Noninformative priors for Pareto distribution

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

[1]Department of Statistics, Kyungpook National University
[2]Department of Applied Statistics, Sangji University
[3]Department of Asset Management, Daegu Haany University

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. S ome 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.

---

[1] Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea,

[2] Associate Professor, Department of Applied Statistics, Sangji University, Kwangwon 220-702, Korea.

[3] Corresponding author: Professor, Department of Asset Management , Daegu Haany University, Kyungpook 712-715, Korea. E-mail: wdlee@dhu.ac.kr

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 $o(n^{-\frac{1}{2}})$, 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 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 α and β. The probability density function of X is given by,

$$f(x; \alpha, \beta) = \alpha \beta^{\alpha} (x + \beta)^{-(\alpha+1)}, x > 0; (\alpha > 0, \beta > 0), \qquad (2.1)$$

where α is the shape parameter and β is the scale parameter.

Including Jeffreys' 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

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

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 of (ω_1, ω_2) per observation is given by

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

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

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

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

The class of first order matching priors is given by

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

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

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

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

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 ω_1 satisfies

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

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

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

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,

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

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 (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 α.

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

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

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

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

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

Theorem 2.1 When the parameter of interest is α, the general form of the above noninformative priors can be written as

$$\pi_{a,b,c,d}(\alpha, \beta) = \alpha^{-a}\beta^{-b}(\alpha + 1)^{-c}(\alpha + 2)^{-d}, \quad a > 0, \ 0 \le b \le 1, \ c \ge 0, \ d \ge 0.$$

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

Proof: The joint posterior distribution under the above prior is

$$\pi_{a,b,c,d}(\alpha, \beta | \mathbf{x}) \propto \alpha^{n-a}\beta^{n\alpha - b}(\alpha + 1)^{-c}(\alpha + 2)^{-d}\prod_{i=1}^{n}(x_i + \beta)^{-(\alpha+1)}.$$

Let $y = minimum(x_1, x_2, \cdots, x_n)$. Then

$$\pi_{a,b,c,d}(\alpha, \beta|\mathbf{x}) \leq \alpha^{n-a}\beta^{n\alpha-b}(\alpha+1)^{-c}(\alpha+2)^{-d}(y+\beta)^{-n(\alpha+1)}.$$

Hence,

$$\int_0^\infty \int_0^\infty \pi_{a,b,c,d}(\alpha, \beta|\mathbf{x})d\beta d\alpha$$

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

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

Letting $\frac{\beta}{y+\beta} = t$, the integration

$$\int_0^\infty \beta^{n\alpha-b}(y+\beta)^{-n(\alpha+1)}d\beta = y^{b-n+1}Beta(n\alpha-b+1, n+b-1),$$

if $n+b-1 > 0$.

For the case: $0 \leq b < 1$,

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

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

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

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

Since,

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

then,

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

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,

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

$$< \int_0^\infty \alpha^{n-a-c-d}(n\alpha+1)^{-(n-1)}d\alpha$$

$$= n^{a+c+d-n-1}\frac{\Gamma(a+c+d-2)\Gamma(n+1-a-c-d)}{\Gamma(n-1)} < \infty,$$

if $a + c + d > 2$ and $n + 1 - a - c - d > 0$.

For the case: $b = 1$,

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

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

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

Since,

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

then,

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

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,

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

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

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

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

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

From the information matrix for (α, β) given in (2.2), following Peers (1965), the first order matching prior π_m for β is the solution of the following partial differential equation.

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

A solution of the above partial differential equation is given by

$$\pi_m(\alpha, \beta) \propto \frac{1}{\beta\alpha^{3/2}\sqrt{\alpha+2}}, \quad \alpha > 0, \ \beta > 0. \tag{2.9}$$

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

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 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

$$\pi_r(\beta, \alpha) = \lim_{i \to \infty} \frac{\pi_i(\beta)\pi_i(\alpha|\beta)}{\pi_i(\beta_0)\pi_i(\alpha_0|\beta_0)} = \frac{1}{\alpha\beta}. \tag{2.10}$$

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 (α, β) is proper if $n - \frac{3}{2} > 0$.

Proof: By Theorem 2.1., it is obvious. $\square$

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

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

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. $\square$

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 α, 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 π_M and π_R match the target coverage probabilities' 0.05 and 0.95, well. But the Jeffreys' 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.

Table 3.1 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

In Table 3.2., when the parameter of interest is β, the coverage probability of matching prior π_m matches the target coverage well. But the Jeffreys' prior does not match the target coverage.

Table 3.2 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

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 α and β under the proposed priors are as follows:

	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

Arnold 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 Arnold and Press (1989b).

The marginal posterior probability densities of α and β 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.
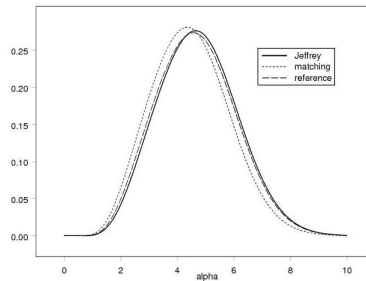

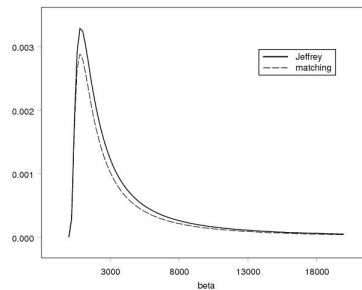**Figure 3.1** Marginal Posterior Distribution of α

Figure 3.2 Marginal Posterior Distribution of β

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

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. (1989a). *Bayesian inference and decision techniques*, Ed. P. Goel and A. Zellner, North-Holland, Amsterdam.

Arnold, B. C. and Press, S. J. (1989b). 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). *Bayesian Statistics IV*, Oxford University Press, Oxford.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of Royal Statistical Society*, Series B, **41,** 113-147.

Cox, D. R. and Reid, N. (1987). Orthogonal parameters and approximate conditional inference (with discussion). *Journal of Royal Statistical Society*, Series 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

Kang, S. G., Kim D. H. and Lee, W. D. (2008). Noninformative priors for the common mean of several inverse gaussian populations. *Journal of Korean Data and Information Science Society*, **19**, 401-411.

Lee, J. and Lee, W. D. (2008). Likelihood based inference for the shape parameter of Pareto distribution. *Journal of Korean Data and Information Science Society*, **19**, 1173-1181.

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, Theory and Methods*, **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*, Series B, **27,** 9-16.

Soliman, A. A. (2001). LINEX and quadratic approximate Bayes estimators applied to the Pareto model. *Communications in Statistics, Simulation and Computation*, **30**, 47-62.

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. L. and Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihood. *Journal of Royal Statistical Society*, Series B, **25**, 318-329.