

55376  
1990  
21

55376  
1990  
21

N° d'ordre : 565

# THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE FLANDRES ARTOIS

pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

par

ABBAR Hassan



## **SUR DIVERSES MÉTHODES D'ESTIMATION DU CONTOUR D'UNE RÉPARTITION PONCTUELLE ALÉATOIRE.**

*Soutenue le 6 Juillet 1990 devant la Commission d'Examen :*

Président : J. GEFFROY, Université Pierre et Marie Curie

Rapporteur : R. MOCHÉ, Université de Lille

Examineurs : M. HALLIN, Université Libre de Bruxelles

P. JACOB, Université de Lille

M. ROUSSIGNOL, Université de Lille

*A mes parents,  
A mes frères et sœurs,  
A mes neveux et nièces,*

*Monsieur le Professeur Jean GEFROY nous fait le grand honneur de présider le jury de cette thèse. Nous tenons à lui exprimer notre respectueuse gratitude.*

*Monsieur le Professeur Pierre JACOB, m'a proposé l'idée originale de ce travail, ma reconnaissance envers lui est profonde pour l'aide efficace et les conseils qu'il n'a jamais cessé de me prodiguer.*

*Je tiens à remercier Messieurs les Professeurs Marc HALLIN, Raymond MOCHÉ et Michel ROUSSIGNOL, qui ont accepté de juger ce travail.*

*Ainsi que Madame Arlette Lengaigne, qui a dactylographié cette thèse avec patience et savoir faire, et toutes les personnes qui ont participé à sa réalisation matérielle.*

## PLAN

### INTRODUCTION.-

#### CHAPITRE I - RÉSULTATS SUR L'ESTIMATION DU SUPPORT D'UNE LOI DE PROBABILITÉ ET D'UN PROCESSUS PONCTUEL.

- 1 - Résultats dans le cas où le support est convexe.
  - 2 - Résultats dans le cas où le support est non convexe.
- Bibliographie.

#### CHAPITRE II - ESTIMATION DU CONTOUR DU SUPPORT D'UN PROCESSUS DE COX.

- 1 - Introduction.
  - 2 - Préliminaires.
  - 3 - Convergence de l'estimateur  $\Phi_n$ .
  - 4 - Loi limite de  $\Phi_n$ .
- Bibliographie.

#### CHAPITRE III - UN ESTIMATEUR SPLINE DU CONTOUR D'UNE RÉPARTITION PONCTUELLE ALÉATOIRE.

- 1 - Introduction.
  - 2 - Préliminaires.
  - 3 - Convergence de l'estimateur du type histogramme.
  - 4 - Construction et convergence de l'estimateur de type spline cubique  $\psi_n$ .
  - 5 - Simulations.
- Bibliographie.

#### CHAPITRE IV - ÉTUDE DE L'ERREUR QUADRATIQUE INTÉGRÉE ET CHOIX DE $k_n$ OPTIMAL.

- 1 - Introduction.
  - 2 - Préliminaires.
  - 3 - Cas où le support est un cercle.
  - 4 - Majoration asymptotique de l'erreur quadratique intégrée.
  - 5 - Exemples de choix de  $k_n$  optimaux.
  - 6 - Etude de l'erreur quadratique intégrée à distance finie dans un cas particulier.
- Bibliographie.

## INTRODUCTION

Dans plusieurs phénomènes concrets où on peut observer des masses ponctuelles aléatoires, il arrive qu'on connaisse déjà un point du support du processus ponctuel. Il est donc naturel de supposer que ce processus se réalise autour de ce point et de définir son support  $S$  en coordonnées polaires par :

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

$\Phi$  est une fonction continue est strictement positive à estimer ; notons que  $S$  est non nécessairement convexe, mais simplement convexe étoilé et que ce modèle est moins restrictif comparé à celui étudié par Geffroy et Gensbittel, puisqu'on suppose que le processus ponctuel  $N$  est un processus de Cox.

Dans le **chapitre I**, on présente des résultats antérieurs sur l'estimation du support d'une loi de probabilité et d'un processus ponctuel.

Dans le **chapitre II**, on suppose qu'il est donné  $n$  copies indépendantes  $N_1, \dots, N_n$  du processus de Cox  $N$  de support  $S$  défini en (1.1). Pour construire l'estimateur  $\Phi_n$  de type histogramme de  $\Phi$ , on partage le support  $S$  en  $k_n$  secteurs et l'estimateur  $\Phi_n$  est défini sur chaque secteur par le rayon du point le plus éloigné de la superposition  $N_1 + \dots + N_n$ . La construction d'un tel estimateur est inspirée d'un article de Geffroy et de travaux ultérieurs de Gensbittel. On établit d'abord une condition suffisante sur  $k_n$  pour que l'estimateur  $\Phi_n$  converge uniformément et presque sûrement vers  $\Phi$ . Cette étude de la convergence est abordée avec un léger progrès car on suppose que l'intensité du processus ponctuel  $N$  peut-être arbitrairement petite au voisinage de la frontière du support  $S$ . Sous certaines conditions que doit vérifier l'intensité du processus ponctuel, on établit que la loi limite de  $\Phi_n$  est une loi de Cambel dont le paramètre dépend de l'intensité du processus ponctuel et de  $\Phi$ .

Dans le **chapitre III**, on aborde le lissage de l'estimateur  $\Phi_n$  par une fonction spline cubique périodique et aléatoire  $\psi_n$  qui l'interpole au milieu de chaque secteur. L'existence et l'unicité de cette fonction  $\psi_n$  sont établies à l'aide de

méthodes voisines de celles utilisées par Berlinet dans l'estimation de la densité. La convergence presque sûre et uniforme de l'estimateur de type spline cubique  $\psi_n$  est établie sous les mêmes conditions que dans le chapitre II. Enfin, on présente quelques expériences de simulation qui montrent que le choix d'un nombre de secteurs bien adapté n'est pas plus simple que pour l'estimation de la densité et qu'on ne pourra donner de résultat qu'en faisant des hypothèses sur la décroissance de la densité de la mesure aléatoire qui dirige le processus ponctuel de Cox  $N$  au voisinage de la frontière de  $S$ .

Dans le **Chapitre IV**, on utilise un moyen classique d'optimisation des estimateurs fonctionnels, la minimisation de l'erreur quadratique intégrée, pour régler au mieux le choix optimal du nombre de secteurs de  $S$ . Pour rassurer sur la validité de la méthode, on commence par étudier un exemple simple et explicite dans lequel on suppose que le support du processus ponctuel est un cercle. Dans ce cas, en minimisant l'erreur quadratique intégrée, on trouve que le meilleur estimateur de type histogramme  $\Phi_n$  est obtenu pour  $k_n = 1$  : c'est-à-dire la plus petite circonférence enveloppant le nuage de points. Dans le cas général, on établit une majoration asymptotique de l'erreur quadratique intégrée et on présente quelques exemples de choix de  $k_n$  optimal. A la fin de ce chapitre et dans le but d'étudier le comportement du  $k_n$  optimal à distance finie et la robustesse de l'estimateur  $\Phi_n$ , on calcule l'erreur quadratique intégrée à distance finie dans un cas particulier et on étudie numériquement cette erreur quadratique intégrée.

Dans ce travail, on n'a pas abordé le problème de la réduction du biais et l'étude des vitesses de convergence des estimateurs  $\Phi_n$  et  $\psi_n$ . Mais nous comptons revenir sur ces questions dans un prochain travail.

## CHAPITRE I

RÉSULTATS SUR L'ESTIMATION DU SUPPORT D'UNE LOI DE  
PROBABILITÉ ET D'UN PROCESSUS PONCTUEL

Ce chapitre est consacré à la présentation de quelques résultats sur l'estimation du support d'une loi de probabilité et d'un processus ponctuel.

Dans le paragraphe 1, ces résultats sont présentés dans le cas où le support est supposé convexe. Par contre, dans le paragraphe 2 les résultats sont donnés dans le cas où le support est non convexe.

## 1 - Résultats dans le cas où le support est convexe.

On sait qu'en dimension 1, l'estimateur du support  $[a, b]$  d'une loi uniforme à partir d'un échantillon  $x = \{x_1, \dots, x_n\}$  repose sur la statistique exhaustive  $(\inf_i x_i, \sup_i x_i)$ .

En dimension supérieure à 1 et par analogie avec le problème en dimension 1, on peut supposer que  $x = \{x_1, \dots, x_n\}$  est un échantillon de la loi uniforme sur un convexe compact  $S$  et baser la construction de l'estimateur sur l'enveloppe convexe de l'échantillon qui est exhaustif. De nombreux auteurs se sont intéressés au calcul de l'espérance mathématique du nombre de sommets, du nombre d'arêtes, du nombre de faces, du périmètre et de la surface de cette enveloppe convexe.

Les premiers travaux sur ce problème remontent à Renyi et Sulanke [6] qui ont établi les résultats suivants :

**Théorème 1.1.-** *Soit  $(P_i)_{1 \leq i \leq n}$  un échantillon de  $n$  points qui obéissent à la même loi de probabilité de support convexe et compact  $K$  de  $\mathbf{R}^2$ . Soit  $L$  le périmètre et  $F$  la surface de  $K$ . Si on désigne par  $L_n$  le périmètre et par  $F_n$  la surface de l'enveloppe convexe  $H_n$  de l'échantillon alors :*

$$E(L_n) = L - \frac{\Gamma(\frac{2}{3})(12F)^{2/3} \int_0^L k^{4/3}(s) ds}{12 n^{2/3}} + o(n^{-1})$$

et

$$E(F_n) = F - \frac{\Gamma(\frac{8}{3})(12F)^{2/3} \int_0^L k^{1/3}(s) ds}{10n^{2/3}} + o(n^{-1})$$

**Théorème 1.2.-** Dans le cas où  $K$  est un carré de côté  $a$  les espérances mathématiques  $E(L_n)$  et  $E(F_n)$  sont données par :

$$E(L_n) = 4a - \frac{a(2-q)\sqrt{2\pi}}{\sqrt{n}} + o\left(\frac{1}{n}\right)$$

avec  $q = \int_0^1 \frac{(\sqrt{1+\tau^2}-1)}{\tau^{3/2}} d\tau$

et 
$$E(F_n) = a^2 - \frac{8a^2 \ell_n(n)}{3n} + o\left(\frac{1}{n}\right).$$

D'après ces formules, on constate que  $L_n$  et  $F_n$  sont des estimateurs asymptotiquement sans biais. Mais à distance finie ils sont biaisés et que le biais dépend du support  $K$ .

On trouve un point de vue similaire dans un article de Efron [2] (1965). Cet auteur est parti d'un échantillon  $W_1, W_2, \dots, W_N$  de  $N$  points tirés selon une loi de probabilité  $\Gamma$  sur  $\mathbf{R}^2$  et il a démontré les résultats suivants : si on suppose que  $\Gamma$  est une loi à symétrie sphérique et si on désigne par  $g_2$  et  $G_2$  respectivement la densité et la fonction de répartition marginales selon la direction de l'axe des abscisses, alors on a :

**Théorème 1.3.-** Si on désigne par :

- $V_N$  le nombre de sommets ou d'arêtes de l'enveloppe convexe  $H_N$  de l'échantillon.
- $P_N$  le périmètre de  $H_N$ .
- $A_N$  la surface de  $H_N$ .
- $(X_1, Y_1)$  et  $(X_2, Y_2)$  les coordonnées des points aléatoires  $W_1$  et  $W_2$ .

alors on a :

$$E(V_N) = 2\pi \binom{N}{2} \int_{-\infty}^{\infty} E[|Y_2 - Y_1| \mid X_1 = p, X_2 = p] G_2^{N-2}(p) g_2^2(p) dp$$

$$E(P_N) = 2\pi \binom{N}{2} \int_{-\infty}^{\infty} E[(Y_2 - Y_1)^2 \mid X_1 = p, X_2 = p] G_2^{N-2}(p) g_2^2(p) dp$$

$$E(A_N) = \pi \binom{N}{2} \int_{-\infty}^{\infty} p E[(Y_2 - Y_1)^2 \mid X_1 = p, X_2 = p] G_2^{N-2}(p) g_2^2(p) dp$$

Les formules du théorème 1.3 sont particulièrement simples dans les cas suivants :

**1<sup>er</sup> cas :** Si  $\Gamma$  admet une densité circulaire et normale :  $g(x, y) = \frac{1}{2\pi} \exp\{-\frac{1}{2}(x^2 + y^2)\}$  alors :

$$E[|Y_2 - Y_1| \mid X_1 = p, X_2 = p] = \frac{2}{\sqrt{\pi}}$$

$$E[(Y_2 - Y_1)^2 \mid X_1 = p, X_2 = p] = 2$$

Ces quantités ne dépendent pas de  $p$  donc, si on pose  $\varphi(x) = (2\pi)^{-1/2} \exp\{-\frac{1}{2}x^2\}$  et  $\Phi(x) = \int_{-\infty}^x \varphi(y)dy$ , on a :

$$E(V_N) = 4\sqrt{\pi} \binom{N}{2} \int_{-\infty}^{\infty} \Phi^{N-2}(p) \varphi^2(p) dp$$

$$E(P_N) = 4\pi \binom{N}{2} \int_{-\infty}^{\infty} \Phi^{N-2}(p) \varphi^2(p) dp$$

$$E(A_N) = 4\pi \binom{N}{2} \int_{-\infty}^{\infty} p \Phi^{N-2}(p) \varphi^2(p) dp$$

**2<sup>ième</sup> cas :** Si  $\Gamma$  est uniforme à l'intérieur du cercle de rayon 1 alors :

$$g_2(p) = \begin{cases} \frac{2}{\pi} \sqrt{1-p^2} & (-1 \leq p \leq 1) \\ 0 & \text{ailleurs} \end{cases}$$

$$G_2(p) = \begin{cases} \frac{1}{2} + \frac{1}{\pi}(p\sqrt{1-p^2} + \sin^{-1}(p)) & \text{si } 0 \leq p \leq 1 \\ 1 & \text{si } p \geq 1 \\ 1 - G_2(-p) & \text{si } p < 0 \end{cases}$$

et par conséquent on a :

$$E[|Y_2 - Y_1| \mid X_1 = p, X_2 = p] = \frac{\pi}{3} g_2(p)$$

$$E[(Y_2 - Y_1)^2 \mid X_1 = p, X_2 = p] = \frac{\pi^2}{6} g_2^2(p)$$

si on remplace ces quantités dans les formules du théorème 1.3., on obtient :

$$E(V_N) = \frac{2\pi^2}{3} \binom{N}{2} \int_{-1}^1 G_2^{N-2}(p) g_2^3(p) dp$$

$$E(P_N) = \frac{\pi^3}{3} \binom{N}{2} \int_{-1}^1 G_2^{N-2}(p) g_2^4(p) dp$$

$$E(A_N) = \frac{\pi^3}{6} \binom{N}{2} \int_{-1}^1 p G_2^{N-2}(p) g_2^4(p) dp$$

Il est à noter qu'en dimension 3, Efron a établi des résultats similaires.

En 1965, Raynaud [5] a étudié dans  $\mathbf{R}^n$  le problème général suivant : étant donné un nuage  $D_N^n$  de  $N$  points pris au hasard dans  $\mathbf{R}^n$ , les  $N$  choix aléatoires sont indépendants les uns des autres et sont effectués suivant une même loi  $\mathcal{L}$ , comment se comporte l'espérance mathématique  $E_N^n$  du nombre des points du polygone d'appui ou de l'enveloppe convexe  $H_N^n$  de ce nuage quand  $N$  tend vers l'infini ?

On appelle polygone d'appui de  $D_N^n$  le plus petit polygone convexe contenant tout le nuage. Un point  $M$  du nuage sera dit sommet du polygone si et seulement si il existe au moins  $(n - 1)$  autres points du nuage tels que, si l'on fait passer un hyperplan par  $M$  et par ces  $(n - 1)$  autres points du nuage, les  $(N - n)$  points du nuage restants se trouvent tous d'un même côté de l'hyperplan. Pour répondre à la question précédente, Raynaud a établi les résultats suivants :

**Théorème 1.4.-** *Si le tirage des points se fait suivant une probabilité uniforme dans la boule de rayon unité, alors :*

$$E_N^n \sim \frac{2}{n!} \left( \frac{B_{n^2-1}}{B_{n-1}} \right)^{\frac{n^2+1}{n+1}} \Gamma\left(\frac{n^2+1}{n+1}\right) (n+1)^{\frac{n(n-1)}{n+1}} N^{\frac{n-1}{n+1}} \quad \text{quand } N \rightarrow \infty$$

où  $B_p = 2 \frac{\pi^{\frac{p}{2}}}{p \Gamma(\frac{p}{2})}$ .

Ce résultat se généralise sous la forme suivante :

**Théorème 1.5.-** *Soit  $\mathcal{F}$  un domaine fermé, strictement convexe, délimité par une variété deux fois continûment différentiable plongée dans  $\mathbf{R}^n$ , définie en tout point par la donnée de ses rayons de courbures principaux, si :*

- $\Sigma$  désigne cette variété différentiable (d'élément  $d\sigma$ )
- $F$  la mesure de Lebesgue à l'intérieur de cette variété
- $R_1, \dots, R_n$  les rayons de courbure principaux en un point  $M$ , ( $\forall i, \forall M, R_i(M) > 0$ )

le tirage au hasard des  $N$  points se fait d'une façon uniforme, à l'intérieur de cette variété, on a :

$$E_N^n \sim K_n \left( \frac{N}{F} \right)^{\frac{n-1}{n+1}} \int_{\Sigma} \frac{d\sigma}{(\prod_i R_i)^{\frac{n+1}{n+1}}} \quad \text{quand } N \rightarrow \infty$$

avec

$$K_n = \Gamma\left(\frac{n^2 + 1}{n + 1}\right) \frac{\Gamma\left(\frac{n^2}{2}\right)}{\Gamma\left(\frac{n^2-1}{2}\right)} \frac{[\Gamma\left(\frac{n-1}{2}\right)]^{\frac{n^2+1}{n+1}}}{[\Gamma\left(\frac{n}{2}\right)]^n} 2^{\frac{n-1}{n+1}} \pi^{\frac{n^2-3n}{2(n+1)}} \frac{(n^2 - 1)^{\frac{n^2-2n-1}{n+1}} (n-1)^{-\frac{n-1}{n+1}}}{n! n^{n-2}}$$

**Théorème 1.6.-** Si la loi de probabilité  $\mathcal{L}$  est une loi normale, alors on a :

$$E_N^n \sim 2^n n^{-\frac{1}{2}} (\text{Log } N)^{\frac{n-1}{2}} \text{ quand } N \rightarrow \infty.$$

Ce résultat est valable en particulier pour une loi normale sphérique (c'est-à-dire avec une distribution de la forme  $\pi^{-(\frac{n}{2})} e^{-\frac{\sum x_i^2}{a}}$ ) et nous permet d'en déduire le

**Théorème 1.7.-** La distribution des sommets du polygone d'appui d'un nuage de  $N$  points tirés au hasard dans  $\mathbf{R}^n$ , indépendamment les uns des autres et suivant la même loi de Gauss :

$$(a\pi)^{-\frac{n}{2}} e^{-\frac{\sum x_i^2}{a}}$$

converge - quand  $N$  tend vers l'infini - vers un processus ponctuel de Poisson homogène sur la sphère de centre l'origine et de rayon  $\sqrt{a \text{Log } N}$  et de densité moyenne finie

$$d = \frac{(n\pi)^{-\frac{1}{2}} 2^{n-1}}{a^{\frac{n-1}{2}} \pi\left(\frac{n}{2}\right)} \text{ points par unité de surface.}$$

En 1977, Ripley et Rasson [7] ont étudié le problème suivant : soit  $m$  la mesure de Lebesgue sur  $\mathbf{R}^2$ ,  $\mathcal{D}$  l'ensemble des domaines compacts et convexes de  $\mathbf{R}^2$  et de mesures positives et  $x_1, \dots, x_N$  la réalisation d'un processus de Poisson homogène sur un domaine  $D \in \mathcal{D}$ . Comment estimer  $D$  ?

Puisque le processus est homogène alors ; conditionnellement à  $N$  ;  $x_1, \dots, x_N$  sont distribués d'une façon indépendante et uniforme.

Si on pose  $x = (x_1, \dots, x_N)$  et si on désigne par  $H(x)$  l'enveloppe convexe du nuage des points, alors  $H(x)$  est une statistique exhaustive et un estimateur du maximum de vraisemblance de  $D$  dans le modèle paramétrique  $\{P_D ; D \in \mathcal{D}\}$  car :  $P_D(x) = \mathbf{1}_D(x) \cdot m(D)^{-N} = \mathbf{1}_D(H(x)) \cdot m(D)^{-N}$ , mais  $H(x)$  est un estimateur biaisé.

Pour estimer  $D$ , Ripley et Rasson ont établi les résultats suivants : soit  $g(D)$  le centre de gravité de  $D$  pour la mesure  $m$  et  $s(D) = D - g(D)$  le domaine centré.

Soit  $V_{N+1}$  (resp.  $V_N$ ) le nombre de sommets de l'enveloppe convexe du nuage  $x_1, \dots, x_{N+1}$  (resp. du nuage  $x_1, \dots, x_N$ ).

Pour estimer  $s(D)$  on a :

**Théorème 1.8.-** *Le domaine centré  $s(D)$  est estimé par  $C \cdot s(H(x))$  avec :*

- si  $m(D)$  est connue  $C = \left[ \frac{m(D)}{m(H(x))} \right]^{\frac{1}{2}}$ .
- si  $m(D)$  est inconnue alors un estimateur sans biais de  $m(D)$  est  $m(H(x))/[1 - E(V_{N+1})/N + 1]$  avec  $E(V_{N+1})/N + 1$  qui est estimé avec un léger biais par  $\frac{V_N}{N}$  et  $C$  est estimée par  $[1/[1 - E(V_{N+1})/N + 1]]^{1/2}$ .

$g(D)$  est estimé par  $g(H(x))$  et finalement l'estimateur proposé par Ripley et Rasson de  $D$  est :

$$g(H(x)) + C \cdot s(H(x))$$

où  $C$  est déterminée par le théorème 1.8.

L'estimateur proposé par Ripley et Rasson est une dilatation homothétique de l'enveloppe convexe de l'échantillon autour de son centre de gravité.

En estimation  $D$  de cette façon, Ripley et Rasson récupère une sorte de non biais car  $E[m(C \cdot s(H(x)))] = m(D)$ .

*Justification :*  $P\{x_{N+1}$  est un point extremum de  $H(x_1, \dots, x_{N+1})/x\} = 1 - \frac{m(H(x))}{m(D)}$  donc  $P\{x_{N+1}$  est un point extremum $\} = 1 - \frac{E[m(H(x))]}{m(D)}$  et par conséquent :

$$E(V_{N+1}) = \sum_{i=1}^{N+1} P\{x_i \text{ est un point extremum}\} = (N+1) \left[ 1 - \frac{E[m(H(x))]}{m(D)} \right]$$

d'où :

$$\begin{aligned} E[m(C \cdot s(H(x)))] &= C^2 E[m(H(x))] \\ &= [1/[1 - E(V_{N+1})/N + 1]] E[m(H(x))] = m(D) \end{aligned}$$

## 2 - Résultats dans le cas où le support est non convexe.

En 1964 et simultanément avec Renyi et Sulanke, Geffroy [4] a étudié le problème suivant : soit  $(X_i, Y_i)_{1 \leq i \leq n}$  un échantillon de  $n$  points qui obéissent à la

même loi de probabilité de support  $D = \{(x, y) \in \mathbf{R}^2 / 0 \leq x \leq 1 ; 0 \leq y < f(x)\}$  où  $f$  est une fonction continue à estimer.

Cet auteur a construit un estimateur de  $f$  de la façon suivante : si on partage  $D$  en  $k_n$  sous domaines  $D_{k_n,r}$  ( $r = 1, \dots, k_n$ ) par des parallèles à  $oy$  équidistantes avec  $(k_n)$  une suite d'entiers positifs qui tend vers l'infini avec  $n$ . Si on désigne par  $I_{n,r}$  l'ensemble des indices des points de l'échantillon qui appartiennent à  $D_{k_n,r}$  et si on pose :

$$\begin{aligned} U_{n,r} &= \sup_{i \in I_{n,r}} (Y_i) & \text{si } I_{n,r} \neq \emptyset \\ U_{n,r} &= 0 & \text{si } I_{n,r} = \emptyset \end{aligned}$$

alors l'estimateur  $f_n$  de  $f$  est défini par :

$$\begin{aligned} f_n(x) &= U_{n,r} \quad \text{pour } x \in \left[ \frac{r-1}{k_n}, \frac{r}{k_n} \right[ \quad (r = 1, \dots, k_n) \\ f_n(1) &= U_{n,k_n}. \end{aligned}$$

Pour l'efficacité de cette estimation dans la topologie de la convergence uniforme sur  $[0, 1]$ , Geffroy a établi les résultats suivants : si on désigne par  $\varphi(x, y)$  la densité de la loi de probabilité par rapport à la mesure de Lebesgue de  $\mathbf{R}^2$  alors on a :

**Théorème 2.1.-** *S'il existe deux constantes positives  $A$  et  $B$  telles que :*

$$A \leq \varphi(x, y) \leq B, \quad \forall (x, y) \in D,$$

*une condition nécessaire et suffisante pour que  $f_n$  converge uniformément vers  $f$  en probabilité est que :*

$$k_n = o(n / \text{Log } n) \quad \text{quand } n \rightarrow \infty.$$

**Théorème 2.2.-** *Lorsque  $f_n$  converge uniformément en probabilité vers  $f$ , cette convergence est presque complète.*

Dans l'hypothèse où  $d(f_n, f)$  converge vers 0, Geffroy a cherché la loi limite de  $d(f_n, f)$  en procédant par ordre de complexité croissante et en examinant les trois cas suivants :

**1<sup>er</sup> cas :**  $f$  et  $\varphi$  constantes.

Si on pose  $Z_n = \sup_{1 \leq r \leq k_n} (M_{k_n,r} - U_{n,r})$  où  $M_{k_n,r}$  est le maximum de  $f(x)$  sur  $\left[ \frac{r-1}{k_n}, \frac{r}{k_n} \right[$  alors on a :

**Théorème 2.3.-** Supposons que l'on ait  $\varphi(x, y) \equiv c$  et  $f(x) \equiv \frac{1}{c}$ , où  $c$  est une constante positive, et que  $k_n = o(n/\text{Log } n)$ . Dans ces conditions, la variable aléatoire  $Z'_n = \frac{n}{k_n} Z_n - \frac{\text{Log } k_n}{c}$  admet une loi-limite qui est la loi de Gumbel ayant pour fonction de répartition  $\exp(-e^{-cz})$ .

Dans ce premier cas  $Z_n = d(f_n, f)$  donc  $\frac{n}{k_n} d(f_n, f) - \frac{\text{Log } k_n}{c}$  admet pour loi-limite la loi de Gumbel de *f.d.r.*  $\exp(-e^{cz})$ .

2<sup>ième</sup> cas :  $f$  variable,  $\varphi$  constante.

Dans ce cas on a :

**Théorème 2.4.-** Supposons que l'on ait  $\varphi(x, y) = c$  (constante)  $\forall (x, y) \in D$ , que  $f$  soit lipschitzienne d'ordre  $\gamma$ , et que  $k_n = o(n/\text{Log } n)$ . Dans ces conditions, si l'on a  $n^{1/1+\gamma} = o(k_n)$ , la variable aléatoire  $Z'_n = \frac{n}{k_n} Z_n - \frac{\text{Log } k_n}{c}$  admet une loi-limite qui est la loi de Gumbel ayant pour fonction de répartition  $\exp(-e^{-cz})$ .

Dans ce deuxième cas  $P\{d(f_n, f) = Z_n\} \rightarrow 1$  quand  $n \rightarrow \infty$  donc la variable aléatoire  $\frac{n}{k_n} d(f_n, f) - \frac{\text{Log } k_n}{c}$  admet pour loi-limite la loi de Gumbel de *f.d.r.*  $\exp(-e^{-cz})$ .

3<sup>ième</sup> cas :  $f$  et  $\varphi$  variables.

Dans ce cas on a :

**Théorème 2.5.-** Supposons que  $f$  et  $\varphi$  soient respectivement lipschitziennes d'ordre  $\gamma$  et  $\delta$ , et qu'il existe deux constantes positives  $A, B$  telles que  $A \leq \varphi(x, y) \leq B \forall (x, y) \in D$ . Si  $k_n$  vérifie  $k_n = o[n/(\text{Log } n)^{1+1/\delta}]$  et  $n^{1/1+\gamma} = o(k_n)$  la variable  $Z_n$  admet une loi-limite du type Gumbel.

Plus précisément, considérons la variable  $Z'_n$  définie par :

$$Z_n = \alpha_n Z'_n + \beta_n$$

où  $\alpha_n = \frac{k_n}{n}$  et  $\beta_n$  est la solution de l'équation :

$$\int_0^1 \exp[\text{Log } k_n - \frac{n}{k_n} \varphi(x, f(x)) \beta_n] dx = 1,$$

quand  $n$  croît indéfiniment,  $Z'_n$  converge vers la variable aléatoire de Gumbel ayant pour fonction de répartition  $\exp(-e^{-cz})$ , où  $c = \inf_{0 \leq x \leq 1} \varphi(x, f(x))$ .

Dans ce troisième cas, on a aussi  $P\{d(f_n, f) = Z_n\} \rightarrow 1$  quand  $n \rightarrow \infty$  donc la variable aléatoire  $\frac{n}{k_n}[d(f, f_n) - \beta_n]$  admet pour loi-limite la loi de Gumbel de *f.d.r.*  $\exp(-e^{-cz})$ .

Tous ces résultats ont été prolongés en 1979 par Gensbittel [3] au cas d'un processus ponctuel de Poisson. **Le travail que nous présentons dans les chapitres qui suivent trouve ici son origine.**

En 1976, Chevalier [1] a établi le résultat suivant : soit  $(\Omega, d)$  un espace métrique,  $\mathcal{A}$  la tribu borélienne associée,  $P$  une probabilité sur  $(\Omega, \mathcal{A})$  de support  $K$  et  $P_*$  la probabilité inférieure associée à  $P$ . Soit  $(X_i)_{1 \leq i \leq n}$  un échantillon d'une variable aléatoire  $X$  de loi  $P$ .

Si on suppose que :

- $\Omega$  est localement compact à base dénombrable,
- tout ensemble borné de  $(\Omega, d)$  est relativement compact,
- pour tout  $x \in \Omega$  et tout  $\varepsilon > 0$ , on a :  $P\{\{M \mid d(x, M) = \varepsilon\}\} = 0$  c'est-à-dire que  $P$  ne charge pas les frontières de la boule de centre  $x$  et de rayon  $\varepsilon$ .
- $K$  est  $\varepsilon$ -convexe c'est-à-dire que  $K$  est égal à son enveloppe  $\varepsilon$ -convexe  $K(\varepsilon) = (((K^\varepsilon)^c)^\varepsilon)^c$  (où  $B^\varepsilon = \{x/d(x, B) < \varepsilon\}$ ).

Si on pose pour tout  $\varepsilon > 0$  et fixé :

$$H'_n = C\left(\bigcup_{x \in I_n} B(x, \varepsilon)\right), \text{ avec}$$

$$I_n = \{x \mid \forall j = 1, \dots, n, X_j \notin B(x, \varepsilon)\}$$

alors on a :

**Théorème 2.6.-** *L'estimateur  $H'_n$  converge p.co. vers  $K$  c'est-à-dire :*

$$\sum_n (1 - P_*(K_{-\varepsilon} \subset H_n \subset K^\varepsilon)) < +\infty \text{ où } K_{-\varepsilon} = \{x/d(x, \bar{K}) \geq \varepsilon\}.$$

Pour estimer des supports non  $\varepsilon$ -convexes, Chevalier a utilisé des sous-ensembles de  $\Omega$  dont le diamètre tend vers 0 avec  $\frac{1}{n}$  et il a construit des estimateurs du support  $K$  par deux méthodes que Gensbittel a prolongé en 1979 au cas d'un processus ponctuel et que l'on va présenter maintenant.

Soit  $N$  un processus ponctuel sur un espace métrique séparable  $(X, d)$ ,  $\lambda$  la fonction d'ensemble sous-additive :  $\lambda(A) = P(N(A) > 0)$  et  $S = \{x \in X \mid$

$\forall A$  ouvert,  $x \in A, \lambda(A) > 0$  } le support de  $N$ . Soit  $N^n = N_1 + \dots + N_n$  une superposition de processus ponctuels indépendants et de même loi que  $N$ .

On dit que  $N$  est à accroissements anticorrélés si, pour tout entier positif  $k$  et pour toute famille  $(A_1, \dots, A_k)$  de boréliens bornés deux à deux disjoints on a :  $P(\bigcap_{j=1}^k \{N(A_j) > 0\}) \leq \prod_{j=1}^k \lambda(A_j)$ .

On dit que  $N$  possède la propriété (H) si, pour tout  $n \in \mathbf{N}^*$ , la superposition  $N^n$  est à accroissements anticorrélés.

– **Méthode des partitions. Cas d'un support compact.**

On suppose ici que  $S$  est compact et que l'on dispose, pour tout entier positif  $n$ , d'une partition  $(X_{n,i})_{i \in F_n}$  de  $X$ , où  $F_n$  est un ensemble dénombrable.

Pour tout  $n \in \mathbf{N}^*$  et pour tout  $i \in F_n$ , on pose :

$$\begin{aligned} \delta_{n,i} &= \text{diamètre de } X_{n,i} \\ \delta_n &= \sup_{i \in F_n} \delta_{n,i} \text{ on supposera que : } \lim_n \delta_n = 0 \end{aligned}$$

Pour tout  $n \in \mathbf{N}^*$ , on pose :

$$\begin{aligned} m(n) &= \inf_{X_{n,i} \subset S} \lambda(X_{n,i}) \\ M(n) &= \sup_{X_{n,i} \subset S} \lambda(X_{n,i}) \\ I_n &= \{i \in F_n / \exists j, 1 \leq j \leq n : N_j(X_{n,i}) > 0\} \\ H_n &= \bigcup_{i \in I_n} X_{n,i} \end{aligned}$$

Gensbittel a établi les résultats suivants :

**Théorème 2.7.-** *Si  $S$  est d'intérieur non vide et si  $N$  vérifie la propriété (H), une condition nécessaire pour que  $H_n$  converge vers  $S$  en probabilité :  $\forall \varepsilon > 0, \lim_n P_*(S_{-\varepsilon} \subset H_n \subset S^\varepsilon) = 1$  (resp. presque complètement :  $\forall \varepsilon > 0, \sum_{n=1}^\infty (1 - P_*(S_{-\varepsilon} \subset H_n \subset S^\varepsilon)) < +\infty$ ) est que la suite  $W_n = (1 - M(n))^n / M(n)$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $W_n$  soit convergente).*

**Théorème 2.8.-** *Si le processus ponctuel  $N$  possède une mesure moyenne  $\mu$  définie par : pour tout borélien  $A$  inclus dans  $S$ ,  $\mu(A) = E(N(A))$ , une condition suffisante pour que  $H_n$  converge vers  $S$  en probabilité (resp. presque complètement) est que la suite  $W'_n = (1 - m(n))^n / m(n)$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $W'_n$  soit convergente).*

– Méthode des partitions. Cas général.

Dans ce cas où  $S$  n'est pas compact, on suppose que  $X$  est séquentiellement compact, et on désigne par  $K = (K_n)_{n \in \mathbf{N}^*}$  une suite croissante de compacts tendant vers  $X$ .

On pose de même, pour tout  $n$  et pour tout  $i \in F_n$  :

$$\delta_{n,i} = \text{diamètre de } X_{n,i}$$

$$\delta_n = \sup_{i \in F_n} \delta_{n,i}$$

on pose pour tout  $n \in \mathbf{N}^*$  :

$$\begin{aligned} S_{(n)} &= S \cap K_n \\ m(n) &= \inf_{X_{n,i} \subset S_{(n)}} \lambda(X_{n,i}) \\ M(n) &= \sup_{X_{n,i} \subset S_{(n)}} \lambda(X_{n,i}) \\ I_n &= \{i \in F_n / \exists j, 1 \leq j \leq n : N_j(X_{n,i}) > 0\} \\ H_n &= \bigcup_{i \in I_n} X_{n,i} \quad H'_n = H_n \cap K_n \end{aligned}$$

Gensbittel a établi les résultats suivants :

**Théorème 2.9.-** *Si  $N$  vérifie (H) et si  $S$  est d'intérieur non-vide, une condition nécessaire pour que  $H_n$  converge vers  $S$  suivant  $K$  en probabilité (resp. presque complètement) est que la suite  $W_n = (1 - M(n))^n / M(n)$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $W_n$  soit convergente).*

La convergence en probabilité de  $H_n$  vers  $S$  suivant  $K$  signifie :

$$\forall \varepsilon > 0 : \lim_n P_* \{S_{(n)-\varepsilon} \subset H'_n \subset S_{(n)}^\varepsilon\} = 1$$

et la convergence presque complète de  $H_n$  vers  $S$  suivant  $K$  signifie :

$$\forall \varepsilon > 0 : \sum_{n=1}^{\infty} \{1 - P_*(S_{(n)-\varepsilon} \subset H'_n \subset S_{(n)}^\varepsilon)\} < +\infty.$$

**Théorème 2.10.-** *Une condition suffisante pour que  $H_n$  converge vers  $S$  suivant  $K$  en probabilité (resp. presque complètement) est que la suite  $W'_n = \nu_n(1 - m(n))^n$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $W'_n$  soit convergente), où  $\nu_n$  est définie par :  $\nu_n = \text{card}\{i \in F_n / X_{n,i} \subset K_n\}$ .*

– Méthode des boules ouvertes dans  $\mathbf{R}^s$ . Cas d'un support compact.

On considère ici  $X = \mathbf{R}^s$ , muni de sa distance usuelle.

On désigne par  $(\rho_n)_{n \in \mathbf{N}^*}$  une suite de réels positifs tendant vers 0 avec  $\frac{1}{n}$ .

On désigne par  $(\rho'_n)_{n \in \mathbf{N}^*}$  et  $(\varepsilon_n)_{n \in \mathbf{N}^*}$  deux suites de réels positifs vérifiant :

$$\forall n \in \mathbf{N}^* : \rho'_n + \varepsilon_n \sqrt{s}/2 < \rho_n$$

de sorte que toute boule  $B(x, \rho_n)$  contienne une boule  $B(y, \rho'_n)$  dont le centre soit sur le réseau  $R(\varepsilon_n)$  des points de  $\mathbf{R}^s$  dont les coordonnées sont des multiples entiers de  $\varepsilon_n$ .

On pose pour tout  $\rho > 0$  et pour tout  $n \in \mathbf{N}^*$  :

$$\begin{aligned} L(\rho) &= \sup_{B(x, \rho) \subset S} \lambda(B(x, \rho)) \\ \ell(\rho) &= \inf_{B(x, \rho) \subset S} \lambda(B(x, \rho)) \\ I_n &= \{x \in \mathbf{R}^s / \exists j, 1 \leq j \leq n : N_j(x) > 0\} \\ H_n &= \bigcup_{x \in I_n} B(x, \rho_n) \end{aligned}$$

**Théorème 2.11.-** *Si  $S$  est d'intérieur non-vide et si  $N$  vérifie (H), une condition nécessaire pour que  $H_n$  converge vers  $S$  en probabilité (resp. presque complètement) est que la suite  $Z_n = (1 - L(\rho_n))^n / L(\rho_n)$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $Z_n$  soit convergente).*

**Théorème 2.12.-** *Si le processus ponctuel  $N$  possède une mesure moyenne, une condition suffisante pour que  $H_n$  converge vers  $S$  en probabilité (resp. presque complètement) est que les suites  $(\rho'_n)_{n \in \mathbf{N}^*}$  et  $(\varepsilon_n)_{n \in \mathbf{N}^*}$  vérifiant :  $\forall n \in \mathbf{N}^* \rho'_n + \varepsilon_n \sqrt{s}/2 < \rho_n$ , puissent être choisies de sorte que la suite  $Z'_n = (\rho_n / \varepsilon_n)^s (1 - \ell(\rho'_n))^n / \ell(\rho'_n)$  tende vers 0 avec  $\frac{1}{n}$  (resp. que la série de terme général  $Z'_n$  soit convergente).*

Dans le cas général et en supposant que l'on dispose d'une suite croissante  $K = (K_n)_{n \in \mathbf{N}^*}$  de compacts tendant vers  $X$ , Gensbittel a établi des conditions nécessaires et suffisantes pour que l'estimateur de  $S$  converge en probabilité et presque complètement vers  $S$ .

## B I B L I O G R A P H I E.

- (1) CHEVALIER, J. *Estimation du support et du contenu du support d'une loi de probabilité.* Ann. Inst. Henri Poincaré, section B, Vol. XII, 4, 1976.
- (2) EFRON, B. *The convex hull of a random set of points.* Biometrika, 52, 331-343, 1965.
- (3) GENSBITTEL, M.H. *Contribution à l'étude statistique de répartitions ponctuelles aléatoires.* Thèse, Université Pierre et Marie Curie, Paris, 1979.
- (4) GEFFROY, J. *Sur un problème d'estimation géométrique.* Pub. I.S.U.P., 13, 191-200, 1964.
- (5) RAYNAUD, H. *Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans  $\mathbf{R}^n$ .* C.R. Acad. Sci. Paris, 262, série A, 623-627, 1965.
- (6) RENYI, A. et SULANKE, R. *Über die konverxe hülle von  $n$  zufällig gewählten punkten II.* Z. Wahrschein., 3, 138-147, 1964.
- (7) RIPLEY, B.D. et RASSON, J.P. *Finding the edge of a Poisson forest.* J. Appl. Prob., 14, 483-491, 1977.



## CHAPITRE II

*Cahiers du C.E.R.O., Volume 31, n°s 3-4, 1989  
Neuvième Rencontre Franco-Belge de Statisticiens*

## ESTIMATING THE EDGE OF A COX PROCESS AREA

by

P. JACOB and H. ABBAR

Université des Sciences et Techniques de Lille Flandres Artois

Let the domain of a Cox point process be a subset of  $\mathbb{R}^2$  defined in polar coordinates by  $S = \{(\rho, \theta) / 0 \leq \rho \leq \phi(\theta); 0 \leq \theta \leq 2\pi\}$ . This paper is devoted to the estimation of the function  $\phi$ .

## 1. INTRODUCTION

The problem of estimating the domain of a point process originates in that of finding the support of a probability distribution, given a sample. In many papers, the interest is focused on the convex hull of the random set of points. The original idea dates back to RENYI and SULANKE [13]. A similar point of view can be found in EFRON [3], RAYNAUD [12], FISCHER [4]. Later, GUILBART [7] settled the question of the continuity of the application which associates any distribution to the convex hull of its support. For a Poisson point process, the estimation of the edge of the convex domain was studied by RIPLEY and RASSON [14].

Simultaneously with [13], GEFFROY considered a planar domain  $S = \{(x,y) \in \mathbb{R}^2 / 0 < x < 1; 0 < y < f(x)\}$  and proposed an estimator  $f_n$  of the function  $f$ , based upon the extrema points of the sample in each cell of a partition of  $S$ . Necessary and sufficient conditions for  $f_n$  to be consistent, and the limit distribution were given in [6]. This has been extended to Poisson point processes in an unpublished work of GENSBITTEL [5]. Further developments may be found in BOSQ [1] and in JACOB [11]. It is worth noticing the paper of CHEVALIER [2], whose methods, though different, also generally make no convexity assumption. The present work is actually a direct descendant of [6].

In various concrete circumstances, the domain of the point process of interest contains a known particular point. A typical instance is provided by the geographical location of customers around a supermarket. For this kind of phenomenon, it seems reasonable to use a point process  $N$  with values around an origin  $O$ , and inside a star-shaped domain  $S$  defined in polar coordinates through an unknown continuous function  $\phi$  :

---

AMS 1980 Subject Classification : 60G55 - 62G05

Key words and phrases - Cox point process, support of a measure, domain of a point process, Gumbei distribution.

$$S = \{(\rho, \theta) / 0 \leq \theta \leq 2\pi; 0 \leq \rho \leq \phi(\theta); \phi(0) = \phi(2\pi)\}. \quad (1.1)$$

Further examination of the previous example indicates that the potential customers do not have the same behaviour on two different days of the week, or on two different periods of the month; moreover this mean behaviour can be directed by random circumstances, such as the weather, or advertising by other supermarkets. Nevertheless, if this mean behaviour is supposed to be given, it seems sound to neglect the possible interactions between customers in comparison with their own personal motivations. If one agrees with this empirical analysis, one can find it natural to suppose that  $N$  behave as a Cox point process.

Alternatively, a Cox point process has a high level of generality among the point processes on a plane, and thus is mathematically interesting by itself !

## 2. PRELIMINAIRES

Let  $\mathcal{B}$  be the ring of bounded Borel sets of  $\mathbb{R}^2$ , and  $\mathbb{E}$  the set of discrete measures, having non negative integers as valued on  $\mathcal{B}$ . On  $\mathbb{E}$  we define the  $\sigma$ -field  $\mathcal{E}$  generated by the coordinate mappings

$$\begin{aligned} \mathbb{E} &\rightarrow \mathbb{N} = \{0, 1, 2, \dots\} \\ m &\rightarrow m(B) \end{aligned} \quad (2.1)$$

where  $B$  ranges over  $\mathcal{B}$ . It is known ([8]) that  $\mathcal{E}$  is also the Borel  $\sigma$ -field of  $\mathbb{E}$  in the vague topology, but this is of no interest here.

A point process is an  $\mathbb{E}$ -valued random variable  $N$  on some probability space  $(\Omega, \mathcal{F}, P)$ . For each  $B \in \mathcal{B}$ ,  $N(B)$  is then an integer-valued random variable on  $(\Omega, \mathcal{F}, P)$ . The set function  $\mu$

$$\begin{aligned} \mathcal{B} &\rightarrow \mathbb{R}^+ = \{0, +\infty\} \\ \mathcal{B} &\rightarrow \mathbb{E} \{N(B)\} = \mu(B) \end{aligned} \quad (2.2)$$

is called the intensity of  $N$ . If, for each  $B \in \mathcal{B}$ ,  $N(B)$  has a finite expectation, the Radon measure  $\mu$  is called the mean measure of  $N$ . As in [9], the zero-probability functional of  $N$  is the mapping  $Z$

$$\begin{aligned} \mathcal{B} &\rightarrow [0,1] \\ \mathcal{B} &\rightarrow Z(B) = P\{N(B) = 0\}. \end{aligned} \quad (2.3)$$

The domain of  $N$  is the set  $S$ , which is the smallest closed set  $F$  such that  $Z(\mathbb{R}^2 - F) = 1$ . If  $N$  has a mean measure  $\mu$ , the domain of  $N$  is also the support of  $\mu$ .

$N$  is a Poisson process with mean measure  $\mu$  if, whenever  $A_1, \dots, A_k \in \mathcal{B}$  are disjoint, the random variables  $N(A_1), \dots, N(A_k)$  are independent, and if for each  $A \in \mathcal{B}$ ,  $N(A)$  has a Poisson distribution with mean  $\mu(A)$ .

Now, let  $\mathbf{M}$  be the set of positive Radon measures on  $\mathbb{R}^2$ , and  $\mathcal{M}$  the  $\sigma$ -field generated by the coordinate mappings :

$$\begin{aligned} \mathbf{M} &\rightarrow \mathbb{R}^+ = [0, +\infty[ \\ m &\rightarrow m(B). \end{aligned} \quad (2.4)$$

A random measure is a  $\mathbf{M}$ -valued random variable  $M$  on  $(\Omega, \mathcal{F}, P)$ .

$N$  is a Cox process directed by  $M$  if, conditional on  $\{M = \mu\}$ ,  $N$  is a Poisson process with mean measure  $\mu$ . It should be noted that the zero-probability functional of  $N$  is then ([9]) :

$$Z(B) = E(\exp(-M(B))). \quad (2.5)$$

We assume in the sequel that we are given i.i.d. copies  $N_1, \dots, N_n$  of a Cox point process  $N$  with domain of the form (1.1). We consider a sequence  $k = k_n$  of integers and for  $r \in \{1, \dots, k\}$ , we introduce the following notations :

$$I_{n,r} = [2\pi(r-1)k^{-1}; 2\pi r k^{-1}[ \quad (2.6)$$

$$D_{n,r} = \{(\rho, \theta) \in S / \theta \in I_{n,r}\} \quad (2.7)$$

$$M_{n,r} = \sup \{\phi(\theta) / \theta \in I_{n,r}\} \quad (2.8)$$

$$m_{n,r} = \inf \{\phi(\theta) / \theta \in I_{n,r}\}. \quad (2.9)$$

For each  $\omega \in \Omega$ ,  $S_n^\omega$  is the discrete support of  $N_1^\omega + \dots + N_n^\omega$ , and  $S_{n,r}^\omega = S_n^\omega \cap D_{n,r}$  : the estimate  $\phi_n^\omega(\theta)$  of  $\phi(\theta)$  is now defined on each cell  $I_{n,r}$  by :

$$U_{n,r}^\omega = \max \{\rho / (\rho, \theta) \in S_{n,r}^\omega\} \quad \text{if } S_{n,r}^\omega \neq \emptyset \quad (2.10)$$

$$U_{n,r}^\omega = 0 \quad \text{if } S_{n,r}^\omega = \emptyset.$$

Our aim is now to prove the almost sure convergence of

$$\Lambda_n = \sup \{|\phi(\theta) - \phi_n(\theta)| / \theta \in [0, 2\pi[ \} \quad (2.11)$$

in the most general framework, and to find the limit distribution.

### 3. CONVERGENCE OF THE ESTIMATOR

In this section, some smoothness assumptions are required about the Cox process  $N$ ; namely, the directing random measure  $M$  is supposed to be absolutely continuous with respect to the 2-dimensional Lebesgue measure  $\lambda$ . We also assume the existence of a continuous version of the random function  $g = \frac{dM}{d\lambda}$ , strictly positive on  $\text{int}(S)$ .

Theorem 3.1. Under the above assumptions, let :

(i)  $k_n \rightarrow \infty$  and (ii)  $k_n = o(n / \text{Log } n)$ . Then  $\Lambda_n \rightarrow 0$  almost surely.

PROOF. Let  $0 < \epsilon < 1$  be arbitrary and consider

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

If the series  $\sum_n \alpha_n$  converges, it follows from the Borel-Cantelli lemma that for all  $n$  large enough :

$$\max_r (M_{nr} - U_{n,r}) \leq (1 - \epsilon) \sup_{\theta} \{\phi(\theta)\}, \quad \text{a.s.} \quad (3.2)$$

By the triangle inequality, we obtain :

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

In view of (3.2) and the continuity of  $\phi$ , this implies :

$$0 \leq \limsup_n \Lambda_n \leq (1 - \epsilon) \sup \{\phi(\theta)\}, \quad \text{a.s.} \quad (3.4)$$

This being true for each  $0 < \epsilon < 1$ , we find  $\lim \Lambda_n = 0$ , a.s.

It is therefore enough to prove that  $\sum_n \alpha_n < +\infty$ . Fix  $0 < \epsilon < \eta < 1$ , and define :

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

Then, in view of (2.5), the following inequalities hold :

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

Next, consider the random variables

$$X_n = \inf \{g(\rho, \theta) / (\rho, \theta) \in A_{n,r}; r = 1, \dots, k\} \quad (3.7)$$

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

By the continuity of  $\phi$ , we have  $\cup A_{n,r} \subset S$  for all  $n$  large enough; thus, there exists a constant  $c$  such that, for  $r = 1, \dots, k$

$$\lambda(A_{n,r}) = \pi(n^2 - \epsilon^2) M_{n,r}^2 k^{-1} \geq c k^{-1} \quad \text{and} \quad M(A_{n,r}) \geq c X_n k^{-1} \geq c Y_n k^{-1}. \quad (3.9)$$

Since  $Y_n k^{-1} \leq k^{-1/2}$ , the inequality  $\exp(-c Y_n k^{-1}) \leq 1 - \frac{c}{2} Y_n k^{-1}$  holds in a deterministic way for all sufficiently large  $n$ . It follows from (3.6) and (3.9) that :

$$\alpha_n \leq k E^n \{ \exp(-c Y_n k^{-1}) \} \leq k \left( 1 - \frac{c}{2} E\{Y_n\} k^{-1} \right)^n \leq k \exp(-n \frac{c}{2} E\{Y_n\} k^{-1}). \quad (3.10)$$

In view of (ii),  $k = (nu)/\text{Log } n$ , with  $u = u_n \rightarrow 0$ , which yields :

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

Clearly,  $\alpha_n$  is the general term of a convergent series provided that  $\liminf_n E\{Y_n\} > 0$ , which remains to be shown. Choose  $\delta > 0$  such that

$${}_n \inf_{\theta} \{\phi(\theta)\} + \delta < \inf_{\theta} \{\phi(\theta)\} \quad (3.12)$$

For all sufficiently large  $n$ , and  $r = 1, \dots, k$ , we have by continuity of  $\phi$  :

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

This leads to the following inclusions :

$$\bigcup_{r=1}^k A_{n,r} \subset \bigcup_{r=1}^k \{(\rho, \theta) \in D_{n,r} / \rho \leq {}_n m_{n,r} + \delta\} \subset \{(\rho, \theta) / \rho \leq {}_n \phi(\theta) + \delta\} = K \quad (3.14)$$

where  $K$  is a compact set which, in view of (12), satisfies :

$$K \subset \text{int}(S) \subset \{g > 0\}. \quad (3.15)$$

Therefore, for all sufficiently large  $n$  :

$$X_n \geq X = \inf \{g(\rho, \theta) / (\rho, \theta) \in K\} > 0. \quad (3.16)$$

Finally, by Fatou's lemma and the obvious remark that  $|X_n - Y_n| \rightarrow 0$  :

$$\liminf_n E\{Y_n\} \geq E\{\liminf_n Y_n\} = E\{\liminf_n X_n\} > E(X) > 0. \quad (3.17)$$

#### 4. LIMIT DISTRIBUTION

Our aim is to study the limit distribution of

$$Z_n = \max_{r=1}^k (M_{n,r} - U_{n,r}). \quad (4.1)$$

For any  $z > 0$ , define

$$D_{n,r}(z) = \{(\rho, \theta) \in S / \theta \in I_{n,r}; \rho > M_{n,r} - z\}. \quad (4.2)$$

Since  $N^n = N_1 + \dots + N_n$  is a Cox process with directing measure  $M^n = M_1 + \dots + M_n$  (see [9]), we obtain the distribution function of  $Z_n$  :

$$\begin{aligned} F_n(z) &= P\{Z_n < z\} = P\left\{\prod_{r=1}^k \{N^n(D_{n,r}(z)) > 0\}\right\} = E\left\{\prod_{r=1}^k P\{N^n(D_{n,r}(z)) > 0/M^n\}\right\} \\ &= E\left\{\prod_{r=1}^k \left[1 - \exp(-M^n(D_{n,r}(z)))\right]\right\} \end{aligned} \quad (4.3)$$

By the bounded convergence theorem, it now suffices to determine a sequence  $\{z_n\}$  such that

$$\prod_{r=1}^k \left[1 - \exp(-M^n(D_{n,r}(z_n)))\right] \quad (4.4)$$

converges almost surely. For that purpose, we need the following useful lemma, where a triangular array  $\{x_{n,r}/r = 1, \dots, k_n; n \geq 0\}$  is called a null array if  $\max_r |x_{n,r}| = \tilde{x}_n \rightarrow 0$ .

LEMMA 4.1. Let  $\{x_{n,r}\}$  be a positive triangular array such that  $\{z_{n,r}\} = \{e^{-x_{n,r}}\}$  is a null array. If  $\prod_{r=1}^k (1 - z_{n,r})$  has a limit  $\varepsilon > 0$ , then  $\prod_{r=1}^k (1 - e^{-x_{n,r} + y_{n,r}})$  has the same limit  $\varepsilon$ ,

whenever  $\{y_{n,r}\}$  is a null array.

PROOF. By the classical inequality

$$s - s^2 \leq \text{Log}(1+s) \leq |s| \quad \text{for} \quad |s| \leq 1/2 \quad (4.5)$$

we obtain that  $\sum_{r=1}^k \text{Log}(1 - z_{n,r}) = \sum_{r=1}^k (-z_{n,r} - t_{n,r})$ , with, for sufficiently large  $n$ ,

$$\sum_{r=1}^k t_{n,r} < \sum_{r=1}^k z_{n,r}^2 < \tilde{z}_n \sum_{r=1}^k z_{n,r}. \quad \text{Since } \tilde{z}_n \rightarrow 0, \text{ it follows that}$$

$$\sum_{r=1}^k z_{n,r} \rightarrow -\text{Log } \varepsilon \quad (4.6)$$

Similarly, we can write  $\sum_{r=1}^k \text{Log}(1 - e^{-x_{n,r} + y_{n,r}}) = \sum_{r=1}^k (-e^{-x_{n,r} + y_{n,r}} - u_{n,r})$ , with, for

sufficiently large  $n$ :

$$\sum_{r=1}^k u_{n,r} < \sum_{r=1}^k e^{-2x_{n,r} + 2y_{n,r}} < 2 \sum_{r=1}^k z_{n,r}^2 < 2 \tilde{z}_n \sum_{r=1}^k z_{n,r}. \quad (4.7)$$

But we also have  $\sum_{r=1}^k e^{-x_{n,r} + y_{n,r}} = \sum_{r=1}^k z_{n,r}(1 + v_{n,r})$ , with, for sufficiently large  $n$ ,

$$|v_{n,r}| \leq 2 |y_{n,r}| \leq 2 \tilde{y}_n, \text{ hence}$$

$$\sum_{r=1}^k z_{n,r} |v_{n,r}| \leq 2 \tilde{y}_n \sum_{r=1}^k z_{n,r}. \quad (4.8)$$

Since  $\tilde{z}_n \rightarrow 0$  and  $\tilde{y}_n \rightarrow 0$ , we finally obtain, by combining (4.6), (4.7), and (4.8):

$$\sum_{r=1}^k \text{Log}(1 - e^{-x_{n,r} + y_{n,r}}) \rightarrow -\text{Log } \varepsilon. \quad (4.9)$$

The next lemma is roughly Geffroy's lemma 3 in [6]. Nevertheless, we reproduce the proof with minor modifications for the sake of completeness.

LEMMA 4.2. Suppose that  $f$  is a positive continuous function on  $[0, 2\pi]$  and that  $k = k_n$  is a sequence of positive numbers. Define the sequence  $\beta = \beta_n$  of positive numbers by the integral equation:

$$\frac{k}{2\pi} \int_0^{2\pi} \exp\left[-\frac{n}{k} \beta f(\theta)\right] d\theta = 1 \quad (4.10)$$

Then, if  $k_n \rightarrow \infty$ , for all  $z > 0$ :

$$\lim_{n \rightarrow \infty} \frac{k}{2\pi} \int_0^{2\pi} \exp \left[ - \left( z + \frac{n}{k} \beta \right) f(\theta) \right] d\theta = e^{-cz} \quad (4.11)$$

with  $c = \inf \{ f(\theta) / 0 \leq \theta \leq 2\pi \}$ .

PROOF. By the Lebesgue dominated convergence theorem, the integral in (4.10) is continuously decreasing from  $k$  to  $0$ , as  $\beta$  goes from  $0$  to  $\infty$ .

Equation (4.10) thus has a unique root if  $k > 1$ . It should also be emphasized that

$$\frac{\log k}{d} \leq \frac{n}{k} \beta \leq \frac{\log k}{c} \quad (4.12)$$

with  $d = \sup \{ f(\theta) / 0 \leq \theta \leq 2\pi \}$ . Now, let us define

$$g_n(\theta) = \frac{k}{2\pi} \exp \left\{ - \frac{n}{k} \beta f(\theta) \right\}, \quad \theta \in [0, 2\pi] \quad (4.13)$$

$$E_\epsilon = \{ \theta / 0 \leq \theta \leq 2\pi ; f(\theta) < c + \epsilon \}, \quad \epsilon > 0 \quad (4.14)$$

$$I_1(\epsilon) = \int_{E_\epsilon} g_n(\theta) d\theta ; I_2(\epsilon) = \int_{[0, 2\pi] - E_\epsilon} g_n(\theta) d\theta, \quad \epsilon > 0. \quad (4.15)$$

We have  $I_2(\epsilon)/I_1(\epsilon) = O(\exp(-\epsilon \frac{n}{k} \beta))$  hence  $I_2(\epsilon)/I_1(\epsilon) \rightarrow 0$  as  $n \rightarrow \infty$ , by (4.12). Since  $I_1(\epsilon) + I_2(\epsilon) < 1$  by (4.10), it follows that  $I_2(\epsilon) \rightarrow 0$ . This gives

$$I_1(\epsilon) = \int_{E_\epsilon} g_n(\theta) d\theta \rightarrow 1, \quad \epsilon > 0 \quad (4.16)$$

Let us return to  $J_n(z) = \frac{k}{2\pi} \int_0^{2\pi} \exp \left[ - \left( z + \frac{n}{k} \beta \right) f(\theta) \right] d\theta$ ; in view of (4.16), we obtain :

$$J_n(z) = \int_{E_\epsilon} e^{-zf(\theta)} g_n(\theta) d\theta + o(1). \quad (4.17)$$

Therefore

$$e^{-(c+\epsilon)z} I_1(\epsilon) + o(1) \leq J_n(z) \leq e^{-cz} I_1(\epsilon) + o(1) \quad (4.18)$$

and

$$e^{-(c+\epsilon)z} \leq \liminf_n J_n(z) \leq \limsup_n J_n(z) \leq e^{-cz}. \quad (4.19)$$

Since  $\epsilon$  was arbitrary,  $\lim_n J_n(z) = e^{-cz}$ , for each  $z > 0$ .

We now consider the special case of a mixed Poisson process.

**THEOREM 4.3.** Let  $\alpha$  be a strictly positive random variable and  $\lambda$  be the 2-dimensional Lebesgue measure. Suppose that  $M = \alpha\lambda$ , that  $\alpha$  has a finite variance and that  $\phi$  is a differentiable function. Let  $\beta = \beta_n$  be the sequence defined by

$$\frac{k}{2\pi} \int_0^{2\pi} \exp \left[ - \frac{n}{k} \beta E\{\alpha\} \phi(\theta) \right] d\theta = 1 \quad (4.20)$$

and  $z_n = \frac{1}{2\pi} \left( \frac{k}{n} z + \beta_n \right)$ ,  $\forall z > 0$ . Then, under the assumptions

(i)  $k = o(n / (\text{Log } n)^2)$  (ii)  $n = o(k^2)$ , we have

$$F_n(z_n) \rightarrow \exp[-\exp(-E\{\alpha\} cz)] \quad (4.21)$$

where  $c = \inf\{\phi(\theta) ; 0 \leq \theta \leq 2\pi\}$ .

PROOF. For every  $z > 0$ , we have

$$2\pi M_{n,r} z - \pi z^2 - \pi(M_{n,r}^2 - m_{n,r}^2) \leq k \lambda\{D_{n,r}(z)\} \leq 2\pi M_{n,r} z - \pi z^2. \quad (4.22)$$

Moreover, the differentiability of  $\phi$  provides a bound on  $(M_{n,r}^2 - m_{n,r}^2)$  :

$$(M_{n,r}^2 - m_{n,r}^2) \leq 2 \sup_{\theta_1, \theta_2 \in I_{n,r}} |\phi(\theta_1) - \phi(\theta_2)| \sup_{\theta} \phi(\theta) \leq \frac{4\pi}{k} \sup_{\theta} |\phi'(\theta)| \sup_{\theta} \phi(\theta). \quad (4.23)$$

Hence, in view of the assumption (ii), we obtain

$$n \lambda\{D_{n,r}(z)\} = \frac{n}{k} 2\pi M_{n,r} z - \frac{n}{k} \pi z^2 + \epsilon_{n,r}, \quad (4.24)$$

where  $\{\epsilon_{n,r}\}$  is a null array. Now, by (4.11) in lemma 4.2, we have  $z_n^2 = O(k^2(\text{Log } k)^2/n^2)$  and  $k(\text{Log } k)^2/n \rightarrow 0$  is easily derived from the assumption (i). Thus  $n \pi z^2/k \rightarrow 0$ , and

$$n \lambda\{D_{n,r}(z_n)\} = (z + \frac{n}{k} \beta) M_{n,r} + \epsilon'_{n,r}. \quad (4.25)$$

where  $\{\epsilon'_{n,r}\}$  is a null array.

For  $r = 1, \dots, k$  choose a  $\theta_{n,r} \in I_{n,r}$  such that

$$\exp[-(z + \frac{n}{k} \beta) E\{\alpha\} \phi(\theta_{n,r})] = \frac{k}{2\pi} \int_{I_{n,r}} \exp[-(z + \frac{n}{k} \beta) E\{\alpha\} \phi(\theta)] d\theta. \quad (4.26)$$

Then,  $M_{n,r} = \phi(\theta_{n,r}) + O(1/k)$  in view of the differentiability of  $\phi$  (see (4.23)), and  $O((z + n\beta/k)/k) = O((n k \text{Log } k)/n k^2) = O(\text{Log } k/k) \rightarrow 0$ . We thus obtain

$$n \lambda\{D_{n,r}(z_n)\} = (z + \frac{n}{k} \beta) \phi(\theta_{n,r}) + \epsilon''_{n,r}, \quad (4.27)$$

where  $\{\epsilon''_{n,r}\}$  is a null array.

Now, by the law of the Iterated Logarithm, we can write  $n^{-1}(\alpha_1 + \dots + \alpha_n)$  in the form  $E\{\alpha\} + \eta_n$ , where  $(\sqrt{n} \eta_n / \sqrt{\text{Log } \text{Log } n})$  is almost surely a bounded sequence. By (4.12), we have that

$$\eta_n n \beta/k = \left[ \sqrt{n} \eta_n / \sqrt{\text{Log } \text{Log } n} \right] O(\sqrt{\text{Log } \text{Log } n} \text{Log } k / \sqrt{n}) \quad (4.28)$$

and since, once again,  $k(\text{Log } k)^2/n \rightarrow 0$ ,  $\{\eta_n(z + \frac{n}{k} \beta) \phi(\theta_{n,r})\}$  is almost surely a null array. Finally, the following representation holds :

$$M^n(D_{n,r}(z_n)) = (E(\alpha) + \eta_n) \left( (z + \frac{n}{k} \beta) \phi(\theta_{n,r}) + \epsilon_{n,r}^n \right) \tag{4.29}$$

By Lemma 4.1, if  $\prod_{r=1}^k (1 - \exp[-(z + n\beta/k) E(\alpha) \phi(\theta_{n,r})])$  has a positive limit  $\ell$ , then

$\prod_{r=1}^k (1 - \exp[-M^n(D_{n,r}(z_n))])$  has the same limit  $\ell$ . But, by (4.26)

$$\sum_{r=1}^k \exp\left[-\left(z + \frac{n}{k} \beta\right) E(\alpha) \phi(\theta_{n,r})\right] = (k/2\pi) \int_0^{2\pi} \exp\left[-\left(z + \frac{n}{k} \beta\right) E(\alpha) \phi(\theta)\right] d\theta. \tag{4.30}$$

Finally, in view of the particular choice of  $\beta$ , Lemma 4.2 carries over in an obvious way : the limit of (4.30) is  $-\text{Log } \ell = \exp[-c E(\alpha) z]$  and

$$\ell = \exp[-\exp(-c E(\alpha) z)]. \tag{4.31}$$

As a further development, we can now assume that  $M$  has a positive continuous density on  $S$

$$g(u) = \frac{dM}{d\lambda}(u), \quad u \in \mathbb{R}^2. \tag{4.32}$$

Contrary to the situation in section 3, we cannot allow  $g$  to be arbitrarily small in the neighbourhood of the edge of  $S$ . We therefore assume that there exists a random variable  $U$  such that

$$0 \leq U \cdot \mathbb{1}_S \leq g, \text{ a.s.} \tag{4.33}$$

Let  $g_1, \dots, g_n$  be independent copies of  $g$ ; just as in the proof of Theorem 4.3, we shall need a L.I.L., but now in the  $C(S)$  setting. A theorem of J. KUELBS (Theorem 4.4 in [10]) gives very general conditions, involving entropy considerations, which lead to the following result : there exists a compact set  $K$  in  $C(S)$  such that

$$P \left\{ \lim_{n \rightarrow \infty} \mathbb{1} \left\| \frac{(g_1 + \dots + g_n) - n E(g)}{\sqrt{2n \text{Log Log } n}} \right\| ; K \right\} = 1. \tag{4.34}$$

Where  $\|f, K\| = \inf \{ \sup \{|f(x) - g(x)| / x \in S\} / g \in K \}$ . As a very coarse application of this fine result we suppose that there exists a random variable  $M$  such that  $E(M^2) < +\infty$  and

$$|g(s) - g(t)| \leq M d(s,t) \tag{4.35}$$

where  $d$  is the natural metric on  $S$ . Let  $N(d, S, \epsilon)$  be the minimal number of balls of  $d$ -radius less than  $\epsilon$  which cover  $S$ , and let  $H(d, S, \epsilon) = \text{Log } N(d, S, \epsilon)$ . Then it is easily seen that the condition

$$\int_0^1 H^{1/2}(d, S, \epsilon) d\epsilon < +\infty \tag{4.36}$$

is fulfilled. One could have hoped for some milder assumption than (4.36), but nevertheless this Lipschitz condition will be useful elsewhere in the proof. The condition (4.37) below will be also necessary (see [10]). There exists a random variable  $V$  such that  $E(V^2) < +\infty$  and

$$0 \leq g \leq V \cdot \mathbb{1}_S \quad (4.37)$$

THEOREM 4.4. Assume that  $M$  has a density  $g$  satisfying (4.33), (4.35) and (4.37), and that  $\phi$  is a differentiable function of  $\theta$ . Let  $u(\theta)$  be the point of  $\mathbb{R}^2$  with polar coordinates  $(\theta, \phi(\theta))$ ; let  $\beta = \beta_n$  be the sequence defined by

$$\frac{k}{2\pi} \int_0^{2\pi} \exp \left[ -\frac{n}{k} \beta E \{g(u(\theta))\} \phi(\theta) \right] d\theta = 1 \quad (4.38)$$

and  $z_n = \frac{1}{2\pi} \left( \frac{k}{n} z + \beta_n \right)$ ,  $\forall z > 0$ . Then, under the assumptions

- (i)  $k = o(n / (\log n)^2)$ ,
- (ii)  $n = o(k^2)$ , we have that

$$F_n(z_n) \rightarrow \exp [- \exp(- cz)] \quad (4.39)$$

where  $c = \min \left\{ E \{ (g(u, \theta)) \} \phi(\theta) / 0 \leq \theta \leq 2\pi \right\}$ .

PROOF. The beginning of the previous proof, [(4.22), (4.23), (4.24), (4.25)], remains valid, but in (4.26) the choice of  $\theta_{n,r} \in I_{n,r}$  must satisfy

$$\begin{aligned} & \exp \left[ - \left( z + \frac{n}{k} \beta \right) \left\{ E \{ g(u(\theta_{n,r})) \} \phi(\theta_{n,r}) \right\} \right] \\ &= \frac{k}{2\pi} \int_{I_{n,r}} \exp \left[ - \left( z + \frac{n}{k} \beta \right) E \{ g(u(\theta)) \} \phi(\theta) \right] d\theta \end{aligned} \quad (4.40)$$

Let us define  $u_{n,r} = u(\theta_{n,r})$ . In each  $D_{n,r}(z_n)$ , we take a random point  $x_{i,n,r}$  such that

$$G_i(x_{i,n,r}) \lambda(D_{n,r}(z_n)) = \int_{D_{n,r}(z_n)} g_i(u) \lambda(du). \quad (4.41)$$

Then there exists a null array  $\{\epsilon_{n,r}\}$  such that

$$\begin{aligned} M^n(D_{n,r}(z_n)) &= \sum_{i=1}^n g_i(x_{i,n,r}) \lambda(D_{n,r}(z_n)) \\ &= \left[ \left( z + \frac{n}{k} \beta \right) \phi(\theta_{n,r}) + \epsilon_{n,r} \right] \frac{1}{n} \sum_{i=1}^n g_i(x_{i,n,r}). \end{aligned} \quad (4.42)$$

Since  $\frac{1}{n} \sum_{i=1}^n g_i(x_{i,n,r}) \leq \frac{1}{n} \sum_{i=1}^n V_i$ , it follows that  $\left\{ \epsilon_{n,r} \cdot \frac{1}{n} \sum_{i=1}^n g_i(x_{i,n,r}) \right\}$  is almost surely a null array. Since  $\phi$  is bounded, it remains to show that

$$\left\{ \left( z + \frac{n}{k} \beta \right) \left( \frac{1}{n} \sum_{i=1}^n g_i(x_{i,n,r}) - E \{ g(u_{n,r}) \} \right) \right\} \quad (4.43)$$

is also almost surely a null array; and the same kind of arguments as in theorem 4.3. will complete the proof. Now,

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i=1}^n g_i(x_{i,n,r}) - E\{g(u_{n,r})\} \right| \\ & \leq \left| \frac{1}{n} \sum_{i=1}^n (g_i(x_{i,n,r}) - g_i(u_{n,r})) \right| + \left| \frac{1}{n} \sum_{i=1}^n g_i(u_{n,r}) - E\{g(u_{n,r})\} \right| \end{aligned} \quad (4.44)$$

and by (4.35), the first term in the sum (4.44) is bounded by

$$\text{diam}(D_{n,r}(z_n)) \cdot \frac{1}{n} \sum_{i=1}^n M_i. \quad (4.45)$$

One can see that  $\max\{\text{diam } D_{n,r}(z_n)/r = 1, \dots, k\} = O(\beta + k^{-1})$ .

Now, by (4.12), we get  $n\beta^2/k < k(\text{Log } k)^2/n$  and  $n\beta/k^2 < (\text{Log } k)/k$ ; thus, in view of the assumptions (i) and (ii),  $(z + n\beta/k) O(\beta + k^{-1}) \rightarrow 0$ . On the other hand, the second term in (4.44) is bounded by

$$\begin{aligned} & \sup \left\{ \left| \frac{1}{n} \sum_{i=1}^n g_i(u(\theta)) - E\{g(u(\theta))\} \right| / 0 \leq \theta \leq 2\pi \right\} \\ & \leq \sup \left\{ \left| \frac{1}{n} \sum_{i=1}^n g_i(u) - E\{g(u)\} \right| / u \in S \right\} \end{aligned} \quad (4.46)$$

$$:= \|\bar{g}_n - E\bar{g}\|. \quad (4.47)$$

By the law of the iterated Logarithm of J. KUELBS, there exists a compact set  $K$  in  $C(S)$  such that, for every  $\epsilon > 0$

$$P \left\{ \frac{\sqrt{n} (\bar{g}_n - E\bar{g})}{\sqrt{2 \text{Log } \text{Log } n}} \notin K^\epsilon, \text{ i.o.} \right\} = 0 \quad (4.48)$$

where  $K^\epsilon = \{f \in C(S) : \|f, K\| \leq \epsilon\}$  is a bounded set. Thus,  $(\sqrt{n} \|\bar{g}_n - E\bar{g}\| / \sqrt{2 \text{Log } \text{Log } n})$  is almost surely a bounded sequence, and the same argument as for (4.28) holds. Thus, we have shown that (4.43) is actually a null array, which concludes the proof.

## 5. CONCLUSION

The model is less restrictive than the models in [5] and [6] since we only require that the point process  $N$  should be a Cox process. Moreover, in section 3, the intensity of  $N$  is allowed to be arbitrarily small on the edge of the domain. However, in the quest for greater generality, many questions are still unsolved: for instance, it can be noticed that for each  $\omega \in \Omega$ , the underlying measure  $M$  has the same support  $N$ .

As a final remark, our estimator looks like an histogram, and some smoothness probably would be welcome. This is the subject of our ongoing work.



## REFERENCES

- [1] BOSQ, D. (1971). Contribution à la théorie de l'estimation fonctionnelle. Thesis, Université Pierre et Marie Curie, Paris.
- [2] CHEVALIER, J. (1976). Estimation du support et du contenu du support d'une loi de probabilité. *Ann. Inst. Henri Poincaré, Section B, Vol. XII, 4*, 339-364.
- [3] EFRON, B. (1965). The convex hull of a random set of points. *Biometrika*, 52, 331-343.
- [4] FISHER, L. (1966). The convex hull of a sample. *Bull. Am. Math. Soc.*, 72, 555-558.
- [5] GENSBITTEL, M.H. (1979). Contribution à l'étude statistique des répartitions ponctuelles aléatoires. Thesis, Université Pierre et Marie Curie, Paris.
- [6] GEFFROY, J. (1964). Sur un problème d'estimation géométrique. *Pub. ISUP*, 13, 191-200.
- [7] 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. *C.R. Acad. Sci. Paris, 277, Série A*, 999-1002.
- [8] KALLENBERG, O. (1983). *Random measures*. Academic Press.
- [9] KARR, A.F. (1986). *Point processes and their statistical inference*. Dekker.
- [10] KUEBLS, J. (1976). A strong convergence theorem for Banach space valued random variables. *Annals of Probability*, 4, 744-771.
- [11] JACOB, P. (1984). Estimation du contour discontinu d'un processus ponctuel sur le plan. *Publications ISUP*, XXIX, 3-4, 1-24.
- [12] RAYNAUD, H. (1965). Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans  $\mathbb{R}^n$ . *C.R. Acad. Sci. Paris, 262, série A*, 627-629.
- [13] RENYI, A. and SUNLANKE, R. (1964). Über die konvexe hülle von  $n$  zufällig gewählten Punkten II. *Z. Wahrschein.*, 3, 138-147.
- [14] RIPLEY, B.D. and RASSON, J.P. (1977). Finding the edge of a Poisson Forest. *J. Appl. Prob.* 14, 483-491.



## CHAPITRE III

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

par

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  $\mathbf{R}^2$  défini en coordonnées polaires par  $S = \{(\rho, \theta) | 0 \leq \rho < \Phi(\theta); 0 \leq \theta \leq 2\pi\}$ . Cet article est consacré à l'estimation de la fonction  $\Phi$  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  $\mathbf{R}^2$  defined in polar coordinates by  $S = \{(\rho, \theta) | 0 \leq \rho < \Phi(\theta); 0 \leq \theta \leq 2\pi\}$ . This paper is devoted to the estimation of the function  $\Phi$  by a random spline.*

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

## 1 - INTRODUCTION

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

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

Dans un article récent avec P. Jacob [14], nous avons étudié la convergence d'un estimateur de  $\Phi$  de type histogramme, ainsi que sa loi limite. Nous définissons dans l'article que voici un estimateur de  $\Phi$  de type spline d'ordre trois, 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 [14], et nous renvoyons pour plus de détails à l'article en question.

## 2 - PRÉLIMINAIRES

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 [11], pour tout borélien borné  $B$  de  $\mathbf{R}^2$  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  $(\Omega, \mathcal{F}, P)$ .

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  $S_n^\omega$  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 [14] est défini sur tout  $I_{n,r}$  par :

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

La construction d'un tel estimateur est inspirée d'un article de Geffroy [8] et de travaux ultérieurs de Gensbittel [7].

On construit ensuite une fonction spline cubique interpolant les points  $\{\theta_{n,r}; r = 1, \dots, k+1\}$ . L'existence et l'unicité de cette fonction  $\psi_n$  sont établies dans la section 4 à l'aide de méthodes voisines de celles utilisées par Berlinet dans l'estimation de la densité [2].

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 tout l'article, on supposera remplies les conditions énoncées dans le paragraphe 2.

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

■

Soit  $0 < \varepsilon < 1$  arbitraire 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 suit d'après le 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 de la continuité de  $\Phi$ , ceci implique :

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

Ceci est vrai pour tout  $\varepsilon$  voisin de 1, donc  $\lim_n \Lambda_n = 0$ , p.s..

Il suffit par conséquent de démontrer que  $\sum_n \alpha_n < +\infty$ . Fixons  $0 < \varepsilon < \eta < 1$ , et définissons :

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

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

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

Considérons, maintenant, les variables aléatoires :

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

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

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

$$(3.9) \quad \begin{aligned} \lambda(A_{n,r}) &= \pi(\eta^2 - \varepsilon^2)M_{n,r}^2 k^{-1} \geq Ck^{-1} \\ \text{et } M(A_{n,r}) &\geq CX_n k^{-1} \geq CY_n k^{-1} \end{aligned}$$

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 :

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

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

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

Il est clair que  $\alpha_n$  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  $\Phi$  :

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

Ceci conduit aux inclusions suivantes :

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

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 en coordonnées polaires.

##### 4.1.- Fonctions splines cubiques en coordonnées polaires.

**a) - Définition.-** Une fonction spline cubique (relativement à  $\theta_{n,r} = \theta_r$  pour  $r = 1, \dots, k+1$ ) est une fonction  $\psi$  de classe  $C^2$  sur  $[0, 2\pi]$  dont la restriction à  $[\theta_r, \theta_{r+1}]$ , pour  $r = 1, \dots, k$ , est un polynôme de degré trois au plus.

On pose :

$$(4.1.1) \quad M_r = \psi''(\theta_r), \text{ pour } r = 1, \dots, k+1.$$

On va voir que si l'on impose à  $\psi$  des conditions d'interpolation :

$$(4.1.2) \quad \begin{aligned} \psi(\theta_r) &= \rho_r, \text{ pour } r = 1, \dots, k+1 \\ \text{avec } \rho_{k+1} &= \rho_1, \end{aligned}$$

la fonction spline  $\psi$  est entièrement déterminée par la donnée de ses "moments"  $(M_r)_{1 \leq r \leq k+1}$  et la condition supplémentaire  $M_{k+1} = M_1$ .

##### b) - Expression de la spline en fonction des $(M_r)$ .

D'après la linéarité de la dérivée seconde de  $\psi$  sur  $[\theta_r, \theta_{r+1}]$  on a :

$$\psi''(\theta) = M_r \frac{k(\theta_{r+1} - \theta)}{2\pi} + M_{r+1} \frac{k(\theta - \theta_r)}{2\pi}, \text{ pour } \theta \in [\theta_r, \theta_{r+1}].$$



Ce système est équivalent au système suivant :

$$\sum \left\{ \begin{array}{l} \frac{1}{2}M_{r-1} + 2M_r + \frac{1}{2}M_{r+1} = \frac{3}{4}\left(\frac{k}{\pi}\right)^2(\rho_{r-1} - 2\rho_r + \rho_{r+1}) \text{ pour } r = 2, \dots, k \\ M_1 + \frac{1}{2}M_2 + \frac{1}{2}M_k + M_{k+1} = \frac{3}{4}\left(\frac{k}{\pi}\right)^2(\rho_2 - \rho_1 - \rho_{k+1} + \rho_k) \end{array} \right.$$

L'existence de  $\psi$  est donc assurée. Son unicité résultera de la condition aux limites  $M_1 = M_{k+1}$ . Le système  $\sum$  se réduit donc au système suivant :

$$\sum^* \left\{ \begin{array}{l} 2M_1 + \frac{1}{2}M_2 + \frac{1}{2}M_k = d_1 \\ 1/2M_1 + \frac{1}{2}M_{k-1} + 2M_k = d_k \\ \frac{1}{2}M_{r-1} + 2M_r + \frac{1}{2}M_{r+1} = d_r \text{ pour } r = 2, \dots, k-1. \end{array} \right.$$

Avec :

$$\begin{aligned} d_1 &= \frac{3}{4}\left(\frac{k}{\pi}\right)^2(\rho_2 - 2\rho_1 + \rho_k) \\ d_k &= \frac{3}{4}\left(\frac{k}{\pi}\right)^2(\rho_{k-1} - 2\rho_k + \rho_1) \\ d_r &= \frac{3}{4}\left(\frac{k}{\pi}\right)^2(\rho_{r-1} - 2\rho_r + \rho_{r+1}) \text{ pour } r = 2, \dots, k-1 \end{aligned}$$

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

$$(4.1.5) \quad \begin{pmatrix} 2 & 1/2 & 0 & \dots & 0 & 0 & 1/2 \\ 1/2 & 2 & 1/2 & \dots & 0 & 0 & 0 \\ 0 & 1/2 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & 1/2 & 0 \\ 0 & 0 & 0 & \dots & 1/2 & 2 & 1/2 \\ 1/2 & 0 & 0 & \dots & 0 & 1/2 & 2 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ \vdots \\ M_{k-1} \\ M_k \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{k-1} \\ d_k \end{pmatrix},$$

ou

$$A \cdot M = d.$$

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

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

$$\sup_{1 \leq r \leq k} |M_r| \leq \sup_{1 \leq r \leq k} |d_r|.$$

■ voir [1] pages 20-21. ■

#### 4.2.- Construction de l'estimateur spline cubique.

Pour tout  $\omega \in \Omega$ , soit  $\psi_n^\omega$  la fonction spline cubique qui interpole l'estimateur de type histogramme  $\Phi_n^\omega$  défini d'après (2.7) en les  $\{\theta_r ; r = 1, \dots, k+1\}$ . On a donc :

$$\psi_n^\omega(\theta_r) = U_r^\omega, \text{ pour } r = 1, \dots, k+1$$

avec

$$U_{k+1}^\omega = U_1^\omega$$

et

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

D'après (4.1.3), on a pour  $\theta \in [\theta_r, \theta_{r+1}]$  :

$$(4.2.1) \quad \begin{aligned} \psi_n(\theta) = & M_r \frac{k(\theta_{r+1} - \theta)^3}{12\pi} + M_{r+1} \frac{k(\theta - \theta_r)^3}{12\pi} + M_r \frac{\pi(\theta - \theta_{r+1})}{3k} \\ & - M_{r+1} \frac{\pi(\theta - \theta_r)}{3k} + \frac{kU_{r+1}(\theta - \theta_r) - kU_r(\theta - \theta_{r+1})}{2\pi}, \end{aligned}$$

où les  $(M_r)$  sont les solutions du système défini d'après (4.1.5) avec :

$$d_1 = \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (U_2 - 2U_1 + U_k)$$

$$d_k = \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (U_{k-1} - 2U_k + U_1)$$

$$d_r = \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (U_{r-1} - 2U_r + U_{r+1}) \text{ pour } r = 2, \dots, k-1$$

#### 4.3.- Convergence de l'estimateur spline cubique.

Notre but maintenant est d'étudier la convergence presque sûre de  $\Gamma_n = \sup\{|\Phi(\theta) - \psi_n(\theta)| \mid \theta \in [0, 2\pi]\}$  vers 0.

Nous allons d'abord démontrer deux propositions utiles. Soit  $\psi$  la fonction spline cubique qui interpole la fonction  $\Phi$  en les  $\{\theta_r ; r = 1, \dots, k+1\}$ . On a donc :

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

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

D'après (4.1.3) on a pour  $\theta \in [\theta_r, \theta_{r+1}]$  :

$$(4.3.1) \quad \begin{aligned} \psi(\theta) = & M_r' \frac{k(\theta_{r+1} - \theta)^3}{12\pi} + M_{r+1}' \frac{k(\theta - \theta_r)^3}{12\pi} + M_r' \frac{\pi(\theta - \theta_{r+1})}{3k} \\ & - M_{r+1}' \frac{\pi(\theta - \theta_r)}{3k} + \frac{k\Phi(\theta_{r+1})(\theta - \theta_r) - k\Phi(\theta_r)(\theta - \theta_{r+1})}{2\pi}, \end{aligned}$$

où les  $(M'_r)$  sont les solutions du système défini d'après (4.1.5) :

$$A \times \begin{pmatrix} M'_1 \\ \vdots \\ \vdots \\ M'_k \end{pmatrix} = \begin{pmatrix} d'_1 \\ \vdots \\ \vdots \\ d'_k \end{pmatrix}, \text{ ou } AM' = d'$$

avec :

$$(4.3.2) \quad \begin{aligned} d'_1 &= \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (\Phi(\theta_2) - 2\Phi(\theta_1) + \Phi(\theta_k)) \\ d'_k &= \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (\Phi(\theta_{k-1}) - 2\Phi(\theta_k) + \Phi(\theta_1)) \\ d'_r &= \frac{3}{4} \left(\frac{k}{\pi}\right)^2 (\Phi(\theta_{r-1}) - 2\Phi(\theta_r) + \Phi(\theta_{r+1})) \text{ pour } r = 2, \dots, k-1 \end{aligned}$$

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

■ Pour  $\theta \in [\theta_r, \theta_{r+1}]$  on a :

$$(4.3.3) \quad \begin{aligned} |\psi(\theta) - \Phi(\theta)| &= \left| M'_r \frac{k(\theta_{r+1} - \theta)^3}{12\pi} + M'_{r+1} \frac{k(\theta - \theta_r)^3}{12\pi} \right. \\ &\quad \left. + M'_r \frac{\pi(\theta - \theta_{r+1})}{3k} - M'_{r+1} \frac{\pi(\theta - \theta_r)}{3k} \right. \\ &\quad \left. + \frac{k\Phi(\theta_{r+1})(\theta - \theta_r) - k\Phi(\theta_r)(\theta - \theta_{r+1})}{2\pi} - \Phi(\theta) \right| \\ &\leq \frac{2\pi^2}{3k^2} |M'_r| + \frac{2\pi^2}{3k^2} |M'_{r+1}| + \frac{2\pi^2}{3k^2} |M'_r| + \frac{2\pi^2}{3k^2} |M'_{r+1}| \\ &\quad + \left| \frac{k}{2\pi} (\theta - \theta_r)(\Phi(\theta_{r+1}) - \Phi(\theta)) - \frac{k}{2\pi} (\theta - \theta_{r+1})(\Phi(\theta_r) - \Phi(\theta)) \right| \\ &\leq \frac{4\pi^2}{3k^2} (|M'_r| + |M'_{r+1}|) + |\Phi(\theta_{r+1}) - \Phi(\theta)| + |\Phi(\theta_r) - \Phi(\theta)| \\ &\leq \frac{8\pi^2}{3k^2} \sup\{|M'_r| \mid 1 \leq r \leq k\} + 2\mu(\Phi, \frac{2\pi}{k}). \end{aligned}$$

D'après le lemme 4.1.1, (4.3.2) et (4.3.3) on a pour  $\theta \in [\theta_r, \theta_{r+1}]$  :

$$|\psi(\theta) - \Phi(\theta)| \leq 6\mu(\Phi, \frac{2\pi}{k}).$$

Par conséquent on a :

$$\sup\{|\Phi(\theta) - \psi(\theta)| \mid \theta \in [0, 2\pi]\} \leq 6\mu(\Phi, \frac{2\pi}{k}).$$

■

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

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

■ Pour  $\theta \in [\theta_r, \theta_{r+1}]$  on a :

$$\begin{aligned} |\psi_n(\theta) - \psi(\theta)| &= \left| \left( \frac{k(\theta_{r+1} - \theta)^3}{12\pi} + \frac{\pi(\theta - \theta_{r+1})}{3k} \right) (M_r - M'_r) \right. \\ &\quad + \left( \frac{k(\theta - \theta_r)^3}{12\pi} - \frac{\pi(\theta - \theta_r)}{3k} \right) (M_{r+1} - M'_{r+1}) \\ &\quad + \frac{k(\theta - \theta_r)}{2\pi} (U_{r+1} - \Phi(\theta_{r+1})) \\ &\quad \left. - \frac{k(\theta - \theta_{r+1})}{2\pi} (U_r - \Phi(\theta_r)) \right| \\ (4.3.4) \quad &\leq \frac{4}{3} \left( \frac{\pi}{k} \right)^2 |M_r - M'_r| + \frac{4}{3} \left( \frac{\pi}{k} \right)^2 |M_{r+1} - M'_{r+1}| \\ &\quad + |U_{r+1} - \Phi(\theta_{r+1})| + |U_r - \Phi(\theta_r)| \\ &\leq \frac{8}{3} \left( \frac{\pi}{k} \right)^2 \sup_{1 \leq r \leq k} \{|M_r - M'_r|\} + 2 \sup_{0 \leq \theta \leq 2\pi} \{|\Phi(\theta) - \Phi_n(\theta)|\} \end{aligned}$$

On sait que  $AM = d$  et  $AM' = d'$ , donc  $M - M' = A^{-1} \cdot (d - d')$ . D'après le lemme 4.1.1, on obtient :

$$(4.3.5) \quad \sup\{|M_r - M'_r| \mid r = 1, \dots, k\} \leq \sup\{|d_r - d'_r| \mid r = 1, \dots, k\}.$$

D'après (4.2.2.) et (4.3.2.), on obtient :

$$(4.3.6) \quad \sup\{|d_r - d'_r| \mid r = 1, \dots, k\} \leq 3 \left( \frac{k}{\pi} \right)^2 \sup\{|\Phi(\theta) - \Phi_n(\theta)| \mid \theta \in [0, 2\pi]\}.$$

D'après (4.3.4), (4.3.5) et (4.3.6) ; on obtient pour  $\theta \in [\theta_r, \theta_{r+1}]$  :

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

et par conséquent on a :

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

■

**Théorème 4.3.1.-** Supposons que (i)  $k_n \rightarrow \infty$  et (ii)  $k_n = o\left(\frac{n}{\log n}\right)$ . Alors  $\Gamma_n$  converge vers 0 presque sûrement.

■ Par l'inégalité triangulaire on a :

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

D'après les propositions 4.3.1 et 4.3.2 on a :

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

D'après la continuité de  $\Phi$  on a :

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

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  $(\rho, \theta)$  est telle que  $\theta$  suit la loi uniforme sur  $[0, 2\pi]$  et  $\rho$  suit la loi uniforme sur  $[0, \Phi(\theta)]$ ,  $\rho$  et  $\theta$  étant indépendants.

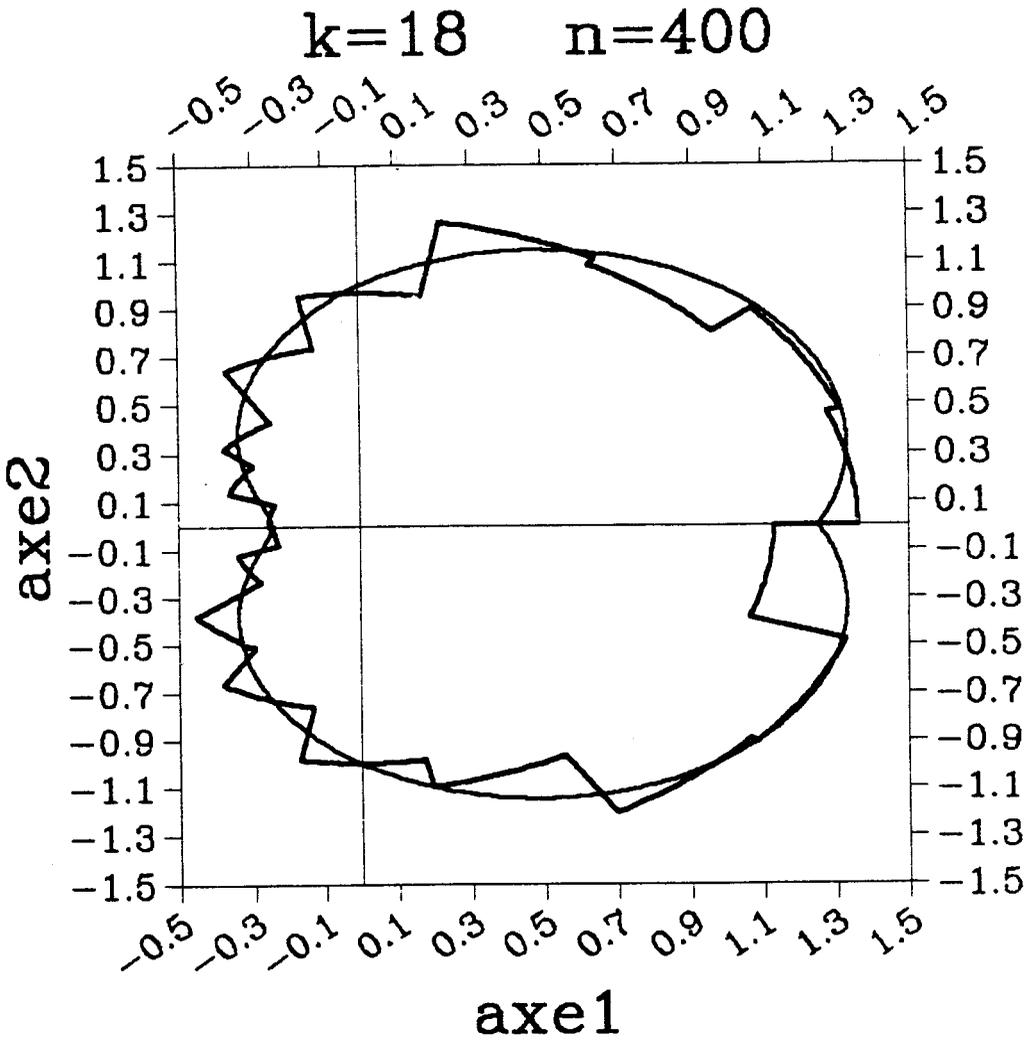
Dans les cas suivants,  $N(S)$  suit une loi de Poisson de paramètre  $n = 400$  (cas 5 et 6) ou  $n = 1000$  (cas 7 et 8), le nombre de secteurs est 9 (cas 5 et 6) ou 18 (cas 7 et 8). La loi de probabilité d'un point aléatoire  $(\rho, \theta)$  est telle que sa densité soit arbitrairement petite au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

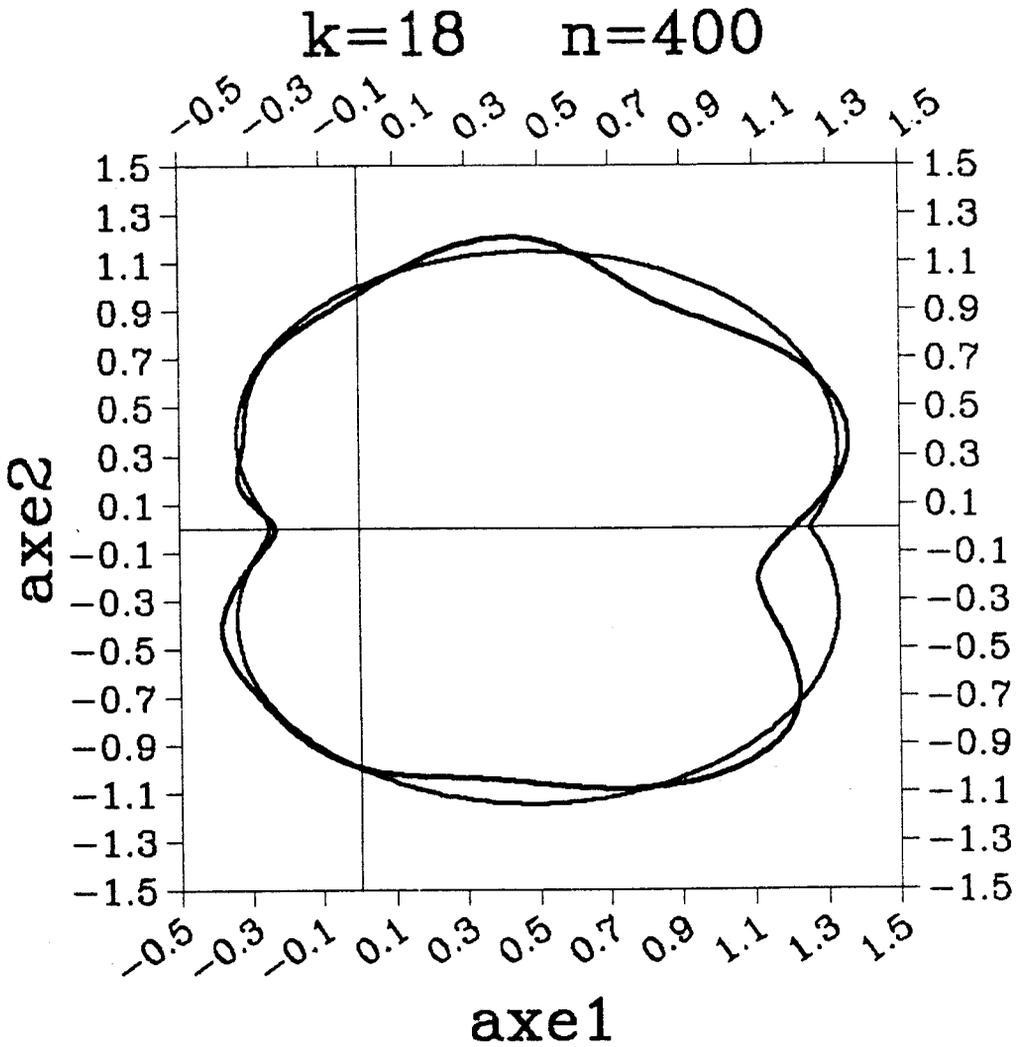
On a choisi un nombre de secteurs plus élevé lorsque la densité ne s'annule pas sur la frontière de  $S$  parce que, intuitivement, il y a dans ce cas plus de points disponibles sur le bord du domaine.

Les figures ayant un numéro impair représentent l'estimateur de type histogramme  $\Phi_n$  et celles de numéro pair, l'estimateur spline  $\psi_n$ . 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'œil nu" que, pour les deux types de lois, un nombre de secteurs trop grand donne une moins bonne estimation de  $\Phi$ , même avec un nombre de points un peu plus grand.

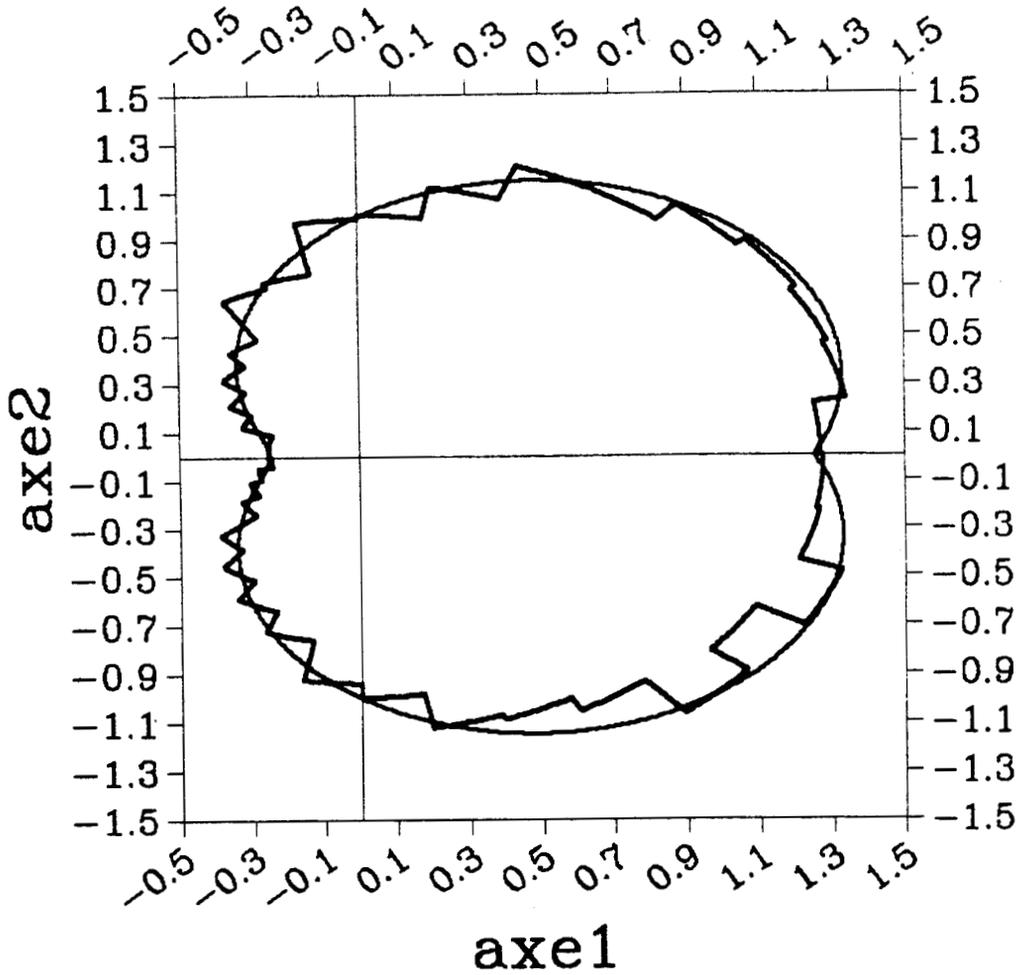
Le problème de la réduction du biais n'est pas abordé. Les figures proposées montrent au moins qu'on ne pourra donner de résultat qu'en faisant des hypothèses précises sur la décroissance de la densité de  $M$  au voisinage de la frontière de  $S$ .

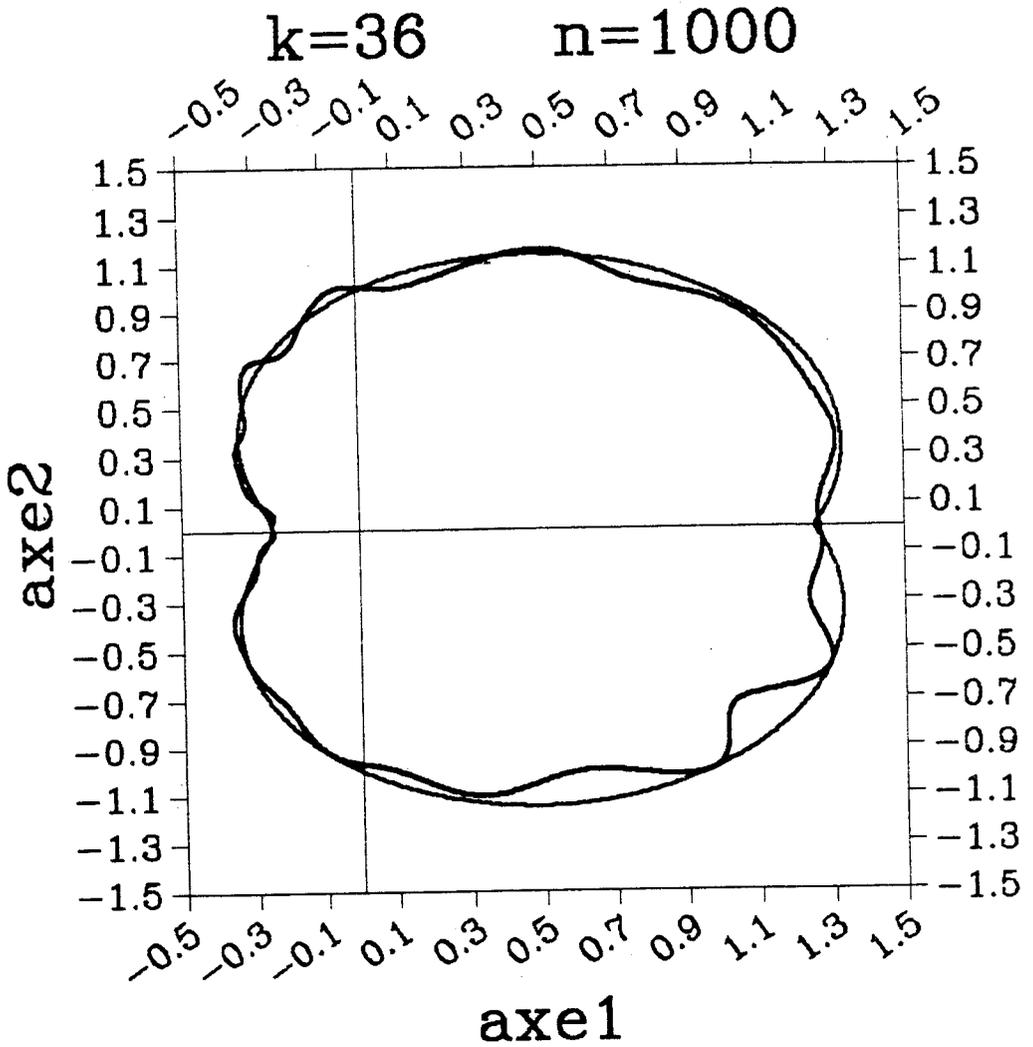
Nous espérons revenir sur ces questions dans un prochain article.

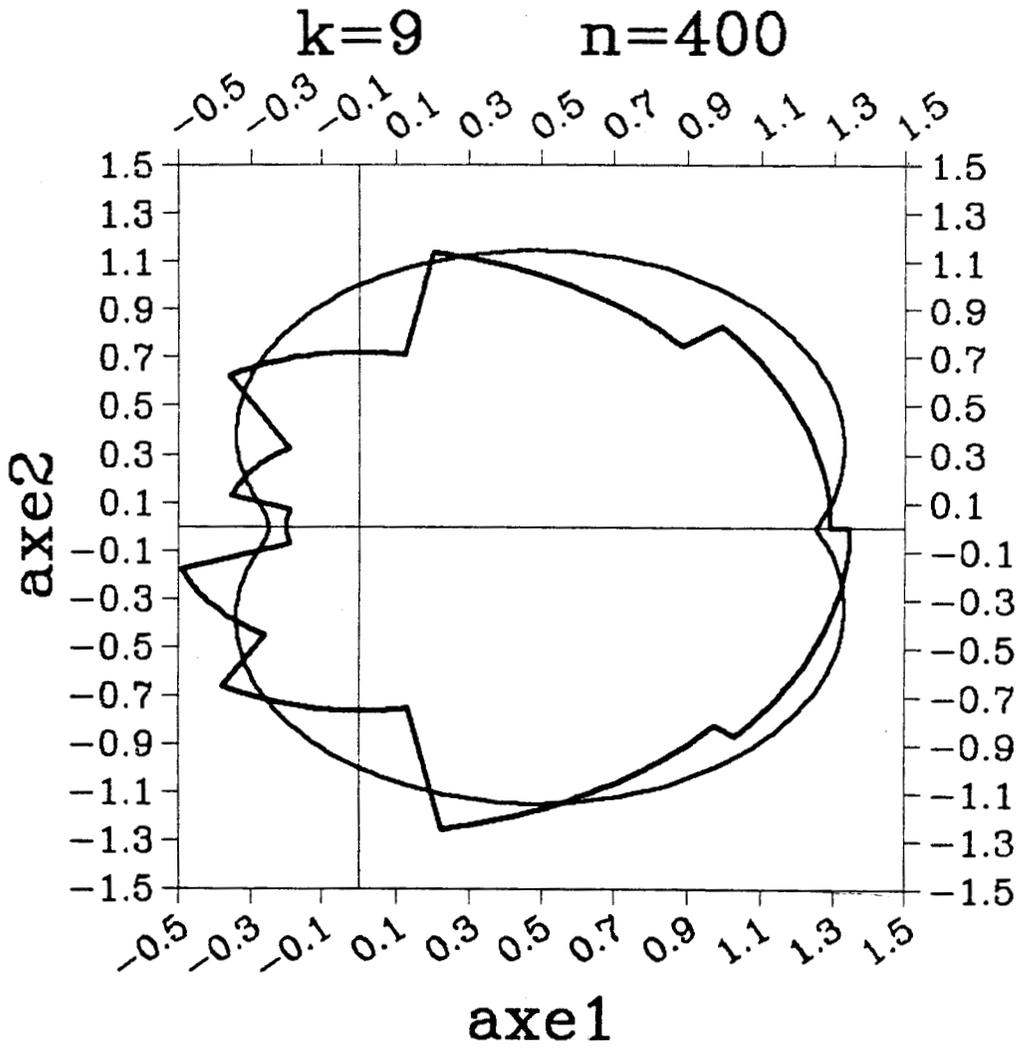


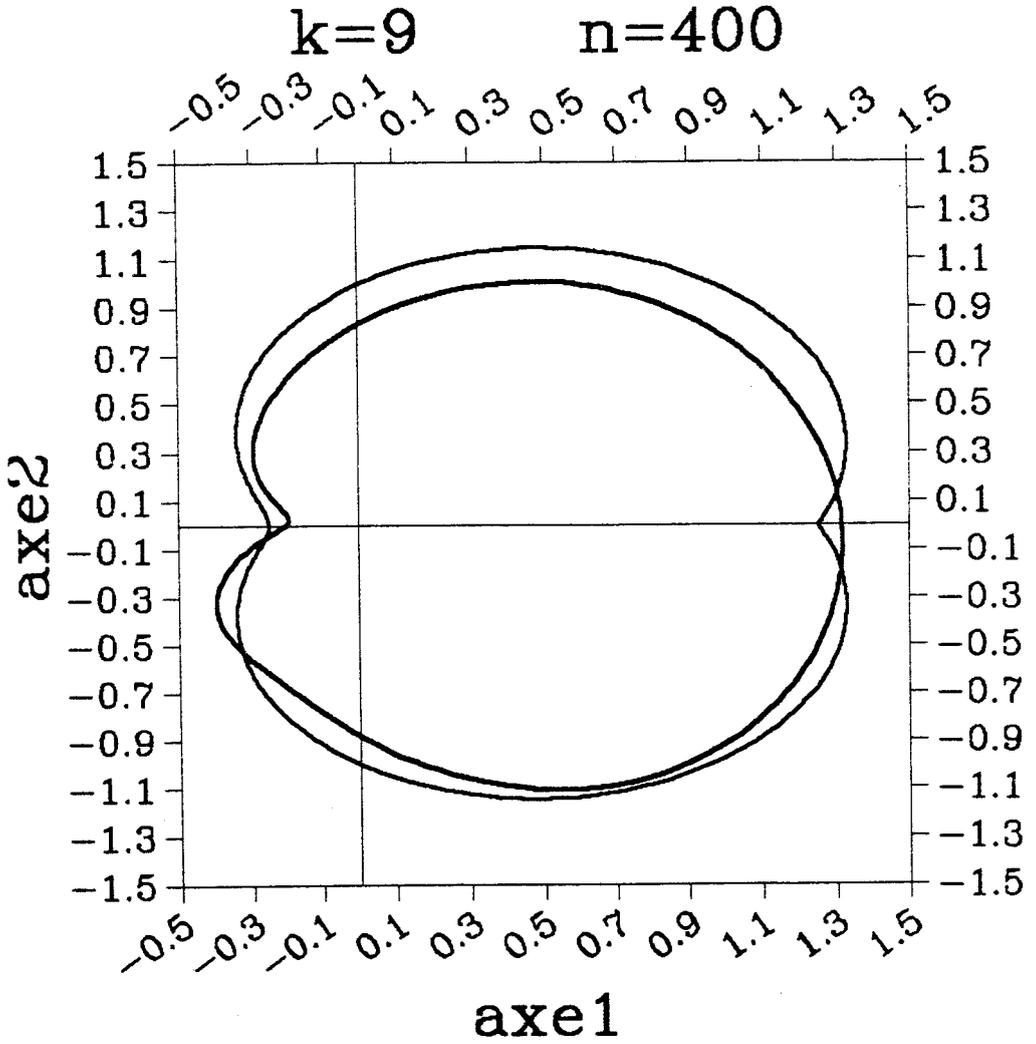


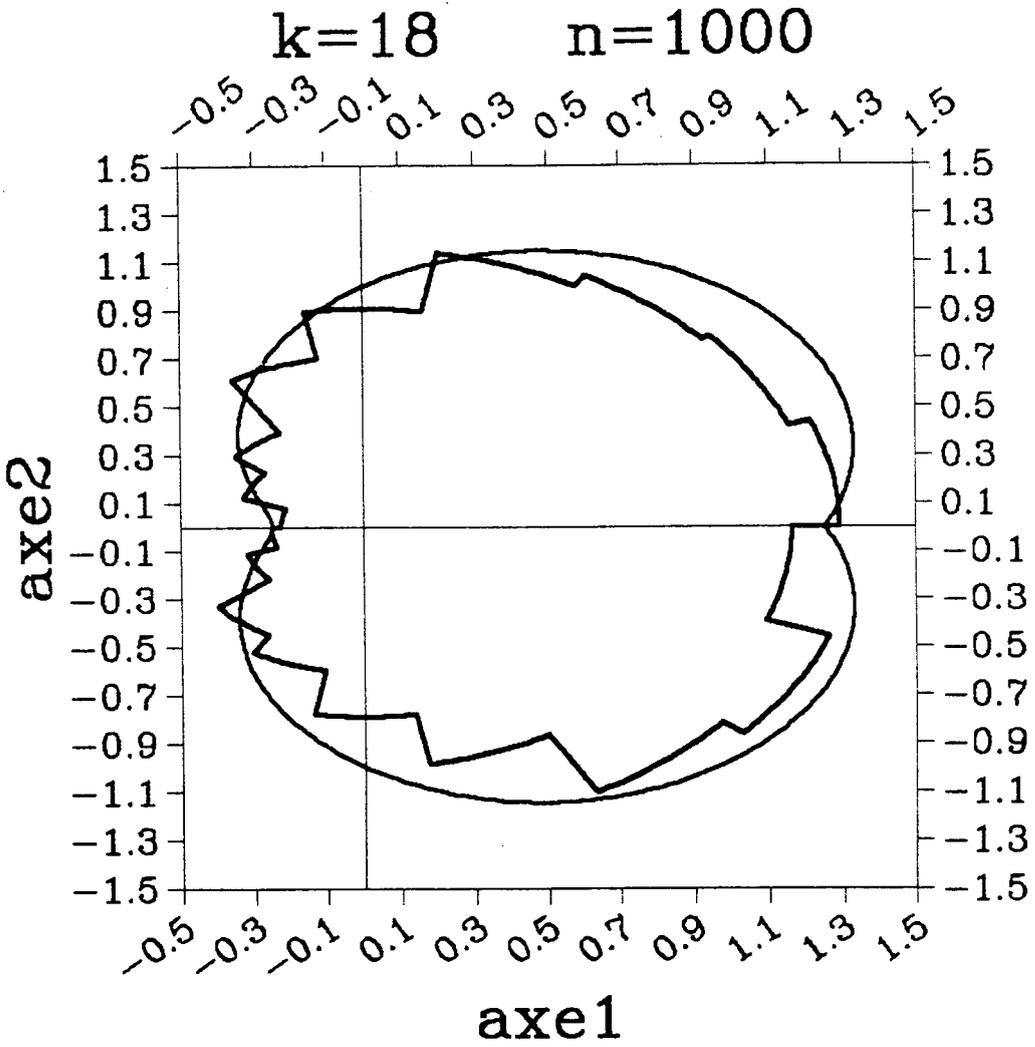
$k=36$        $n=1000$

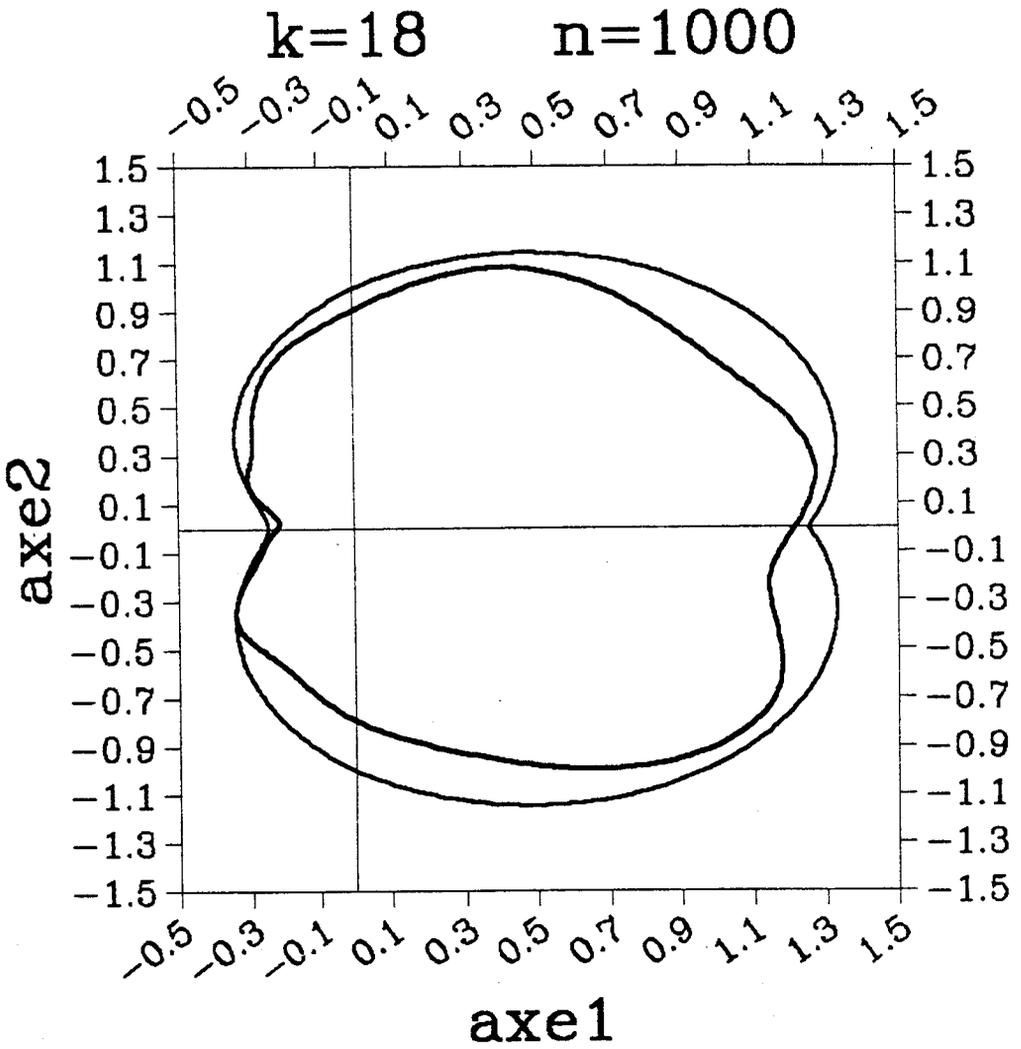














## R E F E R E N C E S.

- (1) AHLBERG, J.H. NILSON, E.N. and WALSH, J.L. *The theory of splines and their applications*. Academic Press, 1977.
- (2) BERLINET, A. *Convergence des estimateurs splines de la densité*. Pub. de l'I.S.U.P., Vol. 26, fasc. 2, 1981.
- (3) BOSQ, D. *Contribution à la théorie de l'estimation fonctionnelle*. Thèse, Université Pierre et Marie Curie, Paris, 1977.
- (4) CHEVALIER, J. *Estimation du support et du contenu du support d'une loi de probabilité*. Ann. Inst. Henri Poincaré, section B, Vol. XII, 4, 1976.
- (5) EFRON, B. *The convex hull of a random set of points*. Biometrika, 52, 331-343, 1965.
- (6) FISHER, L. *The convex hull of a sample*. Bull. An. Math. Soc., 72, 555-558, 1966.
- (7) GENSBITTEL, M.H. *Contribution à l'étude statistique de répartitions ponctuelles aléatoires*. Thèse, Université Pierre et Marie Curie, Paris, 1979.
- (8) GEFFROY, J. *Sur un problème d'estimation géométrique*. Pub. I.S.U.P., 13, 191-200, 1964.
- (9) GUILBART, C. *Etude de la continuité de l'application  $H$  qui à toute mesure de probabilité définie sur  $\mathbf{R}^n$  fait correspondre l'enveloppe convexe fermée du support de cette mesure*. C.R. Acad. Sci. Paris, 277, série A, 999-1002, 1973.
- (10) KALLENBERG, O. *Random measures*. Academic Press.
- (11) KARR, A.F. *Point processes and their statistical inference*. (Dekker), 1986.
- (12) KUELBS, J. *A strong convergence theorem for Banach space valued random variables*. Annals of probability, 4, 744-771, 1976.
- (13) JACOB, P. *Estimation du contour discontinu d'un processus ponctuel sur le plan*. Publications I.S.U.P., XXIX, 3-4, 1-25, 1984.

- 
- (14) JACOB, P. and ABBAR, H. *Estimating the edge of a Cox process area.* Proceedings of the 9<sup>th</sup> Franco-Belgian. Meeting of statisticians (Toulouse). 1988 (à paraître).
- (15) RAYNAUD, H. *Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans  $\mathbf{R}^n$ .* C.R. Acad. Sci. Paris, 262, série A, 627-623, 1965.
- (16) RENYI, A. and SULANKE, R. *Über die konvexe hülle von  $n$  zufällig gewählten punkten II.* Z. Wahrschein., 3, 138-147, 1964.
- (17) RIPLEY, B.D. and RASSON, J.P. *Finding the edge of a Poisson forest.* J. Appl. Pub., 14, 483-491, 1977.

## CHAPITRE IV

### ÉTUDE DE L'ERREUR QUADRATIQUE INTÉGRÉE ET CHOIX DE $k_n$ OPTIMAL

#### 1 - INTRODUCTION

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

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

En pratique, pour construire l'estimateur de  $\Phi$  de type histogramme étudié dans [12] il faut un nombre fini de subdivisions du domaine  $S$ . Pour régler au mieux ce problème, nous utilisons un moyen classique d'optimisation des estimateurs fonctionnels : c'est la minimisation de l'erreur quadratique intégrée.

#### 2 - PRÉLIMINAIRES

Soit  $(\Omega, \mathcal{F}, P)$  un espace probabilisé. On suppose que  $N$  est un processus de Cox dirigé par une mesure aléatoire  $M$  de densité aléatoire  $G$  strictement positive sur l'intérieur de  $S$  et continue sur  $S$ .

Soit  $N_1, \dots, N_n$   $n$  copies indépendantes du processus de Cox  $N$ . Pour tout  $\omega \in \Omega$ , on définit  $S_n^\omega$  comme le support de la mesure aléatoire discrète  $N_1^\omega + \dots + N_n^\omega$ . On considère une suite  $k = k_n$  d'entiers qui tend vers l'infini avec  $n$  et pour  $r \in \{1, \dots, k\}$  on pose :

$$(2.1) \quad I_r = I_{n,r} = [2\pi(r-1)k^{-1}, 2\pi rk^{-1}[$$

$$(2.2) \quad D_r = D_{n,r} = \{(\rho, \theta) \in S \mid \theta \in I_{n,r}\}$$

$$(2.3) \quad S_{n,r}^\omega = S_n^\omega \cap D_r$$

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

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

$$(2.6) \quad \alpha = \sup\{\Phi(\theta) \mid \theta \in [0, 2\pi]\}$$

$$(2.7) \quad m = \inf\{\Phi(\theta) \mid \theta \in [0, 2\pi]\}$$

$$(2.8) \quad \beta = \sup\{\Phi'(\theta) \mid \theta \in [0, 2\pi]\}$$

$$(2.9) \quad H_r(z) = H_{n,r}(z) = \{(\rho, \theta) \in D_r \mid \rho \geq z\}, \text{ pour tout } z \in [0, M_r].$$

Pour tout  $\omega \in \Omega$ , l'estimateur  $\Phi_n^\omega(\theta)$  de  $\Phi(\theta)$  étudié dans [12] est défini sur tout  $I_r$  par :

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

Dans la section 3, nous étudions la minimisation de l'erreur quadratique intégrée dans un cas particulier. Dans la section 4, nous établissons une majoration asymptotique de l'erreur quadratique intégrée dans le cas le plus général. Dans la section 5, nous présentons quelques exemples de choix de  $k_n$  optimal. Dans la section 6, nous étudions dans un cas particulier la minimisation de l'erreur quadratique intégrée à distance finie.

### 3 - CAS OÙ LE SUPPORT EST UN CERCLE

Dans cette section, on suppose que le support du processus ponctuel  $N$  est un cercle :  $\Phi(\theta) = \Phi, \forall \theta \in [0, 2\pi[$ .

Dans ce cas "limite", on se doute que le meilleur estimateur de type  $\Phi_n$  est obtenu pour  $k_n \equiv 1$  : c'est-à-dire la plus petite circonférence enveloppant le nuage de points. C'est d'ailleurs le résultat que l'on obtient en cherchant à minimiser l'erreur quadratique intégrée :

$$\begin{aligned} J_n(\Phi) &= \int_0^{2\pi} E(\Phi(\theta) - \Phi_n(\theta))^2 d\theta \\ &= \sum_{r=1}^k \int_{I_r} E(\Phi(\theta) - U_r)^2 d\theta \\ &= \sum_{r=1}^k \frac{2\pi}{k} E(\Phi - U_r)^2 \end{aligned}$$

En utilisant une intégration par parties, relative aux fonctions de répartition :

$$\begin{aligned} F_r(z) &= F_{n,r}(z) = P(U_{n,r} \leq z), \quad 0 \leq z \leq \Phi \\ \Gamma(z) &= \Phi^2 - (\Phi - z)^2, \quad 0 \leq z \leq \Phi \end{aligned}$$

on obtient :

$$\begin{aligned}
 E(\Phi - U_r)^2 &= \int_0^\Phi (\Phi - z)^2 dF_r(z) \\
 &= \Phi^2 - \int_0^\Phi \Gamma(z) dF_r(z) \\
 &= \Phi^2 - \left[ \Gamma(\Phi)F_r(\Phi) - \Gamma(0)F_r(0) - \int_0^\Phi F_r(z) d\Gamma(z) \right] \\
 &= 2 \int_0^\Phi (\Phi - z) F_r(z) dz
 \end{aligned}$$

Supposons que  $N$  soit un processus ponctuel de Poisson de mesure moyenne  $\mu$  uniforme sur le support :

$$\begin{aligned}
 F_r(z) &= \prod_{i=1}^n P\{N_i(H_r(z)) = 0\} \\
 &= e^{-n\mu(H_r(z))} = e^{-c\frac{n}{k}(\Phi^2 - z^2)}
 \end{aligned}$$

où  $c$  est une constante positive. On en déduit :

$$J_n(\Phi) = 4\pi \int_0^\Phi (\Phi - z) e^{-c\frac{n}{k}(\Phi^2 - z^2)} dz.$$

Il est clair que la plus petite erreur quadratique intégrée est obtenue pour  $k = 1$ . Cet exemple est pris volontairement très simple pour être explicite et rassurer sur la validité de la méthode.

#### 4 - MAJORATION ASYMPTOTIQUE DE L'ERREUR QUADRATIQUE INTÉGRÉE

Dans cette section, on introduit les notations suivantes :

$$(4.1) \quad \varepsilon = \varepsilon_n = 1 - k_n^{-1}$$

$$(4.2) \quad \varepsilon_n S = \{(\rho, \theta) \in S \mid \rho \leq \varepsilon_n \Phi(\theta)\} : \text{homothétique d'ordre } \varepsilon_n \text{ de } S$$

$$(4.3) \quad X_n = \inf\{G(\rho, \theta) \mid (\rho, \theta) \in \varepsilon_n S\}$$

$$J_n(\Phi) = \int_0^{2\pi} E(\Phi(\theta) - \Phi_n(\theta))^2 d\theta : \text{erreur quadratique intégrée.}$$

Et on suppose que  $G$  vérifie pour tout  $(\rho, \theta) \in S$  :

$$(4.5) \quad G(\rho, \theta) \leq B \text{ p.s.,}$$

où  $B$  est une variable aléatoire telle que  $E(B^2) < +\infty$ .

Si on suppose que l'une des hypothèses suivantes soit vérifiée :

$H_1$  : Il existe  $a > 0$  telle que  $E(X_n) \geq a$ .

$H_2$  :  $X_n < b$  où  $b < +\infty$ .

$H_3$  :  $X_n = \frac{A}{\delta_n}$  où  $\delta_n$  est une suite qui tend vers  $\infty$  avec  $n$  et  $A$  une variable aléatoire telle que  $E(A^2) < +\infty$ .

alors on a :

**Théorème 4.1.-** Si  $\Phi$  est dérivable, alors sous les conditions (i)  $k_n = o(n/\text{Log } n)$  et (ii)  $\frac{1}{E(X_n)} = o(\frac{n}{k_n})$  on a la majoration asymptotique suivante :

$$J_n(\Phi) \leq \frac{c_1}{nE(X_n)} + \frac{c_2 k_n^2}{n^2 E^2(X_n)} + \frac{c_3}{k_n^2}$$

où  $c_1, c_2$  et  $c_3$  sont des constantes qui dépendent de la forme de  $S$ .

Avant de démontrer ce théorème, on va d'abord démontrer le lemme utile suivant :

**Lemme 4.1.-** Si on pose  $Y_n = X_n \mathbf{1}_{\{X_n \leq \sqrt{k_n}\}}$  et si on suppose que l'une des hypothèses  $H_1$  ou  $H_2$  ou  $H_3$  soit vérifiée alors :

$$\frac{1}{nE(Y_n)} = \frac{1}{nE(X_n)} + o\left(\frac{1}{nE(X_n)}\right).$$

■  $\frac{1}{nE(Y_n)} - \frac{1}{nE(X_n)} = \frac{1}{nE(X_n)} \left( \frac{E(X_n)}{E(Y_n)} - 1 \right)$ , donc pour que le lemme soit démontré il suffit de montrer que  $\left( \frac{E(X_n)}{E(Y_n)} - 1 \right)$  tend vers 0 avec  $\frac{1}{n}$ .

Si  $H_1$  est vérifiée alors :

$$E(X_n) - E(Y_n) = E[X_n \mathbf{1}_{\{X_n > \sqrt{k_n}\}}] < E[G \mathbf{1}_{\{G > \sqrt{k_n}\}}]$$

et  $E(G \mathbf{1}_{\{G > \sqrt{k_n}\}})$  tend vers 0 avec  $\frac{1}{n}$  car, d'après (4.5),  $G$  est intégrable, donc  $\liminf E(Y_n) \geq a > 0$  et  $E(Y_n) \left( \frac{E(X_n)}{E(Y_n)} - 1 \right) = E(X_n) - E(Y_n)$  tend vers 0 avec  $\frac{1}{n}$  ce qui implique que  $\left( \frac{E(X_n)}{E(Y_n)} - 1 \right)$  tend vers 0 avec  $\frac{1}{n}$ .

Si  $H_2$  est vérifiée alors :

$\left( \frac{E(X_n)}{E(Y_n)} - 1 \right)$  tend vers 0 avec  $\frac{1}{n}$  car pour  $n$  assez grand  $\sqrt{k_n} > b$  et donc  $X_n = Y_n$ .

Si  $H_3$  est vérifiée alors :

$\left( \frac{E(X_n)}{E(Y_n)} - 1 \right)$  tend vers 0 avec  $\frac{1}{n}$  car  $E(Y_n) = E\left[\frac{A}{\delta_n} \mathbf{1}_{\{A \leq \delta_n \sqrt{k_n}\}}\right]$  donc  $\frac{E(X_n)}{E(Y_n)} = \frac{E(A)}{E[A \mathbf{1}_{\{A \leq \delta_n \sqrt{k_n}\}}]}$  tend vers 1 quand  $n$  tend vers l'infini car  $A$  est intégrable. ■

**Démonstration du théorème 4.1.-**

$$\begin{aligned}
J_n(\Phi) &= \int_0^{2\pi} E(\Phi(\theta) - \Phi_n(\theta))^2 d\theta \\
(4.6) \qquad &= \sum_{r=1}^k \int_{I_r} E(\Phi(\theta) - U_r)^2 d\theta
\end{aligned}$$

En utilisant une intégration par partie relative à la fonction de répartition :  $F_r(z) = F_{n,r}(z) = P\{U_{n,r} \leq z\}$  pour  $z \in [0, M_r]$ , et à la fonction :  $\Gamma(z) = (\Phi(\theta) - z)^2$ , pour  $z \in [0, M_r]$ , qu'on peut écrire comme la somme de deux fonctions de répartition l'une croissante et l'autre décroissante, on obtient :

$$\begin{aligned}
E(\Phi(\theta) - U_r)^2 &= \int_0^{M_r} (\Phi(\theta) - z)^2 dF_r(z) \\
(4.7) \qquad &= (\Phi(\theta) - M_r)^2 F_r(M_r) - \Phi^2(\theta) F_r(0) + 2 \int_0^{M_r} (\Phi(\theta) - z) F_r(z) dz
\end{aligned}$$

Puisque  $N$  est un processus de Cox dirigé par la mesure aléatoire  $M$  alors :

$$\begin{aligned}
F_r(z) &= \prod_{i=1}^n P[\{N_i(H_r(z)) = 0\}] \\
(4.8) \qquad &= E^n[e^{-M(H_r(z))}]
\end{aligned}$$

D'après (4.7) et (4.8) on a :

$$(4.9) \qquad E(\Phi(\theta) - U_r)^2 = (\Phi(\theta) - M_r)^2 - \Phi^2(\theta) E^n(e^{-M(D_r)}) + A_r + B_r,$$

avec :

$$\begin{aligned}
A_r &= \int_0^{\varepsilon m_r} 2(\Phi(\theta) - z) E^n[e^{-M(H_r(z))}] dz \\
B_r &= \int_{\varepsilon m_r}^{M_r} 2(\Phi(\theta) - z) E^n[e^{-M(H_r(z))}] dz
\end{aligned}$$

– **Majoration de  $A_r$  :** Si on considère la variable aléatoire :

$$(4.10) \qquad Y_n = X_n \mathbf{1}_{\{X_n \leq \sqrt{k_n}\}}$$

alors on obtient pour  $1 \leq r \leq k$  et  $0 \leq z \leq \varepsilon m_r$  :

$$\begin{aligned}
M(H_r(z)) &\geq X_n \pi (\varepsilon^2 m_r^2 - z^2) k^{-1} \\
(4.11) \qquad &\geq Y_n \pi (\varepsilon^2 m_r^2 - z^2) k^{-1}
\end{aligned}$$

Puisque  $Y_n k^{-1} \leq k^{-1/2}$ , l'inégalité :

$\exp\{-Y_n \pi(\varepsilon^2 m_r^2 - z^2) k^{-1}\} \leq 1 - \frac{Y_n \pi(\varepsilon^2 m_r^2 - z^2) k^{-1}}{2}$  est vérifiée de façon déterministe pour  $n$  assez grand et donc, d'après (4.11), on obtient :

$$\begin{aligned} E^n[e^{-M(H_r(z))}] &\leq \left(1 - \frac{E(Y_n) \pi(\varepsilon^2 m_r^2 - z^2) k^{-1}}{2}\right)^n \\ &\leq \exp\left\{-\frac{n E(Y_n) \pi(\varepsilon^2 m_r^2 - z^2) k^{-1}}{2}\right\}, \end{aligned}$$

d'où

$$A_r \leq \int_0^{\varepsilon m_r} 2(\Phi(\theta) - z) \exp\left\{-\frac{n E(Y_n) \pi \varepsilon m_r (m_r - z) k^{-1}}{2}\right\} dz,$$

puisque  $(\varepsilon^2 m_r^2 - z^2) \geq \varepsilon m_r (\varepsilon m_r - z)$ .

Après intégration on obtient :

$$\begin{aligned} A_r &\leq \frac{4k(\Phi(\theta) - \varepsilon m_r)}{n E(Y_n) \pi \varepsilon m_r} + \frac{8k^2}{n^2 E^2(Y_n) \pi^2 \varepsilon^2 m_r^2} \\ &\quad - \frac{4k\Phi(\theta)}{n E(Y_n) \pi \varepsilon m_r} \exp\left\{-\frac{n E(Y_n) \pi \varepsilon^2 m_r^2 k^{-1}}{2}\right\} \\ &\quad - \frac{8k^2}{n^2 E^2(Y_n) \pi^2 \varepsilon^2 m_r^2} \exp\left\{-\frac{n E(Y_n) \pi \varepsilon^2 m_r^2 k^{-1}}{2}\right\} \end{aligned}$$

Puisque  $\frac{n E(X_n)}{k_n}$  tend vers l'infini, d'après (ii), et  $\varepsilon \geq \frac{1}{2}$  pour  $n$  assez grand, d'après (4.1), et  $\Phi(\theta) - \varepsilon m_r \leq (\alpha + \beta) k^{-1}$ , alors on a :

$$A_r \leq \frac{8(\alpha + \beta)}{n E(Y_n) \pi m} + \frac{32k^2}{n^2 E^2(Y_n) \pi^2 m^2} + o\left(\frac{k^2}{n^2 E^2(Y_n)}\right).$$

Sous les hypothèses  $H_1$ ,  $H_2$  ou  $H_3$  et d'après le lemme 4.1, on peut remplacer  $\frac{8(\alpha + \beta)}{n E(Y_n) \pi m}$  par  $\frac{8(\alpha + \beta)}{n E(X_n) \pi m} + o\left(\frac{1}{n E(X_n)}\right)$  et  $\frac{32k^2}{n^2 E^2(Y_n) \pi^2 m^2}$  par  $\frac{32k^2}{n^2 E^2(X_n) \pi^2 m^2} + o\left(\frac{k^2}{n^2 E^2(X_n)}\right)$  donc :

$$(4.12) \quad A_r \leq \frac{8(\alpha + \beta)}{n E(X_n) \pi m} + \frac{32k^2}{n^2 E^2(X_n) \pi^2 m^2} + o\left(\frac{1}{n E(X_n)}\right) + o\left(\frac{k^2}{n^2 E^2(X_n)}\right).$$

– **Majoration de  $|B_r|$**  : En remarquant que pour  $\varepsilon m_r \leq z \leq M_r$  :  $|\Phi(\theta) - z| \leq M_r - \varepsilon m_r$ , on obtient :

$$|B_r| \leq 2(M_r - \varepsilon m_r)^2 = 2(M_r - m_r + m_r k^{-1})^2$$

et donc :

$$(4.13) \quad |B_r| \leq 2(\alpha + \beta)^2 k^{-2}$$

Il est facile de voir que :

$$(4.14) \quad (\Phi(\theta) - M_r)^2 \leq \beta^2 k^{-2}$$

En faisant le même raisonnement que dans le début de la démonstration de la majoration de  $A_r$  et en remarquant que :  $M(D_r) \geq X_n \pi m^2 k^{-1}$ , on obtient :

$$\Phi^2 E^n [e^{-M(D_r)}] \leq \alpha^2 \exp\left\{-\frac{nE(Y_n)\pi m^2 k^{-1}}{2}\right\}.$$

Donc :

$$(4.15) \quad \Phi^2 E^n [e^{-M(D_r)}] = o\left(\frac{k^2}{n^2 E^2(X_n)}\right).$$

Finalement et d'après (4.15), (4.14), (4.13), (4.12), (4.9) et (4.6) et pour  $n$  assez grand on obtient :

$$J_n(\Phi) \leq \frac{C_1}{nE(X_n)} + \frac{C_2 k_n^2}{n^2 E^2(X_n)} + \frac{C_3}{k_n^2}$$

avec  $C_1 = \frac{16(\alpha+\beta)}{m}$ ,  $c_2 = \frac{64}{\pi m^2}$  et  $c_3 = 2\pi[\beta^2 + 2(\alpha + \beta)^2]$ .

## 5 - EXEMPLES DE CHOIX DE $k_n$ OPTIMAUX

Nous présentons, dans cette section, trois exemples de choix de  $k_n$  optimal. Dans le premier, on suppose que  $N$  est un processus de Cox homogène. Dans le second, on considère que  $N$  est un processus de Cox dirigé par une mesure aléatoire dont la densité ne s'annule pas sur la frontière de  $S$ . Dans le troisième,  $N$  est un processus de Poisson mélangé d'intensité supposée arbitrairement petite au voisinage de la frontière de  $S$ .

**5.1.-** Soit  $N$  un processus de Cox dirigé par une mesure aléatoire  $M = \alpha \lambda$  où  $\alpha$  est une variable aléatoire telle que  $E(\alpha^2) < +\infty$ .

Dans ce cas :  $E(X_n) = E(\alpha)$  donc  $\frac{1}{E(X_n)} = o\left(\frac{n}{k_n}\right)$  puisque  $k_n = o(n/\text{Log } n)$ . L'hypothèse  $H_1$  est vérifiée donc d'après le théorème 4.1 et pour  $n$  assez grand on a :

$$J_n(\Phi) \leq \frac{c_1}{E(\alpha)n} + \frac{c_2 k_n^2}{E^2(\alpha)n^2} + \frac{c_3}{k_n^2}.$$

Il est clair que, dans ce cas, le  $k_n$  optimal est de l'ordre de  $n^{1/2}$ .

**5.2.-** Soit  $N$  un processus de Cox dirigé par une mesure aléatoire  $M$  qui admet une densité aléatoire  $G$  telle que  $0 < A \leq G$  où  $A$  est une variable aléatoire telle que  $E(A^2) < +\infty$ .

Dans ce cas :  $E(X_n) \geq E(A)$  donc  $H_1$  est vérifiée et  $\frac{1}{E(X_n)} = o(\frac{n}{k_n})$  et donc d'après le théorème 4.1. et pour  $n$  assez grand on a :

$$J_n(\Phi) \leq \frac{C_1}{E(A)n} + \frac{C_2 k_n^2}{E^2(A)n^2} + \frac{C_3}{k_n^2}.$$

Dans cet exemple, il est clair aussi que le  $k_n$  optimal est de l'ordre de  $n^{1/2}$ .

**5.3.-** Soit  $N$  un processus de Poisson mélangé d'intensité  $G$  définie par :

$$G(\rho, \theta) = A \frac{(\Phi(\theta) - \rho)}{\pi \Phi^2(\theta)}, \text{ pour tout } (\rho, \theta) \in S,$$

où  $A$  est une variable aléatoire telle que  $E(A^2) < +\infty$ .

Dans ce cas :  $E(X_n) = \frac{E(A)}{\pi \alpha} k_n^{-1}$  donc  $H_3$  est vérifiée et  $\frac{1}{E(X_n)} = \frac{\pi \alpha}{E(A)} k_n$ . Si on suppose de plus que  $k_n^2 = o(n)$  alors  $\frac{1}{E(X_n)} = o(\frac{n}{k_n})$  et donc d'après le théorème 4.1. et pour  $n$  assez grand on obtient :

$$J_n(\Phi) \leq \frac{\pi \alpha c_1 k_n}{E(A)n} + \frac{\pi^2 \alpha^2 c_2 k_n^4}{E^2(A)n^2} + \frac{c_3}{k_n^2}.$$

contrairement aux deux exemples précédents, dans cet exemple le  $k_n$  optimal est de l'ordre de  $n^{1/3}$ .

## 6 - ÉTUDE DE L'ERREUR QUADRATIQUE INTÉGRÉE À DISTANCE FINIE DANS UN CAS PARTICULIER.

Dans le cas général, on a établi des résultats qui sont malheureusement asymptotiques. Dans cette section, on va étudier l'erreur quadratique intégrée à distance finie dans un cas particulier pour voir le comportement des  $k_n$  optimaux et de l'erreur quadratique intégrée pour les petits échantillons.

Pour simplifier les calculs, on suppose que le processus ponctuel  $N$  est un processus de Poisson homogène se réalisant dans un domaine  $S = \{(x, y) / 0 \leq x \leq 1 \text{ et } 0 \leq y < \Phi(x)\}$  où  $\Phi(x) = ax + b$  pour tout  $x \in [0, 1]$  et  $a < 0$ .

On introduit les notations suivantes :

$\Phi_n$ : l'estimateur de type histogramme de  $\Phi$

$$H_r(z) = \{(x, y) \in S / x \in [\frac{r-1}{k}, \frac{r}{k}[ \text{ et } y \geq z\} \text{ pour } 1 \leq r \leq k$$

$$J_n = \int_0^1 E(\Phi(x) - \Phi_n(x))^2 dx : \text{l'erreur quadratique intégrée.}$$

### 6.1.- Calcul de l'erreur quadratique intégrée à distance finie.

Puisque le processus ponctuel  $N$  est un processus de Poisson homogène alors pour  $1 \leq r \leq k$  on a :

$$F_{U_r}(z) = P\{U_r \leq z\} = e^{-nc\lambda(H_r(z))}.$$

Pour  $n$  fixé on a :

si  $0 \leq z \leq m_r$  alors :

$$\begin{aligned}\lambda(H_r(z)) &= \frac{m_r - z}{k} + \frac{M_r - m_r}{2k} \\ &= -\frac{1}{k}\left[z - \left(\frac{m_r + M_r}{2}\right)\right]\end{aligned}$$

en posant  $d = \frac{nc}{k}$  et  $e_r = \frac{m_r + M_r}{2}$  on a :

pour  $1 \leq r \leq k$  et  $0 \leq z \leq m_r$  :

$$\begin{aligned}F_{U_r}(z) &= e^{d(z-e_r)} \\ f_{U_r}(z) &= de^{d(z-e_r)} : \text{densité de } U_r\end{aligned}$$

si  $m_r < z \leq M_r$  alors :

$$\lambda(H_r(z)) = \frac{(M_r - z)^2}{2k(M_r - m_r)}$$

en posant  $C_r = \frac{nc}{2k(m_r - M_r)}$  on a :

pour  $1 \leq r \leq k$  et  $m_r < z \leq M_r$  :

$$\begin{aligned}F_{U_r}(z) &= e^{C_r(z-M_r)^2} \\ f_{U_r}(z) &= 2C_r(z - M_r)e^{C_r(z-M_r)^2}.\end{aligned}$$

$$\begin{aligned}(6.1) \quad J_n &= \sum_{r=1}^k \int_{I_r} E(\Phi(x) - U_r)^2 dx \\ &= \sum_{r=1}^k \left[ \int_{I_r} \Phi^2(x) dx - 2E(U_r) \int_{I_r} \Phi(x) dx + \frac{1}{k} E(U_r^2) \right]\end{aligned}$$

avec :

$$\begin{aligned}E(U_r) &= d \int_0^{m_r} ze^{d(z-e_r)} dz + 2C_r \int_{m_r}^{M_r} (z - M_r) ze^{C_r(z-M_r)^2} dz \\ E(U_r^2) &= d \int_0^{m_r} z^2 e^{d(z-e_r)} dz + 2C_r \int_{m_r}^{M_r} (z - M_r) z^2 e^{C_r(z-M_r)^2} dz\end{aligned}$$

En calculant les intégrales et en remplaçant  $d$ ,  $e_r$  et  $C_r$  par leurs expressions on obtient :

$$(6.2) \quad E(U_r) = M_r - \frac{k}{nc} e^{anc/2k^2} + \frac{k}{nc} e^{-\frac{nc}{k}(\frac{m_r+M_r}{2})} - \sqrt{\frac{\pi a}{2nc}} (e^{anc/2k^2} - 1)$$

$$(6.3) \quad E(U_r^2) = M_r^2 - \left( \frac{2m_r k}{nc} - \frac{2k^2}{(nc)^2} - \frac{2a}{nc} \right) e^{anc/2k^2} - \frac{2k^2}{(nc)^2} e^{-\frac{nc}{k}(\frac{m_r+M_r}{2})} \\ - \frac{2a}{nc} - 2M_r \sqrt{\frac{\pi a}{nc}} (e^{anc/2k^2} - 1)$$

Il est facile de voir que :

$$(6.4) \quad \int_{I_r} \Phi(x) dx = \frac{M_r}{k} + \frac{a}{2k^2}$$

et

$$(6.5) \quad \int_{I_r} \Phi^2(x) dx = \frac{a^2}{3k^3} + \frac{aM_r}{k^2} + \frac{M_r^2}{k}$$

D'après (6.2), (6.3), (6.4) et (6.5) on obtient :

$$(6.6) \quad \int_{I_r} \Phi^2(x) dx - 2E(U_r) \int_{I_r} \Phi(x) dx + \frac{1}{k} E(U_r^2) \\ = \frac{2k}{(nc)^2} e^{anc/2k^2} - \left( \frac{2M_r}{nc} + \frac{a}{knc} + \frac{2k}{(nc)^2} \right) e^{-\frac{nc}{k}(\frac{m_r+M_r}{2})} \\ - \frac{a}{knc} + \frac{a^2}{3k^3} + \frac{a}{k^2} \sqrt{\frac{\pi a}{2nc}} (e^{anc/2k^2} - 1)$$

D'après (6.6) et (6.1) on obtient :

$$J_n = \frac{2k^2}{(nc)^2} e^{anc/2k^2} + \frac{2k^2}{a(nc)^2} e^{-anc/2k^2} \sum_{r=1}^k \left[ \left( -\frac{a}{k} \right) \left( \frac{ncM_r}{k} + \frac{anc}{2k^2} + 1 \right) e^{-\frac{nc}{k} M_r} \right] \\ - \frac{a}{nc} + \frac{a^2}{3k^2} + \frac{a}{k} \sqrt{\frac{\pi a}{2nc}} (e^{anc/2k^2} - 1) \\ \# \frac{2k^2}{(nc)^2} e^{anc/2k^2} + \frac{2k^2}{a(nc)^2} e^{-anc/2k^2} \int_{m-\frac{a}{k}}^{m-a} \left( \frac{nc}{k} x + \frac{anc}{2k^2} + 1 \right) e^{-\frac{nc}{k} x} dx \\ - \frac{a}{nc} + \frac{a^2}{3k^2} + \frac{a}{k} \sqrt{\frac{\pi a}{2nc}} (e^{anc/2k^2} - 1)$$

Après calcul de l'intégrale on obtient finalement :

$$J_n \# \frac{2k^2}{(nc)^2} e^{anc/2k^2} + \frac{2k^2}{a(nc)^2} \left( m - \frac{a}{2k} + \frac{2k}{nc} \right) e^{-\frac{nc}{k} \left( m - \frac{a}{2k} \right)} \\ - \frac{2k^2}{a(nc)^2} \left( m - a + \frac{a}{2k} + \frac{2k}{nc} \right) e^{-\frac{nc}{k} \left( m - a + \frac{a}{2k} \right)} - \frac{a}{nc} + \frac{a^2}{3k^2} \\ + \frac{a}{k} \sqrt{\frac{\pi a}{2nc}} (e^{anc/2k^2} - 1)$$

**6.2.- Minimisation numérique de  $J_n$  et  $k_n$  optimaux.**

Pour  $m = 2$ , pour différentes valeurs de la pente  $a$  comprises entre  $-0,01$  et  $-10$  et pour un nombre moyen de points égal à 50, 100, 500 et 1000 ; on a minimisé numériquement l'erreur quadratique intégrée  $J_n$ . Les valeurs des  $k_n$  optimaux et de  $J_n$  minimales sont dans les tableaux qui suivent. Dans ces tableaux on a mis aussi les erreurs quadratiques intégrées  $J_n^*$  calculées en remplaçant dans  $J_n$ , pour chaque valeur de  $a$ ,  $k_n$  par les  $k_n$  optimaux des autres valeurs de  $a$ . D'après les valeurs de ces  $J_n^*$ , on constate que si on se trompe légèrement sur les  $k_n$  optimaux l'erreur quadratique intégrée  $J_n$  reste encore faible mais dans le cas contraire  $J_n$  prend des valeurs assez grande.

$$m = 2 \quad a = -0,01$$

			(-0,1)	(-0,5)	(-2)	(-4)	(-8)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	1	$0,77 \cdot 10^{-3}$	$0,77 \cdot 10^{-3}$	$0,32 \cdot 10^{-2}$	$0,13 \cdot 10^{-1}$	$0,29 \cdot 10^{-1}$	$0,65 \cdot 10^{-1}$	$0,80 \cdot 10^{-1}$
100	1	$0,18 \cdot 10^{-3}$	$0,18 \cdot 10^{-3}$	$0,18 \cdot 10^{-2}$	$0,72 \cdot 10^{-2}$	$0,16 \cdot 10^{-1}$	$0,29 \cdot 10^{-1}$	$0,39 \cdot 10^{-1}$
500	1	$0,29 \cdot 10^{-6}$	$0,69 \cdot 10^{-4}$	$0,39 \cdot 10^{-3}$	$0,16 \cdot 10^{-2}$	$0,29 \cdot 10^{-2}$	$0,58 \cdot 10^{-2}$	$0,77 \cdot 10^{-2}$
1000	1	$0,12 \cdot 10^{-6}$	$0,30 \cdot 10^{-4}$	$0,20 \cdot 10^{-3}$	$0,72 \cdot 10^{-3}$	$0,14 \cdot 10^{-2}$	$0,30 \cdot 10^{-2}$	$0,37 \cdot 10^{-2}$

$$m = 2 \quad a = -0,1$$

			(-0,01)	(-0,5)	(-2)	(-4)	(-8)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	1	$0,29 \cdot 10^{-4}$	$0,29 \cdot 10^{-4}$	$0,26 \cdot 10^{-2}$	$0,13 \cdot 10^{-1}$	$0,29 \cdot 10^{-1}$	$0,65 \cdot 10^{-1}$	$0,80 \cdot 10^{-1}$
100	1	$0,12 \cdot 10^{-4}$	$0,38 \cdot 10^{-3}$	$0,15 \cdot 10^{-2}$	$0,71 \cdot 10^{-2}$	$0,16 \cdot 10^{-1}$	$0,29 \cdot 10^{-1}$	$0,39 \cdot 10^{-1}$
500	3	$0,23 \cdot 10^{-5}$	$0,18 \cdot 10^{-2}$	$0,34 \cdot 10^{-3}$	$0,15 \cdot 10^{-2}$	$0,29 \cdot 10^{-2}$	$0,58 \cdot 10^{-2}$	$0,77 \cdot 10^{-2}$
1000	4	$0,12 \cdot 10^{-5}$	$0,22 \cdot 10^{-2}$	$0,18 \cdot 10^{-3}$	$0,71 \cdot 10^{-3}$	$0,14 \cdot 10^{-2}$	$0,30 \cdot 10^{-2}$	$0,37 \cdot 10^{-2}$

$$m = 2 \quad a = -0,5$$

			(-0,01)	(-0,1)	(-2)	(-4)	(-8)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	2	$0,12 \cdot 10^{-3}$	$0,31 \cdot 10^{-1}$	$0,31 \cdot 10^{-1}$	$0,95 \cdot 10^{-2}$	$0,27 \cdot 10^{-1}$	$0,64 \cdot 10^{-1}$	$0,79 \cdot 10^{-1}$
100	3	$0,59 \cdot 10^{-4}$	$0,44 \cdot 10^{-1}$	$0,44 \cdot 10^{-1}$	$0,57 \cdot 10^{-2}$	$0,15 \cdot 10^{-1}$	$0,28 \cdot 10^{-1}$	$0,39 \cdot 10^{-1}$
500	7	$0,12 \cdot 10^{-4}$	$0,64 \cdot 10^{-1}$	$0,36 \cdot 10^{-2}$	$0,13 \cdot 10^{-2}$	$0,27 \cdot 10^{-2}$	$0,57 \cdot 10^{-2}$	$0,76 \cdot 10^{-2}$
1000	10	$0,59 \cdot 10^{-5}$	$0,70 \cdot 10^{-1}$	$0,22 \cdot 10^{-2}$	$0,57 \cdot 10^{-3}$	$0,14 \cdot 10^{-2}$	$0,30 \cdot 10^{-2}$	$0,36 \cdot 10^{-2}$

$$m = 2 \quad a = -2$$

			(-0,01)	(-0,1)	(-0,5)	(-4)	(-8)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	4	$0,47 \cdot 10^{-3}$	0,87	0,87	0,12	$0,12 \cdot 10^{-1}$	$0,54 \cdot 10^{-1}$	$0,70 \cdot 10^{-1}$
100	6	$0,23 \cdot 10^{-3}$	0,99	0,99	$0,50 \cdot 10^{-1}$	$0,80 \cdot 10^{-2}$	$0,23 \cdot 10^{-1}$	$0,34 \cdot 10^{-1}$
500	14	$0,47 \cdot 10^{-4}$	$0,12 \cdot 10^1$	$0,99 \cdot 10^{-1}$	$0,86 \cdot 10^{-2}$	$0,12 \cdot 10^{-2}$	$0,46 \cdot 10^{-2}$	$0,67 \cdot 10^{-2}$
1000	19	$0,23 \cdot 10^{-4}$	$0,12 \cdot 10^1$	$0,57 \cdot 10^{-1}$	$0,41 \cdot 10^{-2}$	$0,61 \cdot 10^{-3}$	$0,24 \cdot 10^{-2}$	$0,32 \cdot 10^{-2}$

$$m = 2 \quad a = -4$$

			(-0,01)	(-0,1)	(-0,5)	(-2)	(-8)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	6	$0,94 \cdot 10^{-3}$	$0,40 \cdot 10^1$	$0,40 \cdot 10^1$	0,70	$0,59 \cdot 10^{-1}$	$0,32 \cdot 10^{-1}$	$0,50 \cdot 10^{-1}$
100	9	$0,47 \cdot 10^{-3}$	$0,44 \cdot 10^1$	$0,44 \cdot 10^1$	0,30	$0,21 \cdot 10^{-1}$	$0,12 \cdot 10^{-1}$	$0,24 \cdot 10^{-1}$
500	19	$0,94 \cdot 10^{-4}$	$0,49 \cdot 10^1$	0,45	$0,53 \cdot 10^{-1}$	$0,53 \cdot 10^{-2}$	$0,24 \cdot 10^{-2}$	$0,46 \cdot 10^{-2}$
1000	27	$0,47 \cdot 10^{-4}$	$0,50 \cdot 10^1$	0,26	$0,26 \cdot 10^{-1}$	$0,21 \cdot 10^{-2}$	$0,14 \cdot 10^{-2}$	$0,21 \cdot 10^{-2}$

$$m = 2 \quad a = -8$$

			(-0,01)	(-0,1)	(-0,5)	(-2)	(-4)	(-10)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	9	$0,19 \cdot 10^{-2}$	$0,17 \cdot 10^2$	$0,17 \cdot 10^2$	$0,35 \cdot 10^1$	0,49	$0,85 \cdot 10^{-1}$	$0,11 \cdot 10^{-1}$
100	12	$0,94 \cdot 10^{-3}$	$0,18 \cdot 10^2$	$0,18 \cdot 10^2$	$0,15 \cdot 10^1$	0,20	$0,30 \cdot 10^{-1}$	$0,50 \cdot 10^{-2}$
500	27	$0,19 \cdot 10^{-3}$	$0,20 \cdot 10^2$	$0,20 \cdot 10^1$	0,27	$0,34 \cdot 10^{-1}$	$0,85 \cdot 10^{-2}$	$0,88 \cdot 10^{-3}$
1000	39	$0,94 \cdot 10^{-4}$	$0,20 \cdot 10^2$	$0,11 \cdot 10^1$	0,13	$0,20 \cdot 10^{-1}$	$0,41 \cdot 10^{-2}$	$0,34 \cdot 10^{-3}$

$$m = 2 \quad a = -10$$

			(-0,01)	(-0,1)	(-0,5)	(-2)	(-4)	(-8)
nc	$k_n$ optimal	$J_n$ optimal	$J_{n_1}^*$	$J_{n_2}^*$	$J_{n_3}^*$	$J_{n_4}^*$	$J_{n_5}^*$	$J_{n_6}^*$
50	10	$0,23 \cdot 10^{-2}$	$0,28 \cdot 10^2$	$0,28 \cdot 10^2$	$0,57 \cdot 10^1$	0,88	0,19	$0,61 \cdot 10^{-2}$
100	14	$0,12 \cdot 10^{-2}$	$0,29 \cdot 10^2$	$0,29 \cdot 10^2$	$0,25 \cdot 10^1$	0,36	$0,72 \cdot 10^{-1}$	$0,73 \cdot 10^{-2}$
500	31	$0,23 \cdot 10^{-3}$	$0,31 \cdot 10^2$	$0,31 \cdot 10^1$	0,45	$0,63 \cdot 10^{-1}$	$0,19 \cdot 10^{-1}$	$0,13 \cdot 10^{-2}$
1000	43	$0,12 \cdot 10^{-3}$	$0,32 \cdot 10^2$	$0,18 \cdot 10^1$	0,22	$0,36 \cdot 10^{-1}$	$0,93 \cdot 10^{-2}$	$0,50 \cdot 10^{-3}$

## R E F E R E N C E S.

- (1) **BOSQ, D.** *Contribution à la théorie de l'estimation fonctionnelle.* Thèse, Université Pierre et Marie Curie, Paris, 1977.
- (2) **CHEVALIER, J.** *Estimation du support et du contenu du support d'une loi de probabilité.* Ann. Inst. Henri Poincaré, section B, Vol. XII, 4, 1976.
- (3) **EFRON, B.** *The convex hull of a random set of points.* Biometrika, 52, 331-343, 1965.
- (4) **FISHER, L.** *The convex hull of a sample.* Bull. Am. Math. Soc., 72, 555-558, 1966.
- (5) **GENSBITTEL, M.H.** *Contribution à l'étude statistique de répartitions ponctuelles aléatoires.* Thèse, Université Pierre et Marie Curie, Paris, 1979.
- (6) **GEFFROY, J.** *Sur un problème d'estimation géométrique.* Pub. I.S.U.P., 13, 191-200, 1964.
- (7) **GUILBART, C.** *Etude de la continuité de l'application  $H$  qui à toute mesure de probabilité définie sur  $\mathbf{R}^n$  fait correspondre l'enveloppe convexe fermée du support de cette mesure.* C.R. Acad. Sci. Paris, 277, série A, 999-1002, 1973.
- (8) **KALLENBERG, O.** *Random measures.* Academic Press.
- (9) **KARR, A.F.** *Point processes and their statistical inference.* (Dekker), 1986.
- (10) **KUELBS, J.** *A strong convergence theorem for Banach space valued random variables.* Annals of probability, 4, 744-771, 1976.
- (11) **JACOB, P.** *Estimation du contour discontinu d'un processus ponctuel sur le plan.* Publications I.S.U.P., XXIX, 3-4, 1-25, 1984.
- (12) **JACOB, P. and ABBAR, H.** *Estimating the edge of a Cox process area.* Cahiers du centre d'études de recherche opérationnelle, Vol. 31, n° 3-4.
- (13) **RAYNAUD, H.** *Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans  $\mathbf{R}^n$ .* C.R. Acad. Sci. Paris, 262, série A, 623-627, 1965.

- (14) **RENYI, A. and SULANKE, R.** *Über die konverze hülle von  $n$  zufällig gewählten punkten II.* Z. Wahrschein., 3, 138-147, 1964.
- (15) **RIPLEY, B.D. and RASSON, J.P.** *Finding the edge of a Poisson forest.* J. Appl. Prob., 14, 483-491, 1977.



## RÉSUMÉ

Nous étudions le problème de l'estimation du contour du support  $S$  d'une répartition ponctuelle aléatoire se réalisant dans le plan autour d'un point connu.

Nous proposons pour cela un estimateur de type histogramme de ce contour, et nous en déduisons un nouvel estimateur par un lissage spline cubique. Pour ces deux estimateurs, nous établissons des conditions de convergence, lorsque l'observation est une superposition du processus de Cox. La loi limite de l'estimateur de type histogramme est une loi de Gumbel dont nous trouvons les paramètres.

En pratique, il faut pouvoir déterminer un nombre optimal de subdivisions de  $S$  au sens de l'erreur quadratique intégrée. Nous donnons l'ordre de grandeur asymptotique de ce nombre de subdivisions dans différents cas. Dans un exemple simplifié, nous étudions numériquement le choix optimal de ce nombre à distance finie.

## MOTS CLÉS

VALEURS EXTREMES – PROCESSUS PONCTUELS – ESTIMATION FONCTIONNELLE – SUPPORT D'UNE MESURE – PROCESSUS DE COX – SPLINES CUBIQUES – ERREUR QUADRATIQUE INTEGREE – LOI DE GUMBEL.