

=====
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ DES FRÈRES MENTOURI
FACULTÉ DES SCIENCES EXACTES

=====
DÉPARTEMENT DE MATHÉMATIQUES

N° d'ordre : **10/DS/2023**

N° de série : **02/MAT/2023**

THÈSE

PRÉSENTÉE POUR L'OBTENTION
DU
DIPLOME DE DOCTORAT EN SCIENCES
DE
MATHÉMATIQUES

« Étude asymptotique dans l'inférence statistique »

Par
Madi Meriem

OPTION
Contrôle optimal stochastique

Devant le jury :

Président	I. Lescheb	M.C.A.	Université Des Frères Mentouri Constantine 1
Encadreur	K. Bessila	M. C. A.	Université Des Frères Mentouri Constantine 1
Co-encadreur	I. Laroussi	M.C.A.	Université Des Frères Mentouri Constantine 1
Examineur	F. Messaci	Prof	Université Salah Boubnider Constantine 3
Examineur	S. Kherfouchi	Prof	Université Salah Boubnider Constantine 3
Examineur	M. Boukeloua	M.C.A.	École nationale polytechniques Constantine

Soutenue le :26-03-2023 .

Table des matières

Table des matières	i
Liste des tableaux	iii
Table des figures	iv
Introduction	iii
1 Modes de convergence des suites de variables aléatoires	1
1.1 Modes de convergence usuels	1
1.2 Implications entre les différents modes de convergence	9
1.3 Taux de convergence	20
1.4 Utilisation des concepts de convergence en statistiques	24
2 Estimation à noyau des caractéristiques fonctionnelles d'une loi de probabilité	29
2.1 Méthode d'estimation non paramétrique à noyau	29
2.2 Estimateur à noyau d'une densité de probabilité	33
2.3 Estimateur à noyau de la fonction de répartition	47
2.4 Estimateur à noyau de la fonction de quantile	51
3 Convergence complète en moment du second-ordre avec taux	55
3.1 Convergence complète en moment du second-ordre avec taux	55
3.2 Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de densité	61
3.3 Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de répartition	67
3.4 Convergence c.m.s. de l'estimateur à noyau de la fonction de quantile	73

TABLE DES MATIÈRES

4	Simulation et application	77
4.1	Présentation du noyau de Laplace	77
4.2	Simulation	79
4.3	Application des données réelles	82
4.4	Conclusion	83
	Bibliographie	85

Liste des tableaux

2.1	Quelques noyaux.	31
4.1	MSE et taux de convergences simulés pour la densité.	80
4.2	MSE et taux de convergences simulées de l'estimation de la répartition.	81
4.3	MSE et taux de convergences simulées de l'estimation de quantile.	82

Table des figures

1.1	Résumé des relations d'implications entre les modes de convergence étudiés.	28
2.1	Influence du paramètre de lissage sur l'estimation de la fonction de densité pour une courbe théorique gaussienne.	30
2.2	Noyau d'Epanechnikov ou parabolique.	32
2.3	Noyau Gaussien.	32
2.4	Noyau triangulaire.	32
2.5	Noyau de Silverman.	32
2.6	Noyau uniforme ou de Rozenblatt.	33
2.7	Estimateur à noyau gaussien \hat{f}_6 de la densité en 6 observations.	34
4.1	Noyau de Laplace.	78
4.2	Densité et répartition des données réelles avec ses estimations à noyau.	83

"الحمد لله رب العالمين، الذي جعل لكلّ شيء قدرًا، وجعل لكلّ قدرٍ أجلاً، وجعل لكلّ أجلٍ كتاباً".

Je tiens à remercier monsieur K. Bessila, mon directeur de thèse pour l'aide qu'il m'a procuré.

J'adresse mes plus vifs remerciements mon Co-encadreur de thèse Melle I. Laroussi, pour m'avoir encadrer, pour les nombreuses et fructueuses discussions, pour sa rigueur scientifique, ainsi que pour ces encouragements qui m'ont été d'une grande aide. Je la remercie aussi pour sa générosité, sa gentillesse et sa patience avec moi.

Je remercie infiniment mademoiselle I. Lescheb qui m'a fait l'honneur de présider le jury de ma thèse.

Je remercie chaleureusement aussi les membres du jury madame F. Messaci, madame S. Kherfouchie et monsieur M. Boukeloua d'avoir accepter de juger ma thèse.

Je tiens à remercier toutes les personnes non citées qui auraient contribué d'une manière ou d'une autre à la réalisation de ce travail.

Introduction

L'inférence statistique, dans un cadre paramétrique ou non, aboutie à des résultats de consistance d'estimateurs proposés. Ces résultats font état d'étude de différents types ou modes de convergence qui représentent l'achèvement du travail de recherche. Les modes de convergences usuels comme la convergence en loi, en probabilité, presque sûre et dans L^p évoquent toute formation d'un bon mathématicien probabiliste et statisticien. Les ouvrages qui relatent et développent ces modes ainsi que leurs applications aux variables et aux vecteurs aléatoires sont nombreux, si ce n'est pas dire, incalculables. Mais, certains sont incontournables comme Métivier (1972), Billingsley (1979) et plus récemment Ouvrard (2007). Aussi, dans l'étude de la convergence des suites de variables aléatoires, on s'intéresse à un concept important qui décrit la vitesse à laquelle une suite converge vers la limite, appelé taux de convergence. Dans ce contexte, une question fondamentale est de savoir à quelle vitesse la convergence éclaire les études théoriques sur le sujet qui ont été menées proposant de nouveaux résultats, algorithmes et applications importants pour résoudre ce problème. Le but de ce travail est de présenter un nouveau mode de convergence, que l'on appelle la convergence complète en moment du second-ordre avec un taux et la nouveauté réside ici dans l'introduction du taux de ce type de convergence. Par conséquent, on donne des preuves de quelques propriétés de cette convergence et on démontre la relation d'implication entre elle et celle presque complète. Cette dernière est induite par la convergence en moment dans des contextes différents (voir Chow (1988), Liang et al (2010) et Qiu et Chen (2014) et récemment voir Yu et al (2022)). Le concept de convergence presque complète a été introduit par Hsu et Robbins (1947). Ensuite, il a été utilisé par plusieurs auteurs tels que Gut et Stadtmüller (2011), Gut (1978, 1980), Li et al. (1995), Sung (2007), Sung et Volodine (2006). L'intérêt d'une telle notion réside dans le fait que la convergence presque complète (p.co.) implique la convergence presque sûre (p.s.) due au lemme de Borel-Cantelli. Cependant, dans certaines situations, il est beaucoup plus facile

d'obtenir une convergence complète en moment du second-ordre (c.m.s.) au lieu d'une convergence presque complète.

Comme cadre pratique, on a établi les taux de convergence complète en moment du second-ordre pour les estimateurs du noyau de la densité de probabilité, la fonction de répartition et de la fonction de quantile et on a discuté également de la vitesse à laquelle ces estimateurs convergent. Premièrement, dans le cadre de l'estimation de la fonction de densité de probabilité, de nombreuses études utilisant différentes méthodes ont été proposées. La méthode des noyaux est l'une des meilleures de ces méthodes qui semble être pratique et ne nécessite pas un choix multiple de paramètres. Rosenblatt (1956) a été le premier pionnier de la classe des estimateurs à noyaux de densité, en utilisant deux paramètres, à savoir le noyau K et le paramètre de lissage h . Pour cet estimateur, la convergence en probabilité a été établie par Parzen (1962). Habbema et al. (1974), Hall et Kang (2005), Hall et Wand (1988), Gosh et Chaudhury (2004) et Gosh et Hall (2008) peuvent être consultés pour différents travaux sur le sujet, notamment l'estimation par noyaux classiques de la densité. Dans le cas d'observations indépendantes, les taux optimaux de convergence vers zéro pour l'erreur quadratique moyenne et le biais des estimateurs à noyaux ont été abordés par plusieurs auteurs dans des conditions variables sur le noyau K et la densité f . En guise de contribution, une nouvelle convergence complète en moment du second-ordre presque avec un taux est introduite pour la première fois afin d'améliorer les taux de convergence pour le biais et l'erreur quadratique moyenne (MSE) de l'estimateur à noyau de la densité. Ensuite, la convergence presque complète de l'estimateur à noyau de la densité, sous des conditions plus faibles sur la fonction de densité que celles proposées dans la littérature, est atteinte. En conséquence de la convergence complète en moment du second-ordre avec taux, la convergence presque complète est obtenue avec un meilleur taux.

Deuxièmement, le mode de convergence proposé est appliqué aux estimateurs à noyaux la fonction de répartition et de la fonction de quantile. Notez que pour la fonction de distribution, Nadaraya (1964) a proposé son estimateur par noyau. Tandis que Parzen (1979) a retracé le contexte de Nadaraya (1964) et a construit les estimateurs à noyaux de la fonction de quantile pour lesquels nous avons établi le taux de convergence presque complète. A notre connaissance, il s'agit d'un nouveau résultat obtenu à partir du mode de convergence proposé. Un grand avantage de l'approche proposée est que moins de conditions sont imposées à la fonction du noyau grâce à l'utilisation de l'expression de l'erreur quadratique moyenne.

Cette thèse contient quatre chapitres :

Le premier chapitre, est consacré à présenter les différents modes usuels

de convergence de suites de variables aléatoires réelles, citons la convergence en loi, en probabilité, presque sûre, presque complète et celle en norme complète ou en moment sont définis. Les relations d'implication et implication inverse existant entre ces types de convergence sont étudiées et démontrées dans cette partie. Les règles de calculs concernant la limite de la somme, le produit et la division des suites de variables aléatoires font aussi un objectif de ce chapitre, sans oublier les principaux théorèmes d'application du concept de la convergence en statistique.

Dans le deuxième chapitre, on donne un bref rappel sur la méthode d'estimation à noyau, en donnant quelques exemples sur les noyaux classiques continus les plus utilisés dans la littérature. Ensuite on étudie les propriétés asymptotiques des estimateurs à noyaux des fonctions de densité, de répartition et de quantile, en fixant l'étude sur le biais, la variance, l'erreur moyenne quadratique et la convergence de ces estimateurs.

Le troisième chapitre, présente le nouveau mode de convergence des suites de variables aléatoires réelles, appelé convergence complète en moment du second-ordre avec taux et on démontre quelques propriétés. Ensuite, on applique ce type de convergence sur les estimateurs à noyaux des fonctions de densité, de répartition et de quantile, en démontrant leur convergence vers les fonctions à estimer. Ce nouveau mode de convergence implique la convergence presque complète, ainsi que son taux est inférieur à celui de cette dernière. Tous ces résultats sont obtenus à l'aide de l'expression de l'erreur moyenne quadratique MSE.

Finalement, dans le quatrième chapitre, on réalise une étude de simulation pour présenter la performance des nouveaux taux de convergence des estimateurs à noyaux des fonctions de densité, de répartition et de quantile. On effectue cette étude pour des échantillons des tailles différentes et pour des noyaux distincts, afin de conclure que les nouveaux taux de la convergence complète en moment du second-ordre sont meilleurs que ceux de la convergence presque complète pour les estimateurs étudiés. Une dernière partie de ce chapitre est consacrée à une application des résultats obtenus sur des données réelles recueillies du Centre Hospitalier Universitaire Ben Badis de Constantine.

1 Modes de convergence des suites de variables aléatoires

La convergence des suites de variables aléatoires est un concept important de la théorie des probabilités, avec multiples applications dans le domaine des statistiques et dans l'étude des processus stochastiques. C'est le cas, par exemple, des théorèmes des lois des grands nombres et le théorème central limite. Il existe différents modes de convergence des suites de variables aléatoires comme la convergence en loi, en probabilité, presque sûre, en moments d'ordre p et plus récemment la convergence presque complète. Cette dernière représente l'intérêt principal de notre recherche avec une nouvelle notion de convergence.

L'objectif de ce chapitre est de faire un rappel sur ces modes de convergence des suites de variables aléatoires, les relations entre eux et quelques règles de calculs.

1.1 Modes de convergence usuels

Dans tout ce qui va suivre, on considère l'espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel est défini les variables aléatoires réelles, en abrégé v.a.r., considérées.

1.1.1 Convergence en loi

Généralement considéré comme le mode de convergence des suites de variables aléatoires le plus faible, la convergence en loi (dite aussi conver-

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

gence en répartition) n'implique pas les autres modes de convergence, alors que les autres types de convergence l'impliquent.

Définition 1

Soient X et $(X_n)_{n \in \mathbb{N}^*}$ des v.a.r. définies sur les espaces $(\Omega, \mathcal{A}, \mathbf{P})$ et $(\Omega_n, \mathcal{A}_n, \mathbf{P}_n)$ respectivement. On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers X si et seulement si pour toute fonction réelle bornée et continue Φ de \mathbb{R} dans \mathbb{R} , on a

$$\lim_{n \rightarrow +\infty} \mathbf{E}(\Phi(X_n)) = \mathbf{E}(\Phi(X)),$$

on note

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X.$$

Remarque 1 Gut [20]

Les variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ et X ne sont pas nécessairement définies sur le même espace de probabilité, mais peuvent être définies sur des espaces de probabilité différents, car la convergence en loi est la convergence d'une suite de lois de probabilités des variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ vers la loi de probabilité de variable aléatoire X .

Le théorème suivant montre une propriété d'équivalence entre la définition de la convergence en loi et la convergence des fonctions décrivant la loi de $(X_n)_{n \in \mathbb{N}^*}$ vers celles décrivant la loi de X .

Théorème 1 Métivier [36].

Pour $(X_n)_{n \in \mathbb{N}^*}$ et X définies si dessus, on a

1. Soient $(F_{X_n})_{n \in \mathbb{N}^*}$ et F_X les fonctions des répartitions de $(X_n)_{n \in \mathbb{N}^*}$ et X respectivement, alors

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff F_{X_n}(x) \xrightarrow[n \rightarrow +\infty]{} F_X(x),$$

pour chaque x point de continuité de F_X .

2. Soient $(\Phi_{X_n})_{n \in \mathbb{N}^*}$ et Φ_X les fonctions caractéristiques de $(X_n)_{n \in \mathbb{N}^*}$ et X respectivement, alors

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \forall x \in \mathbb{R}, \Phi_{X_n}(x) \xrightarrow[n \rightarrow +\infty]{} \Phi_X(x),$$

avec Φ_X est continue en 0.

3. Soient $(X_n)_{n \in \mathbb{N}^*}$ à valeurs entières positives et $((G_{X_n})_{n \in \mathbb{N}^*}, G_X)$ les fonctions génératrices de $((X_n)_{n \in \mathbb{N}^*}, X)$ respectivement, alors

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X \iff \forall s, |s| \leq 1, G_{X_n}(s) \xrightarrow[n \rightarrow +\infty]{} G_X(s).$$

Preuve 1 Pour les preuves de ces équivalences, se référer à Métivier [36] pages 176 :184 et plus récemment, se référer à Gut 2005 [20].

La proposition ci-dessous précise les règles de calcul de la limite pour la somme, le produit et la division de deux suites de v.a.r. qui convergent en loi.

Proposition 1 (Théorème de Slutsky) [42]

Soient $(X_n)_{n \in \mathbb{N}^*}$ et $(Y_n)_{n \in \mathbb{N}^*}$ deux suites de variables aléatoires réelles. Alors

pour tout $c \in \mathbb{R}$, tel que $X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ et $Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} c$, on a

1. $(X_n, Y_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} (X, c)$,

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

2. $X_n + Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X + c,$
3. $X_n \cdot Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Xc,$
4. $X_n/Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X/c,$ avec $c \neq 0.$

Preuve 2 Pour la preuve on peut se référer à Polansky [42] page 179.

Remarque 2 Gut [20]

$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$ et $Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y$ n'implique pas que $X_n + Y_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X + Y,$ sauf s'il y'a indépendance entre les variables.

Dans la suite et de la même manière, on définit les autres types de convergence.

1.1.2 Convergence en probabilité, presque sûre et dans L^p

Soient X et $(X_n)_{n \in \mathbb{N}^*}$ des v.a.r. définies sur le même espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$.

Définition 2 La suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers X , lorsque n tend vers l'infini, si et seulement si

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$

On note cette convergence comme suit

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} X.$$

Définition 3 La suite $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X , lorsque n tend vers l'infini, si et seulement si

$$\mathbf{P} \left\{ \omega \in \Omega, \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega) \right\} = 1,$$

on écrit

$$X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X.$$

L'espace $L^p(\Omega, \mathcal{A}, \mathbf{P})$ représente l'espace des variables aléatoires à valeurs réelles, p intégrable, c'est à dire $\mathbf{E}|X|^p < \infty$ avec $p \geq 1$. Communément noté L^p . Sur cet espace on énonce la définition de la convergence en moment d'ordre p comme suit.

Définition 4 Soient X et $(X_n)_{n \in \mathbb{N}^*}$ des v.a.r. dans L^p . On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge vers X dans L^p si et seulement si

$$\lim_{n \rightarrow +\infty} \|X - X_n\|_p = 0,$$

c'est à dire, $\lim_{n \rightarrow +\infty} (\mathbf{E}|X_n - X|^p)^{\frac{1}{p}} = 0$. Et on note

$$X_n \xrightarrow[n \rightarrow +\infty]{L^p} X.$$

Proposition 2 Gut [20]

Soient $(X_n)_{n \in \mathbb{N}^*}$ et $(Y_n)_{n \in \mathbb{N}^*}$ deux suites de v.a.r. qui convergent respectivement vers les v.a.r. X et Y . Alors quelque soit le mode de convergence en probabilité, presque sûre ou dans L^p ; les propriétés suivantes sont vérifiées.

- 1- La limite en probabilité (resp. presque sûre ou dans L^p), si elle existe, est unique p.s.
- 2- $\forall (a, b) \in \mathbb{R}^2$, alors dans ce cas, quelque soit le mode de convergence
 - $aX_n + bY_n \xrightarrow[n \rightarrow +\infty]{} aX + bY$.
 - $X_n \cdot Y_n \xrightarrow[n \rightarrow +\infty]{} X \cdot Y$.
 - $\frac{1}{X_n} \xrightarrow[n \rightarrow +\infty]{} \frac{1}{X}$, si $\mathbf{P}(X_n \neq 0) = 1$.
- 3- Si $\forall n \in \mathbb{N}^*$, $X_n = a$, p.s., alors $X_n \xrightarrow[n \rightarrow +\infty]{} a$.

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

4- La suite $(X_n)_{n \in \mathbb{N}^*}$ est de Cauchy, c'est à dire

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} \mathbf{P}(|X_n - X_m| > \varepsilon) = 0,$$

si et seulement si elle converge en probabilité, presque sûre ou dans L^p .

Preuve 3 Pour la preuve on peut se référer à Gut [20] pages 208 et 247.

La sous section suivante est consacrée à d'autres modes de convergence considérés comme plus fort que ceux mentionnés précédemment.

1.1.3 Convergence presque complète et en moment complète

La convergence presque complète est souvent utilisé pour montrer la convergence presque sûre d'un estimateur sans la nommé, comme fut le cas en 1981 pour les travaux de Földes [15].

Définition 5 Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r.. On dit que $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers la v.a.r. X , si

$$\forall \varepsilon > 0, \sum_{n \in \mathbb{N}^*} \mathbf{P}(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) < \infty,$$

on note ce mode de convergence par

$$X_n \xrightarrow[n \rightarrow +\infty]{p.co} X.$$

Remarque 3 Polansky[42]

La convergence presque complète est plus forte que la convergence en probabilité, puisque cette dernière nécessite seulement que pour chaque $\varepsilon > 0$, la probabilité $\mathbf{P}(|X_n - X| > \varepsilon)$ converge vers 0, mais la convergence presque

complète exige que pour tout $\varepsilon > 0$, les probabilités $\mathbf{P}(|X_n - X| > \varepsilon)$ tendent vers 0, lorsque $n \rightarrow +\infty$. C'est à dire qu'elle exige une vitesse de convergence suffisante pour assurer la convergence de la somme infinie

$$\sum_{n \in \mathbb{N}^*} \mathbf{P}(|X_n - X| > \varepsilon).$$

Par exemple, si $\mathbf{P}(|X_n - X| > \varepsilon) = n^{-1}$, alors la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité, mais pas complètement, lorsque $n \rightarrow +\infty$.

Si $\mathbf{P}(|X_n - X| > \varepsilon) = n^{-2}$ alors la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité et complètement vers X lorsque n tend vers $+\infty$.

Pour le calcul de la limite presque complète de la somme, le produit et l'inverse des suites de variables aléatoires qui convergent presque complètement, on a le résultat qui suit.

Proposition 3 *Ferraty [14]*

Soient $(X_n)_{n \in \mathbb{N}^*}$ et $(Y_n)_{n \in \mathbb{N}^*}$ deux suites de v.a.r. telles que $X_n \xrightarrow[n \rightarrow +\infty]{p.co} l_1$ et $Y_n \xrightarrow[n \rightarrow +\infty]{p.co} l_2$ où l_1, l_2 sont deux nombres réels. Alors on a

$$1- X_n + Y_n \xrightarrow[n \rightarrow +\infty]{p.co} l_1 + l_2.$$

$$2- X_n \cdot Y_n \xrightarrow[n \rightarrow +\infty]{p.co} l_1 \cdot l_2.$$

$$3- \frac{1}{X_n} \xrightarrow[n \rightarrow +\infty]{p.co} \frac{1}{l_1}, \text{ si } l_1 \neq 0.$$

Preuve 4 *On se référer à Ferraty [14], page 230 pour la démonstration de ces propriétés de la convergence presque complète.*

Le lemme suivant dit de Borel-Cantelli, est un résultat immédiat de la convergence presque complète. Utilisé pour démontrer la relation d'implication entre ce mode de convergence et celui de la convergence presque sûre.

Lemme 1 (*Lemme de Borel-Cantelli*) Métivier [36]

Soit $(A_n)_{n \in \mathbb{N}^*}$ une suite d'éléments de \mathcal{A} .

1. $\sum_{n \in \mathbb{N}^*} \mathbf{P}(A_n) < \infty$; alors $\mathbf{P}\left(\limsup_{n \rightarrow +\infty} A_n\right) = 0$.

2. On suppose de plus, les événements $(A_n)_{n \in \mathbb{N}^*}$ indépendants et si

$$\sum_{n \in \mathbb{N}^*} \mathbf{P}(A_n) = \infty,$$

alors

$$\mathbf{P}\left(\limsup_{n \rightarrow +\infty} A_n\right) = 1.$$

Preuve 5 Pour la démonstration voir Métivier [36], page 207.

Considéré comme le mode de convergence le plus fort du fait qu'il implique la convergence presque complète, la convergence en moment complète est définie ci dessous. On introduit deux classes de ce mode de convergence qui sont des versions fortes de la convergence en moment.

Définition 6 La suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge complètement en moment d'ordre $p >$ vers la v.a.r. X , et on dit qu'elle est $s - L^p$ convergente, si pour tout $p > 0$,

$$\sum_{n=1}^{\infty} \mathbf{E}(|X_n - X|^p) < \infty,$$

on note cette convergence comme suit

$$X_n \xrightarrow[n \rightarrow +\infty]{s-L^p} X.$$

Remarque 4 Si $p = 1$ et $X_n \xrightarrow[n \rightarrow +\infty]{s-L^1} X$, on dit que la convergence complète de $(X_n)_{n \in \mathbb{N}^*}$ est en moment d'ordre 1 et on écrit

$$\sum_{n=1}^{\infty} \mathbf{E}(|X_n - X|) < \infty.$$

1.2. Implications entre les différents modes de convergence

Définition 7 *On dit que la suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge en norme complète d'ordre p , pour $p > 0$, vers la v.a.r. X si*

$$\sum_{n=1}^{\infty} \|X_n - X\|_p < \infty,$$

On note

$$X_n \xrightarrow[n \rightarrow +\infty]{s^*-L^p} X.$$

La section précédente a été consacrée à préciser la notion de convergence des suites de variables aléatoires et les propriétés qui en découlent. Dans la suivante on étudie les relations d'implication.

1.2 Implications entre les différents modes de convergence

Dans cette section, on s'intéresse aux relations d'implications qui existent entre les différents modes de convergence, déjà définis, afin de déterminer le mode le plus fort. Soient $(X_n)_{n \in \mathbb{N}^*}$ et X des v.a.r. définies sur l'espace $(\Omega, \mathcal{A}, \mathbf{P})$.

Proposition 4 *Li et Hu [32]*

Soient X et $(X_n)_{n \in \mathbb{N}^}$ des v.a.r., alors, pour tout $p \geq 0$, on a*

1.

$$X_n \xrightarrow[n \rightarrow +\infty]{s-L^\infty} X \implies X_n \xrightarrow[n \rightarrow +\infty]{s-L^p} X.$$

2.

$$X_n \xrightarrow[n \rightarrow +\infty]{s^*-L^p} X \implies X_n \xrightarrow[n \rightarrow +\infty]{L^p} X.$$

3.

$$X_n \xrightarrow[n \rightarrow +\infty]{s-L^p} X \implies X_n \xrightarrow[n \rightarrow +\infty]{L^p} X.$$

4. Si $p > 1$

$$X_n \xrightarrow[n \rightarrow +\infty]{s-L^p} X \implies X_n \xrightarrow[n \rightarrow +\infty]{s^*-L^p} X.$$

5. Si $0 < p < 1$

$$X_n \xrightarrow[n \rightarrow +\infty]{s^*-L^p} X \implies X_n \xrightarrow[n \rightarrow +\infty]{s-L^p} X.$$

Preuve 6 La démonstration de cette proposition est déduite directement des définitions des convergences.

Théorème 2 Li et Hu [32]

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. qui converge complètement en moment d'ordre p vers la v.a.r. X . Alors $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers X .

Preuve 7 D'après l'inégalité de Markov, on a

$$\mathbf{P}(|X_n - X|^p > \varepsilon^p) \leq \frac{\mathbf{E}(|X_n - X|^p)}{\varepsilon^p},$$

ce qui implique

$$\sum_{n=1}^{\infty} \mathbf{P}(|X_n - X|^p > \varepsilon^p) \leq \frac{\sum_{n=1}^{\infty} \mathbf{E}(|X_n - X|^p)}{\varepsilon^p} < \infty \quad \left(\text{car } X_n \xrightarrow[n \rightarrow +\infty]{s.L^p} X \right).$$

Donc

$$\sum_{n=1}^{+\infty} \mathbf{P}(|X_n - X|^p > \varepsilon^p) < \infty.$$

Par conséquent si $p = 1$ on a

$$\sum_{n=1}^{+\infty} \mathbf{P}(|X_n - X| > \varepsilon) < \infty,$$

d'où $X_n \xrightarrow[n \rightarrow +\infty]{p.co} X$. ■

Théorème 3 *Polansky [42]*

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. qui converge presque complètement vers la v.a.r. X . Alors $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X .

Preuve 8 *Polansky [42] page 114.*

D'abord on note que, pour $\omega \in \Omega$,

$$\{|X_m(\omega) - X_m| \leq \varepsilon, \text{ pour tout } m \geq n\}^C = \{|X_m(\omega) - X_m| > \varepsilon, \text{ pour quelques } m \geq n\},$$

donc pour $\varepsilon > 0$

$$\mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\} \right) = \mathbf{P} \left(\bigcup_{m=n}^{\infty} \{|X_m(\omega) - X(\omega)| > \varepsilon\} \right).$$

Ce qui implique

$$\mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\} \right) \leq \sum_{m=n}^{\infty} \mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon\} \right).$$

On a par supposition $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers X , c'est à dire pour tout $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \mathbf{P} \left(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\} \right) < \infty.$$

Pour que cette somme converge, il faut que pour tout $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \sum_{m=n}^{\infty} \mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon\} \right) = 0,$$

c'est-à-dire

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| > \varepsilon, \text{ pour quelques } m \geq n\} \right) = 0.$$

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

Comme $\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$ pour tout événement de \mathcal{A} , alors

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\{\omega : |X_m(\omega) - X(\omega)| \leq \varepsilon, \text{ pour tout } m \geq n\} \right) = 1,$$

ce qui est équivalent à la convergence presque sûre de $(X_n)_{n \in \mathbb{N}^*}$ vers X . ■

Proposition 5 Si la suite $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X alors elle converge en probabilité, c-à-d

$$X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X \implies X_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} X.$$

Preuve 9 Supposons que $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X et montrons la convergence en probabilité. Pour $\varepsilon > 0$ fixé, posons,

$$Y_n = 1_{\{|X_n - X| > \varepsilon\}}.$$

Les fonctions $(Y_n)_{n \in \mathbb{N}^*}$ sont mesurables, positives et majorées par 1. D'après l'hypothèse de la convergence presque sûre, pour presque tout $\omega \in \Omega$, il existe un rang n_0 tel que pour tout $n \geq n_0$,

$$|X_n(\omega) - X(\omega)| \leq \varepsilon,$$

ce qui donne

$$Y_n(\omega) = \begin{cases} 1 & \text{si } |X_n(\omega) - X(\omega)| > \varepsilon \\ 0 & \text{si } |X_n(\omega) - X(\omega)| \leq \varepsilon \end{cases}.$$

Ce qui implique $Y_n(\omega) = 0$ pour tous $n \geq n_0$. C'est l'écriture de la convergence presque sûrement de $(Y_n)_{n \in \mathbb{N}^*}$ vers 0. Aussi, on a

$$\begin{aligned} \mathbf{P}(|X_n - X| > \varepsilon) &= \mathbf{E} \left(1_{\{|X_n - X| > \varepsilon\}} \right) \\ &= \mathbf{E}(Y_n), \end{aligned}$$

1.2. Implications entre les différents modes de convergence

et d'après le théorème de la convergence dominée, on obtient

$$\mathbf{E}(Y_n) \xrightarrow[n \rightarrow \infty]{} \mathbf{E}(Y) = 0.$$

■

Proposition 6 Métivier [36]

Si la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers X alors elle converge en loi. C'est-à-dire

$$X_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} X \implies X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X.$$

Preuve 10 Métivier [36] page 189.

D'après la définition (1), pour Φ une fonction continue et bornée, ce qui implique l'uniformément continuité de Φ , avec \mathbf{P}_{X_n} et \mathbf{P}_X qui représentent les lois image de la probabilité \mathbf{P} par les v.a.r. X_n et X respectivement. On a

$$\left| \int \Phi d\mathbf{P}_{X_n} - \int \Phi d\mathbf{P}_X \right| \leq \int |\Phi(X_n) - \Phi(X)| d\mathbf{P}.$$

Pour $\varepsilon > 0, \exists \delta > 0, \forall x, y \in \mathbb{R} / |x - y| \leq \delta \implies |\Phi(x) - \Phi(y)| < \varepsilon$.

$$\begin{aligned} \int |\Phi(X_n) - \Phi(X)| d\mathbf{P} &= \int_{\{\omega / X_n(\omega) - X(\omega) \leq \delta\}} |\Phi(X_n) - \Phi(X)| d\mathbf{P} \\ &\quad + \int_{\{\omega / X_n(\omega) - X(\omega) > \delta\}} |\Phi(X_n) - \Phi(X)| d\mathbf{P} \\ &\leq \varepsilon + 2 \sup_{x \in \mathbb{R}} |\Phi(x)| \mathbf{P}(|X_n - X| > \delta). \end{aligned}$$

Par hypothèse $\forall \delta > 0, \mathbf{P}(|X_n - X| > \delta) \xrightarrow[n \rightarrow \infty]{} 0$. Donc

$$\int \Phi d\mathbf{P}_{X_n} \xrightarrow[n \rightarrow \infty]{} \int \Phi d\mathbf{P}_X. \text{ Alors } X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X.$$

■

Proposition 7 *Métivier [36]*

Si les v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ et X sont dans L^p , alors la convergence en moment d'ordre p de la suite $(X_n)_{n \in \mathbb{N}^*}$ vers X implique sa convergence en probabilité vers la même limite.

$$X_n \xrightarrow[n \rightarrow \infty]{L^p} X \implies X_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} X.$$

Preuve 11 *Métivier [36] page 190.*

Il suffit d'appliquer l'inégalité de Markov. Soit $\varepsilon > 0$,

$$0 \leq \lim_{n \rightarrow \infty} \mathbf{P}(|X_n - X| > \varepsilon) \leq \lim_{n \rightarrow \infty} \frac{1}{\varepsilon^p} \mathbf{E}(|X_n - X|^p) = 0,$$

donc $X_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} X$. ■

Proposition 8 *Polansky [42]*

Soit $(X_n)_{n \in \mathbb{N}}$ converge dans L^p vers X pour $p > 0$, si

$$\sum_{n=1}^{\infty} \mathbf{E}(|X_n - X|^p) < \infty,$$

alors $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$.

Preuve 12 *Voir Polansky [42], théorème 5.4 page 237.*

On a démontré que la convergence en moment complète d'ordre p implique celles dans L^p et presque complète. Cette dernière entraîne la convergence presque sûre. En suite, on a prouvé que la convergence en probabilité est un résultat de la convergence presque sûre et entraîne la convergence en loi. De plus la convergence dans L^p est plus forte que celle en probabilité. Maintenant on cherche les conditions sous lesquelles on obtient les implications inverses.

Proposition 9 Polansky [42]

Si la suite $(X_n)_{n \in \mathbb{N}^*}$ est convergente en loi vers une constante a dans \mathbb{R} , alors elle converge également en probabilité vers a , c'est-à-dire

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} a \Rightarrow X_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} a.$$

Preuve 13 Polansky [42] page 178.

Soient F_n la fonction de répartition de v.a.r. X_n , pour tout $n \in \mathbb{N}^*$ et F celle de la v.a. constante $X = a$. On a F est continue sur $\mathbb{R} \setminus \{a\}$ et comme $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers X , alors

$$\lim_{n \rightarrow +\infty} F_n(x) = F(x),$$

pour tout x différent de a . Or, $\forall \varepsilon > 0$, on peut écrire

$$\begin{aligned} \mathbf{P}(|X_n - a| > \varepsilon) &= 1 - \mathbf{P}(a - \varepsilon \leq X_n \leq \varepsilon + a) \\ &= 1 - F_n(\varepsilon + a) + F_n(-\varepsilon + a) \end{aligned}$$

Puisque $F_n(\varepsilon + a) \rightarrow F(\varepsilon + a) = 1$ et $F_n(-\varepsilon + a) \rightarrow F(-\varepsilon + a) = 0$, lorsque $n \rightarrow +\infty$, avec $-\varepsilon + a$ et $\varepsilon + a$ appartient à $\mathbb{R} \setminus \{a\}$. Alors

$$\lim_{n \rightarrow +\infty} \mathbf{P}(|X_n - a| > \varepsilon) = 0,$$

donc

$$X_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} a.$$

■

Dans la suite et pour des raisons technique de démonstration, on définit le critère des sous suites qui représente une condition fondamentale pour montrer quelques relations entre les modes de convergence.

Proposition 10 Gut [20]

La suite $(X_n)_{n \in \mathbb{N}^*}$ des v.a.r. converge en probabilité vers X si et seulement si pour toute suite croissante d'entiers $(n)_{n \in \mathbb{N}^*}$ on peut extraire une sous suite $(n_k)_{k \in \mathbb{N}^*}$ telle que, $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X

Preuve 14 Gut [20] page 213

Supposons que $X_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} X$. Pour tout $k \in \mathbb{N}^*$, soit (n_k) le plus petit entier tel que

$$\mathbf{P} \left(|X_{n_k} - X| > \frac{1}{k} \right) < \frac{1}{2^k}.$$

Alors

$$\sum_{k \in \mathbb{N}^*} \mathbf{P} \left(|X_{n_k} - X| \geq \frac{1}{k} \right) < +\infty.$$

Comme $\frac{1}{2^k} < \varepsilon$, pour tout $\varepsilon > 0$, ce qui implique $k > \frac{\log(\frac{1}{\varepsilon})}{\log(2)}$. En appliquant le lemme de Borel-Cantelli, on obtient

$$X_{n_k} \xrightarrow[n \rightarrow +\infty]{p.s.} X.$$

■

L'intégrabilité uniforme est un concept important qui joue un rôle essentiel dans la théorie des probabilité, dans la suite nous avons montré une autre relation entre deux modes de convergence par application du théorème associé.

Définition 8 Une famille $(X_n)_{n \in \mathbb{N}^*}$ de v.a.r. sur $(\Omega, \mathcal{A}, \mathbf{P})$ est uniformément intégrable (U.I) ou équi-intégrable, si pour tout c un réel positif.

$$\lim_{c \rightarrow +\infty} \sup_{n \in \mathbb{N}^*} \mathbf{E} (|X_n| 1_{\{X_n > c\}}) = 0.$$

1.2. Implications entre les différents modes de convergence

On a déjà vu que la convergence dans L^1 implique la convergence en probabilité mais, la proposition suivante montre qu'il y a une relation nécessaire entre ces deux modes.

Proposition 11 *Métivier [36]*

On suppose que $(X_n)_{n \in \mathbb{N}^}$ converge en probabilité vers X .*

Si $(X_n)_{n \in \mathbb{N}^}$ est uniformément intégrable, alors $X \in L^1$ et $(X_n)_{n \in \mathbb{N}^*}$ converge dans L^1 vers X .*

Preuve 15 *Soit $\delta > 0$ et pour tout $c > 0$ on a*

$$\begin{aligned}
 \mathbf{E}(|X_n - X|) &= \int_{\Omega} |X_n - X| d\mathbf{P} \\
 &= \int_{\{|X_n - X| < \delta\}} |X_n - X| d\mathbf{P} + \int_{\{|X_n - X| \geq \delta\}} |X_n - X| d\mathbf{P} \\
 &\leq \delta \mathbf{P}(|X_n - X| < \delta) + \int_{\{|X_n - X| \geq \delta\}} |X| d\mathbf{P} + \int_{\{|X_n - X| \geq \delta\}} |X_n| d\mathbf{P} \\
 &\leq \delta + \int_{\{|X_n - X| \geq \delta\}} |X| d\mathbf{P} + \int_{\{|X_n - X| \geq \delta, |X_n| \geq c\}} |X_n| d\mathbf{P} \\
 &\quad + \int_{\{|X_n - X| \geq \delta, |X_n| < c\}} |X_n| d\mathbf{P}.
 \end{aligned}$$

Soit $\varepsilon > 0$, par uniformément intégrabilité de $(X_n)_{n \in \mathbb{N}^}$, il existe $c > 0$, tel que pour tout $n \in \mathbb{N}$, on a*

$$\int_{\{|X_n - X| \geq \delta, |X_n| \geq c\}} |X_n| d\mathbf{P} \leq \int_{\{|X_n| \geq c\}} |X_n| d\mathbf{P} < \frac{\varepsilon}{4}.$$

Choisissons $\delta \leq \frac{\varepsilon}{4}$ et il existe δ_1 tel que

$$\mathbf{P}(|X_n - X| \geq \delta) \leq \delta_1,$$

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

implique que

$$\int_{\{|X_n - X| \geq \delta\}} |X| d\mathbf{P} \leq \frac{\varepsilon}{4}.$$

Par supposition $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité, alors il existe $n_\varepsilon \in \mathbb{N}$, tel que $\forall n \geq n_\varepsilon$ on a

$$\mathbf{P}(|X_n - X| \geq \delta) \leq \delta_1$$

et

$$\mathbf{P}(|X_n - X| \geq \delta) \leq \frac{\varepsilon}{4c}$$

Etant donné $\varepsilon > 0$, $\exists n_\varepsilon \in \mathbb{N}$, tel que

$$\forall n \geq n_\varepsilon, \int_{\Omega} |X_n - X| d\mathbf{P} < \varepsilon,$$

d'où

$$X_n \xrightarrow[n \rightarrow +\infty]{L^1} X.$$

■

Proposition 12 Soient $p \in [1, \infty[$, $(X_n)_{n \in \mathbb{N}^*}$ une suite des v.a.r. dans L^p et X une variable aléatoire. Si

i- $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$.

ii- Il existe une v.a. $Z \in L^p [0, \infty[$ telle que

$$\forall n \in \mathbb{N}^*, |X_n| \leq Z \quad p.s.$$

Alors $X \in L^p$ et $X_n \xrightarrow[n \rightarrow +\infty]{L^p} X$.

1.2. Implications entre les différents modes de convergence

Preuve 16 *Étant donné que $|X_n| \leq Z$ presque sûrement pour tout n et que $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$, nous avons*

$$|X| \leq Z \quad p.s.$$

Par conséquent, $\mathbf{E}(|X|^p) \leq \mathbf{E}(Z^p) < +\infty$ et donc $X \in L^p$. De plus, pour tout $n \in \mathbb{N}$,

$$|X_n - X|^p \leq (|X_n| + |X|)^p \leq 2^p Z^p \quad p.s.$$

Or $2^p Z^p$ est une v.a.r. intégrable car $Z \in L^p$, donc, d'après le théorème de convergence dominée appliqué à la suite de variables aléatoires $(|X_n - X|^p)_{n \in \mathbb{N}^}$ qui converge presque sûrement vers 0, nous avons*

$$\lim_{n \rightarrow +\infty} \mathbf{E}(|X_n - X|^p) = 0.$$

C'est à dire $X_n \xrightarrow[n \rightarrow +\infty]{L^p} X$. ■

La convergence presque complète entraîne la convergence presque sûre, et la réciproque reste vraie avec une condition, comme le montre les théorèmes suivants.

Théorème 4 *Polansky [42]*

Soit $(X_n)_{n \in \mathbb{N}^}$ une suite de v.a.r. indépendantes et a une constante réelle.*

Si

$$X_n \xrightarrow[n \rightarrow +\infty]{p.s.} a,$$

alors

$$X_n \xrightarrow[n \rightarrow +\infty]{p.co} a.$$

.

Preuve 17 Polansky [42] page 115.

Supposons que $X_n \xrightarrow[n \rightarrow +\infty]{p.s} a$, alors pour chaque $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\{ |X_m - a| > \varepsilon, \text{ pour au moins un } m \geq n \} \right) = 0.$$

Notons que

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P} \left(\{ |X_m - a| > \varepsilon, \text{ pour au moins un } m \geq n \} \right) &= \lim_{n \rightarrow \infty} \mathbf{P} \left(\cup_{m=n}^{\infty} \{ |X_m - a| > \varepsilon \} \right) \\ &= \mathbf{P} \left(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \{ |X_m - a| > \varepsilon \} \right) \\ &= \mathbf{P} \left(\limsup_{n \rightarrow \infty} \{ |X_n - a| > \varepsilon \} \right), \end{aligned}$$

et on a la suite des événements $\{ \cup_{m=n}^{\infty} \{ |X_m - a| > \varepsilon \} \}_{n=1}^{\infty}$ est décroissante de manière monotone, donc d'après le théorème de la monotonie séquentielle, on a

$$\mathbf{P} \left(\limsup_{n \rightarrow \infty} \{ |X_n - a| > \varepsilon \} \right) = 0.$$

La suite $(X_n)_{n \in \mathbb{N}^*}$ est de variables aléatoires indépendantes donc $\{ |X_m - C| > \varepsilon \}_{n=1}^{\infty}$ est une suite des événements indépendants pour tout $\varepsilon > 0$. Donc d'après le lemme de Borell-Cantelli on a pour $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbf{P} (|X_n - a| > \varepsilon) < \infty,$$

c'est à dire $X_n \xrightarrow[n \rightarrow +\infty]{p.co} a.. \blacksquare$

1.3 Taux de convergence

Dans cette section, on définit la notion du taux ou de vitesse de convergence d'une suite réelle ou aléatoire.

Définition 9

1. Soient $(x_n)_{n \in \mathbb{N}^*}$ et $(y_n)_{n \in \mathbb{N}^*}$ deux suites de nombres réels. La notation $x_n = O(y_n)$ quand $n \rightarrow +\infty$, signifie que la suite $\left(\left|\frac{x_n}{y_n}\right|\right)_{n \in \mathbb{N}^*}$ reste bornée quand $n \rightarrow +\infty$.
2. On dit qu'une suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge vers une limite X avec un taux U_n , et on note

$$X_n - X = O(U_n),$$

s'il existe une constante positive M , telle que

$$|X_n - X| \leq MU_n.$$

Pour $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. on donne les définitions des taux de convergences presque complète, presque sûre et en probabilité et on étudie les relations d'implication entre eux. Soit $(U_n)_{n \in \mathbb{N}^*}$ une suite de nombres réels positifs.

Définition 10 On dit que la suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers la v.a.r. X avec un taux U_n , si et seulement si

$$\lim_{m \rightarrow \infty} \overline{\lim}_{n \rightarrow \infty} \mathbf{P}(|X_n - X| < mU_n) = 1,$$

où m est un nombre réel positif et on note

$$X_n - X = O_{\mathbf{P}}(U_n).$$

Définition 11 On dit que la suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers la v.a.r. X avec un taux U_n , si et seulement si

$$\mathbf{P}(\exists c < \infty, \exists n, \forall m > n, |X_m - X| \leq cU_m) = 1.$$

Et on note

$$X_n - X = O_{p.s.}(U_n).$$

Définition 12 *On dit que la suite des v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers la v.a. X avec un taux U_n , si et seulement si*

$$\exists \varepsilon_0 > 0, \sum_{n=0}^{\infty} \mathbf{P}(|X_n - X| \geq \varepsilon_0 U_n) < \infty,$$

et on note

$$X_n - X = O_{p.co.}(U_n).$$

La proposition suivante montre que si une suite $(X_n)_{n \in \mathbb{N}^*}$ de v.a.r. converge presque complètement vers une limite X , avec un taux U_n , alors elle converge presque sûrement et en probabilité vers la même limite X et avec le même taux U_n .

Proposition 13 *Ferraty [14]*

Soit $(X_n)_{n \in \mathbb{N}^}$ une suite de v.a.r., telle que $X_n - X = O_{p.co.}(U_n)$. Alors*

1- $X_n - X = O_{\mathbf{P}}(U_n)$.

2- $X_n - X = O_{p.s.}(U_n)$.

Preuve 18 *Ferraty [14] page 230.*

1. *On démontre ce résultat pour $X = 0$ p.s.*

Supposons qu'il existe $\varepsilon_0 = m > 0$, tel que

$$\sum_{n \geq 1} \mathbf{P}(|X_n| > mU_n) < \infty,$$

ce qui nous permet d'écrire

$$\exists m_0, \forall m > m_0, \sum_n \mathbf{P}(|X_n| > mU_n) < \infty,$$

et d'après le théorème de Borel-Cantelli, on a

$$\exists m_0, \forall m > m_0, \mathbf{P} \left(\limsup_{n \rightarrow \infty} \{|X_n| > mU_n\} \right) = 0.$$

Puis on appliquant le lemme de Fatou, on obtient :

$$\exists m_0, \forall m > m_0, \overline{\lim}_{n \rightarrow \infty} \mathbf{P} (|X_n| > mU_n) = 0,$$

ce qui est équivalent à

$$\exists m_0, \forall m > m_0, \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbf{P} (|X_n| \leq mU_n) = 1.$$

2. De la même manière, et par application du théorème de Borel-Cantelli, on obtient

$$\mathbf{P} \left(\overline{\lim}_{n \rightarrow \infty} \{|X_n| > \varepsilon_0 U_n\} \right) = 0,$$

qui peut être s'écrit

$$\exists \varepsilon_0, \mathbf{P}(\exists n, \forall m > n, |X_m| \leq \varepsilon_0 U_m) = 1,$$

ce qui nous donne le résultat cherché.

■

Un autre résultat concernant les règles de calcul des taux de convergence presque complète est donné par la proposition ci dessous.

Proposition 14 Ferraty [14]

Soient $(X_n)_{n \in \mathbb{N}^*}, (Y_n)_{n \in \mathbb{N}^*}$ deux suites de variables aléatoires, l_X, l_Y deux nombres réels. Si $X_n - l_X = O_{p.co.}(U_n)$ et $Y_n - l_Y = O_{p.co.}(U_n)$ avec $\lim_{n \rightarrow \infty} U_n = 0$. Alors

$$1- (X_n + Y_n) - (l_X + l_Y) = O_{p.co.}(U_n);$$

$$2- (X_n \cdot Y_n) - (l_Y \cdot l_X) = O_{p.co.}(U_n);$$

$$3- \frac{1}{\bar{X}_n} - \frac{1}{l_X} = O_{p.co.}(U_n), \text{ pour } l_X \neq 0$$

Preuve 19 *Pour la démonstration de ces règles de calcul des taux de la convergence presque complète, on se référer à Ferraty [14] page 231.*

Après étudier les différentes relations entre les modes de convergence définis dans la première section, on énonce les théorèmes de convergence les plus utilisés en statistique.

1.4 Utilisation des concepts de convergence en statistiques

Lorsque on fait un sondage aléatoire dans une population, l'augmentation de la taille de l'échantillon implique que la moyenne de l'échantillon se rapproche de plus en plus de celle de la population. D'où l'intérêt des théorèmes des lois des grands nombres. Il existe plusieurs versions de ce théorème et sous différentes conditions.

Théorème 5 *(La loi faible des grands nombres) Métivier [36]*

Soit $(X_n)_{n \in \mathbb{N}^}$ une suite des v.a.r. indépendantes telles que $\mathbf{E}(X_n)$ existe et égal à μ_n , avec*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \mu_k = \mu,$$

aussi la variance $\sigma^2(X_n)$ existe et est notée à σ_n^2 , avec

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sqrt{\sum_{k=1}^n \sigma_k^2} = 0,$$

alors

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} \mu.$$

Preuve 20 On veut montrer la convergence en probabilité de \bar{X}_n vers μ , c'est à dire pour $\varepsilon > 0$, $\lim_{n \rightarrow +\infty} \mathbf{P}(|\bar{X}_n - \mu| > \varepsilon) = 0$, pour cela on applique l'inégalité de Tchebychev

$$\mathbf{P}(|\bar{X}_n - \mathbf{E}(\bar{X}_n)| > \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2}.$$

Or

$$\mathbf{E}(\bar{X}_n) = \mathbf{E}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n} \sum_{k=1}^n \mathbf{E}(X_k) = \mu,$$

et

$$\text{Var}(\bar{X}_n) = \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = \frac{\sum_{k=1}^n \sigma_k^2}{n^2}.$$

Car les v.a.r. sont indépendants par hypothèse. On en déduit que

$$\mathbf{P}(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\sum_{k=1}^n \sigma_k^2}{n^2 \varepsilon^2}.$$

Or

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \frac{\sum_{k=1}^n \sigma_k^2}{n^2 \varepsilon^2} = 0,$$

d'où

$$\lim_{n \rightarrow +\infty} \mathbf{P}(|\bar{X}_n - \mu| > \varepsilon) = 0.$$

La v.a.r. \bar{X}_n converge donc en probabilité vers μ . ■

Théorème 6 (La loi forte des grands nombres) Métivier [36]

Soit $(X_n)_{n \geq 1}$ une suite de v.a.r. indépendantes de variance finie, si $\mathbf{E}(X_n)$ existe et égal à μ_n , avec

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \mu_k = \mu,$$

1. MODES DE CONVERGENCE DES SUITES DE VARIABLES ALÉATOIRES

et si pour tout $n > 0$, $\text{var}(X_n) = \sigma_n^2$, vérifiant

$$\sum_{n=1}^{+\infty} \frac{\sigma_n^2}{n^2} < +\infty,$$

alors

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow +\infty]{p.s.} \mu.$$

Preuve 21 *La démonstration se fait sans difficulté voir Métivier [36], page 204. pour les détails.*

Dans le cas où les v.a.r. $(X_n)_{n \geq 1}$ sont de même loi, on peut énoncer le théorème 3 sans hypothèse sur les variance, comme suit

Théorème 7 (*Loi forte des grands nombres*) Métivier [36]

Si les v.a.r. $(X_n)_{n \geq 1}$ sont indépendantes, de même loi telle que $\mathbf{E}(X_n) = \mu < +\infty$, alors

$$\frac{1}{n} \sum_{i=1}^n X_k \xrightarrow[n \rightarrow +\infty]{p.s.} \mu.$$

Preuve 22 *Ce théorème est un cas particulier du théorème 6. Donc la démonstration se fait de la même manière. Voir Métivier [36], page 205 pour plus de détails.*

Le théorème suivant, appelé théorème central limite, est un résultat important de la convergence en loi d'une suite de variables aléatoires réelles. Il affirme que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire gaussienne. Souvent utilisé pour montrer la normalité asymptotique d'un estimateur.

Théorème 8 (*Théorème central limite*) Ouvrad [38]

Soient $(X_n)_{n \in \mathbb{N}^}$ une suite de v.a.r. indépendantes, définies sur le même espace $(\Omega, \mathcal{A}, \mathbf{P})$, suivant la même loi de moyenne μ et de variance σ^2*

1.4. Utilisation des concepts de convergence en statistiques

supposées finies et X v. a. r. de loi $\mathcal{N}(0, 1)$. Alors la variable aléatoire

$S_n = \sum_{i=1}^n X_i$ vérifie

$$\frac{S_n - n\mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X.$$

Preuve 23 Pour la démonstration voir Ouvrad [38], page 314.

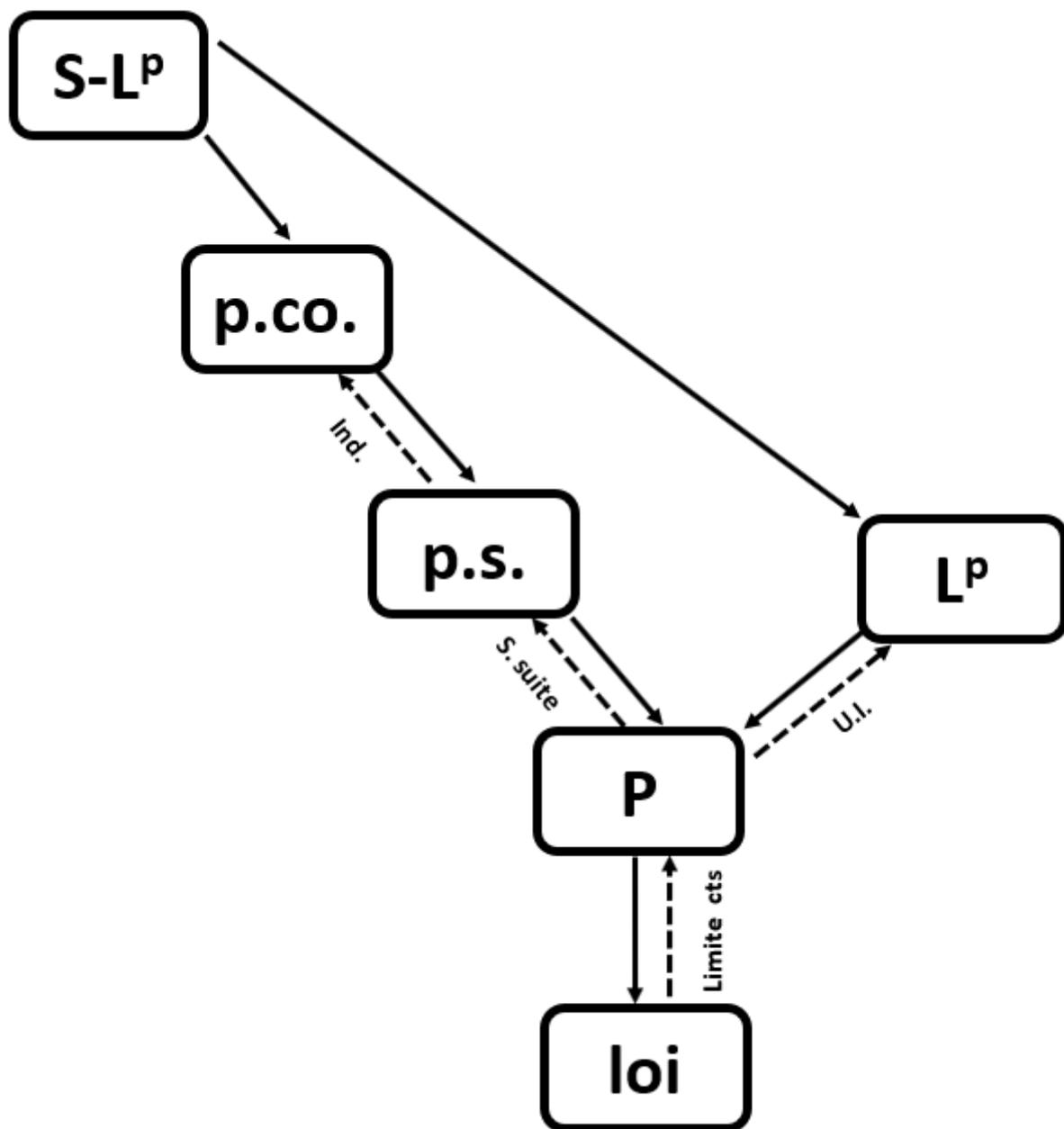


FIGURE 1.1 – Résumé des relations d'implications entre les modes de convergence étudiés.

2 Estimation à noyau des caractéristiques fonctionnelles d'une loi de probabilité

La méthode d'estimation non paramétrique à noyau, appelé aussi de Rosenblatt-Parzen, est introduite en 1956 par Rosenblatt [45] puis améliorée en 1962 par Parzen [40]. En se basant sur un échantillon de la population statistique, cette méthode permet d'approcher la fonction à estimer en tout points du support. Dans ce chapitre on s'intéresse aux propriétés asymptotiques des estimateurs à noyaux des caractéristiques fonctionnelles d'une loi de probabilité inconnue. En fixant l'étude sur le biais, la variance, erreur moyenne quadratique et la convergence presque complète des estimateurs des fonctions de densité de probabilité, de répartition et de quantile.

2.1 Méthode d'estimation non paramétrique à noyau

Les estimateurs à noyaux sont des fonctions de deux paramètres ; une fonction K appelée noyau représentant le poids de chacune des observations dans l'estimation et un paramètre dit fenêtre h qui détermine le degré de lissage de l'estimateur. L'estimation par noyau est l'une des méthodes les plus utilisées dans le domaine non paramétrique car elle semble pratique, robuste et ne nécessite pas un choix multiple des paramètres K et h .

2.1.1 Les paramètres de la méthode d'estimation à noyau

La fenêtre h

C'est un paramètre positif, appelé aussi paramètre de lissage, il détermine si l'estimation est tout à fait lisse ou bien qu'elle est plutôt très irrégulière. Plus h est grand, plus l'estimateur est régulier, contrairement au cas où h est petit, l'estimateur est irrégulier. Le sur-lissage, lorsque $h \rightarrow +\infty$, peut masquer la plupart des propriétés de la vraie fonction à estimer telles que l'asymétrie ou la multimodalité, tandis que le sous-lissage, lorsque $h \rightarrow 0$, fait apparaître des détails artificiels sur le graphique de l'estimateur. la figure suivante illustre le rôle de h dans le lissage de l'estimation.

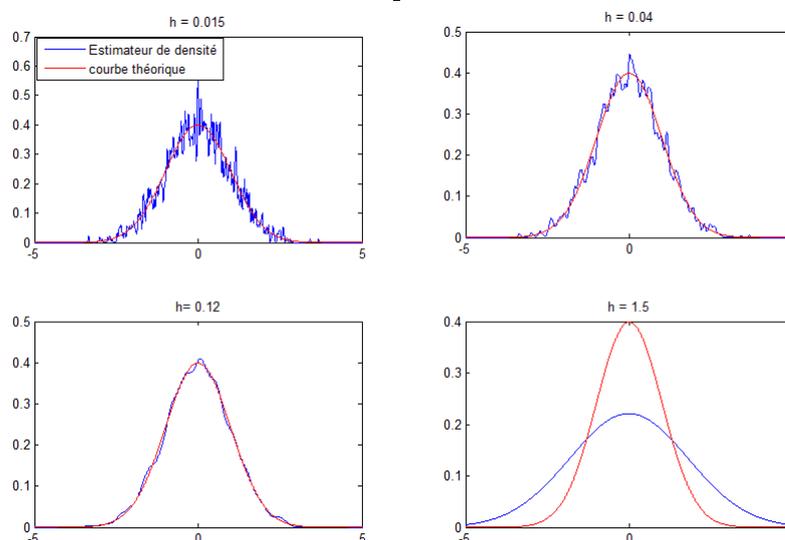


FIGURE 2.1 – Influence du paramètre de lissage sur l'estimation de la fonction de densité pour une courbe théorique gaussienne.

Le but de l'estimation de h est de modifier les données de manière à obtenir des estimateurs dont les caractéristiques convergent aux propriétés des paramètres réels. En général, h est obtenu par des techniques de validation croisée.

Le noyau K

Un noyau continu K est une fonction borélienne, positive et intégrable d'intégrale égale à 1.

2.1. Méthode d'estimation non paramétrique à noyau

Le noyau est l'élément qui détermine la forme des "bosses" qui constituent l'estimation de la densité. Plusieurs travaux ont été effectués pour trouver un noyau préférable d'utilisation à tous les autres, autrement dit un noyau optimal. En 1969, Epanechnikov [9] a proposé sous certains critères, un noyau optimal qui porte son nom. Puis, en 1983, Rao [44] est arrivé à la conclusion que le choix d'un noyau autre que le noyau optimal n'entraînait qu'une légère perte de précision.

L'ordre d'un noyau K représente l'ordre du premier moment non nul, il est paire si le noyau est symétrique. Si le noyau est positif, il est du second ordre.

Les noyaux dont l'ordre est supérieur à 2 ont des parties négatives, par conséquent ils ne sont pas des densités de probabilité. Pour plus de détails on se réfère à Hansen [25].

Les noyaux sont classés suivant le support en deux classes : à support fini, c'est-à-dire des noyaux K dont le support est de la forme $[-a, a]$, pour $a \in \mathbb{R}$. Des noyaux à support infini sont tels que $K(x) \rightarrow 0$ pour $|x| \rightarrow +\infty$.

On distingue deux types de noyaux symétriques et asymétriques. Dans la littérature, la plupart des noyaux utilisés dans un contexte d'estimation de fonction de densité sont symétriques.

Voici quelques exemples sur les noyaux.

Noyaux	Expression mathématique
Epanechnikov	$K(u) = \frac{3}{4}(1 - t^2)1_{[-1,1]}(u)$.
Gaussien	$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right), u \in \mathbb{R}$.
Triangulaire	$K(u) = (1 - u)1_{[-1,1]}(u)$.
Gamma	$K(u) = \frac{u^{\frac{x}{h}} \exp(-\frac{u}{h})}{\Gamma(\frac{x}{h}+1)h^{\frac{x}{h}+1}}, h > 0, u \in \mathbb{R}^+,$
Silverman	$K(u) = \frac{1}{2} \exp\left(\frac{- u }{\sqrt{2}}\right) \sin\left(\frac{ u }{\sqrt{2}} + \frac{\pi}{4}\right), u \in \mathbb{R}$.
Uniforme	$K(u) = \frac{1}{2}1_{[-1,1]}(u)$.

TABLE 2.1 – Quelques noyaux.

Les représentations graphiques des noyaux Epanechnikov, gaussien (normal) et triangulaire, Silverman et uniforme sont données par les figures (2.2), (2.3), (2.4), (2.5) et (2.6).

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

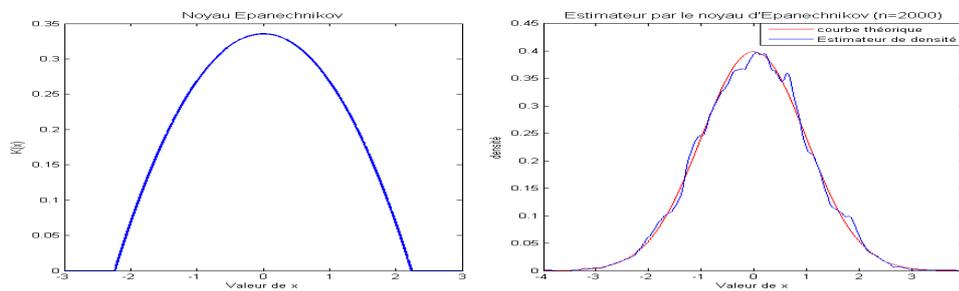


FIGURE 2.2 – Noyau d'Epanechnikov ou parabolique.

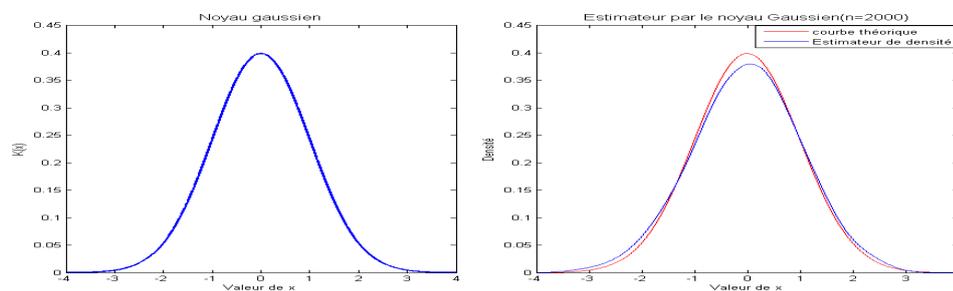


FIGURE 2.3 – Noyau Gaussien.

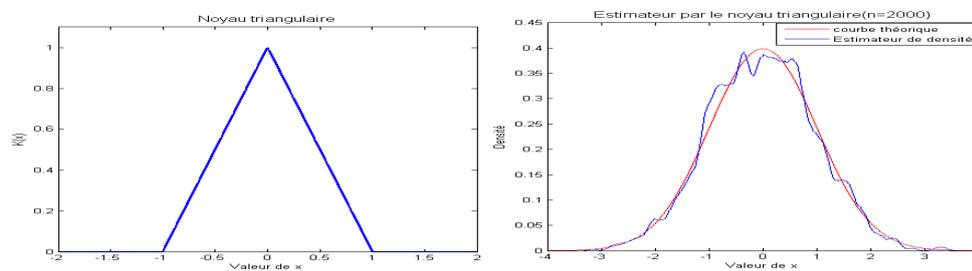


FIGURE 2.4 – Noyau triangulaire.

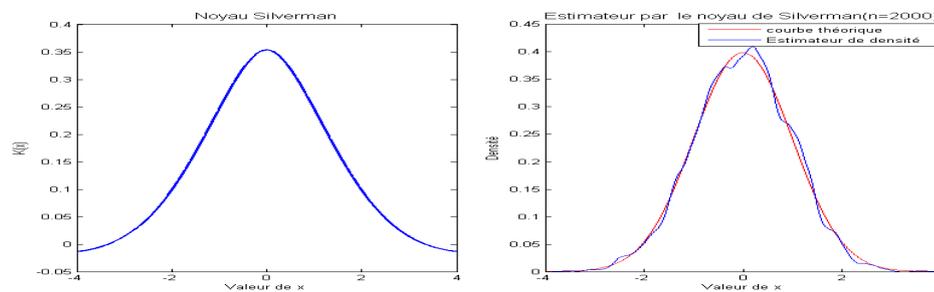


FIGURE 2.5 – Noyau de Silverman.

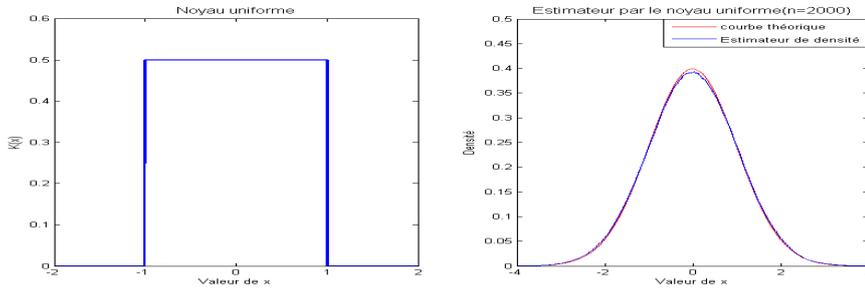


FIGURE 2.6 – Noyau uniforme ou de Rozenblatt.

2.2 Estimateur à noyau d'une densité de probabilité

L'estimation de la fonction de densité de probabilité a une place importante dans l'étude de nombreux phénomènes aléatoires, car la densité permet d'avoir un aperçu rapide des principales caractéristiques de la distribution, comme par exemple : les pics, les cavités, l'asymétrie, etc Plusieurs estimateurs résultats des différentes méthodes d'estimation, ont été proposés pour estimer la densité, ici on s'intéresse aux estimateurs à noyaux.

Considérons une v.a.r. X de densité de probabilité f . Cette dernière donne une description de la distribution de X et permet de déduire des probabilités associées à X de la relation

$$\mathbf{P}(a < X < b) = \int_a^b f(x) dx, \text{ pour tout } a < b.$$

Supposons qu'on a un échantillon X_1, \dots, X_n de v.a.r. de même loi que la v.a.r. X admettant une densité de probabilité f absolument continue par rapport à la mesure de *Lebesgue* sur \mathbb{R} à valeurs dans $[0, +\infty[$, supposée inconnue. Rappelons que la valeur de f au point x , représente la probabilité que la v.a.r. X appartienne à un petit voisinage de x , c'est-à-dire,

$$f(x) = \lim_{h \rightarrow 0} \mathbf{P}(x - h < X < x + h), \quad h > 0.$$

On peut donc estimer cette probabilité par la proportion de l'échantillon X_1, \dots, X_n appartenant à l'intervalle $(x - h, x + h)$. Ainsi un estimateur naturel de f , noté \hat{f}_n , est donné par

$$\begin{aligned} \hat{f}_n(x) &= \frac{1}{2nh} [\text{nombre } X_1, \dots, X_n \in (x - h, x + h)] \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K_0\left(\frac{x - X_i}{h}\right), \end{aligned}$$

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

où K_0 est la fonction du poids définie par

$$K_0(x) = \begin{cases} \frac{1}{2}, & \text{si } |x| < 1 \\ 0, & \text{sinon} \end{cases}.$$

De la définition de K_0 il découle que \hat{f}_n n'est pas continu. Donc pour obtenir la continuité de l'estimateur de la densité continue f , on cherche une généralisation de \hat{f}_n en remplaçant la fonction du poids K_0 par une fonction continue K satisfaisant

$$\int_{\mathbb{R}} K(x) dx = 1,$$

appelée fonction du noyau. Cette densité de probabilité n'est pas obligatoirement symétrique. L'estimateur obtenu est donc de la forme

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right).$$

Cet estimateur est formé par la somme ou plutôt la moyenne des courbes de K , comme le montre la figure suivante pour un noyau gaussien.

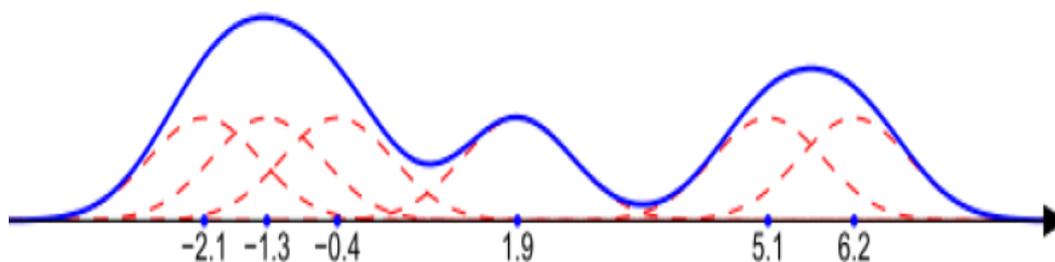


FIGURE 2.7 – Estimateur à noyau gaussien \hat{f}_6 de la densité en 6 observations.

L'estimateur à noyau d'une densité de probabilité inconnue f , noté \hat{f}_n , défini par Parzen (1962) [40], Rosenblatt (1956)[45], est donné par

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad (2.1)$$

où $(h_n)_{n \geq 1}$ est une suite de nombres réels positifs vérifiant $h_n \rightarrow 0$ quand $n \rightarrow \infty$. K est une fonction borélienne, positive et intégrable, d'intégrale sur \mathbb{R} égale à 1, appelée fonction du noyau ou noyau simplement.

2.2.1 Propriétés asymptotiques de l'estimateur à noyau d'une densité de probabilité

Quand on définit un estimateur d'une fonction ou d'un paramètre inconnus, il y a certain nombre de critères qui permettent d'évaluer la similarité de l'estimateur à la fonction ou au paramètre à estimer. Parmi ces critères proposés dans la littérature, on trouve le biais, la variance, l'erreur moyenne quadratique, la convergence ou la consistance d'un estimateur et la normalité asymptotique.

Rappelons que l'estimateur à noyau \hat{f}_n , de la fonction de densité de probabilité, est une fonction de deux paramètres, le noyau K qui établit l'aspect du voisinage de x et le paramètre h_n qui contrôle la taille de ce voisinage. Donc pour obtenir de bonnes propriétés asymptotiques de \hat{f}_n , il faut bien choisir la valeur du paramètre primordial h_n , cependant le noyau K ne doit pas être négligé, car il joue le rôle de réducteur du biais, comme le montrent les résultats du travail de Parzen [40] en 1962.

Pour vérifier les critères de qualité de \hat{f}_n , supposons que le noyau K et le paramètre de lissage h_n vérifient les hypothèses suivantes

$$H_1 - \int_{\mathbb{R}} K(x) dx = 1; \sup_{x \in \mathbb{R}} |K(x)| < \infty; \int_{\mathbb{R}} |K(x)| dx < \infty.$$

$$H_2 - h_n \rightarrow 0 \text{ quand } n \rightarrow +\infty.$$

$$H_3 - nh_n \rightarrow +\infty \text{ quand } n \rightarrow +\infty.$$

$$H_4 - nh_n^2 \rightarrow +\infty \text{ quand } n \rightarrow +\infty.$$

$$H_5 - \frac{nh_n}{\log n} \rightarrow +\infty \text{ quand } n \rightarrow +\infty.$$

Les propriétés du biais et variance de \hat{f}_n sont obtenues en appliquant le théorème de Bochner [5], en 1955, rappelé ci-dessous.

Théorème 9 Soient g et K deux fonctions intégrables, avec K bornée vérifiant $\lim_{|y| \rightarrow \infty} |y| K(y) = 0$. On pose

$$g_n(x) = \frac{1}{h_n} \int_{\mathbb{R}} K\left(\frac{y}{h_n}\right) g(x-y) dy,$$

où $(h_n)_{n \geq 1}$ est une suite de nombres réels positifs vérifiant $h_n \rightarrow 0$ quand

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES
D'UNE LOI DE PROBABILITÉ

$n \rightarrow \infty$. Si g est continue au point $x \in \mathbb{R}$, alors

$$\lim_{n \rightarrow \infty} g_n(x) = g(x) \int_{\mathbb{R}} K(y) dy.$$

Et si g est uniformément continue, alors la convergence de g_n est uniforme.

Biais : Le biais de l'estimateur \hat{f}_n est défini par

$$\begin{aligned} \text{Biais} \left(\hat{f}_n(x) \right) &= \mathbf{E} \left(\hat{f}_n(x) \right) - f(x) \\ &= \frac{1}{h_n} \mathbf{E} \left(K \left(\frac{x - X}{h_n} \right) \right) - f(x) \\ &= \frac{1}{h_n} \int_{\mathbb{R}} K \left(\frac{x - u}{h_n} \right) f_X(u) du - f(x) \end{aligned}$$

Suivant Silverman [47] la valeur approximative du biais est

$$\text{Biais} \left(\hat{f}_n(x) \right) = \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz, \quad (2.2)$$

où $z = \frac{x-u}{h_n}$.

L'estimateur \hat{f}_n de la fonction de densité est asymptotiquement non biaisé, comme le confirme le corollaire suivant.

Corollaire 1 Parzen [40]

Pour f continue et sous les hypothèses H_1, H_2 avec K vérifie $\lim_{|y| \rightarrow \infty} |y| K(y) = 0$, l'estimateur \hat{f}_n est asymptotiquement non biaisé. C'est à dire qu'il satisfait

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\hat{f}_n(x) \right] = f(x).$$

Preuve 24 On a par définition de l'espérance

$$\begin{aligned} \mathbf{E} \left[\hat{f}_n(x) \right] &= \frac{1}{h_n} \int_{\mathbb{R}} K \left(\frac{x - u}{h_n} \right) f_X(u) du \\ &= \int_{\mathbb{R}} K(y) f_X(x - yh) dy, \end{aligned}$$

2.2. Estimateur à noyau d'une densité de probabilité

comme f est intégrable et continue en x et k est intégrable avec $|y| K(y) \rightarrow 0$ lorsque $|y| \rightarrow \infty$. Donc d'après le théorème de Bochner [5], on a

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} K(y) f_X(x - yh) dy = f_X(x) \int_{\mathbb{R}} K(y) dy,$$

et du fait que $\int_{\mathbb{R}} K(x) dx = 1$, on obtient

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\hat{f}_n(x) \right] = f(x).$$

Variance : la variance de l'estimateur \hat{f}_n de la fonction de densité est par définition

$$\begin{aligned} \text{var} \left(\hat{f}_n(x) \right) &= \mathbf{E} \left[\hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right]^2 \\ &= \frac{1}{nh_n^2} \left[\mathbf{E} \left(K^2 \left(\frac{x - X}{h_n} \right) \right) - \mathbf{E}^2 \left(K \left(\frac{x - X}{h_n} \right) \right) \right]. \end{aligned}$$

D'après Silverman [47] la valeur de la variance est

$$\text{var} \left(\hat{f}_n(x) \right) = \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(z) dz. \quad (2.3)$$

Le théorème ci-dessus montre que la variance, de l'estimateur \hat{f}_n , tend vers 0 quand n tend vers l'infinie.

Théorème 10 Parzen [40]

Pour f continue et sous les hypothèses H_1, H_2, H_3 avec K vérifie $\lim_{|y| \rightarrow \infty} |y| K(y) = 0$, la variance de l'estimateur \hat{f}_n satisfait

$$\lim_{n \rightarrow \infty} \text{Var} \left[\hat{f}_n(x) \right] = 0,$$

en tout point x de continuité de f .

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES
D'UNE LOI DE PROBABILITÉ

Preuve 25 On a par définition de la variance de l'estimateur \hat{f}_n

$$\begin{aligned}
 \text{Var}[\hat{f}_n(x)] &= \frac{1}{n^2 h_n^2} \text{Var} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \\
 &= \frac{1}{n h_n^2} \text{Var} \left[K\left(\frac{x - X_i}{h_n}\right) \right] \\
 &= \frac{1}{n h_n^2} \mathbf{E} \left[K^2\left(\frac{x - X_i}{h_n}\right) \right] - \frac{1}{n h_n^2} \mathbf{E}^2 \left[K\left(\frac{x - X_i}{h_n}\right) \right] \\
 &\leq \frac{1}{n h_n^2} \mathbf{E} \left[K^2\left(\frac{x - X}{h_n}\right) \right] \\
 &\leq \frac{1}{n h_n^2} \int_{\mathbb{R}} K^2\left(\frac{x - u}{h_n}\right) f_X(u) du \\
 &\leq \frac{1}{n h_n} \left(\frac{1}{h_n} \int_{\mathbb{R}} K^2\left(\frac{y}{h_n}\right) f_X(x - y) dy \right).
 \end{aligned}$$

Comme K est bornée et intégrable, alors elle est de carrée intégrable. Donc par application du théorème de Bochner [5] et sous H_3 on obtient

$$\lim_{n \rightarrow \infty} \frac{1}{n h_n^2} \int_{\mathbb{R}} K^2\left(\frac{y}{h_n}\right) f_X(x - y) dy = \frac{1}{n h_n} f_X(x) \int_{\mathbb{R}} K^2(y) dy,$$

d'où

$$\lim_{n \rightarrow \infty} \text{Var} \left[\hat{f}_n(x) \right] = 0.$$

Remarque 5 On remarque que le biais ne dépend pas directement de la taille de l'échantillon n , mais plutôt du noyau K et du paramètre de lissage h . Cependant, la variance dépend des ces trois paramètres.

Erreur moyenne quadratique (MSE) ponctuelle : L'erreur moyenne quadratique de l'estimateur \hat{f}_n est définie par

$$\text{MSE}(x) = \mathbf{E} \left[\hat{f}_n(x) - f(x) \right]^2.$$

Le développement de cette expression fait apparaitre la variance et le biais ponctuels de \hat{f}_n , comme suit

$$\begin{aligned} MSE(x) &= \mathbf{E} \left[\hat{f}_n(x)^2 \right] - 2\mathbf{E} \left[\hat{f}_n(x) \right] f(x) + \mathbf{E} [f(x)]^2 \\ &= \mathbf{E} \left[\hat{f}_n(x)^2 \right] + \mathbf{E} \left[\hat{f}_n(x) \right]^2 - \mathbf{E} \left[\hat{f}_n(x) \right]^2 - 2\mathbf{E} \left[\hat{f}_n(x) \right] f(x) + [f(x)]^2 \\ &= var \left(\hat{f}_n(x) \right) + biais^2 \left(\hat{f}_n(x) \right) \end{aligned}$$

La dernière expression montre le compromis pour la minimisation du MSE entre le biais (erreur symétrique) et la variance (erreur aléatoire). A cause du terme $\mathbf{E} \left(\hat{f}_n(x) \right)$, on voit qu'une réduction de biais entraîne une augmentation de la variance et vice versa.

En remplaçant (2.2) et (2.3) dans la dernière expression du MES on obtient

$$MSE(x) = \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(z) dz + \left(\frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz \right)^2.$$

Erreur moyenne quadratique intégrée ($MISE$)

C'est une mesure globale d'efficacité de \hat{f}_n , elle est obtenue en intégrant le MSE sur tout le support de f , elle s'écrit

$$\begin{aligned} MISE(x) &= \int_{\mathbb{R}} MSE(x) dx = \int_{\mathbb{R}} var \left(\hat{f}_n(x) \right) dx + \int_{\mathbb{R}} biais^2 \left(\hat{f}_n(x) \right) dx \\ &= \int_{\mathbb{R}} \left(\frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(z) dz \right) dx + \int_{\mathbb{R}} \left(\frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz \right)^2 dx. \end{aligned}$$

La valeur idéale du paramètre de lissage h_n du point de vue de la minimisation de $MISE$ est obtenu par Parzen [40] (lemme 4A) et est égale à

$$h_{opt} = \left(\int_{\mathbb{R}} z^2 K(z) dz \right)^{-\frac{2}{5}} \left(\int_{\mathbb{R}} K^2(z) dz \right)^{\frac{1}{5}} \left(\int_{\mathbb{R}} f^{(2)}(x)^2 dz \right)^{-\frac{1}{5}} n^{-\frac{1}{5}}.$$

Convergence et consistance

Parmi toutes les qualités que peut avoir un estimateur (ou une suite d'estimateurs), on s'intéresse souvent à sa consistance, c'est-à-dire, au fait qu'une suite d'estimateurs $\left(\hat{f}_n \right)_{n \in \mathbb{N}}$ converge ou non vers f en une distance

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

$d(\hat{f}_n, f)$ donnée. Dans le paragraphe suivant on donne quelques résultats de convergence des estimateurs à noyaux de la densité de probabilité dans la littérature.

Convergence en moyenne quadratique ponctuelle

La convergences en moyenne quadratique ponctuelle (MSE), a été établi par Parzen [40] en 1962 et en 1963, Tiago de Oliveira [53] a montré celle du MISE, comme le montrent les deux théorèmes suivants

Théorème 11 *Parzen [40]*

Soit \hat{f}_n l'estimateur à noyau de la densité de probabilité f . Supposons que le noyau K vérifie H_1 et que le paramètre de lissage h_n satisfait H_2 et H_3 . Alors

$$MSE\left(\hat{f}_n(x)\right) \xrightarrow{n \rightarrow +\infty} 0, \forall x \in \mathbb{R}.$$

Théorème 12 *Tiago de Oliveira [53]*

Soit \hat{f}_n l'estimateur à noyau de la densité de probabilité f satisfaisant $\int |f(x)|^p dx < \infty$. Supposons que le noyau K vérifie H_1 et que le paramètre de lissage h_n satisfait H_2 et H_3 . Alors

$$MISE\left(\hat{f}_n(x)\right) \xrightarrow{n \rightarrow +\infty} 0, \forall x \in \mathbb{R}.$$

Consistances faible et forte

En 1962, Parzen [40] et en 1956, Nadaraya [37] ont montré la consistance faible tandis que Silverman [47] en 1986, a proposé quant à lui la consistance forte.

Théorème 13 *Parzen [40]*

Sous les hypothèses H_1, H_4 et si la transformée de Fourier $\int \exp(-izx) K(x) dx$ est absolument intégrable, alors

$$\sup_{x \in \mathbb{R}} \left| \hat{f}_n(x) - f(x) \right| \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} 0.$$

Théorème 14 *Nadaraya [37]*

Soit f une densité de probabilité uniformément continue et \hat{f}_n son estimateur à noyau K positif et à variations bornées. Pour tout h_n satisfait H_2 tel que $\sum_{n \geq 1} \exp(-\varepsilon n h_n) < +\infty$. Alors

$$\sup_{x \in \mathbb{R}} \left| \hat{f}_n(x) - f(x) \right| \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} 0.$$

Théorème 15 *Silverman [47]*

Soit f une densité de probabilité uniformément continue et \hat{f}_n son estimateur à noyau K positif et à variations bornées. Pour tout h_n satisfait H_2 et H_5 . Alors

$$\sup_{x \in \mathbb{R}} \left| \hat{f}_n(x) - f(x) \right| \xrightarrow[n \rightarrow +\infty]{p.s.} 0.$$

Normalité asymptotique

Il est utile de connaître la loi d'un estimateur, car elle permet de calculer ces caractéristiques et de construire un intervalle de confiance. En général, cette loi est obtenue en utilisant une convergence en loi, ou le théorème centrale limite en particulier. La propriété de la normalité asymptotique affirme que la loi de l'estimateur est gaussien lorsque $n \rightarrow +\infty$. Ce critère de l'estimateur à noyau de la fonction de densité, a été tirée des travaux de Parzen (1962) [40].

Théorème 16 *Parzen [40]*

Soit f une densité de probabilité et \hat{f}_n son estimateur à noyau K vérifiant H_1 . Pour tout h_n satisfait H_2 et H_3 . Alors

$$\frac{\hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right)}{\sqrt{\text{var} \left(\hat{f}_n(x) \right)}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1), \forall x \in \mathbb{R}.$$

2.2.2 Convergence presque complète de l'estimateur à noyau de la densité de probabilité

Intéressons nous maintenant à l'étude de la convergence presque complète de \hat{f}_n , et supposons de plus que la densité de probabilité f , le paramètre de lissage h_n et le noyau K , satisfassent les hypothèses suivantes

H_6 — f est une fonction continue au voisinage de x , où x est un point fixé de \mathbb{R} .

H_7 — K est un noyau d'ordre i au sens de Gasser, c'est-à-dire,

$$\int_{\mathbb{R}} t^j K(t) dt = 0, \forall j = 1, \dots, i-1 \text{ et } 0 < \left| \int_{\mathbb{R}} t^i K(t) dt \right| < +\infty.$$

H_8 — K à support compact.

H_9 — f est k fois continûment dérivable autour du point x .

Dans la suite on démontre la convergence presque complète de l'estimateur \hat{f}_n vers la fonction f , lorsque $n \rightarrow \infty$ et on calcule le taux de cette convergence, en utilisant une inégalité exponentielle de type Bernstein donnée par le lemme suivant

Lemme 2 *Ferraty et Vieu [12]*

Soient X_1, X_2, \dots, X_n des variables aléatoires réelles centrées, indépendantes et de même loi, telles qu'il existe deux réels positifs M et σ^2 vérifiant :

$$|X_1| \leq M \text{ et } \mathbf{E}(X_1^2) \leq \sigma^2.$$

Alors, pour tout $0 < \varepsilon < \frac{\sigma^2}{M}$ on a

$$\mathbf{P} \left(\left| \sum_{i=1}^n X_i \right| > \varepsilon n \right) \leq 2 \exp \left(-\frac{n\varepsilon^2}{4\sigma^2} \right).$$

Dans le corrolaire 1, Parzen (1962) [40] a prouvé que l'estimateur \hat{f}_n est asymptotiquement non biaisé pour les noyaux K vérifiant $\lim_{|y| \rightarrow \infty} |y| K(y) = 0$; le même résultat est donné dans le lemme 3 pour les noyaux à support compact.

Lemme 3 *Ferraty et Vieu [12]*

Sous les hypothèses H_1, H_2, H_6, H_8 on a

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\hat{f}_n(x) \right] = f(x).$$

Preuve 26 *On a par définition*

$$\mathbf{E} \left[\hat{f}_n(x) \right] = \frac{1}{h_n} \int_{\mathbb{R}} K \left(\frac{x-u}{h_n} \right) f(u) du = \int_{\mathbb{R}} K(z) f(x-zh_n) dz.$$

La fonction f est supposée continue sur \mathbb{R} , donc elle est continue sur le support compact du noyau K , et par conséquent f est uniformément continue.

Alors on a

$$\lim_{n \rightarrow \infty} f(x-zh_n) = f(x).$$

Or

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\hat{f}_n(x) \right] = \int_{\mathbb{R}} K(z) f(x) dz = f(x).$$

Lemme 4 *Ferraty et Vieu [12]*

Sous les hypothèses H_1, H_2, H_5, H_6 et H_8 , on a

$$\hat{f}_n(x) - \mathbf{E} \left[\hat{f}_n(x) \right] = O_{p.co.} \left(\sqrt{\frac{\log n}{nh_n}} \right).$$

Preuve 27 *Pour démontrer ce résultat on utilise le lemme 2. On a*

$$\begin{aligned} \hat{f}_n(x) - \mathbf{E} \left[\hat{f}_n(x) \right] &= \frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x-X_i}{h_n} \right) - \mathbf{E} \left[\frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x-X_i}{h_n} \right) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{h_n} K \left(\frac{x-X_i}{h_n} \right) - \frac{1}{h_n} \mathbf{E} \left(K \left(\frac{x-X_i}{h_n} \right) \right) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \Delta_i. \end{aligned}$$

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES
D'UNE LOI DE PROBABILITÉ

Les variables aléatoires $(\Delta_i)_{i=1,\dots,n}$ sont i.i.d. et centrées. Sous H_1 , on pose

K est borné par le nombre réel positif M , alors

$$\left| \frac{1}{h_n} K \left(\frac{x - X_i}{h_n} \right) - \frac{1}{h_n} \mathbf{E} \left(K \left(\frac{x - X_i}{h_n} \right) \right) \right| \leq \frac{2M}{h_n},$$

autrement dit pour tout $i = 1, \dots, n$

$$|\Delta_i| \leq \frac{C}{h_n}.$$

tel que $C = 2M$.

D'autre part

$$\begin{aligned} \mathbf{E} (\Delta_i^2) &= \text{var} \left(\frac{1}{h_n} K \left(\frac{x - X}{h_n} \right) \right) \\ &\leq \frac{1}{h_n^2} \mathbf{E} \left(K^2 \left(\frac{x - X}{h_n} \right) \right) \\ &\leq \frac{1}{h_n^2} \int_{\mathbb{R}} K^2 \left(\frac{x - u}{h_n} \right) f(u) du \end{aligned}$$

En utilisant le changement de variable $z = \frac{x-u}{h_n}$; on obtient

$$\frac{1}{h_n} \int_{\mathbb{R}} K^2(z) f(x - zh_n) dz = \frac{1}{h_n} \int_S K^2(z) f(x - zh_n) dz,$$

où S est le support compact du noyau K .

L'ensemble S est borné, alors il existe $k_1 \geq 0$, tel que pour tout $z \in S$; $|y| \leq k_1$, et comme h_n est borné (car il converge vers 0), il existe $k_2 \geq 0$, tel que pour tout $n \in \mathbb{N}$; $h_n \leq k_2$, donc pour tout $z \in S$ on a

$$x - k_1 k_2 \leq x - zh_n \leq x + k_1 k_2.$$

On a par supposition la fonction f est continue, donc elle est bornée sur le compact $[x - k_1 k_2, x + k_1 k_2]$, c'est à dire qu'il existe un réel positif k_3 tel que pour tout $z \in S$ et pour tout $n \in \mathbb{N}$, on a $f(x - zh_n) \leq k_3$,

d'où

$$\begin{aligned} \mathbf{E}(\Delta_i^2) &\leq \frac{k_3}{h_n} \int_{\mathbb{S}} K^2(z) dz \\ &\leq \frac{k_3 M}{h_n} \int_{\mathbb{S}} K(z) dz = \frac{k_3 M}{h_n} = k a_n^2. \end{aligned}$$

On pose $U_n = \frac{\log n}{n} a_n^2 = \frac{\log n}{n h_n}$, et comme $\lim_{n \rightarrow \infty} \frac{\log n}{n h_n} = 0$ alors

$$\hat{f}_n(x) - \mathbf{E} \left[\hat{f}_n(x) \right] = O_{p.co.} \left(\sqrt{\frac{\log n}{n h_n}} \right).$$

Maintenant à l'aide des lemmes 3 et 4 on peut démontrer le théorème suivant qui confirme la convergence presque complète de l'estimateur à noyau \hat{f}_n vers la densité de probabilité inconnue f , avec un taux de convergence égal à $\sqrt{\frac{\log n}{n h_n}}$.

Théorème 17 *Ferraty et Vieu [12]*

Sous les hypothèses H_1, H_2, H_3, H_5 et H_8 l'estimateur \hat{f}_n converge presque complètement vers f , c'est à dire

$$\hat{f}_n(x) - f(x) = O_{p.co.} \left(\sqrt{\frac{\log n}{n h_n}} \right).$$

Preuve 28 *En se basant sur la décomposition suivante*

$$\hat{f}_n(x) - f(x) = \left[\hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right] - \left[f(x) + \mathbf{E} \left(\hat{f}_n(x) \right) \right].$$

on peut écrire

$$\begin{aligned} \mathbf{P} \left(\left| \hat{f}_n(x) - f(x) \right| > \varepsilon \right) &= \mathbf{P} \left(\left| \left[\hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right] - \left[f(x) + \mathbf{E} \left(\hat{f}_n(x) \right) \right] \right| > \varepsilon \right) \\ &\leq \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \frac{\varepsilon}{2} \right) \\ &\quad + \mathbf{P} \left(\left| f(x) + \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \frac{\varepsilon}{2} \right). \end{aligned}$$

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES
D'UNE LOI DE PROBABILITÉ

D'après le lemme 3, on a

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\hat{f}_n(x) \right] = f(x).$$

C'est à dire qu'il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, pour tout $\varepsilon > 0$,

on a

$$\mathbf{P} \left(\left| f(x) + \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \frac{\varepsilon}{2} \right) = 0,$$

et on obtient

$$\mathbf{P} \left(\left| \hat{f}_n(x) - f(x) \right| > \varepsilon \right) \leq \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \frac{\varepsilon}{2} \right).$$

En suite, d'après le lemme 4, on a

$$\hat{f}_n(x) - \mathbf{E} \left[\hat{f}_n(x) \right] = O_{p.co.} \left(\sqrt{\frac{\log n}{nh_n}} \right),$$

ce qui est équivalent à l'existence de $\varepsilon_0 = 1 > 0$, tel que

$$\sum_{n \in \mathbb{N}^*} \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \sqrt{\frac{\log n}{nh_n}} \right) < \infty.$$

Donc pour $\varepsilon = \sqrt{\frac{\log n}{nh_n}}$, on trouve

$$\begin{aligned} \mathbf{P} \left(\left| \hat{f}_n(x) - f(x) \right| > \sqrt{\frac{\log n}{nh_n}} \right) &\leq \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \frac{1}{2} \sqrt{\frac{\log n}{nh_n}} \right) \\ &\leq \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \sqrt{\frac{\log n}{nh_n}} \right). \end{aligned}$$

D'où

$$\sum_{n \in \mathbb{N}^*} \mathbf{P} \left(\left| \hat{f}_n(x) - f(x) \right| > \sqrt{\frac{\log n}{nh_n}} \right) \leq \sum_{n \in \mathbb{N}^*} \mathbf{P} \left(\left| \hat{f}_n(x) - \mathbf{E} \left(\hat{f}_n(x) \right) \right| > \sqrt{\frac{\log n}{nh_n}} \right) < \infty.$$

La convergence presque complète de l'estimateur \hat{f}_n peut être obtenue en changeant quelques hypothèses, comme le montre le lemme et le théorème qui suivent.

Lemme 5 *Ferraty et Vieu [12]*

Sous les hypothèses H_2, H_5, H_7 et H_9 on a

$$\hat{f}_n(x) - \mathbb{E} \left(\hat{f}_n(x) \right) = O(h_n^k).$$

Théorème 18 *Ferraty et Vieu [12]*

Sous les hypothèses H_1, H_2, H_5, H_8 et H_9 on a

$$\hat{f}_n(x) - \mathbb{E} \left(\hat{f}_n(x) \right) = O(h_n^k) + O \left(\sqrt{\frac{\log n}{nh_n}} \right).$$

2.3 Estimateur à noyau de la fonction de répartition

L'estimation de la fonction de répartition est tout aussi importante que l'estimation de la fonction de densité de probabilité. On a déjà dit que la densité a un avantage sur le plan visuel, tandis que la fonction de répartition contient l'information sur la loi, mais de manière moins visible. Néanmoins, le comportement des estimateurs fonctionnels, la vitesse de convergence et la normalité asymptotique, sont facilement expliquées à l'aide du comportement local de la fonction de répartition, par l'estimation de cette dernière qu'on passe à l'estimation des probabilités des ensembles, par exemple la probabilité qu'une variable aléatoire se cantonne dans un intervalle donné.

Soit X_1, \dots, X_n un échantillon de variables aléatoires identiquement distribuées et indépendantes, définies sur l'espace de probabilité (Ω, \mathcal{A}, P) de même loi que la variable aléatoire X , de fonction de répartition F définie, pour tout $x \in \mathbb{R}$, par

$$F(x) = P(X \leq x).$$

Un estimateur traditionnel de la fonction F , à partir de l'échantillon X_1, \dots, X_n est la fonction de répartition empirique, notée F_n , définie au point x , comme suit

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(X_i \leq x)},$$

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

avec

$$\mathbf{1}_{(X_i \leq x)} = \begin{cases} 0 & \text{si } X_i > x \\ 1 & \text{si } X_i \leq x \end{cases} .$$

La fonction de distribution empirique définie précédemment, n'est pas régulière car à chaque point $X_1 = x_1, \dots, X_n = x_n$, F_n fait un saut égal à $\frac{1}{n}$.

La généralisation ou la version régulière de cet estimateur peut être obtenue à partir de la définition suivante de la fonction de répartition

$$F(x) = \int_{-\infty}^x f(t) dt,$$

où l'estimateur à noyau \hat{f}_n de la densité f peut être intégré pour obtenir un estimateur à noyau de F , que l'on note \hat{F}_n . Cet estimateur a été proposé par Nadaraya [37] en 1964 et est défini par

$$\begin{aligned} \hat{F}_n(x) &= \int_{-\infty}^x \hat{f}_n(t) dt \\ &= \frac{1}{n} \sum_{i=1}^n H\left(\frac{x - X_i}{h_n}\right), \end{aligned}$$

où H est une fonction définie à partir du noyau K comme suit

$$H(x) = \int_{-\infty}^x K(t) dt.$$

H est appelé noyau intégré et h_n est le paramètre de lissage ou la fenêtre de l'estimateur.

2.3.1 Propriétés asymptotiques de l'estimateur à noyau de la fonction de répartition

Les propriétés asymptotiques de l'estimateur \hat{F}_n ont été étudiées par plusieurs auteurs et sous différentes conditions, citons par exemple, Nadaraya [37], en 1964 a prouvé que \hat{F}_n est un estimateur asymptotiquement non biaisé et que sa variance est égale à $\frac{F(x)[1-F(x)]}{n}$ (théorème 1, page 551). Sous la condition de continuité de f , Nadaraya [37] en 1964, Yamato [57] en 1973 et Winter [55] ont démontré que l'estimateur \hat{F}_n est uniformément

convergent vers F avec une probabilité égale à 1. En 1983, sans condition sur f , Singh et al [48] ont prouvés le même résultat. Winter [56] en 1979 a montré que \hat{F}_n vérifie la propriété de Chung-Smirnov suivante

$$\lim_{n \rightarrow \infty} \sup \left\{ \left(\frac{2n}{\log \log n} \right)^{\frac{1}{2}} \sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| \right\} \leq 1,$$

avec une probabilité égale à 1.

La normalité asymptotique de \hat{F}_n a été établi par Watson et Leadbetter (1964) [54]. Falk (1983) [10] a montré que la performance asymptotique de l'estimateur à noyau \hat{F}_n est meilleur que celle de l'estimateur empirique F_n de F .

En 1981, Azzalini [2] a donné une expression asymptotique pour l'erreur moyenne quadratique MSE et a déterminé un paramètre de lissage asymptotiquement optimal pour obtenir un MSE plus petit que celui de l'estimateur empirique et aussi il a trouvé une expression asymptotique de MISE.

En 1988, Swanepoel [52] a trouvé, sous les hypothèses que K est une densité de probabilité bornée, symétrique autour de 0 à support fini, d'ordre 2 et que F est deux fois continûment dérivable avec f' borné et $var(f(X)^2) < \infty$, que l'erreur moyenne quadratique intégrée de \hat{F}_n est

$$\begin{aligned} MISE \left(\hat{F}_n \right) &= \int_{-\infty}^{+\infty} \mathbf{E} \left(\hat{F}_n(x) - F(x) \right)^2 dx \\ &= \frac{1}{6n} + \frac{2hC \int_{-\infty}^{+\infty} f(x)^2 dx}{n} + \frac{h^4 L_n^2 \int_{-\infty}^{+\infty} \left(f(x)' \right)^2 f(x) dx}{4} + o \left(\frac{h}{n} + h^4 \right), \end{aligned}$$

où $C = \int_{-\infty}^{+\infty} sK(x)H(x)dx$ et h est le paramètre de lissage. Il a prouvé, également, que le noyau uniforme est optimal et il a minimisé l'erreur moyenne quadratique intégrée afin d'obtenir un paramètre de lissage optimal, il a affirmé que l'estimateur à noyau \hat{F}_n est asymptotiquement plus efficace que l'estimateur empirique F_n .

Biais, variance et erreur moyenne quadratique de \hat{F}_n

Sous des conditions pas très fortes, en 1964, Nadaraya [37] a prouvé que \hat{F}_n admet, asymptotiquement, la même moyenne et variance que F_n . Pour

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

améliorer cette étude Azzalini [2] en 1981, a proposé d'autres hypothèses, sous les quelles il a obtenu le biais et la variance données ci dessous. Les hypothèses de Azzalini [2] sont :

H_a — la fonction de densité f est une fonction continue au voisinage de x , différentiable de dérivées carrées intégrables.

H_b — $h_n \rightarrow 0$ et $nh_n \rightarrow \infty$, quand $n \rightarrow +\infty$.

H_c — le noyau K est une densité continue, bornée et symétrique au tour de zéro. Donc K satisfait

$$\int_{-\infty}^{+\infty} K(x) dx = 1, \int_{-\infty}^{+\infty} xK(x) dx = 0 \text{ et } \int_{-\infty}^{+\infty} x^2K(x) dx < \infty.$$

Le biais est donné par

$$\text{Biais} \left(\hat{F}_n(x) \right) = \frac{h_n^2 f'(x) \int_{\mathbb{R}} x^2 K(x) dx}{2} + o(h_n^2).$$

La variance est donnée par

$$\text{var}(\hat{F}_n(x)) = \frac{F(x)(1-F(x))}{n} - \frac{2}{n} h_n f(x) \int_{\mathbb{R}} xK(x) H(x) dx + o\left(\frac{h_n}{n}\right).$$

Et L'erreur moyenne quadratique est

$$\begin{aligned} \text{MSE} \left(\hat{F}_n(x) \right) &= \frac{F(x)(1-F(x))}{n} - \frac{2}{n} h_n f(x) \int_{\mathbb{R}} xK(x) H(x) dx \\ &+ \frac{h_n^4 f'^2(x) \left(\int_{\mathbb{R}} x^2 K(x) dx \right)^2}{4} + o\left(h_n^4 + \frac{h_n}{n} \right). \end{aligned}$$

L'expression asymptotique du MSE est donné par

$$\begin{aligned} \text{AMSE} \left(\hat{F}_n(x) \right) &= \frac{F(x)(1-F(x))}{n} - \frac{2}{n} h_n f(x) \int_{\mathbb{R}} xK(x) H(x) dx \\ &+ \frac{h_n^4 f'^2(x) \left(\int_{\mathbb{R}} x^2 K(x) dx \right)^2}{4}. \end{aligned}$$

La valeur de h_n qui minimise $\text{AMSE} \left(\hat{F}_n(x) \right)$ est

$$\hat{h} = \left(\frac{2f(x) \int_{\mathbb{R}} xK(x) H(x) dx}{nf'^2(x) \left(\int_{\mathbb{R}} x^2 K(x) dx \right)^2} \right)^{\frac{1}{3}}.$$

Concernant la convergence presque complète de l'estimateur à noyau \hat{F}_n de la fonction de la répartition, on a la formule 3 de la référence Kitouni et al [28] (page 256) qui donne le taux de la convergence presque complète.

$$\sup \left| \hat{F}_n(x) - F(x) \right| = O_{p.co} \left(\sqrt{\frac{\log n}{n}} \right)$$

L'estimation des quantiles d'une population est d'un grand intérêt, en particulier, lorsque le statisticien ne veut pas présumer une forme paramétrique pour la distribution ou même de supposer que la distribution est symétrique.

2.4 Estimateur à noyau de la fonction de quantile

Les estimateurs non paramétrique populaire des quantiles sont les quantiles d'échantillon. Ces derniers ont de bonnes propriétés statistiques. Cependant, ils ont aussi des inconvénients, ils subissent une perte d'efficacité substantielle pour certaines distributions telles que la distribution normale, à cause de leur expression. Pour améliorer l'efficacité de ces estimateurs, un moyen évident est de former une moyenne pondérée de plusieurs statistiques d'ordre, en utilisant une fonction de pondération appropriée. De tels estimateurs sont appelés "L-estimateurs", combinaisons linéaire des statistiques d'ordres. Le problème devient alors de choisir la fonction de poids ou de pondération. Les estimateurs à noyau des quantiles forment une classe d'estimateurs "L", qui utilise les noyaux, en tant que fonction de poids.

Soit X_1, \dots, X_n un échantillon de variables aléatoires identiquement distribuées et indépendantes, de même loi que la variable aléatoire X de fonction de répartition F absolument continue et inconnue. Soit $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ les statistiques d'ordre correspondantes. On définit la fonction de quantile, notée Q , par l'inverse continu à droite de la fonction F , par

$$Q(p) = \inf\{x : F(x) \geq p\}, 0 < p < 1.$$

L'estimateur empirique de $Q(p)$ est le quantile de l'échantillon défini par l'inverse de la fonction de distribution empirique F_n , comme suit

$$Q_n(p) = \inf\{x : F_n(x) \geq p\} = X_{([np]+1)},$$

où $[np]$ est la partie entière de np .

L'inconvénient majeur des quantiles d'échantillon est qu'ils connaissent un manque partielle d'efficacité causé par la variabilité des statistiques

2. ESTIMATION À NOYAU DES CARACTÉRISTIQUES FONCTIONNELLES D'UNE LOI DE PROBABILITÉ

d'ordre individuelles. Une classe populaire des L-estimateurs est les estimateurs à noyau de la fonction de quantile définis par

$$\hat{Q}_n(p) = \frac{1}{h_n} \sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} K\left(\frac{x-p}{h_n}\right) dx,$$

où K est une fonction de densité symétrique autour de 0 et $h_n \rightarrow 0$ lorsque $n \rightarrow \infty$. Cette forme d'estimateur à noyau de la fonction Q est attribuée à Parzen [41] en 1979, page 113.

2.4.1 Propriétés asymptotiques de l'estimateur à noyau de la fonction de quantile

Concernant les propriétés asymptotique de \hat{Q}_n , en 1984, Falk [11] a montré que la performance asymptotique de \hat{Q}_n est meilleur que celle du quantile empirique de l'échantillon, en étudiant le rapport asymptotique de l'efficacité du quantile de l'échantillon par rapport à \hat{Q}_n . La normalité asymptotique et la consistance moyenne quadratique de cette estimateur est établie par Yang[58], en 1985. En 1986, Padgett [39] a donné la généralisation de la définition de \hat{Q}_n pour les données censurées à droite.

Le théorème suivant donne une expression de l'erreur quadratique moyenne asymptotique de \hat{Q}_n basée sur l' expression de la variance calculée par Falk [11].

Théorème 19 *Sheather et Marron [46]*

Supposons que $Q_X^{(2)}$ est continue au voisinage de p et K est une densité symétrique par rapport à zéro et admet un support compact. L'erreur quadratique moyenne de $\hat{Q}_n(p)$ est donnée par

1. Lorsque $p \neq \frac{1}{2}$ on a

$$\begin{aligned} MSE\{\hat{Q}_n(p)\} &= \frac{p(1-p)}{n} [Q'(p)]^2 - \frac{2h_n}{n} [Q'(p)]^2 \int_{-\infty}^{+\infty} yK(y)H(y)dy \\ &+ \frac{h_n^4}{4} [Q^{(2)}(p)]^2 \left[\int_{-\infty}^{+\infty} y^2 K(y) dy \right]^2 + o\left(\frac{h_n}{n} + h_n^4\right). \end{aligned}$$

2. Lorsque F est symétrique et $p = \frac{1}{2}$

$$MSE\{\hat{Q}_n(0.5)\} = \frac{1}{n} [Q'(0.5)]^2 \left\{ 0.25 + h_n \int_{-\infty}^{+\infty} yK(y)H(y)dy + \frac{1}{nh_n} \int_{-\infty}^{+\infty} K^2(y)dy \right\} + o\left(\frac{h_n}{n}\right) + o(n^{-2}h_n^{-2}).$$

Preuve 29 Pour la preuve on se référer à Sheather et Marron [46].

Corollaire 2 Sheather et Marron [46]

Supposons que les conditions du théorème (19) sont vérifiées. Alors,

1. quand F est symétrique et pour tout p , sauf $p = 0.5$, la fenêtre optimale est donnée par

$$\hat{h}_{opt} = \alpha(K)\beta(Q)n^{-\frac{1}{3}},$$

$$\text{où } \alpha(K) = \left[\frac{2 \int_{-\infty}^{+\infty} yK(y)H(y)dy}{\left\{ \int_{-\infty}^{+\infty} y^2K(y)dy \right\}^2} \right]^{-\frac{1}{3}} \quad \text{et } \beta(Q) = \left[\frac{Q'(p)}{Q^{(2)}(p)} \right]^{\frac{2}{3}}.$$

2. Quand F est symétrique, pour tout $p = 0.5$ et $h_n = O(n^{-\frac{1}{2}})$, on a

$$MSE\{\hat{Q}_n(0.5)\} = \frac{0.25}{n} [Q'(0.5)]^2 + O(n^{-\frac{3}{2}}).$$

3 Convergence complète en moment du second-ordre avec taux

Ce chapitre est consacré à la présentation d'un nouveau mode de convergence des suites de variables aléatoires et son application sur les estimateurs à noyaux des caractéristiques fonctionnelles d'une loi de probabilité inconnue. Ce mode est nommé convergence complète en moment du second-ordre avec taux. Il implique la convergence presque complète, presque sûre et la convergence en moyenne quadratique. Pour les estimateurs à noyaux des fonctions de densité, répartition et de quantile, la convergence en moment du second-ordre avec taux (c.m.s.), nommé en anglais "complete second-order moment convergence with a rate", donne de meilleurs taux de convergence, que ceux obtenus grâce à la convergence presque complète.

3.1 Convergence complète en moment du second-ordre avec taux

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité sur lequel est définies les variables aléatoires réelles (v.a.r.) X et $(X_n)_{n \in \mathbb{N}^*}$.

Définition 13 *Madi et Laroussi [35]*

La suite $(X_n)_{n \in \mathbb{N}^*}$ est dite complètement convergente en moment du second-

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

ordre vers X , avec un taux de convergence $\frac{1}{\sqrt{U_n}}$, si

$$\sum_{n \in \mathbb{N}^*} U_n \mathbf{E} (X_n - X)^2 < \infty,$$

où $(U_n)_{n \in \mathbb{N}^*}$ est une suite de nombres réels positifs. On note ce type de convergence par

$$X_n - X = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right).$$

Le théorème suivant montre que si la suite $(X_n)_{n \in \mathbb{N}^*}$ converge complètement en moment du second-ordre vers X , avec un taux de convergence $\frac{1}{\sqrt{U_n}}$, alors $(X_n)_{n \in \mathbb{N}^*}$ converge presque complètement vers X avec le même taux $\frac{1}{\sqrt{U_n}}$.

Théorème 20 *Madi et Laroussi [35]*

Si la suite $(X_n)_{n \in \mathbb{N}^}$ vérifie*

$$X_n - X = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right),$$

pour $\lim_{n \rightarrow \infty} U_n = +\infty$, alors

$$X_n - X = O_{p.co} \left(\frac{1}{\sqrt{U_n}} \right).$$

Et considérons la suite $(V_n)_{n \in \mathbb{N}^} = (n^\beta U_n)_{n \in \mathbb{N}^*}$ pour $\beta > -1$, $\varepsilon > 0$, on obtient*

$$\sum_{n \in \mathbb{N}^*} n^\beta \mathbf{P} (|X_n - X| > \varepsilon) \leq \sum_{n \in \mathbb{N}^*} n^\beta U_n \mathbf{E} (X_n - X)^2 < \infty.$$

Preuve 30 *Supposons que $\sum_{n \in \mathbb{N}^*} U_n \mathbf{E} (X_n - X)^2 < \infty$, puis en utilisant l'inégalité de Markov, on obtient*

$$\mathbf{P} \left(|X_n - X| \geq \frac{1}{\sqrt{U_n}} \right) \leq U_n \mathbf{E} (X_n - X)^2,$$

3.1. Convergence complète en moment du second-ordre avec taux

donc

$$\sum_{n \in \mathbb{N}^*} \mathbf{P} \left(|X_n - X| \geq \frac{1}{\sqrt{U_n}} \right) \leq \sum_{n \in \mathbb{N}^*} U_n \mathbf{E} (X_n - X)^2 < \infty.$$

Comme $\frac{1}{\sqrt{U_n}}$ tend vers 0 lorsque n tend vers $+\infty$, et qui est équivalent à $\forall \varepsilon > 0$, il existe $n_0 \in \mathbb{N}^*$ tel que pour tous $n \geq n_0$ on a $\frac{1}{\sqrt{U_n}} < \varepsilon$. Alors

$$\mathbf{P}(|X_n - X| > \varepsilon) \leq \mathbf{P}(|X_n - X| \geq \frac{1}{\sqrt{U_n}}),$$

aussi

$$\sum_{n \in \mathbb{N}^*} n^\beta \mathbf{P}(|X_n - X| > \varepsilon) \leq \sum_{n \in \mathbb{N}^*} n^\beta U_n \mathbf{E} (X_n - X)^2 < \infty.$$

Remarque 6 La convergence en moment du second-ordre avec taux implique la convergence presque complète mais la réciproque est fautive. .

Remarque 7 1. La convergence en moyenne quadratique d'une suite $(X_n)_{n \in \mathbb{N}^*}$ de v.a.r. est un résultat immédiat de sa convergence en moment du second-ordre avec taux, c'est-à-dire si

$$\sum_{n \in \mathbb{N}^*} U_n E (X_n - X)^2 < \infty,$$

alors

$$\lim_{n \rightarrow +\infty} E (X_n - X)^2 = 0.$$

2. la convergence complète en moment du second-ordre avec taux implique la convergence complète en moment d'ordre deux, notée $s - L^2$.

En effet,

$$\sum_{n \in \mathbb{N}^*} E (X_n - X)^2 < \sum_{n \in \mathbb{N}^*} U_n E (X_n - X)^2 < \infty,$$

puisque $(U_n)_{n \in \mathbb{N}^*}$ est une suite de nombres réels positifs.

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

On a décidé de présenter quelques propriétés ou règles de calcul pour rendre ce mode de convergence applicable d'une manière simple aux différents opérations mathématique sur les suites, comme fut le cas pour les modes de convergence des suites de variables aléatoires, mentionnés dans le premier chapitre. Soient $(X_n)_{n \in \mathbb{N}^*}$ et $(Y_n)_{n \in \mathbb{N}^*}$ deux suites de v.a.r. définies sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ et $(U_n)_{n \in \mathbb{N}^*}$ une suite de nombres réels positifs.

Proposition 15 *Madi et Laroussi [35]*

Supposons que $\lim_{n \rightarrow \infty} U_n = +\infty$, $X_n - l_X = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$ et $Y_n - l_Y = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$, où l_X et l_Y sont deux nombres réels. Alors on a

1. $(X_n + Y_n) - (l_X + l_Y) = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$.
2. $(X_n \cdot Y_n) - (l_X \cdot l_Y) = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$.
3. $\frac{1}{X_n} - \frac{1}{l_X} = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$, avec $l_X \neq 0$.

Preuve 31 1. Cette propriété est immédiate à partir de l'inégalité suivante : pour X et Y deux v.a.r. on a

$$\mathbf{E}(X + Y)^2 \leq 2^2 (\mathbf{E}(X)^2 + \mathbf{E}(Y)^2).$$

En effet,

$$\begin{aligned} \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}((X_n + Y_n) - (l_X + l_Y))^2 &\leq 2^2 \left(\sum_{n \in \mathbb{N}^*} U_n \mathbf{E}(X_n - l_X)^2 \right. \\ &\quad \left. + \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}(Y_n - l_Y)^2 \right). \end{aligned}$$

et comme $X_n - l_X = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$ et $Y_n - l_Y = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$, on obtient

$$\sum_{n \in \mathbb{N}^*} U_n \mathbf{E}((X_n + Y_n) - (l_X + l_Y))^2 < \infty.$$

3.1. Convergence complète en moment du second-ordre avec taux

2. Pour démontrer la deuxième propriété on utilise la définition de l'espérance mathématique d'une v.a.r. positive X donnée par

$$\mathbf{E}(X) = \int_0^{+\infty} \mathbf{P}(X > t) dt.$$

Appliqué à la v.a.r. positive $((X_n \cdot Y_n) - (l_X \cdot l_Y))^2$, on obtient

$$\begin{aligned} \mathbf{E}(X_n Y_n - l_X l_Y)^2 &= \int_0^{+\infty} \mathbf{P}(|X_n Y_n - l_X l_Y| > \sqrt{t}) dt \\ &= \int_0^{+\infty} \mathbf{P}(|(X_n - l_X)(Y_n - l_Y) \\ &\quad + l_X(Y_n - l_Y) + l_Y(X_n - l_X)| > \sqrt{t}) dt. \end{aligned}$$

En utilisant les inégalités $\mathbf{P}(|X + Y| > t) \leq \mathbf{P}(|X| > \frac{t}{2}) + \mathbf{P}(|Y| > \frac{t}{2})$ et $\mathbf{P}(|X \cdot Y| > t) \leq \mathbf{P}(|X| > \sqrt{t}) + \mathbf{P}(|Y| > \sqrt{t})$ pour toutes v.a.r. X , Y et $t \in \mathbb{R}^+$, on trouve

$$\begin{aligned} \mathbf{E}(X_n Y_n - l_X l_Y)^2 &\leq \mathbf{E}|X_n - l_X|^2 + \mathbf{E}|Y_n - l_Y|^2 + l_X^2 \mathbf{E}|Y_n - l_Y|^2 \\ &\quad + l_Y^2 \mathbf{E}|X_n - l_X|^2, \end{aligned}$$

donc

$$\begin{aligned} \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}((X_n \cdot Y_n) - (l_X \cdot l_Y))^2 &\leq \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}|X_n - l_X|^2 + \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}|Y_n - l_Y|^2 \\ &\quad + l_X^2 \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}|Y_n - l_Y|^2 + l_Y^2 \sum_{n \in \mathbb{N}^*} U_n \mathbf{E}|X_n - l_X|^2. \end{aligned}$$

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

3. Pour $l_X \neq 0$

$$\begin{aligned}
 \mathbf{E} \left(\frac{1}{X_n} - \frac{1}{l_X} \right)^2 &= \int_0^{+\infty} \mathbf{P} \left(\left| \frac{1}{X_n} - \frac{1}{l_X} \right|^2 > t \right) dt \\
 &= \int_0^{+\infty} \mathbf{P} (|X_n - l_X|^2 > t (X_n \cdot l_X)^2) dt \\
 &\leq \int_0^{+\infty} \mathbf{P} (|X_n - l_X|^2 > t) dt \\
 &\leq \mathbf{E} |X_n - l_X|^2,
 \end{aligned}$$

d'où

$$\sum_{n \in \mathbb{N}^*} U_n \mathbf{E} \left(\frac{1}{X_n} - \frac{1}{l_X} \right)^2 \leq \sum_{n \in \mathbb{N}^*} U_n \mathbf{E} |X_n - l_X|^2 < \infty.$$

Le corollaire suivant donne deux propriétés qui sont conséquences directes des règles de calculs précédentes.

Corollaire 3 *Madi et Laroussi [35]*

Supposons que $\lim_{n \rightarrow \infty} U_n = +\infty$, $X_n = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$ et $Y_n - l_Y = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$,

où l_Y est un nombres réel. Alors on a

1. $(X_n \cdot Y_n) = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$.
2. $\frac{X_n}{Y_n} = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right)$, avec $l_Y \neq 0$.

Dans la suite, on démontre la convergence complète en moment du second ordre avec taux des estimateurs à noyau des fonctions de densité de probabilité, de répartition et de quantile. Également, on établi de nouveaux et meilleurs taux de la convergence presque complète pour ces estimateurs, sous mions de conditions sur la fonction du noyau, à l'aide de l'expression de l'erreur moyenne quadratique MSE.

3.2 Convergence c.m.s. avec taux de l'estimateur à noyau de

la fonction de densité

Soit X_1, \dots, X_n un échantillon de variables aléatoires identiquement distribuées et indépendantes (i.i.d.) de même loi que la variable aléatoire X , de fonction f de densité de probabilité supposée inconnue. Rappelons que l'estimateur à noyau de la densité est donné par l'expression suivante

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right),$$

où K est la fonction du noyau et (h_n) est le paramètre de lissage.

Pour démontrer la convergence en moment du second-ordre avec taux de \hat{f}_n , on suppose que le noyau K et la densité de probabilité f vérifient les hypothèses habituelles qui suit

1. $H_1 : \forall x \in \mathbb{R}, K(x) = K(-x)$.
2. $H_2 : \sup_{x \in \mathbb{R}} |K(x)| \leq M < \infty$, pour $M \in \mathbb{R}^+$.
3. $H_3 : \int_{-\infty}^{+\infty} x^2 K(x) dx < \infty$.
4. $H_4 : f \in C^2$ et $f^{(2)}$ est bornée.

Le théorème (21) affirme la convergence complète en moment du second-ordre de l'estimateur à noyau \hat{f}_n de la fonction de densité de probabilité.

Théorème 21 *Madi et Laroussi [35]*

Sous les hypothèses $H_1 - H_4$ et supposons que $\sum_{n \geq 1} U_n h_n^4$ et $\sum_{n \geq 1} \frac{U_n}{nh_n^2}$ sont convergentes, on a

$$\hat{f}_n(x) - f(x) = O_{c.m.s.} \left(\frac{1}{\sqrt{U_n}} \right). \quad (3.1)$$

Preuve 32 *Pour démontrer (3.1), on montre que*

$$\sum_{n \geq 1} U_n \mathbf{E}(\hat{f}_n(x) - f(x))^2 < \infty,$$

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

en utilisant l'expression de l'erreur moyenne quadratique MSE de \hat{f}_n , donnée par

$$MSE(\hat{f}_n(x)) = \mathbf{E}(\hat{f}_n(x) - f(x))^2,$$

qui peut s'écrire sous la forme

$$MSE(\hat{f}_n(x)) = \text{Biais}(\hat{f}_n(x))^2 + \text{Var}(\hat{f}_n(x)).$$

Ainsi pour prouver (3.1), il suffit de montrer

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] < \infty, \quad (3.2)$$

et

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{f}_n(x))^2 < \infty. \quad (3.3)$$

D'abord on démontre l'inégalité (3.2), on a par définition

$$\begin{aligned} \text{Var}[\hat{f}_n(x)] &= \frac{1}{n^2 h_n^2} \text{Var} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right) \\ &= \frac{1}{n h_n^2} \left\{ \mathbf{E} \left[K^2 \left(\frac{x - X}{h_n} \right) \right] - \mathbf{E}^2 \left[K \left(\frac{x - X}{h_n} \right) \right] \right\}, \end{aligned}$$

car les v.a.r. X_1, \dots, X_n sont i.i.d.. De plus la quantité $\mathbf{E}^2 \left[K \left(\frac{x - X}{h_n} \right) \right]$ est positive, ce qui nous permet d'écrire

$$\text{Var}[\hat{f}_n(x)] \leq \frac{1}{n h_n^2} \mathbf{E} \left[K^2 \left(\frac{x - X}{h_n} \right) \right].$$

En utilisant l'hypothèse H_2 , on obtient

$$\text{Var}[\hat{f}_n(x)] \leq \frac{M^2}{n h_n^2}.$$

Par conséquent

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] \leq \sum_{n \geq 1} \frac{U_n M^2}{n h_n^2}.$$

3.2. Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de densité

Sous la supposition de la convergence de la série $\sum_{n \geq 1} \frac{U_n}{nh_n^2}$, on obtient

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] < \infty.$$

D'autre part pour l'inégalité (3.3), on a

$$\begin{aligned} E[\hat{f}_n(x)] &= E \left[\frac{1}{nh_n} \sum_{i=1}^n K \left(\frac{x - X_i}{h_n} \right) \right] \\ &= \frac{1}{h_n} E \left[K \left(\frac{x - X}{h_n} \right) \right] \\ &= \frac{1}{h_n} \int_{\mathbb{R}} K \left(\frac{x - u}{h_n} \right) f(u) du. \end{aligned}$$

En utilisant le changement de variable $z = \frac{x-u}{h_n}$ on obtient $u = x - zh_n$, ce qui implique $du = -h_n dz$, l'espérance de \hat{f}_n devient

$$E[\hat{f}_n(x)] = \int_{\mathbb{R}} K(z) f(x - zh_n) dz.$$

Du fait que f vérifie l'hypothèse H_4 , on peut la développer en série de Taylor d'ordre 3, au voisinage du point x , pour obtenir l'égalité suivante

$$\begin{aligned} E[\hat{f}_n(x)] &= f(x) \int_{\mathbb{R}} K(z) dz - h_n f'(x) \int_{\mathbb{R}} z K(z) dz \\ &+ \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz - \frac{h_n^3}{6} f^{(3)}(\theta) \int_{\mathbb{R}} z^3 K(z) dz, \end{aligned}$$

où θ est un nombre réel entre x et $x - zh_n$.

Comme K est une densité de probabilité, donc $\int_{\mathbb{R}} K(z) dz = 1$, de plus sous les hypothèses H_1 les intégrales $\int_{\mathbb{R}} z K(z) dz$ et $\int_{\mathbb{R}} z^3 K(z) dz$ sont nulles, par conséquent

$$E[\hat{f}_n(x)] = f(x) + \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz.$$

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

Ensuite, $f^{(2)}$ est bornée, c'est-à-dire qu'il existe M_1 un nombre réel positif

tel que $\forall x \in \mathbb{R}, |f^{(2)}(x)| \leq M_1$, ce qui donne

$$E[\hat{f}_n(x)] - f(x) \leq \frac{M_1 h_n^2}{2} \int_{\mathbb{R}} z^2 K(z) dz.$$

Par conséquent

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{f}_n(x))^2 \leq \sum_{n \geq 1} \frac{U_n h_n^4}{4} \left[M_1 \int_{\mathbb{R}} z^2 K(z) dz \right]^2.$$

Puisque le noyau K vérifie H_3 , alors

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{f}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4,$$

où C est une constante réelle positive. Comme la série $\sum_{n \geq 1} U_n h_n^4$ est convergente (par supposition), autrement dit

$$\sum_{n \geq 1} U_n h_n^4 < \infty,$$

alors

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{f}_n(x))^2 < \infty.$$

D'où la convergence complète en moment du second-ordre avec taux de l'estimateur à noyau \hat{f}_n de la fonction de densité.

Remarque 8 Si K est un noyau symétrique à support compact, alors les hypothèses H_2 et H_3 sont satisfaites et on obtient la convergence complète en moment du second-ordre avec taux de \hat{f}_n sous l'hypothèse H_4 seulement.

Maintenant, on donne un exemple dans lequel on montre que les choix des suites $(U_n)_{n \in \mathbb{N}^*}$ et $(h_n)_{n \in \mathbb{N}^*}$ ne sont pas arbitraires, car ses expressions doivent être sélectionnées de sorte que la convergence des séries soit assurée.

3.2. Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de densité

Exemple 1 Choisissons pour tout $n \in \mathbb{N}^*$, $h_n = \frac{(\log n)^4}{n^4}$ et $U_n = \frac{n^2 h_n^4}{\log n}$, on obtient

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] \leq \sum_{n \geq 1} \frac{M^2}{n^3 (\log n)^{-3}},$$

La série de Bertrand à droit est convergente. D'autre part

$$\sum_{n \geq 1} U_n (\text{Biais}[\hat{f}_n(x)])^2 \leq C \sum_{n \geq 1} \frac{1}{n^2 (\log n)^{-3}}.$$

La convergence est obtenue car la série de Bertrand à droit est convergente.

D'après la sous section (2.2.2), l'estimateur à noyaux de la fonction de densité converge presque complètement avec un taux égal à $\sqrt{\frac{\log n}{nh_n}}$. Dans la suite, on propose avec le mode de convergence c.m.s. un meilleur taux dans un sens probabiliste. On obtient une vitesse de convergence plus rapide.

Corollaire 4 Madi et Laroussi [35]

Sous H_1 - H_4 , $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$ et pour $h_n = O\left(\frac{\log^t n}{n^{\alpha+t-1}}\right)$ avec $t > 3$ et $\alpha > 0$

on a

$$\hat{f}_n - f = O_{p.co.} \left(\left(\frac{\log n}{n^{\alpha+1} h_n^4} \right)^{\frac{1}{2}} \right).$$

Preuve 33 Pour démontrer ce résultat, on cherche un U_n pour lequel $\frac{1}{\sqrt{U_n}} < \sqrt{\frac{\log n}{nh_n}}$ pour tout $n \geq 1$, c'est-à-dire qu'on cherche un U_n vérifiant l'inégalité $U_n > \frac{nh_n}{\log n}$.

Soient $U_n = n^\alpha h_n^3 \frac{nh_n}{\log n}$, avec $\alpha > 0$, et $h_n = O\left(\frac{\log^t n}{n^{\alpha+t-1}}\right)$ où $t > 3$. Sous l'hypothèse $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$, on doit vérifier d'abord que $\lim_{n \rightarrow +\infty} U_n = +\infty$ et $\lim_{n \rightarrow +\infty} h_n = 0$. En effet, on a par supposition $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$ et pour

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

$h_n = O\left(\frac{\log^t(n)}{n^{\alpha+t-1}}\right)$, on trouve

$$\begin{aligned} n^\alpha h_n^3 &< n^\alpha h_n \\ &< n^\alpha \frac{\log^t n}{n^{\alpha+t-1}} \\ &< n \left(\frac{\log n}{n}\right)^t, \end{aligned}$$

alors $n^\alpha h_n^3 = O(n)$, d'où notre hypothèse est toujours vérifiée. Pour la limite de h_n on a

$$\begin{aligned} \lim_{n \rightarrow +\infty} h_n &= \lim_{n \rightarrow +\infty} \frac{\log^t(n)}{n^{\alpha+t-1}} \\ &= \lim_{n \rightarrow +\infty} \frac{1}{n^{\alpha-1}} \left(\frac{\log n}{n}\right)^t \\ &= 0. \end{aligned}$$

On termine la vérification des hypothèses par prouver que $\lim_{n \rightarrow +\infty} U_n = +\infty$.

Comme $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$ on a $\left(\frac{\log n}{n^{\alpha+1} h_n^4}\right)^{\frac{1}{2}} > \frac{nh_n}{\log n}$ donc $\lim_{n \rightarrow +\infty} U_n = +\infty$.

Maintenant on montre la convergence c.m.s. de l'estimateur à noyau de la fonction de densité, on a

$$\begin{aligned} \sum_{n \geq 1} U_n h_n^4 &= \sum_{n \geq 1} \frac{n^{\alpha+1} h_n^8}{\log n} \\ &\leq \sum_{n \geq 1} \frac{n^{\alpha+1} h_n}{\log n} \\ &\leq \sum_{n \geq 1} \frac{1}{n^{t-2} (\log n)^{1-t}}. \end{aligned}$$

3.3. Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de répartition

Cette série de Bertrand converge pour $t > 3$.

$$\begin{aligned} \sum_{n \geq 1} \frac{U_n}{nh_n^2} &= \sum_{n \geq 1} \frac{n^\alpha h_n^2}{\log n} \\ &\leq \sum_{n \geq 1} \frac{n^\alpha h_n}{\log n} \\ &\leq \sum_{n \geq 1} \frac{1}{n^{t-1}(\log n)^{1-t}}. \end{aligned}$$

La série à droite converge pour $t > 2$. Combinons les deux conditions de convergence des deux séries, on obtient la convergence complète en moment du second-ordre pour $t > 3$.

3.3 Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de répartition

Soient X_1, X_2, \dots, X_n des variables aléatoires identiquement distribuées et indépendantes, copies de la variable aléatoire X , de densité de probabilité f et fonction de répartition F inconnues. L'estimateur à noyau de la fonction de répartition, noté \hat{F}_n , est donné par

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n H\left(\frac{x - X_i}{h_n}\right),$$

où la fonction H est une fonction de distribution cumulative définie par

$$H(x) = \int_{-\infty}^x K(t)dt.$$

Afin d'obtenir la convergence en moment du second-ordre, on suppose que les hypothèses suivantes sont vérifiées

1. $H_5 : \int_{-\infty}^{+\infty} K(x)H(x)dx < \infty$.
2. $H_6 : \int_{-\infty}^{+\infty} x^2 K(x)H(x)dx < \infty$.
3. $H_7 : \int_{-\infty}^{+\infty} xK(x)H(x)dx \geq 0$.

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

4. $H_8 : F \in C^2$ et f' est bornée.

Le théorème suivant montre que l'estimateur à noyau \hat{F}_n converge complètement en moment du second-ordre avec taux vers la fonction F .

Théorème 22 *Madi et Laroussi [35]*

Sous H_1, H_3, H_5-H_8 et supposons que $\sum_{n \geq 1} U_n h_n^4$ et $\sum_{n \geq 1} \frac{U_n}{nh_n^2}$ sont convergent, on a

$$\hat{F}_n(x) - F(x) = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right). \quad (3.4)$$

Preuve 34 *Pour prouver (3.4) on applique la même idée utilisée dans la démonstration du théorème (21). Premièrement pour le biais, on a*

$$\mathbf{E}[\hat{F}_n(x)] = \mathbf{E} \left[\frac{1}{n} \sum_{i=1}^n H \left(\frac{x - X_i}{h_n} \right) \right] = \int_{\mathbb{R}} H \left(\frac{x - u}{h_n} \right) f(u) du.$$

En utilisant le changement de variable $z = \frac{x-u}{h_n}$, puis en appliquant la formule d'intégration par partie, on obtient

$$\begin{aligned} \mathbf{E}[\hat{F}_n(x)] &= H \left(\frac{x - z}{h_n} \right) F(z) \Big|_{-\infty}^{+\infty} + \frac{1}{h_n} \int_{-\infty}^{+\infty} K \left(\frac{x - z}{h_n} \right) F(z) dz \\ &= \frac{1}{h_n} \int_{-\infty}^{+\infty} K \left(\frac{x - z}{h_n} \right) F(z) dz. \end{aligned}$$

Sous H_8 la fonction F peut être développer en série de Taylor d'ordre 2, au voisinage du point x , afin d'avoir

$$\begin{aligned} \mathbf{E} \left(\hat{F}_n(x) \right) &= \int_{-\infty}^{+\infty} K(y) \left[F(x) - h_n y f(x) + \frac{h_n^2}{2} y^2 f'(x) - \frac{h_n^3}{6} y^3 f''(x) \right] dy \\ &= F(x) \int_{-\infty}^{+\infty} K(y) dy - h_n f(x) \int_{-\infty}^{+\infty} y K(y) dy + \frac{h_n^2}{2} f'(x) \int_{-\infty}^{+\infty} y^2 K(y) dy \\ &\quad - \frac{h_n^3}{6} f''(x) \int_{-\infty}^{+\infty} y^3 K(y) dy. \end{aligned}$$

3.3. Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de répartition

Du fait que K est une densité de probabilité symétrique, il suit que l'expression du Biais devient

$$\text{Biais}(\hat{F}_n(x)) = \frac{h_n^2}{2} f'(x) \int_{\mathbb{R}} y^2 K(y) dy,$$

donc

$$\left(\text{Biais}(\hat{F}_n(x)) \right)^2 = \left(\frac{h_n^2}{2} f'(x) \int_{\mathbb{R}} y^2 K(y) dy \right)^2.$$

L'hypothèse H_3 signifie que $\exists C_1 \in \mathbb{R}^+$ tel que $\int_{\mathbb{R}} y^2 K(y) dy \leq C_1 < \infty$, donc sous H_3 et H_4 , on obtient

$$\left(\text{Biais}(\hat{F}_n(x)) \right)^2 \leq \frac{h_n^4}{4} f'^4(x) C_1^2,$$

par conséquent

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{F}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4 < \infty,$$

où C est une constante réelle positive.

Pour la variance

$$\begin{aligned} \text{Var}[\hat{F}_n(x)] &= \frac{1}{n^2} \text{Var} \sum_{i=1}^n H \left(\frac{x - X_i}{h_n} \right) \\ &\leq \frac{1}{n} \mathbf{E} \left[H^2 \left(\frac{x - X}{h_n} \right) \right]. \end{aligned}$$

Sous H_1 , H_8 , en choisissant le changement de variable $z = \frac{x-u}{h_n}$, en appliquant la formule d'intégration par partie et enfin en utilisant le développement

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

de Taylor-Lagrange d'ordre 2, au voisinage de x , de la fonction F , on trouve

$$\begin{aligned}
\mathbf{E} \left[H^2 \left(\frac{x - X}{h_n} \right) \right] &= \int_{\mathbb{R}} H^2 \left(\frac{x - u}{h_n} \right) f_X(u) du \\
&= \left[H^2 \left(\frac{x - u}{h_n} \right) F(u) \right]_{-\infty}^{+\infty} + \frac{1}{h_n} \int_{\mathbb{R}} 2K \left(\frac{x - u}{h_n} \right) H \left(\frac{x - u}{h_n} \right) F(u) du \\
&= 2 \int_{\mathbb{R}} K(z) H(z) F(x - z.h_n) dz \\
&= 2 \int_{-\infty}^{+\infty} K(z) H(z) \left[F(x) - h_n z f(x) + \frac{h_n^2}{2} z^2 f'(x) + -\frac{h_n^3}{6} y^3 f^2(\theta) \right] dz \\
&= 2 \int_{-\infty}^{+\infty} K(z) H(z) F(x) dz - 2h_n f(x) \int_{-\infty}^{+\infty} z k(z) H(z) dz \\
&\quad + 2 \frac{h_n^2}{2} f'(x) \int_{-\infty}^{+\infty} z^2 K(z) H(z) dz + 2 - \frac{h_n^3}{6} f^2(\theta) \int_{-\infty}^{+\infty} y^3 K(z) H(z) dz \\
&= 2F(x) \int_{-\infty}^{+\infty} K(z) H(z) dz - 2h_n f(x) \int_{-\infty}^{+\infty} z K(z) H(z) dz \\
&\quad + h_n^2 f'(x) \int_{-\infty}^{+\infty} z^2 K(z) H(z) dz + 2 - \frac{h_n^3}{6} f^2(\theta) \int_{-\infty}^{+\infty} y^3 K(z) H(z) dz.
\end{aligned}$$

Pour H_7 satisfaite, on a

$$\mathbf{E} \left[H^2 \left(\frac{x - X}{h_n} \right) \right] \leq 2F(x) \int_{-\infty}^{+\infty} K(z) H(z) dz + h_n^2 f'(x) \int_{-\infty}^{+\infty} z^2 K(z) H(z) dz.$$

Donc

$$\begin{aligned}
\text{Var}[\hat{F}_n(x)] &\leq \frac{4h_n^2 F(x) \int_{-\infty}^{+\infty} K(y) H(y) dy + h_n^4 f'(x) \int_{-\infty}^{+\infty} y^2 K(y) H(y) dy}{nh_n^2} \\
&\leq \frac{C_1 h_n^2 + C_2 h_n^4}{nh_n^2},
\end{aligned}$$

où C_1 et C_2 sont deux constantes réelles.

Comme la suite $(h_n)_{n \in \mathbb{N}^*}$ converge vers 0, alors pour tout $\varepsilon > 0$ il existe un rang n_0 tel que $h_n \leq \varepsilon$ pour tout $n \geq n_0$. D'où

$$\sum_{n \geq 1} U_n \text{Var}[\hat{F}_n(x)] \leq M \sum_{n \geq 1} \frac{U_n}{nh_n^2},$$

3.3. Convergence c.m.s. avec taux de l'estimateur à noyau de la fonction de répartition

où $M = C_1\varepsilon^2 + C_2\varepsilon^4$ et le résultat cherché est obtenu.

Un cas particulier est donné dans la remarque

Remarque 9 Quand K est un noyau du second-ordre, symétrique et à support compact $[-1, 1]$, les propriétés de la fonction H sont données par Baszczyńska (2016) [3] comme suit

1. $\int_{-1}^1 H^2(y)dy \leq \int_{-1}^1 H(y)dy = 1.$
2. $\int_{-1}^1 K(y)H(y)dy = \frac{1}{2}.$
3. $\int_{-1}^1 yK(y)H(y)dy = \frac{1}{2}[1 - \int_{-1}^1 H^2(y)dy].$

Donc les hypothèses H_5 et H_7 sont vérifiées. On peut obtenir le MSE de l'estimateur à noyau de la fonction de répartition \hat{F}_n sous les hypothèses de Azzalini (1981) [2] et quand le noyau satisfait les conditions précédentes, comme suit

$$\begin{aligned} \text{MSE}[\hat{F}_n(x)] &= \frac{h_n^4}{4} f'^2(x) \left(\int_{-1}^1 y^2 K(y) dy \right)^2 + \frac{1}{n} F(x)(1 - F(x)) \\ &\quad - \frac{2}{n} h_n f(x) \int_{-1}^1 y K(y) H(y) dy + o\left(h_n^4 + \frac{h_n}{n}\right). \end{aligned}$$

Le carré du biais est

$$\text{Biais}^2[\hat{F}_n(x)] = \frac{1}{4} h_n^4 f'^2(x) \left(\int_{-1}^1 y^2 K(y) dy \right)^2 + o(h_n^4),$$

donc sous H_8 on obtient

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{F}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4 < \infty,$$

où C est une constante réelle positive.

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

La variance est

$$\begin{aligned} \text{Var}[\hat{F}_n(x)] &= \frac{1}{n}F(x)(1 - F(x)) - \frac{2}{n}h_n f(x) \int_{-1}^1 yK(y)H(y)dy + o\left(\frac{h_n}{n}\right) \\ &\leq \frac{1}{nh_n^2} \left[h_n^2 F(x)(1 - F(x)) - 2h_n^3 f(x) \int_{-1}^1 yK(y)H(y)dy \right], \end{aligned}$$

comme $(h_n)_{n \in \mathbb{N}^*}$ converge vers 0 quand n tend vers ∞ , alors

$$\sum_{n \geq 1} U_n \text{Var}[\hat{F}_n(x)] \leq M \sum_{n \geq 1} \frac{U_n}{nh_n^2},$$

où M est une constante réelle positive. D'où

$$\sum_{n \geq 1} U_n \text{MSE}(\hat{F}_n(x)) < \infty.$$

Le corollaire suivant donne un nouveau taux de la convergence presque complète pour l'estimateur \hat{F}_n vers F .

Corollaire 5 *Madi et Laroussi [35]*

Sous H_1, H_3, H_5-H_8 , $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$, et pour $h_n = O\left(\frac{\log^t n}{n^{\alpha+t-1}}\right)$ avec $t > 3$

et $\alpha > 0$ on a

$$\hat{F}_n - F = O_{p.co.} \left(\sqrt{\frac{\log n}{n^{\alpha+1} h_n^3}} \right).$$

Preuve 35 *Pour démontrer ce résultat on utilise la même idée de la démonstration*

34. Soient $U_n = n^\alpha h_n^3 \frac{n}{\log n}$, pour $\alpha > 0$ et $h_n = \frac{\log^t n}{n^{\alpha+t-1}}$ où $t > 3$. Sous l'hypothèse $\lim_{n \rightarrow +\infty} n^\alpha h_n^3 = +\infty$ on a $U_n > \frac{n}{\log n}$, où $\sqrt{\frac{\log n}{n}}$ est le taux de la convergence presque complète de \hat{F}_n vers F . Afin d'obtenir la convergence complète en moment du second-ordre avec taux, on démontre la convergence

3.4. Convergence c.m.s. de l'estimateur à noyau de la fonction de quantile

des deux séries $\sum_{n \geq 1} U_n h_n^4$ et $\sum_{n \geq 1} \frac{U_n}{n h_n^2}$.

$$\begin{aligned} \sum_{n \geq 1} U_n h_n^4 &= \sum_{n \geq 1} \frac{n^{\alpha+1} h_n^7}{\log n} \\ &\leq \sum_{n \geq 1} \frac{n^{\alpha+1} h_n}{\log n} \\ &\leq \sum_{n \geq 1} \frac{1}{n^{t-2} (\log n)^{1-t}} \end{aligned}$$

Cette série de Bertrand converge pour $t > 3$.

$$\begin{aligned} \sum_{n \geq 1} \frac{U_n}{n h_n^2} &= \sum_{n \geq 1} \frac{n^\alpha h_n}{\log n} \\ &= \sum_{n \geq 1} \frac{1}{n^{t-1} (\log n)^{1-t}} \end{aligned}$$

La série à droite converge pour $t > 2$.

D'où la convergence est obtenue pour $t > 3$.

Exemple 2 Pour le paramètre de lissage $h_n = \frac{(\log n)^4}{n^4}$ et $U_n = \frac{n^2 h_n^3}{\log n}$, on obtient

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{F}_n(x))^2 \leq C \sum_{n \geq 1} \frac{1}{n^2 (\log n)^3},$$

et

$$\sum_{n \geq 1} U_n \text{Var}[\hat{F}_n(x)] \leq M \sum_{n \geq 1} \frac{1}{n^3 (\log n)^3}.$$

Les deux séries à droite sont convergentes.

3.4 Convergence c.m.s. de l'estimateur à noyau de la fonction de quantile

Soit X_1, X_2, \dots, X_n des v.a. indépendantes et identiquement distribuées, copies de la v.a. X de fonction de répartition absolument continue F . No-

3. CONVERGENCE COMPLÈTE EN MOMENT DU SECOND-ORDRE AVEC TAUX

tons par $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ la statistique d'ordre correspondante. L'estimateur à noyau \hat{Q}_n de la fonction de quantile Q , est

$$\hat{Q}_n(p) = \sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h_n} K\left(\frac{x-p}{h_n}\right) dx,$$

où K est la fonction du noyau, et h_n le paramètre de lissage satisfait $h_n \rightarrow 0$ lorsque $n \rightarrow \infty$.

Le résultat de la convergence complète en moment du second-ordre avec taux de \hat{Q}_n est basé sur l'expression du MSE de l'estimateur à noyau de la fonction de quantile, donnée par Sheather and Marron (1990)[46] (théorème 1, p 5), quand p est à l'intérieur de $(0, 1)$, sous les conditions que K est symétrique autour de 0 et a support compact et $Q^{(2)}$ est continue au voisinage de p .

Théorème 23 *Madi et Laroussi [35]*

Sous H_1, H_3, H_5-H_8 et supposons que $\sum_{n \geq 1} U_n h_n^4$ et $\sum_{n \geq 1} \frac{U_n}{n h_n^2}$ sont convergentes, alors

$$\hat{Q}_n(x) - Q(x) = O_{c.m.s} \left(\frac{1}{\sqrt{U_n}} \right). \quad (3.5)$$

Preuve 36 *Suivant Falk (1984) [11] et David (1981) [8], Sheather et Marron (1990) [46] ont donné les expressions du biais et la variance de \hat{Q}_n par*

$$\text{Biais}[\hat{Q}_n(x)] = \frac{1}{2} h_n^2 Q^{(2)}(p) \int_{-\infty}^{+\infty} x^2 K(x) dx + o(h_n^2) + o(n^{-1}),$$

et

$$\text{Var}[\hat{Q}_n(x)] = \frac{1}{n} p(1-p) Q'^2 - \frac{1}{n} h_n Q'^2 \int_{-\infty}^{+\infty} x K(x) H(x) dx + o\left(\frac{h_n}{n}\right),$$

Pour $Q^{(2)}$ borné et $\int_{-\infty}^{+\infty} x^2 K(x) dx < \infty$ on obtient

$$\text{Biais}(\hat{Q}_n(x))^2 \leq C h_n^4,$$

3.4. Convergence c.m.s. de l'estimateur à noyau de la fonction de quantile

où C est une constante réelle positive. Par conséquent

$$\sum_{n \geq 1} U_n \text{Biais}(\hat{Q}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4 < \infty.$$

Pour la variance, sous les hypothèses que $Q^{(2)}$ borné, $\int_{-\infty}^{+\infty} xK(x)H(x)dx < \infty$ et $h_n \rightarrow 0$ lorsque $n \rightarrow \infty$, on trouve

$$\text{Var}[\hat{Q}_n(x)] \leq \frac{M}{nh_n^2},$$

où C est une constante réelle positive, ce qui implique

$$\sum_{n \geq 1} U_n \text{Var}[\hat{Q}_n(x)] \leq M \sum_{n \geq 1} \frac{U_n}{nh_n^2} < \infty.$$

D'où

$$\hat{Q}_n(x) - Q(x) = O_{c.m.s.} \left(\frac{1}{\sqrt{U_n}} \right).$$

Finalement on a obtenu la convergence presque complète et la convergence complète du second-ordre de l'estimateur à noyau \hat{Q}_n de la fonction de quantile.

Remarque 10 Si la fonction du noyau vérifie les même hypothèses de la remarque (9), on obtient la convergence complète en moment du second-ordre avec taux de l'estimateur \hat{Q}_n vers la fonction Q , sous l'hypothèse H_{10} seulement.

4 Simulation et application

L'objectif de ce chapitre est de faire une comparaison des nouveaux taux obtenus par la convergence en moment du second ordre (c.m.s) et ceux obtenus par la convergence presque complète (p.co.). En premier lieu, on présente un nouveau type de noyau dit "Laplace". Puis le reste du chapitre est consacré à une étude de simulation et une application avec des résultats obtenus sur des données réelles recueillies du centre hospitalier universitaire Ben Badis de Constantine en 2018 dans un service de gynécologie.

4.1 Présentation du noyau de Laplace

Les noyaux de Laplace sont une collection de fonctions liées à l'estimation de la densité de probabilité. L'idée de l'utilisation de ce type de noyaux s'inspire du travaux de Chen (2000)[6], quand il a remplacé les noyaux fixes et symétriques, dans l'estimation à noyau standard de la densité, par des noyaux positifs de forme variée de type gamma. Les noyaux de Laplace sont développés par Khan et Akbar (2021)[29], définis comme suit

$$K_{\text{Laplace}(x,h^{1/2})}(u) = \frac{1}{2\sqrt{h}} \exp\left(-\frac{|u-x|}{\sqrt{h}}\right), u \in \mathbb{R},$$

où h est la fenêtre et $x \in \mathbb{R}$.

Pour calculer les valeurs de la densité estimées par le noyau de Laplace, on utilise la commande suivante

Laplace(x=NULL, y, k=NULL, h=NULL).

4. SIMULATION ET APPLICATION

Par exemple

```
y = rexp(100, 1)
xx = seq(min(y) + 0.05, max(y), length = 100)
h = 2
den = Laplace(x = xx, y = y, k = 200, h = h).
```

Les représentation graphiques en utilisant ces noyaux est obtenues à l'aide de commande " plot.Laplace". Par exemple

```
y = rexp(100, 1)
h = 0.79 * IQR(y) * length(y)(- 1/5)
xx = seq(min(y) + 0.05, max(y), length = 100)
den = Laplace(x = xx, y = y, k = 100, h = h)
plot(den, type = "l")
```

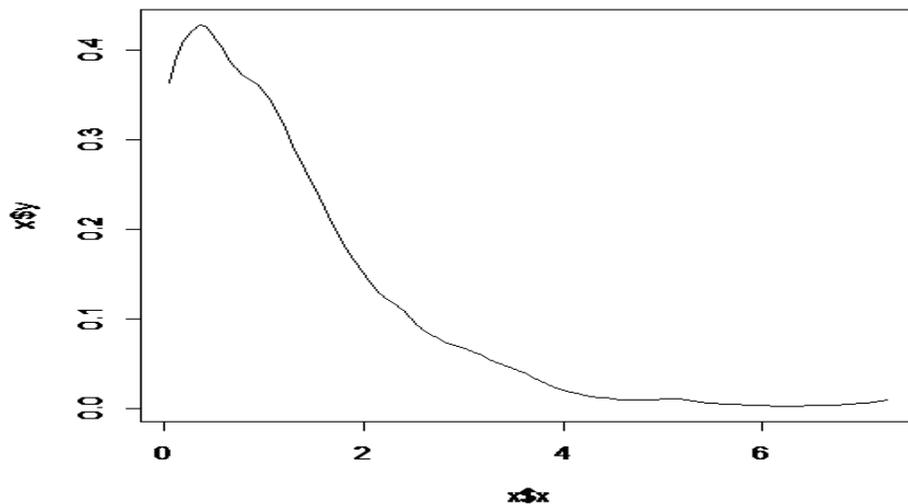


FIGURE 4.1 – Noyau de Laplace.

4.2 Simulation

Dans cette section, on réalise une étude de simulation pour présenter les performances des nouveaux taux de convergence, obtenus par la convergence en moment du second-ordre, des estimateurs à noyaux des fonctions de densité, de répartition et de quantile. Cette étude est effectuée sur des échantillons de taille finie, dans laquelle on donne une impression visuelle de la qualité de la convergence en calculant les valeurs des erreurs moyennes quadratiques (MSE) correspondantes avec les valeurs du taux de convergence complète du moment de second-ordre et celles du taux de convergence presque complète des estimateurs étudiés, en se basant sur un échantillon obtenu à partir de deux modèles théoriques.

Premièrement, pour la fonction de densité on utilise les lois théoriques Gamma et exponentielle et des estimateurs à noyau de Laplace et celui de gamma avec la fenêtre optimale des estimateurs à noyaux de la densité de probabilité.

Dans la deuxième partie, deux estimateurs de la fonction de répartition par noyau Normal et Epanechnikov sont utilisés avec la fenêtre optimale et des tailles différentes des échantillons de loi théorique normale et exponentielle. Enfin, avec le modèle Normal, on effectue des estimations des quantiles (25%, 50%, 75%) pour la même taille d'échantillon.

4.2.1 MSE et taux des convergences de l'estimateur à noyau de la fonction de densité

On propose deux schémas d'estimations à noyau, estimation de la densité par le noyau de Laplace et par le noyau Gamma avec la fenêtre optimale $h = n^{-1/5}$. On effectue des simulations de données d'échantillons à partir de la densité exponentielle et gamma avec des tailles d'échantillon $n = 100, n = 150, 200, 250, 300, 500, 800, 1000$. Les valeurs des MSE et des taux des deux types convergences, complète en moment du second-ordre et presque complète, sont résumés dans la table(4.1).

La première remarque est que les valeurs du MSE de l'estimateur à noyau de Laplace, pour les densités de Gamma et exponentielle, sont inférieures que celle du MSE de l'estimateur à noyau Gamma, pour les différentes tailles d'échantillon, ce qui nous permet de dire que l'estimation à noyau de Laplace est meilleur que celle par le noyau Gamma pour les deux lois théoriques considérées.

4. SIMULATION ET APPLICATION

noyau	taille d'échantillon	Gamma MSE	Exp MSE	taux c.m.s	taux p.co.
Laplace	100	0.043479007	0.001242875	0.0018771917	0.013831660
	150	0.033926113	0.003267882	0.0011938521	0.011722195
	200	0.032411091	0.002689498	0.0008729775	0.010465886
	250	0.027020496	0.004629636	0.0006875970	0.009604571
	300	0.019391081	0.005242242	0.0005671386	0.008964638
	500	0.008215233	0.008617148	0.000334113	0.007428690
	800	0.023472428	0.007204569	0.000207819	0.006286744
	1000	0.026401011	0.006968009	0.000166466	0.005818125
Gamma	100	0.4151408	0.026409057	0.001877192	0.013831660
	150	0.2976568	0.050039754	0.001193852	0.011722195
	200	0.2371582	0.040335448	0.000872978	0.010465886
	250	0.2188774	0.002523508	0.000687597	0.009604571
	300	2.6515305	0.031110228	0.000567139	0.008964638
	500	0.4655139	0.028309673	0.000334113	0.007428690
	800	1.0592910	0.048968084	0.000207819	0.006286744
	1000	0.6732155	0.030791537	0.000166466	0.005818125

TABLE 4.1 – MSE et taux de convergences simulés pour la densité.

Deuxièmement, d'après les résultats obtenus dans le table (4.1), on remarque que les valeurs du taux de la convergence c.m.s. de l'estimateur à noyau de Laplace (resp. noyau de Gamma) sont plus proches des valeurs Exp MSE (resp. noyau de Gamma), que celle du taux de la convergence p.co.. Ainsi que les valeurs du taux de c.m.s. sont inférieures que celle du taux de p.co. pour les deux estimateurs à noyaux de Laplace et de Gamma et dans les deux cas des densités théoriques. Une autre remarque, La valeurs du taux de la convergence c.m.s. se diminuent avec l'augmentation de la taille de l'échantillon et tendent vers 0 plus rapidement que les valeurs du taux de la convergence p.co.. On conclut que le taux de la convergence complète en moment du second-ordre est meilleur que celui de la convergence presque complète pour les deux modèles théoriques.

4.2.2 MSE et taux des convergences de l'estimateur à noyau de la fonction de répartition

On s'intéresse à l'estimation à noyau normale et à noyau et d'Epanechnikov de la fonction de répartition, avec la fenêtre optimale $h = n^{-1/3}$, en

noyau	taille d'échantillon	MSE normale	MSE exp	taux c.m.s.	taux p.co
Normal	100	0.001111	0.00106656	0.0001468843	0.13788621
	200	0.000311	0.00028120	0.0047592004	0.16276236
	300	0.000166	0.00017937	0.0030768525	0.13788621
	500	0.000032	0.00003562	0.0017697362	0.11148639
	800	0.000070	0.00645965	0.0010607176	0.09140987
	1000	0.000052	0.00005699	0.0008311291	0.08311291
Epanechnikov	100	0.000282	0.00028770	0.0001468843	0.13788621
	200	0.000188	0.00024757	0.0047592004	0.16276236
	300	0.000087	0.00004802	0.0030768525	0.13788621
	500	0.000410	0.00043860	0.0017697362	0.11148639
	800	0.000191	0.00003853	0.0010607176	0.09140987
	1000	0.000096	0.00007708	0.0008311291	0.08311291

TABLE 4.2 – MSE et taux de convergences simulées de l'estimation de la répartition.

calculant le MSE et le taux de la convergence c.m.s. (resp. de la convergence p.co.) pour un échantillon de loi normale (resp. de loi exponentielle). La taille d'échantillon varie entre 100 et 1000. Les résultats de la simulation sont donnés par la table (4.2).

D'après les résultats donnés dans le table (4.2), on remarque que, dans les deux cas de l'estimation à noyau normal et d'Epanechnikov, les valeurs du taux de la convergence c.m.s. de l'estimateur de la fonction de répartition à noyau sont plus proches des valeurs du MSE normale et du MSE Exponentielle, respectivement, que celle du taux de la convergence p.co. Ici aussi, on remarque le taux de convergence c.m.s. est plus rapide, lorsqu'il tend vers 0, que le taux de la convergence p.co.. On remarque également sa similitude avec les valeurs de l'MSE dans les deux distributions et qu'il donne de bons résultats .

4.2.3 MSE de l'estimateur à noyau de la fonction de quantile

Maintenant, on calcule l'MSE de l'estimateur à noyau normal de la fonction de quantile. La taille de l'échantillon varie entre 100 et 1000. On résume les calculs numériques dans la table (4.3). D'après ces résultats, on remarque toujours que les valeurs du taux de la convergence c.m.s. sont plus petites que celle du taux de la convergence p.co.

4. SIMULATION ET APPLICATION

taille d'échantillon	MSE 25%	MSE 50%	MSE 75%	taux c.m.s.	taux p.co.
100	0.7203108	0.01220052	0.28767206	0.0001468843	0.1378862
200	0.2117512	0.0127754	0.507382	0.0047592004	0.1627623
300	0.449008368	0.000455053	0.372990054	0.0030768525	0.1378862
500	0.21501058	0.0165624	0.07083821	0.0017697362	0.1114863
800	0.17837084	0.06426792	0.92478425	0.0010607176	0.0914098
1000	0.43201245	0.02671836	0.87488135	0.0008311291	0.0831129

TABLE 4.3 – MSE et taux de convergences simulées de l'estimation de quantile.

Conclusion de la simulation

Finalement, on conclut que, dans tous les cas, le taux de la convergence c.m.s. des estimateurs à noyaux de la densité de probabilité, de la fonction de répartition et de la fonction de quantiles donne de meilleurs résultats que le taux de la convergence p.co. des mêmes estimateurs.

Dans la dernière section, on applique notre étude sur des données réelles recueillies du centre hospitalier universitaire Ben Badis de Constantine en 2018.

4.3 Application des données réelles

Infertilité féminine et IMC

L'indice de Masse Corporelle, IMC, est le résultat de la division du poids en Kg d'une personne par le carré de son taille en m. Il permet d'évaluer le statut pondéral.

Cette étude vise à étudier l'indice de masse corporelle (IMC) des femmes infertiles en âge de procréer. On a utilisé les données de 200 participants du Centre Hospitalier Universitaire Ben Badis de Constantine, Algérie, en 2018.

La première étape consiste à établir le test de conformité entre l'échantillon et une distribution normale. On obtient, $D = 0.079226$ plus petit que p-value 0.1623. Ceci implique l'utilisation de la densité et de la distribution de la loi normale.

La deuxième étape consiste à estimer à noyau la densité et la distribution et à les représenter dans un graphique.

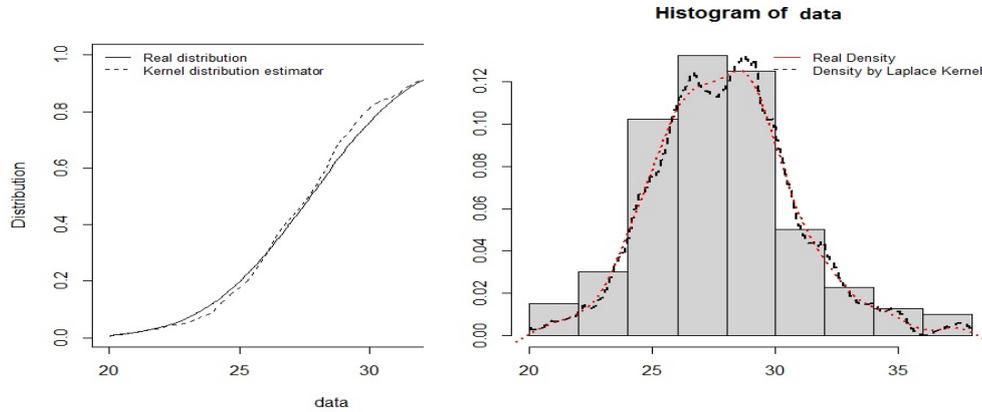


FIGURE 4.2 – Densité et répartition des données réelles avec ses estimations à noyau.

Pour le MSE entre la densité réelle et une densité de noyau normale est égale à 0,003128583 et le taux de convergence c.m.s. est de 0,0008729775 et le taux de la convergence p.co. est de 0,06541179 .

On obtient pour la distribution, $MSE = 0.003128583$ et le taux de convergence c.m.s. est de 0.0047592 et le taux de la convergence p.co. est de 0,1627624 .

Pour le même taux de convergence, on a les quantiles MSE donnés par (25%,0.0260281), (50%, 0.04186125),(75%,0.058767206).

On remarque que les résultats des données réelles sont similaires ou identiques à ceux de la simulation.

4.4 Conclusion

Le présent travail propose une nouvelle méthode pour obtenir le taux de convergence des estimateurs à noyaux des fonctions de densité, de répartition

et de quantile. Cette approche est basée sur l'expression du MSE et le taux obtenu est beaucoup plus efficace dans le cas de la convergence en moment du second-ordre. En effet, le taux de convergence complète en moment du second-ordre donne de meilleurs résultats que celui de la convergence presque complète. Les résultats précédents indiquent que le choix de ce type de convergence en utilisant l'estimation par la méthode du noyau est une bonne alternative de celle presque complète. Nous pouvons appliquer ce type de convergence dans toute estimation nécessitant l'étude de la MSE, comme par exemple le réseau de neurones, la méthode des moindres carrés,.... Elle peut être appliquée dans les statistiques neutrosophiques développées dans Smarandache 2019 [49] et Afzal et al 2021 [1]. De plus, peut-être pouvons-nous appliquer ce type de convergence pour étendre le théorème de la loi des grands nombres.

Bibliographie

- [1] Afzal, U., Alrweili, H., Ahamd, N. et al. (2021). Neutrosophic statistical analysis of resistance depending on the temperature variance of conducting material. *Sci Rep* 11, 23939.
- [2] Azzalini, A., (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, vol. 68, pp. 326-328.
- [3] Baszczynska, A., (2016). Kernel estimation of cumulative distribution function of a random variable with bounded support. *Statistics in transition*, vol. 17, pp. 541-556.
- [4] Billingsley, P., (1979). *Probability and measure*. Wiley series in Probability and statistics.
- [5] Bochner, S., (1955). *Harmonic Analysis and the Theory of probability*. University of Chicago Press, Chicogo, Illinois.
- [6] Chen, S.X., (2000). Probability density function estimation using gamma kernels. *Ann.Inst.Statist.Math.* vol. 52, pp. 471-480.
- [7] Chow, Y.S., (1988). On the rate of moment complete convergence of sample sums and extremes. *Bull.Inst.Math.Acad.Sin.* vol. 16, pp. 177-201.
- [8] David, H.A., (1981). *Order statistics*, 2nd Edition, New York : John Wiley.
- [9] Epanechnikov, V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Theory of Probability and its Applications*, vol. 14, pp. 153-158.
- [10] Falk, M., (1983). Relative deficiency and deficiency of Kernel type estimators of smooth distribution functions. *Statist. Neerlandica.* vol. 73, pp. 73-83.

- [11] Falk, M., (1984). Relative deficiency of Kernel type estimators of quantiles. *The Annals of Statistics*. vol. 12, pp. 261-268.
- [12] Ferraty, F. Vieu P., (2003). *Statistic fonctionnelle : Modèles non paramétrique de régression*, Notes de cours de DEA (2002/2003).
- [13] Földes, A. Rejeto, L. Winter, B. B., (1981). Strong consistency properties of nonparametric estimator for randomly censored data, II : Estimation of density and failure rate, *Periodica Mathematica Hungarica*, vol.12 (1), pp 15-29
- [14] Ferraty, F. Vieu P., (2006). *Nonparametric Functionnel Data Analysis Theory and Practice*, Springer Series in Statistics.
- [15] Földes, A. Rejeto, L. Winter, B. B., (1981). Strong consistency properties of nonparametric estimator for randomly censored data, II : Estimation of density and failure rate, *Periodica Mathematica Hungarica*, vol.12 (1), pp 15-29.
- [16] Ghosh, A.K. Chaudhuri, P. (2004). Optimal smoothing in kernel analysis discriminant, *Statistica Sinica*, vol.14, pp.
- [17] Ghosh, A.K. Hall, P. (2008). On error-rate estimation in nonparametric classification, *Statistica Sinica*, vol. 18, pp. 1081-1100.
- [18] Gut, A., (1978). Marcinkiwicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *Ann. Probab.* vol.6, pp. 469-482.
- [19] Gut, A., (1980). Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *Ann. Probab.* vol.8, pp. 298-313.
- [20] Gut, A., (2005). *Probability : A graduate course*, Springer tests in statistics.
- [21] Gut, A., Stadtmüller, U., (2011). An intermediate Baum-Katz theorem. *Statistics and Probability Letters*, vol.81, pp. 1486-1492.
- [22] Habbema, J.D.F., Hermans, J. Van Den Broek, K. (1974). A stepwise discriminant analysis program using density estimation, *Comp stat 1974, Proceedings in Computational Statistics*, Ed. G. Bruckmann, 101-110, Physica Verlag, Vienna.
- [23] Hall, P. Kang, K.-H. (2005). Bandwidth choice for non parametric classification, *Annals of Statistics*. vol. 33, pp.284-306.

-
- [24] Hall, P., Wand, M.P. (1988). On non parametric discrimination using density differences, *Biometrika*, vol. 75, pp. 541-547.
- [25] Hansen, B.E., (2009.) Lecture Notes on Nonparametrics. University of Wisconsin Spring.
- [26] Harell, F.E., Davis, C.E. (1982). A new distribution-free quantile estimator, *Biometrika*, vol. 69, pp. 635-640.
- [27] Hsu, P.L., Robbins, H., (1947). Complete convergence and the law of large numbers. *Proc. Natl. Acad. Sci. USA*. vol. 33, pp.25-31.
- [28] Kitouni, A., Boukeloua, M., Messaci, F., (2014). Rate of strong consistency for nonparametric estimators based on twice censored data. *Statistics and Probability Letters* vol. 96, pp.255-261.
- [29] Khan, J.A., Akbar, A., (2021). Density estimation by Laplace kernel. Working paper. Departement of Statistis, Bahauddin Zakaria university, Multan, Pakistan.
- [30] Liang, H.Y., Li, D., Rosalsky, A., (2010). Complete moment and integral convergence for sums of negatively associated random variables. *Acta Math. Sin. (Engl. Ser.)* vol.26, pp. 419-432.
- [31] Li, D., Rao, M.B., Jiang, T., Wang, X., (1995). Complete convergence and almost sure convergence of weighted sums of random variables. *Journal of Theoretical Probability* vol. 8,pp.49-76.
- [32] Li, J., Hu, Z.-C., (2017). Toeplitz lemma, Complete convergence and complete moment convergence. *Commun.in Statist. Theory and Methods* vol. 46 :4,pp.1731-1743.
- [33] Lin, Z., Bai, Z., (2010). *Probability Inequalities*, Springer.
- [34] Liu ,X. et al (2022). On the complete convergence for uncertain random variables. *Soft computing*, vol. 26, pp. 1025-1031.
- [35] Madi,M., Laroussi, I., (2022). Rate of complete second-order moment convergence and theoretical applications. *Journal of Taibah University of Science*, vol.16 :1, pp. 566-574.
- [36] Métivier, M., (1972). *Notions fondamentales de la théorie des probabilités, maîtrises de mathématiques*, 2nd edition, Dunod.
- [37] Nadaraya, E.A., (1964). Some new estimates for distribution function. *Theory of Probab.Appl*, vol.9, pp. 497-500.
- [38] Ouvrad, J.Y., 2007. *Probabilités*, Tome 1 Licence, Ellipses.

BIBLIOGRAPHIE

- [39] Padgett, W. J. (1986). A Kernel-Type Estimator of a Quantile Function from Right-Censored Data, *Journal of American Statistical Association*, vol. 81, pp. 215-222.
- [40] Parzen, E., (1962). On the estimation of a probability density function and mode . *Annals of Mathematical Statistics*, vol. 33, pp. 1065-1076.
- [41] Parzen, E., (1979). Non parametric statistical data modelling. *Journal of American Statistical Association*, vol.74, pp. 105-131.
- [42] Polansky, M.A., (2011). *Introduction to Statistical Limit Theory*, Taylor and Francis Group.
- [43] Qui, D., Chen, P., (2014). Complete moment convergence for i.i.d. random variables. *Statistics and Probability Letters*, vol. 91, pp. 76-82.
- [44] Rao, P.B.L., (1983). Nonparametric Functionnal Estimation. *Academic Press*, New York, NY.
- [45] Rosenblatt, M., (1956). Remarks on some non parametric estimates of a density function. *Annals of Mathematical Statistics*, vol.27, pp. 832-837.
- [46] Sheather, S.J., Marron, J. S., (1990). Kernel quantile estimators. *Journal of the American Statistical Association*, vol. 85, pp. 410-416.
- [47] Silverman, B.W., (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall. London.
- [48] Singh, R.S., Gasser, T., and Prasad, B., (1983). Nonparametric estimates of distributions functions. *Communication in Statistics Theory and Methods*, vol12, pp. 2095-2108.
- [49] Smarandache, F., (2019). *Neutrosophic Statistics vs. Classical Statistics, section in Nidus idearum/ Superluminal physics* vol.7, 3rd Edition, p.117.
- [50] Sung, S. H., (2007). Complete convergence for weighted sums of random variables. *Statistics and Probability Letters*, vol. 77, pp. 303-311.
- [51] Sung, S.H., Volodin, A., (2006). On the rate of complete convergence for weighted sums of arrays of random elements. *J.Korean. Math. Soc.*, vol. 43, pp. 815-828.
- [52] Swanepoel, J. W. H., (1988). Mean intgrated squared error properties and optimal kernel when estimating a distribution function. *Comm. Statist. Theory Methods*, vol. 17, pp. 3785 - 3799.

- [53] Tiago de Oliveira, J., (1963). Estatística densidades : resultados assintóticos. *Revista de la Facultad de ciencias, Universidad de Lisboa*, A9, pp. 111-206.
- [54] Watson, G.S., Leadbetter, M.R., (1964). Hazard analysis 2, . *Sankhyd Sed.A*, vol.26 , pp. 101-116.
- [55] Winter, B.B., (1973). Strong uniform consistency of integrals of density estimator. *Canad. J. Statist.*, vol. 1, pp. 247-253.
- [56] Winter, B.B., (1979). Convergence rate of perturbed empirical distribution functions. *Journal. Appli. Probab.*, vol. 16, pp. 163-173.
- [57] Yamato, H., (1973). Uniform convergence of an estimator of a distribution function. *Bull. Math. Statist.*, vol. 15, pp. 69-78.
- [58] Yang, S. S. (1985). A Smooth Nonparametric Estimator of a Quantile Function. *Journal of the American Statistical Association*, vol 80, pp. 1004-1011.
- [59] Yu, Y., et al (2022). On the complete convergence for uncertain random variables. *Soft Comput*, vol 26, pp. 1020-1031.

الملخص

الهدف من هذه الاطروحة هو تقديم طريقة جديدة لتقارب متتاليات المتغيرات العشوائية، وإنشاء معدل أفضل للتقارب شبه الكامل لمقدرات النواة للدوال المميزة لتوزيع احتمالي غير معروف. تسمى هذه الطريقة التقارب الكامل للعزوم من الدرجة الثانية مع المعدل ، و الذي يستلزم التقارب شبه كامل ويعطي معدل تقارب أصغر. في الواقع، هذه الطريقة الجديدة سهلة التطبيق وتعطي أداءً أفضل من التقارب شبه الكامل في حالة المقدرات اللامعلمية مع نواة لدالة الكثافة الاحتمالية ودالة التوزيع والدالة الكمية. من المزايا الكبيرة للنهج المقترح أنه يتم فرض عدد أقل من الشروط على دالة النواة من خلال استخدام عبارة الخطأ متوسط التربيعي.

الكلمات الرئيسية: التقدير اللامعلمي، مقدر نواة الكثافة، مقدر نواة دالة التوزيع، مقدر نواة الدالة الكمية، تقارب العزوم الكامل، معدل التقارب ، التقارب شبه الكامل.

Abstract

The aim of this thesis is to present a novel mode of convergence of sequences of random variables, and to establish a better complete convergence rate's for kernel estimators of functional characteristic's of unknown probability distribution. Complete second order moment convergence with rate is the name of this kind of convergence, which implies almost complete convergence and gives a smaller rate of convergence. Indeed, this mode is easier to obtain and gives better performances than those of the almost complete convergence in the case of the non parametric estimators with kernels of the density function, of the distribution function and of the quantile function. A great advantage of the proposed approach is that less conditions are imposed to the kernel function thanks to the use of the mean squared error expression.

Key Words: Non parametric estimation, Kernel density estimator, Kernel distribution estimator, Kernel quantile estimator, Complete moment convergence, Rate of convergence, Complete convergence.

Résumé

L'objectif de cette thèse est de présenter un nouveau mode de convergence de suites de variables aléatoires, et d'établir un meilleur taux de la convergence presque complète pour les estimateurs à noyau des caractéristiques fonctionnelle d'une loi de probabilité inconnue. On appelle ce mode la convergence complète en moment du second ordre avec taux, qui implique la convergence presque complète et donne un taux de convergence plus petit. En effet, ce mode est plus facile à obtenir et donne de meilleures performances que celles de la convergence presque complète dans le cas des estimateurs non paramétriques à noyaux de la fonction de densité, de la fonction de distribution et de la fonction de quantile. Un grand avantage de l'approche proposée est que moins de conditions sont imposées à la fonction noyau grâce à l'utilisation de l'expression de l'erreur quadratique moyenne.

Mots clés: Estimation non paramétrique, Estimateur à noyau de la densité, Estimateur à noyau de la fonction de répartition, Estimateur à noyau de la fonction de quantile, Convergence en moment complète, Taux de convergence, Convergence presque complète.