Digital Signal Processing ••• (••••) •••–•••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Generation of cubic power-law for high frequency intra-day returns: Maximum Tsallis entropy framework

Dilip Senapati, Karmeshu *

School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, 110067, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Intra-day returns
q-Gaussian distribution
Power-law behavior
Tsallis entropy
Generalized JS divergence
Financial signals

ABSTRACT

A theoretical framework based on the maximum Tsallis entropy is proposed to explain the tail behavior of the intra-day stock returns, providing a rationale for the cubic law behavior for high frequency data. The specification of first two time-dependent moment constraints yields a q-Gaussian distribution for the intra-day stock returns. The value of the parameter q is estimated by minimizing appropriately modified Jensen–Shannon (JS) divergence in Tsallis entropy framework between q-Gaussian distribution and empirical NASDAQ 100 data. The estimated value of q yields the well-known empirically observed cubic law tail behavior of the intra-day stock returns which has been observed for high frequency data sets. To validate the cubic law stylized fact, five more data sets from high frequency NASDAQ 100, S&P 500 and NYSE index have been examined and it is found that the cubic law operates.

1. Introduction

Recent empirical investigations on high frequency stock price data reveal the emergence of cubic law for fat-tailed intra-day returns. Application of signal processing to high frequency financial signals [1] (such as stock price, returns, volatility etc.) has largely remained unexplored [2]. However, Gradojevic and Gencay [3] have pointed out that the well-known signal processing techniques are unable to explain the distributional properties of financial markets, which exhibit long-range interactions. It has been an interesting and challenging problem for researches to find out the true distribution of stock price change. Gerig, Vicente and Fuentes [4] provide a mixture distribution model for intra-day stock returns to capture the tail fluctuation in terms of Student's t-distribution with a certain number of degrees of freedom (df) which is a positive integer. However, t-distribution will not be applicable if the estimation procedure to fit it to empirical data does not lead to a discrete value of the parameter corresponding to degrees of freedom. Researches have applied the non-extensive statistical mechanics to study the tail behavior of price change [5]. Following the Tsallis framework, Borland and Bouchaud [6,7] have modeled the non-Gaussian intra-day stock price change through stochastic differential equation (SDE) involving non-extensive parameter. It has been reported by Stanley and Buchanan [8–10] that across all the stocks in different countries, the intra-day returns (IDR) for high frequency data exhibit cubic law [11–13]. This law has also been validated by Pan and Sinha [13] for Indian financial market. Kumar and Deo [14] used the multifractal detrended fluctuation analysis (MF-DFA) to characterize multifractal structure of Indian financial market. An important problem is to understand the underlying mechanism which leads to such a behavior. This aspect has also been modeled by Cozzolino and Zahner [15]. They make the following observation: "The goal is the probability distribution encoding the decision-maker's uncertainty of the price of the stock at time t, given his prior knowledge concerning its possible future behavior" [15]. Behmardi, Raich and Hero [16] have outlined a new procedure to estimate parameters of Shannon maximum entropy distribution. By employing the maximum entropy framework due to Jaynes [17], the maximization of Shannon entropy subject to first two moments of intra-day returns gives the Gaussian probability distribution. Accordingly maximum Shannon entropy fails to capture the fat tail behavior of intra-day returns. The underlying reason is that Shannon entropy is essentially applicable to systems which are extensive. However, many characteristics of the financial market display non-extensivity. Accordingly, Tsallis entropy may be well-suited to deal with financial applications.

Tsallis [5] has proposed a non-extensive measure of entropy with a parameter q which in the limit $q \to 1$ results in Shannon entropy. This entropy measure has successfully described the phenomenon of multifractal structure and long-range dependence which exhibits a power-law behavior [18]. A notable feature of Tsallis entropy is pseudo-additivity of two independent systems [5,19–21]. In an interesting paper, Borland and Bouchaud [6,7] have

* Corresponding author.
E-mail addresses: senapatidillip@gmail.com (D. Senapati), karmeshu@gmail.com, karmeshu@mail.jnu.ac.in (Karmeshu).

http://dx.doi.org/10.1016/j.dsp.2015.09.018
1051-2004/© 2015 Published by Elsevier Inc.

proposed a non-Gaussian model of stock returns by incorporating the "statistical feedback" process [7], which gives rise to a nonlinear Fokker–Planck equation involving non-extensive parameter q. Gradojevic and Gençay [3] have employed Tsallis entropy to study the behavior of markets, particularly during financial crisis. They examine the behavior of the parameter q to characterize the markets ranging from mature to emerging ones. The parameter q is estimated for normalized stock returns for daily, weekly and monthly data of S&P 500 index. Similar analysis has been carried out by Namakia et al. [22] who have analyzed normalized returns for monthly, weekly and daily for various index data sets TSE, SSE, KS 11, DJIA 30, S&P 500 and NASDAQ 100.

In this paper, we broaden Shannon's maximum entropy framework adopted by Cozzolino and Zahner [15] to the one based on maximum Tsallis entropy. The maximization of Tsallis entropy yields probability distribution of intra-day returns in terms of non-extensive parameter q which can be estimated from the available high frequency stock price data. It may be pertinent to point out that estimation procedures based on Shannon entropy framework may not be valid when applied to problems formulated in Tsallis entropy framework. This limitation is evident in the estimation procedures adopted by Gradojevic and Gencay [3] where they have obtained the estimate of q by minimizing the sum of the squared errors of the logarithms of the q-Gaussian probability density and empirical data density. In the second method, they have estimated the parameter q by maximum likelihood method. Further, Gradojevic and Gencay [3] note that the estimates of q obtained by different procedures are very different. The reason could be on account of use of estimation procedures which are not consistent with Tsallis entropy framework. In light of preceding discussion, the estimation procedure which we have adopted requires minimization of generalized Jensen–Shannon (JS) divergence in Tsallis entropy framework between maximum entropy probability distribution (MEPD) and the reference empirical distribution.

This paper consists of six sections. Starting with introduction in Section 1, we discuss in Section 2 the maximum Tsallis entropy framework to compute probability distribution of intra-day returns involving parameter q. The next Section 3 outlines the procedure based on symmetric JS measure. Using high frequency data Powershares QQQ listed in NASDAQ 100 index, the parameter q is estimated and cubic tail behavior of intra-day returns is demonstrated in Section 4. Section 5 validates the cubic stylized fact for high frequency data based on data of five other stocks viz. General Motors, Coca-Cola, SPDR S&P 500 ETF, NQ Mobile Inc. and Eversource Energy listed in NYSE, NASDAQ 100 and S&P 500 index. The choice of these high frequency datasets, as collected in 2009, is dictated by the fact that these were the only datasets which were freely available on the website www.furturstickdata.com [23] at that point of time. The last section gives the conclusion.

2. Maximum entropy probability distribution

2.1. q-Gaussian distribution of stock return

It is common to model stock intra-day returns $r(t)$ in terms of the logarithmic relative price changes of stock. We have

$$r(t) \equiv \ln\left(\frac{S_t}{S_{t-\Delta t}}\right). \tag{1}$$

The well-known Tsallis entropy [5,20] in terms of non-extensive parameter q is given by

$$S_q(r,t) = \frac{1-\int_{-\infty}^{+\infty}[f_q(r,t)]^q dr}{q-1}, \tag{2}$$

Our aim is to investigate probability distribution of intra-day returns when Tsallis entropy is maximized subject to the following constraints.

The normalization constraint is given by

$$\int_{-\infty}^{+\infty} f_q(r,t)dr = 1. \tag{3}$$

The first two time-dependent moment constraints are specified as

$$\int_{-\infty}^{+\infty} \tilde{f}_{es}(r,t) r dr = \mu t, \tag{4}$$

and

$$\int_{-\infty}^{+\infty} \tilde{f}_{es}(r,t)(r-\mu t)^2 dx = \sigma^2 t, \tag{5}$$

where

$$\tilde{f}_{es}(r,t) = [f_q(r,t)]^q \Big/ \int [f_q(r,t)]^q dr, \tag{6}$$

represents the escort probability distribution [3,22,24]. The justification of using the escort distribution instead of the commonly used probability distribution is given in Abe and Bagci [24]. In the context of a random variable, they have discussed two kinds of expectation viz. ordinary expectation and q-expectation (involving escort distribution). In their paper, they have argued that in non-extensive statistical mechanics that exhibit the power-law distribution, only the q-expectation is consistent with the basic framework of micro-canonical distribution.

We construct Lagrangian as given by,

$$\begin{aligned} L(f_q, r) = {} & \frac{1-\int_{-\infty}^{+\infty}[f_q(r,t)]^q dr}{q-1} + \lambda_1 \left(1 - \int_{-\infty}^{+\infty} f_q(r,t)dr\right) \\ & + \lambda_2 \left(\mu t - \int_{-\infty}^{+\infty} r\left([f_q(r,t)]^q \Big/ \int [f_q(r_1,t)]^q dr_1\right) dr\right) \\ & + \lambda_3 \left(\sigma^2 t - \int_{-\infty}^{+\infty} \left([f_q(r,t)]^q \Big/ \int [f_q(r_1,t)]^q dr_1\right) \right. \\ & \left. \times (r-\mu t)^2 dr\right), \end{aligned} \tag{7}$$

where λ_1, λ_2 and λ_3 are Lagrangian parameters. These can be determined from constraints. Using Euler–Lagrange equation

$$\frac{\partial L}{\partial f_q(r,t)} - \frac{d}{dr}\left(\frac{\partial L}{\partial f_q'(r,t)}\right) = 0\,, \tag{8}$$

Eq. (8) in conjunction with Eqs. (2)–(7) yields the desired maximum entropy probability distribution of intra-day returns. We obtain

$$f_q(r,t) = \frac{1}{Z}\left(1 - \frac{1-q}{(3-q)\sigma^2 t}(r-\mu t)^2\right)^{\frac{1}{1-q}}, \ -\infty < r < \infty, \tag{9}$$

where the normalization constant Z turns out to be

$$Z = \sigma t \sqrt{\frac{(3-q)\pi}{q-1}} \frac{\Gamma(\frac{3-q}{2q-2})}{\Gamma(\frac{1}{q-1})}, \qquad 1 < q < 3. \tag{10}$$