BAB II

TEORI PENUNJANG

2.1. Estimator Tak Bias dan Generalized Jackknife.

Estimator yang mempunyai suku bias merupakan estimator yang kurang baik. Metode Generalized Jackknife yang pertama diperkenalkan oleh Quenuoille merupakan metode untuk mereduksi bias suatu estimator dengan penggunaan ulang sampel sehingga diperoleh estimator baru. Pada subbab ini akan diberikan beberapa pengertian tentang estimator tak bias, metode generalized Jackknife dan cara mèndapatkan estimator – estimator dengan metode Quenuoille serta proses stochastik.

2.1.1. Estimator Tak Bias.

Dalam suatu penaksiran tidak mungkin selalu diperoleh estimator yang tepat sehingga timbul beberapa kriteria tentang baiknya suatu estimator. Salah satu sifat baik estimator adalah tak bias. Disini akan dibahas kriteria tak bias dari estimator untuk menunjang pembahasan selanjutnya.

Definisi 2.1.1.1

Suatu Estimator titik $\hat{\theta}$ disebut estimator tak bias untuk θ jika $E[\hat{\theta}] = \theta$. Kriteria ini menyatakan bahwa rata-rata semua harga $\hat{\theta}$ yang mungkin akan sama dengan

θ. Sedangkan estimator yang tidak tak bias disebut estimator bias. Jika $\hat{\theta}$ adalah estimator bias untuk θ dan bias dari estimator ditulis $B(\hat{\theta})$, maka

$$E[\hat{\theta}] = \theta + B(\hat{\theta}) \qquad .(2.1.1.1)$$

Sehingga bias estimator :

$$B(\hat{\theta}) = E[\hat{\theta}] - \theta$$

$$= E[\hat{\theta} - \theta] \qquad .(2.1.1.2)$$

Dari definisi 2.1.1.1 akan diberikan contoh.

Contoh 2.1.1.1

Misalkan X berdistribusi Uniform dalam interval $(0,\theta)$ dengan fungsi densitas :

$$f(x,\theta) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & x \text{ lainnya} \end{cases} \qquad .(2.1.1.3)$$

akan dibuktikan bahwa

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} Xi \qquad .(2.1.1.4)$$

Merupakan estimator bias untuk θ

Nilai harapan dari X adalah

$$E[X] = \int_0^\theta x \; f(x/\theta)\,dx \qquad .(2.1.1.5)$$

$$= \int_0^\theta x \; \frac{1}{\theta} \; dx$$

$$= \frac{1}{\theta}\left[\frac{x^2}{2}\right]_0^\theta = \frac{\theta}{2} \qquad .(2.1.1.6)$$

sehingga nilai harapan dari $\overline{X}$ adalah

$$E[\overline{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} Xi\right] \qquad (2.1.1.7)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E[Xi]$$

$$= \frac{1}{n} n \frac{\theta}{2}$$

$$= \frac{\theta}{2} = \theta - \frac{\theta}{2} \qquad (2.1.1.8)$$

Dari persamaan (2.1.1.8) diatas dapat dilihat bahwa nilai yang diharapkan dari X tidak sama dengan parameter, sehingga $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} Xi$ merupakan estimator yang bias untuk θ dengan biasnya $-\frac{\theta}{2}$.

2.1.2. Metode Generalized Jackknife.

Salah satu metode untuk mengurangi bias estimator adalah dengan metode Generalized Jackknife. Sedangkan Jackknife merupakan kasus khusus dari Generalized Jackknife. Disini akan diberikan beberapa definisi dan teorema dari Generalized Jackknife.

Definisi 2.1.2.1

Jika diberikan $\hat{\theta}_1$ dan $\hat{\theta}_2$ adalah estimator-estimator untuk θ untuk sembarang bilangan riil $R \neq 1$ didefinisikan Generalized Jackknife

$$G(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\theta}_1 - R\hat{\theta}_2}{1 - R} \qquad .(2.1.2.1)$$

Untuk menunjukan bahwa $G(\hat{\theta}_1, \hat{\theta}_2)$ merupakan estimator tak bias akan diberikan dengan theorema berikut.

Theorema 2.1.2.1

Jika $E\left[\hat{\theta}_k\right] = \theta + b_k(n, \theta), \quad k = 1,2 \qquad .(2.1.2.2)$

Dengan $b_k(n, \theta)$ adalah bias dari suatu estimator untuk θ atas sampel random yang besarnya n pengamatan dengan $b_k(n, \theta) \neq 0$ dan

$$R = \frac{b_1(n, \theta)}{b_2(n, \theta)} \neq 1 \qquad .(2.1.2.3)$$

maka

$$E\left[G(\hat{\theta}_1, \hat{\theta}_2)\right] = \theta \qquad .(2.1.2.4)$$

Bukti

Dari persamaan (2.1.2.1) didefinisikan

$$G(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\theta}_1 - R\,\hat{\theta}_2}{1 - R} \qquad .(2.1.2.5)$$

Kemudian nilai harapan dari $G(\hat{\theta}_1, \hat{\theta}_2)$ adalah :

$$E\left[G(\hat{\theta}_1, \hat{\theta}_2)\right] = \frac{E\left[\hat{\theta}_1\right] - R\,E\left[\hat{\theta}_2\right]}{1 - R} \qquad .(2.1.2.6)$$

$$= \frac{\theta + b_1(n,\theta) - R[\theta + b_2(n,\theta)]}{1-R}$$

$$= \frac{(1-R)\theta + b_1(n,\theta) - Rb_2(n,\theta)}{1-R}$$

$$= \frac{(1-R)\theta + b_1(n,\theta) - \frac{b_1(n,\theta)}{b_2(n,\theta)} b_2(n,\theta)}{1-R}$$

$$= \theta \qquad (2.1.2.7)$$

Maksud theorema 2.1.2.1 adalah jika R diketahui dan diberikan dengan persamaan (2.1.2.3) maka $G(\hat{\theta}_1, \hat{\theta}_2)$ adalah estimator tak bias untuk θ.

Contoh 2.1.2.1

Dengan definisi 2.1.2.1 dimana $G(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\theta}_1 - R\hat{\theta}_2}{1-R}$ dapat diperoleh besarnya bilangan riil R adalah sebagai berikut:

$$R = \frac{\hat{\theta}_1 - G(\hat{\theta}_1, \hat{\theta}_2)}{\hat{\theta}_2 - G(\hat{\theta}_1, \hat{\theta}_2)}, R \neq 1 \qquad (2.1.2.8)$$

Kemudian dengan menggunakan contoh 2.1.1.1 dan diberikan $\hat{\theta}_1 = \frac{\overline{X}}{2}$ dan $\hat{\theta}_2 = \frac{\overline{X}}{3}$

Serta diasumsikan bahwa $G(\hat{\theta}_1, \hat{\theta}_2)$ adalah estimator tak bias. Maka besarnya bilangan riil R adalah :

Dengan menggunakan persamaan (2.1.2.8) yaitu:

$$R = \frac{\hat{\theta}_1 - G(\hat{\theta}_1, \hat{\theta}_2)}{\hat{\theta}_2 - G(\hat{\theta}_1, \hat{\theta}_2)}, R \neq 1$$

$$R = \frac{\overline{X}/2 - G(\overline{X}/2, \overline{X}/3)}{\overline{X}/3 - G(\overline{X}/2, \overline{X}/3)}, R \neq 1$$

Karena $G(\hat{\theta}_1, \hat{\theta}_2)$ adalah estimator tak bias maka nilai harapannya adalah :

$$E\left[G\left(\hat{\theta}_1, \hat{\theta}_2\right)\right] = \theta$$

Sehingga

$$R = \frac{E\left[\overline{X}\right]/2 - E\left[G\left(\overline{X}/2, \overline{X}/3\right)\right]}{E\left[\overline{X}\right]/3 - E\left[G\left(\overline{X}/2, \overline{X}/3\right)\right]}$$

$$R = \frac{\theta/4 - \theta}{\theta/6 - \theta} = \frac{9}{10} \qquad (2.1.2.9)$$

Dari persamaan (2.1.2.9) diperoleh bahwa besarnya R = 9/10.

Dengan menganggap $G\left(\hat{\theta}_1, \hat{\theta}_2\right)$ sebagai suatu estimator untuk θ jika R diketahui, pertanyaan yang timbul adalah bagaimana sebaiknya $\hat{\theta}_1$ dan $\hat{\theta}_2$ dipilih.

Misal $\hat{\theta}_1$ adalah suatu estimator tertentu dari θ atas sampel random $x_1, x_2, x_3, \ldots x_n$ dengan $E\left[\hat{\theta}_1\right] = \theta + b_1(n,\theta) = \theta + b(\theta)f(n)$ . . .(2.1.2.10)

Selanjutnya dimisalkan $\hat{\theta}_2$ adalah estimator lain untuk θ didefinisikan atas subsampel dari $x_1, x_2, x_3, \ldots x_n$ yang diperoleh dengan membuang satu x_i sehingga

$$E\left[\hat{\theta}_2\right] = \theta + b_2(n,\theta) = \theta + b(\theta)f(n-1) \qquad (2.1.2.11)$$

dan $$R = \frac{b_1(n,\theta)}{b_2(n,\theta)} = \frac{b(\theta)f(n)}{b(\theta)f(n-1)} = \frac{f(n)}{f(n-1)} \qquad (2.1.2.12)$$

sehingga jika $f(n)$ diketahui maka R diketahui. Karena pada umumnya R adalah fungsi dari n, maka untuk selanjutnya digunakan notasi R(n) untuk R.

Untuk memilih $\hat{\theta}_1$ dan $\hat{\theta}_2$ dikerjakan dengan metode Quenouille yang akan diberikan pada subbab berikut.

2.1.3. Metode Quenuoille

Akan dibahas prosedur umum untuk mencari $\hat{\theta}_2$ jika $\hat{\theta}_1$ diberikan. Misal $\hat{\theta}_1$ adalah suatu penaksir θ yang didasarkan atas sampel random $x_1, x_2, x_3, \ldots x_n$ kemudian sampel ini dipartisi menjadi N subbagian yang berukuran M sehingga NM = n, dan dibentuk sampel random yang baru dengan menghilangkan sembarang subbagian berukuran M dari sampel semula. Kemudian ditetapkan penaksir $\hat{\theta}^i$ sebagai penaksir θ yang ditetapkan atas subsampel yang didapat setelah subbagian ke-i yang berukuran M dibuang, selanjutnya didefinisikan :

$$J_i(\hat{\theta}) = N\hat{\theta} - (N-1)\hat{\theta}^i,\ i = 1,2,3,\ldots,N \qquad (2.1.3.1)$$

dan

$$J(\hat{\theta}) = \frac{1}{N}\sum_{i=1}^{N} J_i(\theta)$$

$$= N\hat{\theta} - (N-1)\bar{\hat{\theta}}^i \qquad (2.1.3.2)$$

Penaksir $J(\hat{\theta})$ disebut Jackknife dan penaksir $J_i(\hat{\theta})$ disebut nilai samaran (*pseudovalue*) dari Jackknife.

Contoh 2.1.3.1

Misal $X_1, X_2, \ldots, X_n$ adalah variabel random sebanyak n yang berdistribusi normal $N(\mu, \sigma^2)$. Misal $\hat{\sigma}^2$ adalah estimator dari σ^2 yang diperoleh dengan Metode Maksimum Likelihood, maka:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^{\,2} - 2X_i\overline{X} + \overline{X}^2\right)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^{\,2} - 2\sum_{i=1}^{n}X_i\overline{X} + \sum_{i=1}^{n}\overline{X}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i^{\,2} - 2\overline{X}^2 + \overline{X}^2$$

$$= \sum_{i=1}^{n}\frac{X_i^{\,2}}{n} - \overline{X}^2, \text{dimana } \overline{X} = \sum_{i=1}^{n}\frac{X_i}{n}$$

Nilai harapan dari $\hat{\sigma}^2$ adalah:

$$E\left[\hat{\sigma}^2\right] = \sum_{i=1}^{n}E\left[\frac{X_i^{\,2}}{n}\right] - E\left[\overline{X}^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 - \mu^2\right) - \left(\frac{1}{n}\sigma^2 - \mu^2\right)$$

$$= \sigma^2 - \frac{1}{n}\sigma^2$$

maka $\hat{\sigma}^2$ adalah estimator yang bias dari σ^2 dengan bias :

$$B\left(n, \sigma^2\right) = -\frac{1}{n}\sigma^2$$

Jika variabel random itu dipartisi dalam n partisi (M = 1), maka:

$$\hat{\sigma}_i^{\,2} = \frac{1}{n-1}\sum_{\substack{j=1\\ j\neq i}}^{n}X_j^{\,2} - \left(\frac{1}{n-1}\sum_{\substack{j=1\\ j\neq i}}^{n}X_j\right)^2$$

$$= \frac{1}{n-1}\sum_{\substack{j=1\\ j\neq i}}^{n}X_j^{\,2} - \left(\frac{n\overline{X} - X_i}{n-1}\right)^2$$

$$\hat{\sigma}_i{}^2 = \frac{1}{n-1}\left(\sum_{j=1}^{n} X_j{}^2 - X_i{}^2\right) - \frac{n^2\overline{X}^2 - 2n\overline{X}X_i + X_i{}^2}{(n-1)^2}$$

Sehingga

$$\bar{\hat{\sigma}}_i{}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_i{}^2$$

$$= \frac{1}{n(n-1)}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_j{}^2 - \sum_{i=1}^{n} X_i{}^2\right) - \frac{n^2\overline{X}^2 - 2n\overline{X}^2 + \sum_{i=1}\frac{X_i{}^2}{n}}{(n-1)^2}$$

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

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

$$= \frac{n(n-2)}{(n-1)^2}\left(\frac{1}{n}\sum_{i=1}^{n} X_i{}^2 - \overline{X}^2\right)$$

$$= \frac{n(n-2)}{(n-1)^2}\hat{\sigma}^2$$

maka

$$E\left[\bar{\hat{\sigma}}_i{}^2\right] = \frac{n(n-2)}{(n-1)^2}E\left[\hat{\sigma}^2\right]$$

$$= \frac{n(n-2)}{(n-1)^2}\frac{n-1}{n}\sigma^2$$

$$= \sigma^2 - \frac{1}{n-1}\sigma^2$$

sehingga

$$J(\hat{\sigma}^2) = n\hat{\sigma}^2 - (n-1)\bar{\hat{\sigma}}_i^{\,2}$$

$$= n\hat{\sigma}^2 - (n-1)\frac{n(n-2)}{(n-1)^2}\hat{\sigma}^2$$

$$= n\hat{\sigma}^2 - \frac{n(n-2)}{(n-1)}\hat{\sigma}^2$$

$$= n\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) - \frac{n(n-2)}{(n-1)}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right)$$

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

$$= \left(1 - \left(\frac{n-2}{n-1}\right)\right)\sum_{i=1}^{n}(X_i - \overline{X})^2$$

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

Kemudian nilai harapan dari $J(\hat{\sigma}^2)$ adalah:

$$E[J(\hat{\sigma}^2)] = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right]$$

$$= E\left[\frac{1}{n-1}n\hat{\sigma}^2\right]$$

$$= \frac{n}{n-1}E[\hat{\sigma}^2]$$

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

$$= \sigma^2$$

Jadi $J(\hat{\sigma}^2) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ adalah estimator tak bias untuk σ^2

$J(\hat{\theta})$ pada persamaan (2.1.3.2) adalah kejadian khusus dari $G(\hat{\theta}_1, \hat{\theta}_2)$

Dengan :

$$R(N) = \frac{N-1}{N}$$

$$\hat{\theta}_1 = \hat{\theta}$$

dan $\hat{\theta}_2 = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}^i = \bar{\hat{\theta}}^i$

Akan ditunjukan bahwa jackknife adalah kejadian khusus dari $G(\hat{\theta}_1, \hat{\theta}_2)$ yaitu sebagai berikut :

$$J(\hat{\theta}) = N\hat{\theta} - (N-1)\bar{\hat{\theta}}^i$$

$$= N\left[\hat{\theta} - \frac{(N-1)}{N}\bar{\hat{\theta}}^i\right]$$

$$= \frac{\hat{\theta} - \frac{(N-1)}{N}\bar{\hat{\theta}}^i}{\frac{1}{N}}$$

$$= \frac{\hat{\theta} - \frac{(N-1)}{N}\bar{\hat{\theta}}^i}{1 - \frac{N-1}{N}}$$

$$= \frac{\hat{\theta}_1 - R(N)\hat{\theta}_2}{1 - R(N)}$$

Definisi 2.1.3.1

Misalkan $\hat{\theta}_1$ dan $\hat{\theta}_2$ adalah penaksir penaksir dari θ yang didasarkan atas sampel berukuran n sedemikian sehingga

$b_1(n,\theta) = E[\hat{\theta}_1 - \theta] \neq 0$ dan $b_2(n,\theta) = E[\hat{\theta}_2 - \theta] \neq 0$

jika $\left| \lim_{n\to\infty} \frac{b_1(n,\theta)}{b_2(n,\theta)} \right| = 1$ . . . .(2.1.3.3)

maka dikatakan bahwa $\hat{\theta}_1$ dan $\hat{\theta}_2$ penaksir – penaksir θ yang berorder bias sama (*Same Order Bias Estimators*) dan ditulis : $\hat{\theta}_1$ S.O.B.E $\hat{\theta}_2$

jika $0 < \left| \lim_{n\leftarrow\infty} \frac{b_1(n,\theta)}{b_2(n,\theta)} \right| < 1$ . . .(2.1.3.4)

maka dikatakan bahwa $\hat{\theta}_1$ adalah penaksir θ yang berorder bias sama dengan $\hat{\theta}_2$, tetapi $\hat{\theta}_1$ lebih baik dari $\hat{\theta}_2$ *(Better Same Order Bias Estimator)*, ditulis :

$\hat{\theta}_1$ B.S.O.B.E $\hat{\theta}_2$

Definisi 2.1.3.2

Jika $\hat{\theta}_1$ dan $\hat{\theta}_2$ adalah seperti dalam definisi 2.1.1 tanpa pembatas $b_1 = (n,\theta) \neq 0$ dan

$\left| \lim_{n\to\infty} \frac{b_1(n,\theta)}{b_2(n,\theta)} \right| = 0$ . . .(2.1.3.5)

maka dikatakan bahwa $\hat{\theta}_1$ adalah penaksir θ yang berorder bias lebih rendah daripada $\hat{\theta}_2$ *(Lower Order Bias Estimator)*, ditulis : $\hat{\theta}_1$ L.O.B.E $\hat{\theta}_2$.

2.1.4. Proses Stochastik

Proses Stochastik adalah himpunan variabel random dengan fungsi waktu. Proses stochastik sering disebut proses random. Himpunan nilai-nilai yang mungkin untuk suatu variabel random X_n dari suatu proses stochastik $\{X_n , n \geq 1 \}$ disebut ruang state. Ruang state tidak hanya untuk himpunan variabel random yang diskret, juga untuk himpunan variabel random yang kontinu. Setiap harga t dalam $\{X(t), t \in T\}$ akan menghasilkan satu variabel random X(t) dan mempunyai ruang state sendiri.

Dalam hubungan dengan ruang state ada empat macam proses stochastik :

1. Waktu diskrit, ruang state diskrit
2. Waktu diskrit, ruang state kontinu
3. Waktu kontinu, ruang state diskrit
4. Waktu kontinu, ruang state kontinu

Keempat macam proses stochastik ini dapat dinyatakan dengan $\{X(t), t \in T\}$, dalam hal waktu diskrit biasanya digunakan parameter n sehingga di tulis dengan $\{X_n , n = 0,1,2,\ldots\}$

Proses Stochastik dengan incremen independen

Proses Stochastik dikatakan incremen independen jika untuk semua $t_1,t_2\ldots,t_n$ dimana $t_1<t_2<\ldots<t_n$,variabel random $X(t_2)-X(t_1)$, $X(t_3)-X(t_2),\ldots,X(t_n)-X(t_{n-1})$ independen. Sedangkan Proses Stochastik dikatakan incremen independen dan stasionari jika ada tambahan bahwa $X(t_2+h) - X(t_1+h)$ mempunyai distribusi sama dengan $X(t_2)-X(t_1)$ untuk semua t_1, t_2 dan untuk setiap $h > 0$.