Journal of the Korean Data & Information Science Society
2013, **24**(6), 1521–1529

http://dx.doi.org/10.7465/jkdi.2013.24.6.1521
한국데이터정보과학회지

Noninformative priors for the scale parameter in the generalized Pareto distribution†

Sang Gil Kang[1]

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

Received 1 October 2013, revised 23 October 2013, accepted 30 October 2013

Abstract

In this paper, we develop noninformative priors for the generalized Pareto distribution when the scale parameter is of interest. We developed the first order and the second order matching priors. We revealed that the second order matching prior does not exist. It turns out that the reference prior and Jeffrey's prior do not satisfy a first order matching criterion, and Jeffreys' prior, the reference prior and the matching prior are different. Some simulation study is performed and a real example is given.

Keywords: Generalized Pareto distribution, matching prior, reference prior, scale parameter.

1. Introduction

The generalized Pareto distribution (GPD) introduced independently by Pickands (1975) and Balkema and de Haan (1974) as a limiting distribution for scaled excesses over a high threshold has become a central notion in the statistical analysis of extreme events (Smith, 1987; Davison and Smith, 1990; Embrechts *et al.*, 1997). Applications of the GPD to areas such as insurance, reliability, finance, meteorology and environment are widely spread out in the literature.

The GPD is the distribution that contains uniform, exponential and Pareto distribution as special cases. The density function is

$$f(x|\xi,\sigma)=\begin{cases}\sigma^{-1}(1+x\xi/\sigma)^{-(1+\xi)/\xi}, & \xi\neq 0,\\ \sigma^{-1}\exp(-x/\sigma), & \xi=0,\end{cases} \tag{1.1}$$

where $\sigma>0$ is the scale parameter and ξ is the shape parameter. The shape parameter ξ plays an important role of determining the tail shape of the GPD. The support is $x>0$ for $\xi\geq 0$ and $0\leq x\leq -\sigma/\xi$, for $\xi<0$, respectively.

Estimation for GPD parameters is usually done via frequentist methods such as maximum likelihood estimation (see Davison and Smith, 1990), probability weighted moments by Hosking and Wallis (1987) and elemental percentile method by Castillo and Hadi (1997). In the context of estimating GPD parameters based on Bayesian methods, Arnold and Press (1989) explored the use of informative priors for Pareto distribution. Behrens *et al.* (2004)

† This research was supported by Sangji University Research Fund, 2012.
[1] Professor, Department of Computer and Data Information, Sangji University, Wonju 220-702, Korea.

proposed prior elicitation following Coles and Powell (1996). de Zea Bermudez and Amaral Turkman (2003) proposed the use of vague priors. Castellanos and Cabras (2007) proposed using Jeffreys' prior as a default procedure when there is no prior information. Recently, Ho (2010) proposed the matching prior via prediction criterion for the purpose of estimating high quantiles of GPD. Kang *et al.* (2013) derived the reference priors and the matching priors for the shape parameter of GPD.

Estimating upper quantiles in financial risk management is very important. Banks want to model and predict their loss in a certain period and use 99.9% value at risk (VaR) as a basis to compute the required capital. Simply, the 99.9% VaR implies the 99.9% quantile of the loss distribution. Financial loss data are commonly found to have heavy upper tails that correspond to $\xi > 0$ (see Ho, 2010). Thus we only consider GPD with $\xi > 0$ in this paper.

The present paper focuses on developing noninformative priors for the scale parameter of GPD with $\xi > 0$. In the absence of sources of information or past data, Bayesian methods rely on the objective priors or the noninformative priors. We consider Bayesian priors such that the resulting credible intervals for the scale parameter have coverage probabilities equivalent to their frequentist counterparts. Although this matching can be justified only asymptotically, our simulation results indicate that this is indeed achieved for small or moderate sample sizes as well.

This matching idea goes back to Welch and Peers (1963). Interest in such priors revived with the work of Stein (1985) and Tibshirani (1989). Among others, we may cite the work of Mukerjee and Dey (1993), Datta and Ghosh (1995, 1996), Mukerjee and Ghosh (1997). On the other hand, Bernardo (1979) introduced the reference priors which maximizes the Kullback-Leibler divergence between the prior and the posterior. Ghosh and Mukerjee (1992) and Berger and Bernardo (1989,1992) gave a general algorithm to derive a reference prior by splitting the parameters into several groups according to their order of inferential importance. This approach is very successful in various practical problems (Kang, 2011; Kang *et al.*, 2012, 2013). Quite often reference priors satisfy the matching criterion described earlier.

The outline of the remaining sections is as follows. In Section 2, we develop the first order probability matching priors for the scale parameter. Next we derive Fisher information matrix, and also derive the reference priors for the scale parameter. It turns out that Jeffreys' prior, the reference priors and the first order matching priors are different each other. In Section 3, we provide that the propriety of the posterior distribution for a general class of prior distributions which include the reference priors as well as the first order matching prior. In Section 4, simulated frequentist coverage probabilities under the proposed priors are given. A real example is given.

2. The noninformative priors

2.1. The probability matching priors

For a prior π, let $\theta_1^{1-\alpha}(\pi;\mathbf{X})$ denote the $(1-\alpha)$th quantile of the posterior distribution of θ_1, that is,

$$P^{\pi}[\theta_1 \le \theta_1^{1-\alpha}(\pi;\mathbf{X})|\mathbf{X}] = 1-\alpha, \tag{2.1}$$

where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_t)^T$ and θ_1 is the parameter of interest. We want to find priors π for which

$$P_\theta[\theta_1 \le \theta_1^{1-\alpha}(\pi; \mathbf{X})] = 1 - \alpha + o(n^{-u}). \tag{2.2}$$

for some $u > 0$, as n goes to infinity. Priors π satisfying (2.2) are called matching priors. If $u = 1/2$, then π is referred to as a first order matching prior, while if $u = 1$, π is referred to as a second order matching prior.

In order to find such matching priors π, let

$$\theta_1 = \sigma \text{ and } \theta_2 = \sigma \exp\{2\xi\}.$$

With this parametrization, the likelihood function of parameters (θ_1, θ_2) for the model (1.1) is given by

$$L(\theta_1, \theta_2) \propto \frac{1}{\theta_1}\left[1 + \frac{\log(\theta_2/\theta_1)}{2\theta_1}x\right]^{-\frac{2+\log(\theta_2/\theta_1)}{\log(\theta_2/\theta_1)}}. \tag{2.3}$$

Based on (2.3), the Fisher information matrix is given by

$$\mathbf{I}(\theta_1, \theta_2) = \begin{pmatrix} \frac{1}{\theta_1^2[2+\log(\theta_2/\theta_1)]} & 0 \\ 0 & \frac{1}{\theta_2^2[1+\log(\theta_2/\theta_1)][2+\log(\theta_2/\theta_1)]} \end{pmatrix}. \tag{2.4}$$

From the above Fisher information matrix $\mathbf{I}$, θ_1 is orthogonal to θ_2 in the sense of Cox and Reid(1987). Following Tibshirani(1989), the class of first order probability matching prior is characterized by

$$\pi_m^{(1)}(\theta_1, \theta_2) \propto \theta_1^{-1}[2 + \log(\theta_2/\theta_1)]^{-\frac{1}{2}} d(\theta_2), \tag{2.5}$$

where $0 < \theta_1 < \theta_2 < \infty$ and $d(\theta_2) > 0$ is an arbitrary function differentiable in its argument.

The class of prior given in (2.5) can be narrowed down to the second order probability matching priors as given in Mukerjee and Ghosh (1997). A second order probability matching prior is of the form (2.5), and also d must satisfy an additional differential equation (2.10) of Mukerjee and Ghosh (1997), namely

$$\frac{1}{6}d(\theta_2)\frac{\partial}{\partial\theta_1}\{I_{11}^{-\frac{3}{2}}L_{1,1,1}\} + \frac{\partial}{\partial\theta_2}\{I_{11}^{-\frac{1}{2}}L_{112}I^{22}d(\theta_2)\} = 0, \tag{2.6}$$

where

$$L_{1,1,1} = E\left[\left(\frac{\partial \log L}{\partial\theta_1}\right)^3\right] = \frac{\log(\theta_2/\theta_1)[1 - 2\log(\theta_2/\theta_1)] - 2}{\theta_1^3[2 + \log(\theta_2/\theta_1)]^2[2 + 3\log(\theta_2/\theta_1)]},$$

$$L_{112} = E\left[\frac{\partial^3 \log L}{\partial\theta_1^2\partial\theta_2}\right] = \theta_1^{-2}\theta_2^{-1}[1 + \log(\theta_2/\theta_1)]^{-1}[2 + 3\log(\theta_2/\theta_1)]^{-1},$$

$$I_{11} = \frac{1}{\theta_1^2[2 + \log(\theta_2/\theta_1)]}, I^{22} = \theta_2^2[1 + \log(\theta_2/\theta_1)][2 + \log(\theta_2/\theta_1)].$$

Then (2.6) simplifies to

$$\frac{1}{4}g_1(\theta_1,\theta_2) + \frac{1}{d(\theta_2)}\frac{\partial}{\partial\theta_2}\{g_2(\theta_1,\theta_2)d(\theta_2)\} = 0, \tag{2.7}$$

where

$$\begin{aligned} g_1(\theta_1,\theta_2) &= \theta_1^{-1}[2+\log(\theta_2/\theta_1)]^{-\frac{3}{2}}[2+3\log(\theta_2/\theta_1)]^{-2} \\ &\times \{-12+4\log(\theta_2/\theta_1)+13[\log(\theta_2/\theta_1)]^2+2[\log(\theta_2/\theta_1)]^3\} \\ g_2(\theta_1,\theta_2) &= \theta_1^{-1}\theta_2[2+\log(\theta_2/\theta_1)]^{\frac{3}{2}}[2+3\log(\theta_2/\theta_1)]^{-1}. \end{aligned}$$

Note that the first term in the left hand side of (2.7) depends only on the model, whereas the second term involves the prior. Thus there can be no solution to (2.7) unless $d(\theta_2)$ is a function of θ_1 and θ_2. Therefore this rules out the existence of the second order matching priors.

2.2. The reference priors

Reference priors introduced by Bernardo (1979), and extended further by Berger and Bernardo (1992) have become very popular over the years for the development of noninformative priors. In this Section, we derive the reference priors for different groups of ordering of (σ,ξ) by the algorithm of Berger and Bernardo (1992).

From the likelihood (1.1), the Fisher information matrix is given by

$$I(\sigma,\xi) = \begin{pmatrix} \frac{1}{\sigma^2(1+2\xi)} & \frac{1}{\sigma(1+\xi)(1+2\xi)} \\ \frac{1}{\sigma(1+\xi)(1+2\xi)} & \frac{2}{(1+\xi)(1+2\xi)} \end{pmatrix}. \tag{2.8}$$

Firstly, we derive the reference prior for the parameter grouping $\{\sigma,\xi\}$. The compact subsets were taken to be Cartesian products of sets of the form

$$\sigma \in [a_1,b_1], \xi \in [a_2,b_2]. \tag{2.9}$$

In the limit $a_i, i=1,2$ will tend to 0, and $b_i, i=1,2,$ will tend to ∞. For the derivation of the reference prior, from the Fisher information (2.8),

$$h_1 = \frac{1}{2\sigma^2(1+\xi)} \text{ and } h_2 = \frac{2}{(1+\xi)(1+2\xi)}.$$

Here, and below, a subscripted Q denotes a function that is constant and does not depend on any parameters but any Q may depend on the ranges of the parameters.
Step 1. Note that

$$\int_{a_2}^{b_2} h_2^{1/2} d\xi = \int_{a_2}^{b_2}\left[\frac{2}{(1+\xi)(1+2\xi)}\right]^{1/2} d\xi = \sqrt{2}Q_1$$

It follows that

$$\pi_2^l(\xi|\sigma) = Q_1^{-1}(1+\xi)^{-\frac{1}{2}}(1+2\xi)^{-\frac{1}{2}}.$$

Step 2. Now

$$\begin{aligned} E^l\{\log h_1|\xi\} &= \int_{a_2}^{b_2} Q_1^{-1}(1+\xi)^{-\frac{1}{2}}(1+2\xi)^{-\frac{1}{2}} \log\left[\frac{1}{2\sigma^2(1+\xi)}\right] d\xi \\ &= Q_2 - \log\sigma^2. \end{aligned}$$

It follows that

$$\pi_1^l(\sigma) \propto \exp[E^l\{\log h_1|\xi\}/2] = \exp\{Q_2/2\}\sigma^{-1}.$$

Therefore the reference prior is

$$\pi_1(\sigma,\xi) = \lim_{l\to\infty} \frac{\pi_2^l(\xi|\sigma)\pi_1^l(\sigma)}{\pi_2^l(\xi_0|\sigma_0)\pi_1^l(\sigma_0)} \propto \sigma^{-1}(1+\xi)^{-\frac{1}{2}}(1+2\xi)^{-\frac{1}{2}}, \tag{2.10}$$

where σ_0 and ξ_0 are an inner point of the interval $(0,\infty)$.

Notice that the matching priors (2.5) include many different matching priors because of the arbitrary selection of the function d. And for some functions, there does not seem to be any improvement in the coverage probabilities with these posteriors. So we consider a particular first order matching prior where d is θ_2^{-1} in matching priors (2.5). This prior is given by

$$\pi_m(\theta_1,\theta_2) \quad\propto\quad \theta_1^{-1}\theta_2^{-1}[2+\log(\theta_2/\theta_1)]^{-\frac{1}{2}}. \tag{2.11}$$

Remark 2.1 Due to invariance of the matching prior and the reference prior (Berger and Bernardo, 1992; Datta and Ghosh, 1996; Mukerjee and Ghosh, 1997), Jeffreys' prior π_J, the reference prior π_r, and the matching prior π_m in the original parametrization (σ,ξ) are given by

$$\pi_J(\sigma,\xi) \quad\propto\quad \sigma^{-1}(1+\xi)^{-1}(1+2\xi)^{-\frac{1}{2}}, \tag{2.12}$$

$$\pi_r(\sigma,\xi) \quad\propto\quad \sigma^{-1}(1+\xi)^{-\frac{1}{2}}(1+2\xi)^{-\frac{1}{2}}, \tag{2.13}$$

$$\pi_m(\sigma,\xi) \quad\propto\quad \sigma^{-1}(1+\xi)^{-\frac{1}{2}}, \tag{2.14}$$

respectively and Jeffreys', the reference and the matching priors are different each other.

3. Propriety of the posterior distribution

We investigate the propriety of posteriors for a general class of priors which include Jeffreys' prior (2.12), the reference prior (2.13) and the matching prior (2.14). We consider the class of priors

$$\pi(\sigma,\xi) \propto \sigma^{-1}(1+\xi)^{-a}(1+2\xi)^{-b}, \tag{3.1}$$

where $a > 0$ and $b \geq 0$. The following general theorem can be proved.

Theorem 3.1 The posterior distribution of (σ,ξ) under the prior (3.1) is proper if $n+a+b-2 > 0$.

Proof: Note that the joint posterior for σ and ξ given $\mathbf{x}$ is

$$\pi(\sigma,\xi|\mathbf{x}) \quad \propto \quad \sigma^{-n-1}(1+\xi)^{-a}(1+2\xi)^{-b}\prod_{i=1}^{n}\left(1+\frac{\xi}{\sigma}x_i\right)^{-\frac{1+\xi}{\xi}}. \tag{3.2}$$

Then we obtain

$$\pi(\sigma,\xi|\mathbf{x}) \quad \leq \quad \sigma^{-n-1}(1+\xi)^{-a}(1+2\xi)^{-b}\left(1+\frac{\xi}{\sigma}z\right)^{-n\frac{1+\xi}{\xi}} \equiv \pi'(\sigma,\xi|\mathbf{x}), \tag{3.3}$$

where $z=\min\{x_1,\cdots,x_n\}$. Integrating with respect to σ from (3.3), then we get

$$\begin{aligned}
\pi'(\xi|\mathbf{x}) \quad &\propto \quad \xi^{-n}(1+\xi)^{-a}(1+2\xi)^{-b}\frac{\Gamma\left[\frac{n}{\xi}\right]}{\Gamma\left[n+\frac{n}{\xi}\right]} \\
&\propto \quad \frac{\xi^{-n}(1+\xi)^{-a}(1+2\xi)^{-b}}{\frac{n}{\xi}\prod_{i=1}^{n-1}(i+\frac{n}{\xi})} \\
&\leq \quad c_1(n+\xi)^{-(n-1)}(1+\xi)^{-a}(1+2\xi)^{-b} \\
&\leq \quad c_2(1+\xi)^{-(n-1)-a-b}. \qquad (3.4)
\end{aligned}$$

Here c_1 and c_2 are a constants. Therefore the posterior is proper if $n+a+b-2>0$. This completes the proof. □

Theorem 3.2 Under the prior (3.1), the marginal posterior density of σ is given by

$$\pi(\sigma|\mathbf{x}) \quad \propto \quad \int_0^\infty \sigma^{-n-1}(1+\xi)^{-a}(1+2\xi)^{-b}\prod_{i=1}^{n}\left(1+\frac{\xi}{\sigma}x_i\right)^{-\frac{1+\xi}{\xi}} d\xi. \tag{3.5}$$

Note that actually, normalizing constant for the marginal density of σ required two dimensional integration. Therefore we have the marginal posterior density of σ, and so it is easy to compute the marginal moment of σ. In Section 4, we investigate the frequentist coverage probabilities for the π_J, π_r and π_m respectively.

4. Numerical studies

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posteriors density of σ under the noninformative prior π given in (3.1) for several configurations σ,ξ and n. That is to say, the frequentist coverage of a $(1-\alpha)$th posterior quantile should be close to $1-\alpha$. This is done numerically. Table 4.1 gives numerical values of the frequentist coverage probabilities of 0.05 (0.95) posterior quantiles for the our prior. The computation of these numerical values is based on the following algorithm for any fixed true (σ,ξ) and any prespecified probability value α. Here α is 0.05 (0.95). Let $\sigma^\alpha(\pi;\mathbf{X})$ be the αth posterior quantile of σ given $\mathbf{X}$. That is, $F(\sigma^\alpha(\pi;\mathbf{X})|\mathbf{X})=\alpha$, where $F(\cdot|\mathbf{X})$ is the marginal posterior distribution of σ. Then the frequentist coverage probability of this one sided credible interval of σ is

$$P_{(\sigma,\xi)}(\alpha;\sigma) \quad = \quad P_{(\sigma,\xi)}(0<\sigma\leq\sigma^\alpha(\pi;\mathbf{X})). \tag{4.1}$$

Table 4.1 Frequentist coverage probability of 0.05 (0.95) posterior quantiles of σ

σ	ξ	n	π_J	π_r	π_m
0.5	0.5	10	0.029 (0.950)	0.022 (0.945)	0.015 (0.936)
		20	0.041 (0.957)	0.034 (0.952)	0.026 (0.944)
		30	0.051 (0.958)	0.045 (0.954)	0.036 (0.949)
		40	0.053 (0.960)	0.047 (0.956)	0.039 (0.950)
	1.0	10	0.056 (0.959)	0.044 (0.954)	0.029 (0.946)
		20	0.067 (0.960)	0.057 (0.956)	0.045 (0.950)
		30	0.067 (0.959)	0.060 (0.955)	0.050 (0.951)
		40	0.062 (0.958)	0.056 (0.955)	0.049 (0.951)
	3.0	10	0.069 (0.961)	0.055 (0.958)	0.043 (0.954)
		20	0.066 (0.960)	0.057 (0.957)	0.051 (0.953)
		30	0.056 (0.956)	0.051 (0.953)	0.045 (0.949)
		40	0.055 (0.957)	0.051 (0.955)	0.047 (0.952)
1.0	0.5	10	0.027 (0.951)	0.022 (0.945)	0.015 (0.936)
		20	0.041 (0.957)	0.035 (0.952)	0.026 (0.946)
		30	0.050 (0.961)	0.044 (0.957)	0.036 (0.951)
		40	0.058 (0.958)	0.053 (0.955)	0.045 (0.949)
	1.0	10	0.052 (0.957)	0.041 (0.952)	0.028 (0.943)
		20	0.069 (0.963)	0.059 (0.957)	0.048 (0.950)
		30	0.064 (0.960)	0.057 (0.956)	0.047 (0.951)
		40	0.061 (0.958)	0.056 (0.955)	0.048 (0.951)
	3.0	10	0.072 (0.960)	0.059 (0.956)	0.047 (0.952)
		20	0.063 (0.954)	0.057 (0.951)	0.049 (0.948)
		30	0.060 (0.959)	0.055 (0.955)	0.048 (0.953)
		40	0.060 (0.955)	0.056 (0.953)	0.052 (0.949)
5.0	0.5	10	0.031 (0.944)	0.024 (0.937)	0.017 (0.927)
		20	0.039 (0.959)	0.034 (0.955)	0.025 (0.947)
		30	0.051 (0.958)	0.045 (0.954)	0.036 (0.948)
		40	0.056 (0.959)	0.049 (0.955)	0.042 (0.946)
	1.0	10	0.054 (0.959)	0.043 (0.953)	0.029 (0.945)
		20	0.069 (0.961)	0.058 (0.958)	0.048 (0.952)
		30	0.062 (0.957)	0.056 (0.953)	0.045 (0.948)
		40	0.063 (0.960)	0.056 (0.957)	0.049 (0.950)
	3.0	10	0.074 (0.963)	0.062 (0.960)	0.050 (0.955)
		20	0.069 (0.954)	0.063 (0.951)	0.054 (0.948)
		30	0.057 (0.959)	0.053 (0.957)	0.046 (0.953)
		40	0.058 (0.957)	0.054 (0.956)	0.050 (0.953)

The computed $P_{(\sigma,\xi)}(\alpha;\sigma)$ when $\alpha = 0.05(0.95)$ is shown in Table 4.1. In particular, for fixed n and (σ,ξ), we take $10,000$ independent random samples of $\mathbf{X} = (X_1,\cdots,X_n)$ from the generalized Pareto population.

In Table 4.1, we can observe that the reference prior π_r and the matching prior π_m meet well the target coverage probabilities than Jeffreys' prior π_J as the sample size increase. Also note that the results of table are not much sensitive to the change of the values of (σ,ξ). Thus we recommend to use the matching prior and the reference prior.

Example 4.1 This example taken from Giles *et al.* (2011), and involves American Insurance Association data relating insurance losses in excess of 5,000,000 (in 1981 dollars) due to major hurricanes between 1949 and 1980. The data are provided by Hogg and Klugman (1983), and we have subtracted 5 million from each of their sample values, so allowing for the scale of the data reported by those authors, our first datum is 1766.0, etc.

The maximum likelihood estimator (MLE), the Bayes estimates based on Jeffreys' prior, the reference prior and the matching prior for the scale parameter are given by

$$\hat{\sigma}_{MLE} = 68002.58, \hat{\sigma}_J = 73110.09, \hat{\sigma}_r = 71711.36, \hat{\sigma}_m = 71196.57,$$

respectively. The Bayes estimates based on the reference prior and the matching prior give the similar results, and the Bayes estimate based on Jeffreys' prior is larger than other estimates.

5. Concluding remarks

In the generalized Pareto distribution, we have found a prior which is the first and the second order matching priors, and the reference priors for the scale parameter. We revealed the second order matching prior does not exist. It turns out that Jeffreys' prior, the reference prior and the matching prior are different each other. As illustrated in our numerical study, the reference prior and the matching prior seem to be the appropriate results than Jeffreys' prior in the sense of asymptotic frequentist coverage property.

References

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

Balkema, A. A. and de Haan, L. (1974). Residual lifetime at great age. *The Annals of Probability*, **2**, 792-804.

Behrens, C., Lopes, H. F. and Gamerman, D. (2004). Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, **4**, 227-244.

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.

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

Castellanos, M. E. and Cabras, A. (2007). A default Bayesian procedures for the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, **137**, 473-483.

Castillo, E. and Hadi, A. (1997). Fitting the generalized Pareto distribution to data. *Journal of the American Statistical Association*, **92**, 1609-1620.

Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modeling: A review and new developments. *International Statistical Review*, **64**, 119-136.

Datta, G. S. and Ghosh, M. (1995). 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 Annals of Statistics*, **24**, 141-159.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of Royal Statistical Society B*, **52**, 393-442.

de Zea Bermudez, P. and Amaral Turkman, M. A. (2003). Bayesian approach to parameter estimation of the generalized Pareto distribution. *Test*, **12**, 259-277.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling extremal events for insurance and finance*, Springer, Berlin.

Giles, D. E., Feng, H. and Godwin R. T. (2011). *On the bias of the maximum likelihood estimators for the two-parameter Lomax distribution*, Econometrics Working Paper EWP1104, Department of Economics, University of Victoria.

Ghosh, J. K. and Mukerjee, R. (1992). Noninformative priors (with discussion). In *Bayesian Statistics IV*, edited by J. M. Bernardo, et al., Oxford University, Oxford, 195-210.

Ho, K. (2010). A matching prior for extreme quantile estimation of the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, **140**, 1513-1518.

Hogg, R. V. and Klugman, S. A. (1983). On estimation of long-tailed skewed distributions with actuarial applications. *Journal of Econometrics*, **23**, 91-102.

Hosking, J. R. M. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, **29**, 339-349.

Kang, S. G. (2011). Noninformative priors for the common mean in log-normal distributions. *Journal of the Korean Data & Information Science Society*, **22**, 1241-1250.

Kang, S. G., Kim, D. H. and Lee, W. D. (2012). Noninformative priors for the ratio of the scale parameters in the half logistic distributions. *Journal of the Korean Data & Information Science Society*, **23**, 833-841.

Kang, S. G., Kim, D. H. and Lee, W. D. (2013). Noninformative priors for the shape parameter in the generalized Pareto distribution. *Journal of the Korean Data & Information Science Society*, **24**, 171-178.

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. *Journal of the American Statistical Association*, **59**, 665-680.

Pickands, J. (1975). Statistical inferences using extreme order statistics. *The Annals of Statistics*, **3**, 119-131.

Smith, R. L. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, **15**, 1174-1207.

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.

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