METRON - *International Journal of Statistics*
2007, vol. LXV, n. 1, pp. 67-97

ALAIN DESGAGNÉ – JEAN-FRANÇOIS ANGERS

Conflicting information and location parameter inference

Summary - The use of heavy-tailed distributions is a valuable tool in developing robust Bayesian procedures, limiting the influence of conflicting information on posterior inference. In this paper, the behavior of the posterior density of a location parameter is investigated when the sample contains outliers or the prior location is misspecified. Conditions on the tails of the prior and the likelihood are established to determine the proportion of conflicting information that can be rejected by the posterior. It is shown that the posterior distribution converges in law to a density proportional to the product of the densities of the non-conflicting information, as the outliers (and/or the prior location) go to plus or minus infinity, at any given rate. In particular, if the prior is non-conflicting, this limiting density is the posterior that would be obtained from the reduced sample, excluding the outliers. Examples are given to illustrate the results.

Key Words - Bayesian inference; Conflicting information; Outlier; Heavy-tailed modeling; Location parameter; p-credence.

1. Introduction

The use of heavy-tailed distributions is a valuable tool in developing robust Bayesian procedures, limiting the influence of conflicting information on posterior inference. Outlier rejection in Bayesian analysis was first described by De Finetti (1961), where the simplest case with a single observation having mean θ was considered. Theoretical results were given by Dawid (1973) and Hill (1974). O'Hagan (1979) considered outlier rejection in a sample and O'Hagan (1988) considered more general Bayesian modeling based on Student-t distributions. Outlier rejection based on the notion of credence was introduced by O'Hagan (1990) and was generalized to p-credence by Angers (2000) to accommodate a wider class of densities. Other authors approached outlier rejection, see for instance Meinhold and Singpurwalla (1989), Angers and Berger (1991), Carlin and Polson (1991), Angers (1992), Fan and Berger (1992), Geweke (1994) and Angers (1996).

Received February 2005 and revised April 2007.

In Section 2, the behavior of the posterior density of the location parameter is investigated when the sample contains outliers or the location parameter of the prior is misspecified, that is when there is conflicting information. In Section 2.1, the Bayesian context is described. In Section 2.2, conditions of thickness and regularity for the tails of a density, needed when robust inference is expected, are given. In Section 2.3, the main results of this paper are given in Theorem 1. Under certain conditions on the tails of the prior density and the likelihood, asymptotic results are given. Note that the term "asymptotic" is not used in the frequentist sense, where the sample size n tends to infinity, but in the sense that the conflicting values (outliers and/or the prior location) tend to plus or minus infinity. We determine the asymptotic behavior of the marginal density of the observations, the posterior density of the location parameter and the posterior expectation of some functions of the location parameter. In Section 2.4, we present two special cases. The case where the sample contains only one observation is described, which enables us to make the connection with previous papers on robustness. The case where the tail behavior of the prior density and the likelihood are identical is also described for its usefulness in practice. In Section 2.5, conditions are simplified using a measure of tails called left and right p-credences. In Section 3, examples are given to illustrate the results. We conclude in Section 4 and proofs are given in Section 5.

2. Conflicting information

2.1. Context

Consider the following Bayesian context.

i) Let $X_1, \dots, X_n$ be n random variables conditionally independent given θ with the conditional densities of $X_i|\theta$, evaluated at the point x_i, given by $f_i(x_i - \theta)$, where $x_i \in \mathbb{R}, \theta \in \mathbb{R}$, $i = 1, \dots, n$.
ii) The prior density of θ, evaluated at the point θ, is $f_0(x_0 - \theta)$, where $x_0 \in \mathbb{R}$ is a known location parameter.

To emphasize the fact that the prior is an additional source of information, we use the notation $f_0(x_0 - \theta)$ instead of the more usual $\pi_\theta(\theta - x_0)$. Note that the scale parameter is assumed to be known. Extending the results to the location-scale parameters inference is not trivial. The densities $f_0, f_1, \dots, f_n$ are assumed to be proper, positive everywhere and bounded above.

The observations are denoted by $\underset{\sim}{x}_n = (x_1, \dots, x_n)$. Among the $n + 1$ sources of information $(x_0, x_1, \dots, x_n)$, we assume that $k + 1$ of them are fixed and denoted by $(x_{j_0}, x_{j_1}, \dots, x_{j_k})$, $m - k$ of them tend to $-\infty$ and are denoted by $(x_{j_{k+1}}, \dots, x_{j_m})$ and finally $n - m$ of them tend to ∞ and are denoted by $(x_{j_{m+1}}, \ldots, x_{j_n})$, where $(j_0, j_1, \ldots, j_n)$ is any permutation of the set $(0, 1, \ldots, n)$ and $0 \le k \le m \le n$. Note that the location parameter of the prior is treated just as another observation and can be considered as conflicting information if $x_0 \in (x_{j_{k+1}}, \ldots, x_{j_n})$. If x_0 is fixed, we set $x_{j_0} = x_0$ without loss of generality.

Let the posterior density of θ be denoted by $\pi(\theta | \underset{\sim}{x}_n)$ and the marginal density of $X_1, \ldots, X_n$ be denoted by $m(\underset{\sim}{x}_n)$. If we denote $\underset{\sim}{x}_k = (x_{j_1}, \ldots, x_{j_k})$, two other densities are defined as follows:

$$m(\underset{\sim}{x}_k) = \int_{-\infty}^{\infty} \prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)\, d\theta \quad \text{and} \quad \pi(\theta | \underset{\sim}{x}_k) = \frac{\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)}{m(\underset{\sim}{x}_k)}.$$

Note that when the location parameter of the prior is fixed, that is $x_{j_0} = x_0$, $m(\underset{\sim}{x}_k)$ is the marginal density of $X_{j_1}, \ldots, X_{j_k}$ and $\pi(\theta | \underset{\sim}{x}_k)$ is the posterior density of θ, considering only the k fixed observations $\underset{\sim}{x}_k$. If $k = 0$, $m(\underset{\sim}{x}_k)$ is set to 1 and $\pi(\theta | \underset{\sim}{x}_k)$ becomes $f_{j_0}(x_{j_0} - \theta)$. Finally, let the vector of conflicting values be denoted by $\underset{\sim}{\phi} = (-x_{j_{k+1}}, \ldots, -x_{j_m}, x_{j_{m+1}}, \ldots, x_{j_n})$. The notation $\underset{\sim}{\phi} \to \infty$ means that each term of the vector tends to ∞ at any given rate.

In the next section, conditions on the prior density and the likelihood are established to obtain robust Bayesian inference on the location parameter. The influence of the conflicting information on the posterior density is expected to decrease as the conflict increases.

2.2. Conditions of thickness and regularity for the tails of a density

The tails of the densities of the conflicting information must satisfy certain conditions of thickness and regularity when robust inference is expected. Three conditions of thickness and regularity for the tails of a density f are given by conditions C1 to C3 as follows. The density f is assumed to be proper, positive everywhere and bounded above. (Note that the conditions are the same for the left and right tails, except for the support of the density which is given in parentheses for the left tail.)

C1: $\forall \epsilon > 0, \forall h > 0$, there exists a constant $A_1(\epsilon, h) > 0$ such that $z > A_1(\epsilon, h)$ ($z < -A_1(\epsilon, h)$ for the left tail) and $|\theta| \le h \Rightarrow 1 - \epsilon \le \frac{f(z+\theta)}{f(z)} \le 1 + \epsilon$.

For conditions C2 and C3, there exist constants $A_2 > 0$ and $M_2 > 1$ and proper densities f^* and g such that for all $z > A_2$ ($z < -A_2$ for the left tail), C2: $\frac{f^2(z/2)}{f(z)g(z/2)} \leq M_2$,

C3: $\frac{d^2}{dz^2} \log f^*(z) \geq \frac{d^2}{dz^2} \log g(z) \geq 0$,

where f^* can be the density f or a density with the same tail behavior. More formally, the density f^* must satisfy this condition: there exist constants $B > 0$ and $0 < K_1 < K_2 < \infty$ such that $z > B$ ($z < -B$ for the left tail) $\Rightarrow K_1 \leq \frac{f(z)}{f^*(z)} \leq K_2$.

In condition C1, the ratio of the density f measured in two points with any fixed distance approaches 1 when the two points increase in the right tail, which means that a location transformation has no impact on the right tail of the density $f(z)$, as $z \to \infty$. This ensures that the tail is sufficiently heavy. (Note that the interpretation of the conditions is done only for the right tail to ease the text, but it is similar for the left tail.) For example, if $f(z)$ is the density of a normal distribution, $\lim_{z\to\infty} \frac{f(z+1)}{f(z)} = 0$ and condition C1 is not satisfied. If $f(z)$ is the density of a Student distribution, $\lim_{z\to\infty} \frac{f(z+\theta)}{f(z)} = 1$, for any fixed $\theta \in \mathbb{R}$ and condition C1 is satisfied.

For conditions C2 and C3, g can usually be a double Pareto density (two Pareto distributions shifted at 0 and spliced together back-to-back) if the right tail of f is lighter than that of a Pareto, otherwise g can be simply f^*. The density

$$g(z) = \begin{cases} \frac{\alpha}{2}(1+|z|)^{-(1+\alpha)}; & \text{if } \lim_{z\to\infty} \frac{f(z)}{|z|^{-(1+\alpha)}} = 0, \\ f^*(z); & \text{otherwise,} \end{cases}$$

is usually appropriate, for any choice of $\alpha > 0$. The same density $g(z)$ is also usually appropriate when the left tail is considered, except that $\lim_{z\to\infty}$ is replaced by $\lim_{z\to-\infty}$ in the first row.

Condition C2 also ensures that the tail is sufficiently heavy. Condition C3 ensures that the logarithm of the densities f^* and g (in the right tail) are convex, with the log-convexity of $f^*(z)$ more pronounced than that of $g(z)$. It can be shown that a positive function which is logarithmically convex is convex, therefore the right tails of f^* and g are also convex. Using the condition C3 with f^* instead of f is more inclusive, since it allows the inclusion of densities f with non-convex tail. Note that when condition C1 is satisfied, it usually follows that conditions C2 and C3 are also satisfied.

2.3. Rejection of conflicting information

Using the Bayesian context described in Section 2.1 and the conditions of thickness and regularity for the tails of a density described in Section 2.2, the main theorem of this paper is now presented.

Theorem 1. *For any integers k and m such that $0 \leq k \leq m \leq n$ and for any fixed $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$, if conditions* C1 *to* C3 *are satisfied on the left tails of $f_{j_{k+1}}, \ldots, f_{j_m}$ and on the right tails of $f_{j_{m+1}}, \ldots, f_{j_n}$, and if*

$$\lim_{\theta \to -\infty} \frac{\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)}{\prod_{i=k+1}^{m} f_{j_i}(\theta)} = 0 \quad \text{when } k < m, \text{ and} \tag{1}$$

$$\lim_{\theta \to \infty} \frac{\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)}{\prod_{i=m+1}^{n} f_{j_i}(\theta)} = 0 \quad \text{when } m < n\,, \tag{2}$$

then

a) $\lim_{\underset{\sim}{\phi} \to \infty} \frac{m(\underset{\sim}{x}_n)}{m(\underset{\sim}{x}_k) \prod_{i=k+1}^{n} f_{j_i}(x_{j_i})} = 1$,

b) *for any $h > 0$ and for all θ such that $|\theta| \leq h$, $\lim_{\underset{\sim}{\phi} \to \infty} \frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} = 1$,*

c) *for any $h > 0$ and $i \in (k+1, \ldots, n)$, $\lim_{\underset{\sim}{\phi} \to \infty} \Pr[|\theta - x_{j_i}| \leq h|\underset{\sim}{x}_n] = 0$,*

d) *$\theta|\underset{\sim}{x}_n \xrightarrow{\mathcal{L}} \theta|\underset{\sim}{x}_k$ as $\underset{\sim}{\phi} \to \infty$, where the density of the random variables $\theta|\underset{\sim}{x}_n$ and $\theta|\underset{\sim}{x}_k$ evaluated at the point θ are given by $\pi(\theta|\underset{\sim}{x}_n)$ and $\pi(\theta|\underset{\sim}{x}_k)$.*

In addition, for any function $w(\cdot)$ on $\mathbb{R}$ such that $\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] < \infty$ and $|w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)$ is bounded above, if

$$\lim_{\theta \to -\infty} \frac{w(\theta)\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)}{\prod_{i=k+1}^{m} f_{j_i}(\theta)} = 0 \quad \text{when } k < m, \text{ and} \tag{3}$$

$$\lim_{\theta \to \infty} \frac{w(\theta)\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)}{\prod_{i=m+1}^{n} f_{j_i}(\theta)} = 0 \quad \text{when } m < n\,, \tag{4}$$

then

e) $\lim_{\underset{\sim}{\phi} \to \infty} \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[w(\theta)] = \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[w(\theta)]$.

Proof. See Section 5.

Conditions C1 to C3 on a tail ensure that it is logarithmically convex and sufficiently heavy. Conditions given by equations (1) and (2) can be interpreted as follows. The left tail of the density (evaluated at θ) proportional to $\prod_{i=k+1}^{m} f_{j_i}(\theta)$ is heavier than the left tail of the density proportional to $\prod_{i=0}^{k} f_{j_i}(x_{j_i} - \theta)$, also denoted by $\pi(\theta|\underset{\sim}{x}_k)$. In the same way, the right tail of the density proportional to $\prod_{i=m+1}^{n} f_{j_i}(\theta)$ is heavier than the right tail of $\pi(\theta|\underset{\sim}{x}_k)$. Note that there are no conditions on the right tails of $f_{j_{k+1}}, \ldots, f_{j_m}$ or on the left tails of $f_{j_{m+1}}, \ldots, f_{j_n}$.

Asymptotic behavior of the marginal and the posterior densities is established through results a) to d), as $x_{j_i} \to -\infty$ for $i = k+1, \ldots, m$ and as $x_{j_i} \to \infty$ for $i = m+1, \ldots, n$, at any given rate. If the distance between a value x_{j_i} $(i > k)$ and the center of $\pi(\theta|\underset{\sim}{x}_k)$ increases but remains smaller than a certain threshold, the influence of this information on the posterior density usually increases. However, if this distance increases beyond the threshold, a conflict occurs and the influence of x_{j_i} gradually decreases to zero.

2.4. *Special cases*

A first special case is obtained when the information is provided only by the prior and one observation. The conflicting information is either the observation x_1 (if we set $x_{j_0} = x_0, x_{j_1} = x_1$) or the location parameter x_0 (if we set $x_{j_0} = x_1, x_{j_1} = x_0$), depending on which source we trust in case of conflict. We set $k = 0$, $n = 1$ and $\underset{\sim}{x}_n = x_1$. Since $k = 0$, $m(\underset{\sim}{x}_k)$ is set to 1 and $\pi(\theta|\underset{\sim}{x}_k)$ becomes $f_{j_0}(x_{j_0} - \theta)$. If the conflicting value $x_{j_1} \to -\infty$, we set $m = 1$, if $x_{j_1} \to \infty$, we set $m = 0$.

In Theorem 1, conditions C1 to C3 must be satisfied on the left tail of f_{j_1} when conflicting information is on the left ($m = 1$), or on the right tail of f_{j_1} when conflicting information is on the right ($m = 0$). Conditions given by equations (1) and (2) become

$$\lim_{\theta \to -\infty} \frac{f_{j_0}(x_{j_0} - \theta)}{f_{j_1}(\theta)} = 0 \text{ when } m = 1, \text{ and } \lim_{\theta \to \infty} \frac{f_{j_0}(x_{j_0} - \theta)}{f_{j_1}(\theta)} = 0 \text{ when } m = 0.$$

The marginal $m(x_1)$ behaves asymptotically like $f_{j_1}(x_{j_1})$, while the posterior $\pi(\theta|x_1)$ behaves asymptotically like $f_{j_0}(x_{j_0} - \theta)$, where the term asymptotically is understood as $x_{j_1} \to -\infty$ or $x_{j_1} \to \infty$ depending on the case. The mass of the posterior around x_{j_1} converges to 0. The posterior density converges in distribution to the source of information (prior or likelihood) with the lightest tail, if the heaviest tail is sufficiently heavy.

Another interesting special case occurs when the tail behavior is the same, for all densities $f_0, f_1, \ldots, f_n$, and for each of their left and right tails. Therefore, conditions C1 to C3 must be satisfied on each tail. This modeling is useful in practice, since it ensures the same robustness against any extreme sources of information among $x_0, x_1, \ldots, x_n$, for any direction. If the direction of the conflicting values is not specified, it can be shown that the posterior can reject up to $[n/2]$ sources of conflicting information, even if they are all on the same side, where $[a]$ stands for the integer part of a. If the direction is specified, the posterior will be insensitive to the conflicting values, as long as the number of fixed values is larger than both the number of conflicting values on the left and the number of conflicting values on the right.

2.5. Conflicting information using p-credence

The conditions in Theorem 1 can be simplified if the family of densities is restricted for the prior and the likelihood. Conditions C1 to C3 and those given by equations (1) to (4) can be easier to check if they are replaced by slightly stronger conditions stated using p-credence. The left and right p-credences characterize respectively the left and right tails of a density by comparing it to a generalized exponential power (GEP) density, see Desgagné and Angers (2005). If $f(z)$ is the density of a random variable Z, then right p-credence is denoted by p-cred$^+(f)$ or p-cred$^+(Z)$ and is defined as follows.

Definition 1. A density f has right p-credence $(\gamma, \delta, \alpha, \beta)$ if there exists a constant $K > 0$ such that

$$\lim_{z\to\infty} \frac{f(z)}{e^{-\delta|z|^{\gamma}}\, |z|^{-\alpha} \log^{-\beta}|z|} = K\,.$$

Since the GEP density is symmetric, the definition of left p-credence is identical, except that $\lim_{z\to\infty}$ is replaced by $\lim_{z\to-\infty}$. It is denoted by p-cred$^-(f)$ or p-cred$^-(Z)$. Note that p-cred$^-(Z) =$ p-cred$^+(-Z)$.

This definition ensures that the tail of a density f is proportional to the corresponding tail of the GEP density with parameters $(\gamma, \delta, \alpha, \beta)$, which guarantee a certain smoothness in the tail of f. The domain of the parameters is $\gamma \geq 0, \delta \geq 0$ (by convention $\delta = 0$ if $\gamma = 0$), $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$. Left and right p-credences are defined for most of the known densities (see Desgagné and Angers, 2003), which does not restrict too much the family of densities that can be considered.

Once right (or left) p-credence of two densities have been determined, a dominance relation can be established to compare and order their tails, as shown in Proposition 1.

Proposition 1. *Let f and g be two densities such that*

$$\text{p-cred}^{+}(f) = (\gamma, \delta, \alpha, \beta) \quad \textit{and} \quad \text{p-cred}^{+}(g) = \left(\gamma', \delta', \alpha', \beta'\right) .$$

i) *If $\gamma' = \gamma$, $\delta' = \delta$, $\alpha' = \alpha$, $\beta' = \beta$, then we say that right p-credences of f and g are equal, which is denoted by $(\gamma', \delta', \alpha', \beta') = (\gamma, \delta, \alpha, \beta)$. Their right tails are equivalent, which means that* $\lim_{z\to\infty} \frac{f(z)}{g(z)} = k$ *for a positive constant k.*

ii) *We say that right p-credence of g is smaller than that of f, which is denoted by $(\gamma', \delta', \alpha', \beta') < (\gamma, \delta, \alpha, \beta)$, as soon as either one of the following conditions is satisfied*:

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

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

The right tail of g strictly dominates the right tail of f, which means that $\lim_{z\to\infty} \frac{f(z)}{g(z)} = 0$.

The proof of Proposition 1 is given in Angers (2000). The left tails of two densities are compared and ordered in a similar way using left p-credence. The left tail of the density with the smallest left p-credence dominates the left tail of the other density.

It is possible to simplify conditions C1 to C3 with slightly stronger conditions using left and right p-credences, as stated in Proposition 2.

Proposition 2. *If* p-cred$^{+}(f) = (\gamma, \delta, \alpha, \beta)$ *and* $\gamma < 1$, *then conditions* C1 *to* C3 *are satisfied on the right tail of the density f. If* p-cred$^{-}(f) = (\gamma, \delta, \alpha, \beta)$ *and* $\gamma < 1$, *they are satisfied on its left tail.*

The proof is given in Section 5.8.

Conditions given by equations (1) to (4) can be simplified if they are replaced by slightly stronger conditions stated using p-credence. It can be shown that the left and right p-credences of a density proportional to the product of densities can be determined, under certain conditions, as given in Proposition 3.

Proposition 3. *If, for some densities $f_1, \ldots, f_s$,*

a) p-cred$^{+}(f_i) = (\gamma_i, \delta_i, \alpha_i, \beta_i)$, $i = 1, \ldots, s$, *and*

b) *each parameter γ_i is equal either to* 0 *or a constant c, with* $0 < c < 1$, $i = 1, \ldots, s$, *then*

$$\text{p-cred}^{+}(h) = \left(\max_{i=1,\dots,s} \gamma_i, \sum_{i=1}^{s} \delta_i, \sum_{i=1}^{s} \alpha_i, \sum_{i=1}^{s} \beta_i \right),$$

where the density h is such that

$$h(\theta) \propto \prod_{i=1}^{s} f_i(\theta - u_i),$$

for any constants $u_1, \dots, u_s$. By convention, $\delta_i = 0$ if $\gamma_i = 0$.

Using left and right p-credences, we can now state sufficient conditions needed to obtain results a) to d) of Theorem 1.

Theorem 2. *Suppose that* $\text{p-cred}^{-}(f_i) = (\gamma_i^-, \delta_i^-, \alpha_i^-, \beta_i^-)$ *and* $\text{p-cred}^{+}(f_i) = (\gamma_i^+, \delta_i^+, \alpha_i^+, \beta_i^+)$*, where each parameter γ_i^- and γ_i^+ can be equal either to* 0 *or a constant c, with* $0 < c < 1$*,* $i = 0, 1, \dots, n$*. For any integers k and m such that* $0 \le k \le m \le n$ *and for any fixed* $x_{j_0}, x_{j_1}, \dots, x_{j_k}$*, if*

$$\left(\max_{i=k+1,\dots,m} \gamma_{j_i}^-, \sum_{i=k+1}^{m} \delta_{j_i}^-, \sum_{i=k+1}^{m} \alpha_{j_i}^-, \sum_{i=k+1}^{m} \beta_{j_i}^- \right)$$
$$< \left(\max_{i=0,\dots,k} \gamma_{j_i}^+, \sum_{i=0}^{k} \delta_{j_i}^+, \sum_{i=0}^{k} \alpha_{j_i}^+, \sum_{i=0}^{k} \beta_{j_i}^+ \right) \quad \text{when } k < m, \quad \text{and}$$

$$\left(\max_{i=m+1,\dots,n} \gamma_{j_i}^+, \sum_{i=m+1}^{n} \delta_{j_i}^+, \sum_{i=m+1}^{n} \alpha_{j_i}^+, \sum_{i=m+1}^{n} \beta_{j_i}^+ \right)$$
$$< \left(\max_{i=0,\dots,k} \gamma_{j_i}^-, \sum_{i=0}^{k} \delta_{j_i}^-, \sum_{i=0}^{k} \alpha_{j_i}^-, \sum_{i=0}^{k} \beta_{j_i}^- \right) \quad \text{when } m < n,$$

then results a) *to* d) *of Theorem* 1 *are obtained.*

In addition, for any function $w(\cdot)$ on $\mathbb{R}$ such that $\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] < \infty$, $|w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)$ is bounded above, if $\text{p-cred}^{-}(|w|) = (\gamma_w^-, \delta_w^-, \alpha_w^-, \beta_w^-)$ *and* p-cred^{+} $(|w|) = (\gamma_w^+, \delta_w^+, \alpha_w^+, \beta_w^+)$*, where each parameter γ_w^- and γ_w^+ can be equal either* *to* 0 *or the constant c, and if*

$$\left(\max_{i=k+1,\dots,m} \gamma^-_{j_i}, \sum_{i=k+1}^{m} \delta^-_{j_i}, \sum_{i=k+1}^{m} \alpha^-_{j_i}, \sum_{i=k+1}^{m} \beta^-_{j_i}\right)$$

$$< \left(\max_{i=0,\dots,k}(\gamma^+_{j_i}, \gamma^-_w), \sum_{i=0}^{k} \delta^+_{j_i} + \delta^-_w, \sum_{i=0}^{k} \alpha^+_{j_i} + \alpha^-_w, \sum_{i=0}^{k} \beta^+_{j_i} + \beta^-_w\right)$$

when $k < m$, *and*

$$\left(\max_{i=m+1,\dots,n} \gamma^+_{j_i}, \sum_{i=m+1}^{n} \delta^+_{j_i}, \sum_{i=m+1}^{n} \alpha^+_{j_i}, \sum_{i=m+1}^{n} \beta^+_{j_i}\right)$$

$$< \left(\max_{i=0,\dots,k}(\gamma^-_{j_i}, \gamma^+_w), \sum_{i=0}^{k} \delta^-_{j_i} + \delta^+_w, \sum_{i=0}^{k} \alpha^-_{j_i} + \alpha^+_w, \sum_{i=0}^{k} \beta^-_{j_i} + \beta^+_w\right)$$

when $m < n$,

then result e) *of Theorem* 1 *is obtained.*

It is easy to verify Theorem 2 using Propositions 1 to 3.

3. Examples

Suppose that a portfolio manager needs a prediction on the return of a stock index for the next day. He asks 10 experts for their predictions on the return. The manager wants to combine this information with his prior belief using the Bayesian model described in Section 2.1. According to this setting, the manager chooses $f_i(x_i - \theta) = \frac{1}{\sigma_i} h\left(\frac{x_i - \theta}{\sigma_i}\right)$, for $i = 0, 1, \dots, 10$, where $h(\cdot)$ is a Student density with ν degrees of freedom and σ_i is a known scale parameter.

The random variables $X_1, \dots, X_{10}$ represent the prediction of each expert and they are conditionally independent given θ. Each X_i has a distinct distribution, represented by the densities f_i, but they share the same location parameter θ. Since the Student density is symmetric, the parameter θ represents the mean, median and mode of the predictions of all the experts. The portfolio manager will use the posterior expectation of θ as his combined prediction.

The densities f_i in this setting differ only by the scale parameters σ_i. If the standard deviation of X_i is denoted by s_i, then it can be shown that $s_i = \sqrt{\frac{\nu}{\nu - 2}}\,\sigma_i$. The parameters s_i represent the volatility of the predictions of expert i. This means that there is $100(1-\alpha)\%$ chance that the prediction X_i is included in the interval $\theta \pm t_{\alpha/2}(\nu)\sqrt{\frac{\nu-2}{\nu}}\, s_i$, where $t_{\alpha/2}(\nu)$ is the quantile of a Student distribution with ν degrees of freedom and $0 < \alpha < 1$. For example, there is approximately 95% chance that the prediction X_i is included in the interval $\theta \pm 2s_i$, since $t_{\alpha/2}(\nu)\sqrt{\frac{\nu-2}{\nu}}$ is between 1.96 and 2 for all $\nu > 3$ if $\alpha = 0.05$. Inference on θ based only on expert i would produce an estimate of θ given by the observation x_i and an approximate 95% confidence interval for θ given by $x_i \pm 2s_i$.

We arbitrarily chose the degrees of freedom equal to $\nu = 10$ and all standard deviations are assumed equal to $s_i = 1$, or equivalently $\sigma_i = \sqrt{0.8}$, $i = 0, 1, \ldots, 10$. The uncertainty of each source of information (experts and prior) is the same since σ_i is constant, which means that the weight given to each source in the posterior inference will be the same, in case of no conflict. In case of conflict, the credence given to each source will also be the same, since the left and right p-credences of each density are the same, that is $(0, 0, 11, 0)$, where 11 represents the degrees of freedom plus 1. Note that if we wanted to give no weight to the prior information, we could simply set $f_0(x_0 - \theta) \equiv 1$, an improper uniform density defined on $\mathbb{R}$.

Four examples are presented. In each case, we consider that $k+1$ sources of information are fixed, $10-m$ of them are equal to x^* and $m-k$ are equal to $-x^*$, for different values of k and m such that $0 \le k \le m \le 10$. The posterior expectation of θ was computed for values of x^* ranging from -60 to 60 using Monte Carlo simulations with importance sampling, see Desgagné and Angers (2005). Note that it could be estimated using other appropriate numerical methods. The value x^* can be considered as conflicting if its distance with the fixed values is sufficiently large. Note that no distinctions are made between the prior location and the ten observations.

In the first case, $m = k = 9$, and the 10 fixed values are given by $\underset{\sim}{x}_k = (x_{j_0}, x_{j_1}, \ldots, x_{j_9}) = (-2, -2, -1, -1, 0, 0, 1, 1, 2, 2)$. There is only one source of conflicting information denoted by $x_{j_{10}} = x^*$, which can be a prediction of an expert or the location parameter of the prior, x_0. Note that all numbers in this example are expressed in percentage since they represent returns. The posterior expectations of θ are plotted in figure 1 and correspond to the line labeled (10 fixed, 1 x*).

Three other cases are plotted in figure 1. In the second case, $k = 4$, $m = 6$ and the 5 fixed values are given by $(x_{j_0}, x_{j_1}, \ldots, x_{j_4}) = (-1, -1, 0, 1, 1)$. The 2 values (x_{j_5}, x_{j_6}) are equal to $-x^*$ and the 4 values $(x_{j_7}, \ldots, x_{j_{10}})$ are equal to x^*. For example, if $x^* = 4$, the information is given by the vector $(-1, -1, 0, 1, 1, -4, -4, 4, 4, 4, 4)$ and the posterior mean of θ is 0.9 (y-axis), which corresponds approximately to the mode of the line labeled (5 fixed, 4 x*, 2 (-x*)), at the point $x^* = 4$.

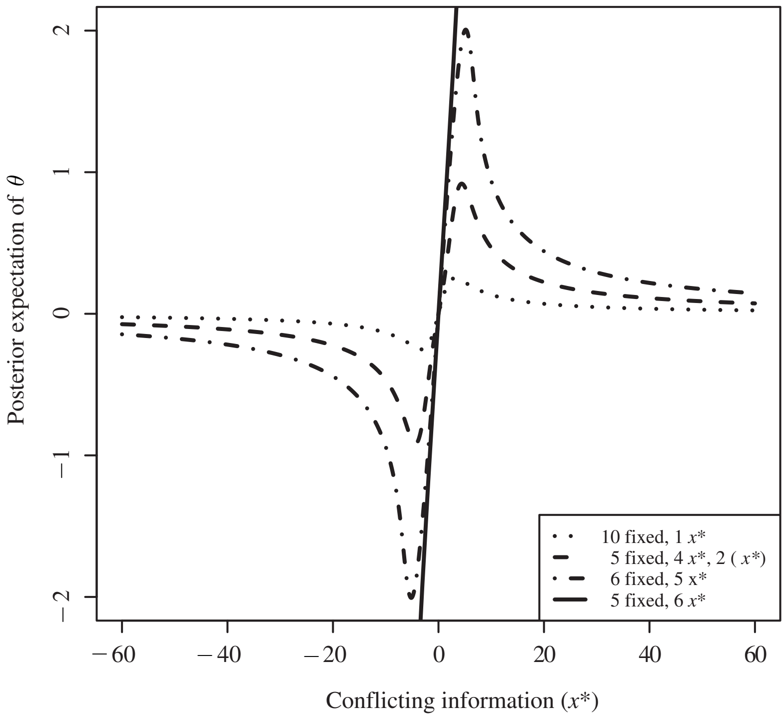

Figure 1. The posterior expectation of θ.

In the third case labeled (6 fixed, 5 x*), $k = 5, m = 5$ and the 6 fixed values are given by $(x_{j_0}, x_{j_1}, \ldots, x_{j_5}) = (-1, -1, 0, 0, 1, 1)$, while the 5 values $(x_{j_6}, \ldots, x_{j_{10}})$ are equal to x^*. In the fourth case labeled (5 fixed, 6 x*), $k = 4, m = 4$ and the 5 fixed values are given by $(x_{j_0}, x_{j_1}, \ldots, x_{j_4}) = (-1, -1, 0, 1, 1)$, while the 6 values $(x_{j_5}, \ldots, x_{j_{10}})$ are equal to x^*.

The left and right p-credences are given by $\text{p-cred}^-(f_i) = \text{p-cred}^+(f_i) = (\gamma_i^-, \delta_i^-, \alpha_i^-, \beta_i^-) = (\gamma_i^+, \delta_i^+, \alpha_i^+, \beta_i^+) = (0, 0, 11, 0)$ for all densities, that is for $i = 0, 1, \ldots, n$. It can be easily verified that conditions C1 to C3 are satisfied for each tail of every density since $\gamma_i^- < 1$ and $\gamma_i^+ < 1$. The tails of the Student density are sufficiently heavy and logarithmically convex. It can be verified that conditions given by equations (1) and (2) of Theorem 1 are satisfied if $(0, 0, \sum_{i=k+1}^{m} \alpha_i^-, 0) < (0, 0, \sum_{i=0}^{k} \alpha_i^+, 0)$ and $(0, 0, \sum_{i=m+1}^{n} \alpha_i^+, 0) < (0, 0, \sum_{i=0}^{k} \alpha_i^-, 0)$, which is equivalent to $\max(m - k, n - m) < k + 1$ since $\alpha_i^- = \alpha_i^+ = 11$ for all i. If the number of conflicting values on each side is smaller than the number of fixed values, then the posterior density will reject the conflicting values as they become extreme. The model can reject up to 5 conflicting values even if they are on the same side and up to 6 conflicting values if they are split on each side, with at least 2 of them on each side.

If we consider the posterior expectation of θ, the function w is $w(\theta) = \theta$ and p-cred$^-(|w|)$ = p-cred$^+(|w|)$ = $(\gamma_w^-, \delta_w^-, \alpha_w^-, \beta_w^-) = (\gamma_w^+, \delta_w^+, \alpha_w^+, \beta_w^+) = (0, 0, -1, 0)$. It can be verified that conditions given by equations (3) and (4) of Theorem 1 are satisfied if $(0, 0, \sum_{i=k+1}^{m} \alpha_i^-, 0) < (0, 0, \sum_{i=0}^{k} \alpha_i^+ + \alpha_w^-, 0)$ and $(0, 0, \sum_{i=m+1}^{n} \alpha_i^+, 0) < (0, 0, \sum_{i=0}^{k} \alpha_i^- + \alpha_i^+, 0)$, which is equivalent to $\max(m - k, n - m) + 1/11 < k + 1$. Clearly, the term 1/11 has no impact, and these conditions are equivalent to $\max(m - k, n - m) < k + 1$. This means that the posterior expectation of θ has the same robustness as the posterior density.

In the first case with 10 fixed values, when $x_{j_{10}}$ increases from 0 to a certain threshold (around 3), the posterior expectation also increases. Beyond this threshold, the influence of $x_{j_{10}}$ gradually decreases to zero, as $x_{j_{10}} \to \infty$. In the limit case, the posterior density considering all observations converges to the density proportional to $\prod_{i=0}^{9} f_{j_i}(x_{j_i} - \theta)$, which is the posterior density excluding the outlier in the particular case where the conflicting value is an observation. The posterior expectation converges to 0 for this case. The interpretation is the same if $x_{j_{10}} \to -\infty$.

The model can reject up to 6 conflicting values, only if they are split on the left and right sides, as can be seen in the second case. However, if the 6 conflicting values are all on the same side, as in the fourth case, the posterior does not reject them. Actually, the conflicting information becomes the 5 fixed values in this case. Finally, the model can reject up to 5 conflicting values, even if they are on the same side, as in the third case.

4. Conclusion

In this paper, the behavior of the posterior density of a location parameter has been investigated when the sample contains outliers or the prior location is misspecified. Conditions on the tails of the prior and the likelihood have been established to determine the proportion of conflicting information that can be rejected by the posterior. It has been shown that the posterior distribution converges in law to a density proportional to the product of the densities of the non-conflicting information, as the outliers (and/or the prior location) go to plus or minus infinity, at any given rate. In particular, if the prior is non-conflicting, this limiting density is the posterior that would be obtained from the reduced sample, excluding the outliers. Examples have been given to illustrate the results.

5. Proofs

The proof of Theorem 1 is given in this section. In Section 5.1, the proof of result a) of Theorem 1 is given. The proof of Lemma 3, needed for this proof, is given in Section 5.2. The proofs of results b) to e) are given through Sections 5.3 to 5.6. The proof of Lemma 10, needed for the proof of result e) in Section 5.6, is given in Section 5.7.

5.1. Proof of result a) *of Theorem* 1

It is assumed that the densities $f_0, f_1, \dots, f_n$ are proper, positive everywhere and bounded above. Without loss of generality, we will assume that $x_{j_i} = x_i$, for $i = 0, 1, \dots, n$, in order to ease the notation. This means that $x_0, x_1, \dots, x_k$ are considered fixed, $x_{k+1}, \dots, x_m$ tend to $-\infty$ and $x_{m+1}, \dots, x_n$ tend to ∞.

If conditions C2 and C3 are satisfied on a tail of a density f, the constants B, K_1, K_2, A_2, M_2 and the density g are determined. Note that we can choose the constants A_2 and M_2 to be as large as we want. To ease the notation later in the proof, we choose $M_2 = \max(M_2, 1/K_1, K_2, \sup_z f(z), g(A_2), 6)$ and $A_2 = \max(A_2, B)$. It follows that f^* is such that $z > A_2$ ($z < -A_2$ for the left tail) $\Rightarrow \frac{1}{M_2} \leq \frac{f(z)}{f^*(z)} \leq M_2$.

It is easy to show that the marginals $m(\underset{\sim}{x}_k)$ and $m(\underset{\sim}{x}_n)$ are positive and bounded above and that the posterior densities $\pi(\theta|\underset{\sim}{x}_k)$ and $\pi(\theta|\underset{\sim}{x}_n)$ are also proper, positive everywhere and bounded above. Note that since the prior location x_0 is considered fixed to ease the notation of the proof, $\pi(\theta|\underset{\sim}{x}_k)$ is the posterior density and $m(\underset{\sim}{x}_k)$ is the marginal. Considering that $0 \leq \int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k)d\theta \leq 1$ for any $h > 0$ and that $\pi(\theta|\underset{\sim}{x}_k)$ depends only on the fixed values $x_0, x_1, \dots, x_k$, it is then possible to show the following lemma.

Lemma 1. $\forall \epsilon > 0$, *there exists a constant* $A_4(\epsilon) > 0$ *such that* $h \geq A_4(\epsilon) \Rightarrow$

$$\int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k)d\theta \geq 1 - \epsilon, \int_{-\infty}^{-h} \pi(\theta|\underset{\sim}{x}_k)d\theta \leq \epsilon \text{ and } \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_k)d\theta \leq \epsilon\,.$$

Assuming that conditions C1 to C3 are satisfied on the right tail of a density f which is proper, positive everywhere and bounded above, two other lemmas needed for the proof are given. Note that if conditions C1 to C3 are satisfied on the left tail of f, the lemmas are the same except for the support, written in parentheses.

Lemma 2. $z > A_2$ *and* $\theta > 0$ $(z < -A_2$ *and* $\theta < 0) \Rightarrow f(z+\theta) \leq (M_2)^2 f(z)$.

Proof. It can be shown that if a positive function is logarithmically convex, then it is also convex. It also can be shown that if the right tail of a proper density is convex, then it is necessarily decreasing. Since f^* is proper and logarithmically convex when $z > A_2$ (see C3), then the right tail of f^* is decreasing, that is $z > A_2 \Rightarrow f^*(z+\theta) < f^*(z), \forall\theta > 0$. Therefore, $z > A_2$

and $\theta > 0 \Rightarrow f(z+\theta) \leq M_2 f^*(z+\theta) \leq M_2 f^*(z) \leq (M_2)^2 f(z)$. Condition C3 is used in the first and last inequalities. The proof for the left tail is similar.

Lemma 3. *$h > A_2$, $z > \max[2h, A_1(1,h)]$ and $\mathbb{D} = [h, \infty)$ ($z < \min[-2h, -A_1(1,h)]$ and $\mathbb{D} = (-\infty, -h]$ for the left tail) $\Rightarrow$*

$$\int_{\mathbb{D}} \frac{f(z-\theta)f(\theta)}{f(z)} d\theta \leq (M_2)^{10} \ \text{ and } \ \frac{f(z-\theta)f(\theta)}{f(z)} \leq (M_2)^{11} \ \text{ for all } \theta \in \mathbb{D}.$$

Proof. See Section 5.2.

Using the fact that the numerators in equations (1) and (2) are proportional to $\pi(\theta|\underset{\sim}{x}_k)$ ($m(\underset{\sim}{x}_k)$ depends only on fixed values $x_0, x_1, \ldots, x_k$), equations (1) and (2) respectively can be rewritten as follows (assuming $k < m$ and $m < n$): $\forall \epsilon > 0$, there exists a constant $A_3(\epsilon) > 0$ such that

$$\theta < -A_3(\epsilon) \Rightarrow \frac{\pi(\theta|\underset{\sim}{x}_k)}{\prod_{i=k+1}^{m} f_i(\theta)} \leq \epsilon \quad \text{and} \quad \theta > A_3(\epsilon) \Rightarrow \frac{\pi(\theta|\underset{\sim}{x}_k)}{\prod_{i=m+1}^{n} f_i(\theta)} \leq \epsilon\,. \tag{5}$$

Denote $y_i = -x_i$ if $i = k+1, \ldots, m$ and $y_i = x_i$ if $i = m+1, \ldots, n$. It follows that $\underset{\sim}{\phi} = (y_{k+1}, \ldots, y_m, y_{m+1}, \ldots, y_n)$ and $\underset{\sim}{\phi} \to \infty$ if $y_i \to \infty$, for $i = k+1, \ldots, n$, at any given rate. Then

$$\begin{aligned}
\frac{m(\underset{\sim}{x}_n)}{m(\underset{\sim}{x}_k) \prod_{i=k+1}^{n} f_i(x_i)} &= \frac{\int_{-\infty}^{\infty} \prod_{i=0}^{n} f_i(x_i - \theta) d\theta}{m(\underset{\sim}{x}_k) \prod_{i=k+1}^{n} f_i(x_i)} \\
&= \frac{\int_{-\infty}^{\infty} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{n} f_i(x_i - \theta) d\theta}{\prod_{i=k+1}^{n} f_i(x_i)} \\
&= \int_{-\infty}^{\infty} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i - \theta)}{f_i(y_i)} d\theta\,.
\end{aligned}$$

Then result a) can be rewritten as follows: $\forall \epsilon > 0$, there exists a constant $A_0(\epsilon)$ such that $y_{k+1} > A_0(\epsilon), \ldots, y_n > A_0(\epsilon) \Rightarrow$

$$1 - \epsilon \leq \int_{-\infty}^{\infty} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i - \theta)}{f_i(y_i)} d\theta \leq 1 + \epsilon\,.$$

First choose any $0 < \epsilon < 1$. Note that if the result is true for $0 < \epsilon < 1$, it is necessarily true for any $\epsilon > 0$. Then define

$$\epsilon_0 = \min\left(\left[(1+\epsilon/3)^{1/(n-k)} - 1\right], \left[1-(1-\epsilon/3)^{1/(n-k+1)}\right], (M_2)^{-11n}\epsilon/3\right).$$

Define $h = \max(A_2, A_3(\epsilon_0), A_4(\epsilon_0))$ and then $A_0(\epsilon) = \max(A_1(\epsilon_0, h), 2h)$. Note that $0 < \epsilon_0 < \frac{1}{3}$, $h > 0$ and $A_0(\epsilon)$ depends only on ϵ. The constant A_1 comes from condition C1, A_2 and M_2 from conditions C2 and C3, A_3 from equation (5) and A_4 from Lemma 1. Actually, there are different constants $A_1(\epsilon, h)$, A_2 and M_2 defined for each density $f_{k+1}, \ldots, f_n$ in conditions C1 to C3. To simplify the notation, $A_1(\epsilon, h)$ corresponds to the largest constant $A_1(\epsilon, h)$, for each specified ϵ and h. In the same way A_2 and M_2 correspond to the largest constant. The integral is divided into three parts: $(-\infty, -h]$, $(-h, h]$ and (h, ∞) and consider that $y_{k+1} > A_0(\epsilon), \ldots, y_n > A_0(\epsilon)$. First consider the integral on $(-h, h]$.

$$\begin{aligned}
&\int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i-\theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i-\theta)}{f_i(y_i)} d\theta \\
&\geq \int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} (1-\epsilon_0) \prod_{i=m+1}^{n} (1-\epsilon_0) d\theta \\
&= (1-\epsilon_0)^{n-k} \int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k) d\theta \geq (1-\epsilon_0)^{n-k+1} \geq 1-\epsilon/3\,.
\end{aligned}$$

Note that C1 is used in the first inequality since $-y_i < -A_1(\epsilon_0, h)$ and $y_i > A_1(\epsilon_0, h)$ and Lemma 1 is used in the second inequality since $h \geq A_4(\epsilon_0)$. In a similar way, it can be shown that

$$\int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i-\theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i-\theta)}{f_i(y_i)} d\theta \leq 1+\epsilon/3\,.$$

Consider now (h, ∞) if $m < n$ is assumed.

$$\begin{aligned}
&\int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i-\theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i-\theta)}{f_i(y_i)} d\theta \\
&\leq (M_2)^{2(m-k)} \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_k) \prod_{i=m+1}^{n} \frac{f_i(y_i-\theta)}{f_i(y_i)} d\theta \\
&\leq \epsilon_0 (M_2)^{2(m-k)} \int_{h}^{\infty} \prod_{i=m+1}^{n} \frac{f_i(y_i-\theta) f_i(\theta)}{f_i(y_i)} d\theta \\
&\leq \epsilon_0 (M_2)^{2(m-k)} (M_2)^{11(n-m-1)} \int_{h}^{\infty} \frac{f_n(y_n-\theta) f_n(\theta)}{f_n(y_n)} d\theta \\
&\leq \epsilon_0 (M_2)^{2(m-k)} (M_2)^{11(n-m-1)} (M_2)^{10} \leq \epsilon_0 (M_2)^{11n} \leq \epsilon/3\,.
\end{aligned}$$

Note that Lemma 2 is used in the first inequality since $-y_i - \theta < -y_i < -A_2$, equation (5) is used in the second one since $\theta \geq A_3(\epsilon_0)$, Lemma 3 is used in the third one since $h > A_2$ and $y_i > \max[2h, A_1(1,h)]$, Lemma 3 is used again in the fourth one, and the inequality $0 \leq k \leq m \leq n$ is used in the fifth one. Consider now (h, ∞) if $m = n$ is assumed.

$$\int_h^\infty \pi(\theta | \underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} d\theta \leq (M_2)^{2(m-k)} \int_h^\infty \pi(\theta | \underset{\sim}{x}_k) d\theta$$
$$\leq \epsilon_0 (M_2)^{2(m-k)} \leq \epsilon_0 (M_2)^{11n} \leq \epsilon/3 \,.$$

Note that Lemma 2 is used in the first inequality since $-y_i - \theta < -y_i < -A_2$ and Lemma 1 is used in the second one since $h > A_4(\epsilon_0)$. In a similar way, it can be shown that

$$\int_{-\infty}^{-h} \pi(\theta | \underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i - \theta)}{f_i(y_i)} d\theta \leq \epsilon/3 \,.$$

Considering the three parts of the integral, we showed that

$$\int_{-\infty}^\infty \pi(\theta | \underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i - \theta)}{f_i(y_i)} d\theta$$
$$\leq (1 + \epsilon/3) + \epsilon/3 + \epsilon/3 = 1 + \epsilon \,,$$

and

$$\int_{-\infty}^\infty \pi(\theta | \underset{\sim}{x}_k) \prod_{i=k+1}^{m} \frac{f_i(-y_i - \theta)}{f_i(-y_i)} \prod_{i=m+1}^{n} \frac{f_i(y_i - \theta)}{f_i(y_i)} d\theta$$
$$\geq (1 - \epsilon/3) + 0 + 0 > 1 - \epsilon \,.$$

5.2. *Proof of Lemma* 3

We first introduce four other lemmas needed to prove Lemma 3. Suppose that conditions C1 to C3 are satisfied on the right tail of a proper density f, positive everywhere and bounded above. (If conditions C1 to C3 are satisfied on its left tail, the lemmas are the same, except for the support given in parentheses. The proofs are given only for the right tail, the proofs for the left tail being similar.)

Lemma 4. $z > A_2$ ($z < -A_2$ *for the left tail*) $\Rightarrow f^*(z) > 0$ *and* $g(z) > 0$.

Proof. If $f^*(z) = 0$ for some $z > A_2$, then $f^*(z)$ cannot have the same tail behavior as f, see condition C3. If $g(z) = 0$ for some $z > A_2$, then condition C2 is not satisfied.

Lemma 5. $z > A_2$ $(z < -A_2) \Rightarrow f(z) \leq (M_2)^3 g(z)$.

Proof. Using Lemma 2, if $z > A_2$ then $f(2z) \leq (M_2)^2 f(z)$. Using C2, if $z > A_2$ then $(M_2)g(z) \geq \frac{f^2(z)}{f(2z)}$. Therefore $z > A_2 \Rightarrow (M_2)^3 g(z) \geq (M_2)^2 \frac{f^2(z)}{f(2z)} = f(z)\frac{(M_2)^2 f(z)}{f(2z)} \geq f(z)$.

Lemma 6. $z > A_2 \Rightarrow g(z) < g(A_2)$ $(z < -A_2 \Rightarrow g(z) < g(-A_2))$.

Proof. Since g is a proper density and it is logarithmically convex (see C3) when $z > A_2$, then the right tail of g is decreasing and then bounded above by $g(A_2)$.

Lemma 7. *For all a, b and z such that* $A_2 \leq a \leq b \leq z - A_2$ $(z + A_2 \leq a \leq b \leq -A_2$ *for the left tail*), $\arg\max_{a \leq \theta \leq b} \frac{f^*(z-\theta) f^*(\theta)}{g(\theta)} \in \{a, b\}$.

Proof. Since the maximum on a range of a convex function is located at its bounds, it is sufficient to show that $\frac{d^2}{d\theta^2} \log \frac{f^*(z-\theta) f^*(\theta)}{g(\theta)} \geq 0$ for any θ such that $A_2 < \theta < z - A_2$, since the convexity of the logarithm of a positive function implies the convexity of the function. Then

$$\frac{d^2}{d\theta^2} \log \frac{f^*(z-\theta) f^*(\theta)}{g(\theta)} = \frac{d^2}{d\theta^2} \log f^*(z-\theta) + \frac{d^2}{d\theta^2} \log f^*(\theta) - \frac{d^2}{d\theta^2} \log g(\theta)\,.$$

Using C3, $\frac{d^2}{d\theta^2} \log f^*(\theta) - \frac{d^2}{d\theta^2} \log g(\theta) \geq 0$ for $\theta > A_2$. It can also be shown that $\frac{d^2}{d\theta^2} \log f^*(z-\theta) = \left(\frac{d^2}{dy^2} \log f^*(y)\right) |_{y=z-\theta}$, and using C3, it is non negative for $z - \theta > A_2$. Then we showed that $\frac{d^2}{d\theta^2} \log \frac{f^*(z-\theta) f^*(\theta)}{g(\theta)} \geq 0$ if $\theta > A_2$ and $z - \theta > A_2$, that is if $A_2 < \theta < z - A_2$. □

To prove Lemma 3, we divide $[h, \infty)$ into three parts: $[h, z/2]$, $(z/2, z-h]$ and $(z-h, \infty)$. Consider that $h > A_2$ and $z > \max[2h, A_1(1, h)]$. The constants A_1 and A_2 respectively come from conditions C1 and C2.

First consider $h \le \theta \le z/2$. Note that $h \le \theta \le z/2$, $h > A_2$ and $z > 2h \Rightarrow z > z - A_2 > z - h \ge z - \theta \ge z/2 \ge \theta \ge h > A_2$. Then

$$\begin{aligned}
\frac{f(z-\theta)f(\theta)}{f(z)} &\le (M_2)^3 \frac{f^*(z-\theta)f^*(\theta)}{f^*(z)} = (M_2)^3 \left(\frac{f^*(z-\theta)f^*(\theta)}{f^*(z)g(\theta)}\right) g(\theta) \\
&\le (M_2)^3 \max\left(\frac{f^*(z-h)f^*(h)}{f^*(z)g(h)}, \frac{f^{*2}(z/2)}{f^*(z)g(z/2)}\right) g(\theta) \\
&\le (M_2)^6 \max\left(\frac{f(z-h)f(h)}{f(z)g(h)}, \frac{f^2(z/2)}{f(z)g(z/2)}\right) g(\theta) \\
&\le (M_2)^6 \max\left(\frac{f(z-h)(M_2)^3}{f(z)}, M_2\right) g(\theta) \\
&\le (M_2)^6 \max\left(2(M_2)^3, M_2\right) g(\theta) = 2(M_2)^9 g(\theta) \\
&\le 2(M_2)^9 g(A_2) \le (M_2)^{11}\,.
\end{aligned}$$

Note that C3 is used in the first inequality since $z - \theta > A_2$, $\theta > A_2$ and $z > A_2$, Lemma 7 is used in the second one since $A_2 < h \le \theta \le z/2 < z - A_2$, C3 is used in the third one since $z - h > A_2, h > A_2, z > A_2$ and $z/2 > A_2$, Lemma 5 is used in the fourth one since $h > A_2$ and C2 is used since $z > A_2$, C1 is used in the fifth one since $z > A_1(1, h)$, Lemma 6 is used in the sixth one since $\theta > A_2$ and finally $2 < M_2$ and $g(A_2) \le M_2$ are used in the last inequality. Furthermore, since $g(\cdot)$ is a proper density,

$$\int_h^{z/2} \frac{f(z-\theta)f(\theta)}{f(z)} d\theta \le 2(M_2)^9 \int_h^{z/2} g(\theta) d\theta \le 2(M_2)^9\,.$$

Consider now $z/2 \le \theta \le z - h$. It is possible to use the preceding results (when $h \le \theta \le z/2$ is considered) if the change of variables $u = z - \theta$ is done, since $h \le u \le z/2$. Then

$$\begin{aligned}
\frac{f(z-\theta)f(\theta)}{f(z)} = \frac{f(u)f(z-u)}{f(z)} &\le 2(M_2)^9 g(u) = 2(M_2)^9 g(z-\theta) \\
&\le 2(M_2)^9 g(A_2) \le (M_2)^{11}\,.
\end{aligned}$$

Lemma 6 is used in the second inequality since $z - \theta > A_2$. Furthermore, since $g(\cdot)$ is a proper density,

$$\int_{z/2}^{z-h} \frac{f(z-\theta)f(\theta)}{f(z)} d\theta \le 2(M_2)^9 \int_{z/2}^{z-h} g(z-\theta) d\theta \le 2(M_2)^9\,.$$

Finally consider $\theta \geq z-h$.

$$\frac{f(z-\theta)f(\theta)}{f(z)} \leq \frac{f(z-\theta)(M_2)^2 f(z-h)}{f(z)} \leq 2(M_2)^2 f(z-\theta)$$
$$\leq 2(M_2)^3 \leq (M_2)^{11}\,.$$

Note that Lemma 2 is used in the first inequality since $\theta \geq z-h > A_2$, C1 is used in the second one since $z > A_1(1,h)$, $\sup_{z\in\mathbb{R}} f(z) \leq M_2$ is used in the third inequality and $2 < M_2$ is used in the last one. Furthermore, since $f(\cdot)$ is a proper density,

$$\int_{z-h}^{\infty} \frac{f(z-\theta)f(\theta)}{f(z)} d\theta \leq 2(M_2)^2 \int_{z-h}^{\infty} f(z-\theta)d\theta \leq 2(M_2)^2 \leq 2(M_2)^9\,.$$

If the integrals on the three domains are considered, then

$$\int_{h}^{\infty} \frac{f(z-\theta)f(\theta)}{f(z)} d\theta \leq 6(M_2)^9 \leq (M_2)^{10}\,.$$

5.3. Proof of result b) *of Theorem* 1

Result b) can be rewritten as follows: $\forall\epsilon > 0, \forall h > 0$ there exists a constant $A_5(\epsilon,h)$ such that $\min[\underset{\sim}{\phi}] > A_5(\epsilon,h)$ and $|\theta| \leq h \Rightarrow 1-\epsilon \leq \frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} \leq 1+\epsilon$. Note that $\min[\underset{\sim}{\phi}]$ stands for $\min[-x_{k+1},\ldots,-x_m, x_{m+1},\ldots,x_n]$. Result a) of Theorem 1 can also be rewritten as follows: $\forall\epsilon > 0$ there exists a constant $A_0(\epsilon)$ such that $\min[\underset{\sim}{\phi}] > A_0(\epsilon) \Rightarrow 1-\epsilon \leq \frac{m(\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i)}{m(\underset{\sim}{x}_n)} \leq 1+\epsilon$.

Choose any $\epsilon > 0$ and any $h > 0$. Then define

$$\epsilon_0 = \min[(1+\epsilon)^{1/(n-k+1)} - 1, 1-(1-\epsilon)^{1/(n-k+1)}]$$

and $A_5(\epsilon,h) = \max[A_0(\epsilon_0), A_1(\epsilon_0,h)]$. The constants A_0 and A_1 respectively come from the proof of result a) of Theorem 1 and from condition C1. Consider that $\min[\underset{\sim}{\phi}] > A_5(\epsilon,h)$ and $|\theta| \leq h$. Then

$$\frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} = \frac{m(\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i-\theta)}{m(\underset{\sim}{x}_n)} = \frac{m(\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i)}{m(\underset{\sim}{x}_n)} \prod_{i=k+1}^{n} \frac{f_i(x_i-\theta)}{f_i(x_i)}$$
$$\leq (1+\epsilon_0)\prod_{i=k+1}^{n} \frac{f_i(x_i-\theta)}{f_i(x_i)} \leq (1+\epsilon_0)\prod_{i=k+1}^{n}(1+\epsilon_0)$$
$$= (1+\epsilon_0)^{n-k+1} \leq 1+\epsilon\,.$$

Result a) is used in the first inequality since $\min[\underset{\sim}{\phi}] > A_0(\epsilon_0)$ and C1 is used in the second one since $\min[\underset{\sim}{\phi}] > A_1(\epsilon_0, h)$. In a similar way, it can be shown that $\frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} \geq 1 - \epsilon$.

5.4. Proof of result c) *of Theorem* 1

Result c) of Theorem 1 says that the posterior density tends to 0, in any finite neighborhood of any conflicting values (or outliers, without loss of generality) x_j, $j \in (k+1, \ldots, n)$. It can be rewritten as follows: $\forall \epsilon > 0, \forall d > 0$, there exists a constant $A_6(\epsilon, d)$ such that $\min[\underset{\sim}{\phi}] > A_6(\epsilon, d)$ and $j \in (k+1, \ldots, n) \Rightarrow \Pr[|\theta - x_j| \leq d|\underset{\sim}{x}_n] \leq \epsilon$. A lemma analogous to Lemma 1 is needed for the proof.

Lemma 8. $\forall \epsilon > 0, \forall h \geq A_4(\epsilon/2)$, *we have* $\min[\underset{\sim}{\phi}] > A_5(\epsilon/2, h) \Rightarrow$

$$\int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_n)d\theta \geq 1 - \epsilon, \int_{-\infty}^{-h} \pi(\theta|\underset{\sim}{x}_n)d\theta \leq \epsilon \text{ and } \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_n)d\theta \leq \epsilon\,.$$

Proof. $\int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_n)d\theta \geq (1 - \epsilon/2) \int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_k)d\theta \geq (1 - \epsilon/2)^2 > 1 - \epsilon$. Result b) of Theorem 1 is used in the first inequality since $\min[\underset{\sim}{\phi}] > A_5(\epsilon/2, h)$ and $|\theta| \leq h$, and Lemma 1 is used in the second one since $h \geq A_4(\epsilon/2)$. Furthermore,

$$\begin{aligned}\int_{-\infty}^{-h} \pi(\theta|\underset{\sim}{x}_n)d\theta + \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_n)d\theta &= \int_{-\infty}^{\infty} \pi(\theta|\underset{\sim}{x}_n)d\theta - \int_{-h}^{h} \pi(\theta|\underset{\sim}{x}_n)d\theta \\ &\leq 1 - (1 - \epsilon) = \epsilon\,.\end{aligned}$$

□

Choose any $\epsilon > 0$ and any $d > 0$. Define $h = A_4(\epsilon/2)$ and define $A_6(\epsilon, d) = \max[A_5(\epsilon/2, h), d + h]$, where the constant A_5 comes from the proof of result b) of Theorem 1. Consider that $\min[\underset{\sim}{\phi}] > A_6(\epsilon, d)$ and $j \in (k+1, \ldots, n)$. Since $|x_j| \in \underset{\sim}{\phi}$, it follows that $|x_j| > d + h$. Then, if $x_j > 0$ (that is for $j = m+1, \ldots, n$),

$$\begin{aligned}\Pr[|\theta - x_j| \leq d|\underset{\sim}{x}_n] &= \int_{x_j - d}^{x_j + d} \pi(\theta|\underset{\sim}{x}_n)d\theta \leq \int_{x_j - d}^{\infty} \pi(\theta|\underset{\sim}{x}_n)d\theta \\ &\leq \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_n)d\theta \leq \epsilon\,.\end{aligned}$$

Lemma 8 is used in the last inequality. The proof for $x_j < 0$ (that is for $j = k+1, \ldots, m$) is similar.

5.5. Proof of result d) *of Theorem* 1

The definition of convergence in law of a sequence of random variables $\{Y_s\}_{s=1,2,3,\ldots}$ to a random variable Y, as $s \to \infty$, is given as follows.

Definition 2. $Y_s \xrightarrow{\mathcal{L}} Y$ if $\lim_{s\to\infty} \Pr[Y_s \leq d] = \Pr[Y \leq d]$, for all d such that $\Pr[Y \leq d]$ is continuous.

In order to use this definition with $Y_s = \theta|\underset{\sim}{x}_n$ and $Y = \theta|\underset{\sim}{x}_k$, the prior location x_0 and the observations $x_1, \ldots, x_n$ are expressed as some functions $h_i(s)$ of the variable s, that is $x_i = h_i(s), i = 0, 1, \ldots, n$, for any functions $h_i(s)$ on $\mathbb{N}$ which satisfy

i) $h_i(s) = c_i$ for all $s \in \mathbb{N}$, where c_i is a constant, if $i = 1, \ldots, k$,
ii) $\lim_{s\to\infty} h_i(s) = -\infty$, if $i = k+1, \ldots, m$,
iii) $\lim_{s\to\infty} h_i(s) = \infty$ if $i = m+1, \ldots, n$.

The density of $Y_s = \theta|\underset{\sim}{x}_n$ evaluated at the point y is then given by

$$\pi(y|\underset{\sim}{x}_n) = \frac{\prod_{i=0}^{n} f_i(x_i - y)}{\int_{-\infty}^{\infty} \prod_{i=0}^{n} f_i(x_i - \theta)\, d\theta}$$

$$= \frac{\prod_{i=0}^{k} f_i(c_i - y) \prod_{i=k+1}^{n} f_i(h_i(s) - y)}{\int_{-\infty}^{\infty} \prod_{i=0}^{k} f_i(c_i - \theta) \prod_{i=k+1}^{n} f_i(h_i(s) - \theta)\, d\theta}$$

and the density of $Y = \theta|\underset{\sim}{x}_k$ evaluated at the point y is given by

$$\pi(y|\underset{\sim}{x}_k) = \frac{\prod_{i=0}^{k} f_i(c_i - y)}{\int_{-\infty}^{\infty} \prod_{i=0}^{k} f_i(c_i - \theta)\, d\theta}.$$

It can be seen that the functions $h_i(s)$ are defined such that $s \to \infty \Leftrightarrow \underset{\sim}{\phi} \to \infty$. Furthermore, it can be seen that the density of $Y = \theta|\underset{\sim}{x}_k$ does not depend on s or $\underset{\sim}{\phi}$. Then $Y_s \xrightarrow{\mathcal{L}} Y$ as $s \to \infty$ for any functions $h_i(s)$ which satisfy i), ii) and iii) $\Leftrightarrow \theta|\underset{\sim}{x}_n \xrightarrow{\mathcal{L}} \theta|\underset{\sim}{x}_k$ as $\underset{\sim}{\phi} \to \infty$ at any given rate.

According to Definition 2, the convergence in law for any functions $h_i(s)$ is obtained if $\lim_{s\to\infty} \Pr[Y_s \le d] = \Pr[Y \le d]$, for all d such that $\Pr[Y \le d]$ is continuous, or equivalently, if $\lim_{\underset{\sim}{\phi}\to\infty} \Pr[\theta \le d|\underset{\sim}{x}_n] = \Pr[\theta \le d|\underset{\sim}{x}_k]$, for all d such that $\Pr[\theta \le d|\underset{\sim}{x}_k]$ is continuous. Therefore, the result d) can be rewritten as follows: $\forall \epsilon > 0$ there exists a constant $A_7(\epsilon)$ such that $\min[\underset{\sim}{\phi}] > A_7(\epsilon)$ and $d \in \mathbb{R} \Rightarrow \left|\Pr[\theta \le d|\underset{\sim}{x}_n] - \Pr[\theta \le d|\underset{\sim}{x}_k]\right| \le \epsilon$.

Choose any $\epsilon > 0$, define $h = A_4(\epsilon/6)$ and $A_7(\epsilon) = A_5(\epsilon/6, h)$. The constants A_4 and A_5 respectively come from Lemma 1 and from the proof of result b) of Theorem 1. The real line is divided into three parts: $(-\infty, -h]$, $(-h, h]$ and (h, ∞), and consider that $\min[\underset{\sim}{\phi}] > A_7(\epsilon)$. First consider $d \le -h$.

$$\Pr[\theta \le d|\underset{\sim}{x}_n] \le \Pr[\theta \le -h|\underset{\sim}{x}_n] = \int_{-\infty}^{-h} \pi(\theta|\underset{\sim}{x}_n) d\theta \le \epsilon/3\,.$$

Lemma 8 is used in the last inequality since $h = A_4(\epsilon/6)$ and $\min[\underset{\sim}{\phi}] > A_5(\epsilon/6, h)$. In the same way, it can be shown, using Lemma 1, that $\Pr[\theta \le d|\underset{\sim}{x}_k] \le \epsilon/6$, since $h = A_4(\epsilon/6)$. From this result and from $\Pr[\theta \le d|\underset{\sim}{x}_n] \le \epsilon/3$, it follows that $\left|\Pr[\theta \le d|\underset{\sim}{x}_n] - \Pr[\theta \le d|\underset{\sim}{x}_k]\right| \le \epsilon/3 < \epsilon$. Now consider $-h < d \le h$.

$$\begin{aligned}
&\left|\Pr[-h < \theta \le d|\underset{\sim}{x}_n] - \Pr[-h < \theta \le d|\underset{\sim}{x}_k]\right| \le \int_{-h}^{d} \left|\pi(\theta|\underset{\sim}{x}_n) - \pi(\theta|\underset{\sim}{x}_k)\right| d\theta \\
&= \int_{-h}^{d} \pi(\theta|\underset{\sim}{x}_k) \left|\frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} - 1\right| d\theta \le \epsilon/6 \int_{-h}^{d} \pi(\theta|\underset{\sim}{x}_k) d\theta \le \epsilon/6\,.
\end{aligned}$$

Result b) of Theorem 1 is used in the second inequality since $\min[\underset{\sim}{\phi}] > A_5(\epsilon/6, h)$ and $|\theta| \le h$. Therefore,

$$\begin{aligned}
&\left|\Pr[\theta \le d|\underset{\sim}{x}_n] - \Pr[\theta \le d|\underset{\sim}{x}_k]\right| \le \left|\Pr[\theta \le -h|\underset{\sim}{x}_n] - \Pr[\theta \le -h|\underset{\sim}{x}_k]\right| \\
&\quad + \left|\Pr[-h < \theta \le d|\underset{\sim}{x}_n] - \Pr[-h < \theta \le d|\underset{\sim}{x}_k]\right| \\
&\le \epsilon/3 + \epsilon/6 = \epsilon/2 < \epsilon\,.
\end{aligned}$$

Finally consider $d > h$.

$$\Pr[h < \theta \le d|\underset{\sim}{x}_n] \le \Pr[\theta > h|\underset{\sim}{x}_n] = \int_{h}^{\infty} \pi(\theta|\underset{\sim}{x}_n) d\theta \le \epsilon/3\,.$$

Lemma 8 is used in the last inequality since $h = A_4(\epsilon/6)$ and $\min[\underset{\sim}{\phi}] > A_5(\epsilon/6, h)$. In the same way, it can be shown, using Lemma 1, that $\Pr[h < \theta \leq d|\underset{\sim}{x}_k] \leq \epsilon/6$, since $h = A_4(\epsilon/6)$. From this result and from $\Pr[h < \theta \leq d|\underset{\sim}{x}_n] \leq \epsilon/3$, it follows that $\left|\Pr[h < \theta \leq d|\underset{\sim}{x}_n] - \Pr[h < \theta \leq d|\underset{\sim}{x}_k]\right| \leq \epsilon/3$. Finally, from this result and from $\left|\Pr[\theta \leq h|\underset{\sim}{x}_n] - \Pr[\theta \leq h|\underset{\sim}{x}_k]\right| \leq \epsilon/2$, it follows that $\left|\Pr[\theta \leq d|\underset{\sim}{x}_n] - \Pr[\theta \leq d|\underset{\sim}{x}_k]\right| \leq \epsilon/2 + \epsilon/3 < \epsilon$.

5.6. Proof of result e) *of Theorem* 1

The condition $\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] < \infty$ can be rewritten as follows: there exists a constant M_1 such that $\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] < M_1$. First we introduce two lemmas needed for the proof. Considering that $0 \leq \int_{-h}^{h} |w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)d\theta < M_1$ and that $|w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)$ depends only on the fixed values $x_0, x_1, \ldots, x_k$, it is then possible to show the following lemma.

Lemma 9. $\forall\epsilon > 0$, *there exists a constant* $A_9(\epsilon) > 0$ *such that* $h \geq A_9(\epsilon) \Rightarrow$

$$\int_{-\infty}^{-h} |w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)d\theta \leq \epsilon \quad \textit{and} \quad \int_{h}^{\infty} |w(\theta)|\,\pi(\theta|\underset{\sim}{x}_k)d\theta \leq \epsilon\,.$$

Another lemma is needed and its proof is given in Section 5.7.

Lemma 10. $\forall\epsilon > 0$, *there exists a constant* $A_8(\epsilon)$ *such that* $\min[\underset{\sim}{\phi}] > A_8(\epsilon) \Rightarrow$

$$\left|\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|] - \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|]\right| < \epsilon\,.$$

Lemma 10 is similar to the result e) of Theorem 1, except that it considers the absolute value of $w(\theta)$. Consider now the result e) of Theorem 1, which can be rewritten as follows: $\forall\epsilon > 0$, there exists a constant $A_{10}(\epsilon)$ such that $\min[\underset{\sim}{\phi}] > A_{10}(\epsilon) \Rightarrow \left|\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[w(\theta)] - \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[w(\theta)]\right| < \epsilon$. Choose any $\epsilon > 0$. Define $\epsilon_0 = \epsilon/6$, $h = A_9(\epsilon_0)$ and $A_{10}(\epsilon) = \max[A_5(\epsilon_0/M_1, h), A_8(\epsilon_0)]$, where the constant $A_5(\epsilon_0/M_1, h)$ comes from the proof of result b) of Theorem 1, which was rewritten as follows: $\forall\epsilon > 0, \forall h > 0$, there exists a constant $A_5(\epsilon, h)$

such that $\min[\underset{\sim}{\phi}] > A_5(\epsilon, h)$ and $|\theta| \le h \Rightarrow \left|\frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} - 1\right| \le \epsilon$. Then

$$\begin{aligned}
&\left|\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[w(\theta)] - \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[w(\theta)]\right| \\
&= \left|\int_{-\infty}^{\infty} w(\theta)\pi(\theta|\underset{\sim}{x}_n)d\theta - \int_{-\infty}^{\infty} w(\theta)\pi(\theta|\underset{\sim}{x}_k)d\theta\right| \\
&\le \int_{-\infty}^{-h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_n)d\theta + \int_{h}^{\infty} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_n)d\theta + \int_{-\infty}^{-h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta \\
&\quad + \int_{h}^{\infty} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta + \int_{-h}^{h} |w(\theta)| \left|\pi(\theta|\underset{\sim}{x}_n) - \pi(\theta|\underset{\sim}{x}_k)\right| d\theta \\
&= \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|] - \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] + \int_{-\infty}^{-h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta \\
&\quad + \int_{h}^{\infty} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta + \int_{-h}^{h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta - \int_{-h}^{h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_n)d\theta \\
&\quad + \int_{-\infty}^{-h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta + \int_{h}^{\infty} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta \\
&\quad + \int_{-h}^{h} |w(\theta)| \left|\pi(\theta|\underset{\sim}{x}_n) - \pi(\theta|\underset{\sim}{x}_k)\right| d\theta \\
&\le \left|\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|] - \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|]\right| + 2\int_{-h}^{h} |w(\theta)| \left|\pi(\theta|\underset{\sim}{x}_n) - \pi(\theta|\underset{\sim}{x}_k)\right| d\theta \\
&\quad + 2\int_{-\infty}^{-h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta + 2\int_{h}^{\infty} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta \\
&\le 6\epsilon_0 = \epsilon\,.
\end{aligned}$$

Lemma 10 is used in the last inequality since $A_{10}(\epsilon) \ge A_8(\epsilon_0)$ and Lemma 9 is used since $h = A_9(\epsilon_0)$. Furthermore

$$\begin{aligned}
&\int_{-h}^{h} |w(\theta)| \left|\pi(\theta|\underset{\sim}{x}_n) - \pi(\theta|\underset{\sim}{x}_k)\right| d\theta = \int_{-h}^{h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k) \left|\frac{\pi(\theta|\underset{\sim}{x}_n)}{\pi(\theta|\underset{\sim}{x}_k)} - 1\right| d\theta \\
&\le \frac{\epsilon_0}{M_1} \int_{-h}^{h} |w(\theta)|\, \pi(\theta|\underset{\sim}{x}_k)d\theta \le \frac{\epsilon_0}{M_1} \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|] \le \epsilon_0\,.
\end{aligned}$$

Result b) of Theorem 1 is used since $A_{10}(\epsilon) \ge A_5(\epsilon_0/M_1, h)$.

5.7. Proof of Lemma 10

We want to show that $\lim_{\underset{\sim}{\phi}\to\infty} \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|] = \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|]$, or equivalently $\lim_{\underset{\sim}{\phi}\to\infty} \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|+1] = \mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|+1]$. Define the density

$$\pi^*(\theta|\underset{\sim}{x}_k) = \frac{(|w(\theta)|+1)\pi(\theta|\underset{\sim}{x}_k)}{\displaystyle\int_{-\infty}^{\infty}(|w(\theta)|+1)\pi(\theta|\underset{\sim}{x}_k)d\theta}$$

$$= \frac{(|w(\theta)|+1)\displaystyle\prod_{i=0}^{k} f_i(x_i-\theta)}{\displaystyle\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{k} f_i(x_i-\theta)d\theta}.$$

Since $\int_{-\infty}^{\infty}|w(\theta)|\pi(\theta|\underset{\sim}{x}_k)d\theta < \infty$ and $|w(\theta)|\pi(\theta|\underset{\sim}{x}_k) < \infty$ as given in the conditions of Theorem 1, it easy to see that $\pi^*(\theta|\underset{\sim}{x}_k)$ is a proper density, positive everywhere and bounded above. We can also define the positive and bounded density $m^*(\underset{\sim}{x}_k) = \int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{k} f_i(x_i-\theta)d\theta$.

It should be noted that all the information concerning the densities f_{j_0}, $f_{j_1}, \ldots, f_{j_k}$ ($f_0, f_1, \ldots, f_k$ for the proof, without loss of generality) in Theorem 1 is given through the densities $\pi(\theta|\underset{\sim}{x}_k)$ and $m(\underset{\sim}{x}_k)$, where $\pi(\theta|\underset{\sim}{x}_k)$ is proper, positive everywhere and bounded.

It is then possible to use the result a) of Theorem 1 using the densities $\pi^*(\theta|\underset{\sim}{x}_k)$ and $m^*(\underset{\sim}{x}_k)$. If the conditions given by equations (1) and (2), given by $\lim_{\theta\to-\infty} \frac{m(\underset{\sim}{x}_k)\pi(\theta|\underset{\sim}{x}_k)}{\prod_{i=k+1}^{m} f_{j_i}(\theta)} = 0$ and $\lim_{\theta\to\infty} \frac{m(\underset{\sim}{x}_k)\pi(\theta|\underset{\sim}{x}_k)}{\prod_{i=m+1}^{n} f_{j_i}(\theta)} = 0$, are used with $\pi^*(\theta|\underset{\sim}{x}_k)$ instead of $\pi(\theta|\underset{\sim}{x}_k)$ and $m^*(\underset{\sim}{x}_k)$ instead of $m(\underset{\sim}{x}_k)$, it can be seen that they are equivalent to the conditions given by equations (1) to (4) using $\pi(\theta|\underset{\sim}{x}_k)$ and $m(\underset{\sim}{x}_k)$. Result a) using $m(\underset{\sim}{x}_k)$ and $\pi(\theta|\underset{\sim}{x}_k)$ is given by

$$\lim_{\underset{\sim}{\phi}\to\infty} \frac{m(\underset{\sim}{x}_k)\displaystyle\prod_{i=k+1}^{n} f_i(x_i)}{\displaystyle\int_{-\infty}^{\infty} m(\underset{\sim}{x}_k)\pi(\theta|\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i-\theta)d\theta} = 1\,,$$

and result a) using $m^*(\underset{\sim}{x}_k)$ and $\pi^*(\theta|\underset{\sim}{x}_k)$ is given by

$$\lim_{\underset{\sim}{\phi}\to\infty} \frac{m^*(\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i)}{\int_{-\infty}^{\infty} m^*(\underset{\sim}{x}_k)\pi^*(\theta|\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i-\theta)d\theta} = 1$$

$$\Leftrightarrow \lim_{\underset{\sim}{\phi}\to\infty} \frac{\left(\prod_{i=k+1}^{n} f_i(x_i)\right)\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{k} f_i(x_i-\theta)d\theta}{\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{n} f_i(x_i-\theta)d\theta} = 1\,.$$

The result can now be shown.

$$\begin{aligned}\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_n)}[|w(\theta)|+1] &= \frac{\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{n} f_i(x_i-\theta)d\theta}{m(\underset{\sim}{x}_n)} \\ &= \frac{\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{n} f_i(x_i-\theta)d\theta}{\left(\prod_{i=k+1}^{n} f_i(x_i)\right)\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{k} f_i(x_i-\theta)d\theta} \\ &\quad\times \frac{m(\underset{\sim}{x}_k)\prod_{i=k+1}^{n} f_i(x_i)}{m(\underset{\sim}{x}_n)} \\ &\quad\times \frac{\int_{-\infty}^{\infty}(|w(\theta)|+1)\prod_{i=0}^{k} f_i(x_i-\theta)d\theta}{m(\underset{\sim}{x}_k)}\,.\end{aligned}$$

If the limit as $\underset{\sim}{\phi}\to\infty$ is taken, the first two terms in the last expression are 1 according to result a) using respectively $\pi^*(\theta|\underset{\sim}{x}_k)$ and $\pi(\theta|\underset{\sim}{x}_k)$. The last term is $\mathbb{E}^{\pi(\theta|\underset{\sim}{x}_k)}[|w(\theta)|+1]$, which proves the result.

5.8. *Proof of Proposition* 2

Firstly we show that condition C1 is satisfied on the right tail of f. Condition C1 can be rewritten as follows. For any constant $h>0$ and for all θ such that $|\theta| \leq h$, $\lim_{z\to\infty} \frac{f(z+\theta)}{f(z)} = 1$. Furthermore, since p-cred$^+(f) = (\gamma, \delta, \alpha, \beta)$, there exists a constant $K_1 > 0$ such that $\lim_{z\to\infty} \frac{f(z)}{e^{-\delta z^\gamma} z^{-\alpha} \log^{-\beta} z} = K_1$. If $h > 0$ and $|\theta| \leq h$, then

$$
\begin{aligned}
\lim_{z\to\infty} \frac{f(z+\theta)}{f(z)} &= \lim_{z\to\infty} \frac{f(z+\theta)}{e^{-\delta(z+\theta)^\gamma}(z+\theta)^{-\alpha}\log^{-\beta}(z+\theta)} \times \frac{e^{-\delta z^\gamma} z^{-\alpha}\log^{-\beta} z}{f(z)} \\
&\quad \times \frac{e^{-\delta(z+\theta)^\gamma}(z+\theta)^{-\alpha}\log^{-\beta}(z+\theta)}{e^{-\delta z^\gamma} z^{-\alpha}\log^{-\beta} z} \\
&= \lim_{z\to\infty} \frac{K_1}{K_1} \frac{e^{-\delta(z+\theta)^\gamma}(z+\theta)^{-\alpha}\log^{-\beta}(z+\theta)}{e^{-\delta z^\gamma} z^{-\alpha}\log^{-\beta} z} \\
&= \lim_{z\to\infty} e^{-\delta((z+\theta)^\gamma - z^\gamma)} \left(\frac{z}{z+\theta}\right)^\alpha \left(\frac{\log z}{\log(z+\theta)}\right)^\beta = 1\,.
\end{aligned}
$$

It is easy to check in the last equality that the last two terms tend to 1 as z tends to infinity. Furthermore, using the Taylor series development of $(z+\theta)^\gamma - z^\gamma$, it can be shown that the last expression tends to 1 as z tends to infinity if and only if $\gamma < 1$.

Secondly, we show that conditions C2 and C3 are satisfied on the right tail of f. Define $f^*(z)$ as a GEP density with the same left and right p-credence as $f(z)$, that is

$$
f^*(z) = \begin{cases} K_2 e^{-\delta|z|^\gamma} |z|^{-\alpha} \log^{-\beta}|z|, & \text{if } |z| > z_0 \ ; \\ K_2 e^{-\delta z_0^\gamma} z_0^{-\alpha} \log^{-\beta} z_0, & \text{if } |z| \leq z_0 \ , \end{cases}
$$

with any $z_0 > 1$, where K_2 is the normalizing constant. The tails behavior of f and f^* are the same and both are proper densities. Define

$$
g(z) = \begin{cases} (1+|z|)^{-3}; & \text{if } \gamma > 0, \delta > 0\,, \\ f^*(z); & \text{if } \gamma = 0, \delta = 0\,. \end{cases}
$$

The density g is proper since $(1+|z|)^{-3}$ and f^* are also proper densities.

Consider the first case, when $0 < \gamma < 1, \delta > 0$ and $g(z) = (1+|z|)^{-3}$. Then

$$
\begin{aligned}
\lim_{z\to\infty} \frac{f^2(z/2)}{f(z)g(z/2)} &= \lim_{z\to\infty} \frac{K_1^2}{K_1} \frac{\left(e^{-\delta(z/2)^\gamma}(z/2)^{-\alpha}\log^{-\beta}(z/2)\right)^2}{\left(e^{-\delta z^\gamma} z^{-\alpha}\log^{-\beta} z\right)\,(1+z/2)^{-3}} \\
&= \lim_{z\to\infty} K_1 e^{-\delta(2^{1-\gamma}-1)z^\gamma}(z/4)^{-\alpha}\left(\frac{\log^2(z/2)}{\log z}\right)^{-\beta}(1+z/2)^3 \\
&= 0\,.
\end{aligned}
$$

The dominant term is the exponential one and it tends to 0 as $z \to \infty$ since $\gamma > 0, \delta > 0$ and $2^{1-\gamma} - 1 > 0 \Leftrightarrow \gamma < 1$. It is sufficient to show that condition C2 is satisfied since f and g are both positive everywhere and bounded above, with monotonous tails.

Furthermore, if $z > z_0$, it can be shown that

$$\begin{aligned}\frac{d^2}{dz^2}\log f^*(z) &= \frac{d^2}{dz^2}\log\left[K_2 e^{-\delta z^\gamma} z^{-\alpha}\log^{-\beta} z\right]\\ &= \frac{d^2}{dz^2}\left[-\delta z^\gamma - \alpha\log z - \beta\log(\log z)\right]\\ &= \gamma(1-\gamma)\delta z^{\gamma-2} + \frac{\alpha}{z^2} + \frac{\beta(\log z + 1)}{z^2\log^2 z}\\ &= \frac{1}{z^2}\left[\gamma(1-\gamma)\delta z^\gamma + \alpha + \frac{\beta}{\log z} + \frac{\beta}{\log^2 z}\right].\end{aligned}$$

It can be shown that $\frac{d^2}{dz^2}\log g(z) = \frac{3}{(1+|z|)^2} > 0$ for any value of z. Finally, if $z > z_0$,

$$\begin{aligned}&\frac{d^2}{dz^2}\log f^*(z) - \frac{d^2}{dz^2}\log g(z)\\ &= \frac{1}{z^2}\left[\gamma(1-\gamma)\delta z^\gamma + \alpha + \frac{\beta}{\log z} + \frac{\beta}{\log^2 z} - \frac{3z^2}{(1+z)^2}\right].\end{aligned}$$

The term in brackets goes to $+\infty$ as $z \to \infty$ if $\gamma(1-\gamma)\delta > 0$, that is if $0 < \gamma < 1$ and $\delta > 0$, which shows that $\frac{d^2}{dz^2}\log f^*(z) - \frac{d^2}{dz^2}\log g(z) \geq 0$ if z is large enough. Then conditions C2 and C3 are satisfied if $\gamma > 0, \delta > 0$.

Consider now the second case, when $\gamma = 0, \delta = 0$ and $g(z) = f^*(z)$.

$$\begin{aligned}&\lim_{z\to\infty}\frac{f^2(z/2)}{f(z)g(z/2)} = \lim_{z\to\infty}\frac{f^2(z/2)}{f(z)f^*(z/2)}\\ &= \lim_{z\to\infty}\frac{f(z/2)}{f(z)}\frac{K_1 e^{-\delta(z/2)^\gamma}(z/2)^{-\alpha}\log^{-\beta}(z/2)}{K_2 e^{-\delta(z/2)^\gamma}(z/2)^{-\alpha}\log^{-\beta}(z/2)} = \lim_{z\to\infty}\frac{K_1 f(z/2)}{K_2 f(z)}\\ &= \lim_{z\to\infty}\frac{K_1}{K_2}\frac{(z/2)^{-\alpha}\log^{-\beta}(z/2)}{z^{-\alpha}\log^{-\beta} z} = \lim_{z\to\infty}\frac{K_1}{K_2}2^\alpha\left(\frac{\log(z/2)}{\log z}\right)^{-\beta} = \frac{K_1}{K_2}2^\alpha.\end{aligned}$$

Furthermore, if $z > z_0$, it can be shown that

$$\frac{d^2}{dz^2}\log f^*(z) = \frac{1}{z^2}\left[\alpha + \frac{\beta}{\log z} + \frac{\beta}{\log^2 z}\right].$$

The term in brackets converge to α as $z \to \infty$. Since f^* is a proper density and $\gamma = \delta = 0$, it follows that $\alpha \geq 1$, which shows that $\frac{d^2}{dz^2} \log f^*(z) = \frac{d^2}{dz^2} \log g(z) \geq 0$ if z is large enough. Then conditions C2 and C3 are also satisfied if $\gamma = 0, \delta = 0$.

ACKNOWLEDGMENTS

The authors are thankful to the NSERC (Natural Sciences and Engineering Research Council of Canada), the FQRNT (Le Fonds québécois de la recherche sur la nature et les technologies) and the SOA (Society of Actuaries) for their financial support.

REFERENCES

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

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

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

DAWID, A. P. (1973) Posterior expectations for large observations, *Biometrika*, 60, 664–667.

DE FINETTI, B. (1961) The Bayesian approach to the rejection of outliers, *Proceedings of the Fourth Berkeley Symposium on Probability and Statistics*, (Vol. 1), Berkeley, University of California Press, 199–210.

DESGAGNÉ, A. and ANGERS, J.-F. (2003) *Computational aspect of the generalized exponential power density*, Technical report CRM-2918, Université de Montréal (http://www.crm.umontreal.ca/pub/Rapports/2900-2999/ 2918.pdf).

DESGAGNÉ, A. and ANGERS, J.-F. (2005) Importance Sampling with the Generalized Exponential Power Density, *Statistics and Computing*, 15, 189–195.

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.

HILL, B. M. (1974) On coherence, inadmissibility and inference about many parameters in the theory of least squares, in: *Studies, Bayesian Econometrics and Statistics*, S. E. Fienberg and A. Zellner (eds.), Amsterdam, North-Holland, 555–584.

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

O'HAGAN, A. (1979) On outlier rejection phenomena in Bayes inference, *Journal of the Royal Statistical Society, Ser. B*, 41, 358–367.

O'Hagan, A. (1988) Modelling with heavy tails, in: *Bayesian Statistic III*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith (eds.), Oxford, Clarendon Press, 345–359.

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

ALAIN DESGAGNÉ
Département de mathématiques
Université du Québec à Montréal
C.P. 8888, Succ. Centre-Ville
Montréal, Québec (Canada)
H3C 3P8
desgagne.alain@uqam.ca

JEAN-FRANÇOIS ANGERS
Département de mathématiques
et de statistique
Université de Montréal
C.P. 6128, Succ. Centre-Ville
Montréal, Québec (Canada)
H3C 3J7
angers@dms.umontreal.ca