

TESTS FOR EQUALITY OF CURVES VIA WAVELETS

CENTRE FOR NEWFOUNDLAND STUDIES

**TOTAL OF 10 PAGES ONLY
MAY BE XEROXED**

(Without Author's Permission)

PENGFEE GUO

TESTS FOR EQUALITY OF CURVES VIA WAVELETS

by

©Pengfei Guo

A thesis submitted to the
School of Graduate Studies
in partial fulfilment of the
requirements for degree of
Master of Science

Department of Mathematics and Statistics
Memorial University of Newfoundland

August 2005 Submitted



St. John's

Newfoundland

Canada

Abstract

In this article, we develop tests for equality of response curves. We assume that the mathematical expression of the response curves are unknown. Consequently, we exploit the adaptive properties of wavelets to construct a wavelet representation of the curves. Then, we develop an approach which combines the False Discovery Rate (FDR) technique and the universal thresholding approach, widely used in data denoising, for detecting differences between the curves. We also discuss some methods based on the general F-test. We consider several examples under a variety of conditions such as unequal variances, unequal number of observations, and distinct design points. The size and power performances of the tests are reported. We present two real examples at the end of this article.

Acknowledgement

I would like to express my sincere gratitude to Dr. A. J. Oyet for granting me the privilege to write this thesis under his supervision and for his support, both financial and otherwise, through out the course of my program. He patiently read through my work several times and offered his very valuable suggestions. This work would have been impossible without his guidance and help.

I thank my professors who have shared with me their wealth of statistical knowledge at various points in the course of my program. The staff of the Math. and Stat. general office were very helpful and nice.

Also, my sincere thanks to my parents, friends and neighbours whose assistance and encouragement have seen me to the completion of my program.

Contents

1	INTRODUCTION	1
1.1	Some Background on Wavelets	3
1.1.1	Wavelet system construction	6
1.1.2	Some important wavelet bases	11
1.2	Some Wavelet Methods for Estimating Functions	15
1.2.1	Wavelet Expansion	15
1.2.2	Discrete Wavelet Transformation (DWT)	15
1.3	Large sample properties of \hat{f}	22
2	FALSE DISCOVERY RATE (FDR) STATISTIC	23
2.1	Test Statistic Construction	25
2.2	Simulation Studies	27
2.3	Robustness of FDR test	34
3	WAVELET EXPANSION STATISTIC	37
3.1	Wavelet Expansion Method for Equal Sample Sizes	37
3.1.1	Construction of F_1 statistic	37
3.1.2	Simulation studies	39
3.1.3	Robustness of F_1 test	43

3.2	Generalized Wavelet Expansion Method	44
3.2.1	Construction of F_2 statistic	44
3.2.2	Simulation Studies	50
3.2.3	Robustness of F_2 test	56
4	APPLICATIONS AND CONCLUDING REMARKS	62
4.1	Application : Two Case Studies	62
4.1.1	Case Studies: Methacholine CRC in Mesenteric Arteries	62
4.1.2	Case Studies: Mouse Lymphoma Assay	68
4.2	Conclusion and Remarks	70

List of Tables

1.1	The filter coefficients	14
2.1	Number of errors committed when testing m null hypotheses	24
2.2	Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance level $q = 0.01, 0.025$ and 0.05	29
2.3	Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance level $q = 0.01, 0.025$ and 0.05	30
2.4	Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance level $q = 0.01, 0.025$ and 0.05	34
2.5	Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance level $q = 0.01, 0.025$ and 0.05	35
2.6	Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance level $q = 0.01, 0.025$ and 0.05	35
2.7	Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance level $q = 0.01, 0.025$ and 0.05	36
3.1	Size of F_1 test based on 1000 replications, the Haar wavelet and $m = 0$ with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05	40

3.2	Power of the F_1 test based on 1000 replications and the Haar ,Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05	41
3.3	Size of F_1 test based on 1000 replications, the Haar wavelets and $m = 0$ with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance levels $q = 0.01, 0.025$ and 0.05	44
3.4	Power of the F_1 test based on 1000 replications and the Haar, Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance levels $q = 0.01, 0.025$ and 0.05	45
3.5	Size of F_1 test based on 1000 replications, the Haar wavelets and $m = 0$ with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon(\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu(\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05	46
3.6	Power of the F_1 test based on 1000 replications and the Haar, Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon(\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu(\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05	47
3.7	Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05	52
3.8	Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05	54
3.9	Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$. at significance levels $q = 0.01, 0.025$ and 0.05	57

3.10	Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$. at significance levels $q = 0.01, 0.025$ and 0.05	58
3.11	Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05	59
3.12	Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05	60
4.1	Results of FDR test for CRC data at significance level $q = 0.01, 0.025$ and 0.05	65
4.2	Results of F_1 test for CRC data with Haar, Daub2, Daub4 and $m = 0$ and 1 at significance level $q = 0.05$	68
4.3	Results of F_2 test for mouse lymphoma assay data with Haar, Daub2, Daub4 and $m = 0, 1$ and 2 at significance level $q = 0.05$	70
4.4	Size of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance levels $q = 0.01$ and 0.05	72
4.5	Power of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$	73
4.6	Power of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$	74
4.7	Size of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$, $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance levels $q = 0.01$ and 0.05	75
4.8	Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$, $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$	76

- 4.9 Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$, $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$ 77
- 4.10 Size of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance levels $q = 0.01$ and 0.05 78
- 4.11 Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$. 79
- 4.12 Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$. 80

List of Figures

1.1	Graphs of scaling functions and primary wavelet from Daubechies' family, $N = 2, 3, 4, 5$ and 8	13
2.1	A plot of (a). the quadratic function $3(x - 0.4)^2$; and (b). the HeaviSine function, $d_3(x)$	28
2.2	A plot of the data $D_i = Y_i - Z_i$ $i = 1, \dots, n$ (solid line), Haar wavelet estimate $\hat{D} = \hat{h}(x) = f \hat{-} g(x)$ (dotted line), and approximate confidence bounds (dotted broken line) for difference $h(x) = f(x) - g(x)$ with parameters (a), (b) $n = 8$ and 16 , respectively, with $\alpha = 0.05$, $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$ and $h(x) = 0$; (c), (d) $n=8$ and 16 , respectively, with $\alpha = 0.025$, $(\sigma_\epsilon, \sigma_\nu) = (3.5, 0.8)$, and $h(x) = 0.5$; (e), (f) $n = 8$ and 16 , respectively with $\alpha = 0.01$, $(\sigma_\epsilon, \sigma_\nu) = (2, 2)$, and $h(x) = \sqrt{1.5}$	32
2.3	A plot of the data $D_i = Y_i - Z_i$ $i = 1, \dots, n$ (solid line), Haar wavelet estimate $\hat{D} = \hat{h}(x) = f \hat{-} g(x)$ (dotted line), and approximate confidence bounds (dotted broken line) for difference $h(x) = f(x) - g(x)$ with parameters (a), (b), (c) $n = 8$, $h(x) = 3(x - 0.4)^2$, and $n = 8$, $h(x) = 3(x - 0.4)^2 + 2.5$, and $n = 16$, $h(x) = 3(x - 0.4)^2 + 2.5$, respectively, with $\alpha = 0.05$, $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$ and Filter = $_4h_k$; (d), (e), (f) $n = 8$, Filter = $_4h_k$ and $n = 8$, Filter = $_8h_k$ and $n = 16$, Filter = $_8h_k$ respectively, with $h(x) = \text{HeaviSine function}$, $\alpha = 0.05$, and $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$	33

4.1	Methacnoline CRC in mesenteric arteries from 4 months(a) and 7 months(b) male rats: Solid line: Observations without L-NAME; dotted line: Observations with L-NAME.	63
4.2	Confidence intervals of $\hat{h}(x)$ for CRC data. (a),(b) and (c): for 4 months; (d), (e) and (f): for 7 months. (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used. Solid line: Observed difference; dotted line: Estimated difference; dashed lines: bounds of 95% confidence interval of $\hat{h}(x)$	64
4.3	CRC Four months data. Solid line: Observed difference $h(x)$; dotted line: Estimated difference $\hat{h}(x)$. (a), (b) and (c): $m = 0$; (d), (e) and (f): $m = 1$; (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used.	66
4.4	CRC Seven months data. Solid line: Observed difference $h(x)$; dotted line: Estimated difference $\hat{h}(x)$. (a), (b) and (c): $m = 0$; (d), (e) and (f): $m = 1$; (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used.	67
4.5	Mouse Lymphoma Assay data with substance A and B	69
4.6	Mouse Lymphoma Assay data with substance A and B: (a), (b) and (c): Haar wavelet is used; (d), (e) and (f): Daub2 wavelet is used. (g), (h) and (i): Daub4 wavelet is used. (a), (d) and (g): $m = 0$; (b), (e) and (h): $m = 1$; (c), (f) and (i): $m = 2$	71

Chapter 1

INTRODUCTION

A common problem in experimental work, such as toxicology and epidemiology, is the comparison of two mean response curves. This comparison becomes complicated when the response functions are unknown.

This problem has been already investigated by a few authors using smooth nonparametric estimates of the regression curve. See Härdle and Marron (1990), Hall and Hart (1990), and King, Hart and Wehrly (1991)). Hall and Hart (1990) discussed a bootstrap procedure for testing the equality of two curves. Under the assumptions that sample sizes are equal ($n_1 = n_2$) and design points are the same ($x_i = u_i$). King et al. (1991) proposed two tests both with common design points, one for normal errors and the other for non-normal errors. It is difficult to justify the asymptotic properties of these statistics, since the statistics which are based on automatically chosen smoothing numbers are computationally demanding. Another restriction of the tests mentioned above is the requirement of equal sample size and same design points, which may be difficult to obtain in practice. In addition, computing p -values for these tests requires several thousands of simulation which an experimenter may not have the time to complete. Consequently, we feel that it is im-

portant to investigate new methods that are relatively easy to implement with power that is comparable to these methods.

Härdle and Marron (1990) propose a semiparametric test. The test involves the method of parametric transformation of axes. They also study estimation and testing of the parameters in the transformations. Delgado (1993) proposes a test which resembles in spirit to the Kolmogorov-Smirnov Statistic. Similar to smoothing estimates methods, these two tests both assume equal designs.

In this study, we propose and investigate three wavelet-based procedures for testing the equality of curves. The first two methods requires equal sample sizes and same design points; whereas the third is a general procedure with no restrictions. In what follows, readers will find that the third procedure requires a larger sample size to achieve high power.

Wavelet theory, which has been developing over the years has proved to be useful in signal processing, fast algorithms for integral transforms in numerical analysis and function representation. For a recent survey on the use of wavelets in signal processing, see Rioul & Vetterli (1991). Due to the flexible nature of wavelet systems, there has been growing interest in the application of wavelets to statistical problems. Successful statistical applications have already been made in areas such as regression, density and function estimation, factor analysis, modeling and forecasting in time series analysis, and spatial statistics.

The problem we consider can be described as follows. The observed data are $\{(Y_i, x_i), i = 1, \dots, n_1\}$ and $\{(Z_j, u_j), j = 1, \dots, n_2\}$ with

$$\begin{aligned} Y_i &= f(x_i) + \epsilon_i, \quad i = 1, \dots, n_1, \\ Z_j &= g(u_j) + \nu_j, \quad j = 1, \dots, n_2, \end{aligned} \tag{1.1}$$

where $\epsilon_i, i = 1, \dots, n_1$ and $\nu_j, j = 1, \dots, n_2$ are independent random errors for the two groups. We assume that in each group the errors are identically distributed with mean 0,

$var(\epsilon) = \sigma_\epsilon^2$ and $var(\nu) = \sigma_\nu^2$, where σ_ϵ^2 is not necessarily equal to σ_ν^2 . However, the ϵ_i 's are assumed to be independent of the ν_j 's. Based on the observed data, an experimenter is interested in determining whether the underlying mean response functions generating Y_i and Z_j are the same. That is, the null and alternative hypotheses of interest to the experimenter are

$$\begin{aligned} H_0 : f(x) &= g(x) \quad \text{for all } x \in [0, 1], \\ H_1 : f(x) &\neq g(x) \quad \text{for some } x \in [0, 1]. \end{aligned} \tag{1.2}$$

In the sections that follow, we discuss some wavelet basics needed for the understanding of the contents of Chapters 2 and 3. It is in these chapters that we develop the test statistics for (1.2). The test statistic in Chapter 2 arises from the false discovery rate(FDR) approach of Abramovich and Benjamini (1995), and the statistic in Chapter 3 is based on wavelet expansion. In the second part of Chapter 3, we generalize the wavelet expansion methods, so that the statistics can be used for the test under any condition, such as unequal sample sizes, different design points and repeated measurements. Two examples the methacholine CRC in mesenteric arteries and the mouse lymphoma assay are used to illustrate the techniques in Chapter 4.

1.1 Some Background on Wavelets

In this section we introduce some definitions and theories on wavelets relevant to our work. More detailed discussions can be found in Mallat (1989), Meyer (1992), Daubechies (1992), Abramovich and Benjamini (1995),(1996), Benjamini and Hochberg (1995), Härdle, Kerkyacharian, Picard, and Tsybakov (1998), and Vidakovic(1998).

A wavelet system is a collection of dilated and translated versions of a scaling function

$\phi(x)$ and a primary wavelet $\psi(x)$ defined by

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$$

and

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), j, k \in \mathbb{Z}$$

respectively. The function $\phi(x)$ and $\psi(x)$ are chosen to satisfy the equations

$$\phi(x) = \sum_{p \in \mathbb{Z}} h_p \phi(2x - p) \quad (1.3)$$

$$\psi(x) = \sum_{r \in \mathbb{Z}} g_r \phi(2x - r) \quad (1.4)$$

and

$$g_r = (-1)^r h_{-r+1} \quad (1.5)$$

for a sequence h_r of constants, called filter coefficients, with

$$\int \phi(x) dx = 1, \int \psi(x) dx = 0, \int \phi^2(x) dx = 1.$$

The condition

$$\sum_{p \in \mathbb{Z}} h_p = 2$$

ensures the existence of a unique solution to equations (1.3) and (1.4). Orthogonality of the translates of $\phi(x)$ is ensured by the condition

$$\sum_{p \in \mathbb{Z}} h_p h_{p-2j} = \delta_j, j \in \mathbb{Z}$$

In the theory of wavelets, the space of square integrable functions, $L_2(\mathbb{R})$, is written as the limit of a sequence of close subspaces V_j where

$$\dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots$$

The nested spaces have an intersection that is trivial and a union that is dense in $\mathbb{L}_2(\mathbb{R})$,

$$\bigcap_j V_j = 0, \overline{\bigcup_j V_j} = \mathbb{L}_2(\mathbb{R}) \quad (1.6)$$

where we denote the closure of a set A by \overline{A} .

Mallat(1989) introduced the notion of a multiresolution analysis, the definition of which we recall here.

Definition 1.1.1 *A multiresolution analysis of $\mathbb{L}_2(\mathbb{R})$ consists of an increasing sequence of closed subspaces V_j , $j \in \mathbb{Z}$ such that*

(a) $\bigcap_j V_j = 0$;

(b) $\overline{\bigcup_j V_j} = \mathbb{L}_2(\mathbb{R})$;

(c) *there exists a scaling function $\phi \in V_0$ such that $\phi(x - k)$, $k \in \mathbb{Z}$ is an orthonormal basis of V_0 ;*

(d) *for all $k \in \mathbb{Z}$, $f(x) \in V_j \iff f(x - k) \in V_j$, and*

(e) $f(x) \in V_j \iff f(2x) \in V_{j+1}$.

The intuitive meaning of (e) is that in passing from V_j to V_{j+1} , the resolution of the approximation is doubled. Mallat(1989) has shown that given any multiresolution analysis, it is possible to derive a function $\psi(x)$ such that the family $\{\psi_{j,k}(x) : j, k \in \mathbb{Z}\}$ is an orthonormal basis of $\mathbb{L}_2(\mathbb{R})$.

To construct $\psi_{j,k}(x)$, we define for each $j \in \mathbb{Z}$ the difference space W_j to be the orthogonal complement of V_j such that

$$W_j \oplus V_j = V_{j+1}, W_j \perp V_j.$$

That is, any function $f(x) \in V_{j+1}$ can be written as a linear combination or direct sum of functions in W_j and V_j . It can be verified that

$$V_j = V_0 \oplus \bigoplus_{i=0}^{j-1} W_i.$$

Iterating this infinitely many times, we find

$$\mathbb{L}_2(\mathbb{R}) = \bigcup_{j=0}^{\infty} V_j = V_0 \oplus \bigoplus_{i=0}^{\infty} W_i.$$

This means that any $f \in \mathbb{L}_2(\mathbb{R})$ can be represented as a series (convergent in $\mathbb{L}_2(\mathbb{R})$):

$$f(x) = \sum_{k \in \mathbb{Z}} c_{j_0 k} \phi_{j_0 k}(x) + \sum_{j=j_0}^{\infty} \sum_{k \in \mathbb{Z}} d_{jk} \psi_{jk}(x), \quad (1.7)$$

where $c_{j_0 k}$, d_{jk} are some coefficients, and $\{\psi_{jk}\}, k \in \mathbb{Z}$ is a basis for W_j . The relation (1.7) is called a multiresolution expansion of f . The space W_j is called resolution level of multiresolution analysis. In Fourier analysis there is only one resolution level. In multiresolution analysis there are many resolution levels which is the origin of its name.

1.1.1 Wavelet system construction

The general framework of wavelet system construction is as follows.

1. Pick a scaling function ϕ such that $\{\phi_{0k}\}$ is an orthonormal system, and (1.6) is satisfied.
2. Find a primary function $\psi \in W_0$, probably using (1.4), such that $\{\psi_{0k}, k \in \mathbb{Z}\} = \{\psi(x - k), k \in \mathbb{Z}\}$ is an orthonormal basis in W_j .
3. Conclude that any $f \in \mathbb{L}_2(\mathbb{Z})$ has the unique representation in terms of an \mathbb{L}_2 -convergent series:

$$f(x) = \sum_{k \in \mathbb{Z}} c_{j_0 k} \phi_{j_0 k}(x) + \sum_{j=j_0}^{\infty} \sum_{k \in \mathbb{Z}} d_{jk} \psi_{jk}(x),$$

where the wavelet coefficients are

$$c_{j_0k} = \int f(x)\overline{\phi_{j_0k}(x)}dx, d_{jk} = \int f(x)\overline{\psi_{jk}(x)}dx.$$

We now outline four constructions of the "scaling function" ϕ found in the literature (see Strang(1989) and Pinheiro and Vidakovic(1997)). Once $\phi(x)$ is known, we can compute the primary wavelet $\psi(x)$ through (1.4).

Construction 1. Iterate $\phi_j(x) = \sum h_k\phi_{j-1}(2x - k)$ with the box function as $\phi_0(x)$. When $h_0 = 2$, the boxes get taller and thinner, approximating the delta function. For $h_0 = h_1 = 1$, the box is invariant: $\phi_1 = \phi_0$. For coefficients $\frac{1}{2}, 1, \frac{1}{2}$, the hat function appears. And $\frac{1}{8}, \frac{4}{8}, \frac{6}{8}, \frac{4}{8}, \frac{1}{8}$ yields the cubic B-spline. An example that will be important in our discussion has coefficients $\frac{1}{4}(1 + \sqrt{3}), \frac{1}{4}(3 + \sqrt{3}), \frac{1}{4}(3 - \sqrt{3}), \frac{1}{4}(1 - \sqrt{3})$. This scaling function leads to orthogonal wavelets.

Construction 2. The second construction takes the Fourier transform of (1.3):

$$\begin{aligned} \hat{\phi}(\xi) &= \sum h_k \int \phi(2x - k)e^{i\xi x} dx \\ &= \frac{1}{2} \left(\sum h_k e^{ik\xi/2} \right) \int \phi(y)e^{iy\xi/2} dy \\ &= P\left(\frac{\xi}{2}\right) \hat{\phi}\left(\frac{\xi}{2}\right) \end{aligned} \tag{1.8}$$

The notation $P(\xi) = \frac{1}{2} \sum h_k e^{ik\xi}$ is a crucial function in this theory. With $\xi = 0$ we find $P(0) = 1$ (see(1.5)). Now repeat (1.8) at $\xi/2, \xi/4, \dots$ and recall $\hat{\phi}(0) = \int \phi(x) = 1$, we get an infinite product:

$$\hat{\phi}(\xi) = P\left(\frac{\xi}{2}\right) \hat{\phi}\left(\frac{\xi}{2}\right) = P\left(\frac{\xi}{2}\right) P\left(\frac{\xi}{4}\right) \hat{\phi}\left(\frac{\xi}{4}\right) = \dots = \prod_{j=1}^{\infty} P\left(\frac{\xi}{2^j}\right). \tag{1.9}$$

For $h_0 = 2$ we find $P \equiv 1$ and $\hat{\phi} \equiv 1$, the transform of the delta function. For $h_0 = h_1 = 1$, the product of the P 's are geometric series:

$$P\left(\frac{\xi}{2}\right) = P\left(\frac{\xi}{4}\right) = \frac{1}{4}(1 + e^{i\xi/2})(1 + e^{i\xi/4}) = \frac{1 - e^{i\xi}}{4(1 - e^{i\xi/4})}$$

As $N \rightarrow \infty$, this approaches the infinite product $(1 - e^{i\xi})(i\xi)$. This is $\int_0^1 e^{i\xi x} dx$, the transform of the box function. The hat function comes from squaring $P(\xi)$ which by (1.9) also squares $\hat{\phi}(\xi)$. The cubic B-spline comes from squaring again.

Construction 3. This construction of ϕ works directly with the recursion (1.3). Suppose ϕ is known at the integer $x = j$, the recursion (1.3) gives ϕ at the half-integers. Then it gives ϕ at the quarter-integers, and ultimately at all dyadic point $x = k/2^j$. This is fast to program.

The values of ϕ at the integers come from an eigenvector. With the wavelet coefficients $(h_0, h_1, h_2, h_3) = (\frac{1}{4}(1 + \sqrt{3}), \frac{1}{4}(3 + \sqrt{3}), \frac{1}{4}(3 - \sqrt{3}), \frac{1}{4}(1 - \sqrt{3}))$, which is one of the wavelet filters in the Daubechies wavelet system and noted as Daub2 in latter section. set $x = 1$ and $x = 2$ in the dilation equation (1.3) and use the fact that $\phi = 0$ unless $0 < x < 3$, we get:

$$\begin{aligned}\phi(1) &= \frac{1}{4}(3 + \sqrt{3})\phi(1) + \frac{1}{4}(1 + \sqrt{3})\phi(2), \\ \phi(2) &= \frac{1}{4}(1 - \sqrt{3})\phi(1) + \frac{1}{4}(3 - \sqrt{3})\phi(2).\end{aligned}$$

This is the eigenvalue problem $\phi = L\phi$, with matrix entries $L_{ij} = h_{2i-j}$. The eigenvalues are 1 and $\frac{1}{2}$, and the corresponding eigenvector for $\lambda = 1$ has components $\phi(1) = \frac{1}{2}(1 + \sqrt{3}), \phi(2) = \frac{1}{2}(1 - \sqrt{3})$. The other eigenvalue $\lambda = \frac{1}{2}$ means that the recursion can be differentiated: $\phi'(x) = \sum h_k 2\phi'(2x - k)$, which leads similarly to $\phi'(1)$ and $\phi'(2)$. For the hat function, the recursion matrix again has $\lambda = 1, \frac{1}{2}$. From the cubic spline, the eigenvalues are $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$.

When $\phi(1)$ and $\phi(2)$ are known, the dilation equation gives $\phi(x)$ at half-integers, such as

$$\begin{aligned}\phi\left(\frac{1}{2}\right) &= \frac{1}{4}(1 + \sqrt{3})\phi(1) = \frac{1}{4}(2 + \sqrt{3}) \\ \phi\left(\frac{3}{2}\right) &= \frac{1}{4}(3 + \sqrt{3})\phi(2) + \frac{1}{4}(3 - \sqrt{3})\phi(1) = 0\end{aligned}$$

Then the equation gives ϕ at quarter-integers as combinations of ϕ at half-integers and so on.

Construction 4. The fourth construction is based on the Daubechies-Lagarias local pyramidal algorithm(see Daubechies and Lagarias(1991,1992)). The Daubechies-Lagarias algorithm enables us to evaluate ϕ and ψ at a point with preassigned precision. We will illustrate the algorithm with wavelets from the Daubechies family; however, the algorithm works for all finite impulse response quadrature mirror filters.

Let ϕ be the scaling function of the D_N wavelet with support $[0, 2N - 1]$. let $x \in (0, 1)$ and define $dyad(x) = d_1, d_2, \dots, d_n, \dots$ as the set of 0 - 1 digits in the dyadic representation of x . That is $x = \sum_{j=1}^{\infty} d_j 2^{-j}$. By $dyad(x, n)$, we denote the subset of the first n digits from $dyad(x)$, i.e., $dyad(x, n) = d_1, d_2, \dots, d_n$. let $h = (h_0, h_1, \dots, h_{2N-1})$ matrices as:

$$T_0 = (\sqrt{2}h_{2i-j-1})_{1 \leq i, j \leq 2N-1}, T_1 = (\sqrt{2}h_{2i-j})_{1 \leq i, j \leq 2N-1}. \quad (1.10)$$

Then the local pyramidal algorithm can be constructed based on Theorem 1.1.1. See Daubechies and Lagarias(1992) or Pinheiro and Vidakovic(1997).

Theorem 1.1.1

$$\lim_{n \rightarrow \infty} T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_n} = \begin{bmatrix} \phi(x) & \phi(x) & \dots & \phi(x) \\ \phi(x+1) & \phi(x+1) & \dots & \phi(x+1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi(x+2N-2) & \phi(x+2N-2) & \dots & \phi(x+2N-2) \end{bmatrix}$$

The convergence of $\|T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_n} - T_{d_1} \cdot T_{d_2} \cdot \dots \cdot T_{d_{n+m}}\|$ to zero for fixed m , is exponential and constructive, i.e., effective decreasing bounds on the error can be established.

Example 1.1.1 Consider the Daub2 scaling function. The corresponding filter is $\mathbf{h} = \left(\frac{1+\sqrt{3}}{4}, \frac{3+\sqrt{3}}{4}, \frac{3-\sqrt{3}}{4}, \frac{1-\sqrt{3}}{4}\right)$. According to (1.10), the matrices T_0 and T_1 are given as

$$T_0 = \begin{bmatrix} \frac{1+\sqrt{3}}{4} & 0 & 0 \\ \frac{3-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{4} & \frac{1+\sqrt{3}}{4} \\ 0 & \frac{1-\sqrt{3}}{4} & \frac{3-\sqrt{3}}{4} \end{bmatrix}, \quad T_1 = \begin{bmatrix} \frac{3+\sqrt{3}}{4} & \frac{1+\sqrt{3}}{4} & 0 \\ \frac{1-\sqrt{3}}{4} & \frac{3-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{4} \\ 0 & 0 & \frac{1-\sqrt{3}}{4} \end{bmatrix}$$

Let us evaluate the scaling function at an arbitrary point, for instance, $x = 0.45$. Twenty "decimals" in the dyadic representation of 0.45 obtained through an S-Plus program are $\text{dyad}(0.45, 20) = 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1$. In addition to the value at 0.45, we get the values at 1.45 and 2.45. The values $\phi(0.45)$, $\phi(1.45)$ and $\phi(2.45)$ may be approximated as averages of the first, second, and third row, respectively in the matrix

$$\prod_{i \in \text{dyad}(0.45, 20)} T_i = \begin{bmatrix} 0.86480582 & 0.86480459 & 0.86480336 \\ 0.08641418 & 0.08641568 & 0.08641719 \\ 0.04878000 & 0.04877973 & 0.04877945 \end{bmatrix}$$

The Daubechies-Lagarias algorithm gives only the values of the scaling function. the following theorem gives the values of the wavelet function.

Theorem 1.1.2 Let x be an arbitrary real number, let the wavelet be given by its filter coefficients $\{h_0, h_1, \dots, h_{2N-1}\}$. Define vector \mathbf{u} with $2N - 1$ components as

$$\mathbf{u}(x) = \{(-1)^{1-[2x]} h_{i+1-[2x]}, i=0, \dots, 2N-2\}.$$

If for some i , the index $i+1-[2x]$ is negative or larger than $2N-1$, then the corresponding components of \mathbf{u} is equal to 0.

Let the vector \mathbf{v} be

$$\mathbf{v}(x, n) = \frac{1}{2N-1} \mathbf{1}' \prod_{i \in \text{dyad}(\{2x\}, n)} T_i,$$

where $\mathbf{1}' = (1, 1, \dots, 1)$ is the row-vector of ones. Then

$$\psi(x) = \lim_{n \rightarrow \infty} \mathbf{v}(x, n),$$

and the limit is constructive.

Computationally, Construction 4 is the easiest to implement. Thus, this construction has been used in this thesis to construct the Daubechies wavelet systems.

1.1.2 Some important wavelet bases

In this section we describe some commonly used families of wavelets: Haar's and Daubechies wavelet systems.

Haar System: The Haar wavelet basis is the simplest example of a wavelet system on $\mathbb{L}^2(\mathbb{R})$. The scaling function is:

$$\phi(x) = I_{[0,1]}(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

The refining relations for the Haar wavelet basis are

$$\phi(x) = \phi(2x - 1) + \phi(2x)$$

and

$$\psi(x) = \psi(2x) - \psi(2x - 1)$$

Daubechies System: Daubechies was the first to construct compactly supported orthogonal wavelets with a preassigned degree of smoothness. The scaling functions and primary wavelets of the Daubechies (1992) wavelet systems, commonly represented as ${}_N\phi(x)$ and ${}_N\psi(x)$ respectively, have no closed forms. They are constructed numerically for different values of the wavelet number N , which identify the number of nonvanishing coefficients in

the “dilation equation” ${}_N\phi(x) = \sum h_k \cdot {}_N\phi(2x - k)$ used in the construction. The choice $N = 1$ yields the Haar wavelets. Once ${}_N\phi(x)$ has been constructed, the corresponding primary wavelet is obtained from ${}_N\psi(x) = \sum_k (-1)^k h_{1-k} \cdot {}_N\phi(2x - k)$. The functions ${}_N\phi(x)$ and ${}_N\psi(x)$ have compact support with vanishing moments of order 1 to N . This property, commonly referred to as a moment condition, guarantees good approximation properties of the corresponding wavelet expansion of a response function $f(x)$ in ${}_N\phi(x)$ and ${}_N\psi(x)$. That is, it determines how quickly the wavelet expansion will converge to the true response $f(x)$. See Härdle et al. (1998) for details.

The algorithm we have used in this thesis is the construction 4 introduced in Section 1.1.1. Table 1.1 list the filter coefficients ${}_N h_k$ for $N = 2$ through 10. In some literatures, the filter coefficients ${}_N h$ is also denoted as $\text{Daub}N$. This notation is also used in this article. Figure 1.1 shows the plots of the corresponding ${}_N\phi$ and ${}_N\psi$, which are the scaling function and the primary function for $\text{Daub}N$, for $N = 2, 3, 4, 5$ and 8. Both ${}_N\phi$ and ${}_N\psi$ have support width $2N - 1$.

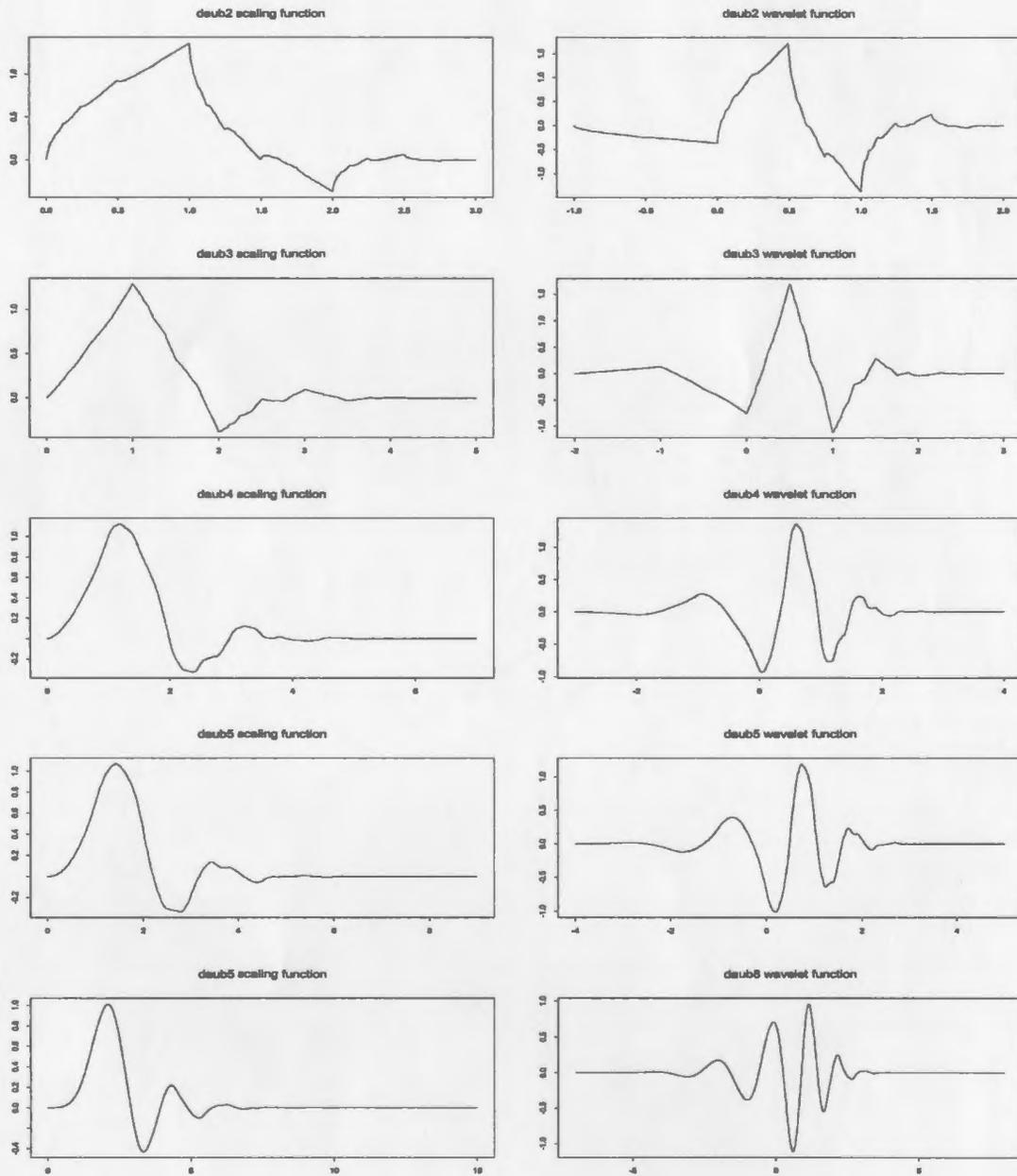


Figure 1.1: Graphs of scaling functions and primary wavelet from Daubechies' family, $N = 2, 3, 4, 5$ and 8 .

Table 1.1: The filter coefficients

k	Duab2	Daub3	Daub4
0	0.4829629131445341	0.3326705529500826	0.2303778133088965
1	0.8365163037378079	0.8068915093110925	0.7148465705529156
2	0.2241438680420133	0.4598775021184915	0.6308807679298589
3	-0.1294095225512603	-0.1350110200102545	-0.0279837694168598
4		-0.0854412738820266	-0.1870348117190930
5		0.0352262918857095	0.0308413818355607
6			0.0328830116668851
7			-0.0105974017850690
k	Duab5	Daub6	Daub7
0	0.1601023979741929	0.1115407433501094	0.0778520540850092
1	0.6038292697971896	0.4946238903984530	0.3965393194819173
2	0.7243085284377729	0.7511339080210953	0.7291320908462351
3	0.1384281459013207	0.3152503517091976	0.4697822874051931
4	-0.2422948870663820	-0.2262646939654398	-0.1439060039285650
5	-0.0322448695846383	-0.1297668675672619	-0.2240361849938750
6	0.0775714938400457	0.0975016055873230	0.0713092192668302
7	-0.0062414902127982	0.0275228655303057	0.0806126091510830
8	-0.0125807519990819	-0.0315820393174860	-0.0380299369350144
9	0.0033357252854737	0.0005538422011614	-0.0165745416306668
10		0.0047772575109455	0.0125509985560998
11		-0.0010773010853084	0.0004295779729213
12			-0.0018016407040474
13			0.0003537137999745
k	Daub8	Daub9	Daub10
0	0.0544158422431070	0.0380779473638783	0.0266700579005555
1	0.3128715909143165	0.2438346746125903	0.1881768000776914
2	0.6756307362973218	0.6048231236901111	0.5272011889317255
3	0.5853546836542239	0.6572880780513005	0.6884590394536035
4	-0.0158291052563724	0.1331973858250075	0.2811723436605774
5	-0.2840155429615815	-0.2932737832791749	-0.2498464243273153
6	0.0004724845739030	-0.0968407832229764	-0.1959462743773770
7	0.1287474266204823	0.1485407493381063	0.1273693403357932
8	-0.0173693010018109	0.0307256814793333	0.0930573646035723
9	-0.0440882539307979	-0.0676328290613299	-0.0713941471663970
10	0.0139810279173996	0.0002509471148314	-0.0294575368218758
11	0.0087460940474065	0.0223616621236790	0.0332126740593410
12	-0.0048703529934519	-0.0047232047577513	0.0036065535669561
13	-0.0003917403733769	-0.0042815036824634	-0.0107331754833305
14	0.0006754494064506	0.0018476468830562	0.0013953517470529
15	-0.0001174767841248	0.0002303857635231	0.0019924052951850
16		-0.0002519631889427	-0.0006858566949597
17		0.0000393473203162	-0.0001164668551292
18			0.0000935886703200
19			-0.0000132642028945

1.2 Some Wavelet Methods for Estimating Functions

There are two main wavelet approaches for estimating functions. These are by (a) wavelet expansion, (b) applying a discrete wavelet transform.

1.2.1 Wavelet Expansion

Mallat (1989) has shown that any $f \in \mathbb{L}_2(\mathbb{R})$ can be represented as a series (convergent in $\mathbb{L}_2(\mathbb{R})$):

$$f(x) = \sum_{k \in \mathbb{Z}} c_{j_0 k} \phi_{j_0 k}(x) + \sum_{j=j_0}^{\infty} \sum_{k \in \mathbb{Z}} d_{jk} \psi_{jk}(x).$$

The coefficients c_{jk} and d_{jk} are then estimated by applying weighted least square (see Oyet and Sutradhar (2003)) or by a nonparametric regression estimator (see Antoniadis, Grégoire and McKeague (1994)). In this article, the wavelet filter we are going to use are Haar wavelet and Daubechies wavelet.

1.2.2 Discrete Wavelet Transformation (DWT)

There are several steps involved in using the DWT to estimate a function. The steps are discussed below.

STEP 1: Apply a discrete wavelet transform to the observed data \mathbf{Y} .

Discrete wavelet transformations map data from the time domain (the original or input data vector) to the wavelet domain. The result is a vector of the same size. Wavelet transformations are linear and they can be defined by matrices of dimension $n \times n$ if they are applied to inputs of size n .

Example 1.2.1 *Let the data vector be $(1, 2)$ and let $M(1, 2)$ be the point in \mathbb{R}^2 with coordinates given by the data vector. The rotation of the coordinate axes by an angle of $\frac{\pi}{4}$ can*

be interpreted as a DWT in the Haar wavelet basis. The rotation matrix is

$$W = \begin{pmatrix} \cos \frac{\pi}{4} & \sin \frac{\pi}{4} \\ \cos \frac{\pi}{4} & -\sin \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

and the discrete wavelet transformation of $(1, 2)'$ is $W \cdot (1, 2)' = (\frac{3}{\sqrt{2}}, -\frac{1}{\sqrt{2}})'$.

The change of basis can be performed by matrix multiplication. Therefore, the procedure of DWT begins with the construction of an orthogonal transformation matrix W . We have already seen a transformation matrix corresponding to Haar's transformation in Example 1.2.1.

In vector notation, we can write the Y_i 's in (1.1) as

$$\mathbf{Y} = \mathbf{f} + \epsilon$$

By applying W to \mathbf{Y} , we have the wavelet image of \mathbf{f} commonly referred to as the wavelet coefficients given by

$$\mathbf{d} = \theta + \epsilon'$$

where $\mathbf{d} = W\mathbf{Y}$, $\theta = W\mathbf{f}$ and $\epsilon' = W\epsilon$. The components of W are constructed as follows.

Let the length of the data be 2^J , let $\mathbf{h} = \{h_s, s \in \mathbb{Z}\}$ be the wavelet filter, and let N be an appropriately chosen constant. Denote by H_k a matrix of size $(2^{J-k} \times 2^{J-k+1})$, $k = 1, \dots$ with entries

$$h_s, \quad s = [(N-1) + (j-1) - 2(i-1)] \text{ modulo } 2^{J-k+1}$$

at the position (i, j) . By analogy, define a matrix G_k , corresponding to the already defined H_k , by replacing h_i by $(-1)^i h_{N+1-i}$. For filters from the Daubechies family, a standard choice for N is the number of vanishing moments.

The matrix $\begin{bmatrix} H_k \\ G_k \end{bmatrix}$ is a basis-changing matrix in 2^{J-k+1} dimensional space; consequently, it is unitary. Therefore,

$$I_{2^{J-k+1}} = \begin{pmatrix} H'_k & G'_k \end{pmatrix} \cdot \begin{pmatrix} H_k \\ G_k \end{pmatrix} = H'_k \cdot H_k + G'_k \cdot G_k,$$

and

$$I_{2^{J-k+1}} = \begin{pmatrix} H_k \\ G_k \end{pmatrix} \cdot \begin{pmatrix} H'_k & G'_k \end{pmatrix} = \begin{pmatrix} H_k \cdot H'_k & H_k \cdot G'_k \\ G_k \cdot H'_k & G_k \cdot G'_k \end{pmatrix}.$$

This implies that

$$H_k \cdot H'_k = I_{2^{J-k}}, G_k \cdot G'_k = I_{2^{J-k}}, H_k \cdot G'_k = G_k \cdot H'_k = 0, \text{ and } H'_k \cdot H_k + G'_k \cdot G_k = I_{2^{J-k+1}}.$$

Now, for a sequence \mathbf{y} the J -step wavelet transformation is $\mathbf{d} = W_J \cdot \mathbf{y}$, where

$$W_1 = \begin{pmatrix} H_1 \\ G_1 \end{pmatrix}, W_2 = \begin{pmatrix} \begin{pmatrix} H_2 \\ G_2 \end{pmatrix} \cdot H_1 \\ G_1 \end{pmatrix},$$

$$W_3 = \begin{pmatrix} \begin{pmatrix} \begin{pmatrix} H_3 \\ G_3 \end{pmatrix} \cdot H_2 \\ G_2 \end{pmatrix} \cdot H_1 \\ G_1 \end{pmatrix}, \dots$$

Example 1.2.2 Suppose that $\mathbf{y} = \{1, 0, -3, 2, 1, 0, 1, 2\}$ and the filter is $\mathbf{h} = (h_0, h_1, h_2, h_3) = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}}\right)$. Then, $J = 3$ and the matrices H_k and G_k are of dimension

$$2^{3-k} \times 2^{3-k+1}.$$

$$H_1 = \begin{pmatrix} h_1 & h_2 & h_3 & 0 & 0 & 0 & 0 & h_0 \\ 0 & h_0 & h_1 & h_2 & h_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 & h_3 & 0 \\ h_3 & 0 & 0 & 0 & 0 & h_0 & h_1 & h_2 \end{pmatrix}$$

$$G_1 = \begin{pmatrix} -h_2 & h_1 & -h_0 & 0 & 0 & 0 & 0 & h_0 \\ 0 & h_3 & -h_2 & h_1 & -h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_3 & -h_2 & h_1 & -h_0 & 0 \\ -h_0 & 0 & 0 & 0 & 0 & h_3 & -h_2 & h_1 \end{pmatrix}$$

Since,

$$H_1 \cdot \mathbf{y} = \{2.19067, -2.19067, 1.67303, 1.15539\}$$

$$G_1 \cdot \mathbf{y} = \{0.96593, 1.86250, -0.96593, 0.96593\}.$$

So the one-step Daub2 DWT of \mathbf{y} is

$$W_1 \mathbf{y} = \{2.19067, -2.19067, 1.67303, 1.15539 | 0.96593, 1.86250, -0.96593, 0.96593\}$$

The two-step Daub2 DWT of \mathbf{y} is

$$W_1 \mathbf{y} = \{1.68301, 0.31699 | -3.28109, -0.18301 | 0.96593, 1.86250, -0.96593, 0.96593\},$$

because

$$H_2 \cdot H_1 \cdot \mathbf{y} = H_2 \cdot \{2.19067, -2.19067, 1.67303, 1.15539\}$$

$$= \{1.68301, 0.31699\}$$

$$G_2 \cdot H_1 \cdot \mathbf{y} = G_2 \cdot \{2.19067, -2.19067, 1.67303, 1.15539\}$$

$$= \{-3.28109, -0.18301\}$$

In this example, due to the lengths of the filter and the data, we can perform the transformation for two steps only, W_1 and W_2 .

STEP 2: The second step recommended by Donoho and Johnstone (1994) and Donoho, Johnstone, Kerkyacharian and Picard (1995) involves a technique called thresholding. First we estimate the error variance σ^2 and use σ to threshold the wavelet coefficients obtained in step 1.

There are several choices for the estimation of σ . Almost all methods use only the wavelet coefficients at the finest scale to estimate the variance of noise. This is based on the fact that the square of signal-to-noise ratio (SNR)² is usually small at high resolutions and, if the signal is not too irregular, the finest scale should contain mainly noise. Moreover, the finest scale contains 50% of all coefficients.

Some choices of the estimator of σ are

$$s = \sqrt{\frac{1}{n/2 - 1} \sum_{i=1}^{n/2} [d_i^{(J-1)} - \bar{d}^{(J-1)}]^2},$$

or a more robust MAD (median absolute deviation from the median) estimator

$$\begin{aligned} \hat{\sigma} &= 1/0.6745 \cdot \text{MAD} [d^{(J-1)}] \\ &= 1.4826 \cdot \text{MEDIAN} [|d^{(J-1)} - \text{MEDIAN}(d^{(J-1)})|] \end{aligned}$$

where $\mathbf{d}^{(J-1)}$ is the vector of finest detail coefficients associated to the multiresolution subspace W_{J-1} .

For all thresholding rules, it is common to set to 0 the coordinates of a vector \mathbf{d} , if they are smaller in absolute value than a fixed non-negative number - the threshold λ .

The most common thresholding policies are *hard* and *soft*. The expression for the hard- and soft- thresholding rules are

$$\delta^h(d, \lambda) = d \cdot \mathbf{1}(|d| > \lambda), \lambda \geq 0, d \in \mathbb{R},$$

and

$$\delta^s(d, \lambda) = (d - \text{sgn}(d) \cdot \lambda) \cdot \mathbf{1}(|d| > \lambda), \lambda \geq 0, d \in \mathbb{R},$$

Donoho and Johnstone (1994) suggest the threshold $\lambda = \sigma\sqrt{2\log n}$, which they call the universal threshold.

Aside from the universal thresholding, an alternative has been suggested by Abramovich and Benjamini (1995).

The process of thresholding of wavelet coefficients can be viewed as a testing problem. For each wavelet coefficient $d_i = \theta_i + \sigma\epsilon_i$, the hypothesis $H_0 : \theta_i = 0$ is tested against the alternative $H_1 : \theta_i \neq 0$. If the hypothesis H_0 is rejected, the coefficient d_i is retained in the model. Otherwise, it is discarded.

The universal threshold can be viewed as a critical value of a similar test in which the level is

$$\begin{aligned} q &= P(|d_i| > \sigma\sqrt{2\log n} | H_0) \\ &= 2\Phi(-\sqrt{2\log n}) \\ &\approx (n\sqrt{\pi\log n})^{-1} \quad [\Phi(-x) \approx \phi(x)/x \text{ when } x \text{ is large}] \end{aligned}$$

It has been shown that the power of the test against the alternative $H_1 : \theta_i = \theta (\neq 0)$ is $O\left(\frac{1}{n\sqrt{\pi\log n}}\right)$ as well.

The universal threshold controls the probability of even one erroneous inclusion of a coefficient. The approximate level of $(n\sqrt{\pi\log n})^{-1}$ tends to zero as n increases. Subsequent severe decreases in power are compensated by accepting almost all H_0 , leading to severe underfitting.

One way to control such dissipation of power is suggested by Abramovich and Benjamini (1995, 1996), which is based on the false discovery rate(FDR) method of Benjamini and Hochberg (1995).

Their procedure has three steps:

(1). For each d_{jk} find its two-sided p -value, p_{jk} in testing $H_0 : \theta_{jk} = 0$,

$$p_{jk} = 2[1 - \Phi(|d_{jk}|/\sigma)].$$

(2). Order the p_{jk} according to their size, $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}$. Find $k = \max\{i | p_{(i)} < (i/n) \cdot q\}$. For this k calculate

$$\lambda_k = \sigma \Phi^{-1}(1 - p(k)/2).$$

(3). Threshold wavelet coefficients at level λ_k .

If ϵ is Gaussian, the above procedure ensures the FDR to be below q . More details about FDR are described in Chapter 2.

After applying the threshold, we obtain thresholded coefficients, $\hat{\mathbf{d}}$, which set to 0 some coordinates of \mathbf{d} , according to the thresholding rules.

STEP 3: Invert the thresholded coefficients, by applying W^{-1} to $\hat{\mathbf{d}}$, to obtain $\hat{\mathbf{f}}$. Let

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ & \cdots & \\ w_{n1} & \cdots & w_{nn} \end{pmatrix}$$

be the matrix performing a discrete wavelet transformation, i.e., $\mathbf{d} = W\mathbf{Y}$. Note that, since W is orthogonal, $W^{-1} = W^T$. In term of W , the wavelet shrinkage estimator of \mathbf{f} is

$$\hat{\mathbf{f}} = W^{-1}(\hat{\sigma}_\epsilon \delta_\lambda(W\mathbf{Y}/\hat{\sigma}_\epsilon)). \quad (1.11)$$

Therefore, the component \hat{f}_i can be written as $\sum_k w_{ki}(\hat{\sigma}_\epsilon \delta_\lambda(W\mathbf{Y}/\hat{\sigma}_\epsilon)_k)$, $i = 1, \dots, n$.

1.3 Large sample properties of \hat{f}

Let $X \sim N(\theta, \sigma^2)$, σ^2 known, and $\delta = \delta(X, \lambda)$ be either a hard- or soft-thresholding rule with threshold λ . Without loss of generality, we may assume that $\sigma^2 = 1$. Bruce, Donoho, Gao, Martin (1994), Gao (1998), and Marron, Adak, Johnstone, Neumann and patil (1998) obtained exact expressions for the expectation and variance of δ under squared error loss for hard- and soft-thresholding rules, which are, respectively:

$$\begin{aligned} M_\lambda^h(\theta) &= \theta + \theta[1 - \Phi(\lambda - \theta) - \Phi(\lambda + \theta)] + \phi(\lambda - \theta) - \phi(\lambda + \theta), \\ V_\lambda^h(\theta) &= (\theta^2 + 1)[1 - \Phi(\lambda - \theta) - \Phi(\lambda + \theta)] \\ &\quad + (\lambda + \theta)\phi(\lambda - \theta) + (\lambda - \theta)\phi(\lambda + \theta) - (M_\lambda^h(\theta))^2, \end{aligned}$$

$$\begin{aligned} M_\lambda^s(\theta) &= M_\lambda^h(\theta) - \lambda[\Phi(\lambda + \theta) - \Phi(\lambda - \theta)], \\ V_\lambda^s(\theta) &= V_\lambda^h(\theta) - \lambda[v(\lambda, \theta) + v(\lambda, -\theta)], \end{aligned}$$

where ϕ and Φ are the standard normal density and cumulative distribution function and

$$v(\lambda, \theta) = [1 + \Phi(\lambda - \theta) - \Phi(\lambda + \theta)] \cdot [(2\theta - \lambda)(1 - \Phi(\lambda - \theta)) + 2\phi(\lambda - \theta)].$$

Under certain conditions, Brillinger (1995) showed that, for each i , \hat{f}_i (see (1.9)) is asymptotically Gaussian with standard errors estimated by

$$s_i = \hat{\sigma} \sqrt{\sum_k w_{ki}^2 V_\lambda \left(\frac{\hat{\theta}_k}{\hat{\sigma}} \right)}.$$

Then we can easily construct approximate confidence intervals for f_i as

$$\left[\hat{f}_i - t_{n-1, q/2} \cdot s_i, \hat{f}_i + t_{n-1, q/2} \cdot s_i \right]$$

where q represents the confidence level.

Chapter 2

FALSE DISCOVERY RATE (FDR) STATISTIC

When pursuing multiple inferences, researchers tend to select the statistically significant ones for emphasis, discussion and support of conclusions. An unguarded use of single-inference procedures results in a greatly increased false positive (significance) rate. To control this multiplicity (selection) effect, classical multiple-comparison procedures (MCPs) aim to control the probability of committing a type I error in families of comparisons under simultaneous consideration. The control of this familywise error rate (FWER) is usually required in a strong sense, i.e. under all configurations of the true and false hypotheses tested (see for example Hochberg and Tamhane (1987)).

Benjamini and Hochberg (1995) suggest a new point of view on the problem of multiplicity. In many multiplicity problems the number of erroneous rejections should be taken into account and not only the question whether any error was made. Yet, at the same time, the seriousness of the loss incurred by erroneous rejections is inversely related to the number of hypotheses rejected. From this points of view, a desirable error rate to control may

Table 2.1: Number of errors committed when testing m null hypotheses

	Declared non-significant	Declared significant	Total
True null hypotheses	U	V	m_0
Non-ture null hypotheses	T	S	$m - m_0$
Total	$m - \mathbf{R}$	R	m

be the expected proportion of errors among the rejected hypotheses, which they term the false discovery rate (FDR).

Consider the problem of testing simultaneously m (null) hypotheses, of which m_0 are true and \mathbf{R} is the number of hypotheses rejected. Table 2.1 summarizes the situation in a traditional form. The specific m hypotheses are assumed to be known in advance; \mathbf{R} is an observable random variable; \mathbf{U} , \mathbf{V} , \mathbf{S} and \mathbf{T} are unobservable random variables. If each individual null hypothesis is tested separately at level α , then $\mathbf{R} = \mathbf{R}(\alpha)$ is increasing in α . We use the equivalent lower case letters for their realized values.

In terms of these random variables, the FWER is $\mathbf{P}(\mathbf{V} \geq 1)$. Testing individually each hypothesis at level α/m guarantees that $\mathbf{P}(\mathbf{V} \geq 1) \leq \alpha$. The proportion of errors committed by falsely rejecting null hypotheses can be viewed through the random variable $\mathbf{Q} = \mathbf{V}/(\mathbf{V} + \mathbf{S})$ – the proportion of the the rejected null hypotheses which are erroneously rejected. The FDR Q_e is defined as the expectation of \mathbf{Q} ,

$$Q_e = E(\mathbf{Q}) = E\{\mathbf{V}/(\mathbf{V} + \mathbf{S})\} = E(\mathbf{V}/\mathbf{R}).$$

If all null hypotheses are ture, the FDR is then equivalent to the FWER. In this case $s = 0$ and $v = r$. So if $v = 0$ then $\mathbf{Q} = 0$, and if $v > 0$ then $\mathbf{Q} = 1$, leading to $\mathbf{P}(\mathbf{V} \geq 1) = E(\mathbf{Q}) = Q_e$. Therefore, control of the FDR implies control of the FWER in the weak sense.

2.1 Test Statistic Construction

The FDR method we are about to describe can be applied only when $n_1 = n_2 = n$ and $x_i = u_i$ in model (1.1). That is, the technique requires equal sample sizes and common design points in (1.1). Let $D_i = Y_i - Z_i = (f(x_i) - g(x_i)) + (\epsilon_i - \nu_i)$. Under the null hypothesis we have $f = g$. It follows that $D_i = (\epsilon_i - \nu_i)$. That is, D_i 's are white noise. By applying the discrete wavelet transformation(DWT) to D_i 's we obtain the corresponding wavelet coefficients d_i 's. Now, following the FDR procedure of Section 1.1.6, we know that, all of the coefficients d_i 's should be rejected. In addition, if we take D_i 's as the observed response and estimate the mean response curve, we should obtain a line at approximately zero parallel to the x axis. This is therefore an approach for verifying the results of the FDR test. The algorithm for the FDR test is outlined below:

STEP 1. Take the differences on both sides of (1.1) to obtain

$$D_i = h(x_i) + \eta_i \quad (2.1)$$

where $D_i = Y_i - Z_i$, $h(x_i) = f(x_i) - g(x_i)$ and $\eta_i = \epsilon_i - \nu_i$. The variance of η_i becomes, $\sigma_\eta^2 = \sigma_\epsilon^2 + \sigma_\nu^2$.

STEP 2. We transform (2.1) into a model in the wavelet domain by applying DWT to (2.1) to obtain

$$\{\mathbf{WD}\}_i = \{\mathbf{Wh}\}_i + \{\mathbf{W}\eta\}_i,$$

where $\mathbf{D} = (D_1, \dots, D_n)^T$, $\mathbf{h} = (h_1, \dots, h_n)^T$ and $\eta = (\eta_1, \dots, \eta_n)^T$ and rewrite (2.1) in the wavelet domain as

$$d_i = \theta_i + \xi_i.$$

where $d_i = \{\mathbf{WD}\}_i$, $\theta_i = \{\mathbf{Wh}\}_i$, and $\xi_i = \{\mathbf{W}\eta\}_i$. It is usual to refer to the d_i 's as

the wavelet coefficients. Then the hypothesis (1.2) is now

$$H_0 : \theta_i = 0 \quad \text{for all } i = 1, \dots, n.$$

in the wavelet domain. Readers who are familiar with wavelet theory would notice immediately that the transformation into the wavelet domain has simplified the test into thresholding of the wavelet coefficients.

STEP 3. For each d_i find its two-sided p -value, p_i for testing $H_0 : \theta_i = 0$.

$$p_i = 2[1 - \Phi(|d_i|/\sigma_\xi)]$$

where $\sigma_\xi^2 = \sigma_{\xi_i}^2 = \sigma_{\eta_i}^2 = \sigma_\epsilon^2 + \sigma_\nu^2$. The second equality holds because the matrix W is orthogonal.

STEP 4. Order the p_i 's in ascending order, $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}$. If $p_{(i)} \geq \frac{i}{n} \cdot q$ for every $1 \leq i \leq n$, reject the null hypothesis (1.2).

As shown in section 1.1.5, The σ_ξ^2 can be estimated by the finest scale $d^{(J-1)}$. For small samples, we may not obtain much information from the data, so the FDR test may sometimes leads to inaccurate results. Hence, when n is small, we use estimates of the differences to validate the FDR test. The main idea is to estimate the difference $h(x) = f(x) - g(x)$ and construct a confidence interval for $\hat{h}(x)$, as described in Section 1.1.6. Our simulation results show that if $\hat{h}(x)$ is close to zero and the interval cover zero, it is an indication that we cannot reject $H_0 : h = 0$. On the other hand, if $\hat{h}(x)$ is not close to zero, then we should reject H_0 . This also depends on the magnitude of σ_η^2 .

The confidence interval is constructed by noting that under H_0 , we have

$$T_i = \frac{\hat{h}_i - h_i}{s_i} = \frac{\hat{h}_i}{s_i} \sim t_{n-1} \quad i = 1, \dots, n,$$

Considering that we conduct n two-sided tests simultaneously, we can construct a Bonferroni-type simultaneous confidence interval of h_i as

$$\left[\hat{h}_i - t_{n-1, \frac{\alpha}{2n}} \cdot s_i, \hat{h}_i + t_{n-1, \frac{\alpha}{2n}} \cdot s_i \right], \quad i = 1, \dots, n.$$

2.2 Simulation Studies

In this section, we examine the size and power performance of the FDR test. The condition under which the simulation study was performed are:

- $n_1 = n_2 = n$.
- $x_i = u_i$, for $i = 1, 2, \dots, n$.
- $n=8, 16$ and 32 .
- $\epsilon \sim N(0, \sigma_\epsilon^2)$ and $\nu \sim N(0, \sigma_\nu^2)$
- $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5), (0.5, 1), (1, 0.5)$ and $(1, 1)$.
- Wavelet filters used in FDR test are Haar wavelet, Daub2, and Daub4.
- Significant level $q = 0.01, 0.025$ and 0.05 .
- In the simulation for the power of the test, $h(x) = \text{constants } (d_1(x))$, quadratic functions $(d_2(x))$ and HeaviSine function $(d_3(x))$. where

$$d_1(x) = (a).\sqrt{1.5}, (b).2.5;$$

$$d_2(x) = (a).3(x - 0.4)^2, (b).3(x - 0.4)^2 + 1.2, (c).3(x - 0.4)^2 + 2;$$

$$d_3(x) = 4 \sin 4\pi x - \text{sgn}(x - 0.3) - \text{sgn}(0.72 - x).$$

- The distributions of error terms, ϵ_i and ν_j , were chosen as

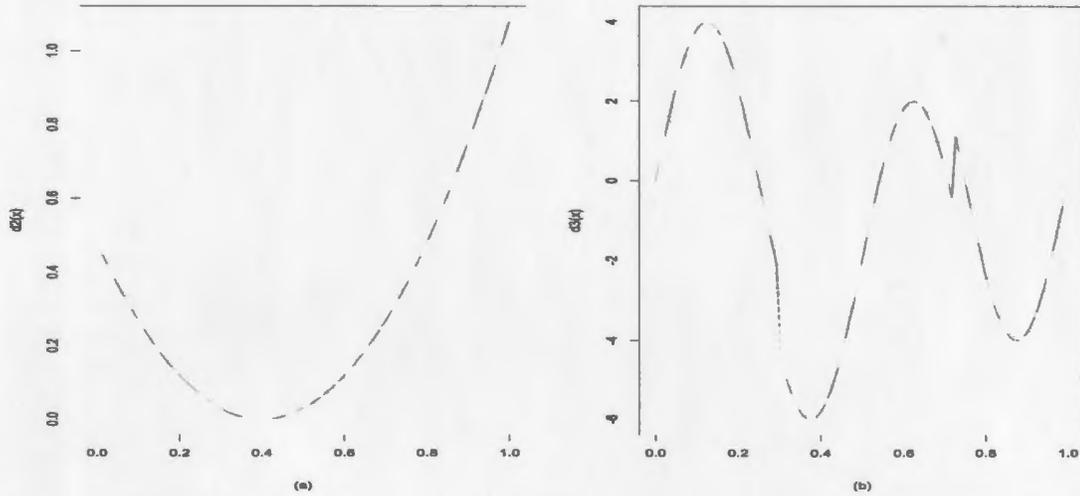


Figure 2.1: A plot of (a). the quadratic function $3(x - 0.4)^2$; and (b). the HeaviSine function, $d_3(x)$.

(a). Normal: $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$;

(b). t : $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$;

(c). Chi-square: $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$;

The choices (a) was used in the size and power simulation studies. Both (b) and (c) were used in the robustness studies in Section 2.4.

To compute the size, we generated observations \mathbf{Y} and \mathbf{Z} of size n following (1.1) and (1.2) with $f(x) = g(x)$. Using these observations we then evaluated $D_i = Y_i - Z_i$ and followed the algorithm for the FDR test to determine whether to reject H_0 or not to reject H_0 . We repeated this process 1000 times and computed the proportion of rejections in 1000 repeated times. This was done for various combinations of σ_ϵ^2 , σ_ν^2 , q and n . The results are

Table 2.2: Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance level $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
8		.071	.097	.137	.091	.105	.130	.083	.111	.130	.085	.094	.105
16		.031	.043	.058	.019	.040	.073	.028	.047	.052	.024	.038	.053
32		.013	.024	.046	.015	.032	.044	.015	.025	.050	.017	.026	.046

shown in Table 2.2. Exactly the same process described above was followed in computing the power in Table 2.3 except that $f(x)$ was different from $g(x)$.

In Table 2.3, the difference between $f(x)$ and $g(x)$, $h(x)$, is taken as the functions $d_1(x)$, $d_2(x)$ and $d_3(x)$ respectively. Compared with the results from Daub2 and Daub4, Haar wavelet gives the best results on sizes and powers. The distributions of errors ϵ_i and ν_j are chosen to be normal. In the robustness simulations, error terms from t and chi-square distributions are used to check the robustness of FDR test.

From Table 2.2 and 2.3, we have following conclusions:

- The performance of the FDR test is affected by sample size and the values of σ_ϵ^2 and σ_ν^2 . The performance of the FDR test improves as the sample size gets larger. For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, $q = 0.01$ and $n = 8$, the size was 0.071. This improved to 0.013 when n increased to 32. We notice the same pattern for $q=0.025$, and 0.05.
- The power of FDR test is mainly determined by the magnitude of the difference between two curves. For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, $q = 0.01$ and $n = 8$, when the difference function was $d_1(a)$, which means $h(x) = \sqrt{1.5}$, the power was 0.636; when $f(x) = d_2(b) = 2.5$, the power increased to 0.992.

Table 2.3: Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance level $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.636	.689	.732	.493	.550	.592	.521	.530	.589	.401	.437	.493
	16		.878	.921	.953	.711	.772	.820	.706	.766	.819	.559	.651	.674
	32		1.00	.997	1.00	.965	.970	.983	.957	.974	.983	.870	.914	.942
$d_1(b)$	8		.992	.998	.999	.948	.973	.984	.948	.962	.984	.885	.930	.938
	16		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.995	.999	.998
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.125	.150	.165	.099	.120	.158	.110	.137	.167	.117	.115	.145
	16		.073	.098	.157	.068	.095	.108	.058	.070	.127	.049	.076	.094
	32		.090	.108	.164	.048	.080	.100	.064	.082	.128	.031	.067	.081
$d_2(b)$	8		.808	.827	.865	.626	.683	.728	.639	.693	.775	.520	.562	.688
	16		.979	.986	.990	.889	.912	.958	.882	.916	.937	.760	.827	.867
	32		1.00	1.00	1.00	.998	1.00	1.00	.997	1.00	1.00	.982	.982	.993
$d_2(c)$	8		.980	.991	.994	.917	.953	.959	.915	.947	.974	.856	.896	.912
	16		1.00	1.00	1.00	.998	1.00	1.00	.998	.999	1.00	.991	.999	.999
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.001	.001
	16		.886	.962	.991	.793	.923	.969	.817	.891	.971	.742	.829	.908
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

- The power of FDR test is also affected by the magnitude of variance of errors. For example, for $q = 0.01$, $n = 8$ and $h(x) = d_1(a)$, when $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, the power was 0.636. This reduced to 0.401 when $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 1)$.
- When the magnitudes of the differences are on the same level, the power of FDR test is also affected by the complexity of the difference function ($h(x)$). For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, $q = 0.01$ and $n = 8$, when $h(x) = d_1(b) = \sqrt{1.5}$, the power was 0.992. This reduced to 0.808 when $h(x) = d_2(b) = 3(x - 0.4)^2 + 1.2$.
- When $n = 8$, the number of coefficients in finest scale, $d^{(J-1)}$, is 4, half of the number of total coefficients. It is therefore difficult to obtain a good estimation of σ_ξ^2 . That explains the poor performance of the FDR method when $n = 8$ and $n = 16$.
- When the difference between f and g is constant, we obtain good power when the constant difference is comparable with the standard deviation $\sigma_\xi = 1$.
- When the difference between $f(x)$ and $g(x)$ is $d_2(a)$, which is not large, considering the magnitude of σ_ϵ and σ_ν , it is natural for the magnitude of the power to be small, especially for small sample. When the difference is taken as $d_2(b)$ and $d_2(c)$, the power of the test improves.
- When the difference is taken as HeaviSine function, although the absolute value of HeaviSine(x) is large, but the complexity of HeaviSine function increases the difficulty of estimating the variances of noise for a small sample. When the sample size increases, the power of FDR and F_1 test improves significantly.
- When the sample size is small, the power is poor. So we need to combine the tests with confidence intervals as shown in Figures 2.2 and 2.3

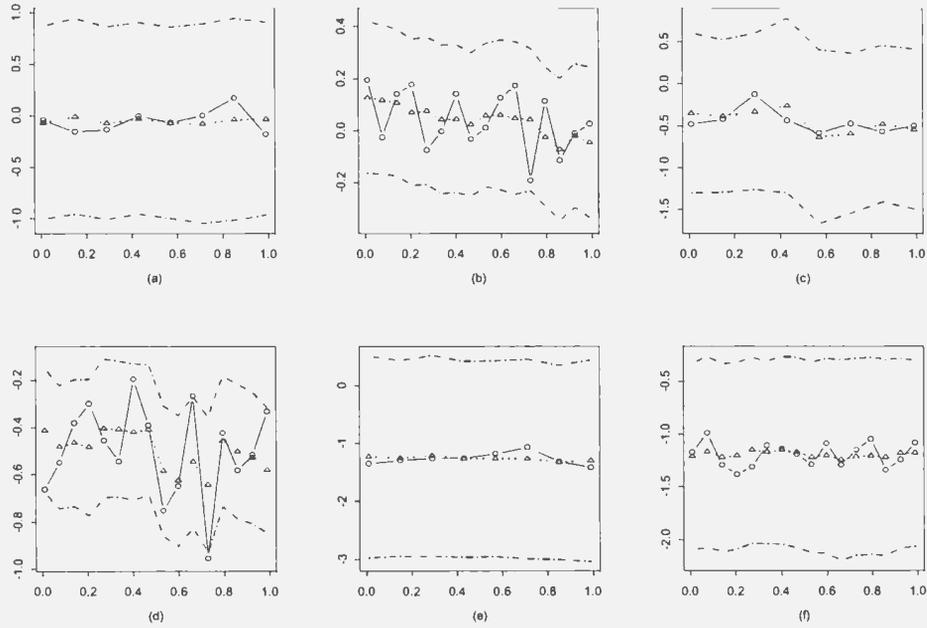


Figure 2.2: A plot of the data $D_i = Y_i - Z_i$ $i = 1, \dots, n$ (solid line), Haar wavelet estimate $\hat{D} = \hat{h}(x) = \hat{f} - \hat{g}(x)$ (dotted line), and approximate confidence bounds (dotted broken line) for difference $h(x) = f(x) - g(x)$ with parameters (a), (b) $n = 8$ and 16 , respectively, with $\alpha = 0.05$, $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$ and $h(x) = 0$; (c), (d) $n=8$ and 16 , respectively, with $\alpha = 0.025$, $(\sigma_\epsilon, \sigma_\nu) = (3.5, 0.8)$, and $h(x) = 0.5$; (e), (f) $n = 8$ and 16 , respectively with $\alpha = 0.01$, $(\sigma_\epsilon, \sigma_\nu) = (2, 2)$, and $h(x) = \sqrt{1.5}$.

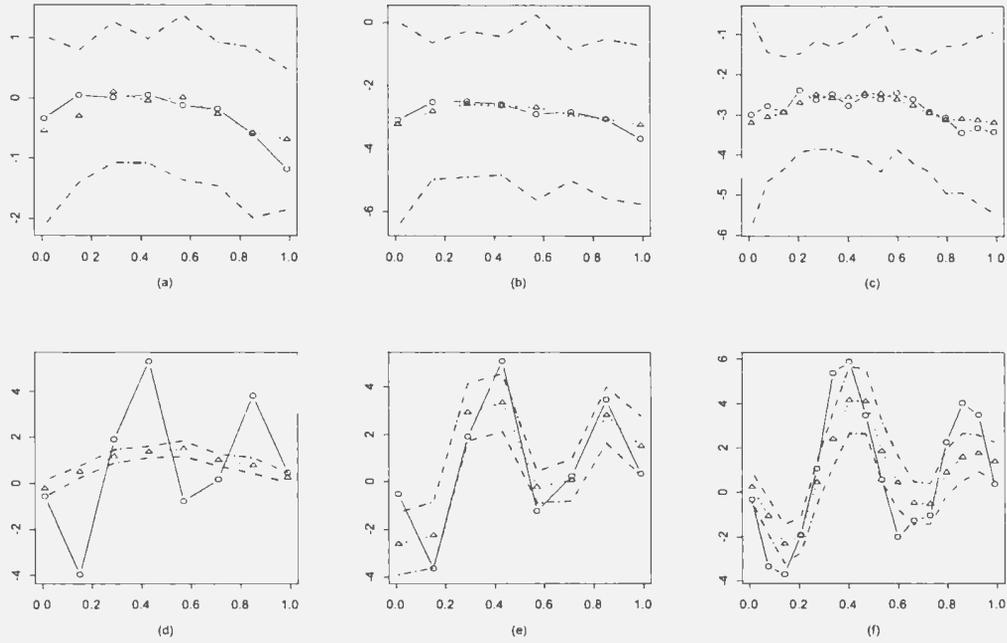


Figure 2.3: A plot of the data $D_i = Y_i - Z_i$ $i = 1, \dots, n$ (solid line), Haar wavelet estimate $\hat{D} = \hat{h}(x) = \hat{f} - \hat{g}(x)$ (dotted line), and approximate confidence bounds (dotted broken line) for difference $h(x) = f(x) - g(x)$ with parameters (a), (b), (c) $n = 8$, $h(x) = 3(x - 0.4)^2$, and $n = 8$, $h(x) = 3(x - 0.4)^2 + 2.5$, and $n = 16$, $h(x) = 3(x - 0.4)^2 + 2.5$, respectively, with $\alpha = 0.05$, $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$ and Filter = ${}_4h_k$; (d), (e), (f) $n = 8$, Filter = ${}_4h_k$ and $n = 8$, Filter = ${}_8h_k$ and $n = 16$, Filter = ${}_8h_k$ respectively, with $h(x) = \text{HeaviSine}$ function, $\alpha = 0.05$, and $(\sigma_\epsilon, \sigma_\nu) = (1.6, 2.5)$.

Table 2.4: Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance level $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$ q	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
8		.074	.102	.110	.076	.112	.114	.067	.093	.119	.081	.099	.140
16		.023	.045	.061	.027	.045	.059	.027	.044	.060	.022	.047	.078
32		.018	.031	.045	.017	.022	.054	.015	.021	.052	.012	.025	.043

2.3 Robustness of FDR test

In the FDR test, it is assumed that both ϵ and ν are normally distributed. In order to check the robustness of FDR test on departures from normal errors, we conduct simulation studies generating errors from t and χ^2 distributions in this section. Tables 2.4 and 2.5 show the size and power of FDR test when the errors is actually from t distribution instead of normal; Tables 2.6 and 2.7 show the size and power of FDR test when errors have χ^2 distribution.

From Table 2.4, 2.5, 2.6 and 2.7, we can see that,

- When the errors are from t distribution, the performance of FDR test on size and power simulation is insensitive to the departure from normal distribution to t distribution.
- When the errors are from a skewed distribution, χ^2 , the performance of FDR test on size is affected slightly. The simulated power remains at a similar level as normal case.

Table 2.5: Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance level $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.619	.699	.730	.475	.543	.593	.479	.520	.594	.367	.484	.484
	16		.873	.922	.939	.716	.770	.834	.712	.784	.813	.535	.605	.682
	32		.998	1.00	.999	.968	.975	.981	.959	.977	.986	.860	.918	.939
$d_1(b)$	8		.991	.997	.996	.954	.965	.978	.939	.973	.979	.886	.934	.953
	16		1.00	1.00	1.00	.999	1.00	1.00	.999	1.00	1.00	.994	.999	.999
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.115	.147	.211	.120	.132	.159	.123	.136	.173	.121	.119	.158
	16		.087	.120	.140	.067	.089	.109	.055	.087	.110	.061	.072	.105
	32		.084	.104	.162	.046	.078	.117	.043	.081	.100	.045	.050	.078
$d_2(b)$	8		.792	.856	.870	.638	.701	.763	.660	.695	.795	.526	.612	.655
	16		.970	.976	.992	.866	.923	.934	.876	.922	.953	.764	.811	.872
	32		1.00	1.00	1.00	.998	.999	.999	.999	.999	1.00	.981	.985	.990
$d_2(c)$	8		.981	.988	.987	.935	.945	.963	.909	.946	.961	.857	.892	.929
	16		1.00	1.00	1.00	.998	.998	1.00	.998	1.00	1.00	.985	.994	1.00
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001
	16		.885	.969	.987	.787	.917	.917	.803	.912	.965	.705	.859	.930
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2.6: Size of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance level $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	q	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
8			.047	.071	.084	.062	.100	.086	.046	.073	.087	.041	.077	.082
16			.015	.033	.059	.017	.031	.070	.026	.025	.051	.023	.046	.054
32			.020	.055	.100	.021	.062	.139	.019	.069	.119	.019	.071	.114

Table 2.7: Power of FDR test based on 1000 replications, the Haar wavelet, $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance level $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.601	.657	.684	.556	.572	.630	.512	.589	.622	.498	.549	.586
	16		.750	.812	.843	.661	.760	.763	.735	.726	.810	.627	.684	.709
	32		.931	.963	.978	.891	.928	.951	.901	.926	.964	.850	.900	.929
$d_1(b)$	8		.943	.949	.963	.898	.924	.933	.906	.940	.946	.874	.919	.926
	16		.989	.998	.998	.986	.992	.995	.992	.997	.998	.977	.989	.993
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.094	.131	.185	.124	.160	.185	.071	.110	.146	.112	.139	.161
	16		.075	.095	.153	.075	.099	.113	.050	.060	.112	.064	.083	.120
	32		.062	.132	.191	.072	.133	.192	.054	.099	.191	.064	.103	.169
$d_2(b)$	8		.708	.766	.808	.661	.719	.737	.672	.711	.783	.618	.680	.717
	16		.867	.916	.930	.819	.848	.872	.839	.876	.898	.777	.814	.869
	32		.986	.994	.997	.955	.977	.990	.984	.981	.996	.956	.973	.977
$d_2(c)$	8		.916	.937	.944	.879	.894	.929	.901	.911	.935	.831	.897	.921
	16		.991	.991	.994	.969	.983	.998	.982	.992	.992	.963	.977	.983
	32		1.00	1.00	1.00	.999	1.00	1.00	1.00	.999	1.00	1.00	1.00	1.00
d_3	8		.000	.001	.000	.000	.003	.002	.001	.001	.000	.000	.002	.002
	16		.824	.910	.945	.793	.885	.936	.763	.897	.921	.764	.866	.915
	32		.997	1.00	1.00	1.00	.999	.999	.993	.997	1.00	.992	1.00	1.00

Chapter 3

WAVELET EXPANSION STATISTIC

3.1 Wavelet Expansion Method for Equal Sample Sizes

In this section, we discuss the simplest case of the wavelet expansion method, which requires the sample sizes to be equal and design points to be the same. In Section 3.1.1, the construction of the F_1 statistic is described in detail. The results of simulation studies on size and power of the F_1 test are reported in Section 3.1.2.

3.1.1 Construction of F_1 statistic

Consider the model (2.1) and expand $h(x)$ using wavelets(see Section 1.1.3) to obtain

$$h(x) = c\phi_0(x) + \sum_{l=0}^m \sum_{k=0}^{2^l-1} d_{lk}\psi_{lk}(x) + r(x)$$

where $r(x)$ are the remainder terms in the wavelet expansion of $h(x)$. That is,

$$r(x) = \sum_{l=m+1}^{\infty} \sum_{k=0}^{2^l-1} d_{lk}\psi_{lk}(x)$$

Assuming the remainders are negligible, let

$$\beta = (c, d_{0,0}, d_{1,0}, d_{1,1}, \dots, d_{m,2^m-1})$$

and

$$q^T(x) = (\phi_0(x), \psi_{0,0}(x), \psi_{1,0}(x), \psi_{1,1}(x), \dots, \psi_{m,2^m-1}(x)).$$

Then, the model (2.1) can be written as

$$D_i = q^T(x_i) \cdot \beta + \eta_i \quad (3.1)$$

In matrix notations, (3.1) becomes

$$\mathbf{D} = \mathbf{Q}\beta + \eta$$

where

$$\mathbf{D} = \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} q^T(x_1) \\ \vdots \\ q^T(x_n) \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix},$$

Oyet and Wiens (2003) note that regression weights are useful in increasing the accuracy of wavelet estimation. Thus, the experimenter may consider weighted least squares estimation of β :

$$\hat{\beta}_{WLS} = (\mathbf{Q}^T \mathbf{W} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{W} \mathbf{D}$$

where $\mathbf{W} = \text{diag}(w_1, \dots, w_n)$ and $w_i = w(x_i) = \int_0^1 \|q_m(s) ds\| / \|q_m(x_i)\|$. Then under the full model (3.1), $\hat{\mathbf{D}}_{F,WLS} = \mathbf{Q} \hat{\beta}_{WLS}$, so that the residual sum of squares (RSS) under the full model is

$$RSS_F = \sum_{i=1}^n w_i e_{F,i}^2,$$

where

$$e_F = \mathbf{D} - \hat{\mathbf{D}}_{F,WLS} = (e_{F,1}, \dots, e_{F,n})^T.$$

Under $H_0 : f = g$ which implies that $h(x) \equiv 0$. Since $h(x) = f(x) - g(x) \approx q^T(x)\beta = 0$ and $q^T(x)$ is not a zero vector, then testing $H_0 : f = g$ is equivalent to testing $H_0 : \beta = 0$.

In this case, the RSS becomes

$$RSS_{H_0} = \sum_{i=1}^n w_i e_{0,i}^2,$$

where

$$e_0 = \mathbf{D} = (D_1, \dots, D_n)^T = (e_{0,1}, \dots, e_{0,n})^T.$$

Then the F-statistic can be constructed as

$$F_1 = \frac{(RSS_{H_0} - RSS_F) / p}{RSS_F / (n - p)}$$

where $p = 2^{m+1}$. Obviously, the null distribution of F_1 will follow a $F_{p, n-p}$ distribution.

3.1.2 Simulation studies

In this section, we examine the size and power performance of the F_1 test.

In Tables 3.1, 3.2 and 3.3, the difference between $f(x)$ and $g(x)$, $h(x)$, is taken as the functions $d_1(x)$, $d_2(x)$ and $d_3(x)$ respectively (see Section 2.2). The filters used in the test are the Haar wavelet, Daub2 and Daub4. The level of decomposition m is taken as 0, 1, 2. The distributions of errors ϵ_i and ν_j are chosen to be normal. In the robustness simulations, error terms from t and chi-square distributions are used to check the robustness of F_1 test.

In the size simulations, different filters and m get similar results. In Table 3.1, we only show the results from Haar wavelet and $m = 0$.

To compute the size, we generated observations Y and Z of size n following (1.1) and (1.2) with $f(x) = g(x)$. Using these observations we then evaluated $D_i = Y_i - Z_i$ and followed the algorithm for the F_1 test to determine whether to reject H_0 or not to reject H_0 . We repeated this process 1000 times and computed the proportion of rejections in 1000 repeated times. This was done for various combinations of σ_ϵ^2 , σ_ν^2 , q and n . The results from Haar wavelet and $m = 0$ are shown in Table 3.1. Exactly the same process described

Table 3.1: Size of F_1 test based on 1000 replications, the Haar wavelet and $m = 0$ with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
8		.009	.028	.052	.010	.018	.048	.011	.024	.042	.011	.030	.044
16		.010	.020	.059	.016	.031	.046	.010	.025	.049	.007	.019	.043
32		.011	.025	.057	.012	.030	.049	.011	.027	.059	.006	.027	.049

above was followed in computing the power in Table 3.2 except that $f(x)$ was different from $g(x)$.

In Table 3.2, we study the power of F_1 test. Under most of the cases, the combination of choices of Haar wavelet and $m = 0$ give the best power. When $h(x)$ is taken as d_3 , (Daub2, $m = 2$) and (Duab4, $m = 2$) also give comparable, or even better results, which are shown in the categories d_3^* and d_3^{**} . This is also the case in the robustness studies. When $m = 2$, the sample size n could only be larger than 8.

From Table 3.1 and 3.2, we draw the following conclusions.

- Similar to the FDR test, the performance of F_1 test is affected by sample size. When the sample size is larger, the performance of the F_1 test is better. However, the F_1 test appear to be able to control the size much better than the FDR test. (see Tables 3.1 and 2.1).
- The power of F_1 test is mainly determined by the magnitude of the difference between two curves. For instance, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, $n = 8$ and $q=0.01$, when the $h(x) = d_1(a) = \sqrt{1.5}$, the power is 0.293; when $h(x) = d_1(b) = 2.5$, the power is 0.943. When the magnitude of difference increased from $\sqrt{1.5}$ to 2.5, the power increases significantly.
- When the magnitudes of the differences are on the same level, the power of F_1 test

Table 3.2: Power of the F_1 test based on 1000 replications and the Haar ,Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.293	.492	.669	.180	.355	.464	.159	.324	.510	.129	.224	.362
	16		.897	.960	.981	.714	.816	.903	.681	.802	.903	.543	.689	.814
	32		1.00	1.00	1.00	.981	.994	1.00	.989	.998	1.00	.948	.981	.991
$d_1(b)$	8		.943	.991	.997	.804	.937	.976	.802	.926	.988	.658	.841	.926
	16		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	1.00	.998
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.013	.041	.098	.014	.038	.076	.016	.035	.085	.013	.026	.068
	16		.044	.100	.165	.025	.057	.136	.034	.063	.118	.027	.060	.114
	32		.144	.202	.305	.059	.138	.214	.066	.169	.234	.046	.110	.180
$d_2(b)$	8		.449	.670	.823	.306	.489	.667	.308	.498	.684	.210	.360	.553
	16		.980	.990	1.00	.888	.956	.987	.865	.941	.984	.738	.867	.940
	32		1.00	1.00	1.00	.999	1.00	1.00	.999	1.00	1.00	.993	.999	.999
$d_2(c)$	8		.881	.957	.991	.706	.881	.954	.720	.884	.950	.566	.779	.882
	16		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.994	.997	1.00
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.010	.002	.005	.025	.003	.005	.017	.003	.015	.040
	16		.828	.973	.994	.703	.909	.972	.719	.898	.967	.613	.820	.924
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3^*	8													
	16		.880	.967	.995	.749	.887	.970	.762	.910	.966	.633	.810	.924
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3^{**}	8													
	16		.886	.976	.994	.771	.916	.958	.727	.920	.975	.641	.807	.933
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

is also affected by the variances of errors. For example, for $h(x) = d_1(a)$, $n = 8$ and $q=0.01$, when $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, the power is 0.293; when $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 1)$, the power reduced to 0.129. Increasing the variances of errors has the same effect as reducing the magnitude of difference. In this case, an increase in the sample size leads to an increase in the power. For example, when n is increased to 16, the powers increases to 0.897 and 0.543 respectively.

- When the magnitudes of the differences are on the same level, the power of F_1 test is also affected by the complexity of the difference function $h(x)$. For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 1)$, $n = 8$ and $q=0.01$, when $h(x) = d_1(b) = 2.5$, the power is 0.658; when $h(x) = d_2(b) = 2(x - 0.4)^2 + 2$, the magnitude of difference is on the comparable level with $d_1(c)$, for this case, the power reduced to 0.566.
- The F_1 test performs well in controlling the size, regardless of the choice of filter and m .
- When the difference function ($h(x)$) is simple (constants and quadratic functions), we tend to obtain higher power with the Haar wavelet filter and $m = 0$. When the difference function is complicated, e.g. $d_3(x)$ (HeaviSine function), we tend to obtain higher power with the choices Daub2 wavelet and $m = 2$ and Daub4 wavelet and $m = 2$ (see Table 3.2 for d_3 , d_3^* and d_3^{**}).
- The choice of filter could be determined by the complexity of the difference function, $h(x)$. When $h(x)$ is more complicated, a longer-length filter is necessary for test.
- The choice of m could be determined by the fluctuation of $h(x)$. The more complicated the fluctuation of $h(x)$ is, the larger m is needed.

- When the difference between f and g is constant, we obtain good power when the constant difference is comparable with the standard deviation σ_ξ .
- When the difference function is taken as $d_2(a)$, which is not large, considering that the magnitude of $h(x)$, it is natural for the power of the test to be small. When the magnitude of $h(x)$ increases ($d_2(b)$ and $d_2(c)$), the power of F_1 test improves.
- When $h(x)$ is taken as the HeaviSine function, the complexity of the HeaviSine function increases the difficulty of estimating the error variances, especially with a small sample. Considering that the degree of freedom of F_1 is $(2^{m+1}, n - 2^{m+1})$, the denominator degree of freedom, $n - 2^{m+1}$, may be very small, when n is small and m is relatively large, which may lead to poor results. From the simulation results in Table 3.2, this happens when $n = 8, m = 1$ and $n = 16, m = 2$.
- Since $n - 2^{m+1}$ must be positive, the value of m must be less than 2 when $n = 8$.

3.1.3 Robustness of F_1 test

In the F_1 test, we assume that ϵ and ν are normal distributed. In this section, we will investigate the robustness of F_1 test on departures from normal errors. The errors we have used are from t and χ^2 distribution instead of normal. Tables 3.3 and 3.4 show the size and power simulation of F_1 test when the errors are from t distribution. Tables 3.5 and 3.6 show the size and power simulation of F_1 test when the errors are from χ^2 distribution. The distributions of errors, for both t and χ^2 , are set to guarantee that the variances of errors remain the same as those of normal case.

From Tables 3.3, 3.4, 3.5 and 3.6, we can see that, the size of F_1 test is slightly but not seriously affected by the non-normal error distribution. The power is a little lower, but

Table 3.3: Size of F_1 test based on 1000 replications, the Haar wavelets and $m = 0$ with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance levels $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$ q	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
8		.005	.021	.043	.008	.020	.039	.012	.018	.045	.009	.024	.042
16		.008	.012	.038	.004	.027	.050	.008	.034	.050	.011	.021	.055
32		.007	.017	.048	.006	.033	.049	.010	.027	.044	.006	.025	.041

still comparable to the results in Table 3.2. This shows that F_1 is robust to departures from normality.

3.2 Generalized Wavelet Expansion Method

Sometimes, it is possible to encounter data for testing equality of curves for which the assumption of equal sample sizes and same design points does not hold. In this section, we modify the wavelet expansion method in Section 3.1, so that it can be used more widely.

3.2.1 Construction of F_2 statistic

Suppose we have repeated observations $\{(Y_{ij}, x_i), i = 1, \dots, n_1, j = 1, 2, \dots, r_i\}$ and $\{(Z_{lk}, u_l), j = 1, \dots, s_l\}$ from

$$\begin{aligned} Y_{ij} &= f(x_i) + \epsilon_{ij}, \\ Z_{lk} &= g(u_l) + \nu_{lk}. \end{aligned} \tag{3.2}$$

If we expand $f(x)$ and $g(u)$ using wavelets, we have

$$\begin{aligned} f(x) &= c\phi_0(x) + \sum_{l=0}^m \sum_{k=0}^{2^l-1} d_{lk}\psi_{lk}(x) + r(x), \\ g(u) &= a\phi_0(u) + \sum_{l=0}^m \sum_{k=0}^{2^l-1} b_{lk}\psi_{lk}(u) + t(u), \end{aligned} \tag{3.3}$$

Table 3.4: Power of the F_1 test based on 1000 replications and the Haar, Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ at significance levels $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.254	.431	.578	.206	.365	.497	.210	.346	.500	.161	.283	.447
	16		.762	.846	.923	.713	.796	.879	.638	.778	.877	.569	.737	.808
	32		.984	.991	.994	.956	.978	.993	.965	.981	.986	.939	.958	.986
$d_1(b)$	8		.846	.928	.968	.785	.894	.939	.769	.904	.950	.714	.846	.919
	16		.994	.998	.999	.995	.998	.998	.994	.995	.996	.992	.992	.996
	32		.999	1.00	1.00	.999	.998	1.00	1.00	.999	1.00	.999	1.00	1.00
$d_2(a)$	8		.021	.030	.081	.010	.042	.065	.010	.042	.085	.015	.042	.082
	16		.040	.068	.149	.028	.061	.114	.026	.072	.145	.025	.061	.120
	32		.097	.142	.260	.074	.130	.235	.076	.143	.210	.059	.110	.177
$d_2(b)$	8		.389	.571	.735	.328	.488	.648	.314	.512	.657	.243	.440	.596
	16		.887	.944	.974	.844	.896	.956	.825	.915	.953	.789	.856	.920
	32		.996	1.00	1.00	.993	.997	.996	.995	.995	.996	.991	.993	.999
$d_2(c)$	8		.761	.904	.953	.706	.840	.924	.701	.862	.925	.619	.795	.875
	16		.992	.995	.999	.983	.990	.999	.990	.994	.995	.971	.982	.995
	32		.999	.999	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00	.999	.999
d_3	8		.000	.005	.014	.002	.008	.018	.000	.005	.029	.001	.010	.028
	16		.744	.906	.965	.680	.858	.947	.666	.869	.958	.617	.834	.916
	32		1.00	.999	1.00	1.00	.999	1.00	1.00	.999	.999	1.00	.999	.999
d_3^*	8													
	16		.774	.914	.967	.709	.859	.942	.721	.877	.944	.653	.795	.910
	32		1.00	1.00	.999	.997	.997	1.00	1.00	1.00	1.00	.998	.999	.998
d_3^{**}	8													
	16		.804	.932	.971	.733	.898	.941	.739	.865	.938	.647	.841	.906
	32		1.00	1.00	.999	1.00	1.00	1.00	.996	1.00	1.00	.998	.999	.999

Table 3.5: Size of F_1 test based on 1000 replications, the Haar wavelets and $m = 0$ with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon(\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu(\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$ q	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
8		.001	.008	.023	.006	.014	.036	.007	.011	.037	.009	.004	.023
16		.007	.014	.030	.004	.021	.040	.004	.013	.042	.006	.018	.045
32		.006	.022	.049	.005	.019	.047	.008	.018	.051	.003	.025	.050

where $r(x)$ and $t(u)$ are the remainder terms in the wavelet expansions of $f(x)$ and $g(u)$. That is,

$$r(x) = \sum_{l=m+1}^{\infty} \sum_{k=0}^{2^l-1} d_{lk} \psi_{lk}(x),$$

$$t(u) = \sum_{l=m+1}^{\infty} \sum_{k=0}^{2^l-1} d_{lk} \psi_{lk}(u).$$

Assuming the remainders are negligible, let

$$\beta = (c, d_{0,0}, d_{1,0}, d_{1,1}, \dots, d_{m,2^m-1}),$$

$$\alpha = (a, b_{0,0}, b_{1,0}, b_{1,1}, \dots, b_{m,2^m-1}),$$

and

$$q^T(x) = (\phi_0(x), \psi_{0,0}(x), \psi_{1,0}(x), \psi_{1,1}(x), \dots, \psi_{m,2^m-1}(x)).$$

Then (3.3) can be written as

$$f(x) = q^T(x) \cdot \beta,$$

$$g(u) = q^T(u) \cdot \alpha,$$

and the model (3.2) can be written as

$$Y_{ij} = q^T(x_i) \cdot \beta + \epsilon_{ij},$$

$$Z_{lk} = q^T(u_l) \cdot \alpha + \nu_{lk}. \tag{3.4}$$

Table 3.6: Power of the F_1 test based on 1000 replications and the Haar, Daub2(*) and Daub4(**) wavelet with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon(\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu(\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025
$d_1(a)$	8		.353	.491	.633	.305	.441	.595	.269	.410	.545	.248	.380	.502
	16		.740	.841	.899	.659	.763	.836	.688	.807	.870	.626	.712	.840
	32		.977	.989	.989	.938	.976	.983	.955	.980	.992	.908	.960	.977
$d_1(b)$	8		.811	.897	.949	.761	.853	.913	.767	.872	.951	.715	.830	.878
	16		.994	1.00	1.00	.977	.992	.997	.993	.996	.999	.982	.996	.996
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.016	.038	.088	.019	.059	.079	.018	.032	.060	.009	.037	.057
	16		.036	.096	.139	.050	.108	.159	.029	.055	.107	.027	.065	.111
	32		.098	.169	.269	.087	.176	.267	.066	.123	.208	.055	.133	.199
$d_2(b)$	8		.469	.625	.765	.377	.559	.683	.384	.560	.690	.343	.517	.610
	16		.876	.919	.964	.808	.859	.925	.817	.924	.949	.750	.843	.916
	32		.997	.996	.998	.992	.995	.993	.995	.999	.999	.981	.985	.999
$d_2(c)$	8		.777	.891	.948	.697	.822	.900	.733	.852	.911	.640	.778	.878
	16		.992	.993	.996	.972	.979	.995	.989	.995	.996	.962	.987	.992
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.006	.013	.002	.003	.026	.000	.005	.013	.001	.007	.022
	16		.755	.896	.965	.717	.872	.943	.724	.863	.926	.638	.846	.915
	32		1.00	1.00	1.00	1.00	1.00	1.00	.998	1.00	1.00	1.00	1.00	1.00
d_3^*	8													
	16		.784	.895	.954	.724	.848	.947	.707	.823	.902	.662	.820	.910
	32		1.00	1.00	1.00	.998	1.00	1.00	1.00	1.00	1.00	.999	1.00	1.00
d_3^{**}	8													
	16		.793	.899	.957	.712	.849	.949	.748	.876	.928	.640	.829	.902
	32		1.00	1.00	1.00	1.00	.999	1.00	.998	1.00	1.00	.999	1.00	1.00

Define

$$\mathbf{V} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1r_1} \\ \vdots \\ Y_{n_11} \\ \vdots \\ Y_{n_1r_{n_1}} \\ Z_{11} \\ \vdots \\ Z_{1s_1} \\ \vdots \\ Z_{n_21} \\ \vdots \\ Z_{n_2s_{n_2}} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} q^T(x_1) & \mathbf{0} \\ \vdots & \vdots \\ q^T(x_1) & \mathbf{0} \\ \vdots & \vdots \\ q^T(x_{n_1}) & \mathbf{0} \\ \vdots & \vdots \\ q^T(x_{n_1}) & \mathbf{0} \\ \mathbf{0} & q^T(u_1) \\ \vdots & \vdots \\ \mathbf{0} & q^T(u_1) \\ \vdots & \vdots \\ \mathbf{0} & q^T(u_{n_2}) \\ \vdots & \vdots \\ \mathbf{0} & q^T(u_{n_2}) \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} \epsilon_{11} \\ \vdots \\ \epsilon_{1r_1} \\ \vdots \\ \epsilon_{n_11} \\ \vdots \\ \epsilon_{n_1r_{n_1}} \\ \nu_{11} \\ \vdots \\ \nu_{1s_1} \\ \vdots \\ \nu_{n_21} \\ \vdots \\ \nu_{n_2s_{n_2}} \end{pmatrix},$$

we have the new model

$$\mathbf{V} = \mathbf{X}\gamma + \mathbf{u} \tag{3.5}$$

where $\gamma = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$, and we have $\mathbf{u} \sim N(0, C)$, where $C = \sigma_\epsilon^2 \Sigma$, and

$$\Sigma = \begin{pmatrix} I_{n_1} & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_x^2}{\sigma_\epsilon^2} I_{n_2} \end{pmatrix}$$

In order to obtain equal variance errors in the model, we multiply both sides of (3.5) by $\Sigma^{-\frac{1}{2}}$

$$\Sigma^{-\frac{1}{2}} \cdot \mathbf{V} = \left(\Sigma^{-\frac{1}{2}} \cdot \mathbf{X} \right) \gamma + \Sigma^{-\frac{1}{2}} \cdot \mathbf{u},$$

which we write as

$$\mathbf{V}^* = \mathbf{X}^* \gamma + \mathbf{u}^*, \quad (3.6)$$

where $\mathbf{V}^* = \Sigma^{-\frac{1}{2}} \cdot \mathbf{V}$, $\mathbf{X}^* = \Sigma^{-\frac{1}{2}} \cdot \mathbf{X}$, and $\mathbf{u}^* = \Sigma^{-\frac{1}{2}} \cdot \mathbf{u}$. $\mathbf{u}^* \sim N(0, \sigma_\epsilon^2 I_{n_1+n_2})$. In (3.6), the weighted least squares estimator (see Oyet and Wiens (2003)) of γ is

$$\hat{\gamma}_{WLS} = (\mathbf{X}^{*T} \mathbf{W} \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{W} \mathbf{V}^*$$

where

$$\begin{aligned} \mathbf{W} &= \text{diag}(\underbrace{w_1, \dots, w_1}_{r_1}, \dots, \underbrace{w_{n_1}, \dots, w_{n_1}}_{r_{n_1}}, \underbrace{w_{n_1+1}, \dots, w_{n_1+1}}_{s_1}, \dots, \underbrace{w_{n_1+n_2}, \dots, w_{n_1+n_2}}_{s_{n_2}}) \\ w_i &= w(x_i) = \int_0^1 \|q_m(s) ds\| / \|q_m(x_i)\| \quad \text{for } i = 1, \dots, n_1, \\ w_i &= w(u_{i-n_1}) = \int_0^1 \|q_m(s) ds\| / \|q_m(u_{i-n_1})\| \quad \text{for } i = n_1 + 1, \dots, n_1 + n_2, \end{aligned}$$

Define $r = \sum_{i=1}^{n_1} r_i + \sum_{l=1}^{n_2} s_l$. Then under the full model (3.5), $\hat{\mathbf{V}}_{F,WLS}^* = \mathbf{X}^* \hat{\gamma}_{WLS}$.

We note that when $H_0 : f = g$ is true, the wavelet expression of f and g will be the same. That is, the null hypothesis becomes $H_0 : \beta = \alpha$ and the expression (3.4) becomes

$$\begin{aligned} Y_{ij} &= q^T(x_i) \cdot \beta + \epsilon_{ij}, \\ Z_{lk} &= q^T(u_l) \cdot \beta + \nu_{lj}. \end{aligned} \quad (3.7)$$

Using matrix notations, we write (3.7) as

$$\mathbf{V} = \mathbf{Q} \beta + \mathbf{u},$$

where $\mathbf{Q} = (q(x_1), \dots, q(x_1), \dots, q(x_{n_1}), \dots, q(x_{n_1}), q(u_1), \dots, q(u_1), \dots, q(u_{n_2}), \dots, q(u_{n_2}))^T$.

Similar to (3.6), we have

$$\mathbf{V}^* = \mathbf{Q}^* \beta + \mathbf{u}^*,$$

with

$$\hat{\beta}_{WLS} = (\mathbf{Q}^{*T} \mathbf{W} \mathbf{Q}^*)^{-1} \mathbf{Q}^{*T} \mathbf{W} \mathbf{V}^*,$$

as an estimate of the regression parameter β and

$$\hat{\mathbf{V}}_{0,WLS}^* = \mathbf{Q}^* \hat{\beta}_{WLS}.$$

as the estimated response under H_0 .

Using the general regression approach, we now construct an F test for $H_0 : \beta = \alpha$ with

$$RSS_{H_0} = \sum_{i=1}^r w_i e_{0,i}^{*2},$$

$$RSS_F = \sum_{i=1}^r w_i e_{F,i}^{*2},$$

where

$$e_0^* = \mathbf{V}^* - \hat{\mathbf{V}}_{0,WLS}^* = (e_{0,1}^*, \dots, e_{0,r}^*)^T,$$

and

$$e_F^* = \mathbf{V}^* - \hat{\mathbf{V}}_{F,WLS}^* = (e_{F,1}^*, \dots, e_{F,r}^*)^T.$$

Then the F-statistic can be constructed as

$$F_2 = \frac{(RSS_{H_0} - RSS_F) / p}{RSS_F / (r - 2p)}$$

where $p = 2^{m+1}$ and the null distribution of F_2 is $F_{p,(r-2p)}$.

3.2.2 Simulation Studies

In this section, we examine the size and power performance of F_2 test. Since we do not take the difference of two group of data, in the size simulation, different signal functions will lead to different results. Hence, in the simulation for the size of the test, we use different signal functions which are indicated below.

In Table 3.7, the conditions under which the size simulation study was performed are:

- We used different signal functions $s_1(x)$, $s_2(x)$ and $s_3(x)$ for $x \in (0, 1)$, where

(a). $s_1(x) = 0$,

(b). $s_2(x) = 3(x - 0.4)^2$, which is $d_2(x)(a)$,

(c). $s_3(x) = 4 \sin 4\pi x - \text{sgn}(x - 0.3) - \text{sgn}(0.72 - x)$, which is $d_3(x)$.

- Sample sizes in two groups are taken as $(n_1 = 12, n_2 = 16)$ and $(n_1 = 16, n_2 = 20)$.
- x_1, x_2, \dots, x_{n_1} are taken as equally spaced points between 0.001 and 0.999. u_1, u_2, \dots, u_{n_2} are taken as equally spaced points between 0.001 and 0.999.
- $\epsilon \sim N(0, \sigma_\epsilon^2)$ and $\nu \sim N(0, \sigma_\nu^2)$.
- The wavelet filter and m used in the F_2 test were (Haar, $m = 0$) and (Daub4, $m = 3$). The combination (Haar, $m = 0$) is used when the difference between f and g is not very complicated, such as constant and quadratic. (Daub4, $m = 3$) is used when the difference is complicated, e.g. HeaviSine.
- The significance levels $q = 0.01, 0.025$, and 0.05 were considered.
- The distributions of error terms, ϵ_i and ν_j , were chosen as
 - (a). Normal: $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$;
 - (b). t : $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$;
 - (c). Chi-square: $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$;

The choice (a) was used in the size and power simulation studies. Both (b) and (c) were used in the robustness studies in Section 3.2.3.

In the size and power simulations, (Haar, $m = 0$) gave the best results for most of the cases and (Daub4, $m = 3$) gave the best results when both $f(x)$ and $g(x)$ are complicated,

Table 3.7: Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

<i>Haar wavelet filter, m=0</i>														
$f = g$	sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	size	q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_1	12, 16		.012	.021	.039	.004	.009	.024	.003	.022	.040	.010	.012	.034
	16, 20		.006	.014	.037	.003	.009	.025	.009	.017	.044	.005	.022	.041
s_2	12, 16		.005	.007	.019	.006	.005	.019	.003	.009	.020	.003	.018	.038
	16, 20		.003	.011	.018	.002	.005	.017	.002	.011	.029	.006	.009	.036
s_3	12, 16		.003	.009	.026	.001	.008	.011	.002	.011	.026	.003	.015	.036
	16, 20		.003	.006	.014	.004	.007	.014	.006	.011	.025	.004	.015	.026
<i>Daub4 wavelet filter, m=3</i>														
$f = g$	sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	size	q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_3	12, 16													
	16, 20		.003	.013	.039	.000	.007	.011	.018	.055	.077	.009	.018	.043

e.g., HeaviSine. So in Table 3.1, we only show the results from (Haar, $m = 0$) and (Daub4, $m = 3$) for $f(x) = g(x) = \text{HeaviSine}$.

To compute the size, we generated observations Y_1, \dots, Y_{n_1} and Z_1, \dots, Z_{n_2} following (3.2) with $f(x) = g(x)$ and observation repeat once at each point. Using these observations we then evaluated $\mathbf{V} = (Y_1, \dots, Y_{n_1}, Z_1, \dots, Z_{n_2})^T$ and followed the algorithm for the F_2 test to determine whether to reject H_0 or not to reject H_0 . We repeated this process 1000 times and computed the proportion of rejections in 1000 repeated times. This was done for various combinations of $\sigma_\epsilon^2, \sigma_\nu^2, q$ and (n_1, n_2) . The results from (Haar, $m = 0$) and (Daub4, $m = 3$) are shown in Table 3.7. Exactly the same process described above was followed in computing the power in Table 3.8 except that $f(x)$ was different from $g(x)$.

We found that the power performance of the F_2 test is affected by the magnitude of

the difference, the complexity of the difference between f and g and the complexity of the signal function. In order to see the effect of these three factors, we take the functions f and g as follows.

(a). $f = s_1$ and $g = s_1 + \sqrt{1.5}$,

(b). $f = s_2$ and $g = s_2 + 2.5$,

(c). $f = s_3$ and $g = s_3 + 2.5$,

(d). $f = s_2$ and $g = 2s_2$,

(e). $f = s_2$ and $g = 2s_2 + 1.2$,

(f). $f = s_2$ and $g = 2s_2 + 2$,

(g). $f = s_3$ and $g = 2s_3$,

The results of power simulation for these different choices of f and g are shown in Table 3.8.

In the simulations, for most of the cases, (Haar, $m = 0$) gave the highest power compared with other combinations, except when the difference between f and g was complicated, such as HeaviSine, we used the combination (Daub4, $m = 3$).

From Tables 3.7 and 3.8, we have the following conclusions:

- The performance of the F_2 test is affected by the complexity of signal function, sample size, values of σ_ϵ^2 and σ_ν^2 . The performance of the F_2 test improves as the signal function is simpler.
- In the simulation for size, we showed the size of F_2 test from combination of (Daub4, $m = 3$) in Table 3.7(g). When using (Daub4, $m = 3$) for $f = g = s_3$, for most of

Table 3.8: Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and $\nu_j \sim N(0, \sigma_\nu^2)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

<i>Haar wavelet filter, m=0</i>														
f, g	sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
(a)	12, 16		.785	.881	.938	.465	.600	.694	.505	.627	.716	.408	.572	.670
	16, 20		.922	.966	.980	.615	.750	.843	.671	.779	.861	.606	.719	.820
(b)	12, 16		.996	1.00	1.00	.948	.974	.984	.968	.982	.990	.973	.987	.996
	16, 20		.998	1.00	1.00	.990	.997	.998	.990	.998	1.00	.996	.997	1.00
(c)	12, 16		.996	.999	.999	.952	.972	.989	.967	.976	.989	.974	.989	.995
	16, 20		.999	1.00	1.00	.987	.992	.999	.989	.998	.998	.998	.998	.999
(d)	12, 16		.013	.032	.065	.008	.024	.045	.016	.034	.057	.012	.014	.053
	16, 20		.018	.058	.089	.013	.028	.055	.015	.048	.071	.017	.041	.052
(e)	12, 16		.880	.934	.964	.612	.745	.827	.579	.727	.805	.543	.690	.788
	16, 20		.971	.988	.994	.790	.880	.939	.770	.850	.900	.778	.879	.929
(f)	12, 16		.994	.999	.997	.923	.973	.980	.916	.954	.970	.937	.976	.992
	16, 20		.999	1.00	1.00	.978	.992	.999	.974	.990	.996	.998	.996	.999
<i>Daub4 wavelet filter, m=3</i>														
(g)*	12, 16													
	16, 20		.585	.795	.918	.263	.454	.681	.594	.827	.942	.309	.549	.716

the cases, the size of F_2 test is conservative. When $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 0.5)$, the sizes are a little larger than the significance levels.

- Similar to the F_1 test, the power of the F_2 test is mainly determined by the magnitude of the difference between two curves. For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, $(n_1, n_2) = (12, 16)$ and $q=0.01$, when the signal function are (a), which means $f = s_1$ and $g = s_1 + \sqrt{1.5}$, the power is 0.785; for (b), which means $f = s_1$ and $g = s_1 + 2.5$, the power is 0.997. When the magnitude of difference increases from $\sqrt{1.5}$ to 2.5, the power increases significantly.
- When the magnitudes of the differences are small and on the same level, the power of F_2 test is also affected by the variances of errors. For example, for case (a), $(n_1, n_2) = (12, 16)$ and $q=0.01$, when $(\sigma_\epsilon^2, \sigma_\nu^2) = (0.5, 0.5)$, the power is 0.785; when $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 1)$, the power reduces to 0.408. Increasing the variances of errors has the same effect as reducing the magnitude of difference. In this case, an increase in the sample size lead to an increase in the power. For example, when (n_1, n_2) is increased to (10,20), the power increases to 0.922 and 0.606 respectively.
- When the magnitudes of the differences are on the same level, the power of F_2 test is also affected by the complexity of the difference between f and g . For example, for $(\sigma_\epsilon^2, \sigma_\nu^2) = (1, 1)$, $(n_1, n_2) = (12, 16)$ and $q=0.01$, for case (b), which means $f = s_1$ and $g = s_1 + 2.5$ and the difference between f and g is constant, the power is 0.973. For case (f), which means $f = s_2$ and $g = 2s_2 + 2$ and the difference is quadratic function, the magnitude of difference is on the comparable level with (b), for this case, the power reduced to 0.937.
- When the difference between f and g is simple (constants and quadratic functions), we obtained higher power with the Haar wavelet filter and $m = 0$, which is shown in

the Table 3.8. When the difference function is complicate (HeaviSine function), we obtain higher power with the choices Daub4 wavelet and $m = 3$.

- The choice of filter could be determined by the complexity of difference between two signal functions, When the difference is complicated, a longer-length filter is necessary for the test.
- The choice of m could be determined by the fluctuation of the difference. The more complicated the fluctuation is, the larger m is needed.
- For case (d), $f = s_2$ and $g = 2s_2$, the difference is s_2 . considering the magnitude of s_2 , ranging form 0.48 to 1.08 on $(0, 1)$, it is natural to get a low power. When the magnitude of difference increases $s_2 + 1.2$ and $s_2 + 2$, in (e) and (f), the power of F_2 increases significantly.
- For case (g), (Daub4, $m = 3$) is more appropriate to be used in F_2 test due to the complexity of HeaviSine function. Although the magnitude of the difference is large, the complexity of signal function and difference, which are both HeaviSine function, reduce the power of F_2 test. Case (d) and (g) are comparable but magnitude of the difference is larger in case (g), which explains why the power in (g) is higher than the power in (d).

3.2.3 Robustness of F_2 test

In the F_2 test, we also assume that ϵ and ν are normally distributed. In this section, we will investigate the robustness of F_2 test on departures from normal errors. The errors we have used are from t and χ^2 distribution instead of normal. Tables 3.9 and 3.10 show the size and power simulation of F_2 test when the errors are from t distribution. Tables 3.11 and 3.12

Table 3.9: Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$. at significance levels $q = 0.01, 0.025$ and 0.05 .

<i>Haar wavelet filter, m=0</i>														
f, g	sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	size	q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_1	12, 16		.004	.010	.027	.003	.010	.015	.004	.014	.024	.005	.014	.031
	16, 20		.009	.015	.021	.006	.035	.030	.006	.014	.034	.003	.016	.035
s_2	12, 16		.001	.006	.020	.002	.008	.016	.002	.011	.019	.002	.007	.026
	16, 20		.001	.009	.011	.000	.015	.011	.004	.009	.022	.003	.009	.016
s_3	12, 16		.000	.003	.011	.001	.009	.008	.002	.012	.017	.001	.012	.016
	16, 20		.002	.006	.012	.000	.013	.010	.002	.005	.022	.006	.007	.019
<i>Daub4 wavelet filter, m=3</i>														
f, g	sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	size	q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_3	12, 16													
	16, 20		.005	.017	.036	.006	.012	.026	.025	.050	.102	.017	.042	.056

Table 3.10: Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$ and $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$. at significance levels $q = 0.01, 0.025$ and 0.05 .

		<i>Haar wavelet filter, m=0</i>												
f, g	sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
(a)	12, 16		.732	.789	.870	.481	.615	.702	.551	.663	.707	.424	.545	.645
	16, 20		.866	.918	.940	.629	.727	.798	.664	.754	.821	.613	.701	.772
(b)	12, 16		.963	.976	.988	.917	.927	.937	.921	.929	.957	.917	.951	.959
	16, 20		.980	.988	.990	.939	.967	.982	.950	.961	.976	.955	.973	.983
(c)	12, 16		.973	.970	.989	.889	.928	.952	.902	.946	.943	.905	.941	.942
	16, 20		.990	.989	.987	.948	.966	.979	.951	.964	.984	.956	.970	.971
(d)	12, 16		.009	.021	.072	.009	.021	.042	.013	.024	.046	.010	.026	.042
	16, 20		.010	.032	.084	.014	.023	.071	.012	.026	.051	.012	.033	.046
(e)	12, 16		.782	.877	.894	.570	.705	.788	.607	.694	.780	.536	.690	.771
	16, 20		.899	.932	.964	.746	.833	.870	.774	.810	.871	.715	.811	.875
(f)	12, 16		.942	.966	.977	.874	.897	.934	.879	.917	.934	.883	.930	.946
	16, 20		.982	.983	.993	.941	.947	.967	.946	.942	.970	.945	.959	.975
		<i>Daub4 wavelet filter, m=3</i>												
(g)*	12, 16													
	16, 20		.603	.813	.917	.367	.532	.722	.637	.837	.934	.396	.588	.762

Table 3.11: Size of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

<i>Haar wavelet filter, m=0</i>														
f, g	sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_1	12, 16		.001	.007	.012	.001	.008	.009	.001	.005	.017	.002	.003	.013
	16, 20		.001	.007	.015	.002	.013	.015	.002	.010	.016	.006	.004	.010
s_2	12, 16		.003	.001	.005	.000	.005	.015	.001	.004	.007	.000	.001	.005
	16, 20		.000	.003	.009	.004	.002	.011	.001	.005	.013	.003	.004	.007
s_3	12, 16		.000	.001	.003	.000	.004	.002	.001	.004	.007	.001	.003	.004
	16, 20		.000	.003	.003	.000	.000	.005	.000	.004	.009	.000	.002	.005
<i>Daub4 wavelet filter, m=3</i>														
$f = g$	sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		q	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
s_3	12, 16													
	16, 20		.041	.049	.078	.027	.041	.069	.087	.118	.162	.067	.073	.133

Table 3.12: Power of F_2 test based on 1000 replications, the combination of (Haar, $m = 0$) and (Daub4, $m = 3$) with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$ and $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ at significance levels $q = 0.01, 0.025$ and 0.05 .

		<i>Haar wavelet filter, m=0</i>												
<i>f, g</i>	sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
			0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05	0.01	0.025	0.05
(a)	12, 16		.640	.710	.776	.487	.547	.613	.503	.566	.679	.389	.501	.615
	16, 20		.756	.816	.872	.606	.644	.713	.643	.728	.787	.529	.647	.707
(b)	12, 16		.599	.726	.758	.483	.560	.627	.395	.536	.645	.365	.474	.565
	16, 20		.762	.793	.862	.558	.656	.696	.534	.699	.751	.524	.604	.697
(c)	12, 16		.582	.703	.743	.414	.521	.606	.381	.481	.604	.323	.438	.555
	16, 20		.726	.793	.846	.562	.637	.688	.539	.655	.748	.470	.595	.646
(d)	12, 16		.913	.953	.956	.818	.846	.861	.877	.901	.919	.824	.875	.908
	16, 20		.968	.976	.979	.876	.899	.936	.936	.950	.952	.917	.933	.958
(e)	12, 16		.912	.930	.951	.801	.828	.873	.873	.881	.909	.833	.859	.892
	16, 20		.967	.974	.977	.889	.892	.902	.932	.941	.968	.912	.930	.947
(f)	12, 16		.897	.922	.931	.821	.866	.870	.845	.874	.930	.817	.852	.885
	16, 20		.962	.968	.985	.862	.892	.923	.914	.945	.965	.901	.935	.937
(g)	12, 16		.006	.005	.030	.013	.023	.037	.003	.006	.014	.007	.014	.021
	16, 20		.009	.015	.047	.008	.033	.064	.001	.003	.017	.003	.009	.022
(h)	12, 16		.701	.784	.829	.563	.640	.703	.544	.659	.740	.476	.604	.699
	16, 20		.847	.889	.906	.687	.731	.770	.686	.787	.832	.657	.729	.775
(i)	12, 16		.889	.926	.937	.796	.811	.863	.827	.875	.891	.782	.815	.873
	16, 20		.942	.963	.973	.856	.878	.909	.900	.929	.948	.874	.916	.932
		<i>Daub4 wavelet filter, m=3</i>												
(j)*	12, 16													
	16, 20		.660	.817	.914	.415	.630	.700	.673	.843	.914	.451	.623	.756

show the size and power simulation of F_2 test when the errors are from χ^2 distribution. The distributions of errors, for both t and χ^2 , are set to guarantee that the variances of errors remain the same as those of normal case.

From Tables 3.9, 3.10, 3.11 and 3.12, we can see that,

- When the errors are from t distribution, the size of F_2 test becomes more conservative. The power is a little lower, but remains at a similar level as normal case.
- When the errors are from a skewed distribution, χ^2 , the size of F_2 test becomes even more conservative. The power is a little lower than those for t distribution, but still comparable.

Chapter 4

APPLICATIONS AND CONCLUDING REMARKS

4.1 Application : Two Case Studies

In this section, we apply the proposed methods to two real data sets. The data set from CRC example satisfies equal sample size and same design points and is used to check the performance of FDR test, F_1 test and the confidence interval method. The data set from mouse lymphoma assay has different sample sizes and design points and the F_2 test is the appropriate choice for this example.

4.1.1 Case Studies: Methacholine CRC in Mesenteric Arteries

This example contains the data from a Methacholine CRC in mesenteric arteries from 4 months male rats and 7 months male rats. For each case(4 months and 7 months), we compare the means response (as shown in Figure 4.1) with L-Name and without L-Name.

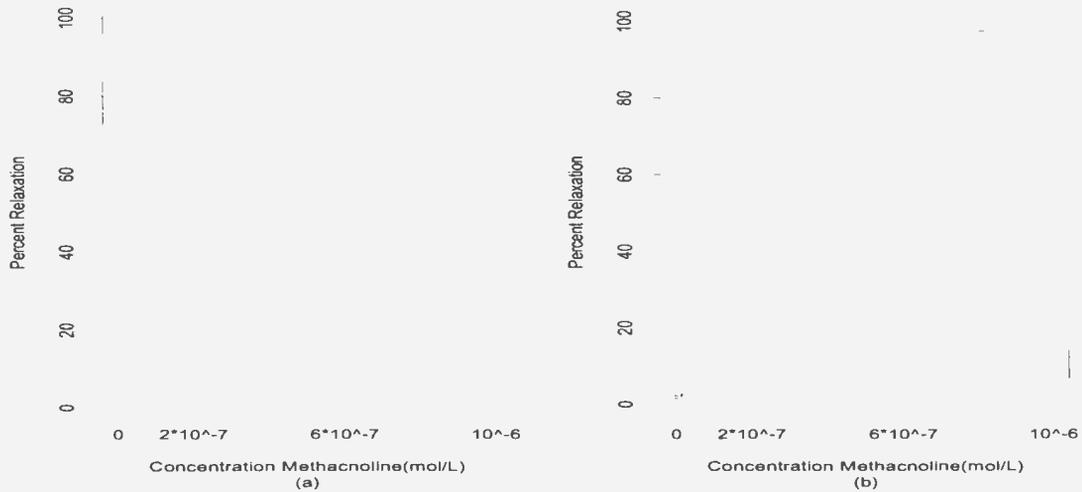


Figure 4.1: Methacoline CRC in mesenteric arteries from 4 months(a) and 7 months(b) male rats:
Solid line: Observations without L-NAME; dotted line: Observations with L-NAME.

Since this data satisfies the conditions of equal sample sizes and same design points, we use the FDR procedure and F_1 test to analyze the data. From our simulation results, we have seen that when the sample size is small, the power of the FDR and F_1 tests are very low especially when the error variances are large. In such cases, confidence interval estimates of the difference was found to be very useful.

Confidence interval method:

From Figure 4.2, we can see that

- For (c), (d) and (e), the s_i goes to zero, so the confidence interval are not visible.
- For four months data, Daub2 gives the best estimation of the difference function.
For seven months data, Daub4 gives the best estimation of difference function. This

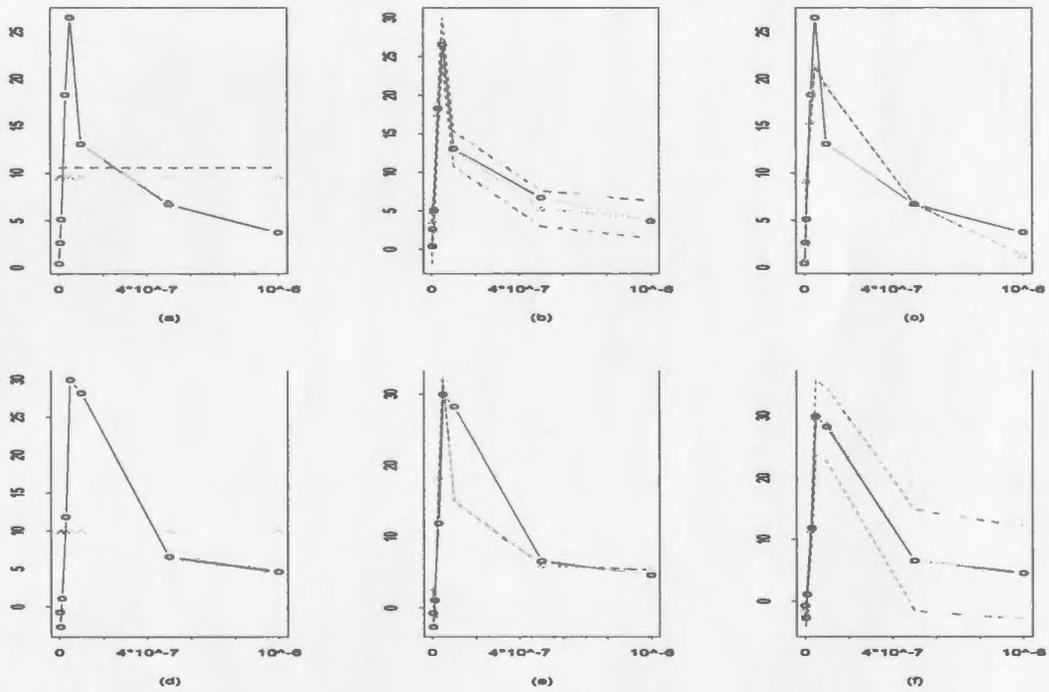


Figure 4.2: Confidence intervals of $\hat{h}(x)$ for CRC data. (a),(b) and (c): for 4 months; (d), (e) and (f): for 7 months. (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used. Solid line: Observed difference; dotted line: Estimated difference; dashed lines: bounds of 95% confidence interval of $\hat{h}(x)$.

Table 4.1: Results of FDR test for CRC data at significance level $q = 0.01, 0.025$ and 0.05 .

Data	filter	Conclusion of FDR test at $q =$		
		0.01	0.025	0.05
4 months	Haar	Reject H_0	Reject H_0	Reject H_0
	Daub2	Reject H_0	Reject H_0	Reject H_0
	Daub4	Reject H_0	Reject H_0	Reject H_0
7 months	Haar	Reject H_0	Reject H_0	Reject H_0
	Daub2	Reject H_0	Reject H_0	Reject H_0
	Daub4	Reject H_0	Reject H_0	Reject H_0

suggested that Daub2 and Daub4 are the appropriate choice for four months data and seven months data.

- For four months data, the confidence interval based on Daub2 is far from including zero. This suggests that $h(x)$ is not zero, which means $f(x)$ and $g(x)$ are not the same. Although Haar and Daub4 gave poor estimation of $\hat{h}(x)$, the confidence intervals based on them are also away from zero.
- For seven months data, the confidence interval based on Daub4 is far from including the zero. This suggests that the $h(x)$ is not zero, which means $f(x)$ and $g(x)$ are not the same. Although Haar and Daub2 gave poor estimation of $\hat{h}(x)$, the confidence intervals based on them are also away from zero.

FDR test:

In Table 4.1, we show the results of the FDR test on CRC data. No matter which wavelet filter is used, Haar, Daub2 or Daub4, FDR test rejects the null hypothesis for both four months and seven months data at significance levels $q = 0.01, 0.025$ and 0.05 .

F_1 test:

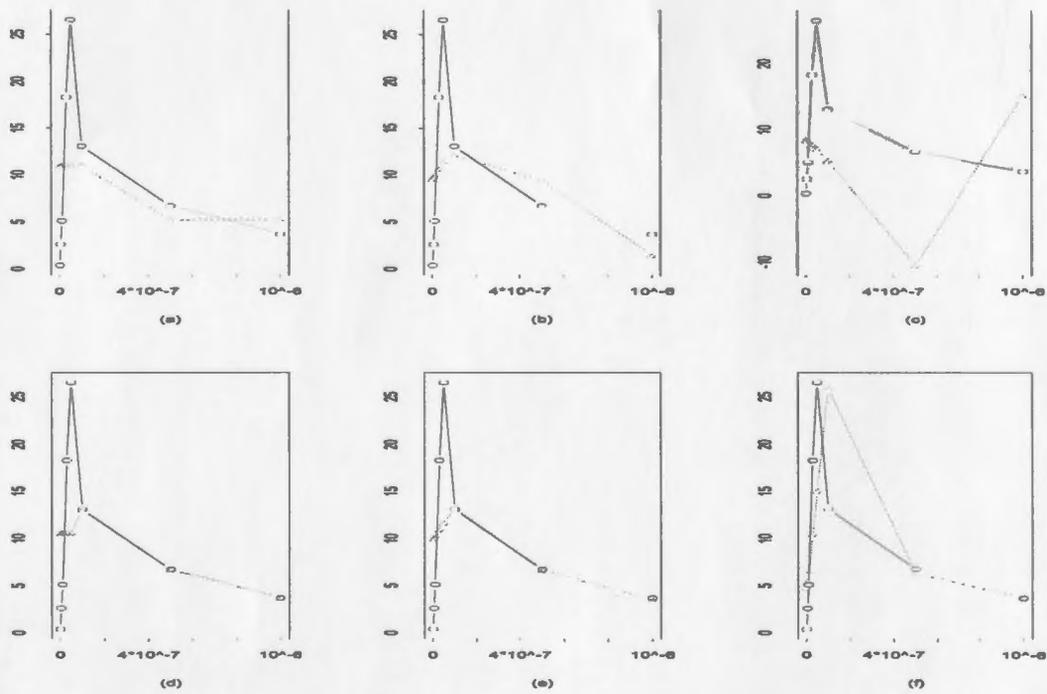


Figure 4.3: CRC Four months data. Solid line: Observed difference $h(x)$; dotted line: Estimated difference $\hat{h}(x)$. (a), (b) and (c): $m = 0$; (d), (e) and (f): $m = 1$; (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used.

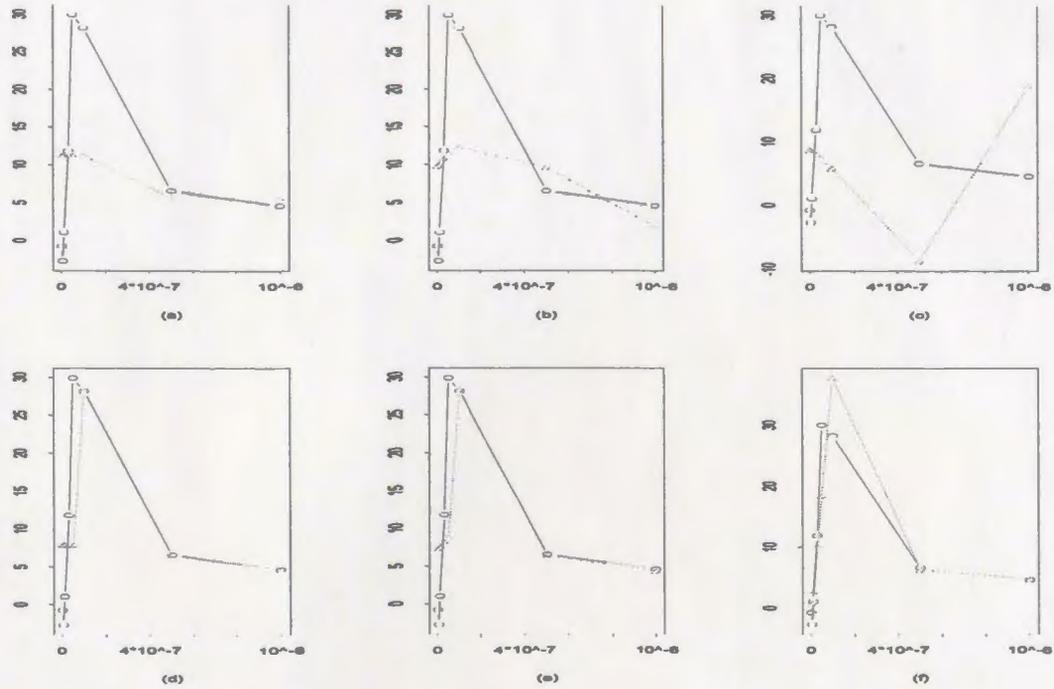


Figure 4.4: CRC Seven months data. Solid line: Observed difference $h(x)$; dotted line: Estimated difference $\hat{h}(x)$. (a), (b) and (c): $m = 0$; (d), (e) and (f): $m = 1$; (a) and (d): Haar wavelet is used. (b) and (e): Daub2 wavelet is used. (c) and (f): Daub4 wavelet is used.

From Figure 4.3 and 4.4, we see that, the estimation of $h(x)$ is better when $m = 1$. Since $h(x)$ is complicated, we need relatively larger value of m to get good estimation. However, the sample size $n = 8$ stop us from using m larger than 1. We recall that the degree of freedom of F_1 is $(2^{m+1}, n - 2^{m+1})$, where the denominator degree of freedom is $n - 2^{m+1} = 4$ when $n = 8$ and $m = 1$. This very small value for the denominator d.f. leads to the poor results in Table 4.2. In order to obtain better results, a larger sample size

Table 4.2: Results of F_1 test for CRC data with Haar, Daub2, Daub4 and $m = 0$ and 1 at significance level $q = 0.05$.

Data	filter	m	F_1	p -value	Conclusion
4 months	Haar	0	4.48	0.064	Do not reject H_0
		1	1.54	0.342	Do not reject H_0
	Daub2	0	5.43	0.045	Reject H_0
		1	1.91	0.273	Do not reject H_0
	Daub4	0	0.59	0.583	Do not reject H_0
		1	2.40	0.209	Do not reject H_0
7 months	Haar	0	2.28	0.184	Do not reject H_0
		1	1.59	0.333	Do not reject H_0
	Daub2	0	2.92	0.130	Do not reject H_0
		1	1.85	0.284	Do not reject H_0
	Daub4	0	0.28	0.762	Do not reject H_0
		1	6.86	0.444	Reject H_0

is needed for F_1 test. From Table 4.2, the F_1 test suggests that H_0 should not be rejected in most cases. This conclusion is not reliable due to the small sample size.

4.1.2 Case Studies: Mouse Lymphoma Assay

The example of mouse lymphoma assay studied the mutant frequency of mouse at different levels of concentrations under substance A and B. The observations are shown in Figure 4.5. In this experiment, repeated measurements are taken at 3 design points for substance A and at 4 design points for substance B.

The estimations of $f(x)$ and $g(x)$ are shown in Figure 4.6. The results of the F_2 test for mouse lymphoma assay data are shown in Table 4.3. We can find the combination of

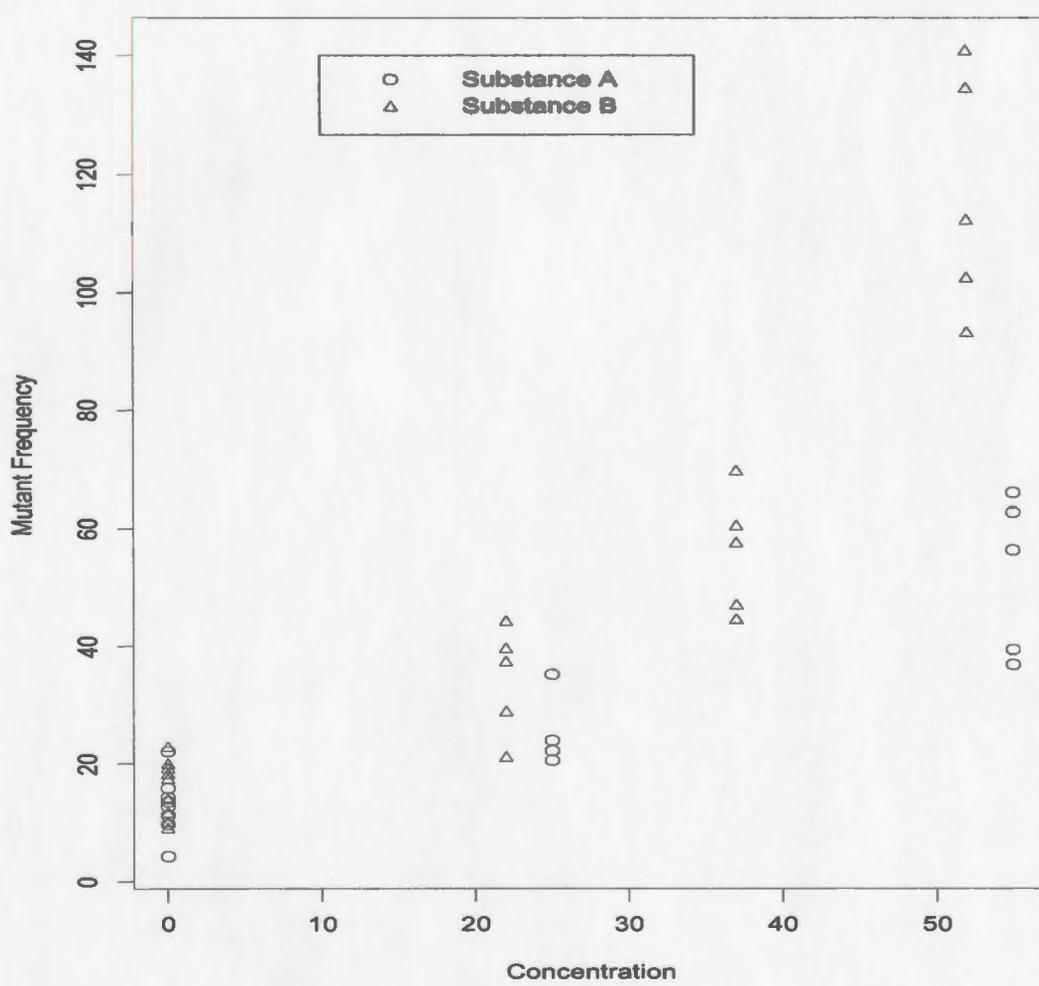


Figure 4.5: Mouse Lymphoma Assay data with substance A and B

Table 4.3: Results of F_2 test for mouse lymphoma assay data with Haar, Daub2, Daub4 and $m = 0, 1$ and 2 at significance level $q = 0.05$.

filter	m	F_2	p-value	Conclusion
Haar	0	5.32	0.009	Reject H_0
	1	25.23	0.000	Reject H_0
	2	9.89	0.000	Reject H_0
Daub2	0	9.76	0.000	Reject H_0
	1	20.23	0.000	Reject H_0
	2	0.07	1.000	Accept H_0
Daub4	0	41.95	0.000	Reject H_0
	1	35.09	0.000	Reject H_0
	2	0.02	1.000	Accept H_0

wavelet filter and m , which give a good estimate for $h(x)$, and then refer to the corresponding results in Table 4.3.

From Figure 4.6, we can find that (d):(Daub2, $m = 0$) and (g):(Daub4, $m = 0$) give relatively better estimations. Then from Table 4.3, we see that for (d), $F_2 = 9.76$ and p -value=0; and for (g), $F_2 = 41.95$ and p -value=0. Both of them suggest rejecting the null hypothesis. The simulation studies shows that we are more likely to commit a type II error unless the difference between the curves is large. The results of the test suggests that the difference is large and therefore H_0 should be rejected.

4.2 Conclusion and Remarks

In this study, we have developed three tests for testing the equality of curves via wavelets. The FDR test involves the discrete wavelet transformation of observations into wavelet do-

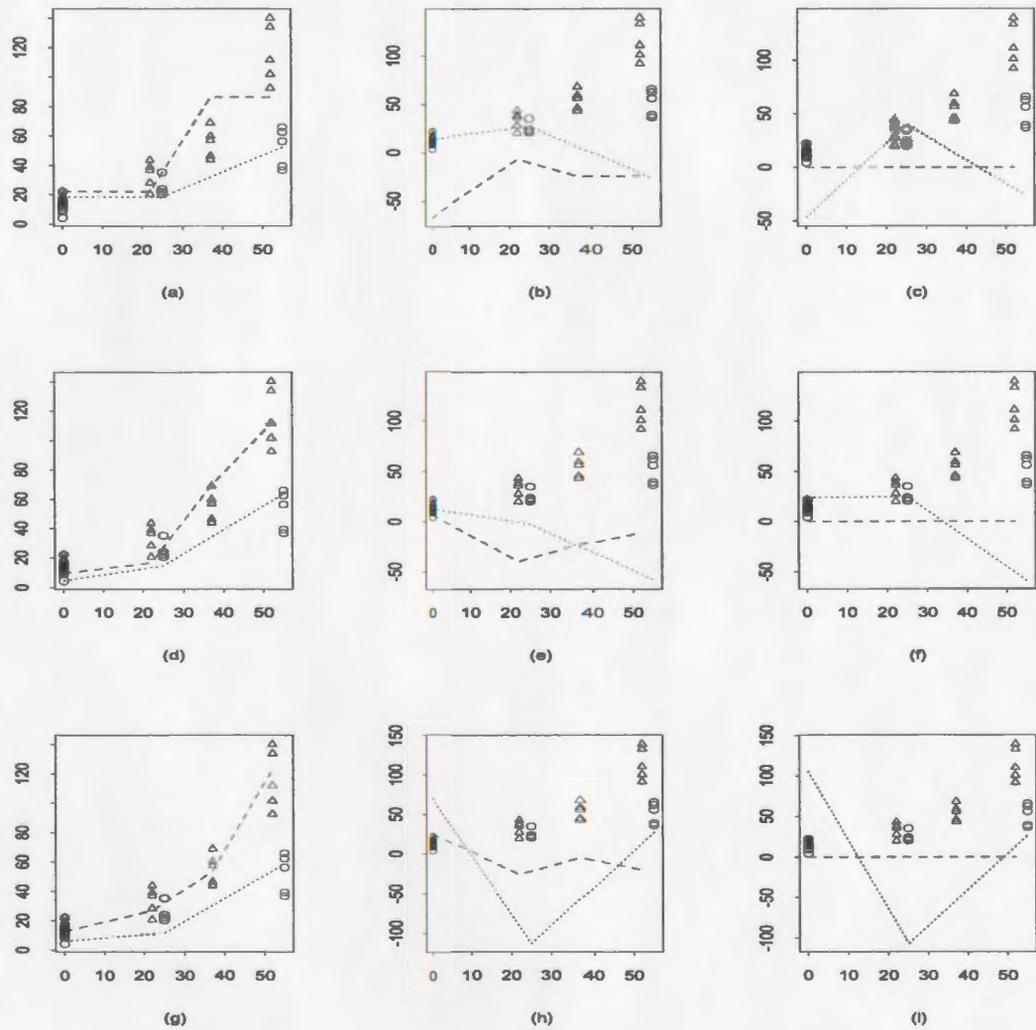


Figure 4.6: Mouse Lymphoma Assay data with substance A and B: (a), (b) and (c): Haar wavelet is used; (d), (e) and (f): Daub2 wavelet is used. (g), (h) and (i): Daub4 wavelet is used. (a), (d) and (g): $m = 0$; (b), (e) and (h): $m = 1$; (c), (f) and (i): $m = 2$.

Table 4.4: Size of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance levels $q = 0.01$ and 0.05 .

		q=0.01											
sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.011	.010	.011	.016	.011	.015	.017	.013	.012	.005	.006	.017
16		.010	.017	.012	.012	.011	.006	.016	.005	.008	.012	.005	.009
32		.011	.009	.014	.009	.009	.007	.014	.009	.005	.014	.006	.013

		q=0.05											
sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.060	.045	.047	.056	.045	.048	.047	.052	.048	.053	.051	.050
16		.053	.051	.041	.055	.031	.051	.049	.043	.063	.038	.048	.046
32		.046	.058	.061	.051	.058	.049	.043	.045	.053	.051	.053	.054

main. The F_1 and F_2 tests exploit the flexibility of wavelet approximations to approximate the unknown response curve by its wavelet expansion.

In Chapter 2, we have described the construction of FDR test in detail. This method is based on the false discovery rate suggested by Benjamini and Hocheberg (1995). We investigate the FDR test by a series of simulation studies on different combinations of sample size, difference function and error variance at different significance levels. The results of simulations show that the performance of FDR test controlling the size improves when sample size increases. We obtain power comparable with the tests proposed in other literature, but is much faster. The robustness of FDR test is also studied in Chapter 2 by a series of simulations. The results of simulations show that FDR test is robust against the departures from normal distribution to t and χ^2 distributions.

In Chapter 3, we proposed two F tests based on the wavelet expansion of functions. First, we constructed the F_1 test, which requires the sample sizes to be equal and the design

Table 4.5: Power of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$.

		$q=0.01$												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.335	.340	.417	.188	.229	.249	.181	.229	.261	.152	.166	.203
	16		.658	.835	.910	.423	.637	.710	.443	.647	.724	.310	.462	.568
	32		.992	.998	.999	.913	.980	.985	.903	.988	.987	.782	.936	.956
$d_1(b)$	8		.914	.931	.966	.766	.822	.857	.741	.803	.854	.604	.676	.740
	16		.999	1.00	1.00	.993	.999	1.00	.991	.999	1.00	.969	.991	.998
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.023	.020	.026	.017	.015	.028	.030	.017	.023	.016	.015	.021
	16		.038	.048	.049	.036	.025	.025	.029	.031	.038	.034	.023	.024
	32		.100	.105	.114	.053	.064	.067	.042	.062	.058	.046	.124	.149
$d_2(b)$	8		.474	.502	.596	.317	.363	.381	.295	.336	.397	.248	.258	.302
	16		.871	.950	.970	.666	.819	.866	.693	.829	.890	.524	.667	.757
	32		1.00	1.00	1.00	.994	.995	.998	.992	.999	.999	.943	.989	.995
$d_2(c)$	8		.849	.881	.928	.696	.757	.772	.670	.733	.790	.536	.600	.651
	16		.998	1.00	1.00	.978	.994	.998	.976	.996	1.00	.924	.980	.990
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	16		.836	.034	.014	.659	.036	.014	.641	.027	.028	.557	.039	.025
	32		1.00	1.00	.985	1.00	.996	.926	1.00	.995	.911	.963	.801	.975

Table 4.6: Power of King's test based on 1000 replications with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\nu_j \sim N(0, \sigma_\nu^2)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$.

		$q=0.05$												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.598	.676	.723	.432	.490	.569	.463	.510	.577	.340	.390	.453
	16		.900	.972	.989	.739	.861	.948	.725	.890	.921	.594	.772	.829
	32		.998	1.00	1.00	.984	.996	1.00	.980	.994	1.00	.943	.966	.995
$d_1(b)$	8		.993	.999	.999	.959	.979	.981	.945	.982	.980	.861	.916	.960
	16		1.00	1.00	1.00	.999	1.00	1.00	1.00	1.00	1.00	.997	1.00	1.00
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.081	.106	.097	.084	.094	.070	.073	.083	.080	.085	.062	.069
	16		.138	.173	.153	.120	.135	.126	.111	.135	.107	.087	.092	.094
	32		.292	.307	.288	.193	.211	.200	.202	.197	.161	.139	.124	.149
$d_2(b)$	8		.776	.840	.865	.613	.648	.734	.597	.662	.760	.491	.541	.647
	16		.975	.998	.997	.891	.972	.985	.885	.971	.976	.802	.912	.939
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	1.00	1.00
$d_2(c)$	8		.972	.990	.999	.920	.957	.961	.899	.940	.974	.847	.883	.920
	16		1.00	1.00	1.00	.999	1.00	1.00	.997	1.00	1.00	.988	.998	1.00
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.000	.000	.000	.001	.000	.000	.002	.002	.001	.002
	16		.998	.381	.256	.984	.321	.222	.986	.285	.237	.953	.282	.214
	32		1.00	1.00	1.00	1.00	1.00	.999	1.00	1.00	.992	1.00	1.00	.975

Table 4.7: Size of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$, $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance levels $q = 0.01$ and 0.05 .

		q=0.01											
sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.010	.007	.006	.007	.009	.011	.013	.009	.007	.010	.008	.014
16		.007	.007	.009	.009	.005	.006	.009	.006	.009	.006	.007	.008
32		.005	.014	.008	.006	.010	.010	.006	.009	.011	.011	.011	.013
		q=0.05											
sample size	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.043	.041	.051	.039	.048	.053	.057	.036	.052	.036	.034	.057
16		.059	.053	.032	.041	.046	.054	.048	.042	.046	.040	.048	.046
32		.048	.052	.045	.057	.044	.042	.052	.050	.042	.040	.043	.045

points to be the same. Then we constructed the F_2 test which relaxes these assumptions and can be used more widely. The simulation studies show that the F_1 test control the size very well. At the same time, F_1 test gave power comparable with FDR test and King's test.

In Tables 4.4-4.12, we show the simulation results of the test proposed in King et. al (1991) for different combination of sample size, difference function, error variance and bandwidth at significance levels $q = 0.01$ and 0.05 . After comparing the results of Table 4.4-4.12 with the corresponding results for FDR test and F_1 test, we find that

- The FDR test can control the size when sample size is large. The F_1 and King's test can control the size very well.
- The FDR test and F_1 test gave comparable power with King's test.
- King's statistic does not have a regular distribution, so a simulation is required to obtain the P -values of the statistic. In King, Hart and Wehrly (1991), they sug-

Table 4.8: Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4, \nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$.

		q=0.01												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.296	.265	.337	.229	.237	.271	.199	.241	.274	.190	.192	.229
	16		.550	.696	.777	.452	.629	.679	.442	.621	.719	.371	.556	.621
	32		.925	.976	.983	.848	.959	.968	.856	.948	.969	.792	.929	.945
$d_1(b)$	8		.775	.855	.881	.746	.780	.831	.737	.777	.831	.691	.740	.776
	16		.986	.992	.999	.975	.986	.992	.975	.985	.991	.951	.975	.979
	32		1.00	1.00	1.00	.999	1.00	1.00	.997	.999	1.00	.999	.999	.998
$d_2(a)$	8		.019	.021	.019	.017	.019	.021	.025	.018	.012	.017	.017	.027
	16		.038	.026	.030	.028	.031	.032	.028	.025	.029	.027	.018	.021
	32		.062	.067	.076	.052	.061	.060	.049	.057	.063	.045	.058	.058
$d_2(b)$	8		.418	.393	.489	.363	.362	.409	.324	.364	.407	.293	.302	.349
	16		.764	.851	.891	.660	.794	.836	.643	.805	.869	.603	.734	.780
	32		.986	.993	.994	.965	.988	.989	.955	.988	.992	.930	.980	.985
$d_2(c)$	8		.705	.794	.817	.685	.705	.768	.680	.699	.777	.619	.663	.711
	16		.978	.984	.989	.948	.981	.986	.952	.975	.982	.919	.956	.964
	32		1.00	1.00	1.00	.998	1.00	.997	.996	.999	1.00	.998	.998	.998
d_3	8		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	16		.705	.026	.022	.679	.033	.024	.622	.028	.029	.545	.036	.028
	32		1.00	.984	.928	1.00	.977	.863	1.00	.980	.868	.999	.961	.820

Table 4.9: Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot t_4$, $\nu_j \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot t_4$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$.

		q=0.05												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.522	.607	.632	.442	.524	.563	.466	.522	.567	.412	.489	.519
	16		.773	.907	.924	.696	.849	.898	.709	.844	.888	.592	.798	.862
	32		.983	.996	.998	.937	.981	.992	.943	.984	.996	.923	.983	.990
$d_1(b)$	8		.951	.960	.974	.910	.940	.955	.907	.947	.960	.893	.925	.946
	16		.994	1.00	1.00	.986	.996	.998	.990	.996	.998	.992	.992	.996
	32		.999	1.00	1.00	1.00	1.00	.999	1.00	1.00	1.00	.999	1.00	1.00
$d_2(a)$	8		.086	.065	.107	.097	.089	.069	.095	.082	.093	.072	.091	.095
	16		.134	.133	.134	.111	.123	.113	.110	.113	.110	.111	.099	.122
	32		.228	.232	.228	.176	.196	.189	.191	.211	.164	.141	.162	.161
$d_2(b)$	8		.687	.749	.794	.607	.678	.701	.626	.662	.715	.558	.629	.657
	16		.915	.968	.979	.863	.937	.952	.847	.933	.947	.777	.901	.940
	32		.998	1.00	.999	.986	.993	.997	.987	.997	.999	.979	.997	.997
$d_2(c)$	8		.927	.943	.965	.876	.907	.930	.873	.922	.943	.847	.895	.918
	16		.991	.998	.998	.982	.995	.997	.984	.995	.997	.983	.986	.996
	32		.999	1.00	1.00	1.00	1.00	.999	1.00	1.00	1.00	.999	1.00	1.00
d_3	8		.001	.000	.000	.001	.001	.001	.000	.001	.000	.000	.000	.002
	16		.982	.354	.235	.959	.340	.224	.969	.316	.221	.948	.293	.233
	32		.999	.996	.988	1.00	.996	.975	1.00	.997	.981	1.00	.991	.963

Table 4.10: Size of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and **0.5** at significance levels $q = 0.01$ and 0.05 .

$q=0.01$													
sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
size	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.006	.004	.006	.007	.004	.003	.002	.006	.001	.005	.002	.003
16		.012	.002	.004	.008	.002	.006	.008	.006	.005	.006	.004	.008
32		.007	.005	.007	.014	.006	.009	.006	.007	.007	.009	.010	.010
$q=0.05$													
sample	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
size	h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
8		.036	.030	.027	.047	.039	.033	.037	.027	.036	.030	.031	.041
16		.039	.025	.034	.037	.043	.038	.037	.036	.040	.031	.032	.051
32		.034	.041	.042	.026	.034	.040	.042	.042	.040	.038	.049	.055

gested 8000 simulations. This greatly reduce the speed of King's test which makes it unattractive to experimenter. Even if only 1000 simulations was used to obtain the p -values for King's statistic, it took the King's test several days to complete computation that can be done in about an hour using the FDR or F_1 test.

Apart from what we have discussed, work is needed on the following directions:

- How to choose the appropriate wavelet filter and level of decomposition m more efficiently.
- How to increase the power when the difference is complicated or sample size is small.

We hope that this work will motivate further research in the direction of construction tests for equality of curves by wavelet methods.

Table 4.11: Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.01$.

		q=0.01												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.327	.384	.421	.301	.344	.381	.252	.287	.345	.227	.275	.348
	16		.546	.674	.767	.499	.644	.685	.481	.638	.690	.452	.589	.633
	32		.897	.968	.970	.844	.918	.945	.860	.955	.972	.803	.904	.925
$d_1(b)$	8		.785	.830	.871	.749	.787	.800	.745	.797	.815	.692	.739	.758
	16		.979	.993	.995	.945	.973	.983	.956	.986	.994	.929	.975	.982
	32		1.00	1.00	1.00	.998	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.020	.018	.020	.027	.027	.033	.011	.007	.007	.020	.017	.011
	16		.043	.030	.039	.048	.040	.036	.025	.019	.022	.032	.024	.033
	32		.077	.095	.075	.087	.077	.093	.056	.069	.065	.045	.057	.068
$d_2(b)$	8		.477	.502	.546	.421	.452	.497	.385	.419	.466	.334	.392	.446
	16		.742	.826	.879	.662	.781	.812	.674	.790	.836	.610	.731	.765
	32		.968	.992	.993	.943	.977	.984	.965	.987	.994	.922	.970	.970
$d_2(c)$	8		.747	.778	.820	.701	.725	.763	.687	.742	.764	.659	.679	.699
	16		.963	.979	.984	.922	.952	.972	.930	.973	.986	.904	.950	.970
	32		.999	1.00	1.00	.998	1.00	1.00	.998	1.00	1.00	.997	1.00	.999
d_3	8		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	16		.714	.035	.019	.636	.030	.018	.677	.027	.021	.583	.036	.025
	32		1.00	.981	.927	.999	.973	.882	.999	.947	.848	1.00	.940	.813

Table 4.12: Power of King's test based on 1000 replications with $\epsilon_i \sim \frac{1}{\sqrt{2}}\sigma_\epsilon \cdot (\chi_1^2 - 1)$, $\nu_i \sim \frac{1}{\sqrt{2}}\sigma_\nu \cdot (\chi_1^2 - 1)$ and bandwidth $h = 0.1, 0.3$ and 0.5 at significance level $q = 0.05$.

		q=0.05												
$h(x)$	n	$(\sigma_\epsilon^2, \sigma_\nu^2)$	(0.5,0.5)			(0.5,1)			(1,0.5)			(1,1)		
		h	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
$d_1(a)$	8		.561	.659	.700	.501	.559	.631	.523	.551	.607	.466	.505	.571
	16		.766	.875	.910	.691	.818	.888	.700	.858	.889	.678	.800	.822
	32		.968	.994	.992	.941	.976	.980	.959	.984	.998	.916	.967	.981
$d_1(b)$	8		.937	.947	.958	.882	.929	.944	.894	.937	.961	.880	.890	.904
	16		.995	1.00	1.00	.980	.994	.999	.991	.998	.999	.980	.995	.994
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_2(a)$	8		.099	.105	.095	.115	.102	.106	.068	.071	.074	.083	.071	.088
	16		.122	.143	.143	.137	.163	.118	.104	.102	.100	.095	.104	.177
	32		.222	.242	.215	.198	.219	.209	.171	.163	.188	.154	.190	.186
$d_2(b)$	8		.693	.752	.795	.625	.666	.723	.671	.683	.732	.594	.626	.692
	16		.881	.944	.959	.820	.907	.943	.846	.946	.946	.808	.893	.910
	32		.992	.998	.999	.977	.998	.995	.994	.999	1.00	.976	.986	.997
$d_2(c)$	8		.915	.931	.937	.851	.896	.923	.866	.912	.940	.844	.856	.877
	16		.991	.997	1.00	.970	.988	.996	.986	.996	.997	.968	.993	.990
	32		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
d_3	8		.000	.000	.002	.000	.000	.001	.000	.000	.001	.001	.001	.000
	16		.967	.355	.274	.952	.313	.218	.944	.378	.228	.927	.328	.255
	32		1.00	.998	.983	1.00	.996	.974	1.00	.991	.960	1.00	.992	.958

REFERENCES

- Abramovich F. & Benjamini, Y. (1995), Thresholding of wavelet coefficients as a multiple hypotheses testing procedure, In Antoniadis A. & Oppenheim, G., *Wavelets and Statistics*, volume 103 of *Lecture Notes in Statistics*, pages 5-14, Springer-Verlag, New York, 1995.
- Abramovich F. & Benjamini, Y. (1996), Adaptive thresholding of wavelet coefficients, *Computational Statistics and Data Analysis*, 22: 351-361.
- Antoniadis, A., Grégoire G., and McKeague, I. W. (1994), Wavelet methods for curve estimation, *Journal of the American Statistical Association*, 89, 1340-1352.
- Benjamini, Y. and Hochberg, Y. (1995), Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society. Series B*, 57, 289-300.
- Brillinger, D. R. (1995), Uses of cumulants in wavelet analysis, *Technical report*, Department of Statistics, University of California at Berkeley.
- Bruce, A. G., Donoho, D. L., Gao, H. Y. & Martin, R. D. (1994), Smoothing and robust wavelet analysis. In *COMPSTAT, Proceedings in Computational Statistics, 11th Symposium*, pages 531-547.
- Bruce, A. G. and Gao, H. Y., Understanding waveletShrink: variance and bias estimation, *Biometrika*, 83, no. 4, 727-745.
- Daubechies, I. (1992), *Ten Lectures on Wavelets*, Number 61 in CBMS-NSF Series in Applied mathematics, Society for Industrial and Applied Mathematics, Philadelphia.

- Daubechies I. and Lagarias, J. C. (1991), Two-scale difference equations. I. Existence and global regularity of solutions, *SIAM Journal on Mathematical Analysis*, 22, no. 5, 1388-1410.
- Daubechies I. and Lagarias, J. C. (1992), Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals, *SIAM Journal on Mathematical Analysis*, 23, no. 4, 1031-1079.
- Delgado, M. A., (1993), Testing the equality of nonparametric regression curves, *Statistics and Probability Letters*, 17, 199-204.
- Donoho, D. L., & Johnstone, I. M. (1994), Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, 81: 425-455.
- Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. & Picard, D. (1995), Wavelet shrinkage: Asymptopia? (with discussion), *Journal of the Royal Statistical Society. Series B*, 57(2):301-369.
- Gao, H-Y (1998), Wavelet shrinkage denoising using the non-negative garrote, *Journal of Computational and Graphical Statistics*, 7(4):469-488.
- Hall, P. & Hart, J. D. (1990), Bootstrap test for difference between means in nonparametric regression, *Journal of the American Statistical Association*, 85, 1039-1049.
- Härdle, W., Kerkyacharian, G., Pickard, D. & Tsybakov, A. (1998), *Wavelets, Approximation, and Statistical Applications*, Lecture Notes in Statistics 129, Springer-Verlag, New York.
- Härdle W. & Marron J. S. (1990), Semiparametric comparisons of regression curves, *The Annals of Statistics*, 18, 63-89.

- Hochberg, Y. and Tamhane, A. C. (1987), *Multiple comparison procedures*, John Wiley and Sons, New York.
- King E. C., Hart J. D. & Wehrly T. E. (1991), Testing the equality of two regression curves using linear smoothers, *Statistics and Probability Letters*, 12, 239-247.
- Koul H. L. and Schick A., Testing for the equality of two nonparametric regression curves, *Journal of Statistical Planning and Inference*, 65, 293-314.
- Kulasekera, K. B. (1995), Comparison of regression curves using quasi-residuals, *Journal of the American Statistical Association*, 90, no. 431, 1085-1093.
- Mallat, S. G. (1989), Multiresolution approximations and wavelet orthonormal bases of $\mathbb{L}^2(\mathbb{R})$, *Transactions of the American Mathematical Society*, 315, 69-87.
- Marron, S. J., Adak, S., Johnstone, I. M. Neumann, M. H. & Patil, P. (1998), Exact risk analysis of wavelet regression. *Journal of Computational and Graphical Statistics*, 7(3):278-309.
- Meyer, Y. (1992), *Wavelets and Operators*, Cambridge University Press, Cambridge.
- Oyet, A. J. (2002), Minimax A- and D-optimal integer-valued wavelet designs for estimation, *The Canadian Journal of Statistics*, 30, no. 2, 301-316.
- Oyet, A. J., and Sutradhar, B. (2003), Testing variances in wavelet regression models. *Statistics and Probability Letters*, 61, 97-109.
- Oyet, A. J., and Wiens, D. P. (2003), On exact minimax wavelet designs obtained by simulated annealing. *Statistics and Probability Letters*, 61, 111-121.

- Pinheiro, A. and Vidakovic, B. (1997), Estimating the square root of a density via compactly supported wavelets, *Computational Statistics and Data Analysis*, 25, no. 4, 399-415.
- Rioul, O., and Vetterli, M. (1991), Wavelets and signal processing, *IEEE Signal Processing Magazine*, 8, 14-38.
- Strang, G. (1989), Wavelets and dilation equations: a brief introduction, *SIAM Review*, 31, no. 4, 614-627.
- Vidakovic, B. (1998), *Statistical Modeling by Wavelets*, Wiley, New York.

