BAB II

TEORI PENUNJANG

`Dalam Bab ini akan dibahas mengenai beberapa teori dasar statistik yang menunjang pembahasan selanjutnya. Antara lain terdiri dari : Metode Maksimum Likelihood Untuk Estimasi Parameter, Fungsi Distribusi Empiris dan Metode Estimasi Densitas Kernel .

2.1. Metode Maksimum Likelihood Untuk Estimasi Parameter

Sebuah metode yang sangat populer untuk memperoleh sebuah estimator parameter adalah metode maksimum likelihood. Pada dasarnya prosedur metode maksimum likelihood menguji apakah estimasi parameter yang tidak diketahui dari suatu sampel nilainya sudah memaksimalkan fungsi itu.

Misal sampel random ($x_1, x_2, \ldots , x_n$) berdistribusi independent dan identik (i.i.d), dan mempunyai fungsi densitas f (x) :

$$X \rightarrow f(\theta, x) \qquad (2.1.1)$$

dengan θ adalah satu atau lebih parameter yang akan dicari, yang nantinya akan menentukan distribusinya.

Maka prosedur metode maksimum likelihood dimulai dengan mendefinisikan fungsi maksimum likelihood sebagai :

$$L(\theta, x) = f(\theta, x_1) \cdot f(\theta, x_2) \cdot \ldots \cdot f(\theta, x_n)$$

$$= \prod_{1}^{n} f(\theta, x_i) \qquad (2.1.2)$$

L (θ; x) adalah fungsi dari θ dan x, dengan x adalah data yang telah ditentukan. Dalam kasus diskrit L (θ; x) adalah probabilitas dari observasi sampel, sedangkan dalam kasus kontinu L (θ; x) adalah pendekatan probabilitas dari sampel yang berada pada suatu interval kecil $(x, x + \Delta)$. Logaritma alam L (θ; x) ditunjukan dengan $l(\theta, x)$, maka logaritma alam dari L (θ; x) dapat dituliskan sebagai :

$$l(\theta, x) = \ln \prod_{1}^{n} f(\theta, x_i)$$

$$= \sum_{1}^{n} l(\theta, x_i) \qquad (2.1.3)$$

Notasi ini sering disebut log-likelihood, dan masing-masing nilai $l(\theta; x_i) = \ln f_\theta(x_i)$ disebut komponen log-likelihood.

Metode maksimum likelihood memilih nilai $\hat{\theta} = \theta$ untuk memaksimalkan $l(\theta, x)$. Dan fungsi likelihood sendiri dapat digunakan untuk menaksir nilai $\hat{\theta}$, yaitu dengan menurunkan log-likelihood terhadap θ :

$$l(\theta, x) = \sum_{1}^{n} l(\theta, x_i)$$

$$\hat{l}(\theta, x) = \frac{dl(\theta, x)}{d\theta} \qquad (2.1.4)$$

Dengan mengasumsikan bahwa nilai maximum likelihood berada di dalam ruang sampel, maka :

$$\hat{l}(\theta, x) = 0 \qquad (2.1.5)$$

Informasi tentang ketelitian θ didapat dari turunan kedua l (θ,x) terhadap θ, yaitu :

$$I(\theta) = \frac{d^2 l(\theta, x)}{d\theta^2} \qquad (2.1.6)$$

dengan I (θ) dievaluasi pada $\hat{\theta} = \theta$. Apabila I (θ) berharga negatif maka $\hat{\theta}$ adalah nilai maksimalnya.

Contoh :

Misalkan X variabel random yang mempunyai probabilitas exponensial, akan ditentukan estimasi untuk λ. Jika diberikan nilai pengamatan dalam sampel random dengan ukuran n dari populasi. Misal n = 5, dan kelima nilai pengamatan itu adalah : 30,4 ; 7,8 ; 1,4 ; 13,1 dan 67,3.

Penyelesaian:

Dari permasalahan tersebut maka fungsi likelihood dapat diperoleh dengan :

$$L(\lambda, x) = \prod_{1}^{n} f(\lambda, x_i)$$

$$= \prod_{1}^{n} (\lambda e^{-\lambda x_i})$$

$$= \prod_{1}^{5} (\lambda e^{-\lambda x_i})$$

$$= \lambda^5 e^{-120\lambda}$$

$$l(\lambda, x) = \ln L(\lambda, x)$$

$$= \ln(\lambda^5 e^{-120\lambda})$$

$$= 5 \ln \lambda - 120\lambda$$

Untuk memaksimalkan nilai $l(\lambda, x)$ maka $l(\lambda, x)$ diturunan terhadap λ, sehingga didapatkan :

$$\frac{dl(\lambda, x)}{d\lambda} = \frac{5}{\lambda} - 120$$

Nilai $\frac{dl(\lambda, x)}{d\lambda} = 0$ diberikan oleh:

$$\frac{5}{\hat{\lambda}} - 120 = 0$$

$$\frac{5}{\hat{\lambda}} = 120$$

$$\hat{\lambda} = 5/120 = 0{,}042$$

Karena nilai $\frac{d^2 l(\lambda,x)}{d\lambda^2} = -\frac{5}{\lambda^2} < 0$, maka nilai $\hat{\lambda}$ adalah nilai maksimum dari $l(\lambda,x)$. Jadi estimasi nilai λ adalah $\hat{\lambda} = 0{,}042$

2.2. Fungsi Distribusi Empiris

Masalah statistik inferensi seringkali berkaitan dengan pengestimasian beberapa aspek dari suatu distribusi probabilitas pada sampel random yang diambil dari distribusi probabilitas tersebut. Ketika distribusi probabilitas diketahui, solusi dari semua pertanyaan yang berkaitan dengan distribusi probabilitas adalah tidak lebih dari perhitungan aritmatik biasa. Tetapi jika disrtibusi probabilitasnya tidak diketahui, maka diperlukan statistik inferensi, yaitu teori statistik modern untuk menarik kesimpulan melalui sifat-sifat populasi dari suatu sampel random.

Misalkan (x_1, x_2,... , x_n) merupakan sampel random yang didistribusikan menurut distribusi F dan mengandung suatu parameter θ.

Maka parameter yang akan dicari tersebut ditulis dengan :

$$\theta = t(F) \tag{2.2.1}$$

θ adalah skalar yang merupakan fungsi dari F.

Karena distribusi probabilitas dari sampel random X_i tidak diketahui dan didistribusikan menurut distribusi F, maka distribusi probabilitasnya dapat diestimasi dengan menggunakan fungsi distribusi empiris, $\hat{F}$.

Estimasi F menggunakan distribusi empiris ini disebut prinsip "plug-in" atau prinsip penggantian, yaitu sebuah metode untuk mengestimasi parameter dari sampel. Estimasi " plug - in " dari parameter $\theta = t(F)$ adalah :

$$\hat{\theta} = t(\hat{F}) \tag{2.2.2}$$

Misalkan untuk mengestimasi beberapa aspek yang penting dari F seperti mean, modus, atau median dapat menggunakan aspek-aspek dari $\hat{F}$.

Fungsi distribusi empiris $\hat{F}$ didefinisikan sebagai distribusi diskret yang mempunyai nilai probabilitas 1 / n pada masing-masing nilai x_i, i = 1, 2, 3, ..., n. Dengan kata lain probabilitas empiris $\hat{F}$ untuk sampel x yang diambil dari himpunan A adalah :

$$\widehat{\text{Prob}}(A) = 1/n\,\{\#\{x_i \in A\}\} \tag{2.2.3}$$

Jadi pada dasarnya prinsip "plug-in" ini mengestimasi parameter θ = t (F) dari F dengan mengestimasi $\hat{\theta}$ = t ($\hat{F}$) dari fungsi distribusi empiris $\hat{F}$.

Contoh :

Pelemparan sebuah dadu sebanyak 100 kali, masing-masing mata dadu yang muncul adalah : (1,2,3,4,5,6) = (13,19,10,17,14,27) kali, akan ditentukan nilai probabilitas pada masing-masing data.

Data selengkapnya adalah pada tabel 2.2.1 berikut ini :

6	3	2	4	6	6	6	5	3	6	2	2	6	2	3	1	5	1	6	6
4	1	5	3	6	6	4	1	4	2	5	6	6	5	5	3	6	2	6	6
6	5	4	1	2	4	2	5	2	2	2	3	6	3	6	1	1	5	1	4
6	2	6	1	2	5	3	5	1	4	5	6	4	3	1	4	4	2	1	4
6	2	4	2	2	1	5	4	2	6	6	4	6	6	4	5	4	3	2	6

Tabel 2.2.1.Data pelemparan sebuah dadu 100 kali.

Tabel 2.2.1 diatas menunjukan sampel random dari n = 100 pemunculan mata dadu pada pelemparan sebuah mata dadu sebanyak 100 kali. X_1= 6, X_2= 3, ... , X_{100}= 6. Distribusi empiris $\hat{F}$ meletakkan probabilitas 1/100 pada masing-masing nilai x_i. Dalam kasus seperti ini dimana ada nilai yang diulang, maka nilai probabilitas empiris dari $\hat{F}$ dapat ditulis sebagai vektor dari frekuensi observasi, $\hat{f}_k$, k = 1,2,...,6.

$$\text{Prob}\,(\,x = k\,) = \hat{f}_k$$

Sehingga nilai $\hat{f}_k$ adalah :

$$\hat{f}_k = 1/n\ \{\#\ \{x_i \in A\ \}\}$$

Untuk kasus ini maka frekuensi observasinya adalah:

$$\hat{f}_1 = 13/100 = 0{,}13$$

$$\hat{f}_2 = 19/100 = 0{,}19$$

.

.

.

$$\hat{f}_6 = 27/100 = 0{,}27$$

Sehingga nilai $\hat{F}$ diberikan oleh :

$$\hat{F} = (\hat{f}_1, \hat{f}_2, \hat{f}_3, \hat{f}_4, \hat{f}_5, \hat{f}_6)$$

$$= (\,0{,}13;\ 0{,}19;\ 0{,}10;\ 0{,}17;\ 0{,}14;\ 0{,}27\,)$$

2.3. Metode Estimasi Densitas Kernel

Dalam menarik kesimpulan dari suatu data observasi, teori probabilitas memegang peranan yang sangat penting. Dengan diketahui nilai probabilitasnya, maka dapat diperoleh informasi dari data tersebut. Dalam hal ini fungsi densitas berguna dalam perhitungan nilai probabilitas.

Sehingga pengetahuan tentang fungsi densitas sangat membantu dalam banyak hal. Salah satu contoh adalah untuk mengetahui bagaimana terjadinya suatu observasi dari sekumpulan observasi X, misalnya dalam interval X = (a,b) , maka :

$$P\ (a < x < b) = \int_{a}^{b} f(x)\,dx$$

Jika nilai probabilitas ini tinggi maka dapat diprediksikan bahwa observasi-observasi tersebut terakumulasi dalam interval itu, sehingga fungsi densitas menginformasikan bahwa observasi-observasi lebih sering terjadi dalam interval tersebut.

Pada umumnya fungsi densitas tidak dapat diketahui secara langsung, sedangkan yang ada hanyalah sekumpulan data dari hasil observasi. Sehingga belum dapat diperoleh informasi dari data-data itu. Untuk mengatasi permasalahan ini perlu dilakukan suatu pendekatan. Ada dua macam pendekatan untuk mengatasi masalah ini. Pertama pendekatan parametrik yaitu suatu pendekatan yang digunakan hanya untuk mengestimasi parameter yang belum diketahui saja sedangkan bentuk fungsi densitasnya sudah diketahui. Kedua adalah pendekatan non parametrik yaitu pendekatan untuk mengestimasi fungsi densitas yang belum diketahui. Salah satu pendekatan yang kedua ini adalah metode estimasi densitas kernel.

Metode estimasi densitas kernel adalah metode pendekatan terhadap fungsi densitas yang juga merupakan metode penghalusan diagram. Metode ini ditentukan oleh suatu parameter yang sangat berpengaruh, yaitu bandwidth dan jenis fungsi kernel itu sendiri.

Bandwidth adalah parameter penghalus yang mengatur derajat kehalusan untuk penghalusan kernel, biasa dilambangkan dengan h. Ada beberapa kemungkinan yang merupakan pengaruh pemilihan nilai bandwidth terhadap grafik penghalusan fungsi densitas kernel , yaitu semakin besar nilai h, maka akan menghasilkan estimasi densitas yang lebih halus, dan sebaliknya semakin kecil nilai h akan menghasilkan estimasi densitas yang semakin kasar atau semakin tidak halus.

Fungsi kernel K secara umum didefinisikan sebagai :

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right) \tag{2.3.1}$$

Jika pengamatan data dilakukan sebanyak i, dengan i = 1,...,n, maka estimator densitas kernelnya menjadi:

$$\begin{aligned} \hat{F}(x) &= \frac{1}{h} \sum_{i=1}^{n} K_h(x - x_i) \\ &= \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \end{aligned} \tag{2.3.2}$$

Fungsi densitas kernel ini dikenal ada 7 macam, yaitu:

Fungsi Kernel	K (u)
Uniform	½ I I (| u | ≤ 1)
Triangel	(1 –|u|) I (| u | ≤ 1)
Epanechnikov	¾ (1 –u^2) I (| u | ≤ 1)
Quartic	15/16 (1 –u^2)^2 I (| u | ≤ 1)
Triweight	35/32 (1 –u^2)^3 I (| u | ≤ 1)
Gausian	$\frac{1}{\sqrt{2\pi}}$ exp (-1/2 u^2)
Cosinus	π/4 cos (π/2 u) I (| u | ≤ 1)