VARIANSI: Journal of Statistics and Its Application on Teaching and Research **ISSN 2684-7590 (Online)**
Vol. 3 No. 3 (2021), 115-125

DOI: 10.35580/variansiunm26374

PERBANDINGAN METODE MOMEN, *MAXIMUM LIKELIHOOD* DAN *BAYES* DALAM MENDUGA PARAMETER DISTRIBUSI PARETO

A. Nurul Amalia[1], Muhammad Arif Tiro[2], Aswi[3]

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: Pareto distribution, Moment Method, *Maximum Likelihood IMethod,* Bayesian Method.

Abstract:

This study examines the estimation of Pareto distribution parameters using three different methods, namely the Moment, *Maximum Likelihoo*d, and Bayesian methods. The Pareto distribution is a continuous distribution with parameters $k > 0$ and $\alpha > 0$. These two parameters are estimated by using three distinct parameter estimation methods. The goodness of fit measure used in choosing the best estimation method is the Mean Square Error (MSE) value. The smallest MSE is the best method. A simulation study is carried out as well as the case study of the data on the number of Gross National Income (GNI) per capita in Southeast Asian countries in 2019. The estimation and simulation results indicate that the best estimation method in estimating the parameters of the Pareto distribution is the *Maximum Likelihood* in terms of MSE value.

1. Pendahuluan

Tiro (2015) mengemukakan bahwa keseluruhan aspek tertentu dari ciri, fenomena, atau konsep yang menjadi pusat perhatian dalam suatu penelitian disebut populasi. Berbagai macam ukuran seperti rataan, median, modus, variansi, dan proporsi yang mengukur populasi dinamakan parameter. Namun, mengukur populasi sulit untuk dilakukan karena tidak memungkinkan untuk mengukur semua elemennya, sehingga dibutuhkan pendugaan populasi dengan menggunakan sampel untuk memudahkan pengukuran itu. Setiap elemen populasi yang diambil sebagai sampel memiliki peluang-peluang tertentu. Hal ini menunjukkan bahwa elemen dalam populasi itu menyebar membentuk distribusi tertentu. Dalam statistika, ada banyak distribusi yang menggambarkan pola-pola penyebaran data. Salah satu distribusi yang sering digunakan adalah distribusi Pareto $f(x; k, \alpha)$. Distribusi Pareto merupakan salah satu distribusi kontinu dengan parameter $k > 0$ dan parameter $\alpha > 0$ (Kleiber & Kotz, 2003). Distribusi Pareto banyak digunakan dalam bidang sosial, bisnis, ekonomi, asuransi, aktuaria, hidrologi dan lain-lain.

Dalam memudahkan pengukuran populasi yang berdistribusi Pareto, sangat penting melakukan pendugaan parameter untuk distribusi ini. Penduga parameter dapat dikatakan sebagai parameter yang baik apabila memiliki ciri tidak bias (nilai penduga sama dengan nilai yang diduganya), efisien (penduga memiliki variansi yang kecil), dan konsisten (nilai penduganya tidak berbanding terbalik dengan bertambahnya jumlah sampel) (Tiro, 2015). Pendugaan parameter dilakukan dengan menggunakan metode-metode tertentu.

* Corresponding author.
e-mail: andinurul.amalia28@gmail.com

Penelitian tentang penduga parameter dengan menggunakan metode Momen, *Maximum Likelihood* dan Bayes pernah dilakukan oleh Quandt (1964) yang meneliti tentang *Old and New Methods of Estimation and the Pareto Distribution*. Yanuar (2020) meneliti tentang pendugaan parameter distribusi Pareto dengan menggunakan Metode Bayes dengan prior konjugat dan nonkonjugat. Warsono (2019) meneliti tentang perbandingan beberapa metode pendugaan parameter distribusi Pareto, dan penelitian dilakukan oleh Setiya (2016) mengenai pendugaan skala dari distribusi Pareto menggunakan Bayes.

Pada penelitian ini, metode yang digunakan untuk menduga parameter distribusi Pareto adalah Metode Momen, *Maximum Likelihood Estimation* (MLE), dan Bayes. Ketiga metode ini kemudian dibandingkan untuk menemukan metode pendugaan yang paling baik untuk menduga parameter distribusi Pareto. Pemilihan metode terbaik diantara ketiga metode ini didasarkan pada perbandingan besarnya kuadrat tengah galat. Semakin kecil kuadrat tengah galat yang dihasilkan, semakin baik pula metode pendugaan yang digunakan karena dapat menggambarkan bahwa sampel yang digunakan memiliki perbedaan yang kecil dengan populasi.

2. Tinjauan Pustaka

Dalam menganalisis data penelitian ini, penulis menyusun langkah-langkah sebagai berikut :

2.1. Distribusi Pareto

Distribusi merupakan konsep yang menjadi dasar pengembangan statistika inferensial, khususnya dalam penaksiran parameter dan hipotesis (Tiro, 2015). Distribusi Pareto pertama kali ditemukan oleh Vilfredo Damaso Pareto pada tahun 1897, seorang sosiolog, ekonom, dan filsuf di Italia. Pareto mengamati bahwa dalam banyak populasi distribusi pendapatan adalah salah satu di mana jumlah individu yang pendapatannya melebihi tingkat tertentu. Fungsi distribusi Pareto klasik didefinisikan sebagai berikut,

$$F(x) = 1 - \left(\frac{k}{x}\right)^{\alpha} ;\ k \le x < \infty ;\ \alpha, k > 0$$

dan fungsi kepadatan peluangnya adalah

$$f(x) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}} ;\ k \le x < \infty ;\ \alpha, k > 0$$

dimana α adalah parameter bentuk dan k adalah parameter skala.

2.2. Metode Momen

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebuah populasi X dengan fungsi kepadatan peluang $f(x; \theta_1, \theta_2, \ldots, \theta_m)$ dimana $\theta_1, \theta_2, \ldots, \theta_m$ adalah m parameter yang tidak diketahui. Misalkan

$$E(X^m) = \int_{-\infty}^{\infty} x^k f(x;\ \theta_1, \theta_2, \ldots, \theta_m) dx$$

menjadi momen populasi ke-p. Lebih lanjut, misalkan

$$M_m = \frac{1}{n}\sum_{i=1}^{n} X_i^m$$

menjadi momen sampel ke-m.

Dalam metode momen, kita mencari penduga untuk parameter $f(x; \theta_1, \theta_2, \ldots, \theta_m)$

$$\begin{gathered} E(X) = M_1 \\ E(X^2) = M_2 \\ E(X^3) = M_3 \\ \vdots \\ E(X^m) = M_m \end{gathered}$$

Metode momen merupakan salah satu metode klasik untuk pendugaan parameter. Metode momen pertama kali ditemukan oleh ahli statistik Inggris Karl Pearson pada tahun 1902 (Sahoo, 2008).

2.3 *Maximum Likelihood*

Metode ini pertama kali digunakan oleh Sir Ronald Fisher pada tahun 1912 untuk menemukan penduga dari parameter yang tidak diketahui. Namun, metode ini berasal dari karya Gauss dan Bernoulli (Sahoo, 2008). Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebuah populasi X dengan fungsi kepadatan peluang $f(x;\theta)$, dimana θ merupakan parameter yang tidak diketahui. Fungsi likelihood $L(\theta)$, adalah distribusi sampel. Itu adalah

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

Definisi ini mengatakan bahwa fungsi *likelihood* dari sampel acak $X_1, X_2, \ldots, X_n$ adalah kepadatan gabungan dari peubah acak $X_1, X_2, \ldots, X_n$.

2.4 Metode Bayes

2.4.1. Distribusi Prior

Misalkan ada beberapa informasi tambahan tentang θ dan informasi tersebut dapat diringkas dalam bentuk distribusi peluang untuk θ, katakanlah $f(\theta)$. Distribusi peluang ini sering disebut distribusi prior untuk θ, dan misalkan rerata dari prior μ_0 dan variansinya adalah σ_0^2. Probabilitas yang terkait dengan distribusi sebelumya sering disebut probabilitas subyektif karena biasanya mencerminkan tingkat kepercayaan analis mengenai nilai sebenarnya dari θ.

Distribusi prior yang akan digunakan adalah distribusi *Uniform* Kontinu. Peubah acak X dikatakan mempunyai sebaran seragam (*uniform*) kontinu dengan parameter a dan b ditulis dengan symbol $X \sim Unif(a, b)$, jika dan hanya jika fungsi kepadatan peluangnya adalah

$$f(x) = \begin{cases} \dfrac{1}{b-a} & ; a < x < b \\ 0 & ; x \text{ yang lain} \end{cases}$$

2.4.2 Distribusi Posterior

Pendekatan Bayesian untuk pendugaan menggunakan distribusi sebelumnya untuk θ, $f(\theta)$, dan distribusi peluang gabungan dari sampel, katakanlah, $f(x_1, x_2, \ldots, x_n | \theta)$ untuk mencari distribusi posterior untuk θ, katakanlah $f(\theta | x_1, x_2, \ldots, x_n)$. Distribusi posterior ini berisi informasi dari sampel dan distribusi sebelumnya untuk θ. Secara konseptual, mudah menemukan distribusi posterior. Distribusi peluang gabungan dari sampel $X_1, X_2, \ldots, X_n$ dan parameter θ adalah

$$f(x_1, x_2, \ldots, x_n, \theta) = f(x_1, x_2, \ldots, x_n | \theta)\, f(\theta)$$

dan fungsi distribusi marginal dari $X_1, X_2, \ldots, X_n$ adalah

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} \sum_{0} f(x_1, x_2, \ldots, x_n, \theta) & \theta \text{ diskrit} \\ \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n, \theta)\, d\theta & \theta \text{ kontinu} \end{cases}$$

Oleh karena itu, distribusi yang diinginkan adalah

$$f(\theta | x_1, x_2, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n, \theta)}{f(x_1, x_2, \ldots, x_n)}$$

Definisikan penduga Bayes dari θ sebagai nilai $\tilde{\theta}$ yang sesuai dengan rerata dari distribusi posterior $f(\theta | x_1, x_2, \ldots, x_n)$. Terkadang rerata dari distribusi posterior θ dapat ditentukan dengan mudah. Sebagai fungsi dari θ, $f(\theta | x_1, x_2, \ldots, x_n)$ adalah fungsi kepadatan peluang dan $x_1, x_2, \ldots, x_n$ hanyalah konstanta. Karena θ masuk ke dalam $f(\theta | x_1, x_2, \ldots, x_n)$ hanya melalui $f(x_1, x_2, \ldots, x_n, \theta)$ jika $f(x_1, x_2, \ldots, x_n, \theta)$ karena fungsi θ dikenali sebagai fungsi peluang yang terkenal, rerata posterior dari θ dapat disimpulkan dari distribusi terkenal tanpa integrasi atau bahkan perhitungan $f(x_1, x_2, \ldots, x_n)$ (Montgomery & Runger, 2014).

2.5. Evaluasi Sifat Penduga

2.5.1 Tidak Bias

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak berukuran n dari sebuah populasi dengan fungsi kepadatan peluang $f(x;\theta)$. Sebuah penduga $\hat{\theta}$ bagi θ adalah sebuah fungsi peubah acak $X_1, X_2, \ldots, X_n$ yang bebas dari parameter θ. Penduga parameter mungkin tidak sama dengan nilai sebenarnya dari parameter untuk setiap realisasi sampel $X_1, X_2, \ldots, X_n$, tapi jika tidak bias maka rata-ratanya akan sama dengan parameter.

2.5.2 Variansinya Minimum

Penduga dengan variansi minimum memiliki variansi terkecil di antara semua penduga untuk parameter yang sama. Misalnya $\hat{\theta}$ sebagai penduga parameter θ memiliki variansi terkecil di antara semua penduga θ, maka $\hat{\theta}$ disebut penduga bervariansi minimum.

2.5.3 Konsisten

Misalkan θ adalah konstanta, $\varepsilon > 0$, dan n adalah indeks dari peubah acak x_n. Jika $\lim_{n\to\infty} P[|x_n - \theta| > \varepsilon] = 0$ untuk $\varepsilon > 0$ dapat dikatakan bahwa x_n adalah peubah konsisten bagi θ dengan notasi *plim* $x_n = \theta$.

3. Metode Penelitian

Dalam menganalisis data penelitian ini, penulis menyusun langkah-langkah sebagai berikut :

3.1. Jenis Penelitian

Jenis penelitian yang digunakan adalah studi pustaka, yaitu mengkaji pustaka-pustaka yang berkaitan dengan penelitian ini. Hal ini berkenaan dengan kegiatan menyeleksi pustaka yang terkait dengan permasalahan, sehingga mampu menyelesaikan permasalahan secara tuntas.

3.2. Sumber Data

Data yang digunakan adalah data bangkitan berdistribusi Pareto di dalam *software* R dan data *Gross National Income* (GNI) Negara Asia Tenggara pada bulan Juli 2019 yang diambil dari *World Bank*.

3.3. Teknik Analisis

Adapun langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Melakukan pendugaan dengan menggunakan metode Momen, *Maximum Likelihood*, dan Bayes.
2. Memeriksa sifat konsisten dan tak bias dari penduga yang diperoleh.
3. Melakukan simulasi data,
   1) Membangkitkan data berdistribusi Pareto sebanyak n = 10, n = 40, n = 100, n = 200 dengan kombinasi parameter α = 1, α = 3, α = 5 dan k = 1, k = 2, dan k = 3;
   2) Melakukan perhitungan nilai penduga parameter dengan bantuan software R;
   3) Melakukan pengulangan langkah 1 untuk masing-masing kombinasi n, α, dan k sebanyak 1000 kali dan langkah ke-2 untuk setiap iterasi;
   4) Menghitung nilai MSE dari pengulangan yang dilakukan sebagai nilai dari sifat variansi yang minimum.
4. Membandingkan MSE dari setiap metode pendugaan.
5. Menentukan metode pendugaan terbaik

4. Hasil dan Pembahasan

4.1. Nilai Harapan dan Variansi Distribusi Pareto

Diketahui bahwa fungsi kepadatan peluang dari distribusi Pareto adalah

$$f(x) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}} \; ; \; k \le x < \infty \; ; \; \alpha, k > 0$$

Maka diperoleh nilai harapan untuk distribusi Pareto adalah sebagai berikut.

$$E(X) = \int_{-\infty}^{\infty} x \, f(x) dx$$

$$= \int_{-\infty}^{k} xf(x)dx + \int_{k}^{\infty} xf(x)\,dx$$
$$= \int_{k}^{\infty} \frac{\alpha k^{\alpha}}{x^{\alpha+1}} x\,dx$$
$$= \frac{\alpha k}{\alpha - 1}, \alpha > 1$$

Kemudian untuk memperoleh variansinya, maka dicari $E(X^2)$ dari distribusi Pareto,

$$E(X^2) = \int_{-\infty}^{k} x^2 f(x)dx + \int_{k}^{\infty} x^2 f(x)dx$$
$$= \int_{k}^{\infty} x^2 \frac{\alpha k^{\alpha}}{x^{\alpha+1}} dx$$
$$= \frac{\alpha k^2}{\alpha - 2}, \alpha > 2$$

Jadi, variansi dari distribusi Pareto diberikan oleh

$$Var(X) = E(X^2) - (E(X))^2$$
$$= \frac{\alpha k^2}{\alpha - 2} - \left(\frac{\alpha k}{\alpha - 1}\right)^2$$
$$= \frac{k^2 \alpha}{(\alpha - 1)^2(\alpha - 2)}, \alpha > 2$$

4.2. Pendugaan Parameter Distribusi Pareto

4.2.1. Metode Momen

Parameter α dapat diduga dengan $E(X) = M_1$, sehingga diperoleh

$$E(X) = \bar{x}$$
$$\hat{\alpha} = \frac{\bar{x}}{(\bar{x} - \hat{k})}$$

Namun, penduga tersebut bergantung pada penduga k yang belum diketahui. Sehingga pendugaan ini tidak bisa digunakan sebagai pendugaan parmeter distribusi Pareto. Oleh karena itu, digunakan fungsi nilai sampel minimum jika disimbolkan dengan x_1 mengingat bahwa k merupakan batas bawah dari x, sehingga diperoleh penduga parameter sebagai berikut.

$$x_1 = \frac{\alpha n k}{\alpha n - 1}$$
$$\hat{k} = \frac{(\hat{\alpha} n - 1)x_1}{\hat{\alpha} n}$$
$$\hat{\alpha} = \frac{n\bar{x} - x_1}{n(\bar{x} - x_1)}$$

4.2.2. *Maximum Likelihood Estimation*

Diberikan fungsi *likelihood* distribusi Pareto sebagai berikut.

$$L(\theta) = f(x_1;\theta)f(x_2;\theta)\ldots f(x_n;\theta)$$
$$= \prod_{i=1}^{n} f(x_i;\theta)$$
$$L(k,\alpha) = \prod_{i=1}^{n} \frac{\alpha k^{\alpha}}{x_i^{\alpha+1}}$$
$$= \alpha^n k^{\alpha n}(x_1 x_2 x_3 \ldots x_n)^{-\alpha-1}$$
$$\ln L(k,\alpha) = \ln \alpha^n + \ln k^{\alpha n} + \ln (x_1 x_2 x_3 \ldots x_n)^{-\alpha-1}$$
$$= n \ln \alpha + \alpha n \ln k - \alpha \sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \ln x_i$$

dimaksimumkan untuk memperoleh penduga α, sehingga diperoleh

$$\frac{\partial L(k,\alpha)}{\partial \alpha} = \frac{n}{\alpha} + n \ln k - \sum_{i=1}^{n} \ln x_i$$

$$\frac{n}{\alpha}+n\ln k-\sum_{i=1}^{n}\ln x_i=0$$

$$\hat{\alpha}=\frac{n}{\sum_{i=1}^{n}\ln x_i-n\ln k}$$

dan untuk penduga $\hat{k}$ tidak dapat diperoleh dengan melakukan diferensial terhadap *k*. Namun, karena k adalah batas bawah dari peubah acak *x*, fungsi tersebut dapat dimaksimalkan jika peubah acak yang digunakan adalah peubah acak terkecil, maka diperoleh

$$\hat{k}=\min x_i$$

4.2.3. Metode Bayes

Metode Bayes digunakan untuk memperoleh penduga parameter α dan k dari distribusi Pareto menggunakan prior distribusi *Uniform* dengan nilai parameter $a = 0$, dan $b = 1$.

$$f(x)=\frac{1}{b-a}=1$$

4.2.3.1 Parameter Bentuk

Fungsi kepadatan peluang posterior dari α yang diketahui adalah

$$f(\alpha|x)=\frac{f(\alpha)L(k,\alpha)}{\int_{-\infty}^{\infty}f(\alpha)L(k,\alpha)d\alpha}$$

Distribusi posterior dapat diperoleh dari hasil bagi berikut.

$$f(\alpha|x)=\frac{\alpha^n k^{\alpha n}\prod_{i=1}^{n}\frac{1}{x_i^{\alpha+1}}}{\int_0^{\infty}\alpha^n k^{n\alpha}\prod_{i=1}^{n}\frac{1}{x_i^{\alpha+1}}\,d\alpha}$$

$$=\frac{\alpha^n k^{\alpha n}\prod_{i=1}^{n}\frac{1}{x_i^{\alpha+1}}}{\frac{1}{\prod_{i=1}^{n}x_i}\Gamma(n+1)\left(\frac{1}{\sum_{i=1}^{n}\ln\frac{x_i}{k}}\right)^{(n+1)}}$$

$$=\frac{\left(\sum_{i=1}^{n}\ln\frac{x_i}{k}\right)^{(n+1)}\alpha^{(n+1)-1}e^{-\alpha\left(\sum_{i=1}^{n}\ln\frac{x_i}{k}\right)}}{\Gamma(n+1)}$$

Maka penduga α diperoleh dari nilai tengah distribusi posterior, hasilnya sebagai berikut.

$$E(\alpha)=\int_0^{\infty}\alpha\,f(\alpha|x)d\alpha$$

$$=\int_0^{\infty}\alpha\,\frac{\left(\sum_{i=1}^{n}\ln\frac{x_i}{k}\right)^{(n+1)}\alpha^n e^{-\alpha\left(\sum_{i=1}^{n}\ln\frac{x_i}{k}\right)}}{\Gamma(n+1)}\,d\alpha$$

$$\hat{\alpha}=\frac{(n+1)}{\sum_{i=1}^{n}\ln\frac{x_i}{k}}$$

4.2.3.2 Parameter Skala

Fungsi kepadatan peluang posterior dari k yang diketahui adalah

$$f(k|x)=\frac{f(k)L(k,\alpha)}{\int_{-\infty}^{\infty}f(k)L(k,\alpha)d\alpha}$$

Distribusi posterior dapat diperoleh dari hasil bagi berikut.

$$f(k|x)=\frac{\alpha^n k^{\alpha n}\prod_{i=1}^{n}\frac{1}{x_i^{\alpha+1}}}{\int_0^{x_1}\alpha^n k^{n\alpha}\prod_{i=1}^{n}\frac{1}{x_i^{\alpha+1}}dk}$$

$$= \frac{\alpha^n k^{\alpha n} \prod_{i=1}^{n} \frac{1}{x_i^{\alpha+1}}}{\prod_{i=1}^{n} \frac{\alpha^n}{x_i^{\alpha+1}(\alpha n+1)} x_1^{\alpha n+1}}$$

$$= \frac{k^{\alpha n}(\alpha n+1)}{x_1^{\alpha n+1}}$$

Maka penduga k diperoleh dari nilai tengah distribusi posterior, hasilnya sebagai berikut.

$$E(k) = \int_0^{x_1} k\, f(k) dk$$

$$= \int_0^{x_1} k \frac{k^{\alpha n}(\alpha+1)}{x_1^{\alpha n+1}} dk$$

$$\hat{k} = \frac{\alpha n+1}{\alpha n+2} x_1$$

4.3. Sifat Penduga Parameter α dan k

4.3.1. Metode Momen

4.3.1.1. Konsisten

Jika *plim* $x_1 = k$, dan jika rata-rata sampel merupakan penduga yang konsisten untuk rata-rata populasi, maka diperoleh,

$$plim\ \hat{\alpha} = plim \left(\frac{n\bar{x} - x_1}{n(\bar{x} - x_1)}\right)$$

$$= plim \left(\alpha + \frac{1-\alpha}{n}\right)$$

$$= \alpha$$

Jadi, $\hat{\alpha}$ merupakan penduga konsisten bagi α,

$$plim\ \hat{k} = plim \left(\frac{(\alpha n - 1)x_1}{\alpha n}\right)$$

$$= plim \left(k - \frac{k}{\alpha n}\right)$$

$$= k$$

dan $\hat{k}$ merupakan penduga konsisten bagi k.

4.3.1.2 Tidak Bias

Jika $\hat{\alpha}$ merupakan penduga tak bias bagi α, maka hal tersebut dapat dibuktikan dengan $E(\hat{\alpha}) = \alpha$ berikut.

$$E\left(\frac{n\bar{x} - x_1}{n(\bar{x} - x_1)}\right) = \frac{\alpha(n-1)}{n-1}$$

$$= \alpha$$

Selanjutnya, jika $\hat{k}$ merupakan penduga tak bias bagi k, maka hal tersebut dapat dibuktikan dengan $E(\hat{k}) = k$ berikut.

$$E\left(\frac{\alpha n - 1}{\alpha n} x_1\right) = \frac{(\alpha n - 1)}{\alpha n} \frac{\alpha n k}{(\alpha n - 1)}$$

$$= k$$

4.3.2. *Maximum Likelihood*

4.3.2.1 Konsisten

Perlu diperhatikan bahwa ketika $\hat{k} = \min x_i$ maka $\hat{k} - \min x_i = 0$. Sehingga, $plim\ |\hat{k} - \min x_i| = 0$ dan merupakan penduga yang konsisten. Untuk $\hat{\alpha}$,

$$plim\ \hat{\alpha} = plim \left(\frac{1}{\ln k + \frac{1}{\alpha} - \ln k}\right)$$

$$= \alpha$$

Jadi, dapat disimpulkan bahwa $\hat{\alpha}$ merupakan penduga konsisten bagi α.

4.3.2.2. Tidak Bias

Jika $\hat{\alpha}$ merupakan penduga tak bias bagi α, maka hal tersebut dapat dibuktikan dengan $E(\hat{\alpha}) = \alpha$ berikut.

$$E\left(\frac{1}{\frac{1}{n}E(\sum_{i=1}^{n}\ln x_i) - \ln k}\right) = \frac{1}{\ln k + \frac{1}{\alpha} - \ln k}$$
$$= \alpha$$

Selanjutnya, jika $\hat{k}$ merupakan penduga tak bias bagi k, maka hal tersebut dapat dibuktikan dengan $E(\hat{k}) = k$ berikut.

$$E(\min x_i) = k$$
$$\left(\frac{\alpha n k}{\alpha n - 1}\right) = k$$
$$\alpha n k = \alpha k n - k$$
$$0 = k$$

Hasil tersebut dapat dikatakan bias karena *k* tidak sama dengan penduganya.

4.3.3. Metode Bayes

4.3.3.1. Konsisten

$$plim\ \hat{\alpha} = plim\ \left(\frac{n+1}{\sum_{i=1}^{n} x_i - \ln k}\right)$$
$$= \alpha$$

Berdasarkan hasil tersebut, dapat disimpulkan bahwa $\hat{\alpha}$ merupakan penduga yang konsisten bagi α.

$$plim\ \hat{k} = plim\left(\frac{n\alpha + 1}{n\alpha + 2}x_1\right)$$
$$= k$$

Jadi, dapat disimpulkan bahwa $\hat{k}$ merupakan penduga yang konsisten bagi *k*.

4.3.3.2. Tidak Bias

$$E(\hat{\alpha}) = \alpha + \frac{\alpha}{n}$$

Berdasarkan hasil tersebut, dapat disimpulkan bahwa $\hat{\alpha}$ merupakan penduga yang bias bagi α.

$$E(\hat{k}) = k\,\frac{(n+1)}{\alpha n - 1}$$

Berdasarkan hasil tersebut, dapat disimpulkan bahwa $\hat{k}$ merupakan penduga bias bagi *k*.

4.4. Simulasi

Hasil simulasi perbandingan metode Momen, *Maximum Likelihood*, dan *Bayes* untuk distribusi Pareto menggunakan *software* R dengan data sebanyak n = 10, n = 40, n = 100, n = 200 dengan kombinasi parameter α = 1, α = 3, α = 5 dan k = 1, k = 3, dan k = 5 dapat dilihat pada tabel berikut.

Tabel 4.1 Nilai Mean Square Error dari Setiap Metode Pendugaan

N		k = 1, α = 1			k = 1, α = 3		
		MM	MLE	Bayes	MM	MLE	Bayes
10	Mean	10,112441	5,139763	-3,404715	1,490600	1,502812	-3.404715
	MSE (α)	0,0001019093	0,0000001576	0,0002305750	0,0000001202	0,0000014192	0,04630445
	MSE (k)	0,0000046582	0,0000000667	0,0000828987	0,0000009505	0,0000000074	0,0000828987
40	Mean	10,087082	-4,843419	-3,404715	1,492767	1,494834	-3.404715
	MSE (α)	0,0000250078	0,0000043902	0,0002305750	0,0000350063	0,0000395121	0,04630445
	MSE (k)	0,0000001848	0,0000000667	0,0000828987	0,0000000384	0,0000000074	0,0000828987
100	Mean	10,198008	10,222372	-3,404715	1,498064	1,499584	-3.404715
	MSE (α)	0,0000430971	0,0000043618	0,0002305750	0,0000049889	0,0000392563	0,04630445
	MSE (k)	0,0000000000	0,0000000667	0,0000828987	0,0000000004	0,0000000074	0,0000828987
200	Mean	70,613873	10,282739	-3,404715	1,499639	1,499832	-3.404715
	MSE (α)	0,0000169061	0,0000013473	0,0002305750	0,0000748681	0,0001212539	0,04630445
	MSE (k)	0,0000000015	0,0000000102	0,0000828987	0,0000000006	0,0000000011	0,0000828987
		k = 1, α = 5			k = 2, α = 1		
10	Mean	1,245842	2,493883	-3.404715	2,224881	10,027953	-3.404715
	MSE (α)	0,0000637814	0,0000039423	0,1723783	0,0001019093	0,0000001576	0,0002305750
	MSE (k)	0,0000003791	0,0000000609	0,0000828987	0,0000186329	0,0000002669	0,01190387
40	Mean	1,246882	2,493982	-3.404715	20,174165	-9,686837	-3.404715

	MSE (α)	0,0001462542	0,0001097558	0,1723783	0,0000250078	0,0000043902	0,0002305750
	MSE (k)	0,0000000144	0,0000000106	0,0000828987	0,0000000739	0,0000002669	0,01190387
100	Mean	1,249295	2,498897	-3.404715	20,396016	20,444743	-3.404715
	MSE (α)	0,0000558923	0,0001090454	0,1723783	0,0000430971	0,0000043618	0,0002305750
	MSE (k)	0,0000000002	0,0000000106	0,0000828987	0,0000000001	0,0000002669	0,01190387
200	Mean	2,499320	2,499450	-3.404715	100,522775	20,565477	-3.404715
	MSE (α)	0,0002814250	0,0003368165	0,1723783	0,0000169061	0,0000134727	0,0002305750
	MSE (k)	0,0000000009	0,0000000016	0,0000828987	0,0000000062	0,0000000409	0,01190387
			k = 2, α = 3			k = 2, α = 5	
10	Mean	2,981200	3,005624	-3.404715	2,491683	2,493883	-3.404715
	MSE (α)	0,0000001202	0,0000014192	0,04630445	0,0000637814	0,0000039423	0,1723783
	MSE (k)	0,0000038021	0,0000000295	0,01190387	0,0000015163	0,0000000106	0,01190387
40	Mean	2,985534	2,989667	-3.404715	2,493765	2,493982	-3.404715
	MSE (α)	0,0000350063	0,0000395121	0,04630445	0,0001462542	0,0001097558	0,1723783
	MSE (k)	0,0000001536	0,0000000295	0,01190387	0,0000000576	0,0000000106	0,01190387
100	Mean	2,996129	2,999168	-3.404715	2,498590	2,498897	-3.404715
	MSE (α)	0,0000049889	0,0000392563	0,04630445	0,0000558923	0,0001090454	0,1723783
	MSE (k)	0,0000000019	0,0000000295	0,01190387	0,0000000009	0,0000000106	0,01190387
200	Mean	2,999278	2,999664	-3.404715	2,499320	2,499450	-3.404715
	MSE (α)	0,0000748681	0,0001212539	0,04630445	0,0002814250	0,0003368165	0,1723783
	MSE (k)	0,0000000024	0,0000000045	0,01190387	0,0000000009	0,0000000016	0,01190387
			k = 3, α = 1			k = 3, α = 3	
10	Mean	30,337322	10,541929	-3.404715	4,471800	4,508436	-3.404715
	MSE (α)	0,0001019093	0,0000001576	0,0002305750	0,0000001202	0,0000014192	0,04630445
	MSE (k)	0,0000419239	0,0000006007	0,04372485	0,0000085547	0,0000000664	0,04372485
40	Mean	30,261247	-10,453026	-3.404715	4,478301	4,484501	-3.404715
	MSE (α)	0,0000250078	0,0000043902	0,0002305750	0,0000350063	0,0000395121	0,04630445
	MSE (k)	0,0000016629	0,0000006007	0,04372485	0,0000003455	0,0000000664	0,04372485
100	Mean	30,594024	30,667115	-3.404715	4,494193	4,498752	-3.404715
	MSE (α)	0,0000430971	0,0000043618	0,0002305750	0,0000049889	0,0000392563	0,04630445
	MSE (k)	0,0000000003	0,0000006007	0,04372485	0,0000000045	0,0000000664	0,04372485
200	Mean	200,284162	30,848216	-3.404715	4,498917	4,499497	-3.404715
	MSE (α)	0,0000169061	0,0000134727	0,0002305750	0,0000748681	0,0001212539	0,04630445
	MSE (k)	0,0000000139	0,0000000919	0,04372485	0,0000000054	0,0000000102	0,04372485
			k = 3, α = 5				
10	Mean	3,737525	3,740825	-3.404715			
	MSE (α)	0,0000637814	0,0000039423	0,1723783			
	MSE (k)	0,0000034116	0,0000000239	0,04372485			
40	Mean	3,740647	3,740974	-3.404715			
	MSE (α)	0,0001462542	0,0001097558	0,1723783			
	MSE (k)	0,0000001296	0,0000000239	0,04372485			
100	Mean	3,747885	3,748346	-3.404715			
	MSE (α)	0,0000558923	0,0001090454	0,1723783			
	MSE (k)	0,0000000020	0,0000000239	0,04372485			
200	Mean	3,748979	3,749174	-3.404715			
	MSE (α)	0,0002814250	0,0003368165	0,1723783			
	MSE (k)	0,0000000021	0,0000000037	0,04372485			

Hasil simulasi tersebut menunjukkan perbandingan nilai *Mean Square Error* parameter α dan k menggunakan metode Momen, *Maximum Likelihood*, dan Bayes. *Maximum Likelihood* memiliki nilai MSE terkecil dibandingkan Momen dan Bayes. Namun, pada kasus α = 1, penduga *likelihood* seringkali memperoleh nilai rerata yang kurang dari nol. Hal ini seharusnya tidak boleh terjadi, mengingat bahwa $0 < k \leq x$.

Berdasarkan hasil simulasi dapat disimpulkan bahwa metode penduga terbaik dalam menduga parameter distribusi Pareto adalah *Maximum Likelihood*.

4.5. Pembahasan

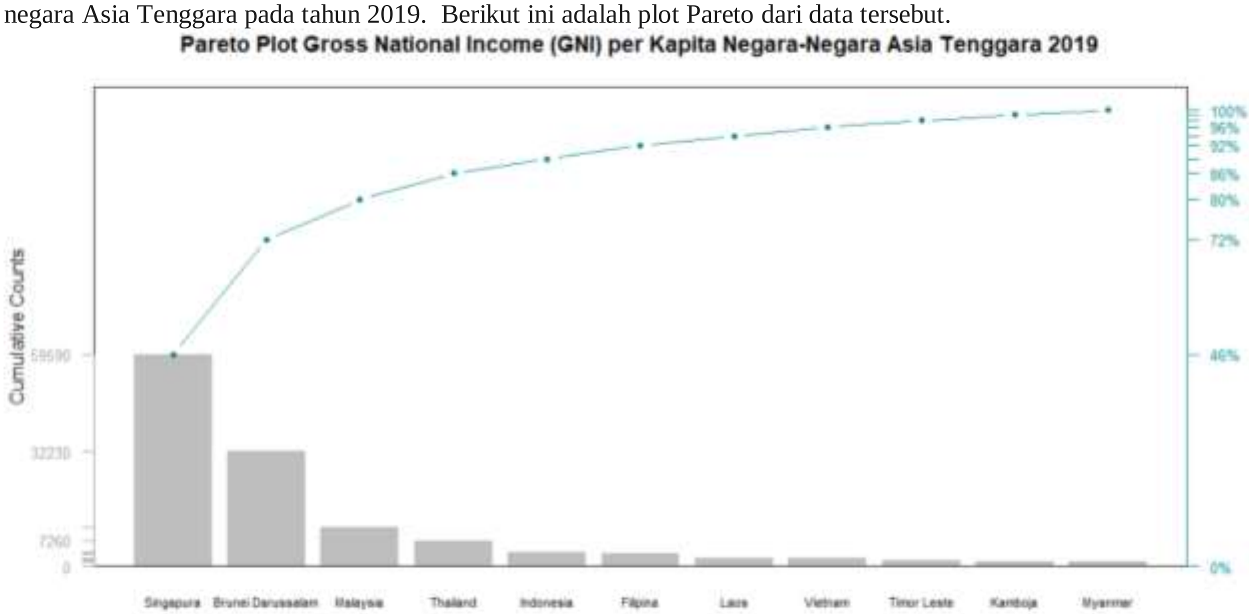
Berdasarkan hasil pendugaan yang diperoleh metode Momen dapat menduga parameter distribusi Pareto dengan penduga yang dapat dihitung oleh penduga parameter $\hat{\alpha} = \frac{n\bar{x}-x_1}{n(\bar{x}-x_1)}$ dan parameter $\hat{k} = \frac{(\hat{\alpha}n-1)x_1}{\hat{\alpha}n}$. Kedua penduga ini merupakan penduga yang tidak bias dan konsisten. Hasil ini sejalan dengan penelitian yang dilakukan oleh Quandt (1964) yang mengatakan bahwa kedua penduga ini merupakan penduga yang konsisten. Berdasarkan hasil simulasi, variansi atau MSE dari penduga Momen ini merupakan MSE yang kecil dibandingkan Bayes.

Di samping itu, pendugaan dengan menggunakan *Maximum Likelihood* menghasilkan penduga parameter $\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln x_i - n \ln k}$ dan $\hat{k} = \min x_i$. Penduga α ini merupakan penduga yang konsisten dan tidak bias berdasarkan pembuktian yang telah dilakukan sebelumnya. Sedangkan untuk penduga k disebut penduga yang konsisten namun bias. Hasil ini sejalan dengan penelitian yang dilakukan oleh Quandt (1964), Warsono (2019) dan Arnold (2015). Berdasarkan hasil simulasi yang dilakukan, penduga *Maximum Likelihood* memiliki MSE yang paling kecil dibandingkan metode Momen dan Bayes.

Pendugaan dengan menggunakan metode Bayes menghasilkan penduga parameter $\hat{\alpha} = \frac{(n+1)}{\sum_{i=1}^{n} \ln \frac{x_i}{k}}$ dan $\hat{k} = \frac{\alpha n+1}{\alpha n+2} x_1$. Hasil pendugaan ini sejalan dengan penelitian yang dilakukan oleh Setiya (2016) dan Yanuar (2020). Penduga ini merupakan penduga yang konsisten namun bias. Berdasarkan hasil simulasi, penduga ini merupakan penduga dengan nilai MSE yang terbesar dibandingkan dengan metode Momen dan *Maximum Likelihood*.

4.6. Penerapan pada Data Real

Pada penelitian ini data yang digunakan adalah data banyaknya *Gross National Income* (GNI) per kapita di negara-negara Asia Tenggara pada tahun 2019. Berikut ini adalah plot Pareto dari data tersebut.



Gambar 4.1 Plot Pareto

Pada Gambar 4.1, terlihat bahwa 80% pendapatan kumulatif negara di Asia Tenggara dipegang oleh Singapura, Brunei Darrussalam dan Malaysia, sedangkan 20% sisanya dipegang oleh negara lain di Asia Tenggara. Berikut ini analisis deskriptif dari data pendapatan per kapita negara di Asia Tenggara menggunakan hasil pendugaan parameter distribusi Pareto dengan *Maximum Likelihood* sebagai metode pendugaan terbaik.

Tabel 4.2 Parameter Data Kasus dengan Menggunakan Penduga *Maximum Likelihood*

α	*k*	*Mean*
0,7683109	1,390	-4.609,419

Berdasarkan Tabel 4.2, terlihat bahwa parameter bentuk dari data pendapatan negara di Asia Tenggara sebesar 0,7683109 sedangkan parameter skalanya sebesar 1.390. Adapun rata-rata pendapatan per kapita negara di Asia Tenggara memperoleh nilai yang kurang dari nol. Hal ini seharusnya tidak boleh terjadi, karena nilai $\alpha < 1$ maka rerata yang diperoleh adalah sebesar -4.609,419. Jika rerata data ini dihitung dengan menggunakan $\bar{x}$ maka hasilnya adalah 11.640,91.

5. Kesimpulan

Pendugaan parameter distribusi Pareto dengan menggunakan metode Momen menghasilkan penduga parameter $\hat{\alpha} = \frac{n\bar{x}-x_1}{n(\bar{x}-x_1)}$ dan parameter $\hat{k} = \frac{(\hat{\alpha}n-1)x_1}{\hat{\alpha}n}$. Pendugaan parameter distribusi Pareto dengan menggunakan *Maximum Likelihood Estimation* menghasilkan penduga parameter $\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\ln x_i - n\ln k}$ dan $\hat{k} = \min x_i$. Pendugaan parameter distribusi Pareto dengan menggunakan metode *Bayes* menghasilkan penduga parameter $\hat{\alpha} = \frac{(n+1)}{\sum_{i=1}^{n}\ln\frac{x_i}{k}}$ dan $\hat{k} = \frac{\alpha n+1}{\alpha n+2}x_1$.Hasil simulasi menunjukkan bahwa penduga *Maximum Likelihood* merupakan penduga yang memiliki nilai MSE yang terkecil. Jadi, metode pendugaan terbaik yang dapat menduga parameter distribusi Pareto adalah *Maximum Likelihood*.

References

Arnold, B. C. (2015). Pareto Distribution Second Edition. Boca Raton: CRC Press.

Bain, L. J., & Engelhardt, M. (1992). Introduction to Probability and Mathematical Statistics: Second Edition. USA: Brooks/Cole.

Casella, G., & Berger, R. L. (2002). Statistical Inference, Second Edition. USA: Thomson Learning.

Chotikapanich, D. (2008). Modelling Income Distribution and Lorenz Curves. New York: Springer.

Hogg, R. V., McKean, J. W., & Craig, A. T. (2005). Introduction to Mathematical Statistics: Sixth Edition. USA: Pearson Prentice Hall.

Kleiber, C., & Kotz, S. (2003). Statistical Size Distributions in Economics and Acturial Sciences. New Jersey: John Wiley & Sons International.

Montgomery, D. C., & Runger, G. C. (2014). Applied Statistics and Probability for Engineers. USA: John Wiley & Sons.

Pangestu, M. A., & Eliyah, S. (2018, Maret 25). Scribd. Retrieved April 4, 2020, from https://www.scribd.com/document/374797219/Distribusi-Pareto

Quandt, R. E. (1964). Old and New Methods of Estimation and the Pareto Distribution. Research Memorandum No. 70, 55-82.

Ridiani, F. (n.d.). Pendugaan Parameter Distribusi Beta dengan Metode Momen dan Metode Maksimum Likelihood. Jurnal Matematika UNAND, Vol. 3 No. 2 Hal. 23-28.

Rosenkrantz, W. (2009). Introduction to Probability and Statistics for Science, Engineering, and Finance. USA: CRC Press.

Sahoo, P. (2008). Probability and Mathematical Statistics. Louisville: University of Louisville.

Setiya, P., Kumar, V., & Pande, M. K. (2016). Bayesian Estimation of Scale Parameter of Pareto Type I Distribution by Two Different Methods. Thailand Statistician, 47-62.

Sudjana. (1986). Metoda Statistika Edisi ke IV. Bandung: Tarsito.

Tiro, M. A. (2015). Dasar-Dasar Statistika Edisi Keempat. Makassar: Andira Publisher.

Tiro, M. A., Sukarna, & Aswi. (2014). Pengantar Teori Peluang. Makassar: Andira Publisher.

Warsono, Gustavia, E., Kurniasari, D., Amanto, & Antonio, Y. (2019). On the Comparison of the Methods of Parameter Estimation for Pareto Distribution. Journal of Physics: Conference Series.

Will, K. (2008, April 3). Advantages of Method of Moments Estimators. Retrieved April 19, 2020, from TAMU Stat: www.stat.tamu.edu/

Yanuar, F., Saputri, C., & Devianto, D. (2020). Bayesian Inference for Pareto Distribution with Prior Conjugate and Prior Non Conjugate. Jurnal Matematika, Statistika & Komputasi, 382-390.