Bayes Estimator as a Function of Some Classical Estimator

S. H. A, Al-Jasim
Professor of Mathematical Statistics
Department of Statistics– Baghdad University

Dr. Mohammed S. Abd-alrazak
Assistant Professor department of statistics– Baghdad University

Abstract

Maximum likelihood estimation method, uniformly minimum variance unbiased estimation method and minimum mean square error estimation, as classical estimation procedures, are frequently used for parameter estimation in statistics, which assuming the parameter is constant , while Bayes method assuming the parameter is random variable and hence the Bayes estimator is an estimator which minimize the Bayes risk for each value the random observable and for square error lose function the Bayes estimator is the posterior mean. It is well known that the Bayesian estimation is hardly used as a parameter estimation technique due to some difficulties to finding a prior distribution.

The interest of this paper is that whether above classical estimators of the parameter for a particular probability distribution can be obtained from Bayes estimator is determined. In this analysis one-parameter Pareto distribution is used to examine the relationship between Bayesian and classical estimators. Considering improper prior distribution for shape parameter of the Pareto distribution of the first kind with known scale parameter which equals one, we have tried to show how the classical estimators can be obtain from Bayes estimator for various choices of hyper parameters of the prior function.

1. Introduction

Bayesian theory and Bayesian probability are named after *Thomas Bayes* (1702 – 1761), who proved a special case of what is now Bayes formula. The Bayesian method in estimation, came in to use only around 1950 and in decision theory and estimation theory, a Bayes estimator is an estimator or decision rule that maximize the posterior expected value of utility function or minimize the posterior expected value of a loss function [1].

Classical statistics, originating with *R. A. Fisher, J. Neyman* and *E. S. Pearson*, this include the technique of point estimation on this bais that the parameter assumed constant. Most well known classical parameter estimation procedures are the maximum likelihood estimation, minimum variance unbiased estimation and minimum mean square error estimation [2]

The Pareto distribution is named after an Italian-born Swiss professor of economics *Vilfredo Pareto* (1848 – 1923), this distribution is first used as a model for the distribution of incomes a model for city population within a given area [4], failure model in reliability theory [3] and a queuing model in operation research [5].

The Pareto distribution is easy to manipulate analytically and provides a good starting point for discussions of more general distribution. For the more it's analytical tractability allows exploration of the relationships between classical and Bayesian estimators.

In this paper we consider the problem of estimating the shape parameter of Pareto distribution using both classical and Bayesian approach. Bayesian estimator derived from posterior distribution has been used to derive the three classical estimators.

The main object of this paper is to examine the classical estimators can be obtained from various choices made within a Bayesian framework for the Pareto distribution for different values of the pair of hyper parameters of the prior function, the Bayes estimator provides three classical estimators.

2. Some classical estimation methods

Consider a random sample of independent observations $x_1, x_2, \ldots, x_n$ from Pareto distribution with density function [4]:

$$f(x,\alpha) = \frac{\alpha}{x^{\alpha+1}} \qquad 1 \le x \text{ and } \alpha > 0 \quad \text{----------}(1)$$

The likelihood function is:

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

$$= \alpha^n \prod_{i=1}^{n} x_i^{-(\alpha+1)}$$

$$\Rightarrow L(\alpha) = \alpha^n \exp\{-(\alpha+1)T\} \quad ----------(2)$$

where $T = \sum_{i=1}^{n} \ln x_i$

$$\therefore \log L(\alpha) = n \log \alpha - (\alpha+1)T$$

$$\frac{\partial \log L(\alpha)}{\partial \alpha} = \frac{n}{\alpha} - T$$

$$\frac{n}{\hat{\alpha}} - T = 0$$

$$\Rightarrow \hat{\alpha} = \frac{n}{T} = \frac{n}{\sum_{i=1}^{n} \ln x_i} \qquad ----------(3)$$

Which is the maximum likelihood estimator (MLE) for *α*

Let $X \sim Pareto(\acute{a})$, **then** $T_i = \ln x_i,\ i = 1,2,\dots\dots\dots,n$ **having exponential distribution*, with parameter *α* and hence** $T = \sum_{i=1}^{n} t_i = \sum_{i=1}^{n} \ln x_i$ **having Gamma distribution with parameters *n* and *α* and the p.d.f for *T* is given by:**

$$f(t) = \frac{\alpha^n}{\Gamma(n)} t^{n-1} e^{-\alpha t}, t \geq 0 \qquad ----------(4)$$

To obtain uniformly minimum variance unbiased estimator (UMVUE) for *α* by *Lehmann Scheffe*, let us set the density function to exponential family as

$$f(x, \alpha) = \alpha^n - \exp[-(\alpha+1)T] = \exp[P(\alpha)K(\alpha) + S(\alpha) + Q(\alpha)]$$

Which provides a complete sufficient statistic [2]

*

$$Y = \ln x \Rightarrow x = e^{Y}$$

$$|J| = \left|\frac{dx}{dy}\right| = e^{Y}$$

$$g(Y) = f(e^{Y})|J| = \frac{\alpha}{(e^{Y})^{\alpha+1}} \bullet e^{Y} = \alpha e^{-\alpha Y}, Y > 0$$

$$\sum_{i=1}^{n} K(x_i) = \sum_{i=1}^{n} \ln x_i = T \qquad for\ \alpha$$

$$Now \quad E\left(\frac{n}{T}\right) = n\int_{0}^{\infty} \frac{\alpha^{n} t^{n-2} e^{-\alpha t}}{\Gamma(n)} dt$$

$$= \frac{n\alpha}{n-1}\int_{0}^{\infty} \frac{\alpha^{n-1} t^{n-2} e^{-\alpha t}}{\Gamma(n-1)} dt$$

$$= \frac{n\alpha}{n-1} \qquad ----------(5)$$

$$\therefore \quad E\left[\frac{n-1}{T}\right] = E\left[\frac{n-1}{n} \bullet \frac{n}{T}\right]$$

$$= \frac{n-1}{n} E\left[\frac{n}{T}\right]$$

$$= \alpha$$

By *Lehmann Scheffe* approach $\alpha^{*} = \dfrac{n-1}{T} = \dfrac{n-1}{\sum_{i=1}^{n} \ln x_i}$ ----------(6)

is the uniformly minimum variance unbiased estimator for α

Let us now consider the form of mean square error $\dfrac{c}{\sum_{i=1}^{n} \ln x_i} = \dfrac{c}{T}$ **. Now minimizing the mean square error**

$$E\left[\left(\frac{c}{T} - \alpha\right)^{2}\right] = c^{2} E\left(\frac{1}{T^{2}}\right) - 2c\alpha\, E\left(\frac{1}{T}\right) + \alpha^{2}$$

w.r.t *c* we get *c=n-2* and hence

$$\tilde{\alpha} = \frac{n-2}{T} = \frac{n-2}{\sum_{i=1}^{n} \ln x_i} \quad ----------(7)$$

Is the minimum mean square error estimator (MMSEE) for *α*

3. Bayes method

To obtain a Bayes estimator for *α*, we will consider the improper prior distribution for *α* of the form

$$\prod(\alpha) = \alpha^{a-1} e^{-b\alpha}; \ \alpha > 0, -\infty < a < \infty, b \geq o \ ----------(8)$$

The prior distribution is the kernel of a gamma distribution when $a \geq 0$**. Now, the posterior p.d.f for random parameter *α* is given by**

$$h(\alpha | x_{1,........,} x_n) = \frac{L(x_{1,........,} x_n | \alpha) \prod(\alpha)}{\int_0^\infty L(x_{1,........,} x_n | \alpha) \prod(\alpha) d\alpha}$$

$$= \frac{\alpha^n \exp[-(\alpha+1)T] \bullet \alpha^{n-1} e^{-b\alpha}}{\int_0^\infty \alpha^{a+n-1} \exp[-(\alpha+1)T - b\alpha] d\alpha}$$

$$= \frac{\alpha^{a+n-1} \exp[-(b+T)\alpha]}{\int_0^\infty \alpha^{a+n-1} \exp[-(b+T)\alpha] d\alpha}$$

$$putting \quad (b+T)\alpha = \lambda, we\ get$$

$$h(\alpha | x_{1,........,} x_n) = \frac{(b+T)^{a+n}}{\Gamma(a+n)} \alpha^{a+n-1} e^{-(b+T)\alpha} \quad ----------(9)$$

$$and\ therefore \quad (\theta | x_{1,........,} x_n) \sim gamma\left(n+a, \frac{1}{b+T}\right)$$

Using squared error lose function $L(\hat{\alpha},\alpha)=(\hat{\alpha},\alpha)^2$ the Bayes estimator for *α* is simply the posterior mean [4] and hence

$$\hat{\theta}_B = E\left(\alpha \middle| x_1,\ldots\ldots,x_n\right)$$

$$=\frac{n+a}{b+T}$$

$$\Rightarrow \hat{\theta}_B = \frac{n+a}{b+\sum_{i=1}^{n}\ln x_i} \qquad \text{----------(10)}$$

4. Conclusion

In this paper we found that (MMSEE) can be explored from Bayes estimator when *a=-2, b=0* which suggests that if prior function for *α* is $\prod(\alpha)= 1/\alpha^3$, the posterior mean coincides with (MMSEE) when *a=-1, b=0* ; the Bayes estimator provides (UMVUE) and in this case prior function becomes $\prod(\alpha)= 1/\alpha^2$, for the case *a=b=0*, this implies prior function is the *Jeffrey's* prior $\prod(\alpha)= 1/\alpha$,a standard non-informative prior as well as improper prior and the Bayes estimator leads (MLE). Finally, we conclude that the Bayes estimator of reliability function for Pareto distribution is $\hat{R}(t)=t^{-\alpha B}$, because the Bayes estimator $\hat{\alpha}_B$ for α in this case will have the properties of (MLE) specially the invariant property.

5. Reference

1) Askoy, S. (2008) : Bayesian decision theory, Bilkent university, department of computer engineering, Saskoy@Cs.bilkent, tr.Cs. 551

2) Lehmann, E.L.(1991): theory of point estimation, Wadsworth and Brooks pacific groove, California

3) Nadarajah, S. & Kots, S.(2003): Reliability in Pareto models, Metron, International, Journal of statistic, vol. LXI, No.2, pp.191-204

4) Podder, C.K. (2004): Comparison of Two Risk Functions Using Pareto distribution, pak. J. States. Vol. 20(3), pp. 369 – 378

5) Shortle, J & Fischer, M. (2005): Using the Pareto distribution in queuing modeling, Submitted Journal of probability and Satistical Science, pp. 1 - 4