A Comparison of the Bayesian Method under Symmetric and Asymmetric Loss Functions to Estimate the Shape Parameter $K$ of Burr Distribution

I Fithriani$^{1,*}$, A R Hakim$^2$ and M Novita$^3$

1 Department of Mathematics, Universitas Indonesia, Depok 16424, Indonesia
2 Undergraduate Program in Mathematics, Universitas Indonesia, Depok 16424, Indonesia

*Corresponding Author: ida.fithriani@gmail.com

Abstract. Burr distribution is one of the most important types of distribution in Burr system and has gained special attention. It has an important role in various disciplines, such as reliability analysis, life testing, survival analysis, actuarial science, economics, forestry, hydrology and meteorology. Thus, the parameter estimation for Burr distribution becomes an important thing to do. The frequentist approach using the maximum likelihood method is the most commonly used way to estimate the parameters of a distribution. In this paper we considered using the Bayesian method to estimate the shape parameter $k$ of Burr distribution using gamma prior which is a conjugate prior. The Bayes estimate for the shape parameter $k$ is obtained under the squared-error loss function (SELF) which is one of the symmetric loss function and the precautionary loss function (PLF) which is one of the asymmetric loss function. Through a simulation study, the comparison was made on the performance of the Bayes estimate for the shape parameter $k$ under these two loss functions with respect to the mean-squared error (MSE) and the posterior risk.

1. Introduction

Burr distribution was first introduced in 1942 by Irving Wingate Burr. Burr distribution was originally known as Burr Type XII distribution, one of the twelve types of the continuous distributions in Burr system. It was discussed in more detail by Burr [1]. Let $X$ be a random variable having Burr distribution with two positive parameters $k$ and $c$, then it has a distribution function defined by

$$F(x) = 1 - (1 + x^c)^{-k}, \quad x \geq 0.$$  \hspace{1cm} (1)

Therefore, its probability density function is defined by

$$f(x) = dF(x)/dx = kc^x^{c-1}(1 + x^c)^{-k-1}, \quad x > 0.$$  \hspace{1cm} (2)

The parameters $k$ and $c$ of Burr distribution are shape parameters [2].

In the past Burr distribution used in various fields of science including reliability analysis [2,3], life testing [4,5], economics [6], forestry [7], survival analysis [8], actuarial science [9], hydrology [10], and meteorology [11]. To find out more about the characteristics of the data having Burr distribution, the parameters of Burr distribution need to be estimated. The parameter estimation of Burr distribution through the frequentist approach is common. Wingo [5], Wang, Keats and Zimmer [3] and Gupta, Gupta and Lvin [4] considered the problem of estimation using the maximum likelihood method. However, the Bayesian approach received more attention. Papadopoulos [12] considered using the Bayesian method to estimate the parameter $k$ under the squared-error loss function (SELF) when the...
parameter $c$ is known. Evans and Ragab [13] assumed that both $k$ and $c$ are unknown and applied the Bayesian method to estimate them under the squared-error loss function. Moore and Papadopoulos [14] obtained the Bayes estimate for the parameter $k$ under various symmetric loss functions. Al-Noor and Al-Ameer [15] compared the Bayesian method under various symmetric and asymmetric loss functions and the other estimation methods to estimate the parameter $k$ by employing the mean-squared error (MSE) of the estimate obtained.

In this paper we considered comparing the Bayesian method under the squared-error loss function (SELF) which is a symmetric loss function and the precautionary loss function (PLF) which is an asymmetric loss function to estimate the unknown parameters $k$ of Burr distribution when the parameter $c$ is known. We considered using the gamma prior distribution as the prior distribution of the parameter $k$ because it is a conjugate prior, so the posterior distribution obtained is also a member of gamma distribution family. We obtained the Bayes estimate for the parameter $k$ in more detail mathematically. To compare these two loss functions in estimating the parameter $k$, we obtained the posterior risk of the estimate under each loss function. A numerical comparison of the estimates was made through a simulation study in terms of the mean-squared error (MSE) and the posterior risk.

2. The gamma prior distribution of the parameter $k$

We consider using prior distribution which is a conjugate prior distribution. A prior distribution is called a conjugate prior distribution for a parameter if and only if the posterior distribution obtained is from the same distribution family as that prior distribution [16]. One of the prior distributions which is a conjugate prior distribution is the gamma prior distribution. Let $K$ be a random variable having the gamma prior distribution with the prior density function defined as

$$
\pi(k) = \frac{1}{\Gamma(a)\beta^a} k^{a-1}e^{-k/\beta}, \quad k > 0.
$$

Bayesian statisticians frequently write that $\pi(k)$ is proportional to $k^{a-1}e^{-k/\beta}$; that is, $\pi(k) \propto k^{a-1}e^{-k/\beta}$.

We can see that $K$ has the gamma distribution with parameters $a, \beta > 0$. According to Hogg and Craig [17], the parameter estimates for $a$ and $\beta$, by using the method of moment, are

$$
\hat{a} = \frac{\sum X_i}{n} \quad \text{and} \quad \hat{\beta} = \frac{S^2}{\sum (X_i - \bar{X})^2 / n},
$$

respectively, where $\bar{X} = \sum X_i / n$ and $S^2 = \sum (X_i - \bar{X})^2 / n$.

3. The posterior distribution of the parameter $k$

Let $X_1, X_2, \ldots, X_n$ denote a random sample from Burr distribution with unknown parameter $k$, where the parameter $c$ is known, then the joint probability density function, or likelihood function, of $X_1, X_2, \ldots, X_n$ given $K = k$, denoted by $L(x_1, x_2, \ldots, x_n|k)$, is

$$
L(x_1, x_2, \ldots, x_n|k) = \prod_{i=1}^{n} f(x_i|k) = k^a e^{\frac{x_i}{\beta}} \left( \prod_{i=1}^{n} \frac{1}{1+x_i^\beta} \right) \left( \prod_{i=1}^{n} (1+x_i^\beta) \right)^{-k}. \tag{5}
$$

From equation (3) and (5), the joint probability density function of $X_1, X_2, \ldots, X_n$ and $K$ is

$$
L(x_1, x_2, \ldots, x_n|k)\pi(k) = \frac{1}{\Gamma(a)\beta^a} k^{a+n-1}e^{-k/\beta} \left( \prod_{i=1}^{n} \frac{x_i^{\alpha-1}}{1+x_i^\beta} \right) \exp \left\{ -\frac{1}{\beta} + \sum_{i=1}^{n} \ln(1+x_i^\beta) \right\} \pi(k). \tag{6}
$$

Therefore, the posterior density function of the parameter $k$, denoted by $\pi(k|x_1, x_2, \ldots, x_n)$, is

$$
\pi(k|x_1, x_2, \ldots, x_n) = \frac{L(x_1, x_2, \ldots, x_n|k)\pi(k)}{\int_0^{\infty} L(x_1, x_2, \ldots, x_n|k)\pi(k)dk}. \tag{7}
$$

By substituting equation (6) to equation (7), we obtain
\[ \pi(k | x_1, x_2, \ldots, x_n) = \frac{1}{\Gamma(a+n)} \left[ \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1+x_i^+) \right]^{-a+n} k^{a+n-1} \exp \left\{ -\left( \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1+x_i^+) \right) k \right\}; \]  

that is, it is the probability density function of the gamma distribution with the parameters \( \alpha^* = \alpha + n \) and \( \beta^* = \left( \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1+x_i^+) \right)^{-1} \). This shows that the gamma prior distribution is a conjugate prior distribution for the parameter \( k \).

4. The Bayes estimate for the parameter \( k \)

To find the Bayes estimate for the parameter \( k \), we need a loss function denoted by \( \mathcal{L}(k, \hat{k}) \). The Bayes estimate is obtained by finding \( \hat{k} \) minimizing the posterior risk defined by

\[ R(\hat{k}) = E[\mathcal{L}(K, \hat{k}) | x_1, x_2, \ldots, x_n]. \]  

In this paper we consider using two different loss functions, i.e. the squared-error loss function (SELF) which is a symmetric loss function and the precautionary loss function (PLF) which is an asymmetric loss function. We use the notation \( \hat{k}_{\text{SELF}} \) to denote the Bayes estimate for the parameter \( k \) under the SELF and the notation \( \hat{k}_{\text{PLF}} \) to denote the Bayes estimate for the parameter \( k \) under the PLF.

4.1. The Bayes estimate for the parameter \( k \) under the SELF

In this case the squared-error loss function (SELF) is defined by

\[ \mathcal{L}(k, \hat{k}_{\text{SELF}}) = (k - \hat{k}_{\text{SELF}})^2. \]  

Then, from equation (10), we obtain the posterior risk

\[ R(\hat{k}_{\text{SELF}}) = E[(K - \hat{k}_{\text{SELF}})^2 | x_1, x_2, \ldots, x_n] \]  

and \( \hat{k}_{\text{SELF}} \) minimizing the posterior risk \( R(\hat{k}_{\text{SELF}}) \) is

\[ \hat{k}_{\text{SELF}} = E(K | x_1, x_2, \ldots, x_n); \]  

that is, \( \hat{k}_{\text{SELF}} \) is the mean, or the first moment, of the posterior distribution of the parameter \( k \) [14]. Therefore, we obtain \( \hat{k}_{\text{SELF}} = \alpha^* \beta^* \); that is,

\[ \hat{k}_{\text{SELF}} = (\alpha + n) \left( \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1+x_i^+) \right)^{-1}. \]  

4.2. The Bayes estimate for the parameter \( k \) under the PLF

In this case the precautionary loss function (PLF) is defined by

\[ \mathcal{L}(k, \hat{k}_{\text{PLF}}) = (k - \hat{k}_{\text{PLF}})^2 / \hat{k}_{\text{PLF}}. \]  

Then, from equation (10), we obtain the posterior risk

\[ R(\hat{k}_{\text{PLF}}) = E[(K - \hat{k}_{\text{PLF}})^2 / \hat{k}_{\text{PLF}} | x_1, x_2, \ldots, x_n] \]  

and \( \hat{k}_{\text{PLF}} \) minimizing the posterior risk \( R(\hat{k}_{\text{PLF}}) \) is

\[ \hat{k}_{\text{PLF}} = [E(K^2 | x_1, x_2, \ldots, x_n)]^{1/2}; \]  

that is, \( \hat{k}_{\text{PLF}} \) is the square root of the second moment of the posterior distribution of the parameter \( k \) [18]. Therefore, we obtain \( \hat{k}_{\text{PLF}} = \beta^* \sqrt{(\alpha^* + 1) \alpha^*} \); that is,
\[
\hat{k}_{PLF} = \sqrt{(\alpha + n + 1)(\alpha + n)} \left[ \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1 + x_i^k) \right]^{-1}.
\] (18)

5. The posterior risk of the Bayes estimate for the parameter \( k \)

5.1. The posterior risk of the Bayes estimate for the parameter \( k \) under the SELF

The posterior risk in equation (12) can be written as

\[
R(\hat{k}_{SELF}) = (\hat{k}_{PLF})^2 - (\hat{k}_{SELF})^2.
\] (19)

By substituting equation (14) and (18) to equation (19), we obtain

\[
R(\hat{k}_{SELF}) = (\alpha + n) \left[ \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1 + x_i^k) \right]^{-2}.
\] (20)

5.2. The posterior risk of the Bayes estimate for the parameter \( k \) under the PLF

The posterior risk \( R(\hat{k}_{PLF}) \) in equation (16) can be written as

\[
R(\hat{k}_{PLF}) = 2(\hat{k}_{PLF} - \hat{k}_{SELF}).
\] (21)

By substituting equation (14) and (18) to equation (21), we obtain

\[
R(\hat{k}_{PLF}) = 2 \left[ \sqrt{(\alpha + n + 1)(\alpha + n)} - (\alpha + n) \right] \left[ \frac{1}{\beta} + \sum_{i=1}^{n} \ln(1 + x_i^k) \right]^{-1}.
\] (22)

5.3. The comparison of the posterior risk of the Bayes estimate for the parameter \( k \) under the SELF and the PLF

Based on the posterior risk in equation (19) and (21), we obtain the ratio

\[
\frac{R(\hat{k}_{SELF})}{R(\hat{k}_{PLF})} = \frac{(\hat{k}_{PLF})^2 - (\hat{k}_{SELF})^2}{2(\hat{k}_{PLF} - \hat{k}_{SELF})} = \frac{\hat{k}_{SELF} + \hat{k}_{PLF}}{2}.
\] (23)

From equation (23), we obtain that

- if \( \hat{k}_{SELF} + \hat{k}_{PLF} < 2 \), we get \( R(\hat{k}_{SELF}) < R(\hat{k}_{PLF}) \); that is, the Bayesian estimation for the parameter \( k \) under the SELF gives better result than under the PLF;
- if \( \hat{k}_{SELF} + \hat{k}_{PLF} = 2 \), we get \( R(\hat{k}_{SELF}) = R(\hat{k}_{PLF}) \); that is, the Bayesian estimation for the parameter \( k \) under the SELF gives result as good as under the PLF; and
- if \( \hat{k}_{SELF} + \hat{k}_{PLF} > 2 \), we get \( R(\hat{k}_{SELF}) > R(\hat{k}_{PLF}) \); that is, the Bayesian estimation for the parameter \( k \) under the PLF gives better result than under the SELF.

6. Simulation Study and Results

In this section we perform a simulation study to compare the performance of each estimate of the parameter \( k \) numerically in terms of the mean-squared error (MSE) and the posterior risk. The simulation study is made by the following steps.

- We determine the values of the shape parameters \( k \) and \( c \) of Burr distribution. We consider \( c = 0.5, 2.0 \) and \( k = 0.1, 0.5, 1.0, 1.5, 2.0 \).
- We consider using the sample size \( n = 10, 50 \) and 100.
- We determine the number of replication, i.e. \( L = 1000 \).
- We generate the random samples of size \( n \) from a uniform distribution \( (U) \) over the interval \((0,1)\), say \( u_1, u_2, \ldots, u_n \). Then, we convert them to the samples having Burr distribution with the parameters \( k \) and \( c \) through the inverse of the distribution function in equation (1), i.e.

\[
x_i = [(1 - u_i)^{-1/k} - 1]^{1/c}, \quad i = 1, 2, \ldots, n.
\] (24)
• By assuming that the parameter $c$ is known, we find the Bayes estimate for the parameter $k$ under the SELF and the PLF according to equation (14) and (18), where the parameters $\alpha$ and $\beta$ of the gamma prior distribution are estimated by $\hat{\alpha}$ and $\hat{\beta}$ in equation (4).

• We make a comparison by calculating the MSE of the estimate according to equation (25)

$$\text{MSE}(k) = \frac{1}{L} \sum_{s=1}^{L} (\hat{k}_s - k)^2,$$

where $\hat{k}_s$ is the estimate of the parameter $k$ at the $s$-th replication.

• In addition to using the MSE, we also make a comparison by calculating the posterior risk according to equation (20) and (22), or equation (19) and (21).

The results of the simulation study are summarized and tabulated in Table 1 below.

| $c$ | $k$ | $n$ | MSE($\hat{k}_{\text{SELF}}$) | MSE($\hat{k}_{\text{PLF}}$) | $R(\hat{k}_{\text{SELF}})$ | $R(\hat{k}_{\text{PLF}})$ |
|----|----|----|----------------|----------------|----------------|----------------|
| 0.1 | 10 | 0.0015572 | 0.008519 | 0.0013717 | 0.0107700 |
| 50 | 0.0002100 | 0.0002185 | 0.001035 | 0.0020218 |
| 100 | 0.0001104 | 0.0001130 | 0.0010093 |
| 0.5 | 10 | 0.0398777 | 0.0471927 | 0.0340126 | 0.0533953 |
| 50 | 0.0052962 | 0.0055429 | 0.0053338 | 0.0101740 |
| 100 | 0.0025734 | 0.0026272 | 0.0050303 |
| 1.0 | 10 | 0.1565890 | 0.1877870 | 0.1393640 | 0.1081660 |
| 50 | 0.0216592 | 0.0225742 | 0.0211730 | 0.0202580 |
| 100 | 0.0105273 | 0.0107741 | 0.0100883 |
| 1.5 | 10 | 0.1835370 | 0.2156420 | 0.1478210 |
| 50 | 0.0516081 | 0.0539030 | 0.0480314 | 0.0304910 |
| 100 | 0.0243331 | 0.0248663 | 0.0151107 |
| 2.0 | 10 | 0.2421290 | 0.2722450 | 0.2538360 | 0.1478210 |
| 50 | 0.0882654 | 0.0922875 | 0.0853002 | 0.0304910 |
| 100 | 0.0372423 | 0.0378557 | 0.0406266 | 0.0200666 |
| 0.1 | 10 | 0.0018893 | 0.0022305 | 0.0014299 | 0.0108862 |
| 50 | 0.0002288 | 0.0002387 | 0.0001231 | 0.0020321 |
| 100 | 0.0001031 | 0.0001053 | 0.0001053 | 0.0020321 |
| 0.5 | 10 | 0.0451579 | 0.0533056 | 0.0341517 | 0.0520079 |
| 50 | 0.0063109 | 0.0066031 | 0.0053816 | 0.0101787 |
| 100 | 0.0025913 | 0.0026252 | 0.0025797 | 0.0050371 |
| 1.0 | 10 | 0.0649506 | 0.0788286 | 0.1028850 | 0.0889534 |
| 50 | 0.0213653 | 0.0223466 | 0.0209623 | 0.0199810 |
| 100 | 0.0104318 | 0.0107050 | 0.0103143 | 0.0100411 |
| 1.5 | 10 | 0.0667565 | 0.0567915 | 0.1485480 | 0.1056750 |
| 50 | 0.0370200 | 0.0386422 | 0.0431799 | 0.0289185 |
| 100 | 0.0217145 | 0.0214849 | 0.0224884 | 0.0147853 |
| 2.0 | 10 | 0.3449830 | 0.2932900 | 0.1730210 | 0.1123570 |
| 50 | 0.0566033 | 0.0544790 | 0.0708417 | 0.0364830 |
| 100 | 0.0314448 | 0.0310473 | 0.0379675 | 0.0191825 |

Based on the tabulated MSE values, the larger sample size we use, the smaller MSE value we obtain, under the SELF as well as under the PLF. We also obtain that under these loss functions, when the true values of the parameter $k$ are smaller, the MSE values are also smaller. In general, the estimates for the parameter $k$ under the SELF have smaller MSE values than under the PLF.
Based on the tabulated posterior risks, the larger sample size we use, the smaller posterior risk we obtain, under the SELF as well as under the PLF. We also obtain that under these loss functions, when the true values of the parameter $k$ are smaller, in general, the posterior risks are also smaller. In general, the estimates for the parameter $k$ under the SELF have smaller posterior risks than under the PLF when the true values of the parameter $k$ are less than 1. Otherwise, when the true values of the parameter $k$ are more than or equal to 1, the estimates under the PLF have the smaller posterior risks.

7. Conclusion
The Bayesian estimation for the shape parameter $k$ of Burr distribution under the squared-error loss function (SELF), which is a symmetric loss function, is found to be superior compared to the precautionary loss function (PLF), which is an asymmetric loss function, in terms of the mean-squared error (MSE) for all sample sizes and all true values of the parameter $k$. However, in terms of the posterior risk, the estimation under the SELF is found to be superior in turn with the PLF; that is, the SELF gives better results than the PLF when the true values of the parameter $k$ are small (less than 1).

8. References
[1] Burr I W 1942 Cumulative frequency functions Annals of Mathematical Statistics 13 215–32
[2] Lewis A W 1981 The Burr distribution as a general parametric family in survivalship and reliability theory applications Ph.D. Thesis Department of Biostatistics University of North Carolina at Chapel Hill
[3] Wang F K, Keats J B and Zimmer W J 1996 Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data Microelectron. Reliab. 36 pp. 359–62
[4] Gupta P L, Gupta R C and Lvin S J 1996 Analysis of failure time data by Burr distribution Communications in Statistics–Theory and Methods 25 pp. 2013–24
[5] Wingo D R 1983 Maximum likelihood methods for fitting the Burr XII distribution to life test data Biometrical Journal 25 77–84
[6] McDonald J B 1984 Some generalized functions for the size distribution of income Econometrica 52 647–63
[7] Lindsay S R, Wood G R and Woollons R C 1996 Modelling the diameter distribution of forest stands using the Burr distribution Journal of Applied Statistics 23 609–20
[8] Githany M E and Al-Awdahi S 2002 Maximum likelihood estimation of Burr XII distribution parameters under random censoring Journal of Applied Statistics 29 955–65
[9] Burnecki K, Härdle W and Weron R 2004 An introduction to simulation of risk processes Encyclopaedia of Actuarial Science 1–7
[10] Shao Q, Wong H, Xia J and Ip W 2004 Models for extremes using the extended three parameter Burr XII system with application to flood frequency analysis Hydrological Sciences Journal 49 685–702
[11] Usta I and Kantar Y M 2013 Analysis of the Burr Type XII distribution for estimation of wind speed distributions Proc. on Science and Engineering (Thailand: I-SEEC) p 302–8
[12] Papadopoulos A S 1978 The Burr distribution as a failure model from a Bayesian approach IEEE Transactions on Reliability 27 369–71
[13] Evans I G and Ragab A S 1983 Bayesian inferences given a type-2 censored sample from Burr distribution Communications in Statistics–Theory and Methods 12 1569–80
[14] Moore D and Papadopoulos A S 2000 The Burr Type XII distribution as a failure model under various loss functions Microelectronics Reliability 40 2117–22
[15] Al-Noor N H and Al-Ameer H A A 2014 Some estimation methods for the shape parameter and reliability function of Burr Type XII distribution / comparison study Mathematical Theory and Modeling 4 63–77
[16] Klugman S A, Panjer H H and Willmot G E 2012 Loss models from data to decisions 4th ed (Canada: John Wiley & Sons, Inc.)
[17] Hogg R V and Craig A T 1995 Introduction to mathematical statistics 5th ed (United States of
Acknowledgments

This work is supported by Hibah PITTA 2018 funded by DRPM Universitas Indonesia No.2308/UN2.R3.1/HKP.05.00/2018.

[18] Norstrom J G 1996 The use of precautionary loss function in risk analysis IEEE Transactions on Reliability 45 400–3