Inference for the Positive Stable Laws Based on a Special Quadratic Distance

Andrew Luong* and Louis G. Doray[1]**

*École d'actuariat, Université Laval,

Cité Universitaire, Ste-Foy, Québec, Canada G1K 7P4

**Département de mathématiques et de statistique, Université de Montréal,

C.P. 6128, Succursale Centre-ville, Montréal, Québec, Canada H3C 3J7

ABSTRACT

Inference methods for the positive stable laws, which have no closed form expression for the density functions are developed based on a special quadratic distance using negative moments. Asymptotic properties of the quadratic distance estimator (QDE) are established. The QDE is shown to have asymptotic relative efficiency close to 1 for almost all the values of the parameter space.

Goodness-of-fit tests are also developed for testing the parametric families and practical numerical techniques are considered for implementing the methods. With simple and efficient methods to estimate the parameters, positive stable laws could find new applications in actuarial science for modelling insurance claims and lifetime data.

Key Words and Phrases: moment-cumulant estimators, negative moments, Laplace transform, quadratic distance, consistency, asymptotic normality, efficiency, goodness-of-fit, Gauss-Newton algorithm.

[1] corresponding author

1 INTRODUCTION

The positive stable laws are defined by their Laplace transform

$$L_X(t) = E(e^{-tX}) = \exp(-ct^{\alpha}), \quad t > 0,$$

where the random variable X is positive, α is the index parameter $(0 < \alpha < 1)$ and $c^{1/\alpha}$ is the scale parameter. Quite often, one encounters insurance claims or lifetime data which display heavy tail behaviors which make positive stable laws good candidates for fitting this type of data.

Theoretically, stable laws have a domain of attraction; they can be viewed as a type of limit laws which can be used to approximate the real law underlying the physical process when the real law is not tractable. For discussions on domain of attraction, see Feller (1971).

Positive stable laws with 2 parameters as given by Brockwell and Brown (1979, 1981) form a subfamily of Hougaard's laws with 3 parameters, which are used in survival analysis for modelling time to event data (see Hougaard (1986)).

The density function of positive stable laws has no closed form expression. It can be represented using infinite series expansions as

$$f(x; c, \alpha) = -\frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{\Gamma(k\alpha + 1)}{k!} (-cx^{-\alpha})^k \sin(\alpha k \pi)$$

(see Feller (1971) or Hougaard (1986)). It should be noted that this makes the maximum likelihood method rather complicated to use for estimating the parameters.

In this paper, a class of quadratic distance estimators using special moments via the Laplace transform are proposed. The method is relatively simple to implement and good efficiency is attainable. The paper is organized as follows. Section 1 gives some properties and motivations for positive stable laws. Some alternative estimators proposed by Brockwell and Brown (1979) are reviewed in section 2. The quadratic distance estimator based on special moments is presented in section 3. Its asymptotic properties are studied in section 4; the estimator is shown to have very high efficiency. Goodness-of-fit test statistics are developed in section 5, where test statistics are shown to have an aymptotic chi-squared distribution. Section 6 presents a numerical algorithm to implement the methods, as well as numerical illustrations. It will be seen that the methods are relatively simple to implement numerically. With simpler methods, we hope that practitioners can use the positive stable laws when it is needed and not let the complexity of its density function be a barrier. Finally, in the last section, we provide some comments and possible extensions.

2 TRADITIONAL ESTIMATORS

Alternative estimators to the maximum likelihood estimators have been proposed in the literature, see for example Brockwell and Brown (1979, 1981), Jarrett (1984).

We will focus here on the moment-cumulant estimators given by Brockwell and Brown (1979). This estimator is consistent but not very efficient as the parameter α tends to 1. The moment-cumulant estimator is defined below. Often, another parametrization is used by defining $\theta_1 = 1/\alpha$ and $\theta_2 = (1/\alpha)\ln c$. The Laplace transform of X then becomes

$$L_X(t) = \exp[-(te^{\theta_2})^{1/\theta_1}], \quad \theta_1 > 1, \quad \theta_2 \in \Re.$$

The original sample $X_1, \ldots, X_n$ is transformed to $Y_i = \ln X_i$, $i = 1, \ldots, n$. The mean and the variance of Y, denoted by $k_1 = E(Y)$ and $k_2 = \mathrm{Var}(Y)$ respectively, are given by Brockwell and Brown (1979) as

$$k_1 = (\alpha^{-1} - 1)\gamma + \alpha^{-1}\ln c = (\theta_1 - 1)\gamma + \theta_2$$

$$k_2 = (\alpha^{-2} - 1)\pi^2/6 = (\theta_1^2 - 1)\pi^2/6,$$

where γ=0.5772157... is Euler's constant. Letting $\bar{k}_1$ and $\bar{k}_2$ represent the sample mean and sample variance respectively, the moment-cumulant estimators $\bar{\theta}_1$ and $\bar{\theta}_2$ are obtained as solutions of $\bar{k}_1 = k_1$ and $\bar{k}_2 = k_2$, a system of equations easy to solve (see subsection 6.1 for the explicit solutions for $\bar{\theta}_1$ and $\bar{\theta}_2$).

Those two authors found that the asymptotic variances of the moment-cumulant estimators $\bar{\theta}_1$ and $\bar{\theta}_2$ were independent of the parameter θ_2; they were dependent only on the value of $\alpha = 1/\theta_1$, $0 < \alpha < 1$, as shown in Table 1.

Using Fourier's series expansion, Brockwell and Brown (1979, 1981) obtained Fisher's information matrix numerically for various values of α or equivalently θ_1. The information matrix is independent of θ_2. By CR_1 and CR_2 in Table 1, we mean the Cramér-Rao lower bound for θ_1 and θ_2, i.e. the variances of the MLE for estimating θ_1 and θ_2 respectively, which are given by the diagonal elements of the inverse of the Fisher's information matrix. Results from Brockwell and Brown (1981) are reproduced in Table 1 (they differ slightly from those in Brockwell and Brown (1979)).

Other estimators were also given by Brockwell and Brown (1979, 1981) such as the estimator based on order statistics and the two-step estimator. The estimator based on order statistics is called the combined estimator by Brockwell and Brown (1979); it is not an efficient estimator for values of α close to 1. The two-step estimator (see Brockwell and Brown (1981)) requires knowledge of the true parameters to be able to obtain high efficiency; it is based on a preliminary estimator obtained by equating the sample average and the expected values of $\exp(-t_j X), t_j \in \Re,\ j = 1, 2$, for some fixed t_1 and t_2. Taking into account the extra variation due to this extra step might lower the reported high efficiency.

Another approach which makes use of the characteristic function is given by Heathcote (1977), but less is known about the question of efficiency of this estimator.

The quadratic distance estimator based on Laplace transform and its asymptotic properties will be given in the next section. The QDE has very high efficiency for practically all the values of the parameters.

3 THE QUADRATIC DISTANCE ESTIMATOR BASED ON LAPLACE TRANSFORM

Using the Laplace transform $L_X(t) = E(e^{-tX}) = \exp(-ct^\alpha), \quad t > 0$, with $\theta_1 = 1/\alpha$ and $\theta_2 = (1/\alpha)\ln c$, Brockwell and Brown (1981) have shown that the negative moments of the positive stable laws possess the property that

$$E(X^{-t}) = e^{-t\theta_2}\Gamma(1+t\theta_1)/\Gamma(1+t),$$

where t is fixed and $\Gamma(\cdot)$ is the gamma function.

Let us define the function $\psi_\theta(t) = E(X^{-t})$; it is natural to use the theoretical moments defined by $\psi_\theta(t)$ and to match them with their empirical counterparts to define the QDE. More precisely, let us choose the points $t_1, t_2, \ldots, t_k > 0$ and let us define the empirical estimator of $\psi_\theta(t_j)$,

$$\psi_n(t_j) = (1/n)\sum_{i=1}^{n} X_i^{-t_j},$$

where $X_1, \ldots, X_n$ are i.i.d. observations from the positive stable laws.

Let us also define the vectors

$Z_n = (\psi_n(t_1), \ldots, \psi_n(t_k))'$

$Z(\theta) = (\psi_\theta(t_1), \ldots, \psi_\theta(t_k))'$, where θ is the parameter vector $[\theta_1, \theta_2]'$.

The QDE based on the Laplace transform, denoted by $\hat{\theta}$, is defined as the value of θ which minimizes the distance

$$d(\theta) = (Z_n - Z(\theta))' Q(\theta)(Z_n - Z(\theta)), \qquad (1)$$

where $Q(\theta)$ is a positive definite matrix which might depend on θ. This is a special distance within the class of quadratic distance introduced by Feuerverger and Mc Dunnough (1984) and fully developed by Luong and Thompson (1987), where a unified theory for estimation and goodness-of-fit is developed. Here, we would like to exploit this quadratic distance to obtain high efficiency estimators for positive stable laws in a way which is more accessible for practioners, and to develop goodness-of-fit tests at the same time.

Feuerverger and McDunnough (1984) noted that by letting $t_1, t_2, \ldots, t_k$ become dense, this estimation technique yields arbitrary high efficiency estimator for positive stable laws. Here, we are approaching the problem using quadratic distance so that the problem of estimation and goodness-of-fit tests are tackled in a unified way. Furthermore, we would like to study the efficiency of the estimator with a finite number of points $t_1, t_2, \ldots, t_k$ and give guidelines to practitioners on how to choose these points (see section 4). Algorithms to implement the techniques are given in section 6.

4 ASYMPTOTIC PROPERTIES OF THE QDE

4.1 Consistency and asymptotic normality

Let θ_0 be the true value of parameter θ. Using lemma (2.4.2) in Luong and Thompson (1987), we can conclude that:
i) $\hat{\theta}_n \xrightarrow{\mathcal{P}} \theta_0$, where $\xrightarrow{\mathcal{P}}$ denotes convergence in probability, i.e. the QDE $\hat{\theta}$ is a consistent estimator of the parameter θ,
ii) $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma_1)$, where $\xrightarrow{\mathcal{L}}$ denotes convergence in law, with

a) $\Sigma_1 = (S'QS)^{-1}S'Q\Sigma QS(S'QS)^{-1}$,

b) $Q = Q(\theta_0)$,

c) $S = \left(\frac{\partial Z_i(\theta)}{\partial \theta_j}\right) = \left(\frac{\partial \psi_\theta(t_i)}{\partial \theta_j}\right)$, a matrix of dimension $k \times 2$, with $i = 1, \ldots, k$ and $j = 1, 2$,

d) $\Sigma = (\sigma_{ij})$ is the variance-covariance matrix of $\sqrt{n}[Z_n - Z(\theta_0)]$ under the hypothesis $\theta = \theta_0$, where

$$\sigma_{ij} = \left(E[X^{-(t_i+t_j)}]\right) - E(X^{-t_i})E(X^{-t_j}) = \psi_\theta(t_i + t_j) - \psi_\theta(t_i)\psi_\theta(t_j).$$

We can also conclude that:

i) $Q(\theta)$ can be replaced by an estimate $\hat{Q}$ in (1) and if $\hat{Q} \xrightarrow{\mathcal{P}} Q(\theta_0) = Q$, we have an estimator asymptotically equivalent to the one obtained by minimizing (1).

ii) The most efficient choice of $Q(\theta)$ is $\Sigma^{-1}(\theta)$, and with this choice of $Q(\theta)$, the corresponding QDE, denoted $\hat{\theta}^*$, has the property

$$\sqrt{n}(\hat{\theta}^* - \theta_0) \xrightarrow{\mathcal{L}} N(0, (S'\Sigma^{-1}S)^{-1}).$$

We will focus on $\hat{\theta}^*$, the most efficient QDE in this class of quadratic distance and note that the choice of points $t_1, \ldots, t_k$ affects $(S'\Sigma^{-1}S)^{-1}$. Consequently, we study the choice of $t_1, \ldots, t_k$ to construct $\hat{\theta}^*$ which are highly efficient, by comparing its variances with the Cramér-Rao lower bounds. Results on the choice of points $t_1, \ldots, t_k$ with k finite are still not available in the literature and the question of choice of a finite number of points $t_1, \ldots, t_k$ to yield high efficiency is important for practitioners who want to use these estimators.

4.2 Efficiency

In Theorem 4.1, we will prove that theoretically the QDE can have very high efficiency; see Feuerverger and McDunnough (1984) for another approach. Before that, we note that since θ_2 is a location parameter, the matrix $(S'\Sigma^{-1}S)^{-1}$ is independent of θ_2; consequently the asymptotic variance-covariance matrix of the QDE $\hat{\theta}^*$, $\text{Var}[\sqrt{n}(\hat{\theta}^* - \theta_0)] = (S'\Sigma^{-1}S)^{-1}$ is independent of θ_2.

Theorem 4.1: The QDE attains full efficiency if $\{t_1, \ldots, t_k\}$ becomes dense with $k \to \infty$.

Proof: Let $l(\theta) = [l_1(\theta), l_2(\theta)]'$ be the individual score function of the positive stable law with

$$l_i(\theta) = \frac{\partial}{\partial \theta_i} \ln f(x; \theta_i).$$

Let $Z = [Z_1, \ldots, Z_k]' = [X^{-t_1} - \psi_\theta(t_1), \ldots, X^{-t_k} - \psi_\theta(t_k)]'$; we want to choose linear functions of Z to approximate $l_1(\theta), l_2(\theta)$ in the mean square error sense. So we want to determine the vectors $a = (a_1, \ldots, a_k)'$ and $b = (b_1, \ldots, b_l)'$ to minimize

$$Q_1(a) = E(l_1(\theta) - a'Z)^2$$

and

$$Q_2(b) = E(l_2(\theta) - b'Z)^2$$

The system of equations

$$\frac{\partial Q_1}{\partial a_1} = 0, \ldots, \frac{\partial Q_1}{\partial a_k} = 0$$

yields

$$E[(a'Z)Z_1] = E[l_1(\theta)Z_1]$$

$$\vdots$$

$$E[(a'Z)Z_k] = E[l_1(\theta)Z_k],$$

or equivalently

$$\Sigma a = \frac{\partial Z}{\partial \theta_1}, \quad \text{since } \frac{\partial}{\partial \theta_1}\psi_\theta(t_j) = E[l_1(\theta)Z_j], \quad j = 1, \ldots, k,$$

and Σ is the variance-covariance matrix of Z.

Similarly, the system of equations

$$\frac{\partial Q_2}{\partial b_1} = 0, \ldots, \frac{\partial Q_2}{\partial b_k} = 0$$

yields

$$\Sigma b = \frac{\partial Z}{\partial \theta_2}.$$

So

$$a^{*\prime} Z = \frac{\partial Z'}{\partial \theta_1} \Sigma^{-1} Z$$

and

$$b^{*\prime} Z = \frac{\partial Z'}{\partial \theta_2} \Sigma^{-1} Z$$

are the best linear functions of Z to approximate the score functions $l_1(\theta), l_2(\theta)$, or $S'\Sigma^{-1}Z$ is the best linear vector function to approximate $l(\theta) = [l_1(\theta), l_2(\theta)]'$. The covariance matrix of $S'\Sigma^{-1}Z$ is $(S'\Sigma^{-1}S)$ and its inverse is the variance-covariance matrix of $\sqrt{n}(\hat{\theta}^* - \theta_0)$, where $\hat{\theta}^*$ is the QDE using $Q = \Sigma^{-1}(\theta)$.

We can view $\hat{\theta}^*$ as being obtained using the approximate score $S'\Sigma^{-1}Z$ which are linear combinations of Z. Since it is based on an approximate score function, $\hat{\theta}^*$ can also be viewed as a form of quasi likelihood estimator (see Heyde (1997) for discussions on quasi score estimation).

Now using the result of Brockwell and Brown (1981), the space spanned by $\{X^{-t_1} - \psi_\theta(t_1), \ldots, X^{-t_k} - \psi_\theta(t_k)\}$ is complete with respect to $l(\theta)$, which means that $S'\Sigma^{-1}Z$ tend to the score function $l(\theta)$ in the mean square error sense, by letting $\{t_1, \ldots, t_k\}$ become dense with $k \to \infty$. Consequently, $\hat{\theta}^*$ attains high efficiency with a proper choice of $t_1, \ldots, t_k$.

□

As suggested by Theorem 4.1, in general the efficiency of the QDE is improved by increasing the number of points t_i. However, from a practical point of view, by choosing more points, the matrix Σ can become near singular and computational problems might arise.

In general, by including points t_i near 0, we gain more efficiency for values of α near 0. So, after extensive numerical computations with MATHEMATICA involving many choices of $\{t_1, \ldots, t_k\}$ for various values of the points t_j and different number of points k, we can recommend that the choices

$$A : t_1 = 0.1, t_2 = 0.2, \ldots, t_{20} = 2 \quad (k = 20),$$

$$B : t_1 = 0.05, t_2 = 0.1, \ldots, t_{20} = 1 \quad (k = 20)$$

$$C : t_1 = 0.1, t_2 = 0.2, \ldots, t_{30} = 3 \quad (k = 30)$$

are among the best, with the matrix Σ^{-1} being relatively simple to obtain.

As we can see from Tables 2, 3 and 4, the QDE based on these choices are almost as efficient as the MLE for practically all the values of the parameter space, and are much more efficient than the moment-cumulant estimators, especially for larger values of α. Note that for $\alpha = 0.9$, with values of $t_i \in B$, the efficiency of $\hat{\theta}^*$ is not as good as with values of $t_i \in A$ or $t_i \in C$ (it is worst for $\hat{\theta}_1^*$ than $\hat{\theta}_2^*$).

5 HYPOTHESIS TESTING

5.1 Simple hypothesis

It is natural to use the quadratic distance to construct test statistics for testing the simple null hypothesis
$H_0 : X_1, \ldots, X_n$ come from a specified positive stable law with Laplace transform

$$L_X(t) = \exp(-c_0 t^{\alpha_0}), \quad \theta_2^0 = (1/\alpha_0) \ln c_0, \quad \theta_1^0 = (1/\alpha_0).$$

Since $\theta_0 = [\theta_1^0, \theta_2^0]'$ is specified, the test statistics

$$nd(\theta_0) = n[Z_n - Z(\theta_0)]'\Sigma^{-1}(\theta_0)[Z_n - Z(\theta_0)]$$

can be used.

It follows from results in Luong and Thompson (1987), that $nd(\theta_0) \xrightarrow{\mathcal{L}} \chi^2_k$, that is a chi-squared test based on the value of the test statistic $nd(\theta_0)$ can be performed.

5.2 Composite hypothesis

To test the composite hypothesis $H_0 : X_1, \ldots, X_n$ come from the family of positive stable laws with Laplace transform

$$L_X(t) = \exp[-ct^\alpha],$$

where the values of the parameters are not specified, we should first calculate the QDE $\hat{\theta}^*$ by minimizing

$$d(\theta) = [Z_n - Z(\theta)]'\Sigma^{-1}(\theta)[Z_n - Z(\theta)]$$

with respect to θ, or an equivalent expression where $\Sigma^{-1}(\theta)$ is replaced by a consistent estimate. The test statistics

$$nd(\hat{\theta}^*) = n[Z_n - Z(\hat{\theta}^*)]'\Sigma^{-1}(\hat{\theta}^*)[Z_n - Z(\hat{\theta}^*)],$$

where $\Sigma^{-1}(\hat{\theta}^*)$ can be replaced by another estimate of $\Sigma^{-1}(\theta)$ if desired, follows an asymptotic distribution χ^2_{k-2}. Similar to the case for the simple null hypothesis, a chi-squared test can be performed using Luong and Thompson (1987).

6 NUMERICAL ALGORITHMS

In this section, we present numerical algoritms to compute the QDE. If a minimization subroutine is not available, an adapted Gauss-Newton method presented below can be used. The method is relatively simple to implement and involves solving linear equations or performing a series of linear regressions. The method involves first-order derivatives matrices. For the traditional Newton-Raphson algorithm, the second-order derivatives matrices which are complicated are needed. See Amemiya (1983) and Wedderburn (1974) for discussions on the Gauss-Newton method. It is widely used to compute nonlinear least-squares estimator or estimator in generalized linear models as given in McCullagh and Nelder (1989).

6.1 Gauss-Newton algorithm

The Gauss-Newton algorithm is presented here for computing the QDE with a general $Q(\theta)$ matrix.

For the QDE with weight matrix $Q(\theta)$, $[\hat{\theta}_1, \hat{\theta}_2]'$ minimizes the distance

$$d(\theta) = [Z_n - Z(\theta)]' Q(\theta) [Z_n - Z(\theta)].$$

Let us define $\psi_\theta(t) = e^{-\theta_2 t}\Gamma(1 + t\theta_1)/\Gamma(1 + t)$, and let us suppose that we have at step i estimate $\hat{\theta}^{(i)} = [\hat{\theta}_1^{(i)}, \hat{\theta}_2^{(i)}]$ for $[\hat{\theta}_1, \hat{\theta}_2]$; using a Taylor's series expansion argument around $[\hat{\theta}_1^{(i)}, \hat{\theta}_2^{(i)}]$, we get

$$\psi_\theta(t) = \psi_{\hat{\theta}^{(i)}}(t) + \left(\frac{\partial \psi_\theta(t)}{\partial \theta_1}, \quad \frac{\partial \psi_\theta(t)}{\partial \theta_2} \right)\Bigg|_{\theta = \hat{\theta}^{(i)}} \begin{pmatrix} \theta_1 - \hat{\theta}_1^{(i)} \\ \theta_2 - \hat{\theta}_2^{(i)} \end{pmatrix}$$

or $\psi_\theta(t) = \psi_{\hat{\theta}^{(i)}}(t) + \left(\frac{\partial \psi_\theta(t)}{\partial \theta_1} \right)\Big|_{\theta = \hat{\theta}^{(i)}} (\theta_1 - \hat{\theta}_1^{(i)}) + \left(\frac{\partial \psi_\theta(t)}{\partial \theta_2} \right)\Big|_{\theta = \hat{\theta}^{(i)}} (\theta_2 - \hat{\theta}_2^{(i)})$.

Now we define

$$X = X(\theta) = \begin{pmatrix} \frac{\partial \psi_\theta(t_1)}{\partial \theta_1} & \frac{\partial \psi_\theta(t_1)}{\partial \theta_2} \\ \vdots & \vdots \\ \frac{\partial \psi_\theta(t_k)}{\partial \theta_1} & \frac{\partial \psi_\theta(t_k)}{\partial \theta_2} \end{pmatrix}$$

a matrix of dimension $k \times 2$ where

$$\frac{\partial \psi_\theta(t)}{\partial \theta_2} = -t e^{-\theta_2 t}\Gamma(1 + t\theta_1)/\Gamma(1 + t)$$

and

$$\frac{\partial \psi_\theta(t)}{\partial \theta_1} = te^{-\theta_2 t}\Gamma'(1+t\theta_1)/\Gamma(1+t),$$

where $\Gamma'(x)$ is the derivative of $\Gamma(x)$.

Using $\psi(x) = \Gamma'(x)/\Gamma(x)$, we obtain

$$\frac{\partial \psi_\theta(t)}{\partial \theta_1} = te^{-\theta_2 t}\psi(1+\theta_1)\Gamma(1+t\theta_1)/\Gamma(1+t).$$

Some package, such as MATHEMATICA, has prewritten routine for calculating $\psi(\alpha)$.

Let $X^{(i)} = X(\hat{\theta}^{(i)})$,

and $Y^{(i)} = \begin{bmatrix} Y_1^{(i)} \\ \vdots \\ Y_k^{(i)} \end{bmatrix}$; the approximation at step $i+1$ of the QDE denoted by $[\hat{\theta}_1^{(i+1)}, \hat{\theta}_2^{(i+1)}]$ minimizes

$$[Y^{(i)} - X^{(i)}\theta]' Q^{-1}(\hat{\theta}^{(i)})[Y^{(i)} - X^{(i)}\theta] \simeq d(\theta).$$

The updated formula is given by

$$\hat{\theta}^{(i+1)} - \hat{\theta}^{(i)} = [X'^{(i)} Q(\hat{\theta}^{(i)}) X^{(i)}]^{-1} X'^{(i)} Q(\hat{\theta}^{(i)}) Y^{(i)},$$

or

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} + [X'^{(i)} Q(\hat{\theta}^{(i)}) X^{(i)}]^{-1} X'^{(i)} Q(\hat{\theta}^{(i)}) Y^{(i)}.$$

The procedure is repeated and updated until convergence; then we obtain the QDE $[\hat{\theta}_1, \hat{\theta}_2]'$ numerically. Often with an initial consistent estimate, the procedure converges quite fast; good starting consistent estimators are given by the moment estimators.

6.2 Numerical illustrations

Kanter (1975) developed a method to generate observations from a stable continuous distribution. We used the simple algorithm presented in Devroye

(1993) where the scale parameter is equal to 1. Remember from section 2 that the information matrix is independent of the scale parameter, i.e. of θ_2.

Let U be a uniform distribution on [0,1] and E be an exponential distribution of mean 1, where U and E are independent. Then

$$S_{\alpha,1} = \left(\frac{\sin(1-\alpha)\pi U}{E \times \sin(\alpha\pi U)}\right)^{\frac{1-\alpha}{\alpha}} \left(\frac{\sin(\alpha\pi U)}{\sin(\pi U)}\right)^{1/\alpha}$$

follows a positive stable distribution with parameters $(\alpha, c = 1)$ and Laplace transform

$$L_X(t) = \exp(-t^\alpha)$$

When the parameter $\alpha = 0.5$, giving $\theta_1 = \alpha^{-1} = 2$, $S_{\alpha,1}$ simplifies to

$$S_{0.5,1} = (1/E)\left(\frac{\sin(0.5\pi U)}{\sin(\pi U)}\right)^2.$$

We generated a sample of 1000 observations from the stable distribution with parameters $\theta_1 = 2$ and $\theta_2 = 0$. As is typical with the stable distribution, some extreme values were observed: the maximum in the sample is 94147.3, while the minimum is 0.0491997. After taking the logarithm of the data, we obtain the sample values

$$\bar{k}_1 = 0.582418, \qquad \bar{k}_2 = 4.46065,$$

from which we directly get the moment-cumulant estimates

$$\bar{\theta}_1 = 1.92659, \qquad \bar{\theta}_2 = 0.0475757.$$

We calculated the QDE $\hat{\theta}$ with various choices for the matrix $Q(\theta)$, the points $t_1, \ldots, t_k$ and the the number of points k.

Table 5 contains the estimates of the QDE; using the moment-cumulant estimates as initial values, the results were instantaneously obtained with the FindMinimum procedure in MATHEMATICA.

We notice from Table 5 that if the range of the t_j's is the same, using a step size of 0.05 or 0.10 produces almost the same results.

Table 6 contains the estimate and estimated standard deviation (in brackets) of the moment-cumulant estimator $\bar{\theta}$ (standard deviation from Table 1 with $\alpha = 0.5$ and $n = 1000$), $\hat{\theta}_A$ obtained with $Q(\theta) = I$ for $t_j \in A$, and $\hat{\theta}^*_A$ obtained with $\hat{Q}(\theta) = \Sigma^{-1}(\hat{\theta}_A)$ for $t_j \in A$. For the parameter θ_1, the best estimator is $\hat{\theta}^*_A$ (smallest bias and standard deviation) and the worst is $\bar{\theta}$ (largest bias and standard deviation), while for parameter θ_2, the standard deviations of the three estimators are comparable, but the bias is the smallest for $\hat{\theta}^*_A$ and the largest for $\bar{\theta}$.

7 CONCLUSION

In this paper, we have presented a method which is simple, to estimate the parameters of the continuous stable positive distribution, by minimizing the distance between the theoretical negative moments and their empirical counterparts. We have shown that it produces consistent, efficient estimators which have an asymptotic normal distribution, from which confidence intervals can be constructed. Using the calculated value of the minimum distance, hypothesis testing on the value of the parameters can be performed at the same time, without having to calculate the value of the Pearson's χ^2 goodness-of-fit test statistics. This would be difficult since the stable distribution does not have a closed-form expression for its density or its cumulative distribution function.

Further improvements of the method might involve the use of a continuous collection of moments, but it will be technically difficult, and possibly discourage applied scientists to use the positive stable laws.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the Committee for Knowledge Extension Research of the Actuarial Foundation, and the support of the Natural Sciences and Engineering Research Coucil of Canada. They also would like to thank Dominique Chabot-Hallé and Alexandre Girouard for computational assistance.

REFERENCES

T. Amemiya, Advanced Econometrics. Harvard University Press, Cambridge, MA, 1985.

P.J. Brockwell and B.M. Brown, Estimation for the Positive Stable Laws. Australian Journal of Statistics. 21 (1979) 139-148.

P.J. Brockwell and B.M. Brown, High-Effciency for the Positive Stable Laws. Journal of the American Statistical Association. 76 (1981) 626-631.

L. Devroye, A triptych of discrete distributions related to the stable law. Statistics and Probability Letters. 18 (1993) 349-351.

W. Feller, An Introduction to Probability, Vol. 2. Wyley, New York, 1971.

A. Feuerverger and P. McDunnough, On Statistical transform methods and their efficiency. The Canadian Journal of Statistics. 12 (1984) 303-317.

C.R. Heathcote, The integrated squared error estimation of parameters. Biometrika. 64 (1977) 255-264.

C.C. Heyde, Quasi-likelihood and its application. Springer-Verlag, 1997.

P. Hougaard, Survival models for heterogeneous populations derived from stable distributions. Biometrika. 73 (1986) 387-396.

R.G. Jarrett, Bounds and expansions for Fisher information when the moments are known. Biometrika. 71 (1984) 101-113.

M. Kanter, Stable densities under change of scale and total variation inequalities. Annals of Probability. 3 (1975) 697-707.

A. Luong and M.E. Thompson, Minimum distance methods based on quadratic distance for transforms. The Canadian Journal of Statistics. 15 (1987) 239-251.

P. McCullagh and J.A. Nelder, Generalized linear models, Chapman and Hall, New York, 1989.

R.W.M. Wedderburn, Quasi-likelihood functions, genralized linear models and the Gauss-Newton method. Biometrika. 61 (1974) 439-447.

Table 1: $n\times$ Asymptotic variance for $\bar{\theta}_2$ and $\bar{\theta}_1$

α	Var($\bar{\theta}_2$)	Var($\bar{\theta}_1$)	CR_2	CR_1
0.1	114.88	109	108.94	58.6
0.2	27.384	26.496	25.921	13.279
0.3	11.247	11.213	10.637	5.054
0.4	5.653	5.859	5.346	2.29
0.5	3.1065	3.375	2.934	1.097
0.6	1.760	2.020	1.648	0.5169
0.7	0.979	1.196	0.8918	0.2231
0.8	0.500	0.655	0.4223	0.07805
0.9	0.196	0.277	0.1334	0.01561

Table 2: Efficiency of QDE $\hat{\theta}^*$ and $\bar{\theta}$ for values of $t_i \in A$

α	V($\hat{\theta}_1^*$)	V($\hat{\theta}_2^*$)	V($\hat{\theta}_1^*$)/CR_1	V($\hat{\theta}_2^*$)/CR_2	V($\bar{\theta}_1$)/CR_1	V($\bar{\theta}_2$)/CR_2
0.1	61.1097	109.451	1.043	1.0047	1.86	1.0545
0.2	13.2917	25.8977	1.001	0.999	1.9953	1.0564
0.3	5.06667	10.6165	1.003	0.998	2.2186	1.0570
0.4	2.28161	5.3433	0.996	1.001	2.5585	1.0574
0.5	1.13631	2.8749	1.0358	0.9798	3.0766	1.0589
0.6	0.518523	1.68588	1.003	1.0229	3.9079	1.0680
0.7	0.22177	0.877312	0.994	0.9837	5.361	1.0977
0.8	0.077463	0.416573	0.992	0.9864	8.392	1.1839
0.9	0.01633	0.136328	1.046	1.022	17.745	1.4693

Table 3: Efficiency of QDE $\hat{\theta}^*$ and $\bar{\theta}$ for values of $t_i \in B$

α	V($\hat{\theta}_1^*$)	V($\hat{\theta}_2^*$)	V($\hat{\theta}_1^*$)$/CR_1$	V($\hat{\theta}_2^*$)$/CR_2$	V($\bar{\theta}_1$)$/CR_1$	V($\bar{\theta}_2$)$/CR_2$
0.1	58.7412	108.834	1.0024	0.9990	1.86	1.0545
0.2	13.2403	26.0017	0.997	1.003	1.9953	1.0564
0.3	4.87249	10.9421	0.964	1.0287	2.2186	1.0570
0.4	2.291	5.48326	1.000	1.0257	2.5585	1.0574
0.5	1.10511	3.0491	1.0074	1.0392	3.0766	1.0589
0.6	0.5203	1.70046	1.0066	1.0318	3.9079	1.0680
0.7	0.22207	0.88409	0.9954	0.9914	5.361	1.0977
0.8	0.07974	0.42806	1.0216	1.0136	8.392	1.1839
0.9	0.02884	0.16991	1.8475	1.2737	17.745	1.4693

Table 4: Efficiency of QDE $\hat{\theta}^*$ and $\bar{\theta}$ for values of $t_i \in C$

α	V($\hat{\theta}_1^*$)	V($\hat{\theta}_2^*$)	V($\hat{\theta}_1^*$)$/CR_1$	V($\hat{\theta}_2^*$)$/CR_2$	V($\bar{\theta}_1$)$/CR_1$	V($\bar{\theta}_2$)$/CR_2$
0.1	61.244	109.479	1.0451	1.0049	1.86	1.0545
0.2	13.2859	25.9223	1.001	1.000	1.9953	1.0564
0.3	5.09473	10.7927	1.008	1.0146	2.2186	1.0570
0.4	2.27181	5.29837	0.992	0.991	2.5585	1.0574
0.5	1.09552	2.89353	0.999	0.986	3.0766	1.0589
0.6	0.5171	1.71039	1.000	1.037	3.9079	1.0680
0.7	0.219617	0.870665	0.9843	0.9763	5.631	1.0977
0.8	0.0789712	0.430226	1.0118	1.0188	8.392	1.1839
0.9	0.0159021	0.136869	1.0187	1.0260	17.745	1.4693

Table 5: Estimated values of the QDE $\hat{\theta}$ with $Q(\theta) = I$

$t_1, \ldots, t_k$	$\hat{\theta}_1$	$\hat{\theta}_2$
$A : 0.1, 0.2, \ldots, 2$	1.9494	0.028844
$C : 0.1, 0.2, \ldots, 3$	1.9294	-0.00185674
$0.1, 0.2, \ldots, 4$	1.8948	-0.0667758
$0.1, 0.2, \ldots, 5$	1.85129	-0.158828
$B : 0.05, 0.10, \ldots, 1$	1.9538	0.0377184
$0.05, 0.10, \ldots, 2$	1.94965	0.0291619
$0.05, 0.10, \ldots, 3$	1.92992	-0.00096794
$0.05, 0.10, \ldots, 5$	1.85212	-0.156985
1, 2	1.95062	0.0308824
$1, 2, \ldots, 5$	1.85972	-0.139435

Table 6: Estimated values of θ with various methods

Estimator	estimate of θ_1 (s.e.)	estimate of $\hat{\theta}_2$ (s.e.)
$\bar{\theta}$	1.92659 (0.05810)	0.0475757 (0.055736)
$\hat{\theta}_A$	1.9494 (0.0424262)	0.028844 (0.0594281)
$\hat{\theta}_A^*$	1.95244 (0.0322591)	0.0191101 (0.0544164)