Noninformative priors for Pareto distribution

Dal Ho Kim\textsuperscript{1}, Sang Gil Kang\textsuperscript{2}, Woo Dong Lee\textsuperscript{3}

\textsuperscript{1}Department of Statistics, Kyungpook National University
\textsuperscript{2}Department of Applied Statistics, Sangji University
\textsuperscript{3}Department of Asset Management, Daegu Haany University

Received 15 September 2009, revised 21 November 2009, accepted 23 November 2009

Abstract

In this paper, we develop noninformative priors for two parameter Pareto distribution. Specially, we derive Jeffreys' prior, probability matching prior and reference prior for the parameter of interest. In our case, the probability matching prior is only a first order matching prior and there does not exist a second order matching prior. Some simulation reveals that the matching prior performs better to achieve the coverage probability. A real example is also considered.

Keywords: Jeffreys' prior, Pareto distribution, probability matching prior, reference prior.

1. Introduction

The Pareto distribution has found applications in modeling 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 model such real life 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 is 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 and Press, 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) and 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 squared error loss and LINEX loss for estimating parameters using subjective priors such as conjugate prior and Gamma-exponential priors.

But there are situations when one is forced to use noninformative priors such as Jeffreys’, reference or matching priors because the prior information for the parameters may not be enough.

There is a great deal of efforts for finding noninformative or objective priors for various statistical models. Jeffreys’ prior was quite successful in many Bayesian inference, but it causes problems when the nuisance parameters are present.

Recently, significant advances are made in the development of noninformative priors via reference or probability matching priors. These noninformative priors work well in many statistical problems when the nuisance parameters are present.

Berger and Bernardo (1989, 1992) extended Bernardo (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 $o(n^{-1})$, 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 Jeffreys’ prior, reference prior and matching prior for the two-parameter Pareto distribution. We show that when the parameter of interest is given, 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. Development of noninformative priors

Let $X$ be the two-parameter Pareto Distribution with parameters $\alpha$ and $\beta$. The probability density function of $X$ is given by,

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

(2.1)
Noninformative priors for Pareto distribution

1215

where $\alpha$ is the shape parameter and $\beta$ is the scale parameter.

Including Jeffreys’ prior, we will find the probability matching prior and reference prior when the parameter of interest is $\alpha$ or $\beta$.

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

$$I(\alpha, \beta) = \begin{pmatrix} \frac{1}{\beta} & -\frac{1}{\beta(\alpha+1)} \\ -\frac{1}{\beta(\alpha+1)} & \frac{\alpha}{\beta^2(\alpha+2)} \end{pmatrix}. \quad (2.2)$$

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

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

Then the information matrix of $(\omega_1, \omega_2)$ per observation is given by

$$I(\omega_1, \omega_2) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\omega_1}{(\omega_1^2 + 2)^2} \end{pmatrix}. \quad (2.3)$$

From the above information matrix, one can find various noninformative priors, given below when the parameter of interest is $\alpha$.

The Jeffreys’ prior $\pi_J(\omega_1, \omega_2)$ is given by

$$\pi_J(\omega_1, \omega_2) \propto \frac{1}{\omega_2(1 + \omega_1)^2\sqrt{\omega_1(2 + \omega_1)}}, \quad \omega_1, \omega_2 > 0. \quad (2.4)$$

The class of first order matching priors is given by

$$\pi_M(\omega_1, \omega_2) \propto \frac{1}{\omega_1(1 + \omega_1)^2} g(\omega_2), \quad \omega_1, \omega_2 > 0, \quad (2.5)$$

where $g(\cdot) > 0$ is arbitrary differentiable function in its argument.

Finally, following Datta and Ghosh (1995b), the reference prior is given by

$$\pi_R(\omega_1, \omega_2) \propto \frac{1}{(1 + \omega_1) \omega_1 \omega_2}, \quad \omega_1, \omega_2 > 0. \quad (2.6)$$

**Remark 2.1** One can easily find the fact that the reference prior satisfies the first order matching criterion (Datta and Ghosh, 1995b).

The class of prior given in (2.6) 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 Mukerjee and Ghosh (1997). A second order probability matching prior for $\omega_1$ satisfies

$$\frac{1}{6}g(\omega_2) \frac{d}{d\omega_1} \left( i_{11}^{-3/2} L_{1,1,1} \right) + \frac{d}{d\omega_2} \left( i_{11}^{-1/2} L_{112} i_{22}^{1/2} g(\omega_2) \right) = 0, \quad (2.7)$$

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

$$L_{1,1,1} = E \left[ \left( \frac{\partial \log L}{\partial \omega_1} \right)^3 \right], \quad L_{112} = E \left[ \frac{\partial^3 \log L}{\partial \omega_1^2 \partial \omega_2} \right],$$

where $E$ represents the expectation.
and

\[ L(\omega_1, \omega_2) = \log \omega_1 + \omega_1 \log \omega_2 + 2\omega_1 \log \omega_1 - \omega_1 \log(1 + \omega_1) \]

\[-(\omega_1 + 1) \log(x + \frac{\omega_1^2 \omega_2}{1 + \omega_1}).\]

In our case,

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

and

\[ L_{1,12} = \frac{2}{\omega_1 \omega_2(1 + \omega_1)(2 + \omega_1)(3 + \omega_1)}. \]

Using the fact that \( i^{22} = (\omega_1 + 2)\omega_2^2 \omega_1 \), the equation (2.7) 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 any 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 \( \alpha \).

The priors (2.4),(2.5) and (2.6) can be re-expressed in terms of the original parameters \( \alpha \) and \( \beta \). For \( \alpha, \beta > 0 \),

\[ \pi_f(\alpha, \beta) \propto \frac{1}{\beta(1 + \alpha)\sqrt{\alpha(\alpha + 2)}} \]

\[ \pi_M(\alpha, \beta) \propto \frac{1}{\alpha^3} \]

\[ \pi_R(\alpha, \beta) \propto \frac{1}{\beta \alpha(1 + \alpha)}. \]

Now, we will prove the propriety of the posterior distributions under Jeffreys, reference and matching priors when the parameter of interest is \( \alpha \).

**Theorem 2.1** When the parameter of interest is \( \alpha \), the general form of the above noninformative priors can be written as

\[ \pi_{a,b,c,d}(\alpha, \beta) = \alpha^{-a} \beta^{-b}(\alpha + 1)^{-c}(\alpha + 2)^{-d}, \quad a > 0, \quad 0 \leq b \leq 1, \quad c \geq 0, \quad d \geq 0. \]

Then, the posterior under the general form of noninformative prior is finite, if (i) for \( 0 \leq b < 1, \quad n + b > 1, \quad a + c + d > 2 \) and \( n + 1 - a - c - d > 0 \) or (ii) for \( b = 1, \quad a + c + d > 1 \) and \( n - a - c - d > 0 \).

**Proof:** The joint posterior distribution under the above prior is

\[ \pi_{a,b,c,d}(\alpha, \beta | x) \propto \alpha^{-a} \beta^{-b}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \prod_{i=1}^{n} (x_i + \beta)^{-(\alpha + 1)}. \]
Let \( y = \text{minimum}(x_1, x_2, \cdots, x_n) \). Then
\[
\pi_{a,b,c,d}(\alpha, \beta|x) \leq \alpha^{-a} \beta^{a-b}(\alpha + 1)^{-c}(\alpha + 2)^{-d}(y + \beta)^{-n(\alpha+1)}.
\]
Hence,
\[
\int_0^\infty \int_0^\infty \pi_{a,b,c,d}(\alpha, \beta|x) \, d\beta \, d\alpha
\leq \int_0^\infty \int_0^\infty \alpha^{-a} \beta^{a-b}(\alpha + 1)^{-c}(\alpha + 2)^{-d}(y + \beta)^{-n(\alpha+1)} \, d\beta \, d\alpha
= \int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \int_0^\infty \beta^{a-b}(y + \beta)^{-n(\alpha+1)} \, d\beta \, d\alpha.
\]
Letting \( \frac{\beta}{y+\beta} = t \), the integration
\[
\int_0^\infty \beta^{a-b}(y + \beta)^{-n(\alpha+1)} \, d\beta = y^{b-n+1} \text{Beta}(n\alpha - b + 1, n + b - 1),
\]
if \( n + b - 1 > 0 \).
For the case: 0 \( \leq b < 1 \),
\[
\int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \int_0^\infty \beta^{a-b}(y + \beta)^{-n(\alpha+1)} \, d\beta \, d\alpha
= y^{b-n+1} \Gamma(n + b - 1) \int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(n\alpha - b + 1)}{\Gamma(n\alpha + n)} \, d\alpha
< y^{b-n+1} \Gamma(n + b - 1) \int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(n\alpha + 1)}{\Gamma(n\alpha + n)} \, d\alpha.
\]
Since,
\[
\Gamma(n\alpha + n) = (n\alpha + n - 1) \cdots (n\alpha + 1)\Gamma(n\alpha + 1),
\]
then,
\[
\int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(n\alpha + 1)}{\Gamma(n\alpha + n)} \, d\alpha = \int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(n\alpha + 1)}{(n\alpha + n - 1) \cdots (n\alpha + 1)} \, d\alpha.
\]
We know that
\[
(\alpha + 1)^{-c}(\alpha + 2)^{-d} < \alpha^{-c-d}
\]
and
\[
[(n\alpha + n - 1) \cdots (n\alpha + 1)]^{-1} < (n\alpha + 1)^{-(n-1)}.
\]
Therefore,
\[
\int_0^\infty \alpha^{-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(n\alpha + 1)}{\Gamma(n\alpha + n)} \, d\alpha
< \int_0^\infty \alpha^{-a-c-d}(n\alpha + 1)^{-(n-1)} \, d\alpha
= n^{a+c+d-n-1} \Gamma(a + c + d - 2) \Gamma(n + 1 - a - c - d) \Gamma(n-1) < \infty,
\]
Noninformative priors for Pareto distribution 1217
if $a + c + d > 2$ and $n + 1 - a - c - d > 0$.

For the case: $b = 1$,

$$
\int_0^\infty \alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \int_0^\infty \beta^{na-1}{y + \beta}^{-n(a+1)}d\beta d\alpha
\]

$$
= y^{-n+2} \int_0^\infty \alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \text{Beta}(na, n) d\alpha
\]

$$
= y^{-n+2} \Gamma(n) \int_0^\infty \alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(na)}{\Gamma(na + n)} d\alpha.
\]

Since,

$$
\Gamma(na + n) = (na + n - 1) \cdots (na) \Gamma(na),
\]

then,

$$
\int_0^\infty \alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(na)}{\Gamma(na + n)} d\alpha = \int_0^\infty \frac{\alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d}}{(na + n - 1) \cdots (na)} d\alpha.
\]

We know that

$$
(\alpha + 1)^{-c}(\alpha + 2)^{-d} < \alpha^{-c-d}
\]

and

$$
[(na + n - 1) \cdots (na + 1)]^{-1} < (na + 1)^{-(n-1)}.
\]

Therefore,

$$
\int_0^\infty \alpha^{n-a}(\alpha + 1)^{-c}(\alpha + 2)^{-d} \frac{\Gamma(na)}{\Gamma(na + n)} d\alpha < \int_0^\infty \frac{1}{n}(\alpha^{n-a-c-d-1}(\alpha + 1)^{-(n-1)}) d\alpha
\]

$$
= n^{a+c+d-n-1} \Gamma(a + c + d - 1) \Gamma(n - a - c - d) \Gamma(n - 1) < \infty,
\]

if $a + c + d > 1$ and $n - a - c - d > 0$. This completes the proof.

When the parameter of interest is $\beta$, we derive the probability matching prior, and reference prior. In this case, we know that Jeffreys’ prior is the same as the case when the parameter of interest is $\alpha$.

From the information matrix for $(\alpha, \beta)$ given in (2.2), following Peers (1965), the first order matching prior $\pi_m$ for $\beta$ is the solution of the following partial differential equation.

$$
\frac{\partial}{\partial \alpha} \alpha^{3/2} \sqrt{\alpha + 2} \pi_m + \frac{\partial}{\partial \beta} \frac{\beta(\alpha + 1) \sqrt{\alpha + 2}}{\sqrt{\alpha}} \pi_m = 0. \quad (2.8)
\]

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}}, \quad \alpha > 0, \beta > 0. \quad (2.9)
\]

This is the first order matching prior when the parameter of interest is $\beta$. 

Remark 2.2 We checked that whether the above first order matching prior (2.9) satisfies the second order matching criterion in Mukerjee and Dey (1993) or not, but, after long calculation, we found that it did not satisfy the criterion.

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

From the information matrix, one can obtain the reference prior for $\alpha$ with given $\beta$ as follows:

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

Choose a sequence of compact sets for $(\beta, \alpha)$ 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 $\alpha$ given $\beta$ is

$$\pi_i(\alpha|\beta) = \int_{k_{1i}}^{k_{2i}} \frac{1}{\alpha} \left( \frac{\alpha}{(\alpha + 1)^2(\alpha + 2)} \right) d\alpha \times \frac{1}{\beta}. $$

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

$$\pi_r(\beta, \alpha) = \lim_{i \to \infty} \pi_i(\beta) \pi_i(\alpha|\beta) = \frac{1}{\alpha \beta}. $$  \hspace{1cm} (2.10)

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

Now, we will show that the first order matching prior given in (2.9) gives a proper posterior distribution but the reference prior given in (2.10) does not.

Theorem 2.2 Under the prior $\pi_m(\alpha, \beta)$, the joint posterior distribution of $(\alpha, \beta)$ is proper if $n - \frac{3}{2} > 0$.

Proof: By Theorem 2.1., it is obvious. □

Theorem 2.3 Under the reference prior $\pi_r(\alpha, \beta)$, the joint posterior is improper.

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

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

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

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

$$= \frac{\Gamma(n)}{nz^n} \prod_{j=1}^{n} \Gamma(n\alpha + n)$$

$$> \frac{\Gamma(n)}{nz^n} \prod_{j=1}^{n} \Gamma(n\alpha + n)$$

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

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

$$= \frac{\Gamma(n)}{nz^n} \prod_{j=1}^{n} \Gamma(n\alpha + n)$$

$$> \frac{\Gamma(n)}{nz^n} \prod_{j=1}^{n} \Gamma(n\alpha + n)$$

Noninformative priors for Pareto distribution 1219
and,
\[
\frac{\Gamma(n)}{n^{z_n}} \int_0^\infty \alpha^{n-2}(n\alpha + n)^{-(n-1)} d\alpha = \frac{\Gamma(n)}{n^{z_n}} \int_0^1 k^{n-2}(1-k)^{-1} dk = \infty.
\]
So, the joint posterior distribution under reference prior is improper. \qed

### 3. Simulations and example

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 \(\alpha\), the coverage probabilities of the priors are given in Table 3.1.

In Table 3.1., under the moderate sample size, one can see that \(\pi_M\) and \(\pi_R\) match the target coverage probabilities’ 0.05 and 0.95, well. But the Jeffreys’ prior \(\pi_J\) does not. Specially, the coverage probability of the reference prior is better than other priors. This is because \(\pi_R\) satisfies the first order matching criterion.

<table>
<thead>
<tr>
<th>(n)</th>
<th>(\pi_M)</th>
<th>(\pi_J)</th>
<th>(\pi_R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0318000</td>
<td>0.0000000</td>
<td>0.9708000</td>
</tr>
<tr>
<td>20</td>
<td>0.0442000</td>
<td>0.0332000</td>
<td>0.9615000</td>
</tr>
<tr>
<td>30</td>
<td>0.0500000</td>
<td>0.0531000</td>
<td>0.9555000</td>
</tr>
<tr>
<td>40</td>
<td>0.0447000</td>
<td>0.0634000</td>
<td>0.9597000</td>
</tr>
<tr>
<td>50</td>
<td>0.0459000</td>
<td>0.0644000</td>
<td>0.9597000</td>
</tr>
</tbody>
</table>

In Table 3.2., when the parameter of interest is \(\beta\), the coverage probability of matching prior \(\pi_m\) matches the target coverage well. But the Jeffreys’ prior does not match the target coverage.

<table>
<thead>
<tr>
<th>(n)</th>
<th>(\pi_m)</th>
<th>(\pi_J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0025000</td>
<td>0.0105000</td>
</tr>
<tr>
<td>20</td>
<td>0.0209000</td>
<td>0.0415000</td>
</tr>
<tr>
<td>30</td>
<td>0.0348000</td>
<td>0.0585000</td>
</tr>
<tr>
<td>40</td>
<td>0.0456000</td>
<td>0.0679000</td>
</tr>
<tr>
<td>50</td>
<td>0.0499000</td>
<td>0.0672000</td>
</tr>
</tbody>
</table>

**Example 3.1** We will illustrate Bayesian analysis using proposed noninformative priors. The data given below are annual wage data. Arnold and Press (1989b) analyzed these data. Specially, Arnold 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 \(\alpha\) and \(\beta\) under the proposed priors are as follows:
Arnold and Press (1989b) gave the estimates of $\alpha$ using the conjugate independent priors and modified Lwin priors as 4.263 and 4.225, respectively. Our noninformative Bayesian analysis gives larger values than those of Arnold and Press (1989b).

The marginal posterior probability densities of $\alpha$ and $\beta$ are depicted in Figure 3.1 and Figure 3.2, respectively.

In Figure 3.1, we can see that the posterior mode of Jeffreys’ prior is slightly larger than the the posterior mode under the other priors.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_J$</td>
<td>4.741</td>
<td>526.731</td>
</tr>
<tr>
<td>$\pi_M$</td>
<td>4.461</td>
<td>518.253</td>
</tr>
<tr>
<td>$\pi_R$</td>
<td>4.671</td>
<td>520.897</td>
</tr>
<tr>
<td>$\pi_m$</td>
<td>4.655</td>
<td>519.574</td>
</tr>
</tbody>
</table>

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, Jeffreys, 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 Jeffreys’ prior did not. We also proved the propriety of first order matching prior.

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


