



HAL
open science

Propriétés des estimateurs par histogrammes et polygones de fréquences de la densité marginale d'un processus à temps continu

François-Xavier Lejeune

► **To cite this version:**

François-Xavier Lejeune. Propriétés des estimateurs par histogrammes et polygones de fréquences de la densité marginale d'un processus à temps continu. *Annales de l'ISUP*, 2006, L (1-2), pp.47-77. hal-03637561

HAL Id: hal-03637561

<https://hal.science/hal-03637561>

Submitted on 11 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
L, fasc. 1-2, 2006, 47 à 77

PROPRIÉTÉS DES ESTIMATEURS PAR HISTOGRAMMES ET POLYONES DE FRÉQUENCES DE LA DENSITÉ MARGINALE D'UN PROCESSUS À TEMPS CONTINU

FRANÇOIS-XAVIER LEJEUNE

L.S.T.A, Université Paris 6, 175, rue du Chevaleret, 75013 Paris, France

Résumé Dans cet article, nous étudions les propriétés de graphiques de fréquences pour estimer la fonction de densité d'un processus réel à temps continu. Plus précisément, sous certaines hypothèses de régularité et d'indépendance asymptotique, l'erreur en moyenne quadratique intégrée de l'histogramme et du polygone de fréquences converge vers zéro avec les mêmes vitesses optimales du cas i.i.d. Avec une condition locale plus faible que celle proposée par Castellana et Leadbetter [9], la vitesse suroptimale en T^{-1} est obtenue. Enfin, nous déduisons d'un résultat de Bosq, Merlevède et Peligrad [5] la normalité asymptotique de ces estimateurs.

Abstract In this article, we study the properties of graphical representations to estimate the density function of a real-valued continuous time process. Thus under mild regularity and asymptotic independence conditions, the mean integrated square error of the histogram and the frequency polygon achieves the same optimal rates of convergence to zero as in the i.i.d. case. With a local assumption weaker than Castellana and Leadbetter's [9], the superoptimal rate T^{-1} is obtained. Finally, we derive the asymptotic normality of both estimators from a result of Bosq, Merlevède and Peligrad [5].

1. Introduction. En temps continu, l'estimation non paramétrique de la densité marginale d'un processus à partir d'une trajectoire observée sur une période $[0, T]$, avec $T \rightarrow \infty$, est l'objet d'une littérature dense (cf. par exemple Bosq [3], Kutoyants [15] et les nombreuses références incluses). Dans cet article, nous estimons cette densité par un histogramme puis par un polygone de fréquences construit en reliant les points centraux des sommets du précédent histogramme. Lorsque les observations sont indépendantes et identiquement distribuées (i.i.d.), les principales propriétés de ces estimateurs sont détaillées dans Scott [23] (chapitres 3 et 4). Dans le cas fortement mélangeant, l'erreur en moyenne quadratique intégrée et la convergence uniforme du polygone sont traitées par Carbon, Garel et Tran [8]. Nous nous

Mots-clefs: Estimation non paramétrique, Histogramme, Polygone de fréquences, α -mélangeance, Erreur en moyenne quadratique intégrée, Normalité asymptotique

AMS Classification: 62G07, 62M09, 62G20

proposons ici d'étendre certains de ces résultats au temps continu. Une nouvelle étude des histogrammes peut paraître quelque peu anachronique étant donné l'existence d'estimateurs plus récents et plus sophistiqués que sont les fonctions orthogonales, les splines ou encore les ondelettes. Néanmoins, du fait de leur simplicité de mise en œuvre, ces graphiques de fréquences sont non seulement toujours populaires en pratique mais ils peuvent aussi s'avérer performants d'un point de vue théorique. Notamment, pour l'erreur en moyenne quadratique intégrée, Scott [22] a montré que la vitesse de convergence de l'estimateur à noyau en $T^{-4/5}$ était également atteinte par le polygone de fréquences. Dans ce travail, la Section 2 porte sur l'erreur en moyenne quadratique intégrée. Sous certaines hypothèses de régularité et de mélange fort, cette erreur converge vers zéro avec les mêmes vitesses optimales du cas i.i.d. Ces vitesses sont de plus minimax. Avec une condition locale de type Castellana-Leadbetter [9], nous améliorons ensuite ces vitesses et l'erreur s'écrit comme un $\mathcal{O}(T^{-1})$. Notons que les vitesses optimale et suroptimale obtenues pour le polygone de fréquences restent comparables à celles de l'estimateur à noyau (cf. Lejeune [18]). Dans la Section 3, une application directe d'un théorème central-limite proposé par Bosq, Merlevède et Peligrad [5] pour les tableaux triangulaires établit la normalité asymptotique de ces estimateurs et permet la construction de régions de confiance. Enfin, les preuves des principaux résultats sont apportées dans la Section 5.

Soit $\mathcal{X} = \{X_t, t \in \mathbb{R}\}$ un processus à temps continu défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, mesurable et à valeurs dans \mathbb{R}^d . On suppose que chaque X_t suit la même loi de probabilité μ de densité f relativement à la mesure de Lebesgue, la fonction f étant inconnue.

Mesure de dépendance. Nous donnons ici quelques rappels sur le coefficient α de mélange fort introduit par Rosenblatt [20]. On pourra se référer à Doukhan [11] pour une lecture plus approfondie du sujet.

DÉFINITION 1.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et soient \mathcal{B} et \mathcal{C} deux sous-tribus de \mathcal{A} . Afin de mesurer la dépendance de \mathcal{B} et \mathcal{C} , on définit le coefficient α de mélange fort par

$$\alpha(\mathcal{B}, \mathcal{C}) = \sup_{B \in \mathcal{B}, C \in \mathcal{C}} |\mathbb{P}(B \cap C) - \mathbb{P}(B)\mathbb{P}(C)|.$$

DÉFINITION 1.2. Le processus \mathcal{X} est dit α -mélangeant si

$$\alpha_{\mathcal{X}}(u) = \sup_{t \in \mathbb{R}} \alpha(\mathcal{F}_{-\infty}^t, \mathcal{F}_{t+u}^{+\infty}) \downarrow_{u \rightarrow \infty} 0,$$

où \mathcal{F}_m^t désigne la σ -algèbre des événements engendrés par $\{X_s, m \leq s \leq t\}$ pour $-\infty \leq m \leq t \leq +\infty$.

Notons $\sigma(X)$ la σ -algèbre des événements engendrés par la variable X et $\|X\|_q = \{\mathbb{E}(X^q)\}^{1/q}$ pour $1 \leq q \leq \infty$. On rappelle pour la suite une forme due à Rio de l'inégalité de Davydov (cf. Bosq [3], p. 21) :

LEMME 1.3. Soient $X \in L^q(\mathbb{P})$ et $Y \in L^r(\mathbb{P})$, si $q > 1$, $r > 1$ et $\frac{1}{r} + \frac{1}{q} = 1 - \frac{1}{p}$, alors

$$|Cov(X, Y)| \leq 2p(2\alpha(\sigma(X), \sigma(Y)))^{\frac{1}{p}} \|X\|_q \|Y\|_r.$$

Critère de convergence. Pour évaluer la précision des estimateurs non paramétriques en un point x , l'erreur en moyenne quadratique est un critère très répandu dans la littérature. En intégrant sur \mathbb{R}^d , nous obtenons une mesure globale de cette erreur quadratique.

DÉFINITION 1.4. Soit \hat{g}_T un estimateur d'une fonction réelle g construit à partir d'une trajectoire de \mathcal{X} observée sur $[0, T]$. L'erreur en moyenne quadratique intégrée (EQI) de \hat{g}_T est définie par

$$EQI(\hat{g}_T) = \mathbb{E} \int_{\mathbb{R}^d} (\hat{g}_T(x) - g(x))^2 dx = \int_{\mathbb{R}^d} \mathbb{E}(\hat{g}_T(x) - g(x))^2 dx.$$

DÉFINITION 1.5. L'estimateur \hat{g}_T est dit convergent en moyenne quadratique intégrée si

$$\lim_{T \rightarrow \infty} EQI(\hat{g}_T) = 0.$$

Notations. Pour abrégier les écritures, on notera $C^k(\mathbb{R}^d)$ l'ensemble des fonctions k fois continûment dérivables sur \mathbb{R}^d et $C_b^k(\mathbb{R}^d)$ le sous-ensemble si toutes les dérivées partielles sont bornées. $L^k(\mathbb{R}^d)$ désignera l'ensemble des fonctions de puissance k intégrable sur \mathbb{R}^d et $f_{(X_s, X_t)}$ la densité du couple (X_s, X_t) . On écrira $g_{s,t} := f_{(X_s, X_t)} - f \otimes f$ et $G(x, y) := \int_0^{+\infty} g_{0,u}(x, y) du$.

2. Estimation et erreur en moyenne quadratique intégrée.

2.1. L'histogramme.

L'histogramme est probablement le plus ancien et le plus connu des estimateurs non paramétriques. L'idée de grouper ainsi les données remonte déjà aux travaux de John Graunt, un recensement statistique de la population londonienne publié en 1662 sous le titre "Natural and Political Observations on the Bills of Mortality". Simple à mettre en œuvre, il est encore fréquemment utilisé en pratique. Dans le cas i.i.d., les principales propriétés de cet estimateur sont développées dans Bosq et Lecoutre [4] (chapitre 6, partie A) et

Scott [23] (chapitre 3). L'étude qui suit propose une extension de certains de ces résultats au cadre continu et fortement mélangeant sous des hypothèses faibles de stationnarité.

Construction de l'estimateur. On obtient un histogramme en réalisant dans un premier temps une partition équilibrée Π_T de \mathbb{R}^d en hypercubes de mesure h_T^d tel que $T \rightarrow \infty$, $h_T \rightarrow 0^+$, $Th_T^d \rightarrow \infty$:

$$\Pi_T = \{\pi_{Tj}, j \in J_T\}, \quad J_T \subset \mathbb{Z},$$

et

$$\pi_{Tj} = \prod_{k=1}^d [b_j^{(k)}, b_{j+1}^{(k)}] = \prod_{k=1}^d \left[c_j^{(k)} - \frac{h_T}{2}, c_j^{(k)} + \frac{h_T}{2} \right], \quad j \in J_T,$$

où $b_j = (b_j^{(1)}, \dots, b_j^{(d)})' \in \mathbb{R}^d$, $b_{j+1}^{(k)} - b_j^{(k)} = h_T$ et $c_j^{(k)} = \frac{b_j^{(k)} + b_{j+1}^{(k)}}{2}$.

DÉFINITION 2.1. Sur chaque hypercube π_{Tj} , l'histogramme est défini par

$$\hat{f}_H(x) = \hat{f}_H(x, \mathcal{X}, T) = \hat{f}_j = \frac{1}{Th_T^d} \int_0^T \mathbf{1}_{\pi_{Tj}}(X_t) dt, \quad x \in \pi_{Tj},$$

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'hypercube π_{Tj} sur la période $[0, T]$ normalisé par la mesure h_T^d de π_{Tj} .

On rappelle la définition de l'intégrale généralisée de Riemann (IGR) qui sera utile pour donner le biais de l'histogramme (cf. Bertrand-Retali [1]) :

DÉFINITION 2.2. Soit g une fonction réelle définie sur \mathbb{R}^d . Si g_n^- et g_n^+ sont deux suites de fonctions Π_n -simples, Riemann-intégrables sur \mathbb{R}^d et telles que

$$g_n^- \leq g_{n+1}^- \leq g \leq g_{n+1}^+ \leq g_n^+,$$

et

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} g_n^-(x) dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} g_n^+(x) dx = I,$$

alors I est l'intégrale généralisée de Riemann de la fonction g . Il en résulte que g est intégrable sur \mathbb{R}^d et que $\int_{\mathbb{R}^d} g(x) dx = I$ et

$$\forall \varepsilon > 0, \quad g_\varepsilon(x) = \sup_{t \in \mathbb{R}^d} \{g(t); \|t - x\| \leq \varepsilon\}$$

est intégrable sur \mathbb{R}^d et $\int_{\mathbb{R}^d} g_\varepsilon(x) dx$ reste bornée quand $\varepsilon \rightarrow 0$.

Les principales hypothèses portant sur le processus sont les suivantes :

Hypothèses H_0 .

- (i) $f \in C_b^2(\mathbb{R}^d)$ et les dérivées partielles et leurs carrés admettent une IGR;
- (ii) f est continue en x et $\|f\|_\infty = M_1 < \infty$;
- (iii) $f(x_s, x_t) = f(x_0, x_{|t-s|})$ pour $s \neq t$.

Hypothèses H_1 . Il existe un borélien de \mathbb{R}^2 : $\Gamma = \{(s, t) \in \mathbb{R}^2, |t - s| \leq u_0, u_0 > 0\}$ tel que

- (i) $\limsup_{T \rightarrow \infty} \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt = \ell_\Gamma < \infty$;
- (ii) $g_{s,t}$ existe pour $(s, t) \notin \Gamma$;
- (iii) $\sup_{(y,z) \in \mathbb{R}^{2d}} f_{X_u/X_0}(z/y) \leq M_2 < \infty$ pour $u \geq u_0$;
- (iv) le processus \mathcal{X} est arithmétiquement fortement mélangeant de coefficient $\alpha_{\mathcal{X}} : \alpha_{\mathcal{X}}(u) \leq a_0 u^{-\varrho}$ pour $u \geq u_0, a_0 > 0$ et $\varrho > 2$.

Hypothèses H_2 .

- (i) $g_{s,t}$ existe pour $s \neq t$;
- (ii) $\int_0^{+\infty} \|g_{0,u}\|_\infty du < \infty$ et $g_{0,u}(\cdot, \cdot)$ est continue en $(x, x) \in \mathbb{R}^{2d}$ pour tout $u > 0$;
- (ii)' $\forall y \in \mathbb{R}^d, \sup_{z \in \mathbb{R}^d} \int_0^{+\infty} |g_{0,u}(y, z)| du \leq \varphi(y)$ où φ est une fonction définie sur \mathbb{R}^d , positive, continue, intégrable et bornée.

Commentaires. Les différentes hypothèses sont classiques en estimation non paramétrique. $H_1(i)$, $H_2(ii)$ et $H_2(ii)'$ sont des conditions locales spécifiques au temps continu. $H_1(i)$ est utilisée de façon analogue par Bosq [3] dans le contexte de l'estimateur à noyau. Ne portant pas directement sur le processus, elle est peu contraignante. Plus restrictive, la condition $H_2(ii)$ de Castellana-Leadbetter [9] est liée à la nature des trajectoires. Celle-ci peut-être affaiblie par la condition $H_2(ii)'$ vérifiée notamment par une large classe de processus de diffusion (cf. Leblanc [16], Veretennikov [26]). $H_0(iii)$ est une condition faible de stationnarité pour le processus et $H_1(iv)$ une condition d'indépendance asymptotique. Nous utiliserons alternativement H_1 et H_2 avec H_0 pour parvenir aux résultats suivants.

On supposera ensuite, pour tout $x \in \mathbb{R}^d$, l'existence d'un indice $j(x, T)$ tel que $x \in \pi_{Tj(x, T)} (=:\pi_{Tj})$.

Convergence dans L^2 . Les résultats suivants apportent des majorations de la variance de l'histogramme et des vitesses de convergence de l'erreur en

moyenne quadratique intégrée. On écrira alors l'*EQI* de \hat{f}_H comme la somme du biais au carré intégré (*BCI*) et de la variance intégrée (*VI*) tels que

$$BCI\{\hat{f}_H\} = \int_{\mathbb{R}^d} (E\{\hat{f}_H(x)\} - f(x))^2 dx \quad \text{et} \quad VI\{\hat{f}_H\} = \int_{\mathbb{R}^d} Var\{\hat{f}_H(x)\} dx.$$

Le lemme suivant dû à Lecoutre [17] nous donne une expression du *BCI* de l'histogramme multivarié :

LEMME 2.3. *Si la condition $H_0(i)$ est vérifiée, alors*

$$BCI\{\hat{f}_H\} = \frac{h_T^2}{12} \Psi_d^2 + o(h_T^2),$$

$$\text{avec } \Psi_d^2 := \int_{\mathbb{R}^d} \sum_{i=1}^d \left(\frac{\partial f}{\partial x_i}(x)\right)^2 dx.$$

Les hypothèses énoncées ci-dessus conduisent à des majorations de la variance de l'histogramme :

THÉORÈME 2.4. *Si les conditions $H_0(ii)(iii)$ et $H_1(ii) - (iv)$ sont vérifiées, alors pour tout $1 < p \leq \varrho - 1$*

$$\begin{aligned} Th_T^d \cdot Var\{\hat{f}_j\} &\leq f(\xi_j)(1 - h_T^d f(\xi_j)) \cdot \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + 2M f(\xi_j) \cdot h_T^\varepsilon \\ &\quad + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} f(\xi_j)^{1 - \frac{1}{p}} \cdot h_T^{\frac{1}{p}\{(d-\varepsilon)(\varrho-p)-d\}}, \end{aligned}$$

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

REMARQUE 1. En changeant $H_1(iv)$ pour la condition plus forte de mélange géométrique : $\alpha_{\mathcal{X}}(u) \leq a_0 \varrho^u$, $u > 0$, $a_0 > 0$, $0 < \varrho < 1$, la condition $1 < p \leq \varrho - 1$ devient $p > 1$.

THÉORÈME 2.5. 1. *Si les conditions $H_0(iii)$ et $H_2(i)(ii)$ sont vérifiées, alors*

$$T \cdot Var\{\hat{f}_j\} \xrightarrow{T \rightarrow \infty} 2G(x, x),$$

2. *Si les conditions $H_0(iii)$ et $H_2(i)(ii)'$ sont vérifiées, alors*

$$T \cdot Var\{\hat{f}_j\} \leq 2\varphi(\eta_j), \quad x \in \pi_{Tj}, \quad \eta_j \in \pi_{Tj}.$$

REMARQUE 2. Sous H_0 et H_1 , on retrouve un résultat semblable à celui obtenu, en particulier, en temps discret par Rio [19] (chapitre 1) dans le cas d'un processus strictement stationnaire et fortement mélangeant.

THÉORÈME 2.6. *Si les conditions $H_0(ii)(iii)$ et $H_1(ii) - (iv)$ sont vérifiées et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R}^d)$ avec $1 < p \leq \varrho - 1$, alors pour $F_p := \int_{\mathbb{R}^d} f^{1-\frac{1}{p}}(x) dx$ et $0 \leq \varepsilon \leq d(1 - \frac{1}{\varrho-p})$:*

$$Th_T^d \cdot VI\{\hat{f}_H\} \leq \left\{ \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(d-\varepsilon)(\varrho-p)-d\}} \right\} (1 + o(1)).$$

THÉORÈME 2.7. *Si les conditions $H_0(iii)$ et $H_2(i)(ii)'$ sont vérifiées, alors pour $\Phi := \int_{\mathbb{R}^d} \varphi(x) dx$:*

$$T \cdot VI\{\hat{f}_H\} \leq 2\Phi + o(1).$$

Par passage à limite et pour des choix ad hoc du paramètre h_T , on obtient les vitesses optimale et suroptimale de l'histogramme :

THÉORÈME 2.8. *Sous les conditions H_0 et H_1 et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R}^d)$ pour $1 < p \leq \varrho - 1$, le choix $h_T = c_T \cdot T^{-\frac{1}{d+2}}$ tel que $c_T \rightarrow c$, $0 < c < \infty$ entraîne*

$$\limsup_{T \rightarrow \infty} T^{\frac{2}{d+2}} \cdot EQI\{\hat{f}_H\} \leq \begin{cases} \frac{c^2}{12} \Psi_d^2 + \frac{1}{c^2} (\ell_\Gamma + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho-p} F_p), \\ \text{si } p = \varrho - 1; \\ \frac{c^2}{12} \Psi_d^2 + \frac{\ell_\Gamma}{c^2}, \text{ si } p < \varrho - 1. \end{cases}$$

REMARQUE 3. Sous H_0 et H_1 , on retrouve des résultats similaires à ceux obtenus dans le cas i.i.d. par Scott [21], Freedman et Diaconis [12] puis Lecoutre [17] en multivarié.

THÉORÈME 2.9. *Si les conditions $H_0(i)(iii)$ et $H_2(i)(ii)'$ sont vérifiées, alors pour $h_T = o(T^{-\frac{1}{2}})$:*

$$\limsup_{T \rightarrow \infty} T \cdot EQI\{\hat{f}_H\} \leq 2\Phi.$$

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 : $dX_t = -(aX_t + b)dt + \sigma dW_t$, $X_0, t \geq 0$ où a, b et σ désignent des réels positifs, de loi de probabilité invariante $\mathcal{N}(a/b, \sigma^2/(2a))$ et $\{\xi_t^{(2)}, t \in \mathbb{R}\}$, solution de l'équation différentielle stochastique : $dX_t = -\theta \text{sgn}(X_t)dt + dW_t$, $\theta > 0, X_0, t \geq 0$ de densité invariante $f(x) = \theta e^{-2\theta|x|}$. Ces deux modèles vérifient notamment l'hypothèse $H_2(ii)'$ d'après Leblanc [16] et appartiennent à la classe des processus qui satisfont aux hypothèses du théorème 2.9.

2.2. Le polygone de fréquences.

Il est possible d'améliorer la vitesse de convergence du biais de l'histogramme en introduisant le polygone de fréquences. Les principales propriétés de cet estimateur ont notamment été explorées par Scott [23] (chapitre 4) en i.i.d., par Carbon, Garel et Tran [8] dans le cas mélangeant et, plus récemment, par Carbon [6, 7] pour les champs aléatoires indexés sur treillis. En particulier, Scott [22] a démontré le premier ce résultat remarquable, à savoir que la vitesse de convergence du noyau en $T^{-4/5}$ était également atteinte par le polygone de fréquences. L'étude qui suit est une extension au temps continu et, ce afin de simplifier la présentation, elle sera considérée dans le cadre *unidimensionnel* ($d = 1$).

Notations. $f^{(k)}$ désignera la dérivée d'ordre k de f et on appellera $R_\alpha(f) := \int_{\mathbb{R}} [f^{(\alpha)}(x)]^2 dx$ le coefficient de rigidité de la densité.

Construction de l'estimateur. On construit 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 2.10. Au point $x \in \mathbb{R}$, on définit l'estimateur \hat{f}_{FP} de f par

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

qui peut s'écrire également sous la forme symétrique de Lagrange

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

ou encore sous la forme des différences divisées de Newton

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

On notera ensuite $\hat{f}_{FP}(x) := \hat{f}_{FP}(x, \mathcal{X}, T)$.

REMARQUE 4. On trouve dans la littérature des variantes du polygone de fréquences en modifiant la méthode d'interpolation entre les rectangles de l'histogramme. Parmi elles, le "Edge Frequency Polygon", introduit par Jones, Samiuddin, Al-Harbey et Maatouk [14], s'obtient en calculant les valeurs moyennes $\{(\hat{f}_{j-1} + \hat{f}_j)/2\}$ aux bords $\{b_j\}$ des intervalles de la partition Π_T puis en reliant ces valeurs par des segments de droite. Le "Generalized Edge Frequency Polygon", proposé ensuite par Dong et Zheng

[10], généralise le précédent estimateur en remplaçant la valeur calculée à chacun des bords $\{b_j\}$ par une moyenne pondérée des $2k$, $k \geq 1$ valeurs $\hat{f}_{j-k}, \dots, \hat{f}_j, \dots, \hat{f}_{j+k-1}$. En i.i.d. et en univarié, notons que l'erreur en moyenne quadratique intégrée asymptotique de ces différentes versions du polygone de fréquences gardent la même forme à des constantes de biais et de variance près :

$$EQI\{\hat{f}_{FP}\} \approx \frac{49}{2880}R_2(f)h_n^4 + \frac{2}{3} \frac{1}{nh_n};$$

$$EQI\{\hat{f}_{EFP}\} \approx \frac{23}{360}R_2(f)h_n^4 + \frac{5}{12} \frac{1}{nh_n};$$

$$EQI\{\hat{f}_{GEFP}\} \approx \frac{23 + 90\tau + 90\tau^2}{360}R_2(f)h_n^4$$

$$+ \left(\frac{1}{4}\omega_1^2 + \frac{1}{6}\omega_k^2 + \frac{1}{6} \sum_{i=1}^{k-1} (\omega_i^2 + \omega_i\omega_{i+1} + \omega_{i+1}^2) \right) \frac{1}{nh_n},$$

où $k \geq 2$ est fixé, $\sum_{i=1}^k \omega_i = 1$, $\omega_i \geq 0$ et $\tau = \sum_{i=1}^k \omega_i(-i + i^2)$. Si $k = 1$ ($\Leftrightarrow \omega_1 = 1$), le "GEFP" coïncide avec le "EFP". Par ailleurs, des choix optimaux du paramètre h_n minimisant ces erreurs sont également proposés dans les références sus-mentionnées.

Pour donner le BCI de \hat{f}_{FP} , nous modifions légèrement les hypothèses H_0 de régularité sur f :

Hypothèses H'_0 .

- (i) $f \in \mathcal{C}^2(\mathbb{R})$, $f(x) > 0$ $\mu - p.p.$, $f^{(2)} \in L^1(\mathbb{R})$ et $f^{(k)} \in L^2(\mathbb{R})$ pour $k = 0, 2$;
- (ii) $|f^{(2)}(x) - f^{(2)}(y)| \leq l_0|x - y|^\nu$ pour $l_0 > 0$, $\nu \in]0, 1]$ et $(x, y) \in \mathbb{R}^2$;
- (iii) f est continue en x et $\|f\|_\infty = M_1 < \infty$;
- (iv) $f_{(X_s, X_t)} = f_{(X_0, X_{|t-s|})}$ pour $s \neq t$.

Convergence dans L^2 . De même que pour l'histogramme, l' EQI de \hat{f}_{FP} s'écrit comme la somme du biais au carré intégré et de la variance intégrée. Le lemme suivant nous donne le BCI du polygone de fréquences. Il s'agit d'une version analogue au résultat de Scott [22] obtenue sous des conditions légèrement différentes. Notons qu'avec la condition höldérienne $H'_0(ii)$, nous n'avons pas besoin de supposer l'existence de la dérivée troisième de f .

LEMME 2.11. Si les conditions $H'_0(i)(ii)$ sont vérifiées, alors

$$BCI\{\hat{f}_{FP}\} = \frac{49}{2880} R_2(f) h_T^4 + o(h_T^4).$$

THÉORÈME 2.12. 1. Si les conditions $H'_0(iii)(iv)$ et $H_1(ii)-(iv)$ sont vérifiées et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R})$ avec $1 < p \leq \varrho - 1$, alors pour $F_p := \int_{\mathbb{R}} f^{1-\frac{1}{p}}(x) dx$ et $0 \leq \varepsilon \leq 1 - \frac{1}{\varrho-p}$:

$$\begin{aligned} Th_T \cdot VI\{\hat{f}_{FP}\} \leq & \left\{ \frac{1}{T} \iint_{[0,T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon \right. \\ & \left. + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho-p} F_p \cdot h_T^{\frac{1}{p}\{(1-\varepsilon)(\varrho-p)-1\}} \right\} (1 + o(1)), \end{aligned}$$

2. Si les conditions $H'_0(iv)$ et $H_2(i)(ii)'$ sont vérifiées, alors pour $\Phi := \int_{\mathbb{R}} \varphi(x) dx$

$$\limsup_{T \rightarrow \infty} T \cdot VI\{\hat{f}_{FP}\} \leq 2\Phi.$$

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

THÉORÈME 2.13. 1. Si les conditions H'_0 et H_1 sont vérifiées et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R})$ pour $1 < p \leq \varrho - 1$, le choix $h_T = k_T \cdot T^{-\frac{1}{5}}$ tel que $k_T \rightarrow k$, $0 < k < \infty$ entraîne

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

2. Si les conditions $H'_0(i)(ii)(iv)$ et $H_2(i)(ii)'$ sont vérifiées et si $h_T = o(T^{-\frac{1}{4}})$, alors

$$\limsup_{T \rightarrow \infty} T \cdot EQI\{\hat{f}_{FP}\} \leq 2\Phi.$$

Commentaires.

1. Dans la première partie du théorème 2.13, nous retrouvons la même vitesse obtenue par Scott [22] dans le cas i.i.d. puis par Carbon, Garel et Tran [8] dans le cas mélangeant. Cette vitesse est optimale au sens où elle est effectivement atteinte par un processus construit de manière analogue à Bosq [3] (p. 96) et vérifiant H'_0 et H_1 . Notons de plus qu'elle est minimax (cf. Bosq [3], chapitre 4).

2. Dans la seconde partie, 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).

REMARQUE 5 (Taille des échantillons). Des simulations sont proposées par Scott [22] dans le contexte i.i.d. Celles-ci mettent en évidence l'intérêt de l'estimateur par polygone de fréquences en ce qui concerne la taille des échantillons. Notamment, lorsque la fonction f théorique est la densité d'une Gaussienne $\mathcal{N}(0, 1)$, pour obtenir une erreur en moyenne quadratique intégrée inférieure à $1/400$: 2254 observations sont nécessaires avec un histogramme, 431 avec un noyau d'Epanechnikov et 546 avec un polygone de fréquences. Simonoff et Hurvich [25] présentent des cas de figure où le polygone de fréquence est meilleur que l'estimateur à noyau et soulignent les bonnes performances du polygone pour estimer les queues lourdes de distribution.

REMARQUE 6 (Positionnement des bords). Il est clair que la construction des graphiques de fréquences dépend directement de la partition Π_T , c'est-à-dire des choix de la position des bords $\{b_j, j \in J_T\}$ et du paramètre de lissage h_T . Lorsque $h_T =: h > 0$ est fixé, la forme des graphiques peut ainsi varier sensiblement avec les bords et induire des interprétations erronées de la densité théorique (cf. Silverman [24] en temps discret, p. 7-11). Néanmoins, l'influence des bords sur les caractéristiques asymptotiques de ces graphiques apparaît négligeable d'après des simulations de Scott [23] (p. 65) en i.i.d. Pour l'histogramme univarié, deux partitions sont notamment considérées où le point $x = 0$ coïncide tantôt avec le milieu d'un intervalle tantôt avec un bord. Le critère EQI exact est minimal dans le premier cas de figure et si la fonction f est la densité d'une loi Gaussienne $\mathcal{N}(0, 1)$, l'écart entre les deux erreurs associées est de 1,29% avec 25 observations et inférieur à 10^{-5} à partir de 100 observations. Pour une loi log-normale, cet écart est aussi inférieur à 10^{-5} à partir de 400 observations. En multivarié, le problème se complique encore avec un choix supplémentaire, celui de l'orientation des axes (ou des hypercubes).

3. Normalité asymptotique. Dans cette section, nous supposons le processus \mathcal{X} strictement stationnaire.

Notation. " $\xrightarrow{\mathcal{L}}$ " désignera la convergence en distribution.

3.1. L'histogramme.

Lorsque les observations sont i.i.d., la convergence en distribution de l'histogramme multivarié vers une Gaussienne est donnée par Lecoutre [4] (p. 153). Pour l'extension au temps continu, on rappelle un théorème central-limite proposé par Bosq, Merlevède et Peligrad [5] pour les tableaux triangulaires multidimensionnels et appliqué pour prouver la normalité asymptotique de l'estimateur à noyau dans les cas discret et continu.

Soit $\mathcal{Y} = \{Y_{n,i}, i \in \mathbb{Z}\}$, $n \geq 1$ un processus à temps discret, strictement stationnaire, à valeurs dans \mathbb{R}^m avec $m \in \mathbb{N}^*$ et tel que $Y_{n,i} = (Y_{n,i}^{(1)}, \dots, Y_{n,i}^{(m)})'$. Le coefficient de mélange fort lié à \mathcal{Y} est donné par

$$\alpha_{\mathcal{Y}}(k) = \sup_{n \geq 1} \alpha(\sigma(Y_{n,i}, i \leq 0), \sigma(Y_{n,i}, i \geq k)).$$

Hypothèses H_3 .

- (i) $\mathbb{E}(Y_{n,i}^{(r)}) = 0$ et $|Y_{n,i}^{(r)}| \leq M_n$ pour tout $n \geq 1$, $i \geq 1$ et $1 \leq r \leq m$, où M_n est une constante qui ne dépend que de n ;
- (ii) $\sup_{i \geq 1, 1 \leq r \leq m} \mathbb{E}(Y_{n,i}^{(r)})^2 < \infty$;
- (iii) pour tout $1 \leq r \leq m$, $1 \leq s \leq m$ et pour toute suite $(b_n)_{n \geq 1}$ telle que $b_n \rightarrow \infty$ quand $n \rightarrow \infty$ et $b_n \leq n$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{b_n} \mathbb{E} \left(\sum_{i=1}^{b_n} Y_{n,i}^{(r)} \sum_{j=1}^{b_n} Y_{n,j}^{(s)} \right) = \sigma_{r,s};$$

- (iv) il existe un réel $1 < a < \infty$ tel que $\sum_{k \geq 1} k \alpha_{\mathcal{Y}}^{\frac{a-1}{a}}(k) < \infty$;

- (v) $M_n \leq C \cdot n^{\frac{a^2}{(3a-1)(2a-1)}}$ pour une constante $C > 0$ et pour tout $n \geq 1$.

THÉORÈME 3.1. *Si les conditions H_3 sont vérifiées alors pour $\Sigma^{(m)} := (\sigma_{r,s})_{1 \leq r, s \leq m}$:*

$$\frac{\sum_{i=1}^n Y_{n,i}}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} N^{(m)} \sim \mathcal{N}(0, \Sigma^{(m)}).$$

La normalité asymptotique de l'histogramme en temps continu se déduit ensuite du théorème 3.1 et des hypothèses de Castellana-Leadbetter [9]. Pour ce faire, nous considérons un m -uplet (x_1, \dots, x_m) de points distincts de \mathbb{R}^d . Au point x_r , $1 \leq r \leq m$, on suppose l'existence d'un indice $j(x_r, T)$ tel que $x_r \in \pi_{Tj(x_r, T)}$ et l'estimateur par histogramme vaut $\hat{f}_H(x_r) = \frac{1}{Th_T^d} \int_0^T \mathbf{1}_{\pi_{Tj(x_r, T)}}(X_t) dt$.

Dans cette section, le coefficient $\alpha_{\mathcal{X}}$ de mélange fort associé au processus \mathcal{X} est donné par

$$\alpha_{\mathcal{X}}(k) = \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_k^{+\infty}) \downarrow_{k \rightarrow \infty} 0.$$

THÉORÈME 3.2. *Supposons les conditions $H_2(i)(ii)$ satisfaites et que, pour tout m -uplet (x_1, \dots, x_m) de points distincts de \mathbb{R}^d tel que $\Sigma^{(m)} = (2G(x_r, x_s))_{1 \leq r, s \leq m}$ soit définie positive, la fonction $g_{0,u}(\cdot, \cdot)$ est continue au point (x_r, x_s) pour chaque $1 \leq r, s \leq m$. Si de plus, il existe un réel $1 < a < \infty$ tel que $\sum_{k \geq 1} k \alpha_{\mathcal{X}}^{(a-1)/a}(k) < \infty$, alors pour $h_T \geq C \cdot T^{-a^2/(d(3a-1)(2a-1))}$, $C > 0$:*

$$\sqrt{T}(\hat{f}_H(x_r) - \mathbb{E}\hat{f}_H(x_r), 1 \leq r \leq m) \xrightarrow[T \rightarrow \infty]{\mathcal{L}} N^{(m)} \sim \mathcal{N}^{(m)}(0, \Sigma^{(m)}).$$

REMARQUE 7. La principale hypothèse du théorème 3.2 repose sur l'existence d'un réel $1 < a < \infty$ tel que le coefficient de mélange fort associé à \mathcal{X} satisfasse la vitesse de convergence $\sum_{k \geq 1} k \alpha_{\mathcal{X}}^{(a-1)/a}(k) < \infty$. Notons que lorsque a est proche de 1, l'hypothèse $H_3(iv)$ est encore vérifiée si le processus \mathcal{X} est géométriquement fortement mélangeant. Enfin, le cas $a = \infty$ est également possible (cf. Bosq, Merlevède et Peligrad [5]).

3.2. Le polygone de fréquences.

De façon analogue à l'histogramme, nous établissons dans le contexte suproptimal la normalité asymptotique du polygone de fréquences en temps continu.

THÉORÈME 3.3. *Si les conditions du théorème 3.2 sont satisfaites avec $d = 1$, alors*

$$\sqrt{T}(\hat{f}_{FP}(x_r) - \mathbb{E}\hat{f}_{FP}(x_r), 1 \leq r \leq m) \xrightarrow[T \rightarrow \infty]{\mathcal{L}} N^{(m)} \sim \mathcal{N}^{(m)}(0, \Sigma^{(m)}).$$

Le corollaire suivant est une version utile du précédent théorème pour construire des régions de confiance du vecteur $(f(x_1), \dots, f(x_m))$.

COROLLAIRE 3.4. *Sous les conditions du théorème 3.3 et si de plus $f \in \mathcal{C}_b^2(\mathbb{R})$, alors pour $h_T = o(T^{-1/4})$:*

$$\sqrt{T}(\hat{f}_{FP}(x_r) - f(x_r), 1 \leq r \leq m) \xrightarrow[T \rightarrow \infty]{\mathcal{L}} N^{(m)} \sim \mathcal{N}^{(m)}(0, \Sigma^{(m)}).$$

REMARQUE 8. Les conditions $h_T = \mathcal{O}(T^{-a^2/((3a-1)(2a-1))})$ et $h_T = o(T^{-1/4})$ sont compatibles pour des valeurs de $a \in]1, (5 + \sqrt{17})/4[$ (cf. Bosq, Merlevède et Peligrad [5]).

Finalement, pour déduire des régions de confiance du corollaire 3.4, nous devons encore chercher à estimer la matrice $\Sigma^{(m)}$ car la valeur de ses éléments est, la plupart du temps, inconnue. Ce problème qui n'est pas abordé ici ouvre une perspective de recherche intéressante pour prolonger ce travail. En attendant, si $d = 1$ et $m = 1$, des travaux récents traitent la question de l'estimation de $G(x, x) = \int_0^{+\infty} g_{0,u}(x, x) dx$ et apportent des réponses adaptées. Guillou et Merlevède [13] proposent ainsi deux estimateurs construits de façons différentes à partir de l'estimateur à noyau. Une autre approche qui utilise l'estimateur du temps local est ensuite examinée par Blanke et Merlevède [2]. Dans la plupart des cas (par exemple, pour le processus d'Ornstein-Uhlenbeck), l'estimateur du temps local atteint de meilleures vitesses de convergence pour l'erreur en moyenne quadratique. Cependant, l'usage des estimateurs à noyau est plus simple en pratique. En choisissant l'une de ces méthodes, nous obtenons déjà un intervalle de confiance pour la valeur $f(x_1)$.

4. Discussion. Nous avons présenté des résultats théoriques qui valident le bon potentiel d'estimateurs basiques fréquemment choisis parmi les praticiens. Le polygone de fréquences est particulièrement intéressant avec des vitesses de convergence comparables à celles des estimateurs à noyau et l'opportunité d'obtenir des régions de fréquences si la matrice des variances et covariances asymptotiques est convenablement estimée. En ce sens, un estimateur de type polygone pour cette matrice pourra être envisagé plus tard. Par ailleurs, notons que le contexte de cette étude, où la trajectoire est parfaitement observée sur $[0, T]$, est évidemment purement théorique mais il donne déjà une bonne indication sur les vitesses idéales. Nous pourrions ensuite essayer d'approcher ces vitesses dans des cadres plus réalistes où, par exemple, le processus n'est observé qu'à des instants discrétisés suivant des schémas aléatoires ou déterministes. Ce cadre est également plus favorable à la mise en œuvre de simulations pour justifier empiriquement du bon comportement de ces estimateurs. Enfin et toujours en temps continu, nous chercherons à généraliser les propriétés de ces estimateurs pour des classes plus vastes de fonctions positives (fonction de hasard, intensité d'un processus de Poisson, etc.). Pour la régression, on pourra aussi étudier une version continue du regressogramme et d'un estimateur analogue construit à partir du polygone de fréquences.

5. Preuves.

5.1. *Preuve du théorème 2.4.* Pour un $x \in \pi_{T_j} \subset \mathbb{R}^d$, l'estimateur par histogramme vaut $\hat{f}_H(x) = \hat{f}_j = \frac{1}{Th_T^d} \int_{[0,T]} \mathbf{1}_{\pi_{T_j}}(X_t) dt$ et

$$\begin{aligned} Th_T^d \cdot \text{Var}\{\hat{f}_j\} &= \frac{1}{Th_T^d} \iint_{[0,T]^2} \text{Cov}(\mathbf{1}_{\pi_{T_j}}(X_s), \mathbf{1}_{\pi_{T_j}}(X_t)) ds dt \\ &= \frac{1}{Th_T^d} \left\{ \underbrace{\iint_{[0,T]^2 \cap \Gamma} \text{Cov}(\mathbf{1}_{\pi_{T_j}}(X_s), \mathbf{1}_{\pi_{T_j}}(X_t)) ds dt}_{=: I_1} \right. \\ &\quad \left. + \underbrace{\iint_{[0,T]^2 \cap \Gamma^c} \text{Cov}(\mathbf{1}_{\pi_{T_j}}(X_s), \mathbf{1}_{\pi_{T_j}}(X_t)) ds dt}_{=: I_2} \right\}. \end{aligned}$$

Supposons ensuite les conditions $H_0(ii)(iii)$ et $H_1(ii) - (iv)$ satisfaites et notons $p_k := \mathbb{P}(X_0 \in \pi_{T_k})$, $k \in J_T$.

Majoration de la première intégrale avec l'inégalité de Cauchy-Schwarz :

$$\begin{aligned} I_1 &\leq \iint_{[0,T]^2 \cap \Gamma} \sqrt{\text{Var}(\mathbf{1}_{\pi_{T_j}}(X_s))} \sqrt{\text{Var}(\mathbf{1}_{\pi_{T_j}}(X_t))} ds dt \\ &\leq \iint_{[0,T]^2 \cap \Gamma} \text{Var}(\mathbf{1}_{\pi_{T_j}}(X_0)) ds dt \leq p_j(1-p_j) \iint_{[0,T]^2 \cap \Gamma} ds dt. \end{aligned}$$

Comme f est continue, il existe un $\xi_j \in \pi_{T_j}$ tel que $\int_{\pi_{T_j}} f(x) dx = h_T^d f(\xi_j)$. D'où

$$I_1 \leq h_T^d f(\xi_j)(1 - h_T^d f(\xi_j)) \iint_{[0,T]^2 \cap \Gamma} ds dt.$$

Majoration de la seconde intégrale avec l'hypothèse de stationnarité $H_0(iii)$ et l'inégalité de Davydov (cf. lemme 1.3) :

$$\begin{aligned} I_2 &= \iint_{[0,T]^2 \cap \Gamma^c} \{ \mathbb{P}(X_s \in \pi_{T_j}, X_t \in \pi_{T_j}) - p_j^2 \} ds dt \\ &\leq 2T \int_{u_0}^{+\infty} | \mathbb{P}(X_0 \in \pi_{T_j}, X_u \in \pi_{T_j}) - p_j^2 | du \\ &\leq 2T \left\{ \underbrace{\int_{u_0}^{h_T^{-d+\varepsilon}} | \mathbb{P}(X_0 \in \pi_{T_j}, X_u \in \pi_{T_j}) - p_j^2 | du}_{=: I_{21}} \right. \\ &\quad \left. + \underbrace{\int_{h_T^{-d+\varepsilon}}^{+\infty} | \mathbb{P}(X_0 \in \pi_{T_j}, X_u \in \pi_{T_j}) - p_j^2 | du}_{=: I_{22}} \right\}, \end{aligned}$$

où ε est un nombre réel dont on précisera les valeurs ensuite.

Comme f est majorée par M_1 , on déduit des hypothèses $H_1(ii)(iii)$:

$$\begin{aligned} I_{21} &\leq \int_{u_0}^{h_T^{-d+\varepsilon}} \iint_{\pi_{Tj}^2} |f(X_0, X_u)(y, z) - f(y)f(z)| dy dz du \\ &\leq \int_{u_0}^{h_T^{-d+\varepsilon}} \iint_{\pi_{Tj}^2} f(y) |f_{u/0}(z/y) - f(z)| dy dz du \\ &\leq h_T^d M \int_{u_0}^{h_T^{-d+\varepsilon}} \int_{\pi_{Tj}} f(y) dy du \quad \text{où } M = \max(M_1, M_2) \\ &\leq (h_T^{-d+\varepsilon} - u_0) \cdot h_T^{2d} M f(\xi_j) \end{aligned}$$

$$\Rightarrow I_{21} \leq h_T^{d+\varepsilon} \cdot M f(\xi_j).$$

En utilisant l'inégalité de Davydov pour un couple $(p, q) \in]1, \varrho - 1] \times [1 + \frac{\varrho}{\varrho-2}, \infty[$ tel que $\frac{2}{q} + \frac{1}{p} = 1$, on a

$$I_{22} \leq \int_{h_T^{-d+\varepsilon}}^{+\infty} 2p \cdot 2^{\frac{1}{p}} \cdot \|1_{\pi_{Tj}}(X_0)\|_q \|1_{\pi_{Tj}}(X_u)\|_q \cdot \alpha_{\mathcal{X}}^{\frac{1}{p}}(u) du,$$

soit avec l'hypothèse de mélange arithmétique $H_1(iv)$:

$$\begin{aligned} I_{22} &\leq 2p \cdot 2^{\frac{1}{p}} \cdot p_j^{\frac{2}{q}} \cdot \int_{h_T^{-d+\varepsilon}}^{+\infty} a_0^{\frac{1}{p}} u^{-\frac{\varrho}{p}} du \\ &\leq 2p (2a_0)^{\frac{1}{p}} \cdot h_T^{\frac{2d}{q}} f(\xi_j)^{\frac{2}{q}} \cdot \frac{p}{\varrho - p} h_T^{\frac{1}{p}((d-\varepsilon)(\varrho-p))} \\ \Rightarrow I_{22} &\leq h_T^{\frac{1}{p}((d-\varepsilon)(\varrho-p)+(p-1)d)} \cdot \frac{2p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} f(\xi_j)^{1-\frac{1}{p}}. \end{aligned}$$

D'où

$$I_2 \leq Th_T^d \cdot \left(h_T^\varepsilon \cdot 2M f(\xi_j) + h_T^{\frac{1}{p}((d-\varepsilon)(\varrho-p)-d)} \cdot \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} f(\xi_j)^{1-\frac{1}{p}} \right).$$

Pour déterminer les valeurs de ε , il suffit de remarquer que l'expression entre parenthèses converge quelque soit ε fixé entre 0 et $d(1 - \frac{1}{\varrho-p})$.

$$\begin{aligned} Th_T^d \cdot Var\{\hat{f}_j\} &\leq f(\xi_j)(1 - h_T^d f(\xi_j)) \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + 2M f(\xi_j) \cdot h_T^\varepsilon \\ &\quad + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} f(\xi_j)^{1-\frac{1}{p}} \cdot h_T^{\frac{1}{p}((d-\varepsilon)(\varrho-p)-d)} < \infty. \end{aligned}$$

5.2. *Preuve du théorème 2.5.* Sous $H_0(iii)$, $H_2(i)$ et avec l'hypothèse $H_2(ii)$ de Castellana-Leadbetter,

$$\begin{aligned} T \cdot \text{Var}\{\hat{f}_j\} &= \frac{1}{Th_T^{2d}} \iint_{[0,T]^2} \text{Cov}(\mathbf{1}_{\pi_{Tj}}(X_{\tau_1}), \mathbf{1}_{\pi_{Tj}}(X_{\tau_2})) d\tau_1 d\tau_2 \\ &= \frac{2}{Th_T^{2d}} \int_0^T \int_u^T \text{Cov}(\mathbf{1}_{\pi_{Tj}}(X_0), \mathbf{1}_{\pi_{Tj}}(X_u)) du du \\ &= \frac{2}{h_T^{2d}} \iint_{\pi_{Tj}^2} \int_0^T (1 - \frac{u}{T}) g_{0,u}(y, z) du dy dz. \end{aligned}$$

On cherche maintenant un équivalent asymptotique de cette expression pour $x \in \pi_{Tj} = \prod_{k=1}^d [b_j^{(k)}, b_{j+1}^{(k)}[=: [b_j, b_{j+1}[$ tel que $b_j \uparrow_{T \rightarrow \infty} x$ et $b_{j+1} = b_j + h_T \mathbf{1}_d \downarrow_{T \rightarrow \infty} x$, où $\mathbf{1}_d$ désigne le vecteur colonne d'ordre d composé de 1.

$$\begin{aligned} T \cdot \text{Var}\{\hat{f}_j\} &= 2 \iint_{[0,1]^{2d}} \int_0^T (1 - \frac{u}{T}) g_{0,u}(b_j + h_{Ts}, b_j + h_{Tt}) du ds dt \\ &= 2 \iint_{[0,1]^{2d}} \{ \Delta_T(b_j + h_{Ts}, b_j + h_{Tt}) \\ &\quad + G(b_j + h_{Ts}, b_j + h_{Tt}) \} ds dt, \end{aligned}$$

avec

$$\Delta_T(y, z) := \int_0^T (1 - \frac{u}{T}) g_{0,u}(y, z) du - \int_0^{+\infty} g_{0,u}(y, z) du,$$

et

$$\begin{aligned} |\Delta_T(b_j + h_{Ts}, b_j + h_{Tt})| &= \left| \int_T^{+\infty} g_{0,u}(b_j + h_{Ts}, b_j + h_{Tt}) du \right. \\ &\quad \left. + \int_0^{+\infty} \mathbf{1}_{[0,T]}(u) \frac{u}{T} g_{0,u}(b_j + h_{Ts}, b_j + h_{Tt}) du \right| \\ &\leq \underbrace{\left| \int_T^{+\infty} g_{0,u}(b_j + h_{Ts}, b_j + h_{Tt}) du \right|}_{\leq \int_T^{+\infty} \|g_{0,u}\|_\infty du \rightarrow 0} \\ &\quad + \left| \int_0^{+\infty} \underbrace{\mathbf{1}_{[0,T]}(u) \frac{u}{T} g_{0,u}(b_j + h_{Ts}, b_j + h_{Tt}) du}_{\rightarrow 0} \right|, \end{aligned}$$

soit $|\Delta_T(b_j + h_{Ts}, b_j + h_{Tt})| \xrightarrow{T \rightarrow \infty} 0$ par le théorème de convergence dominée de Lebesgue. Par ailleurs, comme $G(\cdot, \cdot)$ est continue au point (x, x) :

$$G(b_j + h_{Ts}, b_j + h_{Tt}) \xrightarrow{T \rightarrow \infty} G(x, x).$$

Enfin, on a

$$|\Delta_T(b_j + h_T s, b_j + h_T t) + G(b_j + h_T s, b_j + h_T t)| \leq 3 \int_0^{+\infty} \|g_{0,u}\|_\infty du,$$

tel que $(s, t) \mapsto 3 \int_0^{+\infty} \|g_{0,u}\|_\infty du \in L^1([0, 1]^{2d})$, et nous déduisons du théorème de convergence dominée :

$$\lim_{T \rightarrow \infty} T \cdot \text{Var}\{\hat{f}_j\} = 2 \iint_{[0,1]^{2d}} G(x, x) dy dz = 2G(x, x).$$

En substituant l'hypothèse $H_2(ii)$ par celle plus faible $H_2(ii)'$, nous avons la majoration suivante :

$$\begin{aligned} T \cdot \text{Var}\{\hat{f}_j\} &= \frac{2}{h_T^{2d}} \iint_{\pi_{Tj}^2} \int_0^T (1 - \frac{u}{T}) g_{0,u}(y, z) du dy dz \\ &\leq \frac{2}{h_T^{2d}} \iint_{\pi_{Tj}^2} \underbrace{\sup_{z \in \mathbb{R}^d} \int_0^{+\infty} |g_{0,u}(y, z)| du}_{=:\varphi(y)} dy dz \\ &\leq \frac{2}{h_T^{2d}} h_T^d \cdot h_T^d \varphi(\eta_j) \quad \text{où } \eta_j \in \pi_{Tj}. \end{aligned}$$

Finalement,

$$T \cdot \text{Var}\{\hat{f}_j\} \leq 2\varphi(\eta_j) < \infty.$$

5.3. *Preuve du théorème 2.6.* $VI\{\hat{f}_H\} = \sum_{j \in J_T} \int_{\pi_{Tj}} \text{Var}\{\hat{f}_j\} dx$ et sous les hypothèses du théorème 2.4 avec $0 \leq \varepsilon \leq d(1 - \frac{1}{e-p})$, on a

$$\begin{aligned} Th_T^d \cdot \int_{\pi_{Tj}} \text{Var}\{\hat{f}_j\} dx &\leq h_T^d \left\{ f(\xi_j)(1 - h_T^d f(\xi_j)) \cdot \frac{1}{T} \iint_{[0,T]^{2\Gamma}} ds dt \right. \\ &\quad + 2M f(\xi_j) \cdot h_T^\varepsilon \\ &\quad \left. + \frac{4p^2(2a_0)^{\frac{1}{p}}}{e-p} f(\xi_j)^{1-\frac{1}{p}} \cdot h_T^{\frac{1}{p}\{(d-\varepsilon)(e-p)-d\}} \right\}. \end{aligned}$$

Il reste maintenant à sommer sur l'ensemble des indices $j \in J_T$ de la partition Π_T . Comme les fonctions $\{f^\kappa\}_{\kappa=1-\frac{1}{p}, 1, 2}$ sont supposées Riemann-intégrables sur \mathbb{R}^d , on peut écrire

$$\sum_{j \in J_T} h_T^d f^\kappa(\xi_j) = \int_{\mathbb{R}^d} f^\kappa(x) dx + o(1),$$

et

$$Th_T^d \cdot VI\{\hat{f}_H\} \leq \left\{ \frac{1}{T} \iint_{[0,T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(d-\varepsilon)(\varrho-p)-d\}} \right\} (1 + o(1)).$$

5.4. *Preuve du théorème 2.7.* Pour $x \in \pi T_j$, si les conditions $H_0(iii)$ et $H_2(i)(ii)'$ sont vérifiées, alors

$$T \cdot Var\{\hat{f}_j\} \leq 2\varphi(\eta_j), \quad \eta_j \in \pi T_j,$$

et on en déduit que

$$T \cdot VI\{\hat{f}_H\} \leq 2 \sum_{j \in J_T} h_T^d \varphi(\eta_j) \leq 2\Phi + o(1).$$

5.5. *Preuve du théorème 2.8.* Sous H_0 et H_1 et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R}^d)$ pour $1 < p \leq \varrho - 1$, avec le choix $h_T = c_T \cdot T^{-\frac{1}{d+2}}$ tel que $c_T \rightarrow c$, $0 < c < \infty$, on a d'une part :

$$T^{\frac{2}{d+2}} \cdot BCI\{\hat{f}_H\} = \frac{c_T^2}{12} \Psi_d^2 + o(1) \Rightarrow \lim_{T \rightarrow \infty} T^{\frac{2}{d+2}} \cdot BCI\{\hat{f}_H\} = \frac{c^2}{12} \Psi_d^2.$$

D'autre part, si $p = \varrho - 1$ ($\Leftrightarrow \varepsilon = 0$) :

$$T^{\frac{2}{d+2}} \cdot VI\{\hat{f}_H\} \leq \frac{1}{c_T^d} \left\{ \frac{1}{T} \iint_{[0,T]^2 \cap \Gamma} ds dt + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \right\} (1 + o(1)) \\ \Rightarrow \limsup_{T \rightarrow \infty} T^{\frac{2}{d+2}} \cdot VI\{\hat{f}_H\} \leq \frac{1}{c^d} \left\{ \ell_\Gamma + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \right\},$$

et si $p < \varrho - 1$ ($\Leftrightarrow \varepsilon > 0$) :

$$T^{\frac{2}{d+2}} \cdot VI\{\hat{f}_H\} \leq \frac{1}{c_T^d} \left\{ \frac{1}{T} \iint_{[0,T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(d-\varepsilon)(\varrho-p)-d\}} \right\} (1 + o(1)) \\ \Rightarrow \limsup_{T \rightarrow \infty} T^{\frac{2}{d+2}} \cdot VI\{\hat{f}_H\} \leq \frac{\ell_\Gamma}{c^d}.$$

Finalement, pour $p = \varrho - 1$, on obtient

$$\limsup_{T \rightarrow \infty} T^{\frac{2}{\varrho+2}} \cdot EQI\{\hat{f}_H\} \leq \frac{c^2}{12} \Psi_d^2 + \frac{1}{c^d} \left\{ \ell_\Gamma + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \right\},$$

et pour $p < \varrho - 1$:

$$\limsup_{T \rightarrow \infty} T^{\frac{2}{\varrho+2}} \cdot EQI\{\hat{f}_H\} \leq \frac{c^2}{12} \Psi_d^2 + \frac{\ell_\Gamma}{c^d}.$$

5.6. *Preuve du théorème 2.9.* Sous $H_0(i)(iii)$ et $H_2(i)(ii)'$ et si le paramètre h_T est un $o(T^{-\frac{1}{2}})$, le terme de biais est négligeable devant le terme de variance et la vitesse suroptimale de l'histogramme se déduit immédiatement du théorème 2.7 :

$$\limsup_{T \rightarrow \infty} T \cdot EQI\{\hat{f}_H\} \leq 2\Phi.$$

5.7. *Preuve du lemme 2.11.* Nous reprenons ici les grandes lignes de la preuve apportée par Scott [22] en concluant avec la condition höldérienne $H'_0(ii)$ sur $f^{(2)}$. Soit $x \in [c_j, c_{j+1}[$, $\hat{f}_{FP}(x) = (\frac{x-c_j}{h_T})\hat{f}_{j+1} + (\frac{c_{j+1}-x}{h_T})\hat{f}_j$. Dans un premier temps, on écrit un développement limité de f à l'ordre 2 au voisinage de b_{j+1} ($=: m_j$), milieu de l'intervalle $[c_j, c_{j+1}[$:

$$f(x) = f(m_j) + (x - m_j)f^{(1)}(m_j) + \frac{1}{2}(x - m_j)^2 f^{(2)}(m_j + \theta(x - m_j)), \theta \in]0, 1[.$$

On calcule $\mathbb{E}\{\hat{f}_j\}$ et $\mathbb{E}\{\hat{f}_{j+1}\}$:

$$\begin{aligned} \mathbb{E}\{\hat{f}_j\} &= \frac{1}{h_T} \int_{\pi_{Tj}} f(x) dx = f(m_j) - \frac{f^{(1)}(m_j)}{2} h_T + \frac{f^{(2)}(\xi_j)}{6} h_T^2, \\ \mathbb{E}\{\hat{f}_{j+1}\} &= \frac{1}{h_T} \int_{\pi_{T(j+1)}} f(x) dx = f(m_j) + \frac{f^{(1)}(m_j)}{2} h_T + \frac{f^{(2)}(\xi_{j+1})}{6} h_T^2, \end{aligned}$$

où $\xi_j \in \pi_{Tj}$ et $\xi_{j+1} \in \pi_{T(j+1)}$. On en déduit alors une expression du biais de \hat{f}_{FP} au point x :

$$\begin{aligned} \text{Biais}\{\hat{f}_{FP}(x)\} &= \mathbb{E}\{\hat{f}_{FP}(x)\} - f(x) \\ &= \frac{1}{6}(x - c_j)f^{(2)}(\xi_{j+1})h_T + \frac{1}{6}(c_{j+1} - x)f^{(2)}(\xi_j)h_T \\ &\quad - \frac{1}{2}(x - m_j)^2 f^{(2)}(m_j + \theta(x - m_j)). \end{aligned}$$

On élève ensuite le biais au carré :

$$\begin{aligned}
 & \text{Biais}^2\{\hat{f}_{FP}(x)\} \\
 = & \left(\frac{x-c_j}{6}\right)^2 f^{(2)}(\xi_{j+1})^2 h_T^2 + \left(\frac{c_{j+1}-x}{6}\right)^2 f^{(2)}(\xi_j)^2 h_T^2 \\
 & + \frac{1}{4}(x-m_j)^4 f^{(2)}(m_j + \theta(x-m_j))^2 \\
 & + \frac{(x-c_j)(c_{j+1}-x)}{18} f^{(2)}(\xi_j) f^{(2)}(\xi_{j+1}) h_T^2 \\
 & - \left(\frac{x-c_j}{6}\right)(x-m_j)^2 f^{(2)}(\xi_{j+1}) f^{(2)}(m_j + \theta(x-m_j)) h_T \\
 & - \left(\frac{c_{j+1}-x}{6}\right)(x-m_j)^2 f^{(2)}(\xi_j) f^{(2)}(m_j + \theta(x-m_j)) h_T \\
 =: & Q_1(x, T) + Q_2(x, T) + Q_3(x, T) + Q_4(x, T) - Q_5(x, T) - Q_6(x, T),
 \end{aligned}$$

et on intègre chacun des termes sur $[c_j, c_{j+1}]$:

$$\begin{aligned}
 \int_{c_j}^{c_{j+1}} Q_1(x, T) dx &= \frac{f^{(2)}(\xi_{j+1})^2}{108} h_T^5, \\
 \int_{c_j}^{c_{j+1}} Q_2(x, T) dx &= \frac{f^{(2)}(\xi_j)^2}{108} h_T^5, \\
 \int_{c_j}^{c_{j+1}} Q_3(x, T) dx &= \frac{f^{(2)}(\bar{x}_j)^2}{320} h_T^5 \quad \text{où } \bar{x}_j \in [c_j, c_{j+1}], \\
 \int_{c_j}^{c_{j+1}} Q_4(x, T) dx &= \frac{f^{(2)}(\xi_j) f^{(2)}(\xi_{j+1})}{108} h_T^5, \\
 \int_{c_j}^{c_{j+1}} Q_5(x, T) dx &= \frac{f^{(2)}(\hat{x}_j) f^{(2)}(\xi_{j+1})}{144} h_T^5 \quad \text{où } \hat{x}_j \in [c_j, c_{j+1}], \\
 \int_{c_j}^{c_{j+1}} Q_6(x, T) dx &= \frac{f^{(2)}(\bar{x}_j) f^{(2)}(\xi_j)}{144} h_T^5 \quad \text{où } \bar{x}_j \in [c_j, c_{j+1}].
 \end{aligned}$$

D'où

$$\begin{aligned}
 & \int_{c_j}^{c_{j+1}} \text{Biais}^2\{\hat{f}_{FP}(x)\} dx \\
 = & h_T^5 \left\{ \frac{f^{(2)}(\xi_{j+1})^2}{108} + \frac{f^{(2)}(\xi_j)^2}{108} + \frac{f^{(2)}(\bar{x}_j)^2}{320} + \frac{f^{(2)}(\xi_j) \cdot f^{(2)}(\xi_{j+1})}{108} \right. \\
 & \left. - \frac{f^{(2)}(\hat{x}_j) \cdot f^{(2)}(\xi_{j+1})}{144} - \frac{f^{(2)}(\bar{x}_j) \cdot f^{(2)}(\xi_j)}{144} \right\}.
 \end{aligned}$$

Avec $H'_0(ii)$, on écrit

$$\begin{aligned} f^{(2)}(\xi_{j+1}) &= f^{(2)}(\xi_j) + A_j & \text{où } A_j &= f^{(2)}(\xi_{j+1}) - f^{(2)}(\xi_j) \text{ et } |A_j| \leq l_0 h_T^\nu, \\ f^{(2)}(\hat{x}_j) &= f^{(2)}(\xi_{j+1}) + B_j & \text{où } B_j &= f^{(2)}(\hat{x}_j) - f^{(2)}(\xi_{j+1}) \text{ et } |B_j| \leq l_0 h_T^\nu, \\ f^{(2)}(\bar{x}_j) &= f^{(2)}(\xi_j) + C_j & \text{où } C_j &= f^{(2)}(\bar{x}_j) - f^{(2)}(\xi_j) \text{ et } |C_j| \leq l_0 h_T^\nu, \end{aligned}$$

et en remarquant que

$$\begin{aligned} & \left| h_T^4 \sum_{j \in J_T} h_T \left(\frac{f^{(2)}(\xi_j) \cdot A_j}{108} - \frac{f^{(2)}(\xi_{j+1}) \cdot B_j}{144} - \frac{f^{(2)}(\xi_j) \cdot C_j}{144} \right) \right| \\ & \leq h_T^{4+\nu} \cdot l_0 \left| \frac{1}{108} + \frac{1}{144} + \frac{1}{144} \right| \left(\int_{\mathbb{R}} |f^{(2)}(x)| dx + o(1) \right) \\ & \leq h_T^{4+\nu} \cdot \frac{5l_0 \int_{\mathbb{R}} |f^{(2)}(x)| dx}{216} (1 + o(1)), \end{aligned}$$

on obtient

$$\sum_{j \in J_T} h_T^5 \left(\frac{f^{(2)}(\xi_j) \cdot A_j}{108} - \frac{f^{(2)}(\xi_{j+1}) \cdot B_j}{144} - \frac{f^{(2)}(\xi_j) \cdot C_j}{144} \right) = \mathcal{O}(h_T^{4+\nu}).$$

On en déduit l'expression du biais au carré intégré :

$$BCI\{\hat{f}_{FP}\} = \left\{ \frac{1}{108} + \frac{1}{108} + \frac{1}{320} + \frac{1}{108} - \frac{1}{144} - \frac{1}{144} \right\} R_2(f) h_T^4 + o(h_T^4).$$

Finalement,

$$BCI\{\hat{f}_{FP}\} = \frac{49}{2880} R_2(f) h_T^4 + o(h_T^4). \quad \blacksquare$$

5.8. *Preuve du théorème 2.12.* La preuve est basée sur la décomposition suivante :

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

La variance intégrée de \hat{f}_{FP} s'obtient ensuite en intégrant sur $[c_j, c_{j+1}[$ et en sommant sur tous les indices j de J_T .

Notons \bar{E}_j une expression indépendante de x qui majore $Var\{\hat{f}_j\}$, $j \in J_T$.
On a alors

$$\begin{aligned} |Cov\{\hat{f}_j, \hat{f}_{j+1}\}| &\leq \sqrt{Var\{\hat{f}_j\}}\sqrt{Var\{\hat{f}_{j+1}\}} \leq \frac{1}{2}\{Var\{\hat{f}_j\} + Var\{\hat{f}_{j+1}\}\} \\ &\leq \frac{1}{2}\{\bar{E}_j + \bar{E}_{j+1}\}. \end{aligned}$$

Ainsi,

$$\begin{aligned} \int_{c_j}^{c_{j+1}} Var\{\hat{f}_{FP}(x)\} dx &\leq \frac{h_T}{3}\bar{E}_{j+1} + \frac{h_T}{3}\bar{E}_j + \frac{h_T}{6}\{\bar{E}_j + \bar{E}_{j+1}\} \\ &\leq \frac{h_T}{2}\{\bar{E}_j + \bar{E}_{j+1}\}, \end{aligned}$$

et la suite découle des majorations établies dans les théorèmes 2.4 et 2.5.

Si les conditions $H'_0(iii)(iv)$ et $H_1(ii) - (iv)$ sont vérifiées et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R})$ avec $1 < p \leq \varrho - 1$, alors pour $0 \leq \varepsilon \leq 1 - \frac{1}{\varrho - p}$:

$$\begin{aligned} Th_T \cdot \int_{c_j}^{c_{j+1}} Var\{\hat{f}_{FP}(x)\} dx &\leq \frac{h_T}{2} \left\{ f(\xi_j)(1 - h_T f(\xi_j)) \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + h_T^\varepsilon \cdot 2M f(\xi_j) \right. \\ &\quad \left. + h_T^{\frac{1}{2}\{(1-\varepsilon)(\varrho-p)-1\}} \cdot \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho-p} f(\xi_j)^{1-\frac{1}{p}} \right. \\ &\quad \left. + f(\xi_{j+1})(1 - h_T f(\xi_{j+1})) \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + h_T^\varepsilon \cdot 2M f(\xi_{j+1}) \right. \\ &\quad \left. + h_T^{\frac{1}{2}\{(1-\varepsilon)(\varrho-p)-1\}} \cdot \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho-p} f(\xi_{j+1})^{1-\frac{1}{p}} \right\}. \end{aligned}$$

Si les conditions $H'_0(iv)$ et $H_2(i)(ii)'$ sont vérifiées, alors

$$T \cdot \int_{c_j}^{c_{j+1}} Var\{\hat{f}_{FP}(x)\} dx \leq \frac{h_T}{2} \{2\varphi(\eta_j) + 2\varphi(\eta_{j+1})\} \leq h_T \{\varphi(\eta_j) + \varphi(\eta_{j+1})\}.$$

Il reste maintenant à sommer sur l'ensemble des indices $j \in J_T$ de la partition Π_T . Comme les fonctions $\{f^\kappa\}_{\kappa=1-\frac{1}{p}, 1, 2}$ sont supposées Riemann-intégrables sur \mathbb{R} , on peut écrire

$$\sum_{j \in J_T} h_T f^\kappa(\xi_j) = \int_{\mathbb{R}} f^\kappa(x) dx + o(1).$$

Par conséquent, sous $H'_0(iii)(iv)$ et $H_1(ii) - (iv)$ pour $0 \leq \varepsilon \leq 1 - \frac{1}{\varrho - p}$:

$$Th_T \cdot VI\{\hat{f}_{FP}\} \leq \left\{ \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(1-\varepsilon)(\varrho-p)-1\}} \right\} (1 + o(1)),$$

et sous $H'_0(iv)$ et $H_2(i)(ii)'$:

$$T \cdot VI\{\hat{f}_{FP}\} \leq 2\Phi + o(1).$$

5.9. *Preuve du théorème 2.13.* Par passage à la limite, on obtient les vitesses de convergence de l'erreur en moyenne quadratique intégrée du polygone de fréquences à partir du lemme 2.11 et du théorème 2.12. Sous H'_0 et H_1 et si $f^{1-\frac{1}{p}} \in L^1(\mathbb{R})$ pour $1 < p \leq \varrho - 1$, avec le choix $h_T = k_T \cdot T^{-\frac{1}{5}}$ tel que $k_T \rightarrow k$, $0 < k < \infty$, on a d'une part :

$$\begin{aligned} T^{\frac{4}{5}} \cdot BCI\{\hat{f}_{FP}\} &= \frac{49}{2880} k_T^4 R_2(f) + o(1) \\ &\Rightarrow \lim_{T \rightarrow \infty} T^{\frac{4}{5}} \cdot BCI\{\hat{f}_{FP}\} = \frac{49}{2880} k^4 R_2(f). \end{aligned}$$

D'autre part, si $p = \varrho - 1$ ($\Leftrightarrow \varepsilon = 0$) :

$$\begin{aligned} T^{\frac{4}{5}} \cdot VI\{\hat{f}_{FP}\} &\leq \frac{1}{k_T} \left\{ \frac{1}{T} \iint_{[0, T]^2 \cap \Gamma} ds dt + 2M \cdot h_T^\varepsilon + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \cdot h_T^{\frac{1}{p}\{(1-\varepsilon)(\varrho-p)-1\}} \right\} (1 + o(1)) \\ &\Rightarrow \limsup_{T \rightarrow \infty} T^{\frac{4}{5}} \cdot VI\{\hat{f}_{FP}\} \leq \frac{1}{k} \left\{ \ell_\Gamma + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \right\}, \end{aligned}$$

et si $p < \varrho - 1$, on peut améliorer la constante en choisissant un $\varepsilon > 0$ tel que

$$\limsup_{T \rightarrow \infty} T^{\frac{4}{5}} \cdot VI\{\hat{f}_{FP}\} \leq \frac{\ell_\Gamma}{k}.$$

Par conséquent, pour $p = \varrho - 1$, on obtient

$$\limsup_{T \rightarrow \infty} T^{\frac{4}{5}} \cdot EQI\{\hat{f}_{FP}\} \leq \frac{49}{2880} k^4 R_2(f) + \frac{1}{k} \left\{ \ell_\Gamma + 2M + \frac{4p^2(2a_0)^{\frac{1}{p}}}{\varrho - p} F_p \right\},$$

et pour $p < \varrho - 1$:

$$\limsup_{T \rightarrow \infty} T^{\frac{2}{\varrho+2}} \cdot EQI\{\hat{f}_{FP}\} \leq \frac{49}{2880} k^4 R_2(f) + \frac{\ell_T}{k}.$$

Sous $H'_0(i)(iv)$ et $H_2(i)(ii)'$ et si le paramètre h_T est un $o(T^{-\frac{1}{4}})$, le biais est négligeable devant la variance et on obtient immédiatement la vitesse suroptimale du polygone de fréquences :

$$\limsup_{T \rightarrow \infty} T \cdot EQI\{\hat{f}_{FP}\} \leq 2\Phi.$$

5.10. *Preuve du théorème 3.2.* La preuve du théorème repose sur une application directe du théorème 3.1. On notera $[x]$ le plus grand entier strictement plus petit que le réel x . Considérons un m -uplet (x_1, \dots, x_m) de points distincts de \mathbb{R}^d . Au point x_r , $1 \leq r \leq m$, on suppose l'existence d'un indice $j(x_r, T) =: j(r)$ tel que $x_r \in \pi_{Tj(r)}$, et à partir des valeurs $\hat{f}_H(x_r)$, pour tout $n \geq 1$, on construit une suite strictement stationnaire de vecteurs aléatoires réels centrés $Y_{n,i} = \{(Y_{n,i}^{(1)}, \dots, Y_{n,i}^{(m)})', 1 \leq i \leq n\}$ tels que

$$Y_{n,i}^{(r)} = \frac{1}{\sqrt{\delta} h_T^d} \left\{ \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt - \mathbb{E} \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt \right\},$$

où $n\delta = T_n = T$, $n = [T]$ ($T \geq 1$) et ainsi $1 \leq \delta < 2$. Pour tout $n \geq 1$ et $1 \leq r \leq m$, on a alors

$$\frac{\sum_{i=1}^n Y_{n,i}}{\sqrt{n}} = \{\sqrt{T}(\hat{f}_H(x_1) - \mathbb{E}\hat{f}_H(x_1)), \dots, \sqrt{T}(\hat{f}_H(x_m) - \mathbb{E}\hat{f}_H(x_m))\}.$$

Pour conclure à la convergence en loi de $(\sum_{i=1}^n Y_{n,i}/\sqrt{n})_{n \geq 1}$ vers un vecteur Gaussien, nous devons au préalable vérifier que les hypothèses $H_3(i) - (iv)$ du théorème 3.1 sont satisfaites :

i. Pour tout $n \geq 1$, $i \geq 1$ et $1 \leq r \leq m$,

$$\begin{aligned} |Y_{n,i}^{(r)}| &= \frac{1}{\sqrt{\delta} h_T^d} \left| \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt - \mathbb{E} \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt \right| \\ &= \frac{1}{\sqrt{\delta} h_T^d} \left| \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt - \delta \mathbb{P}(X_0 \in \pi_{Tj(r)}) \right| < \frac{\sqrt{2}}{h_T^d} < \infty. \end{aligned}$$

ii. Pour tout $i \geq 1$ et chaque $1 \leq r \leq m$,

$$\begin{aligned} \mathbb{E}(Y_{n,i}^{(r)})^2 &= \mathbb{E}(Y_{n,1}^{(r)})^2 = \text{Var} \left(\frac{1}{\sqrt{\delta} h_T^d} \int_0^\delta \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt \right) \\ &= \frac{2}{h_T^{2d}} \iint_{\pi_{Tj(r)}^2} \int_0^\delta \left(1 - \frac{u}{\delta}\right) g_{0,u}(y, z) du dy dz, \end{aligned}$$

et si la condition $H_2(ii)$ de Castellana-Leadbetter est vérifiée, on a

$$\mathbb{E}(Y_{n,i}^{(r)})^2 \leq 2 \int_0^{+\infty} \|g_{0,u}\|_\infty du < \infty.$$

iii. Soit a_n une suite de nombres entiers telle que $a_n \rightarrow \infty$ et $a_n \leq n$ pour tout $n \geq 1$. Pour chaque $1 \leq r, s \leq m$,

$$\begin{aligned} &\frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) \\ &= \frac{1}{\delta a_n h_T^{2d}} \iint_{[0, \delta a_n]^2} \text{Cov}(\mathbf{1}_{\pi_{Tj(r)}}(X_{\tau_1}), \mathbf{1}_{\pi_{Tj(s)}}(X_{\tau_2})) d\tau_1 d\tau_2 \\ &= \frac{2}{h_T^{2d}} \iint_{\pi_{Tj(r)} \times \pi_{Tj(s)}} \int_0^{\delta a_n} \left(1 - \frac{u}{\delta a_n}\right) g_{0,u}(y, z) du dy dz \\ &= 2 \iint_{[0,1]^{2d}} \int_0^{\delta a_n} \left(1 - \frac{u}{\delta a_n}\right) g_{0,u}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) du dp dq. \end{aligned}$$

Nous reprenons ensuite les étapes de la preuve du théorème 2.5.

$$\begin{aligned} \frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) &= 2 \iint_{[0,1]^{2d}} \{ \Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) \\ &\quad + G(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) \} du dp dq, \end{aligned}$$

avec

$$\Delta_{a_n}(y, z) := \int_0^{\delta a_n} \left(1 - \frac{u}{\delta a_n}\right) g_{0,u}(y, z) du - \int_0^{+\infty} g_{0,u}(y, z) du,$$

et

$$\begin{aligned} &|\Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ})| \\ &\leq \underbrace{\left| \int_{\delta a_n}^{+\infty} g_{0,u}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) du \right|}_{\leq \int_{\delta a_n}^{+\infty} \|g_{0,u}\|_\infty du \rightarrow 0} \\ &\quad + \underbrace{\left| \int_0^{+\infty} \mathbf{1}_{[0, \delta a_n]}(u) \frac{u}{\delta a_n} g_{0,u}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) du \right|}_{\rightarrow 0}, \end{aligned}$$

soit $|\Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ})| \xrightarrow[n \rightarrow \infty]{} 0$. D'autre part, comme $G(\cdot, \cdot)$ est continue au point (x_r, x_s) :

$$\begin{aligned} (b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) &\xrightarrow{T \rightarrow \infty} (x_r, x_s) \\ \Rightarrow G(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) &\xrightarrow{T \rightarrow \infty} G(x_r, x_s). \end{aligned}$$

Enfin, la quantité $|G(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) + \Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ})|$ est majorée par la fonction $(p, q) \mapsto 3 \int_0^{+\infty} \|g_{0,u}\|_{\infty} du$ qui est intégrable sur le support $[0, 1]^{2d}$ et on déduit du théorème de convergence dominée :

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) = 2G(x_r, x_s) < \infty.$$

iv. Enfin, comme $\alpha_Y(k) \leq \alpha_X(k)$, la condition $\sum_{k \geq 1} k \alpha_X^{(a-1)/a}(k) < \infty$ implique $\sum_{k \geq 1} k \alpha_Y^{(a-1)/a}(k) < \infty$, ce qui achève la démonstration du théorème. ■

5.11. *Preuve du théorème 3.3.* Considérons un m -uplet (x_1, \dots, x_m) de points distincts de \mathbb{R}^d . Au point x_r , $1 \leq r \leq m$, on suppose l'existence d'un indice $j(x_r, T) =: j(r)$ tel que $x_r \in \pi_{Tj(r)}$ et l'estimateur par polygone de fréquences vaut

$$\hat{f}_{FP}(x_r) = \begin{cases} \left(\frac{x_r - c_{j(r)}}{h_T} \right) \hat{f}_{j(r)+1} + \left(\frac{c_{j(r)+1} - x_r}{h_T} \right) \hat{f}_{j(r)} & \text{si } x_r \in \pi_{Tj(r)}^+, \\ \left(\frac{x_r - c_{j(r)-1}}{h_T} \right) \hat{f}_{j(r)} + \left(\frac{c_{j(r)} - x_r}{h_T} \right) \hat{f}_{j(r)-1} & \text{si } x_r \in \pi_{Tj(r)}^-, \end{cases}$$

où $\pi_{Tj(r)}^+ := \pi_{Tj(r)} \cap [c_{j(r)}, c_{j(r)+1}[$ et $\pi_{Tj(r)}^- := \pi_{Tj(r)} \cap]c_{j(r)-1}, c_{j(r)}]$. Sans perte de généralité, on supposera que $x_r \in \pi_{Tj(r)}^+$ et que

$$\hat{f}_{FP}(x_r) = \left(\frac{x_r - c_{j(r)}}{h_T} \right) \hat{f}_{j(r)+1} + \left(\frac{c_{j(r)+1} - x_r}{h_T} \right) \hat{f}_{j(r)}.$$

A partir de $\hat{f}_{FP}(x_r)$, pour tout $n \geq 1$ et chaque $1 \leq r \leq m$, on construit une suite strictement stationnaire de variables aléatoires réelles centrées $\{Y_{n,i}, 1 \leq i \leq n\}$ définies par

$$\begin{aligned} Y_{n,i}^{(r)} &= \frac{(x_r - c_{j(r)})}{\sqrt{\delta h_T^2}} \left\{ \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{T,j(r)+1}}(X_t) dt - \mathbb{E} \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{T,j(r)+1}}(X_t) dt \right\} \\ &+ \frac{(c_{j(r)+1} - x_r)}{\sqrt{\delta h_T^2}} \left\{ \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt - \mathbb{E} \int_{\delta(i-1)}^{\delta i} \mathbf{1}_{\pi_{Tj(r)}}(X_t) dt \right\}, \end{aligned}$$

où $n\delta = T_n = T$, $n = [T]$ ($T \geq 1$) et $1 \leq \delta < 2$.

Pour tout $n \geq 1$ et $1 \leq r \leq m$, on a alors

$$\frac{\sum_{i=1}^n Y_{n,i}}{\sqrt{n}} = \{\sqrt{T}(\hat{f}_{FP}(x_1) - \mathbb{E}\hat{f}_{FP}(x_1)), \dots, \sqrt{T}(\hat{f}_{FP}(x_m) - \mathbb{E}\hat{f}_{FP}(x_m))\}.$$

Il nous reste à vérifier que les hypothèses du théorème 3.1 sont satisfaites :

i. Pour tout $n \geq 1$, $i \geq 1$ et $1 \leq r \leq m$,

$$|Y_{n,i}^{(r)}| \leq \left| \frac{x_r - c_{j(r)}}{h_T} \right| \frac{\sqrt{\delta}}{h_T} + \left| \frac{c_{j(r)+1} - x_r}{h_T} \right| \frac{\sqrt{\delta}}{h_T} < \frac{\sqrt{2}}{h_T} < \infty.$$

ii. Pour tout $i \geq 1$ et chaque $1 \leq r \leq m$, $\mathbb{E}(Y_{n,i}^{(r)})^2 = \mathbb{E}(Y_{n,1}^{(r)})^2$ et

$$\begin{aligned} \mathbb{E}(Y_{n,1}^{(r)})^2 &= \frac{2(x_r - c_{j(r)})^2}{h_T^4} \iint_{\pi_{T,j(r)+1}^2} \int_0^\delta \left(1 - \frac{u}{\delta}\right) g_{0,u}(y, z) \, du \, dy \, dz \\ &+ \frac{2(c_{j(r)+1} - x_r)^2}{h_T^4} \iint_{\pi_{T,j(r)}^2} \int_0^\delta \left(1 - \frac{u}{\delta}\right) g_{0,u}(y, z) \, du \, dy \, dz \\ &+ \frac{4(x_r - c_{j(r)})(c_{j(r)+1} - x_r)}{h_T^4} \\ &\quad \times \iint_{\pi_{T,j(r)} \times \pi_{T,j(r)+1}} \int_0^\delta \left(1 - \frac{u}{\delta}\right) g_{0,u}(y, z) \, du \, dy \, dz, \end{aligned}$$

et si la condition $H_2(ii)$ de Castellana-Leadbetter est satisfaite, on a

$$\begin{aligned} \int_0^\delta \left(1 - \frac{u}{\delta}\right) g_{0,u}(y, z) \, du &\leq \int_0^{+\infty} \|g_{0,u}\|_\infty \, du < \infty \\ \Rightarrow \mathbb{E}(Y_{n,i}^{(r)})^2 &\leq 2 \int_0^{+\infty} \|g_{0,u}\|_\infty \, du < \infty. \end{aligned}$$

iii. Soit a_n une suite de nombres entiers telle que $a_n \rightarrow \infty$ et $a_n \leq n$ pour tout $n \geq 1$. Avec un calcul analogue à celui détaillé dans les preuves des théorèmes 2.5 et 3.2, on a pour chaque $1 \leq r, s \leq m$,

$$\begin{aligned} \frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) &= \frac{2(x_r - c_{j(r)})}{h_T^2} [(x_s - c_{j(s)})J_1 + (c_{j(s)+1} - x_s)J_2] \\ &+ \frac{2(c_{j(r)+1} - x_r)}{h_T^2} [(x_s - c_{j(s)})J_3 + (c_{j(s)+1} - x_s)J_4], \end{aligned}$$

avec

$$\begin{aligned} J_1 &:= \iint_{[0,1]^2} \{ \Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) \\ &\quad + G(b_{j(r)} + h_{TP}, b_{j(s)} + h_{TQ}) \} \, dp \, dq, \end{aligned}$$

$$J_2 := \iint_{[0,1]^2} \{ \Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)+1} + h_{TQ}) \\ + G(b_{j(r)} + h_{TP}, b_{j(s)+1} + h_{TQ}) \} dp dq,$$

$$J_3 := \iint_{[0,1]^2} \{ \Delta_{a_n}(b_{j(r)+1} + h_{TP}, b_{j(s)} + h_{TQ}) \\ + G(b_{j(r)+1} + h_{TP}, b_{j(s)} + h_{TQ}) \} dp dq,$$

$$J_4 := \iint_{[0,1]^2} \{ \Delta_{a_n}(b_{j(r)} + h_{TP}, b_{j(s)+1} + h_{TQ}) \\ + G(b_{j(r)} + h_{TP}, b_{j(s)+1} + h_{TQ}) \} dp dq,$$

et le même $\Delta_{a_n}(\cdot, \cdot)$ défini dans la preuve du théorème 3.2, tels que

$$J_k = G(x_r, x_s)(1 + o(1)), \quad 1 \leq k \leq 4.$$

D'où

$$\frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) = 2G(x_r, x_s) + o(1).$$

Finalement,

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \mathbb{E} \left(\sum_{i=1}^{a_n} Y_{n,i}^{(r)} \sum_{j=1}^{a_n} Y_{n,j}^{(s)} \right) = 2G(x_r, x_s) < \infty.$$

iv. Enfin, pour tout $k \in \mathbb{N}$, on a clairement $\alpha_Y(k) \leq \alpha_X(k)$. Ainsi, la condition $\sum_{k \geq 1} k \alpha_X^{(a-1)/a}(k) < \infty$ implique $\sum_{k \geq 1} k \alpha_Y^{(a-1)/a}(k) < \infty$, ce qui achève la démonstration du théorème. ■

5.12. *Preuve du corollaire 3.4.* Si de plus $f \in \mathcal{C}_b^2(\mathbb{R})$,

$$|\text{Biais}\{\hat{f}_{FP}(x_r)\}| = |\mathbb{E}\hat{f}_{FP}(x_r) - f(x_r)| = \mathcal{O}(h_T^2).$$

Par conséquent, en prenant $h_T = o(T^{-\frac{1}{4}})$, le résultat du corollaire est immédiat. ■

Remerciements. Je remercie Delphine Blanke et le Professeur Denis Bosq de leur soutien et d'avoir contribué par leurs suggestions à la qualité de ce travail.

Références.

- [1] Bertrand-Retali, M. (1978). Convergence uniforme d'un estimateur de la densité par la méthode du noyau. *Rev. Roumaine Math. Pures Appl.*, **23**(3), 361–385.

- [2] Blanke, D. et Merlevède, F. (2000). Estimation of the asymptotic variance of kernel density estimators for continuous time processes. *Math. Methods Statist.*, **9**(3), 270–296.
- [3] Bosq, D. (1998). *Nonparametric statistics for stochastic processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, New York, second edition. Estimation and prediction.
- [4] Bosq, D. et Lecoutre, J.-P. (1987). *Théorie de l'estimation fonctionnelle*. Economie et Statistiques Avancées. Economica, Paris.
- [5] Bosq, D., Merlevède, F., et Peligrad, M. (1999). Asymptotic normality for density kernel estimators in discrete and continuous time. *J. Multivariate Anal.*, **68**(1), 78–95.
- [6] Carbon, M. (2005). Frequency polygons for random fields. *INSEE - Série des Documents de Travail du CREST n° 2005-04*.
- [7] Carbon, M. (2006). Polygones des fréquences pour des champs aléatoires. *C. R. Acad. Sci. Paris Sér. I Math.*, à paraître.
- [8] Carbon, M., Garel, B., et Tran, L. T. (1997). Frequency polygons for weakly dependent processes. *Statist. Probab. Lett.*, **33**(1), 1–13.
- [9] Castellana, J. V. et Leadbetter, M. R. (1986). On smoothed probability density estimation for stationary processes. *Stochastic Process. Appl.*, **21**(2), 179–193.
- [10] Dong, J. et Zheng, C. (2001). Generalized edge frequency polygon for density estimation. *Statist. Probab. Lett.*, **55**(2), 137–145.
- [11] Doukhan, P. (1994). *Mixing*, volume 85 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Properties and examples.
- [12] Freedman, D. et Diaconis, P. (1981). On the histogram as a density estimator : L_2 theory. *Z. Wahrsch. Verw. Gebiete*, **57**(4), 453–476.
- [13] Guillou, A. et Merlevède, F. (2001). Estimation of the asymptotic variance of kernel density estimators for continuous time processes. *J. Multivariate Anal.*, **79**(1), 114–137.
- [14] Jones, M. C., Samiuddin, M., Al-Harbey, A. H., et Maatouk, T. A. H. (1998). The edge frequency polygon. *Biometrika*, **85**(1), 235–239.
- [15] Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London Ltd., London.
- [16] Leblanc, F. (1997). Density estimation for a class of continuous time processes. *Math. Methods Statist.*, **6**(2), 171–199.
- [17] Lecoutre, J.-P. (1985). The L_2 -optimal cell width for the histogram. *Statist. Probab. Lett.*, **3**(6), 303–306.
- [18] Lejeune, F.-X. (2005). Vitesses optimale et suroptimale des polygones de fréquences pour les processus à temps continu. *C. R. Acad. Sci. Paris Sér. I Math.*, **341**, 59–62.

- [19] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin.
- [20] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, **42**, 43–47.
- [21] Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, **66**(3), 605–610.
- [22] Scott, D. W. (1985). Frequency polygons : theory and application. *J. Amer. Statist. Assoc.*, **80**(390), 348–354.
- [23] Scott, D. W. (1992). *Multivariate density estimation*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. John Wiley & Sons Inc., New York. Theory, practice, and visualization, A Wiley-Interscience Publication.
- [24] Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability. Chapman & Hall, London.
- [25] Simonoff, J. S. et Hurvich, C. (1993). A study of the effectiveness of simple density estimation methods. *Comput. Statist.*, **8**, 259–278.
- [26] Veretennikov, A. Y. (1999). On Castellana-Leadbetter's condition for diffusion density estimation. *Stat. Inference Stoch. Process.*, **2**(1), 1–9 (2000).

E-MAIL: fxlejeun@ccr.jussieu.fr

1. INTRODUCTION

Suppose X and Y are correlated exponential random variables with the joint pdf given by

$$f(x, y) = \frac{1}{1-\rho} \exp\left(-\frac{x+y}{1-\rho}\right) h\left(\frac{2\sqrt{xy\rho}}{1-\rho}\right) \quad (1)$$

for $x > 0$, $y > 0$ and $-1 \leq \rho < 1$, where $h(\cdot)$ denotes the modified Bessel function of the first kind of order ν defined by

$$h(\nu) = \frac{e^{-\nu}}{\Gamma(\nu+1)} \sum_{k=0}^{\infty} \frac{1}{\Gamma(\nu+1+k)} \left(\frac{x^2}{4}\right)^k$$

This distribution is due to Downton (1973) and therefore known as Downton's bivariate exponential distribution. The mathematical properties of this distribution (including joint cumulative distribution function, moment generating function, marginal properties and conditional properties) have been studied by several authors. It has also received applications in several areas. Some of them are:

- **Queueing systems:** Consider a single-server queueing system such that the inter-arrival times X and the service times Y have exponential distributions, as is a common assumption in this context. If it is desired to introduce positive correlation (arising from cooperative service) into the model, Downton's distribution is a suitable choice (Chenly and Chou, 1979). Langlois (1984) applied it to queueing systems with infinitely many servers.