Exponential bounds for intensity of jumps

Delphine Blanke, Denis Bosq

EXPONENTIAL BOUNDS FOR INTENSITY OF JUMPS

DELPHINE BLANKE

Avignon University, Laboratoire de Mathématiques d'Avignon, EA2151, 33 rue Louis Pasteur, F-84000 Avignon, France

DENIS BOSQ

Sorbonne Universités, UPMC Univ Paris 06, EA 3124, Laboratoire de Statistique Théorique et Appliquée, F-75005, 4 place Jussieu, Paris, France

Abstract. In this paper, we study intensity of jumps in the context of functional linear processes. The natural space for that is the space $D = D[0,1]$ of cadlag real functions. We begin with limit theorems for ARMAD(1,1) processes. It appears that under some conditions, the functional linear process and its innovation have the same jumps. This nice property allows us to focus on the case of i.i.d. D-valued random variables. For such variables, we estimate the intensity of jumps in various situations : fixed number of jumps, random instants of jumps, random number of instants of jumps, ... We derive exponential rates and limits in distribution.

1. Introduction

A lot of papers are devoted to autoregressive processes with values in separable Hilbert or Banach spaces (see Antoniadis et al., 2012; Besse et al., 2000; Cardot, 1998; Damon and Guillas, 2005; Ferraty and Vieu, 2006; Kargin and Onatski, 2008; Marion and Pumo, 2004; Mas, 2004; Mourid, 2002; Pumo, 1998; Ruiz-Medina, 2012, among many others). It is more difficult to study D-valued linear processes, where $D = D[0,1]$ is the space of cadlag real functions defined on [0,1]. The main reason is that, if D is equipped with the sup-norm, it becomes a non-separable space. In order to obtain separability, it is preferable to use the Skorohod metric (cf Billingsley, 1999). Now, in the framework

E-mail addresses: delphine.blanke@univ-avignon.fr,denis.bosq@upmc.fr.

2000 *Mathematics Subject Classification.* Primary 62M, 60F17.

Key words and phrases. Functional linear processes, Jumps, Estimation of intensity.

of D-valued linear processes there are not many papers (see however El Hajj, 2011; El Hajj, 2013; Bosq, 2014).

Here, our aim is to study continuous time random variables and to estimate the intensity of jumps at random or fixed instants. The process $(X(t),\, t \in [0,1])$ is observed over $[0,1]$, this time interval can be interpreted as one day, one week, one year...

Various applications may be considered. A classical and simple example is the *compound Poisson process*: the holders of an insurance policy are victims of misfortunes at the instants $0 < T_1 < T_2 < \cdots$ following a Poisson process with intensity λ. They obtain the respective payments $\Delta_1, \Delta_2, \cdots$ at instants $T_1, T_2, \cdots$ and one may set

$$X(t) = \sum_{n=1}^{N_t} \Delta_n,\ 0 \le t \le 1$$

where $N_t = \sup\{n\ :\ T_n \le t\}$, $0 \le t \le 1$ and the convention $X(t) = 0$ if $N_t = 0$. A similar example is a particle subjected to impacts at Poissonian instants T_n, where Δ_n denotes the displacement of the particle at time T_n.

Now, another example is the *wind speed* (Jacq et al., 2005) associated with the *mistral gust*: one may construct a D-valued ARMA model and note that, under mild conditions, the model and the strong white noise have the same jumps (see Section 2). Then, since it is difficult to predict the gust intensity, one may suppose that the instants of gusts are *independent*, and the model is *no more Poissonian*. A study of that situation appears in Section 6.

Other models can be exhibited:

- In *finance*, it can be shown that a model with jumps is better than the Black-Scholes model (see Cont and Tankov, 2004; Tankov and Voltchkova, 2009, for details).
- Another example of jumps is associated with *electricity consumption*: clearly, a jump appears early in the morning and late in the evening (Antoniadis et al., 2012; El Hajj, 2013).
- The model invoking *dengue* is slightly different since it involves bifurcation (cf Garba et al., 2008), it is related with dynamical systems but it contains jumps.

The previous examples show that we must consider various distinct situations: they appear below.

In Section 2, we recall some properties of D and give some examples. The next section is devoted to D-valued ARMA(1,1) processes with

$$X_n - \rho(X_{n-1}) = Z_n - \rho'(Z_{n-1}), \quad n \in \mathbb{Z}$$

where (Z_n) is a D-strong white noise and ρ, ρ' are continuous linear operators. It can be shown that limit theorems hold for (Z_n) if and only if they hold for (X_n). Now, if $\rho(D) \subset C[0,1]$ and $\rho'(D) \subset C[0,1]$, it follows that Z_n and X_n have the same jumps. This property leads us to consider i.i.d. D-valued variables in the next section.

Section 4 is devoted to the case where the D-valued random variable X admits k distinct fixed jumps at instants $t_1, \ldots, t_k$. The problem is to estimate the intensity of jump $\mathbb{E}(\Delta_j)$ where $\Delta_j = \left|X(t_j) - X(t_j^-)\right|$, $j = 1, \ldots, k$. Clearly, if $\Delta_{1,j}, \ldots, \Delta_{n,j}$ are i.i.d. copies of Δ_j, or if they satisfy a suitable strong mixing condition, it is easy to obtain limit theorems concerning $(\mathbb{E}(\Delta_j), j = 1, \ldots, k)$ and to estimate the greatest jump. Similar results can be obtained for jumps at increasing random instants: $0 < T_1 < \cdots < T_k < 1$ almost surely. An ordering for intensity of jumps is also available.

In Section 5, we suppose that the number K_i of jumps is random and independent from the $\Delta_{i,k} = \left|X_i(t_k) - X_i(t_k^-)\right|$, $i = 1, \ldots, n$, $k \geq 1$. In order to estimate $\mathbb{E}(\Delta_k)$ from the observed $\Delta_{i,k}$, we set

$$\overline{\Delta}_{n,k} = \frac{\sum_{i=1}^n \Delta_{i,k} \mathbb{1}_{\{K_i \geq k\}}}{\sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}}}, \quad k \geq 1$$

with the convention $\overline{\Delta}_{n,k} = 0$ if $\sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}} = 0$ $(k \geq 0)$. Then, it may be shown that

$$\left|\overline{\Delta}_{n,k} - \mathbb{E}(\Delta_k)\right| = \mathcal{O}\Big(\sqrt{\frac{\ln n}{n}}\Big) \quad \text{a.s..}$$

Some applications and extensions to random ordered instants are given.

Finally, the last section is devoted to the non-ordered case: $T_1, \ldots, T_k$ are independent random instants, thus the scheme is not Poissonian. Each X_i has k jumps $(T_{ij}, j = 1, \ldots, k)$ which are not directly observable. Now, supposing that $\Delta_1, \ldots, \Delta_k$ are independent and noting that

$$\prod_{j=1}^k \big(x - \mathbb{E}(\Delta_j)\big) = \sum_{j=0}^k a_j x^j = 0,$$

one may use a trick for estimating the coefficients $a_0, \ldots, a_{k-1}$ for $k \geq 2$. It follows that one can obtain an equation of the form

$$\sum_{j=0}^k \widehat{a}_{j,n} x^j = 0, \tag{1.1}$$

where $\lim\limits_{n\to\infty} \widehat{a}_{j,n} = a_j$ a.s., $j = 1, \ldots, k$. Then (1.1) can be solved at least by approximation.

Observations in discrete time and numerical applications will appear in a next paper (cf Blanke and Bosq, 2014).

2. Constructing D-valued random variables

In order to study the jumps of the real process $X = \big(X(t), 0 \leq t \leq 1\big)$, it is natural to consider the space $D = D([0,1])$ of cadlag real functions defined over $[0,1]$. If D is equipped with the *sup-norm*: $\|x\| =$

$\sup_{0\le t\le 1} |x(t)|$, it becomes a *non-separable* space. Thus, it is preferable to use the *Skorohod metric* defined as

$$d(x,y) = \inf_{\lambda\in\Lambda} \left\{ |\lambda - I| \vee \|x - y\lambda\| \right\} \; ; \; x, y \in D$$

where Λ is the class of strictly increasing continuous mapping of [0,1] onto [0,1] and I is the identity from $[0,1]$ to $[0,1]$. Then, D equipped with the Skorohod metric is separable, we refer to Billingsley (1999) for a detailed study of D. Now, we denote by $\mathcal{D}$ the σ-algebra generated by the Skorohod metric. We only recall three useful properties of $(D,\mathcal{D})$:

- If $x \in C = C([0,1])$, then

$$d(x_n, x) \xrightarrow[n\to\infty]{} 0 \iff \|x_n - x\| \xrightarrow[n\to\infty]{} 0.$$

- If ρ is a bounded linear operator, i.e.

$$\|\rho\|_{\mathcal{L}} = \sup_{x\in D, \|x\|\le 1} \|\rho(x)\| < \infty,$$

  then it is $\mathcal{D} - \mathcal{D}$ measurable.
- $x \mapsto x(t_0) - x(t_0^-)$ is a continuous linear form on $(D, \|\cdot\|)$.

See Billingsley (1999) and Pestman (1995) for further properties.

Now, let X be a $(D,\mathcal{D})$-valued random variable defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In this paper, we focus on estimation of the intensity of jumps in the following cases:

- a fixed number of jumps at fixed or random times,
- a random number of jumps at fixed times,
- a random number of jumps at random times.

We give below some examples of jumps associated with X corresponding to our framework.

Example 2.1 (k fixed jumps)**.**

$$X(t,w) = \sum_{j=1}^{k} Y_j(t,w) \mathbb{1}_{[t_{j-1},t_j[}(t), \quad t \in [0,1], \; w \in \Omega$$

where t_j's are fixed points with $0 = t_0 < t_1 < \cdots < t_k \le 1$ and Y_j is $\mathcal{D}\otimes\mathcal{A} - \mathcal{B}_{\mathbb{R}}$ measurable, $1 \le j \le k$.

Example 2.2 (k random jumps)**.**
Consider the $k+1$ measurable processes $(Z_j(t,w),\ 1 \le j \le k+1,\ 0 \le t \le 1,\ w \in \Omega)$ with continuous sample paths and k random variables (r.v.) $T_1, \dots, T_k$ with values in]0,1[and such that $\mathbb{P}(T_j = T_{j'}) = 0$, $j \neq j'$. Suppose that $Z_1, \dots, Z_{k+1}$, $T_1, \dots, T_k$ are globally independent. Then, we set

$$X(t,w) = \sum_{j=1}^{k+1} Z_j(t,w) \mathbb{1}_{[T_{j-1}^*(w), T_j^*(w)[}(t), \quad 0 \le t \le 1, w \in \Omega$$

where the T_j^* *are ordered as:* $0 = T_0^* < T_1^* < \cdots T_k^* < T_{k+1}^* = 1$, *and* $X(1) = Z_{k+1}(1)$.

Example 2.3 (Random number of jumps).
Let $0 = T_0 < T_1 < \cdots < T_K < \cdots$ *be a strictly increasing sequence of random variables (almost surely) with* K *a random* $\mathbb{N}$*-valued variable. Let us set* $N_1 = \sum_{k=1}^{\infty} \mathbb{1}_{T_k \le 1}$ *and*

$$X(t,w) = \begin{cases} \sum_{j=1}^{k} Y_j(w) \mathbb{1}_{[T_{j-1}, T_j[}(t) & \text{if } N_1 = k \ge 1,\ 0 \le t \le 1 \\ 0 & \text{if } N_1 = 0 \end{cases}$$

where Y_j *is* $\mathcal{A} - \mathcal{B}_{\mathbb{R}}$ *measurable. Note that an example of such a model is the* compound Poisson process*.*

3. The case of ARMAD processes

3.1. Limit theorems for ARMAD(1,1).

Consider the ARMAD(1,1) process

$$(X_n - m) - \rho(X_{n-1} - m) = Z_n - \rho'(Z_{n-1}),\ n \in \mathbb{Z}$$

where (Z_n) is a D-strong white noise (i.e. (Z_n) is i.i.d., $0 < \mathbb{E}\,\|Z_n\|^2 < \infty$, $\mathbb{E}(Z_n) = 0$), $m \in D$, and $\rho : D \mapsto D$, $\rho' : D \mapsto D$ are linear bounded operators such that $\|\rho^j\|_{\mathcal{L}} < 1$ and $\left\|\rho'^{j'}\right\|_{\mathcal{L}} < 1$ for some integers $j \ge 1$ and $j' \ge 1$. Note that ρ and ρ' are $\mathcal{D} - \mathcal{D}$ measurable (cf Pestman, 1995).

Now, it is easy to show that

$$X_n = m + \sum_{j=0}^{\infty} \rho^j \big(Z_{n-j} - \rho'(Z_{n-j-1})\big),\ n \in \mathbb{Z}$$

almost surely and in L^2. Moreover (Z_n) is the innovation of (X_n) and (X_n) is equidistributed.

A classical example of linear bounded operator in D meeting all our conditions is as follows:

Example 3.1.

$$\rho(x)(t) = \int_0^1 r(s,t) x(s)\, \mathrm{d}s,\ 0 \le t \le 1,\ x \in D,$$

where r *is continuous and* $\max_{0 \le s,t \le 1} |r(s,t)| < 1$. *In addition, one has*

$$\rho(D) \subset C = C([0,1]).$$

Now, we state the law of large numbers.

Proposition 3.1. (X_n) *satisfies the strong law of large numbers (SLLN) if and only if* (Z_n) *satisfies it. The same statement holds for the* L^2 *law of large numbers.*

Proof. We may and do suppose that $m = 0$. Now, set

$$Y_n = X_n - \rho(X_{n-1}) = Z_n - \rho'(Z_{n-1}),\ n \in \mathbb{Z} \tag{3.1}$$

then, we have

$$\overline{X}_n = \frac{\Delta_{n,X}}{n} + (I - \rho)^{-1}\overline{Y}_n, \tag{3.2}$$

where

$$\Delta_{n,X} = (I - \rho)^{-1}\rho(X_0 - X_n),$$

thus, Tchebychev inequality yields

$$\frac{\|\Delta_{n,X}\|}{n} \xrightarrow[n\to\infty]{a.s.} 0$$

then, from (3.1) it follows that

$$\left\|\overline{X}_n\right\| \xrightarrow[n\to\infty]{a.s.} 0 \Longleftrightarrow \left\|\overline{Y}_n\right\| \xrightarrow[n\to\infty]{a.s.} 0. \tag{3.3}$$

Similarly, we may write

$$\overline{Z}_n = \frac{\Delta_{n,Z}}{n} + (I - \rho')^{-1}\overline{Y}_n$$

with

$$\Delta_{n,Z} = (I - \rho')^{-1}\rho'(Z_0 - Z_n),$$

and using again Tchebychev inequality, one obtains

$$\left\|\overline{Z}_n\right\| \xrightarrow[n\to\infty]{a.s.} 0 \Longleftrightarrow \left\|\overline{Y}_n\right\| \xrightarrow[n\to\infty]{a.s.} 0 \tag{3.4}$$

thus, (3.3) and (3.4) give the result.

The proof concerning the L^2 law of large numbers is similar. Details are omitted. □

Example 3.2. *If (Z_n) is convex tight or if it takes its values in the cone of nondecreasing functions over D, then $d(0, \overline{Z}_n) \xrightarrow[n\to\infty]{a.s.} 0$ (cf Daffer and Taylor, 1979) and, since 0 is a continuous function, it is equivalent to write $\left\|\overline{Z}_n\right\| \to 0$ a.s.. In particular, if $Z_n(t) = N_{n+t} - N_n - \lambda t$, $0 \le t \le 1$, $n \ge 1$ where $(N_s,\ s \ge 0)$ is a Poisson process with intensity λ the strong law of large numbers holds.*

We now apply Proposition 3.1 for obtaining consistency of jumps: suppose that X_n has jumps at $t_1, \dots, t_k$ (≥ 1) with intensity $\mathbb{E}\big(X_n(t_j) - X_n(t_j^-)\big)$, $j = 1, \dots, k$. Then, we have :

Corollary 3.1. *If (Z_n) satisfies the SLLN then*

$$\frac{1}{n}\sum_{i=1}^{n}\big(X_i(t_j) - X_i(t_j^-)\big) \xrightarrow[n\to\infty]{a.s.} \mathbb{E}\big(X_i(t_j) - X_i(t_j^-)\big), \tag{3.5}$$

$j = 1, \dots, k$.

Proof. First, if (Z_n) satisfies the SLLN, Proposition 3.1 implies the same property for (X_n). Now, set

$$\varphi_{t_j}(x) = x(t_j) - x(t_j^-),\ x \in D,\ j = 1, \dots, k,$$

it is a continuous linear form on $(D, \|\cdot\|)$ (cf Pestman, 1995). Then, by continuity and linearity of φ_{t_0}, (3.5) follows. □

Now, we make an additional assumption :

Assumption 3.1 (A3.1)**.**

$$\rho(D) \subset C([0,1]),\ \rho'(D) \subset C([0,1])$$

Then:

Corollary 3.2. *Under A3.1, we have*

$$X_n(t_j) - X_n(t_j^-) = Z_n(t_j) - Z_n(t_j^-),\ \ j = 0, \dots, k. \tag{3.6}$$

Proof. Write $X_n = U_n + Z_n$ where $U_n = \rho(X_{n-1}) + \rho'(Z_{n-1})$, then, since $U_n(D) \subset C([0,1])$, (3.6) holds. □

It follows that the jumps of (X_n) are i.i.d. ; that property entails that all results derived in the sequel for i.i.d. jumps are also satisfied by such ARMAD processes!

We now turn to the central limit theorem (CLT).

Proposition 3.2. *The CLT holds for (Z_n) if and only if holds for (X_n).*

Proof. We suppose that $m = 0$ and we use again (3.1) for obtaining

$$\sqrt{n}\,\overline{X}_n = \frac{\Delta_{n,X}}{\sqrt{n}} + (I-\rho)^{-1}\sqrt{n}\overline{Y}_n,$$

and

$$\sqrt{n}\,\overline{Z}_n = \frac{\Delta_{n,Z}}{\sqrt{n}} + (I-\rho')^{-1}\sqrt{n}\overline{Y}_n.$$

Recall that $\Delta_{n,X} = (I-\rho)^{-1}\rho(X_0 - X_n)$ and $\Delta_{n,Z} = (I-\rho')^{-1}\rho'(Z_0 - Z_n)$. Next, as X_0 and X_n are equidistributed we get,

$$\mathbb{P}\Big(\frac{\|\Delta_{n,X}\|}{\sqrt{n}} \ \geq\ \eta\Big) \ \leq\ \frac{\mathbb{E}\,\|\Delta_{n,X}\|}{\eta\sqrt{n}} \ \leq\ \frac{\|(I-\rho)^{-1}\|_{\mathcal{L}}\,\|\rho\|_{\mathcal{L}}\,2\mathbb{E}\,\|X_0\|}{\eta\sqrt{n}},$$

$\eta > 0$. One may clearly obtain a similar bound for $\Delta_{n,Z}$. Since $\frac{\Delta_{n,X}}{\sqrt{n}} \xrightarrow[n\to\infty]{P} 0$ and $\frac{\Delta_{n,Z}}{\sqrt{n}} \xrightarrow[n\to\infty]{P} 0$, the results follow from Billingsley (1999) p. 21 and 27. □

Note that conditions for the CLT can be found in Bloznelis and Paulauskas (1993). Concerning the CLT for jumps, let us set

$$V_{ij} = X_i(t_j) - X_i(t_j^-) - \mathbb{E}\big(X_i(t_j) - X_i(t_j^-)\big), \;\; i = 1, \dots, n, \; j = 1, \dots, k$$

and denote Φ the distribution function of $\mathcal{N}(0,1)$. Then, we have

Corollary 3.3. *Under A3.1,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V_{ij} \Longrightarrow N \sim \mathcal{N}\big(0, \mathbb{E}(V_j^2)\big), \; j = 1, \dots, k.$$

If, in addition $\mathbb{E}\,|V_{ij}|^3 < \infty$*, we have*

$$\sup_{t\in\mathbb{R}} \left| \mathbb{P}\Big(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{V_{ij}}{\sqrt{\mathbb{E}(V_{ij}^2)}} \le t \Big) - \Phi(t) \right| \le \frac{\mathbb{E}\,|V_{ij}|^3}{\big(\mathbb{E}(V_{ij}^2)\big)^{\frac{3}{2}} \sqrt{n}}.$$

Proof. The first part of the proof is clear since one may use directly Corollary 3.2 : the CLT follows since $\big(Z_n(t_j) - Z_n(t_j^-)\big)$ are i.i.d.. Using again Corollary 3.2 we are in a position to apply Berry-Esseen theorem (see e.g. Shiryaev, 1996, p. 374) for the second part. $\square$

Let us conclude this section with some final remarks.

Remark 3.1. *A special case is the model*

$$X_n(t) = a(t) X_{n-1}(t) + Z_n(t), \; 0 \le t \le 1, \; n \in \mathbb{Z}$$

where X_n *and* Z_n *have a jump at* t_0 *and* a *is continuous at* t_0 *and such that* $|a(t_0)| < 1$*. Consequently*

$$X_n(t_0) - X_n(t_0^-) = a(t_0) \cdot \big(X_{n-1}(t_0) - X_{n-1}(t_0^-)\big) + \big(Z_n(t_0) - Z_n(t_0^-)\big),$$

$n \in \mathbb{Z}$*. Then,* $\big(X_n(t_0) - X_n(t_0^-)\big)$ *is a real autoregressive process.*

Remark 3.2. *Note that* $\rho(D) \subset C$ *is not always satisfied in Example 3.1. For example, if* $r(s,t) = a(s)b(t)$ *where* $\int_0^1 a(s)x(s)\,\mathrm{d}s \neq 0$ *and* b *has a jump at* t_0*, one obtains*

$$\rho(x)(t_0) - \rho(x)(t_0^-) = \big(b(t_0) - b(t_0^-)\big) \int_0^1 a(s)x(s)\,\mathrm{d}s \neq 0.$$

Remark 3.3. *A slight modification allows to introduce* exogenous *random variables. Set* $X_n - m = \rho(X_{n-1} - m) + Z_n - \rho'(Z'_n)$*, then, if* (Z_n) *and* (Z'_n) *satisfy the SLLN,* (X_n) *satisfies it.*

4. A fixed number of jumps

4.1. **Case of fixed times of jumps.** We begin with a very simple case. Let X be a $(D, \mathcal{D})$-valued process admitting exactly $k \ge 1$ distinct jumps at times $0 < t_1 < \cdots < t_k < 1$. Now and in all the paper, we make use of the generic notation Δ to denote the intensity of jumps. So we set

$$\Delta_j = \big| X(t_j) - X(t_j^-) \big|, \;\; j = 1, \dots, k$$

and one wants to estimate $\mathbb{E}(\Delta_j)$ (supposed to be finite), $1 \le j \le k$ from n independent copies of Δ_j. In this case, the k jumps are observed therefore known, so one may derive the following immediate results.

Proposition 4.1.
If $\mathbb{E}(\|X\|) < \infty$ *and* $\overline{\Delta}_{j,n} := \frac{1}{n}\sum_{i=1}^{n} \left|X_i(t_j) - X_i(t_j^-)\right|$, $j = 1, \ldots, k$, *we get*

a) $\overline{\Delta}_{j,n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}(\Delta_j)$, $j = 1, \ldots, k$.

b) *If moreover* $\mathbb{E}(\|X\|^2) < \infty$, *then*

$$\Big(\sqrt{n}\big(\overline{\Delta}_{j,n} - \mathbb{E}(\Delta_j)\big),\ j = 1, \ldots, k\Big) \xrightarrow[n\to\infty]{d} \mathcal{N}_k(0, \Sigma)$$

where Σ *is a* $k \times k$ *matrix with elements* $\Sigma_{j,j'} = \mathrm{Cov}\,(\Delta_j, \Delta_{j'})$, $j, j' = 1, \ldots, k$.

It is easy to derive an exponential bound for $\mathbb{P}(\left|\overline{\Delta}_{j,n} - \mathbb{E}(\Delta_j)\right| \ge \varepsilon)$, $j = 1, \ldots, k$ and then, obtain an almost sure rate of convergence in Proposition 4.1-a). For this, one can make use of the following version of Bernstein's inequality:

Proposition 4.2.
Let $\xi_1, \ldots, \xi_n$ *be independent real-valued random variables such that* $\sigma_i^2 := \mathrm{Var}\,(\xi_i)$ *and Bernstein's condition holds for all* $i = 1, \ldots, n$:

$$\mathbb{E}\,|\xi_i - \mathbb{E}(\xi_i)|^m \le (m!/2)\sigma_i^2 H^{m-2},\quad H > 0,\ m = 3, \cdots$$

then for $\varepsilon > 0$:

$$\mathbb{P}\Big(\left|\sum_{i=1}^{n} \xi_i - \mathbb{E}(\xi_i)\right| \ge n\varepsilon\Big) \le 2\exp\Big(-\frac{n^2\varepsilon^2}{2\sum_{i=1}^{n}\sigma_i^2 + 2Hn\varepsilon}\Big).$$

Note that Bernstein's condition is equivalent to the existence of an exponential moment for ξ_i. Indeed, it is true as soon as $|\xi_i - \mathbb{E}(\xi_i)| \le H$ a.s..

If one wants to estimate the greatest jump, it is easy to prove that if $\left|X_i(T_{i,j}) - X_i(T_{i,j}^-)\right|$ fulfills conditions of Proposition 4.2 for each $j = 1, \ldots, k$ (with values $\sigma_j^2 = \mathrm{Var}\,(\left|X_i(t_j) - X_i(t_j^-)\right|)$ and H_j), then

$$\mathbb{P}\Big(\left|\max_{j=1,\ldots,k} \overline{\Delta}_{j,n} - \max_{j=1,\ldots,k} \mathbb{E}(\Delta_j)\right| \ge \varepsilon\Big) \le 2ke^{-nc(\varepsilon)},\ \varepsilon > 0$$

with $c(\varepsilon)^{-1} = \max_{j=1,\ldots,k}(2\sigma_j^2 + 2H_j\varepsilon) > 0$. Actually, it suffices to note that

$$\left|\max_{j=1,\ldots,k} \overline{\Delta}_{j,n} - \max_{j=1,\ldots,k} \mathbb{E}(\Delta_j)\right| \ge \varepsilon \Longrightarrow \max_{j=1,\ldots,k} \left|\overline{\Delta}_{j,n} - \mathbb{E}(\Delta_j)\right| \ge \varepsilon$$

and to deduce the result from Proposition 4.2 for the latter term. Also, Proposition 4.1-b) induces construction of tests for existence of jumps.

An alternative point is the estimation of

$$\mathbb{E}\big(\max_{j=1,\ldots,k}\left|X(t_j)-X(t_j^-)\right|\big) := \mathbb{E}(\Delta_{\max})$$

(remark that $\max_{j=1,\ldots,k}\mathbb{E}(\Delta_j)\le \mathbb{E}(\max_{j=1,\ldots,k}\Delta_j)$). To this end, set

$$\overline{\Delta}_{k,n}=\frac{1}{n}\sum_{i=1}^{n}\max_{j=1,\ldots,k}\left|X_i(t_j)-X_i(t_j^-)\right|.$$

Clearly $\overline{\Delta}_{k,n}\xrightarrow[n\to\infty]{a.s.}\mathbb{E}(\Delta_{\max})$ if $\mathbb{E}(\Delta_{\max})<\infty$ and one gets an exponential rate as soon as $\Delta_{\max}$ admits an exponential moment.

Finally, Proposition 4.1 can be extended to the case of non independent copies of Δ_j, satisfying for example some strong mixing conditions (see e.g. Bradley, 2007). Also, recall that results can be directly applied for some particular functional linear processes considered in Section 3, cf Example 3.2 and Corollary 3.2.

4.2. **Case of k random jumps.** The second step consists in taking random instants $0<T_1<\cdots<T_k<1$ (a.s.) with k fixed. The intensity of jumps is given by $\Delta_j=\left|X(T_j)-X(T_j^-)\right|$, $j=1,\ldots,k$. Again, one wants to estimate $\mathbb{E}(\Delta_j)$, $j=1,\ldots,k$, from i.i.d. copies of Δ_j. The instants of jumps are observed and have the form $0<T_{i,1}<\cdots<T_{i,k}<1$ (a.s.), $i=1,\ldots,n$. Then, the estimator of $\mathbb{E}(\Delta_j)$ is

$$\overline{\Delta}_{j,n}=\frac{1}{n}\sum_{i=1}^{n}\left|X_i(T_{i,j})-X_i(T_{i,j}^-)\right|,\quad j=1,\ldots,k.$$

Clearly, all the above results remain valid: almost sure consistency, exponential rate, estimation of the greatest jump, k-dimensional central limit theorem. Details are left to the reader. Moreover, the next statement shows that it is also possible to classify the jumps according to their respective intensities.

Proposition 4.3.
Suppose that for all $i=1,\ldots,n$, $\left|X_i(T_{i,j})-X_i(T_{i,j}^-)\right|$ fulfills conditions of Proposition 4.2 for each $j=1,\ldots,k$. If $\mathbb{E}(\Delta_{\ell_1})>\cdots>\mathbb{E}(\Delta_{\ell_k})>0$ for some permutation $\{\ell_1,\ldots,\ell_k\}$ of $\{1,\ldots,k\}$, then almost surely for n large enough, one gets $\overline{\Delta}_{\ell_1,n}>\cdots>\overline{\Delta}_{\ell_k,n}$.

Proof. We begin with the study of $\mathbb{P}\big(\bigcup_{j=1}^{k-1}\{\overline{\Delta}_{\ell_j,n}<\overline{\Delta}_{\ell_{j+1},n}\}\big)$. First,

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{j=1}^{k-1}\{\overline{\Delta}_{\ell_j,n}<\overline{\Delta}_{\ell_{j+1},n}\}\Big)&\le\sum_{j=1}^{k-1}\mathbb{P}\big(\overline{\Delta}_{\ell_j,n}<\overline{\Delta}_{\ell_{j+1},n}\big)\\&\le\sum_{j=1}^{k-1}\mathbb{P}\big(\overline{\Delta}_{\ell_{j+1},n}-\overline{\Delta}_{\ell_j,n}-\mathbb{E}(\Delta_{\ell_{j+1}}-\Delta_{\ell_j})>\mathbb{E}(\Delta_{\ell_j}-\Delta_{\ell_{j+1}})\big).\end{aligned}$$

Next, $\mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}}) > 0$ by assumption, so we apply Bernstein's inequality with the property

$$\mathrm{Var}\left(\left|X_i(T_{i,\ell_{j+1}}) - X_i(T^-_{i,\ell_{j+1}})\right| - \left|X_i(T_{i,\ell_j}) - X_i(T^-_{i,\ell_j})\right|\right) \leq 2(\sigma^2_{\ell_{j+1}} + \sigma^2_{\ell_j})$$

where $\sigma_j^2 := \mathrm{Var}\left(\left|X_i(T_{i,j}) - X_i(T^-_{i,j})\right|\right)$ for all $i = 1, \ldots, n$. So there exist $H' > 0$ such that

$$\begin{aligned}\mathbb{P}\big((\overline{\Delta}_{\ell_{j+1},n} - \overline{\Delta}_{\ell_j,n}) - \mathbb{E}(\Delta_{\ell_{j+1}} - \Delta_{\ell_j}) \geq \mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}})\big) \\ \leq \exp\Big(-\frac{n(\mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}}))^2}{4(\sigma^2_{\ell_j} + \sigma^2_{\ell_{j+1}}) + 2H'\mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}})}\Big).\end{aligned}$$

Then, an uniform bound of $j = 1, \ldots, k$ can be obtained by considering $\max_j \mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}})$, $\min_j \mathbb{E}(\Delta_{\ell_j} - \Delta_{\ell_{j+1}})$ as well as the bound $\sigma^2_{\ell_j} + \sigma^2_{\ell_{j+1}} \leq 2\max_j \sigma_j^2$. Next, Borel Cantelli lemma implies that

$$\mathbb{P}\Big(\overline{\lim_{n\to\infty}} \bigcup_{j=1}^{k-1} \{\overline{\Delta}_{\ell_j,n} < \overline{\Delta}_{\ell_{j+1},n}\}\Big) = 0$$

yielding in turn that a.s. for n large enough, $\overline{\Delta}_{\ell_1,n} > \cdots > \overline{\Delta}_{\ell_k,n}$. $\square$

By this way, one may consistently estimate jump's intensities $\mathbb{E}(\Delta_{\ell_j})$ by considering the ordered jumps $\overline{\Delta}_{\ell_j,n}$, $j = 1, \ldots, k$.

5. A random number of fixed jumps

In this part, we consider the bit more intricate case where X takes its values in $(D, \mathcal{D})$ and has K random jumps for some nonnegative r.v. K such that for $k = 0, \cdots$:

$$\mathbb{P}(K = k) = p_k,\ p_0 < 1 \text{ and } p_k \geq 0.$$

We suppose also that K and Δ are independent. If K takes a positive value k then jumps occur at fixed times $0 := t_0 < t_1 < \cdots < t_k < 1$. Then, conditionally on $\{K = k\}$, one gets $\left|X(t_{k_0}) - X(t^-_{k_0})\right| = 0$ for all $k_0 > k$. A possible construction of such a process is given in Example 2.3 in the case of degenerated times $T_k := t_k$. The intensity of the k-th jump is denoted by

$$\Delta_k := \left|X(t_k) - X(t_k^-)\right| \tag{5.1}$$

with the condition $\mathbb{E}(\Delta_k) < \infty$, $k \geq 1$. Note that $\mathbb{E}(\Delta_k) > 0$ as soon as p_k is positive.

5.1. **Estimation of jumps intensities.** We consider an i.i.d. sequence of number of jumps $K_1, \dots, K_n$ independent from the intensities of jumps $\Delta_{i,k} := \left|X_i(t_k) - X_i(t_k^-)\right|$, $i = 1, \dots, n$, $k \geq 1$. For a given value of k, our aim is to estimate $\mathbb{E}(\Delta_k)$ from the observed $\Delta_{i,k} := \left|X_i(t_k) - X_i(t_k^-)\right|$ when $K_i \geq k$, $k \geq 1$. Since the number of jumps is not known and varies with i, we consider the following estimator:

$$\widehat{I}_{k,n} = \begin{cases} \dfrac{\sum_{i=1}^n \left|X_i(t_k) - X_i(t_k^-)\right| \mathbb{1}_{\{K_i \geq k\}}}{\sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}}}, & \text{if } \sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}} > 0 \\ 0, & \text{if } \sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}} = 0, \end{cases} \tag{5.2}$$

which is equivalent to

$$\widehat{I}_{k,n} = \Big(\frac{\sum_{i=1}^n \left|X_i(t_k) - X_i(t_k^-)\right| \mathbb{1}_{\{K_i \geq k\}}}{\sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}}}\Big) \mathbb{1}_{\{\sum\limits_{i=1}^n \mathbb{1}_{\{K_i \geq k\}} > 0\}}$$

for $k \geq 1$, using the convention $\frac{0}{0} = 0$.

Rates of convergence for $\widehat{I}_{k,n}$ toward $\mathbb{E}(\Delta_k)$, $k \geq 1$, are given in the following statement.

Theorem 5.1.
Suppose in addition that for $k \geq 1$ and $i = 1, \dots, n$, $\Delta_{i,k}$ fulfills conditions of Proposition 4.2 with variance σ_k^2 and constant H_k. Then, for all $k = 1, \cdots$, such that $\sum\limits_{i \geq k} p_i > 0$, one obtains

a) for each $c_0 \in]0, 1[$ and all $0 < \varepsilon \leq 2H_k(1 - c_0)$:

$$\mathbb{P}\Big(\left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| \geq \varepsilon\Big) \leq 2 \exp\Big(- \frac{c_0}{2} (\sum_{i \geq k} p_i) \frac{n\varepsilon^2}{\sigma_k^2 + 2H_k^2}\Big); \tag{5.3}$$

b)

$$\overline{\lim_{n\to\infty}} \sqrt{\frac{n}{\ln n}} \left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| \leq \sqrt{2}\sigma_k (\sum_{i \geq k} p_i)^{-\frac{1}{2}} \quad a.s.$$

Proof. We have to study

$$\begin{aligned} &\mathbb{P}\big(\left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| \geq \varepsilon\big) \\ &\qquad = \mathbb{P}\Big(\left|\frac{\sum_{i=1}^n \Delta_{i,k} \mathbb{1}_{\{K_i \geq k\}}}{\sum_{i=1}^n \mathbb{1}_{\{K_i \geq k\}}} \mathbb{1}_{\{\sum\limits_{i=1}^n K_i \geq k\} > 0} - \mathbb{E}(\Delta_k)\right| \geq \varepsilon\big), \end{aligned}$$

$\varepsilon > 0$, which can be written as

$$\mathbb{P}\Big(\Big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\Big| \geq \varepsilon\Big) = \sum_{j=0}^{n} \mathbb{P}\Big(\sum_{i=1}^{n} \mathbb{1}_{\{K_i \geq k\}} = j\Big)$$
$$\times\, \mathbb{P}\Bigg(\left|\frac{\sum_{i=1}^{n} \Delta_{i,k} \mathbb{1}_{\{K_i\geq k\}}}{\sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}}} \mathbb{1}_{\{\sum_{i=1}^{n} K_i \geq k\}>0} - \mathbb{E}(\Delta_k)\right| \geq \varepsilon \Big| \sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}} = j\Bigg).$$

As $\sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}} \sim \mathcal{B}(n, \sum_{i\geq k} p_i)$ and, since $\big\{\sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}} = j\big\}$ is equivalent to have exactly j indicators equal to 1, the i.i.d assumption on the Δ_i's and independence from K give

$$\mathbb{P}\Big(\Big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\Big| \geq \varepsilon\Big) = \mathbb{1}_{\{\varepsilon \leq \mathbb{E}(\Delta_k)\}} \mathbb{P}\Big(\sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}} = 0\Big)$$
$$+ \sum_{j=1}^{n} \mathbb{P}\Big(\sum_{i=1}^{n} \mathbb{1}_{\{K_i\geq k\}} = j\Big) \times \mathbb{P}\Bigg(\left|\frac{\sum_{i=1}^{j} \Delta_{i,k}}{j} - \mathbb{E}(\Delta_k)\right| \geq \varepsilon\Bigg).$$

Now, one may use Bernstein's inequality to obtain:

$$\mathbb{P}\Big(\big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\big| \geq \varepsilon\Big) \leq (1 - \sum_{i\geq k} p_i)^n$$
$$+ 2\sum_{j=1}^{n} \binom{n}{j} (1 - \sum_{i\geq k} p_i)^{n-j} (\sum_{i\geq k} p_i)^j \exp\Big(-\frac{j\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon}\Big)$$

so that,

$$\mathbb{P}\Big(\big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\big| \geq \varepsilon\Big) \leq 2\Big(1 - \sum_{i\geq k} p_i + \sum_{i\geq k} p_i \exp\big(-\frac{\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon}\big)\Big)^n.$$

Since $\ln(1-a) \leq -a$ for $0 < a < 1$ and $1 - e^{-a} \geq a - \frac{a^2}{2}$ for all $a \geq 0$, we successively obtain for all k such that $\sum_{i\geq k} p_i > 0$:

$$\mathbb{P}\Big(\big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\big| \geq \varepsilon\Big) \leq 2\exp\Big(-n\sum_{i\geq k} p_i \big(1 - \exp(-\frac{\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon})\big)\Big)$$
$$\leq 2\exp\Big(-\frac{n\sum_{i\geq k} p_i\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon}\big(1 - \frac{\varepsilon^2}{4\sigma_k^2 + 4H_k\varepsilon}\big)\Big).$$

Next, the condition $0 < c_0 < 1$ and $1 - \frac{1}{2}\frac{\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon} \geq c_0$ entail

$$\mathbb{P}\Big(\big|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\big| \geq \varepsilon\Big) \leq 2\exp\Big(-c_0 n(\sum_{i\geq k} p_i)\frac{\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon}\Big). \tag{5.4}$$

Now, it is easy to verify that $0 < \varepsilon \leq 2H_k(1 - c_0)$ is sufficient to get the condition $1 - \frac{1}{2}\frac{\varepsilon^2}{2\sigma_k^2 + 2H_k\varepsilon} \geq c_0$ and the bound (5.3) is deduced from (5.4) as one has also $\varepsilon \leq 2H_k$. Finally, the rate of convergence is derived from (5.4) with the choice $\varepsilon = c_1\sqrt{\frac{\ln n}{n}}$ with $c_1 > \sqrt{\frac{2\sigma_k^2}{c_0\sum_{i\geq k} p_i}}$ and application of Borel-Cantelli lemma.

□

5.2. **Estimation of the maximal jump.** Theorem 5.1 allows us to estimate the maximal jump of X from i.i.d. copies Δ_i. Suppose that there exists an unique integer $k_{\max}$ such that

$$\mathbb{E}(\Delta_{k_{\max}}) > \max_{\substack{k\geq 1\\ k\neq k_{\max}}} \mathbb{E}(\Delta_k),$$

again, the difficulty is that not all observed sample paths have a number of jumps greater than $k_{\max}$. An estimator of $\mathbb{E}(\Delta_{k_{\max}})$ is given by $\widehat{I}_{\max} = \max\limits_{k=1,\dots,k_n} \widehat{I}_{k,n}$ with $\widehat{I}_{k,n}$ defined by (5.2) and $k_n \to \infty$. We obtain the following result.

Proposition 5.1. *Under the conditions of Theorem 5.1,*

(1) If K has a finite support $\{0,1,\dots,k_0\}$ with $p_0 \neq 1$, then

$$\overline{\lim_{n\to\infty}} \left|\widehat{I}_{\max} - \mathbb{E}(\Delta_{k_{\max}})\right| = \mathcal{O}\Big(\sqrt{\frac{\ln n}{n}}\Big) \quad a.s..$$

(2) If K is a $\mathbb{N}$-valued random variable, and if $k_n \to \infty$ such that $k_n = \mathcal{O}\big(\ln n)^{\kappa}\big)$ for some $\kappa > 0$, then

$$\overline{\lim_{n\to\infty}} \left|\widehat{I}_{\max} - \mathbb{E}(\Delta_{k_{\max}})\right| = \mathcal{O}\Big(\sqrt{\frac{\ln n}{np_{k_n}}}\Big) \quad a.s.$$

with $p_{k_n} = \mathbb{P}(K = k_n)$.

Proof. Observe that

$$\max_{k=1,\dots,k_n} \left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| \geq \left|\max_{k=1,\dots,k_n} \widehat{I}_{k,n} - \max_{k=1,\dots,k_n} \mathbb{E}(\Delta_k)\right|,$$

so,

$$\begin{aligned}\mathbb{P}\Big(\left|\widehat{I}_{\max} - \max_{k=1,\dots,k_n} \mathbb{E}(\Delta_k)\right| > \varepsilon\Big) &\leq \mathbb{P}\Big(\bigcup_{k=1}^{k_n}\Big\{\left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| > \varepsilon\Big\}\Big)\\ &\leq \sum_{k=1}^{k_n} \mathbb{P}\Big(\left|\widehat{I}_{k,n} - \mathbb{E}(\Delta_k)\right| > \varepsilon\Big).\end{aligned}$$

(1) First if K has a finite support $\{0,\dots,k_0\}$, we get that $\widehat{I}_{\max} = \max\limits_{k=1,\dots,k_0} \widehat{I}_{k,n}$ (a.s.) for n large enough such that $k_n \geq k_0$, and in this case, $\max\limits_{k=1,\dots,k_n} \mathbb{E}(\Delta_k) = \max\limits_{k=1,\dots,k_0} \mathbb{E}(\Delta_k)$ and the above summation ends at k_0. By this way, for K with finite support, one

gets under conditions of Theorem 5.1 that for all $0 < c_0 < 1$ and $0 < \varepsilon < 2\min_k H_k(1-c_0)$:

$$\mathbb{P}\Big(\Big|\widehat{I}_{\max} - \max_{k=1,\dots,k_n} \mathbb{E}(\Delta_k)\Big| > \varepsilon\Big) = \mathcal{O}\Big(\exp\big(-c_0 p_{k_0}\frac{n\varepsilon^2}{2\sigma^2+2H\varepsilon}\big)\Big),$$

where $\sigma^2 = \max_k \sigma_k^2$ and $H = \max_k H_k$, yielding in turn that

$$\overline{\lim_{n\to\infty}} \Big|\widehat{I}_{\max} - \mathbb{E}(\Delta_{k_{\max}})\Big| = \mathcal{O}\Big(\sqrt{\frac{\ln n}{n}}\Big) \text{ a.s..}$$

(2) On the other hand, for a $\mathbb{N}$-valued random variable K and n large enough such that $k_n \geq k_{\max}$, one has $\max_{k=1,\dots,k_n} \mathbb{E}(\Delta_k) = \mathbb{E}(\Delta_{k_{\max}})$, and the bound (5.3) gives

$$\mathbb{P}\Big(\Big|\widehat{I}_{\max} - \max_{k=1,\dots,k_n} \mathbb{E}(\Delta_k)\Big| \geq \varepsilon\Big) = \mathcal{O}\Big(k_n \exp\big(-c_0 p_{k_n}\frac{n\varepsilon^2}{2\sigma^2+2H\varepsilon}\big)\Big).$$

The result follows with the choice $\varepsilon = \varepsilon_0\sqrt{\frac{\ln n}{np_{k_n}}}$ for some large enough $\varepsilon_0 > 0$ as soon as k_n has at most a logarithmic order.

□

Note that if K has a infinite support, the obtained rate of convergence depends strongly both on the choice of k_n and its associated value p_{k_n}. We give below two typical examples of expected rates.

Example 5.1. *(a) If $p_{k_n} \asymp k_n^{-\alpha}$ for some $\alpha > 0$, then the choice $k_n \simeq \ln(\ln n)$ gives the same rate as in the finite support case, while one gets a $\mathcal{O}(n^{-\frac{1}{2}}(\ln n)^{\frac{1+\alpha}{2}})$ for $k_n \simeq \ln n$. An example is furnished by the zeta distribution with parameter $q \in]1,+\infty[$ for which $\mathbb{P}(K = k) = \frac{k^{-q}}{\zeta(q)}$, $k = 1, 2, \cdots$.*
(b) For Poisson distribution $\mathcal{P}(\lambda)$, Stirling's approximation gives that $e^{-k_n \ln(k_n)}$ is predominant for p_{k_n}, it is equal to $(\ln n)^{-\ln(\ln(\ln n))}$ for $k_n = \ln(\ln n)$ and the associated a.s. rate of convergence of $\widehat{I}_{\max}$ to $\mathbb{E}(\Delta_{k_{\max}})$ is then of order $o(n^{-\beta})$ for all $0 < \beta < \frac{1}{2}$.

5.3. **Estimation of $k_{\max}$.** Now as soon as $\mathbb{E}(\Delta_{k_{\max}}) > \max_{k\geq 1, k\neq k_{\max}} \mathbb{E}(\Delta_k)$, the existence and uniqueness of $\widehat{k}_{\max} = \arg\max_{k=1,\dots,k_n} \widehat{I}_{k,n}$ are guaranteed, at least for n large enough. Theorem 5.1 allows us to derive the following result.

Proposition 5.2.
Suppose that assumptions of Theorem 5.1 are fulfilled, if k_n and p_{k_n} are such that $\sum_{n\geq 0} k_n \exp(-C_1 n p_{k_n}) < \infty$ for all $C_1 > 0$, one gets that almost surely for n large enough, $\widehat{k}_{\max} = k_{\max}$.

Proof. For n large enough to get $k_n \geq k_{\max}$, one has clearly

$$\mathbb{P}(\widehat{k}_{\max} \neq k_{\max}) \leq \sum_{\substack{k=1\\k\neq k_{\max}}}^{k_n} \mathbb{P}\big(\widehat{I}_{k,n} - \mathbb{E}(\Delta_k) > \widehat{I}_{k_{\max},n} - \mathbb{E}(\Delta_k)\big).$$

But for all $k \geq 0$ and $\varepsilon > 0$, and if $\mathcal{A} = \{\big|\widehat{I}_{k_{\max},n} - \mathbb{E}(\Delta_{k_{\max}})\big| \leq \varepsilon\}$

$$\begin{aligned}\mathbb{P}\big(\widehat{I}_{k,n}-\mathbb{E}(\Delta_k) > \widehat{I}_{k_{\max},n}-\mathbb{E}(\Delta_k)\big) &= \mathbb{P}\big(\widehat{I}_{k,n}-\mathbb{E}(\Delta_k) > \widehat{I}_{k_{\max},n}-\mathbb{E}(\Delta_k), \mathcal{A}\big)\\ &\quad + \mathbb{P}\big(\widehat{I}_{k,n} - \mathbb{E}(\Delta_k) > \widehat{I}_{k_{\max},n} - \mathbb{E}(\Delta_k), \mathcal{A}^c\big).\end{aligned}$$

Next on $\mathcal{A}$, the event $\{\widehat{I}_{k_{\max},n} - \mathbb{E}(\Delta_k) \geq \mathbb{E}(\Delta_{k_{\max}} - \Delta_k) - \varepsilon\}$ holds, so

$$\begin{aligned}\mathbb{P}(\widehat{k}_{\max} \neq k_{\max}) \leq \sum_{\substack{k=1\\k\neq k_{\max}}}^{k_n} \Big\{&\mathbb{P}\big(\widehat{I}_{k,n} - \mathbb{E}(\Delta_k) > \mathbb{E}(\Delta_{k_{\max}} - \Delta_k) - \varepsilon\big)\\ &+ \mathbb{P}\Big(\Big|\widehat{I}_{k_{\max},n} - \mathbb{E}(\Delta_{k_{\max}})\Big| > \varepsilon\Big)\Big\},\end{aligned}$$

the choice $\varepsilon = \frac{1}{2}\mathbb{E}(\Delta_{k_{\max}} - \Delta_k)$ now implies that

$$\mathbb{P}(\widehat{k}_{\max} \neq k_{\max}) \leq \sum_{k=1}^{k_n} \mathbb{P}\big(\widehat{I}_{k,n} - \mathbb{E}(\Delta_k) > \frac{a}{2}\big)$$

with a a positive real such that $\min_{k\geq 1, k\neq k_{\max}} \mathbb{E}\big(\Delta_{k_{\max}} - \Delta_k\big) \geq a > 0$. Finally if k_n and p_{k_n} are such that $\sum_{n\geq 0} k_n \exp(-C_1 n p_{k_n}) < \infty$ for all $C_1 > 0$, then replacing respectively σ_k^2 and H_k by $\sigma^2 = \max_k \sigma_k^2$ and $H = \max_k H_k$ in the bound (5.3), yields together with Borel Cantelli lemma that almost surely for n large enough, $\widehat{k}_{\max} = k_{\max}$. □

6. The case of random jumps

6.1. **The ordered case.** It is noteworthy that all the results of Section 5 remain true if one considers again i.i.d. copies of Δ_k with an arbitrary number of ordered random jumps T_k. A typical case is given in Example 2.3. To this ends, one may consider

$$\widetilde{I}_{k,n} = \Big(\frac{\sum_{i=1}^n \big|X_i(T_{k,i}) - X_i(T_{k,i}^-)\big| \mathbb{1}_{\{K_i\geq k\}}}{\sum_{i=1}^n \mathbb{1}_{\{K_i\geq k\}}}\Big)\mathbb{1}_{\{\sum_{i=1}^n \mathbb{1}_{\{K_i\geq k\}}>0\}}$$

to get a strongly consistent estimator of $\mathbb{E}\,\big|X(T_k) - X(T_k^-)\big|$, the continuous time framework guaranteeing that $\big((\,|X_i(T_{k,i}) - X_i(T_{k,i}^-)|, K_i),\ i = 1, \ldots, n\big)$ is well observed. Details are left to the reader.

6.2. **The non-ordered case.** In this part, we rather suppose that X admits k, $k \geq 2$, independent jumps at independent random instants $T_1, T_2, \dots, T_k$ on $]0,1[$ (so that the $T_k's$ are not necessarily ordered) with $\mathbb{E}(\Delta_1) > \cdots > \mathbb{E}(\Delta_k)$ and $\Delta_j := \left|X(T_j) - X(T_j^-)\right|$, $j = 1, \dots, k$. Our aim is to estimate $\mathbb{E}(\Delta_j)$, $j = 1, \dots, k$, from n i.i.d. copies of Δ_j on the basis of $\left\{ \left|X(T_{ji}^*) - X(T_{ji}^{*-})\right|, j = 1, \dots, k, i = 1, \dots, n \right\}$, where $0 < T_{1i}^* < \cdots < T_{ki}^* < 1$ (a.s.). In this part, we suppose also that $\Delta_1, \dots, \Delta_k$ are independent variables. Clearly, the difficulty is here to identify the jumps. To this end, we follow Bosq (2014)'s methodology for $k = 2$ jumps and generalize it for any arbitrary value of k.

6.2.1. *Case $k = 2$.* Since only strong consistency is established in Bosq (2014), we begin with the case $k = 2$ and make use of Bernstein's inequality to obtain almost sure rates of convergence. The methodology is the following. First remark that $\mathbb{E}(\Delta_1)$ and $\mathbb{E}(\Delta_2)$ are solutions of the quadratic equation $\prod_{j=1}^{2}(x - \mathbb{E}(\Delta_j)) = 0$, which can be written as $x^2 - sx + p = 0$ with $s = \mathbb{E}(\Delta_1) + \mathbb{E}(\Delta_2)$ and $p = \mathbb{E}(\Delta_1)\mathbb{E}(\Delta_2) = \mathbb{E}(\Delta_1\Delta_2)$ by independence of Δ_1 from Δ_2. Solutions $x_1 > x_2$ are given by $x_1 = \frac{1}{2}(s + \sqrt{s^2 - 4p})$ and $x_2 = \frac{1}{2}(s - \sqrt{s^2 - 4p})$. The next result shows that one may consistently estimate the intensities $\mathbb{E}(\Delta_1) > \mathbb{E}(\Delta_2)$ without the knowledge of their corresponding times of arrival and even, without ordering jumps according to their observed intensity! First to estimate $\mathbb{E}(\Delta_1)$ and $\mathbb{E}(\Delta_2)$, we set

$$\widehat{\Delta}_1 = \frac{1}{2}\Big((\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n}) + \sqrt{(\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n})^2 - 4\widetilde{\Delta}_{1:2,n}}\Big)$$
$$\widehat{\Delta}_2 = \frac{1}{2}\Big((\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n}) - \sqrt{(\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n})^2 - 4\widetilde{\Delta}_{1:2,n}}\Big)$$

with the observed

$$\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n} = \frac{1}{n}\sum_{i=1}^{n} \left|X_i(T_{1i}^*) - X(T_{1i}^{*-})\right| + \left|X_i(T_{2i}^*) - X(T_{2i}^{*-})\right|$$
$$\widetilde{\Delta}_{1:2,n} = \frac{1}{n}\sum_{i=1}^{n} \left|X_i(T_{1i}^*) - X(T_{1i}^{*-})\right| \left|X_i(T_{2i}^*) - X(T_{2i}^{*-})\right|.$$

We may derive the following result:

Proposition 6.1. *Suppose that $\left|X_i(T_{1i}) - X_i(T_{1i}^-)\right|$, $\left|X_i(T_{2i}) - X_i(T_{2i}^-)\right|$ and $\left|X_i(T_{1i}) - X_i(T_{1i}^-)\right| \left|X_i(T_{2i}) - X_i(T_{2i}^-)\right|$ fulfill conditions of Proposition 4.2 for all $i = 1, \dots, n$. Then for $j = 1, 2$, we get*

$$\left|\widehat{\Delta}_j - \mathbb{E}(\Delta_j)\right| = \mathcal{O}\Big(\sqrt{\frac{\ln n}{n}}\Big) \quad a.s..$$

Proof. We establish the result for $\widehat{\Delta}_1$, the proof being the same for $\widehat{\Delta}_2$. First, remark that

$$\widetilde{\Delta}_{1n} + \widetilde{\Delta}_{2n} \equiv \frac{1}{n}\sum_{i=1}^{n} \left|X_i(T_{1i}) - X(T_{1i}^-)\right| + \left|X_i(T_{2i}) - X(T_{2i}^-)\right|$$

and

$$\widetilde{\Delta}_{1:2,n} \equiv \frac{1}{n}\sum_{i=1}^{n} \left|X_i(T_{1i}) - X(T_{1i}^-)\right| \left|X_i(T_{2i}) - X(T_{2i}^-)\right|.$$

So, we study the a.s. behaviour of

$$2\widehat{\Delta}_1 = \overline{\Delta}_{1n} + \overline{\Delta}_{2n} + \sqrt{(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - 4\overline{\Delta}_{1:2,n}}$$

where the $\overline{\Delta}_{jn}$ and $\overline{\Delta}_{1:2,n}$ are built on the r.v.'s $\left|X_i(T_{ji}) - X_i(T_{ji}^-)\right|$. Note that as Δ_1 and Δ_2 are independent, we have

$$\begin{aligned} 2\mathbb{E}(\Delta_1) &= \mathbb{E}(\Delta_1 + \Delta_2) + \sqrt{(\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4\mathbb{E}(\Delta_1)\mathbb{E}(\Delta_2)} \\ &= \mathbb{E}(\Delta_1) + \mathbb{E}(\Delta_2) + \sqrt{(\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4\mathbb{E}(\Delta_1\Delta_2)}. \end{aligned}$$

So for $\psi_n = \sqrt{\frac{n}{\ln n}}$, we get the bound

$$\begin{aligned} \psi_n \left|\widehat{\Delta}_1 - \mathbb{E}(\Delta_1)\right| \le \frac{\psi_n}{2}\Big(&\left|\overline{\Delta}_{1,n} - \mathbb{E}\Delta_1\right| + \left|\overline{\Delta}_{2,n} - \mathbb{E}\Delta_2\right| \\ &+ \left|\sqrt{(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - 4\overline{\Delta}_{1:2,n}} - \sqrt{(\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4\mathbb{E}(\Delta_1\Delta_2)}\right| \Big). \end{aligned}$$

For $j = 1, 2$, we handled the terms $\mathbb{P}(\left|\overline{\Delta}_{j,n} - \mathbb{E}\Delta_j\right| \ge \varepsilon_0\psi_n^{-1})$ with Bernstein's inequality and Borel Cantelli's lemma for large enough positive ε_0. For the square-root term, remark that it may be written as

$$\frac{\psi_n}{2}\frac{\left|(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - (\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4(\overline{\Delta}_{1:2,n} - \mathbb{E}(\Delta_1\Delta_2))\right|}{\sqrt{(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - 4\overline{\Delta}_{1:2,n}} + \sqrt{(\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4\mathbb{E}(\Delta_1\Delta_2)}}.$$

The denominator converges almost surely to the positive limit $2(\mathbb{E}\Delta_1 - \mathbb{E}\Delta_2)$. Next, to treat the last term, just observe that

$$\begin{aligned} &\frac{\psi_n}{2}\left|(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - (\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2 - 4(\overline{\Delta}_{1:2,n} - \mathbb{E}(\Delta_1\Delta_2))\right| \\ &\quad\le \frac{\psi_n}{2}\left|(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - (\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2\right| + 2\psi_n\left|\overline{\Delta}_{1:2,n} - \mathbb{E}(\Delta_1\Delta_2)\right| \end{aligned}$$

and that

$$\begin{aligned} &\frac{\psi_n}{2}\left|(\overline{\Delta}_{1n} + \overline{\Delta}_{2n})^2 - (\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)^2\right| \\ &\quad= \frac{\psi_n}{2}\left|(\overline{\Delta}_{1n} + \overline{\Delta}_{2n}) - (\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)\right|\left|\overline{\Delta}_{1n} + \overline{\Delta}_{2n} + \mathbb{E}\Delta_1 + \mathbb{E}\Delta_2\right|. \end{aligned}$$

Again Bernstein's inequality and Borel Cantelli's lemma allow us to control the terms $\psi_n \left|\overline{\Delta}_{1:2,n} - \mathbb{E}(\Delta_1\Delta_2)\right|$ and $\psi_n \left|\overline{\Delta}_{jn} - \mathbb{E}\Delta_j\right|$, $j = 1, 2$ and the result follows since $\left|\overline{\Delta}_{1n} + \overline{\Delta}_{2n} + \mathbb{E}\Delta_1 + \mathbb{E}\Delta_2\right| \xrightarrow[n\to\infty]{a.s.} 2(\mathbb{E}\Delta_1 + \mathbb{E}\Delta_2)$. □

6.2.2. *The general case.* For arbitrary $k \geq 2$, $\mathbb{E}(\Delta_1), \ldots, \mathbb{E}(\Delta_k)$ are again solutions of

$$\prod_{j=1}^{k} \big(x - \mathbb{E}(\Delta_j)\big) = 0 = \sum_{j=0}^{k} a_j x^j,$$

where $a_k = 1$ and for $j = 1, \ldots, k$, Viète's formula gives:

$$a_{k-j} = (-1)^j \sum_{1\leq \ell_1 < \ldots < \ell_j \leq k} \mathbb{E}(\Delta_{\ell_1}) \cdots \mathbb{E}(\Delta_{\ell_j}).$$

Next, roots can be computed by finding the eigenvalues λ_j of the $k \times k$ matrix

$$A = \begin{pmatrix} -\frac{a_1}{a_0} & -\frac{a_2}{a_0} & -\frac{a_3}{a_0} & \ldots & -\frac{a_{k-1}}{a_0} & -\frac{1}{a_0} \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

and taking $\mathbb{E}(\Delta_j) = \lambda_j^{*-1}$, $j = 1, \ldots, k$, with $\lambda_1^* < \cdots < \lambda_k^*$ (see e.g. Pan, 1997). This eigenvalue method can be computationally expensive, but it is known to be fairly robust.

Concerning estimation in the case $k \geq 3$, we require independence between the Δ_j's as it allows to write the coefficients a_{k-j} under the more convenient form:

$$a_{k-j} = (-1)^j \sum_{1\leq \ell_1 < \ldots < \ell_j \leq k} \mathbb{E}\big(\Delta_{\ell_1} \cdots \Delta_{\ell_j}\big)$$

Next, one 'estimates' $\mathbb{E}\big(\Delta_{\ell_1} \cdots \Delta_{\ell_j}\big)$ by

$$\frac{1}{n} \sum_{i=1}^{n} \left|X_i(T_{\ell_1,i}) - X_i(T_{\ell_1,i}^-)\right| \cdots \left|X_i(T_{\ell_j,i}) - X_i(T_{\ell_j,i}^-)\right|.$$

Since these quantities are not observed and all summations are complete in Viète's formula, the trick is again to use observed $D_{\ell_j i} = \left|X(T^*_{\ell_j,i}) - X(T^{*-}_{\ell_j,i})\right|$ for $T^*_{1i} < \cdots < T^*_{ki}$ (a.s.), $i = 1, \ldots, n$, $j =$

$1, \ldots, k$. So the matrix A can be estimated by

$$\widehat{A} = \begin{pmatrix} -\frac{\widehat{a}_{1n}}{\widehat{a}_{0n}} & -\frac{\widehat{a}_{2n}}{\widehat{a}_{0n}} & -\frac{\widehat{a}_{3n}}{\widehat{a}_{0n}} & \ldots & -\frac{\widehat{a}_{k-1,n}}{\widehat{a}_{0,n}} & -\frac{1}{\widehat{a}_{0n}} \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

with

$$\begin{cases} \widehat{a}_{k-1,n} & = -\sum_{j=1}^{k} \overline{D}_{j,n} \\ \widehat{a}_{k-2,n} & = \sum_{1 \le j_1 < j_2 \le k} \overline{D}_{j_1:j_2,n} \\ & \vdots \\ \widehat{a}_{k-\ell,n} & = (-1)^{\ell} \sum_{1 \le j_1 < \ldots < j_\ell \le k} \overline{D}_{j_1:j_\ell,n} \\ & \vdots \\ \widehat{a}_{0,n} & = (-1)^k \prod_{j=1}^{k} \overline{D}_{j,n}. \end{cases}$$

where

$$\overline{D}_{j_1:j_\ell,n} = \frac{1}{n} \sum_{i=1}^{n} \left| X_i(T^*_{j_1,i}) - X_i(T^{*-}_{j_1,i}) \right| \cdots \left| X_i(T^*_{j_\ell,i}) - X_i(T^{*-}_{j_\ell,i}) \right|.$$

First note that positivity of the $D_{j,n}$'s and Descartes' rule of signs (1637) imply that the polynomial $\sum_{j=0}^{k-1} \widehat{a}_{j,n} x^j + x^k$ has 0 negative and at most k positive roots. Also, almost surely for n large enough, $\widehat{a}_{0,n} > 0$ which guarantees the existence of $\widehat{A}$. Moreover, the following proposition shows that we obtain strongly consistent estimators of the coefficients.

Proposition 6.2. *Under the assumption that the* $(\Delta_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, k)$ *are globally independent, we have*

$$\sum_{j=0}^{k-1} \widehat{a}_{j,n} x^j + x^k \xrightarrow[n\to\infty]{a.s.} \sum_{j=0}^{k-1} a_j x^j + x^k = 0.$$

Proof. Clearly, one has for each $\ell = 1, \ldots, k$:

$$\sum_{1 \le j_1 < \ldots < j_\ell \le k} \overline{D}_{j_1:j_\ell,n} \equiv \sum_{1 \le j_1 < \ldots < j_\ell \le k} \overline{\Delta}_{j_1:j_\ell,n} \xrightarrow[n\to\infty]{a.s.} \sum_{1 \le j_1 < \ldots < j_\ell \le k} \mathbb{E}(\Delta_{j_1} \cdots \Delta_{j_\ell}).$$

Then, one gets for $\ell = 1, \ldots, k$,

$$\sum_{1 \le j_1 < \ldots < j_\ell \le k} \mathbb{E}(\Delta_{j_1} \cdots \Delta_{j_\ell}) = \sum_{1 \le j_1 < \ldots < j_\ell \le k} \mathbb{E}(\Delta_{j_1}) \cdots \mathbb{E}(\Delta_{j_\ell}) = (-1)^{\ell} a_{k-\ell}$$

and the result follows from $\sum_{j=0}^{k} a_j x_k = 0$ with $a_k = 1$. $\square$

Consequently, one may expect to recover estimators of the k real roots, at least by approximation.

References

Antoniadis, A., X. Brossat, J. Cugliari, and J.-M. Poggi (2012). Prévision d'un processus à valeurs fonctionnelles en présence de non stationnarités. Application à la consommation d'électricité. *J. SFdS 153*(2), 52–78.

Besse, P., H. Cardot, and D. Stephenson (2000). Autoregressive forecasting of some climatic variation. *Scandinavian J. Statist. 27*(4), 673–687.

Billingsley, P. (1999). *Convergence of probability measures* (Second ed.). Wiley Series in Probability and Statistics: Probability and Statistics. New York: John Wiley & Sons Inc.

Blanke, D. and D. Bosq (2014). Observing jumps in discrete time. Applications. *In preparation.*

Bloznelis, M. and V. Paulauskas (1993). On the central limit theorem in $D[0,1]$. *Statist. Probab. Lett. 17*(2), 105–111.

Bosq, D. (2014). *Estimating and detecting jumps. Applications to D[0,1]-valued linear processes.* Springer. Festschrift in honour of Paul Deheuvels. To appear.

Bradley, R. C. (2007). *Introduction to strong mixing conditions. Vol. 1-3.* Kendrick Press, Heber City, UT.

Cardot, H. (1998). Convergence du lissage spline de la prévision des processus autorégressifs fonctionnels. *C. R. Acad. Sci. Paris Sér. I Math. 326*(6), 755–758.

Cont, R. and P. Tankov (2004). *Financial modelling with jump processes.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL.

Daffer, P. Z. and R. L. Taylor (1979). Laws of large numbers for $D[0,\ 1]$. *Ann. Probab. 7*(1), 85–95.

Damon, J. and S. Guillas (2005). Estimation and simulation of autoregressive hilbertian processes with exogenous variables. *Stat. Inference Stoch. Process. 8*(2), 185–204.

Descartes, R. (1637). *La Géométrie, Livre Troisième.* Leyde: Jan Maire.

El Hajj, L. (2011). Théorèms limites pour les processus autorégressifs à valeurs dans $D[0,1]$. *C. R. Math. Acad. Sci. Paris 349*(13-14), 821–825.

El Hajj, L. (2013). *Théorèmes limites pour les processus autorégressifs à valeurs dans D[0,1].* Ph. D. thesis, Université Paris 6, Paris. Décembre.

Ferraty, F. and P. Vieu (2006). *Nonparametric functional data analysis.* Springer Series in Statistics. New York: Springer. Theory and practice.

Garba, S. M., A. B. Gumel, and M. R. Abu Bakar (2008). Backward bifurcations in dengue transmission dynamics. *Math. Biosci. 215*(1), 11–25.

Jacq, V., P. Albert, and R. Delorme (2005, août). Le mistral. Quelques aspects des connaissances actuelles. *La météorologie 50*, 30–38.

Kargin, V. and A. Onatski (2008). Curve forecasting by functional autoregression. *J. Multivariate Anal. 99*(10), 2508–2526.

Marion, J. M. and B. Pumo (2004). Comparaison des modèles ARH(1) et ARHD(1) sur des données physiologiques. *Ann. I.S.U.P. 48*(3), 29–38.

Mas, A. (2004). Consistance du prédicteur dans le modèle ARH(1): le cas compact. *Ann. I.S.U.P. 48*(3), 39–48.

Mourid, T. (2002). Estimation and prediction of functional autoregressive processes. *Statistics 36*(2), 125–138.

Pan, V. Y. (1997). Solving a polynomial equation: some history and recent progress. *SIAM Rev. 39*(2), 187–220.

Pestman, W. R. (1995). Measurability of linear operators in the Skorokhod topology. *Bull. Belg. Math. Soc. Simon Stevin 2*(4), 381–388.

Pumo, B. (1998). Prediction of continuous time processes by $C_{[0,1]}$-valued autoregressive process. *Stat. Inference Stoch. Process. 1*(3), 297–309.

Ruiz-Medina, M. D. (2012). Spatial functional prediction from spatial autoregressive Hilbertian processes. *Environmetrics 23*(1), 119–128.

Shiryaev, A. N. (1996). *Probability* (Second ed.), Volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York. Translated from the first (1980) Russian edition by R. P. Boas.

Tankov, P. and E. Voltchkova (2009). Jump-diffusion models: a practitioners guide. *Banque et Marchés 99*, 1–24.