A note on exponential bounds for smoothed processes

D Bosq

Pub. Inst. Stat. Univ. Paris
63, fasc.1, 2019, 73-86

A note on exponential bounds for smoothed processes

D. Bosq

Sorbonne Universités, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, F-75005 Paris, France

Abstract

We regularize a modification of the Poisson process and get an exponential bound for sequences of the smoothed process. It is also possible to consider various exponential bounds in the context of density estimation and apply them to the smoothed empirical distribution function. For these smoothed processes, results for the space $C[0,1]$ may be derived as well as for some other Banach spaces.

Keywords: Smooth Poisson process, regular density estimation, exponential rate, Banach spaces, Spline smoothing, Regression.

MSC 2010: 62G07, 62G08, 60G55, 60F10

1 Introduction

Our aim is to regularize the Poisson process and the empirical distribution function. Various exponential rates are envisaged. Also, consistency appears in the context of Banach spaces.

Here $0 = T_0 < T_1 < ... < T_n < ...$ is the arrival time, where $\lambda > 0$ is the intensity of Poisson process with $N_t \sim \mathcal{P}(\lambda t)$. Now we modify the Poisson process by setting

$$N_t(\delta) = \delta . N_t,\ \delta > 0$$

and, by interpolation with $L_t(\delta)$, one obtains

$$\sup_{t \geq 0} |L_t(\delta) - N_t(\delta)| \leq \frac{\delta}{2}$$

(see Section 2).

In Section 3, we study the empirical distribution function and derive exponential bounds from results on $\|f_n - f\|_\infty$. Note that existing exponential rates are rather intricate, then, we obtain a simple result, not necessarily at a higher level but that can be directly derived from Bosq [1998], see Lemma 3.1 and Proposition 3.3.

In Section 4, other classical smoothing techniques, namely based on splines and existence of derivatives, are presented in the context of regression estimation. For recent references concerning derivatives estimation, one may refer to Prakasa Rao [2017], Giné and Sang [2013] with smooth derivative and Chaubey and Shirazi [2015] in the case of wavelets estimation. Concerning Mas and Pumo [2009], they compute the derivatives of the curve data in a functional framework.

In Section 5, in this context of smoothness, we examine the case of $C[0,1]$ which yields immediate results. Results for some general Banach spaces are also possible.

Appendix is postponed until Section 6 for results about density estimation.

For a global bibliographic study, note that smoothing techniques appear in a large number of fields which result in a very large number of publications which cannot be listed exhaustively Thus, we only give here some references. In Smaga [1978]and Efromovich [2001]continuous m-th convergence in derivative is considered. In Reiss [1981], the empirical distribution function and the smooth distribution function are studied. Concerning Sarda [1993], a cross-validation procedure is introduced in order to select the smoothing parameter of the kernel distribution estimate. Also, the polygonal smoothing is a simple and natural method in order to regularize the empirical distribution function (cf. Blanke and Bosq, 2018).

In Servien [2009], various estimation of the distribution function are reviewed. In Cavalier and Hengartner [2009] and Wang et al. [2010], various results are obtained for smooth distribution functions. In Leblanc [2012], jumps discontinuity are smoothed with Berstein polynomials. Finally, in Xue and Wang [2010] sophisticated spline regression is considered. In Chacón et al. [2014] smooth distribution holds in the context of Fourier process.

Concerning exponential bounds, we may consider, for example, results of Roussas [1996a] based on Bennett and Hoeffding inequality; Henriques and Oliveira [2005]with exponential rate for kernel density. In Giné and Nickl [2009] exponential rates are also given with application to adaptive estimation. Of course, the main exponential bound comes from Massart

[1990] and Talagrand [1994].

Finally, a mixing concerning smoothness, density function and Poisson process appears in Chaubey et al. [2011].

2 Smooth Poisson process

Here $0 = T_0 < T_1 < ... < T_n < ...$ is the arrival time, where $\lambda > 0$ is the intensity of Poisson process with $N_t \sim \mathcal{P}(\lambda t)$.

Now, we modify the Poisson process, by setting $\delta > 0$ with

$$N_t(\delta) = \delta . N_t \tag{2.1}$$

then we may define the smoothed Poisson process in the following propostion.

Proposition 2.1. *We put*

$$L_t(\delta) = \frac{\delta t}{2T_1}.1_{\{0 \le t \le T_1\}} + \sum_{k=1}^{\infty} \left(\frac{\delta.(t - T_k)}{(T_{k+1} - T_k)} + \frac{(2k-1)\delta}{2} \right).1_{\{T_k < t \le T_{k+1}\}}$$

and

$$\sup_{t \ge 0} |L_t(\delta) - N_t(\delta)| \le \frac{\delta}{2}$$

Proof. See Figure 1. □

Example 2.2. A refraction with light is envisaged in the context of vertical plan (see Bass, 1978, p.120-121).

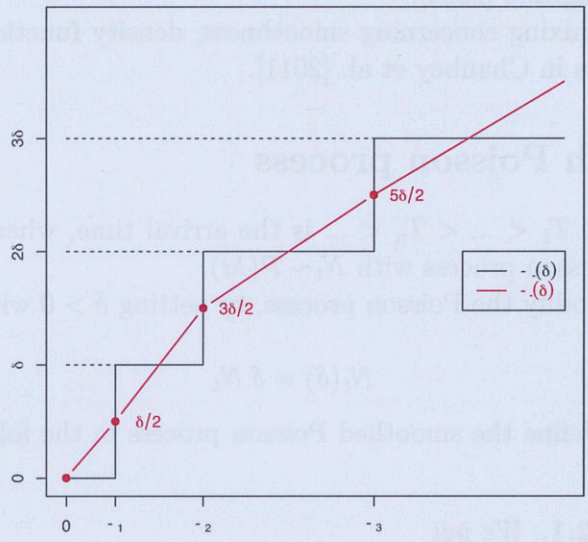

Figure 1: Smooth Poisson process

Lemma 2.3. *We get*

$$EN_t(\delta) = \lambda t\delta$$

and

$$VN_t(\delta) = \lambda t\delta^2$$

Also

$$P(N_t(\delta) = k\delta) = \exp(-\lambda t).\frac{(\lambda t)^k}{k!}, \quad k \in \mathbb{N}$$

Proof. Clear from $N_t(\delta) = \delta.N_t$. □

Another lemma may be useful.

Lemma 2.4. *We have*

$$P(T_k \leq t \leq T_{k+1}) = P(T_k \leq t) - P(T_{k+1} \leq t), \ \ t \geq 0, \ \ k \in \mathbb{N}$$

and also

$$\sum_k P(N_t(\delta) = k\delta) = \sum_k P(T_k \leq t \leq T_{k+1}) < \infty, \quad t > 0, \quad \delta > 0$$

Proof. First, note that

$$P(T_k \leq t) = P(T_k < t), \quad t \geq 0, \, k \in \mathbb{N} \tag{2.}$$

since T_k follows the $\Gamma(k, \lambda)$ distribution. Then

$$\{t < T_{k+1}\} = (\{T_k \leq t\} \cap \{t < T_{k+1}\}) \cup (\{T_k > t\} \cap \{t < T_{k+1}\})$$

thus

$$\{t < T_{k+1}\} = \{T_k \leq t < T_{k+1}\} \cup \{t < T_k < T_{k+1}\}$$

and since $T_k < T_{k+1}$ is trivial one obtains the desired result, cf (2.2).

On the other hand, we get

$$P(N_t(\delta) = k\delta) = P(N_t = k) = P(T_k \leq t < T_{k+1}) = \exp(-\lambda t).\frac{(\lambda t)^k}{k!}$$

and by using Stirling formula, one obtains

$$k! = \sqrt{2\pi k}.\left(\frac{k}{e}\right)^k (1 + \frac{1}{12k} + \mathcal{O}(\frac{1}{k^2}))$$

then

$$P(T_k \leq t < T_{k+1}) = \frac{\exp(-\lambda t)}{\sqrt{2\pi k}} \left(\frac{\lambda t e}{k}\right)^k \frac{1}{1 + \frac{1}{12k} + \mathcal{O}(\frac{1}{k^2})}$$

and

$$\sum_k P(T_k \leq t < T_{k+1}) \leq \sum_k \frac{\exp(-\lambda t)}{\sqrt{2\pi k}} \left(\frac{\lambda t e}{k}\right)^k \frac{1}{1 + \frac{1}{12k}} < \infty.$$

Then, Lemma 2.4 gives

$$P\left(\cap_{n=1}^{\infty} \cup_{k \geq n} \left(T_k \leq t < T_{k+1}\right)\right) = 0.$$

Now, to derive an exponential bound, we consider of i.i.d. modified Poisson processes and put $N_{i,t}(\delta) = \delta.N_i(t)$, $1 \leq i \leq n$.

Proposition 2.5. *There exists $c > 0$ such that*

$$P(|\frac{1}{n}\sum_{i=1}^{n}\left(N_{i,t}(\delta) - \lambda t\delta\right)| \geq \eta) \leq 2\exp\left(-n\frac{\eta^2}{4\lambda t\delta^2 + 2c\eta\delta}\right)$$

with $t > 0, \eta > 0, \delta > 0$.

Proof. First we consider $\delta = 1$, then

$$E\left(\exp aN_t\right) = \sum_{k=0}^{\infty} \exp(ak).\exp(-\lambda)\frac{\lambda^k}{k!} = \exp(-\lambda).\exp^{\exp(a)\lambda} < \infty$$

Now, Cramer's condition (see Remark 2.6) is equivalent to existence of $E(\exp \gamma N_t)$, $\gamma > 0$ (See 1.17 p.24 and p.27 in Bosq, 1998).

Then, from Bernstein's inequality (Bosq, 1998, p.24) one obtain

$$P(|\frac{1}{n}\sum_{i=1}^{n}\left(N_{i,t}(1) - \lambda t\right)| \geq \eta) \leq 2\exp\left(-n\frac{\eta^2}{4\lambda t + 2c\eta}\right).$$

Next by using $N_{i,t}(\delta) = \delta.N_{i,}(t)$ $1 \leq i \leq n$ and Lemma 2.3 we get the desired result. □

Remark 2.6. Here c satisfies

$$E|N_t - \lambda t|^k \leq c^{k-2}k!\lambda t < \infty,\ k \geq 3$$

which is the Cramer's condition.

Next, we may derive the following exponential bound for sequences of i.i.d. smoothed Poisson processes $L_{i,t}(\delta_n)$,$1 \leq i \leq n$, $\delta_n > 0$, based on the $N_{i,t}(\delta_n)$.

Proposition 2.7. *There exists $c > 0$ such that for $\eta > \frac{\delta}{2} > 0$ and $t > 0$:*

$$P(|\frac{1}{n}\sum_{i=1}^{n}\left(L_{i,t}(\delta) - \lambda t\delta\right)| \geq \eta) \leq 2\exp\left(-n\frac{(\eta - \frac{\delta}{2})^2}{4\lambda t\delta^2 + 2c(\eta - \frac{\delta}{2})\delta}\right).$$

Proof. Straightforward from propositions 2.1 and 2.5. □

Proposition 2.8. *For sequences of i.i.d. smoothed Poisson processes with* $\delta = \delta_n \to 0$, *we obtain that almost surely for* n *large enough*

$$\left| \frac{1}{n\delta_n} \sum_{i=1}^{n} L_{i,t}(\delta_n) - \lambda t \right| \leq \frac{1}{2}, \quad t > 0.$$

Proof. In Proposition 2.7, we have $\eta - \frac{\delta_n}{2} > \frac{\eta}{2}$ for n large enough so, for $\eta = \eta_0 \delta_n$ with $\eta_0 > \frac{1}{2}$:

$$P(|\frac{1}{n\delta_n} \sum_{i=1}^{n} L_{i,t}(\delta_n) - \lambda t| \geq \eta_0) \leq 2 \exp\left(-n \frac{\eta_0^2}{16\lambda t + 4c\eta_0} \right)$$

and the result follows with Borel-Cantelli lemma. □

Remark 2.9. Smoothness can also be considered in the context of continuous derivative (see Section 4).

Also, the smooth Poisson process allows to get the space $C[0,1]$ and Section 5 may give direct results in this case.

3 Smooth empirical distribution function and density estimation

For $X_1, \ldots, X_n$ i.i.d. on [0,1] with density f, we may set

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K(\frac{X_i - x}{h_n}), \quad 0 \leq x \leq 1$$

where, for example, $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2}), h_n > 0$.

Next, the smooth estimation of the distribution function $F(t) = \int_0^t f(s)ds$, for $0 \leq t \leq 1$ follows with

$$F_n(t) = \int_0^t f_n(s)ds, \quad 0 \leq t \leq 1.$$

Lemma 3.1. *It holds*

$$\sup_{0 \leq t \leq 1} |F_n(t) - F(t)| \leq \sup_{0 \leq x \leq 1} |f_n(x) - f(x)|$$

Proof. We easily get

$$\sup_{0\le t\le 1}|F_n(t)-F(t)| \le \sup_{0\le t\le 1}\left|\int_0^t (f_n(s)-f(s))ds\right|$$

and

$$\sup_{0\le t\le 1}|F_n(t)-F(t)| \le \int_0^1 \sup_{0\le x\le 1}|f_n(x)-f(x)|\,dx$$

and the result is clear. □

Lemma 3.1 allows to derive exponential bounds from results on density estimation. We may refer to Henriques and Oliveira [2005] and Giné and Nickl [2009], however the exponential rate is rather intricate. Then, in the context of Lemma 3.2 and Proposition 3.3, derived from results in Bosq [1998], we obtain more simple results, not necessarily at a higher level, but sufficient to derive immediate results for F_n.

Lemma 3.2. *If f is of class C_2, then*

$$P(|f_n(x)-f(x)|\ge\eta) \le 2\exp(-\frac{\eta^2}{16\,\|f\|_\infty\,\|K\|_2^2+4\,\|K\|_\infty\,\eta}\,nh_n),\ 0\le x\le 1$$

with $\eta > ah_n^2$ and h_n small enough with $nh_n \to \infty$.

Proof. See Appendix 1. □

Now, we use the following condition satisfied for a gaussian density K:
A1- $|K(v)-K(u)| \le \ell\,|v-u|$ $u,v\in[0,1]$, $\ell>0$.
Then

Proposition 3.3. *From A1 and if f is of class C_2, one obtains*

$$P(\sup_{0\le x\le 1}|f_n(x)-f(x)|\ge\eta) \le 2(k_n+1)\exp(-\frac{\eta^2 nh_n}{64\,\|f\|_\infty\,\|K\|_2^2+8\,\|K\|_\infty\,\eta}),$$

with $\eta > \max(\frac{2\ell\gamma_n}{h_n^2}, ah_n^2)$, $h_n\to 0, \gamma_n\to 0, \gamma_n(k_n+1)=1$, for n large enough.

Proof. See Appendix 2. □

Example 3.4. If $h_n = (\ln n)^a n^{-a}$, $0 < a < 1$ and $\gamma_n = n^{-b}$, $2a < b$ then, we obtain

$$P(\|f_n - f\|_\infty \geq \eta) \leq 2\, n^b \exp(-\frac{\eta^2}{64\,\|f\|_\infty \|K\|_2^2 + 8\,\|K\|_\infty \eta} n^{1-a}(\ln n)^a),$$

$\eta > \max(an^{-2a}(\ln n)^{2a}, 2\ell n^{-b+2a}(\ln n)^{-2a})$. So with $a = \frac{1}{5}$ and $\eta = \eta_0(\frac{\ln n}{nh_n})^{\frac{1}{2}} = \eta_0(\frac{\ln n}{n})^{\frac{2}{5}}$, $\eta_0 > 0$ large enough, we get by Borel-Cantelli lemma

$$\limsup_{n\to\infty}(\frac{n}{\ln n})^{\frac{2}{5}}\ \|f_n - f\|_\infty < \infty \text{ a.s.}$$

Also by Lemma 3.1, the same result holds for $\|F_n - F\|_\infty$.

4 Smoothing Splines

Various smoothing variants can be envisaged with in particular smoothing splines. We give here a small review in the context of functional estimation.

For example, one may consider the model:

$$Y_i = f(X_i) + \varepsilon_i,\ 1 \leq i \leq n \tag{4.1}$$

where ε_i are independent and identically distributed with $E\varepsilon_i = 0$,$\mathrm{Var}(\varepsilon_i) = \sigma^2 > 0$, $1 \leq i \leq n$.

The smoothing spline estimation of f consists in minimizing

$$f_n = \arg\min_f \left[\frac{1}{n}\sum_{i=1}^{n}(Y_i - f(X_i))^2 + \mu \int_0^1 |f^{(p)}(x)|^2 dx\right]$$

with $\mu \geq 0$ and $p \in \mathbb{N}$ (See Efromovich, 1999 p. 347-348).

Another possible model, cf Besse and Cardot [1996], get the following form:

$$y_j = f(t_j) + \eta_j,\ 1 \leq j \leq q.$$

Now, we suppose that $t \to f(t)$ belongs to a Sobolev space, say

$$W^2[0,1] = \left\{ f : [0,1] \to \mathbb{R};\, f, f' \text{a.c.}, \int_0^1 |f''(t)|^2 dt \prec \infty \right\}$$

equipped with the scalar product

$$\langle f, g \rangle = \int_0^1 (f(t)\, g(t) + f''(t)\, g''(t))dt \; f, g \in W^2$$

Then the spline approximation is the solution of

$$\min_{f \in W^2[0,1]} \left\{ \frac{1}{q} \sum_{j=1}^{q} (f(t_j) - y_j)^2 + \ell \int_0^1 |f''(t)^2 dt \right\}$$

where ℓ is a smoothing parameter

Remark 4.1. In the context of (4.1), one may derive results for spline distribution function. We may write for the empirical distribution function $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq x\}}$ as

$$F_n(X_i) = F(X_i) + \varepsilon_i, \;\; i = 1 \leq i \leq n$$

and

$$F_n = \arg \min_{F \in C^p[0,1]} \left[\frac{1}{n} \sum_{i=1}^{n} (F_n(X_i) - F(X_i))^2 + \mu \int_0^1 |F^{(p)}(x)|^2 dx \right], \; p \in \mathbb{N}$$

with μ strictly positive. Now results holding for splines may be applied, see for example Rice and Rosenblatt [1983] and Xue and Wang [2010] for monotonic distribution function estimation.

Remark 4.2. Generalizations with various splines smoothing may be found in Berlinet and Thomas-Agnan [2004], see chapter 3, and also in Wahba [1990].

5 Banach Space

Since $M = (L_t(\delta), 0 \leq t \leq 1)$ is a smooth Poisson process, one may consider that M takes its values in $C[0,1]$ and various results appear for i.i.d. sequences of such processes.

Now we put

$$\bar{M}_n(t, \delta) = \frac{1}{n} \sum_{i=1}^{n} (L_i(t, \delta) - \lambda t \delta), \; 0 \leq t \leq 1$$

Then, some simple results appear by considering $C[0,1]$:

Proposition 5.1. *We get*

- $n^{1/4}(\log n)^{-\beta}\bar{M}_n(t,\delta) \to 0\ a.s., 0 \leq t \leq 1, \beta > \frac{1}{2}$
- $n\,E(\bar{M}_n^2(t,\delta)) \to \sum_{k=-\infty}^{+\infty} E(X_0(\delta,t)\,X_k(\delta,t))$ where $X_k(\delta,t)) = (L_k(t,\delta) - \lambda t\delta)$, $0 \leq t \leq 1$, $k \in \mathbb{N}$.
- Also $\frac{1}{\sqrt{n}}\bar{M}_n(t,\delta) \Rightarrow N$where N is standard gaussian.

Proof. See Bosq [2000] cf p.168, Theorem 6.11, Theorem 6.12, Theorem 6.15 (see p.169-175). □

Remark 5.2. It is also possible to obtain results in Banach space of type 2 as well as the law of the iterated logarithm (see Ledoux and Talagrand, 1991).

6 Appendix

6.1 Proof of Lemma 3.2

Consider Bernstein inequality, see for example (1.18) in Bosq, 1998, that is:

$$P(|f_n(x) - Ef_n(x)| > \eta) \leq 2\exp(-\frac{\eta^2}{4\,\|f\|_\infty \|K\|_2^2 + 2\,\|K\|_\infty \eta} nh_n) \quad (6.1)$$

Now we want to replace $Ef_n(x)$ with $f(x)$. By using Bosq, 1998 in (2.22) p.48, one obtains a classical decomposition of the bias:

$$|Ef_n(x) - f(x)| \leq ah_n^2 \quad (6.2)$$

where a does not depend on n and f is of class C_2.
Then, we have

$$|f_n(x) - f(x)| \leq |f_n(x) - Ef_n(x)| + |Ef_n(x) - f(x)|$$

and

$$|f_n(x) - f(x)| \leq |f_n(x) - Ef_n(x)| + ah_n^2$$

Next, we get

$$P(|f_n(x) - f(x)| > \eta) \leq P(|f_n(x) - f(x)| > \eta - a\frac{h_n^2}{2})$$

and since $h_n \to 0$, (6.1) gives the result with $\eta > ah_n^2$ and $\eta - a\frac{h_n^2}{2} > \frac{\eta}{2}$ for n large enough.

6.2 Proof of Proposition 3.3

Now, we consider discretization by setting $j\gamma_n, 0 \leq j \leq k_n$where $(k_n+1)\gamma_n = 1$ and k_n is an integer. Then

$$f_n(x_j) = \frac{1}{nh_n}\sum_{i=1}^{n} K(\frac{X_i - x_j}{h_n}),\ j\gamma_n \leq x_j \leq (j+1)\gamma_n,\ 0 \leq j \leq k_n$$

and

$$f_n(j\gamma_n) = \frac{1}{nh_n}\sum_{i=1}^{n} K(\frac{X_i - j\gamma_n}{h_n}),\ 0 \leq j \leq k_n$$

Then, by using A1, we easily get

$$|f_n(x_j) - f_n(j\gamma_n)| \leq \frac{\ell\gamma_n}{h_n^2}, \tag{6.3}$$

and

$$|Ef_n(x_j) - Ef_n(j\gamma_n)| \leq \frac{\ell\gamma_n}{h_n^2}, \tag{6.4}$$

Now we put

$$f_n(x)-Ef_n(x) = [f_n(x)-f_n(j\gamma_n)]+[f_n(j\gamma_n)-Ef_n(j\gamma_n)]+[Ef_n(j\gamma_n)-Ef_n(x)]$$

where $j\gamma_n \leq x \leq (j+1)\gamma_n$.
Then, from (6.3) and (6.4), one obtains

$$|(f_n(x) - Ef_n(x)) - (f_n(j\gamma_n) - Ef_n(j\gamma_n))| \leq 2\frac{\ell\gamma_n}{h_n^2}. \tag{6.5}$$

Now, let us put

$$A = \max_{0\leq j\leq k_n} |f_n(j\gamma_n) - Ef_n(j\gamma_n)|$$

and

$$B = \sup_{0\leq x\leq 1} |f_n(x) - Ef_n(x)|$$

Then, from (6.5), for $j\gamma_n \leq x \leq (j+1)\gamma_n$

$$|f_n(x) - Ef_n(x)| \leq |f_n(j\gamma_n) - Ef_n(j\gamma_n)| + 2\frac{\ell\gamma_n}{h_n^2}$$

thus

$$B \leq A + 2\frac{\ell\gamma_n}{h_n^2}$$

and

$$p_n = P(B \geq \varepsilon) \leq P(A \geq \varepsilon - 2\frac{\ell\gamma_n}{h_n^2})$$

where $\varepsilon - 2\frac{\ell\gamma_n}{h_n^2} > 0$ and $\frac{\gamma_n}{h_n^2} \to 0$.
Now we have

$$p_n \leq \sum_{j=0}^{k_n} P(|f_n(j\gamma_n) - Ef_n(j\gamma_n)| \geq \varepsilon - 2\frac{\ell\gamma_n}{h_n^2})$$

and from (6.1) and n large enough with $\varepsilon - 2\frac{\ell\gamma_n}{h_n^2} > 0$:

$$p_n \leq 2(k_n + 1)\exp(-\frac{\varepsilon^2}{16\|f\|_\infty \|K\|_2^2 + 4\|K\|_\infty \varepsilon} nh_n).$$

The proof for the bias is then similar to Lemma 3.2.

References

J. Bass. *Cours de Mathématique.* Tome 2. Masson, 1978.

A. Berlinet and C. Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics.* Kluwer academic publishers, Dordrecht, 2004.

P. Besse and H. Cardot. Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Canad. Journ. Statist.*, 24(4):467–487, 1996.

D. Blanke and D. Bosq. Polygonal smoothing of the empirical distribution function. *Stat. Inference Stoch. Process.*, 21(2):263–287, 2018.

D. Bosq. *Nonparametric statistics for stochastic processes. Estimation and prediction.* Lecture Notes in Statistics, 110. Springer-Verlag, New-York, Second edition, 1998.