Asymptotic Equivalence of Continuously and Discretely Sampled Jump-Diffusion Models

Becheri, Irene G.
Tilburg University, Econometrics and Operations Research
Warandelaan 2
5037 AB Tilburg, The Netherlands
E-mail: i.g.becheri@uvt.nl

Drost, Feike C.
Tilburg University, Econometrics and Operations Research
Warandelaan 2
5037 AB Tilburg, The Netherlands
E-mail: f.c.drost@uvt.nl

Werker, Bas J.M.
Tilburg University, Econometrics and Operations Research
Warandelaan 2
5037 AB Tilburg, The Netherlands
E-mail: b.j.m.werker@uvt.nl

Introduction

Models specified in continuous time have attracted much attention over the past decades in various fields of applications, for instance, physics and finance. Concrete specifications of continuous time models generally contain unknown parameters that have to be estimated. For this inference problem, the existing literature either assumes continuous-time observations to be given, or discusses inference based on observations sampled discretely in time. Several authors have studied the problem of estimating parameters when a diffusion process has been observed continuously. Reviews can be found in Basawa and Prakasa Rao (1980), Liptser and Shiryaev (1978) and Kutoyants (2003). Notable papers with continuous-time observations are Sørensen (1991), discussing the inference for diffusion processes with jumps from continuous observations under the setting which includes the non-ergodic case, and Akritas and Johnson (1981), in which the asymptotic inference for general Lévy processes has been studied based on likelihood theory. Estimation problems for discretely observed diffusion processes have been studied by many authors as well, for instance see Genon-Catalot and Jacod (1993), Dacunha-Castelle and Florens-Zmirou (1986) and Bibby and Sørensen (1995). Among all possible observation schemes, we are specifically interested in high-frequency sampling schemes. These schemes assume the number of observations, n, increasing to infinity, the distance between observations, h_n, tending to zero and the whole observational interval, nh_n, going to infinity. The first work on this scheme is Prakasa-Rao (1983), proposing a least squares approach in case where the diffusion coefficient is constant and the parameter of interest is one-dimensional. To prove asymptotic normality of this estimator Prakasa-Rao (1983) assumed the time between observations tends to zero sufficiently fast that $nh_n^2 \to 0$, and referred to this condition as a rapidly increasing experimental design assumption.

The present paper gives sufficient conditions under which, from a statistical point of view, discrete observations from a high-frequency sampling scheme (satisfying the rapidly increasing experimental design condition) contain as much information as continuous-time observations. It does so in a fairly general jump-diffusion setting with time-varying drift and jump intensity. Our model assumes the volatility function of the diffusion term to be fully specified, that is, it does not contain unknown parameters of interest. The reason for this is that, with a continuous-time observations, the volatility function is essentially observable using the quadratic variation of the process and the inference problem of the diffusion term is degenerated. Some results for estimating the volatility function are given in Doob (1953) and Genon-Catalot and Jacod (1994).

There is a large literature on the probabilistic convergence of discrete-time processes to continuous-time jump-diffusions. However, it is important to note that such results do not imply that the statistical inference problems based on both types of processes are close in any sense. A prime example is given in Wang (2002). This paper shows that while discrete-time GARCH processes are known to converge to a continuous-time diffusion (see, Wang (2002)), the inference problems are different. The reason is that for (asymptotic) equivalence of the inference problems, the likelihood-ratio processes for various parameter values need to converge, not the process itself. This is another reason why we assume throughout this paper that the volatility function does not contain unknown parameters.

This paper offers three contributions. First of all, we provide a Local Asymptotic Normality result for a continuous-time observation from a jump-diffusion process. This result extends both Kutoyants (2003) in which the case of diffusion processes is discussed and, Jacod and Shiryaev (2002) where conditions for LAN to hold are given for statistical models based on counting processes. Secondly, we give a sufficient condition on a general jump-identification mechanism that allows one to reconstruct the continuous-time central sequence from discrete-time observations only. Thirdly, we discuss an existing technique to identify jumps, proposed by Shimizu and Yoshida (2006), and show that this fulfills the required condition.

LAN for continuous-time observations

Let $\Theta \subset \mathbb{R}^d$ be an open parameter space. We consider, for $\boldsymbol{\theta} \in \Theta$, the strong solution of

$$(1) \quad \mathrm{d}X_t = \mu(\boldsymbol{\theta}, X_{t-})\,\mathrm{d}t + \sigma(X_{t-})\,\mathrm{d}W_t + \mathrm{d}J_t^{\boldsymbol{\theta}},$$

where $J_t^{\boldsymbol{\theta}}$ is defined by

$$J_t^{\boldsymbol{\theta}} = \sum_{i=1}^{N_t^{\boldsymbol{\theta}}} U_i.$$

This strong solution lives on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_{\boldsymbol{\theta}})$, where $\mathbb{P}_{\boldsymbol{\theta}}$ is such that W is a standard Brownian motion (with respect to $(\mathcal{F}_t)_{t\geq 0}$) and, conditionally on the path of X, $(N_t^{\boldsymbol{\theta}})_{t\geq 0}$ is a Poisson process with time varying intensity $(\lambda(\boldsymbol{\theta}, X_{t-}))_{t\geq 0}$ and the U_i are i.i.d. random variables with density $f(\boldsymbol{\theta}, \cdot)$. Moreover, the jump U_{N_t} of $J^{\boldsymbol{\theta}}$ at time t is supposed to be independent of $\mathcal{F}_{t-}$.

We suppose, for all $\boldsymbol{\theta} \in \Theta$, the functions $\mu(\boldsymbol{\theta}, \cdot)$, $\sigma(\cdot)$, $\lambda(\boldsymbol{\theta}, \cdot)$, and $f(\boldsymbol{\theta}, \cdot)$ satisfy the conditions of the existence of a unique solution to (1) and the necessary conditions to ensure that $(X_t)_{t\geq 0}$ is ergodic. On the conditions of the ergodicity of diffusion processes with jumps, see Kwon and Lee (1999). In the following ξ will be a random variable whose distribution, under $\mathbb{P}_{\boldsymbol{\theta}}$, is that of the invariant measure associated with the ergodic process X. Expectations under $\mathbb{P}_{\boldsymbol{\theta}}$ are denoted by $\mathbb{E}_{\boldsymbol{\theta}}$. Finally, we denote by $\mathbb{P}_{\boldsymbol{\theta}}^{(T)}$ the probability measure induced by the process $X^T = \{X_t, 0 \leq t \leq T\}$, under $\mathbb{P}_{\boldsymbol{\theta}}$, in the measurable space of right-continuous functions on $[0, T]$.

In this section, we consider the problem of inference about the parameter $\boldsymbol{\theta}$ based on an observation X^T where the sample path is observed in continuous time. In particular, we establish Local Asymptotic Normality (LAN) for the family of probability measures $\left(\mathbb{P}_{\boldsymbol{\theta}}^{(T)}, \boldsymbol{\theta} \in \Theta\right)$.

Definition 1. *A family of measures* $\left(\mathbb{P}_{\boldsymbol{\theta}}^{(T)}, \boldsymbol{\theta} \in \Theta\right)$ *is called Locally Asymptotically Normal (LAN) at* *a point* $\boldsymbol{\theta} \in \Theta$ *if, for any* $\boldsymbol{h} \in \mathbb{R}^d$*, the likelihood ratio*

$$L_T(\boldsymbol{h}) = \frac{\mathrm{d}\mathbb{P}^{(T)}_{\boldsymbol{\theta}+\frac{\boldsymbol{h}}{\sqrt{T}};X^T}}{\mathrm{d}\mathbb{P}^{(T)}_{\boldsymbol{\theta};X^T}},$$

admits the representation

$$L_T(\boldsymbol{h}) = \exp\left[\boldsymbol{h}^*\,\Delta_T(\boldsymbol{\theta},X^T) - \frac{1}{2}\boldsymbol{h}^* I(\boldsymbol{\theta})\,\boldsymbol{h} + r_T(\boldsymbol{\theta},\boldsymbol{h},X^T)\right], \tag{2}$$

where, under $\mathbb{P}^{(T)}_{\boldsymbol{\theta}}$*,*

$$\Delta_T(\boldsymbol{\theta},X^T) \xrightarrow{\mathcal{L}} \mathcal{N}(0,I(\boldsymbol{\theta})), \quad r_T(\boldsymbol{\theta},\boldsymbol{h},X^T) = o_{\mathbb{P}_{\boldsymbol{\theta}}}(1).$$

We say that the family $\left(\mathbb{P}^{(T)}_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\right)$ *is LAN if it is LAN at every point* $\boldsymbol{\theta} \in \Theta$*.*

Note that in our model, given $\boldsymbol{\theta}$, W^T, $N^{\boldsymbol{\theta},T}$, and U_i, for $i = 1,\dots,N_T^{\boldsymbol{\theta},T}$, are in the σ-algebra generated by X^T.

Proposition 2. *Suppose that* $\mu(\boldsymbol{\theta},\cdot)/\sigma(\cdot)$*,* $\lambda(\boldsymbol{\theta},\cdot)$*, and* $f(\boldsymbol{\theta},\cdot)$ *satisfy suitable smoothness assumptions. Then the family of measures* $\left(\mathbb{P}^{(T)}_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\right)$ *is LAN for each* $\boldsymbol{\theta} \in \Theta$ *with central sequence,*

$$\begin{aligned}\Delta_T(\boldsymbol{\theta},X^T) = {} & \int_0^T \frac{\nabla_{\boldsymbol{\theta}}\mu(\boldsymbol{\theta},X_{u-})}{\sigma^2(X_{u-})}\,\mathrm{d}W_u + \int_0^T \frac{\nabla_{\boldsymbol{\theta}}\lambda(\boldsymbol{\theta},X_{u-})}{\lambda(\boldsymbol{\theta},X_{u-})}\left(\mathrm{d}N_u^{\boldsymbol{\theta}} - \lambda(\boldsymbol{\theta},X_{u-})\,\mathrm{d}u\right)\\ & + \sum_{i=1}^{N_T^{\boldsymbol{\theta}}} \frac{\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta},U_i)}{f(\boldsymbol{\theta},U_i)}\end{aligned} \tag{3}$$

and Fisher information matrix

$$I(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}}\left[\frac{\nabla_{\boldsymbol{\theta}}\mu(\boldsymbol{\theta},\xi)(\nabla_{\boldsymbol{\theta}}\mu(\boldsymbol{\theta},\xi))^*}{\sigma^2(\xi)} + \frac{\nabla_{\boldsymbol{\theta}}\lambda(\boldsymbol{\theta},\xi)(\nabla_{\boldsymbol{\theta}}\lambda(\boldsymbol{\theta},\xi))^*}{\lambda(\boldsymbol{\theta},\xi)} + \lambda(\boldsymbol{\theta},\xi)\frac{\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta},U_1)\,(\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta},U_1))^*}{(f(\boldsymbol{\theta},U_1))^2}\right]. \tag{4}$$

LAN for discrete-time observations

Jump-diffusion models of the type (1) are widely used in applications. Examples are the modeling of security prices in financial markets, risks in insurance, etc. In practise, often, a continuous-time observation X^T is not available but the process is recorded in discrete time. More precisely, assume that we observe the process X at times t_i^n, $i = 1,\dots,n$, only. Assuming an appropriate high-frequency sampling scheme, to be defined precisely below, we show that, from an asymptotic and local point-of-view, the discrete-time observations $X_{t_i^n}$ contain as much information as the continuous sample path X^T.

We assume that the sampling scheme is regularly spaced and high frequency. This means that we observe the process X at time points $t_i^n = ih_n$, for $i = 0,\dots,n$, where h_n is the length of the observational intervals, where we assume $nh_n \to \infty$ and $nh_n^2 \to 0$ as $n \to \infty$. In order to prove that the limit experiment in the case of discrete-time observations equals that of continuous-time observations, we show that the central sequence (3) in Proposition 2 can be obtained, up to $o_p(1)$-terms, based on the discrete-time observations $X_{t_0^n},\dots,X_{t_n^n}$ only. In particular this implies that the deficiency of the continuous-time observations experiments with respect to the discrete-time observation experiments tends to zero.

We prove that using high frequency data we can build a central sequence that differs from the continuous central sequence (3) in terms of order $o_{\mathbb{P}}(1)$. In this respect, we need to recall the concept of Predictably Uniformly Tight (P-UT). To have an insight into P-UT see Jacod and Shiryaev (2002) Section VI.6.

Theorem 3. *The statistical information in* $(X_{t_i^n})_{0\le i\le n}$ *for inference about* $\boldsymbol{\theta}$ *in the model* (1) *equals, as* $n\to\infty$*, that in* X^{nh_n}*. To be precise the* $(X_{t_i^n})_{0\le i\le n}$*-measurable sequence*

$$\frac{1}{\sqrt{nh_n}}\int_0^{nh_n}\frac{\nabla\mu(\boldsymbol{\theta},X_u^n)}{\sigma(X_u^n)}\,\mathrm{d}X_u^n-\frac{1}{\sqrt{nh_n}}\int_0^{nh_n}\left(\frac{\nabla\mu(\boldsymbol{\theta},X_u^n)}{\sigma(X_u^n)}\mu(\boldsymbol{\theta},X_u^n)+\nabla\lambda(\boldsymbol{\theta},X_u^n)\right)\mathrm{d}u$$

$$(5)\qquad -\frac{1}{\sqrt{nh_n}}\int_0^{nh_n}\left(\frac{\nabla\mu(\boldsymbol{\theta},X_u^n)}{\sigma(X_u^n)}\Delta X_u^n+\frac{\nabla\lambda(\boldsymbol{\theta},X_u^n)}{\lambda(\boldsymbol{\theta},X_u^n)}+\frac{\nabla f(\boldsymbol{\theta},\Delta X_u^n)}{f(\boldsymbol{\theta},\Delta X_u^n)}\right)\mathrm{d}N_u^{\boldsymbol{\theta},n},$$

is a central sequence for $(\mathbb{P}_{\boldsymbol{\theta}}^{(T)})$*, where the sequence* (X_t^n) *is defined by:*

$$(6)\qquad X_t^n=X_{t_i^n}\quad t\in[t_i^n,t_{i+1}^n).$$

and $N^{\boldsymbol{\theta},n}$ *is a sequence of semimartingales Predictably Uniformly Tight (P-UT) and such that converges in probability to* $N^{\boldsymbol{\theta}}$*.*

More formally, that means that (5) *differs from* (3) *in terms of order* $o_{\mathbb{P}_{\boldsymbol{\theta}}^{(T)}}(1)$

Moreover, using a jump identification mechanism proposed by Shimizu and Yoshida (2006), we are able to show a possible choice for $(N_t^{\boldsymbol{\theta},n})_{t\ge0}$ that fulfills the requirement of Theorem 3.

REFERENCES

M. Akritas and R. Johnson (1981) Asymptotic inference in Lévy processes of disconinuous type, *Ann. Statist.*, 9, 604–614.

I. Basawa and B. Prakasa Rao (1980) *Statistical Inference for Stochastic Processes*, Probability and Mathematical Statistics, Academic Press, Inc., London-New York.

B. Bibby and M. Sørensen (1995) Martingale estimation functions for discretely observed diffusion processes, *Bernoulli*, 1, 17-39.

D. Dacunha-Castelle and D. Florens-Zmirou (1986) Estimation of the coefficient of a diffusion from discrete observations, *Stochastics*, 19, 263–284.

J. Doob (1953) *Stochastic Processes.* John Wiley, New York.

V. Genon-Catalot and J. Jacod (1993) On the estimation of the diffusion coefficient for multidimensional diffusion processes, *Ann. Inst. Henri Poincaré Prob. Statist.*, 29, 119–151.

V. Genon-Catalot and J. Jacod (1994) On the estimation of diffusion coefficient for diffusion processes, *Scand. J. Statist.* 21, 193–221.

J. Jacod and A. Shiryaev (2002) Limit Theorems for Stochastic Processes. Springer.

Y. Kutoyants (2003)*Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag, New York-Berlin.

Y. Kwon and C. Lee (1999) Strong feller property and irreducibility of diffusions with jumps, *Stoch. Stoch. Rep.* 67, 147–157

R. Liptser and A. Shiryaev (1978). Statistics of Random Processes II: Applications. Springer-Verlag, Berlin

B. Prakasa-Rao (1983) Asymptotic theory for non-linear least square estimator for diffusion processes, *Math. Oper. forsch. Statist. Ser. Statist.* 14, 195–209.

Y. Shimizu and N.Yoshida (2006) Estimation of parameters for diffusion processes with jumps form discrete observations,*Stat. Inference Stoch. Proc.* 9, 227–277.

M. Sørensen (1991) Likelihood methods for diffusions with jumps, *Stat. Inference Stoch. Proc.* 67–105

Y. Wang (2002) Asymptotic nonequivalence of GARCH models and diffusions, *Ann. Statist.*. 30 754–783.

ABSTRACT

We establish Local Asymptotic Normality (LAN) for continuous observations from jump-diffusion models with time-varying drift and jump intensity, but known volatility. We show that discrete-time high-frequency observations from the same model contain, in an asymptotic and local sense, the same information about the parameters of interest. More precisely, we provide sufficient conditions on a jump identification mechanism that allows us to construct a central sequence for the continuous-time model, using discrete-time observations only.