Multiple change-point estimation: model with non zero jumps sum

Alioune TOP

Université du Maine, Le Mans

Laboratoire Manceau de Mathématique

DYNSTOCH, Rennes June 10 2016

Plan

1. Introduction
2. Statement of the problem and some preliminaries
3. Asymptotic properties of Bayesian estimator and MLE
   - Mains results
   - Proofs of theorems
4. Simulations

Introduction

The estimation problem with discontinuous density with shift parameter was considered by many authors

- The i.i.d case

The first study was initiated in the work of Chernov and Rubin (discontinuous density with one jump).

Chernov, H. and Rubin, H., (1956), The estimation of the location of a discontinuity in density,
Proc.Third Berkeley Symp. Math.Statist. and Prob., 1, 19-38.

The case of many discontinuities was studied in the work of Rubin,

Rubin, H. (1961) The estimation of the discontinuities in multivariate densities, and related problems in stochastic process,,
Proc.Fourth Berkeley Symp.Math.Statist. and Prob., 1, 563-574.

See as well Ermakov

Ermakov, M. S., (1977), Asymptotic behavior of statistical estimates of parameters of multidimensional discontinuous density,
Zap. LOMI, **74**, 83–107 (in Russian).

Further development can be found in the works of Ibragimov and Khaminski, Strasser and Pflug.

Ibragimov, I. A. and Khasminskii, R. Z.,(1981) Statistical Estimation.
Asymptotic Theory, Springer, New York

Strasser, H. (1982) Lacal asymptotic minimax properties of Pitman estimates,
Z.Whrscheinlichkeitstheory verw. Gebiete 60,223-247.

Pflug, G. C., (1983) The limiting log-likelihood process for discontinuous density families,
Z. Wahrsch. Verw. Geb., **64**, 15–35.

- The case of Poisson process

Gal'tchouk and Rozovskii considered the disorder-type hypothesis testing problem

Gal'tchouk, L. I. and Rozovskii, B. L., (1971) The disorder problem for a Poisson process,
Theor. Probab. Appl., **16**, 712–716.

The problem of parameter estimation (consistency, limit distributions, convergence of moments, asymptotic efficiency) was considered by Kutoyants

Kutoyants, Yu. A., (1984) Parameter Estimation for Stochastic Processes
Heldermann-Verlag, Berlin.

Kutoyants, Yu. A.,(1998) Statistical Inference for Spatial Poisson Processes
Lecture Notes in Statistics **134**, *Springer-Verlag, New York*.

Note as well the related statistical problems in the works

Deshayes, J., (1983) Ruptures de modèles en statistique,
Thèse d'État, Université Paris-Sud.

Akman, V.E. and Raftery, A.E., (1986) Asymptotic inference for a change-point Poisson process,
Ann. Statist., **14**, 4, 1583–1590.

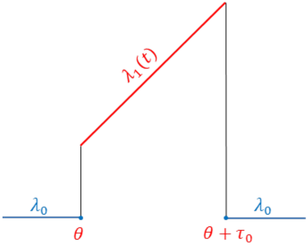
Dabye, A.S. (1999) Estimation par la méthode du maximum de vraisemblance pour un processus de Poisson d'intensité discontinue,
Comptes Rendue de l'Accadémie des Sciences, Séries 1, Mathématics, 329, 4, 335-338.

The Model

We suppose that the observations $X^{(n)} = (X_1, \ldots, X_n)$ are n independent inhomogeneous Poisson processes $X_j = \{X_j(t),\ 0 \leq t \leq T\}$, $j = 1, \ldots, n$ with the same intensity function

$$\lambda(\theta, t) = \lambda_0 + \lambda_1(t) 1\!\!1_{\{\theta \leq t \leq \theta + \tau_0\}}, \qquad 0 \leq t \leq \tau, \quad \theta \in \Theta = (\alpha, \beta)$$

Here $\tau = T - \tau_0, \quad 0 < \alpha < \beta < \beta + \tau_0 < \tau, \quad \inf_{\theta \in \Theta} |\lambda_1(\theta + \tau_0) - \lambda_1(\theta)| > 0$.
Under this condition we have the two jumps of the intensity function on the interval of observations for $\theta \in \Theta$.



The Model

- Our goal

The parameter θ is supposed to be unknown and we have to estimate it by the observations $X^{(n)}$. We are interested by the asymptotic ($n \to \infty$) behavior of the MLE and the BE.

- The interest of the model

The considered model of observation with intensity function $\lambda\left(\theta_0, t\right) = \lambda_0 + \lambda_1(t) 1\!\!1_{\{\theta_0 \leq t \leq \theta_0 + \tau_0\}}$ is typical for statistical radiophysics and this problem of detection of poissonian signal in poissonian noise comes from Grant of RSF devoted to this class of problems

- $\lambda_1(t) 1\!\!1_{\{\theta_0 \leq t \leq \theta_0 + \tau_0\}}$ is an signal of length $\tau_0 > 0$
- $\lambda_0 > 0$ is some Poissonian noise

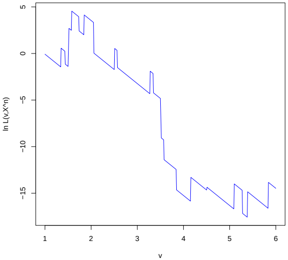
Therefore the problem of estimation of the parameter θ corresponds to the evaluation of the moment of arriving of the signal.
In optical communication theory : the parameter (information) θ is a transmitted through the Poissonian channel with modulated intensity where λ_0 is the intensity of the noise.

Denote by $\mathbf{P}_\theta^{(n)}$ the measure induced in the space of observation by n realizations of the Poisson process with the intensity function $\lambda(\theta, t)$, $0 \le t \le \tau$. As $\lambda_0 > 0$ and $\lambda_1(t)$ is bounded the measures $\mathbf{P}_\theta^{(n)}$, $\theta \in \Theta$ are equivalent and the likelihood ratio function is

$$
\begin{aligned}
L\left(\theta, \theta_1, X^{(n)}\right) \quad = \quad & \frac{d\mathbf{P}_\theta^{(n)}}{d\mathbf{P}_{\theta_1}^{(n)}}(X^{(n)}) = \exp\Bigg\{ \sum_{j=1}^{n} \int_0^\tau \ln\Big(\frac{\lambda_0 + \lambda_1(t) 1\!\!1_{\{\theta \le t \le \theta + \tau_0\}}}{\lambda_0 + \lambda_1(t) 1\!\!1_{\{\theta_1 \le t \le \theta_1 + \tau_0\}}} \Big) dX_j(t) \\
& - n \int_0^\tau \left(\lambda_1(t) 1\!\!1_{\{\theta \le t \le \theta + \tau_0\}} - \lambda_1(t) 1\!\!1_{\{\theta_1 \le t \le \theta_1 + \tau_0\}} \right) dt \Bigg\}.
\end{aligned}
$$

Here $\theta_1 \in \Theta$ is some fixed value

A realization of such log likelihood ratio in the case $n = 1$ and $\theta_0 = 2$ is given below.



Defintion

As the likelihood ratio $L\left(\theta, \theta_1, X^{(n)}\right)$ is a discontinuous function of θ, we define the MLE $\hat{\theta}_n$ as a solution of the following equation

$$\max\left\{L\left(\hat{\theta}_n+, \theta_1, X^{(n)}\right), L\left(\hat{\theta}_n-, \theta_1, X^{(n)}\right)\right\} = \sup_{\theta\in\Theta} L\left(\theta, \theta_1, X^{(n)}\right).$$

Here $L\left(\hat{\theta}_n+, \theta_1, X^{(n)}\right)$ and $L\left(\hat{\theta}_n-, \theta_1, X^{(n)}\right)$ are the left and the right limits of the function $L\left(\theta, \theta_1, X^{(n)}\right)$ at the point $\hat{\theta}_n$ respectively.
To introduce the Bayesian estimator we suppose that the unknown parameter is a random variable with known, positive, continuous density function $p(\theta)$, $\theta \in \Theta$. Then BE $\widetilde{\theta}_n$ is a conditional expectation, which can be written as follows

$$\widetilde{\theta}_n = \mathbf{E}\left(\theta / X^{(n)}\right) = \int_\alpha^\beta \theta p(\theta) L\left(\theta, X^{(n)}\right)\, d\theta \left(\int_\alpha^\beta p(\theta) L\left(\theta, X^{(n)}\right)\, d\theta\right)^{-1}.$$

Notations

Introduce the process

$$Z_{\theta_0}(u) = \begin{cases} \exp\left\{\rho_1(\theta_0)\ X^+(u) + \rho_2(\theta_0)\ Y^+(u) - r(\theta_0)u\right\}, & u \geq 0 \\ \exp\left\{-\rho_1(\theta_0)\ X^-(-u) - \rho_2(\theta_0)\ Y^-(-u) - r(\theta_0)u\right\}, & u < 0, \end{cases}$$

where $X^+(\cdot)$, $X^-(\cdot)$, $Y^+(\cdot)$ and $Y^-(\cdot)$ are independent Poisson processes (IPP) on $\mathbb{R}_+$ of the constant intensities $\lambda_0 + \lambda_1(\theta_0)$, λ_0, λ_0 and $\lambda_0 + \lambda_1(\theta_0 + \tau_0)$ respectively. The parameters $\rho_1(\theta_0)$, $\rho_2(\theta_0)$ and $r(\theta_0)$ are defined as follows

$$\rho_1(\theta_0) = \ln\frac{\lambda_0}{\lambda_0 + \lambda_1(\theta_0)}, \quad \rho_2(\theta_0) = \ln\frac{\lambda_0 + \lambda_1(\theta_0 + \tau_0)}{\lambda_0}, \quad r(\theta_0) = \lambda_1(\theta_0 + \tau_0) - \lambda_1(\theta_0).$$

Denote $\rho_1 = \rho_1(\theta_0)$, $\rho_2 = \rho_2(\theta_0)$, $r = r(\theta_0)$. Indeed, by putting $u = \frac{v}{r}$, $X_1^{\pm}(v) = X^{\pm}(\frac{v}{r})$ and $Y_1^{\pm}(v) = Y^{\pm}(\frac{v}{r})$ we get

$$Z_\rho^*(v) := \begin{cases} \exp\left\{\rho_1 X_1^+(v) + \rho_2 Y_1^+(v) - v\right\}, & v \geq 0 \\ \exp\left\{-\rho_1 X_1^-(-v) - \rho_2 Y_1^-(-v) - v\right\}, & v < 0, \end{cases}$$

where $X_1^+(\cdot)$, $X_1^-(\cdot)$, $Y_1^+(\cdot)$ and $Y_1^-(\cdot)$ are IPP on $\mathbb{R}_+$ of intensities $\frac{\lambda_0 e^{-\rho_1}}{r}$, $\frac{\lambda_0}{r}$, $\frac{\lambda_0}{r}$ and $\frac{\lambda_0 e^{\rho_2}}{r}$ respectively.

Introduce the random variables $\hat{u}$, $\hat{u}_\rho$, $\tilde{u}$ and $\tilde{u}_\rho$ by the equations

$$\max\left\{Z_{\theta_0}(\hat{u}-), Z_{\theta_0}(\hat{u}+)\right\} = \sup_{u\in\mathbb{R}} Z_{\theta_0}(u),$$

$$\max\left\{Z_\rho^*(\hat{u}_\rho-), Z_\rho^*(\hat{u}_\rho+)\right\} = \sup_{v\in\mathbb{R}} Z_\rho^*(v),$$

$$\tilde{u} = \int_{-\infty}^{+\infty} u Z_{\theta_0}(u)\, du \left(\int_{-\infty}^{+\infty} Z_{\theta_0}(u)\, du\right)^{-1}$$

and

$$\tilde{u}_\rho = \int_{-\infty}^{+\infty} v Z_\rho^*(v)\, dv \left(\int_{-\infty}^{+\infty} Z_\rho^*(v)\, dv\right)^{-1}.$$

Let us note that $\hat{u} \equiv \frac{\hat{u}_\rho}{r}$ and $\tilde{u} \equiv \frac{\tilde{u}_\rho}{r}$.

Mains results

Introduce the conditions $\mathbf{C_0}$:

- *The constants λ_0 and τ_0 are strictly positive and known.*
- *The function $\lambda_1(\cdot)$, $t \in [0,\tau]$ is strictly increasing, strictly positive and continuous.*

The first result gives us the lower bound on the risk of all the estimators.

Theorem 1

Let the conditions $\mathbf{C_0}$ be fulfilled. Then for all $\theta_0 \in \Theta$

$$\lim_{\delta\to 0} \varliminf_{n\to+\infty} \inf_{\bar\theta_n} \sup_{|\theta-\theta_0|<\delta} n^2 \mathbf{E}_\theta \left(\bar\theta_n - \theta\right)^2 \geq \mathbf{E}_{\theta_0} \tilde u^2 = \frac{\mathbf{E}_{\theta_0}\left(\tilde u_\rho^2\right)}{r^2}. \qquad (1)$$

Here the *inf* is taken over all possible estimators $\bar\theta_n$ of the parameter θ.
The inequality (1) allows us to give the following definition.
Let the conditions $\mathbf{C_0}$ be satisfied, we say that an estimator $\bar\theta_n$ is asymptotically efficient, if for all $\theta_0 \in \Theta$ we have

$$\lim_{\delta\to 0} \lim_{n\to+\infty} \sup_{|\theta-\theta_0|<\delta} n^2 \mathbf{E}_\theta \left(\bar\theta_n - \theta\right)^2 = \frac{\mathbf{E}_{\theta_0}\left(\tilde u_\rho^2\right)}{r^2}$$

Mains results

Denote $\mathbf{K} \subset \Theta$ a compact set.

Theorem 2

Let the conditions $\mathbf{C_0}$ *be fulfilled. Then the Bayesian estimator* $\tilde{\theta}_n$ *and the maximum likelihood estimator* $\hat{\theta}_n$ *verify uniformly on* $\theta_0 \in \mathbf{K}$ *the relations :*
they are consistent

$$\mathbf{P}_{\theta_0} - \lim_{n\to+\infty} \tilde{\theta}_n = \theta_0, \quad \mathbf{P}_{\theta_0} - \lim_{n\to+\infty} \hat{\theta}_n = \theta_0$$

converge in Law

$$\mathcal{L}_{\theta_0}\left\{n\left(\tilde{\theta}_n - \theta_0\right)\right\} \Rightarrow \mathcal{L}\left(\frac{\tilde{u}_\rho}{r}\right), \quad \mathcal{L}_{\theta_0}\left\{n\left(\hat{\theta}_n - \theta_0\right)\right\} \Rightarrow \mathcal{L}\left(\frac{\hat{u}_\rho}{r}\right).$$

For any $p > 0$ *the moments of estimators converge*

$$\lim_{n\to+\infty} \mathbf{E}_{\theta_0}|n\left(\tilde{\theta}_n - \theta_0\right)|^p = \mathbf{E}_{\theta_0}\frac{|\tilde{u}_\rho|^p}{|r|^p}, \quad \lim_{n\to+\infty} \mathbf{E}_{\theta_0}|n\left(\hat{\theta}_n - \theta_0\right)|^p = \frac{\mathbf{E}_{\theta_0}|\hat{u}_\rho|^p}{|r|^p}.$$

The BE is asymptotically efficient.

Proofs of theorems

The presented proofs are based on the general results of Ibragimov and Khasminski (1981) and on the development in case of Poisson process given by Kutoyants(1984, 1998) .
To apply it we study the normalized likelihood ratio process of the model

$$
\begin{aligned}
Z_{\theta_0,n}(u) &\equiv L\left(\theta_0+\frac{u}{n},\theta_0,X^{(n)}\right)\\
&= \exp\left\{\sum_{j=1}^{n}\int_0^{\tau}\ln\left(\frac{\lambda_0+\lambda_1(t)1\!\!1_{\{\theta_0+\frac{u}{n}\leq t\leq\theta_0+\frac{u}{n}+\tau_0\}}}{\lambda_0+\lambda_1(t)1\!\!1_{\{\theta_0\leq t\leq\theta_0+\tau_0\}}}\right)dX_j(t)\right.\\
&\quad \left. -n\int_0^{\tau}\left(\lambda_1(t)1\!\!1_{\{\theta_0+\frac{u}{n}\leq t\leq\theta_0+\frac{u}{n}+\tau_0\}}-\lambda_1(t)1\!\!1_{\{\theta_0\leq t\leq\theta_0+\tau_0\}}\right)dt\right\}
\end{aligned}
$$

where $u\in U_n=\left(n\left(\alpha-\theta_0\right),n\left(\beta-\theta_0\right)\right)$.

The factor of normalization is n.

Lemma 1

Let the conditions C_0 be satisfied, then the finite dimensional distributions of the process $Z_{\theta_0,n}(u)$ converge to the finite dimensional distributions of the process $Z_{\theta_0}(u)$ and this convergence is uniform with respect to $\theta_0 \in \mathbf{K}$.

Lemma 2

Let the conditions C_0 be satisfied, then there exists a constant $C > 0$ such that

$$\mathbf{E}_{\theta_0} \mid Z_{\theta_0,n}^{1/2}(u_1) - Z_{\theta_0,n}^{1/2}(u_2) \mid^2 \leq C \mid u_1 - u_2 \mid;$$

for all $n \in \mathbb{N}$, $u_1, u_2 \in U_n$ and $\theta_0 \in \mathbf{K}$

Lemma 3

Let the conditions C_0 be satisfied, then there exists a constant $c > 0$ such that

$$\mathbf{E}_{\theta_0} Z_{\theta_0,n}^{1/2}(u) \leq e^{-c|u|}$$

For all $n \in \mathbb{N}$, $u \in U_n$ and $\theta_0 \in \mathbf{K}$

For $\theta = \theta_0 + \frac{u}{n}$, the Bayesian estimator can be written as

$$\tilde{\theta}_n = \frac{\int_\alpha^\beta \theta p(\theta) L\left(\theta, X^{(n)}\right)\ d\theta}{\int_\alpha^\beta p(\theta) L\left(\theta, X^{(n)}\right)\ d\theta} = \theta_0 + \frac{1}{n} \frac{\int_{\mathbb{U}_n} u p(\theta_0 + \frac{u}{n}) L\left(\theta_0 + \frac{u}{n}, X^{(n)}\right)\ du}{\int_{\mathbb{U}_n} p(\theta_0 + \frac{u}{n}) L\left(\theta_0 + \frac{u}{n}, X^{(n)}\right)\ du}$$

Therefore

$$n(\tilde{\theta}_n - \theta_0) = \frac{\int_{\mathbb{U}_n} u p(\theta_0 + \frac{u}{n}) Z_{n,\theta_0}(u)\ du}{\int_{\mathbb{U}_n} p(\theta_0 + \frac{u}{n}) Z_{n,\theta_0}(u)\ du}.$$

In view of Lemmas 1, 2 and 3 we can, referring to Theorem A.22 (see Ibragimov and Khasminski) assert that the right hand term converges to

$$\tilde{u} = \frac{\int_{\mathbb{R}} u Z_{\theta_0}(u) du}{\int_{\mathbb{R}} Z_{\theta_0}(u) du} \qquad i.e. \qquad n(\tilde{\theta}_n - \theta_0) \Rightarrow \tilde{u}.$$

The consistency and the convergence of the moments of $\tilde{\theta}_n$ also hold.

Construction of the lower bound

The uniform convergence of moments of the BE and the continuity of the limit risk allow us to obtain the inequality of the Theorem 1. We have

$$\sup_{|\theta-\theta_0|<\delta} n^2 \mathbf{E}_\theta(\bar\theta_n-\theta)^2 \geq n^2 \int_{\theta_0-\delta}^{\theta_0+\delta} \mathbf{E}_\theta(\bar\theta_n-\theta)^2 p_\delta(\theta)d\theta.$$

Here we introduced a density function $(p_\delta(\theta), \theta_0-\delta<\theta<\theta_0+\delta)$. Let us denote by $\tilde\theta_{\delta,n}$ the BE which corresponds to this density function. Then we have the inequality

$$\int_{\theta_0-\delta}^{\theta_0+\delta} \mathbf{E}_\theta(\bar\theta_n-\theta)^2 p_\delta(\theta)d\theta \geq \int_{\theta_0-\delta}^{\theta_0+\delta} \mathbf{E}_\theta(\tilde\theta_{\delta,n}-\theta)^2 p_\delta(\theta)d\theta.$$

As we have a uniform convergence of moments for this BE, we obtain the limit

$$\lim_{n\to\infty} n^2 \int_{\theta_0-\delta}^{\theta_0+\delta} \mathbf{E}_\theta(\tilde\theta_{\delta,n}-\theta)^2 p_\delta(\theta)d\theta = \int_{\theta_0-\delta}^{\theta_0+\delta} \frac{\mathbf{E}_\theta(\tilde u_\rho^2)}{r(\theta)^2} p_\delta(\theta)d\theta.$$

Recall that $r(\theta)=\lambda_1(\theta+\tau_0)-\lambda_1(\theta)$ and $\mathbf{E}_\theta(\tilde u_\rho^2)$ are continuous functions of θ. Therefore it is possible to verify that

$$\lim_{\delta\to 0} \int_{\theta_0-\delta}^{\theta_0+\delta} \frac{\mathbf{E}_\theta(\tilde u_\rho^2)}{r(\theta)^2} p_\delta(\theta)d\theta = \frac{\mathbf{E}_{\theta_0}(\tilde u_\rho^2)}{r^2}.$$

Weak convergence in Skorohod metric

Introduce the space $\mathbf{D}_0(\mathbb{R})$ of functions $\varphi(u)$ without discontinuities of the second kind defined on $\mathbb{R}$ and such that $\lim_{|u|\to+\infty} \varphi(u)=0$. We assume that all the functions $\varphi(u) \in \mathbf{D}_0(\mathbb{R})$ are continuous from the right, and have limits from the left (càdlàg).
Let φ_1 and φ_2 be two functions belonging to $\mathbf{D}_0(\mathbb{R})$. The Skorohod distance between them is defined as follows

$$d(\varphi_1, \varphi_2) = \inf_{\mu} \left[\sup_{\mathbb{R}} |\varphi_1(u) - \varphi_2(\mu(u))| + \sup_{\mathbb{R}} |u - \mu(u)| \right],$$

where the inf is taken over all the increasing continuous one-to-one mappings $\mu : \mathbb{R} \longrightarrow \mathbb{R}$. This metric space $(\mathbf{D}_0(\mathbb{R}), d(\cdot,\cdot))$ is complete and separable. For $z \in \mathbf{D}_0(\mathbb{R})$, we put

$$\begin{aligned} \Delta_h(z) &= \sup_{u\in\mathbb{R}} \sup_{u-h\le u' < u < u'' \le u+h} \left[\min\left\{ \left| z(u') - z(u) \right|, \left| z(u'') - z(u) \right| \right\} \right] \\ &+ \sup_{|u|>h^{-1}} |z(u)|. \end{aligned}$$

Weak convergence in Skorohod metric

For all $\theta \in \Theta$, suppose that we have a sequence $(z_{n,\theta})_{n\geq 1}$ of stochastic processes $z_{n,\theta} = \left\{z_{n,\theta}(u), u \in \mathbb{R}\right\}$ and a process $z_\theta = \left\{z_\theta(u), u \in \mathbb{R}\right\}$ such that realizations of these processes belong to the space $\mathbf{D}_0(\mathbb{R})$. Denote $\mathbf{Q}_\theta^n$ and $\mathbf{Q}_\theta$ the distributions (which we suppose depending on a parameter $\theta \in \Theta$) induced on the measurable space $(\mathbf{D}_0(\mathbb{R}), \mathcal{B}(\mathbb{R}))$ by the processes $z_{n,\theta}$ and z_θ respectively. Here $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra of the metric space $\mathbf{D}_0(\mathbb{R})$. A criterion of weak convergence in $\mathbf{D}_0(\mathbb{R})$ is given in the following lemma.

Lemma 4

Let the following two conditions be satisfied :

1- *the finite dimensional distributions of the process $z_{n,\theta}$ converge to the finite dimensional distributions of the process z_θ uniformly in $\theta \in \mathbf{K} \subset \Theta$.*

2- *For any $\epsilon > 0$, we have*

$$\lim_{h\to 0} \sup_{n\in\mathbb{N}} \sup_{\theta\in\mathbf{K}} \mathbf{Q}_\theta^n \left\{\Delta_h(z_{n,\theta}) > \epsilon\right\} = 0. \tag{2}$$

Then for all functionals $\phi(\cdot) \in \mathbf{D}_0(\mathbb{R})$ the distribution of $\phi(z_{n,\theta})$ converges to the distribution of $\phi(z_\theta)$ uniformly in $\theta \in \mathbf{K}$, that is, $z_{n,\theta}$ converges weakly uniformly to z_θ.

Consistency and convergence in law

We need the weak convergence of the likelihood ratio $Z_{n,\theta_0}(\cdot)$ to the process $Z_{\theta_0}(\cdot)$ in the space $\mathbf{D}_0(\mathbb{R})$. Suppose that we already proved this convergence.
For any set $B \in \mathcal{B}(\mathbb{R})$, we define on $\mathbf{D}_0(\mathbb{R})$ the functionals $\Phi_B(\cdot)$ and $\Psi_B(\cdot)$ by

$$\Phi_B(\varphi) = \sup_{u \in B} \varphi(u) \quad \text{and} \quad \Psi_B(\varphi) = \sup_{u \in B^c} \varphi(u)$$

respectively. Thus, the functionals $\Phi_B(\cdot)$ and $\Psi_B(\cdot)$ are continuous in the the Skorohod metric.
Put $\hat{u}_n = n(\hat{\theta}_n - \theta_0)$. We obtain

$$\begin{aligned} \mathbb{P}_{\theta_0}^{(n)}\left(\hat{u}_n \in B\right) &= \mathbb{P}_{\theta_0}^{(n)}\left\{(\Phi_B(Z_{n,\theta}) > \Psi_B(Z_{n,\theta_0})\right\} \\ &\longrightarrow \mathbb{P}_{\theta_0}\left(\Phi_B(Z_{\theta_0}) > \Psi_B(Z_{\theta_0})\right) = \mathbb{P}_{\theta_0}\left(\hat{u}_{\Psi_B(Z_{\theta_0})} \in B\right). \end{aligned}$$

Hence the consistency and convergence in law of the MLE are proved.

Proof of Lemma 4

The convergence of the finite dimensional distributions is already checked by Lemma 1.
Recall that $U_n = ((\alpha - \theta_0)\, n, (\beta - \theta_0)\, n)$, and put

$$V_n = \Big((\alpha - \theta_0)\, n - 1, (\beta - \theta_0)\, n + 1 \Big).$$

The process $Z_{n,\theta_0}(u)$ is defined on the set U_n. We extend it over the entire V_n such that it is continuously decreasing to zero in the bands of width 1 but still keeps the discontinuous points in u. Outside V_n we define the process $Z_{n,\theta}(\cdot)$=0 . Now the process $Z_{n,\theta_0}(\cdot)$ is defined on the whole real line for all n, and the realizations of the process $Z_{n,\theta_0}(\cdot)$ belong to the space $\mathbb{D}_0(\mathbb{R})$ with probability 1.
We set for $z \in \mathbb{D}_0(\mathbb{R})$,

$$\begin{aligned} \Delta_h^l(z) &= \sup_{u,u',u'' \in \delta_l} \left[\min\left\{ \left| z(u') - z(u) \right|, \left| z(u'') - z(u) \right| \right\} \right] \\ &+ \sup_{l \le u \le l+h} |z(u) - z(l)| + \sup_{l+1-h \le u \le l+1} |z(u) - z(l+1)| . \end{aligned}$$

Here $l > 0$ and $u, u', u'' \in \delta_l$ means that $l \le u - h \le u' < u < u'' \le u + h \le l + 1$.

- First we estimate the probability $\mathbf{P}_{\theta_0}^{(n)}\left(\Delta_h^l(Z_{n,\theta_0}^{1/4}) > h^{1/8}\right)$

Notations

- $\mathbb{D}$ be the event that on the interval $[l, l+1]$ there exist at least two jumps of the process $Z_{n,\theta_0}(u)$ such that the distance between them is less than $2h$.
- $\mathbb{D}_p$ the event that the process $Z_{n,\theta_0}(u)$ has at least p jumps on the interval $(u, u+h)$ and $(u+\tau_0, u+\tau_0+h)$.
-

$$\mathbb{C}_h = \left\{ u \in \delta_l : \sup_{u^{'}, u^{"} \in \delta_l} \left[\min \left\{ \left| Z_{n,\theta_0}^{\frac{1}{4}}(u^{'}) - Z_{n,\theta_0}^{\frac{1}{4}}(u) \right|, \left| Z_{n,\theta_0}^{\frac{1}{4}}(u^{"}) - Z_{n,\theta_0}^{\frac{1}{4}}(u) \right| \right\} \geq h^{1/8} \right] \right\}.$$

Lemma 5

Let the conditions C_0 be satisfied, then there exists a constant $C > 0$ such that

$$\sup_{\theta_0 \in \mathbf{K}} \mathbf{P}_{\theta_0}^{(n)}(\mathbb{D}_1) \leq Ch \quad \text{and} \quad \sup_{\theta_0 \in \mathbf{K}} \mathbf{P}_{\theta_0}^{(n)}(\mathbb{D}_2) \leq C^2 h^2.$$

Consequently

- If the event $\mathbb{D}$ occurs ; then $\mathbf{P}_{\theta_0}^{(n)}(\mathbb{D}) \leq Ch$.
- If the event $\mathbb{D}^c$ occurs then $\mathbf{P}_{\theta_0}^{n}(\mathbb{C}_h) \leq Ch^{\frac{3}{8}}$.

The others terms of the modulus $\Delta_h^l(z)$ can be estimated in a similar way. This gives us the estimate

$$\begin{aligned} \mathbf{P}_{\theta_0}^{n}(\Delta_h^l(Z_{n,\theta_0}^{\frac{1}{4}}) > h^{\frac{1}{8}}) &\leq \mathbf{P}_{\theta_0}^{n}(\mathbb{D}) + \mathbf{P}_{\theta_0}^{n}\left(\Delta_h^l(Z_{n,\theta_0}^{\frac{1}{4}}) > h^{\frac{1}{8}}, \mathbb{D}^c\right) \\ &\leq Ch + Dh^{\frac{1}{8}} \leq \gamma h^{\frac{3}{8}}. \end{aligned} \tag{3}$$

To end the proof we need also the following lemma

Lemma 6

Let

$$M_n = \sup_{|u|<L} Z_{n,\theta_0}^{\frac{3}{4}}(u),$$

then we have

$$\mathbf{P}_{\theta_0}^{n}\left\{M_n > h^{\frac{-1}{16}}\right\} \leq \kappa h^{\frac{1}{128}}$$

Simulations

We suppose that the observations $X^{(n)} = (X_1, \ldots, X_n)$ are n independent inhomogeneous Poisson processes $X_j = \{X_j(t),\ 0 \le t \le 10\}$, $j = 1, \ldots, n$ with the same intensity function

$$\lambda(\theta, t) = 1 + 2t1\!\!1_{\{\theta \le t \le \theta+2\}}, \qquad 0 \le t \le \tau$$

with $\theta \in (1, 6)$ and $\tau = 8$. The true value of the parameter is θ_0 =2. Then we have

$$\begin{aligned} L\left(\theta, X^{(n)}\right) &= \exp\left\{\sum_{j=1}^{n} \int_0^8 \ln\left(1 + 2t1\!\!1_{\{\theta \le t \le \theta+2\}}\right) dX_j(t) - 4n(\theta+1)\right\} \\ &= \exp\left\{\sum_{j=1}^{n} \sum_{\theta \le t_j^i \le \theta+2} \ln\left(1 + 2t_j^i\right) - 4n(\theta+1)\right\}, \end{aligned} \tag{4}$$

where $\{t_j^i\}_{j=1,\cdots,N_j}$ $(N_j = X_j(10))$ are the events of the process X_j with intensity function $\lambda(2, t)$.

The second sum in (4) is equal to zero when there is no event of the observed process.

asymptotic behavior of estimators

n	10	30	50	100	120	140
$\hat{\theta}_n$	3.02	1.23	2.21	1.99	1.99	2.001

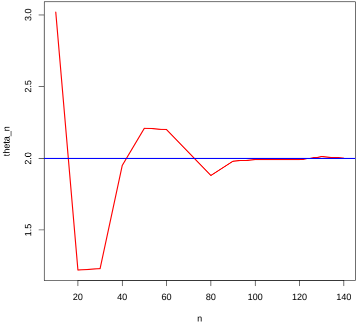


FIGURE: evolution of $\hat{\theta}_n$ with respect to n

For large values of n, the estimator $\hat{\theta}_n$ approachs reasonably to the true value $\theta = 2$

Behavior of limit process

Remaind that $\theta_0 = 2$, $\tau_0 = 2$, $\lambda_0 = 1$ and $\lambda_1(t) = 2t$ for $t \in (0, 8)$. Therefore we obtain $\rho_1 = -\ln 5$, $\rho_2 = \ln 9$ and $r = 4$.

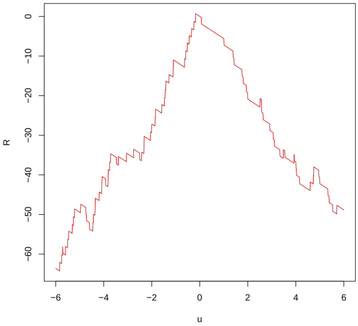


FIGURE: A sample path of the process $\ln Z(u)$

To estimate the limit variances of the MLE and BE we made 10^4 simulations of these variables and the results are

$$\sigma^2_{MLE} \approx \frac{1}{N}\sum_{l=1}^{N} \hat{u}_l^2 = 1.33 \quad and \quad \sigma^2_{BE} \approx \frac{1}{N}\sum_{l=1}^{N} \tilde{u}_l^2 = 0.58.$$

This confirms that

$$\sigma^2_{MLE} > \sigma^2_{BE}.$$

These values concur with the theoretical results that the Bayesian estimator outperforms the MLE. It concur also the i.i.d. case with one point of singularity (see Ibragimov and Khasminski (1981) and Kutoyants(1998)) where it was mentioned that the Bayesian estimators are generally more efficient that the MLE estimators in Change-Point estimation.

Thank You !