BAB II

MENGESTIMASI SUATU PARAMETER DENGAN MENGGUNAKAN TITIK ESTIMASI

1. METODE UNTUK MENDAPATKAN ESTIMATOR * ESTIMATOR

Misalkan $X_1,\ldots,X_n$ adalah n sampel random dari suatu density f $(.;\theta)$ yang diketahui, tetapi parameter nya (yaitu θ) tak diketahui. Kemudian jika θ adalah suatu vektor dari bilangan-bilangan real, katakanlah $\theta=(\theta_1,\ldots\theta_k)$, dimana $\theta_1,\ldots,\theta_k$ adalah k parameter (disini k=1) Dan jika $\underline{\Theta}$ adalah ruang parameter, yang menunjukan himpunan dari harga-harga yang mungkin dari parameter θ tersebut maka yang diinginkan adalah untuk mendapatkan suatu statistik yang digunakan sebagai estimator dari fungsi-fungsi parameter dari θ, katakanlah $\tau_1(\theta),\tau_2(\theta),\ldots$ $\tau_k(\theta)$ dari $\theta=(\theta_1,\ldots,\theta_k)$.

Catatan :

Statistik didefinisikan sebagai suatu fungsi dari random variabel yang dapat diobservasi dan tak memuat parameter yang tak diketahui. Selanjutnya suatu estimator akan didefinisikan seperti dibawah ini.

DEFINISI 1 ESTIMATOR.

Sebarang statistik yang harganya digunakan untuk mengestimasi $\tau(\theta)$, dimana $\tau(\theta)$ adalah fungsi dari parameter θ didefinisikan sebagai estimator dari $\tau(\theta)$. Jika X_1 $\ldots,X_n$ adalah n sampel random dari density f $(.;\theta)$ yang digunakan untuk mengestimasi $\tau(\theta)$ (dimana $\tau(.)$ adalah fungsi dari θ) dan jika $t(X_1,\ldots,X_n)$ adalah estimator dari $\tau(\theta)$ maka $t(X_1,\ldots,X_n)$ dapat dipikirkan dalam 2 cara :

1. Sebagai random variabel katakanlah $T=t(X_1..X_n)$.

2. Sebagai fungsi t (........)

Catatan

T menyatakan random variabel t $(X_1, ...X_n)$

T menyatakan fungsi t (, ...)

t menyatakan harga dari T sehingga $t = t(x_1....x_n)$

Suatu "Statistik" digunakan sebagai estimator dan "Harga dari statistik" tersebut digunakan untuk mengestimasi parameter yang takdiketahui. Jadi kata "Estimator" untuk fungsi dan kata "Estimasi" untuk suatu harga dari fungsi tersebut sebagai contoh :

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ adalah suatu estimator untuk mean

Jadi T adalah $\bar{X}_n$, $t = \bar{x}_n$ dan t (........) adalah fungsi yang didefinisikan dengan jumlahan dari argumen argumennya yang kemudian dibagi dengan n.

NOTASI

$\hat{\theta}$ digunakan untuk mengestimasi θ, yang lebih umum lagi adalah vektor $(\hat{\theta}_1, ... \hat{\theta}_k)$ digunakan untuk mengestimasi $(\theta_1......\theta_k)$, dimana setiap $\hat{\theta}_j$ mengestimasi θ_j, untuk $J = 1, ..., k$. Jadi jika $\hat{\hat{\theta}}$ adalah estimasi dari θ, maka $\hat{\theta}$ adalah estimator dari θ yang sesuai.

1.1. METODE MOMENT

Misalkan $f(..;\theta_1,.....\theta_k)$ adalah suatu density dari suatu random variabel X yang mempunyai K parameter yaitu $\theta_1, ...\theta_k$.

Dan misalkan μ'_r menunjukkan moment ke r mendekati 0, sehingga $\mu'_r = E\left[X^r\right]$

Biasanya μ'_r adalah suatu fungsi dari K parameter sehingga $\mu'_r = \mu'_r \ (\theta_1, \ldots\ldots\ldots \theta_k)$.

Jika $X_1 \ldots\ldots X_n$ adalah n random sampel dari density $f(\cdot;\theta_1, \ldots\ldots \theta_k)$ dan jika M' adalah moment sampel ke J maka

$$M'_J = \frac{1}{n} \sum_{i=1}^{n} X_i^J$$

Dari M'_J akan didapatkan k persamaan :

$M'_J = M'_J \ (\theta_1, \ldots\ldots \theta_k)$ dimana j= 1,..k ... (1)

dengan k variabel yaitu $\theta_1, \ldots\ldots \theta_k$.

Dan jika $\hat{\theta}_1, \ldots\ldots \hat{\theta}_k$ adalah solusi dari persamaan (1) (misalkan solusi tunggal) maka estimator $(\hat{\theta}_1, \ldots\ldots \hat{\theta}_k)$ (dengan $\hat{\theta}_j$ adalah estimasi dari θ_j) adalah estimator dari $(\theta_1, \ldots\ldots \theta_k)$ yang dapat diperoleh dengan metode moment tersebut diatas.

Estimator - estimator diperoleh dengan mengganti moment populasi dengan moment sampel.

Sebagai contoh.

Jika $X_1, \ldots\ldots X_n$ adalah n random sampel dari suatu distribusi uniform pada $(\mu - \sqrt{3}\,\sigma, \mu + \sqrt{3}\,\sigma)$ maka parameter yang tak diketahui adalah μ dan σ , yang manamerupakan mean dan standard deviasidari populasi.

Sehingga persamaan metode moment adalah :

$M'_1 = \mu'_1 = \mu'_1 \ (\mu, \sigma) = \mu$ dan

$M'_2 = \mu'_2 = \mu'_2 \ (\mu, \sigma) = \sigma^2 + \mu^2$

Jadi estimator metode moment adalah $\overline{X}$ untuk μ dan

$\sqrt{\frac{1}{n}\sum X_i^2 - \overline{X}^2} = \sqrt{\frac{1}{n}\sum (X_i - \overline{X})^2}$ untuk σ

1.2. METODE LIKELIHOOD MAXIMUM.

Untuk mencari estimator dengan metode likelihood maximum akan didefinisikan lebih dulu fungsi likelihood dibawah ini.

DEFINISI 2 FUNGSI LIKELIHOOD.

Fungsi likelihood dari n random variabel $X_1, \ldots, X_n$ didefinisikan sebagai density berserikat dari n random variabel, katakanlah $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n,\theta)$ yang merupakan fungsi dari θ.

Khususnya jika $X_1,\ldots,X_n$ adalah n sampel random dari density $f(x;\theta)$ maka fungsi likelihood adalah :

$$f(X_1;\theta).\ f(X_2;\theta)\ldots.\ f(X_n,\theta).$$

NOTASI

Untuk fungsi likelihood yang merupakan fungsi dari θ digunakan notasi $L(\theta;x_1..x_n;\theta)$ atau $L(.,x_1,..,x_n)$.
Jika $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta)$ adalah density berserikat dari n random variabel, dimana θ tak diketahui.
Dan jika harga yang diamati adalah $x_1',\ldots,x_n'$ maka untuk mendapatkan harga dari θ dalam $\underline{\Theta}$ yang ditunjukkan oleh $\hat{\theta}$, yang memaximalkan fungsi likelihood .... $L(\theta,x_1',\ldots\ldots,x_n')$. Harga $\hat{\theta}$ yang memaximalkan fungsi likelihood adalah fungsi dari $x_1, x_2, \ldots, x_n$; katakanlah $\hat{\theta} = \hat{\vartheta}(x_1,\ldots,x_n)$. Dan random variabel .... $\hat{\theta} = \hat{\vartheta}(X_1,\ldots,X_n)$ dinamakan estimator likelihood maximum dari θ.

DEFINISI 3 ESTIMATOR LIKELIHOOD MAXIMUM.

Misalkan $L(\theta) = L(\theta;x_1,\ldots,x_n)$ adalah fungsi likelihood untuk random variabel $X_1,..,X_n$. Jika $\hat{\theta}$ adalah harga-harga dari θ dalam $\underline{\Theta}$ yang memaksimumkan $L(\theta)$

Maka $\hat{\theta} = \hat{\vartheta}(X_1,\ldots,X_n)$ adalah estimator likelihood maksimum dari θ, dan $\hat{\theta} = \hat{\vartheta}(x_1,\ldots,x_n)$ adalah estimasi likelihood maximum dari θ untuk sampel $x_1,\ldots,x_n$.

Catatan :

$\hat{\theta} = \hat{\vartheta}(x_1,\ldots,x_n)$ merupakan fungsi dari observasi $x_1,\ldots,x_n$, jika $X_1,\ldots,X_n$ adalah n sampel random dari beberapa fungsi density $f(x;\theta)$ maka fungsi likelihood adalah $L(\theta) = f(x_1,\theta)\, f(x_2,\theta) \ldots f(x_n,\theta)$. Jadi estimator likelihood maximum merupakan solusi dari persamaan $\dfrac{\partial L(\theta)}{\partial \theta} = 0$

Jika $L(\theta)$ dan $\log L(\theta)$ masing-masing merupakan fungsi likelihood dan log naturalis dari fungsi likelihood maka kedua-duanya mempunyai titik maksimum yang harganya sama dengan 0.

Jika fungsi likelihood mempunyai k parameter sehingga $L(\theta_1,\ldots,\theta_k) = \prod_{i=1}^{n} f(x_i;\theta_1,\ldots,\theta_k)$ maka estimator-estimator likelihood maksimum dari parameter-parameter $\theta_1,\ldots\theta_k$ adalah random variabel $\hat{\theta}_1 = \hat{\vartheta}_1(X_1,\ldots,X_n)$, $\hat{\theta}_2 = \hat{\vartheta}_2(X_1,\ldots,X_n)$, $\hat{\theta}_k = \hat{\vartheta}_k(X_1,\ldots,X_n)$, dimana $\hat{\theta}_1,\ldots,\hat{\theta}_k$ adalah harga-harga dalam $\bar{\theta}$ yang memaksimumkan $L(\theta_1,\ldots,\theta_k)$, jadi estimator-estimator likelihood maksimum adalah solusi dari k persamaan di bawah ini :

$$\frac{\partial L(\theta_1,\ldots,\theta_k)}{\partial \theta_1} = 0$$

$$\frac{\partial L(\theta_1,\ldots,\theta_k)}{\partial \theta_k} = 0$$

sebagai gambaran akan kita berikan contoh dibawah ini.

Jika $X_1, \ldots, X_n$ adalah n observasi yang ditarik dari suatu populasi normal dengan mean yang sama yaitu μ dan varian yang berbeda yaitu $\sigma_1^2, \ldots, \sigma_n^2$ maka apakah mungkin untuk mengestimasi semua parameter tersebut ?

Dan jika σ_i^2 diketahui maka ditanyakan estimator like lihood maximum dari μ ?

Penyelesaian.

Fungsi density dari populasi normal adalah.

$$f(x_i;\mu,\sigma_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma_i}\right)^2\right),$$

dengan $i = 1, \ldots, n$

Fungsi likelihood maximum adalah :

$$L = L(\mu;\sigma_1^2, \ldots, \sigma_n^2 ; x_1, \ldots, x_n)$$

$$= \prod_{i=1}^{n} f(x_i;\mu,\sigma_i)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\tfrac{1}{2}\left(\frac{x_i-\mu}{\sigma_i}\right)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma_1\cdots\sigma_n} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma_i}\right)^2\right]$$

$$\frac{\partial L}{\partial \mu} = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma_1\cdots\sigma_n}\right) \exp\left[-\tfrac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma_i}\right)^2\right]$$

$$\cdot \left[\sum_{i=1}^{n} \frac{x_i-\mu}{\sigma_i^2}\right]$$

$$\frac{\partial L}{\partial \mu} = 0 \Longrightarrow \sum_{i=1}^{n} \frac{x_i-\mu}{\sigma_i^2} = 0 \quad \ldots\ldots\ldots (1)$$

$$\frac{\partial L}{\partial \sigma_1^2} = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma_1^3\,\sigma_2 \cdots \sigma_n} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma_i}\right)^2\right] +$$

$$\left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{(-\frac{1}{2})}{\sigma_1 \cdots\cdots \sigma_n} \exp\left[-\frac{1}{2}\sum\left(\frac{x_i-\mu}{\sigma_i}\right)^2\right] \cdot$$

$$(-\tfrac{1}{2})\,(x_i-\mu)^2 \quad \left(\frac{-1}{\sigma_1^4}\right)$$

$$\frac{\partial L}{\partial \sigma_1^2} = 0 \Longrightarrow \frac{-1}{\sigma_1^2} + \frac{(x_1-\mu)}{\sigma_1^4} = 0$$

$$\therefore\ \sigma_1^2 = (x_1-\mu)^2 \quad \ldots\ldots\ldots\ (2)$$

Dengan cara yang sama maka :

$$\sigma_2^2 = (x_2-\mu)^2 \quad \ldots\ldots\ldots\ldots\ (3)$$

*

*

*

*

$$\sigma_n^2 = (x_n-\mu)^2 \quad \ldots\ldots\ldots\ldots(n+1)$$

Dari persamaan (1) sampai dengan persamaan (n+1), terlihat bahwa μ dan setiap σ_i^2 untuk i = 1,2,..... ,n tak mungkin kita estimasi secara bersama-sama. Jadi dengan kata lain parameternya tidak dapat diestimasi semuanya, sebab parameter-parameter tersebut saling bergantungan. Jika σ_i^2 diketahui, dimana i = 1,2, ...,n dan dari persamaan (1) yaitu $\sum_{i=1}^{n} \frac{x_i-\mu}{\sigma_i^2} = 0$ maka

$$\mu = \frac{\sum_{i=1}^{n} x_i/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2}$$

Jadi estimator likelihood maximum dari μ adalah :

$$\frac{\sum_{i=1}^{n} X_i/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2}$$ dimana setiap σ_i diketahui untuk $i = 1,2, \ldots ,n$.

THEOREMA 1 SIFAT INVARIAN DARI ESTIMATOR-ESTIMATOR LIKE HOOD MAXIMUM.

Misal $\hat{\Theta} = \hat{\varphi}(X_1, \ldots ,X_n)$ adalah estimator likelihood maximum dari θ dalam density $f(x ; \theta)$, dimana θ berdimensi satu. Jika $\tau(.)$ adalah suatu fungsi dengan inversnya berharga satu, maka estimator-estimator likelihood maximum dari $\tau(\theta)$ adalah $\tau(\hat{\theta})$

B u k t i :

Ambil $\hat{\theta}$ adalah estimasi likelihood maximum dari θ maka yang dibuktikan adalah:

$$L(\tau(\theta); x_1, \ldots ,x_n) \leq L(\tau(\hat{\theta}); x_1, \ldots ,x_n)$$

misal τ^{-1} adalah fungsi invers dari $\tau(.)$ yang harganya 1 maka $\tau^{-1} = 1$.

$$\begin{aligned}
L(\tau(\theta); x_1, \ldots,x_n) &= L(\tau^{-1}(\tau(\theta); x_1, \ldots ,x_n) \\
&= L(\tau^{-1}(\tau(\theta); x_1, \ldots ,x_n) \\
&= L(\theta; x_1, \ldots ,x_n) \\
&\leq L(\theta; x_1, \ldots ,x_n) \text{ sebab } \hat{\theta} \text{ merupakan estimasi likelihood maksimum.} \\
&= L(\tau^{-1}\tau(\hat{\theta}); x_1, \ldots ,x_n) \\
&= L(\tau(\hat{\theta}); x_1, \ldots ,x_n)
\end{aligned}$$

Jadi terbuktilah bahwa :

$$L\,(\tau(\theta);\, x_1, \ldots, x_n) \leqslant L\,(\tau(\hat{\theta});\, x_1, \ldots, x_n).$$

Jadi estimator likelihood maximum dari $\tau(\theta)$ adalah - $\tau(\hat{\theta})$ terbukti.

THEOREMA 2.

Misal $\hat{\Theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$, dimana $\hat{\Theta}_j = \hat{\varphi}_j\ (X_1, \ldots X_n)$ adalah suatu estimator likelihood maximum dari $\theta = (\theta_1, \ldots, \theta_k)$ dalam density $f\,(.\,;\theta_1, \ldots, \theta_k)$ Jika $\tau(\theta) = (\tau_1,(\theta), \ldots, \tau_k\,(\theta))$ untuk $1 \leqslant r \leqslant k$ adalah suatu satu transformasi dari ruang parameter Θ, maka suatu estimator likelihood maximum dari $\tau(\theta) = (\tau_1\,(\theta), \ldots\ldots\ \tau_r(\theta))$ adalah $\tau\,(\hat{\theta}) = (\tau_1(\hat{\theta}), \ldots \tau_r(\hat{\theta}))$.

B U K T I :

Ambil $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ adalah estimasi likelihood maximum dari $\theta = (\theta_1, \ldots, \theta_k)$. Cukup menunjukkan bahwa $M\,(\tau\,(\hat{\theta};\, x_1, \ldots, x_n) \geqslant M\,(\tau;\, x_1, \ldots, x_n)$ untuk sebarang $\tau \in T$

$$M\,(\tau; x_1, \ldots, x_n) = \sup_{\{\theta;\,\tau(\theta)=\tau\}} L\,(\theta;\, x_1, \ldots\ldots, x_n)$$

$$\leqslant \sup_{\theta \in \Theta} L\,(\theta;\, x_1, \ldots\ldots, x_n)$$

$$= L\,(\hat{\theta};\, x_1, \ldots, x_n)$$

$$= \sup_{\{\theta;\,\tau(\theta)=\tau(\hat{\theta})\}} L\,(\theta;\, x_1, \ldots\ldots, x_n)$$

$$= M\,(\tau(\hat{\theta})\,;\, x_1, \ldots\ldots, x_n)$$

Karena $M\,(\tau; x_1, \ldots, x_n) \leqslant M\,(\tau\,(\hat{\theta});\, x_1, \ldots\ldots, x_n)$ terbuktilah bahwa estimator likelihood maximum dari

$\tau(\theta) = (\tau_1(\theta), \ldots, \tau_r(\theta))$ adalah

$\tau(\hat{\theta}) = (\tau_1(\hat{\theta}), \ldots, \tau_r(\hat{\theta}))$

Catatan:

Jika $\tau_j(\theta) = \tau_j(\theta_1, \ldots, \theta_k)$ maka estimator likelihood maximum dari $\tau_j(\theta_1, \ldots, \theta_k)$ adalah $\tau_j(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ dimana J = 1, 2, ...,r .

2. SUFFICIENSI DARI SUATU STATISTIK.

Untuk mendapatkan estimator-estimator yang terbaik maka dalam fasal ini akan dibahas statistik - statistik yang sufficien. Dimana statistik yang sufficien- tersebut dapat digunakan untuk mengestimasi parameter O yang diinginkan. Selanjutnya pengertian dan definisi - yang berhubungan dengan statistik yang sufficien tersebut akan dibahas dalam fasal-fasal berikut dibawah ini.

2.1. STATISTIK YANG SUFFICIEN

Jika $X_1, \ldots, X_n$ adalah n sampel random dari - suatu density $f(\cdot\,;\theta)$ maka untuk mendefinisiksn suatu statistik yang sufficien yaitu merupakan suatu jenis khusus dari statistik yang mana cukup untuk tujuan estimasi suatu parameter.

DEFINISI 4 STATISTIK YANG SUFFICIEN.

Jika $X_1, \ldots, X_n$ adalah n random sampel dari density $f(\cdot\,;\theta)$ dimana para meter θ mungkin suatu vektor maka statistik $S = s(X_1, \ldots, X_n)$ didefinisikan sebagai suatu statistik yang sufficien jika dan hanya jika distribusi bersyarat dari $X_1, \ldots, X_n$ dengan syarat $S = s$ tidak tergantung pada θ untuk setiap harga s dari S.

C a t a t a n.

Untuk menunjukkan statistik yang sufficien kita gunakan statistik $S = s(X_1, .., X_n)$.

Definisi 4 diatas menyatakan bahwa suatu statistik $S = s(X_1, ..., X_n)$ adalah sufficien jika distribusi bersyarat dari sampel dengan syarat harga statistik S tak tergantung dari θ.

Diffinisi diatas mempunyai suatu kelemahan, yaitu :

(1). Definisi tersebut tidak memberi tahukan statistik mana yang bisa menjadi statistik yang sufficien.

(2). Definisi tersebut menekankan untuk memperoleh suatu distribusi bersyarat yang mungkin tidaklah mudah, khususnya untuk random variabel kontinu.

Maka dari itu akan didefinisikan statistik yang sufficien lainnya dibawah ini:

DEFINISI 5 STATISTIK YANG SUFFICIEN.

Misalkan jika $X_1, .., X_n$ adalah n sampel random dari fungsi density $f(. ;\theta)$ maka suatu statistik $S = s(X_1, .., X_n)$ didefinisikan sebagai suatu statistik yang sufficien jika dan hanya jika distribusi bersyarat dari T dengan syarat S tidak tergantung pada θ untuk sebarang statistik $T = t(X_1, ..., X_n)$.

Definisi 5 khususnya berguna dalam menunjukkan bahwa statistik yang khusus adalah tidak sufficien. Misalnya untuk membuktikan bahwa suatu statistik $T' = t'(X_1, .., X_n)$ untuk mana distribusi bersyarat dari T dengan syarat T' tergantung pada θ. Maka $T' = t'(X_1, ..., X_n)$ adalah Statistik yang tidak sufficien

Untuk beberapa masalah, tak terdapat hanya satu statis tik yang sufficien saja, tetapi akan terdapat bebera pa statistik-statistik yang sufficien dan berserikat.

DEFINISI 6 STATISTIK YANG SUFFICIEN BERSERIKAT.

Misalkan $X_1, .. ,X_n$ adalah n sampel random dari density f (. ; θ) statistik-statistik $S_1, .. ,S_r$ didefinisikan sebagai statistik-statistik yang sufficien dan berserikat jika dan hanya jika dis tribusi bersyarat dari $X_1, .. ,X_n$ dengan syarat $S_1 = s_1, .. ,S_r = s_r$ tak tergantung dari θ.

Jadi dengan menggunakan definisi 6 maka sampel X_1 $..,X_n$ adalah statistik yang sufficien berserikat sebab distribusi bersyarat dari sampel dengan syarat sampel nya sendiri tidak tergantung pada θ.

THEOREMA 3.

Jika $S_1 = s_1 (X_1, .. ,X_n),\ldots , S_n = s_n (X_1, \ldots ,X_n)$ adalah himpunan dari statistik- statistik yang sufficien berserikat maka setiap himpunan dari fungsi satu-satu, atau transformasi dari $S_1 ,\ldots,S_n$ adalah juga sufficien berserikat.

B UK T I :

Misalkan $X_1, .. ,X_n$ adalah n random sampel dari density f (. ;θ). Ambil sebarang himpunan $Y_1, .. . ,Y_n$ yang dibentuk dengan $S_i \rightarrow f(S_i) = Y_i$ adalah fungsi-satu-satu dengan i = 1, ... ,n dan andaikan bahwa $Y_1,\ldots,Y_n$ adalah himpunan dari statistik yang tidak suffi cien maka $f_{X_1,\ldots,X_n / Y_1, \ldots,Y_n}(X_1,\ldots,X_n)/(Y_1,\ldots,Y_n ;\theta)$ tergantung pada θ. Dan oleh karena setiap Y_i adalah fungsi dari S_i dan untuk S_i adalah merupakan -

fungsi dari S_i nya sendiri, dimana $i = 1, 2, \ldots, n$ maka :

$$f_{X_1,\ldots,X_n / S_1,\ldots,S_n}(x_1,\ldots,x_n / s_1,\ldots s_n;\, \theta)$$

pun tergantung pada θ. Dan karena $S_1,\ldots,S_n$ adalah himpunan dari statistik - statistik yang sufficien dan berserikat maka :

$$f_{X_1,\ldots,X_n / S_1,\ldots,S_n}(x_1,\ldots,x_n / s_1,\ldots,s_n;\, \theta)$$

tak tergantung pada θ.

Hal ini berlawanan dengan yang diketahui, jadi pengandaian diingkar, maka yang benar adalah bahwa $Y_1, \ldots, Y_n$ adalah himpunan dari statistik-statistik yang sufficien dan berserikat. Dan karena $Y_1,\ldots,Y_n$ sebarang, maka setiap himpunan dari fungsi satu-satu dari $S_1,\ldots,S_n$ adalah sufficien dan berserikat. Jadi Theorama terbukti.

Contoh :

Misalkan X adalah suatu observasi tunggal dari $N(0\,;\,\theta)$ dimana $\theta = \sigma^2$

Ditanyakan : Apakah X suatu statistik yang sufficien dan tentukan estimator likelihood maksimum dari θ.

Penyelesaian :

$$f_X(x;\theta) = \frac{1}{\sqrt{2\pi\theta}}\, e^{-\frac{1}{2}(x^2/\theta)}$$

$$f_{X/X=x}(x/x;\theta) = \frac{f_X(x)}{f_X(x)} = 1$$ Jadi tak tergantung pada θ

Jadi X adalah suatu statistik yang sufficien.

Fungsi likelihood untuk random variabel X adalah :

$$L = \frac{1}{\sqrt{2\pi\theta}}\, e^{-\frac{1}{2}(x^2/\theta)}$$

$$\frac{\partial L}{\partial\theta} = \frac{-\frac{1}{2}}{\sqrt{2\pi\theta^3}}\, e^{-\frac{1}{2}(x^2/\theta)} + \frac{1}{\sqrt{2\pi\theta}}\, e^{-\frac{1}{2}(x^2/\theta)} \cdot \left(-\frac{x^2}{2}\right) \cdot \left(-\theta^{2}\right)$$

$$\frac{\partial L}{\partial \theta} = 0 \Longrightarrow \frac{-\frac{1}{2}}{\sqrt{2\pi\theta^3}}\, e^{-\frac{1}{2}(x^2/\theta)} + \frac{1}{\sqrt{2\pi\theta^5}}\, e^{-\frac{1}{2}(x^2/\theta)} \cdot \frac{x^2}{2}$$

$$\theta^{-3/2} = \theta^{-5/2}\, x^2$$

$$\theta = x^2 \rightarrow \sqrt{\theta} = x$$

Jadi estimator likelihood makximum dari θ adalah X

THEOREMA 4

Suatu estimator likelihood maximum atau himpunan-estimator-estimator likelihood maximum tergantung pada sampel yang merupakan fungsi dari himpunan statistik statistik yang sufficien dan berserikat.

B U K T I :

Jika $S_1 = s_1(x_1,\ldots,x_n),\ldots\ S_k = s_k(x_1,\ldots,x_n)$ adalah statistik yang sufficien berserikat maka fungsi likelihood dapat ditulis sebagai.

$$L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$= g(s_1(x_1, \ldots, x_n), \ldots, s_k(x_1, \ldots, x_n); \theta)$$

$$\cdot\, h(x_1, \ldots, x_n).$$

Yang merupakan fungsi dari θ, $L(\theta; X_1, \ldots, X_n)$ akan mempunyai titik maximum pada kedudukan yang sama dengan titik maksimumnya $g(S_1, \ldots, S_k; \theta)$, tetapi kedudukan di mana g mencapai maximum dapat tergantung pada $x_1, \ldots, x_n$ saja, oleh karena g merupakan fungsi dari $S_1, \ldots, S_k$.

2.2. STATISTIK - STATISTIK MINIMAL YANG SUFFICIEN.

Jika suatu himpunan dari statistik - statistik - mempunyai himpunan bagian yang lebih kecil dimana -

Dalam partisi yang dihasilkannya dari pertisi himpunan - himpunan statistik yang lain, maka statistik yang pertama dikatakan lebih padat dari pada statistik yang terakhir.

Suatu himpunan dari statistik yang sufficien minimal adalah suatu himpunan dari statistik - statistik yang sufficien, yang mempunyai himpunan bagian yang lebih sedikit dalam partisinya dari pada partisi yang dihasil sebarang himpunan statistik - statistik yang sufficien-lainnya. Dengan kata lain dikatakan bahwa suatu himpunan dari statistik yang sufficien adalah minimal jika tidak ada himpunan dari statistik yang sufficien lain nya yang lebih padat dari pada statistik yang sufficien pertama.

DEFINISI 7 STATISTIK YANG SUFFICIEN MINIMAL

Suatu himpunan dari statistik - statistik yang sufficien berserikat didefinisikan sebagai statistik yang sufficien minimal jika dan hanya jika ia merupakan fungsi dari setiap himpunan dari statistik yang sufficien lainnya.

3. ESTIMASI YANG UNBIASE (UNBIASE ESTIMATION).

Oleh karena estimator - estimator dengan rata-rata kesalahan kwadrat jarang terdapat, suatu cara yang - baik adalah membatasi klas dari fungsi estimasi dan mencari estimator - estimator dengan rata-rata kesalahan kwadrat minimum uniform didalam klas yang terbatas tersebut.

Satu cara untuk membatasi klas dari fungsi estimasi yang akan dibicarakan hanyalah estimator-estimator yang unbiase saja dan diantara klas dari estimator-estimator yang unbiase, lalu untuk menyelidiki suatu estimator-estimator dengan rata-rata kesalahan kwadrat minimum

Rata - rata kesalahan kwadrat dari suatu estimator T terhadap $\tau(\theta)$ dapat ditulis sebagai :

$$E_\theta\left[(T-(\tau(\theta))^2\right] = Var_\theta(T) + \left\{\tau(\theta) - E_\theta(T)\right\}^2$$

dan jika T adalah suatu estimator yang unbiase dari (θ), maka $E_\theta(T) = \tau(\theta)$ dan karena itu $E_\theta\left[(T-\tau(\theta))^2\right] = Var_\theta(T)$

Jadi untuk mencari suatu estimator dengan rata - rata kesalahan kwadrat minimum uniform diantara estimator- estimator yang unbiase adalah sama dengan untuk mencari suatu estimator dengan varian minimum uniformdiantara estimator - estimator yang unbiase.

DEFINISI 8 UNBIASE ESTIMATOR VARIAN MINIMUM UNIFORM (UMVUE).

Misalkan $X_1, .., X_n$ adalah n random sample dari $f(.;\theta)$ suatu estimator $T^* = t^*(X_1, .., X_n)$ dari $\tau(\theta)$ didefinisikan sebagai suatu unbiase estimator Varian minimum uniform dari $\tau(\theta)$ jika dan hanya jika :

i) $E_\theta(T^*) = \tau(\theta)$, sehingga T^* adalah unbiase

ii) $Var_\theta(T^*) = Var_\theta(T)$ untuk setiap estimator

$T = t(X_1,..,X_n)$ dari $\tau(\theta)$ yang lain, yang memenuhi $E_\theta(T) = \tau(\theta)$.

DEFINISI 9 KELUARGA LENGKAP DARI DENSITY.

Misalkan $X_1, ..., X_n$ adalah n random sample dari $f(.;\theta)$ dengan ruang parameter Θ dan misalkan

$T = t\ (X_1, \ldots ,X_n)$ adalah suatu statistik. Keluarga density dari T didefinisikan adalah lengkap jika dan hanya jika $E_\theta \left[Z\ (r) \right] = \theta$ untuk semua $\theta \in \bar{\underline{\theta}}$. Berarti bahwa $P_\theta \left[Z\ (T) = \theta \right] = 1$ untuk semua $\theta \in \bar{\underline{\theta}}$, dimana Z (T) adalah suatu statistik.

DEFINISI 10 STATISTIK YANG LENGKAP.

Statistik T didefinisikan adalah lengkap jika dan hanya jika keluarga dari densitynya adalah lengkap.

THEOREMA 5.

Misal $X_1, \ldots ,X_n$ adalah random sample dari density $f\ (.;\theta)$, $\theta \in \bar{\underline{\theta}}$. dimana $\bar{\underline{\theta}}$ adalah suatu interval (mungkin infinite).

Jika $f\ (x\ ;\ \theta) = a\ (\theta).b\ (x).\ \exp \left[c\ (\theta).\ d\ (x) \right]$, sehingga $f\ (.\ ;\ \theta)$ adalah anggota dari keluarga eksponensial berparameter satu, maka $\sum d(x_i)$ adalah suatu statistik lengkap dan sufficien yang minimal.

B u k t i : $\prod_{i=1}^{n} f(x_i;\theta) = a^n(\theta) \left[\prod_{i=1}^{n} b(x_i)\ \exp\ (c(\theta) \sum_{i=1}^{n} d(x_i)) \right]$

$$= \left\{ a^n(\theta) \exp \left[c(\theta) \sum_{i=1}^{n} d(x_i) \right] \right\} \left\{ \prod_{i=1}^{n} b(x_i) \right\}$$

maka $\prod_{i=1}^{n} f(x_i;\theta)$ dapat dipandang sebagai

$g(S(x_1, \ldots ,x_n);\ \theta).\ h(x_1, \ldots ,x_n)$ dengan

$g(S(x_1, \ldots ,x_n);\theta) = \left\{ a^n(\theta)\ \exp \left[c(\theta) \sum_{i=1}^{n} d(x_i) \right] \right\}$ dan

$h(x_1, \ldots ,x_n) = \prod_{i=1}^{n} b\ (x_i)$

jadi $S\ (x_1, \ldots ,x_n) = \sum (d\ (x_i)\ )$

Dan menurut teorema faktorisasi maka $\Sigma d(x_i)$ merupakan statistik yang sufficien pandang $\Sigma d(x_i)$ maka $\Sigma d(x_i) = d(\Sigma(x_i))$. Jadi $\Sigma d(x_i)$ merupakan fungsidari Σx_i dimana Σx_i merupakan statistik yang sufficien.

Berarti $\Sigma d(x_i)$ adalah statistik minimal, jadi $\Sigma d(x_i)$ adalah statistik minimal yang sufficien. Pandanglah $Z(T) = 0$, dimana $Z(T)$ merupakan fungsi dari T dengan $T = \Sigma d(x_i)$ maka $E_\theta(Z(T)) \equiv 0$ untuk setiap $\theta \in \bar{\theta}$,yang berarti pula bahwa $P_\theta(Z(T) = \theta) = 1$ untuk setiap $\theta \in \bar{\theta}$ .jadi keluarga dari $f(x;\theta) = a(\theta).b(x) \exp(c(\theta).d(x))$ adalah lengkap. Karena keluarga $f(x;\theta)$ adalah lengkap maka $\Sigma d(x_i)$ adalah lengkap. Jadi $\Sigma d(x_i)$ adalah statistik-minimal yang sufficien dan lengkap.

THEOREMA 6

Misalkan $X_1, .. ,X_n$ adalah n sample random dari $S = s(X_1, .. ,X_n)$ adalah statistik yang sufficien-dan lengkap dan jika $T^* = t^*(S)$ yangmerupakanfungsi dari S, adalah suatu unbiase estimator dari $\tau(\theta)$, maka T^* adalah suatu umvue dari $\tau(\theta)$.

B U K T I :

Misal T adalah sebarang unbiase estimator dari $\tau(\theta)$ yang merupakan fungsi dari S, sehingga $T^* = t(\ )$maka $E_\theta[T^* - T'] = 0$ untuk semua $\theta \in \bar{\theta}$ dan T^*-T' adalah fungsi dari S, karena S adalah lengkap, maka $P_\theta[t^*(S) = t'(S)] = 1$

untuk semua $\theta \in \bar{\theta}$ oleh karena itu hanya ada satu unbiase estimator dari $\tau(\theta)$ yang merupakan fungsi dari S sebarang misalkan T adalah sebarang unbiase estimator dari $\tau(\theta)$.

Maka T^* haruslah sama dengan $E(T/S)$ oleh karena $E(T/S)$

Adalah unbiase estimotor dari $\tau(\theta)$ yang tergantung pada S. Menurut teorema 6 (Rao- Blackwell) , maka $Var_\theta\ (T^*)S \leqslant Var_\theta\ (T)$ untuk semua $\theta \in \bar{\Theta}$; jadi T^* adalah UMVUE dari $\tau(\theta)$.

Jadi teorema 6 diatas mengatakan :

1. Jika terdapat suatu statistik S yang sufficiecien dan lengkap dan jika terdapat suatu unbiase estimator untuk $\tau(\theta)$, maka terdapat suatu UMVUE.
2. UMVUE adalah unbiase estimator yang tunggal dari $\tau(\theta)$ yang merupakan fungsi dari S.

DEFINISI 11 KELENGKAPAN BERSERIKAT

Jika $X_1, \ldots, X_n$ adalah n sampel random dari suatu density $f(x;\theta_1, \ldots, \theta_n)$ dan jika $(T_1, \ldots, T_m)$ adalah himpunan dari statistik - statistik $T_1, \ldots, T_m$, maka $(T_1, \ldots, T_m)$ didefinisikan adalah lengkap - berserikat jika dan hanya jika.

$E_\theta\left[Z(T_1, \ldots, T_m)\right] = 0$ untuksemua $\theta \in \bar{\Theta}$, berarti - $P_\theta\left[Z(T_1, \ldots, T_m) = 0\right] = 1$ untuk semua $\theta \in \bar{\Theta}$ Dimana $Z(T_1, \ldots, T_m)$ adalah suatu statistik.

THEOREMA 7

Misalkan $X_1, \ldots, X_n$ adalah n sampel random dari suatu density $f(x; \theta_1, \ldots, \theta_k)$.

Jika $f(x; \theta_1, \ldots, \theta_k) = a(\theta_1, \ldots, \theta_k)\, b(x) \exp \left[\sum_{j=1}^{k} C_j(\theta_1, \ldots, \theta_k)\, d_j(x)\right]$, sehingga $f(x;\theta_1, \ldots, \theta_k)$ adalah anggota keluarga exponensial berparameter K, maka $\left(\sum_{i=1}^{n} d_1(x_i), \ldots, \sum_{i=1}^{n} d_k(x_i)\right)$ adalah suatu himpunan yang minimal dari statistik lengkap dan sufficien yang berserikat.

B U K T I :

$$\prod_{i=1}^{n} f\,(x_i;\theta_1,\,..\,,\theta_k) = a^n\,(\theta_1,\,...,\theta_k)\prod_{i=1}^{n} b\,(x_i)$$

$$\exp\left[\sum_{j=1}^{k} C_j\,(\theta_1,\,..\,,\theta_k)\sum_{i=1}^{n} d_j\,(x_i)\right]$$

$$= \left\{a^n\,(\theta_1,....,\theta_k)\,\exp\left[\sum_{j=1}^{k} C_j\,(\theta_1,....,\theta_k)\sum_{i=1}^{n} d_j(x_i)\right]\right\}$$

$$\left\{\prod_{i=1}^{n} b\,(x_i)\right\}$$

dengan menggunakan kriteria faktorisasi maka

$\sum_{i=1}^{n} d_j\,(x_i)$, dimana $j = 1,2,\,..\,,k$, merupakan statistik yang sufficien.

Karena $\Sigma\,x_i$ merupakan statistik yang sufficien maka.

$\sum_{i=1}^{n} d_j\,(x_i) = d_j\,(\sum_{i=1}^{n} x_i\,)$, ini adalah merupakan-

fungsi dari $\Sigma\,x_i$, karena setiap $\sum_{i=1}^{n} d\,(x_i)$ untuk $j=1,..,k$ merupakan

fungsi dari $\Sigma\,x_i$ maka himpunan $(\sum_{i=1}^{n} d_j\,(x_i)$ merupakanhimpunan dari statistik-statistik yang sufficien berserikat dan minimal.

Sekarang dibuktikan lengkapnya

Ambil $Z\,(\sum_{i=1}^{n} d_j\,(x_i)) = 0$ maka

$$E_\theta\left[Z\,(\sum_{i=1}^{n} d_j\,(x_i)\,)\right] \equiv 0 \text{ untuk setiap } \theta\in\bar{\theta} \text{ dengan}$$

$j = 1,2,\,........,k$

Jadi $\left[E_\theta(Z\,(\sum_{i=1}^{n} d_1(x_i),\,...,E_\theta(Z\,(\sum_{i=1}^{n} d_k(x_i)\,)\equiv\,0\right.$

Yang berarti juga bahwa

$$P_\theta \left[Z\left(\sum_{i=1}^{n} d_1 (x_i), \ldots, \sum_{i=1}^{n} d_k (x_i) \right) = 0 \right] \equiv 1$$

Untuk setiap $\theta \in \Theta$

Jadi : $\left(\sum_{i=1}^{n} d_1 (x_i), \ldots, \sum_{i=1}^{n} d_k (x_i) \right)$ adalah himpunan

statistik yang lengkap.

Jadi $\left(\sum_{i=1}^{n} d_1 (x_i), \ldots, \sum_{i=1}^{n} d_k (x_i) \right)$ adalah himpunan

yang minimal dari statistik-statistik yang sufficien , lengkap dan berserikat.

THEOREMA 8.

Misal $X_1, \ldots, X_n$ adalah n sampel random dari $f(x;\theta_1, \ldots, \theta_k)$ dan $\theta = (\theta_1, \ldots, \theta_k)$
Jika $S_1 = S_1(x_1, \ldots, x_n), \ldots, S_m = S_m(x_1, \ldots, x_n)$, adalah suatu himpunan dari statistik-statistik yang sufficien, lengkap dan berserikat dan jika terdapat suatu unbiase estimator dari $(\tau_1(\theta), \ldots, \tau_l(\theta))$ maka terdapat dengan tunggal suatu unbiase estimator dari $(\tau_1(\theta), \ldots, \tau_l(\theta))$, yaitu $T_1^* = t_1^*(S_1, \ldots, S_m), \ldots, T_r^* = t_r^*(S_1, \ldots, S_m)$.

B U K T I :

Ambil sebarang unbiase estimator dari $(\tau_1(\theta), \ldots, \tau_r(\theta))$ yaitu $(T_j', \ldots, T_r')$ dimana setiap T_j' berkorespondensi dengan setiap $\tau_j(\theta)$, untuk $j = 1,2, \ldots, r$ juga setiap T_j' merupakan fungsi dari $S_1, \ldots, S_m$ jadi $T_j' = C_j(S_1, \ldots, S_m)$.

Maka $E_\theta (T_j^* - T_j') \equiv 0$ untuk setiap $\theta \in \Theta$ dan T_j' adalah fungsi dari $S_1, \ldots, S_m$.

Karena $S_1, \ldots, S_m$ adalah himpunan dari statistik yang lengkap maka

$P_\theta\left[t_j^*(S_1, \ldots, S_m) - t_j(S_1, \ldots, S_m)\right] = 1$ untuk setiap $\theta \in \Theta$. Jadi hanya ada satu unbiase estimato r $\tau_j(\theta)$, dimana $j = 1,2,\ldots,r$, yang merupakanfungsi dari $S_1, \ldots, S_m$.

Misal T_j adalah sebarang unbiase estimator dari $\tau_j(\theta)$ untuk $j = 1,2,\ldots,r$; maka T_j^* haruslah sama dengan $E(T_j/S_1,\ldots,S_m)$, sebab $E(E(T_j/S_1, \ldots, S_m)) = E(T_j) = \tau_j(\theta)$.

Jadi T_j adalah unbiase estimator dari $\tau_j(\theta)$ yang merupakan fungsi dari $S_1, \ldots, S_m$ dimana $j=1.2,\ldots r$.

Jadi $T_j = T_j^*$ untuk $j = 1.2,\ldots,r$.

Maka terdapatlah dengan tunggal suatu unbiase estimator T_j dari $\tau_j(\theta)$ untuk $j = 1,2, \ldots r$.

Jadi terdapatlah dengan tunggal unbiase esti mator $(T_1^*, \ldots, T_r^*)$ dari $(\tau_1(\theta), \ldots, \tau_r(\theta))$, dimana setiap T_j^* merupakan fungsi dari $S_1, \ldots, S_m$.