Econometric Theory, **20**, 2004, 927–942. Printed in the United States of America.
DOI: 10.1017/S0266466604205060

ESTIMATING THE SKEWNESS IN DISCRETELY OBSERVED LÉVY PROCESSES

JEANNETTE H.C. WOERNER
University of Göttingen

We consider models for financial data by Lévy processes, including hyperbolic, normal inverse Gaussian, and Carr, Geman, Madan, and Yor (CGMY) processes. They are given by their Lévy triplet $(\mu(\theta), \sigma^2, e^{\theta x} g(x)\nu(dx))$, where μ denotes the drift, σ^2 the diffusion, and $e^{\theta x} g(x)\nu(dx)$ the Lévy measure, and the unknown parameter θ models the skewness of the process. We provide local asymptotic normality results and construct efficient estimators for the skewness parameter θ taking into account different discrete sampling schemes.

1. INTRODUCTION

Lévy processes, processes with stationary independent increments, became popular for modeling financial data during the last decade. However, the earliest attempt to model the stock behavior by a Lévy process, the Brownian motion, was by Bachelier (1900) in his Ph.D. thesis. More recently there has been a focus on Lévy processes with jumps. Hyperbolic Lévy motions (cf. Eberlein and Keller, 1995; Keller, 1997), generalized hyperbolic Lévy motions (cf. Prause, 1999; Raible, 2000), normal inverse Gaussian processes (cf. Barndorff-Nielsen, 1998; Rydberg, 1997), stable processes (cf. Rachev and Mittnik, 2000), variance gamma processes (cf. Madan and Senata, 1990), and CGMY processes, also called truncated Lévy flights (cf. Carr, Geman, Madan, and Yor, 2002) yield good models for log-return processes of prices and exchange rates. These are models of the form $\log S_t = X_t$, where S_t is the price process. Models based on these processes are less restrictive than the traditional ones; they allow jumps and include both finite and infinite activity and also bounded and unbounded variation. Furthermore, the empirical facts of excess kurtosis, skewness, and fat tails can be modeled more realistically.

One parameter that is especially important for modeling is the skewness parameter. The skewness of a distribution is modeled by multiplying with an exponential term $e^{\theta x}$, $\theta \in \mathbb{R}$. For $\theta > 0$ the resulting distribution is right/positive skewed depending on the size of θ; i.e., the bigger θ is the more weight is put on larger x. This parameter is an important parameter in finance, because

I thank Prof. Dr. L. Rüschendorf for his steady encouragement, the referees for helpful comments, and the German National Scholarship Foundation for financial support. Address correspondence to Jeannette H.C. Woerner, Institute for Mathematical Stochastics, University of Göttingen, Maschmühlenweg 8-10, D-37073 Göttingen, Germany; e-mail: woerner@math.uni-goettingen.de.

 0266-4666/04 $12.00

according to the empirical data the distribution of the log-return prices is mostly skewed (cf. Prause, 1999). Carr et al. (2002) perform a detailed analysis of skewness, finding that statistical data (i.e., the time series of stock returns) are significantly skewed either right or left, whereas risk-neutral data (i.e., data derived from option prices) are consistently left skewed. The skewness parameter is of course not the same as the financial term *skewness*, the appropriately normed third centered moment of a distribution. It measures the same effect, namely, the derivation from the symmetric distribution, but especially for fitting data we need the skewness parameter itself.

The main problem for estimating parameters entering a Lévy process X_t is that in general the process is given by the Lévy–Khinchin formula or in other words the characteristic function,

$$Ee^{iuX_t} = e^{iu\mu(\theta)-(u^2/2)\sigma^2(\theta)+\int(e^{iux}-1-iuh(x))g(x,\theta)\nu(dx)},$$

where μ denotes the drift, σ^2 the diffusion, and g the density w.r.t. ν of the Lévy measure, satisfying $\int(1\wedge x^2)g(x,\theta)\nu(dx)<\infty$, and $h(x)$ is some truncation function, which behaves like x in the neighborhood of zero and ensures integrability in the characteristic function. Common examples are $h(x) = x1_{|x|\le 1}(x)$ and $h(x) = x/(x^2+1)$. However, the density of the process itself is unknown and cannot be calculated explicitly, as for most stable, hyperbolic, generalized hyperbolic and CGMY processes. Hence we have to find conditions on the Lévy triplet $(\mu(\theta),\sigma^2(\theta),g(x,\theta)\nu(dx))$ that allow us to construct efficient estimators explicitly.

We look at the special case where our unknown parameter is the skewness, $g(x,\theta) = e^{\theta x}g(x)$, and focus on the concept of asymptotic statistics. We show that under very mild regularity conditions we obtain local asymptotic normality for the skewness parameter, giving us the maximal rate of convergence of a sequence of regular estimators and the minimal asymptotic variance, which turns out to be fully explicit, only involving the quantities of the Lévy triplet. Furthermore, we can then construct efficient estimators and show the relation to martingale estimating functions.

For all financial applications (e.g., fitting of models and also pricing derivatives and quantifying risk) it is important to have good estimators of the underlying parameters when the data are given at discrete time points, $X_\Delta, X_{2\Delta},\ldots,X_{n\Delta}$, because continuous data are hardly available in practice or not economical to observe.

We face two different sampling schemes; either we let the distance between the observations Δ be fixed and the number of observations n tend to infinity, or we let $n\Delta\to\infty$ as $\Delta\to 0$ and $n\to\infty$. The first sampling scheme seems to be of more practical interest, because the distance of the observations can be large. The second one is an approximation to the continuously observed model. The third possible discrete sampling scheme where $n\Delta = \text{const.} < \infty$, $n\to\infty$, and $\Delta\to 0$, which would be the classical framework of high-frequency data, is not possible in our setting. Heuristically, this can easily be seen when looking at the continuously observed model. In the continuously observed model we

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

obtain local asymptotic normality for the skewness parameter (cf. Akritas and Johnson, 1981) when the observed time T tends to infinity. Hence we also need in the discretely observed model that $\Delta n \to \infty$, because T may be identified with $n\Delta$. This has some important implications. We cannot infer the skewness parameter by high-frequency data over a fixed period of time. On the other hand, it has the advantage that it makes sense to infer the skewness parameter, even in the presence of other unknown parameters such as diffusion or scale, because they have a faster rate of convergence.

The outline of the paper is the following. In Section 2 we will review the concepts of local asymptotic normality and measure changes in Lévy processes and the result for continuously observed models. In Section 3 we will prove local asymptotic normality, and in Section 4 we construct efficient estimators and view them in the context of martingale estimating functions. Section 5 concludes.

2. PRELIMINARY RESULTS

The theory of local asymptotic statistics and the related efficiency results are established by Le Cam (1960) and Hájek (1972) and extended by Jeganathan (1981, 1983) and others.

This concept provides answers to important questions in estimation theory, e.g., how to characterize optimal estimators. Having local asymptotic normality we can specify the maximal rate of convergence of a sequence of estimators and the minimal asymptotic variance. Furthermore, it not only allows us to decide if a given sequence of estimators is efficient but also allows us to construct efficient estimators from suboptimal estimators by a one-step improvement, described, e.g., in Le Cam and Young (1990).

Let us recall the definition of local asymptotic normality (LAN). Let $p_n(X_0,\dots,X_n;\theta)$ be the joint density of $(X_0,\dots,X_n)$ under $\theta \in \Theta$ and $l_n(\theta + \delta_n h,\theta)$ the log-likelihood around θ, i.e.,

$$l_n(\theta + \delta_n h,\theta) = \log \frac{p_n(X_0,\dots,X_n;\theta + \delta_n h)}{p_n(X_0,\dots,X_n;\theta)},$$

where $\delta_n, h > 0$.

DEFINITION (LAN). *Let Θ be an open subset of $\mathbb{R}^d$ and $\mathcal{E}_n = (\Omega_n, \mathcal{F}_n, \{P_{n,\theta}: \theta \in \Theta\}), n \geq 1$ a sequence of experiments.*

For fixed $\theta \in \Theta$ $(\mathcal{E}_n)_n$ is called LAN *in θ, if*

(i) There exist $\delta_n = \delta_n(\theta) \downarrow 0$.

$\exists$ *$\mathcal{F}_n$-measurable $\Lambda_n : \Omega_n \to \mathbb{R}^d$*

$$\Gamma_n : \Omega_n \to \mathbb{R}^{d\times d},$$

Γ_n symmetric, strictly positive definite, such that for all $h \in \mathbb{R}^d$

$$l_n(\theta + \delta_n h,\theta) = h^T\Lambda_n - \frac{1}{2}\, h^T\Gamma_n h + o_{P_{n,\theta}}(1) \quad \text{as } n \to \infty.$$

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

(ii) There exists a finite, positive semidefinite, nonrandom matrix Γ*, such that*

$$\Gamma_n \xrightarrow{P} \Gamma,$$

$$\Lambda_n \xrightarrow{\mathcal{D}} N \xrightarrow[n\to\infty]{\mathcal{D}} (0,\Gamma),$$

as $n \to \infty$.

LAN im$\xrightarrow[n\to\infty]{\mathcal{D}}$plies that a sequence of estimators cannot converge to the true parameter value θ_0 at a rate faster than δ_n^{-1} and the asymptotic variance of a δ_n^{-1}-consistent estimator is bounded from below by Γ^{-1}.

The LAN property is often established by proving L^2-convergence, which implies the appropriate expansion of the log-likelihood function only involving first derivatives. However, in our case of the skewness parameter it turns out that it is easier to take a different way, looking at the Taylor expansion of the log-likelihood function up to the second order, because second derivatives exist. For the sampling scheme with fixed distance of observations Δ we are in the framework of independent and identically distributed (i.i.d.) random variables, which is well studied (cf. Witting, 1985; Janssen, 1992; van der Vaart, 1998). For the other sampling scheme we are in the framework of triangular arrays because of the dependence of the densities on n through Δ_n, and we have to establish an appropriate central limit theorem (CLT) and law of large numbers (LLN) for the first and second terms of the Taylor expansion of the log-likelihood function around the true parameter θ_0. For more details see the proof of Theorem 2 in Section 3.

Our aim is to prove LAN for discretely observed Lévy processes. However, for continuously observed Lévy processes there are some results known. Akritas and Johnson (1981) consider general purely discontinuous Lévy processes. We will state their result for the special case of the skewness parameter. We recall this result because the continuously observed model builds a natural benchmark for the model with discrete observations. Especially the sampling scheme with $n\Delta \to \infty$ as $\Delta \to 0$ and $n \to \infty$ may be well compared with the continuous model, as $n\Delta$ corresponds to t.

However, the continuous model only provides an optimality bound when the underlying measures are absolutely continuous. This is quite a strong restriction to the possible variation of parameters as we can see in the following theorem. Skorokhod (1957) derived this result first; for a detailed account, see, e.g., Shiryaev (1999).

THEOREM 1. *Let* X_t *be a Lévy process with triplet* $(\mu, \sigma^2, g(x)\nu(dx))$ *under some probability measure P. Then the following two conditions are equivalent.*

(1) *There is a probability measure* $Q \overset{loc}{\approx} P$ *such that* X_t *is a Q-Lévy process with triplet* $(\bar{\mu}, \bar{\sigma}^2, \bar{g}(x)\bar{\nu}(dx))$.

(2) *All of the following four conditions hold.*
- $\bar{g}(x)\bar{\nu}(dx) = k(x)g(x)\nu(dx)$ *for some Borel function* $k : \mathbb{R} \to (0,\infty)$.
- $\bar{\mu} = \mu + \int h(x)(k(x)-1)g(x)\nu(dx) + \sigma\beta$ *for some* $\beta \in \mathbb{R}$.

- $\bar{\sigma} = \sigma$.
- $\int(1 - \sqrt{k(x)})^2 g(x)\nu(dx) < \infty$.

This theorem implies that we cannot have the LAN property in the continuous model when we aim to estimate the diffusion, or a scalar factor in front of an infinite Lévy measure, because the underlying measures cannot be absolutely continuous. However, for the skewness parameter, i.e., $k(x) = \exp\{\theta x\}$, we can have an absolutely continuous change of measures and LAN for the continuously observed model with $\delta = 1/\sqrt{T}$ and $\Gamma = \int x^2 e^{\theta_0 x} g(x)\nu(dx)$, where T is the observed time that tends to infinity and $\sigma = 0$ (cf. Akritas and Johnson, 1981).

3. LOCAL ASYMPTOTIC NORMALITY

Let us now assume that we are given a Lévy process with a skewed Lévy measure $e^{\theta x} g(x)\nu(dx)$ and we aim to estimate the skewness parameter θ. First of all we are interested how the multiplicative term in the Lévy measure changes the distribution of the underlying Lévy process. We obtain the following lemma. Though we shall only look at Lévy processes with unbounded support, the same calculations hold for processes with bounded support, especially for subordinators.

LEMMA 1. *Denote by $p_t(x)$ the density of the Lévy process w.r.t. m, given by the triplet $(\mu, \sigma^2, g(x)\nu(dx))$ and by $p_t(x,\theta)$ the density corresponding to the process with the skewed Lévy measure $e^{\theta x} g(x)\nu(dx)$. Furthermore, assume that for all $\theta \in U(\theta_0) \subset \Theta$, where $U(\theta_0)$ is a neighborhood of the true parameter θ_0,*

$$\int_{|x|\geq 1} e^{\theta x} g(x)\nu(dx) < \infty. \tag{1}$$

Then we obtain

$$p_t(x,\theta) = \frac{e^{\theta x} p_t(x)}{\int e^{\theta x} p_t(x) m(dx)}, \tag{2}$$

and the corresponding drift is $\bar{\mu} = \mu + \int h(x)(e^{\theta x} - 1) g(x)\nu(dx)$.

Remark. (1) For processes with finite variation, i.e., $\int(|x| \wedge 1) g(x)\nu(dx) < \infty$, e.g., all compound Poisson processes and subordinators, we may take $h(x) = 0$ and obtain $\bar{\mu} = \mu$.

(2) We consider densities p_t w.r.t. to some measure m. However, of most practical interest, except for compound Poisson processes, is when m equals the Lebesgue measure, as is the case in our examples. Even though the density p_t might not be known explicitly, conditions on the Lévy measure may ensure

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

its existence; e.g., as was shown in Tucker (1962), infiniteness of the Lévy measure together with a Lebesgue density of the Lévy measure already ensures the existence of a Lebesgue density p_t. However, the relation between the existence of a density p_t and the existence of a density of the corresponding Lévy measure and its mass near zero is much more complex. A detailed outline is given in Sato (1999).

Proof. Under the assumption (1) the denominator in (2) is well defined. We assume that the characteristic function corresponding to p_t is

$$e^{t\left(iu\mu-(1/2)\sigma^2+\int(e^{iux}-1-iuh(x))g(x)\nu(dx)\right)}.$$

Hence we can calculate the characteristic function of the skewed distribution

$$\begin{aligned}
&\int e^{iux}\frac{e^{\theta x}p_t(x)}{\int e^{\theta y}p_t(y)m(dy)}m(dx)\\
&\quad=\exp\left\{t\left(iu\mu+\theta\mu-\frac{1}{2}\sigma^2u^2+\frac{1}{2}\sigma^2\theta^2\right.\right.\\
&\qquad\qquad+\int(e^{iux+\theta x}-1-iuh(x)-\theta h(x))g(x)\nu(dx)-\theta\mu\\
&\qquad\qquad\left.\left.-\frac{1}{2}\sigma^2\theta^2-\int(e^{\theta x}-1-\theta h(x))g(x)\nu(dx)\right)\right\}\\
&\quad=\exp\left\{t\left(iu\left(\mu+\int h(x)(e^{\theta x}-1)g(x)\nu(dx)\right)-\frac{1}{2}\sigma^2u^2\right.\right.\\
&\qquad\qquad\left.\left.+\int(e^{iux}-1-iuh(x))e^{\theta x}g(x)\nu(dx)\right)\right\},
\end{aligned}$$

which yields the desired result. ∎

Of course this result is not new. The principle of multiplying with an exponential factor is well established in different disciplines of stochastics but is named differently. In statistics the family of distributions or processes derived by varying the skewness parameter is called the natural exponential family (cf. Janssen, 1992; Küchler and Sørensen, 1997). In finance the the distribution obtained by adding the skewness is called the Esscher transform (cf. Shiryaev, 1999). This concept is also an easy way of explicitly calculating equivalent martingale measures.

Using (2) we can prove LAN for the skewness parameter. Though we start with a general Lévy process with unknown density, we obtain a fully explicit result of the minimal asymptotic variance, depending only on the quantities of the Lévy triplet, for both sampling schemes. Furthermore, our minimal asymptotic variance for the sampling scheme with $\Delta \to 0$ turns out to be the same as in the continuously observed model.

THEOREM 2. *Assume condition (1) for $\theta \in U(\theta_0) \subset \Theta$, a neighborhood of θ_0; then we obtain* LAN

(i) with $\delta_n = 1/\sqrt{n}$ and

$$\Gamma = \Delta\left(\sigma^2 + \int x^2 e^{\theta_0 x} g(x)\nu(dx)\right),$$

as $n \to \infty$, under the sampling scheme with fixed Δ,

(ii) with $\delta_n = 1\big/\sqrt{n\Delta}$ and

$$\Gamma = \sigma^2 + \int x^2 e^{\theta_0 x} g(x)\nu(dx),$$

as $n\Delta \to \infty$, under the sampling scheme with $\Delta \to 0$ and $n \to \infty$.

Proof. Because we have i.i.d. increments as a result of the structure of the Lévy processes, we can use the results for i.i.d. random variables.

First, we have to calculate the derivatives. We may interchange integration and differentiation if θ is in the interior of Θ and obtain

$$\begin{aligned}
\frac{\partial}{\partial\theta} p_t(x,\theta) &= \frac{\partial}{\partial\theta} \frac{e^{\theta x} p_t(x)}{\int e^{\theta x} p_t(x) m(dx)} \\
&= \frac{x e^{\theta x} p_t(x) \int e^{\theta x} p_t(x) m(dx) - e^{\theta x} p_t(x) \left(\frac{\partial}{\partial\theta} \int e^{\theta x} p_t(x) m(dx)\right)}{\left(\int e^{\theta x} p_t(x) m(dx)\right)^2} \\
&= x p_t(x,\theta) - \int x p_t(x,\theta) m(dx)\, p_t(x,\theta) \\
\frac{\partial^2}{\partial\theta^2} p_t(x,\theta) &= x \frac{\partial}{\partial\theta} p_t(x,\theta) - \int x \frac{\partial}{\partial\theta} p_t(x,\theta) m(dx) p_t(x,\theta) \\
&\quad - \int x p_t(x,\theta) m(dx) \frac{\partial}{\partial\theta} p_t(x,\theta) \\
&= x^2 p_t(x,\theta) - 2\int x p_t(x,\theta) m(dx) x p_t(x,\theta) \\
&\quad - \int x^2 p_t(x,\theta) m(dx) p_t(x,\theta) \\
&\quad + 2\left(\int x p_t(x,\theta) m(dx)\right)^2 p_t(x,\theta).
\end{aligned}$$

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

(i) For the sampling scheme with fixed distance of observations Δ, we are in the framework of i.i.d. random variables with densities; hence conditions under which LAN in θ_0 holds are well known (see, e.g., Witting, 1985, p. 179; Le Cam and Young, 1990, Ch. 6.3). The conditions are that the density is positive m a.s. and continuously differentiable w.r.t. the unknown parameter θ in a neighborhood $U(\theta_0)$. Then we obtain LAN with maximal rate of convergence $\delta_n = 1/\sqrt{n}$ and

$$\Gamma(\theta_0) = \int \frac{\dot{p}_\Delta^2(x,\theta_0)}{p_\Delta(x,\theta_0)}\, m(dx),$$

which is the inverse of the minimal asymptotic variance, provided $\Gamma(\theta)$ is finite for all $\theta \in U(\theta_0)$ and continuous in θ_0. This result applies as for $\theta \in U(\theta_0)$, $p_t(x,\theta)$ is continuously differentiable in θ and $p_t(x,\theta) > 0\ [m]$. We obtain LAN with $\delta_n = 1\big/\sqrt{n}$ and

$$\begin{aligned}\Gamma(\theta_0) &= \int \frac{(\dot{p}_\Delta(x,\theta_0))^2}{p_\Delta(x,\theta_0)}\, m(dx)\\ &= \int x^2 p_\Delta(x,\theta_0) m(dx) - \left(\int x p_\Delta(x,\theta_0) m(dx)\right)^2\\ &= \Delta\left(\sigma^2 + \int x^2 e^{\theta_0 x} g(x)\nu(dx)\right),\end{aligned}$$

because $\Gamma(\theta)$ is obviously finite for $\theta \in U(\theta_0)$ and continuous in θ_0. For the last equation we use the well-known moment representation, by the derivatives of the characteristic function, $\hat{p}_\Delta^{(n)}(0) = (i)^n \int x^n p_\Delta(x) m(dx)$.

(ii) For the sampling scheme with $\Delta \to 0$ we need a few more conditions to ensure the LAN property, which results basically from conditions needed to perform the CLT and LLN for triangular schemes (see Woerner, 2001; Gnedenko and Kolmogorov, 1968). The conditions are as follows. Assume that there exists a neighborhood $U(\theta_0)$ of θ_0 such that the density $p_t(x,\theta)$ w.r.t. m of the Lévy process is two times continuously differentiable w.r.t. the unknown parameter θ and regular (i.e., $\int \dot{p}_t(x,\theta) m(dx) = (\partial/\partial\theta) \int p_t(x,\theta) m(dx) = 0$, $\int \ddot{p}_t(x,\theta) m(dx) = (\partial^2/\partial\theta^2) \int p_t(x,\theta) m(dx) = 0$) for all $\theta \in U(\theta_0)$. Denote by $\theta_n = \theta_n(x)$ a measurable function from $\mathbb{R} \to \left[\theta_0, \theta_0 + h\big/\sqrt{n\Delta}\right]$. Furthermore, assume that for all $\epsilon > 0$ as $n \to \infty$ and $\Delta \to 0$, such that $n\Delta \to \infty$,

$$\text{(i)}\quad \frac{1}{\Delta}\int \frac{\dot{p}_\Delta^2(x,\theta_n(x))}{p_\Delta^2(x,\theta_n(x))}\, p_\Delta(x,\theta_0) m(dx) \to \Gamma(\theta_0) < \infty,$$

$$\text{(ii)}\quad \frac{1}{\Delta}\int \frac{\ddot{p}_\Delta(x,\theta_n(x))}{p_\Delta(x,\theta_n(x))}\, p_\Delta(x,\theta_0) m(dx) \to 0,$$

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

(iii) $\displaystyle \frac{1}{\Delta}\int_{\left|\frac{\dot p_\Delta(x,\theta_0)}{p_\Delta(x,\theta_0)}\right|\geq\epsilon\sqrt{n\Delta}} \frac{\dot p_\Delta^2(x,\theta_0)}{p_\Delta(x,\theta_0)}\, m(dx)\to 0,$

(iv) $\displaystyle \int_{\left|\frac{1}{\Delta}\frac{\dot p_\Delta^2(x,\theta_n(x))}{p_\Delta^2(x,\theta_n(x))}-\Gamma(\theta_0)\right|>\epsilon n}\left|\frac{1}{\Delta}\frac{\dot p_\Delta^2(x,\theta_n(x))}{p_\Delta^2(x,\theta_n(x))}-\Gamma(\theta_0)\right| p_\Delta(x,\theta_0)m(dx)\to 0,$

(v) $\displaystyle \frac{1}{\Delta}\int_{\frac{1}{\Delta}\left|\frac{\ddot p_\Delta(x,\theta_n(x))}{p_\Delta(x,\theta_n(x))}\right|>\epsilon n}\left|\frac{\ddot p_\Delta(x,\theta_n(x))}{p_\Delta(x,\theta_n(x))}\right| p_\Delta(x,\theta_0)m(dx)\to 0.$

Then we obtain LAN with $\delta_n = 1/\sqrt{n\Delta}$ and $\Gamma=\Gamma(\theta_0)$ for the sampling scheme with $\Delta\to 0$ as $n\to\infty$.

Finding dominating functions to ensure these conditions is straightforward because of the special simple structure that θ only enters in the exponential term and Δ only as a polynomial factor, by applying the moment representation to the integrals w.r.t. $p_\Delta(x,\theta)m(dx)$. We look at the details of condition (i); the others can be checked analogously.

$$\begin{aligned}
&\frac{1}{\Delta}\int\frac{\dot p_\Delta^2(x,\theta_n(x))}{p_\Delta^2(x,\theta_n(x))}\,p_\Delta(x,\theta_0)m(dx)\\
&\quad=\frac{1}{\Delta}\Bigg(\int x^2p_\Delta(x,\theta_0)m(dx)-2\int xp_\Delta(x,\theta_0)m(dx)\int xp_\Delta(x,\theta_n(x))m(dx)\\
&\qquad+\bigg(\int xp_\Delta(x,\theta_n(x))m(dx)\bigg)^2\Bigg)\\
&\quad=\sigma^2+\int x^2g(x,\theta_0)\nu(dx)\\
&\qquad-2\Delta\bigg(\mu+\int h(x)(e^{\theta_0x}-1)g(x)\nu(dx)+\int(x-h(x))g(x,\theta_0)\nu(dx)\bigg)\\
&\qquad\times\bigg(\mu+\int h(x)(e^{\theta_n(x)x}-1)g(x)\nu(dx)+\int(x-h(x))g(x,\theta_n(x))\nu(dx)\bigg)\\
&\qquad+\Delta\bigg(\mu+\int h(x)(e^{\theta_n(x)x}-1)g(x)\nu(dx)+\int(x-h(x))g(x,\theta_n(x))\nu(dx)\bigg)^2\\
&\quad\to\sigma^2+\int x^2g(x,\theta_0)\nu(dx),
\end{aligned}$$

as $n\to\infty$, $\Delta\to 0$, and $n\Delta\to\infty$. ■

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

Example 1 (Normal inverse Gaussian process)

The normal inverse Gaussian process is characterized by the Lévy triplet

$$\left(\mu, 0, \frac{\delta\alpha K_1(\alpha|x|)e^{\beta x}}{\pi|x|}\right),$$

where K_1 denotes the modified Bessel function of third order with index 1 and $\mu \in I\!R$, $\alpha, \delta > 0$, $0 \leq |\beta| < \alpha$. Because as $|x| \to 0$

$$\frac{\delta\alpha K_1(\alpha|x|)e^{\beta x}}{\pi|x|} \sim \text{const.}\ \frac{1}{x^2},$$

the process is of unbounded variation and possesses infinitely many jumps. Hence it also has a density w.r.t. the Lebesgue measure,

$$p_t(x, \alpha, \beta, \delta, \mu) = \frac{\alpha e^{t\delta\sqrt{\alpha^2-\beta^2}-t\beta\mu}}{\pi} \frac{t\delta}{\sqrt{t^2\delta^2 + (x - t\mu)^2}} \times K_1(\alpha\sqrt{t^2\delta^2 + (x - t\mu)^2})e^{\beta x}.$$

As pointed out in Barndorff-Nielsen (1998) this process is used both for modeling turbulence, in particular when the Reynolds number is high, and in finance. This is due to some special properties of the normal inverse Gaussian process, such as possible asymmetry modeled by the skewness parameter β, unbounded variation, and semiheavy tails; namely, as $|x| \to \infty$

$$\frac{\delta\alpha K_1(\alpha|x|)e^{\beta x}}{\pi|x|} \sim \text{const.}\ |x|^{-3/2}e^{-\alpha|x|+\beta x}.$$

The parameter β may be estimated according to Theorem 2. Here assumption (1) is satisfied. Hence for the sampling scheme with Δ fixed and $n \to \infty$ we have $\delta_n = 1/\sqrt{n}$ and

$$\Gamma_\beta = \int x^2 p_\Delta(x, \alpha, \beta, \delta)\, dx - \left(\int x p_\Delta(x, \alpha, \beta, \delta)\, dx\right)^2$$

$$= \Delta \int x^2 g(x, \alpha, \beta, \delta)\, dx$$

$$= \Delta \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{5/2}},$$

which is indeed the same result as using the density and the results for i.i.d. random variables.

Example 2 (Gamma process)

The gamma process is characterized by the Lévy triplet

$$\left(0,0,\frac{\alpha e^{-\beta x}}{x}\right),$$

with $\alpha, \beta, x > 0$. Hence the process is a subordinator and only possesses non-negative jumps. The density w.r.t. the Lebesgue measure can be calculated, and we can see that the name reflects the fact that the increments are distributed according to a gamma function,

$$p_t(x,\alpha,\beta) = \frac{\beta^{\alpha t} x^{\alpha t-1} e^{-\beta x}}{\Gamma(\alpha t)}.$$

Analogously to Example 1 β can be estimated, and we obtain for the sampling scheme with Δ fixed and $n \to \infty$, $\delta_n = 1/\sqrt{n}$ and

$$\Gamma_\beta = \Delta \int x^2 g(x,\alpha,\beta)\,dx = \Delta\,\frac{\alpha}{\beta^2},$$

which is again the same result as using the density and the i.i.d. results.

Example 3 (Hyperbolic Lévy motion)

The hyperbolic Lévy motion which was introduced by Barndorff-Nielsen (1977) for modeling mass-size distributions of aeolin sand deposits, has also been applied to some other areas of interest, e.g., turbulence data (cf. Barndorff-Nielsen, 1996) and to financial data (cf. Eberlein and Keller, 1995; Keller, 1997; Rydberg, 1997; Prause, 1999; Raible, 2000).

The hyperbolic Lévy motion may be characterized by the Lévy triplet

$$\left(\mu,0,\frac{e^{\beta x}}{|x|}\left(\int_0^\infty \frac{e^{-\sqrt{2y+\alpha^2}|x|}}{\pi^2 y(J_1^2(\delta\sqrt{2y}) + Y_1^2(\delta\sqrt{2y}))}\,dy + e^{-\alpha|x|}\right)\right),$$

where J_1 denotes the Bessel function of the first order with index 1 and Y_1 the Bessel function of the second order with index 1. Furthermore, we have $\alpha, \delta > 0$ and $0 \le |\beta| < \alpha$. Keller (1997) has established that the density of the Lévy measure behaves like x^{-2} at the origin; hence the process is of unbounded variation. Though the process possesses a density, we cannot calculate it analytically. Only the distribution for $t = 1$ can be written down explicitly; it is the hyperbolic distribution

$$\frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2-\beta^2})}\, e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)},$$

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

where K_1 denotes the modified Bessel function of third order with index 1. We can estimate β analogously to Example 1 and obtain for the sampling scheme with Δ fixed and $n \to \infty$, $\delta_n = 1/\sqrt{n}$ and

$$\Gamma_\beta = \Delta \int x^2 g(x,\alpha,\beta)\,dx$$
$$= \delta^2 \left(\frac{K_2(\chi)}{\chi K_1(\chi)} + \frac{\beta^2}{\alpha^2 - \beta^2} \left[\frac{K_3(\chi)}{K_1(\chi)} - \left(\frac{K_2(\chi)}{K_1(\chi)} \right)^2 \right] \right),$$

where $\chi = \delta\sqrt{\alpha^2 - \beta^2}$ using Keller (1997).

Example 4 (CGMY process)

The CGMY process, named after Carr, Geman, Madan, and Yor, or in physical literature called truncated stable or truncated Lévy flight, is given by the Lévy triplet

$$\left(0,0,C\,\frac{\exp\{-G|x|\}}{|x|^{1+Y}}\,1_{x<0} + C\,\frac{\exp\{-M|x|\}}{|x|^{1+Y}}\,1_{x>0}\right),$$

where $C > 0$, $G \geq 0$, $M \geq 0$, and $Y < 2$. This class of processes is a flexible model for index dynamics and also for the dynamics of individual stocks, because by varying the parameters it allows all features of finite and infinite activity, bounded and unbounded variation, and also skewness to be modeled directly by the characteristic function (cf. Carr et al., 2002). The variance gamma process (cf. Madan and Senata, 1990) is a special case of the CGMY process. In general the density of the process is not known explicitly, but as in the previous examples we can infer the parameters M or G in the special one-sided case, when either $G = \infty$ or $M = \infty$, only by the knowledge of the Lévy measure. When $G = \infty$ we obtain $(0,0,C(\exp\{-M|x|\}/|x|^{1+Y})1_{x>0})$ and for the sampling scheme with Δ fixed and $n \to \infty$, $\delta_n = 1/\sqrt{n}$ and

$$\Gamma_M = \Delta C \int_0^\infty x^2\,\frac{\exp(-Mx)}{x^{Y+1}}\,dx = \Delta C \Gamma(2 - Y) M^{Y-2}.$$

4. EFFICIENT ESTIMATORS

With this explicit result for the minimal asymptotic variance we can now try to find a sequence of estimators that is efficient. Because the continuous model is a benchmark for the discretely observed model, we look at the continuous likelihood function to get an idea of what discrete estimation functions may look like. In the continuous model the likelihood function is given by

$$L_t(\theta) = \exp\{\theta Y_t - t \log E e^{\theta Y_1}\},$$

where Y_t denotes the process under the measure $p_t(x)m(dx)$. Changing to the process X_t under the measure $e^{\theta x}p_t(x)m(dx)/(\int e^{\theta x}p_t(x)m(dx))$ and discretizing the time, i.e., $t = n\Delta$, leads to the log-likelihood function

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

$$l_n(\theta) = X_{n\Delta} - n\Delta E_\theta X_1, \tag{3}$$

which indeed provides an appropriate estimating function.

We obtain that only the knowledge of the last observation is important for the estimation procedure. In other words, the last observation contains all necessary information. This is however not surprising, because from the theory of exponential families it is well known that Y_t or $X_{n\Delta}$, respectively, is sufficient for $\{P_t^\theta, \theta \in \Theta\}$.

THEOREM 3 (Optimal estimators).

(i) Let $\hat{\theta}_n^\Delta$ be the solution of $l_n(\theta) = X_{n\Delta} - n\Delta E_\theta X_1 = 0$; then

$$\sqrt{n}(\hat{\theta}_n^\Delta - \theta) \xrightarrow{\mathcal{D}} N\left(0, \left[\Delta\left(\sigma^2 + \int x^2 e^{\theta x} g(x)\nu(dx)\right)\right]^{-1}\right)$$

as $n \to \infty$. Hence it is efficient for the sampling scheme with fixed Δ and $n \to \infty$.

(ii) Let $\hat{\theta}_{n\Delta}$ be the solution of $l_n(\theta) = X_{n\Delta} - n\Delta E_\theta X_1 = 0$; then

$$\sqrt{n\Delta}(\hat{\theta}_{n\Delta} - \theta) \xrightarrow{\mathcal{D}} N\left(0, \left(\sigma^2 + \int x^2 e^{\theta x} g(x)\nu(dx)\right)^{-1}\right),$$

as $n\Delta \to \infty$, where $\Delta \to 0$ and $n \to \infty$. Hence it is efficient for the sampling scheme when $n\Delta \to \infty$ as $\Delta \to 0$ and $n \to \infty$.

Proof.

(i) Using $X_0 = 0$ and $n\Delta E_\theta X_1 = E_\theta X_{n\Delta}$, we can rewrite

$$\begin{aligned} l_n(\theta) &= X_{n\Delta} - n\Delta E_\theta X_1 \\ &= \sum_{i=1}^n [X_{i\Delta} - X_{(i-1)\Delta} - E_\theta[X_{i\Delta} - X_{(i-1)\Delta}]]. \end{aligned} \tag{4}$$

This yields asymptotic normality by inserting

$$\frac{1}{\sqrt{n}} l_n(\theta) \xrightarrow{\mathcal{D}} N(0, \Delta \mathrm{Var}_\theta(X_1))$$

as $n \to \infty$ by the CLT and

$$\begin{aligned} \frac{1}{n} \dot{l}_n(\theta) &= -\Delta \frac{\partial}{\partial\theta}\left(\int h(x)(e^{\theta x} - 1)g(x)\nu(dx)\right. \\ &\qquad \left. - \int (x - h(x))e^{\theta x} g(x)\nu(dx)\right) \\ &= -\Delta \int x^2 e^{\theta x} g(x)\nu(dx) = -\Delta \mathrm{Var}_\theta(X_1) \end{aligned}$$

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

in the expansion

$$l_n(\theta) = -\dot{l}_n(\theta_n)(\hat{\theta}_n - \theta),$$

where $\theta_n \in (\theta, \hat{\theta}_n)$.

(ii) This is analogous to using the CLT for triangular schemes as in Theorem 2 for the convergence in distribution and calculating the explicit form of $\dot{l}_n(\theta)/(n\Delta)$ as in (i). ■

With Theorem 3 equation (3) now provides an easily computable estimating function that indeed leads to efficient estimators. It is especially simple because it only involves the observations and the first moment. Depending on the form of the first moment the equation can sometimes even be solved analytically.

Example 5 (Gamma process)

For the gamma process equation (3) is

$$X_{n\Delta} - n\Delta \frac{\alpha}{\beta} = 0.$$

Hence we obtain $\hat{\beta}_n^{\Delta} = (n\Delta/X_{n\Delta})\alpha$ for the sequence of efficient estimators as $n \to \infty$ for the sampling scheme with fixed Δ.

Equation (4) can also be viewed as the simplest form of a martingale estimating function, having the general form

$$G_n(\theta) = \sum_{i=1}^{n} \left(f(Y_i, \theta) - \int f(x, \theta) p_\Delta(x, \theta) m(dx) \right),$$

where $Y_i = X_{i\Delta} - X_{(i-1)\Delta}$ i.i.d. distributed according to $p_\Delta(.,\theta)$. We can now show that our sequence of estimators is not only optimal in the sense of local asymptotic statistics but also in the sense of Godambe and Heyde (1987), which is the classical optimality concept for martingale estimating functions.

Let us first give a short review on the concept of martingale estimating functions and the definition of optimality by Godambe and Heyde (1987).

The basic problem is that we want to draw inference for discretely observed stochastic processes when the likelihood function is unknown. Because in general the maximum likelihood estimator performs quite well, the idea is to approximate the unknown score function to obtain an approximate maximum likelihood estimator. Using an approximation, the problem might occur that we perhaps do not have mean zero and hence eventually will get biased estimates, especially when the distance of observations Δ is bounded away from zero. A solution is to approximate the score function by a zero mean martingale w.r.t. the filtration generated by the observations. This implies that we obtain consistent and asymptotically normal estimators.

The optimality concept of Godambe and Heyde (1987) and Heyde (1988) formalizes the heuristics that the optimal element in a given class of martingale estimating functions is the one whose L^2-distance to the true score function is minimal, or in the partial order of nonnegative matrices the distance to the score function is minimal.

DEFINITION (O_A-optimality). *Let*

$$G_n^* \in \mathcal{M}_1 \subset \mathcal{M} = \{G_n | G_n \text{ martingale}, EG_n = 0, \langle G, G^T \rangle, \bar{G}_n \text{ nonsingular}\},$$

where $\bar{G}_n$ denotes the compensator of $\dot{G}_n$ and $\langle G, G^T \rangle$ the quadratic characteristic of G_n. Then $G_n^* \in \mathcal{M}_1$ is called O_A-optimal in $\mathcal{M}_1$ if and only if

$$(\bar{G}_n)^{-1} \langle G, G^{*T} \rangle_n = (\bar{G}_n^*)^{-1} \langle G^*, G^{*T} \rangle_n \quad \forall G_n \in \mathcal{M}_1.$$

As a result of the special structure of i.i.d. increments, all considerations simplify greatly for our model compared to general stochastic processes. By straightforward calculations we obtain $f^*(x) = x$ for the optimal G^*. Hence we also have optimality in the sense of Godambe and Heyde (1987) for the estimators in Theorem 3.

5. CONCLUSION

We derived local asymptotic normality for the skewness parameter of Lévy processes that are observed at discrete time points only. This provides an efficiency criterion in terms of the maximal rate of convergence and the minimal asymptotic variance for a sequence of estimators. Furthermore, we obtained easily computable estimating functions that lead to efficient estimators both in the sense of asymptotic statistics and in the sense of Godambe and Heyde (1987) for martingale estimating functions. Hence our results enable us to optimally infer the skewness parameter from discrete observations for the popular Lévy process models, such as generalized hyperbolic, normal inverse Gaussian, and CGMY, even when other unknown parameters are involved.

REFERENCES

Akritas, M.G. & R.A. Johnson (1981) Asymptotic inference in Levy processes of the discontinuous type. *Annals of Statistics* 9, 604–614.

Bachelier, L. (1900) *Theorie de la speculation*. Gauthier-Villar.

Barndorff-Nielsen, O.E. (1977) Exponentially decreasing distributions for the logarithm of particle size. In *Proceedings of the Royal Society London A*, vol. 353, pp. 401–419.

Barndorff-Nielsen, O.E. (1996) Probability and statistics: Selfdecomposability, finance and turbulence. In L. Accardi et al. (ed.), *Proceedings of the Conference "Probability towards 2000,"* Centre for Applied Probability, Cornell University.

Barndorff-Nielsen, O.E. (1998) Processes of normal inverse Gaussian type. *Finance and Stochastics* 2, 41–68.

Carr, P., H. Geman, D. B. Madan, & M. Yor (2002) The fine structure of asset returns: An empirical investigation. *Journal of Business* 75, 305–332.

Eberlein, E. & U. Keller (1995) Hyperbolic distributions in finance. *Bernoulli* 1, 281–299.

Gnedenko, B.V. & A.N. Kolmogorov (1968) *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley.

https://doi.org/10.1017/S0266466604205060 Published online by Cambridge University Press

Godambe, V.P. & C.C. Heyde (1987) Quasi-likelihood and optimal estimation. *International Statistical Review* 55, 231–244.

Hájek, J. (1972) Local asymptotic minimax and admissibility in estimation. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* vol. 1, pp. 175–194.

Heyde, C.C. (1988) Fixed sample and asymptotic optimality for classes of estimating functions. *Contemporary Mathematics* 80, 241–247.

Janssen, A. (1992) Conditions for local asymptotic normality of exponential families. *Statistics and Decisions* 10, 173–182.

Jeganathan, P. (1981) On a decomposition of the limit distribution of a sequence of estimators. *Sankhyā A* 43, 26–36.

Jeganathan, P. (1983) Some properties of risk functions in estimation when the limit of the experiment is mixed normal. *Sankhyā A* 45, 66–86.

Keller, U. (1997) *Realistic Modelling of Financial Derivatives*. Ph.D. thesis, University of Freiburg.

Küchler, U. & M. Sørensen (1997) *Exponential Families of Stochastic Processes*. Springer-Verlag.

Le Cam, L. (1960) Locally asymptotically normal families of distributions. *University of California Publications in Statistics* 3, 37–98.

Le Cam, L. & Yang, G. (1990) *Asymptotics in Statistics*. Springer.

Madan, D.B. & E. Senata (1990) The variance gamma (VG) model for share market returns. *Journal of Business* 63, 511–524.

Prause, K. (1999) *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*. Ph.D. thesis, University of Freiburg.

Rachev, S.T. & S. Mittnik (2000) *Stable Paretian Models in Finance*. Wiley.

Raible, S. (2000) *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. Ph.D. thesis, University of Freiburg.

Rydberg, T. (1997) The normal inverse Gaussian Levy process: Simulation and approximation. *Communications in Statistics, Stochastic Models* 13, 887–910.

Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Shiryaev, A.N. (1999) *Essentials of Stochastic Finance: Facts, Models and Theory*. World Scientific.

Skorokhod, A.V. (1957) On the differentiability of measures which correspond to stochastic processes, part I. Processes with independent increments. *Theory of Probability and Applications* 2, 407–432.

Tucker, H.G. (1962) Absolute continuity of infinitely divisible distributions. *Pacific Journal of Mathematics* 12, 1125–1129.

van der Vaart, A.W. (1998) *Asymptotic Statistics*. Cambridge Series on Statistical and Probabilistic Mathematics. Cambridge University Press.

Witting, H. (1985) *Mathematische Statistik I*. B.G. Teubner.

Woerner, J.H.C. (2001) *Statistical Analysis for Discretely Observed Lévy Processes*. Ph.D. thesis, University of Freiburg.