Int. J. Nonlinear Anal. Appl. 13 (2022) 1, 1873-1877
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.5819

Estimation of parameter for the Pareto distribution based on right censoring

Rana Hasan Shamkhi[a,*], Wisam Kamil Ghafil[b], Aseel Ali Jaaze[a]

[a]*Department of pharmacognosy and medicinal plants, University of Basrah, Iraq*
[b]*Department of Mathematics, University of Thi-Qar, Iraq*

(Communicated by Madjid Eshaghi Gordji)

Abstract

In this paper, we found the estimation of the unknown parameter δ when ϑ is a known parameter in the Pareto distribution. First, we get the maximum probability estimators(MLEs) for unknown parameters. We have obtained the Bayes Estimators of unknown parameter δ using Lindley's approximation. A Monte Carlo simulation is performed and used a programming language R to compare the performance of the method used, and the data set was analyzed for illustration purposes.

Keywords: Estimation Pareto distribution, right censoring

1. Introduction

Pareto distribution is a continuous distribution. Refer to Pareto distributions are useful modeling and predicting tools in a wide variety of socioeconomic contexts, see in [11] and [13] Rytgaard, Mette.(1990) the probability density function (pdf), Cumulative distribution function (cdf)and survival function for any random variable from Pareto with parameters δ and ϑ , respectively are

$$f(x) =\frac{\delta\vartheta^{\delta}}{x^{\delta+1}}, \delta > 0, \vartheta > 0, > x \geq \vartheta \tag{1.1}$$

$$F(x) =1 - (\frac{\vartheta}{x})^{\delta}, \delta > 0, \vartheta > 0, > x \geq \vartheta \tag{1.2}$$

$$S(x) =(\frac{\vartheta}{x})^{\delta}, \delta > 0, \vartheta > 0, x \geq \vartheta \tag{1.3}$$

*Corresponding author
Email addresses: rana.jbr.sci@uobasrah.edu.iq (Rana Hasan Shamkhi), Wisam.kamil@utq.edu.iq (Wisam Kamil Ghafil), aseel.jaaze@uobasrah.edu.iq (Aseel Ali Jaaze)

Received: May 2021 *Accepted:* September 2021

The unknown parameter $\delta > 0$, and known parameter $\vartheta = 1$. If $x < \vartheta$, the pdf is zero .If x has pdf (1.1)we write $P(\delta, \vartheta)$.The probability density function(pdf) and the cumulative distribution function(cdf) become as follows

$$f(x) = \frac{\delta}{x^{\delta+1}}, \delta > 0, > x \geq \vartheta \tag{1.4}$$

$$F(x) = 1 - \frac{1}{(x)^{\delta}}, \delta > 0, x \geq \vartheta \tag{1.5}$$

[14] proposed a study about progressive Type-II censored data of the Pareto distribution and [16]their study was about the combined-unified under hybrid censored samples from pareto distribution,[9]have studied a new approach under the regression framework is suggested the estimation of unknown parameters the Pareto distribution based on the progressive Type-II censoring data.In [4] Estimation of the population mean under right censored observations is studied As well[8] estimation of the Parameter of the Pareto distribution by used a pivotal quantity ,in [12]have studied about Bayes Estimators for the Parameter of Pareto Type I Distribution using Generalized Square Error Loss Function.The Pareto distribution It is used in quality control, scientific, actuarial and many other kinds of observable phenomena. It was mainly used to describe the distribution of wealth in a society , It can be used in many other situations.like, it can be used to model the life of an item manufactured with a certain warranty period and it have many applications in survival analysis, economics, life testing, finance, telecommunication, reliability analysis, engineering and physics .

2. Maximum likelihood estimation

censored Data is the data that at least has a single one of the observations is incomplete , as in clinical trials, maybe it will be that only a minimum time of the event can be observed because the study finishes before the event occurs, or subjects leave the study. That is called right-censoring see in [15] and [2]describe parameter estimation of Burr distribution for right censored data using Bayes method . Let Z_j reference the time of the event and C_j reference the time of the censoring for unit $j = 1, \cdots, n$.based on right censoring, the data that is observed contain the pair.

$$T_j = \min(Z_j, C_j), D_j = l(Z_j \leq C_j), j = 1, \cdots, n \tag{2.1}$$

we notice here $l(.)$ is the indicator function, therefore D_i identifies either T_j is the time of the event or the time of the censoring. Let $Y = (T_j, D_j) : j = 1, \cdots, n$ reference the observable data. For data $y = (t_j, d_j)$ observed from a random based on Type I right-censored data generating process,[3] is give the likelihood function,

$$\begin{aligned} L_y(\delta) &\propto \prod_{j=1}^{n} f_\delta(t_j)^{d_j} (1 - F_\delta(t_j))^{1-d_j} \\ &= \prod_{j=1}^{n} \Big(\frac{\delta}{(t_j)^{\delta+1}}\Big)^{d_j} \Big(\frac{1}{t_j}\Big)^{\delta(1-d_j)} \end{aligned} \tag{2.2}$$

with corresponding log-likelihood function

$$l(\delta|D) \propto \sum_{j=1}^{n} d_j(\log\delta - (\delta+1)\log t_j + \delta \sum_{j=1}^{n}(1-d_j)\log(\frac{1}{t_j}) \tag{2.3}$$

Now, the first partial derivatives of the log-likelihood for δ can be written as

$$l_\delta \equiv \frac{\partial l(\delta|D)}{\partial \delta} = \frac{1}{\delta}\sum_{j=1}^{n} d_j - \sum_{j=1}^{n} d_i \log(t_j) + \sum_{j=1}^{n}(1-d_j)\log(\frac{1}{t_j}) \tag{2.4}$$

To compute the MLEs of the unknown parameter δ , we need to solve normal equation $l_\delta = 0$, where l_δ is given in (2.4) .

$$\frac{1}{\delta}\sum_{j=1}^{n} d_j - \sum_{j=1}^{n} d_j \log(t_j) + \sum_{j=1}^{n}(1-d_j)\log(\frac{1}{t_j}) = 0 \tag{2.5}$$

We get

$$\hat{\delta} = \frac{\sum_{j=1}^{n} d_j}{\sum_{j=1}^{n} d_j \log(t_j) + \sum_{j=1}^{n}(1-d_j)\log(\frac{1}{t_j})}$$

3. Bayesian Estimation Using Lindley's Approximation

Bayesian estimation is a more workable method than the classical estimation methods, and has drew attention from many researchers like [17]and [7] were used maximum likelihood and Bayesian methods to get the estimators of the entropy for a two-parameter of Burr type XII distribution based on progressive type-II censored data , and [5]have studied Bayesian estimators of Weibull mixture in intensely censored data setting. Many people use the Bayesian method to estimate parameters and connected functions for various distributions. In this section, we estimate the unknown parameter δ of Pareto distribution .If ϑ were known, then a prior distribution for δ , is Gamma distribution . it is used for its flexibility. [1] studied the Bayesian estimation and prediction for Pareto data . In specific, we assume throughout this section that δ have independent $G(c,d)$, prior distribution with $c > 0$,$d > 0$ i.e.,

$$\pi(\delta) \propto \delta^{c-1} e^{-\delta d}$$

the joint pdf is given by

$$h(x,\delta) = \prod_{i=1}^{n}\Big(\frac{\delta}{(t_i)^{\delta+1}}\Big)^{d_i}\Big(\frac{1}{t_i}\Big)^{\delta(1-d_i)}\delta^{c-1}e^{-\delta d}$$

The marginal pdf of δ is obtain by

$$\begin{aligned} h(x) &= \int_0^\infty h(x,\delta)d\delta \\ &= \int_0^\infty \prod_{i=1}^{n}\Big(\frac{\delta}{(t_i)^{\delta+1}}\Big)^{d_i}\Big(\frac{1}{t_i}\Big)^{\delta(1-d_i)}\delta^{c-1}e^{-\delta d}d\delta \end{aligned} \tag{3.1}$$

Theprior distribution combined with the likelihood function is used to derive the posterior distribution. The posterior PDF is given by

$$h(\delta|x) = \frac{h(x,\delta)}{h(x)} \tag{3.2}$$

$$= \frac{\prod_{i=1}^{n}\Big(\frac{\delta}{(t_i)^{\delta+1}}\Big)^{d_i}\Big(\frac{1}{t_i}\Big)^{\delta(1-d_i)}\delta^{c-1}e^{-\delta d}}{\int_0^\infty \prod_{i=1}^{n}\Big(\frac{\delta}{(t_i)^{\delta+1}}\Big)^{d_i}\Big(\frac{1}{t_i}\Big)^{\delta(1-d_i)}\delta^{c-1}e^{-\delta d}d\delta} \tag{3.3}$$

Numerical integration procedure can be used to compute, some approximation . One of the most widely used numerical techniques in Bayseian inference is Lindleys approximation proposed in[10] and[6]

$$l_{\delta\delta} \equiv \frac{\partial^2 l(\delta|D)}{\partial \delta^2} = -\frac{\sum_{i=1}^{n} d_i}{\delta^2} \tag{3.4}$$

$$l_{\delta\delta\delta} \equiv \frac{\partial^3 l(\delta|D)}{\partial \delta^3} = \frac{2\sum_{i=1}^{n} d_i}{\delta^3} \tag{3.5}$$

$$\begin{aligned}
v(\delta) =& \delta^{c-1} e^{-\delta d} \\
p =& \log(v(\delta)) \\
=& (c-1)\log\delta - \delta d \\
p_1 =& \frac{\partial p}{\partial \delta} = \frac{c-1}{\delta} - d \\
u =& \delta, u_1 = 1, u_2 = 0 \\
\xi^2 =& \Big(- \frac{\partial^3 l(\delta|D)}{\partial \delta^3} \Big)^{-1} \\
\hat{\delta} =& [u + \frac{1}{2}(u_2 + 2u_1 p_1)\xi^2 + \frac{\xi^4}{2}\frac{\partial^3 l(\delta|D)}{\partial \delta^3} u_1] \\
=& [\delta + (\frac{a-1}{\delta} - b)\xi^2 + \frac{\xi^4}{2}\frac{\partial^3 l(\delta|D)}{\partial \delta^3}]
\end{aligned}$$

3.1. simulation

In this section, we perform a Monte Carlo simulation to compare the estimated performance obtained In the simulation,we study the following

- Generate $u \sim U(0,1)$.
- The failure times T were generated by $T = (\frac{1}{1-u_1})^{1/\delta}$.
- The true values ,δ were taken to be 0.4,and 0.8.
- The sample sizes were taken to be $n = 100$ and 200.
- The results include percentagesof right -censored observations the estimated bias (Bias)given by

$$Bias(\delta) = \frac{1}{k}\sum_{i=1}^{k}(\delta_i - \delta_0)$$

the sample standard error (SSE)given by

$$SSE(\delta) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\delta_i - \bar{\delta})^2}$$

the average of the estimated standard errors (ESE)using observed information matrix the empirical coverage probability(CP).
where δ_0 is the true value of δ,$\bar{\delta}$ is the mean of the estimates ofδ and k is the number of replications.

n	δ	Bias	SSE	ESE	CP
100	0.4	-0.055	0.317	0.433	0.956
100	0.8	-0.085	0.391	0.282	0.961
200	0.4	0.043	0.311	0.272	0.952
200	0.8	-0.087	0.215	0.223	0.956

3.2. Conclusion

In this paper, we examine both classical and Bayesian analysis of survival time observations on right censoring when the lifetime of the items follows a Pareto distribution. The MLEs have been calculated and they are working well, as we used the Lindley approximation, to calculate the Bayes estimates we have considered gamma prior on the parameterδ.The results include percentages of right-censored observations the estimated bias ,the sample standard error (SSE) and the estimated standard errors (ESE) where as the difference between the true value and the estimated value is very small, so the results are good.

References

[1] B.C. Arnold and S.J. Press, *Bayesian estimation and prediction for pareto data*, J. Amer. Stat. Assoc. 84(408) (1989) 1079–1084.

[2] L. Briliano, S. Abdullah and I. Fithriani, *Parameter estimation of burr distribution for right censored data with bayes method*, AIP Conf. Proc. 2242 (2020) 030004.

[3] J. Cahoon and R. Martin, *Generalized inferential models for censored data*, Int. J. Appr. Reas. 137 (2021) 51–66.

[4] S. Datta, *Estimating the mean life time using right censored data*, Stat. Method. 2(1) (2005) 65–69.

[5] F. Ducros and P. Pamphile, *Bayesian estimation of weibull mixture in heavily censored data setting*, Reliability Engin. Syst. Safety, 180 (2018) 453–462.

[6] W.K. Ghafil, T. J. Khraibet and A.A. Alwan, *Maximum likelihood and bayesian estimation of rayleigh with partly interval-censored case-i data*, Neuro Quantol. 18(5) (2020) 26–28.

[7] W.K. Ghafil and R.H. Shamkhi, *Parametric regression analysis of bivariate the proportional hazards model with current status data*, Int. J. Nonlinear Anal. Appl. 12(2)(2021) 1591–1598.

[8] J.H. Kim, S. Ahn and S. Ahn, *Parameter estimation of the pareto distribution using a pivotal quantity*, J. Korean Stat. Soc. 46(3) (2017) 438–450.

[9] N. Kim, *On the maximum likelihood estimation for a normal distribution under random censoring*, Commun. Stat. Appl. Meth. 25(6) (2018) 647–658.

[10] A. Lavanya and T.L. Alexander, *Estimation of parameters using lindley's method*, Int. J. Adv. Res. 4(12) (2016) 1767–1778.

[11] R.E. Quandt, *Old and mew methods op estimation and the pareto distribution*, Metrika 10 (1964) 55-–82.

[12] A. Rasheed and A. Al-Gazi, *Bayes estimators for the shape parameter of pareto type ii distribution under generalized square error loss function*, Math. Theory Model. 6(11) (2014) 20–32.

[13] M. Rytgaard, *Estimation in the pareto distribution*, ASTIN Bulletin: The Journal of the IAA 20(2) (1990) 201–216.

[14] J.I. Seo, S.B. Kang and H.Y. Kim, *New approach for analysis of progressive type-II censored data from the pareto distribution*, Commun. Stat. Appl. Meth. 25(5)(2018) 569–575.

[15] R. Shinohara, *Estimation of Survival of Left Truncated and Right Censored Data Under Increasing Hazard*, Master of Science, Mc Gill University, Montreal, Quebec, 2007.

[16] K.S. Sultan and W. Emam, *The combined-unified hybrid censored samples from pareto distribution: Estimation and properties*, Appl. Sci. 11(13) (2021) 6000.

[17] X. Wang and W. Gui, *Bayesian estimation of entropy for burr type xii distribution under progressive type-ii censored data*, Math. 9(4)(2021) 313.