JKAU: Sci., **Vol. 18,** pp: 93-103 (2006 A.D. / 1427 A.H.)

Robust Bayesian Estimation of Location for Symmetric Stable Distributions

Ateq A. Al-Ghamedi
Dept. of Statistics, Faculty of Science,
King Abdulaziz University, Jeddah, Saudi Arabia
E-mail: aamalghamdi@kau.edu.sa

Abstract. A robust Bayesian estimation of location parameter θ of symmetric stable distributions $\alpha \in$ (2, 1.5, 1, 0.5) used to estimate the location parameter θ for the posterior distribution. Normal distribution is used as special case form the family of symmetric stable distributions to show the derivation of the asymptotic estimation of location parameter. Computer simulation used to investigate the performance of our robust procedure. Our robust procedure can be adopted to investigate other symmetric distribution.

Keywords: Order statistics; Qantiles; Ratio of the symmetric differences statistic; Tail thickness.

Introduction

Stable distributions are rich class of distributions that allow skewness and heavy tails. The class was characterized by Paul Levy (1954), in his study of sums of independent identically distributed terms. The general stable distribution is described by four parameters: an index of stability $\alpha \in$ (0, 2], a skewness parameter $\beta \in$ [−1, 1], a scale parameter $\gamma > 0$ and a location parameter $\beta \in \mathbb{R}$. The lack of closed formulas for densities and distribution functions for all but a few stable distributions (Gaussian, Cauchy and Levy) has been a major drawback to the use of stable distributions by practitioners[1].

There are multiple parameterizations for stable distributions. We will use the following:

Definition: A random variable X is stable $S(\alpha, \beta, \gamma, \theta)$ if and only if

$$X^d = AZ + B , \qquad (1.1)$$

where $0 < \alpha \le 2$, $-1 \le \beta \le 1$, $A \ge 0$, $B \in \mathbb{R}$, $Z = Z(\alpha, \beta)$ is a random variable with characteristic function

$$E[\exp(iuZ)] = \begin{cases} \exp(-|u|^{\alpha}\ [1 + i\,\beta \tan\frac{\pi\alpha}{2}(sign\ \ u)(|u|^{1-\alpha} - 1)]) & \alpha \ne 1 \\ \exp(-|u|[1 + i\,\beta\frac{2}{\pi}(sign\ \ u)\ \ \ln|u|]) & \alpha = 1. \end{cases} \quad (1.2)$$

and

$$sign\ \ u = \begin{cases} -1 & u < 0 \\ 0 & u = 0 \\ 1 & u > 0. \end{cases}$$

Some properties of stable distributions are:

1. Every stable distribution has a mode.
2. It is clear that $Z(2, 0) \sim N(0, 2)$, $Z(1, 0) \sim$ *Cauchy* (0, 1) and $Z(\frac{1}{2}, 0) \sim$ *Levy* (0, 1).
3. When $Z(\alpha, -\beta) \overset{d}{=} -Z(\alpha, \beta)$ then the distribution is symmetric.
4. When $1 < \alpha \le 2$, the mean of $X \sim S(\alpha, \beta, \gamma, \theta)$ is $\mu = E(X) = \theta - \beta\gamma \tan \frac{\pi\alpha}{2}$. As $\alpha \downarrow 1$, it has a mean of $\mu = \beta \tan \frac{\pi\alpha}{2}$. When $\beta = 0$, the mean is always 0. When $\beta > 0$, the mean tends to $+\infty$ because both tail are getting heavier, the right tail is heavier than left. By symmetry, the $\beta < 0$ case has the mean tends to $-\infty$. Finally, when α reaches 1, the tails are too heavy for the integral $E(X) = \int_{-\infty}^{\infty} xf(x)dx,$ to converge.
5. Tail approximation: let $X \sim S(\alpha, \beta)$ with $0 < \alpha \le 2$, $-1 \le \beta \le 1$. Then for large X,

$$\begin{aligned} P(X > x) &= C_{\alpha}(1 + \beta)\, x^{-\alpha}, \\ f(x / \alpha, \beta) &\sim \alpha C_{\alpha}(1 + \beta)\, x^{-(\alpha+1)}, \end{aligned} \quad (1.3)$$

where $C_{\alpha} = \Gamma(\alpha)(\sin\frac{\pi\alpha}{2})/\pi$. Using the symmetry property, the lower tail properties are similar.

6. The class of distribution functions, D, selected for the study consists of all df's $F(x - \theta)$, $-\infty < \theta < \infty$, where $F(x); = S_{\alpha}(x)$; and $S_{\alpha}(x)$ is the df of the symmetric stable distribution with index of stability α. The probability density function (pdf) $f(x) = F'(x)$ is represented by the inverse Fourier transform of the cdf (with the location parameter equal to zero and the scale parameter equal to one) as:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-|t|^{\alpha}} e^{-itx} dt = \frac{1}{\pi} \int_{0}^{\infty} e^{-t^{\alpha}} \cos(tx) dt \quad -\infty < x < \infty. \qquad (1.4)$$

Note that for $\alpha = 2$, we have a normal distribution with mean zero and variance 2[2, 3].

7. The quantile function $\xi(q) = F^{-1}(q)$, $0 < q < 1$, is uniquely defined[4].

Tables 1 and 2 show some of quantiles and density values of stable distributions (α = 2, 1.5, 1 and 0.5). Also Fig. 1 and 2 show distribution functions and densities[4].

Table 1. $f(x_q)$ of some stable distributions.

α \ q	**0.50**	**0.75**	**0.95**
2.00	0.282095	0.224702	0.072928
1.50	0.287353	0.206242	0.030029
1.00	0.318310	0.159155	0.007790
0.50	0.636620	0.065480	0.000413

Table 2. Quantiles of stable distributions.

α \ q	**0.60**	**0.75**	**0.90**	**0.95**	**0.99**
2.00	0.35827	0.95387	1.81238	2.32617	3.28995
1.50	0.35334	0.96893	2.06146	0.05194	7.73644
1.00	0.32492	1.00000	3.07768	6.31375	31.82052
0.50	0.20889	1.28383	12.7413	57.30403	1559.72610

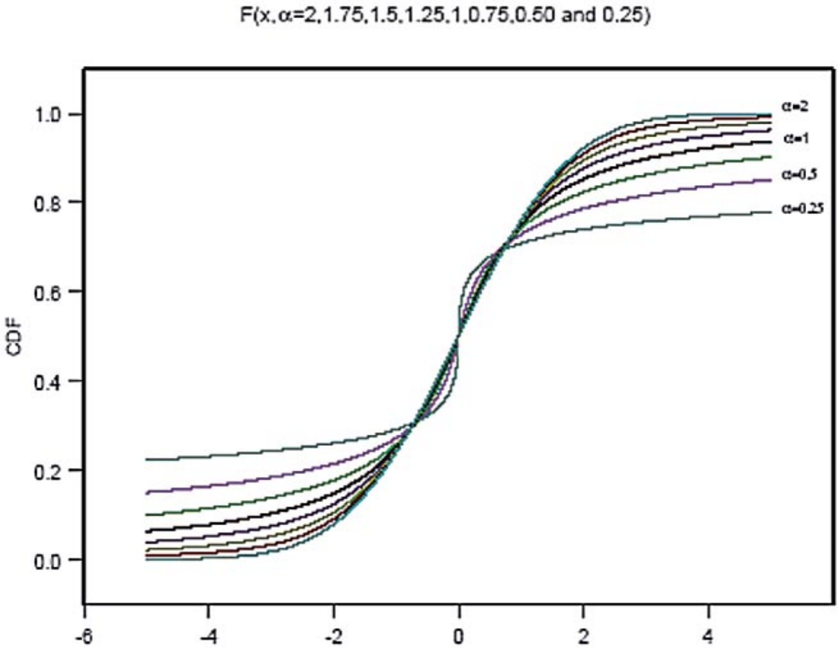

Fig. 1. Symmetric stable distribution functions.

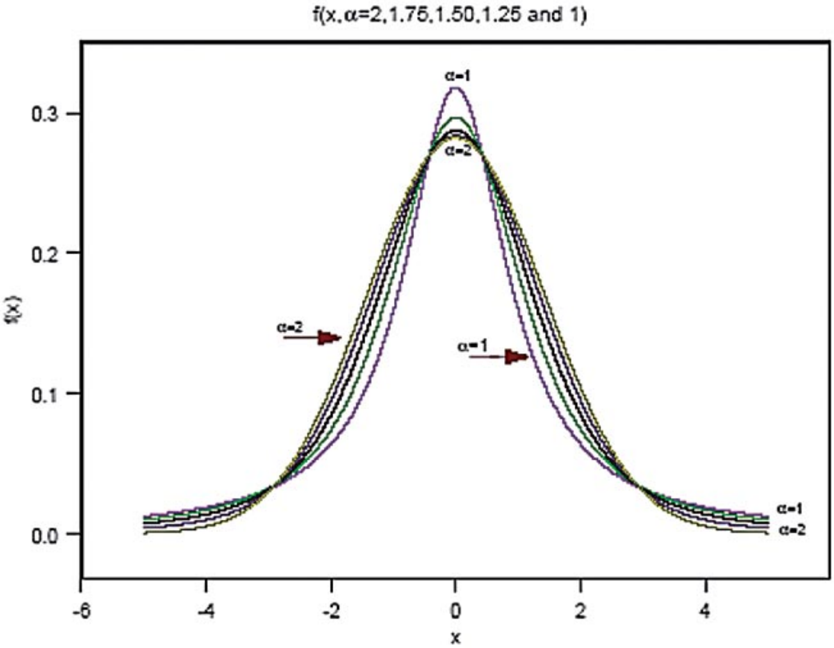

Fig. 2. Symmetric stable densities.

Bayesian Estimation

Let $X_1, X_2, ..., X_n$ be an observable random variable with density $f(x / \theta)$, $\theta \in \mathbb{R}$.

$$f(x_i / \theta) = \frac{1}{\pi} \int_0^\infty e^{-t^\alpha} \cos(t(x-\theta))dt \quad -\infty < x < \infty. \tag{2.1}$$

where $\alpha = 2, 1.5, 1$ or 0.5. A prior distribution on the unobservable parameter θ has to be subjectively elicited

$$p(\theta) = \frac{1}{\pi} \frac{1}{(1+\theta^2)} - \infty < \theta < \infty \quad . \tag{2.2}$$

Cauchy distribution has been chosen since it has a heavy tail, so the estimates will be robust. The posterior expectation of the quantity $m(\theta)$, a function of the location parameter, is of interest, if it does exist. If it does not exist then the other measure of location such as the median or the trimmed mean will be used instead. The posterior distribution of the location parameter θ is given the variables $x_1, x_2, ..., x_n$ is:

$$f(\theta / x_1, ..., x_n) = \frac{[\prod_{i=1}^n f(x_i / \theta)] p(\theta)}{\int [\prod_{i=1}^n f(x_i / \theta)] p(\theta) d\theta} \overset{\theta}{\alpha} \frac{1}{1+\theta^2} \prod_{i=1}^n [\int_0^\infty e^{-t^\alpha} \cos(t(x_i - \theta))dt] \tag{2.3}$$

Under the squared error loss function, the Bayes estimation of $m(\theta)$ is the posterior mean which is given by:

$$E(m(\theta) / \underline{x}) = \frac{\int \frac{m(\theta)}{1+\theta^2} [\prod_{i=1}^n f(x_i / \theta)] d\theta}{\int \frac{1}{1+\theta^2} [\prod_{i=1}^n f(x_i / \theta)] d\theta} \quad . \tag{2.4}$$

Since this distribution is unknown and the ratio of integrals in (2.4) does not seem to take a closed form and can not be approximated by a known distribution, the Importance sampling method will be used to sample from the posterior and then study the samples properties[5, 6].

Results

Since the densities of stable distribution do not have a close form for all α, a special case will be chosen to show the method estimation ($\alpha = 2$).

$$f(\theta/\underline{x}) = \frac{\dfrac{1}{1+\theta^2} e^{-\sum(x_i-\theta)^2/4}}{\int \dfrac{1}{1+\theta^2} e^{-\sum(x_i-\theta)^2/4} d\theta} . \tag{3.1}$$

and

$$E(\theta/\underline{x}) = \frac{\int \dfrac{\theta}{1+\theta^2} e^{-\sum(x_i-\theta)^2/4} d\theta}{\int \dfrac{1}{1+\theta^2} e^{-\sum(x_i-\theta)^2/4} d\theta} . \tag{3.2}$$

then a numerical estimation of the expectation is given by

$$E(\theta/\underline{x}) \simeq \frac{\sum_{i=1}^{n} \dfrac{\theta}{1+\theta^2}}{\sum_{i=1}^{n} \dfrac{1}{1+\theta^2}} \tag{3.3}$$

Table 3 shows a summary of the samples that were generated from posterior distributions. It seems that the median in all distributions gave a good approximation for the location parameter. The first and third quartiles are good estimates of the real quartiles in all the distributions. Also, Fig. 3 to 12 show the shape of the posterior distributions and histograms of the samples. We can say that the median is the best measure of location for the family of symmetric stable distributions. However, in some cases such as $\alpha = 2$ the mean is better.

Table 3. Summary of samples from posterior distributions.

q / α	**Min**	***Q*1**	**Med**	**Mean**	***Q*3**	**Max**
2.00	–18.711	– 1.026	–0.004	0.003	1.147	17.060
1.50	–85.408	– 2.393	0.041	–0.271	1.642	159.381
1.00	–1406.2	–11.985	0.148	–	14.136	1540.8
0.50	–	–29307	–0.943	–	194740	–

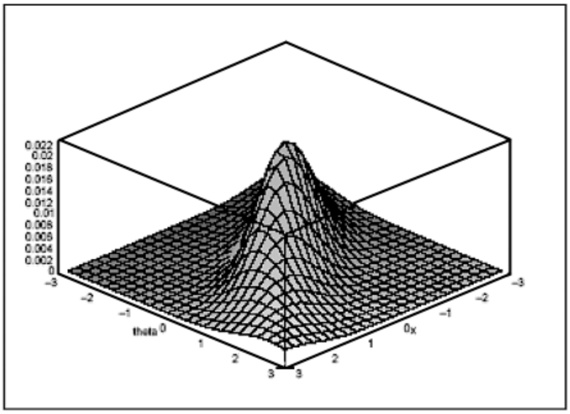

Fig. 3. Posterior shape when $\alpha = 2$.

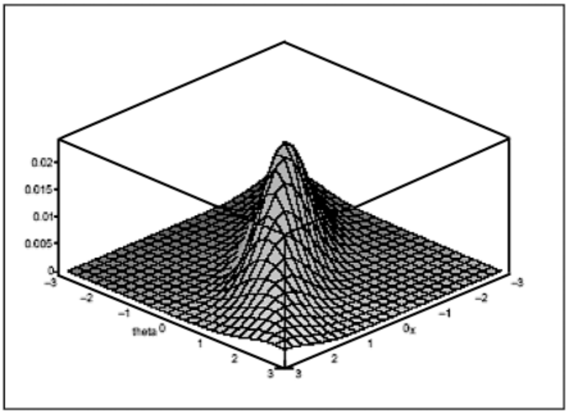

Fig. 4. Posterior shape when $\alpha = 1.5$.

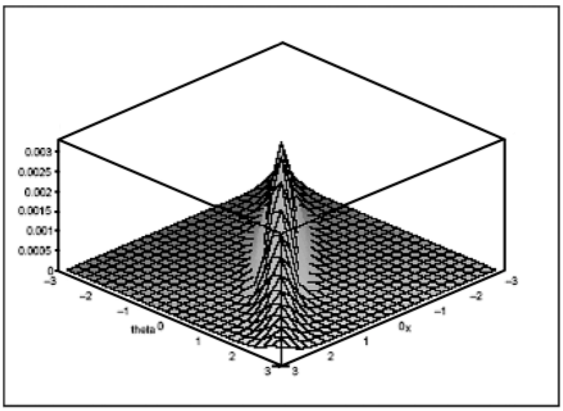

Fig. 5. Posterior shape when $\alpha = 1$.

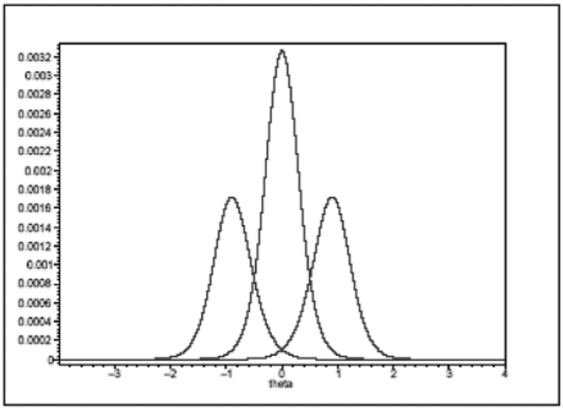

Fig. 6. Posterior shape when $\alpha = 1$.

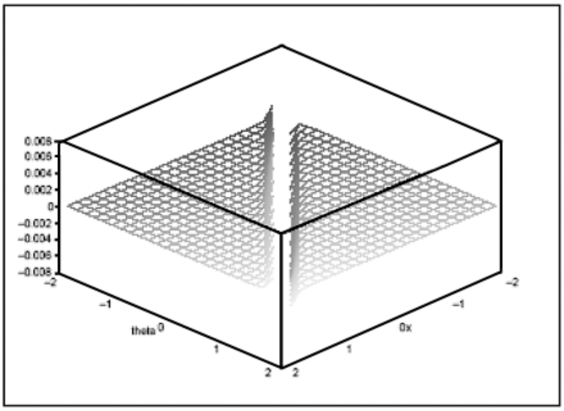

Fig. 7. Posterior shape when $\alpha = 0.5$.

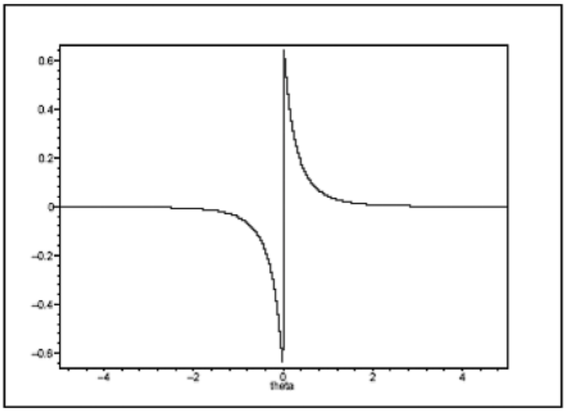

Fig. 8. Posterior shape when $\alpha = 0.5$.

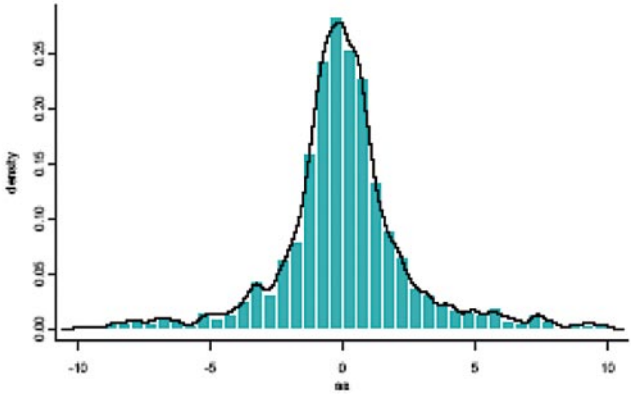

Fig. 9. Simulation of posterior shape when $\alpha = 2$.

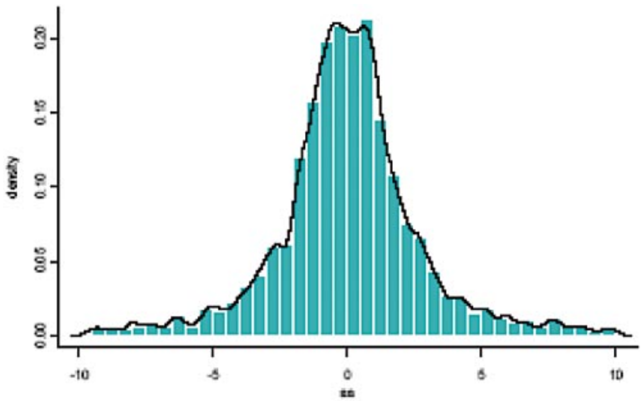

Fig. 10. Simulation of posterior shape when $\alpha = 1.5$.

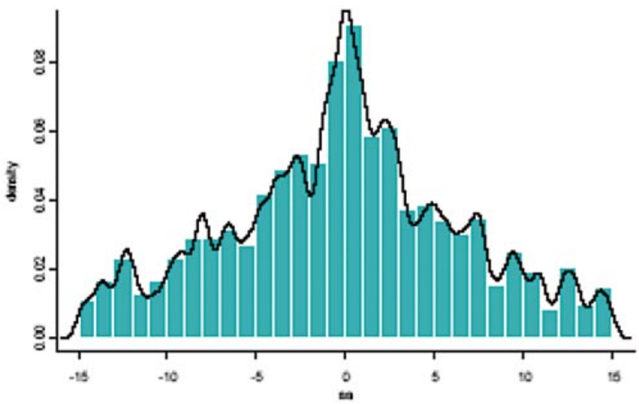

Fig. 11. Simulation of posterior shape when $\alpha = 1$.

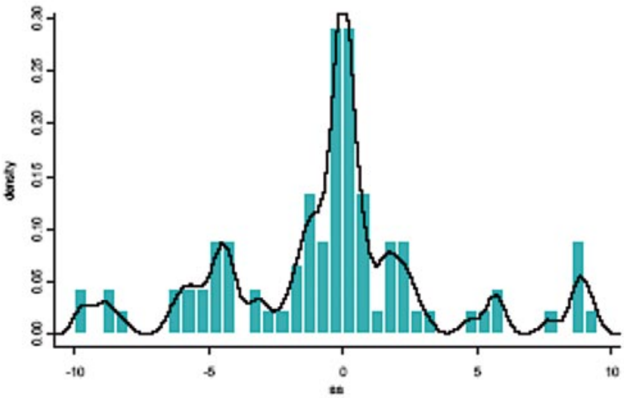

Fig. 12. Simulation of posterior shape when $\alpha = 0.5$.

References

[1] **Nolan, J.P.,** "Numerical Calculation of Stable Densities and Distribution Functions", *Communications in Statistics-Stochastic Models*, **13:** 163-170 (1997).

[2] **Badahdah, S.O.** and **Siddiqui, M.M.,** "Adaptive Robust Estimation of Location for Stable Distributions", *Communications in Statistics- Theory and Methods*, **20:** 631-652 (1991).

[3] **Crow, E.L.** and **Siddiqui, M.M.,** "Robust Estimation of Location", *Journal of the American Statistical Association*, **62:** 353-389 (1967).

[4] **Al-Gamedi, A.A.,** *"Robust Estimation and Testing of Location for Symmetric Stable Distributions"*, Ph.D. Dissertation. Colorado State University, Fort Collins (2002).

[5] **Bernardo, J.M.** and **Smith, A.F.,** *"Bayesian Theory"*, J. Wiley Sons, New York (1994).

[6] **Insua, D.R.** and **Ruggeri, F.,** *"Robust-Bayesian Analysis"*, Springer-Verlag, New York (2000).

تقديرات مقاومة بيزية للمعلمة المكانية في التوزيعات المتزنة المتماثلة

عتيق أحمد الغامدي

قسم الإحصاء ، كلية العلوم ، جامعة الملك عبدالعزيز

جـــدة – المملكة العربية السعودية

المستخلص. بما أن التوزيعات المتزنة (Stable distributions) ليس لها صيغة رياضية محددة ، فإنه من الصعب التعامل معها من قبل الباحثين . ويحجم كثير من الباحثين عن استخدامها رغم أهميتها في تمثيل الكثير من الظواهر . كما أن خاصية الذيل العريض الذي تشتهر به تحاكي كثيراً من الظواهر الطبيعية . وفي هذه الورقة تم استخدام طريقة المحاكاة لتصوير شكل التوزيعات البعدية للتوزيعات المتزنة ، ومن ثم استخدام إحدى طرق المعاينة (Importance sampling) لتوليد بيانات من التوزيع البعدي (Posterior distribution) . وبعد ذلك استخدمت هذه البيانات لتقدير المعالم المكانية للتوزيعات . كما استخدم التوزيع الطبيعي كحالة خاصة من عائلة التوزيعات المتزنة لعرض الطرق الرياضية لتقدير المعلمة المكانية . ويمكن تعميم استخدام هذه الطريقة لتقدير المعالم المكانية لأي نوع من أنواع التوزيعات المتماثلة .