Computational Aspect of the Generalized Exponential Power Density

Alain Desgagné[*][†] Jean-François Angers[*][‡]

CRM-2918

April 2003

[*]Département de mathématiques et de statistique, Université de Montréal, C.P. 6128, Succ. "Centre-ville", Montréal, Québec H3C 3J7

[†]Email: desgagne@dms.umontreal.ca and

[‡]Email: angers@dms.umontreal.ca.

Abstract

In this paper, the generalized exponential power (GEP) distribution is studied. This family encompasses a vast majority of the usual distributions and some others using a simple transformation. The rich variety of its tails behavior makes the GEP density a natural benchmark to characterize and order tails of a large class of densities. Analytic formulas and numerical methods are proposed to evaluate its normalizing constant, its moments and its cdf. Furthermore, two methods to simulate observations from the GEP distribution are proposed. Some examples of simulations are presented. A numerical method using the GEP density is also given for the estimation of posterior moments when the prior and the likelihood are symmetric densities defined on the real line. Some examples of the behavior of the posterior density when an observation is an outlier are presented. It can be seen that the use of heavy-tailed distributions is a valuable tool in developing robust Bayesian procedures, limiting the influence that extreme information sources can have on posterior inferences.

Key words: Bayesian inference; heavy tail density, credence, importance sampling, outlier

Résumé

Nous voulons étudier dans cet article la famille de puissances d'exponentielle généralisée (GEP). Cette famille comprend la grande majorité des densités usuelles et quelques autres à une transformation près. La riche diversité du comportement des queues de la densité GEP en fait une densité de référence naturelle pour caractériser et ordonner les queues d'une grande classe de densités.

Des méthodes analytiques et numériques sont proposées afin d'évaluer sa constante de normalisation, ses moments et sa fonction de répartition. De plus, deux méthodes sont proposées pour simuler des observations provenant de la densité GEP. Quelques exemples de simulations sont présentées.

Une méthode numérique basée sur la densité GEP est aussi présentée pour l'estimation des moments *a posteriori* lorsque la loi *a priori* et la vraisemblance sont des densités symétriques définies sur les réels. Des exemples du comportement de la densité *a posteriori* lorsqu'on est en présence d'une valeur aberrante est présentée. On peut voir que l'utilisation de distributions à queues épaisses est un outil précieux pour développer des procédures bayésiennes robustes, limitant l'influence qu'une source d'information extrême peut avoir sur l'inférence *a posteriori*.

1 Introduction

The use of heavy-tailed distributions is a valuable tool in developing robust Bayesian procedures, limiting the influence that extreme information sources can have on posterior inferences, see for instance Meinhold and Singpurwalla (1989); O'Hagan (1990); Angers and Berger (1991); Carlin and Polson (1991); Angers (1992); Fan and Berger (1992); Geweke (1994); Angers (1996). O'Hagan (1990) introduced the notion of credence to characterize tails of a symmetric density on the real line. This notion has been generalized in Angers (2000), called p-credence, to accommodate a wider class of densities. Essentially, p-credence of a density is determined by comparing its tails to a reference density introduced in Angers (2000), called the generalized exponential power (GEP) density. (By using the notion of p-credence, a dominance relation can be established to compare two symmetric densities on $\mathbb{R}$. In case of conflict between the prior and the likelihood information, the Bayes rule (under the squared error loss) will collapse to the mean of the source of information with the heaviest tails, if they are sufficiently heavy.)

In this paper, we want to study the generalized exponential power distribution. This family encompasses a vast majority of the usual distributions (see Table 1) and some others using simple transformations. The rich variety of its tails behavior makes the GEP density a natural benchmark to characterize and order tails of a large class of densities. Of course, the modelling of tails by the GEP density is also possible, even though this subject is not tackled in this paper. The GEP density may also be a good candidate as the importance function in Monte Carlo simulations, especially for the estimation of posterior moments in Bayesian inference, as suggested in Section 6. Indeed, the simulation of observations from the GEP distribution is possible and it is addressed in this paper. Furthermore, characterization of the posterior tails using p-credence can be done (see Angers, 2000), which gives us the opportunity to choose the parameters of the GEP density in order to obtain an importance function with heavier tails than the posterior ones.

In Section 2 of this paper, we present the GEP density and we define the notion of p-credence to characterize and order tails of densities. In Sections 3 and 4, we determine its normalizing constant, moments and cdf. For some choices of parameters, analytic formulas are possible, but in general, numerical methods are needed. We also present two methods to simulate observations from the GEP distribution. For different choices of parameters, the inverse transformation method or the rejection method is used. Some examples of simulations are presented in Section 5.

Finally in Section 6, we see some applications of the GEP density and p-credence. We present a numerical method for the estimation of posterior moments when the prior and the likelihood are symmetric densities defined on the real line with their p-credence defined. We also present some examples of the behavior of the posterior density when an observation is an outlier. If we observe only one data point and if certain conditions are satisfied, we show that the posterior density will collapse to the prior density or to the likelihood when this observation goes to infinity. For a sample of several observations and under certain conditions, we show that if one observation goes to infinity, the posterior will collapse to the posterior density obtained from the sample where the outlier is omitted.

2 Generalized exponential power density and dominance relation using p-credence

In Section 2.1, we present the generalized exponential power density defined on the real line. We define the notion of p-credence in order to characterize the tails of a symmetric density by comparison to the GEP density. We also show how the densities can be ordered using a dominance relation. In

Section 2.2, we present the GEP density defined on (z_0, ∞) and we extend the notion of p-credence to characterize the right tail of a density.

2.1 Generalized exponential power density defined on the real line

The generalized exponential power family was introduced in Angers (2000). This class of distributions encompasses exponential, polynomial and logarithmic tails behavior (see Table 1). The general form of the density is given by

$$\begin{aligned} p(z|\gamma,\delta,\alpha,\beta,z_0) &\propto \exp\{-\delta \max(|z|,z_0)^\gamma\} \max(|z|,z_0)^{-\alpha} \log^{-\beta}[\max(|z|,z_0)] \\ &= \begin{cases} e^{-\delta|z|^\gamma} |z|^{-\alpha} \log^{-\beta}|z|; & \text{if } |z| > z_0, \\ e^{-\delta z_0^\gamma} z_0^{-\alpha} \log^{-\beta} z_0; & \text{if } |z| \leq z_0, \end{cases} \end{aligned} \tag{1}$$

where $z \in \mathbb{R}$, $\gamma \geq 0$, $\delta > 0$ (by convention we define $\delta = 0$ if $\gamma = 0$), $\alpha \in \mathbb{R}, \beta \in \mathbb{R}$ and $z_0 \geq 0$. (Note that the parameters α and β of the GEP density defined in Angers (2000) have been changed respectively to $-\alpha$ and $-\beta$ in order to ease the comparison of p-credences.) In addition, the parameters γ, δ, α, β and z_0 must satisfy the following conditions:

C1: $z_0 > \begin{cases} 1; & \text{if } \beta \neq 0, \\ 0; & \text{if } \alpha \neq 0, \beta = 0; \end{cases}$

C2: $\alpha + \frac{\beta}{\log(z_0)} \geq -\delta\gamma z_0^\gamma$;

C3: $\alpha \geq 1$ if $\gamma = 0$;

C4: $\beta > 1$ if $\gamma = 0, \alpha = 1$.

The first condition is needed in order for the density to be strictly positive and bounded. The second condition guarantees the unimodality of the density and it is always satisfied if z_0 is chosen to be large enough. The third and fourth conditions ensure that it is a proper density. The GEP density is symmetric with respect to the origin and is constant for $-z_0 \leq z \leq z_0$.

The GEP family was introduced to provide a benchmark for the characterization of the tails behavior of a density. Such characterization is addressed by the notion of p-credence, defined in Angers (2000) as follows:

Definition 1. *A density f on $\mathbb{R}$ has p-credence $(\gamma, \delta, \alpha, \beta)$, denoted by* p-cred$(f) = (\gamma, \delta, \alpha, \beta)$*, if there exist constants $0 < k \leq K < \infty$ such that for all $z \in \mathbb{R}$*

$$k \leq \frac{f(z)}{p(z|\gamma,\delta,\alpha,\beta,z_0)} \leq K,$$

where $p(z|\gamma,\delta,\alpha,\beta,z_0)$ is given by equation (1). We also write p-cred$(Z) = (\gamma, \delta, \alpha, \beta)$ *if the density of Z has p-credence $(\gamma, \delta, \alpha, \beta)$.*

The notion of p-credence characterizes the tails behavior of a density by comparing it to a GEP density. Essentially this definition ensures that $f(z)$ is of order $e^{-\delta|z|^\gamma}|z|^{-\alpha}\log^{-\beta}|z|$ for large values of $|z|$. Note that the parameter z_0 is not listed as an argument in p-credence since it has no influence on the tails behavior. By Definition 1, it is trivial to see that p-credence of $p(z|\gamma,\delta,\alpha,\beta,z_0)$ is $(\gamma, \delta, \alpha, \beta)$. It should also be noted that allowing γ to be negative would provide no more generality in the tails behavior of a density. Furthermore, most of the usual symmetric densities on $\mathbb{R}$ (such as the normal, Student-t, Laplace, logistic densities) are covered by this definition of p-credence.

Once the tails behavior of densities have been characterized by p-credence, a dominance relation can be established to compare them.

Definition 2. *Let f and g be any two densities on $\mathbb{R}$. We say that*

i) f dominates g, denoted by $f \succeq g$, if there exists a constant $k > 0$ such that

$$f(z) \geq kg(z)\, \forall z \in \mathbb{R};$$

ii) f is equivalent to g, denoted by $f \approx g$, if both $f \succeq g$ and $g \succeq f$, or written differently, if there exist constants $0 < k \leq K < \infty$ such that

$$k \leq \frac{f(z)}{g(z)} \leq K, \forall z \in \mathbb{R};$$

iii) f strictly dominates g, denoted by $f \succ g$, if $f \succeq g$ but $g \not\succeq f$.

Note that if p-cred$(f) = (\gamma, \delta, \alpha, \beta)$ then $f \approx p(\cdot|\gamma, \delta, \alpha, \beta, z_0)$, where $p(\cdot|\gamma, \delta, \alpha, \beta, z_0)$ is given by equation (1). The densities are ordered by the dominance relation as shown in Proposition 1.

Proposition 1. *Let f and g be two densities on $\mathbb{R}$ such that* p-cred$(f) = (\gamma', \delta', \alpha', \beta')$ *and* p-cred$(g) = (\gamma, \delta, \alpha, \beta)$*, then*

i) $f \approx g$ if $\gamma' = \gamma$, $\delta' = \delta$, $\alpha' = \alpha$ and $\beta' = \beta$;

ii) $f \succ g$ if:

a) $\gamma' < \gamma$;

b) $\gamma' = \gamma$, $\delta' < \delta$;

c) $\gamma' = \gamma$, $\delta' = \delta$, $\alpha' < \alpha$;

d) $\gamma' = \gamma$, $\delta' = \delta$, $\alpha' = \alpha$, $\beta' < \beta$.

If $f \approx g$, we say that f and g have the same p-credence. If $f \succ g$, we say that p-credence of g is greater than p-credence of f. In Bayesian inference for location parameter, Angers (2000) shows that under certain conditions, when a conflict between the prior and the likelihood informations occurs, the conflict is solved in favor of the density with the largest p-credence. Thus, the term p-credence has the sense of credibility given to a density in case of conflict (see O'Hagan, 1990).

The reason we defined α and β differently from Angers (2000) is to ease the comparison of p-credences. In fact, when we compare p-cred$(f) = (\gamma', \delta', \alpha', \beta')$ and p-cred$(g) = (\gamma, \delta, \alpha, \beta)$ using Proposition 1, we compare parameters from left to right. As soon as an inequality between two parameters occurs, we say that the density with the largest parameter has the largest p-credence.

2.2 Generalized exponential power density defined on (z_0, ∞)

The GEP density can be modified in order to be defined on (z_0, ∞) by considering only the right tail of the GEP density defined on $\mathbb{R}$. The density is then given by

$$p_1(z|\gamma, \delta, \alpha, \beta, z_0) \propto \begin{cases} \mathrm{e}^{-\delta z^{\gamma}} z^{-\alpha} \log^{-\beta} z; & z > z_0, \\ 0; & \text{otherwise}, \end{cases} \tag{2}$$

with the same conditions on parameters. However the condition C1 can be relaxed as:

C5: $z_0 \begin{cases} \geq 1; & \text{if } \beta < 1, \beta \neq 0, \\ > 1; & \text{if } \beta \geq 1, \\ > 0; & \text{if } \alpha \geq 1, \beta = 0. \end{cases}$

Condition C5 allows $\lim_{z \to z_0^+} p_1(z|\gamma, \delta, \alpha, \beta, z_0)$ to be 0 or infinity without making the density improper.

When the density is defined on (z_0, ∞), we are interested to the right tail of the distribution. The notion of p-credence being defined for symmetric densities on $\mathbb{R}$, it must be extended to the notion of right p-credence as follows:

Definition 3. *A density f defined on (a, ∞), $a \in \mathbb{R}$, has right p-credence $(\gamma, \delta, \alpha, \beta)$, denoted by* $\text{p-cred}^+(f) = (\gamma, \delta, \alpha, \beta)$, *if there exist constants $0 < k \leq K < \infty$ and $z_0 > \max(a, 0)$ such that for all $z > z_0$*

$$k \leq \frac{f(z)}{p_1(z|\gamma, \delta, \alpha, \beta, z_0)} \leq K,$$

where $p_1(z|\gamma, \delta, \alpha, \beta, z_0)$ is given by equation (2). We also write $\text{p-cred}^+(Z) = (\gamma, \delta, \alpha, \beta)$ *if the density of Z has right p-credence $(\gamma, \delta, \alpha, \beta)$.*

The notion of left p-credence, denoted by $\text{p-cred}^-(\cdot)$, can be defined in a similar way to characterize the left tail of a distribution. It is also possible to order the right tails of two densities by comparison of the parameters of their right p-credence from left to right as in Proposition 1. As soon as an inequality between two parameters occurs, we say that the density with the largest parameter has the largest right p-credence. The density with the smallest right p-credence has the heaviest right tail.

Another parameterization for the GEP density may be used if we translate $p_1(z|\gamma, \delta, \alpha, \beta, z_0)$ to the left up to 0. The density is then given by

$$p_2(z|\gamma, \delta, \alpha, \beta, z_0) \propto \mathrm{e}^{-\delta(z+z_0)^\gamma}(z + z_0)^{-\alpha} \log^{-\beta}(z + z_0), z > 0.$$

A variation of the GEP density defined on $\mathbb{R}$ is also possible if $p_2(z|\gamma, \delta, \alpha, \beta,\ z_0)$ is flipped by symmetry on the negative line. The density is then given by

$$p_3(z|\gamma, \delta, \alpha, \beta, z_0) \propto \mathrm{e}^{-\delta(|z|+z_0)^\gamma}(|z| + z_0)^{-\alpha} \log^{-\beta}(|z| + z_0), z \in \mathbb{R}\,.$$

It is essentially the GEP density on $\mathbb{R}$ defined earlier by equation (1), with the uniform part on $[-z_0, z_0]$ removed and with the tails translated to 0. The two densities $p_2(z|\gamma, \delta, \alpha, \beta, z_0)$ and $p_3(z|\gamma, \delta, \alpha, \beta, z_0)$ are not considered in this paper, but it is trivial to extend the theory to these densities.

3 Generalized exponential power density when $\gamma = 0, \delta = 0$ or $\beta = 0$

In Sections 3 and 4, we want to evaluate the normalizing constant, the moments and the cdf of the GEP distribution. We also want to simulate observations from it. To achieve this, we consider three different sets of parameters.

The first two cases considered in this section are $\gamma = 0, \delta = 0$ and $\beta = 0$. For these cases, the normalizing constant, the moments and the cdf of the distribution can be evaluated analytically.

Furthermore, the algorithm to generate random variables is a direct application of the inverse transformation method.

The last case, considered in Section 4, is $\gamma > 0, \delta > 0, \beta \neq 0$. In this case, Monte Carlo with importance sampling is needed to evaluate the normalizing constant, the moments and the cdf, and the rejection method is needed to simulate observations.

3.1 Calculation of the normalizing constant, the moments and the cumulative distribution function

For $j \geq 0$, let $H_j(\gamma, \delta, \alpha, \beta, z_0)$ and $I_j(\gamma, \delta, \alpha, \beta, z_0)$ defined as

$$H_j(\gamma, \delta, \alpha, \beta, z_0) = \int_{z_0}^{\infty} z^j \, \mathrm{e}^{-\delta z^\gamma} \, z^{-\alpha} \log^{-\beta} z \, dz,$$

and

$$\begin{aligned} &I_j(\gamma, \delta, \alpha, \beta, z_0) \\ &\quad = \int_{-\infty}^{\infty} |z|^j \exp\left\{-\delta \max(|z|, z_0)^\gamma\right\} \max(|z|, z_0)^{-\alpha} \log^{-\beta}\left[\max(|z|, z_0)\right] dz \\ &\quad = 2\left[\frac{1}{j+1} \, \mathrm{e}^{-\delta z_0^\gamma} \, z_0^{1-\alpha+j} \log^{-\beta} z_0 + H_j(\gamma, \delta, \alpha, \beta, z_0)\right]. \end{aligned} \tag{3}$$

If the GEP density is defined on $\mathbb{R}$, then the normalizing constant is equal to $I_0^{-1}(\gamma, \delta, \alpha, \beta, z_0)$ and the j^{th} moment of $|Z|$ is given by

$$\mathbb{E}\left(|Z|^j\right) = \frac{I_j(\gamma, \delta, \alpha, \beta, z_0)}{I_0(\gamma, \delta, \alpha, \beta, z_0)}, \; j > 0.$$

Note that if $\mathbb{E}\left(|Z|^j\right) < \infty$ and j is an integer,

$$\mathbb{E}\left(Z^j\right) = \begin{cases} 0; & \text{if } j \text{ is odd,} \\ \mathbb{E}\left(|Z|^j\right); & \text{if } j \text{ is even.} \end{cases}$$

Furthermore, we can show that the cdf is given by

$$\Pr[Z \leq k] = \begin{cases} \frac{1}{2}\left[1 + \frac{k}{z_0}\left(1 - \frac{H_0(\gamma,\delta,\alpha,\beta,z_0)}{\mathrm{e}^{-\delta z_0^\gamma} z_0^{1-\alpha} \log^{-\beta} z_0 + H_0(\gamma,\delta,\alpha,\beta,z_0)}\right)\right]; & \text{if } 0 \leq k \leq z_0, \\ 1 - \frac{1}{2}\left(\frac{H_0(\gamma,\delta,\alpha,\beta,k)}{\mathrm{e}^{-\delta z_0^\gamma} z_0^{1-\alpha} \log^{-\beta} z_0 + H_0(\gamma,\delta,\alpha,\beta,z_0)}\right); & \text{if } k > z_0. \end{cases}$$

Since the density is symmetric with respect to 0, it is easy to obtain $\Pr[Z \leq k]$ for $k < 0$.

If the GEP density is defined on (z_0, ∞) then the normalizing constant is equal to $H_0^{-1}(\gamma, \delta, \alpha, \beta, z_0)$ and the j^{th} moment of Z is given by

$$\mathbb{E}\left(Z^j\right) = \frac{H_j(\gamma, \delta, \alpha, \beta, z_0)}{H_0(\gamma, \delta, \alpha, \beta, z_0)}.$$

Furthermore, we can show that the cdf is given by

$$\Pr_1[Z \leq k] = 1 - \frac{H_0(\gamma, \delta, \alpha, \beta, k)}{H_0(\gamma, \delta, \alpha, \beta, z_0)}.$$

The challenge is to compute $H_j(\gamma, \delta, \alpha, \beta, z_0)$. If $\gamma = 0, \delta = 0$ or $\beta = 0$, and the conditions C3 to C5 are respected, we can show, for $j \geq 0$, that

$$H_j(\gamma, \delta, \alpha, \beta, z_0) =$$

$$\begin{cases} \frac{\Gamma(\frac{1-\alpha+j}{\gamma}, \delta z_0^\gamma)}{\gamma \delta^{\frac{1-\alpha+j}{\gamma}}}; & \text{if } \gamma > 0, \delta > 0, \beta = 0, j \geq 0, \\ \frac{z_o^{-(\alpha-j-1)}}{\alpha-j-1}; & \text{if } \gamma = 0, \delta = 0, \beta = 0, j < \alpha - 1, \\ \frac{\Gamma(1-\beta, [\alpha-j-1]\log z_0)}{(\alpha-j-1)^{1-\beta}}; & \text{if } \gamma = 0, \delta = 0, \beta \neq 0, j < \alpha - 1, \\ \frac{(\log z_0)^{-(\beta-1)}}{\beta-1}; & \text{if } \gamma = 0, \delta = 0, \beta > 1, j = \alpha - 1, \\ \infty; & \text{otherwise,} \end{cases} \tag{4}$$

where $\Gamma(\lambda, a)$ is the incomplete gamma function defined by

$$\Gamma(\lambda, a) = \int_a^\infty \mathrm{e}^{-u}\, u^{\lambda-1} du,$$

$\lambda \in \mathbb{R}, a > 0\, (a \geq 0$ if $\lambda > 0)$. In particular, when $a = 0$ and $\lambda > 0$, $\Gamma(\lambda, 0)$ is the gamma function and is denoted by $\Gamma(\lambda)$.

3.2 Special cases of the GEP distribution

Many known distributions (see Johnson, Kotz and Balakrishnan, 1994) are special cases of the GEP distribution. They are given in Table 1 along with their sets of parameters, support, normalizing constant and moments.

The first three entries correspond to the exponential power (see Box and Tiao, 1962), normal and Laplace densities, each of them defined on $\mathbb{R}$. The five following entries correspond to the generalized gamma distribution with four of its special cases, that is the Weibull, Rayleigh, Maxwell-Boltzmann and gamma distributions, each of them defined on $(0, \infty)$. The last three entries listed in Table 1 are heavy tail distributions defined on (z_0, ∞), that is the Pareto, log-gamma and log-Pareto distributions. For each of them, the parameters γ and δ of the GEP density are set to 0. The log-Pareto (see Reiss and Thomas, 1997) is the GEP distribution with the heaviest right-tail. Though it is a proper density, all j^{th} moments, $j > 0$, are infinite.

Other known distributions correspond to a transformation of the GEP distribution.

i) If X is GEP$(\gamma, \delta, 1-\gamma, 0, 0)$ with $\gamma, \delta, x > 0$ (Weibull), then $Z = 1/X$ has a Fréchet distribution, defined on $(0, \infty)$, with right p-credence $(0, 0, 1 + \gamma, 0)$.

ii) If X is GEP$(1, 1, 0, 0, 0)$, $x > 0$ (exponential), then $Z = -\log X$ has a Gumbel distribution, defined on $\mathbb{R}$, with right p-credence $(1, 1, 0, 0)$.

iii) If X is GEP$(1, \delta, 1 - \tau, 0, 0)$ with $\delta, \tau, x > 0$ (gamma), then $Z = 1/X$ has an inverse-gamma distribution, defined on $(0, \infty)$, with right p-credence $(0, 0, 1 + \tau, 0)$.

iv) If X is GEP$(0,0,2,0,1)$, $x>1$ (Pareto), then $Z=\log(X-1)$ has a logistic distribution, defined on $\mathbb{R}$, with p-credence $(1,1,0,0)$.

v) If X is GEP$(0,0,1+\alpha,0,1)$, $\alpha>0, x>1$ (Pareto), then $Z=(X-1)^{1/\beta}$, $\beta>0$, has a Burr distribution, defined on $(0,\infty)$, with right p-credence $(0,0,\alpha\beta+1,0)$.

Note that the Gumbel, Fréchet and Weibull are respectively the extreme value distributions of type I, II and III (the parameterization of the Weibull distribution differs in extreme value theory, see Reiss and Thomas, 1997 or Kotz and Nadarajah, 2000).

Though some distributions are not a special case of the GEP distribution, their p-credence is still defined. For example p-credence of the Student-t distribution with ν degrees of freedom is $(0,0,\nu+1,0)$ and right p-credence of a F-distribution with r_1 and r_2 degrees of freedom is $(0,0,1+r_2/2,0)$.

——————— Table 1 should appear here ———————

3.3 Simulation of observations

We first determine an algorithm to generate random variables from the GEP density defined on the real line. It is natural to divide the support into the three intervals $(-\infty,-z_0]$, $(-z_0,z_0]$ and (z_0,∞) since the GEP density is symmetric with respect to 0 and is constant between $-z_0$ and z_0. Let $Z\sim p(z|\gamma,\delta,\alpha,\beta,z_0)$. The algorithm to generate Z is divided in five steps as follows:

1) Compute

$$q_0=\Pr[-z_0<Z\le z_0]=\begin{cases}\left[1+\frac{H_0(\gamma,\delta,\alpha,\beta,z_0)}{e^{-\delta z_0^\gamma}\,z_0^{1-\alpha}\log^{-\beta}z_0}\right]^{-1}; & \text{if } z_0>0,\\ 0; & \text{if } z_0=0.\end{cases}$$

2) Generate u from a uniform distribution on $(0,1)$ and

   i) if $0\le u\le\frac{1-q_0}{2}$ then go to step 5 to simulate an observation on $(-\infty,-z_0]$;

   ii) if $\frac{1-q_0}{2}<u\le\frac{1+q_0}{2}$ then go to step 3 to simulate an observation on $(-z_0,z_0]$;

   iii) if $\frac{1+q_0}{2}<u\le 1$ then go to step 4 to simulate an observation on (z_0,∞).

3) Let $z=\frac{z_0(2u-1)}{q_0}$. Return to step 2 to simulate another observation.

4) Let $w=\frac{2u-1-q_0}{1-q_0}$. If conditions C3 to C5 are respected, then let

$$z=\begin{cases}\left[\frac{1}{\delta}F^{-1}_{\left(\frac{1-\alpha}{\gamma}\right)}\left(w+(1-w)F_{\left(\frac{1-\alpha}{\gamma}\right)}(\delta z_0^\gamma)\right)\right]^{1/\gamma}; & \\ & \text{if } \gamma>0,\delta>0,\alpha<1,\beta=0,\\ \frac{z_0}{w^{\frac{1}{\alpha-1}}}; & \text{if } \gamma=0,\delta=0,\beta=0,\\ \exp\left\{\frac{F^{-1}_{(1-\beta)}\left(w+(1-w)F_{(1-\beta)}((\alpha-1)\log z_0)\right)}{\alpha-1}\right\}; & \\ & \text{if } \gamma=0,\delta=0,\beta<1(\beta\neq 0),\\ \exp\left\{\frac{\log z_0}{w^{\frac{1}{\beta-1}}}\right\}; & \text{if } \gamma=0,\delta=0,\alpha=1,\beta>1,\end{cases}\tag{5}$$

where $F_{(\lambda)}(\cdot)$ is the cdf of a gamma distribution with shape and scale parameters respectively equal to $\lambda > 0$ and 1. $F^{-1}_{(\lambda)}(\cdot)$ is its inverse cdf. Return to step 2 to simulate another observation.

5) Let $w = \frac{2u}{1-q_0}$. Obtain z using equation (5) and change its sign. Return to step 2 to simulate another observation.

The inverse transformation method is used in step four. Here, to generate z, only one uniform random variable is needed.

Note that the two cases with parameters set to $\gamma > 0, \delta > 0, \alpha \geq 1, \beta = 0$ and $\gamma = 0, \delta = 0, \beta \geq 1$ are not considered in the algorithm. These cases are considered in Section 4.3 with the rejection method.

It is also possible to simulate an observation from the GEP distribution defined on the support (z_0, ∞). First generate randomly w from a uniform distribution on $(0, 1)$ and obtain z using equation (5).

Most of the distributions listed in Table 1 satisfy the first condition of equation (5), except the Pareto, log-gamma and log-Pareto distributions, which satisfy respectively the second, third and fourth conditions.

4 Generalized exponential power density when $\gamma > 0, \delta > 0, \beta \neq 0$

In this section we study the GEP distribution when $\gamma > 0, \delta > 0, \beta \neq 0, (\alpha \in \mathbb{R}, z_0 > 1)$. The normalizing constant, the moments and the cdf cannot be evaluated analytically. Furthermore the inverse transformation method is not appropriate to generate random variables. Instead of it, Monte Carlo simulations with importance sampling is used to evaluate the normalizing constant, the moments and the cdf, and the rejection method is used to generate random variables. Both methods require the use of a proposal distribution from which we can generate directly. The choice of this distribution is the subject of Section 4.1. Monte Carlo and rejection methods are addressed respectively in Sections 4.2 and 4.3. Two examples of these methods are discussed in Section 5.

4.1 The choice of a proposal distribution

We first consider the GEP density defined on $\mathbb{R}$, given by $p(z|\gamma, \delta, \alpha, \beta, z_0)$. We search a proposal distribution g defined on $\mathbb{R}$ to generate random variables directly. Ideally, $g(z)$ should dominate and be as close as possible to $p(z|\gamma, \delta, \alpha, \beta, z_0)$. A criteria which reflects these objectives consists in choosing g such that

$$w = \sup_z \frac{p(z|\gamma, \delta, \alpha, \beta, z_0)}{g(z)}$$

is as small as possible. Note that $w \geq 1$ since both $p(z|\gamma, \delta, \alpha, \beta, z_0)$ and $g(z)$ are proper densities, with $w = 1$ if $g(z) = p(z|\gamma, \delta, \alpha, \beta, z_0)$ $\forall z \in \mathbb{R}$. Furthermore, g must dominate $p(z|\gamma, \delta, \alpha, \beta, z_0)$ according to the criteria, otherwise w is infinite. This criteria ensures that the probability of acceptation in the rejection method is maximized (see Robert, 1996). This also ensures that the supremum of the weight function in the Monte Carlo simulations with importance sampling is minimized.

We propose the GEP density given by $p(z|\gamma, \delta, \alpha^*, 0, z_0)$ as the proposal distribution $g(z)$. We require the condition $\alpha^* < 1$ to ensure that we can generate random variables directly from $p(z|\gamma, \delta, \alpha^*, 0, z_0)$ using the algorithm of Section 3.3.

In order to make w as small as possible, the next step is to choose α^* (under the constraint $\alpha^* < 1$) as the parameter α_0 which minimizes $w(\alpha_0)$, where

$$\begin{aligned} w(\alpha_0) &= \sup_z \frac{p(z|\gamma,\delta,\alpha,\beta,z_0)}{p(z|\gamma,\delta,\alpha_0,0,z_0)} \\ &= \sup_z \frac{I_0(\gamma,\delta,\alpha_0,0,z_0)}{I_0(\gamma,\delta,\alpha,\beta,z_0)} \max[|z|\,,z_0]^{\alpha_0-\alpha} \log^{-\beta} \max[|z|\,,z_0]. \end{aligned}$$

Proposition 2. *The parameter α_0 which minimizes $w(\alpha_0)$ is given by*

$$\alpha^* = \begin{cases} \arg\min_{\alpha_0 \in \left[\alpha + \frac{\beta}{\log z_0}, \min(1,\alpha)\right)} I_0(\gamma,\delta,\alpha_0,0,z_0) \left(\frac{\beta}{\alpha_0-\alpha}\right)^{-\beta}; & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textit{if } \alpha + \frac{\beta}{\log z_0} < \min(1,\alpha), \\ \min(1-\epsilon,\alpha); \qquad\qquad\qquad\qquad\qquad\qquad \textit{otherwise}\,, \end{cases}$$

where $\epsilon > 0$ and $I_0(\gamma,\delta,\alpha_0,0,z_0)$ is given by equations (3) and (4). Furthermore, $w(\alpha^*)$ is given by

$$w(\alpha^*) = \begin{cases} \frac{I_0(\gamma,\delta,\alpha^*,0,z_0)}{I_0(\gamma,\delta,\alpha,\beta,z_0)} \mathrm{e}^\beta \left(\frac{\beta}{\alpha^*-\alpha}\right)^{-\beta}; & \textit{if } \alpha + \frac{\beta}{\log z_0} < \min(1,\alpha), \\ \frac{I_0(\gamma,\delta,\alpha^*,0,z_0)}{I_0(\gamma,\delta,\alpha,\beta,z_0)} z_0^{\alpha^*-\alpha} \log^{-\beta} z_0; & \textit{otherwise.} \end{cases}$$

In practice, if $\alpha^* = 1-\epsilon$ and ϵ is chosen too close to 0, numerical problems can arise in the generation process. (For most sets of parameters considered in this paper, $\epsilon = 0.01$ seems appropriate.)

The minimization of $I_0(\gamma,\delta,\alpha_0,0,z_0)\left(\frac{\beta}{\alpha_0-\alpha}\right)^{-\beta}$ with respect to α_0 has to be done numerically since no analytic answer is available. However it can be shown that this function is strictly convex.

Note that $\alpha^* = \min(1-\epsilon,\alpha)$ if $\beta > 0$ and $\alpha^* = 1-\epsilon$ if $\beta < 0$ and $\alpha + \frac{\beta}{\log z_0} \geq \min(1,\alpha)$. Furthermore, we can see, using Proposition 1, that $p(z|\gamma,\delta,\alpha,\beta,z_0)$ is dominated by $p(z|\gamma,\delta,\alpha^*,0,z_0)$ since $\alpha > \alpha^*$ or $\alpha = \alpha^*, \beta > 0$.

If the GEP density is defined on (z_0,∞), the choice of α^* is also given by Proposition 2, except that $I_0(\gamma,\delta,\alpha_0,0,z_0)$ is replaced by $H_0(\gamma,\delta,\alpha_0,0,z_0)$.

4.2 Monte Carlo with importance sampling

In this section, we first want to estimate the normalizing constant, moments and cdf of the GEP density defined on $\mathbb{R}$. Using $p(z|\gamma,\delta,\alpha^*,0,z_0)$ as the importance function, the algorithm of Monte Carlo simulations is divided in three steps.

1) Generate $z_1, z_2, ..., z_m$ from the density $p(z|\gamma,\delta,\alpha^*,0,z_0)$.

2) Compute the weight function for $i = 1, ..., m$ as

$$\begin{aligned} v(z_i) &= \frac{\exp\left\{-\delta \max(|z_i|\,,z_0)^\gamma\right\} \max(|z_i|\,,z_0)^{-\alpha} \log^{-\beta}\left[\max(|z_i|\,,z_0)\right]}{p(z_i|\gamma,\delta,\alpha^*,0,z_0)} \\ &= I_0(\gamma,\delta,\alpha^*,0,z_0) \max(|z_i|\,,z_0)^{\alpha^*-\alpha} \log^{-\beta}\left[\max(|z_i|\,,z_0)\right]. \end{aligned}$$

3) - Estimate the normalizing constant $I_0^{-1}(\gamma,\delta,\alpha,\beta,z_0)$ by
$\left[\frac{1}{m}\sum_{i=1}^m v(z_i)\right]^{-1}$,

- Estimate $\mathbb{E}\left(|Z|^j\right)$ by $\frac{\sum_{i=1}^m |z_i|^j v(z_i)}{\sum_{i=1}^m v(z_i)}$,
- Estimate $\Pr[Z \leq k]$ by $\frac{\sum_{i=1}^m \mathbb{I}_{[z_i \leq k]} v(z_i)}{\sum_{i=1}^m v(z_i)}$, where $\mathbb{I}_{[z_i \leq k]}$ is the indicator function of the set $\{z_i \leq k\}$.

Note that when we chose $\alpha^* = \arg\min_{\alpha_0} w(\alpha_0)$ in Proposition 2, we also minimized $\sup_z v(z)$, which makes the weights more stable.

If the GEP density is defined on (z_0, ∞), the algorithm is the same, except the z_i's are generated from $p_1(z|\gamma, \delta, \alpha^*, 0, z_0)$ and $I_0(\gamma, \delta, \alpha^*, 0, z_0)$ is replaced by $H_0(\gamma, \delta, \alpha^*, 0, z_0)$ in $v(z_i)$. Note that α^* should be chosen according to Proposition 2 replacing $I_0(\gamma, \delta, \alpha_0, 0, z_0)$ by $H_0(\gamma, \delta, \alpha_0, 0, z_0)$.

4.3 Rejection method

In this section, we want to simulate observations from $p(z|\gamma, \delta, \alpha, \beta, z_0)$. We use the algorithm of the rejection method (see Ross, 1997 or Lange, 1999) with $p(z|\gamma, \delta, \alpha^*, 0, z_0)$ as the proposal density, as follows:

1) Generate z^* from $p(z|\gamma, \delta, \alpha^*, 0, z_0)$.

2) Generate u from a uniform distribution on $(0, 1)$.

3) If $0 < u < B(z^*)$ then set $z = z^*$, otherwise return to step 1,

where

$$\begin{aligned} B(z^*) &= \frac{1}{w(\alpha^*)} \frac{p(z^*|\gamma, \delta, \alpha, \beta, z_0)}{p(z^*|\gamma, \delta, \alpha^*, 0, z_0)} \\ &= \begin{cases} \frac{\max[|z^*|, z_0]^{\alpha^*-\alpha} \log^{-\beta} \max[|z^*|, z_0]}{e^{\beta}\left(\frac{\beta}{\alpha^*-\alpha}\right)^{-\beta}}; & \text{if } \alpha + \frac{\beta}{\log z_0} < \min(1, \alpha), \\ \frac{\max[|z^*|, z_0]^{\alpha^*-\alpha} \log^{-\beta} \max[|z^*|, z_0]}{z_0^{\alpha^*-\alpha} \log^{-\beta} z_0}; & \text{otherwise ,} \end{cases} \end{aligned}$$

with $w(\alpha^*) = \sup_z \frac{p(z|\gamma, \delta, \alpha, \beta, z_0)}{p(z|\gamma, \delta, \alpha^*, 0, z_0)}$ as given in Proposition 2.

As expected, it can be shown that $0 \leq B(z^*) \leq 1$, $\forall z^* \in \mathbb{R}$. Since the average probability of accepting is $\frac{1}{w(\alpha^*)}$, choosing $\alpha^* = \arg\min_{\alpha_0} w(\alpha_0)$ in Proposition 2 also maximizes the probability of accepting the value in step 3.

We mentioned in Section 3.3 that the two cases with parameters set to $\gamma > 0, \delta > 0, \alpha \geq 1, \beta = 0$ and $\gamma = 0, \delta = 0, \beta \geq 1$ are considered in this section. For the first case, we simulate observations using the rejection method with $p(z|\gamma, \delta, \alpha^*, 0, z_0)$ as the proposal density and using $\alpha^* = 1 - \epsilon$. For the second case, we must first generate z' from $p(z'|1, 1, \beta, 0, (\alpha - 1)\log z_0)$ using the rejection method with $p(z'|1, 1, \alpha^*, 0, (\alpha - 1)\log z_0)$ as the proposal density and using $\alpha^* = 1 - \epsilon$. Then we set $z = \exp\left(\frac{z'}{\alpha - 1}\right)$.

If we want to simulate observations from the GEP density defined on (z_0, ∞), that is $p_1(z|\gamma, \delta, \alpha, \beta, z_0)$, then the algorithm is the same, except that z^* is generated from $p_1(z|\gamma, \delta, \alpha^*, 0, z_0)$ in step 1. Note that α^* should be chosen according to Proposition 2 replacing $I_0(\gamma, \delta, \alpha_0, 0, z_0)$ by $H_0(\gamma, \delta, \alpha_0, 0, z_0)$.

5 Examples of simulations of observations

5.1 First example

The density of interest in this example is the GEP density defined on $\mathbb{R}$ with parameters set to $\gamma = 1, \delta = 1, \alpha = 1, \beta = -10, z_0 = 5.122$. (We chose z_0 to be the smallest number satisfying condition C2. Furthermore we chose $\beta < 0$ and $\alpha = 1$ in order to have $\alpha^* \in \left[\alpha + \frac{\beta}{\log(z_0)}, \min(1, \alpha)\right)$ in Proposition 2.)

Since $\gamma > 0, \delta > 0, \beta \neq 0$, a proposal distribution must be chosen in order to evaluate the normalizing constant and the moments using Monte Carlo simulations or to simulate observations using the rejection method. The proposal density is $p(z|1, 1, \alpha^*, 0, 5.122)$, where α^* is given by Proposition 2. We see in Figure 1 that the function $w(\alpha_0)$ is strictly convex and minimized at $\alpha_0 = -4.687$, hence we choose $\alpha^* = -4.687$.

——————— Figure 1 should appear here ———————

Using Monte Carlo simulations with the proposal density as the importance function, we estimate the normalizing constant to 0.401 and the standard deviation of Z to 4.920 after 1000 simulations. The density of interest $p(z|1, 1, 1, -10, 5.122)$ and the proposal density $p(z|1, 1, -4.687, 0, 5.122)$ are shown in Figure 2. Both distributions are very close to each other, the proposal (dash line) having heavier tails than the density of interest (solid line).

——————— Figure 2 should appear here ———————

We simulated 1000 observations from the density of interest $p(1, 1, 1, -10, 5.122)$ using the rejection method. We needed to simulate 1060 observations from the proposal in order to accept 1000 observations. The average probability of accepting z^* in the rejection method is $\frac{1}{w(\alpha^*)} = 94.2\%$, where $w(\alpha^*)$ is given in Proposition 2.

A graph of the function $B(z^*)$ is shown in Figure 3. For $0 \leq |z^*| \leq 5.122$, $B(z^*)$ is equal to 0.974. Then, the probability $B(z^*)$ increases to 1 and then decreases to 0 as $|z^*|$ goes to infinity.

——————— Figure 3 should appear here ———————

5.2 Second example

The density of interest in the second example is the GEP density with parameters set to $\gamma = 1, \delta = 1, \alpha = 2, \beta = 2, z_0 = 1.5$. (Any choice of $z_0 > 1$ gives a unimodal density so we set $z_0 = 1.5$ arbitrarily. We chose $\beta > 0$ and $\alpha > 1$ in order to have $\alpha^* = 1 - \epsilon$ in Proposition 2.) The proposal density is $p(z|1, 1, \alpha^*, 0, 1.5)$, where $\alpha^* = 1 - \epsilon$, as specified in Proposition 2. We set $\epsilon = 0.01$.

Using Monte Carlo simulations with the proposal as the importance function, we estimate the normalizing constant to 0.479 and the standard deviation of Z to 1.047 after 1000 simulations. The density of interest $p(z|1, 1, 2, 2, 1.5)$ and the proposal density $p(z|1, 1, 0.99, 0, 1.5)$ are shown in Figure 4. We can see that the proposal distribution (dash line) has heavier tails than the density of interest (solid line).

——————— Figure 4 should appear here ———————

We simulated 1000 observations from the density of interest $p(1, 1, 2, 2, 1.5)$ using the rejection method. We needed to simulate 1258 observations from the proposal in order to accept 1000 observations. The average probability of accepting z^* in the rejection method is $\frac{1}{w(\alpha^*)} = 79.6\%$, where $w(\alpha^*)$ is given in Proposition 2.

A graph of the function $B(z^*)$ is shown in Figure 5. For $0 \leq |z^*| \leq 1.5$, $B(z^*)$ is equal to 1. Then, the probability $B(z^*)$ decreases to 0 as $|z^*|$ goes to infinity.

——————— Figure 5 should appear here ———————

6 Applications of the GEP density and p-credence

In this section, we consider Bayesian inference for location parameter when the prior and the likelihood are symmetric densities defined on the real line with their p-credence defined.

In Section 6.1, we evaluate the posterior moments using Monte Carlo simulations with importance sampling. We show how the GEP density and p-credence can contribute to determine an effective importance function. In Section 6.2, we discuss a first example with one observation where p-credence of the prior is larger than p-credence of the likelihood. We see how the posterior behaves when a conflict occurs between the prior and the likelihood. In Section 6.3, we present a second example, again with one observation, but this time p-credence of the likelihood is the largest. Finally in Section 6.4, we consider an example with several observations. We discuss, using p-credence, how the posterior behaves when one or several observations are outliers.

6.1 Estimation of posterior moments using Monte Carlo simulations with importance sampling

In this section, we want to evaluate the posterior moments using Monte Carlo simulations with importance sampling, when the prior and the likelihood are symmetric densities defined on the real line with their p-credence defined. The selection of an appropriate importance function is the main issue of this section and is addressed with the GEP density and the notion of p-credence.

Consider $\theta \sim \pi(\theta)$ and $X|\theta \sim f(x-\theta)$, two symmetric densities defined on $\mathbb{R}$ with p-cred$(\theta) = (\gamma, \delta, \alpha, \beta)$ and p-cred$(X-\theta|\theta) = (\gamma', \delta', \alpha', \beta')$. We want to evaluate $\mathbb{E}(\theta^j|\mathrm{x})$ if we observe $\mathrm{x} = (x_1, ..., x_n)$.

The algorithm of Monte Carlo consists in generating $\theta_1, ..., \theta_m$ from an importance function $g(\theta)$ and to estimate the j^{th} posterior moment by

$$\widehat{\mathbb{E}}(\theta^j|\mathrm{x}) = \frac{\sum_{k=1}^m \theta_k^j v(\theta_k)}{\sum_{k=1}^m v(\theta_k)},$$

where the weight function $v(\theta_k)$ is given by

$$v(\theta_k) \propto \frac{\pi(\theta_k)\prod_{i=1}^n f(x_i - \theta_k)}{g(\theta_k)}.$$

Note that the weight function can be multiplied by any constant since the constant vanish in the calculation of $\widehat{\mathbb{E}}(\theta^j|\mathrm{x})$.

The posterior density can be multimodal, with possible modes around the prior location and around each observation. It is thus difficult to choose an importance function $g(\theta)$ close to the posterior. We propose to choose the importance function as $g(\theta) = p(\theta - \mu^*|\gamma^*, \delta^*, \alpha^*, \beta^*, z_0^*)$, a GEP density defined on $\mathbb{R}$ by equation (1). The parameters μ^* and z_0^* will be chosen to ensure that the uniform part of the GEP density is covering most of the prior and the likelihood (expressed in function of θ). Furthermore, the parameters $\gamma^*, \delta^*, \alpha^*$ and β^* will be chosen to ensure that the importance function is dominating the posterior density or equivalently that the tails of $g(\theta)$ are heavier than those of $\pi(\theta|\mathrm{x})$.

The tails of $g(\theta)$ are set with the help of p-credence. If there is only one observation, that is $n = 1$, we know (see Angers, 2000) that, $\forall z_0$ satisfying condition C1,

$$\pi(\theta|x) \preceq p(\theta - x_{**}|(\max(\gamma', \gamma), \delta_{**}, \alpha' + \alpha, \beta' + \beta, z_0),$$

where

$$\delta_{**} = \begin{cases} \delta'; & \text{if } \gamma' > \gamma, \\ \delta; & \text{if } \gamma' < \gamma, \end{cases} \qquad x_{**} = \begin{cases} x; & \text{if } \gamma' > \gamma, \\ 0; & \text{if } \gamma' < \gamma. \end{cases}$$

We can generalize this result for $n > 1$ as follows:

$$\pi(\theta|\mathrm{x}) \preceq p(\theta - \mu|(\max(\gamma', \gamma), \delta_{***} - \tau\, \mathbb{I}_{[\max(\gamma',\gamma)>1]}, n\alpha' + \alpha, n\beta' + \beta, z_0),$$

where

$$\delta_{***} = \begin{cases} n\delta'; & \text{if } \gamma' > \gamma, \\ \delta; & \text{if } \gamma' < \gamma, \\ n\delta' + \delta; & \text{if } \gamma' = \gamma, \end{cases} \tag{6}$$

$\mu \in \mathbb{R}$, $\tau \in (0, \delta_{***})$, z_0 is satisfying condition C1 and $\mathbb{I}_{[\max(\gamma',\gamma)>1]}$ is the indicator function of the set $\{\max(\gamma', \gamma) > 1\}$.

We can now define the parameters of the importance function by Proposition 3. But let first define respectively q_p and q'_p as the p^{th} percentiles of $\pi(\theta)$ and $f(x - \theta)$, and let define $x_{(i)}$ as the i^{th} ordered observation.

Proposition 3. *The parameters of the importance function $g(\theta) = p(\theta - \mu^*|\gamma^*, \delta^*, \alpha^*, \beta^*, z_0^*)$ which ensure that the uniform part of the GEP density is covering most of the prior and the likelihood and that tails of $g(\theta)$ are heavier than those of $\pi(\theta|\mathrm{x})$ are given by:*

$\mu^* = \frac{m_1+m_2}{2}$, $z_0^* = \frac{m_2-m_1}{2}$, $\gamma^* = \max(\gamma', \gamma)$, $\delta^* = \delta_{***} - \tau\, \mathbb{I}_{[\gamma^*>1]}$,

$$\alpha^* = \begin{cases} n\alpha' + \alpha; & \textit{if } \gamma^* = 0, \delta^* = 0, \\ \min(n\alpha' + \alpha, n\alpha' + \alpha + \frac{n\beta'+\beta}{\log z_0^*}, 1 - \epsilon); & \textit{otherwise,} \end{cases}$$

$$\beta^* = \begin{cases} \min(n\beta' + \beta, 1 - \epsilon); & \textit{if } \gamma^* = 0, \delta^* = 0, \\ 0; & \textit{otherwise,} \end{cases}$$

where $m_1 = \min(-q_p, x_{(1)} - q'_p)$, $m_2 = \max(q_p, x_{(n)} + q'_p)$, δ_{***} is given by equation (6), $\tau \in (0, \delta_{***})$ and $0.5 < p < 1$.

Note that p and τ must be specified. In practice, $p = 0.99$ seems appropriate to cover a sufficient part of the prior and the likelihood. The choice of p being arbitrary, a gross approximation of the quantile is sufficient. Furthermore, $\tau = 0.01$ seems also appropriate. A larger value of τ would give an importance function with heavier tails, which is not necessary. We can simulate observations from $g(\theta)$ with the inverse transformation method seen in Section 3.3.

6.2 First example

Consider $\theta \sim \pi(\theta)$ and $x|\theta \sim f(x - \theta)$. In this example, the densities are $\pi(\theta) \propto \max[|\theta|, 0.1]^{-10}$, $\theta \in \mathbb{R}$ and $f(x - \theta) \propto \max[|x - \theta|, 0.1]^{-5}$, $x \in \mathbb{R}$. These are GEP densities with p-credences of $(0, 0, 10, 0)$ and $(0, 0, 5, 0)$ respectively. Using Proposition 1, we see that the prior has the largest p-credence or that the likelihood has heavier tails. Standard deviations of π and f are respectively 0.065 and 0.082.

Figure 6 shows how the posterior behaves when the observation x goes away from the prior mean fixed to 0. Results are shown in Table 2. Figures 6a shows the distributions when $x = 0$, that is when both sources of information agree perfectly. Not surprisingly, the posterior is centered at 0. (Note that the likelihood is expressed as a function of θ centered in x). Figure 6b shows the distributions when $x = 0.22$. Both sources of information are compatible and the posterior averages these informations. Figure 6c and 6d show the distributions when $x = 0.5$ and $x = 2$. We can see that most of the information from the likelihood has been discarded, reflecting a conflict between the prior and the likelihood. Since p-credence of the prior is the largest, the posterior mean converges to the prior mean as x goes to infinity. ——————— Figure 6 should appear here ———————

——————— Table 2 should appear here ———————

The observation $x = 0.22$ in Figure 6b corresponds to a change point. We see in Figure 7 the posterior mean (solid line) for values of x ranging from 0 to 2. As x increases from 0 to 0.22, the posterior mean also increases, reflecting the compromise between both sources of information. However, if x increases above 0.22, the posterior mean decreases, reflecting that a portion of information from the likelihood is discarded, this portion is increasing as x increases.

——————— Figure 7 should appear here ———————

In Figure 7, the posterior standard deviation is also shown as the dotted line for values of x ranging from 0 to 2. We see that the posterior standard deviation decreases as x increases from 0 to 0.18, then increases to the smallest standard deviation between the prior and the likelihood as x goes to infinity. As the difference between both sources of information increases, the intersection of their densities decreases, which reduces the variability in the posterior. However, if the difference exceeds the threshold of 0.18, this reduction of variability is offset by the increase in the posterior variability resulting from the discarded likelihood information. The posterior standard deviation is always smaller than the standard deviation of both the prior and the likelihood.

6.3 Second example

For the second example, the densities are $\pi(\theta) \propto \max[|\theta|, 0.1]^{-5}$, $\theta \in \mathbb{R}$ and $f(x-\theta) \propto \max[|x-\theta|, 0.1]^{-10}$, $x \in \mathbb{R}$. This is simply the first example when the densities of the prior and the likelihood are interchanged. Here the likelihood has the largest p-credence or equivalently the prior has heavier tails.

We can show, when the prior and the likelihood are interchanged for a location parameter inference with one observation, that the posterior mean of one case is equal to x minus the posterior mean of the other case. Furthermore, all the posterior centered moments of even order are the same. All posterior centered moments of odd order for one case are the negative of the other. The second example is then symmetric to the first one, with the posterior converging to the likelihood as x goes to infinity. Results for different values of x are shown in Table 3. The difference between x and the posterior mean for different values of x is also represented by the solid line in Figure 7. As x increases from 0 to 0.22, the posterior mean also increases in a way reflecting the compromise between both sources of information. However, if x increases above 0.22, $x - \mathbb{E}(\theta|x)$ decreases, reflecting that a portion of information from the prior is discarded. As x goes to infinity, the posterior mean goes to x, that is the mean of the likelihood information.

——————— Table 3 should appear here ———————

6.4 Third example

For the last example, we consider four observations, independent and identically distributed from the density $f(x-\theta) \propto \max[|x-\theta|, 0.1]^{-5}$. The prior density of θ is given by $\pi(\theta) \propto \max[|\theta|, 0.1]^{-10}$, $\theta \in \mathbb{R}$. These are GEP densities with p-credence of $(0, 0, 10, 0)$ for the prior and $(0, 0, 5, 0)$ for

$f(x-\theta)$. According to Angers (2000), p-credence of the likelihood is then $(0,0,20,0)$. Though p-credence of the density for one observation is smaller than p-credence of the prior, p-credence of the likelihood for four observations is larger than p-credence of the prior.

We see in Figure 8a the prior, the likelihood (expressed as a function of θ) and the posterior when $x_1 = 0.15$, $x_2 = 0.20$, $x_3 = 0.25$ and $x_4 = 0.30$. There is no conflict between the prior and the likelihood, so the posterior compromises between them. In Figure 8b we use the same setting as in Figure 8a but 0.25 is added to all observations. There is an apparent conflict between the prior and the likelihood, so the posterior converges to the source of information with the largest p-credence, that is the likelihood.

——————— Figure 8 should appear here ———————

In Figure 9, we see the posterior densities when $x_1 = 0.40$, $x_2 = 0.45$ and $x_3 = 0.50$ and for different values of x_4. The dashed line is the posterior when $x_4 = 0.55$, as in Figure 8b. The dotted line is the posterior when $x_4 = 1.50$ and the solid line is the posterior without x_4. We want to illustrate that as x_4 goes to infinity, the posterior converges to the posterior without the outlier x_4. Note that the posterior also has a mode around the location of the prior. Furthermore, when $x_4 = 1.5$, there is another mode around 1.5. However these modes have a negligible mass compared to the rest of the posterior. Note that the posterior also discard most of the information from the prior. Actually, there is three groups of conflicting information: the prior with p-credence of (0,0,5,0), the likelihood without x_4 with p-credence (see Angers, 2000) of (0,0,15,0) and the outlier x_4 with p-credence of (0,0,5,0). The posterior converges to the group with the largest p-credence, that is the likelihood without x_4.

——————— Figure 9 should appear here ———————

Finally we can see in Figure 10 the posterior (dashed line) and the likelihood (dotted line) means for different values of x_4, when $x_1 = 0.40$, $x_2 = 0.45$ and $x_3 = 0.50$. As x_4 increases from 0.55 to 0.61, the posterior mean also increases. After this change point, it decreases and converges to 0.397, that is the posterior mean without the outlier x_4, reflecting the fact that an increasing portion of x_4 is discarded as x_4 goes to infinity. The same pattern apply to the likelihood mean, except the change point is around 0.65 and the likelihood means converge to 0.450 as x_4 goes to infinity, that is the likelihood mean without the outlier x_4. The likelihood begins to discard information later than the posterior since the prior, centered in 0, is considered in the posterior.

——————— Figure 10 should appear here ———————

7 Conclusion

In this paper, we have considered some computational aspects of the generalized exponential power density. In particular, some methods for the evaluation of the normalizing constant, the moments and the cumulative distribution function have been proposed. For some special cases, analytic evaluation can be computed, but in general numerical methods are required. We proposed two methods, depending on the sets of parameters, to simulate observations from the GEP distribution. We also proposed a numerical method for the estimation of posterior moments when the prior and the likelihood are symmetric densities defined on the real line. Finally, we presented some examples of the behavior of the posterior density when an observation is an outlier.

8 Acknowledgments

The authors are thankful to the NSERC (Natural Sciences and Engineering Research Council of Canada), the FCAR (Fonds pour la Formation de chercheurs et l'aide à la recherche) and the SOA (Society of Actuaries) for their financial support.

9 References

ANGERS, J.-F. (1992) *Use of Student-t prior for the estimation of normal means: A computational approach, in Bayesian Statistic IV*, eds. Bernardo, J.M., Berger, J.O., David, A.P., and Smith, A.F.M., New York: Oxford University Press, pp. 567-575.

ANGERS, J.-F. (1996) *Protection against outliers using a symmetric stable law prior, in IMS Lecture Notes - Monograph Series*, Vol. **29**, pp. 273-283

ANGERS, J.-F. (2000), P-Credence and outliers, *Metron* **58**, 81-108.

ANGERS, J.-F. and BERGER, J.O. (1991) Robust hierarchical Bayes estimation of exchangeable means, *The Canadian Journal of Statistics*, **19**, 39-56.

BOX, G. and TIAO, G. (1962) A further look at robustness via Bayes's theorem, *Biometrika*, **49**, 419-432.

CARLIN, B. and POLSON, N. (1991) Inference for nonconjugate Bayesian models using the Gibbs sampler, *The Canadian Journal of Statistics*, **19**, 399-405.

FAN, T.H. and BERGER, J.O. (1992) Behaviour of the posterior distribution and inferences for a normal means with t prior distributions, *Statistics & Decisions*, **10**, 99-120.

GEWEKE, J. (1994) Priors for macroeconomic time series and their applications, *Econometric Theory*, **10**, 609-632.

JOHNSON, N.L., KOTZ, S. AND BALAKRISHNAN, N. (1994), *Continuous univariate distributions*, Vol. 1 (second edition), Wiley, New York.

KOTZ, S. and NADARAJAH, S. (2000) *Extreme value distributions: theory and applications*, Imperial College Press, London.

LANGE, K. (1999) *Numerical analysis for statisticians*, Springer-Verlag, New York.

MEINHOLD, R. and SINGPURWALLA, N. (1989) Robustification of Kalman filter models, *Journal of the American Statistical Association*, **84**, 479-486.

O'HAGAN, A. (1990) Outliers and credence for location parameter inference, *Journal of the American Statistical Association*, **85**, 172-176.

REISS, R.-D. and THOMAS, M. (1997) *Statistical analysis of extreme values*, Birkhauser Verlag, Basel.

ROBERT, C. (1996) *Méthodes de Monte Carlo par chaînes de Markov*, Economica, Paris.

ROSS, S.M. (1997) *Simulation* (second edition), Academic Press, San Diego

Table 1: Special cases of the GEP distribution

Density	$\gamma, \delta, \alpha, \beta, z_0$	**Parameters and support**	**Normalizing constant**	$\mathbb{E}(\lvert Z\rvert^j)$ $(j > 0)$
exponential power	$\frac{2}{1+\beta}, \frac{1}{2}, 0, 0, 0$	$\beta > -1$, $z \in \mathbb{R}$	$\frac{2^{-\left(\frac{1+\beta}{2}\right)}}{(1+\beta)\Gamma\left(\frac{1+\beta}{2}\right)}$	$\frac{\Gamma\left(\frac{(1+j)(1+\beta)}{2}\right)2^{\frac{j(1+\beta)}{2}}}{\Gamma\left(\frac{1+\beta}{2}\right)}$
normal	$2, \frac{1}{2\sigma^2}, 0, 0, 0$	$\sigma > 0$, $z \in \mathbb{R}$	$\frac{1}{\sigma\sqrt{2\pi}}$	$\frac{\left(2\sigma^2\right)^{j/2}\Gamma\left(\frac{j+1}{2}\right)}{\sqrt{\pi}}$
Laplace	$1, \delta, 0, 0, 0$	$\delta > 0$, $z \in \mathbb{R}$	$\frac{\delta}{2}$	$\frac{\Gamma(1+j)}{\delta^j}$
generalized gamma	$\gamma, \delta, \alpha, 0, 0$	$\gamma > 0, \delta > 0$, $\alpha < 1, z > 0$	$\frac{\gamma\delta^{\frac{1-\alpha}{\gamma}}}{\Gamma(\frac{1-\alpha}{\gamma})}$	$\frac{\Gamma(\frac{1-\alpha+j}{\gamma})}{\Gamma(\frac{1-\alpha}{\gamma})\delta^{\frac{j}{\gamma}}}$
Weibull	$\gamma, \delta, 1-\gamma, 0, 0$	$\gamma > 0, \delta > 0$, $z > 0$	$\gamma\delta$	$\frac{\Gamma(1+\frac{j}{\gamma})}{\delta^{\frac{j}{\gamma}}}$
Rayleigh	$2, \frac{1}{2\sigma^2}, -1, 0, 0$	$\sigma > 0$, $z > 0$	$\frac{1}{\sigma^2}$	$\left(2\sigma^2\right)^{j/2}\Gamma\left(1+\frac{j}{2}\right)$
Maxwell-Boltzmann	$2, \frac{1}{2\sigma^2}, -2, 0, 0$	$\sigma > 0$, $z > 0$	$\frac{1}{\sigma^3}\sqrt{\frac{2}{\pi}}$	$\frac{2}{\sqrt{\pi}}\left(2\sigma^2\right)^{j/2}\Gamma\left(\frac{j+3}{2}\right)$
gamma	$1, \delta, 1-\tau, 0, 0$	$\delta > 0, \tau > 0$, $z > 0$	$\frac{\delta^\tau}{\Gamma(\tau)}$	$\frac{\Gamma(\tau+j)}{\Gamma(\tau)\delta^j}$
Pareto	$0, 0, \alpha, 0, z_0$	$\alpha > 1, z_0 > 0$, $z > z_0$	$(\alpha-1)z_0^{\alpha-1}$	$\frac{z_0^j(\alpha-1)}{(\alpha-j-1)}$ $(0 < j < \alpha-1)$
log-gamma	$0, 0, \alpha, \beta, 1$	$\alpha > 1, \beta < 1$, $z > 1$	$\frac{(\alpha-1)^{1-\beta}}{\Gamma(1-\beta)}$	$\left(\frac{\alpha-1}{\alpha-j-1}\right)^{1-\beta}$ $(0 < j < \alpha-1)$
log-Pareto	$0, 0, 1, \beta, z_0$	$\beta > 1, z_0 > 1$, $z > z_0$	$\frac{\beta-1}{(\log z_0)^{1-\beta}}$	∞

Table 2: Results of the first example

x	posterior mean	posterior standard deviation
0	0	0.062
0.22	0.085	0.039
0.50	0.043	0.061
2.00	0.011	0.065

Table 3: Results of the second example

x	posterior mean	posterior standard deviation
0	0	0.062
0.22	0.133	0.039
0.50	0.458	0.061
2.00	1.989	0.065

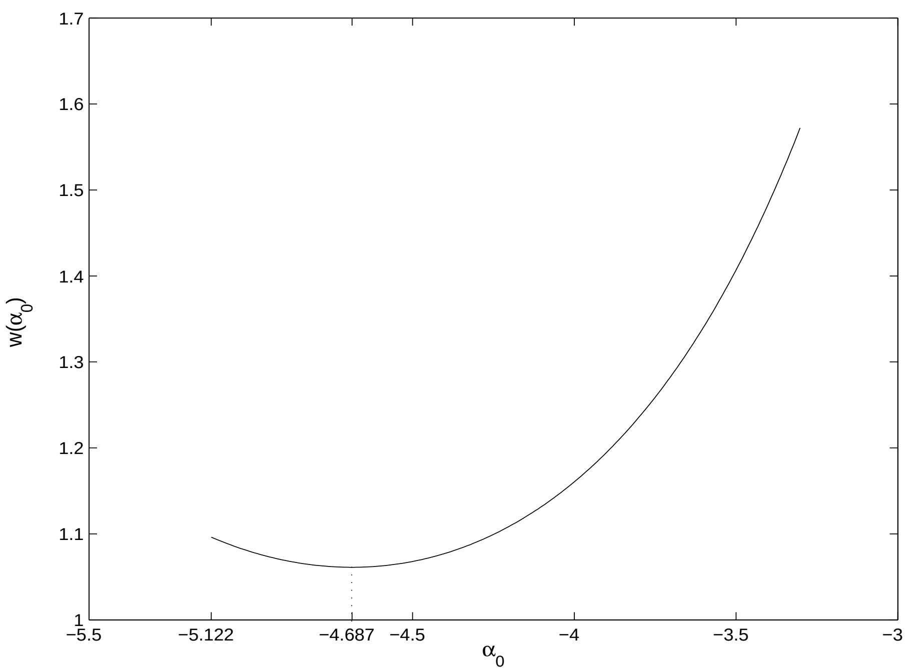

Figure 1: Choice of α^*: the parameter α_0 which minimizes $w(\alpha_0)$ (example 5.1)

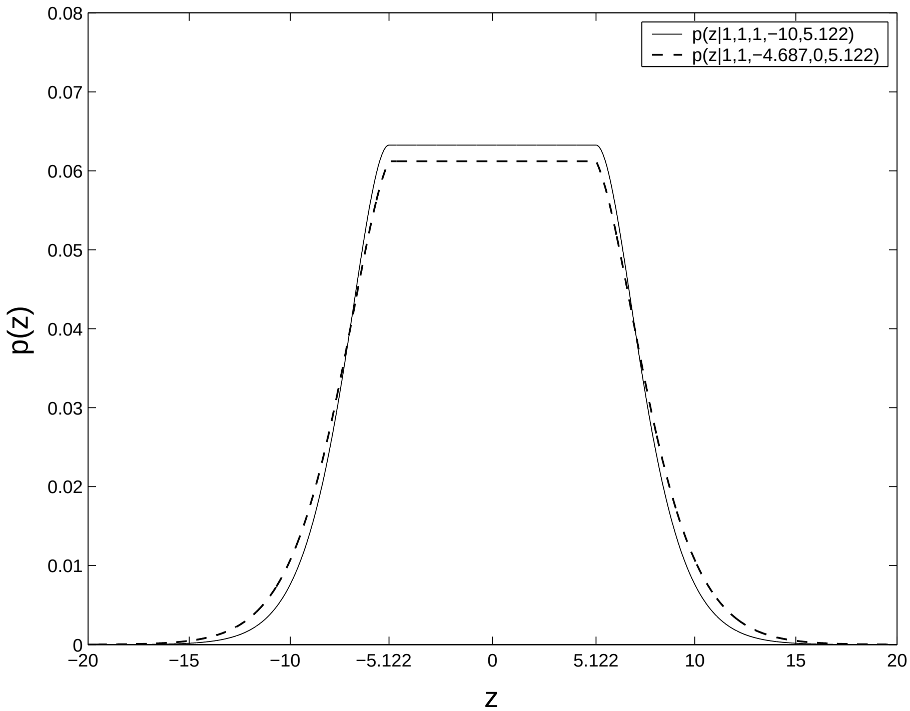

Figure 2: The density of interest and the proposal (example 5.1)

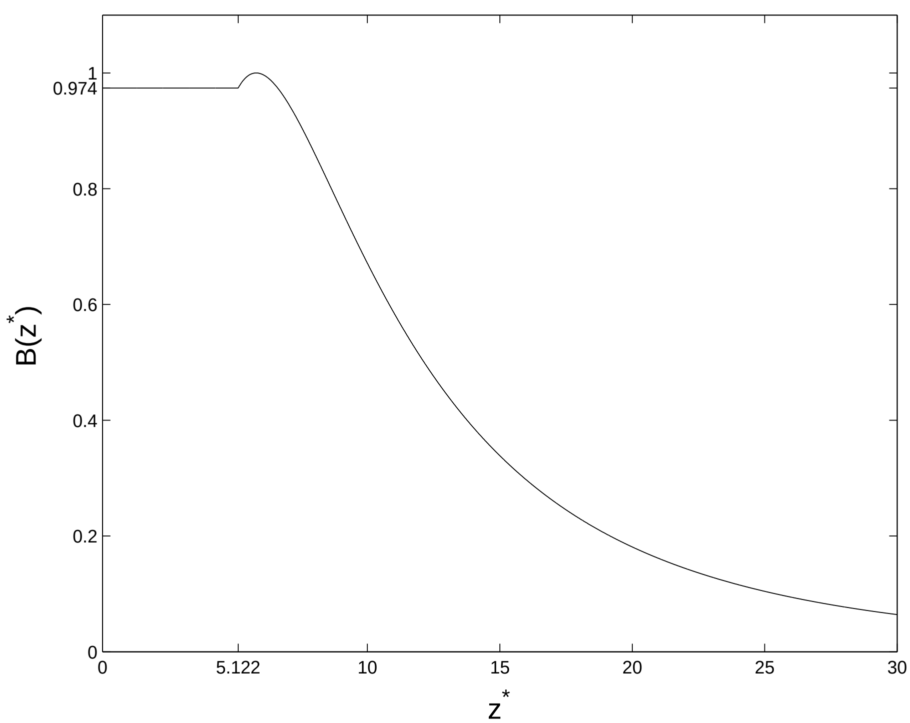

Figure 3: The function $B(z^*)$ in the rejection method (example 5.1)

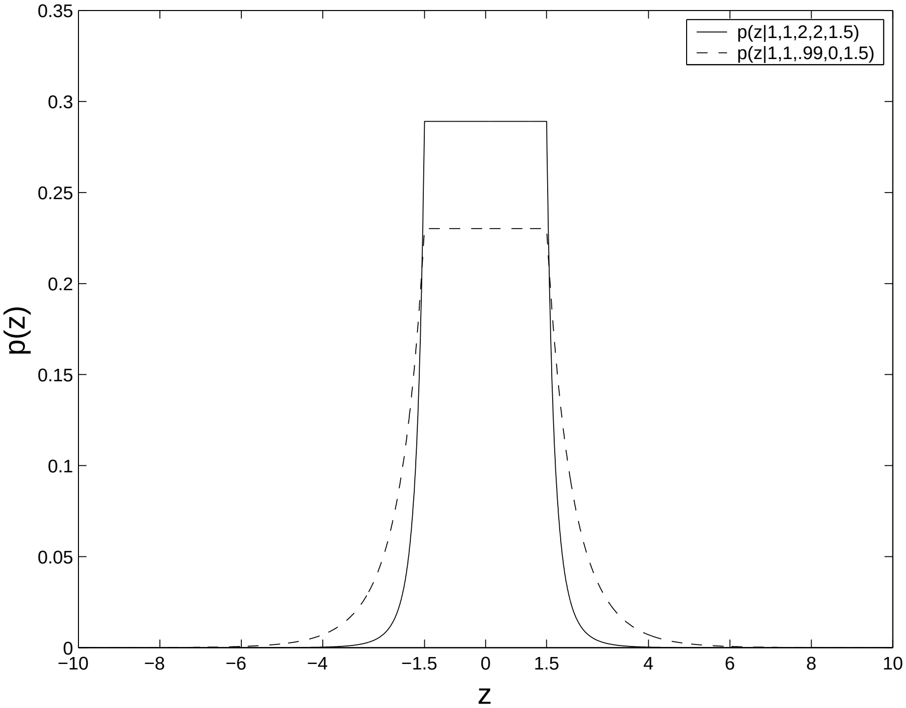

Figure 4: The density of interest and the proposal (example 5.2)

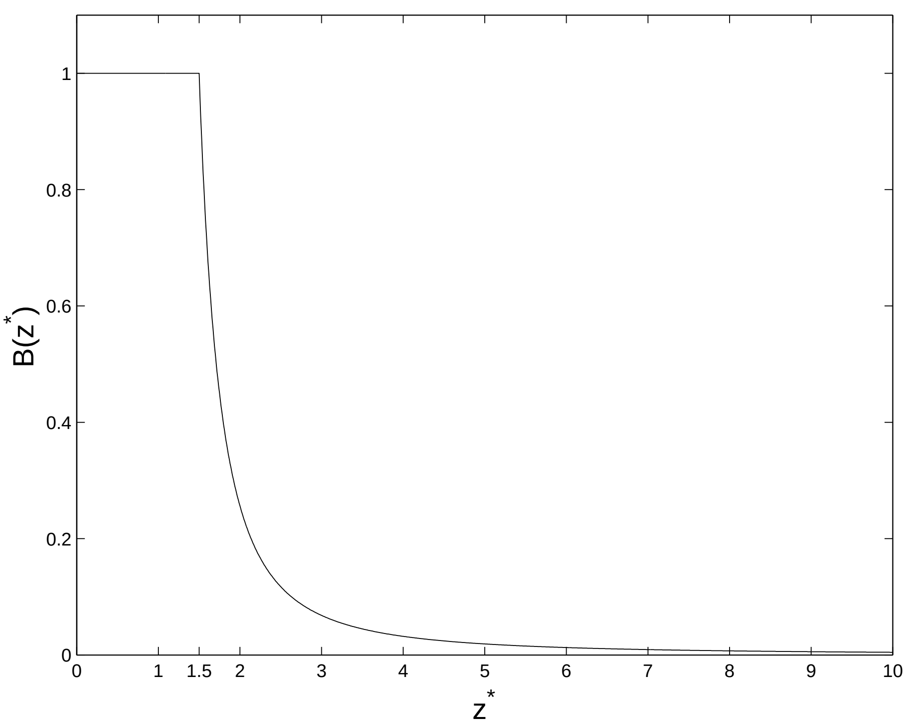

Figure 5: The function $B(z^*)$ in the rejection method (example 5.2)

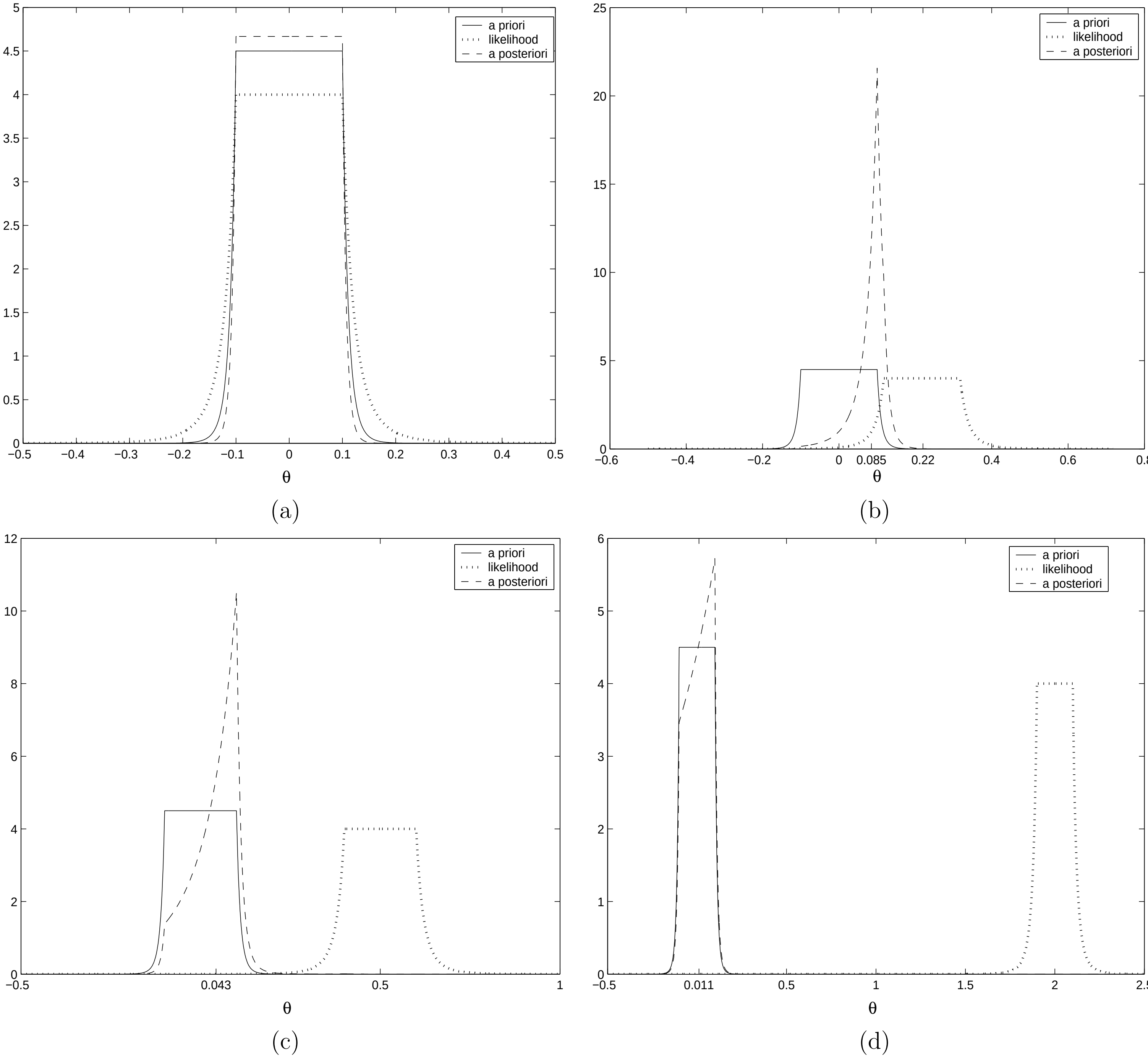

Figure 6: The densities when (a) $x = 0$, (b) $x = 0.22$ (c) $x = 0.5$ (d) $x = 2$ (example 6.2)

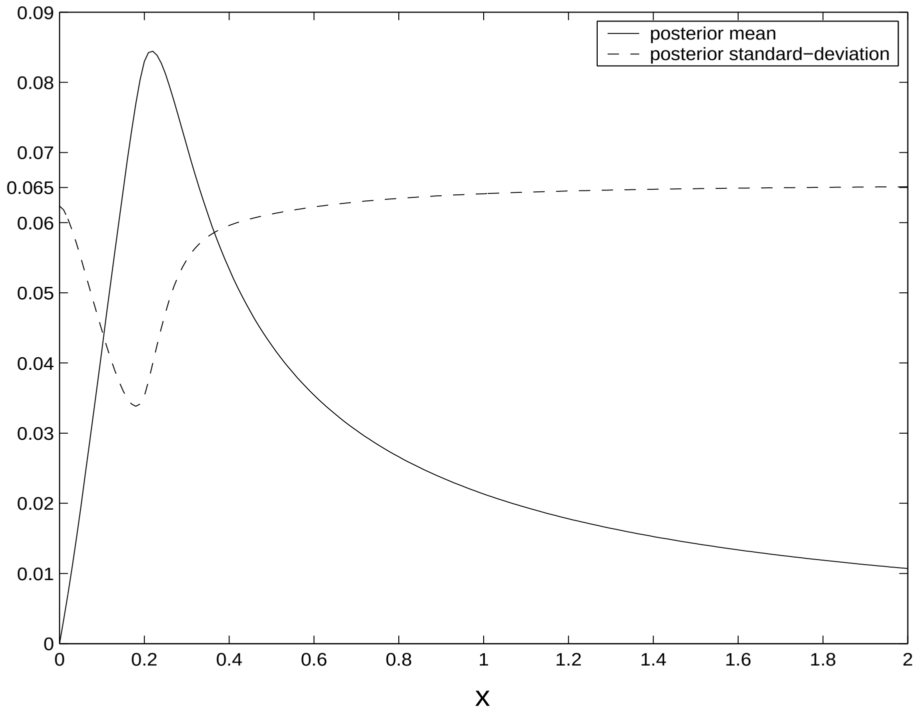

Figure 7: Posterior mean and standard deviation for different values of x (example 6.2)

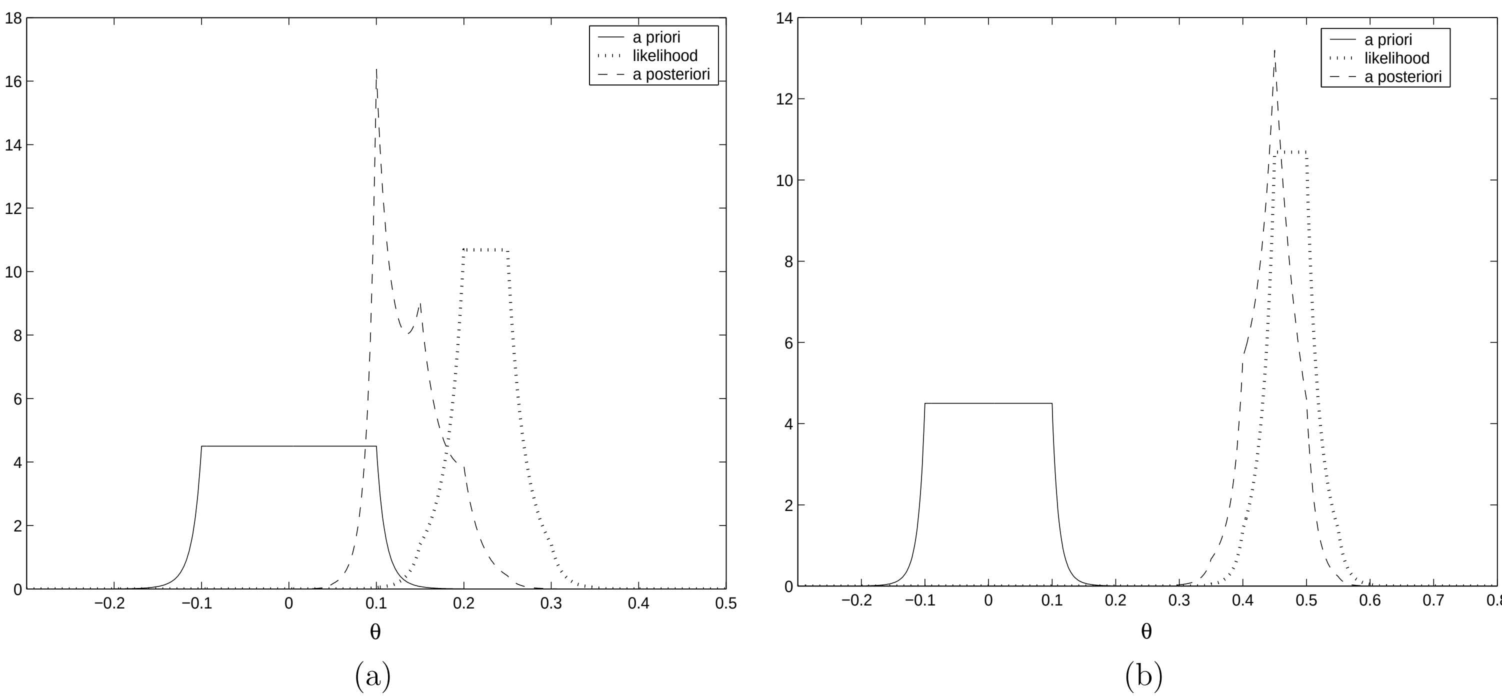

Figure 8: The densities when (a) $x_1 = 0.15$, $x_2 = 0.20$, $x_3 = 0.25$, $x_4 = 0.30$ (b) $x_1 = 0.40$, $x_2 = 0.45$, $x_3 = 0.50$, $x_4 = 0.55$ (example 6.4)

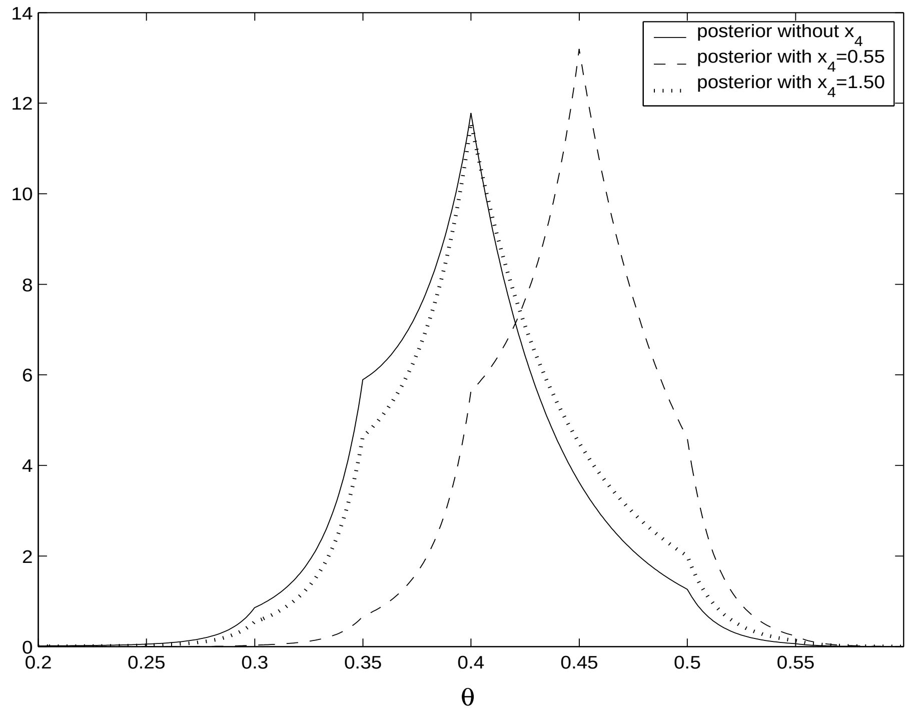

Figure 9: The posterior densities when $x_1 = 0.40$, $x_2 = 0.45$, $x_3 = 0.50$ and for different values of x_4 (example 6.4)

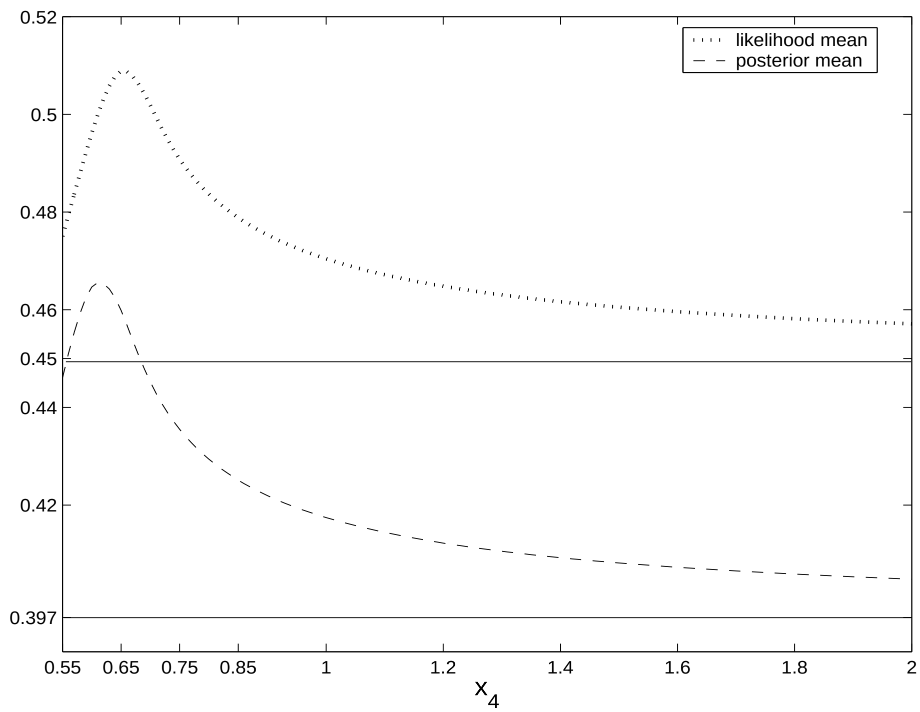

Figure 10: The posterior and likelihood means when $x_1 = 0.40$, $x_2 = 0.45$, $x_3 = 0.50$ and for different values of x_4 (example 6.4)