Semi-parametric modelling in finance: theoretical foundations

N H Bingham[1] and Rüdiger Kiesel[2,3]

[1] Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, UK

[2] Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, UK

E-mail: nick.bingham@brunel.ac.uk and r.t.kiesel@lse.ac.uk

Abstract

The benchmark theory of mathematical finance is the Black–Scholes–Merton theory, based on Brownian motion as the driving noise process for asset prices. Here the distributions of returns of the assets in a portfolio are multivariate normal. The two most obvious limitations here concern symmetry and thin tails, neither being consistent with real data. The most common replacements for the multinormal are parametric—stable, generalized hyperbolic, variance gamma. In this paper we advocate the use of semi-parametric models for distributions, where the mean vector μ and covariance Σ are parametric components and the so-called density generator (function) $\mathbf{g}$ is the non-parametric component. We work mainly within the family of elliptically contoured distributions, focusing particularly on normal variance mixtures with self-decomposable mixing distributions. We show how the parametric cases can be treated in a unified, systematic way within the non-parametric framework and obtain the density generators for the most important cases.

1. Introduction

The benchmark theory in mathematical finance—the Black–Scholes–Merton theory—is based on normal (or Gaussian) driving noise. This theory leads to Gaussian asset return distributions—*normal models* or *Gaussian models*—and has much to recommend it, by way of mathematical tractability, familiarity and completeness of markets (at least in the standard set-up). Its limitations are equally well known: it commits one to *symmetry* and to *very thin tails*, neither consistent with the reality of financial data. This has motivated numerous proposals for alternative models. The first—*stable models*—suggests replacing the Gaussian noise (Wiener process) by a stable process, thereby giving tails that decay like a power rather than log–quadratic decay. This proposal goes back

to Mandelbrot (1963); for monograph treatments, see e.g. Mandelbrot (1997), part IV or Rachev and Mittnik (2000).

The length of the time-interval over which returns are calculated comes into play here—long intervals lead to the Gaussian model, at least approximately, by *aggregational Gaussianity*. Stable distributions preserve the type of distribution under time aggregation, a perfectly natural property from the economic standpoint. However, stable distributions are so heavy-tailed that the second moment is infinite, a fact that is inconsistent with empirical findings for most financial time series. An alternative uses models based on the generalized hyperbolic distribution, advocated by Eberlein and others, see e.g. Eberlein (2001) for an overview. Here a driving Lévy process—*hyperbolic Lévy motion*—is constructed with one-period return distributions from the class of generalized hyperbolic distributions fitted

---

[3] Author to whom any correspondence should be addressed.

to the empirical observed returns. As well as allowing for heavy-tailed and skewed returns, these models are also reasonably consistent under time aggregation. Let us mention that the variance gamma model, proposed by Madan and Seneta (1990) and extended in Carr *et al* (2000), where the driving Brownian motion is time-changed by a gamma process, can be subsumed within the above framework. All these are examples of parametric models. These have the advantages of mathematical tractability, but the drawback that complex reality cannot be captured accurately by a small number of parameters.

More generally, one can work non-parametrically, with an arbitrary infinitely-divisible distribution, and correspondingly, by the Lévy–Khintchine formula, with an arbitrary *Lévy process* as driving noise, in place of the Gaussian noise in the classical case. This leads to the notion of general *Lévy models* for asset returns. The principal drawbacks here are greater mathematical complexity, and the *independent increments* assumption involved in using Lévy processes to model the driving noise. In reality, the noise terms over contiguous intervals will certainly be dependent (particularly in periods of market turbulence, which are atypical but crucial). Lévy processes are also well adapted to the modelling of *stochastic volatility*; for a treatment focusing on the dynamic (time-series) aspects here, we refer to Barndorff-Nielsen and Shephard (2001a).

The aim of the paper is to set up a general framework for Lévy-type models, with particular reference to the multidimensional case. This is relevant to portfolio theory, and to questions of asset allocation. We advocate a *semi-parametric* approach to combine the advantages of the parametric and non-parametric approaches. We use a *parametric* component, incorporating the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and a *non-parametric* component, modelling the *shape* of the distribution, specifically, questions of tail-decay (kurtosis). Here, shape, which we can think of as a density on $[0, \infty)$, incorporates what remains when we work up to location and scatter, that is modulo affine transformations, while $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents the affine part. Hodgson *et al* (2000) also study semi-parametric modelling in a similar setting, from the point of view of empirical testing of the capital asset pricing model (CAPM).

In section 2 we discuss normal variance-mean mixtures as a general class for the underlying return distributions. We show that all parametric classes of distributions used to set up the above-mentioned Lévy models fall within this class.

In section 3 we turn to the framework of *elliptical models* or of *elliptically contoured distributions* (Bickel *et al* 1998, sections 4.2, 6.3, 7.2, 7.8, Fang *et al* 1990, ch 2, 3, see section 3). This has a number of important advantages.

(1) The roles of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are preserved, and the shape component is flexible enough to handle anisotropy (though not general asymmetry) and any desired tail decay.

(2) The benchmark normal/Gaussian theory, and the major special case ($\boldsymbol{\beta} = 0$, section 2) of its principal competitor, the hyperbolic/normal inverse Gaussian (NIG) theory, are conveniently contained and extensively generalized.

(3) *Adaptive* estimation is possible, that is ignorance of one of the parametric and non-parametric components need extract no price in efficiency when estimating the other (Bickel *et al* 1998).

(4) Elliptical models have *linear regression* (Fang *et al* 1990, Embrechts *et al* 2001). This enables the methods and results of classical statistics—the linear model and regression—to be brought to bear.

Elliptic models are well adapted to incorporate *infinite divisibility*, a concept which we discuss in section 4. Both the normal/Gaussian and hyperbolic/NIG models are infinitely divisible. Infinite divisibility is the key feature of a distribution that allows the introduction of Lévy processes to model the driving noise, generalizing the r-dimensional Brownian motion in the Black–Scholes–Merton model (albeit at the cost of market completeness). Within the family of infinitely divisible laws, the *self-decomposable* laws are particularly important. Our main focus is on self-decomposable elliptic distributions; both the normal/Gaussian and hyperbolic/NIG families (for $\boldsymbol{\beta} = 0$—see below) are included here (Halgreen 1979). Self-decomposable distributions are also important in the time-series (dynamic) aspects, rather than the static distributional aspects studied here. We turn in section 5 to the important class of distributions generated as normal variance mixtures and characterize those distributions which can be used to construct Lévy-type models. In section 6 we discuss the class of generalized hyperbolic distributions within our framework and obtain an explicit expression for the density generator of distributions within this class. We also provide the decay rate for the density generator, an important quantity for value-at-risk (VaR) considerations. We close in section 7 with an outline of how to implement and apply the semi-parametric model. A detailed analysis of the econometric properties and applications to risk management will be contained in a forthcoming paper (Bingham *et al* 2002).

2. Normal variance-mean mixtures

If U is a random variable on $[0, \infty)$ with law F, and

$$\boldsymbol{X} | (U = u) \sim N_r(\boldsymbol{\mu} + u\boldsymbol{\beta}, u\Delta),$$

where Δ is a symmetric positive definite $r \times r$ matrix with determinant one and $\boldsymbol{\mu}$, $\boldsymbol{\beta}$ are r vectors, then the distribution of $\boldsymbol{X}$ is called a *normal variance-mean mixture* with position $\boldsymbol{\mu}$, drift $\boldsymbol{\beta}$, structure matrix Δ ($|\Delta| = 1$ is imposed to ensure identifiability) and mixing distribution F. If $\boldsymbol{\beta} = 0$, a case studied in more detail later, X is a *normal variance mixture*. The density $f_X(\boldsymbol{x})$ of $\boldsymbol{X}$ is

$$\exp\{(\boldsymbol{x} - \boldsymbol{\mu})^T \Delta^{-1} \boldsymbol{\beta}\}$$
$$\times \int_0^\infty \frac{1}{(2\pi u)^{\frac{1}{2}r}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T (u\Delta)^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \right.$$
$$\left. -\frac{1}{2}u\boldsymbol{\beta}^T \Delta^{-1} \boldsymbol{\beta} \right\} \mathrm{d}F(u),$$

and its characteristic function (CF) is

$$\psi_X(t) = \int_{\mathbb{R}^r} \exp\{it^T x\} f_X(x)\, dx$$

$$= \int_{\mathbb{R}^r} dx \exp\{it^T x\} \int_0^\infty f_{X|U}(x|u)\, dF(u)$$

$$= \int_0^\infty dF(u) \exp\{it^T \mu + it^T \beta u - \tfrac{1}{2} t^T \Delta u t\},$$

interchanging the order of integration and using $X|(U = u) \sim N_r(\mu + u\beta, u\Delta)$. Thus if

$$\phi(s) := \int_0^\infty e^{-su}\, dF(u) \qquad (s > 0)$$

is the Laplace–Stieltjes transform of U,

$$\psi_X(t) = \exp\{it^T \mu\} \phi(\tfrac{1}{2} t^T \Delta t - it^T \beta). \qquad \text{(NVMM)}$$

The distribution is isotropic iff $\beta = 0$ and $\Delta = I$.

We see immediately that ψ is *infinitely divisible* iff ϕ is. *Self-decomposability* transfers from F, ϕ to ψ only for $\beta = 0$ in general.

For $\beta = 0$, the distribution has *elliptical symmetry*, as with the multivariate normal. This is very convenient technically; see below. However, this comes at the price of being able to model asymmetry.

One prime example is that of F the generalized inverse Gaussian (GIG) distribution. Recall the Bessel function K_λ of the third kind (or Macdonald function) and its integral representation

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty u^{\lambda-1} \exp\left\{-\frac{1}{2} x\left(u + \frac{1}{u}\right)\right\} du \qquad (x > 0)$$

(Watson 1944, sections 3.7 and 6.22, (8)). Observe that $K_\lambda = K_{-\lambda}$, on writing $1/u$ for u; also $K_{\pm\frac{1}{2}}(x) = \sqrt{\pi/(2x)}\, e^{-x}$. Then for $\lambda \in \mathbb{R}$, $\delta, \gamma > 0$

$$f(x) = \frac{(\gamma/\delta)^\lambda}{2K_\lambda(\delta\gamma)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\gamma^2 x + \frac{\delta^2}{x}\right)\right\} \qquad (x > 0)$$

is a probability density on $\mathbb{R}_+$. Its Laplace transform is

$$\zeta(s) = \left(\frac{\delta^2}{\gamma^2 + 2s}\right)^{\frac{1}{2}\lambda} \frac{K_\lambda(\delta\sqrt{\gamma^2 + 2s})}{K_\lambda(\delta\gamma)} \qquad (s > 0), \quad (K_\lambda)$$

which we shall refer to as the (K_λ)-*formula*. Of course, this functional form is to be expected since f belongs to an *exponential family*; see e.g. Barndorff-Nielsen (1978), section 8.1. This law is called the GIG or $\text{GIG}_{\lambda,\delta,\gamma}$ (and we use $d_{GIG}(x; \lambda, \delta, \gamma)$ as notation for its density). The *inverse Gaussian* case is $\lambda = 1$: $IG_{\delta,\gamma} = \text{GIG}_{1,\delta,\gamma}$. The name arises because these are the first hitting-time distributions for drifting Brownian motion; see Barndorff-Nielsen *et al* (1978) for hitting times, Jørgensen (1982) for a monograph treatment of statistical aspects. It was shown by Barndorff-Nielsen and Halgreen (1977) that GIG is infinitely divisible, and by Halgreen (1979) that GIG is self-decomposable.

We now form normal mean-variance mixtures, using the GIG laws as mixing distributions. The resulting normal mean-variance mixture has density

$$f(x) = \frac{(\gamma/\delta)^\lambda}{(2\pi)^{\frac{1}{2}r} \alpha^{\lambda-\frac{r}{2}} K_\lambda(\delta\gamma)}$$
$$\times (\delta^2 + (x-\mu)^T \Delta^{-1}(x-\mu))^{(\lambda-\frac{r}{2})/2}$$
$$\times K_{\lambda-\frac{1}{2}r}\left(\alpha\sqrt{\delta^2 + (x-\mu)^T \Delta^{-1}(x-\mu)}\right)$$
$$\times \exp\{\beta^T(x-\mu)\}$$

where

$$\alpha^2 := \gamma^2 + \beta^T \Delta^{-1}\beta.$$

The class of distributions with these densities is called the *generalized hyperbolic distributions*. In the univariate case ($r = 1$) the densities depend on five parameters (recall $|\Delta| = 1$) with the following interpretation: $\alpha > 0$ determines the shape; $0 < |\beta| < \alpha$ is a skewness parameter; $\mu \in \mathbb{R}$ determines the location; δ is a scaling parameter comparable to σ; $\lambda \in \mathbb{R}$ describes some subclasses. Affine-invariant (no change under change of location and scale) parametrizations are given by

$$\zeta = \delta\sqrt{\alpha^2 - \beta^2}, \qquad \rho = \frac{\beta}{\alpha},$$

and

$$\xi = (1 + \zeta)^{-\frac{1}{2}}, \qquad \chi = \xi\rho.$$

Since $0 \leqslant |\chi| < \xi < 1$ the distribution parametrized by χ and ξ can be represented by points of a triangle, the so-called shape triangle (see Shiryaev (1999) for further discussion). The simplest case is the univariate case $r = 1$ with $\lambda = 1$. Then as $K_{\frac{1}{2}}(x) = \sqrt{\pi/(2x)}e^{-x}$, f simplifies to

$$f(x) = \frac{\gamma}{2\alpha\delta K_1(\delta\gamma)} \exp\left\{-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)\right\}.$$

The log-density has graph

$$y = c - \alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)$$

for some constant c, which is the lower branch of a *hyperbola*, with *asymptotes*

$$y - c = (\beta \pm \alpha)(x - \mu).$$

Note that $\beta \neq 0$ is needed to obtain asymptotes with *asymmetric slopes*, that is not of the form $\pm\alpha$.

Now the empirical findings of Bagnold (1941) were that, for the distribution of sizes of particles of sand, if log-density is plotted against log-size, one obtains an approximation to a smooth unimodal curve approaching two linear asymptotes (with asymmetric slopes in general) at $\pm\infty$. The simplest such curve is the lower branch of a hyperbola, and this motivates the definition of the generalized hyperbolic distribution above and its interpretation as a normal variance-mean mixture.

A further important case is the NIG, where $\lambda = -1/2$ and the density is

$$d_{NIG}(x) = \frac{\alpha}{\pi} \exp\left\{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right\}$$

$$\times \frac{K_1\left(\alpha\delta\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}}.$$

The NIG distributions form the only subclass of the generalized hyperbolic laws which are closed under convolution:

$$\text{NIG}(\alpha, \beta, \delta_1, \mu_1) * \text{NIG}(\alpha, \beta, \delta_2, \mu_2)$$
$$= \text{NIG}(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2).$$

Finally let us mention the *variance gamma distribution* obtained by using a gamma distribution as its mixing distribution, which is the special case $\lambda > 0$, $\delta = 0$ and $\gamma > 0$ in the above framework:

$$d_\Gamma(x) = \left(\frac{\gamma^2}{2}\right)^\lambda \frac{1}{\Gamma(\lambda)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\gamma^2 x\right\}, \qquad x > 0.$$

These distributions have been used in the context of variance-gamma models in Madan and Seneta (1990), Carr *et al* (1998).

We now have some choice as to how to proceed. We can use normal mean-variance mixtures (general β). This gives a semi-parametric model, extending the hyperbolic model of Bagnold, Barndorff-Nielsen, Eberlein and others above. This allows us to model asymmetry, but complicates the estimation theory. Alternatively, we can specialize to $\beta = 0$, that is to normal variance mixtures. This restricts us to ellipsoidal symmetry (as with the multivariate normal)—though not to isotropy. However, because the distribution is now elliptically contoured, it is much easier to handle and the estimation is much simplified.

We restrict attention to the elliptically contoured case $\beta = 0$ below. Here we can give a fairly complete treatment. The case of general β—as in the extension of Korsholm (2000) to higher dimensions, and the dynamic aspects, we propose to discuss elsewhere.

3. Elliptical models

An r-dimensional distribution is *spherically symmetric* if it is invariant under the action of the orthogonal group $O(r)$. The density f is then a function of $x := \|\boldsymbol{x}\|$ rather than of $\boldsymbol{x}$, and similarly the CF ψ is a function of $t := \|\boldsymbol{t}\|$ rather than of $\boldsymbol{t}$.

An r-dimensional distribution is *elliptically contoured* if it is the image of a spherically symmetric distribution under an affine transformation. We shall confine attention, for simplicity, to the absolutely continuous, full-rank, L_2 case. Then the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ are defined, $\boldsymbol{\Sigma}$ is non-singular, and the density f is a function of the quadratic form $Q := (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$:

$$f(\boldsymbol{x}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g(Q) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g((\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})). \quad \text{(EC)}$$

We write $f \sim EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, and call $g : \mathbb{R}_+ \to \mathbb{R}_+$ the *density generator* of f, or 'shape'. Then $\theta := (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is

the *parameter*, or parametric part, of the model, g the non-parametric part. The CF ψ of f is of the form

$$\psi_X(\boldsymbol{t}) := \mathbb{E} \exp\{\mathrm{i}\boldsymbol{t}^T \boldsymbol{X}\} = \exp\{\mathrm{i}\boldsymbol{t}^T \boldsymbol{\mu}\} \phi(\boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t}) \qquad \text{(EC')}$$

for some scalar function ϕ called the *characteristic generator* of ψ, or f (Fang *et al* 1990 ch 2, Cambanis *et al* 1981). It is convenient to write here $f \sim EC_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \phi)$ also.

To see the link between g in (EC) and ϕ in (EC'), pass to the *standardized* variable $\boldsymbol{Y} := \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{X} - \boldsymbol{\mu})$ with mean vector $\boldsymbol{0}$ and covariance matrix $\boldsymbol{I}$. Then

$$\phi(t^2) = \phi(\boldsymbol{t}^T \boldsymbol{t}) = \psi_Y(\boldsymbol{t}) = \int_{\mathbb{R}^r} \exp\{\mathrm{i}\boldsymbol{t}^T \boldsymbol{y}\} g(\boldsymbol{y}^T \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}$$

$$= \int_{\mathbb{R}^r} \exp\{\mathrm{i}\boldsymbol{t}^T \boldsymbol{y}\} g(y^2) \, \mathrm{d}\boldsymbol{y}.$$

Thus $g(y^2)$ is the density $f_Y(\boldsymbol{y})$ of $\boldsymbol{Y}$, which being a function of y only we can write as $f_Y(y)$. The density generator thus satisfies

$$g(y^2) = f_Y(\boldsymbol{y}) = f_Y(y). \qquad \text{(DG)}$$

One has the *stochastic representation*

$$\boldsymbol{x} = \boldsymbol{\mu} + R A^T \boldsymbol{u}, \qquad A^T A = \boldsymbol{\Sigma}, \qquad \text{(SR)}$$

where $\boldsymbol{u}$ is uniformly distributed on the unit r sphere and $R > 0$ is a scalar random variable with

$$Q(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = R^2 = \|\boldsymbol{y}\|^2 = y^2, \quad \text{(QF)}$$

in terms of the standardized variables above. Thus (EC) says that the quadratic form $Q = R^2$ is what matters. Its density h is given in terms of the density generator g by

$$h(u) = \frac{\pi^{r/2}}{\Gamma(r/2)} u^{r/2-1} g(u). \qquad \text{(hg)}$$

Scaling and identifiability. In (EC), one can absorb a scale-factor $c > 0$ in the quadratic form Q on the right, adjusting g accordingly. If one passes from one representation $(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ of f to another, say $(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*, g^*)$, in this way, the transformation formulae are

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}, \qquad \boldsymbol{\Sigma}^* = c\boldsymbol{\Sigma},$$
$$\phi^*(.) = \phi(./c), \qquad g^*(.) = c^{\frac{1}{2}r} g(c.).$$

The covariance matrix and the matrix parameter $\boldsymbol{\Sigma}$ are linked by

$$\text{cov}(\boldsymbol{x}) = -2\phi'(0)\boldsymbol{\Sigma} = \frac{\mathbb{E}(R^2)}{r}\boldsymbol{\Sigma} = \frac{\mathbb{E}(Q)}{r}\boldsymbol{\Sigma}.$$

One thus has some choice as to scaling. Because we wish to maintain the direct link to the Markowitz mean-variance theory, the natural choice of scale for us is to take, throughout,

$$\text{cov}(\boldsymbol{x}) = \boldsymbol{\Sigma}$$

(or $E(R^2) = r$) and we shall do this. The above ambiguity of scale then disappears, and the model (EC) becomes identifiable.

The elliptically contoured framework applies more generally than above (densities need not exist, nor need means

and covariances, Σ may be singular). We restrict as above for convenience and interpretability; see Fang *et al* (1990) section 2.5 for further detail and references.

Curse of dimensionality. The dimensionality in which we work is the number of different assets in our portfolio, which may be large. Even reducing from individual stocks to equity indices may still leave the dimensionality high. For instance, work on asset allocation for a world equity portfolio, Bodie *et al* (1999) considers 48 equity indices. One of the principal advantages of our elliptically contoured model is that, although our insistence on retaining the Markowitzian interpretations of μ and Σ dictates the dimensionality of the parametric part, the harder—non-parametric—part g of our model, being a *scalar* function, *escapes the curse of dimensionality*. This is extremely convenient from a computational viewpoint; see the simulation studies in Bingham and Kiesel (2001a) and the treatment of real financial data in Bingham *et al* (2002).

Radial characteristic functions. There are two ways to handle the vector u in (SR) uniformly distributed on the unit sphere. One way is to use $u^T u = 1$, leading to (QF) and (DG) above. The other, to which we now turn, is to integrate over the unit sphere.

In the presence of radial symmetry, the r-dimensional Fourier transform above decomposes into two parts, the radial part, which carries the information, and the spherical part. This decomposition gives rise to a variant on the Hankel transform, an integral transform with kernel the modified Bessel function

$$\Lambda_\nu(u) := \frac{J_\nu(u)\Gamma(\nu+1)}{(\frac{1}{2}u)^\nu}, \qquad \nu := \tfrac{1}{2}(r-2).$$

Then (see Bochner and Chandrasekharan 1949 IV.5, Kingman 1963) the CF is

$$\mathbb{E}\Lambda_\nu(tY) = \int_0^\infty \Lambda_\nu(ty)\, \mathrm{d}H(y),$$

where H is the distribution function of Y. This is called the *radial CF*, $\Psi(t)$ say, of the *radial distribution H*:

$$\Psi(t) = \int_0^\infty \Lambda_\nu(ty)\, \mathrm{d}H(y).$$

Then

$$\Psi(t) = \phi(t^2) = \int_{\mathbb{R}^r} \exp\{\mathrm{i}t^T y\} g(y^2)\, \mathrm{d}y$$
$$= \int_0^\infty \Lambda_\nu(ty)\, \mathrm{d}H(y).$$

Unimodality. We are interested here in modelling densities $f_X(x)$ that have a unique mode at $x = \mu$. This is ensured by restricting the density generator g in (EC) to be *decreasing* on $[0, \infty)$, and we shall do this throughout (equivalently, we restrict to $f_Y(y)$ decreasing in y).

Decomposition. The elliptically contoured form (EC) conveniently splits the functional form of f into two parts, which behave differently with respect to affine transformations. The density generator g—the non-parametric part—governs the *shape* of f, which is what remains when we work modulo affine transformations. The parametric part $\theta = (\mu, \Sigma)$ may be thought of as the affine part of the model, and a

desirable property of estimators of it is *affine equivariance*, that is commutability with affine transformations, or changes of location and scatter. For background, see e.g. Lopuhaä and Rousseeuw (1991), Lopuhaä (1999).

Examples.

(1) The *normal/Gaussian* case. The density generator of the multivariate normal distribution is given by

$$g(u) = (2\pi)^{-\frac{1}{2}r} \exp\{-\tfrac{1}{2}u\}.$$

The CF is

$$\psi(t) = \exp\{\mathrm{i}t^T\mu - \tfrac{1}{2}t^T\Sigma t\},$$

so

$$\phi(u) = \exp\{-\tfrac{1}{2}u\}.$$

The multivariate normal is a member (the $N = s = 1, t = \frac{1}{2}$ case) of the class of symmetric Kotz-type distributions, which are characterized by exponentially decaying density generators of form

$$g(u) = C_r u^{N-1} \exp\{-tu^s\}, \qquad s, t > 0, 2N + r > 2,$$

where C_r is a constant.

(2) The *multivariate t distribution*. For the multivariate t distribution with m degrees of freedom the density generator exhibits power decay. It is a member (the $N = \frac{1}{2}(r+m), m$ an integer case) of the class of symmetric multivariate Pearson type VII distributions with density generators

$$g(u) = (\pi m)^{-\frac{1}{2}r} \frac{\Gamma(N)}{\Gamma(N - r/2)} \left(1 + \frac{u}{m}\right)^{-N},$$
$$N > r/2, m > 0.$$

Limitations. We point out three limitations of the elliptical model.

(1) No real data will be exactly elliptical (though we can test for ellipsoidal symmetry: Beran (1979), Li *et al* (1997)). Within the family of all distributions, any neighbourhood of an elliptical distribution will contain non-elliptical ones. Because of this, Hampel *et al* (1986, remark 2, p 273) go so far as to describe the generalization beyond the parametric case as 'spurious'. This criticism could, in principle, be levelled against any semi-parametric model. We prefer to make use of the modelling advantages of the elliptical approach, while bearing robustness questions in mind when estimating the component parts of the model.

(2) Elliptical distributions have linear regression. While this is technically very convenient, and reasonable for portfolios of stocks, it is less suitable for portfolios containing options. Since options have value—they convey rights, but not obligations (and so are examples of 'contingent claims')—they are assets in their own right and can be traded. Indeed, the prices at which options are traded in the market provide us with the means to estimate the 'implied volatility' of the underlying

stocks. However, the pay-off functions of options are highly nonlinear: $(S - K)_+$, $(K - S)_+$ in the commonest cases, of European calls and puts. The Black–Scholes value of an option, being a smoothed version of the pay-off after suitable expectation, inherits this pronounced nonlinearity. Consequently the theory presented here is not suitable for use with portfolios containing a significant proportion of options rather than stocks.

(3) As noted earlier, in the important special case of normal variance-mean mixtures, only the special case $\beta = 0$ of normal variance mixtures is elliptically contoured, but only the case $\beta \neq 0$, for the prototypical example of the hyperbolic case, can model asymmetry.

4. Infinite divisibility and self-decomposability

We recall that infinitely divisible distributions are characterized through their CFs via the Lévy–Khintchine formula, and are the marginal laws of Lévy processes, i.e. stochastic processes with stationary independent increments. For a full account of both aspects, we refer to the monographs of Bertoin (1996) and Sato (1999).

Infinitely divisible laws are those obtained as limit laws of asymptotically negligible triangular arrays. Consequently, infinite divisibility is preserved under affine transformations. Thus in testing an elliptically contoured distribution for infinite divisibility, we may without loss take $\boldsymbol{\mu} = 0$, $\boldsymbol{\Sigma} = \boldsymbol{I}$, which reduces matters to g or ϕ. So an elliptically contoured distribution is infinitely divisible iff, for each $k = 1, 2, \ldots$, its characteristic generator satisfies

$$\phi = \phi_k^k$$

with ϕ_k also a characteristic generator (example: the normal case, with $\phi_k = \exp(-\frac{1}{2}u/k)$).

Unfortunately, we have no such way of testing for infinite divisibility using the density generator g, although this is of more direct concern from the modelling point of view and more accessible statistically. Indeed, we know from the form of the Lévy–Khintchine formula that describing the class of infinite divisible laws requires transform language.

A subclass of the infinitely divisible laws is obtained by specializing the two-index triangular array to one index, arising from a single sequence of affine transformations. Such infinitely divisible laws are called *self-decomposable* (or of Lévy's class L). They may be characterized alternatively by the property that, for each $\rho \in (0, 1)$,

$$\psi(t) = \psi(\rho t)\psi_\rho(t),$$

with ψ_ρ again a CF (here ψ_ρ is uniquely determined, and also infinitely divisible). For details, see Sato (1999), section 15, 17, 53, Feller (1971), XVII 8.

For elliptically contoured distributions, the self-decomposability condition becomes, for characteristic generators,

$$\phi(u) = \phi(\rho u)\phi_\rho(u) \qquad \text{(SD)}$$

for each $\rho \in (0, 1)$, with ϕ_ρ again a characteristic generator (example: the normal case, with $\phi_\rho(u) = \exp(-\frac{1}{2}(1 - \rho)u)$).

We restrict attention to self-decomposable distributions here, for four reasons.

(1) Self-decomposable distributions are absolutely continuous and unimodal (Sato 1999, theorems 28.4 and 53.1). Thus, for self-decomposable elliptically contoured distributions, the density generator g exists and is decreasing, as we wish (section 3).

(2) Self-decomposable laws arise as marginal laws in autoregressive time-series models

$$X_t = \rho X_{t-1} + \epsilon_t,$$

where the innovation s (or errors) ϵ_t are independent, of each other and of X_s, $s < t$ (Bondesson 1981; cf (SD)).

(3) Self-decomposability is transferred from the mixing distribution to a normal variance mixture (section 5 below). Our principal specific examples are of this type. Also, normal variance mixtures are capable of modelling a wide variety of real data sets, for structural reasons (Romanowski 1979, Barndorff-Nielsen *et al* 1982).

(4) Barndorff-Nielsen *et al* (1988) have recently studied time series arising in finance (and turbulence) with various types of dependence structure and NIG marginals (section 6). They use the theory of stochastic processes of Ornstein–Uhlenbeck type (Sato (1999, section 17)). This depends on self-decomposability, rather than the specifics of the parametric NIG model, a theme developed further in Barndorff-Nielsen and Shephard (2001b).

Radial CF. The infinitely divisible radial CFs $\Psi(t) = \psi(t)$ are known (Kingman 1963, section 6): they are those of the form

$$\Psi(t) = \exp\left\{-c \int_0^\infty (1 - \Lambda_v(tx))\left(\frac{(1 + x^2)}{x^2}\right) \mathrm{d}G(x)\right\}$$

with $c > 0$ and G a probability distribution on $\mathbb{R}_+$. This is analogous to the form $\phi(s) = \exp\{-c \int_0^\infty (1 - \mathrm{e}^{-sx})((1 + x)/x) \mathrm{d}G(x)\}$ for infinitely divisible Laplace–Stieltjes transforms on $\mathbb{R}_+$ (section 5 below and Feller 1971, XIII.7). Both are instances of more general formulae for Urbanik's theory of generalized convolution algebras. For details and references, see Urbanik (1964), Bingham (1971, 1984).

Of course, recognizing infinite divisibility from this criterion may be difficult given a specific Ψ or ψ, and impossible given a specific g. This underlines the importance of structural aspects such as self-decomposability, which ensures infinite divisibility, or the normal variance mixture property (section 5), which makes it easier to recognize.

Lévy-type models of financial markets. Recall the standard Black–Scholes model defined via the SDE

$$\mathrm{d}S_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}W_t),$$

with constant coefficients and a standard Brownian motion W. The solution of the SDE is

$$S_t = S_0 \exp\left\{\mu t - \frac{\sigma^2}{2}t + \sigma W_t\right\}.$$

One can consider a general exponential Lévy process model for stock prices

$$S_t = S_0 \exp(L_t),$$

with a Lévy process L. Now the distribution of a Lévy process is uniquely determined by any of its one-dimensional marginal distributions, so we may use the distribution of L_1, i.e. F_1. Then F_1 is infinitely divisible and its CF is given by the Lévy–Khintchine formula as

$$\mathbb{E}(\exp\{iuL_1\}) = \exp\{-\psi(u)\},$$

with the Lévy exponent ψ of L_1 given by

$$\psi(u) = \frac{c^2}{2}u^2 - i\alpha u + \int_{\{|x|<1\}}(1 - e^{-iux} - iux)\mu(dx)$$
$$+ \int_{\{|x|\geq 1\}}(1 - e^{-iux})\mu(dx),$$

with $\alpha, c \in \mathbb{R}$ and μ a σ-finite measure on $\mathbb{R}/\{0\}$, the Lévy measure, satisfying

$$\int \min\{1, x^2\}\mu(dx) < \infty.$$

The Lévy–Khintchine formula can now be used to read off properties of the corresponding Lévy process L_t, see Protter (1992), I.4. Examples of such models using the distributions introduced above are

 (i) hyperbolic distributions, see Eberlein and Keller (1995) and Eberlein *et al* (1998).
 (ii) NIG distributions, see Barndorff-Nielsen (1997, 1998).
(iii) Generalized hyperbolic distributions, see Eberlein (2001).
(iv) Variance-gamma model, see Madan and Seneta (1990) and Carr *et al* (1998) and its extension the CGMY model, see Carr *et al* (2000).

The normal variance-mean framework unifies the above approaches. Instead of specifying a parametric family, such as the GIG laws for the underlying mixing distribution F, we can estimate F (or its density) non-parametrically. The corresponding estimation problem has been treated by Korsholm (2000) in the one-parametric case.

5. Normal variance mixtures

Suppose that U has law F on $(0, \infty)$, and that given $U = u$, x is r-variate normal with mean μ and covariance matrix $u\Sigma$:

$$X|(U = u) \sim N_r(\mu, u\Sigma).$$

Then X is called a *normal variance mixture*. Such mixtures have been studied in, e.g. Barndorff-Nielsen *et al* (1982)

(taking $\beta = 0$ there) and Fang *et al* (1990) section 2.6. From (NVMM) we find that their CFs have the form

$$\psi_X(t) = \exp\{it^T\mu\}\phi_M(\tfrac{1}{2}t^T\Sigma t), \qquad \text{(NVM)}$$

where $\phi_M(s) := \int_0^\infty e^{-su}dF(u)(s > 0)$ (we use the suffix M for 'mixture', rather than introduce another Greek letter) is the Laplace–Stieltjes transform of U. Comparing with (EC'), we see that *all normal variance mixtures are elliptically contoured*, with characteristic generator

$$\phi(s) = \phi_M(s/2).$$

Conversely, any elliptically contoured distribution whose characteristic generator is completely monotone (or, by Bernstein's theorem, is a Laplace–Stieltjes transform) can arise in this way. This is essentially the content of Schoenberg's theorem (see Bingham 1973). The possible ϕ, or ϕ_M, that can arise belong to

$$\Phi_\infty := \cap_{n=1}^\infty \Phi_n,$$

where

$$\Phi_n := t\{h : h(t_1^2 + \cdots + t_n^2) \text{ is an } n\text{-dimensional}$$
$$\text{characteristic function}\}.$$

Then $\phi \in \Phi_\infty$ iff ϕ is of the form

$$\phi(x) = \int_0^\infty e^{-xr^2}dG(r)$$

for some law G on $(0, \infty)$ (that is under the change of variable $r^2 = u$, ϕ is completely monotone, as above).

Infinite divisibility. The question of infinite divisibility, or self-decomposability, of normal variance mixtures thus transfers, by (NVM), to ϕ_M, or the mixing law F: the law of X is infinitely divisible if F is, etc. Now the infinitely **Q. 1** divisible laws on $[0, \infty)$ are known: their Laplace transforms are of the form $\phi(s) = \exp\{-h(s)\}$, where $h(0) = 0$ and h' is completely monotone (see e.g. Feller 1971, XIII.7). That is, the infinitely divisible ϕ are

$$\phi(s) = \exp\left\{-\int_0^\infty (1 - e^{-sx}\,d\mu(x))\right\},$$

where the Lévy measure μ satisfies $\int_0^1 x\,d\mu(x) < \infty$. Infinite divisibility of normal variance mixtures was studied by Kelker (1971).

Self-decomposability. The further restriction to self-decomposability requires further that the Lévy measure be absolutely continuous, with density $\nu(x) = k(x)/x$ with $k(.)$ decreasing (Santo 1999, corollary 15.11).

Simulation. Given any decreasing k on $\mathbb{R}_+$ integrable at the origin, we can form the Lévy measure $k(x)\,dx/x$, and simulate random variables U from it, using the method of Bondesson (1982) involving shot-noise processes. We can simulate $Z \sim N_r(0, I)$ as a vector $(Z_1, \ldots, Z_r)$ of independent standard normal variates, and then simulate X as $\mu + U^{\frac{1}{2}}\Sigma^{\frac{1}{2}}Z$.

We thus have a wealth of examples of self-decomposable normal variance mixtures to hand, one for each decreasing k, from which we can simulate at will.

6. The hyperbolic and NIG models

We now return to the discussion of the *generalized hyperbolic distributions*, GH $= \mathrm{GH}_{l,\gamma,\delta;\mu,\Sigma}$ in the framework of normal variance mixtures, using the GIG laws as mixing distributions (with $\beta = 0$, see section 2). By sections 3–5, the CF is

$$\psi(t) = \exp\{it^T\mu\}\zeta(t^T\Sigma t)$$

with ζ as in the K_λ formula, and the laws GH are elliptically contoured, infinitely divisible and self-decomposable.

In the notation of section 2, we have $\phi = \zeta$, $F = \mathrm{GIG}$. From section 5, the infinite divisibility and self-decomposability of GIG transfer to GH: *the generalized hyperbolic laws* GH *are infinitely divisible and self-decomposable.* Being normal variance mixtures, their characteristic generators are *completely monotone*, and in particular are *decreasing*.

Density generator. Since $X|(U = u) \sim N_r(\mu, u\Sigma)$, one has, writing $Q := (x - \mu)^T\Sigma^{-1}(x - \mu)$ as before,

$$f_{X|U}(x|u) = \frac{1}{(2\pi)^{\frac{1}{2}r}|\Sigma|^{\frac{1}{2}}}u^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}u^{-1}Q\right\}.$$

So

$$f_X(x) = \int_0^\infty f_{X,U}(x, u)\,du = \int_0^\infty f_{X|U}(x|u)f_U(u)\,du$$

$$= \frac{1}{(2\pi)^{r/2}|\Sigma|^{1/2}}$$

$$\times \int_0^\infty u^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}u^{-1}Q\right\}\frac{(\gamma/\delta)^\lambda}{2K_\lambda(\gamma\delta)}$$

$$\times u^{\lambda-1}\exp\left\{-\frac{1}{2}\left(\gamma^2 u + \frac{\delta^2}{u}\right)\right\}du$$

$$= \frac{1}{(2\pi)^{r/2}|\Sigma|^{1/2}}\frac{(\gamma/\delta)^\lambda}{2K_\lambda(\gamma\delta)}$$

$$\times \int_0^\infty u^{\lambda-\frac{3}{2}}\exp\left\{-\frac{1}{2}\left(\gamma^2 u + \frac{(\delta^2 + Q)}{u}\right)\right\}du.$$

The integrand is the density of $\mathrm{GIG}_{\lambda-\frac{1}{2},\gamma,\sqrt{\delta^2+Q}}$ to within normalization, so the integral is $2K_{\lambda-\frac{1}{2}}(\gamma\sqrt{\delta^2+Q})\times(\gamma/\sqrt{\delta^2+Q})^{-(\lambda-\frac{1}{2})}$:

$$f_X(x) = \frac{1}{(2\pi)^{r/2}|\Sigma|^{1/2}}\frac{(\gamma/\delta)^\lambda}{K_\lambda(\gamma\delta)}\frac{K_{\lambda-\frac{1}{2}}(\gamma\sqrt{\delta^2+Q})}{(\gamma/\sqrt{\delta^2+Q})^{(\lambda-\frac{1}{2})}}.$$

From (EC), $f_X(x) = |\Sigma|^{-\frac{1}{2}}g(Q)$, with g the density generator. Comparing, the density generator of the generalized hyperbolic law GH $= \mathrm{GH}_{\lambda,\gamma,\delta,\mu,\Sigma}$ is

$$g(u) = \frac{1}{(2\pi)^{r/2}}\frac{(\gamma/\delta)^\lambda}{K_\lambda(\gamma\delta)}\frac{K_{\lambda-\frac{1}{2}}(\gamma\sqrt{\delta^2+u})}{(\gamma/\sqrt{\delta^2+u})^{(\lambda-\frac{1}{2})}},\qquad u > 0.$$

Now the function ζ above, being a Laplace transform, is completely monotone, so in particular is decreasing. To within constants, g has the same functional form as ζ (with $\lambda - \frac{1}{2}$ in place of λ), apart from interchange of

(i) $K_{\lambda-\frac{1}{2}}$ with $K_{-(\lambda-\frac{1}{2})}$,

(ii) γ with δ.

Now (i) has no effect (recall $K_\nu = K_{-\nu}$ from the integral representation) and (ii) amounts to a reparametrization. Thus g inherits the complete monotonicity of ζ: *the density generator g of the generalized hyperbolic law* GH *is completely monotone, in particular is decreasing.*

One can read off the decay rate of g:

$$K_\nu(x) \sim \sqrt{\frac{\pi}{2x}}e^{-x}\qquad(x \to \infty)$$

(Watson 1944, section 3.71 (13)): thus all K_ν decay like $K_{\pm\frac{1}{2}}(x) = \sqrt{\pi/(2x)}e^{-x}$), so

$$g(u) \sim \mathrm{const}\, u^{\frac{1}{2}\lambda-\frac{3}{4}}\exp\left\{-\gamma\sqrt{\delta^2+u}\right\}\qquad(u \to \infty).$$

Since the argument u replaces the quadratic form $Q = (x - \mu)^T\Sigma^{-1}(x - \mu)$, this shows again that GH has log–linear tail-decay. This should be compared with the log–quadratic tail decay in the normal/Gaussian case, where $g(u) = \mathrm{const}\exp(-\frac{1}{2}u)$.

Quadric surfaces. We can now clearly see the two main qualitative features of the generalized hyperbolic distribution, which motivate its definition and place it in context.

(i) The density $f(x)$, being a smooth and decreasing function of $Q = (x - \mu)^T\Sigma^{-1}(x - \mu)$, is a unimodal r-variate density with mode μ, and is elliptically contoured (infinitely divisible, self-decomposable).

(ii) The tail decay is dominated by the term $\exp\{-\gamma\sqrt{\delta^2+Q}\}$. Normalizing this term to obtain a density, the corresponding *log-density*

$$y = \mathrm{const} - \gamma\sqrt{\delta^2+Q}$$
$$= \mathrm{const} - \gamma\sqrt{\delta^2 + (x-\mu)^T\Sigma^{-1}(x-\mu)}$$

gives the lower sheet of a *hyperboloid of two sheets*, a quadric surface in $\mathbb{R}^{r+1}$. This explains the name, and the interplay between the 'elliptic' and 'hyperbolic' aspects. For background on quadrics, see e.g. Kendall (1961, sections 34–38), Brannan *et al* (1999, section 1.4).

In the univariate case $r = 1$, the most symmetrical case gives

$$f(x) = \mathrm{const}\exp\left\{-\xi\sqrt{1+(x/\delta)^2}\right\},$$

called the *hyperbolic* distribution $hyp_{\xi,\delta}$ in Bingham and Kiesel (2001b).

Asymptotic cone. As $x \to \infty$, $y \sim -\gamma\sqrt{Q}$. Thus y is asymptotic to the cone

$$y = -\gamma\sqrt{(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

(a right circular cone, if we make the affine transformation to standardized variables). This is the analogue in $(r+1)$-space of the two linear asymptotes that the log-density of the univariate hyperbolic distribution approaches.

NIG distribution. This is a relative of the univariate hyperbolic distribution, introduced by Barndorff-Nielsen (1998,

1997) and studied by Rydberg (1997). It is defined as the distribution of one coordinate of a (drifting) bivariate Brownian motion when the other coordinate first hits a linear barrier. It seems that the NIG and hyperbolic models are broadly similar, with NIG providing if anything a better empirical fit to data. Since our main emphasis is not specific to any parametric model, we refer to the sources above for further detail.

7. Conclusions

The problem addressed here is the modelling of stock-price and asset-return distributions in higher dimensions, motivated by questions of portfolio selection and risk management in finance. We propose a semi-parametric model, which uses elliptically contoured distributions, specifically normal variance mixtures with self-decomposable mixing distributions, with particular emphasis on the density generator g. For applications, we have to estimate two components, the parametric part (μ, Σ) and the non-parametric part g. For (μ, Σ) standard methods suffice; to estimate the density generator g we rely on non-parametric function estimation, see e.g. Härdle (1990). It turns out that the tail decay of g is an indicator of the likelihood of extreme events and thus the whole density generator g is informative about risk and risk-management questions assuming that risk is measured by the portfolio VaR (which is coherent in the sense of Artzner *et al* (1999) for elliptically distributed risks). An application of the proposed technique, together with a detailed analysis of the econometric properties of the model, is contained in a forthcoming paper Bingham *et al* (2002).

References

Artzner P, Delbaen F, Eber J-M and Heath D 1999 Coherent measures of risk *Math. Finance* **9** 203–28

Bagnold R A 1941 *The Physics of Blown Sand and Desert Dunes* (London: Matthew)

Barndorff-Nielsen O E 1978 Hyperbolic distributions and distributions on hyperbolae *Scandinavian J. Statistics* **5** 151–7

Barndorff-Nielsen O E 1997 Normal inverse Gaussian distributions and stochastic volatility modelling *Scandinavian J. Statistics* **24** 1–13

Barndorff-Nielsen O E 1998 Processes of normal inverse Gaussian type *Finance Stochastics* **2** 41–68

Barndorff-Nielsen O E, Blaesild P and Halgreen C 1978 First-hitting time models for the generalized inverse Gaussian distribution *Stochastic Process Applications* **7** 49–54

Barndorff-Nielsen O E and Halgreen O 1977 Infinite diversibility of the hyperbolic and generalized inverse Gaussian distributions *Z. Wahrschein.* **38** 309–12

Barndorff-Nielsen O E, J L Jensen and Sørensen M 1998 Some stationary processes in discrete and continuous time *Adv. Appl. Probab.* **30** 989–1007

Barndorff-Nielsen O E, Kent J T and Sørensen M 1982 Normal variance-mean mixtures and z-distributions *Int. Stat. Rev.* **50** 145–59

Barndorff-Nielsen O E and Shephard N 2001a Modelling by Lévy processes for financial econometrics ed O E Barndorff-Nielsen, T Mikosch and S Resnick *Lévy Processes: Theory and Applications* (Boston, MA: Birkhäuser)

Barndorff-Nielsen O E and Shephard N 2001b Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics *J. R. Stat. Soc.* B **63** 167–241

Beran R 1979 Testing for ellipsoidal symmetry of a multivariate density *Ann. Stat.* **7** 150–62

Bertoin J 1996 *Lévy Processes, Cambridge Tracts in Mathematics* vol 21 (Cambridge: Cambridge University Press)

Bickel P J, Klaasen C A J, Ritov Y and Wellner J A 1998 *Efficient and Adaptive Estimation for Semi-Parametric Models* (New York: Springer) (1st edn 1993 (Baltimore, MD: Johns Hopkins University Press))

Bingham N H 1971 Factorization theory and domains of attraction for generalized convolution algebras *Proc. Lond. Math. Soc.* **23** 16–30

Bingham N H 1973 Positive definite functions on spheres *Proc. Camb. Phil. Soc.* **73** 145–56

Bingham N H 1984 On a theorem of Kłosowska about generalized convolutions *Colloq. Math.* **48** 117–25

Bingham N H and Kiesel R 2001a Hyperbolic and semiparametric models in finance *Disordered and Complex Systems* ed P Sollich, A C C Coolen, L P Hughston and R F Streater (New York: American Institute of Physics)

Bingham N H and R Kiesel 2001b Modelling asset return with hyperbolic distributions *Asset Return Distributions* ed J Knight and S Satchell (Butterworth-Heinemann)

Bingham N H, Kiesel R and Schmidt R 2002 *Semi-Parametric Models in Finance: Econometric Applications* (working paper)

Bochner S and Chandrasekharan K 1949 Fourier transforms *Annals of Mathematical Studies* vol 19 (Princeton, NJ: Princeton University Press)

Bodie Z, Kane A and Marcus A J 1999 *Investments* 4th edn (New York: McGraw-Hill)

Bondesson L 1981 Contribution to the discussion of the paper by D R Cox: statistical analysis of time series: some recent developments *Scandinavian J. Statistics* **8** 93–115

Bondesson L 1982 On simulation from infinitely divisible distributions *Adv. Appl. Probab.* **14** 855–69

Brannan D A, Esplen M F and Gray J J 1999 *Geometry* (Cambridge: Cambridge University Press)

Cambanis S, Huang S and Simons G S 1981 On the theory of elliptically contoured distributions *J. Multivariate Anal.* **11** 368–85

Carr P, Chang E and Madan D B 1998 The variance-gamma process and option pricing *Euro. Finance Rev.* **2** 79–105

Carr P, Geman H, Madan D B and Yor M 2000 *The Fine Structure of Asset Returns: An Empirical Investigation* University of Maryland (working paper)

Eberlein E 2001 Applications of generalized hyperbolic Lévy motions to finance *Lévy Processes: Theory and Applications* ed O E Barndorff-Nielsen, T Mikosch and S Resnick (Boston, MA: Birkhäuser)

Eberlein E and U Keller 1995 Hyperbolic distributions in finance *Bernoulli* **1** 281–99

Eberlein E, Keller U and Prause K 1998 New insights into smile, mispricing and value-at-risk: the hyperbolic model *J. Business* **71** 371–406

Embrechts P, McNeil A and Straumann D 2001 Correlation and dependency in risk management: properties and pitfalls *Risk Management: Value at Risk Beyond* ed M Dempster and H K Moffat (Cambridge: Cambridge University Press)

Fang K-T, Kotz S and Ng K-W 1990 *Symmetric Multivariate and Related Distributions* (London: Chapman and Hall)

Feller W 1971 *An Introduction to Probability Theory and its Applications* vol 2 2nd edn (Chichester: Wiley)

Halgreen C 1979 Self–decomposability of the generalized inverse Gaussian and hyperbolic distribution functions *Z. Wahrschein.* **47** 13–18

Hampel F R, Rousseeuw P J, Ronchetti E M and Stahel W A 1986 *Robust Statistics. The Approach Based on Influence Functions* (Chichester: Wiley)

Härdle W 1990 *Smoothing Techniques: With Implementation in S* (Berlin: Springer)

Q. 2

Hodgson D J, Linton O and Vorkink K 2000 Testing the capital asset pricing model efficiently under elliptical symmetry: a semi-parametric approach *Preprint* LSE

Jørgensen B 1982 *Statistical Properties of the Generalized Inverse Gaussian Distribution Function (Springer Lecture Notes in Statistics vol 9)* (Berlin: Springer)

Kelker D 1971 Infinite divisibility and variance mixtures of the normal distribution *Annals Math. Stat.* **42** 802–8

Kendall M G 1961 *A Course in the Geometry of n Dimensions (Griffin's Stat. Monogr. Courses vol 8)* (London: Charles Griffin)

Kingman J F C 1963, Random walks with spherical symmetry *Acta Math.* **109** 11–53

Korsholm L 2000 The semi-parametric normal mean-variance mixture model *Scandinavian J. Statistics* **27** 227–61

Li R, Fang K-T and Zhu L 1997 Some Q–Q probability plots to test spherical and elliptical symmetry *J. Comp. Graph. Stat.* **6** 435–50

Lopuhaä H P 1999 Asymptotics of reweighted estimators of multivariate location and scatter *Ann. Statistics* **27** 1638–65

Lopuhaä H P and Rousseeuw P J 1991 Breakdown properties of affine equivariant estimators of multivariate location and scatter *Ann. Statistics* **19** 229–48

Madan D B and Seneta E 1990 The variance-gamma (VG) model for share market returns *J. Business* 511–24

Mandelbrot B 1963 The variation of certain speculative prices, *J. Business* **36** 394–419

Mandelbrot B B 1997 *Fractals and Scaling in Finance. Discontinuity, Concentration, Risk* (Berlin: Springer) Selected works of Benoit B Mandelbrot

Protter P 1992 *Stochastic Integration and Differential Equations* 2nd edn (New York: Springer)

Rachev S and Mittnik S 2000 *Stable Paretian Models in Finance* (Chichester: Wiley)

Romanowski M 1979 *Random Errors in Observations and the Influence of Modulation on their Distributions* (Stuttgart: Verlag Konrad Wittwer)

Rydberg T H 1997 The normal inverse Gaussian Lévy process: simulation and approximation *Research Report* Department of Theoretical Statistics, Institute of Mathematics, University of Århus

Sato K-I 1999 *Lévy Processes and Infinite Divisibility (Cambridge Studies in Advanced Mathematics vol 68)* (Cambridge: Cambridge University Press)

Shiryaev A N 1999 *Essentials of Stochastic Finance Advanced Series of Statistical Science & Applied Probability* vol 3 (Singapore: World Scientific)

Urbanik K 1964 Generalized convolutions *Stud. Math.* **23** 217–45

Watson G N) 1944 *A Treatise on the Theory of Bessel Functions* 2nd edn (Cambridge: Cambridge University Press) (1st edn 1922)

Q. 3

Queries for IOP paper 129222

Journal: **Quant**
Author: **N H Bingham and Rüdiger Kiesel**
Short title:

Page 7

Query 1:
    Author: This sentence seems incomplete. Please check.

Page 9

Query 2:
    Author: Please supply city

Page 10

Query 3:
    Author: Please provide volume number for the reference Madan and Seneta (1990).