Commun Nonlinear Sci Numer Simulat 17 (2012) 5088–5096

Contents lists available at SciVerse ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Stable and generalized-*t* distributions and applications

P.N. Rathie *, M. Coutinho, T.R. Sousa, G.S. Rodrigues, T.B. Carrijo

Department of Statistics, University of Brasília, 70910-900 Brasília, D.F., Brazil

ARTICLE INFO

Article history:
Received 24 December 2011
Received in revised form 9 May 2012
Accepted 9 May 2012
Available online 24 May 2012

Keywords:
Stable and generalized-*t* distributions
χ^2-Divergence
Time series analysis
Financial data
Sunspots
Heavy tails

ABSTRACT

In this paper a generalized-*t* distribution is introduced and used as an alternative to the symmetric stable distribution. To do so, the χ^2-divergence is presented and minimized to approximate the symmetric stable distribution, as accurately as possible, by the generalized-*t* distribution. Kth moments for the generalized-*t* distribution function are given. The stable distribution is defined in terms of generalized hypergeometric functions. Five applications with natural data (sunspots activity), and financial data (stock exchange in Brazil, South Africa and Venezuela, and daily variation of Petrobras stock market) are analyzed. A time series analysis is used to eliminate data correlation in each data set, and then the distributions are used to fit the residuals of these models.

1. Introduction

Since 1963 [9,10], stable distributions had been used to fit financial data in economics. They are valuable models for data sets covering extreme events, which usually exhibit heavy tails. The class of stable distribution was described by Lévy [7] in his investigation of the behavior of sums of i.i.d random variables.

There are several different methods to estimate the parameters. Usually, the estimation technique is based on some properties of the characteristic function, or based on numerical calculations of the stable density function. The lack of a simple closed form for the density function is problematic for the task of parameter estimation, but fortunately there are now several methods to carry out the estimation with real life problems. The maximum likelihood procedure proposed by Nolan [15] is usually the most accurate method, and has been widely used by researchers when working with stable distribution estimation.

In this paper a generalized-*t* distribution is introduced, and used as an alternative to the symmetric stable distribution. To do so, the χ^2-divergence is presented and minimized to approximate the symmetric stable distribution, as accurately as possible, by the generalized-*t* distribution. It can be shown that there is not much difference between these distributions for a set of parameters; this means we can use the generalized-*t* distribution instead of the stable distribution in the symmetric case.

Finally, five practical applications with real data are presented using the generalized-*t* distribution, and the asymmetric stable distribution. When the data do not show strong skewness, it could be noticed that the generalized-*t* distribution can be used instead of the stable distribution.

* Corresponding author.
E-mail addresses: rathie@unb.br, pushpanrathie@yahoo.com (P.N. Rathie).

1007-5704/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cnsns.2012.05.013

2. Materials and methods

To compare the generalized-t distribution and the stable distribution we used the five datasets listed below:

1. Data from the 2007 stock exchange in Brazil.
2. Data from the 2001 stock exchange in South Africa.
3. Data from the 2000 stock exchange in Venezuela.
4. Daily variation of Petrobras stock market.
5. Average number of sunspots observed each year during 1990 to 2008.

These databases have strong autocorrelation, so a time series analysis was performed in each case before modeling the data with the stable and generalized-t distributions. For more information on time series analysis, refer to [2,5].

We used the program called STABLE developed by Nolan [14] to estimate the parameters of the stable distribution. No optional feature was used, and direct search for Maximum Likelihood Estimator (MLE) was performed. In short, this program uses the McCulloch [12] quantile estimator as an initial approximation to the parameters and then a constrained quasi-Newton method is applied to perform the maximization. For more details, see Nolan [15]. To estimate the parameters of the generalized-t distribution, the log-likelihood function was maximized computationally.

3. Theory and calculation

In this section, we present the generalized-t distribution and the stable distribution. Also, we use the χ^2 divergence to approximate, for a set of parameters, the stable distribution by the generalized-t distribution.

3.1. Generalized-t distribution

In this section, we define the generalized-t distribution and obtain its k-th absolute moments and likelihood function. A random variable X has the generalized-t distribution if its density function is given by [19]

$$f_t(x;a,b,d) = \frac{c}{(1+a|x|^b)^d}, \quad a,b,d>0, \quad bd>1, \quad -\infty<x<\infty, \tag{1}$$

where

$$c = \frac{a^{1/b}b}{2B(1/b, d-1/b)} \tag{2}$$

and $B(\cdot)$ represents a beta function.

As the distribution is symmetric about the origin, all odd absolute moments are zeros. The k-th absolute moment about the origin is given by

$$E(|X|^k) = \frac{B\left(\frac{k+1}{b}, d-\frac{k+1}{b}\right)}{a^{\frac{k}{b}}B(1/b, d-1/b)}, \tag{3}$$

for $0<k<bd-1$.

The parameters a, b, and d of expression (1) can be estimated by the method of maximum likelihood. The log-likelihood function is given by

$$\ln L(a,b,d|\mathbf{x}) = n\left[\frac{\ln a}{b} + \ln b - \ln 2 - \ln B\left(\frac{1}{b}, d-\frac{1}{b}\right)\right] - d\sum_{i=1}^{n}\ln\left(1+a|x_i|^b\right). \tag{4}$$

Computationally, it is possible to determine the values of a, b, and d that maximize (4). For this, one can apply any suitable multivariate optimization algorithm, such as Nelder and Mead [13], utilizing a proper statistical or mathematical software. Initial values for the parameters are usually required, and play an important role in the maximization process. Depending on the chosen method, the algorithm can return a local, but not global, maximum. It is recommended to repeat the procedure based on different starting points, and then compare the results for validation. In this paper, we have used the function *Optim*, in the statistical software *R* [23], to execute the maximization process.

3.2. Symmetric and asymmetric Lévy distribution

The class of stable distribution was defined by Lévy in 1924 [7] in his investigation of the behavior of sums of i.i.d random variables. A random variable X is said to be stable if its characteristic function is given by

$$E(e^{i\theta X}) = exp\{-|\sigma\theta|^\alpha[1-i\beta sgn(\theta)w(\alpha,\theta)+i\mu\theta]\}, \tag{5}$$