LAN property for diffusion processes with jumps with discrete observations

Arturo Kohatsu-Higa

Ritsumeikan University, Japan

joint work with Tran Ngoc Khue (University of Paris 13)
& Eulalia Nualart (University Pompeu Fabra, Barcelona)

Tokyo Workshop

Tokyo, 3 September, 2013

Outline

1. Introduction to the LAMN and LAN property
2. LAN property for a linear model with jumps
3. LAN property for a diffusion process with jumps

Parametric statistical model

Consider a parametric statistical model $(\mathfrak{X}^n, \mathcal{B}(\mathfrak{X}^n), \{\mathrm{P}^n_\theta, \theta \in \Theta\})$:

- A probability space $(\Omega, \mathcal{F}, \mathrm{P})$,
- A parameter space Θ : a closed rectangle of $\mathbb{R}^k$, for some integer $k \geq 1$,
- Random vector

$$\begin{aligned} X^n = (X_1, X_2, \ldots, X_n) : \Omega \times \Theta &\to \mathfrak{X}^n \subset \mathbb{R}^n \\ (\omega, \theta) &\mapsto X^n(\omega, \theta), \end{aligned}$$

- $\mathcal{B}(\mathfrak{X}^n)$: Borel σ-algebra of observable events,
- P^n_θ : probability measure on $(\mathfrak{X}^n, \mathcal{B}(\mathfrak{X}^n))$ induced by X^n under θ.
- Suppose that X^n has a density $p_n(x; \theta), x \in \mathbb{R}^n$ for all $\theta \in \Theta$.
- An estimator $T : \mathfrak{X}^n \to \Theta : x \mapsto T(x)$.

Motivation : LAMN and LAN property

- Interest of parametric estimation based on continuous-time and discrete-time observations of diffusion processes with jumps.
- Our objectives : solve the problem of asymptotic efficiency of the estimators.
- More precisely, this problem is closely linked to the LAMN and LAN property.
- Efficiency of an unbiased estimator : its variance achieves the Cramér-Rao lower bound in the Cramér-Rao's inequality.

Cramér-Rao's inequality via Malliavin calculus

Consider a random vector $X^{(n)} = (X_1, \dots, X_n)$.

Theorem (Corcuera and Kohatsu-Higa'11)

- Suppose that $X_i \in \mathbb{D}^{1,2}$ and there exists a stochastic process $u \in \mathrm{Dom}(\delta)$ such that

$$\int_0^T D_t X_i u(t) dt = \partial_\theta X_i \quad \text{for all } i = 1, \dots, n. \tag{1}$$

- Let T be an unbiased estimator of θ.
- Under regularity hypothesis on the parametric statistical model :

$$Var_\theta(T(X^{(n)})) \geq \frac{1}{Var_\theta(\mathrm{E}_\theta[\delta(u)|X^{(n)}])}. \tag{2}$$

- Furthermore, if $X^{(n)}$ admits a density $p_n(x; \theta)$ then

$$\mathrm{E}_\theta(\delta(u)|X^{(n)} = x) = \partial_\theta \log p_n(x; \theta). \tag{3}$$

Same formula explained in other form

An IBP formula (in ∞-dimensions) is for any $f \in C_b^\infty$ and any $G \in L^1(\Omega)$ there exists a $L^1(\Omega)$ random variable $H(F, G)$ such that

$$\mathrm{E}[f'(F)G] = \mathrm{E}[f(F)H(F,G)]$$

These formulas have been found for various (∞-dimensional) stochastic differential equations (with jumps, or even with correlation structure in the driving noise).
Then if $T : C[0,1] \to \mathbb{R}$ is a unbiased estimator of a parameter μ then

$$\begin{aligned} 1 &= \partial_\mu \mathrm{E}[T(X)] = \mathrm{E}[< T'(X), \partial_\mu X >] \\ &= \mathrm{E}[T(X)H(X, \partial_\mu X)] \leq V(T(X))V(H(X, \partial_\mu X)). \end{aligned}$$

Therefore the Cramer-Rao bound will follow. Interestingly controlling, estimating and approximating $H(X, \partial_\mu X)$ is a matter that it is well studied in this area. This as explained before is clearly related to the logarithmic derivatives of the density.

Definition : LAMN and LAN property

Let Θ be a closed rectangle of $\mathbb{R}^k$, for some integer $k \geq 1$.

1. Definition. The sequence of $(\mathfrak{X}^n, \mathcal{B}(\mathfrak{X}^n), \{\mathrm{P}^n_\theta, \theta \in \Theta\})$ has the local asymptotic mixed normality (LAMN) property at θ if there exist positive definite $k \times k$ matrix $\varphi_n(\theta)$ satisfying that $\varphi_n(\theta) \to 0$ as $n \to \infty$, and $k \times k$ symmetric positive definite random matrix $\Gamma(\theta)$: for any $u \in \mathbb{R}^k$, $n \to \infty$,

$$\log \frac{d\mathrm{P}^n_{\theta+\varphi_n(\theta)u}}{d\mathrm{P}^n_\theta}(X^n) \xrightarrow{\mathcal{L}(\mathrm{P}_\theta)} u^{\mathsf{T}}\Gamma(\theta)^{1/2}\mathcal{N}(0, I_k) - \frac{1}{2}u^{\mathsf{T}}\Gamma(\theta)u, \tag{4}$$

   $\mathcal{N}(0, I_k)$: a centered $\mathbb{R}^k$-valued Gaussian variable, independent of $\Gamma(\theta)$.
   $\Gamma(\theta)$: asymptotic Fisher information matrix.

2. When $\Gamma(\theta)$ is deterministic, we have the local asymptotic normality (LAN) property at θ.

Consequences of the LAMN property

- Reduce to study the convergence in law under P_θ of log-likelihood ratio :

$$\log \frac{d\mathrm{P}^n_{\theta+\varphi_n(\theta)u}}{d\mathrm{P}^n_\theta}(X^n) = \log \frac{p_n(X^n;\theta+\varphi_n(\theta)u)}{p_n(X^n;\theta)}.$$

- Conditional convolution theorem : Suppose that the LAMN property holds at θ. If $(\tilde{\theta}_n)_{n\geq 1}$ is a regular sequence of estimators of θ : $\forall u \in \mathbb{R}^k$,

$$\varphi_n(\theta)^{-1}\left(\tilde{\theta}_n - (\theta+\varphi_n(\theta)u)\right) \overset{\mathcal{L}(\mathrm{P}_{\theta+\varphi_n(\theta)u})}{\underset{n\to\infty}{\longrightarrow}} V(\theta),$$

  for some $\mathbb{R}^k$-valued r.v $V(\theta)$, then $\mathcal{L}\left(V(\theta)|\Gamma(\theta)\right) = \mathcal{N}\left(0,\Gamma(\theta)^{-1}\right) \star G_{\Gamma(\theta)}$.

- Therefore, $(\tilde{\theta}_n)_{n\geq 1}$ is called asymptotically efficient if as $n\to\infty$,

$$\varphi_n^{-1}(\theta)\left(\tilde{\theta}_n - \theta\right) \overset{\mathcal{L}(\mathrm{P}_\theta)}{\longrightarrow} \Gamma(\theta)^{-1/2}\mathcal{N}(0,I_k). \tag{5}$$

- Minimax theorem implies that $\Gamma(\theta)^{-1}$ gives the lower bound for the asymptotic variance of estimators.

Several references

1. Gobet'01 derives the LAMN property in the non-ergodic case :

$$X_t^\theta = X_0 + \int_0^t b(\theta, s, X_s^\theta)ds + \int_0^t \sigma(\theta, s, X_s^\theta)dB_s, \quad t \in [0, 1]. \qquad (6)$$

2. Gobet'02 shows the LAN property in the ergodic case :

$$X_t^{\alpha,\beta} = X_0 + \int_0^t b(\alpha, X_s^{\alpha,\beta})ds + \int_0^t \sigma(\beta, X_s^{\alpha,\beta})dB_s, \quad t \geq 0. \qquad (7)$$

3. Delattre and *al.*'11 have established the LAMN property :

$$X_t^\lambda = X_0 + \int_0^t b(s, X_s^\lambda)ds + \int_0^t a(s, X_s^\lambda)dB_s + \sum_{k:T_k \leq t} c(X_{T_k-}^\lambda, \lambda_k), \qquad (8)$$

for $t \in [0, 1]$, where the jump times $T_1, T_2, \ldots, T_K$ are given.

For the proof of these results, they use tools of Malliavin calculus and upper and lower Gaussian type estimates of the transition densities of the diffusion processes. In the jump case, these estimates are not satisfied !

The density estimates

One has the tendency to believe that "good" upper and lower estimates of the density should essentially solve the problem.
We started computing some of these estimates and they do not work. In fact, even in simpler one dimensional situations (Gaussian type jumps) the upper density estimate is of the type

$$\frac{C}{\sqrt{t}} \exp\left(-c|y-x|\sqrt{|\ln \frac{|y-x|}{t}|}\right).$$

The lower density estimates are of the type for $x \neq y$ with a different estimate over the diagonal.

$$Ce^{-\lambda t} \exp\left(-c|y-x|\sqrt{|\ln(\frac{|y-x|}{t})|}\right)$$

The last result shows that in general the estimate for large/small $|y-x|$ should be different. This is a first negative result !

A linear model with jumps

$$X_t = x + \theta t + B_t + N_t - \lambda t. \quad (9)$$

- $B = (B_t, t \geq 0)$ is a standard Brownian motion,
- $N = (N_t, t \geq 0)$ is a Poisson process with intensity $\lambda > 0$.
- $(\theta, \lambda) \in \Theta \times \Lambda \subset \mathbb{R} \times \mathbb{R}_+^*$ are unknown parameters to be estimated.
- High frequency observation $X^n = (X_{t_0}, X_{t_1}, ..., X_{t_n})$, where $t_k = k\Delta_n$:
  - Number of observations $n \to \infty$,
  - Distance between observations $\Delta_n \to 0$,
  - Horizon $n\Delta_n \to \infty$.
- X^n admits a density $p_n(x; (\theta, \lambda))$.
- $p^{(\theta,\lambda)}(t, \cdot, \cdot)$: transition density of X_t conditionally on X_0 under (θ, λ).

Theorem

For all $(\theta, \lambda) \in \Theta \times \Lambda$ *and* $(u, v) \in \mathbb{R}^2$*, as* $n \to \infty$*,*

$$\log \frac{p_n\left(X^n; \left(\theta + \frac{u}{\sqrt{n\Delta_n}}, \lambda + \frac{v}{\sqrt{n\Delta_n}}\right)\right)}{p_n(X^n; (\theta, \lambda))}$$

$$\xrightarrow{\mathcal{L}(\mathrm{P}_{(\theta,\lambda)})} \begin{pmatrix} u \\ v \end{pmatrix}^{\mathsf{T}} \mathcal{N}(0, \Gamma(\theta, \lambda)) - \frac{1}{2} \begin{pmatrix} u \\ v \end{pmatrix}^{\mathsf{T}} \Gamma(\theta, \lambda) \begin{pmatrix} u \\ v \end{pmatrix},$$

where $\mathcal{N}(0, \Gamma(\theta, \lambda))$ *is a centered* $\mathbb{R}^2$*-valued Gaussian variable with covariance matrix*

$$\Gamma(\theta, \lambda) = \begin{pmatrix} 1 & -1 \\ -1 & 1 + \frac{1}{\lambda} \end{pmatrix}.$$

For simplicity let $(\theta_0(n), \lambda_0(n)) := \left(\theta + \frac{u}{\sqrt{n\Delta_n}}, \lambda + \frac{v}{\sqrt{n\Delta_n}}\right)$.

Sketch of the proof

- **Step 1.** By Markov property,

$$
\begin{aligned}
&\log \frac{p_n\left(X^n;(\theta_0(n),\lambda_0(n))\right)}{p_n(X^n;(\theta,\lambda))}\\
&= \sum_{k=0}^{n-1} \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\theta p^{\left(\theta_0(n,\ell),\lambda\right)}}{p^{\left(\theta_0(n,\ell),\lambda\right)}}(\Delta_n, X_{t_k}, X_{t_{k+1}})d\ell\\
&\quad + \sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n}} \int_0^1 \frac{\partial_\lambda p^{(\theta_0(n),\lambda_0(n,\ell))}}{p^{(\theta_0(n),\lambda_0(n,\ell))}}(\Delta_n, X_{t_k}, X_{t_{k+1}})d\ell.
\end{aligned}
$$

$\theta_0(n,\ell) := \theta_0 + \frac{\ell u}{\sqrt{n\Delta_n}}$, $\lambda_0(n,\ell) := \lambda + \frac{\ell v}{\sqrt{n\Delta_n}}$.

Sketch of the proof (2)

- **Step 2.** Use the integration by parts formula of the Malliavin calculus,

$$\log \frac{p_n(X^n;(\theta_0(n),\lambda_0(n)))}{p_n(X^n;(\theta,\lambda))} = u\frac{B_{n\Delta_n}}{\sqrt{n\Delta_n}} - v\frac{B_{n\Delta_n}}{\sqrt{n\Delta_n}} - \frac{u^2}{2} - \frac{v^2}{2} + uv + \sum_{k=0}^{n-1} R_{k,n}$$

$$+ \sum_{k=0}^{n-1} \frac{v}{\sqrt{n\Delta_n}} \int_0^1 M(\theta_0(n),\lambda_0(n,\ell),X_{t_k}.X_{t_{k+1}})d\ell.$$

$$M(\theta_0(n),\lambda_0(n,\ell),X_{t_k}.X_{t_{k+1}}) := \mathrm{E}_{X_{t_k}}^{(\theta_0(n),\lambda_0(n,\ell))}\left[\frac{\tilde{N}_{t_{k+1}} - \tilde{N}_{t_k}}{\lambda_0(n,\ell)}\Big| X_{t_{k+1}}^{(\theta_0(n),\lambda_0(n,\ell))} = X_{t_{k+1}}\right].$$

- **Step 3.** Finally, apply the central limit theorem for triangular arrays of random variables to show the convergence in law.

- In order to do this let us consider the main term which is the variance. Note the term $\sqrt{n\Delta_n}^{-2}$. In that case, the important residue term is

$$\mathrm{E}_{X_{t_k}}^{(\theta,\lambda)}\left[\left(M(\theta_0(n),\lambda_0(n,\ell),X_{t_k}.X_{t_{k+1}}) - \mathrm{E}_{X_{t_k}}^{(\theta,\lambda)}\left[M(\theta_0(n),\lambda_0(n,\ell),X_{t_k}.X_{t_{k+1}})\right]\right)^2\right]$$

- In order to show that $\sum_{k=0}^{n-1} R_{k,n}$ converges to zero in probability, we use the decomposition method of jumps by considering the events J_i which count the number of jumps that have occured in the interval $[t_k, t_{k+1}]$:

$$J_0 = \{N_{t_{k+1}} - N_{t_k} = 0\},$$
$$J_1 = \{N_{t_{k+1}} - N_{t_k} = 1\},$$
$$J_2 = \{N_{t_{k+1}} - N_{t_k} \geq 2\}.$$

- Therefore in the previous double expectation we have 9 different cases to treat. That is, for $i, j \in \{0, 1, 2\}$ consider

$$\mathrm{E}_{X_{t_k}}^{(\theta,\lambda)}\left[1_{J_i}\left(M(\theta_0(n), \lambda_0(n,\ell), X_{t_k}.X_{t_{k+1}}) - \mathrm{E}_{X_{t_k}}^{(\theta,\lambda)}\left[1_{J_j} M(\theta_0(n), \lambda_0(n,\ell), X_{t_k}.X_{t_{k+1}})\right]\right)^2\right]$$

The most difficult case is $j = 1$, $i = 1$. The other cases are handled due to different reasons ! This is why the argument in Gobet does not work exactly the same way !

In the $j = 1$, $i = 1$ case one needs to use $\mathrm{E}_{X_{t_k}}^{(\theta,\lambda)}\left[1_{J_j}\right] M(\theta_0(n), \lambda_0(n,\ell), X_{t_k}.X_{t_{k+1}})$ then use that $1_{J_j} = 1_{J_0} + 1_{J_2}$ and the “continuity” of M.

A diffusion process with jumps

Consider a diffusion process with jumps $X = (X_t)_{t\geq 0}$ satisfying

$$dX_t = b(\theta, X_t)dt + \sigma(X_t)dB_t + \int_{\mathbb{R}_0} c(X_{t-}, z)\left(N(dt, dz) - \nu(dz)dt\right) \quad (10)$$

- $\theta \in \Theta \subset \mathbb{R}$ determines the drift coefficient,
- $N(dt, dz)$ is a Poisson random measure with intensity measure $\nu(dz)dt$,
- High frequency observation $X^n = (X_{t_0}, X_{t_1}, ..., X_{t_n})$, where $t_k = k\Delta_n$,
- $p_n(\cdot; \theta)$: density of X^n under θ,
- $p^\theta(t, x, y)$: transition density of X_t conditionally on $X_0 = x$ under θ.
- Estimators for this setting have been proposed by Shimizu-Yoshida and Ogiwara-Yoshida. These estimators will essentially be optimal.
- The difficult part in the present analysis is the fact that proving optimality requires careful study of residual terms that do not need to be analyzed in the case of proving only asymptotic normality.

Assumptions on the model

- Regularity conditions on the coefficients and Lévy measure $\nu(dz)$.
- The drift coefficient $b(\theta, x)$ is uniformly bounded on $\Theta \times \mathbb{R}$.
- The jumps are bounded. (In order to apply convergence arguments. The general case will be treated in the near future).
- Ergodicity : there exists a unique invariant probability measure $\pi_\theta(dx)$:

$$\frac{1}{T}\int_0^T g(\theta, X_t)dt \xrightarrow{\mathrm{P}_\theta} \int_{\mathbb{R}} g(\theta, x)\pi_\theta(dx), \tag{11}$$

as $T \to \infty$, for any π_θ-integrable function g, uniformly in $\theta \in \Theta$.

LAN property

Theorem

Under regularity and ergodicity conditions, for all $\theta \in \Theta$ and $u \in \mathbb{R}$, as $n \to \infty$,

$$\log \frac{p_n(X^n; \theta_0(n))}{p_n(X^n; \theta)} \xrightarrow{\mathcal{L}(\mathrm{P}_\theta)} u\mathcal{N}\left(0, \Gamma(\theta)\right) - \frac{u^2}{2}\Gamma\left(\theta\right),$$

where $\Gamma\left(\theta\right)$ is given by

$$\Gamma\left(\theta\right) = \int_{\mathbb{R}} \left(\frac{\partial_\theta b(\theta, x)}{\sigma(x)} \right)^2 \pi_\theta(dx).$$

Sketch of the proof

- **Step 1.** Use tools of Malliavin calculus,

$$\log \frac{p(X^n;\theta_0(n))}{p(X^n;\theta)} = \sum_{k=0}^{n-1} \xi_{k,n},$$

where

$$\xi_{k,n} := \frac{u}{\sqrt{n\Delta_n}} \int_0^1 \frac{1}{\Delta_n} \mathrm{E}_{X_{t_k}}^{\theta_0(n,\ell)} \left[\delta \left(\partial_\theta X_{t_{k+1}}^{\theta_0(n,\ell)} U \right) \middle| X_{t_{k+1}}^{\theta_0(n,\ell)} = X_{t_{k+1}} \right] d\ell.$$

Here, δ denotes Skorohod integral of the Brownian motion and

$$U(t) = \frac{1}{D_t X_{t_{k+1}}^{\theta_0(n,\ell)}} = \sigma^{-1}\left(X_t^{\theta_0(n,\ell)}\right) \partial_x X_t^{\theta_0(n,\ell)} \left(\partial_x X_{t_{k+1}}^{\theta_0(n,\ell)} \right)^{-1}.$$

Sketch of the proof (2)

- **Step 2.** Apply the central limit theorem for triangular arrays of random variables,

$$\sum_{k=0}^{n-1} \mathrm{E}_\theta \left[\xi_{k,n} | \mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}_\theta} -\frac{u^2}{2}\Gamma(\theta),$$

$$\sum_{k=0}^{n-1} \left(\mathrm{E}_\theta \left[\xi_{k,n}^2 | \mathcal{F}_{t_k}\right] - \left(\mathrm{E}_\theta \left[\xi_{k,n} | \mathcal{F}_{t_k}\right]\right)^2\right) \xrightarrow{\mathrm{P}_\theta} u^2\Gamma(\theta),$$

$$\sum_{k=0}^{n-1} \mathrm{E}_\theta \left[\xi_{k,n}^4 | \mathcal{F}_{t_k}\right] \xrightarrow{\mathrm{P}_\theta} 0.$$

The more demanding proof is the last one. Taking into account two problems : 1. Recall that $\mathrm{E}_\theta[N_t^k] = O(t)$ for any $k \geq 1$ and that usual asymptotic expansions are not so easy to use. 2. The upper and lower bounds for densities are important here as one needs to change measures from $\theta + \frac{u}{n\Delta_n}$ to θ. Therefore just conditioned integration by parts formulas will not suffice.

The general question is : In proving these type of LAN properties any upper and lower bound are enough ?

The general question is : In proving these type of LAN properties any upper and lower bound are enough ? In Gobet (2001, 2002) essentially upper and lower Gaussian type bounds are used.
In our case, upper and lower bounds are not the same. Still, by conditioning on the number of jumps then within each class comparisons can be made efficiently.
We find that the analysis has to be divided in cases. In the no-jump case and when $|X_{t_k} - y| \geq \Delta_n^{1/2-\epsilon}$ for $\epsilon > 0$ small enough, the lower bound is

$$p^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y)^2 \geq \left(\int_{\mathbb{R}_0} q_{(1)}^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y; a)\mu(da) e^{-\lambda\Delta_n} \lambda\Delta_n \right)^2$$

For the opposite case $|X_{t_k} - y| \leq \Delta_n^{1/2-\epsilon}$ one uses

$$p^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y)^2 \geq q^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y) e^{-\lambda\Delta_n} \int_{\mathbb{R}_0} q_{(1)}^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y; a)\mu(da) e^{-\lambda\Delta_n} \lambda\Delta_n.$$

Here μ is the jump distribution, q is the Gaussian type density. On the other hand given the structure of the problem once we know that there is one jump specific Gaussian upper bounds can also be obtained.
In the one jump case wit $|X_{t_k} - y| \leq \Delta_n^{1/2-\epsilon}$, one has to use

$$p^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y) \geq q^{\theta_0(n,\ell)}(\Delta_n, X_{t_k}, y) e^{-\lambda\Delta_n}$$

Therefore this becomes like a stratification method.

In short...

- We apply the Girsanov's theorem and the discrete-time ergodic theorem to get that as $n \to \infty$,

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{\partial_\theta b(\theta, X_{t_k})}{\sigma(X_{t_k})}\right)^2 \xrightarrow{\mathrm{P}_\theta} \int_{\mathbb{R}}\left(\frac{\partial_\theta b(\theta, x)}{\sigma(x)}\right)^2 \pi_\theta(dx).$$

- We use the decomposition method of jumps and the estimates of the transition densities of a diffusion process without jumps.

Work in progress

1. A class of ergodic SDE with jumps with unbounded drift coefficient. Requires a specific bound for densities of corresponding continuous SDEs
2. The case of unbounded jumps requires a convergence argument. Some conditions will naturally arise.
3. The case where θ determines the drift and jump coefficients :
$$dX_t = b(\theta, X_t)dt + \sigma(X_t)dB_t + \int_{\mathbb{R}_0} c(\theta, X_{t-}, z)\,(N(dt, dz) - \nu_\theta(dz)dt)\,.$$
4. The case when $\sigma(\beta, X_t)$ depends again on Gaussian estimates of densities. But the measures are strongly singular between them (therefore we have to stop thinking of Girsanov's theorem). At the same time, variances can be compared (therefore an analytic approach is needed). Here again a stratification in needed. Not done yet.

References

1. Clément, E., Delattre, S. and Gloter, A. (2013), Asymptotic lower bounds in estimating jumps, *to appear in Bernoulli*.
2. Gobet, E. (2001), Local asymptotic mixed normality property for elliptic diffusions : a Malliavin calculus approach, *Bernoulli*, **7**, 899-912.
3. Gobet, E. (2002), LAN property for ergodic diffusions with discrete observations, *Ann. I. H. Poincaré*, **38**, 711-737.
4. Ogiwara T. and Yoshida N. Quasi-likelihood analysis for the stochastic differential equation with jumps. Stat Inference Stoch Process (2011) 14 :189-229
5. Shimizu Y. and Yoshida N. Estimation of Parameters for Diffusion Processes with Jumps from Discrete Observations . Statistical Inference for Stochastic Processes (2006) 9 :227-277.

Thank you for your attention !