Part II

CONTRIBUTED PAPERS

Bayesian Robustness
IMS Lecture Notes - Monograph Series (1996) Volume 29

PROTECTION AGAINST OUTLIERS USING A SYMMETRIC STABLE LAW PRIOR[1]

BY JEAN-FRANÇOIS ANGERS

Université de Montréal

Estimation of the expectation of a multivariate normal random vector is considered. The components of the mean vector are assumed to be exchangeable. This information is modeled using a hierarchical prior with independent symmetric stable law at the first stage level. It is shown that the first stage Bayes estimator has an analytic expression and that it is robust with respect to the presence of outlying observations.

1. Introduction. Let $\vec{X} = (X_1, X_2, \ldots, X_p)^t$ have a p-variate normal distribution with mean vector $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^t$ and covariance matrix $\sigma^2 \mathbf{I}_p$, where σ^2 is assumed to be known. The components of $\vec{\theta}$ are believed to be exchangeable, and hence "shrinkage" estimation of them is desired.

The assumption of exchangeability of the components of $\vec{\theta}$ can be modeled easily in a hierarchical Bayesian fashion with a two stage prior. The use of hierarchical models provides a way to consider more complex models since it allows the location and the structural information to be modeled separately. In this paper, the first stage prior is of the form $\pi_1(\vec{\theta}|\mu, \tau) = \prod_{j=1}^{p} \pi_{1,\alpha}(\theta_j|\mu, \tau)$ where $\pi_{1,\alpha}$ denotes the density of a symmetric stable law with parameters α, μ and τ. The hyperparameter α controls the tail behavior of the prior and it is assumed to be known. (As it will be seen at the end of the third section, α does not need to be specified very accurately.) The hyperparameters μ and τ represent the common location and scale parameters of $\pi_{1,\alpha}$ and their prior density is denoted by $\pi_2(\mu, \tau)$. (Applications to finance of the symmetric stable law are discussed in Press (1975).)

In Angers and Berger (1991), a generalized Bayes estimator was developed for the special case $\alpha = 1$. In Angers (1992), a similar problem was considered but a Student-t density with $2k + 1$ ($k \in N$) degrees of freedom was used as the first stage prior.

The goal of this paper is to present an estimator which is insensitive to the presence of outlying coordinates, that is, one bad data point will not overly influence all estimates. Furthermore, this estimator can be written in an analytic form involving only hypergeometric functions. In Section 2, the hierarchical Bayesian formulation of the problem is presented and the proposed robust estimator is derived. In Section 3, its robust behavior is studied. In the last section, a numerical example comparing the estimator for several values of α is discussed.

---

[1] *AMS 1991 Subject Classifications:* Primary 62F35; secondary 62F15, 62E25
Key words and phrases: Robust estimator, symmetric stable law, outliers.

2. The robust hierarchical Bayes estimator. Let $\vec{X} \sim N_p(\vec{\theta}, \sigma^2 \mathbf{I}_p)$ with mean vector $\vec{\theta}$ and covariance matrix $\sigma^2 \mathbf{I}_p$, where σ^2 is assumed to be known. The components of $\vec{\theta}$ are believed to be exchangeable which is typically modeled via a two-stage prior:

$$\vec{\theta}|\mu, \tau \sim \pi_1(\vec{\theta}|\mu, \tau) = \prod_{j=1}^{p} \pi_{1,\alpha}(\theta_j|\mu, \tau) \text{ and } (\mu, \tau) \sim \pi_2(\mu, \tau).$$

2.1. Preliminaries. Let $\widehat{\pi}_{1,\alpha}$ be the Fourier transform of the first stage prior density $\pi_{1,\alpha}(\theta_j|0,1)$ and $\widehat{f}^{(l)}$ be the l^{th} derivative of the Fourier transform of the likelihood function with $\theta_j = 0\,\forall j$ and $\sigma^2 = 1$. It can be shown that, under the squared error loss (*cf.* Angers, 1995) we have:

$$\begin{aligned}
m(x_j|\mu,\tau) &= \text{the first stage marginal density of } x_j \\
&= \frac{1}{\sigma}\mathcal{I}_0\left(\tau/\sigma, [x_j-\mu]/\sigma\right), \\
\widehat{\theta}_j(x_j|\mu,\tau) &= \text{the first stage Bayes estimator of } \theta_j \\
&= x_j - \frac{i}{2\pi}\frac{\mathcal{I}_1\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}{\mathcal{I}_0\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}(x_j-\mu), \\
\rho(\pi,\widehat{\theta}_j|x_j,\mu,\tau) &= \text{the posterior expected loss of } \widehat{\theta}_j(x_j|\mu,\tau) \\
&= \frac{\sigma^2}{4\pi^2}\left[\left(\frac{\mathcal{I}_1\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}{\mathcal{I}_0\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}\frac{[x_j-\mu]}{\sigma}\right)^2\right. \\
&\qquad \left. - \frac{\mathcal{I}_2\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}{\mathcal{I}_0\left(\tau/\sigma,[x_j-\mu]/\sigma\right)}\right],
\end{aligned}$$

where $i = \sqrt{-1}$, $\mathcal{I}_l(a,b) = \int_{-\infty}^{\infty} \widehat{f}^{(l)}(s)\widehat{\pi}_1(as)e^{i2\pi bs}ds$ for $l = 0,1,2$.

Hence, the hierarchical Bayes estimator of θ_j is given by

$$(1)\ \widehat{\theta}_j = x_j - \frac{i}{2\pi} \times \frac{\int_0^\infty\int_{-\infty}^{\infty}[x_j-\mu]\mathcal{I}_1\left(\frac{\tau}{\sigma},\frac{x_j-\mu}{\sigma}\right)\left[\prod_{l\neq j}\mathcal{I}_0\left(\frac{\tau}{\sigma},\frac{x_l-\mu}{\sigma}\right)\right]\pi_2(\mu,\tau)d\mu d\tau}{\int_0^\infty\int_{-\infty}^{\infty}\left[\prod_{l=1}^{p}\mathcal{I}_0\left(\frac{\tau}{\sigma},\frac{x_l-\mu}{\sigma}\right)\right]\pi_2(\mu,\tau)d\mu d\tau},$$

and its posterior expected loss is given by

$$\begin{aligned}
\rho(\pi,\widehat{\theta}_j|\vec{x}) &= -\frac{\sigma^2}{4\pi^2}\frac{\int_0^\infty\int_{-\infty}^{\infty}\mathcal{I}_2\left(\frac{\tau}{\sigma},\frac{x_j-\mu}{\sigma}\right)\left[\prod_{l\neq j}\mathcal{I}_0\left(\frac{\tau}{\sigma},\frac{x_l-\mu}{\sigma}\right)\right]\pi_2(\mu,\tau)d\mu d\tau}{\int_0^\infty\int_{-\infty}^{\infty}\left[\prod_{l=1}^{p}\mathcal{I}_0\left(\frac{\tau}{\sigma},\frac{x_l-\mu}{\sigma}\right)\right]\pi_2(\mu,\tau)d\mu d\tau} \\
&\quad -(x_j-\widehat{\theta}_j(\vec{x}))^2.
\end{aligned}$$

In this paper, we consider the choice of the independent symmetric stable law, denoted by $S_\alpha(\mu,\tau)$, as the first stage prior. This model was chosen because of its range of tail behavior (from Cauchy to normal tails) and also because the first stage Bayes estimator of θ_j has an analytic expression. Except for $\alpha = 1$ (Cauchy density) and $\alpha = 2$ (normal density), the density of $S_\alpha(\mu, \tau)$ can not be written in closed form, but it is known (*cf.* Feller, 1966) that the Fourier transform of $\pi_{1,\alpha}(\theta)$ is given by $\widehat{\pi}_{1,\alpha}(s) = \exp\{-(2\pi|s|)^\alpha\}$.

2.2. The first stage Bayes estimator.

THEOREM 1 *Suppose that $X \sim N(\theta, \sigma^2)$ and $\theta \sim S_\alpha(\mu, \tau)$ where σ, μ, τ and α are known. Then,*

$$\begin{aligned}
\mathcal{I}_0(a,b) &= \frac{1}{\sqrt{2\pi}}\mathrm{G}_0(a,b),\\
\mathcal{I}_1(a,b) &= -i2\sqrt{2}b\mathrm{G}_1(a,b),\\
\mathcal{I}_2(a,b) &= -2\sqrt{2}\pi\left[\mathrm{G}_0(a,b) - 2\mathrm{G}_2(a,b)\right],\\
\text{where}\quad \mathrm{G}_j(a,b) &= \mathrm{e}^{-b^2/2}\sum_{l=0}^{\infty}(-1)^l\frac{\Gamma(\frac{l\alpha+j+k+1}{2})}{l!}(\sqrt{2}a)^{l\alpha}\\
&\quad \times {}_1\mathrm{F}_1\left(-\frac{[l\alpha+j-k]}{2}, k+\frac{1}{2}, \frac{b^2}{2}\right),
\end{aligned}$$

${}_1\mathrm{F}_1(\gamma, \delta, z)$ denotes the degenerate hypergeometric function (cf. Gradishteyn and Ryzhik, 1980) and where $k = 0$ if j is even and 1 otherwise.

PROOF 1 *We know that $\widehat{f}(s) = \mathrm{e}^{-2\pi^2 s^2}$ and $\widehat{\pi}_{1,\alpha}(s) = \mathrm{e}^{-(2\pi|s|)^\alpha}$. Hence $\widehat{f}^{(1)}(s) = -4\pi^2 s\widehat{f}(s)$ and $\widehat{f}^{(2)}(s) = -4\pi^2[1 - 4\pi^2 s^2]\widehat{f}(s)$. Consequently, in order to compute the $\mathcal{I}_l$'s, we need to evaluate*

$$\mathrm{h}_j(a,b) = \int_{-\infty}^{\infty} s^j\widehat{f}(s)\widehat{\pi}_1(as)\mathrm{e}^{i2\pi bs}\mathrm{d}s.$$

Using some algebra and special functions, we can show that

$$\begin{aligned}
\mathrm{h}_j(a,b) &= \tfrac{\mathrm{e}^{-b^2/2}}{(\sqrt{2\pi})^{j+1}}\textstyle\sum_{l=0}^{\infty}(-1)^l\frac{\Gamma([l\alpha+j+1]/2)}{l!}(\sqrt{2}a)^{l\alpha}\\
&\quad \times {}_1\mathrm{F}_1\left(-\tfrac{[l\alpha+j]}{2}, \tfrac{1}{2}, \tfrac{b^2}{2}\right)\\
&= \tfrac{1}{(\sqrt{2\pi})^{j+1}}\mathrm{G}_j(a,b), \ \textit{if } j \textit{ is even}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{h}_j(a,b) &= \tfrac{i\sqrt{2}b\mathrm{e}^{-b^2/2}}{(\sqrt{2\pi})^{j+1}}\textstyle\sum_{l=0}^{\infty}(-1)^l\frac{\Gamma([l\alpha+j+2]/2)}{l!}(\sqrt{2}a)^{l\alpha}\\
&\quad \times {}_1\mathrm{F}_1\left(-\tfrac{[l\alpha+j-1]}{2}, \tfrac{3}{2}, \tfrac{b^2}{2}\right)\\
&= \tfrac{i\sqrt{2}b}{(\sqrt{2\pi})^{j+1}}\mathrm{G}_j(a,b), \ \textit{if } j \textit{ is odd.}
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathcal{I}_0(a,b) &= \mathrm{h}_0(a,b) = \tfrac{1}{\sqrt{2\pi}}\mathrm{G}_0(a,b),\\
\mathcal{I}_1(a,b) &= -4\pi^2\mathrm{h}_1(a,b) = -i2\sqrt{2}b\mathrm{G}_1(a,b),\\
\mathcal{I}_2(a,b) &= -4\pi^2\left[\mathrm{h}_0(a,b) - 4\pi^2\mathrm{h}_2(a,b)\right]\\
&= -2\sqrt{2}\pi\left[\mathrm{G}_0(a,b) - 2\mathrm{G}_2(a,b)\right].
\end{aligned}$$

THEOREM 2 *If $X_j \sim N(\theta_j, \sigma^2)$ and $\theta_j \sim S_\alpha(\mu, \tau)$ independently for $j = 1, \ldots, p$ where σ, μ, τ and α are known and if $L(\vec{\theta}, \widehat{\vec{\theta}}) = \sum_{j=1}^{p}(\theta_j - \widehat{\theta}_j)^2$, then*

$$
\begin{aligned}
m(x_j|\mu,\tau) &= \frac{1}{\sqrt{2\pi}\sigma}\mathrm{G}_0(\tau/\sigma, [x_j-\mu]/\sigma), \\
\widehat{\theta}_j(x_j|\mu,\tau) &= x_j - 2\frac{\mathrm{G}_1(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}(x_j-\mu), \\
\rho(\pi,\widehat{\theta}_j|\mu,\tau) &= \sigma^2\left[1 - 2\frac{\mathrm{G}_2(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}\right. \\
&\quad \left. - \left(2\frac{\mathrm{G}_1(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}\frac{[x_j-\mu]}{\sigma}\right)^2\right].
\end{aligned}
\tag{2}
$$

2.3. The hierarchical Bayes estimators. From equation (1), it is clear that $\widehat{\theta}_j$ is obtained by integrating equation (2) with respect to

$$
\pi_2(\mu,\tau|\vec{x}) = \frac{\left[\prod_{j=1}^{p}\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)\right]\pi_2(\mu,\tau)}{\int_0^\infty\int_{-\infty}^{\infty}\left[\prod_{j=1}^{p}\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)\right]\pi_2(\mu,\tau)\mathrm{d}\mu\mathrm{d}\tau}. \tag{3}
$$

THEOREM 3 *Under the model given by Theorem 2 and if $\pi_2(\mu,\tau)$ is the prior density of (μ,τ), then, provided that all the integrals exist,*

$$
\begin{aligned}
m(\vec{x}) &= \frac{1}{(\sqrt{2\pi}\sigma)^p}\int_0^\infty\int_{-\infty}^{\infty}\left[\prod_{j=1}^{p}\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)\right]\pi_2(\mu,\tau)\mathrm{d}\mu\mathrm{d}\tau, \\
\widehat{\theta}_j(\vec{x}) &= x_j - 2E^{\pi_2(\mu,\tau|\vec{x})}\left[\frac{\mathrm{G}_1(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}(x_j-\mu)\right], \\
\rho(\pi,\widehat{\theta}_j|\vec{x}) &= \sigma^2 - 2\sigma^2E^{\pi_2(\mu,\tau|\vec{x})}\left[\frac{\mathrm{G}_2(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}\right] \\
&\quad -4\left(E^{\pi_2(\mu,\tau|\vec{x})}\left[\frac{\mathrm{G}_1(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}(x_j-\mu)\right]\right)^2,
\end{aligned}
$$

where $\pi_2(\mu,\tau|\vec{x})$ is given by equation (3),

THEOREM 4 *If $\pi_2(\mu,\tau) \equiv 1$, then all the integrals given in Theorem 3 exist provided that $p > 1 + 2/\alpha$.*

PROOF 2 *Using algebra, it can be shown that*

$$
\int_{-\infty}^{\infty} s^j \mathrm{e}^{-s^2/2}\mathrm{e}^{-(a|s|)^\alpha}\mathrm{e}^{ibs}\mathrm{d}s \le \frac{c_1\sigma[x-\mu]^j}{(\sigma^2[1+2a^\alpha]/2+[x-\mu]^2)^{[j+1]/2}},
$$

where c_1 is an appropriate normalizing constant. Using Angers and Berger (1991), we obtain

$$\left|\int_{-\infty}^{\infty} s^j \mathrm{e}^{-s^2/2}\mathrm{e}^{-(a|s|)^\alpha}\mathrm{e}^{ibs}\mathrm{d}s\right| \leq \frac{c_1\sigma|[x-\mu_0]-[\mu-\mu_0]|^j}{(\sigma^2[1+2a^\alpha]/2+([x-\mu_0]-[\mu-\mu_0])^2)^{[j+1]/2}}$$
$$\leq c_1\sigma\,|[x-\mu_0]-[\mu-\mu_0]|^j\left(\frac{(1+|x-\mu_0|)^2}{\sigma^2[1+2a^\alpha]/2+[\mu-\mu_0]^2}\right)^{[j+1]/2}$$
$$= \frac{c_1\sigma(1+|x-\mu_0|)^{j+1}}{(\sigma^2[1+2a^\alpha]/2+[\mu-\mu_0]^2)^{[j+1]/2}}\sum_{m=0}^{j}\binom{j}{m}|x-\mu_0|^{j-m}|\mu-\mu_0|^m,$$

where $\mu_0 \in R$. Hence $E^{\pi_2(\mu,\tau|\vec{x})}[\theta^j] < \infty$ if

$$\left|\int_0^\infty\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}s^j\mathrm{e}^{-s^2/2}\mathrm{e}^{-(a|s|)^\alpha}\mathrm{e}^{ib_ls}\mathrm{d}s\right]\prod_{k\neq l}\int_{-\infty}^{\infty}\mathrm{e}^{-s^2/2}\mathrm{e}^{-(a|s|)^\alpha}\mathrm{e}^{ib_ks}\mathrm{d}s\,\mathrm{d}\mu\mathrm{d}\tau\right|$$
$$\leq c_1^p\sigma^p[1+|x_l-\mu_0|]^{j+1}\left(\prod_{k\neq l}[1+|x_k-\mu_0|]\right)\sum_{m=0}^{j}\binom{j}{m}|x_l-\mu_0|^{j-m}$$
$$\times\int_0^\infty\int_{-\infty}^{\infty}\frac{|\mu-\mu_0|^m}{(\sigma^2[1+2a^\alpha]/2+|\mu-\mu_0|^2)^{(p+j)/2}}\mathrm{d}\mu\mathrm{d}\tau$$
$$\frac{c_1^p 2^{[p+j-1]/2}[1+|x_l-\mu_0|]\left(\prod_{k\neq l}[1+|x_k-\mu_0|]\right)\Gamma(1/\alpha)}{2^{1/\alpha}\sigma^{j-2}\Gamma([p+j]/2)}$$
$$\times\sum_{m=0}^{j}\binom{j}{m}|x_l-\mu_0|^{j-m}\left(\frac{\sigma}{\sqrt{2}}\right)^m\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{p+j-m-1}{2}-\frac{1}{\alpha}\right)$$
$< \infty$ *if* $p > 1 + 2/\alpha$.

Note that if the tail of the $\pi_2(\mu,\tau)$ behaves as τ^{-c}, it can be shown that the integrals given in Theorem 3 exist provided that $p > 1 + 2(1-c)/\alpha$.

One way to evaluate the different quantities given in Theorem 3 is to use Monte Carlo method with importance sampling. Since

$$m(\vec{x}|\mu,\tau) \;\leq\; c_1^p\prod_{k=1}^{p}\left(\frac{(1+|x_k-\mu_0|)^2}{\sigma^2[1+2a^\alpha]/2+[\mu-\mu_0]^2}\right)^{1/2} \tag{4}$$

$$\propto\;\left(\sigma^2[1+2a^\alpha]/2+[\mu-\mu_0]^2\right)^{-p/2}, \tag{5}$$

we can choose, as importance sampling function,

$$(6)\mu|a \;\sim\; \mathcal{T}_{p-1}\left(\mu_0,\sqrt{\frac{\sigma^2[1+2a^\alpha]}{2(p-1)}}\right),$$

$$(7)\;\; a \;\sim\; \text{Burr}\left(\frac{p-1}{2}-\frac{1}{\alpha},\frac{1}{2},\alpha\right),\ \text{i.e. } g(a) = \frac{\alpha\left(\frac{p-1}{2}-\frac{1}{\alpha}\right)a^{\alpha-1}}{(1+2a^\alpha)^{([p-1]/2)-[1/\alpha])}}.$$

3. Behavior of $\widehat{\theta}_j(\vec{x})$ in presence of outlier. In order to study the behavior of $\widehat{\theta}_j(\vec{x})$ in presence of outliers, we need to know the asymptotic behavior of the hypergeometric function. In Olver (1974), it is shown that ${}_1\mathrm{F}_1(\gamma,\delta,z) = \frac{\Gamma(\delta)}{\Gamma(\gamma)}\mathrm{e}^z z^{\gamma-\delta}\times[1+O(z^{-1})]$ if γ is not a negative integer. Otherwise, it can be shown that ${}_1\mathrm{F}_1(\gamma,\delta,z) = (-1)^\gamma\frac{\Gamma(\delta)}{\Gamma(\delta-\gamma)}z^{-\gamma}\times[1+O(z^{-1})]$. Using this result, the following lemma can be easily proven.

LEMMA 1 *If $1 \leq \alpha < 2$ and if $|b|$ is large, then*

$$\begin{aligned} \mathrm{G}_j(a,b) &\approx \frac{\Gamma(k+1/2)}{\sqrt{\pi}2^{[j+k-1]/2}|b|^{[j+k+1]/2}} \sum_{l=1}^{\infty}(-1)^{l-1}\sin([l\alpha+j-k]\pi/2) \\ &\quad \times \frac{\Gamma(l\alpha+j+1)}{l!}\left(\frac{a}{|b|}\right)^{l\alpha}, \end{aligned}$$

where $k = 0$ if j is even and 1 otherwise. If $\alpha = 2$, then

$$\begin{aligned} \mathrm{G}_0(a,b) &= \sqrt{\pi}\mathrm{e}^{-(0.5-a^2)b^2}, \\ \mathrm{G}_1(a,b) &= \mathrm{G}_0(a,b)/2, \\ \mathrm{G}_2(a,b) &= \left[\sqrt{\pi}\mathrm{e}^{-b^2/2} - b^2\mathrm{G}_0(a,b)\right]/2. \end{aligned}$$

In the next theorem, the behavior of the first stage Bayes estimator (and its posterior expected loss) is given for large values of $|x-\mu|$.

THEOREM 5 *If $|x-\mu| >> 0$ and if $1 \leq \alpha < 2$, then*

$$\begin{aligned} m(x|\mu,\tau) &= \frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)}{\pi}\frac{\tau^\alpha}{|x-\mu|^{\alpha+1}} + O(|x-\mu|^{-[2\alpha+1]}), \\ \widehat{\theta}(x|\mu,\tau) &= x - \frac{\sigma^2(\alpha+1)}{2|x-\mu|} + O(|x-\mu|^{-2\alpha}), \\ \rho(\pi,\widehat{\theta}|x,\mu,\tau) &= \sigma^2\left[1+\frac{(\alpha+1)(\alpha+3)}{[x-\mu]^2}\right] + O(|x-\mu|^{-(2\alpha+1)}). \end{aligned}$$

PROOF 3 *If $|b|$ is large and if $\alpha < 2$, using the previous lemma, we have that,*

$$\begin{aligned} \mathrm{G}_0(a,b) &= \sqrt{2}\sin(\alpha\pi/2)\Gamma(\alpha+1)\frac{a^\alpha}{|b|^{\alpha+1}} + O(|b|^{-[2\alpha+1]}), \\ \mathrm{G}_1(a,b) &= \frac{\sin(\alpha\pi/2)\Gamma(\alpha+2)}{2\sqrt{2}}\frac{a^\alpha}{|b|^{\alpha+3}} + O(|b|^{-[2\alpha+3]}), \\ \mathrm{G}_2(a,b) &= -\frac{\sin(\alpha\pi/2)\Gamma(\alpha+3)}{\sqrt{2}}\frac{a^\alpha}{|b|^{\alpha+3}} + O(|b|^{-[2\alpha+3]}). \end{aligned}$$

Consequently, using Theorem 2, one can show that

$$\begin{aligned} m(x|\mu,\tau) &= \frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)}{\pi\sigma}\frac{a^\alpha}{|b|^{\alpha+1}} + O(|b|^{-[2\alpha+1]}), \\ \widehat{\theta}(x|\mu,\tau) &= x - \frac{\sigma(\alpha+1)}{2}\frac{1}{|b|} + O(|b|^{-2\alpha}), \\ \rho(\pi,\widehat{\theta}|\mu,\tau) &= \sigma^2\left[1+\frac{(\alpha+1)(\alpha+3)}{4}\frac{1}{b^2} + O(|b|^{-[2\alpha+1]})\right]. \end{aligned}$$

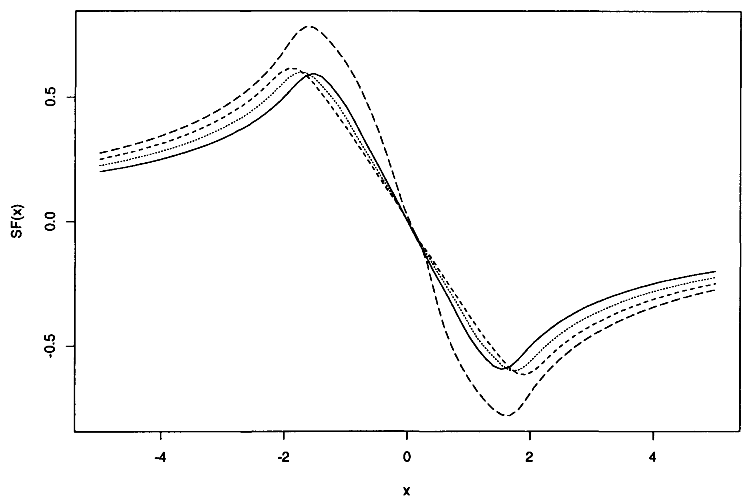

Figure 1: *Behavior of $SF(x) = (\hat{\theta}(x \mid \mu, \tau) - x]/\sigma$ for outlying values of x for $\mu = 0, \tau = \sigma = 1$ and $\sigma = 1$ (solid line), 1.25 (dotted line), 1.5 ("en" dashed line) and 1.75 ("em" dashed line).*

In Figure 1, the behavior of $\mathrm{SF}(x) = \left[\widehat{\theta}(x|\mu,\tau) - x\right]/\sigma$ is illustrated for several values of α. It can be seen from this figure that the chosen value of α does not affect much the behavior of $\mathrm{SF}(x)$. For all values of α considered, $\mathrm{SF}(x)$ is almost linear for small values of $|x-\mu|/\sigma$. However, if $|x-\mu|/\sigma$ is large, the shrinkage function is monotonically decreasing towards 0. The rate of decrease of $\mathrm{SF}(x)$ is controlled by α. The smaller α is, the faster $\mathrm{SF}(x)$ goes to 0.

In the next theorem, the robustness behavior of $\widehat{\theta}_j(\vec{x})$ is given.

THEOREM 6 *For $j = 1, 2, \ldots, l-1, l+1, \ldots, p$, let $X_j \sim N(\theta_j, \sigma^2)$ (σ^2 known), and $\theta_j \sim S_\alpha(\mu, \tau)$ where $1 \le \alpha < 2$ and $p > 3 + 2/\alpha$. If $|x_l - \mu| \to \infty$, that is x_l is an outlier, then*

$$\begin{aligned}
\pi_2(\mu,\tau|\vec{x}) &\to \pi_2^*(\mu,\tau|\vec{x}_{(-l)}), \\
\widehat{\theta}_l(\vec{x}) &\to x_l, \\
\widehat{\theta}_j(\vec{x}) &\to E^{\pi_2^*(\mu,\tau|\vec{x}_{(-l)})}\left[\widehat{\theta}_j(x_j|\mu,\tau)\right] \ \textit{for } j \ne l, \\
\rho(\pi,\widehat{\theta}_l|\vec{x}) &\to \sigma^2, \\
\rho(\pi,\widehat{\theta}_j|\vec{x}) &\to \sigma^2 - 2\sigma^2 E^{\pi_2^*(\mu,\tau|\vec{x}_{(-l)})}\left[\frac{\mathrm{G}_2(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}\right]
\end{aligned}$$

$$-4\left(E^{\pi_2^*(\mu,\tau|\vec{x}_{(-l)})}\left[\frac{\mathrm{G}_1(\tau/\sigma,[x_j-\mu]/\sigma)}{\mathrm{G}_0(\tau/\sigma,[x_j-\mu]/\sigma)}(x_j-\mu)\right]\right)^2 \quad \textit{for } j\neq l,$$

where $\pi_2^*(\mu,\tau|\vec{x}_{(-l)}) \propto \tau^\alpha \prod_{j\neq l} m(x_j|\mu,\tau)$.

PROOF 4 *Let* $d_l = \min_{j\neq l}|x_l - x_j|$. *Using the dominated convergence theorem and Theorem 5, we have:*

$$\begin{aligned}
&\lim_{|x_l|\to\infty}\int_0^\infty\int_{-\infty}^\infty \left[|x_l|^{\alpha+1}m(x_l|\mu,\tau)\right]\left[\textstyle\prod_{j\neq l} m(x_j|\mu,\tau)\right]\mathrm{d}\mu\mathrm{d}\tau \\
=&\lim_{|x_l|\to\infty}\int_0^\infty\left\{\frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)\tau^\alpha}{\pi}\int_{-\infty}^{x_l-d_l/2}\left|\frac{x_l}{x_l-\mu}\right|^{\alpha+1}\left[\textstyle\prod_{j\neq l} m(x_j|\mu,\tau)\right]\mathrm{d}\mu\right. \\
&+\left(\frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)\tau^\alpha}{\pi}\right)^{p-1}\int_{-d_l/2}^{d_l/2}\frac{|x_l|^{\alpha+1}m(x_l|x_l+\mu,\tau)}{\prod_{j\neq l}|x_j+x_l-\mu|^{\alpha+1}}\mathrm{d}\mu \\
&\left.+\left(\frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)\tau^\alpha}{\pi}\right)^{p}\int_{d_l/2}^\infty\frac{|x_l|^{\alpha+1}}{\prod_{j=1}^p|x_j+x_l-\mu|^{\alpha+1}}\mathrm{d}\mu\right\}\mathrm{d}\tau \\
=&\frac{\sin(\alpha\pi/2)\Gamma(\alpha+1)}{\pi}\int_0^\infty\tau^\alpha\int_{-\infty}^\infty\textstyle\prod_{j\neq l} m(x_j|\mu,\tau)\mathrm{d}\mu\mathrm{d}\tau<\infty
\end{aligned}$$

if $p > 3 + 2/\alpha$. *The rest of the proof follows directly from this result.*

Note that, if there is k outliers $x_{l_1},\ldots x_{l_k}$, then

$$\pi_2^*(\mu,\tau|\vec{x}_{(-l_1,\ldots,-l_k)}) \propto \tau^{k\alpha}\pi_2(\mu,\tau)\prod_{j\notin\{l_1,\ldots,l_k\}} m(x_j|\mu,\tau).$$

If $\pi_2(\mu,\tau)\equiv 1$, then all the quantities given in Theorem 6 exist if $p > 2k+1+2/\alpha$.

4. Numerical example. For the numerical example discussed in this section, the second stage prior was chosen to be $\pi_2(\mu,\tau)\equiv 1$. To evaluate the integrals given in Theorem 3, the Monte Carlo method with importance sampling was used. The importance sampling functions are given by equations (6) and (7). The observations vector used in this section has been generated according to the following scheme:

1. generate $\theta_j \sim S_1(0,1)$, for $j = 1,\ldots,10$;
2. generate $X_j \sim N(\theta_j,1)$, for $j = 1,\ldots,10$.

Using these observations we computed $\widehat{\theta}_j(\vec{x})$, $j = 1,\ldots,10$ with $\alpha = 1.0$, 1.25, 1.5, 1.75 and 2.0. These values are given in Table 1 along with their posterior expected loss (numbers in parentheses). From this table, one can see that the choice of α does not influence much on $\widehat{\theta}_j(\vec{x})$ as long as $\alpha < 2$. Furthermore, $\widehat{\theta}_j(\vec{x})$ (for $\alpha < 2$) has a limited shrinkage for large values of x_j, that is x_1, and x_{10}. However, this does not prevent the shrinkage for the other coordinates.

TABLE 1: *Values of* $\widehat{\theta}(\vec{x})$ *and* $\rho(\pi, \widehat{\theta}|\vec{x})$

x_j	$\alpha = 1.0$	$\alpha = 1.25$	$\alpha = 1.5$	$\alpha = 1.75$	$\alpha = 2.0$
-4.55	-4.35 (1.16)	-4.34 (1.22)	-4.31 (1.26)	-4.31 (1.27)	-3.32 (0.83)
-1.88	-1.51 (1.18)	-1.45 (1.17)	-1.41 (1.26)	-1.41 (1.91)	-1.27 (0.80)
-0.54	-0.20 (0.87)	-0.16 (0.80)	-0.13 (0.79)	-0.16 (1.32)	-0.24 (0.80)
-0.44	-0.12 (0.86)	-0.08 (0.78)	-0.05 (0.79)	-0.08 (1.22)	-0.17 (0.80)
0.54	0.58 (0.87)	0.61 (0.86)	0.61 (0.82)	0.77 (0.66)	0.59 (0.80)
0.91	0.83 (0.85)	0.86 (0.86)	0.85 (0.84)	0.97 (0.80)	0.87 (0.81)
1.09	0.95 (0.85)	0.97 (0.85)	0.97 (0.83)	1.08 (0.84)	1.01 (0.81)
2.31	1.92 (0.94)	1.87 (0.83)	1.84 (0.83)	1.89 (0.83)	1.94 (0.82)
2.58	2.19 (1.01)	2.13 (0.92)	2.08 (0.88)	2.21 (0.89)	2.15 (0.83)
4.16	3.87 (1.25)	3.83 (1.30)	3.79 (1.43)	3.75 (1.29)	3.36 (0.86)

REFERENCES

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

ANGERS, J.-F., 1995. Fourier transform and Bayes estimator of a location parameter. *Statistics & Probability Letters.* (to appear).

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

FELLER, W., 1966. *An introduction to probability theory and its applications*, Volume II. John Wiley and Sons, New York.

GRADISHTEYN, L. S. AND RYZHIK, I. M., 1980. *Table of integrals, series and products.* Academic Press, New York.

OLVER, F., 1974. *Asymptotics and special functions.* Academic Press, New York.

PRESS, S. J., 1975. Modelling with heavy tails. In *A Modern Course on Statistical Distributions in Scientific Work and the International Conference on Characterizations of Statistical Distribution with Application*, pages 87–102. D. Reidel Publishing Co.

DÉP. DE MATHÉMATIQUES ET DE STATISTIQUE
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCC. CENTRE-VILLE
MONTRÉAL, QC H3C 3J7