Posterior propriety in objective Bayesian extreme value analyses

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 extreme value distribution we consider objective Bayesian inference using so-called “noninformative” or “default” prior distributions for the model parameters, specifically a Jeffreys prior, the maximal data information (MDI) prior and independent uniform priors on separate model parameters. We investigate the important issue of 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.

1 Introduction

Extreme value theory provides asymptotic justification for particular families of models for extreme data. Let $X_1, X_2, \ldots, X_N$ be a sequence of independent 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 GP$(\sigma_u, \xi)$ distribution, with threshold-dependent scale parameter $\sigma_u$, shape parameter $\xi$ and distribution function

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

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

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

Research Report number 323, Department of Statistical Science, University College London. Date: January 2014

†Address for correspondence: Paul Northrop, Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK. E-mail: p.northrop@ucl.ac.uk
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 $b \to \infty$ 


Commonly-used frequentist methods of inference for extreme value distributions are maximum likelihood estimation (MLE) and probability-weighted moments (PWM). However, conditions on $\xi$ 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 et al., 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 $\xi$ 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 objective analyses (Berger 2006). In the latter, a prior is constructed using a formal rule, providing an objective prior when no subjective information is to be incorporated into the analysis.

Many such formal rules have been proposed: Kass and Wasserman 1996 provides a comprehensive review. In this paper 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 improper; that is, they do not integrate to a finite number and therefore do not correspond to a proper probability distribution. 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 (Hobert and Casella 1996).

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($\mu, \sigma, \xi$) model for annual maximum sea-levels, placing independent normal priors on $\mu$, log $\sigma$ and $\xi$ 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 objective priors we consider in this paper. In section 3 we investigate whether or not these priors yield a proper posterior distribution when updated using a random sample $z = (z_1, \ldots, z_m)$ 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 $y = (y_1, \ldots, y_n)$ from a GEV distribution in section 4.

Proofs of results are presented in the appendix.
2 Objective priors for extreme value distributions

Let \( Y \) is a random variable with density function \( f(Y | \phi) \), indexed by a parameter vector \( \phi \) and define the Fisher information matrix \( I(\phi) \) by \( I_{ij} = \mathbb{E}[-\partial^2 \ln f(Y | \phi) / \partial \phi_i \partial \phi_j] \).

**Uniform priors.** Priors that are flat, i.e. equal to a positive constant, suffer from the problem that they are not automatically invariant to reparameterisation: for example, if we give \( \log \sigma \) a uniform distribution then \( \sigma \) is not uniform. Thus, it matters which particular parameterization is used to define the prior.

**Jeffreys priors.** Jeffreys’ “general rule” [Jeffreys 1961] is

\[
\pi_J(\phi) \propto \text{det}(I(\phi))^{1/2}.
\]

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 will follow this modification, which is summarised on page 1345 of [Kass and Wasserman 1996]. If there is no location parameter then (3) is used. If there is a location parameter \( \mu \), say, then \( \phi = (\mu, \theta) \) and

\[
\pi_J(\mu, \theta) \propto \text{det}(I(\theta))^{1/2},
\]

where \( I(\theta) \) is calculated holding \( \mu \) 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

\[
\pi_M(\phi) \propto \exp \{ \mathbb{E}[\log f(Y | \phi)] \}.
\]

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 \( \sigma \) rather than \( \sigma_u \). We consider a class of priors of the form \( \pi(\sigma, \xi) \propto \pi_\xi(\xi) / \sigma, \sigma > 0, \xi \in \mathbb{R} \), that is, a priori \( \sigma \) and \( \xi \) are independent and \( \log \sigma \) has an improper uniform prior over the real line.

The posterior is given by

\[
\pi_\sigma(\sigma, \xi | y) = C_m^{-1} \pi_\xi(\xi) \sigma^{-(m+1)} \prod_{i=1}^{m} (1 + \xi z_i / \sigma)^{-(1+1/\xi)}, \quad \sigma > 0, \xi > -\sigma / z_m,
\]

where

\[
C_m = \int_{-\infty}^{\infty} \int_{\max(0,-\xi z_m)}^{\infty} \pi_\xi(\xi) \sigma^{-(m+1)} \prod_{i=1}^{m} (1 + \xi z_i / \sigma)^{-(1+1/\xi)} \, d\sigma \, d\xi
\]

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 \( (3) \) with \( \phi = (\sigma, \xi) \) gives the Jeffrey’s prior

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

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

Using \( (5) \) gives the MDI prior

\[
\pi_{M,GP}(\sigma, \xi) \propto \frac{1}{\sigma} e^{-(\xi+1)} \propto \frac{1}{\sigma} e^{-\xi} \quad \sigma > 0, \xi \in \mathbb{R}.
\]

(7)

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

Placing independent uniform priors on \( \log \sigma \) and \( \xi \) gives the prior

\[
\pi_{U,GP}(\sigma, \xi) \propto \frac{1}{\sigma}, \quad \sigma > 0, \xi \in \mathbb{R},.
\]

(8)

This prior was proposed by Pickands (1994).

Figure 1 shows the Jeffreys and MDI priors for GP parameters as a functions of \( \xi \). The MDI prior increases without limit as \( \xi \to -\infty \).

![Scaled Jeffreys and MDI GP prior densities against \( \xi \).](image)

Figure 1: Scaled Jeffreys and MDI GP prior densities against \( \xi \).

3.2 Results

**Theorem 1.** A sufficient condition for the prior \( \pi(\sigma, \xi) \propto \pi_\xi(\xi)/\sigma, \sigma > 0, \xi \in \mathbb{R} \) to yield a proper posterior density function is that \( \pi_\xi(\xi) \) is (proportional to) a proper density function.
Proof. This trivial extension of the proof of theorem 1 in [Eugenia Castellanos and Cabras (2007)] is given in the appendix.

The MDI prior (7) 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 (7) yields a proper posterior density function.

**Proof.** See appendix.

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

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

that is, a (proper) unit exponential prior on \( \xi + 1 \). Any finite lower bound on \( \xi \) ensures propriety of the posterior but \( \xi = -1 \), for which the GP distribution reduces to a uniform distribution on \((0, \sigma)\), seems sensible. For smaller values of \( \xi \) the GP density is increasing over its support, which seems unlikely in an application where interest is in modelling the upper tail of a distribution.

**Theorem 3.** The truncated MDI prior (9) yields a proper posterior density function for \( m \geq 1 \).

**Proof.** This follows directly from theorem 1.

**Theorem 4.** A sufficient condition for the uniform prior (8) to yield a proper posterior density function is that \( m \geq 3 \).

**Proof.** See appendix.

### 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 \( \pi(\mu, \sigma, \xi) \propto \pi_\xi(\xi)/\sigma, \sigma > 0, \mu, \xi \in \mathbb{R} \) that is, a priori \( \mu, \sigma \) and \( \xi \) are independent and \( \mu \) and \( \log \sigma \) have a improper uniform priors over the real line.

Based on a random sample \( y_1, \ldots, y_n \) the posterior density for \( (\mu, \sigma, \xi) \) is proportional to

\[
\sigma^{-(n+1)} \pi_\xi(\xi) \left\{ \prod_{i=1}^{n} \left[ 1 + \xi \left( \frac{y_i - \mu}{\sigma} \right) \right]^{-1/(1+\xi)} \right\} \exp \left\{ - \sum_{i=1}^{n} \left[ 1 + \xi \left( \frac{y_i - \mu}{\sigma} \right) \right]^{-1/(1+\xi)} \right\},
\]

where \( \sigma > 0 \) and 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, page 63) give the Fisher information matrix for the GEV distribution. Using (4) with \( \phi = (\mu, \sigma, \xi) \) gives the Jeffrey’s prior

\[
\pi_{J,GEV}(\mu, \sigma, \xi) = \frac{1}{\sigma \xi^2} \left\{ \left[ 1 - 2\Gamma(2 + \xi) + \frac{\pi^2}{6} + \left( 1 - \frac{1}{\xi} \right)^2 \right] \left[ \frac{2q}{\xi} - p \right] \right\}^{1/2}, \mu \in \mathbb{R}, \sigma > 0, \xi > -1/2, \tag{11}
\]

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 \( \gamma \approx 1.57722 \) is Euler’s constant. van Noortwijk et al. (2004) give an alternative form for the Jeffrey’s prior, based on (3).

Beirlant et al. (2004, page 435) give the form of the MDI prior:

\[
\pi_{M,GEV}(\mu, \sigma, \xi) = \frac{1}{\sigma} e^{-\gamma(\xi+1+1/\gamma)} \propto \frac{1}{\sigma} e^{-\gamma(1+\xi)} \sigma > 0, \mu, \xi \in \mathbb{R}. \tag{12}
\]

Placing independent uniform priors on \( \mu, \log \sigma \) and \( \xi \) gives the prior

\[
\pi_{U,GEV}(\mu, \sigma, \xi) \propto \frac{1}{\sigma}, \sigma > 0, \mu, \xi \in \mathbb{R}. \tag{13}
\]

Figure 2 shows the Jeffreys and MDI priors for GEV parameters as a functions of \( \xi \). The MDI prior increases without limit as \( \xi \to -\infty \) and the Jeffrey’s prior increases without limit as \( \xi \to \infty \).

\[
\text{Jeffreys} \quad \text{MDI}
\]

Figure 2: Scaled Jeffreys and MDI GEV prior densities against \( \xi \).

4.2 Results

**Theorem 5.** For the prior \( \pi(\mu, \sigma, \xi) \propto \pi_\xi(\xi)/\sigma, \sigma > 0, \mu, \xi \in \mathbb{R} \) to yield a proper posterior density function it is necessary that \( n \geq 2 \) and, in that event, it is sufficient that \( \pi_\xi(\xi) \) is (proportional to) a proper density function.
Proof. See appendix.

**Theorem 6.** There is no sample size for which the independence Jeffreys prior (11) yields a proper posterior density function.

**Proof.** See appendix.

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 5 requires only that $\int_{-1/2}^{\xi_+} \pi_\xi(\xi) \, d\xi$ is finite. From the proof of theorem 6 we have $\pi_\xi(\xi) < 2 \left[ \frac{\pi^2}{6} + (1 - \gamma)^2 \right]^{1/2} (1 + 2\xi)^{-1/2}$ for $\xi \in (-1/2, -1/2 + \epsilon)$, where $\epsilon > 0$. Therefore,

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

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

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

**Theorem 7.** There is no sample size for which the MDI prior (12) yields a proper posterior density function.

**Proof.** See appendix.

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

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

(14)

is a sensible way to yield a proper posterior distribution.

**Theorem 8.** The truncated MDI prior (14) yields a proper posterior density function for $n \geq 2$.

**Proof.** This follows directly from theorem 5.

**Theorem 9.** A sufficient condition for the uniform prior (13) to yield a proper posterior density function is that $n \geq 4$.

**Proof.** See appendix.

## 5 Discussion

We have shown that some of the objective 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 $\xi$ priori. An interesting aspect of our findings is that the Jeffreys prior (11) for GEV parameters fails to yield a proper posterior, whereas the uniform prior (13) requires only weak conditions to ensure posterior propriety. This is the opposite of more general experience, summarised by Berger (2006, page 393) and Yang and Berger (1998, page 5), that 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 $\pi_\xi(\xi)$ of this prior increases for large $\xi$. An
alternative prior based on Jeffreys’ general rule \cite{vanNoortwijk2004} also has this property.

The conditions sufficient for posterior propriety under the uniform priors \cite{Beirlant2004} and \cite{Smith1989} 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 ones posterior is proper, with a sufficiently informative sample posterior impropriety might not present a practical problem \cite{Kass1996}. This may explain why \cite{Beirlant2004} 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 $p_u = P(X > u)$. Objective priors for a binomial probability have been studied extensively, see, for example, Tuyl et al. \cite{Tuyl2009}. An approximately equivalent approach is the non-homogeneous Poisson process (NHPP) model \cite{Smith1989}, which is parameterized in terms of GEV parameters $\mu, \sigma$ and $\xi$ 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 $(\mu, \sigma, \xi)$ is proportional to

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

(15)

where $n$ is the (notional) number of blocks into which the data are divided in defining $(\mu, \sigma, \xi)$. Without loss of generality, we take $n = m$. The exponential term in (15) 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 (15) is less than

$$
\sigma^{-(m+1)} \pi_{\xi}(\xi) \exp \left\{ -\sum_{i=1}^{m} \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \prod_{i=1}^{m} \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-1/\xi},
$$

(16)

Equation (16) is of the same form as (10), with $n = m$ and $y_i = x_i, i = 1, \ldots, n$. Therefore, theorems 5 and 9 apply to the NHPP model, that is, for posterior propriety it is sufficient that either (a) $n \geq 2$ and $\pi(\mu, \sigma, \xi) \propto \pi_{\xi}(\xi)/\sigma$, for $\sigma > 0, \mu, \xi \in \mathbb{R}$, where $\int_{\xi} \pi_{\xi}(\xi) \, d\xi$ 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 \cite{Roy2013} use GEV regression modelling to analyze reliability data. They prove posterior propriety under relatively strong conditions on the prior distribution placed on $(\sigma, \xi)$. The conditions required in our theorem 9 are weaker. We expect that the proof of theorem 7 in Roy and Dey \cite{Roy2013} could be adapted using approach we use to prove our theorem 9. Another extension is to explore other formal rules for constructing priors, such as reference priors \cite{Berger2009} and probability matching priors \cite{Datta2009}. Ho \cite{Ho2010} considers the latter for the GP distribution.
Acknowledgements

One of us (NA) was funded by an Engineering and Physical Sciences Research Council studentship while carrying out this work.

References


**Appendix**

**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]

$$
E(Z^r) = \frac{r! \sigma^r}{\prod_{i=1}^{r} (1 - i \xi)}.
$$

(17)

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

$$
E(Z^{-a/\xi}) = \int_{0}^{\sigma/\xi} v^{-a/\xi} \frac{1}{\sigma} \left(1 + \frac{\xi v}{\sigma}\right)^{-(1+1/\xi)} dv,
$$

$$
= (-\xi)^{a/\xi - 1} \sigma^{-a/\xi} \int_{0}^{1} x^{-a/\xi} (1 - x)^{-(1+1/\xi)} dx,
$$

$$
= (-\xi)^{a/\xi - 1} \sigma^{-a/\xi} \frac{\Gamma(1 - a/\xi) \Gamma(-1/\xi)}{\Gamma(1 - (a + 1)/\xi)}.
$$

(18)
where we have used integral number 1 in section 3.251 on page 324 of [Gradsteyn and Ryzhik (2007)], namely
\[
\int_0^1 x^{\mu-1} (1-x)^{\nu-1} \, dx = \frac{1}{\lambda} B \left( \frac{\mu}{\lambda}, \nu \right) = \frac{\Gamma(\mu/\lambda)\Gamma(\nu)}{\Gamma(\mu/\lambda + \nu)} \quad \lambda > 0, \nu > 0, \mu > 0,
\]
with \(\lambda = 1, \mu = 1 - a/\xi\) and \(\nu = -1/\xi\).

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

**Proof of theorem 1**

Suppose \(m = 1\), with an observation \(z\). The normalizing constant \(C\) of the posterior distribution is given by
\[
C_1 = \int_{-\infty}^{\infty} \pi(\xi) \int_{\frac{z}{\sigma}}^{\infty} \sigma^{-2} (1 + \xi z/\sigma)^{-(1+1/\xi)} \, d\sigma \, d\xi + \int_0^\infty \pi(\xi) \int_0^{\infty} \sigma^{-2} (1 + \xi z/\sigma)^{-(1+1/\xi)} \, d\sigma \, d\xi,
\]
\[
= \frac{1}{z} \int_{-\infty}^{\infty} \pi(\xi) \, d\xi.
\]

If the latter integral is finite, that is, \(\pi(\xi)\) 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 \(m \geq 1\).

**Proof of theorem 2**

Let \(A(\xi) = e^{-\xi}\) and \(B(\sigma, \xi) = \sigma^{-(m+1)} \prod_{i=1}^m (1 + \xi z_i/\sigma)^{-(1+1/\xi)}\). Then, from (6) we have
\[
C_m = \int_{-\infty}^{\infty} A(\xi) \int_{\max(0, -\xi z_m)}^{\infty} B(\sigma, \xi) \, d\sigma \, d\xi
\]
\[
= \int_{-\infty}^{\frac{z}{\sigma}} \int_{-\xi z_m}^{\infty} A(\xi) \, d\sigma \, d\xi + \int_0^0 \int_{-\xi z_m}^{\infty} B(\sigma, \xi) \, d\sigma \, d\xi + \int_0^0 \int_{-\xi z_m}^{\infty} B(\sigma, \xi) \, d\sigma \, d\xi.
\]

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(\(\sigma, \xi\)) density. Therefore \(B(\sigma, \xi) \geq \sigma^{-(m+1)}\). Thus, the first integral above satisfies
\[
\int_{-\infty}^{\frac{z}{\sigma}} \int_{-\xi z_m}^{\infty} A(\xi) \, d\sigma \, d\xi > \int_{-\infty}^{\frac{z}{\sigma}} \int_{-\xi z_m}^{\infty} \sigma^{-(m+1)} \, d\sigma \, d\xi,
\]
\[
= \int_{-\infty}^{\frac{z}{\sigma}} A(\xi) \left[ -\frac{1}{m} \sigma^{-m} \right]_{-\xi z_m}^{\infty} \, d\xi,
\]
\[
= \int_{-\infty}^{\frac{z}{\sigma}} A(\xi) \frac{1}{m} [-\xi z_m]^{-m} \, d\xi,
\]
\[
= \frac{1}{m z_m} \int_1^{\infty} v^{-m} e^v \, dv,
\]
where \(v = -\xi\). This integral is divergent for all \(m \geq 1\), so there is no sample size for which the posterior is proper.
Proof of theorem 4

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

$$ I_1 = \int_{-\infty}^{-1} \int_{-\xi z_3}^{\infty} B(\sigma, \xi) \; d\sigma \; d\xi, \quad I_2 = \int_{0}^{1} \int_{-\xi z_3}^{\infty} B(\sigma, \xi) \; d\sigma \; d\xi, \quad I_3 = \int_{1}^{\infty} \int_{-\xi z_3}^{\infty} B(\sigma, \xi) \; d\sigma \; d\xi $$

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

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

$$ (1 + \rho) (1 + \rho z_2)(1 + \rho z_3) \geq (\rho z_3 + \rho z_1)(\rho z_3 + \rho z_2)(1 + \rho z_3), $$

$$ = (\rho)^2 (z_3 - z_1)(z_3 - z_2)(1 + \rho z_3), $$

$$ = (-\xi)^{2} \sigma^{-2}(z_3 - z_1)(z_3 - z_2)(1 + \rho z_3). $$

Therefore,

$$ \prod_{i=1}^{3} \left(\frac{1 + \xi z_i}{\sigma}\right)^{-(1+1/\xi)} < (-\xi)^{2(1+1/\xi)} \sigma^{2(1+1/\xi)} \left[(z_3 - z_2)(z_3 - z_1)\right]^{-(1+1/\xi)} I_{1\sigma} \; d\xi. $$

Thus,

$$ I_1 \leq \int_{-\infty}^{-1} (-\xi)^{2(1+1/\xi)} \left[(z_3 - z_2)(z_3 - z_1)\right]^{-(1+1/\xi)} I_{1\sigma} \; d\xi, $$

where

$$ I_{1\sigma} = \int_{-\xi z_3}^{\infty} \sigma^{-4} \sigma^{-2(1+1/\xi)} \left(1 + \frac{\xi z_3}{\sigma}\right)^{-(1+1/\xi)} \; d\sigma, $$

$$ = z_3^{-1} \int_{0}^{1/\xi z_3} v^{-2/\xi} \left(1 + \frac{\xi v}{z_3}\right)^{-(1+1/\xi)} \; dv, $$

$$ = (-\xi)^{2(1-2/\xi)} z_3^{-(1-2/\xi)} \Gamma(1-2/\xi) \Gamma(-1/\xi) \Gamma(1-3/\xi). $$

where $v = 1/\sigma$ and the last line follows from (18) on noting that the integrand is $E(V^{-2/\xi})$, where $V \sim \text{GP}(z_3^{-1}, \xi)$. Therefore,

$$ I_1 \leq \int_{-\infty}^{-1} (-\xi)^{3} \left[(z_3 - z_2)(z_3 - z_1)\right]^{(1+1/\xi)} z_3^{-(1-2/\xi)} \Gamma(1-2/\xi) \Gamma(-1/\xi) \Gamma(1-3/\xi) \; d\xi, $$

$$ = [z_3(z_3 - z_2)(z_3 - z_1)]^{-1} \int_{-\infty}^{-1} (-\xi)^{3} \left(1 - \frac{z_2}{z_3}\right)^{x} \left(1 - \frac{z_1}{z_3}\right)^{-1/\xi} \frac{\Gamma(1-2/\xi) \Gamma(-1/\xi) \Gamma(1-3/\xi)}{\Gamma(1 + 3x)} \; d\xi, $$

$$ = [z_3(z_3 - z_2)(z_3 - z_1)]^{-1} \int_{0}^{1} x \left(1 - \frac{z_2}{z_3}\right)^{x} \left(1 - \frac{z_1}{z_3}\right)^{-1/\xi} \frac{\Gamma(1 + 2x) \Gamma(x)}{\Gamma(1 + 3x)} \; dx, $$

$$ = [z_3(z_3 - z_2)(z_3 - z_1)]^{-1} \int_{0}^{1} \left(1 - \frac{z_2}{z_3}\right)^{x} \left(1 - \frac{z_1}{z_3}\right)^{-1/\xi} \frac{\Gamma(1 + 2x) \Gamma(1 + x)}{\Gamma(1 + 3x)} \; dx, $$

where $x = -1/\xi$ and we have used the relation $\Gamma(1 + x) = x \Gamma(x)$. The integrand in (21) 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,

$$I_2 = \int_0^\infty \int_{-\infty}^0 \sigma^{-4} \prod_{i=1}^3 \left(1 + \frac{\xi z_i}{\sigma}\right)^{-(1+1/\xi)} \, d\sigma \, d\xi,$$

$$\leq \int_0^\infty \int_{-\infty}^0 \sigma^{-4} \left(1 + \frac{\xi z_1}{\sigma}\right)^{-(1+1/\xi)} \, d\sigma \, d\xi,$$

$$= \int_{-1}^0 \int_0^\infty \sigma^{-4} \left(1 + \frac{\xi z_3}{\sigma}\right)^{-(1+1/\xi)} \, d\sigma \, d\xi,$$

$$= \int_{-1}^0 \int_0^\infty \left(1 + \frac{\xi v}{z_1^{-1}}\right)^{-(1+1/\xi)} \, dv \, d\xi,$$

$$= z_1^{-1} \int_{-1}^0 \int_0^\infty 2z_1^{-2} \, dv \, d\xi,$$

$$= 2z_1^{-3} \int_{-1}^0 \left\{\left(\frac{1}{2} - \xi\right)^{-1} - (1 - \xi)^{-1}\right\} \, d\xi,$$

$$= 2z_1^{-3} \ln(3/2).$$

Proof that $I_3$ is finite

We have $\xi > 0$ and so $-(1 + 1/\xi) < 0$. We let $g_n = (\prod_{i=1}^n z_i)^{1/n}$. Using Mitrinović [1964, page 130):

$$\prod_{k=1}^n (1 + a_k) \geq (1 + b)^n, \quad a_k > 0; \quad \prod_{k=1}^n a_k = b^n, \quad (22)$$

with $a_k = \xi z_k/\sigma$ 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 $\beta = \alpha/\xi g_3 = 1/(3 + 2\xi)g_3$. We note that for $\xi > 0$, $\alpha < 1/2$ and so this inner integral is the second moment of a GP($\beta, \alpha$) distribution.
Therefore,
\[
I_3 \leq \int_0^\infty \frac{2\beta^2}{(1 - \alpha)(1 - 2\alpha)} \, d\xi,
\]
\[
= 2\int_0^\infty \frac{1}{(\xi + 3)(2\xi + 3)} \, d\xi,
\]
\[
= 2\int_0^\infty \left( \frac{1}{\xi + 3/2} - \frac{1}{\xi + 3} \right) \, d\xi,
\]
\[
= 2\frac{1}{9}g_3^{-3} \ln 2.
\]

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

**Proof of theorem 5**

Throughout the following proofs we define \(\delta_i = y_i - y_1, i = 2, \ldots, n\).

We make the parameter transformation \(\phi = \mu - \sigma/\xi\). Then the posterior density for \((\phi, \sigma, \xi)\) is given by
\[
\pi(\phi, \sigma, \xi) = K_n^{-1}\pi_\xi(\xi)|\xi|^{-n(1+1/\xi)}G_n(\phi, \sigma),
\]
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 \(b = |\xi|^{-1/\xi} \sum_{i=1}^n |y_i - \phi|^{-1/\xi}\) and \(v = \sigma^{1/\xi}\). The normalizing constant \(K_n\) is given by
\[
K_n = \int_{-\infty}^\infty \int_{0}^\infty \pi_\xi(\xi)|\xi|^{-n(1+1/\xi)}G_n(\phi, \sigma) \, d\sigma \, d\phi \, d\xi,
\]
\[
= \int_{-\infty}^\infty \pi_\xi(\xi)|\xi|^{-n(1+1/\xi)} \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} \int_0^\infty \sigma^{n/\xi - 1} \exp \left\{ -b\sigma^{1/\xi} \right\} \, d\sigma \, d\phi \, d\xi,
\]
\[
= \int_{-\infty}^\infty \pi_\xi(\xi)|\xi|^{-n(1+1/\xi)} \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} \int_0^\infty v^{n-1} \exp \left\{ -bv \right\} |\xi| \, dv \, d\phi \, d\xi,
\]
\[
= \int_{-\infty}^\infty \pi_\xi(\xi)|\xi|^{-n(1+1/\xi)} \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} \Gamma(n)b^{-n} |\xi| \, d\phi \, d\xi,
\]
\[
= \int_{-\infty}^\infty \pi_\xi(\xi)|\xi|^{-n(1+1/\xi)} \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} (n - 1)!|\xi|^{n/\xi + 1} \left\{ \sum_{i=1}^n |y_i - \phi|^{-1/\xi} \right\}^{-n} \, d\phi \, d\xi,
\]
\[
= (n - 1)! \int_{-\infty}^\infty \pi_\xi(\xi)|\xi|^{-1-n} \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} \left\{ \sum_{i=1}^n |y_i - \phi|^{-1/\xi} \right\}^{-n} \, d\phi \, d\xi,
\]
\[
= (n - 1)! \left\{ \prod_{i=1}^n |y_i - \phi|^{-(1+1/\xi)} \right\} \left\{ \sum_{i=1}^n |y_i - \phi|^{-1/\xi} \right\}^{-n} \, d\phi, \quad (23)
\]

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 (23) gives

\[
K_2^+ = \int_0^\infty \pi_\xi(\xi) \xi^{-1} \int_{-\infty}^{y_1} \frac{(y_1 - \phi)^{-(1+1/\xi)}(y_2 - \phi)^{-(1+1/\xi)}}{(y_1 - \phi)^{-1/\xi} + (y_2 - \phi)^{-1/\xi}}^2 \, d\phi \, d\xi,
\]

\[
\int_0^\infty \pi_\xi(\xi) \xi^{-1} \int_0^\infty u^{-2} \left(1 + \delta_2 u\right)^{-(1+1/\xi)} \, du \, d\xi,
\]

\[
\frac{1}{2} \delta_2^{-1} \int_0^\infty \pi_\xi(\xi) \, d\xi,
\]

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 \( \xi \) 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 (23) gives

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

Therefore,

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

Thus, \( K_2 \) is finite if \( \int_{-\infty}^\infty \pi_\xi(\xi) \, d\xi \) is finite, and the result follows. \( \square \)

**Proof of theorem 6**

The crucial aspects are the rates at which \( \pi_\xi(\xi) \to \infty \) as \( \xi \downarrow -1/2 \) and as \( \xi \to \infty \).

The component \( \pi_\xi(\xi) \) of (11) involving \( \xi \) can be expressed as

\[
\pi_\xi^2(\xi) = \frac{1}{\xi^4} (T_1 + T_2),
\]

where

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

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

Firstly, we derive a lower bound for \( \pi_\xi(\xi) \) that holds for \( \xi > 3 \). Using the duplication formula (Abramowitz and Stegun, 1972, page 256; 6.1.18)

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

with \( z = 1/2 + \xi \) in (25) we have

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

We note that

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

16
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,

$$T_1 > \left[\frac{\pi^2}{6} + (1 - \gamma)^2\right] \pi^{-1/2} (1 + \xi)^4 [\Gamma(1 + \xi)]^2. \tag{27}$$

Completing the square in (26) gives

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

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 $\xi$ and so $f(\xi)$ decreases with $\xi$. 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 page 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

$$T_2 > -(1 + \xi)^4 [\Gamma(1 + \xi)]^2. \tag{28}$$

Substituting (27) and (28) in (24) gives, for $\xi > 3$,

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

$$> c \Gamma[1 + \xi]^2,$$

$$> c (1 + \xi)^2 (\lambda - \gamma)$$

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 $\xi$ component of the Jeffreys prior (11) is given by

$$\pi^2(\xi) > c^{1/2} (1 + \xi)^{\lambda - \gamma}, \text{ for } \xi > 3. \tag{29}$$

[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 (23) gives

$$K_n^+ = (n - 1)! \int_2^\infty \pi^2(\xi) \xi^{1-n} \int_0^\infty u^{n-2} \prod_{i=1}^n \left[1 + \sum_{i=2}^n (1 + \xi u)^{-1/\xi}\right] \text{ du } \text{ d}x. \tag{30}$$
For $\xi > 0$ we have $1 + \sum_{i=2}^{n} (1 + \delta_i u)^{-1/\xi} \leq n$ and $\prod_{i=1}^{n} (1 + \delta_i u)^{-(1+1/\xi)} \geq (1 + \delta_n u)^{-(n-1)(1+1/\xi)}$. Applying these inequalities to (30) gives

$$K_n^+ \geq n^{-n}(n-1)! \prod_{i=1}^{n-2} \frac{1}{1-i\alpha} = (n-2)! \frac{n^{-2}\delta_n - n^{-2}}{(n-2)(n-1)}.$$

Substituting (32) into (31) gives

$$K_n^+ \geq n^{-n}(n-1)! (n-2)! \delta_n^{-1-n} \int_{2}^{\infty} \frac{1}{1-i\alpha} \prod_{i=1}^{n-2} \frac{1}{n-2-i\alpha} \pi_\xi(\xi) \, d\xi,$$

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 $\text{GP}(\beta, \alpha)$ distribution and so, using (17) with $r = n-2$, the integral over $u$ is given by

$$\int_{-\infty}^{\infty} e^{-\gamma(n+1/\xi)} \pi_\xi(-\xi) \, d\xi.$$

We show that the integral $K_n^+$, giving the contribution to the normalising constant from $\xi < 0$, diverges. From the proof of theorem 3 we have

$$K_n^+ = (n-1)! \int_{-\infty}^{1} e^{-\gamma(n+1/\xi)} (-\xi)^{-1-n} \int_{y_n}^{\infty} \left\{ \prod_{i=1}^{n} |y_i - \phi|^{-(1+1/\xi)} \right\} \left\{ \sum_{i=1}^{n} |y_i - \phi|^{-1/\xi} \right\}^{-n} \, d\phi \, d\xi.$$
For \( \xi < -1 \) we have \(-(1 + 1/\xi) < 0 \) and \(-1/\xi > 0\). Therefore, for \( i = 2, \ldots, n, \)
\((\phi - y_i)^{-1(1+1/\xi)} > (\phi - y_i)^{-1(1+1/\xi)} \) and \((\phi - y_i)^{-1/\xi} < (\phi - y_i)^{-1/\xi} \), and thus the \( \phi \)-integrand
is greater than \( n^{-n}(\phi - y_i)^{-n} \). Therefore,
\[
K_n^- > (n-1)! \int_{-\infty}^{-1} e^{-\gamma(1+\xi)} (-\xi)^{1-n} \int_{y_n}^{\infty} n^{-n}(\phi - y_i)^{-n} \, d\phi \, d\xi,
\]
\[
= (n-1)! n^{-n}(n-1)^{-1}(y_n - y_1)^{1-n} \int_{-\infty}^{-1} e^{-\gamma(1+\xi)} (-\xi)^{1-n} \, d\xi,
\]
\[
= (n-2)! n^{-n}(y_n - y_1)^{1-n} e^{-\gamma} \int_{1}^{\infty} x^{1-n} \, e^x \, dx.
\]

For all samples sizes \( n \) this integral diverges so the result follows.

**Proof of theorem 9**

We need to show that \( K_4 \) is finite. We split the range of integration over \( \xi \) in \([23]\) so that
\( K_4 = J_1 + J_2 + J_3 \), with respective contributions from \( \xi < -1, -1 \leq \xi \leq 0 \) and \( \xi > 0 \).

**Proof that \( J_1 \) is finite**

We use the substitution \( u = (\phi - y_i)^{-1} \) in \([23]\) to give
\[
J_1 = 3! \int_{-\infty}^{-1} (-\xi)^{-3} \int_{y_4}^{\infty} \left\{ \prod_{i=1}^{4} (\phi - y_i)^{-1(1+1/\xi)} \right\} \left\{ \sum_{i=1}^{4} (\phi - y_i)^{-1/\xi} \right\}^{-n} \, d\phi \, d\xi,
\]
\[
= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \int_{0}^{1/\delta_4} u^2 \prod_{i=2}^{4} (1 - \delta_i u)^{-(1+1/\xi)} \left\{ 1 + \sum_{i=2}^{4} (1 - \delta_i u)^{-1/\xi} \right\}^{-4} \, du \, d\xi.
\]

A similar calculation to \([20]\) gives
\[
\prod_{i=2}^{4} (1 - \delta_i u)^{-(1+1/\xi)} \leq u^{-2(1+1/\xi)} \left\{ \prod_{i=2}^{4} (\delta_i - \delta_i) \right\}^{-(1+1/\xi)} (1 - \delta_i u)^{-(1+1/\xi)}.
\]

Noting also that \( 1 + \sum_{i=2}^{4} (1 - \delta_i u)^{-1/\xi} \geq 1 \) we have
\[
J_1 \leq 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{4} (\delta_i - \delta_i) \right\}^{-(1+1/\xi)} \int_{0}^{1/\delta_4} u^{-2/\xi(1 - \delta_i u)^{-(1+1/\xi)}} \, du \, d\xi,
\]
\[
= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{4} (\delta_i - \delta_i) \right\}^{-(1+1/\xi)} \beta \int_{0}^{1/\delta_4} u^{-2/\xi} \frac{1}{\beta} \left( 1 + \frac{\xi u}{\beta} \right)^{-(1+1/\xi)} \, du \, d\xi,
\]
\[
= 3! \int_{-\infty}^{-1} (-\xi)^{-3} \left\{ \prod_{i=2}^{4} (\delta_i - \delta_i) \right\}^{-(1+1/\xi)} \delta_i^{2/\xi - 1} \frac{\Gamma(1 - 2/\xi)\Gamma(-1/\xi)}{\Gamma(1 - 3/\xi)} \, d\xi,
\]

19
where \( \beta = -\xi/\delta_4 \) and the last line follows from \([18]\) on noting that the integrand is \( E(U^{-2/\xi}) \), where \( U \sim \text{GP}(\beta, \xi) \). Therefore,

\[
J_1 \leq 3! \int_{-\infty}^{-1} (-\xi)^{-3} (y_4 - y_1)^{2/\xi - 1} (y_4 - y_2)^{-1/\xi} (y_4 - y_3)^{-1/\xi} \frac{\Gamma(1 - 2/\xi) \Gamma(-1/\xi)}{\Gamma(1 - 3/\xi)} \, d\xi, 
\]

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

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

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

(33)

where \( x = -1/\xi \) and we have used the relation \( \Gamma(1 + x) = x \Gamma(x) \). The integrand in (33) 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 (23) gives

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

For \(-1 \leq \xi \leq 0\) we have \(-1 + 1/\xi \geq 0\). Noting that \( 0 < 1 - \delta_i u < 1 \) gives

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

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

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

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

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

\[
= 3!(y_4 - y_1)^{-3} \ln(3/2)
\]

where \( \beta = -\xi/\delta_4 \) and the penultimate last line follows from \([18]\) on noting that the integrand is \( E(U^2) \), where \( U \sim \text{GP}(\beta, \xi) \).

**Proof that \( J_3 \) is finite**

Using the substitution \( u = (y_1 - \phi)^{-1} \) in (23) gives

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

\[
= 3! \int_{0}^{\infty} \xi^3 \int_{0}^{\infty} u^2 \prod_{i=2}^{4} (1 + \delta_i u)^{-(1+1/\xi)} \left\{ 1 + \sum_{i=2}^{4} (1 + \delta_i u)^{-1/\xi} \right\}^{-4} \, du \, d\xi.
\]
Noting that for $\xi > 0$ we have $-(1 + 1/\xi) < 0$, using (22) with $a_k = \delta_k u$ gives

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

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

$$J_3 \leq 3! \int_0^\infty \xi^3 \int_0^\infty u^2 (1 + g u)^{-3\left(1+1/\xi\right)} \, du \, d\xi,$$

$$\leq 3! \int_0^\infty \xi^3 \beta \int_0^\infty u^2 \frac{1}{\beta} \left(1 + \frac{\alpha u}{\beta}\right)^{-\left(1+1/\alpha\right)} \, du \, d\xi,$$

where $\alpha = \xi/(2\xi + 3)$ and $\beta = \alpha/g$. Therefore,

$$J_3 \leq 3! \int_0^\infty \xi^3 \beta \frac{2\beta^2}{(1 - \alpha)(1 - 2\alpha)} \, d\xi,$$

$$= 4g^{-3} \int_0^\infty \frac{1}{(\xi + 3)(2\xi + 3)} \, d\xi,$$

$$= \frac{4}{3} g^{-3} \int_0^\infty \left( \frac{1}{\xi + 3/2} - \frac{1}{\xi + 3} \right) \, d\xi,$$

$$= \frac{4}{3} g^{-3} \ln 2.$$

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