MODELLING AND PREDICTION OF FINANCIAL TIME SERIES

N. H. BINGHAM

Abstract

We consider statistical aspects of the modelling and prediction theory of time series in one and many dimensions. We discuss Lévy-based and general models, and the stationary and non-stationary cases. Our starting point is the recent pair of surveys, *Szegö's theorem and its probabilistic descendants* and *Multivariate prediction and matrix Szegö theory*, by this author.

Keywords and phrases: Lévy process, semi-parametric model, elliptic distribution, stationary, ergodic diffusion, time series, prediction, Ornstein-Uhlenbeck process

1. Introduction.

Mathematical finance provides us with a rich source of interesting mathematical and statistical problems. The difficulty of predicting asset prices, combined with the importance of being able to do so, motivates the search for stochastic models for risky asset prices; see e.g. [BinKi2], or any other book in the area. And the area abounds with real data, and so one's models can be checked and compared.

The prototypical model for asset prices is the Black-Scholes model, dating back to 1973; for a text-book treatment in the multi-dimensional (or multi-asset) case, see e.g. [KarShr], Ch. 1. An investor holds some cash risklessly, at interest rate r say (see below), and some risky assets, $S_1(t), \dots S_d(t)$ say. The (Black-Scholes) stochastic differential equation (SDE) driving the asset prices contains a vector $\mu = (\mu_1, \dots, \mu_d)$, where μ_i is the *return rate* on the ith asset, and a covariance matrix Σ, giving the covariances between the assets. By Markowitzian diversification ([Mar1], [Mar2]), we should

(a) think of risk and return together, not separately – that is, we should think of μ and Σ together;

(b) hold a *balanced portfolio* of a number of assets, 'balanced' meaning that there is lots of *negative correlation*.

Thus any proper treatment contains (μ, Σ), which we regard as the *parametric component* of our model. The number d of assets could be quite large (e.g., we might diversify between the dozen or so principal sectors in the economy, and within sectors, diversify further between several leading firms). The negative correlation means that, when the economic climate changes (as it will), losses on some assets will tend to be compensated by gains on others ('what one loses on the swings, one gains on the roundabouts').

There remains the main question: how to model the 'driving noise' – the randomness driving the risky asset prices. The first choice concerns the dimensionality of the driving noise process. The world is exceedingly complex (effectively infinite-dimensional); our data is d-dimensional – but, some of the principal features of the ambient financial and business environment are reflected in the 'business cycle', which can plausibly be taken one-dimensional. With a one-dimensional noise process as our 'risk driver', we can
(a) use a diffusion-based model, leading to a complete market (see e.g. [BinFrKi] §4);
(b) use a model based on Lévy processes; see e.g. [BinFrKi] §3, [ConTan]. This allows for jumps, present in real financial data, but leads to an incomplete market model.
In either case, one has to model the driving noise process; this involves the Lévy measure in the Lévy case, and the stationary distribution in the diffusion case. This is the *non-parametric* component of our model.

Large financial data sets, such as multivariate time series as here, typically need preliminary treatment – filtering ('cleaning'), de-seasonalization, etc. We suppose this done here, as in e.g. [BreyDE].

2. Semi-parametric models

We thus seek a *semi-parametric* model, combining a parametric component (μ, Σ) à la Markowitz, modelling the portfolio, and a non-parametric component, modelling the noise or dynamics.

We assume for simplicity that the process $X = (X_t)$ ($t = 0, 1, 2, \ldots$: we work in discrete time, until further notice) belongs to the *elliptic family* (see e.g. [FanKN]). The distributions depend only on the quadratic form

$$Q(x) := (x - \mu)^T \Sigma^{-1} (x - \mu),$$

with d-vector argument x; we restrict attention to the full-rank case, when Σ is invertible. See e.g. [BinKi1], [BinKiS] for the distributional (or static) theory, [BinFrKi] for the dynamic theory – *multivariate elliptic processes*. Of course, this assumption of elliptical symmetry needs testing; this can be done, e.g. as in [ManPQ].

In any compound statistical model such as this, one encounters the question of the cost, when estimating one part of the model, of ignorance of the other part. This is often considerable: the prototype is the normal model $N(\mu, \sigma)$, where estimating μ with σ known leads to the normal distribution but estimating μ with σ unknown leads to the Student t-distribution. One might expect similar difficulties here, but these do not in fact occur. The reason for this is the connection of the elliptical model with the *affine group*. This theme is developed in detail in [BiKRW], §4.2, §6.3, for i.i.d. errors, Drost et al. [DroKW] for time-series errors.

The first task is to estimate (μ, Σ). As our model will be elliptic, and so the affine group will be relevant, it is sensible to choose our estimators to respect the action of the affine group, or in statistical language to be *affine equivariant* (cf. [GruRoc]). Such a location estimator $\hat{\mu}$ for μ is given by the *Oja median*; see e.g. [Oja], [OjaOR]. For a survey of multivariate medians, see [Sma], and for robust estimators of location, [Lop]. Affine equivariant estimators $\hat{\Sigma}$ of Σ are then given by [KosMO] and [LopRou]; see also [XiWu].

We point out that, even in the univariate case, estimation of means is not possible with the same precision as estimation of variances. This phenomenon is called *mean blur*; see e.g. Merton [Mer], Luenberger [Lue], §8.5, Kuhn et al. [KuhPRF].

3. Stationarity and non-stationarity.

We recall the martingale approach to the pricing of risky assets (see e.g. [BinKi2], Preface): one should
(i) *discount* everything: pass from nominal prices S_t to real prices $\tilde{S}_t := e^{-rt} S_t$, with r the riskless interest rate;
(ii) pass from the real probability measure P to the *equivalent martingale measure (EMM)*, or *risk-neutral measure*, P^*, and then take conditional expectations.
The point of the discounting in (i) is to induce stationarity (in the Black-Scholes model, one can solve the relevant SDE, obtaining an exponential martingale (cf. [RevYor] IV.3, VIII), from which stationarity follows). As is well known, particularly from the work of C. W. J. (Sir Clive) Granger (see e.g. [GraHat], [Bin2] §1), it is of prime practical importance to avoid use of statistical methods designed for stationary situations in non-stationary ones, as this can lead to misleading conclusions, via the phenomenon of *spurious regression.*

In practice, this leads to problems of implementation. As in all time-series methods (for which our basic reference is [BroDav]), one needs a long series of data to be able to predict with any accuracy (cf. [DahGir] here). On the other hand, the world in general, and the financial world in particular, is changeable, and interest rates change over time. Furthermore, one has a range of possible interest rates by which to discount. Two obvious candidates are the bank rate (in the UK, the Bank of England rate) and the Libor (London inter-bank offered rate). Libor in particular is widely used to price financial derivatives. During the liquidity crisis of 2007-8 (and subsequently), banks were reluctant to lend, so reluctant to quote (Libor is calculated as a weighted average of the rates quoted by a range of leading banks), and this caused problems in pricing [1].

Apart from discounting, we note briefly the other common way of inducing stationarity – differencing. One tests for stationarity; if the test is failed, one differences and tests again, etc. This is part of the standard Box-Jenkins approach, for which we refer to [BoxJeRe].

The basis of the prediction theory of stationary time-series is the *Kolmogorov Isomorphism Theorem (KIT)* ([Kol]; see e.g. [Bin2], §2). There is a random measure Y with orthogonal increments, the *Cramér process* or *spectral process* ([Cra], [CraLea] §7.5; see e.g. [Kal] for background on random measures) and a probability measure m on T, the *spectral measure*, together with an isomorphism

$$X_t \leftrightarrow e^{it\cdot} \qquad (KIT)$$

between the Hilbert spaces $\mathcal{H}$ (the L_2-space of the process $X = (X_t)$) and $L_2(m)$, which maps maps between the *time domain* on the left and the *frequency domain* on the right. One has the *Cramér representation*

$$X_t = \int_T e^{it\theta} dY(\theta) \qquad (CR)$$

(the integral is round the unit circle T),

$$E[(dY(\theta))^2] = dm(\theta),$$

and (taking $E[X_n] = 0$, $var(X_n) = 1$ for simplicity, so covariance becomes correlation) the correlation is given by the *spectral representation*

$$E[X_n \bar{X}_0] = \int_T e^{-in\theta} dm(\theta). \qquad (SR)$$

[1] Worse: it has recently emerged that Libor quotes were manipulated at times, leading to the current so-called "Lie-bor scandal".

An alternative to discounting (or differencing) is to use the non-stationary form of KIT, for which see e.g. [Kak]. Here one has, instead of a random measure $Y(.)$, a random *bimeasure* $Y(.,.)$. As with the stationary case, this rests on the Karhunen-Loève expansion (and so on Mercer's theorem, on the eigenexpansion of the positive definite covariance function, and Stone's theorem, on spectral representation of groups of unitary operators on Hilbert space); see [Karh], [Loe] II §37 for background and details.

For a further approach, in terms of *locally stationary processes*, see Dahlhaus [Dah1], [Dah2].

4. Discrete v. continuous time.

We restrict here to the stationary case, for simplicity. Typically, one has data from a (here, financial) time series, which are discrete. We take the time points equally spaced, say integer – these may be daily closing prices, daily or monthly returns, high-frequency 'tick' data, etc. (tick data show special features, because of market microstructure – see e.g. [BinS]). By Pontryagin (or Fourier) duality, the discreteness of time corresponds to the compactness of the unit circle in the spectral representation above. However, the underlying models for stock prices are typically SDEs, so in continuous time; also, there is a well-developed theory of continuous-time econometrics; see e.g. Bergstrom [Berg]. It is thus useful to be able to pass at will between discrete and continuous time (cf. [JaPr]). We can do this, simply by passing from integer time t to continuous time t in (SR) (indeed, we have anticipated this in our notation, by not using n for integer time).

Mathematically, discrete time and integrals on the unit circle involve Hardy spaces on the unit disc D, while continuous time involves spectral representations on the real line, and Hardy spaces on the upper half-plane; see e.g. [Bin2], §8.11 for background and references. Discrete time involves orthogonal polynomials on the unit circle (OPUC), continuous time involves orthogonal polynomials on the real line (OPRL); see Simon [Sim1] for OPUC, [Sim2] for OPUC and OPRL together. Levinson and McKean [LevMcK] and Dym and McKean [DymMcK] deal with continuous time; for discrete-time analogues of some of their results, see [KasBi].

By the Cramér representation (CR), we may pass at will between X_t at integer times, Y, and X_t with time continuous – that is, we may recover the whole process $X = (X_t)$ from the values of X *sampled* at integer times $t = n$. This allows us to *interpolate* from integer to non-integer times. This interpolation is in fact extremely regular – indeed, it is as regular as one could imagine. What results here is a *random entire function*, indeed one of exponential type π (see e.g. Boas [Boa] 2.1, 6.8), by the Paley-Wiener theorem [PaWi].

The length 2π of the interval $[0, 2\pi]$ (identified with the unit circle T under $\theta \leftrightarrow e^{i\theta}$) – or rather, its ratio to the time-interval 1 between the integer times n – is crucial here. The *sampling theorem* states that, under suitable conditions, one may recover the full function from values sampled at a certain minimum rate, the *Nyquist rate*, but not at less frequent values. See e.g. Partington [Par], §7,2, where the result is formulated in terms of Paley-Wiener spaces. For K a compact subset of the real line, write $PW(K)$ for the space of functions f whose Fourier transform $\hat{f}$ vanishes off K and is in L_2; for $b > 0$ write $PW(b)$ for $PW([-b, b])$. The functions in $PW(b)$ are entire of exponential type b, and belong to L_2. By the sampling theorem, f can be recovered by sampling at equal intervals π/b:

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\pi/b) \frac{\sin b(t - n\pi/b)}{b(t - n\pi/b)},$$

the series converging uniformly and in L_2 (the sampling rate π/b is called the *Nyquist rate*). Thus we have $b = \pi$ and Nyquist rate 1. The sampling theorem can be traced back to Cauchy in 1841, E. T. Whittaker in 1915, V. A. Kotelnikov in 1933, J. M. Whittaker [Whi] in 1935, G. H. Hardy in 1941 and C. E. Shannon in 1949; see also [Hig1] §2.6.3, [Hig2], [Hig3]. For stochastic processes, as here, our main source is S. P. Lloyd in 1959 [Llo]; cf. Quinn and Hannan [QuHa], (2.23). For background, references and a multiplicity of viewpoints and applications, see e.g. Benedetto and Ferreira [BenFe]. We note in particular that sampling at irregularly spaced time-points is possible.

5. Stochastic volatility

It is the volatility σ, or Σ, in the Black-Scholes model that is most problematic, and has to be estimated; as is well known, this can be done as *historic volatility* by time-series methods, or *implied volatility* by inferring it from option prices via the Black-Scholes formula (cf. e.g. [BinKi2] §7.3). Volatility in fact varies, e.g. with strike price (leading to the so-called volatility smile, or smirk – see e.g. [Gat]). One also observes *volatility clustering*: high volatility is associated with periods of economic stress, uncertainty or turmoil, and these are interspersed with more normal periods with lower volatility (see Aldous [Ald] for a general theoretical view on such clumping phenomena). This has led to the extensive use of *stochastic volatility (SV)* models. In discrete time, prominent here are the *ARCH* (autoregressive conditional heteroscedasticity) and *GARCH* (generalised ARCH) models; see e.g. [Gou].

As above, one can proceed here in discrete or continuous time. The continuous-time acronyms for ARMA and GARCH are CARMA and COGARCH; for CARMA see e.g. [Bro], [BroFeKl], [BroDavYa2], and for COGARCH [KluMaSz], [KluLiMa]; cf. [DroW].

6. Diffusions

Models based on diffusions may be preferred, on mathematical grounds, because they are continuous, or because they lead to complete market models. Ergodic diffusions may be used as risk drivers in the multivariate elliptic processes of §2 ([BinFrKi] §4). See e.g. Forman and Sorensen [ForSor] and the references there for details, Kutoyants [Kut] for background on estimation for ergodic diffusions.

7. Lévy models

Lévy processes have been widely used for modelling of stock prices. They allow jumps (at the cost of leading to incomplete market models); the case for doing so has been convincingly argued in, e.g., [ConTan] Ch. 1. They allow for infinite activity – infinitely many jumps in finite time, resulting from an infinite Lévy measure; this is very useful for modelling 'jitter' – the very large number of very small jumps observed when the price process of a heavily traded stock in normal market conditions is observed in detail. And of course they allow a Brownian component, which may be convenient even when not necessary. For a detailed recent study of univariate, high-frequency data, see [AitSJ1] – [AitSJ4].

Ornstein-Uhlenbeck SDEs with driving Lévy noise, which lead to self-decomposable (SD) stationary distributions, have been widely used in this context; see e.g. [KluLiMa], [BroDavYa1]. In the multivariate elliptic processes of [BinFrKi] §3, one has

$$X_t - \mu = R_t A^T U_t = R_t \Sigma^{\frac{1}{2}} U_t, \tag{MEP}$$

where Σ has Cholesky decomposition $\Sigma = A^T A$ (so $A = \Sigma^{\frac{1}{2}}$, the usual matrix square root of the positive definite matrix Σ), $U = (U_t)$ is Brownian motion on the d-dimensional sphere, and $R = (R_t)$ is the *risk driver*, the solution of the SDE

$$dR_t = -cR_t dt + dZ_t, \tag{OU}$$

of Ornstein-Uhlenbeck (OU) type, with driving noise $Z = (Z_t)$ a subordinator (positive Lévy process), $c > 0$ ($c = 1$ if convenient). Subject to the log-integrability condition $\int \log_+(|x|)d\nu(x) < \infty$, where ν is the Lévy measure of Z, (OU) has a unique strong solution $R = (R_t)$, which is positive, stationary and has a limit distribution R_∞ as $t \to \infty$; this is self-decomposable (SD). Conversely, every SD law is the limit law of such a process of OU type. See Sato [Sat] §15-17 and §33 for background, and [BinFrKi] §3; see also [Bin1] for the useful property of being of *type G* (G for Gaussian). One has

$$R_t^2 = (X_t - \mu)^T \Sigma^{-1} (X_t - \mu), \tag{MEP'}$$

which enables one to estimate the density of R_∞ by standard density-estimation methods (see e.g. [Sil], [VenR] §5.6). One can pass between the Lévy measures of Z_1 and R_∞ by [Sat] Th. 17.5. Consistent statistical inference is possible in this setting: see [JonMV] (non-parametric), [JonM] (parametric).

From (MEP), one has ([BinFrKi], §3)

$$var(X_t | R_t) = R_t^2 \Sigma, \qquad var(X_t) = E[R_t^2]\Sigma.$$

This gives *volatility clustering*, one of the stylized facts of mathematical finance, subject to weak conditions on the covariance of R (e.g., being able to approximate by the first few terms of the eigenexpansion). On the other hand, using a one-dimensional risk-driver, though mathematically convenient, is clearly an over-simplification.

For Sato processes – additive processes, with independent but not necessarily stationary increments, see [Sat] Ch. 2, and for financial applications, [EbeMad], [MadYor].

8. Prediction

For prediction theory in continuous time, we refer to Dym and McKean [DymMcK].

For prediction theory for ARMA processes, we refer to [BroDav], Ch. 5.

What allows the reduction to a one-dimensional risk driver, as in (MEP'), is the spherical symmetry of the Brownian motion U in (MEP). Without this, one needs a d-dimensional treatment (below).

In one dimension, the recent survey [Bin2] gives an account, with references, of finite and of infinite predictors (where one predicts the future given a finite section of the past or of the whole past), and of the convergence of finite-predictor coefficients to infinite-predictor coefficients ([Bin2], §§4,5).

The methods of [Bin3] allow a direct multidimensional approach along the same lines.

We note that the sequence of Verblunsky coefficients $\alpha = (\alpha_n)_{n=1}^{\infty}$ (for which see also [BinInKa]) allows an attractive unrestricted parametrization ($\alpha_n \in D$ in the scalar case, with D the unit disc, $\|\alpha_n\| < 1$ in the matrix case) as an alternative to the ARMA model.

Verblunsky's theorem (see e.g. [Bin2] §3, [Bin3] §3) gives a bijection between the set of sequences α with terms in D (or sequences of matrices of norm < 1) and the set m of probability measures on T. One thus has a choice between estimating m, or at least its density w, by density-estimation and spectral methods, and estimating $\alpha = (\alpha_n)$, the partial autocorrelation function (PACF). The α_n are the diagonal elements in the triangular matrix of finite-predictor coefficients, via the Levinson-Durbin algorithm (see e.g. [Bin2], §3). These can be estimated, as in [BroDav] §§3.4, 8.2 (sample PACF; cf. [Deg1], [Deg2], [McLZ]).

It would be interesting to compare and contrast the performance of the methods reviewed here on a variety of real data, particularly financial data and particularly in many dimensions. But time and space do not allow us to pursue this programme here.

Acknowledgement. This is an extended version of my talk at the Conference on Markov and Semi-Markov Processes at the University of Thessaloniki in September 2011. I thank the organisers for their invitation and hospitality. I also thank Pierre Blacque-Florentin for many interesting discussions.

References

[AitSJ1] Y. Ait-Sahalia and J. Jacod, Testing for jumps in a discretely observed process. *Ann. Stat.* **37** (2009), 184-222.
[AitSJ2] Y. Ait-Sahalia and J. Jacod, Estimating the degree of activity of jumps in high-frequency data. *Ann. Stat.* **37** (2009), 2204-2244.
[AitSJ3] Y. Ait-Sahalia and J. Jacod, Is Brownian motion necessary to model high-frequency data? *Ann. Stat.* **38** (2010), 3093-3128.
[AitSJ4] Y. Ait-Sahalia and J. Jacod, Testing whether jumps have finite or infinite activity. *Ann. Stat.* **39** (2011), 1689-1719.

[Ald] D. Aldous, *Probability approximations via the Poisson clumping heuristic*. Springer, 1989.
[BenFe] J. J. Benedetto and P. J. S. G. Ferreira (ed.), *Modern sampling theory*. Birkhäuser, 2001.
[Berg] A. R. Bergstrom, *Continuous-time econometric modelling*. Oxford University Press, 1990.
[BicKRW] P. J. Bickel, C. A. J. Klaassen, Y. Ritov and J. A. Wellner, *Efficient and adaptive estimation for semi-parametric models*. Springer, 1998.
[Bin1] Lévy processes and self-decomposability in finance. *Probability and Mathematical Statistics* **26** (2006), 367-378.
[Bin2] Szegö's theorem and its probabilistic descendants. *Probability Surveys* **9** (2012), 287-324.
[Bin3] Multivariate prediction and matrix Szegö theory. *Probability Surveys* **9** (2012), 325-339.
[BinFrKi] N. H. Bingham, J. M. Fry and R. Kiesel, Multivariate elliptic processes. *Statistica Neerlandica* **64** (2010), 352-366.
[BinInKa] N. H. Bingham, A. Inoue and Y. Kasahara, An explicit representation of Verblunsky coefficients. *Stat. Prob. Letters* **82** (2012), 403-410.
[BinKi1] N. H. Bingham and R. Kiesel, Semi-parametric modelling in finance: theoretical foundations. *Quantitative Finance* **2** (2002), 241-250.
[BinKi2] N. H. Bingham and Rüdiger Kiesel), *Risk-neutral Valuation: Pricing and Hedging of Financial Derivatives*, 2nd ed., 2004 (1st ed. 1998).
[BinKiS] N. H. Bingham, R. Kiesel and R. Schmidt, A semi-parametric approach to risk management. *Quantitative Finance* **3** (2003), 426-441.
[BinS] N. H. Bingham and R. Schmidt, Distributional and temporal dependence structure of high-frequency financial data: A copula approach. Pages 69-92 in [KabLiSt].
[Boa] R. P. Boas, *Entire functions*, Academic Press, 1954.
[BoxJeRe] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time-series analysis. Forecasting and control* (4th ed.). Wiley, 2008.
[Bra] L. de Branges, *Hilbert spaces of entire functions*. Prentice-Hall, 1968.
[BreyDE] W. Breymann, A. Dias and P. Embrechts, Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance* **3** (2003), 1-16.
[Bro] P. J. Brockwell, Continuous-time ARMA processes. Pages 249-276 in [ShaR].
[BroDav] P. J. Brockwell and R. A. Davis, *Time series: Theory and methods* (2nd ed.), Springer, New York, 1991 (1st ed. 1987).

[BroDavYa1] P. J. Brockwell, R. A. Davis and Y. Yang, Estimation for non-negative Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Prob.* **44** (2007), 977-989.
[BroDavYa2] P. J. Brockwell, R. A. Davis and Y. Yang, Estimation for non-negative Lévy-driven CARMA processes. *J. Bus. Econ. Stat.* **29** (2011), 250-259.
[BroFeKl] P. J. Brockwell, V. Ferrazzano and C. Klüppelberg, High-frequency sampling of continuous-time ARMA processes. *J. Time Series Analysis* **33** (2012), 152-160.
[ConTan] R. Cont and P. Tankov, *Financial modelling with jump processes*. Chapman and Hall, 2004.
[Cra] H. Cramér, On harmonic analysis in certain function spaces. *Ark. Mat. Astr. Fys.* **28B** (1942), 1-7, reprinted in *Collected Works of Harald Cramér* Volume II, 941-947, Springer, 1994.
[CraLea] H. Cramér and R. Leadbetter, *Stationary and related stochastic processes*. Wiley, 1967.
[Dah1] R. Dahlhaus, Fitting time-series models to non-stationary processes. *Ann. Stat.* **25** (1977), 1-25.
[Dah2] R. Dahlhaus, A likelihood approximation for locally stationary processes. *Ann. Stat.* **28** (2000), 1762-1794.
[DahGir] R. Dahlhaus and L. Giraitis, On the optimal segment length for parameter estimates for locally stationary time series. *J. Time Series Anal.* **19** (1998), 629-656.
[Deg1] S. Dégerine, Sample partial autocorrelation function. *IEEE Trans. Signal Processing* **41** (1993), 403-407.
[Deg2] S. Dégerine, Sample partial autocorrelation function of a multivariate time series. *J. Multivariate Analysis* **50** (1994), 294-313.
[DroKW] F. C. Drost, C. A. J. Klaassen and B. J. M. Werker, Adaptive estimation in time-series models. *Ann. Statist.* **25** (1977), 786-817.
[DroW] F. C. Drost and B. J. M. Werker, Closing the GARCH gap: continuous-time GARCH modelling. *J. Econometrics 74* (1996), 31-57.
[DutFGR] R. Dutter, P. Filzmoser, U. Gather and P. J. Rousseeuw (ed.), *Developments in robust statistics*. Springer, 2002.
[DymMcK] H. Dym and H. P. McKean, *Gaussian processes, function theory and the inverse spectral problem*. Academic Press, 1976.
[EbeMad] E. Eberlein and D. B. Madan, Sato processes and the valuation of structured products. *Quantitative Finance* **9** (2009), 27-42.
[FanKN] K.-T. Fang, S. Kotz and K.-W. Ng, *Symmetric multivariate and*

related topics. Chapman and Hall, 1990.
[ForSor] J. L. Forman, M. Sorensen, The Pearson diffusions: a class of statistically tractable diffusion processes. *Scand. J. Stat.* **35** (2008), 438-465.
[Gat] J. Gatheral, *The volatility surface: A practitioner's guide*. Wiley, 2006.
[Gou] C. Gouriéroux, *ARCH models and financial applications*. Springer, 1997.
[GraHat] C. W. J. Granger and M. Hatanaka, *Spectral analysis of economic time series*, Princeton University Press, 1964.
[GruRoc] R. Grübel and D. M. Rocke, On the cumulants of affine equivariant estimators in elliptical families. *J. Multiv. Anal.* **35** (1990), 203-222.
[Hig1] J. R. Higgins, *Completeness and basis properties of sets of special functions*. Cambridge Tracts in Math. **72**, Cambridge University Press, 1977.
[Hig2] J. R. Higgins, Five short stories about the cardinal series. *Bull. Amer. Math. Soc.* **12** (1985), 45-89.
[Hig3] J. R. Higgins, *Sampling theory in Fourier analysis and signal analysis*. Oxford University Press, 1996.
[JaPr] J. Jacod and P. Protter, *Discretization of processes*. Springer, 2012.
[JonM] G. Jongbloed and F. H. van der Meulen, Parametric estimation for subordinators and induced OU processes. *Scand. J. Stat.* **33** (2006), 825-847.
[JonMV] G. Jongbloed, F. H. van der Meulen and A. W. van der Vaart, Non-parametric inference for Lévy-driven OU processes. *Bernoulli* **11** (2005), 759-791.
[KabLiSt] Yu. Kabanov, R. Liptser and J. Stoyanov (ed.), *The Shiryaev Festschrift: From stochastic calculus to mathematical finance*. Springer, 2006.
[Kak] Y. Kakihara, The Kolmogorov isomorphism theorem and extensions to some non-stationary processes. Pages 443-470 in [ShaR].
[Kal] O. Kallenberg, *Random measures*, 4th ed. Academic Press, 1986.
[KarShr] I. Karatzas and S. E. Shreve, *Methods of mathematical finance*. Springer, 1998.
[Karh] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fenn.* **37** (1947), 1-79.
[KasBi] Y. Kasahara and N. H. Bingham, Verblunsky coefficients and Nehari sequences. *Trans. Amer. Math. Soc.*, in press.
[KluLiMa] C. Klüppelberg, A. Lindner and R. Maller, Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. Pages 393-419 in [KabLiSt].

[KluMaSz] C. Klüppelberg, R. A. Maller and A. Szimayer, The COGARCH: a review, with news on option pricing and statistical inference. *Surveys in stochastic processes* 29-58, EMS Ser. Congr. Rep., Eur. Math. Soc., Zürich, 2011.

[Kol] A. N. Kolmogorov, Stationary sequences in Hilbert space. *Bull. Moskov. Gos. Univ. Mat.* **2** (1941), 1-40 (in Russian; reprinted, *Selected works of A. N. Kolmogorov, Vol. 2: Theory of probability and mathematical statistics*, Nauka, Moskva, 1986, 215-255).

[KosMO] G. Koshevoy, J. Möttönen and H. Oja, Scatter matrix estimation based on the zonotope. *Ann. Statist.* **31** (2003), 1439-1459.

[KuhPRF] D. Kuhn, P. Parpas, B. Rustem and R. Fonseca, Dynamic mean-variance portfolio analysis under model risk. *J. Computational Finance* **12** (2009), 91-115.

[Kut] Yu. A. Kutoyants, *Statistical inference for ergodic diffusion processes.* Springer, 2004.

[LevMcK] N. Levinson and H. P. McKean, Weighted trigonometrical approximation on R^1 with application to the germ field of a stationary Gaussian noise. *Acta Math.* **112** (1964), 99-143.

[Llo] S. P. Lloyd, A sampling theorem for stationary (wide sense) stochastic processes. *Trans. Amer. Math. Soc.* **92** (1959), 1-12.

[Loe] M. Loève, *Probability theory*, Volumes I, II, 4th ed., Springer, 1978.

[Lop] H. P. Lopuhaä, Highly efficient estimators of multivariate location with high breakdown point. *Ann. Statist.* **20** (1992), 398-413.

[LopRou] H. P. Lopuhaä and P. J. Rousseeuw, Breakdown points of affine equivariant estimation of multivariate location and covariance matrices. *Ann. Statist.* **19** (1991), 229-248.

[Lue] D. G. Luenberger, *Investment science.* Oxford University Press, 1998.

[MadYor] D. B. Madan and M. Yor, The S & P 500 index as a Sato process: travelling at the speed of the VIX. *Applied Math. Finance* **18** (2011), 227-244.

[ManPQ] A. Manzotti, F. J. Perez and A. J. Quiroz, A statistic for testing the null hypothesis of elliptical symmetry. *J. Multivariate Analysis* **81** (2002), 274-285.

[Mar1] H. Markowitz, Portfolio selection. *J. Finance* **7** (1952), 77-91.

[Mar2] H. Markowitz, *Portfolio selection: Efficient diversification of investments.* Wiley, 1959.

[McLZ] A. I. McLeod and Y. Zhang, Partial autocorrelation parametrization for subset autoregression. *J. Time Series Anal.* **27** (2006), 599-612.

[Mer] R. C. Merton, On estimating the expected return on the market: an exploratory investigation. *J. Financial Economics* **8** (1980), 323-361.
[Oja] H. Oja, Affine invariant multivariate sign and rank tests and corresponding estimates. A review. *Scand. J. Statist.* **26** (1999), 319-343.
[OjaOR] H. Oja, P. Orponen and T. Ronkainan, Computation of the multivariate Oja median. Pages 344-359 in [DutFGR].
[Par] J. R. Partington, *Interpolation, identification and sampling.* Oxford University Press, 1997.
[PalW] R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain.* AMS Colloq. Publ. XIX, Amer. Math. Soc., 1934.
[QuHa] B. G. Quinn and E. J. Hannan, *The estimation and tracking of frequency.* Cambridge University Press, 2001.
[RevYor] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, 3rd ed. Grundl. math. Wiss **293**, Springer, 1999.
[Sat] K.-I. Sato, *Lévy processes and infinitely divisible distributions.* Cambridge University Press, 1999.
[ShaR] D. N. Shanbhag and C. R. Rao (ed.), *Stochastic processes: Theory and methods. Handbook of Statistics* **19**, North-Holland, 2001.
[Sil] B. W. Silverman, *Density estimation for statistics and data analysis.* Chapman and Hall, 1986.
[Sim1] B. Simon, *Orthogonal polynomials on the unit circle. Part 1: Classical theory. Part 2: Spectral theory.* AMS Colloquium Publications 54.1, 54.2, American Math. Soc., Providence RI, 2005.
[Sim2] B. Simon, *Szegö's theorem and its descendants: Spectral theory for L^2 perturbations of orthogonal polynomials.* Princeton University Press, 2011.
[Sma] C. G. Small, A survey of multidimensional medians. *Internat. Stat. Review* **58** (1990), 263-277.
[VenR] W. N. Venables and B. D. Ripley, *Modern applied statistics with S*, 4th ed. Springer, 2002.
[Whi] J. M. Whittaker, *Interpolatory function theory*, Cambridge Tracts in Math. **33**, Cambridge University Press, 1935.
[XiWu] H. Xiao and W.-B. Wu, Covariance matrix estimation for stationary time series. *Ann. Statist.* **40** (2012), 466-493.

Mathematics Department, Imperial College, London SW7 2AZ, UK
n.bingham@ic.ac.uk   nick.bingham@btinternet.com