## BAB II

## MODEL LAIRD-WARE DENGAN ERROR AR(1)

Pada bab II ini akan dibahas mengenai model Laird-Ware dengan error AR(1) yaitu suatu model untuk data longitudinal yang saling berkorelasi AR(1). Jadi pada bagian ini akan dijelaskan hal-hal yang berhubungan dengan model Laird-Ware dengan error AR(1) yaitu penjelasan tentang proses AR(1), data longitudinal, model Laird-Ware dan model Laird-Ware dengan error AR(1).

### 2.1. Proses AR(1)

Pokok pembahasan dalam tugas akhir ini adalah penaksiran parameter dari model Laird-Ware dengan error AR(1). Untuk itu terlebih dahulu akan dijelaskan tentang proses AR(1) secara umum.

Proses AR(1) dapat dipandang sebagai sebuah proses dengan model

$$Z_t = \phi\, Z_{t-1} + a_t \tag{2.1}$$

dengan suku sesatan $a_t$ berdistribusi normal dengan rata-rata nol dan kovarian $\sigma_a^2$ dan model ini dianggap stasioner. Proses AR(1) merupakan sebuah proses dimana nilai sekarang suatu proses dinyatakan sebagai nilai yang lalu ditambah satu sesatan random sekarang. Jadi dapat dipandang $Z_t$ dibuat dalam model regresi pada satu nilai Z yang lalu ($Z_{t-1}$).

Karena $a_t$ independen dengan $Z_{t-1}$, maka variansinya adalah:

$$\text{var}(Z_t) = \phi^2 \,\text{var}(Z_{t-1}) + \text{var}(a_t)$$

dengan $\text{var}(Z_t) = \text{var}(Z_{t-1}) = \gamma_0$ dan $\text{var}(a_t) = \sigma_a^2$, maka didapat

$$\gamma_0 = \phi^2 \;\gamma_0 \;+ \sigma_a^2$$

$$(1 - \phi^2)\;\gamma_0 = \sigma_a^2$$

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi^2} \qquad (2.2)$$

dan supaya $\gamma_0$ berhingga dan tidak negatif, haruslah

$$-1 < \phi < 1 \qquad (2.3)$$

Ketidaksamaan (2.3) inilah yang merupakan syarat supaya runtun waktunya stasioner. Jadi solusi dari $Z_t = \phi\, Z_{t-1} + a_t$ akan stasioner jika dan hanya jika $|\phi| < 1$. Syarat $|\phi| < 1$ ini adalah syarat stasioneritas. Umumnya $\phi$ adalah dalam range/jangkauan $0 \leq \phi < 1$, tetapi nilai negatif dari $\phi$ adalah juga mungkin terjadi ketika pengamatan-pengamatan yang berdampingan adalah berkorelasi negatif. Sebuah proses stasioner mempunyai suatu varian yang konstan atas waktu. Persamaan (2.2) merupakan persamaan hubungan antara varian proses $Z_t$ dan varian sesatan $a_t$.

Fungsi kovarian dari suatu runtun waktu stasioner hanya tergantung pada selisih waktu antara dua pengamatan. Fungsi kovarian didefinisikan sebagai:

$$\gamma_h = E(Z_t\, Z_{t-h})$$

dimana h adalah selisih waktu antara dua pengamatan yang disebut "lag" waktu. Fungsi kovarian adalah simetrik, yaitu $\gamma_{-h} = \gamma_h$. Fungsi kovarian pada lag nol (selisih waktu antara dua pengamatan adalah nol) merupakan varian proses, jadi :

$$E(Z_t^2) = \gamma_0$$

Fungsi korelasi dari suatu runtun waktu stasioner adalah fungsi kovarian dibagi dengan varian proses. Dengan mengalikan kedua sisi persamaan (2.1) dengan $Z_{t-h}$ ($h = 1, 2, \ldots$) dan mengambil nilai ekspektasinya diperoleh:

$$E(Z_t Z_{t-h}) = \phi\, E(Z_{t-1} Z_{t-h}) + E(a_t Z_{t-h})$$

Karena runtun waktu diasumsikan stasioner dan karena $a_t$ independen dengan $Z_{t-h}$, maka didapat

$$\gamma_h = \phi\, \gamma_{h-1} + 0$$

$$\gamma_h = \phi\, \gamma_{h-1} \qquad \text{untuk } h = 1, 2, \ldots$$

Untuk $h = 1$, didapat $\gamma_1 = \phi\, \gamma_0$

Untuk $h = 2$ didapat $\gamma_2 = \phi\, \gamma_1 = \phi^2\, \gamma_0$

Untuk $h = 3$ didapat $\gamma_3 = \phi\, \gamma_2 = \phi^3\, \gamma_0$

Sehingga secara umum $\gamma_h = \phi^h\, \gamma_0$

Maka jika fungsi korelasi dari suatu runtun waktu stasioner adalah fungsi kovarian dibagi dengan varian proses, maka didapat

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{\phi^h\, \gamma_0}{\gamma_0} = \phi^h \quad \text{untuk } h = 0, 1, 2, \ldots \qquad (2.4)$$

dimana

$\rho_h$ : korelasi antara pengamatan lag h

$\gamma_h$ : kovarian antara pengamatan lag h

$\gamma_0$ : kovarian antara pengamatan lag 0 = varian proses

Sekarang akan dibahas penaksiran parameter $\phi$ dari model AR(1)

$$Z_t = \phi\, Z_{t-1} + a_t$$

berdasarkan pada runtun waktu sampel teramati $Z_1, Z_2, \ldots, Z_n$. Untuk model ini, dari persamaan (2.4) dipunyai hubungan

$$\rho_1 = \phi.$$

Dalam penaksiran parameter, $\rho_1$ disamakan dengan $r_1$, autokorelasi sampel lag 1. Maka dapat ditaksir $\phi$ dengan:

$$\phi = r_1 \qquad (2.5)$$

dimana

$$r_1 = \frac{C_1}{C_0}$$

$$r_1 = \frac{\frac{1}{n}\sum_{t=1}^{n} (Z_t - \bar{Z})(Z_{t-1} - \bar{Z})}{\frac{1}{n}\sum_{t=1}^{n} (Z_t - \bar{Z})^2} \qquad (2.6)$$

dengan

$C_1 = \frac{1}{n}\sum_{t=1}^{n} (Z_t - \bar{Z})(Z_{t-1} - \bar{Z})$, adalah kovarian sampel lag 1

$C_0 = \frac{1}{n}\sum_{t=1}^{n} (Z_t - \bar{Z})^2$, adalah varian sampel

$\overline{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t$ , adalah rata-rata $Z_t$

## 2.2. Data Longitudinal

Pada bagian ini akan dibahas mengenai pengertian data longitudinal, serta asumsi-asumsi pada data longitudinal yang sangat berpengaruh dalam bentuk pemodelannya yaitu pada bentuk model Laird-Ware yang akan digunakan.

Data longitudinal adalah suatu data pengamatan yang dilakukan dalam bentangan waktu pendek yang telah ditentukan. Data longitudinal diperoleh dari pengamatan-pengamatan dengan kriteria-kriteria sebagai berikut : perlakuan dinyatakan sebagai grup, tiap grup diamati subyek-subyek yang dikenai perlakuan dan pengaruh grup terhadap subyek (respon dari tiap subyek dalam grup) diamati secara berkelanjutan (berulang-ulang) dalam bentang waktu pendek yang telah ditentukan. Subyek-subyek tersebut diasumsikan merupakan sampel random dari suatu populasi.

Data longitudinal dapat dinyatakan dengan $Y_{ijk}$ , dimana indek i menyatakan grup ke-i, indek j menyatakan subyek ke-j dalam grup, dan indek k menyatakan urutan pengamatan ke-k. Nilai-nilai indek i, j dan k adalah:

$i = 1, 2, \ldots, a$ dengan a : jumlah grup

$j = 1, 2, \ldots, b$ dengan b : jumlah subyek

$k = 1, 2, \ldots, n$ dengan n : jumlah pengamatan

Diasumsikan grup-grup saling bebas satu sama lain, subyek-subyek didalam grup juga saling bebas, dan data $Y_{ijk}$ diamati secara terurut sesuai urutan pengamatan. Data-data $Y_{ijk}$ ada yang saling bebas namun ada juga yang saling berkorelasi. Jika data-data $Y_{ijk}$ saling berkorelasi, maka data-data $Y_{ijk}$ merupakan suatu runtun waktu pendek. Karena subyek-subyek di dalam grup saling bebas, maka data-data pengamatan antar subyek-subyek tersebut juga saling bebas. Jadi kemungkinan adanya data-data yang saling berkorelasi hanya pada data-data pengamatan dari satu subyek saja.

Karena pada data longitudinal tiap subyek diamati lebih dari sekali secara terurut sesuai waktu pengamatan, maka mungkin bahwa data hasil pengamatan berulang tersebut ($Y_{ijk}$) saling berkorelasi runtun artinya data pengamatan pada suatu waktu dipengaruhi oleh data pengamatan waktu sebelumnya, misalnya seperti yang akan dibahas dalam Tugas Akhir ini, yaitu data-data pengamatan $Y_{ijk}$ saling berkorelasi runtun AR(1). Tiap runtun waktu terdiri dari pengamatan-pengamatan yang diambil dari sebuah subyek. Jadi dalam analisis data longitudinal yang saling berkorelasi struktur datanya adalah sejumlah runtun waktu pendek dari pengamatan berulang sejumlah subyek. Namun data pengamatan dari subyek-subyek yang berbeda (data antar subyek) diasumsikan independen (saling bebas), karena subyek-subyek tersebut saling bebas. Jadi tidak ada korelasi antara data-data pengamatan dari subyek-subyek yang berbeda. Misalnya data-data pengamalan $Y_{i1k}$ dan data-data pengamatan $Y_{i2k}$ akan saling bebas karena dalam grup i tersebut, subyek 1 dan subyek 2 saling bebas.

Disini akan diberikan contoh data longitudinal dari pengamatan 2 grup yaitu grup A dan grup B dengan masing-masing grup akan diamati 3 subyek. Untuk grup A terdiri dari 3 subyek yaitu subyek 1A, subyek 2A dan subyek 3A. Sedangkan untuk grup B terdiri dari pengamatan untuk subyek 1B, subyek 2B dan subyek 3B. Tiap-tiap subyek dalam grup-grup tersebut diamati 4 kali pengamatan. Contoh data longitudinal ini dapat digambarkan dalam tabel 2.1 berikut:

Tabel 2.1. Gambaran Data Longitudinal

| | | Urutan Pengamatan | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| Grup A | Subyek 1A | $Y_{111}$ | $Y_{112}$ | $Y_{113}$ | $Y_{114}$ |
| | Subyek 2A | $Y_{121}$ | $Y_{122}$ | $Y_{123}$ | $Y_{124}$ |
| | Subyek 3A | $Y_{131}$ | $Y_{132}$ | $Y_{133}$ | $Y_{134}$ |
| Grup B | Subyek 1B | $Y_{211}$ | $Y_{212}$ | $Y_{213}$ | $Y_{214}$ |
| | Subyek 2B | $Y_{221}$ | $Y_{222}$ | $Y_{223}$ | $Y_{224}$ |
| | Subyek 3B | $Y_{231}$ | $Y_{232}$ | $Y_{233}$ | $Y_{234}$ |

## 2.3. Model Laird-Ware

Suatu model linier mixed (campuran) yang umum untuk data longitudinal telah dikemukakan oleh Laird dan Ware, yang kemudian model tersebut disebut model Laird-Ware. Model Laird-Ware untuk model pengamatan pada subyek ke-j berbentuk:

$$\mathbf{Y}_j = \mathbf{X}_j\,\beta + \mathbf{Z}_j\,\gamma_j + \varepsilon_j \qquad (2.7)$$

dimana:

$\mathbf{Y}_j$ : vektor kolom $n_j$ x 1 dari variabel respon (pengamatan) pada subyek ke-j, dimana $n_j$ adalah jumlah pengamatan pada subyek ke-j.

$\mathbf{X}_j$ : matrik $n_j$ x b sebagai koefisien parameter $\beta$.

$\beta$ : vektor b x 1 sebagai efek tetap, dimana b adalah jumlah efek tetap.

$\mathbf{Z}_j$ : matrik $n_j$ x g sebagai koefisien parameter efek random $\gamma_j$

$\gamma_j$ : vektor g x 1 sebagai efek random, dimana g adalah jumlah efek random. $\gamma_j$ diasumsikan berdistribusi normal independen atas subyek-subyek dengan rata-rata nol dan matrik kovarian $\sigma^2$B, dimana B adalah matrik kovarian antar subyek.

$\varepsilon_j$ : vektor kolom $n_j$ x 1 dari error pengamatan dalam subyek ke-j. $\varepsilon_j$ diasumsikan berdistribusi normal dengan rata-rata nol dan matrik kovarian $\sigma^2\mathbf{W}_j$, dimana $\mathbf{W}_j$ adalah matrik kovarian dari $\varepsilon_j$. Error-error pengamatan dalam subyek ke-j ($\varepsilon_j$ ) ini ada yang saling bebas, namun juga bisa terjadi $\varepsilon_j$ ini saling berkorelasi runtun. Korelasi runtun dari $\mathbf{Y}_j$ pada pemodelan akan menjadi bagian dari struktur error modelnya. Jadi jika data-data pengamatan pada subyek ke-j ($\mathbf{Y}_j$ ) saling bebas, maka error model pengamatan dalam subyek ke-j ($\varepsilon_j$ ) juga saling bebas. Jika $\varepsilon_j$ ini saling bebas maka $\mathbf{W}_j$ = I, suatu matrik identitas. Tapi jika data-data pengamatan pada subyek ke-j ($\mathbf{Y}_j$ ) saling berkorelasi, maka error model pengamatan dalam subyek ke-j ($\varepsilon_j$) juga akan saling berkorelasi. Jadi jika $Y_j$ saling berkorelasi AR(1), maka $\varepsilon_j$ juga saling berkorelasi AR(1). Sehingga $\phi_j$ pada proses AR(1) untuk $\varepsilon_j$ sama dengan $\phi_j$ untuk proses AR(1) untuk $Y_j$.

## 2.4. Model Laird-Ware dengan Error AR(1)

Karena pada data longitudinal tiap subyek diamati lebih dari sekali, maka mungkin bahwa data hasil pengamatan berulang tersebut berkorelasi runtun. Tiap runtun waktu terdiri dari pengamatan-pengamatan yang diambil pada sebuah subyek. Korelasi runtun ini dalam pemodelan nanti adalah bagian dari struktur error. Pada bagian ini akan dibahas model Laird-Ware dengan $\varepsilon_j$ saling berkorelasi runtun AR(1), yaitu $\varepsilon_j$ merupakan suatu proses Autoregresi orde pertama atau AR(1), karena data pengamatannya saling berkorelasi runtun AR(1).

Pada model Laird-Ware dengan error-error pengamatan dalam subyek ke-j ($\varepsilon_j$) merupakan proses AR(1), ditetapkan tidak ada kovariansi antar subyek sehingga $\mathbf{B} = \mathbf{0}$. Hal ini dikarenakan adanya korelasi runtun AR(1) pada $\varepsilon_j$ hanya terjadi pada pengamatan-pengamatan dari sebuah subyek saja, yaitu korelasi runtun AR(1) pada $\varepsilon_j$ hanya terjadi pada pengamatan-pengamatan dari subyek ke-j saja. Jadi dalam hal ini tidak melibatkan pembahasan dalam analisa data dengan pengamatan-pengamatan dari banyak subyek. Sehingga karena yang dibahas adalah pengamatan-pengamatan dari satu subyek saja, maka disini tidak ada matrik kovariansi antar subyek sehingga

$\mathbf{B} = \mathbf{0}$. Penetapan $\mathbf{B} = \mathbf{0}$ ini menyebabkan hilangnya bagian $\mathbf{Z}_j \gamma_j$ dari model (2.7). Sehingga model Laird-Ware untuk model pengamatan pada subyek ke-j, dengan error-error pengamatan dalam subyek ke-j ($\varepsilon_j$ ) berkorelasi runtun AR(1) menjadi berbentuk

$$\mathbf{Y}_j = \mathbf{1}\beta_j + \varepsilon_j \tag{2.8}$$

dengan $\varepsilon_j$ diasumsikan berdistribusi normal dengan rata-rata nol dan matrik kovarian $\sigma^2\mathbf{W}_j$. Pada model (2.8) tersebut $\beta_j$ adalah rata-rata pengamatan dari populasi subyek ke-j.

$\mathbf{W}_j$ adalah matrik kovarian error dalam subyek ke-j ($\varepsilon_j$) yang berupa suatu matrik korelasi. Misalkan untuk sebuah subyek dengan 4 kali pengamatan dimana $\varepsilon_j$ berstruktur AR(1), maka $\mathbf{W}_j$ akan berbentuk:

$$\mathbf{W}_j = \begin{bmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{1,3} & \rho_{1,4} \\ \rho_{2,1} & \rho_{2,2} & \rho_{2,3} & \rho_{2,4} \\ \rho_{3,1} & \rho_{3,2} & \rho_{3,3} & \rho_{3,4} \\ \rho_{4,1} & \rho_{4,2} & \rho_{4,3} & \rho_{4,4} \end{bmatrix} \tag{2.9}$$

dimana elemen-elemen dalam matrik $\mathbf{W}_j$ pada persamaan (2.9) adalah $\rho_{a,b}$ yaitu korelasi antara pengamatan ke-a dan pengamatan ke-b. $\rho_{a,b} = \rho_h$ ,dimana $h = | a - b |$, sehingga matrik $\mathbf{W}_j$ pada (2.9) menjadi:

$$\mathbf{W}_j = \begin{bmatrix} \rho_0 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & \rho_0 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_0 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & \rho_0 \end{bmatrix} \tag{2.10}$$

Elemen-elemen dalam matrik $\mathbf{W_j}$ pada persamaan (2.10) diatas adalah $\rho_h$ yaitu korelasi antara pengamatan-pengamatan yang berselisih waktu antar pengamatan h atau korelasi antara pengamatan lag h. Dengan menggunakan rumus pada persamaan (2.4) maka matrik $\mathbf{W_j}$ pada persamaan (2.10) menjadi:

$$\mathbf{W_j} = \begin{bmatrix} 1 & \phi_j & \phi_j^2 & \phi_j^3 \\ \phi_j & 1 & \phi_j & \phi_j^2 \\ \phi_j^2 & \phi_j & 1 & \phi_j \\ \phi_j^3 & \phi_j^2 & \phi_j & 1 \end{bmatrix} \qquad (2.11)$$

dimana $\phi_j$ dalam matrik $\mathbf{W_j}$ pada persamaan (2.11) diatas adalah koefisien autoregresi dari proses AR(1), dan $\phi_j$ dalam interval $-1 < \phi_j < 1$.