C. R. Acad. Sci. Paris, t. 332, Série I, p. 761–764, 2001
Statistique/*Statistics*
(Systèmes dynamiques/*Dynamical Systems*)

Estimation de la densité invariante de systèmes dynamiques en dimension 1

Clémentine PRIEUR

Mathématiques, Université de Cergy-Pontoise, site Saint-Martin, 2, avenue Adolphe-Chauvin, 95302 Cergy-Pontoise cedex, France
Courriel : prieur@math.u-cergy.fr

(Reçu le 22 décembre 2000, accepté le 9 février 2001)

Résumé. Dans cette Note nous établissons un théorème limite central pour les estimateurs à noyau de la densité invariante de certains systèmes dynamiques en dimension 1. La preuve de ce théorème se fait en deux étapes essentielles : l'étude de la vitesse de convergence de la variance de l'estimateur puis l'obtention du théorème limite central par une variante de la méthode de Lindeberg–Rio [6]. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Density estimation for one-dimensional dynamical systems

Abstract. *In this Note we obtain a central limit theorem for standard kernel invariant density estimates of one-dimensional dynamical systems. The two main steps in the proof of this theorem are the following: the study of speed of convergence for the variance of the estimator and then a variation on the Lindeberg–Rio method* [6]. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Un problème classique est celui de l'estimation de la densité marginale f d'une suite stationnaire $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires dépendantes. On sait que si la suite $(X_n)_{n\in\mathbb{N}}$ satisfait des conditions de mélange [7] ou de dépendance faible [3], on a le résultat suivant :

$$\sqrt{nb_n}\,\big[\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x)\big] \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, f(x)\int_{-\infty}^{+\infty} K^2(t)\,\mathrm{d}t\right), \tag{1}$$

où $\widehat{f}_n(x)$ est un estimateur classique à noyau défini de la façon suivante :

$$\widehat{f}_n(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x - X_k}{b_n}\right), \tag{2}$$

Note présentée par Paul DEHEUVELS.

S0764-4442(01)01909-7/FLA

avec une suite $(b_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ et un noyau $K : \mathbb{R} \to \mathbb{R}_+$ à support compact D tels que :

$$b_n \xrightarrow[n\to+\infty]{} 0, \quad nb_n \xrightarrow[n\to+\infty]{} +\infty, \quad \text{et} \quad \int_D K(t)\,\mathrm{d}t = 1, \quad \int_D K^2(t)\,\mathrm{d}t > 0.$$

Dans cette Note, nous nous restreignons à la classe de systèmes dynamiques :

$$\forall\, n \geqslant 1, \quad X_n = T^n X_0, \tag{3}$$

où T est une fonction d'un intervalle $I \subset \mathbb{R}$ dans lui-même admettant une mesure de probabilité invariante μ_0 absolument continue par rapport à la mesure de Lebesgue, et X_0 une variable aléatoire de loi μ_0. Notons f la densité de μ_0 par rapport à la mesure de Lebesgue. Le but est d'estimer f. Nous travaillons sous une hypothèse de contrôle des corrélations. Avant de l'écrire, définissons l'ensemble de fonctions suivant :

DÉFINITION 1.1. – $\mathcal{VB}$ est l'ensemble des fonctions h de I dans $\mathbb{R}$ à variation bornée et de norme $\mathbb{L}^1$ finie (*voir* [8] par exemple). Si $\|h\|_{\mathcal{VB}} := \mathcal{V}(h) + \|h\|_1$, où $\mathcal{V}(h)$ est la variation de h et $\|\cdot\|_1$ la norme standard sur $\mathbb{L}^1$, alors $\|\cdot\|_{\mathcal{VB}}$ est une norme et $\mathcal{VB}$ muni de cette norme est un espace de Banach.

Nous pouvons maintenant supposer :

$$\begin{aligned} &\exists\, \kappa > 0,\ \exists (d_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}},\ \sum_{k=0}^{+\infty} d_k < \infty \text{ tels que } \forall\, n \geqslant 0,\ \forall h,\ k \in \mathcal{VB}, \\ &\bigl|\operatorname{Cov}\bigl(h(X_0), k(X_n)\bigr)\bigr| \leqslant \kappa\, \|k\|_1\, \|h\|_{\mathcal{VB}}\, d_n, \end{aligned} \tag{4}$$

où Cov désigne la covariance par rapport à la mesure invariante μ_0. Ce contrôle de covariances permet de conclure à l'existence d'un théorème de type (1). C'est l'objet du théorème 3.1 énoncé dans le paragraphe 3.

2. Hypothèses sur le modèle et exemples

Dans cette partie nous donnons les hypothèses techniques requises pour montrer le théorème limite central. Prenons $K \in \mathcal{VB}$. Nous considérons ensuite un intervalle $I \subset \mathbb{R}$, et T une application de I dans lui-même. Notons λ la mesure de Lebesgue, et $\operatorname{int}(I)$ l'intérieur de I. Nous supposons :

- pour tout k dans $\mathbb{N}$, pour tout x dans $\operatorname{int}(I)$, $\lim_{t\to 0^-} T^k(x-t) =: T^k(x^+)$ et $\lim_{t\to 0^+} T^k(x-t) =: T^k(x^-)$ existent ;
- pour tout k dans $\mathbb{N}^*$, posons $D_-^k := \{x \in \operatorname{int}(I),\ T^k(x^-) = x\}$ et $D_+^k := \{x \in \operatorname{int}(I),\ T^k(x^+) = x\}$. Soit $I_1 := \operatorname{int}(I) \setminus \bigcup_{k\in\mathbb{N}^*} (D_-^k \cup D_+^k) =: \operatorname{int}(I) \setminus B$. Alors $\lambda(B) = 0$;
- T admet au moins une mesure de probabilité invariante μ_0 qui est absolument continue par rapport à la mesure de Lebesgue : $\mathrm{d}\mu_0 = f\,\mathrm{d}\lambda$;
- soit $S := \operatorname{supp}(\mu_0)$ le support de μ_0. Si $I_2 := \{x \in S \subset I,\ f \text{ est continue en } x \text{ et } f(x) > 0\}$, alors $\lambda(S \setminus I_2) = 0$.

X_0 est une variable aléatoire de loi μ_0. Nous rappelons que $(X_n)_{n\in\mathbb{N}}$ est définie par (3). Enfin, nous faisons l'hypothèse de contrôle des corrélations décrite par l'inégalité (4) de l'introduction.

Dans le cas où la décroissance dans (4) est exponentielle, il existe des exemples simples (qui peuvent s'étendre par conjugaison) :

- *les fonctions de type Lasota–Yorke*. – Pour une définition plus précise de ces fonctions, nous renvoyons au chapitre 3 de Viana [8] ou à [5].
  - Les fonctions « tentes » (*voir* figure 1, où le sommet est le point $[0.5, 0.75]$).
  - Les fonctions r-adiques (*voir* figure 2, où $r = 8/3$).
- *Les fonctions à nombre infini dénombrable d'intervalles de monotonicité*. – La fonction de Gauss par exemple. Elle est définie sur $[0,1]$ par $T(x) = \frac{1}{x} - \left[\frac{1}{x}\right]$ si $x \neq 0$ et $T(0) = 0$.

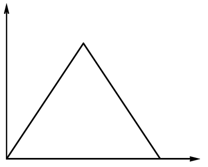

Figure 1. – Fonction « tente ».

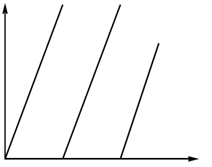

Figure 2. – Fonction r-adique

3. Énoncé du théorème limite central et éléments de preuve

Dans un premier temps, énonçons le résultat qui fait l'objet de cette Note. C'est un théorème limite central portant sur le processus d'estimation centré : $Y_n(x) := \sqrt{nb_n}\big(\widehat{f}_n(x) - \mathbf{E}\,\widehat{f}_n(x)\big)$, où $\widehat{f}_n(x)$ est défini à partir de $(X_n)_{n\in\mathbb{N}}$ comme dans (2).

THÉORÈME 3.1. – *Supposons que T satisfait les hypothèses du paragraphe 2, et que pour tout $1 \leqslant i \leqslant \ell$, $x_i \in I_1 \cap I_2$. Alors les répartitions fini-dimensionnelles $(\overline{Y_n}(x_1), \dots, \overline{Y_n}(x_\ell))$, du processus*

$$\overline{Y_n}(x) \equiv \frac{Y_n(x)}{\sqrt{f(x)\int_{-\infty}^{+\infty} K^2(t)\,\mathrm{d}t}}$$

convergent en loi, quand n tend vers $+\infty$, vers $\mathcal{N}(0,\, I_\ell)$.

Remarque 3.1. – Pour $\ell = 1$ ce théorème est un théorème limite central. Si l arbitraire, si J est un intervalle compact inclus dans I, ce résultat montre que la suite des processus $\{\overline{Y_n}(x);\ x \in J\}_{n\geqslant 1}$ n'est pas tendue dans $\mathcal{C}(J)$; ses répartitions fini-dimensionnelles convergent en effet vers celles du bruit blanc gaussien $\dot{W}$.

Pour montrer ce théorème nous commençons par montrer un résultat de convergence en moyenne quadratique pour cet estimateur.

LEMME 3.1. – *Supposons que T satisfait les hypothèses du paragraphe 2. Soit $x \in I_1 \cap I_2$. Alors il existe une suite $(\varepsilon_n)_{n\in\mathbb{N}} \in (\mathbb{R})^{\mathbb{N}}$ qui tend vers 0 quand n tend vers l'infini telle que* :

$$\operatorname{Var}\big(\widehat{f}_n(x)\big) = \frac{f(x)\int_D K^2(s)\,\mathrm{d}s}{nb_n}\,[1 + \varepsilon_n].$$

Remarque 3.2. – Une première évaluation de la vitesse de convergence de $\widehat{f}_n(x)$ (en $\mathcal{O}(1/(nb_n^2))$) est donnée dans [2] (ou plus récemment dans [4]). Notre résultat donne lieu au théorème limite central, il fournit donc « la » vitesse de convergence.

Pour montrer un tel résultat dans un cadre de dépendance faible, il est tout à fait raisonnable de supposer l'existence de densités jointes régulières pour les couples $(X_0, X_k)_{k\in\mathbb{N}^*}$. Ici, les lois jointes sont singulières et cette absence de régularité dissocie l'étude de la dépendance faible de celle des systèmes dynamiques. Pour montrer le lemme 3.1, nous utilisons ici les deux lemmes ci-dessous :

LEMME 3.2. – *Supposons que T satisfait les hypothèses du paragraphe 2. Soit $x \in I_1 \cap I_2$. Alors pour chaque k dans $\mathbb{N}^*$, il existe une suite $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ telle que*

$$\operatorname{Cov}\left(K\left(\frac{x - X_0}{b_n}\right), K\left(\frac{x - X_k}{b_n}\right)\right) = b_n\,\varepsilon(n,k).$$

LEMME 3.3. – *Supposons que T satisfait les hypothèses du paragraphe* 2. *Soit* $x \in I_1 \cap I_2$. *Alors*

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\operatorname{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) \xrightarrow[n\to\infty]{} 0.$$

Le point essentiel dans la preuve du lemme 3.2 est l'étude, pour $k \geqslant 1$ de $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right]$. Pour n grand, pour $k \geqslant 1$ fixé nous avons :

$$\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right] \leqslant b_n \int_D K(t)f(x-tb_n)K\left(\frac{x-T^k(x-tb_n)}{b_n}\right)\mathrm{d}t.$$

Comme K est à support D compact, nous avons

$$x-T^k(x-tb_n) \xrightarrow[n\to+\infty]{} x-T^k(x\mp) \neq 0, \quad \text{uniformément en } t \in D,\ t \geqslant 0 \text{ ou } t \in D,\ t \leqslant 0.$$

Donc, pour k fixé, $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)K\left(\frac{x-X_k}{b_n}\right)\right] = \mathrm{o}(b_n)$, les termes $\mathbb{E}\left[K\left(\frac{x-X_0}{b_n}\right)\right]\mathbb{E}\left[K\left(\frac{x-X_k}{b_n}\right)\right]$ se traitent de façon classique. Nous utilisons ensuite le lemme 3.2 ainsi que le contrôle des corrélations (4) pour montrer le lemme 3.3. La fin de la preuve du théorème 3.1 (*cf.* [5]) repose (comme dans [3]) sur une variation de la méthode de Lindeberg–Rio.

Remarque 3.3. – L'étude du biais est classique, elle fait appel à la régularité de f qui dans certains cas [1] se déduit de celle de T. D'autre part, nous pouvons étendre nos résultats au cas où X_0 ne suit plus forcément la loi invariante. $(X_n)_{n\in\mathbb{N}}$ n'est alors plus stationnaire. L'étude du cas non stationnaire repose essentiellement sur les propriétés de l'opérateur de Perron–Frobenius associé à T défini de la façon suivante :

$$\mathcal{L} : \begin{cases} \mathcal{BV} \longrightarrow \mathcal{BV}, \\ \omega \longmapsto \mathcal{L}\omega(x) = \displaystyle\sum_{Ty=x}\frac{\omega(y)}{|T'(y)|}. \end{cases}$$

Références bibliographiques

[1] Amroun A., Systèmes dynamiques perturbés. Sur une classe de fonctions ζ dynamiques, Thèse de doctorat de l'Université Paris-6, 1995.

[2] Bosq D., Guégan D., Nonparametric estimation of the chaotic function and the invariant measure of a dynamical system, Statis. & Probab. Let. 25 (1995) 201–212.

[3] Coulon-Prieur C., Doukhan P., A triangular central limit theorem under a new weak dependence condition, Statis. & Probab. Let. 47 (2000) 61–68.

[4] Maës J., Statistique non paramétrique des processus dynamiques réels en temps discret, Thèse de doctorat de l'Université Paris-6, 1999.

[5] Prieur C., Functional estimation of dynamical systems, Preprint de l'Université de Cergy-Pontoise, n° 7/00, 2000.

[6] Rio E., Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes, Probab. Theory Relat. Fields 104 (1996) 255–282.

[7] Robinson P.M., Nonparametric estimators for time series, J. Time Ser. Anal. 4–3 (1983) 185–207.

[8] Viana M., Stochastic dynamics of deterministic systems, Instituto de Matematica Pura e Aplicada, IMPA 21 (1997).

Résumé d'un texte qui sera conservé pendant cinq ans dans les Archives de l'Académie et dont copie peut être obtenue.