Two problems of statistical estimation for stochastic processes

Samvel B. GASPARYAN

PhD advisor Yu.A. Kutoyants

Université du Maine, Le Mans, France
Yerevan State University, Yerevan, Armenia

GasparyanSB@gmail.com

12 December, 2016
Le Mans

Outline

Inhomogeneous Poisson process

- In non-parametric estimation the unknown object is a function.

Inhomogeneous Poisson process

- In non-parametric estimation the unknown object is a function.
- We observe a periodic Poisson process with a known period τ

$$X^T = \{X(t),\, t \in [0, T]\},\ T = n\tau.$$

Inhomogeneous Poisson process

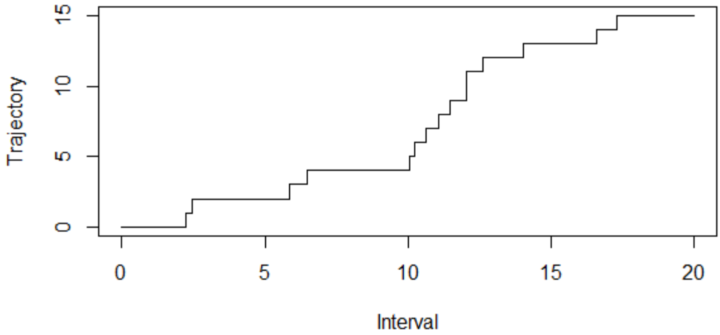
- In non-parametric estimation the unknown object is a function.
- We observe a periodic Poisson process with a known period τ

$$X^T = \{X(t),\ t \in [0, T]\},\ T = n\tau.$$

- $X(0) = 0$, has independent increments and there exists a non-negative, increasing function $\Lambda(t)$ s.t. for all $t \in [0, T]$

$$\mathbf{P}(X(t) = k) = \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)},\ k = 0, 1, \cdots .$$

Trajectory of a Poisson (counting) process



Models description

- In the definition

$$\mathbf{P}(X(t) = k) = \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)}, \; k = 0, 1, \cdots .$$

Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots .$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t)=\Lambda(t)$.

Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots.$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t)=\Lambda(t)$.
- We consider the case were $\Lambda(\cdot)$ is absolutely continuous $\Lambda(t)=\int_0^t \lambda(s)\mathrm{d}s$.

Models description

- In the definition

$$\mathbf{P}(X(t)=k)=\frac{[\Lambda(t)]^k}{k!}e^{-\Lambda(t)},\ k=0,1,\cdots.$$

- The function $\Lambda(\cdot)$ is called the mean function since $\mathbf{E}X(t)=\Lambda(t)$.
- We consider the case were $\Lambda(\cdot)$ is absolutely continuous $\Lambda(t)=\int_0^t \lambda(s)\mathrm{d}s$.
- The non-negative function $\lambda(\cdot)$ is called the intensity function and the periodicity of a Poisson process means the periodicity of its intensity function

$$\lambda(t)=\lambda(t+k\tau),\ t\in[0,\tau],\ k\in\mathcal{Z}_+.$$

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1, \cdots, n,$$

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1,\cdots,n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1, \cdots, n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.
- Estimation problems of $\lambda(t),\ t \in [0,\tau]$ and $\Lambda(t),\ t \in [0,\tau]$ are completely different.

Mean and the Intensity functions

- With the notations

$$X_j(t) = X((j-1)\tau + t) - X((j-1)\tau),\ t \in [0,\tau],$$
$$X_j = \{X_j(t),\ t \in [0,\tau]\},\ j = 1, \cdots, n,$$

- We get an i.i.d. model $X^n = (X_1, X_2, \cdots, X_n)$ generated from a Poisson process.
- Estimation problems of $\lambda(t),\ t \in [0,\tau]$ and $\Lambda(t),\ t \in [0,\tau]$ are completely different.
- We would like to have Hájek-Le Cam type lower bounds for function estimation.

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0, \tau]\}$.

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0,\tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0, \tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.
- To assess the quality of an estimator we use the MISE

$$\mathbf{E}_\Lambda ||\bar{\Lambda}_n - \Lambda||^2 = \mathbf{E}_\Lambda \int_0^\tau (\bar{\Lambda}_n(t) - \Lambda(t))^2 \mathrm{d}t.$$

Mean function estimation

- We consider the estimation problem of the mean function $\{\Lambda(t),\ t \in [0, \tau]\}$.
- Each measurable function $\bar{\Lambda}_n(t) = \bar{\Lambda}_n(t, X^n)$ of observations is an estimator for $\Lambda(t)$.
- To assess the quality of an estimator we use the MISE

$$\mathbf{E}_\Lambda ||\bar{\Lambda}_n - \Lambda||^2 = \mathbf{E}_\Lambda \int_0^\tau (\bar{\Lambda}_n(t) - \Lambda(t))^2 \mathrm{d}t.$$

- The simplest estimator is the *empirical mean function*

$$\hat{\Lambda}_n(t) = \frac{1}{n} \sum_{j=1}^n X_j(t),\ t \in [0, \tau].$$

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_{\Lambda}||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^{\tau} \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.
2. The asymptotic error (which is non-asymptotic) of the EMF is $\int_0^{\tau} \Lambda(t)\mathrm{d}t$.

The basic equality for the EMF

- The following basic equality for the EMF implies two things

$$\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 = \int_0^\tau \Lambda(t)\mathrm{d}t.$$

1. The EMF is consistent with the classical rate of convergence $\sqrt{n}$.
2. The asymptotic error (which is non-asymptotic) of the EMF is $\int_0^\tau \Lambda(t)\mathrm{d}t$.

- Can we have better rate of convergence or smaller asymptotic error for an estimator?

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$, (Kutoyants' results [Kutoyants, 2016])

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$, (Kutoyants' results [Kutoyants, 2016])
  1. for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta \downarrow 0} \varliminf_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda (\sqrt{n}(\bar{\Lambda}_n - \Lambda))^2 \geq \Lambda^*(t), \; t \in [0, \tau],$$

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$, (Kutoyants' results [Kutoyants, 2016])
  1. for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta\downarrow 0} \underline{\lim}_{n\to+\infty} \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda(\sqrt{n}(\bar{\Lambda}_n - \Lambda))^2 \geq \Lambda^*(t),\ t\in[0,\tau],$$

  2.

$$\lim_{\delta\downarrow 0} \underline{\lim}_{n\to+\infty} \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t)\mathrm{d}t,$$

  with $V_\delta = \{\Lambda : \sup_{0\leq t\leq\tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}$, $\delta > 0$.

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$, (Kutoyants' results [Kutoyants, 2016])
  1. for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta \downarrow 0} \underline{\lim}_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda (\sqrt{n}(\bar{\Lambda}_n - \Lambda))^2 \geq \Lambda^*(t),\ t \in [0, \tau],$$

  2.

$$\lim_{\delta \downarrow 0} \underline{\lim}_{n \to +\infty} \sup_{\Lambda \in V_\delta} \mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t) \mathrm{d}t,$$

  with $V_\delta = \{\Lambda : \sup_{0 \leq t \leq \tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}$, $\delta > 0$.
- Reformulation

$$\underline{\lim}_{n \to +\infty} \sup_{\Lambda \in \mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) \geq 0.$$

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar\Lambda_n(t)$, (Kutoyants' results [Kutoyants, 2016])
  1. for all estimators $\bar\Lambda_n(t)$
  $$\lim_{\delta\downarrow 0}\ \underline{\lim}_{n\to+\infty}\ \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda(\sqrt{n}(\bar\Lambda_n-\Lambda))^2 \geq \Lambda^*(t),\ t\in[0,\tau],$$
  2. $$\lim_{\delta\downarrow 0}\ \underline{\lim}_{n\to+\infty}\ \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda||\sqrt{n}(\bar\Lambda_n-\Lambda)||^2 \geq \int_0^\tau \Lambda^*(t)\mathrm{d}t,$$

  with $V_\delta=\{\Lambda:\ \sup_{0\leq t\leq\tau}|\Lambda(t)-\Lambda^*(t)|\leq\delta\}$, $\delta>0$.
- Reformulation
$$\underline{\lim}_{n\to+\infty}\ \sup_{\Lambda\in\mathcal{F}}\left(\mathbf{E}_\Lambda||\sqrt{n}(\bar\Lambda_n-\Lambda)||^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq 0.$$
- $\mathcal{F}\subset \mathrm{L}_2[0,\tau]$ is a sufficiently "rich", bounded set.

As. efficiency of the EMF

- The EMF is asymptotically efficient among all estimators $\bar{\Lambda}_n(t)$, (Kutoyants' results [Kutoyants, 2016])
  1. for all estimators $\bar{\Lambda}_n(t)$

$$\lim_{\delta\downarrow 0} \underline{\lim}_{n\to+\infty} \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda(\sqrt{n}(\bar{\Lambda}_n - \Lambda))^2 \geq \Lambda^*(t),\ t\in[0,\tau],$$

  2.

$$\lim_{\delta\downarrow 0} \underline{\lim}_{n\to+\infty} \sup_{\Lambda\in V_\delta} \mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 \geq \int_0^\tau \Lambda^*(t)\mathrm{d}t,$$

  with $V_\delta = \{\Lambda : \sup_{0\leq t\leq\tau} |\Lambda(t) - \Lambda^*(t)| \leq \delta\}$, $\delta > 0$.
- Reformulation

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) \geq 0.$$

- $\mathcal{F}\subset \mathrm{L}_2[0,\tau]$ is a sufficiently "rich", bounded set.
- Can we have other asymptotically efficient estimators?

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.
- In other words, it depends on the choice of the set $\mathcal{F}$ in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0.$$

Efficient estimators

- Existence of other as. efficient estimators depends on the regularity conditions imposed on unknown $\Lambda(\cdot)$.
- In other words, it depends on the choice of the set $\mathcal{F}$ in

$$\underline{\lim}_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_{\Lambda} ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^{\tau} \Lambda(t)\mathrm{d}t \right) \geq 0.$$

- Demanding existence of derivatives of higher order of the unknown function, we can enlarge the class of as. efficient estimators.

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : \|\Lambda - \Lambda^*\|^2 \leq R,\ \Lambda^*(\tau) = \Lambda(\tau)\}.$$

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : \|\Lambda - \Lambda^*\|^2 \leq R, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- Consider a kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

First results

- At first, consider the L_2 ball with a center Λ^*

$$\mathcal{B}(R) = \{\Lambda : ||\Lambda - \Lambda^*||^2 \leq R, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

- Consider a kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- Kernels satisfy

$$K_n(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K_n(u)\mathrm{d}u = 1,\ n \in \mathcal{N},$$

and we continue them τ periodically on the whole real line **R**

$$K_n(u) = K_n(-u),\quad K_n(u) = K_n(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

Kernel-type estimator

- Consider the trigonometric basis in $\mathrm{L}_2[0,\tau]$

$$\phi_1(t)=\sqrt{\frac{1}{\tau}},\ \phi_{2l}(t)=\sqrt{\frac{2}{\tau}}\cos\frac{2\pi l}{\tau}t,\ \phi_{2l+1}(t)=\sqrt{\frac{2}{\tau}}\sin\frac{2\pi l}{\tau}t.$$

Kernel-type estimator

- Consider the trigonometric basis in $\mathrm{L}_2[0,\tau]$

$$\phi_1(t)=\sqrt{\frac{1}{\tau}},\ \phi_{2l}(t)=\sqrt{\frac{2}{\tau}}\cos\frac{2\pi l}{\tau}t,\ \phi_{2l+1}(t)=\sqrt{\frac{2}{\tau}}\sin\frac{2\pi l}{\tau}t.$$

- Coefficients of the kernel-type estimator w.r.t. this basis

$$\tilde{\Lambda}_{1,n}=\hat{\Lambda}_{1,n},\ \tilde{\Lambda}_{2l,n}=\sqrt{\frac{\tau}{2}}K_{2l,n}(\hat{\Lambda}_{2l,n}-\Lambda^*_{2l})+\Lambda^*_{2l},$$

$$\tilde{\Lambda}_{2l+1,n}=\sqrt{\frac{\tau}{2}}K_{2l,n}(\hat{\Lambda}_{2l+1,n}-\Lambda^*_{2l+1})+\Lambda^*_{2l+1},\ l\in\mathcal{N},$$

where $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF.

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^{\tau} K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- with a kernel satisfying the condition

$$n \sup_{l \geq 1} \left| \sqrt{\frac{\tau}{2}} K_{2l,n} - 1 \right|^2 \longrightarrow 0,$$

Efficiency over a ball

- A kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^\tau K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

- with a kernel satisfying the condition

$$n \sup_{l\geq 1} \left| \sqrt{\frac{\tau}{2}} K_{2l,n} - 1 \right|^2 \longrightarrow 0,$$

- is asymptotically efficient over a ball

$$\lim_{n\to+\infty} \sup_{\Lambda\in\mathcal{B}(R)} \left(\mathbf{E}_\Lambda || \sqrt{n}(\tilde{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) = 0.$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function and relax the conditions on the kernel:

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function and relax the conditions on the kernel:
  1. the mean function belongs to $\Sigma(R)$

$$\Sigma(R) = \{\Lambda : \ ||\lambda - \lambda^*||^2 \leq R, \ \Lambda^*(\tau) = \Lambda(\tau)\}.$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function and relax the conditions on the kernel:
  1. the mean function belongs to $\Sigma(R)$

$$\Sigma(R) = \{\Lambda : \ ||\lambda - \lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

  2. the kernel satisfies

$$n \sup_{l \geq 1} \left| \frac{\sqrt{\frac{\tau}{2}} K_{2l,n} - 1}{\frac{2\pi l}{\tau}} \right|^2 \longrightarrow 0,$$

Efficiency over a compact set

- Now we impose additional conditions of regularity on the unknown mean function and relax the conditions on the kernel:
  1. the mean function belongs to $\Sigma(R)$

     $$\Sigma(R) = \{\Lambda : \ ||\lambda - \lambda^*||^2 \leq R, \, \Lambda^*(\tau) = \Lambda(\tau)\}.$$

  2. the kernel satisfies

     $$n \sup_{l \geq 1} \left| \frac{\sqrt{\frac{\tau}{2}} K_{2l,n} - 1}{\frac{2\pi l}{\tau}} \right|^2 \longrightarrow 0,$$

- The kernel-type estimator is asymptotically efficient over $\Sigma(R)$

$$\lim_{n \to +\infty} \sup_{\Lambda \in \Sigma(R)} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\tilde{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t) \mathrm{d}t \right) = 0.$$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u), \quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u), \quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty.$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u), \quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty$.
- Then, the following kernels satisfy the previous condition

$$K_n(u) = \frac{1}{h_n} K\left(\frac{u}{h_n}\right) \mathbb{1}\left\{|u| \leq \frac{\tau}{2} h_n\right\}$$

Example of another as. effective estimator

- Consider a kernel

$$K(u) \geq 0,\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} K(u)\mathrm{d}u = 1,$$

$$K(u) = K(-u), \quad K(u) = K(u + k\tau),\ u \in \left[-\frac{\tau}{2}, \frac{\tau}{2}\right],\ k \in \mathcal{Z}.$$

- A sequence $0 \leq h_n \leq 1$ be s.t. $h_n^2 n \longrightarrow 0,\ n \to +\infty$.
- Then, the following kernels satisfy the previous condition

$$K_n(u) = \frac{1}{h_n} K\left(\frac{u}{h_n}\right) \mathbb{1}\left\{|u| \leq \frac{\tau}{2} h_n\right\}$$

- the corresponding kernel-type estimator

$$\tilde{\Lambda}_n(t) = \int_0^{\tau} K_n(s-t)(\hat{\Lambda}_n(s) - \Lambda_*(s))\mathrm{d}s + \Lambda_*(t).$$

is as. efficient over $\Sigma(R)$.

Second order efficiency

- How to compare as. efficient estimators?

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) \geq 0,$$

- that is, the sequence $\gamma_n \longrightarrow +\infty$ s.t.

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \gamma_n \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) \geq \mathrm{C},$$

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty}\sup_{\Lambda\in\mathcal{F}}\left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n-\Lambda)||^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq 0,$$

- that is, the sequence $\gamma_n\longrightarrow+\infty$ s.t.

$$\varliminf_{n\to+\infty}\sup_{\Lambda\in\mathcal{F}}\gamma_n\left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n-\Lambda)||^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq \mathrm{C},$$

- Then, to construct an estimator which attains this bound.

Second order efficiency

- How to compare as. efficient estimators?
- The first step would be to find the rate of convergence in

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq 0,$$

- that is, the sequence $\gamma_n \longrightarrow +\infty$ s.t.

$$\varliminf_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}} \gamma_n \left(\mathbf{E}_\Lambda ||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t \right) \geq \mathrm{C},$$

- Then, to construct an estimator which attains this bound.
- Calculate the constant C.

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].
- Asymptotic efficiency in non-parametric estimation problems was done for the first time in [Pinsker M.S., 1980],

Related works

- Second order estimation was introduced by [Golubev G.K. and Levit B.Ya., 1996] in the distribution function estimation problem.
- For other models second order efficiency was proved by [Dalalyan A.S. and Kutoyants Yu.A., 2004], [Golubev G.K. and Härdle W., 2000].
- Asymptotic efficiency in non-parametric estimation problems was done for the first time in [Pinsker M.S., 1980],
- where the analogue of the inverse of the Fisher information in non-parametric estimation problem was calculated (Pinsker's constant).

Main theorems (joint work with Yu.A. Kutoyants)

- Introduce

$$\mathcal{F}_m^{per}(R,S)=\left\{\Lambda(\cdot):\int_0^\tau[\lambda^{(m-1)}(t)]^2\mathrm{d}t\leq R,\,\Lambda(0)=0,\,\Lambda(\tau)=S\right\}$$

where $R>0$, $S>0, m>1$, are given constants.

Main theorems (joint work with Yu.A. Kutoyants)

- Introduce

$$\mathcal{F}_m^{per}(R,S)=\left\{\Lambda(\cdot):\int_0^\tau[\lambda^{(m-1)}(t)]^2\mathrm{d}t\leq R,\,\Lambda(0)=0,\,\Lambda(\tau)=S\right\}$$

where $R>0$, $S>0$, $m>1$, are given constants.

- For all estimators $\bar{\Lambda}_n(t)$ of the mean function $\Lambda(t)$, following lower bound holds

$$\varliminf_{n\to+\infty}\sup_{\Lambda\in\mathcal{F}_m^{per}(R,S)} n^{\frac{1}{2m-1}}\left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n-\Lambda)||^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq-\Pi,$$

Main theorems (joint work with Yu.A. Kutoyants)

- Introduce

$$\mathcal{F}_m^{per}(R,S)=\left\{\Lambda(\cdot):\ \int_0^\tau [\lambda^{(m-1)}(t)]^2\mathrm{d}t\leq R,\ \Lambda(0)=0,\ \Lambda(\tau)=S\right\}$$

where $R>0$, $S>0$, $m>1$, are given constants.

- For all estimators $\bar\Lambda_n(t)$ of the mean function $\Lambda(t)$, following lower bound holds

$$\varliminf_{n\to+\infty}\ \sup_{\Lambda\in\mathcal{F}_m^{per}(R,S)} n^{\frac{1}{2m-1}}\left(\mathbf{E}_\Lambda||\sqrt{n}(\bar\Lambda_n-\Lambda)||^2-\int_0^\tau\Lambda(t)\mathrm{d}t\right)\geq-\Pi,$$

- where

$$\Pi=\Pi_m(R,S)=(2m-1)R\left(\frac{S}{\pi R}\frac{m}{(2m-1)(m-1)}\right)^{\frac{2m}{2m-1}},$$

plays the role of the Pinsker's constant.

Second order as. efficient estimator

- Consider

$$\Lambda_n^*(t) = \hat{\Lambda}_{0,n}\phi_0(t) + \sum_{l=1}^{N_n} \tilde{K}_{l,n}\hat{\Lambda}_{l,n}\phi_l(t),$$

where $\{\phi_l\}_{l=0}^{+\infty}$ is the trigonometric cosine basis, $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF w.r.t. this basis and

$$\tilde{K}_{l,n} = \left(1 - \left|\frac{\pi l}{\tau}\right|^m \alpha_n^*\right)_+, \quad \alpha_n^* = \left[\frac{S}{nR}\frac{\tau}{\pi}\frac{m}{(2m-1)(m-1)}\right]^{\frac{m}{2m-1}},$$

$$N_n = \frac{\tau}{\pi}(\alpha_n^*)^{-\frac{1}{m}} \approx \mathrm{C} n^{\frac{1}{2m-1}}, \quad x_+ = \max(x,0),\ x \in \mathbf{R}.$$

Second order as. efficient estimator

- Consider

$$\Lambda_n^*(t) = \hat{\Lambda}_{0,n}\phi_0(t) + \sum_{l=1}^{N_n} \tilde{K}_{l,n}\hat{\Lambda}_{l,n}\phi_l(t),$$

where $\{\phi_l\}_{l=0}^{+\infty}$ is the trigonometric cosine basis, $\hat{\Lambda}_{l,n}$ are the Fourier coefficients of the EMF w.r.t. this basis and

$$\tilde{K}_{l,n} = \left(1 - \left|\frac{\pi l}{\tau}\right|^m \alpha_n^*\right)_+, \quad \alpha_n^* = \left[\frac{S}{nR}\frac{\tau}{\pi}\frac{m}{(2m-1)(m-1)}\right]^{\frac{m}{2m-1}},$$

$$N_n = \frac{\tau}{\pi}(\alpha_n^*)^{-\frac{1}{m}} \approx \mathrm{C}n^{\frac{1}{2m-1}}, \quad x_+ = \max(x,0),\ x \in \mathbf{R}.$$

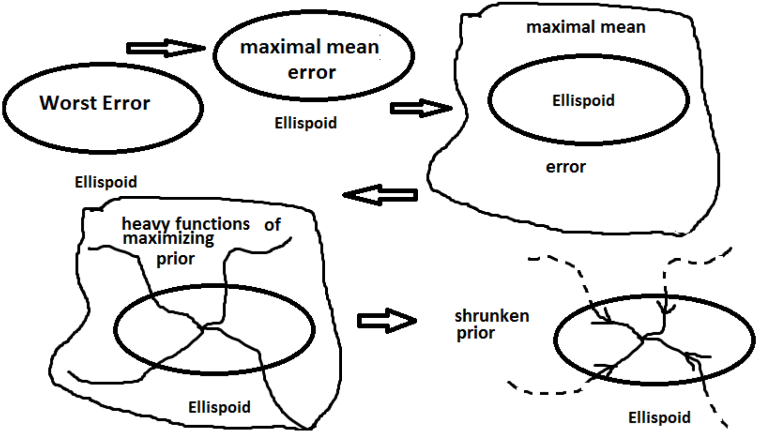
- The estimator $\Lambda_n^*(t)$ attains the lower bound described above, that is,

$$\lim_{n\to+\infty} \sup_{\Lambda\in\mathcal{F}_m(R,S)} n^{\frac{1}{2m-1}} \left(\mathbf{E}_\Lambda||\sqrt{n}(\bar{\Lambda}_n - \Lambda)||^2 - \int_0^\tau \Lambda(t)\mathrm{d}t\right) = -\Pi.$$

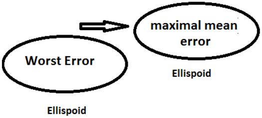
Sketch of the proof (Lower bound)

- Proof consists of several steps:

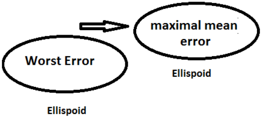
Sketch of the proof (Lower bound)

- Proof consists of several steps:
  - 

First step



First step

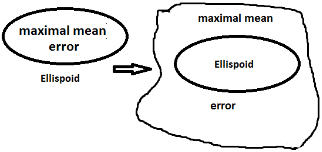


- Reduce the minimax problem to a Bayes risk maximization problem

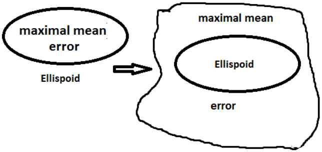
$$\sup_{\Lambda\in\mathcal{F}_m^{(per)}(R,S)} \left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right)\geq$$

$$\sup_{\mathbf{Q}\in\mathcal{P}}\int_{\mathcal{F}_m^{(per)}(R,S)}\left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right)\mathrm{d}\mathbf{Q}.$$

Second step


maximal mean
error
Ellispoid
maximal mean
Ellispoid
error

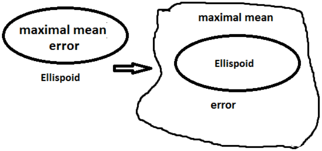
Second step



- In the maximization problem replace the set of probabilities $\mathcal{P}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$

$$\sup_{\mathbf{Q}\in\mathcal{P}(\mathcal{F})} \int_{\mathcal{F}_m^{(per)}(R,S)} \left(\mathbf{E}_\Lambda||\bar{\Lambda}_n-\Lambda||^2-\mathbf{E}_\Lambda||\hat{\Lambda}_n-\Lambda||^2\right) \mathrm{d}\mathbf{Q},$$

Second step

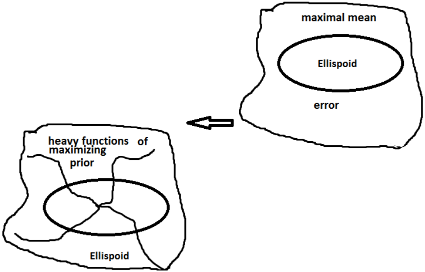


- In the maximization problem replace the set of probabilities $\mathcal{P}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$

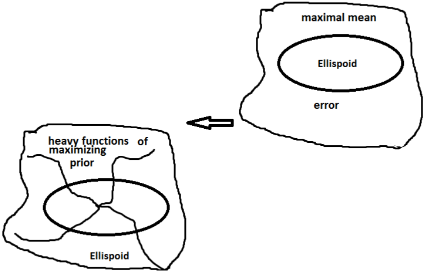
$$\sup_{\mathbf{Q}\in\mathcal{P}(\mathcal{F})} \int_{\mathcal{F}_m^{(per)}(R,S)} \left(\mathbf{E}_\Lambda||\bar{\Lambda}_n - \Lambda||^2 - \mathbf{E}_\Lambda||\hat{\Lambda}_n - \Lambda||^2\right) \mathrm{d}\mathbf{Q},$$

- by the set of probabilities $\mathbf{E}(\mathcal{F})$ concentrated on $\mathcal{F}_m^{(per)}(R,S)$ in mean.

Third step


maximal mean
Ellispoid
error
heavy functions of maximizing prior
Ellispoid

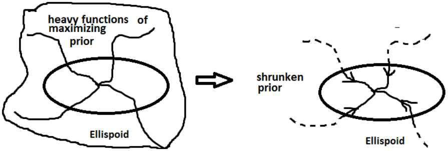
Third step

- 

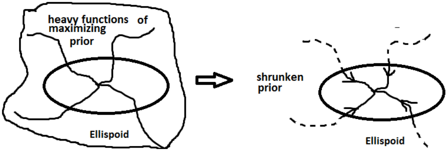
- Replace the ellipsoid by the least favorable parametric family (heavy functions)

$$\sup_{\Lambda_\theta \in \mathcal{F}_m^{(per)}(R,S)} \int_\Theta \left(\mathbf{E}_\theta ||\bar{\Lambda}_n - \Lambda_\theta||^2 - \mathbf{E}_\theta ||\hat{\Lambda}_n - \Lambda_\theta||^2 \right) \mathrm{d}\mathbf{Q}.$$

Fourth step


heavy functions of maximizing prior
Ellispoid
shrunken prior
Ellispoid

Fourth step



- Shrink the heavy functions and the least favorable prior distribution to fit the ellipsoid

$$\mathbf{Q}\{\theta : \Lambda_\theta \notin \mathcal{F}_m^{(per)}(R,S)\} = o(n^{-2}).$$

Sketch of the proof (Upper bound)

- Specificity of L_2 minimax estimation problem is that the efficient estimator is linear

$$\tilde{\Lambda}_{l,n} = K_{l,n}\hat{\Lambda}_{l,n}.$$

Sketch of the proof (Upper bound)

- Specificity of L_2 minimax estimation problem is that the efficient estimator is linear

$$\tilde{\Lambda}_{l,n} = K_{l,n}\hat{\Lambda}_{l,n}.$$

- Since we demanded existence of the $m-$th derivative, then the class of as. efficient estimators is

$$n \sup_{l\geq 1} \left| \frac{K_{l,n} - 1}{\left(\frac{\pi l}{\tau}\right)^{2m}} \right|^2 \longrightarrow 0,$$

Sketch of the proof (Upper bound) 2

- Consider the class

$$\mathrm{K} = \left\{ K_{l,n} : \left| \frac{K_{l,n} - 1}{\left(\frac{\pi l}{\tau}\right)^{2m}} \right|^2 \leq \alpha_n^2 \right\}, \ \alpha_n^2 \longrightarrow 0,$$

Sketch of the proof (Upper bound) 2

- Consider the class

$$\mathrm{K} = \left\{ K_{l,n} : \left| \frac{K_{l,n} - 1}{\left(\frac{\pi l}{\tau}\right)^{2m}} \right|^2 \leq \alpha_n^2 \right\},\ \alpha_n^2 \longrightarrow 0,$$

- $$\tilde{K}_{l,n} = \min_{K_{l,n} \in \mathrm{K}} |K_{l,n}| = \left(1 - \left(\frac{\pi l}{\tau}\right)^m \alpha_n\right)_+ .$$

Sketch of the proof (Upper bound) 2

- Consider the class

$$\mathrm{K} = \left\{ K_{l,n} : \left| \frac{K_{l,n} - 1}{\left(\frac{\pi l}{\tau}\right)^{2m}} \right|^2 \leq \alpha_n^2 \right\}, \; \alpha_n^2 \longrightarrow 0,$$

- $$\tilde{K}_{l,n} = \min_{K_{l,n} \in \mathrm{K}} |K_{l,n}| = \left(1 - \left(\frac{\pi l}{\tau} \right)^m \alpha_n \right)_+ .$$

- Minimizing the second order risk w.r.t. α_n we get

$$\alpha_n^* = \left[\frac{S}{nR} \frac{\tau}{\pi} \frac{m}{(2m-1)(m-1)} \right]^{\frac{m}{2m-1}} \approx \mathrm{C} n^{-\frac{m}{2m-1}} .$$

Sketch of the proof (Upper bound) 2

- Consider the class

$$\mathrm{K} = \left\{ K_{l,n} : \left| \frac{K_{l,n} - 1}{\left(\frac{\pi l}{\tau}\right)^{2m}} \right|^2 \leq \alpha_n^2 \right\}, \; \alpha_n^2 \longrightarrow 0,$$

-

$$\tilde{K}_{l,n} = \min_{K_{l,n} \in \mathrm{K}} |K_{l,n}| = \left(1 - \left(\frac{\pi l}{\tau} \right)^m \alpha_n \right)_+ .$$

- Minimizing the second order risk w.r.t. α_n we get

$$\alpha_n^* = \left[\frac{S}{nR} \frac{\tau}{\pi} \frac{m}{(2m-1)(m-1)} \right]^{\frac{m}{2m-1}} \approx \mathrm{C} n^{-\frac{m}{2m-1}} .$$

- (the condition $n(\alpha_n^*)^2 = \mathrm{C} n^{-\frac{1}{2m-1}} \longrightarrow 0$ is also satisfied).

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$
  1. have to be replaced by $\Lambda(\tau) \leq S$.

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$
  1. have to be replaced by $\Lambda(\tau) \leq S$.
  2. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

$$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \text{ data=signal+"noise"}$$

and the variance of the noise is $\frac{1}{n}\Lambda(t)$.

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$
  1. have to be replaced by $\Lambda(\tau) \leq S$.
  2. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

     $$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \quad \text{data=signal+"noise"}$$

     and the variance of the noise is $\frac{1}{n}\Lambda(t)$.
  3. (Simultaneous estimation of the function and the variance of noise).

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$
  1. have to be replaced by $\Lambda(\tau) \leq S$.
  2. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get
     $$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \text{ data=signal+“noise”}$$
     and the variance of the noise is $\frac{1}{n}\Lambda(t)$.
  3. (Simultaneous estimation of the function and the variance of noise).
- Adaptive estimation: construct an estimator that does not depend on m, S, R.

What can be done or what had to be done?

- The condition $\Lambda(\tau) = S$ in the definition of the set $\mathcal{F}_m^{(per)}(R, S)$
  1. have to be replaced by $\Lambda(\tau) \leq S$.
  2. The last one cannot be thrown out since with a notation $\pi_j(t) = X_j(t) - \Lambda(t)$ we get

     $$\hat{\Lambda}_n(t) = \Lambda(t) + \frac{1}{n}\sum_{j=1}^{n} \pi_j(t), \text{ data=signal+“noise”}$$

     and the variance of the noise is $\frac{1}{n}\Lambda(t)$.
  3. (Simultaneous estimation of the function and the variance of noise).
- Adaptive estimation: construct an estimator that does not depend on m, S, R.
- Consider other models or formulate a general result for non-parametric LAN.

Forward-Backward SDE

- For a given (*Forward*) Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \quad X_0 = x_0,\ t \in [0, T],$$

and functions $f(t, x, y, z), \Phi(x)$,

Forward-Backward SDE

- For a given (*Forward*) Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \quad X_0 = x_0,\ t \in [0, T],$$

  and functions $f(t, x, y, z), \Phi(x)$,

- we are searching a couple of processes (Y_t, Z_t)

Forward-Backward SDE

- For a given (*Forward*) Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \quad X_0 = x_0, \; t \in [0, T],$$

  and functions $f(t, x, y, z), \Phi(x)$,
- we are searching a couple of processes (Y_t, Z_t)
- which satisfies the stochastic differential equation

$$\mathrm{d}Y_t = -f(t, X_t, Y_t, Z_t)\mathrm{d}t + Z_t\mathrm{d}W_t, \; 0 \le t \le T,$$

  with the terminal condition $Y_T = \Phi(X_T)$.
  [Pardoux, Peng, 1992].

Solution of a FBSDE

- If the function $u = u(t, x)$ satisfies the following partial differential equation

$$\frac{\partial u}{\partial t} + S(t,x)\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 u}{\partial x^2} = -f\left(t, x, u, \sigma(t,x)\frac{\partial u}{\partial x}\right),$$

with the terminal condition $u(T, x) = \Phi(x)$,

Solution of a FBSDE

- If the function $u = u(t,x)$ satisfies the following partial differential equation

$$\frac{\partial u}{\partial t} + S(t,x)\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 u}{\partial x^2} = -f\left(t, x, u, \sigma(t,x)\frac{\partial u}{\partial x}\right),$$

with the terminal condition $u(T,x) = \Phi(x)$,

- then the processes

$$Y_t = u(t, X_t),\ Z_t = \sigma(t, X_t)u'_x(t, X_t),$$

satisfy (by the Itô's formula) the following SDE

$$\mathrm{d}Y_t = \left[\frac{\partial u}{\partial t}(t,X_t) + S(t,X_t)\frac{\partial u}{\partial x}(t,X_t) + \frac{1}{2}\sigma^2(t,X_t)\frac{\partial^2 u}{\partial x^2}(t,X_t)\right]\mathrm{d}t$$
$$+ \sigma(t,X_t)\frac{\partial u}{\partial x}(t,X_t)\mathrm{d}W_t,$$

hence $\mathrm{d}Y_t = -f(t, X_t, Y_t, Z_t)\mathrm{d}t + Z_t\mathrm{d}W_t, \quad Y_T = \Phi(X_T).$

Statistical Estimation

- Suppose that the coefficient $\sigma(\cdot)$ of the forward equation depends on an unknown parameter $\vartheta \in \Theta = (\alpha, \beta)$,

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \quad X_0 = x_0,\ t \in [0, T],$$

Statistical Estimation

- Suppose that the coefficient $\sigma(\cdot)$ of the forward equation depends on an unknown parameter $\vartheta \in \Theta = (\alpha, \beta)$,

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \quad X_0 = x_0, \ t \in [0, T],$$

- We have discrete time observations of the solution $\{X_t, \ t \in [0, T]\}$ of the forward SDE

$$\mathbf{X}^n = (X_{t_0}, X_{t_1}, \cdots, X_{t_n}), \ t_j = j\frac{T}{n}, j = 0, 1, \cdots, n.$$

Statistical Estimation

- Suppose that the coefficient $\sigma(\cdot)$ of the forward equation depends on an unknown parameter $\vartheta \in \Theta = (\alpha, \beta)$,

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \quad X_0 = x_0, \ t \in [0, T],$$

- We have discrete time observations of the solution $\{X_t, \ t \in [0, T]\}$ of the forward SDE

$$\mathbf{X}^n = (X_{t_0}, X_{t_1}, \cdots, X_{t_n}), \ t_j = j\frac{T}{n}, j = 0, 1, \cdots, n.$$

- Our goal is, based on this observations, estimate the solution of the FBSDE

$$Y_t = u(\vartheta, t, X_t), \quad Z_t = \sigma(\vartheta, t, X_t)u'_x(\vartheta, t, X_t).$$

which also depends on the unknown parameter.

Heuristics

- We have to construct an estimator $\vartheta^*_{t,n}$ of ϑ, based on the observations up to time t

$$\mathbf{X}^k = (X_{t_0}, X_{t_1}, \cdots, X_{t_k}),\ k = \left[\frac{t}{T}n\right],\ t_k = \frac{k}{n}T,\ \delta = t_j - t_{j-1} = \frac{T}{n},$$

which

Heuristics

- We have to construct an estimator $\vartheta^*_{t,n}$ of ϑ, based on the observations up to time t

$$\mathbf{X}^k = (X_{t_0}, X_{t_1}, \cdots, X_{t_k}),\ k = \left[\frac{t}{T} n\right],\ t_k = \frac{k}{n} T,\ \delta = t_j - t_{j-1} = \frac{T}{n},$$

which

1. have to be easily calculable for every t,

Heuristics

- We have to construct an estimator $\vartheta^*_{t,n}$ of ϑ, based on the observations up to time t

$$\mathbf{X}^k = (X_{t_0}, X_{t_1}, \cdots, X_{t_k}),\ k = \left[\frac{t}{T}n\right],\ t_k = \frac{k}{n}T,\ \delta = t_j - t_{j-1} = \frac{T}{n},$$

which

1. have to be easily calculable for every t,
2. the estimator $Y^*_{t,n} = u(\vartheta^*_{t,n}, t, X_{t_k})$ must have asymptotically the smallest quadratic error

$$\mathbf{E}_\vartheta (Y^*_{t,n} - Y_t)^2 \longrightarrow \min, \quad n \longrightarrow +\infty.$$

Parameter in Diffusion

- How to construct a consistent estimator for the unknown parameter in the diffusion coefficient of a SDE?

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \quad X_0 = x_0, \; t \in [0, T],$$

Parameter in Diffusion

- How to construct a consistent estimator for the unknown parameter in the diffusion coefficient of a SDE?

$$\mathrm{d}X_t = S(t, X_t)\mathrm{d}t + \sigma(\vartheta, t, X_t)\mathrm{d}W_t, \quad X_0 = x_0, \ t \in [0, T],$$

- [Genon-Catalot, Jacod, 1993] We are considering the minimum contrast estimator

$$\hat{\vartheta}_{t,n} = \arg\min_{\vartheta \in \Theta} U_k(\vartheta),$$

where the contrast function is

$$U_k(\vartheta) = \sum_{j=1}^{k} \left[\delta \ln \sigma^2(\vartheta, t_{j-1}, X_{t_{j-1}}) + \frac{(X_{t_j} - X_{t_{j-1}})^2}{\sigma^2(\vartheta, t_{j-1}, X_{t_{j-1}})} \right].$$

- This estimator is consistent and asymptotically mixed normal

$$\hat{\vartheta}_{t,n} \longrightarrow \vartheta_0, \quad \delta^{-1/2}(\hat{\vartheta}_{t,n} - \vartheta_0) \Longrightarrow \xi_t(\vartheta_0),$$

where the convergence is stable.

- This estimator is consistent and asymptotically mixed normal

$$\hat{\vartheta}_{t,n} \longrightarrow \vartheta_0, \quad \delta^{-1/2}(\hat{\vartheta}_{t,n} - \vartheta_0) \Longrightarrow \xi_t(\vartheta_0),$$

where the convergence is stable.

- $$\xi_t(\vartheta_0) = \frac{\sqrt{2}}{\mathrm{I}_t(\vartheta_0)} \int_0^t \frac{\dot{\sigma}(\vartheta_0, s, X_s)}{\sigma(\vartheta_0, s, X_s)} \mathrm{d}w(s),$$
$$\mathrm{I}_t(\vartheta_0) = 2 \int_0^t \frac{\dot{\sigma}(\vartheta_0, s, X_s)^2}{\sigma(\vartheta_0, s, X_s)^2} \mathrm{d}s,$$

$w(t)$ is a Wiener process so that $\{w(s), s \le t\}$ is independent from $\{X(s), s \le t\}$.

- This estimator is consistent and asymptotically mixed normal

$$\hat{\vartheta}_{t,n} \longrightarrow \vartheta_0, \quad \delta^{-1/2}(\hat{\vartheta}_{t,n} - \vartheta_0) \Longrightarrow \xi_t(\vartheta_0),$$

where the convergence is stable.

- $$\xi_t(\vartheta_0) = \frac{\sqrt{2}}{\mathrm{I}_t(\vartheta_0)} \int_0^t \frac{\dot\sigma(\vartheta_0, s, X_s)}{\sigma(\vartheta_0, s, X_s)} \mathrm{d}w(s),$$
$$\mathrm{I}_t(\vartheta_0) = 2 \int_0^t \frac{\dot\sigma(\vartheta_0, s, X_s)^2}{\sigma(\vartheta_0, s, X_s)^2} \mathrm{d}s,$$

$w(t)$ is a Wiener process so that $\{w(s),\ s \leq t\}$ is independent from $\{X(s),\ s \leq t\}$.
- Hence, $\hat{Y}_{t,n} = u(\hat{\vartheta}_{t,n}, t, X_{t_k})$ will be a consistent estimator for the process Y_t for each fixed $t \in (0, T]$.

- This estimator is consistent and asymptotically mixed normal

$$\hat{\vartheta}_{t,n} \longrightarrow \vartheta_0, \quad \delta^{-1/2}(\hat{\vartheta}_{t,n} - \vartheta_0) \Longrightarrow \xi_t(\vartheta_0),$$

where the convergence is stable.

- $$\xi_t(\vartheta_0) = \frac{\sqrt{2}}{\mathrm{I}_t(\vartheta_0)} \int_0^t \frac{\dot{\sigma}(\vartheta_0, s, X_s)}{\sigma(\vartheta_0, s, X_s)} \mathrm{d}w(s),$$
$$\mathrm{I}_t(\vartheta_0) = 2 \int_0^t \frac{\dot{\sigma}(\vartheta_0, s, X_s)^2}{\sigma(\vartheta_0, s, X_s)^2} \mathrm{d}s,$$

$w(t)$ is a Wiener process so that $\{w(s),\, s \leq t\}$ is independent from $\{X(s),\, s \leq t\}$.

- Hence, $\hat{Y}_{t,n} = u(\hat{\vartheta}_{t,n}, t, X_{t_k})$ will be a consistent estimator for the process Y_t for each fixed $t \in (0, T]$.
- But this estimator has a problem: **it is difficult to calculate.**

Estimation of the solution of a FBSDE (joint work with Yu. A. Kutoyants)

Example: Consider the following SDE

$$\mathrm{d}X_t = S(X_t)\mathrm{d}t + \sqrt{\vartheta}h(X_t)\mathrm{d}W_t, \quad X_0 = 0.$$

Suppose the problem is to estimate the unknown positive parameter ϑ based on continuous time observations $\{X_t,\, t \in [0, T]\}$. It is well known that by continuous time observations the unknown parameter in the diffusion coefficient can be estimated without an error.

Estimation without an Error

- Applying the Itô formula to the function $G(x) = x^2$, we get

$$X_t^2 = 2\int_0^t X_s \mathrm{d}X_s + \vartheta \int_0^t h^2(X_s)\mathrm{d}s,$$

hence, with the probability 1

Estimation without an Error

- Applying the Itô formula to the function $G(x) = x^2$, we get

$$X_t^2 = 2\int_0^t X_s \mathrm{d}X_s + \vartheta\int_0^t h^2(X_s)\mathrm{d}s,$$

hence, with the probability 1

- $$\vartheta = \frac{X_t^2 - 2\int_0^t X_s \mathrm{d}X_s}{\int_0^t h^2(X_s)\mathrm{d}s}, \quad \forall 0 < t \leq T.$$

Estimation without an Error

- Applying the Itô formula to the function $G(x) = x^2$, we get

$$X_t^2 = 2\int_0^t X_s \mathrm{d}X_s + \vartheta \int_0^t h^2(X_s)\mathrm{d}s,$$

  hence, with the probability 1

- $$\vartheta = \frac{X_t^2 - 2\int_0^t X_s \mathrm{d}X_s}{\int_0^t h^2(X_s)\mathrm{d}s}, \quad \forall 0 < t \leq T.$$

- So, if one has continuous time observations even in very small interval, then the unknown parameter can be found for almost all realizations.

The First Estimator

Suppose that $\tau > 0$ is a small number and we have to estimate the solution Y_t at the points $t \in [\tau, T]$. Denote $m = \left[\frac{\tau}{T} n\right]$, that is

$$0 = t_0 < \cdots < t_m \leq \tau < t_{m+1} < \cdots < t_k < t < t_{k+1} < \cdots < t_n = T.$$

Consider

$$\tilde{\vartheta}_{\tau,n} = \arg\min_{\vartheta \in \Theta} U_m(\vartheta),$$

where $U_m(\vartheta)$ is the same contrast function. Then, by previous arguments about estimation without an error, we have that the estimator

$$\tilde{Y}_{t,n} = u(\tilde{\vartheta}_{\tau,n}, t, X_{t_k})$$

is consistent. This estimator is the simplest one but it is not asymptotically efficient.

Efficiency

Before finding the best estimator we have to define what is the best estimator. [Ibragimov, Hasminskii, 1981], [Kutoyants, 2004] For this we need a theorem that compares all estimators as $n \longrightarrow +\infty$. We suppose that the function $u(\vartheta, t, X_t)$ is sufficiently smooth w.r.t. ϑ. Our first result is

Theorem (Lower bound)

For all estimators $\bar{Y}_{t_k,n}$ of the process Y_t the following inequality holds

$$\underline{\lim}_{\varepsilon\to 0}\ \underline{\lim}_{n\to+\infty}\ \sup_{|\vartheta-\vartheta_0|<\varepsilon} \mathbf{E}_\vartheta \ell\left(\delta^{-1/2}\left(\bar{Y}_{t_k,n}-Y_{t_k}\right)\right) \geq$$
$$\geq \mathbf{E}_{\vartheta_0}\ell(\dot{u}(\vartheta_0,t,X_t)\xi_t(\theta_0)), \quad \ell(u)=|u|^p, p>0.$$

Remark. Here we compare the estimators with the unknown process in the point t_k instead of point t.

Ideas of the Proof.

The proof of this theorem based on the fact [Dohnal, 1987], [Gobet, 2001] that under some regularity conditions the probability measures $\{\mathcal{P}_{\vartheta}^{n,k}, \vartheta \in \Theta\}$ induced by the observations

$$\mathbf{X}^k = (X_{t_0}, X_{t_1}, \cdots, X_{t_k})$$

$$\mathcal{P}_{\vartheta_0}^{n,k}(B) = \mathcal{P}(\mathbf{X}^k \in B), \forall B \in \mathcal{B}(\mathbf{R}^k)$$

satisfy the Local Asymptotic Mixed Normality (LAMN) [Jeganathan, 1982] condition: for all $\vartheta_0 \in \Theta$,

$$\ln \frac{\mathrm{d}\mathcal{P}_{\vartheta_0+\sqrt{\delta}v}^{n,k}}{\mathrm{d}\mathcal{P}_{\vartheta_0}^{n,k}}(\mathbf{X}^k) = v\Delta_{n,k}(\vartheta_0) - \frac{1}{2}v^2\mathrm{I}_{n,k}(\vartheta_0) + r_{n,k}(v,\vartheta_0),$$

where $r_{n,k}(v,\vartheta_0) \longrightarrow 0$, in $\mathcal{P}_{\vartheta_0}^{n,k}$ probability, for every $v \in R$,

$$\Delta_{n,k}(\vartheta_0) = \sqrt{2}\sum_{j=1}^{k} \frac{\dot\sigma(\vartheta_0, t_{j-1}, X_{t_{j-1}})}{\sigma(\vartheta_0, t_{j-1}, X_{t_{j-1}})^2}(X_{t_j} - X_{t_{j-1}} - S(t_{j-1}, X_{t_{j-1}})\delta)$$

$$\Longrightarrow \Delta_t(\vartheta_0) = \sqrt{2}\int_0^t \frac{\dot\sigma(\theta_0, s, X_s)}{\sigma(\theta_0, s, X_s)} dw(s),$$

stably in $\mathcal{P}_{\vartheta_0}^{n,k}$ law, $w(t)$ is a Wiener process independent from $\{X_s, 0 \leq s \leq t\}$ and, in $\mathcal{P}_{\vartheta_0}^{n,k}$ probability,

$$\mathrm{I}_{n,k}(\vartheta_0) = 2\sum_{j=1}^{k} \frac{\dot\sigma(\vartheta_0, t_{j-1}, X_{t_{j-1}})^2}{\sigma(\vartheta_0, t_{j-1}, X_{t_{j-1}})^2}\delta$$

$$\longrightarrow \mathrm{I}_t(\vartheta_0) = 2\int_0^t \frac{\dot\sigma(\vartheta_0, s, X_s)^2}{\sigma(\vartheta_0, s, X_s)^2} \mathrm{d}s,$$

$$\xi_t(\theta_0) = \frac{\Delta_t(\vartheta_0)}{\mathrm{I}_t(\vartheta_0)}.$$

Asymptotically Efficient Estimator

To find asymptotically efficient estimator we use the idea of **Le Cam**'s one-step Maximum Likelihood Estimator (MLE). For FBSDEs this idea was first implemented in the paper [Kutoyants, Zhou, 2014].

We take the preliminary estimator $\tilde{\vartheta}_{\tau,n}$ constructed by the observations

$$0 = t_0 < t_1 < \cdots < t_m \leq \tau$$

and improve it in the following way

$$\vartheta^*_{t,n} = \tilde{\vartheta}_{\tau,n} + \sqrt{\delta}\frac{\Delta_{n,k}(\tilde{\vartheta}_{\tau,n})}{\mathrm{I}_{n,k}(\tilde{\vartheta}_{\tau,n})}.$$

The Second Estimator

Then, the estimator $Y^*_{t,n} = u(\vartheta^*_{t,n}, t, X_{t_k})$ is consistent and asymptotically mixed normal

$$Y^*_{t_k,n} \longrightarrow Y_t, \quad \delta^{-\frac{1}{2}}(Y^*_{t_k,n} - Y_{t_k}) \Longrightarrow \dot u(\vartheta_0, t, X_t)\xi_t(\theta_0),$$

as $n \longrightarrow +\infty$ and

Theorem

The estimator $Y^*_{t,n} = u(\vartheta^*_{t,n}, t, X_{t_k})$ is asymptotically efficient

$$\lim_{\varepsilon\to 0}\lim_{n\to+\infty}\sup_{|\vartheta-\vartheta_0|<\varepsilon} \mathbf{E}_\vartheta \ell\left(\delta^{-1/2}\left(Y^*_{t_k,n} - Y_{t_k}\right)\right) =$$

$$= \mathbf{E}_{\vartheta_0}\ell(\dot u(\vartheta_0, t, X_t)\xi_t(\theta_0)), \quad \ell(u) = |u|^p, p > 0.$$

References

Gasparyan S.B.,
Second order asymptotical efficiency for a Poisson process.
Journal of Contemporary Mathematical Analysis, 50(2), 98-106. (2015).

Gasparyan S.B. and Kutoyants Yu.A.,
On the lower bound in second order asymptotically efficient estimation for Poisson processes.
Submitted. (2016).

Gasparyan S.B. and Kutoyants Yu.A.,
On Approximation of the BSDE with Unknown Volatility in Forward Equation.
Armenian Journal of Mathematics 7.1, 59-79. (2015).

Gasparyan S.B. and Kutoyants Yu.A.,
An example of one-step MLE-process in volatility estimation problem.
Izvestiya Natsionalnoi Akademii Nauk 50, 3, 71-76., Armenia. (2015).

References

Gill R.D. and Levit B.Ya.,
Applications of the van Trees inequality: a Bayesian Cramér-Rao bound.
Bernoulli, 1(1-2), 59-79. (1995).

Golubev G.K. and Levit B.Ya.,
On the second order minimax estimation of distribution functions.
Math. Methods Statist., 5(1), 1-31. (1996).

Golubev G.K. and Härdle W.,
Second order minimax estimation in partial linear models.
Math. Methods Statist., 9(2), 160-175. (2000).

Ibragimov, I., Hasminskii R.,
Statistical Estimation: Asymptotic Theory.
Springer, New York (1981).

References

Dalalyan A.S. and Kutoyants Yu.A.,
On second order minimax estimation of invariant density for ergodic diffusion.
Statistics & Decisions, 22(1), 17-42. (2004).

Kutoyants, Yu. A.,
Introduction to Statistics of Poisson Processes.
To appear. (2016).

Kutoyants, Yu. A.,
On Multi-Step MLE-Process for Ergodic Diffusion.
Submitted. (2016).

Pinsker M.S.,
Optimal filtering of square-integrable signals in Gaussian noise.
Probl. Inf. Transm., 16(2), 120-133. (1980).

References

Dohnal, G.,
On estimating the diffusion coefficient.
J. Appl. Probab., 24, 1, 105-114 (1987).

Genon-Catalot, V., Jacod, J.,
On the estimation of diffusion coefficient for multi-dimensional diffusion.
Ann. IHP, Sec. B,29, 1, 119-151. (1993).

Gobet, E.,
LAMN property for elliptic diffusion: a Malliavin calculus approach.
Bernoulli 7.6, 899-912. (2001).

Jeganathan P.,
On the asymptotic theory of estimation.
Sankhya:The Indian Journal of Statistics, 44.2, 173-212. (1982).

References

Kutoyants, Yu. A.,
Statistical Inference for Ergodic Diffusion Processes.
Springer, London (2004).

Kutoyants, Yu., Zhou, L.,
On approximation of backward stochastic differential equation.
J. Stat. Plann. Infer. 150, 111-123. (2014).

Liptser, R., Shiryaev, A.N.,
Statistics of Random Processes.
Springer, New York (2005).

Pardoux, E., Peng, S.,
BSDEs and Quasilinear Parabolic Partial Differential Equations, in
SDEs and their applications, Lect. Not. Cont. Inf. Sci., vol.176, 200-217, Springer (1992).

It's Over!