ISBN:1858-1382

J M S K

Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 2, No.2, 90-96, Januari 2006

Penaksir Tak Bias Variansi Minimum Seragam

Georgina M. Tinungki[†]

Abstrak

Apabila terdapat lebih dari satu penaksir tak bias, maka ada tak berhingga banyaknya penaksir tak bias, akan tetapi mungkin pula tidak ada satu pun penaksir tak bias. Sehingga dapat muncul dua pertanyaan, yaitu bagaimana cara memilih salah satu penaksir tak bias untuk dipakai, jika terdapat beberapa atau banyak penaksir tak bias, dan penaksir bias bagaimana yang harus digunakan jika tidak terdapat penaksir tak bias. Beberapa sifat dari penaksir kemungkinan maksimum dapat menjawab pertanyaan yang muncul, dengan melihat sifat ketak biasan, keselarasan dan bentuknya yang lain. Penaksir tak bias optimum dengan penaksir tak bias varian minimum seragam mempunyai varian yang minimum di antara semua kemungkinan penaksir yang tak bias. Penaksir tak bias varian minimum seragam yang ada biasanya mencapai batas bawah variansi Rao-Cramer, meskipun penaksir tak bias bervariansi minimum ini bisa saja tidak dapat ditentukan. Suatu metode mencari penaksir tak bias varian minimum seragam adalah dengan mencari suatu fungsi tak bias dari suatu statistik cukup yang lengkap. Bentuk lainnya adalah dengan mencari fungsi tak bias jika diketahui adanya suatu statistik cukup yang lengkap.

Keywords: *Penaksirtakbias, variansi minimum seragam, statistikcukup*.

1. Pendahuluan

Gagasan dasar dari penaksir statistik, khususnya penaksir titik yaitu dengan mengamati peubah acak $X_1, X_2,...., X_n$, dalam hal ini tidak perlu bebas maupun berdistribusi identik, dimana fungsi distribusi $X_1, X_2,...., X_n$ bergantung pada parameter θ yang tidak diketahui, seperti misalnya rataan suatu distribusi normal yang mungkin tidak diketahui, namun diketahui termasuk dalam suatu himpunan tertentu Θ. Masalah dalam menaksir Θ merupakan permasalah penaksir titik. Dalam menaksir Θ digunakan suatu fungsi yang sesuai dengan pengamatan $X_1, X_2,...., X_n$ yang disebut sebagai 'statistik' atau suatu 'penaksir'.

Jika terdapat lebih dari satu penaksir tak bias, maka ada tak berhingga banyaknya penaksir tak bias, akan tetapi mungkin pula tidak ada satu pun penaksir tak bias. Dalam hal ini akan muncul dua pertanyaan. Pertama, jika terdapat banyak penaksir tak bias, bagaimana cara memilih salah satu untuk dipakai. Kedua, jika tidak terdapat penaksir tak bias, penaksir bias macam apa yang sebaiknya dipergunakan? Beberapa sifat dari penaksir kemungkinan maksimum yang dapat menjawab pertanyaan-pertanyaan yang muncul, dengan melihat sifat ketak biasan, keselarasan dan bentuknya yang lain.

Dalam banyak permasalahan terapan, yang diinginkan bukanlah menaksir parameter Θ itu sendiri, tetapi suatu fungsi dari parameter tersebut. Sebagai ilustrasi, jika ingin membandingkan distribusi pendapatan dua kota, jika Δ adalah prosentase dalam distribusi pendapatan yang tumpang tindih, maka untuk menaksir Δ tidak saja harus menaksir μ, υ, dan σ,

[†]*Staf Pengajar pada Jurusan Matematika FMIPA Universitas Hasanuddin Makassar*

tetapi juga suatu fungsi dari ketiga parameter tersebut. Penggunaan prinsip ketakragaman dari penaksir kemungkinana maksimum akan memecahkan masalah ini.

Penaksir tak bias optimum dengan penaksir tak bias varian minimum seragam mempunyai varian yang minimum diantara semua kemungkinan penaksir yang tak bias. Walaupun penaksir tak bias varian minimum seragam tidak selalu ada, tetapi sering pula ada dan mencapai batas bawah variansi Cramer-Rao dari suatu penaksir tak bias. Suatu metode mencari penaksir tak bias varian minimum seragam, dengan mencari suatu fungsi tak bias dari suatu statistik cukup yang lengkap. Bentuk lainnya adalah mencari fungsi tak bias jika diketahui statistik cukup lengkap.

2. Landasan Teori

2.1. Penaksir Tak Bias

Jika terdapatpeubahacak $X_1, X_2,...., X_n$ (tidakperlubebasmaupun berdistribusi identik), fungsidistribusidari $X_1, X_2,...., X_n$ bergantungpada parameter θ yang tidakdiketahui (misalnya rataan suatu distribusi normal mungkin tidak diketahui), tetapi diketahui termasuk dalam suatuhimpunantertentu θ. Persoalanmenaksir θ merupakanpermasalahanpenaksir titik.Jika dipunyai ruang sampel dasar Ω, percobaan pada dasarnya memilih suatu w € Ω, yang pengamatannya $X_1(w) = x_1,..., X_n(w) = x_n$, danfungsidistribusidari $X_1, X_2,...., X_n$, misalnya $F_{x_1},...,_{x_n}(y_1,...., y_n|\theta)$. Jikainginmenaksir θ, tentunyadalampenaksiran θ menggunakansuatufungsi yang sesuaidaripengamatan $X_1, X_2,...., X_n$. Fungsi seperti ini disebut suatu penaksir.

***Definisi* 1**. Setiap fungsi peubah acak yang diamati $t_n(X_1,.., X_n)$, disebut penaksir dari θ dimana $X_1, X_2,...., X_n$ p.a maka fungsi itu juga suatu p.a. Suatu nilai tertentu dari penaksir $t_n(X_1,.., X_n)$ disebut taksiran dari θ. Jika diinginkan penaksir θ yang pada rata-ratanya tepat sama dengan θ, dan jika demikian penaksir θ disebut **"**tak bias**".**

***Definisi* 2**. Dikatakan $t_n(X_1,...., X_n)$ suatu penaksir tak biasdari θ jika $E_\theta t_n(X_1,....., X_n) \equiv \theta$, untuk semua $\theta \in \Theta$. Subscript θ pada tanda harapan menyatakan fungsi distribusi $F_{x_1,......., x_n}(y_1,..., y_n|\theta)$ yang akan digunakan dalam menghitung nilainya.

Sifat lain yang dimiliki suatu penaksir θ ialah kekonvergenan dalam peluang ke θ (yakni, peluangnya 'mengumpul' dekat ke θ), jika banyaknya pengamatan $n \to \infty$. Jika $X_1, X_2,...., X_n$ adalah sampel random dari populasi berdistribusi Normal dengan mean μdan variansi σ^2*, maka penduga* titik untuk μadalah

$$\hat{\mu} = \text{g(X)} = \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Untuk n cukup besar, maka $\overline{X}_n$ adalah dekat dgn μdalam arti bahwa :

$$P[\left|\overline{X}_n - \mu\right| \le \delta] = \left|\int_{-\sqrt{n}\delta/\sigma}^{\sqrt{n}\delta/\sigma} \frac{1}{\sqrt{2\pi}} e^{-t^{2/2}} dt \longrightarrow 1\right.$$

untuk $n \to \infty$, sehingga $E(\overline{X}) = \mu$, berapapun harga μ dan σ^2.

2.2. Metode Kemungkinan Maksimum (*Maksimum Likelihood Estimator*)

Metode untuk menghitung estimasi titik yang paling populer adalah metode Kemungkinan Maksimum (*a maximum likelihood estimator*) yang memilih P $\in$ P , yang merupakan pilihan "terbaik" untuk data observasi yang sebenarnya. Anggapan bahwa variabel random $X_1, X_2, ..., X_n$ mempunyai distribusi bersama P dengan parameter θ.

Metode estimasi kemungkinan maksimum berusaha menentukan Modus(harga yang memaksimumkan fkp dari distribusi). Namun kanyataannyamodus merupakan estimasi yang lebih jelek dari mean atau median, yang menjelaskan mengapa sifat-sifat sampel kecil metode ini umumnya jelek. Tetapi untuk sampel yang besar, modus cenderung mendekati mean dan median, sehingga metode ini mempunyai sifat-sifat sampel besar yang "baik".

Formulasi Prinsip Metode Maksimum Likelihood

Pandang sampel random $X_1, X_2, \ldots, Xn$ dari distribusi yang mempunyai pdf $f(x;\theta)$, $\theta \in \Omega$ .Fkp (pdf) bersama dari $X_1, X_2, \ldots, Xn$ adalah $f(x_1;\Theta), f(x_2;\Theta) \ldots f(x_n;\Theta)$.

Pdf bersama dapat dipandang sebagai suatufungsidari θ, yang dikatakansebagai**fungsi likelihood***L* dari sampel random, dan ditulis sebagai :

$$L(\theta; x_1, x_2, ..., x_n) = f(x_1;\theta)\, f(x_2;\theta) ......f(x_n;\theta), \quad \theta \in \Omega$$

Dari pendugaan tersebut diperoleh fungsi yang tidak berarti (*nontrivial function*), katakan $u(x_1, x_2, \ldots, x_n)$, sedemikianhinggabila θ digantioleh$u(x_1, x_2, \ldots, x_n)$, fungsi likelihood L adalah *maksimum*. Atau, dapat dikatakan

$$L[u(x_1, x_2, \ldots, x_n); x_1, x_2, \ldots, x_n)] < L(\Theta; x_1, x_2, .., x_n) \ \forall \quad \theta \in \Omega$$

Maka statistik $u(X_1, X_2, \ldots, X_n)$, disebut sebagai ***Penduga Maksimum Likelihood*** dari θ, dan dan dinyatakan dengan simbol $\hat{\theta} = u(X_1, X_2, ..., X_n)$

Sebagai contoh, jika $X_1, X_2, \ldots, X_n$ menyatakan sampel random dari distribusi dengan fkp

$$f(x) = \Theta^x (1-\Theta)^{1-x,} \quad , \quad x = 0,1$$
$$= 0 \quad , \quad \text{lainnya, dimana } 0 \le \Theta \le 1.$$

Maka fkp bersama dari distribusi di atas

$$\Theta^{x1}(1-\Theta)^{1-x1}\Theta^{x2}(1-\Theta)^{1-x2} \ldots \Theta^{xn}(1-\Theta)^{1-xn} = \Theta^{\sum xi}\Theta^{xi}(1-\Theta)^{n-\sum xi}$$

dimana x_i sama-sama 0 atau 1, $i = 1,2,\ldots,n$. Sehingga fungsi $L(\Theta)$,yang disebut sebagai *fungsi likelihood*, dituliskan menjadi

$$L(\Theta) = \Theta^{\sum xi}\Theta^{x\,\mathrm{i}}(1-\Theta)^{n-\sum x\mathrm{i}}, \quad 0 \le \Theta \le 1$$

Peluang maksimum $L(\Theta)$dari sampel x_1, x_2, …,$x_n$adalah θ. Jelas, nilaimaksimumdari θ merupakanestimasi yang baikdari θ sebabpeluangnyalebihbesar dari pada sampel yang utama. Bagaimanapun karena fungsi likelihood $L(\Theta)$ beserta logaritmanya $ln\ L(\Theta)$adalah maksimumuntuknilai θ yang sama, maka salah satu dari $L(\Theta)$ atau $ln\ L(\Theta)$ dapat digunakan. Dalam hal ini,

$$\ln L(\theta) = \left(\sum_1^n x_i\right)\ln\theta + \left(n - \sum_1^n x_i\right)\ln(1-\theta)$$

Jadi

$$\frac{d\ln L(\theta)}{d\theta} = \frac{\sum x_i}{\theta} - \frac{n-\sum x_i}{1-\theta} = 0$$

Terlihatbahwa θ tidaksamadengan 0 atau 1. Ini ekivalen untuk persamaan

$$(1-\theta)\sum_1^n x_i = \theta\left(n - \sum_1^n x_i\right)$$

Sehingga diperolehpersamaanuntuk θ adalah $\sum_1^n x_i / n$ dimana maksimum $\sum_1^n x_i / n$ itu ada, sehingga$L(\theta)$dan$ln\ L(\theta)$dapatdiperiksa. , dalam kasus $\sum_1^n x_i / n$ semua x_1, x_2, …,$x_n$bersama-sama adalah 0 dan 1. $\sum_1^n x_i / n$ adalahnilaimaksimum$L(\theta)$dari θ, atau secara statistik dikatakan $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n X_i$ adalah Estimasi Maximum Likelihood dari θ.

2.3. Kecukupan Statistik

Apabila diamati suatu peubah acak X yang distribusinya bergantung pada suatu parameter θ yang tidak diketahui (misalnya $X = (X_1, \ldots, X_n)$, dengan $X_1, \ldots, X_n$peubah acak bebas dan berdistribusi identik, masing-masing mempunyai fungsi kepadatan $f(x/\theta)$, sedangkan θ tidak diketahui, $\theta \in \Theta$), maka X dapat disederhanakan melalui suatu fungsi tanpa kehilangan suatu informasi tentang θ. Sebagai ilustrasi, jika $t(X) = t(X_1, \ldots, Xn) = (X_1, \ldots, X_n)/n)$,rataan sampel, mungkin dari berbagai hal mengandung semua informasi tentang θ yang relevan, dalah hal tersebut $t(X)$ disebut suatu **statistik cukup.**

***Definisi* 3**. Statistik $t(X_1, \ldots, Xn)$adalah statistik cukup untuk θ jika $P_\theta[(X_1,\ldots, X_n) \in \vartheta | t(X_1,\ldots, X_n) = c] = l(\vartheta|c)$, $\forall \vartheta \subseteq \Re^n$. Jika bagian kiri persamaan diatas suatu fungsi dari ϑ dan c, misalnya $l(\vartheta|c)$, tapi bukan fungsi dari θ (disiniθ dan c mungkin vektor).

2.4. Kelengkapan Statistik

***Definisi* 4**. Apabila suatu statistik cukup T pada suatu masalah memiliki keluarga fungsi kepadatan peluang $g(x/\theta)$ yang lengkap, maka T disebut **statistik cukup lengkap.**

Sebagaiilustrasi, misalkan $X_1,...., X_n$ peubahacakbebas, masing-masing seragam pada [0, θ) dengan $\theta > 0$. Y_n = maks ($X_1,...., X_n$) suatustatistikcukup untuk θ dan mempunyai fkp

$$f_{Y_n}(x) = \frac{nx^{n-1}}{\theta^n} I_{[0,\theta]}(x)$$

Dalam menyelidiki kelengkapan keluarga ini, misalkan $u(.)$ sembarangfungsi, maka $E_\theta u(Y_n) = 0$, yang berarti

$$\int_0^\theta \frac{u(x)nx^{n-1}}{\theta^n} dx = 0$$

Jika kedua ruas dikalikan dengan θ/n yang > 0, maka berlaku

$$\int_0^\theta u(x)x^{n-1} = 0$$

Sehingga $u(\theta)$ = 0, $\forall \theta$, ataudengan kata lain $u(x)$ = 0, $\forall x > 0$. Sehingga $Y_n$merupakan **statistik cukup lengkap untuk *θ*.**

2.5. Penaksir Tak Bias Bervariansi Minimum Seragam

Penaksir tak bias bervariansi minimum merupakan penaksir yang mempunyai variansi yang minimum di antara semua kemungkinan penaksir tak bias, walaupun penaksir tidak selalu ada. Namun demikian, suatu penaksir tak bias bisa mencapai batas bawah ketaksamaan Cramer–Rao. Batas bawah ini diperoleh dengan mencari suatu fungsi tak bias dari suatu statistik cukup yang lengkap.

***Definisi* 5**. Jika dimisalkan $E_\theta t_i \Big|_{\equiv}^{\theta} \theta_i$ untuk i =1,2,…, r, dan di antara semua penaksir tak bias θ_i , t_i mempunyai varian terkecil untuk setiap kemungkinan nilai θ, maka t_i disebut **penaksir takbias dengan variansi minimum seragam (PTVMS) dari *θ*.**

Perhatikan Teorema Rao-Blackwell berikut, yang akan memberikan teknik mencari penaksir tak bias Y, dimana diketahui suatu penaksir takbias $\varphi(X)$yang baru dari θ sedemikian rupa sehingga $\varphi(X)$ mempunyai variansi yang lebih kecil dari variansi Y secara menyeluruh (seragam).

***Teorema* 1:** Teorema **Rao-Blackwell.** Apabila X dan Y peubah acak yang memenuhi $E(Y)=\mu$dan $Var(Y) = \sigma^2_Y < +\infty$. Andaikan $\varphi(x) = E(Y|X = x)$, maka $E\varphi(X) = \mu$ dan $Var\varphi(X) \le \sigma^2 Y$. Kesamaan iniberlakujika $P[Y = \varphi(X)] = 1$.

Bukti:
Misalkan $f(x,y)$menyatakan fungsi kepadatan gabungan dari (X,Y), dimana $f_1(x)$ fungsi padat dari$X$, $f_2(y)$ fungsi padat dari Y, dan $h(y/x)$ fungsi padat bersyarat dari Y jika diketahui X = x. Pertama akandiperlihatkanbahwa $E\varphi(X) = \mu.$ Mengingat

$$E(Y/X=x)=\int_{-\infty}^{\infty} yh(y/x)dy=\frac{\int_{-\infty}^{\infty} yf(x,y)dy}{f_1(x)}=\varphi(x)$$

maka

$$\int_{-\infty}^{\infty} yf(x,y)dy=\varphi(x)f_1(x)$$

sehingga

$$E\varphi(X)=\int_{-\infty}^{\infty}\varphi(x)f_1(x)dx=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} yf(x,y)dy\right]dx$$
$$=\int_{-\infty}^{\infty} y\left[\int_{-\infty}^{\infty} f(x,y)dx\right]dy=\int_{-\infty}^{\infty} yf_2(y)dy=\mu$$

Selanjutnya akandiperlihatkanbahwa $Var\varphi(X)\le\sigma_Y{}^2$; kesamaanberlakujika $P[Y=\varphi(X)]=1$. Makadiperoleh

$$\sigma_Y{}^2=E(Y-\mu)^2=E[Y-\varphi(X))+(\varphi(X)-\mu)]^2$$
$$=E(Y-\varphi(X))^2+E(\varphi(X)-\mu)^2+2E(Y-\varphi(X))(\varphi(X)-\mu)$$
$$=E(Y-\varphi(X))^2+Var\varphi(X)+0.$$

Sehingga

$$Var\varphi(X)=\sigma_Y{}^2-EY-\varphi(X))^2\le\sigma_Y{}^2,$$

Kesamaanberlakujika $E(Y-\varphi(X))^2=0$, danterakhirbenarjika $P[Y=\varphi(X)=1$.□

Sebagai ilustrasi perhatikan contoh berikutini, andaikan $X_1,\ldots,X_n$ peubahacakbebas, masing-masing $N(\mu,\sigma^2)$, μ dan σ^2 keduanya tidak diketahui. PTVMS dapat dicari dengan terlebih dahulumenunjukkanbahwa $X_1,\ldots,X_n$ peubahacakbebas, dimana $(\varphi_1(X_1,\ldots,X_n),\ldots,\varphi_r(X_1,\ldots,X_n))$ yang merupakanstatistik cukup minimum untuk θ, $\theta=(\theta_1,\ldots,\theta_r)$, makadiperolehstatistikcukup minimum lengkap $T(X_1,\ldots,X_n)=(\sum X_i,\sum X_i^2)$ untuk (μ,σ^2). Juga $\overline{X}_n=\sum X_i/n$ dan $s_n{}^2=\sum(X_i-\overline{X}_n)/(n-1)$ adalahfungsi takbias dari statistik cukuplengkap :

$$E(\overline{X}_n)=\mu$$
$$E(s_n^2)=\sigma^2$$

sehingga

$\overline{X}_n$ suatu PTVMS untuk μ

s_n^2 suatu PTVMS untuk σ^2 .

Daftar Pustaka

[1]. Dudewicz E.J., and Mishra, S.N., 1988, "*Modern Mathematical Statistics*",John Wiley & Sons Inc., New York.

[2]. Hogg, R.V., and Craig, E., 1995, "*Introduction to Mathematical Statistics*, 5th *Edition*",Prentice-Hall, New York.