

Sur quelques résultats de convergence dans l'inférence d'une classe de processus ponctuels

Aboubakary Diakhaby^{a,b}

^a UFR de mathématiques appliquées et d'informatique, Université Gaston Berger, BP 234, Saint-Louis, Sénégal

^b Équipe PRISME, Université René Descartes, UFR de mathématiques et informatique, 45, rue des Saints-Pères, 75270 Paris cedex 06, France

Reçu le 20 mars 2001 ; accepté après révision le 28 janvier 2002

Note présentée par Paul Deheuvels.

Résumé

Dans [7], nous avons étudié les conditions suffisantes de convergence uniforme presque complète et la vitesse de convergence, d'estimateurs non paramétriques par la méthode du noyau de la densité de la mesure moyenne et de la régression dans le cas d'un processus de Poisson ; cette étude généralise ces résultats à une classe plus large de processus ponctuels, notée \mathcal{F} . *Pour citer cet article* : A. Diakhaby, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 597–602. © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Convergence results for point processes estimators

Abstract

Our aim is to generalize some results obtained for a Poisson point process in [7], to a general point process. Those results are in field of complete convergence of two like Parzen–Rosenblatt estimates of density of mean measure function and regression curves. Those estimates are defined from the superposition of n i.i.d. point processes as:

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^m K\left(\frac{x - X_i^{(n)}}{h(n)}\right) \quad \text{and} \quad \hat{\Psi}_n(x) = \frac{\sum_{i=1}^m Y_i K\left(\frac{x - X_i^{(n)}}{h(n)}\right)}{\sum_{i=1}^m K\left(\frac{x - X_i^{(n)}}{h(n)}\right)},$$

where m is the number of seem generics points of the superposition. We give some sufficient conditions for the convergence of those kernel-like estimators. *To cite this article*: A. Diakhaby, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 597–602. © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Abridged English version

Let f_0 be a point process defined on a probability space (Ω, \mathcal{A}, P) which values in R^2 . For any Borel set $A \subset R^2$, denote by $N(f_0, A)$ the number of points of the process lying in A and set $p = N(f_0, R^2)$. Denote by f^* the derivative with respect to Lebesgue measure of the mean measure μ of f_0 and by λ the modulus of continuity of the kernel K . We suppose that the point process belongs on the class \mathcal{F} defined in the

Adresses e-mail : diakhaby@ugb.sn (A. Diakhaby); aboubakary.Diakhaby@math-info.univ-paris5.fr (A. Diakhaby).

sequel. For $\theta > 0$, we set $\chi_n = (\min(h^{-1}, \sqrt{nh^2})/(\log(n))^\theta$, and $r(x) = \int_R y f^*(x, y) dy$, for $0 < \varepsilon < \frac{a+b}{2}$, $\Delta(\varepsilon) = [a + \varepsilon, b - \varepsilon]$. Define $\Psi(x) = E(Y_j/X_j = x)$, $j = 1, \dots, p_0$, for $p > 0$ and $\Psi(x) \equiv 0$ for $p = 0$; and $f(x) = \int_R f^*(x, y) dy$. We can estimate $f(x)$ by $\hat{f}_n(x)$ and $\Psi(x)$ by $\hat{\Psi}_n(x)$ and formulate now our main results.

THEOREM 0.1. –

(a) If λ satisfies the condition (i) then

$$h^{-2} = o\left(\frac{n}{\log(n)}\right) \implies \lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |\hat{f}_n(x) - f(x)| = 0 \text{ almost completely (a.c.).}$$

(b) Moreover if r is continuous in Δ and if Y_i , $i = 1, \dots, p$, are uniformly bounded then for all ε such that

$$\inf_{x \in \Delta(\varepsilon)} f(x) > 0 \text{ and } \sup_{x \in \Delta(\varepsilon)} |\Psi(x)| < \infty,$$

$$h^{-2} = o\left(\frac{n}{\log(n)}\right) \implies \lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |\hat{\Psi}_n(x) - \Psi(x)| = 0 \text{ a.c.}$$

THEOREM 0.2. –

(a) If f is differentiable with bounded derivative in Δ and if λ satisfies the condition (i') then for $\theta > \frac{1}{2}$

$$h^{-2} = o\left(\frac{n}{\log(n)}\right) \implies \lim_{n \rightarrow \infty} \chi_n \sup_{x \in \Delta(\varepsilon)} |\hat{f}_n(x) - f(x)| = 0 \text{ a.c.}$$

(b) Moreover if r is continuous in Δ then for all ε such that $\inf_{x \in \Delta(\varepsilon)} f(x) > 0$ and $\sup_{x \in \Delta(\varepsilon)} |\Psi(x)| < \infty$ and for $\theta > \frac{1}{2}$

$$h^{-2} = o\left(\frac{n}{\log(n)}\right) \implies \lim_{n \rightarrow \infty} \chi_n \sup_{x \in \Delta(\varepsilon)} |\hat{\Psi}_n(x) - \Psi(x)| = 0 \text{ a.c.}$$

1. Introduction

Dans [7], nous avons étudié les conditions suffisantes de convergence uniforme presque complète et la vitesse de convergence, d'estimateurs non paramétriques par la méthode du noyau de la densité de la mesure moyenne et de la régression dans le cas d'un processus de Poisson ; cette étude généralise ces résultats à une classe plus large de processus ponctuels, notée \mathcal{F} . Les processus ponctuels constituent un cadre général pour étudier la répartition aléatoire de points dans un certain domaine ; ces points peuvent avoir chacun plusieurs attributs. Un exemple simple est une image qui est caractérisée par les points qui la constituent ainsi que la couleur de chaque point. Les motivations d'une telle étude sont réelles et nombreuses [8] : on trouve des applications en traitement d'images, en biométrie, dans l'étude de la performance des systèmes et réseaux informatiques. Dans ce dernier cas, si des éléments actifs d'un réseau, par exemple des routeurs et des switches, sont localisés en des points d'un espace X et chaque élément actif reçoit des paquets de plusieurs protocoles ou qui ont des adresses de destination différentes, alors ces différentes possibilités prennent des valeurs dans un espace produit $Y = \prod_{i=1}^n Y_i$, ou les Y_i sont des ensembles d'adresses IP par exemple ou des plages de ports. On obtient ainsi un processus ponctuel à valeur dans l'espace $X \times Y$. La modélisation d'un tel système par un processus de Poisson est une approximation grossière car ne prenant pas en compte les phénomènes de grappes. Dans l'étude de ce type de problèmes, les paramètres densité de la mesure moyenne et fonction de régression sont d'un grand intérêt.

Les résultats sur la convergence des estimateurs de type Parzen–Rosenblatt dans le cadre des processus ponctuels concernent le plus souvent les processus de Poisson [3] et [7]. D'autres types d'estimateurs ont cependant fait l'objet d'études dans un cadre plus large, concernant notamment la densité de la mesure

moyenne : Marion [10] utilise la méthode de l’histogramme et Mendes [9], celle de l’histogramme et de la fenetre mobile. Dia [4], utilise le régressogramme pour estimer la fonction de régression et propose dans [5] l’estimateur de la densité qui sera la base de notre étude. Une description de ces différentes méthodes d’estimation non paramétrique est proposée dans le livre [1].

2. Notations et hypothèses

2.1. Notations

Etant donné un processus ponctuel f_0 sur (Ω, \mathcal{A}, P) à valeur dans R^2 , nous noterons l’effectif aléatoire de f_0 , sur tout borélien $A \subset R^2$, par $N(f_0, A)$ et l’effectif aléatoire total sur R^2 par $p = N(f_0, R^2)$. On définit ainsi une famille de variables aléatoires à valeurs entières, indexée par les éléments de la tribu borélienne de R^2 , notée \mathcal{B} ; $N(f_0, \cdot) : \Omega \times \mathcal{B} \rightarrow \overline{N}$. Pour un borélien A fixé, $N(f_0, A)$ est une variable aléatoire tandis que pour presque tout $\omega \in \Omega$, $N(f_0^\omega, \cdot)$ sera une mesure. Pour $p \neq 0$, les p points génériques du processus seront notés par : $(X_1, Y_1), \dots, (X_p, Y_p)$. Soit $f_{(n)}$ la superposition de n processus ponctuels $\{N(f_i, \cdot), i = 1, \dots, n\}$, indépendants définis sur (Ω, \mathcal{A}, P) de même distribution que f_0 ; $N(f_{(n)}, \cdot)$ est défini pour tout borélien A par $N(f_{(n)}, A) = \sum_{i=1}^n N(f_i, A)$. On désignera par m l’effectif total de ce processus sur R^2 et si $m \neq 0$, par $(X_1^{(n)}, Y_1^{(n)}), \dots, (X_m^{(n)}, Y_m^{(n)})$ les m observations de la superposition. La mesure moyenne d’un processus f_0 est définie pour tout borélien A par $\mu(A) = EN(f_0, A)$. Le processus marginal première projection sur R de f_0 , noté f_{01} , admet une mesure moyenne μ_1 ayant pour dérivée de Radon–Nikodym la fonction, notée f , définie par : $f(x) = \int_{-\infty}^{\infty} f^*(x, y) dy$ où f^* désigne la dérivée par rapport à la mesure de Lebesgue de la mesure moyenne μ de f_0 . Pour $p = p_0 \geq 1$, on supposera :

- (a) $E(Y_j / X_1 = x_1, \dots, X_{p_0} = x_{p_0}) = E(Y_j / X_j = x_j), j = 1, \dots, p_0$;
- (b) La fonction $\Psi(x) = E(Y_j / X_j = x), j = 1, \dots, p_0$, est indépendante de j et de p_0 . Si $p = 0$, on pose $\Psi \equiv 0$.

2.2. Estimateurs

Nous estimons f par \hat{f}_n et Ψ par $\hat{\Psi}_n$ définis par

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^m K\left(\frac{x - X_i^{(n)}}{h(n)}\right), \quad \hat{\Psi}_n(x) = \frac{\sum_{i=1}^m Y_i K\left(\frac{x - X_i^{(n)}}{h(n)}\right)}{\sum_{i=1}^m K\left(\frac{x - X_i^{(n)}}{h(n)}\right)},$$

où K désigne un noyau de convolution [1], c’est-à-dire une densité de probabilité telle que :

$$\int_{-\infty}^{\infty} |K(y)| dy < \infty, \quad \sup_{-\infty < y < \infty} |yK(y)| < \infty, \quad \lim_{y \rightarrow +\infty} |yK(y)| = 0,$$

et où h , le pas de discrétisation, est une fonction de n telle que $\lim_{n \rightarrow +\infty} h = 0$ et $\lim_{n \rightarrow +\infty} nh = +\infty$.

Pour plus de détails sur la construction de ces estimateurs, on peut se référer à [6,8]. Dans le cadre de l’échantillon de vecteurs aléatoires le livre de Bosq et Lecoutre [1] est une référence.

2.3. Hypothèses

2.3.1. Hypothèses générales

Nous noterons par \mathcal{F} la classe des processus ponctuels tels que :

HG1 : La mesure moyenne μ est finie, et on note $\|\mu_1\| = \mu_1(R)$. L’effectif total p vérifie la propriété \mathcal{P} suivante : p admet des moments de tous ordres ; et

$$\exists \Gamma > 0, \exists \Lambda > 0, \forall n \geq 2, \quad |E(p - \|\mu_1\|)^n| \leq \frac{n!}{2} \Gamma^{n-2} \Lambda^2.$$

HG2 : La densité f de la mesure moyenne μ_1 est continue sur son support noté $\Delta = [a, b], a, b \in R$.

Remarque. – Cette classe contient notamment les processus de Poisson et les résultats obtenus constituent une généralisation de ceux connus pour les variables aléatoires. En effet, soit le processus ponctuel f_0 associé à une variable aléatoire $Z = (X, Y)$ de densité continue g sur un compact de R^2 . Nous avons $f_0 \in \mathcal{F}$ avec $p = 1$ presque sûrement, la densité de la mesure moyenne μ_1 est ici la densité de probabilité f de la variable aléatoire $X : \forall A \in B_R, EN(f_{01}, A) = P(X \in A)$.

2.3.2. *Hypothèses sur le noyau K*

On supposera dans tout ce qui suit que :

K est continu à support compact tel que $\int K(y) dy = 1$.

Il existe une fonction $\lambda : R^+ \rightarrow R^+$ décroissante vers zéro telle que : $\forall x, x', |K(x) - K(x')| \leq \lambda(|x - x'|)$. On considère les hypothèses suivantes sur λ :

- (i) $\lambda(1/n^\alpha) = o(h)$, pour un α positif;
- (i') $\lambda(1/n^\alpha) = o(h^2)$, pour un α positif.

3. Résultats

On suppose dans la suite que les processus sont dans la classe \mathcal{F} .

Posons pour $\theta > 0, \chi_n = (\min(h^{-1}, \sqrt{nh^2})) / (\log(n))^\theta, r(x) = \int_R y f^*(x, y) dy$ et $\Delta(\varepsilon) = [a + \varepsilon, b - \varepsilon]$.

3.1. Sur la convergence

THÉORÈME 1. –

- (a) Si λ vérifie l'hypothèse (i), alors la relation : $h^{-2} = o(\frac{n}{\log(n)})$, implique pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |\widehat{f}_n(x) - f(x)| = 0 \text{ presque complètement (p.co.).}$$

- (b) Si de plus, r est continue sur Δ , les variables $Y_i, i = 1, \dots, p$, sont uniformément bornées et s'il existe un ε positif tel que : $\inf_{x \in \Delta(\varepsilon)} f(x) > 0$ et $\sup_{x \in \Delta(\varepsilon)} |\Psi(x)| < \infty$, alors on a :

$$h^{-2} = o\left(\frac{n}{\log(n)}\right) \implies \lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |\widehat{\Psi}_n(x) - \Psi(x)| = 0 \text{ p.co.}$$

3.1.1. *Preuve*

Principe : Nous utiliserons essentiellement la même démarche que dans [8] ainsi que le lemme suivant qui donne une inégalité de type Bernstein–Fréchet pour les processus qui sont dans la classe \mathcal{F} .

LEMME 1. – Soit ϕ une fonction borélienne et bornée de R^2 dans R , et pour $m > 0, (X_1^{(n)}, Y_1^{(n)}), \dots, (X_m^{(n)}, Y_m^{(n)})$ les variables restreintes génériques de la superposition de n processus ponctuels iid de même loi que f_0 . Si f_0 vérifie l'hypothèse HG1, $\forall t, 0 \leq t < \Gamma^2/\Lambda$,

$$P\left(\left|\sum_{i=1}^m \left[\phi(X_i^{(n)}, Y_i^{(n)}) - E\left(\sum_{i=1}^m \phi(X_i^{(n)}, Y_i^{(n)})\right)\right]\right| \geq nt\right) \leq 2 \exp\left(-\frac{nt^2}{4\Gamma^2}\right).$$

Démonstration. – Posons $Z = \sum_{i=1}^p \phi(X_i, Y_i) - E(\sum_{i=1}^p \phi(X_i, Y_i))$.

Montrons que si p vérifie la propriété \mathcal{P} alors Z vérifie aussi cette propriété. Nous avons, après une double sommation sur les différents processus de la superposition dans un premier temps et ensuite sur les points génériques de chaque processus, que :

$$\sum_{i=1}^m \phi(X_i^{(n)}, Y_i^{(n)}) - E\left(\sum_{i=1}^m \phi(X_i^{(n)}, Y_i^{(n)})\right) = \sum_{i=1}^n Z_i,$$

où les $Z_i, i = 1, \dots, n$, sont des variables aléatoires iid de même loi que Z . De la majoration $|Z| \leq 2Mp$, où M étant un majorant de $|\phi|$, nous avons $E(|Z|^k) \leq 2^k M^k E(p^k)$. D'autre part, $p = (p - \|\mu_1\|) + \|\mu_1\|$ d'où $E(p^k) = \sum_{j=0}^k C_k^j E(p - \|\mu_1\|)^j \|\mu_1\|^{k-j}$ et

$$\exists \Gamma > 0, \exists \Lambda > 0, \forall n \geq 2, \quad |E(p - \|\mu_1\|)^n| \leq \frac{n!}{2} \Gamma^{n-2} \Lambda^2.$$

Donc $\exists H > 0, E(p^k) \leq 2H^k k!$, par exemple pour $H = \max(\Gamma, \|\mu_1\|, \Lambda)$. Par suite pour tout $k \geq 2$, $E(|Z|^k) \leq \frac{k!}{2} (2MH)^{k-2} (4MH)^2$ et ainsi un lemme de Collomb [2] permet de conclure. \square

Suite de la preuve du théorème. – (a) En utilisant la majoration :

$$|\widehat{f}_n(x) - f(x)| \leq |E(\widehat{f}_n(x)) - f(x)| + |E(\widehat{f}_n(x)) - \widehat{f}_n(x)|$$

nous avons grâce à une version du lemme de Bochner (Lemme II.2 de [6])

$$\sup_{x \in \Delta(\varepsilon)} |E(\widehat{f}_n(x)) - f(x)| \rightarrow 0.$$

Appliquons le lemme pour obtenir

$$P\left(|E(\widehat{f}_n(x)) - \widehat{f}_n(x)| > \frac{\varepsilon}{2}\right) \leq 2 \exp\left(-\frac{\varepsilon^2 nh^2}{16\Lambda}\right).$$

En utilisant la décomposition de $\Delta(\varepsilon)$ comme dans la preuve des théorèmes 1 et 2 de [6], on obtient

$$P\left(\sup_{x \in \Delta(\varepsilon)} |E(\widehat{f}_n(x)) - \widehat{f}_n(x)| > \frac{\varepsilon}{2}\right) \leq \frac{2n^\alpha}{h} \exp\left(-\frac{\varepsilon^2 nh^2}{16\Lambda}\right).$$

Le second membre est le terme général d'une série convergente.

(b) Posons $r_n(x) = \frac{1}{nh} \sum_{i=1}^m Y_i^{(n)} K(x - X_i^{(n)}/h)$.

Nous avons la majoration suivante :

$$\sup_{x \in \Delta(\varepsilon)} |\widehat{\Psi}_n(x) - \Psi(x)| \leq \frac{1}{\inf_{x \in \Delta(\varepsilon)} \widehat{f}_n(x)} \left\{ \sup_{x \in \Delta(\varepsilon)} |\Psi(x)| \sup_{x \in \Delta(\varepsilon)} |\widehat{f}_n(x) - f(x)| + \sup_{x \in \Delta(\varepsilon)} |r_n(x) - r(x)| \right\}.$$

Le Lemme 2 de [7] donne $\lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |r_n(x) - E(r_n(x))| = 0$ p.co., et en utilisant le lemme de Bochner, on montre (comme dans [8], Lemme II.2, Chapitre IV) que : $\lim_{n \rightarrow \infty} \sup_{x \in \Delta(\varepsilon)} |r(x) - E(r_n(x))| = 0$. Ces deux résultats assurent la convergence uniforme (p.co.) de r_n vers r . La preuve se termine comme au Théorème 3 de [7].

3.2. Sur la vitesse de convergence

THÉORÈME 2. –

(a) Si f est différentiable de dérivée bornée sur Δ et λ vérifie l'hypothèse (i'), alors pour tout $\theta > \frac{1}{2}$, la relation : $h^{-2} = o\left(\frac{n}{\log(n)}\right)$, implique pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \chi_n \sup_{x \in \Delta(\varepsilon)} |\widehat{f}_n(x) - f(x)| = 0 \quad \text{p.co.}$$

(b) Si de plus, r est continue sur Δ et si il existe un ε positif tel que

$$\inf_{x \in \Delta(\varepsilon)} f(x) > 0 \quad \text{et} \quad \sup_{x \in \Delta(\varepsilon)} |\Psi(x)| < \infty,$$

alors pour tout $\theta > \frac{1}{2}$, la relation : $h^{-2} = o\left(\frac{n}{\log(n)}\right)$, implique pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \chi_n \sup_{x \in \Delta(\varepsilon)} |\widehat{\Psi}_n(x) - \Psi(x)| = 0 \quad \text{p.co.}$$

3.2.1. *Preuve*

(a) Pour tout x et x' de Δ , $|f(x) - f(x')| < D|x - x'|$ et donc nous avons uniformément en $x \in \Delta$, $|E(\widehat{f}_n(x)) - f(x)| < D(|a| + |b|)h$.

Pour n grand, nous avons d'une part $D(|a| + |b|)h < D\varepsilon/2\chi_n$ donc $\chi_n \sup_{x \in \Delta(\varepsilon)} |E(\widehat{f}_n(x)) - f(x)| < D\varepsilon/2$; et d'autre part, nous avons déjà obtenu que :

$$P\left(\chi_n |E(\widehat{f}_n(x)) - \widehat{f}_n(x)| > \frac{D\varepsilon}{2}\right) \leq 2 \exp\left(-\frac{\varepsilon^2 D^2 (\log(n))^{2\theta}}{16\Lambda}\right).$$

Après la décomposition de Δ comme au théorème précédent, on obtient :

$$P\left(\chi_n \sup_{x \in \Delta(\varepsilon)} |E(\widehat{f}_n(x)) - \widehat{f}_n(x)| > \frac{D\varepsilon}{2}\right) \leq \frac{2n^\alpha}{h} \exp\left(-\frac{\varepsilon^2 D^2 (\log(n))^{2\theta}}{16\Lambda}\right).$$

Le second membre est le terme général d'une série convergente.

(b) On utilise la même démarche que pour la partie a) en substituant, respectivement, f et \widehat{f}_n par r et r_n . La partie (b) du Théorème 1 nous permet ensuite de conclure.

Références bibliographiques

- [1] D. Bosq, J.P. Lecoutre, Théorie de l'estimation fonctionnelle, Economica, Paris, 1987.
- [2] G. Collomb, Conditions nécessaires et suffisantes de convergence uniforme d'un estimateur de la régression, estimation des dérivées de la régression, C. R. Acad. Sci. Paris 228 (1979) 161–163.
- [3] M. Curioni, Estimation de la densité de la densité des couleurs d'un processus de Poisson non homogène, Thèse de doctorat de troisième cycle, Université Paris VI, France, 1977.
- [4] G. Dia, Étude d'un estimateur de la fonction de régression pour un processus ponctuel valeurs dans $R_+^s \times R$ ($s \geq 1$), SERDICA Bulg. Math. Publ. 13 (1987) 383–395.
- [5] G. Dia, Non-parametric estimation of the density of a point process, Statist. Probab. Lett. 10 (1990) 397–405.
- [6] G. Dia, A. Diakhaby, Sur l'estimation de la densité et de la courbe de régression d'un processus ponctuel, Rapport technique 151, LSTA, Université Paris VI, France, 1991.
- [7] G. Dia, A. Diakhaby, Estimation non paramétrique de la densité et de la régression pour un processus ponctuel, C. R. Acad. Sci. Paris 321 (12) (1995) 1627–1630.
- [8] A. Diakhaby, Estimation de la densité et de la régression dans les processus ponctuels chromatiques, Thèse de doctorat de troisième cycle, Université C.A.D. de Dakar, Sénégal, 1992.
- [9] M. Lopez, Problème d'estimation dans les processus ponctuels chromatiques aléatoires, Thèse de doctorat de troisième cycle, Université Paris VI, France, 1982.
- [10] J.M. Marion, Estimation de la densité de la répartition locale moyenne des couleurs d'un processus ponctuel chromatique, Thèse de doctorat de troisième cycle, Université Paris VI, France, 1985.