II. LANDASAN TEORI

Dalam proses penelitian penduga parameter dari suatu distribusi diperlukan beberapa konsep dan teori yang mendukung dari ilmu statistika. Berikut ini akan dijelaskan beberapa konsep dan teori yang berkaitan dengan penduga parameter distribusi GB2 menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan bantuan *Software* R.

2.1 Fungsi Gamma

Fungsi Gamma merupakan salah satu dari beberapa fungsi khusus di dalam matematika. Fungsi Gamma merupakan perluasan dari transformasi laplace yang sangat penting dalam matematika dan sebagai dasar dalam perkembangan teknologi dan sains modern.

Definisi 2.1

Fungsi Gamma yang dinotasikan dengan $\Gamma(p)$ yaitu:

$$\Gamma(p) = \int_{0}^{\infty} e^{-y}\, y^{p-1}\, dy, \;\; p > 0$$

(Hogg n Craig, 1995)

Pada sub-bab selanjutnya akan dijelaskan tentang fungsi digamma yang akan digunakan uantuk menyelesaikan persamaan yang nantinya akan dijumpai pada saat melakukan turunan (derivatif) pada saat mencari penduga suatu parameter.

2.2 Fungsi Digamma

Fungsi Digamma merupakan hasil turunan (derivatif) pertama dari fungsi Gamma.

Definisi 2.2

Fungsi Digamma didefinisikan sebagai berikut:

$$\psi(p) = \frac{\partial\,(\ln \Gamma(p))}{\partial p} = \frac{\Gamma'(p)}{\Gamma(p)}, \qquad p > 0$$

(Abramowitz, and Stegun,1972)

Sub-bab berikutnya akan dijelaskan tentang fungsi polygamma yang merupakan hasil turunan (derivatif) pertama dari fungsi digamma

2.3 Fungsi Polygamma

Fungsi Polygamma merupakan fungsi yang diperoleh dari turunan ke-n fungsi Gamma.

Definisi 2.3

Fungsi Polygamma didefinisikan sebagai berikut:

$$\psi^{(n)}(p) = \frac{\partial^n(\psi)(p)}{\partial p^n} = \frac{\partial^{(n+1)}(\ln\Gamma(p))}{\partial p^{(n+1)}}, p > 0$$

(Abramowitz, and Stegun,1972)

Selain fungsi Gamma, Digamma maupun Polygamma pada penelitian ini juga menggunakan fungsi Beta yang digunakan untuk menyelesaikan integral khusus.

2.4 Fungsi Beta

Fungsi Beta juga merupakam salah satu fungsi khusus yang ada di dalam matematika. Fungsi Beta digunakan untuk mengevaluasi integral tentu.

Definisi 2.4

Fungsi Beta adalah suatu fungsi bernilai real dengan dua peubah, didefinisikan oleh suatu bentuk integral, yaitu:

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx,\ \ p > 0, q > 0$$

$$B(p,q) = \int_0^\infty \frac{x^{p-1}}{(1+x)^{p+q}}\,dx,\ \ p > 0, q > 0$$

(Ross, 2010)

Sifat: $B(p,q) = B(q,p)$

Bukti

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\ dx$$

$$= \int_0^1 (1-y)^{p-1} y^{q-1}\ dy$$

$$= \int_0^1 y^{q-1}\ (1-y)^{p-1}\ dy$$

$$= B(q,p) \qquad \blacksquare$$

Fungsi Beta dapat juga dinyatakan dengan fungsi Gamma yaitu:

Dengan $B(p,q) = \frac{\Gamma(p).\Gamma(\mathrm{q})}{\Gamma(\mathrm{p+q})}$

(Ross, 2010)

2.5 Distribusi *Generalized Beta* II GB2)

Definisi 2.5

Suatu variabel acak dikatakan memiliki distribusi *Generalized Beta* II (GB2) dengan parameter (a,b,p,q) jika fungsi kepekatan peluangnya adalah:

$$f(x) = \frac{a\ (x_i)^{ap-1}}{b^{ap}\ B(p,q)\left(1 + \left(\frac{x_i}{b}\right)^{a}\right)^{p+q}},\quad x > 0$$

dengan:

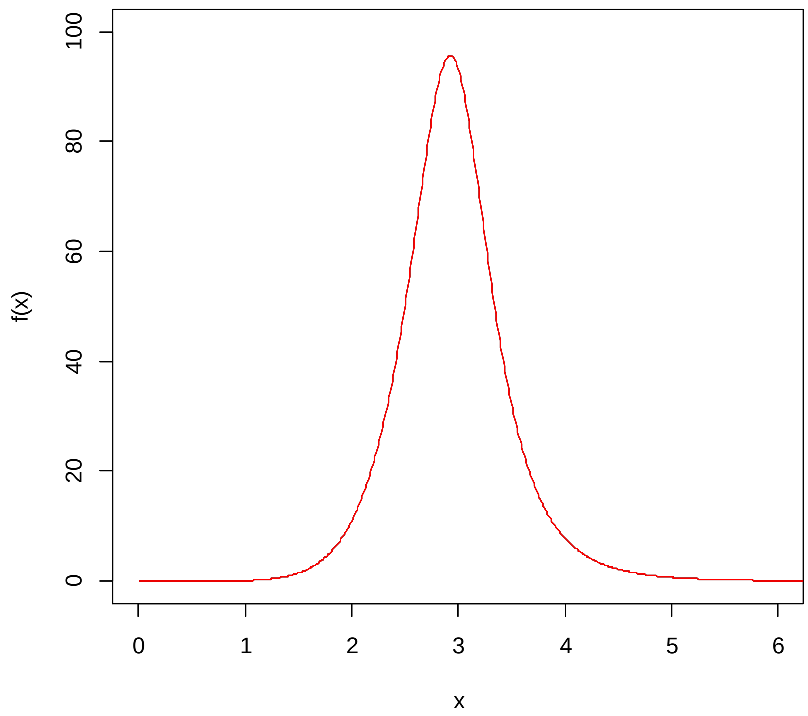
a, b, p, q adalah bilangan positif

$B(p, q)$ adalah fungsi beta

b adalah parameter skala

a, p, q adalah parameter bentuk

(Kleiber, and Kotz, 2003)



Gambar 1. Grafik Fungsi Kepekatan Peluang *Generalized Beta* II (GB2) dengan $a = 2, b = 3, p = 4, q = 5$

Statistik inferensia terdiri dari pengujian hipotesis dan pendugaan. Pada penelitian ini akan dilakukan pendugaan parameter. Pendugaan parameter dilakukan untuk menduga ukuran dari suatu populasi yang belum diketahui. Definisi pendugaan parameter akan dijelaskan pada Subbab 2.6.

2.6 Penduga Parameter

Dalam statistik inferensia dibutuhkan pemahaman mengenai kaidah-kaidah pengambilan kesimpulan tentang suatu parameter populasi berdasarkan karakteristik sampel. Hal ini membangun apa yang disebut dengan pendugaan titik dari fungsi kepekatan peluang parameter yang tidak diketahui.

Definisi 2.6

Misal suatu peubah acak X memiliki fungsi kepekatan peluang yang bergantung pada suatu parameter tak diketahui θ dengan sembarang nilai dari suatu himpunan ruang parameter Ω, maka dinotasikan dengan

$f(x; \theta), \theta \in \Omega$.

Definisi 2.7

Misal $X_1, X_2, ..., X_n$ berdistribusi bebas stokastik identik dengan fungsi kepekatan peluang $f(x; \theta), \theta \in \Omega$. Suatu statistik $U(X_1, X_2, ..., X_n) = U(X)$ yang digunakan untuk menduga $g(\theta)$ disebut sebagai penduga bagi $g(\theta)$.

Untuk menduga parameter dari suatu distribusi dapat dilakukan dengan beberapa metode. Dalam penelitian ini pendugaan parameter distribusi GB2 akan dilakukan dengan menggunakan metode MLE. Definisi metode MLE akan dijelaskan pada Subbab 2.7.

2.7 Metode Pendugaan *Maximum Likelihood Estimation* (MLE)

Definisi 2.8

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak berukuran n yang saling bebas stokastik identik dari suatu distribusi yang mempunyai fungsi kepekatan peluang $f(x; \theta)$, $\theta \in \Omega$. Fungsi kepekatan peluang bersama dari $X_1, X_2, \ldots, X_n$ adalah $f(x_1; \theta)\, f(x_2; \theta) \ldots f(x_n; \theta)$ yang merupakan fungsi kemungkinan (*Likelihood Function*).

Untuk $x_1, x_2, \ldots, x_n$ tetap, fungsi kemungkinan merupakan fungsi dari θ dan dilambangkan dengan $L(\theta)$ dan dinotasikan sebagai berikut:

$$
\begin{aligned}
L(\theta) &= f(\bar{x}\ ;\ \theta) \\
&= f(x_1, x_2, \ldots, x_n\ ;\ \theta) \\
&= f(x_1; \theta)\, f(x_2; \theta) \ldots f(x_n; \theta),\ \theta \in \Omega \\
&= \prod_{i=1}^{n} f\ (x_i;\ \theta)
\end{aligned}
$$

Definisi 2.9

$L(\theta) = f(x_1, x_2, ..., x_n\ ;\ \theta)$, $\theta \in \Omega$ merupakan fungsi kepekatan peluang dari $x_1, x_2, ..., x_n$. Untuk hasil pengamatan $x_1, x_2, ..., x_n$, nilai $\hat{\theta}$ berada dalam Ω ($\hat{\theta} \in \Omega$), dimana $L(\theta)$ maksimum yang disebut sebagai *Maximum Likelihood Estimation* (MLE) dari θ. Jadi, $\hat{\theta}$ merupakan penduga dari θ.

Jika $f(x_1, x_2, ..., x_n\ ;\ \hat{\theta}) = \max f(x_1, x_2, ..., x_n\ ;\ \theta)$, $\theta \in \Omega$ maka untuk memaksimumkan $L(\theta)$ terhadap parameternya. Biasanya mencari turunan dari $L(\theta)$ terhadap parameternya relatif sulit, sehingga dalam penyelesainnya dapat diatasi dengan menggunakan logaritma atau fungsi ln dari $L(\theta)$ yaitu: ln $L(\theta) = \sum_{i=1}^{n} \ln f(x_i\ ;\ \theta)$. Untuk memaksimumkan ln $L(\theta)$ adalah dengan mencari turunan dari ln $L(\theta)$ terhadap parameternya, kemudian hasil turunannya dibuat sama dengan nol.

$$\frac{\partial \ln \mathrm{L}(\theta)}{\partial \theta} = 0 \quad \text{(Hogg and Craig, 1995)}$$

Dalam penduga parameter dari suatu distribusi ada penduga parameter yang tidak dapat diselesaikan secara analitik, sehingga perlu diselesikan dengan cara numerik. Salah satu cara yang digunakan adalah dengan teknik iteratif yaitu Metode Newton Raphson. Metode Newton Raphson sering digunakan karena metode ini lebih sederhana dn mempunyai konvergensi yang cepat. Subbab 2.8 akan menjelaskan tentang definisi Metode Newton Raphson.

2.8 Metode Newton Raphson

Apabila proses pendugaan parameter didapat persamaan akhir yang non linear maka tidak mudah memperoleh pendugaan parameter tersebut, sehingga diperlukan suatu metode numerik untuk memecahkan persamaan non linear tersebut. Salah satu metode yang digunakan untuk memecahkan sistem persamaan non linear adalah Metode Newton Raphson. Metode Newton Raphson adalah metode untuk menyelesaikan persamaan non linear secara iteratif. Jika θ_0 merupakan nilai awal (inisialisasi) dari θ atau θ_0 merupakan nilai ke 1 dari θ, maka dapat dimisalkan $\theta_0 = \theta_i$ dan $\theta_1 = \theta_{i+1}$ dengan i awal 0.

Metode ini dapat diperluas untuk menyelesaikan sistem persamaan dengan lebih dari satu parameter. Misal θ_1, θ_2, . . . θ_p ... maka iterasinya sebagai berikut:

$\theta_{i+1} = \theta_i - [H^{-1} g]$

Dengan $\hat{\theta}_{i+1} = \begin{bmatrix} \hat{\theta}_{i+1} \\ \vdots \\ \hat{\theta}_{p+1} \end{bmatrix}$ dan $\hat{\theta}_i = \begin{bmatrix} \hat{\theta}_{1i} \\ \vdots \\ \hat{\theta}_{pi} \end{bmatrix}$

Vektor gradien atau vektor turunan pertama terhadap parameternya dan lambangnya dengan $g(x)$ yaitu:

$$g(\theta) = \frac{\partial \, ln \, \mathrm{L} \, (\theta)}{\partial x} = \begin{bmatrix} \frac{\partial \, ln \, \mathrm{L} \, (\theta)}{\partial \theta_1} \\ \frac{\partial \, ln \, \mathrm{L} \, (\theta)}{\partial \theta_2} \\ \vdots \\ \frac{\partial \, ln \, \mathrm{L} \, (\theta)}{\partial \theta_n} \end{bmatrix}$$

Matriks Hessian atau matriks turunan kedua dari fungsi logaritma natural terhadap parameter a, b, p dan q dilambangkan dengan $\mathrm{H}(\theta)$ yaitu:

$$\mathrm{H}(\theta) = \frac{\partial^2 \, ln \, \mathrm{L}\,(\theta)}{\partial\theta\partial\theta'}$$

$$= \begin{bmatrix} \frac{\partial^2 \, ln \, \mathrm{L}\,(\theta)}{\partial\theta_1^2} & \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_1 \, \partial\theta_2} & \cdots & \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_1 \, \partial\theta_n} \\ \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_2 \, \partial\theta_1} & \frac{\partial^2 \, ln \, \mathrm{L}\,(\theta)}{\partial\theta_2^2} & \cdots & \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_2 \, \partial\theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_n \, \partial\theta_1} & \frac{\partial^2 \ln \mathrm{L}\,(\theta)}{\partial\theta_n \, \partial\theta_2} & \cdots & \frac{\partial^2 \, ln \, \mathrm{L}\,(\theta)}{\partial\theta_n^2} \end{bmatrix}$$

(Seber and Wild, 2003)

Untuk memudahkan melakukan proses iterasi dengan metode Newton Raphson akan digunakan *Software* R. Penjelasan mengenai *Software* R akan dijelaskan pada Subbab 2.9.

2.9 *Software* R

Software R adalah perangkat lunak bebas untuk komputasi statistik dan grafik. *Software* R merupakan proyek GNU *General Public License Free Software Fundation* yang mirip bahasa S yang dikembangkan di Bell Laboratories oleh John Chambers dan rekannya. *Software* R menyediakan berbagai statistik seperti linear dan non linear *modeling*, pengujian analisis klasik, analisis *time-series*, klasifikasi dan lainnya, serta perangkat lunak yang umumnya digunakan untuk manipulasi data, perhitungan, dan tampilan grafik yang mencakup antara lain:

a. Penanganan data yang efektif dan penyimpangan data
b. Rangkaian operator untuk perhitungan array dalam matrik tertentu
c. Fasilitas grafik untuk analisis data dan menampilkannya, baik pada layar maupun *hardcopy*
d. Bahasa pemrograman yang sederhana, berkembang dengan baik dan efektif