Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Acad. Sci. Paris, Ser. I 341 (2005) 59–62
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Statistique/Probabilités

Vitesses optimale et suroptimale des polygones de fréquences pour les processus à temps continu

François-Xavier Lejeune

LSTA, université Paris 6, 175, rue du Chevaleret, 75013 Paris, France

Reçu le 30 janvier 2005 ; accepté après révision le 13 mai 2005

Disponible sur Internet le 20 juin 2005

Présenté par Paul Deheuvels

Résumé

Cette Note porte sur les vitesses de convergence d'un estimateur non-paramétrique de la densité d'un processus à temps continu. Plus précisément, sous certaines hypothèses de régularité et d'indépendance asymptotique, l'erreur quadratique intégrée du polygone de fréquences converge vers zéro à la vitesse optimale $T^{-4/5}$ du cas i.i.d. Avec une condition locale plus faible que celle de Castellana–Leadbetter [Stochastic Process. Appl. 21 (1986) 179–193], la vitesse « suroptimale » T^{-1} est obtenue. ***Pour citer cet article : F.-X. Lejeune, C. R. Acad. Sci. Paris, Ser. I 341 (2005).***

Abstract

Optimal and superoptimal rates of frequency polygons for continuous-time processes. This Note deals with density estimation in continuous-time. Then under mild regularity and asymptotic independence conditions, the mean integrated square error achieves the same optimal rate $T^{-4/5}$ of convergence to zero as in the i.i.d. case. Under a local assumption weaker than Castellana–Leadbetter's [Stochastic Process. Appl. 21 (1986) 179–193], we obtain the parametric rate T^{-1}. ***To cite this article: F.-X. Lejeune, C. R. Acad. Sci. Paris, Ser. I 341 (2005).***

1. Introduction

Soit $\{X_t, t \in \mathbb{R}\}$ un processus stochastique réel, mesurable et défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, chaque X_t ayant la même loi μ de densité f relativement à la mesure de Lebesgue. L'estimation non-paramétrique de f lorsque la trajectoire est observée sur $[0, T]$, avec $T \to \infty$, est l'objet d'une littérature abondante, voir par exemple [1,5] et les nombreuses références incluses. En particulier, le polygone de fréquences, qui s'obtient en reliant tous les sommets d'un histogramme en leur milieu, n'a pas encore été étudié en temps continu. Dans cette Note, nous

Adresse e-mail : fxlejeun@ccr.jussieu.fr (F.-X. Lejeune).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2005.05.003

établissons des vitesses de convergence optimale et suroptimale comparables à celles de l'estimateur à noyau, et meilleures que l'histogramme en terme d'erreur quadratique intégrée (EQI). En ce sens, et du fait de sa simplicité de mise en œuvre, le polygone de fréquences est une alternative satisfaisante de l'histogramme fréquemment utilisé en pratique.

2. Construction de l'estimateur et hypothèses

On obtient le polygone de fréquences en réalisant dans un premier temps un histogramme sur une partition équilibrée Π_T de $\mathbb{R}$ en intervalles de longueur h_T tel que $T \to \infty, h_T \to 0, Th_T \to \infty : \Pi_T = \{\pi_{Tj},\ j \in J_T\},\ J_T \subset \mathbb{Z}$ et $\pi_{Tj} = [b_j, b_{j+1}[= [c_j - \frac{h_T}{2}, c_j + \frac{h_T}{2}[\ ,\ j \in J_T$ où b_j et b_{j+1} désignent les bords de π_{Tj} et $c_j = (b_j + b_{j+1})/2$. Sur chaque intervalle π_{Tj}, l'histogramme vaut : $\hat{f}_H(x) = \hat{f}_j = \frac{1}{Th_T} \int_0^T \mathbf{1}_{\pi_{Tj}}(X_t)\,\mathrm{d}t,\ x \in \pi_{Tj}$ où $\mathbf{1}_A$ est la fonction indicatrice de l'ensemble A, et s'interprète comme le temps moyen que passe le processus dans l'intervalle π_{Tj} sur la période $[0, T]$ divisé par la longueur h_T de π_{Tj}. On construit ensuite le polygone de fréquences en traçant une droite segmentée reliant les points centraux des sommets de chaque rectangle de l'histogramme. L'estimateur ainsi obtenu est continu et dérivable en dehors des points $\{c_j,\ j \in J_T\}$.

Définition 2.1. On définit l'estimateur $\hat{f}_{FP}$ de f par :

$$\hat{f}_{FP}(x) = \sum_{j \in J_T} \left\{ \left(\frac{x - c_j}{h_T} \right) \hat{f}_{j+1} + \left(\frac{c_{j+1} - x}{h_T} \right) \hat{f}_j \right\} \mathbf{1}_{[c_j, c_{j+1}[}(x), \quad x \in \mathbb{R}.$$

Pour abréger les écritures, $\mathcal{C}^k(\mathbb{R})$ désignera l'ensemble des fonctions k fois continûment dérivables sur $\mathbb{R}$, $f^{(k)}$ la dérivée d'ordre k de f et $f_{(X_s,X_t)}$ la densité du couple (X_s, X_t). On notera : $g_{s,t} = f_{(X_s,X_t)} - f \otimes f$ et $R_\alpha(f) = \int_{\mathbb{R}} [f^{(\alpha)}(x)]^2\,\mathrm{d}x$. Les hypothèses portant sur le processus sont les suivantes :

- **Hypothèses** H_0 :
  - (i) f est continue en x et $\|f\|_\infty = M_1 < \infty$;
  - (ii) $f_{(X_s,X_t)} = f_{(X_0,X_{|t-s|})}$ pour $s \neq t$;
  - (iii) $f \in \mathcal{C}^2(\mathbb{R})$, $\inf_{x \in \mathbb{R}} f(x) > 0$, $f^{(2)} \in L^1(\mathbb{R})$, $f^{(k)} \in L^2(\mathbb{R})$ pour $k = 0, 2$ et $|f^{(2)}(x) - f^{(2)}(y)| \leqslant L_0|x - y|^\nu$ pour $L_0 > 0$, $\nu \in]0, 1]$ et $(x, y) \in \mathbb{R}^2$.
- **Hypothèses** $H_1(\Gamma, \varrho)$: il existe un borélien de $\mathbb{R}^2$: $\Gamma = \{(s, t) \in \mathbb{R}^2,\ |t - s| \leqslant u_0,\ u_0 > 0\}$ tel que
  - (i) $\limsup_{T \to \infty} \frac{1}{T} \int_{[0,T]^2 \cap \Gamma} \mathrm{d}s\,\mathrm{d}t = \ell_\Gamma < \infty$;
  - (ii) $g_{s,t}$ existe pour $(s, t) \notin \Gamma$;
  - (iii) $\sup_{(y,z) \in \mathbb{R}^2} f_{X_u/X_0}(z/y) \leqslant M_2 < \infty$ pour $u \geqslant u_0$;
  - (iv) $\{X_t,\ t \in \mathbb{R}\}$ est un processus arithmétiquement fortement mélangeant de cœfficient α (voir [4]) : $\alpha(u) \leqslant a_0 u^{-\varrho}$ pour $u \geqslant u_0$, $a_0 > 0$ et $\varrho > 2$.
- **Hypothèses** H_2 :
  - (i) $g_{s,t}$ existe pour $s \neq t$;
  - (ii) $\forall y \in \mathbb{R}, \sup_{z \in \mathbb{R}} \int_0^{+\infty} |g_{0,u}(y, z)|\,\mathrm{d}u \leqslant \varphi(y)$ où φ est une fonction définie sur $\mathbb{R}$, positive, continue, intégrable et bornée.

Commentaires. $H_1(\Gamma, \varrho)$(i) et H_2(ii) sont des conditions spécifiques au temps continu. $H_1(\Gamma, \varrho)$(i) est utilisée de façon analogue dans le contexte de l'estimateur à noyau (voir [1]). Ne portant pas directement sur le processus, elle est peu contraignante. Plus restrictive, H_2(ii) est une condition liée à la nature des trajectoires et vérifiée notamment par une large classe de processus de diffusion (voir [6]). Cette condition affaiblit celle de Castellana et Leadbetter [3] : $\int_0^{+\infty} \|g_{0,u}\|_\infty\,\mathrm{d}u < \infty$. Les autres conditions sont classiques en estimation non-paramétrique. Enfin, nous utiliserons alternativement $H_1(\Gamma, \varrho)$ et H_2 avec H_0 pour parvenir à nos résultats.

3. Polygone de fréquences : critère EQI

On suppose que pour tout $x \in \mathbb{R}$, il existe un indice $j(x,T)$ tel que $x \in \pi_{Tj(x,T)}$ ($:= \pi_{Tj}$). Les hypothèses énoncées ci-dessus conduisent à une majoration de la variance de l'histogramme utile pour la suite :

Théorème 3.1. (i) *Si les conditions* H_0(i)–(ii) *et* $H_1(\Gamma,\varrho)$(ii)–(iv) *sont vérifiées, alors pour tout* $1 < p \leqslant \varrho - 1$

$$Th_T \cdot \operatorname{Var}\{\hat{f}_j\} \leqslant f(\xi_j)\big(1 - h_T f(\xi_j)\big) \cdot \frac{1}{T} \int\limits_{[0,T]^2 \cap \Gamma} \mathrm{d}s\,\mathrm{d}t + 2Mf(\xi_j) \cdot h_T^{\varepsilon} + \frac{4p^2(2a_0)^{1/p}}{\varrho - p} f(\xi_j)^{1-1/p} \cdot h_T^{\frac{1}{p}\{(1-\varepsilon)(\varrho-p)-1\}},$$

avec $M = \max(M_1, M_2)$ *et* $0 \leqslant \varepsilon \leqslant 1 - \frac{1}{\varrho - p}$, $x \in \pi_{Tj}$, $\xi_j \in \pi_{Tj}$;

(ii) *Si les conditions* H_0(ii) *et* H_2 *sont vérifiées, alors* $T \cdot \operatorname{Var}\{\hat{f}_j\} \leqslant 2\varphi(\eta_j)$, $x \in \pi_{Tj}$, $\eta_j \in \pi_{Tj}$.

Eléments de preuve. (i) On obtient le premier terme de la majoration en utilisant l'inégalité de Cauchy–Schwarz. Le deuxième terme vient de $H_1(\Gamma,\varrho)$(ii)(iii). Enfin, le troisième terme découle de l'inégalité de Davydov (voir [1, p. 21]) et de l'hypothèse de mélangeance $H_1(\Gamma,\varrho)$(iv).

(ii) La seconde majoration se déduit des hypothèses H_2. □

Remarque 1. On retrouve sous H_0 et $H_1(\Gamma,\varrho)$ un résultat semblable à celui obtenu par Rio ([7], Chapitre 1), en temps discret, dans le cas d'un processus strictement stationnaire et fortement mélangeant. Si on remplace H_2(ii) par la condition de Castellana et Leadbetter avec $g_{0,u}$ continue en (x,x), on obtient le comportement limite exact : $\lim_{T\to\infty} T \cdot \operatorname{Var}\{\hat{f}_j\} = 2\int_0^{+\infty} g_{0,u}\,\mathrm{d}u$.

Sous H_0, il est connu que l'on peut améliorer la vitesse de convergence du biais de l'histogramme en introduisant l'estimateur par polygone de fréquences. L'erreur quadratique intégrée du polygone de fréquences a déjà été étudiée dans les cas i.i.d. [8] et mélangeant [2]. Les résultats suivants apportent une majoration et des vitesses de convergence pour l'EQI en temps continu. On écrira alors l'EQI de $\hat{f}_{FP}$ comme la somme du biais carré intégré (BCI) et de la variance intégrée (VI), tels que $BCI\{\hat{f}_{FP}\} = \int_{\mathbb{R}} (E\{\hat{f}_{FP}(x)\} - f(x))^2\,\mathrm{d}x$ et $VI\{\hat{f}_{FP}\} = \int_{\mathbb{R}} \operatorname{Var}\{\hat{f}_{FP}(x)\}\,\mathrm{d}x$.

Lemme 3.2. *Sous* H_0(iii), *on obtient* $BCI\{\hat{f}_{FP}\} = \frac{49}{2880} R_2(f) h_T^4 + \mathrm{o}(h_T^4)$.

Preuve. Analogue à [8] en utilisant H_0(iii). □

Théorème 3.3. (i) *Si les conditions* H_0(i)–(ii) *et* $H_1(\Gamma,\varrho)$(ii)–(iv) *sont vérifiées et si* $f^{1-1/p} \in L^1(\mathbb{R})$ *pour un* p *tel que* $1 < p \leqslant \varrho - 1$, *alors pour* $F_p = \int_{\mathbb{R}} f^{1-1/p}(x)\,\mathrm{d}x$ *et* $0 \leqslant \varepsilon \leqslant 1 - \frac{1}{\varrho - p}$:

$$Th_T \cdot VI\{\hat{f}_{FP}\} \leqslant \left\{ \frac{1}{T} \int\limits_{[0,T]^2 \cap \Gamma} \mathrm{d}s\,\mathrm{d}t + 2M \cdot h_T^{\varepsilon} + \frac{4p^2(2a_0)^{1/p}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(1-\varepsilon)(\varrho-p)-1\}} \right\} (1 + \mathrm{o}(1)) ;$$

(ii) *Si les conditions* H_0(ii) *et* H_2 *sont vérifiées, alors* $T \cdot VI\{\hat{f}_{FP}\} \leqslant 2\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x + \mathrm{o}(1)$.

Preuve partielle. La preuve est basée sur la décomposition suivante pour $x \in [c_j, c_{j+1}[$:

$$\operatorname{Var}\big\{\hat{f}_{FP}(x)\big\} = \left(\frac{x - c_j}{h_T}\right)^2 \operatorname{Var}\{\hat{f}_{j+1}\} + \left(\frac{c_{j+1} - x}{h_T}\right)^2 \operatorname{Var}\{\hat{f}_j\} + 2\left(\frac{x - c_j}{h_T}\right)\left(\frac{c_{j+1} - x}{h_T}\right) \operatorname{Cov}\{\hat{f}_j, \hat{f}_{j+1}\}.$$

□

4. Vitesses optimale et suroptimale

Finalement, pour des choix ad hoc du paramètre h_T, on obtient les vitesses optimale et suroptimale du polygone de fréquences :

Théorème 4.1. (i) *Si les conditions* H_0 *et* $H_1(\Gamma, \varrho)$ *sont vérifiées et si* $f^{1-1/p} \in L^1(\mathbb{R})$ *pour* $1 < p \leqslant \varrho - 1$, *le choix* $h_T = k_T \cdot T^{-1/5}$ *tel que* $k_T \to k, 0 < k < \infty$ *entraîne*

$$\limsup_{T\to\infty} T^{4/5} \cdot EQI\{\hat{f}_{FP}\} \leqslant \begin{cases} \frac{49}{2880}k^4 R_2(f) + \frac{1}{k}\big(\ell_\Gamma + 2M + 4p^2(2a_0)^{1/p}F_p\big) & \text{si } p = \varrho - 1, \\ \frac{49}{2880}k^4 R_2(f) + \frac{1}{k}\ell_\Gamma & \text{si } p < \varrho - 1; \end{cases}$$

(ii) *Si les conditions* H_0(ii)(iii) *et* H_2 *sont vérifiées et si* $h_T = \mathrm{o}(T^{-1/4})$, *alors*

$$\limsup_{T\to\infty} T \cdot EQI\{\hat{f}_{FP}\} \leqslant 2\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x.$$

Remarque 2. (1) Dans le Théorème 4.1(i), on retrouve la vitesse du cas i.i.d. Cette vitesse est optimale au sens où elle est atteinte par un processus construit de manière analogue à [1, p. 96], et vérifiant H_0 et $H_1(\Gamma, \varrho)$. Notons de plus qu'elle est minimax (voir [1], Chapitre 4).

(2) Pour la partie (ii), on obtient la vitesse suroptimale propre au temps continu. En effet, sous les conditions H_0 et H_2, la variance asymptotique ne dépend pas de h_T (contrairement au cas discret).

Exemple 1. On considère deux modèles de processus de diffusion homogène : $\{\xi_t^{(1)},\ t \in \mathbb{R}\}$, le processus d'Ornstein–Uhlenbeck vérifiant l'équation de Langevin : $\mathrm{d}X_t = -(aX_t + b)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, X_0, t \geqslant 0$ où a, b et σ désignent des réels positifs, de loi de probabilité invariante $\mathcal{N}(\frac{b}{a}, \frac{\sigma^2}{2a})$ et $\{\xi_t^{(2)},\ t \in \mathbb{R}\}$, solution de l'équation différentielle stochastique : $\mathrm{d}X_t = -\theta\,\mathrm{sgn}(X_t)\,\mathrm{d}t + \mathrm{d}W_t,\ \theta > 0,\ X_0, t \geqslant 0$ de densité invariante $f(x) = \theta\,\mathrm{e}^{-2\theta|x|}$. Ces deux modèles vérifient notamment l'hypothèse H_2(ii) d'après [6] et appartiennent à la classe des processus qui satisfont aux hypothèses du Théorème 4.1.

Remerciements

J'exprime ma gratitude à Delphine Blanke et au Professeur Denis Bosq pour l'intérêt qu'ils portent à mes travaux. Je tiens également à remercier les rapporteurs pour les améliorations suggérées.

Références

[1] D. Bosq, Nonparametric Statistics for Stochastic Processes. Estimation and Prediction, second ed., Lecture Notes in Statist., vol. 110, Springer-Verlag, New York, 1998.
[2] M. Carbon, B. Garel, L.T. Tran, Frequency polygons for weakly dependent processes, Statist. Probab. Lett. 33 (1997) 1–13.
[3] J.V. Castellana, M.R. Leadbetter, On smoothed probability density estimation for stationary processes, Stochastic Process. Appl. 21 (1986) 179–193.
[4] P. Doukhan, Mixing: Properties and Examples, Lecture Notes in Statist., vol. 85, Springer-Verlag, New York, 1994.
[5] Y.A. Kutoyants, Statistical Inference for Ergodic Diffusion Processes, Springer Ser. Statist., Springer, 2003.
[6] F. Leblanc, Density estimation for a class of continuous time processes, Math. Methods Statist. 6 (1997) 171–199.
[7] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants, Collect. Math. Appl., vol. 31, Springer-Verlag, Berlin, 2000.
[8] D.W. Scott, Frequency polygons: theory and application, J. Amer. Statist. Assoc. 80 (1985) 348–354.