Semiparametric Estimation Theory for Discretely Observed Lévy Processes

Chris A.J. Klaassen
Enno Veerman

Korteweg-de Vries Institute for Mathematics
University of Amsterdam

EURANDOM
August 29, 2011

Discretely observed Lévy processes

Let $\{Y_t \,:\, t \geq 0\}$ be a Lévy process; sample paths are càdlàg; stationary independent increments.
Observe this process at times $t = 0, 1, 2, \ldots$ and base inference on

$$X_i = Y_i - Y_{i-1}\,, \quad i = 1, \ldots, n.$$

Since $\{Y_t \,:\, t \geq 0\}$ is a Lévy process, the observations $X_1, \ldots, X_n$ are i.i.d. with infinitely divisible distribution.

Discretely observed Lévy processes

Infinitely divisible

The observations $X_1, \ldots, X_n$ are i.i.d. with infinitely divisible distribution $P_{\mu,\sigma,\nu}$ and characteristic function

$$E\left(e^{itX}\right) = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2 + \int \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right] d\nu(x)\right),$$

where $\mu \in \mathbb{R}$, $\sigma \geq 0$, and the Lévy measure $\nu(\cdot)$ is a measure on $\mathbb{R} \setminus \{0\}$ satisfying

$$\int [x^2 \wedge 1]\, d\nu(x) < \infty.$$

Discretely observed Lévy processes

Infinitely divisible

The observations $X_1, \ldots, X_n$ are i.i.d. with infinitely divisible distribution in $\mathcal{P} = \{P_{\mu,\sigma,\nu} \,:\, \mu \in \mathbb{R},\, \sigma \geq 0,\; \nu(\cdot) \text{ Lévy measure}\}$. $\mathcal{P}$ defines a semiparametric model with μ and σ as Euclidean parameters, and $\nu(\cdot)$ as Banach parameter.

Parameter of interest

$$\theta \,:\, \mathcal{P} \to \mathbb{R}^k$$

Outline

Outline

1. Basics Semiparametrics
2. Efficient Estimation for Discretely Observed Lévy Processes
3. Further comments

Crash Course Semiparametrically Efficient Estimation

1. Asymptotic bound on performance of estimators in a regular parametric model (Local Asymptotic Normality):
    - Hájek-LeCam Convolution Theorem
    - Local Asymptotic Minimax Theorem
    - Local Asymptotic *Spread* Theorem
2. Regular parametric submodels of semiparametric model
3. Least favorable parametric submodel $\Rightarrow$ semiparametric bound
   Techniques to obtain semiparam. efficient influence function:
    - Projection of influence function on tangent space
    - Projection of score function on subspace of tangent space determined by nuisance parameters
4. Construction of estimator attaining bounds; i.e., of estimator that is asymptotically linear in the efficient influence function

Hájek-LeCam Convolution Theorem

In a regular parametric model one has Local Asymptotic Normality

$$\sum_{i=1}^{n} \log\left[\frac{p(X_i;\theta_n)}{p(X_i;\theta_0)}\right] = \frac{h}{\sqrt{n}}\sum_{i=1}^{n}\dot{\ell}_{\theta_0}(X_i) - \frac{1}{2}h^T I(\theta_0)h + o_P(1)$$

under θ_0 with $\theta_n = \theta_0 + h/\sqrt{n}$, where $\dot{\ell}_{\theta_0}(\cdot)$ is the *score function*.

Convolution theorem; under LAN

$$\forall h\ \sqrt{n}\,(T_n - q(\theta_n)) \xrightarrow{\mathcal{D}}_{\theta_n} L \ \Rightarrow L = \mathcal{N}\left(0, \dot{q}(\theta_0)I^{-1}(\theta_0)\dot{q}^T(\theta_0)\right) * M$$

and $L = \mathcal{N}\left(0, \dot{q}(\theta_0)I^{-1}(\theta_0)\dot{q}^T(\theta_0)\right)$ iff

$$\sqrt{n}\left\{T_n - \left[q(\theta_0) + \frac{1}{n}\sum_{i=1}^{n}\dot{q}(\theta_0)I^{-1}(\theta_0)\dot{\ell}_{\theta_0}(X_i)\right]\right\} \xrightarrow{P}_{\theta_0} 0$$

Hájek-LeCam Convolution Theorem

Efficiency

(T_n) is called (asymptotically) efficient iff

$$\sqrt{n}\left\{ T_n - \left[q(\theta_0) + \frac{1}{n}\sum_{i=1}^{n} \dot{q}(\theta_0) I^{-1}(\theta_0)\dot{\ell}_{\theta_0}(X_i) \right] \right\} \xrightarrow{P}_{\theta_0} 0$$

- Taking $q(\theta) = (I, 0)\,\theta$ one can study efficiency in presence of nuisance parameters.
- Taking regular parametric submodels of semiparametric models one can study efficiency in presence of infinite-dimensional nuisance parameters; try to get $\dot{q}(\theta_0) I^{-1}(\theta_0)\dot{q}^T(\theta_0)$ as large as possible.

Geometric Interpretation

Efficiency

(T_n) is called (asymptotically) efficient iff

$$\sqrt{n}\left\{T_n - \left[q(\theta_0) + \frac{1}{n}\sum_{i=1}^{n}\tilde{\ell}(X_i)\right]\right\} \xrightarrow{P}_{\theta_0} 0$$

with *the efficient influence function* being

$$\tilde{\ell}(\cdot) = \dot{q}(\theta_0) I^{-1}(\theta_0)\dot{\ell}_{\theta_0}(\cdot)$$

$$\tilde{\ell} \in [\,\dot{\ell}\,] = \dot{\mathcal{P}} \subset L_2^0(P_0)\,, \quad P_0 \leftrightsquigarrow \theta_0, \quad \dot{\ell} = \dot{\ell}_{\theta_0}, \quad E_{P_0}\dot{\ell} = 0$$

The closed linear span of the components of $\dot{\ell}$ (stemming from all regular parametric submodels) is denoted by $[\,\dot{\ell}\,] = \dot{\mathcal{P}}$ and is called the tangent space of $\mathcal{P}$ at P_0.

Geometric Interpretation

Efficiency and linearity

(T_n) is called (asymptotically) linear iff

$$\sqrt{n}\left\{T_n - \left[q(\theta_0) + \frac{1}{n}\sum_{i=1}^{n}\psi(X_i)\right]\right\} \xrightarrow{P}_{\theta_0} 0$$

with $\psi(\cdot)$ the *influence function.*
(T_n) is called (asymptotically) efficient iff $\psi = \tilde{\ell} = \dot{q}(\theta_0)I^{-1}(\theta_0)\dot{\ell}_{\theta_0}$ the *efficient influence function.* ($\theta(P) \leftrightsquigarrow q(\theta)$ pathwise diff.)
Theorem For any model $\mathcal{P}$ with tangent space $\dot{\mathcal{P}}$ at P_0, and $\forall\ \psi$

$$\psi - \tilde{\ell} \perp \dot{\mathcal{P}} \quad \text{or} \quad \tilde{\ell} = \prod\left(\psi \,\middle|\, \dot{\mathcal{P}}\right)$$

Geometric Interpretation

Efficient influence function and tangent space

$$\tilde{\ell} \in [\dot{\ell}] = \dot{\mathcal{P}} \subset L_2^0(P_0)$$

- Let $\mathcal{P}$ be a nonparametric, semiparametric, or parametric model.
- Let $P_0 \in \mathcal{P}$ and let $\tilde{\ell} \in \dot{\mathcal{P}}$ be the corresponding efficient influence function.
- Let $\mathcal{P}_s$ be a submodel, parametric or not, with $P_0 \in \mathcal{P}_s$, and let $\tilde{\ell}_s \in \dot{\mathcal{P}}_s$ denote the corresponding efficient influence function.

Geometry

$$P_0 \in \mathcal{P}_s \subset \mathcal{P}, \quad \dot{\mathcal{P}}_s \subset \dot{\mathcal{P}}$$

Geometric Interpretation

Projection efficient influence functions

$$P_0 \in \mathcal{P}_s \subset \mathcal{P}, \quad \tilde{\ell}_s \in \dot{\mathcal{P}}_s \subset \dot{\mathcal{P}}, \quad \tilde{\ell} \in \dot{\mathcal{P}} \subset L_2^0(P_0)$$

Theorem

$$\tilde{\ell}_s = \prod \left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s \right)$$

Proof From the preceding Theorem we know

$$\forall\, \psi \quad \tilde{\ell} = \prod \left(\psi \,\middle|\, \dot{\mathcal{P}} \right)$$

and hence in view of $\dot{\mathcal{P}}_s \subset \dot{\mathcal{P}}$

$$\tilde{\ell}_s = \prod \left(\psi \,\middle|\, \dot{\mathcal{P}}_s \right) = \prod \left(\prod \left(\psi \,\middle|\, \dot{\mathcal{P}} \right) \,\middle|\, \dot{\mathcal{P}}_s \right) = \prod \left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s \right) \qquad \square$$

Geometric Interpretation

Projection efficient influence functions

$$P_0 \in \mathcal{P}_s \subset \mathcal{P}, \quad \tilde{\ell}_s \in \dot{\mathcal{P}}_s \subset \dot{\mathcal{P}}, \quad \tilde{\ell} \in \dot{\mathcal{P}} \subset L_2^0(P_0)$$

Theorem

$$\tilde{\ell}_s = \prod \left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s \right)$$

Increments Lévy process

- P_0 some infinitely divisible distribution
- $\mathcal{P}_s$ all infinitely divisible distributions
- $\mathcal{P}$ all distributions

$$\theta : \mathcal{P} \to \mathbb{R}^k, \quad \theta(P) = \int g\, dP, \ F_P^{-1}(u)$$

Geometric Interpretation

Nonparametric tangent space

Lemma $P_0 \in \mathcal{P}$, all distributions.

$$\dot{\mathcal{P}} = L_2^0(P_0)$$

Proof Let $h \in L_2^0(P_0)$, and choose $\chi : \mathbb{R} \to (0,2)$, $\chi(0) = \chi'(0) = 1, \quad 0 < \chi'/\chi < 2$. E.g. $\chi(x) = 2/(1+e^{-x})$.

$$\eta \mapsto \frac{dP_\eta}{dP_0}(\cdot) = \frac{\chi(\eta h(\cdot))}{\int \chi(\eta h(x))\, dP_0(x)}$$

defines a regular parametric submodel with score function

$$\dot{\ell}_\eta(x)\Big|_{\eta=0} = \frac{\chi'}{\chi}(\eta h(x))\, h(x) - \frac{\int \chi'(\eta h) h\, dP_0}{\int \chi(\eta h)\, dP_0}\Big|_{\eta=0} = h(x). \qquad \square$$

Nonparametric efficient estimation

$P_0 \in \mathcal{P}$, all distributions, $\dot{\mathcal{P}} = L_2^0(P_0)$

$$\theta(P) = \int g \, dP, \quad \int g^2 \, dP < \infty$$

Linear, asymptotically efficient estimator

$$T_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i) = \theta(P_0) + \frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) - \int g \, dP_0\right]$$

Indeed,

$$\psi = g - \int g \, dP_0 \in L_2^0(P_0) = \dot{\mathcal{P}} \Rightarrow \psi = \tilde{\ell} = g - \int g \, dP_0$$

Outline

1. Basics Semiparametrics
2. Efficient Estimation for Discretely Observed Lévy Processes
3. Further comments

Geometry

Increments Lévy process

- P_0 some infinitely divisible distribution
- $\mathcal{P}_s$ all infinitely divisible distributions
- $\mathcal{P}$ all distributions; $\dot{\mathcal{P}} = L_2^0(P_0)$

$$\theta : \mathcal{P} \to \mathbb{R}^k, \quad \theta(P) = \int g\,dP, \quad \tilde{\ell} = g - \int g\,dP_0 \in \dot{\mathcal{P}}$$

Projection efficient influence functions

$$P_0 \in \mathcal{P}_s \subset \mathcal{P}, \quad \tilde{\ell}_s \in \dot{\mathcal{P}}_s \subset \dot{\mathcal{P}}, \quad \tilde{\ell} \in \dot{\mathcal{P}} \subset L_2^0(P_0)$$

Theorem

$$\tilde{\ell}_s = \prod\left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s\right)$$

Efficient estimator for discretely observed Lévy process

Main theorem

Theorem

$$\text{If } \sigma > 0, \text{ then} \quad \dot{\mathcal{P}}_s = L_2^0(P_0) = \dot{\mathcal{P}}$$

and hence

$$\tilde{\ell}_s = \prod\left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s\right) = \prod\left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}\right) = \tilde{\ell} = g - \int g \, dP_0$$

and hence

$$T_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i) = \theta(P) + \frac{1}{n}\sum_{i=1}^{n}\left[g(X_i) - \int g \, dP\right]$$

is asymptotically efficient (under all asymptotically linear estimators) in estimating $\theta(P) = \int g \, dP$ within the model $\mathcal{P}_s$ of all infinitely divisible distributions.

Proof main theorem; score functions

Main theorem

Theorem If $\sigma > 0$, then

$$\dot{\mathcal{P}}_s = L_2^0(P_0) = \dot{\mathcal{P}}$$

Proof Fix $\mu_0 \in \mathbb{R}, \sigma > 0$, and Lévy measure ν, corresponding to $P_0 \in \mathcal{P}_s$. Choose a probability measure Q on $\mathbb{R} \setminus \{0\}$. Let distribution $P_{\mu,\eta}$ have characteristic function

$$\phi_{\mu,\eta}(t) = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2 + \int \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]} \right] d(\nu + \eta Q)(x) \right)$$

Note $P_{\mu_0,0} = P_0$ and $P_{\mu,\eta}$ has an everywhere positive density w.r.t. Lebesgue measure, $f_{\mu,\eta}$ say. Write $\phi_0 = \phi_{\mu_0,0}$, $f_0 = f_{\mu_0,0}$. Note

$$f_{\mu,\eta}(x) = \frac{1}{2\pi} \int e^{-itx} \phi_{\mu,\eta}(t)\, dt$$

Proof main theorem; score functions

$$\phi_{\mu,\eta}(t) = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2 + \int \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]} \right] d(\nu + \eta Q)(x) \right)$$

$$f_{\mu,\eta}(x) = \frac{1}{2\pi} \int e^{-itx} \phi_{\mu,\eta}(t)\, dt$$

Score function for location

$$\frac{\partial}{\partial \mu} \log\left(f_{\mu,\eta}(x)\right)\Big|_{\mu=\mu_0,\eta=0} = -\frac{f_0'}{f_0}(x)$$

$$= \frac{\partial}{\partial \mu} \log\left(\int e^{-itx} \phi_{\mu,0}(t) dt \right)\Big|_{\mu=\mu_0} = \frac{\int it\, e^{-itx} \phi_0(t)\, dt}{\int e^{-itx} \phi_0(t)\, dt}$$

Proof main theorem; score functions

$$\phi_{\mu,\eta}(t) = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2 + \int \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]} \right] d(\nu + \eta Q)(x) \right)$$

$$f_{\mu,\eta}(x) = \frac{1}{2\pi} \int e^{-itx} \phi_{\mu,\eta}(t)\, dt, \quad -\frac{f_0'}{f_0}(x) = \frac{\int it\, e^{-itx} \phi_0(t)\, dt}{\int e^{-itx} \phi_0(t)\, dt}$$

Score function for Lévy measure ν in direction Q

$$\left. \frac{\int \left\{ \int \left[e^{ity} - 1 - ity\mathbf{1}_{[|y|<1]} \right] dQ(y) \right\} e^{-itx} \phi_{\mu_0,\eta}(t)\, dt}{\int e^{-itx} \phi_{\mu_0,\eta}(t)\, dt} \right|_{\eta=0}$$

$$= \frac{\int \left\{ \phi_Q(t) - 1 - it\mu_Q \right\} e^{-itx} \phi_0(t)\, dt}{\int e^{-itx} \phi_0(t)\, dt}$$

$$= \frac{f_{P_0 \star Q}}{f_0}(x) - 1 + \mu_Q \frac{f_0'}{f_0}(x)$$

Proof main theorem; score functions

Score function for location is $-\frac{f_0'}{f_0}(x)$. Score function for Lévy measure ν in direction Q is $\frac{f_{P_0 \star Q}}{f_0}(x) - 1 + \mu_Q \frac{f_0'}{f_0}(x)$.
With Q degenerate at $y \neq 0$ this becomes

$$\frac{f_0(x-y)}{f_0(x)} - 1 + \mu_Q \frac{f_0'}{f_0}(x).$$

Conclusion

$$\left[-\frac{f_0'}{f_0}(\cdot), \frac{f_0(\cdot - y)}{f_0(\cdot)} - 1 + \mu_Q \frac{f_0'}{f_0}(\cdot)\,;\; y \in \mathbb{R} \right]$$
$$= \left[-\frac{f_0'}{f_0}(\cdot), \frac{f_0(\cdot - y)}{f_0(\cdot)} - 1\,;\; y \in \mathbb{R} \right] \subset \dot{\mathcal{P}}_s$$

Proof main theorem; orthogonality

To prove

$$\dot{\mathcal{P}}_s = L_2^0(P_0)$$

We have shown

$$\left[-\frac{f_0'}{f_0}(\cdot), \frac{f_0(\cdot - y)}{f_0(\cdot)} - 1 \,;\ y \in \mathbb{R} \right] \subset \dot{\mathcal{P}}_s$$

We will prove

$$L_2^0(P_0) \ni \ g \perp \dot{\mathcal{P}}_s \quad \Rightarrow \quad g = 0$$

more precisely

$$\forall y \quad g \perp \frac{f_0(\cdot - y)}{f_0(\cdot)} - 1 \quad \Rightarrow \quad g = 0$$

Proof main theorem; completeness

To prove for $g \in L_2^0(P_0)$

$$\forall y \int g(x) \left\{ \frac{f_0(x-y)}{f_0(x)} - 1 \right\} dP_0(x) = 0 \Rightarrow g(x) = 0 \text{ Lebesgue a.a. } x$$

or

$$\forall y \in \mathbb{R} \int g(x+y)\, dP_0(x) = 0 \quad \Rightarrow \quad g = 0 \text{ Lebesgue a.e.}$$

This is related to completeness of the location family of P_0.

Proof main theorem; annihilating signed measures

Choose $0 < \epsilon < 1$. For Lévy measure ν define

$$c_\epsilon = \int_{\epsilon \le |x|} d\nu(x), \ d_\epsilon = \int_{\epsilon \le |x| < 1} x \, d\nu(x), \ G_\epsilon(y) = \frac{1}{c_\epsilon} \int_{x \le y, \, \epsilon \le |x|} d\nu(x)$$

c_ϵ and d_ϵ are finite, G_ϵ is distribution function. Define H_ϵ by

$$H_\epsilon(z) = \sum_{j=0}^{\infty} e^{-c_\epsilon} \frac{c_\epsilon^j}{j!} G_\epsilon^{*j}(z + d_\epsilon) \, .$$ Then

$$\begin{aligned} \int e^{itz} \, dH_\epsilon(z) &= \sum_{j=0}^{\infty} e^{-c_\epsilon} \frac{c_\epsilon^j}{j!} \int e^{itz - itd_\epsilon} dG_\epsilon^{*j}(z) \\ &= \exp\left(c_\epsilon \, E_{G_\epsilon} \left(e^{itY} - 1 \right) - itd_\epsilon \right) \\ &= \exp\left(\int_{\epsilon \le |x|} \left[e^{itx} - 1 - itx \mathbf{1}_{[|x|<1]} \right] \, d\nu(x) \right) . \end{aligned}$$

Proof main theorem; annihilating signed measures

So, with

$$H_\epsilon(z) = \sum_{j=0}^{\infty} e^{-c_\epsilon} \frac{c_\epsilon^j}{j!} G_\epsilon^{*j}\left(z + d_\epsilon\right) \text{ we have}$$

$$\int e^{itz}\, dH_\epsilon(z) = \exp\left(\int_{\epsilon \le |x|} \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right] d\nu(x)\right)$$

Similarly (Enno), with

$$H_\epsilon^-(z) = \sum_{j=0}^{\infty} e^{c_\epsilon} \frac{(-c_\epsilon)^j}{j!} G_\epsilon^{*j}\left(z - d_\epsilon\right) \text{ we have}$$

$$\int e^{itz}\, dH_\epsilon^-(z) = \exp\left(-\int_{\epsilon \le |x|} \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right] d\nu(x)\right)$$

Proof main theorem; annihilating signed measures

By multiplication we see that the Fourier-Stieltjes transform of the convolution of the measure defined by H_ϵ and the signed measure induced by H_ϵ^- equals 1.
This means that the convolution corresponds to unit point mass at 0.
In a sense one could say that the signed measure induced by H_ϵ^- annihilates H_ϵ.

Proof main theorem; completeness

$$X = \mu_0 + \sigma U + Y_\epsilon + Z_\epsilon \ \sim \ P_0$$

U, Y_ϵ, and Z_ϵ are independent
U is a standard normal random variable
Y_ϵ has characteristic function

$$E\left(e^{itY_\epsilon}\right) = \exp\left(\int_{0<|x|<\epsilon} \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right] d\nu(x)\right)$$

Z_ϵ has characteristic function

$$E\left(e^{itZ_\epsilon}\right) = \int e^{itz}\, dH_\epsilon(z) = \exp\left(\int_{\epsilon\le|x|} \left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right] d\nu(x)\right)$$

To prove for $g \in L_2^0(P_0)$

$$\forall y \in \mathbb{R} \quad Eg(X+y) = 0 \quad \Rightarrow \quad g = 0 \text{ Lebesgue a.e.}$$

Proof main theorem; completeness

$$X = \mu_0 + \sigma U + Y_\epsilon + Z_\epsilon \; \sim \; P_0$$

Define $g^*(z) = Eg(\mu_0 + \sigma U + Y_\epsilon + z)$. Then for all y

$$0 = Eg(X + y) = Eg^*(Z_\epsilon + y)$$

and hence for all $a \in \mathbb{R}$ $(y = w + a)$

$$\begin{aligned} 0 &= \int Eg^*(Z_\epsilon + w + a)\, dH_\epsilon^-(w) \\ &= \int\int g^*(z + w + a)\, dH_\epsilon(z)\, dH_\epsilon^-(w) \\ &= \int g^*(v + a)\, dH_\epsilon \star H_\epsilon^-(v) = g^*(a) \end{aligned}$$

Here we use $g \in L_2^0(P_0)$.

Proof main theorem; completeness

We have

$$0 = g^*(a) = Eg(\mu_0 + \sigma U + Y_\epsilon + a)$$

Define

$$\tilde{g}(z) = Eg(\mu_0 + \sigma U + z)$$

Then

$$0 = g^*(a) = E\tilde{g}(Y_\epsilon + a)$$

Proof main theorem; completeness

$$0 = E\tilde{g}(Y_\epsilon + a)$$

Let Y_ϵ and Y_ϵ^* be i.i.d., let U, Y_ϵ, Y_ϵ^*, and Z_ϵ be independent, and denote $Y_\epsilon + Z_\epsilon = V$.
Fix $b \in \mathbb{R}$ and $\delta > 0$.
In view of $E|\tilde{g}(V+b)| \leq E|g(X+b)| < \infty$ holds, there exists a continuous function $\chi(\cdot)$ with compact support satisfying

$$E\left|\tilde{g}(V+b) - \chi(V+b)\right| < \delta$$

Proof main theorem; completeness

$$0 = E\tilde{g}(Y_\epsilon + a),\ E\left|\tilde{g}(V+b) - \chi(V+b)\right| < \delta,\ V = Y_\epsilon + Z_\epsilon$$

$$\begin{aligned}
E\left|\tilde{g}(V+b)\right| &= E\left\{\int \left|\tilde{g}\left(Y_\epsilon + z + b\right) - E\tilde{g}\left(Y_\epsilon^* + z + b\right)\right| dH_\epsilon(z)\right\} \\
&\leq E\Big\{\int \left|\tilde{g}\left(Y_\epsilon + z + b\right) - \chi\left(Y_\epsilon + z + b\right)\right| \\
&\qquad + E\left|\tilde{g}\left(Y_\epsilon^* + z + b\right) - \chi\left(Y_\epsilon^* + z + b\right)\right| \\
&\qquad + \Big|\chi\left(Y_\epsilon + z + b\right) - E\chi\left(Y_\epsilon^* + z + b\right)\Big| dH_\epsilon(z)\Big\} \\
&< 2\delta + E\left|\chi\left(Y_\epsilon + Z_\epsilon + b\right) - \chi\left(Y_\epsilon^* + Z_\epsilon + b\right)\right|.
\end{aligned}$$

Proof main theorem; completeness

$$E\,|\tilde{g}(V+b)| < 2\delta + E\,|\chi\,(Y_\epsilon + Z_\epsilon + b) - \chi\,(Y_\epsilon^* + Z_\epsilon + b)|$$

By

$$E\left(e^{itY_\epsilon}\right) = \exp\left(\int_{0<|x|<\epsilon}\left[e^{itx} - 1 - itx\mathbf{1}_{[|x|<1]}\right]d\nu(x)\right)$$

it follows that Y_ϵ converges to 0 in probability as $\epsilon \downarrow 0$, and hence $(Y_\epsilon, Y_\epsilon^*, Z_\epsilon) = (Y_\epsilon, Y_\epsilon^*, V - Y_\epsilon)$ converges in distribution to $(0, 0, V)$. Since $\chi(\cdot)$ is bounded and continuous this implies

$$\lim_{\epsilon\downarrow 0} E\,|\chi\,(Y_\epsilon + Z_\epsilon + b) - \chi\,(Y_\epsilon^* + Z_\epsilon + b)| = 0$$

So,

$$E\,|\tilde{g}(V+b)| < 2\delta \text{ arbitrarily small}$$

Proof main theorem; completeness

$$E|\tilde{g}(V+b)| = 0 \text{ for all } \; b \in \mathbb{R}$$

Hence, we have e.g. $E|\tilde{g}(V+U)| = 0$.
Because $V+U$ has a positive density with respect to Lebesgue measure, this implies

$$\tilde{g}(y) = Eg(\mu_0 + \sigma U + y) = 0$$

for Lebesgue almost all $y \in \mathbb{R}$. By completeness of the normal location family

$$g(\mu_0 + \sigma U + y) = 0$$

holds a.s. for all $y \in \mathbb{R}$ and hence

$$g(\mu_0 + \sigma U + V) = g(X) = 0$$

holds a.s. □

Efficient estimator for discretely observed Lévy process

Main theorem

Theorem

$$\text{If } \sigma > 0, \text{ then } \quad \dot{\mathcal{P}}_s = L_2^0(P_0) = \dot{\mathcal{P}}$$

and hence

$$\tilde{\ell}_s = \prod \left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}}_s \right) = \prod \left(\tilde{\ell} \,\middle|\, \dot{\mathcal{P}} \right) = \tilde{\ell} = g - \int g \, dP_0$$

and hence

$$T_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i) = \theta(P) + \frac{1}{n} \sum_{i=1}^{n} \left[g(X_i) - \int g \, dP \right]$$

is asymptotically efficient (under all asymptotically linear estimators) in estimating $\theta(P) = \int g \, dP$ within the model $\mathcal{P}_s$ of all infinitely divisible distributions.

Outline

Efficient estimator for discretely observed Lévy process

1. Compound Poisson case has been treated by Enno Veerman
2. Remaining case, namely $\sigma = 0$ and $\nu(\{|x| < \epsilon\}) > 0$ for all $\epsilon > 0$, still conjecture
3. Further research needed for case of nonequidistant time points

Finite sample spread inequality

Definitions

ϑ random variable on $\mathbb{R}$ with density $w(\cdot)$
Given $\vartheta = \theta, \quad X_1, \ldots, X_n$ i.i.d. with parameter θ

$$H(z) = P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_{\vartheta}(X_i) + \frac{1}{\sqrt{n}} \frac{w'}{w}(\vartheta) \leq z\right)$$

is the distribution function of the score statistic

$$G(y) = P\left(\sqrt{n}(T_n - \vartheta) \leq y\right)$$

is the weighted distribution function of any estimator

Finite sample spread inequality

Definitions

ϑ random variable on $\mathbb{R}$ with density $w(\cdot)$

Given $\vartheta = \theta$, $X_1, \ldots, X_n$ i.i.d. with parameter θ

$$H(z) = P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_{\vartheta}(X_i) + \frac{1}{\sqrt{n}} \frac{w'}{w}(\vartheta) \leq z \right)$$

$$G(y) = P\left(\sqrt{n}(T_n - \vartheta) \leq y \right)$$

Spread inequality

$$G^{-1}(v) - G^{-1}(u) \geq K^{-1}(v) - K^{-1}(u) = \int_u^v \frac{1}{\int_s^1 H^{-1}(t)dt} ds$$

Local asymptotic spread inequality

Fix $\theta_0 \in \mathbb{R}$ write $\vartheta = \theta_0 + \frac{\sigma}{\sqrt{n}}\zeta$ with ζ random, density $w_0(\cdot)$

$$H_{n\sigma}(z) = P\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{\ell}_{\theta_0+\frac{\sigma}{\sqrt{n}}\zeta}(X_i) + \frac{1}{\sigma}\frac{w_0'}{w_0}(\zeta) \le z\right)$$

$$G_{n\sigma}(y) = P\left(\sqrt{n}\left(T_n - \theta_0 - \frac{\sigma}{\sqrt{n}}\zeta\right) \le y\right)$$

Local asymptotic spread inequality

$$\liminf_{\sigma\to\infty}\liminf_{n\to\infty}\left[G_{n\sigma}^{-1}(v) - G_{n\sigma}^{-1}(u)\right] \ge \lim_{\sigma,n\to\infty}\int_u^v \frac{1}{\int_s^1 H_{n\sigma}^{-1}(t)dt}ds$$

$$= \frac{1}{\sqrt{I(\theta_0)}}\left[\Phi^{-1}(v) - \Phi^{-1}(u)\right]$$

Local asymptotic spread inequality

Local asymptotic spread theorem

$$
\begin{aligned}
& \limsup_{\sigma\to\infty} \limsup_{n\to\infty} \left[G_{n\sigma}^{-1}(v) - G_{n\sigma}^{-1}(u) \right] \\
\geq\ & \liminf_{\sigma\to\infty} \liminf_{n\to\infty} \left[G_{n\sigma}^{-1}(v) - G_{n\sigma}^{-1}(u) \right] \\
\geq\ & \lim_{\sigma,n\to\infty} \int_u^v \frac{1}{\int_s^1 H_{n\sigma}^{-1}(t)\,dt} ds = \frac{1}{\sqrt{I(\theta_0)}} \left[\Phi^{-1}(v) - \Phi^{-1}(u) \right]
\end{aligned}
$$

with equalities for all $0 < u < v < 1$ iff

$$
\sqrt{n} \left\{ T_n - \theta_0 - \frac{1}{n} \sum_{i=1}^{n} \frac{1}{I(\theta_0)} \dot{\ell}_{\theta_0}(X_i) \right\} \to_{P_{\theta_0}} 0
$$