Nonparametric estimation of the characteristic triplet of a discretely observed Lévy process

Shota Gugushvili

EURANDOM
Technische Universiteit Eindhoven

Eindhoven, 15-17 July 2009

Outline

Outline

Characterisation of Lévy processes

- Let $X = (X_t)_{t \geq 0}$ be a Lévy process.
- Marginal distributions of X are infinitely divisible and are determined by the distribution of X_1. Conversely, given an infinitely divisible distribution μ, one can construct a Lévy process, such that $P_{X_1} = \mu$.
- With this in mind, the Lévy-Khintchine formula provides us with unique means for characterisation of any Lévy process:

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1 - izx 1_{[|x|<1]})\nu(dx) \right].$$

Here $\gamma \in \mathbb{R}, \sigma \geq 0$ and ν is a measure concentrated on $\mathbb{R}\backslash\{0\}$, such that $\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < \infty$.

Characterisation of Lévy processes

- Let $X = (X_t)_{t \geq 0}$ be a Lévy process.
- Marginal distributions of X are infinitely divisible and are determined by the distribution of X_1. Conversely, given an infinitely divisible distribution μ, one can construct a Lévy process, such that $P_{X_1} = \mu$.
- With this in mind, the Lévy-Khintchine formula provides us with unique means for characterisation of any Lévy process:

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1 - izx 1_{[|x|<1]}) \nu(dx) \right].$$

Here $\gamma \in \mathbb{R}, \sigma \geq 0$ and ν is a measure concentrated on $\mathbb{R} \backslash \{0\}$, such that $\int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < \infty$.

Characterisation of Lévy processes

- Let $X = (X_t)_{t \geq 0}$ be a Lévy process.
- Marginal distributions of X are infinitely divisible and are determined by the distribution of X_1. Conversely, given an infinitely divisible distribution μ, one can construct a Lévy process, such that $P_{X_1} = \mu$.
- With this in mind, the Lévy-Khintchine formula provides us with unique means for characterisation of any Lévy process:

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1 - izx1_{[|x|<1]})\nu(dx) \right].$$

Here $\gamma \in \mathbb{R}, \sigma \geq 0$ and ν is a measure concentrated on $\mathbb{R} \backslash \{0\}$, such that $\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < \infty$.

Characteristic triplet

- The measure ν is called the Lévy measure corresponding to the Lévy process X.
- The triple (γ, σ, ν) is referred to as Lévy or characteristic triplet of X.
- The representation in terms of the triplet (γ, σ, ν) is unique.
- Consequently, the Lévy triplet provides us with means for unique characterisation of a law of any Lévy process.

Characteristic triplet

- The measure ν is called the Lévy measure corresponding to the Lévy process X.
- The triple (γ, σ, ν) is referred to as Lévy or characteristic triplet of X.
- The representation in terms of the triplet (γ, σ, ν) is unique.
- Consequently, the Lévy triplet provides us with means for unique characterisation of a law of any Lévy process.

Characteristic triplet

- The measure ν is called the Lévy measure corresponding to the Lévy process X.
- The triple (γ, σ, ν) is referred to as Lévy or characteristic triplet of X.
- The representation in terms of the triplet (γ, σ, ν) is unique.
- Consequently, the Lévy triplet provides us with means for unique characterisation of a law of any Lévy process.

Characteristic triplet

- The measure ν is called the Lévy measure corresponding to the Lévy process X.
- The triple (γ, σ, ν) is referred to as Lévy or characteristic triplet of X.
- The representation in terms of the triplet (γ, σ, ν) is unique.
- Consequently, the Lévy triplet provides us with means for unique characterisation of a law of any Lévy process.

Inference procedures for Lévy processes

- Many statistical problems for Lévy processes can be reduced to inference on the Lévy triplet.
- Depending on the parametrisation of the Lévy measure ν (or its density) there are several ways to approach estimation problems for Lévy processes: parametric, nonparametric and semiparametric approaches.
- In the parametric case the Lévy measure is parametrised by a Euclidean parameter. E.g. one can consider the class of gamma processes and in this case the corresponding Lévy measure can be parametrised as

$$\nu(dx) = \alpha x^{-1} e^{-x/\beta} 1_{[x>0]} dx.$$

Inference procedures for Lévy processes

- Many statistical problems for Lévy processes can be reduced to inference on the Lévy triplet.
- Depending on the parametrisation of the Lévy measure ν (or its density) there are several ways to approach estimation problems for Lévy processes: parametric, nonparametric and semiparametric approaches.
- In the parametric case the Lévy measure is parametrised by a Euclidean parameter. E.g. one can consider the class of gamma processes and in this case the corresponding Lévy measure can be parametrised as

$$\nu(dx) = \alpha x^{-1} e^{-x/\beta} 1_{[x>0]} dx.$$

Inference procedures for Lévy processes

- Many statistical problems for Lévy processes can be reduced to inference on the Lévy triplet.
- Depending on the parametrisation of the Lévy measure ν (or its density) there are several ways to approach estimation problems for Lévy processes: parametric, nonparametric and semiparametric approaches.
- In the parametric case the Lévy measure is parametrised by a Euclidean parameter. E.g. one can consider the class of gamma processes and in this case the corresponding Lévy measure can be parametrised as

$$\nu(dx) = \alpha x^{-1} e^{-x/\beta} 1_{[x>0]} dx.$$

Nonparametric inference for Lévy processes

- In the nonparametric case one does not impose parametric assumptions on the given family of Lévy measures, but only some smoothness assumptions.
- Doubts have been expressed in the literature whether Lévy processes parametrised by a small number of parameters (two, three or four) can adequately represent complex realities of financial markets.
- Nonparametric techniques might come in handy when one is interested e.g. in the shape of a Lévy density.
- They are not necessarily competitors of the parametric approach, but often complement it.

Nonparametric inference for Lévy processes

- In the nonparametric case one does not impose parametric assumptions on the given family of Lévy measures, but only some smoothness assumptions.
- Doubts have been expressed in the literature whether Lévy processes parametrised by a small number of parameters (two, three or four) can adequately represent complex realities of financial markets.
- Nonparametric techniques might come in handy when one is interested e.g. in the shape of a Lévy density.
- They are not necessarily competitors of the parametric approach, but often complement it.

Nonparametric inference for Lévy processes

- In the nonparametric case one does not impose parametric assumptions on the given family of Lévy measures, but only some smoothness assumptions.
- Doubts have been expressed in the literature whether Lévy processes parametrised by a small number of parameters (two, three or four) can adequately represent complex realities of financial markets.
- Nonparametric techniques might come in handy when one is interested e.g. in the shape of a Lévy density.
- They are not necessarily competitors of the parametric approach, but often complement it.

Nonparametric inference for Lévy processes

- In the nonparametric case one does not impose parametric assumptions on the given family of Lévy measures, but only some smoothness assumptions.
- Doubts have been expressed in the literature whether Lévy processes parametrised by a small number of parameters (two, three or four) can adequately represent complex realities of financial markets.
- Nonparametric techniques might come in handy when one is interested e.g. in the shape of a Lévy density.
- They are not necessarily competitors of the parametric approach, but often complement it.

Inference when low frequency data are available only

- Most of existing literature deals with estimation of the characteristic triplet under the assumption that either a continuous record of observations is available over time interval $[0, T]$, or X is observed at time instances $\Delta_n, 2\Delta_n, \ldots, n\Delta_n$ with $\Delta_n \to 0$ as $n \to \infty$ and $n\Delta_n \to \infty$.
- It is equally interesting to study inference procedures under the assumption that $\Delta_n = \Delta$ remains fixed as $n \to \infty$.
- Related references include Buchmann and Grübel (2003), Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Inference when low frequency data are available only

- Most of existing literature deals with estimation of the characteristic triplet under the assumption that either a continuous record of observations is available over time interval $[0, T]$, or X is observed at time instances $\Delta_n, 2\Delta_n, \ldots, n\Delta_n$ with $\Delta_n \to 0$ as $n \to \infty$ and $n\Delta_n \to \infty$.
- It is equally interesting to study inference procedures under the assumption that $\Delta_n = \Delta$ remains fixed as $n \to \infty$.
- Related references include Buchmann and Grübel (2003), Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Inference when low frequency data are available only

- Most of existing literature deals with estimation of the characteristic triplet under the assumption that either a continuous record of observations is available over time interval $[0, T]$, or X is observed at time instances $\Delta_n, 2\Delta_n, \ldots, n\Delta_n$ with $\Delta_n \to 0$ as $n \to \infty$ and $n\Delta_n \to \infty$.
- It is equally interesting to study inference procedures under the assumption that $\Delta_n = \Delta$ remains fixed as $n \to \infty$.
- Related references include Buchmann and Grübel (2003), Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Outline

Drift, compound Poisson process, Brownian motion

- We will concentrate on nonparametric inference for Lévy processes that are of finite jump activity and have absolutely continuous Lévy measures.
- In essence this means that we consider a superposition of a drift term, a compound Poisson process and an independent Brownian motion.
- The Lévy-Khintchine formula in our case takes the form

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1)\rho(x)dx\right], \quad (1)$$

where the Lévy density ρ is such that $\lambda := \int_{-\infty}^{\infty} \rho(x)dx < \infty$.

Drift, compound Poisson process, Brownian motion

- We will concentrate on nonparametric inference for Lévy processes that are of finite jump activity and have absolutely continuous Lévy measures.
- In essence this means that we consider a superposition of a drift term, a compound Poisson process and an independent Brownian motion.
- The Lévy-Khintchine formula in our case takes the form

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1)\rho(x)dx\right], \quad (1)$$

where the Lévy density ρ is such that $\lambda := \int_{-\infty}^{\infty} \rho(x)dx < \infty$.

Drift, compound Poisson process, Brownian motion

- We will concentrate on nonparametric inference for Lévy processes that are of finite jump activity and have absolutely continuous Lévy measures.
- In essence this means that we consider a superposition of a drift term, a compound Poisson process and an independent Brownian motion.
- The Lévy-Khintchine formula in our case takes the form

$$\phi_{X_1}(z) = \exp\left[i\gamma z - \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{izx} - 1)\rho(x)dx \right], \quad (1)$$

where the Lévy density ρ is such that $\lambda := \int_{-\infty}^{\infty} \rho(x)dx < \infty$.

Reformulation of the problem

- Suppose we dispose a sample $X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$ from the process X.
- By a rescaling argument, without loss of generality, we may take $\Delta = 1$.
- Based on this sample, our goal is to infer the characteristic triplet (γ, σ^2, ρ), as well as λ.
- The problem is equivalent to the following one: let $X_1, \ldots, X_n$ be i.i.d. copies of a random variable X with characteristic function given by (1). Based on these observations, estimate γ, σ^2, ρ and λ.

Reformulation of the problem

- Suppose we dispose a sample $X_{\Delta}, X_{2\Delta}, \ldots, X_{n\Delta}$ from the process X.
- By a rescaling argument, without loss of generality, we may take $\Delta = 1$.
- Based on this sample, our goal is to infer the characteristic triplet (γ, σ^2, ρ), as well as λ.
- The problem is equivalent to the following one: let $X_1, \ldots, X_n$ be i.i.d. copies of a random variable X with characteristic function given by (1). Based on these observations, estimate γ, σ^2, ρ and λ.

Reformulation of the problem

- Suppose we dispose a sample $X_{\Delta}, X_{2\Delta}, \ldots, X_{n\Delta}$ from the process X.
- By a rescaling argument, without loss of generality, we may take $\Delta = 1$.
- Based on this sample, our goal is to infer the characteristic triplet (γ, σ^2, ρ), as well as λ.
- The problem is equivalent to the following one: let $X_1, \ldots, X_n$ be i.i.d. copies of a random variable X with characteristic function given by (1). Based on these observations, estimate γ, σ^2, ρ and λ.

Reformulation of the problem

- Suppose we dispose a sample $X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$ from the process X.
- By a rescaling argument, without loss of generality, we may take $\Delta = 1$.
- Based on this sample, our goal is to infer the characteristic triplet (γ, σ^2, ρ), as well as λ.
- The problem is equivalent to the following one: let $X_1, \ldots, X_n$ be i.i.d. copies of a random variable X with characteristic function given by (1). Based on these observations, estimate γ, σ^2, ρ and λ.

Outline

Estimation method

- Many statisticians are addicted to some favourite statistical tools.
- One of such tools is a plug-in device.
- To use a plug-in device, we need to explicitly express the Lévy density in terms of the distribution of X.
- One possible approach is to base estimation procedures on an appropriate inversion of the Lévy-Khintchine formula, cf. Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Estimation method

- Many statisticians are addicted to some favourite statistical tools.
- One of such tools is a plug-in device.
- To use a plug-in device, we need to explicitly express the Lévy density in terms of the distribution of X.
- One possible approach is to base estimation procedures on an appropriate inversion of the Lévy-Khintchine formula, cf. Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Estimation method

- Many statisticians are addicted to some favourite statistical tools.
- One of such tools is a plug-in device.
- To use a plug-in device, we need to explicitly express the Lévy density in terms of the distribution of X.
- One possible approach is to base estimation procedures on an appropriate inversion of the Lévy-Khintchine formula, cf. Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Estimation method

- Many statisticians are addicted to some favourite statistical tools.
- One of such tools is a plug-in device.
- To use a plug-in device, we need to explicitly express the Lévy density in terms of the distribution of X.
- One possible approach is to base estimation procedures on an appropriate inversion of the Lévy-Khintchine formula, cf. Van Es et al. (2007), Genon-Catalot and Comte (2008) and Neumann and Reiß (2009).

Outline

Condition on ρ

In the sequel we will find it convenient to use the jump size density $f(x) := \rho(x)/\lambda$.

Condition on ρ

Let the unknown density ρ belong to the class

$$W(\beta, L, \Lambda, K) = \Big\{ \rho : \rho(x) = \lambda f(x), \int_{-\infty}^{\infty} x^2 f(x) dx \leq K, \int_{-\infty}^{\infty} |t|^{\beta} |\phi_f(t)| dt \leq L, \lambda \in (0, \Lambda] \Big\},$$

where β, L, Λ and K are strictly positive numbers.

Condition on ρ

In the sequel we will find it convenient to use the jump size density $f(x) := \rho(x)/\lambda$.

Condition on ρ

Let the unknown density ρ belong to the class

$$W(\beta, L, \Lambda, K) = \Big\{\rho : \rho(x) = \lambda f(x), \int_{-\infty}^{\infty} x^2 f(x) dx \leq K, \int_{-\infty}^{\infty} |t|^{\beta} |\phi_f(t)| dt \leq L, \lambda \in (0, \Lambda] \Big\},$$

where β, L, Λ and K are strictly positive numbers.

Conditions on γ and σ

Condition on σ

Let σ be such that $\sigma \in (0, \Sigma]$, where Σ is a strictly positive number.

Condition on γ

Let γ be such that $|\gamma| \leq \Gamma$, where Γ denotes a positive number.

Conditions on γ and σ

Condition on σ

Let σ be such that $\sigma \in (0, \Sigma]$, where Σ is a strictly positive number.

Condition on γ

Let γ be such that $|\gamma| \leq \Gamma$, where Γ denotes a positive number.

Outline

Inversion of the ch.f. ϕ_X

- Let $\Re(z)$ and $\Im(z)$ denote the real and the imaginary parts of a complex number z, respectively.
- From (1) we have

$$\log\left(|\phi_X(t)|\right) = -\lambda + \lambda\Re(\phi_f(t)) - \frac{\sigma^2 t^2}{2}. \qquad (2)$$

- Let v^h be a kernel that depends on a bandwidth h and is such that

$$\int_{-1/h}^{1/h} v^h(t)dt = 0, \quad \int_{-1/h}^{1/h} \left(-\frac{t^2}{2}\right) v^h(t)dt = 1.$$

Inversion of the ch.f. ϕ_X

- Let $\Re(z)$ and $\Im(z)$ denote the real and the imaginary parts of a complex number z, respectively.
- From (1) we have

$$\log(|\phi_X(t)|) = -\lambda + \lambda\Re(\phi_f(t)) - \frac{\sigma^2 t^2}{2}. \qquad (2)$$

- Let v^h be a kernel that depends on a bandwidth h and is such that

$$\int_{-1/h}^{1/h} v^h(t)dt = 0, \quad \int_{-1/h}^{1/h} \left(-\frac{t^2}{2}\right) v^h(t)dt = 1.$$

Inversion of the ch.f. ϕ_X

- Let $\Re(z)$ and $\Im(z)$ denote the real and the imaginary parts of a complex number z, respectively.
- From (1) we have

$$\log\left(|\phi_X(t)|\right) = -\lambda + \lambda\Re(\phi_f(t)) - \frac{\sigma^2 t^2}{2}. \tag{2}$$

- Let v^h be a kernel that depends on a bandwidth h and is such that

$$\int_{-1/h}^{1/h} v^h(t)dt = 0, \quad \int_{-1/h}^{1/h} \left(-\frac{t^2}{2}\right) v^h(t)dt = 1.$$

Inversion of the ch.f. ϕ_X (continued)

- In view of (2)

$$\int_{-1/h}^{1/h} \log(|\phi_X(t)|) v^h(t) dt = \lambda \int_{-1/h}^{1/h} \Re(\phi_f(t)) v^h(t) dt + \sigma^2. \tag{3}$$

- Provided enough assumptions on v^h, one can achieve that the right-hand side of (3) tends to σ^2 as $h \to 0$.
- A natural way to construct an estimator of σ^2 then is to replace in (3) $\log(|\phi_X(t)|)$ by its estimator $\log(|\phi_{emp}(t)|)$.
- This approach is close in spirit to the one in Belomestny and Reiß(2006).

Inversion of the ch.f. ϕ_X (continued)

- In view of (2)

$$\int_{-1/h}^{1/h} \log(|\phi_X(t)|) v^h(t) dt = \lambda \int_{-1/h}^{1/h} \Re(\phi_f(t)) v^h(t) dt + \sigma^2. \tag{3}$$

- Provided enough assumptions on v^h, one can achieve that the right-hand side of (3) tends to σ^2 as $h \to 0$.
- A natural way to construct an estimator of σ^2 then is to replace in (3) $\log(|\phi_X(t)|)$ by its estimator $\log(|\phi_{emp}(t)|)$.
- This approach is close in spirit to the one in Belomestny and Reiß(2006).

Inversion of the ch.f. ϕ_X (continued)

- In view of (2)

$$\int_{-1/h}^{1/h} \log(|\phi_X(t)|)v^h(t)dt = \lambda \int_{-1/h}^{1/h} \Re(\phi_f(t))v^h(t)dt + \sigma^2. \tag{3}$$

- Provided enough assumptions on v^h, one can achieve that the right-hand side of (3) tends to σ^2 as $h \to 0$.
- A natural way to construct an estimator of σ^2 then is to replace in (3) $\log(|\phi_X(t)|)$ by its estimator $\log(|\phi_{emp}(t)|)$.
- This approach is close in spirit to the one in Belomestny and Reiß(2006).

Inversion of the ch.f. ϕ_X (continued)

- In view of (2)

$$\int_{-1/h}^{1/h} \log(|\phi_X(t)|) v^h(t) dt = \lambda \int_{-1/h}^{1/h} \Re(\phi_f(t)) v^h(t) dt + \sigma^2. \tag{3}$$

- Provided enough assumptions on v^h, one can achieve that the right-hand side of (3) tends to σ^2 as $h \to 0$.
- A natural way to construct an estimator of σ^2 then is to replace in (3) $\log(|\phi_X(t)|)$ by its estimator $\log(|\phi_{emp}(t)|)$.
- This approach is close in spirit to the one in Belomestny and Reiß(2006).

Estimator of σ^2

- We propose

$$\tilde{\sigma}_n^2 = \int_{-1/h}^{1/h} \max\{\min\{M_n, \log(|\phi_{emp}(t)|)\}, -M_n\} v^h(t) dt \quad (4)$$

as an estimator of σ^2.

- Here M_n denotes a sequence of positive numbers diverging to infinity at a suitable rate.
- The truncation in (4) is introduced due to technical reasons in order to work out asymptotics.

Estimator of σ^2

- We propose

$$\tilde{\sigma}_n^2 = \int_{-1/h}^{1/h} \max\{\min\{M_n, \log(|\phi_{emp}(t)|)\}, -M_n\} v^h(t) dt \quad (4)$$

  as an estimator of σ^2.
- Here M_n denotes a sequence of positive numbers diverging to infinity at a suitable rate.
- The truncation in (4) is introduced due to technical reasons in order to work out asymptotics.

Estimator of σ^2

- We propose

$$\tilde{\sigma}_n^2 = \int_{-1/h}^{1/h} \max\{\min\{M_n, \log(|\phi_{emp}(t)|)\}, -M_n\} v^h(t) dt \quad (4)$$

  as an estimator of σ^2.
- Here M_n denotes a sequence of positive numbers diverging to infinity at a suitable rate.
- The truncation in (4) is introduced due to technical reasons in order to work out asymptotics.

Conditions on v^h and h

Condition on v^h

Let the kernel $v^h(t) = h^3 v(ht)$, where v is continuous and real-valued, $\operatorname{supp} v = [-1, 1]$ and

$$\int_{-1}^{1} v(t)dt = 0, \quad \int_{-1}^{1} \left(-\frac{t^2}{2}\right) v(t)dt = 1, \quad v(t) = O(t^\beta) \text{ as } t \to 0.$$

Condition on h

Let the bandwidth h depend on n and be such that $h_n = (\eta \log n)^{-1/2}$ with $0 < \eta < \Sigma^{-2}$.

Conditions on v^h and h

Condition on v^h

Let the kernel $v^h(t) = h^3 v(ht)$, where v is continuous and real-valued, $\operatorname{supp} v = [-1, 1]$ and

$$\int_{-1}^{1} v(t)dt = 0, \quad \int_{-1}^{1} \left(-\frac{t^2}{2}\right) v(t)dt = 1, \quad v(t) = O(t^\beta) \text{ as } t \to 0.$$

Condition on h

Let the bandwidth h depend on n and be such that $h_n = (\eta \log n)^{-1/2}$ with $0 < \eta < \Sigma^{-2}$.

Condition on M_n

Condition on M_n

Let the truncating sequence $M = (M_n)_{n \geq 1}$ be such that $M_n = h^{-2} \log \log n$.

Condition on M_n

Condition on M_n

Let the truncating sequence $M = (M_n)_{n \geq 1}$ be such that $M_n = h^{-2} \log\log n$.

Asymptotics of $\tilde{\sigma}_n^2$

Uniform consistency of $\tilde{\sigma}_n^2$

We have

$$\sup_{|\gamma|\leq\Gamma} \sup_{\sigma\in(0,\Sigma]} \sup_{\rho\in W(\beta,L,\Lambda,K)} \mathrm{E}\,[(\tilde{\sigma}_n^2-\sigma^2)^2] \lesssim (\log n)^{-\beta-3}.$$

Asymptotics of $\tilde{\sigma}_n^2$

Uniform consistency of $\tilde{\sigma}_n^2$

We have

$$\sup_{|\gamma|\leq\Gamma}\ \sup_{\sigma\in(0,\Sigma]}\ \sup_{\rho\in W(\beta,L,\Lambda,K)} \mathrm{E}\left[(\tilde{\sigma}_n^2-\sigma^2)^2\right]\lesssim(\log n)^{-\beta-3}.$$

Outline

Asymptotics of $\tilde{\sigma}_n^2$

- Estimators of λ and γ can be constructed via similar methods. Results comparable to that for $\tilde{\sigma}_n^2$ were obtained for these estimators $\tilde{\lambda}_n$ and $\tilde{\gamma}_n$.
- In particular, the obtained convergence rates were again logarithmic, albeit slower than that for $\tilde{\sigma}_n^2$.
- We do not go into details due to the lack of time.

Asymptotics of $\tilde{\sigma}_n^2$

- Estimators of λ and γ can be constructed via similar methods. Results comparable to that for $\tilde{\sigma}_n^2$ were obtained for these estimators $\tilde{\lambda}_n$ and $\tilde{\gamma}_n$.
- In particular, the obtained convergence rates were again logarithmic, albeit slower than that for $\tilde{\sigma}_n^2$.
- We do not go into details due to the lack of time.

Asymptotics of $\tilde{\sigma}_n^2$

- Estimators of λ and γ can be constructed via similar methods. Results comparable to that for $\tilde{\sigma}_n^2$ were obtained for these estimators $\tilde{\lambda}_n$ and $\tilde{\gamma}_n$.
- In particular, the obtained convergence rates were again logarithmic, albeit slower than that for $\tilde{\sigma}_n^2$.
- We do not go into details due to the lack of time.

Outline

Another inversion of ϕ_X

- Solving for ϕ_ρ in (1), we get

$$\phi_\rho(t) = \operatorname{Log}\left(\frac{\phi_X(t)}{e^{i\gamma t} e^{-\lambda} e^{-\sigma^2 t^2/2}} \right). \tag{5}$$

  Here Log denotes the *distinguished* logarithm.

- By Fourier inversion

$$\rho(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \operatorname{Log}\left(\frac{\phi_X(t)}{e^{i\gamma t} e^{-\lambda} e^{-\sigma^2 t^2/2}} \right) dt.$$

Another inversion of ϕ_X

- Solving for ϕ_ρ in (1), we get

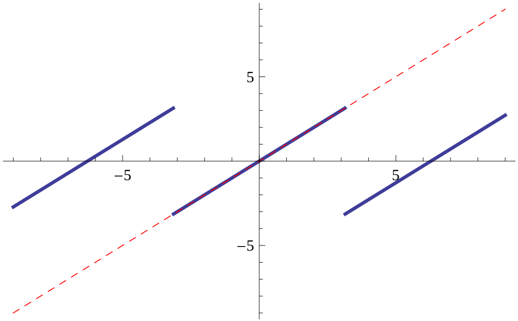
$$\phi_\rho(t) = \operatorname{Log}\left(\frac{\phi_X(t)}{e^{i\gamma t}e^{-\lambda}e^{-\sigma^2 t^2/2}}\right). \tag{5}$$

  Here Log denotes the *distinguished* logarithm.

- By Fourier inversion

$$\rho(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} \operatorname{Log}\left(\frac{\phi_X(t)}{e^{i\gamma t}e^{-\lambda}e^{-\sigma^2 t^2/2}}\right) dt.$$

Logarithm of exp(it)



Estimator ϕ_X

- Let k be a symmetric kernel with Fourier transform ϕ_k supported on $[-1, 1]$ and nonzero there, and let $h > 0$ be a bandwidth.
- Since the characteristic function ϕ_X is integrable, there exists a density q of X, and moreover, it is continuous and bounded.
- This density can be estimated by a kernel density estimator

$$q_n(x) = \frac{1}{nh} \sum_{j=1}^{n} k\left(\frac{x - X_j}{h}\right).$$

- Its characteristic function $\phi_{emp}(t)\phi_k(ht)$ will then serve as an estimator of $\phi_X(t)$.

Estimator ϕ_X

- Let k be a symmetric kernel with Fourier transform ϕ_k supported on $[-1,1]$ and nonzero there, and let $h > 0$ be a bandwidth.
- Since the characteristic function ϕ_X is integrable, there exists a density q of X, and moreover, it is continuous and bounded.
- This density can be estimated by a kernel density estimator

$$q_n(x) = \frac{1}{nh} \sum_{j=1}^{n} k\left(\frac{x - X_j}{h}\right).$$

- Its characteristic function $\phi_{emp}(t)\phi_k(ht)$ will then serve as an estimator of $\phi_X(t)$.

Estimator ϕ_X

- Let k be a symmetric kernel with Fourier transform ϕ_k supported on $[-1, 1]$ and nonzero there, and let $h > 0$ be a bandwidth.
- Since the characteristic function ϕ_X is integrable, there exists a density q of X, and moreover, it is continuous and bounded.
- This density can be estimated by a kernel density estimator

$$q_n(x) = \frac{1}{nh} \sum_{j=1}^{n} k\left(\frac{x - X_j}{h}\right).$$

- Its characteristic function $\phi_{emp}(t)\phi_k(ht)$ will then serve as an estimator of $\phi_X(t)$.

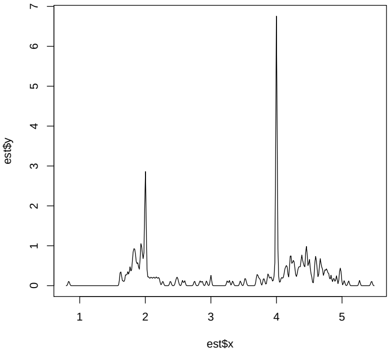
Estimator ϕ_X

- Let k be a symmetric kernel with Fourier transform ϕ_k supported on $[-1,1]$ and nonzero there, and let $h>0$ be a bandwidth.
- Since the characteristic function ϕ_X is integrable, there exists a density q of X, and moreover, it is continuous and bounded.
- This density can be estimated by a kernel density estimator

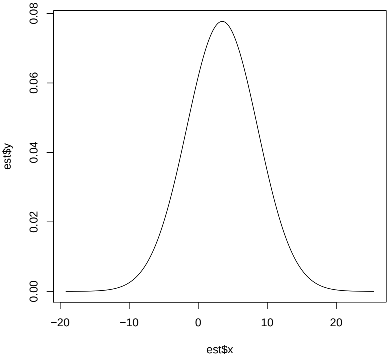
$$q_n(x) = \frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{x-X_j}{h}\right).$$

- Its characteristic function $\phi_{emp}(t)\phi_k(ht)$ will then serve as an estimator of $\phi_X(t)$.

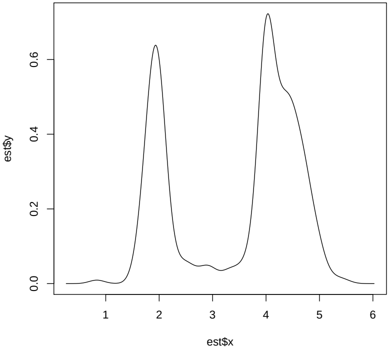
Undersmoothed KDE



Oversmoothed KDE



Just right KDE



Naive estimator of ρ

- For those ω's from the sample space Ω, for which the distinguished logarithm in the integral below is well-defined, ρ can be estimated by the plug-in type estimator,

$$\rho_n(x) = \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-itx} \operatorname{Log}\left(\frac{\phi_{emp}(t)\phi_k(ht)}{e^{i\tilde{\gamma}_n t} e^{-\tilde{\lambda}_n} e^{-\tilde{\sigma}_n^2 t^2/2}} \right) dt,$$

  while for those ω's, for which the distinguished logarithm cannot be defined, we can assign an arbitrary value to $\rho_n(x)$, e.g. zero.

- Notice that the estimator is real-valued, which can be seen by changing the integration variable from t into $-t$.

Naive estimator of ρ

- For those ω's from the sample space Ω, for which the distinguished logarithm in the integral below is well-defined, ρ can be estimated by the plug-in type estimator,

$$\rho_n(x) = \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-itx} \operatorname{Log} \left(\frac{\phi_{emp}(t)\phi_k(ht)}{e^{i\tilde{\gamma}_n t} e^{-\tilde{\lambda}_n} e^{-\tilde{\sigma}_n^2 t^2/2}} \right) dt,$$

  while for those ω's, for which the distinguished logarithm cannot be defined, we can assign an arbitrary value to $\rho_n(x)$, e.g. zero.
- Notice that the estimator is real-valued, which can be seen by changing the integration variable from t into $-t$.

Estimator of ρ

We need to introduce truncation in the definition of ρ_n

$$
\begin{aligned}
\hat{\rho}_n(x) = & -i\tilde{\gamma}_n \frac{1}{2\pi}\int_{-1/h}^{1/h} e^{-itx} t dt \\
& + \tilde{\lambda}_n \frac{1}{2\pi}\int_{-1/h}^{1/h} e^{-itx} dt + \frac{\tilde{\sigma}_n^2}{2}\frac{1}{2\pi}\int_{-1/h}^{1/h} e^{-itx} t^2 dt \\
& + \frac{1}{2\pi}\int_{-1/h}^{1/h} e^{-itx} \max\{\min\{M_n, \log(|\phi_{emp}(t)\phi_k(ht)|)\}, -M_n\} dt \\
& + i\frac{1}{2\pi}\int_{-1/h}^{1/h} e^{-itx} \max\{\min\{M_n, \arg(\phi_{emp}(t)\phi_k(ht))\}, -M_n\} dt.
\end{aligned}
$$

Asymptotics of $\hat{\rho}_n$

Risk bound for $\hat{\rho}_n$

If $k(x) = \sin x/(\pi x)$, the sinc kernel, then we have

$$\sup_{|\gamma|\leq\Gamma} \sup_{\sigma\in(0,\Sigma]} \sup_{\rho\in W^*_{sym}(\beta,L,C,\Lambda,K)} \mathrm{MISE}[\hat{\rho}_n] \lesssim (\log n)^{-\beta},$$

where $W^*_{sym}(\beta, L, C, \Lambda, K)$ denotes the class of Lévy densities ρ, such that $\rho \in W(\beta, L, \Lambda, K)$, ρ is symmetric, and additionally

$$\int_{-\infty}^{\infty} |t|^{2\beta} |\phi_f(t)|^2 dt \leq C.$$

Asymptotics of $\hat{\rho}_n$

Risk bound for $\hat{\rho}_n$

If $k(x) = \sin x/(\pi x)$, the sinc kernel, then we have

$$\sup_{|\gamma|\leq\Gamma} \sup_{\sigma\in(0,\Sigma]} \sup_{\rho\in W^*_{sym}(\beta,L,C,\Lambda,K)} \text{MISE}[\hat{\rho}_n] \lesssim (\log n)^{-\beta},$$

where $W^*_{sym}(\beta, L, C, \Lambda, K)$ denotes the class of Lévy densities ρ, such that $\rho \in W(\beta, L, \Lambda, K)$, ρ is symmetric, and additionally

$$\int_{-\infty}^{\infty} |t|^{2\beta} |\phi_f(t)|^2 dt \leq C.$$

Outline

Can we do better?

- The logarithmic convergence rate for estimation of ρ can be easily understood on an intuitive level when comparing our problem to the deconvolution density estimation.
- In the latter case it is well-known that if the distribution of the error is normal, and if the class of the target densities is sufficiently large, e.g. some Hölder class, the minimax convergence rate for estimation of the target density will be logarithmic for both the mean squared error and mean integrated squared error as measures of risk.
- We will provide a similar result for the problem of estimation of a Lévy density ρ.

Can we do better?

- The logarithmic convergence rate for estimation of ρ can be easily understood on an intuitive level when comparing our problem to the deconvolution density estimation.
- In the latter case it is well-known that if the distribution of the error is normal, and if the class of the target densities is sufficiently large, e.g. some Hölder class, the minimax convergence rate for estimation of the target density will be logarithmic for both the mean squared error and mean integrated squared error as measures of risk.
- We will provide a similar result for the problem of estimation of a Lévy density ρ.

Can we do better?

- The logarithmic convergence rate for estimation of ρ can be easily understood on an intuitive level when comparing our problem to the deconvolution density estimation.
- In the latter case it is well-known that if the distribution of the error is normal, and if the class of the target densities is sufficiently large, e.g. some Hölder class, the minimax convergence rate for estimation of the target density will be logarithmic for both the mean squared error and mean integrated squared error as measures of risk.
- We will provide a similar result for the problem of estimation of a Lévy density ρ.

Actually we cannot do better

Denote by T an arbitrary Lévy triplet (γ, σ^2, ρ), such that $|\gamma| \leq \Gamma, \sigma \in (0, \Sigma], \lambda \in (0, \Lambda]$ and let

$$\int_{-\infty}^{\infty} |t|^{2\beta} |\phi_f(t)|^2 dt \leq C$$

for $\beta \geq 1/2$. Let $\mathcal{T}$ be a collection of all such triplets. Then

$$\inf_{\widetilde{\rho}_n} \sup_{\mathcal{T}} \mathrm{MISE}[\widetilde{\rho}_n] \gtrsim (\log n)^{-\beta},$$

where the infimum is taken over all estimators $\widetilde{\rho}_n$ based on observations $X_1, \ldots, X_n$.

Outline

No formal proofs, but some intuition

- It is expected that lower bounds of the logarithmic order can be obtained for estimation of γ, σ^2 and λ as well.
- Such results are actually not surprising.
- E.g. recall for σ^2 comparable results from Butucea and Matias (2005) for estimation of the error variance in the supersmooth deconvolution problem.
- Another paper containing examples of the breakdown of the usual (in regular parametric models) root-n convergence rate for estimation of a finite-dimensional parameter is Ishwaran (1999).

No formal proofs, but some intuition

- It is expected that lower bounds of the logarithmic order can be obtained for estimation of γ, σ^2 and λ as well.
- Such results are actually not surprising.
- E.g. recall for σ^2 comparable results from Butucea and Matias (2005) for estimation of the error variance in the supersmooth deconvolution problem.
- Another paper containing examples of the breakdown of the usual (in regular parametric models) root-n convergence rate for estimation of a finite-dimensional parameter is Ishwaran (1999).

No formal proofs, but some intuition

- It is expected that lower bounds of the logarithmic order can be obtained for estimation of γ, σ^2 and λ as well.
- Such results are actually not surprising.
- E.g. recall for σ^2 comparable results from Butucea and Matias (2005) for estimation of the error variance in the supersmooth deconvolution problem.
- Another paper containing examples of the breakdown of the usual (in regular parametric models) root-n convergence rate for estimation of a finite-dimensional parameter is Ishwaran (1999).

No formal proofs, but some intuition

- It is expected that lower bounds of the logarithmic order can be obtained for estimation of γ, σ^2 and λ as well.
- Such results are actually not surprising.
- E.g. recall for σ^2 comparable results from Butucea and Matias (2005) for estimation of the error variance in the supersmooth deconvolution problem.
- Another paper containing examples of the breakdown of the usual (in regular parametric models) root-n convergence rate for estimation of a finite-dimensional parameter is Ishwaran (1999).