Physica A 486 (2017) 628–637

Contents lists available at ScienceDirect

Physica A

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

Bivariate sub-Gaussian model for stock index returns

Matylda Jabłońska-Sabuka [a], Marek Teuerle [b], Agnieszka Wyłomańska [b,*]

[a] Lappeenranta University of Technology, PO Box 20, 53851 Lappeenranta, Finland
[b] Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wroclaw University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

HIGHLIGHTS

- Bivariate α-stable sub-Gaussian random vectors are considered.
- New estimation parameters' estimation method is proposed.
- The new testing procedure is introduced.
- The proposed methods are used for simulated as well as real financial time series.

ARTICLE INFO

Article history:
Received 22 December 2015
Received in revised form 9 May 2017
Available online 12 June 2017

Keywords:
Nonparametric methods
Characteristic function
Bivariate sub-Gaussian distribution
α-stable process

ABSTRACT

Financial time series are commonly modeled with methods assuming data normality. However, the real distribution can be nontrivial, also not having an explicitly formulated probability density function. In this work we introduce novel parameter estimation and high-powered distribution testing methods which do not rely on closed form densities, but use the characteristic functions for comparison. The approach applied to a pair of stock index returns demonstrates that such a bivariate vector can be a sample coming from a bivariate sub-Gaussian distribution. The methods presented here can be applied to any nontrivially distributed financial data, among others.

1. Introduction

A large part of currently performed research in many fields relies heavily on proper model identification, whether it concerns model type or its order, see [1]. That applies to both deterministic and stochastic problems. For instance, in deterministic case a wrong choice of model family can lead to a very good fit to the data but accompanied by a very misleading out-of-sample forecast. Also, incorrect choice of model order within one family may cause a completely erroneous extrapolation out of the available data set.

Similar problems may occur in stochastic approaches. The key to modeling time series is the proper identification of the distribution that drives the data. Many classical models, e.g. in finance, assume data or the model error to be normally distributed. However, the reality is far from perfect, and the range of possible distributions to be considered is vast. Therefore, creating appropriate and universal testing tools for stochastic processes is of researchers' interest. Only the right knowledge of the true distribution of the considered data can allow correct estimation of model parameters. And only then one can actually perform simulations, for instance, to replace costly real experiments in engineering, physics, mining, and many

* Corresponding author.
E-mail addresses: matylda.jablonska-sabuka@lut.fi (M. Jabłońska-Sabuka), marek.teuerle@pwr.edu.pl (M. Teuerle), agnieszka.wylomanska@pwr.edu.pl (A. Wyłomańska).

http://dx.doi.org/10.1016/j.physa.2017.05.080
0378-4371/

more fields as discussed by [2,3], among others. Monte Carlo simulations for risk analysis are of no value then the assumption about the underlying process is incorrect.

Current approaches make their best effort not to base tests on the assumption that we know the explicit formula of the cumulative distribution function (CDF), as for many families of distributions a closed form of CDF is not available. For instance, [4] have introduced a non-parametric approach for generic diffusion hypothesis. Simulation studies have shown that the test performed acceptably well in finite samples, both uni- and multivariate. Another examples are testing procedures based on the distribution property (but not explicit form of CDF). For instance, [5] proposed a simple test based on the property of α-stable distribution such that the sum of independent α-stable random variables still has stable distribution and the index of stability has the same value. See also [6].

In this work we present a testing procedure for bivariate stock index data verifying that they represent the sub-Gaussian family of distributions. The same data has been previously studied by [7,8], where authors were illustrating methods of general multivariate stable vectors' estimation without proposing a testing procedure for the goodness-of-fit. The main advantage of the testing procedure presented in our work is that it relies on distribution's characteristic function rather than closed form CDF. A similar method was proposed in one-dimensional case for testing tempered-stable distribution of indoor air quality parameters, see [1]. It is worth to mention, the sub-Gaussian family of distributions belong to the $\alpha-$stable family. Since Mandelbrot introduced the α-stable distribution in modeling of financial asset returns [9,10], numerous empirical studies have been done in both natural and economic sciences. The works of Rachev and Mittnik [11] and Rachev et al. [12] have focused attention on a general framework for market and credit risk management, option pricing, and portfolio selection based on the α-stable distribution. The heavy tailed or leptokurtic character of the distribution of price changes has been observed in various markets, see [12–16], therefore many research papers concentrate on the problem of financial data analysis by using heavy tailed-based models, [17–21].

Stock market prices and price indices are found among some of the most commonly studied financial time series. That is not surprising, considering the huge volumes of trades performed every day in stock markets across the world. For instance, [22] demonstrated that some of the asymmetric leptokurtic features presented in some of the Chinese stock index returns can be captured by an α-stable law. Other works look for bivariate dependencies between stocks and other commodity prices, like oil, see [23]. Some studies on bivariate data sets have been done as well. In [24] the authors investigated a combination of VARMA(u, v)-EGARCH(p, q) models with normal errors and a constant conditional correlation using MSCI domestic and world-ex-domestic index pairs for the Emu, Japan, the United Kingdom, and the United States. [25] have also compared the relative importance of the conditional mean and volatility components of a VAR-EGARCH model of bivariate returns to international stock market indices and found that an EGARCH-only model without a conditional mean specification yields better-behaved residuals and no less accurate return and volatility predictions than a VAR-EGARCH model.

In this work, we verify that other than normal distribution families should be considered when analyzing multivariate index data. As our results demonstrate, the given index pair can be described though a sub-Gaussian distribution. Our study proposes a new estimation method of sub-Gaussian distribution based on the form of the characteristic function as well as the testing procedure if given multivariate vector of observations constitute a sample from this distribution. Our theoretical results are confirmed by simulation study and real data analysis.

There are many papers where authors analyze the empirical characteristic function as the main statistic used for parameters estimation and testing [26–29]. However, to the best of our knowledge, this is the first time that the distance between empirical and theoretical characteristic functions is used in order to estimate the parameters of sub-Gaussian bivariate distribution. We also provide a goodness-of-fit testing procedure for the sub-Gaussian vectors, which is also novel. It is based on bootstrap methodology and we use the empirical and theoretical characteristic functions as the counterpart of the Kolmogorov–Smirnov test statistics. This is contrary to the classical version of the Kolmogorov–Smirnov statistic, where the test function is a distance between empirical and theoretical cumulative distribution functions. In the literature there is limited number of papers where the bivariate (or multivariate) sub-Gaussian distribution is considered however there is almost no results concerning goodness-of-fit test for such distribution.

This article is structured as follows. Section 2 introduces the α-stable vectors and their special case, sub-Gaussian vectors, together with the related main definitions. Next, in Section 3 we present the estimation procedure for bivariate vectors of observations from sub-Gaussian distribution. Section 4 sketches the testing procedure for verifying whether bivariate observations can be considered as a vector from sub-Gaussian distribution. Section 5 presents the simulation study for introduced estimation and testing procedures, while in Section 6 we examine a real bivariate data set. Section 7 concludes.

2. Multidimensional α-stable random vectors

In general, the α-stable random variables play an important role in the probability theory. Namely, the classical central limit theorem, see [30] states that the only possible limit for the appropriately shifted and renormalized sum of identical and independently distributed (i.i.d.) sequence of random variables with finite mean and variance is a Gaussian (normal) random variable. In a more generalized framework, it appears that α-stable random variables are the only possible non-degenerate limits for renormalized sum of any i.i.d. sequences, see [30,31].