Measures of Dependence for Multivariate Lévy Distributions[1]

J. Boland*, T. R. Hurd*, M. Pivato† and L. Seco†

*Dept. of Mathematics and Statistics, McMaster University
Hamilton, Canada, L8S 4K1
†Dept. of Mathematics, University of Toronto
Toronto, Canada, M5S 3G3

Abstract. Recent statistical analysis of a number of financial databases is summarized. Increasing agreement is found that logarithmic equity returns show a certain type of asymptotic behaviour of the largest events, namely that the probability density functions have power law tails with an exponent $\alpha \approx 3.0$. This behaviour does not vary much over different stock exchanges or over time, despite large variations in trading environments. The present paper proposes a class of multivariate distributions which generalizes the observed qualities of univariate time series. A new consequence of the proposed class is the "spectral measure" which completely characterizes the multivariate dependences of the extreme tails of the distribution. This measure on the unit sphere in M–dimensions, in principle completely general, can be determined empirically by looking at extreme events. If it can be observed and determined, it will prove to be of importance for scenario generation in portfolio risk management.

I INTRODUCTION

Much research over the past forty years has examined statistical properties of stock returns with an aim to find models which improve over the Brownian motion models which began with Bachelier's work in the early 1900s. Bachelier's theory of speculation, adapted to logarithmic returns, assumes that successive increments of the logarithmic returns $X(t, \delta t) = \log[S(t + \delta t)/S(t)]$ of a time series of prices $S(t)$ are (a) random, (b) statistically independent, (c) identically distributed, and (d) Gaussian with zero mean. Bachelier himself pointed to assumption (d) as the weakest conceptual link, and subsequent researchers have noted that by replacing Gaussians with more general distributions the remaining three assumptions stand up relatively well. Amongst the models which have gained some popularity in the literature are stable–Lévy processes [11], [13], truncated stable–Lévy processes

[1] Research supported by the Natural Sciences and Engineering Research Council of Canada and MITACS, Canada

[3], fractional Brownian motion [12], hyperbolic distributions [7], [8] and Bessel distributions [10], [9].

Of the shortcomings of geometric Brownian motion, the most egregious is the fat tail problem: everyone agrees empirical time series for financial returns are always leptokurtic, i.e. they have tails which are thicker than those of any Gaussian. Amongst the fat–tailers who regard this fact as very important we can identify two camps: the "exponential tailers" and the "power law tailers" who see a probability density function which is asymptotically proportional to a negative power of the event size.

The theoretical justification for power law tails rests on the experience of physicists, geologists and others who study the mathematics of complex systems which consist of an enormous number of units connected by finite range interactions. For example, highly turbulent fluids are described qualitatively by a cascade of energy from large distance scales through to a short distance dissipative scale. The spectrum of excitations exhibits "Kolmogorov scaling" which is a clear power law. The concept of "self–organized criticality" has been proposed as an explanation for the ubiquity of power laws in such complex non-equilibrium systems [1]. A key consequence of this concept is the "universality" of the exponents of the power laws in question, i.e. their invariance under continuous changes in the microscopic interaction law.

An equity market can certainly be viewed as a large system of traders in interaction with each other and the external economy. Thus it is natural to imagine that prices, being a measure of the state of the system, might exhibit self–organized criticality and consequently the type of universal power law seen in some physical systems. Of course, finance is not physics, and any tidy picture from physics will be severely muddied by problems of changing market conditions (stemming for example from evolving technology), irrationality and other psychological factors, the wide variation in trading styles, etc. Nevertheless, if self–organized criticality is applicable in finance it leads naturally to an optimistic prediction, consequences of which we will explore in this paper:

Scaling Prediction: *The log returns within equity markets should exhibit power law tails with a universal exponent α. This exponent will the same in different sectors within a given stock market, and will be the same in different stock markets around the world.*

In section 2 we review some recent work on the statistics of financial time series. The evidence does appear to offer some support to the validity of the scaling prediction: the log returns of a great number of different equities and equity indices, over time intervals from one minute to several days, show power law tails all with exponents $\alpha \approx 3.0$. Furthermore, there is some evidence for universal behaviour beyond that asserted in the scaling prediction: the probability density functions for returns over time intervals in the range from five minutes to several days change only by an overall scale, having the same universal shape.

The present review takes this empirical evidence seriously, and explores the consequences of the scaling prediction. The main difficulty in considering non-Gaussian distributions is that the dependence structure is not easily characterized in a unique way. In the Gaussian case, the correlation matrix determines the multivariate distribution once the marginals are known. Moreover, despite the fact that determining the full correlation structure exactly from market data is often difficult, the Gaussian distribution is simple enough that determining a few of its principal components is often sufficient for risk management purposes, and these can easily be obtained from market data.

We will characterize a general class of multivariate distributions consistent with the scaling prediction for its marginals, and derive their associated asymptotic dependence structure ("tail dependence"). This tail dependence will turn out to be entirely determined by a measure on a high-dimensional sphere, and we will obtain estimates for the impact of this measure on the conditional correlations in the tails of the marginals. In a precise sense, this "spectral measure" is the natural analog for Lévy distributions of the Gaussian correlation matrix. The results presented here are proven in the working paper [2]. The issue of the determination of the dependence structure from historical data will be addressed in future work.

II REVIEW OF EMPIRICAL STUDIES

Here we briefly summarize some empirical studies done over the past forty years which have been important for researchers studying the distributions of price log-returns $X(t, \delta t)$. For an extended and readable discussion, see [5].

Benoit Mandelbrot [11] was an early pioneer in studying the returns of cotton price fluctuations, and proposed nearly forty years ago that the distribution of log returns $X(t, \delta t)$ can be modelled by a stable Lévy distribution. More precisely, he proposed that the probability that $X(t, \delta t)$ is bigger than x is asymptotically $Cx^{-\alpha}$ as $x \to \infty$, where $\alpha \approx 1.7$. His proposal was based on two elements: first, that the distribution of price returns is decidedly non-Gaussian, and second, that the functional form of these distributions doesn't depend on the increment size δt. Models based on this stable ($\alpha < 2$) Lévy hypothesis have the difficulty that their distributions do not have a finite second moment, but stability guarantees that there is a well-developed theory generalizing many features of Gaussian distributions [13].

But we now know this picture is too simplistic, since as δt grows large, the distributions converge to a Gaussian [3]. Two papers, [6] examine the stock price log-returns of individual companies and the log-returns of the S&P 500 index. They present evidence that this crossover to Gaussian behavior occurs for $\delta t \approx 4$ days in the case of the S&P 500 index, and for $\delta t \approx 16$ days in the case of individual stocks. They emphasize however, that for δt less than these crossover points, the distributions are essentially independent of δt.

This brings us to the second departure from Mandlebrot's hypothesis. While the return distributions are power laws as he suggested, the work just cited shows that for a great range of x values (from 3 to 50 standard deviations in the case of the

S&P 500, and from 2 to 80 standard deviations in the case of individual companies), the exponent α is approximately 3, not 1.7. This is true for any of the 1000 stocks chosen, as well as the Hang-Seng and NIKKEI indices, and does not depend on the size of δt (below the crossover mentioned above). Dacorogna et al [4] have found $\alpha \gtrsim 3$ power laws in foreign exchange markets as well. These findings suggest that return distributions for financial fluctuations should not be modelled with Lévy stable laws, but with more general Lévy distributions characterized by $\alpha > 2$. Our main goal in this paper is to extend this idea to the multivariate setting, where we assume that there is a fixed $\alpha \approx 3$ so that in every direction, the cumulative density function has power law decay with exponent α.

III MODELLING ASSUMPTIONS

In the remainder of this paper we focus on a collection of M different equity prices. We let $S_t = \{S_t^k\}_{k=1,\ldots,M}$ be the vector stock price process where S_t^k is the price of the kth equity at time $t > 0$. We focus on the sequence of returns sampled at N discrete times $t_j = j\delta t,\ j = 0, 1, \ldots, N$, where $\delta t > 0$ is a fixed time interval. We define the vector of (log) returns from t_{j-1} to t_j to be $X_j = \{X_j^k\}_{k=1,\ldots,M}$ with $X_j^k = \log(S_{t_j}^k / S_{t_{j-1}}^k)$. In general, and in the remainder of the paper, we are concerned with determining a set of mathematically natural conditions on the process X_t which are compatible with the observed regularities in the data.

Remark: We defer to a subsequent paper the question of how the distribution of the returns vary with the value of δt. In the present paper, δt should be taken to be a value intermediate between 1 minute and 4 days.

We list a set of modelling assumptions which will characterize a certain class of multivariate distributions. We state these assumptions as applicable directly to the time series of returns: more realistically, one should think of these assumptions as applicable to the set of underlying risk factors in the model. Observed time series would then be modelled as a process driven by these factors. Nonetheless, when applied directly to observed time series our modelling assumptions do a surprisingly reasonable job.

H1 *The random variables X_j are independent and identically distributed. We call the underlying M–dimensional distribution X, thus $X_j \stackrel{d}{=} X$ for all j. X has finite covariance and is infinitely divisible.*

We assume finite covariance since this does seem to be true in observed time series. Infinite divisibility is a consequence of thinking of X as the increment $X_{\delta t}$ of an underlying continuous time process X_t. From this condition the Lévy–Khintchine theorem implies the following representation for the log–characteristic function $\Psi_X(u) = -\log(E[e^{i(u,X)}])$ of X:

$$\Psi_X(u) = -i(u,m) + (u,Qu)/2 + \int_{\mathbf{R}^M} [e^{i(u,a)} - 1 - i(u,a)\chi_{|a|<1}(a)]d\mu(a) \qquad (1)$$

Here $u, m \in \mathbf{R}^M$, Q is a positive semi-definite quadratic form on $\mathbf{R}^M$, the Lévy–Khintchine (LK) measure μ is a positive measure on $\mathbf{R}^M$ which satisfies a certain integrability condition, and for any set $A \subset \mathbf{R}^M$, χ_A denotes the indicator function for that set.

The LK–measure has an interpretation which we find very useful in guiding our work. A Poisson process with jump size $a \in \mathbf{R}^M$ and rate $\lambda > 0$ has log–characteristic function $\lambda[e^{i(u,a)} - 1]$. Roughly, the last term of (1) is interpreted as a (continuous) compounding of Poisson jump processes in which jumps of size in the set $[a, a + da]$ occur at the rate $d\mu(a)$.

The second assumption restricts the form of the log–characteristic function to be that of a "purely discontinuous" process:

H2 *We assume* $Q = 0$.

The final assumption embodies and abstracts the observation that observed equities as well as numerous *stock indices* seem to have power law tails with the same exponent.

H3 *There is a constant* $\alpha > 0$ *such that every generalized (univariate) marginal* $Y = \xi^T X$, $\xi \in \mathbf{R}^M$ *of the* $\mathbf{R}^M$*–valued random variable* X *has a probability density function* ρ_Y *with power law asymptotics:*

$$\rho_Y(y) \sim g_\pm(\xi)|y|^{-1-\alpha} \quad \text{as } y \to \pm\infty$$

for some constants $g_\pm(\xi) \geq 0$. *We say that* X *has universal exponent* α.

IV A CLASS OF MULTIVARIATE DISTRIBUTIONS

In this section we discuss a general family which satisfies **H1-H3**.

Definitions: Given $\alpha > 2$, we say an M–dimensional random variable is of class α if its log–characteristic function has the form (1) with $Q = 0$ and such that the Lévy-Khintchine measure $d\mu_X(a)$ is asymptotically in separated form:

$$d\mu_X(a) \sim r^{-1-\alpha}\, dr\, d\Gamma(\hat{a}), \quad r \to \infty \tag{2}$$

where $r = |a|$ and $\hat{a} = a/r$. The "spectral measure" Γ is a general (positive, finite) measure on the unit sphere $S \in \mathbf{R}^M$. Let $\mathcal{H}_{M,\alpha}$ denote the family of all M–dimensional class α random variables, and let $\mathcal{H}_\alpha = \cup_{M>0}\mathcal{H}_{M,\alpha}$.

The main structural result concerning the family $\mathcal{H}_\alpha$ is that it is closed under linear transformations.

Theorem 1 *Let* $X \in \mathcal{H}_{M,\alpha}$, *and let* $B : \mathbf{R}^M \to \mathbf{R}^k$ *be a non–zero linear transformation. Then* $B(X) \in \mathcal{H}_{k,\alpha}$.

Theorem 1 implies that **H3** is true for $X \in \mathcal{H}_{M,\alpha}$. To see this, we need the following univariate result:

Proposition 2 *Suppose the univariate X has LK measure μ which satisfies*

$$\mu(a) \sim g_{\pm}|a|^{-1-\alpha} \quad as\ a \to \pm\infty \tag{3}$$

Then the probability density function of X has the same asymptotics as μ.

A direct asymptotic calculation now shows that class α distributions satisfy **H3**:

Corollary 3 *Let $X \in \mathcal{H}_{M,\alpha}$. Then for any $\xi \in \mathbf{R}^M$, the generalized marginal of X in the direction of ξ, $Y = (\xi, X)$, has power law asymptotics with exponent α and constants*

$$g_{\pm}(\xi) = \int ((\xi, \hat{a}))_{\pm}^{\alpha} d\Gamma(\hat{a}) \tag{4}$$

where $(x)_{\pm} = \max(\pm x, 0)$.

We can use a multivariate version of Proposition 2 to strengthen the connection between the LK measure and the probability density function of X.

Theorem 4 *Let $X \in \mathcal{H}_{M,\alpha}$. Then the probability density function of X has the same asymptotics as the LK measure of X.*

Thus the behaviour of large values of a random vector $X \in \mathcal{H}_{M,\alpha}$ is determined by Γ. We introduce a set of "large events" parameterized by a radius $R > 0$: $B_R = \{X : |X| > R\}$, where $|\cdot|$ denotes Euclidean norm. Now let$X_R = X|R$ be X conditioned on the event B_R. The probability density function of X_R is $\rho(x)\chi_R(x)/\int \rho(x)\chi_R(x)\,dx$. To emphasize that the following result is very different from what will happen with Gaussians, observe that if X is a multivariate Gaussian, then for R large X_R becomes concentrated on the maximal eigenspace of the covariance matrix of X. Thus the dependences possible for large Gaussian events are very restricted.

Theorem 5 *Suppose $X \in \mathcal{H}_{\alpha,M}$ with spectral measure Γ and vanishing mean. Then the correlation matrix of X_R is asymptotic to the correlation of Γ:*

$$C_{R,ij} = \frac{E[(X_{Ri} - \bar{X}_{Ri})(X_{Rj} - \bar{X}_{Rj})]}{\sqrt{var(X_{Ri})var(X_{Rj})}} \sim \frac{E_{\Gamma}[a_i - \bar{a}_i)(a_j - \bar{a}_j)]}{\sqrt{var_{\Gamma}(a_i)var_{\Gamma}(a_j)}}$$

Example: A well-studied fat–tailed distribution in economics is the univariate t–distribution with *3* degrees of freedom which has probability density function $C(1+x^2)^{-2}$ and log–characteristic function $\Psi(u) = |u| - \log(1+|u|)$. We can prove that this distribution satisfies **H1,H2**. Now we use this distribution to generate a family of multivariate t–distributions satisfying **H1–H3** with tail exponent $\alpha = 3$. Let $X = (X_1, X_2, \ldots, X_M)$ be a random vector whose components are i.i.d. random variables with this distribution. Then define the family to consist of all $Y = B(X)$ for any value M and any linear transformation B. Such distributions have the asymptotic behaviour (2). The spectral measure of Y consists of a finite number of "atoms" (delta functions) at the points $B(e_i)$ with weights $|B(e_i)|^{\alpha}$, where e_i is the i–th standard basis vector in $\mathbf{R}^M$.

V CONCLUSIONS

We have seen that the scaling prediction leads for each M to a broad class of M–dimensional distributions which are characterized by an asymptotic scaling in both the LK measure and in the probability density function. For such random vectors, the joint distribution of large events is determined by an exponent $\alpha > 2$ and the spectral measure Γ, an arbitrary finite measure on the unit sphere in $\mathbf{R}^M$. Normalized moments of the conditional random vector X_R are given by the moments of Γ if R is large enough.

Such behaviour is completely different in character from what is possible in the Gaussian universe.

It remains to be seen whether the spectral measure can really be observed in real data. In principle, this will require processing huge amounts of data to sift out the large events. One is looking for the normalized moments of X_R to stabilize as R gets large. Equivalently, one can project X_R radially onto the unit sphere: the resulting measure should be approximated by Γ for R large. If the spectral measure does exist in real equity data, this will be a strong verification that some form of self–organized criticality is true in financial markets.

REFERENCES

1. P. Bak. *How nature works: the science of self–organized criticality* Copernicus, Springer-Verlag, 1996
2. J. Boland, T. R. Hurd, M. Pivato and L. Seco. Extreme value dependences in multivariate financial time series working paper, 2000. `http://www.phimac.org`
3. J. P. Bouchaud and M. Potters. *Theory of financial risk.* To appear Aug. 2000
4. M. Dacorogna, U. Müller, O. Pictet and C. de Vries. Olsen & Associates, preprint, 1998. `http://www.olsen.ch/library/research/oa_working.html`
5. J. Farmer. *Computing in Science and Engineering (IEEE)*, Nov.-Dec., 1999, 26–39.
6. P. Gopikrishnan, V. Plerou, L. Amaral, M. Meyer, H. Stanley. *Phys. Rev.* E 60: 5305, 1999; and
   V. Plerou, P. Gopikrishnan, L. Amaral, M. Meyer, H. Stanley. *Phys. Rev.* E 60: 6519, 1999
7. U. Keller and E. Eberlein. *Bernoulli*, 1:281–299, 1995.
8. R. Kiesel and N .H. Bingham. Modelling asset returns with hyperbolic distributions. preprint, 1999.
9. F. Longstaff. Stochastic volatility and option valuation, UCLA School of Management preprint 1995.
10. D. B. Madan, P. Carr, and E. Chang. *European Finance review*, 2:79–105, 1998.
11. B. Mandelbrot. *J. of Business*, 36:394–419, 1963.
12. B. Mandelbrot and J. W. Van Ness. *SIAM Review*, 10:422–437, 1968.
13. G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes* Chapman & Hall, New York, 1994