Optimal estimation of certain random quantities

Mark Podolskij

joint work with J. Ivanovs

Aarhus University, Denmark

Risk and Statistics, Ulm

Topic of the talk

Let $(X_t)_{t\in[0,1]}$ be a stochastic process (Brownian motion, Lévy process, SDE etc.). Given the observations

$$X_0, X_{\Delta_n}, X_{2\Delta_n}, \ldots, X_{\lfloor 1/\Delta_n \rfloor \Delta_n} \qquad \text{with } \Delta_n \to 0$$

and the **random** parameter of interest Q, what is the **optimal** estimator of Q?

Low vs. high frequency data

Low frequency data	*High frequency data*
Observed data	***Observed data***
$X_1, X_2, ..., X_n \;\; i.i.d. \sim F$	$X_0(\omega), X_{\Delta_n}(\omega), ..., X_{\lfloor 1/\Delta_n \rfloor \Delta_n}(\omega)$
Asymptotic knowledge	***Asymptotic knowledge***
distribution function F	$(X_t(\omega))_{t \in [0,1]}$
Identifiable objects	***Identifiable objects***
functionals of F	functionals of $(X_t(\omega))_{t \in [0,1]}$

Background

- In the classical test theory the model parameters are **deterministic** objects. There exist numerous approaches to access the optimality of estimators: Cramer-Rao bounds, maximum likelihood theory, minimax approach, Le Cam theory, etc.
- However, in the high frequency setting the objects of interests are often **random**. Examples include quadratic variation, realised jumps, supremum/infimum of a process, local times, occupation time measures etc.
- In this framework very little is known about how to construct optimal estimates.

Let X be a continuous semimartingale of the form

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s \qquad t \geq 0$$

where a and σ are stochastic processes, and W is a Brownian motion. An important result in the theory of high frequency data is the following theorem.

Theorem (Jacod(94))

It holds that

$$\Delta_n^{-1/2} \left(\sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} \left(X_{i\Delta_n} - X_{(i-1)\Delta_n} \right)^2 - \int_0^1 \sigma_s^2 ds \right) \xrightarrow{d_{st}} \mathcal{MN}\left(0, 2\int_0^1 \sigma_s^4 ds \right)$$

Recently, Clement, Delattre & Gloter (13) have proved that the above estimator is **asymptotically efficient** applying an infinite dimensional LAMN property.

- The results of Clement, Delattre & Gloter (13) only cover estimation problems for volatility functionals. In this talk we will rather focus on the following random objects:

$$\overline{X} := \sup_{s\in[0,1]} X_s$$

$$l(x) := \lim_{\epsilon\downarrow 0} \frac{1}{2\epsilon} \int_0^1 1_{(-\epsilon,\epsilon)}(X_s - x)ds$$

$$L(x) := \int_0^1 1_{(x,\infty)}(X_s)ds$$

which is the supremum, local time and occupation time measure of the process X, respectively.

- We are interested in **optimal estimation** of these objects given high frequency data $(X_{i\Delta_n})_{0\leq i\leq\lfloor 1/\Delta_n\rfloor}$.

A remark on optimality

We will see that many naive estimators are rate optimal, but not efficient! In fact, efficient estimators are easy to introduce.

Let $Q = \Phi((X_s)_{s\in[0,1]})$ be a random variable of interest. An optimal estimator of Q is given as

(i) in L^2-sense: $\mathbb{E}[Q|\ (X_{i\Delta_n})_{0\le i\le \lfloor 1/\Delta_n \rfloor}]$

(ii) in L^1-sense: $\text{median}[Q|\ (X_{i\Delta_n})_{0\le i\le \lfloor 1/\Delta_n \rfloor}]$

We will investigate the asymptotic theory for these type of estimates in the setting of supremum, local time and occupation time measure of the process X, where X is a Brownian motion, stable Lévy process or a continuous diffusion process.

Naive estimator of the supremum

- It is rather simple to propose the following estimate for the supremum

$$M_n := \max_{i=1,\ldots,\lfloor 1/\Delta_n \rfloor} X_{i\Delta_n} \xrightarrow{\mathbb{P}} \overline{X}$$

  where the consistency holds for all Lévy processes X.

- The asymptotic theory for the maximum has been studied in several papers including Asmussen, Glynn & Pitman (95) (Brownian motion) and Ivanovs (18) (general Lévy processes).

- Since $M_n < \overline{X}$, the estimator M_n is downward biased and there were several attempts to correct the bias.

A result on zooming-in at supremum

The following result from the theory of Lévy processes will be extremely useful for our asymptotic theory.

Theorem (Ivanovs (18))

Let X be an α-stable Lévy process with $\alpha \in (0,2]$. Denote by τ the time of the supremum of X on the interval $[0,1]$. Then we obtain the functional stable convergence

$$(Z_t^n)_{t\in\mathbb{R}} := \left(\Delta_n^{-1/\alpha}\left(X_{\tau+t\Delta_n} - X_\tau\right)\right)_{t\in\mathbb{R}} \stackrel{d_{st}}{\to} \left(\widehat{X}_t\right)_{t\in\mathbb{R}}$$

*where $\widehat{X}$ is the so called **Lévy process conditioned to stay negative**, which is independent of $\mathcal{F}$. When X is a Brownian motion, we deduce the identity*

$$\widehat{X}_t = -\|B_t\|$$

where B is a 3-dimensional Brownian motion.

Application to estimation of the supremum

The previous result has the following consequence.

Theorem (Ivanovs (18))

Let X be an α-stable Lévy process with $\alpha \in (0,2]$. Then it holds that

$$\Delta_n^{-1/\alpha}\left(M_n - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U})$$

where $U \sim \mathcal{U}(0,1)$ is independent of $\widehat{X}$ and $\mathcal{F}$.

Sketch of proof: Note that

$$\Delta_n^{-1/\alpha}\left(X_{(\lceil\tau/\Delta_n\rceil+i)\Delta_n} - X_\tau\right) = Z^n_{i+\{\tau/\Delta_n\}}$$

Recall that $\{\tau/\Delta_n\} \xrightarrow{d_{st}} U \sim \mathcal{U}(0,1)$. Since $Z^n \xrightarrow{d_{st}} \widehat{X}$, we conclude that

$$\Delta_n^{-1/\alpha}\left(M_n - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U})$$

□

Computation of the optimal estimator: The Brownian case

The basis of our approach is the computation of the conditional probability

$$H_n(x) := \mathbb{P}\left(\overline{X} \leq x \mid (X_{i\Delta_n})_{0 \leq i \leq \lfloor 1/\Delta_n \rfloor}\right) \qquad x > 0.$$

Due to Markov and self-similarity property of X, we easily see that

$$H_n(x) = \prod_{i=1}^{n} F\left(\Delta_n^{-1/2}(x - X_{\frac{i-1}{n}}), \Delta_n^{-1/2}\Delta_i^n X\right)$$

where $F(x,y) = \mathbb{P}\left(\overline{X} \leq x \mid X_1 = y\right) = 1 - \exp(-2x(x-y))$. After rescaling we deduce the stable convergence

$$H_n\left(\Delta_n^{1/2}x + M_n\right) = \prod_{i \in \mathbb{Z}} F\left(x + \Delta_n^{-1/2}(M_n - X_{(i-1)\Delta_n}), \Delta_n^{-1/2}\Delta_i^n X\right)$$

$$\xrightarrow{d_{st}} G(x) := \prod_{i \in \mathbb{Z}} F\left(x + \max_{j \in \mathbb{Z}} \widehat{X}_{j+U} - \widehat{X}_{i+U}, \widehat{X}_{i+1+U} - \widehat{X}_{i+U}\right).$$

Conditional mean and conditional median

- For the conditional mean $T_n^{(2)} := \mathbb{E}\left[\overline{X} \middle| (X_{i\Delta_n})_i\right]$ we obtain the formula

$$T_n^{(2)} - \overline{X} = (M_n - \overline{X}) + \Delta_n^{1/2} \int_0^\infty \left(1 - H_n\left(\Delta_n^{1/2} x + M_n\right)\right) dx$$

Hence, the probabilistic structure of X only affects the second order term.

- Similarly, for the conditional median $T_n^{(1)} := \text{median}\left[\overline{X} \middle| (X_{i\Delta_n})_i\right]$ we deduce the identity

$$T_n^{(1)} - \overline{X} = (M_n - \overline{X}) + \Delta_n^{1/2} H_n\left(\Delta_n^{1/2} \cdot + M_n\right)^{-1}(1/2)$$

and again the probabilistic structure of X only affects the second order term.

Asymptotic theory for the optimal estimators: Brownian case

Theorem (Ivanovs & P. (19))

Define the estimates

$$T_n^{(1)} = median\left[\overline{X} \middle| (X_{i\Delta_n})_i\right], \qquad T_n^{(2)} = \mathbb{E}\left[\overline{X} \middle| (X_{i\Delta_n})_i\right].$$

(i) *It holds that*

$$\Delta_n^{-1/2}\left(T_n^{(1)} - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U}) + G^{-1}(1/2).$$

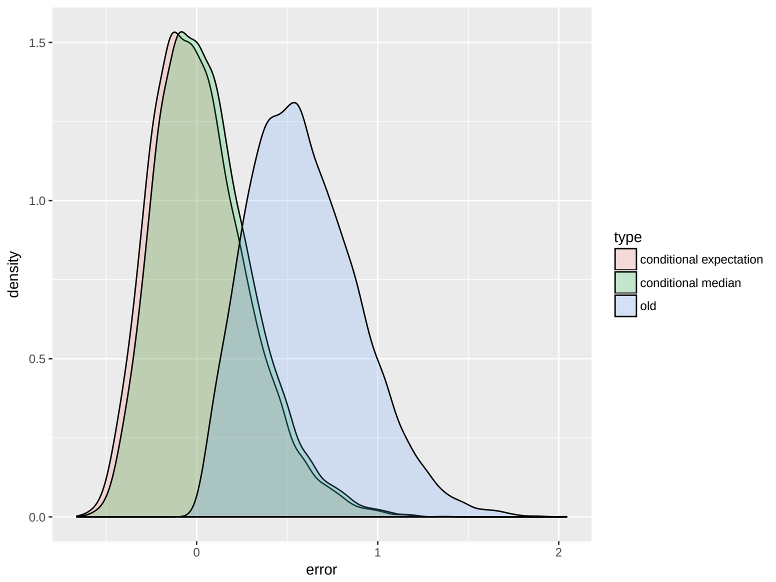
(ii) *Furthermore,*

$$\Delta_n^{-1/2}\left(T_n^{(2)} - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U}) + \int_0^\infty (1 - G(y))dy.$$

In particular, we have that

$$\frac{MSE(M_n)}{MSE(T_n^{(2)})} \approx 6.25\ !$$

Simulation of asymptotic distributions



Theorem (Ivanovs & P. (19))

Let X be a α-stable Lévy motion with $\alpha \in (0,2)$.

(i) *Define $T_n^{(1)} = median[\overline{X} | \ (X_{i\Delta_n})_i]$. Then we obtain*

$$\Delta_n^{-1/\alpha}\left(T_n^{(1)} - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U}) + G^{-1}(1/2).$$

and the estimator is L^1-optimal for $\alpha \in (1,2)$.

(ii) *Define $T_n^{(2)} = \mathbb{E}[\overline{X} | \ (X_{i\Delta_n})_i]$ for $\alpha \in (1,2)$. Then it holds that*

$$\Delta_n^{-1/\alpha}\left(T_n^{(2)} - \overline{X}\right) \xrightarrow{d} \max_{j\in\mathbb{Z}}(\widehat{X}_{j+U}) + \int_0^\infty (1 - G(y))dy.$$

- In this chapter we assume that X is a Brownian motion. Recall the definition of local time:

$$l(x) = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int_0^1 1_{(-\epsilon,\epsilon)}(X_s - x) ds$$

where $x \in \mathbb{R}$.

- A straightforward estimator of $l(x)$ is given as

$$l^n(x) := a_n \Delta_n \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} g\left(a_n(X_{i\Delta_n} - x)\right) \xrightarrow{\mathbb{P}} l(x)$$

where g is a kernel satisfying $\int_{\mathbb{R}} g(x)dx = 1$, and $a_n \to \infty$ with $a_n \Delta_n \to 0$.

- We will focus on a more general class of statistics:

$$V(h,x)^n := a_n \Delta_n \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} h\left(a_n(X_{i\Delta_n} - x), \Delta_n^{-1/2} \Delta_i^n X\right)$$

Theorem (Borodin (86), Jacod (98))

Assume that $a_n = \Delta_n^{-1/2}$ and h satisfies the condition $|h(y,z)| \leq h_1(y)\exp(\lambda|z|)$ for some $\lambda > 0$ and $\int_{\mathbb{R}} |y|^p h_1(y)dy < \infty$ for some $p > 3$. Then it holds that

$$V(h,x)^n \xrightarrow{\mathbb{P}} c_h l(x)$$

where $c_h = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} h(y,z)\varphi(z)dz\right) dy$ and φ denotes the density of the standard normal distribution. Furthermore, we obtain the stable convergence

$$\Delta_n^{-1/4}\left(V(h,x)^n - c_h l(x)\right) \xrightarrow{d_{st}} \mathcal{MN}(0, v_h l(x))$$

for a certain constant $v_h > 0$.

An interesting example is the **number of crossings at level 0** which corresponds to $x = 0$ and $h(y,z) = 1_{(-\infty,0)}(y(y+z))$.

As we mentioned earlier, the L^2-optimal estimator of the local time is given by

$$\widehat{l}^n(x) = \mathbb{E}\left[l(x) \middle| \left(X_{i\Delta_n}\right)_{1\le i\le \lfloor 1/\Delta_n \rfloor}\right]$$

The following distributional identity connects the law of local times to the law of the supremum:

$$\left(l_t(0), |X_t|\right)_{t\in\mathbb{R}} = \left(\overline{X}_t, \overline{X}_t - X_t\right)_{t\in\mathbb{R}}$$

Applying the Markov and self-similarity property of the Brownian motion we deduce that

$$\widehat{l}^n(x) = V(h_0, x)^n \qquad \text{with} \qquad a_n = \Delta_n^{-1/2}$$

and

$$h_0(y,z) = 2|y|e^{z^2/2}\int_0^1 s^{-3/2}e^{-y^2/(2s)}\overline{\Phi}\left(\frac{|y+z|}{\sqrt{1-s}}\right)ds$$

Here $\overline{\Phi}$ denotes the tail distribution of the standard normal law.

Asymptotic theory for $V(h,x)^n$

Theorem (Ivanovs & P. (19))

We obtain the stable convergence

$$\Delta_n^{-1/4}\left(V(h_0,x)^n - I(x)\right) \xrightarrow{d_{st}} \mathcal{MN}(0, v_{h_0} I(x))$$

We conjecture that this result can be extended to continuous stochastic differential equations.

Occupation time measure

- In this part we consider a Brownian motion X. The object of interest is the occupation time measure

$$L(x) = \int_0^1 1_{(x,\infty)}(X_s)ds$$

which turns out to be easier to treat than the previous two cases.

- We will again compute the conditional mean estimator

$$L^n(x) := \mathbb{E}\left[L(x)|\ (X_{i\Delta_n})_{1\le i\le \lfloor 1/\Delta_n\rfloor}\right]$$

Define $L^i_{i-1}(x) = \int_{(i-1)\Delta_n}^{i\Delta_n} 1_{(x,\infty)}(X_s)ds$ and observe the identity

$$\mathbb{E}\left[L^i_{i-1}(x)|X_{(i-1)\Delta_n}, \Delta_n^{-1/2}\Delta_i^n X\right]$$
$$= \Delta_n \int_0^1 \overline{\Phi}_{t(1-t)}\left(\Delta_n^{-1/2}(x - X_{(i-1)\Delta_n} - t\Delta_i^n X)\right) dt$$

where $\overline{\Phi}_t$ is the tail distribution of $\mathcal{N}(0,t)$.

Computation of $L^n(x)$

Using again the Markov property of the Brownian motion we obtain the formula

$$L^n(x) = \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} \mathbb{E}\left[L^i_{i-1}(x) \middle| \, (X_{i\Delta_n})_{1 \le i \le \lfloor 1/\Delta_n \rfloor} \right]$$

$$= \Delta_n \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} f\left(\Delta_n^{-1/2}(x - X_{(i-1)\Delta_n}), \Delta_n^{-1/2} \Delta_i^n X \right)$$

with

$$f(y,z) = \int_0^1 \overline{\Phi}_{t(1-t)}\left(y - tz\right) dt$$

Theorem (Ivanovs & P. (19))

We obtain the stable convergence

$$\Delta_n^{-3/4}\left(L^n(x) - \int_0^1 1_{(x,\infty)}(X_s)ds\right) \xrightarrow{d_{st}} \mathcal{MN}(0, v_f l(x))$$

where $v_f > 0$ *is a certain constant.*

The rate optimality of the rate $\Delta_n^{-3/4}$ has been shown in Ngo & Ogawa (11) in the setting of continuous diffusion models.

Thank you very much for your attention!