

STATISTIQUE ET ANALYSE DES DONNÉES

H. ABBAR

Un estimateur spline du contour d'une répartition ponctuelle aléatoire

Statistique et analyse des données, tome 15, n° 3 (1990), p. 1-19

http://www.numdam.org/item?id=SAD_1990__15_3_1_0

© Association pour la statistique et ses utilisations, 1990, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

*Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques*

<http://www.numdam.org/>

**UN ESTIMATEUR SPLINE DU CONTOUR
D'UNE RÉPARTITION PONCTUELLE ALÉATOIRE.**

H. ABBAR

Laboratoire de Statistique et Probabilités
U.F.R. de Mathématiques
Université de Lille Flandres Artois
F-59655 VILLENEUVE D'ASCQ CEDEX

Résumé : *Supposons que le support d'un processus ponctuel soit un sous-ensemble de \mathbb{R}^2 défini en coordonnées polaires par $S = \{(\rho, \theta) \mid 0 \leq \rho < \Phi(\theta) ; 0 \leq \theta \leq 2\pi\}$. Cet article est consacré à l'estimation de la fonction Φ par une fonction spline aléatoire.*

Mots clés : Processus ponctuel, support d'une mesure, support d'un processus ponctuel, splines cubiques.

Classification AMS : 60G55 - 62G05

Summary : *Let the domain of a point process be a subset of \mathbb{R}^2 defined in polar coordinates by $S = \{(\rho, \theta) \mid 0 \leq \rho < \Phi(\theta) ; 0 \leq \theta \leq 2\pi\}$. This paper is devoted to the estimation of the function Φ by a random spline.*

Key words and phrases : Point process, support of a measure, domain of a point process, cubic spline.

1 - INTRODUCTION

Il arrive souvent qu'un processus ponctuel se réalise autour d'un point connu qui fait naturellement partie de son support. Par exemple, en Ecologie, l'étude des lieux de capture de proies par un oiseau prédateur autour de son nid donne une idée de son territoire. On peut aussi penser à la zone de chalandise d'un centre commercial ou à la répartition des arbres morts autour d'une industrie polluante.

On considère un processus ponctuel N se réalisant dans un domaine S de \mathbb{R}^2 défini en coordonnées polaires par une fonction Φ continue qu'il s'agit d'estimer :

$$(1.1) \quad S = \{(\rho, \theta) \mid 0 \leq \theta \leq 2\pi ; 0 \leq \rho < \Phi(\theta) ; \Phi(0) = \Phi(2\pi)\}.$$

On trouve en ophtalmologie une méthode de détermination de la zone de la rétine atteinte par une anomalie qui correspond assez bien à ce modèle. Cela consiste à demander au patient de fixer un point déterminé, tandis que des points choisis au hasard s'allument un bref instant dans son champ de vision. Le patient doit indiquer ceux qu'il a pu percevoir. L'expérience est répétée plusieurs fois.

Evidemment, le modèle serait plus général si nous pouvions envisager simultanément le choix de 0 et l'estimation de Φ , c'est-à-dire le cas où aucun point particulier ne s'impose naturellement. Cependant, le problème de la détermination de l'ensemble des points 0 tels que toute demi-droite passant par 0 ne rencontre la frontière du domaine S du processus ponctuel qu'en un point au plus ne peut être résolu de façon simple. En particulier, le barycentre du nuage de points ne peut convenir. D'ailleurs même le barycentre de la mesure moyenne EN n'est pas généralement un point 0 vis à vis duquel le support S soit étoilé, et on peut même imaginer des exemples où ce point serait à l'extérieur de S . Il n'en reste pas moins qu'un bon choix de l'origine permettrait d'éviter des biais ponctuels trop importants. On pourrait naturellement envisager aussi d'autres modèles où la fonction Φ serait par exemple une fonction multivoque.

Dans un article récent avec P. Jacob [Jacob, Abbar, 1989], nous avons étudié la convergence d'un estimateur de Φ de type histogramme, ainsi que sa loi limite. Nous définissons dans l'article que voici un estimateur de Φ de type spline cubique, nous établissons la convergence presque sûre de cet estimateur, et nous présentons enfin quelques expériences de simulation. Pour faciliter la lecture, nous reprenons dans la

section 3 la démonstration de la convergence de l'estimateur défini dans [Jacob, Abbar, 1989], et nous renvoyons pour plus de détails à l'article en question.

2 - PRELIMINAIRES

On considère S de la forme (1.1) et une fonction aléatoire G strictement positive sur l'intérieur de S et continue sur S . Par ailleurs, G est identiquement nulle à l'extérieur de S , et on ne fait pas d'hypothèse particulière quant aux valeurs qu'elle prend sur la frontière de S .

On suppose qu'une observation élémentaire est constituée par un processus ponctuel N dont la loi, conditionnée par $\{G = g\}$, est celle d'un processus de Poisson d'intensité g . Autrement dit, N est un processus de Cox dirigé par une mesure aléatoire M dont G est la densité.

Il est à noter que la "zéro-probability functional" de N est définie, comme dans [Karr, 1986], pour tout borélien borné B de \mathbb{R}^2 par :

$$(2.1) \quad Z(B) = P\{N(B) = 0\} = E\{\exp(-M(B))\},$$

où M est la mesure aléatoire qui dirige le processus N . Tous les éléments aléatoires envisagés dans l'article sont supposés définis sur le même espace probabilisé (Ω, \mathcal{F}, P) .

On suppose dans la suite qu'il est donné n copies indépendantes N_1, \dots, N_n du processus de Cox N dont le support est défini d'après (1.1). On considère une suite $k = k_n$ d'entiers et pour $r \in \{1, \dots, k\}$, on introduit les notations suivantes :

$$(2.2) \quad I_{n,r} = [2\pi(r-1)k^{-1}, 2\pi rk^{-1}[$$

$$(2.3) \quad D_{n,r} = \{(\rho, \theta) \in S \mid \theta \in I_{n,r}\}$$

$$(2.4) \quad M_{n,r} = \sup\{\Phi(\theta) \mid \theta \in I_{n,r}\}$$

$$(2.5) \quad m_{n,r} = \inf\{\Phi(\theta) \mid \theta \in I_{n,r}\}$$

$$(2.6) \quad \theta_{n,r} = 2\pi(r-1)k^{-1} + \pi k^{-1} \text{ (milieu de } I_{n,r} \text{) pour } r \in \{1, \dots, k+1\}.$$

Pour tout $\omega \in \Omega$, on définit S_n^ω comme le support de la mesure aléatoire discrète $N_1^\omega + \dots + N_n^\omega$ et on pose $S_{n,r}^\omega = S_n^\omega \cap D_{n,r}$: l'estimateur $\Phi_n^\omega(\theta)$ de $\Phi(\theta)$ étudié dans [Jacob, Abbar, 1988] est défini sur tout $I_{n,r}$ par :

$$(2.7) \quad \begin{aligned} U_{n,r}^{\omega} &= \max \{ \rho \mid (\rho, \theta) \in S_{n,r}^{\omega} \} \text{ si } S_{n,r}^{\omega} \neq \emptyset \\ U_{n,r}^{\omega} &= 0 \quad \quad \quad \text{si } S_{n,r}^{\omega} = \emptyset \end{aligned}$$

La construction d'un tel estimateur est inspirée d'un article de Geffroy [1964] et de travaux ultérieurs de Gensbittel [1979].

On construit ensuite une fonction spline cubique périodique ψ_n^{ω} interpolant la fonction Φ_n^{ω} aux points $\{\theta_{n,r}; r = 1, \dots, k+1\}$. L'existence et l'unicité de cette fonction ψ_n sont établies dans la section 4 à l'aide de méthodes voisines de celles utilisées par Berlinet dans l'estimation de la densité [Berlinet, 1981].

Pour la clarté de l'exposé, nous reprenons d'abord la convergence presque sûre de $\Lambda_n = \sup \{ |\Phi(\theta) - \Phi_n(\theta)| \mid \theta \in [0, 2\pi] \}$ vers 0.

3 - CONVERGENCE DE L'ESTIMATEUR DE TYPE HISTOGRAMME

Dans cette section, comme dans tous l'article, on supposera remplies les conditions énoncées dans le paragraphe 2.

THEOREME 3.1.- *Supposons que (i) $k_n \rightarrow \infty$ et (ii) $k_n = o(\frac{n}{\log n})$. Alors Λ_n converge vers 0 presque sûrement.*

DEMONSTRATION : Fixons ε dans $]0, 1[$ et considérons :

$$(3.1) \quad \alpha_n = P \left\{ \bigcup_{r=1}^k \{ \varepsilon M_{n,r} \geq U_{n,r} \} \right\}.$$

Si la série $\sum_n \alpha_n$ converge, il résulte du lemme de Borel-Cantelli que pour n assez grand :

$$(3.2) \quad \max_r (M_{n,r} - U_{n,r}) \leq (1-\varepsilon) \sup_{\theta} \{ \Phi(\theta) \}, \text{ p.s.}$$

Par l'inégalité triangulaire, on obtient :

$$(3.3) \quad \Lambda_n \leq \max_r (M_{n,r} - U_{n,r}) + \max_r (M_{n,r} - m_{n,r})$$

D'après (3.2) et la continuité de Φ , ceci implique :

$$(3.4) \quad 0 \leq \limsup_n \Lambda_n \leq (1-\epsilon) \sup_{\theta} \{\Phi(\theta)\} , \text{ p.s.}$$

Si ceci est vrai pour tout ϵ de $]0,1[$, on a $\lim_n \Lambda_n = 0$, p.s.,

Il suffit par conséquent de démontrer que $\sum_n \alpha_n < +\infty$. Fixons ϵ et η tels que $0 < \epsilon < \eta < 1$ et définissons :

$$(3.5) \quad A_{n,r} = \{(\rho, \theta) \in D_{n,r} \mid \epsilon M_{n,r} \leq \rho \leq \eta M_{n,r}\} .$$

Alors, d'après (2.1), on a les inégalités suivantes :

$$(3.6) \quad \begin{aligned} \alpha_n &\leq \sum_{r=1}^k P\left\{ \bigcap_{i=1}^n [N_i(A_{n,r}) = 0] \right\} = \sum_{r=1}^k Z^P(A_{n,r}) \\ &\leq \sum_{r=1}^k E^n \{ \exp(-M(A_{n,r})) \} . \end{aligned}$$

Considérons, maintenant, les variables aléatoires :

$$(3.7) \quad X_n = \inf\{G(\rho, \theta) \mid (\rho, \theta) \in A_{n,r} \ r = 1, \dots, k\} .$$

$$(3.8) \quad Y_n = X_n \cdot 1_{\{X_n \leq \sqrt{k}\}} .$$

Par la continuité de Φ , on a $\bigcup_r A_{n,r} \subset S$ pour n assez grand ; ainsi, il existe une constance $C > 0$ telle que, pour $r = 1, \dots, k$:

$$(3.9) \quad \lambda(A_{n,r}) = \pi(\eta^2 - \epsilon + 2)M_{n,r}^2 k^{-1} \geq Ck^{-1}$$

$$\text{et } M(A_{n,r}) \geq CX_n k^{-1} \geq CY_n k^{-1} .$$

Puisque $Y_n k^{-1} \leq k^{-1/2}$, l'inégalité $\exp(-CY_n k^{-1}) \leq 1 - \frac{CY_n k^{-1}}{2}$ est vérifiée de façon déterministe pour n assez grand. D'après (3.6) et (3.9) il suit que pour n assez grand :

$$\begin{aligned}
 \alpha_n &\leq kE^n\{\exp(-CY_n k^{-1})\} \\
 (3.10) \quad &\leq k\left(1 - \frac{CE(Y_n)k^{-1}}{2}\right)^n \\
 &\leq k \exp(-nCE(Y_n)k^{-1}/2).
 \end{aligned}$$

D'après (ii), $k = (nu)/\log n$, avec $u = u_n \rightarrow 0$, on obtient pour n assez grand :

$$(3.11) \quad \alpha_n \leq \frac{u}{\log n} n^{1-CE(Y_n)/2u}.$$

Il est clair que α_n est le terme général d'une série convergente si on démontre que $\liminf_n E(Y_n) > 0$. Choisissons $\delta > 0$ tel que :

$$(3.12) \quad \delta < \inf_{\theta} \{\Phi(\theta)\} - \eta \inf_{\theta} \{\Phi(\theta)\}.$$

Pour n assez grand, et $r = 1, \dots, k$, on a par la continuité de Φ :

$$(3.13) \quad \eta M_{n,r} \leq \eta(m_{n,r} + \delta) < \eta m_{n,r} + \delta.$$

Ceci conduit aux inclusions suivantes :

$$\begin{aligned}
 (3.14) \quad \bigcup_{r=1}^k A_{n,r} &\subset \bigcup_{r=1}^k \{(\rho, \theta) \in D_{n,r} \mid \rho \leq \eta m_{n,r} + \delta\} \\
 &\subset \{(\rho, \theta) \mid \rho \leq \eta \Phi(\theta) + \delta\} = K,
 \end{aligned}$$

où K est un compact qui satisfait, d'après (3.12) :

$$K \subset \text{int}(S) \subset \{G > 0\}.$$

Par conséquent, pour n assez grand :

$$X_n \geq X = \inf\{G(\rho, \theta) \mid (\rho, \theta) \in K\} > 0.$$

Finalement, par le lemme de Fatou et la remarque évidente $|X_n - Y_n| \rightarrow 0$

$$\begin{aligned}
 \liminf_n E(Y_n) &\geq E(\liminf_n Y_n) \\
 &= E(\liminf_n X_n) \geq E(X) > 0.
 \end{aligned}$$

4. CONSTRUCTION ET CONVERGENCE DE L'ESTIMATEUR DE TYPE SPLINE CUBIQUE

Avant d'aborder la construction et la convergence de l'estimateur spline cubique, nous allons définir les fonctions splines cubiques périodiques.

4.1. Fonctions splines cubiques périodiques.

a) **Définition.**- Une fonction spline cubique périodique (relativement à $\theta_{n,r} = \theta_r$ pour $r = 1, \dots, k+1$) est une fonction ψ périodique de classe C^2 dont la restriction à $[\theta_r, \theta_{r+1}]$, pour $r = 1, \dots, k$, est un polynôme de degré trois au plus.

On pose :

$$M_r = \psi''(\theta_r), \text{ pour } r = 1, \dots, k+1$$

$$h = \frac{2\pi}{k}$$

b) Expression de ψ en fonction des (M_r) .

Comme pour les splines classiques [Ahlberg, Nilson, Walsh, 1977], en imposant à ψ les conditions d'interpolation :

$$\psi(\theta_r) = \rho_r \text{ pour } r = 1, \dots, k+1$$

avec un dernier noeud

$$\begin{cases} \theta_{k+1} = \theta_1 + 2\pi \\ \rho_{k+1} = \rho_1 \end{cases}$$

on obtient pour $\theta \in [\theta_r, \theta_{r+1}]$:

$$(4.1.1) \quad \psi(\theta) = \frac{1}{6} M_r \frac{(\theta_{r+1} - \theta)^3}{h} + \frac{1}{6} M_{r+1} \frac{(\theta - \theta_r)^3}{h} - \frac{M_r}{6} h(\theta_{r+1} - \theta) - \frac{M_{r+1}}{6} h(\theta - \theta_r) + \frac{(\theta - \theta_r)\rho_{r+1} + (\theta_{r+1} - \theta)\rho_r}{h}$$

c) Calcul des (M_r) .

Comme pour les splines cubiques classiques, on construit le système de $k-1$ équations en $k+1$ inconnues M_1, \dots, M_{k+1} qui traduit la continuité de la dérivée première de ψ en $\theta_2, \dots, \theta_k$ et on ajoute à ce système les 2 équations :

$$\begin{cases} M_{k+1} = M_1 \\ \psi'(\theta_{k+1}^+) = \psi'(\theta_1^-) \end{cases} \text{ donc la périodicité est bien supposée,}$$

les (M_r) sont solutions du système suivant :

$$\sum \begin{cases} 2M_1 + \frac{1}{2} M_2 + \frac{1}{2} M_k & = d_1 \\ \frac{1}{2} M_1 + \frac{1}{2} M_{k-1} + 2M_k & = d_k \\ \frac{1}{2} M_{r-1} + 2M_r + \frac{1}{2} M_{r+1} & = d_r \text{ pour } r = 2, \dots, k-1. \end{cases}$$

Avec :

$$d_1 = \frac{3}{h^2} (\rho_2 - 2\rho_1 + \rho_k)$$

$$d_k = \frac{3}{h^2} (\rho_{k-1} - 2\rho_k + \rho_1)$$

$$d_r = \frac{3}{h^2} (\rho_{r-1} - 2\rho_r + \rho_{r+1}) \text{ pour } r = 2, \dots, k-1.$$

Le système \sum peut s'écrire sous la forme matricielle suivante :

$$(4.1.2) \quad \begin{pmatrix} 2 & 1/2 & 0 & \dots & 0 & 0 & 1/2 \\ 1/2 & 2 & 1/2 & \dots & 0 & 0 & 0 \\ 0 & 1/2 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & 1/2 & 0 \\ 0 & 0 & 0 & \dots & 1/2 & 2 & 1/2 \\ 1/2 & 0 & 0 & \dots & 0 & 1/2 & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ \vdots \\ M_{k-1} \\ M_k \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{k-1} \\ d_k \end{pmatrix}$$

où

$$A.M = d$$

La matrice $A = [a_{ij}]$ est une matrice carrée d'ordre k .

LEMME 4.1.1.- Si l'on munit \mathbb{R}^k de la norme du sup, la matrice A est inversible et la norme de A^{-1} est majorée par 1. On aura donc :

$$\sup_{1 \leq i \leq k} |M_i| \leq \sup_{1 \leq i \leq k} |d_i| .$$

DEMONSTRATION : Cette démonstration est en fait tout à fait classique [Ahlberg, Nilson, Walsh, 1977]. Soit $x \in \mathbb{R}^k$ et $y = Ax$ avec :

$$\|x\| = \sup_{1 \leq i \leq k} |x_i| = |x_m| \text{ et } \|y\| = \sup_{1 \leq i \leq k} |y_i|$$

$$\|y\| \geq |a_{mm}x_m + \sum_{j \neq m} a_{mj}x_j|$$

$$\geq |a_{mm}|x_m| - \left| \sum_{j \neq m} a_{mj}x_j \right|$$

A est une matrice à diagonale dominante : $a_{mm} > \sum_{j \neq m} a_{mj}$ donc

$$|a_{mm}|x_m| - \left| \sum_{j \neq m} a_{mj}x_j \right| = a_{mm} |x_m| - \left| \sum_{j \neq m} a_{mj}x_j \right|$$

$$\geq a_{mm}|x_m| - |x_m| \sum_{j \neq m} a_{mj}$$

D'où

$$\frac{\|x\|}{\|y\|} \leq \frac{1}{a_{mm} - \sum_{j \neq m} a_{mj}} = 1$$

Donc

$$\|A^{-1}\| = \sup_{\|y\|} \frac{\|A^{-1}y\|}{\|y\|} = \sup_{\|y\|} \frac{\|x\|}{\|y\|} \leq 1.$$

4.2.- Construction de l'estimateur spline cubique.

Pour tout $\omega \in \Omega$, soit ψ_n^ω la fonction spline cubique périodique qui interpole l'estimateur de type histogramme Φ_n^ω défini d'après (2.7) en les $\{\Phi_r ; r = 1, \dots, k+1\}$. D'après (4.1.1), on a pour $\theta \in [\theta_r, \theta_{r+1}]$:

$$\begin{aligned} \psi_n(\theta) = & \frac{1}{6} M_r \frac{(\theta_{r+1} - \theta)^3}{h} + \frac{1}{6} M_{r+1} \frac{(\theta - \theta_r)^3}{h} - \frac{M_r}{6} h(\theta_{r+1} - \theta) \\ & - \frac{M_{r+1}}{6} h(\theta - \theta_r) + \frac{(\theta - \theta_r)u_{r+1} + (\theta_{r+1} - \theta)U_r}{h}, \end{aligned}$$

où les (M_r) sont les solutions du système défini d'après (4.1.2) avec :

$$d_1 = \frac{3}{h^2} (U_2 - 2U_1 + U_k)$$

$$d_k = \frac{3}{h^2} (U_{k-1} - 2U_k + U_1)$$

$$d_r = \frac{3}{h^2} (U_{r-1} - 2U_r + U_{r+1}) \text{ pour } r = 2, \dots, k-1.$$

4.3. Convergence de l'estimateur spline cubique.

La convergence de l'estimateur spline est déduite de celle de l'estimateur du type histogramme. Ceci résulte des propositions suivantes : soit ψ la fonction spline cubique périodique qui interpole la fonction Φ en les $\{\theta_r ; r = 1, \dots, K+1\}$.

PROPOSITION 4.3.1.- $\sup\{|\Phi(\theta) - \psi(\theta)| \mid \theta \in [0, 2\pi]\} \leq 6\mu(\Phi, h)$ où $\mu(\Phi, \varepsilon) = \sup\{|\Phi(\theta) - \Phi(\theta')| \mid |\theta - \theta'| \leq \varepsilon\}$ est le module de continuité de Φ .

PROPOSITION 4.3.2.- Pour tout $\omega \in \Omega$ on a :

$$\sup\{|\psi_n^\omega(\theta) - \psi(\theta)| \mid \theta \in [0, 2\pi]\} \leq 10 \sup\{|\Phi(\theta) - \Phi_n^\omega(\theta)| \mid \theta \in [0, 2\pi]\}.$$

DÉMONSTRATIONS : Les démonstration de ces deux propositions sont basées sur l'écriture de la spline cubique (interpolant (θ_r, ρ_r)). Si $\theta \in [\theta_{r+1}, \theta_r]$ alors :

$$(*) \quad \begin{cases} \psi(\theta) = \frac{1}{6} M_r \frac{(\theta_{r+1} - \theta)^3}{h} + \frac{1}{6} M_{r+1} \frac{(\theta - \theta_r)^3}{h} - \frac{M_r}{6} h(\theta_{r+1} - \theta) \\ - \frac{M_{r+1}}{6} h(\theta - \theta_r) + \frac{(\theta - \theta_r)\rho_{r+1} - (\theta_{r+1} - \theta)\rho_r}{h} \\ \text{où les } M_r \text{ sont les solutions de } A \begin{bmatrix} M_1 \\ \vdots \\ M_k \end{bmatrix} = \begin{bmatrix} \vdots \\ \frac{3}{h^2} (\rho_{r-1} - 2\rho_r + \rho_{r+1}) \\ \vdots \end{bmatrix} \end{cases}$$

En remplaçant chaque ρ_n par $\phi(\theta_r)$ dans l'expression (*), on obtient :

$$|\psi(\theta) - \Phi(\theta)| \leq \frac{h^2}{3} |M_r| + \frac{h^2}{3} |M_{r+1}| + 2 \max \{ |\Phi(\theta_r) - \Phi(\theta)|, |\Phi(\theta_{r+1}) - \Phi(\theta)| \}$$

et comme $|\Phi(\theta_{r-1}) - \Phi(\theta_r)| \leq \mu(\Phi, h)$, on a $|M_r| \leq \frac{6}{h^2} \mu(\Phi, h)$ d'où nous déduisons la proposition 4.3.1.

En remplaçant cette fois chaque ρ_r par $(\Phi(\theta_r) - \Phi_n(\theta_r))$ dans l'expression (*), (ce qui modifie les M_r), on obtient :

$$|\psi_n(\theta) - \psi(\theta)| \leq \frac{h^2}{3} |M_r| + \frac{h^2}{3} |M_{r+1}| + 2 \sup_{0 \leq \theta \leq 2\pi} |\Phi(\theta) - \Phi_n(\theta)|$$

et comme

$$|M_r| \leq \frac{3}{h^2} 4 \sup_{0 \leq \theta \leq 2\pi} |\Phi(\theta) - \Phi_n(\theta)|$$

on a

$$|\psi_n(\theta) - \psi(\theta)| \leq 10 \sup_{0 \leq \theta \leq 2\pi} |\Phi(\theta) - \Phi_n(\theta)|$$

ce qui établit la proposition 4.3.2.

THEOREME 4.3.1.- *Supposons que (i) $k_n \rightarrow \infty$ et (ii) $k_n = o(\frac{n}{\log n})$. Alors $\Gamma_n = \sup\{|\Phi(\theta) - \psi_n(\theta)| \mid \theta \in [0, 2\pi]\}$ converge vers 0 presque sûrement.*

DEMONSTRATION : Par l'inégalité triangulaire on a :

$$\begin{aligned} \Gamma_n &\leq \sup \{ |\Phi(\theta) - \psi(\theta)| \mid \theta \in [0, 2\pi] \} \\ &\quad + \sup \{ |\psi(\theta) - \psi_n(\theta)| \mid \theta \in [0, 2\pi] \} . \end{aligned}$$

D'après les propositions 4.3.1 et 4.3.2 on a :

$$\Gamma_n \leq 10 \sup \{|\Phi(\theta) - \Phi_n(\theta)| \mid \theta \in [0, 2\pi]\} + 6\mu(\Phi, h).$$

Φ est une fonction continue donc :

$$0 \leq \limsup_n \Gamma_n \leq 10 \lim_n \Lambda_n .$$

D'après le théorème 3.1, on a $\lim_n \Lambda_n = 0$, p.s. et par conséquent $\Gamma_n \rightarrow 0$ p.s.

5 - SIMULATIONS

Nous présentons dans ce paragraphe les résultats de deux expériences de simulation dans lesquelles la fonction à estimer est de la forme :

$$\Phi(\theta) = -\frac{\theta^2}{\pi^2} + \frac{2\theta}{\pi} + \frac{1}{4} + \cos \theta$$

Dans les quatre premiers cas, le processus ponctuel est un processus de Poisson N . $N(S)$ suit une loi de Poisson de paramètre $n = 400$ dans les cas 1 et 2, de paramètre $n = 1000$ dans les cas 3 et 4. Dans les figures 1 et 2, le nombre k de secteurs est égal à 18, dans les figures 3 et 4, k est égal à 36. La loi de probabilité d'un point aléatoire (ρ, θ) est telle que θ suit la loi uniforme sur $[0, 2\pi]$ et ρ suit la loi uniforme sur $[0, \Phi(\theta)]$, ρ et θ étant indépendants.

Dans les cas suivants, $N(S)$ suit une loi de Poisson de paramètre $n = 400$ (cas 5 et 6) ou $n = 1000$ (cas 7 et 8), le nombre de secteurs est 9 (cas 5 et 6) ou 18 (cas 7 et 8). La loi de probabilité d'un point aléatoire (ρ, θ) est telle que sa densité soit arbitrairement petite au voisinage de la frontière du support S . Plus précisément, ρ et θ sont indépendants, θ suit la loi uniforme sur $[0, 2\pi]$ et ρ suit une loi de densité $(2\Phi(\theta) - 2x) \cdot \Phi^{-2}(\theta)$, $0 \leq x \leq \Phi(\theta)$.

On a choisi un nombre de secteurs plus élevé lorsque la densité ne s'annule pas sur la frontière de S parce que, intuitivement, il y a dans ce cas plus de points disponibles sur le bord du domaine.

Les figures ayant un numéro impair représentent l'estimateur de type histogramme Φ_n et celles de numéro pair, l'estimateur spline ψ_n . Le choix d'un nombre de secteurs bien adapté n'est pas plus simple que pour l'estimation de la densité. On voit en effet "à l'oeil nu" que, pour les deux types de lois, un nombre de secteurs trop grand donne une moins bonne estimation de Φ , même avec un nombre de points un peu plus grand.

Le problème de la réduction du biais n'est pas abordé. Les figures proposées montrent au moins qu'on ne pourra donner de résultat qu'en faisant des hypothèses précises sur la décroissance de la densité de M au voisinage de la frontière de S .

Enfin, il est à noter que le théorème 4.3.1. est vrai pour de nombreux autres modes d'approximations de l'estimateur de type histogramme ; en fait toutes les approximations $A_h f$ d'une fonction f qui satisfont des majorations similaires à celles des propositions 4.3.1 et 4.3.2. C'est le cas, par exemple, des quasi-interpolants $A_h f = \sum \beta_j B_j$ d'une fonction f connue sur une grille de pas h , où les β_j sont les valeurs de la fonction f en les points donnés et les B_j sont des B-splines d'un ordre donné [De Boor, 1978]. Quand la fonction f a un certain nombre de dérivées, ces quasi-interpolants ne donnent pas des ordres de convergence aussi élevés que les splines d'ordre supérieur.

Nous espérons revenir sur ces questions dans un prochain article.

Figure 1

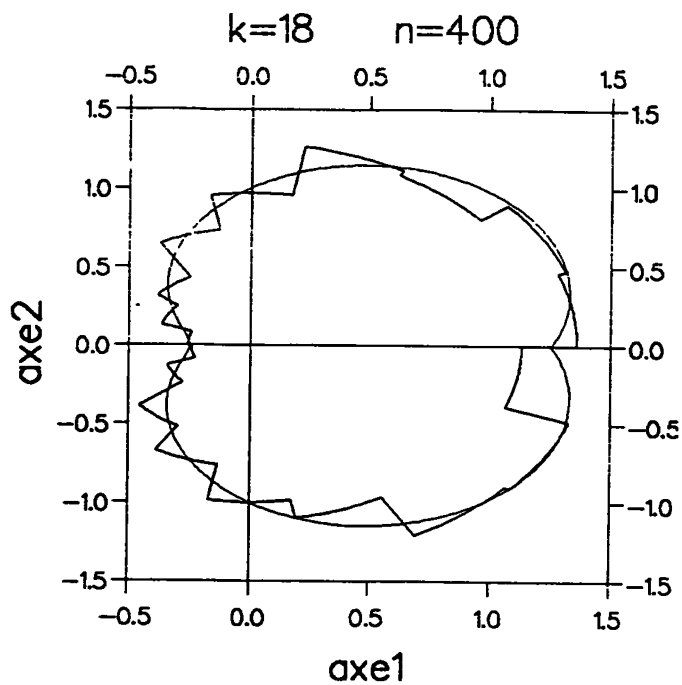


Figure 2

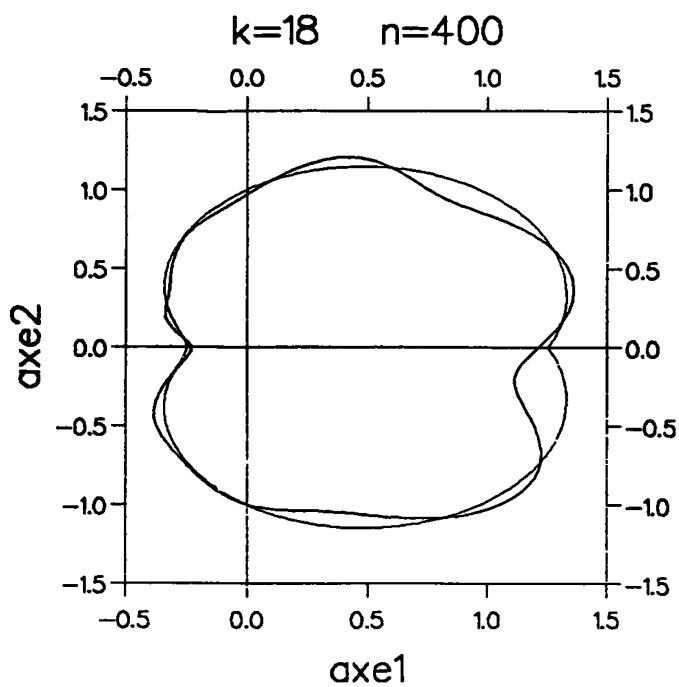


Figure 3

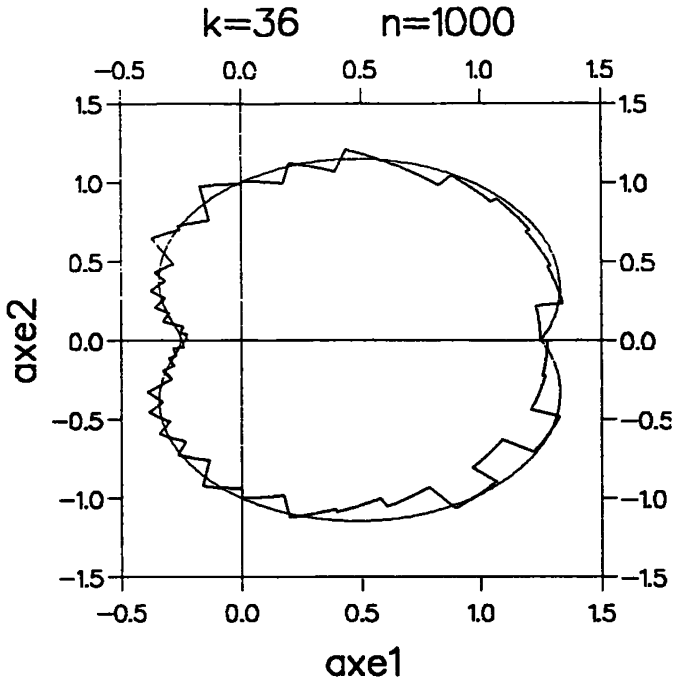


Figure 4

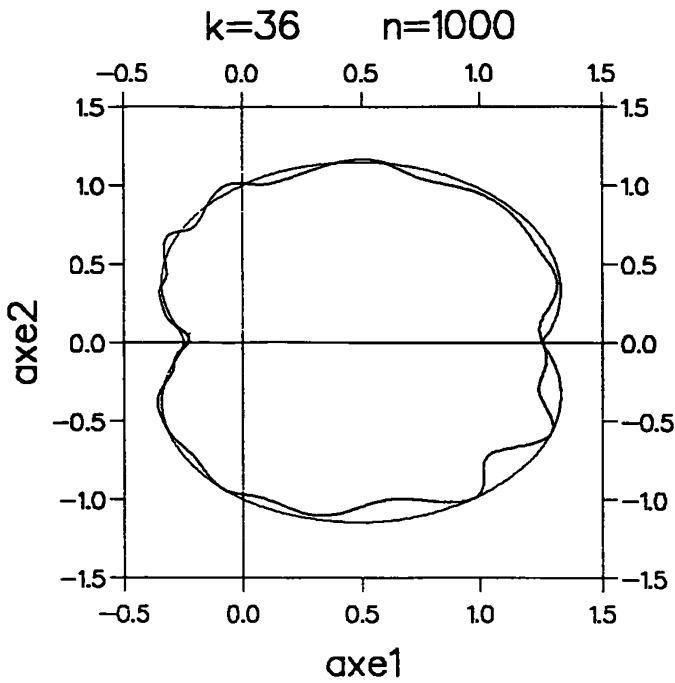


Figure 5

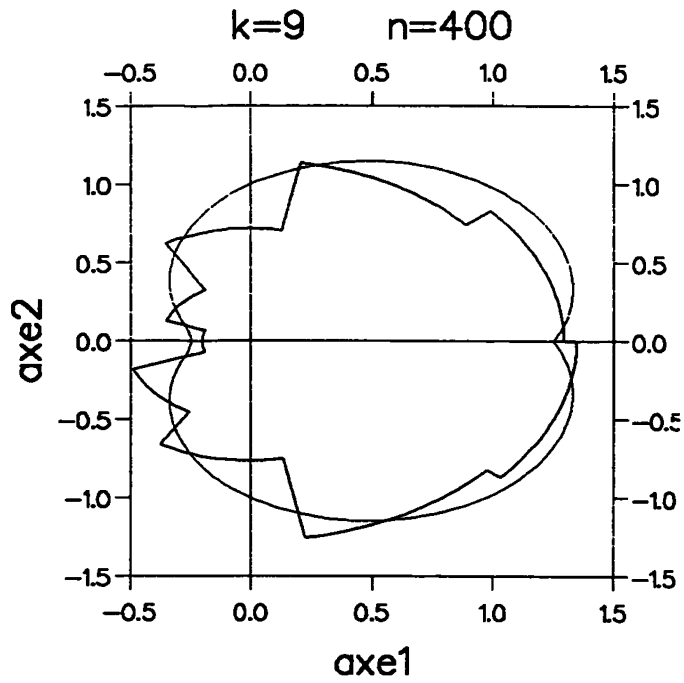
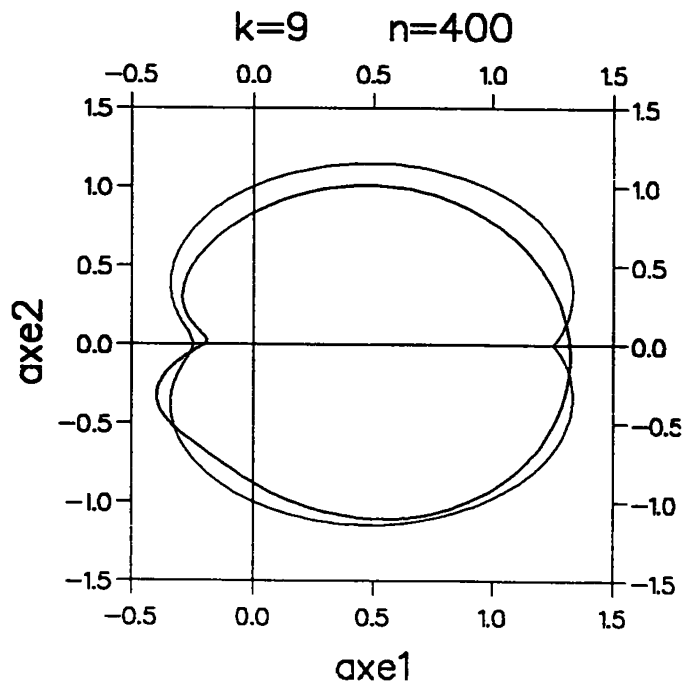
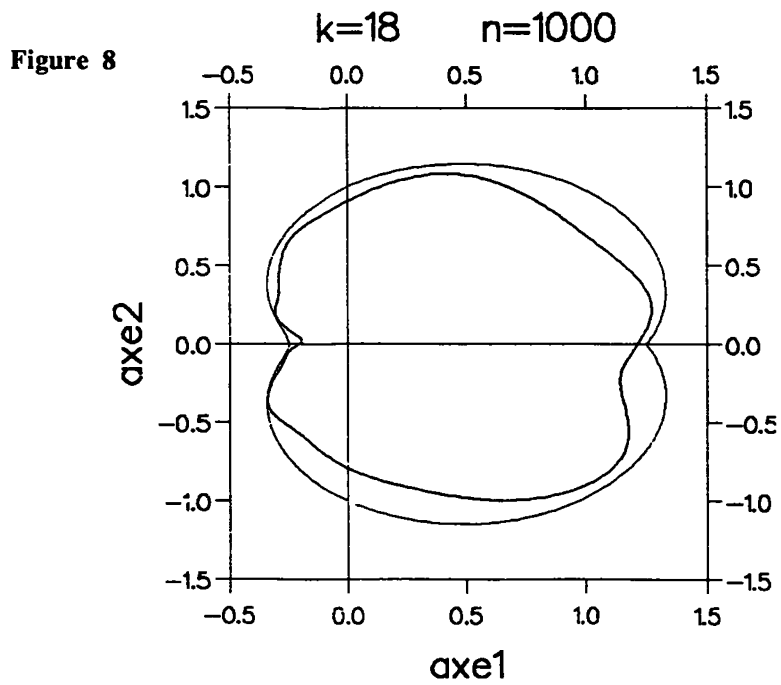
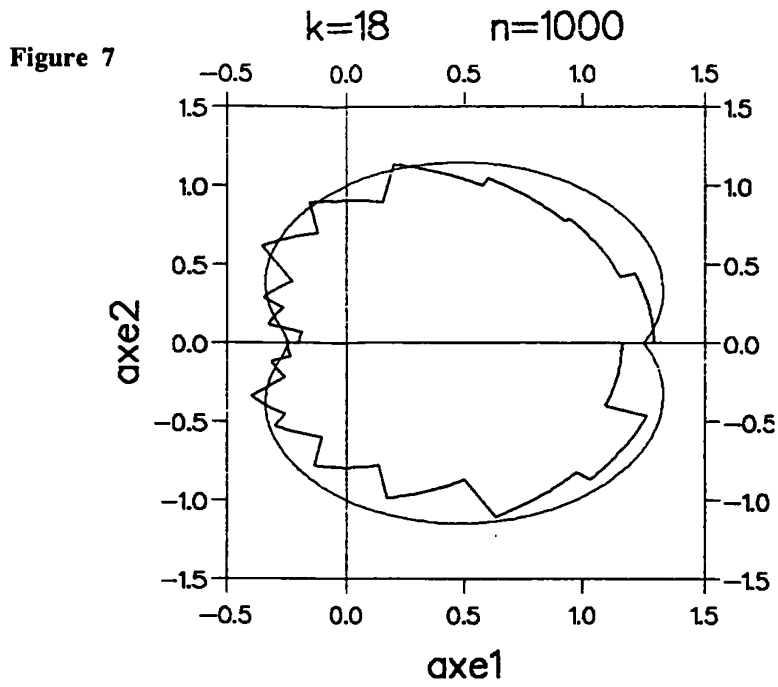


Figure 6





REFERENCES.

- Ahlberg, J.H., Nilson, E.N. et Walsh, J.L. (1977). The theory of splines and their applications, *Academic Press*.
- Berlinet, A. (1981). Convergence des estimateurs splines de la densité, *Publications de l'Institut de Statistique des Universités de Paris, Vol. 26, fasc. 2*.
- Bosq, D. (1977). Contribution à la théorie de l'estimation fonctionnelle. *Thèse, Université Pierre et Marie Curie, Paris*.
- Chevalier, J. (1976). Estimation du support et du contenu du support d'une loi de probabilité. *Annales Institut Henri Poincaré, section B, Vol. XII, 4*.
- De Boor, C. (1978). A practical guide to splines. *Applied Mathematical Science 27, Springer Verlag, New-York*.
- Efron, B. (1965). The convex hull of a random set of points. *Biometrika, 52, 331-343*.
- Fisher, L. (1966). The convex hull of a sample. *Bulletin of American Mathematical Society 72, 555-558*.
- Gensbittel, M.H. (1979). Contribution à l'étude statistique de répartitions ponctuelles aléatoires. *Thèse, Université Pierre et Marie Curie, Paris*.
- Geffroy, J. (1964). Sur un problème d'estimation géométrique. *Publications de l'Institut de Statistique des Universités de Paris, 13, 191-200*.
- Guilbart, C. (1973). Etude de la continuité de l'application H qui à toute mesure de probabilité définie sur \mathbb{R}^n fait correspondre l'enveloppe convexe fermée du support de cette mesure. *Compte Rendu de l'Académie des Sciences, Paris, 277, série A, 999-1002*.
- Kallenberg, O. Random measures. *Academic Press*.
- Karr, A.F. (1986). Point processes and their statistical inference. (*Dekker*).
- Kuelbs, J. (1976). A strong convergence theorem for Banach space valued random variables, *Annals of probability, 4, 744-771*.
- Jacob, P. (1984). Estimation du contour discontinu d'un processus ponctuel sur le plan. *Publications de l'Institut de Statistique des Universités de Paris, XXIX, 3-4, 1-25*.
- Jacob, P. et Abbar, H. (1989). Estimating the edge of a Cox process area. *Cahiers du Centre d'Etudes de Recherche Opérationnelle, Vol. 31, n° 3-4*.
- Raynaud, H. (1965). Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans \mathbb{R}^n . *Compte Rendu de l'Académie des Sciences, Paris, 262, série A, 623-627*.

Renyi, A. et Sulanke, R. (1964). Über die konvexe hülle von n zufällig gewählten punkten II. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 3, 138-147.

Ripley, B.D. et Rassin, J.P. (1977). Finding the edge of a Poisson forest. *Journal of Applied Probability*, 14, 483-491.