POSTERIOR PROPRIETY IN BAYESIAN EXTREME VALUE ANALYSES USING REFERENCE PRIORS

Paul J. Northrop and Nicolas Attalides

University College London

Abstract: The Generalized Pareto (GP) and Generalized extreme value (GEV) distributions play an important role in extreme value analyses as models for threshold excesses and block maxima, respectively. For each of these distributions we consider Bayesian inference using "reference" prior distributions (in the general sense of priors constructed using formal rules) for the model parameters, specifically a Jeffreys prior, the maximal data information (MDI) prior and independent uniform priors on separate model parameters. We investigate whether these improper priors lead to proper posterior distributions. We show that, in the GP and GEV cases, the MDI prior, unless modified, never yields a proper posterior and that in the GEV case this also applies to the Jeffreys prior. We also show that a sample size of three (four) is sufficient for independent uniform priors to yield a proper posterior distribution in the GP (GEV) case.

Key words and phrases: Extreme value theory, generalized extreme value distribution, generalized Pareto distribution, posterior propriety, reference prior.

1. Introduction

Extreme value theory provides an asymptotic justification for particular families of models for extreme data. Let $X_1, \ldots X_N$ be a sequence of independently and identically distributed random variables. Let u_N be a threshold, increasing with N. Pickands (1975) showed that if there is a non-degenerate limiting distribution for appropriately linearly rescaled excesses of u_N then this limit is a Generalized Pareto (GP) distribution. In practice, a suitably high threshold u is chosen empirically. Given that there is an exceedance of u, the excess $Z = X - u$ is modelled by a $\text{GP}(\sigma_u, \xi)$ distribution, with threshold-dependent scale parameter σ_u, shape parameter ξ and distribution function

$$F_{GP}(z) = \begin{cases} 1 - \left(1 + \frac{\xi z}{\sigma_u}\right)_+^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-\frac{z}{\sigma_u}), & \xi = 0, \end{cases} \tag{1.1}$$

where $z > 0$, $z_+ = \max(z, 0)$, $\sigma_u > 0$ and $\xi \in \mathbb{R}$. The use of the generalized extreme value (GEV) distribution Jenkinson (1955), with distribution function

$$F_{GEV}(y) = \begin{cases} \exp\left\{-\left[1 + \frac{\xi(y-\mu)}{\sigma}\right]_+^{-1/\xi}\right\}, & \xi \neq 0, \\ \exp\left\{-\exp[-\frac{(y-\mu)}{\sigma}]\right\}, & \xi = 0, \end{cases} \qquad (1.2)$$

where $\sigma > 0$ and $\mu, \xi \in \mathbb{R}$, as a model for block maxima is motivated by considering the behaviour of $Y = \max\{X_1, \ldots, X_b\}$ as $b \to \infty$ Fisher and Tippett (1928); Leadbetter, Lindgren, and Rootzén (1983).

Commonly-used frequentist methods of inference for extreme value distributions are maximum likelihood estimation (MLE) and probability-weighted moments (PWM). However, conditions on ξ are required for the asymptotic theory on which inferences are based to apply: $\xi > -1/2$ for MLE Smith (1984, 1985) and $\xi < 1/2$ for PWM Hosking, Wallis, and Wood (1985); Hosking and Wallis (1987). Alternatively, a Bayesian approach Coles (2001); Coles and Powell (1996); Stephenson and Tawn (2004) can avoid conditions on the value of ξ and performs predictive inference about future observations naturally and conveniently using Markov chain Monte Carlo (MCMC) output. A distinction can be made between *subjective* analyses, in which the prior distribution supplies information from an expert Coles and Tawn (1996) or more general experience of the quantity under study Martins and Stedinger (2000, 2001), and so-called *objective* analyses Berger (2006). In the latter, a prior is constructed using a formal rule, for use when no subjective information is to be incorporated into the analysis. There is disagreement about appropriate terminology for such priors: we follow Kass and Wasserman (1996) in using the term *reference prior*.

Many formal rules have been proposed; Kass and Wasserman (1996) provide a comprehensive review. We consider three priors that have been used in extreme value analyses: the Jeffreys prior Eugenia Castellanos and Cabras (2007); Beirlant et al. (2004), the maximal data information (MDI) prior Beirlant et al. (2004), and the uniform prior Pickands (1994). These priors are not proper probability distributions. An improper prior can lead to an improper posterior, which is clearly undesirable. There is no general theory providing simple conditions under which an improper prior yields a proper posterior for a particular model, so this must be investigated case-by-case. Eugenia Castellanos and Cabras (2007) establish that Jeffreys prior for the GP distribution always yields a proper posterior, but no such results exist for the other improper priors we consider. It is important that posterior propriety is established because impropriety may not create obvious numerical problems; for example, MCMC output may appear perfectly reasonable Athreya and Roy (2014).

One way to ensure posterior propriety is to use a diffuse proper prior, such as a normal prior with a large variance Coles and Tawn (2005); Smith (2005) or by truncating an improper prior Smith and Goodman (2000). For example,

Coles (2001, chapter 9) uses a GEV(μ, σ, ξ) model for annual maximum sea-levels, placing independent normal priors on μ, $\log \sigma$, and ξ with respective variances $10^4, 10^4$, and 100. However, one needs to check that the posterior is not sensitive to the choice of proper prior and, as Bayarri and Berger (2004) note "...these posteriors will essentially be meaningless if the limiting improper objective prior would have resulted in an improper posterior distribution." Therefore, independent uniform priors on separate model parameters are of interest in their own right and as the limiting case of independent diffuse normal priors.

In Section 2 we give the general form of the three priors we consider. In Section 3 we investigate whether or not these priors yield a proper posterior distribution given a random sample $\boldsymbol{z} = (z_1, \ldots, z_m)$ from the GP distribution and, in cases where propriety is possible, we derive sufficient conditions for this to occur. We repeat this for a random sample $\boldsymbol{y} = (y_1, \ldots, y_n)$ from a GEV distribution in Section 4. In Section 5 we discuss some implications of these results and possible extensions. Proofs of results are presented in the appendix.

2. Reference Priors for Extreme Value Distributions

Let Y be a random variable with density function $f(Y \mid \phi)$, indexed by a parameter vector ϕ, and define the Fisher information matrix $I(\phi)$ by $I(\phi)_{ij} = \mathrm{E}\left[-\partial^2 \ln f(Y \mid \phi)/\partial\phi_i\partial\phi_j\right]$.

Uniform priors. Priors that are flat suffer from the problem that they are not automatically invariant to reparameterisation; for example, if we give $\log \sigma$ a uniform distributon then σ is not uniform. Thus, it matters which particular parameterization is used to define the prior.

Jeffreys priors. Jeffreys' "general rule" Jeffreys (1961) is

$$\pi_J(\phi) \propto \det(I(\phi))^{1/2}. \tag{2.1}$$

An attractive property of this rule is that it produces a prior that is invariant to reparameterization. Jeffreys suggested a modification of this rule for use in location-scale problems. We follow this modification, which is summarised on page 1345 of Kass and Wasserman (1996). If there is no location parameter then (2.1) is used. If there is a location parameter μ, say, then $\phi = (\mu, \theta)$ and

$$\pi_J(\mu, \theta) \propto \det(I(\theta))^{1/2}, \tag{2.2}$$

where $I(\theta)$ is calculated holding μ fixed. In the current context the GP distribution does not have a location parameter whereas the GEV distribution does.

MDI prior. The MDI prior Zellner (1971) is defined as

$$\pi_M(\phi) \propto \exp\left\{\mathrm{E}[\log f(Y \mid \phi)]\right\}. \tag{2.3}$$

This is the prior for which the increase in average information, provided by the data via the likelihood function, is maximised. For further information, see Zellner (1998).

3. Generalized Pareto (GP) Distribution

Without loss of generality we take the m threshold excesses to be ordered; $z_1 < \cdots < z_m$. For simplicity we denote the GP scale parameter by σ rather than σ_u. We consider a class of priors of the form $\pi(\sigma, \xi) \propto \pi(\xi)/\sigma, \sigma > 0, \xi \in \mathbb{R}$, where $\pi(\xi)$ is a function depending only on ξ; thus, *a priori* σ and ξ are independent and $\log \sigma$ has an improper uniform prior over the real line.

The posterior is given by

$$\pi_G(\sigma, \xi \mid \boldsymbol{z}) = C_m^{-1} \pi(\xi)\, \sigma^{-(m+1)} \prod_{i=1}^{m} \left(1 + \frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)}, \quad \sigma > 0, \; \xi > -\frac{\sigma}{z_m},$$

where

$$C_m = \int_{-\infty}^{\infty} \int_{\max(0, -\xi z_m)}^{\infty} \pi(\xi)\, \sigma^{-(m+1)} \prod_{i=1}^{m} \left(1 + \frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)} \, \mathrm{d}\sigma \, \mathrm{d}\xi \quad (3.1)$$

and the inequality $\xi > -\sigma/z_m$ comes from the constraints $1 + \xi z_i/\sigma > 0, i = 1, \ldots, m$ in the likelihood.

3.1. Prior densities

Using (2.1) with $\phi = (\sigma, \xi)$ gives the Jeffreys prior

$$\pi_{J,GP}(\sigma, \xi) \propto \frac{1}{\sigma(1+\xi)(1+2\xi)^{1/2}}, \quad \sigma > 0, \; \xi > -\frac{1}{2}.$$

Eugenia Castellanos and Cabras (2007) show that a proper posterior density results for $m \geq 1$.

Using (2.3) gives the MDI prior

$$\pi_{M,GP}(\sigma, \xi) \propto \frac{1}{\sigma}\, \mathrm{e}^{-(\xi+1)} \propto \frac{1}{\sigma}\, \mathrm{e}^{-\xi} \quad \sigma > 0, \; \xi \in \mathbb{R}. \tag{3.2}$$

Beirlant et al. (2004, p.447) use this prior but they do not investigate the propriety of the posterior.

Placing independent uniform priors on $\log \sigma$ and ξ gives the prior

$$\pi_{U,GP}(\sigma, \xi) \propto \frac{1}{\sigma}, \qquad \sigma > 0, \; \xi \in \mathbb{R}, \tag{3.3}$$

proposed by Pickands (1994).

Figure 1 shows the Jeffreys and MDI priors for GP parameters as functions of ξ. The MDI prior increases without limit as $\xi \to -\infty$ and the Jeffreys prior increases without limit as $\xi \downarrow -1/2$.
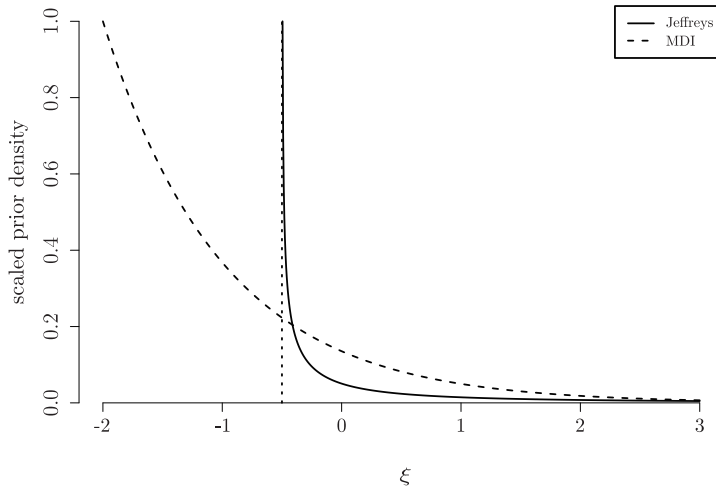
Figure 1.   Scaled Jeffreys and MDI GP prior densities against ξ.

3.2. Results

Theorem 1. *A sufficient condition for the prior $\pi(\sigma, \xi) \propto \pi(\xi)/\sigma, \sigma > 0, \xi \in \mathbb{R}$ to yield a proper posterior density function is that $\pi(\xi)$ is (proportional to) a proper density function.*

The MDI prior (3.2) does not satisfy the condition in Theorem 1 because $\exp\{-(\xi + 1)\}$ is not a proper density function on $\xi \in \mathbb{R}$.

Theorem 2. *There is no sample size for which the MDI prior (3.2) yields a proper posterior density function.*

The problem with the MDI prior is due to its behaviour for negative ξ so a simple solution is to place a lower bound on ξ *a priori.* This approach is common in extreme value analyses, for example, Martins and Stedinger (2001) constrain ξ to $(-1/2, 1/2)$ *a priori.* We suggest

$$\pi'_{M,GP}(\sigma, \xi) = \frac{1}{\sigma} e^{-(\xi+1)}, \, \xi \geq -1, \tag{3.4}$$

a (proper) unit exponential prior on $\xi + 1$. Any finite lower bound on ξ ensures propriety of the posterior but $\xi = -1$, for which the GP distribution reduces to a uniform distribution on $(0, \sigma)$, seems less arbitrary than other choices as it corresponds to a change in the behaviour of the GP density. For $\xi > -1$, the GP density $f_{GP}(z)$ decreases in z, which is what one anticipates when conducting an extreme value analysis to make inferences about future large, rare values. For $\xi < -1$, $f_{GP}(z)$ increases without limit as it approaches its mode at the upper end point $-\sigma/\xi$, behaviour that is not expected in such analyses.

Corollary to Theorem 1. *The truncated MDI prior* (3.4) *yields a proper posterior density function for $m \geq 1$.*

Theorem 3. *A sufficient condition for the uniform prior* (3.3) *to yield a proper posterior density function is that $m \geq 3$.*

4. Generalized Extreme Value (GEV) Distribution

Without loss of generality we take the n block maxima to be ordered; $y_1 < \cdots < y_n$. We consider a class of priors of the form $\pi(\mu, \sigma, \xi) \propto \pi(\xi)/\sigma$, $\sigma > 0, \mu, \xi \in \mathbb{R}$; thus, *a priori* μ, σ, and ξ are independent and μ and $\log \sigma$ have improper uniform priors over the real line.

Based on a random sample $y_1, \ldots, y_n$ the posterior density for (μ, σ, ξ) is proportional to

$$\sigma^{-(n+1)} \pi(\xi) \exp \left\{ -\sum_{i=1}^{n} w_i^{-1/\xi} \right\} \prod_{i=1}^{n} w_i^{-(1+1/\xi)}, \tag{4.1}$$

where $w_i = 1 + \xi(y_i - \mu)/\sigma$ and $\sigma > 0$. If $\xi > 0$ then $\mu - \sigma/\xi < y_1$, and if $\xi < 0$ then $\mu - \sigma/\xi > y_n$.

4.1. Prior densities

Kotz and Nadarajah (2000, p.63) give the Fisher information matrix for the GEV distribution (1.2). Using (2.2) with $\phi = (\mu, \sigma, \xi)$ gives the Jeffreys prior

$$\pi_{J,GEV}(\mu, \sigma, \xi) = \frac{1}{\sigma \xi^2} \left\{ [1 - 2\Gamma(2 + \xi) + p] \left[\frac{\pi^2}{6} + \left(1 - \gamma + \frac{1}{\xi} \right)^2 - \frac{2q}{\xi} + \frac{p}{\xi^2} \right] \right.$$
$$\left. - \left[1 - \gamma + \frac{1}{\xi} - \frac{1}{\xi} \Gamma(2 + \xi) - q + \frac{p}{\xi} \right]^2 \right\}^{1/2}, \ \mu \in \mathbb{R}, \sigma > 0, \ \xi > -\frac{1}{2}, \tag{4.2}$$

where $p = (1 + \xi)^2 \Gamma(1 + 2\xi)$, $q = \Gamma(2 + \xi)\{\psi(1 + \xi) + (1 + \xi)/\xi\}$, $\psi(r) = \partial \log \Gamma(r)/\partial r$, and $\gamma \approx 0.57722$ is Euler's constant. van Noortwijk, Kalk, and Chbab (2004) give an alternative form for the Jeffreys prior based on (2.1).

Beirlant et al. (2004, p.435) give the form of the MDI prior as

$$\pi_{M,GEV}(\mu, \sigma, \xi) = \frac{1}{\sigma} e^{-\gamma(\xi + 1 + 1/\gamma)} \propto \frac{1}{\sigma} e^{-\gamma(1+\xi)}, \quad \sigma > 0, \mu, \xi \in \mathbb{R}. \tag{4.3}$$

Placing independent uniform priors on μ, $\log \sigma$, and ξ gives the prior

$$\pi_{U,GEV}(\mu, \sigma, \xi) \propto \frac{1}{\sigma}, \quad \sigma > 0, \mu, \xi \in \mathbb{R}. \tag{4.4}$$

Figure 2 shows the Jeffreys and MDI priors for GEV parameters as a functions of ξ. The MDI prior increases without limit as $\xi \to -\infty$ and the Jeffreys prior increases without limit as $\xi \to \infty$ and as $\xi \downarrow -1/2$.
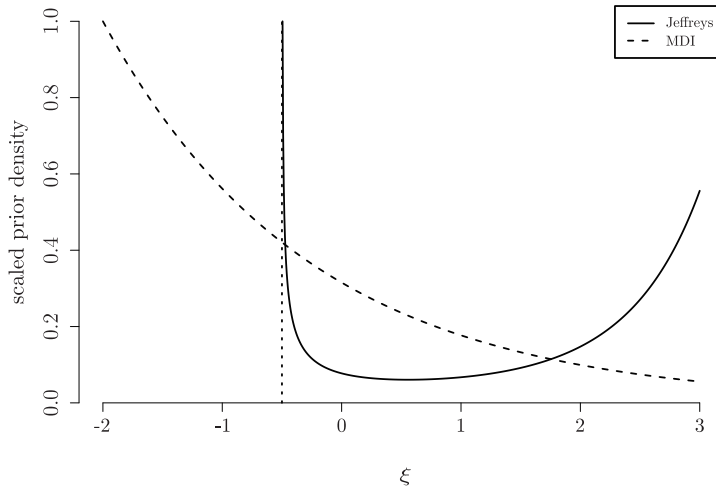
Figure 2.   Scaled Jeffreys and MDI GEV prior densities against ξ.

4.2. Results

Theorem 4. *For the prior $\pi(\mu, \sigma, \xi) \propto \pi(\xi)/\sigma, \sigma > 0, \mu, \xi \in \mathbb{R}$ to yield a proper posterior density function it is necessary that $n \geq 2$ and, in that event, it is sufficient that $\pi(\xi)$ is (proportional to) a proper density function.*

Theorem 5. *There is no sample size for which the Jeffreys prior (4.2) yields a proper posterior density function.*

Truncation of the independence Jeffreys prior to $\xi \leq \xi_+$ would yield a proper posterior density function if $n \geq 2$. In this event Theorem 4 requires only that $\int_{-1/2}^{\xi_+} \pi(\xi) \, \mathrm{d}\xi$ is finite, where $\pi(\xi) = \sigma\pi_{J,GEV}(\mu, \sigma, \xi)$ (see (4.2)). From the proof of Theorem 5 we have $\pi(\xi) < 2 \left[\pi^2/6 + (1 - \gamma)^2\right]^{1/2} (1 + 2\xi)^{-1/2}$ for $\xi \in (-1/2, -1/2 + \epsilon)$, where $\epsilon > 0$. Therefore,

$$\int_{-1/2}^{-1/2+\epsilon} \pi(\xi) \, \mathrm{d}\xi < 2 \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right]^{1/2} \int_{-1/2}^{-1/2+\epsilon} (1 + 2\xi)^{-1/2} \, \mathrm{d}\xi$$

$$= 2^{3/2} \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right]^{1/2} \epsilon^{1/2}.$$

The integral over $(-1/2 + \epsilon, \xi_+)$ is also finite. However, the choice of an *a priori* upper limit for ξ may be less obvious than the choice of a lower limit.

Theorem 6. *There is no sample size for which the MDI prior (4.3) yields a proper posterior density function.*

As in the GP case, truncating the MDI prior to $\xi \geq -1$,

$$\pi'_{M,GEV}(\mu, \sigma, \xi) \propto \frac{1}{\sigma}\, e^{-\gamma(1+\xi)} \quad \mu \in \mathbb{R}, \sigma > 0, \xi \geq -1, \tag{4.5}$$

is one way to yield a proper posterior distribution.

Corollary to Theorem 4. *The truncated MDI prior* (4.5) *yields a proper posterior density function for $n \geq 2$.*

Theorem 7. *A sufficient condition for the uniform prior* (4.4) *to yield a proper posterior density function is that $n \geq 4$.*

5. Discussion

We have shown that some of the reference priors used, or proposed for use, in extreme value modelling do not yield a proper posterior distribution unless we are willing to truncate the possible values of ξ *priori*. An interesting aspect of our findings is that the Jeffreys prior (4.2) for GEV parameters fails to yield a proper posterior, whereas the uniform prior (4.4) requires only weak conditions to ensure posterior propriety. This is the opposite of more general experience, summarised by Berger (2006, p.393) and Yang and Berger (1998, p.5), that the Jeffreys prior almost always yields a proper posterior whereas a uniform prior often fails to do so. The impropriety of the posterior under the Jeffreys prior is due to the high rate at which the component $\pi(\xi)$ of this prior increases for large ξ. An alternative prior based on Jeffreys' general rule (2.1) van Noortwijk, Kalk, and Chbab (2004) also has this property.

The conditions sufficient for posterior propriety under the uniform priors (3.3) and (4.4) are weak. Therefore, a posterior yielded by a diffuse normal priors is meaningful but such a prior could be replaced by an improper uniform prior. Although it is reassuring to know that a posterior is proper, with a sufficiently informative sample posterior impropriety might not present a practical problem (Kass and Wasserman (1996, section 5.2)). This may explain why Beirlant et al. (2004, p.435 and 447) obtain sensible results using (untruncated) MDI priors. However, the posterior impropriety may be evident for smaller sample sizes.

In making inferences about high quantiles of the marginal distribution of X, the GP model for threshold excesses is combined with a binomial(N, p_u) model for the number of excesses, where $p_u = P(X > u)$. Reference priors for a binomial probability have been studied extensively, see, for example, Tuyl, Gerlachy, and Mengersenz (2009). An approximately equivalent approach is the non-homogeneous Poisson process (NHPP) model Smith (1989), which is parameterized in terms of GEV parameters μ, σ, and ξ relating to the distribution of

$\max\{X,\ldots,X_b\}$. Suppose that m observations $x_1,\ldots,x_m$ exceed u. Under the NHPP the posterior density for (μ,σ,ξ) is proportional to

$$\sigma^{-(m+1)}\pi(\xi)\exp\left\{-n\left[1+\xi\left(\frac{u-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\}\prod_{i=1}^{m}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]_+^{-(1+1/\xi)}, \quad (5.1)$$

where n is the (notional) number of blocks into which the data are divided in defining (μ,σ,ξ). Without loss of generality, we take $n=m$. The exponential term in (5.1) is an increasing function of u, and $x_i > u, i = 1,\ldots,m$. Therefore,

$$\exp\left\{-n\left[1+\xi\left(\frac{u-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\} < \exp\left\{-\sum_{i=1}^{m}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\}$$

and (5.1) is less than

$$\sigma^{-(m+1)}\pi(\xi)\exp\left\{-\sum_{i=1}^{m}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]_+^{-1/\xi}\right\}\prod_{i=1}^{m}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]_+^{-(1+1/\xi)}. \quad (5.2)$$

Equation (5.2) is of the same form as (4.1), with $n=m$ and $y_i = x_i, i = 1,\ldots,n$. Therefore, theorems 4 and 7 apply to the NHPP model: for posterior propriety it is sufficient that either (a) $n \geq 2$ and $\pi(\mu,\sigma,\xi) \propto \pi(\xi)/\sigma$, for $\sigma > 0, \mu,\xi \in \mathbb{R}$, where $\int_\xi \pi(\xi)\,\mathrm{d}\xi$ is finite, or (b) $n \geq 4$ and $\pi(\mu,\sigma,\xi) \propto 1/\sigma$, for $\sigma > 0, \mu,\xi \in \mathbb{R}$.

One possible extension of our work is to regression modelling using extreme value response distributions. For example, Roy and Dey (2014) use GEV regression modelling to analyze reliability data. They prove posterior propriety under conditions on the prior for (σ,ξ) that are stronger than those in our theorems 4 and 7. Future work will investigate our conjecture that the conditions in Roy and Dey (2014) can be weakened. Another extension is to explore other formal rules for constructing priors, such as reference priors Berger, Bernardo, and Sun (2009) and probability matching priors Datta et al. (2009). Ho (2010) considers the latter for the GP distribution.

Acknowledgements

Appendix

A.1. Moments of a GP distribution

We give some moments of the GP distribution for later use. Suppose that $Z \sim GP(\sigma,\xi)$, where $\xi < 1/r$. Then Giles and Feng (2009)

$$\mathrm{E}(Z^r) = \frac{r!\sigma^r}{\prod_{i=1}^{r}(1 - i\xi)}, \qquad r = 1, 2, \dots. \tag{A.1}$$

Now suppose that $\xi < 0$. Then, for a constant $a > \xi$, and using the substitution $x = -\xi v/\sigma$, we have

$$
\begin{aligned}
\mathrm{E}(Z^{-a/\xi}) &= \int_0^{-\sigma/\xi} v^{-a/\xi} \frac{1}{\sigma}\left(1 + \frac{\xi v}{\sigma}\right)^{-(1+1/\xi)} \mathrm{d}v \\
&= (-\xi)^{a/\xi - 1}\sigma^{-a/\xi} \int_0^1 x^{-a/\xi}(1 - x)^{-(1+1/\xi)} \, \mathrm{d}x \\
&= (-\xi)^{a/\xi - 1}\sigma^{-a/\xi} \frac{\Gamma(1 - a/\xi)\Gamma(-1/\xi)}{\Gamma(1 - (a+1)/\xi)},
\end{aligned}
\tag{A.2}
$$

where we have used integral number 1 in section 3.251 on page 324 of Gradshteyn and Ryzhik (2007), namely

$$\int_0^1 x^{\mu-1}(1 - x^\lambda)^{\nu-1} \, \mathrm{d}x = \frac{1}{\lambda}\mathrm{Beta}\left(\frac{\mu}{\lambda}, \nu\right) = \frac{\Gamma(\mu/\lambda)\Gamma(\nu)}{\Gamma(\mu/\lambda + \nu)} \qquad \lambda > 0, \nu > 0, \mu > 0,$$

with $\lambda = 1, \mu = 1 - a/\xi$ and $v = -1/\xi$.

In the following proofs we use the generic notation $\pi(\xi)$ for the component of the prior relating to ξ; the form of $\pi(\xi)$ varies depending on the prior being considered.

A.2. Proof of Theorem 1

This is a trivial extension of the proof of Theorem 1 in Eugenia Castellanos and Cabras (2007). Suppose $m = 1$, with an observation z. The normalizing constant C of the posterior distribution is given by

$$
\begin{aligned}
C_1 &= \int_{-\infty}^0 \pi(\xi) \int_{-\xi z}^\infty \sigma^{-2}\left(1 + \frac{\xi z}{\sigma}\right)^{-(1+1/\xi)} \mathrm{d}\sigma \, \mathrm{d}\xi \\
&\quad + \int_0^\infty \pi(\xi) \int_0^\infty \sigma^{-2}\left(1 + \frac{\xi z}{\sigma}\right)^{-(1+1/\xi)} \mathrm{d}\sigma \, \mathrm{d}\xi \\
&= \frac{1}{z} \int_{-\infty}^\infty \pi(\xi) \, \mathrm{d}\xi.
\end{aligned}
$$

If the latter integral is finite, $\pi(\xi)$ is proportional to a proper density function, then the posterior distribution is proper for $m = 1$ and therefore, by successive iterations of Bayes' theorem, it is proper for $m \geq 1$. The corollary follows directly.

A.3. Proof of Theorem 2

Let $A(\xi) = \mathrm{e}^{-\xi}$ and $B_m(\sigma, \xi) = \sigma^{-(m+1)} \prod_{i=1}^m \left(1 + \xi z_i/\sigma\right)^{-(1+1/\xi)}$. Then, from (3.1) we have

$$C_m = \int_{-\infty}^{\infty} A(\xi) \int_{\max(0,-\xi z_m)}^{\infty} B_m(\sigma,\xi) \, \mathrm{d}\sigma \, \mathrm{d}\xi$$

$$= \int_{-\infty}^{-1} A(\xi) \int_{-\xi z_m}^{\infty} B_m(\sigma,\xi) \, \mathrm{d}\sigma \, \mathrm{d}\xi + \int_{-1}^{0} A(\xi) \int_{-\xi z_m}^{\infty} B_m(\sigma,\xi) \, \mathrm{d}\sigma \, \mathrm{d}\xi$$

$$+ \int_{0}^{\infty} A(\xi) \int_{0}^{\infty} B_m(\sigma,\xi) \, \mathrm{d}\sigma \, \mathrm{d}\xi.$$

The latter two integrals converge for $m \geq 1$. However, the first integral diverges for all samples sizes. For $\xi < -1$, $(1+\xi z/\sigma)^{-(1+1/\xi)} > 1$ when z is in the support $(0, -\sigma/\xi)$ of the GP(σ,ξ) density. Therefore $B_m(\sigma,\xi) > \sigma^{-(m+1)}$. Thus, the first integral above satisfies

$$\int_{-\infty}^{-1} A(\xi) \int_{-\xi z_m}^{\infty} B_m(\sigma,\xi) \, \mathrm{d}\sigma \, \mathrm{d}\xi > \int_{-\infty}^{-1} A(\xi) \int_{-\xi z_m}^{\infty} \sigma^{-(m+1)} \, \mathrm{d}\sigma \, \mathrm{d}\xi$$

$$= \int_{-\infty}^{-1} A(\xi) \left[-\frac{1}{m}\sigma^{-m} \right]_{-\xi z_m}^{\infty} \mathrm{d}\xi$$

$$= \int_{-\infty}^{-1} A(\xi)\frac{1}{m} \left[-\xi z_m\right]^{-m} \, \mathrm{d}\xi$$

$$= \frac{1}{m z_m^m} \int_{1}^{\infty} v^{-m} \mathrm{e}^v \, \mathrm{d}v,$$

where $v = -\xi$. This integral is divergent for all $m \geq 1$, so there is no sample size for which the posterior is proper.

A.4. Proof of Theorem 3

We need to show that C_3 is finite. We split the range of integration over ξ so that $C_3 = I_1 + I_2 + I_3$, where

$$I_1 = \int_{-\infty}^{-1} \int_{-\xi z_3}^{\infty} B_3 \, \mathrm{d}\sigma \, \mathrm{d}\xi, \quad I_2 = \int_{-1}^{0} \int_{-\xi z_3}^{\infty} B_3 \, \mathrm{d}\sigma \, \mathrm{d}\xi, \quad I_3 = \int_{0}^{\infty} \int_{0}^{\infty} B_3 \, \mathrm{d}\sigma \, \mathrm{d}\xi,$$

and $B_3 = B_3(\sigma,\xi) = \sigma^{-4} \prod_{i=1}^{3} (1 + \xi z_i/\sigma)^{-(1+1/\xi)}$. For convenience we let $\rho = \xi/\sigma$ and $\zeta_i = z_3 - z_i, i = 1, 2$.

Proof that I_1 is finite. We have $\xi < -1$ and so $-(1 + 1/\xi) < 0$, $\rho < 0$ and $0 < 1 + \rho z_i < 1$ for $i = 1, 2, 3$. Noting that $-\rho z_3 < 1$ gives

$$(1 + \rho z_1)(1 + \rho z_2)(1 + \rho z_3) > (-\rho z_3 + \rho z_1)(-\rho z_3 + \rho z_2)(1 + \rho z_3)$$

$$= (-\rho)^2 \zeta_1 \zeta_2 (1 + \rho z_3)$$

$$= (-\xi)^2 \sigma^{-2} \zeta_1 \zeta_2 (1 + \rho z_3). \tag{A.3}$$

Therefore,

$$\prod_{i=1}^{3}\left(1+\frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)} < (-\xi)^{-2(1+1/\xi)}\sigma^{2(1+1/\xi)}\left[\zeta_2\,\zeta_1\left(1+\frac{\xi z_3}{\sigma}\right)\right]^{-(1+1/\xi)}.$$

Thus, $I_1 \le \int_{-\infty}^{-1}(-\xi)^{-2(1+1/\xi)}\left[\zeta_2\,\zeta_1\right]^{-(1+1/\xi)} I_{1\sigma}\,\mathrm{d}\xi$, where

$$\begin{aligned}
I_{1\sigma} &= \int_{-\xi z_3}^{\infty}\sigma^{-4}\sigma^{2(1+1/\xi)}\left(1+\frac{\xi z_3}{\sigma}\right)^{-(1+1/\xi)}\,\mathrm{d}\sigma\\
&= z_3^{-1}\int_0^{-1/\xi z_3}v^{-2/\xi}\frac{1}{z_3^{-1}}\left(1+\frac{\xi v}{z_3^{-1}}\right)^{-(1+1/\xi)}\,\mathrm{d}v\\
&= (-\xi)^{2/\xi-1}z_3^{-(1-2/\xi)}\frac{\Gamma(1-2/\xi)\Gamma(-1/\xi)}{\Gamma(1-3/\xi)},
\end{aligned}$$

where $v = 1/\sigma$ and the last line follows from (A.2) with $a = 2$ and $\sigma = z_3^{-1}$. Therefore,

$$\begin{aligned}
I_1 &\le \int_{-\infty}^{-1}(-\xi)^{-3}\left[\zeta_2\,\zeta_1\right]^{-(1+1/\xi)}z_3^{-(1-2/\xi)}\frac{\Gamma(1-2/\xi)\Gamma(-1/\xi)}{\Gamma(1-3/\xi)}\,\mathrm{d}\xi\\
&= [z_3\zeta_2\,\zeta_1]^{-1}\int_{-\infty}^{-1}(-\xi)^{-3}\left(\frac{\zeta_2}{z_3}\right)^{-1/\xi}\left(\frac{\zeta_1}{z_3}\right)^{-1/\xi}\frac{\Gamma(1-2/\xi)\Gamma(-1/\xi)}{\Gamma(1-3/\xi)}\,\mathrm{d}\xi\\
&= [z_3\zeta_2\,\zeta_1]^{-1}\int_0^1 x\left(\frac{\zeta_2}{z_3}\right)^x\left(\frac{\zeta_1}{z_3}\right)^x\frac{\Gamma(1+2x)\Gamma(x)}{\Gamma(1+3x)}\,\mathrm{d}x\\
&= [z_3\zeta_2\,\zeta_1]^{-1}\int_0^1\left(\frac{\zeta_2}{z_3}\right)^x\left(\frac{\zeta_1}{z_3}\right)^x\frac{\Gamma(1+2x)\Gamma(1+x)}{\Gamma(1+3x)}\,\mathrm{d}x, \qquad (A.4)
\end{aligned}$$

where $x = -1/\xi$ and we have used the relation $\Gamma(1+x) = x\,\Gamma(x)$. The integrand in (A.4) is finite over the range of integration so this integral is finite and therefore I_1 is finite.

Proof that I_2 is finite. We have $-1 < \xi < 0$, so $-(1+1/\xi) > 0$ and $(1+\xi z/\sigma)^{-(1+1/\xi)} < 1$ and decreases in z over $(0, -\sigma/\xi)$. Therefore,

$$\begin{aligned}
I_2 &= \int_{-1}^{0}\int_{-\xi z_3}^{\infty}\sigma^{-4}\prod_{i=1}^{3}\left(1+\frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)}\,\mathrm{d}\sigma\,\mathrm{d}\xi\\
&\le \int_{-1}^{0}\int_{-\xi z_3}^{\infty}\sigma^{-4}\left(1+\frac{\xi z_3}{\sigma}\right)^{-(1+1/\xi)}\,\mathrm{d}\sigma\,\mathrm{d}\xi\\
&= \int_{-1}^{0}z_3^{-1}\int_0^{-1/\xi z_3}v^2\frac{1}{z_3^{-1}}\left(1+\frac{\xi v}{z_3^{-1}}\right)^{-(1+1/\xi)}\,\mathrm{d}v\,\mathrm{d}\xi
\end{aligned}$$

$$= z_3^{-1} \int_{-1}^0 \frac{2z_3^{-2}}{(1-\xi)(1-2\xi)} \, d\xi$$

$$= 2z_3^{-3} \int_{-1}^0 \left\{ \left(\frac{1}{2} - \xi\right)^{-1} - (1-\xi)^{-1} \right\} \, d\xi$$

$$= 2z_3^{-3} \ln\left(\frac{3}{2}\right),$$

where the integral over v follows from (A.1) with $r = 2$ and $\sigma = z_3^{-1}$.

Proof that I_3 is finite. We have $\xi > 0$ so $-(1+1/\xi) < 0$. Let $g_n = (\prod_{i=1}^n z_i)^{1/n}$. From Mitrinović (1964, p.130):

$$\prod_{k=1}^n (1 + a_k) \geq (1+b)^n, \qquad a_k > 0; \quad \prod_{k=1}^n a_k = b^n, \qquad (A.5)$$

with $a_k = \xi z_k/\sigma$ and $b = \xi g_3/\sigma$ gives

$$\prod_{i=1}^3 \left(1 + \frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)} \leq \left(1 + \frac{\xi g_3}{\sigma}\right)^{-3(1+1/\xi)},$$

and therefore

$$I_3 = \int_0^\infty \int_0^\infty \sigma^{-4} \prod_{i=1}^3 \left(1 + \frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)} \, d\sigma \, d\xi$$

$$\leq \int_0^\infty \int_0^\infty \sigma^{-4} \left(1 + \frac{\xi g_3}{\sigma}\right)^{-3(1+1/\xi)} \, d\sigma \, d\xi$$

$$= \int_0^\infty \beta \int_0^\infty v^2 \frac{1}{\beta} \left(1 + \frac{\alpha v}{\beta}\right)^{-(1+1/\alpha)} \, dv \, d\xi,$$

where $v = 1/\sigma$, $\alpha = 1/(2 + 3/\xi)$ and $\beta = \alpha/\xi g_3 = 1/(3 + 2\xi)g_3$. For $\xi > 0$, $\alpha < 1/2$ so using (A.1) with $r = 2, \sigma = \beta$ and $\xi = \alpha$ gives

$$I_3 \leq \int_0^\infty \beta \frac{2\beta^2}{(1-\alpha)(1-2\alpha)} \, d\xi$$

$$= \frac{2}{3} g_3^{-3} \int_0^\infty \frac{1}{(\xi + 3)(2\xi + 3)} \, d\xi$$

$$= \frac{2}{9} g_3^{-3} \int_0^\infty \left(\frac{1}{\xi + 3/2} - \frac{1}{\xi + 3}\right) \, d\xi$$

$$= \frac{2}{9} g_3^{-3} \ln 2.$$

The normalizing constant C_3 is finite, so $\pi_{U,GP}(\sigma, \xi)$ yields a proper posterior density for $m = 3$ and therefore does so for $m \geq 3$.

A.5. Proof of Theorem 4

Throughout the following proofs we take $\delta_i = y_i - y_1, i = 2, \ldots, n$. We make the parameter transformation $\phi = \mu - \sigma/\xi$. Then the posterior density for (ϕ, σ, ξ) is given by

$$\pi(\phi, \sigma, \xi) = K_n^{-1} \pi(\xi) |\xi|^{-n(1+1/\xi)} G_n(\phi, \sigma),$$

where

$$G_n(\phi, \sigma) = \sigma^{n/\xi-1} \left\{ \prod_{i=1}^{n} |y_i - \phi|^{-(1+1/\xi)} \right\} \exp \left\{ -|\xi|^{-1/\xi} \sigma^{1/\xi} \sum_{i=1}^{n} |y_i - \phi|^{-1/\xi} \right\}$$

and if $\xi > 0$ then $\phi < y_1$, and if $\xi < 0$ then $\phi > y_n$.

We let $v = \sigma^{1/\xi}$, $H = H(\phi, \xi) = |\xi|^{-1/\xi} \sum_{i=1}^{n} |y_i - \phi|^{-1/\xi}$ and $J = J(\phi, \xi) = \prod_{i=1}^{n} |y_i - \phi|^{-(1+1/\xi)}$. The normalizing constant K_n is given by

$$\begin{aligned}
K_n &= \int_{-\infty}^{\infty} \int \int_{0}^{\infty} \pi(\xi) |\xi|^{-n(1+1/\xi)} G_n(\phi, \sigma) \, d\sigma \, d\phi \, d\xi \\
&= \int_{-\infty}^{\infty} \pi(\xi) |\xi|^{-n(1+1/\xi)} \int J \int_{0}^{\infty} \sigma^{n/\xi-1} \exp\left\{ -H\sigma^{1/\xi} \right\} \, d\sigma \, d\phi \, d\xi \\
&= \int_{-\infty}^{\infty} \pi(\xi) |\xi|^{-n(1+1/\xi)} \int J \int_{0}^{\infty} v^{n-1} \exp\{-Hv\} \, |\xi| \, dv \, d\phi \, d\xi \\
&= \int_{-\infty}^{\infty} \pi(\xi) |\xi|^{-n(1+1/\xi)} \int J \Gamma(n) H^{-n} \, |\xi| \, d\phi \, d\xi \\
&= \int_{-\infty}^{\infty} \pi(\xi) |\xi|^{-n(1+1/\xi)} \int J (n-1)! |\xi|^{n/\xi+1} \left\{ \sum_{i=1}^{n} |y_i - \phi|^{-1/\xi} \right\}^{-n} d\phi \, d\xi \\
&= (n-1)! \int_{-\infty}^{\infty} \pi(\xi) |\xi|^{1-n} \int J(\phi, \xi) \left\{ \sum_{i=1}^{n} |y_i - \phi|^{-1/\xi} \right\}^{-n} d\phi \, d\xi. \qquad \text{(A.6)}
\end{aligned}$$

For $n = 1$ the integral $\int_{\phi:\xi(y_1-\phi)>0} |y_1 - \phi|^{-1} \, d\phi$ is divergent so if $n = 1$ the posterior is not proper for any prior in this class.

Now we take $n = 2$ and for clarity consider the cases $\xi > 0$ and $\xi < 0$ separately, with respective contributions K_2^+ and K_2^- to K_2. For $\xi > 0$, using the substitution $u = (y_1 - \phi)^{-1}$ in (A.6) gives

$$\begin{aligned}
K_2^+ &= \int_{0}^{\infty} \pi(\xi) \, \xi^{-1} \int_{-\infty}^{y_1} \frac{(y_1 - \phi)^{-(1+1/\xi)}(y_2 - \phi)^{-(1+1/\xi)}}{\left\{ (y_1 - \phi)^{-1/\xi} + (y_2 - \phi)^{-1/\xi} \right\}^2} \, d\phi \, d\xi \\
&= \int_{0}^{\infty} \pi(\xi) \, \xi^{-1} \int_{0}^{\infty} \frac{(1 + \delta_2 u)^{-(1+1/\xi)}}{\left\{ 1 + (1 + \delta_2 u)^{-1/\xi} \right\}^2} \, du \, d\xi \\
&= \frac{1}{2} \delta_2^{-1} \int_{0}^{\infty} \pi(\xi) \, d\xi,
\end{aligned}$$

the final step following because the u-integrand is a multiple $(\xi \delta_2^{-1})$ of a shifted log-logistic density function with location, scale and shape parameters of $0, \xi \delta_2^{-1}$ and ξ respectively, and the location of this distribution equals the median. For $\xi < 0$ an analogous calculation using the substitution $v = (y_n - \phi)^{-1}$ in (A.6) gives

$$K_2^- = \frac{1}{2} \delta_2^{-1} \int_{-\infty}^0 \pi(\xi) \, d\xi.$$

Therefore,

$$K_2 = K_2^+ + K_2^- = \frac{1}{2} \delta_2^{-1} \int_{-\infty}^\infty \pi(\xi) \, d\xi.$$

Thus, K_2 is finite if $\int_{-\infty}^\infty \pi(\xi) \, d\xi$ is finite, and the result follows. The corollary follows directly.

A.6. Proof of Theorem 5

The crucial aspects are the rates at which $\pi(\xi) \to \infty$ as $\xi \downarrow -1/2$ and as $\xi \to \infty$. The component $\pi(\xi)$ of (4.2) involving ξ can be expressed as

$$\pi_\xi^2(\xi) = \frac{1}{\xi^4}(T_1 + T_2), \tag{A.7}$$

where

$$T_1 = \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right](1 + \xi)^2 \Gamma(1 + 2\xi), \tag{A.8}$$

$$T_2 = \frac{\pi^2}{6} + \left[2(1-\gamma)(\gamma+\psi(1+\xi)) - \frac{\pi^2}{3}\right]\Gamma(2+\xi)$$
$$- [1 + \psi(1+\xi)]^2 [\Gamma(2+\xi)]^2. \tag{A.9}$$

First, we derive a lower bound for $\pi(\xi)$ that holds for $\xi > 3$. Using the duplication formula (Abramowitz and Stegun (1972, p.256; 6.1.18))

$$\Gamma(2z) = (2\pi)^{-1/2} \, 2^{2z-1/2} \, \Gamma(z) \, \Gamma(z + \frac{1}{2}),$$

with $z = 1/2 + \xi$ in (A.8) we have

$$T_1 = \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right](1 + \xi)^2 \, \pi^{-1/2} 2^{2\xi} \, \Gamma(1/2 + \xi) \, \Gamma(1 + \xi).$$

Then

$$\Gamma(1/2 + \xi) = \frac{\Gamma(3/2 + \xi)}{1/2 + \xi} > \frac{\Gamma(1 + \xi)}{1/2 + \xi} = \frac{2\Gamma(1 + \xi)}{1 + 2\xi} > \frac{\Gamma(1 + \xi)}{1 + \xi},$$

where for the first inequality to hold it is sufficient that $\xi > 1/2$; and that, for $\xi > 3$, $2^{2\xi} > (1 + \xi)^3$. Therefore,

$$T_1 > \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right] \pi^{-1/2} (1 + \xi)^4 [\Gamma(1 + \xi)]^2. \qquad \text{(A.10)}$$

Completing the square in (A.9) gives

$$T_2 = - \left\{[1 + \psi(1 + \xi)] \Gamma(2 + \xi) + f(\xi)\right\}^2 + [f(\xi)]^2 + \frac{\pi^2}{6},$$

where

$$f(\xi) = \frac{\pi^2/6 - (1 - \gamma)(\gamma + \psi(1 + \xi))}{1 + \psi(1 + \xi)} = \frac{\pi^2/6 + (1 - \gamma)^2}{1 + \psi(1 + \xi)} - (1 - \gamma)$$

and $[f(\xi)]^2 + \pi^2/6 > 0$.

For $\xi > 0$, $\psi(1 + \xi)$ increases with ξ and so $f(\xi)$ decreases with ξ. Therefore, for $\xi > 3$, $f(\xi) < f(3) \approx 0.39$ and

$$T_2 > - \left\{[1 + \psi(1 + \xi)] \Gamma(2 + \xi) + f(3)\right\}^2.$$

For $\xi > 0$, we have $\psi(1 + \xi) < \ln(1 + \xi) - (1 + \xi)^{-1}/2$ (Qiu and Vuorinen (2004, Theorem C)) and $\ln(1 + \xi) \leq \xi$ (Abramowitz and Stegun (1972, p.68; 4.1.33)). Therefore, noting that $\Gamma(2 + \xi) = (1 + \xi)\Gamma(1 + \xi)$, we have

$$T_2 > - \left\{(1 + \xi)^2 \Gamma(1 + \xi) - \frac{1}{2}\Gamma(1 + \xi) + f(3)\right\}^2.$$

For $\xi > 3$, $f(3) - \Gamma(1 + \xi)/2 < 0$ so

$$T_2 > -(1 + \xi)^4 [\Gamma(1 + \xi)]^2. \qquad \text{(A.11)}$$

Substituting (A.10) and (A.11) in (A.7) gives, for $\xi > 3$,

$$\begin{aligned}
\pi_\xi^2(\xi) &> \frac{(1 + \xi)^4}{\xi^4} \left\{\left[\frac{\pi^2}{6} + (1 - \gamma)^2\right] \pi^{-1/2} - 1\right\} [\Gamma(1 + \xi)]^2 \\
&> c[\Gamma(1 + \xi)]^2 \\
&> c(1 + \xi)^{2(\lambda\xi - \gamma)},
\end{aligned}$$

where $c = (4/3)^4\{[\pi^2/6 + (1 - \gamma)^2]\pi^{-1/2} - 1\} \approx 0.0913$ and the final step uses the inequality $\Gamma(x) > x^{\lambda(x-1)-\gamma}$, for $x > 0$ Alzer (1999), where $\lambda = (\pi^2/6 - \gamma)/2 \approx 0.534$. Thus, a lower bound for the ξ component of the Jeffreys prior (4.2) is given by

$$\pi(\xi) > c^{1/2}(1 + \xi)^{\lambda\xi - \gamma}, \qquad \text{for } \xi > 3. \qquad \text{(A.12)}$$

(In fact, numerical work shows that this lower bound holds for $\xi > -1/2$.)

Let K_n^+ denote the contribution to K_n for $\xi > 3$. Using the substitution $u = (y_1 - \phi)^{-1}$ in (A.6) gives

$$K_n^+ = (n-1)! \int_3^\infty \pi(\xi)\,\xi^{1-n} \int_0^\infty u^{n-2} \frac{\prod_{i=1}^n (1+\delta_i u)^{-(1+1/\xi)}}{\left\{1 + \sum_{i=2}^n (1+\delta_i u)^{-1/\xi}\right\}^n}\, \mathrm{d}u\,\mathrm{d}\xi. \quad \text{(A.13)}$$

For $\xi > 0$ we have $1 + \sum_{i=2}^n (1+\delta_i u)^{-1/\xi} \le n$ and $\prod_{i=1}^n (1+\delta_i u)^{-(1+1/\xi)} \ge (1+\delta_n u)^{-(n-1)(1+1/\xi)}$. Applying these inequalities to (A.13) gives

$$K_n^+ \ge n^{-n}(n-1)! \int_3^\infty \pi(\xi)\,\xi^{1-n} \int_0^\infty u^{n-2}(1+\delta_n u)^{-(n-1)(1+1/\xi)}\,\mathrm{d}u\,\mathrm{d}\xi$$

$$= n^{-n}(n-1)! \int_3^\infty \pi(\xi)\,\xi^{1-n}\beta \int_0^\infty u^{n-2}\frac{1}{\beta}\left(1+\frac{\alpha u}{\beta}\right)^{-(1+1/\alpha)}\,\mathrm{d}u\,\mathrm{d}\xi, \quad \text{(A.14)}$$

where $\beta = \alpha/\delta_n$ and $\alpha = [n-2+(n-1)/\xi]^{-1}$ and $0 < \alpha < (n-2)^{-1}$. The u-integrand is the density function of a $\mathrm{GP}(\beta, \alpha)$ distribution and so, using (A.1) with $r = n-2$, the integral over u is given by

$$(n-2)!\,\beta^{n-2}\prod_{i=1}^{n-2}\frac{1}{1-i\alpha} = (n-2)!\,\xi^{n-2}\delta_n^{2-n}\prod_{i=1}^{n-2}\frac{1}{(n-2-i)\xi+n-1}. \quad \text{(A.15)}$$

Substituting (A.15) into (A.14) gives

$$K_n^+ \ge n^{-n}(n-1)!(n-2)!\,\delta_n^{1-n}\int_3^\infty \frac{1}{(n-2)\xi+n-1}\prod_{i=1}^{n-2}\frac{1}{(n-2-i)\xi+n-1}\,\pi(\xi)\,\mathrm{d}\xi$$

$$= n^{-n}(n-1)!(n-2)!\,\delta_n^{1-n}\int_3^\infty \prod_{i=0}^{n-2}\frac{1}{(n-2-i)\xi+n-1}\,\pi(\xi)\,\mathrm{d}\xi$$

$$= n^{-n}(n-1)!(n-2)!\,\delta_n^{1-n}(n-1)^{1-n}\int_3^\infty \prod_{i=0}^{n-2}\frac{1}{1+\frac{i}{n-1}\xi}\,\pi(\xi)\,\mathrm{d}\xi$$

$$> C(n)\int_3^\infty \frac{1}{(1+\xi)^{n-2}}\,\pi(\xi)\,\mathrm{d}\xi,$$

where $C(n) = n^{-n}(n-1)!(n-2)!\,\delta_n^{1-n}(n-1)^{1-n}$. Applying (A.12) gives

$$K_n^+ > C(n)\,c^{1/2}\int_3^\infty (1+\xi)^{2-n+\lambda\xi-\gamma}\,\mathrm{d}\xi.$$

For any sample size n the integrand $\to \infty$ as $\xi \to \infty$. Therefore, the integral diverges and the result follows.

Now we derive an upper bound for $\pi_\xi(\xi)$ that applies for ξ close to $-1/2$. We note that for $-1/2 < \xi < 0$ we have $\Gamma(1+2\xi) = \Gamma(2+2\xi)/(1+2\xi) < (1+2\xi)^{-1}$. From (A.7) we have

$$\pi_\xi^2(\xi) = \left[\frac{\pi^2}{6} + (1-\gamma)^2\right]\left(\frac{1+\xi}{\xi^2}\right)^2 \Gamma(1+2\xi) + \frac{T_2}{\xi^4},$$

where $T_2 \to -3.039$ as $\xi \downarrow -1/2$. Noting that $(1+\xi)^2/\xi^4 \to 4$ as $\xi \downarrow -1/2$ shows that $\pi(\xi) < 2\left[\pi^2/6 + (1-\gamma)^2\right]^{1/2}(1+2\xi)^{-1/2}$ for $\xi \in (-1/2, -1/2+\epsilon)$, for some $\epsilon > 0$. In fact numerical work shows that $\epsilon \approx 1.29$.

A.7. Proof of Theorem 6

We show that the integral K_n^-, giving the contribution to the normalising constant from $\xi < -1$, diverges. From the proof of Theorem 4 we have

$$K_n^- = (n-1)! \int_{-\infty}^{-1} e^{-\gamma(1+\xi)}(-\xi)^{1-n} \int_{y_n}^{\infty} J(\phi,\xi)\left\{\sum_{i=1}^{n} |y_i - \phi|^{-1/\xi}\right\}^{-n} d\phi \, d\xi,$$

where $J(\phi,\xi) = \prod_{i=1}^{n} |y_i - \phi|^{-(1+1/\xi)}$. For $\xi < -1$ we have $-(1+1/\xi) < 0$ and $-1/\xi > 0$. Therefore, for $i = 2, \ldots, n$, $(\phi - y_i)^{-(1+1/\xi)} > (\phi - y_1)^{-(1+1/\xi)}$ and $(\phi - y_i)^{-1/\xi} < (\phi - y_1)^{-1/\xi}$, and thus the ϕ-integrand is greater than $n^{-n}(\phi - y_1)^{-n}$. Therefore,

$$K_n^- > (n-1)! \int_{-\infty}^{-1} e^{-\gamma(1+\xi)}(-\xi)^{1-n} \int_{y_n}^{\infty} n^{-n}(\phi - y_1)^{-n} \, d\phi \, d\xi$$

$$= (n-1)! \, n^{-n}(n-1)^{-1}(y_n - y_1)^{1-n} \int_{-\infty}^{-1} e^{-\gamma(1+\xi)}(-\xi)^{1-n} \, d\xi$$

$$= (n-2)! \, n^{-n}(y_n - y_1)^{1-n} e^{-\gamma} \int_{1}^{\infty} x^{1-n} e^{\gamma x} \, dx,$$

where $x = -\xi$. For all n this integral diverges so the result follows.

A.8. Proof of Theorem 7

We need to show that K_4 is finite. We split the range of integration over ξ in (A.6) so that $K_4 = J_1 + J_2 + J_3$, with respective contributions from $\xi < -1$, $-1 \le \xi \le 0$ and $\xi > 0$.

Proof that J_1 is finite. We use the substitution $u = (\phi - y_1)^{-1}$ in (A.6) to give

$$J_1 = 3! \int_{-\infty}^{-1}(-\xi)^{-3} \int_{y_4}^{\infty}\left\{\prod_{i=1}^{4}(\phi - y_i)^{-(1+1/\xi)}\right\}\left\{\sum_{i=1}^{4}(\phi - y_i)^{-1/\xi}\right\}^{-n} d\phi \, d\xi$$

$$= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \int_{0}^{1/\delta_4} u^2 \prod_{i=2}^{4} (1 - \delta_i u)^{-(1+1/\xi)} \left\{ 1 + \sum_{i=2}^{4} (1 - \delta_i u)^{-1/\xi} \right\}^{-4} \mathrm{d}u \, \mathrm{d}\xi.$$

A similar calculation to (A.3) gives

$$\prod_{i=2}^{4} (1 - \delta_i u)^{-(1+1/\xi)} \leq u^{-2(1+1/\xi)} \left\{ \prod_{i=2}^{3} (\delta_4 - \delta_i) \right\}^{-(1+1/\xi)} (1 - \delta_4 u)^{-(1+1/\xi)}.$$

Noting also that $1 + \sum_{i=2}^{4} (1 - \delta_i u)^{-1/\xi} \geq 1$ we have

$$J_1 \leq 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{3} (\delta_4 - \delta_i) \right\}^{-(1+1/\xi)} \int_{0}^{1/\delta_4} u^{-2/\xi} (1 - \delta_4 u)^{-(1+1/\xi)} \, \mathrm{d}u \, \mathrm{d}\xi$$

$$= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{3} (\delta_4 - \delta_i) \right\}^{-(1+1/\xi)} \beta \int_{0}^{1/\delta_4} u^{-2/\xi} \frac{1}{\beta} \left(1 + \frac{\xi u}{\beta} \right)^{-(1+1/\xi)} \mathrm{d}u \, \mathrm{d}\xi$$

$$= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{3} (\delta_4 - \delta_i) \right\}^{-(1+1/\xi)} \delta_n^{2/\xi-1} \frac{\Gamma(1 - 2/\xi)\Gamma(-1/\xi)}{\Gamma(1 - 3/\xi)} \, \mathrm{d}\xi,$$

where $\beta = -\xi/\delta_4$ and the last line follows from (A.2) with $a = 2$ and $\sigma = \beta$.

Therefore,

$$J_1 \leq 3! \int_{-\infty}^{-1} (-\xi)^{-3} (y_4 - y_1)^{2/\xi-1} \prod_{i=2}^{3} (y_4 - y_i)^{-(1+1/\xi)} \frac{\Gamma(1 - 2/\xi)\Gamma(-1/\xi)}{\Gamma(1 - 3/\xi)} \, \mathrm{d}\xi$$

$$= 3! \prod_{i=1}^{3} (y_4 - y_i)^{-1} \int_{-\infty}^{-1} (-\xi)^{-3} \left(\prod_{i=2}^{3} \frac{y_4 - y_i}{y_4 - y_1} \right)^{-1/\xi} \frac{\Gamma(1 - 2/\xi)\Gamma(-1/\xi)}{\Gamma(1 - 3/\xi)} \, \mathrm{d}\xi$$

$$= 3! \prod_{i=1}^{3} (y_4 - y_i)^{-1} \int_{0}^{1} x \left(\prod_{i=2}^{3} \frac{y_4 - y_i}{y_4 - y_1} \right)^{x} \frac{\Gamma(1 + 2x)\Gamma(x)}{\Gamma(1 + 3x)} \, \mathrm{d}x$$

$$= 3! \prod_{i=1}^{3} (y_4 - y_i)^{-1} \int_{0}^{1} \left(\prod_{i=2}^{3} \frac{y_4 - y_i}{y_4 - y_1} \right)^{x} \frac{\Gamma(1 + 2x)\Gamma(1 + x)}{\Gamma(1 + 3x)} \, \mathrm{d}x, \qquad \text{(A.16)}$$

where $x = -1/\xi$ and we have used the relation $\Gamma(1 + x) = x\,\Gamma(x)$. The integrand in (A.16) is finite over the range of integration so this integral is finite and therefore J_1 is finite.

Proof that J_2 is finite. Using the substitution $u = (\phi - y_1)^{-1}$ in (A.6) gives

$$J_2 = 3! \int_{-1}^{0} (-\xi)^{-3} \int_{0}^{1/\delta_4} u^2 \prod_{i=2}^{4} (1 - \delta_i u)^{-(1+1/\xi)} \left\{ 1 + \sum_{i=2}^{4} (1 - \delta_i u)^{-1/\xi} \right\}^{-4} \mathrm{d}u \, \mathrm{d}\xi.$$

For $-1 \le \xi \le 0$ we have $-(1 + 1/\xi) \ge 0$. Noting that $0 < 1 - \delta_i u < 1$ gives

$$\prod_{i=2}^{4}(1 - \delta_i u)^{-(1+1/\xi)} \le (1 - \delta_4 u)^{-(1+1/\xi)}.$$

Noting also that $1 + \sum_{i=2}^{4}(1 - \delta_i u)^{-1/\xi} \ge 1$ we have

$$J_2 \le 3! \int_{-1}^{0}(-\xi)^{-3} \int_{0}^{1/\delta_4} u^2(1 - \delta_4 u)^{-(1+1/\xi)} \, du \, d\xi$$

$$= 3! \int_{-1}^{0}(-\xi)^{-3}\beta \int_{0}^{1/\delta_4} u^2 \frac{1}{\beta}\left(1 + \frac{\xi u}{\beta}\right)^{-(1+1/\xi)} \, du \, d\xi$$

$$= 3\delta_4^{-3} \int_{-1}^{0} \frac{2}{(1-\xi)(1-2\xi)} \, d\xi$$

$$= 12(y_4 - y_1)^{-3}\ln(\frac{3}{2})$$

where $\beta = -\xi/\delta_4$ and the penultimate line follows from (A.2) with $r = 2$ and $\sigma = \beta$.

Proof that J_3 is finite. Using the substitution $u = (y_1 - \phi)^{-1}$ in (A.6) gives

$$J_3 = 3! \int_{0}^{\infty}\xi^{-3}\int_{-\infty}^{y_1}\left\{\prod_{i=1}^{4}(y_i - \phi)^{-(1+1/\xi)}\right\}\left\{\sum_{i=1}^{4}(y_i - \phi)^{-1/\xi}\right\}^{-4} d\phi \, d\xi$$

$$= 3! \int_{0}^{\infty}\xi^{-3}\int_{0}^{\infty} u^2 \prod_{i=2}^{4}(1 + \delta_i u)^{-(1+1/\xi)}\left\{1 + \sum_{i=2}^{4}(1 + \delta_i u)^{-1/\xi}\right\}^{-4} du \, d\xi.$$

Noting that for $\xi > 0$ we have $-(1 + 1/\xi) < 0$, using (A.5) with $a_k = \delta_k u$ gives

$$\prod_{i=2}^{4}(1 + \delta_i u)^{-(1+1/\xi)} \le (1 + gu)^{-3(1+1/\xi)},$$

where $g = (\delta_2\delta_3\delta_4)^{1/3}$. Noting also that $1 + \sum_{i=2}^{4}(1 + \delta_i u)^{-1/\xi} \ge 1$ we have

$$J_3 \le 3! \int_{0}^{\infty}\xi^{-3}\int_{0}^{\infty} u^2(1 + gu)^{-3(1+1/\xi)} \, du \, d\xi$$

$$\le 3! \int_{0}^{\infty}\xi^{-3}\beta \int_{0}^{\infty} u^2 \frac{1}{\beta}\left(1 + \frac{\alpha u}{\beta}\right)^{-(1+1/\alpha)} \, du \, d\xi,$$

where $\alpha = \xi/(2\xi + 3)$ and $\beta = \alpha/g$. Therefore, (A.1) with $r = 2$, $\sigma = \beta$ and $\xi = \alpha$ gives

$$J_3 \le 3! \int_{0}^{\infty}\xi^{-3}\beta\frac{2\beta^2}{(1 - \alpha)(1 - 2\alpha)} \, d\xi$$

$$
\begin{aligned}
&= 4g^{-3} \int_0^\infty \frac{1}{(\xi + 3)(2\xi + 3)} \, \mathrm{d}\xi \\
&= \frac{4}{3} g^{-3} \int_0^\infty \left(\frac{1}{\xi + 3/2} - \frac{1}{\xi + 3} \right) \, \mathrm{d}\xi \\
&= \frac{4}{3} g^{-3} \ln 2.
\end{aligned}
$$

The normalizing constant K_4 is finite, so $\pi_{U,GEV}(\mu, \sigma, \xi)$ yields a proper posterior density for $n = 4$ and therefore does so for $n \geq 4$.

References

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications.

Alzer, H. (1999). Inequalities for the gamma function. *Proc. Amer. Math. Soc.* **128**, 141-147.

Athreya, K. B. and Roy, V. (2014). Monte Carlo methods for improper target distributions. *Electron. J. Statist.* **8**, 2664-2692.

Bayarri, M. J. and Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statist. Sci.* **19**, 58-80.

Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., De Waal, D. and Ferro, C. (2004). *Statistics of Extremes.* Wiley.

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Anal.* **1**, 385-402.

Berger, J., Bernardo, J. and Sun, D. (2009). The formal definition of reference priors. *Ann. Statist.* **37**, 905-938.

Coles, S. G. (2001). *An Introduction to Statistical Modelling of Extreme Values.* Springer, London.

Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modelling: a review and new developments. *Int. Statist. Rev.* **64**, 119-136.

Coles, S. G. and Tawn, J. A. (1996). A Bayesian analysis of extreme rainfall data. *J. Roy. Statist. Soc. Ser. C* **45**, 463-478.

Coles, S. G. and Tawn, J. A. (2005). Bayesian modelling of extreme surges on the UK east coast. *Phil. Trans. Roy. Soc. Ser. A.* **363**, 1387-1406.

Datta, G., Mukerjee, R., Ghosh, M. and Sweeting, T. (2009). Bayesian prediction with approximate frequentist validity. *Ann. Statist.* **28**, 1414-1426.

Eugenia Castellanos, M. and Cabras, S. (2007). A default Bayesian procedure for the generalized Pareto distribution. *J. Statist. Plann. Inference* **137**, 473-483.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.* **24**, 180-190.

Giles, D. E. and Feng, H. (2009). Bias-corrected maximum likelihood estimation of the parameters of the generalized Pareto distribution. Econometric Working Paper EWP0902, University of Victoria.

Gradshteyn, I. S. and Ryzhik, I. W. (2007). *Table of Integrals, Series and Products.* 7th edition. Academic Press, New York.

Ho, K.-W. (2010). A matching prior for extreme quantile estimation of the generalized Pareto distribution. *J. Statist. Plann. Inference* **140**, 1513-1518.

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

Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27**, 251-261.

Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **81**, 158-171.

Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *J. Amer. Statist. Assoc.* **91**, 1343-1370.

Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. Imperial College Press, London.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

Martins, E. S. and Stedinger, J. R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour. Res.* **36**, 737.

Martins, E. S. and Stedinger, J. R. (2001). Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resour. Res.* **37**, 2551.

Mitrinović, D. A. (1964). *Elementary Inequalities*. Noordhoff, Groningen.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119-131.

Pickands, J. (1994). Bayes quantile estimation and threshold selection for the generalized Pareto family. In *Extreme Value Theory and Applications* (Edited by J. Galambos, J. Lechner, and E. Simiu), 123-138. Springer, US.

Qiu, S. L. and Vuorinen, M. (2004). Some properties of the gamma and psi functions, with applications. *Math. Comp.* **74**, 723-742.

Roy, V. and Dey, D. K. (2014). Propriety of posterior distributions arising in categorical and survival models under generalized extreme value distribution. *Statist. Sinica* **24**, 699-722.

Smith, E. (2005). Bayesian modelling of extreme rainfall data. Ph.D. thesis, Newcastle University.

Smith, R. L. (1984). Threshold methods for sample extremes. In *Statistical Extremes and Applications* (Edited by J. Oliveira), 621-638. Springer, Netherlands.

Smith, R. L. (1985). Maximum likelihood estimation in a class of non-regular cases. *Biometrika* **72**, 67-92.

Smith, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statist. Sci.* **4**, 367-377.

Smith, R. L. and Goodman, D. J. (2000). Bayesian risk analysis. In *Extremes and Integrated Risk Management* (Edited by P. Embrechts), 235-251. Risk Books, London.

Stephenson, A. and Tawn, J. A. (2004). Inference for extremes: accounting for the three extremal types. *Extremes* **7**, 291-307.

Tuyl, F., Gerlachy, R. and Mengersenz, K. (2009). Posterior predictive arguments in favor of the Bayes-Laplace prior as the consensus prior for binomial and multinomial parameters. *Bayesian Anal.* **4**, 151-158.

van Noortwijk, J. M., Kalk, H. J. and Chbab, E. H. (2004). Bayesian estimation of design loads. *HERON* **49**, 189-205.

Yang, R. and Berger, J. (1998). A catalog of noninformative priors. ISDS Discussion paper 97-42, Duke University. Available at `http://www.stats.org.uk/priors/noninformative/YangBerger1998.pdf`.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. Wiley.

Zellner, A. (1998). Past and recent results on maximal data information priors. *J. Statist. Res.* **32**, 1-22.

Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK.

E-mail: p.northrop@ucl.ac.uk

Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK.

E-mail: nicolas.attalides.10@ucl.ac.uk