N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
ACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT

Présentée par :

Henien Djelloul Saiah Aicha

Domaine : Mathématiques-Informatiques.

Filière : Mathématiques.

Intitulé de la formation : Statistique, Mathématiques Appliquées à l'économie et à la finance.

Intitulée

Sur le test d'hétéroscédasticité en analyse statistique des données fonctionnelles

Soutenue le : 05 Mars 2019

Devant le jury composé de :

Président : Pr. Benaissa Samir, Univ. Djillali Liabès, Sidi Bel Abbès, Algérie.

Examinateurs : Pr. Benchikh Toufik , Univ. Djillali Liabès, Sidi Bel Abbès, Algérie.

Dr. Mechab Boubaker, Univ. Djillali Liabès, Sidi Bel Abbès, Algérie.
Pr. Guendouzi Toufik , Univ. Saida, Algérie.

Directeur de thèse : Pr. Rachdi Mustapha, Univ. Grenoble Alpes, France.

Co-Directeur de thèse : Pr. Laksaci Ali, Univ. Djillali Liabès, Sidi Bel Abbès, Algérie.

Année universitaire : 2018 / 2019.

Contents

Abstract

The work presented in this thesis covers two issues, those concerned the **heteroscedasticity** test in functional nonparametric estimation. We will concentrate in the first one, on the nonparametric functional model, which corresponds to a nonparametric regression when the regressor is a functional variable. Once, in the second work we focused on the parametric case with a nonlinear regression operator, in this case the functional regression assumed to belong to a parametric family of functions.

To construct our tests statistics we based on the empirical version of the moment condition by evaluating the difference between the conditional variance and unconditional variance.

By adding some standard assumptions, our tests statistics have a asymptotic normal distribution under the homoscedasticity's hypothesis, we also established that those tests can detect local alternatives distinct from the null.

It worth to noting that, as well as the kernel estimation, some tools have been used here as the small ball probabilities $P(X \in d(x, X))$ where it appears in the asymptotic developments where X$\in \mathcal{F}$ and d is a semi-metric, in addition to the degenerate and non-degenerate U-statistic theories.

Finally, to testing our results some simulated data examples are presented .

Acknowledgments

I dedicate this thesis

To my parents, whom I can not thank enough,

To my son Ilyane,

To my husband Mourad,

To my sister,

To my brothers Abd raouf, Abd ennour and Youcef,

To all my family.

To my Masters

To our Master and Jury President

Professor Benaissa Samir

University of Sidi Bel Abbes

You do me the honor of chairing our thesis jury.

I benefited from your knowledge during your teachings.

Your simplicity and your humility as well as your rigor for the work amazed me.

Please accept my gratitude.

To our Master and Judge

Professor Benchikh Toufik

University of Sidi Bel Abbes

You do me a great honor to judge this thesis.

The extent of your knowledge, your kindness and your availability have always aroused my admiration.

Please find, dear Master, in this work the expression of my esteem and my consideration.

To our Master and Judge

Professor Guendouzi Toufik

University of Saida

Thank you for having honored me with your presence.

You kindly agreed to judge this thesis.

This honor touches me infinitely and I wish to express my deep gratitude to you.

Please accept, dear Master, in this work the assurance of my esteem and deep respect.

To our Master and Judge

Professor Mechab Boubakeur

University of Sidi Bel Abbes.

You do me the great honor to accept to sit on this jury.

Your human and professional qualities are recognized by everyone.

Thank you for the respect and consideration you have shown me since our first meeting.

Please find in this work all my admiration.

To our Masters and Thesis Directors

Professor LAKSACI ALI

University of sidi bel abbes.

Professor RACHDI MUSTAPHA

University Grenoble Alps, laboratory AGEIS EA 7407, France.

You have done me the honor to confide to me the subject of this thesis.

Thank you for the kindness and spontaneity with which you have kindly directed this work.

I had the great pleasure of working under your direction, and I found with you the advisers and the guides.

You have received me under all circumstances with sympathy and kindness.

Your competence, your dynamism, your rigor, your modesty and your human and professional qualities have given me great admiration and respect.

Dear Masters, please find here the testimony of my sincere gratitude and deep gratitude.

Chapter 1: Introduction

Contents

1.1 Contribution of the thesis

1.1.1 The thesis plan

Our main goal is to propose a consistent nonparametric tests of heteroscedasticity in the functional space. For this dissertation, we follow the steps below :

We decompose our work into four Chapters. Indeed, the first one is an introductory one in where we give a brief history to the functional statistics and nonparametric approaches, in addition to the functional, nonparametric and finally heteroscedasticity tests. We also present in the last Section the results that we have obtained in the next to Chapters.

The second Chapter is concerning the heteroscedasticity test in the nonparametric regression. Indeed,we transform the problem of the heteroscedasticity test into a problem of comparison between the variance and the conditional variance. We based on the kernel estimation of the conditional variance to construct our test statistic. We demonstrate the asymptotic distribution of the test statistic and we quantify the robustness of the test by giving the convergence rate in probability when the null hypothesis is disturbed and we finish by giving some simulated data examples.

For the third Chapter, we consider the parametric case by taking into account the setup just as the second Chapter. Actually, we extend our results to the nonlinear regression. In both Chapters (the second and the third), the proofs of our results are offered in the last of each Chapter.

We finally conclude our dissertation by summarizing the results obtained and discuss the main directions of the prospective works.

1.1.2 What heteroscedasticity is ?

Let's point out that, in this work, our results are given in the infinite dimensional space, meaning that, when the data are of functional kind. Indeed, there exist a various areas in which the functional approaches have been used, either at physiological, biological, climatological. . . , to illustrate our problematic, let us take an example in the chemometric one.

Noting that, during the chopped meat packaging, the grease level must been putting on. To give the level of grease existed in a piece of meat, it possible to make an accurate chemical analysis, however, this process takes a large of times, money and deteriorates the piece studied. Consequently, required another useful processes. From here considered to predicting grease content from spectrometric curves.
The spectrometric data have a functional kind which can be summarizing by curves. Such that, those data are obtained by measuring the absorbance of lights of different wavelengths in each piece of meat. We are interested in a dataset from the study by chemical analysis and spectrometry of 215 pieces of meat (this dataset existed at http://lib.stat.cmu.edu/datasets/tecator). Thus, 215 spectrometric curves and corresponding grease levels are available (see Figure 1.1).

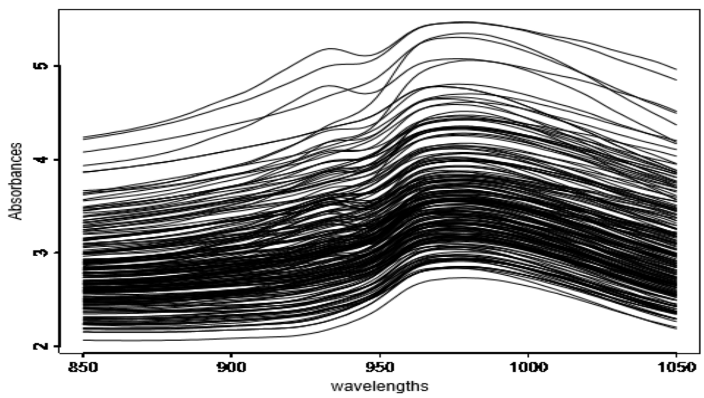

Figure 1.1: *Spectrometric curves obtained from the 215 pieces of meat studied*

By using neural network methods, **Borggaard** and ***Thodberg*** [1992] have been the first ones interested to divine how much the piece of meat contain of grease from the spectrometric curve associated with it. Actually, our work devoted to the prediction of a real variable from a functional one. Thus, it could be used to predict grease levels from of the spectrometric curve. Several authors make an assumption to know how the grease content and the spectrometric curve are connected. Thus, we can answer to this kind of question and pose our hypothesis test, is there **heteroscedasticity** or not ?

Assume that we have the couple (X, Y) in which X takes values in infinite dimensional space while Y a real random variables and those are connected by the following relations :

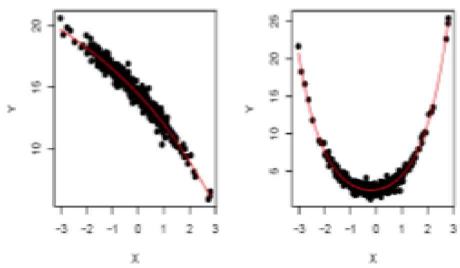

Figure : Cas 1 Cas 2

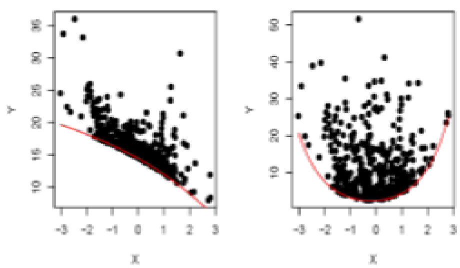

Figure : Cas 1 Cas 2

Case 1: homoscedastic framework

$$Y = r(X) + \epsilon.$$

Case 2: heteroscedastic framework

$$Y = r(X) + \sigma(X)\epsilon.$$

For the first case we make the assumption that $V(\epsilon_i) = \sigma^2$ for all i. That is, the variance of the error term is constant. If the error terms do not have constant variance, they are said to be heteroscedastic.

1.2 Bibliographic

1.2.1 Functional data

History

Keeping up with the age of speed the data are more being recorded continuously during the time interval, typically, its have a massive volume consisting of billion or trillions of records which was so difficult to process using traditional techniques, consequently, there is a need to propose an alternative methods adapted to this kind of data. From here required a new statistic branch called Functional Statistic, which treat the observations like a functional random elements valued in an infinite dimensional space.

It's worth noting that, statistical inference for Functional Data Analysis [F.D.A] has been the focus of several investigations, the area has an older history backs to the fiftieth, precisely in 1950. Actually, appeared in the Swedish statistician ***Grenander***'s thesis entitled "stochastic processes and statistical inference", in which, he showed the possibility of applying statistical concepts and methods of inference to stochastic processes and obtaining practically working methods of this kind by studying special cases of inference. In the same year, ***C.R. Rao***, at that time was at the start of his graduate studies, attracted to the foregoing considered subject, thence, he decided to prepare for the necessary mathematics to appreciate and possibly make some contributions to the area, and so on, it already got. In 1958, ***Rao***, taking into consideration the principal component analysis and the factorial analysis of the data, studied some statistics methods to compare some growth curves, similarly, ***Tucker***[1958] established parameters of the functional relation between two variables by using some factorial analysis techniques. In the beginning of the sixties, the climatologists ***Obhukov***, ***Holmstrom*** and ***Buell*** [1960, and 1971 respectively] were interested in pressure fields, thus, they proposed the first factor analysis of functional variables.

And so on, it was increasingly common to yield functional data in scientific studies as was the case, ***Deville***[1974] in economy, ***Molenaar*** and ***Boomsma***[1987] then ***Kirkpatrick*** and ***Heckman*** [1989] in the field of genetics, in where, they were interesting to observations in the form of trajectories. ***Borggaard*** and ***Thodberg*** [1992] provide interesting applications of the functional principal component analysis to chemistry and many recent applications of functional data analysis are discussed in ***ferraty*** [2011].

As we know, the ***Hotelling*** and ***Pabst***'s publication in 1936 has been the inception of nonparametric statistics, in addition, its methods enabled to be used in a various situations and dominant in most statistical journals [a special issues of Statistical Science 2004], it's wellknown in F.D.A that, this methods are more appropriate than the parametric one due to the useless of the graphical tool where it becomes very hard to exploring the relationship between the co-variables.

In the current years, the nonparametric modeling takes a large place in the F.D.A. Actually, the first model that has been estimated nonparametrically was in 1998, in which ***Gasser et al*** have proposed an estimator of the mode with verifying the fractal condition, after few years [2000], ***Ferraty*** and ***Vieu*** have explored, by taking the same fractal condition, the almost complete convergence of the kernel estimation of the regression function, and then, ***Dabo-Niang***[2002] has established the almost sure convergence and the asymptotic normality of histogram type estimator of the functional density, while, ***Rhomari*** and ***Dabo-Niang***[2004] have obtained the convergence in L^p norm, in the independent case, of the kernel estimation of the nonparametric regression. For the mixing case, ***Ferraty et al*** [2004] have studied the almost complete convergence, while, ***Masry*** [2005] presented the asymptotic normality.

For the basic results on the nonparametric approaches, there exist interesting books as : ***Ferraty*** and ***Vieu*** [2006] or ***Cuevas*** [2014] and ***Goia*** and ***Vieu*** [2016] for recent advanced and references.

1.2.2 The functional regression

The prediction of scalar response given explanatory functional random variable is an important subject in the modern statistics. The regression operator is the most preferred model in this prediction problem.
Indeed, the first model has been introduced and studied in 1991 by ***Ramsay*** and ***Delzell*** , in which, they show how the theory of L-splines can support generalizations of linear modeling and principal components analysis to samples drawn from random functions. After two years, ***Hastie*** and ***Mallows*** discussed the overview of ***Frank*** and ***Friedman*** [1993] , indeed, they Compare the partial least squares method with similar techniques such as ridge regression and principal components regression. In 2000, by using the regularity condition, and without asking for the regression operator form, ***Ferraty*** and ***Vieu*** introduced the first results concerned a generalized estimator.

1.2.3 Tests : functional statistic

Testing hypothesis in the functional statistics takes an important place in the literature, however, those tests are restricted. Indeed,to comparing two groups of curves some tests have been proposed in 1998 by ***Fan and Lin***. Actually, their functional data have been decomposed to series expansions by either Wavelet or Fourier. ***Viele*** [2001] has been offered a test that validate or not the stochastic modulation of the probability law that existed the data. In 2004, an ANOVA-like test has been proposed by ***Cuevas et al*** in the case of independent curves, their assumptions was about the noise and the observation of the curves. ***Ferraty et al*** [2007] have been introduced a test by considered some covariances hypothesis, this test based on the factors analysis compared different groups of functional data. In the same year, ***Chen and Zhang***, take into account the model framework just as ***Cuevas et al*** [2004] and discused in two cases an L^2-norm test: the curves are tainted with error measurement and are taken from a concentrate grid of points. ***Hall*** and ***Vial*** [2006b] suggest a test on dimension reduction of the studied variable. The test of ***James*** and ***Sood*** [2006] and ***Mas*** [2007a] carried on the expectation form of the studied functional random variable. In 2010, the t-test and a global L^2-norm have been deeply discussed by ***Zhang et al*** in the situation of two samples of curves. The inferential methods through the bootstrap are studied in 2O12 by ***Crainiceanu et al*** for the mean profiles in addition to the likelihood-ration statistics where have been offered by ***Staicu et al*** in 2014.

Furthermore, we find also a proposed tests to the prediction of scalar response given explanatory functional random variable by using the regression operator. In the literature, this kind of model is very few. ***Gadiaga*** and ***Ignaccolo*** [2005] introduce a non-effect test of the explanatory variable based on Methods of projection.***Cardot et al*** [2003, 2004] consider a test in the linear functional model, ***Chiou*** and ***Müller*** [2007] propose a heuristic suitability test based on the Decomposition into functional principal components of the explanatory variable. ***Shen*** and ***Faraway*** [2004] have been tested in linear model the functional variable influence when the variable response is functional and the explanatory one is a vector. ***Cardot et al*** [2007] offer a structural tests based on permutation methods. ***Mas*** [2000] propose a nullity test of the autocorrelation operator of an ARH. ***Antoniadis*** and ***Sapatinas***'s test [2007], when the two variables are functional, a model with a functional mixed effect.

1.2.4 Nonparametric tests

The first nonparametric test appeared in 1710 in the works of ***J. Arbuthnot*** who introduced the sign test. But most nonparametric tests were developed between 1940 and 1955. We make special mention of the articles of:***Andrey Nikolacvich Kolmogorov*** in 1933 who invented the goodness-of-fit test for a sample, The rank sum test was introduced by ***Frank Wilcoxon*** in 1945, in which it used to determine whether two independent samples were selected from populations having the same distribution. ***Hanry Mann*** and ***Donald Whitney*** published their proposal in 1947 [on a test whether one of two random variables is stochastically larger than the other]. The test however, is older, it was introduced at least six times in addition to ***Wilcoxon*** [1945] and ***Mann*** and ***Whitney*** [1947]. The first who proposed the test was ***Gustav deuchler*** in 1914. The test was popularized by ***Sidney Siegel*** [1956] in his influential textbook on nonparametric statistics. In 1950, ***Mood. A.M*** wrote an introduction to the theory of statistics, Two years later, ***Kruskal.W.H*** and ***Wallis.W.A*** offered a test using mostly to know if there is variance between the means in the population. Later many other articles were added to this list. ***Savage, I.R*** [1962] published a bibliography of about 3000 articles, written before in 1962, concerning nonparametric tests.

1.2.5 Heteroscedasticity test

It is widely known in the standard regression model that the variances are the same for all observation, even so, it could be found in practical applications that the errors have a different variance, in this case they are said to be heteroscedastic. However, the model is not efficiency in the presence of this situation, it necessitates the homoscedasticity of the data. Thus, it is utmost interesting to detect the heteroscedasticity for data.

To testing this phenomenon several authors have proposed either parametric or nonparametric methods in the literature. Indeed, In 1965, ***S.Goldfeld*** and ***R.Quandt*** have proposed a parametric and nonparametric test for detecting the homoscedasticity of the residuals. ***Ramsey*** [1968] developed a test to detect the existing of specification error in part of his Ph.D. thesis (called Ramsey Regression Equation Specification Error Test (RESET)), for the linear regression model. ***Rutemiller*** and ***Bowers*** [1968], suggested an estimation method in a Heteroscedastic Regression Model for obtaining a multiple linear regression equation which permits either the variance or the mean, of normally distributed random variables. ***Glejser*** [1969], presented his heteroscedasticity test in which he based on ordinary least-squares to obtain the used residuals in the regression. In 1974, ***Harvey*** and ***Phillips*** proposed their parametric test constructed on recursive residuals obtained by a simple calculate in the general linear model . ***Harvey*** in

1976 presented a paper on estimating regression models with multiplicative heteroscedasticity (where he suggested a likelihood ratio test for heteroscedasticity). ***Bickel*** studied in 1978 the two tests proposed by ***Anscombe***. Actually, he examined in the linear model , the heteroscedasticity and non-linearity asymptotic power of the tests. ***Breusch*** and ***Pagan*** [1979] offered a test in linear regression model by utilizing the same structure then of lagrangian multiplier test. ***White*** [1980] presented a consistent parametric test of heteroscedasticity in linear regression model. ***Cook*** and ***Weisberg*** [1983] and ***Tsai*** [1986] has obtained the score test statistic with parametric variance function and non-constant variance, which has been modified by ***Simonoff*** and ***Tsai*** [1994] for linear models.

Furthermore, there have been in the literature a widely papers devoted to state heteroscedasticity in the case of the nonparametric regression. Indeed, under the normality assumption on error term, ***Eubank, Thomas*** and ***Muller, Zhao*** [1993, 1995.resp] have presented a test of heteroscedasticity constructed by basing on the kernel methods in nonparametric regression models. In 1998, the suggested test of ***Dette*** and ***Munk*** was based on the best L^2-approximation of the variance function estimator which was extended to partially linear regression models in 2005 by ***You*** and ***Chen***. In 2001, ***Naito*** offered by taking into account marked empirical process of the squared residuals, a test in nonparametric model. ***Zhu, Fujikoshi*** and ***Naito*** [2001] offered test in parametric and nonparametric regression models. To construct their own test they based on the integrated difference between the conditional variance and unconditional variance weighted . ***Dette*** [2002] has checked the heteroscedasticity in nonparametric regression by extending the idea of ***Zheng*** [1996] of goodness of fit of the mean regression. ***Stute*** and ***Zhu*** [2005] tested the SIM structure. By using estimators of the distribution of residuals under the null and alternative, ***Dette et al*** [2007] searched closely , in nonparametric regression, for the issues that could be existed in testing the conditional variance once it have a parametric form. ***Zhang*** and ***Mei*** [2008] obtained a test for the constant variance of the model errors based on residual analysis.

1.3 Main results

In this section we state the results obtained under some standard assumptions, the first case concerning the heteroscedasticity test when the functional regression has a nonparametric form, while, the second one it about the parametric case with a nonlinear regression operator.

- **For the first results**

(H1) There exists $m \geq 2$ such that $\mathbb{E}[Y^m|X = x] < \delta_m(x) < C < \infty$ with $\delta_m(.)$ continuous on $\mathcal{S}$.

(H2) $\mathbb{E}[\epsilon^8|X = x] \leq b(x)$ with $b(x)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[b^2(X)] < \infty$.

(H3) The kernel K is a differentiable function supported on $[0, 1]$ such that

$$K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds > 0 \quad \text{and} \quad K(1) - \int_0^1 (K(s))'\tau(s)ds \neq 0.$$

(H4) The bandwidth parameter $h := h(n)$ is strictly positive such that:

$$n \to 0, n\,\phi(h) \to \infty, n\sqrt{\phi(h)}\max\left(h^{4\beta}, \frac{1}{\log^2 n}\right) \to 0 \text{ and } \frac{(\log n)^2}{n\,\phi(h)} < \psi_{\mathcal{S}}\left(\frac{\log n}{n}\right) < \frac{n\,\phi(h)}{\log n}$$

when n tends to infinity.

We state the following results:

- Under H_0

Theorem 1.1. *When (H1)-(H4) and (2)-(6) hold we have*

$$n\sqrt{\phi(h)}\,W_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \text{ as } n \to \infty$$

where $s^2 = 2\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)\mathbb{E}\left[f(X)V^2\left[\epsilon_2^2|X\right]\right]$.

Moreover,

$$T_n = n\sqrt{\phi(h)}\frac{W_n}{\widehat{s}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \to \infty$$

where

$$\widehat{s}^2 = \frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} K\left(\frac{d(X_i, X_j)}{h}\right)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)^2(\widehat{\epsilon_i}^2 - \widehat{\sigma}^2)^2.$$

- Under H_1

Theorem 1.2. *When (H1)-(H4) and (2)-(6) hold we have*

$$\frac{T_n}{n\sqrt{\phi(h)}} \longrightarrow \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 f(X)]/s_1, \quad \textit{In probability}$$

$$\textit{where } s_1^2 = \frac{\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)}{\left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)} \mathbb{E}[(V[\epsilon^2|X] + (V[\epsilon|X] - \sigma^2)^2)f(X)].$$

For the robustness of the test we introduce the following sequence of local alternatives

$$H_{1n}: \quad V[\epsilon|x] - \sigma^2 = \delta_n g(x)$$

Corollary 1.1. *Given (H1)-(H4) and (2)-(6), we have, under H_{1n} with $\delta_n = n^{-1/2}\phi^{-1/4}(h)$*

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, 1) \textit{ as } n \to \infty$$

where $\mu = \left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)\mathbb{E}[g^2(X)f(X)]/s.$

- **For the second results**

(H1) $\mathbb{E}[\epsilon^8|X = x] \leq b(x)$ with $b(x)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[b^2(X)] < \infty$.

(H2) The parameter space Θ is a compact and convex subset of $\mathbb{R}$. $f(X, \partial)$ is a Borel measurable function on $\mathcal{F} \times \mathbb{R}$ for each ∂ and a twice continuously differentiable real function on Θ for each $X \in \mathcal{F}$. Moreover, $E\left[\sup_{\partial\in\Theta} f(X_i, \partial)^2\right] < \infty$, and

$$E\left[\sup_{\partial\in\Theta}\left|\frac{\delta^2 f(X_i,\partial)}{\delta\partial}\right|\right] < \infty,$$
$$E\left[\sup_{\partial\in\Theta}\left|(y_i - f(X_i,\partial))^2\frac{\delta^2 f(X_i,\partial)}{\delta\partial}\right|\right] < \infty.$$

(H3) $E\left[(y_i - f(X_i, \partial))^2\right]$ takes a unique minimum at $\partial_0 \in \Theta$.

(H4) The kernel K is a differentiable function supported on $[0, 1]$ such that

$$K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds > 0 \quad \text{and} \quad K(1) - \int_0^1 (K(s))'\tau(s)ds \neq 0.$$

(H5) The bandwidth parameter $h := h(n)$ is strictly positive such that:

$$h \to 0, n\,\varphi(h) \ \to \infty,\ n\,\sqrt{\varphi(h)}\,\max\left(h^{4\beta}, \frac{1}{\log^2 n}\right) \to 0 \ \text{ and } \ \frac{(\log n)^2}{n\,\varphi(h)} < \psi_S\left(\frac{\log n}{n}\right) < \frac{n\,\varphi(h)}{\log n}$$

when n tends to infinity.

- Under H_0

Theorem 1.3. *If (H1)-(H6), (3) and (5) hold, then we get*

$$n\,\sqrt{\varphi(h)}\,W_n \ \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \ as\ n \to \infty$$

Where $s^2 = 2\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)\mathbb{E}\left[f(X)V^2\left[\epsilon_2^2|X\right]\right]$.

Moreover,

$$T_n = n\,\sqrt{\varphi(h)}\,\frac{W_n}{\widehat{s}} \ \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \ as\ n \to \infty$$

where

$$\widehat{s}^2 = \frac{1}{n(n-1)\varphi(h)}\sum_{i=1}^{n}\sum_{j\neq i,=1}^{n} K\left(d(X_i, X_j)h^{-1}\right)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)^2(\widehat{\epsilon_i}^2 - \widehat{\sigma}^2)^2.$$

- Under H_1

Theorem 1.4. *When (H1)-(H6), (3) and (5) hold, then we get*

$$\frac{T_n}{n\,\sqrt{\varphi(h)}} \longrightarrow \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 p(X)]/s_1, \quad \textit{In probability}$$

$$\textit{where } s_1^2 = \frac{\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)}{\left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)}\mathbb{E}[(V[\epsilon^2|X] + (V[\epsilon|X] - \sigma^2)^2)p(X)].$$

Now, we examine the robustness of the test and introduce the following sequence of local alternatives

$$H_{1n}: \quad V[\epsilon|x] - \sigma^2 = \delta_n g(x)$$

So, we obtain the following Corollary

Corollary 1.2. *Given (H1)-(H4), (3) and (5) we get, under H_{1n} with $\delta_n = n^{-1/2}\phi^{-1/4}(h)$*

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, 1) \text{ as } n \to \infty$$

where $\mu = \left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)\mathbb{E}[g^2(X)p(X)]/s.$

Bibliography

[1] Antoniadis, A. et Sapatinas, T. (2007) Estimation and inference in functional mixed-effect models Comp. Stat. and Data Anal. 51 (10) 4793-4813.

[2] Breusch, T. S. and Pagan, A. R.(1979) A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1294.

[3] Bickel, P. L. (1978). Using residuals robustly. I. Tests for heteroscedasticity and nonlinearity. Annals of Statistics, 6, 266-291

[4] Buell. CE. (1971). Integral equation representation for factor analysis. Journal of the Atmospheric Sciences.

[5] Boomsma. DI , PCM. Molenaar. (1987). The genetic analysis of repeated measures. I. Simplex models. Behavior genetics. 17 111â123.

[6] Borggaard. C, HH. Thodberg. (1992). Optimal minimal neural interpretation of spectra. Analytical chemistry. 64 545â551.

[7] Cardot, H., Goia, A. et Sarda, P. (2004) Testing for no effect in functional linear regression models, some computational approaches. *Comm. Statist. Simulation Comput.*, **33** , 179-199.

[8] Cardot, H., Crambes, C., Kneip, A. and Sarda, P (2007) Smoothing splines estimators in functional linear regression with errors-in-variables. Computational Statistics and Data Analysis, special issue on functional data analysis 51 (10) 4832-4848.

[9] Chiou, J.M. and Müller H.-G. (2007) Diagnostics for functional regression via residual processes Computational Statistics and Data Analysis 51, (10) 4849- 4863.

[10] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference,* **147**, 1-23.

[11] Cuevas, A., Febrero, M., and Fraiman, R. (2004). An anova test for functional 575 data. *Computational statistics and data analysis*, **47** , 111-122.

[12] Cook, R. D and Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression, Biometrika, 70, 1-10.

[13] Crainiceanu, Ciprian M., A.M Staicu ,S. Ray , N. Punjabi (2012) Bootstrap-based inference on the difference in the means of two correlated functional processes. 31 3223-3240.

[14] Dette, H. and Munk, A. (1998). Testing heteroscedasticity in nonparametric regression. *Journal of the Royal Statistical Society,* Series B,**60**, 693-708.

[15] Dette, H., Neumeyer, N. and Van Keilegom, I. (2007). A new test for the parametric form of the variance function in nonparametric regression, Journal of the Royal Statistical Society: Series B, 69(5), 903-917.

[16] Dabo-Niang, S. (2002). On Density estimation in an infinite-dimensional space : application to diffusion processes] PhD Paris VI.

[17] Dabo-Niang, S. and Rhomari, N. (2003) Kernel regression estimation when the regressor takes values in metric space, C. R. Math. Acad. Sci. Paris 336 (1) 75-80.

[18] Dette, H.(2002). A consistent test for heteroscedasticity in nonparametric regression based on the kernel method, Journal of Statistical Planning and Inference, 103, 311-329.

[19] Deville. JC. (1974). Méthodes statistiques et numériques de l'analyse harmonique. Annales de l'INSEE. 3, 5-101.

[20] Eubank, R. L. and Thomas, W.(1993). Detecting heteroscedasticity in non-parametric regression, Journal of the Royal Statistical Society: Series B, 55(1),145-155. bibitem22 Frank. I.E. et Friedman. J.H. (1993). A statistical view of some chemometrics regression tools. Technometrics 35 109-135.

[21] Fan. J, Lin, SK,(1998). Test of significance when data are curves, Journal of the American Statistical Association 93 1007-1021.

[22] Ferraty, F., Mas, A. and P. Vieu (2007). Nonparametric regression on functional data: inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**, 267-286.

[23] Ferraty. F.(2011). Recent advances in functional data analysis and related topics. book

[24] Ferraty F. et Vieu P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. Comptes Rendus de l'Académie des Sciences- Series I-Mathematics, 330(2) :139â142.

[25] Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statistic., 16, 111-127.

[26] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer-Verlag, New York, 2006.

[27] Gasser.T, Hall.P, Presnell. B. (1998). Nonparametric estimation of the mode of a distribution of random curves. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 60(4) :681â691.

[28] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *J. Mutiv. Anal.*, **146**, 1-6.

[29] Gadiaga. D, R. Ignaccolo. (2005) Test of No-effect Hypothesis. Afrika Statistika, 1 67-76.

[30] Grenander. U.(1950). Stochastic processes and statistical inference.Arkiv för matematik. 1 195-277.

[31] Goldfeld. S. and R.Quandt. (1965). Some Tests for Homoscedasticity. Journal of the American Statistical Association. 60 539-547

[32] Glejser. H.(1969). A New Test for Heteroskedasticity. Journal of the American Statistical Association. 64 316-323.

[33] Harvey. A.C., G.D.A. Phillips. (1974). A comparison of the power of some tests for heteroskedasticity in the general linear model. Journal of Econometrics. 2 307-316.

[34] Harvey.A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. Econometrica. 44 461-465.

[35] Hastie, T. et Mallows, C. (1993) Discussion of A statistical view of some chemometrics regression tools. by Frank, I.E. and Friedman, J.H. Technometrics 35 140-143.

[36] Hall, P. et Vial, C. (2006a) Assessing extrema of empirical principal component functions. Ann. Statist. 34 (3) 1518-1544.

[37] Hotelling. H, MR. Pabst. (1936). Rank correlation and tests of significance involving no assumption of normality. The Annals of Mathematical Statistics. 7 29-43.

[38] James. G.M. et Sood, A. (2006) Performing hypothesis tests on the shape of functional data. (English summary) Comp. Stat. and Data Anal. 50 (7) 1774-1792.

[39] Kruskal. W.H, Wallis. W.A.(1952).Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association. 47 583-621.

[40] Kirkpatrick. M, N. Heckman. (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. Journal of mathematical biology. 27 429â450.

[41] Lin, J. G. and Qu, X. Y. (2012). A consistent test for heteroscedasticity in semiparametric regression with nonparametric variance function based on the kernel method, *Statistics*, **46**, 565-576.

[42])Müller, H. G. and Zhao, P.L. (1995). On a semi-parametric variance function model and a test for heteroscedasticity, Annals of Statistics, 23, 946-967.

[43] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.

[44] Mas. A. (2007a) Testing for the mean of random curves : a penalization approach Statistical Inference for Stochastic Processes 10 (2) 147-163.

[45] Mas, A. (2000) Estimation d'opérateurs de corrélation de processus fonctionnels : lois limites, tests, déviations modérées Thése de doctorat de l'université de Paris 6.

[46] Mann. H. B, D. R. Whitney. (1947). on a test whether one of two random variables is stochastically larger than the other .The Annals of Mathematical Statistics. 18 50-60.

[47] Ramsay, J. and Dalzell, C. (1991) Some tools for functional data analysis J. R. Statist. Soc. B. 53 539-572.

[48] Ramsey. J. B.(1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. Journal of the Royal Statistical Society. Series B (Methodological). 31 350-371.

[49] Rao. CR.(1958). Some statistical methods for comparison of growth curves. Biometrics. 14 1-17.

[50] Rutemiller. H C, David A. Bowers. (1968). Estimation in a Heteroscedastic Regression Model. Journal of the American Statistical Association. 63 552-557.

[51] Shen, Q. et Faraway, J. (2004) An F test for linear models with functional responses. (English summary) Statist. Sinica 14 (4) 1239-1257.

[52] Siegel. S. (1956). Nonparametric statistics. for the behavioral sciences. book

[53] Savage, I.R. (1962). Bibliography of nonparametric statistics. Book.

[54] Staicu. AM, Y. Li, C.M. Crainiceanu , D. Ruppert.(2014) Likelihood Ratio Tests for Dependent Data with Applications to Longitudinal and Functional Data Analysis. 41 932-949.

[55] Stute, W. and Zhu, L. X. (2005). Nnonparametric checks for single-index models, Annals of Statistics, 33, 1048-1083.

[56] Simonoff, J. S. and Tsai, C. L. (1994). Improved tests for nonconstant variance in regression based on the modified profile likelihood, Journal of Appmied Statistics, 43, 357-370.

[57] Tsai, C. L. (1986). Score test for the first-order autoregressive with heterosedasticity, Biometrika, 73, 61-70.

[58] Tucker. LR.(1958). Determination of parameters of a functional relation by factor analysis. Psychometrika. 23 19â23.

[59] Viele, K. (2001) Evaluating fit in functional data analysis using model embeddings. (English, French summary) Canad. J. Statist. 29 (1) 51-66.

[60] White. H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica. 48 817-838.

[61] Wilcoxon. F. (1945). Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1 80-83.

[62] You, J. H. and Chen, G. M. (2005). Testing heteroscedasticity in partially linear regression models, Statistics and Probability Letters, 73, 61-70.

[63] Zheng, J. X. (2009). A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics, **75**, 263-289.

[64] Zhu, L. X., Fujikoshi, Y. and Naito, K. (2001). Heteroscedasticity checks for regression models, Science in China (Series A), 44, 1236-1252.

[65] Zhang, J.-T. et Chen, J. (2007) Statistical inferences for functional data. Ann. Statist. 35 (3) 1052-1079.

[66] Zhang, J.-T. et Chen, J. (2010) Statistical inferences for functional data. Ann. Statist. 35 (3) 1052-1079.

[67] Zheng, J. X. (2009). Testing heteroscedasticity in nonlinear and nonparametric regressions, Canadian Journal of Statistics, **37**, 282-300.

Chapter 2: Heteroscedasticity test when the covariables are curves

The contents of the following Chapter is the one of the recent paper ***Henien, Laksaci et al.*** [2018][1] accepted for publication. Indeed, in this Chapter we are interested in the **heteroscedasticity** test in nonparametric regression on functional variable. We firstly present our functional framework in Section 2, thence, we construct the test statistic in section 3. The asymptotic behavior of this test is studied in Section 4. In practice, some simulated data examples are reported in Section 5.

[1]The corresponding authors:
Aicha Henien : Laboratory of Statistic and Process Stochastic, University Djillali Liabes, Sidi Bel Abbes.
Larbi Ait-Hennani :University Lille 2, Law and Health, IUT C, Roubaix, France.
Jacques Demongeot : University Grenoble Alps, laboratory AGEIS EA 7407, France.
Ali Laksaci : Department of Mathematics, College of Science, King Khalid University, Abha, Saudi Arabia.
Mustapha Rachdid : University Grenoble Alps, laboratory AGEIS EA 7407, France.

Contents

Abstract

We present in this paper a consistent nonparametric test for heteroscedasticity when data are of functional kind. The latter is constructed by evaluating the difference between the conditional and unconditional variances. We show the asymptotic normality of the statistical test under the null hypothesis. In addition, we prove that this test is consistent against all deviations from homoscedasticity condition.

2.1 Introduction

The prediction of scalar response given explanatory functional random variable is an important subject in the modern statistics. The regression operator is the most preferred model in this prediction problem. However, this model is not efficiency in the heteroscedasticity case, it requires the homoscedasticity of the data which cannot be guaranteed a priory. In this paper we shall construct a test statistic to detect the heteroscedasticity for functional data.

Checking the heteroscedasticity of the data has received lot of attention in in the literature of multivariate statistics. Several authors have proposed parametric or nonparametric methods to testing this phenomenon in vectorial statistics. We return to ***Breusch and Pagan***[1979], ***Koenker and Bassett***[1982] or Diblasi and ***Bowman***[1997] (among others) for some parametric procedures and ***Dette and Munk***[1998] , ***Zheng***[2009] or ***Lin et al.***[2012] (and the references therein) for some nonparametric approaches. However, much less attention has been paid to Testing heteroscedasticity in functional statistics. To the best of our knowledge this problem has not been addressed so far.

Noting that statistical inference for Functional Data Analysis (FDA) has been the focus of several investigations (***Ramsay and Silverman***[2005], ***Horváth and Kokoszka***[2012], ***Zhang***[2013], ***Bongiorno et al.***[2014] and ***Hsing and Eubank***[2015], for recent advances). In this context, it is well known that the nonparametric methods is more appropriate than the parametric approaches, because the graphical tool for exploring the relationship between the explanatory variables and the scalar response is not available, and hence it becomes very hard to have some informations on the shape of the relationship between the functional variable and the scalar response. Recently, the nonparametric modeling takes a large place in the FDA literature (see ***Ferraty and Vieu***[2006] for basic results or ***Cuevas***[2014] and ***Goia and Vieu***[2016] for recent advanced and references). In particular, there is an extensive literature on the nonparametric estimation of the regression function. We cite for instance ***Masry***[2005] for the asymptotic normality, ***Ferraty et al.***[2007] for the L_2 consistency, ***Ferraty, Laksaci et al.***[2010] for the uniform almost complete convergence. More recently ***Kara, Laksaci et al.***[2015] have established the strong convergence (with rates), uniformly in bandwidth parameters, of the kernel estimator of the regression operator. On the other hand testing hypotheses in functional data has been widely developed by considering various test problems in both situations parametric or nonparametric structures (see, ***Cardot et al.*** [2003], ***Cuevas et al.*** [2004], ***Delsol*** [2013] , ***Zhang et al.*** [2010], ***Hilgert et al.***[2013], ***Staicu et al.*** [2014] to cite a few).

In this paper we propose a consistent nonparametric test for heteroscedasticity, based on kernel estimate of the nonparametric regression. We establish the asymptotic normality of the construct test statistic under the null hypothesis of homoscedasticity and can detect local alternatives distinct from the null. Thus, this work can be considered as generalization to infinite dimensional of the results of ***Zheng*** [2009] in the multivariate case. Moreover, this generalization is obtained under some standard conditions in nonparametric functional statistics allowing to avoid the problem of the curse of dimensionality in multivariate case. It should be noted that, the intrinsic dimensionality of these data poses challenges both for theory and computation, but the infinite dimensional structure of the data is an interesting source of information, which brings many opportunities for all statistical analysis. Finally let us point out that as in the vectorial case the heteroscedasticity checking is an important preliminary step before making the regression analysis accurate and efficient for functional data. Finally let us point out that these question of infinite.

This paper is organized as follows. We present our functional framework in the following Section. We construct the test statistic in Section 3. The asymptotic behavior of this test is studied in Section4. Some simulated data examples are reported in Section 5. All proofs are put into the Appendix.

2.2 Functional data framework

Let (X_i, Y_i) for $i = 1, \dots, n$, be a sample of independent and identically distributed pairs as (X, Y) which is a random vector valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space. In the following d is a semi-metric on $\mathcal{F}$, x is a fixed point in $\mathcal{F}$, N_x is a fixed neighborhood of x and the closed ball centered at x and of radius a is denoted

$$B(x, a) = \{y \in \mathcal{F} \text{ such that } d(y, x) \leq a\}.$$

Furthermore, we assume that X and Y are connected by the following relation

$$Y = r(X) + \epsilon, \tag{1}$$

where r is an operator from $\mathcal{F}$ to $\mathcal{R}$ and ϵ is a random error variable such that $\mathbb{E}[\epsilon|X] = 0$. Thereafter, we suppose that the operator r is belong to a functional space characterized the by following regularity condition

$$\forall x_1, x_2 \in N_x, \quad |r(x_1) - r(x_2)| \leq C d^{\beta}(x_1, x_2), \quad C > 0 \quad \beta > 0. \tag{2}$$

It is well documented that all the asymptotic analysis in nonparametric statistics for functional variables is closely related to the concentration properties of the probability measure of the explanatory random variable X. The same thing of testing hypotheses we suppose that the regressor X such that

There exists a nonnegative continuous functions ϕ and f such that

$$\mathbb{P}(X \in B(x, a)) = \phi(a).f(x) + o(\phi(a)). \tag{3}$$

We point out that this version of the small ball probability function is standard in this context of nonparametric functional data analysis. Precisely, this condition has been introduced by Masry (2005) and since has widely been used thereafter. Typically, such decomposition of the small ball probability function is verified of several usual case. In particular the function f and ϕ can be expressed through the Onsager-Machlup function (see, Ferraty, Laksaci et al. (2010) for more discussion in this question). In this contribution we suppose that the function ϕ such that:

$$\text{For all } s \in [0,1], \qquad \lim_{a\to 0} \frac{\phi(sa)}{\phi(a)} = \tau(s) \tag{4}$$

The function express the variation of ϕ. The latter has been explicitly expressed for all the usual cases by Ferraty et al. (2007). Concerning the function f, we assume that the Kolmogorov's ϵ-entropy [2] $\psi_{\mathcal{S}}$ of the support $\mathcal{S}$ of f such that

$$\sum_{n=1}^{\infty} \exp\left\{(1-\eta)\psi_{\mathcal{S}}\left(\frac{\log n}{n}\right)\right\} < \infty, \qquad \text{for some } \eta > 1. \tag{5}$$

The same thing here this consideration is classic in nonparametric functional data analysis, especially, to state the uniform consistency. This assumption was introduced by ***Ferraty et al.*** [2010] to establish the uniform almost complete consistency of kernel estimate of the regression function. We refer to this cited work of some examples of $\mathcal{S}$ and $\mathcal{F}$ for which $\psi_{\mathcal{S}}$ is explicitly given.

Finally, it worth to noting all these functions (ϕ, $\psi_{\mathcal{S}}$, τ) are closely linked to the topological structure of the functional space , in sense that all these functions increase or decrease with the choice of the semi-metric d. Thus we will see thereafter that the choice of the semi-metric has an important influence on the power as well as the robustness of the test.

[2] Let $\epsilon > 0$ be given. A finite set of points $x_1, x_2, \ldots, x_N$ in $\mathcal{F}$ is called an ϵ-net for $\mathcal{S}$ if $S \subset \bigcup_{k=1}^{N} B(x_k, \epsilon)$. The quantity $\psi_{\mathcal{S}}(\epsilon) = \log(N_\epsilon(S))$, where $N_\epsilon(S)$ is the minimal number of open balls in $\mathcal{F}$ of radius ϵ which is necessary to cover S, is called the Kolmogorov's ϵ-entropy of the set S.

2.3 Construction of the test statistic

Recall that our main aim is to test the heteroscedasticity of the model (1). Typically, we test

$$H_0 : \quad V[\epsilon|X] = \sigma^2$$

versus

$$H_1 : \quad V[\epsilon|X] \neq \sigma^2.$$

In order to construct test statistic, we suppose that the function f such that

$$f(X) > 0 \quad \text{almost surly,} \quad \text{and} \quad \mathbb{E}[f(X)] < \infty. \tag{6}$$

The latter allows to show that H_0 is equivalent to write

$$S_1 := \mathbb{E}[(\epsilon^2 - \sigma^2)\mathbb{E}[(\epsilon^2 - \sigma^2)|X]f(X)] = \mathbb{E}[\mathbb{E}^2[(\epsilon^2 - \sigma^2)|X]f(X)] = 0$$

whereas H_1 is equivalent to

$$S_1 = \mathbb{E}[(\epsilon^2 - \sigma^2)\mathbb{E}[(\epsilon^2 - \sigma^2)|X]f(X)] = \mathbb{E}[\mathbb{E}^2[(\epsilon^2 - \sigma^2)|X]f(X)] = \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 f(X)] > 0$$

Thus, our test's problem is equivalent to test

$$S_1 = 0 \quad \text{vs} \quad S_1 > 0$$

Combing the ideas of ***Zheng*** [2009] to those of ***Ferraty and Vieu*** [2006] to construct a nonparametric estimate S_1. Indeed, we put $\Delta_i = \mathbb{E}[(\epsilon_i^2 - \sigma^2)|X_i]f(X_i)$ and we consider the empirical version of S_1 denoted by

$$\widehat{S_1} = \frac{1}{n}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2)\Delta_i.$$

According to ***Ezzahrioui and Ould-Saïd***[2008] the function f can be estimates by

$$\widehat{f}(x) = \frac{1}{n\phi(h)}\sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{h}\right).$$

It follows that the natural estimate of S_1 is

$$W_n = \frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} (\widehat{\epsilon_i}^2 - \widehat{\sigma}^2) K\left(\frac{d(X_i, X_j)}{h}\right)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)$$

where $\widehat{\epsilon_i}$ and $\widehat{\sigma}^2$ are estimators of ϵ_i and σ respectively. These estimators are defined by

$$\widehat{\epsilon_i} = Y_i - \widehat{r}(X_i) \quad \text{and } \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \widehat{\epsilon_i}^2$$

with

$$\widehat{r}(x) = \frac{\sum_{i==1}^{n} K\left(\frac{d(x, X_i)}{h}\right) Y_i}{\sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{h}\right)}.$$

Now, we establish the asymptotic distribution of this statistic. To do that, we need the following assumptions

(H1) There exists $m \geq 2$ such that $\mathbb{E}[Y^m|X = x] < \delta_m(x) < C < \infty$ with $\delta_m(.)$ continuous on $\mathcal{S}$.

(H2) $\mathbb{E}[\epsilon^8|X = x] \leq b(x)$ with $b(x)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[b^2(X)] < \infty$.

(H3) The kernel K is a differentiable function supported on $[0, 1]$ such that

$$K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds > 0 \quad \text{and} \quad K(1) - \int_0^1 (K(s))'\tau(s)ds \neq 0.$$

(H4) The bandwidth parameter $h := h(n)$ is strictly positive such that:

$$n \to 0, n\,\phi(h) \to \infty,\ n\sqrt{\phi(h)} \max\left(h^{4\beta}, \frac{1}{\log^2 n}\right) \to 0 \text{ and } \frac{(\log n)^2}{n\,\phi(h)} < \psi_S\left(\frac{\log n}{n}\right) < \frac{n\,\phi(h)}{\log n}$$

when n tends to infinity.

All these assumptions are standard in this context, because they are the same as those classically used in the heteroscedasticity or in the nonparametric functional data analysis. Specifically (H1), (H3), (H4) are the same as those used by ***Ferraty, laksaci et al.*** [2010] for the uniform consistency in functional statistics while (H2) is the same as in ***Zheng*** [2009] for the heteroscedasticity in multivariate case.

2.4 The main result

We state the following results:

- Under H_0

Theorem 2.5. *When (H1)-(H4) and (2)-(6) hold we have*

$$n\sqrt{\phi(h)}\,W_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \text{ as } n \to \infty$$

where $s^2 = 2\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)\mathbb{E}\left[f(X)V^2\left[\epsilon_2^2|X\right]\right]$.

Moreover,

$$T_n = n\sqrt{\phi(h)}\,\frac{W_n}{\widehat{s}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \to \infty$$

where

$$\widehat{s}^2 = \frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i,=1}^{n} K\left(\frac{d(X_i, X_j)}{h}\right)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)^2(\widehat{\epsilon_i}^2 - \widehat{\sigma}^2)^2.$$

- Under H_1

Theorem 2.6. *When (H1)-(H4) and (2)-(6) hold we have*

$$\frac{T_n}{n\sqrt{\phi(h)}} \longrightarrow \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 f(X)]/s_1, \quad \text{In probability}$$

$$\text{where } s_1^2 = \frac{\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)}{\left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)}\mathbb{E}[(V[\epsilon^2|X] + (V[\epsilon|X] - \sigma^2)^2)f(X)].$$

Now, we examine the robustness of the test against all possible departures from homoscedasticity. To do that we introduce the following sequence of local alternatives

$$H_{1n}: \quad V[\epsilon|x] - \sigma^2 = \delta_n g(x)$$

where the known function $g(\cdot)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[g^2(X)] < \infty$. So, we obtain the following Corollary

Corollary 2.3. *Given (H1)-(H4) and (2)-(6), we have, under* H_{1n} *with* $\delta_n = n^{-1/2}\phi^{-1/4}(h)$

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, 1) \text{ as } n \to \infty$$

where $\mu = \left(K(1) - \int_0^1 (K(s))' \tau(s) ds\right) \mathbb{E}[g^2(X) f(X)]/s$.

PROOF. The detail of the proof is given in appendix. It is based on the same decomposition of ***Zheng*** [2009] and the following preliminary results.

Lemma 2.4.1. *(see,* ***Ferraty, laksaci et al.*** *[2010])*

Under the hypotheses (H1), (H3), (H4) (2), (3) and (5), we have

$$\sup_{x \in \mathcal{S}} |\widehat{f}(x) - f(x)| = o(1) \quad \textit{Almost completly} \tag{7}$$

and

$$\sup_{x \in \mathcal{S}} |\widehat{r}(x) - r(x)| = O(h^\beta) + O\left(\sqrt{\frac{\psi_{\mathcal{S}}\left(\frac{\log n}{n}\right)}{n\phi(h)}}\right) \quad \textit{Almost completely.} \tag{8}$$

Lemma 2.4.2. *(see,* ***Zheng*** *[1996])*

Let U_n *a second-order* U*-statistic*[3] *of kernel* $H_n(\cdot,\cdot)$ *such that* $\mathbb{E}[H_n^2(Z_i, Z_j)] = o(n)$*, then*

$$U_n - \mathbb{E}[H_n(Z_i, Z_j)] = o(1) \quad \textit{In probability.}$$

Lemma 2.4.3. *(see,* ***Zheng*** *[1996])*

Assume that $\mathbb{E}[H_n(Z_1, Z_2)|Z_1] = 0$ *and* $\mathbb{E}[H_n^2(Z_1, Z_2)] < \infty$ *for each* n*. If*

$$\frac{\mathbb{E}[G_n^2(Z_1, Z_2)] + n^{-1}\mathbb{E}[H_n^4(Z_1, Z_2)]}{\mathbb{E}^2[H_n^2(Z_1, Z_2)]} \to 0$$

where $G_n(Z_1, Z_2) = \mathbb{E}\left[H_n(Z_1, Z_3) H_n(Z_3, Z_2)|Z_1, Z_2\right]$*, then*

$$\frac{nU_n}{2\mathbb{E}^{1/2}[H_n^2(Z_1, Z_2)]} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \text{ as } n \to \infty.$$

[3]The general second-order U-statistic is of the form

$$U_n = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i, =1}^{n} H_n(Z_i, Z_j)$$

where (Z_i) are i.i.d. sample and H_n such that $H_n(Z_i, Z_j) = H_n(Z_j, Z_i)$. Further, a U-statistic is said to be degenerate if $\mathbb{E}[H_n(Z_i, Z_j)|Z_i] = 0$ almost surly for all $i \neq j$

2.5 On the finite sample performance of the test

The main goal of this Section is to show how we can implement easily and rapidly our test statistic in practice. For this purpose, we consider the following regression model

$$Y = r(X) + \sigma(X)\varepsilon$$

where the explanatory curves are defined by :

$$X_i(t) = a_i \sin(4(b_i - t)) + b_i + \eta_{i,t}\,, \quad \forall t \in (0,1) \text{ and } i = 1, 2, \ldots, n$$

where b_i (respectively, $\eta_{i,t}$) is distributed as $\mathcal{N}(0,3)$, (respectively, $\mathcal{N}(0,0.5)$), while the n random variables a_i's are generated according to a $\mathcal{N}(4,3)$ distribution. All the curves X_i's are discretized on the same grid generated from 100 equispaced measurements in $(0,1)$ (cf. Figure 2.1).

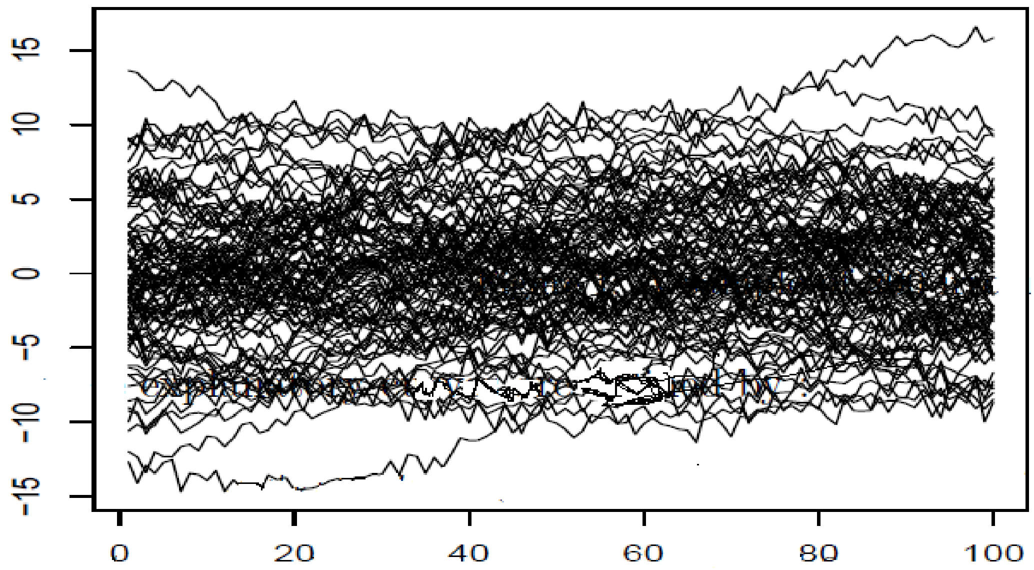

Figure 2.1: *A sample of* 200 *irregular curves*

The operator r, is defined by:

$$r(X_i) = \int_0^1 \frac{dt}{1 + X_i^2(t)} \text{ for } i = 1, \ldots, n$$ 9

In order to highlight the consistency of the proposed test statistic we compare the p-values of this test with various degree of heteroscedasticity. Specifically we consider four operators $\sigma(\cdot)$

Model 1	$\sigma(X) \equiv 1$	Homoscedasticity case
Model 2	$\sigma(X) = 4 * exp\left(\sqrt{\int_0^1 X^2(t)dt}\right)$	Exponential case
Model 3	$\sigma(X) = 2 * cos\left(\pi\sqrt{\int_0^1 X^2(t)dt}\right)$	Consensus case
Model 4	$\sigma(X) = \left(1 + \sqrt{\int_0^1 X^2(t)dt}\right)^2$	Polynomial case

The nonparametric structure of the test T_n is an important feature in this context of functional data where it is difficult to detect the form of the relationship between the variable of interest and the regressor. Moreover, it is well known that the nonparametric test has more advantage than parametric one which is not consistent against all deviation from the null hypothesis (see, ***Zheng*** [1996] for more discussion in this subject for the multivariate case). On the other hand, it is shown in this last work that this type of test statistic is simpler to compute than Bierens's test, because it contains less parameters. In the present test the principal parameter is the bandwidth h. Now, to emphasize the importance of this parameter in the efficiency of this test we control the sensitivity of our approach to the smoothing parameter choice. Specifically, we compare the two selectors methods proposed by ***Benhenni et al.*** [2010] of the regression function (the local selection method and the global one).

For practical purposes, we select the local optimal bandwidth $h_{\mathcal{X}}$ obtained by the following local cross-validation criterion

$$h_{\mathcal{X}} = \arg\min_{h\in H_n(\mathcal{X})} LCV_{\mathcal{X}} \quad where \quad LCV_{\mathcal{X}} = n^{-1}\sum_{j=1}^{n}(Y_j - \widehat{r}(X_j))W_{\mathcal{X}}(X_j)$$

with

$$W_{\mathcal{X}}(t) = \begin{cases} 1 & \text{if } d(t,\mathcal{X}) < a(\mathcal{X}) \\ 0 & \text{otherwise} \end{cases}$$

and $H_n(\mathcal{X})$ the set for $a(\mathcal{X})$ such that the ball centered at $\mathcal{X}$ with radius $a(\mathcal{X})$ contains exactly k neighbors of x. Concerning the global choice we take the same smoothing parameter for all observations that is

$$h = \arg\min_{h \in H_n} GCV \quad where \quad GCV = n^{-1} \sum_{j=1}^{n} (Y_j - \widehat{r}(X_j)) \tag{10}$$

with H_n is the subset is the quantiles of order q of the vector of all distances between the curves.

We use the following algorithm:

- Step 1: For each curve X_i we calculate the estimate ϵ_i^{Loc} and $\widehat{\sigma^{Loc}}$ by using the bandwidth parameter h_{X_i}.

- Step 3: We calculate the estimate ϵ_i^{Glob} and $\widehat{\sigma^{Glob}}$ by using the global bandwidth parameter h of (10).

- Step 4: We calculate the test statistic for both case

$$T_n^{Loc} = \sqrt{\frac{n}{2(n-1)}} \frac{\sum_{i=1}^n \sum_{j \neq i,=1}^n K\left(\frac{d(X_i,X_j)}{h_{X_j}}\right)(\widehat{\epsilon_j^{Loc}}^2 - \widehat{\sigma^{Loc}}^2)(\widehat{\epsilon_i^{Loc}}^2 - \widehat{\sigma^{Loc}})}{\sqrt{\sum_{i=1}^n \sum_{j \neq i,=1}^n K^2\left(\frac{d(X_i,X_j)}{h_{X_j}}\right)(\widehat{\epsilon_j^{Loc}}^2 - \widehat{\sigma^{Loc}}^2)^2(\widehat{\epsilon_i^{Loc}}^2 - \widehat{\sigma^{Loc}}^2)^2}}$$

and

$$T_n^{Glob} = \sqrt{\frac{n}{2(n-1)}} \frac{\sum_{i=1}^n \sum_{j \neq i,=1}^n K\left(\frac{d(X_i,X_j)}{h}\right)(\widehat{\epsilon_j^{Glob}}^2 - \widehat{\sigma^{Glob}}^2)(\widehat{\epsilon_i^{Glob}}^2 - \widehat{\sigma^{Glob}}^2)}{\sqrt{\sum_{i=1}^n \sum_{j \neq i,=1}^n K^2\left(\frac{d(X_i,X_j)}{h}\right)(\widehat{\epsilon_j^{Glob}}^2 - \widehat{\sigma^{Glob}}^2)^2(\widehat{\epsilon_i^{Glob}}^2 - \widehat{\sigma^{Glob}}^2)^2}}$$

- Step 5: We determine the P-values of T_n^{Loc} and T_n^{Glob}

In this illustration study we have used the semi-metric based on the m first eigenfunctions of the empirical covariance operator associated with the $m = 3$ greatest eigenvalues (cf. ***Benhenni et al.*** [2007] for more discussions in this choice) and the quadratic kernel. The results of the four selected sample sizes, $n \in \{50, 100, 200, 500\}$, are gathered in the following table

Model	n	Loc. Selector	Glob. Selector
Model 1	50	0.09	0.07
	100	0.26	0.28
	200	0.30	0.28
	500	0.33	0.32
Model 2	50	0.1	0.16
	100	0.07	0.09
	200	0.01	0.03
	500	0.01	0.01
Model 3	50	0.2	0.22
	100	0.12	0.18
	200	0.04	0.08
	500	0.02	0.06
Model 4	50	0.15	0.18
	100	0.09	0.11
	200	0.05	0.07
	500	0.02	0.04

Table 1: The P-Values of the test.

Table 1 presents the P-values of T_n for various sample sizes $n = 50, 100, 200, 500$. We can see clearly that the performance of our test is varied with the type of heteroscedasticity and the sample size n. However, the result of the local choice is better than the global one. Moreover, we point out that this approach is faster even if the sample sizes is large. It worth to noting that the result of the small size can be improved by using the bootstrapping approach. This question is an important prospect of this work.

2.6 Appendix

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. Moreover, we put, for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$:

$$K_{ij} = K(h^{-1}d(X_i, X_j)),\ R_i = r(X_i),\ \widehat{R}_i = \widehat{r}(X_i),\ \epsilon_i = Y_i - R_i$$

and

$$f_i = f(X_i),\ u_i = \epsilon_i^2 - \sigma^2, \widehat{f}_i = \widehat{f}(X_i).$$

Proof of Theorem 2.5

By writing

$$\widehat{\epsilon}_i^2 - \widehat{\sigma}^2 = u_i - 2\epsilon_i(\widehat{R}_i - R_i) + (\widehat{R}_i - R_i)^2 + (\sigma^2 - \widehat{\sigma}^2)$$

we have

$$\begin{aligned}
W_n \quad &= \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_iu_j\right\}}_{W_{1n}} + 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i\epsilon_j(\widehat{R}_i - R_i)(\widehat{R}_j - R_j)\right\}}_{W_{2n}} \\
&+ \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(\widehat{R}_i - R_i)^2(\widehat{R}_j - R_j)^2\right\}}_{W_{3n}} + \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(\widehat{\sigma}^2 - \sigma^2)^2\right\}}_{W_{4n}} \\
&- 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i\epsilon_j(\widehat{R}_j - R_j)\right\}}_{W_{5n}} + 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i(\widehat{R}_j - R_j)^2\right\}}_{W_{6n}} \\
&- 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i(\widehat{\sigma} - \sigma^2)\right\}}_{W_{7n}} - 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i(\widehat{R}_i - R_i)(\widehat{R}_j - R_j)^2\right\}}_{W_{8n}} \\
&+ 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i(\widehat{R}_i - R_i)(\widehat{\sigma}^2 - \sigma^2)\right\}}_{W_{9n}} - 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(\widehat{R}_i - R_i)^2(\widehat{\sigma}^2 - \sigma^2)\right\}}_{W_{10n}}
\end{aligned}$$

then W_n is decomposed into 10 terms W_{kn}, $k = 1, \ldots 10$

$$W_n = W_{1n} + 4W_{2n} + W_{3n} + W_{4n} - 4W_{5n} + 2W_{6n} - 2W_{7n} - 4W_{8n} + 4W_{9n} - 2W_{10n}.$$

This it suffices to prove that under H_0

$$n\sqrt{\phi(h)}\, W_{1n} \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \text{ as } n \to \infty \tag{11}$$

and

$$n\sqrt{\phi(h)}W_{2n} = o(1) \quad \text{In probability,} \quad n\sqrt{\phi(h)}W_{3n} = o(1) \quad \text{In probability,}$$
$$n\sqrt{\phi(h)}W_{4n} = o(1) \quad \text{In probability,} \quad n\sqrt{\phi(h)}W_{5n} = o(1) \quad \text{In probability,}$$
$$n\sqrt{\phi(h)}W_{6n} = o(1) \quad \text{In probability,} \quad n\sqrt{\phi(h)}W_{7n} = o(1) \quad \text{In probability,}$$
$$n\sqrt{\phi(h)}W_{8n} = o(1) \quad \text{In probability,} \quad n\sqrt{\phi(h)}W_{9n} = o(1) \quad \text{In probability,}$$

$$\text{and} \quad n\sqrt{\phi(h)}W_{10n} = o(1) \quad \text{In probability.}$$

Firstly, we show (11). To do that we rewrite W_{1n} as U-statistic of the following form

$$W_{1n} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} H_n(Z_i, Z_j)$$

where

$$H_n(Z_i, Z_j) = \frac{1}{\phi(h)} K_{ij} u_i u_j \quad \text{with} \quad Z_i = (X_i, \epsilon_i).$$

Employing properties of the conditional expectation to prove, under H_0, that

$$\mathbb{E}[H_n(Z_1, Z_2)|Z_1] = \frac{1}{\phi(h)} u_1 K_{12} \mathbb{E}\left[u_2|X_2\right] = 0.$$

So, we can apply Lemma2.4.3 on U_{1n} as degenerate U-statistic. For this purpose, we must evaluate the following quantities

$$\mathbb{E}\left[G_n^2(Z_1, Z_2)\right], \quad \mathbb{E}\left[H_n^4(Z_1, Z_2)\right] \quad \text{and} \quad \mathbb{E}\left[H_n^2(Z_1, Z_2)\right].$$

Concerning the first quantity we use (3) and (4) together with (H3) to write that

$$\begin{aligned}
\mathbb{E}\left[G_n^2(Z_1,Z_2)\right] &= \mathbb{E}\left[\mathbb{E}\left[H_n(Z_1,Z_3)H_n(Z_3,Z_2)|Z_1,Z_2\right]^2\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[u_1^2u_2^2\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[K_{13}K_{23}u_3^2|X_3\right]|X_1,X_2\right]^2\right]\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[u_1^2u_2^2\mathbb{E}\left[K_{13}K_{32}\mathbb{E}\left[u_3^2|X_3\right]|X_1,X_2\right]^2\right] \\
&= \frac{1}{\phi^4(h)}\int\int\left[\int K(h^{-1}d(x_1,x_3))K(h^{-1}d(x_3,x_2))\right. \\
&\qquad \left.\mathbb{E}\left[u_3^2|x_3\right]dP_{X_3}(x_3)\right]^2\mathbb{E}\left[u_1^2|X_1\right]\mathbb{E}\left[u_2^2|X_2\right]dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&= \frac{1}{\phi^4(h)}\iint_{D=\{x_1,x_2,\ d(x_1,x_2)\leq 2h\}}\left[\int_{B(x_1,h)\cap B(x_2,h)}K(h^{-1}d(x_1,x_3))K(h^{-1}d(x_3,x_2))\right. \\
&\qquad \left.\mathbb{E}\left[u_3^2|x_3\right]dP_{X_3}(x_3)\right]^2\mathbb{E}\left[u_1^2|X_1\right]\mathbb{E}\left[u_2^2|X_2\right]dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\leq \frac{C}{\phi^4(h)}\iint_{D=\{x_1,x_2,\ d(x_1,x_2)\leq 2h\}}\left[\int_{B(x_1,h)\cap B(x_2,h)}dP_{X_3}(x_3)\right]^2dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\leq \frac{C}{\phi(h)}.
\end{aligned}$$

We conclude that

$$\mathbb{E}\left[G_n^2(Z_1,Z_2)\right] = O\left(\frac{1}{\phi(h)}\right). \qquad (12)$$

Now, for the second quantity, we write

$$\begin{aligned}
\mathbb{E}\left[H^4(Z_1,Z_2)\right] &= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4u_1^4u_2^4\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4u_1^4|X_2,X_1\right]\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4|X_2\right]\mathbb{E}\left[u_1^4|X_1\right]\right] \\
&\leq \frac{C}{\phi^4(h)}\int\int_{B(x_2,h)}dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\leq \frac{C}{\phi^3(h)}\int f(x)dP_X(x).
\end{aligned}$$

Therefore

$$\mathbb{E}\left[H^4(Z_1,Z_2)\right] = O\left(\frac{1}{\phi^3(h)}\right). \qquad (13)$$

Next, for the last term we use the continuity of the operator $\mathbb{E}[\epsilon^4|X=\cdot]$ to get

$$\begin{aligned}
\mathbb{E}\left[H^2(Z_1,Z_2)\right] &= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2u_1^2u_2^2\right] \\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2^2u_1^2|X_2,X_1\right]\right] \\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2^2|X_2\right]\mathbb{E}\left[u_1^2|X_1\right]\right] \\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2^2|X_2\right]\right] + o\left(\frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2^2|X_2\right]\right]\right) \\
&= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\int f(x_2)\mathbb{E}^2\left[u_2^2|x_2\right]dP_{X_2}(x_2) + o\left(\frac{1}{\phi(h)}\right) \\
&= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\int f(x_2)\left(V\left[\epsilon^2|x_2\right]\right)^2 dP_{X_2}(x_2) + o\left(\frac{1}{\phi(h)}\right)
\end{aligned}$$

$$= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right]\right)^2\right] + o\left(\frac{1}{\phi(h)}\right). \quad (14)$$

It follows that

$$\mathbb{E}\left[H^2(Z_1,Z_2)\right] = O\left(\frac{1}{\phi(h)}\right).$$

Combining this last evaluation to (7) and (8) to obtained

$$\frac{\mathbb{E}\left[G_n^2(Z_1,Z_2)\right] + n^{-1}\mathbb{E}\left[H^4(Z_1,Z_2)\right]}{\mathbb{E}^2\left[H^2(Z_1,Z_2)\right]} = O(\phi(h)) + O\left(\frac{1}{n\phi(h)}\right) = o(1)$$

Hence, from Lemma 2.4.3 we have

$$\frac{nW_{1n}}{2\mathbb{E}^{1/2}\left[H^2(Z_1,Z_2)\right]} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \text{ as } n \to \infty$$

which complete the proof of (11).

Secondly, we treat $n\sqrt{\phi(h)}W_{2n}$. To do that, we write

$$W_{2n} = \underbrace{W_{2n}\mathbb{1}_{\{\prod_{i=1}^n f(X_i)\neq 0\}}}_{U_{2n}} + W_{2n}\mathbb{1}_{\{\prod_{i=1}^n f(X_i)=0\}} \quad (15)$$

where $1\!\!1_A$ is indicator function of A. For U_{2n}, we introduce the quantity

$$Q_{2n} = \left\{ \frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{h} K_{ij} \epsilon_i \epsilon_j (\widehat{R_i} - R_i)(\widehat{R_j} - R_j) \frac{\widehat{f_i} \widehat{f_j}}{f_i f_j} \right\}.$$

The uniform consistency of Lemma 2.4.1 allows to write that

$$U_{2n} = Q_{2n} + o\left(Q_{2n}\right).$$

So, it suffices to evaluate the leading term that is $n\sqrt{\phi(h)}Q_{2n}$. Indeed,

$$\begin{aligned} \mathbb{E}\left[Q_{2n}^2\right] \quad &= \mathbb{E}\Bigg[\frac{1}{n^6(n-1)^2\phi^6(h)} \sum_i \sum_{j \neq i} \sum_k \sum_l \sum_{i'} \sum_{j' \neq i'} \sum_{k'} \sum_{l'} \frac{1}{f_i f_j f_{i'} f_{j'}} \\ & K_{ij} K_{ik} K_{jl} K_{i'j'} K_{i'k'} K_{j'l'} (Y_k - R_i)(Y_l - R_j)(Y_{k'} - R_{i'})(Y_{l'} - R_{j'}) \epsilon_i \epsilon_j \epsilon_{i'} \epsilon_{j'} \Bigg]. \end{aligned}$$

Observe that the terms of this summation are non null only if

$$\{i, j\} = \{i', j'\}$$

or

$$\{i, j, i', j'\} = \{k, l, k', l'\} \quad \text{with} \quad i \neq k, j \neq l, \ i' \neq k', j' \neq l'.$$

It is clear that the first case is the leading part of this sum. Thus,

$$\begin{aligned} \mathbb{E}\left[Q_{2n}^2\right] &= \frac{2}{n^6(n-1)^2\phi^6(h)} n(n-1)(n-2)^2(n-3)^2 \mathbb{E}\Bigg[\frac{1}{f_1^2 f_2^2} K_{12}^2 K_{13} K_{24} K_{15} K_{26} \\ & (Y_3 - R_1)(Y_4 - R_2)(Y_5 - R_1)(Y_6 - R_2) \epsilon_1^2 \epsilon_2^2 \Bigg] + o\left((n^2\phi(h))^{-1}\right). \\ &= \frac{2}{n^6(n-1)^2\phi^6(h)} n(n-1)(n-2)^2(n-3)^2 \mathbb{E}\Bigg[\frac{1}{f_1^2 f_2^2} K_{12}^2 K_{13} K_{24} K_{15} K_{26} \\ & (R_3 - R_1)(R_4 - R_2)(R_5 - R_1)(R_6 - R_2) \sigma^4 \Bigg] + o\left((n^2\phi(h))^{-1}\right). \end{aligned}$$

From (2), we get

$$\mathbb{E}\left[Q_{2n}^2\right] \leq \frac{2Ch^{4\beta}}{n^6(n-1)^2\phi^6(h)} n(n-1)(n-2)^2(n-3)^2 \mathbb{E}\left[\frac{1}{f_1^2 f_2^2} K_{12}^2 K_{13} K_{24} K_{15} K_{26} \right]$$

Further, using the fact that K has compact support $(0,1)$ to write that

$$\mathbb{E}\left[\frac{1}{f_1^2 f_2^2}K_{12}^2 K_{13}K_{24}K_{15}K_{26}\right] = \int\int\int_0^1\int_0^1\int_0^1\int_0^1 \frac{1}{f^2(x_1)f^2(x_2)}K^2(h^{-1}d(x_1,x_2))$$
$$K(t_3)K(t_4)K(t_5)K(t_6)dP^{X_1}(x_1)dP^{X_2}(x_2)dP^{h^{-1}d(x_1,X_3)}(t_3)$$
$$dP^{h^{-1}d(x_2,X_4)}(t_4)dP^{h^{-1}d(x_1,X_5)}(t_5)dP^{h^{-1}d(x_2,X_6)}(t_6)$$
$$\leq C\int\int \frac{1}{f^2(x_1)f^2(x_2)}K^2(h^{-1}d(x_1,x_2))$$
$$\int_0^1 dP^{h^{-1}d(x_1,X_3)}(t_3)\int_0^1 dP^{h^{-1}d(x_2,X_4)}(t_4)$$
$$\int_0^1 dP^{h^{-1}d(x_1,X_5)}(t_5)\int_0^1 dP^{h^{-1}d(x_2,X_6)}(t_6)dP^{X_1}(x_1)dP^{X_2}(x_2)$$
$$\leq C\phi^5(h)\int f(x)dP^X(x) + o(\phi^5(h))$$

Finally, we obtain

$$\mathbb{E}\left[Q_{2n}^2\right] = O\left(\frac{h^{4\beta}}{n^2\phi(h)}\right).$$

Consequently, the Chebyshiev's inequality permits to conclude that

$$n\sqrt{\phi(h)}Q_{2n} \to 0 \quad \text{In probability.}$$

Hence

$$n\sqrt{\phi(h)}U_{2n} \to 0 \quad \text{In probability.}$$

About the second term of (15), we use (6) to write, for all $\varepsilon > 0$

$$P\left\{n\sqrt{\phi(h)}\left|W_{2n}1\!\!1_{\{\prod_{i=1}^n f(X_i)=0\}}\right| > \varepsilon\right\} \leq P\left\{\prod_{i=1}^n f(X_i) = 0\right\} \leq nP\{f(X)=0\} = 0. \qquad (16)$$

which finish the proof of the second limit.

Thirdly, for the term W_{3n} we use a similar decomposition of (15) to write that

$$W_{3n} = \underbrace{W_{3n}1\!\!1_{\{\prod_{i=1}^n f(X_i)\neq 0\}}}_{U_{3n}} + W_{3n}1\!\!1_{\{\prod_{i=1}^n f(X_i)=0\}}.$$

The second term is evaluated by the same fashion that in (16). While for the first term we have

$$U_{3n} \leq \sup_{x \in \mathcal{S}} |\widehat{R}(x) - R(x)|^4 \frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij}$$

Recall that, the second part of Lemma 2.4.1 gives

$$\sup_{x \in \mathcal{S}} |\widehat{R}(x) - R(x)| = O(h^{\beta}) + O\left(\sqrt{\frac{\psi(n^{-1} \log n)}{n\phi(h)}} \right) \quad \text{In probability.}$$

On the other hand by a simple manipulation we can show that

$$\frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij} = O(1) \quad \text{In probability.}$$

By (H4), we get

$$n\sqrt{\phi(h)}\, U_{3n} = O\left(n\sqrt{\phi(h)} \max\left(h^{4\beta}, \frac{1}{\log^2 n} \right) \right) = o(1).$$

Thus, we conclude the required limit.

Fourthly, the term W_{4n} is evaluated by the same arguments as those used in W_{3n}. Formally, it suffices to prove that

$$n\sqrt{\phi(h)}(\widehat{\sigma}^2 - \sigma^2)^2 = o(1) \quad \text{In probability.}$$

Indeed, we have

$$\widehat{\sigma}^2 - \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (\epsilon_i^2 - \sigma^2) + \frac{2}{n} \sum_{i=1}^{n} \epsilon_i (\widehat{R}_i - R_i) + \frac{1}{n} \sum_{i=1}^{n} (\widehat{R}_i - R_i)^2.$$

It is easy to see that

$$V\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\epsilon_i^2 - \sigma^2) \right] = V\left[\epsilon^2 \right], \qquad \mathbb{E}\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\epsilon_i^2 - \sigma^2) \right] = 0$$

and

$$V\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \right] = V\left[\epsilon \right], \qquad \mathbb{E}\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \right] = 0.$$

Thus

$$\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i^2-\sigma^2)=o\left(\frac{1}{\sqrt{n\sqrt{\phi(h)}}}\right)\quad \text{In probability}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(\widehat{R}_i-R_i)=o\left(\frac{1}{\sqrt{n\sqrt{\phi(h)}}}\right)\quad \text{In probability.}$$

Similarly to U_{3n}, we use the convergence rate of the uniform consistency of Lemma (2.4.1) to prove that

$$\sqrt{n\sqrt{\phi(h)}}\sup_{x\in\mathcal{S}}|\widehat{R}(x)-R(x)|=o(1)\quad \text{In probability.}$$

It follows that

$$\sqrt{n\sqrt{\phi(h)}}(\widehat{\sigma}^2-\sigma^2)=o(1)\quad \text{In probability.}$$

Hence

$$n\sqrt{\phi(h)}W_{4n}=o(1)\quad \text{In probability.}$$

Finally, the proofs of the others terms $W_{5n},\dots W_{10n}$ are very similar to the treated cases. So, their proofs are omitted. The proof of the first part is completed.

Now, we proceed to state the second part of Theorem 2.5. Once again we use the Slutsky lemma. So, it suffices to show that

$$\widehat{s}^2\longrightarrow s^2\quad \text{In probability.}$$

Indeed, we define the following U-statistic

$$\widehat{s}_0^2=\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i,=1}^{n}K_{ij}^2u_i^2u_j^2.$$

We use the uniform consistency of $\widehat{R}(\cdot)-R(\cdot)$ and $\widehat{\sigma}-\sigma$ to deduce that

$$\widehat{s}^2=\widehat{s}_0^2+o(1)\quad \text{In probability.}$$

In addition, from (9) and (8) we obtain

$$\mathbb{E}\left[H'^2(Z_i, Z_j)\right] = \phi^2(h)\mathbb{E}\left[H^4(Z_i, Z_j)\right] = O\left(\frac{1}{\phi(h)}\right) = o(n)$$

and

$$\mathbb{E}\left[H'(Z_i, Z_j)\right] = \phi(h)\mathbb{E}\left[H^2(Z_i, Z_j)\right] = s^2$$

Therefore, Lemma2.4.2 yields the proof of the second part of this Theorem.

Proof of Theorem 2.6

The proof follows the same lines as that of the second part of Theorem 2.5. In particular, the uniform consistency of $\widehat{R}(\cdot) - R(\cdot)$ and $\widehat{\sigma} - \sigma$ allow to write that

$$W_n = W_{1n} + o(1) \quad \text{In probability.}$$

So, it suffices to state, under H_1, the consistency of W_{1n} and $\widehat{s_0^2}$. Once again, we apply Lemma 2.4.3 to W_{1n} and $\widehat{s_0^2}$ as U-statistics. So, it suffices to evaluate, under H_1, the following quantities

$$\mathbb{E}\left[H'^2(Z_i, Z_j)\right], \quad \mathbb{E}\left[H'(Z_i, Z_j)\right], \quad \mathbb{E}\left[H^2(Z_i, Z_j)\right] \text{and} \quad \mathbb{E}\left[H(Z_i, Z_j)\right]$$

It is shown in the proof of (11) that

$$\begin{aligned}
\mathbb{E}\left[H^2(Z_1, Z_2)\right] &= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\int f(x_2)\mathbb{E}^2\left[u_2^2|x_2\right]dP_{X_2}(x_2) + o\left(\frac{1}{\phi(h)}\right) \\
&= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(\mathbb{E}\left[u_2^2|X\right]\right)^2\right] + o\left(\frac{1}{\phi(h)}\right) \\
&= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right] + o\left(\frac{1}{\phi(h)}\right).
\end{aligned}$$

Therefore

$$\mathbb{E}\left[H^2(Z_1, Z_2)\right] = O\left(\frac{1}{\phi(h)}\right) = o(n)$$

and

$$\mathbb{E}\left[H'(Z_i, Z_j)\right] = \left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right] + o(1).$$

By the same way

$$\begin{aligned}
n^{-1}\mathbb{E}\left[H'^2(Z_1,Z_2)\right] &= \frac{1}{n\phi^2(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4|X_2\right]\mathbb{E}\left[u_1^4|X_1\right]\right] \\
&\leq \frac{C}{n\phi^2(h)}\int\int_{B(x_2,h)} dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\leq \frac{C}{n\phi(h)}\int f(x)dP_X(x) \to 0.
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[H(Z_1,Z_2)\right] &= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}u_1u_2\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2u_1|X_2,X_1\right]|\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2|X_2\right]\mathbb{E}\left[u_1|X_1\right]\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2|X_2\right]\right] + o\left(\frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2|X_2\right]\right]\right) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\int f(x_2)\mathbb{E}^2\left[u_2|x_2\right]dP_{X_2}(x_2) + o(1) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\mathbb{E}^2\left[u|X\right]\right] + o(1) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon|X\right] - \sigma^2)^2\right] + o(1)
\end{aligned}$$

So, from Lemma 2.4.3 we get

$$W_n \longrightarrow \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon_2|X\right] - \sigma^2)^2\right] \quad \text{In probability.}$$

and

$$\widehat{s}^2 \longrightarrow 2\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right] + o(1) \quad \text{In probability.}$$

Consequently

$$\frac{T_n}{n\phi(h)} \to \frac{\left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon_2|X\right] - \sigma^2)^2\right]}{2\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right]} \quad \text{In probability.}$$

Proof of Corollary2.3

Similarly to previous Theorem, it suffices to prove the asymptotic normality of $n\varphi^{1/2}(h)W_n$ and the convergence in probability of $\widehat{s}^2$. For the first purpose we use the same decomposition of 2.5 to write

$$W_{1n} = W_{1n} + o((n\varphi^{1/2}(h))^{-1})$$

Thereafter, we introduce $u_i' = u_i - \delta_n g(X_i)$, thus

$$\begin{aligned}
W_{1n} &= \frac{1}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j \neq i} K_{ij} u_i u_j \\
&= \frac{1}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j \neq i} K_{ij} (u_i' + \delta_n g(X_i))(u_j' + \delta_n g(X_j)) \\
&= \frac{1}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j \neq i} K_{ij} u_i' u_j' \\
&+ \frac{2\delta_n}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j \neq i} K_{ij} u_i' g(X_j) \\
&+ \frac{\delta_n^2}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j \neq i} K_{ij} g(X_i) g(X_j) \\
&= U_{11n} + 2\delta_n U_{12n} + \delta_n^2 U_{13n}.
\end{aligned}$$

Observe that $\mathbb{E}[u_i'|X_i] = 0$, then we can treat U_{11n} as degenerate U-statistic similarly to (11). Therefore,

$$n\sqrt{\varphi(h)}U_{11n} \to N(0, s^2)$$

and by the same argument as those used in the consistency of $\widehat{s}^2$ we get for $\delta_n = n^{-1/2}\varphi^{-1/4}(h)$

$$n\sqrt{\varphi(h)}\delta_n^2 U_{13n} = U_{13n} \to \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}[g^2(X)f(X)], \quad \text{In probability}$$

and

$$n\sqrt{\varphi(h)}\delta_n U_{12n} = \sqrt{n}(\varphi(h))^{1/4} U_{12n} \to 0, \quad \text{In probability.}$$

Now, it suffices to use the the Slutsky lemma to conclude the claimed result.

Bibliography

[1] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data. *Computational statistics and data analysis*, **22**, 353-369.

[2] Bongiorno E., Salinelli, E., Goia, A. and Vieu, P. (Eds.) (2014). An overview of IWFOS'2014. In *Contributions in infinite-dimensional statistics and related topics*, 1-6, Esculapio, Bologna.

[3] Breusch, T. S. and Pagan, A. R.(1979) A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1294.

[4] Cardot, H., Goia, A. et Sarda, P. (2004) Testing for no efect in functional linear regression models, some computational approaches. *Comm. Statist. Simulation Comput.*, **33** , 179-199.

[5] Cuevas, A., Febrero, M., and Fraiman, R. (2004). An anova test for functional 575 data. *Computational statistics and data analysis*, **47** , 111-122.

[6] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference*, **147**, 1-23.

[7] Delsol, L. (2013). No effect tests in regression on functional variable and some applications to spectrometric studies.*Comput. Statist.*, **28** , 1775–1811.

[8] Dette, H. and Munk, A. (1998). Testing heteroscedasticity in nonparametric regression. *Journal of the Royal Statistical Society*, Series B,**60**, 693-708.

[9] Diblasi, A. and Bowman, A (1997). Testing for constant variance in a linear model. *Statistics and Probability Letters*, **33**, 95-103.

[10] Ezzahrioui, M., Ould-Saïd, E. (2008). Asymptotic normality of nonparametric estimator of the conditional mode for functional data.*J. of Nonparametric Statist.*, **20**, 3-18.

[11] Ferraty, F., Mas, A. and P. Vieu (2007). Nonparametric regression on functional data: inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**, 267-286.

[12] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer-Verlag, New York, 2006.

[13] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference*, **140**, 335-352.

[14] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *J. Mutiv. Anal.*, **146**, 1-6.

[15] Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer Series in Statistics. Springer, New York.

[16] Hsing, T. and Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester.

[17] Karaa, L-Z. Laksacia, A, Rachdi, M. and Vieu, P. (2015) Uniform in bandwidth consistency results for kernel estimates in various statistical models with functional. *J. Nonparametr. Statist.* (accepted).

[18] Koenker, B. and Bassett, G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, **50**, 43-61.

[19] Lin, J. G. and Qu, X. Y. (2012). A consistent test for heteroscedasticity in semiparametric regression with nonparametric variance function based on the kernel method, *Statistics*, **46**, 565-576.

[20] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.

[21] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, New York.

[22] Staicu, A., Li, Y. Crainiceanu, C. M. and Ruppert, D. (2014). Likelihood ratio tests for dependent data with applications to longitudinal and functional data analysis. *Scand. J. Stat.*, **41**, 932-949.

[23] Zhang, J.-T., Liang, X., and Xiao, S. (2010). On the two-sample behrens-fisher problem for functional data. *Journal of Statistical Theory and Practice*, **4** , 571-587.

[24] Zhang, J. (2013). *Analysis of variance for functional data.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

[25] Zheng, J. X. (2009). A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics, **75**, 263-289.

[26] Zheng, J. X. (2009). Testing heteroscedasticity in nonlinear and nonparametric regressions, Canadian Journal of Statistics, **37**, 282-300.

Chapter 3: Testing heteroscedasticity in nonlinear regression

In the current Chapter, we focused on the parametric case, therefore, we consider a nonlinear regression models where the explanatory variable valued in an infinite dimensional space $\mathcal{F}$. Thereafter, we follow the same steps just as the previous Chapter.

Contents

In what comes below, we present a consistent nonparametric test for **heteroscedasticity**. Thereafter, we show that our test has asymptotic normality under the homoscedasticity hypothesis. Furthermore, we establish the consistency of the constructed test against any form of **heteroscedasticity**.

3.1 Introduction

Over several years, the multivariate data has been used for modeling. However, keeping up with the age of speed the data are more being recorded continuously during the time interval. Thus, the classical methods of multivariate data analysis do not seem suitable for studying this kind of data. From here required a new statistic branch called Functional Statistic, which treat the observations like a functional random elements valued in an infinite dimensional space. This kind of data has received a lot of attention in the literature either at physiologic (***Brumback and Rice*** [1998]), biologic (***Müller and Stadtmüller*** [2005]), demographics (***Chiou and Müller***[2009]),...

We concentrate in this Chapter on a functional regression model precisely where the response variable Y is real valued whereas the explanatory variable X belongs to a functional space. Moreover, we propose a consistent nonparametric test for **heteroscedasticity** depending on the kernel estimations of the nonlinear regression. It's worth noting that, this model is not efficiency in the heteroscedasticity case, it requires the homoscedasticity of the data which cannot be guaranteed a priori. Thus, testing this phenomenon is of utmost importance. To the best of our knowledge, testing **heteroscedasticity** in functional statistics has not been addressed so far, whereas, we more recently have been the first who established this phenomenon in nonparametric regression (***Henien, Laksaci et al***[2018]).

Similar to the last work, we based on the nonparametric approaches such that it is more appropriate than the parametric one. In the literature, this methods take a large place in Functional Data Analysis (FDA), thus, it has been the preferred one of several authors for modeling (see ***Ferraty and Vieu*** [2006] for basic results or ***Cuevas***[2014] and ***Goia and Vieu***[2016] for recent advanced and references). Moreover, in the functional statistics literature, a limited number of tests are found, thus, a various test problems are considered either in parametric or nonparametric structures (see: ***Cardot et al***[2003], ***Cuevas et al*** [2004], ***Delsol***[2013], ***Zhang et al***[2010], ***Hilgert et al***[2013], ***Staicu et al***[2014]). In addition, there have been also a proposed tests in the case when the response variable is real and the explanatory is functional one. However, this kind of model are very few (see: ***Cardot et al.***[2004] ***Gadiaga and Ignaccolo*** [2005], ***Chiou and Müller*** [2007]). In multivariate statistics, testing **heteroscedasticity**has received a lot of attention, several authors have proposed either parametric or nonparametric methods in the literature, citing ***Zheng***[2009], ***Lin et al.***[2012] or ***Breusch and Pagan***[1979], ***Diblasi and Bowman*** [1997] and among others for the proposed approaches.

In what follows, we extend the previous work of ***Zheng***[2009] in multivariate statistics to the infinite dimensional space. Indeed, we aim to present a consistent nonparametric test for **heteroscedasticity**, based on the kernel estimation of the nonlinear regression. Thereafter, we show that our test has asymptotic normality under the homoscedasticity hypothesis and can be also consistent against any form of **heteroscedasticity**. It should be noted that, to avoid the problem of the curse of dimensionality in multivariate case, some standard conditions in nonparametric functional statistics are used through this paper.

We organize our work as follows. In the following Section we construct our test statistic. The main results of this paper are given in Section 3, in which, we study the asymptotic behavior of the considered test statistics under the null (respectively alternative) hypothesis. All proofs are put into the Appendix.

3.2 Construction of the test statistic

We consider a set of independent pairs of random variables $(X_i, Y_i)_{1\leq i\leq n}$ identically distributed as (X, Y) which is a random vector valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space. x is a fixed point in $\mathcal{F}$, N_x is a fixed neighborhood of x and the closed ball centered at x and of radius a is denoted

$$B(x, a) = \{y \in \mathcal{F} \text{ such that } d(y, x) \leq a\}.$$

Moreover, the nonlinear regression operator r of Y on X is defined by $E[Y_i \mid X_i = X] = r(X)$ where $r(X)$ assumed to belong to a parametric family of known real functions $f(X, \partial)$ on $F \times \Theta$ where $\Theta \subset \mathbb{R}$. Indeed, r characterized the by following regularity condition

$$\forall x_1, x_2 \in N_x, \quad |r(x_1) - r(x_2)| \leq Cd^{\beta}(x_1, x_2), \quad C > 0 \quad \beta > 0. \tag{1}$$

Furthermore, we suppose that the regressor X such that, there exists a nonnegative continuous functions ϕ and f such that

$$\mathbb{P}(X \in B(x, a)) = \phi(a).f(x) + o(\phi(a)). \tag{2}$$

Thence, we suppose that the function ϕ such that:

$$\text{For all } s \in [0, 1], \qquad \lim_{a\to 0} \frac{\phi(sa)}{\phi(a)} = \tau(s) \tag{3}$$

Concerning the function f, we assume that the Kolmogorov's ϵ-entropy $\psi_{\mathcal{S}}$ of the support $\mathcal{S}$ of f such that

$$\sum_{n=1}^{\infty} \exp\left\{(1-\eta)\psi_{\mathcal{S}}\left(\frac{\log n}{n}\right)\right\} < \infty, \qquad \text{for some } \eta > 1. \tag{4}$$

We point out that this considerations is classic in nonparametric functional data analysis (see, ***Masry***[2005] ***Ferraty, Laksaci et al.***[2010], for more discussion in this question).

Therefore, we present our **heteroscedasticity** test as flow,

$$H_0 : \quad V[\epsilon|X] = \sigma^2 \qquad \text{versus} \qquad H_1 : \quad V[\epsilon|X] \neq \sigma^2$$

Where ϵ is a random error variable such that $\mathbb{E}[\epsilon|X] = 0$. To construct test statistic, we denote $\epsilon_i = Y_i - f(X_i, \partial_0)$ and suppose that the function p such that

$$p(X) > 0 \quad \text{almost surly,} \quad \text{and} \quad \mathbb{E}[p(X)] < \infty. \tag{5}$$

Then, under H_0 , since $E\left[\epsilon^2 \mid X\right] = V[\epsilon \mid X] = \sigma^2$, we can show that

$$W = \mathbb{E}[(\epsilon^2 - \sigma^2)\mathbb{E}[(\epsilon^2 - \sigma^2)|X]p(X)] = \mathbb{E}[\mathbb{E}^2[(\epsilon^2 - \sigma^2)|X]p(X)] = 0.$$

While H_1 is equivalent to write

$$W = \mathbb{E}[(\epsilon^2 - \sigma^2)\mathbb{E}[(\epsilon^2 - \sigma^2)|X]p(X)] = \mathbb{E}[\mathbb{E}^2[(\epsilon^2 - \sigma^2)|X]p(X)] = \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 p(X)] > 0.$$

Therefore, we consider the empirical version of W denoted by

$$\widehat{W} = \frac{1}{n}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2)\Lambda_i.$$

Where,

$$\Lambda_i = \mathbb{E}[(\epsilon_i^2 - \sigma^2)|X_i]p(X_i)$$

Now it sufficient to estimate $E(\epsilon_i^2 \mid X_i)$ by the kernel method :

$$\widehat{E}(\epsilon^2 \mid X_i) = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\phi(h)}K(d(X_i, X_j)h^{-1})\epsilon_j^2}{\widehat{p}(X_i)}$$

While, p can be estimates by (see, ***Ezzahrioui and Ould-Saïd***[2008]) :

$$\widehat{p}(x) = \frac{1}{n\varphi(h)} \sum_{i=1}^{n} K\left(d(X_i, X_j)h^{-1}\right).$$

Replacing ϵ_i and σ^2 by their estimators $\widehat{\epsilon_i} = Y_i - f(X_i, \widehat{\partial}$[1] and $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \widehat{\epsilon_i}^2$, it follows that the natural estimate of W is

$$W_n = \frac{1}{n(n-1)\varphi(h)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} K\left(d(X_i, X_j)h^{-1}\right)(\widehat{\epsilon_i}^2 - \widehat{\sigma}^2)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)$$

3.3 The main result

In order to establish our asymptotic results, we impose the following regularity assumptions.

(H1) $\mathbb{E}[\epsilon^8 | X = x] \leq b(x)$ with $b(x)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[b^2(X)] < \infty$.

(H2) The parameter space Θ is a compact and convex subset of $\mathbb{R}$. $f(X, \partial)$ is a Borel measurable function on $\mathcal{F} \times \mathbb{R}$ for each ∂ and a twice continuously differentiable real function on Θ for each $X \in \mathcal{F}$. Moreover, $E\left[\sup_{\partial\in\Theta} f(X_i, \partial)^2\right] < \infty$, and

$$E\left[\sup_{\partial\in\Theta}\left|\frac{\delta^2 f(X_i, \partial)}{\delta\partial}\right|\right] < \infty,$$
$$E\left[\sup_{\partial\in\Theta}\left|(y_i - f(X_i, \partial))^2 \frac{\delta^2 f(X_i, \partial)}{\delta\partial}\right|\right] < \infty.$$

(H3) $E\left[(y_i - f(X_i, \partial))^2\right]$ takes a unique minimum at $\partial_0 \in \Theta$.

(H4) The kernel K is a differentiable function supported on $[0, 1]$ such that

$$K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds > 0 \quad \text{and} \quad K(1) - \int_0^1 (K(s))'\tau(s)ds \neq 0.$$

(H5) The bandwidth parameter $h := h(n)$ is strictly positive such that:

$$n \to 0, n\,\varphi(h) \to \infty,\ n\sqrt{\varphi(h)}\max\left(h^{4\beta}, \frac{1}{\log^2 n}\right) \to 0 \ \text{ and } \ \frac{(\log n)^2}{n\,\varphi(h)} < \psi_S\left(\frac{\log n}{n}\right) < \frac{n\,\varphi(h)}{\log n}$$

when n tends to infinity.

[1] $\widehat{\partial}$ is the nonlinear least square estimator of ∂_0 where this later can be estimated by any $\sqrt{n}$–consistent method.

Assumptions 2 and 3 are standard for ensuring the consistency and asymptotic normality of nonlinear least squares estimators. Assumption 4 and 5 are the same as those used by ***Ferraty, laksaci et al.***[2010] for the uniform consistency in functional statistics. While, the assumption 1 is the same as in ***Zheng***[2009] for the **heteroscedasticity** in multivariate case.

- Under H_0

Theorem 3.7. *If (H1)-(H6), (3) and (5) hold, then we get*

$$n\sqrt{\phi(h)}\,W_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \text{ as } n \to \infty$$

Where $s^2 = 2\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)\mathbb{E}\left[f(X)V^2\left[\epsilon_2^2|X\right]\right]$.

Moreover,

$$T_n = n\sqrt{\phi(h)}\,\frac{W_n}{\widehat{s}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \to \infty$$

where

$$\widehat{s}^2 = \frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i,=1}^{n} K\left(d(X_i, X_j)h^{-1}\right)(\widehat{\epsilon_j}^2 - \widehat{\sigma}^2)^2(\widehat{\epsilon_i}^2 - \widehat{\sigma}^2)^2.$$

- Under H_1

Theorem 3.8. *When (H1)-(H6), (3) and (5) hold, then we get*

$$\frac{T_n}{n\sqrt{\phi(h)}} \longrightarrow \mathbb{E}[(V[\epsilon|X] - \sigma^2)^2 p(X)]/s_1, \quad \text{In probability}$$

$$\text{where } s_1^2 = \frac{\left(K^2(1) - \int_0^1 (K^2(s))'\tau(s)ds\right)}{\left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)}\mathbb{E}[(V[\epsilon^2|X] + (V[\epsilon|X] - \sigma^2)^2)p(X)].$$

Now, we have to show that our test robust against all possible departures from homoscedasticity. Thence, we consider a sequence of local alternatives

$$H_{1n}: \quad V[\epsilon|x] - \sigma^2 = \delta_n g(x)$$

where the known function $g(\cdot)$ is continuous on $\mathcal{S}$ such that $\mathbb{E}[g^2(X)] < \infty$. So, we obtain the following Corollary

Corollary 3.4. *Given (H1)-(H4), (3) and (5) we get, under H_{1n} with $\delta_n = n^{-1/2}\phi^{-1/4}(h)$*

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(\mu, 1) \text{ as } n \to \infty$$

where $\mu = \left(K(1) - \int_0^1 (K(s))'\tau(s)ds\right)\mathbb{E}[g^2(X)p(X)]/s.$

3.4 Appendix

To prove our results, we based on the same decomposition of ***Zheng***[2009] and the following Lemmas.

Lemma 3.4.1. *(see, **Zheng**[2009])*

Let U_n a second-order U-statistic of kernel $H_n(\cdot,\cdot)$ such that $\mathbb{E}[H_n^2(Z_i, Z_j)] = o(n)$, then

$$U_n - \mathbb{E}[H_n(Z_i, Z_j)] = o(1) \quad \textit{In probability.}$$

Lemma 3.4.2. *(see, **Zheng**[2009])*

Assume that $\mathbb{E}[H_n(Z_1, Z_2)|Z_1] = 0$ and $\mathbb{E}[H_n^2(Z_1, Z_2)] < \infty$ for each n. If

$$\frac{\mathbb{E}[G_n^2(Z_1, Z_2)] + n^{-1}\mathbb{E}[H_n^4(Z_1, Z_2)]}{\mathbb{E}^2[H_n^2(Z_1, Z_2)]} \to 0$$

where $G_n(Z_1, Z_2) = \mathbb{E}\left[H_n(Z_1, Z_3)H_n(Z_3, Z_2)|Z_1, Z_2\right]$, *then*

$$\frac{nU_n}{2\mathbb{E}^{1/2}[H_n^2(Z_1, Z_2)]} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \text{ as } n \to \infty.$$

Lemma 3.4.3. *(see **Jennrich, 1969, White**[1981,1982]),*

$$\sqrt{n}(\widehat{\partial} - \partial_0) = O_p(1).$$

Over what come behind, we point out for convenience notation that, C and C' some strictly positive generic constants, as well, for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$:

$$K_{ij} = K(h^{-1}d(X_i, X_j)), \; \epsilon_i = Y_i - f(X_i, \partial_0)$$

and

$$p_i = p(X_i), \; u_i = \epsilon_i^2 - \sigma^2, \; \widehat{p_i} = \widehat{p}(X_i).$$

Proof of Theorem 3.7

To make W_n's treatment easier, we write

$\widehat{\epsilon}_i^2 - \widehat{\sigma}^2 = u_i - 2\epsilon_i(f(X,\widehat{\eth}) - f(X,\eth_0)) + (f(X,\widehat{\eth}) - f(X,\eth_0))^2 + (\sigma^2 - \widehat{\sigma}^2)$

Therefore, W_n will be decomposed as follows

$$
\begin{aligned}
W_n &= \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_iu_j\right\}}_{W_{1n}} \\
&+ 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i\epsilon_j(f(X_i,\widehat{\eth}) - f(X_i,\eth_0))(f(X_j,\widehat{\eth}) - f(X_j,\eth_0))\right\}}_{W_{2n}} \\
&+ \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(f(X_i,\widehat{\eth}) - f(X_i,\eth_0))^2(f(X_j,\widehat{\eth}) - f(X_j,\eth_0))^2\right\}}_{W_{3n}} \\
&+ \underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(\widehat{\sigma}^2 - \sigma^2)^2\right\}}_{W_{4n}} \\
&- 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i\epsilon_j(f(X_j,\widehat{\eth}) - f(X_j,\eth_0))\right\}}_{W_{5n}} \\
&+ 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i(f(X_j,\widehat{\eth}) - f(X_j,\eth_0))^2\right\}}_{W_{6n}} \\
&- 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i(\widehat{\sigma} - \sigma^2)\right\}}_{W_{7n}} \\
&- 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i(f(X_i,\widehat{\eth}) - f(X_i,\eth_0))(f(X_j,\widehat{\eth}) - f(X_j,\eth_0))^2\right\}}_{W_{8n}} \\
&+ 4\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}\epsilon_i(f(X_i,\widehat{\eth}) - f(X_i,\eth_0))(\widehat{\sigma}^2 - \sigma^2)\right\}}_{W_{9n}} \\
&- 2\underbrace{\left\{\frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(f(X_i,\widehat{\eth}) - f(X_i,\eth_0))^2(\widehat{\sigma}^2 - \sigma^2)\right\}}_{W_{10n}} \\
&\equiv W_{1n} + 4W_{2n} + W_{3n} + W_{4n} - 4W_{5n} + 2W_{6n} - 2W_{7n} - 4W_{8n} + 4W_{9n} - 2W_{10n}.
\end{aligned}
$$

Thus, it remains to show that under the null

$$n\sqrt{\phi(h)}\,W_{1n} \xrightarrow{\mathcal{D}} \mathcal{N}(0, s^2) \text{ as } n \to \infty \tag{6}$$

and

$$n\sqrt{\phi(h)}W_{in} = o_p(1), i = 2, \ldots, 10.$$

Foremost we prove (6). To this point, W_{1n} can be written as U-statistic of the form below

$$W_{1n} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} H_n(Z_i, Z_j)$$

whither

$$H_n(Z_i, Z_j) = \frac{1}{\phi(h)} K_{ij} u_i u_j \quad \text{with} \quad Z_i = (X_i, \epsilon_i).$$

By using the conditional expectation properties we prove that, under H_1,

$$\mathbb{E}[H_n(Z_1, Z_2)|Z_1] = \frac{1}{\phi(h)} u_1 K_{12} \mathbb{E}\left[u_2|X_2\right] = 0.$$

Consequently, U_{1n} is a degenerate U-statistic, and so on, we can apply Lemma3.4.2 by considering the quantities

$$\mathbb{E}\left[G_n^2(Z_1, Z_2)\right], \quad \mathbb{E}\left[H_n^4(Z_1, Z_2)\right] \quad \text{and} \quad \mathbb{E}\left[H_n^2(Z_1, Z_2)\right].$$

Firstly, we write under (H4), (2) and (3) that

$$\begin{aligned}
\mathbb{E}\left[G_n^2(Z_1,Z_2)\right] &= \mathbb{E}\left[\mathbb{E}\left[H_n(Z_1,Z_3)H_n(Z_3,Z_2)|Z_1,Z_2\right]^2\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[u_1^2u_2^2\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[K_{13}K_{23}u_3^2|X_3\right]|X_1,X_2\right]^2\right]\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[u_1^2u_2^2\mathbb{E}\left[K_{13}K_{32}\mathbb{E}\left[u_3^2|X_3\right]|X_1,X_2\right]^2\right] \\
&= \frac{1}{\phi^4(h)}\int\int\left[\int K(h^{-1}d(x_1,x_3))K(h^{-1}d(x_3,x_2))\right. \\
&\qquad \left.\mathbb{E}\left[u_3^2|x_3\right]dP_{X_3}(x_3)\right]^2\mathbb{E}\left[u_1^2|X_1\right]\mathbb{E}\left[u_2^2|X_2\right]dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&= \frac{1}{\phi^4(h)}\iint_{D=\{x_1,x_2,\ d(x_1,x_2)\le 2h\}}\left[\int_{B(x_1,h)\cap B(x_2,h)}K(h^{-1}d(x_1,x_3))K(h^{-1}d(x_3,x_1))\right. \\
&\qquad \left.\mathbb{E}\left[u_3^2|x_3\right]dP_{X_3}(x_3)\right]^2\mathbb{E}\left[u_1^2|X_1\right]\mathbb{E}\left[u_2^2|X_2\right]dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\le \frac{C}{\phi^4(h)}\iint_{D=\{x_1,x_2,\ d(x_1,x_2)\le 2h\}}\left[\int_{B(x_1,h)\cap B(x_2,h)}dP_{X_3}(x_3)\right]^2 dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\le \frac{C}{\phi(h)}.
\end{aligned}$$

Therefore

$$\mathbb{E}\left[G_n^2(Z_1,Z_2)\right] = O\left(\frac{1}{\phi(h)}\right). \tag{7}$$

Secondly, we write

$$\begin{aligned}
\mathbb{E}\left[H^4(Z_1,Z_2)\right] &= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4u_1^4u_2^4\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4u_1^4|X_2,X_1\right]\right] \\
&= \frac{1}{\phi^4(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4|X_2\right]\mathbb{E}\left[u_1^4|X_1\right]\right] \\
&\le \frac{C}{\phi^4(h)}\int\int_{B(x_2,h)}dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\le \frac{C}{\phi^3(h)}\int f(x)dP_X(x).
\end{aligned}$$

Thus

$$\mathbb{E}\left[H^4(Z_1,Z_2)\right] = O\left(\frac{1}{\phi^3(h)}\right). \tag{8}$$

Now, we require the continuity of the operator $\mathbb{E}[\epsilon^4|X=\cdot]$ to the remainder term, and so, we get

$$\begin{aligned}
\mathbb{E}\left[H^2(Z_1,Z_2)\right] &= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2u_1^2u_2^2\right]\\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2^2u_1^2|X_2,X_1\right]\right]\\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2^2|X_2\right]\mathbb{E}\left[u_1^2|X_1\right]\right]\\
&= \frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2^2|X_2\right]\right]+o\left(\frac{1}{\phi^2(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2^2|X_2\right]\right]\right)\\
&= \frac{1}{\phi(h)}\left(K^2(1)-\int_0^1(K^2)'(s)\tau(s)ds\right)\int f(x_2)\mathbb{E}^2\left[u_2^2|x_2\right]dP_{X_2}(x_2)+o\left(\frac{1}{\phi(h)}\right)\\
&= \frac{1}{\phi(h)}\left(K^2(1)-\int_0^1(K^2)'(s)\tau(s)ds\right)\int f(x_2)\left(V\left[\epsilon^2|x_2\right]\right)^2dP_{X_2}(x_2)+o\left(\frac{1}{\phi(h)}\right)
\end{aligned}$$

$$= \frac{1}{\phi(h)}\left(K^2(1)-\int_0^1(K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right]\right)^2\right]+o\left(\frac{1}{\phi(h)}\right).$$

9

As a result

$$\mathbb{E}\left[H^2(Z_1,Z_2)\right]=O\left(\frac{1}{\phi(h)}\right).$$

Combining this last evaluation to (7) and (8) to obtained

$$\frac{\mathbb{E}\left[G_n^2(Z_1,Z_2)\right]+n^{-1}\mathbb{E}\left[H^4(Z_1,Z_2)\right]}{\mathbb{E}^2\left[H^2(Z_1,Z_2)\right]}=O(\phi(h))+O\left(\frac{1}{n\phi(h)}\right)=o(1)$$

Hence, from Lemma (3.4.2) we have

$$\frac{nW_{1n}}{2\mathbb{E}^{1/2}\left[H^2(Z_1,Z_2)\right]}\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\text{ as }n\to\infty$$

which complete the proof of (6).

Now, we have to proof that $n\sqrt{\phi(h)}W_{2n}=o(1)$. For be done, we use the Taylor expansion to write an approximation of W_{2n} as

$$\begin{aligned}
W_{2n} &= (\widehat{\partial}-\partial_0)\frac{1}{n(n-1)}\sum_{i=1}^n\sum_{j\neq i}^n\frac{1}{\phi(h)}K_{ij}\epsilon_i\epsilon_j\times\frac{\delta f(X_i,\xi_1)}{\delta\partial}\frac{\delta f(X_j,\xi_2)}{\delta\partial}(\hat{\partial}-\partial_0)\\
&= (\hat{\partial}-\partial_0)S_{1n}(\hat{\partial}-\partial_0)
\end{aligned}$$

where ξ valued in interval $[\partial_0,\widehat{\partial}]$ (ξ_1 depending on X_i and ξ_2 depending on X_j).

Thence, we have

$$\begin{aligned} E[|S_{1n}|] &\leq E\left[\frac{1}{\varphi(h)}K_{ij}\sigma(X_i)\sigma(X_j)\left|\frac{\delta f(X_i,\xi_1)}{\delta\partial}\frac{\delta f(X_j,\xi_2)}{\delta\partial}\right|\right] \\ &= \int\frac{1}{\varphi(h)}K_{ij}\sigma(X_i)\sigma(X_j)\left|\frac{\delta f(X_i,\xi_1)}{\delta\partial}\frac{\delta f(X_j,\xi_2)}{\delta\partial}\right|dP^{X_i}(x_i)dP^{X_j}(x_j) \\ &= \int K(u)\sigma(X_i)\sigma(X_j)\left|\frac{\delta f(X_i,\xi_1)}{\delta\partial}\frac{\delta f(X_j,\xi_2)}{\delta\partial}\right|f(x_j)dP^{X_i}(x_i) \\ &= O(1). \end{aligned}$$

Whereas, $\sigma(X)$ refer to $E(|\ \epsilon\ |\ |X)$. By applying the lemma3.4.3 we have

$$\begin{aligned} W_{2n} &= O_p(1/\sqrt{n})O_p(1)O_p(1/\sqrt{n}) \\ &= O_p(1/n). \end{aligned}$$

Thus

$$n\sqrt{\varphi(h)}W_{2n} = O(\sqrt{\varphi(h)}) \xrightarrow{P} 0.$$

Just as the same way, W_{3n} can be also rewrite such

$$\begin{aligned} W_{3n} &= (\widehat{\partial}-\partial_0)^2\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{1}{\varphi(h)}K_{ij}\left(\frac{\delta f(X_i,\xi_1)}{\delta\partial}\frac{\delta f(X_j,\xi_2)}{\delta\partial}\right)^2(\hat{\partial}-\partial_0)^2 \\ &= (\widehat{\partial}-\partial_0)^2 S_{2n}(\hat{\partial}-\partial_0)^2. \end{aligned}$$

Similar to the proof of S_{1n}, we can show that

$$W_{3n} = O_p(1/n).O_p(1).O_p(1/n) = O_p(1/n).$$

Finally, we have

$$n\sqrt{\varphi(h)}W_{3n} = O_p(\sqrt{\varphi(h)}) \xrightarrow{P} 0.$$

For the term W_{4n} it suffices to prove that

$$n\sqrt{\varphi(h)}(\widehat{\sigma}^2-\sigma^2)^2 = o(1) \quad \text{In probability.}$$

In fact, we use the following decomposition

$$\widehat{\sigma}^2 - \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2) + \frac{2}{n}\sum_{i=1}^{n}\epsilon_i(f(X_i,\hat{\partial}) - f(X_i,\partial_0)) + \frac{1}{n}\sum_{i=1}^{n}(f(X_i,\hat{\partial}) - f(X_i,\partial_0))^2.$$

Thence, we can easily achieve that

$$V\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2)\right] = V\left[\epsilon^2\right], \qquad \mathbb{E}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2)\right] = 0$$

and

$$V\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_i\right] = V\left[\epsilon\right], \qquad \mathbb{E}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_i\right] = 0.$$

Thus

$$\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i^2 - \sigma^2) = o\left(\frac{1}{\sqrt{n\sqrt{\phi(h)}}}\right) \quad \text{In probability}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(f(X_i,\hat{\partial}) - f(X_i,\partial_0)) = o\left(\frac{1}{\sqrt{n\sqrt{\phi(h)}}}\right) \quad \text{In probability.}$$

Similarly to the proof of W_{2n} and W_{3n}

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left(f(X_i,\hat{\partial}) - f(X_i,\partial_0)\right)^2 &= \frac{1}{n}\sum_{i=1}^{n}\left((\hat{\partial} - \partial_0)\frac{\delta f(X_i,\xi_1)}{\delta\partial}\right)^2 \\ &= O(1/n).\end{aligned}$$

Now, we can easily show that

$$\sqrt{n\sqrt{\phi(h)}}(\widehat{\sigma}^2 - \sigma^2) = o(1) \quad \text{In probability.}$$

Hence, we deduce that

$$n\sqrt{\phi(h)}W_{4n} = o(1) \quad \text{In probability.}$$

In the same way as the proofs of the treated cases, we can easily get

$$n\sqrt{\phi(h)}W_{in} = o(1) \quad \text{In probability, for} \quad i = 5, \ldots, 10.$$

Thus, the proof of the first part is completed.

The second part of Theorem 3.7 can be proved by using the Slutsky lemma. Indeed, it suffices to show that

$$\widehat{s}^2 \longrightarrow s^2 \quad \text{In probability.}$$

In similar way of the proof of W_{1n}, we get

$$\begin{aligned} \widehat{s}_0^2 &= \frac{1}{n(n-1)\phi(h)} \sum_{i=1}^{n} \sum_{j\neq i,=1}^{n} K_{ij}^2 u_i^2 u_j^2. \\ &= \widehat{s}_0^2 + o(1) \end{aligned}$$

Moreover, from (9) and (8) we obtain

$$\mathbb{E}\left[H'^2(Z_i, Z_j)\right] = \phi^2(h)\mathbb{E}\left[H^4(Z_i, Z_j)\right] = O\left(\frac{1}{\phi(h)}\right) = o(n)$$

and

$$\mathbb{E}\left[H'(Z_i, Z_j)\right] = \phi(h)\mathbb{E}\left[H^2(Z_i, Z_j)\right] = s^2$$

Thus, Lemma3.4.2 yields the proof of the second part of this Theorem.

Proof of Theorem 3.8

Following the same procedures as the second part of theorem 3.7, theorem 3.8 can be proved.

Actually, we have just to evaluate, under H_1, the following quantities

$$\mathbb{E}\left[H'^2(Z_i, Z_j)\right], \quad \mathbb{E}\left[H'(Z_i, Z_j)\right], \quad \mathbb{E}\left[H^2(Z_i, Z_j)\right] \text{and} \quad \mathbb{E}\left[H(Z_i, Z_j)\right]$$

It is shown in the proof of (11) that

$$\begin{aligned} \mathbb{E}\left[H^2(Z_1, Z_2)\right] &= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right) \int f(x_2)\mathbb{E}^2\left[u_2^2|x_2\right] dP_{X_2}(x_2) + o\left(\frac{1}{\phi(h)}\right) \\ &= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(\mathbb{E}\left[u_2^2|X\right]\right)^2\right] + o\left(\frac{1}{\phi(h)}\right) \\ &= \frac{1}{\phi(h)}\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right] + o\left(\frac{1}{\phi(h)}\right). \end{aligned}$$

Therefore

$$\mathbb{E}\left[H^2(Z_1,Z_2)\right] = O\left(\frac{1}{\phi(h)}\right) = o(n)$$

and

$$\mathbb{E}\left[H'(Z_i,Z_j)\right] = \left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X]-\sigma^2)^2\right)^2\right] + o(1).$$

By the same way

$$\begin{aligned}
n^{-1}\mathbb{E}\left[H'^2(Z_1,Z_2)\right] &= \frac{1}{n\phi^2(h)}\mathbb{E}\left[K_{12}^4\mathbb{E}\left[u_2^4|X_2\right]\mathbb{E}\left[u_1^4|X_1\right]\right] \\
&\leq \frac{C}{n\phi^2(h)}\int\int_{B(x_2,h)} dP_{X_1}(x_1)dP_{X_2}(x_2) \\
&\leq \frac{C}{n\phi(h)}\int f(x)dP_X(x) \to 0.
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[H(Z_1,Z_2)\right] &= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}u_1u_2\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2u_1|X_2,X_1\right]|\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}\left[u_2|X_2\right]\mathbb{E}\left[u_1|X_1\right]\right] \\
&= \frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2|X_2\right]\right] + o\left(\frac{1}{\phi(h)}\mathbb{E}\left[K_{12}^2\mathbb{E}^2\left[u_2|X_2\right]\right]\right) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\int f(x_2)\mathbb{E}^2\left[u_2|x_2\right]dP_{X_2}(x_2) + o(1) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\mathbb{E}^2\left[u|X\right]\right] + o(1) \\
&= \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon|X\right]-\sigma^2)^2\right] + o(1)
\end{aligned}$$

So, from Lemma (3.4.2) we get

$$W_n \longrightarrow \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon_2|X\right]-\sigma^2)^2\right] \quad \text{In probability.}$$

and

$$\widehat{s}^2 \xrightarrow{P} 2\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right] + o(1)$$

Consequently

$$\frac{T_n}{n\phi(h)} \xrightarrow{P} \frac{\left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)(V\left[\epsilon_2|X\right] - \sigma^2)^2\right]}{2\left(K^2(1) - \int_0^1 (K^2)'(s)\tau(s)ds\right)\mathbb{E}\left[f(X)\left(V\left[\epsilon^2|X\right] + (V[\epsilon|X] - \sigma^2)^2\right)^2\right]}$$

Proof of Corollary3.4

Like the previous proofs, by using the same decomposition of Theorem 3.7, we can show that

$$W_n = W_{1n} + o((n\phi^{1/2}(h))^{-1})$$

Thence, let we assume that $u_i' = u_i - \delta_n g(X_i)$, thus, W_{1n} can be decompose like

$$\begin{aligned}
W_{1n} &= \frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i u_j \\
&= \frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}(u_i' + \delta_n g(X_i))(u_j' + \delta_n g(X_j)) \\
&= \frac{1}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i' u_j' \\
&\quad + \frac{2\delta_n}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}u_i' g(X_j) \\
&\quad + \frac{\delta_n^2}{n(n-1)\phi(h)}\sum_{i=1}^{n}\sum_{j\neq i} K_{ij}g(X_i)g(X_j) \\
&= U_{11n} + 2\delta_n U_{12n} + \delta_n^2 U_{13n}.
\end{aligned}$$

We can clearly see that $\mathbb{E}[u_i'|X_i] = 0$, for this reason we can treat U_{11n} as degenerate U-statistic similarly to (11). Therefore,

$$n\sqrt{\phi(h)}U_{11n} \rightarrow N(0, s^2).$$

And by the same argument as those used in the consistency of $\widehat{s}^2$ we get for $\delta_n = n^{-1/2}\phi^{-1/4}(h)$:

$$n\sqrt{\phi(h)}\delta_n^2 U_{13n} = U_{13n} \to \left(K(1) - \int_0^1 (K)'(s)\tau(s)ds\right)\mathbb{E}[g^2(X)f(X)], \quad \text{In probability}$$

and

$$n\sqrt{\phi(h)}\delta_n U_{12n} = \sqrt{n}(\phi(h))^{1/4}U_{12n} \to 0, \quad \text{In probability.}$$

Finally, to conclude our proof, we just have to apply the Slutsky lemma.

Bibliography

[1] Breusch, T. S. and Pagan, A. R.(1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1294.

[2] Brumback. B. A, Rice. J. A. (1998). Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves. Journal of the American Statistical Association. 93 961-976.

[3] Cardot, H., Goia, A. et Sarda, P. (2004). Testing for no efect in functional linear regression models, some computational approaches. *Comm. Statist. Simulation Comput.*, **33** , 179-199.

[4] Cardot. H, Ferraty. F, Mas. A, Sarda. P.(2003). Testing Hypotheses in the Functional Linear Model. Scandinavian Journal of Statistics. 30 241-255.

[5] Chiou. J-M, Müller. H-G.(2009). Modeling hazard rates as functional data for the analysis of cohort lifetables and mortality forecasting. Journal of the American Statistical Association. 104 572-585.

[6] Chiou, J.M. and Müller H.-G. (2007). Diagnostics for functional regression via residual processes Computational Statistics and Data Analysis 51, (10) 4849- 4863.

[7] Cuevas, A., Febrero, M., and Fraiman, R. (2004). An anova test for functional 575 data. *Computational statistics and data analysis*, **47** , 111-122.

[8] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference*, **147**, 1-23.

[9] Delsol, L. (2013). No effect tests in regression on functional variable and some applications to spectrometric studies.*Comput. Statist.*, **28** , 1775–1811.

[10] Diblasi, A. and Bowman, A (1997). Testing for constant variance in a linear model. *Statistics and Probability Letters*, **33**, 95-103.

[11] Ezzahrioui, M., Ould-Saïd, E. (2008). Asymptotic normality of nonparametric estimator of the conditional mode for functional data.*J. of Nonparametric Statist.*, **20**, 3-18.

[12] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice.* Springer-Verlag, New York, 2006.

[13] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plann. Inference, 140, 335- 352.

[14] Gadiaga. D, R. Ignaccolo. (2005). Test of No-effect Hypothesis. Afrika Statistika, 1 67-76.

[15] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *J. Mutiv. Anal.*, **146**, 1-6.

[16] Henien. A, Laksaci. A, Rachdi. M, Ait-Hennani. L, Demongeot. J.(2018). Test d'hétéroscédasticité quand les covariables sont fonctionnelles (Heteroscedasticity test when the covariables are functionals).Comptes Rendus Mathematique. (accepted).

[17] Hilgert. N, Mas. A, Verzelen. N. (2013). Minimax adaptive tests for the functional linear model. The Annals of Statistics. 41 838-869.

[18] Jennrich. RI.(1969). Asymptotic Properties of Non-Linear Least Squares Estimators. The Annals of Mathematical Statistics. 40 633-643.

[19] Lin, J. G. and Qu, X. Y. (2012). A consistent test for heteroscedasticity in semiparametric regression with nonparametric variance function based on the kernel method, *Statistics*, **46**, 565-576.

[20] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. Stoch. Proc. and their Appl., 115, 155-177.

[21] Müller. H-G, Stadtmüller. U (2005). Generalized functional linear models. The Annals of Statistics. 33 774-805.

[22] Staicu, A., Li, Y. Crainiceanu, C. M. and Ruppert, D. (2014). Likelihood ratio tests for dependent data with applications to longitudinal and functional data analysis. *Scand. J. Stat.*, **41**, 932-949.

[23] White, H.(1981). Consequences and detection of misspecified nonlinear regression models. Journal of the American Statistical Association. 76 419-433.

[24] White, H.(1982). Maximum likelihood estimation of misspecified models. Econometrics. 50 1-25.

[25] Zhang, J.-T., Liang, X., and Xiao, S. (2010). On the two-sample behrens-fisher problem for functional data. *Journal of Statistical Theory and Practice*, **4** , 571-587.

[26] Zheng, J. X. (2009). Testing heteroscedasticity in nonlinear and nonparametric regressions, Canadian Journal of Statistics, **37**, 282-300.

[27] Zheng, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics, **75**, 263-289.

Chapter 4: Conclusions

In this final chapter, we resume the contributions of this dissertation and the possible impact as we see it and we talk over the main directions of the prospective work.

Contents

4.1 Summary of the thesis

The work presented in this thesis provide new theoretical results on the **heteroscedasticity** test in the context of functional statistical spaces. Indeed, where the explanatory variable valued in an infinite dimensional space. Thus, the proposed tests are innovative in this context. Whereas, they are consistent either at nonlinear or nonparametric regression models. Actually, our tests are based on the nonparametric estimation techniques and constructed by evaluating the difference between the conditional variance and unconditional variance. They are allowed to test if the variance of the error terms is constant which means the presence of homoscedasticity. In addition, we proved that those tests are also consistent against all deviations from the null hypothesis.

We point out that, our work can be considered as a generalization to infinite dimensional of some results established in multivariate case. As far as we know, we have been the first who established this results, meaning that, testing **heteroscedasticity** in functional statistics. It should be noted that, in order to avoid the problem of the curse of dimensionality in multivariate case, some standard conditions in nonparametric functional statistics are used through this work. Therefore, we have established in Chapter 1 and Chapter 2 the asymptotic behavior of the constructed tests statistics under the homoscedasticity (resp. alternative) hypothesis. In addition, the corollary 5 and 6 are introduced to examine the robustness of the tests against all possible departures from homoscedasticity.

Finally, for practical purposes, some simulated data examples are presented in Chapter 2, to evaluate the performance of the developed test. Whereas, this type of test statistics is simpler to calculate due to the less of parameters, the bandwidth needs to be chosen. Moreover, it has more advantage because of the nonparametric form which is better than the parametric one thus it is not consistent against all deviation from the null hypothesis.

4.2 Direction for future work

In the desire to continue our work, we show in what be come below, the extent of the subject studied and its implication in other statistical problems. Thus, we present some interesting prospects and open questions of this thesis.

- We have interested in our prediction problem on scalar response given explanatory functional random variable, which is the most preferred in this context, hence, it could be fascinating to extending our results to the case of functional response.

- **On another conditional models :**

  To construct our hereroscedasticity tests, we based on the conditional variance. While, there are another conditional quantities like the mode, density,quantiles,... that it would be interesting to be used.

- **The dependent case :**

  In this thesis, our results have been obtained for the independent case. For the next work, it could be interesting to study an other cases as α-mixing or β-mixing variables. The literature on dependent functional kernel estimators has been fairly well developed (see ***Masry*** [2005] or ***Ferraty and Vieu*** [2006]). In the other hand, this generalization can pose the problem of bootstrapping dependent variables (see for instance ***Politis and Romano*** [1994]).

- Additionally, it would be an interesting challenge to construct a specific semi-metrics suitable to the image study. Furthermore, we can trying to use the results to study a sequence of independent (resp; dependent) images.

- **Spatial prediction :**

  We suppose that we want to predict a real characteristic of the Spatio-Temporal process $Z_{t,s\in I\times S\subset R^3}$ (where t time and s the geographical position) in a future time t_0 to a fixed zone s_0. It could be interesting to construct a test statistics to detect the heteroscedasticity for the data by using **Spatio-Temporal** variables.

- **Another estimation methods :**

  we based on the kernel estimation to construct our heteroscedasticity tests, it seems possible to used an others estimation methods as:

  -**Robust estimation :** This estimation method has been extensively studied in vectorial statistics. We refer to***Laksaci et al***[2008] and ***Attouch***[2009] for the function case.

  -**The local linear estimation :** this method has been introduced for functional statistics not long ago, in 2009 by ***Baillo and Grané.***. We can name for example: ***Demongeot et al***[2011, 2013, 2014 and 2016] for the analysis of the estimator in the conditional distribution framework or functional response variable, ***Wang et al***[2015] for the dependent data. The tools developed in this thesis should be able to adapt to this estimation methods.

- **The semi-parametric models :**

  In addition to the models that we focused on in our work, there have an other interesting models, the semi-parametric models as partially linear models, to extend our idea (see ***Aneiros- Perez and Vieu***[2006, 2008] or ***Lian***[2011]), the single-index functional models (see ***Goia and Vieu***[2015]) or the projection models (see ***Chen and collab)***[2011]), which recently, this field take a wide interest on functional data analysis.

Bibliography

[1] Antoniadis, A. et Sapatinas, T. (2007) Estimation and inference in functional mixed-effect models Comp. Stat. and Data Anal. 51 (10) 4793-4813.

[2] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data. *Computational statistics and data analysis*, **22**, 353-369.

[3] Bickel, P. L. (1978). Using residuals robustly. I. Tests for heteroscedasticity and nonlinearity. Annals of Statistics, 6, 266-291

[4] Bongiorno E., Salinelli, E., Goia, A. and Vieu, P. (Eds.) (2014). An overview of IWFOS'2014. In *Contributions in infinite-dimensional statistics and related topics*, 1-6, Esculapio, Bologna.

[5] Boomsma. DI , PCM. Molenaar. (1987). The genetic analysis of repeated measures. I. Simplex models. Behavior genetics. 17 111â123.

[6] Borggaard. C, HH. Thodberg. (1992). Optimal minimal neural interpretation of spectra. Analytical chemistry. 64 545â551.

[7] Breusch, T. S. and Pagan, A. R.(1979) A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1294.

[8] Brumback. B. A, Rice. J. A. (1998). Smoothing Spline Models for the Analysis of Nested and Crossed Samples of Curves. Journal of the American Statistical Association. 93 961-976.

[9] Buell. CE. (1971). Integral equation representation for factor analysis. Journal of the Atmospheric Sciences.

[10] Cardot, H., Goia, A. et Sarda, P. (2004) Testing for no efect in functional linear regression models, some computational approaches. *Comm. Statist. Simulation Comput.*, **33** , 179-199.

[11] Cardot, H., Crambes, C., Kneip, A. and Sarda, P (2007) Smoothing splines estimators in functional linear regression with errors-in-variables. Computational Statistics and Data Analysis, special issue on functional data analysis 51 (10) 4832-4848.

[12] Cardot. H, Ferraty. F, Mas. A, Sarda. P.(2003). Testing Hypotheses in the Functional Linear Model. Scandinavian Journal of Statistics. 30 241-255.

[13] Chiou, J.M. and Müller H.-G. (2007) Diagnostics for functional regression via residual processes Computational Statistics and Data Analysis 51, (10) 4849- 4863.

[14] Chiou. J-M, Müller. H-G.(2009). Modeling hazard rates as functional data for the analysis of cohort lifetables and mortality forecasting. Journal of the American Statistical Association. 104 572-585.

[15] Cook, R. D and Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression, Biometrika, 70, 1-10.

[16] Crainiceanu, Ciprian M., A.M Staicu ,S. Ray , N. Punjabi (2012) Bootstrap-based inference on the difference in the means of two correlated functional processes. 31 3223-3240.

[17] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference*, **147**, 1-23.

[18] Cuevas, A., Febrero, M., and Fraiman, R. (2004). An anova test for functional 575 data. *Computational statistics and data analysis*, **47** , 111-122.

[19] Dabo-Niang, S. (2002). On Density estimation in an infinite-dimensional space : application to diffusion processes] PhD Paris VI.

[20] Dabo-Niang, S. and Rhomari, N. (2003) Kernel regression estimation when the regressor takes values in metric space, C. R. Math. Acad. Sci. Paris 336 (1) 75â80.

[21] Delsol, L. (2013). No effect tests in regression on functional variable and some applications to spectrometric studies.*Comput. Statist.*, **28** , 1775–1811.

[22] Dette, H. and Munk, A. (1998). Testing heteroscedasticity in nonparametric regression. *Journal of the Royal Statistical Society,* Series B,**60**, 693-708.

[23] Dette, H., Neumeyer, N. and Van Keilegom, I. (2007). A new test for the parametric form of the variance function in nonparametric regression, Journal of the Royal Statistical Society: Series B, 69(5), 903-917.

[24] Dette, H.(2002). A consistent test for heteroscedasticity in nonparametric regression based on the kernel method, Journal of Statistical Planning and Inference, 103, 311-329.

[25] Deville. JC. (1974). MÃ©thodes statistiques et numÃ©riques de l'analyse harmonique. Annales de l'INSEE. 3, 5-101.

[26] Diblasi, A. and Bowman, A (1997). Testing for constant variance in a linear model. *Statistics and Probability Letters,* **33**, 95-103.

[27] Eubank, R. L. and Thomas, W.(1993). Detecting heteroscedasticity in nonparametric regression, Journal of the Royal Statistical Society: Series B, 55(1),145-155.

[28] Ezzahrioui, M., Ould-Saïd, E. (2008). Asymptotic normality of nonparametric estimator of the conditional mode for functional data.*J. of Nonparametric Statist.*, **20**, 3-18.

[29] Fan. J, Lin, SK,(1998). Test of significance when data are curves, Journal of the American Statistical Association 93 1007-1021.

[30] Ferraty, F., Mas, A. and P. Vieu (2007). Nonparametric regression on functional data: inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**, 267-286.

[31] Ferraty. F.(2011). Recent advances in functional data analysis and related topics. book

[32] Ferraty F. et Vieu P. (2000). Dimension fractale et estimation de la rÃ©gression dans des espaces vectoriels semi-normÃ©s. Comptes Rendus de lâAcadÃ©mie des Sciences- Series I-Mathematics, 330(2) :139â142.

[33] Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist., 16, 111-127.

[34] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference*, **140**, 335-352.

[35] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer-Verlag, New York, 2006.

[36] Frank. I.E. et Friedman. J.H. (1993). A statistical view of some chemometrics regression tools. Technometrics 35 109-135.

[37] Gadiaga. D, R. Ignaccolo. (2005) Test of No-effect Hypothesis. Afrika Statistika, 1 67-76.

[38] Gasser.T, Hall.P, Presnell. B. (1998). Nonparametric estimation of the mode of a distribution of random curves. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 60(4) :681â691.

[39] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *J. Mutiv. Anal.*, **146**, 1-6.

[40] Goldfeld. S. and R.Quandt. (1965). Some Tests for Homoscedasticity. Journal of the American Statistical Association. 60 539-547

[41] Glejser. H.(1969). A New Test for Heteroskedasticity. Journal of the American Statistical Association. 64 316-323.

[42] Grenander. U.(1950). Stochastic processes and statistical inference.Arkiv fÃĪr matematik. 1 195â277.

[43] Hall, P. et Vial, C. (2006a) Assessing extrema of empirical principal component functions. Ann. Statist. 34 (3) 1518-1544.

[44] Harvey. A.C., G.D.A. Phillips. (1974). A comparison of the power of some tests for heteroskedasticity in the general linear model. Journal of Econometrics. 2 307-316.

[45] Harvey.A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. Econometrica. 44 461-465.

[46] Hastie, T. et Mallows, C. (1993) Discussion of âA statistical view of some chemometrics regression tools.â by Frank, I.E. and Friedman, J.H. Technometrics 35 140-143.

[47] Henien. A, Laksaci. A, Rachdi. M, Ait-Hennani. L, Demongeot. J.(2018). Test d'hÃ©tÃ©roscÃ©dasticitÃ© quand les covariables sont fonctionnelles (Heteroscedasticity test when the covariables are functionals).Comptes Rendus Mathematique. (accepted).

[48] Hilgert. N, Mas. A, Verzelen. N. (2013). Minimax adaptive tests for the functional linear model. The Annals of Statistics. 41 838-869.

[49] Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer Series in Statistics. Springer, New York.

[50] Hsing, T. and Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester.

[51] Hotelling. H, MR. Pabst. (1936). Rank correlation and tests of significance involving no assumption of normality. The Annals of Mathematical Statistics. 7 29-43.

[52] James. G.M. et Sood, A. (2006) Performing hypothesis tests on the shape of functional data. (English summary) Comp. Stat. and Data Anal. 50 (7) 1774-1792.

[53] Jennrich. RI.(1969). Asymptotic Properties of Non-Linear Least Squares Estimators. The Annals of Mathematical Statistics. 40 633-643.

[54] Karaa, L-Z. Laksacia, A, Rachdi, M. and Vieu, P. (2015) Uniform in bandwidth consistency results for kernel estimates in various statistical models with functional. *J. Nonparametr. Statist.* (accepted).

[55] Kirkpatrick. M, N. Heckman. (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. Journal of mathematical biology. 27 429â450.

[56] Koenker, B. and Bassett, G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, **50**, 43-61.

[57] Kruskal. W.H, Wallis. W.A.(1952).Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association. 47 583-621.

[58] Lin, J. G. and Qu, X. Y. (2012). A consistent test for heteroscedasticity in semiparametric regression with nonparametric variance function based on the kernel method, *Statistics*, **46**, 565-576.

[59] Mann. H. B, D. R. Whitney. (1947). on a test whether one of two random variables is stochastically larger than the other .The Annals of Mathematical Statistics. 18 50-60.

[60] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.

[61] Mas. A. (2007a) Testing for the mean of random curves : a penalization approach Statistical Inference for Stochastic Processes 10 (2) 147-163.

[62] Mas, A. (2000) Estimation dâopÂ´erateurs de corrÂ´elation de processus fonctionnels : lois limites, tests, dÂ´eviations modÂ´erÂ´ees Th‘ese de doctorat de lâuniversitÂ´e de Paris 6.

[63])Müller, H. G. and Zhao, P.L. (1995). On a semi-parametric variance function model and a test for heteroscedasticity, Annals of Statistics, 23, 946-967.

[64] Müller. H-G, Stadtmüller. U (2005). Generalized functional linear models. The Annals of Statistics. 33 774-805.

[65] Ramsay, J. and Dalzell, C. (1991) Some tools for functional data analysis J. R. Statist. Soc. B. 53 539-572.

[66] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, New York.

[67] Ramsey. J. B.(1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. Journal of the Royal Statistical Society. Series B (Methodological). 31 350-371.

[68] Rao. CR.(1958). Some statistical methods for comparison of growth curves. Biometrics. 14 1-17.

[69] Rutemiller. H C, David A. Bowers. (1968). Estimation in a Heteroscedastic Regression Model. Journal of the American Statistical Association. 63 552-557.

[70] Savage, I.R. (1962). Bibliography of nonparametric statistics. Book.

[71] Shen, Q. et Faraway, J. (2004) An F test for linear models with functional responses. (English summary) Statist. Sinica 14 (4) 1239-1257.

[72] Siegel. S. (1956). Nonparametric statistics. for the behavioral sciences. book

[73] Simonoff, J. S. and Tsai, C. L. (1994). Improved tests for nonconstant variance in regression based on the modified profile likelihood, Journal of Appmied Statistics, 43, 357-370.

[74] Staicu. AM, Y. Li, C.M. Crainiceanu , D. Ruppert.(2014) Likelihood Ratio Tests for Dependent Data with Applications to Longitudinal and Functional Data Analysis. 41 932-949.

[75] Stute, W. and Zhu, L. X. (2005). Nnonparametric checks for single-index models, Annals of Statistics, 33, 1048-1083.

[76] Tsai, C. L. (1986). Score test for the first-order autoregressive with heterosedasticity, Biometrika, 73, 61-70.

[77] Tucker. LR.(1958). Determination of parameters of a functional relation by factor analysis. Psychometrika. 23 19â23.

[78] Viele, K. (2001) Evaluating fit in functional data analysis using model embeddings. (English, French summary) Canad. J. Statist. 29 (1) 51-66.

[79] White. H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica. 48 817-838.

[80] White, H.(1981). Consequences and detection of misspecified nonlinear regression models. Journal of the American Statistical Association. 76 419-433.

[81] White, H.(1982). Maximum likelihood estimation of misspecified models. Econometrics. 50 1-25.

[82] Wilcoxon. F. (1945). Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1 80-83.

[83] You, J. H. and Chen, G. M. (2005). Testing heteroscedasticity in partially linear regression models, Statistics and Probability Letters, 73, 61-70.

[84] Zhang, J.-T. et Chen, J. (2007) Statistical inferences for functional data. Ann. Statist. 35 (3) 1052-1079.

[85] Zhang, J.-T. et Chen, J. (2010) Statistical inferences for functional data. Ann. Statist. 35 (3) 1052-1079.

[86] Zhang, J. (2013). *Analysis of variance for functional data*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

[87] Zheng, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. Journal of Econometrics, **75**, 263-289.

[88] Zheng, J. X. (2009). Testing heteroscedasticity in nonlinear and nonparametric regressions, Canadian Journal of Statistics, **37**, 282-300.

[89] Zhu, L. X., Fujikoshi, Y. and Naito, K. (2001). Heteroscedasticity checks for regression models, Science in China (Series A), 44, 1236-1252.