BAB II

MATERI PENUNJANG

2.1. Metode Kuadrat Terkecil Biasa (OLS)

Metode kuadrat terkecil biasa dikemukakan oleh Carl Friedrich Gauss, seorang ahli matematika bangsa Jerman. Untuk menjelaskan pendekatan cara Gauss, didefinisikan fungsi regresi populasi dua variabel,

$$Yi = \beta_o + \beta_1 X_i + \varepsilon_i \tag{2.1.1}$$

dengan asumsi-asumsi sebagai berikut :

Asumsi 1 : $E(\varepsilon_i \mid X_i) = 0$ (2.1.2)

Asumsi 1 menyatakan bahwa nilai bersyarat dari ε_i pada Xi tertentu adalah nol.

Asumsi 2 : $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j))$

$= E(\varepsilon_i\, \varepsilon_j)$ karena asumsi 1

$= 0 \quad i \neq j$ (2.1.3)

dengan i dan j dua pengamatan yang berbeda

Asumsi 2 menyatakan bahwa gangguan ε_i dan ε_j tidak berkorelasi atau tidak ada korelasi berurutan, atau tidak ada autokorelasi.

Asumsi 3 : $\mathrm{Var}(\varepsilon_i \mid X_i) = E(\varepsilon_i - E(\varepsilon_i))^2$

$= E(\varepsilon_i^2)$ karena asumsi 1

$= \sigma^2$ (2.1.4)

Asumsi 3 menyatakan bahwa varians ε_i untuk tiap Xi adalah konstan sebesar σ^2. Secara teknis, menyatakan asumsi homoskedastisitas, atau penyebaran sama, atau varians sama.

Asumsi 4 : $\text{Cov}(\varepsilon_i, X_i) = E(\varepsilon_i - E(\varepsilon_i))(X_i - E(X_i))$

$$= 0 \tag{2.1.5}$$

Asumsi 4 menyatakan bahwa gangguan ε_i dan variabel yang menjelaskan X tidak berkorelasi.

Untuk asumsi 4 secara otomatis terpenuhi jika variabel X tak random atau tak stokastik dan asumsi 1 dipenuhi karena sebagian besar teori regresi persamaan didasarkan pada asumsi bahwa variabel X tak stokastik.

Suatu model regresi yang memenuhi keempat asumsi diatas dikenal dengan model regresi klasik standar, atau linear umum. Dari fungsi regresi populasi (2.1.1) diperoleh fungsi regresi sampel,

$$Y_i = \hat{\beta}_o + \hat{\beta}_1 X_i + e_i \tag{2.1.6}$$

$$= \hat{Y}_i + e_i \tag{2.1.7}$$

dengan $\hat{Y}_i$ adalah nilai taksiran (rata-rata bersyarat) dari Y_i. Dengan alternatif lain dinyatakan sebagai

$$e_i = Y_i - \hat{Y}_i$$

$$= Y_i - \hat{\beta}_o + \hat{\beta}_1 X_i \tag{2.1.8}$$

e_i (residual) merupakan perbedaan antara nilai Y_i sebenarnya dangan nilai Y_i yang ditaksir.

Misalkan diberikan n pasang observasi atas Y dan X tertentu, dan ditentukan fungsi regresi sampel yang sedekat mungkin dengan nilai sebenarnya. Adapun caranya dengan menjumlahkan residual (sisa) $\sum_{i=1}^{n} e_i = (Y_i - \hat{Y}_i)$. Untuk menghindarkan masalah digunakan metode kuadrat terkecil yang ditetapkan dengan meminimalkan $\sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$

$$= \sum_{i=1}^{n} (Y_i - \hat{\beta}_o - \hat{\beta}_1 X_i)^2 \tag{2.1.9}$$

Prinsip kuadrat terkecil memilih $\hat{\beta}_o$ dan $\hat{\beta}_1$ sedemikian rupa sehingga untuk suatu sampel tertentu $\sum_{i=1}^{n} e_i^{\,2}$ sekecil mungkin.

Adapun caranya dengan meminimalkan $\sum_{i=1}^{n} e_i$ yaitu menurunkan $\sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n}$ (Y$_i$ - $\hat{\beta}_o - \hat{\beta}_1 x_i)^2$ secara parsial terhadap $\hat{\beta}_o$ dan $\hat{\beta}_1$ untuk memperoleh

$$\frac{\partial\left(\sum_{i=1}^{n} e_i^{\,2}\right)}{\partial \hat{\beta}_o} = 2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_o - \hat{\beta}_1 X_i\right) \cdot (-1)$$

$$= -2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_o - \hat{\beta}_1 X_i\right) \tag{2.1.10}$$

$$\frac{\partial\left(\sum_{i=1}^{n} e_i^{\,2}\right)}{\partial \hat{\beta}_1} = 2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_o - \hat{\beta}_1 X_i\right) \cdot (-X_i)$$

$$= -2\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_o - \hat{\beta}_1 X_i\right) X_i \tag{2.1.11}$$

Dengan menyamakan kedua persamaan (2.1.10) dan (2.1.11) dengan nol, setelah penyederhaan secara aljabar dan manipulasi, menghasilkan penaksir yang diberikan untuk menaksir βo dan β_1 :

$$\sum_{i=1}^{n} Y_i = n\hat{\beta}_o + \hat{\beta}_1 \sum_{i=1}^{n} X_i \qquad (2.1.12)$$

$$\sum_{i=1}^{n} Y_i X_i = \hat{\beta}_o \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 \qquad (2.1.13)$$

dengan n adalah besarnya sampel. Persamaan (2.1.12) dan (2.1.13) dikenal sebagai persamaan normal.

Dengan memecahkan persamaan normal secara simultan, diperoleh :

$$|\Delta| = \begin{vmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{vmatrix} = n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} (X_i)\right)^2$$

$$\hat{\beta}_o = \frac{\begin{vmatrix} \sum_{i=1}^{n} Y_i & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} Y_i X_i & \sum_{i=1}^{n} X_i^2 \end{vmatrix}}{|\Delta|} = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \Sigma X_i}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} (X_i)\right)^2}$$

$$= \overline{Y} - \hat{\beta}_1 \overline{X} \qquad (2.1.14)$$

$$\hat{\beta}_1 = \frac{\begin{vmatrix} n & \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} Y_i X_i \end{vmatrix}}{|\Delta|} = \frac{n \sum_{i=1}^{n} Y_i X_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} (X_i) \right)^2}$$

$$= \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2} \qquad (2.1.15)$$

Penaksir yang diperoleh (2.1.14 dan 2.1.15) dikenal sebagai penaksir kuadrat terkecil.

Dengan mengingat asumsi mengenai ε_i, maka sebagai akibatnya $E(\hat{\beta}_1) = \beta_1 + \sum_{i=1}^{n} k_i E(\varepsilon_i) = \beta_1$, yang menunjukkan bahwa $\hat{\beta}_1$ adalah penaksir tak bias. Dan sesuai dengan definisi $Var(\hat{\beta}_1) = E(\hat{\beta}_1 - E(\hat{\beta}_1))^2$ maka

$$Var(\hat{\beta}_1) = E(\hat{\beta}_1 - \beta_1)^2$$

$$= E(\beta_1 + \sum_{i=1}^{n} k_i x_i - \beta_1)^2$$

$$= E(\sum_{i=1}^{n} k_i \varepsilon_i)^2$$

Karena sesuai dengan asumsi, $E(\varepsilon_i^2) = \sigma^2$ untuk tiap i, dan $E(\varepsilon_i \varepsilon_j) = 0$ untuk $i \neq j$, Sebagai akibatnya

$$Var(\hat{\beta}_1) = \sigma^2 \sum_{i=1}^{n} k_i^2$$

$$= \frac{\sigma^2}{\sum_{i=1}^{n} ki^2}$$ dengan menggunakan definisi k_i^2

Untuk varians $\hat{\beta}_0$ dapat diperoleh dengan mengikuti prosedur diatas.

Hasilnya, $$\mathrm{Var}(\hat{\beta}_0) = \frac{Xi^2}{n\sum_{i=1}^{n} xi^2}\sigma^2$$

dengan $x_i = X_i - \overline{X}$.

Adapun kesalahan standar taksiran OLS dapat diperoleh sebagai berikut :

$$\mathrm{Se}(\hat{\beta}_1) = \frac{\sigma}{\sqrt{\sum_{i=1}^{n} Xi^2}}$$

$$\mathrm{Se}(\hat{\beta}_0) = \sqrt{\frac{\sum_{i=1}^{n} Xi^2}{n\sum_{i=1}^{n} x_i^2}}\sigma$$

Adapun sifat-sifat penaksir kuadrat terkecil adalah :

Theorema Gauss Markov : dengan melihat asumsi model regresi linear klasik, penaksir kuadrat terkecil, dalam kelas penaksir linear tak bias, mempunyai varians minimum, yaitu penaksir tadi BLUE (Best Linear Unbiased Estimator).

2.2. Sifat-sifat Penaksir Titik

Sifat-sifat penaksir titik dibagi menjadi dua kategori yaitu sifat sampel kecil atau sampel terbatas dan sifat sampel besar atau asimptotik.

2.2.1. Sifat Sampel Kecil

1. Ketakbiasan (unbiasednes)

Suatu penaksir $\hat{\theta}$ dikatakan penaksir tak bias dari θ jika nilai yang diharapkan dari $\hat{\theta}$ sama dengan θ sebenarnya ; yaitu, $E(\hat{\theta}) = \theta$.

2. Varians minimum

Suatu penaksir $\hat{\theta}_1$ dikatakan sebagai penaksir dengan varians minimum dari θ jika varians dari $\hat{\theta}_1$ lebih kecil atau sama dengan varians dari $\hat{\theta}_2$, untuk setiap $\hat{\theta}_2$ yang merupakan penaksir lain dari θ.

3. Efisiensi

Suatu penaksir $\hat{\theta}_1$ dikatakan efisien jika $\hat{\theta}_1$ dan $\hat{\theta}_2$ adalah dua penaksir tak bias dari θ, dan $\hat{\theta}_1$ variansnya lebih kecil atau sama dengan varians $\hat{\theta}_2$, untuk semua $\hat{\theta}_2$ estimator tak bias dari θ.

4. Kelinieran (Linearity)

Suatu penaksir $\hat{\theta}$ dikatakan penaksir linier dari θ jika merupakan suatu fungsi linear dari observasi sampel.

5. Penaksir tak bias linier terbaik (Best Linier Unbias Estimator atau BLUE)

Suatu penaksir $\hat{\theta}$ dikatakan BLUE jika $\hat{\theta}$ adalah linier, tak bias dan mempunyai varians minimum.

2.2.2. Sifat Sampel Besar (asimptotik)

1. Ketakbiasan asimptotik (asymptotic unbiasedness)

Suatu penaksir $\hat{\theta}$ dikatakan takbias asimtotik jika $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$.

1. Konsistensi

Suatu penaksir $\hat{\theta}$ dikatakan konsisten jika $\hat{\theta}$ mendekati nilai θ yang sebenarnya dengan semakin besarnya ukuran sampel. Secara formal, suatu penaksir $\hat{\theta}$ dikatakan penaksir konsisten dari θ jika probabilitas dari nilai absolut untuk perbedaan $\hat{\theta}$ dan θ menjadi lebih kecil dari δ (suatu kuantitas positif yang kecil secara albitrality) mendekati satu atau

$$\lim_{n\to\infty} P\left\{\left|\hat{\theta}_n - \theta\right| < \delta\right\} = 1 \qquad \delta > 0$$

Ini seringkali dinyatakan sebagai :

$$P\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$$

Untuk mengetahui apakah suatu penaksir konsisten, kondisi cukup adalah bahwa $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$ dan $\lim_{n\to\infty} \mathrm{Var}(\hat{\theta}_n) = 0$.

Yaitu, $\hat{\theta}$ tak bias secara asimtotik dan dengan meningkatnya n secara tak terbatas varians mendekati nol. Dengan kata lain, distribusi dari $\hat{\theta}$ menjadi semakin terkonsentrasi pada nilai θ dengan meningkatnya n secara tak terbatas.

2. Kenormalan asimtotik

Suatu penaksir $\hat{\theta}$ dikatakan normal asimtotik jika distribusi samplingnya cenderung mendekati distribusi normal dengan meningkatnya ukuran sampel n secara tak terbatas.

2.3. Matriks dan Determinan untuk ordo 2 x 2

Suatu matriks adalah suatu barisan segiempat dari angka-angka atas unsur yang diatur dalam baris dan kolom. Matriks dengan ordo atau dimensi M dan N (ditulis sebagai M x N) adalah kumpulan dari M x N unsur yang disusun dalam M baris dan N kolom. Jadi matriks A (M x N) dinyatakan sebagai

$$A=\left[a_{ij}\right]=\begin{bmatrix} a_{11}\, a_{12}\, a_{13} \ldots\ldots a_{1N} \\ a_{21}\, a_{22}\, a_{23} \ldots\ldots a_{2N} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{M1}\, a_{M2}\, a_{M3} \ldots a_{MN} \end{bmatrix}$$

dengan a_{ij} adalah unsur yang muncul dalam baris ke i dan kolom ke j dari A

$[a_{ij}]$ adalah pernyataan ringkas untuk matriks A yang unsur khasnya adalah (aij)

Ordo, atau dimensi dari suatu matriks adalah banyaknya baris dan kolom, yang seringkali ditulis dibawah matriksnya untuk memudahkan .

Matriks singular adalah matriks bujursangkar yang determinannya sama dengan 0 dan matriks nonsingular jika determinannya tidak sama dengan 0.

Untuk suatu matriks bujursangkar A ada angka (skalar) yang bersesuaian dengan A yang dikenal determinan suatu matriks yang dinyatakan det Δ atau simbol $|\Delta|$ dimana $|\;|$ berarti determinan.

Contoh:

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{13} & a_{14} \end{bmatrix} \quad \text{maka} \quad |A| == a_{11}a_{14}-a_{12}a_{13}$$

Rumus Cramer :

Misal $a_{11}x + a_{21}y = b_1$

$a_{21}x + a_{22}y = b_2$

dengan determinan matrik koefisien tak sama dengan nol adalah

$$x = \frac{\det\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}}{\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}, \quad y = \frac{\det\begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix}}{\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}$$

Tingkat (rank) suatu matriks adalah ordo dari submatriks bujursangkar yang terbesar dan determinannya tidak sama dengan nol.

2.4. Korelasi r

r didefinisikan sebagai korelasi untuk mengukur hubungan linier antara X dan Y, dengan

$$r = \frac{[n\Sigma X_i Y_i - (\Sigma X_i)(\Sigma Y_i)]}{\sqrt{n\Sigma X_i{}^2 - (\Sigma X_i)^2}\sqrt{n\Sigma Y_i{}^2 - (\Sigma Y_i)^2}}$$

sifat – sifat r adalah:

Batasnya adalah $-1 \leq r \leq 1$

r = 1 berarti korelasi positif sempurna (kuat).

r = 0 berarti tidak berkorelasi.

r = -1 bearti korelasi negatif sempurna (kuat).

2.5. Koefisien Determinasi R^2

R^2 didefinisikan sebagai koefisien determinasi (sampel) untuk mengukur kebaikan suai (goodness of fit) garis regresi. Secara verbal, R^2 mengukur proporsi (bagian) atau prosentase sumbangan variabel yang menjelaskan terhadap variabel terikat..

$$R^2 = \frac{\Sigma(\hat{Y}_i - \overline{Y})^2}{\Sigma(Y_i - \overline{Y})^2}$$

$$= \frac{(\Sigma x_i y_i)^2}{(\Sigma x_i^{\,2})(\Sigma y_i^{\,2})}$$

$$= \frac{[n\Sigma X_i Y_i - (\Sigma X_i)(\Sigma Y_i)]^2}{[n\Sigma X_i^{\,2} - (\Sigma X_i)^2][n\Sigma Y_i^{\,2} - (\Sigma Y_i)^2]}$$

Sifat – sifat R^2 adalah:

1. R^2 merupakan besaran non negatif.
2. Batasnya adalah $0 \leq R^2 \leq 1$