Statistical Methodology 21 (2014) 49–58

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

# On estimation for the Pareto distribution

Hui He, Na Zhou, Ruiming Zhang *

*College of Science, Northwest A&F University, Yangling, Shaanxi 712100, PR China*

ARTICLE INFO

*Article history:*
Received 1 June 2013
Received in revised form
22 January 2014
Accepted 12 March 2014

*Keywords:*
Pareto distribution
MLE
UMVUE
Probability density function
Reliability function

ABSTRACT

In this work, we obtain the $r$-th raw moments of the probability density function (PDF) and reliability function (RF) for the Pareto distribution under the maximum likelihood estimation (MLE) and uniform minimum variance unbiased estimation (UMVUE). We derive some large sample properties of the estimators, the MLE and UMVUE of the PDF as well as RF. Two examples are provided to compute the efficient estimations of PDF and RF numerically. Our results indicate that there are no absolute superiorities of MLEs over the UMVUEs of PDF and RF and vice versa.


## 1. Introduction

We consider a random variable $X$ has the Pareto distribution (PD) with PDF

$$f(x) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, \tag{1}$$

and RF

$$G(x) = \text{Prob}\{X > x\} = \left(\frac{k}{x}\right)^{\alpha}, \tag{2}$$

where $\alpha$ is a shape parameter ($\alpha > 0$), and $k$ is a scale parameter (known, and $x > k > 0$). In economics, $k$ usually represents some minimum income with a known value, see [1].

PD was applied by Pareto [7] to model the allocation of wealth among individuals and the distribution of incomes. It has been widely used in economics, insurance (general liability, commercial

* Corresponding author. Tel.: +86 13032906582.
*E-mail addresses:* hhznby@163.com (H. He), zhouna871208@163.com (N. Zhou), ruiming_nwsuaf@163.com, ruimingzhang@yahoo.com (R. Zhang).

http://dx.doi.org/10.1016/j.stamet.2014.03.002
1572-3127/

auto [9]), geography (sizes of human settlements [8]), physical sciences (sizes of sand particles or meteorites [8], clusters of Bose–Einstein condensate near absolute zero [5]), chemical sciences (distributions of electrolytic powder production [4]). Asrabadi [1] established the UMVUEs for the PDF and cumulative distribution function (CDF) of PD. Based on the work of Asrabadi [1], Dixit and Jabbari Nooghabi [2] tried to study the mean square errors (MSEs) of the MLEs and UMVUEs for the PDF and CDF of PD and their results seem to show that the MLEs are more efficient than the UMVUEs of PDF and CDF. Unfortunately, their work are seriously flawed. Most of their main claims in [2] are wrong, and their conclusion, the MLEs are more efficient than the UMVUEs of PDF and CDF, is unreasonable.

We present our main results in Section 2. Most of the results in Section 2.1 are corrected versions of the wrong results of [2]. We also notice that the exact expressions of the MSEs of estimators of PDF and RF may not be useful in case of large scale samples and large scale numerical computations. For this reason we have derived the asymptotic expressions of the $r$-th raw moments and MSEs in Section 2.2. Two numerical examples are provided in Section 2.3 to show how to compute the efficient estimations of PDF and RF. In Section 2.4 we expose the fatal errors in [2].

## 2. Main results

As a notational convenience, let

$$z = z(x) = \log\frac{x}{k}, \qquad z_x = z_x(x) = \frac{dz}{dx} = \frac{1}{x},$$

through the rest of this paper. It is known that the UMVUEs of $f(x)$ and $G(x)$ are given by [1]

$$\hat{f}(x) = \frac{n-1}{s} z_x \left(1 - \frac{z}{s}\right)^{n-2}, \qquad \hat{G}(x) = \left(1 - \frac{z}{s}\right)^{n-1},$$

where $z < s,\ s = \sum_{i=1}^{n} z(x_i)$ and $s$ follow the Gamma distribution $Ga(n, \alpha)$. Note that the UMVUE of $\alpha$ is $\hat{\alpha} = (n-1)/s$.

The MLEs of $f(x)$ and $G(x)$ can be computed easily, they are

$$\tilde{f}(x) = \tilde{\alpha} z_x e^{-\tilde{\alpha} z}, \qquad \tilde{G}(x) = e^{-\tilde{\alpha} z},$$

where $\tilde{\alpha} = \frac{n}{s}$ is the MLE of $\alpha$. Note that the PDF of $s$ is given by $h(s) = \frac{\alpha^n s^{n-1}}{\Gamma(n)} \exp(-\alpha s)$.

### *2.1. The r-th raw moments of estimations*

**Theorem 1.** *For $n > r > 0$, the $r$-th raw moments of $\tilde{f}(x)$ and $\tilde{G}(x)$ are given by*

$$E(\tilde{f}(x))^r = \frac{2}{\Gamma(n)} (n\alpha z_x)^r \left(\sqrt{nr\alpha z}\right)^{n-r} K_{n-r}\left(2\sqrt{nr\alpha z}\right), \tag{3}$$

$$E(\tilde{G}(x))^r = \frac{2}{\Gamma(n)} \left(\sqrt{nr\alpha z}\right)^{n} K_{n}\left(2\sqrt{nr\alpha z}\right), \tag{4}$$

*where $K_\nu(x)$ is the modified Bessel function* [6].

**Proof.** For the proof we just need to note the well-known integral representation [6],

$$K_\nu(x) = \frac{1}{2}\left(\frac{x}{2}\right)^{\nu} \int_0^\infty \exp\left(-t - \frac{x^2}{4t}\right) \frac{dt}{t^{\nu+1}}. \quad \square$$

**Corollary 1.** *The mean square errors of $\tilde{f}(x)$ and $\tilde{G}(x)$ are given by*

$$\begin{aligned} \mathrm{MSE}(\tilde{f}(x)) = {} & \frac{2(n\alpha z_x)^2}{\Gamma(n)} \left(\sqrt{2n\alpha z}\right)^{n-2} K_{n-2}\left(2\sqrt{2n\alpha z}\right) \\ & - \frac{4n\alpha z_x}{\Gamma(n)} f(x) \left(\sqrt{n\alpha z}\right)^{n-1} K_{n-1}\left(2\sqrt{n\alpha z}\right) + f^2(x). \end{aligned} \tag{5}$$