{\sigma^2}{\left(\sum_{l=1}^n x_l^2\right)^2} \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 + \rho x_2 + \dots + \rho^{n-1} x_n \\ \rho x_1 + x_2 + \dots + \rho^{n-2} x_n \\ \vdots \\ \rho^{n-1} x_1 + \rho^{n-2} x_2 + \dots + x_n \end{bmatrix} \\
&= \frac{\sigma^2}{\left(\sum_{l=1}^n x_l^2\right)^2} \left[ x_1(x_1 + \rho x_2 + \dots + \rho^{n-1} x_n) + \right. \\
&\quad x_2(\rho x_1 + x_2 + \dots + \rho^{n-2} x_n) + \dots + x_n(\rho^{n-1} x_1 + \\
&\quad \left. \rho^{n-2} x_2 + \dots + x_n) \right] \\
&= \frac{\sigma^2}{\left(\sum_{l=1}^n x_l^2\right)^2} \left[ x_1^2 + \rho x_1 x_2 + \dots + \rho^{n-1} x_1 x_n + \right. \\
&\quad \left. \rho x_1 x_2 + x_2^2 + \dots + \rho^{n-2} x_2 x_n + \dots + \rho^{n-1} x_1 x_n \right. \\
&\quad \left. + \rho^{n-2} x_2 x_n + \dots + x_n^2 \right] \\
&= \frac{\sigma^2}{\left(\sum_{l=1}^n x_l^2\right)^2} \left[ x_1^2 + \dots + x_n^2 + 2\rho x_1 x_2 + \dots + 2\rho x_{n-1} x_n \right. \\
&\quad \left. + \dots + 2\rho^2 x_1 x_3 + \dots + 2\rho^2 x_{n-2} x_n + \dots + 2\rho^{n-2} x_2 x_n \right. \\
&\quad \left. + 2\rho^{n-1} x_1 x_n \right] \\
&= \frac{\sigma^2}{\left(\sum_{l=1}^n x_l^2\right)^2} \left[ \sum_{l=1}^n x_l^2 + 2\rho \sum_{l=1}^{n-1} x_l x_{l+1} + 2\rho^2 \sum_{l=1}^{n-2} x_l x_{l+2} \right. \\
&\quad \left. + \dots + 2\rho^{n-1} x_1 x_n \right]
\end{aligned}$$

$$\text{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{t=1}^n x_t^2} \left[ 1 + 2\rho \frac{\sum_{t=1}^{n-1} x_t x_{t+1}}{\sum_{t=1}^n x_t^2} + 2\rho^2 \frac{\sum_{t=1}^{n-2} x_t x_{t+2}}{\sum_{t=1}^n x_t^2} \right. \\ \left. + \dots + 2\rho^{n-1} \frac{x_1 x_n}{\sum_{t=1}^n x_t^2} \right] \quad (3.2.10)$$

yang selanjutnya dinotasikan dengan  $\text{Var}(\hat{\beta})_{\text{AR}(1)}$ , yaitu varians  $\hat{\beta}$  bila kesalahan berkorelasi dengan autoregresif derajat pertama.

Dalam praktek, kebanyakan data menunjukkan  $\rho$  positif, sehingga

$$\text{Var}(\hat{\beta})_{\text{AR}(1)} > \text{Var}(\hat{\beta})$$

Artinya varians taksiran kuadrat terkecil bukan merupakan varians yang minimum lagi bila kesalahan berkorelasi.

### 3.2.2 Pengaruh Autokorelasi Terhadap Interval Kepercayaan dan Uji Hipotesa

Karena perhitungan interval kepercayaan dan pengujian hipotesa menggunakan penaksir varians  $\hat{\sigma}^2$ , maka pengaruh autokorelasi terhadap interval kepercayaan dan pengujian hipotesa akan ditinjau melalui sifat penaksir varians  $\hat{\sigma}^2$  yang seharusnya merupakan penaksir tak bias dari  $\sigma^2$ .

Dari (2.2.13) diketahui :

$$\hat{\sigma}^2 = \frac{e'e}{n-p}$$

Untuk model sederhana :

$$y_i = \beta x_i + \varepsilon_i^* = \hat{\beta} x_i + e_i$$

dimana  $y_i = Y_i - \bar{Y}$  ;  $x_i = X_i - \bar{X}$  ;  $\varepsilon_i^* = \varepsilon_i - \bar{\varepsilon}$

Penaksir  $\sigma^2$  yaitu  $\hat{\sigma}^2$ , menjadi

$$\begin{aligned} \hat{\sigma}^2 &= \frac{e'e}{n-1} \\ &= \frac{\sum e^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^n (y_i - \hat{\beta} x_i)^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n (\beta x_i + \varepsilon_i^* - \hat{\beta} x_i)^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n [-(\hat{\beta} - \beta) x_i + \varepsilon_i^*]^2 \\ &= \frac{1}{n-1} [(\hat{\beta} - \beta)^2 \sum x_i^2 + \sum \varepsilon_i^{*2} - 2(\hat{\beta} - \beta) \sum x_i \varepsilon_i^*] \\ &= \frac{1}{n-1} [\sum \varepsilon_i^{*2} + (\hat{\beta} - \beta)^2 \sum x_i^2 - 2(\hat{\beta} - \beta) \sum x_i (y_i - \beta x_i)] \\ &= \frac{1}{n-1} [\sum \varepsilon_i^{*2} + (\hat{\beta} - \beta)^2 \sum x_i^2 - 2(\hat{\beta} - \beta) \sum x_i (\hat{\beta} x_i + e_i - \beta x_i)] \\ &= \frac{1}{n-1} [\sum \varepsilon_i^{*2} + (\hat{\beta} - \beta)^2 \sum x_i^2 - 2(\hat{\beta} - \beta) \sum x_i [(\hat{\beta} - \beta) x_i + e_i]] \\ &= \frac{1}{n-1} [\sum \varepsilon_i^{*2} + (\hat{\beta} - \beta)^2 \sum x_i^2 - 2(\hat{\beta} - \beta)^2 \sum x_i^2 - 2(\hat{\beta} - \beta) \sum x_i e_i] \\ &= \frac{1}{n-1} [\sum \varepsilon_i^{*2} - (\hat{\beta} - \beta)^2 \sum x_i^2] \quad (3.2.11) \end{aligned}$$

Sekarang akan dihitung  $E(\hat{\sigma}^2)$

$$E(\hat{\sigma}^2) = \frac{1}{n-1} \left[ \sum_{t=1}^n E(\varepsilon_t^{*2}) - (\sum x_t^2) \text{Var}(\hat{\beta})_{AR(1)} \right]$$

dimana dari (3.2.10) diketahui bahwa :

$$\begin{aligned} \text{Var}(\hat{\beta})_{AR(1)} &= \frac{\sigma^2}{\sum_{t=1}^n x_t^2} \left[ 1 + 2\rho \frac{\sum_{t=1}^{n-1} x_t x_{t+1}}{\sum_{t=1}^n x_t^2} + 2\rho^2 \frac{\sum_{t=1}^{n-2} x_t x_{t+2}}{\sum_{t=1}^n x_t^2} \right. \\ &\quad \left. + \dots + 2\rho^{n-1} \frac{x_1 x_n}{\sum_{t=1}^n x_t^2} \right] \\ \text{Var}(\hat{\beta})_{AR(1)} &= \frac{\sigma^2}{\sum_{t=1}^n x_t^2} + \frac{2\sigma^2}{\left( \sum_{t=1}^n x_t^2 \right)^2} \left[ \rho \sum_{t=1}^{n-1} x_t x_{t+1} + \rho^2 \sum_{t=1}^{n-2} x_t x_{t+2} \right. \\ &\quad \left. + \dots + \rho^{n-1} x_1 x_n \right] \quad (i) \end{aligned}$$

Langkah selanjutnya adalah menghitung  $\sum_{t=1}^n E(\varepsilon_t^{*2})$

$$\begin{aligned} E(\varepsilon_t^{*2}) &= E(\varepsilon_t - \bar{\varepsilon})^2 = E(\varepsilon_t^2 + \bar{\varepsilon}^2 - 2\varepsilon_t \bar{\varepsilon}) \\ &= E(\varepsilon_t^2) + E\left[\frac{1}{n} \sum \varepsilon_t\right]^2 - 2E\left[\varepsilon_t \left(\frac{1}{n} \sum \varepsilon_t\right)\right] \quad (ii) \end{aligned}$$

$$\begin{aligned} E\left[\frac{1}{n} \sum \varepsilon_t\right]^2 &= E\left[\frac{1}{n} (\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n)\right]^2 \\ &= \frac{1}{n^2} E[\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2 + 2\varepsilon_1 \varepsilon_2 + \dots + 2\varepsilon_1 \varepsilon_n \\ &\quad + 2\varepsilon_2 \varepsilon_3 + \dots + 2\varepsilon_2 \varepsilon_n + \dots + 2\varepsilon_{n-1} \varepsilon_n] \end{aligned}$$

$$= \frac{1}{n^2} \left\{ E(\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2) + [2E(\varepsilon_1 \varepsilon_2) + 2E(\varepsilon_2 \varepsilon_3) + \dots + 2E(\varepsilon_{n-1} \varepsilon_n)] + [2E(\varepsilon_1 \varepsilon_3) + 2E(\varepsilon_2 \varepsilon_4) + \dots + 2E(\varepsilon_{n-2} \varepsilon_n)] + \dots + 2E(\varepsilon_1 \varepsilon_n) \right\}$$

$$E\left[\frac{1}{n} \sum \varepsilon_t\right]^2 = \frac{1}{n^2} [n\sigma^2 + 2(n-1)\rho\sigma^2 + 2(n-2)\rho^2\sigma^2 + \dots + 2\rho^{n-1}\sigma^2]$$

Hasil ini dimasukkan ke dalam persamaan (ii), sehingga

$$\begin{aligned} E(\varepsilon_t^{*2}) &= E(\varepsilon_t^2) + \frac{1}{n^2} [n\sigma^2 + 2(n-1)\rho\sigma^2 + 2(n-2)\rho^2\sigma^2 + \dots + 2\rho^{n-1}\sigma^2] \\ &\quad - \frac{2}{n} E[\varepsilon_t(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t + \dots + \varepsilon_n)] \\ &= \sigma^2 + \frac{\sigma^2}{n} + \frac{2\sigma^2}{n^2} [(n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1}] \\ &\quad - \frac{2}{n} [E(\varepsilon_t \varepsilon_1) + E(\varepsilon_t \varepsilon_2) + \dots + E(\varepsilon_t)^2 + \dots + E(\varepsilon_t \varepsilon_n)] \\ &= \frac{(n+1)\sigma^2}{n} + \frac{2\sigma^2}{n^2} [(n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1}] \\ &\quad - \frac{2}{n} [\rho^{t-1}\sigma^2 + \rho^{t-2}\sigma^2 + \dots + \sigma^2 + \dots + \rho^{n-t}\sigma^2] \\ &= \frac{(n+1)\sigma^2}{n} + \frac{2\sigma^2}{n^2} [(n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1}] \\ &\quad - \frac{2\sigma^2}{n} [\rho^{t-1} + \rho^{t-2} + \dots + 1 + \dots + \rho^{n-t}] \end{aligned}$$

Maka

$$\sum_{t=1}^n E(\varepsilon_t^{*2}) = \sum_{t=1}^n \left\{ \frac{(n+1)\sigma^2}{n} + \frac{2\sigma^2}{n^2} [(n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1}] - \frac{2\sigma^2}{n} [\rho^{t-1} + \rho^{t-2} + \dots + 1 + \dots + \rho^{n-t}] \right\}$$

$$\begin{aligned}
& - \frac{2\sigma^2}{n} [\rho^{t-1} + \rho^{t-2} + \dots + 1 + \dots + \rho^{n-t}] \} \\
& = (n+1)\sigma^2 + \frac{2\sigma^2}{n} [(n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1}] \\
& - \frac{2\sigma^2}{n} \sum_{t=1}^n [\rho^{t-1} + \rho^{t-2} + \dots + 1 + \dots + \rho^{n-t}]
\end{aligned}$$

Akan dihitung suku terakhir

$$\begin{aligned}
& \sum_{t=1}^n [\rho^{t-1} + \rho^{t-2} + \dots + 1 + \dots + \rho^{n-t}] \\
& = (1 + \rho + \rho^2 + \dots + \rho^{n-1}) + (\rho + 1 + \rho + \dots + \\
& \rho^{n-2}) + (\rho^2 + \rho + 1 + \dots + \rho^{n-3}) + \dots + \\
& (\rho^{n-1} + \rho^{n-2} + \rho^{n-3} + \dots + 1) \\
& = (1+1 + \dots + 1) + (\rho + \rho + \dots + \rho) + \dots + \\
& (\rho^{n-2} + \dots + \rho^{n-2}) + \rho^{n-1} + \rho^{n-1} \\
& = n + 2(n-1)\rho + 2(n-2)\rho^2 + \dots + 4\rho^{n-2} + 2\rho^{n-1}
\end{aligned}$$

Hasil ini dimasukkan ke persamaan  $\sum_{t=1}^n E(\varepsilon_t^{*2})$ , sehingga

$$\begin{aligned}
\sum_{t=1}^n E(\varepsilon_t^{*2}) & = (n+1)\sigma^2 + \frac{2\sigma^2}{n} \left\{ (n-1)\rho + (n-2)\rho^2 + \dots + \rho^{n-1} \right. \\
& \quad \left. - [n + 2(n-1)\rho + 2(n-2)\rho^2 + \dots + 4\rho^{n-2} + 2\rho^{n-1}] \right\} \\
& = (n+1)\sigma^2 + \frac{2\sigma^2}{n} [-n - (n-1)\rho - (n-2)\rho^2 - \dots - 2\rho^{n-2} - \rho^{n-1}] \\
& = (n+1)\sigma^2 - \frac{2\sigma^2}{n} [n + n(n-1)\rho + (n-2)\rho^2 + \dots + 2\rho^{n-2} + \rho^{n-1}]
\end{aligned}$$

Misalkan suku terakhir dari persamaan diatas

$$n + (n-1)\rho + (n-2)\rho^2 + \dots + 2\rho^{n-2} + \rho^{n-1} = A$$

Jika

$$\begin{aligned}
 A - \rho A &= n + (n-1)\rho + (n-2)\rho^2 + \dots + 2\rho^{n-2} + \rho^{n-1} - n\rho - (n-1)\rho^2 \\
 &\quad - (n-2)\rho^3 - \dots - 2\rho^{n-1} - \rho^n \\
 A(1-\rho) &= n - \rho - \rho^2 - \dots - \rho^{n-1} - \rho^n \\
 &= n - (\rho + \rho^2 + \dots + \rho^n) \\
 &= n - \frac{\rho(1-\rho^n)}{(1-\rho)} \\
 &= \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)} \\
 A &= \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2}
 \end{aligned}$$

Sehingga

$$\begin{aligned}
 \sum_{t=1}^n E(\varepsilon_t^{*2}) &= (n+1)\sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} \right] \\
 &= n\sigma^2 + \sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} \right] \\
 &= n\sigma^2 + \sigma^2 - 2\sigma^2 + 2\sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} \right] \\
 &= n\sigma^2 - \sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} - n \right] \\
 &= (n-1)\sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n(1-\rho) - \rho(1-\rho^n) - n(1-\rho)^2}{(1-\rho)^2} \right] \\
 \sum_{t=1}^n E(\varepsilon_t^{*2}) &= (n-1)\sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n\rho(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} \right] \quad \text{(iii)}
 \end{aligned}$$

Kemudian hasil persamaan (iii) dan (i) dimasukkan pada  $E(\hat{\sigma}^2)$ , sehingga :

$$\begin{aligned}
 E(\hat{\sigma}^2) &= \frac{1}{n-1} \left\{ \left[ (n-1)\sigma^2 - \frac{2\sigma^2}{n} \left[ \frac{n\rho(1-\rho) - \rho(1-\rho^n)}{(1-\rho)^2} \right] \right] \right. \\
 &\quad - \left[ \left( \sum_{t=1}^n x_t^2 \right) \left[ \frac{\sigma^2}{n \sum_{t=1}^n x_t^2} + \frac{2\sigma^2}{\left( \sum_{t=1}^n x_t^2 \right)^2} \left[ \rho \sum_{t=1}^{n-1} x_t x_{t+1} + \dots + \right. \right. \right. \\
 &\quad \left. \left. \left. \rho^{n-1} x_1 x_n \right] \right] \right\} \\
 &= \frac{(n-1)\sigma^2}{(n-1)} - \frac{2\sigma^2}{(n-1)} \left[ \frac{n\rho(1-\rho) - \rho(1-\rho^n)}{n(1-\rho)^2} \right] - \frac{\sigma^2}{n-1} \\
 &\quad - \frac{2\sigma^2}{(n-1) \sum_{t=1}^n x_t^2} \left[ \rho \sum_{t=1}^{n-1} x_t x_{t+1} + \dots + \rho^{n-1} x_1 x_n \right] \\
 &= \sigma^2 - \frac{2\sigma^2}{(n-1)} \left[ \frac{n\rho(1-\rho) - \rho(1-\rho^n)}{n(1-\rho)^2} + \frac{1}{2} \right] \\
 &\quad - \frac{2\sigma^2}{(n-1) \sum_{t=1}^n x_t^2} \left[ \rho \sum_{t=1}^{n-1} x_t x_{t+1} + \dots + \rho^{n-1} x_1 x_n \right] \\
 &= \sigma^2 - \frac{2\sigma^2}{(n-1)} \left[ \frac{n\rho(1-\rho) - \rho(1-\rho^n) + \frac{1}{2}(n(1-\rho)^2)}{n(1-\rho)^2} \right]
 \end{aligned}$$



$$= \frac{2\sigma^2}{(n-1)\sum x_l^2} \left[ \rho \sum_{l=1}^{n-1} x_l x_{l+1} + \dots + \rho^{n-1} x_1 x_n \right]$$

$$= \sigma^2 - \frac{2\sigma^2}{(n-1)} \left[ \frac{\frac{1}{2}(n(1-\rho^2)) - \rho(1-\rho^n)}{n(1-\rho)^2} \right]$$

$$- \frac{2\sigma^2}{(n-1)\sum x_l^2} \left[ \rho \sum_{l=1}^{n-1} x_l x_{l+1} + \dots + \rho^{n-1} x_1 x_n \right]$$

$$E(\hat{\sigma}^2) = \sigma^2 - \frac{2\sigma^2}{(n-1)} \left\{ \left[ \frac{\frac{1}{2}(n(1-\rho^2)) - \rho(1-\rho^n)}{n(1-\rho)^2} \right] + \left[ \frac{\rho \sum_{l=1}^{n-1} x_l x_{l+1} + \dots + \rho^{n-1} x_1 x_n}{\sum x_l^2} \right] \right\} \quad (3.2.12)$$

Jadi

$$E(\hat{\sigma}^2) \neq \sigma^2$$

Ini berarti penaksir varians  $\hat{\sigma}^2$  pada (3.2.11) bukan merupakan penaksir tak bias dari  $\sigma^2$ .

Akibatnya interval kepercayaan dan pengujian hipotesa dari taksiran kuadrat terkecil akan membuat kesimpulan yang salah. Artinya interval kepercayaan atau daerah penerimaan akan menjadi lebih pendek atau lebih lebar, tergantung arah biasnya negatif atau positif.

Arah bias negatif bila selisih  $E(\hat{\sigma}^2) - \sigma^2$  adalah negatif, sehingga interval kepercayaan atau daerah penerimaan akan menjadi lebih pendek. Sedang arah bias

positif bila selisih  $E(\hat{\sigma}^2) - \sigma^2$  adalah positif, sehingga interval kepercayaan atau daerah penerimaan akan menjadi lebih besar dari yang seharusnya.

