C. R. Acad. Sci. Paris, Ser. I 351 (2013) 139–141

Contents lists available at SciVerse ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Probability Theory/Statistics

Remark on the finite-dimensional character of certain results of functional statistics

Remarque sur le caractère fini-dimensionnel de certains résultats de statistique fonctionnelle

Jean-Marc Azaïs [a], Jean-Claude Fort [b]

[a] *Université de Toulouse, IMT, ESP, 31062 Toulouse cedex 9, France*
[b] *Université Paris-Descartes, SPC, MAP5, 45, rue des Saints-Pères, 75006 Paris, France*

A R T I C L E   I N F O

Article history:
Received 20 November 2012
Accepted after revision 12 February 2013
Available online 27 February 2013

Presented by Jean-Pierre Kahane

A B S T R A C T

This note shows that some assumption on small balls probability, frequently used in the domain of functional statistics, implies that the considered functional space is of finite dimension. To complete this result an example of L^2 process is given that does not fulfill this assumption.

© 2013 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

R É S U M É

Cette note montre qu'une hypothèse concernant les probabilités de petites boules, fréquemment utilisée en statistique fonctionnelle, implique que la dimension de l'espace fonctionnel considéré est finie. Un exemple de processus L^2, ne vérifiant pas cette hypothèse, vient compléter ce résultat.

© 2013 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. The result

In several functional statistics papers (cf. [1,2,4,6,5,3]), the following hypothesis is used:

(H) *Let x be a point of the space $\mathcal{X}$ where a functional variable X lives. The space $\mathcal{X}$ is equipped with a semi-distance and $B(x,h)$ is the ball with center x and radius $h > 0$. We set $\varphi_x(h) = \mathbb{P}(X \in B(x,h))$ and we assume*:

$$\inf_{h\in[0,1]} \int_0^1 \varphi_x(ht)\,\mathrm{d}t/\varphi_x(h) \geqslant \theta_x > 0,$$

where the parameter θ_x is locally bounded away from zero.

It is well known that a semi-distance is in fact a distance on a quotient space, say $\overline{\mathcal{X}}$. The aim of this note is to prove that (H) implies that $\overline{\mathcal{X}}$ is of finite dimension.

E-mail addresses: jean-marc.azais@math.univ-toulouse.fr (J.-M. Azaïs), jean-claude.fort@parisdescartes.fr (J.-C. Fort).

1631-073X/$ – see front matter © 2013 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.crma.2013.02.004

Proof. Without loss of generality, we can assume that $\theta_x < 1/2$ and that the space is equipped with a true distance. Letting $F_x(h) = \int_0^h \varphi_x(t)\,\mathrm{d}t$, we have:

$$\frac{1}{h}\int_0^h \varphi_x(t)\,\mathrm{d}t/\varphi_x(h) = \frac{F_x(h)}{hF'_x(h)} \geqslant \theta_x.$$

By integration, we obtain:

$$F_x(h)/F_x(1) \geqslant h^{1/\theta_x}. \tag{1}$$

Since φ is non-decreasing, we have:

$$F_x(h) \leqslant h\varphi_x(h)$$

and thus

$$\varphi_x(h) \geqslant h^{1/\theta_x - 1} F_x(1).$$

Let $x \in \mathcal{X}$ such that the parameter θ_y has a positive lower bound for y in the ball $B(x, h_0)$. This implies that $\varphi_x(h_0) > 0$. By a scaling, we can assume for simplicity that $h_0 = 1$. For all y in $B(x, 1/4)$, and for all $h \in [1/2, 1]$,

$$\varphi_y(h) \geqslant \varphi_x(1/4) \geqslant \left(\frac{1}{4}\right)^{1/\theta - 1} F_x(1),$$

where θ is the uniform lower bound for θ_y for y in $B(x, 1)$. By integration:

$$F_y(1) \geqslant 1/2\left(\frac{1}{4}\right)^{1/\theta - 1} F_x(1)$$

and

$$\varphi_y(h) \geqslant h^{1/\theta - 1} F_y(1) \geqslant 1/2\left(\frac{1}{4}\right)^{1/\theta - 1} h^{1/\theta - 1} F_x(1).$$

This implies that there exist at most $\mathcal{O}(h^{1-\theta})$ disjoint balls of radius h in $B(x, 1/4)$. Then the same set of balls but with radius $2h$ is a covering, which implies in turn that the box (or entropy) dimension of $B(x, 1/4)$ is finite. Since the Hausdorff dimension is smaller than the box dimension, it is also finite. □

Remark 1. Suppose, in addition, that the probability distribution function of X satisfying (H) admits a density with respect to some natural positive measure μ and suppose that this density is locally upper- and lower-bounded by M and m, respectively:

$$m\mu\big(B(x, h)\big) \leqslant \varphi_x(h) \leqslant M\mu\big(B(x, h)\big).$$

We have then:

$$\exists d > 0,\ C_2 > 0,\ h_0 > 0, \quad 0 \leqslant h \leqslant h_0 \quad \Longrightarrow \quad 1/C_2 h^d \leqslant \mu\big(B(x, h)\big) \leqslant C_2 h^d,$$

which means that μ shares the same property as the distribution of X. For example, this property does not hold for the Wiener measure.

We may note that the classical random processes, i.e. Brownian motion and more general Gaussian processes, for which the probability of small balls is known, do not satisfy (H) (see [7] for more details).

2. Random series of functions

Here we consider $X(t)$ a stochastic process constructed as a random series of functions. More precisely:

$$X(t) = \sum_{n=1}^{\infty} \alpha_n Z_n \varphi_n(t) \tag{2}$$

where

– $(Z_n, n \geqslant 1)$ is a sequence of independent real random variables, with mean 0 and variance 1, with absolutely continuous densities $(f_n, n \geqslant 1)$ *w.r.t.* the Lebesgue measure. This is the case for most of the usual continuous distributions of probability: exponential, normal, polynomial, gamma, beta, etc.;
– $(1; \varphi_n, n \geqslant 1)$ is an orthonormal basis of $L^2([0,1])$;
– $(\alpha_n, n \geqslant 1)$ is a sequence of positive real numbers $\sum_{n=1}^{+\infty} \alpha_n^2 < \infty$.

The sum (2) converges in $L^2([0,1])$.

In particular, the form (2) covers all the Gaussian processes through the Karhunen–Loève decomposition.

Then we have:

Proposition 2.1. $\lim_{h\to 0} h^d \mathbb{P}(\|X\|_2 \leqslant h) = 0$ *for any* $d \geqslant 0$, *so that the process* $X(t)$ *does not fulfill the assumption* (H) *for the* L^2 *norm at the point zero.*

The proof is based on the properties of the convolution:

Let f and g the absolutely continuous densities of probability of two independent random variables U and V. Then U^2 (resp. V^2) has the density:

$$p_{U^2}(u) = \frac{\tilde{f}(\sqrt{u})}{\sqrt{u}}, \quad \text{resp. } p_{V^2}(v) = \frac{\tilde{g}(\sqrt{v})}{\sqrt{v}},$$

where $\tilde{f}$ and $\tilde{g}$ are the symmetrized of f and g, $\tilde{f} = 1/2(f(x) + f(-x))$. It follows that $U^2 + V^2$ has for density:

$$C(x) = \int_0^x \frac{\tilde{f}(\sqrt{u})}{\sqrt{u}} \frac{\tilde{g}(\sqrt{x-u})}{\sqrt{x-u}} \, \mathrm{d}u = \int_0^1 \frac{1}{\sqrt{v(1-v)}} \tilde{g}\big(\sqrt{(1-v)x}\big) \tilde{f}(\sqrt{vx}) \, \mathrm{d}v,$$

for $x \geqslant 0$ and 0 elsewhere.

It is easy to see that for any $0 < A < B$, the function C is Lipschitz on $[A, B]$. At $x = 0$ it takes the value 0 with right limit $\beta(\frac{1}{2}, \frac{1}{2}) f(0) g(0)$. Now, if we make the convolution product of two such functions C_1 and C_2, we obtain a function vanishing for $x \leqslant 0$, continuous at 0 and Lipschitz on any compact interval of $\mathbb{R}$, thus absolutely continuous. Using a classical result, making the convolution product of k such absolutely continuous functions yield a $\mathcal{C}^{k-1}$ function whose $(k-1)$-th derivative is absolutely continuous.

Then, applying iteratively this result, we conclude that the density of the variable $\|X\|_2^2 = \sum_{n=1}^{\infty} \alpha_n^2 Z_n^2$ is infinitely differentiable with all its derivatives vanishing at 0. The claimed property follows. □

Remark 2. If the process X lives in $L^p[0,1]$, $p \in (2, \infty]$, we have $\|X\|_2^2 \leqslant \|X\|_p^2$, thus:

$$\mathbb{P}\big(\|X\|_p \leqslant h\big) \leqslant \mathbb{P}\big(\|X\|_2 \leqslant h\big).$$

References

[1] Florent Burba, Frédéric Ferraty, Philippe Vieu, *k*-Nearest neighbour method in functional nonparametric regression, J. Nonparametr. Stat. 21 (4) (2009) 453–469.

[2] Sophie Dabo-Niang, Estimation du mode dans un espace vectoriel semi-normé (Mode estimation in a semi-normed vector space), C. R. Acad. Sci. Paris, Ser. I 339 (9) (2004) 659–662 (in French).

[3] Frédéric Ferraty, Aldo Goia, Philippe Vieu, Régression non-paramétrique pour des variables aléatoires fonctionnelles mélangeantes, C. R. Acad. Sci. Paris, Ser. I 334 (3) (2002) 217–220.

[4] Frédéric Ferraty, Nadia Kudraszow, Philippe Vieu, Nonparametric estimation of a surrogate density function in infinite-dimensional spaces, J. Nonparametr. Stat. 24 (2) (2012) 447–464.

[5] Frédéric Ferraty, Ali Laksaci, Amel Tadj, Philippe Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, J. Stat. Plan. Infer. 140 (2) (2010) 335–352.

[6] F. Ferraty, I. Van Keilegom, P. Vieu, Regression when both response and predictor are functions, J. Multivariate Anal. 109 (2012) 10–28.

[7] Wenbo V. Li, Qi-Man Shao, Gaussian processes: Inequalities, small ball probabilities and applications, in: C.R. Rao, D. Shanbhag (Eds.), Stochastic Processes: Theory and Methods, Handbook of Statistics, vol. 19, Elsevier, New York, 2001, pp. 533–598.