Journal of the Korean Data & Information Science Society
2012, **23**(6), 1299–1308

http://dx.doi.org/10.7465/jkdi.2012.23.6.1299
한국데이터정보과학회지

Default Bayesian hypothesis testing for the scale parameters in nonregular Pareto distributions †

Sang Gil Kang[1]

[1]Department of Computer and Data Information, Sangji University

Received 22 October 2012, revised 13 November 2012, accepted 18 November 2012

Abstract

This article deals with the problem of testing the equality of the scale parameters in nonregular Pareto distributions. We propose Bayesian hypothesis testing procedures for the equality of the scale parameters under the noninformative prior. The noninformative prior is usually improper which yields a calibration problem that makes the Bayes factor to be defined up to a multiplicative constant. So we propose the default Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors under the reference priors. Simulation study and a real data example are provided.

Keywords: Fractional Bayes factor, intrinsic Bayes factor, Pareto distribution, reference prior, scale parameter.

1. Introduction

Consider X and Y are independently distributed random variables according to the Pareto distribution $\mathcal{P}(\alpha, \lambda_1)$ with the shape parameter α and the scale parameter λ_1, and the Pareto distribution $\mathcal{P}(\beta, \lambda_2)$ with the shape parameter β and the scale parameter λ_2. Then the Pareto distributions of X and Y are given by

$$f(x|\alpha, \lambda_1) = \alpha\lambda_1^{\alpha} x^{-(\alpha+1)}, x \geq \lambda_1 > 0, \alpha > 0, \tag{1.1}$$

and

$$f(y|\beta, \lambda_2) = \beta\lambda_2^{\beta} y^{-(\beta+1)}, y \geq \lambda_2 > 0, \beta > 0, \tag{1.2}$$

respectively. The present paper focuses on testing the equality of the scale parameters in the Pareto distributions.

In Bayesian model selection or testing problem, the Bayes factors under proper priors or informative priors have been very successful. However, limited information and time constraints often require the use of noninformative priors. Since noninformative priors such as Jeffreys' prior or reference prior (Berger and Bernardo, 1989, 1992) are typically improper

† This research was supported by Sangji University Research Fund, 2011.
[1] Associate professor, Department of Computer and Data Information, Sangji University, Wonju 220-702, Korea. E-mail: sangkg@sangji.ac.kr

so that they are only defined up to arbitrary constants which affects the values of Bayes factors. Spiegelhalter and Smith (1982), O'Hagan (1995) and Berger and Pericchi (1996) have made efforts to compensate for that arbitrariness.

Spiegelhalter and Smith (1982) used the device of imaginary training sample in the context of linear model comparisons to choose the arbitrary constants. But the choice of imaginary training sample depends on the models under comparison, and so there is no guarantee that the Bayes factor of Spiegelhalter and Smith (1982) is coherent for multiple model comparisons. Berger and Pericchi (1996) introduced the intrinsic Bayes factor using a data-splitting idea, which would eliminate the arbitrariness of improper prior. O'Hagan (1995) proposed the fractional Bayes factor. For removing the arbitrariness he used a portion of the likelihood with a so-called the fraction b. These approaches have shown to be quite useful in many statistical areas (Kang *et al.*, 2010, 2011; Cho, 2010). An excellent exposition of the objective Bayesian method to model selection is proposed in Berger and Pericchi (2001).

The Pareto distribution provides a statistical model which has an extensive variety of applications. It has been found in describing distributions of studies of income, property values, insurance risk, stock prices fluctuations, migration, size of cities and firms, word frequencies, occurrences of natural resources, business failures, service time in queuing systems, error clustering in communications circuits and lifetime data, etc (Arnold and Press, 1983; Fernández, 2008). The Pareto distribution has been used by many authors in a Bayesian viewpoint (e.g., Arnold and Press, 1983, 1989; Geisser, 1984, 1985; Lwin, 1972; Nigm and Hamdy, 1987; Tiwari *et al.*, 1996; Ko and Kim, 1999; Fernández, 2008; Kim *et al.*, 2009; Kang, 2010). For the common scale parameter, Elfessi and Jin (1996) derived a class of improved estimators which uniformly dominates the MLE under a class of convex scale invariant loss functions. In many problems we are interested in comparing the variabilities of two populations. Usually this can be done by comparing the scales of two populations. However the problem of testing for the equality of the scale parameters has not been considered. Therefore there is a necessity for developing hypothesis testing procedure for the equality of the scale parameters.

In this paper, we propose the objective Bayesian hypothesis testing procedures for the equality of the scale parameters in Pareto distributions based on the Bayes factors. The outline of the remaining sections is as follows. In Section 2, we introduce the Bayesian hypothesis testing based on the Bayes factor. In Section 3, using the reference priors, we provide the Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors. In Section 4, simulation study and a real data example are given.

2. Intrinsic and fractional Bayes factors

Suppose that hypotheses H_1,H_2 ,$\cdots$,H_q are under consideration, with the data $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ having probability density function $f_i(\mathbf{x}|\theta_i)$ under hypothesis H_i. The parameter vector θ_i is unknown. Let $\pi_i(\theta_i)$ be the prior distribution of parameter θ_i under hypothesis H_i, and let p_i be the prior probability of hypothesis H_i,$i = 1, 2, \cdots, q$. Then the posterior probability of that the hypothesis H_i is true is

$$P(H_i|\mathbf{x}) = \left(\sum_{j=1}^{q} \frac{p_j}{p_i} \cdot B_{ji} \right)^{-1}, \tag{2.1}$$

where B_{ji} is the Bayes factor of hypothesis H_j to hypothesis H_i defined by

$$B_{ji} = \frac{\int f_j(\mathbf{x}|\theta_j)\pi_j(\theta_j)d\theta_j}{\int f_i(\mathbf{x}|\theta_i)\pi_i(\theta_i)d\theta_i} = \frac{m_j(\mathbf{x})}{m_i(\mathbf{x})}. \tag{2.2}$$

The B_{ji} can be interpreted as the comparative support of the data for H_j versus H_i. The computation of B_{ji} needs specification of the prior distribution $\pi_i(\theta_i)$ and $\pi_j(\theta_j)$. Often in Bayesian analysis, one can use noninformative priors π_i^N. Common choices are the uniform prior, Jeffreys' prior and the reference prior. The noninformative prior π_i^N is typically improper. Hence the use of noninformative prior π_i^N in (2.2) causes the B_{ji} to contain unspecified constants. To solve this problem, Berger and Pericchi (1996) proposed the intrinsic Bayes factor, and O'Hagan (1995) proposed the fractional Bayes factor.

One solution to this indeterminacy problem is to use part of the data as a training sample. Let $\mathbf{x}(l)$ denote the part of the data to be so used and let $\mathbf{x}(-l)$ be the remainder of the data, such that

$$0 < m_i^N(\mathbf{x}(l)) < \infty, i = 1, \cdots, q. \tag{2.3}$$

In (2.3), the posteriors $\pi_i^N(\theta_i|\mathbf{x}(l))$ are well defined. Now, consider the Bayes factor, $B_{ji}(l)$, with the remainder of the data $\mathbf{x}(-l)$, using $\pi_i^N(\theta_i|\mathbf{x}(l))$ as the prior:

$$B_{ji}(l) = \frac{\int f(\mathbf{x}(-l)|\theta_j, \mathbf{x}(l))\pi_j^N(\theta_j|\mathbf{x}(l))d\theta_j}{\int f(\mathbf{x}(-l)|\theta_i, \mathbf{x}(l))\pi_i^N(\theta_i|\mathbf{x}(l))d\theta_i} = B_{ji}^N \cdot B_{ij}^N(\mathbf{x}(l)) \tag{2.4}$$

where

$$B_{ji}^N = B_{ji}^N(\mathbf{x}) = \frac{m_j^N(\mathbf{x})}{m_i^N(\mathbf{x})}$$

and

$$B_{ij}^N(\mathbf{x}(l)) = \frac{m_i^N(\mathbf{x}(l))}{m_j^N(\mathbf{x}(l))}$$

are the Bayes factors that would be obtained for the full data $\mathbf{x}$ and training samples $\mathbf{x}(l)$, respectively.

Berger and Pericchi (1996) proposed the use of a minimal training sample to compute $B_{ij}^N(\mathbf{x}(l))$. Then, an average over all the possible minimal training samples contained in the sample is computed. Thus the arithmetic intrinsic Bayes factor (AIBF) of H_j to H_i is

$$B_{ji}^{AI} = B_{ji}^N \times \frac{1}{L}\sum_{l=1}^{L} B_{ij}^N(\mathbf{x}(l)), \tag{2.5}$$

where L is the number of all possible minimal training samples. Also the median intrinsic Bayes factor (MIBF) by Berger and Pericchi (1998) of H_j to H_i is

$$B_{ji}^{MI} = B_{ji}^N \times ME[B_{ij}^N(\mathbf{x}(l))], \tag{2.6}$$

where ME indicates the median for all the training sample Bayes factors.

Therefore we can also calculate the posterior probability of H_i using (2.1), where B_{ji} is replaced by B_{ji}^{AI} and B_{ji}^{MI} from (2.5) and (2.6), respectively.

The fractional Bayes factor (O'Hagan, 1995) is based on a similar intuition to that behind the intrinsic Bayes factor but, instead of using part of the data to turn noninformative priors into proper priors, it uses a fraction, b, of each likelihood function, $L(\theta_i) = f_i(\mathbf{x}|\theta_i)$, with the remaining $1-b$ fraction of the likelihood used for model discrimination. Then the fractional Bayes factor (FBF) of hypothesis H_j versus hypothesis H_i is

$$B_{ji}^F = B_{ji}^N \cdot \frac{\int L^b(\mathbf{x}|\theta_i)\pi_i^N(\theta_i)d\theta_i}{\int L^b(\mathbf{x}|\theta_j)\pi_j^N(\theta_j)d\theta_j} = B_{ji}^N \cdot \frac{m_i^b(\mathbf{x})}{m_j^b(\mathbf{x})}. \tag{2.7}$$

O'Hagan (1995) proposed three ways for the choice of the fraction b. One common choice of b is $b = m/n$, where m is the size of the minimal training sample, assuming that this number is uniquely defined. See O'Hagan (1995, 1997) and the discussion by Berger and Mortera in O'Hagan (1995).

3. Bayesian hypothesis testing procedures

Let $X_i, i = 1, \cdots, n$ denote observations from the Pareto distribution $\mathcal{P}(\alpha, \lambda_1)$, and $Y_i, i = 1, \cdots, m$ denote observations from the Pareto distribution $\mathcal{P}(\beta, \lambda_2)$. Then likelihood function is given by

$$f(\mathbf{x}, \mathbf{y}|\alpha, \beta, \lambda_1, \lambda_2) = \alpha^n \beta^m \lambda_1^{n\alpha} \lambda_2^{m\beta} \prod_{i=1}^n x_i^{-(\alpha+1)} \prod_{i=1}^m y_i^{-(\beta+1)}, x_i \geq \lambda_1, y_i \geq \lambda_2, \tag{3.1}$$

where $\mathbf{x} = (x_1, \cdots, x_n)$, $\mathbf{y} = (y_1, \cdots, y_m)$, $\alpha > 0, \beta > 0$ and $\lambda_1, \lambda_2 > 0$.

We are interested in testing the hypotheses $H_1 : \lambda_1 = \lambda_2$ versus $H_2 : \lambda_1 \neq \lambda_2$ based on the fractional Bayes factor and the intrinsic Bayes factors.

3.1. Bayesian hypothesis testing procedure based on the fractional Bayes factor

From (3.1) the likelihood function under the hypothesis $H_1 : \lambda_1 = \lambda_2 \equiv \lambda$ is

$$L_1(\lambda, \alpha, \beta|\mathbf{x}, \mathbf{y}) = \alpha^n \beta^m \lambda^{n\alpha+m\beta} \prod_{i=1}^n x_i^{-(\alpha+1)} \prod_{i=1}^m y_i^{-(\beta+1)}. \tag{3.2}$$

And under the hypothesis H_1, the reference prior for (λ, α, β) derived by Kang (2010) and is

$$\pi_1^N(\lambda, \alpha, \beta) \propto \lambda^{-1}\alpha^{-1}\beta^{-1}. \tag{3.3}$$

Then from the likelihood (3.2) and the reference prior (3.3) , the element $m_1^b(\mathbf{x}, \mathbf{y})$ of the FBF under H_1 is given by

$$\begin{aligned} & m_1^b(\mathbf{x}, \mathbf{y}) \qquad (3.4) \\ &= \int_0^{z_{(1)}} \int_0^\infty \int_0^\infty L_1^b(\lambda, \alpha, \beta|\mathbf{x}, \mathbf{y})\pi_1^N(\lambda, \alpha, \beta)d\alpha d\beta d\lambda \\ &= \Gamma[bn]\Gamma[bm] \prod_{i=1}^n x_i^{-b} \prod_{i=1}^m y_i^{-b} \int_0^{z_{(1)}} \lambda^{-1} \left[b \sum_{i=1}^n \log\left(\frac{x_i}{\lambda}\right)\right]^{-bn} \left[b \sum_{i=1}^m \log\left(\frac{y_i}{\lambda}\right)\right]^{-bm} d\lambda, \end{aligned}$$

where $z_{(1)} = \min\{x_1, \cdots, x_n, y_1, \cdots, y_m\}$. For the hypothesis $H_2 : \lambda_1 \neq \lambda_2$, the reference prior for $(\lambda_1, \lambda_2, \alpha, \beta)$ is

$$\pi^N(\lambda_1, \lambda_2, \alpha, \beta) \propto \lambda_1^{-1}\lambda_2^{-1}\alpha^{-1}\beta^{-1}. \tag{3.5}$$

The likelihood function under the hypothesis H_2 is

$$L_2(\lambda_1, \lambda_2, \alpha, \beta|\mathbf{x}, \mathbf{y}) = \alpha^n\beta^m\lambda_1^{n\alpha}\lambda_2^{m\beta}\prod_{i=1}^{n} x_i^{-(\alpha+1)}\prod_{i=1}^{m} y_i^{-(\beta+1)}. \tag{3.6}$$

Thus from the likelihood (3.6) and the reference prior (3.5), the element $m_2^b(\mathbf{x}, \mathbf{y})$ of FBF under H_2 is given as follows.

$$\begin{aligned} m_2^b(\mathbf{x}, \mathbf{y}) &= \int_0^{y_{(1)}}\int_0^{x_{(1)}}\int_0^{\infty}\int_0^{\infty} L_2^b(\lambda_1, \lambda_2, \alpha, \beta|\mathbf{x}, \mathbf{y})\pi_2^N(\lambda_1, \lambda_2, \alpha, \beta)d\alpha d\beta d\lambda_1 d\lambda_2 \\ &= \Gamma[bn]\Gamma[bm]\prod_{i=1}^{n} x_i^{-b}\prod_{i=1}^{m} y_i^{-b}\int_0^{y_{(1)}}\int_0^{x_{(1)}} \lambda_1^{-1}\lambda_2^{-1} \\ &\quad \times\left[b\sum_{i=1}^{n}\log\left(\frac{x_i}{\lambda_1}\right)\right]^{-bn}\left[b\sum_{i=1}^{m}\log\left(\frac{y_i}{\lambda_2}\right)\right]^{-bm} d\lambda_1 d\lambda_2, \end{aligned} \tag{3.7}$$

where $x_{(1)} = \min\{x_1, \cdots, x_n\}$ and $y_{(1)} = \min\{y_1, \cdots, y_m\}$. Therefore the element B_{21}^N of FBF is given by

$$B_{21}^N = \frac{S_2(\mathbf{x}, \mathbf{y})}{S_1(\mathbf{x}, \mathbf{y})}, \tag{3.8}$$

where

$$S_1(\mathbf{x}, \mathbf{y}) = \int_0^{z_{(1)}} \lambda^{-1}\left[\sum_{i=1}^{n}\log\left(\frac{x_i}{\lambda}\right)\right]^{-n}\left[\sum_{i=1}^{m}\log\left(\frac{y_i}{\lambda}\right)\right]^{-m} d\lambda$$

and

$$S_2(\mathbf{x}, \mathbf{y}) = \int_0^{y_{(1)}}\int_0^{x_{(1)}} \lambda_1^{-1}\lambda_2^{-1}\left[\sum_{i=1}^{n}\log\left(\frac{x_i}{\lambda_1}\right)\right]^{-n}\left[\sum_{i=1}^{m}\log\left(\frac{y_i}{\lambda_2}\right)\right]^{-m} d\lambda_1 d\lambda_2.$$

And the ratio of marginal densities with fraction b is

$$\frac{m_1^b(\mathbf{x}, \mathbf{y})}{m_2^b(\mathbf{x}, \mathbf{y})} = \frac{S_1(\mathbf{x}, \mathbf{y}; b)}{S_2(\mathbf{x}, \mathbf{y}; b)}, \tag{3.9}$$

where

$$S_1(\mathbf{x}, \mathbf{y}; b) = \int_0^{z_{(1)}} \lambda^{-1}\left[\sum_{i=1}^{n}\log\left(\frac{x_i}{\lambda}\right)\right]^{-bn}\left[\sum_{i=1}^{m}\log\left(\frac{y_i}{\lambda}\right)\right]^{-bm} d\lambda$$

and

$$S_2(\mathbf{x}, \mathbf{y}; b) = \int_0^{y_{(1)}}\int_0^{x_{(1)}} \lambda_1^{-1}\lambda_2^{-1}\left[\sum_{i=1}^{n}\log\left(\frac{x_i}{\lambda_1}\right)\right]^{-bn}\left[\sum_{i=1}^{m}\log\left(\frac{y_i}{\lambda_2}\right)\right]^{-bm} d\lambda_1 d\lambda_2.$$

Thus the FBF of H_2 versus H_1 is given by

$$B_{21}^F = \frac{S_2(\mathbf{x},\mathbf{y})}{S_1(\mathbf{x},\mathbf{y})} \cdot \frac{S_1(\mathbf{x},\mathbf{y};b)}{S_2(\mathbf{x},\mathbf{y};b)}. \tag{3.10}$$

Note that the calculation of the FBF of H_2 versus H_1 requires only two dimensional integration.

3.2. Bayesian hypothesis testing procedure based on the intrinsic Bayes factor

The element B_{21}^N of the intrinsic Bayes factor is computed in the fractional Bayes factor. So under minimal training sample, we only calculate the marginal densities for the hypotheses H_1 and H_2, respectively. The marginal density of (X_{j_1}, X_{j_2}) and (Y_{k_1}, Y_{k_2}) is finite for all $1 \le j_1 < j_2 \le n$ and $1 \le k_1 < k_2 \le m$ under each hypothesis. Thus we conclude that any training sample of size 4 is a minimal training sample.

The marginal density $m_1^N(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})$ under H_1 is given by

$$\begin{aligned} m_1^N(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}) &= \int_0^{z_{(j_1)}} \int_0^\infty \int_0^\infty f(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}|\lambda,\alpha,\beta)\pi_1^N(\lambda,\alpha,\beta)d\alpha d\beta d\lambda \\ &= \int_0^{z_{(j_1)}} (x_{j_1}x_{j_2})^{-1}(y_{k_1}y_{k_2})^{-1}\lambda^{-1} \\ &\quad \times \left[\log\left(\frac{x_{j_1}}{\lambda}\right) + \log\left(\frac{x_{j_2}}{\lambda}\right)\right]^{-2} \left[\log\left(\frac{y_{k_1}}{\lambda}\right) + \log\left(\frac{y_{k_2}}{\lambda}\right)\right]^{-2} d\lambda, \end{aligned}$$

where $z_{(j_1)} = \min\{x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}\}$, And the marginal density $m_2^N(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})$ under H_2 is given by

$$\begin{aligned} & m_2^N(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}) \\ &= \int_0^{y_{(k_1)}} \int_0^{x_{(j_1)}} \int_0^\infty \int_0^\infty f(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}|\lambda_1,\lambda_2,\alpha,\beta)\pi_2^N(\lambda_1,\lambda_2,\alpha,\beta)d\alpha d\beta d\lambda_1 d\lambda_2 \\ &= \int_0^{y_{(k_1)}} \int_0^{x_{(j_1)}} (x_{j_1}x_{j_2})^{-1}(y_{k_1}y_{k_2})^{-1}\lambda_1^{-1}\lambda_2^{-1} \\ &\times \left[\log\left(\frac{x_{j_1}}{\lambda_1}\right) + \log\left(\frac{x_{j_2}}{\lambda_1}\right)\right]^{-2} \left[\log\left(\frac{y_{k_1}}{\lambda_2}\right) + \log\left(\frac{y_{k_2}}{\lambda_2}\right)\right]^{-2} d\lambda_1 d\lambda_2, \end{aligned}$$

where $x_{(j_1)} = \min\{x_{j_1}, x_{j_2}\}$ and $y_{(k_1)} = \min\{y_{k_1}, y_{k_2}\}$. Therefore the AIBF of H_2 versus H_1 is given by

$$B_{21}^{AI} = \frac{S_2(\mathbf{x},\mathbf{y})}{S_1(\mathbf{x},\mathbf{y})} \left[\frac{1}{L}\sum_{j_1,j_2}^{n}\sum_{k_1,k_2}^{m} \frac{T_1(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})}{T_2(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})}\right], \tag{3.11}$$

where $L = [nm(n-1)(m-1)]/4$,

$$T_1(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}) = \int_0^{z_{(j_1)}} \lambda^{-1} \left[\log\left(\frac{x_{j_1}}{\lambda}\right) + \log\left(\frac{x_{j_2}}{\lambda}\right)\right]^{-2} \left[\log\left(\frac{y_{k_1}}{\lambda}\right) + \log\left(\frac{y_{k_2}}{\lambda}\right)\right]^{-2} d\lambda$$

and

$$\begin{aligned} T_2(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2}) &= \int_0^{y_{(k_1)}} \int_0^{x_{(j_1)}} \lambda_1^{-1}\lambda_2^{-1} \\ &\times \left[\log\left(\frac{x_{j_1}}{\lambda_1}\right) + \log\left(\frac{x_{j_2}}{\lambda_1}\right)\right]^{-2} \left[\log\left(\frac{y_{k_1}}{\lambda_2}\right) + \log\left(\frac{y_{k_2}}{\lambda_2}\right)\right]^{-2} d\lambda_1 d\lambda_2. \end{aligned}$$

Also the MIBF of H_2 versus H_1 is given by

$$B_{21}^{MI} = \frac{S_2(\mathbf{x}, \mathbf{y})}{S_1(\mathbf{x}, \mathbf{y})} ME\left[\frac{T_1(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})}{T_2(x_{j_1}, x_{j_2}, y_{k_1}, y_{k_2})}\right]. \tag{3.12}$$

Note that the calculations of the AIBF and the MIBF of H_2 versus H_1 require only two dimensional integration.

4. Numerical studies

4.1. Simulation study

In order to assess the Bayesian hypothesis testing procedures, we evaluate the posterior probability for several configurations of (α, λ_1), (β, λ_2) and (n, m). In particular, for fixed (α, λ_1) and (β, λ_2), we take 1,000 independent random samples of X_i and Y_i with sample size n and m from the models (1.1) and (1.2), respectively. We want to test the hypotheses $H_1 : \lambda_1 = \lambda_2$ versus $H_2 : \lambda_1 \neq \lambda_2$. The posterior probabilities of H_1 being true are computed assuming equal prior probabilities. Table 4.1 shows the results of the averages and the standard deviations in parentheses of posterior probabilities. In Table 4.1, $P^F(\cdot)$,$P^{AI}(\cdot)$ and $P^{MI}(\cdot)$ are the posterior probabilities of the hypothesis H_1 being true based on FBF, AIBF and MIBF, respectively. From Table 4.1, the FBF, the AIBF and the MIBF give fairly reasonable answers for all configurations. Also the FBF, the AIBF and the MIBF give a similar behavior for all sample sizes. However the FBF and the AIBF slightly favor the hypothesis H_2 than the MIBF.

4.2. Example

This example taken from Arnold and Press (1989). Dyer (1981) reported annual wage data (in multiples of 100 US dollars) of a random sample of 30 production line workers in a large industrial firm for which the Pareto distribution appeared to be adequate. To test the equality of the scale parameters, we randomly divided this data into two groups. The data sets are given by

Group 1: 119, 112, 156, 123, 115, 119, 132, 107, 103, 105, 158, 111, 101, 157, 112.
Group 2: 112, 154, 108, 103, 107, 125, 128, 151, 104, 116, 140, 108, 104, 119, 115.

For this data sets, the maximum likelihood estimates of α and λ_1 in group 1 are 5.6323 and 101.0, and for group 2, the maximum likelihood estimates of β and λ_2 are 7.1022 and 103.0.

We want to test the hypotheses $H_1 : \lambda_1 = \lambda_2$ versus $H_2 : \lambda_1 \neq \lambda_2$. The values of the Bayes factors and the posterior probabilities of H_1 are given in Table 4.2. From the results of Table 4.2, the posterior probabilities based on various Bayes factors give the same answer,

Table 4.1 The averages and the standard deviations in parentheses of posterior probabilities

λ_1	λ_2	(n,m)	$P^F(H_1\|\mathbf{x},\mathbf{y})$	$P^{AI}(H_1\|\mathbf{x},\mathbf{y})$	$P^{MI}(H_1\|\mathbf{x},\mathbf{y})$
			$\alpha=0.5,\beta=2.0$		
5.0	1.0	5,5	0.2604(0.1350)	0.3274 (0.1614)	0.3406 (0.1697)
		5,10	0.2449(0.1412)	0.3298 (0.1588)	0.3419 (0.1639)
		10,10	0.0538(0.0586)	0.0809 (0.0848)	0.0863 (0.0897)
		10,20	0.0523(0.0568)	0.0827 (0.0830)	0.0889 (0.0889)
5.0	3.5	5,5	0.5753(0.1509)	0.6033 (0.1486)	0.6363 (0.1568)
		5,10	0.5847(0.1895)	0.5844 (0.1661)	0.6190 (0.1729)
		10,10	0.5457(0.1877)	0.5760 (0.1879)	0.6190 (0.1936)
		10,20	0.5850(0.1961)	0.5791 (0.1889)	0.6219 (0.1950)
5.0	4.8	5,5	0.6615(0.1539)	0.6762 (0.1573)	0.7180 (0.1650)
		5,10	0.7156(0.1610)	0.6937 (0.1504)	0.7364 (0.1551)
		10,10	0.7806(0.1518)	0.7890 (0.1495)	0.8271 (0.1449)
		10,20	0.8007(0.1705)	0.7808 (0.1679)	0.8188 (0.1634)
5.0	5.0	5,5	0.6589(0.1526)	0.6737 (0.1534)	0.7159 (0.1600)
		5,10	0.7061(0.1792)	0.6875 (0.1671)	0.7292 (0.1733)
		10,10	0.7805(0.1630)	0.7906 (0.1628)	0.8270 (0.1594)
		10,20	0.8190(0.1660)	0.8019 (0.1654)	0.8370 (0.1613)
5.0	5.2	5,5	0.6623(0.1517)	0.6704 (0.1610)	0.7154 (0.1656)
		5,10	0.7134(0.1710)	0.6850 (0.1644)	0.7291 (0.1688)
		10,10	0.7690(0.1592)	0.7768 (0.1606)	0.8170 (0.1541)
		10,20	0.7999(0.1763)	0.7766 (0.1828)	0.8157 (0.1763)
5.0	6.5	5,5	0.5731(0.1921)	0.5692 (0.2159)	0.6123 (0.2187)
		5,10	0.5549(0.2681)	0.5144 (0.2788)	0.5569 (0.2872)
		10,10	0.4578(0.2988)	0.4418 (0.3135)	0.4863 (0.3172)
		10,20	0.3524(0.3676)	0.3272 (0.3646)	0.3528 (0.3780)
5.0	9.0	5,5	0.3938(0.2359)	0.3689 (0.2617)	0.4111 (0.2671)
		5,10	0.2711(0.3096)	0.2428 (0.3051)	0.2653 (0.3219)
		10,10	0.1134(0.1996)	0.1015 (0.2005)	0.1182 (0.2153)
		10,20	0.0575(0.1955)	0.0530 (0.1879)	0.0573 (0.1986)
			$\alpha=1.0,\beta=1.0$		
5.0	1.0	5,5	0.1548(0.1128)	0.1628 (0.1272)	0.1920 (0.1380)
		5,10	0.1224(0.1052)	0.1495 (0.1233)	0.1799 (0.1366)
		10,10	0.0066(0.0116)	0.0072 (0.0135)	0.0091 (0.0163)
		10,20	0.0056(0.0106)	0.0075 (0.0142)	0.0096 (0.0174)
5.0	3.5	5,5	0.5555(0.1804)	0.5408 (0.2047)	0.5872 (0.2020)
		5,10	0.5594(0.1994)	0.5185 (0.2042)	0.5741 (0.2014)
		10,10	0.4214(0.2247)	0.3998 (0.2337)	0.4526 (0.2405)
		10,20	0.4024(0.2066)	0.3610 (0.2011)	0.4177 (0.2121)
5.0	4.8	5,5	0.6779(0.1488)	0.6752 (0.1698)	0.7182 (0.1653)
		5,10	0.7476(0.1704)	0.7192 (0.1819)	0.7628 (0.1743)
		10,10	0.7784(0.1671)	0.7737 (0.1805)	0.8124 (0.1707)
		10,20	0.8302(0.1574)	0.8058 (0.1721)	0.8412 (0.1598)
5.0	5.0	5,5	0.6753(0.1565)	0.6750 (0.1768)	0.7192 (0.1741)
		5,10	0.7474(0.1691)	0.7200 (0.1852)	0.7632 (0.1772)
		10,10	0.7950(0.1587)	0.7909 (0.1712)	0.8285 (0.1601)
		10,20	0.8364(0.1577)	0.8127 (0.1734)	0.8471 (0.1612)
5.0	5.2	5,5	0.6743(0.1603)	0.6739 (0.1816)	0.7174 (0.1792)
		5,10	0.7556(0.1571)	0.7264 (0.1750)	0.7693 (0.1676)
		10,10	0.7851(0.1621)	0.7813 (0.1760)	0.8197 (0.1662)
		10,20	0.8357(0.1512)	0.8117 (0.1666)	0.8471 (0.1536)
5.0	6.5	5,5	0.5918(0.1685)	0.5782 (0.1954)	0.6256 (0.1910)
		5,10	0.5918(0.2173)	0.5486 (0.2346)	0.5930 (0.2345)
		10,10	0.5389(0.2230)	0.5194 (0.2360)	0.5731 (0.2354)
		10,20	0.3794(0.2696)	0.3319 (0.2694)	0.3753 (0.2793)
5.0	9.0	5,5	0.4388(0.1867)	0.4191 (0.2073)	0.4668 (0.2064)
		5,10	0.2826(0.2228)	0.2419 (0.2242)	0.2751 (0.2368)
		10,10	0.1925(0.1595)	0.1784 (0.1620)	0.2140 (0.1788)
		10,20	0.0323(0.0935)	0.0254 (0.0840)	0.0309 (0.0948)
			$\alpha=1.0,\beta=3.0$		
5.0	1.0	5,5	0.1228(0.0936)	0.1676 (0.1246)	0.1782 (0.1319)
		5,10	0.1024(0.0895)	0.1635 (0.1190)	0.1732 (0.1245)
		10,10	0.0045(0.0076)	0.0076 (0.0128)	0.0082 (0.0140)
		10,20	0.0044(0.0081)	0.0082 (0.0141)	0.0088 (0.0150)
5.0	3.5	5,5	0.4893(0.1634)	0.5110 (0.1724)	0.5440 (0.1772)
		5,10	0.4987(0.1960)	0.4930 (0.1801)	0.5286 (0.1851)
		10,10	0.3216(0.1786)	0.3461 (0.1891)	0.3827 (0.2008)
		10,20	0.3578(0.1955)	0.3511 (0.1905)	0.3884 (0.2023)
5.0	4.8	5,5	0.6613(0.1610)	0.6610 (0.1690)	0.7056 (0.1705)
		5,10	0.7376(0.1694)	0.6957 (0.1624)	0.7399 (0.1619)
		10,10	0.7725(0.1601)	0.7728 (0.1648)	0.8138 (0.1571)
		10,20	0.8197(0.1575)	0.7836 (0.1611)	0.8237 (0.1553)
5.0	5.0	5,5	0.6833(0.1574)	0.6849 (0.1673)	0.7288 (0.1706)
		5,10	0.7428(0.1719)	0.7010 (0.1703)	0.7459 (0.1701)
		10,10	0.7849(0.1664)	0.7829 (0.1721)	0.8230 (0.1632)
		10,20	0.8401(0.1467)	0.8065 (0.1559)	0.8444 (0.1464)
5.0	5.2	5,5	0.6709(0.1567)	0.6728 (0.1692)	0.7179 (0.1676)
		5,10	0.7369(0.1686)	0.6912 (0.1741)	0.7355 (0.1728)
		10,10	0.7767(0.1603)	0.7735 (0.1707)	0.8156 (0.1619)
		10,20	0.7943(0.1857)	0.7491 (0.2039)	0.7928 (0.1939)
5.0	6.5	5,5	0.4919(0.2259)	0.4624 (0.2521)	0.5107 (0.2537)
		5,10	0.3869(0.3102)	0.3313 (0.3081)	0.3637 (0.3223)
		10,10	0.2371(0.2524)	0.2116 (0.2542)	0.2450 (0.2691)
		10,20	0.1193(0.2610)	0.1039 (0.2476)	0.1141 (0.2609)
5.0	9.0	5,5	0.2158(0.1929)	0.1835 (0.2012)	0.2195 (0.2114)
		5,10	0.0874(0.2021)	0.0694 (0.1853)	0.0781 (0.1986)
		10,10	0.0187(0.0666)	0.0145 (0.0616)	0.0186 (0.0694)
		10,20	0.0047(0.0576)	0.0041 (0.0545)	0.0045 (0.0568)

Table 4.2 Bayes factor and posterior probabilities of $H_1 : \lambda_1 = \lambda_2$

B_{21}^F	$P^F(H_1\|\mathbf{x},\mathbf{y})$	B_{21}^{AI}	$P^{AI}(H_1\|\mathbf{x},\mathbf{y})$	B_{21}^{MI}	$P^{MI}(H_1\|\mathbf{x},\mathbf{y})$
0.2752	0.7842	0.3253	0.7545	0.2412	0.8057

and select the hypothesis H_1. The AIBF has the smallest posterior probability than any other posterior probabilities based on the FBF and the MIBF. The MIBF slightly seems to favor the complex hypothesis than the AIBF and the FBF.

5. Concluding remarks

In this paper, we developed the objective Bayesian hypothesis testing procedures based on the fractional Bayes factor and the intrinsic Bayes factors for the equality of the scale parameters in Pareto distributions under the reference priors. From our numerical results, the developed hypothesis testing procedures give fairly reasonable answers for all parameter configurations. However the FBF and the AIBF slightly favor the hypothesis H_2 than the MIBF. From our simulation and example, we recommend the use of the FBF and the AIBF than the MIBF for practical application in view of its simplicity and ease of implementation.

References

Arnold, B. C. and Press, S. J. (1983). Bayesian inference for Pareto populations. *Journal of Econometrics*, **21**, 287-306.

Arnold, B. C. and Press, S. J. (1989). Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association*, **84**, 1079-1084.

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association*, **84**, 200-207.

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors (with discussion). In *Bayesian Statistics IV*, edited by J. M. Bernardo et al., Oxford University Press, Oxford, 35-60.

Berger, J. O. and Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, **91**, 109-122.

Berger, J. O. and Pericchi, L. R. (1998). Accurate and stable Bayesian model selection: The median intrinsic Bayes factor. *Sankya B*, **60**, 1-18.

Berger, J. O. and Pericchi, L. R. (2001). Objective Bayesian methods for model selection: Introduction and comparison (with discussion). In *Model Selection, Institute of Mathematical Statistics Lecture Notes-Monograph Series*, **38**, edited by P. Lahiri, Beachwood, Ohio, 135-207.

Cho, J. S. (2010). Bayesian multiple comparisons in Freund's bivariate exponential populations with type I censored data. *Journal of the Korean Data & Information Science Society*, **21**, 569-574.

Dyer, D. (1981). Structural probability bounds for the strong Pareto law. *Canadian Journal of Statistics*, **9**, 71-77.

Elfessi, A. and Jin, C. (1996). On robust estimation of the common scale parameter of several Pareto distributions. *Statistics & Probability Letters*, **29**, 345-352.

Fernández, A. J. (2008). Highest posterior density estimation form multiply censored Pareto data. *Statistical Papers*, **49**, 333-341.

Geisser, S. (1984). Prediction Pareto and exponential observables. *Canadian Journal of Statistics*, **12**, 143-152.

Geisser, S. (1985). Interval prediction for Pareto and exponential observables. *Journal of Econometrics*, **29**, 173-185.

Kang, S. G. (2010). Noninformative priors for the common scale parameter in Pareto distributions. *Journal of the Korean Data & Information Science Society*, **21**, 335-343.

Kang, S. G., Kim, D. H. and Lee, W. D. (2010). Default Bayesian testing for normal mean with known coefficient of variation. *Journal of the Korean Data & Information Science Society*, **21**, 297-308.

Kang, S. G., Kim, D. H. and Lee, W. D. (2011). Default Bayesian testing for the bivariate normal correlation coefficient. *Journal of the Korean Data & Information Science Society*, **22**, 1007-1016.

Kim, D. H., Kang, S. G. and Lee, W. D. (2009). Noninformative priors for Pareto distribution. *Journal of the Korean Data & Information Science Society*, **20**, 1213-1223.

Ko, J. H. and Kim, Y. H. (1999). Bayesian prediction inference for censored Pareto model. *Journal of the Korean Data & Information Science Society*, **10**, 147-154.

Lwin, T. (1972). Estimation of the tail of the Paretian law. *Scandinavian Actuarial Journal*, **55**, 170-178.

Nigm, A. M. and Hamdy, H. L. (1987). Bayesian prediction bounds for the Pareto lifetime model. *Communications in Statistics: Theory and Methods*, **16**, 1761-1772.

O'Hagan, A. (1995). Fractional Bayes factors for model comparison (with discussion). *Journal of Royal Statistical Society B*, **57**, 99-118.

O'Hagan, A. (1997). Properties of intrinsic and fractional Bayes factors. *Test*, **6**, 101-118.

Spiegelhalter, D. J. and Smith, A. F. M. (1982). Bayes factors for linear and log-linear models with vague prior information. *Journal of Royal Statistical Society B*, **44**, 377-387.

Tiwari, R. C., Yang, Y. and Zalkikar, J. N. (1996). Bayes estimation for the Pareto failure-model using gibbs sampling. *IEEE Transactions on Reliability*, **45**, 471-476.