International Journal of Pure and Applied Mathematics
Volume 50 No. 2 2009, 245-250

ESTIMATION OF EXTREME VALUES OF RETURNS USING THE ZIPF-MANDELBROT LAW

Vladimir Gisin[1 §], Andrey Markov[2], Igor Vinukov[3]
[1,2,3]Finance Academy under the Government of the Russian Federation
49, Leningradsky Prospect, Moscow, 125468, RUSSIA
[1]e-mail: vgisin@online.ru

Abstract: Many empirical distributions in finance have so-called fat tails which exhibit power-low behavior. This phenomenon can be explained by means of the Zipf-Mandelbrot law. In this paper we estimate the extreme values of stock return using this low.

AMS Subject Classification: 62P20, 91B28, 91B84
Key Words: Zipf law, fat tails, financial market

1. Introduction

Many social, economic, human and other systems may be successfully studied using the model developed and investigated by Zipf. In his book [9] Zipf has formulated the principle of the least effort. According to this principle the law established in linguistics (later called the Zipf's law) can be extended to other fields. Phenomena similar to Zipf's law have been found out in economics (Pareto), biology (Jule) and many other areas.

The Zipf's law has numerous mathematical, physical and philosophical justifications. It was revised in its generic form by Mandelbrot (see [6]). The Zipf-Mandelbrot law (ZM law for short) appears when a certain random variable is the sum of a big number of disparate random variables. Moreover, according to [8], the ZM law is a manifestation of extreme disparity (the principle of minimal symmetry), when there exists a singled out archetype of the normal state of a system, described by these random variables.

Received: August 14, 2008 © 2009 Academic Publications
§Correspondence author

Recently the ZM law has been discovered in many social and economic areas (see [5], [7], [2], [1], [3]). In the context of financial markets the ZM law is attributed to power ("fat") tales of distributions describing financial time series.

2. The Zipf-Mandelbrot Law and Stock Returns

Basically the ZM law may be presented in the following way. Suppose N is a certain class of objects, and any object may be ranged in one of K disjoint taxons. Each taxon has its rank i, $i = 1, 2, \ldots, K$. It may be supposed that the first taxon contains the most typical objects and when the rank increases the objects become less typical. Let M be a family of objects from N. Let n_i be the number of objects from M in taxon i, $i = 1, 2, \ldots, K$. We say that the family M complies with the Zipf-Mandelbrot law if the share of objects from M which are in the taxon i is equal to

$$\frac{A}{(B+i)^{\gamma}}, \tag{1}$$

where A, B and γ are constants.

Consider the asset price evolution $P = (P_t)$. We fix a time interval Δt and define return R_t over Δt by

$$R_t = \ln \frac{P_t}{P_{t-\Delta t}}.$$

Normally the distribution of R_t has power "tales" (see [3]):

$$P(|R_t| > x) \propto x^{-\zeta}. \tag{2}$$

This formula is in a sense a continuous version of the ZM law. Indeed, using (2) we have

$$P(|R_t| \in [x-h, x]) \propto h \cdot x^{-(\zeta+1)}. \tag{3}$$

In the last formula x plays the same role as the rank i in (1).

Now consider the evolution of R_t over the time interval T. Let r be the mean value of R_t and let $Q_t = |R_t - r|$. Then (2) may be adapted for Q_t:

$$P(Q_t > x) \propto (x+b)^{-\zeta}. \tag{4}$$

We suppose that $\operatorname{esssup} Q_t < \infty$ on time horizon T. We divide the range of Q_t into K intervals of length h and enumerate these intervals so that

$$I_i = [(i-1)h, ih], \quad i = 1, 2, \ldots, K.$$

In what follows we identify taxons with intervals I_i. Actually using (3) and (4) we get

$$P(Q_t \in I_i) = A(i+B)^{-\gamma}. \tag{5}$$

The aim of this paper is to estimate the upper bound of $|R_t|$.

3. Estimation of the Number of Taxons

The hypothesis of fat tails says that (5) is valid for big values of i. Normally (5) holds approximately for all values of i. To simplify computations we assume later on that (5) is fulfilled for all $i = 1, 2, \ldots, K$. This assumption allows us to use the techniques developed in [4].

From now on we suppose that t takes discrete values $\Delta t, 2\Delta t, \ldots, n\Delta t$. Let

$$p_i = A(i+B)^{-\gamma}, \quad i = 1, 2, \ldots, K, \tag{6}$$

where K is the least nonnegative integer such that $Q_t \in [0, Kh]$ (a.s.).

Let $P_n(m,i)$ be the probability of the event "exactly m values of Q_t are in I_i". We suppose that

$$P_n(m,i) = \binom{n}{m} p_i^m (1-p_i)^{n-m}. \tag{7}$$

Since n is big and p_i are relatively small we replace (7) by

$$P_n(m,i) = \frac{n{p_i}^m}{m!} e^{-np_i}.$$

Let k_m be the number of intervals I_i which contain exactly m values of Q_t. It is clear that

$$E(k_m) = \sum_{i=1}^{K} P_n(m,i).$$

Putting

$$f_m(x) = \left(\frac{An}{(x+B)^\gamma}\right)^m \cdot e^{-\frac{An}{(x+B)^\gamma}}$$

we get

$$E(k_m) = \sum_{i=1}^{K} \frac{f_m(i)}{m!}. \tag{8}$$

We are going to replace the sum

$$\sum_{i=1}^{K} f_m(i)$$

from (8) by the integral

$$\int_{\alpha}^{\alpha+K} f_m(x)\,dx$$

with $0<\alpha<1$. First we expand $f_m(x)$ as a power series in $x-i$ in the interval

$$(\alpha+i-1;\alpha+i).$$

We get

$$f_m(x)=\sum_{j=0}^{\infty}\frac{1}{j!}f_m^{(j)}(i)(x-i)^j. \tag{9}$$

Since all derivatives $f_m^{(j)}(i)$ are jointly bounded the sum on the right hand side in (9) is absolutely convergent. So we have

$$\int_{\alpha+i-1}^{\alpha+i} f_m(x)\,dx=\sum_{j=0}^{\infty}\int_{\alpha+i-1}^{\alpha+i}\frac{1}{j!}f_m^{(j)}(i)(x-i)^j\,dx$$
$$=\sum_{j=0}^{\infty}\frac{f_m^{(j)}(i)}{(j+1)!}\left(\alpha^{j+1}-(\alpha-1)^{j+1}\right).$$

Now summing over i we arrive at

$$\int_{\alpha}^{\alpha+K} f_m(x)\,dx=\sum_{i=1}^{K}\sum_{j=0}^{\infty}\int_{\alpha+i-1}^{\alpha+i}\frac{1}{j!}f_m^{(j)}(i)(x-i)^j\,dx$$
$$=\sum_{j=0}^{\infty}\sum_{i=1}^{K}\frac{f_m^{(j)}(i)}{(j+1)!}\left(\alpha^{j+1}-(\alpha-1)^{j+1}\right).$$

So up to higher order terms we get

$$\int_{\alpha}^{\alpha+K} f_m(x)\,dx\approx\sum_{i=1}^{K}f_m(i)+(\alpha-1/2)(f_m(\alpha+K)-f_m(\alpha).$$

If $\alpha\to -B$ then $f_m(\alpha)\to 0$ and hence

$$\sum_{i=1}^{K}f_m(i)\approx\int_{-B}^{-B+K} f_m(x)\,dx+(B+1/2)f_m(-B+K). \tag{10}$$

Using (8), (10) and substituting $y=An(B+x)^{-\gamma}$ we come to the following:

$$\frac{E(k_m)}{K}\approx\frac{(An)^{1/\gamma}}{K\gamma m!}\int_{AnK^{-\gamma}}^{\infty} y^{m-\frac{1}{\gamma}-1}e^{-y}\,dy. \tag{11}$$

Put

$$\rho=\frac{An}{K^{\gamma}},\quad \mu=m-\frac{1}{\gamma},\quad \nu_m=\frac{\rho^{1/\gamma}}{\gamma m!}\int_{\rho}^{\infty} y^{\mu-1}e^{-y}\,dy. \tag{12}$$

It can be checked that

$$\frac{\nu_{m+2} - \dfrac{\mu+1}{m+2}\nu_{m+1}}{\nu_{m+1} - \dfrac{\mu}{m+1}\nu_{m+1}} = \frac{\rho}{m+2}.$$

In fact with

$$\Gamma(\mu, \rho) = \int_\rho^\infty y^{\mu-1} e^{-y}\, dy$$

we have

$$\Gamma(\mu+1, \rho) = \rho^\mu e^{-\rho} + \mu\Gamma(\mu, \rho)$$

whence it follows that

$$\nu_{m+1} = \frac{\rho^{1/\gamma}}{\gamma(m+1)!}\Gamma(\mu+1, \rho) = \frac{\mu}{m+1}\nu_m + \frac{1}{\gamma} \cdot \frac{\rho^m e^{-\rho}}{(m+1)!}.$$

Then according to (11)

$$\rho \approx (m+1)\frac{(m+2)E(k_{m+2}) - (m+1-\dfrac{1}{\gamma})E(k_{m+1})}{(m+1)E(k_{m+1}) - (m-\dfrac{1}{\gamma})E(k_m)}. \tag{13}$$

Note that the error in (13) increases with m increasing. So (13) should be applied for small m such as $m = 1, 2, 3$. For example

$$\rho \approx 2 \cdot \frac{3E(k_3) - (2-\dfrac{1}{\gamma})E(k_2)}{2E(k_2) - (1-\dfrac{1}{\gamma})E(k_1)}. \tag{14}$$

Given A and γ from the ZM law and ρ from (14) we can calculate K using (12) as follows:

$$K \approx \left(\frac{An}{\rho}\right)^{1/\gamma}.$$

Now this value of K may be applied to estimate the range of R_t:

$$r - Kh < R_t < r + Kh \quad (a.s.).$$

Acknowledgements

This work was supported by Russian Foundation for Basic Research under grant 08-06-00283.

References

[1] R.I. Axtell, Zipf distribution of US firm sizes, *Science*, **293** (2001), 1818-1821.

[2] S. Brakman, H. Garretsen, C. Van Marrewijk, M. Van den Berg, The return of Zipf: Towards a further uderstanding of the rank-size distribution, *Journal of Regional Science*, **39** (1999), 739-767.

[3] X. Gabaix, P. Gopikrishnan, V. Plerou, H.E. Stanley, Quantifying and understanding the economics of large financial movements, *Journal of Economic Dynamics and Control*, **32** (2008), 303-319.

[4] V.M. Kalinin, Functionals related to Poisson distribution and the statistical structure of a text, *Trudi Matematicheskogo Instituta Steklova*, **79** (1965), 182-197, In Russian.

[5] S.A. Krashakov, A.B. Teslyuk, L.N. Shyur, On the generality of the order distribution of web sites popularity, *Bull. of RFFI*, **35**, No. 1 (1981), 46-66, In Russian.

[6] B.B. Mandelbrot, Contribution a la theorie mathematique des jeux de communication, *Publ. Inst. Stat. Univ. Paris*, **2** (1953), 1-124.

[7] A.V. Podlazov, Competitors' distribution, scale invariance of state and models of linear growth, *Izvestija Vuzov. Applied nonlinear dynamics*, **10**, No-s: 1-2 (2002), 20-43, In Russian.

[8] U.A. Shreider, A.A. Sharov, *Systems and Models*, Radio i Sviaz, Moscow (1982), In Russian.

[9] G.K. Zipf, *Human Behavior and the Principle of the Least Effort*, Cambridge University Press, Cambridge (1949).