Un test d'adéquation global pour la fonction de répartition conditionnelle

Sandie Ferrigno

ACADÉMIE DE MONTPELLIER

UNIVERSITÉ MONTPELLIER II

— SCIENCES ET TECHNIQUES DU LANGUEDOC —

Thèse

présentée à l'Université Montpellier II - Sciences et Techniques du Languedoc
pour l'obtention du

DIPLÔME DE DOCTORAT

Spécialité : *Mathématiques appliquées et applications des Mathématiques*
Section CNU : 26
Formation Doctorale : *Biostatistique*
Ecole Doctorale : *Information, Structures, Systèmes*

Un test d'adéquation global pour la fonction de répartition conditionnelle

par

Sandie FERRIGNO

Soutenue publiquement le **17 Décembre 2004** devant le jury composé de :

M. Jean-Noël BACRO	Université Montpellier II	Président du jury
M. Alain BERLINET	Université Montpellier II	Examinateur
M. Gilles DUCHARME	Université Montpellier II	Directeur de thèse
M. Michel HAREL	IUFM du Limousin	Rapporteur
M. Jérôme SARACCO	Université de Bourgogne	Examinateur
M. Pascal SARDA	Université Toulouse II	Rapporteur

à ma petite cousine Vickie

Remerciements

Ce travail n'aurait jamais vu le jour sans l'aide et le soutien de diverses personnes qui ont joué un rôle important auprès de moi durant ces 4 années et je profite donc de ces quelques mots pour les remercier.

Je voudrai commencer par remercier Gilles Ducharme pour tout d'abord m'avoir proposé un sujet de thèse aussi intéressant. Je le remercie aussi pour toute son aide, ses nombreuses lectures et relectures pendant les derniers mois de rédaction et pour ce goût pour la recherche qu'il a su me transmettre durant ces 4 années.

Je voudrai poursuivre en remerciant vivement les deux rapporteurs de cette thèse, messieurs Michel Harel et Pascal Sarda. Je les remercie en particulier de tout l'intérêt qu'ils ont porté à mon travail.

Un grand merci à Jean-Noël Bacro pour avoir accepté d'être le Président de mon jury de thèse et pour sa gentillesse à mon égard.

Je remercie également Jérôme Saracco d'avoir accepté de faire partie de mon jury de thèse mais aussi pour toute la gentillesse qu'il a eu à mon égard durant ces années où j'ai pu le cotoyer à Montpellier. Je le remercie également pour tous les précieux conseils qu'il a pu me donner et pour la confiance qu'il m'a témoignée en me proposant divers enseignements durant mon année d'Attachée Temporaire d'Enseignement et de Recherche à Montpellier.

Je remercie aussi Alain Berlinet pour l'intérêt qu'il a porté à mon travail en acceptant de faire partie de mon jury de thèse.

Je souhaiterai à présent remercier Pierre Jacob, directeur de l'Equipe Probabilités et Statistique. Je le remercie en particulier de m'avoir confié des enseignements en Biostatistiques durant mon année d'Attachée Temporaire d'Enseignement et de Recherche à Montpellier. Je le remercie également pour les nombreux conseils qu'il a pu me donner et pour toute sa gentillesse.

Je voudrai à présent remercier les autres membres de l'Equipe Probabilités et Statistique qui à leur manière m'ont beaucoup apporté durant ces 4 années. Je commencerai par remercier Benoît Cadre pour toute l'aide qu'il m'a apportée pour cette thèse mais aussi pour son soutien moral et pour m'avoir motivée dans les moments difficiles. Je remercie aussi Ludovic Menneteau pour son aide et son soutien surtout lors de mes premières Journées de Statistique à Bruxelles en 2002. Merci aussi à Mohammed Mellouk et Ali Gannoun pour toute leur gentillesse et merci de m'avoir confié, chacun leur tour, des enseignements de Statistique en Deug B. Je remercie aussi Gérard Biau pour ses précieux conseils ainsi que André Mas pour toute la sympathie qu'il m'a témoignée depuis son arrivée à Montpellier.

Je voudrai maintenant remercier toutes les personnes travaillant à l'IUT STID de Carcassonne qui m'accueillent en tant qu'Attachée Temporaire d'Enseignement et de Recherche pour cette année universitaire 2004/2005. Je remercie en particulier son directeur, Denys Chaffardon ainsi que Sylvie Viguier-Pla pour toute sa gentillesse depuis mon arrivée à l'IUT. Merci aussi à Michel Meste, Mouna Kamel, Thierry Spinosa, Evelyne Armstrong, Gabriel

Fraisse et Jean-Marc Hugounnet pour toute la gentillesse et les encouragements qu'ils m'ont témoignés pour cette thèse.

Je voudrai poursuivre en remerciant Baptiste Chapuisat pour toute son aide en informatique mais surtout pour sa patience et sa gentillesse à mon égard pendant toutes ces années. Je remercie aussi chaleureusement Pierrette Arnaud et Bernadette Lacan pour leur disponibilité, leur soutien dans les moments difficiles et toute leur gentillesse. Merci aussi à Nicole Grachet, Geneviève Piard, Mireille Piquet et Eric Hugounenk. Merci enfin à Sarah Dubaquier et Cathy Araspin pour leur gentillesse et leur disponibilité.

Je voudrai poursuivre en remerciant mes amis doctorants et docteurs qui ont rendu cette thèse plus agréable. Je remercie notamment Marie-José Martinez pour son amitié depuis déjà deux ans ! Je la remercie aussi d'avoir partagé des moments détentes sportifs où elle a su me remonter le moral dans les moments difficiles ! Je remercie aussi Laurent Rouvière pour toute sa gentillesse. Une petite pensée aussi pour Mohannad Hassan, Thomas Verron mais aussi pour les "anciens" Biostatisticiens Olivier Gimenez, Pierre Lafaye de Michaux, Benoit Frichot et Mariem Mint-El-Mouvid. Je n'oublie pas aussi de remercier Guillaume Ricotta pour toute l'amitié qu'il me témoigne depuis quelques années déjà.

Viennent enfin les remerciements du coeur. Je voudrai commencer par remercier mes parents pour toute leur gentillesse et leur compréhension tout au long de mes études. Merci de m'avoir supportée même dans les moments difficiles et encore merci d'avoir toujours été là pour moi tout simplement. Un grand merci à mon petit frère Jérémie qui a toujours été un rayon de soleil dans ma vie. Je souhaite le remercier pour tout le soutien qu'il m'apporte depuis toujours et pour sa présence si précieuse à Montpellier près de moi.

Je voudrai aussi remercier chaleureusement mon papi et ma mamie pour leur coup de fil quotidien qui m'a souvent remonté le moral et pour leur soutien permanent. Merci d'avoir Scoubi près de vous, il a, à sa manière, embelli mes weekends à Sète. Merci aussi à ma grand-mère pour toutes ses pensées et ses prières pour moi.

Je remercie énormément Julien, mon soutien permanent, qui a vécu cette thèse au jour le jour à mes côtés. Je ne le remercierai jamais assez pour avoir été simplement près de moi surtout la dernière année qui a été la plus difficile pour moi moralement. Il a su m'écouter, me rassurer et a toujours cru en moi ce qui a fait ma force.

Je terminerai en remerciant d'abord ma famille Aveyronnaise et Savoyarde d'être venue à ma soutenance de thèse et de m'avoir encouragée durant toutes mes études. Je souhaite la bienvenue à ma petite cousine Vickie à qui je dédie cette thèse et je remercie du fond du coeur sa maman Christel pour tout ce qu'elle a fait pour moi depuis toujours. Merci aussi à ma marraine Elisabeth et à mon cousin Julien pour toute leur gentillesse.

Je remercie enfin mes amis Sétois et Montpelliérains ainsi que toutes les personnes qui sont venus assister à ma soutenance de thèse. Cela m'a beaucoup touchée. Merci en particulier à Pierre, Sophie, Nathalie, Magali et Sophie.

Table des matières

Introduction

Soient X et Y, deux variables aléatoires réelles. Dans le but d'expliquer la variable Y à partir de la variable X, il est courant que le statisticien ou la statisticienne propose d'ajuster un modèle. Cependant, il ou elle doit s'assurer de la validité du modèle proposé. Pour ce faire, il ou elle cherchera à ajuster le modèle à un échantillon $(X_i, Y_i)_{i=1,\ldots,n}$ de copies du couple de variables (X, Y) de départ. Par exemple, choisissons comme modèle de référence le modèle de régression :

$$Y_i = g(X_i, \theta) + \sigma(X_i)\epsilon_i, i = 1, \ldots, n. \tag{1}$$

Le fait de proposer un tel modèle nous amène à émettre un certain nombre d'hypothèses. Certaines, par exemple, s'articulent autour de la forme fonctionnelle de la quantité $g(., \theta)$. D'autres sont relatives à la forme fonctionnelle de la quantité $\sigma(.)$ et on doit aussi s'interroger quant à la distribution asymptotique du terme ϵ_i. D'autres questions relatives à ce modèle se profilent aussi rapidement. On peut notamment s'interroger sur l'exactitude de l'additivité du terme d'erreur mais aussi sur la distribution de probabilité de la variable X et surtout de son indépendance avec le terme ϵ. En pratique, toutes ces hypothèses ont une certaine importance sur la qualité du modèle proposé et la non validation de l'une ou l'autre de ces hypothèses pourrait rendre le modèle complètement incongru. Il est donc nécessaire de développer des méthodes permettant d'asseoir la validité du modèle par rapport à l'échantillon considéré afin de pouvoir l'utiliser par la suite avec un certain degré de confiance.

Les méthodes statistiques mises en place pour répondre à ce genre de demande sont les tests d'adéquation. De nombreux tests ont déjà été proposés dans la littérature pour assurer la validité de la plupart des hypothèses sous jacentes à l'utilisation du modèle (1). Mais tous ces tests sont "directionnels", dans le sens où ils permettent de ne tester qu'un seul aspect à la fois du modèle considéré. De plus, leur mise en place nécessite souvent la validité de certaines des autres hypothèses. Les procédures permettant d'asseoir complètement un modèle statistique comme celui de (1) sont donc de ce fait assez lourdes et il n'est donc pas déraisonnable de songer à mettre en place un test plus global qui permettrait de valider d'un seul coup la structure du modèle considéré. L'objet de cette thèse, qui s'articule autour de 6 chapitres, est le développement et l'étude d'un tel test.

Le premier chapitre de ce travail est consacré à la mise en place de la procédure de test d'adéquation qui va permettre de tester globalement toutes les hypothèses associées au modèle considéré. On utilise pour cela une quantité qui permet de capter toute l'information contenue dans le modèle proposé et d'expliquer globalement la variable Y en fonction de la variable X. Il s'agit de la fonction de répartition conditionnelle de Y sachant $\{X = x\}$ donnée par

$$F(y|x) = P(Y \leq y|X = x). \tag{2}$$

Si le modèle est exact, la fonction de répartition conditionnelle prendra une certaine valeur $F_0(y|x)$ et le problème de valider ce modèle se ramène alors à considérer le test de l'hypothèse nulle :

$$H_0 : F(y|x) = F_0(y|x). \tag{3}$$

Dans ce travail, nous proposons donc de construire un test global basé sur cette fonction de répartition. Notre approche s'appuie sur l'article précurseur de Bickel et Rosenblatt (1973), dont les idées ont été exploitées par de nombreux auteurs dont, entre autres, Härdle et Mammen (1993) ou bien encore Alcalá *et al.* (1999).

Notre test est basé sur une "opposition" standard qui consiste à comparer un estimateur non paramétrique de $F(y|x)$ à sa valeur supposée $F_0(y|x)$ et à rejeter l'hypothèse nulle (3) si la "distance" entre ces deux quantités dépasse un seuil critique. Ici, l'estimateur non paramétrique sera basé sur l'approche polynômiale locale, déjà étudiée par Fan et Gijbels (1996) dans un contexte de régression. On propose aussi d'utiliser une "distance" de type Cramér-von Mises (Cramér, 1928 et von Mises, 1941) généralisée comme statistique de test.

Le reste de la thèse s'articule de la manière suivante. Dans la première partie, on se place dans le cas où la distribution à tester sous l'hypothèse nulle (3), soit $F_0(y|x)$, est entièrement spécifiée. Ce cas est certes peu courant en pratique mais il permet d'obtenir plus commodément certains résultats asymptotiques. On considère au Chapitre 2 le cas particulier où la fonction de répartition conditionnelle $F(y|x)$ est un polynôme de degré fixé en x. Dans ce contexte, on propose une décomposition de la statistique de test qui va permettre de mettre en exergue une U-statistique dont les propriétés asymptotiques nous aident à déduire une région critique pour l'hypothèse nulle considérée. Dans le troisième chapitre, on généralise ces résultats au cas où la fonction de répartition conditionnelle supposée sous l'hypothèse nulle est une fonction plus générale. Moyennant quelques hypothèses supplémentaires sur l'estimation de la fonction de répartition, on obtient le même comportement asymptotique du test que dans le cas précédent.

Dans la deuxième partie de ce travail, on étend les résultats asymptotiques de la première partie au cas où la fonction de répartition supposée sous l'hypothèse nulle contient un certain nombre de paramètres inconnus. Ce cas de figure est beaucoup plus proche du contexte pratique dans lequel on rencontre ce type de modèle. On propose dans le Chapitre 4 une nouvelle décomposition de la statistique de test qui va permettre, en utilisant entre autres les résultats de la première partie, de déterminer une région critique pour ce nouveau genre d'hypothèses.

L'existence de plusieurs tests d'adéquation pour un même modèle pose la question parallèle de savoir si l'un d'eux se comporte mieux suivant un critère donné. Un des critères les plus répandus est la puissance. Les études par simulations ne représentent qu'une image partielle de la puissance des tests et de ce fait, on propose, dans le Chapitre 5, d'en faire une étude plus fine par le biais de la puissance asymptotique locale. Celle-ci passe par l'étude du comportement asymptotique de notre test sous des suites d'alternatives contigües, que l'on introduit au Chapitre 5. Enfin, l'ensemble des résultats obtenus dans ce travail est influencé par un certain nombre de paramètres. Cela nous amène à discuter au Chapitre 6 de critères de choix de ces divers paramètres qui interviennent implicitement ou explicitement dans l'expression de la statistique de test. On y propose en particulier un critère de choix de la fenêtre d'ajustement qui, comme plusieurs auteurs l'ont déjà noté dans d'autres contextes, joue un rôle important.

Ce travail se termine par un résumé des différents résultats obtenus et par la présentation de quelques pistes de recherches futures.

Chapitre 1

Test d'adéquation pour une fonction de répartition conditionnelle

1.1 Introduction

Soit (X, Y), un couple de variables aléatoires à valeurs dans $\mathbb{R}^2$. Pour étudier la relation liant X à Y, il est courant d'utiliser la "fonction de régression" définie par l'espérance conditionnelle de Y sachant $\{X = x\}$ et dénotée $E(Y|X = x)$. Cette fonction de régression est un résumé pratique de l'effet de X sur le comportement de Y.

Une autre quantité, toute aussi intéressante et contenant, elle, toute l'information liant X à Y est la fonction de répartition conditionnelle de Y sachant $\{X = x\}$, dénotée par

$$F(y|x) = P(Y \leq y|X = x). \tag{1.1}$$

L'intérêt de cette fonction provient du fait que la plupart des quantités statistiques utilisées en pratique pour comprendre la relation liant X et Y sont des fonctionnelles de cette fonction de répartition conditionnelle : outre l'espérance conditionnelle déjà évoquée, citons la médiane conditionnelle et plus généralement les quantiles conditionnels, ou encore la variance conditionnelle pour n'en donner que quelques unes.

De nombreuses procédures statistiques se basent sur l'hypothèse que la loi de Y sachant X provient d'une certaine distribution. Par exemple, un modèle statistique courant est celui du modèle de régression

$$Y = m(X) + \sigma(X)\epsilon,$$

où $E(\epsilon) = 0$ et $Var(\epsilon) = 1$. Supposons ϵ indépendant de X, entraînant alors

$$m(x) = E(Y|X = x), \sigma^2(x) = Var(Y|X = x).$$

On suppose aussi très souvent en pratique que $\epsilon \sim N(0,1)$, la loi Normale d'espérance 0 et de variance 1. Si, en plus, on suppose $m(x) = \beta_0 + \beta_1 x$ et $\sigma^2(x) = \sigma^2$, on obtient alors le modèle de régression linéaire simple

$$\mathcal{L}(Y|X=x) = N\left(\beta_0 + \beta_1 x, \sigma^2\right), \tag{1.2}$$

où $\mathcal{L}(Y|X=x)$[1] désigne la loi conditionnelle de Y étant donnée X. Ce modèle est d'une grande importance dans les applications de la statistique et au fil des ans, on a mis au point une multitude de méthodes statistiques pour estimer et tester ses paramètres. Ces méthodes reposent plus ou moins lourdement sur les suppositions qui sont faites. Il semble donc important de pouvoir valider ces suppositions à l'aide d'un test. On appelle ce type de procédure statistique un test d'adéquation.

Pour ce faire, une approche consiste à vérifier chacune de ces suppositions individuellement en s'appuyant sur un échantillon $(X_1, Y_1), \ldots, (X_n, Y_n)$, de n copies indépendantes de (X, Y). Par exemple, pour le modèle (1.2), on pourrait d'abord vérifier la normalité de ϵ par l'un ou l'autre des tests qui ont été développés à cet effet (D'Agostino et Stephens, 1986, Templet, 1995). On pourrait ensuite tester la linéarité de $m(x)$, encore là par l'un ou l'autre des tests conçus à cette fin (Härdle et Mammen, 1993, Alcalá *et al.*, 1999). Il conviendrait aussi de s'assurer que $\sigma^2(x) = \sigma$ en utilisant par exemple le test de Liero (2003). On devra par la suite tester l'indépendance entre X et ϵ. Ici, le problème est déjà plus difficile étant donné que ϵ est inobservable et à notre connaissance, il n'existe pas de test pour ce problème. Enfin, il faudra s'assurer de l'additivité du terme d'erreur, supposition pour laquelle il n'existe pas non plus de procédure, à notre connaissance.

Dans ce travail, on va développer une approche plus globale où toutes les hypothèses faites dans le but d'obtenir un modèle pour $F(y|x)$ sont testées simultanément. Plus précisément, le problème dont nous allons traiter est de construire un test pour valider ou infirmer l'hypothèse nulle :

$$H_0 : F(y|x) = F_0(y|x) \qquad \forall (x, y), \tag{1.3}$$

où $F_0(y|x)$ est la fonction de répartition conditionnelle correspondant au modèle supposé, par exemple dans le modèle (1.2), la fonction de répartition d'une $N(\beta_0 + \beta_1 x, \sigma^2)$.

En général $F_0(y|x)$ fait intervenir un certain nombre de paramètres. Dans l'exemple (1.2), ces paramètres sont les quantités β_0, β_1 et σ^2. Pour le test de (1.3), il arrive de rares cas où on peut supposer que ces paramètres sont connus, auquel cas $F_0(y|x)$ de (1.3) est

[1] L'existence de la loi conditionnelle de Y étant donnée X est due au Théorème de Jirina (Dudley, 1989).

dite entièrement spécifiée parce que tous les paramètres sont donnés. Mais généralement en pratique ces paramètres sont inconnus. Si on rassemble ces paramètres inconnus dans un vecteur θ appartenant à un espace paramétrique Θ et si on réécrit la fonction de répartition conditionnelle sous la forme $F_0(y|x;\theta)$, alors (1.3) devient l'hypothèse nulle composite

$$H_0 : F(\,.\,|\,.\,) \in \{F_0(\,.\,|\,.\,;\theta), \theta \in \Theta\}. \tag{1.4}$$

Le problème (1.4) étant généralement plus difficile, nous allons tout d'abord considérer le problème de tester (1.3) pour repousser l'analyse de (1.4) au Chapitre 4.

Une approche possible est de comparer $F_0(y|x)$ à un estimateur de $F(y|x)$ et de rejeter H_0 si la "distance" entre $F_0(y|x)$ et cet estimateur est trop grande. C'est l'approche que nous allons développer dans ce travail. Il existe différentes méthodes pour estimer $F(y|x)$ à partir d'un échantillon (X_i, Y_i), $i = 1 \ldots n$ de n copies indépendantes de (X, Y). Ici, nous allons nous servir de méthodes d'estimation non paramétriques. En particulier, l'une d'entre elles, très présente dans la littérature actuelle, consiste à travailler avec des "polynômes locaux" et s'appelle de ce fait l'estimation polynômiale locale. C'est cette méthode que nous allons utiliser dans la suite.

L'estimation polynômiale locale a été introduite par Stone (1977, 1980, 1982) et Cleveland (1979). Elle a, jusqu'à présent, surtout été utilisée pour l'estimation de l'espérance conditionnelle et a mené aux techniques dites de "régression polynômiale locale". Ces méthodes sont des versions sophistiquées de la régression polynômiale classique, où on applique localement cette dernière. Plus précisément, pour tout point x, la fonction de régression $m(x) = E(Y|X = x)$ est estimée par un polynôme de degré fixé calculé à partir des données "proches" de x.

Historiquement, la régression polynômiale locale a connu ses premières applications quand Stone (1977) a considéré la régression linéaire locale dans son schéma général de "régression non paramétrique". Cleveland (1979) a repris l'idée en introduisant la possibilité de pondérer plus lourdement les points les plus proches de x. Ensuite, Lejeune (1985) a étudié le comportement de l'erreur quadratique moyenne. Il a montré une équivalence entre la régression polynômiale locale et la régression par noyaux optimaux, qui toutes les deux éliminent le biais jusqu'à un ordre donné et sont de variance minimale. Fan (1992, 1993) et Fan et Gijbels (1996) ont montré que la régression polynômiale locale présente plusieurs avantages sur les autres méthodes de régression non paramétriques : ils ont remarqué sa capacité à contourner le principal écueil des méthodes à noyau classiques en corrigeant automatiquement les effets de bord tout en conservant les propriétés d'optimalité théoriques.

Ruppert et Wand (1994) ont montré que son extension au cas multivarié est aisée. En résumé, les dix dernières années ont été marquées par un développement prodigieux de la régression polynômiale locale, tant au point de vue théorique qu'appliqué.

Malgré le succès remarquable de l'estimation polynômiale locale de la fonction $m(x)$, les applications à d'autres fonctionnelles de $F(y|x)$ ont été plutôt timides. Tsybakhov (1986) a introduit l'estimateur linéaire local de quantiles conditionnels. Le cas de la variance conditionnelle apparaît d'abord comme une étape dans la procédure de choix de la fenêtre proposée par Fan et Gijbels (1995). Puis, Ruppert *et al.* (1997) proposent un estimateur polynômial local de cette variance conditionnelle et montrent que la propriété de correction automatique des effets de bord reste valable. Härdle et Tsybakov (1997) étudient aussi l'estimateur polynômial local de la variance conditionnelle mais pour un modèle autorégressif. Fan et Yao (1998) proposent un deuxième estimateur de la variance conditionnelle. Ils montrent que, sans connaître la fonction de régression, cette variance peut être estimée en effectuant une régression linéaire locale sur les résidus au carré du modèle. Dans un autre registre, Mint El Mouvid (2000) a utilisé les polynômes locaux pour estimer la fonction de répartition conditionnelle. Elle retrouve les mêmes propriétés avantageuses que présente la régression polynômiale locale et obtient le comportement asymptotique local et global de son estimateur.

Ayant fait ses preuves sur le plan théorique, la régression polynômiale locale est maintenant utilisée dans de nombreux domaines où l'on applique la statistique comme en témoignent, entre autres, les travaux de Fan et Gijbels (1996), Fan *et al.* (1998) et Yu et Jones (1998).

1.2 L'estimateur polynômial local

Soient (X_i, Y_i), $i = 1 \dots n$, n copies indépendantes de (X, Y), de fonction de répartition conjointe $F(x, y) = P[X \leq x, Y \leq y]$. Nous allons présenter l'estimateur polynômial local de :

$$\begin{aligned} F(y|x) &= P(Y \leq y|X = x), & (1.5)\\ &= E\left(I_{\{Y \leq y\}}|X = x\right), & (1.6) \end{aligned}$$

où $I_{\{A\}}$ dénote la fonction indicatrice de l'ensemble A. L'idée sous jacente à l'estimation polynômiale locale est en fait un simple développement de Taylor. Considérons x et y fixés et supposons que la fonction qui à z associe $F(y|z)$ est suffisamment régulière dans un voisinage de x. Alors, par une approximation de Taylor,

$$F(y|z) \approx F(y|x) + (z-x)F^{(1)}(y|x) + \ldots + (z-x)^p \frac{F^{(p)}(y|x)}{p!}, \tag{1.7}$$

où $F^{(\nu)}(y|x)$ est la $\nu^{\text{ième}}$ dérivée de $F(y|z)$ par rapport à z, évaluée en x. Ce développement permet d'approximer localement la fonction $F(y|z)$ par un polynôme de degré p en $z - x$, et donc en z, sans requérir que cette fonction appartienne à un modèle paramétrique quelconque. Rappelons aussi que dans ce contexte, la fonction $F(y|z)$ minimise en $m(y|z)$ l'expression :

$$E\left(\left(I_{\{Y\leq y\}} - m(y|z)\right)^2 |Z = z\right). \tag{1.8}$$

L'expression (1.8) et l'approximation (1.7) nous amènent au problème des moindres carrés pondérés suivant :

$$\text{trouver} \quad \hat{\beta} = \arg\min_\beta \sum_{i=1}^{n} \left[I_{\{Y_i\leq y\}} - \sum_{\nu=0}^{p} \beta_\nu (X_i - x)^\nu\right]^2 K\left(\frac{X_i - x}{h}\right), \tag{1.9}$$

où β est le vecteur de dimension $(p+1)$ dont les composantes sont les β_ν pour $\nu = 0, \ldots, p$. Dans cette expression, $K(.)$ est une fonction, appelée le noyau d'ajustement qui est supposée pour le moment positive et d'intégrale finie. Elle est utilisée pour diriger la minimisation de (1.9) en mettant un poids plus important aux données X_i proches de x. La quantité $h > 0$ est appelée la fenêtre d'ajustement et détermine la largeur du voisinage considéré autour de x.

Pour expliciter la solution $\hat{\beta}$ de (1.9), nous introduisons la notation matricielle suivante. Soient,

$$X_{n\times(p+1)} = \begin{pmatrix} 1 & (X_1 - x) & \ldots & (X_1 - x)^p \\ 1 & (X_2 - x) & \ldots & (X_2 - x)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (X_n - x) & \ldots & (X_n - x)^p \end{pmatrix}, \tag{1.10}$$

$$P = diag\left(K\left(\frac{X_i - x}{h}\right)\right)_{n\times n} = \begin{pmatrix} K\left(\frac{X_1 - x}{h}\right) & & 0 \\ & \ddots & \\ 0 & & K\left(\frac{X_n - x}{h}\right) \end{pmatrix}, \tag{1.11}$$

et

$$I_{\{Y \leq y\}_{n\times 1}} = \begin{pmatrix} I_{\{Y_1 \leq y\}} \\ I_{\{Y_2 \leq y\}} \\ \vdots \\ I_{\{Y_n \leq y\}} \end{pmatrix}. \tag{1.12}$$

En termes matriciels, le problème de minimisation (1.9) est équivalent à trouver $\hat{\beta}$, le vecteur de dimension $(p+1)$ qui minimise en β la fonction

$$\Big(I_{\{Y \leq y\}} - X\beta\Big)^T P \Big(I_{\{Y \leq y\}} - X\beta\Big), \tag{1.13}$$

où ^{T} désigne la transposée d'un vecteur ou d'une matrice. Par les résultats classiques de la méthode des moindres carrés pondérés, en supposant que la matrice $X^T P X$ soit inversible, la solution du système (1.13) est donnée par (Mint El Mouvid, 2000, p22) :

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = (X^T P X)^{-1} X^T P I_{\{Y \leq y\}}. \tag{1.14}$$

En comparant (1.7) et (1.9), on constate que $\hat{\beta}_\nu$ est un estimateur de $\dfrac{F^{(\nu)}(y|x)}{\nu!}$ et, en particulier, $\hat{\beta}_0$ estime $F(y|x)$. Cependant, dans ce travail, nous nous intéressons à l'estimation de $F(y|x)$ dans le but de construire un test de H_0 donnée en (1.3). Nous nous limiterons donc ici au cas où $\nu = 0$. Dénotons par $\hat{F}_n(y|x) = \hat{\beta}_0$ l'estimateur polynômial local de cette fonction de répartition conditionnelle. On va maintenant exprimer $\hat{\beta}_0$ de façon plus explicite. Pour ce faire, soit e_0, le vecteur de dimension $(p+1)$ ayant un "1" à la première composante et des 0 partout ailleurs. Des calculs simples d'algèbre linéaire montrent que (Mint El Mouvid, 2000, p22) :

$$\hat{F}_n(y|x) = \hat{\beta}_0 = e_0^T \hat{\beta} = \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) I_{\{Y_i \leq y\}}, \tag{1.15}$$

avec

$$W^n(t) = e_0^T (X^T P X)^{-1} (1, ht, \ldots, (ht)^p)^T K(t). \tag{1.16}$$

Si on pose $p = 0$, l'estimateur $\hat{\beta}_0$ de (1.15) coïncide avec celui étudié par Collomb (1980), qui s'inspire à son tour de l'estimateur de la fonction de régression de Nadaraya-Watson (Nadaraya, 1964 et Watson, 1964) que l'on appelle parfois l'"ajustement constant". Pour

$p = 1$, on obtient l'estimateur linéaire local étudié par Ducharme et Mint El Mouvid (2001).

La fonction de poids $W^n(.)$, définie en (1.16), dépend de l'échantillon $(X_1, \ldots, X_n)$ et satisfait (Fan et Gijbels, 1996, p103),

$$\sum_{i=1}^{n}(X_i - x)^k W^n\left(\frac{X_i - x}{h}\right) = \begin{cases} 1 & \text{si } k = 0, \\ 0 & \text{si } 0 < k \leq p. \end{cases} \tag{1.17}$$

Cette fonction est difficile à manipuler mathématiquement puisque sa forme fonctionnelle dépend des X_i. Notre premier travail consiste à contourner ce problème en remplaçant $W^n(.)$ par une approximation plus "stable". Avant de se lancer dans cette approximation, on propose d'introduire un certain nombre de notations et d'hypothèses, classiques pour la plupart en estimation non paramétrique et qui vont nous être utiles pour obtenir l'approximation du noyau $W^n(.)$.

1.3 Notations et Hypothèses

1.3.1 Notations et hypothèses sur le noyau

On définit $K_h(.)$, appelé généralement "noyau de convolution" par :

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right),$$

où $x \in \mathbb{R}$ et $h > 0$. Dans cette expression, $K(.)$ est le noyau dit d'ajustement. On définit maintenant la matrice S de dimension $(p + 1) \times (p + 1)$ dont les éléments sont les $(\mu_{i+j-2})_{1 \leq i,j \leq p+1}$, moments du noyau $K(.)$ définis par :

$$\mu_k = \int_{\mathbb{R}} x^k K(x) dx, \ \forall k = 0, \ldots, 2p. \tag{1.18}$$

On suppose que :

- $(K.0)$ Le noyau $K(.)$ est une densité de probabilité, symétrique autour de 0, bornée, dérivable et de dérivée bornée. On suppose de plus qu'il est à support compact. Sans perte de généralité, nous posons ce support compact comme étant $[-1, 1]$. De plus, on suppose que la matrice S est inversible.

1.3.2 Hypothèses sur les fonctions

- $(X.0)$ La variable aléatoire X a un support compact $[c, d]$. Sans perte de généralité et pour des raisons de commodité, nous posons ce support compact comme étant $[-1, 2]$.

– $(f.0)$ La densité marginale $f(.)$ de X existe et est bornée pour tout $x \in [-1,2]$. Il existe donc des constantes m et M telles que $0 < m < f(x) < M$ pour tout $x \in [-1,2]$. Elle est aussi continûment dérivable dans un voisinage de tout $x \in]-1,2[$ et de dérivée bornée.

– $(F_0.y.1)$ La fonction qui à y associe $F_0(y|x)$ est continûment dérivable en y pour tout $x \in [-1,2]$ ce qui assure l'existence de la densité conditionnelle $f_0(y|x)$ de Y sachant $\{X = x\}$.

– $(F_0^{(k)}.x.1)$ La fonction qui à x associe $F_0(y|x)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage de tout $x \in]-1,2[$ pour tout $y \in \mathbb{R}$.

– $(F_0^{(k)}.x.2)$ On suppose $(F_0^{(k)}.x.1)$ et on suppose de plus que pour tout u dans un voisinage $V(x)$ de $x \in]-1,2[$ et pour tout $y \in \mathbb{R}$,

$$\sup_{u \in V(x)} \sup_{y \in \mathbb{R}} \left| F_0^{(k)}(y|u) \right| \leq M,$$

où $F_0^{(k)}(y|x)$ est la $k^{\text{ième}}$ dérivée de la fonction $F(y|x)$ par rapport à x.

A ce stade, il semble important de clarifier certaines notations qui vont être utilisées dans la suite. Tout d'abord, on dénotera par $E_F(X)$, l'espérance de X relative à la loi $F(.)$ de X, de densité $f(.)$. De la même manière, on fera la distinction entre les notations $E_{F_0}(.)$ et $E_{F_0}(.\,|X)$ qui désigneront respectivement l'espérance relative à la loi conjointe $F_0(.,.)$ et l'espérance relative à la loi conditionnelle $F_0(.\,|\,.)$. On utilisera ces mêmes notations pour le calcul de variances.

1.3.3 Hypothèses sur la fenêtre h

– $(H.0)$ On suppose que $h = h(n) \longrightarrow 0$ lorsque $n \longrightarrow +\infty$ et que $\forall n, h < 1$.

- $(H.1)$ On suppose $(H.0)$ et on suppose de plus que $nh \longrightarrow +\infty$, lorsque $n \longrightarrow +\infty$.

- $(H.2)$ On suppose $(H.0)$ et $(H.1)$ et on suppose de plus que $\dfrac{\sqrt{\log n}}{\sqrt{n}h^2} \longrightarrow 0$, lorsque $n \longrightarrow +\infty$.

- $(H.3)$ On suppose $(H.0)$, $(H.1)$ et $(H.2)$ et on suppose de plus que $nh^{2p+\frac{5}{2}} \longrightarrow 0$, lorsque $n \longrightarrow +\infty$.

Ces hypothèses étant mises en place, on peut maintenant s'intéresser à l'approximation du noyau $W^n(.)$.

1.4 Approximation du noyau $W^n(.)$

On revient maintenant à notre premier objectif qui consiste à approximer le noyau $W^n(.)$ par une fonction plus “stable”. Suite à son expression donnée en (1.16), on constate qu'en introduisant la matrice $H = diag(1, \ldots, h^p)$ de dimension $(p+1) \times (p+1)$, alors

$$\begin{aligned} W^n(t) &= e_0^T \left(X^T P X\right)^{-1} (1, ht, \ldots, (ht)^p)^T K(t), \\ &= \frac{1}{nh} e_0^T H \left(\frac{1}{nh} X^T P X\right)^{-1} H(1, t, \ldots, t^p)^T K(t), \\ &= \frac{1}{nh} e_0^T S_n^{-1}(x)(1, t, \ldots, t^p)^T K(t), \end{aligned} \tag{1.19}$$

où on a posé

$$S_n(x) = H^{-1} \left(\frac{1}{nh} X^T P X\right) H^{-1}. \tag{1.20}$$

Afin d'approximer cette expression de $W^n(.)$, définissons maintenant :

$$K^*(t) = e_0^T S^{-1} (1, t, \ldots, t^p)^T K(t), \tag{1.21}$$

où S est la matrice de dimension $(p+1) \times (p+1)$, dont les éléments sont les $(\mu_{i+j-2})_{1 \leq i,j \leq p+1}$ définis en (1.18). Sous l'hypothèse $(K.0)$, ces quantités sont finies et cette matrice est supposée inversible. Cette fonction $K^*(.)$, qui est appelée *le noyau équivalent*, satisfait (Fan et Gijbels, 1996, p104) la version "théorique" de (1.17),

$$\int_{\mathbb{R}} t^k K^*(t)dt = \begin{cases} 1 & \text{si } k = 0, \\ 0 & \text{si } 0 < k \leq p. \end{cases} \tag{1.22}$$

On s'intéresse à présent au lien existant entre $W^n(.)$ et $K^*(.)$. Ce lien passe par l'étude de la structure de la matrice $S_n(x)$ et en particulier par son approximation. Dénotons par $(S_{n,i+j-2}(x))_{1\leq i,j\leq p+1}$, les éléments de la matrice $S_n(x)$ avec :

$$S_{n,k}(x) = \frac{1}{nh}\sum_{i=1}^{n}\left(\frac{X_i - x}{h}\right)^k K\left(\frac{X_i - x}{h}\right),\ k = 0,\ldots,2p. \tag{1.23}$$

On commence par approximer chacun des termes $S_{n,k}(x)$, pour $k = 0,\ldots,2p$. On donne pour cela le lemme suivant, qui constitue une légère amélioration du résultat démontré par Huang et Fan (1999, p939) :

Lemme 1.4.1 *Soit* $[a,b] \in]-1,2[$. *Alors, sous les hypothèses* $(K.0)$, $(X.0)$, $(f.0)$, $(H.2)$, *on a que presque sûrement* :

$$\sup_{x\in[a,b]} |S_{n,k}(x) - f(x)\mu_k| = O(h). \tag{1.24}$$

Preuve Voir Annexe A. ♣

Pour expliquer les choses plus précisément et éviter tout risque de confusion dans la suite, on donne quelques précisions concernant la notation utilisée dans l'énoncé de ce lemme. Soit X, de loi $F(.)$ sur l'espace de probabilité $(\Omega, \Im, P)$ et $\{X_n, n \geq 1\}$, une suite de copies indépendantes de X définies sur le même espace de probabilité. Alors, $\forall w \in \Omega - \mathcal{N}$ où $P(\mathcal{N}) = 0$, on a que $S_{n,k}(x)$, en tant que fonction de w via $X_1(\omega),\ldots,X_n(\omega)$, satisfait (1.24). Dans la suite on rencontre fréquemment des énoncés du type du Lemme 1.4.1. Pour simplifier, on utilise la notation $O_{ps}(1)$ pour désigner une variable aléatoire, fonction de $X_1, X_2,\ldots,X_n$ qui est bornée pour presque toutes ces suites $X_1(\omega), X_2(\omega),\ldots,X_n(\omega),\ldots$, quand n est suffisamment grand. Aussi, on notera $O_{ps}(h^q) = h^q O_{ps}(1)$. Ainsi, le Lemme 1.4.1 se réécrit plus simplement en disant que sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$ et $(H.2)$, on a :

$$\sup_{x\in[a,b]} |S_{n,k}(x) - f(x)\mu_k| = O_{ps}(h). \tag{1.25}$$

On écrira aussi, encore plus simplement,

$$S_{n,k}(x) = f(x)\mu_k + O^x_{ps}(h), \tag{1.26}$$

où maintenant $O_{ps}^x(h) = hO_{ps}^x(1)$ et le terme $O_{ps}^x(1)$ désigne un processus stochastique en x, fonction de $X_1, \ldots, X_n$ qui est borné sur $x \in [a, b]$ avec probabilité 1 (on dira "presque sûrement borné" en $x \in [a, b]$). Il faudra cependant tenir compte de la structure plus complexe du terme $O_{ps}^x(h)$ dans (1.26) quand on utilisera cette notation.

Il importe de signaler que le sup dans (1.24) doit être pris sur un sous domaine de $]-1, 2[$ et qu'en aucun cas on ne peut avoir $a = -1$ ou $b = 2$ car sinon le résultat (1.24) ne tient plus. Cette remarque aura des répercussions plus loin (voir à la Section 1.5) sur le problème de tester (1.3). Maintenant, (1.24) permet de déduire que quand x est restreint à $[a, b]$, chacun des éléments de la matrice $S_n(x)$ peut s'écrire sous la forme :

$$S_{n,k}(x) = f(x)\mu_k + hb_{h,k}(x), \tag{1.27}$$

où les $b_{h,k}(x)$ sont des quantités aléatoires presque sûrement bornées en $x \in [a, b]$. Ainsi, on a l'écriture matricielle :

$$S_n(x) = f(x)\left(I + hB_h(x)S^{-1}\right)S, \tag{1.28}$$

où $B_h(.)$ est la matrice aléatoire de dimension $(p+1) \times (p+1)$ dont les éléments sont les $(b_{h,i+j-2}(x))_{1 \leq i,j \leq p+1} / f(x)$. Etant donnée l'hypothèse $(f.0)$, on en déduit directement que si h est suffisamment petit (n suffisamment grand) pour assurer que l'inverse de cette matrice existe,

$$S_n^{-1}(x) = \frac{1}{f(x)} S^{-1}\left(I + hB_h(x)S^{-1}\right)^{-1}, \tag{1.29}$$

où I dénote est la matrice identité de dimension $(p+1) \times (p+1)$. De plus, par un résultat classique donné dans Bilodeau et Brenner (1999, p13), on a :

$$\begin{aligned}
\left(I + hB_h(x)S^{-1}\right)^{-1} &= I - hB_h(x)S^{-1} + h^2\left(B_h(x)S^{-1}\right)^2\left(I + hB_h(x)S^{-1}\right)^{-1}, \\
&= I + h\left(-B_h(x)S^{-1} + h\left(B_h(x)S^{-1}\right)^2\left(I + hB_h(x)S^{-1}\right)^{-1}\right), \\
&= I + hC_h(x),
\end{aligned} \tag{1.30}$$

où $C_h(x)$ est la matrice de dimension $(p+1) \times (p+1)$ dont chaque élément est aléatoire et preque sûrement borné en $x \in [a, b]$. On en déduit donc l'approximation :

$$S_n^{-1}(x) = \frac{S^{-1}}{f(x)}(I + hC_h(x)). \tag{1.31}$$

Si on substitue cette approximation dans la définition de $W^n(.)$, donnée en (1.19), on obtient, pour n suffisamment grand :

$$\begin{aligned}
W^n(t) &= \frac{1}{nhf(x)} e_0^T S^{-1} \left(I + hC_h(x)\right)(1, t, \ldots, t^p)^T K(t), \\
&= \frac{1}{nhf(x)} e_0^T S^{-1}(1, t, \ldots, t^p)^T K(t) \\
&\quad + h\frac{1}{nhf(x)} e_0^T S^{-1} C_h(x)(1, t, \ldots, t^p)^T K(t), \\
&= \frac{K^*(t)}{nhf(x)} + h\frac{1}{nhf(x)} e_0^T D_h(x)(1, t, \ldots, t^p)^T K(t), \qquad (1.32)
\end{aligned}$$

où $D_h(x) = S^{-1}C_h(x)$ est une matrice aléatoire de dimension $(p+1) \times (p+1)$ dont chacun des éléments est presque sûrement borné en $x \in [a, b]$. Regardons à présent la structure du terme $e_0^T D_h(x)(1, t, \ldots, t^p)^T$. Si les $(d_{1j,h}(x))_{j=0,\ldots,p}$ définissent les variables aléatoires de la première ligne de la matrice $D_h(x)$, alors on peut écrire :

$$\begin{aligned}
e_0^T D_h(x)(1, t, \ldots, t^p)^T &= \sum_{j=0}^{p} t^j d_{1j,h}(x), \\
&= Z_n(t, x) \sum_{j=0}^{p} \alpha_j t^j, \\
&= Z_n(t, x) P(t), \qquad (1.33)
\end{aligned}$$

où $Z_n(t, x)$ est un processus stochastique presque sûrement borné en $x \in [a, b]$ et en $t \in [-1, 1]$ de sorte que l'on puisse écrire à l'instar de (1.26), $Z_n(t, x) = O_{ps}^{(t,x)}(1)$ et où $P(t)$ est un polynôme de degré p quelconque en t que l'on suppose cependant non nul pour tout $t \in [-1, 1]$. Pour plus de confort dans la suite, on choisit aussi ce polynôme comme étant symétrique, c'est à dire $P(t) = P(-t)$. Si maintenant on pose $\bar{K}(t) = P(t)K(t)$, on en déduit que :

$$\begin{aligned}
W^n(t) &= \frac{K^*(t)}{nhf(x)} + hZ_n(t, x)\frac{\bar{K}(t)}{nhf(x)}, \\
&= \tilde{W}^n(t, x, h) + hO_{ps}^{(t,x)}(1)\bar{W}^n(t, x, h), \qquad (1.34)
\end{aligned}$$

où on a posé

$$\tilde{W}^n(t, x, h) = \frac{K^*(t)}{nhf(x)}, \qquad (1.35)$$

$$\bar{W}^n(t,x,h) = \frac{\bar{K}(t)}{nhf(x)}. \tag{1.36}$$

Cette expression est plus facile à manipuler mathématiquement que (1.16) puisque les formes fonctionnelles de $K^*(.)$ et de $\bar{K}(t)$ ne dépendent pas des X_i et correspondent à des estimateurs de type “noyau”. Signalons que dans la suite on calculera l'estimateur polynômial local à partir du noyau (1.16) mais que ses propriétés mathématiques seront étudiées via l'expression (1.34) de son noyau. Cette relation (1.34) permet aussi de retrouver en cas particulier le résultat du Lemme 1.2.1 de Huang et Fan (1999, p930) qui annoncent que presque sûrement,

$$\sup_{t\,\in\,[-1,1]}\ \sup_{x\,\in\,[a,b]} \left| nhW^n(t) - \frac{K^*(t)}{f(x)} \right| = o(1). \tag{1.37}$$

Dans la suite, on aura besoin d'utiliser l'approximation (1.24) qui est plus fine par rapport au résultat (1.37). On s'intéresse à présent aux propriétés vérifiées par le noyau équivalent $K^*(.)$ de (1.21). Elles sont données dans la proposition suivante :

Proposition 1.4.2 *Sous l'hypothèse* (K.0), *le noyau équivalent* $K^*(.)$ *est symétrique autour de 0, borné, dérivable et de dérivée bornée. De plus, son support est le compact* $[-1,1]$.

Preuve Voir Annexe B.♣

Il est intéressant de signaler que si on se place dans le cas d'un ajustement linéaire local ($p=1$), les noyaux $K(t)$ et $K^*(t)$ sont identiques (Fan et Gijbels, 1996, p66). On peut aussi noter que les propriétés données dans la Proposition 1.4.2 ne sont plus toutes vérifiées si on se place dans un contexte d'estimation des dérivées de la fonction de répartition conditionnelle (Fan et Gijbels, 1996, p65).

On définit maintenant le produit de convolution du noyau $K^*(.)$ par lui même par :

$$K^{*^{(2)}}(.) = K^* * K^*(.) = \int_{\mathbb{R}} K^*(.-x)K^*(x)dx. \tag{1.38}$$

Sous l'hypothèse de la Proposition 1.4.2, $K^{*^{(2)}}(.)$ est symétrique et à support compact $[-2,2]$. Enfin, on aura besoin de définir $K^{*^{(4)}}(.)$ par :

$$K^{*^{(4)}}(.) = K^{*^{(2)}} * K^{*^{(2)}}(.) = \int_{\mathbb{R}} K^{*^{(2)}}(.-x)K^{*^{(2)}}(x)dx. \tag{1.39}$$

Les divers paramètres intervenant dans l'estimation de la fonction de répartition conditionnelle étant introduits, on peut à présent s'intéresser à la construction du test de (1.3).

1.5 Construction du test d'adéquation

Pour construire le test de (1.3), notre idée consiste à se donner une "distance" entre $\hat{F}_n(y|x)$, l'estimateur polynômial local de $F(y|x)$, et $F_0(y|x)$, la valeur de $F(y|x)$ supposée sous H_0. Cette distance sera notre statistique de test. Il existe de nombreuses mesures de distances applicables aux fonctions de répartition. Nous allons utiliser ici, à l'instar de Härdle et Mammen (1993), Stute et González-Manteiga (1996) ou bien encore Alcalá *et al.* (1999), une distance de type L^2. Nous proposons donc de travailler avec la distance dite de Cramér-von Mises (Cramér, 1928 et von Mises, 1941) généralisée pour construire notre test.

Pour présenter cette distance, on va d'abord considérer le cas non conditionnel d'observations $Y_1, \dots, Y_n$ $iid \sim F(.)$. Supposons que l'on veuille tester :

$$H_0 : F(.) = F_0(.).$$

Dans ce contexte, un estimateur non paramétrique de $F(y)$ est

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^{n} I_{\{Y_i \leq y\}},$$

la fonction de répartition empirique. La "distance" de Cramér-von Mises entre $\hat{F}_n(.)$ et $F_0(.)$, aussi appelée statistique de Cramér-von Mises est donnée par

$$T = n \int_{\mathbb{R}} \left(\hat{F}_n(y) - F_0(y) \right)^2 F_0(dy). \tag{1.40}$$

La statistique de Cramér-von Mises généralisée (Anderson et Darling, 1954) est alors donnée par :

$$T = n \int_{\mathbb{R}} \left(\hat{F}_n(y) - F_0(y) \right)^2 w(y) F_0(dy), \tag{1.41}$$

où $w(y)$ est une fonction de poids satisfaisant certaines contraintes qu'on va préciser ultérieurement. Notons en particulier que si on pose $w(y) = 1$ dans (1.41), on retrouve la statistique (1.40). Si on choisit $w(y) = (F_0(y)(1 - F_0(y)))^{-1}$, on obtient la statistique dite d'Anderson-Darling (Anderson et Darling, 1954).

L'expression (1.41) suggère de considérer, pour le problème de tester (1.3), la statistique de test

$$n \int_{-1}^{2} \int_{\mathbb{R}} \left(\hat{F}_n(y|x) - F_0(y|x) \right)^2 w_x(y) F_0(dy|x)\ w(x) dx.$$

Mais, en regard des résultats du Lemme 1.4.1, il nous faut être un peu moins gourmand et considérer :

$$n \int_a^b \int_{\mathbb{R}} \left(\hat{F}_n(y|x) - F_0(y|x) \right)^2 w_x(y) F_0(dy|x) \; w(x)dx,$$

où $[a,b] \in]-1,2[$. Sans perte de généralité dans la suite, nous prendrons pour des raisons de commodité $[a,b] = [0,1]$. Ainsi, nous allons devoir nous contenter de tester une version "allégée" de (1.3) où (x,y) est restreint à $[0,1] \times \mathbb{R}$. Dans cette expression, $\hat{F}_n(y|x)$ représente l'estimateur polynômial local de $F(y|x)$. Il faut aussi, pour des raisons techniques, changer la constante n devant l'intégrale par la valeur $n\sqrt{h}$. Enfin, dans le but d'obtenir une classe de tests adaptés à diverses situations, nous introduisons dans cette statistique de test les fonctions de poids $w(x)$ et $w_x(y)$. Elles vérifient les conditions suivantes :

- $(P.0)$ La fonction qui à x associe $w(x)$ est positive et bornée dans un voisinage de tout $x \in [0,1]$.

- $(P.1)$ La fonction qui à (x,y) associe $w_x(y)$ est positive et bornée dans un voisinage de tout $x \in [0,1]$ et pour tout $y \in \mathbb{R}$.

Au final, sous ces hypothèses et la condition $(F_0.y.1)$, la statistique de test s'écrit sous la forme :

$$T = n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \left(\hat{F}_n(y|x) - F_0(y|x) \right)^2 w_x(y) f_0(y|x) dy \; w(x)dx. \tag{1.42}$$

Remarquons que si on pose $w(x) = \delta_x(x_0)$ pour un $x_0 \in [0,1]$, où $\delta_x(x_0)$ vaut 1 si $x = x_0$ et 0 sinon, on retrouve, à la constante $\sqrt{h}$ près, la forme de la statistique de Cramér-von Mises généralisée :

$$T_{x_0} = n\sqrt{h} \int_{\mathbb{R}} \left(\hat{F}_n(y|x_0) - F_0(y|x_0) \right)^2 w_{x_0}(y) F_0(dy|x_0), \tag{1.43}$$

pour le test de $H_0 : F(y|x_0) = F_0(y|x_0)$ en un point précis x_0, avec cependant la différence que $\hat{F}_n(y|x)$ n'est plus la fonction de répartition empirique mais l'estimateur polynômial local de $F(y|x)$. On voit donc que (1.42) peut se réécrire

$$T = \int_0^1 T_x w(x) dx$$

et constitue une moyenne de T_x sur une plage choisie des valeurs possibles de x. Par un choix approprié de $w(x)$, on peut mettre l'accent sur certaines valeurs ou plages de valeurs de x afin de détecter avec plus de puissance les écarts à la version "allégée" de (1.3) aux endroits qui paraîtraient à priori les plus intéressants à l'expérimentateur.

Dans les chapitres suivants, nous allons étudier le comportement asymptotique de la statistique de test (1.42). On se place d'abord dans le contexte où la fonction à tester sous H_0 est entièrement spécifiée. On reviendra ensuite au Chapitre 4 au cas de figure plus réaliste où celle-ci comporte un ou plusieurs paramètres inconnus.

Chapitre 2

Etude du comportement asymptotique de T sous H_0 dans le cas où $F_0(y|x)$ est un polynôme de degré inférieur ou égal à p en x

2.1 Introduction

Dans ce chapitre on va étudier le comportement asymptotique de la statistique de test T donnée en (1.42) pour tester l'hypothèse nulle,

$$H_0 : F(y|x) = F_0(y|x) \quad \forall y \in \mathbb{R}, \forall x \in [0,1]. \tag{2.1}$$

Dans leur article précurseur, Alcalá *et al.* (1999) se sont limités au cadre de la régression polynômiale locale et à l'étude d'une statistique de test ayant quelques relations avec la notre dans le cas où la fonction de régression à tester sous H_0 est un polynôme. Ceci peut être raisonnable dans un contexte de régression où les polynômes fournissent de bonnes approximations à bien des fonctions, mais l'est moins dans un contexte d'estimation de $F(y|x)$ qui doit être une fonction croissante en y. Cependant, pour faciliter le lien avec les travaux d'Alcalá *et al.* (1999), nous allons d'abord traiter ce cas où $F_0(y|x)$ est un polynôme de degré $\leq p$ en x. De plus, le comportement asymptotique de T est plus facile à obtenir dans ce cadre. Le cas où $F_0(y|x)$ n'est pas un polynôme découle partiellement des résultats que nous donnons ici et sera traité au Chapitre 3.

2.2 Décomposition de la statistique de test

Afin d'étudier au mieux le comportement asymptotique de la statistique de test donnée en (1.42), on propose d'abord une décomposition de celle-ci. Comme la fonction qui à x

associe $F_0(y|x)$ est un polynôme de degré inférieur ou égal à p, elle est p fois dérivable en x pour tout $y \in \mathbb{R}$. Pour tout $i \in \{1, \dots, n\}$, un développement de Taylor donne :

$$F_0(y|X_i) = \sum_{j=0}^{p} C_x^j(y)(X_i - x)^j, \tag{2.2}$$

avec

$$C_x^j(y) = \frac{F_0^{(j)}(y|x)}{j!},$$

et où certains termes $F_0^{(j)}(y|x)$ peuvent être nuls si le degré de $F_0(y|x)$ est inférieur à p. Ainsi,

$$\begin{aligned}
\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) F_0(y|X_i) &= \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) \left[\sum_{j=0}^{p} C_x^j(y)(X_i - x)^j\right], \\
&= \sum_{j=0}^{p} C_x^j(y) \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) (X_i - x)^j, \\
&= C_x^0(y), \\
&= F_0(y|x),
\end{aligned} \tag{2.3}$$

en utilisant la propriété (1.17). Par (1.15) et (2.3) on a de plus,

$$\begin{aligned}
\hat{F}_n(y|x) - F_0(y|x) &= \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) I_{\{Y_i \le y\}} - \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) F_0(y|X_i), \\
&= \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right) \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right),
\end{aligned} \tag{2.4}$$

où le noyau $W^n(.)$ est défini en (1.16). Suite aux remarques faites à la Section 1.2.1 concernant la structure complexe de ce noyau, on va pour le moment travailler avec un noyau plus "stable" et de ce fait, on considère pour le moment la quantité

$$\sum_{i=1}^{n} \tilde{W}^n(X_i, x, h) \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right), \tag{2.5}$$

où on signale que

$$\tilde{W}^n(X_i, x, h) = \frac{K^*\left(\frac{X_i - x}{h}\right)}{nhf(x)}, \tag{2.6}$$

où $K^*(.)$ est le *noyau équivalent* introduit en (1.21) et dont la "structure" de type noyau classique va permettre de "manipuler" plus facilement la quantité donnée en (2.5).

Si maintenant on élève (2.5) au carré et qu'on l'intègre de la même façon que la statistique de test (1.42), on obtient une nouvelle statistique

$$\begin{aligned}\tilde{T} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left(\sum_{i=1}^n \tilde{W}^n(X_i,x,h)\left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)\right)^2 w(x)w_x(y)f_0(y|x)dydx,\\ &= T_1+T_2, \qquad (2.7)\end{aligned}$$

où

$$\begin{aligned}T_1 &= n\sqrt{h}\sum_{i=1}^n\int_0^1\int_{\mathbb{R}}\tilde{W}^n(X_i,x,h)^2\left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)^2\\ &\quad\times w(x)w_x(y)f_0(y|x)dydx, \qquad (2.8)\end{aligned}$$

et

$$\begin{aligned}T_2 &= n\sqrt{h}\sum_{1\leq i\neq j\leq n}\int_0^1\int_{\mathbb{R}}\tilde{W}^n(X_i,x,h)\tilde{W}^n(X_j,x,h)\left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)\\ &\quad\times\left(I_{\{Y_j\leq y\}}-F_0(y|X_j)\right)w(x)w_x(y)f_0(y|x)dydx. \qquad (2.9)\end{aligned}$$

On va s'intéresser pour le moment au comportement asymptotique de cette nouvelle statistique car celui-ci est beaucoup plus simple à obtenir que celui de la statistique de test initiale donnée en (1.42). On montrera ultérieurement (à la Section 2.5) que leur comportement asymptotique est identique.

La décomposition (2.7) montre que le comportement asymptotique de la statistique $\tilde{T}$ va dépendre du comportement asymptotique des statistiques T_1 et T_2 qui s'écrivent aussi sous la forme :

$$T_1 = n\sqrt{h}\sum_{i=1}^n u_i, \qquad (2.10)$$

et

$$T_2 = \sum_{1\leq i,j\leq n} k_{ijn}, \qquad (2.11)$$

où u_i est la variable aléatoire fonction de (X_i, Y_i), de n et de h donnée par

$$u_i = \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_i, x, h)^2 \left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right)^2 w(x) w_x(y) f_0(y|x) dy dx, \tag{2.12}$$

et où k_{ijn} est la variable aléatoire fonction de (X_i, Y_i), (X_j, Y_j), de n et de h donnée par

$$k_{ijn} = \begin{cases} 0 & \text{si } i = j, \\ n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_i, x, h) \tilde{W}^n(X_j, x, h) \left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right) & \\ \times \left(I_{\{Y_j \leq y\}} - F_0(y|X_j)\right) w(x) w_x(y) f_0(y|x) dy dx & \text{si } i \neq j. \end{cases} \tag{2.13}$$

Nous allons à présent étudier le comportement asymptotique de chacun de ces deux termes individuellement.

2.3 Etude de T_1

Calculons d'abord $E_{F_0}(T_1)$, l'espérance de T_1 relative à la loi conjointe $F_0(.,.)$. Nous montrerons ensuite que $T_1 - E_{F_0}(T_1)$ converge en probabilité vers 0.

Proposition 2.3.1 *Supposons que $F_0(y|x)$ vérifie les hypothèses $(F_0.y.1)$ et $(F_0^{(1)}.x.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(H.0)$, $(P.0)$ et $(P.1)$, on a :*

$$E_{F_0}(T_1) = h^{-\frac{1}{2}} a_0 + O\left(h^{\frac{1}{2}}\right), \tag{2.14}$$

où

$$a_0 = K^{*(2)}(0) \int_0^1 \int_{\mathbb{R}} \frac{F_0(y|x)(1 - F_0(y|x))}{f(x)} w(x) w_x(y) f_0(y|x) dy dx. \tag{2.15}$$

Preuve Signalons que comme $F_0(y|x)$ est un polynôme de degré $\leq p$ en x, la condition $(F_0^{(1)}.x.1)$ est automatiquement satisfaite. Sous les hypothèses $(X.0)$, $(f.0)$, $(F_0.y.1)$ et par le théorème de Fubini, on obtient, partant de (2.10),

$$\begin{aligned} E_{F_0}(T_1) &= n^2 \sqrt{h} E_{F_0}(u_1) \\ &= n^2 \sqrt{h} \int_{\mathbb{R}} \int_0^1 \int_{-1}^2 \tilde{W}^n(x_1, x, h)^2 \left(\int_{\mathbb{R}} \left(I_{\{y_1 \leq y\}} - F_0(y|x_1)\right)^2 f_0(y_1|x_1) dy_1\right) \end{aligned}$$

$$\times\; w(x)w_x(y)f(x_1)dx_1 f_0(y|x)dxdy, \tag{2.16}$$

puisque les u_i de (2.12) sont *iid*. Or, $\forall y \in \mathbb{R}$ et $\forall x_1 \in [-1,2]$,

$$\int_{\mathbb{R}} \left(I_{\{y_1 \le y\}} - F_0(y|x_1)\right)^2 f_0(y_1|x_1)dy_1 = F_0(y|x_1)\left(1 - F_0(y|x_1)\right). \tag{2.17}$$

En substituant (2.17) dans (2.16) et en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6), on trouve

$$\begin{aligned} E_{F_0}(T_1) &= h^{-\frac{3}{2}} \int_{\mathbb{R}} \int_0^1 \int_{-1}^2 K^*\left(\frac{x_1 - x}{h}\right)^2 F_0(y|x_1)\left(1 - F_0(y|x_1)\right) f(x_1)dx_1 \\ &\quad \times f^{-2}(x)w(x)w_x(y)f_0(y|x)dydx. \end{aligned} \tag{2.18}$$

On pose à présent le changement de variable $x_1 = uh + x$. On obtient :

$$\begin{aligned} E_{F_0}(T_1) &= h^{-\frac{1}{2}} \int_{\mathbb{R}} \int_0^1 \left[\int_{\frac{-1-x}{h}}^{\frac{2-x}{h}} K^*(u)^2 F_0(y|uh+x)\left(1 - F_0(y|uh+x)\right)\right. \\ &\quad \left. \times f(uh+x)du\right] \times f^{-2}(x)w(x)w_x(y)f_0(y|x)dydx. \end{aligned} \tag{2.19}$$

Si on pose,

$$g_y(t) = F_0(y|t)\left(1 - F_0(y|t)\right) f(t),$$

l'expression (2.19) s'écrit, puisque $h < 1$, $x \in [0,1]$ et $K^*(.)$ a pour support $[-1,1]$,

$$\begin{aligned} E_{F_0}(T_1) &= h^{-\frac{1}{2}} \int_{\mathbb{R}} \int_0^1 \left[\int_{-1}^1 K^*(u)^2 g_y(uh+x)du\right] f^{-2}(x) \\ &\quad \times w(x)w_x(y)f_0(y|x)dydx. \end{aligned} \tag{2.20}$$

Puisque $x \in [0,1]$, la fonction qui à t associe $g_y(t+x)$ est continue et dérivable sur $t \in (0, uh)$, pour tout $u \in [-1,1]$. Par le théorème de la moyenne, il existe $\alpha_1 \in (0, uh)$ tel que :

$$g_y(uh+x) = g_y(x) + uhg_y^{'}(\alpha_1 + x),$$

où $g_y^{'}(\alpha_1 + x)$ est la dérivée de la fonction $g_y(t+x)$ par rapport à t, évaluée au point $t = \alpha_1$. On en déduit que sous l'hypothèse $(K.0)$,

$$\int_{-1}^1 K^*(u)^2 g_y(uh+x)du \;=\; g_y(x)\int_{-1}^1 K^*(u)^2 du$$

$$
\begin{aligned}
& + h \int_{-1}^{1} K^*(u)^2 u g_y^{'}(\alpha_1 + x) du, \\
= \; & g_y(x) K^{*^{(2)}}(0) + h \int_{-1}^{1} K^*(u)^2 u g_y^{'}(\alpha_1 + x) du, \qquad (2.21)
\end{aligned}
$$

puisque selon la Proposition 1.4.2, $K^*(.)$ est symétrique sur $[-1, 1]$. On en déduit aussi que :

$$
\begin{aligned}
E_{F_0}(T_1) \; = \; & h^{-\frac{1}{2}} a_0 \\
& + h^{\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \int_{-1}^{1} u K^*(u)^2 g_y^{'}(\alpha_1 + x) f^{-2}(x) w(x) w_x(y) \\
& \times f_0(y|x) du dy dx. \qquad (2.22)
\end{aligned}
$$

Dans tout le reste de ce travail, M dénotera une constante majorante pouvant varier d'une ligne à l'autre. Maintenant, sous l'hypothèse $(f.0)$,

$$
g_y'(t) = F_0^{(1)}(y|t) \left(1 - 2F_0(y|t)\right) f(t) + F_0(y|t) \left(1 - F_0(y|t)\right) f'(t),
$$

de sorte que, sous l'hypothèse $(F_0^{(1)}.x.2)$ et uniformément en y, t,

$$
\begin{aligned}
|g_y'(t)| \; \leq \; & |F_0^{(1)}(y|t)| \times |1 - 2F_0(y|t)| \times f(t) + F_0(y|t) \left(1 - F_0(y|t)\right) \times |f'(t)|, \\
\leq \; & M.
\end{aligned}
$$

En particulier, comme $\alpha_1 + x$ est dans un voisinage de $x \in [0, 1]$, $|g_y'(\alpha_1 + x)| \leq M$, uniformément en u, y et x. On en déduit que sous les hypothèses supplémentaires $(K.0)$, $(X.0)$, $(P.0)$ et $(P.1)$,

$$
\begin{aligned}
& \left| h^{\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \int_{-1}^{1} u K^*(u)^2 g_y^{'}(\alpha_1 + x) f^{-2}(x) w(x) w_x(y) f_0(y|x) du dy dx \right| \\
\leq \; & h^{\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \int_{-1}^{1} |u| K^*(u)^2 |g_y^{'}(\alpha_1 + x)| f^{-2}(x) w(x) w_x(y) f_0(y|x) du dy dx, \\
\leq \; & M h^{\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \int_{-1}^{1} |u| K^*(u)^2 f_0(y|x) du dy dx, \\
\leq \; & M h^{\frac{1}{2}}. \qquad (2.23)
\end{aligned}
$$

On en déduit que sous l'hypothèse $(H.0)$,

$$E_{F_0}(T_1) \quad = \quad h^{-\frac{1}{2}}a_0 + O\left(h^{\frac{1}{2}}\right).\clubsuit$$

On va maintenant montrer que T_1 converge en probabilité vers son espérance sous H_0. Plus précisément, on a le résultat suivant, dans lequel $\xrightarrow{P}$ désigne la convergence en probabilité :

Proposition 2.3.2 *Sous les hypothèses de la Proposition* 2.3.1 *et sous l'hypothèse supplémentaire* $(H.1)$, *on a :*

$$T_1 - h^{-\frac{1}{2}}a_0 \xrightarrow{P} 0, \tag{2.24}$$

où a_0 *est donnée en* (2.15).

Preuve Par l'inégalité de Bienaymé-Tchebytchev,

$$\forall\ \epsilon > 0, P\left(|T_1 - E_{F_0}(T_1)| > \epsilon\right) \leq \frac{Var_{F_0}(T_1)}{\epsilon^2}.$$

Il faut donc calculer $Var_{F_0}(T_1)$, la variance relative à la loi conjointe $F_0(.\,,.)$, et montrer que celle-ci tend vers 0. Partant de (2.10), et comme les u_i de (2.12) sont iid, nous avons,

$$Var_{F_0}(T_1) = n^3h\left(E_{F_0}(u_1^2) - E_{F_0}^2(u_1)\right). \tag{2.25}$$

Mais comme

$$E_{F_0}^2(u_1) = \frac{1}{n^4h}E_{F_0}^2(T_1), \tag{2.26}$$

la Proposition 2.3.1 permet d'écrire,

$$\begin{aligned} E_{F_0}^2(u_1) &= \frac{1}{n^4h}\left(h^{-\frac{1}{2}}a_0 + O\left(h^{\frac{1}{2}}\right)\right)^2, \\ &= \frac{1}{n^4h}\left(h^{-1}a_0^2 + 2h^{-\frac{1}{2}}a_0O\left(h^{\frac{1}{2}}\right) + O(h)\right), \\ &= \frac{1}{n^4h}\left(O\left(h^{-1}\right) + O(1) + o(1)\right), \end{aligned} \tag{2.27}$$

de sorte que,

$$\begin{aligned} Var_{F_0}(T_1) &= n^3hE_{F_0}(u_1^2) - \frac{n^3h}{n^4h}\left(O\left(h^{-1}\right) + O(1) + o(1)\right), \\ &= n^3hE_{F_0}(u_1^2) + O\left(\frac{1}{nh}\right) + O\left(\frac{1}{n}\right) + o\left(\frac{1}{n}\right). \end{aligned} \tag{2.28}$$

Ainsi, le comportement asymptotique de $Var_{F_0}(T_1)$ dépend essentiellement de celui de $n^3hE_{F_0}(u_1^2)$. Or,

$$\begin{aligned} u_1^2 &= \int_0^1\int_0^1\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{W}^n(X_1,x,h)^2\tilde{W}^n(X_1,x',h)^2\Big(I_{\{Y_1\leq y\}}-F_0(y|X_1)\Big)^2 \\ &\quad \times \Big(I_{\{Y_1\leq y'\}}-F_0(y'|X_1)\Big)^2 w(x)w(x')w_x(y)w_{x'}(y')f_0(y|x)f_0(y'|x')dydy'dxdx', \end{aligned}$$

et, en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6) et le théorème de Fubini, on a,

$$\begin{aligned} 0\leq n^3hE_{F_0}(u_1^2) &= n^{-1}h^{-3}\int_{-1}^2\int_0^1\int_0^1\int_{\mathbb{R}}\int_{\mathbb{R}} K^*\left(\frac{x_1-x}{h}\right)^2 K^*\left(\frac{x_1-x'}{h}\right)^2 \\ &\quad \times\left[\int_{\mathbb{R}}\Big(I_{\{y_1\leq y\}}-F_0(y|x_1)\Big)^2\Big(I_{\{y_1\leq y'\}}-F_0(y'|x_1)\Big)^2\right. \\ &\quad \left.\times f_0(y_1|x_1)dy_1\right]\times f^{-2}(x)f^{-2}(x')w(x)w(x') \\ &\quad \times w_x(y)w_{x'}(y')f_0(y|x)f_0(y'|x')dydy'dxdx'f(x_1)dx_1. \end{aligned} \tag{2.29}$$

Maintenant, $\forall y,y'\in\mathbb{R}$, $\forall x_1\in[-1,2]$,

$$0\leq\int_{\mathbb{R}}\Big(I_{\{y_1\leq y\}}-F_0(y|x_1)\Big)^2\Big(I_{\{y_1\leq y'\}}-F_0(y'|x_1)\Big)^2 f_0(y_1|x_1)dy_1\leq M,$$

et sous les hypothèses $(P.0)$, $(P.1)$ et $(f.0)$, les fonctions de poids et la fonction $f(.)$ sont bornées. Ainsi,

$$\begin{aligned} n^3hE_{F_0}(u_1^2) &\leq Mn^{-1}h^{-3}\int_{-1}^2\int_0^1\int_0^1 K^*\left(\frac{x_1-x}{h}\right)^2 K^*\left(\frac{x_1-x'}{h}\right)^2 \\ &\quad\times\left(\int_{\mathbb{R}} f_0(y|x)dy\right)\left(\int_{\mathbb{R}} f_0(y'|x')dy'\right)dxdx'dx_1, \\ &= Mn^{-1}h^{-3}\int_{-1}^2\int_0^1\int_0^1 K^*\left(\frac{x_1-x}{h}\right)^2 K^*\left(\frac{x_1-x'}{h}\right)^2 dxdx'dx_1. \end{aligned}$$

On effectue à présent les changements de variables $x=-uh+x_1$ et $x'=-vh+x_1$ de Jacobien inverse h^2 pour obtenir, comme $h<1$ et $x_1\in[-1,2]$,

$$\begin{aligned}
n^3 h E_{F_0}(u_1^2) &\leq Mn^{-1}h^{-1}\int_{-1}^{2}\int_{\frac{x_1-1}{h}}^{\frac{x_1}{h}}\int_{\frac{x_1-1}{h}}^{\frac{x_1}{h}} K^*(u)^2 K^*(v)^2 dudvdx_1,\\
&\leq Mn^{-1}h^{-1}\int_{-1}^{2}\left(\int_{\mathbb{R}} K^*(u)^2 du\right)\left(\int_{\mathbb{R}} K^*(v)^2 dv\right)dx_1,\\
&\leq Mn^{-1}h^{-1}\left[K^{*^{(2)}}(0)\right]^2,\\
&\leq \frac{M}{nh} = o(1),
\end{aligned}$$

sous l'hypothèse $(H.1)$. En combinant ceci avec (2.28), on déduit alors que $T_1 - E_{F_0}(T_1) \xrightarrow{P} 0$ et donc que

$$T_1 - h^{-\frac{1}{2}} a_0 \xrightarrow{P} 0.\clubsuit \tag{2.30}$$

2.4 Etude de T_2

Au préalable, rappelons la définition d'une statistique de forme "quadratique".

Définition 2.4.1 (*de Jong*, 1987, *p*261) *Soient X_i, $i = 1, \ldots, n$, des variables aléatoires indépendantes de fonction de répartition F à valeurs dans $\mathbb{R}$. On définit par forme "quadratique" toute variable aléatoire de la forme :*

$$\begin{aligned}
T_n &= n^{-1}\sum_{1\leq i,j\leq n} Q_{ijn}(X_i, X_j),\\
&= n^{-1}\sum_{1\leq i,j\leq n} Q_{ijn},
\end{aligned}$$

où les $Q_{ijn}(.,.)$ sont des fonctions symétriques en leurs arguments (appelées "noyaux") et telles que :

$$\int_{\mathbb{R}}\int_{\mathbb{R}} Q_{ijn}^2(x,y)dF(x)\times dF(y) < +\infty.$$

De (2.11), T_2 s'écrit sous la forme

$$T_2 = \sum_{1\leq i,j\leq n} k_{ijn}, \tag{2.31}$$

où les k_{ijn} ont été définis en (2.13). Cette écriture permet de voir que T_2 est, à la constante n^{-1} près, de forme "quadratique". De plus, comme $k_{iin} = 0$ et $k_{ijn} = k_{jin}$, on a aussi l'écriture suivante :

$$\begin{aligned} T_2 &= \sum_{1 \leq i \neq j \leq n} k_{ijn}, \\ &= 2 \sum_{1 \leq i < j \leq n} k_{ijn}, \end{aligned} \tag{2.32}$$

On remarque que T_2 est aussi, à une constante près, une U-statistique de noyau k_{ijn}.

2.4.1 Le Théorème de de Jong (1987)

On a maintenant besoin de la définition et du théorème suivants, dûs à de Jong (1987).

Définition 2.4.2 (*de Jong*, 1987, *p*263) *Soient* (X_i, Y_i), $i = 1, \ldots, n$, *des couples iid de variables aléatoires de loi conjointe* F. *Considérons* T_n, *une forme "quadratique" de noyaux* Q_{ijn}. T_n *est dite "propre" ("clean") si*

$$E_F\left(Q_{ijn} | (X_i, Y_i)\right) = 0, \quad \forall \ 1 \leq i, j \leq n.$$

Cette définition permet d'énoncer à présent le théorème de de Jong (1987), pour lequel $\xrightarrow{\mathcal{L}}$ désigne la convergence en loi.

Théorème 2.4.3 (*de Jong*, 1987, *p*264) *Soit* T_n, *une forme "quadratique propre" de noyaux* Q_{ijn} *et de variance* $Var_F(T_n)$. *Si* T_n *vérifie les conditions :*

$$a) \qquad \frac{\max_{1 \leq i \leq n} \sum_{j=1}^{n} Var_F(Q_{ijn})}{Var_F(T_n)} \longrightarrow 0 \text{ quand } n \longrightarrow +\infty,$$

$$b) \qquad \frac{E_F(T_n^4)}{Var_F^2(T_n)} \longrightarrow 3 \text{ quand } n \longrightarrow +\infty,$$

alors

$$T_n \xrightarrow{\mathcal{L}} N(0, \sigma^2), \tag{2.33}$$

où

$$\sigma^2 = \lim_{n \to +\infty} Var_F(T_n).$$

Preuve Voir de Jong (1987, p269).♣

2.4.2 T_2, forme quadratique "propre"

Sous H_0, T_2 est "propre" au sens de la Définition 2.4.2. En effet, par le théorème de Fubini, et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned} E_{F_0}\left(k_{ijn}|(X_i,Y_i)\right) &= E_{F_0}\left(k_{12n}|(X_1,Y_1)\right) \\ &= n\sqrt{h}\int_{-1}^{2}\int_0^1\int_{\mathbb{R}}\tilde{W}^n(X_1,x,h)\tilde{W}^n(x_2,x,h)\left(I_{\{Y_1\le y\}}-F_0(y|X_1)\right) \\ &\quad\times\left(\int_{\mathbb{R}}\left(I_{\{y_2\le y\}}-F_0(y|x_2)\right)f_0(y_2|x_2)dy_2\right) \\ &\quad\times w(x)w_x(y)f_0(y|x)dydxf(x_2)dx_2. \end{aligned} \tag{2.34}$$

Or, $\forall y\in\mathbb{R}$ et $\forall x_2\in[-1,2]$,

$$\int_{\mathbb{R}}\left(I_{\{y_2\le y\}}-F_0(y|x_2)\right)f_0(y_2|x_2)dy_2=0, \tag{2.35}$$

de sorte que $E_{F_0}\left(k_{12n}|(X_1,Y_1)\right)=0$. De plus, $k_{11n}=0$ par construction ce qui permet de déduire que T_2 est propre.

2.4.3 Calcul de la variance de T_2

La variance asymptotique de T_2 est donnée dans la proposition suivante :

Proposition 2.4.4 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(H.0)$, $(P.0)$ et $(P.1)$, on a :*

$$\lim_{n\to+\infty}Var_{F_0}(T_2)=\sigma_0^2, \tag{2.36}$$

avec

$$\begin{aligned} \sigma_0^2 &= 2K^{*(4)}(0)\int_0^1\int_{\mathbb{R}^2}\left[\frac{F_0(y\wedge y'|x')-F_0(y|x')F_0(y'|x')}{f(x')}\right]^2 \\ &\quad\times w^2(x')w_{x'}(y)w_{x'}(y')f_0(y|x')f_0(y'|x')dydy'dx', \end{aligned} \tag{2.37}$$

où $y\wedge y'=\min\{y,y'\}$.

Preuve A partir de l'expression de T_2 donnée en (2.32),

$$\begin{aligned} Var_{F_0}(T_2) &= Var_{F_0}\left(2\sum_{1\leq i<j\leq n} k_{ijn}\right), \\ &= 4E_{F_0}\left(\left(\sum_{1\leq i<j\leq n} k_{ijn}\right)^2\right), \end{aligned} \tag{2.38}$$

puisque la statistique T_2 est "propre". De plus, de de Jong (1987, p266), on a

$$\begin{aligned} \left(\sum_{1\leq i<j\leq n} k_{ijn}\right)^2 &= \sum_{1\leq i<j\leq n} k_{ijn}^2 + 6\sum_{1\leq i<j<l\leq n} k_{ijn}k_{iln} \\ &\quad + 6\sum_{1\leq i<j<k<l\leq n} k_{ijn}k_{kln}, \end{aligned} \tag{2.39}$$

de sorte que

$$\begin{aligned} Var_{F_0}(T_2) &= 4\Bigg[\sum_{1\leq i<j\leq n} E_{F_0}(k_{ijn}^2) + 6\sum_{1\leq i<j<l\leq n} E_{F_0}(k_{ijn}k_{iln}) \\ &\quad + 6\sum_{1\leq i<j<k<l\leq n} E_{F_0}(k_{ijn}k_{kln})\Bigg]. \end{aligned} \tag{2.40}$$

Maintenant, comme k_{ijn} et k_{kln} sont indépendantes quand $1 \leq i < j < k < l \leq n$,

$$\begin{aligned} E_{F_0}(k_{ijn}k_{kln}) &= E_{F_0}(k_{ijn})E_{F_0}(k_{kln}), \\ &= 0. \end{aligned}$$

Aussi, si $1 \leq i < j < l \leq n$,

$$\begin{aligned} E_{F_0}(k_{ijn}k_{iln}) &= E_{F_0}\left(E_{F_0}(k_{ijn}k_{iln}|(X_i,Y_i))\right), \\ &= E_{F_0}\left(E_{F_0}(k_{ijn}|(X_i,Y_i))E_{F_0}(k_{iln}|(X_i,Y_i))\right), \\ &= 0. \end{aligned}$$

Enfin, si $1 \leq i < j \leq n$, $E_{F_0}(k_{ijn}^2) = E_{F_0}(k_{12n}^2)$. Ceci permet de déduire que

$$Var_{F_0}(T_2) = 2n(n-1)E_{F_0}(k_{12n}^2). \tag{2.41}$$

Par (2.13) on obtient,

$$k_{12n}^2 = n^2h\int_{\mathbb{R}^2}\int_0^1\int_0^1 \tilde{W}^n(X_1,x,h)\tilde{W}^n(X_2,x,h)\tilde{W}^n(X_1,x',h)\tilde{W}^n(X_2,x',h)$$

$$\times \left(I_{\{Y_1 \leq y\}} - F_0(y|X_1)\right)\left(I_{\{Y_1 \leq y'\}} - F_0(y'|X_1)\right)$$

$$\times \left(I_{\{Y_2 \leq y\}} - F_0(y|X_2)\right)\left(I_{\{Y_2 \leq y'\}} - F_0(y'|X_2)\right)$$

$$\times\, w(x)w_x(y)w(x')w_{x'}(y')f_0(y|x)f_0(y'|x')dxdx'dydy',$$

de sorte que, par le théorème de Fubini et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned} E_{F_0}(k_{12n}^2) &= n^2 h \int_{\mathbb{R}^2}\int_{-1}^{2}\int_{-1}^{2}\int_0^1\int_0^1 \tilde{W}^n(x_1,x,h)\tilde{W}^n(x_2,x,h)\tilde{W}^n(x_1,x',h)\tilde{W}^n(x_2,x',h) \\ &\quad \times \left(\int_{\mathbb{R}} \left(I_{\{y_1 \leq y\}} - F_0(y|x_1)\right)\left(I_{\{y_1 \leq y'\}} - F_0(y'|x_1)\right) f_0(y_1|x_1)dy_1\right) \\ &\quad \times \left(\int_{\mathbb{R}} \left(I_{\{y_2 \leq y\}} - F_0(y|x_2)\right)\left(I_{\{y_2 \leq y'\}} - F_0(y'|x_2)\right) f_0(y_2|x_2)dy_2\right) \\ &\quad \times\, w(x)w_x(y)w(x')w_{x'}(y')f_0(y|x)f_0(y'|x')dxdx' \\ &\quad \times\, f(x_1)f(x_2)dx_1dx_2dydy'. \end{aligned} \tag{2.42}$$

Or, $\forall y, y' \in \mathbb{R}$ et $\forall x_1 \in [-1, 2]$,

$$\begin{aligned} &\int_{\mathbb{R}} \left(I_{\{y_1 \leq y\}} - F_0(y|x_1)\right)\left(I_{\{y_1 \leq y'\}} - F_0(y'|x_1)\right) f_0(y_1|x_1)dy_1 \\ &= \int_{\mathbb{R}} I_{\{y_1 \leq y\}} I_{\{y_1 \leq y'\}} f_0(y_1|x_1)dy_1 - F_0(y|x_1)\int_{\mathbb{R}} I_{\{y_1 \leq y'\}} f_0(y_1|x_1)dy_1 \\ &\quad - F_0(y'|x_1)\int_{\mathbb{R}} I_{\{y_1 \leq y\}} f_0(y_1|x_1)dy_1 + F_0(y|x_1)F_0(y'|x_1), \\ &= F_0(y \wedge y'|x_1) - F_0(y|x_1)F_0(y'|x_1) = g_{y,y'}(x_1). \end{aligned} \tag{2.43}$$

De la même manière, $\forall y, y' \in \mathbb{R}$ et $\forall x_2 \in [-1, 2]$,

$$\int_{\mathbb{R}} \left(I_{\{y_2 \leq y\}} - F_0(y|x_2)\right)\left(I_{\{y_2 \leq y'\}} - F_0(y'|x_2)\right) f_0(y_2|x_2)dy_2 = g_{y,y'}(x_2). \tag{2.44}$$

En substituant (2.43) et (2.44) dans (2.42) et en utilisant la définition de $\tilde{W}^n(.,.,.)$, donnée en (2.6), on obtient :

$$\begin{aligned}
E_{F_0}(k_{12n}^2) &= n^{-2}h^{-3}\int_{\mathbb{R}^2}\int_{-1}^{2}\int_{-1}^{2}\int_0^1\int_0^1 K^*\left(\frac{x_1-x}{h}\right)K^*\left(\frac{x_2-x}{h}\right) \\
&\quad \times K^*\left(\frac{x_1-x'}{h}\right)K^*\left(\frac{x_2-x'}{h}\right)g_{y,y'}(x_1)g_{y,y'}(x_2)f^{-2}(x)f^{-2}(x') \\
&\quad \times w(x)w_x(y)w(x')w_{x'}(y')f_0(y|x)f_0(y'|x') \\
&\quad \times f(x_1)f(x_2)dxdx'dx_1dx_2dydy'. \qquad (2.45)
\end{aligned}$$

Pour simplifier les calculs qui vont suivre, posons :

$$\alpha_y(t) = f^{-2}(t)w(t)w_t(y)f_0(y|t). \qquad (2.46)$$

L'expression (2.45) devient alors :

$$\begin{aligned}
E_{F_0}(k_{12n}^2) &= n^{-2}h^{-3}\int_{\mathbb{R}^2}\int_{-1}^{2}\int_{-1}^{2}\int_0^1\int_0^1 K^*\left(\frac{x_1-x}{h}\right) \\
&\quad \times K^*\left(\frac{x_2-x}{h}\right)K^*\left(\frac{x_1-x'}{h}\right)K^*\left(\frac{x_2-x'}{h}\right)g_{y,y'}(x_1)g_{y,y'}(x_2) \\
&\quad \times \alpha_y(x)\alpha_{y'}(x')f(x_1)f(x_2)dxdx'dx_1dx_2dydy'. \qquad (2.47)
\end{aligned}$$

On effectue dans (2.47) les changements de variables $x_1 = uh + x'$ et $x_2 = vh + x'$, de Jacobien inverse h^2 pour obtenir, comme $h < 1$ et $x' \in [0,1]$,

$$\begin{aligned}
E_{F_0}(k_{12n}^2) &= n^{-2}h^{-1}\int_{\mathbb{R}^2}\int_0^1\int_0^1\int_{\frac{-1-x'}{h}}^{\frac{2-x'}{h}}\int_{\frac{-1-x'}{h}}^{\frac{2-x'}{h}} K^*\left(\frac{x'-x}{h}+u\right)K^*\left(\frac{x'-x}{h}+v\right) \\
&\quad \times K^*(u)K^*(v)g_{y,y'}(uh+x')g_{y,y'}(vh+x')\alpha_y(x)\alpha_{y'}(x') \\
&\quad \times f(uh+x')f(vh+x')dudvdxdx'dydy', \\
&= n^{-2}h^{-1}\int_{\mathbb{R}^2}\int_0^1\int_0^1\int_{-1}^{1}\int_{-1}^{1} K^*\left(\frac{x'-x}{h}+u\right)K^*\left(\frac{x'-x}{h}+v\right) \\
&\quad \times K^*(u)K^*(v)L_{y,y'}(uh+x')L_{y,y'}(vh+x') \\
&\quad \times \alpha_y(x)\alpha_{y'}(x')dudvdxdx'dydy', \qquad (2.48)
\end{aligned}$$

où on a posé

$$L_{y,y'}(.) = g_{y,y'}(.)f(.).$$

On effectue à nouveau un changement de variable dans (2.48) en posant $x = -th + x'$ pour obtenir :

$$\begin{aligned} E_{F_0}(k_{12n}^2) &= n^{-2}\int_{\mathbb{R}^2}\int_0^1\int_{\frac{x'-1}{h}}^{\frac{x'}{h}}\int_{-1}^1\int_{-1}^1 K^*(t+u)K^*(t+v) \\ &\quad \times K^*(u)K^*(v)L_{y,y'}(uh+x')L_{y,y'}(vh+x') \\ &\quad \times \alpha_y(-th+x')\alpha_{y'}(x')dudvdtdx'dydy'. \end{aligned} \tag{2.49}$$

En substituant le résultat (2.49) dans l'expression (2.41), on en déduit que :

$$Var_{F_0}(T_2) = \frac{(n-1)}{n}\sigma_h^2, \tag{2.50}$$

avec

$$\begin{aligned} \sigma_h^2 &= 2\int_{\mathbb{R}^2}\int_0^1\int_{\frac{x'-1}{h}}^{\frac{x'}{h}}\int_{-1}^1\int_{-1}^1 K^*(t+u)K^*(t+v)K^*(u)K^*(v)L_{y,y'}(uh+x') \\ &\quad \times L_{y,y'}(vh+x')\alpha_y(-th+x')\alpha_{y'}(x')dudvdtdx'dydy'. \end{aligned} \tag{2.51}$$

Maintenant, sous l'hypothèse $(f.0)$, la fonction qui à u associe $L_{y,y'}(uh+x')$ est positive et continue sur le compact $[-1,1]$. Ceci est dû à la continuité des fonctions $F_0(.|x')$ et $f(x')$ dans un voisinage de $x' \in [0,1]$. Cette fonction est donc bornée et atteint ses bornes. Pour chaque $y, y' \in \mathbb{R}$ et $x' \in [0,1]$, on peut donc trouver $u_1 \in [-1,1]$ tel que :

$$0 \le L_{y,y'}(uh+x') \le L_{y,y'}(u_1h+x'). \tag{2.52}$$

Par le même raisonnement, on peut donc trouver $v_1 \in [-1,1]$ tel que :

$$0 \le L_{y,y'}(vh+x') \le L_{y,y'}(v_1h+x'). \tag{2.53}$$

Enfin, sous les hypothèses $(f.0)$, $(P.0)$ et $(P.1)$, la fonction qui à t associe $\alpha_y(-th+x')$ est bornée. Pour chaque $y \in \mathbb{R}$ et $x' \in [0,1]$, on peut donc trouver t_1 tel que

$$0 \le \alpha_y(-th+x') \le \alpha_y(-t_1h+x'). \tag{2.54}$$

Ceci nous permet de déduire que :

$$|2K^*(t+u)K^*(t+v)K^*(u)K^*(v)L_{y,y'}(uh+x')L_{y,y'}(vh+x')\alpha_y(-th+x')\alpha_{y'}(x')|$$

$$\leq 2|K^*(t+u)K^*(t+v)K^*(u)K^*(v)|L_{y,y'}(u_1h+x')L_{y,y'}(v_1h+x')\alpha_y(-t_1h+x')\alpha_{y'}(x'). \quad (2.55)$$

Le membre de droite de (2.55) est intégrable par rapport à u, v, t, x', y et y'. En effet, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$,

$$\begin{aligned}
&\int_{\mathbb{R}^2}\int_0^1 \left(\int_{\frac{x'-1}{h}}^{\frac{x'}{h}}\int_{-1}^1\int_{-1}^1 2|K^*(t+u)K^*(t+v)K^*(u)K^*(v)|dudvdt\right) \\
&\times L_{y,y'}(u_1h+x')L_{y,y'}(v_1h+x')\alpha_y(-t_1h+x')\alpha_{y'}(x')dx'dydy' \\
\leq\ & 2\int_{\mathbb{R}^2}\int_0^1\left(\int_{\mathbb{R}}\left[|K^*|^{(2)}(t)\right]^2 dt\right) L_{y,y'}(u_1h+x')L_{y,y'}(v_1h+x') \\
&\times \alpha_y(-t_1h+x')\alpha_{y'}(x')dx'dydy', \\
=\ & 2|K^*|^{(4)}(0)\int_{\mathbb{R}^2}\int_0^1 L_{y,y'}(u_1h+x')L_{y,y'}(v_1h+x') \\
&\times \alpha_y(-t_1h+x')\alpha_{y'}(x')dx'dydy', \\
\leq\ & M\int_{\mathbb{R}^2}\int_0^1 L_{y,y'}(u_1h+x')L_{y,y'}(v_1h+x')\alpha_y(-t_1h+x')\alpha_{y'}(x')dx'dydy', \\
\leq\ & M\int_{\mathbb{R}^2}\int_0^1 \alpha_y(-t_1h+x')\alpha_{y'}(x')dx'dydy', \\
\leq\ & M\int_0^1\left(\int_{\mathbb{R}} f_0(y|-t_1h+x')dy\right)\left(\int_{\mathbb{R}} f_0(y'|x')dy'\right)dx', \\
\leq\ & M. \qquad (2.56)
\end{aligned}$$

On a aussi

$$\begin{aligned}
&\lim_{h\to 0} 2K^*(t+u)K^*(t+v)K^*(u)K^*(v)L_{y,y'}(uh+x')L_{y,y'}(vh+x')\alpha_y(-th+x')\alpha_{y'}(x') \\
&\quad = 2K^*(t+u)K^*(t+v)K^*(u)K^*(v)\left[L_{y,y'}(x')\right]^2\alpha_y(x')\alpha_{y'}(x'). \qquad (2.57)
\end{aligned}$$

Les résultats (2.56) et (2.57) nous permettent d'appliquer le théorème de convergence dominée pour obtenir :

$$\lim_{h\to 0}\sigma_h^2 = \lim_{h\to 0} 2\int_{\mathbb{R}^2}\int_0^1\int_{\mathbb{R}}\int_{-1}^1\int_{-1}^1 K^*(t+u)K^*(t+v)K^*(u)K^*(v)L_{y,y'}(uh+x')$$

$$
\begin{aligned}
& \times L_{y,y'}(vh+x')\alpha_y(-th+x')\alpha_{y'}(x')dudvdtdx'dydy', \\
= \quad & 2\int_{\mathbb{R}^2}\int_0^1\int_{\mathbb{R}}\int_{-1}^1\int_{-1}^1 K^*(t+u)K^*(t+v)K^*(u)K^*(v) \\
& \times \left[L_{y,y'}(x')\right]^2 \alpha_y(x')\alpha_{y'}(x')dudvdtdx'dydy', \\
= \quad & 2\int_{\mathbb{R}^2}\int_0^1\int_{\mathbb{R}}\left(\int_{-1}^1 K^*(t+u)K^*(u)du\right)\left(\int_{-1}^1 K^*(t+v)K^*(v)dv\right) \\
& \times \left[L_{y,y'}(x')\right]^2 \alpha_y(x')\alpha_{y'}(x')dtdx'dydy', \\
= \quad & 2\int_{\mathbb{R}^2}\int_0^1\left(\int_{\mathbb{R}}\left[K^{*^{(2)}}(t)\right]^2 dt\right)\left[L_{y,y'}(x')\right]^2 \alpha_y(x')\alpha_{y'}(x')dx'dydy', \\
= \quad & \sigma_0^2, \qquad (2.58)
\end{aligned}
$$

avec

$$
\begin{aligned}
\sigma_0^2 \quad = \quad & 2K^{*^{(4)}}(0)\int_0^1\int_{\mathbb{R}^2}\left[\frac{F_0(y\wedge y'|x')-F_0(y|x')F_0(y'|x')}{f(x')}\right]^2 \\
& \times w^2(x')w_{x'}(y)w_{x'}(y')f_0(y|x')f_0(y'|x')dydy'dx'. \qquad (2.59)
\end{aligned}
$$

On en déduit que :

$$
\begin{aligned}
\lim_{n\to+\infty} Var_{F_0}(T_2) \quad & = \quad \lim_{n\to+\infty}\frac{n-1}{n}\sigma_h^2, \\
& = \quad \sigma_0^2.\clubsuit \qquad (2.60)
\end{aligned}
$$

2.4.4 Vérification de la première hypothèse du Théorème de de Jong (1987)

Comme T_2 est "propre", $E_{F_0}(k_{ijn}) = 0 \quad \forall\ 1 \le i, j \le n$ et on en déduit que

$$
Var_{F_0}(k_{ijn}) = E_{F_0}(k_{ijn}^2). \qquad (2.61)
$$

De plus, des expressions (2.41) et (2.50), $E_{F_0}(k_{ijn}^2) = \dfrac{\sigma_h^2}{2n^2}, \quad \forall\ 1 \le i < j \le n$. Comme $k_{iin} = 0$, on en déduit que,

$$
\max_{1\le i\le n}\sum_{j=1}^n Var_{F_0}(k_{ijn}) \quad = \quad \sum_{j=2}^n Var_{F_0}(k_{1jn}),
$$

$$= \frac{(n-1)\sigma_h^2}{2n^2},$$

$$= \frac{1}{2n} Var(T_2),$$

de sorte que,

$$\lim_{n\to+\infty} \frac{\max_{1\leq i\leq n} \sum_{j=1}^{n} Var_{F_0}(k_{ijn})}{Var_{F_0}(T_2)} = \lim_{n\to+\infty} \frac{1}{2n} = 0.$$

La première hypothèse du Théorème de de Jong (1987) est donc bien vérifiée.

2.4.5 Vérification de la deuxième hypothèse du Théorème de de Jong (1987)

Calculons $E_{F_0}(T_2^4)$. En utilisant l'expression (2.32) de T_2 et la décomposition donnée par de Jong (1987, p266) du terme $\left(\sum_{1\leq i<j\leq n} k_{ijn}\right)^4$, on a,

$$\begin{aligned} E_{F_0}(T_2^4) &= E_{F_0}\left(\left(2\sum_{1\leq i<j\leq n} k_{ijn}\right)^4\right), \\ &= 16(G_1 + 6G_2 + 12G_3 + 24G_4 + 6G_5), \\ &= 16G_1 + 96G_2 + 192G_3 + 384G_4 + 96G_5, \end{aligned} \tag{2.62}$$

avec,

- $G_1 = \sum_{1\leq i<j\leq n} E_{F_0}(k_{ijn}^4)$,

- $G_2 = \sum_{1\leq i<j<k\leq n} \left(E_{F_0}(k_{ijn}^2 k_{ikn}^2) + E_{F_0}(k_{jin}^2 k_{jkn}^2) + E_{F_0}(k_{kin}^2 k_{kjn}^2)\right)$,

- $G_3 = \sum_{1\leq i<j<k\leq n} \left(E_{F_0}(k_{ijn}^2 k_{kjn} k_{kin}) + E_{F_0}(k_{ikn}^2 k_{jkn} k_{jin}) + E_{F_0}(k_{jkn}^2 k_{ijn} k_{ikn})\right)$,

- $G_4 = \sum_{1\leq i<j<k<l\leq n} (E_{F_0}(k_{ijn} k_{ikn} k_{ljn} k_{lkn}) + E_{F_0}(k_{ijn} k_{iln} k_{kjn} k_{kln}) + E_{F_0}(k_{ikn} k_{iln} k_{jkn} k_{jln}))$,

– $G_5 = \sum_{1\leq i<j<k<l\leq n}\Big(E_{F_0}(k_{ijn}^2 k_{kln}^2) + E_{F_0}(k_{ikn}^2 k_{jln}^2) + E_{F_0}(k_{iln}^2 k_{jkn}^2)\Big).$

On étudie à présent chacun des termes de (2.62). Le premier s'écrit,

$$\begin{aligned} 16G_1 &= 16 \sum_{1\leq i<j\leq n} E_{F_0}(k_{ijn}^4), \\ &= 8n(n-1)E_{F_0}(k_{12n}^4). \end{aligned}$$

On a maintenant besoin du lemme suivant :

Lemme 2.4.5 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(H.1)$, $(P.0)$ et $(P.1)$, on a :*

$$E_{F_0}(k_{12n}^4) = o(n^{-2}).$$

Preuve Voir l'Annexe C. ♣

On en déduit,

$$16G_1 = o(1). \tag{2.63}$$

Le deuxième terme de (2.62) est :

$$\begin{aligned} &96 \sum_{1\leq i<j<k\leq n} \Big(E_{F_0}(k_{12n}^2 k_{13n}^2) + E_{F_0}(k_{21n}^2 k_{23n}^2) + E_{F_0}(k_{31n}^2 k_{32n}^2)\Big) \\ &= 96 \times \frac{n(n-1)(n-2)}{3!}\Big(E_{F_0}(k_{12n}^2 k_{13n}^2) + E_{F_0}(k_{21n}^2 k_{23n}^2) + E_{F_0}(k_{31n}^2 k_{32n}^2)\Big), \\ &= 48n(n-1)(n-2)E_{F_0}(k_{12n}^2 k_{13n}^2), \end{aligned}$$

les trois espérances étant identiques puisque d'une part les k_{ijn} sont symétriques en leurs arguments et d'autre part les variables intervenant dans les k_{ijn} sont identiquement distribuées. On utilise le lemme suivant :

Lemme 2.4.6 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$, on a :*

$$E_{F_0}(k_{12n}^2 k_{13n}^2) = o(n^{-3}).$$

Preuve Voir l'Annexe D. ♣

On en déduit,

$$96G_2 = o(1).$$

Passons à l'étude du troisième terme de (2.62). Le résultat (3) de De Jong (1987, p267) annonce que :

$$|G_3| \leq G_2.$$

On en déduit,

$$192G_3 = o(1). \tag{2.64}$$

Pour l'avant dernier terme de (2.62), on a :

$$\begin{aligned}
& 384 \sum_{1 \leq i<j<k<l \leq n} \Big(E_{F_0}(k_{ijn}k_{ikn}k_{ljn}k_{lkn}) + E_{F_0}(k_{ijn}k_{iln}k_{kjn}k_{kln}) \\
& + E_{F_0}(k_{ikn}k_{iln}k_{jkn}k_{jln}) \Big) \\
= \;& 384 \times \frac{n(n-1)(n-2)(n-3)}{4!} \Big(E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n}) + E_{F_0}(k_{12n}k_{14n}k_{32n}k_{34n}) \\
& + E_{F_0}(k_{13n}k_{14n}k_{23n}k_{24n}) \Big), \\
= \;& 48n(n-1)(n-2)(n-3)E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n}),
\end{aligned}$$

pour les mêmes raisons que dans le cas précédent. On utilise ensuite le lemme suivant :

Lemme 2.4.7 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(H.0)$, $(P.0)$ et $(P.1)$, on a :*

$$E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n}) = o(n^{-4}).$$

Preuve Voir l'Annexe E. ♣

On en déduit alors que

$$384G_4 = o(1).$$

Il ne reste plus maintenant qu'à étudier le dernier terme de (2.62). Nous avons,

$$96G_5 \;=\; 96 \sum_{1 \leq i<j<k<l \leq n} \left(E_{F_0}(k_{ijn}^2 k_{kln}^2) + E_{F_0}(k_{ikn}^2 k_{jln}^2) + E_{F_0}(k_{iln}^2 k_{jkn}^2) \right),$$

$$= 96 \times \frac{n(n-1)(n-2)(n-3)}{4!} \times 3E_{F_0}(k_{12n}^2 k_{34n}^2),$$

$$= 12n(n-1)(n-2)(n-3)E_{F_0}^2(k_{12n}^2), \qquad (2.65)$$

car k_{12n} est indépendant de k_{34n}. Or, on a déjà montré en (2.41) que

$$E_{F_0}(k_{12n}^2) = \frac{Var_{F_0}(T_2)}{2n(n-1)}. \qquad (2.66)$$

Ainsi,

$$96G_5 = 3 \times \frac{n(n-1)(n-2)(n-3)}{n^2(n-1)^2} Var_{F_0}^2(T_2),$$

$$= 3Var_{F_0}^2(T_2)(1+o(1)).$$

Au total, on trouve que

$$E_{F_0}(T_2^4) = 3Var_{F_0}^2(T_2)(1+o(1)) + o(1).$$

Ainsi,

$$\frac{E_{F_0}(T_2^4)}{Var_{F_0}^2(T_2)} = \frac{3Var_{F_0}^2(T_2)(1+o(1))+o(1)}{Var_{F_0}^2(T_2)} \longrightarrow 3 \text{ quand } n \longrightarrow +\infty. \qquad (2.67)$$

Les deux hypothèses du Théorème de de Jong (1987) étant vérifiées, on a donc démontré le théorème suivant :

Théorème 2.4.8 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(H.1)$, $(P.0)$ et $(P.1)$, on a :*

$$T_2 \xrightarrow{\mathcal{L}} N(0, \sigma_0^2) \qquad (2.68)$$

avec

$$\sigma_0^2 = 2K^{*(4)}(0) \int_0^1 \int_{\mathbb{R}^2} \left[\frac{F_0(y \wedge y'|x') - F_0(y|x')F_0(y'|x')}{f(x')} \right]^2$$

$$\times\, w^2(x') w_{x'}(y) w_{x'}(y') f_0(y|x') f_0(y'|x') dy dy' dx'. \qquad (2.69)$$

Enfin, revenons à la statistique (2.7). On avait obtenu la décomposition,

$$\tilde{T} = T_1 + T_2.$$

En combinant les résultats (2.30) et (2.68), on obtient le corollaire suivant :

Corollaire 2.4.9 *Sous les hypothèses du Théorème 2.4.8 et sous l'hypothèse supplémentaire* $(F_0^{(1)}.x.2)$, *on a :*

$$\frac{\tilde{T} - h^{-\frac{1}{2}} a_0}{\sigma_0} \xrightarrow{\mathcal{L}} N(0,1) \tag{2.70}$$

où les quantités a_0 *et* σ_0^2 *sont définies respectivement en* (2.15) *et* (2.69).

2.5 Comportement asymptotique de la statistique de test

On va à présent faire le lien entre la statistique $\tilde{T}$ dont on vient d'obtenir le comportement asymptotique en (2.70) et notre statistique de test, donnée en (1.42). Suite à la définition de $\hat{F}_n(y|x)$ donnée en (1.15) et à partir de la décomposition du noyau $W^n(.)$, obtenue en (1.34), on a :

$$\begin{aligned}
\hat{F}_n(y|x) - F_0(y|x) &= \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right) \\
&= \sum_{i=1}^{n} \tilde{W}^n\left(X_i, x, h\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right) \\
&\quad + h \sum_{i=1}^{n} Z_n\left(\frac{X_i - x}{h}, x\right) \bar{W}^n(X_i, x, h)(I_{\{Y_i \leq y\}} - F_0(y|X_i)),
\end{aligned}$$

où le noyau $\tilde{W}^n(.,.,.)$ a été défini en (2.6) et où on signale que

$$\bar{W}^n(X_i, x, h) = \frac{\bar{K}\left(\dfrac{X_i - x}{h}\right)}{nhf(x)}. \tag{2.71}$$

On rappelle aussi que dans cette expression, $Z_n\left(\dfrac{X_i - x}{h}, x\right)$ est presque sûrement borné en $x \in [0,1]$ si n est assez grand. Maintenant, si on élève ce terme au carré, on obtient :

$$\begin{aligned}
\left[\hat{F}_n(y|x) - F_0(y|x)\right]^2 &= \left[\sum_{i=1}^{n} \tilde{W}^n(X_i, x, h)(I_{\{Y_i \leq y\}} - F_0(y|X_i))\right]^2 \\
&\quad + h^2 \left[\sum_{i=1}^{n} Z_n\left(\frac{X_i - x}{h}, x\right) \bar{W}^n(X_i, x, h)(I_{\{Y_i \leq y\}} - F_0(y|X_i))\right]^2
\end{aligned}$$

$$+ 2h \left[\sum_{i=1}^{n} \tilde{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i)) \right]$$
$$\times \left[\sum_{i=1}^{n} Z_n \left(\frac{X_i - x}{h}, x \right) \bar{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i)) \right]$$

qui, injecté dans l'expression (1.42) de la statistique de test donne :

$$T = \tilde{T} + h^2 \bar{T} + 2h \tilde{\bar{T}}, \tag{2.72}$$

où $\tilde{T}$ a été défini en (2.7) et où on a posé

$$\begin{aligned} \bar{T} &= n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \left(\sum_{i=1}^{n} Z_n \left(\frac{X_i - x}{h}, x \right) \bar{W}^n(X_i, x, h) \left(I_{\{Y_i \le y\}} - F_0(y|X_i) \right) \right)^2 \\ &\quad \times w(x) w_x(y) f_0(y|x) dy dx, \end{aligned} \tag{2.73}$$

et

$$\begin{aligned} \tilde{\bar{T}} &= n\sqrt{h} \sum_{i=1}^{n} \sum_{j=1}^{n} \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_i, x, h) Z_n \left(\frac{X_j - x}{h}, x \right) \bar{W}^n(X_j, x, h) \\ &\quad \times \left(I_{\{Y_i \le y\}} - F_0(y|X_i) \right) \left(I_{\{Y_j \le y\}} - F_0(y|X_j) \right) \\ &\quad \times w(x) w_x(y) f_0(y|x) dy dx. \end{aligned} \tag{2.74}$$

Commençons par étudier le comportement asymptotique de la statistique $\bar{T}$. Celle-ci se décompose à son tour sous la forme

$$\bar{T} = \bar{T}_1 + \bar{T}_2, \tag{2.75}$$

avec

$$\begin{aligned} \bar{T}_1 &= n\sqrt{h} \sum_{i=1}^{n} \int_0^1 \int_{\mathbb{R}} Z_n \left(\frac{X_i - x}{h}, x \right)^2 \bar{W}^n(X_i, x, h)^2 \left(I_{\{Y_i \le y\}} - F_0(y|X_i) \right)^2 \\ &\quad \times w(x) w_x(y) f_0(y|x) dy dx, \end{aligned} \tag{2.76}$$

et

$$\bar{T}_2 = n\sqrt{h} \sum_{1 \le i \ne j \le n} \int_0^1 \int_{\mathbb{R}} Z_n \left(\frac{X_i - x}{h}, x \right) Z_n \left(\frac{X_j - x}{h}, x \right) \bar{W}^n(X_i, x, h) \bar{W}^n(X_j, x, h)$$

$$
\begin{aligned}
&\times \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right)\left(I_{\{Y_j \le y\}} - F_0(y|X_j)\right) w(x)w_x(y)f_0(y|x)dydx, \\
= &\sum_{1 \le i \ne j \le n} l_{ijn}, \qquad (2.77)
\end{aligned}
$$

avec

$$
\begin{aligned}
l_{ijn} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_i - x}{h}, x\right) Z_n\left(\frac{X_j - x}{h}, x\right) \bar{W}^n(X_i, x, h)\bar{W}^n(X_j, x, h) \\
&\quad \times \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right)\left(I_{\{Y_j \le y\}} - F_0(y|X_j)\right) w(x)w_x(y)f_0(y|x)dydx. \qquad (2.78)
\end{aligned}
$$

Tout d'abord, puisque $|Z_n(.,.)| \le M$,

$$
\begin{aligned}
0 \le \bar{T}_1 &\le Mn\sqrt{h}\sum_{i=1}^{n}\int_0^1\int_{\mathbb{R}} \bar{W}^n(X_i, x, h)^2 \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right)^2 \\
&\quad \times w(x)w_x(y)f_0(y|x)dydx. \qquad (2.79)
\end{aligned}
$$

De plus, on peut recycler la preuve de la Proposition 2.3.2 pour montrer que

$$
\begin{aligned}
&n\sqrt{h}\sum_{i=1}^{n}\int_0^1\int_{\mathbb{R}} \bar{W}^n(X_i, x, h)^2 \left(I_{\{Y_i \le y\}} - F_0(y|X_i)\right)^2 w(x)w_x(y)f_0(y|x)dydx \\
= &\ O\left(h^{-\frac{1}{2}}\right) + o_p(1).
\end{aligned}
$$

En effet, cette statistique est analogue à la statistique T_1 introduite en (2.8). La seule différence avec celle-ci est la structure du noyau qui la caractérise, mais cette structure ne jouant aucun rôle dans la preuve de la Proposition 2.3.2, on peut conclure directement quant au résultat. On déduit alors que :

$$
\bar{T}_1 = O\left(h^{-\frac{1}{2}}\right) + o_p(1). \qquad (2.80)
$$

On s'intéresse à présent au comportement asymptotique de la statistique $\bar{T}_2$. Par l'inégalité de Bienaymé-Tchebytchev, on a :

$$\forall k > 0, P\left(\left|\bar{T}_2 - E_{F_0}\left(\bar{T}_2\right)\right| > k\right) \leq \frac{Var_{F_0}\left(\bar{T}_2\right)}{k^2}. \tag{2.81}$$

De plus,

$$\begin{aligned} E_{F_0}\left(\bar{T}_2\right) &= n(n-1)E_{F_0}(l_{12n}), \\ &= n(n-1)E_F\left(E_{F_0}(l_{12n}|X_1, X_2, \ldots, X_n)\right). \end{aligned} \tag{2.82}$$

Aussi,

$$\begin{aligned} E_{F_0}(l_{12n}|X_1, X_2, \ldots, X_n) &= n\sqrt{h}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_1 - x}{h}, x\right) Z_n\left(\frac{X_2 - x}{h}, x\right) \\ &\quad \times \bar{W}^n(X_1, x, h)\bar{W}^n(X_2, x, h) \\ &\quad \times \left(I_{\{y_1 \leq y\}} - F_0(y|X_1)\right)\left(I_{\{y_2 \leq y\}} - F_0(y|X_2)\right) \\ &\quad \times w(x)w_x(y)f_0(y|x)dydxf_0(y_1|X_1)dy_1 f_0(y_2|X_2)dy_2, \\ &= 0, \end{aligned} \tag{2.83}$$

car $\int_{\mathbb{R}}\left(I_{\{y_1 \leq y\}} - F_0(y|X_1)\right) f_0(y_1|X_1)dy_1 = 0, \forall y \in \mathbb{R}, \forall X_1 \in [-1, 2]$. On en déduit que $E_{F_0}(l_{12n}) = 0$ et donc

$$E_{F_0}\left(\bar{T}_2\right) = 0. \tag{2.84}$$

Maintenant, comme $E_{F_0}\left(\bar{T}_2|X_1, X_2, \ldots, X_n\right) = 0$,

$$\begin{aligned} Var_{F_0}\left(\bar{T}_2\right) &= E_F\left(Var_{F_0}\left(\bar{T}_2|X_1, X_2, \ldots, X_n\right)\right) + Var_F\left(E_{F_0}\left(\bar{T}_2|X_1, X_2, \ldots, X_n\right)\right), \\ &= E_F\left(Var_{F_0}\left(\bar{T}_2|X_1, X_2, \ldots, X_n\right)\right). \end{aligned} \tag{2.85}$$

Or, par le même raisonnement qu'en (2.41) appliqué ici au cas conditionnel, on a :

$$Var_{F_0}\left(\bar{T}_2|X_1, X_2, \ldots, X_n\right) = 2n(n-1)E_{F_0}\left(l_{12n}^2|X_1, X_2, \ldots, X_n\right). \tag{2.86}$$

De plus,

$$
\begin{aligned}
E_{F_0}\left(l_{12n}^2|X_1,X_2,\ldots,X_n\right) &= n^2h\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_1-x}{h},x\right)Z_n\left(\frac{X_1-x'}{h},x\right)\\
&\times Z_n\left(\frac{X_2-x}{h},x\right)Z_n\left(\frac{X_2-x'}{h},x\right)\\
&\times \bar{W}^n(X_1,x,h)\bar{W}^n(X_2,x,h)\bar{W}^n(X_1,x',h)\bar{W}^n(X_2,x',h)\\
&\times \left(I_{\{y_1\le y\}}-F_0(y|X_1)\right)\left(I_{\{y_2\le y\}}-F_0(y|X_2)\right)\\
&\times \left(I_{\{y_1\le y'\}}-F_0(y'|X_1)\right)\left(I_{\{y_2\le y'\}}-F_0(y'|X_2)\right)\\
&\times w(x)w_x(y)f_0(y|x)dydxw(x')w_{x'}(y')f_0(y'|x')dy'dx'\\
&\times f_0(y_1|X_1)dy_1f_0(y_2|X_2)dy_2,\\
&\le n^2h\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}|Z_n\left(\frac{X_1-x}{h},x\right)||Z_n\left(\frac{X_1-x'}{h},x\right)|\\
&\times |Z_n\left(\frac{X_2-x}{h},x\right)||Z_n\left(\frac{X_2-x'}{h},x\right)|\\
&\times |\bar{W}^n(X_1,x,h)||\bar{W}^n(X_2,x,h)||\bar{W}^n(X_1,x',h)||\bar{W}^n(X_2,x',h)|\\
&\times \left(\int_{\mathbb{R}}\left|I_{\{y_1\le y\}}-F_0(y|X_1)\right|\left|I_{\{y_1\le y'\}}-F_0(y'|X_1)\right|f_0(y_1|X_1)dy_1\right)\\
&\times \left(\int_{\mathbb{R}}\left|I_{\{y_2\le y\}}-F_0(y|X_2)\right|\left|I_{\{y_2\le y'\}}-F_0(y'|X_2)\right|f_0(y_2|X_2)dy_2\right)\\
&\times w(x)w_x(y)f_0(y|x)dydxw(x')w_{x'}(y')f_0(y'|x')dy'dx',\\
&\le Mn^2h\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}|\bar{W}^n(X_1,x,h)||\bar{W}^n(X_2,x,h)|\\
&\times |\bar{W}^n(X_1,x',h)||\bar{W}^n(X_2,x',h)|dxdx',
\end{aligned}
$$

$$\leq \frac{M}{n^2h^3}\int_0^1\int_0^1 |\bar{K}\left(\frac{X_1-x}{h}\right)||\bar{K}\left(\frac{X_1-x'}{h}\right)| \times |\bar{K}\left(\frac{X_2-x}{h}\right)||\bar{K}\left(\frac{X_2-x'}{h}\right)|dxdx', \tag{2.87}$$

en utilisant la définition de $\bar{W}^n(.,.,.)$ donnée en (2.71). De plus, si on pose les changements de variables $x=-uh+X_1$ et $x'=-vh+X_1$ de Jacobien $\frac{1}{h^2}$, on obtient, comme $h<1$ et $X_1\in[-1,2]$,

$$\begin{aligned}
E_{F_0}\left(l^2_{12n}|X_1,X_2,\ldots,X_n\right) &\leq \frac{M}{n^2h}\int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}}\int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}} |\bar{K}\left(\frac{X_2-X_1}{h}+u\right)||\bar{K}\left(u\right)| \\
&\quad\times |\bar{K}\left(\frac{X_2-X_1}{h}+v\right)||\bar{K}\left(v\right)|dudv, \\
&\leq \frac{M}{n^2h}\left(\int_{\mathbb{R}}|\bar{K}\left(\frac{X_2-X_1}{h}+u\right)||\bar{K}\left(u\right)|du\right) \\
&\quad\times\left(\int_{\mathbb{R}}|\bar{K}\left(\frac{X_2-X_1}{h}+v\right)||\bar{K}\left(v\right)|dv\right), \\
&\leq \frac{M}{n^2h}\left[|\bar{K}|^{(2)}\left(\frac{X_1-X_2}{h}\right)\right]^2.
\end{aligned}\tag{2.88}$$

On en déduit que sous l'hypothèse $(f.0)$,

$$\begin{aligned}
E_F\left(E_{F_0}\left(l^2_{12n}|X_1,\ldots,X_n\right)\right) &\leq E_F\left(\frac{M}{n^2h}\left[|\bar{K}|^{(2)}\left(\frac{X_1-X_2}{h}\right)\right]^2\right), \\
&= \frac{M}{n^2h}\int_{-1}^2\int_{-1}^2\left[|\bar{K}|^{(2)}\left(\frac{x_1-x_2}{h}\right)\right]^2 f(x_1)f(x_2)dx_1dx_2, \\
&\leq \frac{M}{n^2h}\int_{-1}^2\int_{-1}^2\left[|\bar{K}|^{(2)}\left(\frac{x_1-x_2}{h}\right)\right]^2 dx_1dx_2, \\
&= \frac{M}{n^2}\int_{-1}^2\int_{\frac{-1-x_2}{h}}^{\frac{2-x_2}{h}}\left[|\bar{K}|^{(2)}\left(u\right)\right]^2 dudx_2, \\
&\leq \frac{M}{n^2}\int_{-1}^2\int_{\mathbb{R}}\left[|\bar{K}|^{(2)}\left(u\right)\right]^2 dudx_2, \\
&\leq \frac{M}{n^2}.
\end{aligned}\tag{2.89}$$

On en déduit alors que

$$Var_{F_0}\left(\bar{T}_2\right) \leq 2n(n-1) \times \frac{M}{n^2} = O(1). \tag{2.90}$$

Si on substitue maintenant les résultats (2.84) et (2.90) dans l'inégalité (2.81), on en déduit que pour tout $k > 0$,

$$\begin{aligned} P\left(\left|\bar{T}_2\right| \leq k\right) &= 1 - P\left(\left|\bar{T}_2\right| > k\right), \\ &\geq 1 - \frac{M}{k^2}, \end{aligned}$$

ce qui implique que :

$$\bar{T}_2 = O_p(1). \tag{2.91}$$

On s'intéresse à présent au comportement asymptotique de la statistique $\tilde{\bar{T}}$ dont on rappelle l'expression :

$$\begin{aligned} \tilde{\bar{T}} &= n\sqrt{h}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_0^1\int_{\mathbb{R}} \tilde{W}^n(X_i,x,h)Z_n\left(\frac{X_j - x}{h}, x\right)\bar{W}^n(X_j,x,h) \\ &\times \left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right)\left(I_{\{Y_j \leq y\}} - F_0(y|X_j)\right) \\ &\times w(x)w_x(y)f_0(y|x)dydx. \end{aligned} \tag{2.92}$$

Cette statistique se décompose elle aussi en deux statistiques :

$$\tilde{\bar{T}} = \tilde{\bar{T}}_1 + \tilde{\bar{T}}_2, \tag{2.93}$$

avec

$$\begin{aligned} \tilde{\bar{T}}_1 &= n\sqrt{h}\sum_{i=1}^{n}\int_0^1\int_{\mathbb{R}} \tilde{W}^n(X_i,x,h)Z_n\left(\frac{X_i - x}{h}, x\right)\bar{W}^n(X_i,x,h)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right)^2 \\ &\times w(x)w_x(y)f_0(y|x)dydx, \end{aligned} \tag{2.94}$$

et

$$\begin{aligned} \tilde{\bar{T}}_2 &= n\sqrt{h}\sum_{1\leq i\neq j\leq n}\int_0^1\int_{\mathbb{R}} \tilde{W}^n(X_i,x,h)Z_n\left(\frac{X_j - x}{h}, x\right)\bar{W}^n(X_j,x,h) \\ &\times \left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right)\left(I_{\{Y_j \leq y\}} - F_0(y|X_j)\right) \\ &\times w(x)w_x(y)f_0(y|x)dydx, \\ &= \sum_{1\leq i\neq j\leq n} t_{ijn}, \end{aligned} \tag{2.95}$$

avec

$$\begin{aligned} t_{ijn} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}} \tilde{W}^n(X_i,x,h)Z_n\left(\frac{X_j-x}{h},x\right)\bar{W}^n(X_j,x,h) \\ &\times \left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)\left(I_{\{Y_j\leq y\}}-F_0(y|X_j)\right) \\ &\times w(x)w_x(y)f_0(y|x)dydx. \end{aligned} \tag{2.96}$$

Tout d'abord, comme $|Z_n(.,.)|\leq M$,

$$\begin{aligned} \left|\tilde{\bar{T}}_1\right| &\leq n\sqrt{h}\sum_{i=1}^n\int_0^1\int_{\mathbb{R}} |\tilde{W}^n(X_i,x,h)||Z_n\left(\frac{X_i-x}{h},x\right)||\bar{W}^n(X_i,x,h)| \\ &\times \left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)^2 w(x)w_x(y)f_0(y|x)dydx, \\ &\leq Mn\sqrt{h}\sum_{i=1}^n\int_0^1\int_{\mathbb{R}} |\tilde{\bar{W}}^n(X_i,x,h)|\left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)^2 \\ &\times w(x)w_x(y)f_0(y|x)dydx, \end{aligned} \tag{2.97}$$

où on a posé $\tilde{\bar{W}}^n(X_i,x,h)=\tilde{W}^n(X_i,x,h)\bar{W}^n(X_i,x,h)$. De plus, on peut recycler la preuve de la Proposition 2.3.2 pour montrer que

$$\begin{aligned} &n\sqrt{h}\sum_{i=1}^n\int_0^1\int_{\mathbb{R}} |\tilde{\bar{W}}^n(X_i,x,h)|\left(I_{\{Y_i\leq y\}}-F_0(y|X_i)\right)^2 w(x)w_x(y)f_0(y|x)dydx \\ &= O\left(h^{-\frac{1}{2}}\right)+o_p(1). \end{aligned}$$

En effet, cette statistique est quasiment analogue à la statistique T_1 donnée en (2.8). La seule différence avec celle-ci est la forme du "double" noyau $|\tilde{\bar{W}}^n(.,.,.)|$ qui la caractérise. Or cette structure ne joue aucun rôle dans la preuve de la Proposition 2.3.2 ce qui permet de conclure directement quant au résultat. On déduit alors que :

$$\tilde{\bar{T}}_1 = O\left(h^{-\frac{1}{2}}\right)+o_p(1). \tag{2.98}$$

On s'intéresse à présent au comportement asymptotique de la statistique $\tilde{\bar{T}}_2$. Contrairement à la statistique $\bar{T}_2$, cette statistique possède un noyau t_{ijn} qui n'est pas symétrique en ses arguments. On la symétrise en posant :

$$\tilde{\bar{T}}_2 = \sum_{1 \le i \ne j \le n} T_{ijn}, \tag{2.99}$$

où le nouveau noyau T_{ijn} est défini par :

$$T_{ijn} = \frac{1}{2}\left(t_{ijn} + t_{jin}\right), \tag{2.100}$$

et est symétrique en ses arguments, ce qui permet de proposer aussi l'écriture suivante pour $\tilde{\bar{T}}_2$:

$$\tilde{\bar{T}}_2 = 2 \sum_{1 \le i \le j \le n} T_{ijn}. \tag{2.101}$$

Maintenant,

$$\begin{aligned} E_{F_0}\left(\tilde{\bar{T}}_2\right) &= n(n-1)E_{F_0}(T_{12n}), \\ &= n(n-1)E_F\left(E_{F_0}(T_{12n}|X_1, X_2, \ldots, X_n)\right). \end{aligned} \tag{2.102}$$

De plus,

$$E_{F_0}(T_{12n}|X_1, X_2, \ldots, X_n) = \frac{1}{2}\left[E_{F_0}(t_{12n}|X_1, X_2, \ldots, X_n) + E_{F_0}(t_{21n}|X_1, X_2, \ldots, X_n)\right]. \tag{2.103}$$

Aussi,

$$\begin{aligned} E_{F_0}(t_{12n}|X_1, X_2, \ldots, X_n) &= n\sqrt{h}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} \tilde{W}^n(X_1, x, h) Z_n\left(\frac{X_2 - x}{h}, x\right) \bar{W}^n(X_2, x, h) \\ &\quad \times \left(I_{\{y_1 \le y\}} - F_0(y|X_1)\right)\left(I_{\{y_2 \le y\}} - F_0(y|X_2)\right) \\ &\quad \times w(x) w_x(y) f_0(y|x) dy dx f_0(y_1|X_1) dy_1 f_0(y_2|X_2) dy_2, \\ &= 0, \end{aligned} \tag{2.104}$$

car $\int_{\mathbb{R}} \left(I_{\{y_1 \leq y\}} - F_0(y|X_1)\right) f_0(y_1|X_1)dy_1 = 0$, $\forall\, y \in \mathbb{R}, \forall X_1 \in [-1, 2]$. De la même manière, on montre que $E_{F_0}\left(t_{21n}|X_1, X_2, \ldots, X_n\right) = 0$ ce qui permet de déduire :

$$E_{F_0}\left(\tilde{\bar{T}}_2\right) = 0. \tag{2.105}$$

Maintenant, comme $E_{F_0}\left(\tilde{\bar{T}}_2|X_1, X_2, \ldots, X_n\right) = 0$,

$$\begin{aligned} Var_{F_0}\left(\tilde{\bar{T}}_2\right) &= E_F\left(Var_{F_0}\left(\tilde{\bar{T}}_2|X_1, X_2, \ldots, X_n\right)\right) + Var_F\left(E_{F_0}\left(\tilde{\bar{T}}_2|X_1, X_2, \ldots, X_n\right)\right), \\ &= E_F\left(Var_{F_0}\left(\tilde{\bar{T}}_2|X_1, X_2, \ldots, X_n\right)\right). \end{aligned} \tag{2.106}$$

Or, par le même raisonnement qu'en (2.41) appliqué ici au cas conditionnel, on a :

$$Var_{F_0}\left(\tilde{\bar{T}}_2|X_1, X_2, \ldots, X_n\right) = 2n(n-1)E_{F_0}\left(T^2_{12n}|X_1, X_2, \ldots, X_n\right). \tag{2.107}$$

De plus,

$$\begin{aligned} E_{F_0}\left(T^2_{12n}|X_1, X_2, \ldots, X_n\right) &= \frac{1}{4}\Big[E_{F_0}\left(t^2_{12n}|X_1, X_2, \ldots, X_n\right) + E_{F_0}\left(t^2_{21n}|X_1, X_2, \ldots, X_n\right) \\ &\quad + 2E_{F_0}\left(t_{12n}t_{21n}|X_1, X_2, \ldots, X_n\right)\Big]. \end{aligned} \tag{2.108}$$

Or,

$$\begin{aligned} E_{F_0}\left(t^2_{12n}|X_1, X_2, \ldots, X_n\right) &= n^2 h \int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_2 - x}{h}, x\right) Z_n\left(\frac{X_2 - x'}{h}, x'\right) \\ &\quad \times \tilde{W}^n(X_1, x, h)\bar{W}^n(X_2, x, h)\tilde{W}^n(X_1, x', h)\bar{W}^n(X_2, x', h) \\ &\quad \times \left(I_{\{y_1 \leq y\}} - F_0(y|X_1)\right)\left(I_{\{y_2 \leq y\}} - F_0(y|X_2)\right) \\ &\quad \times \left(I_{\{y_1 \leq y'\}} - F_0(y'|X_1)\right)\left(I_{\{y_2 \leq y'\}} - F_0(y'|X_2)\right) \\ &\quad \times w(x)w_x(y)f_0(y|x)dydxw(x')w_{x'}(y')f_0(y'|x')dy'dx' \\ &\quad \times f_0(y_1|X_1)dy_1 f_0(y_2|X_2)dy_2, \end{aligned}$$

$$\begin{aligned}
&\leq n^2 h \int_0^1 \int_{\mathbb{R}} \int_0^1 \int_{\mathbb{R}} \left|Z_n\left(\frac{X_2-x}{h},x\right)\right| \left|Z_n\left(\frac{X_2-x'}{h},x'\right)\right| \\
&\quad \times |\tilde{W}^n(X_1,x,h)||\bar{W}^n(X_2,x,h)||\tilde{W}^n(X_1,x',h)||\bar{W}^n(X_2,x',h)| \\
&\quad \times \left(\int_{\mathbb{R}} \left|I_{\{y_1\leq y\}} - F_0(y|X_1)\right| \left|I_{\{y_1\leq y'\}} - F_0(y'|X_1)\right| f_0(y_1|X_1)dy_1\right) \\
&\quad \times \left(\int_{\mathbb{R}} \left|I_{\{y_2\leq y\}} - F_0(y|X_2)\right| \left|I_{\{y_2\leq y'\}} - F_0(y'|X_2)\right| f_0(y_2|X_2)dy_2\right) \\
&\quad \times w(x)w_x(y)f_0(y|x)dydxw(x')w_{x'}(y')f_0(y'|x')dy'dx', \\
&\leq Mn^2 h \int_0^1 \int_{\mathbb{R}} \int_0^1 \int_{\mathbb{R}} |\tilde{W}^n(X_1,x,h)||\bar{W}^n(X_2,x,h)| \\
&\quad \times |\tilde{W}^n(X_1,x',h)||\bar{W}^n(X_2,x',h)|dxdx', \\
&\leq \frac{M}{n^2h^3} \int_0^1 \int_0^1 \left|K^*\left(\frac{X_1-x}{h}\right)\right| \left|K^*\left(\frac{X_1-x'}{h}\right)\right| \\
&\quad \times \left|\bar{K}\left(\frac{X_2-x}{h}\right)\right| \left|\bar{K}\left(\frac{X_2-x'}{h}\right)\right| dxdx', \qquad (2.109)
\end{aligned}$$

en utilisant les définitions de $\tilde{W}^n(.,.,.)$ et de $\bar{W}^n(.,.,.)$ données respectivement en (2.6) et (2.71). Maintenant, on pose les changements de variables $x = -uh + X_1$ et $x' = -vh + X_2$ de Jacobien $\frac{1}{h^2}$ pour obtenir,

$$\begin{aligned}
E_{F_0}\left(t_{12n}^2|X_1,X_2,\ldots,X_n\right) &\leq \frac{M}{n^2h} \int_{\frac{X_2-1}{h}}^{\frac{X_2}{h}} \int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}} \left|\bar{K}\left(\frac{X_2-X_1}{h}+u\right)\right| \left|K^*(u)\right| \\
&\quad \times \left|K^*\left(\frac{X_1-X_2}{h}+v\right)\right| \left|\bar{K}(v)\right| dudv, \\
&\leq \frac{M}{n^2h} \int_{\mathbb{R}} \int_{\mathbb{R}} \left|\bar{K}\left(\frac{X_2-X_1}{h}+u\right)\right| \left|K^*(u)\right| \\
&\quad \times \left|K^*\left(\frac{X_1-X_2}{h}+v\right)\right| \left|\bar{K}(v)\right| dudv. \qquad (2.110)
\end{aligned}$$

On en déduit que

$$E_F\left(E_{F_0}(t_{12n}^2|X_1,X_2,\ldots,X_n)\right)$$

$$
\begin{aligned}
&\leq E_F\left(\frac{M}{n^2h}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}\left(\frac{X_2-X_1}{h}+u\right)\right|\left|K^*(u)\right|\left|K^*\left(\frac{X_1-X_2}{h}+v\right)\right|\left|\bar{K}(v)\right|dudv\right),\\
&= \frac{M}{n^2h}\int_{-1}^{2}\int_{-1}^{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}\left(\frac{x_2-x_1}{h}+u\right)\right|\left|K^*(u)\right|\\
&\quad\times\left|K^*\left(\frac{x_1-x_2}{h}+v\right)\right|\left|\bar{K}(v)\right|dudvf(x_1)f(x_2)dx_1dx_2,\\
&\leq \frac{M}{n^2h}\int_{-1}^{2}\int_{-1}^{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}\left(\frac{x_2-x_1}{h}+u\right)\right|\left|K^*(u)\right|\\
&\quad\times\left|K^*\left(\frac{x_1-x_2}{h}+v\right)\right|\left|\bar{K}(v)\right|dudvdx_1dx_2,\\
&= \frac{M}{n^2}\int_{-1}^{2}\int_{\frac{-1-x_1}{h}}^{\frac{2-x_1}{h}}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}(-t+u)\right|\left|K^*(u)\right|\left|K^*(t+v)\right|\left|\bar{K}(v)\right|dudvdtdx_1,\\
&\leq \frac{M}{n^2}\int_{-1}^{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}(-t+u)\right|\left|K^*(u)\right|\left|K^*(t+v)\right|\left|\bar{K}(v)\right|dudvdtdx_1,\\
&\leq \frac{M}{n^2}.
\end{aligned}
\tag{2.111}
$$

On majore de la même manière les quantités $E_F\left(E_{F_0}(t_{21n}^2|X_1,X_2,\ldots,X_n)\right)$ et $E_F\left(E_{F_0}\left(t_{12n}t_{21n}|X_1,X_2,\ldots,X_n\right)\right)$, ce qui permet de déduire que :

$$
Var_{F_0}\left(\tilde{\tilde{T}}_2\right) \leq 2n(n-1)\times\frac{M}{n^2} = O(1). \tag{2.112}
$$

Comme $E_{F_0}(\tilde{\tilde{T}}_2)=0$ et $Var_{F_0}(\tilde{\tilde{T}}_2)=O(1)$, on en déduit, à l'instar de (2.91),

$$
\tilde{\tilde{T}}_2 = O_p(1). \tag{2.113}
$$

Si maintenant on subsitue les résultats (2.80), (2.91), (2.98) et (2.113) dans la décomposition (2.72) de la statistique de test, on obtient, sous l'hypothèse $(H.0)$,

$$
\begin{aligned}
T &= \tilde{T}+h^2\left[O\left(h^{-\frac{1}{2}}\right)+o_p(1)+O_p(1)\right]+2h\left[O\left(h^{-\frac{1}{2}}\right)+o_p(1)+O_p(1)\right],\\
&= \tilde{T}+O\left(h^{\frac{3}{2}}\right)+o_p\left(h^2\right)+O_p\left(h^2\right)+O\left(h^{\frac{1}{2}}\right)+o_p(h)+O_p(h),
\end{aligned}
$$

$$= \tilde{T} + o_p(1), \tag{2.114}$$

de sorte qu'on en déduit :

$$\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} = \frac{\tilde{T} - h^{-\frac{1}{2}}a_0}{\sigma_0} + o_p(1), \tag{2.115}$$

et donc $\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0}$ et $\frac{\tilde{T} - h^{-\frac{1}{2}}a_0}{\sigma_0}$ ont le même comportement asymptotique donné dans le corollaire suivant :

Corollaire 2.5.1 *Sous les hypothèses du Théorème* 2.4.8 *et sous l'hypothèse supplémentaire* $(F_0^{(1)}.x.2)$, *on a :*

$$\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} \xrightarrow{\mathcal{L}} N(0,1) \tag{2.116}$$

où les quantités a_0 *et* σ_0^2 *sont définies respectivement en* (2.15) *et* (2.69).

Du Corollaire 2.5.1 découle alors une stratégie de test de niveau α pour l'hypothèse nulle donnée en (2.1) : on rejettera H_0 si

$$\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite. La pertinence d'un test unilatéral pour tester H_0 donnée en (2.1) sera expliquée au Chapitre 5.

Notons que les expressions de a_0 et σ_0^2 font intervenir la densité marginale $f(x)$ de X. Or, souvent en pratique, cette densité est inconnue. Cette stratégie de test est alors inutilisable. On est donc amené à considérer la statistique de test

$$\frac{T - h^{-\frac{1}{2}}a_1}{\sigma_1}, \tag{2.117}$$

où a_1 et σ_1 sont des estimateurs convergents de a_0 et σ_0. Une possibilité est de remplacer dans ces termes la densité $f(x)$ par un estimateur $\hat{f}_n(x)$.

De (2.15), a_0 peut s'écrire

$$a_0 = \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{f(x)} dy dx, \tag{2.118}$$

avec

$$\phi(x,y) = K^{*^{(2)}}(0)F_0(y|x)\left(1 - F_0(y|x)\right)w_x(y)f_0(y|x)w(x), \tag{2.119}$$

et de façon similaire, de (2.69),

$$\sigma_0^2 = \int_0^1 \int_{\mathbb{R}^2} \frac{\phi_1(x,y,y')}{f(x)^2} dydy'dx, \tag{2.120}$$

avec

$$\begin{aligned} \phi_1(x,y,y') &= 2K^{*^{(4)}}(0)\left[F_0(y \wedge y'|x) - F_0(y|x)F_0(y'|x)\right]^2 \\ &\quad \times w^2(x)w_x(y)w_x(y')f_0(y|x)f_0(y'|x). \end{aligned} \tag{2.121}$$

Soit $\hat{f}_n(x)$ un estimateur de $f(.)$ et définissons

$$a_1 = \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{\hat{f}_n(x)} dydx, \tag{2.122}$$

et

$$\sigma_1^2 = \int_0^1 \int_{\mathbb{R}^2} \frac{\phi_1(x,y,y')}{\hat{f}_n(x)^2} dydy'dx. \tag{2.123}$$

Maintenant,

$$\frac{T - h^{-\frac{1}{2}}a_1}{\sigma_1} = \left[\frac{T - h^{-\frac{1}{2}}a_0 + h^{-\frac{1}{2}}(a_0 - a_1)}{\sigma_0}\right] \times \frac{\sigma_0}{\sigma_1}. \tag{2.124}$$

Or,

$$\begin{aligned} h^{-\frac{1}{2}}|a_0 - a_1| &= h^{-\frac{1}{2}} \left| \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{f(x)} dydx - \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{\hat{f}_n(x)} dydx \right|, \\ &= h^{-\frac{1}{2}} \left| \int_0^1 \int_{\mathbb{R}} \phi(x,y) \left[\frac{1}{f(x)} - \frac{1}{\hat{f}_n(x)}\right] dydx \right|, \\ &\leq h^{-\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{f(x)} \left|1 - \frac{f(x)}{\hat{f}_n(x)}\right| dydx, \\ &\leq h^{-\frac{1}{2}} \sup_{x\in[0,1]} \left|1 - \frac{f(x)}{\hat{f}_n(x)}\right| \int_0^1 \int_{\mathbb{R}} \frac{\phi(x,y)}{f(x)} dydx, \\ &= a_0 h^{-\frac{1}{2}} \sup_{x\in[0,1]} \left|1 - \frac{f(x)}{\hat{f}_n(x)}\right|, \\ &= o(1), \end{aligned} \tag{2.125}$$

si $h^{-\frac{1}{2}} \sup\limits_{x\in[0,1]} \left|\hat{f}_n(x) - f(x)\right| = o_p(1)$ et en rappelant que $0 < m < f(x) < M \ \forall x \in [0,1]$. De la même façon,

$$\begin{aligned}
\left|\sigma_0^2 - \sigma_1^2\right| &= \left|\int_0^1 \int_{\mathbb{R}^2} \phi_1(x,y,y')\Big[\frac{1}{f(x)^2} - \frac{1}{\hat{f}_n(x)^2}\Big]dydy'dx\right|, \\
&\leq \int_0^1 \int_{\mathbb{R}^2} \frac{\phi_1(x,y,y')}{f(x)^2}\left|1 - \frac{f(x)^2}{\hat{f}_n(x)^2}\right| dydy'dx, \\
&\leq \sup_{x\in[0,1]}\left|1 - \frac{f(x)^2}{\hat{f}_n(x)^2}\right| \int_0^1 \int_{\mathbb{R}^2} \frac{\phi_1(x,y,y')}{f(x)^2} dydy'dx, \\
&= \sigma_0^2 \sup_{x\in[0,1]}\left|1 - \frac{f(x)^2}{\hat{f}_n(x)^2}\right|, \\
&= o(1), \qquad (2.126)
\end{aligned}$$

si $\sup\limits_{x\in[0,1]} \left|\hat{f}_n(x) - f(x)\right| = o_p(1)$. On en déduit que $\sigma_0 - \sigma_1 = o_p(1)$, c'est à dire

$$\frac{\sigma_0}{\sigma_1} = 1 + o_p(1). \qquad (2.127)$$

Ainsi, si on substitue les résultats (2.125) et (2.127) dans l'expression (2.124),

$$\frac{T - h^{-\frac{1}{2}}a_1}{\sigma_1} = \frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} + o_p(1), \qquad (2.128)$$

et $\dfrac{T - h^{-\frac{1}{2}}a_1}{\sigma_1}$ et $\dfrac{T - h^{-\frac{1}{2}}a_0}{\sigma_0}$ ont le même comportement asymptotique. On a donc montré le corollaire suivant :

Corollaire 2.5.2 *Si $\hat{f}_n(.)$ est un estimateur de la densité $f(.)$ de X tel que*

$$h^{-\frac{1}{2}} \sup_{x\in[0,1]} \left|\hat{f}_n(x) - f(x)\right| = o_p(1),$$

alors sous les hypothèses du Corollaire 2.5.1, on a :

$$\frac{T - h^{-\frac{1}{2}}a_1}{\sigma_1} \xrightarrow{\mathcal{L}} N(0,1) \qquad (2.129)$$

où les quantités a_1 et σ_1 sont données respectivement en (2.122) et (2.123).

Ainsi, une stratégie de test de niveau α de H_0 donnée en (2.1) consiste à rejeter H_0 si

$$\frac{T - h^{-\frac{1}{2}} a_1}{\sigma_1} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Note Si $\hat{f}_n(x)$ est l'estimateur à noyau H basé sur $X_1, \ldots, X_n$ de fenêtre k, c'est à dire :

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

alors de Rao (1983, p48),

$$h^{-\frac{1}{2}} \sup_{x \in [0,1]} \left|\hat{f}_n(x) - f(x)\right| = h^{-\frac{1}{2}} \left[O_{ps}\left(n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}} k^{-1}\right) + O(k)\right],$$

de sorte que pour que $h^{-\frac{1}{2}} \sup_{x \in [0,1]} \left|\hat{f}_n(x) - f(x)\right|$ tende vers 0, on doit choisir k tel que :

$$\frac{\sqrt{\log n}}{\sqrt{n} k \sqrt{h}} \longrightarrow 0 \text{ et } h^{-\frac{1}{2}} k \longrightarrow 0.$$

Plus précisément, si on pose $h = O(n^{-\delta_1})$ et $k = O(n^{-\delta_2})$ avec $0 < \delta_1, \delta_2 < 1$, les deux conditions précédentes impliquent que :

$$\frac{\delta_1}{2} < \delta_2 < \frac{1 - \delta_1 - \epsilon}{2},$$

où $\epsilon > 0$ est petit. Cette relation entre δ_1 et δ_2 assure un choix assez large de valeurs pour δ_2. En particulier, suite aux hypothèses données à la Section 1.3.3 concernant la fenêtre h, il est toujours possible de choisir la valeur de k optimale pour l'estimation à noyau de la densité, c'est à dire $k = O\left(n^{-\frac{1}{5}}\right)$.

Les résultats obtenus dans ce chapitre concernant le comportement asymptotique de la statistique de test dans le cas où $F_0(y|x)$ est un polynôme de degré $\leq p$ en x vont être utilisés au chapitre suivant qui considère le cas plus réel où $F_0(y|x)$ est une fonction de répartition quelconque.

Chapitre 3

Etude du comportement asymptotique de la statistique de test sous H_0 dans le cas où $F_0(y|x)$ est une fonction quelconque

3.1 Introduction

A l'instar du Chapitre 2, on étudie dans ce chapitre le comportement asymptotique de la statistique de test T donnée en (1.42) pour tester l'hypothèse nulle

$$H_0 : F(y|x) = F_0(y|x) \quad \forall y \in \mathbb{R}, \forall x \in [0,1]. \tag{3.1}$$

Nous allons supposer ici que $F_0(y|x)$ est une fonction de répartition quelconque mais satisfaisant néanmoins certaines conditions de régularité, ce qui est plus pertinent en pratique que le cas étudié au Chapitre 2 où $F_0(y|x)$ était un polynôme en x de degré inférieur ou égal à p. Le comportement asymptotique de la statistique de test dans ce cas de figure découle partiellement des résultats qui ont été obtenus au chapitre précédent. Il convient cependant de donner une nouvelle décomposition de cette statistique de test. Cette décomposition nécessite le calcul du biais conditionnel de l'estimateur polynômial local de $F(y|x)$ (qui est nul dans le contexte du Chapitre 2), biais qui est contrôlé par le choix du paramètre p, l'ordre de l'ajustement polynômial local de $F(y|x)$. Nous discutons de ce choix dans le prochain paragraphe.

3.2 Choix du paramètre p

Le paramètre p est l'un des trois paramètres qui apparaissent de manière explicite dans le problème de minimisation (1.9), puis de manière implicite dans l'expression de l'estimateur polynômial local de la fonction de répartition conditionnelle (1.15). Cette expression de l'estimateur polynômial local $\hat{F}_n(y|x)$ permet aussi de voir que plus p est grand, plus l'expression de l'estimateur devient complexe. Pour $p = 0$, on retrouve l'estimateur de la fonction de répartition conditionnelle étudié par Collomb(1980), aussi appelé l'ajustement constant. Pour $p = 1$, on a l'estimateur linéaire local étudié par Ducharme et Mint El Mouvid (2001).

Dans un contexte d'estimation de la fonction de régression $E(Y|X = x)$, Fan et Gijbels (1996, p62) remarquent que bien que le biais de l'estimateur soit en priorité contrôlé par la fenêtre h, le choix de l'ordre du polynôme local a une certaine importance. En effet, ils montrent que le biais de l'estimateur polynômial local de $E(Y|X = x)$ passe de l'ordre $O(h^{p+1})$ à l'ordre $O(h^{p+2})$ selon que p est respectivement impair ou pair. Par contre, si p est pair, un terme supplémentaire apparaît dans l'expression de ce biais. En pratique, ce terme supplémentaire, fonction de $\dfrac{f'(x)}{f(x)}$ est plutôt gênant puisque les fonctions $f'(x)$ et $f(x)$ étant inconnues, il augmente le nombre de termes inconnus dans l'expression du biais. Ainsi, la plupart des auteurs recommandent de choisir p impair pour l'estimation par les polynômes locaux de la fonction de régression.

On choisit ici d'emboîter le pas à ces auteurs et dans la suite de ce travail, on se limite aux cas où le degré p du polynôme utilisé pour l'estimation de $F(y|x)$ est impair. Dans ce contexte qu'on ne précisera plus, on étudie à présent le biais de $\hat{F}_n(y|x)$.

3.3 Calcul du biais de $\hat{F}_n(y|x)$, conditionnellement aux X observés

La décomposition de la statistique de test passe d'abord par la détermination de l'espérance de $\hat{F}_n(y|x)$ conditionnellement aux X observés ou plus précisément de l'ordre de ce biais lorsque p est impair.

Proposition 3.3.1 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0^{(p+1)}.x.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$ et $(f.0)$,*

$$E_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n) = F_0(y|x) + O_{ps}^{(x,y)}(h^{p+1}). \tag{3.2}$$

Dans cette expression, $O_{ps}^{(x,y)}(h^{p+1}) = h^{p+1}O_{ps}^{(x,y)}(1)$ et le terme $O_{ps}^{(x,y)}(1)$, déjà défini en (1.26), est borné en $x \in [0,1]$ et en $y \in \mathbb{R}$.

Preuve En utilisant la définition de $\hat{F}_n(y|x)$ donnée en (1.15), on a :

$$\begin{aligned}
E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n) &= E_{F_0}\left(\sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right) I_{\{Y_i\le y\}}|X_1,\ldots,X_n\right), \\
&= \sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right) E_{F_0}(I_{\{Y_i\le y\}}|X_1,\ldots,X_n), \\
&= \sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right) F_0(y|X_i). \qquad (3.3)
\end{aligned}$$

De plus, sous l'hypothèse $(F_0^{(p+1)}.x.2)$, la fonction qui à z associe $F_0(y|z)$ est $p+1$ fois continûment dérivable dans un voisinage de tout $x\in]-1,2[$. On effectue alors un développement de Taylor d'ordre $p+1$ de $F_0(y|X_i)$ autour de $x\in[0,1]$ et pour tout $i=1,\ldots,n$, il existe $x_i^*\in(X_i,x)$ tel que :

$$\begin{aligned}
F_0(y|X_i) &= F_0(y|x)+(X_i-x)F_0^{(1)}(y|x)+\ldots+(X_i-x)^p\frac{F_0^{(p)}(y|x)}{p!} \\
&\quad +(X_i-x)^{p+1}\frac{F_0^{(p+1)}(y|x_i^*)}{(p+1)!}, \qquad (3.4)
\end{aligned}$$

où on rappelle que $F_0^{(j)}(y|x)$ est la $j^{\text{ième}}$ dérivée de la fonction $F_0(y|x)$ par rapport à x. Si on substitue ce résultat dans l'expression (3.3), on obtient, en utilisant les propriétés du noyau $W^n(.)$, données en (1.17),

$$\begin{aligned}
E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n) &= \sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right)\left[F_0(y|x)+\sum_{j=1}^p (X_i-x)^j\frac{F_0^{(j)}(y|x)}{j!}\right. \\
&\quad \left. +(X_i-x)^{p+1}\frac{F_0^{(p+1)}(y|x_i^*)}{(p+1)!}\right], \\
&= F_0(y|x)\sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right) \\
&\quad +\sum_{j=1}^p \frac{F_0^{(j)}(y|x)}{j!}\sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right)(X_i-x)^j \\
&\quad +\sum_{i=1}^n W^n\left(\frac{X_i-x}{h}\right)(X_i-x)^{p+1}\frac{F_0^{(p+1)}(y|x_i^*)}{(p+1)!}, \\
&= F_0(y|x)+R_{xy}, \qquad (3.5)
\end{aligned}$$

où R_{xy} est le "reste" associé au développement de Taylor (3.4) défini par :

$$R_{xy} = \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1}\frac{F_0^{(p+1)}(y|x_i^*)}{(p+1)!}. \tag{3.6}$$

Dans ce qui suit, M est une constante majorante pouvant varier d'une ligne à l'autre. Maintenant, suite à la décomposition du noyau $W^n(.)$ donnée en (1.34),

$$\begin{aligned}
|R_{xy}| &= \left|\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1}\frac{F_0^{(p+1)}(y|x_i^*)}{(p+1)!}\right|,\\
&\leq \sum_{i=1}^{n}\left|\tilde{W}^n(X_i,x,h)\right|(X_i - x)^{p+1}\frac{\left|F_0^{(p+1)}(y|x_i^*)\right|}{(p+1)!},\\
&\quad + h\sum_{i=1}^{n}\left|Z_n\left(\frac{X_i - x}{h},x\right)\right|\left|\bar{W}^n(X_i,x,h)\right|(X_i - x)^{p+1}\frac{\left|F_0^{(p+1)}(y|x_i^*)\right|}{(p+1)!},
\end{aligned} \tag{3.7}$$

où les noyaux $\tilde{W}^n(.,.,.)$ et $\bar{W}^n(.,.,.)$ sont définis respectivement en (2.6) et (2.71). Maintenant, comme $\left|Z_n\left(\frac{X_i - x}{h},x\right)\right| \leq M$ et sous l'hypothèse $(F_0^{(p+1)}.x.2)$,

$$\begin{aligned}
|R_{xy}| &\leq \sup_{y\in\mathbb{R}}\sup_{x\in[0,1]}\frac{|F_0^{(p+1)}(y|x)|}{(p+1)!}\times\sum_{i=1}^{n}\left|\tilde{W}^n(X_i,x,h)\right|(X_i - x)^{p+1},\\
&\quad + Mh\sup_{y\in\mathbb{R}}\sup_{x\in[0,1]}\frac{|F_0^{(p+1)}(y|x)|}{(p+1)!}\times\sum_{i=1}^{n}\left|\bar{W}^n(X_i,x,h)\right|(X_i - x)^{p+1},\\
&\leq M\sum_{i=1}^{n}\left|\tilde{W}^n(X_i,x,h)\right|(X_i - x)^{p+1},\\
&\quad + Mh\sum_{i=1}^{n}\left|\bar{W}^n(X_i,x,h)\right|(X_i - x)^{p+1},\\
&= \frac{Mh^{p+1}}{f(x)}\times\frac{1}{nh}\sum_{i=1}^{n}\left|K^*\left(\frac{X_i - x}{h}\right)\right|\left(\frac{X_i - x}{h}\right)^{p+1},\\
&\quad + \frac{Mh^{p+2}}{f(x)}\times\frac{1}{nh}\sum_{i=1}^{n}\left|\bar{K}\left(\frac{X_i - x}{h}\right)\right|\left(\frac{X_i - x}{h}\right)^{p+1}.
\end{aligned} \tag{3.8}$$

De plus, en recyclant le résultat (1.23) aux statistiques $\frac{1}{nh}\sum_{i=1}^{n}\left|K^*\left(\frac{X_i - x}{h}\right)\right|\left(\frac{X_i - x}{h}\right)^{p+1}$ et $\frac{1}{nh}\sum_{i=1}^{n}\left|\bar{K}\left(\frac{X_i - x}{h}\right)\right|\left(\frac{X_i - x}{h}\right)^{p+1}$, on obtient, de (1.26) et sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$ et $(H.2)$:

$$\frac{1}{nh}\sum_{i=1}^{n}\left|K^*\left(\frac{X_i-x}{h}\right)\right|\left(\frac{X_i-x}{h}\right)^{p+1}=O(1)+hO_{ps}^x(1), \tag{3.9}$$

et

$$\frac{1}{nh}\sum_{i=1}^{n}\left|\bar{K}\left(\frac{X_i-x}{h}\right)\right|\left(\frac{X_i-x}{h}\right)^{p+1}=O(1)+hO_{ps}^x(1), \tag{3.10}$$

où les termes $O_{ps}^x(.)$, déjà définis en (1.23), sont presque sûrement bornés en $x \in [0,1]$. On en déduit que sous l'hypothèse $(f.0)$,

$$\begin{aligned} |R_{xy}| &\leq O(h^{p+1})+O_{ps}(h^{p+2})+O(h^{p+2})+O_{ps}(h^{p+3}), \\ &= O_{ps}(h^{p+1}). \end{aligned} \tag{3.11}$$

Il est important de signaler que cette majoration tient uniformément pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$. On en déduit alors le résultat cherché. ♣

3.4 Décomposition de la statistique de test

Tout d'abord, on peut écrire que :

$$\begin{aligned} \hat{F}_n(y|x)-F_0(y|x) &= \left(\hat{F}_n(y|x)-E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n)\right) \\ &\quad+\left(E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n)-F_0(y|x)\right). \end{aligned} \tag{3.12}$$

De plus, comme on l'a montré en (3.3),

$$E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n)=\sum_{i=1}^{n}W^n\left(\frac{X_i-x}{h}\right)F_0(y|X_i),$$

de sorte que,

$$\hat{F}_n(y|x)-E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n)=\sum_{i=1}^{n}W^n\left(\frac{X_i-x}{h}\right)(I_{\{Y_i\leq y\}}-F_0(y|X_i)). \tag{3.13}$$

D'autre part, sous les hypothèses de la Proposition 3.3.1, on a que presque sûrement, si n est suffisamment grand,

$$E_{F_0}(\hat{F}_n(y|x)|X_1,\ldots,X_n)-F_0(y|x)=h^{p+1}R_n(x,y), \tag{3.14}$$

où $R_n(x,y)$ est presque sûrement borné en $x \in [0,1]$ et en $y \in \mathbb{R}$. Ainsi,

$$\begin{aligned}\hat{F}_n(y|x) - F_0(y|x) &= \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right) \\ &\quad + h^{p+1} R_n(x,y).\end{aligned} \tag{3.15}$$

Si on élève ce terme au carré, on obtient :

$$\begin{aligned}\left[\hat{F}_n(y|x) - F_0(y|x)\right]^2 &= \left[\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right)\right]^2 + h^{2(p+1)} R_n^2(x,y) \\ &\quad + 2h^{p+1} R_n(x,y) \sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right),\end{aligned}$$

qui, injecté dans (1.42), donne à l'instar de (2.72) obtenue dans le cas où $F_0(y|x)$ est un polynôme en x, la décomposition suivante :

$$T = T^* + T_3 + T_4, \tag{3.16}$$

où T^* est semblable à la statistique donnée en (2.72), dont le comportement asymptotique a été étudié au chapitre précédent,

$$T_3 = nh^{2p+\frac{5}{2}} \int_0^1 \int_{\mathbb{R}} R_n^2(x,y) w(x) w_x(y) f_0(y|x)\, dy dx, \tag{3.17}$$

$$\begin{aligned}T_4 &= 2nh^{p+\frac{3}{2}} \sum_{i=1}^{n} \int_0^1 \int_{\mathbb{R}} R_n(x,y) W^n\left(\frac{X_i - x}{h}\right)\left(I_{\{Y_i \leq y\}} - F_0(y|X_i)\right) \\ &\quad \times w(x) w_x(y) f_0(y|x)\, dy dx.\end{aligned} \tag{3.18}$$

La décomposition (3.16) permet de remarquer que le comportement asymptotique de la statistique de test T dépend du comportement asymptotique de la statistique T^* et du comportement asymptotique des statistiques T_3 et T_4 dont nous allons faire à présent l'étude.

3.5 Comportement asymptotique de la statistique de test

Etude de T^*

Le comportement asymptotique de la statistique T^* se déduit directement du Corollaire 2.5.1. En effet, ce résultat se base uniquement sur le fait que la fonction à tester sous H_0 est p fois dérivable, ce qui est le cas dans le présent contexte puisqu'on la suppose en particulier $p+1$ fois dérivable. On peut donc recycler ici le résultat du Corollaire 2.5.1.

Etude de T_3 et T_4

Proposition 3.5.1 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(P.0)$, $(P.1)$ et $(H.3)$, on a :*

$$T_3 = o_p(1). \tag{3.19}$$

Preuve On a,

$$\begin{aligned} T_3 &= nh^{2p+\frac{5}{2}} \int_0^1 \int_{\mathbb{R}} R_n^2(x,y)w(x)w_x(y)f_0(y|x)dydx, \\ &\leq Mnh^{2p+\frac{5}{2}} \int_0^1 \int_{\mathbb{R}} f_0(y|x)dydx, \\ &= o_p(1).\clubsuit \end{aligned}$$

Proposition 3.5.2 *Supposons que $F_0(y|x)$ vérifie l'hypothèse $(F_0.y.1)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$ et $(H.3)$,*

$$T_4 = o_p(1). \tag{3.20}$$

Preuve De (3.18) et en utilisant la décomposition du noyau $W^n(.)$, donnée en (1.34), on peut écrire que :

$$T_4 = 2\tilde{T}_4 + 2h\bar{T}_4, \tag{3.21}$$

où on a posé

$$\tilde{T}_4 = nh^{p+\frac{3}{2}} \sum_{i=1}^n v_i,$$

et

$$\bar{T}_4 = nh^{p+\frac{3}{2}} \sum_{i=1}^n w_i,$$

avec

$$v_i = \int_0^1 \int_{\mathbb{R}} R_n(x,y)\tilde{W}^n(X_i,x,h)(I_{\{Y_i\leq y\}} - F_0(y|X_i))w(x)w_x(y)f_0(y|x)dydx,$$

et

$$w_i = \int_0^1 \int_{\mathbb{R}} R_n(x,y)Z_n\left(\frac{X_i - x}{h}, x\right)\bar{W}^n(X_i,x,h)(I_{\{Y_i\leq y\}} - F_0(y|X_i))w(x)w_x(y)f_0(y|x)dydx,$$

où les noyaux $\tilde{W}^n(.,.,.)$ et $\bar{W}^n(.,.,.)$ sont définis respectivement en (2.6) et (2.71) et les termes $R_n(x,y)$ et $Z_n\left(\dfrac{X_i-x}{h},x\right)$ sont presque sûrement bornés en $x\in[0,1]$ et en $y\in\mathbb{R}$. Par le théorème de Fubini et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned}
E_{F_0}(v_1|X_1,\ldots,X_n) &= \int_0^1\int_{\mathbb{R}} R_n(x,y)\tilde{W}^n(X_1,x,h)\left(\int_{\mathbb{R}}(I_{\{y_1\le y\}}-F_0(y|X_1))f_0(y_1|X_1)dy_1\right)\\
&\quad\times w(x)w_x(y)f_0(y|x)dydx,\\
&= 0,
\end{aligned}$$

car $\int_{\mathbb{R}}(I_{\{y_1\le y\}}-F_0(y|X_1))f_0(y_1|X_1)dy_1 = 0$, $\forall y\in\mathbb{R}$, $\forall X_1\in[-1,2]$. Il s'en suit que $E_{F_0}(v_1)=0$ et donc $E_{F_0}\left(\tilde{T}_4\right)=0$. Maintenant, comme $E_{F_0}(v_1|X_1,\ldots,X_n)=0$,

$$\begin{aligned}
Var_{F_0}(\tilde{T}_4) &= n^2h^{2p+3}Var_{F_0}\left(\sum_{i=1}^n v_i\right),\\
&= n^2h^{2p+3}E_F\left(Var_{F_0}\left(\sum_{i=1}^n v_i|X_1,X_2,\ldots,X_n\right)\right),\\
&= n^3h^{2p+3}E_F\left(E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right)\right). \qquad (3.22)
\end{aligned}$$

Calculons à présent $E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right)$. Nous avons,

$$\begin{aligned}
v_1^2 &= \int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} R_n(x,y)R_n(x',y')\tilde{W}^n(X_1,x,h)\tilde{W}^n(X_1,x',h)\\
&\quad\times(I_{\{Y_1\le y\}}-F_0(y|X_1))(I_{\{Y_1\le y'\}}-F_0(y'|X_1))w(x)w(x')w_x(y)w_{x'}(y')\\
&\quad\times f_0(y|x)f_0(y'|x')dydxdy'dx',
\end{aligned}$$

de sorte que,

$$\begin{aligned}
E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right) &\le \int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1 |R_n(x,y)||R_n(x',y')|\left|\tilde{W}^n(X_1,x,h)\right|\left|\tilde{W}^n(X_1,x',h)\right|\\
&\quad\times\left(\int_{\mathbb{R}}\left|I_{\{y_1\le y\}}-F_0(y|X_1)\right|\left|I_{\{y_1\le y'\}}-F_0(y'|X_1)\right|f_0(y_1|X_1)dy_1\right)\\
&\quad\times w(x)w(x')w_x(y)w_{x'}(y')f_0(y|x)f_0(y'|x')dydxdy'dx'. \qquad (3.23)
\end{aligned}$$

De plus, $\forall y,y'\in\mathbb{R}$ et $\forall X_1\in[-1,2]$,

$$\int_{\mathbb{R}}\left|I_{\{y_1\le y\}}-F_0(y|X_1)\right|\left|I_{\{y_1\le y'\}}-F_0(y'|X_1)\right|f_0(y_1|X_1)dy_1\le M,$$

de sorte que, sous les hypothèses $(P.0)$ et $(P.1)$,

$$\begin{aligned}
E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right) &\leq M\int_0^1\int_0^1\left|\tilde{W}^n(X_1,x,h)\right|\left|\tilde{W}^n(X_1,x',h)\right| \\
&\quad\times\left(\int_{\mathbb{R}} f_0(y|x)dy\right)\left(\int_{\mathbb{R}} f_0(y'|x')dy'\right)dxdx', \\
&\leq \frac{M}{n^2h^2}\int_0^1\int_0^1 |K^*\left(\frac{X_1-x}{h}\right)||K^*\left(\frac{X_1-x'}{h}\right)|dxdx', \quad (3.24)
\end{aligned}$$

en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6). On effectue les changements de variables $x=-uh+X_1$ et $x'=-vh+X_1$ de Jacobien $\dfrac{1}{h^2}$ pour obtenir, comme $h<1$ et $X_1\in[-1,2]$,

$$\begin{aligned}
E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right) &\leq \frac{M}{n^2}\int_{\frac{X_1-x}{h}}^{\frac{X_1}{h}}\int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}} |K^*(u)|\,|K^*(v)|\,dudv, \\
&= \frac{M}{n^2}\left(\int_{\mathbb{R}}|K^*(u)|du\right)\left(\int_{\mathbb{R}}|K^*(v)|dv\right), \\
&= \frac{M}{n^2},
\end{aligned}$$

sous l'hypothèse $(K.0)$. Ainsi,

$$\begin{aligned}
0\leq Var_{F_0}\left(\tilde{T}_4\right) &\leq n^3h^{2p+3}\times\frac{M}{n^2}, \\
&= Mnh^{2p+3}, \\
&= o(1),
\end{aligned}$$

ce qui permet de déduire que

$$\tilde{T}_4 = o_p(1). \quad (3.25)$$

Etudions maintenant le comportement asymptotique de $\bar{T}_4$. Par le théorème de Fubini et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned}
E_{F_0}(w_1|X_1,\ldots,X_n) &= \int_0^1\int_{\mathbb{R}} R_n(x,y)Z_n\left(\frac{X_1-x}{h},x\right)\bar{W}^n(X_1,x,h) \\
&\quad\times\left(\int_{\mathbb{R}}(I_{\{y_1\leq y\}}-F_0(y|X_1))f_0(y_1|X_1)dy_1\right)
\end{aligned}$$

$$\times\ w(x)w_x(y)f_0(y|x)dydx,$$

$$= \quad 0,$$

car $\int_{\mathbb{R}}(I_{\{y_1\leq y\}} - F_0(y|X_1))f_0(y_1|X_1)dy_1 = 0,\ \forall y \in \mathbb{R}, \forall X_1 \in [-1,2]$. Il s'en suit que $E_{F_0}(w_1) = 0$ et donc $E_{F_0}\left(\bar{T}_4\right) = 0$. Maintenant, comme $E_{F_0}(w_1|X_1,\ldots,X_n) = 0$,

$$\begin{aligned} Var_{F_0}(\bar{T}_4) &= n^2h^{2p+3}Var_{F_0}\left(\sum_{i=1}^n w_i\right), \\ &= n^2h^{2p+3}E_F\left(Var_{F_0}\left(\sum_{i=1}^n w_i|X_1,X_2,\ldots,X_n\right)\right), \\ &= n^3h^{2p+3}E_F\left(E_{F_0}\left(w_1^2|X_1,\ldots,X_n\right)\right). \end{aligned} \tag{3.26}$$

Calculons à présent $E_{F_0}\left(w_1^2|X_1,\ldots,X_n\right)$. Nous avons,

$$\begin{aligned} w_1^2 = & \int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} R_n(x,y)R_n(x',y')Z_n\left(\frac{X_1-x}{h},x\right)Z_n\left(\frac{X_1-x'}{h},x'\right) \\ & \times \bar{W}^n(X_1,x,h)\bar{W}^n(X_1,x',h) \\ & \times (I_{\{Y_1\leq y\}} - F_0(y|X_1))(I_{\{Y_1\leq y'\}} - F_0(y'|X_1)) \\ & \times w(x)w(x')w_x(y)w_{x'}(y')f_0(y|x)f_0(y'|x')dydxdy'dx', \end{aligned}$$

de sorte que,

$$\begin{aligned} E_{F_0}\left(w_1^2|X_1,\ldots,X_n\right) \leq & \int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1 |R_n(x,y)||R_n(x',y')||Z_n\left(\frac{X_1-x}{h},x\right)||Z_n\left(\frac{X_1-x'}{h},x'\right)| \\ & \times \left|\bar{W}^n(X_1,x,h)\right|\left|\bar{W}^n(X_1,x',h)\right| \\ & \times \left(\int_{\mathbb{R}}\left|I_{\{y_1\leq y\}} - F_0(y|X_1)\right|\left|I_{\{y_1\leq y'\}} - F_0(y'|X_1)\right|f_0(y_1|X_1)dy_1\right) \\ & \times w(x)w(x')w_x(y)w_{x'}(y')f_0(y|x)f_0(y'|x')dydxdy'dx'. \end{aligned} \tag{3.27}$$

De plus, $\forall y, y' \in \mathbb{R}$ et $\forall X_1 \in [-1,2]$,

$$\int_{\mathbb{R}}\left|I_{\{y_1\leq y\}} - F_0(y|X_1)\right|\left|I_{\{y_1\leq y'\}} - F_0(y'|X_1)\right|f_0(y_1|X_1)dy_1 \leq M,$$

de sorte que, sous les hypothèses $(P.0)$ et $(P.1)$,

$$\begin{aligned}
E_{F_0}\left(w_1^2|X_1,\ldots,X_n\right) &\le M\int_0^1\int_0^1\left|\bar{W}^n(X_1,x,h)\right|\left|\bar{W}^n(X_1,x',h)\right| \\
&\quad\times\left(\int_{\mathbb{R}} f_0(y|x)dy\right)\left(\int_{\mathbb{R}} f_0(y'|x')dy'\right)dxdx', \\
&\le \frac{M}{n^2h^2}\int_0^1\int_0^1|\bar{K}\left(\frac{X_1-x}{h}\right)||\bar{K}\left(\frac{X_1-x'}{h}\right)|dxdx', \quad (3.28)
\end{aligned}$$

en utilisant la définition de $\bar{W}^n(.,.,.)$ donnée en (2.71). On effectue les changements de variables $x=-uh+X_1$ et $x'=-vh+X_1$ de Jacobien $\frac{1}{h^2}$ pour obtenir, comme $h<1$ et $X_1\in[-1,2]$,

$$\begin{aligned}
0\le E_{F_0}\left(w_1^2|X_1,\ldots,X_n\right) &\le \frac{M}{n^2}\int_{\frac{X_1-x}{h}}^{\frac{X_1}{h}}\int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}}\left|\bar{K}(u)\right|\left|\bar{K}(v)\right|dudv, \\
&= \frac{M}{n^2}\left(\int_{\mathbb{R}}|\bar{K}(u)|du\right)\left(\int_{\mathbb{R}}|\bar{K}(v)|dv\right), \\
&= \frac{M}{n^2},
\end{aligned}$$

sous l'hypothèse $(K.0)$. Ainsi,

$$\begin{aligned}
0\le Var_{F_0}\left(\bar{T}_4\right) &\le n^3h^{2p+3}\times\frac{M}{n^2}, \\
&= Mnh^{2p+3}, \\
&= o(1),
\end{aligned}$$

ce qui permet de déduire que

$$\bar{T}_4=o_p(1). \quad (3.29)$$

Si on substitue maintenant (3.25) et (3.29) dans la décomposition (3.21) de T_4 on a, sous l'hypothèse $(H.0)$,

$$\begin{aligned}
T_4 &= o_p(1)+ho_p(1), \\
&= o_p(1).\clubsuit
\end{aligned}$$

Revenons maintenant à la décomposition de la statistique de test obtenue en (3.16). Les Propositions 3.5.1 et 3.5.2 permettent de déduire que :

$$T = T^* + o_p(1), \tag{3.30}$$

de sorte que les statistiques T et T^* ont le même comportement asymptotique résumé dans le théorème suivant :

Théorème 3.5.3 *Sous les hypothèses du Corollaire* 2.5.1 *et sous les hypothèses supplémentaires* $(F_0^{(p+1)}.x.2)$ *et* $(H.3)$, *on a :*

$$\frac{T - h^{-\frac{1}{2}} a_0}{\sigma_0} \xrightarrow{\mathcal{L}} N(0,1), \tag{3.31}$$

où les quantités a_0 *et* σ_0^2 *sont données respectivement en* (2.15) et (2.69).

Ainsi, un test de niveau α de l'hypothèse nulle donnée en (3.1) consiste à rejeter H_0 si

$$\frac{T - h^{-\frac{1}{2}} a_0}{\sigma_0} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Comme dans le chapitre précédent, les expressions de a_0 et σ_0^2 intervenant dans cette étude asymptotique comportent comme paramètre de nuisance la densité $f(.)$ qui est généralement inconnue en pratique. Il est alors nécessaire de les estimer en remplaçant ce paramètre de nuisance par un estimateur. En s'appuyant directement sur le raisonnement proposé à la Section 2.5, on donne le résultat de convergence suivant pour le test de l'hypothèse nulle donnée en (3.1) :

Théorème 3.5.4 *Sous les hypothèses du Théorème* 3.5.3 *et du Corollaire* 2.5.2, *on a :*

$$\frac{T^* - h^{-\frac{1}{2}} a_1}{\sigma_1} \xrightarrow{\mathcal{L}} N(0,1), \tag{3.32}$$

où les quantités a_1 *et* σ_1^2 *sont données respectivement en* (2.122) et (2.123).

Ces résultats nous permettent de constater que moyennant l'ajout des hypothèses supplémentaires $(H.3)$ portant sur la fenêtre d'ajustement h et $(F_0^{(p+1)}.x.2)$, le comportement asymptotique de la statistique de test est le même que dans le cas où $F_0(y|x)$ est un polynôme de degré $\leq p$ en x. Nous allons à présent nous placer dans le contexte plus pratique où $F_0(y|x)$ n'est plus entièrement spécifiée mais comporte des paramètres inconnus. Cette étude fait l'objet du chapitre suivant.

Chapitre 4

Test de la fonction de répartition conditionnelle dans le cas où celle-ci comporte des paramètres inconnus

4.1 Introduction

Dans les deux chapitres précédents, on a construit un test d'adéquation pour la fonction de répartition conditionnelle dans le cas où la distribution à tester sous H_0, soit $F_0(y|x)$, est entièrement spécifiée. En pratique, cette situation est assez rare. C'est pourquoi, dans ce chapitre, on considère le cas plus général où $F_0(y|x)$ fait intervenir un certain nombre de paramètres inconnus, comme dans l'exemple (1.2) où ces paramètres sont les quantités β_0, β_1 de $m(x)$ et σ^2 de $\sigma^2(x)$. Si on rassemble ces paramètres inconnus dans un vecteur θ appartenant à l'espace paramétrique Θ que l'on suppose un sous-ensemble ouvert de $\mathbb{R}^d, d \geq 1$ et si on réécrit dans ce contexte la fonction de répartition conditionnelle sous la forme $F_0(y|x;\theta)$, alors l'hypothèse nulle simple (1.3) devient l'hypothèse nulle composite :

$$H_0 : F(\,.\,|\,.\,) \in \{F_0(\,.\,|\,.\,;\,\theta), \theta \in \Theta\}, \tag{4.1}$$

que l'on veut confronter à l'alternative :

$$H_1 : F(\,.\,|\,.\,) \notin \{F_0(\,.\,|\,.\,;\theta), \theta \in \Theta\}. \tag{4.2}$$

Comme dans les chapitres précédents, on doit être moins ambitieux et se limiter à $x \in \mathcal{C}$, où $\mathcal{C}$ est un compact appartenant au support $[-1, 2]$ de la densité $f(.)$ de X. Sans perte de généralité, on supposera ici aussi que $\mathcal{C} = [0, 1]$. Dans la suite, on dénotera par $\theta_0 \in \Theta$ la "vraie" valeur (inconnue) du paramètre θ, de sorte que tester l'hypothèse nulle (4.1) restreinte à $\mathcal{C}$ équivaut à tester :

$$H_0 : \exists \theta_0 \in \Theta \text{ tel que } F(y|x) = F_0(y|x;\theta_0) \;\; \forall y \in \mathbb{R}, \forall x \in \mathcal{C}. \tag{4.3}$$

Notons aussi que, dans ce qui suit, on suppose que X ne contient pas d'information supplémentaire sur la "vraie" valeur θ_0 de θ. Ceci correspond au cas où la variable X est dite "exogène" pour θ, selon la terminologie de la Définition 1.4.9 de Gouriéroux et Montfort (1989, p30).

L'adaptation de notre procédure de test à ces nouvelles conditions nécessite certaines notations et hypothèses que nous introduisons dans la section suivante.

4.2 Notations et Hypothèses

Pour les raisons données à la Section 3.2, on utilise un ajustement polynômial local d'ordre p impair pour l'estimation de $F(y|x)$. Notons de plus que dans tout ce qui va suivre, M est une constante majorante pouvant varier d'une ligne à l'autre. On pose alors les conditions suivantes, en plus de celles déjà introduites à la Section 1.3 :

- $(F_0^*.y.1)$ Pour tout $x \in [-1, 2]$ et tout θ dans un voisinage $V(\theta_0)$ de θ_0, la fonction qui à y associe $F_0(y|x;\theta)$ est continûment différentiable en y, ce qui assure l'existence de la densité conditionnelle $f_0(y|x;\theta)$ de Y sachant $\{X = x\}$.

- $(F_0^{*(k)}.x.1)$ Pour tout $y \in \mathbb{R}$ et tout θ dans un voisinage $V(\theta_0)$ de θ_0, la fonction qui à x associe $F_0(y|x;\theta)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage $V(x)$ de $x \in]-1, 2[$.

- $(F_0^{*(k)}.x.2)$ On suppose $(F_0^{*(k)}.x.1)$ et on suppose de plus que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, pour tout u dans un voisinage $V(x)$ de $x \in]-1, 2[$ et tout $y \in \mathbb{R}$,
$$\sup_{\theta \in V(\theta_0)} \sup_{u \in V(x)} \sup_{y \in \mathbb{R}} \left| F_0^{(k)}(y|x;\theta) \right| \leq M,$$
où $F_0^{(k)}(y|x;\theta)$ est la $k^{ième}$ dérivée de la fonction $F_0(y|x;\theta)$ par rapport à x.

Notons que ces 3 conditions remplaceront maintenant les conditions $(F_0.y.1)$, $(F_0^{(k)}.x.1)$ et $(F_0^{(k)}.x.2)$ de la Section 1.3.2.

– $(F_0.\theta.1)$ Pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $F_0(y|x;\theta)$ est deux fois continûment différentiable pour tout θ dans un voisinage $V(\theta_0)$ de θ_0.

– $(F_0.\theta.2)$ On suppose $(F_0.\theta.1)$ et on suppose que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$,

$$\left|\left|\dot{F}_0(y|x;\theta)\right|\right| \leq M \text{ et } \left|\left|\ddot{F}_0(y|x;\theta)\right|\right|_F \leq M,$$

où $\dot{F}_0(y|x;\theta)$ et $\ddot{F}_0(y|x;\theta)$ désignent respectivement les différentielles d'ordre 1 et 2 de la fonction $F_0(y|x;\theta)$ par rapport à θ, où $||.||$ désigne la norme euclidienne associée à l'espace paramétrique Θ et où $||.||_F$ désigne la norme matricielle de Frobénius définie, pour toute matrice réelle A, par :

$$||A||_F = \sqrt{trace(A^T A)}.$$

– $(f_0.\theta.1)$ Pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $f_0(y|x;\theta)$ est continûment différentiable pour tout θ dans un voisinage $V(\theta_0)$ de θ_0.

– $(f_0.\theta.2)$ On suppose $(f_0.\theta.1)$ et on suppose de plus qu'il existe une fonction $h(x,y)$ positive vérifiant $\int_{\mathbb{R}} h(x,y)dy \leq M \ \forall x \in [-1,2]$ et telle que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$,

$$\left|\left|\dot{f}_0(y|x;\theta)\right|\right| \leq h(x,y),$$

où $\dot{f}_0(y|x;\theta)$ désigne la différentielle de la fonction $f_0(y|x;\theta)$ par rapport à θ.

Notons que la condition $(f_0.\theta.2)$ est une condition “classique” en théorie de l'estimation.

4.3 La procédure de test

La forme de la statistique du test d'adéquation donnée en (1.42) fait intervenir de façon explicite l'expression de la distribution $F_0(y|x)$ testée sous l'hypothèse nulle H_0. Le problème est qu'on a maintenant une famille paramétrée de distributions où la vraie valeur θ_0 du paramètre θ est inconnue. Une solution consiste alors à remplacer ce paramètre inconnu par un estimateur convergent. Plus précisément, si on remplace dans (1.42), $F_0(y|x)$ par $F_0(y|x;\hat{\theta})$, où $\hat{\theta}$ est un estimateur convergent de θ_0, la statistique de test s'écrit :

$$T = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\hat{\theta})\right]^2 w(x)w_x(y)f_0(y|x;\hat{\theta})dydx. \tag{4.4}$$

Cependant, ce changement au sein de la statistique de test est susceptible de modifier son comportement asymptotique et c'est ce point que nous allons étudier ici. Concernant le choix de l'estimateur $\hat{\theta}$ de θ_0, la théorie "classique" de l'estimation statistique propose plusieurs méthodes pour générer des estimateurs convergents d'un paramètre inconnu θ_0. On choisit ici d'utiliser un estimateur $\hat{\theta}$ de θ_0 qui vérifie les hypothèses :

- $(E.1)$ $\sqrt{n}(\hat{\theta} - \theta_0) = O_p(1)$.

- $(E.2)$ $\hat{\theta} \xrightarrow{ps} \theta_0$,

où $\xrightarrow{ps}$ désigne la convergence presque sûre. Concernant $(E.1)$, on dira que $\hat{\theta}$ est $\sqrt{n}$-convergent pour θ_0. Cette propriété, partagée par la plupart des estimateurs classiques, dont par exemple l'estimateur du maximum de vraisemblance de θ_0, a son importance ici.

Notons aussi que sous l'hypothèse $(F_0.\theta.2)$, la fonction qui à θ associe $F_0(y|x;\theta)$ est (localement) Lipschitzienne dans un voisinage $V(\theta_0)$ de θ_0. En effet, si $\theta_1 \in V(\theta_0)$, par un développement de Taylor d'ordre 1, il existe $\theta^{**} \in (\theta_0, \theta_1)$ tel que :

$$F_0(y|x;\theta_1) = F_0(y|x;\theta_0) + (\theta_1 - \theta_0)^T \dot{F}_0(y|x;\theta^{**}),$$

de sorte que, pour tout $x \in [-1, 2]$ et tout $y \in \mathbb{R}$, on a par Cauchy-Schwarz,

$$\begin{aligned}|F_0(y|x;\theta_1) - F_0(y|x;\theta_0)| &\leq ||\theta_1 - \theta_0|| \times ||\dot{F}_0(y|x;\theta^{**})||,\\ &\leq M||\theta_1 - \theta_0||,\end{aligned}$$

sous l'hypothèse $(F_0.\theta.2)$. En particulier, par $(E.2)$, $\hat{\theta} \in V(\theta_0)$ avec la probabilité 1, de sorte que, si n est suffisamment grand et avec probabilité 1,

$$|F_0(y|x;\hat{\theta}) - F_0(y|x;\theta_0)| \leq M||\hat{\theta} - \theta_0||. \tag{4.5}$$

Signalons pour finir que dans la suite, on suppose que n est toujours suffisamment grand pour que l'on puisse supposer que $\hat{\theta} \in V(\theta_0)$ avec probabilité 1.

4.4 Décomposition de la statistique de test

Pour simplifier l'étude du comportement asymptotique de la statistique de test (4.4), supposons pour le moment que $\theta \in \Theta \subset \mathbb{R}$. Donnons-en ensuite une décomposition, ce qui permettra d'étudier plus facilement ce comportement asymptotique. Tout d'abord,

$$\hat{F}_n(y|x) - F_0(y|x;\hat{\theta}) = \hat{F}_n(y|x) - F_0(y|x;\theta_0) + F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta}),$$

de sorte que,

$$\begin{aligned}\Big[\hat{F}_n(y|x) - F_0(y|x;\hat{\theta})\Big]^2 &= \Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]^2 + \Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]^2 \\ &\quad + 2\Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]\Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]. \qquad (4.6)\end{aligned}$$

De plus, par l'hypothèse $(f_0.\theta.1)$ et le fait que $\hat{\theta}$ soit avec probabilité 1 dans un voisinage de θ_0, on peut invoquer le théorème de la moyenne pour annoncer qu'il existe un $\theta^* \in (\theta_0, \hat{\theta})$ tel que, pour n suffisamment grand :

$$f_0(y|x;\hat{\theta}) = f_0(y|x;\theta_0) + (\hat{\theta} - \theta_0)\dot{f}_0(y|x;\theta^*). \qquad (4.7)$$

Si on substitue (4.6) et (4.7) dans l'expression de la statistique de test (4.4), on obtient la décomposition :

$$T = T_1 + T_2 + T_3 + T_4 + T_5 + T_6, \qquad (4.8)$$

avec

$$T_1 = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]^2 w(x)w_x(y)f_0(y|x;\theta_0)dydx, \qquad (4.9)$$

$$\begin{aligned}T_2 &= n\sqrt{h}(\hat{\theta} - \theta_0)\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]^2 w(x)w_x(y) \\ &\quad \times \dot{f}_0(y|x;\theta^*)dydx, \qquad (4.10)\end{aligned}$$

$$T_3 = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]^2 w(x)w_x(y)f_0(y|x;\theta_0)dydx, \qquad (4.11)$$

$$\begin{aligned}T_4 &= n\sqrt{h}(\hat{\theta} - \theta_0)\int_0^1\int_{\mathbb{R}}\Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]^2 w(x)w_x(y) \\ &\quad \times \dot{f}_0(y|x;\theta^*)dydx, \qquad (4.12)\end{aligned}$$

$$\begin{aligned} T_5 &= 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\right] \\ &\quad \times w(x)w_x(y)f_0(y|x;\theta_0)dydx, \end{aligned} \tag{4.13}$$

$$\begin{aligned} T_6 &= 2n\sqrt{h}(\hat{\theta}-\theta_0)\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\right] \\ &\quad \times w(x)w_x(y)\dot{f}_0(y|x;\theta^*)dydx. \end{aligned} \tag{4.14}$$

Ainsi, le comportement asymptotique de la statistique de test T, donnée en (4.4) va dépendre du comportement asymptotique de la somme des 6 statistiques T_1, T_2, T_3, T_4, T_5 et T_6. Nous allons montrer que seul le comportement asymptotique de T_1 importe.

A l'instar de la remarque faite à la Section 1.3, on va utiliser dans ce chapitre la notation $E_{F_0}(.)$ pour désigner l'espérance associée à la loi conjointe $F_0(.\,,\,.\,;\,\theta)$ et $E_{F_0}(.\,|X)$ dénotera l'espérance associée à la loi conditionnelle $F_0(.\,|\,.\,;\,\theta)$. Comme dans les précédents chapitres, $E_F(X)$ désignera l'espérance de X, associée à la loi marginale $F(.)$ de X, de densité $f(.)$. Notons aussi que ces mêmes notations seront transposées à celles concernant la variance.

4.5 Etude du comportement asymptotique de la statistique de test

Etude de T_1

La statistique T_1 est analogue à la statistique de test T, introduite en (1.42) et dont le comportement asymptotique est étudié au Chapitre 3 avec pour seules différences, l'utilisation de $F_0(y|x;\theta_0)$ à la place de $F_0(y|x)$ et de $f_0(y|x;\theta_0)$ à la place de $f_0(y|x)$. Cette différence n'en est pas vraiment une et du coup, son comportement asymptotique se déduit automatiquement de celui obtenu au Théorème 3.5.3, que l'on rappelle dans la proposition suivante :

Proposition 4.5.1 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$ et $(F_0^{*(p+1)}.x.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$ et $(H.3)$, on a :*

$$\frac{T_1 - h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} \xrightarrow{\mathcal{L}} N(0,1), \tag{4.15}$$

avec

$$a_{\theta_0} = K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}} \frac{F_0(y|x;\theta_0)(1-F_0(y|x;\theta_0))}{f(x)} w_x(y)f_0(y|x;\theta_0)dy\ w(x)dx, \qquad (4.16)$$

et

$$\sigma^2_{\theta_0} = 2K^{*^{(4)}}(0)\int_0^1\int_{\mathbb{R}^2}\left[\frac{F_0(y\wedge y'|x';\theta_0)-F_0(y|x';\theta_0)F_0(y'|x';\theta_0)}{f(x')}\right]^2$$

$$\times\ w^2(x')w_{x'}(y)w_{x'}(y')f_0(y|x';\theta_0)f_0(y'|x';\theta_0)dydy'dx'. \qquad (4.17)$$

Preuve On recycle la preuve du Théorème 3.5.3 puisque les hypothèses $(F_0^*.y.1)$ et $(F_0^{*^{(p+1)}}.x.2)$ impliquent les anciennes versions de celles-ci.

Etude de T_2

Proposition 4.5.2 *Supposons qu'en plus de* $(f_0.\theta.2)$, $F_0(y|x;\theta)$ *vérifie les hypothèses de la Proposition 4.5.1. Alors, sous les hypothèses* $(E.1)$ *et* $(E.2)$, *on a :*

$$T_2 = o_p(1). \qquad (4.18)$$

Preuve Sous les hypothèses $(f_0.\theta.2)$, $(P.0)$ et $(P.1)$ et puisque par $(E.2)$, $\theta^*\in V(\theta_0)$ avec probabilité 1, si n est suffisamment grand,

$$\begin{aligned} |T_2| &\leq n\sqrt{h}|\hat{\theta}-\theta_0|\times\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\right]^2 w(x)w_x(y)|\dot{f}_0(y|x;\theta^*)|dydx,\\ &\leq |\hat{\theta}-\theta_0|\times n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\right]^2 w(x)w_x(y)h(x,y)dydx,\\ &= |\hat{\theta}-\theta_0|\times V, \end{aligned} \qquad (4.19)$$

où

$$V = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\right]^2 w(x)w_x(y)h(x,y)dydx. \qquad (4.20)$$

Le Lemme suivant donne le comportement asymptotique de la statistique V.

Lemme 4.5.3 *Sous les hypothèses de la Proposition* 4.5.2, *on a :*

$$V = O\left(h^{-\frac{1}{2}}\right) + O_p(1). \qquad (4.21)$$

Preuve La statistique V est semblable à la statistique T_1 avec pour seule différence l'utilisation de la fonction $h(x,y)$ à la place de la densité $f_0(y|x,\theta_0)$. La convergence de T_1 décrite à la Proposition 4.5.1 est obtenue de la preuve du Théorème 3.5.3 en utilisant en particulier le fait que $\int_{\mathbb{R}} f_0(y|x,\theta_0)dy < +\infty$ pour tout $x \in [-1,2]$ puisqu'il s'agit d'une densité. Or, la condition $(f_0.\theta.2)$ impose que $\int_{\mathbb{R}} h(x,y)dy < +\infty$ pour tout $x \in [-1,2]$ de sorte qu'on obtient la convergence (4.21) pour la statistique V. ♣

Revenons maintenant à la preuve de la Proposition 4.5.2. Sous l'hypothèse $(E.1)$, $\hat{\theta}-\theta_0 = O_p\left(\frac{1}{\sqrt{n}}\right)$ de sorte que :

$$
\begin{aligned}
|T_2| &\leq |\hat{\theta}-\theta_0| \times \left[O\left(h^{-\frac{1}{2}}\right) + O_p(1)\right], \\
&= O_p\left(\frac{1}{\sqrt{nh}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right), \\
&= o_p(1),
\end{aligned}
$$

sous l'hypothèse $(H.1)$. ♣

Etude de T_3

Proposition 4.5.4 *Supposons que $F_0(y|x;\theta)$ vérifie l'hypothèse $(F_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.1)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_3 = o_p(1). \tag{4.22}$$

Preuve Suite à la remarque (4.5),

$$
\begin{aligned}
0 \leq T_3 &= n\sqrt{h}\int_0^1 \int_{\mathbb{R}} \left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]^2 w(x)w_x(y)f_0(y|x;\theta_0)dydx, \\
&\leq M\sqrt{h} \times n(\hat{\theta}-\theta_0)^2.
\end{aligned}
$$

Sous l'hypothèse $(E.1)$, on obtient :

$$T_3 \leq M\sqrt{h}O_p(1) = o_p(1).♣ \tag{4.23}$$

Etude de T_4

Proposition 4.5.5 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(H.0)$, $(P.0)$, $(P.1)$, $(E.1)$ et $(E.2)$, on a :*

$$T_4 = o_p(1). \tag{4.24}$$

Preuve Sous les hypothèses $(F_0.\theta.2)$, $(f_0.\theta.2)$, $(P.0)$ et $(P.1)$ et suite à la remarque (4.5),

$$\begin{aligned}
|T_4| &= |n\sqrt{h}(\hat{\theta}-\theta_0)\int_0^1\int_{\mathbb{R}}\left[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\right]^2 w(x)w_x(y)\dot{f}_0(y|x;\theta^*)dydx|,\\
&\leq M\sqrt{h}|\hat{\theta}-\theta_0|\times n(\hat{\theta}-\theta_0)^2\int_0^1\int_{\mathbb{R}}|\dot{f}_0(y|x;\theta^*)|dydx,\\
&\leq M\sqrt{h}|\hat{\theta}-\theta_0|\times n(\hat{\theta}-\theta_0)^2\int_0^1\int_{\mathbb{R}}h(x,y)dydx,\\
&\leq M\sqrt{h}|\hat{\theta}-\theta_0|\times n(\hat{\theta}-\theta_0)^2,\\
&\leq M\sqrt{h}o_p(1)O_p(1) = o_p(1),
\end{aligned}$$

sous l'hypothèse $(H.0)$.♣

Etude de T_5

Proposition 4.5.6 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$ et $(F_0.\theta.2)$. Alors sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$T_5 = o_p(1). \tag{4.25}$$

Preuve Sous l'hypothèse $(F_0.\theta.2)$, la fonction qui à θ associe $F_0(y|x;\theta)$ est deux fois continûment dérivable en θ. Il existe donc un $\theta^{**}\in(\theta_0,\hat{\theta})$ tel que :

$$F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta}) = (\theta_0-\hat{\theta})\dot{F}_0(y|x;\theta_0)+\frac{1}{2}(\hat{\theta}-\theta_0)^2\ddot{F}_0(y|x;\theta^{**}).$$

Si on substitue ce résultat dans l'expression (4.13) de T_5, on obtient :

$$T_5 = T_5^* + T_5^{**}, \tag{4.26}$$

où

$$T_5^* = 2\sqrt{n}(\theta_0 - \hat{\theta}) \times \sqrt{nh} \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right] \dot{F}_0(y|x;\theta_0) \times w(x) w_x(y) f_0(y|x;\theta_0) dy dx, \quad (4.27)$$

et

$$T_5^{**} = \sqrt{h} \times n(\hat{\theta} - \theta_0)^2 \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right] \ddot{F}_0(y|x;\theta^{**}) \times w(x) w_x(y) f_0(y|x;\theta_0) dy dx. \quad (4.28)$$

Sous les hypothèses $(F_0.\theta.2)$, $(P.0)$, $(P.1)$, $(E.1)$ et $(E.2)$, on déduit directement que

$$\begin{aligned} |T_5^{**}| &\leq \sqrt{h} \times n(\hat{\theta} - \theta_0)^2 \int_0^1 \int_{\mathbb{R}} \left|\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right| \times |\ddot{F}_0(y|x;\theta^{**})| \\ &\quad \times w(x) w_x(y) f_0(y|x;\theta_0) dy dx, \\ &\leq M\sqrt{h} \times n(\hat{\theta} - \theta_0)^2 \int_0^1 \int_{\mathbb{R}} f_0(y|x;\theta_0) dy dx, \\ &= O_p\left(\sqrt{h}\right) = o_p(1), \end{aligned} \quad (4.29)$$

sous l'hypothèse $(H.0)$. L'étude du comportement asymptotique de T_5^* nécessite d'introduire la proposition suivante.

Lemme 4.5.7 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$ et $(F_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$ et $(H.3)$, on a :*

$$\sqrt{nh} \int_0^1 \int_{\mathbb{R}} (\hat{F}_n(y|x) - F_0(y|x;\theta_0)) \dot{F}_0(y|x;\theta_0) w(x) w_x(y) f_0(y|x;\theta_0) dy dx = o_p(1). \quad (4.30)$$

Preuve Voir Annexe F. ♣

Le Lemme 4.5.7 et l'hypothèse $(E.1)$ permettent de déduire que :

$$T_5^* = o_p(1), \quad (4.31)$$

de sorte que, associé à (4.29), on conclut :

$$T_5 = o_p(1).♣. \quad (4.32)$$

Etude de T_6

Proposition 4.5.8 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.1)$, $(H.0)$, $(E.1)$ et $(E.2)$, on a :*

$$T_6 = o_p(1). \tag{4.33}$$

Preuve Sous les hypothèses $(F_0.\theta.2)$, $(f_0.\theta.2)$, $(P.0)$, $(P.1)$ et $(E.2)$ et suite à la remarque (4.5),

$$\begin{aligned}
|T_6| &\leq 2n\sqrt{h}|\hat{\theta}-\theta_0| \int_0^1 \int_{\mathbb{R}} |\hat{F}_n(y|x) - F_0(y|x;\theta_0)| \times |F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})| \\
&\quad \times w(x)w_x(y) \times |\dot{f}_0(y|x;\theta^*)|dydx, \\
&\leq Mn\sqrt{h}|\hat{\theta}-\theta_0| \times \int_0^1 \int_{\mathbb{R}} |F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})| \times h(x,y)dydx, \\
&\leq M\sqrt{h} \times n(\hat{\theta}-\theta_0)^2 \int_0^1 \int_{\mathbb{R}} h(x,y)dydx, \\
&= O_p\left(\sqrt{h}\right) = o_p(1),
\end{aligned}$$

sous les hypothèses $(E.1)$ et $(H.0)$.♣

Les Propositions 4.5.1 à 4.5.8 montrent que seul le comportement asymptotique de T_1 importe dans le comportement asymptotique de la statistique de test. On en déduit alors le théorème suivant :

Théorème 4.5.9 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$\frac{T - h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} \xrightarrow{\mathcal{L}} N(0,1), \tag{4.34}$$

où les quantités a_{θ_0} et $\sigma^2_{\theta_0}$ sont données respectivement en (4.16) et (4.17).

Preuve On avait obtenu en (4.8) la décomposition :

$$T = T_1 + T_2 + T_3 + T_4 + T_5 + T_6,$$

de sorte que

$$\begin{aligned}\frac{T-h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} &= \frac{T_1-h^{-\frac{1}{2}}a_{\theta_0}+T_2+T_3+T_4+T_5+T_6}{\sigma_{\theta_0}}, \\ &= \frac{T_1-h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}}+o_p(1). \qquad (4.35)\end{aligned}$$

La Proposition 4.5.1 permet de conclure directement quant à la convergence en loi de la statistique de test (4.4) sous l'hypothèse H_0 donnée en (4.3). ♣

Ce comportement asymptotiquement normal est donc le même dans le présent contexte que lorsque la fonction à tester sous H_0 ne comporte aucun paramètre inconnu. Du Théorème 4.5.9 découle alors une stratégie de test de niveau asymptotique α pour l'hypothèse nulle donnée en (4.3) : on rejettera H_0 si

$$\frac{T-h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Notons cependant que dans cette expression, les quantités a_{θ_0} et σ_{θ_0} font intervenir la valeur de θ_0 qui est inconnue. Cette stratégie de test est donc inutilisable en pratique. Une possibilité est de remplacer θ_0 par un estimateur convergent, par exemple ici, $\hat{\theta}$ vérifiant les hypothèses $(E.1)$ et $(E.2)$. On est donc amené à considérer la statistique de test

$$\frac{T-h^{-\frac{1}{2}}a_{\hat{\theta}}}{\sigma_{\hat{\theta}}}.$$

Il se pose alors la question de savoir si un test basé sur cette statistique est asymptotiquement valide. Or,

$$\frac{T-h^{-\frac{1}{2}}a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} = \left[\frac{T-h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}}+\frac{h^{-\frac{1}{2}}(a_{\theta_0}-a_{\hat{\theta}})}{\sigma_{\theta_0}}\right]\times\frac{\sigma_{\theta_0}}{\sigma_{\hat{\theta}}}. \qquad (4.36)$$

De plus, sous les hypothèses $(F_0.\theta.1)$ et $(f_0.\theta.1)$, la fonction qui à θ associe

$$R_\theta(x',y,y') = \left[F_0(y\wedge y'|x';\theta)-F_0(y|x';\theta)F_0(y'|x';\theta)\right]^2 f_0(y|x';\theta)f_0(y'|x';\theta),$$

est continue pour tout θ dans un voisinage $V(\theta_0)$ de θ_0. On en déduit que la fonction qui à θ associe

$$\sigma_\theta^2 = 2K^{*(4)}(0)\int_0^1\int_{\mathrm{IR}^2} R_\theta(x',y,y')\frac{w(x)w_x(y)}{f(x)}dydy'dx',$$

est continue pour tout θ dans ce même voisinage et que donc σ_θ l'est aussi. Sous l'hypothèse $(E.2)$, $\hat{\theta}$ appartient à ce voisinage avec probabilité 1 et on obtient :

$$\frac{\sigma_{\theta_0}}{\sigma_{\hat{\theta}}} = 1 + o_p(1).$$

Comme $\sigma_{\theta_0} > 0$, le terme $\dfrac{1}{\sigma_{\theta_0}}$ est majoré et il ne reste donc plus qu'à montrer que $h^{-\frac{1}{2}}(a_{\theta_0} - a_{\hat{\theta}})$ est asymptotiquement négligeable. Sous les hypothèses $(F_0.\theta.1)$ et $(f_0.\theta.1)$, la fonction qui à θ associe a_θ est continue et dérivable sur $(\theta_0, \hat{\theta})$, puisque sous l'hypothèse $(E.2)$, $\hat{\theta}$ est dans un voisinage de θ_0 avec probabilité 1. Par le théorème de la moyenne, il existe donc $\theta_1 \in (\theta_0, \hat{\theta})$ tel que :

$$a_{\theta_0} - a_{\hat{\theta}} = (\theta_0 - \hat{\theta})\dot{a}_{\theta_1}, \tag{4.37}$$

où

$$\dot{a}_{\theta_1} = \frac{\partial a_\theta}{\partial \theta}\Big|_{\theta_1}.$$

Ainsi,

$$\begin{aligned}
\frac{h^{-\frac{1}{2}}(a_{\theta_0} - a_{\hat{\theta}})}{\sigma_{\theta_0}} &= h^{-\frac{1}{2}}(\theta_0 - \hat{\theta})\frac{\dot{a}_{\theta_1}}{\sigma_{\theta_0}}, \\
&= (nh)^{-\frac{1}{2}} \times \sqrt{n}(\theta_0 - \hat{\theta})\frac{\dot{a}_{\theta_1}}{\sigma_{\theta_0}}.
\end{aligned}$$

De plus, la quantité $|\dot{a}_\theta|$ est majorée pour tout θ dans un voisinage $V(\theta_0)$ de θ_0. En effet, sous les hypothèses $(F_0.\theta.1)$ et $(f_0.\theta.1)$, les fonctions qui à θ associent $F_0(y|x;\theta)$ et $f_0(y|x;\theta)$ sont dérivables dans un voisinage $V(\theta_0)$ de θ_0 et sous les hypothèses supplémentaires $(F_0.\theta.2)$, $(f_0.\theta.2)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$, on peut inverser dérivée et intégrale pour obtenir :

$$\begin{aligned}
|\dot{a}_\theta| &= \left|\frac{\partial}{\partial\theta}\left[K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}} \frac{F_0(y|x;\theta)(1 - F_0(y|x;\theta))}{f(x)} w_x(y) f_0(y|x;\theta)dy\ w(x)dx\right]\right|, \\
&\leq K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}}\left|\frac{\partial}{\partial\theta}\big[F_0(y|x;\theta)(1 - F_0(y|x;\theta))f_0(y|x;\theta)\big]\right|\frac{w(x)w_x(y)}{f(x)}dydx, \\
&\leq M\int_0^1\int_{\mathbb{R}}\Big[|\dot{F}_0(y|x;\theta)| \times |1 - 2F_0(y|x;\theta)| \times f_0(y|x;\theta) + F_0(y|x;\theta)(1 - F_0(y|x;\theta)) \\
&\quad \times |\dot{f}_0(y|x;\theta)|\Big]dydx,
\end{aligned}$$

$$\begin{aligned}
&\leq \quad M\left[\int_0^1 \int_{\mathbb{R}} f_0(y|x;\theta)dydx + \int_0^1 \int_{\mathbb{R}} h(x,y)dydx\right], \\
&\leq \quad M. \qquad (4.38)
\end{aligned}$$

Comme θ_1 appartient au voisinage $V(\theta_0)$ de θ_0 si n est suffisamment grand, on en déduit en particulier que :

$$\frac{|\dot{a}_{\theta_1}|}{\sigma_{\theta_0}} \leq M,$$

et on obtient :

$$\begin{aligned}
\frac{h^{-\frac{1}{2}}|a_{\theta_0} - a_{\hat{\theta}}|}{\sigma_{\theta_0}} &= \frac{1}{\sqrt{nh}} \times \sqrt{n}|\theta_0 - \hat{\theta}| \times \frac{|\dot{a}_{\theta_1}|}{\sigma_{\theta_0}}, \\
&\leq \frac{M}{\sqrt{nh}}|\sqrt{n}(\theta_0 - \hat{\theta})|, \\
&= o_p(1), \qquad (4.39)
\end{aligned}$$

sous les hypothèses $(E.1)$ et $(H.1)$. Si on substitue ce résultat dans (4.36), on en déduit que les statistiques $\dfrac{T - h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}}$ et $\dfrac{T - h^{-\frac{1}{2}}a_{\hat{\theta}}}{\sigma_{\hat{\theta}}}$ ont le même comportement asymptotique. On vient donc de montrer le théorème suivant :

Théorème 4.5.10 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$\frac{T - h^{-\frac{1}{2}}a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} \xrightarrow{\mathcal{L}} N(0,1), \qquad (4.40)$$

où

$$a_{\hat{\theta}} = K^{*(2)}(0)\int_0^1 \int_{\mathbb{R}} \frac{F_0(y|x;\hat{\theta})(1 - F_0(y|x;\hat{\theta}))}{f(x)} w_x(y) f_0(y|x;\hat{\theta})dy\; w(x)dx, \qquad (4.41)$$

et

$$\begin{aligned}
\sigma_{\hat{\theta}}^2 &= 2K^{*(4)}(0)\int_0^1 \int_{\mathbb{R}^2} \left[\frac{F_0(y \wedge y'|x';\hat{\theta}) - F_0(y|x';\hat{\theta})F_0(y'|x';\hat{\theta})}{f(x)}\right]^2 \\
&\quad \times w^2(x')w_{x'}(y)w_{x'}(y')f_0(y|x';\hat{\theta})f_0(y'|x';\hat{\theta})dydy'dx'. \qquad (4.42)
\end{aligned}$$

Ainsi, un test de (4.3) de niveau asymptotique α consiste à rejeter H_0 si

$$\frac{T - h^{-\frac{1}{2}} a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Il est rare en pratique que $F_0(y|x;\theta)$ ne comporte qu'un seul paramètre inconnu, comme le montre l'exemple (1.2). Il faut donc regarder ce que devient ce comportement lorsque $\theta \in \Theta \subset \mathbb{R}^d$ avec $d > 1$.

4.6 Cas multidimensionnel où θ est un vecteur de dimension d

4.6.1 Décomposition de la statistique de test

De la même façon qu'à la Section 4.4 on a le développement

$$\begin{aligned}\Big[\hat{F}_n(y|x) - F_0(y|x;\hat{\theta})\Big]^2 &= \Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]^2 + \Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]^2 \\ &\quad + 2\Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big] \\ &\quad \times \Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]. \end{aligned} \tag{4.43}$$

Si n est suffisamment grand, il existe aussi avec probabilité 1 un $\theta^* \in (\theta_0, \hat{\theta})$ (c'est à dire sur la droite reliant θ_0 et $\hat{\theta}$) tel que :

$$f_0(y|x;\hat{\theta}) = f_0(y|x;\theta_0) + (\hat{\theta} - \theta_0)^T \dot{f}_0(y|x;\theta^*), \tag{4.44}$$

où on signale que maintenant

$$\dot{f}_0(y|x;\theta^*) = \begin{pmatrix} \dfrac{\partial f_0(y|x;\theta^*)}{\partial \theta_1} \\ \vdots \\ \dfrac{\partial f_0(y|x;\theta^*)}{\partial \theta_d} \end{pmatrix}. \tag{4.45}$$

En suivant les étapes données à la Section 4.4, on obtient

$$T = T_1 + T_2^* + T_3 + T_4^* + T_5 + T_6^*, \tag{4.46}$$

avec T_1, T_3 et T_5 données respectivement en (4.9), (4.11) et (4.13) et

$$\begin{aligned} T_2^* &= n\sqrt{h}(\hat{\theta}-\theta_0)^T \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 w(x)w_x(y) \\ &\quad \times \dot{f}_0(y|x;\theta^*)dydx, \end{aligned} \tag{4.47}$$

$$\begin{aligned} T_4^* &= n\sqrt{h}(\hat{\theta}-\theta_0)^T \int_0^1 \int_{\mathbb{R}} \left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]^2 w(x)w_x(y) \\ &\quad \times \dot{f}_0(y|x;\theta^*)dydx, \end{aligned} \tag{4.48}$$

et

$$\begin{aligned} T_6^* &= 2n\sqrt{h}(\hat{\theta}-\theta_0)^T \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right] \\ &\quad \times w(x)w_x(y)\dot{f}_0(y|x;\theta^*)dydx. \end{aligned} \tag{4.49}$$

4.6.2 Comportement asymptotique de la statistique de test

L'adaptation des résultats obtenus dans le cas univarié permet directement de conclure quant au comportement asymptotique du test, donné dans le théorème suivant :

Théorème 4.6.1 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$\frac{T - h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} \xrightarrow{\mathcal{L}} N(0,1), \tag{4.50}$$

où les expressions de a_{θ_0} et $\sigma^2_{\theta_0}$ sont données respectivement en (4.16) et (4.17).

Comme pour le cas univarié, les expressions a_{θ_0} et $\sigma^2_{\theta_0}$ données dans le Théorème 4.6.1 comportent le paramètre inconnu θ_0. Ce résultat n'est donc pas directement utilisable en pratique. Si on utilise cependant le même raisonnement que pour le Théorème 4.5.10, on en déduit le théorème suivant dont la preuve est escamotée :

Théorème 4.6.2 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(E.1)$,*

$(E.2)$ et $(H.3)$, on a :

$$\frac{T - h^{-\frac{1}{2}} a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} \xrightarrow{\mathcal{L}} N(0,1), \tag{4.51}$$

où les expressions de $a_{\hat{\theta}}$ et *$\sigma^2_{\hat{\theta}}$ sont données respectivement en* (4.41) *et* (4.42).

Ainsi, un test de l'hypothèse nulle donnée en (4.3) de niveau asymptotique α consiste à rejeter H_0 si

$$\frac{T - h^{-\frac{1}{2}} a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} > z_{1-\alpha},$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

4.7 Cas particulier où la fonction de poids $w_x(y)$ dépend de paramètre(s) inconnu(s)

Jusqu'à présent, on a considéré que la fonction de poids $w_x(y)$ qui intervient dans la statistique de test (4.4) n'est pas fonction de paramètre(s) inconnu(s). Comme en pratique il n'y a pas de raison de ne pas autoriser $w_x(y)$ à dépendre de θ, nous étudions maintenant cette possibilité. Ce serait par exemple le cas si on souhaite imiter la statistique d'Anderson-Darling (Anderson et Darling, 1954) en considérant une fonction de poids du type $F_0^{-1}(y|x;\theta)(1-F_0(y|x;\theta))^{-1}$. Supposons donc que la fonction de poids $w_x(y)$ comporte un ou plusieurs paramètres inconnus que l'on retrouve dans le vecteur θ et dénotons alors cette fonction par $w(y|x;\theta)$. Pour simplifier les calculs, on va supposer que θ est de dimension 1. Supposons aussi :

- $(P.2)$ Pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $w(y|x;\theta)$ est positive, bornée et continûment différentiable en θ dans un voisinage $V(\theta_0)$ de θ_0. On suppose de plus que pour tout θ appartenant à $V(\theta_0)$, pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

  $$|\dot{w}(y|x;\theta)| \leq M,$$

  où $\dot{w}(y|x;\theta)$ est la différentielle de la fonction $w(y|x;\theta)$ par rapport à θ.

Notons que sous l'hypothèse $(P.2)$, la fonction qui à θ associe $w(y|x;\theta)$ est (localement) Lipschitzienne dans un voisinage $V(\theta_0)$ de θ_0. En particulier, par le même raisonnement que celui menant à (4.5), pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

$$|w(y|x;\hat{\theta}) - w(y|x;\theta_0)| \quad \leq \quad M|\hat{\theta} - \theta_0|. \tag{4.52}$$

Dans ce contexte, la statistique de test (4.4) s'écrit alors :

$$T = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\hat{\theta})\right]^2 w(x)w(y|x;\hat{\theta})f_0(y|x;\hat{\theta})dydx. \tag{4.53}$$

Nous allons montrer que cette modification de la statistique de test ne change en rien son comportement asymptotique. Donnons en d'abord une décomposition. Notons que

$$w(y|x;\hat{\theta}) = w(y|x;\theta_0) + \left[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\right], \tag{4.54}$$

de sorte que, par le même cheminement qu'à la Section 4.4, on arrive à la décomposition suivante pour la statistique de test :

$$T = \sum_{i=1}^{12} T_i^{**}, \tag{4.55}$$

avec

$$T_1^{**} = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 w(x)w(y|x;\theta_0)f_0(y|x;\theta_0)dydx, \tag{4.56}$$

$$\begin{aligned} T_2^{**} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 w(x)w(y|x;\theta_0) \\ &\quad \times (\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx, \end{aligned} \tag{4.57}$$

$$\begin{aligned} T_3^{**} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]^2 w(x)w(y|x;\theta_0) \\ &\quad \times f_0(y|x;\theta_0)dydx, \end{aligned} \tag{4.58}$$

$$\begin{aligned} T_4^{**} &= n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]^2 w(x)w(y|x;\theta_0) \\ &\quad \times (\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx, \end{aligned} \tag{4.59}$$

$$\begin{aligned} T_5^{**} &= 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]w(x) \\ &\quad \times w(y|x;\theta_0)f_0(y|x;\theta_0)dydx, \end{aligned} \tag{4.60}$$

$$T_6^{**} = 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]w(x)$$

$$\times\, w(y|x;\theta_0)(\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx, \tag{4.61}$$

$$T_7^{**} = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\Big]^2\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]$$

$$\times\, w(x)f_0(y|x;\theta_0)dydx, \tag{4.62}$$

$$T_8^{**} = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\Big]^2\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]$$

$$\times\, w(x)(\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx, \tag{4.63}$$

$$T_9^{**} = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\Big]^2\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]$$

$$\times\, w(x)f_0(y|x;\theta_0)dydx, \tag{4.64}$$

$$T_{10}^{**} = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\Big]^2\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]$$

$$\times\, w(x)(\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx, \tag{4.65}$$

$$T_{11}^{**} = 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\Big]\times\Big[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\Big]$$

$$\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]w(x)f_0(y|x;\theta_0)dydx, \tag{4.66}$$

$$T_{12}^{**} = 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x)-F_0(y|x;\theta_0)\Big]\times\Big[F_0(y|x;\theta_0)-F_0(y|x;\hat{\theta})\Big]$$

$$\times\Big[w(y|x;\hat{\theta})-w(y|x;\theta_0)\Big]w(x)(\theta_0-\hat{\theta})\dot{f}_0(y|x;\theta^*)dydx. \tag{4.67}$$

Les statistiques T_1^{**} à T_6^{**} sont semblables aux 6 statistiques intervenant dans la décomposition de la statistique de test (4.8). En effet, la seule différence avec celles-ci est l'utilisation de la fonction de poids $w(y|x;\theta_0)$ à la place de $w_x(y)$. On peut donc directement recycler les preuves des Propositions 4.5.1 à 4.5.8 pour obtenir leur comportement asymptotique. Regardons à présent ce qu'il advient des six dernières statistiques.

Etude du comportement asymptotique des statistiques T_7^{**} à T_{12}^{**}

Proposition 4.7.1 *Sous les hypothèses de la Proposition* 4.5.1 *et sous l'hypothèse supplémentaire* $(P.2)$, *on a :*

$$T_7^{**} = o_p(1). \tag{4.68}$$

Preuve Sous les hypothèses $(P.0)$, $(P.2)$ et suite à la remarque (4.52),

$$\begin{aligned}
|T_7^{**}| &\leq n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 \times |w(y|x;\hat{\theta}) - w(y|x;\theta_0)| \\
&\quad \times w(x)f_0(y|x;\theta_0)dydx, \\
&\leq M|\hat{\theta}-\theta_0| \times n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 w(x)f_0(y|x;\theta_0)dydx, \\
&= M|\hat{\theta}-\theta_0| \times W,
\end{aligned}$$

où on a posé

$$W = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 w(x)f_0(y|x;\theta_0)dydx.$$

La statistique W est semblable à la statistique T_1 dans le cas où on choisit une fonction de poids $w_x(y)$ égale à 1. Du coup, son comportement asymptotique se déduit directement de celui déjà obtenu pour T_1 et sous les hypothèses de la Proposition 4.5.1, on a en particulier :

$$W = O\left(h^{-\frac{1}{2}}\right) + O_p(1).$$

Enfin, sous l'hypothèse $(E.1)$, $\hat{\theta} - \theta_0 = O_p\left(\frac{1}{\sqrt{n}}\right)$ de sorte que :

$$\begin{aligned}
|T_7^{**}| &\leq M|\hat{\theta}-\theta_0| \times \left[O\left(h^{-\frac{1}{2}}\right) + O_p(1)\right], \\
&= O_p\left(\frac{1}{\sqrt{nh}}\right) + O_p\left(\frac{1}{\sqrt{n}}\right), \\
&= o_p(1),
\end{aligned} \tag{4.69}$$

sous l'hypothèse $(H.1)$. ♣

Proposition 4.7.2 *Supposons que $F_0(y|x;\theta)$ vérifie l'hypothèse $(f_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_8^{**} = o_p(1). \tag{4.70}$$

Preuve Suite à la remarque (4.52),

$$\begin{aligned}
|T_8^{**}| &= |n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 \times \left[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\right] \\
&\quad \times w(x) \times (\hat{\theta} - \theta_0) \times \dot{f}_0(y|x;\theta^*)dydx|, \\
&\leq Mn\sqrt{h}(\hat{\theta} - \theta_0)^2\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]^2 \times h(x,y)dydx, \\
&\leq M \times \sqrt{h}O_p(1),
\end{aligned}$$

de sorte que par $(H.0)$,

$$|T_8^{**}| = o_p(1).\clubsuit \tag{4.71}$$

Proposition 4.7.3 *Supposons que $F_0(y|x;\theta)$ vérifie l'hypothèse $(F_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_9^{**} = o_p(1). \tag{4.72}$$

Preuve Suite aux remarques (4.5) et (4.52),

$$\begin{aligned}
|T_9^{**}| &= |n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right]^2 \times \left[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\right] \\
&\quad \times w(x)f_0(y|x;\theta_0)dydx|, \\
&\leq Mn\sqrt{h}|\hat{\theta} - \theta_0|^3 \times \int_0^1\int_{\mathbb{R}} f_0(y|x;\theta_0)dydx, \\
&\leq o_p(1).\clubsuit
\end{aligned}$$

Proposition 4.7.4 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_{10}^{**} = o_p(1). \tag{4.73}$$

Preuve Suite aux remarques (4.5) et (4.52),

$$\begin{aligned}
|T_{10}^{**}| &= |n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big]^2 \times \Big[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\Big] \\
&\quad \times w(x)(\hat{\theta}-\theta_0)\dot{f}_0(y|x;\theta^*)dydx|, \\
&\leq Mn\sqrt{h}(\hat{\theta}-\theta_0)^4\int_0^1\int_{\mathbb{R}}|\dot{f}_0(y|x;\theta^*)|dydx, \\
&\leq Mn\sqrt{h}(\hat{\theta}-\theta_0)^4 = o_p(1).\clubsuit
\end{aligned}$$

Proposition 4.7.5 *Supposons que $F_0(y|x;\theta)$ vérifie l'hypothèse $(F_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_{11}^{**} = o_p(1). \tag{4.74}$$

Preuve Suite aux remarques (4.5) et (4.52),

$$\begin{aligned}
|T_{11}^{**}| &= |2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\Big[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\Big]\Big[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\Big] \\
&\quad \times \Big[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\Big]w(x)f_0(y|x;\theta_0)dydx|, \\
&\leq 2n\sqrt{h}\int_0^1\int_{\mathbb{R}}|\hat{F}_n(y|x) - F_0(y|x;\theta_0)| \times |F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})| \\
&\quad \times |w(y|x;\hat{\theta}) - w(y|x;\theta_0)| \times w(x)f_0(y|x;\theta_0)dydx, \\
&\leq Mn\sqrt{h}(\hat{\theta}-\theta_0)^2\int_0^1\int_{\mathbb{R}}|\hat{F}_n(y|x) - F_0(y|x;\theta_0)| \times f_0(y|x;\theta_0)dydx, \\
&= M\sqrt{h}\left(\sqrt{n}(\hat{\theta}-\theta_0)\right)^2 = o_p(1).\clubsuit
\end{aligned}$$

Proposition 4.7.6 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.0)$, on a :*

$$T_{12}^{**} = o_p(1). \tag{4.75}$$

Preuve Encore là, suite aux remarques (4.5) et (4.52),

$$\begin{aligned}
|T_{12}^{**}| &= |2n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x;\theta_0)\right]\left[F_0(y|x;\theta_0) - F_0(y|x;\hat{\theta})\right] \\
&\quad \times \left[w(y|x;\hat{\theta}) - w(y|x;\theta_0)\right]w(x)(\theta_0 - \hat{\theta})\dot{f}_0(y|x;\theta^*)dydx|, \\
&\leq Mn\sqrt{h}|\hat{\theta} - \theta_0|^3\int_0^1\int_{\mathbb{R}}|\hat{F}_n(y|x) - F_0(y|x;\theta_0)| \times h(x,y)dydx, \\
&\leq Mn\sqrt{h}|\hat{\theta} - \theta_0|^3 = o_p\left(\sqrt{h}\right) = o_p(1).\clubsuit
\end{aligned}$$

Ces dernières propositions permettent de remarquer que seul le comportement asymptotique de T_1^{**} intervient dans le comportement asymptotique de la statistique de test. Celui-ci est donné dans le théorème suivant :

Théorème 4.7.7 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$\frac{T - h^{-\frac{1}{2}}a_{\theta_0}}{\sigma_{\theta_0}} \xrightarrow{\mathcal{L}} N(0,1), \tag{4.76}$$

où

$$a_{\theta_0} = K^{*(2)}(0)\int_0^1\int_{\mathbb{R}}\frac{F_0(y|x;\theta_0)(1 - F_0(y|x;\theta_0))}{f(x)}w(y|x;\theta_0)f_0(y|x;\theta_0)dy\ w(x)dx, \tag{4.77}$$

et

$$\begin{aligned}
\sigma_{\theta_0}^2 &= 2K^{*(4)}(0)\int_0^1\int_{\mathbb{R}^2}\left[\frac{F_0(y\wedge y'|x';\theta_0) - F_0(y|x';\theta_0)F_0(y'|x';\theta_0)}{f(x')}\right]^2 \\
&\quad \times w^2(x')w(y|x';\theta_0)w(y'|x';\theta_0)f_0(y|x';\theta_0)f_0(y'|x';\theta_0)dydy'dx'.
\end{aligned} \tag{4.78}$$

Preuve On avait obtenu en (4.55) la décomposition suivante :

$$T = T_1^{**} + \sum_{i=2}^{12} T_i^{**},$$

de sorte que

$$\frac{T - h^{-\frac{1}{2}} a_{\theta_0}}{\sigma_{\theta_0}} = \frac{T_1^{**} - h^{-\frac{1}{2}} a_{\theta_0} + \sum_{i=2}^{12} T_i^{**}}{\sigma_{\theta_0}},$$

$$= \frac{T_1^{**} - h^{-\frac{1}{2}} a_{\theta_0}}{\sigma_{\theta_0}} + o_p(1). \quad (4.79)$$

L'adaptation de la Proposition 4.5.1 permet directement de déduire la convergence asymptotique du test. ♣

Comme dans les contextes précédents, ce résultat est inutilisable en pratique puisque les quantités a_{θ_0} et $\sigma^2_{\theta_0}$ dépendent de θ_0 qui est inconnu. Par le même raisonnement que pour le Théorème 4.6.2, on en déduit le théorème suivant, dont la preuve est escamotée :

Théorème 4.7.8 *Supposons que $F_0(y|x;\theta)$ vérifie les hypothèses $(F_0^*.y.1)$, $(F_0^{*(p+1)}.x.2)$, $(F_0.\theta.2)$ et $(f_0.\theta.2)$. Alors, sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.2)$, $(E.1)$, $(E.2)$ et $(H.3)$, on a :*

$$\frac{T - h^{-\frac{1}{2}} a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} \xrightarrow{\mathcal{L}} N(0,1), \quad (4.80)$$

où $a_{\hat{\theta}}$ et $\sigma^2_{\hat{\theta}}$ sont donnés par (4.77) et (4.78) respectivement avec $\hat{\theta}$ remplaçant θ_0.

Ainsi, un test de niveau asymptotique α de l'hypothèse nulle donnée en (4.3), nous amène à rejeter H_0 si

$$\frac{T - h^{-\frac{1}{2}} a_{\hat{\theta}}}{\sigma_{\hat{\theta}}} > z_{1-\alpha}, \quad (4.81)$$

où $z_{1-\alpha}$ est le fractile d'ordre $1-\alpha$ de la loi normale centrée réduite.

Le Théorème 4.7.8 permet donc de constater que le fait de faire dépendre la fonction de poids $w_x(y)$ de paramètres inconnus ne modifie en rien le comportement asymptotique de la statistique de test.

Il semble à présent intéressant d'étudier la puissance asymptotique de nos procédures de test. Plus particulièrement, on va s'intéresser à la puissance asymptotique locale du test donné en (1.42). Cette étude fait l'objet du prochain chapitre. On propose d'étudier cette puissance asymptotique locale dans le contexte des Chapitres 2 et 3, c'est à dire lorsque la fonction à tester sous H_0 est entièrement connue.

Chapitre 5

Puissance asymptotique locale du test

5.1 Introduction

L'existence de plusieurs tests d'adéquation pour un même problème pose la question parallèle de savoir si l'un d'eux se comporte mieux que les autres suivant un critère donné. Concernant un test statistique, un des critères les plus répandus est la puissance. Il existe cependant différentes manières de calculer la puissance d'un test d'adéquation. On peut en théorie la calculer directement selon sa définition mais ce calcul est rarement faisable en pratique. Pour des tailles d'échantillons finies, il est possible de la simuler par la méthode de Monte-Carlo, mais en pratique, seul un petit nombre d'alternatives peuvent être étudiées donnant ainsi une image partielle du comportement de ces tests. Plusieurs autres techniques ont donc été développées. Une revue de ces différentes approches, avec leurs points forts et leurs inconvénients, est donnée par Inglot et Ledwina (2001). L'une des plus utilisées propose d'affiner l'analyse de puissance en étudiant son comportement asymptotique local. Cette fonction de puissance asymptotique locale a souvent une forme simple qui permet les comparaisons. Son calcul passe par la détermination de la distribution asymptotique des statistiques de test sous une suite d'hypothèses alternatives, appelées "hypothèses contigües" et dénotées dans la suite par H_{1n}.

Dans le présent contexte, il nous semble intéressant de déterminer la puissance asymptotique locale du test basé sur la statistique introduite en (1.42). Afin de simplifier les calculs, on revient au contexte où la fonction à tester est entièrement spécifiée sous l'hypothèse nulle. Le comportement asymptotique de notre statistique de test vers la loi normale a déjà été obtenu sous l'hypothèse nulle ; ce comportement est donné dans le Théorème 3.5.3. On veut à présent déterminer ce comportement sous des suites d'hypothèses contigües, ce qui nous amène d'abord à définir plus exactement ces alternatives dans le paragraphe suivant. On rappelle en outre que dans tout ce qui suit, M est une constante majorante pouvant varier d'une ligne à l'autre.

5.2 Hypothèses contigües

On va supposer que la condition $(F_0.y.1)$ s'étend maintenant à $F(y|x)$, la "vraie" fonction de répartition conditionnelle et à $F_0(y|x)$, la fonction de répartition conditionnelle supposée sous H_0, de sorte que leur densité conditionnelle respective, soit $f(y|x)$ et $f_0(y|x)$, existe avec

$$F(y|x) = \int_{-\infty}^{y} f(t|x)dt \text{ et } F_0(y|x) = \int_{-\infty}^{y} f_0(t|x)dt. \tag{5.1}$$

Etant donné ce lien entre les densités et les fonctions de répartition conditionnelles, l'hypothèse (3.1) équivaut à

$$H_0 : f(y|x) = f_0(y|x), \ \forall y \in \mathbb{R}, \forall x \in [0,1]. \tag{5.2}$$

On considère à présent des suites d'alternatives contigües dont les densités conditionnelles ont la forme :

$$f_{1n}(y|x) = f_0(y|x) + c_n g_x(y), \tag{5.3}$$

où $g_x(y)$ est une fonction définie sur $y \in \mathbb{R}$ et $x \in [-1,2]$. Plus précisément, on considère une série de problèmes de tests ayant la structure triangulaire suivante.

Pour chaque $n \geq 1$, considérons $(X_{1n}, Y_{1n}), \ldots, (X_{nn}, Y_{nn})$, *iid* de fonction de répartition conjointe $F(x,y)$ associée à la densité conjointe $f(x,y)$. Si H_0 est vraie, $F(x,y) = F_0(x,y)$ de densité conjointe $f_0(x,y)$. Si maintenant H_{1n} est vraie, $F(x,y) = F_{1n}(x,y)$ de densité conjointe $f_{1n}(x,y)$. On suppose pour simplifier que la densité marginale de X, soit $f(.)$, est la même sous H_0 comme sous H_{1n}, de sorte que :

$$f_0(x,y) = f_0(y|x)f(x), \tag{5.4}$$

et

$$f_{1n}(x,y) = f_{1n}(y|x)f(x), \tag{5.5}$$

où $f_{1n}(y|x)$ est donnée en (5.3). Si on écrit $f(x,y) = f(y|x)f(x)$, alors

$$\text{si } H_0 \text{ est vraie, } f(y|x) = f_0(y|x), \tag{5.6}$$

$$\text{si } H_{1n} \text{ est vraie, } f(y|x) = f_{1n}(y|x). \tag{5.7}$$

Pour les raisons exposées au Chapitre 1, on doit cependant se limiter à un sous domaine $[a, b]$ de $]-1, 2[$. Sans perte de généralité, on prend ce domaine comme étant $[0, 1]$. Ceci nous amène donc à considérer la série de problèmes de test des hypothèses :

$$H_0 : f(y|x) = f_0(y|x), \ \forall y \in \mathbb{R}, \forall x \in [0, 1], \tag{5.8}$$

contre

$$H_{1n} : f(y|x) = f_{1,n}(y|x) = f_0(y|x) + c_n g_x(y). \tag{5.9}$$

Dans l'expression (5.3), la densité $f_{1n}(y|x)$ de l'alternative locale H_{1n} doit converger vers $f_0(y|x)$; c'est pourquoi on choisit c_n comme une suite de constantes positives tendant vers 0. On précisera le choix de c_n plus loin.

Pour l'instant, par le fait que $f_{1n}(y|x)$ est une densité, on doit imposer les conditions suivantes sur la fonction $g_x(y)$:

- $(g.0)$ $f_0(y|x) + c_n g_x(y) \geq 0, \forall x \in [-1, 2]$ et $\forall y \in \mathbb{R}$.
- $(g.1)$ $\int_{\mathbb{R}} g_x(y) dy = 0, \forall x \in [-1, 2]$.

Pour des raisons techniques, on suppose aussi que :

- $(g.2)$ $\int_{\mathbb{R}} |g_x(y)| dy \leq M, \forall x \in [-1, 2]$.

De plus, si on considère les relations données en (5.1), l'hypothèse (5.9) équivaut à :

$$H_{1n} : F(y|x) = F_{1n}(y|x) = F_0(y|x) + c_n G(y|x), \tag{5.10}$$

où $G(y|x)$ est définie telle que :

$$G(y|x) = \int_{-\infty}^{y} g_x(t) dt. \tag{5.11}$$

En raison des hypothèses sur la fonction $g_x(t)$, $G(y|x)$ est continûment différentiable en y pour tout $x \in [-1, 2]$. De plus, $|G(y|x)|$ est majorée pour tout $x \in [-1, 2]$ et $y \in \mathbb{R}$. Le fait que $F_{1n}(y|x)$ soit une fonction de répartition entraîne aussi que $G(-\infty|x) = G(+\infty|x) = 0$ (on l'obtient directement par la définition (5.11)). On suppose aussi que :

– $(G^{(k)}.x.1)$ La fonction qui à x associe $G(y|x)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage de tout $x \in]-1,2[$ pour tout $y \in \mathbb{R}$.

– $(G^{(k)}.x.2)$ On suppose $(G^{(k)}.x.1)$ et on suppose de plus que pour tout u dans un voisinage $V(x)$ de $x \in]-1,2[$ et pour tout $y \in \mathbb{R}$,

$$\sup_{u \in V(x)} \sup_{y \in \mathbb{R}} \left| G^{(k)}(y|u) \right| \leq M,$$

où $G^{(k)}(y|x)$ est la $k^{\text{ième}}$ dérivée de la fonction $G(y|x)$ par rapport à x.

Afin de simplifier la procédure mise en place dans ce chapitre, on revient pour le moment au contexte du Chapitre 2 où $F_0(y|x)$ est un polynôme de degré $\leq p$ en x. La fonction de répartition $F_{1n}(y|x)$ doit donc être un polynôme de degré $\leq p$ en x et c'est pourquoi on suppose, en plus des hypothèses émises au préalable, que la fonction $G(y|x)$ est un polynôme de degré $\leq p$ en x.

Ces suites d'hypothèses contigües étant établies, on peut maintenant s'intéresser au comportement asymptotique de la statistique de test sous H_{1n}. Une première étape de cette étude passe par une nouvelle décomposition de cette statistique, ce qui fait l'objet du prochain paragraphe.

5.3 Décomposition de la statistique de test

On considère, à l'instar de la statistique de test T donnée en (1.42), la statistique de test suivante :

$$T = n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \left(\hat{F}_n(y|x) - F_0(y|x)\right)^2 w(x) w_x(y) f_0(y|x) dy dx, \tag{5.12}$$

où $w(x)$ et $w_x(y)$ sont les fonctions de poids dont les propriétés sont données en $(P.0)$ et $(P.1)$ respectivement. De (5.10), on peut écrire :

$$\hat{F}_n(y|x) - F_0(y|x) = \hat{F}_n(y|x) - F_{1n}(y|x) + c_n G(y|x),$$

de sorte que, si on élève ce terme au carré, on obtient :

$$\begin{aligned} \left[\hat{F}_n(y|x) - F_0(y|x)\right]^2 &= \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right]^2 + c_n^2 G^2(y|x) \\ &\quad + 2c_n \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x). \end{aligned} \tag{5.13}$$

De (5.3), on a aussi :

$$f_0(y|x) = f_{1n}(y|x) - c_n g_x(y). \tag{5.14}$$

Si on substitue les résultats (5.13) et (5.14) dans la statistique de test (5.12), on obtient la décomposition suivante :

$$T = T_1 + T_2 + T_3 + T_4 + T_5 + T_6, \tag{5.15}$$

où

$$T_1 = n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_{1n}(y|x)\right]^2 w(x)w_x(y)f_{1n}(y|x)dydx, \tag{5.16}$$

$$T_2 = -n\sqrt{h}c_n\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_{1n}(y|x)\right]^2 w(x)w_x(y)g_x(y)dydx, \tag{5.17}$$

$$T_3 = n\sqrt{h}c_n^2\int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)f_{1n}(y|x)dydx, \tag{5.18}$$

$$T_4 = -n\sqrt{h}c_n^3\int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)g_x(y)dydx, \tag{5.19}$$

$$\begin{aligned} T_5 = {} & 2n\sqrt{h}c_n\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x) \\ & \times w_x(y)f_{1n}(y|x)dydx, \end{aligned} \tag{5.20}$$

$$T_6 = -2n\sqrt{h}c_n^2\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)g_x(y)dydx. \tag{5.21}$$

La décomposition (5.15) montre que l'étude du comportement asymptotique de la statistique de test T sous H_{1n} passe par l'étude du comportement asymptotique de la somme des 6 statistiques T_1 à T_6.

5.4 Comportement asymptotique de la statistique de test

Suite à la décomposition (5.15) de la statistique de test, on étudie dans cette partie le comportement asymptotique de chacune des six statistiques qui la composent. Il se pose alors la question du choix de la suite c_n.

La plupart des études de puissances asymptotiques classiques se basent sur des alternatives locale H_{1n} qui convergent vers l'hypothèse nulle à la vitesse $\frac{1}{\sqrt{n}}$ (Pitman, 1949). Ceci suggère de travailler pour notre test avec $c_n = (n\sqrt{h})^{-\frac{1}{2}}$, suite de constantes positives qui tend bien vers 0 quand $n \longrightarrow +\infty$. Ce choix permet d'assurer la convergence de l'hypothèse alternative H_{1n} vers l'hypothèse nulle H_0. D'autre part, on montrera ultérieurement que ce choix de c_n permet d'obtenir une puissance locale non triviale c'est à dire supérieure au niveau α du test.

A l'instar de la remarque faite à la Section 1.3, on utilise les notations $E_{F_{1n}}(.)$ et $E_{F_{1n}}(.\,|X)$ pour désigner l'espérance associée à la loi conjointe $F_{1n}(.\,,\,.)$ et à la loi conditionnelle $F_{1n}(.\,|\,.)$ respectivement. Comme dans les précédents chapitres, $E_F(X)$ désignera l'espérance de X associée à la loi marginale $F(.)$ de X, de loi $f(.)$. Notons aussi que ces mêmes notations seront transposées à celles concernant la variance.

On rappelle que dans ce qui suit, $(X_{1n}, Y_{1n}), \ldots, (X_{nn}, Y_{nn})$ sont $iid \sim F(x,y)$.

Etude de T_1

Proposition 5.4.1 *Sous les hypothèses du Corollaire* 2.5.1, *sous les hypothèses supplémentaires* $(g.0)$, $(g.1)$, $(g.2)$ *et si l'hypothèse* (5.9) *est vraie, on a :*

$$\frac{T_1 - h^{-\frac{1}{2}} a_{1n}}{\sigma_{1n}} \xrightarrow{\mathcal{L}} N(0,1), \tag{5.22}$$

où

$$a_{1n} = K^{*^{(2)}}(0) \int_0^1 \int_{\mathbb{R}} \frac{F_{1n}(y|x)(1 - F_{1n}(y|x))}{f(x)} w_x(y) f_{1n}(y|x) dy w(x) dx, \tag{5.23}$$

et

$$\begin{aligned} \sigma_{1n}^2 &= 2K^{*^{(4)}}(0) \int_0^1 \int_{\mathbb{R}^2} \left[\frac{F_{1n}(y \wedge y'|x') - F_{1n}(y|x') F_{1n}(y'|x')}{f(x')} \right]^2 \\ &\quad \times w^2(x') w_{x'}(y) w_{x'}(y') f_{1n}(y|x') f_{1n}(y'|x') dy dy' dx'. \end{aligned} \tag{5.24}$$

Preuve Comme $F_{1n}(y|x)$ est un polynôme de degré $\leq p$ en x, on a, à l'instar de (2.3) :

$$\begin{aligned} T_1 &= n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \left(\sum_{i=1}^n W^n \left(\frac{X_{in} - x}{h} \right) \left(I_{\{Y_{in} \leq y\}} - F_{1n}(y|X_{in}) \right) \right)^2 \\ &\quad \times w(x) w_x(y) f_{1n}(y|x) dy dx, \end{aligned} \tag{5.25}$$

où le noyau $W^n(.)$ est défini en (1.16). Même en travaillant avec un tableau triangulaire de variables aléatoires, on peut reprendre la démarche de la Section 2.5 pour obtenir, à l'instar de (2.114),

$$T_1 = \tilde{T}_1 + o_p(1), \tag{5.26}$$

avec

$$\tilde{T}_1 = n\sqrt{h}\int_0^1 \int_{\mathbb{R}} \left(\sum_{i=1}^n \tilde{W}^n(X_{in}, x, h)\left(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})\right)\right)^2 w(x)w_x(y)f_{1n}(y|x)dydx, \tag{5.27}$$

où le noyau $\tilde{W}^n(.,.,.)$ a été défini en (2.6). Cette dernière statistique se décompose, à l'instar de (2.7), en deux statistiques :

$$\tilde{T}_1 = \tilde{T}_1^* + \tilde{T}_1^{**}, \tag{5.28}$$

où

$$\begin{aligned} \tilde{T}_1^* &= n\sqrt{h}\sum_{i=1}^n \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_{in}, x, h)^2 \left(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})\right)^2 \\ &\quad \times w(x)w_x(y)f_{1n}(y|x)dydx, \end{aligned} \tag{5.29}$$

et

$$\begin{aligned} \tilde{T}_1^{**} &= n\sqrt{h}\sum_{1\leq i\neq j\leq n} \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_{in}, x, h)\tilde{W}^n(X_{jn}, x, h)\left(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})\right) \\ &\quad \times \left(I_{\{Y_{jn}\leq y\}} - F_{1n}(y|X_{jn})\right) w(x)w_x(y)f_{1n}(y|x)dydx. \end{aligned} \tag{5.30}$$

Le comportement asymptotique de la statistique $\tilde{T}_1^*$ s'obtient de la même manière que celui de T_1 définie en (2.13) même si on travaille cette fois avec un tableau triangulaire $(X_{i,n}, Y_{i,n}), i = 1, \ldots, n, n \geq 1$ de variables aléatoires. En effet, en recyclant les preuves des Propositions 2.3.1 et 2.3.2, on a :

$$E_{F_{1n}}\left(\tilde{T}_1^*\right) = h^{-\frac{1}{2}}a_{1n} + O\left(h^{\frac{1}{2}}\right),$$

et

$$Var_{F_{1n}}\left(\tilde{T}_1^*\right) \leq O\left(\frac{1}{nh}\right) + o\left(\frac{1}{nh}\right),$$

où a_{1n} est donnée en (5.23) pour obtenir, par l'inégalité de Bienaymé-Tchebytchev,

$$\tilde{T}_1^* - h^{-\frac{1}{2}} a_{1n} \xrightarrow{P} 0. \tag{5.31}$$

La convergence en loi de $\tilde{T}_1^{**}$ demande cependant un peu plus d'attention. En effet, cette statistique possède la structure :

$$\tilde{T}_1^{**} = \sum_{1 \leq i,j \leq n} Q_{ijn}\left((X_{in}, Y_{in}), (X_{jn}, Y_{jn})\right), \tag{5.32}$$

où les noyaux $Q_{ijn}\left((X_{in}, Y_{in}), (X_{jn}, Y_{jn})\right)$ sont définis par :

$$Q_{ijn}\left((X_{in}, Y_{in}), (X_{jn}, Y_{jn})\right)$$

$$= \begin{cases} 0 & \text{si } i = j, \\ n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_{in}, x, h) \tilde{W}^n(X_{jn}, x, h) \left(I_{\{Y_{in} \leq y\}} - F_{1n}(y|X_{in})\right) & \\ \times \left(I_{\{Y_{jn} \leq y\}} - F_{1n}(y|X_{jn})\right) w(x) w_x(y) f_{1n}(y|x) dy dx & \text{si } i \neq j. \end{cases} \tag{5.33}$$

Cette structure ne permet pas d'utiliser directement le théorème de de Jong (1987) qui s'applique à une suite infinie de variables aléatoires *iid* et pas à un tableau triangulaire. On peut cependant se ramener au cas de ce théorème par le raisonnement suivant. Considérons $Z_i = (X_i, Y_i), i = 1, \ldots, n$ des couples de variables aléatoires *iid* de fonction de répartition conjointe $F_0(x, y)$. Pour chaque $n \geq 1$ et pour $i = 1, \ldots, n$, soit

$$\begin{aligned} Z_{in}^* &= \left(X_i, F_{1n}^{-1}\left(F_0(Y_i|X_i)|X_i\right)\right), \\ &= \left(X_i, G_n(Z_i)\right). \end{aligned} \tag{5.34}$$

Pour chaque $n \geq 1$, les $Z_{in}^*, i = 1, \ldots, n$, sont *iid* de fonction de répartition conjointe $F_{1n}(x, y)$. On peut donc écrire que :

$$\mathcal{L}(\tilde{T}_1^{**}) = \mathcal{L}\left(\sum_{1 \leq i,j \leq n} Q_{ijn}\left(Z_{in}^*, Z_{jn}^*\right)\right), \tag{5.35}$$

$$
\begin{aligned}
&= \mathcal{L}\left(\sum_{1\leq i,j\leq n} Q_{ijn}\left((X_i, G_n(Z_i)), (X_j, G_n(Z_j))\right)\right), \\
&= \mathcal{L}\left(\sum_{1\leq i,j\leq n} Q^*_{ijn}\left(Z_i, Z_{j,}\right)\right), \qquad (5.36)
\end{aligned}
$$

où d'une manière générale, $\mathcal{L}(T)$ désigne la loi de la statistique T et où maintenant, les noyaux $Q^*_{ijn}(Z_i, Z_j)$ sont définis par :

$$
Q^*_{ijn}\left(Z_i, Z_j\right) = \begin{cases} 0 & \text{si } i = j, \\ n\sqrt{h}\int_0^1\int_{\mathrm{IR}} \tilde{W}^n(X_i, x, h)\tilde{W}^n(X_j, x, h)\left(I_{\{G_n(Z_i)\leq y\}} - F_{1n}(y|X_i)\right) & \\ \times\left(I_{\{G_n(Z_j)\leq y\}} - F_{1n}(y|X_j)\right) w(x)w_x(y)f_{1n}(y|x)dydx & \text{si } i \neq j. \end{cases} \qquad (5.37)
$$

La représentation (5.36) correspond au contexte de variables aléatoires *iid* de de Jong (1987) car celui-ci autorise les fonctions noyaux Q^*_{ijn} à varier avec n. Il suffit alors de vérifier que les conditions de de Jong (1987) s'appliquent à cette représentation pour en déduire le comportement asymptotique de notre statistique (5.35). En effet, ces conditions s'appliquent aux moments de $\tilde{T}_1^{**}$ de (5.36) et donc de Q^*_{ijn}. Mais ceux-ci sont les mêmes que ceux de Q_{ijn}. Une adaptation des preuves de ces conditions données à la Section 2.4 permet cette vérification. On en conclut que :

$$
\frac{\tilde{T}_1^{**}}{\sigma_{1n}} \xrightarrow{\mathcal{L}} N(0,1) \qquad (5.38)
$$

où σ^2_{1n} est défini en (5.24). Si maintenant, on substitue l'ensemble de ces résultats dans la décomposition de T_1 donnée en (5.31), on obtient directement le résultat cherché.♣

Note Les quantités a_{1n} et σ^2_{1n} convergent respectivement vers les quantités a_0 et σ^2_0, données en (2.15) et (2.69). De plus,

$$
\begin{aligned}
\frac{T_1 - h^{-\frac{1}{2}}a_0}{\sigma_0} &= \left[\frac{T_1 - h^{-\frac{1}{2}}a_{1n} + h^{-\frac{1}{2}}\left(a_{1n} - a_0\right)}{\sigma_{1n}}\right] \times \frac{\sigma_{1n}}{\sigma_0}, \\
&= \left[\frac{T_1 - h^{-\frac{1}{2}}a_{1n}}{\sigma_{1n}}\right] \times (1 + o(1)) + O\left(h^{-\frac{1}{2}}(a_{1n} - a_0)\right). \qquad (5.39)
\end{aligned}
$$

Regardons à présent le comportement asymptotique de $h^{-\frac{1}{2}}(a_{1n}-a_0)$. De (5.3) et (5.10), on a :

$$\begin{aligned}
a_{1n} &= K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}} F_{1n}(y|x)(1-F_{1n}(y|x))f_{1n}(y|x)f^{-1}(x)w_x(y)dyw(x)dx,\\
&= K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}} [F_0(y|x)+c_nG(y|x)]\,[1-F_0(y|x)-c_nG(y|x)]\\
&\quad\times[f_0(y|x)+c_ng_x(y)]\times f^{-1}(x)w_x(y)dyw(x)dx,\\
&= a_0+c_nK^{*^{(2)}}(0)\Big[-2\int_0^1\int_{\mathbb{R}} F_0(y|x)G(y|x)f_0(y|x)f^{-1}(x)w_x(y)dyw(x)dx\\
&\quad+\int_0^1\int_{\mathbb{R}} G(y|x)f_0(y|x)f^{-1}(x)w_x(y)dyw(x)dx\\
&\quad+\int_0^1\int_{\mathbb{R}} g_x(y)F_0(y|x)(1-F_0(y|x))f^{-1}(x)w_x(y)dyw(x)dx\Big]\\
&\quad+c_n^2K^{*^{(2)}}(0)\Big[-\int_0^1\int_{\mathbb{R}} G^2(y|x)f_0(y|x)f^{-1}(x)w_x(y)dyw(x)dx\\
&\quad-2\int_0^1\int_{\mathbb{R}} g_x(y)F_0(y|x)G(y|x)f^{-1}(x)w_x(y)dyw(x)dx\\
&\quad+\int_0^1\int_{\mathbb{R}} g_x(y)G(y|x)f^{-1}(x)w_x(y)dyw(x)dx\Big]\\
&\quad-c_n^3K^{*^{(2)}}(0)\int_0^1\int_{\mathbb{R}} g_x(y)G^2(y|x)f^{-1}(x)w_x(y)dyw(x)dx.
\end{aligned}$$

Puisque sous les hypothèses $(K.0)$, $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$ et $(g.2)$ ces intégrales sont majorées et comme sous $(H.1)$, $c_n^2h^{-\frac{1}{2}}$ et $c_n^3h^{-\frac{1}{2}}$ sont $o(1)$, on obtient directement que :

$$\begin{aligned}
h^{-\frac{1}{2}}|a_{1n}-a_0| &\leq Mh^{-\frac{1}{2}}c_n+o(1),\\
&= O\left(\frac{1}{\sqrt{nh^{\frac{3}{2}}}}\right)+o(1),
\end{aligned}$$

$$= \quad o(1), \tag{5.40}$$

sous l'hypothèse $(H.2)$. Si maintenant on substitue le résultat (5.40) dans l'expression (5.39), on obtient :

$$\frac{T_1 - h^{-\frac{1}{2}} a_0}{\sigma_0} = \frac{T_1 - h^{-\frac{1}{2}} a_{1n}}{\sigma_{1n}} + o_p(1), \tag{5.41}$$

et donc $\dfrac{T_1 - h^{-\frac{1}{2}} a_{1n}}{\sigma_{1n}}$ et $\dfrac{T_1 - h^{-\frac{1}{2}} a_0}{\sigma_0}$ ont le même comportement asymptotique sous H_{1n}.

Etude de T_2

Proposition 5.4.2 *Sous les hypothèses de la Proposition* 5.4.1, *on a :*

$$T_2 = o_p(1). \tag{5.42}$$

Preuve On a,

$$\begin{aligned} |T_2| &= \left| -n\sqrt{h} c_n \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x) \right]^2 w(x) w_x(y) g_x(y) dy dx \right|, \\ &\leq c_n n \sqrt{h} \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x) \right]^2 w(x) w_x(y) |g_x(y)| dy dx. \end{aligned} \tag{5.43}$$

De plus,

$$n\sqrt{h} \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x) \right]^2 w(x) w_x(y) |g_x(y)| dy dx = O\left(h^{-\frac{1}{2}}\right) + O_p(1). \tag{5.44}$$

En effet, cette statistique est semblable à la statistique T_1 définie en (5.16) avec pour seule différence l'utilisation de la fonction $|g_x(y)|$ à la place de la densité $f_{1n}(y|x)$. Cette différence ne nous empêche pas de recycler la preuve de la Proposition 5.4.1 puisque la condition $(g.2)$ assure que $\int_{\mathbb{R}} |g_x(y)| dy < +\infty$ pour tout $x \in [-1, 2]$ qui est nécessaire pour l'obtention du résultat. On obtient alors :

$$\begin{aligned} |T_2| &\leq c_n \left[O\left(h^{-\frac{1}{2}}\right) + O_p(1) \right], \\ &= o_p(1), \end{aligned} \tag{5.45}$$

sous l'hypothèse $(H.2)$.♣

Etude de T_3

$$\begin{aligned}
T_3 &= n\sqrt{h}c_n^2\int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)f_{1n}(y|x)dydx,\\
&= \int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)f_{1n}(y|x)dydx = b_{1n},\\
&\longrightarrow \int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)f_0(y|x)dydx = b_0 \geq 0,\ \text{quand } n \longrightarrow +\infty.
\end{aligned}$$

Etude de T_4

Proposition 5.4.3 *Sous les hypothèses* $(P.0)$, $(P.1)$, $(g.0)$, $(g.1)$, $(g.2)$ *et* $(H.1)$,

$$T_4 = o(1). \tag{5.46}$$

Preuve On a,

$$\begin{aligned}
|T_4| &= \left|-n\sqrt{h}c_n^3\int_0^1\int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)g_x(y)dydx\right|,\\
&\leq Mc_n\int_0^1\int_{\mathbb{R}} G^2(y|x)|g_x(y)|dydx,\\
&\leq M\left((n\sqrt{h})^{-\frac{1}{2}}\right) = o(1),
\end{aligned}$$

sous l'hypothèse $(H.1)$.♣

Etude de T_5

Proposition 5.4.4 *Sous les hypothèses de la Proposition* 5.4.1, *on a :*

$$T_5 = o_p(1). \tag{5.47}$$

Preuve On a,

$$\begin{aligned}
T_5 &= 2n\sqrt{h}c_n\int_0^1\int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)f_{1n}(y|x)dydx,\\
&= 2(n\sqrt{h})^{\frac{1}{2}}\int_0^1\int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)f_{1n}(y|x)dydx.
\end{aligned}$$

Maintenant, la décomposition du noyau $W^n(.)$ donnée en (1.34) permet d'obtenir, sachant que les X_{in} sont *iid* de fonction de répartition $F(.)$,

$$\begin{aligned}\hat{F}_n(y|x) - F_{1n}(y|x) &= \sum_{i=1}^{n} \tilde{W}^n(X_{in}, x, h)(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})) \\ &+ h\sum_{i=1}^{n} Z_n\left(\frac{X_{in}-x}{h}, x\right)\bar{W}^n(X_{in}, x, h) \\ &\times (I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})), \qquad (5.48)\end{aligned}$$

où les noyau $\tilde{W}^n(.,.,.)$ et $\bar{W}^n(.,.,.)$ sont définis respectivement en (2.6) et (2.71). Dans cette expression, on rappelle que $Z_n\left(\frac{X_{in}-x}{h}, x\right)$, défini en (1.33), est presque sûrement borné en $x \in [0,1]$. On en déduit que :

$$\begin{aligned}&\left(n\sqrt{h}\right)^{\frac{1}{2}} \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)f_{1n}(y|x)dydx \\ &= \left(n\sqrt{h}\right)^{\frac{1}{2}} \sum_{i=1}^{n} r_i + h\left(n\sqrt{h}\right)^{\frac{1}{2}} \sum_{i=1}^{n} s_i, \qquad (5.49)\end{aligned}$$

où on a posé

$$r_i = \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_{in}, x, h)\left(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})\right) G(y|x)w(x)w_x(y)f_{1n}(y|x)dydx, \qquad (5.50)$$

et

$$\begin{aligned}s_i &= \int_0^1 \int_{\mathbb{R}} Z_n\left(\frac{X_{in}-x}{h}, x\right)\bar{W}^n(X_{in}, x, h)\left(I_{\{Y_{in}\leq y\}} - F_{1n}(y|X_{in})\right) \\ &\times G(y|x)w(x)w_x(y)f_{1n}(y|x)dydx. \qquad (5.51)\end{aligned}$$

Etudions à présent le comportement asymptotique de la statistique

$$\left(n\sqrt{h}\right)^{\frac{1}{2}} \sum_{i=1}^{n} r_i.$$

Par le théorème de Fubini,

$$\begin{aligned}E_{F_{1n}}(r_1) &= \int_{-1}^{2} \int_{\mathbb{R}} \int_0^1 \tilde{W}^n(x_{1n}, x, h)\left(\int_{\mathbb{R}} (I_{\{y_{1n}\leq y\}} - F_{1n}(y|x_{1n}))f_{1n}(y_{1n}|x_{1n})dy_{1n}\right) \\ &\times f_{1n}(y|x)G(y|x)w(x)w_x(y)dydxf(x_{1n})dx_{1n}, \\ &= 0,\end{aligned}$$

car $\int_{\mathbb{R}} \left(I_{\{y_{1n} \leq y\}} - F_{1n}(y|x_{1n}) \right) f_{1n}(y_{1n}|x_{1n}) dy_{1n} = 0$ $\forall y \in \mathbb{R}$ et $\forall x_{1n} \in [-1, 2]$. On en déduit que $E_{F_{1n}} \left(\left(n\sqrt{h} \right)^{\frac{1}{2}} \sum_{i=1}^{n} r_i \right) = 0$. Comme les r_i sont *iid* on a aussi

$$Var_{F_{1n}} \left(\left(n\sqrt{h} \right)^{\frac{1}{2}} \sum_{i=1}^{n} r_i \right) = n^2 \sqrt{h} E_{F_{1n}}(r_1^2). \tag{5.52}$$

De plus, en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6),

$$\begin{aligned} r_1^2 &= \frac{1}{n^2h^2} \int_0^1 \int_{\mathbb{R}} \int_0^1 \int_{\mathbb{R}} K^* \left(\frac{X_{1n} - x}{h} \right) K^* \left(\frac{X_{1n} - x'}{h} \right) f^{-1}(x) f^{-1}(x') \\ &\quad \times (I_{\{Y_{1n} \leq y\}} - F_{1n}(y|X_{1n}))(I_{\{Y_{1n} \leq y'\}} - F_{1n}(y'|X_{1n}))G(y|x)G(y'|x') \\ &\quad \times w(x)w(x')w_x(y)w_{x'}(y') f_{1n}(y|x) f_{1n}(y'|x') dy dx dy' dx', \end{aligned}$$

de sorte que, sous les hypothèses $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(g.0)$, $(g.1)$ et $(g.2)$,

$$\begin{aligned} 0 \leq E_{F_{1n}}(r_1^2) &\leq \frac{1}{n^2h^2} \int_{-1}^2 \int_{\mathbb{R}} \int_0^1 \int_{\mathbb{R}} \int_0^1 \left| K^* \left(\frac{x_{1n} - x}{h} \right) \right| \left| K^* \left(\frac{x_{1n} - x'}{h} \right) \right| f^{-1}(x) f^{-1}(x') \\ &\quad \times \left(\int_{\mathbb{R}} \left| I_{\{y_{1n} \leq y\}} - F_{1n}(y|x_{1n}) \right| \left| I_{\{y_{1n} \leq y'\}} - F_{1n}(y'|x_{1n}) \right| f_{1n}(y_{1n}|x_{1n}) dy_{1n} \right) \\ &\quad \times |G(y|x)| \times |G(y'|x')| \, w(x)w(x')w_x(y)w_{x'}(y') \\ &\quad \times f_{1n}(y|x) f_{1n}(y'|x') dy dx dy' dx' f(x_{1n}) dx_{1n}, \\ &\leq \frac{M}{n^2h^2} \int_{-1}^2 \int_0^1 \int_0^1 \left| K^* \left(\frac{x_{1n} - x}{h} \right) \right| \left| K^* \left(\frac{x_{1n} - x'}{h} \right) \right| \\ &\quad \times \left(\int_{\mathbb{R}} |G(y|x)| \, f_{1n}(y|x) dy \right) \left(\int_{\mathbb{R}} |G(y'|x')| \, f_{1n}(y'|x') dy' \right) dx dx' f(x_{1n}) dx_{1n}, \\ &\leq \frac{M}{n^2h^2} \int_{-1}^2 \int_0^1 \int_0^1 \left| K^* \left(\frac{x_{1n} - x}{h} \right) \right| \left| K^* \left(\frac{x_{1n} - x'}{h} \right) \right| dx dx' f(x_{1n}) dx_{1n}. \end{aligned}$$

On effectue à présent les changements de variables $x = -uh + x_{1n}$ et $x' = -vh + x_{1n}$ de Jacobien $\dfrac{1}{h^2}$ pour obtenir, sous l'hypothèse $(K.0)$,

$$E_{F_{1,n}}(r_1^2) \leq \frac{M}{n^2} \int_{-1}^2 \int_{\mathbb{R}} \int_{\mathbb{R}} |K^*(u)| \, |K^*(v)| \, du dv f(x_{1,n}) dx_{1,n},$$

$$\leq \frac{M}{n^2}.$$

Si on substitue ce résultat dans (5.52), on a

$$Var_{F_{1n}}\left(\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} r_i\right) \leq \sqrt{h}M = o(1), \tag{5.53}$$

sous l'hypothèse $(H.0)$. Par l'inégalité de Bienaymé-Tchebychev,

$$\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} r_i = o_p(1). \tag{5.54}$$

On étudie pour finir le deuxième terme de la décomposition (5.49), à savoir

$$\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} s_i.$$

On a,

$$\begin{aligned} E_{F_{1n}}(s_1|X_{1n},\dots,X_{nn}) &= \int_{\mathbb{R}}\int_0^1 Z_n\left(\frac{X_{1n}-x}{h},x\right)\bar{W}^n(X_{1n},x,h) \\ &\quad \times \left(\int_{\mathbb{R}}(I_{\{y_{1n}\leq y\}} - F_{1n}(y|X_{1n}))f_{1n}(y_{1n}|X_{1n})dy_{1n}\right) \\ &\quad \times f_{1n}(y|x)G(y|x)w(x)w_x(y)dydx, \\ &= 0. \end{aligned}$$

On en déduit que $E_{F_{1n}}\left(\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} s_i\right) = 0$. Maintenant, comme $E_{F_{1n}}(s_1|X_{1n},\dots,X_{nn}) = 0$,

$$\begin{aligned} Var_{F_{1n}}\left(\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} s_i\right) &= n\sqrt{h}Var_{F_{1n}}\left(\sum_{i=1}^{n} s_i\right), \\ &= n\sqrt{h}E_F\left(Var_{F_{1n}}\left(\sum_{i=1}^{n} s_i|X_{1n},X_{2n},\dots,X_{nn}\right)\right), \\ &= n^2\sqrt{h}E_F\left(E_{F_{1n}}\left(s_1^2|X_{1n},\dots,X_{nn}\right)\right). \end{aligned} \tag{5.55}$$

De plus, en utilisant la définition de $\bar{W}^n(.,.,.)$ donnée en (2.71),

$$s_1^2 = \frac{1}{n^2h^2}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_{1n}-x}{h},x\right)Z_n\left(\frac{X_{1n}-x'}{h},x'\right)\bar{K}\left(\frac{X_{1n}-x}{h}\right)\bar{K}\left(\frac{X_{1n}-x'}{h}\right)$$

$$\times\, (I_{\{Y_{1n}\leq y\}} - F_{1n}(y|X_{1n}))(I_{\{Y_{1n}\leq y'\}} - F_{1n}(y'|X_{1n}))G(y|x)G(y'|x')$$

$$\times\, f^{-1}(x)f^{-1}(x')w(x)w(x')w_x(y)w_{x'}(y')f_{1n}(y|x)f_{1n}(y'|x')dydxdy'dx',$$

de sorte que, comme $|Z_n(.,.)| \leq M$ et sous les hypothèses $(X.0)$, $(f.0)$, $(P.0)$, $(P.1)$, $(g.0)$, $(g.1)$ et $(g.2)$,

$$\begin{aligned}
E_{F_{1n}}\left(s_1^2|X_{1n},\ldots,X_{nn}\right) \;\leq\; & \frac{1}{n^2h^2}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1 |Z_n\left(\frac{X_{1n}-x}{h},x\right)||Z_n\left(\frac{X_{1n}-x'}{h},x'\right)| \\
& \times |\bar{K}\left(\frac{X_{1n}-x}{h}\right)||\bar{K}\left(\frac{X_{1n}-x'}{h}\right)| \\
& \times \left(\int_{\mathbb{R}}\left|I_{\{y_{1n}\leq y\}} - F_{1n}(y|X_{1n})\right|\left|I_{\{y_{1n}\leq y'\}} - F_{1n}(y'|X_{1n})\right|\right. \\
& \left.\times\, f_{1n}(y_{1n}|X_{1n})dy_{1n}\right) \\
& \times\, f^{-1}(x)f^{-1}(x')|G(y|x)||G(y'|x')| \\
& \times\, w(x)w(x')w_x(y)w_{x'}(y')f_{1n}(y|x)f_{1n}(y'|x')dydxdy'dx', \\
\leq\; & \frac{M}{n^2h^2}\int_0^1\int_0^1\left|\bar{K}\left(\frac{X_{1n}-x}{h}\right)\right|\left|\bar{K}\left(\frac{X_{1n}-x'}{h}\right)\right|dxdx'.
\end{aligned}$$

On effectue à présent les changements de variables $x = -uh + X_{1n}$ et $x' = -vh + X_{1n}$ de Jacobien $\dfrac{1}{h^2}$ pour obtenir, sous l'hypothèse $(K.0)$,

$$\begin{aligned}
E_{F_{1n}}\left(s_1^2|X_{1n},\ldots,X_{nn}\right) \;&\leq\; \frac{M}{n^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}(u)\right|\left|\bar{K}(v)\right|dudv, \\
&\leq\; \frac{M}{n^2}.
\end{aligned}$$

Si on substitue ce résultat dans (5.55), on obtient :

$$Var_{F_{1n}}\left(\left(n\sqrt{h}\right)^{\frac{1}{2}}\sum_{i=1}^{n} s_i\right) \leq \sqrt{h}M = o(1), \tag{5.56}$$

sous l'hypothèse $(H.0)$. Par l'inégalité de Bienaymé-Tchebychev,

$$\left(n\sqrt{h}\right)^{\frac{1}{2}} \sum_{i=1}^{n} s_i = o_p(1). \tag{5.57}$$

Si on substitue maintenant les résultats (5.54) et (5.57) dans la décomposition (5.49), on obtient :

$$T_5 \quad = \quad o_p(1).\clubsuit \tag{5.58}$$

Etude de T_6

Proposition 5.4.5 *Sous les hypothèses de la Proposition* 5.4.1, *on a :*

$$T_6 = o_p(1). \tag{5.59}$$

Preuve On a,

$$\begin{aligned} T_6 &= -2n\sqrt{h}c_n^2 \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)g_x(y)dydx, \\ &= -2 \int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)g_x(y)dydx. \end{aligned}$$

De façon similaire à (5.49), on a :

$$\int_0^1 \int_{\mathbb{R}} \left[\hat{F}_n(y|x) - F_{1n}(y|x)\right] G(y|x)w(x)w_x(y)g_x(y)dydx \quad = \quad \sum_{i=1}^{n} w_i + h\sum_{i=1}^{n} z_i, \tag{5.60}$$

avec

$$w_i = \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_{in}, x, h)(I_{\{Y_{in} \leq y\}} - F_{1n}(y|X_{in}))G(y|x)w(x)w_x(y)g_x(y)dydx, \tag{5.61}$$

et

$$\begin{aligned} z_i \quad &= \quad \int_0^1 \int_{\mathbb{R}} Z_n\left(\frac{X_{in} - x}{h}, x\right) \bar{W}^n(X_{in}, x, h)(I_{\{Y_{in} \leq y\}} - F_{1n}(y|X_{in})) \\ &\quad \times G(y|x)w(x)w_x(y)g_x(y)dydx, \end{aligned} \tag{5.62}$$

où les noyaux $\tilde{W}^n(.,.,.)$ et $\bar{W}^n(.,.,.)$ sont définis respectivement en (2.6) et (2.71). On rappelle aussi que le terme $Z_n\left(\dfrac{X_{in} - x}{h}, x\right)$, défini en (1.33), est presque sûrement borné en $x \in [0,1]$. Passons maintenant à l'étude du comportement asymptotique du premier terme de (5.60). On a

$$\begin{aligned} E_{F_{1n}}(w_1) &= \int_{-1}^{2}\int_{\mathbb{R}}\int_0^1 \tilde{W}^n(x_{1n},x,h)\left(\int_{\mathbb{R}}(I_{\{y_{1n}\leq y\}}-F_{1n}(y|x_{1n}))f_{1n}(y_{1n}|x_{1n})dy_{1n}\right) \\ &\quad \times G(y|x)w(x)w_x(y)g_x(y)dydxf(x_{1n})dx_{1n}, \\ &= 0. \end{aligned}$$

On en déduit que $E_{F_{1n}}\left(\sum_{i=1}^n w_i\right)=0$. Comme les w_i sont *iid* d'espérance nulle, on a aussi

$$Var_{F_{1n}}\left(\sum_{i=1}^n w_i\right) = nE_{F_{1n}}\left(w_1^2\right), \tag{5.63}$$

de sorte que, en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6),

$$\begin{aligned} w_1^2 &= \frac{1}{n^2h^2}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} K^*\left(\frac{X_{1n}-x}{h}\right)K^*\left(\frac{X_{1n}-x'}{h}\right)f^{-1}(x)f^{-1}(x') \\ &\quad \times (I_{\{Y_{1n}\leq y\}}-F_{1n}(y|X_{1n}))(I_{\{Y_{1n}\leq y'\}}-F_{1n}(y'|X_{1n})) \\ &\quad \times G(y|x)G(y'|x')w(x)w(x')w_x(y)w_{x'}(y')g_x(y)g_{x'}(y')dydxdy'dx'. \end{aligned}$$

On en déduit que :

$$\begin{aligned} 0 \leq E_{F_{1n}}(w_1^2) &\leq \frac{M}{n^2h^2}\int_{-1}^{2}\int_0^1\int_0^1\left|K^*\left(\frac{x_{1n}-x}{h}\right)\right|\left|K^*\left(\frac{x_{1n}-x'}{h}\right)\right|dxdx'f(x_{1n})dx_{1n}, \\ &\leq \frac{M}{n^2}\int_{-1}^{2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|K^*(u)\right|\left|K^*(v)\right|dudvf(x_{1n})dx_{1n}, \\ &\leq \frac{M}{n^2}. \end{aligned}$$

Si on substitue ce résultat dans (5.63), on a

$$Var_{F_{1n}}\left(\sum_{i=1}^n w_i\right) \leq \frac{M}{n} = o(1), \tag{5.64}$$

et on en déduit que :

$$\sum_{i=1}^n w_i = o_p(1). \tag{5.65}$$

Etudions maintenant le dernier terme de (5.60). On a,

$$\begin{aligned} E_{F_{1n}}(z_1|X_{1n},\ldots,X_{nn}) &= \int_{\mathbb{R}}\int_0^1 Z_n\left(\frac{X_{1n}-x}{h},x\right)\bar{W}^n(X_{1n},x,h) \\ &\times \left(\int_{\mathbb{R}}(I_{\{y_{1n}\leq y\}}-F_{1n}(y|X_{1n}))f_{1n}(y_{1n}|X_{1n})dy_{1n}\right) \\ &\times G(y|x)w(x)w_x(y)g_x(y)dydx, \\ &= 0. \end{aligned}$$

On en déduit que $E_{F_{1n}}\left(\sum_{i=1}^n z_i\right)=0$. Maintenant, comme $E_{F_{1n}}(z_1|X_{1n},\ldots,X_{nn})=0$,

$$\begin{aligned} Var_{F_{1n}}\left(\sum_{i=1}^n z_i\right) &= Var_{F_{1n}}\left(\sum_{i=1}^n z_i\right), \\ &= E_F\left(Var_{F_{1n}}\left(\sum_{i=1}^n z_i|X_{1n},X_{2n},\ldots,X_{nn}\right)\right), \\ &= nE_F\left(E_{F_{1n}}\left(z_1^2|X_{1n},\ldots,X_{nn}\right)\right). \end{aligned} \tag{5.66}$$

De plus, en utilisant la définition de $\bar{W}^n(.,.,.)$ donnée en (2.71),

$$\begin{aligned} z_1^2 \;=\; & \frac{1}{n^2h^2}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_{1n}-x}{h},x\right)Z_n\left(\frac{X_{1n}-x'}{h},x'\right)\bar{K}\left(\frac{X_{1n}-x}{h}\right)\bar{K}\left(\frac{X_{1n}-x'}{h}\right) \\ & \times (I_{\{Y_{1n}\leq y\}}-F_{1n}(y|X_{1n}))(I_{\{Y_{1n}\leq y'\}}-F_{1n}(y'|X_{1n}))f^{-1}(x)f^{-1}(x') \\ & \times G(y|x)G(y'|x')w(x)w(x')w_x(y)w_{x'}(y')g_x(y)g_{x'}(y')dydxdy'dx', \end{aligned}$$

de sorte que,

$$\begin{aligned} E_{F_{1n}}(z_1^2|X_{1n},\ldots,X_{nn}) &\leq \frac{M}{n^2h^2}\int_0^1\int_0^1\left|\bar{K}\left(\frac{X_{1n}-x}{h}\right)\right|\left|\bar{K}\left(\frac{X_{1n}-x'}{h}\right)\right|dxdx', \\ &\leq \frac{M}{n^2h^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\bar{K}(u)\right|\left|\bar{K}(v)\right|dudv, \\ &\leq \frac{M}{n^2}. \end{aligned}$$

Si on substitue ce résultat dans (5.66), on a

$$Var_{F_{1n}}\left(\sum_{i=1}^{n} z_i\right) \leq \frac{M}{n} = o(1). \tag{5.67}$$

Par l'inégalité de Bienaymé-Tchebytchev, on déduit que

$$\sum_{i=1}^{n} z_i = o_p(1). \tag{5.68}$$

On substitue maintenant les résultats (5.65) et (5.68) dans la décomposition (5.60) pour obtenir :

$$T_6 = o_p(1).\clubsuit \tag{5.69}$$

Les Propositions 5.4.1 à 5.4.5 permettent de remarquer que seul le comportement asymptotique de T_1 et T_3 intervient dans le comportement asymptotique de la statistique de test. Celui-ci est résumé dans le théorème suivant :

Théorème 5.4.6 *Sous les hypothèses de la Proposition* 5.4.1, *on a :*

$$\frac{T - h^{-\frac{1}{2}}a_0 - b_0}{\sigma_0} \xrightarrow{\mathcal{L}} N(0,1), \tag{5.70}$$

où les quantités a_0 et σ_0^2 *sont données respectivement en* (2.15) *et* (2.69) *et où la quantité* b_0 *est définie par :*

$$b_0 = \int_0^1 \int_{\mathbb{R}} G^2(y|x)w(x)w_x(y)f_0(y|x)dydx. \tag{5.71}$$

Le Théorème 5.4.6 permet de remarquer que la puissance asymptotique locale du test basé sur T est constante pour des régions de la forme

$$\{F_0(y|x) + \left(n\sqrt{h}\right)^{-\frac{1}{2}} G(y|x)\}, \tag{5.72}$$

où la fonction $G(y|x)$ est telle que (5.71) est constant. De plus, cette puissance asymptotique locale est supérieure au niveau α du test puisque $b_0 \geq 0$. En effet, on peut réécrire (5.70) sous la forme

$$\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} \xrightarrow{\mathcal{L}} N\left(\frac{b_0}{\sigma_0}, 1\right), \tag{5.73}$$

de sorte que :

$$P\left[\frac{T - h^{-\frac{1}{2}}a_0}{\sigma_0} > z_{1-\alpha}\right] \quad \longrightarrow \quad P\left[N\left(0,1\right) > z_{1-\alpha} - \frac{b_0}{\sigma_0}\right] \geq \alpha. \tag{5.74}$$

Cette remarque permet donc de conforter le choix que nous avons fait au Chapitre 2 de considérer un test unilatéral.

Notons maintenant que le résultat obtenu au Théorème 5.4.6 peut s'étendre, à l'instar des résultats obtenus au Chapitre 3, au cas où la fonction de répartition $F_0(y|x)$ est une fonction quelconque satisfaisant néanmoins certaines conditions de régularité. La fonction $G(y|x)$ devra alors vérifier l'hypothèse supplémentaire $(G^{(p+1)}.x.2)$ pour pouvoir adapter les résultats obtenus au Chapitre 3.

Chapitre 6

Un critère de choix de la fenêtre d'ajustement

6.1 Introduction

Le but de ce chapitre est de discuter des critères de choix des différents paramètres intervenant lors de l'estimation polynômiale de la fonction de répartition conditionnelle $F(y|x)$ et qui apparaissent de manière implicite ou explicite dans les différents résultats que nous avons obtenus tout au long de ce travail. En particulier, l'estimateur polynômial local de la fonction de répartition conditionnelle dépend de trois paramètres : le degré p du polynôme local, le noyau $K(.)$ et la fenêtre d'ajustement h. Le choix de p a été discuté à la Section 3.2 où nous avons convenu de choisir une valeur de p impaire. Concernant le choix du noyau d'ajustement $K(.)$, Fan *et al.* (1997) ont donné un critère de choix dans le contexte d'estimation polynômiale locale de la fonction de régression. Nous reviendrons en fin de chapitre sur ce point mais, pour le moment, il n'apparaît pas déraisonnable de prendre par exemple le noyau d'Epanechnikov défini par :

$$K(t) = \frac{3}{4}(1 - t^2) I_{\{-1 \leq t \leq 1\}}.$$

Maintenant, le choix de la fenêtre d'ajustement est extrêmement important et c'est de ce choix dont nous allons discuter dans l'essentiel de ce chapitre.

6.2 Critère de choix de la fenêtre d'ajustement

La fenêtre h contrôle la largeur du voisinage considéré autour de x lors de l'ajustement polynômial de la fonction de répartition conditionnelle et détermine de ce fait la complexité de l'estimateur puisqu'il détermine le degré de lissage de celui-ci. Le choix d'une fenêtre trop petite reproduit presque intégralement les données et dans ce contexte l'estimateur "sous-lisse" et est donc très variable. En effet, si on considère le cas $p = 1$, quand $h = 0$, l'estimateur linéaire local de $F(y|x)$ pour y fixe est une interpolation des données $(I_{\{Y_i \leq y\}}, X_i)$ alors que pour x fixe, $\hat{F}_n(y|x) = 0$ si $X_i \neq x$ et vaut $I_{\{Y_i \leq y\}}$ si $X_i = x$. A contrario, un choix du paramètre de lissage trop grand entraîne un sur-lissage ce qui se traduit par une augmentation du biais de l'estimateur. On peut citer en exemple le cas de la régression linéaire locale où, si on choisit $h = +\infty$, la courbe de régression linéaire locale se confond avec la régression linéaire simple sur l'ensemble des points. Le but est donc de choisir une fenêtre qui équilibre asymptotiquement le biais et la variance de l'estimateur.

Une importante littérature a été consacrée au choix de h dans le contexte des méthodes d'ajustement non paramétriques. Par exemple, Hall *et al.* (1991), Sheather et Jones (1991), Chiu (1996), Jones *et al.* (1996) et plus récemment Futschik et Clarke (2004) ont discuté du choix de la fenêtre pour l'estimation de la densité. Brockmann *et al.* (1993) ont proposé un critère de sélection de la fenêtre dans le cadre de la régression. Dans un contexte plus particulier d'ajustement polynômial local, Fan et Gijbels (1992, 1995, 1996), ont aussi proposé des critères de choix de la fenêtre dans le cadre de la régression.

Il existe deux types de fenêtres : les fenêtres constantes et les fenêtres variables. Fan et Gijbels (1996) donnent, dans le cadre de la régression polynômiale locale, un critère de choix théorique pour ce dernier type de fenêtre. En adaptant ces travaux de Fan et Gijbels (1996) à notre contexte, un critère de choix pour une fenêtre d'ajustement variable peut être obtenu en considérant une mesure de perte que l'on appelle l'erreur quadratique moyenne (ou MSE= "Mean Squared Error") au point (x, y) et définie par :

$$\Big[Biais_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n)\Big]^2 + Var_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n), \tag{6.1}$$

où le premier terme est le biais de $\hat{F}_n(y|x)$ conditionnellement aux X observés et relativement à la loi conditionnelle $F_0(.\,|\,.)$ et le second est la variance de $\hat{F}_n(y|x)$, toujours conditionnellement aux X observés et relativement à la même loi. Dans ce contexte, une fenêtre "optimale" peut être celle qui minimise l'expression (6.1).

Cependant, cette fenêtre optimale est fonction de x, y et des données $X_1, X_2, \ldots, X_n$. Elle ne pourra donc pas être utilisée dans notre contexte de tests d'adéquation puisqu'on travaille avec une fenêtre d'ajustement constante. Toujours dans le cadre de la régression polynômiale locale, Fan et Gijbels (1996) proposent d'autres critères de choix pour obtenir une fenêtre optimale constante. Notre but ici est d'adapter un de ces critères à notre contexte d'estimation de la fonction de répartition conditionnelle.

On va donc considérer l'erreur quadratique moyenne intégrée (ou MISE="Mean Integrated Squared Error") définie dans notre contexte par :

$$\int_0^1 \int_{\mathbb{R}} \left[\left(Biais_{F_0} \left(\hat{F}_n(y|x)|X_1, \ldots, X_n \right) \right)^2 + Var_{F_0} \left(\hat{F}_n(y|x)|X_1, \ldots, X_n \right) \right] \pi(x)\pi_x(y) f_0(y|x) dy dx, \tag{6.2}$$

où $\pi(x)$ et $\pi_x(y)$ sont des fonctions de poids positives, qui permettent de pondérer le terme entre crochets qui correspond à l'erreur quadratique moyenne évaluée au point (x, y). Une fenêtre "optimale" peut alors être obtenue en minimisant cette fonction. Etant donné que le calcul exact de (6.2) est difficile à effectuer, on propose ici de déterminer la fenêtre optimale constante asymptotique, c'est à dire la fenêtre qui va minimiser l'expression asymptotique de la MISE.

Dans le but de simplifier notre démarche, on suppose que $F_0(y|x)$ est entièrement spécifiée. On suppose aussi que x appartient toujours à l'espace compact $[0, 1]$, lui-même appartenant au support $[-1, 2]$ de la densité $f(.)$. La proposition suivante donne l'espérance conditionnelle de l'estimateur polynômial local de la fonction de répartition conditionnelle sous de telles conditions.

Proposition 6.2.1 *Supposons que $F_0(y|x)$ vérifie les hypothèses $(F_0^{(p+1)}.x.2)$ et $(F_0^{(p+2)}.x.2)$. Alors, sous les hypothèses supplémentaires $(K.0)$, $(X.0)$, $(f.0)$ et $(H.1)$, l'espérance de $\hat{F}_n(y|x)$ conditionnellement aux X observés est donnée par :*

$$\begin{aligned} E_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n) &= F_0(y|x) + \frac{F_0^{(p+1)}(y|x)}{(p+1)!} h^{p+1} \gamma_{p+1} \\ &\quad + O_{ps}^{(x,y)}(h^{p+2}) + O_{ps}^{(x,y)}(h^{p+3}), \end{aligned} \tag{6.3}$$

où $\gamma_{p+1} = \int_{\mathbb{R}} u^k K^*(u) du$.

Preuve En utilisant la définition de $\hat{F}_n(y|x)$ donnée en (1.15), on a :

$$\begin{aligned} E_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n) &= \sum_{i=1}^n W^n \left(\frac{X_i - x}{h} \right) E_{F_0}(I_{\{Y_i \leq y\}}|X_1, \ldots, X_n), \\ &= \sum_{i=1}^n W^n \left(\frac{X_i - x}{h} \right) F_0(y|X_i). \end{aligned} \tag{6.4}$$

De plus, sous l'hypothèse $(F_0^{(p+2)}.x.2)$, la fonction qui à x associe $F_0(y|x)$ est $p + 2$ fois continûment dérivable. On effectue alors un développement de Taylor d'ordre $p + 2$ de

$F_0(y|X_i)$ autour de x : pour tout $i = 1, \ldots, n$ et il existe $x_i^* \in (X_i, x)$ tel que :

$$\begin{aligned} F_0(y|X_i) &= F_0(y|x) + (X_i - x)F_0^{(1)}(y|x) + \ldots + (X_i - x)^p \frac{F_0^{(p)}(y|x)}{p!} \\ &\quad + (X_i - x)^{p+1} \frac{F_0^{(p+1)}(y|x)}{(p+1)!} + (X_i - x)^{p+2} \frac{F_0^{(p+2)}(y|x_i^*)}{(p+2)!}, \end{aligned} \tag{6.5}$$

où on rappelle que $F_0^{(j)}(y|x)$ est la j ième dérivée de la fonction $F_0(y|x)$ par rapport à x. Si on substitue ce résultat dans l'expression (6.4) de l'espérance conditionnelle de l'estimateur, on obtient, en utilisant les propriétés de la fonction $W^n(.)$, données en (1.17),

$$\begin{aligned} E_{F_0}(\hat{F}_n(y|x)|X_1, \ldots, X_n) &= F_0(y|x) \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right) \\ &\quad + \sum_{j=1}^p \frac{F_0^{(j)}(y|x)}{j!} \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^j \\ &\quad + \frac{F_0^{(p+1)}(y|x)}{(p+1)!} \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1} \\ &\quad + \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+2} \frac{F_0^{(p+2)}(y|x_i^*)}{(p+2)!}, \\ &= F_0(y|x) + Q_{xy} + R_{xy}, \end{aligned} \tag{6.6}$$

où

$$Q_{xy} = \frac{F_0^{(p+1)}(y|x)}{(p+1)!} \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1}, \tag{6.7}$$

et

$$R_{xy} = \sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+2} \frac{F_0^{(p+2)}(y|x_i^*)}{(p+2)!}. \tag{6.8}$$

Dans la suite, à l'instar de Ruppert et Wand (1994), on va négliger le terme de reste R_{xy}. On s'intéresse donc essentiellement au terme Q_{xy} qui va “contrôler” le biais ou plutôt l'ordre du biais de l'estimateur. A partir de la définition de $W^n(.)$ donnée en (1.16) et si on pose $H = diag(1, \ldots, h^p)$,

$$\sum_{i=1}^n W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1}$$

$$
\begin{aligned}
&= e_0^T \left(X^T P X\right)^{-1} \begin{bmatrix} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1} \\ \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+2} \\ \vdots \\ \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)(X_i - x)^{2p+1} \end{bmatrix}, \\
&= e_0^T \left(\frac{1}{nh} X^T P X\right)^{-1} \times \frac{1}{nh} X^T P \begin{bmatrix} (X_1 - x)^{p+1} \\ (X_2 - x)^{p+1} \\ \vdots \\ (X_n - x)^{p+1} \end{bmatrix}, \\
&= e_0^T H \left(\frac{1}{nh} X^T P X\right)^{-1} H \times \frac{1}{nh} H^{-1} X^T P \begin{bmatrix} (X_1 - x)^{p+1} \\ (X_2 - x)^{p+1} \\ \vdots \\ (X_n - x)^{p+1} \end{bmatrix}, \\
&= e_0^T S_n^{-1}(x) \times \frac{1}{nh} H^{-1} X^T P \begin{bmatrix} (X_1 - x)^{p+1} \\ (X_2 - x)^{p+1} \\ \vdots \\ (X_n - x)^{p+1} \end{bmatrix},
\end{aligned} \tag{6.9}
$$

où $S_n(x) = H^{-1}\left(\frac{1}{nh} X^T P X\right) H^{-1}$ est la matrice de dimension $(p+1) \times (p+1)$ dont les éléments sont les $(S_{n,i+j-2}(x))_{1 \le i,j \le p+1}$ définis en (1.23). Comme à la Section 1.3.1, dénotons à présent par S la matrice de dimension $(p+1)\times(p+1)$ dont les éléments sont les $(\mu_{i+j-2})_{1\le i,j\le p+1}$ définis en (1.18). On a montré en (1.31) que si n est suffisamment grand,

$$
S_n^{-1}(x) = f^{-1}(x) S^{-1} + h O_{ps}^x(1), \tag{6.10}
$$

où $O_{ps}^x(1)$ est une matrice aléatoire de dimension $(p+1) \times (p+1)$ dont chacune des composantes est presque sûrement bornée en $x \in [0,1]$, de sorte que

$$
e_0^T S_n^{-1}(x) = f^{-1}(x) e_0^T S^{-1} + h O_{ps}^x(1)^T, \tag{6.11}
$$

où le terme $O_{ps}^x(1)$ est un vecteur aléatoire de dimension $p+1$ dont chacune des composantes est presque sûrement bornée en $x \in [0,1]$. Suite à l'expression (1.26) on a aussi :

$$
\frac{1}{nh} H^{-1} X^T P \begin{bmatrix} (X_1 - x)^{p+1} \\ (X_2 - x)^{p+1} \\ \vdots \\ (X_n - x)^{p+1} \end{bmatrix} = h^{p+1} \begin{bmatrix} S_{n,p+1}(x) \\ S_{n,p+2}(x) \\ \vdots \\ S_{n,2p+1}(x) \end{bmatrix},
$$

$$= \quad h^{p+1}\left[f(x)\mu + hO^x_{ps}(1)\right], \tag{6.12}$$

où $\mu = (\mu_{p+1}, \ldots, \mu_{2p+1})^T$. Si on substitue les résultats (6.11) et (6.12) dans l'expression (6.9), on obtient,

$$\begin{aligned}\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1} &= \left[f^{-1}(x)e_0^T S^{-1} + hO^x_{ps}(1)^T\right] \\ &\quad \times h^{p+1}\left[f(x)\mu + hO^x_{ps}(1)\right], \\ &= h^{p+1}e_0^T S^{-1}\mu + O^x_{ps}(h^{p+2}) \\ &\quad + O^x_{ps}(h^{p+3}). \end{aligned} \tag{6.13}$$

De plus,

$$\begin{aligned} e_0^T S^{-1}\mu &= e_0^T S^{-1}[\mu_{p+1}, \ldots, \mu_{2p+1}]^T, \\ &= \int_{\mathbb{R}} e_0^T S^{-1}[1, \ldots, t^p]^T K(t) t^{p+1} dt, \\ &= \int_{\mathbb{R}} K^*(t) t^{p+1} dt, \\ &= \gamma_{p+1}, \end{aligned} \tag{6.14}$$

par la définition de $K^*(.)$ donnée en (1.21). Si on substitue ce résultat dans (6.13), on obtient :

$$\begin{aligned}\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1} &= h^{p+1}\gamma_{p+1} + O^x_{ps}(h^{p+2}) \\ &\quad + O^x_{ps}(h^{p+3}), \end{aligned} \tag{6.15}$$

de sorte que, sous l'hypothèse $(F_0^{(p+1)}.x.2)$,

$$\begin{aligned} Q_{xy} &= \frac{F_0^{(p+1)}(y|x)}{(p+1)!}\sum_{i=1}^{n} W^n\left(\frac{X_i - x}{h}\right)(X_i - x)^{p+1}, \\ &= h^{p+1}\gamma_{p+1}\frac{F_0^{(p+1)}(y|x)}{(p+1)!} + O^{(x,y)}_{ps}(h^{p+2}) + O^{(x,y)}_{ps}(h^{p+3}). \end{aligned} \tag{6.16}$$

Comme l'ordre du terme de reste est négligeable par rapport à l'ordre des termes issus de l'étude asymptotique de Q_{xy}, on en déduit alors le résultat. ♣

Passons maintenant à l'étude de la variance conditionnelle de l'estimateur polynômial de $F(y|x)$. Son expression est donnée dans la proposition suivante :

Proposition 6.2.2 *Sous les hypothèses* $(K.0)$, $(X.0)$, $(f.0)$ *et* $(H.1)$,

$$Var_{F_0}\left(\hat{F}_n(y|x)|X_1,\ldots,X_n\right) = K^{*^{(2)}}(0)\left[\frac{F_0(y|x)(1-F_0(y|x))}{nhf(x)}\right] + o_p\left(\frac{1}{nh}\right). \tag{6.17}$$

Preuve A partir de sa définition donnée en (1.14),

$$\hat{F}_n(y|x) = e_0^T\left(X^TPX\right)^{-1}X^TPI_{\{Y\le y\}},$$

où les matrices X, P et le vecteur $I_{\{Y\le y\}}$ ont été définis respectivement en (1.10), (1.11) et (1.12). On en déduit que :

$$\begin{aligned}
&Var_{F_0}\left(\hat{F}_n(y|x)|X_1,\ldots,X_n\right)\\
&= e_0^T\left(X^TPX\right)^{-1}X^TP\,Var_{F_0}\left(I_{\{Y\le y\}}|X_1,\ldots,X_n\right)PX\left(X^TPX\right)^{-1}e_0,
\end{aligned}$$

car la matrice P est diagonale et les matrices P et X dépendent des variables X_i, $i=1,\ldots,n$. De plus,

$$Var_{F_0}\left(I_{\{Y\le y\}}|X_1,\ldots,X_n\right) = Q,$$

où Q est la matrice de dimension $n\times n$ diagonale (puisque les variables Y_i sont *iid*), dont les éléments diagonaux sont donnés par :

$$Q_i = F_0(y|X_i)(1-F_0(y|X_i)), \quad \forall i=1,\ldots,n.$$

Si maintenant on pose $\Sigma = PQP$, la matrice diagonale de dimension $n\times n$ (les matrices P et Q étant diagonales), dont les éléments sont :

$$\Sigma_i = K^2\left(\frac{X_i-x}{h}\right)\times\left[F_0(y|X_i)(1-F_0(y|X_i))\right], \quad \forall i=1,\ldots,n,$$

on obtient :

$$\begin{aligned}
Var_{F_0}\left(\hat{F}_n(y|x)|X_1,\ldots,X_n\right) &= e_0^T\left(X^TPX\right)^{-1}X^T\Sigma X\left(X^TPX\right)^{-1}e_0,\\
&= \frac{1}{nh}e_0^T\left(\frac{1}{nh}X^TPX\right)^{-1}\frac{1}{nh}X^T\Sigma X\left(\frac{1}{nh}X^TPX\right)^{-1}e_0,
\end{aligned}$$

de sorte que, si on pose $S_n^*(x,y) = H^{-1}\left(\frac{1}{nh}X^T\Sigma X\right)H^{-1}$, la matrice de dimension $(p+1)\times(p+1)$ dont les éléments sont définis par $(S_{n,i+j-2}^*(x,y))_{1\leq i,j\leq p+1}$, avec

$$S_{n,k}^*(x,y) = \frac{1}{nh}\sum_{i=1}^n K^2\left(\frac{X_i-x}{h}\right)\left(\frac{X_i-x}{h}\right)^k\Big[F_0(y|X_i)(1-F_0(y|X_i))\Big], \quad \forall k=0,\ldots,2p,$$

et si on utilise la matrice $S_n(x)$, déjà introduite dans la preuve de la Proposition 6.2.1,

$$Var_{F_0}\left(\hat{F}_n(y|x)|X_1,\ldots,X_n\right) = \frac{1}{nh}e_0^T S_n^{-1}(x)S_n^*(x,y)S_n^{-1}(x)e_0.$$

Cette expression de la variance conditionnelle n'est cependant pas directement utilisable malgré le fait qu'elle soit exacte. On cherche donc à approximer les matrices $S_n(x)$ et $S_n^*(x,y)$. On a déjà noté en (1.31) que :

$$S_n^{-1}(x) = \frac{S^{-1}}{f(x)} + hO_{ps}^x(1), \tag{6.18}$$

où le terme $O_{ps}^x(1)$ est presque sûrement borné en $x\in[0,1]$. Il reste donc à approximer la matrice $S_n^*(x,y)$. Cette approximation nécessite au préalable le résultat suivant :

Lemme 6.2.3 *Sous les hypothèses* $(K.0)$, $(X.0)$, $(f.0)$ *et* $(H.1)$,

$$S_{n,k}^*(x,y) = \nu_k f(x)F_0(y|x)(1-F_0(y|x)) + o_p(1),$$

où $\nu_k = \int_{\mathbb{R}} u^k K^2(u)du.$

Preuve Voir Annexe G. ♣

Ce résultat permet de déduire que :

$$S_n^*(x,y) = f(x)F_0(y|x)(1-F_0(y|x))S^* + o_p(1), \tag{6.19}$$

où S^* est la matrice de dimension $(p+1)\times(p+1)$ dont les éléments sont les $(\nu_{i+j-2})_{1\leq i,j\leq p+1}$ et où $o_p(1)$ est la matrice aléatoire de dimension $(p+1)\times(p+1)$ dont chacun des éléments converge en probabilité vers 0 pour tout $x\in[0,1]$ et tout $y\in\mathbb{R}$. On en déduit que :

$$\begin{aligned} Var_{F_0}\left(\hat{F}_n(y|x)|X_1,\ldots,X_n\right) &= \frac{1}{nh}e_0^T\left[\frac{S^{-1}}{f(x)} + hO_{ps}^x(1)\right] \\ &\quad \times\Big[f(x)F_0(y|x)(1-F_0(y|x))S^* + o_p(1)\Big] \end{aligned}$$

$$
\begin{aligned}
& \times \left[\frac{S^{-1}}{f(x)} + hO^x_{ps}(1)\right] e_0, \\
= \; & \frac{1}{nhf(x)} F_0(y|x)(1 - F_0(y|x)) e_0^T S^{-1} S^* S^{-1} e_0 \\
& + o_p\left(\frac{1}{nh}\right). \qquad (6.20)
\end{aligned}
$$

Enfin,

$$
\begin{aligned}
e_0^T S^{-1} S^* S^{-1} e_0 &= e_0^T S^{-1} \int_{\mathbb{R}} K^2(t) \begin{pmatrix} 1 & t & \dots & t^p \\ t & t^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ t^p & \dots & \dots & t^{2p} \end{pmatrix} dt S^{-1} e_0, \\
&= \int_{\mathbb{R}} e_0^T S^{-1} (1, t, \dots, t^p)^T K^2(t)(1, t, \dots, t^p) S^{-1} e_0 dt, \\
&= \int_{\mathbb{R}} K^{*^2}(t) dt, \\
&= K^{*^{(2)}}(0), \qquad (6.21)
\end{aligned}
$$

puisque, suite à la Proposition 1.4.2, le noyau $K^*(.)$ est symétrique. Enfin, si on substitue ce résultat dans l'expression (6.20), on obtient :

$$
Var_{F_0}(\hat{F}_n(y|x)|X_1, \dots, X_n) = \left[\frac{F_0(y|x)(1 - F_0(y|x))}{nhf(x)}\right] K^{*^{(2)}}(0) + o_p\left(\frac{1}{nh}\right). \clubsuit \qquad (6.22)
$$

On remarque que cette expression de la variance asymptotique est indépendante du choix du paramètre p contrairement à l'expression de l'espérance conditionnelle de l'estimateur obtenue en (6.3).

Les Propositions 6.2.1 et 6.2.2 permettent à présent de donner un critère de choix optimal de la fenêtre h. Tout d'abord, par la Proposition 6.2.1 et sous les hypothèses données,

$$
\left[Biais_{F_0}(\hat{F}_n(y|x)|X_1, \dots, X_n)\right]^2 \approx \left[\frac{F_0^{(p+1)}(y|x)}{(p+1)!}\right]^2 h^{2(p+1)} \gamma_{p+1}^2. \qquad (6.23)
$$

Si on substitue le terme dominant de (6.22) et (6.23) dans l'expression (6.2), on est amené à chercher la valeur de h qui minimise la fonction :

$$\int_0^1 \int_{\mathbb{R}} \left[\gamma_{p+1}^2 \left[\frac{F_0^{(p+1)}(y|x)}{(p+1)!} \right]^2 h^{2(p+1)} + K^{*^{(2)}}(0) \left[\frac{F_0(y|x)(1-F_0(y|x))}{nhf(x)} \right] \right] \pi(x)\pi_x(y) f_0(y|x) dy dx. \tag{6.24}$$

Il faut donc trouver h_{opt} qui résoud l'équation :

$$\frac{d}{dh} \left[\int_{\mathbb{R}} \int_0^1 H(x,y,h)\pi(x)\pi_x(y) f_0(y|x) dy dx \right] |_{h_{opt}} = 0,$$

où

$$H(x,y,h) = \gamma_{p+1}^2 \left[\frac{F_0^{(p+1)}(y|x)}{(p+1)!} \right]^2 h^{2(p+1)} + K^{*^{(2)}}(0) \left[\frac{F_0(y|x)(1-F_0(y|x))}{nhf(x)} \right].$$

Quelques manipulations algébriques montrent que :

$$h_{opt} = C_p(K) \left[\frac{\int_0^1 \int_{\mathbb{R}} \left[\frac{F_0(y|x)(1-F_0(y|x))}{f(x)} \right] \pi(x)\pi_x(y) f_0(y|x) dy dx}{\int_0^1 \int_{\mathbb{R}} \left[F_0^{(p+1)}(y|x) \right]^2 \pi(x)\pi_x(y) f_0(y|x) dy dx} \right]^{\frac{1}{2p+3}} n^{-\frac{1}{2p+3}},$$

où $C_p(K)$ est la constante définie par :

$$C_p(K) = \left[\frac{(p+1)! K^{*^{(2)}}(0)}{2(p+1)\gamma_{p+1}^2} \right]^{\frac{1}{2p+3}}. \tag{6.25}$$

Dans cette expression, on constate que par un choix judicieux et approprié des fonctions de poids, les intégrales sont finies et le dénominateur est non nul. En revanche, cette expression théorique de la fenêtre asymptotique optimale n'est pas directement utilisable en pratique. En effet, elle comporte entre autres le "paramètre" de nuisance $f(x)$ qui est généralement inconnu en pratique. Une solution consiste alors à remplacer cette quantité inconnue par un estimateur non paramétrique comme celui déjà introduit à la Section 2.5 et dénoté par $\hat{f}_n(x)$, pour obtenir la fenêtre optimale asymptotique approchée suivante :

$$\hat{h}_{opt} = C_p(K) \left[\frac{\int_0^1 \int_{\mathbb{R}} \left[\frac{F_0(y|x)(1-F_0(y|x))}{\hat{f}_n(x)} \right] \pi(x)\pi_x(y) f_0(y|x) dy dx}{\int_0^1 \int_{\mathbb{R}} \left[F_0^{(p+1)}(y|x) \right]^2 \pi(x)\pi_x(y) f_0(y|x) dy dx} \right]^{\frac{1}{2p+3}} n^{-\frac{1}{2p+3}}.$$

Si on adapte maintenant les résultats obtenus au contexte où $F_0(y|x)$ n'est plus entièrement spécifiée mais comporte un certain nombre de paramètres inconnus que l'on retrouve dans le vecteur θ_0, la fenêtre asymptotique optimale théorique constante est alors donnée par :

$$\hat{h}_{opt} = C_p(K) \left[\frac{\int_0^1 \int_{\mathbb{R}} \left[\frac{F_0(y|x;\theta_0)(1-F_0(y|x;\theta_0))}{\hat{f}_n(x)} \right] \pi(x)\pi_x(y) f_0(y|x;\theta_0) dy dx}{\int_0^1 \int_{\mathbb{R}} [F_0^{(p+1)}(y|x;\theta_0)]^2 \pi(x)\pi_x(y) f_0(y|x;\theta_0) dy dx} \right]^{\frac{1}{2p+3}} n^{-\frac{1}{2p+3}}.$$

Cette expression n'est là encore pas directement utilisable en pratique car elle comporte le paramètre θ_0 qui est inconnu. Une solution, proposée dans d'autres contextes, comme celui de l'estimation de la fonction de régression, consiste à remplacer ce paramètre inconnu par un estimateur : c'est un exemple de ces méthodes dites "plug-in" qui permettent d'obtenir une valeur approchée de la fenêtre h optimale. Ces méthodes offrent en général de bonnes performances comparé aux méthodes classiques comme le montrent Jones *et al.* (1996) ou bien encore Chiu (1996) dans un contexte d'estimation de la densité. Maintenant, les arguments favorisant ce type de méthodes ont été plus récemment discutés. En effet, comme en fait état Loader (1999), toujours dans le cadre de l'estimation de la densité, le comportement des méthodes de type "plug-in" dépend lourdement du choix de l'estimateur pilote $\hat{f}_n(.)$ qui, mal choisi, peut mener à une fenêtre complètement inappropriée.

La littérature abonde d'autres méthodes permettant d'obtenir une fenêtre raisonnable pour des contextes tels que l'estimation de la densité ou de la fonction de régression. On peut notamment citer les méthodes de validations croisées utilisées entre autres par Hall *et al.* (1992). Celles-ci présentent aussi des inconvénients car elles sous-lissent fréquemment. Jones *et al.* (1996) présentent et comparent les différentes méthodes de choix de fenêtres dans un contexte d'estimation de la densité. Plus récemment, Futschik et Clarke (2004) proposent un nouveau critère de choix du paramètre de lissage toujours dans ce même contexte. Ils suggèrent une méthode qui permet de palier à l'ensemble des inconvénients liés à chacune des méthodes que nous avons évoquées précédemment.

Il serait donc judicieux, en perspective de ce travail, de faire un inventaire de ces différentes méthodes et d'essayer d'adapter certaines d'entre elles à notre contexte d'estimation de la fonction de répartition conditionnelle. Notons aussi que la fenêtre optimale obtenue dans ce chapitre s'inscrit dans le contexte d'estimation de la fonction de répartition. Or, rien ne dit que la fenêtre optimale pour l'estimation est la même que la fenêtre optimale pour le test. Cela pourrait faire partie d'une recherche future tout comme le choix d'un noyau d'ajustement optimal $K(.)$ pour le test. En effet, pour le problème de test, Ghosh et Huang (1991) ont montré que d'autres critères que la MISE peuvent être employés pour choisir de façon optimale le noyau et que ce noyau optimal peut différer du noyau optimal dans un contexte d'estimation. Maintenant, si on s'attache à déterminer le noyau optimal pour le problème d'estimation, il faut alors choisir celui qui minimise la constante (6.25). Cela nous amènerait donc à travailler avec le noyau équivalent $K^*(.)$, défini en (1.21), et dont la structure a déjà été étudiée dans de nombreux travaux, notamment ceux de Berlinet (1993) et Berlinet et Thomas-Agnan (2004), pour lesquels ce noyau s'inscrit dans une terminologie de noyau d'ordre $p+1$.

Cela nous amène à présent à donner un bilan de ce qui a été fait dans ce travail et à discuter des nombreuses pistes de recherche qui pourraient faire l'objet d'une suite pour celui-ci.

Conclusion et Perspectives

Le point de départ de ce travail est un modèle posé par le statisticien pour tenter d'expliquer le comportement d'une variable aléatoire Y à partir d'une autre variable X. Nous avons proposé une procédure de test d'adéquation qui permet de tester globalement toutes les hypothèses incorporées dans le modèle considéré.

Les principaux résultats que nous avons obtenus dans cette thèse sont les suivants. Le comportement asymptotique de notre statistique de test converge toujours vers une loi normale standard dans tous les cas de figure que nous avons considérés. Nous avons aussi fait l'étude de la puissance asymptotique locale de notre test. Cela nous a permis d'obtenir que cette puissance locale relève à nouveau d'une loi normale, cette fois-ci excentrée. Enfin, nous avons proposé un critère de choix optimal pour la fenêtre d'ajustement.

Cette étude est cependant loin d'être complète et de nombreuses perspectives de travail peuvent être envisagées. Tout d'abord, concernant le choix de la fenêtre d'ajustement h, le critère de choix que nous avons proposé s'insère dans un contexte d'estimation non paramétrique. Or, rien ne dit que la fenêtre optimale pour le test est la même que la fenêtre optimale pour l'estimation. Cela reste donc une piste de recherche à explorer. Dans un même registre, il serait intéressant de voir si un choix judicieux du noyau d'ajustement n'aurait pas une influence positive sur la puissance du test.

C'est pourquoi il serait aussi pertinent de mettre en place quelques études de simulations. Cela permettrait aussi de vérifier la justesse des approximations asymptotiques obtenues et de mettre l'accent sur des thèmes que nous n'avons pas abordés dans ce travail, notamment concernant le choix des fonctions de poids introduites dans les différentes statistiques de test. Dans un contexte de régression, de nombreux auteurs (Alcalà *et al.*, 1999 et Härdle et Mammen, 1993) ont montré que les comportements asymptotiques de leurs procédures de test ont une faible vitesse de convergence. Différentes techniques de rééchantillonnage ont été développées, notamment le Wild Bootstrap (Härdle et Mammen, 1993) pour palier à ce problème. Il serait donc intéressant de voir si ces techniques sont pertinentes dans le présent contexte et, si tel est le cas, de réfléchir à comment les adapter à notre procédure de test.

Dans ce travail, nous avons choisi d'utiliser des statistiques de tests basées sur la "distance" de type Cramér-von Mises (Cramér, 1928 et von Mises, 1941) généralisée. Aussi, un axe de recherche qu'il serait intéressant d'envisager est l'utilisation d'autres types de "distance" dans notre procédure. On pourrait par exemple considérer des semi-normes basées sur le processus

$$\sqrt{nh}\left(\hat{F}_n(y|x) - F_0(y|x)\right).$$

Certaines d'entre elles, fréquemment utilisées dans la littérature, pourraient être choisies comme statistiques de test comme par exemple la "distance du sup" de Kolmogorov (1933), qui mène à la statistique de test :

$$\sqrt{nh}\sup_{x\in[0,1]}\sup_{y\in\mathbb{R}}\left|\hat{F}_n(y|x) - F_0(y|x)\right|.$$

On pourrait aussi proposer une statistique de test inspirée de celle proposée par Watson (1961) qui reprend la statistique de Cramér-von Mises (Cramér, 1928 et von Mises, 1941) mais propose de centrer $\hat{F}_n(y|x) - F_0(y|x)$ pour obtenir une statistique de test de la forme :

$$n\sqrt{h}\int_0^1\int_{\mathbb{R}}\left[\hat{F}_n(y|x) - F_0(y|x) - \int_0^1\int_{\mathbb{R}}\left(\hat{F}_n(y|x) - F_0(y|x)\right)F_0(dy|x)\right]^2 F_0(dy|x)dx.$$

Enfin, certains auteurs, comme par exemple Gregory (1977), proposent encore d'autres statistiques de test qu'il serait peut-être intéressant d'examiner dans notre contexte. Il serait aussi opportun d'essayer de trouver une décomposition de la statistique de test globale en un certain nombre de composantes directionnelles qui permettraient de tester ces hypothèses spécifiques et donner des indications sur les aspects de $F_0(y|x)$ qui ne collent pas aux données lorsque H_0 est rejetée.

Un axe de recherche qu'il serait intéressant d'explorer est la généralisation de notre procédure de test au cas où la variable X est multivariée. Plus particulièrement, on pourrait envisager des adaptations "pratiques" de notre procédure de test à des modèles de type

GLM ("Generalized Linear Models"). En effet, cette classe de modèles généralisant les modèles linéaires classiques permet l'analyse de données discrètes et continues pour lesquelles la loi normale n'est plus adaptée. Ce genre de modèles, et l'ensemble des modèles qui en sont issus, occupent maintenant une place importante dans la modélisation statistique et trouvent un intérêt dans de nombreux domaines d'applications.

Enfin, nous avons étudié la puissance asymptotique locale de notre test en considérant des alternatives inspirées de la famille d'alternatives locales de type Pitman (1949). Il serait intéressant de voir ce que devient cette fonction de puissance asymptotique locale en utilisant d'autres types d'alternatives comme par exemple, une de celles référencées dans l'article de Inglot et Ledwina (2001).

Annexe A

Preuve du Lemme 1.4.1 Soit, pour $k \in \{0, 1, \ldots, 2p\}$,

$$S_{n,k}(x) = \frac{1}{nh} \sum_{i=1}^{n} \left(\frac{X_i - x}{h} \right)^k K \left(\frac{X_i - x}{h} \right), \tag{A.1}$$

où les $X_1, \ldots, X_n$ sont $iid \sim F(.)$. On considère maintenant :

$$\begin{aligned} \sup_{x \in [a,b]} |S_{n,k}(x) - f(x)\mu_k| &= \sup_{x \in [a,b]} |S_{n,k}(x) - E_F(S_{n,k}(x)) + E_F(S_{n,k}(x)) - f(x)\mu_k|, \\ &\leq \sup_{x \in [a,b]} |S_{n,k}(x) - E_F(S_{n,k}(x))| \\ &\quad + \sup_{x \in [a,b]} |E_F(S_{n,k}(x)) - f(x)\mu_k|, \end{aligned} \tag{A.2}$$

où

$$\begin{aligned} E_F(S_{n,k}(x)) &= \frac{1}{h} E_F \left[\left(\frac{X_i - x}{h} \right)^k K \left(\frac{X_i - x}{h} \right) \right], \\ &= \frac{1}{h} \int_{-1}^{2} \left(\frac{y - x}{h} \right)^k K \left(\frac{y - x}{h} \right) dF(y). \end{aligned} \tag{A.3}$$

Si on injecte (A.3) dans le premier terme de (A.2), en signalant que :

$$S_{n,k}(x) = \frac{1}{h} \int_{-1}^{2} \left(\frac{y - x}{h} \right)^k K \left(\frac{y - x}{h} \right) d\hat{F}_n(y), \tag{A.4}$$

où $\hat{F}_n(y)$ est la fonction de répartition empirique, on obtient :

$$\sup_{x \in [a,b]} |S_{n,k}(x) - E_F(S_{n,k}(x))| = \sup_{x \in [a,b]} \frac{1}{h} \left| \int_{-1}^{2} \left(\frac{y - x}{h} \right)^k K \left(\frac{y - x}{h} \right) d\left(\hat{F}_n(y) - F(y) \right) \right|.$$

On intègre par parties en posant $u_k(y) = \left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right)$ et $v(y) = \hat{F}_n(y) - F(y)$, pour obtenir :

$$
\begin{aligned}
\int_{-1}^{2}\left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right) d\left(\hat{F}_n(y)-F(y)\right) &= [u_k(y)v(y)]_{-1}^{2} - \int_{-1}^{2}\left(\hat{F}_n(y)-F(y)\right) du_k(y), \\
&= \left[\left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right)\left(\hat{F}_n(y)-F(y)\right)\right]_{-1}^{2} \\
&\quad - \int_{-1}^{2}\left(\hat{F}_n(y)-F(y)\right) du_k(y). \qquad \text{(A.5)}
\end{aligned}
$$

Mais,

$$
\begin{aligned}
\left[\left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right)\left(\hat{F}_n(y)-F(y)\right)\right]_{-1}^{2} &= \left(\frac{2-x}{h}\right)^k K\left(\frac{2-x}{h}\right)\left(\hat{F}_n(2)-F(2)\right) \\
&\quad - \left(\frac{-1-x}{h}\right)^k K\left(\frac{-1-x}{h}\right)\left(\hat{F}_n(-1)-F(-1)\right), \\
&= 0,
\end{aligned}
$$

de sorte que,

$$
\int_{-1}^{2}\left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right) d\left(\hat{F}_n(y)-F(y)\right) = -\int_{-1}^{2}\left(\hat{F}_n(y)-F(y)\right) du_k(y). \qquad \text{(A.6)}
$$

Maintenant,

$$
\begin{aligned}
du_k(y) &= \frac{d}{dy}\left(\frac{y-x}{h}\right)^k K\left(\frac{y-x}{h}\right), \\
&= \frac{1}{h}\left(\frac{y-x}{h}\right)^{k-1}\left[kK\left(\frac{y-x}{h}\right) + \left(\frac{y-x}{h}\right)K'\left(\frac{y-x}{h}\right)\right]. \qquad \text{(A.7)}
\end{aligned}
$$

Ainsi, en injectant (A.7) dans (A.5), on obtient

$$
\sup_{x\in[a,b]} |S_{n,k}(x) - E_F(S_{n,k}(x))|
$$

$$
\begin{aligned}
&= \sup_{x\in[a,b]} \frac{1}{h}\left|\int_{-1}^{2}\left(\hat{F}_n(y)-F(y)\right)\left[\frac{1}{h}\left(\frac{y-x}{h}\right)^{k-1}\left[kK\left(\frac{y-x}{h}\right)+\left(\frac{y-x}{h}\right)K'\left(\frac{y-x}{h}\right)\right]dy\right]\right|,\\
&\leq \frac{1}{h}\sup_{y\in[-1,2]}\left|\hat{F}_n(y)-F(y)\right| \sup_{x\in[a,b]}\int_{-1}^{2}\frac{1}{h}\left|\left(\frac{y-x}{h}\right)^{k-1}\left[kK\left(\frac{y-x}{h}\right)+\left(\frac{y-x}{h}\right)K'\left(\frac{y-x}{h}\right)\right]\right|dy.
\end{aligned}
$$

On effectue à présent le changement de variables $w=\dfrac{y-x}{h}$, de sorte que, comme $h<1$ et $x\in[a,b]$,

$$
\begin{aligned}
&\sup_{x\in[a,b]}\int_{-1}^{2}\frac{1}{h}\left|\left(\frac{y-x}{h}\right)^{k-1}\left[kK\left(\frac{y-x}{h}\right)+\left(\frac{y-x}{h}\right)K'\left(\frac{y-x}{h}\right)\right]\right|dy\\
&= \sup_{x\in[a,b]}\int_{\frac{-1-x}{h}}^{\frac{2-x}{h}}\left|w^{k-1}\left[kK(w)+wK'(w)\right]\right|dw,\\
&\leq \int_{\mathbb{R}}|w|^{k-1}\left[kK(w)+|w|\,|K'(w)|\right]dw,\\
&= \alpha_k<+\infty, \qquad\qquad (A.8)
\end{aligned}
$$

sous l'hypothèse $(K.0)$. De plus, de Rao (1983, p48), on a que presque sûrement :

$$
\sup_{y\in[-1,2]}\left|\hat{F}_n(y)-F(y)\right|\leq\frac{\sqrt{\log n}}{\sqrt{n}}. \qquad (A.9)
$$

On en déduit qu'avec probabilité 1, si n est suffisamment grand,

$$
\sup_{x\in[a,b]}\left|S_{n,k}(x)-E_F\left(S_{n,k}(x)\right)\right|\leq\frac{\alpha_k\sqrt{\log n}}{h\sqrt{n}}. \qquad (A.10)
$$

Considérons maintenant le deuxième terme de (A.2). On a, si h est suffisamment petit,

$$
\begin{aligned}
\sup_{x\in[a,b]}\left|E_F\left(S_{n,k}(x)\right)-f(x)\mu_k\right| &= \sup_{x\in[a,b]}\left|\frac{1}{h}\int_{-1}^{2}\left(\frac{y-x}{h}\right)^{k}K\left(\frac{y-x}{h}\right)f(y)dy-f(x)\mu_k\right|,\\
&= \sup_{x\in[a,b]}\left|\int_{\frac{-1-x}{h}}^{\frac{2-x}{h}}w^kK(w)f(x+wh)dw-f(x)\int_{-1}^{1}w^kK(w)dw\right|,\\
&= \sup_{x\in[a,b]}\left|\int_{-1}^{1}w^kK(w)f(x+wh)dw-f(x)\int_{-1}^{1}w^kK(w)dw\right|,
\end{aligned}
$$

en supposant que h satisfasse :

$$\frac{-1-a}{h} < -1 \text{ et } \frac{2-b}{h} > 1,$$

c'est à dire

$$h < a+1 \text{ et } h < 2-b.$$

Maintenant, puisque par $(f.0)$ la densité $f(.)$ est Lipschitzienne,

$$\begin{aligned}
\sup_{x\in[a,b]} \left|E_F\left(S_{n,k}(x)\right) - f(x)\mu_k\right| &\leq \sup_{x\in[a,b]} \left|\int_{-1}^{1} w^k K(w)(f(x+wh)-f(x))dw\right|, \\
&\leq \sup_{x\in[a,b]} \int_{-1}^{1} |w|^k K(w)|f(x+wh)-f(x)|dw, \\
&\leq Mh \int_{-1}^{1} |w|^{k+1} K(w)dw, \\
&= Mh\kappa_{k+1} < +\infty. \qquad \text{(A.11)}
\end{aligned}$$

En combinant (A.10) et (A.11), on obtient que presque sûrement et si n est suffisamment grand :

$$\begin{aligned}
\sup_{x\in[a,b]} \left|S_{n,k}(x) - f(x)\mu_k\right| &\leq \frac{\alpha_k\sqrt{\log n}}{h\sqrt{n}} + Mh\kappa_{k+1}, \\
&= h\left(\frac{\alpha_k\sqrt{\log n}}{h^2\sqrt{n}} + M\kappa_{k+1}\right), \\
&= h\left(o(1) + M\kappa_{k+1}\right), \qquad \text{(A.12)}
\end{aligned}$$

sous l'hypothèse $(H.2)$. ♣

Annexe B

Preuve de la Proposition 1.4.2 Par (1.21),

$$K^*(t) = e_0^T S^{-1}(1, t, \ldots, t^p)^T K(t), \tag{B.1}$$

où S est la matrice de dimension $(p+1) \times (p+1)$ dont les éléments sont les $(\mu_{i+j-2})_{1 \leq i,j \leq p+1}$ donnés en (1.18). Une expression plus simple de $K^*(t)$ est obtenue en étudiant la structure de la matrice S^{-1}. Le noyau d'ajustement $K(.)$ étant symétrique et satisfaisant $(K.0)$, la matrice S possède la structure particulière

$$S = \begin{pmatrix} \times & 0 & \times & \ldots \\ 0 & \times & 0 & \ldots \\ \times & 0 & \times & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{B.2}$$

où "$\times$" désigne un nombre quelconque. Fan *et al.* (1996, p124) ont montré que la matrice S^{-1} possède alors la même structure particulière que S. On en déduit que $e_0^T S^{-1}$ est de la forme $(d_1, 0, d_3, 0, \ldots)$. Pour obtenir $K^*(t)$, il ne reste plus qu'à multiplier ce vecteur terme à terme par $(K(t), tK(t), \ldots, t^p K(t))^T$. On obtient alors

$$K^*(t) = \sum_{k=1,\ k\,:\,impair}^{p+1} d_k t^{k-1} K(t). \tag{B.3}$$

Cette dernière expression permet de voir que sous (K.0), le noyau équivalent $K^*(t)$ est symétrique autour de 0. De plus, il est à support compact $[-1, 1]$, borné, dérivable et de dérivée borné. ♣

Annexe C

Preuve du Lemme 2.4.5 On va montrer que $E_{F_0}(k_{12n}^4) = o(n^{-2})$. On rappelle que dans ce qui suit, M est une constante majorante qui peut varier d'une ligne à l'autre. Par les définitions de $\tilde{W}^n(.,.,.)$ et de k_{ijn} données respectivement en (2.6) et (2.13) et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned}
0 \leq E_{F_0}(k_{12n}^4) &= n^{-4}h^{-6}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4}\int_{\mathbb{R}^4} K^*\left(\frac{x_1-x}{h}\right)K^*\left(\frac{x_2-x}{h}\right)\\
&\times K^*\left(\frac{x_1-x^{'}}{h}\right)K^*\left(\frac{x_2-x^{'}}{h}\right)K^*\left(\frac{x_1-x^{''}}{h}\right)\\
&\times K^*\left(\frac{x_2-x^{''}}{h}\right)K^*\left(\frac{x_1-x^{'''}}{h}\right)K^*\left(\frac{x_2-x^{'''}}{h}\right)\\
&\times \left(\int_{\mathbb{R}}(I_{\{y_1\leq y\}}-F_0(y|x_1))(I_{\{y_1\leq y'\}}-F_0(y'|x_1))\right.\\
&\left.\times (I_{\{y_1\leq y^{''}\}}-F_0(y^{''}|x_1))(I_{\{y_1\leq y^{'''}\}}-F_0(y^{'''}|x_1))f_0(y_1|x_1)dy_1\right)\\
&\times \left(\int_{\mathbb{R}}(I_{\{y_2\leq y\}}-F_0(y|x_2))(I_{\{y_2\leq y'\}}-F_0(y'|x_2))\right.\\
&\left.\times (I_{\{y_2\leq y^{''}\}}-F_0(y^{''}|x_2))(I_{\{y_2\leq y^{'''}\}}-F_0(y^{'''}|x_2))f_0(y_2|x_2)dy_2\right)\\
&\times w(x)w(x^{'})w(x^{''})w(x^{'''})w_x(y)w_{x'}(y^{'})w_{x''}(y^{''})w_{x'''}(y^{'''})\\
&\times f_0(y|x)f_0(y'|x')f_0(y''|x'')f_0(y'''|x''')dy\,dy^{'}\,dy^{''}\,dy^{'''}\\
&\times f^{-2}(x)f^{-2}(x^{'})f^{-2}(x^{''})f^{-2}(x^{'''})dxdx'dx''dx'''
\end{aligned}$$

$$\times f(x_1)f(x_2)dx_1dx_2,$$

$$\leq n^{-4}h^{-6}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4}\int_{\mathbb{R}^4}|K^*\left(\frac{x_1-x}{h}\right)K^*\left(\frac{x_2-x}{h}\right)$$

$$\times K^*\left(\frac{x_1-x^{'}}{h}\right)K^*\left(\frac{x_2-x^{'}}{h}\right)K^*\left(\frac{x_1-x^{''}}{h}\right)$$

$$\times K^*\left(\frac{x_2-x^{''}}{h}\right)K^*\left(\frac{x_1-x^{'''}}{h}\right)K^*\left(\frac{x_2-x^{'''}}{h}\right)|$$

$$\times\left(\int_{\mathbb{R}}|I_{\{y_1\leq y\}}-F_0(y|x_1)||I_{\{y_1\leq y'\}}-F_0(y'|x_1)|\right.$$

$$\left.\times|I_{\{y_1\leq y''\}}-F_0(y^{''}|x_1)||I_{\{y_1\leq y'''\}}-F_0(y^{'''}|x_1)|f_0(y_1|x_1)dy_1\right)$$

$$\times\left(\int_{\mathbb{R}}|I_{\{y_2\leq y\}}-F_0(y|x_2)||I_{\{y_2\leq y'\}}-F_0(y'|x_2)|\right.$$

$$\left.\times|I_{\{y_2\leq y''\}}-F_0(y^{''}|x_2)||I_{\{y_2\leq y'''\}}-F_0(y^{'''}|x_2)|f_0(y_2|x_2)dy_2\right)$$

$$\times w(x)w(x^{'})w(x^{''})w(x^{'''})w_x(y)w_{x'}(y^{'})w_{x''}(y^{''})w_{x'''}(y^{'''})$$

$$\times f_0(y|x)f_0(y'|x')f_0(y''|x'')f_0(y'''|x''')dy\,dy^{'}\,dy^{''}\,dy^{'''}$$

$$\times f^{-2}(x)f^{-2}(x^{'})f^{-2}(x^{''})f^{-2}(x^{'''})dxdx'dx''dx'''$$

$$\times f(x_1)f(x_2)dx_1dx_2.$$

De plus, $\forall y,y',y'',y'''\in\mathbb{R}$ et $\forall x_1,x_2\in[-1,2]$,

$$\int_{\mathbb{R}}|I_{\{y_1\leq y\}}-F_0(y|x_1)||I_{\{y_1\leq y'\}}-F_0(y^{'}|x_1)||I_{\{y_1\leq y''\}}-F_0(y^{''}|x_1)|$$

$$\times|I_{\{y_1\leq y'''\}}-F_0(y^{'''}|x_1)|f_0(y_1|x_1)dy_1\leq M,$$

et

$$\int_{\mathbb{R}} |I_{\{y_2 \leq y\}} - F_0(y|x_2)||I_{\{y_2 \leq y'\}} - F_0(y'|x_2)||I_{\{y_2 \leq y''\}} - F_0(y''|x_2)|$$

$$\times |I_{\{y_2 \leq y'''\}} - F_0(y'''|x_2)| f_0(y_2|x_2) dy_2 \leq M.$$

Par $(f.0)$, $(P.0)$ et $(P.1)$ respectivement, la densité $f(.)$ et les différentes fonctions de poids sont bornées de sorte que,

$$\begin{aligned}
0 \leq E_{F_0}(k_{12n}^4) \quad &\leq \quad M n^{-4} h^{-6} \int_{-1}^{2} \int_{-1}^{2} \int_{[0,1]^4} |K^*\left(\frac{x_1 - x}{h}\right) K^*\left(\frac{x_2 - x}{h}\right) \\
&\times K^*\left(\frac{x_1 - x'}{h}\right) K^*\left(\frac{x_2 - x'}{h}\right) K^*\left(\frac{x_1 - x''}{h}\right) \\
&\times K^*\left(\frac{x_2 - x''}{h}\right) K^*\left(\frac{x_1 - x'''}{h}\right) K^*\left(\frac{x_2 - x'''}{h}\right)| \\
&\times dx\, dx'\, dx''\, dx''' dx_1 dx_2.
\end{aligned}$$

Pour alléger la notation des calculs qui suivent, on pose $K^+(.) = |K^*(.)|$ et $K_h^+(.) = \dfrac{1}{h} K^+\left(\dfrac{.}{h}\right)$. Sous l'hypothèse $(K.0)$, la fonction $K^+(.)$ est positive, symétrique et à support compact $[-1, 1]$. On obtient alors,

$$\begin{aligned}
0 \leq E_{F_0}(k_{12n}^4) \quad &\leq \quad M n^{-4} h^{-6} \int_{-1}^{2} \int_{-1}^{2} \int_{[0,1]^4} K^+\left(\frac{x_1 - x}{h}\right) K^+\left(\frac{x_2 - x}{h}\right) \\
&\times K^+\left(\frac{x_1 - x'}{h}\right) K^+\left(\frac{x_2 - x'}{h}\right) K^+\left(\frac{x_1 - x''}{h}\right) \\
&\times K^+\left(\frac{x_2 - x''}{h}\right) K^+\left(\frac{x_1 - x'''}{h}\right) K^+\left(\frac{x_2 - x'''}{h}\right) dx\, dx'\, dx''\, dx''' dx_1 dx_2, \\
&= \quad M n^{-4} h^{2} \int_{-1}^{2} \int_{-1}^{2} \int_{[0,1]^4} K_h^+(x_1 - x) K_h^+(x_2 - x) K_h^+(x_1 - x') K_h^+(x_2 - x') \\
&\times K_h^+(x_1 - x'') K_h^+(x_2 - x'') K_h^+(x_1 - x''') K_h^+(x_2 - x''')
\end{aligned}$$

$$
\begin{aligned}
& \times\, dx\, dx^{'}\, dx^{''}\, dx^{'''} dx_1 dx_2, \\
= \quad & M n^{-4} h^2 \int_{-1}^{2} \int_{-1}^{2} \left(\int_0^1 K_h^+(x_1 - x) K_h^+(x_2 - x) dx \right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x^{'}) K_h^+(x_2 - x^{'}) dx^{'} \right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x^{''}) K_h^+(x_2 - x^{''}) dx^{''} \right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x^{'''}) K_h^+(x_2 - x^{'''}) dx^{'''} \right) dx_1 dx_2. \qquad (C.1)
\end{aligned}
$$

De plus,

$$
\begin{aligned}
\int_0^1 K_h^+(x_1 - x) K_h^+(x_2 - x) dx &= \int_0^1 K_h^+(x_1 - x_2 + x_2 - x) K_h^+(x_2 - x) dx, \\
&= \int_{x_2 - 1}^{x_2} K_h^+(x_1 - x_2 + u) K_h^+(u) du, \\
&\leq \int_{\mathbb{R}} K_h^+(x_1 - x_2 + u) K_h^+(u) du, \\
&= K_h^{+^{(2)}}(x_2 - x_1), \\
&= \frac{1}{h} K^{+^{(2)}} \left(\frac{x_2 - x_1}{h} \right). \qquad (C.2)
\end{aligned}
$$

Les autres intégrales sont majorées de la même façon pour obtenir finalement :

$$
\begin{aligned}
0 \leq E_{F_0}(k_{12n}^4) &\leq M n^{-4} h^2 \int_{-1}^{2} \int_{-1}^{2} \left[K_h^{+^{(2)}}(x_2 - x_1) \right]^4 dx_1 dx_2, \\
&= M n^{-4} h^{-2} \int_{-1}^{2} \int_{-1}^{2} \left[K^{+^{(2)}} \left(\frac{x_2 - x_1}{h} \right) \right]^4 dx_1 dx_2, \\
&= M n^{-4} h^{-1} \int_{-1}^{2} \int_{\frac{-1 - x_1}{h}}^{\frac{2 - x_1}{h}} \left[K^{+^{(2)}}(t) \right]^4 dt dx_1, \\
&= o(n^{-2}), \qquad (C.3)
\end{aligned}
$$

sous les hypothèses $(K.0)$ et $(H.1)$. ♣

Annexe D

Preuve du Lemme 2.4.6 On va montrer que $E_{F_0}(k_{12n}^2 k_{13n}^2) = o(n^{-3})$. Par les définitions de $\tilde{W}^n(.,.,.)$ et de k_{ijn} données respectivement en (2.6) et (2.13) et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$
\begin{aligned}
0 \leq E_{F_0}(k_{12n}^2 k_{13n}^2) \quad = \quad & n^{-4}h^{-6} \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4}\int_{\mathrm{I\!R}^4} K^*\left(\frac{x_1 - x}{h}\right) K^*\left(\frac{x_2 - x}{h}\right) \\
& \times K^*\left(\frac{x_1 - x^{'}}{h}\right) K^*\left(\frac{x_2 - x^{'}}{h}\right) K^*\left(\frac{x_1 - x^{''}}{h}\right) \\
& \times K^*\left(\frac{x_3 - x^{''}}{h}\right) K^*\left(\frac{x_1 - x^{'''}}{h}\right) K^*\left(\frac{x_3 - x^{'''}}{h}\right) \\
& \times \left(\int_{\mathrm{I\!R}} (I_{\{y_1 \leq y\}} - F_0(y|x_1))(I_{\{y_1 \leq y^{'}\}} - F_0(y^{'}|x_1)) \right. \\
& \left. \times (I_{\{y_1 \leq y^{''}\}} - F_0(y^{''}|x_1))(I_{\{y_1 \leq y^{'''}\}} - F_0(y^{'''}|x_1)) f_0(y_1|x_1) dy_1 \right) \\
& \times \left(\int_{\mathrm{I\!R}} (I_{\{y_2 \leq y\}} - F_0(y|x_2))(I_{\{y_2 \leq y^{'}\}} - F_0(y^{'}|x_2)) f_0(y_2|x_2) dy_2 \right) \\
& \times \left(\int_{\mathrm{I\!R}} (I_{\{y_3 \leq y^{''}\}} - F_0(y^{''}|x_3))(I_{\{y_3 \leq y^{'''}\}} - F_0(y^{'''}|x_3)) f_0(y_3|x_3) dy_3 \right) \\
& \times w(x)w(x^{'})w(x^{''})w(x^{'''}) w_x(y) w_{x^{'}}(y^{'}) w_{x^{''}}(y^{''}) w_{x^{'''}}(y^{'''}) \\
& \times f_0(y|x) f_0(y'|x') f_0(y''|x'') f_0(y'''|x''') dy\, dy^{'}\, dy^{''}\, dy^{'''} f^{-2}(x) f^{-2}(x^{'}) \\
& \times f^{-2}(x^{''}) f^{-2}(x^{'''}) dx\, dx^{'}\, dx^{''}\, dx^{'''} f(x_1)dx_1 f(x_2)dx_2 f(x_3)dx_3,
\end{aligned}
$$

$$
\begin{aligned}
\leq\ & n^{-4}h^{-6}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4}\int_{\mathbb{R}^4}|K^*\left(\frac{x_1-x}{h}\right)K^*\left(\frac{x_2-x}{h}\right) \\
& \times K^*\left(\frac{x_1-x^{'}}{h}\right)K^*\left(\frac{x_2-x^{'}}{h}\right)K^*\left(\frac{x_1-x^{''}}{h}\right) \\
& \times K^*\left(\frac{x_3-x^{''}}{h}\right)K^*\left(\frac{x_1-x^{'''}}{h}\right)K^*\left(\frac{x_3-x^{'''}}{h}\right)| \\
& \times \Bigg(\int_{\mathbb{R}}|I_{\{y_1\leq y\}}-F_0(y|x_1)||I_{\{y_1\leq y'\}}-F_0(y^{'}|x_1)| \\
& \times |I_{\{y_1\leq y''\}}-F_0(y^{''}|x_1)||I_{\{y_1\leq y'''\}}-F_0(y^{'''}|x_1)|f_0(y_1|x_1)dy_1\Bigg) \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_2\leq y\}}-F_0(y|x_2)||I_{\{y_2\leq y'\}}-F_0(y^{'}|x_2)|f_0(y_2|x_2)dy_2\right) \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_3\leq y''\}}-F_0(y^{''}|x_3)||I_{\{y_3\leq y'''\}}-F_0(y^{'''}|x_3)|f_0(y_3|x_3)dy_3\right) \\
& \times w(x)w(x^{'})w(x^{''})w(x^{'''})w_x(y)w_{x'}(y^{'})w_{x''}(y^{''})w_{x'''}(y^{'''}) \\
& \times f_0(y|x)f_0(y'|x')f_0(y''|x'')f_0(y'''|x''')dy\,dy^{'}\,dy^{''}\,dy^{'''}f^{-2}(x)f^{-2}(x^{'}) \\
& \times f^{-2}(x^{''})f^{-2}(x^{'''})dx\,dx^{'}\,dx^{''}\,dx^{'''}f(x_1)dx_1f(x_2)dx_2f(x_3)dx_3.
\end{aligned}
$$

De plus, $\forall y, y', y'', y''' \in \mathbb{R}$ et $\forall x_1, x_2, x_3 \in [-1, 2]$,

$$
\begin{aligned}
& \int_{\mathbb{R}}|I_{\{y_1\leq y\}}-F_0(y|x_1)||I_{\{y_1\leq y'\}}-F_0(y^{'}|x_1)||I_{\{y_1\leq y''\}}-F_0(y^{''}|x_1)| \\
& \times |I_{\{y_1\leq y'''\}}-F_0(y^{'''}|x_1)|f_0(y_1|x_1)dy_1\leq M,
\end{aligned}
$$

$$
\int_{\mathbb{R}}|I_{\{y_2\leq y\}}-F_0(y|x_2)||I_{\{y_2\leq y'\}}-F_0(y^{'}|x_2)|f(y_2|x_2)dy_2\leq M,
$$

et

$$
\int_{\mathbb{R}}|I_{\{y_3\leq y''\}}-F_0(y''|x_3)||I_{\{y_3\leq y'''\}}-F_0(y^{'''}|x_3)|f(y_3|x_3)dy_3\leq M.
$$

Par $(f.0)$, $(P.0)$, $(P.1)$ respectivement, la fonction $f(.)$ et les fonctions de poids sont bornées. Ainsi,

$$\begin{aligned}
E_{F_0}(k_{12n}^2 k_{13n}^2) \leq\ & Mn^{-4}h^{-6} \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4} \left| K^*\left(\frac{x_1 - x}{h}\right) K^*\left(\frac{x_2 - x}{h}\right) K^*\left(\frac{x_1 - x'}{h}\right)\right. \\
& \times K^*\left(\frac{x_2 - x'}{h}\right) K^*\left(\frac{x_1 - x''}{h}\right) K^*\left(\frac{x_3 - x''}{h}\right) K^*\left(\frac{x_1 - x'''}{h}\right) \\
& \times \left. K^*\left(\frac{x_3 - x'''}{h}\right)\right| dx\, dx'\, dx''\, dx''' dx_1 dx_2 dx_3.
\end{aligned}$$

Comme dans l'Annexe C, on pose $K^+(.) = |K^*(.)|$. On obtient alors :

$$\begin{aligned}
E_{F_0}(k_{12n}^2 k_{13n}^2) \leq\ & Mn^{-4}h^{-6} \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4} K^+\left(\frac{x_1 - x}{h}\right) K^+\left(\frac{x_2 - x}{h}\right) K^+\left(\frac{x_1 - x'}{h}\right) \\
& \times K^+\left(\frac{x_2 - x'}{h}\right) K^+\left(\frac{x_1 - x''}{h}\right) K^+\left(\frac{x_3 - x''}{h}\right) K^+\left(\frac{x_1 - x'''}{h}\right) \\
& \times K^+\left(\frac{x_3 - x'''}{h}\right) dx\, dx'\, dx''\, dx''' dx_1 dx_2 dx_3, \\
=\ & Mn^{-4}h^{2} \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4} K_h^+(x_1 - x) K_h^+(x_2 - x) K_h^+(x_1 - x') \\
& \times K_h^+(x_2 - x') K_h^+(x_1 - x'') K_h^+(x_3 - x'') K_h^+(x_1 - x''') \\
& \times K_h^+(x_3 - x''') dx\, dx'\, dx''\, dx''' dx_1 dx_2 dx_3, \\
=\ & Mn^{-4}h^{2} \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2} \left(\int_0^1 K_h^+(x_1 - x) K_h^+(x_2 - x) dx\right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x') K_h^+(x_2 - x') dx'\right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x'') K_h^+(x_3 - x'') dx''\right) \\
& \times \left(\int_0^1 K_h^+(x_1 - x''') K_h^+(x_3 - x''') dx'''\right) dx_1 dx_2 dx_3. \qquad \text{(D.1)}
\end{aligned}$$

De plus,

$$\begin{aligned}
\int_0^1 K_h^+(x_1-x)K_h^+(x_2-x)dx &= \int_0^1 K_h^+(x_1-x_2+x_2-x)K_h^+(x_2-x)dx, \\
&= \int_{x_2-1}^{x_2} K_h^+(x_1-x_2+u)K_h^+(u)du, \\
&\leq \int_{\mathbb{R}} K_h^+(x_1-x_2+u)K_h^+(u)du, \\
&= K_h^{+^{(2)}}(x_2-x_1), \\
&= \frac{1}{h}K^{+^{(2)}}\left(\frac{x_2-x_1}{h}\right) \qquad \text{(D.2)}
\end{aligned}$$

Les autres intégrales sont majorées de la même façon pour obtenir :

$$\begin{aligned}
E_{F_0}(k_{12n}^2k_{13n}^2) &\leq Mn^{-4}h^2\int_{-1}^2\int_{-1}^2\int_{-1}^2\left[K_h^{+^{(2)}}(x_2-x_1)\right]^2\left[K_h^{+^{(2)}}(x_3-x_1)\right]^2dx_1dx_2dx_3, \\
&= Mn^{-4}h^2\int_{-1}^2\left(\int_{-1}^2\left[K_h^{+^{(2)}}(x_2-x_1)\right]^2dx_2\right) \\
&\quad \times\left(\int_{-1}^2\left[K_h^{+^{(2)}}(x_3-x_1)\right]^2dx_3\right)dx_1. \qquad \text{(D.3)}
\end{aligned}$$

De plus, comme $K_h^{+^{(2)}}(.)$ est symétrique,

$$\begin{aligned}
\int_{-1}^2\left[K_h^{+^{(2)}}(x_2-x_1)\right]^2dx_2 &= \int_{-1-x_1}^{2-x_1}\left[K_h^{+^{(2)}}(u)\right]^2du, \\
&\leq \int_{\mathbb{R}}\left[K_h^{+^{(2)}}(u)\right]^2du, \\
&= K_h^{+^{(4)}}(0), \\
&= \frac{1}{h}K^{+^{(4)}}\left(\frac{0}{h}\right) \qquad \text{(D.4)}
\end{aligned}$$

L'autre intégrale est majorée de la même manière de sorte que :

$$0 \leq E_{F_0}(k_{12n}^2k_{13n}^2) \leq Mn^{-4}h^2\left[K_h^{+^{(4)}}(0)\right]^2,$$

$$\begin{aligned} &\leq Mn^{-4}\left[K^{+^{(4)}}(0)\right]^2, \\ &\leq Mn^{-4}, \\ &= o\left(n^{-3}\right), \end{aligned} \tag{D.5}$$

sous l'hypothèse $(K.0)$. ♣

Annexe E

Preuve du Lemme 2.4.7 Par les définitions de $\tilde{W}^n(.,.,.)$ et de k_{ijn}, données respectivement en (2.6) et en (2.13), on obtient, de la même façon qu'à l'Annexe D et sous les hypothèses $(F_0.y.1)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$,

$$
\begin{aligned}
& |E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n})| \\
\leq \quad & n^{-4}h^{-6}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4}\int_{\mathbb{R}^4} |K^*\left(\frac{x_1-x}{h}\right)K^*\left(\frac{x_2-x}{h}\right)K^*\left(\frac{x_1-x^{'}}{h}\right) \\
& \times K^*\left(\frac{x_3-x^{'}}{h}\right)K^*\left(\frac{x_4-x^{''}}{h}\right)K^*\left(\frac{x_2-x^{''}}{h}\right)K^*\left(\frac{x_4-x^{'''}}{h}\right)K^*\left(\frac{x_3-x^{'''}}{h}\right)| \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_1\leq y\}}-F_0(y|x_1)|\times|I_{\{y_1\leq y'\}}-F_0(y^{'}|x_1)|f_0(y_1|x_1)dy_1\right) \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_2\leq y\}}-F_0(y|x_2)|\times|I_{\{y_2\leq y''\}}-F_0(y^{''}|x_2)|f_0(y_2|x_2)dy_2\right) \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_3\leq y'\}}-F_0(y^{'}|x_3)|\times|I_{\{y_3\leq y'''\}}-F_0(y^{'''}|x_3)|f_0(y_3|x_3)dy_3\right) \\
& \times \left(\int_{\mathbb{R}}|I_{\{y_4\leq y''\}}-F_0(y^{''}|x_4)|\times|I_{\{y_4\leq y'''\}}-F_0(y^{'''}|x_4)|f_0(y_4|x_4)dy_4\right) \\
& \times w_x(y)w_{x'}(y^{'})w_{x''}(y^{''})w_{x'''}(y^{'''})f_0(y|x)f_0(y'|x')f_0(y''|x'')f_0(y'''|x''')dy\,dy^{'}\,dy^{''}\,dy^{'''} \\
& \times f^{-2}(x)f^{-2}(x^{'})f^{-2}(x^{''})f^{-2}(x^{'''})w(x)w(x^{'})w(x^{''})w(x^{'''})dx\,dx^{'}\,dx^{''}\,dx^{'''} \\
& \times f(x_1)f(x_2)f(x_3)f(x_4)dx_1dx_2dx_3dx_4,
\end{aligned}
$$

$$\leq \quad Mn^{-4}h^{-6}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4} K^{+}\left(\frac{x_1-x}{h}\right)K^{+}\left(\frac{x_2-x}{h}\right)K^{+}\left(\frac{x_1-x^{'}}{h}\right)$$

$$\times\, K^{+}\left(\frac{x_3-x^{'}}{h}\right)K^{+}\left(\frac{x_4-x^{''}}{h}\right)K^{+}\left(\frac{x_2-x^{''}}{h}\right)K^{+}\left(\frac{x_4-x^{'''}}{h}\right)K^{+}\left(\frac{x_3-x^{'''}}{h}\right)$$

$$\times\, dx\, dx^{'}\, dx^{''}\, dx^{'''} dx_1 dx_2 dx_3 dx_4, \tag{E.1}$$

où on a posé $K^{+}(.) = |K^{*}(.)|$, de sorte que,

$$|E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n})|$$

$$\leq \quad Mn^{-4}h^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{[0,1]^4} K_h^{+}(x_1-x)K_h^{+}(x_2-x)K_h^{+}(x_1-x^{'})$$

$$\times\, K_h^{+}(x_3-x^{'})K_h^{+}(x_4-x^{''})K_h^{+}(x_2-x^{''})K_h^{+}(x_4-x^{'''})K_h(x_3-x^{'''})$$

$$\times\, dx\, dx^{'}\, dx^{''}\, dx^{'''} dx_1 dx_2 dx_3 dx_4,$$

$$= \quad Mn^{-4}h^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\left(\int_0^1 K_h^{+}(x_1-x)K_h^{+}(x_2-x)dx\right)$$

$$\times\left(\int_0^1 K_h^{+}(x_1-x^{'})K_h^{+}(x_3-x^{'})dx'\right)\left(\int_0^1 K_h^{+}(x_4-x^{''})K_h^{+}(x_2-x^{''})dx''\right)$$

$$\times\left(\int_0^1 K_h^{+}(x_4-x^{'''})K_h(x_3-x^{'''})dx'''\right)dx_1dx_2dx_3dx_4. \tag{E.2}$$

De plus,

$$\begin{aligned}\int_0^1 K_h^{+}(x_1-x)K_h^{+}(x_2-x)dx &= \int_0^1 K_h^{+}(x_1-x_2+x_2-x)K_h^{+}(x_2-x)dx,\\ &= \int_{x_2-1}^{x_2} K_h^{+}(x_1-x_2+u)K_h^{+}(u)du,\\ &\leq \int_{\mathbb{R}} K_h^{+}(x_1-x_2+u)K_h^{+}(u)du,\\ &= K_h^{+^{(2)}}(x_2-x_1),\\ &= \frac{1}{h}K^{+^{(2)}}\left(\frac{x_2-x_1}{h}\right).\end{aligned} \tag{E.3}$$

Les autres intégrales sont majorées de la même manière pour obtenir :

$$\begin{aligned}
|E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n})| &\leq Mn^{-4}h^2 \int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2}\int_{-1}^{2} K_h^{+^{(2)}}(x_2-x_1)K_h^{+^{(2)}}(x_3-x_1) \\
&\quad \times K_h^{+^{(2)}}(x_2-x_4)K_h^{+^{(2)}}(x_3-x_4)dx_1dx_2dx_3dx_4, \\
&= Mn^{-4}h^2 \int_{-1}^{2}\int_{-1}^{2} \left(\int_{-1}^{2} K_h^{+^{(2)}}(x_2-x_1)K_h^{+^{(2)}}(x_3-x_1)dx_1\right) \\
&\quad \times \left(\int_{-1}^{2} K_h^{+^{(2)}}(x_2-x_4)K_h^{+^{(2)}}(x_3-x_4)dx_4\right) dx_2dx_3. \qquad \text{(E.4)}
\end{aligned}$$

De la même manière,

$$\begin{aligned}
\int_{-1}^{2} K_h^{+^{(2)}}(x_2-x_1)K_h^{+^{(2)}}(x_3-x_1)dx_1 &= \int_{-1}^{2} K_h^{+^{(2)}}(x_2-x_3+x_3-x_1)K_h^{+^{(2)}}(x_3-x_1)dx_1, \\
&= \int_{x_3-2}^{x_3+1} K_h^{+^{(2)}}(x_2-x_3+u)K_h^{+^{(2)}}(u)du, \\
&\leq \int_{\mathbb{R}} K_h^{+^{(2)}}(x_2-x_3+u)K_h^{+^{(2)}}(u)du, \\
&= K_h^{+^{(4)}}(x_3-x_2), \\
&= \frac{1}{h}K^{+^{(4)}}\left(\frac{x_3-x_2}{h}\right). \qquad \text{(E.5)}
\end{aligned}$$

L'autre intégrale de est majorée de la même façon pour obtenir, comme $K_h^{+^{(4)}}(.)$ est symétrique,

$$\begin{aligned}
|E_{F_0}(k_{12n}k_{13n}k_{42n}k_{43n})| &\leq Mn^{-4}h^2 \int_{-1}^{2}\int_{-1}^{2} \left[K_h^{+^{(4)}}(x_3-x_2)\right]^2 dx_3dx_2, \\
&= Mn^{-4}h^2 \int_{-1}^{2}\int_{-1-x_2}^{2-x_2} \left[K_h^{+^{(4)}}(t)\right]^2 dtdx_2, \\
&\leq Mn^{-4}h^2 \int_{-1}^{2}\int_{\mathbb{R}} \left[K_h^{+^{(4)}}(t)\right]^2 dtdx_2, \\
&\leq Mn^{-4}h^2 K_h^{+^{(8)}}(0),
\end{aligned}$$

$$= \quad Mn^{-4}hK^{+^{(8)}}(0),$$

$$= \quad o(n^{-4}), \tag{E.6}$$

sous les hypothèses $(K.0)$ et $(H.0)$.♣

Annexe F

Preuve du Lemme 4.5.7 Par le même cheminement qu'en (3.15) et sous les nouvelles hypothèses introduites à la Section 4.2, on a le développement suivant :

$$
\begin{aligned}
\hat{F}_n(y|x) - F_0(y|x;\theta_0) &= \sum_{i=1}^{n} W^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i;\theta_0)) \\
&\quad + h^{p+1} R_n^*(x, y; \theta_0),
\end{aligned}
$$

où $R_n^*(x,y;\theta_0)$, à l'instar de $R_n(x,y)$ déjà défini en (3.14), est presque sûrement borné en $x \in [0,1]$, $y \in \mathbb{R}$ et $\theta_0 \in \Theta$ si n est suffisamment grand. En utilisant maintenant la décomposition du noyau $W^n(.)$, donnée en (1.34), on obtient :

$$
\begin{aligned}
\hat{F}_n(y|x) - F_0(y|x;\theta_0) &= \sum_{i=1}^{n} \tilde{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i;\theta_0)) \\
&\quad + h \sum_{i=1}^{n} Z_n\left(\frac{X_i - x}{h}, x\right) \bar{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i;\theta_0)) \\
&\quad + h^{p+1} R_n^*(x, y; \theta_0),
\end{aligned}
$$

où les noyaux $\tilde{W}^n(.,.,.)$ et $\bar{W}^n(.,.,.)$ sont définis respectivement en (2.6) et (2.71). De plus, dans cette expression, $Z_n\left(\frac{X_i - x}{h}, x\right)$, déjà défini en (1.33), est presque sûrement borné en $x \in [0,1]$ si n est suffisamment grand. Si on intègre cette quantité en x et en y, on obtient :

$$
\begin{aligned}
&\sqrt{nh} \int_0^1 \int_{\mathbb{R}} (\hat{F}_n(y|x) - F_0(y|x;\theta_0)) \dot{F}_0(y|x;\theta_0) w(x) w_x(y) f_0(y|x;\theta_0) dy dx \\
&= U + hV \\
&\quad + \sqrt{nh^{2p+3}} \int_0^1 \int_{\mathbb{R}} R_n^*(x, y; \theta_0) \dot{F}_0(y|x;\theta_0) w(x) w_x(y) f_0(y|x;\theta_0) dy dx, \qquad \text{(F.1)}
\end{aligned}
$$

où on a posé

$$U = \sqrt{nh}\sum_{i=1}^{n} u_i, \tag{F.2}$$

$$V = \sqrt{nh}\sum_{i=1}^{n} v_i, \tag{F.3}$$

avec

$$\begin{aligned} u_i &= \int_0^1 \int_{\mathbb{R}} \tilde{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i;\theta_0))\dot{F}_0(y|x;\theta_0) \\ &\quad \times w(x)w_x(y)f_0(y|x;\theta_0)dydx, \end{aligned} \tag{F.4}$$

et

$$\begin{aligned} v_i &= \int_0^1 \int_{\mathbb{R}} Z_n\left(\frac{X_i - x}{h}, x\right) \bar{W}^n(X_i, x, h)(I_{\{Y_i \le y\}} - F_0(y|X_i;\theta_0))\dot{F}_0(y|x;\theta_0) \\ &\quad \times w(x)w_x(y)f_0(y|x;\theta_0)dydx, \end{aligned} \tag{F.5}$$

De plus, sous les hypothèses $(F_0.\theta.2)$, $(P.0)$ et $(P.1)$,

$$\begin{aligned} &\left|\int_0^1 \int_{\mathbb{R}} R_n^*(x,y;\theta_0)\dot{F}_0(y|x;\theta_0)w(x)w_x(y)f_0(y|x;\theta_0)dydx\right| \\ \le\ &\int_0^1 \int_{\mathbb{R}} |R_n^*(x,y;\theta_0)|\,|\dot{F}_0(y|x;\theta_0)|w(x)w_x(y)f_0(y|x;\theta_0)dydx, \\ \le\ &M, \end{aligned} \tag{F.6}$$

de sorte que

$$\begin{aligned} &\sqrt{nh}\int_0^1 \int_{\mathbb{R}} (\hat{F}_n(y|x) - F_0(y|x;\theta_0))\dot{F}_0(y|x;\theta_0)w(x)w_x(y)f_0(y|x;\theta_0)dydx \\ =\ &U + hV + O_p\left(\sqrt{nh^{2p+3}}\right). \end{aligned} \tag{F.7}$$

On étudie à présent le comportement asymptotique de U. Sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$, on a

$$E_{F_0}(u_1) = \int_{-1}^{2}\int_{\mathbb{R}}\int_0^1 \tilde{W}^n(x_1, x, h)\left(\int_{\mathbb{R}} (I_{\{y_1 \le y\}} - F_0(y|x_1;\theta_0))f_0(y_1|x_1;\theta_0)dy_1\right)$$

$$\times \dot{F}_0(y|x;\theta_0)w(x)w_x(y)f_0(y|x;\theta_0)dydxf(x_1)dx_1,$$

$$= \quad 0,$$

car $\int_{\mathbb{R}}(I_{\{y_1\leq y\}} - F_0(y|x_1;\theta_0))f_0(y_1|x_1;\theta_0)dy_1 = 0, \forall y \in \mathbb{R}$ et $\forall x_1 \in [-1,2]$. On en déduit que $E_{F_0}(U) = 0$. Maintenant, comme les u_i sont *iid* et d'espérance nulle,

$$Var_{F_0}(U) = n^2hE_{F_0}(u_1^2). \tag{F.8}$$

Or,

$$\begin{aligned}
u_1^2 &= \frac{1}{n^2h^2}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} K^*\left(\frac{X_1-x}{h}\right)K^*\left(\frac{X_1-x'}{h}\right)f^{-1}(x)f^{-1}(x') \\
&\quad \times (I_{\{Y_1\leq y\}} - F_0(y|X_1;\theta_0)) \times (I_{\{Y_1\leq y'\}} - F_0(y'|X_1;\theta_0)) \\
&\quad \times \dot{F}_0(y|x;\theta_0)w(x)w_x(y)\dot{F}_0(y'|x';\theta_0)w(x')w_{x'}(y') \\
&\quad \times f_0(y|x;\theta_0)f_0(y'|x';\theta_0)dydxdy'dx',
\end{aligned}$$

en utilisant la définition de $\tilde{W}^n(.,.,.)$ donnée en (2.6). On en déduit, par le théorème de Fubini et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned}
E_{F_0}(u_1^2) &\leq \frac{1}{n^2h^2}\int_{-1}^2\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1 \left|K^*\left(\frac{x_1-x}{h}\right)\right|\left|K^*\left(\frac{x_1-x'}{h}\right)\right|f^{-1}(x)f^{-1}(x') \\
&\quad \times \left(\int_{\mathbb{R}}\left|I_{\{y_1\leq y\}} - F_0(y|x_1;\theta_0)\right|\left|I_{\{y_1\leq y'\}} - F_0(y'|x_1;\theta_0)\right|f_0(y_1|x_1;\theta_0)dy_1\right) \\
&\quad \times \left|\dot{F}_0(y|x;\theta_0)\right|w(x)w_x(y)\left|\dot{F}_0(y'|x';\theta_0)\right|w(x')w_{x'}(y') \\
&\quad \times f_0(y|x;\theta_0)f_0(y'|x';\theta_0)dydxdy'dx'f(x_1)dy_1dx_1.
\end{aligned}$$

De plus, $\forall y, y' \in \mathbb{R}$ et $\forall x_1 \in [-1,2]$,

$$\int_{\mathbb{R}}\left|I_{\{y_1\leq y\}} - F_0(y|x_1;\theta_0)\right|\left|I_{\{y_1\leq y'\}} - F_0(y'|x_1;\theta_0)\right|f_0(y_1|x_1;\theta_0)dy_1 \leq M. \tag{F.9}$$

Alors, sous les hypothèses $(F_0.y.1)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$ et par le théorème de Fubini,

$$E_{F_0}(u_1^2) \leq \frac{M}{n^2h^2}\int_{-2}^1\int_0^1\int_0^1\left|K^*\left(\frac{x_1-x}{h}\right)\right|\left|K^*\left(\frac{x_1-x'}{h}\right)\right|$$

$$\times \left(\int_{\mathbb{R}} \left|\dot{F}_0(y|x;\theta_0)\right| f_0(y|x;\theta_0)dy\right)$$

$$\times \left(\int_{\mathbb{R}} \left|\dot{F}_0(y'|x';\theta_0)\right| f_0(y'|x';\theta_0)dy'\right) dxdx'f(x_1)dx_1.$$

De plus, sous les hypothèses $(F_0.y.1)$ et $(F_0.\theta.2)$ et pour tout $x \in [0,1]$,

$$\begin{aligned}\int_{\mathbb{R}} \left|\dot{F}_0(y|x;\theta_0)\right| f_0(y|x;\theta_0)dy &\leq M\int_{\mathbb{R}} f_0(y|x;\theta_0)dy,\\ &= M.\end{aligned}$$

On en déduit que :

$$E_{F_0}(u_1^2) \leq \frac{M}{n^2h^2}\int_{-1}^{2}\int_0^1\int_0^1 \left|K^*\left(\frac{x_1-x}{h}\right)\right|\left|K^*\left(\frac{x_1-x'}{h}\right)\right| dxdx'f(x_1)dx_1.$$

On effectue à présent les changements de variables $x = -uh + x_1$ et $x' = -vh + x_1$ de Jacobien $\frac{1}{h^2}$ et on obtient, sous l'hypothèse $(K.0)$, comme $h < 1$ et $x_1 \in [-1,2]$,

$$\begin{aligned}E_{F_0}(u_1^2) &\leq \frac{M}{n^2}\int_{-1}^{2}\int_{\frac{x_1-1}{h}}^{\frac{x_1}{h}}\int_{\frac{x_1-1}{h}}^{\frac{x_1}{h}} |K^*(u)|\,|K^*(v)|\,dudvf(x_1)dx_1,\\ &\leq \frac{M}{n^2}\int_{-1}^{2}\int_{\mathbb{R}}\int_{\mathbb{R}} |K^*(u)|\,|K^*(v)|\,dudvf(x_1)dx_1,\\ &= \frac{M}{n^2}.\end{aligned}$$

Ainsi,

$$Var_{F_0}(U) \leq n^2h \times \frac{M}{n^2} = Mh = o(1), \tag{F.10}$$

sous l'hypothèse $(H.0)$. L'inégalité de Bienaymé-Tchebytchev permet de déduire que

$$U = o_p(1). \tag{F.11}$$

Etudions à présent le comportement asymptotique de la statistique V. Sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$ et par le théorème de Fubini, on a

$$\begin{aligned}E_{F_0}(v_1|X_1,\ldots,X_n) &= \int_{\mathbb{R}}\int_0^1 Z_n\left(\frac{X_1-x}{h},x\right)\bar{W}^n(X_1,x,h)\\ &\quad\times\left(\int_{\mathbb{R}}(I_{\{y_1\leq y\}} - F_0(y|X_1;\theta_0))f_0(y_1|X_1;\theta_0)dy_1\right)\end{aligned}$$

$$\times \dot{F}_0(y|x;\theta_0)w(x)w_x(y)f_0(y|x;\theta_0)dydx,$$

$$= \quad 0,$$

car $\int_{\mathbb{R}}(I_{\{y_1\leq y\}} - F_0(y|X_1;\theta_0))f_0(y_1|X_1;\theta_0)dy_1 = 0, \forall y \in \mathbb{R}$ et $\forall X_1 \in [-1,2]$. On en déduit que $E_{F_0}(v_1) = 0$ et donc $E_{F_0}(V) = 0$. Maintenant, comme $E_{F_0}(v_1|X_1,\ldots,X_n) = 0$,

$$\begin{aligned}
Var_{F_0}(V) &= nhVar_{F_0}\left(\sum_{i=1}^{n} v_i\right),\\
&= nhE_F\left(Var_{F_0}\left(\sum_{i=1}^{n} v_i|X_1,X_2,\ldots,X_n\right)\right),\\
&= n^2hE_F\left(E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right)\right). \qquad \text{(F.12)}
\end{aligned}$$

Or,

$$\begin{aligned}
v_1^2 \;=\; & \frac{1}{n^2h^2}\int_0^1\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}} Z_n\left(\frac{X_1-x}{h},x\right)Z_n\left(\frac{X_1-x'}{h},x'\right)\bar{K}\left(\frac{X_1-x}{h}\right)\bar{K}\left(\frac{X_1-x'}{h}\right)\\
& \times (I_{\{Y_1\leq y\}} - F_0(y|X_1;\theta_0)) \times (I_{\{Y_1\leq y'\}} - F_0(y'|X_1;\theta_0))\\
& \times \dot{F}_0(y|x;\theta_0)w(x)w_x(y)\dot{F}_0(y'|x';\theta_0)w(x')w_{x'}(y')\\
& \times f^{-1}(x)f^{-1}(x')f_0(y|x;\theta_0)f_0(y'|x';\theta_0)dydxdy'dx',
\end{aligned}$$

en utilisant la définition de $\bar{W}^n(.,.,.)$ donnée en (2.71). On en déduit, par le théorème de Fubini et sous les hypothèses $(F_0.y.1)$, $(X.0)$ et $(f.0)$,

$$\begin{aligned}
E_{F_0}\left(v_1^2|X_1,\ldots,X_n\right) \;\leq\; & \frac{1}{n^2h^2}\int_{\mathbb{R}}\int_0^1\int_{\mathbb{R}}\int_0^1 |Z_n\left(\frac{X_1-x}{h},x\right)||Z_n\left(\frac{X_1-x'}{h},x'\right)|\\
& \times \left|\bar{K}\left(\frac{X_1-x}{h}\right)\right|\left|\bar{K}\left(\frac{X_1-x'}{h}\right)\right|\\
& \times \left(\int_{\mathbb{R}}\left|I_{\{y_1\leq y\}} - F_0(y|X_1;\theta_0)\right|\left|I_{\{y_1\leq y'\}} - F_0(y'|X_1;\theta_0)\right| f_0(y_1|X_1;\theta_0)dy_1\right)\\
& \times f^{-1}(x)f^{-1}(x')\left|\dot{F}_0(y|x;\theta_0)\right|\left|\dot{F}_0(y'|x';\theta_0)\right|\\
& \times w(x)w_x(y)w(x')w_{x'}(y')f_0(y|x;\theta_0)f_0(y'|x';\theta_0)dydxdy'.
\end{aligned}$$

De plus, $\forall y, y' \in \mathbb{R}$ et $\forall X_1 \in [-1, 2]$,

$$\int_{\mathbb{R}} \left| I_{\{y_1 \leq y\}} - F_0(y|X_1;\theta_0) \right| \left| I_{\{y_1 \leq y'\}} - F_0(y'|X_1;\theta_0) \right| f_0(y_1|X_1;\theta_0) dy_1 \leq M. \tag{F.13}$$

Alors, sous les hypothèses $(F_0.y.1)$, $(X.0)$, $(f.0)$, $(P.0)$ et $(P.1)$ et par le théorème de Fubini,

$$\begin{aligned} E_{F_0}(v_1^2|X_1,\ldots,X_n) &\leq \frac{M}{n^2h^2} \int_0^1 \int_0^1 \left| \bar{K}\left(\frac{X_1 - x}{h}\right) \right| \left| \bar{K}\left(\frac{X_1 - x'}{h}\right) \right| \\ &\times \left(\int_{\mathbb{R}} \left| \dot{F}_0(y|x;\theta_0) \right| f_0(y|x;\theta_0) dy \right) \\ &\times \left(\int_{\mathbb{R}} \left| \dot{F}_0(y'|x';\theta_0) \right| f_0(y'|x';\theta_0) dy' \right) dx dx'. \end{aligned}$$

De plus, sous les hypothèses $(F_0.y.1)$ et $(F_0.\theta.2)$ et pour tout $x \in [0, 1]$,

$$\begin{aligned} \int_{\mathbb{R}} \left| \dot{F}_0(y|x;\theta_0) \right| f_0(y|x;\theta_0) dy &\leq M \int_{\mathbb{R}} f_0(y|x;\theta_0) dy, \\ &= M. \end{aligned}$$

On en déduit que :

$$E_{F_0}(v_1^2|X_1,\ldots,X_n) \leq \frac{M}{n^2h^2} \int_0^1 \int_0^1 \left| \bar{K}\left(\frac{X_1 - x}{h}\right) \right| \left| \bar{K}\left(\frac{X_1 - x'}{h}\right) \right| dx dx'.$$

On effectue à présent les changements de variables $x = -uh + X_1$ et $x' = -vh + X_1$ de Jacobien $\frac{1}{h^2}$ et on obtient, sous l'hypothèse $(K.0)$, comme $h < 1$ et $X_1 \in [-1, 2]$,

$$\begin{aligned} 0 \leq E_{F_0}(v_1^2|X_1,\ldots,X_n) &\leq \frac{M}{n^2} \int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}} \int_{\frac{X_1-1}{h}}^{\frac{X_1}{h}} \left| \bar{K}(u) \right| \left| \bar{K}(v) \right| du dv, \\ &\leq \frac{M}{n^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \bar{K}(u) \right| \left| \bar{K}(v) \right| du dv, \\ &= \frac{M}{n^2}. \end{aligned}$$

Ainsi,

$$Var_{F_0}(V) \leq n^2 h \times \frac{M}{n^2} = Mh = o(1), \tag{F.14}$$

sous l'hypothèse $(H.0)$. L'inégalité de Bienaymé-Tchebytchev permet de déduire que

$$V = o_p(1). \tag{F.15}$$

En substituant les résultats (F.11) et (F.15) dans (F.7), on en conclut que sous l'hypothèse $(H.3)$:

$$\begin{aligned}
&\sqrt{nh}\int_0^1\int_{\mathrm{IR}}(\hat{F}_n(y|x) - F_0(y|x;\theta_0))\dot{F}_0(y|x;\theta_0)w(x)w_x(y)f_0(y|x;\theta_0)dydx \\
&= \; o_p(1) + ho_p(1) + O_p\left(\sqrt{nh^{2p+3}}\right), \\
&= \; o_p(1).\clubsuit \qquad \text{(F.16)}
\end{aligned}$$

Annexe G

Preuve du Lemme 6.2.3 On rappelle que :

$$S^*_{n,k}(x,y) = \frac{1}{nh}\sum_{i=1}^{n} u_i, \tag{G.1}$$

avec

$$u_i = K\left(\frac{X_i - x}{h}\right)^2 \left(\frac{X_i - x}{h}\right)^k \left[F_0(y|X_i)(1 - F_0(y|X_i))\right], \tag{G.2}$$

où les X_i sont $iid \sim F(.)$. Comme les u_i sont iid, on a :

$$\begin{aligned} E_F(S^*_{n,k}(x,y)) &= \frac{1}{h}E_F(u_1), \\ &= \frac{1}{h}\int_{-1}^{2} K\left(\frac{x_1 - x}{h}\right)^2 \left(\frac{x_1 - x}{h}\right)^k \\ &\quad \times F_0(y|x_1)(1 - F_0(y|x_1))f(x_1)dx_1, \end{aligned} \tag{G.3}$$

de sorte que, si on effectue le changement de variable $x_1 = -uh + x$, on obtient, comme $h < 1$, $x \in [0,1]$ et $K(.)$ a le support $[-1,1]$,

$$\begin{aligned} E_F(S^*_{n,k}(x,y)) &= \int_{\frac{-1-x}{h}}^{\frac{2-x}{h}} K(u)^2 u^k F_0(y|uh + x)(1 - F_0(y|uh + x))f(uh + x)du, \\ &= \int_{-1}^{1} K(u)^2 u^k g_y(uh + x)du, \end{aligned} \tag{G.4}$$

où $g_y(.)$ est définie à la Section 2.3. De plus, on a déjà montré dans cette même section qu'il existe $\alpha_1 \in (0, uh)$ tel que :

$$g_y(uh + x) = g_y(x) + uhg'_y(\alpha_1 + x),$$

où $g'_y(\alpha_1 + x)$ est telle que $|g'_y(\alpha_1 + x)| \leq M$, uniformément en u, x et y. Cette dernière remarque permet de déduire que, sous l'hypothèse $(H.0)$ et pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

$$\begin{aligned} E_F(S^*_{n,k}(x,y)) &= g_y(x)\int_{-1}^{1} K(u)^2u^k du + h\int_{-1}^{1} K(u)^2u^k g'_y(\alpha_1+x)du, \\ &= g_y(x)\nu_k + hO(1), \\ &= F_0(y|x)(1-F_0(y|x))f(x)\nu_k + o(1), \end{aligned} \tag{G.5}$$

où ν_k est défini dans l'énoncé du Lemme 6.2.3. D'autre part, comme les u_i sont *iid*, on a aussi :

$$\begin{aligned} Var_F(S^*_{n,k}(x,y)) &= \frac{1}{nh^2}Var_F(u_1), \\ &= \frac{1}{nh^2}E_F(u_1^2) - \frac{1}{nh^2}E_F^2(u_1), \\ &= \frac{1}{nh^2}E_F(u_1^2) + o(1). \end{aligned} \tag{G.6}$$

De plus, pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

$$\begin{aligned} 0 \leq E_F(u_1^2) &= \int_{-1}^{2} K\left(\frac{x_1-x}{h}\right)^4 \left(\frac{x_1-x}{h}\right)^{2k} F_0(y|x_1)^2(1-F_0(y|x_1))^2 f(x_1)dx_1, \\ &\leq M\int_{-1}^{2} K\left(\frac{x_1-x}{h}\right)^4 \left(\frac{x_1-x}{h}\right)^{2k} dx_1, \end{aligned} \tag{G.7}$$

de sorte que, si on pose le changement de variable $x_1 = -uh + x$, on obtient, comme $h < 1$, $x \in [0,1]$ et $K(.)$ a le support $[-1,1]$,

$$\begin{aligned} E_F(u_1^2) &\leq Mh\int_{\frac{-1-x}{h}}^{\frac{2-x}{h}} K(u)^4u^{2k}du, \\ &\leq Mh\int_{\mathbb{R}} K^4(u)u^{2k}du, \\ &\leq Mh, \end{aligned} \tag{G.8}$$

sous l'hypothèse $(K.0)$. Si on substitue ce résultat dans (G.6), on en déduit que pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

$$Var_F(S^*_{n,k}(x,y)) = o(1). \tag{G.9}$$

Par l'inégalité de Bienaymé-Tchebytchev, le résultat suit. ♣

Résumé des différentes hypothèses utilisées dans la thèse

Hypothèses relatives à tous les Chapitres

- $(K.0)$ Le noyau $K(.)$ est une densité de probabilité, symétrique autour de 0, bornée, dérivable et de dérivée bornée. On suppose de plus qu'il est à support compact. Sans perte de généralité, on pose ce support compact comme étant $[-1,1]$. On suppose de plus que la matrice S, définie à la Section 1.3.1 est inversible.

- $(X.0)$ La variable aléatoire X a un support compact $[c,d]$. Sans perte de généralité, on pose ce support compact comme étant $[-1,2]$.

- $(f.0)$ La densité marginale $f(.)$ de X existe et est bornée pour tout $x \in [-1,2]$. Il existe donc des constantes m et M telles que $0 < m < f(x) < M$ pour tout $x \in [-1,2]$. Elle est aussi continûment dérivable dans un voisinage de tout $x \in]-1,2[$ et de dérivée bornée.

- $(H.0)$ On suppose que $h = h(n) \longrightarrow 0$ lorsque $n \longrightarrow +\infty$ et que $\forall n, h < 1$.

- $(H.1)$ On suppose $(H.0)$ et on suppose de plus que $nh \longrightarrow +\infty$, lorsque $n \longrightarrow +\infty$.

- $(H.2)$ On suppose $(H.0)$ et $(H.1)$ et on suppose de plus que $\dfrac{\sqrt{\log n}}{\sqrt{n}h^2} \longrightarrow 0$, lorsque $n \longrightarrow +\infty$.

- $(H.3)$ On suppose $(H.0)$, $(H.1)$ et $(H.2)$ et on suppose de plus que $nh^{2p+\frac{5}{2}} \longrightarrow 0$, lorsque $n \longrightarrow +\infty$.

- $(P.0)$ La fonction qui à x associe $w(x)$ est positive et bornée dans un voisinage de tout

$x \in [0, 1]$.

– $(P.1)$ La fonction qui à (x, y) associe $w_x(y)$ est positive et bornée dans un voisinage de tout $x \in [0, 1]$ et pour tout $y \in \mathbb{R}$.

Hypothèses relatives aux Chapitres 2, 3, 5 et 6

– $(F_0.y.1)$ La fonction qui à y associe $F_0(y|x)$ est continûment dérivable en y pour tout $x \in [-1, 2]$ ce qui assure l'existence de la densité conditionnelle $f_0(y|x)$ de Y sachant $\{X = x\}$.

– $(F_0^{(k)}.x.1)$ La fonction qui à x associe $F_0(y|x)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage $V(x)$ de $x \in]-1, 2[$, pour tout $y \in \mathbb{R}$.

– $(F_0^{(k)}.x.2)$ On suppose $(F_0^{(k)}.x.1)$ et on suppose de plus que pour tout u dans un voisinage $V(x)$ de $x \in]-1, 2[$ et tout $y \in \mathbb{R}$,

$$\sup_{u \in V(x)} \sup_{y \in \mathbb{R}} |F_0^{(k)}(y|x)| \leq M,$$

où $F_0^{(k)}(y|x)$ est la $k^{\text{ième}}$ dérivée de la fonction $F(y|.)$ par rapport à x.

– $(g.0)$ $f_0(y|x) + c_n g_x(y) \geq 0, \forall x \in [-1, 2]$ et $\forall y \in \mathbb{R}$.

– $(g.1)$ $\int_{\mathbb{R}} g_x(y)dy = 0, \forall x \in [-1, 2]$.

– $(g.2)$ $\int_{\mathbb{R}} |g_x(y)|dy \leq M, \forall x \in [-1, 2]$.

– $(G^{(k)}.x.1)$ La fonction qui à x associe $G(y|x)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage de tout $x \in]-1,2[$ pour tout $y \in \mathbb{R}$.

- $(G^{(k)}.x.2)$ On suppose $(G^{(k)}.x.1)$ et on suppose de plus que pour tout u dans un voisinage $V(x)$ de $x \in]-1,2[$ et pour tout $y \in \mathbb{R}$,
$$\sup_{u \in V(x)} \sup_{y \in \mathbb{R}} \left| G^{(k)}(y|u) \right| \leq M,$$
où $G^{(k)}(y|x)$ est la $k^{\text{ième}}$ dérivée de la fonction $G(y|x)$ par rapport à x.

Hypothèses relatives au Chapitre 4

- $(F_0^*.y.1)$ Pour tout $x \in [-1,2]$ et tout θ dans un voisinage $V(\theta_0)$ de θ_0, la fonction qui à y associe $F_0(y|x;\theta)$ est continûment différentiable en y, ce qui assure l'existence de la densité conditionnelle $f_0(y|x;\theta)$ de Y sachant $\{X = x\}$.

- $(F_0^{*(k)}.x.1)$ Pour tout $y \in \mathbb{R}$ et tout θ dans un voisinage $V(\theta_0)$ de θ_0, la fonction qui à x associe $F_0(y|x;\theta)$ est k $(k \geq 1)$ fois continûment dérivable dans un voisinage $V(x)$ de $x \in]-1,2[$.

- $(F_0^{*(k)}.x.2)$ On suppose $(F_0^{*(k)}.x.1)$ et on suppose de plus que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, pour tout u dans un voisinage $V(x)$ de $x \in]-1,2[$ et tout $y \in \mathbb{R}$,
$$\sup_{\theta \in V(\theta_0)} \sup_{u \in V(x)} \sup_{y \in \mathbb{R}} |F_0^{(k)}(y|x;\theta)| \leq M,$$
où $F_0^{(k)}(y|x;\theta)$ est la $k^{\text{ième}}$ dérivée de la fonction $F_0(y|x;\theta)$ par rapport à x.

- $(F_0.\theta.1)$ Pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $F_0(y|x;\theta)$ est deux fois continûment différentiable pour tout θ dans un voisinage $V(\theta_0)$ de θ_0.

- $(F_0.\theta.2)$ On suppose $(F_0.\theta.1)$ et on suppose que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$,
$$||\dot{F}_0(y|x;\theta)|| \leq M \quad \text{et} \quad ||\ddot{F}_0(y|x;\theta)||_F \leq M,$$

où $\dot{F}_0(y|x;\theta)$ et $\ddot{F}_0(y|x;\theta)$ sont respectivement les différentielles d'ordre 1 et 2 de la fonction $F_0(y|x;\theta)$ par rapport à θ, où $||.||$ désigne la norme euclidienne associée à l'espace paramétrique Θ et où $||.||_F$ désigne la norme matricielle de Frobénius définie, pour toute matrice réelle A par :

$$||A||_F = \sqrt{trace(A^T A)}.$$

- $(f_0.\theta.1)$ Pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $f_0(y|x;\theta)$ est continûment différentiable pour tout θ dans un voisinage $V(\theta_0)$ de θ_0.

- $(f_0.\theta.2)$ On suppose $(f_0.\theta.1)$ et on suppose de plus qu'il existe une fonction $h(x,y)$ positive vérifiant $\int_{\mathbb{R}} h(x,y)dy \leq M \ \forall x \in [-1,2]$ et telle que pour tout θ dans un voisinage $V(\theta_0)$ de θ_0, pour tout $x \in [-1,2]$ et tout $y \in \mathbb{R}$,

$$||\dot{f}_0(y|x;\theta)|| \leq h(x,y),$$

où $\dot{f}_0(y|x;\theta)$ est la différentielle de la fonction $f_0(y|x;\theta)$ par rapport à θ.

- $(E.1)$ $\sqrt{n}(\hat{\theta} - \theta_0) = O_p(1)$.

- $(E.2)$ $\hat{\theta} \xrightarrow{ps} \theta_0$.

- $(P.2)$ Pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$, la fonction qui à θ associe $w(y|x;\theta)$ est positive, bornée et continûment différentiable pour tout θ dans un voisinage $V(\theta_0)$ de θ_0. On suppose de plus que pour tout θ appartenant à ce même voisinage, pour tout $x \in [0,1]$ et tout $y \in \mathbb{R}$,

$$|\dot{w}(y|x;\theta)| \leq M,$$

où $\dot{w}(y|x;\theta)$ est la différentielle de la fonction $w(y|x;\theta)$ par rapport à θ.

Liste des auteurs

Abdous, B. & Bensaid, E. (2000).
Abdous, B., Berlinet, A. & Hengartner, N. (2002).
Aït-Sahalia, Y., Bickel, P. & Stoker, T.M. (1994).
Alcalá, J.T., Cristóbal, J.A. & González-Manteiga, W. (1999).
Anderson, T.W. & Darling, D.A. (1952, 1954).
Azzalini, A. & Bowman, A. (1993).
Baraud, Y., Huet, S. & Laurent, B. (2003).
Berlinet, A. (1993).
Berlinet, A., Gannoun, A. & Matzner-Løber, E. (2001).
Berlinet, A. & Thomas-Agnan, C. (2004).
Bickel, P.J. & Rosenblatt, M. (1973).
Biedermann, S. & Dette, H. (2000, 2001).
Bierens, H.J. (1982, 1990).
Bilodeau, M. & Brenner, D. (1999).
Bosq, D. & Lecoutre, J.P. (1987).
Bowman, A.W., Jones, M.C. & Gijbels, I. (1998).
Breidt, F.J. & Opsomer, J.D. (2000).
Brockmann, M., Gasser, T. & Herrmann, E. (1993).
Cabaña, A. & Cabaña, E.M. (2001).
Cheng, M-Y. (1997).
Chiu, S-T. (1996).
Cleveland, W.S. (1979).
Cleveland, W.S. & Devlin, S.J. (1988).
Collomb, G. (1980).
Collomb, G. & Härdle, W. (1986).
Cramér, H. (1928).
Cristóbal, J.A. & Alcalá, J.T. (2001).
D'Agostino, R.B. & Stephens, M.A. (1986).
Delgado, M.A. & González-Manteiga, W. (2001).
Dette, H. (1999, 2000, 2002).
Dette, H. & Munk, A. (1998).
De Wet, T. & Randles, R.H. (1987).
Diebolt, J. (1995).
Ducharme, G.R. & Ferrigno, S. (2002, 2004).
Ducharme, G.R. & Mint El Mouvid, M. (2001).

Nadaraya, E.A. (1964).
Neuhaus, G. (1976).
Opsomer, J.D. & Ruppert, D. (1997).
Pitman, E.J.G (1949).
Powell, J.L., Stock, J.H. & Stoker, T.M. (1989).
Prakasa Rao, P.L.S. (1983).
Ramil Novo, L.A. & González-Manteiga, W. (1998).
Randles, R.H. (1982).
Robinson, P.M. (1989).
Rodríguez-Campos, M., González-Manteiga, W. & Cao, R. (1998).
Rosenblatt, M. (1976).
Ruppert, D. (1997).
Ruppert, D., Sheather, S.J. & Wand, M.P. (1995).
Ruppert, D. & Wand, M.P. (1994).
Ruppert, D., Wand, M.P., Holst, U. & Hössjer, O. (1997).
Serfling, R.J. (1971, 1980).
Sheather, S.J. & Jones, M.C. (1991).
Shorack, G.R. & Wellner, J.A. (1986).
Stoker, T.M. (1989).
Stone, C.J. (1977, 1980, 1982).
Stute, W. (1991, 1997).
Stute, W. & González-Manteiga, W. (1996).
Stute, W., González-Manteiga, W. & Presedo Quindimil, M. (1998).
Templet, I. (1995).
Tenreiro, C. (1996).
Tsybakov, A.B. (1986).
Vilar-Fernández, J.M. & González-Manteiga, W. (2000).
Wand, M.P. & Jones, M.C. (1995).
Watson, G.S. (1961, 1964).
Yatchew, A.J. (1992).
Yu, K. & Jones, M.G. (1998).
Zheng, J.X. (1996).

Bibliographie

[1] **Abdous, B. & Bensaid, E.** (2000). Multivariate local polynomial fitting for a probability distribution function and its partial derivatives. *Nonparametric Statistics*, **13**, pp. 77-94.

[2] **Abdous, B., Berlinet, A. & Hengartner, N.** (2002). A general theory for kernel estimation of smooth functionals of the distribution function and their derivatives. *Revue Roumaine de Mathématiques Pures et Appliquées*, **48**(3), pp. 217-232.

[3] **Aït-Sahalia, Y., Bickel, P.J. & Stoker, T.M.** (1994). Goodness-of-fit tests for regression using kernels methods. *Working paper* (*Sloan school of Management*), *Massachusetts Institute of Technology.*

[4] **Alcalá, J.T., Cristóbal, J.A. & González-Manteiga, W.** (1999). Goodness-of-fit tests for linear models based on local polynomials. *Statistics & Probability Letters*, **42**, pp. 39-46.

[5] **Anderson, T.W. & Darling, D.A.** (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic process. *The Annals of Mathematical Statistics*, **23**, pp. 193-212.

[6] **Anderson, T.W. & Darling, D.A.** (1954). A test of goodness-of-fit. *Journal of the American Statistical Association*, **49**, pp. 765-769.

[7] **Azzalini, A. & Bowman, A.** (1993). On the use of nonparametric regression for checking linear relationships. *Journal of the Royal Statistical Society* (*Series B*), **55**(2), pp. 549-557.

[8] **Baraud, Y., Huet, S. & Laurent, B.** (2003). Adaptive tests of linear hypotheses by model selection. *The Annals of Statistics*, **31**(1), pp. 225-251.

[9] **Baraud, Y., Huet, S. & Laurent, B.** (2003). Adaptive tests of qualitatives hypotheses. *ESAIM. Probability and Statistics*, **7**, pp. 147-159.

[10] **Berlinet, A.** (1993). Hierarchies of higher order kernels. *Probability Theory and Related Fields*, **94**, pp. 489-504.

[11] **Berlinet, A., Gannoun, A. & Matzner-Løber, E.** (2001). Asymptotic normality of convergent estimates of conditional quantiles. *Statistics,* **35**, pp. 139-169.

[12] **Berlinet, A. & Thomas-Agan, C.** (2004). *Reproducing kernel hilbert spaces in probability and statistics.* Kluwer Academic Publishers, London, 355p.

[13] **Bickel, P.J. & Rosenblatt, M.** (1973). On some global measures of the deviations of density function estimates. *The Annals of Statistics*, **1**(6), pp. 1071-1095.

[14] **Biedermann, S. & Dette, H.** (2000). Testing linearity of regression models with dependant errors by kernel based methods. *Test,* **9**(2), pp. 417-438.

[15] **Biedermann, S. & Dette, H.** (2001). Optimal designs for testing the functional form of a regression via nonparametric estimation techniques. *Statistics* & *Probability Letters,* **52**, pp. 215-224.

[16] **Bierens, H.J.** (1982). Consistent model specification tests. *Journal of Econometrics*, **20**(1), pp. 105-134.

[17] **Bierens, H.J.** (1990). A consistent conditional moment test of functional form. *Econometrica*, **58**(6), pp. 1443-1458.

[18] **Bilodeau, M. & Brenner, D.** (1999). *Theory of multivariate statistics.* Springer Texts in Statistics. Springer-Verlag, New York, 288p.

[19] **Bosq, D. & Lecoutre, J.P.** (1987). *Théorie de l'estimation fonctionnelle.* Economica, Paris, 342p.

[20] **Bowman, A.W., Jones, M.C. & Gijbels, I.** (1998). Testing monotonicity of regression. *Journal of Computational and Graphical Statistics*, **7**(4), pp. 489-500.

[21] **Breidt, F.J. & Opsomer, J.D.** (2000). Local polynomial regression estimators in survey sampling. *The Annals of Statistics*, **28**(4), pp. 1026-1053.

[22] **Brockmann, M., Gasser, T. & Herrmann, E.** (1993). Locally adaptive bandwidth choice for kernel regression estimators. *Journal of the American Statistical Association*, **88**(424), pp. 1302-1309.

[23] **Cabaña, A. & Cabaña, E.M.** (2001). Goodness-of-fit tests based on quadratic functionals of transformed empirical process. *Statistics*, **35**(2), pp. 171-189.

[24] **Cheng, M-Y.** (1997). A bandwidth selector for local linear density estimators. *The Annals of Statistics*, **25**(3), pp. 1001-1013.

[25] **Chiu, S-T.** (1996). A comparative review of bandwidth selection for kernel density estimation. *Statistica Sinica*, **6**(1), pp. 129-145.

[26] **Cleveland, W.S.** (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74**(368), pp. 829-836.

[27] **Cleveland, W.S. & Devlin, S.J.** (1988). Locally weighted regression : an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, **83**(403), pp. 596-610.

[28] **Collomb, G.** (1980). Estimation non paramétrique de probabilités conditionnelles. *Comptes Rendus de l'Académie des Sciences, Paris*, **t.291**, *Série A*, pp. 427-430.

[29] **Collomb, G. & Härdle, W.** (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction : kernel regression estimation from

dependent observations. *Stochastic Processes and their Applications*, **23**, pp. 77-89.

[30] **Cramér, H.** (1928). On the composition of elementary errors. *Skandinavisk Aktuarietidskrift*, **11**, pp. 13-74/141-180.

[31] **Cristóbal, J.A. & Alcalá, J.T.** (2001). An overview of nonparametric contributions to the problem of functionnal estimation from biased data. *Test*, **10**(2), pp. 309-332.

[32] **D'Agostino, R.B. & Stephens, M.A.** (1986). *Goodness-of-fit techniques.* Marcel Dekker, Inc., New York, 560p.

[33] **Delgado, M.A. & González-Manteiga, W.** (2001). Significance testing in nonparametric regression based on the bootstrap. *The Annals of Statistics*, **29**(5), pp. 1469-1507.

[34] **Dette, H.** (1999). A consistent test for the functional form of a regression based on a difference of variance estimators. *The Annals of Statistics*, **27**(3), pp. 1012-1040.

[35] **Dette, H.** (2000). On a nonparametric test for linear relationships. *Statistics & Probability Letters* , **46**, pp. 307-316.

[36] **Dette, H.** (2002). A consistent test for heteroscedasticity in nonparametric regression based on the kernel method. *Journal of Statistical Planning and Inference*, **103**, pp. 311-329.

[37] **Dette, H. & Munk, A.** (1998). Validation of linear regression models. *The Annals of Statistics*, **26**(2), pp. 778-800.

[38] **De Wet, T. & Randles, R.H.** (1987). On the effect of substituting parameter estimators in limiting χ^2 U and V statistics. *The Annals of Statistics*, **15**(1), pp. 398-412.

[39] **Diebolt, J.** (1995). A nonparametric test for the regression function : asymptotic theory. *Journal of Statistical Planning and Inference*, **44**(1), pp. 1-17.

[40] **Ducharme, G.R. & Ferrigno, S.** (2002). Polynômes locaux et tests d'adéquation conditionnels. *Actes des XXXIV^{ièmes} Journées de Statistique, Bruxelles-Louvain-la-Neuve,* 13-17 *Mai, p*224.

[41] **Ducharme, G.R. & Ferrigno, S.** (2004). Un test non paramétrique de la fonction de répartition conditionnelle. *Actes des XXXVI^{ièmes} Journées de Statistique, Montpellier,* 24-28 *Mai, p*21.

[42] **Ducharme, G.R. & Ferrigno, S.** (2004). A nonparametric test of the conditional distribution function. *Proceedings of the 6th World Congress of the Bernoulli Society for Mathematical Statistics and Probability, Barcelona, July,* 26-31, *p*101.

[43] **Ducharme, G.R. & Mint El Mouvid, M.** (2001). Convergence presque sûre de l'estimateur linéaire local de la fonction de répartition conditionnelle. *Comptes Rendus de l'Académie des Sciences, Paris*, **t.333**, *Série I*, pp. 873-876.

[44] **Dudley, R.M.** (1989). *Real analysis and probability.* Chapman & Hall, New York, 435p.

[45] **Eubank, R.L. & Hart, J.D.** (1992). Testing goodness-of-fit in regression via order selection criteria. *The Annals of Statistics*, **20**(3), pp. 1412-1425.

[46] **Eubank, R.L., Hart, J.D. & LaRiccia, V.N.** (1993). Testing goodness-of-fit via nonparametric function estimation techniques. *Communications in Statistics. Theory and Methods*, **22**(12), pp. 3327-3354.

[47] **Eubank, R.L. & LaRiccia, V.N.** (1993). Testing for no effect in nonparametric regression. *Journal of Statistical Planning and Inference*, **36**(1), pp. 1-14.

[48] **Fan, J.** (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, **87**(420), pp. 998-1004.

[49] **Fan, J.** (1993). Local linear regression smoothers and their minimax efficiencies. *The Annals of Statistics*, **21**(1), pp. 196-216.

[50] **Fan, J.** (2000). Prospects of nonparametric modelling. *Journal of the American Statistical Association*, **95**(452), pp. 1296-1300.

[51] **Fan, J., Farmen, M. & Gijbels, I.** (1998). Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society* (*Series B*), **60**(3), pp. 591-608.

[52] **Fan, J., Gasser, T., Gijbels, I., Brockmann, M. & Engel, J.** (1997). Local polynomial regression : optimal kernels and asymptotic minimax efficiency. *Annals of the Institute of Statistical Mathematics*, **49**(1), pp. 79-99.

[53] **Fan, J. & Gijbels, I.** (1992). Variable bandwidth and local linear regression smoothers. *The Annals of Statistics*, **20**(4), pp. 2008-2036.

[54] **Fan, J. & Gijbels, I.** (1995). Adaptive order polynomial fitting : bandwidth robustification and bias reduction. *Journal of Computational and Graphical Statistics*, **4**(3), pp. 213-227.

[55] **Fan, J. & Gijbels, I.** (1995). Data-driven bandwidth selection in local polynomial fitting : variable bandwidth and spatial adaptation. *Journal of the Royal Statistic Society* (*Series B*), **57**(2), pp. 371-394.

[56] **Fan, J. & Gijbels, I.** (1996). *Local polynomial modelling and its applications.* Chapman & Hall, London, 341p.

[57] **Fan, J., Gijbels, I., Hu, T-C. & Huang, L-S.** (1996). A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, **6**, pp. 113-127.

[58] **Fan, J., Heckman, N.E. & Wand, M.P.** (1995). Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association*, **90**(429), pp. 141-150.

[59] **Fan, J. & Huang, L-S.** (2001). Goodness-of-fit tests for parametric regression models. *Journal of the American Statistical Association*, **96**(454), pp. 640-652.

[60] **Fan, J. & Yao, Q.** (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, **85**(3), pp. 645-660.

[61] **Fan, Y.** (1994). Testing the goodness-of-fit of a parametric density function by kernel method. *Econometric Theory*, **10**, pp. 316-356.

[62] **Fan, Y. & Li, Q.** (1996). Consistent model specification tests : omitted variables and semiparametric functional forms. *Econometrica*, **64**(4), pp. 865-890.

[63] **Fan, Y. & Li, Q.** (1999) Central limit theorem for degenerate U-statistics of absolutely regular processes with application to model specification testing. *Nonparametric Statistics*, **10**, pp. 245-271.

[64] **Feller, W.** (1966). *An introduction to probability theory and its applications. Vol.II.* Wiley & Sons, New York, 636p.

[65] **Ferguson, T.S.** (1996). *A course in large sample theory.* Chapman & Hall, London, 245p.

[66] **Frichot, B.** (2000). *Optimalité asymptotique et à distance finie de tests d'ajustement.* Thèse, Université Montpellier II, Sciences et Techniques du Languedoc, France, 158p.

[67] **Futschik, A. & Clarke, B.R.** (2004). Improving bandwidth selection methods by adding qualitative constraints. *Computational Statistics*, **19**, pp. 445-453.

[68] **Ghosal, S., Sen, A. & van der Vaart, A.W.** (2000). Testing Monotonicity of regression. *The Annals of Statistics*, **28**(4), pp. 1054-1082.

[69] **Ghosh, B.K. & Huang, W-M.** (1991). The power and optimal kernel of the Bickel-Rosenblatt test for goodness-of-fit. *The Annals of Statistics*, **19**(2), pp. 999-1009.

[70] **Gijbels, I. & Rousson, V.** (2001). A nonparametric least-squares test for checking a polynomial relationship. *Statistics & Probability Letters*, **51**, pp. 253-261.

[71] **González-Manteiga, W. & Cao, R.** (1993). Testing the hypothesis of a general linear model using nonparametric regression estimation. *Test*, **2**(1-2), pp. 161-188.

[72] **González-Manteiga, W. & Vilar-Fernandez, J.M.** (1995). Testing linear regression models using nonparametric regression estimators when errors are non-independent. *Computational Statistics & Data Analysis*, **20**, pp. 521-541.

[73] **Gouriéroux, C. & Montfort, A.** (1989). *Statistique et modèles économétriques.* Economica, Paris, 480p.

[74] **Gouriéroux, C. & Tenreiro, C.** (2001). Local power properties of kernel based goodness-of-fit tests. *Journal of Multivariate Analysis*, **78**(2), pp. 161-190.

[75] **Gozalo, P.L.** (1993). A consistent model specification test for nonparametric estimation of regression functions models. *Econometric Theory*, **9**(3), pp. 451-477.

[76] **Grams, W.F. & Serfling, R.J.** (1973). Convergences rates for U-statistics and related statistics. *The Annals of Statistics*, **1**, pp. 153-160.

[77] **Gregory, G.G.** (1977). Large sample theory for U-statistics and tests of fit. *The Annals of Statistics*, **5**(1), pp. 110-123.

[78] **Gregory, G.G.** (1980). On efficiency and optimality of quadratic tests. *The Annals of Statistics*, **8**(1), pp. 116-131.

[79] **Guerre, E. & Lavergne, P.** (2002). Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory*, **18**(5), pp. 1139-1171.

[80] **Hájek, J. & Sidák, Z.** (1967). *Theory of rank test.* Academic Press, New York, 297p.

[81] **Hall, P.** (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis,* **14**(1), pp. 1-16.

[82] **Hall, P. & Marron, J.S.** (1991). Lower bounds for bandwidth selection in density estimation. *Probability Theory and Related Fields,* **90**(2), pp. 149-173.

[83] **Hall, P., Marron, J.S. & Park, B.U.** (1992). Smoothed cross-validation. *Probability Theory and Related Fields,* **92**(1), pp. 1-20.

[84] **Hall, P., Sheather, S.J., Jones, M.C. & Marron, J.S.** (1991). An optimal data-based bandwidth selection in kernel density estimation. *Biometrika,* **78**(2), pp. 263-269.

[85] **Härdle, W.** (1990). *Applied Nonparametric Regression.* Cambridge University Press, Cambridge, 333p.

[86] **Härdle, W. & Mammen, E.** (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, **21**(4), pp. 1926-1947.

[87] **Härdle, W. & Stoker, T.M.** (1989). Investigating smooth multiple regression by the method of average derivatives. *Journal of the American Statistical Association*, **84**(408), pp. 986-995.

[88] **Härdle, W. & Tsybakov, A.** (1997). Local polynomial estimators of the volatility function in nonparametric autoregression. *Journal of Econometrics*, **81**(1), pp. 223-242.

[89] **Hjellvik, H. & Tjøstheim, D.** (1996). Nonparametric statistics for testing of linearity and serial independence. *Nonparametric Statistics*, **6**, pp. 223-251.

[90] **Hjellvik, V., Yao, Q. & Tjøstheim, D.** (1998). Linearity testing using local polynomial approximation. *Journal of Statistical Planning and Inference*, **68**(2), pp. 295-321.

[91] **Hoeffding, W.** (1948a). A class of statistics with asymptotic normal distributions. *The Annals of Mathematical Statistics*, **19**, pp. 293-325.

[92] **Hoeffding, W.** (1948b). A nonparametric test of independence. *The Annals of Mathematical Statistics*, **19**, pp. 546-557.

[93] **Hoeffding, W.** (1961). The strong law of large numbers for U-statistics. *Institute of Statistics Mimeo Series*, **302**, *University of North Carolina.*

[94] **Honda, T.** (1998). Testing the goodness-of-fit of a linear model by kernel regression. *Communications in Statistics. Theory and Methods*, **27**(3), pp. 529-546.

[95] **Hong, Y.** (1993). Consistent specification testing using optimal nonparametric kernel estimation. *Working paper, Cornell University.*

[96] **Hong, Y. & White, H.** (1995). Consistent specification testing via nonparametric series regression. *Econometrica*, **63**(5), pp. 1133-1159.

[97] **Horowitz, J.L. & Härdle, W.** (1994). Testing a parametric model against a semiparametric alternative. *Econometric Theory*, **10**, pp. 821-848.

[98] **Horowitz, J.L. & Spokoiny, V.G.** (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, **69**(3), pp. 599-631.

[99] **Huang, L-S.** (2001). Testing the adequacy of a linear model via critical smoothing. *Journal of Statistical Computation and Simulation*, **68**(3), pp. 281-294.

[100] **Huang, L-S. & Fan, J.** (1999). Nonparametric estimation of quadratic regression functionals. *Bernoulli*, **5**(5), pp. 927-949.

[101] **Inglot, T. & Ledwina, T.** (2001). Intermediate approach to comparison of some goodness-of-fit tests. *The Annals of the Institute of Statistical Mathematics*, **53**(4), pp. 810-834.

[102] **Jones, M.C., Marron, J.S. & Sheather, S.J.** (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*,

91(433), pp. 401-407.

[103] **de Jong, P.** (1987). A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields*, **75**, pp. 261-277.

[104] **Kendall, M.G. & Stuart, A.** (1966). *The advanced theory of statistics. Vol.*3 *: Design and analysis, and time Series.* Hafner Publishing Co., New York, 552p.

[105] **Khashimov, Sh.A.** (1988). On limit theorems for distributions of generalized von Mises functionals. *Theory of Probability and its Applications*, **33**(3), pp. 566-571.

[106] **Kim, J.T.** (2000). An order selection criterion for testing goodness-of-fit. *Journal of the American Statistical Association*, **95**(451), pp. 829-835.

[107] **Kolmogorov, A.N.** (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Instituto Italiano degli Attuari*, **4**, pp. 83-91.

[108] **Kozek, A.S.** (1991). A nonparametric test of fit of a parametric model. *Journal of Multivariate Analysis*, **37**, pp. 66-75.

[109] **Lavergne, P. & Vuong, Q.** (2000). Nonparametric signifiance testing. *Econometric Theory*, **16**, pp. 576-601.

[110] **Le Cam, L.** (1986). *Asymptotic methods in statistical decision theory.* Springer-Verlag, New York, 742p.

[111] **Lecoutre, J-P. & Tassi, P.** (1987). *Statistique non paramétrique et robustesse.* Economica, Paris, 455p.

[112] **Lee, A.J.** (1990). *U-statistics. Theory and practice.* Marcel Dekker, Inc., New York, 302p.

[113] **Lee T-H. & Ullah, A.** (2001). Nonparametric bootstrap tests for neglected nonlinearity in time series regression models. *Nonparametric Statistics*, **13**, pp.

425-451.

[114] **Lehmann, E.L.** (1959). *Testing statistical hypotheses.* Wiley & Sons, New York, 369p.

[115] **Lejeune, M.** (1985). Estimation non-paramétrique par noyaux : régression polynômiale mobile. *Revue de Statistiques Appliquées*, **33**(3), pp. 43-67.

[116] **Li, Q. & Wang, S.** (1998). A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics*, **87**, pp. 145-165.

[117] **Liero, H.** (2003). Testing homoscedasticity in nonparametric regression. *Journal of Nonparametric Statistics*, **15**(1), pp. 31-51.

[118] **Liero, H., Laüter, H. & Konakov, V.** (1998). Nonparametric versus parametric goodness-of-fit. *Statistics*, **31**(2), pp. 115-149.

[119] **Liu, Z., Stengos, T. & Li, Q.** (2000). Nonparametric model check based on local polynomial fitting. *Statistic & Probability Letters*, **48**, pp. 327-334.

[120] **Loader, C.R.** (1999). Bandwidth selection : classical or plug-in ? *The Annals of Statistics*, **27**(2), pp. 415-438.

[121] **Masry, E.** (1996). Multivariate local polynomial regression for time series : uniform strong consistency and rates. *Journal of Times Series Analysis*, **17**(6), pp. 571-599.

[122] **Masry, E.** (1996). Multivariate regression estimation. Local polynomial fitting for time series. *Stochastic Processes and their Applications*, **65**(1), pp. 81-101.

[123] **Mint El Mouvid, M.** (2000). *Sur l'estimateur linéaire local de la fonction de répartition conditionnelle.* Thèse, Université Montpellier II, Sciences et Techniques du Languedoc, France, 160p.

[124] **von Mises, R.** (1941). On the foundations of probability and statistics. *The Annals of Mathematical Statistics*, **12**, pp. 195-205.

[125] **Mugdadi, A.R. & Ahmad, I.A.** (2004). A bandwidth selection for kernel density estimation of functions of random variables. *Computational Statistics and Data Analysis*, **47**, pp. 49-62.

[126] **Muller, H.G.** (1987). Weighted local regression and kernel methods for nonparametric curve fitting. *Journal of the American Statistical Association*, **82**(397), pp. 231-238.

[127] **Munk, A. & Czado, C.** (1998). Nonparametric validation of similar distribution and assessement of goodness-of-fit. *Journal of the Royal Statistical Society* (*Series B*), **60**(1), pp. 223-241.

[128] **Nadaraya, E.A.** (1964). On estimating regression. *Theory of Probability and its Applications*, **9**, pp. 141-142.

[129] **Neuhaus, G.** (1976). Asymptotic power properties of the Cramér-von Mises test under contiguous alternatives. *Journal of Multivariate Analysis*, **6**, pp. 95-110.

[130] **Opsomer, J.D. & Ruppert, D.** (1997). Fitting a bivariate additive model by local polynomial regression. *The Annals of Statistics*, **25**(1), pp. 186-211.

[131] **Pitman, E.J.G.** (1949). *Lecture notes on nonparametric statistical inference.* Columbia University, New York, 234p.

[132] **Powell, J.L., Stock, J.H. & Stoker, T.M.** (1989). Semiparametric estimation of index coefficient. *Econometrica*, **57**(6), pp. 1403-1430.

[133] **Prakasa Rao, B.L.S.** (1983). *Nonparametric functional estimation.* Academic Press, London, 522p.

[134] **Ramil Novo, L.A. & González-Manteiga, W.** (1998). χ^2 goodness-of-fit tests for polynomial regression. *Communications in Statistics. Simulation and Computation*,

27(1), pp. 229-258.

[135] **Randles, R.H.** (1982). On the asymptotic normality of statistics with estimated parameters. *The Annals of Statistics*, **10**(2), pp. 462-474.

[136] **Robinson, P.M.** (1989). Hypothesis testing in semiparametric and nonparametric models for econometric time series. *Review of Economic Studies*, **56**(4), pp. 511-534.

[137] **Rodríguez-Campos, M., González-Manteiga, W. & Cao, R.** (1998). Testing the hypothesis of a generalized linear regression model using nonparametric regression estimation. *Journal of Statistical Planning and Inference*, **67**, pp. 99-122.

[138] **Rosenblatt, M.** (1976). On the maximal deviation of k-dimensional density estimates. *The Annals of Probability*, **4**(6), pp. 1009-1015.

[139] **Ruppert, D.** (1997). Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation. *Journal of the American Statistical Association*, **92**(439), pp. 1049-1062.

[140] **Ruppert, D., Sheather, S.J. & Wand, M.P.** (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, **90**(432), pp. 1257-1270.

[141] **Ruppert, D. & Wand, M.P.** (1994). Multivariate locally weighted least squares regression. *The Annals of Statistics*, **22**(3), pp. 1346-1370.

[142] **Ruppert, D., Wand, M.P., Holst, U. & Hössjer, O.** (1997). Local polynomial variance-function estimation. *Technometrics*, **39**(3), pp. 262-273.

[143] **Serfling, R.J.** (1971). The law of the iterated logarithm for U-statistics and related von Mises statistics. *The Annals of Mathematical Statistics*, **42,** pp. 17-94.

[144] **Serfling, R.J.** (1980). *Approximation theorems of mathematical statistics.* Wiley & Sons, New York, 371p.

[145] **Sheather, S.J. & Jones, M.C.** (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistic Society* (*Series B*), **53**(3), pp. 683-690.

[146] **Shorack, G.R. & Wellner, J.A.** (1986). *Empirical processes with applications to statistics.* Wiley & Sons, New York, 938p.

[147] **Stoker, T.M.** (1989). Tests of additive derivative constraints. *Review of Economic Studies*, **56**(4), pp. 535-552.

[148] **Stone, C.J.** (1977). Consistent nonparametric regression. *The Annals of Statistics*, **5**(4), pp. 595-645.

[149] **Stone, C.J.** (1980). Optimal rates of convergence for nonparametric estimators. *The Annals of Statistics*, **8**(6), pp. 1348-1360.

[150] **Stone, C.J.** (1982). Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, **10**(4), pp. 1040-1053.

[151] **Stute, W.** (1991). Conditional U-statistics. *The Annals of Probability*, **19**(2), pp. 812-825.

[152] **Stute, W.** (1997). Nonparametric model checks for regression. *The Annals of Statistics*, **25**(2), pp. 613-641.

[153] **Stute, W. & González-Manteiga, W.** (1996). NN goodness-of-fit tests for linear models. *Journal of Statistical Planning and Inference*, **53**(1), pp. 75-92.

[154] **Stute, W., González-Manteiga, W. & Presedo Quindimil, M.** (1998). Bootstrap approximations in models checks for regression. *Journal of the American Statistical Association* , **93**(441), pp. 141-149.

[155] **Templet, I.** (1995). *Test d'adéquation des résidus d'un modèle de régression.* Rapport de stage de DEA de Biostatistique, Université Montpellier II, Sciences et

Techniques du Languedoc, France, 42p.

[156] **Tenreiro, C.** (1996). Tests d'ajustement à une densité fondés sur un estimateur non paramétrique à noyau pour des observations dépendantes. *Annales d'Économie et de Statistique*, **43**, pp. 129-148.

[157] **Tsybakov, A.B.** (1986). Robust construction of functions by local approximation method (Russian). *Problemy Peredachi Informatsii*, **22**(2), pp. 69-84.

[158] **Vilar Fernández, J.M. & González-Manteiga, W.** (2000). Resampling for checking linear regression models via non-parametric regression estimation. *Computational Statistics and Data Analysis*, **35**, pp. 211-231.

[159] **Wand, M.P. & Jones, M.C.** (1995). *Kernel Smoothing.* Chapman & Hall, London, 212p.

[160] **Watson, G.S.** (1961). Goodness-of-fit tests on a circle. *Biometrika*, **48**, pp. 109-114.

[161] **Watson, G.S.** (1964). Smooth regression analysis. *Sankhyā,* (*Series A*), **26**, pp. 359-372.

[162] **Yatchew, A.J.** (1992). Nonparametric regression tests based on least squares. *Econometric Theory*, **8**(4), pp. 435-451.

[163] **Yu, K. & Jones, M.G.** (1998). Local linear quantile regression. *Journal of the American Statistical Association*, **93**(441), pp. 228-237.

[164] **Zheng, J.X.** (1996). A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics*, **75**(2), pp. 263-289.

Résumé : Soient X et Y, deux variables aléatoires. De nombreuses procédures statistiques permettent d'ajuster un modèle à ces données dans le but d'expliquer Y à partir de X. La mise en place d'un tel modèle fait généralement appel à diverses hypothèses que l'on doit valider pour justifier son utilisation. Dans ce travail, on propose une approche globale où toutes les hypothèses faites pour asseoir ce modèle sont testées simultanément. Plus précisément, on construit un test basé sur une quantité qui permet de canaliser toute l'information liant X à Y : la fonction de répartition conditionnelle de Y sachant $\{X = x\}$ définie par $F(y|x) = P(Y \leq y|X = x)$. Notre test compare la valeur prise par l'estimateur polynômial local de $F(y|x)$ à une estimation paramétrique du modèle supposé et rejette sa validité si la "distance" entre ces deux quantités est trop grande. Dans un premier temps, on considère le cas où la fonction de répartition supposée est entièrement spécifiée et, dans ce contexte, on établit le comportement asymptotique du test. Dans la deuxième partie du travail, on généralise ce résultat au cas plus courant en pratique où le modèle supposé contient un certain nombre de paramètres inconnus. On étudie ensuite la puissance locale du test en déterminant son comportement asymptotique local sous des suites d'hypothèses contigües. Enfin, on propose un critère de choix de la fenêtre d'ajustement qui intervient lors de l'étape d'estimation polynômiale locale de la fonction de répartition conditionnelle.

Mots-clés : Alternatives contigües, estimation non paramétrique, estimation polynômiale locale, fenêtre d'ajustement, fonction de répartition conditionnelle, puissance asymptotique locale, statistique de Cramér-von Mises, U-statistique, test d'adéquation.

Abstract : Let (X, Y) be a random vector of $\mathbb{R}^2$. Many statistical procedures allow to fit models to such data but this usually requires many assumptions that must somehow be validated before the model can be used with some degree of confidence. In this work, we propose a global test that can assess the validity of the supposed model by testing simultaneously all the hypotheses made about the model. This test is based on a quantity that embodies all the information about the joint behaviour of X and Y : the conditional distribution function $F(y|x) = P(Y \leq y|X = x)$ and on a standard paradigm that consists in comparing the local polynomial estimator of $F(y|x)$ with a parametric one and rejecting the model if the distance between these two quantities exceeds a critical value. In the first part, we give the asymptotic behaviour of our test statistic when the supposed conditional distribution function does not involve unknown parameters. In a second part, we generalize to the case where such parameters are involved. Then, we give the local asymptotic power of the test by studying its asymptotic behaviour under contiguous alternatives. Finally, we propose a choice for the bandwidth parameter used in the local polynomial estimation of the distribution function.

Keywords : Bandwidth, conditional distribution function, contiguous hypotheses, Cramer-von Mises statistic, goodness-of-fit test, local asymptotic power, local polynomial estimation, nonparametric estimation, U-statistic.

Résumé : Soient X et Y, deux variables aléatoires. De nombreuses procédures statistiques permettent d'ajuster un modèle à ces données dans le but d'expliquer Y à partir de X. La mise en place d'un tel modèle fait généralement appel à diverses hypothèses que l'on doit valider pour justifier son utilisation. Dans ce travail, on propose une approche globale où toutes les hypothèses faites pour asseoir ce modèle sont testées simultanément. Plus précisément, on construit un test basé sur une quantité qui permet de canaliser toute l'information liant X à Y : la fonction de répartition conditionnelle de Y sachant $\{X = x\}$ définie par $F(y|x) = P(Y \leq y|X = x)$. Notre test compare la valeur prise par l'estimateur polynômial local de $F(y|x)$ à une estimation paramétrique du modèle supposé et rejette sa validité si la "distance" entre ces deux quantités est trop grande. Dans un premier temps, on considère le cas où la fonction de répartition supposée est entièrement spécifiée et, dans ce contexte, on établit le comportement asymptotique du test. Dans la deuxième partie du travail, on généralise ce résultat au cas plus courant en pratique où le modèle supposé contient un certain nombre de paramètres inconnus. On étudie ensuite la puissance locale du test en déterminant son comportement asymptotique local sous des suites d'hypothèses contigües. Enfin, on propose un critère de choix de la fenêtre d'ajustement qui intervient lors de l'étape d'estimation polynômiale locale de la fonction de répartition conditionnelle.

Mots-clés : Alternatives contigües, estimation non paramétrique, estimation polynômiale locale, fenêtre d'ajustement, fonction de répartition conditionnelle, puissance asymptotique locale, statistique de Cramér-von Mises, U-statistique, test d'adéquation.

Abstract : Let (X, Y) be a random vector of $\mathbb{R}^2$. Many statistical procedures allow to fit models to such data but this usually requires many assumptions that must somehow be validated before the model can be used with some degree of confidence. In this work, we propose a global test that can assess the validity of the supposed model by testing simultaneously all the hypotheses made about the model. This test is based on a quantity that embodies all the information about the joint behaviour of X and Y : the conditional distribution function $F(y|x) = P(Y \leq y|X = x)$ and on a standard paradigm that consists in comparing the local polynomial estimator of $F(y|x)$ with a parametric one and rejecting the model if the distance between these two quantities exceeds a critical value. In the first part, we give the asymptotic behaviour of our test statistic when the supposed conditional distribution function does not involve unknown parameters. In a second part, we generalize to the case where such parameters are involved. Then, we give the local asymptotic power of the test by studying its asymptotic behaviour under contiguous alternatives. Finally, we propose a choice for the bandwidth parameter used in the local polynomial estimation of the distribution function.

Discipline : Mathématiques appliquées et applications des Mathématiques.

Unité d'accueil : Equipe Probabilités et Statistique. Institut de Mathématiques et de Modélisation de Montpellier. UMR CNRS 5149. CC051 - Université Montpellier II. Place E.Bataillon. 34095 Montpellier Cedex 05.