Int. J. Nonlinear Anal. Appl. 13 (2022) No. 1, 871-880
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.5605

Approximation (M.S.E) of the shape parameter for Pareto distribution by using the standard Bayes estimator

Adwea Naji Atewi[a,*], Safaa Jawad Ali[a], Bushra Swedan Naser[a], Muataz Adnan Harhoosh[a]

[a] *Institute of Medical Technology- Al-Mansour, Middle Technical University, Baghdad, Iraq*

(Communicated by Madjid Eshaghi Gordji)

Abstract

In this paper, Pareto distribution was studied using a standard Bayes estimator. A Pareto distribution of two parameters is investigated to find the approximation of (M.S.E) of the shape parameter by depending on Tyler series of two variables to propose a model mathematically.

Keywords: Approximation, Bayes estimator, Pareto distribution, Shape parameter, Jefferys prior.

1. Introduction

The principle of Pareto distribution is used now in many applications, such as oil trading markets and incomes rule. This distribution also proves that the level of the output and the level of input is not the same all the times depending on the rule of 80/20, see [3, 2]. Moreover, this principle is used in economics, population distribution and finance, see [10, 14]. In this work, many theories of Pareto are used to obtain Pareto distribution from the standard Bayes estimator throughout Tyler series of two variables, see [13, 11, 6]. Noninformative prior (p.d.f) depending on Jefferys prior rule is used to estimate parameters for this distribution, see [5, 15].

*Corresponding author

Email addresses: adwea-naji@mtu.edu.iq (Adwea Naji Atewi), safajawad1970@gmail.com (Safaa Jawad Ali), bushrasewdan@gmail.com (Bushra Swedan Naser), Muataz_adnan@mtu.edu.iq (Muataz Adnan Harhoosh)

Received: August 2021 *Accepted:* September 2021

1.1. preliminaries [13]

Let be a random variable r.v that follows a continuous Pareto distribution which is denoted by Where two parameters the scale and shape parameters respectively , has (p.d.f) probability density function:

$$f(x) = \begin{cases} \frac{\alpha c^{\alpha}}{x^{\alpha+1}} & c > 0, \quad \alpha > 0 \\ 0 & Otherwise \end{cases} \tag{1.1}$$

If $x = c$, then f above reaches the greater value and implies that:

$$f(\alpha, c) = \begin{cases} \frac{\alpha}{c} & c > 0, \quad \alpha > 0 \\ 0 & Otherwise \end{cases}$$

Figure 1 shows some Pareto p.d.f's:

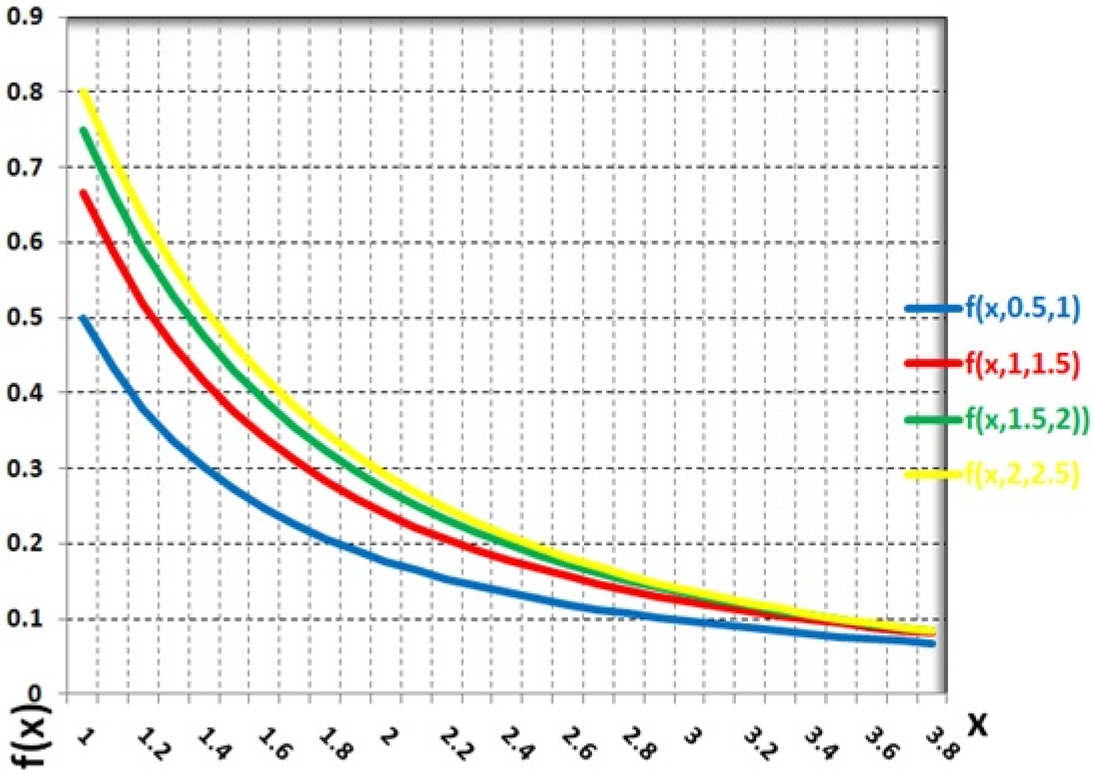

Figure 1: Pareto p.d.f's

1.2. The cumulative distribution function (c.d.f) [13]

(c.d.f) of the r.v X is According to the following formula:

$$F(x) = P(X \leq x) = \int_{c}^{x} f(u)du = 1 - \left(\frac{c}{x}\right)^{\alpha} \tag{1.2}$$

1.3. Hazard (Failure) Function

$h\left(x\right) = \frac{f(x)}{R(x)}$, where $R\left(x\right)$ is the reliability function $R(x) = \left(\frac{c}{x}\right)^{\alpha}$, and hence:

$$h(x) = \frac{\alpha}{x} \tag{1.3}$$

1.4. The cumulative hazard function

$$H(x) = -\ln R(x) = -\ln\left(\frac{c}{x}\right)^{\alpha} = -\alpha\left(\ln c - \ln x\right) \tag{1.4}$$

1.5. Mean and variance [11]

1.5.1. The population mean of X is:

$$E(x) = \frac{\alpha c}{\alpha - 1} = \int_c^{\infty} x f(x) dx$$

Simple simplifying implies that:

$$E(x) = \frac{\alpha c}{\alpha - 1} \tag{1.5}$$

1.5.2. Variance of Pareto distribution

$$Var(x) = \frac{c^2\alpha}{(\alpha-1)^2(\alpha-2)}$$
$$E(x^2) = \int_c^{\infty} x^2\ f(x)dx$$

Again, by solving this integral implies that:

$$E(x^2) = \frac{\alpha c^2}{\alpha - 2}$$
$$Var(x) = E(x^2) - [E(x)]^2 \tag{1.6}$$

Substituting (1.5) and (1.6) in that last equation then:

$$Var(x) = \frac{c^2\alpha}{(\alpha-1)^2(\alpha-2)} \tag{1.7}$$

1.6. Median and mode

1.6.1. Median for Pareto distribution [6]

A median for any variables are the middle number in a sorted, or is the value of x of r.v X such that $F(X) = p(X \leq x) = \frac{1}{2}$. For Pareto distribution case, $x_{md} = c^{\alpha}\sqrt{2}$

$$F(X) = \frac{1}{2}$$
$$x_{md} = c^{\alpha}\sqrt{2} \tag{1.8}$$

1.6.2. Mode for Pareto distribution

A mode is known in statistics, it is the most frequently occurring value in a set of data, or in a probability space, the mode for distribution might have more than one or may not exist.

For Pareto distribution the mode is:

$$Mode = c \tag{1.9}$$

1.7. Related theories [4, 12, 1]

Theorem 1.1. *A random sample of size n, let it be* $x_1,\ x_2,\ x_3,\ \ldots,\ x_n$ *from* $par(\alpha, c)$ *then:-*

(i) The first order statistics
(ii) The limiting distribution of

Proof .
(i) by using equations (1.1), (1.2), *ee have:*

$$y_1 = \min\{x_1, x_2, \ldots, x_n\}$$
$$P(y_1 \le y) = \begin{cases} 1 - P(y_1 > y), & y < c \\ 0, & y \ge c \end{cases}$$
$$\because P(x_i \le y) = 1 - P(x_i > y), \qquad i = 1, 2, \ldots, n$$

we have

$$P(x_i \le y) = \int_c^y f(x_i, c, \alpha)\ dx = 1 - \left(\frac{c}{y}\right)^{\alpha}$$

Therefore

$$P(Y_1 \ge y) = \Pi_{i=1}^{n} p(x_i \ge y) = \left(\frac{c}{y}\right)^{n\alpha}$$
$$P(Y_1 < y) = 1 - \left(\frac{c}{y}\right)^{n\alpha}$$
$$h(y) = \begin{cases} \frac{d}{dy}\left[1 - \left(\frac{c}{y}\right)^{n\alpha}\right], & c \le y < \infty \\ 0 & e.w \end{cases} = \begin{cases} \frac{v\alpha c^{n\alpha}}{y^{1+n\alpha}}, & c \le y < \infty \\ 0 & e.w \end{cases}$$

That is $Y_1 \sim par(n\alpha, c)$
(ii) $\because Y_1 \sim par\,(n\alpha, c)$, *then c.d.f of* Y_1 *is:*

$$h(Y_1) = \begin{cases} 0 & Y_1 \le c \\ 1 - \left[1 - \frac{c}{Y_1}\right]^{n\alpha} & c < Y_1 < \infty \\ 1 & Y_1 \to \infty \end{cases}$$

Then

$$\lim_{n\to\infty} h(Y_1) = \begin{cases} 0 & Y_1 \le c \\ 1 & Y_1 > c \end{cases}$$
$$E(Y_1) = \frac{n\alpha c}{n\alpha - 1} \tag{1.10}$$
$$Var(Y_1) = \frac{n\alpha c^2}{(n\alpha - 1)^2(n\alpha - 2)} \tag{1.11}$$

□

Theorem 1.2. *$X \sim par(\alpha, c)$, where the random variable c is unknown parameter therefore:*

$$Y = \ln X \sim Exp(\beta, \phi), \text{ where } \beta = \ln c \text{ and } \phi = 1/\alpha$$

Proof .

$$\begin{aligned}
X &= e^Y \\
|J| &= \left|\frac{\partial x}{\partial Y}\right| = e^Y, \qquad J \text{ is the jacobian} \\
\because f(x) &= \frac{\alpha c^\alpha}{x^{1+\alpha}}, \qquad x > c \\
g(Y) &= f(e^Y).|J| = \frac{\alpha c^\alpha}{(e^Y)^{1+\alpha}}.e^Y = \alpha c^\alpha e^{-\alpha Y}
\end{aligned}$$

Since $\beta = \ln c$, $\phi = 1/\alpha \Longrightarrow c = e^\beta$, $\alpha = 1/\phi$, then $\qquad g(Y) = (1/\phi).e^{-(Y-\beta)/\phi}, \quad Y > 0$
Hence, $Y \sim Exp(\beta, \phi)$. Then $\mu_r = E(Y) \& Var(Y)$

$$E(Y) = \phi + \beta \Longrightarrow \mu_r = E(Y) = \ln c + 1/\alpha \tag{1.12}$$
$$Var(Y) = \phi^2 \Longrightarrow Var(Y) = 1/\alpha^2 \tag{1.13}$$

□

1.8. Standard Bayes Estimator [7]

The idea of Bayes estimation is relies on the assumption that the parameter is a random variable, and that there is prior information (Prior P.d.f) about this parameter that is combined with the Posterior information (Posterior. P.d.f):

$$\begin{aligned}
&h(\phi|x_, \ldots, x_n) \propto L(x_1, \ldots, x_n : \phi) \quad where \quad L(x_1, \ldots, x_n : \phi) \quad \text{is Likelihood function} \\
&h(\phi|x_, \ldots, x_n) = \kappa L(x_1, \ldots, x_n : \phi)\ J(\phi), \ \ where \ \ J(\phi) \ \ textisPriorP.d.f \\
&\kappa^{-1} = \int_{\forall\phi} \kappa L(x_1, \ldots, x_n : \phi)\ J(\phi)\ d\phi; \ where\ \kappa\ is\ constant
\end{aligned}$$

Suppose we use the quadratic loss function $L(\hat{\phi}, \phi) = (\hat{\phi}^2 - \phi)^2$ and the risk function $E\{L(\hat{\phi}, \phi)\}$

$$\begin{aligned}
E\{L(\hat{\phi}, \phi)\} &= \int_{\forall\phi} (\hat{\phi} - \phi)^2\ h(\phi|x_, \ldots, x_n)\ d\phi \\
&= \int_{\forall\phi} \left((\hat{\phi})^2 - 2\hat{\phi}\phi + \phi^2\right)\ h(\phi|x_, \ldots, x_n)\ d\phi \\
&= \hat{\phi}^2 - 2\hat{\phi}E(\phi|x_1, \ldots, x_n) + E(\phi^2|x_1, \ldots, x_n) \\
\frac{\partial}{\partial\hat{\phi}} E\{L(\hat{\phi}, \phi)\} &= 2\hat{\phi} - 2E(\phi|x_1, \ldots, x_n) = 0 \Longrightarrow \hat{\phi} = E(\phi|x_1, \ldots, x_n) \\
\because \frac{\partial^2}{\partial\hat{\phi}^2} E\{L(\hat{\phi}, \phi)^2\} &= 2 > 0
\end{aligned}$$

Then the critical point is $\hat{\phi}_B = E(\phi|x_1, \ldots, x_n)$ a local minimum point.

1.8.1. Using noninformative prior (p.d.f) of estimation parameters for Pareto distribution:

Relying on the Jeffreys prior method see [5, 15], prior probability density function which is each of the two parameters α and c, so we can assume:

$$J_1(\alpha) \propto 1/\alpha^s \ , \qquad \alpha > 0$$
$$J_2(c) \propto 1/c^r, \qquad c > 0$$
$$J(\alpha, c) \propto J_1(\alpha).J_2(c)$$
$$J(\alpha, c) \propto 1/\alpha^s c^r$$

A random sample of size n, let it be $x_1, \ x_2, \ x_3, \ \dots, \ x_n$ from $par(\alpha, c)$, Hence the likelihood function is:

$$L(x_1, x_2, \dots, x_n; \alpha, c) = \Pi_{i=1}^n f(x_i/\alpha, c) = \Pi_{i=1}^n \frac{\alpha c^\alpha}{x_i^{\alpha+1}} = \Pi_{i=1}^n \frac{\alpha c^\alpha}{x_i^{\alpha+1}}.$$

we get

$$h(x_1, x_2, \dots, x_n; \alpha, c) \propto \frac{1}{\alpha^s c^r}.\frac{\alpha^n c^{n\alpha}}{(\Pi_{i=1}^n x_i)^{\aleph+1}}$$
$$h(x_1, x_2, \dots, x_n; \alpha, c) \propto \alpha^{n-s} c^{n\alpha-r} Exp\left[-(\alpha+1)\sum_{i=1}^n \ln x_i\right]$$

Let $r = 1$

$$h(x_1, x_2, \dots, x_n; \alpha, c) = \kappa \alpha^{n-s} c^{n\alpha-1} Exp\left[-(\alpha+1)\sum_{i=1}^n \ln x_i\right]$$

$$\begin{aligned}
\kappa^{-1} &= \int_0^\infty \int_0^{x_{(1)}} \alpha^{n-s} c^{n\alpha-1} Exp\left[-(\alpha+1)\sum_{i=1}^n \ln x_i\right] dc\, d\alpha \\
&= Exp\left(-\sum_{i=1}^n \ln x_i\right) \int_0^\infty \left[\frac{c^{n\alpha}}{n\alpha}\right]_0^{x_{(1)}} \alpha^{n-s} Exp\left(-\alpha \sum_{i=1}^n \ln x_i\right) d\alpha \\
&= Exp\left(-\sum_{i=1}^n \ln x_i\right) \int_0^\infty \alpha^{n-s} \frac{x_{(1)}^{n\alpha}}{n\alpha} Exp\left(-\alpha \sum_{i=1}^n \ln x_i\right) d\alpha \\
&= \frac{Exp\left(-\sum_{i=1}^n \ln x_i\right)}{n} \int_0^\infty \alpha^{n-s-1} Exp\left[-\alpha\left(\sum_{i=1}^n \ln x_i - n \ln x_{(1)}\right)\right] d\alpha \\
&= \frac{Exp\left(-\sum_{i=1}^n \ln x_i\right)}{n} \int_0^\infty \alpha^{n-s-1} Exp\left[-\alpha \sum_{i=1}^n \ln\left(\frac{x_i}{x_{(1)}}\right)\right] d\alpha \\
\kappa^{-1} &= \frac{Exp\left(-\sum_{i=1}^n \ln x_i\right)}{n} \frac{\Gamma(n-s)}{\left[\sum_{i=1}^n \ln\left(\frac{x_i}{x_{(1)}}\right)\right]^{n-1}} \\
\kappa &= \frac{n}{Exp\left(-\sum_{i=1}^n \ln x_i\right)} \frac{\left[\sum_{i=1}^n \ln\left(\frac{x_i}{x_{(1)}}\right)\right]^{n-1}}{\Gamma(n-s)}
\end{aligned}$$

$$h(\alpha, c, X_1, \ldots, X_n) = \frac{n\alpha^{n-s} c^{n\alpha-1} Exp\left(-\alpha \sum_{i=1}^{n} \ln x_i\right)}{\Gamma(n-s)\left[\sum_{i=1}^{n} \ln\left(\frac{x_i}{x_{(1)}}\right)\right]^{-(n-s)}}, \quad 0 < c < x_{(1)} < x < \infty, \quad \alpha > 0.$$

$$\hat{c} = x_{(1)} = \min\{x_1, x_2, \ldots, x_n\} \tag{1.14}$$

$$M(\alpha|X_1, X_2, \ldots, X_n) = \int_c h(\alpha, c, X_1, \ldots, X_n)\, dx$$

$$M(\alpha|X_1, X_2, \ldots, X_n) = \frac{\alpha^{n-s-1} e^{\left(-\alpha \sum_{i=1}^{n} \ln x_i\right)}}{\Gamma(n-s)\left[\sum_{i=1}^{n} \ln\left(\frac{x_i}{x_{(1)}}\right)\right]^{n-s}}$$

Suppose we use the quadratic loss function so Bayes estimator for is mean of posterior:

$$\hat{\alpha} = E\left(\frac{\alpha}{x}\right) = \frac{n-s}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{c}\right)}$$

If $s = 1$, then the standard Bayes estimator is maximum likelihood.
If $s = 2$, therefore standard Bayes is unbiased & min var of parameter α [9].

$$\hat{\alpha} = \frac{n-s}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{c}\right)} \tag{1.15}$$

1.9. Approximation (M.S.E) of the standard Bayes estimators [8]

Let $X_1,\ X_2,\ \ldots,\ X_n$be a r.v of size n follows Pareto distribution with two parameters. We will approximation (M.S.E) of the standard Bayes estimator, by depending on Tyler series this series was named after the English mathematician Brook Taylor creating Taylor theorem (1685 – 1731). For Tyler series several uses, but the most important it allows to express any mathematical function through many borders can we find approximate solutions to the problem of whether the exact solution intractable. Then we take Tyler series of two - variables $R,\ S$ at point $(\mu_1,\ \ \mu_2)$ up to second degree by using the following equations:

$$\begin{aligned} E[g(R,S)] = g(\mu_R, \mu_S) &+ \frac{1}{2} Var[R]\ \frac{\partial^2}{\partial R^2}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}} + \frac{1}{2} Var[S]\ \frac{\partial^2}{\partial S^2}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}} \\ &+ Cov[R,S]\left[\frac{\partial^2}{\partial R \partial S}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}}\right] \end{aligned} \tag{1.16}$$

$$\begin{aligned} Var[g(R,S)] = Var[R]\ &\left[\frac{\partial}{\partial R}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}}\right]^2 + Var[S]\ \left[\frac{\partial}{\partial S}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}}\right]^2 \\ &+ 2Cov[R,S]\left[\frac{\partial}{\partial R}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}}\right]\left[\frac{\partial}{\partial S}\ g(R,S)\bigg|_{\substack{\mu_R\\ \mu_S}}\right] \end{aligned} \tag{1.17}$$

by equations (1.14) and (1.15): $\hat{c} = x_{(1)} = \min\{x_1, x_2, \ldots, x_n\}$ and $\hat{\alpha} = \frac{n-s}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{c}\right)}$ By Theorem 1.1, then:

$$E[\hat{c}] = E[Y_1] = \frac{n\alpha c}{n\alpha - 1} \tag{1.18}$$

$\lim_{n\to\infty} E[\hat{c}] = \lim_{n\to\infty} \frac{n\alpha c}{v\alpha - 1} = c$. Shows that the is asymptotically and unbiased for c.

$$Var[\hat{c}] = Var[Y_1] = \frac{n\alpha c^2}{(n\alpha-1)^2(n\alpha-2)} \tag{1.19}$$

$\hat{\alpha} = \frac{n-s}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{c}\right)} = \frac{n-s}{\sum_{i=1}^{n} \ln x_i - n \ln y_i}$ Let, $M_i = \ln x_i$ *and* $W = \ln y_i$, we get $\hat{\alpha} = \frac{n-s}{\frac{n}{m}\sum M_i - nW}$.

Let $R = \overline{M},\ S = W$, in (1.16) and (1.17), then:

$$E[g(\overline{M}, S)] = g(\mu_{\overline{M}}, \mu_S) + \frac{1}{2} Var[\overline{M}]\ \frac{\partial^2}{\partial \overline{M}^2}\ g(R, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} + \frac{1}{2} Var[S]\ \frac{\partial^2}{\partial S^2}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}}$$
$$+ Cov[\overline{M}, S] \left[\frac{\partial^2}{\partial \overline{M} \partial S}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} \right]$$

and

$$Var[g(\overline{M}, S)] = Var[\overline{M}] \left[\frac{\partial}{\partial \overline{M}}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} \right]^2 + Var[S] \left[\frac{\partial}{\partial S}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} \right]^2$$
$$+ 2Cov[\overline{M}, S] \left[\frac{\partial}{\partial \overline{M}}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} \right] \left[\frac{\partial}{\partial S}\ g(\overline{M}, S)\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} \right].$$

By equations (1.12) and (1.13), we get:

$$\mu_{\overline{M}} = \ln c + \frac{1}{\alpha} \tag{1.20}$$
$$Var(\overline{M}) = \frac{1}{n\alpha^2} \tag{1.21}$$
$$\mu_W = \ln c + \frac{1}{n\alpha} \tag{1.22}$$
$$Var(W) = \frac{1}{n^2\alpha^2} \tag{1.23}$$

We has c.d.f from eq.(1.2), that is:

$$H(W) = \begin{cases} 0 & W \leq \ln c \\ 1 - \left(\frac{c}{e^W}\right)^{n\alpha} & \ln c \leq W \leq \infty \\ 1 & W = \infty \end{cases} \tag{1.24}$$

Shows that the r.v W converges randomly (stochastically) to $\ln c$

$$g(\overline{M}, W) = \hat{\alpha} = \frac{n-2}{n(\overline{M} - W)} \tag{1.25}$$

From Eq.(1.20) and Eq.(1.22), we get

$$g(\overline{M}, W)\big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} = \frac{n-2}{n\left(\ln c + \frac{1}{\alpha} - \ln c - \frac{1}{n\alpha}\right)} = \frac{\alpha(n-2)}{n-1} \tag{1.26}$$
$$\frac{\partial g(\overline{M}, W)}{\partial \overline{M}}\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} = \frac{-(n-2)}{n(\overline{M} - W)^2} = \frac{-n(n-2)\alpha^2}{(n-1)^2} \tag{1.27}$$
$$\frac{\partial^2 g(\overline{M}, W)}{\partial \overline{M}^2}\Big|_{\substack{\mu_{\overline{M}} \\ \mu_S}} = \frac{2(n-2)}{n(\overline{M} - W)^2} = \frac{2(n-2)}{n\left(\ln c + \frac{1}{\alpha} - \ln c - \frac{1}{n\alpha}\right)^2} = \frac{2(n-2)n^2\alpha^2}{(n-1)^2} \tag{1.28}$$

$$\left.\frac{\partial g(\overline{M},W)}{\partial W}\right|_{\mu_{\overline{M}} \atop \mu_S} = \frac{(n-2)}{n(\overline{M}-W)^2} = \frac{(n-2)}{n\left(\ln c + \frac{1}{\alpha} - \ln c - \frac{1}{n\alpha}\right)^2} = \frac{n(n-2)\alpha^2}{(n-1)^2} \tag{1.29}$$

$$\left.\frac{\partial^2 g(\overline{M},W)}{\partial W^2}\right|_{\mu_{\overline{M}} \atop \mu_S} = \frac{2(n-2)}{n(\overline{M}-W)^3} = \frac{2(n-2)}{n\left(\ln c + \frac{1}{\alpha} - \ln c - \frac{1}{n\alpha}\right)^3} = \frac{2(n-2)n^2\alpha^3}{(n-1)^3} \tag{1.30}$$

$$\left.\frac{\partial^2 g(\overline{M},W)}{\partial \overline{M}\partial W}\right|_{\mu_{\overline{M}} \atop \mu_S} = \frac{-2(n-2)}{n(\overline{M}-W)^3} = \frac{-2(n-2)}{n\left(\ln c + \frac{1}{\alpha} - \ln c - \frac{1}{n\alpha}\right)^3} = \frac{-2(n-2)n^2\alpha^3}{(n-1)^3} \tag{1.31}$$

Since $\overline{M} \xrightarrow{stochastically} \mu_{\overline{M}}$ *and* $W \xrightarrow{stochastically} \ln c$. So $\overline{M}W \xrightarrow{stochastically} \mu_{\overline{M}} \ln c$, therefore

$$E[\overline{M}W] \cong \mu_{\overline{M}} \ln c = \left[\ln c + \frac{1}{\alpha}\right] \ln c$$

$$Cov[\overline{M}W] \cong E[\overline{M}W] - E[\overline{M}]E[W] \cong \left(\ln c + \frac{1}{\alpha}\right) \ln c - \left(\ln c + \frac{1}{\alpha}\right)\left(\ln c + \frac{1}{n\alpha}\right)$$

$$Cov[\overline{M}W] = \frac{-(\alpha \ln c + 1)}{n\alpha^2} \tag{1.32}$$

$$E[\hat{\alpha}] \cong E[g(\overline{M},W)] \cong \frac{\alpha(n-2)}{(n-1)} + \frac{1}{2}\left[\frac{1}{n\alpha^2}\right]\left[\frac{2\alpha^3 n^2(n-2)}{(n-1)^3}\right] + \frac{1}{2}\left[\frac{1}{n^2\alpha^2}\right]\left[\frac{2\alpha^3 n^2(n-2)}{(n-1)^3}\right] + \left[\frac{-(\alpha \ln c + 1)}{n\alpha^2}\right]\left[\frac{2\alpha^3 n^2(n-2)}{(n-1)^3}\right]$$

$$E[\hat{\alpha}] = \frac{\alpha(n-2)}{n-1}\left[1 + \frac{1}{(n-1)^2}(\ 3n + 1 + 2\ln c\ )\right] \tag{1.33}$$

$$Var[\hat{\alpha}] = Var[g(\overline{M},W)] = \frac{1}{n\alpha^2}\left[\frac{-n\alpha^2(n-2}{(n-1)^2}\right]^2 + \frac{1}{n^2\alpha^2}\left[\frac{n\alpha^2(n-2}{(n-1)^2}\right]^2 + 2\left[\frac{-(\alpha \ln c + 1)}{n\alpha^2}\right]\left[\frac{-n\alpha^2(n-2}{(n-1)^2}\right]^2\left[\frac{n\alpha^2(n-2)}{(n-1)^2}\right]^2$$

$$Var[\hat{\alpha}] = \frac{n\alpha^2(n-2)^2}{(n-1)^4}\left[2\alpha \ln c + 3 + \frac{1}{n}\right] \tag{1.34}$$

$$M.S.E(\hat{\alpha}) = Var[\hat{\alpha}] + [bias\ (\hat{\alpha})]^2 \tag{1.35}$$

and

$$\text{bias}(\hat{\alpha}) = E(\hat{\alpha}) - \alpha \tag{1.36}$$

Put Eq. (1.34) in Eq. (1.36), then we get

$$\text{bias}(\hat{\alpha}) = \frac{\alpha(n-2)}{(n-1)}\left[1 + \frac{1}{(n-1)^2}(\ 3n + 1 + 2n\alpha \ln c\ )\right] - \alpha \tag{1.37}$$

Then put Eq. (1.34) and Eq. (1.37) in Eq. (1.35)

$$M.S.E(\hat{\alpha}) = \frac{n\alpha^2(n-2)^2}{(n-1)^4}\left[2\alpha \ln c + 3 + \frac{1}{n}\right] + \left[\frac{\alpha(n-2)}{(n-1)}\left[1 + \frac{1}{(n-1)^2}(\ 3n + 1 + 2n\alpha \ln c\ )\right] - \alpha\right]^2 \tag{1.38}$$

2. Conclusion

From our theoretical result a above, one can conclude that this method of approximate can be extended to applying other types of Pareto distribution and applying this method to find an approximation (M.S.E.) of estimation methods for other distributions. Moreover, the approximate mean squared error become more exact if we used the higher degree of approximation .

References

[1] M.P. Alan,*Introduction to statistical limit theory*, 1st Edition, Chapman and Hall, 2011.
[2] B.C. Arnold, *Pareto Distributions*, Second Edition, CRC Press, Boca Raton, Florida, 2015.
[3] J. Chu, O. Dickin and S. Nadarajah, *A review of goodness of fit test for Pareto distribution*, J. Comput. Appl. Math. 36 (2019) 13–41.
[4] R. V. Hogg and S.A. Klugman,*Loss Distribution*, John Wiley and Sons New York, 1984.
[5] D.-H. Kim, Sang.-G. Kang and W.-D. Lee, *Non information prior for Pareto distribution*, J. Korean Data Inf. Sci. Soc. 20 (2009) 1213–1223.
[6] A.A. Khaliq, *Census of Life, Arabic Edition*, Ammaan-Jordan, Dar Yazouri, 2011.
[7] E.L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., John Wiley and Sons, NY., 2003.
[8] H. Levy and H.M. Markowitz,*Approximating expected utility by a function of mean and variance*, Amer. Econ. Rev. 69 (1979) 308–317.
[9] H.J. Malik, *Estimation of the parameters of the Pareto distribution*, Metrika Int. J. Theor. Appl.Stat. 15 (1970) 126–132.
[10] S. G. Meintanis and Y. Bassiakrs, *Data – transformation and test of fit for the generalized Pareto hypothesis*, Commun. Stat. Theory Methods 13 (2007) 1965–1979.
[11] A.M. Mood, F.A. Graybill and D.C. Boess, *Introduction to The Theory of Statistics*, Third Edition, McGraw-Hill, London, 1974.
[12] S. Nadarajah and S. Kotz,*Reliability for Pareto Models*, Metro Int. J. Stat. LXI(2) (2003) 191–204.
[13] L. Norman, J.K. Samuel and N. Balakrishnan, *Continuous Univariate Distributions*, Second Edition, Wiley, 1994.
[14] M. Obradovic, M. Jovanovic and B. Milosevie, *Goodness-of-fit tests for Pareto distribution based on a characterization and their asymptotics Statistics*, J. Theor. Appl. Stat. 49 (2015) 1026–1041.
[15] W. Dong, D.-H. Kim and S.-G. Kang, *Non information prior for linear combinations of exponential means*, J. Korean Data Inf. Sci. Soc. 27 (2016) 565–575.