NONINFORMATIVE PRIORS FOR PARETO DISTRIBUTION : REGULAR CASE

Dal Ho Kim [1] Woo Dong Lee [2] Sang Gil Kang [3]

Abstract

In this paper, we develop noninformative priors for two parameter Pareto distribution. Specially, we derive Jeffrey's prior, probability matching prior and reference prior for the parameter of interest.

In our case, the probability matching prior is only a first order and there does not exist a second order matching prior. Some simulation reveals that the matching prior performs better to achieve the coverage probability. And a real example will be given.

1 INTRODUCTION

The Pareto distribution has found applications in modelling problems involving distributions of income when incomes exceed a certain limit.Many socio-economic and other naturally occurring quantities are distributed according to certain statistical distributions with very long right tails. Examples of some of these empirical phenomena are distributions of city population sizes, occurrence of natural resources, stock price fluctuations, size of firms, personal incomes, and error clustering in communication circuits.

Many distributions have been developed in an attempt to explain these empirical data. The Pareto and lognormal distributions have played a major part in these investigations. It has been observed that while the fit of the Pareto curve may be rather good at the extremities of the income range, the fit over the whole range is often rather poor. On the other hand, the lognormal distribution fits well over a larger part of the income range but diverges markedly at the extremities.

For economists concerned with upper tails of distributions, the Pareto distribution is probably more useful than the lognormal which generally gives poor fit in the tails.

The Pareto distribution has reverse J shaped and positively skewed with a decreasing hazard rate. Although the family was originally applied to analyzing certain socio-economic and nature phenomena with observations in long tails, the family has potential for modelling reliability and life time data as well (Arnold, 1983).

A number of authors have studied Bayesian inference procedures for this distribution, e.g., Arnold and Press (1983,1989a,1989b), Geisser (1984,1985), Lwin (1972), Nigam and Hamdy (1987), Tiwari, Yang and Zalkikar (1996).

Arnold and Press (1989b) studied the Bayesian estimation problem using the independent conjugate prior and modified Lwin prior. Recently, Soliman (2001) studied the Bayesian estimation of Pareto distribution with scale and shape parameters in various situations. He considered squarred error loss and LINEX loss for estimating parameters using subjective priors such as conjugate prior and Gamma-exponential priors.

But there is a situation that one forces to use noninformative priors such as Jeffrey's, reference or matching priors because the prior information for the distribution is rare.

There has been an amount of efforts for finding a noninformative or objective priors under certain statistical models. Jeffrey's prior was quite successful in many Bayesian inference, but it causes problems when the nuisance parameters are present.

[1] Department of Statistics, Kyungpook National University
[2] Faculty of Information and Science, Kyungsan University
[3] Department of Statistics , Sangji University

Recently, much effort has been done for finding reference or probability matching prior. These noninformative priors work well in many statistical problems when the nuisance parameters are present.

Berger and Bernado (1989,1992) extended Bernardo's (1979) reference prior approach, giving a general algorithm to derive a reference prior by splitting the parameters into several groups according to their order of inferential importance. On the other hand, Welch and Peers (1963), Peers(1965) and Stein (1985) found a prior which meet the frequentist coverage probability of the posterior region of a real-valued parametric function to match the nominal level with a reminder of $o(n^{-\frac{1}{2}})$, where n is the sample size. Tibshirani (1989) reconsidered the case when the real valued parameter of interest is orthogonal (in the sense of Cox and Reid (1987)) to the nuisance parameter vector. These priors, as usually referred to as 'first order ' matching priors, were further studied in Datta and Ghosh (1995a, 1995b, 1996). Recently, Mukerjee and Ghosh(1997) developed a 'second order' , that is $o(n^{-1})$, matching prior. They extended the result in Mukerjee and Dey (1993) to the case of multiple nuisance parameters based on quantiles, and also developed a second order matching prior based on distribution function.

In this paper, we derive the Jeffreys', reference and matching priors for the two parameters Pareto distribution. We show that when the parameter of interest is the model parameter itself, there does not exist a second order matching priors. Posterior propriety under the proposed noninformative priors will be given. And also, some examples are given including coverage probabilities using artificial data.

2   NONINFORMATIVE PRIORS FOR PARETO DISTRIBUTION

The Pareto distribution with shape parameter α and scale parameter β is given by

$$f(x; \alpha, \beta) = \alpha\beta^\alpha(x + \beta)^{-(\alpha+1)}, x > 0; (\alpha > 0, \beta > 0). \tag{1}$$

Soliman (2001) considered Bayes estimators of parameters, reliability, and failure rate functions for the above density (4).

Including Jeffrey's prior, we will find the probability matching prior and reference prior when the parameter of interest is α or β.

Let $I(\alpha, \beta)$ be the information matrix of α and β per observation. Then

$$I(\alpha, \beta) = \begin{pmatrix} \frac{1}{\alpha^2} & -\frac{1}{\beta(\alpha+1)} \\ -\frac{1}{\beta(\alpha+1)} & \frac{\alpha}{\beta^2(\alpha+2)} \end{pmatrix} \tag{2}$$

Suppose that we are interested in estimating the parameter α. Then α is the parameter of interest and β is the nuisance parameter. Consider the following reparametrization, which give the orthogonality of parameters in the sense of Cox and Reid (1987). Let

$$\omega_1 = \alpha, \omega_2 = \beta\frac{1+\alpha}{\alpha^2}.$$

Then the information matrix for (ω_1, ω_2) per an observation is given by

$$I(\omega_1, \omega_2) = \begin{pmatrix} \frac{1}{(1+\omega_1)^2\omega_1^2} & 0 \\ 0 & \frac{\omega_1}{(\omega_1+2)\omega_2^2} \end{pmatrix} \tag{3}$$

From the above information matrix, one can find the various noninformative priors when the parameter of interest is α, as follows.

The Jeffrey's prior $\pi_M(\omega_1, \omega_2)$ is given by

$$\pi_J(\omega_1, \omega_2) \propto \frac{1}{\omega_2(1 + \omega_1)\sqrt{\omega_1(2 + \omega_1)}}, \omega_1, \omega_2 > 0. \tag{4}$$

The class of first order matching prior is given by

$$\pi_M(\omega_1, \omega_2) \propto \frac{1}{\omega_1(1 + \omega_1)} g(\omega_2), \omega_1, \omega_2 > 0, \tag{5}$$

where $g(\cdot)$ is arbitrary differentiable function in its argument.

Finally, the reference prior is given by

$$\pi_R(\omega_1, \omega_2) \propto \frac{1}{(1 + \omega_1)\omega_1\omega_2}, \omega_1, \omega_2 > 0. \tag{6}$$

REMARK 1. One can easily find the fact that the reference prior satisfies the first order matching criterion.

The class of prior given in (6) is quite large, and it is necessary to narrow down this class of priors. To this end, we consider the class of second order probability matching priors as given in Murkerjee and Ghosh (1997). A second probability matching prior is a prior which satisfies the following equation. When the parameter of interest, ω_1, is orthogonal to the nuisance parameter ω_2, a second order matching prior satisfies

$$\frac{1}{6}g(\omega_2)\frac{d}{d\omega_1}\left(i_{11}^{-3/2}L_{1,1,1}\right) + \frac{d}{d\omega_2}\left(i_{11}^{-1/2}L_{112}i^{22}g(\omega_2)\right) = 0, \tag{7}$$

where L is a likelihood function of (ω_1, ω_2), i^{ab} is the a-th row and b-th column element of inverse of information matrix $I(\omega_1, \omega_2)$,

$$L_{1,1,1} = E\left[\left(\frac{\partial \log L}{\partial \omega_1}\right)^3\right],$$

and

$$L_{112} = E\left[\frac{\partial^3 \log L}{\partial \omega_1^2 \partial \omega_2}\right].$$

From the pdf (1), the likelihood function of (ω_1, ω_2) is given by

$$L(\omega_1, \omega_2) = \log \omega_1 + \omega_1 \log \omega_2 + 2\omega_1 \log \omega_1 - \omega_1 \log(1 + \omega_1) - (\omega_1 + 1)\log(x + \frac{\omega_1^2\omega_2}{1 + \omega_1}).$$

And from the above likelihood function, one can get

$$L_{1,1,1} = -\frac{2(3 + 5\omega_1)}{\omega_1^3(1 + \omega_1)^3(\omega_1 + 3)},$$

and

$$L_{112} = \frac{2}{\omega_1\omega_2(1 + \omega_1)(2 + \omega_1)(3 + \omega_1)}.$$

Using the fact that $i^{22} = \frac{(\omega_1 + 2)\omega_2^2}{\omega_1}$, the equation (8) reduces to

$$g(\omega_2)\frac{2}{\omega_1 + 3} + \frac{1}{\omega_1}\frac{d}{d\omega_2}(\omega_2 g(\omega_2)) = 0.$$

There does not exist a function $g(\omega_2)$ which satisfies the above differential equation. So, there does not exist a second order matching prior when the parameter of interest is α.

The priors (5),(6) and (7) can be re-expressed as original parameters α and β: For $\alpha, \beta > 0$,

$$\pi_J(\alpha, \beta) \quad \propto \quad \frac{1}{\beta(1+\alpha)\sqrt{\alpha(\alpha+2)}}$$

$$\pi_M(\alpha, \beta) \quad \propto \quad \frac{1}{\alpha^3}$$

$$\pi_R(\alpha, \beta) \quad \propto \quad \frac{1}{\beta\alpha(1+\alpha)}.$$

Now, we will prove the propriety of posterior distribution under the Jeffrey's, reference and matching priors when the parameter of interest is α.

Theorem 1. Under the prior $\pi_M(\alpha, \beta)$, the joint posterior distribution of (α, β) is proper if $n - 2 > 0$.

Proof. The joint posterior distribution of $\pi_M(\alpha, \beta|\mathbf{x})$ is proportional to

$$\pi_M(\alpha, \beta|\mathbf{x}) \propto \alpha^{n-3}\beta^{n\alpha} \prod_{i=1}^{n}(x_i + \beta)^{-(\alpha+1)}.$$

Let y be the minimum of $x_1, x_2, \cdots, x_n$. Then

$$\int_0^\infty \int_0^\infty \alpha^{n-3}\beta^{n\alpha} \prod_{i=1}^{n}(x_i + \beta)^{-(\alpha+1)}d\alpha d\beta \quad \leq \quad \int_0^\infty \int_0^\infty \alpha^{n-3}\beta^{n\alpha}(y+\beta)^{-n(\alpha+1)}d\alpha d\beta$$

$$= \quad y^{-n+1}\Gamma(n-1)\int_0^\infty \alpha^{n-3}\frac{\Gamma(n\alpha+1)}{\Gamma(n\alpha+n)}d\alpha.$$

And the integrand of the last term is bounded by

$$\alpha^{n-3}\frac{\Gamma(n\alpha+1)}{\Gamma(n\alpha+n)} \quad = \quad \frac{\alpha^{n-3}}{(n\alpha+n-1)(n\alpha+n-2)\cdots(n\alpha+1)}$$

$$\leq \quad \frac{\alpha^{n-3}}{(n\alpha+1)^{n-1}}.$$

Now, the integral

$$\int_0^\infty \frac{\alpha^{n-3}}{(n\alpha+1)^{n-1}}d\alpha = n^{-(n-2)}Beta(n-2, 1),$$

where $Beta(n-2, 1) = \frac{\Gamma(n-2)\Gamma(1)}{\Gamma(n-1)}$ is finite if $n - 2 > 0$. This completes the proof.

Theorem 2. Under the prior $\pi_J(\alpha, \beta)$, the joint posterior distribution of (α, β) is proper.

Proof. The joint posterior distribution is proportional to

$$\pi_J(\alpha, \beta|\mathbf{x}) \propto \alpha^{n-1/2}(1+\alpha)^{-1}(2+\alpha)^{-1/2}\beta^{n\alpha-1} \prod_{i=1}^{n}(x_i + \beta)^{-(\alpha+1)}.$$

Let y be the minimum of $x_1, x_2, \cdots, x_n$. Then

$$\int_0^\infty \int_0^\infty \pi_J(\alpha, \beta|\mathbf{x})d\alpha d\beta \quad \leq \quad \int_0^\infty \int_0^\infty \alpha^{n-1/2}(1+\alpha)^{-1}(2+\alpha)^{-1/2}$$

$$\times \beta^{n\alpha-1}(y+\beta)^{-n(\alpha+1)}d\alpha d\beta$$

$$= \quad y^{-n}\int_0^\infty \alpha^{n-1/2}(1+\alpha)^{-1}(2+\alpha)^{-1/2}\frac{\Gamma(n\alpha)\Gamma(n)}{\Gamma(n\alpha+n)}d\alpha.$$

Now, consider the integration for α.

$$
\begin{aligned}
\int_0^\infty \frac{\alpha^{n-1/2}\Gamma(n\alpha)}{(1+\alpha)(2+\alpha)^{1/2}\Gamma(n\alpha+n)}d\alpha &= \int_0^\infty \frac{\alpha^{n-1/2}(1+\alpha)^{-1}(2+\alpha)^{-1/2}}{(n\alpha+n-1)(n\alpha+n-2)\cdots n\alpha}d\alpha \\
&\leq \int_0^\infty \frac{\alpha^{n-1/2}}{(1+\alpha)(2+\alpha)^{1/2}(n\alpha)^n}d\alpha \\
&= \frac{1}{n^n}\int_0^\infty \frac{1}{(1+\alpha)\sqrt{(\alpha+2)\alpha}}d\alpha \\
&= \frac{\pi}{2n^n} < \infty.
\end{aligned}
$$

This completes the proof.

Theorem 3. Under the prior $\pi_R(\alpha,\beta)$, the joint posterior distribution of (α,β) is proper if $n > 2$.

Proof. The joint posterior distribution of (α,β) under reference prior is given by

$$
\pi_R(\alpha,\beta|\mathbf{x}) \propto \frac{\alpha^{n-1}\beta^{n\alpha-1}}{1+\alpha}\prod_{i=1}^n (x_i+\beta)^{-(\alpha+1)}.
$$

Let y be the minimum of $x_1, x_2, \cdots, x_n$. Then

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

Note that,

$$
\begin{aligned}
\int_0^\infty\int_0^\infty \frac{\alpha^{n-1}\beta^{n\alpha-1}(y+\beta)^{-n(\alpha+1)}}{1+\alpha}d\alpha d\beta &= y^{-n}\int_0^\infty \frac{\alpha^{n-1}}{(1+\alpha)}\frac{\Gamma(n\alpha)\Gamma(n)}{\Gamma(n\alpha+n)}d\alpha \\
&= y^{-n}\Gamma(n)\int_0^\infty \frac{\alpha^{n-2}}{(1+\alpha)n} \times \\
&\qquad \frac{1}{(n\alpha+n-1)(n\alpha+n-2)\cdots(n\alpha+1)}d\alpha \\
&\leq \frac{y^{-n}\Gamma(n)}{n}\int_0^\infty \frac{\alpha^{n-2}}{(1+\alpha)(n\alpha+1)^{n-1}}d\alpha.
\end{aligned}
$$

Since $\frac{\alpha}{1+\alpha} < 1$, the integration in the last inequality reduced to

$$
\int_0^\infty \frac{\alpha^{n-3}}{(n\alpha+1)^{n-1}}d\alpha = n^{-(n-2)}Beta(n-2,1).
$$

And $Beta(n-2,1)$ is finite if $n > 2$. This completes the proof.

When the parameter of interest is β, we derive the probability matching prior, and reference prior. In this case, we know that Jefferey's prior is the same as the case when the parameter of interest is α.

From the information matrix for (α,β) given in (3), following Peers (1965), the first order matching prior π_r for β is the solution of the following partial differential equation.

$$
\frac{\partial}{\partial\alpha}\alpha^{3/2}\sqrt{\alpha+2}\pi_r + \frac{\partial}{\partial\beta}\frac{\beta(\alpha+1)\sqrt{\alpha+2}}{\sqrt{\alpha}}\pi_r = 0. \tag{8}
$$

A solution of the above partial differential equation is given by

$$
\pi_m(\alpha,\beta) \propto \frac{1}{\beta\alpha^{3/2}\sqrt{\alpha+2}}, \alpha > 0, \beta > 0. \tag{9}
$$

This is the first order matching prior when the parameter of interest is β.

Now, we will check that whether the above first order matching prior is the second order matching prior or not. Let θ_1 be the parameter of interest and θ_2 be a nuisance parameter. Let

$$\Sigma = \begin{pmatrix} a_{20} & a_{11} \\ a_{11} & a_{02} \end{pmatrix} \tag{10}$$

be the Fisher information matrix for (θ_1, θ_2).

When θ_1 is not orthogonal to θ_2, Murkerjee and Dey (1993) show that the second order matching prior for $\theta_1 (= \beta)$ satisfies the above differential equation (9) and additional following equation:

$$T(\pi; \theta_1, \theta_2) \equiv \sum_{i=1}^{4} L_i(\pi; \theta_1, \theta_2) = 0,$$

where

$$L_1(\pi; \theta_1, \theta_2) = \frac{1}{2}\{D_1^2 B^{-1} - 2D_1^1 D_2^1 (a_{11} a_{02}^{-1} B^{-1}) + D_2^2 (a_{11}^2 a_{02}^{-2} B^{-1})\},$$

$$L_2(\pi; \theta_1, \theta_2) = -\frac{D_1^2 \pi - 2a_{11} a_{02}^{-1} D_1 D_2 \pi + a_{11}^2 a_{02}^{-2} D_2^2 \pi}{2\pi B},$$

$$L_3(\pi; \theta_1, \theta_2) = -\frac{D_2^1 \{(a_{02}B)^{-1}(K_{21} - 2a_{11}a_{02}^{-1}K_{12} + a_{11}^2 a_{02}^{-2} K_{03})\pi\}}{2\pi},$$

$$L_4(\pi; \theta_1, \theta_2) = -\frac{D_1^1 \psi(\pi; \theta_1, \theta_2) - D_2^1 \{a_{11} a_{02}^{-1} \psi(\pi; \theta_1, \theta_2)\}}{\pi}.$$

Here $D_i^k = \partial^k / \partial \theta_i^k$,

$$K_{ij} = E\{D_1^i D_2^j \log p(X; \theta_1, \theta_2)\},$$

$$\psi(\pi; \theta_1, \theta_2) = \frac{(K_{30} - 3a_{11}a_{02}^{-1}K_{21} + 3a_{11}^2 a_{02}^{-2} K_{12} - a_{11}^3 a_{02}^{-3} K_{03})\pi}{6B^2}$$

$$= \frac{D_1^1 \pi - a_{11} a_{02}^{-1} D_2^1 \pi}{B},$$

and $B = a_{20} - a_{11}^2 / a_{02}$.

After some calculation, we get

$$L_1(\pi; \beta, \alpha) = 10\alpha^2 + 16\alpha + 2/\alpha + 7,$$

$$L_2(\pi; \beta, \alpha) = -\frac{12\alpha^4 + 48\alpha^3 + 67\alpha^2 + 36\alpha + 8}{2\alpha^2 + 4\alpha},$$

$$L_3(\pi; \beta, \alpha) = -\frac{\alpha + 1}{\alpha + 2},$$

and

$$L_4(\pi; \beta, \alpha) = \frac{2(2\alpha^5 + 18\alpha^4 + 55\alpha^3 + 63\alpha^2 + 11\alpha - 13)}{(\alpha + 2)(\alpha + 3)^2}.$$

Hence

$$T(\pi; \theta_1, \theta_2) = \frac{5\alpha^3 + 12\alpha^2 + \alpha - 2}{2(\alpha + 2)(\alpha + 3)^2} \neq 0.$$

This means that the probability matching prior $\pi_m(\alpha, \beta)$ is not a second order matching prior but a first order matching prior.

Berger and Bernardo (1989) gave an algorithm for deriving a reference prior for problems with nuisance parameters. We will derive the reference prior when β is of interest.

From the information matrix, one can obtain the reference prior for α with given β as follows:

$$\pi(\alpha|\beta) = \frac{1}{\alpha}.$$

Choose a sequence of compact sets for (β, α) by $(l_{1i}, l_{2i}) \times (k_{1i}, k_{2i})$, so that $l_{1i}, k_{1i} \to 0$ and $l_{2i}, k_{2i} \to \infty$ as $i \to \infty$. Let I_A be the indicator of a set A. The conditional prior of α given β is

$$\pi_i(\alpha|\beta) = \frac{\frac{1}{\alpha}I_{(k_{1i},k_{2i})}(\alpha)}{\int_{k_{1i}}^{k_{2i}} \frac{1}{\alpha}d\alpha} = \frac{I_{(k_{1i},k_{2i})}(\alpha)}{\alpha[\log(k_{2i}) - \log(k_{1i})]},$$

and the marginal reference prior for β is

$$\pi_i(\beta) = \exp\left\{\frac{1}{2[\log(k_{2i}) - \log(k_{1i})]} \int_{k_{2i}}^{k_{1i}} \frac{1}{\alpha} \log\left(\frac{\alpha}{(\alpha+1)^2(\alpha+2)}\right) d\alpha\right\} \times \frac{1}{\beta}.$$

Following Berger and Bernardo (1989), the reference prior for $\{\beta, \alpha\}$ is

$$\pi_r(\beta, \alpha) = \lim_{i \to \infty} \frac{\pi_i(\beta)\pi_i(\alpha|\beta)}{\pi_i(\beta_0)\pi_i(\alpha_0|\beta_0)} = \frac{1}{\alpha\beta}. \tag{11}$$

Here $(\alpha_0, \beta_0) = (1, 1)$.

Now, we will show that the first order matching prior given in (10) gives a proper posterior distribution but the reference prior given in (12) does not.

Theorem 4. Under the prior $\pi_m(\alpha, \beta)$, the joint posterior distribution of (α, β) is proper if $n - \frac{3}{2} > 0$.

Proof. Under the prior $\pi_m(\alpha, \beta)$, the joint posterior distribution of α and β is given by

$$\pi_m(\alpha, \beta|\mathbf{x}) \propto \alpha^{n-\frac{3}{2}} \beta^{n\alpha-1}(\alpha+2)^{-\frac{1}{2}} \prod_{i=1}^{n}(x_i + \beta)^{-(\alpha+1)}.$$

So the integral with respect to α and β is

$$\int_0^\infty \int_0^\infty \frac{\alpha^{n-\frac{3}{2}}\beta^{n\alpha-1}}{\sqrt{\alpha+2}} \prod_{i=1}^{n}(x_i+\beta)^{-(\alpha+1)} d\alpha d\beta \leq \int_0^\infty \int_0^\infty \frac{\alpha^{n-\frac{3}{2}}\beta^{n\alpha-1}}{\sqrt{\alpha+2}}(y+\beta)^{-n(\alpha+1)} d\alpha d\beta,$$

where y is the minimum of $x_1, x_2, \cdots, x_n$. The integral of the righthand side of the above inequality with respect to β results in

$$y^{-n}\Gamma(n) \int_0^\infty \frac{\alpha^{n-3/2}}{\sqrt{\alpha+2}(n\alpha+n-1)(n\alpha+n-2)\cdots(n\alpha+1)(n\alpha)} d\alpha.$$

The above quantity is less than or equal to

$$\frac{y^{-n}\Gamma(n)}{n} \int_0^\infty \frac{\alpha^{n-5/2}}{(n\alpha+1)^{n-1}} d\alpha,$$

since $\frac{1}{\sqrt{\alpha+2}} < 1$. Now the integral with respect to α gives

$$y^{-n}\Gamma(n)n^{1/2-n}Beta(n - \frac{3}{2}, \frac{1}{2}),$$

which is finite if $n > \frac{3}{2}$. This completes the proof.

Theorem 5. Under the reference prior $\pi_r(\alpha, \beta)$, the joint posterior is improper.

Table 1: Coverage Probabilities of Priors for α

n	π_M		π_J		π_R	
10	.0318000	1.0000000	.0000000	.9708000	.0000000	.9558000
20	.0442000	.9905000	.0332000	.9615000	.0144000	.9469000
30	.0500000	.9673000	.0531000	.9555000	.0364000	.9450000
40	.0447000	.9558000	.0634000	.9597000	.0452000	.9494000
50	.0459000	.9517000	.0644000	.9597000	.0492000	.9506000

Proof. The joint posterior distribution under the reference prior is proportional to

$$\pi_r(\alpha, \beta | \mathbf{x}) \propto \alpha^{n-1} \beta^{n\alpha-1} \prod_{i=1}^{n} (x_i + \beta)^{-(\alpha+1)}.$$

Now, let z be the maximum of $x_1, x_2, \cdots, x_n$. Then

$$\int_0^\infty \int_0^\infty \alpha^{n-1} \beta^{n\alpha-1} \prod_{i=1}^{n} (x_i + \beta)^{-(\alpha+1)} d\alpha d\beta \;\geq\; \int_0^\infty \int_0^\infty \frac{\alpha^{n-1}\beta^{n\alpha-1}}{(z+\beta)^{n(\alpha+1)}} d\alpha d\beta$$

$$= \frac{\Gamma(n)}{nz^n} \int_0^\infty \alpha^{n-2} \prod_{j=1}^{n-1} (n\alpha + j)^{-1} d\alpha$$

$$> \frac{\Gamma(n)}{nz^n} \int_0^\infty \alpha^{n-2} (n\alpha + n)^{-(n-1)} d\alpha,$$

and,

$$\frac{\Gamma(n)}{nz^n} \int_0^\infty \alpha^{n-2} (n\alpha + n)^{-(n-1)} d\alpha = \frac{\Gamma(n)}{n^n z^n} \int_0^1 k^{n-2} (1-k)^{-1} dk = \infty.$$

So, the joint posterior distribution under reference prior is improper.

3   SIMULATIONS AND EXAMPLES

In this section, we will compare the coverage probability of the proposed noninformative priors. Using the proposed noninformative priors, we will analyze the real data.

When the parameter of interest is α, the coverage probabilities of the priors are given in Table 1.

In Table 1., under the moderate sample size, one can see that π_M and π_R match the target coverage probability well. But the Jeffrey's prior π_J does not. Specially, the coverage probability of the reference prior is better than other priors. This is because π_R satisfies the first order matching criterion.

In Table 2., when the parameter of interest is β, the coverage probability of matching prior π_m matches the target coverage well. But the Jeffrey's prior does not match the target coverage.

Example. We will illustrate Bayesian analysis using proposed noninformative priors. The data given below are annual wage data. Dyer (1981) and Arnord and Press (1989b) analyzed these data. Specially, Arnord and Press (1989b) used Bayesian set up for analyzing the data. Annual wage data (in multiples of 100 U.S. dollars) of a random sample 30 production-line workers in a large industrial firm were as follows:

112	154	119	108	112	156	123	103	115	107	125	119	128
132	107	151	103	104	116	140	108	105	158	104	119	111
101	157	112	115									

The Bayes estimators of α and β under the proposed priors are as follows:

Table 2: Coverage Probabilities of Priors for β

n	π_m		π_J	
10	.0025000	.9524000	.0105000	.9751000
20	.0209000	.9486000	.0415000	.9651000
30	.0348000	.9463000	.0585000	.9613000
40	.0456000	.9480000	.0679000	.9612000
50	.0499000	.9486000	.0672000	.9605000

	α	β
π_J	4.741	526.731
π_M	4.461	518.253
π_R	4.671	520.897
π_m	4.655	519.574

Arnord and Press (1989b) gave the estimates of α using the conjugate independent priors and modified Lwin priors as 4.263 and 4.225, respectively. Our noninformaive Bayesian analysis gives larger values than those of Arnord and Press (1989b).

The marginal posterior probability densities of α and β are depicted in Figure 1 and Figure 2, respectively.

In Figure 1, we can see that the posterior mode of Jeffrey's prior is slightly larger than the the posterior mode under the other priors.

4 CONCLUDING REMARKS

We developed the noninformative priors for the Pareto distribution with scale and shape parameters.

When the parameter of interest is the shape parameter, Jeffrey's, reference and probability matching priors are developed. We showed that the reference prior also satisfies first order matching criterion. And there does not exist second order matching prior. We showed the propriety of posterior under proposed noninformative priors. And some simulation for comparing frequentist coverage probability showed that reference prior and first order matching prior matched the target coverage probability. The reference prior is slightly better than probability matching prior in terms of coverage probability when the sample size is moderate.

When the parameter of interest is the scale parameter, we developed reference prior and first order matching prior. This first order matching prior does not satisfy the second order matching criterion, and the reference prior does not give proper posterior distribution. Some simulation revealed that the coverage probability of first order matching prior matched the target coverage probability. But the Jeffrey's prior did not. We also proved the propriety of first order matching prior.

We developed the probability matching prior for the non-regular case of Pareto distribution, the posterior distribution under this prior was improper.

And we provided a real data example. We calculated the Bayes estimators, and depicted marginal posterior distributions.

As a consequence, we recommend the use of reference prior when the parameter of interest is the shape parameter, and the use of probability matching prior when the parameter of interest is the scale parameter.

References

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

36

Arnold, B.C. and Press, S.J. (1989a),Bayesian Inference and Decision Techniques, eds. P. Goel and A. Zellner, Amsterdam:North-Holland, pp. 157-173.

Arnold, B.C. and Press, S.J. (1989b),Bayesian Estimation and Prediction for Pareto Data, 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), *Bayesian Statistics IV*, J. M. Bernardo, et. al., Oxford University Press, Oxford, 35-60.

Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference (with discussion), *Journal of Royal Statistical Society*, B, **41**, 113-147.

Byer, D. (1981), Structural Probability Bounds for the Strong Pareto Law, Canadian Journal of Statistics, 9,71-77.

Cox, D. R. and Reid, N. (1987). Orthogonal Parameters and Approximate Conditional Inference (with discussion), *Journal of Royal Statistical Society*, B, **49**, 1-39.

Datta, G. S. and Ghosh, J. K. (1995a). On Priors Providing Frequentist Validity for Bayesian Inference, *Biometrika*, **82**, 37-45.

Datta, G. S. and Ghosh, M. (1995b). Some Remarks on Noninformative Priors, *Journal of the American Statistical Association*, **90**, 1357-1363.

Datta, G. S. and Ghosh, M. (1996). On the Invariance of Noninformative Priors, *The Annal of Statistics*, **24**, 141-159.

Geisser, S. (1984), Predicting 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

Ghosal, S. (1999), Probability Matching Priors for Non-Regular Cases, Biometrica, 86, 4, 956-964.

Lwin, T. (1972), Estimation of the Tail of the Paretian Law, Scandinavian Actuarial Journal, 55, 170-178.

Mukerjee, R. and Dey, D. K. (1993). Frequentist validity of Posterior Quantiles in the Presence of a Nuisance Parameter : Higher Order Asymptotics, *Biometrika*, **80**, 499-505.

Mukerjee, R. and Ghosh, M. (1997). Second Order Probability Matching Priors, *Biometrika*, **84**, 970-975.

Nigam, A.M. and Hamdy, H.L. (1987), Bayesian Prediction Bounds for the Pareto Lifetime Model, Communications in Statistics, 16, 1761-1772.

Peers, H. W. (1965). On Confidence Sets and Bayesian Probability Points in the Case of Several Parameters, *Journal of Royal Statistical Society*, B, **27**, 9-16.

Stein, C. (1985). On the Coverage Probability of Confidence Sets based on a Prior Distribution, *Sequential Methods in Statistics*, Banach Center Publications, **16**, 485-514.

Tibshirani, R. (1989). Noninformative Priors for One Parameter of Many, *Biometrika*, **76**, 604-608.

Tiwari, R.C. Yang, Y. and Zalkikar, J.N. (1996), Bayes Estimation for the Pareto Failure-Model Using Gibbs Sampling, IEEE Trans. Reliability, R-45,471-476.

Welch, B. N. and Peers, B. (1963). On Formulae for Confidence Points based on Integrals of Weighted Likelihood, *Journal of Royal Statistical Society*, **35**, 318-3298.
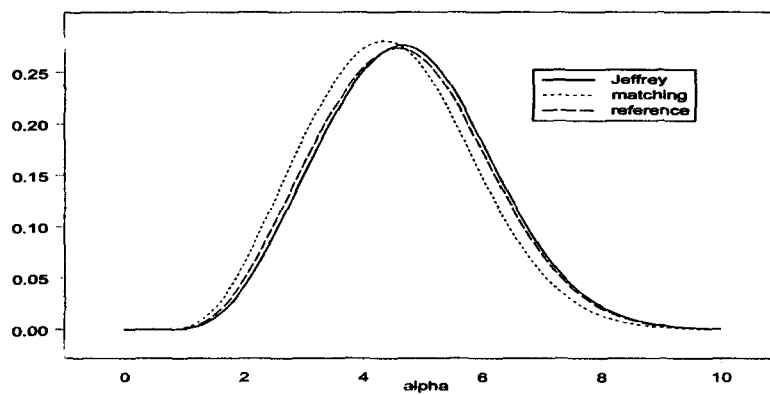
Figure 1: Marginal Posterior Distribution of α
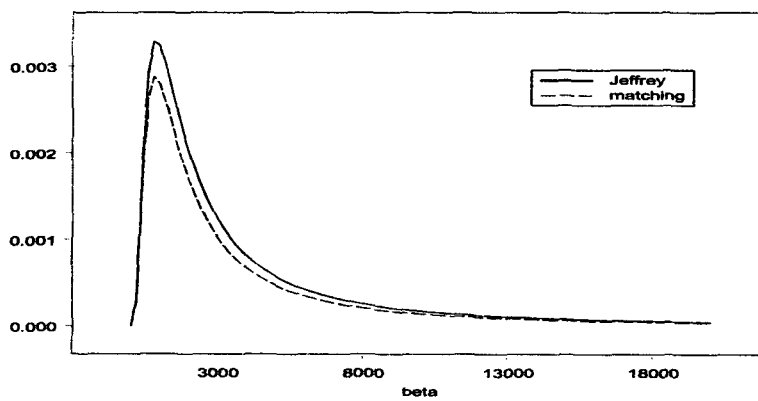
Figure 2: Marginal Posterior Distribution of β