Modeling financial returns by discrete stable distributions

Lenka Slámová[1], Lev B. Klebanov[2]

Abstract. This paper develops a new approach to modeling financial returns by introducing discrete stable distributions. It is well known that the financial returns are not normally distributed, extremal events occur more often than the Gaussian distribution suggests. Already in the sixties Benoit Mandelbrot suggested a hypothesis that returns follow a stable Paretian law. Inspired by the discrete nature of prices appearing on the markets we model the financial returns by discrete analogues of absolutely continuous stable distributions. The known discrete stability of random variables on $\mathbb{N}$ is generalized to the case of random variables on $\mathbb{Z}$. We give brief introduction to the theory of discrete stability on $\mathbb{Z}$, show connection of discrete stable random variables to their absolutely continuous counterparts and focus mainly on methods of estimation of parameters of these distributions from the real data of financial returns.

Keywords: discrete stable distributions, parameter estimation, M-estimator

JEL classification: C13
AMS classification: 60E07

1 Introduction

The observation that financial returns have heavy tails and a hypothesis that they follow stable Paretian law were suggested by Benoît Mandelbrot. Since then many mathematicians and economists were studying the implications of this hypothesis, and stable distributions became a very important field of probability for both theoreticians and practitioners.

The stable Paretian hypothesis assumes that the financial returns are continuous random variables, which is the case of logarithmic returns ($\log(V_1/V_0)$) or arithmetic returns ($(V_1 - V_0)/V_0$). However sometimes it is more convenient to consider simple returns computed as a difference of buy price and sell price. These prices are quoted on the market on a discrete grid (sometimes called ticks). Then the returns are discrete random variable and thus the assumption they follow a stable Paretian law is incorrect. A discrete distribution that allows for heavy tails and have the stability property is needed for modeling such returns.

The notion of discrete stability for lattice random variables on positive integers was introduced in [5] and further studied in [1]. In [4] an extension of discrete stability for random variables on $\mathbb{Z}$ was proposed. A discrete analogue of stable Paretian distribution was introduced and a connection to absolutely continuous stable distributions was shown.

In this article we propose a method of parameter estimation for discrete stable distribution family and compare it with a well known empirical characteristic function method, that was reviewed in [6]. We illustrate the quality of both methods by estimating parameters of simulated data. Finally we do an empirical study on market data of futures prices and compare the performance of the fit with the normal and continuous stable distribution fit.

[1]Charles University in Prague, Department of probability and mathematical statistics, Sokolovská 83, 18675 Praha 8, e-mail: slamova.lenka@gmail.com
[2]Charles University in Prague, Department of probability and mathematical statistics, Sokolovská 83, 18675 Praha 8, e-mail: klebanov@chello.cz

2 Discrete stable distributions

A random variable X is strictly stable if for all $n \in \mathbb{N}$ there is a constant a_n such that $a_n \sum_{i=1}^{n} X_i \stackrel{d}{=} X$, where X_i's are independent copies of X. When defining stability for discrete random variables one has to choose a different normalization, because the normalized sum $a_n \sum_{i=1}^{n} X_i$ is not generally integer valued.

A definition of positive discrete stable random variables was introduced in [5], where a normalization $\sum_{i=1}^{n} \tilde{X}_i(p_n) \stackrel{d}{=} X$ was used, with $\tilde{X}(p_n) = \sum_{j=1}^{X} \varepsilon_j$, where $\mathbb{P}(\varepsilon_i = 1) = 1 - \mathbb{P}(\varepsilon_i = 0) = p_n$ with $p_n \in (0,1)$. A positive random variable X with characteristic function

$$f(t) = \exp\left\{-\lambda(1 - e^{\mathrm{i}t})^{\gamma}\right\}, \quad \lambda > 0, \gamma \in (0,1]$$

is called *positive discrete stable* and is denoted by $\mathscr{PDS}(\gamma,\lambda)$. This distribution has its support in $\mathbb{N}_0$ and in case of $\gamma = 1$ corresponds to Poisson distribution. [1] showed that positive discrete stable distribution with $\gamma < 1$ belongs to the domain of normal attraction of the absolutely continuous positive stable distribution (stable distribution with index of stability $\alpha = \gamma$ and with skewness parameter $\beta = 1$), whose Laplace transform is given by $\exp(-\lambda t^{\gamma})$. The case of $\gamma = 1$ is a degenerate one, where the normalized sum of Poisson random variables converges to a constant λ.

The discrete stability for random variables on $\mathbb{Z}$ was defined in [4]. The first approach to symmetric discrete random variables uses the following normalization:

$$\sum_{j=1}^{n} \tilde{X}_j(p_n) \xrightarrow{d} X, \quad \text{where} \quad \tilde{X}(p_n) = \sum_{i=1}^{|X|} \varepsilon_i(p_n)$$

and $\varepsilon_i(p_n)$ are taking values ± 1 with probability p_n and 0 with probability $1 - 2p_n$, with $p_n \downarrow 0$. This normalization leads to a distribution with characteristic function

$$f(t) = \exp\left\{-\lambda\left(1 - \frac{1}{2}\left(e^{\mathrm{i}t} + e^{-\mathrm{i}t}\right)\right)^{\gamma}\right\} \quad \text{with} \quad \lambda > 0, \gamma \in (0,1].$$

Such a distribution is then called *symmetric discrete stable* and is denoted by $\mathscr{SDS}(\gamma,\lambda)$. This distribution belongs to the domain of normal attraction of the absolutely continuous symmetric stable distribution with index of stability $\alpha = 2\gamma$. Hence the case $\gamma = 1$ can be considered as a discrete version of Gaussian distribution.

The general case of discrete random variables on $\mathbb{Z}$ requires once again a different normalization, namely

$$\sum_{j=1}^{n} \tilde{X}_j(p_n^1, p_n^2) \xrightarrow{d} X, \quad \text{where} \quad \tilde{X}(p_n^1, p_n^2) = \begin{cases} \sum_{i=1}^{X} \varepsilon_i(p_n^1), & X \geq 0 \\ -\sum_{i=1}^{|X|} \varepsilon_i(p_n^2), & X < 0 \end{cases}$$

and $\varepsilon_i(p)$ are taking values 1 with probability p and 0 with probability $1 - p$, where $p_n^i \downarrow 0$. The characteristic function of such distribution takes the following form

$$f(t) = \exp\left\{-\lambda_1\left(1 - e^{\mathrm{i}t}\right)^{\gamma} - \lambda_2\left(1 - e^{-\mathrm{i}t}\right)^{\gamma}\right\}, \quad \text{with} \quad \lambda_1, \lambda_2 > 0, \gamma \in (0,1].$$

Such a distribution is called *discrete stable distribution* and is denoted by $\mathscr{DS}(\gamma,\lambda_1,\lambda_2)$. The discrete stable distribution for $\gamma < 1$ is in the domain of normal attraction of absolutely continuous stable distribution with index of stability $\alpha = \gamma$ and for $\gamma = 1$ in the domain of normal attraction of symmetric stable distribution with index of stability $\alpha = 2$.

3 Parameter estimation

Statistical methods of estimation in case of discrete stable distributions has several problems that inhibit to use most of the methods - like maximum likelihood or method of moments. These methods are based on assumptions like availability of a closed form of probability function or existence of moments up to some order. However this is not the case of discrete stable distributions. We describe here and use ECF method using empirical characteristic function and AML method which use M-estimator for estimation of the parameters of discrete stable distributions.

3.1 Empirical characteristic function method

This method was studied broadly as an alternative method of estimation when the cumulative distribution function is not known in a closed form, but we have a closed form for characteristic function. [3] introduced an ECF method known as "$k - L$ procedure". A review of different approaches to ECF method can be found in [6].

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions on a probability space $(\mathcal{X}, \mathcal{M})$, where the parametric space $\Theta = \mathbb{R}^d$. The characteristic function of the distribution P_θ is defined by $f(t, \theta) = \mathbb{E}_\theta \left[e^{\mathrm{i}tX} \right]$ and the empirical characteristic function (ECF) based on a sample of observations $x_1, \ldots, x_n$ is defined by

$$\widehat{f}_n(t) = \frac{1}{n} \sum_{i=1}^{n} e^{\mathrm{i}tx_i}.$$

The general idea for ECF estimation is to minimize distance between f and $\widehat{f}_n$. Since discrete stable distributions are lattice distributions, their characteristic function is periodic with period 2π, it is enough to to minimize the distance only on the interval $[-\pi, \pi]$. Choose a discrete grid of points $\{t_1, \ldots, t_k\} \in [-\pi, \pi]$ and denote by

$$V_n = \left(\widehat{f}_n(t_1), \ldots, \widehat{f}_n(t_k)\right)' \quad \text{and} \quad V_\theta = \left(f(t_1), \ldots, f(t_k)\right)'.$$

The estimate that solves $\min_\theta (V_n - V_\theta)'(V_n - V_\theta)$ corresponds to the nonlinear OLS regression of V_n of V_θ. [6] argues that this estimator is not efficient and [3] suggests to use nonlinear GLS regression as follows. Denote by Ω the covariance matrix of V_n, then an efficient estimator can be obtained as a solution of

$$\min_\theta (V_n - V_\theta)' \widehat{\Omega}^{-1} (V_n - V_\theta),$$

where $\widehat{\Omega}$ is a consistent estimate of Ω.

3.2 Approximate maximum likelihood method

Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions on a probability space $(\mathcal{X}, \mathcal{M})$, where the parametric space $\Theta = \mathbb{R}^d$. In the maximum likelihood estimation one assumes the existence of a density $p(x, \theta)$ and of a function

$$\mathbf{J}(x, \theta) = \left(\frac{\frac{\partial p}{\partial \theta_i}(x, \theta)}{p(x, \theta)} \right)_{i=1,\cdots,d}.$$

The maximum likelihood estimator θ^* of the parameter θ, given a set of n observations $x_1, \cdots, x_n$, is a solution of $\sum_{j=1}^{n} \mathbf{J}(x_j, \theta) = \mathbf{0}$. However if the density does not exist or it is not known in an analytical form, this method cannot be used. Let us denote by $\mathcal{L}_k$ a linear space generated by set of functions $\{1, \varphi_1(x), \cdots, \varphi_k(x)\}$ on $(\mathcal{X}, \mathcal{M})$. We assume we know the functionals of our distribution, namely

$$\begin{aligned} \mathbb{E}_\theta \varphi_i(X) &= \pi_i(\theta), \quad i = 0, \ldots, k, \\ \mathbb{E}_\theta \varphi_i(X)\varphi_j(X) &= \pi_{ij}(\theta), \quad i, j = 0, \ldots, k. \end{aligned}$$

Instead of the function $\mathbf{J}$ we define a new function $\widehat{\mathbf{J}}(x, \theta)$ as a projection of $\mathbf{J}$ onto the space $\mathcal{L}_k$. Then $\widehat{\mathbf{J}}$ has to take the following form

$$\widehat{\mathbf{J}}(x, \theta) = \mathbf{c}(\theta)\varphi(x),$$

where $\varphi(x) = \left(\varphi_j(x), j = 0, \ldots, k\right)$, $\mathbf{c}(\theta) = \left(c_{ij}(\theta), i = 1, \ldots, d, j = 0, \cdots, k\right)$.

Since $\widehat{\mathbf{J}}$ is a projection of $\mathbf{J}$ onto $\mathcal{L}_k$, it has to hold

$$\mathbb{E}_\theta \left(J_i(X, \theta) - \widehat{J}_i(X, \theta) \right) \varphi_j(x) = 0, \quad i = 1, \ldots, d; j = 0, \ldots, k.$$

From this it follows that it has to hold

$$\frac{\partial \pi_j}{\partial \theta_i}(\theta) = \sum_{m=0}^{k} c_{im}(\theta) \pi_{mj}(\theta), \quad i = 1, \ldots, d; j = 0, \ldots, k.$$

If we use a matrix notation, $\nabla\pi(\theta) = \mathbf{c}(\theta)\Pi(\theta)$. Hence if the inverse of $\Pi(\theta)$ exists, then we are able to compute the matrix $\mathbf{c}(\theta) = \nabla\pi(\theta)\left(\Pi(\theta)\right)^{-1}$. Now the M-estimator θ^* of the parameter θ, given n observations $x_1, \ldots, x_n$, is the solution of

$$\sum_{k=1}^{n} \widehat{\mathbf{J}}(x_k, \theta) = \mathbf{0} \quad \text{or} \quad \sum_{m=1}^{n}\sum_{j=0}^{k} c_{ij}(\theta)\varphi_j(x_m) = 0, \quad \text{for} \quad i = 1, \ldots, d.$$

The question that arises with this method is the choice of functions φ. One possibility is as follows. Choose $k \in \mathbb{N}$ and then $z_1, \ldots, z_k \in \mathbb{Z}$ and define $\varphi_i(x) = z_i^x$ for $i = 1, \ldots, k$. Then the functionals $\pi_i(\theta) = \mathbb{E}_\theta(z_i^X) = \mathcal{P}(z_i)$, where $\mathcal{P}$ is the probability generating function of our distribution. A wise choice of k and z_i's is such that the variance of the resulting estimator is minimal.

4 Simulation from discrete stable family

The simulation from positive discrete stable distribution uses the stochastic representation emphasized by [2] stating that

$$\mathscr{PDS}(\gamma, \lambda) \stackrel{d}{=} \mathscr{P}\big(\lambda^{1/\gamma}\mathscr{S}(\gamma, 1, \sigma, 0)\big),$$

where $\mathscr{P}(\lambda)$ is Poisson r.v. with parameter λ, $\mathscr{S}(\alpha, \beta, \sigma, \mu)$ is stable r.v. with index of stability α, skewness β, scale σ and location μ, and $\sigma = \left(\cos(\gamma\pi/2)\right)^{1/\gamma}$.

For simulation from discrete stable distribution we use the fact that

$$\mathscr{DS}(\gamma, \lambda_1, \lambda_2) \stackrel{d}{=} \mathscr{PDS}(\gamma, \lambda_1) - \mathscr{PDS}(\gamma, \lambda_2).$$

Symmetric discrete stable distribution have a similar stochastic representation as the positive case.

Theorem 1. *A symmetric discrete stable random variable $\mathscr{SDS}(\gamma, \lambda)$ is distributed as a compound Poisson random variable Y with intensity $\lambda^{1/\gamma}\mathscr{S}(\gamma, 1, \sigma, 0)$, jumps taking values ± 1 with the same probability and $\sigma = \left(\cos(\gamma\pi/2)\right)^{1/\gamma}$.*

Proof. The Laplace transform of a random variable $\mathscr{S}(\gamma, 1, \sigma, 0)$ is $\mathbb{E}\exp(-u\mathscr{S}(\gamma, 1, \sigma, 0)) = \exp(-u^\gamma)$. The characteristic function of a compound Poisson variable with intensity λ and jumps taking values ± 1 with the same probability is $\exp(-\lambda(1 - 1/2(e^{\mathrm{i}t} + e^{-\mathrm{i}t})))$. Then it follows easily that the characteristic function of the random variable Y is

$$\mathbb{E}\left[e^{\mathrm{i}tY}\right] = \mathbb{E}\exp\left\{-\lambda^{1/\gamma}\mathscr{S}(\gamma, 1, \sigma, 0)\left(1 - \frac{1}{2}\left(e^{\mathrm{i}t} + e^{-\mathrm{i}t}\right)\right)\right\} = \exp\left\{-\lambda\left(1 - \frac{1}{2}\left(e^{\mathrm{i}t} + e^{-\mathrm{i}t}\right)\right)^{\gamma}\right\}.$$

And in this we recognize characteristic function of $\mathscr{SDS}(\gamma, \lambda)$. □

This representation gives us a useful tool for simulation from symmetric discrete stable distribution since stable and compound Poisson random variable may be easily generated.

5 Simulation study

In this section we will compare the performance of the two methods, ECF and AML, on simulated data. We simulate samples from positive discrete stable, symmetric discrete stable and discrete stable distributions with different values of parameters. The results are in Table 1. We can see that the AML method is better for estimating the index of stability γ, however the performance of the ECF method is better with the scale parameter λ. This is evident especially in the $\mathscr{DS}$ case, where the ECF method gives a very biased estimate of γ.

$\mathscr{PDS}$	(γ, λ)	(1, 0.5)	(0.8, 2)	(0.5, 4)	(0.2, 1)
Method	ECF	(1.0044, 0.5011)	(0.8052, 2.0220)	(0.4997, 3.9903)	(0.2052, 1.0098)
	AML	(1.0000, 0.4999	(0.8043, 2.0242)	(0.4998, 3.9725)	(0.2050, 1.0093)

$\mathscr{SDS}$	(γ, λ)	(1, 0.5)	(0.8, 2)	(0.5, 4)	(0.2, 1)
Method	ECF	(1.016, 0.4995)	(0.7886, 2.0389)	(0.4968, 4.0751)	(0.2041, 0.9960)
	AML	(1.000, 0.4972)	(0.8054, 2.0810)	(0.4977, 4.0823)	(0.1997, 0.9887)

$\mathscr{DS}$	$(\gamma, \lambda_1, \lambda_2)$	(1, 1, 2)	(0.5, 3, 0.5)	(0.3, 1, 1)
Method	ECF	(0.499, 0.981, 1.965)	(0.246, 2.934, 0.467)	(0.151, 0.987, 1.027)
	AML	(1.000, 0.856, 1.876)	(0.492, 2.935, 0.477)	(0.299, 1.000, 1.004)

Table 1: Estimated parameters of $\mathscr{PDS}$, $\mathscr{SDS}$ and $\mathscr{DS}$ distribution from simulated data

6 Empirical application to market data

As we mentioned in the introduction, one of the motivation for studying discrete stable distributions is an effort to have a more appropriate tool for modeling discrete financial returns. In this section we take a look at real data and show that discrete stable distributions are able to capture the nature of the data really well. We work with Bund futures prices (i.e. futures on German government bonds with maturity 8.5 to 10.5 years) from May 2010 and we consider intraday returns in ticks over different time periods (30 seconds, 1 minute, 2 minutes, 5 minutes, 15 minutes, 30 minutes, 1 hour). Such short returns are of interest for example for market makers and high frequency trading systems. The estimated parameters of the $\mathscr{SDS}$ and $\mathscr{DS}$ distribution by the AML method are displayed in Table 2. It is interesting to notice that the returns over different periods keep the same index of stability and the scale parameter changes, what suggest that the data have the stability property.

$\mathscr{SDS}$	period	10 s	30 s	1 min	2 min
Method	AML	(0.888, 0.760)	(0.852, 1.552)	(0.834, 2.495)	(0.854, 4.551)
$\mathscr{SDS}$	period	5 min	15 min	30 min	1 hour
Method	AML	(0.836, 8.863)	(0.796, 16.303)	(0.779, 25.946)	(0.764, 40.083)

$\mathscr{DS}$	period	10 s	30 s	1 min	2 min
Method	AML	(0.989, 0.913, 0.656)	(0.955, 0.816, 0.808)	(0.959, 1.414, 1.393)	(0.968, 2.577, 2.515)
$\mathscr{DS}$	period	5 min	15 min	30 min	1 hour
Method	AML	(0.928, 3.445, 3.401)	(0.864, 4.112, 4.057)	(0.816, 4.695, 4.414)	(0.777, 5.199, 4.679)

Table 2: Estimated parameters of $\mathscr{SDS}(\gamma, \lambda)$ and $\mathscr{DS}(\gamma, \lambda_1, \lambda_2)$ from real data

The quality of the fit of the empirical data with symmetric discrete stable, discrete stable and stable distribution is displayed at Figure 1. The mean square error of the fit with $\mathscr{SDS}$ distribution is 1.4e-5, for $\mathscr{DS}$ is 1.6e-5 and with $\mathscr{S}$ is 3.0e-5.

Acknowledgements

Authors were partially supported by Czech Science Foundation (grant P402/12/G097 and grant P203/12/0665).

References

[1] Christoph, G., and Schreiber, K.: Discrete stable random variables. *Statistics & Probability Letters* **37** (1998), 243–247.

[2] Devroye, L.: A triptych of discrete distributions connected to stable law. *Statistics & Probability Letters* **18** (1993), 349–351.

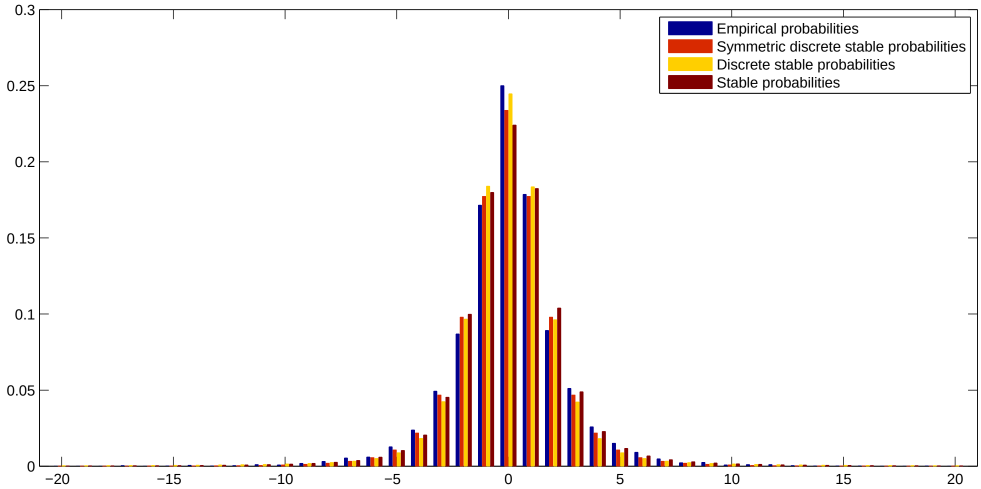

Figure 1: Empirical probabilities together with the theoretical probabilities of the fitted distributions

[3] Feuerverger, A., and McDunnough, P.: On the efficiency of empirical characteristic function procedure. *Journal of the Royal Statistical Society B* **43** (1981), 20–27.

[4] Klebanov, L., and Slámová, L.: Integer valued stable random variables. *Submitted for publication* (2012).

[5] Steutel, F. W., and van Harn, K. : Discrete analogues of self-decomposability and stability. *Annals of Probabability* **7** (1979), 893–899.

[6] Yu, J.: Empirical characteristic function estimation and its applications. *Econometric reviews* **23** (2004), 93–123.