Imprecise Estimation of The Probability Density Function

Bilal Nehme, Kevin Loquin, Olivier Strauss

Estimation imprécise de la densité de probabilité

Imprecise estimation of the probability density function

B.Nehme[1] K. Loquin[2] O.Strauss[3]
LIRMM, Université Montpellier II, 161 rue Ada, 34392 Montpellier cedex 5, France

[1] Bilal.Nehme@lirmm.fr
[2] Kevin.Loquin@lirmm.fr
[3] Olivier.Strauss@lirmm.fr

Résumé :

Dans cet article, nous proposons une adaptation de l'estimateur de densité de probabilité de Parzen Rosenblatt utilisant des noyaux maxitifs. Le résultat de cet estimateur, en chaque point du domaine de la densité à estimer, est un intervalle au lieu d'une valeur ponctuelle. Notre approche est cohérente avec l'estimation de Parzen Rosenblatt, car sous certaines conditions explicitées dans cet article, notre estimateur contient cette estimation.

Mots-clés :

Parzen Rosenblatt, densité de probabilité, noyau maxitif, estimation imprécise.

Abstract:

In this paper, we propose an adaptation of the Parzen Rosenblatt density estimator that uses maxitive kernels. The result of this estimator, on every point of the domain of the density to be estimated, is interval valued. We prove the consistency of our approach with the Parzen Rosenblatt estimator, since, according to conditions exposed in this paper, our estimate contains this estimation.

Keywords:

Parzen Rosenblatt, probability density function, maxitive kernel, imprecise estimation.

1 Introduction.

La plupart des outils d'analyse, de filtrage, d'agrégation et de tests statistiques s'appuient sur une connaissance plus ou moins exacte de la densité de probabilité sous-jacente à un ensemble d'observations. L'estimation de la densité de probabilité est donc un problème fondamental qui a fait l'objet d'une très vaste littérature [3, 4]. Cette estimation s'appuie généralement sur un ensemble fini d'observations $(x_1, ..., x_n)$ de n variables aléatoires indépendantes et identiquement distribuées $(X_1, ..., X_n)$ de loi f à estimer. Cependant, ce type d'estimation n'est fiable que dans des conditions asymptotiques, c.à.d. avec nombre infini d'observations. Vouloir obtenir une estimation précise de la densité de probabilité parait optimiste au vue de l'incomplétude des observations ou du manque d'information. Une estimation imprécise de cette densité de probabilité, plus robuste vis à vis de l'incomplétude des données, apparait alors plus judicieuse. Nous nous intéressons dans cet article à l'estimateur à noyau de Parzen Rosenblatt [1, 2] donné en tout point x de Ω par :

$$\hat{f}_\kappa(x) = \frac{1}{n}\sum_{i=1}^{n} \kappa(x - x_i), \qquad (1)$$

où κ est un noyau sommatif [5]. De même qu'un noyau sommatif est un voisinage pondéré probabiliste, de même un noyau maxitif est un voisinage pondéré possibiliste. Un noyau maxitif π est équivalent à une famille de noyaux sommatifs, notée $core(\pi)$ [5]. Dans un précédent article [6], nous avons proposé une méthode d'estimation imprécise cohérente de la fonction de répartition F, utilisant un noyau maxitif π. Cet estimateur est cohérent car l'intervalle obtenu avec π, $\overline{F}_\pi(x) = [\underline{F}_\pi(x), \overline{F}_\pi(x)]$, contient les estimations de Parzen Rosenblatt $\hat{F}_\kappa(x)$ obtenus avec tous les noyaux sommatifs contenus dans $core(\pi)$. Cette approche prend en compte les défauts de modélisation de l'estimateur de Parzen Rosenblatt en rendant "plus souple" le choix du noyau. On montre dans cet article qu'une adaptation directe de la méthode proposée dans [6] pour une estimation imprécise

de la densité de probabilité est impossible. En s'appuyant sur le fait que f est la dérivée de la fonction de répartition F, nous proposons une adaptation détournée de la méthode de [6], qui consiste à réaliser une estimation imprécise de la dérivée de la fonction de répartition. On obtient ainsi un intervalle d'estimation $\overline{\underline{f}}_\pi(x) = [\underline{f}_\pi(x), \overline{f}_\pi(x)]$ cohérent par rapport aux estimateurs de Parzen Rosenblatt $\hat{f}_\kappa(x)$ obtenus avec des noyaux de la famille représentée par π.
L'article est organisé comme suit. Dans la section 2, nous rappelons certains concepts sur les distributions, l'estimateur de Parzen Rosenblatt et les noyaux maxitifs. L'estimation imprécise de la densité de probabilité est donnée dans la section 3. La section 4 est dédiée aux expérimentations. La section 5 conclut l'article.

2 Notions préliminaires.

2.1 Rappels sur les distributions.

Soit Ω un sous-ensemble ouvert de $\mathbb{R}$. Soit s une application L_1 de Ω dans $\mathbb{R}$ associée à une distribution (au sens de Schwartz [7]) et κ une fonction à support compact de Ω dans $\mathbb{R}$. On appelle produit de convolution de s et κ la fonction $\widehat{s}_\kappa(x)$de Ω dans $\mathbb{R}$ telle que :

$$\widehat{s}_\kappa(x) = \int_\Omega s(w)\kappa(x-w)dw. \tag{2}$$

En notant κ_x , la fonction translatée de κ en x :

$$\forall \omega \in \Omega, \kappa_x(w) = \kappa(x-w), \tag{3}$$

ce produit de convolution peut être noté :

$$\widehat{s}_\kappa(x) = \langle s, \kappa_x \rangle = \int_\Omega s(w)\kappa_x(w)dw, \tag{4}$$

où $\langle ., . \rangle$ est le produit scalaire défini pour les fonctions L_1.
Si κ est une fois différentiable, alors un corolaire simple de la dérivation au sens des distribution nous permet de relier ds, la dérivée de s au sens des distributions, à $d\kappa$, la dérivée de κ au sens des fonctions par :

$$\langle ds, \kappa_x \rangle = - \langle s, d\kappa_x \rangle . \tag{5}$$

2.2 Rappel sur l'estimateur de Parzen Rosenblatt.

La méthode de Parzen Rosenblatt [1] consiste à réaliser une estimation non paramétrique de la densité de probabilité sous-jacente à un ensemble de n observations $(x_1, ..., x_n)$ de n variables aléatoires indépendantes et identiquement distribuées $(X_1, ..., X_n)$ de loi f à estimer. La réalisation de cette estimation, en tout point de Ω , nécessite la définition d'un voisinage pondéré sous la forme d'une fonctionnelle intégrable à 1 que nous appelons noyaux sommatif. Un noyau sommatif est une fonction κ à valeur dans $\mathbb{R}^+$ définie sur Ω vérifiant la propriété de sommativité :

$$\int_\Omega \kappa(w)dw = 1. \tag{6}$$

Pour un noyau sommatif donné κ , l'estimation de $\widehat{f}_\kappa$ est donnée en tout point $x \in \Omega$ par :

$$\widehat{f}_\kappa(x) = \frac{1}{n}\sum_{i=1}^{n} \kappa(x - x_i). \tag{7}$$

La plupart des noyaux sommatifs couramment utilisés en estimation fonctionnelle sont monomodaux, symétriques et centrés (c'est à dire définissant un voisinage autour de l'origine). Grâce à l'expression (3), on peut réécrire la formule (7) comme suit :

$$\widehat{f}_\kappa(x) = \frac{1}{n}\sum_{i=1}^{n} \kappa_x(x_i). \tag{8}$$

L'estimation $\widehat{f}_\kappa$ en chaque point $x \in \Omega$ peut être réécrite comme le produit de scalaire du noyau κ_x avec la mesure empirique e_n :

$$\widehat{f}_\kappa(x) = \langle e_n, \kappa_x \rangle , \tag{9}$$

avec

$$e_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}, \tag{10}$$

où δ_{x_i} est l'impulsion de Dirac translatée en x_i En effet,

$$\begin{aligned}\langle e_n, \kappa_x \rangle &= \int_\Omega \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}(w)\kappa_x(w)dw \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_\Omega \delta_{x_i}(w)\kappa_x(w)dw \\ &= \frac{1}{n}\sum_{i=1}^{n}\kappa_x(x_i) = \widehat{f}_\kappa(x).\end{aligned}$$

Un noyau sommatif peut être vu comme une distribution de probabilité, induisant une mesure de confiance P_κ définie par :

$$\forall A \subseteq \Omega, P_\kappa(A) = \int_A \kappa(w)dw. \qquad (11)$$

L'estimation $\widehat{f}_\kappa$ en chaque point $x \in \Omega$ peut alors être interprétée comme l'espérance de la mesure empirique e_n dans le voisinage probabiliste défini par le noyau κ_x :

$$\widehat{f}_\kappa(x) = \mathbb{E}_{\kappa_x}(e_n) = \langle e_n, \kappa_x \rangle . \qquad (12)$$

De la même façon, on peut réaliser, pour un noyau sommatif η donné, une estimation de la fonction de répartition par :

$$\widehat{F}_\eta(x) = \mathbb{E}_{\eta_x}(E_n) = \langle E_n, \eta_x \rangle , \qquad (13)$$

E_n étant la fonction de répartition empirique définie par :

$$E_n(x) = \frac{1}{n}\sum_{i=1}^{n} H(x - x_i), \qquad (14)$$

où H est l'échelon d'Heaviside défini par :

$$H(x) = \begin{cases} 1 & si \;\; x \geq 0 \\ 0 & sinon \end{cases} \qquad (15)$$

En remarquant que e_n n'est autre que la dérivée au sens des distributions de E_n , et en utilisant l'expression (5), on peut récrire l'expression (12) sous la forme :

$$\widehat{f}_\kappa(x) = \langle dE_n, \kappa_x \rangle = \langle E_n, -d\kappa_x \rangle . \qquad (16)$$

2.3 Rappels sur les noyaus maxitifs.

Un noyau maxitif est une fonction permettant de définir un autre type de voisinage pondéré autour de chaque point de Ω . Un noyau maxitif π est une application à valeur dans $[0, 1]$ définie sur Ω vérifiant la propriété de maxitivité :

$$\sup_{x\in\Omega} \{\pi(x)\} = 1. \qquad (17)$$

Nous nous intéressons, dans cet article, aux noyaux maxitifs monomodaux. De même qu'un noyau sommatif définit une mesure de probabilité, de même un noyau maxitif définit une mesure de possibilité [8, 9, 10], notée Π_π, définie par :

$$\forall A \subseteq \Omega, \Pi_\pi(A) = \sup_{x\in A} \{\pi(x)\} . \qquad (18)$$

La mesure de possibilité étant non-additive, le noyau maxitif π définit une mesure de confiance duale à la mesure de possibilité, appelée mesure de nécessité, notée N_π, définie par :

$$\forall A \subseteq \Omega, N_\pi(A) = 1 - \Pi_\pi(A^c), \qquad (19)$$

A^c étant le complémentaire de A dans Ω.
Dû à la non-addivité des mesures de possibilité et de nécessité, l'opérateur espérance, utilisé avec des mesures de probabilités, doit être remplacé par sa généralisation qu'est l'intégrale de Choquet.
On dit qu'un noyau maxitif π domine un noyau sommatif κ, si la mesure de possibilité Π_π domine la mesure de probabilité P_κ, c'est à dire :

$$\forall A \subseteq \Omega, P_\kappa(A) \leq \Pi_\pi(A). \qquad (20)$$

En ce sens, un noyau maxitif définit l'ensemble des noyaux sommatifs qu'il domine. Cet ensemble est noté $core(\pi)$.
Une propriété fondamentale des noyaux maxitif, découlant des travaux de Schmeidler [11] et Denneberg [12], permet de répercuter la propriété de domination des noyaux sur les estimations. En effet, si s est une fonction bornée de Ω , π un noyau maxitif et κ un noyau sommatif dominé par π , on a la relation suivante :

$$\mathbb{C}_{N_\pi}(s) \leq \mathbb{E}_\kappa(s) \leq \mathbb{C}_{\Pi_\pi}(s), \qquad (21)$$

où $\mathbb{C}_{\Pi_\pi}(s)$ (rsp. $\mathbb{C}_{N_\pi}(s)$) est l'intégrale de Choquet de la fonction s par raport à la mesure de confiance Π_π (rsp. N_π). Plus précisément, $\mathbb{C}_{\Pi_\pi}(s)$ (rsp. $\mathbb{C}_{N_\pi}(s)$) est la borne supérieure (rsp. inférieure) de l'ensemble des espérances $\mathbb{E}_\kappa(s)$ avec $\kappa \in core(\pi)$. En notant $\overline{\mathbb{E}}_\pi(s)$ (rsp. $\underline{\mathbb{E}}_\pi(s)$) la borne supérieure (rsp. inférieure) de ces espérances, on obtient un opérateur d'estimation intervalliste de s dans le voisinage défini par le noyau maxitif π de la forme :

$$\overline{\underline{\mathbb{E}}}_\pi(s) = \left[\underline{\mathbb{E}}_\pi(s), \overline{\mathbb{E}}_\pi(s)\right], \qquad (22)$$

tel que :

$$\forall S \in \overline{\underline{\mathbb{E}}}_\pi(s), \exists \kappa \in core(\pi)/\mathbb{E}_\kappa(s) = S. \quad (23)$$

Cet estimateur intervalliste représente l'ensemble des valeurs $\mathbb{E}_\kappa(s)$ obtenues pour tous les noyaux sommatifs $\kappa \in core(\pi)$.

2.4 Choix du noyau maxitif.

En statistiques non paramétriques précises, la question du choix du noyau est généralement écartée par les utilisateurs de l'estimateur de Parzen Rozenblatt [1, 2]. En effet, l'étude du comportement asymptotique des estimateurs de densité (quand $n \to +\infty$) montre que la convergence de $\hat{f}_\kappa(x)$ vers la vraie densité f dépend plus de la largeur de bande du noyau que de sa forme. Dans ce contexte, le noyau d'Epanechnikov est généralement choisi car étant celui qui minimise le critère $MISE$[1] [3, 4].
Cependant, la condition asymptotique est irréalisable techniquement, car le nombre d'observations reste fini. En condition non asymptotique, l'estimation $\hat{f}_\kappa(x)$ dépend à la fois de la largeur de bande et de la forme du noyau. Dans le cas d'estimations imprécises avec noyaux maxitifs, l'impact de l'a priori sur le choix du noyau maxitif est moindre que dans le cas sommatif. En effet, choisir un noyau maxitif, c'est faire le choix d'une famille de noyau sommatifs avec des caractéristiques communes. Illustrons cela sur un exemple. Si l'utilisateur pense pouvoir s'appuyer sur un noyau sommatif particulier, par exemple celui d'Epanechnikov, pour lequel il peut spécifier la largeur de bande maximale Δ_{max}, alors l'utilisation de la transformation probabilité possibilité [13, 14], dite objective, permet de définir le noyau maxitif le plus spécifique dominant tous les noyaux ayant la forme désirée et de largeur de bande inférieure ou égale à Δ_{max}. L'intervalle d'estimation obtenu en utilisant l'expression (22) est alors l'intervalle le plus spécifique contenant l'ensemble des estimations qu'on aurait obtenues avec des noyaux sommatifs dominés par ce noyau maxitif. Dans le cas où l'utilisateur doute de son propre jugement, il souhaitera une estimation intervalliste prenant en compte plus de noyaux sommatifs. Il peut alors utiliser la transformation dite subjective (moins spécifique) de Dubois et Prade [13, 14] et obtenir un intervalle d'estimation contenant celui obtenu avec la transformation dite objective. Enfin, s'il ne peut rien spécifier d'autre que le support (ou largeur de bande) maximal du noyau sommatif à utiliser et le fait que ce noyau est symétrique et monomodal, il est alors possible d'utiliser le noyau maxitif linéaire (triangulaire) puisqu'il domine tout noyau sommatif monomodal symétrique borné ayant une largeur de bande inférieure à la largeur de bande spécifiée [13].

[1] Mean Integrated Squared Error : distance entre l'estimateur et la densité f

3 Estimation imprécise de la densité de probabilité.

3.1 Position du problème.

L'estimateur de Parzen Rosenblatt de la fonction de repartition peut s'ecrire $\widehat{F}_\eta(x) = \mathbb{E}_{\eta_x}(E_n)$. Dans un article précédent [6], nous avons utilisé l'expression (22) pour réaliser une estimation imprécise de la fonction de répartition d'une variable aléatoire, via un noyau maxitif π, vérifiant $\eta \in core(\pi)$, à partir d'un échantillon d'observations de cette variable. Cette estimation est obtenue simplement

par :

$$\begin{aligned}\left[\underline{F}_{\pi}(x), \overline{F}_{\pi}(x)\right] &= \overline{\underline{\mathbb{E}}}_{\pi_x}(E_n) \\ &= \left[\underline{\mathbb{E}}_{\pi_x}(E_n), \overline{\mathbb{E}}_{\pi_x}(E_n)\right],\end{aligned}$$

où π_x est le noyau maxitif translaté de π en x :

$$\forall w \in \Omega, \pi_x(w) = \pi(x - w).$$

Dû à la propriété de domination des noyaux maxitifs, cet intervalle est le plus spécifique contenant l'ensemble des valeurs d'estimation $F_\eta(x)$ obtenues par la méthode de Parzen Rosenblatt pour tout noyau η dominé par le noyau π.
Une interprétation simpliste de cet opérateur d'estimation laisserait à penser que l'on peut facilement obtenir une estimation imprécise de f, la densité de probabilité, sous la forme

$$\begin{aligned}\left[\underline{f}_{\pi}(x), \overline{f}_{\pi}(x)\right] &= \overline{\underline{\mathbb{E}}}_{\pi_x}(e_n) \\ &= \left[\underline{\mathbb{E}}_{\pi_x}(e_n), \overline{\mathbb{E}}_{\pi_x}(e_n)\right]. \quad (24)\end{aligned}$$

Cependant, l'estimation imprécise (22) fait intervenir une intégrale de Choquet, nécessitant le fait que la fonction à intégrer soit bornée. e_n étant une combinaison linéaire d'impulsions de Dirac, cette propriété n'est pas vérifiée, et donc l'expression (24) n'a pas de sens.
Nous proposons, dans cet article, de contourner en partie cette difficulté en s'appuyant sur la relation (16).

3.2 Estimation imprécise de f.

La dérivée d'un noyau sommatif κ peut s'écrire comme la combinaison linéaire de deux noyaux sommatifs :

$$-d\kappa = A^+\eta^+ - A^-\eta^-. \quad (25)$$

En effet,

$$-d\kappa = d\kappa^+ - d\kappa^-,$$

où $d\kappa^+(x) = max(0, -d\kappa(x))$ et $d\kappa^-(x) = max(0, d\kappa(x))$. Si l'on pose :

$$A^+ = \int_\Omega d\kappa^+(w)dw \quad \text{et} \quad A^- = \int_\Omega d\kappa^-(w)dw$$

et

$$\eta^+(w) = \frac{d\kappa^+(w)}{A^+} \quad \text{et} \quad \eta^-(w) = \frac{d\kappa^-(w)}{A^-}$$

l'expression (25) est vérifiée et η^+ et η^- sont sommatifs.

Théorème 1 *Soit κ un noyau sommatif. Soient π^+ et π^- deux noyaux maxitifs tels que $\eta^+ \in core(\pi^+)$ et $\eta^- \in core(\pi^-)$ alors pour tout $x \in \Omega$:*

$$\widehat{f}_\kappa(x) \in A^+\overline{\underline{\mathbb{E}}}_{\pi_x^+}(E_n) \ominus A^-\overline{\underline{\mathbb{E}}}_{\pi_x^-}(E_n),$$

où $\ominus$ est l'extention de la soustraction aux intervalles.

Preuve D'après (16) et (25), en s'appuyant sur la linéarité du produit scalaire, on a :

$$\widehat{f}_\kappa(x) = A^+\left\langle E_n, \eta_x^+\right\rangle - A^-\left\langle E_n, \eta_x^-\right\rangle.$$

D'aprés (13) on a que :

$$\widehat{f}_\kappa(x) = A^+\widehat{F}_{\eta^+}(x) - A^-\widehat{F}_{\eta^-}(x).$$

En s'appuyant sur les résultats de [6] présentés dans la section 3.1, on a que :

$$\text{pour } \eta^+ \in core(\pi),\ \widehat{F}_{\eta^+}(x) \in \overline{\underline{\mathbb{E}}}_{\pi_x^+}(E_n),$$

$$\text{pour } \eta^- \in core(\pi),\ \widehat{F}_{\eta^-}(x) \in \overline{\underline{\mathbb{E}}}_{\pi_x^-}(E_n),$$

d'où

$$\widehat{f}_\kappa(x) \in A^+\ \overline{\underline{\mathbb{E}}}_{\pi_x^+}(E_n) \ominus A^-\ \overline{\underline{\mathbb{E}}}_{\pi_x^-}(E_n).$$

Le calcul des bornes inférieures et supérieures de l'estimation maxitive de E_n via un noyau π fait intervenir deux intégrales de Choquet. Nous en donnons les expressions :

$$\mathbb{C}_{\Pi_\pi}(E_n) = \frac{1}{n}\sum_{i=1}^{n}\left(\pi(x_i)H(x_i - x) + H(x - x_i)\right),$$

$$\mathbb{C}_{N_\pi}(E_n) = \frac{1}{n}\sum_{i=1}^{n}\left((1 - \pi(x_i))H(x - x_i)\right).$$

Pour plus de détails sur le calcul des bornes voir [6].

4 Expérimentations.

Nous proposons, dans cette section, deux expérimentations afin d'illustrer tant les qualités que les défauts de la méthode d'estimation que nous proposons. Nous basons nos expérimentations sur des observations simulées d'une variable aléatoire dont la densité de probabilité est une distribution bimodale obtenue en contaminant une loi normale de variance 4 centrée en 8 par une loi normale de variance unitaire et centrée en 3. Nous avons choisi d'utiliser, comme noyau sommatif de référence, le noyau d'Epanechnikov car c'est le plus utilisé en estimation de Parzen Rosenblatt d'une part, et c'est aussi celui qui a le comportement le plus singulier dans l'approche que nous proposons.
Dans la première expérience, nous avons simulé 1000 observations $(x_1, ..., x_{1000})$. Nous avons réalisé l'estimation de Parzen-Rosenblatt précise en utilisant un noyau d'Epanechnikov dont la largeur de bande est optimale (c'est à dire minimise la distance MISE). Nous avons d'autre part réalisé l'estimation imprécise de cette distribution en dominant le noyau dérivé, comme indiqué par les hypothèses du théorème 1 grâce aux transformations probabilité-possibilité objective (la plus spécifique) d'une part (Figure 2) et subjective d'autre part (Figure 1). Nous avons reporté sur chaque figure la courbe de la densité simulée.
On peut voir, sur la Figure 2, que l'utilisation de la domination la plus spécifique conduit à une estimation inférieure égale à l'estimation par noyau sommatif. Cette propriété est une des particularités du noyau d'Epanechnikov. Par contre, l'examen de la Figure 1 montre qu'une domination moins spécifique des noyaux de dérivation conduit à une estimation précise qui semble à mi-chemin des bornes de l'estimation imprécise. La seconde expérience tend à mettre en valeur la robustesse induite par l'utilisation d'un estimation imprécise. Nous proposons de caractériser cette robustesse par deux types d'indice. Le premier montre la corrélation entre l'écart $(\overline{f}_\pi(x) - \underline{f}_\pi(x))$ d'une part et la variance d'estimation

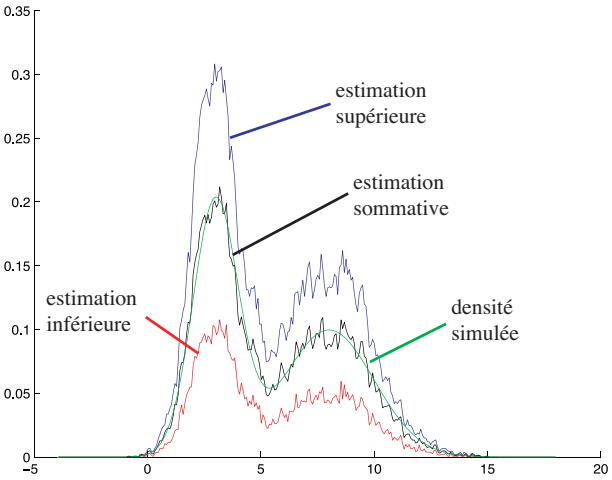

Figure 1 – estimation précise et imprécise, domination avec la transformation subjective.

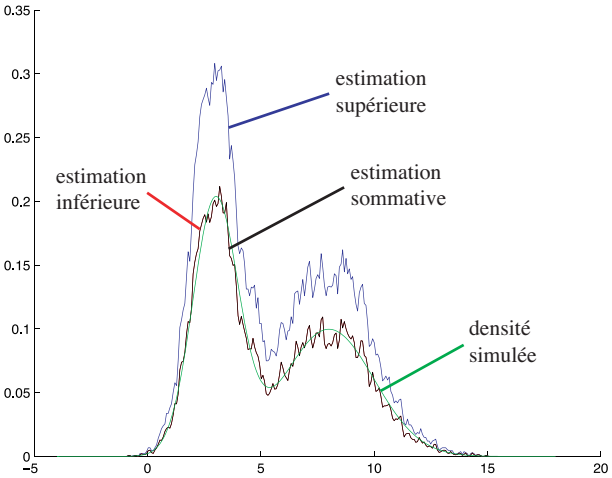

Figure 2 – estimation précise et imprécise, domination avec la transformation objective.

de $\widehat{f}_\kappa(x)$ d'autre part. Le second montre l'aptitude de l'estimation imprécise $\left[\underline{f}_\pi(x), \overline{f}_\pi(x)\right]$ obtenue avec une seule expérience à contenir la vraie densité $f(x)$.

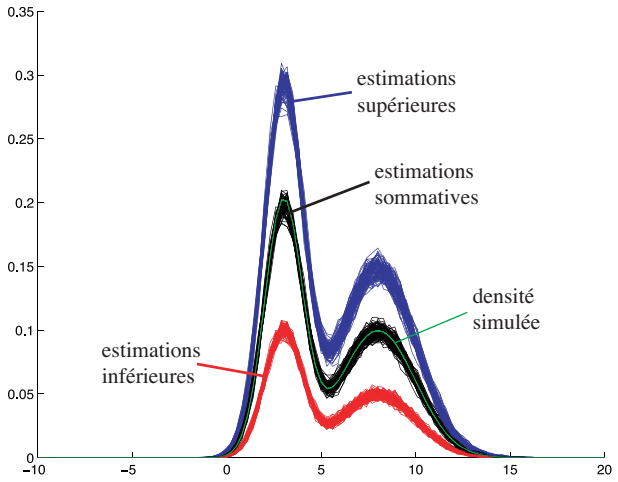

Figure 3 – Superposition des 100 estimations pour une largeur de bande de 0.8.

L'expérience se déroule de la manière suivante. On effectue 500000 tirages indépendants issus de la loi bimodale simulée. On divise l'ensemble de ces observations en 100 échantillons de 5000 observations. Pour chacun des 100 échantillons, on dispose d'une estimation précise $\widehat{f}_\kappa$ et d'une estimation imprécise $\left[\underline{f}_\pi, \overline{f}_\pi\right]$. Ces estimations sont évaluées en 100 points d'échantillonnage de la droite réelle sur l'intervalle $[-10, 20]$. En chaque point d'échantillonnage w, on calcule la variance $v_{\widehat{f}_\kappa}(w)$ de l'estimation $\widehat{f}_\kappa(w)$ d'une part et, pour chaque groupe d'observation, l'ecart $(\overline{f}_\pi(w) - \underline{f}_\pi(w))$. Nous calculons ensuite les indices de corrélations de Pearson, Kendal et Spearman entre ces deux séries de valeurs pour des largeurs de bandes variant entre 0.2 et 1.6. Notons qu'une étude du comportement de l'estimateur nous a permis de déterminer que la largeur de bande optimale, c'est à dire celle minimisant la distance MISE pour 5000 observations, est de 0.8.

A titre indicatif, la Figure 3 montre la superposition des estimations précises et imprécises obtenues avec les 100 échantillons d'observations pour une largeur de bande de 0.8 et pour une transformation probabilité-possibilité subjective.

D'autre part, nous nous intéressons à l'aptitude de l'intervalle $\left[\underline{f}_\pi(w), \overline{f}_\pi(w)\right]$ à contenir la vraie distribution $f(w)$. Pour caractériser cette aptitude, nous calculons, en chaque point, le nombre d'intervalles, obtenu grâce à un échantillon, contenant la vraie valeur de $f(w)$. Ce nombre est rapporté au nombre de valeurs testées pour obtenir le taux d'intervalle de prédiction correctes.

La Figure 4 montre l'évolution de ces quatre

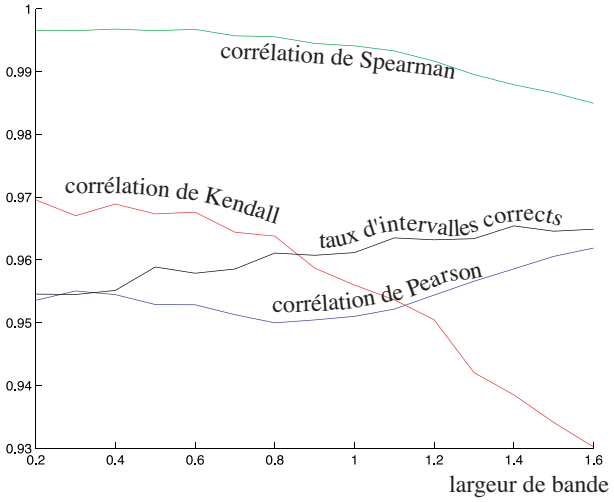

Figure 4 – Courbes de performance de l'estimateur imprécis en fonction de la largeur de bande.

indices pour différentes largeurs de bande. Concernant les indices de corrélation, on peut constater une très forte corrélation entre la variance d'estimation précise de la densité et l'imprécision de l'estimation imprécise de la densité. Cette corrélation décroit lorsqu'on s'écarte de la largeur de bande optimale, mais reste cependant très élevée (supérieure à 0.8 dans la pratique). De même, on peut constater, via l'indice de cohérence, la très bonne aptitude de l'estimation intervalliste à prédire la vraie densité, même lorsque la largeur de bande n'est pas optimale.

On peut cependant déplorer que ces très bons résultats soient contrebalancés par le manque de spécificité de l'estimation (comme l'illustre la Figure 3). Ce manque de spécificité est dû au fait que les estimations $\overline{\underline{\mathbb{E}}}_{\pi_x^+}(E_n)$ et $\overline{\underline{\mathbb{E}}}_{\pi_x^-}(E_n)$ sont considérées comme indépendantes dans la méthode proposée, alors qu'elles ne le sont pas. Il serait donc maintenant interessant de trouver le couple de noyaux maxitifs (π^-, π^+) permettant une domination de la dérivée du noyau κ aboutissant à une estimation plus spécifique de f, ou encore de modifier l'extention choisie de l'opération de soustraction de façon à prendre en compte cette dépendance.

5 Conclusion.

Dans cet article, nous avons proposé un estimateur de densité de probabilité s'appuyant sur une représentation du voisinage par noyau maxitif. Cet estimateur s'apparente très fortement à celui de Parzen Rosenblatt, avec, comme principale différence, le fait que l'estimation produite en chaque point de la droite réelle est disponible sous forme d'intervalle en lieu et place d'une valeur précise. Un utilisateur d'estimateur de densité pourrait s'interroger sur les raisons qui pourraient l'amener à passer d'un estimateur précis à un estimateur imprécis. Nous avons deux arguments principaux. Premier argument : l'utilisation d'un voisinage maxitif permet de représenter une mauvaise connaissance du noyau optimal à utiliser. Ce défaut de connaissance est directement impacté sur l'imprécision de l'estimation. Second argument : l'écart de l'intervalle d'estimation de densité obtenu avec un noyau maxitif est très fortement corrélé avec la variance des estimations qui auraient été obtenues avec un noyau sommatif et différents échantillons. On dispose alors d'une mesure objective de l'erreur d'estimation. Un des points clef de notre travail est l'utilisation de la relation entre densité de probabilité et fonction de répartition d'une variable aléatoire (l'une est la dérivée de l'autre). Ce constat nous a permis de transformer l'estimation ponctuelle de la densité de probabilité avec un voisinage sommatif en une différence d'estimation imprécise de la densité cumulée. On doit cependant noter que l'estimation obtenue n'est pour l'instant pas assez spécifique. Ce manque de spécificité est principalement du au fait que la différence de Minkowski utilisée ne permet pas de prendre en compte la dépendance des noyaux sommatifs dominés. En modifiant l'opérateur de différence de Minkowski, il devrait être possible de recouvrer cette spécificité.

Références

[1] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33 : 1065-1076, 1962.

[2] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 27 : 832-837, 1956.

[3] J.S. Simonoff. *Smoothing Methods in Statistics*. Springer-Verlag, 1996.

[4] D.W. Scott. *Multivariate Density Estimation*. Wiley Interscience, 1992.

[5] K. Loquin, O. Strauss. On the granularity of summative kernels. *Fuzzy Sets and Systems*, manuscrit accepté.

[6] K. Loquin, O. Strauss. Imprecise functional estimation : the cumulative distribution case. *Soft Methods in Probability and Statistic (SMPS), 2008*, accepté.

[7] L. Schwartz. *Théorie des distributions*. Hermann, Paris, 1950.

[8] D. Dubois, H. Prade. *Théorie des possibilités : Applications à la représentation des connaissances en informatique*. Masson, 1988.

[9] D. Dubois, H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49 : 65-74, 1992.

[10] G. de Cooman. Possibility theory. Part I : Measure- and integral-theoretic groundwork ; Part II : Conditional possibility ; Part III : Possibilistic independence. *Int. J. of General Systems*, 25 : 291-371, 1997.

[11] D. Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97 : 255-261, 1986.

[12] D. Denneberg. *Non Additive Measure and Integral*. Kluwer Academic Publishers, 1994.

[13] D. Dubois, H. Prade, L. Foulloy, G. Mauris. Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities. *Reliable Computing*, 10 : 273-297, 2004.

[14] D. Dubois, H. Prade, S. Sandri. On possibility/probability transformations. *Fuzzy Logic. State of the Art*, pp. 103-112, Kuwer Acad. Publ., S. Lowen, R. and Roubens, M. (ed.) 1993.