

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU

FACULTE DES SCIENCES

DEPARTEMENT DE MATHEMATIQUES

Mémoire de Master

SPECIALITE : MATHEMATIQUES OPTION :

PROBABILITES ET STATISTIQUE

Présenté par :

Ait Ouarab Malika

Thème

PROPRIÉTÉS ASYMPTOTIQUES DE QUELQUES
ESTIMATEURS À NOYAUX ASYMÉTRIQUES DE
LA FONCTION DE DENSITÉ

Devant le jury d'examen composé de :

| | | | |
|---------------------|-----------|--------------|------------|
| M. Yousfi Smail ; | M. C. B ; | U.M.M.T.O. ; | Président |
| M. Graiche Farid ; | M. C. B ; | U.M.M.T.O ; | Rapporteur |
| M. Mehiri Mohamed ; | M. A. A ; | U.M.M.T.O. ; | Examineur |

Soutenu le :

Remerciements

La louange est à Allah

Avant tout je remercie Allah le tout-puissant, de m'avoir donné la force, le courage, la volonté et la patience pour terminer ce mémoire.

Je tiens à présenter mes sincères remerciements à monsieur Graiche Farid, pour avoir accepté de m'encadrer et orienté, aidé et encouragé tout au long de la réalisation de ce travail. Pour ses conseils permanents, et sa disponibilité. J'ai beaucoup appris à ses côtés et je suis très honoré de l'avoir comme encadreur.

Je remercie également les membres du jury pour l'intérêt qu'ils ont porté à mon travail. Je remercie vivement monsieur Yousfi Smail qui m'a fait l'honneur de présider le jury de ce mémoire, monsieur Mehiri Mohamed d'avoir accepté d'examiner ce travail.

J'exprime toute ma reconnaissance et gratitude à l'administration et à l'ensemble du corps enseignant de l'université Mouloud Mammeri de Tizi-Ouzou pour leurs efforts à me garantir la continuité et l'aboutissement de ce programme de master.

Enfin, je tiens à remercier, mes parents, mes frères et soeurs pour leurs soutiens, conseils, et leur aide, ainsi que toutes les personnes ayant contribués de près ou de loin à l'élaboration de ce travail ainsi que mes ami(e)s avec lesquels j'ai partagée les bons moments inoubliables de mon cursus universitaire.

TABLE DES MATIÈRES

| | |
|---|-----------|
| Introduction | 3 |
| 1 Rappels et définitions | 6 |
| 1.1 Espérance d'une variable aléatoire | 6 |
| 1.2 Variance d'une variable aléatoire | 7 |
| 1.3 Inégalité de Bienaymé-Tchebychev | 7 |
| 1.4 Estimation d'un paramètre inconnu | 7 |
| 1.4.1 Biais d'un estimateur | 8 |
| 1.4.2 Estimateur convergent | 8 |
| 1.4.3 L'erreur quadratique moyenne (MSE) | 8 |
| 1.4.4 L'erreur quadratique moyenne intégrée (MISE) | 9 |
| 1.5 Fonction de densité Gamma | 9 |
| 1.6 Fonction de densité Bêta | 11 |
| 2 Estimation de la fonction de densité par noyau associé | 13 |
| 2.1 Noyaux | 13 |
| 2.1.1 Noyau symétrique | 13 |
| 2.1.2 Noyau asymétrique | 15 |
| 2.2 Estimateur à noyau | 17 |
| 2.2.1 Estimateur à noyau symétrique | 18 |
| 2.2.2 Estimateur à noyaux asymétriques | 19 |

| | | |
|----------|--|-----------|
| 2.2.3 | Choix du paramètre de lissage | 20 |
| 3 | Propriétés asymptotiques de quelques noyaux asymétriques | 21 |
| 3.1 | Estimateur à noyau associé gamma | 21 |
| 3.1.1 | Biais | 22 |
| 3.1.2 | Variance | 23 |
| 3.1.3 | Erreur quadratique moyenne (MSE) | 24 |
| 3.1.4 | Erreur quadratique moyenne intégrée (MISE) | 25 |
| 3.2 | Estimateur à noyau associé Bêta | 25 |
| 3.2.1 | Biais | 26 |
| 3.2.2 | Variance | 27 |
| 3.2.3 | Erreur quadratique moyenne (MSE) | 28 |
| 3.2.4 | Erreur quadratique moyenne intégrée (MISE) | 28 |
| 4 | Application numérique | 29 |
| 4.1 | Introduction | 29 |
| 4.2 | Application | 29 |
| 4.2.1 | Calcul du biais de $\hat{f}_h(x)$ | 31 |
| 4.2.2 | Calcul de $MSE(\hat{f}_h(x))$ | 31 |
| 4.2.3 | Calcul de $MISE(\hat{f}_h(x))$ et $MISE^*(\hat{f}_h(x))$ | 33 |
| | Conclusion | 37 |
| | Bibliographie | 38 |

INTRODUCTION

L'estimation fonctionnelle est un domaine très important de la statistique mathématique. Elle est divisée en deux approches principales, à savoir, l'estimation paramétrique et non paramétrique.

Dans l'approche paramétrique, l'objectif est d'estimer les paramètres d'une distribution connue par des méthodes d'estimation comme la méthode des moments et la méthode du maximum de vraisemblance.

Dans l'approche non paramétrique, la distribution n'est pas connue et on estime la densité à partir de l'information disponible sur l'ensemble des observations.

Il existe plusieurs méthodes d'estimation non paramétrique comme la méthode de l'histogramme, la méthode de séries orthogonales et la méthode splines mais la plus utilisée est la méthode du noyau, qui sera utilisée dans ce mémoire.

La méthode d'estimation non paramétrique du noyau fut introduite par Rosenblatt en 1956, puis améliorée par Parzen en 1962.

La mise en oeuvre de cette méthode nécessite le choix d'un noyau K et d'un paramètre h dit paramètre de lissage ou fenêtre qui dépend essentiellement de la taille n de l'échantillon.

Quand le support de la densité est non borné, on utilise des noyaux symétriques (dit aussi classiques). Nous citons comme exemples, les noyaux rectangulaire, triangulaire, parabolique (Epanechnikov), biweight et gaussien. Cependant, lorsqu'on veut estimer des densités à support borné au moins d'un côté, l'estimateur à noyau classique devient non consistant, à cause des effets du bord car les noyaux symétriques assignent des poids à l'extérieur du support de la densité.

Pour remédier à ce problème, Certain auteurs ont proposé l'utilisation de noyaux dont le support coïncide avec celui de la densité estimée afin de résoudre le problème des effets de bord. C'est notamment le cas de Chen (2000) avec les noyaux gamma pour estimer des densités à support dans $[0, +\infty[$ et les noyaux bêta pour les densités à support dans $[0, 1]$.

Ce mémoire est composé de quatre chapitres :

Dans le chapitre 1, nous rappelons les notions essentielles qu'on utilisera dans les chapitres suivants comme le biais et l'erreur quadratique moyenne d'un estimateur ainsi que les densités gamma et bêta.

Dans le chapitre 2, nous introduisons la notion de noyau ainsi que l'estimateur à noyau dans les deux cas, symétrique (pour les densités à support non borné) et asymétrique (pour les densités à support borné) avec des exemples dans chaque cas.

Dans le chapitre 3, nous étudions les propriétés asymptotiques (biais, erreur quadratique moyenne et erreur quadratique moyenne intégrée) des estimateurs à noyau asymétriques gamma et bêta introduits par Chen (2000).

Dans le dernier chapitre, nous effectuons une application numérique de ces propriétés asymptotiques dans le cas où les variables aléatoires suivent une loi exponentielle.

CHAPITRE 1

RAPPELS ET DÉFINITIONS

1.1 Espérance d'une variable aléatoire

L'espérance d'une variable aléatoire X notée $E[X]$ représente la valeur moyenne prise par la variable X .

Si X est une variable discrète à valeurs dans $D = \{x_1, \dots, x_n\}$, son espérance est

$$E[X] = x_1P(X = x_1) + \dots + x_nP(X = x_n) = \sum_{i=1}^n x_iP(X = x_i)$$

Lorsque X est à valeurs dans l'ensemble infini $D = \{x_i : i \geq 1\}$, son espérance est

$$E[X] = \sum_{i=1}^{+\infty} x_iP(X = x_i)$$

Si X est une variable continue de densité f , lorsque l'intégrale est bien définie, son espérance est

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

Lorsque $E[X] = 0$, on dit que la variable est centrée.

1.2 Variance d'une variable aléatoire

La variance d'une variable aléatoire X , notée $Var(X)$, est définie par

$$Var(X) = E[(X - E[X])^2]$$

La variance s'écrit aussi

$$Var(X) = E[X^2] - E[X]^2$$

L'écart type est la racine carrée de la variance

$$\sigma(X) = \sqrt{Var(X)}$$

Lorsqu'une variable X vérifie $Var(X) = 1$, on dit que la variable est réduite.

1.3 Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire positive dont l'espérance mathématique et la variance existent, alors

$$\forall \varepsilon > 0, P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

1.4 Estimation d'un paramètre inconnu

Un estimateur d'un paramètre inconnu θ est une valeur $\hat{\theta}$ calculée sur un échantillon $\{X_1, \dots, X_n\}$ tiré au hasard.

La valeur $\hat{\theta} = h(X_1, \dots, X_n)$ est donc une variable aléatoire possédant une espérance $E(\hat{\theta})$ et une variance $Var(\hat{\theta})$.

Pour chaque réalisation (x_1, \dots, x_n) de l'échantillon aléatoire, la valeur observée $\hat{\theta}_n = h(x_1, \dots, x_n)$ de l'estimateur $\hat{\theta}$ est appelé estimation de θ .

1.4.1 Biais d'un estimateur

Le biais d'un estimateur $\hat{\theta}$ est définie par

$$b_{\theta}(\hat{\theta}) = E(\hat{\theta}) - \theta$$

Un estimateur $\hat{\theta}$ est dit sans biais pour θ si

$$E(\hat{\theta}) = \theta \Leftrightarrow b_{\theta}(\hat{\theta}) = 0$$

Il est dit biaisé si

$$b_{\theta}(\hat{\theta}) \neq 0$$

Un estimateur $\hat{\theta}$ est dit asymptotiquement sans biais pour θ si

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$$

1.4.2 Estimateur convergent

Un estimateur $\hat{\theta}$ est dit convergent vers θ s'il converge en probabilité vers θ c'est à dire

$$\lim_{n \rightarrow \infty} p(|\hat{\theta} - \theta| \geq \varepsilon) = 0, \forall \varepsilon > 0.$$

D'après Bienaymé-Tchebychev pour qu'un estimateur $\hat{\theta}$ converge vers θ , il suffit que $\hat{\theta}$ soit asymptotiquement sans biais et

$$\lim_{n \rightarrow \infty} V(\hat{\theta}) = 0$$

1.4.3 L'erreur quadratique moyenne (MSE)

L'erreur quadratique moyenne d'un estimateur $\hat{\theta}$ par rapport au paramètre θ est une mesure caractérisant la "précision" de cet estimateur. Elle indique la différence quadratique moyenne entre les valeurs des données observées

et les valeurs des données estimées. Cette erreur est définie par

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

Elle peut s'écrire aussi

$$\begin{aligned} MSE(\hat{\theta}) &= Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 \\ &= Var(\hat{\theta}) + b_{\theta}^2(\hat{\theta}) \end{aligned}$$

Un bon estimateur est un estimateur qui minimise le MSE (sans biais et à variance minimale).

1.4.4 L'erreur quadratique moyenne intégrée (MISE)

L'erreur quadratique moyenne intégrée est la mesure théorique commune la plus utilisée pour évaluer l'erreur entre θ et $\hat{\theta}$.

Elle est définie par

$$\begin{aligned} MISE(\hat{\theta}) &= \int_{\mathbb{R}} MSE(\hat{\theta}) \\ &= \int_{\mathbb{R}} Var(\hat{\theta}) + \int_{\mathbb{R}} b_{\theta}^2(\hat{\theta}) \end{aligned}$$

1.5 Fonction de densité Gamma

La loi gamma de paramètres α et β (strictement positifs) est une loi continue asymétrique définie sur $\mathfrak{X} = \mathbb{R}^+$ de densité de probabilité

$$f(x, \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x) 1_{\mathbb{R}^+}(x)$$

Avec

$$\Gamma(\alpha) = \int_{\mathbb{R}^+} e^{-x} x^{\alpha-1} dx$$

$1_{\mathbb{R}^+}(x)$ est la fonction indicatrice sur \mathbb{R}^+ définie par

$$1_{\mathbb{R}^+}(x) = \begin{cases} 1 & \text{si } x \in \mathbb{R}^+ \\ 0 & \text{sinon} \end{cases}$$

Si X est une variable aléatoire suivant la loi gamma ($X \sim \Gamma(\alpha, \beta)$), alors

$$E(X) = \alpha\beta$$

$$V(X) = \alpha\beta^2$$

Les graphes ci-dessous représentent la densité gamma pour différentes valeurs de α avec $\beta = 2$.

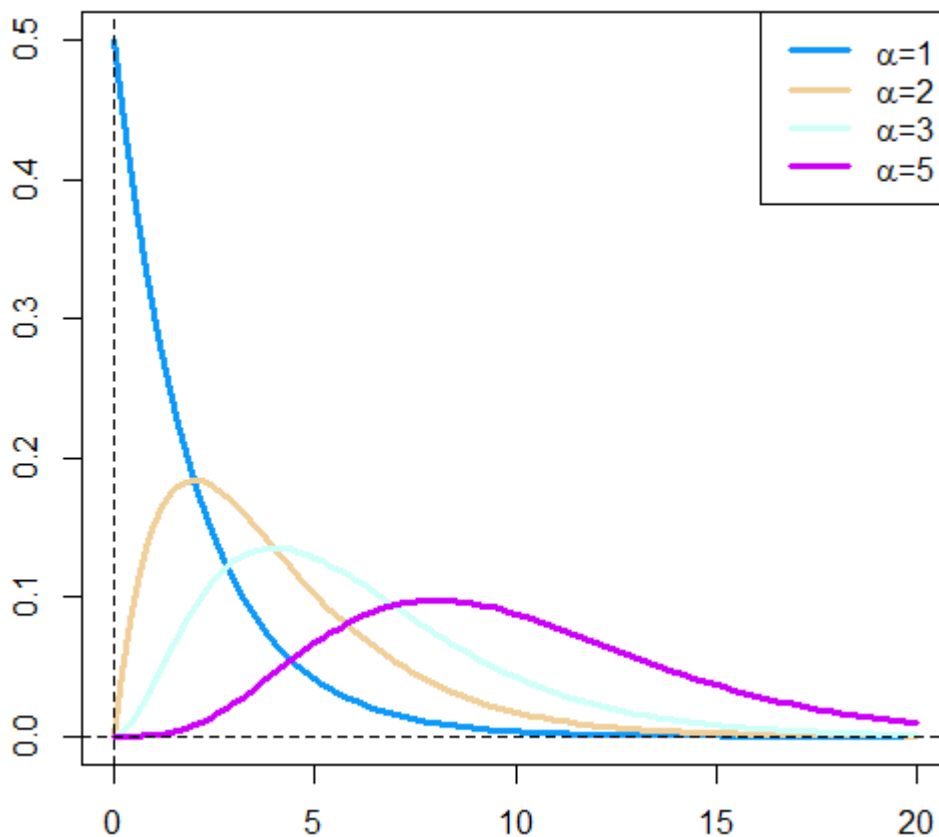


Figure 1.1. Densité Gamma pour $\beta = 2$

1.6 Fonction de densité Bêta

La loi bêta de paramètres α et β (strictement positifs) est une loi continue asymétrique définie sur $[0, 1]$, de densité de probabilité

$$f(x, \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} 1_{[0,1]}(x)$$

Où

$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du$$

On peut écrire aussi

$$B(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

Si X est une variable aléatoire suivant la loi bêta ($X \sim B(\alpha, \beta)$), alors

$$E(X) = \frac{\alpha}{\alpha + \beta}$$

$$V(X) = \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

Les graphes ci-dessous représentent la densité bêta pour différentes valeurs de α avec $\beta = 2$.

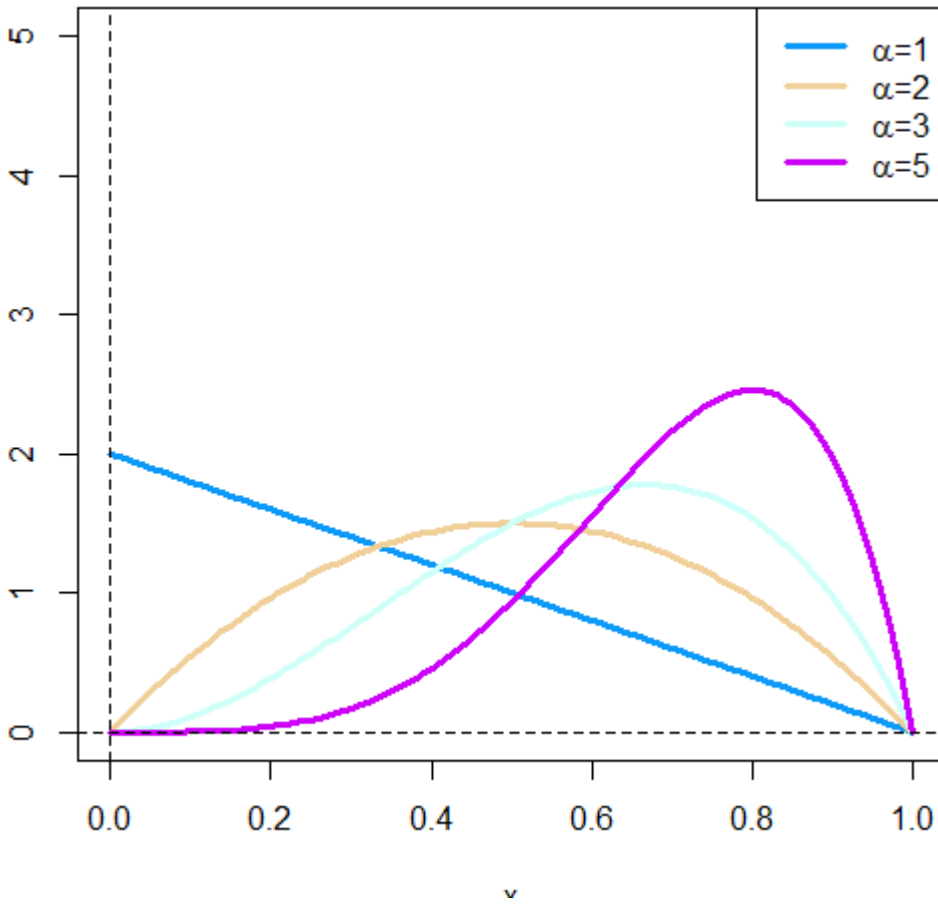


Figure 1.2. Densité Bêta pour $\beta = 2$.

CHAPITRE 2

ESTIMATION DE LA FONCTION DE DENSITÉ PAR NOYAU ASSOCIÉ

2.1 Noyaux

Dans cette partie, nous nous intéressons à la notion de noyau qu'on utilise pour estimer la densité de probabilité inconnue d'une variable aléatoire.

Définition

Soit $K : \mathbb{R} \rightarrow \mathbb{R}$, on dit que K est un noyau si et seulement si :

$$\int_{\mathbb{R}} K(u) du = 1$$

$$K(u) \geq 0, \forall u \in \mathbb{R}$$

2.1.1 Noyau symétrique

Un noyau est dit symétrique si, pour tout u dans son ensemble de définition $K(u) = K(-u)$.

Chapitre 2. Estimation de la fonction de densité par noyau associé 14

K satisfait les conditions suivantes :

$$\int_{\mathbb{R}} uK(u) du = 0$$

,

$$\int_{\mathbb{R}} u^2 K(u) du = \sigma_K^2 < \infty$$

.

Exemples de noyaux symétriques

Le tableau suivant donne les formules pour quelques noyaux symétriques.

| Noyaux | Supports | Densités |
|---------------|--------------|--|
| Epanechnikov | $[-1, 1]$ | $K(u) = \frac{3}{4}(1 - u^2)$ |
| Gaussien | \mathbb{R} | $K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right)$ |
| Rectangulaire | $[-1, 1]$ | $K(u) = \frac{1}{2}$ |
| Triangulaire | $[-1, 1]$ | $K(u) = 1 - u $ |
| Biweight | $[-1, 1]$ | $K(u) = \frac{15}{16}(1 - u^2)^2$ |

Table 2.1. Noyaux symétriques

Les graphes des noyaux précédents sont représentés dans la figure suivante :

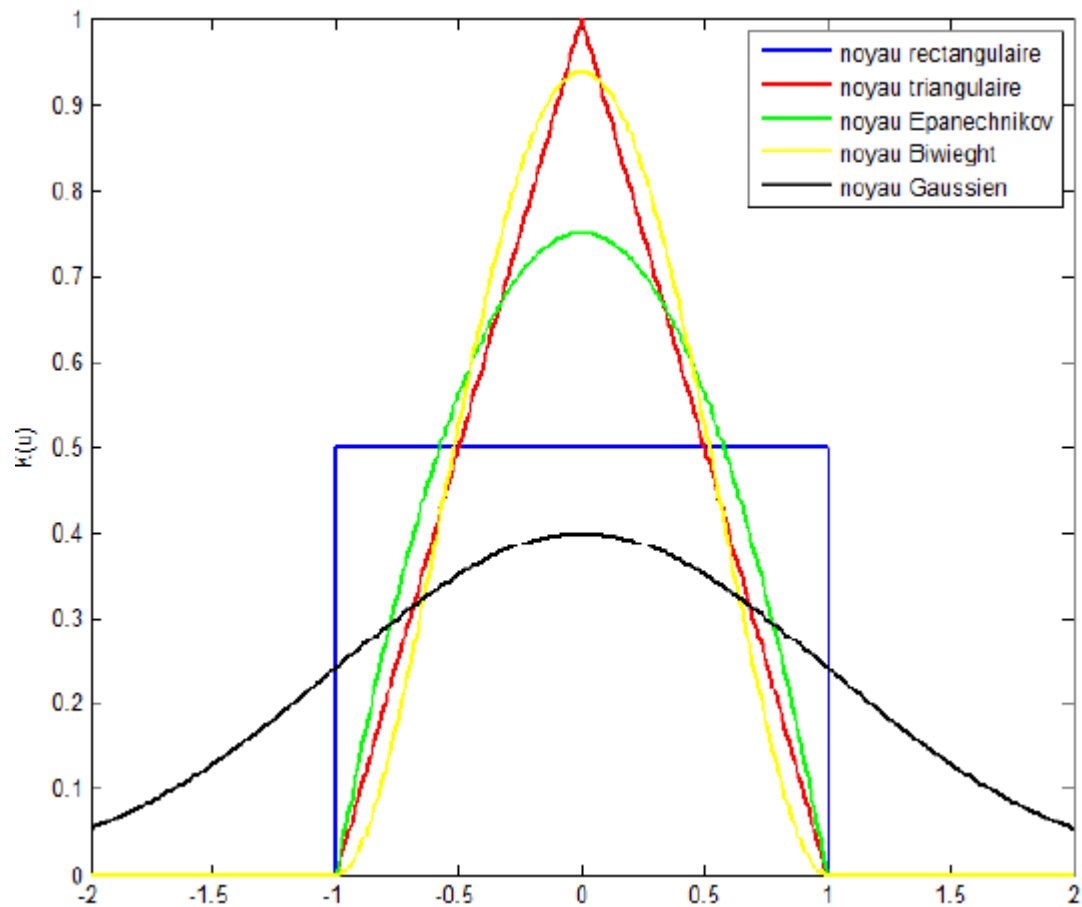


Figure 2.1. Noyaux symétriques

2.1.2 Noyau asymétrique

Lorsqu'on veut estimer des densités à support borné au moins d'un coté, l'estimateur à noyau classique (symétrique) devient non consistant à cause des effets du bord. La solution la plus simple consiste à remplacer le noyau symétrique par un noyau asymétrique.

Exemples de noyaux asymétriques

Le tableau ci-dessous donne les formules de deux noyaux asymétriques.

Chapitre 2. Estimation de la fonction de densité par noyau associé 16

| Noyaux | Supports | Densités |
|--------|----------------|---|
| Gamma | \mathbb{R}^+ | $K_{\frac{x}{h}+1, h}(u) = \frac{u^{\frac{x}{h}} \exp\left(-\frac{u}{h}\right)}{h^{\frac{x}{h}+1} \Gamma\left(\frac{x}{h}+1\right)}$ |
| Bêta | $[0, 1]$ | $K_{\frac{x}{h}+1, \left(\frac{1-x}{h}\right)+1}(u) = \frac{u^{\frac{x}{h}} (1-u)^{\frac{1-x}{h}}}{\beta\left(\frac{x}{h}+1, \frac{1-x}{h}+1\right)}$ |

Table 2.2. Noyaux gamma et bêta

Les courbes des deux noyaux précédents sont représentées ci dessous pour $x = 0.5$ avec différentes valeurs de h .

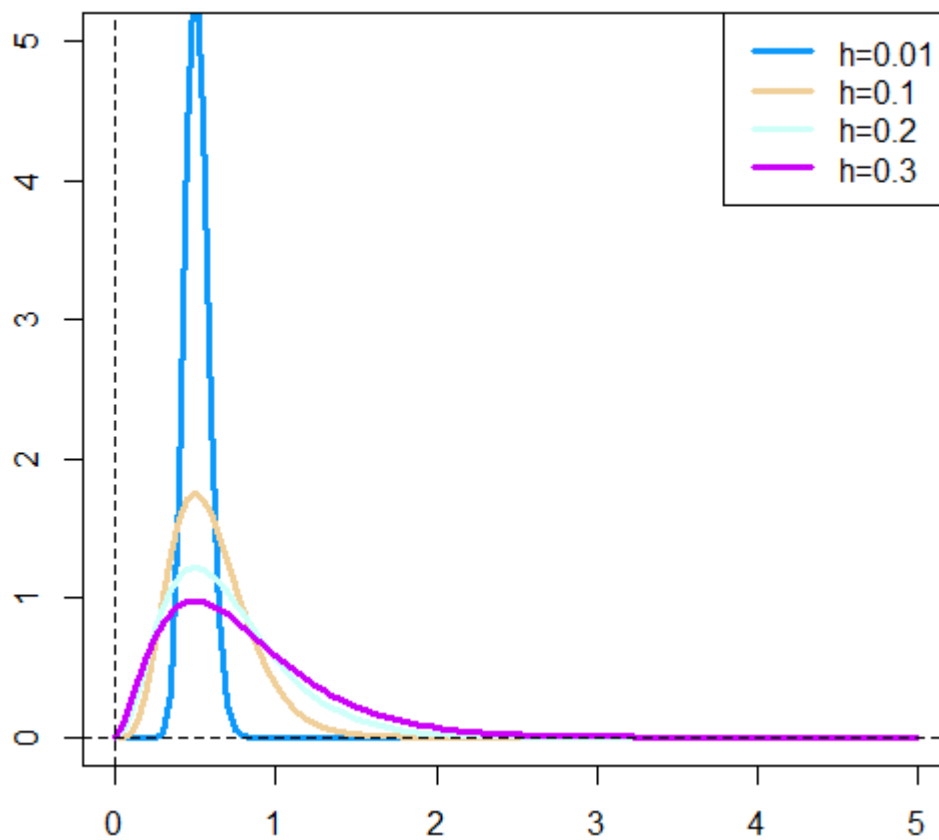


Figure 2.2. Noyau Gamma pour $x=0.5$

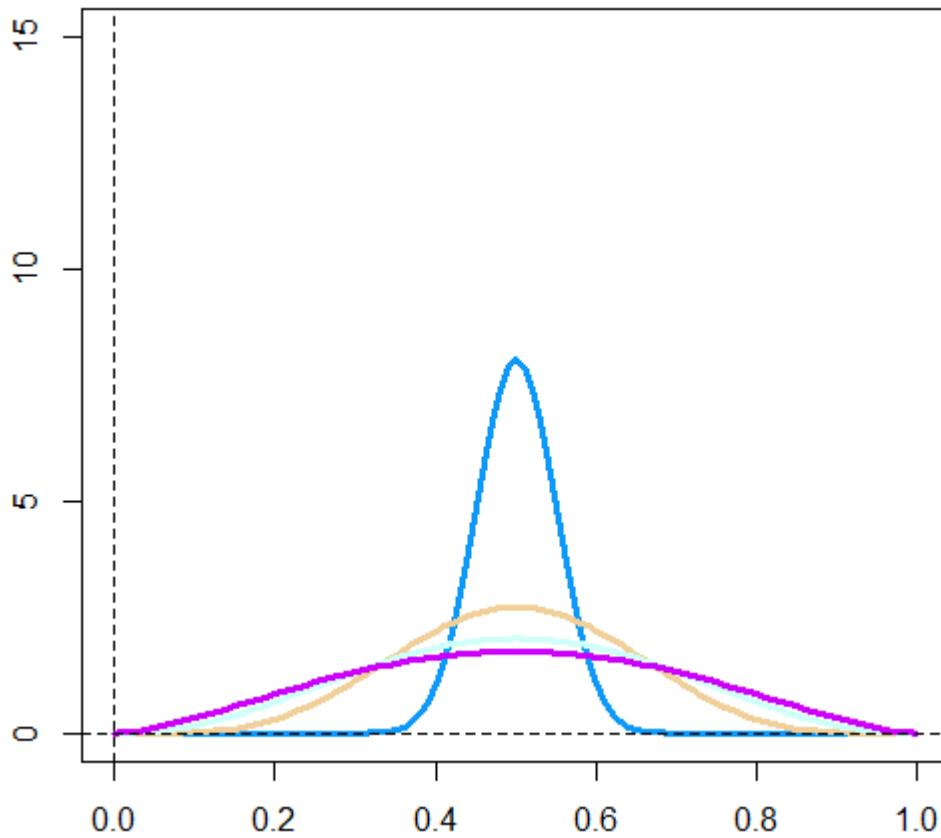


Figure 2.3. Noyau Beta pour $x=0.5$

2.2 Estimateur à noyau

Soit X_1, \dots, X_n un échantillon de densité f sur \mathbb{R} et de fonction de répartition $F(x) = \int_{-\infty}^x f(t)dt$.

On appelle fonction de répartition empirique associé à X_1, \dots, X_n la fonction aléatoire $F_n(x) : \mathbb{R} \rightarrow [0, 1]$ définie par :

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

Chapitre 2. Estimation de la fonction de densité par noyau associé 18

A partir de la définition d'une densité de probabilité, on définit

$$\hat{f}(x) = \frac{F_n(x+h) - F_n(x-h)}{2h}, \quad \text{avec } h \rightarrow 0$$

Cette dernière peut être réécrite en ses points de continuité, sous la forme suivante :

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

Où

$$K(u) = \begin{cases} \frac{1}{2} & \text{si } -1 \leq u \leq 1 \\ 0 & \text{sinon} \end{cases}$$

L'estimateur \hat{f} est dit estimateur à noyau de Rosenblatt. C'est le premier exemple d'estimateur à noyau construit à l'aide du noyau $K(u)$.

2.2.1 Estimateur à noyau symétrique

Si X_1, X_2, \dots, X_n est un échantillon de variables aléatoires indépendantes, de densité f inconnue, alors l'estimateur à noyau symétrique de f peut être défini comme suit :

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (2.1)$$

où K est le noyau associé symétrique et $h = h_n$ est appelé paramètre de lissage.

La fonction $\hat{f}(x)$ ainsi définie est une densité de probabilité.

Exemples d'estimateurs à noyaux symétriques

| Noyaux | Supports | Densités | Estimateur |
|---------------|--------------|---|--|
| Epanechnikov | $[-1, 1]$ | $K(u) = \frac{3}{4}(1 - u^2)$ | $\hat{f}_h(x) = \frac{3}{4nh} \sum_{i=1}^n [1 - (\frac{x-X_i}{h})^2]$ |
| Gaussien | \mathbb{R} | $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2})$ | $\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp(-\frac{(x-X_i)^2}{2h^2})$ |
| Rectangulaire | $[-1, 1]$ | $K(u) = \frac{1}{2}$ | $\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K(\frac{x-X_i}{h})$ |
| Triangulaire | $[-1, 1]$ | $K(u) = 1 - u $ | $\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n [1 - \frac{x-X_i}{h}]$ |
| Biweight | $[-1, 1]$ | $K(u) = \frac{15}{16}(1 - u^2)^2$ | $\hat{f}_h(x) = \frac{15}{16nh} \sum_{i=1}^n [1 - (\frac{x-X_i}{h})^2]^2$ |

Table 2.3. Estimateurs à noyaux symétrique

2.2.2 Estimateur à noyaux asymétriques

Lorsque le support de la densité f est borné, le biais de l'estimateur (2.1) présente un problème au voisinage de la borne car le noyau symétrique assigne un poids en dehors du support lorsque le lissage est pris en compte près de la borne.

Pour y remédier, on remplace le noyau symétrique par un noyau asymétrique et l'estimateur aura la forme suivante :

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_{x,h}(X_i), x \in \mathbb{R} \tag{2.2}$$

avec $K_{x,h}$ un noyau asymétrique de cible x et de fenêtre h .

Exemples d'estimateurs à noyaux asymétriques

| Noyaux | Supports | Densités | estimateur |
|--------|----------------|---|---|
| Gamma | \mathbb{R}^+ | $K_{\frac{x}{h}+1,h}(u) = \frac{u^{\frac{x}{h}} \exp^{-\frac{u}{h}}}{h^{\frac{x}{h}+1} \Gamma(\frac{x}{h}+1)}$ | $\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_{\frac{x}{h}+1,h}(X_i)$ |
| Bêta | $[0, 1]$ | $K_{\frac{x}{h}+1, \frac{(1-x)}{h}+1}(u) = \frac{u^{\frac{x}{h}} (1-u)^{\frac{1-x}{h}}}{\beta(\frac{x}{h}+1, \frac{1-x}{h}+1)}$ | $\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(X_i)$ |

Table 2.4. Estimateurs à noyaux gamma et bêta

2.2.3 Choix du paramètre de lissage

Le paramètre de lissage h a une grande influence sur la performance de l'estimateur \hat{f} .

Le h optimal est celui qui minimise l'erreur quadratique moyenne intégrée (MISE).

Exemple 2.1

Soient X_1, X_2, \dots, X_n un échantillon de variables aléatoires indépendantes de loi normale $N(0, \sigma^2)$.

Si le noyau est $K \sim N(0, 1)$, on obtient $h_{opt} = 1.06\hat{\sigma}n^{-\frac{1}{5}}$, où $\hat{\sigma}$ est l'estimateur de σ .

CHAPITRE 3

PROPRIÉTÉS ASYMPTOTIQUES DE QUELQUES NOYAUX ASYMÉTRIQUES

Dans cette partie nous présentons les différentes propriétés de quelques estimateurs à noyau asymétriques (gamma et bêta) introduits par Chen (2000) telles que le biais, la variance, le MSE et le MISE.

3.1 Estimateur à noyau associé gamma

Soit X_1, X_2, \dots, X_n un échantillon de variables aléatoires indépendantes identiquement distribuées, d'une densité de probabilité continue f inconnue à support positif ($[0, +\infty[$) et deux fois continûment dérivable ($f \in C^2([0, +\infty])$).

Nous considérons l'estimateur \hat{f}_G à noyau gamma

$$\hat{f}_G(x) = \frac{1}{n} \sum_{i=1}^n K_{\frac{x}{h}+1, h}(X_i)$$

Chapitre 3. Propriétés asymptotiques de quelques noyaux asymétriques 22

Avec

$$K_{\frac{x}{h}+1, h}(u) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\Gamma(\frac{x}{h} + 1)} \times \frac{u^{\frac{x}{h}} e^{-\frac{u}{h}}}{h^{\frac{x}{h}+1}},$$

où $h > 0$ est le paramètre de lissage et K est le noyau associé à une variable aléatoire de loi gamma de paramètres $\frac{x}{h} + 1$ et h .

On suppose de plus que

$$\lim_{n \rightarrow \infty} h = 0 \text{ et } \lim_{n \rightarrow \infty} nh = +\infty$$

3.1.1 Biais

$$\text{Biais}(\hat{f}_G(x)) = E(\hat{f}_G(x)) - f(x)$$

$$\begin{aligned} E(\hat{f}_G(x)) &= E(K_{\frac{x}{h}+1, h}(u)) = \int_0^{\infty} K_{\frac{x}{h}+1, h}(y) f(y) dy \\ &= E(f(\varepsilon_x)) \end{aligned}$$

Avec ε_x une variable aléatoire suivant une loi gamma de paramètres $\frac{x}{h} + 1$ et h .

On a

$$u_x = E(\varepsilon_x) = x + h, \text{ Var}(\varepsilon_x) = xh + h^2$$

Par la formule de Taylor $f(\varepsilon_x)$ au voisinage du point u_x , on a

$$f(\varepsilon_x) = f(u_x) + f'(u_x)(\varepsilon_x - u_x) + \frac{1}{2} f''(u_x)(\varepsilon_x - u_x)^2 + o(h)$$

Ceci implique que

$$\begin{aligned} E(\hat{f}_G(x)) &= E(f(u_x)) + \frac{1}{2} f''(u_x) E(\varepsilon_x - u_x)^2 + o(h) \\ &= f(u_x) + \frac{1}{2} f''(u_x) \text{Var}(\varepsilon_x) + o(h) \\ &= f(x + h) + \frac{1}{2} f''(x + h)(xh + h^2) + o(h) \\ &= f(x) + hf'(x) + \frac{1}{2} x f''(x) + o(h) \end{aligned}$$

Le biais de \hat{f}_G est

$$\begin{aligned} \text{Biais}(\hat{f}_G(x)) &= E(\hat{f}_G(x)) - f(x) \\ &= h(f'(x) + \frac{1}{2}x f''(x)) + o(h) \quad (3.1) \end{aligned}$$

Remarque 3.1

Lorsque n tend vers $+\infty$ (h tend vers 0), alors le Biais de \hat{f}_G tend vers 0, ce que veut dire que \hat{f}_G est asymptotiquement sans biais pour f .

3.1.2 Variance

$$\begin{aligned} \text{Var}(\hat{f}_G(x)) &= \frac{1}{n} \text{Var}(K_{\frac{x}{h}+1, h}(u)) \\ &= \frac{1}{n} (E(K_{\frac{x}{h}+1, h}(u))^2) - \frac{1}{n} (E(K_{\frac{x}{h}+1, h}(u)))^2 \\ &\leq \frac{1}{n} (E(K_{\frac{x}{h}+1, h}(u))^2) \\ &= \frac{1}{n} \int_0^\infty K_{\frac{x}{h}+1, h}^2(t) f(t) dt \\ &= \frac{1}{n} \int_0^\infty \left(\frac{t^{\frac{x}{h}} e^{-\frac{t}{h}}}{h^{\frac{x}{h}+1} \Gamma(\frac{x}{h} + 1)} \right)^2 f(t) dt \end{aligned}$$

D'où

$$\begin{aligned} \text{Var}(\hat{f}_G(x)) &\leq \frac{1}{n} \int_0^\infty \frac{t^{\frac{2x}{h}} e^{-\frac{2t}{h}}}{h^{\frac{2x}{h}+2} \Gamma^2(\frac{x}{h} + 1)} f(t) dt \\ &= \frac{1}{n} \int_0^\infty \frac{t^{\frac{2x}{h}} e^{-\frac{t}{\frac{h}{2}}}}{\left(\frac{h}{2}\right)^{\frac{2x}{h}+1} \Gamma(\frac{2x}{h} + 1)} \times \frac{\Gamma(\frac{2x}{h} + 1)}{h \Gamma^2(\frac{x}{h} + 1) 2^{\frac{2x}{h}+1}} f(t) dt \\ &= \frac{1}{n} \times \frac{\Gamma(\frac{2x}{h} + 1)}{h \Gamma^2(\frac{x}{h} + 1) 2^{\frac{2x}{h}+1}} \int_0^\infty K_{\frac{2x}{h}+1, \frac{h}{2}}(t) f(t) dt \end{aligned}$$

Ce qui donne

$$V(\hat{f}_G(x)) \leq \frac{1}{n} \left(\frac{h^{-1} \Gamma(\frac{2x}{h} + 1)}{2^{\frac{2x}{h}+1} \Gamma^2(\frac{x}{h} + 1)} \right) E(f(\eta_x))$$

Chapitre 3. Propriétés asymptotiques de quelques noyaux asymétriques 24

Où η_x est une variable aléatoire suivant la loi Gamma de paramètres $\frac{2x}{h} + 1$ et $\frac{h}{2}$

En notant

$$B_h(x) = \frac{h^{-1} \Gamma(\frac{2x}{h} + 1)}{2^{\frac{2x}{h}+1} \Gamma^2(\frac{x}{h} + 1)}$$

On aura

$$Var(\hat{f}_G(x)) \leq n^{-1} B_h(x) E(f(\eta_x))$$

D'après Chen (2010), on a pour h suffisamment petit

$$B_h(x) \sim \begin{cases} \frac{1}{2\sqrt{\pi}} h^{-\frac{1}{2}} x^{-\frac{1}{2}} & \text{si } \frac{x}{h} \rightarrow \infty \\ \frac{\Gamma(2k+1)}{2^{2k+1} \Gamma^2(k+1)} h^{-1} & \text{si } \frac{x}{h} \rightarrow k \end{cases}$$

Ce qui implique que

$$Var(\hat{f}_G(x)) \sim \begin{cases} \frac{1}{2\sqrt{\pi}} n^{-1} h^{-\frac{1}{2}} x^{-\frac{1}{2}} f(x) & \text{si } \frac{x}{h} \rightarrow \infty \\ \frac{\Gamma(2k+1)}{2^{1+2k} \Gamma^2(k+1)} n^{-1} h^{-1} f(x) & \text{si } \frac{x}{h} \rightarrow k \end{cases} \quad (3.2)$$

3.1.3 Erreur quadratique moyenne (MSE)

$$\begin{aligned} MSE(\hat{f}_G(x)) &= \text{Biais}^2(\hat{f}_G(x)) + Var(\hat{f}_G(x)) \\ &= [h(f'(x) + \frac{1}{2} x f''(x))]^2 + \frac{1}{2\sqrt{\pi}} n^{-1} h^{-\frac{1}{2}} x^{-\frac{1}{2}} f(x) + o(h^2 + n^{-1} h^{-\frac{1}{2}}) \end{aligned} \quad (3.3)$$

Remarque 3.2

Quand $n \rightarrow \infty$ ($h \rightarrow 0$) et $nh \rightarrow \infty$, on a $MSE(\hat{f}_G(x))$ tend vers 0.

3.1.4 Erreur quadratique moyenne intégrée (MISE)

$$\begin{aligned}
 MISE(\hat{f}_G(x)) &= \int_0^\infty MSE(\hat{f}_G(x)) \\
 &= h^2 \int_0^\infty \left\{ x f'(x) + \frac{1}{2} x f''(x) \right\}^2 dx + \frac{1}{2\sqrt{\pi}} n^{-1} h^{-\frac{1}{2}} \int_0^\infty x^{-\frac{1}{2}} f(x) dx \\
 &+ o(n^{-1} h^{-\frac{1}{2}} + h^2) \quad (3.4)
 \end{aligned}$$

Le paramètre h qui minimise $MISE(\hat{f}_G(x))$ est donné par

$$h^* = \frac{\left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-\frac{1}{2}} f(x) dx \right]^{\frac{2}{5}}}{4^{\frac{2}{5}} \left[\int_0^\infty \left\{ x f'(x) + \frac{1}{2} x f''(x) \right\}^2 dx \right]^{\frac{2}{5}}} n^{-\frac{2}{5}} \quad (3.5)$$

Le MISE optimal (pour h^*) est donné par

$$MISE^*(\hat{f}_G(x)) = \frac{5}{4^{\frac{2}{5}}} \left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-\frac{1}{2}} f(x) dx \right]^{\frac{4}{5}} \left[\int_0^\infty \left\{ x f'(x) + \frac{1}{2} x f''(x) \right\}^2 dx \right]^{\frac{1}{5}} n^{-\frac{4}{5}} \quad (3.6)$$

3.2 Estimateur à noyau associé Bêta

Soit X_1, X_2, \dots, X_n un échantillon de variables aléatoires indépendantes identiquement distribuées d'une densité de probabilité continue f inconnue à support $[0,1]$ et deux fois continûment dérivable ($f \in C^2([0, +\infty])$).

Nous considérons l'estimateur \hat{f}_B à noyau Bêta

$$\hat{f}_B(x) = \frac{1}{n} \sum_{i=1}^n K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(X_i)$$

avec

$$K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u) = \frac{1}{n} \sum_{i=1}^n \frac{1}{B\left(\frac{x}{h}+1, \frac{1-x}{h}+1\right)} u^{\frac{x}{h}} (1-u)^{\frac{1-x}{h}},$$

où $h > 0$ est le paramètre de lissage et K est le noyau associé à une variable aléatoire de loi bêta de paramètres $\frac{x}{h} + 1$ et $\frac{1-x}{h} + 1$

On suppose de plus que

$$\lim_{n \rightarrow \infty} h = 0 \text{ et } \lim_{n \rightarrow \infty} nh = +\infty$$

3.2.1 Biais

$$\text{Biais}(\hat{f}_B(x)) = E(\hat{f}_B(x)) - f(x)$$

$$\begin{aligned} E(\hat{f}_B(x)) &= E(K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u)) \\ &= \int_0^1 K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(y) f(y) dy = E(f(\varepsilon_x)) \end{aligned}$$

Avec ε_x une variable aléatoire suivant une loi bêta de paramètres $\frac{x}{h} + 1$ et $\frac{1-x}{h} + 1$

On a

$$u_x = E(\varepsilon_x) = (1 - 2x)h, V(\varepsilon_x) = hx(1 - x)$$

Par la formule de Taylor de $f(\varepsilon_x)$ au voisinage du point u_x , on a

$$f(\varepsilon_x) = f(u_x) + f'(u_x)(\varepsilon_x - u_x) + \frac{1}{2}f''(u_x)(\varepsilon_x - u_x)^2 + o(h)$$

Ceci implique que

$$\begin{aligned} E(\hat{f}_B(x)) &= E(f(u_x)) + \frac{1}{2}f''(u_x)E(\varepsilon_x - u_x)^2 + o(h) \\ &= f(u_x) + \frac{1}{2}f''(u_x)Var(\varepsilon_x) + o(h) \\ &= f(x) + \{(1 - 2x)f'(x) + \frac{1}{2}x(1 - x)f''(x)\}h + o(h) \end{aligned}$$

Le Biais de \hat{f}_B est

$$\begin{aligned} \text{Biais}(\hat{f}_B(x)) &= E(\hat{f}_B(x)) - f(x) \\ &= \{(1 - 2x)f'(x) + \frac{1}{2}x(1 - x)f''(x)\}h + o(h) \quad (3.7) \end{aligned}$$

Remarque 3.3

Lorsque n tend vers $+\infty$ (h tend vers 0), alors Biais ($\hat{f}_B(x)$) tend vers 0, ce qui veut dire que \hat{f}_B est asymptotiquement sans biais pour f .

3.2.2 Variance

$$\begin{aligned} Var(\hat{f}_B(x)) &= \frac{1}{n} Var(K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u)) \\ &= \frac{1}{n} (E(K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u))^2) - \frac{1}{n} (E(K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u)))^2 \end{aligned}$$

On a

$$\begin{aligned} (K_{\frac{x}{h}+1, \frac{1-x}{h}+1}(u))^2 &= \frac{1}{B^2(\frac{x}{h}+1, \frac{1-x}{h}+1)} u^{\frac{2x}{h}} (1-u)^{\frac{2(1-x)}{h}} \\ &= \frac{B(\frac{2x}{h}+1, \frac{2(1-x)}{h}+1)}{B^2(\frac{x}{h}+1, \frac{1-x}{h}+1)} \times \frac{u^{\frac{2x}{h}} (1-u)^{\frac{2(1-x)}{h}}}{B(\frac{2x}{h}+1, \frac{2(1-x)}{h}+1)} \\ &= \frac{B(\frac{2x}{h}+1, \frac{2(1-x)}{h}+1)}{B^2(\frac{x}{h}+1, \frac{1-x}{h}+1)} \times K_{\frac{2x}{h}+1, \frac{2(1-x)}{h}+1}(u) \end{aligned}$$

Ce qui donne

$$E(k_{\frac{x}{h}+1, \frac{(1-x)}{h}+1}(u))^2 = A_h(x) E\{f(y_x)(u)\}$$

Où y_x est une variable aléatoire suivant une loi bêta de paramètres $\frac{2x}{h} + 1$

et $\frac{2(1-x)}{h} + 1$ et $A_h(x) = \frac{B(\frac{2x}{h}+1, \frac{2(1-x)}{h}+1)}{B^2(\frac{x}{h}+1, \frac{1-x}{h}+1)}$

D'après Chen (2010), on a pour h suffisamment petit

$$A_h(x) \sim \begin{cases} \frac{1}{2\sqrt{\pi}} \{x(1-x)\}^{-\frac{1}{2}} h^{-\frac{1}{2}} \text{ si } \frac{x}{h} \text{ et } \frac{1-x}{h} \rightarrow \infty \\ \frac{\Gamma(2k+1)}{2^{2k+1} \Gamma^2(k+1)} h^{-1} \text{ si } \frac{x}{h} \text{ ou } \frac{1-x}{h} \rightarrow k \end{cases}$$

Ce qui implique que

$$Var(\hat{f}_B(x)) \sim \begin{cases} \frac{1}{2\sqrt{\pi}} \times \frac{n^{-1}h^{-\frac{1}{2}}}{\{x(x-1)\}^{\frac{1}{2}}} \{f(x) + o(n^{-1})\} & \text{si } \frac{x}{h} \text{ et } \frac{1-x}{h} \rightarrow \infty \\ \frac{\Gamma(2k+1)}{2^{2k+1}\Gamma^2(k+1)} \times \frac{1}{nh} \{f(x) + o(n^{-1})\} & \text{si } \frac{x}{h} \rightarrow k \text{ ou } \frac{1-x}{h} \rightarrow k \end{cases} \quad (3.8)$$

3.2.3 Erreur quadratique moyenne (MSE)

$$\begin{aligned} MSE(\hat{f}_B(x)) &= Bias^2(\hat{f}_B(x)) + V(\hat{f}_B(x)) \\ &= [\{(1-2x)f'(x) + \frac{1}{2}x(1-x)f''(x)\}h + o(h)]^2 + \frac{1}{2\sqrt{\pi}} \times \frac{n^{-1}h^{-\frac{1}{2}}}{\{x(1-x)\}^{\frac{1}{2}}} \{f(x) + o(n^{-1})\} \\ &= [h(1-2x)f'(x) + \frac{1}{2}x(1-x)f''(x)]^2 + \frac{1}{2\sqrt{\pi}} \times \frac{n^{-1}h^{-\frac{1}{2}}}{\{x(1-x)\}^{\frac{1}{2}}} f(x) + o(n^{-1}h^{-\frac{1}{2}} + h^2) \end{aligned} \quad (3.9)$$

Remarque 3.4

Quand $n \rightarrow \infty$ ($h \rightarrow 0$) et $nh \rightarrow \infty$, on a MSE tend vers 0.

3.2.4 Erreur quadratique moyenne intégrée (MISE)

$$\begin{aligned} MISE(\hat{f}_B(x)) &= \int_0^1 MSE(\hat{f}_B(x)) \\ &= h^2 \int_0^1 \{(1-2x)f'(x) + \frac{1}{2}x(1-x)f''(x)\}^2 dx + \frac{1}{2\sqrt{\pi}} n^{-1}h^{-\frac{1}{2}} \int_0^1 \{x(1-x)\}^{-\frac{1}{2}} f(x) dx \\ &+ o(n^{-1}h^{-\frac{1}{2}} + h^2) \end{aligned} \quad (3.10)$$

Le paramètre h qui minimise $MISE(\hat{f}_B(x))$ est

$$h^* = \frac{[\frac{1}{2\sqrt{\pi}} \int_0^1 \{x(1-x)\}^{-\frac{1}{2}} f(x) dx]^{\frac{2}{5}}}{4^{\frac{2}{5}} [\int_0^1 (1-2x)f'(x) + \frac{1}{2}x(1-x)f''(x) dx]^{\frac{2}{5}}} \times n^{-\frac{2}{5}} \quad (3.11)$$

Le MISE optimal (pour h^*) est donné par

$$MISE^*(\hat{f}_B(x)) = \frac{5}{4^{\frac{4}{5}}} \left[\frac{1}{2\sqrt{\pi}} \int_0^1 \frac{f(x)}{\{x(1-x)\}^{\frac{1}{2}}} dx \right]^{\frac{4}{5}} \left[\int_0^1 (1-2x)f'(x) + \frac{1}{2}x(1-x)f''(x) dx \right]^{\frac{1}{5}} n^{-\frac{4}{5}} \quad (3.12)$$

CHAPITRE 4

APPLICATION NUMÉRIQUE

4.1 Introduction

Dans ce chapitre, on va utiliser les formules (3.1), (3.3), (3.5) et (3.6) pour calculer les valeurs du biais, du MSE et du MISE d'un estimateur à noyau gamma, dans le cas où les observations suivent une loi exponentielle de paramètre $\frac{1}{2}$.

4.2 Application

Soient X_1, X_2, \dots, X_n des variables aléatoires indépendantes identiquement distribuées suit une loi exponentielle de paramètre $\frac{1}{2}$ de densité f définie par

$$f(x) = \frac{1}{2} e^{-\frac{1}{2}x}, \quad x \geq 0$$

Comme le support de la densité f est $[0, +\infty[$, on utilise l'estimateur à noyau Gamma défini par

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_{\frac{x}{h}+1, h}(X_i)$$

avec $K_{\frac{x}{h}+1,h}$ un noyau gamma de paramètres $\frac{x}{h} + 1$ et h .

Dans la figure suivante, on représente le graphe de l'estimateur \hat{f}_h en rouge et celui de la densité f en bleu pour différents n .

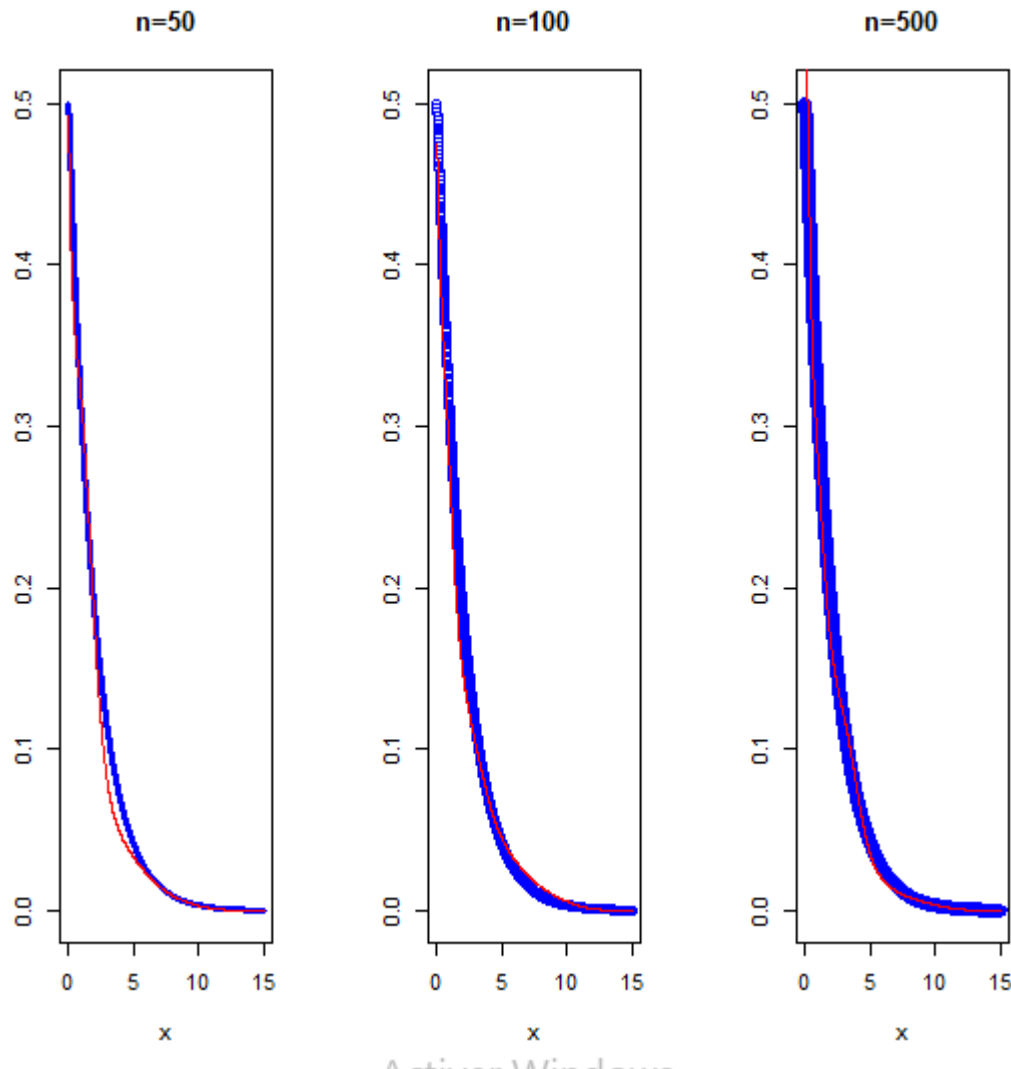


Figure 4.1. Courbes de \hat{f}_h et f

On voit sur la figure que quand n augmente, la courbe de \hat{f}_h s'approche de celle de la densité théorique f .

4.2.1 Calcul du biais de $\hat{f}_h(x)$

En utilisant la formule (3.1), on calcule les valeurs du biais de \hat{f}_h pour différents x et h et on obtient le tableau suivant

| h | $Biais(\hat{f}_h(0.1))$ | $Biais(\hat{f}_h(0.5))$ | $Biais(\hat{f}_h(2))$ |
|-----|-------------------------|-------------------------|-----------------------|
| 0.1 | -0.02318622 | -0.01703627 | -0.004598493 |
| 0.2 | -0.04637243 | -0.03407253 | -0.009196986 |
| 0.5 | -0.11593109 | -0.08518134 | -0.022992465 |
| 0.7 | -0.16230352 | -0.11925387 | -0.032189451 |
| 0.8 | -0.18548974 | -0.13629014 | -0.036787944 |

Table 4.1. Valeurs de $Biais(\hat{f}_h(x))$

On remarque que quand h augmente, la valeur absolue du biais de $\hat{f}_h(x)$ augmente.

4.2.2 Calcul de $MSE(\hat{f}_h(x))$

Le tableau suivant donne pour différentes valeurs de n , les valeurs de $MSE(\hat{f}_h(x))$ pour $x = 0.1, 0.5$ et 2 et $h = 0.1, 0.3$ et 0.5 en utilisant la formule (3.3).

| h | n | $MSE(\hat{f}_h(0.1))$ | $MSE(\hat{f}_h(0.5))$ | $MSE(\hat{f}_h(2))$ |
|-----|-----|-----------------------|-----------------------|---------------------|
| 0.1 | 10 | 0.422 | 0.154 | 0.036 |
| | 30 | 0.141 | 0.051 | 0.012 |
| | 50 | 0.084 | 0.031 | 0.007 |
| | 100 | 0.042 | 0.015 | 0.003 |
| | 200 | 0.021 | 0.008 | 0.001 |
| 0.3 | 10 | 0.248 | 0.091 | 0.021 |
| | 30 | 0.085 | 0.032 | 0.007 |
| | 50 | 0.053 | 0.020 | 0.004 |
| | 100 | 0.029 | 0.011 | 0.002 |
| | 200 | 0.017 | 0.007 | 0.001 |
| 0.5 | 10 | 0.201 | 0.076 | 0.016 |
| | 30 | 0.076 | 0.030 | 0.005 |
| | 50 | 0.051 | 0.021 | 0.003 |
| | 100 | 0.032 | 0.014 | 0.002 |
| | 200 | 0.022 | 0.010 | 0.001 |

Table 4.2. Valeurs de $MSE(\hat{f}_h(x))$

On remarque que pour x et h fixés, $MSE(\hat{f}_h(x))$ diminue quand n augmente.

Le graphe suivant donne pour h fixé, les variations de $MSE(\hat{f}_h(x))$ en fonction de x pour différentes valeurs de n .

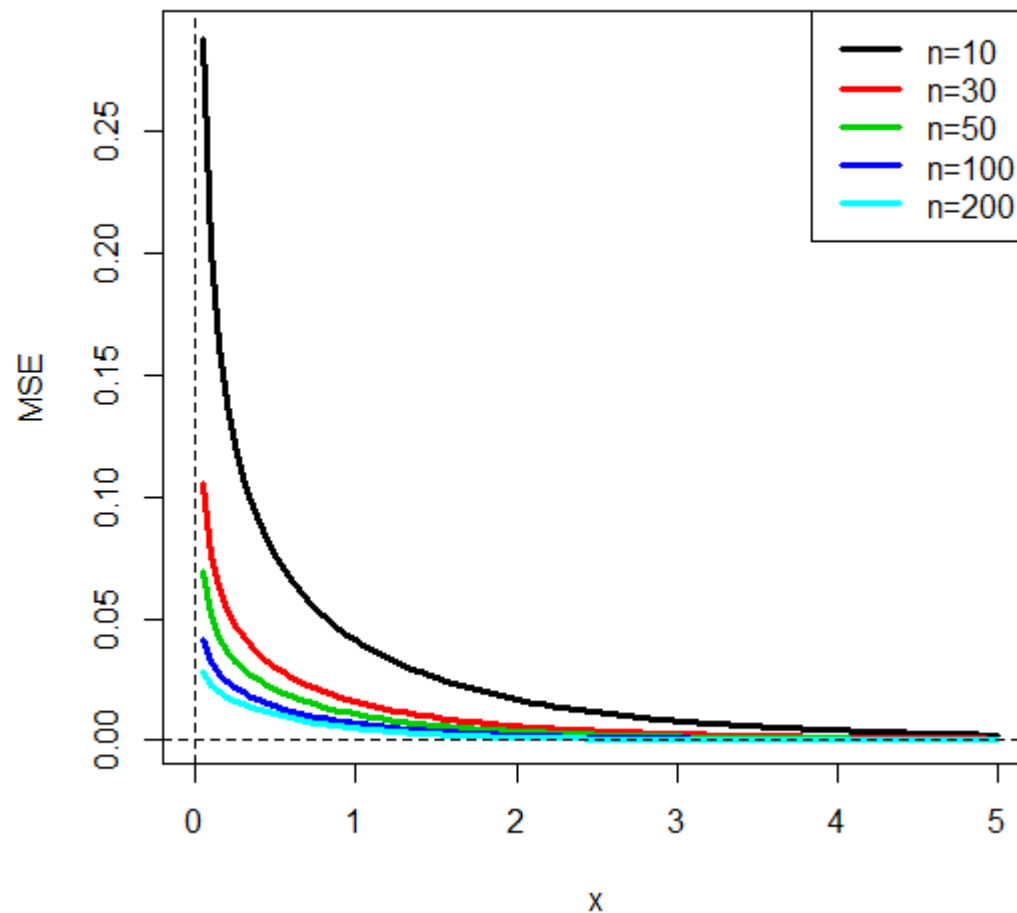


Figure 4.2. Courbes de $MSE(\hat{f}_h(x))$ en fonction de x

On voit que pour h fixé, $MSE(\hat{f}_h(x))$ diminue quand n augmente pour n'importe quelle valeur de x .

4.2.3 Calcul de $MISE(\hat{f}_h(x))$ et $MISE^*(\hat{f}_h(x))$

Le tableau suivant donne pour différentes valeurs de n , les valeurs de $MISE(\hat{f}_h(x))$ pour $h = 0.1, 0.3$ et 0.5 en utilisant la formule (3.5).

| h | n | $MISE(\hat{f}_h(x))$ |
|-----|-----|----------------------|
| 0.1 | 10 | 0.353 |
| | 30 | 0.118 |
| | 50 | 0.071 |
| | 100 | 0.036 |
| | 200 | 0.019 |
| 0.3 | 10 | 0.216 |
| | 30 | 0.081 |
| | 50 | 0.054 |
| | 100 | 0.034 |
| | 200 | 0.024 |
| 0.5 | 10 | 0.196 |
| | 30 | 0.091 |
| | 50 | 0.070 |
| | 100 | 0.054 |
| | 200 | 0.046 |

Table 4.3. Valeurs de $MISE(\hat{f}_h(x))$

On remarque que pour h fixé, $MISE(\hat{f}_h(x))$ diminue quand n augmente.

Les variations de $MISE(\hat{f}_h(x))$ en fonction de h sont représentées dans le graphe suivant

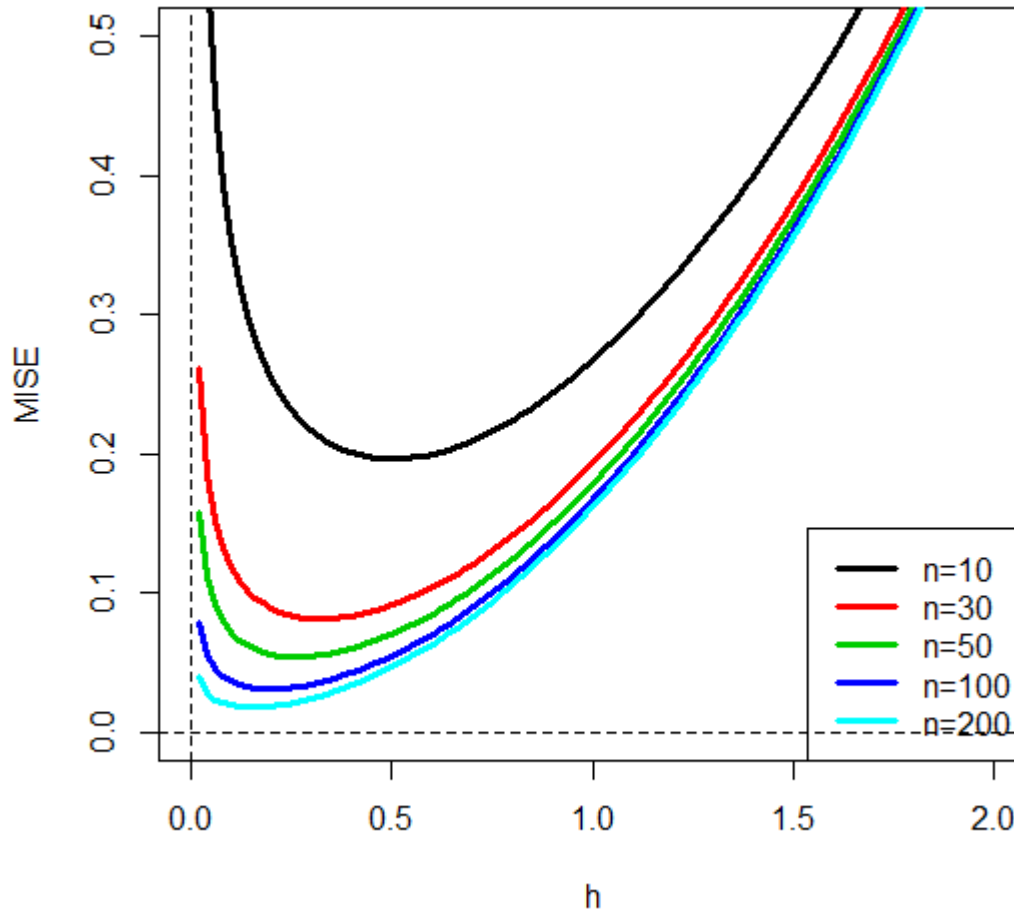


Figure 4.2. Courbes de $MISE(\hat{f}_h(x))$ en fonction de h

On voit que $MISE(\hat{f}_h(x))$ diminue quand n augmente pour n'importe quelle valeur de h .

En utilisant la formule (3.6), on calcule les valeurs de h^* et $MISE^*(\hat{f}_h(x))$ en fonction de n et on aura le tableau suivant.

| n | h^* | $MISE^*(\hat{f}_h(x))$ |
|-----|-------|------------------------|
| 10 | 1.050 | 0.280 |
| 30 | 0.677 | 0.116 |
| 50 | 0.551 | 0.077 |
| 100 | 0.418 | 0.044 |
| 200 | 0.316 | 0.025 |

Table 4.4. Valeurs de h^* et $MISE^*(\hat{f}_h(x))$

Les valeurs de h^* et $MISE^*(\hat{f}_h(x))$ diminuent quand n augmente.

CONCLUSION

Dans ce travail, nous avons étudié les propriétés asymptotiques telles que le biais, l'erreur quadratique moyenne et l'erreur quadratique moyenne intégrée de deux estimateurs à noyau gamma et bêta à supports dans $[0, +\infty[$ et $[0, 1]$ respectivement. Ensuite nous avons donné une application numérique de ces propriétés pour différentes valeurs de x , n et h .

Dans la première partie, nous avons fait un rappel de quelques notions essentielles comme le biais, l'erreur quadratique moyenne et l'erreur quadratique moyenne intégrée d'un estimateur.

Dans la partie 2, nous avons défini l'estimateur à noyau dans le deux cas, symétrique et asymétrique.

La partie 3 est consacrée à l'étude des propriétés asymptotiques telles que le biais, l'erreur quadratique moyenne et l'erreur quadratique moyenne intégrée des estimateurs à noyau asymétriques gamma et bêta.

Enfin, dans la dernière partie, nous avons calculé pour différentes valeurs de x , n et h les valeurs numériques du biais, le MSE, le MISE et MISE optimal. Les résultats montrent que :

Quand n augmente, le biais, le MSE, le MISE et le MISE optimal diminuent.

RÉFÉRENCES

- [1] Bouezmarni, T. & scaillet, O.(2004). consistency of asymmetric Kernel density Estimators and smoothed histograms with Application to income Data.Econometric theory, 21, 390-412.
- [2] Chen, S. X. (1999). Beta kernels estimators for density function. computational statistics and data analysis 31, 131-145.
- [3] Chen,S. X.(2000). Gamma kernels estimators for density function. Annals of the institute of statical mathematics 52,471-480.
- [4] Chen, S. X. (2000). Beta kernel smoothers for regression curves. Statistica Sinica, 10, pages 73-91.
- [5] Epanechnikov, V. A. (1969). Non parametric estimation of a multivariate probability density, theory of probability and its applications 14,153-158.
- [6] Nadaraya, E. A. (1965). On nonparametric estimation of density function and regression. Theory. Probab. Appl, 10, pages 186-190.
- [7] Parzen, E. (1962). On estimation of probability density function and mode.

Annals of mathematical statistics, vol. 33, no.3,p.pp.1065-1076.

[8] Rosenblat, M. (1956). Remarks on some non parametric estimates of a density function. Annals of mathematical statistics 27,832-837.

[9] Silvermann, B. W. (1986). Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

[10] Tsbakov, A. B.(2004). Introduction a l'estimation non paramétrique. Springer, Paris.

[11] Wand, M. P. & Jones, M. C. (1995). Kernel Smoothing. Chapman and Hall, London.

[12] Zougab, N. (2007). Etude comparative des méthodes de sélection du paramètre de lissage dans l'estimation de la densité de probabilité par la méthode du noyau. Mémoire de de Magister, Université de Béjaia.

[13] Zougab, N. (2013). Approche Bayésienne dans l'estimation non paramétrique de la densité de probabilité et la courbe de régression de la moyenne. Thèse de Doctorat en Mathématiques Appliquées, Université de Béjaia.

Résumé

L'objectif de ce travail est l'étude des propriétés asymptotiques telles que le biais, le MSE et le MISE de quelques estimateurs à noyaux asymétriques utilisés pour estimer une densité à support non borné.

Les résultats théoriques et l'application numérique ont montré que lorsque la taille de l'échantillon augmente, l'estimateur a de meilleures propriétés.

Mots clés : Estimateur à noyau, noyau gamma, noyau bêta, MSE, MISE.

Abstract

The objective of this work is the study of the asymptotic proprieties like the bias, the MSE and the MISE of some asymmetric kernel estimators used for estimation of density with non bounded supports.

The theoretical and the numeric application prove that when the size of the sample increase, the estimator have better proprieties.

Key words : Kernel estimator, gamma kernel, beta kernel, MSE, MISE.