Using Jeffrey prior information to estimate the shape parameter $k$ of Burr distribution

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

$^{1,2,3}$ Department of Mathematics, Universitas Indonesia, Depok 16424, Indonesia

E-mail: $^a$ariefr.rahman41@sci.ui.ac.id, $^b$mila.novita@sci.ui.ac.id, $^c$ida.fithriani@gmail.com

Abstract. Burr distribution with two parameters was first introduced by Burr. This distribution has gained special attention and has been applied in various disciplines. The maximum likelihood method is the most commonly used to estimate its parameters. However, the Bayesian method receives more attention. The parameter estimation using the Bayesian method not only uses the information from the sample data but also combines it with the prior information for the parameter. Jeffrey prior information is one of the prior information we can use. This prior is a noninformative prior. It is proportional to the square root of the Fisher information for the parameter. In this paper we use Jeffrey prior information to estimate the shape parameter $k$ of Burr distribution. As a comparison, we also use an extension of Jeffrey prior information which is proportional to the Fisher information raised by a positive constant. The comparison is made through a simulation with respect to the mean-squared error (MSE) and the posterior risk. The results of the comparison show that the Bayesian estimation for the shape parameter $k$ under Jeffrey prior information gives better results in turn with the extended Jeffrey prior information.

1. Introduction

In 1942 Irving Wingate Burr introduced Burr system consisting of twelve types of continuous distributions [1]. One of them is Burr Type XII distribution known as Burr distribution. A random variable $X$ having Burr distribution with two parameters $k > 0$ and $c > 0$ has a distribution function defined by

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

Therefore, its probability density function is defined by

$$f(x) = \frac{dF(x)}{dx} = kcx^{-1}(1 + x^c)^{-k-1}, \quad x > 0. \quad (2)$$

The parameters $k$ and $c$ are shape parameters. Therefore, its probability density function can be either decreasing or unimodal [2].

Burr distribution had been discussed in more detail by several authors due to its importance in various disciplines, especially in reliability and failure time analyses. The discussion on these disciplines was made by Lewis [2], Wingo [3], Wang, Keats and Zimmer [4] and Gupta, Gupta and Lvin [5].

In recent years the parameter estimation for Burr distribution has been performed by several authors through a frequentist approach as well as a Bayesian approach, but the Bayesian approach gains more attention. In the Bayesian approach the main problem is how to determine a prior information for the parameter. The determination of the prior information is subjective. When we
have no strong belief in determining the prior because of no (or minimal) information about the parameter, we need a noninformative prior. One of the noninformative prior we can use is a uniform prior introduced by Laplace in 1812, where all possible values of the parameter are given equal weight. However, the uniform prior does not satisfy the invariance property under a one-to-one transformation. Alternatively, we can use the other noninformative prior which has been used more widely, i.e. Jeffrey prior introduced by Jeffreys in 1961. Jeffrey prior is proportional to the square root of the Fisher information for the parameter in a random sample. This prior satisfies the invariance property under a one-to-one transformation [6].

There are several authors who had performed the Bayesian estimation under Jeffrey prior information, as well as the uniform prior information. Papadopoulos [7] used the Bayesian method under the uniform prior, compared to the gamma prior which is a conjugate prior, to estimate the parameter k of Burr distribution. Moore and Papadopoulos [8] obtained the estimator for the parameter k under Jeffrey prior, compared to the gamma prior. In this paper we estimate the parameter k using the Bayesian method under Jeffrey prior compared to an extension of Jeffrey prior. The extended Jeffrey prior is proportional to the Fisher information raised by a positive constant. The comparison of these priors had been made by Nasir and Al-Anber [9] and Al-Noor and Al-Ameer [10] through a simulation with respect to the mean-squared error (MSE) of the estimator. However, in this paper we also make the comparison with respect to the posterior risk of the estimator and make the estimation in more detail. Furthermore, we also present the computational time of the experiment on a simulation.

2. Jeffrey prior and the extended Jeffrey prior information

Let k be a realization of a positive-valued random variable K of the continuous type. Jeffrey prior, denoted by \( \pi_J(k) \), for the parameter k is defined by

\[
\pi_J(k) = \lambda \sqrt{I_n(k)},
\]

(3)

where \( \lambda \) is a positive constant [9]. \( I_n(k) \) in equation (3) is the Fisher information for the parameter k in a random sample \( X_1, X_2, \ldots, X_n \) of size n and defined by

\[
I_n(k) = E\left[ \left( \frac{\partial \ln L(X_1, X_2, \ldots, X_n | k)}{\partial k} \right)^2 \right] \quad \text{or} \quad I_n(k) = -E\left[ \frac{\partial^2 \ln L(X_1, X_2, \ldots, X_n | k)}{\partial k^2} \right],
\]

(4)

where \( L(x_1, x_2, \ldots, x_n | k) \) is the joint probability density function, or the likelihood function, of the random sample \( X_1, X_2, \ldots, X_n \), given \( K = k \) [11]. Bayesian statisticians often write Jeffrey prior \( \pi_j(k) \) is proportional to the square root of the Fisher information \( I_n(k) \); that is,

\[
\pi_j(k) \propto \sqrt{I_n(k)}.
\]

(5)

In this paper we also use the extended Jeffrey prior. It is denoted by \( \pi_{EJ}(k) \) and defined by

\[
\pi_{EJ}(k) = \lambda |I_n(k)|^p,
\]

(6)

where \( p \) is a positive constant [9]. We can write \( \pi_{EJ}(k) \) is proportional to the Fisher information \( I_n(k) \) raised by \( p \); that is,

\[
\pi_{EJ}(k) \propto |I_n(k)|^p.
\]

(7)

We can see that if \( p = 0.5 \), \( \pi_{EJ}(k) \) is equal to Jeffrey prior \( \pi_J(k) \).

3. The Fisher information for the parameter k

Let a random variable \( X \) have Burr distribution with unknown parameter k, where the parameter c is known. The probability density function of \( X \), given \( K = k \), can be denoted by \( f(x | k) \). Let \( X_1, X_2, \ldots, X_n \) denote a random sample of size n from that distribution, then its joint probability density function, or its likelihood function, given \( K = k \), is defined by
\[ L(x_1, x_2, \ldots, x_n|k) = \prod_{i=1}^{n} f(x_i|k) = k^n c^n \left( \prod_{i=1}^{n} \frac{x_i^{n-1}}{1 + x_i^n} \right) \left( \prod_{i=1}^{n} (1 + x_i^n) \right)^{-k}. \] (8)

From \( L(x_1, x_2, \ldots, x_n|k) \) in equation (8), we obtain
\[ \ln L(x_1, x_2, \ldots, x_n|k) = n \ln k + n \ln c + \sum_{i=1}^{n} \ln \left( \frac{x_i^{n-1}}{1 + x_i^n} \right) - k \sum_{i=1}^{n} \ln(1 + x_i^n), \]
then
\[ \frac{\partial \ln L(x_1, x_2, \ldots, x_n|k)}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n} \ln(1 + x_i^n) \quad \text{and} \quad \frac{\partial^2 \ln L(x_1, x_2, \ldots, x_n|k)}{\partial k^2} = -\frac{n}{k^2}. \]

From equation (4), the Fisher information for the parameter \( k \) in the random sample \( X_1, X_2, \ldots, X_n \) is
\[ I_n(k) = -E \left[ \frac{\partial^2 \ln L(X_1, X_2, \ldots, X_n|k)}{\partial k^2} \right] = -E \left( -\frac{n}{k^2} \right) = \frac{n}{k^2}. \] (9)

4. The Bayesian estimation for the parameter \( k \) under Jeffrey prior information

4.1. Jeffrey prior for the parameter \( k \)
By substituting \( I_n(k) \) in equation (9) to equation (3), we obtain Jeffrey prior
\[ \pi_j(k) = \frac{\lambda}{k}, \] (10)
or we can write
\[ \pi_j(k) \propto \frac{1}{k}. \] (11)

4.2. The posterior distribution for the parameter \( k \) under Jeffrey prior
Through the Bayes\' theorem, we combine the prior information with the information from the sample data represented by the likelihood function. Then, the result is expressed as a posterior distribution whose probability density function is denoted by \( \pi(k|x_1, x_2, \ldots, x_n) \) and defined by
\[ \pi(k|x_1, x_2, \ldots, x_n) = \frac{L(x_1, x_2, \ldots, x_n|k) \pi(k)}{\int_0^\infty L(x_1, x_2, \ldots, x_n|k) \pi(k) dk}. \] (12)

From \( \pi_j(k) \) in equation (10) and \( L(x_1, x_2, \ldots, x_n|k) \) in equation (8), we obtain
\[ L(x_1, x_2, \ldots, x_n|k) \pi_j(k) = \lambda k^{n-1} c^n \sqrt{n} \left( \prod_{i=1}^{n} \frac{x_i^{n-1}}{1 + x_i^n} \right) \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^n) \right]. \] (13)

By substituting \( L(x_1, x_2, \ldots, x_n|k) \pi_j(k) \) in equation (13) to equation (12), we obtain
\[ \pi_j(k|x_1, x_2, \ldots, x_n) = \frac{k^{n-1} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^n) \right]}{\int_0^\infty k^{n-1} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^n) \right] dk}. \]

To evaluate the integral form in the denominator, let
\[ u = k \sum_{i=1}^{n} \ln(1 + x_i^n) \quad \text{or} \quad k = \left[ \sum_{i=1}^{n} \ln(1 + x_i^n) \right]^{-1} u, \] (14)
then
\[
dk = \left[ \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right]^{-1} du
\]
and we obtain
\[
\int_{0}^{n} k^{-1} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right] dk = \left[ \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right]^{-1} \Gamma(n).
\]
Therefore, we obtain the posterior density function
\[
\pi_{1}(k|x_{1},x_{2},\ldots,x_{n}) = \frac{1}{\Gamma(n)} \left[ \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right]^{-1} k^{-1} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right].
\]
We can see that the posterior density function \( \pi_{1}(k|x_{1},x_{2},\ldots,x_{n}) \) in equation (15) is the probability density function of the gamma distribution whose parameters are
\[
\alpha_{j} = n \quad \text{and} \quad \beta_{j} = \left[ \sum_{i=1}^{n} \ln(1 + x_{i}^{c}) \right]^{-1}.
\]

4.3. The Bayesian estimation for the parameter \( k \) under Jeffrey prior

After we obtain the posterior distribution, we find the estimator for the parameter \( k \). To find it, we need a loss function \( \mathcal{L}(k,\hat{k}) \) and the posterior risk \( R(\hat{k}) \) defined by
\[
R(\hat{k}) = E \left[ \mathcal{L}(K,\hat{k}) \right| x_{1},x_{2},\ldots,x_{n}].
\]

The Bayes estimator for the parameter \( k \) is obtained by finding \( \hat{k} \) minimizing the posterior risk \( R(\hat{k}) \).

In this paper we consider the most common and easy-to-use loss function, i.e. the squared-error loss function (SELF) defined by
\[
\mathcal{L}(k,\hat{k}) = (k - \hat{k})^{2}.
\]
Then, from equation (17), we obtain the posterior risk
\[
R(\hat{k}) = E \left[ (K - \hat{k})^{2} \right| x_{1},x_{2},\ldots,x_{n}].
\]
Under the SELF, \( \hat{k} \) minimizing the posterior risk \( R(\hat{k}) \) is
\[
\hat{k} = E(K|x_{1},x_{2},\ldots,x_{n});
\]
that is, \( \hat{k} \) is the mean, or the first moment, of the posterior distribution for the parameter \( k \) [7]. Because the posterior distribution is the gamma distribution whose parameters are \( \alpha_{j} \) and \( \beta_{j} \) given in equation (16), we obtain the estimator \( \hat{k}_{j} = \alpha_{j}/\beta_{j} \); that is,
\[
\hat{k}_{j} = \frac{n}{\sum_{i=1}^{n} \ln(1 + x_{i}^{c})}.
\]

4.4. The posterior risk of the Bayes estimator for the parameter \( k \) under Jeffrey prior

From the posterior risk \( R(\hat{k}) \) in equation (19) and the estimator \( \hat{k} \) in equation (20), we can write
\[
R(\hat{k}) = E \left[ (K - E(K|x_{1},x_{2},\ldots,x_{n}))^{2} \right| x_{1},x_{2},\ldots,x_{n}].
\]
and we obtain
\[
R(\hat{k}) = \operatorname{var}(K|x_{1},x_{2},\ldots,x_{n});
\]
that is, the posterior risk under the SELF is the variance of the posterior distribution for the parameter $k$. Because the posterior distribution is the gamma distribution whose parameters are $\alpha$ and $\beta$, given in equation (16), we obtain the posterior risk $R(\hat{k}) = \alpha/\beta$: that is,

$$R(\hat{k}) = \frac{n}{\sum_{i=1}^{n} \ln(1 + x_i^c)}.$$  

(23)

5. The Bayesian estimation for the parameter $k$ under the extended Jeffrey prior information

5.1. The extended Jeffrey prior for the parameter $k$

By substituting $I_n(k)$ in equation (9) to the equation (6), we obtain the extended Jeffrey prior

$$\pi_E(k) \propto \frac{\lambda n^{m}}{k^{2p}},$$  

(24)

or we can write

$$\pi_E(k) \propto \frac{1}{k^{2p}}.$$  

(25)

5.2. The posterior distribution for the parameter $k$ under the extended Jeffrey prior

From $\pi_E(k)$ in equation (24) and $L(x_1, x_2, \ldots, x_n|k)$ in equation (8), we obtain

$$L(x_1, x_2, \ldots, x_n|k)\pi_E(k) = \lambda k^{n-2p} c^{m} n^{p} \left( \prod_{i=1}^{n} \frac{x_i^{c-1}}{1 + x_i^c} \right) \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^c) \right].$$

(26)

By substituting $L(x_1, x_2, \ldots, x_n|k)\pi_E(k)$ in equation (26) to equation (12), we obtain

$$\pi_E(k|x_1, x_2, \ldots, x_n) = \frac{k^{n-2p} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^c) \right]}{\int_{0}^{\infty} k^{n-2p} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^c) \right] dk}.$$

From equation (14), we can evaluate the integral form in the denominator and obtain

$$\int_{0}^{\infty} k^{n-2p} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^c) \right] dk = \left[ \sum_{i=1}^{n} \ln(1 + x_i^c) \right]^{(n-2p+1)} \Gamma(n - 2p + 1).$$

Therefore, we obtain the posterior density function

$$\pi_E(k|x_1, x_2, \ldots, x_n) = \frac{1}{\Gamma(n - 2p + 1)} \left[ \sum_{i=1}^{n} \ln(1 + x_i^c) \right]^{(n-2p+1)} k^{(n-2p+1)-1} \exp \left[ -k \sum_{i=1}^{n} \ln(1 + x_i^c) \right].$$

(27)

We can see that the posterior density function $\pi_E(k|x_1, x_2, \ldots, x_n)$ in equation (27) is the probability density function of the gamma distribution whose parameters are

$$\alpha_E = n - 2p + 1 \quad \text{and} \quad \beta_E = \left[ \sum_{i=1}^{n} \ln(1 + x_i^c) \right]^{-1}.$$  

(28)

5.3. The Bayesian estimation for the parameter $k$ under the extended Jeffrey prior

As the previous discussion, under the SELF, the Bayes estimator for the parameter $k$ is the mean, or the first moment, of the posterior distribution for the parameter $k$. Under the extended Jeffrey prior, the posterior distribution is the gamma distribution whose parameters are $\alpha_E$ and $\beta_E$ given in equation (28). Therefore, we obtain the estimator $\hat{k}_E = \alpha_E/\beta_E$: that is,
\[ \hat{k}_E = \frac{n - 2p + 1}{\sum_{i=1}^{n} \ln(1 + x_i^c)} \]  

(29)

5.4. The posterior risk of the Bayes estimator for the parameter k under the extended Jeffrey prior

As the previous discussion, the posterior risk under the SELF is the variance of the posterior distribution for the parameter k. Because the posterior distribution is the gamma distribution whose parameters are \( \alpha_k \) and \( \beta_k \) given in equation (28), we obtain the posterior risk \( R(\hat{k}_E) = \alpha_k \hat{\beta}_k^2 \); that is,

\[ R(\hat{k}_E) = \frac{n - 2p + 1}{\left[ \sum_{i=1}^{n} \ln(1 + x_i^c) \right]^2}. \]  

(30)

6. Simulation and Results

To compare the performance of the estimator for the parameter k under Jeffrey prior and the extended Jeffrey prior numerically, we perform a simulation. The comparison is made with respect to the mean-squared error (MSE) and the posterior risk of the estimator. The simulation we make follows the steps below.

- We determine the values of the parameters k and c. These values are varied to observe the effect of the parameters on the estimator. We consider determining \( c = 0.5, 2.0 \) and \( k = 0.1, 0.5, 1.0, 1.5, 2.0 \).
- We determine the sample size \( n \). To represent small, medium and large sample size, we consider determining \( n = 10, 50 \) and 100.
- We determine the number of replications, i.e. \( M = 1000 \); that is, the process is repeated 1000 times to obtain 1000 independent random samples of size \( n \).
- We determine the power \( p \) for the extended Jeffrey prior to differentiate it to Jeffrey prior. We consider determining \( p = 0.3 \) which is less than 0.5 and \( p = 0.8 \) which is more than 0.5.
- We generate the random samples of size \( n \) from a uniform distribution \( (U) \) over the interval \((0,1)\); that is, we obtain \( u_1, u_2, \ldots, u_n \). Then, we transform 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.

\[ x_i = [(1-u_i)^{1/k} - 1]^{1/c}, \quad i = 1, 2, \ldots, n. \]  

(31)

- By assuming that the parameter \( c \) is fixed, we find the Bayes estimator for the parameter \( k \) under Jeffrey prior and the extended Jeffrey prior based on each sample of size \( n \); that is, \( \hat{k}_j \) and \( \hat{k}_E \) according to equation (21) and (29).
- After we obtain the Bayes estimator for the parameter \( k \) under each prior, we make a comparison by calculating the MSE according to equation

\[ \text{MSE}(\hat{k}) = \frac{1}{M} \sum_{h=1}^{M} (\hat{k}_h - k)^2, \]  

(32)

where \( \hat{k}_h \) is the estimator for the parameter \( k \) at the \( h \)-th replication, \( h = 1, 2, \ldots, M \). An estimator is said to be the best if its MSE has the smallest value.
- In addition to using the MSE, we also make a comparison by calculating the posterior risk \( R(\hat{k}_j) \) and \( R(\hat{k}_E) \) according to equation (23) and (30). An estimator is said to be the best if its posterior risk has the smallest value.

We summarize and tabulate the results of the simulation in Table 1 below.
Based on the tabulated MSE values as presented in Table 1, the larger sample size we use, the smaller MSE value we obtain, not only the estimation under Jeffrey prior but also under the extended Jeffrey prior. We also obtain that under these priors, for the same sample size, when the true values of the parameter $k$ are smaller, the MSE values are also smaller. In general, the estimators for the parameter $k$ under Jeffrey prior have smaller MSE values than under the extended Jeffrey prior when the power $p$ is 0.3 (less than 0.5). Otherwise, when the power $p$ is 0.8 (more than 0.5), the estimators under the extended Jeffrey prior have smaller MSE values than under Jeffrey prior.

Based on the tabulated posterior risks as presented in Table 1, the larger sample size we use, the smaller posterior risk we obtain, not only the estimation under Jeffrey prior but also under the extended Jeffrey prior. We also obtain that under these priors, for the same sample size, when the true values of the parameter $k$ are smaller, the posterior risks are also smaller. In general, the estimators for the parameter $k$ under Jeffrey prior have smaller posterior risks than under the extended Jeffrey prior when the power $p$ is 0.3 (less than 0.5). Otherwise, when the power $p$ is 0.8 (more than 0.5), the estimators under the extended Jeffrey prior have smaller posterior risks than under Jeffrey prior.

$$
\begin{array}{cccccccccc}
| c | k | n | \text{Jeffrey Prior} & \text{The Extended Jeffrey Prior} & \text{Jeffrey Prior} & \text{The Extended Jeffrey Prior} & \text{Computational Time (in sec)} \\
|---|---|---|------------------|------------------|------------------|------------------|------------------|------------------|
| 0.5 | 0.1 | 10 | 0.001535 | 0.00001773 | 0.001261 | 0.001387 | 0.001442 | 0.001303 | 0.035446 |
|   |    | 50 | 0.002460 | 0.00002555 | 0.002325 | 0.002166 | 0.002027 | 0.002132 | 0.182032 |
|   |    | 100| 0.000113 | 0.00001116 | 0.000110 | 0.000105 | 0.000105 | 0.000104 | 0.250246 |
| 0.5 | 10  | 0.035795 | 0.041156 | 0.029762 | 0.033484 | 0.034823 | 0.031475 | 0.040751 |
| 50  | 0.005324 | 0.005461 | 0.005181 | 0.005195 | 0.005236 | 0.005132 | 0.156265 |
| 100 | 0.002575 | 0.002618 | 0.002525 | 0.002571 | 0.002582 | 0.002556 | 0.250084 |
| 1.0 | 10  | 0.149116 | 0.171023 | 0.124323 | 0.134477 | 0.139857 | 0.126409 | 0.036109 |
| 50  | 0.021537 | 0.022279 | 0.020679 | 0.021254 | 0.021424 | 0.020999 | 0.136043 |
| 100 | 0.010170 | 0.010359 | 0.009949 | 0.010329 | 0.010370 | 0.010267 | 0.248364 |
| 1.5 | 10  | 0.408065 | 0.465450 | 0.340891 | 0.315054 | 0.327656 | 0.296150 | 0.037178 |
| 50  | 0.051967 | 0.053591 | 0.050103 | 0.047638 | 0.048020 | 0.047067 | 0.155156 |
| 100 | 0.025544 | 0.026057 | 0.024916 | 0.023432 | 0.023526 | 0.023291 | 0.304977 |
| 2.0 | 10  | 0.564961 | 0.648791 | 0.471126 | 0.531805 | 0.553077 | 0.499896 | 0.041989 |
| 50  | 0.087924 | 0.090651 | 0.084845 | 0.083845 | 0.085060 | 0.083372 | 0.132457 |
| 100 | 0.041418 | 0.042203 | 0.040488 | 0.041383 | 0.041549 | 0.041135 | 0.295812 |
| 2.0 | 0.1 | 10  | 0.001907 | 0.002175 | 0.001150 | 0.001423 | 0.001480 | 0.001337 | 0.034776 |
| 50  | 0.000216 | 0.000223 | 0.000207 | 0.000213 | 0.000214 | 0.000210 | 0.129348 |
| 100 | 0.000108 | 0.000109 | 0.000106 | 0.000102 | 0.000103 | 0.000102 | 0.244397 |
| 0.5 | 10  | 0.042393 | 0.048605 | 0.035168 | 0.034895 | 0.036291 | 0.032801 | 0.042162 |
| 50  | 0.005674 | 0.005871 | 0.005441 | 0.005338 | 0.005380 | 0.005274 | 0.129727 |
| 100 | 0.002605 | 0.002646 | 0.002558 | 0.002567 | 0.002577 | 0.002552 | 0.252931 |
| 1.0 | 10  | 0.175736 | 0.200696 | 0.146651 | 0.139256 | 0.144826 | 0.130901 | 0.040481 |
| 50  | 0.020949 | 0.021618 | 0.020198 | 0.021086 | 0.021255 | 0.020833 | 0.129640 |
| 100 | 0.011099 | 0.011291 | 0.010873 | 0.010328 | 0.010369 | 0.010266 | 0.294799 |
| 1.5 | 10  | 0.397639 | 0.457980 | 0.326517 | 0.323163 | 0.336089 | 0.303773 | 0.042887 |
| 50  | 0.047416 | 0.048898 | 0.045761 | 0.047379 | 0.047758 | 0.046810 | 0.132863 |
| 100 | 0.024399 | 0.024812 | 0.023920 | 0.023194 | 0.023287 | 0.023055 | 0.271170 |
| 2.0 | 10  | 0.836536 | 0.955787 | 0.693110 | 0.590840 | 0.614474 | 0.555390 | 0.042558 |
| 50  | 0.088464 | 0.091485 | 0.084954 | 0.085103 | 0.085784 | 0.084082 | 0.138303 |
| 100 | 0.042187 | 0.043191 | 0.040932 | 0.041921 | 0.042089 | 0.041669 | 0.252456 |

Table 1. The MSE and the posterior risk of the estimator for the parameter $k$ of Burr distribution.
Furthermore, in the last column of Table 1, we also present the computational time of the experiment. For the same true value of the parameter $k$, when the sample sizes are larger, the computational time are longer.

7. Conclusion
With respect to the MSE as well as the posterior risk, the Bayesian estimation for the parameter $k$ of Burr distribution under Jeffrey prior is found to be superior in turn with the extended Jeffrey prior for all sample sizes and all true values of the parameter $k$. That is, the estimation under the extended Jeffrey prior with the power $p$ more than 0.5 gives better results than the estimation under Jeffrey prior; otherwise, when the power $p$ is less than 0.5, the better results are given by the estimation under Jeffrey prior.

Acknowledgments
This work is supported by Hibah PITTA 2018 funded by DRPM Universitas Indonesia No.2283/UN2.R3.1/HKP.05.00/2018.

References
[1] Burr I W 1942 Cumulative frequency functions Annals of Mathematic. Stat. 13 215–32
[2] Lewis A W 1981 The Burr distribution as a general parametric family in survivorship and reliability theory applications Ph.D. Thesis Department of Biostatistics University of North Carolina at Chapel Hill
[3] Wingo D R 1983 Maximum likelihood methods for fitting the Burr XII distribution to life test data Biometric. J. 25 77–84
[4] 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. 36 359–62
[5] Gupta P L, Gupta R C and Lvin S J 1996 Analysis of failure time data by Burr distribution Commun. Stat.–Theory and Meth. 25 2013–24
[6] Berger J O 1985 Statistical decision theory and Bayesian analysis 2nd ed (New York: Springer-Verlag New York Inc.)
[7] Papadopoulos A S 1978 The Burr distribution as a failure model from a Bayesian approach IEEE Transactions on Reliab. 27 369–71
[8] Moore D and Papadopoulos A S 2000 The Burr Type XII distribution as a failure model under various loss functions Microelectron. Reliab. 40 2117–22
[9] Nasir S A and Al-Anber N J 2012 A comparison of the Bayesian and other methods for estimation of reliability function for Burr-XII distribution J. Math. and Stat. 8 42–8
[10] 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 Mathematic. Theory and Modeling 4 63–77
[11] Hogg R V and Craig A T 1995 Introduction to mathematical statistics 5th ed (United States of America: Prentice Hall, Inc.)