

Carlos Manuel Rebelo Tenreiro da Cruz

Thèse présentée à la Faculdade de Ciências
e Tecnologia da Universidade de Coimbra, pour obtenir
le grade de Doutor em Matemática (Matemática Aplicada).

Estimation Fonctionnelle:
Applications aux Tests d'Adéquation
et de Paramètre Constant

Soutenue le 27 novembre 1995

devant le Jury composé de:

Président: A. S. Alves, Université de Coimbra
Rapporteurs: C. Gouriéroux, Directeur de Thèse, Université de Paris IX
M. Ivette Gomes, Université de Lisbonne
Membres: J. A. Fernandes de Carvalho, Université de Coimbra
N. Mendes Lopes, Université de Coimbra
E. Gonçalves, Université de Coimbra

Departamento de Matemática

Universidade de Coimbra

1995

A Laura

A mes parents et frères

Cette dissertation a été réalisée sous la direction de Monsieur Christian Gouriéroux. Sans son aide précieuse et ses nombreux commentaires et suggestions ce travail n'aurait sans doute pu être mené à bien. Je tiens à lui exprimer ma profonde reconnaissance.

Je remercie Madame Esmeralda Gonçalves, qui a toujours fait preuve d'une grande disponibilité et dont les commentaires ont permis d'améliorer la présentation de cette thèse.

Mes remerciements vont aussi à Madame Nazaré Mendes Lopes, sous la direction de qui j'ai commencé en 1987 l'étude des estimateurs non paramétriques de la densité, et à Messieurs Denis Bosq, Michel Delecroix et Alain Monfort, pour leurs suggestions.

Je remercie sincèrement Emmanuel Guerre, avec qui j'ai discuté de certaines idées qui sont développées dans ce travail. Son amitié m'a été précieuse.

Je remercie le Département de Mathématiques de l'Université de Coimbra et le Centre de Recherche en Économie et Statistique de Paris pour toutes les facilités dont j'ai pu profiter pendant la réalisation de cette thèse.

Ce travail a bénéficié des soutiens de la J.N.I.C.T. (Junta Nacional de Investigação Científica e Tecnológica), du Gouvernement Français, du C.M.U.C. (Centro de Matemática da Universidade de Coimbra) et de la Fondation Calouste Gulbenkian.

Table des Matières

Introduction	v
1 Estimation d'une fonction de moments conditionnels sous des conditions locales	1
1.1 Définition de l'estimateur $\phi_n(x)$. Quelques notations	1
1.2 Hypothèses principales	2
1.2.1 Indépendance asymptotique	2
1.2.2 Moments conditionnels croisés	5
1.2.3 Le cas Markovien	6
1.3 L'estimateur à noyau $\bar{g}_n(x)$	7
1.3.1 Comportement asymptotique de $cov(\bar{g}_n(x), \bar{g}_n(y))$	7
1.3.2 Convergence en moyenne quadratique de $\bar{g}_n(x)$	12
1.3.3 Normalité asymptotique de $(\bar{g}_n(x_j) - E[\bar{g}_n(x_j)]; j = 1, \dots, T)$	12
1.3.4 Normalité asymptotique de $(\bar{g}_n(x_j) - m(g, x_j)f(x_j); j = 1, \dots, T)$	19
1.4 Estimation d'une fonction de moments	21
1.4.1 Convergence en probabilité de $\phi_n(x)$	22
1.4.2 Normalité asymptotique de $(\phi_n(x_j) - \phi(x_j); j = 1, \dots, T)$	22
1.4.3 Quelques commentaires	24
1.5 Applications	25
1.5.1 Estimation de la densité	25
1.5.2 Estimation de la moyenne conditionnelle	26
1.5.3 Estimation de la variance conditionnelle	27

1.5.4	Estimation des coefficients d'asymétrie conditionnelle et de kurtosis conditionnelle	28
2	M-estimateurs à noyau: diagnostic non paramétrique de modèles structurels	31
2.1	Exemple introductif	31
2.2	M-estimateur à noyau: définition et propriétés asymptotiques	33
2.2.1	Définition	33
2.2.2	Convergence presque sûre	34
2.2.3	Normalité asymptotique	42
2.2.4	Conditions de normalité asymptotique dans le cas convexe	48
2.3	Diagnostic non paramétrique de modèles structurels	50
2.3.1	Régions ponctuelles de confiance asymptotique	51
2.3.2	M-estimateur local à noyau	53
2.3.3	Quelques exemples d'application	54
2.3.4	Une illustration	58
3	Théorèmes de limite centrale pour des U-statistiques dégénérées	67
3.1	Définitions et notations	67
3.2	Processus β -mélangeants. Quelques inégalités	68
3.3	Théorème de limite centrale pour des suites triangulaires	71
3.4	Théorèmes de limite centrale pour \mathcal{G}_n , \mathcal{H}_n et $\mathcal{G}_n + \mathcal{H}_n$	72
3.4.1	Loi asymptotique de \mathcal{G}_n	72
3.4.2	Loi asymptotique de \mathcal{H}_n	75
3.4.3	Loi asymptotique de $\mathcal{G}_n + \mathcal{H}_n$	82
4	Tests d'ajustement à une densité fondés sur un estimateur non paramétrique à noyau	85
4.1	Les écarts quadratiques I_n^1 et I_n^2	85
4.2	Développements asymptotiques de l'erreur quadratique moyenne intégrée	86
4.3	Quelques tests convergents	92
4.3.1	Le test fondé sur $T_n^{1,1}$	94
4.3.2	Le test fondé sur $T_n^{2,1}$	97

4.3.3	Comparaison des statistiques de test $T_n^{1,1}$ et $T_n^{2,1}$	102
4.3.4	Statistiques modifiées	102
4.4	Puissance locale des tests	104
4.4.1	Hypothèses sur la suite d'alternatives locales (AL)	105
4.4.2	Propriétés asymptotiques des écarts I_n^1 et I_n^2 pour une suite d'alternatives locales	106
4.4.3	Quelques remarques et définitions	112
4.4.4	Calculs de puissance pour les tests $C_n^{1,k}$, $k = 1, 2$	114
4.4.5	Calculs de puissance pour les tests $C_n^{2,k}$, $k = 1, 2$	118
4.5	Sur le choix du noyau	121
4.6	Applications des statistiques $T_n^{2,k}$, $k = 1, 2$, au cas d'un test paramétrique	122
4.6.1	Description des hypothèses	123
4.6.2	Test fondé sur une statistique multivariée à noyau	124
4.6.3	Calculs de puissance pour le test $C_n(\Pi^P)$	125
4.6.4	Choix optimal de Π^P	127

Introduction

Considérons une variable aléatoire réelle X admettant une densité f , qu'on suppose inconnue, par rapport à la mesure de Lebesgue. La loi de X étant caractérisée par f , un problème important est l'estimation de f à partir de l'observation de n copies $X_i, i = 1 \dots, n$ de X .

Supposons maintenant que nous voulons expliquer les valeurs prises par une nouvelle variable réelle Y à partir des valeurs prises par X . Ce problème est souvent modélisé par $Y = r(X) + U$, où $E[U|X] = 0$ et r est une application mesurable réelle. La fonction de régression $r(\cdot) = E[Y|X = \cdot]$, apportant de l'information sur la relation de dépendance inconnue de Y et X , un autre problème important est l'estimation de r à partir de l'observation de n copies $(X_i, Y_i), i = 1 \dots, n$ de la variable (X, Y) .

Pour chacun des problèmes précédents, nous pouvons concevoir deux approches possibles: une approche paramétrique, où la fonction à estimer (densité dans le premier cas et fonction de régression dans le deuxième) appartient à une famille paramétrique de courbes fixée au départ, et une approche non paramétrique, où nous ne restreignons pas la forme fonctionnelle des applications que nous voulons estimer.

Dans la démarche paramétrique, la forme fonctionnelle des courbes f et r est connue et seulement les paramètres sont inconnus; le problème d'estimation de f et r se réduit alors à un problème d'estimation des paramètres. Nous pouvons, par exemple imposer que la loi commune des X_1, \dots, X_n est normale, de moyenne m et de variance σ^2 inconnues, c'est-à-dire,

$$f(x) = f(x; m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), x \in \mathbb{R},$$

ou que les valeurs prises par Y sont une fonction affine des valeurs prises par X , c'est-à-dire, qu'il existe des réels a et b tels que

$$r(x) = r(x; a, b) = ax + b, x \in \mathbb{R}.$$

Les estimateurs des fonctions densité et régression sont, dans ces cas, donnés, pour $x \in \mathbb{R}$, par

$$f(x; \hat{m}_n, \hat{\sigma}_n^2) \text{ et } r(x; \hat{a}_n, \hat{b}_n),$$

où $\hat{m}_n, \hat{\sigma}_n^2, \hat{a}_n$ et \hat{b}_n sont, respectivement, des estimateurs des paramètres m, σ^2, a et b .

Dans la démarche non paramétrique nous n'admettons pas l'existence d'un paramètre appartenant à un espace vectoriel de dimension finie à partir duquel nous puissions estimer f et r . Ces fonctions sont alors estimées "directement". Les "paramètres" de chacun des problèmes d'estimation précédents sont ainsi les fonctions f et r elles mêmes, de sorte que nous disons dans ce cas que la densité et la fonction de régression sont des paramètres fonctionnels.

Une méthode d'estimation de la densité a été proposée et étudiée dans les importants travaux de Rosenblatt [83] et Parzen [75] qui datent, respectivement, de 1956 et 1962, et elle a été ensuite appliquée à l'estimation de la fonction de régression, indépendamment par Watson [109] et Nadaraya [70] en 1964. Nous parlons de la méthode du noyau de convolution.

Dès ces travaux pionniers, une vaste bibliographie sur l'estimation fonctionnelle, communément appelée estimation non paramétrique, est apparue. Malgré l'introduction d'autres méthodes d'estimation non paramétrique de la densité et de la régression, (comme l'estimation par l'histogramme et régressogramme, la méthode des points les plus proches ou la méthode des fonctions orthogonales), la plus grande partie de cette littérature porte sur les estimateurs à noyau.

Les propriétés asymptotiques de tels estimateurs, au niveau ponctuel ou au niveau global, suivant divers critères et modes de convergence stochastique, ont été obtenues dans une première étape pour des observations réelles, indépendantes et identiquement distribuées, puis généralisées au cadre de processus stochastiques multivariés fortement stationnaires satisfaisant des conditions de dépendance. Telles propriétés sont obtenues sous des conditions assez générales ce qui justifie l'utilisation de plus en plus fréquente de tels estimateurs en divers domaines de la statistique.

Les estimateurs non paramétriques de la densité et de la régression obtenus par la méthode du noyau, ainsi que des estimateurs d'autres paramètres fonctionnels conditionnels, dont les propriétés locales nous étudions dans le premier chapitre de ce travail, nous permettent une connaissance détaillée des lois de dimension finie du processus observé ou de certaines de ses caractéristiques. Par exemple, dans le cas particulier de l'estimation non paramétrique de la

densité on peut ainsi détecter la nature non gaussienne éventuelle du processus observé, ou dans le cas de l'estimation non paramétrique de la régression certaines caractéristiques non linéaires.

Reprenons l'exemple introductif, et admettons que nous voulons juger l'adéquation du modèle de régression considéré $Y = aX + b + U$, avec $E[U|X] = 0$, à la situation qu'il se propose décrire. L'importance de cette étape de diagnostic est évidente. Dans le cas où l'adéquation est confirmée, nous concluons que l'estimateur de la fonction de régression $r(x; \hat{a}_n, \hat{b}_n)$, qui prophète des bonnes propriétés des estimateurs paramétriques \hat{a}_n et \hat{b}_n , doit être préféré par rapport aux estimateurs non paramétriques, car il a, pour chaque x , une vitesse de convergence vers $E[Y|X = x]$ supérieure à celle de tels estimateurs. En effet, d'un point de vue théorique, la vitesse de convergence d'un estimateur fonctionnel est inférieure à la vitesse "paramétrique" \sqrt{n} . Ceci signifie que, les estimateurs non paramétriques demandent, pour atteindre la même précision que les estimateurs paramétriques, un nombre d'observations plus élevé. Dans le cas où l'adéquation du modèle paramétrique n'est pas confirmée, nous évitons les problèmes qui résultent de l'utilisation d'estimateurs paramétriques en modèles mal spécifiés.

Bien qu'un diagnostic sur la convenable spécification du modèle de régression considéré ci-dessus puisse être élaboré d'après la comparaison ponctuelle entre l'estimateur à noyau de la régression et l'estimateur "paramétrique" $r(x; \hat{a}_n, \hat{b}_n)$, nous verrons qu'il est préférable d'utiliser une procédure de diagnostic développée dans le deuxième chapitre. Une telle procédure est fondée sur une classe d'estimateurs fonctionnels appelés M-estimateurs à noyau, où "M" est relatif au fait que ces estimateurs sont obtenus par minimisation ou maximisation d'une fonction objectif.

L'obtention de régions ponctuelles asymptotiques de confiance pour la densité marginale du processus observé, à partir de la normalité asymptotique ponctuelle de l'estimateur à noyau, nous permet de formuler un diagnostic sur l'hypothèse simple $H_0 : f = f_0$, d'égalité de la densité des variables observées à une densité de probabilité f_0 fixée au départ ou, plus généralement, sur une hypothèse composite de la forme $H'_0 : f \in \{f_0(\cdot; \theta) : \theta \in \Theta \subset \mathbb{R}^k\}$, imposant à f d'appartenir à un ensemble de densités paramétré par θ et fixé au départ. Reprenant l'exemple initial, nous pouvons nous demander si la forme fonctionnelle prise pour f , que nous avons à priori supposée normale de moyenne et variance inconnues, est ou non adéquate. Cependant, le caractère local des résultats ne nous permet pas de formuler un test de niveau asymptotique α

et asymptotiquement convergent pour une telle hypothèse. Afin de contourner cette difficulté nous présentons dans le dernier chapitre de cette thèse, des tests d'ajustement fondés sur des mesures quadratiques de l'écart entre la densité de la vraie loi sous l'hypothèse nulle et un estimateur par noyau de cette fonction.

Les tests d'ajustement de Kolmogorov-Smirnov ou de Cramér-von Mises fondés sur la distribution limite de la fonction de répartition empirique sont usuellement préférés aux tests fondés sur la distribution limite des estimateurs non paramétriques de la densité. Cependant, nous pouvons citer quelques avantages de l'utilisation de tests d'ajustement de ce dernier type. En effet, dans un cadre de dépendance faible dans les variables observées ou dans un cadre multivarié, il n'existe pas de tests analogues fondés sur la fonction de répartition empirique. De plus, nous pouvons étendre notre approche au test de l'hypothèse composite H'_0 spécifiée ci-dessus en remplaçant, dans les statistiques de test que nous allons considérer et qui nous conduisent à des tests uniformément sans biais, le paramètre inconnu θ par un estimateur paramétrique convergeant $\hat{\theta}_n$, puisque les théorèmes limites obtenus sous l'hypothèse simple, restent valables dans ce cas (cf. Bickel et Rosenblatt [4]). Ceci n'est plus vrai pour les théorèmes limites associés aux statistiques de Kolmogorov-Smirnov ou de Cramér-von Mises.

La brève description que nous avons donnée de notre travail montre que nous nous intéressons à des problèmes d'estimation, de test et de validation de modèles structurels. Afin de préciser notre démarche d'analyse de ces divers problèmes, nous allons maintenant décrire en détail les sujets de cette thèse, chapitre par chapitre.

Dans le premier chapitre nous considérons un processus stochastique $(X_i, i \in \mathbb{Z})$ à valeurs dans \mathbb{R}^d , fortement stationnaire et α -mélangeant (cf. Rosenblatt [84]), et pour $i \in \mathbb{Z}$, nous notons V_i et W_i les vecteurs aléatoires $(X_{i+s}, \dots, X_{i+1})$ et $(X_{i-s_0}, X_{i-s_1}, \dots, X_{i-s_q})$, où les entiers positifs q, s, s_0, \dots, s_q , fixés au départ, satisfont $0 \equiv s_0 < s_1 < \dots < s_q$ et $s > 0$. On suppose que W_0 admet une densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}^{d'}$, avec $d' = d(q+1)$. Si $g = (g_1, \dots, g_p)$ et Φ sont des applications mesurables connues de $(\mathbb{R}^d)^s$ dans \mathbb{R}^p et de \mathbb{R}^p dans $\mathbb{R}^{p'}$, respectivement, nous nous intéressons aux estimateurs du paramètre fonctionnel $\phi(x) = \Phi(m(g, x)f(x))$, de la forme $\phi_n(x) = \Phi(\bar{g}_n(x))$, avec $m(g, x) = E[g(V_0)|W_0 = x]$ et

$$\bar{g}_n(x) = \left(\frac{1}{nh_1^{d'}(n)} \sum_{i=1}^n g_1(V_i) K_1\left(\frac{x - W_i}{h_1(n)}\right), \dots, \frac{1}{nh_p^{d'}(n)} \sum_{i=1}^n g_p(V_i) K_p\left(\frac{x - W_i}{h_p(n)}\right) \right),$$

où $x \in \mathbb{R}^{d'}$. Pour $r = 1, \dots, p$, $(h_r(n))$ est une suite de réels strictement positifs, tendant vers zéro lorsque n tend vers l'infini, et K_r est un noyau sur $\mathbb{R}^{d'}$, c'est-à-dire, une application intégrable de $\mathbb{R}^{d'}$ dans \mathbb{R} telle que $\int_{\mathbb{R}^{d'}} K_r(x) dx = 1$.

D'abord proposé par Rosenblatt [83] et étudié par Parzen [75], dans le cas de l'estimation non paramétrique de la densité dans un cadre d'échantillonnage, l'estimateur par noyau a ensuite été appliqué à l'estimation de la fonction de régression par Watson [109] et Nadaraya [71]. L'étude de l'estimateur de la densité, sous des conditions de dépendance entre variables observées, a été menée par Roussas [92] pour les processus de Markov satisfaisant la condition (D_0) , puis généralisée par Rosenblatt [85] à la classe des processus de Markov ρ -mélangeants (sur la condition (D_0) voir aussi Doob [27] et sur les processus de Markov mélangeants voir aussi Rosenblatt [86]). L'estimation non paramétrique de la régression pour les processus de Markov ρ -mélangeants est abordée par Yakowitz [111]. Les conditions de Markov et de ρ -mélangeance ont été affaiblies dans le cas réel par Robinson [81]. Dans un article plus récent Roussas [94] obtient, dans le cas multivarié, des résultats de convergence pour l'estimateur de la densité marginale du processus. D'autres mesures de dépendance sur le processus $(X_i, i \in \mathbb{Z})$ peuvent être trouvées dans les travaux de Castellana et Leadbetter [17] et dans le chapitre VII de Rosenblatt [90].

Les propriétés asymptotiques de l'estimateur à noyau $\bar{g}_n(\cdot)$ que nous obtenons dans 1.3, nous permettent de dériver la convergence en probabilité (ou en moyenne quadratique) de l'estimateur $\phi_n(x)$, pour $x \in \mathbb{R}^{d'}$ (Théorème 1.4.1), ainsi que la convergence des lois de dimension finie du processus empirique $(\phi_n(x), x \in \mathbb{R}^{d'})$ (Théorème 1.4.2). Ces résultats généralisent au cas multivarié ceux de Robinson [81], sur l'estimation non paramétrique de la densité de la loi de W_0 ($g = 1$, $\Phi(u) = u, u \in \mathbb{R}$), et de l'espérance conditionnelle $E[h(V_0)|W_0 = x]$ ($g = (1, h)$, $\Phi(u, v) = \frac{v}{u}, u \in \mathbb{R} \setminus \{0\}, v \in \mathbb{R}$), et nous permettent d'obtenir de façon simple des résultats de convergence pour d'autres paramètres fonctionnels conditionnels: rapports de moments conditionnels, variance conditionnelle, coefficient d'asymétrie conditionnel, kurtosis conditionnelle,....

Suivant la ligne de Robinson [81], et en utilisant une technique de démonstration classique (voir aussi Rosenblatt [85] et Bosq et Lecoutre [9]), ces résultats sont dérivés en imposant des hypothèses locales sur (des versions) des paramètres fonctionnels qu'on veut estimer, ainsi

que sur d'autres paramètres liés, au lieu de conditions globales comme celles qu'on trouve par exemple dans les travaux déjà cités de Yakowitz [111] et Roussas [94]. Pour cela, nous avons été conduits à utiliser dans la construction de l'estimateur à noyau, des noyaux à support borné. De plus, l'introduction d'hypothèses locales sur certains moments conditionnels croisés où intervient l'application g , nous permet d'obtenir les divers résultats avec des conditions sur la vitesse de convergence vers zéro des coefficients de mélange plus faibles que celles considérées dans Robinson [81], dans le cas où $\|g\|$ n'est pas bornée. Ces hypothèses sont discutés en 1.2.

Diverses applications sont proposées en 1.5. Nous considérons le problème classique de l'estimation de la densité de probabilité de W_0 qui contient comme cas particulier l'estimation de la densité marginale du processus $(X_i, i \in \mathbb{Z})$, puis celui de l'estimation de la moyenne conditionnelle qui inclut l'estimation de la régression. Dans ce dernier cas, en utilisant la forme de la précision asymptotique de l'estimateur, nous justifions la pratique courante consistant à retenir dans l'estimateur de la régression un seul noyau et une seule suite. Nous terminons en énonçant des résultats sur trois autres paramètres conditionnels: la variance conditionnelle, et les coefficients d'asymétrie et de kurtosis conditionnels. Si l'estimation de la variance conditionnelle peut permettre de détecter la présence éventuelle d'un effet hétéroscédastique dans la série des observations, la construction de régions ponctuelles de confiance asymptotiques pour les deux derniers paramètres, permet de tester une conséquence de l'hypothèse de normalité de la loi conditionnelle, en comparant respectivement à 0 et à 3 les coefficients estimés d'asymétrie et de kurtosis conditionnels. Une telle démarche peut être utilisée pour effectuer des recherches sur l'adéquation de l'hypothèse de normalité conditionnelle sur les erreurs ARCH(1) d'un processus autorégressif d'ordre un (cf. Gouriéroux [35] pg. 35-41).

Le deuxième chapitre de cette dissertation présente les résultats de Gouriéroux, Monfort et Tenreiro [37]. Nous étudions une classe d'estimateurs non paramétriques, usuellement désignés par M-estimateurs à noyau, et expliquons comment l'utiliser pour le diagnostic non paramétrique de modèles structurels. Ces estimateurs sont définis comme solution d'un problème d'optimisation du type

$$\hat{\theta}_n(s) = \text{Arg min}_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \psi(Y_i, \hat{\alpha}_n; s; \theta), s \in \mathbb{R}^d,$$

où S et ψ sont des applications connues à valeurs dans \mathbb{R}^d et \mathbb{R} , définissant la variable condi-

tionante et la fonction critère, respectivement, K est un noyau sur \mathbb{R}^d , $Y_i, i = 1, \dots, n$, sont les variables observées, et la suite de variables aléatoires $\widehat{\alpha}_n$ converge vers une constante α_∞ , lorsque n tend vers l'infini. Aucune condition de convexité en θ est imposée au départ sur ψ .

Les M-estimateurs à noyau ont été introduits pour des formes particulières de la fonction critère ψ . Robinson [82] considère des M-estimateurs de paramètres conditionnels de position pour lesquels $\psi = \rho(Y_i - \theta)$, où ρ est une fonction convexe (voir aussi Härdle [46]). Dans le cas particulier où $\psi = (Y_i - \theta)^2$, nous obtenons l'estimateur à noyau usuel de $E[Y|S = s]$. Quand $\psi = (Y_i - \theta_0 - \theta_1 S_i - \dots - \theta_p S_i^p)^2$, nous obtenons les coefficients de l'estimateur de la régression $E[Y|S = s]$ défini par $\widehat{\theta}_{0,n}(s) + \widehat{\theta}_{1,n}(s)s + \dots + \widehat{\theta}_{p,n}(s)s^p$ (cf. Cleveland [18]). En prenant $\psi = (g_1(Y_i) - \theta_1)^2 + \dots + (g_p(Y_i) - \theta_p)^2$, où $g = (g_1, \dots, g_p)$ est une fonction connue, nous obtenons l'estimateur à noyau usuel de $(E[g_1(Y)|S = s], \dots, E[g_p(Y)|S = s])$ que nous étudions en détail dans le premier chapitre.

La définition et l'application des M-estimateurs à noyau au diagnostic non paramétrique de modèles structurels sont motivées par un exemple avec lequel nous commençons le deuxième chapitre. Nous établissons ensuite les propriétés ponctuelles de convergence presque sûre (Théorème 2.2.6) et de normalité asymptotique de ces estimateurs (Théorème 2.2.8) qui peuvent être obtenues en utilisant la démarche usuelle développée pour les M-estimateurs paramétriques (cf. Wald [108], Jennrich [58], Huber [53], Burguete, Gallant et Souza [14], White [110], Gouriéroux et Monfort [36]), à partir des propriétés des estimateurs à noyau appliquées à la fonction objectif du problème d'optimisation défini ci-dessus, et à ses première et seconde dérivées par rapport à θ . Ces résultats sont obtenus pour une classe de processus fortement stationnaires et géométriquement α -mélangeants. Cependant, nous verrons que cette classe peut être nettement élargie dans le cas particulier où la fonction critère ψ est convexe en θ .

Dans la dernière partie du deuxième chapitre, nous introduisons une approximation de l'estimateur fonctionnel $\widehat{\theta}_n(\cdot)$ dans un voisinage de l'hypothèse $H_0 = \{\lim_n \widehat{\theta}_n(s) = 0, \forall s\}$, et expliquons comment nous pouvons l'utiliser dans la formulation de diagnostics non paramétriques de modèles structurels. Ce sont par exemple des diagnostics sur la forme paramétrique d'un modèle de régression, sur une conséquence de l'hypothèse d'adéquation de la densité conditionnelle à une famille paramétrique de courbes, sur l'oubli d'une variable dans une régression, ou sur la forme d'une hétéroscédasticité conditionnelle. En particulier, on montrera que dans la

construction de régions ponctuelles de confiance asymptotique pour le paramètre fonctionnel $\lim_n \widehat{\theta}_n(\cdot)$ sous H_0 , il est préférable de considérer l'estimateur

$$\widehat{\theta}_n(s) = \text{Arg min}_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \widehat{\alpha}_n)}{h_n}\right) \psi(Y_i, \widehat{\alpha}_n; S(Y_i, \widehat{\alpha}_n); \theta), s \in \mathbb{R}^d.$$

Cet estimateur satisfait $\lim_n \widehat{\theta}_n(s) = \lim_n \widehat{\theta}_n(s) \forall s$, et il a la même précision asymptotique que $\widehat{\theta}_n(\cdot)$. Cependant, le terme le plus significatif du biais asymptotique de $\widehat{\theta}_n(\cdot)$, caractéristique des estimateurs à noyau, est nul sous l'hypothèse nulle, ce qui permet d'attendre un meilleur comportement à distance finie de cet estimateur.

Une expérience sur données simulées est aussi proposée.

Le troisième chapitre contient des résultats techniques, notamment de théorèmes limites pour des formes quadratiques intervenant dans les développements des statistiques que nous introduisons au Chapitre 4. Nous obtenons ainsi des théorèmes de limite centrale pour des suites de variables aléatoires réelles du type

$$\mathcal{G}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n g_n(X_{in}),$$

$$\mathcal{H}_n = \frac{1}{n} \sum_{1 \leq j < i \leq n} \{h_n(X_{in}, X_{jn}) - E h_n(X_{in}, X_{jn})\} \text{ et}$$

$$(\mathcal{G}_n, \mathcal{H}_n),$$

(Théorèmes 3.4.5, 3.4.11 et 3.4.12) où $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$ est une suite de processus stochastiques à valeurs dans \mathbb{R}^d , fortement stationnaires, géométriquement absolument réguliers de façon uniforme par rapport à n (cf. Volkonskiï et Rozanov [106]). Pour $n \in \mathbb{N}^*$, $g_n(\cdot)$ et $h_n(\cdot, \cdot)$ sont des applications mesurables respectivement de \mathbb{R}^d et $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R} , et nous admettons que $E[g_n(X_{0n})] = 0$, $E[h_n(X_{0n}, y)] = 0$ et $h_n(x, y) = h_n(y, x)$, pour $x, y \in \mathbb{R}^d$. \mathcal{G}_n et \mathcal{H}_n sont donc des U-statistiques dégénérées d'ordres 1 et 2, respectivement.

La théorie des U-statistiques, d'abord développée par Hoeffding [51], a fait, dans les dernières années, l'objet de nombreux travaux de recherche. Divers résultats établissant la normalité asymptotique des U-statistiques dégénérées peuvent être trouvés dans la littérature. Nous pouvons citer, par exemple, les travaux de Hall [42], Jammalamadaka et Jason [57], de Jong [23] et O'Neil [74] dans un cadre d'échantillonnage, et de Khashimov [59] et Yoshihara [113] et [114] dans le cas dépendant. Le théorème de limite centrale que nous obtenons pour la U-statistique

dégénérée de second ordre \mathcal{H}_n , dont la technique de démonstration est basée en Takahata et Yoshihara [99], généralise au cas mélangeant le Théorème 1 de Hall [42]. L'ensemble de conditions suffisantes que nous obtenons pour la normalité, n'est sûrement pas le plus général, mais comme le souligne Hall à propos de son Théorème 1, il établit un compromis entre la généralité et la simplicité. Ce résultat, ainsi que ceux obtenus pour \mathcal{G}_n et $(\mathcal{G}_n, \mathcal{H}_n)$, sont fondés sur une inégalité de Yoshihara [113] pour les processus absolument réguliers et sur un théorème de limite centrale de Dvoretzky [29] pour des suites triangulaires. Nous commençons par rappeler ces résultats en 3.2 et 3.3.

Bien qu'utilisés dans cette thèse uniquement dans la détermination de la loi asymptotique de formes quadratiques basées sur l'estimateur à noyau de la densité (Chapitre 4), les résultats du Chapitre 3 peuvent être aussi appliqués à la détermination de la loi asymptotique de formes quadratiques basées sur l'estimateur à noyau de la fonction de régression (cf. Tenreiro [102], [103]).

Dans le quatrième chapitre de ce travail, nous nous intéressons au problème de test de l'hypothèse

$$H_0 : f = f_0,$$

au niveau asymptotique α , où f_0 est une densité fixée au départ, et f désigne la densité commune des variables aléatoires observées X_1, \dots, X_n .

Une méthode classique proposée par Karl Pearson en 1900 pour tester l'hypothèse qu'un n échantillon provient d'une population à densité fixée, s'appuie sur la comparaison de l'histogramme empirique basé sur une partition de \mathbb{R} avec l'histogramme de f_0 , en utilisant un test du qui-deux (voir Cochran [19]). Plus tard, les tests fondés sur la comparaison de la fonction de répartition empirique avec la fonction de répartition correspondante à f_0 sont apparus. Ce sont les cas des bien connus tests de Kolmogorov-Smirnov et de Crámer-von Mises (voir Anderson [1] et Milbrodt [66]). Plus récemment, en profitant de l'introduction des estimateurs à noyau de la densité de probabilité, Bickel et Rosenblatt [4] ont proposé des tests fondés sur la comparaison entre f_0 et cet estimateur. Cette même idée a été suivie par Eubank et LaRiccia [30] qui ont utilisé un autre estimateur de la densité.

Les résultats que nous présentons dans le Chapitre 4 sont des généralisations, au cas multivarié et dans un cadre d'indépendance asymptotique, des résultats obtenus par Bickel et

Rosenblatt [4]. Les tests que nous proposons sont fondés sur l'estimateur à noyau f_n de la densité marginale du processus des observations $(X_i, i \in \mathbb{Z})$, que nous supposons de dimension d , fortement stationnaire, et de loi marginale absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d . Un tel estimateur est défini par

$$f_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), x \in \mathbb{R}^d,$$

où (h_n) est une suite de nombres réels strictement positifs, convergeant vers zéro lorsque n tend vers l'infini, et K est un noyau sur \mathbb{R}^d .

Dans un premier temps, il est naturel de construire des statistiques de test à partir de l'écart entre l'estimateur par noyau f_n et sa moyenne sous l'hypothèse nulle, l'estimateur étant asymptotiquement sans biais, ou à partir de l'écart entre f_n et la vraie densité f_0 sous l'hypothèse nulle. De tels écarts peuvent être mesurés par $\int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\}^2 \pi(x) dx$ ou par $\int_{\mathbb{R}^d} \{f_n(x) - f_0(x)\}^2 \pi(x) dx$, où π est une fonction de poids et $E_0 f_n(x)$ désigne l'espérance de $f_n(x)$ sous H_0 (des tests fondés sur le maximum de l'écart normalisé de f_n à f_0 ont été aussi considérés dans [4]). Nous allons ainsi étudier les propriétés de procédures de test fondés sur les écarts

$$I_n^1 = \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\}^2 \pi(x) dx \text{ et } I_n^2 = \int_{\mathbb{R}^d} \{f_n(x) - f_0(x)\}^2 \pi(x) dx.$$

La loi limite de ces deux variables a été étudiée par divers auteurs. Dans le cas réel $d = 1$, Bickel et Rosenblatt [4] ont utilisé une technique fondée sur l'approximation forte de Brillinger, plus tard améliorée par Komlós, Major et Tusnády, pour la fonction de répartition empirique (pour les résultats et références voir Bickel et Rosenblatt [4] et Rosenblatt [89]). Une autre démarche qui se sert d'une technique de Poissonisation de la taille n de l'échantillon, a été utilisée dans le cas bivarié par Rosenblatt [88]. Ces procédures imposent cependant des conditions restrictives sur la densité f , et sur la suite (h_n) . Hall [42] montre que le problème précédent peut être aussi abordé dans le cadre des U-statistiques dégénérées. Il obtient les lois limites des statistiques I_n^1 et I_n^2 , avec $\pi = 1$, pour le cas multivarié, sans imposer de conditions restrictives ni sur f , ni sur (h_n) . Tous ces résultats ont été obtenus dans un cadre d'échantillonnage.

La construction de tests de niveau asymptotique α et asymptotiquement convergents pour tester $f = f_0$ contre $f \neq f_0$, qu'on présente en 4.3 (Théorèmes 4.3.3 et 4.3.10), est basée sur les

variables aléatoires

$$nh_n^{d/2} \{I_n^1 - E_0 I_n^1\} \text{ et } d(n) \{I_n^2 - E_0 I_n^2\},$$

où $(d(n))$ est une suite tendant vers l'infini, dépendant de la fenêtre h_n , de l'ordre du noyau et de l'ordre de dérivabilité de f_0 . La normalité asymptotique de ces variables a été obtenue par Takahata et Yoshihara [99] (avec $\pi = 1$) pour une sous-classe des processus fortement stationnaires et absolument réguliers. Nous présentons une preuve de leur résultat en employant des techniques de U-statistiques décrites au Chapitre 3. On remarque que les conditions imposées sur la suite (h_n) sont un peu plus restrictives que dans le cas d'échantillonnage. Dans le cas des variables dépendantes, comme les termes déterministes $E_0 I_n^1$ et $E_0 I_n^2$ ne peuvent pas être évalués avec la seule connaissance de f_0 , nous sommes conduits à introduire un développement asymptotique de l'erreur quadratique moyenne intégrée qui nous permettra de corriger les statistiques naturelles I_n^1 et I_n^2 de leurs effets de biais asymptotique, et nous conduira à différents tests asymptotiques. On considère en particulier ceux fondés sur les statistiques

$$T_n^{1,1} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f_0(x) \pi(x + uh_n) dx du \right\},$$

$$T_n^{1,2} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du \right\},$$

$$T_n^{2,1} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f_0(x) \pi(x + uh_n) dx du - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \right\}, \text{ et}$$

$$T_n^{2,2} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \right\}.$$

En suivant la démarche employée par Bickel et Rosenblatt [4], la comparaison des différentes procédures de test proposées, ainsi que l'évaluation de leurs puissances locales, est analysée en 4.4, en étudiant le comportement asymptotique de chacune des statistiques de test sous une suite de processus stochastiques fortement stationnaires, dont la densité marginale g_n , est de la forme (suite d'alternatives locales pour f_0)

$$g_n(x) = f_0(x) + a_n \gamma(x) + o(a_n) \gamma_n(x),$$

pour $x \in \mathbb{R}^d$, où (a_n) est une suite de réels positifs convergeant vers zéro, lorsque n tend vers l'infini. Dans un cadre paramétrique, on s'intéresse au cas $a_n = \frac{1}{\sqrt{n}}$. Cependant, les tests que nous étudions sont fondés sur des estimateurs non paramétriques pour lesquels la

vitesse paramétrique \sqrt{n} ne peut pas être obtenue. Par rapport à une hypothèse alternative si vaste, nous trouvons, pour les tests asymptotiquement convergents et strictement sans biais (cf. Bickel et Rosenblatt [4] pg. 1082) fondés sur les statistiques proposées, une vitesse de séparation maximale entre $H_{1n} : f = g_n$ et H_0 égale à $\frac{1}{\sqrt{n} h_n^{d/4}}$, qui est inférieure à la vitesse usuelle obtenue dans les tests paramétriques.

La description du comportement asymptotique des écarts I_n^1 et I_n^2 sous une telle suite d'alternatives locales (Théorèmes 4.4.4 et 4.4.7), nous permettra de détecter distinctes propriétés de puissance des tests fondés sur I_n^1 et sur I_n^2 . De plus, on conclut que les corrections asymptotiquement équivalentes sous l'hypothèse nulle, introduites pour les termes de biais $E_0 I_n^1$ et $E_0 I_n^2$, ne le sont déjà plus sous la suite d'alternatives locales considérée. Ceci justifie les différentes propriétés de puissance trouvées pour les tests fondés sur $T_n^{1,1}$ et $T_n^{1,2}$, et sur $T_n^{2,1}$ et $T_n^{2,2}$. Uniformément sur les alternatives locales considérées, nous concluons que les tests fondés sur $T_n^{1,1}$, $T_n^{2,1}$ et $T_n^{2,2}$, n'atteignent la puissance locale du test fondé sur $T_n^{1,2}$ que pour des choix particuliers de la suite (h_n) .

Finalement, en tenant en compte que les tests fondés sur I_n^2 peuvent atteindre une puissance asymptotique non triviale pour certaines alternatives locales avec $a_n = \frac{1}{\sqrt{n}}$, nous discutons en 4.6 et nous plaçant dans un cadre d'échantillonnage, la construction d'un test asymptotiquement convergent fondé sur une statistique multivariée de tels écarts quadratiques évaluées pour des différentes pondérations π , pour tester l'hypothèse simple $H_0 : f = g(\cdot; \theta_0)$, où θ_0 est à l'intérieur de $\Theta \subset \mathbb{R}^K$, contre une hypothèse alternative paramétrique $H_0^c : f \in \{g(\cdot; \theta) \mid \theta \in \Theta \setminus \{\theta_0\}\}$, en obtenant la puissance paramétrique. Nous montrerons encore qu'un choix adéquat des différentes fonctions de poids permettra d'obtenir un test asymptotique strictement sans biais pour tester H_0 contre H_0^c , dont la puissance asymptotique est identique à celle du test du rapport de vraisemblance.

Après avoir terminé la rédaction de cette thèse, nous avons pris connaissance du travail de Y. Fan [32] publié pendant l'année 1994, où l'auteur étudie des tests d'ajustement d'une hypothèse composite, fondés sur les statistiques I_n^1 et I_n^2 avec $\pi = 1$, dans un cadre d'échantillonnage. Malgré les naturels points communs avec l'étude que nous avons développée dans le quatrième chapitre de cette thèse, nous ne voulons pas terminer la présente introduction sans une référence à quelques points où les abordages développés dans les deux travaux ne coïncident pas.

Le premier de ces points concerne l'hypothèse d'indépendance entre les variables observées qui a été considérée par Fan [32]. Une telle hypothèse permet à l'auteur d'utiliser, dans l'étude du comportement asymptotique des statistiques de test, les théorèmes limites sur les U-statistiques dégénérées obtenus par Hall [42]. En admettant que les variables aléatoires observées peuvent satisfaire à une structure de dépendance, nous avons eu besoin, dans une première étape, d'étendre tels théorèmes limites au cas de dépendance considéré. Deuxièmement, notons que Fan [32] développe son étude dans le cas général d'une hypothèse nulle composite, n'ayant pas cette démarche été suivie dans cette thèse. En effet, ayant considéré comme principal objectif de l'étude que nous avons développée, l'obtention de tests uniformément strictement sans biais, et sachant (cf. Bickel et Rosenblatt [4]) que tels tests peuvent s'obtenir, dans le cas d'une hypothèse nulle composite, à partir des tests correspondants pour une hypothèse nulle simple, comme nous l'avons déjà souligné, nous avons uniquement considéré l'étude de ce dernier cas. Finalement, notons que le fait d'avoir considéré, dans cette thèse, de tests fondés sur des écarts quadratiques pondérés, si bien que, apparemment, sans importance dans le test d'une hypothèse simple contre une alternative non paramétrique, il se révèle intéressant dans le test d'une hypothèse simple contre une hypothèse alternative paramétrique.

Chapitre 1

Estimation d'une fonction de moments conditionnels sous des conditions locales

Dans ce chapitre, nous établissons, sous des conditions locales, la convergence ponctuelle et la normalité asymptotique d'une classe d'estimateurs par noyau fondés sur un processus strictement stationnaire et mélangeant. Cette classe contient les estimateurs par noyau usuels de la densité et de l'espérance conditionnelle, ainsi que ceux de toute fonction suffisamment régulière de telles quantités: variance conditionnelle, coefficient d'asymétrie conditionnel, kurtosis conditionnelle,....

1.1 Définition de l'estimateur $\phi_n(x)$. Quelques notations

Soit $(X_i, i \in \mathbb{Z})$ un processus stochastique à valeurs dans \mathbb{R}^d , fortement stationnaire. Considérons pour $i \in \mathbb{Z}$,

$$V_i = (X_{i+s}, \dots, X_{i+1}) \text{ et } W_i = (X_{i-s_0}, X_{i-s_1}, \dots, X_{i-s_q}), \quad (1.1.1)$$

où $q \in \mathbb{N}$ et $s, s_0, \dots, s_q \in \mathbb{N}$ sont fixés au départ, avec

$$0 \equiv s_0 < s_1 < \dots < s_q \text{ et } s > 0.$$

On suppose que W_0 admet une densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}^{d'}$, avec $d' = d(q+1)$. Si $g = (g_1, \dots, g_p)$ et Φ sont des applications mesurables connues de $(\mathbb{R}^d)^s$

dans \mathbb{R}^p et de \mathbb{R}^p dans $\mathbb{R}^{p'}$, respectivement, nous nous intéressons aux estimateurs du paramètre fonctionnel $\phi(x) = \Phi(m(g, x)f(x))$ de la forme $\phi_n(x) = \Phi(\bar{g}_n(x))$, où $m(g, x)$ désigne l'espérance conditionnelle $E[g(V_0)|W_0 = x]$, et

$$\bar{g}_n(x) = \left(\frac{1}{nh_1^{d'}(n)} \sum_{i=1}^n g_1(V_i) K_1 \left(\frac{x - W_i}{h_1(n)} \right), \dots, \frac{1}{nh_p^{d'}(n)} \sum_{i=1}^n g_p(V_i) K_p \left(\frac{x - W_i}{h_p(n)} \right) \right), \quad (1.1.2)$$

où $x \in \mathbb{R}^{d'}$. Pour $r = 1, \dots, p$, $(h_r(n))$ est une suite de réels strictement positifs, tendant vers zéro lorsque n tend vers l'infini, et K_r est un noyau sur $\mathbb{R}^{d'}$, c'est-à-dire, une application de $\mathbb{R}^{d'}$ dans \mathbb{R} intégrable telle que $\int_{\mathbb{R}^{d'}} K_r(x) dx = 1$.

Les estimateurs à noyau précédents dont nous étudions les propriétés de convergence en probabilité (ou en moyenne quadratique) et de convergence des lois de dimension finie du processus empirique associé, apparaissent de façon naturelle dans l'estimation non paramétrique de la densité de probabilité de la loi de W_0 ($g = 1$, $\Phi(u) = u$, $u \in \mathbb{R}$), dans celle de l'espérance conditionnelle $E[h(V_0)|W_0 = x]$ ($g = (1, h)$, $\Phi(u, v) = \frac{v}{u}$, $u \in \mathbb{R} \setminus \{0\}$, $v \in \mathbb{R}$), et aussi dans celle de toute fonction suffisamment régulière de telles quantités comme, par exemple, la variance conditionnelle, et les coefficients d'asymétrie et de kurtosis conditionnels.

Par la suite, nous désignons par: \mathcal{B}_x l'ensemble des applications bornées dans un voisinage de x ; \mathcal{C}_x l'ensemble des applications continues en x ; $\mathcal{A}_x(\alpha)$ ($\alpha \in]0, 1]$) l'ensemble des applications α -lipschitziennes dans un voisinage de x ; $\mathcal{D}_x(m)$ ($m \in \mathbb{N}^*$) l'ensemble des applications dont les composantes admettent des dérivées partielles d'ordre m continues dans un voisinage de x ; $\mathcal{B}_x(M, \delta)$ l'ensemble des applications bornées par $M > 0$ dans un voisinage $\delta > 0$ de x .

1.2 Hypothèses principales

En plus de la stationnarité forte, nous imposons diverses hypothèses de régularité sur le processus $(X_i, i \in \mathbb{Z})$. Celles-ci pourront concerner son comportement asymptotique ou la régularité de ses moments conditionnels.

1.2.1 Indépendance asymptotique

Soit F_n^m , $n, m \in \mathbb{Z} \cup \{-\infty, +\infty\}$ la tribu engendrée par les variables $\{X_i, n \leq i \leq m\}$. La dépendance entre le passé $F_{-\infty}^0$ et le futur $F_i^{+\infty}$ peut être mesurée par le nombre (cf. Rosenblatt [84])

$$\alpha(i) = \sup_{\substack{A \in \mathcal{F}_{-\infty}^0 \\ B \in \mathcal{F}_i^{+\infty}}} |P(A \cap B) - P(A)P(B)|. \quad (1.2.1)$$

Définition 1.2.2 On dit que le processus $(X_i, i \in \mathbb{Z})$ est de type *Alpha*(γ), avec $\gamma \in]0, 1]$, s'il existe $r \in]0, \frac{\gamma}{1+\gamma}[$ tel que

$$\sum_{i=1}^{+\infty} \alpha^r(i) < \infty.$$

Un processus de type *Alpha*(γ) pour un $\gamma \in]0, 1]$, est évidemment α -mélangeant (ou fortement mélangeant), c'est-à-dire, $\alpha(i) \rightarrow 0$, si $i \rightarrow +\infty$. La condition *Alpha*(γ) est impliquée par d'autres conditions sur la vitesse à laquelle $\alpha(i)$ tend vers zéro. Elle l'est, par exemple, si: $\alpha(i) = O(\rho^i)$ avec $0 < \rho < 1$, cas de certains processus ARMA multivariés ou bilinéaires réels fondés sur des bruits blancs i.i.d. (cf. Pham et Tran [76] et Mokkadem [67]) et de certains processus autorégressifs d'ordre 1, avec des erreurs qui admettent des représentations ARCH, TARARCH ou β -ARCH (cf. Diebolt et Guégan [26]); si $\alpha(i) = O(i^r \rho^i)$ avec $0 < \rho < 1$ et $r > 0$, cas des processus autorégressifs d'ordre 1, fondés sur un bruit blanc i.i.d. de Cauchy (cf. Gastwirth et Rubin [33]); si $\alpha(i) = O(\frac{1}{i^\epsilon})$ avec $\epsilon > 1 + \frac{1}{\gamma}$, cas de certains processus gaussiens dont la covariance satisfait $cov(X_0, X_i) = O(|i|^{-a})$, avec $a > 2 + \frac{1}{\gamma}$ (cf. Doukhan [28] pg. 59); ou si $\alpha(i) = 0$ pour $i > m$, cas de processus qui admettent une représentation moyenne mobile d'ordre m , ou plus généralement, de processus m -dépendants. Par contre, si $\alpha(i) = \frac{C}{i^\epsilon}$, avec $C > 0$ et $\epsilon \leq 1 + \frac{1}{\gamma}$, la condition *Alpha*(γ) n'est pas satisfaite.

La proposition suivante établit la vitesse de convergence vers zéro du reste d'ordre n de la série $\sum_{i=1}^{\infty} \alpha^t(i)$, pour $t > \frac{\gamma}{1+\gamma}$, où $\alpha(i)$ est le coefficient de mélange d'ordre i d'un processus de type *Alpha*(γ).

Proposition 1.2.3 Soit $\gamma \in]0, 1]$, fixé. Les propositions suivantes sont équivalentes:

- i) $(X_i, i \in \mathbb{Z})$ est de type *Alpha*(γ);
- ii) étant donné $s, t > 0$ avec $\frac{t}{1+s} = \frac{\gamma}{1+\gamma}$, il existe $\eta > 0$ tel que

$$n^{s+\eta} \sum_{i=n}^{+\infty} \alpha^t(i) \rightarrow 0, \text{ lorsque } n \rightarrow +\infty,$$

où $\alpha(i)$ est défini par (1.2.1).

Démonstration: Soit $\gamma \in]0, 1]$, fixé. Soient $s, t > 0$ tels que $\frac{t}{1+s} = \frac{\gamma}{1+\gamma}$, et $r \in]0, \frac{\gamma}{1+\gamma}[$ tel que $\sum_{i=1}^{+\infty} \alpha^r(i) < \infty$. On a $i\alpha^r(i) \rightarrow 0, i \rightarrow +\infty$, puisque $(\alpha(i))$ est décroissante. Soit alors $\alpha(i) = \frac{\epsilon_i}{i^{1/r}}$, où $\epsilon_i \rightarrow 0, i \rightarrow +\infty$. Notons $\xi_i = \sup_{j \geq i} \epsilon_j^t$, pour $i \in \mathbb{N}^*$. La condition ii) est alors satisfaite avec $\eta = t\left(\frac{1}{r} - \frac{1+\gamma}{\gamma}\right) > 0$. En effet,

$$\begin{aligned} n^{s+\eta} \sum_{i=n}^{+\infty} \alpha^t(i) &= n^{s+\eta} \sum_{i=n}^{+\infty} \frac{\epsilon_i^t}{i^{t/r}} \\ &\leq \xi_n n^{s+\eta} \sum_{i=n}^{+\infty} \frac{1}{i^{t/r}} \\ &= O(\xi_n) = o(1). \end{aligned}$$

Inversement, soit $\eta > 0$ tel que pour $s = \frac{1+\gamma}{\gamma} - 1$, on a $n^{s+\eta} \sum_{i=n}^{+\infty} \alpha(i) \rightarrow 0, n \rightarrow +\infty$. Si $m > n$, comme la suite $(\alpha(i))$ est décroissante, on a

$$\sum_{i=n}^{+\infty} \alpha(i) - \sum_{i=m}^{+\infty} \alpha(i) = \sum_{i=n}^{m-1} \alpha(i) \geq (m-n)\alpha(m).$$

En faisant $m = 2n$, on a, d'après les expressions précédentes

$$\begin{aligned} (2n)^{1+s+\eta} \alpha(2n) &\leq 2(2n)^{s+\eta} \left(\sum_{i=n}^{+\infty} \alpha(i) - \sum_{i=2n}^{+\infty} \alpha(i) \right) \\ &= 2^{1+s+\eta} n^{s+\eta} \sum_{i=n}^{+\infty} \alpha(i) - 2(2n)^{s+\eta} \sum_{i=2n}^{+\infty} \alpha(i), \end{aligned}$$

et donc $(2n)^{1+s+\eta} \alpha(2n) \rightarrow 0, n \rightarrow +\infty$. De façon analogue, la même conclusion est valable pour $m = 2n + 1$, ce qui permet de conclure que $n^{\frac{1+\gamma}{\gamma}+\eta} \alpha(n) \rightarrow 0, n \rightarrow +\infty$. Le processus $(X_i, i \in \mathbb{Z})$ est ainsi de type *Alpha*(γ). ■

D'autres conditions de dépendance sur le processus peuvent être considérées. Les plus usuelles concernent la vitesse de convergence vers zéro, lorsque i tend vers l'infini, des coefficients $\rho(i) = \sup_{\substack{f \in L^2(F_{-\infty}^0) \\ g \in L^2(F_i^{+\infty})}} \frac{\text{cov}(f,g)}{(\text{Var}(f))^{1/2}(\text{Var}(g))^{1/2}}$ ou $\phi(i) = \sup_{\substack{A \in F_{-\infty}^0, P(A) > 0 \\ B \in F_i^{+\infty}}} |P(B|A) - P(B)|$ (cf. Bradley [12]). Dans ce cas nous disons que $(X_i, i \in \mathbb{Z})$ est ρ ou ϕ -mélangeant, respectivement. Des rapports entre ces divers types de mélange sont déduits des inégalités $4\alpha(i) \leq \rho(i) \leq 2\sqrt{\phi(i)}$, pour $i \in \mathbb{N}^*$. On rappelle encore que dans le cas d'un processus gaussien, les coefficients $\alpha(i)$ et $\rho(i)$ sont équivalents (cf. Kolmogorov et Rozanov [60]). Nous discuterons en §1.4.3 ce que deviennent les résultats des sections suivantes lorsque ces autres notions de dépendance sont utilisées.

1.2.2 Moments conditionnels croisés

Etant donnés x_1, \dots, x_T , T points de $\mathbb{R}^{d'}$, notons par $\mathcal{N}_p(x_1, \dots, x_T)$, avec $p \in \mathbb{N}^*$, l'ensemble des applications g de $(\mathbb{R}^d)^s$ dans \mathbb{R}^p , pour lesquelles il existe $M > 0$ et $\delta > 0$ tels que

$$E[\|g(V_i)\| \|g(V_0)\| \mid X_i = \cdot, W_0 = \cdot] f_{(X_i, W_0)}(\cdot, \cdot) \in \bigcap_{a,b=1}^T \mathcal{B}_{(\pi(x_a), x_b)}, \text{ si } 1 \leq i \leq s_q,$$

$$E[\|g(V_i)\| \|g(V_0)\| \mid W_i = \cdot, W_0 = \cdot] f_{(W_i, W_0)}(\cdot, \cdot) \in \bigcap_{a,b=1}^T \mathcal{B}_{(x_a, x_b)}(M, \delta), \text{ si } i > s_q,$$

où $\|\cdot\|$ est une norme quelconque dans \mathbb{R}^d , $\pi(x) = x_0$, si $x = (x_0, x_1, \dots, x_q) \in (\mathbb{R}^d)^{q+1}$, et $f_{(W_i, W_0)}$ ($i > s_q$) et $f_{(X_i, W_0)}$ ($1 \leq i \leq s_q$) désignent respectivement les densités de probabilité des vecteurs (W_i, W_0) et (X_i, W_0) , dont on admet l'existence. De plus, notons $\mathcal{P}_p(x)$, pour $p \in \mathbb{N}^*$ et $x \in \mathbb{R}^{d'}$, l'ensemble des applications $g = (g_1, \dots, g_p)$ de $(\mathbb{R}^d)^s$ dans \mathbb{R}^p , telles que $g_r \in \mathcal{N}_1(x)$, pour $r = 1, \dots, p$.

Les deux ensembles d'indices $i > s_q$ et $1 \leq i \leq s_q$ considérés ci-dessus, correspondent respectivement à la situation où il n'y a pas de variables communes à W_i et W_0 , et celle où on ne peut pas toujours admettre l'existence de la densité du couple (W_i, W_0) . Ce dernier cas peut être illustré lorsqu'on estime $E[X_1 \mid X_0 = \cdot, X_{-2} = \cdot]$ où $(X_i, i \in \mathbb{Z})$ est un processus réel. On a $W_i = (X_i, X_{i-2})$, et pour $i = s_q = 2$ la densité du vecteur (W_i, W_0) n'existe pas.

La condition précédente sur les moments conditionnels croisés apparaît comme une extension de la condition, sur les densités jointes, utilisée dans Robinson [81] pg. 191, ou de celle considérée dans Rosenblatt [90] pg. 199. Lorsqu'il existe $R > 0$ tel que $P(\|g(V_0)\| < R) = 1$, ce qui est en particulier vrai si g est bornée ou si f_{X_0} est à support borné et g est continue, la condition $g \in \mathcal{N}_p(x_1, \dots, x_T)$ se réduit en effet à $1 \in \mathcal{N}_p(x_1, \dots, x_T)$, c'est-à-dire, il existe $M > 0$ et $\delta > 0$ tels que

$$f_{(X_i, W_0)} \in \bigcap_{a,b=1}^T \mathcal{B}_{(\pi(x_a), x_b)}, \text{ si } 1 \leq i \leq s_q,$$

$$f_{(W_i, W_0)} \in \bigcap_{a,b=1}^T \mathcal{B}_{(x_a, x_b)}(M, \delta), \text{ si } i > s_q.$$

La restriction la plus importante introduite, soit par la condition $1 \in \mathcal{N}_p(x_1, \dots, x_T)$, soit par $g \in \mathcal{N}_p(x_1, \dots, x_T)$, n'est pas le fait de demander que les applications qui y sont introduites soient bornées dans un voisinage des points signalés, mais surtout que cela soit uniforme en i . Dans le cas des processus m -dépendants et tenant compte de l'indépendance entre les variables W_i et W_0 pour $i > s_q + m$, la condition d'uniformité introduite par $1 \in \mathcal{N}_p(x_1, \dots, x_T)$ est

satisfaite si f_{W_0} est bornée dans un voisinage des points x_1, \dots, x_T , tandis que celle introduite par $g \in \mathcal{N}_p(x_1, \dots, x_T)$ est satisfaite si $m(\|g\|^2, \cdot)$ et f_{W_0} sont bornées dans un voisinage des points x_1, \dots, x_T , car pour $i > s_q + s + m$ et $u, v \in \mathbb{R}^{d'}$, on a

$$\begin{aligned} & E[\|g(V_i)\| \|g(V_0)\| \mid W_i = v, W_0 = u] f_{(W_i, W_0)}(v, u) \\ & \leq E^{1/2}[\|g(V_i)\|^2 \mid W_i = v, W_0 = u] E^{1/2}[\|g(V_0)\|^2 \mid W_i = v, W_0 = u] f_{(W_i, W_0)}(v, u) \\ & = (m(\|g\|^2, v) m(\|g\|^2, u))^{1/2} f_{W_0}(v) f_{W_0}(u). \end{aligned}$$

Lorsque que les densités de probabilité $f_{(V_i, W_i, V_0, W_0)}$ des vecteurs (V_i, W_i, V_0, W_0) existent pour $i > s + s_q$, ce qui se vérifie par exemple, dans le cas des processus gaussiens, on remarque que cette condition d'uniformité porte sur les densités jointes des vecteurs (V_i, W_i) et (V_0, W_0) pour i assez grand, car pour $u, v \in \mathbb{R}^{d'}$, on a

$$\begin{aligned} & E[\|g(V_i)\| \|g(V_0)\| \mid W_i = v, W_0 = u] f_{(W_i, W_0)}(v, u) \\ & = \int_{(\mathbb{R}^d)^s} \int_{(\mathbb{R}^d)^s} \|g(x)\| \|g(y)\| f_{(V_i, W_i, V_0, W_0)}(x, v, y, u) dx dy. \end{aligned}$$

1.2.3 Le cas Markovien

Si $(X_i, i \in \mathbb{Z})$ est un processus de Markov d'ordre 1, et si nous nous intéressons à l'estimation de $\phi(x) = \Phi(E[g(X_1) \mid X_0 = x] f(x))$, nous discutons maintenant les hypothèses $1 \in \mathcal{N}_p(x_1, \dots, x_T)$ et $g \in \mathcal{N}_p(x_1, \dots, x_T)$, où x_1, \dots, x_T sont des points dans \mathbb{R}^d . La condition $1 \in \mathcal{N}_p(x_1, \dots, x_T)$, qui se réduit à: il existe $M > 0$ et $\delta > 0$ tels que $f_{(X_i, X_0)}(\cdot, \cdot) \in \bigcap_{a, b=1}^T \mathcal{B}_{(x_a, x_b)}(M, \delta)$, si $i \geq 1$, est satisfaite si $f_{X_0}(\cdot)$ et $\sup_{x \in \mathbb{R}^d} f_{X_1 \mid X_0=x}(\cdot)$ sont bornées dans un voisinage des points x_1, \dots, x_T , où $f_{X_1 \mid X_0=x}(\cdot)$ désigne la densité de la loi X_1 conditionnée par $X_0 = x$. En effet, nous avons pour $i \geq 1$ et $u, v \in \mathbb{R}^d$,

$$\begin{aligned} f_{(X_i, X_0)}(v, u) &= f_{X_0}(u) \int_{\mathbb{R}^d} f_{X_i \mid X_{i-1}=x}(v) f_{X_{i-1} \mid X_0=u}(x) dx \\ &\leq f_{X_0}(u) \sup_{x \in \mathbb{R}^d} f_{X_1 \mid X_0=x}(v). \end{aligned}$$

La condition $g \in \mathcal{N}_p(x_1, \dots, x_T)$, qui se réduit à: il existe $M > 0$ et $\delta > 0$ tels que $E[\|g(X_{i+1})\| \|g(X_1)\| \mid X_i = \cdot, X_0 = \cdot] f_{(X_i, X_0)}(\cdot, \cdot) \in \bigcap_{a, b=1}^T \mathcal{B}_{(x_a, x_b)}(M, \delta)$, si $i \geq 1$, est satisfaite si les applications $\|g\|$, $E[\|g(X_1)\| \mid X_0 = \cdot]$, $f_{X_0}(\cdot)$ et $\sup_{x \in \mathbb{R}^d} f_{X_1 \mid X_0=x}(\cdot)$ sont bornées dans un voisinage des points x_1, \dots, x_T . En effet, pour $i = 1$ et $u, v \in \mathbb{R}^d$, il suffit de remarquer que

$E[\|g(X_{i+1})\|g(X_1)\| \mid X_i = v, X_0 = u] = \|g(v)\|E[\|g(X_1)\| \mid X_0 = v]$. Pour $i > 1$ et $u, v \in \mathbb{R}^d$, nous avons

$$\begin{aligned} & E[\|g(X_{i+1})\|g(X_1)\| \mid X_i = v, X_0 = u]f_{(X_i, X_0)}(v, u) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \|g(y)\|g(x)\|f_{X_{i+1}|X_i=v}(y)f_{X_i|X_1=x}(v)f_{X_1|X_0=u}(x)f_{X_0}(u)dx dy \\ &\leq f_{X_0}(u) \sup_{x \in \mathbb{R}^d} f_{X_1|X_0=x}(v)E[\|g(X_1)\| \mid X_0 = u]E[\|g(X_1)\| \mid X_0 = v]. \end{aligned}$$

1.3 L'estimateur à noyau $\bar{g}_n(x)$

Dans les paragraphes suivants nous présentons quelques propriétés asymptotiques de l'estimateur à noyau $\bar{g}_n(\cdot)$, que nous utilisons plus tard pour dériver des résultats de convergence pour l'estimateur $\phi_n(\cdot)$. Nous décrivons ensuite le comportement asymptotique de la covariance $cov(\bar{g}_n(x), \bar{g}_n(y))$, avec x et y points de $\mathbb{R}^{d'}$. En conséquence de ce résultat nous dérivons la convergence en moyenne quadratique de l'estimateur $\bar{g}_n(x)$. Puis nous établissons la normalité asymptotique des lois de dimension finie du processus empirique $(\bar{g}_n(x), x \in \mathbb{R}^{d'})$.

Notons pour $x \in \mathbb{R}^{d'}$ et $n \in \mathbb{N}^*$, $\bar{g}_n(x) = (g_{1,n}(x), \dots, g_{p,n}(x))$, où pour $r = 1, \dots, p$, et d'après (1.1.2), on a

$$g_{r,n}(x) = \frac{1}{nh_r^{d'}(n)} \sum_{i=1}^n g_r(V_i)K_r\left(\frac{x - W_i}{h_r(n)}\right), \quad (1.3.1)$$

où les variables aléatoires V_i et W_i , pour $i \in \mathbb{Z}$, sont définies par (1.1.1).

1.3.1 Comportement asymptotique de $cov(\bar{g}_n(x), \bar{g}_n(y))$

Pour $x, y \in \mathbb{R}^{d'}$, on a

$$cov(\bar{g}_n(x), \bar{g}_n(y)) = \left[cov(g_{r,n}(x), g_{r',n}(y)) \right]_{r,r'=1,\dots,p},$$

où pour $r, r' = 1, \dots, p$,

$$cov(g_{r,n}(x), g_{r',n}(y)) = E \left[\frac{1}{nh_r^{d'}(n)} \sum_{i=1}^n g_{r,x}(Z_i) \frac{1}{nh_{r'}^{d'}(n)} \sum_{j=1}^n g_{r',y}(Z_j) \right],$$

avec $Z_i = (V_i, W_i)$, $i \in \mathbb{Z}$, et

$$g_{r,x}(Z_i) = g_r(V_i)K_r\left(\frac{x - W_i}{h_r(n)}\right) - E \left[g_r(V_i)K_r\left(\frac{x - W_i}{h_r(n)}\right) \right].$$

Ainsi, d'après la stationnarité forte du processus $(Z_i, i \in \mathbb{Z})$, on peut écrire

$$\begin{aligned} \text{cov}(g_{r,n}(x), g_{r',n}(y)) & \quad (1.3.2) \\ &= \frac{1}{nh_r^{d'}(n)} \left(h_{r'}^{-d'}(n) E[g_{r,x}(Z_0)g_{r',y}(Z_0)] \right. \\ & \quad \left. + h_{r'}^{-d'} \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) \{E[g_{r,x}(Z_i)g_{r',y}(Z_0)] + E[g_{r',y}(Z_i)g_{r,x}(Z_0)]\} \right). \end{aligned}$$

L'étude du terme de covariance de l'égalité précédente nous conduit à l'introduction de conditions sur la vitesse de convergence vers zéro du coefficient de mélange et aussi à l'imposition de conditions de régularité sur certains moments conditionnels croisés. Nous le ferons à l'aide du lemme suivant:

Lemme 1.3.3 *Soient x un point de \mathbb{R}^t , $(g_i)_{i \in \mathbb{N}^*}$ une famille d'applications intégrables de \mathbb{R}^t dans \mathbb{R} , et φ une application de \mathbb{R}^t dans \mathbb{R} telle que: i) $\sup_{y \in \mathbb{R}^t} |\varphi(y)| < \infty$, ii) $\int_{\mathbb{R}^t} |\varphi(y)| dy < \infty$, et iii) $\lim_{\|y\| \rightarrow +\infty} \|y\|^t \varphi(y) = 0$. Si $\sup_{i \in \mathbb{N}^*} \int_{\mathbb{R}^t} |g_i(y)| dy < \infty$, et si les applications $g_i, i \in \mathbb{N}^*$ sont, par rapport à i , uniformément bornées dans un voisinage $\delta > 0$ de x , alors*

$$\sup_{i \in \mathbb{N}^*} \frac{1}{h^t} \int_{\mathbb{R}^t} g_i(y) \varphi\left(\frac{x-y}{h}\right) dy = O(1), \text{ lorsque } h \rightarrow 0.$$

Démonstration: Conséquence immédiate de l'inégalité suivante, valable pour tout $\delta > 0$, $h > 0$, et toute l'application g intégrable de \mathbb{R}^t dans \mathbb{R} (cf. Parzen [75] pg. 1067, Cacoullos [15] pg. 181)

$$\begin{aligned} \left| \frac{1}{h^t} \int_{\mathbb{R}^t} g(y) \varphi\left(\frac{x-y}{h}\right) dy - g(x) \int_{\mathbb{R}^t} \varphi(y) dy \right| & \leq \sup_{\|y\| \leq \delta} |g(x-y) - g(x)| \int_{\mathbb{R}^t} |\varphi(y)| dy \\ & \quad + \frac{1}{\delta^t} \sup_{\|y\| > \frac{\delta}{h}} \|y\|^t |\varphi(y)| \int_{\mathbb{R}^t} |g(y)| dy \\ & \quad + |g(x)| \int_{\|y\| > \frac{\delta}{h}} |\varphi(y)| dy. \quad \blacksquare \end{aligned}$$

Notons $\mathcal{K}^{d'}$ la classe des noyaux K sur $\mathbb{R}^{d'}$ qui satisfont les conditions i), ii) et iii) du lemme précédent. En particulier, si on note par $\mathcal{K}_c^{d'}$ la classe des noyaux sur $\mathbb{R}^{d'}$ bornés à support borné, on a $\mathcal{K}_c^{d'} \subset \mathcal{K}^{d'}$.

Proposition 1.3.4 *Soient $x, y \in \mathbb{R}^{d'}$, et (h_n) une suite de nombres réels strictement positifs telle que $h_n \rightarrow 0, n \rightarrow +\infty$. On suppose:*

- i) $E\|g(V_0)\|^{2+\delta} < \infty$, pour un $\delta > 0$;
- ii) $m(g'g, \cdot)f(\cdot) \in \mathcal{C}_x \cap \mathcal{C}_y$;
- iii) $m(\|g\|^{2+\delta}, \cdot)f(\cdot) \in \mathcal{B}_x \cap \mathcal{B}_y$;
- iv) $g \in \mathcal{N}_p(x, y)$;
- v) $(X_i, i \in \mathbb{Z})$ est de type $\text{Alpha}(\frac{\delta}{2+\delta})$;
- vi) pour $r = 1, \dots, p$, $h_r(n) = c_r h_n$, $c_r > 0$;
- vii) les noyaux K_r , $r = 1, \dots, p$, appartiennent à la classe $\mathcal{K}_C^{d'}$.

Alors, pour $r, r' = 1, \dots, p$,

$$\lim_{n \rightarrow +\infty} nh_n^{d'} \text{cov}(g_{r,n}(x), g_{r',n}(y)) = m(g_r g_{r'}, x) f(x) \left(\frac{1}{c_r c_{r'}} \right)^{d'} \int_{\mathbb{R}^{d'}} K_r \left(\frac{u}{c_r} \right) K_{r'} \left(\frac{u}{c_{r'}} \right) du \delta_{x-y},$$

où δ_x est la fonction indicatrice de l'ensemble $\{0\}$ évaluée en x .

Démonstration: Etant donnés $x, y \in \mathbb{R}^{d'}$ et $r, r' = 1, \dots, p$, on a d'après l'inégalité (1.3.2),

$$\begin{aligned} & nh_n^{d'} \text{cov}(g_{r,n}(x), g_{r',n}(y)) \\ &= \left(\frac{1}{c_r c_{r'}} \right)^{d'} \left(h_n^{-d'} E[g_{r,x}(Z_0)g_{r',y}(Z_0)] \right. \\ &\quad \left. + h_n^{-d'} \sum_{i=1}^{n-1} \left(1 - \frac{i}{n} \right) \{ E[g_{r,x}(Z_i)g_{r',y}(Z_0)] + E[g_{r',y}(Z_i)g_{r,x}(Z_0)] \} \right). \end{aligned} \quad (1.3.5)$$

Le comportement asymptotique du terme de variance de l'égalité précédente s'obtient en utilisant la démarche classique d'après les hypothèses i)-iii) (cf. Bosq et Lecoutre [9] pg. 87).

Ainsi, nous avons

$$\lim_{n \rightarrow +\infty} h_n^{-d'} E[g_{r,x}(Z_0)g_{r',y}(Z_0)] = m(g_r g_{r'}, x) f(x) \int_{\mathbb{R}^{d'}} K_r \left(\frac{u}{c_r} \right) K_{r'} \left(\frac{u}{c_{r'}} \right) du \delta_{x-y}. \quad (1.3.6)$$

Les hypothèses iii)-vii) nous allons permettre de conclure que le terme de variance de l'égalité (1.3.5) est plus significatif que celui de covariance. Etablissons donc la convergence vers zéro, lorsque n tend vers l'infini, de ce dernier terme. Pour cela, montrons que

$$\lim_{n \rightarrow +\infty} h_n^{-d'} \sum_{i=1}^{n-1} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| = 0. \quad (1.3.7)$$

Nous allons majorer le terme général de la série précédente, en utilisant soit l'inégalité pour les covariances entre fonctions de composantes écartées dans le temps d'un processus α -mélangeant obtenue par Davydov [21] (une démonstration de ce résultat est donnée dans

Hall [44] pg. 278), soit l'hypothèse $g \in \mathcal{N}_p(x, y)$. Ainsi, pour $i > s + s_q$, d'après le Lemme 1.3.3 et l'hypothèse iii), on a

$$\begin{aligned} & |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| \\ & \leq 8\alpha^{\frac{\delta}{2+\delta}}(i-s-s_q)E^{\frac{1}{2+\delta}}\left|g_r(V_i)K_r\left(\frac{x-W_i}{h_r(n)}\right)\right|^{2+\delta}E^{\frac{1}{2+\delta}}\left|g_{r'}(V_0)K_{r'}\left(\frac{y-W_0}{h_{r'}(n)}\right)\right|^{2+\delta} \\ & = O\left(\alpha^{\frac{\delta}{2+\delta}}(i-s-s_q)(h_r^{d'}(n)h_{r'}^{d'}(n))^{\frac{1}{2+\delta}}\right), \\ & = O\left(\alpha^{\frac{\delta}{2+\delta}}(i-s-s_q)(h_n^{d'})^{\frac{2}{2+\delta}}\right). \end{aligned}$$

D'autre part, on a pour $i \geq 1$,

$$|E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| = \left|E\left[g_r(V_i)K_r\left(\frac{x-W_i}{h_r(n)}\right)g_{r'}(V_0)K_{r'}\left(\frac{y-W_0}{h_{r'}(n)}\right)\right]\right| + O\left(h_n^{2d'}\right). \quad (1.3.8)$$

Le Lemme 1.3.3 et les hypothèses iv), vi) et vii), nous permettent d'obtenir les développements suivants pour le premier terme de la partie droite de l'égalité précédente dans les cas où $i > s_q$ et $i \leq s_q$. Soient $x = (x_0, x_1, \dots, x_q)$ et pour $u \in \mathbb{R}^d$, $\varphi_r(u) = \sup_{x_1, \dots, x_q \in \mathbb{R}^d} |K_r(u, x_1, \dots, x_q)|$.

Nous avons

$$\begin{aligned} & \sup_{i > s_q} \left|E\left[g_r(V_i)K_r\left(\frac{x-W_i}{h_r(n)}\right)g_{r'}(V_0)K_{r'}\left(\frac{y-W_0}{h_{r'}(n)}\right)\right]\right| \\ & \leq \sup_{i > s_q} \int_{\mathbb{R}^{d'} \times \mathbb{R}^{d'}} E[|g_r(V_i)||g_{r'}(V_0)| \mid W_i = u, W_0 = v] f_{(W_i, W_0)}(u, v) \\ & \quad \times |K_r|\left(\frac{x-u}{h_r(n)}\right) |K_{r'}|\left(\frac{y-v}{h_{r'}(n)}\right) dudv \\ & = O(h_n^{2d'}), \end{aligned}$$

et pour $1 \leq i \leq s_q$,

$$\begin{aligned} & \left|E\left[g_r(V_i)K_r\left(\frac{x-W_i}{h_r(n)}\right)g_{r'}(V_0)K_{r'}\left(\frac{y-W_0}{h_{r'}(n)}\right)\right]\right| \\ & \leq \int_{\mathbb{R}^d \times \mathbb{R}^{d'}} E[|g_r(V_i)||g_{r'}(V_0)| \mid X_i = u, W_0 = v] f_{(X_i, W_0)}(u, v) \\ & \quad \times \varphi_r\left(\frac{x_0-u}{h_r(n)}\right) |K_{r'}|\left(\frac{y-v}{h_{r'}(n)}\right) dudv \\ & = O(h_n^{d'+d}). \end{aligned}$$

D'après les majorations précédentes et en désignant par $[x]$ la partie entière de x , on a pour $0 < \xi < 1$,

$$h_n^{-d'} \sum_{i=1}^{+\infty} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]|$$

$$\begin{aligned}
&= \sum_{i=1}^{s_q} h_n^{-d'} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| + \sum_{i=s_q+1}^{[(h_n^{d'})^{\xi-1}] + s + s_q - 1} h_n^{-d'} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| \\
&\quad + \sum_{i=[(h_n^{d'})^{\xi-1}] + s + s_q}^{+\infty} h_n^{-d'} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]| \\
&= O\left(h_n^d + h_n^{d'} (h_n^{d'})^{\xi-1} + h_n^{-d'} \sum_{i=[(h_n^{d'})^{\xi-1}] + s + s_q}^{+\infty} \alpha^{\frac{\delta}{2+\delta}}(i-s-s_q) (h_n^{d'})^{\frac{2}{2+\delta}}\right) \\
&= O\left(h_n^d + (h_n^{d'})^\xi + \sum_{i=[(h_n^{d'})^{\xi-1}] + s + s_q}^{+\infty} \frac{\alpha^{\frac{\delta}{2+\delta}}(i)}{(h_n^{d'})^{\frac{\delta}{2+\delta}}}\right).
\end{aligned}$$

La conclusion découle maintenant de l'hypothèse v), car d'après la Proposition 1.2.3, il existe $\eta > 0$ tel que $n^{\frac{\delta}{2+\delta} + \eta} \sum_{i=n}^{+\infty} \alpha^{\frac{\delta}{2+\delta}}(i)$ converge vers zéro, lorsque $n \rightarrow +\infty$. Il suffit donc de prendre $\xi = \eta / (\frac{\delta}{2+\delta} + \eta)$ dans le développement précédent. \blacksquare

Remarquons que la proposition précédente reste valable si l'hypothèse sur les moments conditionnels croisés, $g \in \mathcal{N}_p(x, y)$, est remplacée par l'hypothèse $1 \in \mathcal{N}_p(x, y)$ qui se réduit à une hypothèse sur les densités jointes, et si de plus on admet que le processus $(X_i, i \in \mathbb{Z})$ est de type $Alpha(\frac{\delta}{4+\delta})$, c'est-à-dire, si nous exigeons une plus grande vitesse de convergence vers zéro du coefficient de mélange. De façon à justifier cette affirmation, vérifions que dans ces conditions, l'égalité (1.3.7) reste encore valable.

D'après l'inégalité de Hölder, on obtient pour $i \geq 1$,

$$\begin{aligned}
&\left| E \left[g_r(V_i) K_r \left(\frac{x - W_i}{h_r(n)} \right) g_{r'}(V_0) K_{r'} \left(\frac{y - W_0}{h_{r'}(n)} \right) \right] \right| \leq E^{\frac{1}{2+\delta}} \left| g_r^{2+\delta}(V_0) K_r \left(\frac{x - W_0}{h_r(n)} \right) \right| \\
&\quad \times E^{\frac{1}{2+\delta}} \left| g_{r'}^{2+\delta}(V_0) K_{r'} \left(\frac{y - W_0}{h_{r'}(n)} \right) \right| E^{\frac{\delta}{2+\delta}} \left| K_r \left(\frac{x - W_i}{h_r(n)} \right) K_{r'} \left(\frac{y - W_0}{h_{r'}(n)} \right) \right|^{\frac{1+\delta}{\delta}}.
\end{aligned}$$

En tenant compte de l'hypothèse $1 \in \mathcal{N}_p(x, y)$, nous pouvons conclure d'après le Lemme 1.3.3 et l'égalité (1.3.8), que le terme $|E[g_{r,x}(Z_i)g_{r',y}(Z_0)]|$ est, uniformément en i , de l'ordre de $h_n^{d'+d\frac{\delta}{2+\delta}}$ si $1 \leq i \leq s_q$, et de l'ordre de $h_n^{d'(1+\frac{\delta}{2+\delta})}$ si $i > s_q$.

D'autre part, si le processus est de type $Alpha(\frac{\delta}{4+\delta})$, on conclut d'après la Proposition 1.2.3 qu'il existe $\eta > 0$ tel que $n^{1+\eta} \sum_{i=n}^{+\infty} \alpha^{\frac{\delta}{2+\delta}}(i) \rightarrow 0, n \rightarrow +\infty$, ou encore,

$$\sum_{i=[(h_n^{d'})^{\frac{\delta}{2+\delta}(\xi-1)}]}^{+\infty} \frac{\alpha^{\frac{\delta}{2+\delta}}(i)}{(h_n^{d'})^{\frac{\delta}{2+\delta}}} \rightarrow 0, n \rightarrow +\infty, \text{ avec } \xi = \frac{\eta}{1+\eta}.$$

On trouve ainsi pour le terme $h_n^{-d'} \sum_{i=1}^{+\infty} |E[g_{r,x}(Z_i)g_{r',y}(Z_0)]|$ la majoration

$$C \left((h_n^d)^{\frac{\delta}{2+\delta}} + (h_n^{d'})^{\frac{\delta}{2+\delta}} \xi + \sum_{i=\lfloor (h_n^{d'})^{\frac{\delta}{2+\delta}}(\xi-1) \rfloor}^{+\infty} \frac{\alpha^{\frac{\delta}{2+\delta}}(i)}{(h_n^{d'})^{\frac{\delta}{2+\delta}}} \right),$$

valable pour un $C > 0$, ce qui permet de conclure.

1.3.2 Convergence en moyenne quadratique de $\bar{g}_n(x)$

La décomposition de l'erreur quadratique moyenne dans la somme d'un terme de variance et d'un terme de carré du biais (cf. Parzen [75] pg. 1069), nous permet d'obtenir la convergence en moyenne quadratique de l'estimateur $\bar{g}_n(x)$ vers $m(g, x)f(x)$, à partir de la convergence vers zéro de la variance de chacun des estimateurs $g_{r,n}(x)$ pour $r = 1, \dots, p$, qu'on peut déduire en suivant la démonstration de la Proposition 1.3.4 et en admettant que $nh_r^{d'}(n) \rightarrow +\infty$, $n \rightarrow +\infty$, pour $r = 1, \dots, p$, et à partir de la convergence vers zéro de leurs biais qu'on peut dériver de la façon usuelle, d'après le Lemme de Bochner (cf. Parzen [75] pg. 1067, et Cacoullos [15] pg. 180). Nous avons donc le

Théorème 1.3.9 *Soit x un point de $\mathbb{R}^{d'}$. On suppose:*

- i) $E\|g(V_0)\|^{2+\delta} < \infty$ pour un $\delta > 0$;
- ii) $m(g, \cdot)f(\cdot) \in \mathcal{C}_x$;
- iii) $m(\|g\|^{2+\delta}, \cdot)f(\cdot) \in \mathcal{B}_x$;
- iv) $g \in \mathcal{P}_p(x)$;
- v) $(X_i, i \in \mathbb{Z})$ est de type $\text{Alpha}(\frac{\delta}{2+\delta})$;
- vi) pour $r = 1, \dots, p$, $h_r(n) \rightarrow 0$, $nh_r^{d'}(n) \rightarrow +\infty$, si $n \rightarrow +\infty$;
- vii) les noyaux $K_r, r = 1, \dots, p$, appartiennent à la classe $\mathcal{K}_c^{d'}$.

Alors

$$\bar{g}_n(x) \rightarrow m(g, x)f(x), \text{ en moyenne quadratique, si } n \rightarrow +\infty.$$

1.3.3 Normalité asymptotique de $(\bar{g}_n(x_j) - E[\bar{g}_n(x_j)]); j = 1, \dots, T)$

Nous allons maintenant nous intéresser au comportement asymptotique de la suite de processus $(\bar{g}_n(x) - E[\bar{g}_n(x)], x \in \mathbb{R}^{d'})$, et notamment à la convergence de ses lois de dimension finie. Le théorème suivant, dont la démonstration est fondée sur la démonstration correspondante donnée

par Rosenblatt [85] pour l'estimateur de la densité, établit la normalité asymptotique de telles lois de dimension finie.

Théorème 1.3.10 *Soient x_1, \dots, x_T , T points distincts de $\mathbb{R}^{d'}$, et (h_n) une suite de nombres réels strictement positifs telle que $h_n \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty$, lorsque $n \rightarrow +\infty$. On suppose:*

- i) $E\|g(Y_0)\|^{2+\delta} < \infty$ pour un $\delta > 0$;
- ii) $m(g'g, \cdot)f(\cdot) \in \bigcap_{j=1}^T \mathcal{C}_{x_j}$;
- iii) $m(\|g\|^{2+\delta}, \cdot)f(\cdot) \in \bigcap_{j=1}^T \mathcal{B}_{x_j}$;
- iv) $g \in \mathcal{N}_p(x_1, \dots, x_T)$;
- v) $(X_i, i \in \mathbb{Z})$ est de type $\text{Alpha}(\frac{\delta}{2+\delta})$;
- vi) pour $r = 1, \dots, p$, $h_r(n) = c_r h_n$, $c_r > 0$;
- vii) les noyaux $K_r, r = 1, \dots, p$, appartiennent à la classe $\mathcal{K}_c^{d'}$.

Alors le vecteur ligne

$$\sqrt{nh_n^{d'}} (\bar{g}_n(x_j) - E[\bar{g}_n(x_j)] ; j = 1, \dots, T),$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de matrice de covariance $\Sigma = [\Sigma_{i,j}]_{i,j=1,\dots,T}$, où $\Sigma_{i,j}$ a pour éléments

$$\sigma_{r,r'}^{i,j} = m(g_r g_{r'}, x_j) f(x_j) \left(\frac{1}{c_r c_{r'}} \right)^{d'} \int_{\mathbb{R}^{d'}} K_r \left(\frac{u}{c_r} \right) K_{r'} \left(\frac{u}{c_{r'}} \right) du \delta_{x_i - x_j}, \quad (1.3.11)$$

pour $r, r' = 1, \dots, p$.

Démonstration: Soit $C' = (C_1, \dots, C_{pT})$ un vecteur quelconque de \mathbb{R}^{pT} , et désignons par V_n le vecteur ligne $\sqrt{nh_n^{d'}} (\bar{g}_n(x_j) - E[\bar{g}_n(x_j)] ; j = 1, \dots, T)$. D'après le théorème de Cramer-Wold (cf. [68] pg. 162), nous montrerons que $V_n C'$ est asymptotiquement normal de moyenne zéro et de variance $C' \Sigma C$. La variable $V_n C$ s'écrit sous la forme

$$V_n C = \sum_{i=1}^n S_{n,i},$$

où pour $i = 1, \dots, n$,

$$S_{n,i} = \sum_{r=1}^p \sum_{j=1}^T \frac{C_{(j-1)p+r}}{c_r^{d'/2}} \frac{1}{(nh_r^{d'}(n))^{1/2}} g_{r,x_j}(Z_i),$$

et pour $r = 1, \dots, p$ et $j = 1, \dots, T$,

$$g_{r,x_j}(Z_i) = g_r(V_i)K_r\left(\frac{x_j - W_i}{h_r(n)}\right) - E\left[g_r(V_0)K_r\left(\frac{x_j - W_0}{h_r(n)}\right)\right].$$

Soient $r(n) = r$, $m(n) = m$ et $k(n) = k$ des suites de nombres entiers convergeant vers l'infini, telles que

$$\frac{m}{r} \rightarrow +\infty, \text{ et } k = \left[\frac{n}{m+r}\right], \quad (1.3.12)$$

où $[x]$ désigne la partie entière de x .

Considérons la décomposition

$$\sum_{i=1}^n S_{n,i} = \sum_{l=1}^k (A_l + B_l) + H,$$

où pour $l = 1, \dots, k$ et $r'(n) = r' = r(n) + s + s_q - 1$

$$A_l = \sum_{i=(l-1)(m+r')+1}^{lm+(l-1)r'} S_{n,i}, \quad B_l = \sum_{i=lm+(l-1)r'+1}^{l(m+r')} S_{n,i} \text{ et } H = \sum_{i=k(m+r')+1}^n S_{n,i}.$$

A) Convergence en moyenne quadratique vers zéro des variables $\sum_{l=1}^k B_l$ et H

Comme

$$E^{\frac{1}{2}}\left[\sum_{l=1}^k B_l\right]^2 \leq \sum_{r=1}^p \sum_{j=1}^T \frac{|C_{(j-1)p+r}|}{|c_r|^{d_r/2}} E^{\frac{1}{2}}\left[\frac{1}{(nh_r^{d_r}(n))^{1/2}} \sum_{l=1}^k \sum_{i=lm+(l-1)r'+1}^{l(m+r')} g_{r,x_j}(Z_i)\right]^2,$$

il suffit de montrer que pour $r = 1, \dots, p$ et $j = 1, \dots, T$,

$$E\left[\frac{1}{(nh_r^{d_r}(n))^{1/2}} \sum_{l=1}^k U_l^j\right]^2 \rightarrow 0, n \rightarrow +\infty,$$

où pour $l = 1, \dots, k$,

$$U_l^j = \sum_{i=lm+(l-1)r'+1}^{l(m+r')} g_{r,x_j}(Z_i).$$

En effet, comme

$$E\left[\frac{1}{(nh_r^{d_r}(n))^{1/2}} \sum_{l=1}^k U_l^j\right]^2 = \frac{k}{nh_r^{d_r}(n)} E(U_1^j)^2 + \frac{2}{nh_r^{d_r}(n)} \sum_{l=1}^{k-1} (k-l) E(U_{l+1}^j U_1^j),$$

où

$$\frac{k}{nh_r^{d_r}(n)} E(U_1^j)^2 \leq \frac{kr'}{n} \frac{E(g_{r,x_j}(Z_0))^2}{h_r^{d_r}(n)} + 2\frac{kr'}{n} \sum_{i=1}^{\infty} \frac{|E[g_{r,x_j}(Z_i)g_{r,x_j}(Z_0)]|}{h_r^{d_r}(n)},$$

et

$$\left| \frac{1}{nh_r^{d'}(n)} \sum_{l=1}^{k-1} (k-l) E(U_{l+1}^j U_1^j) \right| \leq \frac{kr'}{n} \sum_{i=1}^{\infty} \frac{|E[g_{r,x_j}(Z_i)g_{r,x_j}(Z_0)]|}{h_r^{d'}(n)},$$

on a la conclusion désirée d'après (1.3.6) et (1.3.7), en notant que $\frac{kr'}{n} \rightarrow 0$, lorsque n tend vers l'infini.

De façon analogue on démontrerait la convergence en moyenne quadratique vers zéro de H .

B) Décomposition de la variable $\sum_{l=1}^k A_l$

Soit maintenant (L_n) une suite de nombres positifs convergeant vers l'infini, et considérons pour $i = 1, \dots, n$, la décomposition

$$S_{n,i} = \sum_{r=1}^p \sum_{j=1}^T \frac{C_{(j-1)p+r}}{c_r^{d'/2}} \frac{1}{(nh_r^{d'}(n))^{1/2}} \{g'_{r,x_j}(Z_i) + g''_{r,x_j}(Z_i)\} = S'_{n,i} + S''_{n,i},$$

où pour $r = 1, \dots, p$ et $j = 1, \dots, T$,

$$g'_{r,x_j}(Z_i) = g_r(V_i) \mathbb{1}_{|g_r(V_i)| \leq L_n} K_r \left(\frac{x_j - W_i}{h_r(n)} \right) - E \left[g_r(V_0) \mathbb{1}_{|g_r(V_0)| \leq L_n} K_r \left(\frac{x_j - W_0}{h_r(n)} \right) \right] \text{ et}$$

$$g''_{r,x_j}(Z_i) = g_{r,x_j}(Z_i) - g'_{r,x_j}(Z_i).$$

Soient encore, pour $l = 1, \dots, k$,

$$A'_l = \sum_{i=(l-1)(m+r')+1}^{lm+(l-1)r'} S'_{n,i}, \text{ et } A''_l = A_l - A'_l.$$

La variable aléatoire $\sum_{l=1}^k A''_l$ converge en moyenne quadratique vers zéro, lorsque n tend vers l'infini. En effet, en procédant comme dans A), il suffira de montrer que $h_r^{-d'}(n) E(g''_{r,x_j}(Z_0))^2$ est négligeable pour pouvoir conclure que pour $r = 1, \dots, p$ et $j = 1, \dots, T$,

$$E \left[\frac{1}{(nh_r^{d'}(n))^{1/2}} \sum_{l=1}^k V_l^j \right]^2 \rightarrow 0, n \rightarrow +\infty,$$

où pour $l = 1, \dots, k$,

$$V_l^j = \sum_{i=(l-1)(m+r')+1}^{lm+(l-1)r'} g''_{r,x_j}(Z_i).$$

D'après l'inégalité de Bienaymé-Chebyshev et le Lemme 1.3.3, nous avons

$$\begin{aligned} & h_r^{-d'}(n) E(g''_{r,x_j}(Z_0))^2 \\ &= \frac{1}{h_r^{d'}(n)} \int_{\mathbb{R}^{d'}} E[g_r^2(V_0) \mathbb{1}_{|g_r(V_0)| > L_n} | W_0 = z] K_r^2 \left(\frac{x_j - z}{h_r(n)} \right) f(z) dz + O(h_r^{d'}(n)) \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{h_r^{d'}(n)} \int_{\mathbb{R}^{d'}} E^{\frac{2}{2+\delta}} [|g_r^{2+\delta}(V_0)| \mid W_0 = z] E^{\frac{\delta}{2+\delta}} [\mathbb{1}_{|g_r(V_0)| > L_n} \mid W_0 = z] \\
&\quad \times K_r^2 \left(\frac{x_j - z}{h_r(n)} \right) f(z) dz + O(h_r^{d'}(n)) \\
&\leq \frac{1}{L_n^\delta} \frac{1}{h_r^{d'}(n)} \int_{\mathbb{R}^{d'}} m(|g_r|^{2+\delta}, z) K_r^2 \left(\frac{x_j - z}{h_r(n)} \right) f(z) dz + O(h_r^{d'}(n)), \\
&= O\left(\frac{1}{L_n^\delta}\right) + O(h_r^{d'}(n)) = o(1).
\end{aligned}$$

D'après les points précédents il suffit donc d'étudier la loi limite de la variable $\sum_{l=1}^k A'_l$.

C) Normalité asymptotique de la variable $\sum_{l=1}^k A'_l$

D'après Robinson [81] (pg. 199), soient

$$m(n) = [n^{\frac{1}{3}} \eta_n] \text{ et } r(n) = [n^{\frac{1}{3}} \eta_n^2],$$

où $\eta_n \rightarrow 0, n \rightarrow +\infty$, avec $\eta_n > \max\left(n^{-\frac{1}{24}}, \epsilon_{n^{1/4}}^{\frac{1}{6}}\right)$, et $\epsilon_n = \sup_{j \geq n} j^2 \alpha(j) \downarrow 0, n \rightarrow +\infty$. Les conditions (1.3.12) sont évidemment satisfaites.

Montrons maintenant que les variables $(A'_l), l = 1, \dots, k$, peuvent être considérées comme des variables indépendantes. Nous avons (cf. Rosenblatt [85] pg. 207)

$$\begin{aligned}
B_n &= \left| E\left(\exp\left(it \sum_{l=1}^k A'_l\right)\right) - \prod_{l=1}^k E(\exp(itA'_l)) \right| \\
&\leq \sum_{j=1}^{k-1} \left| E\left(\exp\left(it \sum_{l=1}^j A'_l\right) \exp(itA'_{j+1})\right) - E\left(\exp\left(it \sum_{l=1}^j A'_l\right)\right) E(\exp(itA'_{j+1})) \right|.
\end{aligned}$$

Comme pour $j = 1, \dots, k-1$,

$$\exp\left(it \sum_{l=1}^j A'_l\right) \in F_{-\infty}^{j(m+r')-r'+s} \text{ et } \exp(itA'_{j+1}) \in F_{j(m+r')+1-s_q}^{+\infty}$$

et en notant que ces variables sont en module bornées par 1, alors d'après l'inégalité de Ibragimov [54] (Lemme 1.2, pg. 352), on obtient

$$\begin{aligned}
B_n &\leq \sum_{j=1}^{k-1} 4\alpha(1 - s_q + r' - s) \\
&= 4(k(n) - 1)\alpha(r(n)), \text{ car } r'(n) = r(n) + s + s_q - 1, \\
&= O(\eta_n^{-5} \epsilon_{r(n)}) = O(\eta_n^{-5} \epsilon_{n^{1/4}}) \\
&= O\left(\epsilon_{n^{1/4}}^{\frac{1}{6}}\right) = o(1).
\end{aligned}$$

(L'inégalité de Ibragimov [54], à laquelle nous avons fait référence, a été d'abord obtenue, dans une forme moins précise, par Volkonskiï et Rozanov [106]).

La normalité asymptotique de $\sum_{l=1}^k A'_l$ découlera maintenant de l'application du théorème de limite centrale de Liapounov (cf. [68] pg. 168). Pour cela, nous allons établir que

$$1) \lim_{n \rightarrow +\infty} \sum_{l=1}^k E[A'_l]^2 = C' \Sigma C \text{ et } 2) \lim_{n \rightarrow +\infty} \sum_{l=1}^k E[A'_l]^4 = 0. \quad (1.3.13)$$

D) Vérification de la condition 1) de (1.3.13)

Comme

$$\begin{aligned} \sum_{l=1}^k E[A'_l]^2 &= kE[A'_1]^2 \\ &= kE[A_1]^2 + kE[A_1'']^2 + O\left((kE[A_1]^2)^{1/2} (kE[A_1'']^2)^{1/2}\right), \end{aligned}$$

nous allons montrer que $\lim_{n \rightarrow +\infty} kE[A_1]^2 = C' \Sigma C$ et que $\lim_{n \rightarrow +\infty} kE[A_1'']^2 = 0$.

Le terme $kE[A_1]^2$ admet le développement

$$\begin{aligned} kE[A_1]^2 &= \sum_{r_1, r_2=1}^p \sum_{j_1, j_2=1}^T \frac{C_{(j_1-1)p+r_1} C_{(j_2-1)p+r_2}}{(c_{r_1} c_{r_2})^{d'/2}} \\ &\quad \times kE \left[\frac{1}{(nh_{r_1}^{d'}(n))^{1/2}} \sum_{i=1}^m g_{r_1, x_{j_1}}(Z_i) \frac{1}{(nh_{r_2}^{d'}(n))^{1/2}} \sum_{j=1}^m g_{r_2, x_{j_2}}(Z_j) \right] \\ &= \sum_{r_1, r_2=1}^p \sum_{j_1, j_2=1}^T \frac{C_{(j_1-1)p+r_1} C_{(j_2-1)p+r_2}}{(c_{r_1} c_{r_2})^{d'/2}} \frac{1}{(h_{r_1}(n) h_{r_2}(n))^{d'/2}} \\ &\quad \times \left\{ \frac{km}{n} E[g_{r_1, x_{j_1}}(Z_0) g_{r_2, x_{j_2}}(Z_0)] \right. \\ &\quad \left. + \frac{k}{n} \sum_{i=1}^m (m-i) \{ E[g_{r_1, x_{j_1}}(Z_i) g_{r_2, x_{j_2}}(Z_0)] + E[g_{r_2, x_{j_2}}(Z_i) g_{r_1, x_{j_1}}(Z_0)] \} \right\} \\ &= \sum_{r_1, r_2=1}^p \sum_{j_1, j_2=1}^T \frac{C_{(j_1-1)p+r_1} C_{(j_2-1)p+r_2}}{(c_{r_1} c_{r_2})^{d'/2}} \\ &\quad \times \frac{km}{n} \left(\frac{h_{r_1}(n)}{h_{r_2}(n)} \right)^{\frac{d'}{2}} \frac{1}{h_{r_1}^{d'}(n)} E[g_{r_1, x_{j_1}}(Z_0) g_{r_2, x_{j_2}}(Z_0)] + o(1), \end{aligned}$$

puisque d'après (1.3.7)

$$\begin{aligned} &\left| \frac{k}{n(h_{r_1}(n) h_{r_2}(n))^{d'/2}} \sum_{i=1}^m (m-i) \{ E[g_{r_1, x_{j_1}}(Z_i) g_{r_2, x_{j_2}}(Z_0)] + E[g_{r_2, x_{j_2}}(Z_i) g_{r_1, x_{j_1}}(Z_0)] \} \right| \\ &\leq \frac{km}{n} \left(\frac{h_{r_1}(n)}{h_{r_2}(n)} \right)^{\frac{d'}{2}} \sum_{i=1}^{\infty} \frac{|E[g_{r_1, x_{j_1}}(Z_i) g_{r_2, x_{j_2}}(Z_0)]| + |E[g_{r_2, x_{j_2}}(Z_i) g_{r_1, x_{j_1}}(Z_0)]|}{h_{r_1}^{d'}(n)} = o(1). \end{aligned}$$

En tenant compte de (1.3.6), du fait que $\frac{km}{n} \rightarrow 1, n \rightarrow +\infty$, et de la définition (1.3.11) de la matrice Σ , on peut donc écrire

$$\begin{aligned} \lim_{n \rightarrow +\infty} kE[A_1]^2 &= \sum_{r_1, r_2=1}^p \sum_{j_1, j_2=1}^T C_{(j_1-1)p+r_1} C_{(j_2-1)p+r_2} \\ &\quad \times \left(\frac{1}{c_{r_1} c_{r_2}} \right)^{d'} f(x_{j_1}) m(g_{r_1} g_{r_2}, x_{j_1}) \int_{\mathbb{R}^{d'}} K_{r_1} \left(\frac{v}{c_{r_1}} \right) K_{r_2} \left(\frac{v}{c_{r_2}} \right) dv \delta_{x_{j_1} - x_{j_2}} \\ &= \sum_{r_1, r_2=1}^p \sum_{j_1, j_2=1}^T C_{(j_1-1)p+r_1} C_{(j_2-1)p+r_2} \sigma_{r_1, r_2}^{j_1, j_2} \\ &= C' \Sigma C. \end{aligned}$$

En utilisant la technique de démonstration employée dans le point B) précédent, on conclut que $kE[A_1'']^2$ converge vers zéro, lorsque n tend vers l'infini.

E) Vérification de la condition 2) de (1.3.13)

Considérons la majoration

$$\begin{aligned} \left(\sum_{l=1}^k E[A_l']^4 \right)^{1/4} &= (kE[A_1']^4)^{1/4} \\ &\leq \sum_{r=1}^p \sum_{j=1}^T \frac{|C_{(j-1)p+r}|}{|c_r|^{d'/2}} \left(kE \left[\frac{1}{(nh_r^{d'}(n))^{1/2}} \sum_{i=1}^m g'_{r, x_j}(Z_i) \right]^4 \right)^{1/4}, \end{aligned}$$

où pour $r = 1, \dots, p$ et $j = 1, \dots, T$,

$$\begin{aligned} kE \left[\frac{1}{(nh_r^{d'}(n))^{1/2}} \sum_{i=1}^m g'_{r, x_j}(Z_i) \right]^4 &= \frac{k}{(nh_r^{d'}(n))^2} \sum_{i_1} \left\{ E[g'_{r, x_j}(Z_{i_1})]^4 \right. \\ &\quad + \sum_{i_1 \neq i_2} E[g_{r, x_j}'^2(Z_{i_1}) g'_{r, x_j}(Z_{i_2}) (g'_{r, x_j}(Z_{i_1}) + g'_{r, x_j}(Z_{i_2}))] \\ &\quad + \sum_{i_1 \neq i_3 \neq i_2} \left(E[g_{r, x_j}'^2(Z_{i_1}) g'_{r, x_j}(Z_{i_2}) g'_{r, x_j}(Z_{i_3})] \right. \\ &\quad \left. \left. + \sum_{\substack{i_1 \neq i_4 \neq i_2 \\ i_4 \neq i_3}} E[g'_{r, x_j}(Z_{i_1}) g'_{r, x_j}(Z_{i_2}) g'_{r, x_j}(Z_{i_3}) g'_{r, x_j}(Z_{i_4})] \right) \right\}. \end{aligned}$$

Etant donné que d'après l'hypothèse $g \in \mathcal{N}(x_1, \dots, x_T)$, on a $E[g'_{r, x_j}(Z_0)]^4 = O(L_n^2 h_r^{d'}(n))$, et que, pour $i > s_q$ et uniformément en i , $E[g'_{r, x_j}(Z_i) g'_{r, x_j}(Z_0)]^2 = O(L_n^2 h_r^{2d'}(n))$, on obtient

finaleme

$$\begin{aligned} \sum_{l=1}^k E[A_l^4] &= O\left(\frac{k}{(nh_r^{d'}(n))^2} L_n^2(mh_r^{d'}(n) + m^2 h_r^{3d'/2}(n) + m^4 h_r^{2d'}(n))\right) \\ &= O\left(L_n^2 \frac{m^3}{n}\right) + o(1) \\ &= O(L_n^2 m_n^3) + o(1). \end{aligned}$$

Il suffit maintenant, pour conclure la démonstration du Théorème 1.3.10, de choisir L_n tel que $L_n^2 m_n^3 = o(1)$. \blacksquare

1.3.4 Normalité asymptotique de $(\bar{g}_n(x_j) - m(g, x_j)f(x_j); j = 1, \dots, T)$

La normalité asymptotique du vecteur aléatoire

$$\sqrt{nh_n^{d'}} (\bar{g}_n(x_j) - m(g, x_j)f(x_j); j = 1, \dots, T)$$

qui admet la décomposition

$$\sqrt{nh_n^{d'}} (\bar{g}_n(x_j) - E[\bar{g}_n(x_j)]; j = 1, \dots, T) + \sqrt{nh_n^{d'}} (E[\bar{g}_n(x_j)] - m(g, x_j)f(x_j); j = 1, \dots, T),$$

est une conséquence du Théorème 1.3.10 et du comportement asymptotique du terme déterministe de la somme précédente. L'inclusion du facteur $\sqrt{nh_n^{d'}}$ dans le terme de biais, nous oblige à préciser sa vitesse de convergence vers zéro. Pour cela, des hypothèses supplémentaires sur les noyaux, sur $m(g, \cdot)f(\cdot)$ et sur (h_n) sont nécessaires. L'emploi de noyaux bornés à support borné, nous permettra de préciser cette vitesse en utilisant des hypothèses locales sur le paramètre fonctionnel $m(g, \cdot)f(\cdot)$.

Pour $m \in \mathbb{N}^*$, désignons par $\mathcal{K}^{d'}(m)$ la classe des noyaux K sur $\mathbb{R}^{d'}$ d'ordre m , c'est-à-dire, tels que

$$\int_{\mathbb{R}^{d'}} \|z\|^m |K(z)| dz < \infty \text{ et } \int_{\mathbb{R}^{d'}} z_1^{a_1} \dots z_{d'}^{a_{d'}} K(z) dz = 0,$$

avec $a_1, \dots, a_{d'} \in \mathbb{N}^*$, tels que $0 < \sum_{i=1}^{d'} a_i < m$ et $z = (z_1, z_2, \dots, z_{d'})$. Soit de plus $\mathcal{K}_c^{d'}(m)$ la sous-classe de $\mathcal{K}^{d'}(m)$ des noyaux sur $\mathbb{R}^{d'}$ d'ordre m , bornés à support borné. Pour $m \geq 1$, on a $\mathcal{K}_c^{d'}(m) \subset \mathcal{K}^{d'}(m)$ et en particulier $\mathcal{K}_c^{d'}(1) = \mathcal{K}_c^{d'}$.

Le Théorème suivant établit le comportement asymptotique des lois de dimension finie de la suite de processus $(\bar{g}_n(x) - m(g, x)f(x); x \in \mathbb{R}^{d'})$.

Théorème 1.3.14 Soient x_1, \dots, x_T , T points distincts de $\mathbb{R}^{d'}$, et (h_n) une suite de nombres réels strictement positifs telle que $h_n \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty$, lorsque $n \rightarrow +\infty$. Supposons satisfaites les hypothèses i)-vi) du Théorème 1.3.10. Considérons les trois systèmes d'hypothèses:

H₁) $m(g, \cdot)f(\cdot) \in \bigcap_{j=1}^T \mathcal{A}_{x_j}(\alpha)$, pour un $\alpha \in]0, 1]$, $nh_n^{d'+2\alpha} \rightarrow 0$, $n \rightarrow +\infty$, et $K_r \in \mathcal{K}_c^{d'}$, pour $r = 1, \dots, p$.

H₂) $m(g, \cdot)f(\cdot) \in \bigcap_{j=1}^T \mathcal{D}_{x_j}(m)$, pour un $m \geq 2$, $nh_n^{d'+2m} \rightarrow 0$, $n \rightarrow +\infty$, et $K_r \in \mathcal{K}_c^{d'}(m)$, pour $r = 1, \dots, p$.

H₃) $m(g, \cdot)f(\cdot) \in \bigcap_{j=1}^T \mathcal{D}_{x_j}(m)$, pour un $m \geq 1$, $nh_n^{d'+2m} \rightarrow 1$, $n \rightarrow +\infty$, $K_r \in \mathcal{K}_c^{d'}(m)$, pour $r = 1, \dots, p$.

a) Si l'un des systèmes d'hypothèses **H₁**) ou **H₂**) est satisfait, alors le vecteur ligne

$$\sqrt{nh_n^{d'}} (\bar{g}_n(x_j) - m(g, x_j)f(x_j); j = 1, \dots, T), \quad (1.3.15)$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de matrice de covariance Σ définie par (1.3.11).

b) Si le système d'hypothèses **H₃**) est satisfait, alors le vecteur ligne (1.3.15) converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne

$$\left(\Delta^m(m(g, \cdot)f(\cdot), K)(x_j) \begin{bmatrix} c_1^m & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & c_p^m \end{bmatrix}; j = 1, \dots, T \right),$$

et de matrice de covariance Σ définie par (1.3.11), où pour $x \in \mathbb{R}^{d'}$, $\Delta^m(m(g, \cdot)f(\cdot), K)(x)$ est le vecteur ligne

$$\Delta^m(m(g, \cdot)f(\cdot), K)(x) = \left(\Delta^m(m(g_r, \cdot)f(\cdot), K_r)(x); r = 1, \dots, p \right), \quad (1.3.16)$$

avec

$$\Delta^m(m(g_r, \cdot)f(\cdot), K_r)(x) = \frac{(-1)^m}{m!} \sum_{i_1, \dots, i_m=1}^{d'} \int_{\mathbb{R}^{d'}} u_{i_1} \dots u_{i_m} K_r(u) du \frac{\partial^m (m(g_r, \cdot)f(\cdot))}{\partial x_{i_1} \dots \partial x_{i_m}}(x),$$

et $u = (u_1, \dots, u_{d'})$.

Démonstration: Conséquence immédiate du Théorème 1.3.10, en tenant compte du fait que pour $r = 1, \dots, p$ et $x \in \mathbb{R}^{d'}$, le biais $E[g_{r,n}(x)] - m(g_r, x)f(x)$ admet les développements suivants, selon le système d'hypothèses considéré.

Sous H_1) et pour n assez grand, on a

$$|E[g_{r,n}(x)] - m(g_r, x)f(x)| = O\left(\int_{\mathbb{R}^{d'}} \|yh_r(n)\|^\alpha |K_r(y)| dy\right) = O(h_r^\alpha(n)).$$

Sous H_2) ou H_3), un développement de Taylor de $m(g_r, \cdot)f(\cdot)$ dans un voisinage de x nous conduit à (voir aussi Bosq et Lecoutre [9] pg. 88)

$$\begin{aligned} E[g_{r,n}(x)] - m(g_r, x)f(x) &= \int_{\mathbb{R}^{d'}} K_r(u) [m(g_r, x - uh_r(n))f(x - uh_r(n)) - m(g_r, x)f(x)] du \\ &= \int_{\mathbb{R}^{d'}} K_r(u) \left\{ \sum_{j=1}^{m-1} \frac{(-1)^j}{j!} h_r^j(n) \sum_{i_1, \dots, i_j=1}^{d'} u_{i_1} \dots u_{i_j} \frac{\partial^j (m(g_r, \cdot)f(\cdot))}{\partial x_{i_1} \dots \partial x_{i_j}}(x) \right\} du \\ &\quad + \int_{\mathbb{R}^{d'}} K_r(u) \left\{ \frac{(-1)^m}{(m-1)!} h_r^m(n) \sum_{i_1, \dots, i_m=1}^{d'} u_{i_1} \dots u_{i_m} \right. \\ &\quad \left. \times \int_0^1 \frac{\partial^m (m(g_r, \cdot)f(\cdot))}{\partial x_{i_1} \dots \partial x_{i_m}}(x - h_r(n)ut)(1-t)^{m-1} dt \right\} du. \end{aligned}$$

D'après le théorème de la convergence dominée, et les conditions imposées sur le noyau K_r , on obtient

$$E[g_{r,n}(x)] - m(g_r, x)f(x) = h_r^m(n) \Delta^m(m(g_r, \cdot)f(\cdot), K_r)(x) + o(h_r^m(n)).$$

Les hypothèses sur la suite (h_n) nous permettent maintenant d'obtenir le résultat annoncé. ■

1.4 Estimation d'une fonction de moments

D'après les résultats précédents, nous pouvons facilement déduire des estimateurs par noyau d'une fonction de $m(g, x)f(x)$, pour $x \in \mathbb{R}^{d'}$. Etant donnée une telle fonction

$$\phi(x) = \Phi(m(g_1, x)f(x), \dots, m(g_p, x)f(x)),$$

où on suppose toujours que $m(g, x)f(x)$ appartient à l'intérieur du domaine de Φ , nous pouvons l'estimer par

$$\phi_n(x) = \Phi(g_{1,n}(x), \dots, g_{p,n}(x)),$$

où $(g_{1,n}(x), \dots, g_{p,n}(x)) = \bar{g}_n(x)$ est défini par (1.1.2). Ainsi, la convergence en probabilité de $\phi_n(x)$ vers $\phi(x)$ (si Φ est continue) et sa normalité asymptotique (si Φ est différentiable), peuvent être obtenues en tenant compte la convergence et la normalité asymptotique de l'estimateur $\bar{g}_n(x)$. Ce sont ces résultats complémentaires que nous énonçons ci-dessous.

1.4.1 Convergence en probabilité de $\phi_n(x)$

Théorème 1.4.1 *Soit x un point de $\mathbb{R}^{d'}$. Sous les conditions du Théorème 1.3.9, si Φ est une application de \mathbb{R}^p dans $\mathbb{R}^{p'}$ continue en $m(g, x)f(x)$, alors*

$$\phi_n(x) \rightarrow \phi(x), \text{ en probabilité, si } n \rightarrow +\infty.$$

Démonstration: D'après le Théorème 1.3.9, et pour $x \in \mathbb{R}^{d'}$, nous avons

$$\bar{g}_n(x) \rightarrow m(g, x)f(x), \text{ en probabilité, si } n \rightarrow +\infty.$$

En tenant compte des égalités $\phi_n(x) = \Phi(\bar{g}_n(x))$ et $\phi(x) = \Phi(m(g, x)f(x))$, où Φ est continue en $m(g, x)f(x)$, on conclut d'après le Théorème de Slutsky (cf. [68] pg. 169) que $\phi_n(x)$ converge en probabilité vers $\phi(x)$, lorsque n tend vers l'infini. ■

Notons que si l'application Φ est linéaire, nous pouvons, sous les conditions du Théorème 1.4.1, déduire la convergence en moyenne quadratique de $\phi_n(x)$ vers $\phi(x)$, lorsque n tend vers l'infini.

1.4.2 Normalité asymptotique de $(\phi_n(x_j) - \phi(x_j); j = 1, \dots, T)$

Les systèmes d'hypothèses $\mathbf{H}_1)$, $\mathbf{H}_2)$ et $\mathbf{H}_3)$ auxquels nous faisons référence dans la suite sont ceux définis dans le Théorème 1.3.14.

Théorème 1.4.2 *Soient x_1, \dots, x_T , T points distincts de $\mathbb{R}^{d'}$ et Φ une application de \mathbb{R}^p dans $\mathbb{R}^{p'}$ différentiable en $m(g, x_1)f(x_1), \dots, m(g, x_T)f(x_T)$, dont la matrice jacobienne de taille (p', p) au point y sera désignée par $\frac{\partial \Phi}{\partial y}(y)$. Sous les conditions du Théorème 1.3.14, on a:*

a) *Si l'un des systèmes d'hypothèses $\mathbf{H}_1)$ ou $\mathbf{H}_2)$ est satisfait, alors le vecteur ligne*

$$\sqrt{nh_n^{d'}} (\phi_n(x_j) - \phi(x_j); j = 1, \dots, T), \quad (1.4.3)$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de matrice de covariance

$$\left[\frac{\partial \Phi}{\partial y}(m(g, x_i)f(x_i)) \Sigma_{i,j} \left(\frac{\partial \Phi}{\partial y} \right)' (m(g, x_j)f(x_j)) \right]_{i,j=1,\dots,T} \quad (1.4.4)$$

où la matrice $\Sigma = [\Sigma_{i,j}]_{i,j=1,\dots,T}$ est définie par (1.3.11).

b) Si le système d'hypothèses \mathbf{H}_3) est satisfait, alors le vecteur ligne (1.4.3) converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne

$$\left(\Delta^m(m(g, \cdot)f(\cdot), K)(x_j) \begin{bmatrix} c_1^m & \dots & 0 \\ 0 & & c_p^m \end{bmatrix} \left(\frac{\partial \Phi}{\partial y} \right)' (m(g, x_j)f(x_j)); j = 1, \dots, T \right),$$

et de matrice de covariance donnée par (1.4.4), où $\Delta^m(m(g, \cdot)f(\cdot), K)$ est défini par (1.3.16).

Notons que la dynamique du processus n'intervient pas dans l'expression de la matrice de covariance asymptotique (1.4.4) et qu'il y a indépendance asymptotique des composantes du vecteur étudié. Ceci résulte du caractère local du paramètre fonctionnel $m(g, \cdot)f(\cdot)$, mais aussi et surtout de la condition de mélange imposée au processus, comme on peut le déduire de Rosenblatt [91] pg. 84-87.

Démonstration du Théorème 1.4.2: On a

$$\begin{aligned} & (\phi_n(x_j) - \phi(x_j); j = 1, \dots, T) & (1.4.5) \\ & = (\Phi(\bar{g}_n(x_j)) - \Phi(m(g, x_j)f(x_j)); j = 1, \dots, T) \\ & = (\Phi(\bar{g}_n(x_1)), \dots, \Phi(\bar{g}_n(x_T))) - (\Phi(m(g, x_1)f(x_1)), \dots, \Phi(m(g, x_T)f(x_T))) \\ & = \Phi^*(\bar{g}_n(x_1), \dots, \bar{g}_n(x_T)) - \Phi^*(m(g, x_1)f(x_1), \dots, m(g, x_T)f(x_T)), \end{aligned}$$

où Φ^* définie par $\Phi^*(y_1, \dots, y_T) = (\Phi(y_1), \dots, \Phi(y_T))$, pour $y_1, \dots, y_T \in \mathbb{R}^p$, est différentiable en $(m(g, x_j)f(x_j); j = 1, \dots, T)$, puisque par hypothèse Φ est différentiable dans les points $m(g, x_1)f(x_1), \dots, m(g, x_T)f(x_T)$.

a) D'après le Théorème 1.3.14 a), on conclut, en tenant compte de (1.4.5), que le vecteur (1.4.3) converge, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de matrice de covariance (cf. [68], Corollaire 7, pg. 167)

$$\begin{aligned} & \frac{\partial \Phi^*}{\partial y} (m(g, x_j)f(x_j); j = 1, \dots, T) \Sigma \left(\frac{\partial \Phi^*}{\partial y} \right)' (m(g, x_j)f(x_j); j = 1, \dots, T) \\ & = \left[\frac{\partial \Phi}{\partial y} (m(g, x_j)f(x_j)) \Sigma_{i,j} \left(\frac{\partial \Phi}{\partial y} \right)' (m(g, x_j)f(x_j)) \right]_{i,j=1,\dots,T}. \end{aligned}$$

b) De façon analogue, on conclut d'après le Théorème 1.3.14 b) et (1.4.5), que le vecteur (1.4.3) converge, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne

$$\left(\Delta^m(m(g, \cdot)f(\cdot), K)(x_j) \begin{bmatrix} c_1^m & \dots & 0 \\ 0 & & c_p^m \end{bmatrix}; j = 1, \dots, T \right)$$

$$\begin{aligned}
& \times \left(\frac{\partial \Phi^*}{\partial y} \right)' (m(g, x_j) f(x_j); j = 1, \dots, T) \\
= & \left(\Delta^m(m(g, \cdot) f(\cdot), K)(x_j) \begin{bmatrix} c_1^m & \cdot & 0 \\ 0 & \cdot & c_p^m \end{bmatrix} \right) \left(\frac{\partial \Phi}{\partial y} \right)' (m(g, x_j) f(x_j); j = 1, \dots, T),
\end{aligned}$$

et de matrice de covariance définie par (1.4.4). ■

1.4.3 Quelques commentaires

On rassemble dans la suite quelques commentaires sur les résultats précédents:

(1) En suivant l'observation faite à propos de la Proposition 1.3.4, les Théorèmes 1.4.1 et 1.4.2 restent valables si on remplace les hypothèses $g \in \mathcal{P}_p(x)$ et $g \in \mathcal{N}_p(x_1, \dots, x_T)$, par $1 \in \mathcal{P}_p(x)$ et $1 \in \mathcal{N}_p(x_1, \dots, x_T)$, respectivement, et si on demande une plus grande vitesse de convergence vers zéro des coefficients de mélange, en admettant que le processus $(X_i, i \in \mathbb{Z})$ est de type $Alpha(\frac{\delta}{4+\delta})$. On retrouve ainsi les conditions utilisées dans Robinson [81] pg. 189.

(2) L'observation précédente est également applicable au Théorème 1.4.1 (resp. 1.4.2), si l'une des conditions $\sum_{i=1}^{+\infty} \rho(i) < \infty$ ou $\sum_{i=1}^{+\infty} \phi^{\frac{1+\delta}{2+\delta}}(i) < \infty$ (resp. $\rho(i) = o(i^{-2})$ ou $\phi(i) = o(i^{-2})$), est satisfaite. Ces conditions sur les coefficients de ρ et ϕ -mélange, sont conséquence des différentes majorations qu'on peut obtenir pour $|E(XY) - E(X)E(Y)|$, où $X \in F_{-\infty}^0$ et $Y \in F_i^{+\infty}$, c'est-à-dire, des majorations pour les covariances entre des variables appartenant au passé et au futur du processus observé (cf. [6] pour le coefficient ϕ).

(3) Lorsqu'il existe $R > 0$ tel que $P(\|g(V_0)\| < R) = 1$, on peut prendre δ aussi grand que l'on veut: il suffit alors que la condition " $(X_i, i \in \mathbb{Z})$ est de type $Alpha(1)$ ", soit satisfaite au lieu de la condition " $(X_i, i \in \mathbb{Z})$ est de type $Alpha(\frac{\delta}{2+\delta})$ ".

(4) En utilisant formellement les notations introduites le long de ce chapitre, si $((V_i, W_i), i \in \mathbb{Z})$ est un processus stochastique fortement stationnaire à valeurs dans $\mathbb{R}^s \times \mathbb{R}^d$, les Théorèmes 1.4.1 et 1.4.2 restent valables si la condition de mélange est maintenant imposée au processus $((V_i, W_i), i \in \mathbb{Z})$, et si $\mathcal{N}_p(x_1, \dots, x_T)$, pour $p \in \mathbb{N}^*$ et $x_1, \dots, x_T \in \mathbb{R}^d$, désigne l'ensemble des applications g de \mathbb{R}^s dans \mathbb{R}^p , pour lesquelles il existe $M > 0$ et $\delta > 0$ tels que

$$E[\|g(V_i)\| \|g(V_0)\| \mid W_i = \cdot, W_0 = \cdot] f_{(W_i, W_0)}(\cdot, \cdot) \in \bigcap_{a,b=1}^T \mathcal{B}_{(x_a, x_b)}(M, \delta), \text{ si } i \geq 1.$$

1.5 Applications

Les résultats que nous présentons dans les paragraphes suivants sur l'estimation de la densité de probabilité, de la moyenne conditionnelle, de la variance conditionnelle et des coefficients d'asymétrie et de kurtosis conditionnels, sont des conséquences du Théorème 1.4.2. Tenant compte du Théorème 1.4.1, nous pouvons obtenir des estimateurs convergents pour les différentes variances asymptotiques, et donc construire des régions ponctuelles asymptotiques de confiance pour les divers paramètres d'intérêt. Par simplicité d'exposition nous ne considérons que des hypothèses de différentiabilité sur les divers paramètres. Dans le cas de l'estimation de la densité ou de la régression, telles régions ponctuelles pourront respectivement, permettre de détecter la nature non gaussienne éventuelle du processus observé, ou certaines caractéristiques non linéaires. Si l'estimation de la variance conditionnelle peut permettre de détecter la présence éventuelle d'un effet hétéroscédastique dans la série des observations, dans le cas de l'estimation des coefficients d'asymétrie et de kurtosis conditionnels, que nous savons être égaux, respectivement, à 0 et à 3 pour toute la famille de lois conditionnelles normales, nous pouvons tester deux conséquences de l'hypothèse de normalité de la distribution conditionnelle, en comparant à 0 et à 3, les coefficients estimés. Une telle démarche peut, par exemple, être utilisée pour effectuer des recherches sur l'adéquation de l'hypothèse de normalité conditionnelle sur les erreurs ARCH(1) d'un processus autorégressif d'ordre un (cf. Gouriéroux [35] pg. 35-41). Remarquons que dans ce dernier cas, le fait qu'il ne s'agisse pas de vrais résidus demande des hypothèses supplémentaires sur le noyau et sur la suite h_n que celles considérées par la suite. Une procédure de ce type est illustrée dans le Chapitre 2.

Dans cette section on notera par h_n un réel strictement positif, dépendant de n , et par K un noyau sur $\mathbb{R}^{d'}$. On notera encore par $\xrightarrow[n \rightarrow +\infty]{d} N(0, \sigma^2)$ la convergence vers la loi normale de moyenne zéro et de variance σ^2 .

1.5.1 Estimation de la densité

Considérons l'estimateur par noyau de la densité de probabilité f de W_0 , évaluée au point $x \in \mathbb{R}^{d'}$, défini, pour $n \in \mathbb{N}^*$, par

$$f_n(x) = \frac{1}{nh_n^{d'}} \sum_{i=1}^n K\left(\frac{x - W_i}{h_n}\right). \quad (1.5.1)$$

Théorème 1.5.2 Soit x un point de $\mathbb{R}^{d'}$. On suppose:

- i) $f \in \mathcal{D}_x(m)$, pour un $m \geq 1$;
- ii) $1 \in \mathcal{N}_1(x)$;
- iii) $(X_i, i \in \mathbb{Z})$ est de type Alpha(1);
- iv) $nh_n^{d'+2m} \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty$, $n \rightarrow +\infty$;
- v) $K \in \mathcal{K}_c^{d'}(m)$.

Alors

$$\sqrt{nh_n^{d'}} (f_n(x) - f(x)) \xrightarrow[n \rightarrow +\infty]{d} N\left(0, f(x) \int_{\mathbb{R}^{d'}} K^2(u) du\right).$$

Si on prend $W_i = X_i$ dans (1.5.1), $f_n(\cdot)$ est l'estimateur à noyau de la densité marginale du processus observé $(X_i, i \in \mathbb{Z})$. Dans ce cas, la condition $1 \in \mathcal{N}_1(x)$ se réduit à: il existe $M > 0$ et $\delta > 0$ tels que $f_{(X_i, X_0)} \in \mathcal{B}_{(x, x)}(M, \delta)$, pour $i \geq 1$.

Si les variables $(X_i, i \in \mathbb{Z})$ sont indépendantes, l'hypothèse iii) est satisfaite, et l'hypothèse ii) est une conséquence de l'hypothèse i).

1.5.2 Estimation de la moyenne conditionnelle

Soit g une application réelle mesurable sur $(\mathbb{R}^{d'})^s$. Considérons pour $x \in \mathbb{R}^{d'}$, l'estimateur de la moyenne conditionnelle $m(g, x)$, défini par

$$m_n(g, x) = \left\{ \frac{1}{nh_2^{d'}(n)} \sum_{i=1}^n K_2\left(\frac{x - W_i}{h_2(n)}\right) \right\}^{-1} \left\{ \frac{1}{nh_1^{d'}(n)} \sum_{i=1}^n g(V_i) K_1\left(\frac{x - W_i}{h_1(n)}\right) \right\}, \quad (1.5.3)$$

où pour $r = 1, 2$, $h_r(n) = c_r h_n$, avec $c_r > 0$, et K_r est un noyau sur $\mathbb{R}^{d'}$.

Théorème 1.5.4 Soit x un point de $\mathbb{R}^{d'}$ où $f(\cdot)$ est non nulle. On suppose:

- i) $E|g(V_0)|^{2+\delta} < \infty$ pour un $\delta > 0$;
- ii) $f(\cdot)$, $m(g, \cdot) \in \mathcal{D}_x(m)$, pour un $m \geq 1$;
- iii) $m(g^2, \cdot) \in \mathcal{C}_x$;
- iv) $m(|g|^{2+\delta}, \cdot) \in \mathcal{B}_x$;
- v) $(1, g) \in \mathcal{N}_2(x)$;
- vi) $(X_i, i \in \mathbb{Z})$ est de type Alpha($\frac{\delta}{2+\delta}$);
- vii) $nh_n^{d'+2m} \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty$, $n \rightarrow +\infty$;
- viii) $K_1, K_2 \in \mathcal{K}_c^{d'}(m)$.

Alors

$$\sqrt{nh_n^{d'}} (m_n(g, x) - m(g, x)),$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de variance

$$\frac{1}{f(x)} \left\{ m(g^2, x) \int_{\mathbb{R}^{d'}} K_{1,c_1}^2(u) du + m^2(g, x) \int_{\mathbb{R}^{d'}} [K_{2,c_2}^2(u) - 2K_{1,c_1}(u)K_{2,c_2}(u)] du \right\},$$

où $K_{r,c_r}(\cdot) = K_r(\cdot/c_r)/c_r^{d'}$.

La forme générale prise pour l'estimateur (1.5.3) avec K_1 et $h_1(n)$ éventuellement différents de K_2 et $h_2(n)$, respectivement, nous permet de conclure que d'un point de vue asymptotique, nous pouvons, dans l'estimation de $m(g, \cdot)$, nous restreindre aux estimateurs construits avec un seul noyau et une seule suite. En effet, en notant $\sigma^2(K_{1,c_1}, K_{2,c_2})$ la variance asymptotique précédente, nous concluons facilement que $\sigma^2(K_{1,c_1}, K_{1,c_1}) \leq \sigma^2(K_{1,c_1}, K_{2,c_2})$. Le choix des noyaux doit être donc effectué de façon que $K_{1,c_1} = K_{2,c_2}$. Nous justifions ainsi cette pratique courante. Si on prend $K_1 = K_2 = K$ et $h_1(n) = h_2(n) = h_n$, nous obtenons le résultat classique

$$\sqrt{nh_n^{d'}} (m_n(g, x) - m(g, x)) \xrightarrow[n \rightarrow +\infty]{d} N\left(0, \frac{\text{Var}[g(V_0) | W_0 = x]}{f(x)} \int_{\mathbb{R}^{d'}} K^2(u) du\right).$$

1.5.3 Estimation de la variance conditionnelle

Soit g une application réelle mesurable sur $(\mathbb{R}^{d'})^s$ et considérons pour $x \in \mathbb{R}^{d'}$, l'estimateur de la variance conditionnelle $\text{Var}[g(V_0) | W_0 = x]$ défini par

$$v_n(g, x) = \left\{ \frac{1}{nh_n^{d'}} \sum_{i=1}^n g^2(V_i) K\left(\frac{x - W_i}{h_n}\right) \right\} / f_n(x) - m_n^2(g, x),$$

où $f_n(x)$ est défini par (1.5.1) et $m_n(g, x)$ est défini par (1.5.3) avec $K_1 = K_2 = K$ et $h_1(n) = h_2(n) = h_n$.

Théorème 1.5.5 Soit x un point de $\mathbb{R}^{d'}$ où $f(\cdot)$ est non nulle. On suppose:

- i) $E|g(V_0)|^{4+\delta} < \infty$ pour un $\delta > 0$;
- ii) $f(\cdot)$, $m(g, \cdot)$, $m(g^2, \cdot) \in \mathcal{D}_x(m)$, pour un $m \geq 1$;
- iii) $m(g^3, \cdot)$, $m(g^4, \cdot) \in \mathcal{C}_x$;
- iv) $m(|g|^{4+\delta}, \cdot) \in \mathcal{B}_x$;
- v) $(1, g, g^2) \in \mathcal{N}_3(x)$;
- vi) $(X_i, i \in \mathbb{Z})$ est de type $\text{Alpha}(\frac{\delta}{4+\delta})$;

vii) $nh_n^{d'+2m} \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty$, $n \rightarrow +\infty$;

viii) $K \in \mathcal{K}_c^{d'}(m)$.

Alors

$$\sqrt{nh_n^{d'}} (v_n(g, x) - \text{Var}[g(V_0) | W_0 = x]),$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de variance

$$\frac{1}{f(x)} \int_{\mathbb{R}^{d'}} K^2(u) du \text{Var}[(g(V_0) - m(g, x))^2 | W_0 = x].$$

1.5.4 Estimation des coefficients d'asymétrie conditionnelle et de kurtosis conditionnelle

Si $(X_i, i \in \mathbb{Z})$ est un processus réel, les coefficients d'asymétrie et de kurtosis conditionnels sont définis pour $x \in \mathbb{R}^{d'}$, avec $d' = q + 1$, respectivement par

$$\gamma_1(x) = \frac{E[(X_1 - E[X_1 | W_0 = x])^3 | W_0 = x]}{E^{3/2}[(X_1 - E[X_1 | W_0 = x])^2 | W_0 = x]} \text{ et}$$

$$\gamma_2(x) = \frac{E[(X_1 - E[X_1 | W_0 = x])^4 | W_0 = x]}{E^2[(X_1 - E[X_1 | W_0 = x])^2 | W_0 = x]}.$$

Comme application du Théorème 1.4.2, nous allons énoncer les propriétés de normalité asymptotique d'estimateurs de ces paramètres. Etant donnée la complexité des calculs dans le cas général, on se limite ici à la présentation du cas où $E[X_1 | W_0 = x] = 0$. Nous pouvons donc écrire

$$\gamma_1(x) = \Psi(f(x), E[X_1^2 | W_0 = x]f(x), E[X_1^3 | W_0 = x]f(x)) \text{ et}$$

$$\gamma_2(x) = \Phi(f(x), E[X_1^2 | W_0 = x]f(x), E[X_1^4 | W_0 = x]f(x)),$$

où $\Psi(u, v, w) = u^{1/2}w/v^{3/2}$ et $\Phi(u, v, w) = uw/v^2$, pour $u, v, w \in \mathbb{R}$ et $v \neq 0$.

Nous considérons les estimateurs à noyau des paramètres fonctionnels précédents définis par

$$\gamma_{1,n}(x) = \frac{\left(\sum_{j=1}^n K\left(\frac{x-W_j}{h_n}\right)\right)^{1/2} \sum_{i=1}^n X_{i+1}^3 K\left(\frac{x-W_i}{h_n}\right)}{\left(\sum_{i=1}^n X_{i+1}^2 K\left(\frac{x-W_i}{h_n}\right)\right)^{3/2}} \text{ et}$$

$$\gamma_{2,n}(x) = \frac{\sum_{j=1}^n K\left(\frac{x-W_j}{h_n}\right) \sum_{i=1}^n X_{i+1}^4 K\left(\frac{x-W_i}{h_n}\right)}{\left(\sum_{i=1}^n X_{i+1}^2 K\left(\frac{x-W_i}{h_n}\right)\right)^2}.$$

On notera par *id* l'application identité dans \mathbb{R} .

Théorème 1.5.6 Soit x un point de $\mathbb{R}^{d'}$ où $f(\cdot)$ et $E[X_1^2|W_0 = \cdot]$ sont non nulles. On suppose:

- i) $E|X_1|^{6+\delta} < \infty$ pour un $\delta > 0$;
- ii) $f(\cdot), E[X_1^2|W_0 = \cdot], E[X_1^3|W_0 = \cdot] \in \mathcal{D}_x(m)$, pour un $m \geq 1$;
- iii) $E[X_1^5|W_0 = \cdot], E[X_1^6|W_0 = \cdot] \in \mathcal{C}_x$;
- iv) $E[|X_1|^{6+\delta}|W_0 = \cdot] \in \mathcal{B}_x$;
- v) $(1, id^2, id^3) \in \mathcal{N}_3(x)$;
- vi) $(X_i, i \in \mathbb{Z})$ est de type $Alpha(\frac{\delta}{6+\delta})$;
- vii) $nh_n^{d'+2m} \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty, n \rightarrow +\infty$;
- viii) $K \in \mathcal{K}_c^{d'}(m)$.

Alors

$$\sqrt{nh_n^{d'}} (\gamma_{1,n}(x) - \gamma_1(x)),$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de variance

$$\frac{1}{f(x)} \int_{\mathbb{R}^{d'}} K^2(u) du \left\{ \frac{E[X_1^6|W_0 = x]}{E^3[X_1^2|W_0 = x]} - 3 \frac{E[X_1^5|W_0 = x]}{E^{5/2}[X_1^2|W_0 = x]} \gamma_1(x) + \frac{1}{4} (9\gamma_2(x) - 1) \gamma_1^2(x) \right\}.$$

Théorème 1.5.7 Soit x un point de $\mathbb{R}^{d'}$ où $f(\cdot)$ et $E[X_1^2|W_0 = \cdot]$ sont non nulles. On suppose:

- i) $E|X_1|^{8+\delta} < \infty$ pour un $\delta > 0$;
- ii) $f(\cdot), E[X_1^2|W_0 = \cdot], E[X_1^4|W_0 = \cdot] \in \mathcal{D}_x(m)$, pour un $m \geq 1$;
- iii) $E[X_1^6|W_0 = \cdot], E[X_1^8|W_0 = \cdot] \in \mathcal{C}_x$;
- iv) $E[|X_1|^{8+\delta}|W_0 = \cdot] \in \mathcal{B}_x$;
- v) $(1, id^2, id^4) \in \mathcal{N}_3(x)$;
- vi) $(X_i, i \in \mathbb{Z})$ est de type $Alpha(\frac{\delta}{8+\delta})$;
- vii) $nh_n^{d'+2m} \rightarrow 0$ et $nh_n^{d'} \rightarrow +\infty, n \rightarrow +\infty$;
- viii) $K \in \mathcal{K}_c^{d'}(m)$.

Alors

$$\sqrt{nh_n^{d'}} (\gamma_{2,n}(x) - \gamma_2(x)),$$

converge en loi, lorsque $n \rightarrow +\infty$, vers la loi normale de moyenne zéro et de variance

$$\frac{1}{f(x)} \int_{\mathbb{R}^{d'}} K^2(u) du \left\{ \frac{E[X_1^8|W_0 = x]}{E^4[X_1^2|W_0 = x]} - 4 \frac{E[X_1^6|W_0 = x]}{E^3[X_1^2|W_0 = x]} \gamma_2(x) - \gamma_2^2(x) + 4\gamma_2^3(x) \right\}.$$

Pour $x \in \mathbb{R}^{d'}$ fixé, les résultats suivants nous permettent de fonder des tests de normalité pour la loi conditionnelle de X_1 sachant que $W_0 = x$. Soit $f_n(x)$ défini par (1.5.1).

Corollaire 1.5.8 *Dans les conditions du Théorème 1.5.6, et sous l'hypothèse de normalité de la loi de X_1 conditionnée par $W_0 = x$, on a*

$$\sqrt{nh_n^{d'}} \frac{\gamma_{1,n}(x)}{\sqrt{15 \int_{\mathbb{R}^{d'}} K^2(u) du}} \sqrt{f_n(x)} \xrightarrow[n \rightarrow +\infty]{d} N(0, 1).$$

Corollaire 1.5.9 *Dans les conditions du Théorème 1.5.7, et sous l'hypothèse de normalité de la loi de X_1 conditionnée par $W_0 = x$, on a*

$$\sqrt{nh_n^{d'}} \frac{\gamma_{2,n}(x) - 3}{\sqrt{24 \int_{\mathbb{R}^{d'}} K^2(u) du}} \sqrt{f_n(x)} \xrightarrow[n \rightarrow +\infty]{d} N(0, 1).$$

Chapitre 2

M-estimateurs à noyau: diagnostic non paramétrique de modèles structurels

Nous présentons dans ce chapitre les résultats de Gouriéroux, Monfort et Tenreiro [37]. Nous considérons une classe de M-estimateurs à noyau, que nous proposons pour effectuer des recherches non paramétriques de spécifications paramétriques. Nous donnons les propriétés ponctuelles de convergence presque sûre et normalité asymptotique de tels estimateurs, et expliquons comment les utiliser pour effectuer de telles recherches. Nous décrivons aussi quelques problèmes où les techniques développées peuvent être appliquées.

2.1 Exemple introductif

L'exemple suivant motivera la définition d'une classe d'estimateurs fonctionnels désignés par M-estimateurs à noyau, et illustre les applications de ces estimateurs au diagnostic non paramétrique de modèles structurels.

Supposons que nous voulons tester l'adéquation d'un modèle de régression $Y = a_0X + b_0 + U$, avec $E[U|X] = 0$ (a_0 et b_0 sont des nombres réels connus), ou de façon équivalente, tester l'hypothèse

$$H_0 : E[Y | X = x] = a_0x + b_0, \forall x \in \mathbb{R},$$

quand on observe n copies (X_i, Y_i) , $i = 1, \dots, n$, de la variable aléatoire (X, Y) .

En considérant pour $x \in \mathbb{R}$ et $\theta \in \Theta \subset \mathbb{R}$,

$$L_\infty(x; \theta) = E[(Y - (a_0X + b_0) - \theta)^2 | X = x]f_X(x),$$

où f_X désigne la densité de probabilité de X , l'hypothèse H_0 peut s'écrire de façon équivalente

$$H_0 : \theta_\infty(x) = \text{Arg min}_\theta L_\infty(x; \theta) = 0, \forall x \in \mathbb{R}.$$

En tenant compte du fait que, sous certaines conditions de régularité, $L_\infty(x; \theta)$ peut être estimé, pour chaque θ et $x \in \mathbb{R}$, par

$$L_n(x; \theta) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right) (Y_i - a_0X_i - b_0 - \theta)^2,$$

où (h_n) est une suite de nombres réels strictement positifs convergent vers zéro, lorsque n tend vers l'infini, et K est un noyau sur \mathbb{R} , il est donc naturel de proposer

$$\hat{\theta}_n(x) = \text{Arg min}_\theta L_n(x; \theta),$$

comme estimateur du paramètre d'intérêt $\theta_\infty(x)$, et de fonder un diagnostic sur H_0 d'après la comparaison à zéro de cet estimateur, qu'on désigne par M-estimateur à noyau.

De façon analogue, un diagnostic sur l'hypothèse d'adéquation d'un modèle linéaire de régression, ou de façon équivalente sur l'hypothèse

$$H'_0 : \exists a, b \in \mathbb{R} : E[Y | X = x] = ax + b, \forall x \in \mathbb{R},$$

peut être élaboré à partir d'une procédure en deux étapes. Dans une première étape d'estimation paramétrique on prend le modèle de régression linéaire $Y = aX + b + U$, avec $E[U|X] = 0$, et on considère les estimateurs des moindres carrés \hat{a}_n et \hat{b}_n des paramètres a et b . Dans une deuxième étape d'estimation non paramétrique, on considère pour $x \in \mathbb{R}$, le problème d'optimisation considéré ci-dessus mais où les paramètres a_0 et b_0 sont remplacés par les estimateurs obtenus dans la première étape d'estimation

$$\text{min}_\theta \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right) (Y_i - \hat{a}_nX_i - \hat{b}_n - \theta)^2.$$

Nous finissons cette deuxième étape d'estimation non paramétrique en comparant à zéro la solution $\hat{\theta}_n(\cdot)$ de ce problème. Dans le cas présent, $\theta_\infty(\cdot)$ est défini, pour $x \in \mathbb{R}$, par

$$\begin{aligned} \theta_\infty(x) &= \text{Arg min}_\theta E[(Y - (a_\infty X - b_\infty) - \theta)^2 | X = x]f_X(x) \\ &= E[Y|X = x] - a_\infty x - b_\infty, \end{aligned}$$

où $a_\infty = \lim_n \widehat{a}_n$ et $b_\infty = \lim_n \widehat{b}_n$, presque sûrement.

En interprétant $\theta_\infty(\cdot)$ comme un résidu fonctionnel relatif au modèle paramétrique considéré sous l'hypothèse nulle, le graphique de $\widehat{\theta}_n(\cdot)$ est donc un graphique *résidu* \times *variable* qui concerne le modèle structurel paramétré, et qui est construit par une démarche non paramétrique.

Pour concrétiser la démarche que nous venons de décrire, nous introduisons ensuite et dans un cadre général la définition d'un M-estimateur à noyau et nous décrivons ses propriétés asymptotiques, en spécifiant en particulier, la loi asymptotique de la variable aléatoire $\widehat{\theta}_n(x) - \theta_\infty(x)$, pour chaque x .

2.2 M-estimateur à noyau: définition et propriétés asymptotiques

2.2.1 Définition

Soient Θ une partie compacte de \mathbb{R}^p munie de sa tribu borélienne, \mathcal{Y} un espace mesurable, et \mathcal{A} un sous-ensemble ouvert de \mathbb{R}^q muni de sa tribu borélienne. Soient Y_0 une variable aléatoire à valeurs dans \mathcal{Y} , et s et α_∞ des points fixés en \mathbb{R}^d et \mathcal{A} , respectivement.

Soient S une application mesurable de $\mathcal{Y} \times \mathcal{A}$ dans \mathbb{R}^d , et $\psi(\cdot, \cdot; s; \cdot)$ une application de $\mathcal{Y} \times \mathcal{A} \times \Theta$ dans \mathbb{R} telle que:

- i) l'application partielle $\psi(\cdot, \cdot; s; \theta) : (Y, \alpha) \rightarrow \psi(Y, \alpha; s; \theta)$ est mesurable pour tout $\theta \in \Theta$;
- ii) l'application partielle $\psi(Y, \alpha; s; \cdot) : \theta \rightarrow \psi(Y, \alpha; s; \theta)$ est continue dans Θ , pour tout $(Y, \alpha) \in \mathcal{Y} \times \mathcal{A}$.

Les hypothèses précédentes sur S et ψ sont désignées par \mathbf{C}_0 .

Pour $\theta \in \Theta$, admettons que l'application $L_\infty(\cdot; s; \theta)$, définie pour $z \in \mathbb{R}^d$ par

$$L_\infty(\cdot; s; \theta) : z \rightarrow L_\infty(z; s; \theta) = E[\psi(Y_0, \alpha_\infty; s; \theta) \mid S(Y_0, \alpha_\infty) = z] f_S(z),$$

est continue en s , où $E|\psi(Y_0, \alpha_\infty; s; \theta)| < \infty$, $E[\psi(Y_0, \alpha_\infty; s; \theta) \mid S(Y_0, \alpha_\infty) = z]$ est une version de l'espérance conditionnelle de $\psi(Y_0, \alpha_\infty; s; \theta)$ sachant $S(Y_0, \alpha_\infty) = z$, et f_S désigne la densité de probabilité de la variable aléatoire $S(Y_0, \alpha_\infty)$, dont la loi est supposée absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d . On note par simplicité $L_\infty(s; s; \theta) = L_\infty(s; \theta)$.

Définition du paramètre fonctionnel d'intérêt $\theta_\infty(\cdot)$

Désignons par $\theta_\infty(\cdot)$ le paramètre fonctionnel défini pour $s \in \mathbb{R}^d$, comme la solution, qu'on suppose unique (**condition d'identifiabilité asymptotique**), du problème d'optimisation

$$\min_{\theta \in \Theta} L_\infty(s; \theta), \quad (2.2.1)$$

c'est-à-dire,

$$\theta_\infty(s) = \text{Arg} \min_{\theta \in \Theta} L_\infty(s; \theta). \quad (2.2.2)$$

Définition d'un M-estimateur à noyau

Soient $(Y_i, i \in \mathbb{Z})$ un processus stochastique fortement stationnaire à valeurs dans \mathcal{Y} , K un noyau sur \mathbb{R}^d , (h_n) une suite de nombres réels strictement positifs et $(\hat{\alpha}_n)$ une suite de variables aléatoires à valeurs dans \mathcal{A} . Alors, d'après le Lemme 2 de Jennrich [58] pg. 637, il existe, pour $s \in \mathbb{R}^d$ et $n \in \mathbb{N}^*$, une variable aléatoire $\hat{\theta}_n(s)$ à valeurs dans Θ telle que

$$\hat{\theta}_n(s) = \text{Arg} \min_{\theta \in \Theta} L_n(s; \theta),$$

où

$$L_n(s; \theta) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \psi(Y_i, \hat{\alpha}_n; s; \theta). \quad (2.2.3)$$

L'estimateur $\hat{\theta}_n(\cdot)$, défini pour $s \in \mathbb{R}^d$, comme une des solutions du problème d'optimisation

$$\min_{\theta \in \Theta} L_n(s; \theta), \quad (2.2.4)$$

est appelé M-estimateur à noyau. On dira aussi que $\hat{\theta}_n(\cdot)$ est un M-estimateur à noyau associé à $L_n(\cdot; \theta)$.

Nous étudions dans les paragraphes suivants les propriétés ponctuelles de convergence presque sûre et de normalité asymptotique de cet estimateur.

2.2.2 Convergence presque sûre

Nous nous intéressons maintenant à la convergence ponctuelle presque sûre d'un M-estimateur à noyau $\hat{\theta}_n(\cdot)$. Pour $s \in \mathbb{R}^d$, cette convergence sera obtenue en utilisant les arguments classiques (cf. White [110] et Gouriéroux et Monfort [36]), à partir des propriétés asymptotiques des

estimateurs à noyau appliquées à la fonction objectif $L_n(s; \theta)$ du problème d'optimisation à distance finie (2.2.4). Ainsi, nous montrerons la convergence uniforme presque sûre (p.s.) de $L_n(s; \theta)$ vers la fonction objectif $L_\infty(s; \theta)$ du problème d'optimisation asymptotique (2.2.1), ce qui permet de conclure que la solution $\hat{\theta}_n(s)$ du problème à distance finie converge vers la solution unique $\theta_\infty(s)$ du problème asymptotique (cf. Gouriéroux et Monfort [36], vol. 2, pg. 431-432). Notons que le caractère d'uniformité concerne le paramètre θ et non s , comme il est usuel dans l'estimation non paramétrique.

Les hypothèses générales suivantes concernent le processus observé, la suite (h_n) , le noyau K et la suite de variables aléatoires $(\hat{\alpha}_n)$.

Hypothèses générales (GA)

Hypothèses sur le processus

$(Y_i, i \in \mathbb{Z})$ est un processus stochastique fortement stationnaire géométriquement α -mélangeant. Nous supposons donc qu'il existe $C > 0$ et $\rho \in]0, 1[$ tels que

$$\alpha(i) \leq C\rho^i, \text{ pour } i \in \mathbb{N}^*,$$

où $\alpha(i)$ est le coefficient de mélange d'ordre i du processus $Y_i, i \in \mathbb{Z}$, défini par (1.2.1).

Hypothèses sur le noyau K

K est un noyau différentiable sur \mathbb{R}^d , qu'on suppose borné à support borné. De plus, nous admettons que les dérivées partielles, $\frac{\partial K}{\partial u_j}, j = 1, \dots, d$, sont bornées.

Hypothèses sur (h_n)

On suppose que

$$h_n \rightarrow 0, \text{ lorsque } n \rightarrow +\infty,$$

et qu'il existe $\beta \in]0, 1[$ tel que

$$\frac{n^{(1-\beta)/2} h_n^d}{\log n} \rightarrow +\infty, \text{ lorsque } n \rightarrow +\infty.$$

Hypothèses sur $(\hat{\alpha}_n)$

$\hat{\alpha}_n$ converge presque sûrement vers une constante α_∞ , lorsque n tend vers l'infini, et

$$\hat{\alpha}_n - \alpha_\infty = o(h_n^{d+1}).$$

Nous avons besoin de considérer des conditions de régularité sur les applications ψ , S et L_∞ , autres que la condition \mathbf{C}_0) introduite dans §2.2.1. De la même façon que \mathbf{C}_0), ces conditions concernent le point s de \mathbb{R}^d , fixé au départ. Elles seront donc désignées par “hypothèses locales”. Si $x = \alpha_1 e_1 + \dots + \alpha_N e_N$ est un élément d’un espace vectoriel réel de dimension finie dont une base est $\{e_1, \dots, e_N\}$, nous désignons $\|x\|$ une norme quelconque du vecteur $(\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N$.

Hypothèses locales

\mathbf{C}_1) Il existe $\delta > 0$ tel que $E \left[\sup_{\theta \in \Theta} |\psi(Y_0, \alpha_\infty; s; \theta)| \right]^{\frac{2}{\beta} + \delta} < \infty$, où β est le paramètre intervenant dans les hypothèses sur (h_n) , et de plus

$$E \left[\left| \sup_{\theta \in \Theta} |\psi(Y_0, \alpha_\infty; s; \theta)|^2 \mid S(Y_0, \alpha_\infty) = \cdot \right. \right] f_S(\cdot) \in \mathcal{B}_s;$$

\mathbf{C}_2) Pour tout $\epsilon > 0$, il existe un voisinage $V(s)$ de s tel que

$$\sup_{z \in V(s)} \sup_{\theta \in \Theta} |L_\infty(z; s; \theta) - L_\infty(s; s; \theta)| < \epsilon.$$

Nous admettons aussi que les applications partielles $\theta \rightarrow \psi(Y, \alpha_\infty; s; \theta)$, $\alpha \rightarrow \psi(Y, \alpha; s; \theta)$ et $\alpha \rightarrow S(Y, \alpha)$ satisfont pour tout $Y \in \mathcal{Y}$ et $\theta \in \Theta$, les conditions de Lipschitz suivantes:

\mathbf{C}_3) Pour tout $\theta_1, \theta_2 \in \Theta$,

$$|\psi(Y, \alpha_\infty; s; \theta_1) - \psi(Y, \alpha_\infty; s; \theta_2)| \leq M(Y, \alpha_\infty; s) \|\theta_1 - \theta_2\|, \text{ où } E|M(Y_0, \alpha_\infty; s)| < \infty;$$

\mathbf{C}_4) Pour tout $\alpha \in \mathcal{A}$,

$$|\psi(Y, \alpha; s; \theta) - \psi(Y, \alpha_\infty; s; \theta)| \leq N_\psi(Y, \alpha_\infty; s; \theta) \|\alpha - \alpha_\infty\|, \text{ où } E \left[\sup_{\theta \in \Theta} N_\psi(Y_0, \alpha_\infty; s; \theta) \right] < \infty;$$

\mathbf{C}_5) Pour tout $\alpha \in \mathcal{A}$,

$$\|S(Y, \alpha) - S(Y, \alpha_\infty)\| \leq N(Y, \alpha_\infty) \|\alpha - \alpha_\infty\|, \text{ où } E \left[\sup_{\theta \in \Theta} |\psi(Y_0, \alpha_\infty; s; \theta)| N(Y_0, \alpha_\infty) \right] < \infty.$$

Remarquons que si Θ est convexe et l’application partielle $\theta \rightarrow \psi(Y, \alpha_\infty; s; \theta)$ est différentiable dans $\text{int}(\Theta)$ et continue dans Θ , la condition de Lipschitz donnée dans \mathbf{C}_3) est satisfaite avec (cf. Jennrich [58] pg. 638)

$$M(Y, \alpha_\infty; s) = \sup_{\theta \in \text{int}(\Theta)} \left\| \frac{\partial \psi}{\partial \theta}(Y, \alpha_\infty; s; \theta) \right\|.$$

De façon analogue, si \mathcal{A} est convexe et les applications partielles $\alpha \rightarrow \psi(Y, \alpha; s; \theta)$ et $\alpha \rightarrow S(Y, \alpha)$ sont différentiables dans \mathcal{A} , les conditions de Lipschitz $C_4)$ et $C_5)$ sont satisfaites avec

$$N_\psi(Y, \alpha_\infty; s; \theta) = \sup_{\alpha \in \mathcal{A}} \left\| \frac{\partial \psi}{\partial \alpha}(Y, \alpha; s; \theta) \right\| \text{ et } N(Y, \alpha_\infty) = \sup_{\alpha \in \mathcal{A}} \left\| \frac{\partial S}{\partial \alpha}(Y, \alpha) \right\|.$$

Les conditions imposées sur la vitesse de convergence vers zéro du coefficient de mélange et de la suite (h_n) sont liées à l'utilisation de l'inégalité exponentielle suivante (cf. Bosq [10] et Carbon [16]), dans le théorème de convergence qu'on établit ci-dessous pour un M-estimateur à noyau.

Lemme 2.2.5 *Soit $(Z_t, t \in \mathbb{Z})$ un processus réel, centré, géométriquement α -mélangeant, tel que*

$$\sup_{1 \leq t \leq n} |Z_t| \leq N_n < +\infty, \text{ pour } n \in \mathbb{N}^*.$$

Alors, $\forall \lambda \in]0, 1[$, $\exists n_0(\lambda) : \forall n \geq n_0(\lambda)$, $\forall \epsilon > 0$:

$$P\left(\left|\sum_{t=1}^n Z_t\right| > n\epsilon\right) \leq 3 \exp\left(-\frac{\epsilon}{14} \frac{n^{(1-\lambda)/2}}{N_n}\right).$$

Nous pouvons maintenant énoncer le théorème de convergence déjà cité.

Théorème 2.2.6 *Soit Θ une partie compacte de \mathbb{R}^p . Sous les hypothèses générales **(GA)**, étant donné un point s de \mathbb{R}^d pour lequel nous supposons valables les hypothèses locales **C₀**, ..., **C₅**, soit pour $\theta \in \Theta$, $L_n(s; \theta)$ donnée par (2.2.3). Si $\hat{\theta}_n(\cdot)$ est un M-estimateur à noyau associé à $L_n(\cdot; \theta)$, alors*

$$\lim_{n \rightarrow +\infty} \hat{\theta}_n(s) = \theta_\infty(s), \text{ presque sûrement,}$$

où $\theta_\infty(s)$ est défini par (2.2.2).

Dans les conditions du théorème précédent, et de la même façon que pour les M-estimateurs paramétriques, l'hypothèse de compacité sur Θ peut être remplacée par la condition: Θ est d'intérieur non vide et $\theta_\infty(s) \in \text{int}(\Theta)$.

Démonstration du Théorème 2.2.6: Soit $s \in \mathbb{R}^d$ et $\theta \in \Theta$ dans les conditions de l'énoncé. Considérons la décomposition

$$L_n(s; \theta) - L_\infty(s; \theta) = [L_n(s; \theta) - \tilde{L}_n(s; \theta)] + [\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta)] + [E\tilde{L}_n(s; \theta) - L_\infty(s; \theta)],$$

où

$$\begin{aligned}\tilde{L}_n(s; \theta) &= \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta), \text{ et} \\ L_\infty(\cdot; s; \theta) &= E[\psi(Y_0, \alpha_\infty; s; \theta) \mid S(Y_0, \alpha_\infty) = \cdot] f_S(\cdot).\end{aligned}$$

Chacun des termes de la décomposition précédente est étudié dans les points suivants.

A) Convergence uniforme de $E\tilde{L}_n(s; \cdot)$

Comme $h_n \rightarrow 0, n \rightarrow +\infty$, et K est à support borné, on conclut que si $W(s)$ est un voisinage quelconque de s , il existe $n_0 \in \mathbb{N}^*$ tel que, pour $n \geq n_0$ on a

$$\begin{aligned}\sup_{\theta \in \Theta} |E\tilde{L}_n(s; \theta) - L_\infty(s; \theta)| &\leq \sup_{\theta \in \Theta} \int_{\mathbb{R}^d} |L_\infty(s - yh_n; s; \theta) - L_\infty(s; s; \theta)| |K(y)| dy \\ &\leq \sup_{z \in W(s)} \sup_{\theta \in \Theta} |L_\infty(z; s; \theta) - L_\infty(s; s; \theta)| \int_{\mathbb{R}^d} |K(y)| dy.\end{aligned}$$

Ainsi, d'après l'hypothèse C₂), on obtient $\lim_{n \rightarrow +\infty} \sup_{\theta \in \Theta} |E\tilde{L}_n(s; \theta) - L_\infty(s; \theta)| = 0$.

B) Décomposition de $\tilde{L}_n(s; \cdot) - E\tilde{L}_n(s; \cdot)$

Pour $\theta \in \Theta$, considérons la décomposition suivante de $\tilde{L}_n(s; \theta)$,

$$\begin{aligned}\tilde{L}_n(s; \theta) &= \tilde{L}'_n(s; \theta) + \tilde{L}''_n(s; \theta) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta) \mathbb{I}_{|\psi(Y_i, \alpha_\infty; s; \theta)| \leq M_n} \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta) \mathbb{I}_{|\psi(Y_i, \alpha_\infty; s; \theta)| > M_n},\end{aligned}$$

où $M_n = n^\gamma$, avec $\gamma \in](\frac{2}{\beta} + \delta)^{-1}, (\frac{2}{\beta})^{-1}[$, et β et δ sont donnés dans les hypothèses sur (h_n) et dans C₁). Le choix de la suite M_n permet d'établir la convergence

$$\lim_{n \rightarrow +\infty} \sup_{\theta \in \Theta} |\tilde{L}''_n(s; \theta) - E\tilde{L}''_n(s; \theta)| = 0, \text{ presque sûrement (p.s.)}$$

En effet, d'après l'inégalité de Bienaymé-Chebyshev et l'hypothèse C₁), on a, pour $\epsilon > 0$

$$P\left(\sup_{\theta \in \Theta} |\psi(Y_n, \alpha_\infty; s; \theta)| > \epsilon M_n\right) \leq \epsilon^{-(\frac{2}{\beta} + \delta)} M_n^{-(\frac{2}{\beta} + \delta)} E\left[\sup_{\theta \in \Theta} |\psi(Y_0, \alpha_\infty; s; \theta)|\right]^{\frac{2}{\beta} + \delta},$$

et donc, on conclut d'après la définition de M_n avec $\gamma > (\frac{2}{\beta} + \delta)^{-1}$, et le théorème de Borel-Cantelli, que

$$\frac{\sup_{\theta \in \Theta} |\psi(Y_n, \alpha_\infty; s; \theta)|}{M_n} \rightarrow 0, n \rightarrow +\infty, \text{ p.s..}$$

Ainsi, comme la suite M_n est croissante, on obtient (cf. Mack et Silverman [64] pg. 408)

$$\limsup_n \frac{\sup_{j=1, \dots, n} \sup_{\theta \in \Theta} |\psi(Y_j, \alpha_\infty; s; \theta)|}{M_n} \leq 1, \text{ p.s.},$$

et donc

$$\sup_{\theta \in \Theta} |\tilde{L}_n''(s; \theta)| = 0, \text{ p.s., pour } n \text{ assez grand.}$$

D'autre part, et encore d'après l'hypothèse C_1), on a

$$\sup_{\theta \in \Theta} E|\tilde{L}_n''(s; \theta)| = O\left(\frac{1}{M_n}\right),$$

ce qui permet d'obtenir la convergence souhaitée.

C) Convergence uniforme presque sûre de $\tilde{L}_n(s; \cdot) - E\tilde{L}_n(s; \cdot)$ vers zéro

Considérons un recouvrement fini $\Theta_\ell, \ell = 1, \dots, N$, du compact Θ , dont les éléments sont de diamètre (par rapport à la norme du maximum) inférieur ou égal à $1/n$: $\text{diam}(\Theta_\ell) \leq 1/n$. Ce recouvrement peut être choisi de façon que $N < (n \text{diam}(\Theta) + 1)^p$.

D'après le point B), et pour n assez grand, on a

$$\begin{aligned} & \sup_{\theta \in \Theta} |\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta)| \\ &= \max_{\ell=1, \dots, N} \sup_{\theta \in \Theta_\ell} |\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta)| \\ &\leq \max_{\ell=1, \dots, N} \sup_{\theta \in \Theta_\ell} |\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta) - \tilde{L}_n(s; \theta_\ell) + E\tilde{L}_n(s; \theta_\ell)| \\ &\quad + \max_{\ell=1, \dots, N} |\tilde{L}_n'(s; \theta_\ell) - E\tilde{L}_n'(s; \theta_\ell)| + o(1), \end{aligned}$$

où $\theta_\ell \in \Theta_\ell$, pour $\ell = 1, \dots, N$.

D'après l'hypothèse C_3), on a pour $\theta_1, \theta_2 \in \Theta$,

$$|\tilde{L}_n(s; \theta_1) - \tilde{L}_n(s; \theta_2)| \leq \sup_{u \in \mathbb{R}^d} |K(u)| \frac{\|\theta_1 - \theta_2\|}{h_n^d} \frac{1}{n} \sum_{i=1}^n M(Y_i, \alpha_\infty; s),$$

et donc de l'application du théorème ergodique, et pour n assez grand, on a

$$\begin{aligned} & \sup_{\theta \in \Theta} |\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta)| \\ &\leq \max_{\ell=1, \dots, N} \sup_{\theta \in \Theta_\ell} O\left(\frac{\|\theta - \theta_\ell\|}{h_n^d}\right) + \max_{\ell=1, \dots, N} |\tilde{L}_n'(s; \theta_\ell) - E\tilde{L}_n'(s; \theta_\ell)| + o(1) \\ &= O\left(\frac{1}{nh_n^d}\right) + \max_{\ell=1, \dots, N} |\tilde{L}_n'(s; \theta_\ell) - E\tilde{L}_n'(s; \theta_\ell)| + o(1). \end{aligned}$$

Comme $nh_n^d \rightarrow +\infty, n \rightarrow +\infty$, la convergence presque sûre de $\sup_{\theta \in \Theta} |\tilde{L}_n(s; \theta) - E\tilde{L}_n(s; \theta)|$ vers zéro s'obtient de la convergence presque sûre de $\max_{\ell=1, \dots, N} |\tilde{L}'_n(s; \theta_\ell) - E\tilde{L}'_n(s; \theta_\ell)|$ vers zéro, qu'on établira en conséquence du Lemme 2.2.5 et d'après les hypothèses générales (GA).

Pour $\ell = 1, \dots, N$,

$$\tilde{L}'_n(s; \theta_\ell) - E\tilde{L}'_n(s; \theta_\ell) = \frac{1}{n} \sum_{i=1}^n Z_{i,n},$$

où pour $i = 1, \dots, n$,

$$\begin{aligned} |Z_{i,n}| &= \left| \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta_\ell) \mathbb{I}_{|\psi(Y_i, \alpha_\infty; s; \theta_\ell)| \leq M_n} \right. \\ &\quad \left. - E\left[\frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta_\ell) \mathbb{I}_{|\psi(Y_i, \alpha_\infty; s; \theta_\ell)| \leq M_n} \right] \right| \\ &\leq \frac{2}{h_n^d} M_n \sup_{u \in \mathbb{R}^d} |K(u)|. \end{aligned}$$

Ainsi, l'application du Lemme 2.2.5, avec $\lambda \in]0, 1[$ choisi tel que $(1 - \lambda)/2 - \gamma \geq (1 - \beta)/2$ (toujours possible puisque $\gamma < \beta/2$), permet de conclure que pour n assez grand, et pour tout $\epsilon > 0$,

$$\begin{aligned} P\left(\max_{\ell=1, \dots, N} |\tilde{L}'_n(s; \theta_\ell) - E\tilde{L}'_n(s; \theta_\ell)| > \epsilon\right) &\leq \sum_{\ell=1}^N P\left(|\tilde{L}'_n(s; \theta_\ell) - E\tilde{L}'_n(s; \theta_\ell)| > \epsilon\right) \\ &\leq 3N \exp\left(-\frac{\epsilon}{28} \frac{n^{(1-\lambda)/2} h_n^d}{M_n \sup_{u \in \mathbb{R}^d} |K(u)|}\right) \\ &\leq 3N \exp\left(-\frac{\epsilon}{28} \frac{n^{(1-\beta)/2} h_n^d}{\sup_{u \in \mathbb{R}^d} |K(u)|}\right). \end{aligned}$$

D'après la condition $\lim_{n \rightarrow +\infty} \frac{n^{(1-\beta)/2} h_n^d}{\log n} = +\infty$, et en tenant en compte que $N = O(n^p)$, nous obtenons alors

$$\sum_{n=1}^{\infty} P\left(\max_{\ell=1, \dots, N} |\tilde{L}'_n(s; \theta_\ell) - E\tilde{L}'_n(s; \theta_\ell)| > \epsilon\right) < +\infty,$$

ce qui entraîne la convergence presque sûre souhaitée.

D) Convergence uniforme presque sûre de $L_n(s; \cdot) - \tilde{L}_n(s; \cdot)$ vers zéro

Les hypothèses C₄) et C₅), le fait que K et ses dérivées partielles soient bornées, et l'application du théorème ergodique, nous permettent d'écrire

$$\begin{aligned} &\sup_{\theta \in \Theta} |L_n(s; \theta) - \tilde{L}_n(s; \theta)| \\ &= \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} \left\{ K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \psi(Y_i, \hat{\alpha}_n; s; \theta) - K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \psi(Y_i, \alpha_\infty; s; \theta) \right\} \right| \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{nh_n^d} \sum_{i=1}^n \left\{ C_1 \sup_{\theta \in \Theta} N_\psi(Y_i, \alpha_\infty; s; \theta) \|\hat{\alpha}_n - \alpha_\infty\| \right. \\
&\quad \left. + C_2 \sup_{\theta \in \Theta} |\psi(Y_i, \alpha_\infty; s; \theta)| N(Y_i, \alpha_\infty) \frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n} \right\}, \text{ avec } C_1, C_2 > 0, \\
&= O\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n^{d+1}}\right).
\end{aligned}$$

Ainsi, d'après les hypothèses sur $\hat{\alpha}_n$, nous avons $\lim_{n \rightarrow +\infty} \sup_{\theta \in \Theta} |L_n(s; \theta) - \tilde{L}_n(s; \theta)| = 0$, p.s..

D'après les points précédents nous obtenons uniformément en Θ , la convergence presque sûre de la fonction objectif du problème d'optimisation à distance finie $L_n(s; \theta)$, vers la fonction objectif $L_\infty(s; \theta)$, du problème asymptotique. Ceci permet d'obtenir le résultat annoncé. ■

Dans le cas particulier où la variable conditionante ne dépend pas de α , c'est-à-dire, si $S(\cdot, \alpha) = S(\cdot)$ pour tout $\alpha \in \mathcal{A}$, observons que la convergence presque sûre vers zéro de $\sup_{\theta \in \Theta} |L_n(s; \theta) - \tilde{L}_n(s; \theta)|$ établie dans la démonstration précédente, et donc le Théorème 2.2.6, peuvent être obtenus sous des conditions plus faibles sur le noyau K et sur la vitesse de convergence de $\hat{\alpha}_n - \alpha_\infty$ vers zéro. En effet, en admettant que K est un noyau borné à support borné, que $E[\sup_{\theta \in \Theta} N_\psi(Y_0, \alpha_\infty; s; \theta)]^{\frac{2}{\beta} + \delta} < \infty$, pour un $\delta > 0$, où β est donné dans les hypothèses sur la suite (h_n) , et que $E[\sup_{\theta \in \Theta} N_\psi(Y_0, \alpha_\infty; s; \theta) \mid S(Y_0) = \cdot] f_S(\cdot) \in \mathcal{B}_s$, nous obtenons d'après les propriétés de convergence presque sûre de l'estimateur à noyau, la majoration

$$\begin{aligned}
\sup_{\theta \in \Theta} |L_n(s; \theta) - \tilde{L}_n(s; \theta)| &\leq \frac{1}{nh_n^d} \sum_{i=1}^n \left| K\left(\frac{s - S(Y_i)}{h_n}\right) \right| \sup_{\theta \in \Theta} N_\psi(Y_i, \alpha_\infty; s; \theta) \|\hat{\alpha}_n - \alpha_\infty\| \\
&= O(\|\hat{\alpha}_n - \alpha_\infty\|).
\end{aligned}$$

Contrairement au cas général où la variable conditionante dépend de α , nous n'avons pas besoin d'introduire des conditions sur la vitesse de convergence de $\hat{\alpha}_n$ vers α_∞ . Par conséquent, si la variable conditionante ne dépend pas de α , et si $h_n = O(n^{-\gamma})$, pour $\gamma > 0$, les conditions imposées sur h_n sont satisfaites si $0 < \gamma < \frac{1}{2d}$. Dans le cas général, et en admettant que la suite de variables aléatoires $\hat{\alpha}_n$ satisfait une loi du logarithme itéré, c'est-à-dire, que la suite de variables $\sqrt{\frac{n}{\log \log n}} (\hat{\alpha}_n - \alpha_\infty)$ est presque sûrement bornée, les conditions imposées sur h_n sont satisfaites dès que $0 < \gamma < \frac{1}{2(d+1)}$.

2.2.3 Normalité asymptotique

La normalité asymptotique d'un M-estimateur à noyau $\hat{\theta}_n(s)$, pour $s \in \mathbb{R}^d$, est obtenue à partir des propriétés asymptotiques des estimateurs à noyau appliquées à la première et seconde dérivées de la fonction objectif $L_n(s; \theta)$ par rapport à θ , en considérant le développement de la condition du premier ordre

$$\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \hat{\alpha}_n; s; \hat{\theta}_n(s)) = 0,$$

dans un voisinage de $\theta_\infty(s)$ et de α_∞ , où on suppose que $\theta_\infty(s)$ est intérieur à Θ .

Pour cela, on suppose valables les hypothèses générales introduites dans le paragraphe précédent, ainsi que les hypothèses suivantes:

Hypothèses générales additionnelles (GB)

Hypothèses additionnelles sur $(\hat{\alpha}_n)$

$\hat{\alpha}_n$ converge en probabilité vers α_∞ et

$$\hat{\alpha}_n - \alpha_\infty = O_p(h_n^{d+2}) \text{ et } \hat{\alpha}_n - \alpha_\infty = o_p\left(\frac{1}{\sqrt{nh_n^d}}\right).$$

Ainsi que dans l'étude de la convergence presque sûre des M-estimateurs à noyau, quelques hypothèses techniques, concernant le point s de \mathbb{R}^d fixé au départ, sont aussi nécessaires.

Hypothèses locales additionnelles

On suppose que $\theta_\infty(s) \in \text{int}(\Theta)$, et soit $V(\theta_\infty(s)) \subset \Theta$ un voisinage ouvert de $\theta_\infty(s)$ qui satisfait les conditions suivantes:

N₀) L'application ψ satisfait les conditions i) et ii) de l'hypothèse **C₀**) et:

iii) l'application partielle $\psi(Y, \alpha; s; \cdot) : \theta \rightarrow \psi(Y, \alpha; s; \theta)$ est deux fois différentiable dans $V(\theta_\infty(s))$, pour tout $Y \in \mathcal{Y}$ et $\alpha \in \mathcal{A}$;

iv) les applications partielles $S(Y, \cdot) : \alpha \rightarrow S(Y, \alpha)$ et $\frac{\partial \psi}{\partial \theta}(Y, \cdot; s; \theta) : \alpha \rightarrow \frac{\partial \psi}{\partial \theta}(Y, \alpha; s; \theta)$ sont différentiables en α_∞ , pour tout $Y \in \mathcal{Y}$ et $\theta \in V(\theta_\infty(s))$.

Désignons maintenant par **N₁**), ..., **N₅**), les hypothèses **C₁**), ..., **C₅**) introduites en §2.2.2 où on suppose $\Theta = V(\theta_\infty(s))$.

Afin de développer la condition du premier ordre dans un voisinage de α_∞ et de $\theta_\infty(s)$, admettons que (voir 1.1 pour la définition des ensembles d'applications \mathcal{B}_s , \mathcal{C}_s et $\mathcal{D}_s(m)$):

N₆) Pour tout $Y \in \mathcal{Y}$, $\alpha \in \mathcal{A}$, $\theta \in V(\theta_\infty(s))$ et $h > 0$, on a

$$\begin{aligned} & \left\| K\left(\frac{s - S(Y, \alpha)}{h}\right) \frac{\partial \psi}{\partial \theta}(Y, \alpha; s; \theta) - K\left(\frac{s - S(Y, \alpha_\infty)}{h}\right) \frac{\partial \psi}{\partial \theta}(Y, \alpha_\infty; s; \theta) \right. \\ & \quad + \sum_{j=1}^d \frac{\partial K}{\partial u_j}\left(\frac{s - S(Y, \alpha_\infty)}{h}\right) \frac{\partial \psi}{\partial \theta}(Y, \alpha_\infty; s; \theta) \frac{\partial S_j}{\partial \alpha'}(Y, \alpha_\infty) \frac{\alpha - \alpha_\infty}{h} \\ & \quad \left. - K\left(\frac{s - S(Y, \alpha_\infty)}{h}\right) \frac{\partial^2 \psi}{\partial \theta \partial \alpha'}(Y, \alpha_\infty; s; \theta)(\alpha - \alpha_\infty) \right\| \leq \frac{\|\alpha - \alpha_\infty\|^2}{h^2} P(Y, \alpha_\infty; s; \theta), \end{aligned} \quad (2.2.7)$$

où les variables aléatoires $\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial^2 \psi}{\partial \theta \partial \alpha'}(Y_0, \alpha_\infty; s; \theta) \right\|$ et $\sup_{\theta \in V(\theta_\infty(s))} P(Y_0, \alpha_\infty; s; \theta)$ sont intégrables, et $E\left[\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial^2 \psi}{\partial \theta \partial \alpha'}(Y_0, \alpha_\infty; s; \theta) \right\| \mid S(Y_0, \alpha_\infty) = \cdot\right] f_S(\cdot) \in \mathcal{B}_s$.

N₇) Pour tout $Y \in \mathcal{Y}$ et $\theta \in V(\theta_\infty(s))$,

$$\begin{aligned} & \left\| \frac{\partial \psi}{\partial \theta}(Y, \alpha_\infty; s; \theta) - \frac{\partial \psi}{\partial \theta}(Y, \alpha_\infty; s; \theta_\infty(s)) - \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y, \alpha_\infty; s; \theta_\infty(s))(\theta - \theta_\infty(s)) \right\| \\ & \leq \|\theta - \theta_\infty(s)\|^2 Q(Y, \alpha_\infty; s; \theta_\infty(s)), \end{aligned}$$

où les variables aléatoires $\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_0, \alpha_\infty; s; \theta) \right\| \left\| \frac{\partial S}{\partial \alpha}(Y_0, \alpha_\infty) \right\|$ et $Q(Y_0, \alpha_\infty; s; \theta_\infty(s))$ sont intégrables, et les applications données par $E[Q(Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = \cdot] f_S(\cdot)$ et $E\left[\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_0, \alpha_\infty; s; \theta) \right\| \left\| \frac{\partial S}{\partial \alpha}(Y_0, \alpha_\infty) \right\| \mid S(Y_0, \alpha_\infty) = \cdot\right] f_S(\cdot)$ appartiennent à \mathcal{B}_s .

La normalité asymptotique des M-estimateurs à noyau étant obtenue à partir des propriétés des estimateurs à noyau étudiées dans le premier chapitre, et en particulier en tenant compte des observations (1) et (4) du §1.4.3, nous considérons finalement les hypothèses suivantes:

N₈) Pour $i \geq 1$, les applications $f_{(S(Y_i, \alpha_\infty), S(Y_0, \alpha_\infty))}$ sont uniformément bornées dans un voisinage de (s, s) , où $f_{(S(Y_i, \alpha_\infty), S(Y_0, \alpha_\infty))}$ désigne la densité de probabilité du vecteur aléatoire $(S(Y_i, \alpha_\infty), S(Y_0, \alpha_\infty))$, qu'on suppose de loi absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d \times \mathbb{R}^d$.

N₉)

- i) $E\left[\left\| \frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta_\infty(s)) \right\|^{2+\delta}\right] < \infty$, pour un $\delta > 0$;
- ii) $E\left[\frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial \psi}{\partial \theta'}(Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = \cdot\right] f_S(\cdot) \in \mathcal{C}_s$;
- iii) $E\left[\left\| \frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta_\infty(s)) \right\|^{2+\delta} \mid S(Y_0, \alpha_\infty) = \cdot\right] f_S(\cdot) \in \mathcal{B}_s$.

N₁₀)

- i) $E \left[\left\| \frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \right\| \right]^{2+\delta} < \infty$, pour un $\delta > 0$;
- ii) $E \left[\frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{C}_s$;
- iii) $E \left[\left\| \frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \right\|^{2+\delta} \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{B}_s$.

N₁₁)

- i) $E \left[\left\| \frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S}{\partial \alpha'} (Y_0, \alpha_\infty) \right\| \right]^{2+\delta} < \infty$, pour un $\delta > 0$;
- ii) $E \left[\frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S}{\partial \alpha'} (Y_0, \alpha_\infty) \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{D}_s(1)$;
- iii) $E \left[\left\| \frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S}{\partial \alpha'} (Y_0, \alpha_\infty) \right\|^{2+\delta} \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{B}_s$.

Théorème 2.2.8 Soit Θ une partie compacte de \mathbb{R}^p . Sous les hypothèses générales **(GA)** et **(GB)**, soit s un point de \mathbb{R}^d tel que $\theta_\infty(s) \in \text{int}(\Theta)$, où $\theta_\infty(s)$ est défini par (2.2.2), et pour lequel on suppose valables les hypothèses locales **N₀**),...,**N₁₁**). Soit pour $\theta \in \Theta$, $L_n(\cdot; \theta)$ donnée par (2.2.3). Si $\hat{\theta}_n(\cdot)$ est un M-estimateur à noyau associé à $L_n(\cdot; \theta)$, et si

$$A(s) = E \left[- \frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = s \right] f_S(s),$$

est inversible, la variable aléatoire

$$\sqrt{nh_n^d} (\hat{\theta}_n(s) - \theta_\infty(s) - A(s)^{-1} b_n(s)),$$

avec

$$b_n(s) = \int_{\mathbb{R}^d} K(u) \frac{\partial L_\infty}{\partial \theta} (s - uh_n; s; \theta_\infty(s)) du,$$

est asymptotiquement normale de moyenne zéro et de matrice de covariance

$$C(s) = A(s)^{-1} B(s) A(s)^{-1} \int_{\mathbb{R}^d} K^2(u) du,$$

où

$$B(s) = E \left[\frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial \psi}{\partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = s \right] f_S(s).$$

Remarques 2.2.9 (1) Nous obtenons l'expression classique $A(s)^{-1} B(s) A(s)^{-1}$ pour la précision asymptotique d'un M-estimateur paramétrique (cf. White [110]), mais où chaque matrice est évaluée conditionnellement à $S(Y_0, \alpha_\infty) = s$ et où ce terme est multiplié par l'effet provenant du noyau.

(2) Comme les matrices $A(s)$ et $B(s)$ sont évaluées conditionnellement à $S(Y_0, \alpha_\infty) = s$, on peut conclure que les estimateurs $\widehat{\theta}_n(s)$ et

$$\widehat{\theta}_n(s) = \mathop{\text{Arg min}}_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \widehat{\alpha}_n)}{h_n}\right) \psi(Y_i, \widehat{\alpha}_n; S(Y_i, \widehat{\alpha}_n); \theta),$$

ont la même précision asymptotique. En étant un cas particulier de l'estimateur précédent, $\widehat{\theta}_n(\cdot)$ sera étudié en détail plus tard.

Démonstration du Théorème 2.2.8: Soit $s \in \mathbb{R}^d$, fixe, et admettons satisfaites les conditions de l'énoncé. Comme par hypothèse $\theta_\infty(s)$ est à l'intérieur de l'espace des paramètres et $\lim_{n \rightarrow +\infty} \widehat{\theta}_n(s) = \theta_\infty(s)$ p.s. d'après le Théorème 2.2.6, on conclut que $\widehat{\theta}_n(s)$ satisfait asymptotiquement et presque sûrement la condition du premier ordre (cf. Gouriéroux et Monfort [36] pg. 438)

$$\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \widehat{\alpha}_n)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \widehat{\alpha}_n; s; \widehat{\theta}_n(s)) = 0, \quad (2.2.10)$$

car l'application partielle $\theta \rightarrow \psi(Y, \alpha; s; \theta)$ est différentiable dans un voisinage ouvert de $\theta_\infty(s)$, $V(\theta_\infty(s)) \subset \Theta$, pour tout $Y \in \mathcal{Y}$ et $\alpha \in \mathcal{A}$.

Nous allons dans un premier temps développer l'égalité précédente dans un voisinage de α_∞ . D'après l'hypothèse N₆), (2.2.10) s'écrit

$$\begin{aligned} & \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \widehat{\theta}_n(s)) \\ & - \sum_{j=1}^d \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} \frac{\partial K}{\partial u_j}\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \widehat{\theta}_n(s)) \frac{\partial S_j}{\partial \alpha'}(Y_i, \alpha_\infty) \frac{\widehat{\alpha}_n - \alpha_\infty}{h_n} \\ & + \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \alpha'}(Y_i, \alpha_\infty; s; \widehat{\theta}_n(s)) (\widehat{\alpha}_n - \alpha_\infty) \\ & + O_p\left(\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} \frac{\|\widehat{\alpha}_n - \alpha_\infty\|^2}{h_n^2} P(Y_i, \alpha_\infty; s; \widehat{\theta}_n(s))\right) = 0, \end{aligned}$$

où les deux derniers termes sont encore d'après l'hypothèse N₆), de l'ordre de $\sqrt{nh_n^d} \|\widehat{\alpha}_n - \alpha_\infty\|$, en tenant compte de $\widehat{\alpha}_n - \alpha_\infty = O_p(h_n^{d+2})$.

D'après l'hypothèse N₇) et le théorème de la moyenne, un deuxième développement dans un voisinage de $\theta_\infty(s)$ nous conduit à

$$\begin{aligned} & \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \\ & + \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \alpha_\infty; s; \theta_\infty(s)) (\widehat{\theta}_n(s) - \theta_\infty(s)) \end{aligned}$$

$$\begin{aligned}
& +O_p\left(\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) Q(Y_i, \alpha_\infty; s; \theta_\infty(s)) \|\hat{\theta}_n(s) - \theta_\infty(s)\|^2\right) \\
& - \sum_{j=1}^d \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} \frac{\partial K}{\partial u_j}\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'}(Y_i, \alpha_\infty) \frac{\hat{\alpha}_n - \alpha_\infty}{h_n} \\
& - \sum_{j=1}^d \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} \frac{\partial K}{\partial u_j}\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \alpha_\infty; s; \tilde{\theta}_n(s)) (\hat{\theta}_n(s) - \theta_\infty(s)) \\
& \quad \times \frac{\partial S_j}{\partial \alpha'}(Y_i, \alpha_\infty) \frac{\hat{\alpha}_n - \alpha_\infty}{h_n} \\
& +O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) = 0,
\end{aligned}$$

où $\tilde{\theta}_n(s)$ appartient au segment de droite $[\hat{\theta}_n(s), \theta_\infty(s)]$.

L'idée de la démonstration est de montrer que les différents termes de la décomposition précédente, autres que les deux premiers, sont négligeables. L'hypothèse N₇) permet immédiatement de conclure que le troisième et le cinquième termes sont, respectivement,

$$\begin{aligned}
O_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s) - \theta_\infty(s)\|^2\right) &= o_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s) - \theta_\infty(s)\|\right), \text{ et} \\
O_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s) - \theta_\infty(s)\| \frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n}\right) &= o_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s) - \theta_\infty(s)\|\right).
\end{aligned}$$

Ainsi, pour n assez grand, on a l'égalité

$$\begin{aligned}
& \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \\
& + \sum_{i=1}^n \frac{1}{nh_n^d} K\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \sqrt{nh_n^d} (\hat{\theta}_n(s) - \theta_\infty(s)) \\
& - \sum_{j=1}^d \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} \frac{\partial K}{\partial u_j}\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'}(Y_i, \alpha_\infty) \frac{\hat{\alpha}_n - \alpha_\infty}{h_n} \\
& +O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) + o_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s) - \theta_\infty(s)\|\right) = 0.
\end{aligned} \tag{2.2.11}$$

Chacun des termes précédents est étudié dans la suite.

A) Etude du premier terme

Le premier terme de (2.2.11) admet la décomposition

$$\begin{aligned}
& \frac{1}{\sqrt{nh_n^d}} \sum_{i=1}^n \left\{ K\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \right. \\
& \quad \left. - E\left[K\left(\frac{s-S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \theta_\infty(s)) \right] \right\}
\end{aligned}$$

$$+\sqrt{nh_n^d} E \left[\frac{1}{h_n^d} K \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial \psi}{\partial \theta} (Y_i, \alpha_\infty; s; \theta_\infty(s)) \right],$$

où le premier terme de cette somme est, d'après le Théorème 1.3.10 et les hypothèses N₈) et N₉), asymptotiquement normal de moyenne zéro et de matrice de covariance

$$E \left[\frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial \psi}{\partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = s \right] f_S(s) \int_{\mathbb{R}^d} K^2(u) du.$$

De plus, d'après N₃), on conclut que

$$\begin{aligned} & E \left[\frac{1}{h_n^d} K \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial \psi}{\partial \theta} (Y_i, \alpha_\infty; s; \theta_\infty(s)) \right] \\ &= \int_{\mathbb{R}^d} K(u) E \left[\frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = s - uh_n \right] f_S(s - uh_n) du \\ &= \int_{\mathbb{R}^d} K(u) \frac{\partial L_\infty}{\partial \theta} (s - uh_n; s; \theta_\infty(s)) du = b_n(s). \end{aligned}$$

B) Etude du deuxième terme

D'après les hypothèses N₈) et N₁₀), la convergence en probabilité de la variable aléatoire

$$\sum_{i=1}^n \frac{1}{nh_n^d} K \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_i, \alpha_\infty; s; \theta_\infty(s)),$$

vers

$$-A(s) = E \left[\frac{\partial^2 \psi}{\partial \theta \partial \theta'} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \mid S(Y_0, \alpha_\infty) = s \right] f_S(s),$$

lorsque n tend vers l'infini, est une conséquence immédiate du Théorème 1.3.9.

C) Etude du troisième terme

Comme K est un noyau différentiable à support borné, on a

$$\int_{\mathbb{R}^d} \frac{\partial K}{\partial u_j} (u) du = 0, \text{ pour } j = 1, \dots, d.$$

Ainsi, au facteur $-\frac{\hat{\alpha}_n - \alpha_\infty}{h_n}$ près, le troisième terme de (2.2.11) admet la décomposition

$$\begin{aligned} & \sum_{j=1}^d \frac{1}{\sqrt{nh_n^d}} \sum_{i=1}^n \left\{ \frac{\partial K}{\partial u_j} \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial \psi}{\partial \theta} (Y_i, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'} (Y_i, \alpha_\infty) \right. \\ & \quad \left. - E \left[\frac{\partial K}{\partial u_j} \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial \psi}{\partial \theta} (Y_i, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'} (Y_i, \alpha_\infty) \right] \right\} \\ & + \sum_{j=1}^d \sqrt{nh_n^d} \left\{ E \left[\frac{1}{h_n^d} \frac{\partial K}{\partial u_j} \left(\frac{s - S(Y_i, \alpha_\infty)}{h_n} \right) \frac{\partial \psi}{\partial \theta} (Y_i, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'} (Y_i, \alpha_\infty) \right] \right. \\ & \quad \left. - E \left[\frac{\partial \psi}{\partial \theta} (Y_0, \alpha_\infty; s; \theta_\infty(s)) \frac{\partial S_j}{\partial \alpha'} (Y_0, \alpha_\infty) \mid S(Y_0, \alpha_\infty) = s \right] f_S(s) \int_{\mathbb{R}^d} \frac{\partial K}{\partial u_j} (u) du \right\}. \end{aligned}$$

En tenant compte du Théorème 1.3.10 et de la démonstration du Théorème 1.3.14, on conclut que le terme en étude est négligeable. En effet, d'après les hypothèses sur $(\hat{\alpha}_n)$ il s'écrit sous la forme

$$O_p\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n}\right) + O_p\left(\sqrt{nh_n^d} h_n \frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n}\right) = o_p(1).$$

Finalement, d'après les points précédents, l'inversibilité de la matrice $A(s)$ permet de conclure la démonstration du Théorème 2.2.8. \blacksquare

Dans le cas particulier où la variable conditionante ne dépend pas de α , admettons satisfaite l'hypothèse $E\left[\sup_{\theta \in V(\theta_\infty(s))} N_{\frac{\partial \psi}{\partial \theta}}(Y_0, \alpha_\infty; s; \theta) | S(Y_0) = \cdot\right] f_S(\cdot) \in \mathcal{B}_s$ et remplaçons l'hypothèse N_6) par l'hypothèse N_4) avec $\psi = \frac{\partial \psi}{\partial \theta}$. La condition du premier ordre (2.2.10) devient

$$\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; s; \hat{\theta}_n(s)) + O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) = 0.$$

Le Théorème 2.2.8 est donc valable en admettant la convergence presque sûre de $\hat{\alpha}_n$ vers α_∞ et que $\hat{\alpha}_n - \alpha_\infty = o_p\left(\frac{1}{\sqrt{nh_n^d}}\right)$. Ainsi, si la variable aléatoire $\sqrt{n}(\hat{\alpha}_n - \alpha_\infty)$ est bornée en probabilité, et $h_n = O(n^{-\gamma})$, pour un $\gamma > 0$, les conditions imposées sur (h_n) sont satisfaites si $0 < \gamma < \frac{1}{2d}$. Dans le cas général où la variable conditionante dépend éventuellement de α , et en admettant valable une loi du logarithme itéré pour la suite $(\hat{\alpha}_n)$, les conditions sur (h_n) sont satisfaites si $0 < \gamma \leq \frac{1}{2(d+2)}$.

2.2.4 Conditions de normalité asymptotique dans le cas convexe

Nous discutons maintenant le cas particulier où la fonction critère est convexe en θ . Nous verrons qu'en utilisant des arguments classiques (cf. Huber [52]), la normalité asymptotique des M-estimateurs à noyau, décrite dans le Théorème 2.2.8, peut être obtenue avec des conditions moins restrictives sur la vitesse de convergence vers zéro du coefficient de mélange et de la suite (h_n) . En particulier, la convergence presque sûre de $\hat{\alpha}_n - \alpha_\infty$ vers zéro peut être maintenant remplacée par une convergence en probabilité.

Considérons les hypothèses générales suivantes:

Hypothèses générales (GC)

Hypothèses sur le processus

$(Y_i, i \in \mathbb{Z})$ est un processus stochastique fortement stationnaire de type $Alpha(\frac{\delta}{4+\delta})$, pour un $\delta > 0$.

Hypothèses sur le noyau K

K est un noyau positif sur \mathbb{R}^d , borné à support borné.

Hypothèses sur (h_n)

$$h_n \rightarrow 0 \text{ et } nh_n^d \rightarrow +\infty, \text{ lorsque } n \rightarrow +\infty.$$

Hypothèses sur $(\hat{\alpha}_n)$

$\hat{\alpha}_n$ satisfait les conditions imposées dans les hypothèses générales (GB).

Considérons encore l'hypothèse

Hypothèse locale additionnelle

N₁₂) Il existe un voisinage $V(\theta_\infty(s))$ de $\theta_\infty(s)$ tel que:

- i) $E \left[\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta) \right\| \right]^{2+\delta} < \infty$;
- ii) $E \left[\frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta) \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{C}_s$, pour tout $\theta \in V(\theta_\infty(s))$;
- iii) $E \left[\sup_{\theta \in V(\theta_\infty(s))} \left\| \frac{\partial \psi}{\partial \theta}(Y_0, \alpha_\infty; s; \theta) \right\|^{2+\delta} \mid S(Y_0, \alpha_\infty) = \cdot \right] f_S(\cdot) \in \mathcal{B}_s$.

Le résultat suivant peut être obtenu pour des processus $(Y_i, i \in \mathbb{Z})$ de type $Alpha(\frac{\delta}{2+\delta})$, pour un $\delta > 0$, si des conditions convenables sont imposées sur certains moments conditionnels croisés. Par simplicité d'exposition nous considérons seulement le cas des processus de type $Alpha(\frac{\delta}{4+\delta})$, pour un $\delta > 0$ (voir le commentaire (1) du §1.4.3).

Théorème 2.2.12 *Soit Θ une partie compacte de \mathbb{R}^p . Sous les hypothèses générales (GC), soit pour $\theta \in \Theta$, $L_n(\cdot; \theta)$ donnée par (2.2.3) et $\hat{\theta}_n(\cdot)$ un M -estimateur à noyau associé à $L_n(\cdot; \theta)$. Pour $s \in \mathbb{R}^d$, et $\theta_\infty(s)$ défini par (2.2.2), admettons qu'il existe un voisinage ouvert $V(\theta_\infty(s)) \subset \Theta$ de $\theta_\infty(s)$ tel que $\hat{\theta}_n(s) \in V(\theta_\infty(s))$, et sur lequel l'application partielle $\psi(Y, \alpha; s; \cdot) : \theta \rightarrow \psi(Y, \alpha; s; \theta)$ est convexe, pour tout $Y \in \mathcal{Y}$ et $\alpha \in \mathcal{A}$. Supposons valables les hypothèses locales $\mathbf{N}_0), \mathbf{N}_3), \mathbf{N}_6), \dots, \mathbf{N}_{12})$ (par rapport à $V(\theta_\infty(s))$ et à la valeur δ donnée dans les hypothèses sur le processus). Alors, la variable aléatoire*

$$\sqrt{nh_n^d} (\hat{\theta}_n(s) - \theta_\infty(s) - A(s)^{-1}b_n(s)),$$

est asymptotiquement normale de moyenne zéro et de matrice de covariance $C(s)$, où $b_n(s)$, $A(s)$ et $C(s)$ sont définis dans le Théorème 2.2.8.

Démonstration: Soit $s \in \mathbb{R}^d$ dans les conditions de l'énoncé. En tenant compte des hypothèses $N_0), N_6), \dots, N_{11})$, la normalité asymptotique de la variable $\sqrt{nh_n^d}(\widehat{\theta}_n(s) - \theta_\infty(s) - A(s)^{-1}b_n(s))$ est obtenue, comme dans le Théorème 2.2.8, à partir du développement de la condition du premier ordre (2.2.10). La validité d'un tel développement est fondée sur la convergence en probabilité de $\widehat{\theta}_n(s)$ vers $\theta_\infty(s)$, que nous établissons dans la suite. La convexité en θ de la fonction objectif $L_n(s; \theta)$, que nous pouvons déduire d'après la convexité de l'application partielle $\psi(Y, \alpha; s; \cdot) : \theta \rightarrow \psi(Y, \alpha; s; \theta)$ et de la non négativité du noyau K , permet d'obtenir la convergence désirée. Pour cela, étant donné $\epsilon > 0$, il suffira de noter que

$$\begin{aligned} & P\left(\|\widehat{\theta}_n(s) - \theta_\infty(s)\| > \epsilon\right) \\ & \leq \sum_{i=1}^p \left[P\left(\widehat{\theta}_{n,i}(s) > \theta_{\infty,i}(s) + \epsilon\right) + P\left(\widehat{\theta}_{n,i}(s) < \theta_{\infty,i}(s) - \epsilon\right) \right] \\ & = \sum_{i=1}^p \left[P\left(\frac{\partial L_n}{\partial \theta_i}(s; \theta_\infty(s) + \epsilon e_i) \leq 0\right) + P\left(\frac{\partial L_n}{\partial \theta_i}(s; \theta_\infty(s) - \epsilon e_i) \geq 0\right) \right], \end{aligned}$$

où $\widehat{\theta}_{n,i}(s)$ et $\theta_{\infty,i}(s)$ désignent les composantes d'ordre i des vecteurs $\widehat{\theta}_n(s)$ et $\theta_\infty(s)$, respectivement, et e_i le $i^{\text{ème}}$ vecteur de la base canonique de \mathbb{R}^p . Ainsi, la convergence en probabilité annoncée s'obtient d'après les hypothèses $N_3), N_6), N_8)$ et $N_{12})$, car pour $i = 1, \dots, p$, $\frac{\partial L_n}{\partial \theta_i}(s; \theta_\infty(s) + \epsilon e_i)$ et $\frac{\partial L_n}{\partial \theta_i}(s; \theta_\infty(s) - \epsilon e_i)$ convergent en probabilité, lorsque n tend vers l'infini, respectivement vers $\frac{\partial L_\infty}{\partial \theta_i}(s; \theta_\infty(s) + \epsilon e_i) > 0$ et $\frac{\partial L_\infty}{\partial \theta_i}(s; \theta_\infty(s) - \epsilon e_i) < 0$. ■

2.3 Diagnostic non paramétrique de modèles structurels

Les résultats obtenus en 2.2 sont maintenant utilisés dans la formulation de diagnostics sur l'hypothèse

$$H_0 = \{\theta_\infty(s) = 0, \forall s \in \mathbb{R}^d\}.$$

Cependant, le caractère local de tels résultats ne nous permet pas de présenter des tests de niveau asymptotique α , avec $\alpha \in]0, 1[$ fixe, de l'hypothèse H_0 . Nous nous confinerons à présenter des régions ponctuelles de confiance au niveau asymptotique α , pour le paramètre fonctionnel $\theta_\infty(\cdot)$ sous l'hypothèse nulle.

Nous expliquons dans la suite comment utiliser le Théorème 2.2.8 pour construire ces régions ponctuelles asymptotiques de confiance, et proposons un estimateur asymptotiquement équiva-

lent à $\widehat{\theta}_n(\cdot)$ sous l'hypothèse nulle, mais plus facile à calculer dans un cas général. On termine cette section en présentant quelques exemples d'application et une expérience sur données simulées.

2.3.1 Régions ponctuelles de confiance asymptotique

La méconnaissance de la fonction $z \rightarrow \frac{\partial L_\infty}{\partial \theta}(z; s; \theta_\infty(s))$, intervenant dans le terme

$$b_n(s) = \int_{\mathbb{R}^d} K(u) \frac{\partial L_\infty}{\partial \theta}(s - uh_n; s; \theta_\infty(s)) du, \quad s \in \mathbb{R}^d,$$

introduit dans le Théorème 2.2.8, ne nous permet pas d'utiliser directement ce théorème pour obtenir des régions ponctuelles de confiance asymptotique pour le paramètre fonctionnel $\theta_\infty(\cdot)$. Un développement de Taylor de la fonction $\frac{\partial L_\infty}{\partial \theta}(\cdot; s; \theta_\infty(s))$ dans un voisinage du point s , nous permet de contourner cette difficulté. Celui-ci est l'objectif de la

Proposition 2.3.1 *Si pour $s \in \mathbb{R}^d$, $\frac{\partial L_\infty}{\partial \theta}(\cdot; s; \theta_\infty(s)) \in \mathcal{D}_s(m)$ pour un $m \geq 1$, et si $K \in \mathcal{K}_c^d(m)$, alors le terme $b_n(s)$ admet le développement*

$$b_n(s) = h_n^m \Delta^m \left(\frac{\partial L_\infty}{\partial \theta}(\cdot; s; \theta_\infty(s)), K \right)(s) + o(h_n^m),$$

où $\Delta^m \left(\frac{\partial L_\infty}{\partial \theta}(\cdot; s; \theta_\infty(s)), K \right)(\cdot)$ est défini par (1.3.16).

Démonstration: Conséquence immédiate de l'égalité $\frac{\partial L_\infty}{\partial \theta}(s; s; \theta_\infty(s)) = 0$, et des hypothèses sur $\frac{\partial L_\infty}{\partial \theta}(\cdot; s; \theta_\infty(s))$ et sur le noyau K (voir la démonstration du Théorème 1.3.14). ■

Sous les conditions du Théorème 2.2.8 et d'après la proposition précédente, le terme de biais $A(s)^{-1}b_n(s)$ peut être asymptotiquement négligé, si on admet que la suite (h_n) satisfait la condition supplémentaire $nh_n^{d+2m} \rightarrow 0$, $n \rightarrow +\infty$. Dans ces conditions la variable aléatoire

$$\sqrt{nh_n^d} (\widehat{\theta}_n(s) - \theta_\infty(s)),$$

est asymptotiquement normale de moyenne zéro et de matrice de covariance $C(s)$. Si $\widehat{C}_n(s)$ est un estimateur convergent de $C(s)$, nous pouvons construire des régions ponctuelles asymptotiques de confiance pour le paramètre fonctionnel $\theta_\infty(\cdot)$, sous l'hypothèse H_0 , en approchant la loi de $\widehat{\theta}_n(s)$ par une loi normale de moyenne zéro et de matrice de covariance $\frac{\widehat{C}_n(s)}{nh_n^d}$:

$$\widehat{\theta}_n(s) \sim N \left(0, \frac{\widehat{C}_n(s)}{nh_n^d} \right).$$

Un premier désavantage de cette démarche concerne les conditions assez restrictives qu'on a besoin d'imposer sur la suite (h_n) . Si $h_n = O(n^{-\gamma})$, pour $\gamma > 0$, telles conditions sont satisfaites si $\frac{1}{d+2m} < \gamma \leq \frac{1}{2(d+2)}$. Nous devons alors choisir $m > \frac{d}{2} + 2$, ce qui exclut la possibilité de prendre pour K un noyau positif. Cette condition de positivité est cependant essentielle pour obtenir la normalité asymptotique d'un M-estimateur à noyau dans le cas convexe, sous les conditions décrites en §2.2.4. Deuxièmement, la variable aléatoire $\hat{\theta}_n(s)$ est, à distance finie, affecté d'un biais égal à $A(s)^{-1}b_n(s)$, qui n'est pas considéré dans la loi asymptotique précédente.

Nous allons voir que ce problème peut être résolu si on utilise dans la construction des régions ponctuelles asymptotiques de confiance pour le paramètre fonctionnel $\theta_\infty(\cdot)$, sous l'hypothèse H_0 , l'estimateur

$$\widehat{\hat{\theta}}_n(s) = \text{Arg min}_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \psi(Y_i, \hat{\alpha}_n; S(Y_i, \hat{\alpha}_n); \theta),$$

au lieu de $\hat{\theta}_n(s)$. Rappelons que (cf. Remarque 2.2.9 (2)) les estimateurs $\hat{\theta}_n(s)$ et $\widehat{\hat{\theta}}_n(s)$ ont la même précision asymptotique.

Proposition 2.3.2 *Admettons que la fonction critère $\psi(Y, \alpha; s; \theta)$ est indépendante de s et satisfait l'hypothèse N_3 . Si $\theta_\infty(s) = 0$, pour tout $s \in \mathbb{R}^d$, alors $b_n(s) = 0$, pour tout $n \in \mathbb{N}^*$ et $s \in \mathbb{R}^d$.*

Démonstration: Notons $\psi(Y, \alpha; s; \theta) = \tilde{\psi}(Y, \alpha; \theta)$. D'après la définition de $\theta_\infty(s)$, et l'hypothèse N_3 , on a

$$0 = \frac{\partial L_\infty}{\partial \theta}(s; s; \theta_\infty(s)) = E\left[\frac{\partial \tilde{\psi}}{\partial \theta}(Y_0, \alpha_\infty; 0) \mid S(Y_0, \alpha_\infty) = s\right] f_S(s), \text{ pour tout } s \in \mathbb{R}^d.$$

Ainsi, l'application $z \rightarrow E\left[\frac{\partial \tilde{\psi}}{\partial \theta}(Y_0, \alpha_\infty; 0) \mid S(Y_0, \alpha_\infty) = z\right] f_S(z)$ est identiquement nulle ce qui permet de conclure que, pour $n \in \mathbb{N}^*$ et $s \in \mathbb{R}^d$,

$$b_n(s) = \int_{\mathbb{R}^d} K(u) E\left[\frac{\partial \tilde{\psi}}{\partial \theta}(Y_0, \alpha_\infty; 0) \mid S(Y_0, \alpha_\infty) = s - uh_n\right] f_S(s - uh_n) du = 0. \quad \blacksquare$$

En résumé, dans la construction de régions ponctuelles asymptotiques de confiance pour le paramètre $\theta_\infty(\cdot)$ sous H_0 , il est donc préférable de considérer l'estimateur $\widehat{\hat{\theta}}_n(s)$, solution d'un problème d'optimisation dont la fonction critère est égale à $\psi(Y, \alpha; S(Y, \alpha); \theta)$, à l'estimateur $\hat{\theta}_n(s)$ solution d'un problème d'optimisation de fonction critère $\psi(Y, \alpha; s; \theta)$. En

effet, le Théorème 2.2.8 et les propositions précédentes nous conduisent à approcher, sous l'hypothèse H_0 , la loi des estimateurs $\widehat{\theta}_n(s)$ et $\widehat{\widehat{\theta}}_n(s)$ par une loi normale de moyenne zéro et de matrice de covariance $\frac{C(s)}{nh_n^d}$. Cependant, la loi de $\widehat{\widehat{\theta}}_n(s)$, est à distance finie, plus “proche” de cette loi asymptotique que celle de $\widehat{\theta}_n(s)$.

2.3.2 M-estimateur local à noyau

Dans la suite, et en tenant compte des résultats du paragraphe précédent, nous admettons que la fonction critère est indépendante de s . Nous la désignerons par $\psi(Y, \alpha; \theta)$. Le calcul de l'estimateur $\widehat{\theta}_n(\cdot)$ peut être difficile dans la pratique, puisque il demande un grand nombre d'optimisations (autant que la grille de valeurs de s considérée) dès que la fonction critère est de forme complexe. Il sera donc utile de proposer un estimateur facile à calculer et asymptotiquement équivalent à $\widehat{\theta}_n(\cdot)$ sous l'hypothèse nulle.

Un tel estimateur peut être facilement déduit à partir des développements asymptotiques obtenus pour $\widehat{\theta}_n(\cdot)$ dans la démonstration du Théorème 2.2.8. Des formes alternatives du résultat suivant peuvent être obtenues dans les cas particuliers où la variable conditionante ne dépend pas de α ou où la fonction critère est convexe en θ .

Proposition 2.3.3 *Sous les conditions du Théorème 2.2.8, soit s un point de \mathbb{R}^d tel que $\theta_\infty(s) = 0$. Si on suppose que les hypothèses locales \mathbf{N}_4) et \mathbf{N}_5) sont aussi valables pour $\psi = \frac{\partial^2 \psi}{\partial \theta \partial \theta'}$, l'estimateur*

$$\begin{aligned} \widetilde{\theta}_n(s) &= \left\{ - \sum_{i=1}^n \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \widehat{\alpha}_n; 0) K\left(\frac{s - S(Y_i, \widehat{\alpha}_n)}{h_n}\right) \right\}^{-1} \\ &\quad \times \sum_{i=1}^n \frac{\partial \psi}{\partial \theta}(Y_i, \widehat{\alpha}_n; 0) K\left(\frac{s - S(Y_i, \widehat{\alpha}_n)}{h_n}\right), \end{aligned}$$

est asymptotiquement équivalent à $\widehat{\theta}_n(s)$:

$$\widetilde{\theta}_n(s) - \widehat{\theta}_n(s) = o_p\left(\frac{1}{\sqrt{nh_n^d}}\right).$$

Démonstration: Si s est un point de \mathbb{R}^d tel que $\theta_\infty(s) = 0$, d'après la démonstration du Théorème 2.2.8 on conclut que

$$\sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; 0)$$

$$\begin{aligned}
& + \sum_{i=1}^n \frac{1}{nh_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \alpha_\infty; 0) \sqrt{nh_n^d} \hat{\theta}_n(s) \\
& + o_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s)\|\right) + O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) + O_p\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n}\right) = 0.
\end{aligned}$$

Les hypothèses locales N₄) et N₅) étant valables pour $\psi = \frac{\partial^2 \psi}{\partial \theta \partial \theta'}$, on peut écrire

$$\begin{aligned}
& \sum_{i=1}^n \frac{1}{nh_n^d} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \alpha_\infty; 0) \\
& = \sum_{i=1}^n \frac{1}{nh_n^d} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \frac{\partial^2 \psi}{\partial \theta \partial \theta'}(Y_i, \hat{\alpha}_n; 0) + O_p\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n^{d+1}}\right).
\end{aligned}$$

De plus, d'après l'hypothèse N₆) et en tenant compte de la discussion faite dans la démonstration du Théorème 2.2.8, on a

$$\begin{aligned}
& \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \alpha_\infty)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \alpha_\infty; 0) \\
& = \sum_{i=1}^n \frac{1}{\sqrt{nh_n^d}} K\left(\frac{s - S(Y_i, \hat{\alpha}_n)}{h_n}\right) \frac{\partial \psi}{\partial \theta}(Y_i, \hat{\alpha}_n; 0) \\
& \quad + O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) + O_p\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n}\right).
\end{aligned}$$

L'égalité

$$\sqrt{nh_n^d} (\hat{\theta}_n(s) - \tilde{\theta}_n(s)) = o_p\left(\sqrt{nh_n^d} \|\hat{\theta}_n(s)\|\right) + O_p\left(\sqrt{nh_n^d} \|\hat{\alpha}_n - \alpha_\infty\|\right) + O_p\left(\frac{\|\hat{\alpha}_n - \alpha_\infty\|}{h_n^{d+1}}\right),$$

qu'on dérive d'après les développements précédents, la définition de $\tilde{\theta}_n(s)$ et l'inversibilité de la matrice $A(s)$, permet d'obtenir l'équivalence souhaitée, compte tenu de la vitesse de convergence vers α_∞ imposée à la suite $(\hat{\alpha}_n)$. ■

2.3.3 Quelques exemples d'application

Nous décrivons maintenant quelques problèmes où les techniques développées peuvent être appliquées.

Exemple 2.3.4 (Adéquation de la régression à une famille paramétrique de courbes)

Soit (X, Y) un vecteur aléatoire à valeurs dans $\mathbb{R}^d \times \mathbb{R}$. Nous voulons formuler un diagnostic sur l'hypothèse d'adéquation de la fonction de régression $E[Y | X = \cdot]$ à une famille paramétrique de courbes donnée au départ:

$$H_0 : E[Y | X = \cdot] \in \{\mu(\cdot; m), m \in \mathcal{M}\}, \text{ où } \mathcal{M} \subset \mathbb{R}^k.$$

En étant donné un estimateur paramétrique \hat{m}_n de m , nous pouvons pour $x \in \mathbb{R}^d$, considérer l'estimateur

$$\hat{\theta}_n(x) = \text{Arg min}_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{x - X_i}{h_n}\right) (Y_i - \mu(X_i, \hat{m}_n) - \theta)^2.$$

Pour \hat{m}_n nous pouvons en particulier prendre l'estimateur des moindres carrés non linéaires pondérés (cf. [36] pg. 237-238), c'est-à-dire,

$$\hat{m}_n = \text{Arg min}_{m \in \mathcal{M}} \sum_{i=1}^n (Y_i - \mu(X_i; m))^2 \omega(X_i),$$

où ω est une fonction strictement positive donnée. Dans un modèle linéaire à k paramètres,

$$\mu(\cdot; m) = m_1 h_1(\cdot) + \dots + m_k h_k(\cdot),$$

où $h = (h_1, \dots, h_k)'$ est une fonction à valeurs dans \mathbb{R}^k , \hat{m}_n est l'estimateur des moindres carrés linéaires pondérés,

$$\hat{m}_n = \left\{ \sum_{i=1}^n \omega(X_i) h(X_i) h'(X_i) \right\}^{-1} \sum_{i=1}^n \omega(X_i) h(X_i) Y_i,$$

à condition que la première matrice soit inversible.

L'hypothèse nulle s'écrit alors de façon équivalente

$$H_0 : \{\theta_{\infty}(\cdot) = 0\},$$

où $\theta_{\infty}(\cdot) = E[Y|X = \cdot] - \mu(\cdot; m_{\infty})$, et $m_{\infty} = \lim_{n \rightarrow +\infty} \hat{m}_n$, presque sûrement.

Dans ce cas, les estimateurs $\hat{\theta}_n(\cdot)$ et $\tilde{\theta}_n(\cdot)$ définis dans la Proposition 2.3.3 sont égaux, et convergent vers zéro sous l'hypothèse nulle. Pour $x \in \mathbb{R}^d$, ils sont définis par

$$\hat{\theta}_n(x) = \tilde{\theta}_n(x) = \left\{ \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \right\}^{-1} \left\{ \sum_{i=1}^n (Y_i - \mu(X_i; \hat{m}_n)) K\left(\frac{x - X_i}{h_n}\right) \right\}.$$

Les diagnostics fondés sur une procédure d'estimation en deux étapes comme celle que nous venons de décrire, avec une première étape d'estimation paramétrique, et une deuxième non paramétrique, peuvent être comparés avec la solution usuellement proposée pour le problème précédent. Elle concerne l'estimation non paramétrique de la régression en utilisant l'estimateur à noyau $m_n(x)$ défini par

$$m_n(x) = \left\{ \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \right\}^{-1} \left\{ \sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h_n}\right) \right\},$$

et sa comparaison à $\mu(x; \widehat{m}_n)$. Cette procédure équivaut au calcul de

$$\bar{\theta}_n(x) = \text{Arg min}_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{x - X_i}{h_n}\right) (Y_i - \mu(x; \widehat{m}_n) - \theta)^2,$$

(puisque $\bar{\theta}_n(x) = m_n(x) - \mu(x; \widehat{m}_n)$), et dans la comparaison à zéro de cette fonction. D'après §2.3.1, la démarche fondée sur $\widehat{\theta}_n(\cdot)$ est préférable à celle fondée sur $\bar{\theta}_n(\cdot)$.

Exemple 2.3.5 (Sur une conséquence de l'adéquation de la densité conditionnelle à une famille paramétrique de densités conditionnelles) Soit (X, Y) un vecteur aléatoire

à valeurs dans $\mathbb{R}^d \times \mathbb{R}$ et désignons par $f_0(\cdot | x)$, $x \in \mathbb{R}^d$, la densité de la loi conditionnelle de $Y | X = x$, qu'on suppose inconnue. Considérons une famille paramétrique de densités conditionnelles $\{f(\cdot | \cdot; c), c \in \mathcal{C}\}$, où $\mathcal{C} \subset \mathbb{R}^k$. Soit pour $x \in \mathbb{R}^d$, $c_{\infty}(x)$ la pseudo-valeur vraie de c au sens de la distance de Kullback,

$$c_{\infty}(x) = \text{Arg min}_c \int_{\mathbb{R}^d} \log \frac{f_0(y | x)}{f(y | x; c)} f_0(y | x) dy.$$

Nous voulons inférer sur l'hypothèse $H_0 = \{c_{\infty}(x) \text{ est indépendant de } x\}$.

Si \widehat{c}_n est l'estimateur de la pseudo-maximum de vraisemblance de c ,

$$\widehat{c}_n = \text{Arg max}_{c \in \mathcal{C}} \sum_{i=1}^n \log f(Y_i | X_i; c),$$

et comme l'hypothèse H_0 est $\{\theta_{\infty}(x) = 0, \forall x \in \mathbb{R}^d\}$, elle peut être analysée en comparant à zéro l'estimateur

$$\widehat{\theta}_n(x) = \text{Arg max}_{\theta} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{x - X_i}{h_n}\right) \log f(Y_i | X_i; \widehat{c}_n + \theta).$$

Remarquons qu'en particulier le rejet de H_0 entraîne le rejet de l'hypothèse H'_0 d'adéquation de la densité conditionnelle $f_0(\cdot | \cdot)$ à la famille $\{f(\cdot | \cdot; c), c \in \mathcal{C}\}$: $H'_0 = \{\exists c \in \mathcal{C} : f_0(\cdot | \cdot) = f(\cdot | \cdot; c)\}$. Admettons que, pour $x \in \mathbb{R}$, $Y | X = x \sim N(xa_0, \omega^2(x))$ où $a_0 \in \mathbb{R}$ est fixé. Si on prend la famille de densités conditionnelles $f(y|x; a) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-xa)^2}{2}\right)$, paramétrée par $a \in \mathbb{R}$, l'hypothèse H_0 s'écrit $H'_0 = \{\exists a : Y | X = x \sim N(xa, 1)\}$. Lorsque H_0 est satisfaite la démarche précédente ne permet pas de détecter l'erreur de spécification. Si on considère maintenant la famille paramétrique $f(y|x; a, b) = \frac{1}{\sqrt{2\pi b^2}} \exp\left(-\frac{(y-xa)^2}{2b^2}\right)$ pour $a \in \mathbb{R}$ et $b \in \mathbb{R} \setminus \{0\}$, nous avons $H'_0 = \{\exists a, b : Y | X = x \sim N(xa, b^2)\}$. Comme dans ce cas $c_{\infty}(\cdot)$ n'est pas constant, nous pouvons détecter l'erreur de spécification.

Exemple 2.3.6 (Détection de l'oubli de variable dans une régression) Considérons le modèle de régression linéaire: $Y_i = aX_i + b + u_i, i \in \mathbb{Z}$, et soient \hat{a}_n et \hat{b}_n les estimateurs des moindres carrés de a et b , respectivement. On se demande s'il est nécessaire d'introduire une nouvelle variable Z dans ce modèle de régression. Cette variable peut être introduite avec un effet additif (c'est-à-dire dans le paramètre b), ou avec un effet multiplicatif (c'est-à-dire, dans le paramètre a). Ces deux possibilités sont analysées en simultané si on considère, pour $z \in \mathbb{R}$, le problème

$$\min_{\theta_1, \theta_2} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{z - Z_i}{h_n}\right) (Y_i - \hat{a}_n X_i - \hat{b}_n - \theta_1 X_i - \theta_2)^2,$$

dont la solution est donnée par

$$\begin{bmatrix} \hat{\theta}_{1,n}(z) \\ \hat{\theta}_{2,n}(z) \end{bmatrix} = \left\{ \sum_{i=1}^n K\left(\frac{z - Z_i}{h_n}\right) \begin{bmatrix} X_i^2 & X_i \\ X_i & 1 \end{bmatrix} \right\}^{-1} \sum_{i=1}^n K\left(\frac{z - Z_i}{h_n}\right) \begin{bmatrix} X_i \hat{u}_i \\ \hat{u}_i \end{bmatrix}, \quad (2.3.7)$$

où $\hat{u}_i = Y_i - \hat{a}_n X_i - \hat{b}_n$.

L'hypothèse d'absence d'effet multiplicatif est $\{\theta_{1,\infty}(z) = 0, \forall z\}$, et d'absence d'effet additif est $\{\theta_{2,\infty}(z) = 0, \forall z\}$.

Exemple 2.3.8 (Spécifications ARCH(1) pour les erreurs dans un modèle de régression) Considérons le modèle de régression $Y_i = bX_i + \varepsilon_i, i \in \mathbb{Z}$. Nous voulons "tester" des spécifications de type ARCH(1) pour le processus $(\varepsilon_i, i \in \mathbb{Z})$ (cf. Gouriéroux [35] pg. 61), c'est-à-dire,

$$\begin{aligned} E[\varepsilon_i | \varepsilon_{i-1}] &= 0, \\ V[\varepsilon_i | \varepsilon_{i-1}] &= c + a\varepsilon_{i-1}^2, \text{ pour } i \in \mathbb{Z}. \end{aligned}$$

Les divers paramètres b, c et a , peuvent être estimés par la méthode du pseudo-maximum de vraisemblance (cf. [35] Chap. IV). Notons ces estimateurs par \hat{b}_n, \hat{c}_n et \hat{a}_n , et par $\hat{\varepsilon}_i = Y_i - \hat{b}_n X_i$. Les spécifications précédentes peuvent être analysées en simultané d'après la résolution, pour $\varepsilon \in \mathbb{R}$, du problème

$$\min_{m, \sigma^2} \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n} K\left(\frac{\varepsilon - \hat{\varepsilon}_i}{h_n}\right) \left\{ \log(\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2 + \sigma^2) + \frac{(\hat{\varepsilon}_{i+1} - m)^2}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2 + \sigma^2} \right\},$$

et d'après la comparaison à zéro de sa solution $(\hat{m}_n(\varepsilon), \hat{\sigma}_n^2(\varepsilon))$.

L'hypothèse nulle d'une spécification correcte de la moyenne conditionnelle est $\{m_\infty(\varepsilon) = 0, \forall \varepsilon\}$ et celle d'une spécification correcte de la variance conditionnelle est

$\{\sigma_\infty^2(\varepsilon) = 0, \forall \varepsilon\}$. Notons que, dans ce cas, la variable conditionante dépend de la première étape d'estimation et que nous n'obtenons pas une forme explicite pour les estimateurs $\hat{m}_n(\varepsilon)$ et $\hat{\sigma}_n^2(\varepsilon)$. Il sera donc préférable d'utiliser les estimateurs modifiés $\tilde{m}_n(\varepsilon)$ et $\tilde{\sigma}_n^2(\varepsilon)$ (cf. Proposition 2.3.3). Ils sont définis par

$$\begin{aligned} (\tilde{m}_n(\varepsilon), \tilde{\sigma}_n^2(\varepsilon))' &= \left\{ \sum_{i=1}^n K\left(\frac{\varepsilon - \hat{\varepsilon}_i}{h_n}\right) \left[\begin{array}{cc} \frac{1}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2} & \frac{\hat{\varepsilon}_{i+1}^2}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2} \\ \frac{\hat{\varepsilon}_{i+1}^2}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2} & \frac{\hat{\varepsilon}_{i+1}^2}{(\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2)^3} - \frac{1}{2} \frac{1}{(\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2)^2} \end{array} \right] \right\}^{-1} \\ &\times \sum_{i=1}^n K\left(\frac{\varepsilon - \hat{\varepsilon}_i}{h_n}\right) \left(\frac{\hat{\varepsilon}_{i+1}^2}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2}, -\frac{1}{2} \frac{1}{\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2} + \frac{1}{2} \frac{\hat{\varepsilon}_{i+1}^2}{(\hat{c}_n + \hat{a}_n \hat{\varepsilon}_i^2)^2} \right)'. \end{aligned}$$

2.3.4 Une illustration

La démarche développée dans les sections précédentes est maintenant illustrée par une expérience sur données simulées, en reprenant l'Exemple 2.3.6 de détection de l'oubli de variable dans un modèle de régression.

Données simulées

Nous considérons des données simulées engendrées de la façon suivante

$$Y_i = a_0(Z_i)X_i + b_0 + u_i, \quad i = 1, \dots, n,$$

où (X_i) , (Z_i) et (u_i) sont des bruits blancs gaussiens standard indépendants, $a_0(\cdot)$ est une application réelle et b_0 un nombre réel. Dans notre étude nous considérons les deux cas suivants:

- I) $a_0(Z_i) = Z_i, b_0 = 1$;
- II) $a_0(Z_i) = Z_i^2, b_0 = 1$.

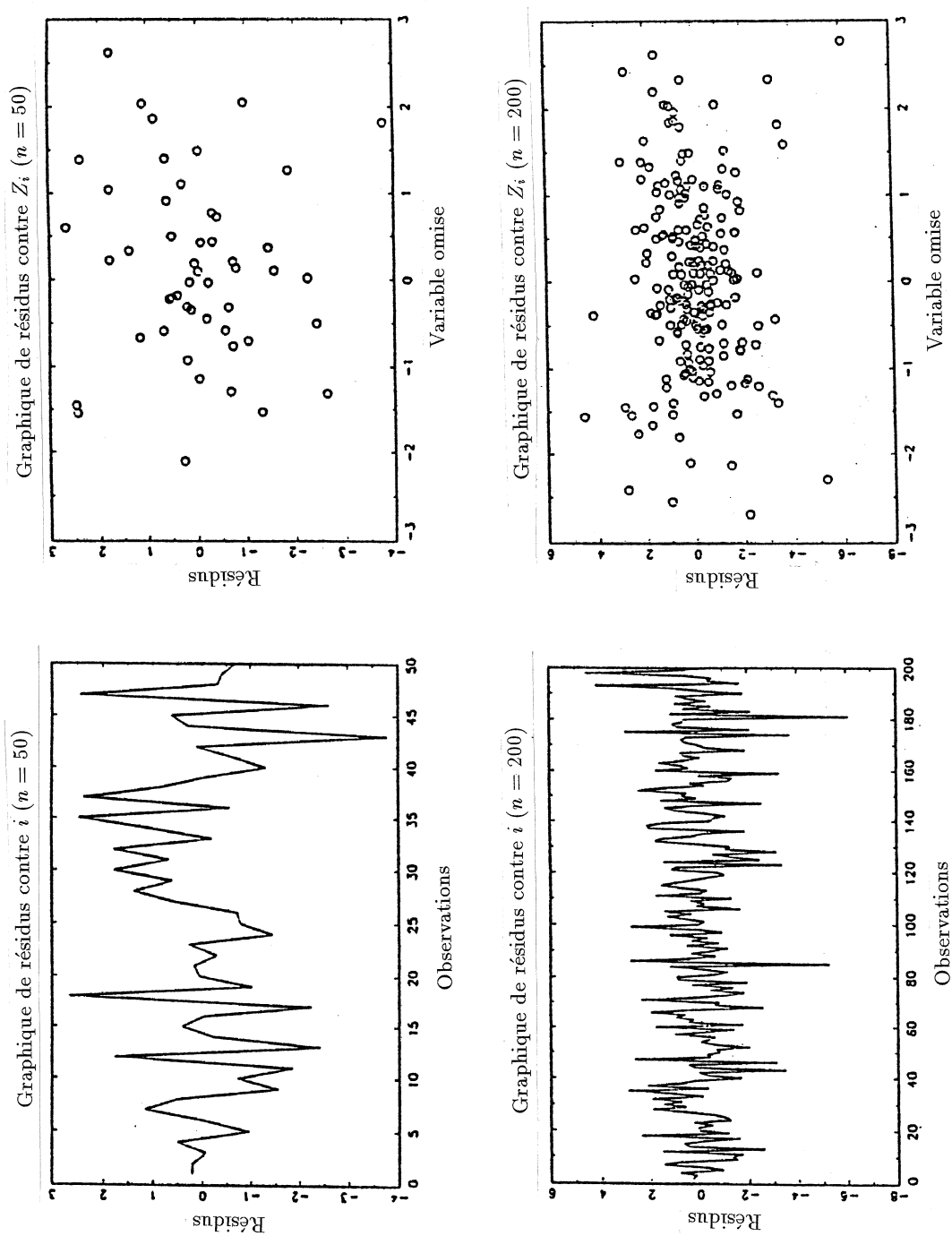
Régression linéaire

Supposons que nous estimons un modèle de la forme

$$Y_i = aX_i + b + v_i, \quad i \in \mathbb{Z}.$$

Les estimateurs des moindres carrés ordinaires \hat{a}_n et \hat{b}_n des paramètres a et b , convergent respectivement vers les pseudo-valeurs vrais $Ea_0(Z_i)$ et b_0 . Par rapport aux données considérées, nous avons

- I) $Ea_0(Z_i) = EZ_i = 0, b_0 = 1$;
- II) $Ea_0(Z_i) = EZ_i^2 = 1, b_0 = 1$.

Figure 2.3.9: Graphiques de résidus contre i et contre Z_i pour I), avec $n = 50$ et 200.

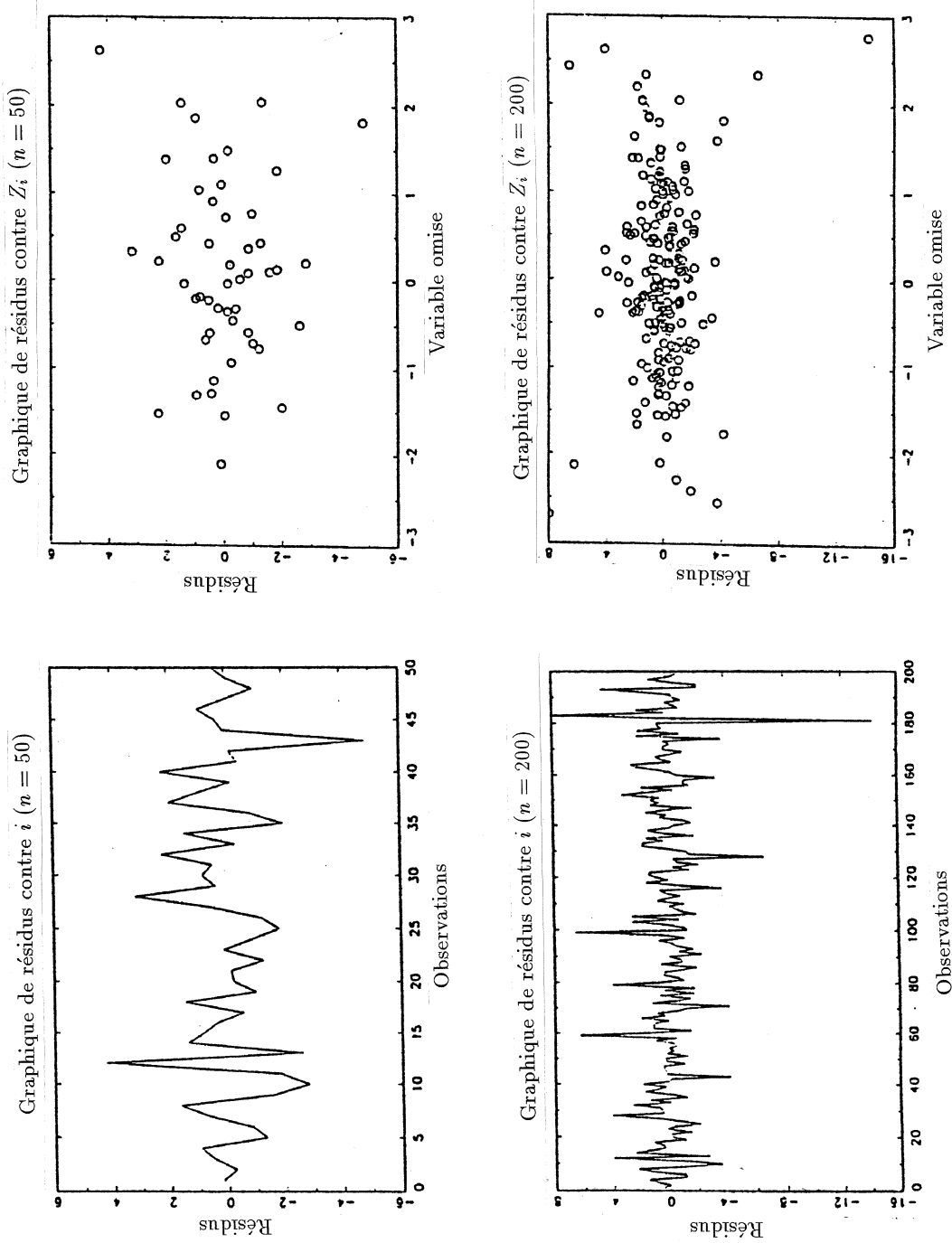


Figure 2.3.10: Graphiques de résidus contre i et contre Z_i pour II), avec $n = 50$ et 200 .

Si on considère le résidu $\widehat{v}_i = Y_i - \widehat{a}_n X_i - \widehat{b}_n$, on voit que pour chaque i , \widehat{v}_i converge en probabilité, lorsque n tend vers l'infini, vers $v_i^\infty = [a_0(Z_i) - E a_0(Z_i)] X_i + u_i$. Ceci entraîne que $E v_i^\infty Z_i = E [(a_0(Z_i) - E a_0(Z_i)) Z_i] E X_i = 0$, puisque $E X_i = 0$, et donc, un test usuel ne va pas détecter la présence de la variable Z . De plus, comme $v_i^\infty, i \in \mathbb{Z}$, est un bruit blanc et $E[v_i^\infty | Z_i]$ est égal à zéro, un graphique de résidus de \widehat{v}_i contre i , ou contre Z_i , ne permet pas, probablement, de détecter aucune anomalie.

Les graphiques de résidus contre i et contre Z_i , pour I) et II), et pour $n = 50$ et $n = 200$, que nous présentons dans les Figures 2.3.9 et 2.3.10 confirment les prévisions précédentes.

M-estimateurs à noyau

Nous considérons maintenant les M-estimateurs à noyau $\widehat{\theta}_{1,n}(\cdot)$ et $\widehat{\theta}_{2,n}(\cdot)$ définis par (2.3.7), des paramètres $\theta_{1,\infty}(\cdot)$ et $\theta_{2,\infty}(\cdot)$ qui mesurent respectivement, l'existence d'un effet multiplicatif et d'un effet additif de la variable Z dans le modèle de régression linéaire considéré. Sous le modèle simulé, et pour $z \in \mathbb{R}$, $(\widehat{\theta}_{1,n}(z), \widehat{\theta}_{2,n}(z))$ converge en probabilité, lorsque n tend vers l'infini, vers $(\theta_{1,\infty}(z), \theta_{2,\infty}(z)) = (a_0(z) - E a_0(Z_i), 0)$.

Les estimateurs $\widehat{\theta}_{1,n}(\cdot)$ et $\widehat{\theta}_{2,n}(\cdot)$ dépendent du noyau K , de la fenêtre h_n , et du nombre d'observations n . Nous avons considéré trois noyaux: le noyau d'Epanechnikov ($K(u) = \frac{3}{4}(1 - u^2)\mathbb{I}_{|u| \leq 1}$), le noyau quadratique ($K(u) = \frac{15}{16}(1 - u^2)^2\mathbb{I}_{|u| \leq 1}$), et le noyau gaussien ($K(u) = (2\pi)^{-1/2} \exp(-\frac{u^2}{2})$). Les résultats obtenus avec ces noyaux ont été très semblables et dans la suite nous considérons uniquement le noyau d'Epanechnikov.

Nous nous intéressons par l'hypothèse nulle $H_0 = \{(\theta_{1,\infty}(z), \theta_{2,\infty}(z)) = (0, 0), \forall z \in \mathbb{R}\}$. Sous H_0 , la matrice de covariance asymptotique du vecteur aléatoire $\sqrt{nh_n}(\widehat{\theta}_{1,n}(z), \widehat{\theta}_{2,n}(z))$ est donnée par (cf. Théorème 2.2.8 et Proposition 2.3.2)

$$C(z) = A(z)^{-1} B(z) A(z)^{-1} \int_{\mathbb{R}} K^2(u) du,$$

où $A(z)$ et $B(z)$ peuvent être estimés de façon convergente respectivement par

$$-2 \frac{X'X}{n} f_Z(z) \text{ et } 4 \widehat{\sigma}_n^2 \frac{X'X}{n} f_Z(z),$$

où X est la matrice dont les lignes sont égales à $(X_i, 1)$ pour $i = 1, \dots, n$, $\widehat{\sigma}_n^2$ est l'estimateur des moindres carrés de la variance de v_i , et $f_Z(\cdot)$ est l'estimateur de la densité de Z où on utilise le

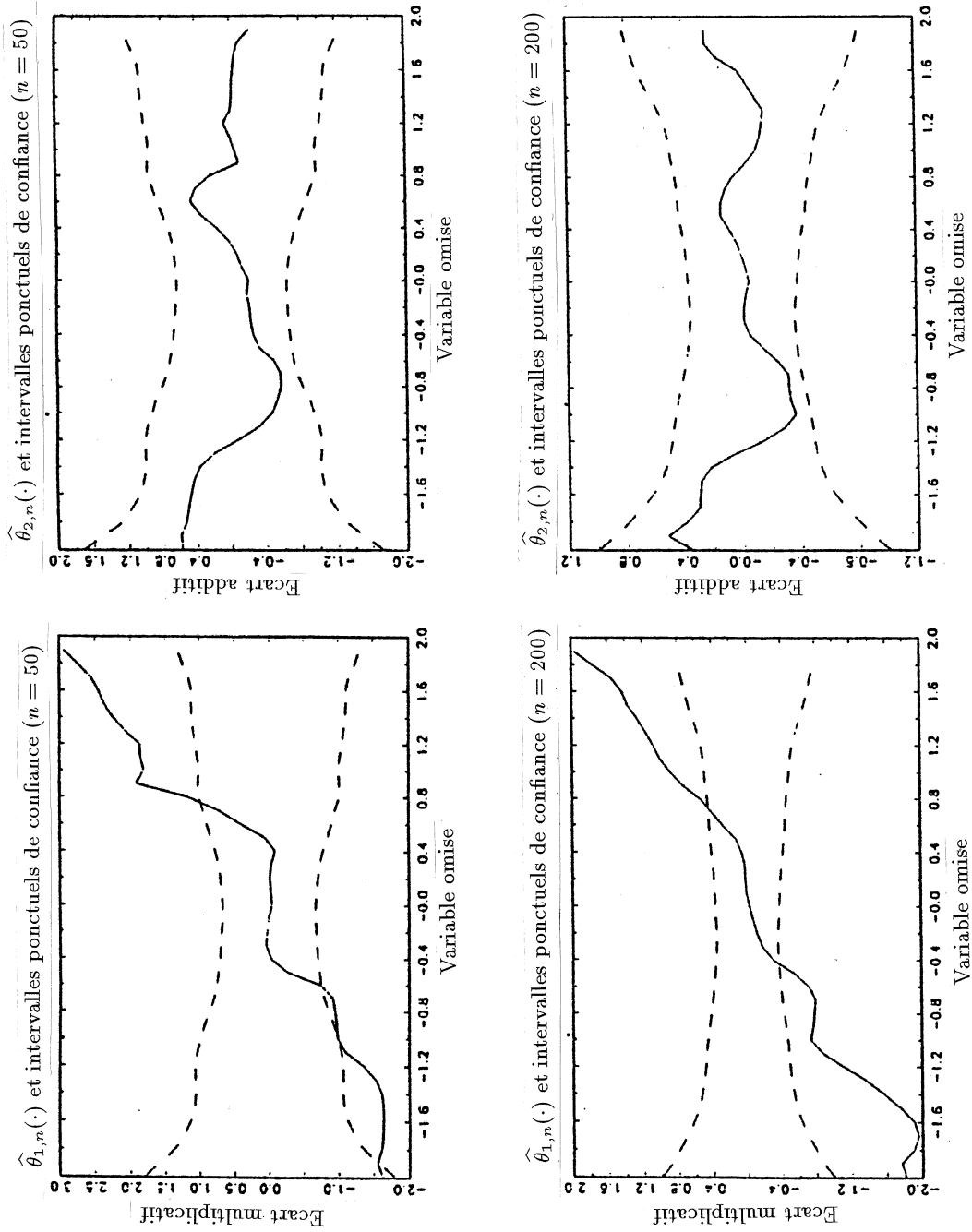


Figure 2.3.11: Estimateurs $\hat{\theta}_{1,n}(\cdot)$, $\hat{\theta}_{2,n}(\cdot)$, et intervalles ponctuels de confiance sous H_0 au niveau asymptotique 0.05 pour I), avec $h_n = 0.5$ et $n = 50$ et 200.

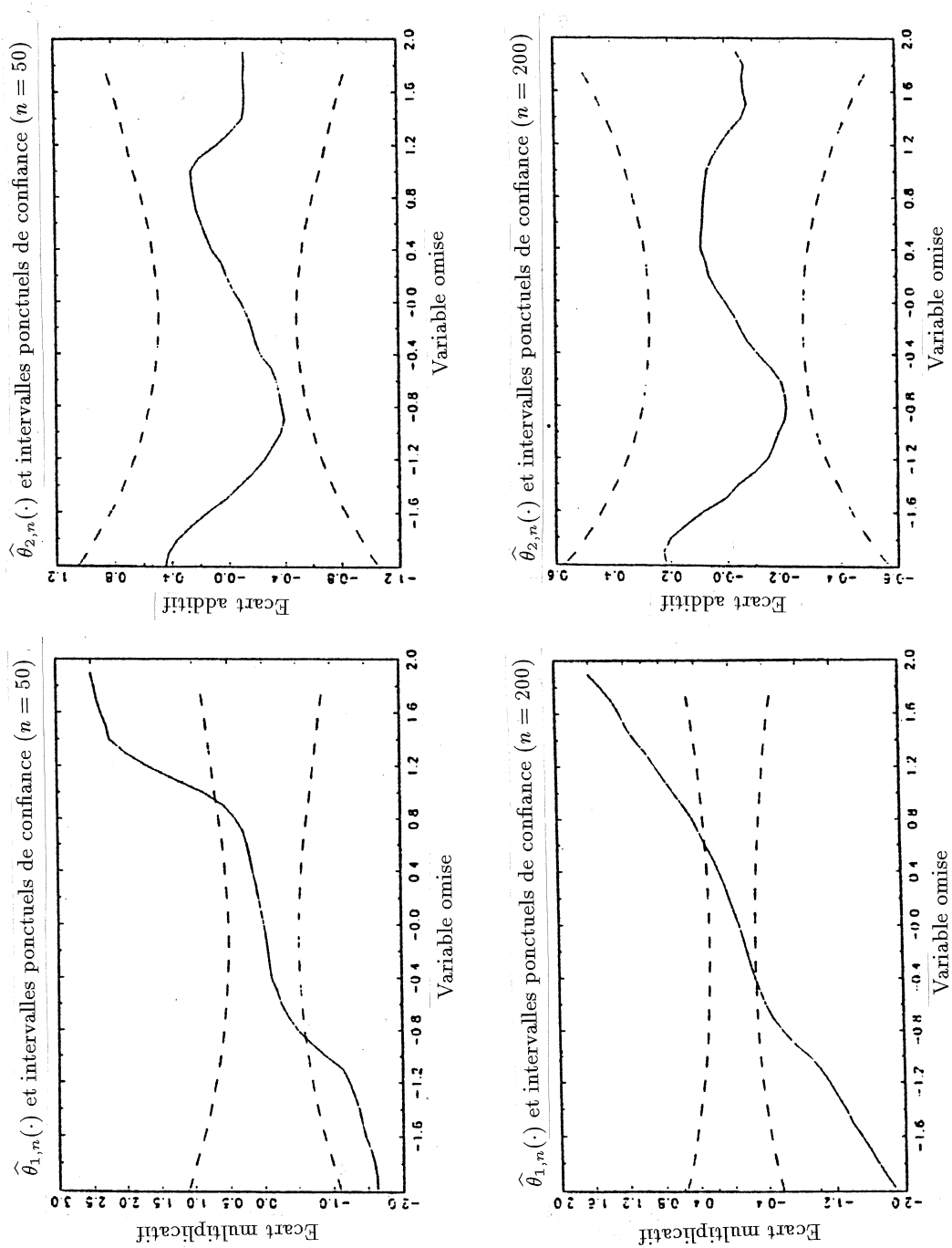


Figure 2.3.12: Estimateurs $\hat{\theta}_{1,n}(\cdot)$, $\hat{\theta}_{2,n}(\cdot)$, et intervalles ponctuels de confiance sous H_0 au niveau asymptotique 0.05 pour I), avec $h_n = 1$ et $n = 50$ et 200.

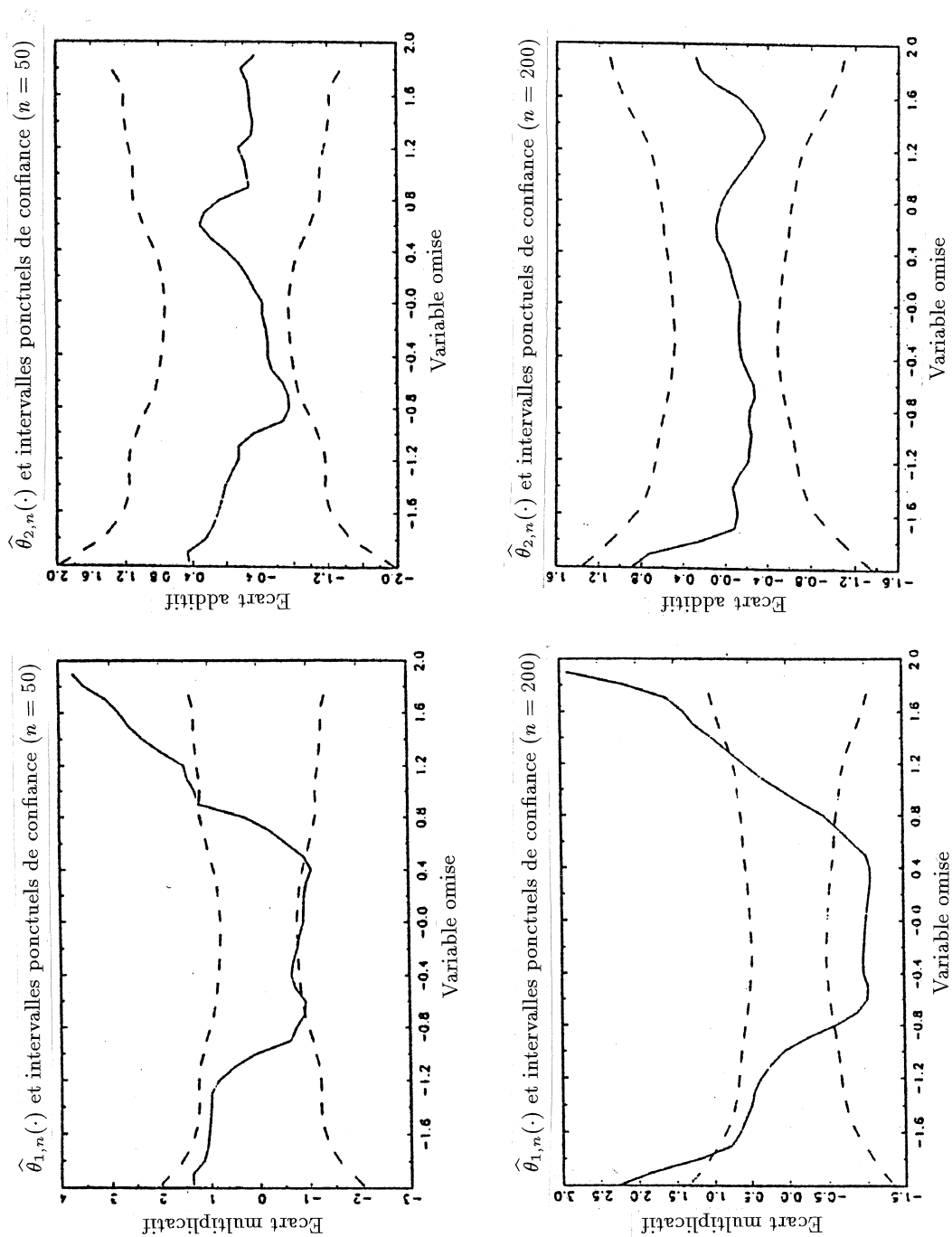


Figure 2.3.13: Estimateurs $\hat{\theta}_{1,n}(\cdot)$, $\hat{\theta}_{2,n}(\cdot)$, et intervalles ponctuels de confiance sous H_0 au niveau asymptotique 0.05 pour II), avec $h_n = 0.5$ et $n = 50$ et 200.

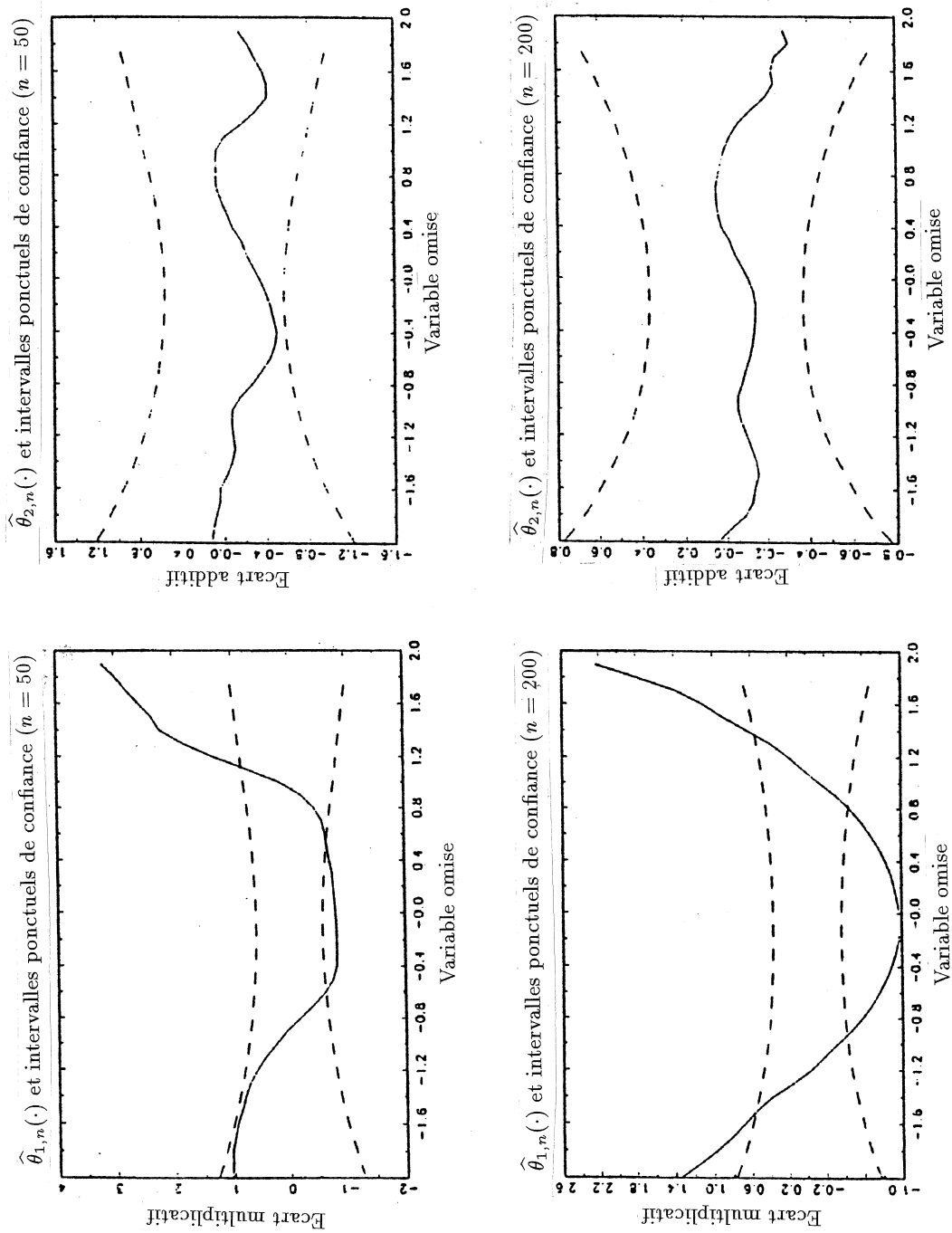


Figure 2.3.14: Estimateurs $\hat{\theta}_{1,n}(\cdot)$, $\hat{\theta}_{2,n}(\cdot)$, et intervalles ponctuels de confiance sous H_0 au niveau asymptotique 0.05 pour II), avec $h_n = 1$ et $n = 50$ et 200.

même noyau et la même fenêtre que dans l'estimateur $(\widehat{\theta}_{1,n}(\cdot), \widehat{\theta}_{2,n}(\cdot))$. De plus $\int_{\mathbb{R}} K^2(u) du = \frac{3}{5}$ pour le noyau d'Epanechnikov. Ainsi, sous H_0 , la distribution du vecteur $(\widehat{\theta}_{1,n}(z), \widehat{\theta}_{2,n}(z))$ est, pour $z \in \mathbb{R}$, approchée par

$$N\left(0, \frac{3\widehat{\sigma}_n^2}{5f_Z(z)h_n}(X'X)^{-1}\right),$$

ce qui permet d'obtenir des intervalles asymptotiques de confiance pour les valeurs $\theta_{1,\infty}(z)$ et $\theta_{2,\infty}(z)$, des paramètres fonctionnels $\theta_{1,\infty}(\cdot)$ et $\theta_{2,\infty}(\cdot)$ au point z .

Dans les Figures 2.3.11, 2.3.12, 2.3.13 et 2.3.14, on trouve les graphiques des estimateurs $\widehat{\theta}_{1,n}(\cdot)$ et $\widehat{\theta}_{2,n}(\cdot)$, ainsi que les intervalles ponctuels de confiance au niveau asymptotique 0.05, pour les problèmes I) et II). Pour chacun de ces problèmes nous avons considéré deux valeurs de h_n : 0.5 et 1, et deux valeurs de n : 50 et 200.

D'après les graphiques, on voit que $\widehat{\theta}_{2,n}(\cdot)$ est pour toutes les valeurs de z , dans l'intervalle asymptotique de confiance, tandis que $\widehat{\theta}_{1,n}(\cdot)$ est, pour quelques valeurs de z , au-dehors de cet intervalle, pour toutes les valeurs de h_n et n considérées. On peut aussi remarquer, que les curves sont plus régulières pour $h_n = 1$ que pour $h_n = 0.5$. Sa régularité augmente aussi avec n .

Cette démarche ne nous permet pas simplement de détecter l'oubli de la variable Z dans le modèle de régression linéaire considéré, mais elle permet aussi de conclure que cette variable doit être introduite dans le modèle à partir du paramètre b . De plus, le graphique de $\widehat{\theta}_{1,n}(\cdot)$ nous suggère qu'un tel effet multiplicatif est de type quadratique.

Chapitre 3

Théorèmes de limite centrale pour des U-statistiques dégénérées

Nous dérivons dans ce chapitre des conditions suffisantes pour la normalité asymptotique de la somme de U-statistiques dégénérées d'ordre 1 et 2, engendrées par une suite de processus stochastiques fortement stationnaires et géométriquement absolument réguliers. Ces résultats généralisent ceux de Hall [42] obtenus dans un cadre d'échantillonnage.

3.1 Définitions et notations

Soient $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$ une suite de processus stochastiques à valeurs dans \mathbb{R}^d , fortement stationnaires. Considérons les U-statistiques d'ordre 1 et 2 (cf. Hoeffding [51]),

$$\mathcal{G}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n g_n(X_{in}), \quad (3.1.1)$$

où $g_n(\cdot)$ est une application mesurable de \mathbb{R}^d dans \mathbb{R} , avec $E[g_n(X_{0n})] = 0$, et

$$\mathcal{H}_n = \frac{1}{n} \sum_{1 \leq j < i \leq n} \{h_n(X_{in}, X_{jn}) - E[h_n(X_{in}, X_{jn})]\}, \quad (3.1.2)$$

où $h_n(\cdot, \cdot)$ est une application mesurable symétrique de $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R} .

Nous admettons que la U-statistique (3.1.2) est dégénérée, c'est-à-dire,

$$E[h_n(X_{0n}, \cdot)] = 0, P_{X_{0n}} \text{ presque sûrement.}$$

En interprétant la condition $E[g_n(X_{0n})] = 0$ comme un cas limite de la condition précédente, nous dirons aussi que la U-statistique (3.1.1) est dégénérée.

Nous nous intéressons dans ce chapitre par l'étude du comportement asymptotique de la suite de vecteurs aléatoires $(\mathcal{G}_n, \mathcal{H}_n)$.

Dans un cadre d'indépendance asymptotique, des théorèmes de limite centrale pour la variable aléatoire \mathcal{G}_n sont bien connus. Pour cela nous faisons uniquement référence au travail de Ibragimov [54]. Une dérivation différente du résultat de Ibragimov se trouve dans Hall et Heyde [44] pg. 132. Dans le cas de la variable \mathcal{H}_n , les travaux de Khashimov [59] et Yoshihara [113] et [114] établissent des conditions suffisantes pour sa normalité asymptotique. Si dans Khashimov [59] les conditions de mélange imposées sont assez restrictives, dans Yoshihara [113] ces conditions sont beaucoup plus faibles, mais ils restent d'autres hypothèses assez restrictives pour les applications envisagées dans cette thèse. Le théorème de limite centrale que nous obtenons pour la U-statistique dégénérée de second ordre \mathcal{H}_n , dont la technique de démonstration est basée en Takahata et Yoshihara [99], généralise au cas mélangeant le Théorème 1 de Hall [42]. L'ensemble de conditions suffisantes que nous obtenons pour la normalité, n'est sûrement pas le plus général mais il établit un compromis entre la généralité et la simplicité. Ce résultat, ainsi que ceux obtenus pour \mathcal{G}_n et $(\mathcal{G}_n, \mathcal{H}_n)$, sont fondés sur une inégalité de Yoshihara [113] pour les processus absolument réguliers, et sur un théorème de limite centrale de Dvoretzky [29] pour des suites triangulaires. Nous commençons par rappeler ces résultats dans les deux sections suivantes.

3.2 Processus β -mélangeants. Quelques inégalités

Soit $(X_i, i \in \mathbb{Z})$ un processus stochastique fortement stationnaire à valeurs dans \mathbb{R}^d . Jusqu'à maintenant nous avons utilisé le coefficient de mélange $\alpha(i)$ comme mesure de dépendance entre le passé $F_{-\infty}^0$ et le futur $F_i^{+\infty}$ du processus $X_i, i \in \mathbb{Z}$. Dans ce chapitre nous considérons une autre mesure fondée sur le coefficient $\beta(i)$ dont l'introduction est attribuée en Volkonskiï et Rozanov [106] à A.N. Kolmogorov. Il est défini, pour $i \in \mathbb{N}^*$, par

$$\beta(i) = E \left[\sup_{A \in F_i^{+\infty}} |P(A|F_{-\infty}^0) - P(A)| \right]. \quad (3.2.1)$$

Nous disons que le processus $(X_i, i \in \mathbb{Z})$ est β -mélangeant (ou absolument régulier) si $\beta(i) \rightarrow 0$,

lorsque $i \rightarrow +\infty$. Des rapports entre ce type de mélange et ceux considérés dans les chapitres précédents peuvent se déduire des inégalités $2\alpha(i) \leq \beta(i) \leq \phi(i)$, pour $i \in \mathbb{N}^*$ (cf. Bradley [12]). Le cas de certains processus ARMA multivariés ou bilinéaires réels fondés sur des bruits blancs i.i.d. déjà cités en §1.2.1, sont des exemples de processus β -mélangeants. De plus, leurs coefficients de mélange de type β convergent vers zéro de façon exponentielle. D'autre part, on rappelle qu'une caractérisation des processus gaussiens β -mélangeants en terme de la densité spectral a été obtenue par Ibragimov et Solev [56].

Les inégalités que nous obtenons dans ce paragraphe, concernant le coefficient de mélange introduit ci-dessus, ont été obtenues dans les travaux de Yoshihara [112] et [113]. Contrairement à celles qui concernent les coefficients de type α ou ϕ , les inégalités pour le coefficient de mélange de type β ne sont pas usuelles dans la littérature. Pour cela, et parce qu'elles seront souvent utilisées dans ce chapitre, nous passons à leur présentation.

En étant donnés

$$i_1 < i_2 < \dots < i_j < i_{j+1} < \dots < i_k,$$

on désigne par Q , R et P respectivement, les lois des vecteurs aléatoires $Y = (X_{i_1}, \dots, X_{i_j})$, $Z = (X_{i_{j+1}}, \dots, X_{i_k})$ et $(X_{i_1}, \dots, X_{i_j}, X_{i_{j+1}}, \dots, X_{i_k})$, et par $Q \times R$ la mesure produit de Q par R .

Le résultat suivant est dû à Yoshihara [113]. Nous l'obtenons ici, en employant la technique de démonstration du Lemme 1 de Yoshihara [112] (Corollaire 3.2.3 suivant).

Lemme 3.2.2 *Soit h une application mesurable. S'il existe $M > 0$ et $\gamma > 0$ tels que*

$$\int_{(\mathbb{R}^d)^k} |h(x_1, \dots, x_k)|^{1+\gamma} dP \leq M \text{ et } \int_{(\mathbb{R}^d)^k} |h(x_1, \dots, x_k)|^{1+\gamma} d(Q \times R) \leq M,$$

alors

$$E|E[h(Y, Z)|Y] - H(Y)| \leq 4 M^{\frac{1}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j),$$

où $H(u) = E[h(u, Z)]$.

Démonstration: Désignons par $S(\cdot|y)$ la loi de Z conditionnée par $Y = y$, et par $|S(\cdot|y) - R(\cdot)|$ la variation totale de $S(\cdot|y) - R(\cdot)$ (cf. Halmos [45] pg. 122). On a

$$E|E[h(Y, Z)|Y] - H(Y)|$$

$$\begin{aligned}
&= \int_{(\mathbb{R}^d)^j} \left| \int_{(\mathbb{R}^d)^{k-j}} h(y, z) S(dz|y) - \int_{(\mathbb{R}^d)^{k-j}} h(y, z) R(dz) \right| Q(dy) \\
&\leq \int_{(\mathbb{R}^d)^j} \int_{(\mathbb{R}^d)^{k-j}} |h(y, z)| |S(\cdot|y) - R(\cdot)| (dz) Q(dy).
\end{aligned}$$

Soient maintenant $D = \{(x_1, \dots, x_k) \in (\mathbb{R}^d)^k \mid |h(x_1, \dots, x_k)| \leq M^{\frac{1}{1+\gamma}} \beta^{-\frac{1}{1+\gamma}} (i_{j+1} - i_j)\}$, D^c le complémentaire de D dans $(\mathbb{R}^d)^k$ et $\mathcal{B}(E)$ la tribu borélienne de E . Nous avons,

$$\begin{aligned}
&\int \int_D |h(y, z)| |S(\cdot|y) - R(\cdot)| (dz) Q(dy) \\
&\leq M^{\frac{1}{1+\gamma}} \beta^{-\frac{1}{1+\gamma}} (i_{j+1} - i_j) \int_{(\mathbb{R}^d)^j} |S(\cdot|y) - R(\cdot)| ((\mathbb{R}^d)^{k-j}) Q(dy) \\
&\leq 2M^{\frac{1}{1+\gamma}} \beta^{-\frac{1}{1+\gamma}} (i_{j+1} - i_j) \int_{(\mathbb{R}^d)^j} \sup_{B \in \mathcal{B}((\mathbb{R}^d)^{k-j})} |S(B|y) - R(B)| Q(dy) \\
&= 2M^{\frac{1}{1+\gamma}} \beta^{-\frac{1}{1+\gamma}} (i_{j+1} - i_j) E \left[\sup_{B \in \mathcal{F}_{i_{j+1}}^{i_j}} |P(B|F_{i_1}^{i_j}) - P(B)| \right] \\
&\leq 2M^{\frac{1}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j),
\end{aligned}$$

et aussi,

$$\begin{aligned}
&\int \int_{D^c} |h(y, z)| |S(\cdot|y) - R(\cdot)| (dz) Q(dy) \\
&= \int \int_{D^c} |h(y, z)|^{1+\gamma} |h(y, z)|^{-\gamma} |S(\cdot|y) - R(\cdot)| (dz) Q(dy) \\
&\leq M^{-\frac{\gamma}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j) \left\{ \int \int_{D^c} |h(y, z)|^{1+\gamma} S(dz|y) Q(dy) \right. \\
&\quad \left. + \int \int_{D^c} |h(y, z)|^{1+\gamma} R(dz) Q(dy) \right\} \\
&\leq 2M^{\frac{1}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j).
\end{aligned}$$

Ainsi, d'après les inégalités précédentes, on obtient

$$E|E[h(Y, Z)|Y] - H(Y)| \leq 4M^{\frac{1}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j). \quad \blacksquare$$

En reprenant la démonstration précédente, on conclut facilement que sous les conditions du Lemme 3.2.2, si l'application h satisfait $|h| \leq C$, avec $C > 0$, alors l'inégalité y présentée devient

$$E|E[h(Y, Z)|Y] - H(Y)| \leq 2C\beta(i_{j+1} - i_j).$$

Les inégalités suivantes sont des conséquences immédiates du Lemme 3.2.2.

Corollaire 3.2.3 *Sous les hypothèses du Lemme 3.2.2, on a*

$$|E[h(Y, Z)] - E[H(Y)]| \leq 4M^{\frac{1}{1+\gamma}} \beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j).$$

Corollaire 3.2.4 *Soient g et h des applications mesurables telles que $E|g(Y)|^s < \infty$ et $E|h(Z)|^t < \infty$, avec $s, t > 1$ et $\frac{1}{s} + \frac{1}{t} < 1$. Alors*

$$|E[g(Y)h(Z)] - E[g(Y)]E[h(Z)]| \leq 4\beta^{1-\frac{1}{s}-\frac{1}{t}} (i_{j+1} - i_j) E^{\frac{1}{s}}|g(Y)|^s E^{\frac{1}{t}}|h(Z)|^t.$$

Corollaire 3.2.5 *Soit g une application mesurable telle que $E|g(Z)|^{1+\gamma} < \infty$, pour un $\gamma > 0$. Alors*

$$E|E[g(Z)|Y] - E[g(Z)]| \leq 4\beta^{\frac{\gamma}{1+\gamma}} (i_{j+1} - i_j) E^{\frac{1}{1+\gamma}}|g(Z)|^{1+\gamma}.$$

3.3 Théorème de limite centrale pour des suites triangulaires

Le théorème de limite centrale pour des suites triangulaires que nous présentons dans la suite, joue un rôle fondamental dans l'étude de la loi asymptotique des U-statistiques dégénérées d'ordre 2. Nous utiliserons le Théorème 2.2 présenté dans Dvoretzky [29] (voir aussi le troisième chapitre de Hall et Heyde [44]) en choisissant une suite de tribus plus fine que celle engendrée par les sommes partielles (voir remarque 3.1 de Dvoretzky) et une condition iii) impliquant la condition de Lindeberg conditionnelle imposée dans le Théorème 2.2 de Dvoretzky.

Théorème 3.3.1 *Soient X_{kn} , $k=1, 2, \dots, k_n$, $n \geq 1$, des variables aléatoires \mathcal{F}_{kn} -mesurables, où les tribus \mathcal{F}_{kn} satisfont $\mathcal{F}_{k-1,n} \subset \mathcal{F}_{kn}$, $k=1, 2, \dots, k_n$, $n \geq 1$ (on prend pour \mathcal{F}_{0n} la tribu grossière). On suppose:*

- i) $\sum_{k=1}^{k_n} E[X_{kn}|\mathcal{F}_{k-1,n}] = o_p(1)$;
- ii) $\sum_{k=1}^{k_n} \text{Var}[X_{kn}|\mathcal{F}_{k-1,n}] = \sigma^2 + o_p(1)$;
- iii) *il existe $\delta > 0$ tel que $\sum_{k=1}^{k_n} E|X_{kn}|^{2+\delta} = o(1)$.*

Alors $S_n = \sum_{k=1}^{k_n} X_{kn}$ est asymptotiquement normal de moyenne zéro et de variance σ^2 .

3.4 Théorèmes de limite centrale pour \mathcal{G}_n , \mathcal{H}_n et $\mathcal{G}_n + \mathcal{H}_n$

Dans cette section nous désignons par $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$ une suite de processus stochastiques de dimension d , fortement stationnaires. On suppose qu'il existe $C > 0$ et $\rho \in]0, 1[$, tels que

$$\sup_{n \in \mathbb{N}^*} \beta_n(i) \leq C\rho^i, \forall i \in \mathbb{N}^*,$$

avec

$$\beta_n(i) = E \left[\sup_{A \in F_i^{+\infty}(n)} |P(A|F_{-\infty}^0(n)) - P(A)| \right], \quad (3.4.1)$$

où pour $i, j \in \mathbb{Z} \cup \{-\infty, +\infty\}$, $F_i^j(n)$ désigne la tribu engendrée par les variables aléatoires $\{X_{kn}, i \leq k \leq j\}$.

Nous établissons dans la suite, des théorèmes de limite centrale pour les suites de variables aléatoires \mathcal{G}_n , \mathcal{H}_n et $\mathcal{G}_n + \mathcal{H}_n$ définies dans 3.1. Notons pour $r > 0$, $\|\xi\|_r = E^{1/r}|\xi|^r$, pour toute variable aléatoire ξ telle que $E|\xi|^r < \infty$.

3.4.1 Loi asymptotique de \mathcal{G}_n

Dans le lemme suivant nous présentons un développement asymptotique de la U-statistique dégénérée d'ordre un, \mathcal{G}_n , définie par (3.1.1).

Lemme 3.4.2 *On suppose qu'il existe $\delta_0 > 0$ tel que:*

- i) $\|g_n(X_{0n})\|_{2+\delta_0} = O(1)$;
- ii) $E[g_n(X_{jn})g_n(X_{0n})] = c_j + o(1)$, pour tout $j = 0, 1, 2, \dots$

Alors la variable aléatoire \mathcal{G}_n est asymptotiquement équivalente à $\sum_{\alpha=1}^k T_{\alpha,n}$, où pour $\alpha = 1, 2, \dots, k$,

$$T_{\alpha,n} = \frac{1}{\sqrt{n}} \sum_{i=\alpha(r+m)-m+1}^{\alpha(r+m)} g_n(X_{in}), \quad (3.4.3)$$

avec $k = k(n) = \lfloor \frac{n}{m+r} \rfloor$, $m = m(n) = \lfloor n^{\delta_1} \rfloor$, $r = r(n) = \lfloor n^{\delta_2} \rfloor$ ($\lfloor x \rfloor$ désigne la partie entière du réel x), et $0 < \delta_2 < \delta_1 < \frac{1}{2}(\frac{\delta_0}{1+\delta_0})$.

De plus, on a:

$$\begin{aligned} D1) \quad & \sum_{\alpha=1}^k E[T_{\alpha,n}|F_{\alpha-1,n}] = o_p(1), \\ D2) \quad & \sum_{\alpha=1}^k \text{Var}[T_{\alpha,n}|F_{\alpha-1,n}] = \sigma_1^2 + o_p(1), \text{ et} \end{aligned} \quad (3.4.4)$$

$$D3) \quad \sum_{\alpha=1}^k E|T_{\alpha,n}|^{2+\delta_0} = o(1),$$

où σ_1^2 est donné par $c_0 + 2 \sum_{j=1}^{\infty} c_j$, et $F_{\alpha,n}$ désigne la tribu $F_1^{\alpha(r+m)+1}(n)$, pour $\alpha=1, 2, \dots, k$.

Démonstration: Considérons la décomposition

$$\mathcal{G}_n = \frac{1}{\sqrt{n}} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} g_n(X_{in}) + \frac{1}{\sqrt{n}} \sum_{\alpha=0}^k \sum_{i=b_{\alpha+1}}^{a_{\alpha+1}-1} g_n(X_{in}),$$

où $b_0 = 0$, $a_i = b_{i-1} + r + 1$, $b_i = a_i + m - 1$ ($i = 1, 2, \dots, k$) et $a_{k+1} = n + 1$, et $k(n)$, $m(n)$ et $r(n)$ sont définis dans l'énoncé. Cette décomposition introduit des sommes comportant respectivement des paquets de termes successifs de tailles m et r respectivement et distants de r et m . On conclut facilement que le terme correspondant aux paquets de taille petite converge en probabilité vers zéro, lorsque n tend vers l'infini. Ainsi, on a

$$\mathcal{G}_n = \sum_{\alpha=1}^k T_{\alpha,n} + o_p(1),$$

où pour $\alpha=1, 2, \dots, k$, $T_{\alpha,n}$ est défini par (3.4.3).

Etablissons maintenant les conditions D1), D2) et D3) de (3.4.4).

A) Etablissement de la condition D3)

D'après la stationnarité du processus X_{in} , $i \in \mathbb{Z}$, l'hypothèse i) et le choix de δ_1 avec $\delta_1 < \frac{1}{2}(\frac{\delta_0}{1+\delta_0})$, on a

$$\begin{aligned} \sum_{\alpha=1}^k E|T_{\alpha,n}|^{2+\delta_0} &= kE|T_{1,n}|^{2+\delta_0} \\ &\leq k \left(\frac{1}{\sqrt{n}} \sum_{i=1}^m \|g_n(X_{in})\|_{2+\delta_0} \right)^{2+\delta_0} \\ &= O \left(k \left(\frac{m}{\sqrt{n}} \right)^{2+\delta_0} \right) \\ &= O \left(n^{\delta_1(1+\delta_0)-\delta_0/2} \right) = o(1). \end{aligned}$$

B) Etablissement de la condition D1)

Nous montrons la convergence L_1 qui impliquera la convergence en probabilité. Comme d'après le point précédent $\|T_{\alpha,n}\|_{2+\delta_0} = O(\frac{m}{\sqrt{n}})$, l'application du Corollaire 3.2.5 avec $\gamma = 1 + \delta_0$,

nous permet d'écrire

$$\begin{aligned} E \left| \sum_{\alpha=1}^k E[T_{\alpha,n} | F_{\alpha-1,n}] \right| &\leq \sum_{\alpha=1}^k E |E[T_{\alpha,n} | F_{\alpha-1,n}] - E[T_{\alpha,n}]| \\ &= O \left(\sqrt{n} \beta_n^{\frac{1+\delta_0}{2+\delta_0}}(r) \right), \end{aligned}$$

et le majorant est négligeable d'après la convergence exponentielle vers zéro des coefficients de mélange, et du fait que r tend vers l'infini avec n .

C) Etablissement de la condition D2)

Il résulte du point précédent que $\sum_{\alpha=1}^k |E[T_{\alpha,n} | F_{\alpha-1,n}]| = o_p(1)$. Il suffit alors pour établir la condition D2), de montrer que $\sum_{\alpha=1}^k E[T_{\alpha,n}^2 | F_{\alpha-1,n}] = \sigma_1^2 + o(1)$. Pour cela, nous allons montrer que

$$\sum_{\alpha=1}^k E[T_{\alpha,n}^2] = \sigma_1^2 + o(1) \text{ et } \sum_{\alpha=1}^k \{E[T_{\alpha,n}^2 | F_{\alpha-1,n}] - E[T_{\alpha,n}^2]\} = o_p(1).$$

En effet, d'après les hypothèses i) et ii),

$$\begin{aligned} \sum_{\alpha=1}^k E[T_{\alpha,n}^2] &= kE[T_{1,n}^2] = \frac{km}{n} E \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m g_n(X_{in}) \right)^2 \\ &= \frac{km}{n} \left(E[g_n(X_{0n})]^2 + \frac{2}{m} \sum_{j=1}^{m-1} (m-j) E[g_n(X_{jn})g_n(X_{0n})] \right) \\ &= c_0 + 2 \sum_{j=1}^{\infty} c_j + o(1) = \sigma_1^2 + o(1), \end{aligned}$$

puisque $\frac{km}{n} \rightarrow 1, n \rightarrow +\infty$. De plus, en tenant en compte que $\|T_{\alpha,n}^2\|_{1+\delta_0/2} = O\left(\frac{m^2}{n}\right)$, l'application du Corollaire 3.2.5 avec $\gamma = \delta_0/2$, nous permet d'écrire

$$\begin{aligned} E \left| \sum_{\alpha=1}^k \{E[T_{\alpha,n}^2 | F_{\alpha-1,n}] - E[T_{\alpha,n}^2]\} \right| &\leq \sum_{\alpha=1}^k E |E[T_{\alpha,n}^2 | F_{\alpha-1,n}] - E[T_{\alpha,n}^2]| \\ &= O \left(\frac{km^2}{n} \beta_n^{\frac{\delta_0}{2+\delta_0}}(r) \right) = o(1), \end{aligned}$$

d'après la convergence exponentielle vers zéro des coefficients de mélange. ■

La normalité asymptotique de la variable aléatoire \mathcal{G}_n , découle maintenant du lemme précédent et de l'application du Théorème 3.3.1 à la suite doublement indexée de variables $T_{\alpha,n}$ et de tribus $F_{\alpha,n}$, pour $\alpha=1, 2, \dots, k$.

Théorème 3.4.5 *Sous les conditions du Lemme 3.4.2, la variable aléatoire \mathcal{G}_n est asymptotiquement normale de moyenne zéro et de variance σ_1^2 donnée par $c_0 + 2 \sum_{j=1}^{\infty} c_j$.*

3.4.2 Loi asymptotique de \mathcal{H}_n

Soit, pour $n \in \mathbb{N}^*$, \mathcal{H}_n la U-statistique dégénérée d'ordre 2 définie par (3.1.2). Pour $n \in \mathbb{N}^*$, notons par $(\bar{X}_{in}, i \in \mathbb{Z})$ une copie du processus $(X_{in}, i \in \mathbb{Z})$ qui lui est indépendante. Pour $r > 0$, introduisons les notations suivantes:

$$\begin{aligned} u_n(r) &= \max \left\{ \max_{1 \leq i \leq n} \|h_n(X_{in}, X_{0n})\|_r, \|h_n(X_{0n}, \bar{X}_{0n})\|_r \right\} \\ v_n(r) &= \max \left\{ \max_{1 \leq i \leq n} \|G_{n0}(X_{in}, X_{0n})\|_r, \|G_{n0}(X_{0n}, \bar{X}_{0n})\|_r \right\} \\ w_n(r) &= \|G_{n0}(X_{0n}, X_{0n})\|_r \\ z_n(r) &= \max_{0 \leq i \leq n, 1 \leq j \leq n} \left\{ \|G_{nj}(X_{in}, X_{0n})\|_r, \|G_{nj}(X_{0n}, X_{in})\|_r, \|G_{nj}(X_{0n}, \bar{X}_{0n})\|_r \right\} \end{aligned} \quad (3.4.6)$$

où $G_{nj}(u, v) = E[h_n(X_{jn}, u)h_n(X_{0n}, v)]$, pour $j \in \mathbb{N}$ et $u, v \in \mathbb{R}^d$.

Nous présentons dans le lemme suivant un développement asymptotique de la variable \mathcal{H}_n . La technique de démonstration que nous employons est fondée en Takahata et Yoshihara [99]. Si dans un cadre d'indépendance la variable aléatoire \mathcal{H}_n est une martingale (cf. Hall [42]), dans le cas présent d'indépendance asymptotique, il découle du lemme suivant qu'elle est asymptotiquement équivalente à une martingale.

Lemme 3.4.7 *On suppose qu'il existe $\delta_0 > 0$, $\gamma_0 < \frac{1}{2}$ et $\gamma_1 > 0$, tels que:*

- i) $u_n(4 + \delta_0) = O(n^{\gamma_0})$;
- ii) $v_n(2 + \frac{\delta_0}{2}) = o(1)$;
- iii) $w_n(2 + \frac{\delta_0}{2}) = o(n^{\frac{1}{2}})$;
- iv) $z_n(2)n^{\gamma_1} = O(1)$;
- v) $E[G_{n0}(X_{0n}, X_{0n})] = 2\sigma_2^2 + o(1)$.

Alors la variable aléatoire \mathcal{H}_n est asymptotiquement équivalente à $\sum_{\alpha=1}^k U_{\alpha,n}$, où pour $\alpha = 1, 2, \dots, k$,

$$U_{\alpha,n} = \frac{1}{n} \sum_{i=\alpha(r+m)-m+1}^{\alpha(r+m)} \sum_{j=1}^{(\alpha-1)(r+m)+1} h_n(X_{in}, X_{jn}), \quad (3.4.8)$$

avec $k = k(n) = \lfloor \frac{n}{m+r} \rfloor$, $m = m(n) = \lfloor n^{\delta_1} \rfloor$, $r = r(n) = \lfloor n^{\delta_2} \rfloor$ et $0 < \delta_2 < \delta_1 < \min(\gamma_1/2, (1 - 2\gamma_0)/3)$.

De plus, on a:

$$\begin{aligned}
D1) \quad & \sum_{\alpha=1}^k E[U_{\alpha,n}|F_{\alpha-1,n}] = o_p(1), \\
D2) \quad & \sum_{\alpha=1}^k \text{Var}[U_{\alpha,n}|F_{\alpha-1,n}] = \sigma_2^2 + o_p(1), \text{ et} \\
D3) \quad & \sum_{\alpha=1}^k E|U_{\alpha,n}|^4 = o(1),
\end{aligned} \tag{3.4.9}$$

où $F_{\alpha,n}$ est la tribu $F_1^{\alpha(r+m)+1}(n)$, pour $\alpha=1, 2, \dots, k$.

Démonstration: Prenons la décomposition

$$\begin{aligned}
\mathcal{H}_n &= \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} \bar{h}_n(X_{in}, X_{jn}) + \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=a_\alpha-r+1}^{i-1} \bar{h}_n(X_{in}, X_{jn}) \\
&+ \frac{1}{n} \sum_{\alpha=0}^k \sum_{i=b_\alpha+1}^{a_{\alpha+1}-1} \sum_{j=1}^{b_\alpha+1-r} \bar{h}_n(X_{in}, X_{jn}) + \frac{1}{n} \sum_{\alpha=0}^k \sum_{i=b_\alpha+1}^{a_{\alpha+1}-1} \sum_{j=b_\alpha+2-r}^{i-1} \bar{h}_n(X_{in}, X_{jn}),
\end{aligned}$$

où $\bar{h}_n(X_{in}, X_{jn}) = h_n(X_{in}, X_{jn}) - E[h_n(X_{in}, X_{jn})]$, $b_0 = 0$, $a_i = b_{i-1} + r + 1$, $b_i = a_i + m - 1$ ($i=1, 2, \dots, k$) et $a_{k+1} = n + 1$. Les suites $k(n)$, $m(n)$ et $r(n)$ sont définies dans l'énoncé.

Les trois derniers termes de l'égalité précédente contiennent des paquets de petite taille. Une démarche analogue à celle employée dans Yoshihara [113] nous permet de conclure que d'après les hypothèses i), iv), v) et du choix de δ_1 et δ_2 , ces termes convergent en probabilité vers zéro.

Ainsi, nous avons

$$\mathcal{H}_n = \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} \{h_n(X_{in}, X_{jn}) - E[h_n(X_{in}, X_{jn})]\} + o_p(1).$$

L'application du Corollaire 3.2.3 avec $\gamma = 3 + \delta_0$, l'hypothèse i) et la convergence exponentielle vers zéro du coefficient de mélange permettent de conclure que

$$\begin{aligned}
& \left| \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E[h_n(X_{in}, X_{jn})] \right| \\
& \leq \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} |E[h_n(X_{in}, X_{jn})]| \\
& = \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} |E[h_n(X_{in}, X_{jn})] - E[h_n(X_{in}, \bar{X}_{jn})]|
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} 4u_n(4+\delta_0)\beta_n^{\frac{3+\delta_0}{4+\delta_0}}(r) \\
&= O\left(\frac{kmnu_n(4+\delta_0)\beta_n^{\frac{3+\delta_0}{4+\delta_0}}(r)}{n}\right) = o(1).
\end{aligned}$$

On a alors

$$\mathcal{H}_n = \sum_{\alpha=1}^k U_{\alpha,n} + o_p(1),$$

où pour $\alpha=1, 2, \dots, k$, $U_{\alpha,n}$ est défini par (3.4.8).

Etablissons maintenant les conditions D1), D2) et D3) de (3.4.9).

A) Etablissement de la condition D1)

En appliquant le Lemme 3.2.2 avec $\gamma = 3 + \delta_0$, on obtient, comme ci-dessus,

$$\begin{aligned}
&E\left|\sum_{\alpha=1}^k E[U_{\alpha,n}|F_{\alpha-1,n}]\right| \\
&\leq \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E|E[h_n(X_{in}, X_{jn})|F_{\alpha-1,n}]| \\
&= \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E|E[h_n(X_{in}, X_{jn})|F_{\alpha-1,n}] - E[h_n(X_{in}, \bar{X}_{jn})]| \\
&\leq \frac{1}{n} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} 4u_n(4+\delta_0)\beta_n^{\frac{3+\delta_0}{4+\delta_0}}(r),
\end{aligned}$$

et le majorant est négligeable d'après l'hypothèse i) et la convergence exponentielle vers zéro des coefficients de mélange.

B) Etablissement de la condition D2)

Il résulte du point précédent que $\sum_{\alpha=1}^k |E[U_{\alpha,n}|F_{\alpha-1,n}]| = o_p(1)$, et donc, il suffit, pour établir la condition D2), de montrer que

$$\sum_{\alpha=1}^k E[U_{\alpha,n}^2|F_{\alpha-1,n}] = \sigma_2^2 + o_p(1). \quad (3.4.10)$$

Notons pour $\alpha = 1, \dots, k$,

$$\tilde{E}[U_{\alpha,n}^2|F_{\alpha-1,n}] = \frac{1}{n^2} \sum_{i,i'=a_\alpha}^{b_\alpha} \sum_{j,j'=1}^{a_\alpha-r} G_{n,|i-i'|}(X_{jn}, X_{j'n}),$$

où G_{ni} est défini pour $u, v \in \mathbb{R}^d$ par $G_{ni}(u, v) = E[h_n(X_{in}, u)h_n(X_{0n}, v)]$.

L'application du Lemme 3.2.2 avec $\gamma = 2 + \delta_0/2$, permet de conclure que

$$\begin{aligned} & E \left| \sum_{\alpha=1}^k E[U_{\alpha,n}^2 | F_{\alpha-1,n}] - \sum_{\alpha=1}^k \tilde{E}[U_{\alpha,n}^2 | F_{\alpha-1,n}] \right| \\ & \leq \frac{1}{n^2} \sum_{\alpha=1}^k \sum_{i,i'=a_\alpha}^{b_\alpha} \sum_{j,j'=1}^{a_\alpha-r} E | E[h_n(X_{in}, X_{jn})h_n(X_{i'n}, X_{j'n}) | F_{\alpha-1,n}] - G_{n,|i-i'|}(X_{jn}, X_{j'n}) | \\ & = O \left(km^2 u_n^2 (4 + \delta_0) \beta_n^{\frac{4+\delta_0}{6+\delta_0}}(r) \right) = o(1). \end{aligned}$$

Il suffit donc, de montrer (3.4.10) avec $E[U_{\alpha,n}^2 | F_{\alpha-1,n}]$ remplacé par $\tilde{E}[U_{\alpha,n}^2 | F_{\alpha-1,n}]$.

Considérons la décomposition

$$\sum_{\alpha=1}^k \tilde{E}[U_{\alpha,n}^2 | F_{\alpha-1,n}] = S_{1n} + S_{2n} + 2S_{3n},$$

avec

$$\begin{aligned} S_{1n} &= \frac{m}{n^2} \sum_{\alpha=1}^k \sum_{j=1}^{a_\alpha-r} G_{n0}(X_{jn}, X_{jn}), \quad S_{2n} = \frac{m}{n^2} \sum_{\alpha=1}^k \sum_{\substack{j,j'=1 \\ j \neq j'}}^{a_\alpha-r} G_{n0}(X_{jn}, X_{j'n}) \text{ et} \\ S_{3n} &= \frac{1}{n^2} \sum_{\alpha=1}^k \sum_{i=1}^{m-1} (m-i) \sum_{j,j'=1}^{a_\alpha-r} G_{ni}(X_{jn}, X_{j'n}). \end{aligned}$$

On montrera dans la suite que chacun des termes $S_{1n} - E[S_{1n}]$, S_{2n} et S_{3n} converge en probabilité vers zéro, lorsque n tend vers l'infini, ce qui permettra de conclure, puisque d'après l'hypothèse v) on a

$$\begin{aligned} E[S_{1n}] &= \frac{m}{n^2} \sum_{\alpha=1}^k \sum_{j=1}^{a_\alpha-r} E[G_{n0}(X_{0n}, X_{0n})] = \frac{m}{n^2} \sum_{\alpha=1}^k a_\alpha E[G_{n0}(X_{0n}, X_{0n})] + O\left(\frac{r}{n}\right) \\ &= \frac{m(m+r)(1+k)k}{2n^2} E[G_{n0}(X_{0n}, X_{0n})] + o(1) = \sigma_2^2 + o(1). \end{aligned}$$

$$\mathbf{B}_1) \quad \underline{S_{1n} - E[S_{1n}] = o_p(1)}.$$

On a

$$E[S_{1n} - E[S_{1n}]]^2 = \frac{m^2}{n^4} \sum_{\alpha,\alpha'=1}^k \sum_{j=1}^{a_\alpha-r} \sum_{j'=1}^{a_{\alpha'}-r} E[\bar{G}_{n0}(X_{jn})\bar{G}_{n0}(X_{j'n})],$$

où pour $u \in \mathbb{R}^d$, on note $\bar{G}_{n0}(u) = G_{n0}(u, u) - E[G_{n0}(X_{0n}, X_{0n})]$. L'application du Corollaire 3.2.4 avec $s = t = 2 + \delta_0/2$ nous permet d'écrire

$$\begin{aligned} |E[\bar{G}_{n0}(X_{jn})\bar{G}_{n0}(X_{j'n})]| &\leq 4\beta_n^{\frac{\delta_0}{4+\delta_0}} (|j - j'|) \|\bar{G}_{n0}(X_{jn})\|_{2+\delta_0/2} \|\bar{G}_{n0}(X_{j'n})\|_{2+\delta_0/2} \\ &\leq 16\beta_n^{\frac{\delta_0}{4+\delta_0}} (|j - j'|) w_n^2 (2 + \frac{\delta_0}{2}), \end{aligned}$$

et donc, d'après l'hypothèse iii) on obtient pour $E[S_{1n} - E[S_{1n}]]^2$ la majoration

$$O\left(\frac{m^2 k^2 n w_n^2 (2 + \frac{\delta_0}{2})}{n^4}\right) = O\left(\frac{w_n^2 (2 + \frac{\delta_0}{2})}{n}\right) = o(1).$$

B₂) $\underline{S_{2n} = o_p(1)}$.

On a

$$E[S_{2n}]^2 = \frac{4m^2}{n^4} \sum_{\alpha, \alpha'=1}^k \sum_{\substack{j_1, j_2=1 \\ j_1 < j_2}}^{a_\alpha - r} \sum_{\substack{j_3, j_4=1 \\ j_3 < j_4}}^{a_{\alpha'} - r} E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})].$$

On détermine maintenant des majorations pour le terme général précédent, où on suppose que $j_1 \leq j_3 \leq j_2 \leq j_4$ (de façon analogue on étudie les autres cas).

a) Si $\max(j_3 - j_1, j_4 - j_2) > r$, on trouve d'après l'application du Corollaire 3.2.3 avec $\gamma = \delta_0/4$, la majoration

$$|E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})]| \leq 4v_n^2 (2 + \frac{\delta_0}{2}) \beta_n^{\frac{\delta_0}{4 + \delta_0}}(r),$$

puisque $E[G_{n0}(\bar{X}_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})] = E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, \bar{X}_{j_4 n})] = 0$.

b) Si $\max(j_3 - j_1, j_4 - j_2) \leq r$ et $j_2 - j_3 > r$, on considère la décomposition

$$\begin{aligned} & |E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})]| \\ & \leq |E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})] - E[G_{n0}(\bar{X}_{j_1 n}, X_{j_2 n}) G_{n0}(\bar{X}_{j_3 n}, X_{j_4 n})]| \\ & \quad + |E[G_{n0}(\bar{X}_{j_1 n}, X_{j_2 n}) G_{n0}(\bar{X}_{j_3 n}, X_{j_4 n})]|. \end{aligned}$$

Pour le premier terme on obtient la même majoration que dans a). Pour le deuxième, on montre qu'il est majoré par $\|G_{n, j_3 - j_1}(X_{0n}, \bar{X}_{0n})\|_2 \|G_{n, j_4 - j_2}(X_{0n}, \bar{X}_{0n})\|_2$, et donc par $z_n^2(2)$ si $\min(j_3 - j_1, j_4 - j_2) > 0$, par $v_n^2(2)$ si $\max(j_3 - j_1, j_4 - j_2) = 0$, et par $z_n(2)v_n(2)$ en cas contraire.

c) Finalement, si $\max(j_3 - j_1, j_2 - j_3, j_4 - j_2) \leq r$, on a

$$|E[G_{n0}(X_{j_1 n}, X_{j_2 n}) G_{n0}(X_{j_3 n}, X_{j_4 n})]| \leq v_n^2(2).$$

Les majorations trouvées dans les points précédents nous permettent de conclure, d'après les hypothèses ii) et iv) et le choix de δ_1 avec $\delta_1 < \max(\gamma_1, 1/3)$, que $E[S_{2n}]^2 = o(1)$.

B₃) $S_{3n} = o_p(1)$.

On a

$$E[S_{3n}]^2 = \frac{1}{n^4} \sum_{\alpha, \alpha'=1}^k \sum_{i, i'=1}^{m-1} \sum_{j_1, j_2=1}^{a_\alpha-r} \sum_{j_3, j_4=1}^{a_{\alpha'}-r} (m-i)(m-i') E[G_{ni}(X_{j_1n}, X_{j_2n}) G_{ni'}(X_{j_3n}, X_{j_4n})].$$

Soient $s_1 \leq s_2 \leq s_3 \leq s_4$ les indices j_1, j_2, j_3, j_4 écrits par ordre croissant.

a) Si $\max(s_2 - s_1, s_4 - s_3) > r$, on a d'après le Corollaire 3.2.3 avec $\gamma = \delta_0/4$,

$$|E[G_{ni}(X_{j_1n}, X_{j_2n}) G_{ni'}(X_{j_3n}, X_{j_4n})]| \leq 4u_n^4 (4 + \delta_0) \beta_n^{\frac{\delta_0}{4+\delta_0}}(r).$$

b) Si $\max(s_2 - s_1, s_4 - s_3) \leq r$, on a

$$|E[G_{ni}(X_{j_1n}, X_{j_2n}) G_{ni'}(X_{j_3n}, X_{j_4n})]| \leq z_n^2(2).$$

Ces majorations permettent de conclure, d'après l'hypothèse iv) et le choix de δ_1 et δ_2 avec $\delta_1, \delta_2 < \gamma_1/2$, que $E[S_{3n}]^2$ est négligeable.

C) Etablissement de la condition D3)

Finalement

$$\begin{aligned} \sum_{\alpha=1}^k E[U_\alpha]^4 &\leq \frac{1}{n^4} \sum_{\alpha=1}^k \sum_{i_1, \dots, i_4=a_\alpha}^{b_\alpha} \sum_{j_1, \dots, j_4=1}^{a_\alpha-r} \left| E \left[\prod_{t=1}^4 h_n(X_{i_tn}, X_{j_tn}) \right] - E \left[\prod_{t=1}^4 h_n(X_{i_tn}, \bar{X}_{j_tn}) \right] \right| \\ &\quad + \frac{1}{n^4} \sum_{\alpha=1}^k \sum_{i_1, \dots, i_4=a_\alpha}^{b_\alpha} \sum_{j_1, \dots, j_4=1}^{a_\alpha-r} \left| E \left[\prod_{t=1}^4 h_n(X_{i_tn}, \bar{X}_{j_tn}) \right] \right|. \end{aligned}$$

Le premier terme de la somme précédente est négligeable puisque d'après l'application du Corollaire 3.2.3 avec $\gamma = \delta_0/4$, on conclut qu'il est d'ordre $km^4 u_n^4 (4 + \delta_0) \beta_n^{\frac{\delta_0}{4+\delta_0}}(r)$, car

$$\max \left\{ \left\| \prod_{t=1}^4 h_n(X_{i_tn}, X_{j_tn}) \right\|_{1+\delta_0/4}, \left\| \prod_{t=1}^4 h_n(X_{i_tn}, \bar{X}_{j_tn}) \right\|_{1+\delta_0/4} \right\} \leq u_n^4 (4 + \delta_0).$$

Pour l'étude du second terme, il nous faut analyser divers cas. Il est égale à

$$\frac{1}{n^4} \sum_{\alpha=1}^k \sum_{i_1, \dots, i_4=a_\alpha}^{b_\alpha} \sum_{j_1, \dots, j_4=1}^{a_\alpha-r} |E[g_{i_1, \dots, i_4}(X_{j_1n}, \dots, X_{j_4n})]|,$$

où pour $u_1, \dots, u_4 \in \mathbb{R}^d$,

$$g_{i_1, \dots, i_4}(u_1, \dots, u_4) = E \left[\prod_{t=1}^4 h_n(X_{i_tn}, u_t) \right].$$

En supposant $j_1 \leq j_2 \leq j_3 \leq j_4$, on a:

a) Si $\max(j_2 - j_1, j_4 - j_3) > r$, on trouve comme ci-dessus la majoration

$$|E[g_{i_1, \dots, i_4}(X_{j_1 n}, \dots, X_{j_4 n})]| \leq 4u_n^4(4 + \delta_0)\beta_n^{\frac{\delta_0}{4+\delta_0}}(r).$$

b) Si $\max(j_2 - j_1, j_4 - j_3) \leq r$ et $j_3 - j_2 > r$, on considère la décomposition

$$\begin{aligned} & |E[g_{i_1, \dots, i_4}(X_{j_1 n}, \dots, X_{j_4 n})]| \\ & \leq |E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, X_{j_3 n}, X_{j_4 n})] - E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, \bar{X}_{j_3 n}, \bar{X}_{j_4 n})]| \\ & \quad + |E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, \bar{X}_{j_3 n}, \bar{X}_{j_4 n})]|. \end{aligned}$$

Pour le premier terme on obtient la même majoration que dans a). Pour le deuxième, qui est égal à $|E[G_{n, j_2 - j_1}(X_{i_2 n}, X_{i_1 n})G_{n, j_4 - j_3}(X_{i_4 n}, X_{i_3 n})]|$, on trouve:

b₁) Si $\min(j_2 - j_1, j_4 - j_3) > 0$,

$$|E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, \bar{X}_{j_3 n}, \bar{X}_{j_4 n})]| \leq z_n^2(2).$$

b₂) Si $0 = j_2 - j_1 < j_4 - j_3$ (resp. $0 = j_4 - j_3 < j_2 - j_1$),

$$|E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, \bar{X}_{j_3 n}, \bar{X}_{j_4 n})]| \leq \begin{cases} z_n(2)w_n(2), & \text{si } i_1 = i_2 \text{ (resp. } i_3 = i_4), \\ z_n(2)v_n(2), & \text{si } i_1 \neq i_2 \text{ (resp. } i_3 \neq i_4). \end{cases}$$

b₃) Si $\max(j_2 - j_1, j_4 - j_3) = 0$,

$$\begin{aligned} & |E[g_{i_1, \dots, i_4}(X_{j_1 n}, X_{j_2 n}, \bar{X}_{j_3 n}, \bar{X}_{j_4 n})]| \\ & \leq \begin{cases} v_n^2(2), & \text{si } i_1 \neq i_2 \text{ et } i_3 \neq i_4, \\ v_n(2)w_n(2), & \text{si } i_1 \neq i_2 \text{ et } i_3 = i_4 \text{ (ou si } i_1 = i_2 \text{ et } i_3 \neq i_4), \\ w_n^2(2), & \text{si } i_1 = i_2 = i_3 = i_4, \\ 4\beta_n^{\frac{\delta_0}{4+\delta_0}}(|i_3 - i_2|)w_n^2(2 + \frac{\delta_0}{2}) \\ \quad + E^2[G_n(X_{0n}, X_{0n})], & \text{si } i_1 = i_2 \neq i_3 = i_4. \end{cases} \end{aligned}$$

Les majorations trouvées dans les points a) et b), les hypothèses i)-v) et le choix de δ_1 et δ_2 , nous permettent de conclure que la somme des termes correspondants est négligeable.

c) Si $\max(j_2 - j_1, j_3 - j_2, j_4 - j_3) \leq r$, on déduit

$$|E[g_{i_1, \dots, i_4}(X_{j_1 n}, \dots, X_{j_4 n})]| \leq u_n^4(4 + \delta_0).$$

D'après l'hypothèse i) et le choix de δ_1 et δ_2 avec $\delta_1, \delta_2 < (1 - 2\gamma_0)/3$, on obtient pour la somme des termes correspondants, la majoration

$$O\left(\frac{m^3 r^3 u_n^4 (4 + \delta_0)}{n^2}\right) = O(n^{3\delta_1 + 3\delta_2 + 4\gamma_0 - 2}) = o(1).$$

Comme l'ordre imposé aux indices j_1, \dots, j_4 ne restreint pas la généralité de l'analyse précédente, la condition D3) est donc satisfaite. ■

Le lemme précédent et le Théorème 3.3.1 nous permettent de déduire le

Théorème 3.4.11 *Sous les conditions du Lemme 3.4.7, la variable aléatoire \mathcal{H}_n est asymptotiquement normale de moyenne zéro et de variance σ_2^2 .*

Le théorème précédent généralise au cas mélangeant le Théorème 1 de Hall [42]. En effet, dans le cadre d'échantillonnage analysé par Hall, on peut prendre $\delta_0 = 0$, et comme, dans ce cas, $z_n(r) = 0$, pour $n \in \mathbb{N}^*$ et $r > 0$, il suffit pour obtenir la normalité asymptotique de la variable \mathcal{H}_n , de vérifier l'existence de $\gamma_0 < 1/2$ tel que $\|h_n(X_{1n}, X_{0n})\|_4 = O(n^{\gamma_0})$, $\|G_{n0}(X_{1n}, X_{0n})\|_2 = o(1)$, $\|G_{n0}(X_{0n}, X_{0n})\|_2 = o(n^{\frac{1}{2}})$ et $E[G_{n0}(X_{0n}, X_{0n})] = 2\sigma_2^2 + o(1)$. Cet ensemble d'hypothèses est en effet moins restrictif que celui utilisé dans Hall [42].

3.4.3 Loi asymptotique de $\mathcal{G}_n + \mathcal{H}_n$

Finalement, en conséquence des résultats précédents, on établit la normalité asymptotique de $\mathcal{G}_n + \mathcal{H}_n$, ou, de façon équivalente, celle du vecteur $(\mathcal{G}_n, \mathcal{H}_n)$.

Théorème 3.4.12 *Si on suppose satisfaites les hypothèses du Lemme 3.4.2 avec $\delta_0 = 2$, et celles du Lemme 3.4.7, alors $\mathcal{G}_n + \mathcal{H}_n$ est asymptotiquement normale de moyenne zéro et de variance donnée par $\sigma_1^2 + \sigma_2^2$.*

Démonstration: D'après les Lemmes 3.4.2 et 3.4.7, nous savons que

$$\mathcal{G}_n + \mathcal{H}_n = \sum_{\alpha=1}^k (T_{\alpha,n} + U_{\alpha,n}) + o_p(1), \quad (3.4.13)$$

où pour $\alpha = 1, \dots, k$, les variables aléatoires $T_{\alpha,n}$ et $U_{\alpha,n}$ sont définies respectivement par (3.4.3) et (3.4.8), avec $k = k(n) = \lfloor \frac{n}{m+r} \rfloor$, $m = m(n) = \lfloor n^{\delta_1} \rfloor$ et $r = r(n) = \lfloor n^{\delta_2} \rfloor$, et δ_1 et δ_2 sont

tels que $0 < \delta_2 < \delta_1 < \min(\gamma_1/2, (1 - 2\gamma_0)/3)$, où γ_0 et γ_1 sont donnés dans les hypothèses i) et iv) du Lemme 3.4.7. De plus, les conditions i) et iii) du Théorème 3.3.1 appliqué à la suite doublement indexée de variables aléatoires $T_{\alpha,n} + U_{\alpha,n}$ et de tribus $F_1^{b_\alpha}(n)$ pour $\alpha = 1, \dots, k$, sont satisfaites avec $\delta = 2$, et la condition ii) est satisfaite avec $\sigma^2 = \sigma_1^2 + \sigma_2^2$ si la somme des covariances conditionnelles est négligeable, c'est-à-dire, si

$$\sum_{\alpha=1}^k E[T_{\alpha,n}U_{\alpha,n}|F_{\alpha-1,n}] = o_p(1) \text{ et } \sum_{\alpha=1}^k E[T_{\alpha,n}|F_{\alpha-1,n}]E[U_{\alpha,n}|F_{\alpha-1,n}] = o_p(1),$$

où $F_{\alpha,n}$ est la tribu $F_1^{b_\alpha}(n)$ pour $\alpha = 1, \dots, k$.

D'une part, la première de ces inégalités, découle d'après les hypothèses i) des Lemmes 3.4.2 et 3.4.7 et de l'application du Lemme 3.2.2 avec $\gamma = 1$. En effet,

$$\begin{aligned} & E \left| \sum_{\alpha=1}^k E[T_{\alpha,n}U_{\alpha,n}|F_{\alpha-1,n}] \right| \\ & \leq \frac{1}{n\sqrt{n}} \sum_{\alpha=1}^k \sum_{i,i'=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E |E[h_n(X_{in}, X_{jn})g_n(X_{i'n})|F_{\alpha-1,n}]| \\ & = \frac{1}{n\sqrt{n}} \sum_{\alpha=1}^k \sum_{i,i'=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E |E[h_n(X_{in}, X_{jn})g_n(X_{i'n})|F_{\alpha-1,n}] - E[h_n(X_{in}, \bar{X}_{jn})g_n(X_{i'n})]| \\ & = O \left(mn^{\frac{1}{2}+\gamma_0} \beta_n^{1/2}(r) \right) = o(1). \end{aligned}$$

D'autre part, de façon analogue à ci-dessus

$$\begin{aligned} & E \left| \sum_{\alpha=1}^k E[T_{\alpha,n}|F_{\alpha-1,n}]E[U_{\alpha,n}|F_{\alpha-1,n}] \right| \\ & \leq \frac{1}{n\sqrt{n}} \sum_{\alpha=1}^k \sum_{i,i'=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-r} E |E[g_n(X_{i'n})E[h_n(X_{in}, X_{jn})|F_{\alpha-1,n}] | F_{\alpha-1,n}]| = o(1). \end{aligned}$$

Le Théorème 3.3.1 et (3.4.13) nous permettent maintenant de conclure. ■

Chapitre 4

Tests d'ajustement à une densité fondés sur un estimateur non paramétrique à noyau

Dans ce chapitre nous proposons des tests d'ajustement à une densité donnée, fondés sur des mesures quadratiques d'écart entre cette densité et un estimateur non paramétrique à noyau. Les résultats établis ici, en utilisant les techniques de U-statistiques dégénérées développées dans le troisième chapitre, sont des généralisations, au cas multivarié et dans un cadre d'indépendance asymptotique, des résultats obtenus par Bickel et Rosenblatt [4].

4.1 Les écarts quadratiques I_n^1 et I_n^2

Soit $(X_i, i \in \mathbb{Z})$ un processus stochastique de dimension d fortement stationnaire, de loi marginale absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d . On note f la densité marginale du processus. Dans ce chapitre on considère le problème de test de l'hypothèse que X_1, \dots, X_n , provient d'un processus stationnaire à densité marginale f_0 , c'est-à-dire, on s'intéresse au problème de test de l'hypothèse

$$H_0 : f = f_0,$$

au niveau asymptotique α , où f_0 est une densité supposée connue. Suivant la ligne de Bickel et Rosenblatt [4], les tests proposés sont fondés sur l'estimateur par noyau f_n de la densité f commune aux variables aléatoires observées X_1, \dots, X_n , défini par (cf. Rosenblatt [83], Parzen

[75] et Cacoullos [15])

$$f_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad (4.1.1)$$

où (h_n) est une suite de nombres réels strictement positifs convergeant vers zéro, lorsque n tend vers l'infini, et K est un noyau sur \mathbb{R}^d (cf. §1.5.1).

Il est dans un premier temps naturel de construire des statistiques de test à partir de l'écart entre l'estimateur par noyau f_n et sa moyenne sous l'hypothèse nulle, l'estimateur étant asymptotiquement sans biais, ou par l'écart entre f_n et la vraie densité f_0 sous l'hypothèse nulle. De tels écarts peuvent être mesurés par

$$I_n^1 = \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\}^2 \pi(x) dx,$$

ou par

$$I_n^2 = \int_{\mathbb{R}^d} \{f_n(x) - f_0(x)\}^2 \pi(x) dx,$$

où π est une fonction de poids et $E_0 f_n(x)$ désigne l'espérance de $f_n(x)$ sous H_0 .

La construction de tests de niveau asymptotique α et asymptotiquement convergents pour tester H_0 contre H_0^c , où on désigne par H_0^c l'hypothèse alternative $f \neq f_0$, est basée sur les variables aléatoires

$$nh_n^{d/2} \{I_n^1 - E_0 I_n^1\} \text{ et } d(n) \{I_n^2 - E_0 I_n^2\},$$

dont la normalité asymptotique est établie en employant des techniques de U-statistiques décrites au Chapitre 3, et où $(d(n))$ est une suite de nombres réels positifs tendant vers l'infini, dépendant de la fenêtre h_n , de l'ordre du noyau, et de l'ordre de dérivabilité de f_0 . Puisque dans le cas des variables dépendantes, les termes déterministes $E_0 I_n^1$ et $E_0 I_n^2$ ne peuvent pas être évalués avec la seule connaissance de f_0 , nous introduisons dans la section suivante des développements asymptotiques de ces quantités qui permettront de corriger les statistiques naturelles I_n^1 et I_n^2 de tels effets de biais asymptotique.

4.2 Développements asymptotiques de l'erreur quadratique moyenne intégrée

Les hypothèses introduites dans les points suivants sont désignées par (P) et (N), respectivement, selon qu'elles concernent le processus, ou le noyau.

Hypothèses sur le processus (P)

Soit $(X_i, i \in \mathbb{Z})$ un processus stochastique de dimension d , fortement stationnaire de loi marginale absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d . On note f sa densité, et on suppose que son coefficient de mélange de type β défini par (3.2.1) décroît vers zéro de façon exponentielle.

On suppose aussi que les lois des vecteurs (X_i, X_0) , $i \geq 1$, admettent des versions des densités $f_{(X_i, X_0)}$ qui satisfont les contraintes

$$\sup_{x \in \mathbb{R}^d} f(x) < \infty \text{ et}$$

$$\sup_{i \in \mathbb{N}^*} \sup_{\substack{x \in \{u \in \mathbb{R}^d | f(u) > 0\} \\ y \in \mathbb{R}^d}} f_i(y|x) < \infty,$$

où $f_i(y|x) = \frac{f_{(X_i, X_0)}(y, x)}{f(x)}$. La première de ces conditions est souvent utilisée dans l'obtention de théorèmes de limite centrale pour l'erreur quadratique intégrée comme on peut le voir dans Bickel et Rosenblatt [4], Hall [42] et Takahata et Yoshihara [99]. La deuxième condition, qui se réduit à la première dans un cadre d'échantillonnage, est dans cette étude avec dépendance une condition importante permettant, par exemple, d'obtenir le développement asymptotique de l'erreur quadratique moyenne intégrée que nous allons présenter ci-dessous. Elle est par exemple satisfaite dans le cas d'un processus gaussien stationnaire, parce que la variance conditionnelle ne dépend pas de la variable conditionante.

Hypothèse sur le noyau (N)

On suppose que K est un noyau borné sur \mathbb{R}^d .

Il est classique de décomposer l'erreur quadratique moyenne dans les termes de variance et de carré de biais. Nous rappelons cette décomposition dans le théorème suivant et présentons un développement asymptotique du terme de variance.

Théorème 4.2.1 *On suppose satisfaites les hypothèses (P) et (N), et soit π une application mesurable et bornée de \mathbb{R}^d dans \mathbb{R} . Alors*

$$E \int_{\mathbb{R}^d} \{f_n(x) - f(x)\}^2 \pi(x) dx = E \int_{\mathbb{R}^d} \{f_n(x) - E f_n(x)\}^2 \pi(x) dx + \int_{\mathbb{R}^d} \{E f_n(x) - f(x)\}^2 \pi(x) dx,$$

où, si $h_n \rightarrow 0$, $n \rightarrow +\infty$, le terme de variance est tel que

$$E \int_{\mathbb{R}^d} \{f_n(x) - E f_n(x)\}^2 \pi(x) dx = \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(x) \pi(x + uh_n) dx du + o\left(\frac{1}{nh_n^{d/2}}\right).$$

Démonstration: Pour $x \in \mathbb{R}^d$,

$$f_n(x) - Ef_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^n \left\{ K\left(\frac{x - X_i}{h_n}\right) - EK\left(\frac{x - X_i}{h_n}\right) \right\}.$$

Définissant, pour $u, v \in \mathbb{R}^d$,

$$\begin{aligned} H_n(u, v) & \quad (4.2.2) \\ &= \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x - u}{h_n}\right) - EK\left(\frac{x - X_0}{h_n}\right) \right\} \left\{ K\left(\frac{x - v}{h_n}\right) - EK\left(\frac{x - X_0}{h_n}\right) \right\} \pi(x) dx, \end{aligned}$$

on obtient, en utilisant la stationnarité de $(X_i, i \in \mathbb{Z})$ et la symétrie de la fonction H_n , l'égalité

$$E \int_{\mathbb{R}^d} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx = \frac{1}{nh_n^{d/2}} E[H_n(X_0, X_0)] + \frac{2}{n^2 h_n^{d/2}} \sum_{1 \leq j < i \leq n} E[H_n(X_i, X_j)].$$

Chacun de ces termes est analysé dans les points suivants.

A) Développement asymptotique de $E[H_n(X_0, X_0)]$

On a

$$\begin{aligned} E[H_n(X_0, X_0)] &= \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} \left\{ EK^2\left(\frac{x - X_0}{h_n}\right) - E^2 K\left(\frac{x - X_0}{h_n}\right) \right\} \pi(x) dx \\ &= \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2\left(\frac{x - y}{h_n}\right) f(y) dy \pi(x) dx - \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} E^2 K\left(\frac{x - X_0}{h_n}\right) \pi(x) dx \\ &= \frac{1}{h_n^{d/2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(y) \pi(y + uh_n) dy du + O(h_n^{d/2}), \end{aligned}$$

puisque, en désignant par $K * \bar{K}$ le produit de convolution de K par \bar{K} où pour $u \in \mathbb{R}^d$, $\bar{K}(u) = K(-u)$ désigne le symétrisé du noyau K , on obtient la majoration

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} E^2 K\left(\frac{x - X_0}{h_n}\right) \pi(x) dx \right| \\ &= \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K\left(\frac{x - y}{h_n}\right) f(y) dy \int_{\mathbb{R}^d} K\left(\frac{x - z}{h_n}\right) f(z) dz \pi(x) dx \right| \\ &= h_n^d \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y) f(z) \int_{\mathbb{R}^d} K(v) K\left(v + \frac{y - z}{h_n}\right) dv \pi(y + vh_n) dy dz \right| \\ &\leq h_n^d \sup_{x \in \mathbb{R}^d} |\pi(x)| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y) f(z) |K * \bar{K}|\left(\frac{y - z}{h_n}\right) dy dz \\ &= h_n^{2d} \sup_{x \in \mathbb{R}^d} |\pi(x)| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(z + uh_n) f(z) |K * \bar{K}|(u) du dz \\ &\leq h_n^{2d} \sup_{x \in \mathbb{R}^d} |\pi(x)| \sup_{x \in \mathbb{R}^d} f(x) \int_{\mathbb{R}^d} |K * \bar{K}|(u) du. \end{aligned}$$

Il reste maintenant à analyser le terme de covariance $\frac{2}{n^2 h_n^{d/2}} \sum_{1 \leq j < i \leq n} E[H_n(X_i, X_j)]$, et à montrer que ce terme est d'ordre inférieur à $\frac{1}{n h_n^{d/2}}$. Si jamais les variables $X_i, i \in \mathbb{Z}$, sont indépendantes, ce terme est identiquement nul. L'idée de la démonstration est alors d'examiner l'écart entre ce cas et celui de la dépendance. Pour cela nous utiliserons le Corollaire 3.2.3. Notons \bar{X}_0 une copie de X_0 indépendante du processus $(X_i, i \in \mathbb{Z})$. Etablissons alors la majoration adéquate pour

$$u_n(r) = \max \left\{ \max_{1 \leq i \leq n} \|H_n(X_i, X_0)\|_r, \|H_n(X_0, \bar{X}_0)\|_r \right\}, \text{ avec } r \geq 1.$$

B) Majoration de $u_n(r)$

D'après la définition de $H_n(\cdot, \cdot)$, on conclut que pour $u, v \in \mathbb{R}^d$,

$$\begin{aligned} H_n(u, v) &= \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-u}{h_n}\right) K\left(\frac{x-v}{h_n}\right) \pi(x) dx - \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-u}{h_n}\right) EK\left(\frac{x-X_0}{h_n}\right) \pi(x) dx \\ &\quad - \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-v}{h_n}\right) EK\left(\frac{x-X_0}{h_n}\right) \pi(x) dx + \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} E^2 K\left(\frac{x-X_0}{h_n}\right) \pi(x) dx, \end{aligned}$$

où chacun de ces derniers termes est majoré par (voir le point A))

$$h_n^{d/2} \sup_{x \in \mathbb{R}^d} |\pi(x)| \sup_{x \in \mathbb{R}^d} f(x) \int_{\mathbb{R}^d} |K| * |\bar{K}|(u) du.$$

Uniformément par rapport à $u, v \in \mathbb{R}^d$, on peut donc écrire

$$H_n(u, v) = \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-u}{h_n}\right) K\left(\frac{x-v}{h_n}\right) \pi(x) dx + O(h_n^{d/2}). \quad (4.2.3)$$

Il est donc évident que pour $r \geq 1$ et $i \geq 1$, on a

$$h_n^{3d/2} \|H_n(X_i, X_0)\|_r \leq E^{\frac{1}{r}} \left| \int_{\mathbb{R}^d} K\left(\frac{x-X_i}{h_n}\right) K\left(\frac{x-X_0}{h_n}\right) \pi(x) dx \right|^r + O(h_n^{2d}),$$

où

$$\begin{aligned} &E \left| \int_{\mathbb{R}^d} K\left(\frac{x-X_i}{h_n}\right) K\left(\frac{x-X_0}{h_n}\right) \pi(x) dx \right|^r \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} K\left(\frac{x-v}{h_n}\right) K\left(\frac{x-u}{h_n}\right) \pi(x) dx \right|^r f_i(v|u) f(u) dv du \\ &\leq (h_n^d)^r \sup_{x \in \mathbb{R}^d} |\pi(x)|^r \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|K| * |\bar{K}|)^r \left(\frac{v-u}{h_n}\right) f_i(v|u) f(u) dv du \\ &= (h_n^d)^{r+1} \sup_{x \in \mathbb{R}^d} |\pi(x)|^r \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (|K| * |\bar{K}|)^r(z) f_i(u+h_n z|u) f(u) dz du \\ &\leq (h_n^d)^{r+1} \sup_{x \in \mathbb{R}^d} |\pi(x)|^r \sup_i \sup_{x, y} f_i(y|x) \int_{\mathbb{R}^d} (|K| * |\bar{K}|)^r(z) dz. \end{aligned}$$

Nous en déduisons que $h_n^{3d/2} \|H_n(X_i, X_0)\|_r \leq C(h_n^d)^{1+\frac{1}{r}}$, avec $C > 0$. Une majoration du même ordre peut être obtenue pour $h_n^{3d/2} \|H_n(X_0, \bar{X}_0)\|_r$, si on remplace dans les développements précédents $f_i(v|u)$ par $f(v)$. Ceci permet de conclure que pour $r \geq 1$, il existe $C > 0$ tel que, pour n assez grand,

$$u_n(r) \leq C(h_n^d)^{\frac{1}{r}-\frac{1}{2}}. \quad (4.2.4)$$

C) Ordre du terme de covariance

En utilisant le Corollaire 3.2.3 et l'inégalité (4.2.4), nous concluons que pour $i \geq 1$ et $r > 1$

$$\begin{aligned} |E[H_n(X_i, X_0)]| &= |E[H_n(X_i, X_0)] - E[H_n(X_0, \bar{X}_0)]| \\ &\leq 4\beta^{\frac{r-1}{r}}(i)u_n(r) \\ &\leq 4C\beta^{\frac{r-1}{r}}(i)(h_n^d)^{\frac{1}{r}-\frac{1}{2}}, \end{aligned}$$

et donc

$$\begin{aligned} \left| nh_n^{d/2} \frac{1}{n^2 h_n^{d/2}} \sum_{1 \leq j < i \leq n} E[H_n(X_i, X_j)] \right| &\leq \frac{1}{n} \sum_{i=1}^{n-1} (n-i) |E[H_n(X_i, X_0)]| \\ &\leq \sum_{i=1}^{n-1} |E[H_n(X_i, X_0)]| \\ &\leq 4C(h_n^d)^{\frac{1}{r}-\frac{1}{2}} \sum_{i=1}^{n-1} \beta^{\frac{r-1}{r}}(i). \end{aligned}$$

Puisque d'après l'hypothèse sur les coefficients de mélange, la somme $\sum_{i=1}^{n-1} \beta^{\frac{r-1}{r}}(i)$ est convergente, nous en déduisons que le majorant est négligeable dès que $r < 2$. ■

Remarquons que le résultat précédent peut s'obtenir avec des conditions plus faibles sur les coefficients de mélange. En effet, il suffit d'admettre la convergence de la série $\sum_{i=1}^{+\infty} \beta^r(i)$ pour un $r \in]0, \frac{1}{2}[$. Sous des conditions plus générales, un résultat analogue, mais d'un ordre moins précis a été obtenu par Meloche [65].

Remarquons que, si les variables aléatoires $(X_i, i \in \mathbb{Z})$ sont indépendantes, la parcelle résiduelle du terme de variance ne comporte pas d'effet de corrélation temporelle, en étant donc d'ordre $1/n$.

Sous certaines conditions supplémentaires, on peut trouver des développements du terme $\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(x) \pi(x + uh_n) dx du$ qui permettent d'éviter le calcul explicite de cette intégrale

double. Notons d'abord que pour $\pi = 1$, nous trouvons simplement

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(x) \pi(x + uh_n) dx du = \int_{\mathbb{R}^d} K^2(u) du.$$

Dans le cas plus général où π est éventuellement différente de 1, nous pouvons encore écrire

$$\begin{aligned} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(x) \pi(x + uh_n) dx du &= E \int_{\mathbb{R}^d} K^2(u) \pi(X_0 + uh_n) du \\ &= \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du + O_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

ce qui permet d'énoncer le corollaire suivant,

Corollaire 4.2.5 *Sous les conditions du Théorème 4.2.1, si de plus $nh_n^d \rightarrow +\infty$, $n \rightarrow +\infty$, alors*

$$E \int_{\mathbb{R}^d} \{f_n(x) - E f_n(x)\}^2 \pi(x) dx = \frac{1}{nh_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du + o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

Supposons maintenant que f admet des dérivées partielles jusqu'à l'ordre m continues, bornées et intégrables, et que le noyau K satisfait $\int_{\mathbb{R}^d} \|u\|^m K^2(u) du < \infty$. En admettant que π est presque partout continue (c'est-à-dire, l'ensemble de ses points de discontinuité est de mesure de Lebesgue nulle), un développement de Taylor et l'application du théorème de la convergence dominée nous conduisent à

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f(x) \pi(x + uh_n) dx du = \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f(x) \pi(x) dx + M(f, h_n, m) + o(h_n^m),$$

où on note $M(f, h_n, 0) = 0$ et pour $m \geq 1$,

$$M(f, h_n, m) = \sum_{j=1}^m \frac{(-1)^j}{j!} h_n^j \sum_{i_1, \dots, i_j=1}^d \int_{\mathbb{R}^d} K^2(u) u_{i_1} \dots u_{i_j} du \int_{\mathbb{R}^d} \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x) \pi(x) dx, \quad (4.2.6)$$

avec $u = (u_1, u_2, \dots, u_d)$.

Ainsi,

$$\begin{aligned} E \int_{\mathbb{R}^d} \{f_n(x) - E f_n(x)\}^2 \pi(x) dx \\ = \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f(x) \pi(x) dx + \frac{1}{nh_n^d} M(f, h_n, m) + o\left(\frac{h_n^m}{nh_n^d}\right). \end{aligned}$$

Corollaire 4.2.7 *Sous les conditions du Théorème 4.2.1, si la densité f admet des dérivées partielles jusqu'à l'ordre m continues, bornées et intégrables, si le noyau K satisfait $\int_{\mathbb{R}^d} \|u\|^m K^2(u) du < \infty$, et si π est presque partout continue, on a, pour $m \geq \frac{d}{2}$*

$$\begin{aligned} E \int_{\mathbb{R}^d} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx \\ = \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f(x) \pi(x) dx + \frac{1}{nh_n^d} M(f, h_n, m) + o\left(\frac{1}{nh_n^{d/2}}\right), \end{aligned}$$

et pour $m > \frac{d}{2}$

$$\begin{aligned} E \int_{\mathbb{R}^d} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx \\ = \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f(x) \pi(x) dx + \frac{1}{nh_n^d} M(f, h_n, m-1) + o\left(\frac{1}{nh_n^{d/2}}\right). \end{aligned}$$

On retrouve évidemment la simplification correspondante au cas $\pi = 1$. En effet, on déduit que $M(f, h_n, m) = 0$, car $\int_{\mathbb{R}^d} \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x) dx = 0$, pour $j = 1, \dots, m$, et $i_1, \dots, i_j = 1, \dots, d$.

Notons que dans le cas réel $d = 1$, si on prend $m = 1$ ou $m = 2$, on obtient

$$E \int_{\mathbb{R}} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx = \frac{1}{nh_n} \int_{\mathbb{R}} K^2(u) du \int_{\mathbb{R}} f(x) \pi(x) dx + o\left(\frac{1}{nh_n^{1/2}}\right).$$

Dans le cas $d = 2$, nous pouvons également prendre $m = 1$. Nous obtenons alors

$$\begin{aligned} E \int_{\mathbb{R}^2} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx &= \frac{1}{nh_n^2} \int_{\mathbb{R}^2} K^2(u) du \int_{\mathbb{R}^2} f(x) \pi(x) dx \\ &- \frac{1}{nh_n} \left\{ \int_{\mathbb{R}^2} K^2(u) u_1 du \int_{\mathbb{R}^2} \frac{\partial f}{\partial x_1}(x) \pi(x) dx + \int_{\mathbb{R}^2} K^2(u) u_2 du \int_{\mathbb{R}^2} \frac{\partial f}{\partial x_2}(x) \pi(x) dx \right\} + o\left(\frac{1}{nh_n}\right). \end{aligned}$$

Nous pouvons remarquer que si le noyau est symétrique, par exemple s'il s'agit d'un noyau gaussien, les termes du type $\int_{\mathbb{R}^2} K^2(u) u_j du$ sont nuls et l'approximation se réduit à

$$E \int_{\mathbb{R}^2} \{f_n(x) - Ef_n(x)\}^2 \pi(x) dx = \frac{1}{nh_n^2} \int_{\mathbb{R}^2} K^2(u) du \int_{\mathbb{R}^2} f(x) \pi(x) dx + o\left(\frac{1}{nh_n}\right).$$

4.3 Quelques tests convergents

En plus des hypothèses (P) et (N), nous considérons dans la suite des hypothèses sur la suite h_n et sur la fonction de poids, désignées par (S) et (π), respectivement.

Hypothèses sur la suite h_n (S)

On suppose que

$$h_n \rightarrow 0 \text{ et } nh_n^d \rightarrow +\infty, \text{ lorsque } n \rightarrow +\infty, \quad (4.3.1)$$

et qu'il existe $\gamma \in]0, 1[$ tel que

$$\limsup_n n^\gamma h_n^d < \infty. \quad (4.3.2)$$

Les conditions (4.3.1) sont classiques dans l'étude des estimateurs par noyau. La condition (4.3.2) est peu restrictive et est vérifiée par exemple si $h_n = O(n^{-\delta})$, $0 < \delta < \frac{1}{d}$, dès qu'on choisit $\gamma \in]0, \delta d]$.

Hypothèses sur la fonction de poids π (π)

On suppose que π est une application non négative et presque partout continue de \mathbb{R}^d dans \mathbb{R} , satisfaisant $\sup_{x \in \mathbb{R}^d} \pi(x) < \infty$, et $\pi(x) > 0$ pour presque tout x dans le support de f_0 .

Les résultats suivants présentent la région critique de tests de niveau asymptotique α et asymptotiquement convergents fondés sur les statistiques $T_n^{1,1}$ et $T_n^{2,1}$ définies par

$$T_n^{1,1} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f_0(x) \pi(x + uh_n) dx du \right\} \text{ et}$$

$$T_n^{2,1} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f_0(x) \pi(x + uh_n) dx du - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \right\},$$

où f_n est défini par (4.1.1) et $(d(n))$ est une suite tendant vers l'infini que nous préciserons ci-après.

Les statistiques de test précédentes sont obtenues à partir des écarts I_n^1 et I_n^2 respectivement, par la correction donnée dans le Théorème 4.2.1 de leurs biais asymptotiques. Nous les comparerons sous l'hypothèse nulle et discuterons ensuite des équivalents de ces statistiques déduits de développements de l'erreur quadratique moyenne intégrée.

On désigne par ϕ la fonction de répartition de la loi normale standard, et par $K * \bar{K}$ le produit de convolution de K par \bar{K} , où \bar{K} est le symétrisé du noyau K , défini pour $u \in \mathbb{R}^d$ par $\bar{K}(u) = K(-u)$.

4.3.1 Le test fondé sur $T_n^{1,1}$

Théorème 4.3.3 *On suppose satisfaites les hypothèses (P), (N), (S) et (π) . Soit f_0 bornée sur \mathbb{R}^d . Alors la statistique $T_n^{1,1}$ est asymptotiquement normale sous l'hypothèse nulle, de moyenne zéro et de variance $2\nu_1^2$ avec*

$$\nu_1^2 = \int_{\mathbb{R}^d} f_0^2(x) \pi^2(x) dx \int_{\mathbb{R}^d} (K * \bar{K})^2(z) dz.$$

On en déduit que

$$\left\{ T_n^{1,1} \geq \phi^{-1}(1 - \alpha)(2\nu_1^2)^{1/2} \right\},$$

est la région critique d'un test de niveau asymptotique α . Ce test est asymptotiquement convergent pour tester H_0 contre H_0^c .

Démonstration: Dans les points suivants on établit la normalité asymptotique de $T_n^{1,1}$ sous H_0 et la convergence asymptotique du test proposé.

A) Equivalent asymptotique de $T_n^{1,1}$

Notons $U_n = nh_n^{d/2} \{I_n^1 - E_0 I_n^1\}$. D'après le Théorème 4.2.1 U_n est asymptotiquement équivalente à $T_n^{1,1}$. En étant $H_n(\cdot, \cdot)$ défini par (4.2.2) sous l'hypothèse nulle, considérons la décomposition

$$\begin{aligned} I_n^1 &= \frac{1}{n^2 h_n^{d/2}} \sum_{i=1}^n H_n(X_i, X_i) + \frac{2}{n^2 h_n^{d/2}} \sum_{1 \leq j < i \leq n} H_n(X_i, X_j) \\ &= I_{n1}^1 + I_{n2}^1. \end{aligned}$$

Nous pouvons alors écrire

$$U_n = nh_n^{d/2} \{I_{n1}^1 - E_0 I_{n1}^1\} + nh_n^{d/2} \{I_{n2}^1 - E_0 I_{n2}^1\}. \quad (4.3.4)$$

Nous allons montrer que le premier terme de cette décomposition est négligeable et que le second terme est asymptotiquement normal.

B) Etude de $nh_n^{d/2} \{I_{n1}^1 - E_0 I_{n1}^1\}$

D'après la définition de $H_n(u, u)$, il est facile de conclure qu'il existe $C > 0$ tel que $\sup_{u \in \mathbb{R}^d} |H_n(u, u)| \leq C(h_n^{d/2})^{-1}$. Nous pouvons alors majorer le moment d'ordre deux de

$nh_n^{d/2}\{I_{n1}^1 - E_0I_{n1}^1\}$. En désignant par $\bar{H}_n(\cdot) = H_n(\cdot, \cdot) - E_0H_n(X_0, X_0)$ et en tenant compte de la majoration précédente et du Corollaire 3.2.4, nous avons

$$\begin{aligned} E \left[nh_n^{d/2}\{I_{n1}^1 - E_0I_{n1}^1\} \right]^2 &= \frac{1}{n} E[\bar{H}_n(X_0)]^2 + \frac{2}{n^2} \sum_{i=1}^{n-1} (n-i) E[\bar{H}_n(X_i)\bar{H}_n(X_0)] \\ &\leq \frac{4C^2}{nh_n^d} \left(1 + \sum_{i=1}^{n-1} \beta(i) \right) \\ &= O\left(\frac{1}{nh_n^d}\right) = o(1), \text{ puisque } nh_n^d \rightarrow +\infty, n \rightarrow +\infty. \end{aligned}$$

On conclut alors que

$$nh_n^{d/2}\{I_{n1}^1 - E_0I_{n1}^1\} = o_p(1).$$

C) Etude de $nh_n^{d/2}\{I_{n2}^1 - E_0I_{n2}^1\}$

Il reste maintenant à étudier le comportement asymptotique de la variable aléatoire

$$nh_n^{d/2}\{I_{n2}^1 - E_0I_{n2}^1\} = \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_n(X_i, X_j) - E_0H_n(X_i, X_j)\}.$$

La normalité asymptotique de cette variable découlera de l'application du Théorème 3.4.11. Soient pour $r > 0$ et $n \in \mathbb{N}^*$, $u_n(r)$, $v_n(r)$, $w_n(r)$ et $z_n(r)$ définis par (3.4.6) pour le processus $(X_i, i \in \mathbb{Z})$, avec $h_n = H_n$. Sous les hypothèses (P), (N) et (π) , et pour $r \geq 1$, il existe $C > 0$ tel que, pour n assez grand

$$\begin{aligned} u_n(r) &\leq C(h_n^d)^{\frac{1}{r} - \frac{1}{2}}, \\ v_n(r) &\leq C(h_n^d)^{\frac{1}{r}}, \\ w_n(r) &\leq C, \text{ et} \\ z_n(r) &\leq Ch_n^d. \end{aligned} \tag{4.3.5}$$

La première inégalité est l'inégalité (4.2.4). Les autres inégalités s'obtiennent avec une technique analogue.

Les quatre premières conditions du Théorème 3.4.11 sont alors satisfaites avec $\delta_0 > 0$ fixé, $\gamma_0 = \frac{2+\delta_0}{8+2\delta_0} < \frac{1}{2}$ et $\gamma_1 \in]0, \gamma]$, où γ est tel que $\limsup_n n^\gamma h_n^d < \infty$ (hypothèse (4.3.2)). En effet, d'après les hypothèses (4.3.1), nous vérifions que

$$\begin{aligned} \frac{u_n(4 + \delta_0)}{n^{\gamma_0}} &= O\left(\frac{(h_n^d)^{\frac{1}{4+\delta_0} - \frac{1}{2}}}{n^{\gamma_0}}\right) = O\left(\frac{1}{(nh_n^d)^{\gamma_0}}\right) = o(1), \\ v_n(2 + \frac{\delta_0}{2}) &= O\left((h_n^d)^{\frac{1}{2+\delta_0/2}}\right) = o(1), \end{aligned}$$

$$\begin{aligned} \frac{w_n(2 + \frac{\delta_0}{2})}{n^{1/2}} &= O\left(\frac{1}{n^{1/2}}\right) = o(1), \text{ et} \\ z_n(2)n^{\gamma_1} &= O(n^{\gamma_1}h_n^d) = O(n^\gamma h_n^d) = O(1). \end{aligned}$$

Il nous reste à déterminer la forme de la variance asymptotique. Nous allons montrer que $E_0\left[\frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-X_0}{h_n}\right)K\left(\frac{x-\bar{X}_0}{h_n}\right)\pi(x)dx\right]^2$ tend vers une limite. D'après (4.2.3) celle-ci coïncidera alors avec celle de $E_0[G_{n0}(X_0, X_0)]$, où pour $u, v \in \mathbb{R}^d$, $G_{n0}(u, v) = E_0[H_n(X_0, u)H_n(X_0, v)]$. Puisque toute fonction de carré intégrable est limite dans $L^2(\mathbb{R}^d)$ d'une suite de fonctions continues, il suffit de considérer f_0 continue dans \mathbb{R}^d . D'après le théorème de la convergence dominée, on a

$$\begin{aligned} E_0 &\left[\frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} K\left(\frac{x-X_0}{h_n}\right)K\left(\frac{x-\bar{X}_0}{h_n}\right)\pi(x)dx\right]^2 \\ &= \frac{1}{h_n^{3d}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K\left(\frac{x-u}{h_n}\right)K\left(\frac{x-v}{h_n}\right)\pi(x)dx\right)^2 f_0(u)f_0(v)dudv \\ &= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K(y)K\left(y + \frac{u-v}{h_n}\right)\pi(u+yh_n)dy\right)^2 f_0(u)f_0(v)dudv \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K(y)K(y+z)\pi(v+(y+z)h_n)dy\right)^2 f_0(v+zh_n)f_0(v)dzdv \\ &= \int_{\mathbb{R}^d} f_0^2(x)\pi^2(x)dx \int_{\mathbb{R}^d} (K * \bar{K})^2(z)dz + o(1). \end{aligned}$$

Ainsi,

$$\lim_{n \rightarrow +\infty} E_0[G_{n0}(X_0, X_0)] = \nu_1^2 = \int_{\mathbb{R}^d} f_0^2(x)\pi^2(x)dx \int_{\mathbb{R}^d} (K * \bar{K})^2(z)dz. \quad (4.3.6)$$

D'après le Théorème 3.4.11 et l'égalité (4.3.4), on conclut que U_n et donc $T_n^{1,1}$, convergent, lorsque n tend vers l'infini, vers la loi normale de moyenne zéro et de variance $2\nu_1^2$.

D) Convergence asymptotique du test

On suppose maintenant que f_1 est la densité marginale du processus $(X_i, i \in \mathbb{Z})$. Nous avons

$$\begin{aligned} I_n^1 &= 2 \int_{\mathbb{R}^d} \{f_n(x) - E_1 f_n(x)\} \{E_1 f_n(x) - E_0 f_n(x)\} \pi(x) dx \\ &\quad + \int_{\mathbb{R}^d} \{f_n(x) - E_1 f_n(x)\}^2 \pi(x) dx + \int_{\mathbb{R}^d} \{E_1 f_n(x) - E_0 f_n(x)\}^2 \pi(x) dx, \end{aligned}$$

où, d'après le Théorème 4.2.1, les deux premiers termes convergent en probabilité vers zéro, lorsque n tend vers l'infini. Pour le dernier terme, en utilisant les arguments de densité employés

auparavant, il suffit de considérer f_0 et f_1 continues dans \mathbb{R}^d . D'après le théorème de la convergence dominée, on a

$$\begin{aligned}
& \int_{\mathbb{R}^d} \{E_1 f_n(x) - E_0 f_n(x)\}^2 \pi(x) dx \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \frac{1}{h_n^d} K\left(\frac{x-u}{h_n}\right) (f_1 - f_0)(u) du \right)^2 \pi(x) dx \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \frac{1}{h_n^d} K(v) (f_1 - f_0)(x - v h_n) dv \right)^2 \pi(x) dx \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u) K(v) (f_1 - f_0)(x - u h_n) (f_1 - f_0)(x - v h_n) \pi(x) dx du dv \\
&= \int_{\mathbb{R}^d} \{f_1(x) - f_0(x)\}^2 \pi(x) dx + o(1).
\end{aligned}$$

Si $f_1 \neq f_0$, on obtient,

$$\begin{aligned}
\frac{T_n^{1,1}}{n h_n^{d/2}} &= I_n^1 - \frac{1}{n h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) f_0(x - u h_n) \pi(x) dx du \\
&= \int_{\mathbb{R}^d} \{f_1(x) - f_0(x)\}^2 \pi(x) dx + o_p(1),
\end{aligned}$$

et donc, $T_n^{1,1}$ converge en probabilité vers $+\infty$, si $n \rightarrow +\infty$, puisque d'après l'hypothèse (π) , π est strictement positive dans le support de f_0 . Ceci entraîne la convergence du test. ■

4.3.2 Le test fondé sur $T_n^{2,1}$

Si on compare les statistiques I_n^1 et I_n^2 on note qu'elles ne diffèrent que par le remplacement par f_0 de l'espérance de f_n prise sous l'hypothèse nulle. On analyse dans la suite cet écart. Pour cela on a besoin de quelques hypothèses supplémentaires sur la densité f_0 et sur le noyau K .

On désigne par $\mathcal{W}^d(m)$ ($m \in \mathbb{N}^*$), l'ensemble des densités de probabilité sur \mathbb{R}^d qui admettent des dérivées partielles d'ordre m continues, bornées et intégrables dans \mathbb{R}^d .

Si $f_0 \in \mathcal{W}^d(m)$ et $K \in \mathcal{K}^d(m)$ (cf. §1.3.4) pour un $m \in \mathbb{N}^*$, nous avons vu dans la démonstration du Théorème 1.3.14 que pour $x \in \mathbb{R}^d$,

$$E_0 f_n(x) - f_0(x) = h_n^m \Delta_n^m(f_0, K)(x), \quad (4.3.7)$$

où

$$\begin{aligned}
& \Delta_n^m(f_0, K)(x) \\
&= \frac{(-1)^m}{(m-1)!} \sum_{i_1, \dots, i_m=1}^d \int_{\mathbb{R}^d} u_{i_1} \dots u_{i_m} K(u) \int_0^1 \frac{\partial^m f_0}{\partial x_{i_1} \dots \partial x_{i_m}}(x - h_n u t) (1-t)^{m-1} dt du.
\end{aligned} \quad (4.3.8)$$

Etant donné que les dérivées partielles d'ordre m de f_0 sont continues et bornées, on conclut d'après le théorème de la convergence dominée que, pour $x \in \mathbb{R}^d$,

$$\lim_{n \rightarrow +\infty} \Delta_n^m(f_0, K)(x) = \Delta^m(f_0, K)(x),$$

avec

$$\Delta^m(f_0, K)(x) = \frac{(-1)^m}{m!} \sum_{i_1, \dots, i_m=1}^d \int_{\mathbb{R}^d} u_{i_1} \dots u_{i_m} K(u) du \frac{\partial^m f_0}{\partial x_{i_1} \dots \partial x_{i_m}}(x). \quad (4.3.9)$$

Par simplicité nous notons $\Delta_n^m f_0 = \Delta_n^m(f_0, K)$ et $\Delta^m f_0 = \Delta^m(f_0, K)$.

Théorème 4.3.10 *On suppose satisfaites les hypothèses (P), (N), (S) et (π) . Soient f_0 bornée sur \mathbb{R}^d , et $f_0 \in \mathcal{W}^d(m)$ et $K \in \mathcal{K}^d(m)$, pour un $m \in \mathbb{N}^*$. Notons*

$$\nu_1^2 = \int_{\mathbb{R}^d} f_0^2(x) \pi^2(x) dx \int_{\mathbb{R}^d} (K * \bar{K})^2(z) dz, \quad (4.3.11)$$

$$\nu_2^2 = \text{Var}_0((\Delta^m f_0 \cdot \pi)(X_0)) + 2 \sum_{j=1}^{+\infty} \text{cov}_0((\Delta^m f_0 \cdot \pi)(X_j), (\Delta^m f_0 \cdot \pi)(X_0)), \quad (4.3.12)$$

(Var_0 et cov_0 désignent la variance et la covariance évaluées sous l'hypothèse H_0), et choisissons la statistique $T_n^{2,1}$ avec un coefficient $d(n)$ donné par

$$d(n) = \begin{cases} nh_n^{d/2}, & \text{si } \lambda \in [0, +\infty[\\ \sqrt{n} h_n^{-m}, & \text{si } \lambda = +\infty \end{cases} \quad \text{avec } \lambda = \lim_{n \rightarrow +\infty} nh_n^{d+2m}. \quad (4.3.13)$$

i) Si $nh_n^{d+2m} \rightarrow \lambda \in [0, +\infty[$, $n \rightarrow +\infty$, ce qui implique $nh_n^{d/2} = O(\sqrt{n} h_n^{-m})$, alors la statistique $T_n^{2,1}$ est asymptotiquement normale sous l'hypothèse nulle, de moyenne zéro et de variance $2\nu_1^2 + 4\lambda\nu_2^2$. On en déduit que

$$\left\{ T_n^{2,1} \geq \phi^{-1}(1 - \alpha)(2\nu_1^2 + 4\lambda\nu_2^2)^{1/2} \right\},$$

est la région critique d'un test de niveau asymptotique α . Ce test est asymptotiquement convergent pour tester H_0 contre H_0^c .

ii) Si $nh_n^{d+2m} \rightarrow \lambda = +\infty$, $n \rightarrow +\infty$, ce qui implique $\sqrt{n} h_n^{-m} = o(nh_n^{d/2})$, alors la statistique $T_n^{2,1}$ est asymptotiquement normale sous l'hypothèse nulle, de moyenne zéro et de variance $4\nu_2^2$. On en déduit que

$$\left\{ T_n^{2,1} \geq 2\phi^{-1}(1 - \alpha)|\nu_2| \right\},$$

est la région critique d'un test de niveau asymptotique α . Ce test est asymptotiquement convergent pour tester H_0 contre H_0^c .

Remarques 4.3.14 (1) Dans un cadre d'échantillonnage le terme de variance ν_2^2 qui se réduit à $Var_0((\Delta^m f_0 \cdot \pi)(X_0))$ peut être évaluée sous H_0 . Dans le cas général la seule connaissance de f_0 ne permet pas d'évaluer ν_2^2 . Cette quantité devra, dans ce cas, être estimée de façon convergente. Soit

$$\hat{\nu}_{2,n}^2 = Var_0((\Delta^m f_0 \cdot \pi)(X_0)) + 2 \sum_{j=1}^{m(n)} \left\{ \frac{1}{n-j} \sum_{i=1}^{n-j} (\Delta^m f_0 \cdot \pi)(X_i) (\Delta^m f_0 \cdot \pi)(X_{i+j}) - \left(\frac{1}{n} \sum_{i=1}^n (\Delta^m f_0 \cdot \pi)(X_i) \right)^2 \right\},$$

où $m(n)$ est une suite d'entiers convergeant vers l'infini. Si $\frac{m^3(n)}{n} \rightarrow 0$, lorsque $n \rightarrow +\infty$, on peut montrer la convergence en probabilité de $\hat{\nu}_{2,n}^2$ vers ν_2^2 , lorsque n tend vers l'infini. Le théorème précédent reste alors valable, si on remplace ν_2^2 par cet estimateur, dans la définition des régions critiques.

(2) De plus, remarquons que si les variables aléatoires $X_i, i \in \mathbb{Z}$, sont indépendantes, les résultats précédents peuvent être obtenus en remplaçant l'hypothèse (S) par la seule hypothèse (4.3.1). En effet, l'hypothèse additionnelle (4.3.2) est utilisée pour déduire la condition iv) du Lemme 3.4.7, et cette condition est toujours vérifiée dans un cadre d'échantillonnage.

Démonstration du Théorème 4.3.10: Dans les points suivants on établit la normalité asymptotique de $T_n^{2,1}$ sous H_0 et la convergence asymptotique du test.

A) Equivalence asymptotique de $T_n^{2,1}$

Notons $V_n = d(n)\{I_n^2 - E_0 I_n^2\}$. D'après le Théorème 4.2.1, V_n est asymptotiquement équivalente à $T_n^{2,1}$, car $d(n) = O(nh_n^{d/2})$. On a, alors,

$$\frac{1}{d(n)} V_n = \frac{1}{nh_n^{d/2}} U_n + \frac{1}{\sqrt{n} h_n^{-m}} W_n, \quad (4.3.15)$$

où U_n a été défini dans la démonstration du Théorème 4.3.3 par $U_n = nh_n^{d/2}\{I_n^1 - E_0 I_n^1\}$, et

$$W_n = 2\sqrt{n} h_n^{-m} \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx.$$

B) Normalité asymptotique de U_n

Dans la démonstration du Théorème 4.3.3, nous avons établi que U_n et $T_n^{1,1}$ sont asymptotiquement

tiquement équivalentes et que

$$U_n = \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_n(X_i, X_j) - E_0 H_n(X_i, X_j)\} + o_p(1),$$

où $H_n(\cdot, \cdot)$ est défini par (4.2.2). Il découle aussi de la démonstration de ce théorème que les conditions du Théorème 3.4.11 sont satisfaites permettant ainsi de conclure que U_n tend vers la loi normale de moyenne zéro et de variance $2\nu_1^2$.

C) Normalité asymptotique de W_n

D'après (4.3.7) et pour $\Delta_n^m f_0$ définie par (4.3.8), la variable aléatoire W_n s'écrit

$$\begin{aligned} W_n &= 2\sqrt{n} h_n^{-m} \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx \\ &= 2\sqrt{n} h_n^{-m} \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\} h_n^m (\Delta_n^m f_0 \cdot \pi)(x) dx \\ &= \frac{2}{\sqrt{n}} \sum_{i=1}^n \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x - X_i}{h_n}\right) - E_0 K\left(\frac{x - X_0}{h_n}\right) \right\} (\Delta_n^m f_0 \cdot \pi)(x) dx \\ &= \frac{2}{\sqrt{n}} \sum_{i=1}^n G_n(X_i), \end{aligned}$$

où pour $u \in \mathbb{R}^d$,

$$G_n(u) = \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x - u}{h_n}\right) - E_0 K\left(\frac{x - X_0}{h_n}\right) \right\} (\Delta_n^m f_0 \cdot \pi)(x) dx.$$

Vérifions maintenant que les conditions du Lemme 3.4.2 sont satisfaites avec $\delta_0 = 2$ et $c_j = cov_0((\Delta^m f_0 \cdot \pi)(X_j), (\Delta^m f_0 \cdot \pi)(X_0))$, pour $j = 0, 1, 2, \dots$, et $\Delta^m f_0$ définie par (4.3.9).

Comme π est bornée ainsi que les dérivées partielles d'ordre m de f_0 , il est alors facile de conclure que $\sup_{n \in \mathbb{N}^*} \sup_{u \in \mathbb{R}^d} |G_n(u)| < \infty$, et donc la condition i) du Lemme 3.4.2 est satisfaite. Etablissons maintenant la forme de la variance asymptotique, en démontrant que pour $j = 0, 1, 2, \dots$,

$$\lim_{n \rightarrow +\infty} E_0 [G_n(X_j) G_n(X_0)] = cov_0((\Delta^m f_0 \cdot \pi)(X_j), (\Delta^m f_0 \cdot \pi)(X_0)).$$

D'après la définition de G_n ,

$$\begin{aligned} &E_0 [G_n(X_j) G_n(X_0)] \\ &= E_0 \left[\frac{1}{h_n^{2d}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K\left(\frac{x - X_j}{h_n}\right) K\left(\frac{y - X_0}{h_n}\right) (\Delta_n^m f_0 \cdot \pi)(x) (\Delta_n^m f_0 \cdot \pi)(y) dx dy \right] \end{aligned}$$

$$\begin{aligned}
& -E_0^2 \left[\frac{1}{h_n^d} \int_{\mathbb{R}^d} K\left(\frac{x-X_0}{h_n}\right) (\Delta_n^m f_0 \cdot \pi)(x) dx \right] \\
= & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u)K(v) E_0 [(\Delta_n^m f_0 \cdot \pi)(X_j + uh_n)(\Delta_n^m f_0 \cdot \pi)(X_0 + vh_n)] dudv \\
& - \left(\int_{\mathbb{R}^d} K(u) E_0 [(\Delta_n^m f_0 \cdot \pi)(X_0 + uh_n)] du \right)^2.
\end{aligned}$$

Pour presque tout $u, v \in \mathbb{R}^d$, les termes

$$E_0 [(\Delta_n^m f_0 \cdot \pi)(X_j + uh_n)(\Delta_n^m f_0 \cdot \pi)(X_0 + vh_n)] \text{ et } E_0 [(\Delta_n^m f_0 \cdot \pi)(X_0 + uh_n)],$$

convergent, d'après le théorème de la convergence dominée, respectivement vers

$$E_0 [(\Delta^m f_0 \cdot \pi)(X_j)(\Delta^m f_0 \cdot \pi)(X_0)] \text{ et } E_0 [(\Delta^m f_0 \cdot \pi)(X_0)].$$

Une nouvelle application de ce théorème nous permet de conclure.

On conclut alors, d'après le Théorème 3.4.5, que W_n est asymptotiquement normale de moyenne zéro et de variance $4\nu_2^2$.

D) Normalité asymptotique de V_n

L'égalité (4.3.15) s'écrit encore

$$V_n = \frac{d(n)}{nh_n^{d/2}} U_n + \frac{d(n)}{\sqrt{n}h_n^{-m}} W_n,$$

où d'après les points précédents et le Théorème 3.4.12, on conclut que V_n et donc $T_n^{2,1}$, convergent, lorsque n tend vers l'infini, vers la loi normale de moyenne zéro et de variance donnée par $2\nu_1^2 + 4\lambda\nu_2^2$, si $\lambda \in [0, +\infty[$ et par $4\nu_2^2$, si $\lambda = +\infty$.

E) Convergence asymptotique du test

Supposons maintenant que f_1 soit la densité marginale du processus $(X_i, i \in \mathbb{Z})$. D'après le point D) de la démonstration du Théorème 4.3.3, nous avons

$$\begin{aligned}
I_n^2 & - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \\
= & I_n^1 + 2 \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx \\
= & \int_{\mathbb{R}^d} \{f_1(x) - f_0(x)\}^2 \pi(x) dx + 2 \int_{\mathbb{R}^d} \{f_n(x) - E_0 f_n(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx + o_p(1).
\end{aligned}$$

Comme $f_0 \in \mathcal{W}^d(m)$ et $K \in \mathcal{K}^d(m)$, on a

$$\int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx = h_n^{2m} \int_{\mathbb{R}^d} (\Delta_n^m f_0(x))^2 \pi(x) dx,$$

ce qui permet de conclure que

$$\frac{T_n^{2,1}}{d(n)} = \int_{\mathbb{R}^d} \{f_1(x) - f_0(x)\}^2 \pi(x) dx + o_p(1).$$

Donc, si $f_1 \neq f_0$, $T_n^{2,1}$ converge en probabilité vers $+\infty$, si $n \rightarrow +\infty$, puisque π est strictement positive dans le support de f_0 . Ceci entraîne la convergence du test. ■

4.3.3 Comparaison des statistiques de test $T_n^{1,1}$ et $T_n^{2,1}$

Les résultats concernant les lois asymptotiques reposent sur des développements des statistiques de test sous l'hypothèse nulle. Ces développements sont résumés ci-dessous :

$$T_n^{1,1} = \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_n(X_i, X_j) - E_0 H_n(X_i, X_j)\} + o_p(1), \text{ et}$$

$$T_n^{2,1} = \frac{d(n)}{nh_n^{d/2}} \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_n(X_i, X_j) - E_0 H_n(X_i, X_j)\} + \frac{d(n)}{\sqrt{nh_n^{-m}}} \frac{2}{\sqrt{n}} \sum_{i=1}^n G_n(X_i) + o_p(1),$$

où H_n et G_n sont données pour $u, v \in \mathbb{R}^d$, par

$$H_n(u, v) = \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x-u}{h_n}\right) - E_0 K\left(\frac{x-X_0}{h_n}\right) \right\} \left\{ K\left(\frac{x-v}{h_n}\right) - E_0 K\left(\frac{x-X_0}{h_n}\right) \right\} \pi(x) dx,$$

$$\text{et } G_n(u) = \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x-u}{h_n}\right) - E_0 K\left(\frac{x-X_0}{h_n}\right) \right\} (\Delta_n^m f_0 \cdot \pi)(x) dx,$$

avec $\Delta_n^m f_0$ définie par (4.3.8).

Le couple

$$\left(\frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_n(X_i, X_j) - E_0 H_n(X_i, X_j)\}, \frac{2}{\sqrt{n}} \sum_{i=1}^n G_n(X_i) \right),$$

est asymptotiquement gaussien, ce qui explique les résultats obtenus dans les Théorèmes 4.3.3 et 4.3.10. A partir de ces développements on déduit aussi le résultat suivant.

Théorème 4.3.16 *Sous les conditions du Théorème 4.3.10, si $nh_n^{d+2m} \rightarrow \lambda = 0$, $n \rightarrow +\infty$, les statistiques de test $T_n^{1,1}$ et $T_n^{2,1}$ sont asymptotiquement équivalentes sous l'hypothèse nulle.*

4.3.4 Statistiques modifiées

En tenant en compte les divers développements des termes $E_0 I_n^1$ et $E_0 I_n^2$ obtenus d'après les Corollaires 4.2.5 et 4.2.7, nous discutons dans la suite des modifications asymptotiquement

équivalentes sous l'hypothèse nulle, des statistiques de test $T_n^{1,1}$ et $T_n^{2,1}$. Les résultats des paragraphes précédents restent valables pour ces modifications.

D'après le Corollaire 4.2.5, les statistiques $T_n^{1,1}$ et

$$T_n^{1,2} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du \right\},$$

ainsi que les statistiques $T_n^{2,1}$ et

$$T_n^{2,2} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^d} K^2(u) \pi(X_i + uh_n) du - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \right\},$$

sont asymptotiquement équivalentes sous l'hypothèse nulle.

De façon analogue, sous les conditions du Corollaire 4.2.7, les statistiques $T_n^{1,3}$ et $T_n^{1,4}$ définies par

$$T_n^{1,3} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f_0(x) \pi(x) dx - \frac{1}{nh_n^d} M(f_0, h_n, m) \right\}, \text{ si } m \geq \frac{d}{2}, \text{ et}$$

$$T_n^{1,4} = nh_n^{d/2} \left\{ I_n^1 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f_0(x) \pi(x) dx - \frac{1}{nh_n^d} M(f_0, h_n, m-1) \right\}, \text{ si } m > \frac{d}{2},$$

et les statistiques $T_n^{2,3}$ et $T_n^{2,4}$ définies par

$$T_n^{2,3} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f_0(x) \pi(x) dx - \frac{1}{nh_n^d} M(f_0, h_n, m) - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f(x)\}^2 \pi(x) dx \right\}, \text{ si } m \geq \frac{d}{2}, \text{ et}$$

$$T_n^{2,4} = d(n) \left\{ I_n^2 - \frac{1}{nh_n^d} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} f_0(x) \pi(x) dx - \frac{1}{nh_n^d} M(f_0, h_n, m-1) - \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx \right\}, \text{ si } m > \frac{d}{2},$$

sont, sous l'hypothèse nulle, asymptotiquement équivalentes à $T_n^{1,1}$ et à $T_n^{2,1}$, respectivement.

Dans le cas réel $d = 1$, si on prend $m = 1$, $T_n^{1,4}$ est précisément une des statistiques considérées dans Bickel et Rosenblatt [4].

Finalement, remarquons que si la suite (h_n) est choisie de façon que $nh_n^{d/2+2m} \rightarrow 0$, ou que $nh_n^{2d+2m} \rightarrow +\infty$ lorsque $n \rightarrow +\infty$, nous pouvons, à partir de la statistique $T_n^{2,1}$ et de chacune des statistiques précédentes qui lui sont asymptotiquement équivalentes, obtenir d'autres statistiques encore asymptotiquement équivalentes à $T_n^{2,1}$. En effet, pour h_n satisfaisant les conditions précédentes, nous avons, respectivement, $\int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx = o\left(\frac{1}{nh_n^{d/2}}\right)$

(cf. (4.3.7)) et $\frac{1}{nh_n^d} = o\left(\frac{1}{\sqrt{nh_n^{-m}}}\right)$. Si par exemple dans le cas réel $d = 1$ on prend $m = 2$, la statistique

$$nh_n^{1/2} \left\{ I_n^2 - \frac{1}{nh_n} \int_{\mathbb{R}} K^2(u) du \int_{\mathbb{R}} f_0(x) \pi(x) dx \right\},$$

obtenue à partir des statistiques $T_n^{2,3}$ ou $T_n^{2,4}$, est asymptotiquement équivalente à $T_n^{2,1}$ si $h_n = o(n^{-\frac{2}{5}})$. Cette statistique a été aussi considérée dans [4].

Puisque les différences $T_n^{k,1} - T_n^{k,3}$ et $T_n^{k,1} - T_n^{k,4}$, pour $k = 1, 2$, sont déterministes et négligeables, nous nous limitons dans la section suivante à l'étude de la puissance locale des tests asymptotiques définis par les régions critiques

$$C_n^{1,k} = \left\{ T_n^{1,k} \geq \phi^{-1}(1 - \alpha)(2\nu_1^2)^{1/2} \right\}, \quad k = 1, 2,$$

$$C_n^{2,k} = \left\{ T_n^{2,k} \geq \phi^{-1}(1 - \alpha)(2\nu_1^2 + 4\lambda\nu_2^2)^{1/2} \right\}, \quad k = 1, 2, \text{ et } \lambda \in [0, +\infty[, \text{ et}$$

$$C_n^{2,k} = \left\{ T_n^{2,k} \geq 2\phi^{-1}(1 - \alpha)|\nu_2| \right\}, \quad k = 1, 2, \text{ et } \lambda = +\infty,$$

où ν_1^2 , ν_2^2 et λ , sont définis respectivement par (4.3.11), (4.3.12) et (4.3.13).

4.4 Puissance locale des tests

De façon à décrire les propriétés de puissance des tests jusqu'à maintenant introduits, nous nous intéressons au comportement asymptotique des écarts I_n^1 et I_n^2 , sous une suite, notée $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$, de processus stochastiques fortement stationnaires dont la suite de densités marginales $(g_n)_{n \in \mathbb{N}^*}$ (suite d'alternatives locales pour f_0), est de la forme

$$g_n(x) = f_0(x) + a_n \gamma(x) + o(a_n) \gamma_n(x), \quad (4.4.1)$$

pour $x \in \mathbb{R}^d$, où

$$\sup_{x \in \mathbb{R}^d} |\gamma(x)| < \infty, \quad \sup_{n \in \mathbb{N}^*} \sup_{x \in \mathbb{R}^d} |\gamma_n(x)| < \infty, \quad (4.4.2)$$

$$\int_{\mathbb{R}^d} |\gamma(x)| dx < \infty, \quad \sup_{n \in \mathbb{N}^*} \int_{\mathbb{R}^d} |\gamma_n(x)| dx < \infty, \quad (4.4.3)$$

et (a_n) est une suite de nombres réels positifs convergeant vers zéro, lorsque n tend vers l'infini.

Les hypothèses introduites dans le paragraphe suivant sont désignées par (AL).

4.4.1 Hypothèses sur la suite d'alternatives locales (AL)

Admettons que $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$, est une suite de processus stochastiques de dimension d , fortement stationnaires dont les coefficients de mélange de type β , définis par (3.2.1), décroissent vers zéro de façon exponentielle, uniformément par rapport à n . Pour $n \in \mathbb{N}^*$, soit g_n , définie par (4.4.1), la densité marginale du processus $(X_{in}, i \in \mathbb{Z})$, et désignons par $g_{in}(\cdot|x)$, la densité de la loi conditionnelle de $X_{in}|X_{0n} = x$, qu'on suppose satisfaire la contrainte

$$\sup_{n, i \in \mathbb{N}^*} \sup_{\substack{x \in \{u \in \mathbb{R}^d | g_n(u) > 0\} \\ y \in \mathbb{R}^d}} g_{in}(y|x) < \infty.$$

Discussion des hypothèses précédentes

Il est à ce niveau utile de discuter ces hypothèses et notamment de donner leur interprétation en terme de variables aléatoires. Plus précisément, nous allons montrer que les hypothèses (AL) sont satisfaites dans le cas d'une suite de processus définis pour $n \in \mathbb{N}^*$, par

$$X_{in} = X_i + \delta_n Z_i, \quad i \in \mathbb{Z},$$

où $(X_i, i \in \mathbb{Z})$ satisfait l'hypothèse (P) avec $f = f_0$, (δ_n) est une suite de nombres positifs qui tend vers zéro, lorsque n tend vers l'infini, et $(Z_i, i \in \mathbb{Z})$ est un processus stochastique fortement stationnaire de dimension d , indépendant de $(X_i, i \in \mathbb{Z})$, avec un coefficient de mélange décroissant vers zéro de façon exponentielle.

D'après l'indépendance entre les processus $(X_i, i \in \mathbb{Z})$ et $(Z_i, i \in \mathbb{Z})$, on peut déduire que les coefficients de mélange de la suite de processus $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$ satisfont aussi la condition de décroissance exponentielle (cf. Bradley [12] pg. 173-174). De plus pour $i, n \in \mathbb{N}^*$, $y \in \mathbb{R}^d$ et $x \in \{u \in \mathbb{R}^d | g_n(u) > 0\}$, nous avons

$$g_{in}(y|x) \leq \sup_{i \in \mathbb{N}^*} \sup_{x, y} f_i(y|x),$$

puisque, pour $x, y \in \mathbb{R}^d$,

$$g_n(x) = E[f_0(x - \delta_n Z_0)],$$

$$f_{(X_{in}, X_{0n})}(y, x) = E[f_0(x - \delta_n Z_0) f_i(y - \delta_n Z_i | x - \delta_n Z_0)],$$

et donc

$$g_{in}(y|x) = \frac{f_{(X_{in}, X_{0n})}(y, x)}{g_n(x)},$$

apparaît alors comme une moyenne pondérée des $f_i(y - \delta_n Z_i | x - \delta_n Z_0)$, d'où la majoration.

Les formes des alternatives locales diffèrent selon les hypothèses concernant la loi marginale de Z_0 . Il faut distinguer deux cas selon que ces variables sont ou non centrées.

Si $E[Z_0] \neq 0$ et $E|Z_{0i}|^2 < \infty$, pour $i = 1, 2, \dots, d$, où Z_{0i} désigne la $i^{\text{ème}}$ composante de Z_0 , et si f_0 admet des dérivées partielles jusqu'à l'ordre 2 continues, bornées et intégrables dans \mathbb{R}^d , on a pour $x \in \mathbb{R}^d$,

$$\begin{aligned} g_n(x) &= E[f_0(x - \delta_n Z_0)] \\ &= f_0(x) + \delta_n \gamma(x) + \delta_n^2 \gamma_n(x), \end{aligned}$$

avec

$$\begin{aligned} \gamma(x) &= - \sum_{i=1}^d E[Z_{0i}] \frac{\partial f_0}{\partial x_i}(x), \text{ et} \\ \gamma_n(x) &= \sum_{i,j=1}^d E \left[Z_{0i} Z_{0j} \int_0^1 \frac{\partial^2 f_0}{\partial x_i \partial x_j}(x - \delta_n Z_0 t)(1-t) dt \right]. \end{aligned}$$

Si $E[Z_0] = 0$ et $E|Z_{0i}|^3 < \infty$, pour $i = 1, 2, \dots, d$, le développement doit être mené à l'ordre supérieur; si f_0 admet des dérivées partielles jusqu'à l'ordre 3 continues, bornées et intégrables dans \mathbb{R}^d , on a pour $x \in \mathbb{R}^d$,

$$g_n(x) = f_0(x) + \delta_n^2 \gamma(x) + \delta_n^3 \gamma_n(x),$$

avec

$$\begin{aligned} \gamma(x) &= \frac{1}{2} \sum_{i,j=1}^d E[Z_{0i} Z_{0j}] \frac{\partial^2 f_0}{\partial x_i \partial x_j}(x), \text{ et} \\ \gamma_n(x) &= -\frac{1}{2} \sum_{i,j,k=1}^d E \left[Z_{0i} Z_{0j} Z_{0k} \int_0^1 \frac{\partial^3 f_0}{\partial x_i \partial x_j \partial x_k}(x - \delta_n Z_0 t)(1-t)^2 dt \right]. \end{aligned}$$

On conclut que si Z_0 est de moyenne non nulle, $a_n = \delta_n$, et si Z_0 est de moyenne nulle, $a_n = \delta_n^2$, où a_n est donné dans (4.4.1). Ceci permet de relier les ordres des alternatives locales exprimés en terme de densités (a_n) et en terme de variables (δ_n).

4.4.2 Propriétés asymptotiques des écarts I_n^1 et I_n^2 pour une suite d'alternatives locales

Nous considérons les écarts étudiés précédemment, en introduisant un double indice n pour insister sur le fait que la loi du processus varie avec n le long d'une alternative locale. Ces

écarts sont

$$I_{nn}^1 = \int_{\mathbb{R}^d} \{f_{nn}(x) - E_0 f_n(x)\}^2 \pi(x) dx,$$

et

$$I_{nn}^2 = \int_{\mathbb{R}^d} \{f_{nn}(x) - f_0(x)\}^2 \pi(x) dx,$$

où $f_{nn}(x)$, pour $x \in \mathbb{R}^d$ et $n \in \mathbb{N}^*$, est défini par

$$f_{nn}(x) = \frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{x - X_{in}}{h_n}\right).$$

Dans les résultats suivants, nous commençons par mettre en évidence des expressions asymptotiquement équivalentes de ces écarts. Dans le théorème ci-dessous ν_1^2 est défini par la formule (4.3.11).

Théorème 4.4.4 *On suppose satisfaites les hypothèses (AL), (N), (S) et (π) . Soit f_0 bornée sur \mathbb{R}^d . On a*

$$\begin{aligned} I_{nn}^1 &= \frac{1}{nh_n^{d/2}} \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} \\ &\quad + a_n^2 \int_{\mathbb{R}^d} \gamma^2(x) \pi(x) dx + \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) g_n(x) \pi(x + uh_n) dx du \\ &\quad + o_p(a_n^2) + o_p\left(\frac{1}{nh_n^{d/2}}\right), \end{aligned}$$

où pour $u, v \in \mathbb{R}^d$,

$$\begin{aligned} &H_{nn}(u, v) \\ &= \frac{1}{h_n^{3d/2}} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x-u}{h_n}\right) - EK\left(\frac{x-X_{0n}}{h_n}\right) \right\} \left\{ K\left(\frac{x-v}{h_n}\right) - EK\left(\frac{x-X_{0n}}{h_n}\right) \right\} \pi(x) dx. \end{aligned} \tag{4.4.5}$$

De plus, la variable aléatoire

$$\frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\},$$

converge, lorsque n tend vers l'infini, vers la loi normale de moyenne zéro et de variance $2\nu_1^2$.

Démonstration: Considérons la décomposition

$$I_{nn}^1 = \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\}^2 \pi(x) dx - E \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\}^2 \pi(x) dx$$

$$\begin{aligned}
& + E \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\}^2 \pi(x) dx \\
& + 2 \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\} \{Ef_{nn}(x) - E_0f_n(x)\} \pi(x) dx \\
& + \int_{\mathbb{R}^d} \{Ef_{nn}(x) - E_0f_n(x)\}^2 \pi(x) dx \\
= & \frac{1}{nh_n^{d/2}} U_{nn} + E_{nn}^1 + 2E_{nn}^2 + E_{nn}^3.
\end{aligned}$$

Nous analysons dans la suite chacun de ces termes.

A) Etude de la variable U_{nn}

De façon analogue à la démonstration du Théorème 4.3.3, on déduit que la variable aléatoire U_{nn} s'écrit

$$U_{nn} = \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} + o_p(1),$$

où H_{nn} est défini par (4.4.5). De plus, on déduit que les hypothèses du Théorème 3.4.11 sont satisfaites et que la variance asymptotique de U_{nn} est égale à $2\nu_1^2$ (voir la preuve des relations (4.3.5) et (4.3.6)).

B) Etude de la variable E_{nn}^1

De façon analogue au développement établi pour le terme de variance dans la démonstration du Théorème 4.2.1, on a

$$\begin{aligned}
E_{nn}^1 & = E \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\}^2 dx \\
& = \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) g_n(x) \pi(x + uh_n) dx du + o\left(\frac{1}{nh_n^{d/2}}\right).
\end{aligned}$$

C) Etude de la variable E_{nn}^2

On a

$$\begin{aligned}
\frac{\sqrt{n}}{a_n} E_{nn}^2 & = \frac{\sqrt{n}}{a_n} \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\} \{Ef_{nn}(x) - E_0f_n(x)\} \pi(x) dx \\
& = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x - X_{in}}{h_n}\right) - EK\left(\frac{x - X_{0n}}{h_n}\right) \right\} p_n(x) \pi(x) dx,
\end{aligned}$$

en désignant par $p_n(x) = a_n^{-1} \{Ef_{nn}(x) - E_0f_n(x)\} = \frac{1}{h_n^d} \int_{\mathbb{R}^d} K\left(\frac{x-y}{h_n}\right) (\gamma(y) + o(1)\gamma_n(y)) dy$.

D'après les conditions (4.4.2) sur γ et γ_n , on a $\sup_{n \in \mathbb{N}^*} \sup_{x \in \mathbb{R}^d} |p_n(x)| < \infty$, et donc

$$\sup_{u \in \mathbb{R}^d} \left| \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x-u}{h_n}\right) - EK\left(\frac{x - X_{0n}}{h_n}\right) \right\} p_n(x) \pi(x) dx \right| = O(1),$$

ce qui permet, de la même façon que dans le point B) de la démonstration du Théorème 4.3.3, d'obtenir $E\left[\frac{\sqrt{n}}{a_n}E_{nn}^2\right]^2 = O(1)$. Ainsi,

$$E_{nn}^2 = O_p\left(\frac{a_n}{\sqrt{n}}\right).$$

D) Etude de la variable E_{nn}^3

En supposant que γ est continue dans \mathbb{R}^d , on a d'après les conditions (4.4.1), (4.4.2) et (4.4.3), et le théorème de la convergence dominée, que

$$\begin{aligned} E_{nn}^3 &= \int_{\mathbb{R}^d} \{Ef_{nn}(x) - E_0f_n(x)\}^2 \pi(x) dx \\ &= \int_{\mathbb{R}^d} \left\{ \frac{1}{h_n^d} \int_{\mathbb{R}^d} K\left(\frac{x-y}{h_n}\right) (g_n(y) - f_0(y)) dy \right\}^2 \pi(x) dx \\ &= a_n^2 \int_{\mathbb{R}^d} \left\{ \frac{1}{h_n^d} \int_{\mathbb{R}^d} K\left(\frac{x-y}{h_n}\right) \gamma(y) dy \right\}^2 \pi(x) dx + o(a_n^2) \\ &= a_n^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u)K(v)\gamma(y)\gamma(y+(u-v)h_n)\pi(y+uh_n) dy dudv + o(a_n^2) \\ &= a_n^2 \int_{\mathbb{R}^d} \gamma^2(x)\pi(x) dx + o(a_n^2). \end{aligned}$$

Des arguments de densité permettent d'étendre le résultat pour γ bornée et intégrable dans \mathbb{R}^d (hypothèses (4.4.2) et (4.4.3)).

D'après les points précédents, on déduit finalement le résultat énoncé. ■

Dans l'analyse de l'écart I_{nn}^2 admettons que $f_0 \in W^d(m)$, pour un $m \in \mathbb{N}^*$. Nous supposons en outre que la suite d'alternatives locales vérifie la condition supplémentaire suivante:

Hypothèse additionnelle (AL1)

On dit que la condition (AL1) est satisfaite, si pour $i \in \mathbb{N}^*$, il existe $e_i \in \mathbb{R}$ tel que, pour deux suites quelconques u_n et v_n sur \mathbb{R}^d , qui convergent vers zéro, lorsque n tend vers l'infini, on ait

$$E\left[(\Delta_n^m f_0 \cdot \pi)(X_{in} + u_n)(\Delta_n^m f_0 \cdot \pi)(X_{0n} + v_n)\right] \rightarrow e_i, \text{ si } n \rightarrow +\infty, \quad (4.4.6)$$

où $\Delta_n^m f_0$ est définie par (4.3.8).

Théorème 4.4.7 *On suppose satisfaites les hypothèses (AL), (N), (S) et (π) . Soit f_0 bornée sur \mathbb{R}^d , et on suppose $f_0 \in \mathcal{W}^d(m)$ et $K \in \mathcal{K}^d(m)$, pour un $m \in \mathbb{N}^*$. On a*

$$\begin{aligned} I_{nn}^2 &= \frac{1}{nh_n^{d/2}} \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} + \frac{1}{\sqrt{nh_n^{-m}}} \frac{2}{\sqrt{n}} \sum_{i=1}^n G_{nn}(X_{in}) \\ &+ a_n^2 \int_{\mathbb{R}^d} \gamma^2(x) \pi(x) dx + \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) g_n(x) \pi(x + uh_n) dx du \\ &+ 2h_n^m \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u) (g_n(x) - f_0(x)) (\Delta_n^m f_0 \cdot \pi)(x + uh_n) dx du \\ &+ \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx + o_p(a_n^2) + o_p\left(\frac{1}{nh_n^{d/2}}\right), \end{aligned}$$

où $\Delta_n^m f_0$ est défini par (4.3.8), H_{nn} est défini par (4.4.5) et pour $u \in \mathbb{R}^d$,

$$G_{nn}(u) = \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left\{ K\left(\frac{x-u}{h_n}\right) - EK\left(\frac{x-X_{0n}}{h_n}\right) \right\} (\Delta_n^m f_0 \cdot \pi)(x) dx. \quad (4.4.8)$$

De plus, sous l'hypothèse supplémentaire (AL1), le vecteur aléatoire

$$\left(\frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\}, \frac{2}{\sqrt{n}} \sum_{i=1}^n G_{nn}(X_{in}) \right),$$

converge, lorsque n tend vers l'infini, vers la loi normale de moyenne zéro et de matrice de covariance $\begin{pmatrix} 2\nu_1^2 & 0 \\ 0 & 4\nu_2^{*2} \end{pmatrix}$, où ν_1^2 est défini par la formule (4.3.11) et

$$\nu_2^{*2} = \text{Var}_0((\Delta^m f_0 \cdot \pi)(X_0)) + 2 \sum_{j=1}^{+\infty} (e_j - E_0^2[(\Delta^m f_0 \cdot \pi)(X_0)]). \quad (4.4.9)$$

Remarques 4.4.10 (1) D'après les conditions de mélangeance imposées à la suite de processus $(X_{in}, i \in \mathbb{Z})_{n \in \mathbb{N}^*}$, on conclut que la suite $(e_i - E_0^2[(\Delta^m f_0 \cdot \pi)(X_0)])_{i \in \mathbb{N}^*}$ est absolument convergente.

(2) Si pour $i \in \mathbb{N}^*$, la suite de densités $f_{(X_{in}, X_{0n})}$, convergeant vers une densité $f_{(X_{i\infty}, X_{0\infty})}$, lorsque $n \rightarrow +\infty$, est sous les conditions du théorème de la convergence dominée, alors la condition (AL1) est satisfaite avec $e_i = E[(\Delta^m f_0 \cdot \pi)(X_{i\infty})(\Delta^m f_0 \cdot \pi)(X_{0\infty})]$.

(3) Admettons que π est continue et que, pour $x \in \mathbb{R}^d$, les dérivées partielles d'ordre m de f_0 appartiennent à $\mathcal{A}_x(\alpha)$. Si pour chaque $i \in \mathbb{N}^*$, la loi du couple (X_{in}, X_{0n}) converge, lorsque $n \rightarrow +\infty$, vers la loi d'un couple noté $(X_{i\infty}, X_{0\infty})$, alors nous sommes dans les conditions du Théorème 3.1 de Ranga Rao [78] ce qui permet de conclure que la condition (AL1) est satisfaite avec e_i donné dans (2).

(4) Dans un cadre d'échantillonnage et d'après la forme retenue pour la densité marginale de la suite d'alternatives locales, la condition (AL1) est toujours vérifiée avec $e_i = 0$, pour $i \in \mathbb{N}^*$. Dans ce cas $\nu_2^{*2} = \nu_2^2 = \text{Var}_0((\Delta^m f_0 \cdot \pi)(X_0))$. L'égalité $\nu_2^{*2} = \nu_2^2$, est vérifiée dans les cas (2) et (3) précédents si, par exemple, pour tout $i \in \mathbb{N}^*$, la loi du couple $(X_{i\infty}, X_{0\infty})$ coïncide avec celle de (X_i, X_0) sous H_0 .

Démonstration du Théorème 4.4.7: Considérons la décomposition,

$$\begin{aligned} I_{nn}^2 &= I_{nn}^1 + 2 \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx \\ &\quad + 2 \int_{\mathbb{R}^d} \{Ef_{nn}(x) - E_0 f_n(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx \\ &\quad + \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx. \end{aligned}$$

A) Développement asymptotique de I_{nn}^2

D'après le développement trouvé pour I_{nn}^1 dans le Théorème 4.4.4, on peut écrire

$$\begin{aligned} I_{nn}^2 &= \frac{1}{nh_n^{d/2}} U_{nn} + \frac{1}{\sqrt{nh_n^{-m}}} W_{nn} \\ &\quad + a_n^2 \int_{\mathbb{R}^d} \gamma^2(x) \pi(x) dx + \frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) g_n(x) \pi(x + uh_n) dx du \\ &\quad + 2h_n^m \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u) (g_n(x) - f_0(x)) (\Delta_n^m f_0 \cdot \pi)(x + uh_n) dx du \\ &\quad + \int_{\mathbb{R}^d} \{E_0 f_n(x) - f_0(x)\}^2 \pi(x) dx + o_p(a_n^2) + o_p\left(\frac{1}{nh_n^{d/2}}\right), \end{aligned}$$

où $U_{nn} = \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} + o_p(1)$, et

$$\begin{aligned} W_{nn} &= 2\sqrt{n} h_n^{-m} \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\} \{E_0 f_n(x) - f_0(x)\} \pi(x) dx \\ &= 2\sqrt{n} \int_{\mathbb{R}^d} \{f_{nn}(x) - Ef_{nn}(x)\} (\Delta_n^m f_0 \cdot \pi)(x) dx \\ &= 2\sqrt{n} \frac{1}{nh_n^d} \sum_{i=1}^n \int_{\mathbb{R}^d} \left\{ K\left(\frac{x - X_{in}}{h_n}\right) - EK\left(\frac{x - X_{in}}{h_n}\right) \right\} (\Delta_n^m f_0 \cdot \pi)(x) dx \\ &= \frac{2}{\sqrt{n}} \sum_{i=1}^n G_{nn}(X_{in}). \end{aligned}$$

B) Normalité asymptotique du vecteur (U_{nn}, W_{nn})

D'après la démonstration du Théorème 4.4.4, nous savons que la variable U_{nn} satisfait les hypothèses du Théorème 3.4.11, ce qui a permis de conclure sa convergence vers la loi normale

de moyenne zéro et de variance égale à $2\nu_1^2$. Suivant le point C) de la démonstration du Théorème 4.3.10 on conclut que sous l'hypothèse (AL1), la variable W_{nn} satisfait les conditions du Théorème 3.4.5 avec $\delta_0 = 2$, $c_0 = \text{Var}_0((\Delta^m f_0 \cdot \pi)(X_0))$ et $c_j = e_j - E_0^2[(\Delta^m f_0 \cdot \pi)(X_0)]$, pour $j = 1, 2, \dots$, où $\Delta^m f_0$ est définie par (4.3.9). On en déduit que W_{nn} converge en loi vers la loi normale de moyenne zéro et de variance $4\nu_2^{*2}$.

La convergence en loi annoncée pour le couple (U_{nn}, W_{nn}) découle maintenant de l'application du Théorème 3.4.12. ■

4.4.3 Quelques remarques et définitions

Asymptotiquement équivalents sous l'hypothèse nulle, les estimateurs introduits pour les termes de biais $E_0 I_n^1$ et $E_0 I_n^2$, ne le sont déjà plus sous la suite d'alternatives locales. En effet, l'approximation introduite pour ces termes dans le Théorème 4.2.1, n'est pas toujours adéquate pour corriger le terme $\frac{1}{nh_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K^2(u) g_n(x) \pi(x + uh_n) dx du$, intervenant dans les biais des écarts I_{nn}^1 et I_{nn}^2 , lequel ne résulte pas de l'éloignement de l'alternative locale g_n par rapport à f_0 . De ce fait, les tests définis par les régions critiques $C_n^{1,1}$ et $C_n^{1,2}$, ou $C_n^{2,1}$ et $C_n^{2,2}$, présentent différentes propriétés de puissance. En particulier, on trouve dans les statistiques de test $T_n^{1,1}$ et $T_n^{2,1}$, un effet de biais égal à $\frac{a_n}{h_n^{d/2}} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} \gamma(x) \pi(x) dx$, si $h_n^{d/2} = O(a_n)$. Remarquons aussi que le terme $2h_n^m \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(u) (g_n(x) - f_0(x)) (\Delta_n^m f_0 \cdot \pi)(x + uh_n) dx du$ intervenant dans $E I_{nn}^2$, et qui sous l'hypothèse H_0 est égal à zéro ($g_n = f_0$, pour tout n), sous la suite d'alternatives locales considérée, introduit dans les statistiques fondées sur l'écart I_n^2 un effet de biais égal à $2a_n h_n^m d(n) \int_{\mathbb{R}^d} \Delta^m f_0(x) \gamma(x) \pi(x) dx$, si $\frac{1}{\sqrt{n}} = O(a_n)$.

Les définitions que nous introduisons dans la suite, nous permettent une description plus précise des caractéristiques des tests précédents au niveau de la puissance locale. Nous notons C_n la région critique d'un test de niveau asymptotique α et asymptotiquement convergent pour tester H_0 contre H_0^c , et par P_{g_n} la probabilité prise sous la suite d'alternatives locales (4.4.1).

Définition 4.4.11 *On dit que la suite (\underline{a}_n) de nombres réels positifs convergeant vers zéro, lorsque n tend vers l'infini, est une suite inférieure du test asymptotique défini par la région critique C_n , si pour toute suite (g_n) d'alternatives locales de la forme (4.4.1) avec $\gamma \neq 0$, on a*

$$\lim_{n \rightarrow +\infty} P_{g_n}(C_n) = \alpha, \text{ si } \frac{a_n}{\underline{a}_n} = o(1).$$

De plus, si la suite inférieure (\underline{a}_n^I) satisfait $\underline{a}_n = O(\underline{a}_n^I)$, pour toute suite inférieure (\underline{a}_n) , on dit que (\underline{a}_n^I) est une suite limite inférieure du test asymptotique défini par la région critique C_n .

Les suites inférieures caractérisent les vitesses de convergence vers f_0 de la suite d'alternatives locales g_n pour lesquelles la probabilité asymptotique de la région critique est, uniformément par rapport aux alternatives locales considérées, la même que sous l'hypothèse nulle.

Définition 4.4.12 On dit que la suite (\bar{a}_n) de nombres réels positifs convergeant vers zéro, lorsque n tend vers l'infini, est une suite supérieure du test asymptotique défini par la région critique C_n , si pour toute suite d'alternatives locales (g_n) de la forme (4.4.1) avec $\gamma \neq 0$, on a

$$\lim_{n \rightarrow +\infty} P_{g_n}(C_n) = 1, \text{ si } \frac{\bar{a}_n}{a_n} = o(1).$$

De plus, si la suite supérieure (\bar{a}_n^I) satisfait $\bar{a}_n^I = O(\bar{a}_n)$, pour toute suite supérieure (\bar{a}_n) , on dit que (\bar{a}_n^I) est une suite limite supérieure du test asymptotique défini par la région critique C_n .

Les suites supérieures caractérisent les vitesses de convergence vers f_0 de la suite d'alternatives locales g_n , pour lesquelles la probabilité asymptotique de la région critique est, uniformément par rapport aux alternatives locales considérées, la même que sous une hypothèse alternative fixe.

Définition 4.4.13 On dit que la suite (a_n^I) est une suite limite du test asymptotique défini par la région critique C_n , s'il existe des suites limite inférieure (\underline{a}_n^I) et limite supérieure (\bar{a}_n^I) telles que $a_n^I = \underline{a}_n^I = \bar{a}_n^I$.

Nous nous intéressons en particulier aux tests définis ci-dessous (cf. Bickel et Rosenblatt [4] pg. 1082):

Définition 4.4.14 Le test asymptotique défini par la région critique C_n est asymptotiquement convergent et uniformément strictement sans biais (resp. uniformément sans biais) au seuil α contre des alternatives locales (USSB) (resp. (USB)), s'il existe une suite limite (a_n^I) telle que pour toute suite d'alternatives locales (g_n) de la forme (4.4.1), on a

$$\lim_{n \rightarrow +\infty} P_{g_n}(C_n) > \alpha, \text{ (resp. } \geq \alpha) \text{ si } a_n = a_n^I.$$

Nous déterminons dans les paragraphes suivants les suites limites des tests définis par les régions critiques $C_n^{j,k}$, $j, k = 1, 2$. Ces suites sont trouvées en examinant directement la loi limite de la statistique de test sous la suite d'alternatives locales.

4.4.4 Calculs de puissance pour les tests $C_n^{1,k}$, $k = 1, 2$

En désignant par $T_{nn}^{1,1}$ et $T_{nn}^{1,2}$ les statistiques de test sous la suite d'alternatives locales considérée, nous obtenons, sous les conditions du Théorème 4.4.4, les développements

$$\begin{aligned} T_{nn}^{1,2} &= \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} \\ &\quad + nh_n^{d/2} a_n^2 \int_{\mathbb{R}^d} \gamma^2(x) \pi(x) dx + o_p\left(nh_n^{d/2} a_n^2\right) + o_p(1), \text{ et} \\ T_{nn}^{1,1} &= T_{nn}^{1,2} + \frac{a_n}{h_n^{d/2}} \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} \gamma(x) \pi(x) dx + o\left(\frac{a_n}{h_n^{d/2}}\right) + o_p(1), \end{aligned}$$

où H_{nn} est défini par (4.4.5).

En utilisant les développements précédents et en tenant compte du Théorème 4.4.4, nous déterminons maintenant les suites limites des tests définis par les régions critiques $C_n^{1,k}$, $k = 1, 2$, en admettant satisfaite l'hypothèse

Hypothèse supplémentaire sur π

La fonction de poids π satisfait $\pi(x) > 0$, pour presque tout $x \in \mathbb{R}^d$.

Ces suites dépendent du choix de la fenêtre h_n . De façon à simplifier la discussion et illustrer les résultats à l'aide de figures, nous considérons le cas où h_n est de la forme

$$h_n = n^{-a},$$

avec $0 < a < \frac{1}{d}$, d'après les hypothèses (S). Dans ce cas, les suites limite inférieure et supérieure sont aussi de la forme

$$\underline{a}_n^I = n^{-\underline{b}} \text{ et } \bar{a}_n^I = n^{-\bar{b}}.$$

Nous trouvons pour le test de région critique $C_n^{1,2}$

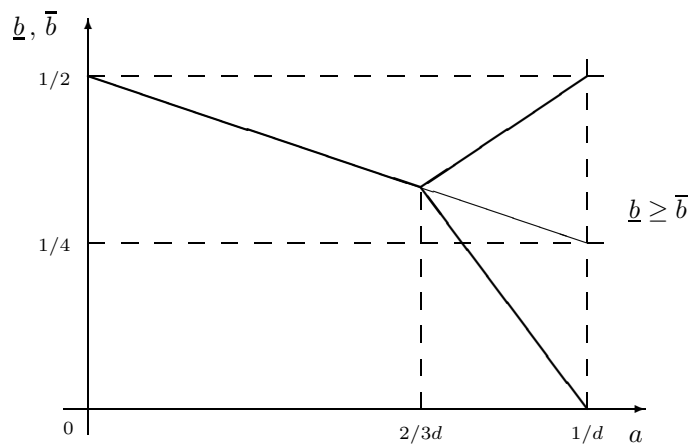
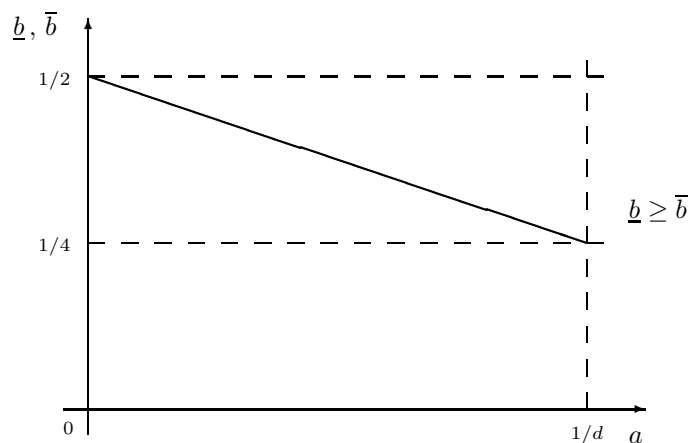
$$\underline{b} = \bar{b} = \frac{1}{2} - \frac{ad}{4},$$

ce qui équivaut à la suite limite

$$a_n^I = \frac{1}{\sqrt{n} h_n^{d/4}},$$

et pour le test de région critique $C_n^{1,1}$,

$$\underline{b} = \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{a \leq \frac{2}{3d}\}} + \frac{ad}{2} \mathbb{I}_{\{a > \frac{2}{3d}\}}, \text{ et}$$

Figure 4.4.15: Ordres des suites limites pour le test de région critique $C_n^{1,1}$ Figure 4.4.16: Ordres des suites limites pour le test de région critique $C_n^{1,2}$

$$\bar{b} = \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{a \leq \frac{2}{3d}\}} + (1 - ad) \mathbb{I}_{\{a > \frac{2}{3d}\}},$$

ce qui équivaut aux suites limites,

$$\underline{a}_n^I = \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{n^{-2/3d} = O(h_n)\}} + h_n^{d/2} \mathbb{I}_{\{h_n = o(n^{-2/3d})\}}, \text{ et}$$

$$\bar{a}_n^I = \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{n^{-2/3d} = O(h_n)\}} + \frac{1}{n h_n^d} \mathbb{I}_{\{h_n = o(n^{-2/3d})\}}.$$

Ces ordres sont résumés dans les Figures 4.4.15 et 4.4.16.

On rassemble dans la suite quelques commentaires.

(1) L'ordre $a_n^I = \frac{1}{\sqrt{n}}$ associé aux tests paramétriques ne peut pas être obtenu. Cependant, il est possible d'être très proche de cet ordre si on choisit une large fenêtre ($a \simeq 0$), ou pour les tests du type $C_n^{1,1}$ une petite fenêtre ($a \simeq \frac{1}{d}$) en considérant seulement des alternatives particulières.

(2) Si on considère $a_n = a_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}$, la puissance asymptotique du test $C_n^{1,2}$ est une fonction croissante de $\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx$. En effet, en tenant compte du fait que $T_{nn}^{1,2}$ est asymptotiquement normale de moyenne $\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx$ et de variance $2\nu_1^2$, on a

$$\lim_{n \rightarrow +\infty} P_{g_n}(W_n^{1,2}) = 1 - \phi\left(\phi^{-1}(1 - \alpha) - \frac{\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx}{(2\nu_1^2)^{1/2}}\right),$$

pour toute suite d'alternatives locales (g_n) de la forme (4.4.1). Le test $C_n^{1,2}$ est donc USSB, d'après la positivité stricte de $\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx$ pour tout γ non nulle.

(3) L'observation précédente reste valable pour le test $C_n^{1,1}$ si la fenêtre h_n est choisie de façon convenable. En effet, lorsque $n^{-2/3d} = o(h_n)$ ($a < \frac{2}{3d}$), on a $T_{nn}^{1,1} = T_{nn}^{1,2} + o_p(1)$. Si on choisit h_n tel que $h_n = O(n^{-2/3d})$ ($a \geq \frac{2}{3d}$), le test $C_n^{1,1}$ n'est plus USSB.

Analysons un peu plus en détaille la puissance asymptotique du test défini par la région critique $C_n^{1,1}$ lorsque h_n satisfait $h_n = O(n^{-2/3d})$.

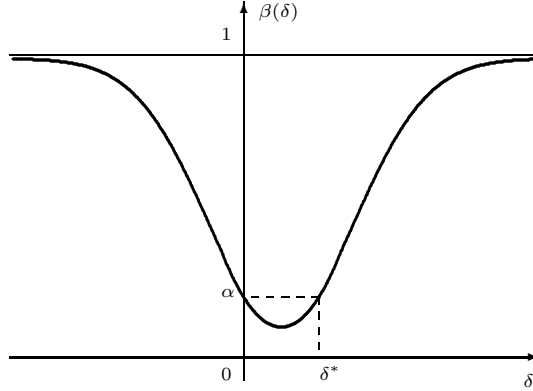
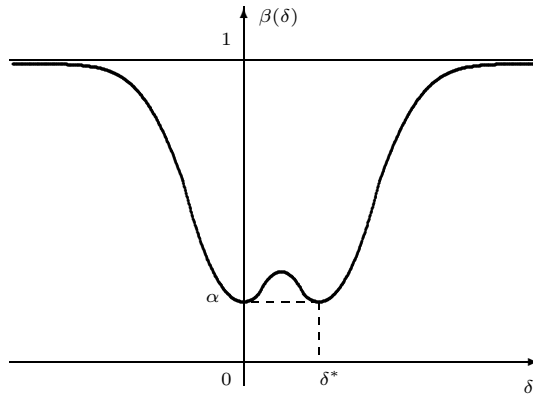
Si $h_n = o(n^{-2/3d})$ ($a > \frac{2}{3d}$), le test $C_n^{1,1}$ n'admet pas une suite limite. Sa puissance asymptotique $\lim_{n \rightarrow +\infty} P_{g_n}(C_n^{1,1})$ peut être de même inférieure à α . C'est par exemple le cas où $a_n = \bar{a}_n^I$ (resp. $a_n = o(\bar{a}_n^I)$ et $\underline{a}_n^I = o(a_n)$), car la statistique de test converge soit vers $+\infty$, soit vers $-\infty$, dépendant, respectivement, du signe positif ou négatif de

$$\int_{\mathbb{R}^d} K^2(u)du \int_{\mathbb{R}^d} \gamma(x)\pi(x)dx + \int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx \text{ (resp. } \int_{\mathbb{R}^d} \gamma(x)\pi(x)dx).$$

Si on considère h_n tel que $nh_n^{3d/2} \rightarrow \mu \in]0, +\infty[$, $n \rightarrow +\infty$ ($a = \frac{2}{3d}$), il existe une suite limite $a_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}$. Cependant, il est dans ce cas préférable de considérer une version bilatérale du test fondé sur $T_n^{1,1}$, si nous voulons obtenir un test asymptotique USB. Pour mieux comprendre le problème, considérons une suite d'alternatives locales (g_n) avec $a_n = a_n^I$, $\gamma(x) = \delta\gamma_0(x)$, pour $x \in \mathbb{R}^d$, où γ_0 est fixé et $\delta \in \mathbb{R}$. Sous cette suite nous avons

$$T_n^{1,1} \xrightarrow[n \rightarrow +\infty]{d} N\left(\delta \frac{1}{\sqrt{\mu}} \int_{\mathbb{R}^d} K^2(u)du \int_{\mathbb{R}^d} \gamma_0(x)\pi(x)dx + \delta^2 \int_{\mathbb{R}^d} \gamma_0^2(x)\pi(x)dx, 2\nu_1^2\right).$$

Considérons la fonction puissance correspondante à ces alternatives locales. Elle est définie pour $\delta \in \mathbb{R}$, par

Figure 4.4.17: Fonction puissance du test unilatérale $C_n^{1,1}$ ($\mu \in]0, +\infty[$ et $a_n = a_n^I$)Figure 4.4.18: Fonction puissance du test bilatérale fondé sur $T_n^{1,1}$ ($\mu \in]0, +\infty[$ et $a_n = a_n^I$)

$$\beta(\delta) = \lim_{n \rightarrow +\infty} P_{g_n}(C_n^{1,1}).$$

Nous avons,

$$\beta(\delta) = 1 - \phi \left(\phi^{-1}(1 - \alpha) - \frac{\delta \int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} \gamma_0(x) \pi(x) dx + \delta^2 \sqrt{\mu} \int_{\mathbb{R}^d} \gamma_0^2(x) \pi(x) dx}{(2\nu_1^2 \mu)^{1/2}} \right).$$

Pour $\delta = \delta^* = \frac{-\int_{\mathbb{R}^d} K^2(u) du \int_{\mathbb{R}^d} \gamma_0(x) \pi(x) dx}{\sqrt{\mu} \int_{\mathbb{R}^d} \gamma_0^2(x) \pi(x) dx}$ positif, cette fonction a la forme donnée dans la Figure 4.4.17. On remarque aussi que l'introduction d'une version bilatérale du test fondé sur $T_n^{1,1}$ n'est pas une bonne solution ne nous permettant pas d'obtenir un test USSB (cf. Figure 4.4.18).

4.4.5 Calculs de puissance pour les tests $C_n^{2,k}$, $k = 1, 2$

En désignant par $T_{nn}^{2,1}$ et $T_{nn}^{2,2}$ les statistiques de test sous la suite d'alternatives locales considérée, nous obtenons, sous les conditions du Théorème 4.4.7, les développements

$$\begin{aligned} T_{nn}^{2,2} &= \frac{d(n)}{nh_n^{d/2}} \frac{2}{n} \sum_{1 \leq j < i \leq n} \{H_{nn}(X_{in}, X_{jn}) - EH_{nn}(X_{in}, X_{jn})\} + \frac{d(n)}{\sqrt{nh_n^{-m}}} \frac{2}{\sqrt{n}} \sum_{i=1}^n G_{nn}(X_{in}) \\ &\quad + d(n)a_n^2 \int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx + 2d(n)h_n^m a_n \int_{\mathbb{R}^d} \Delta^m f_0(x)\gamma(x)\pi(x)dx \\ &\quad + o_p(d(n)a_n^2) + o(d(n)h_n^m a_n) + o_p(1), \text{ et} \end{aligned}$$

$$T_{nn}^{2,1} = T_{nn}^{2,2} + \frac{d(n)a_n}{nh_n^d} \int_{\mathbb{R}^d} \gamma(x)\pi(x)dx \int_{\mathbb{R}^d} K^2(u)du + o\left(\frac{d(n)a_n}{nh_n^d}\right) + o_p(1),$$

où la suite $(d(n))$ est définie par (4.3.13), et H_{nn} et G_{nn} sont définis par (4.4.5) et (4.4.8), respectivement.

Dans le cas où le paramètre ν_2^2 , intervenant dans la définition des régions critiques des tests fondés sur I_n^2 , ne peut pas être évalué sous H_0 , celui-ci doit être remplacé par $\hat{\nu}_{2,n}^2$, l'estimateur introduit dans la Remarque 4.3.14 (1). En étant ν_2^2 la variance asymptotique de la variable $\frac{1}{\sqrt{n}} \sum_{i=1}^n G_n(X_i)$ définie dans §4.3.3, nous allons admettre dans l'étude de la puissance asymptotique des tests $C_n^{2,k}$, $k = 1, 2$, que l'estimateur $\hat{\nu}_{2,n}^2$ est encore sous la suite d'alternatives locales considérée, un estimateur convergent de la variance asymptotique de la variable aléatoire précédente. Nous admettons donc que, sous la suite d'alternatives locales,

$$\hat{\nu}_{2,n}^2 \rightarrow \nu_2^{*2}, \text{ en probabilité, } n \rightarrow +\infty,$$

où ν_2^{*2} est défini par (4.4.9).

La convergence précédente aura lieu si la suite d'alternatives locales satisfait la condition

$$E[(\Delta^m f_0 \cdot \pi)(X_{in})(\Delta^m f_0 \cdot \pi)(X_{0n})] \rightarrow e_i, n \rightarrow +\infty,$$

où e_i est défini par (4.4.6). Cette condition est satisfaite sous les conditions décrites dans les Remarques 4.4.10.

De la même façon que pour les tests définis par les régions critiques $C_n^{1,k}$, $k = 1, 2$, et en admettant que l'hypothèse supplémentaire sur π introduite en §4.4.4 est satisfaite, nous déterminons en fonction de la fenêtre $h_n = n^{-a}$, $0 < a < \frac{1}{d}$, les ordres des suites limite

inférieure $\underline{a}_n^I = n^{-b}$ et supérieure $\bar{a}_n^I = n^{-\bar{b}}$, des tests asymptotiques $C_n^{2,k}$, $k = 1, 2$. Nous le ferons à l'aide des développements précédents et du Théorème 4.4.7.

Nous trouvons pour le test de région critique $C_n^{2,2}$,

$$\underline{b} = \frac{1}{2} \mathbb{I}_{\{a \leq \frac{1}{d+2m}\}} + \left(1 - \frac{a}{2}(d+2m)\right) \mathbb{I}_{\{\frac{1}{d+2m} < a \leq \frac{1}{d/2+2m}\}} + \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{a > \frac{1}{d/2+2m}\}}, \text{ et}$$

$$\bar{b} = am \mathbb{I}_{\{a \leq \frac{1}{d/2+2m}\}} + \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{a > \frac{1}{d/2+2m}\}},$$

ce qui équivaut aux suites limite

$$\underline{a}_n^I = \frac{1}{\sqrt{n}} \mathbb{I}_{\{n^{-\frac{1}{d+2m}} = O(h_n)\}} + \frac{1}{nh_n^{d/2+m}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d+2m}}) \text{ et } n^{-\frac{1}{d/2+2m}} = O(h_n)\}} \\ + \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d/2+2m}})\}}, \text{ et}$$

$$\bar{a}_n^I = h_n^m \mathbb{I}_{\{n^{-\frac{1}{d/2+2m}} = O(h_n)\}} + \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d/2+2m}})\}}.$$

Pour le test de région critique $C_n^{2,1}$, nous avons

i) si $d < 2m$:

$$\underline{b} = \frac{1}{2} \mathbb{I}_{\{a \leq \frac{1}{d+2m}\}} + \left(1 - \frac{a}{2}(d+2m)\right) \mathbb{I}_{\{\frac{1}{d+2m} < a \leq \frac{1}{d/2+2m}\}} \\ + \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{\frac{1}{d/2+2m} < a \leq \frac{2}{3d}\}} + \frac{ad}{2} \mathbb{I}_{\{a > \frac{2}{3d}\}}, \text{ et}$$

$$\bar{b} = am \mathbb{I}_{\{a \leq \frac{1}{d/2+2m}\}} + \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{I}_{\{\frac{1}{d/2+2m} < a \leq \frac{2}{3d}\}} + (1-ad) \mathbb{I}_{\{a > \frac{2}{3d}\}},$$

ce qui équivaut aux suites limite

$$\underline{a}_n^I = \frac{1}{\sqrt{n}} \mathbb{I}_{\{n^{-\frac{1}{d+2m}} = O(h_n)\}} + \frac{1}{nh_n^{d/2+m}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d+2m}}) \text{ et } n^{-\frac{1}{d/2+2m}} = O(h_n)\}} \\ + \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d/2+2m}}) \text{ et } n^{-2/3d} = O(h_n)\}} + h_n^{d/2} \mathbb{I}_{\{h_n = o(n^{-2/3d})\}}, \text{ et}$$

$$\bar{a}_n^I = h_n^m \mathbb{I}_{\{n^{-\frac{1}{d/2+2m}} = O(h_n)\}} + \frac{1}{\sqrt{n} h_n^{d/4}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d/2+2m}}) \text{ et } n^{-2/3d} = O(h_n)\}} + \frac{1}{nh_n^d} \mathbb{I}_{\{h_n = o(n^{-2/3d})\}}.$$

ii) si $d \geq 2m$:

$$\underline{b} = \frac{1}{2} \mathbb{I}_{\{a \leq \frac{1}{d+2m}\}} + \left(1 - \frac{a}{2}(d+2m)\right) \mathbb{I}_{\{\frac{1}{d+2m} < a \leq \frac{1}{d+m}\}} + \frac{ad}{2} \mathbb{I}_{\{a > \frac{1}{d+m}\}}, \text{ et}$$

$$\bar{b} = am \mathbb{I}_{\{a \leq \frac{1}{d+m}\}} + (1-ad) \mathbb{I}_{\{a > \frac{1}{d+m}\}},$$

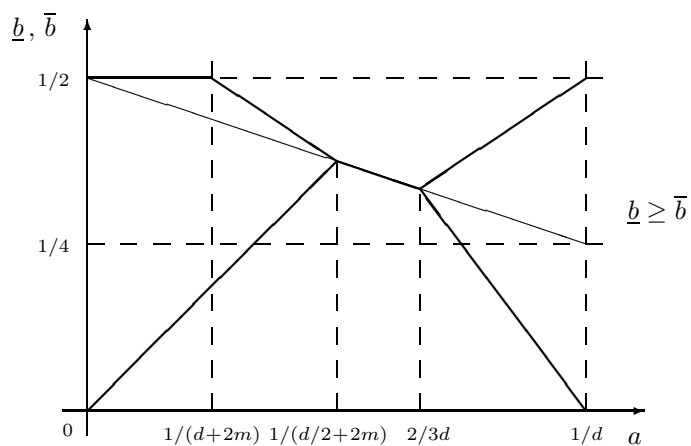


Figure 4.4.19: Ordres des suites limites pour le test de région critique $C_n^{2,1}$ avec $d < 2m$

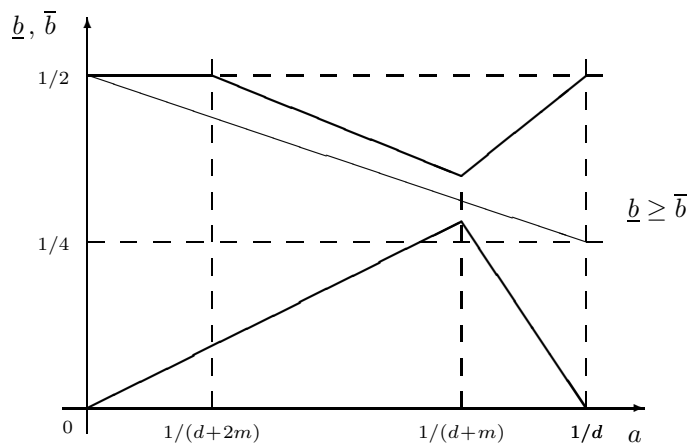


Figure 4.4.20: Ordres des suites limites pour le test de région critique $C_n^{2,1}$ avec $d \geq 2m$

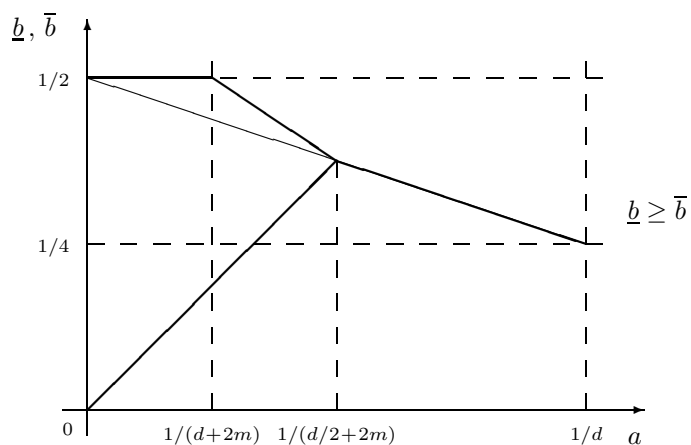


Figure 4.4.21: Ordres des suites limites pour le test de région critique $C_n^{2,2}$

ce qui équivaut aux suites limite

$$\begin{aligned} \underline{a}_n^I &= \frac{1}{\sqrt{n}} \mathbb{I}_{\{n^{-\frac{1}{d+2m}} = O(h_n)\}} + \frac{1}{nh_n^{d/2+m}} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d+2m}}) \text{ et } n^{-\frac{1}{d+m}} = O(h_n)\}} \\ &\quad + h_n^{d/2} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d+m}})\}}, \text{ et} \\ \bar{a}_n^I &= h_n^m \mathbb{I}_{\{n^{-\frac{1}{d+m}} = O(h_n)\}} + \frac{1}{nh_n^d} \mathbb{I}_{\{h_n = o(n^{-\frac{1}{d+m}})\}}. \end{aligned}$$

Ces ordres sont résumés dans les Figures 4.4.19, 4.4.20 et 4.4.21.

Les observations faites pour les tests fondés sur $T_n^{1,k}$, $k = 1, 2$, peuvent d'une façon générale être ici adaptées. On souligne simplement quelques points:

(1) Sous les conditions du Théorème 4.4.7, le test défini par la région critique $C_n^{2,2}$ est U.S.S.B. si on choisit h_n tel que $h_n = o(n^{-\frac{1}{d/2+2m}})$ ($a > \frac{1}{d/2+2m}$). Le même résultat est valable pour le test défini par $C_n^{2,1}$ dans le cas $d < 2m$, dès que la suite (h_n) satisfasse $h_n = o(n^{-\frac{2}{d+m}})$ et $n^{-\frac{2}{3d}} = o(h_n)$ ($\frac{1}{d/2+2m} < a < \frac{2}{3d}$). En effet, les variables $T_{nn}^{2,1}$ et $T_{nn}^{2,2}$ sont, dans les conditions précédentes, asymptotiquement équivalentes à $T_{nn}^{1,2}$. Si $d \geq 2m$, il n'existe pas un choix de la fenêtre h_n qui rend le test $C_n^{2,1}$ USB. En résumé, on conclut qu'uniformément sur les alternatives locales considérées, les tests définis par les régions critiques $C_n^{1,1}$, $C_n^{2,1}$ et $C_n^{2,2}$ n'atteignent la puissance locale du test $C_n^{1,2}$, que pour des choix particuliers de la suite (h_n) .

(2) La remarque la plus importante porte sur la vitesse $\frac{1}{\sqrt{n}}$ obtenue pour la suite limite inférieure des tests fondés sur I_n^2 , dans le cas où $nh_n^{d+2m} \rightarrow \lambda \in]0, +\infty[$, $n \rightarrow +\infty$ ($a \leq \frac{1}{d+2m}$). On conclut facilement que la puissance asymptotique de ces tests sous la suite d'alternatives locales de la forme (4.4.1) avec $a_n = \frac{1}{\sqrt{n}}$ est dans le cas $\lambda = +\infty$ indépendante du choix du noyau K , et supérieure à celle du cas $\lambda \in]0, +\infty[$.

4.5 Sur le choix du noyau

Le développement de procédures de test fondés sur les écarts I_n^1 et I_n^2 , dépend du choix du noyau K . Nous considérons dans la discussion suivante uniquement le cas des tests USSB. La puissance asymptotique de ces tests sous une suite d'alternatives locales avec $a_n = a_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}$, est maximale quand le noyau K est choisi de façon à minimiser la fonctionnelle

$$J(K) = \int_{\mathbb{R}^d} (K * \bar{K})^2(z) dz.$$

Ghosh et Huang [34] déterminent dans le cas réel $d = 1$, le noyau optimal dans ce sens, en se restreignent à la classe $\mathcal{K}^1(\xi, \tau)$, où $\xi \in \mathbb{R}$ et $\tau > 0$, des noyaux positifs K_0 sur \mathbb{R} tels que

$$\int_{\mathbb{R}} tK_0(t)dt = \xi \text{ et } \int_{\mathbb{R}} (t - \xi)^2 K_0(t)dt = \tau^2.$$

Dans l'estimation de la densité, il est usuel de considérer $\xi = 0$ et $\tau = 1$ (cf. Bosq et Lecoutre [9] pg. 82, ou Silverman [97] pg. 40). Ce résultat peut être facilement étendu au cas multivarié $d > 1$. Considérons la classe $\mathcal{K}^d(\xi, \tau)$ des noyaux sur \mathbb{R}^d de la forme

$$K(x) = \prod_{i=1}^d K_0(x_i), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d, \quad (4.5.1)$$

où $K_0 \in \mathcal{K}^1(\xi, \tau)$. Nous avons alors le résultat suivant:

Proposition 4.5.2 *Soient $\xi \in \mathbb{R}$ et $\tau > 0$, fixes. Alors*

$$\min_{K \in \mathcal{K}^d(\xi, \tau)} J(K) = J(K^*) = (3\sqrt{3}\tau)^{-d},$$

où

$$K^*(x_1, \dots, x_d) = \left(\frac{1}{2\sqrt{3}\tau} \right)^d \prod_{i=1}^d \mathbb{I}_{\{|x_i - \xi| \leq \tau\sqrt{3}\}}, \quad (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Démonstration: Conséquence immédiate du Théorème 1.1 de Ghosh et Huang [34] et de l'égalité

$$J(K) = \prod_{i=1}^d \int_{\mathbb{R}} (K_0 * \bar{K}_0)^2(x_i) dx_i,$$

valable pour un noyau $K \in \mathcal{K}^d(\xi, \tau)$ de la forme (4.5.1). ■

4.6 Applications des statistiques $T_n^{2,k}$, $k = 1, 2$, au cas d'un test paramétrique

Quand l'hypothèse alternative est non contrainte, nous avons vu que la vitesse des suites limites des tests USSB fondés sur les statistiques $T_n^{j,k}$, $j, k = 1, 2$, est égale à $\frac{1}{\sqrt{n} h_n^{d/4}}$, inférieure à la vitesse usuellement obtenue dans les tests paramétriques. Cependant, la vitesse $\frac{1}{\sqrt{n}}$ est obtenue pour la suite limite inférieure des tests fondés sur $T_n^{2,k}$, $k = 1, 2$, quand la fenêtre (h_n) est choisie de façon convenable. Bien que ces tests ne soient pas USSB, nous proposons de regarder dans cette section, comment utiliser les statistiques correspondantes pour tester une hypothèse simple paramétrique contre une hypothèse alternative également paramétrique. Nous montrerons qu'à

partir de méthodes de noyau, nous pouvons construire un test USSB au seuil α dont la suite limite correspond à la vitesse paramétrique $\frac{1}{\sqrt{n}}$.

4.6.1 Description des hypothèses

Les hypothèses introduites dans la suite sur le modèle paramétrique et sur la suite d'alternatives locales, seront désignées, comme auparavant, respectivement par (P) et (AL).

Sur le modèle paramétrique (P)

Considérons une famille paramétrique de densités de probabilité (par rapport à la mesure de Lebesgue sur \mathbb{R}^d) bornées et strictement positives $g(\cdot; \theta)$, où $\theta \in \Theta \subset \mathbb{R}^K$, satisfaisant la condition d'identifiabilité $\forall \theta_1, \theta_2 \in \Theta : g(\cdot; \theta_1) = g(\cdot; \theta_2) \Rightarrow \theta_1 = \theta_2$.

Dans cette section on admet que les variables aléatoires X_1, \dots, X_n sont indépendantes et identiquement distribuées de densité $g(\cdot; \theta)$, pour un $\theta \in \Theta$. Pour θ_0 à l'intérieur de Θ , nous voulons tester l'hypothèse

$$H_0 : f = g(\cdot; \theta_0) \text{ contre } H_0^c : f = g(\cdot; \theta), \text{ avec } \theta \in \Theta \setminus \{\theta_0\}.$$

Admettons encore que cette famille de densités satisfait les conditions de dérivation sous le signe intégral, et que la matrice d'information de Fisher définie par

$$I(\theta) = E_\theta \left[\frac{\partial \log g}{\partial \theta}(X_0; \theta) \frac{\partial \log g}{\partial \theta'}(X_0; \theta) \right],$$

est régulière pour $\theta = \theta_0$.

Sur la suite d'alternatives locales (AL)

Les alternatives locales paramétriques sont définies en fonction du paramètre θ par

$$\theta_n = \theta_0 + a_n \delta, \tag{4.6.1}$$

où $\delta = (\delta_1, \dots, \delta_K)' \in \mathbb{R}^K$ et (a_n) est une suite de réels positifs convergeant vers zéro lorsque n tend vers l'infini. En terme de densités, ceci correspond à la suite d'alternatives locales

$$g_n(x) = g(x; \theta_n) = g(x; \theta_0) + a_n \delta' \frac{\partial g}{\partial \theta}(x; \theta_0) + a_n^2 \int_0^1 \delta' \frac{\partial^2 g}{\partial \theta \partial \theta'}(x; \theta_0 + t a_n \delta) \delta (1-t) dt,$$

où on suppose satisfaites les hypothèses permettant d'effectuer le développement limité. Ainsi, avec les notations de la Section 4.4, nous avons une alternative locale convergeant vers $g(x; \theta_0)$

de la forme (4.4.1), avec

$$\gamma(x) = \delta' \frac{\partial g}{\partial \theta}(x; \theta_0), \quad x \in \mathbb{R}^d.$$

En tenant compte de la régularité de la matrice $I(\theta_0)$, l'ensemble des applications γ de la forme précédente, obtenu d'après la variation de δ dans \mathbb{R}^K , est, donc, un espace vectoriel de dimension finie K , engendré par les dérivées partielles $\frac{\partial g}{\partial \theta_1}(\cdot; \theta_0), \dots, \frac{\partial g}{\partial \theta_K}(\cdot; \theta_0)$. Ce fait est à la base des résultats que nous allons obtenir dans cette section.

Afin de construire un test USSB, avec $a_n^I = \frac{1}{\sqrt{n}}$, nous sommes conduits à retenir les statistiques $T_n^{2,k}$, $k = 1, 2$, et à étudier leur comportement sous la suite d'alternatives locales (4.6.1). Supposons que $g(\cdot; \theta_0) \in W^d(m)$ et $K \in K^d(m)$ pour un $m \in \mathbb{N}^*$, et considérons les statistiques $T_n^{2,k}$, $k = 1, 2$, avec une suite (h_n) satisfaisant $nh_n^{d+2m} \rightarrow +\infty, n \rightarrow +\infty$. En tenant compte de leur équivalence asymptotique nous les désignons par T_n^2 . Sous la suite d'alternatives locales (4.6.1), on a d'après le Théorème 4.4.7

$$\begin{aligned} T_n^2 = & Z_n + \sqrt{n}h_n^{-m} \left\{ a_n^2 \delta' \int_{\mathbb{R}^d} \frac{\partial g}{\partial \theta}(x; \theta_0) \frac{\partial g}{\partial \theta'}(x; \theta_0) \pi(x) dx \right. \\ & \left. + 2h_n^m a_n \int_{\mathbb{R}^d} \Delta^m g(x; \theta_0) \frac{\partial g}{\partial \theta'}(x; \theta_0) \pi(x) dx \right\} \delta \\ & + o_p(\sqrt{n}h_n^{-m}a_n^2) + o(\sqrt{n}a_n) + o_p(1), \end{aligned} \quad (4.6.2)$$

où $\Delta^m g(\cdot; \theta_0)$ est défini par (4.3.9) et Z_n est asymptotiquement normale de moyenne zéro et de variance $4\nu_2^2 = 4Var_0(\Delta^m g(X_0; \theta_0)\pi(X_0))$.

4.6.2 Test fondé sur une statistique multivariée à noyau

Considérons pour $p \in \mathbb{N}^*$, fixé, la statistique multivariée

$$T_n(\Pi^p) = (T_n^2(\pi_1), T_n^2(\pi_2), \dots, T_n^2(\pi_p))',$$

où on note $\Pi^p = (\pi_1, \pi_2, \dots, \pi_p)'$ et $T_n^2(\pi_i)$ désigne la statistique T_n^2 où on utilise la fonction de poids π_i .

Introduisons sur le vecteur des fonctions de poids Π^p , l'hypothèse suivante notée (II):

Hypothèse sur le vecteur des fonctions de poids (II)

On suppose que les fonctions composantes du vecteur des fonctions de poids sont presque partout continues et bornées sur \mathbb{R}^d . De plus on admet qu'une de ces fonctions, π_j , satisfait

$\pi_j(x) > 0$ pour presque tout $x \in \mathbb{R}^d$.

Sous H_0 et sous les conditions du Théorème 4.3.10, où l'hypothèse (π) est remplacée par l'hypothèse (Π) , il est facile de conclure que

$$T_n(\Pi^p)/2 \xrightarrow[n \rightarrow +\infty]{d} N(0, A(\Pi^p)),$$

avec

$$A(\Pi^p) = \text{Var}_0[\Pi^p(X_0)\Delta^m g(X_0; \theta_0)].$$

Théorème 4.6.3 *Supposons $A(\Pi^p)$ inversible pour un $p \in \mathbb{N}^*$ et admettons satisfaites les hypothèses (P), (N), (S) et (Π) . Soient $g(\cdot; \theta_0) \in W^d(m)$ et $K \in K^d(m)$ pour un $m \in \mathbb{N}^*$. Si $h_n^{-1} = o(n^{\frac{1}{d+2m}})$, alors la région critique*

$$C_n(\Pi^p) = \left\{ T_n(\Pi^p)' A(\Pi^p)^{-1} T_n(\Pi^p) \geq 4 \chi_{1-\alpha}^2(p) \right\},$$

où $\chi_{1-\alpha}^2(p)$ est le quantile d'ordre $1-\alpha$ de la loi du $\chi^2(p)$, définit un test de niveau asymptotique α et asymptotiquement convergent pour tester H_0 contre H_0^c .

Démonstration: Le test défini par la région critique $C_n(\Pi^p)$ est en effet de niveau asymptotique α d'après la normalité asymptotique du vecteur aléatoire $T_n(\Pi^p)$, et de la régularité imposée à $A(\Pi^p)$. Plaçons nous maintenant sous une alternative fixe et notons $g(\cdot; \theta_1)$ la densité commune des variables aléatoires observées. D'après la démonstration du Théorème 4.3.10 nous avons

$$\frac{1}{\sqrt{n}h_n^{-m}} T_n(\Pi^p) = \int_{\mathbb{R}^d} \{g(x; \theta_1) - g(x; \theta_0)\}^2 \Pi^p(x) dx + o_p(1).$$

En tenant compte de la condition d'identifiabilité imposée au modèle paramétrique et de l'hypothèse (Π) , nous concluons que $\|T_n(\Pi^p)\|$ converge en probabilité vers $+\infty$, lorsque n tend vers l'infini, si $\theta_1 \neq \theta_0$. Ceci entraîne la convergence asymptotique du test. ■

4.6.3 Calculs de puissance pour le test $C_n(\Pi^p)$

On commence par l'étude des propriétés asymptotiques de $\xi_n(\Pi^p) = T_n(\Pi^p)' A(\Pi^p)^{-1} T_n(\Pi^p)/4$ sous une suite d'alternatives locales. On le fera à l'aide du lemme suivant qui est une conséquence immédiate du développement (4.6.2):

Lemme 4.6.4 *Supposons $A(\Pi^p)$ inversible pour un $p \in \mathbb{N}^*$ et admettons satisfaites les hypothèses (AL), (N), (S) et (II). Soient $g(\cdot; \theta_0) \in W^d(m)$ et $K \in K^d(m)$ pour un $m \in \mathbb{N}^*$. Soit de plus $h_n^{-1} = o(n^{\frac{1}{d+2m}})$. Sous la suite d'alternatives locales (4.6.1), nous avons*

$$T_n(\Pi^p)/2 = \mathcal{Z}_n + \sqrt{n}h_n^{-m} \left\{ a_n^2 \frac{C(\Pi^p; \delta)}{2} + h_n^m a_n B(\Pi^p) \right\} \delta + o_p(\sqrt{n}h_n^{-m} a_n^2) + o(\sqrt{n}a_n) + o_p(1),$$

où

$$B(\Pi^p) = E_0 \left[\Pi^p(X_0) \Delta^m g(X_0; \theta_0) \frac{\partial \log g}{\partial \theta'}(X_0; \theta_0) \right],$$

$$C(\Pi^p; \delta) = E_0 \left[\Pi^p(X_0) \delta' \frac{\partial g}{\partial \theta}(X_0; \theta_0) \frac{\partial \log g}{\partial \theta'}(X_0; \theta_0) \right],$$

et \mathcal{Z}_n est un vecteur aléatoire asymptotiquement normal de moyenne zéro et de matrice de covariance $A(\Pi^p)$.

Nous pouvons maintenant établir le comportement asymptotique de $\xi_n(\Pi^p)$ pour toute suite d'alternatives locales avec $\delta \neq 0$, suivant les différents ordres de convergence vers zéro de la suite a_n . Nous obtenons alors:

a) $\xi_n(\Pi^p) \xrightarrow[n \rightarrow +\infty]{d} \chi^2(p)$, si $a_n = o\left(\frac{1}{\sqrt{n}}\right)$;

b) $\xi_n(\Pi^p) \xrightarrow[n \rightarrow +\infty]{d} \chi^2(p; \delta' B(\Pi^p)' A(\Pi^p)^{-1} B(\Pi^p) \delta)$, si $a_n = \frac{1}{\sqrt{n}}$, où le paramètre de décentrage est non nul pour tout $\delta \neq 0$, si et seulement si, le rang de $B(\Pi^p)$ est égal à K : $\text{rg}B(\Pi^p) = K$;

c) $\xi_n(\Pi^p) \rightarrow +\infty$, en probabilité, si $\frac{1}{\sqrt{n}} = o(a_n)$ et $a_n \neq h_n^m$, puisque d'après l'hypothèse (II), $\|C(\Pi^p; \delta)\delta\| > 0$, pour tout $\delta \neq 0$.

Finalement,

d) $\xi_n(\Pi^p) \rightarrow +\infty$, en probabilité, si $a_n = h_n^m$, si et seulement si, la condition

$$(\mathbf{LU}) : \left[\frac{C(\Pi^p; \delta)}{2} + B(\Pi^p) \right] \delta = 0 \Rightarrow \delta = 0,$$

est satisfaite.

Si la vérification de la condition $\text{rg}B(\Pi^p) = K$ entraîne $p \geq K$, par contre la condition (LU) peut être satisfaite pour des valeurs de p inférieures à K . En effet, si par exemple, $K = 2$ (resp. $K = 3$) et $p = 1$, la condition (LU) est satisfaite si la conique (resp. quadrique) $[C(\Pi^p; \delta)/2 + B(\Pi^p)]\delta = 0$ se réduit à l'origine. Remarquons aussi, que la vérification conjointe des conditions $\text{rg}B(\Pi^p) = K$ et (LU) peut nous obliger à prendre p strictement supérieur à K .

C'est par exemple le cas où $K = 1$, où pour $p = 1$ la condition $\text{rg}B(\Pi^p) = 1$ équivaut à $B(\Pi^1) \neq 0$, tandis que la condition (LU) équivaut à $B(\Pi^1) = 0$.

Nous avons donc le résultat suivant:

Théorème 4.6.5 *Plaçons nous dans les conditions du Lemme 4.6.4. Le test défini par la région critique $C_n(\Pi^p)$ est USSB avec $a_n^I = \frac{1}{\sqrt{n}}$, si et seulement si $\text{rg}B(\Pi^p) = K$ et la condition (LU) est satisfaite. De plus, la puissance asymptotique du test est mesurée par la matrice*

$$J(\Pi^p) = B(\Pi^p)' A(\Pi^p)^{-1} B(\Pi^p),$$

qui satisfait la relation

$$J(\Pi^p) \ll I(\theta_0),$$

où \ll est la relation d'ordre usuelle sur les matrices symétriques.

Démonstration: D'après la discussion précédente, il suffit de voir que la relation $J(\Pi^p) \ll I(\theta_0)$ est une conséquence immédiate de l'interprétation de $I(\theta_0)$ comme borne d'efficacité asymptotique. ■

4.6.4 Choix optimal de Π^p

D'après le Théorème 4.6.5, nous concluons que les différentes régions critiques $C_n(\Pi^p)$ peuvent être comparées à partir de la matrice $J(\Pi^p)$; en particulier, nous obtenons une puissance optimale si et seulement si, pour un $p \geq K$,

$$J(\Pi^p) = I(\theta_0).$$

Cette puissance optimale peut être obtenue si le vecteur des fonctions de poids est choisi de façon convenable.

Théorème 4.6.6 *Admettons que pour un $p \geq K$, la fonction $\Pi^p = (\Pi^K, \Pi^{p-K})'$, à valeurs dans $\mathbb{R}^K \times \mathbb{R}^{p-K}$, satisfait (II) et est telle que $B(\Pi^{p-K}) = 0$ et $\Pi^K(\cdot) = C \Pi^*(\cdot)$, p.s., où C est une matrice inversible et Π^* est la fonction vectorielle définie pour $x \in \mathbb{R}^d$, par*

$$\Pi^*(x) = \frac{\partial \log g}{\partial \theta}(x; \theta_0) \{ \Delta^m g(x; \theta_0) \}^{-1} \mathbb{1}_{\{ \Delta^m g(x; \theta_0) \neq 0 \}}.$$

Si la fonction $\Delta^m g(\cdot; \theta_0)$ est presque sûrement non nulle dans \mathbb{R}^d , alors

$$J(\Pi^p) = I(\theta_0).$$

Démonstration: D'après la définition de Π^K , on a

$$\Pi^K(x) \Delta^m g(x; \theta_0) = C \frac{\partial \log g}{\partial \theta}(x; \theta_0),$$

pour presque tout $x \in \mathbb{R}^d$. Ainsi,

$$B(\Pi^K) = E_0 \left[C \frac{\partial \log g}{\partial \theta}(X_0; \theta_0) \frac{\partial \log g}{\partial \theta'}(X_0; \theta_0) \right] = C I(\theta_0) \text{ et}$$

$$A(\Pi^K) = \text{Var}_0 \left[C \frac{\partial \log g}{\partial \theta}(X_0; \theta_0) \right] = C I(\theta_0) C'.$$

Finalement, comme $B(\Pi^{p-K}) = 0$, nous avons

$$J(\Pi^p) = B(\Pi^K)' A(\Pi^K)^{-1} B(\Pi^K) = I(\theta_0). \quad \blacksquare$$

En tenant en compte que le vecteur des fonctions de poids $\Pi^p = (\Pi^K, \Pi^{p-K})'$ satisfait l'hypothèse (II), la fonction $\Pi^* = C^{-1} \Pi^K$, définie dans le théorème précédent, doit en particulier être bornée. Si ce n'est pas le cas, nous pouvons prendre une version tronquée $\Pi_L^*(x) = \Pi^*(x) \mathbb{I}_{\{\|\Pi^*(x)\| \leq L\}}$, $x \in \mathbb{R}^d$, pour $L > 0$, et la perte de puissance peut être prise aussi petite que l'on veut.

Le vecteur des pondérations $\Pi^K = (\pi_1, \dots, \pi_K)'$, défini dans le Théorème 4.6.6, est dans la suite explicite dans les cas particuliers de tests sur un paramètre de position ou d'échelle.

Exemple 4.6.7 (Paramètre de position) Dans le cas où $g(x; \theta) = g(x - \theta)$, avec $d = 1$ et $\theta \in \mathbb{R}$ ($K = 1$), et on veut tester $\theta = \theta_0$ contre $\theta \neq \theta_0$, on peut choisir

$$\pi_1(x) = g'(x - \theta_0) [g(x - \theta_0) g^{(m)}(x - \theta_0)]^{-1},$$

si $g^{(m)}(x - \theta_0) \neq 0$ et $\pi_1(x) = 0$, sinon. Si on prend $m = 1$, on trouve pour x tel que $g'(x - \theta_0) \neq 0$,

$$\pi_1(x) = [g(x - \theta_0)]^{-1}.$$

On notera que cette pondération est exactement celle intervenant dans une procédure de test de type test du khi-deux. La statistique $T_n^2(\pi_1)$ correspondante, apparaît ainsi comme une analogue de cette procédure usuelle fondée sur un estimateur à noyau de la densité.

Exemple 4.6.8 (Paramètre de position et d'échelle) Dans le cas où $g(x; \mu, \sigma) = \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right)$, pour $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ et $\sigma > 0$ ($K=2$), et on veut tester $(\mu, \sigma^2) = (\mu_0, \sigma_0^2)$ contre $(\mu, \sigma^2) \neq (\mu_0, \sigma_0^2)$, les fonctions π_1 et π_2 peuvent être choisies égales à

$$\pi_1(x) = \frac{1}{\sigma_0} g'\left(\frac{x-\mu_0}{\sigma_0}\right) \left[g\left(\frac{x-\mu_0}{\sigma_0}\right) g^{(m)}\left(\frac{x-\mu_0}{\sigma_0}\right) \right]^{-1} \text{ et}$$

$$\pi_2(x) = \left\{ \frac{1}{\sigma_0} + \frac{x-\mu_0}{\sigma_0^2} g'\left(\frac{x-\mu_0}{\sigma_0}\right) \left[g\left(\frac{x-\mu_0}{\sigma_0}\right) \right]^{-1} \right\} \left[g^{(m)}\left(\frac{x-\mu_0}{\sigma_0}\right) \right]^{-1},$$

si $g^{(m)}\left(\frac{x-\mu_0}{\sigma_0}\right) \neq 0$, et égales à zéro, sinon.

D'après les Théorèmes 4.6.5 et 4.6.6, nous atteindrons la puissance asymptotique optimale si le vecteur des fonctions de poids $\Pi^p = (\Pi^K, \Pi^{p-K})$ est choisi de façon que les conditions (II) et (LU) soient satisfaites, où Π^K est défini dans le Théorème 4.6.6, et Π^{p-K} est tel que $\text{rg}A(\Pi^{p-K}) = p - K$ et $B(\Pi^{p-K}) = 0$. La complexité des calculs associés à la vérification de telles conditions rend la méthodologie que nous venons de décrire de difficile application pratique.

Dans le cas plus simple $K = 1$ et en reprenant l'Exemple 4.6.7, nous verrons qu'il est toujours possible de choisir l'ordre m et une fonction de poids π_2 de façon que les conditions précédentes soient satisfaites.

Exemple 4.6.7 (cont.) Nous avons

$$\begin{aligned} B(\pi_1) &= -c_K \int_{\mathbb{R}} [g'(x)]^2 [g(x)]^{-1} dx = -c_K I(\theta_0), \\ C(\pi_1; \delta) &= \delta \int_{\mathbb{R}} [g'(x)]^3 [g^{(m)}(x)]^{-1} [g(x)]^{-1} dx, \\ B(\pi_2) &= -c_K \int_{\mathbb{R}} g^{(m)}(x - \theta_0) g'(x - \theta_0) \pi_2(x) dx, \text{ et} \\ C(\pi_2; \delta) &= \delta \int_{\mathbb{R}} [g'(x - \theta_0)]^2 \pi_2(x) dx, \end{aligned}$$

où $c_K = \frac{(-1)^m}{m!} \int_{\mathbb{R}} u^m K(u) du$. Indépendamment du choix de π_2 , remarquons que si on choisit $m = 1$, nous avons $B(\pi_1) = -c_K C(\pi_1; \delta) / \delta$ et $B(\pi_2) = -c_K C(\pi_2; \delta) / \delta$, et donc la condition (LU) n'est pas satisfaite. Ainsi, la pondération $\pi_1(x) = [g(x - \theta_0)]^{-1}$ n'est pas une pondération optimale. Nous allons maintenant considérer deux possibles choix de π_2 :

$$\pi_2(x) = 1 \text{ et } \pi_2(x) = [g(x - \theta_0)]^{-1}, \text{ pour } x \in \mathbb{R}.$$

En chacun de ces cas, le vecteur des fonctions de poids $\Pi^2 = (\pi_1, \pi_2)'$ satisfait les conditions imposées par l'hypothèse (II) sauf la condition d'être borné. Nous savons, dans ce cas, que la perte de puissance peut être prise aussi petite que l'on veut. Si $\pi_2(\cdot) = 1$, et si on choisit $m = 2$ (ou m pair) nous avons $B(\pi_2) = 0$. Comme $C(\pi_2; \delta)$ est non nul, pour tout $\delta \neq 0$, on conclut que la condition (LU) est satisfaite. Finalement, la matrice $A(\Pi^2)$ est inversible, puisque elle est diagonale d'éléments diagonaux non nuls. La même conclusion peut être obtenue dans le cas $\pi_2(\cdot) = [g(\cdot - \theta_0)]^{-1}$, si on admet que $g(x) = g(-x)$, $x \in \mathbb{R}$.

Bibliographie

- [1] Anderson, T.W. et Darling, D.A. (1952). Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *Ann. Math. Statist.*, **23**, 193-212.
- [2] Altman, N.S. (1990). Kernel smoothing of data with correlated errors. *J. Amer. Statist. Assoc.*, **85**, 749-759.
- [3] Azzalini, A., Bowman, A. et Härdle, W. (1989). On the use of nonparametric regression for model checking. *Biometrika*, **76**, 1-12.
- [4] Bickel, P.J. et Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.*, **1**, 1071-1095.
- [5] Bierens, H.J. (1987). Kernel Estimators of Regression Functions. *In: Advances in Econometrics (Fifth World Congress of the Econometric Society, 1985, Boston)*, Vol. 1, ed. Bewley, T.F., Cambridge University Press, 99-144.
- [6] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- [7] Billingsley, P. (1986). *Probability and Measure*. Wiley, New York.
- [8] Bosq, D. (1983). Lois limites et efficacité asymptotique des tests hilbertiens de dimension finie sous des hypothèses adjacentes. *Statist. Analyse de données*, **8**, 1-40.
- [9] Bosq, D. et Lecoutre, J-P. (1987). *Théorie de l'Estimation Fonctionnelle*. Economica, Paris.
- [10] Bosq, D. (1988). Prédiction nonparamétrique d'un processus stationnaire non borné. DP. LSTA 81, Université Paris VI.
- [11] Bosq, D. (1989). Tests du χ^2 généralisé. Comparaison avec le test du χ^2 classique. *Rev. Statist. Appliquée*, **38**, 43-52.
- [12] Bradley, R.C. (1986). Basic properties of strong mixing conditions. *In: Dependence in Probability and Statistics, a survey of recent results*, ed. Eberlein, E. et Taqqu, M., Birkhäuser, Boston, 165-192.
- [13] Brown, B.M. (1971). Martingale central limit theorems. *Ann. Math. Statist.*, **42**, 59-66.
- [14] Burguete, J., Gallant, R. et Souza, G. (1982). On unification of the asymptotic theory of nonlinear econometric methods. *Econometric Review*, **1**, 151-190.
- [15] Cacoullos, T. (1966). Estimation of a multivariate density. *Ann. Inst. Statist. Math.*, **18**, 179-189.
- [16] Carbon, M. (1988) *Inégalités de Grandes Déviations dans les Processus. Applications à l'Estimation Fonctionnelle*. Thèse de Doctorat de l'Université de Paris VI.

- [17] Castellana, J.V. et Leadbetter, M.R. (1986). On smoothed probability density estimation for stationary processes. *Stoch. Proc. Appl.*, **21**, 179-193.
- [18] Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.*, **74**, 829-836.
- [19] Cochran, W.G. (1952). The χ^2 test of goodness of fit. *Ann. Math. Statist.*, **23**, 315-345.
- [20] Csörgő, M. et Révész, P. (1981) *Strong Approximations in Probability and Statistics*. Academic Press, New York.
- [21] Davydov, YU.A. (1968). Convergence of distributions generated by stationary stochastic processes. *Theory Probab. Appl.*, **13**, 691-696.
- [22] Davydov, YU.A. (1973). Mixing conditions for Markov chains. *Theory Probab. Appl.*, **18**, 312-328.
- [23] de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Rel. Fields*, **75**, 261-277.
- [24] Delecroix, M. (1987). *Sur l'estimation et la prévision non-paramétrique des processus esgodiques*. Thèse de Doctorat de l'Université des Sciences et Techniques de Lille Flandres-Artois.
- [25] Devroye, L. et Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View*. Wiley, New York.
- [26] Diebolt, J. et Guégan, D. (1990). Probabilistic properties of the general nonlinear markovian process of order one and applications to time series modelling. Rapport Technique 125, CNRS, Paris.
- [27] Doob, J.L. (1953). *Stochastic Processes*. Wiley, New York.
- [28] Doukhan, P. (1994). *Mixing: properties and examples*. Lect. Notes in Statist., 85, Springer-Verlag, New York.
- [29] Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables. *In: Proceedings of the 6th Berkeley Sympo. Math. Statist. Probab.*, Univ. Calif. 1970, Vol. 2, 513-535.
- [30] Eubank, R.L. et LaRiccia, V.N. (1992). Asymptotic comparison of Crámer-von Mises and non-parametric function estimation techniques for testing goodness-of-fit. *Ann. Statist.*, **20**, 2071-2086.
- [31] Fan, J. (1992). Design-adaptative Nonparametric Regression. *J. Amer. Statist. Assoc.*, **87**, 998-1004.
- [32] Fan, Y. (1994). Testing the goodness of fit of a parametric density function by kernel method. *Econometric Theory*, **10**, 316-356.
- [33] Gastwirth, J.L. et Rubin, H. (1975). The asymptotic distribution theory of the empirical CDF for mixing processes. *Ann. Statist.*, **3**, 809-824.
- [34] Ghosh, B.K. et Huang, W.-M. (1991). The power and optimal kernel of the Bickel-Rosenblatt test for goodness of fit. *Ann. Statist.*, **19**, 999-1009.
- [35] Gouriéroux, C. (1992). *Modeles ARCH et Applications Financières*. Economica, Paris.
- [36] Gouriéroux, C. et Monfort, A. (1989). *Statistique et Modèles Econométriques*. Economica, Paris.
- [37] Gouriéroux, C., Monfort, A. et Tenreiro, C. (1994). Kernel M-estimators: nonparametric diagnostics for structural models. Document de travail 9405, CEPREMAP, Paris.

-
- [38] Gouriéroux, C., Monfort, A. et Trognon, A. (1984). Pseudo maximum likelihood methods: theory. *Econometrica*, **52**, 681-700.
- [39] Gozalo, P. et Linton, O. (1994). Local nonlinear least squares estimation: using parametric information nonparametrically. Cowles foundation discussion paper 1075, Yale University.
- [40] Gregory, G.G. (1977). Large sample theory for U-statistics and tests of fit. *Ann. Statist.*, **5**, 110-123.
- [41] Györfi, L., Härdle, W., Sarde, P. et Vieu, P. (1989). *Nonparametric Curve Estimation from Time Series*. Lect. Notes in Statist., 60, Springer-Verlag, New York.
- [42] Hall, P. (1984). Central limit theorem for integrated square error properties of multivariate nonparametric density estimators. *J. Multivariate Anal.*, **14**, 1-16.
- [43] Hall, P. (1984). Integrated square error properties of kernel estimators of regression functions. *Ann. Statist.*, **12**, 241-260.
- [44] Hall, P. et Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.
- [45] Halmos, P.R. (1950). *Measure Theory*. D.Van Nostrand, New York.
- [46] Härdle, W. (1984). Robust regression function estimation. *J. Multivariate Anal.*, **14**, 169-180.
- [47] Härdle, W. (1989). Asymptotic maximal deviation of M-smoothers. *J. Multivariate Anal.*, **29**, 163-179.
- [48] Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- [49] Härdle, W. (1990). *Smoothing Techniques: with implementation in S*. Cambridge University Press.
- [50] Härdle, W. et Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.*, **21**, 1926-1947.
- [51] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, **19**, 293-325.
- [52] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **33**, 73-101.
- [53] Huber, P.J. (1981). *Robust Statistics*. Wiley, New York.
- [54] Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Theory Probab. Appl.*, **7**, 349-382.
- [55] Ibragimov, I.A. et Rozanov, Yu.A. (1974). *Processus Aléatoires Gaussiens*. Editions Mir, Moscou.
- [56] Ibragimov, I.A. et Solev, V.N. (1969). A condition for regularity of a gaussian stationary process. *Soviet Math. Dokl.*, **10**, 371-375.
- [57] Jammalamadaka, R.S. et Janson, S. (1986). Limit theorems for a triangular scheme of U-statistics with applications to inter-point distances. *Ann. Probab.*, **14**, 1347-1358.
- [58] Jennrich, R. (1969). Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.*, **40**, 633-643.
- [59] Khashimov, Sh.A. (1987). On the asymptotic distribution of the generalized U-statistic for dependent variables. *Theory Probab. Appl.*, **32**, 373-375.

- [60] Kolmogorov, A.N. et Rozanov, Yu.A. (1960). On strong mixing conditions for stationary gaussian processes. *Theory Probab. Appl.*, **5**, 204-208.
- [61] Konakov, V.D. (1977). On a global measure of deviation for an estimate of the regression line. *Theory Probab. Appl.*, **22**, 858-868.
- [62] Loève, M. (1977). *Probability Theory I*. 4 ed., Springer-Verlag, New York.
- [63] Loève, M. (1978). *Probability Theory II*. 4 ed., Springer-Verlag, New York.
- [64] Mack, Y.P. et Silverman, B. (1982). Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. Verw. Gebiete*, **61**, 405-415.
- [65] Meloche, J. (1990). Asymptotic behaviour of the mean integrated squared error of kernel density estimators for dependent observations. *Canad. J. Statist.*, **18**, 205-211.
- [66] Milbrodt, H. et Stasser, H. (1990). On the asymptotic power of the two-sided Kolmogorov-Smirnov test. *J. Statist. Plann. Inference*, **26**, 1-23.
- [67] Mokkaïem, A. (1990). Propriétés de mélange des processus autorégressifs polynomiaux. *Ann. Inst. H. Poincaré*, **26**, 219-260.
- [68] Monfort, A. (1980). *Cours de Probabilités*. Economica, Paris.
- [69] Monfort, A. (1980). *Cours de Statistique Mathématique*. Economica, Paris.
- [70] Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.*, **9**, 141-142.
- [71] Nadaraya, E.A. (1965). On non-parametric estimates of density functions and regression curves. *Theory Probab. Appl.*, **10**, 186-190.
- [72] Nadaraya, E.A. (1983). A limit distribution of the square error deviation of nonparametric estimators of the regression function. *Z. Wahrsch. Verw. Gebiete*, **64**, 37-48.
- [73] Nguyen, H.T. (1984). Recursive nonparametric estimation in stationary Markov processes. *Pub. Inst. Stat. Univ. Paris*, **29**, 65-84.
- [74] O'Neil, K.A. et Redner, R.A. (1993). Asymptotic distribution of weighted U-statistics of degree 2. *Ann. Math. Statist.*, **21**, 1159-1169.
- [75] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**, 1065-1076.
- [76] Pham, T.D. et Tran, L.T. (1985). Some mixing properties of time series models. *Stochastic Process. Appl.*, **19**, 297-303.
- [77] Prakasa Rao, B.L.S. (1983). *Nonparametric Functional Estimation*. Academic Press, New York.
- [78] Ranga Rao, R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.*, **33**, 659-680.
- [79] Révész, P. (1972). On empirical density function. *Periodica Mathematica Hungarica*, **2**, 85-110.
- [80] Révész, P. (1978). A strong law of the empirical density function. *Periodica Mathematica Hungarica*, **9**, 317-324.
- [81] Robinson, P.M. (1983). Nonparametric estimators for time series. *J. Time Series Anal.*, **4**, 185-207.

-
- [82] Robinson, P.M. (1984). Robust nonparametric autoregression. *In: Robust and Nonlinear Time Series Analysis*, Lect. Notes in Statist., 26, 247-255.
- [83] Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.*, **27**, 832-837.
- [84] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA.*, **42**, 43-47.
- [85] Rosenblatt, M. (1970). Density estimates and Markov sequences. *In: Proc. First International Symposium on Nonparametric Techniques, Indiana University*, ed. Madan Puri, 199-210.
- [86] Rosenblatt, M. (1971). *Markov Processes Structure and Asymptotic Behavior*. Springer-Verlag, New York.
- [87] Rosenblatt, M. (1974). *Random Processes*. Springer-Verlag, New York.
- [88] Rosenblatt, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.*, **3**, 1-14. [correction **10** (1982), 646].
- [89] Rosenblatt, M. (1976). On the maximal deviation of k-dimensional density estimates. *Ann. Probab.*, **4**, 1009-1015.
- [90] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser, Boston.
- [91] Rosenblatt, M. (1991). *Stochastic Curve Estimation*. NSF-CBMS Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, Hayward.
- [92] Roussas, G.G. (1969). Nonparametric estimation in Markov processes. *Ann. Inst. Statist. Math.*, **21**, 73-87.
- [93] Roussas, G.G. (1988). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann. Inference*, **18**, 135-149.
- [94] Roussas, G.G. (1990). Asymptotic normality of the kernel estimate under dependence conditions: application to hazard rate. *J. Statist. Plann. Inference*, **25**, 81-104.
- [95] Roussas, G.G. (1990). Nonparametric regression estimation under mixing conditions. *Stochastic Process. Appl.*, **36**, 107-116.
- [96] Schuster, E.F. (1972). Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. *Ann. Math. Statist.*, **43**, 84-88.
- [97] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- [98] Stone, C.J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.*, **8**, 1348-1360.
- [99] Takahata, H. et Yoshihara, K. (1987). Central limit theorems for integrated square error of non-parametric density estimators based on absolutely regular random sequences. *Yokohama Math. J.*, **35**, 95-111.
- [100] Tenreiro, C. (1994). O teste de ajustamento de Bickel-Rosenblatt generalizado a processos misturadores. *In: A Estatística e o Futuro, e o Futuro da Estatística (Actas do I Congresso Anual da SPE, 1993, Vimeiro)*, ed. Pestana, D. et al., Edições Salamandra, Lisboa, 89-98.
- [101] Tenreiro, C. (1994). Estimation non paramétrique d'une fonction de moments conditionnels sous des conditions locales. *C. R. Acad. Sci. Paris, Ser. I*, **318**, 1149-1152.

- [102] Tenreiro, C. (1994). Um teste de adequação dum modelo paramétrico de regressão sob condições de dependência. A aparecer nas Actas do II Congresso Anual da SPE, Luso, 20 a 23 de Junho de 1994.
- [103] Tenreiro, C. (1995). Théorèmes limites pour les erreurs quadratiques intégrées des estimateurs à noyau de la densité et de la régression sous des conditions de dépendance. A aparecer em *C. R. Acad. Sci. Paris*, Ser. I, **320**.
- [104] Tiago de Oliveira, J. (1963). Estatística de densidades, resultados assintóticos. *Rev. Fac. Ciências Lisboa*, 2 sér., A, **9**, 111-206.
- [105] Tusnády, G. (1973). An asymptotically efficient test for fitting a distribution. *Periodica Mathematica Hungarica*, **3**, 157-165.
- [106] Volkonskiĭ, V.A. et Rozanov, Yu.A. (1959). Some limit theorems for random functions. I. *Theory Probab. Appl.*, **4**, 178-197.
- [107] Volkonskiĭ, V.A. et Rozanov, Yu.A. (1961). Some limit theorems for random functions. II. *Theory Probab. Appl.*, **6**, 186-198.
- [108] Wald, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.*, **20**, 595-601.
- [109] Watson, G.S. (1964). Smooth regression analysis. *Sankhya A*, **26**, 359-372.
- [110] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **50**, 1-25.
- [111] Yakowitz, S.J. (1985). Nonparametric density estimation prediction and regression for Markov processes. *J. Amer. Statist. Assoc.*, **80**, 215-221.
- [112] Yoshihara, K. (1976). Limiting behavior of U-statistics for stationary, absolutely regular processes. *Z. Wahrsch. Verw. Gebiete*, **35**, 237-252.
- [113] Yoshihara, K. (1989). Limiting behavior of generalized quadratic forms generated by absolutely regular sequences. In: *Proc. Fourth Prague Conference on Asymptotic Statistics*, 539-547.
- [114] Yoshihara, K. (1989). Limiting behavior of generalized quadratic forms generated by absolutely regular sequences II. *Yokohama Math. J.*, **37**, 109-123.