

# Test basado en Wavelet para correlación serial en Panel de Datos

MÓNICA SOFÍA MONTILLA RODRÍGUEZ  
RICARDO TOVAR

UNIVERSIDAD EAFIT  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
MAESTRÍA EN MATEMÁTICAS APLICADAS  
MEDELLÍN  
2015



# Test basado en Wavelet para correlación serial en Panel de Datos

MÓNICA SOFÍA MONTILLA RODRÍGUEZ  
RICARDO TOVAR

Trabajo de investigación presentado como requisito parcial para optar el  
título de Magíster en Matemáticas Aplicadas

Director  
Javier Martínez Plazas  
Magíster en Matemáticas Aplicadas

UNIVERSIDAD EAFIT  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
MAESTRÍA EN MATEMÁTICAS APLICADAS  
MEDELLÍN  
2015



Dedicatorias

*A mi Madre Graciela Q.E.P., a mi esposa Merideni T., a mis hijos Erick y Angie con todo mi amor y cariño.*

Ricardo Tovar

*A mi esposo Fernando, a mi hijo Luis Alejandro con todo mi amor y a mis padres Luis Enrique y Gladys.*

Mónica Sofía Montilla Rodríguez



## Agradecimientos

A las universidades, de la Amazonia y EAFIT, por darnos la oportunidad de concluir con satisfacción la formación académica obtenida durante el desarrollo de nuestra maestría; así mismo, resaltar el apoyo incondicional al MS.c. Javier Martinez Plazas quien con su asesoría y con su conocimiento facilitó el desarrollo de este trabajo de grado; y por último al MS.c. Freddy Hernan Marín y al PhD Jairo Villegas por su orientación en los procesos académicos que apoyaron nuestra formación.





---

## Índice general

---

|   |           |
|---|-----------|
| <b>1. Preliminares</b>  | <b>3</b>  |
| 1.1. Terminología . . . . .                                     | 3         |
| 1.2. Espacio de probabilidad . . . . .                          | 5         |
| 1.2.1. Variable aleatoria . . . . .                             | 6         |
| 1.2.2. Procesos estocásticos . . . . .                          | 8         |
| 1.3. Teoría asintótica . . . . .                                | 9         |
| 1.3.1. Convergencia en probabilidad. . . . .                    | 9         |
| 1.3.2. Convergencia en distribución y distribución límite . . . | 10        |
| 1.3.3. Distribuciones asintóticas . . . . .                     | 10        |
| <b>2. Wavelets</b>  | <b>13</b> |
| 2.1. Introducción a los Wavelets . . . . .                      | 13        |
| 2.2. Transformada de Fourier . . . . .                          | 14        |
| 2.2.1. Serie de Fourier . . . . .                               | 16        |
| 2.3. Transformadas wavelets . . . . .                           | 17        |
| 2.3.1. Transformada wavelet continua . . . . .                  | 17        |
| 2.3.2. Transformada wavelet discreta . . . . .                  | 21        |
| 2.4. Filtros Lineales . . . . .                                 | 22        |
| 2.4.1. Filtros con dominio en el tiempo . . . . .               | 23        |
| 2.4.2. Filtros con Dominio en la Frecuencia . . . . .           | 25        |
| 2.4.3. Filtros Pasa Baja y Pasa Alta . . . . .                  | 26        |
| 2.5. Filtro Haar . . . . .                                      | 27        |

|  |           |
|--|-----------|
| 2.6. Filtro Daubechies . . . . .                                     | 28        |
| <b>3. Panel de Datos</b>   | <b>31</b> |
| 3.1. Panel de datos . . . . .  | 32        |
| 3.2. Modelo general de un panel de datos . . . . .                   | 32        |
| 3.3. Criterios para la selección del modelo . . . . .                | 36        |
| 3.4. Modelo de efectos fijos . . . . .                               | 36        |
| 3.5. Contraste de significatividad de los efectos de grupo . . . . . | 38        |
| 3.6. Los estimadores intra y entre grupos . . . . .                  | 39        |
| 3.7. Paneles no balanceados y efectos fijos . . . . .                | 41        |
| 3.8. Efectos aleatorios . . . . .                                    | 41        |
| 3.9. Heterocedasticidad . . . . .                                    | 43        |
| 3.9.1. Detección de la heterocedasticidad . . . . .                  | 43        |
| 3.10. Autocorrelación . . . . .                                      | 44        |
| 3.10.1. Detección de la autocorrelación . . . . .                    | 45        |
| <b>4. Caso de Aplicación</b>   | <b>47</b> |
| 4.1. Verificación de la Demostración de $\hat{W}_1$ . . . . .        | 47        |
| 4.1.1. Acotamiento del factor de traslación y los rezagos . . . . .  | 50        |
| 4.1.2. Convergencia en Probabilidad . . . . .                        | 51        |
| 4.1.3. Convergencia en Distribución Normal . . . . .                 | 57        |
| 4.1.4. Igualdad de varianzas y acotamiento . . . . .                 | 60        |
| 4.1.5. Monotonía para $a_i$ en $Q(f_i, f_{io})$ . . . . .            | 68        |
| 4.1.6. Convergencia en probabilidad y en distribución . . . . .      | 69        |
| 4.1.7. Convergencia entre densidades espectrales . . . . .           | 73        |
| 4.2. Verificación de las hipótesis . . . . .                         | 77        |
| 4.3. Contraste de Resultados . . . . .                               | 81        |
| 4.4. Conclusiones . . . . .  | 83        |
| <b>5. Notación</b>   | <b>85</b> |
| <b>Bibliografía</b>  | <b>89</b> |

---

## Introducción

---

A partir del método construido por J. B. Fourier en el siglo XIX, para solucionar ecuaciones diferenciales parciales lineales que representaban la transferencia de calor en sólidos, se abrió una gran ventana de posibilidades en el análisis de fenómenos en diferentes campos como la ingeniería, la economía, la física, la medicina, entre otros. A pesar de lo espléndida de la teoría de Fourier, el avance de las tecnologías y la necesidad de análisis cada vez más precisos, han dejado ver cierta debilidad de dicho análisis. La poca adecuación de las funciones trigonométricas para reflejar fenómenos muy localizados, es una de ellas, i.e., una mínima perturbación en la información en un momento determinado genera un impacto global en toda su representación de Fourier. De esta manera cuando hay superposición de varios eventos complejos bien diferenciados, el análisis de Fourier no es efectivo en su totalidad.

Al tratar de mejorar el análisis de Fourier, surge un nuevo método, el análisis *wavelet*. Su objetivo cubre todo un campo de estudio denominado análisis tiempo-frecuencia. En este sentido, la idea base del análisis *wavelet* es construir a partir de una única función, un conjunto o familia de funciones que posibilite el estudio de manera más precisa.

En economía, la información puede presentarse como un conjunto de datos en el mismo periodo de tiempo para diferentes individuos, a los cuales se les conoce como datos de *corte transversal*; también es posible que se presenten en diversos periodos de tiempo para un mismo individuo, en este caso, se denominan *series de tiempo*. Un panel de datos, es la combinación de datos de corte transversal y series de tiempo. En este sentido, los datos varían en

## 2 Introducción

---

frecuencia y tiempo, por lo que los métodos de análisis para cada tipo de datos son deficientes.

En la actualidad el análisis wavelet está siendo utilizado como método de estudio de los panel de datos. En 2006 Hong y Kao [32] propusieron dos test para determinar la correlación serial de información económica registrada en panel de datos.

El test  $\hat{W}_1 = \left( \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \hat{\alpha}_{ijk}^2 - \hat{M} \right) / \hat{V}^{\frac{1}{2}}$  analiza la heterocedasticidad consistente, la cual surge para diferentes varianzas  $\sigma_i^2$  y escalas finas  $J_i$ ; mientras que el test  $\hat{W}_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left( 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \hat{\alpha}_{ijk}^2 - \hat{M} \right) / V_{i0}^{\frac{1}{2}}$  analiza la heterocedasticidad corregida.

De acuerdo a la tesis de maestría de [42] se plantea que los test de [32] no son consistentes y su confiabilidad queda entre dicho. Con el propósito de resolver dicha dualidad se planteó el presente trabajo de investigación, y se estudió la estructura matemática del test  $\hat{W}_1$  y se volvió a programar en Matlab la función *wavetest* agregando parámetros y verificándolo con la base de datos proporcionada por [77].

El documento está organizado en cuatro capítulos. En el primer capítulo se presenta en forma general la terminología básica para el fundamento teórico de los capítulos siguientes, tales como el análisis de Fourier, series de tiempo y teoría de probabilidad. En el segundo capítulo se discuten conceptos sobre *wavelet* para el análisis multirresolución de series de tiempo que permiten construir wavelet con mejores propiedades de aproximación. En el tercer capítulo se plantean los conceptos sobre *panel de datos*, sus ventajas y desventajas, los test que determinan los modelos econométricos a utilizar y algunos aspectos metodológicos de la técnica de panel. Finalmente, en el capítulo cuarto se presentan: *i*) la demostración del test  $\hat{W}_1$  estructurada en siete ítem, tales como la acotación de los factores de traslación, la convergencia en probabilidad y/o en distribución, y la convergencia de las densidades espectrales; *ii*) los valores de los test calculados  $\hat{W}_1$  a partir de los datos reales proporcionados por [77] con los cuales se verificaron las hipótesis. El capítulo termina presentando las conclusiones del trabajo.

Al final del documento se presenta un apartado con la notación utilizada en el interior de este, la cual facilita la lectura de la demostración del test  $\hat{W}_1$ .

# CAPÍTULO 1

---

## Preliminares

---

En este corto capítulo se presentará alguna terminología necesaria para la lectura de esta monografía. Se hace un corto repaso de temas de análisis, teoría de probabilidad, procesos estocásticos y teoría asintótica (p.e., [6], [14], [17], [30], [51], [57], [71]).

### 1.1. Terminología

Las funciones integrables  $L_1(\mathbb{R})$  es el espacio de todas las funciones  $f : \mathbb{R} \rightarrow \mathbb{C}$ , tal que  $\int_{\mathbb{R}} |f(t)| dt = \|f\|_{L_1} < \infty$ . De igual forma se tiene  $L_2(\mathbb{R})$ , el espacio las funciones cuadrado-integrables, cuya norma es

$$\|f\|_{L_2} = \left( \int_{\mathbb{R}} |f(t)|^2 dt \right)^{1/2} < \infty.$$

Este espacio se dota con el producto escalar

$$\langle f, g \rangle_{L_2} = \int_{\mathbb{R}} f(t) \overline{g(t)} dt,$$

donde  $\overline{g(t)}$  denota el conjugado complejo de  $g(t)$ . Con este producto interno el espacio  $L_2(\mathbb{R})$  es de Hilbert. Las funciones  $f, g \in L_2(\mathbb{R})$  son ortogonales si

## 4 Preliminares

---

$\langle f, g \rangle_{L_2} = 0$ . En general,  $L_p(\mathbb{R})$  ( $p \geq 1$ ), es el espacio de todas las funciones (clases de equivalencia)  $f : \mathbb{R} \rightarrow \mathbb{C}$ , tal que  $\int_{\mathbb{R}} |f(t)|^p dt = \|f\|_{L_p}^p < \infty$ , acá

$$\|f\|_{L_p} = \left( \int_{\mathbb{R}} |f(t)|^p dt \right)^{1/p}$$

es la norma de  $f$  en  $L_p(\mathbb{R})$ . Otro espacio que se utilizará es  $\ell_2(\mathbb{Z})$ , el de las sucesiones  $(x_j)$ ,  $j \in \mathbb{Z}$ , tal que  $\sum_j |x_j|^2 < \infty$ .

Sea  $F = \mathbb{C}$  o  $\mathbb{R}$ ,  $X$  y  $Y$  espacios normados (espacios vectoriales equipados con una norma). Un operador lineal es una función  $T : X \rightarrow Y$  tal que  $T(au + bv) = aT(u) + bT(v)$ , para cada  $a, b \in F$  y cada  $u, v \in X$ . El operador  $T$  es continuo en  $u_0$  si para cada  $\epsilon > 0$  existe  $\delta > 0$  tal que si

$$\|u - u_0\|_X < \delta \quad \text{entonces} \quad \|Tu - Tu_0\|_Y < \epsilon. \quad (1.1.1)$$

Si (1.1.1) se cumple para cada  $u_0 \in X$  se dice que  $T$  es continuo en  $X$ . Si  $\delta$  no depende del punto  $u_0$  se dice que  $T$  es uniformemente continuo en  $X$ .

El operador  $T$  es acotado si y sólo si existe una constante  $c > 0$  tal que  $\|Tu\|_Y \leq c\|u\|_X$  para cada  $u \in X$ .

Si  $f, g \in L_1(\mathbb{R})$ , entonces la convolución de  $f$  y  $g$ , denotada  $f * g$ , se define por

$$(f * g)(t) = \int_{\mathbb{R}} f(t - z)g(z)dz.$$

Un sistema de funciones  $\{\phi_j, j \in \mathbb{Z}\}$ ,  $\phi_j \in L_2(\mathbb{R})$ , se llama ortonormal si

$$\int_{\mathbb{R}} \phi_j(t)\overline{\phi_k(t)}dt = \delta_{jk},$$

donde  $\delta_{jk}$  es la delta de Kronecker. Es decir,

$$\delta_{jk} = \begin{cases} 1, & \text{si } j = k; \\ 0, & \text{si } j \neq k. \end{cases}$$

Un sistema ortonormal se llama una base en un subespacio  $V$  de  $L_2(\mathbb{R})$  si cualquier función  $f \in V$  tiene una representación de la forma

$$f(t) = \sum_j c_j \phi_j(t),$$

donde los coeficientes  $c_j$  satisfacen  $\sum_j |c_j|^2 < \infty$ . En lo que sigue se utilizará la notación  $\sum_j = \sum_{j=-\infty}^{\infty}$ ,  $\int_{\mathbb{R}} = \int_{-\infty}^{\infty}$ ,  $\|f\|_{L_2} = \|f\|_2$  y  $\langle \cdot, \cdot \rangle_2$ .

La función característica del conjunto  $A$ ,  $\chi_A$ , se define por

$$\chi_A(t) = \begin{cases} 1, & t \in A; \\ 0, & t \notin A. \end{cases}$$

También se utilizará la notación  $I\{A\}$  para denotar esta función y la llaman función indicadora.

## 1.2. Espacio de probabilidad

A partir de la teoría de la medida, la teoría de probabilidad ha alcanzado un alto grado de formalización. En las siguientes líneas se presentan algunos elementos básicos sobre el tema, para un estudio profundo se puede consultar [6].

**Definición 1.2.1.** *Sea  $\Omega$  un conjunto no vacío y  $\mathcal{A}$  una colección de subconjuntos de  $\Omega$ .  $\mathcal{A}$  es una  $\sigma$ -álgebra sobre  $\Omega$  si y sólo si se satisfacen las siguientes condiciones*

- i)  $\Omega \in \mathcal{A}$
- ii) Si  $A_1, A_2, \dots$  es una sucesión contable de elementos de  $\mathcal{A}$ , entonces  $\bigcup A_n \in \mathcal{A}$
- iii) Si  $A \in \mathcal{A}$ , entonces  $A^c \in \mathcal{A}$ , donde  $A^c$  es el complemento de  $A$  en  $\Omega$ .

La pareja  $(\Omega, \mathcal{A})$  se llama espacio medible y a los elementos de  $\mathcal{A}$ , conjuntos medibles.

**Definición 1.2.2.** *Sea  $\mathcal{C}$  una colección de subconjuntos de  $\Omega$ . Por  $\sigma$ -álgebra minimal que contiene a  $\mathcal{C}$  o la  $\sigma$ -álgebra que genera a  $\mathcal{C}$ , denotada  $\sigma(\mathcal{C})$ , se entiende una  $\sigma$ -álgebra de subconjuntos de  $\Omega$  tal que si  $\mathcal{K}$  es otra  $\sigma$ -álgebra que contiene a  $\mathcal{C}$ , entonces  $\mathcal{C} \subset \sigma(\mathcal{C}) \subset \mathcal{K}$ .*

La  $\sigma$ -álgebra  $\mathcal{B}$  generada por todos los conjuntos abiertos de  $\mathbb{R}^n$ , se llama álgebra de Borel y los elementos en  $\mathcal{B}$  se llaman conjuntos de Borel. Esta  $\sigma$ -álgebra es de gran interés en diversos campos de la matemática, en particular en la teoría de probabilidades.

**Definición 1.2.3.** *Una probabilidad  $P$  es una medida normalizada sobre un espacio medible  $(\Omega, \mathcal{A})$ ; esto es,  $P$  es una función de valor real la cual asigna a todo  $A \in \mathcal{A}$  el número  $P(A)$  tal que*

## 6 Preliminares

---

i)  $P(\Omega) = 1$

ii) Si  $A_1, A_2, \dots$  es una sucesión contable de elementos de  $\mathcal{A}$  disjuntos dos a dos, entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

iii)  $P(A) \geq 0$  para todo  $A \in \mathcal{A}$ .

La tripla  $(\Omega, \mathcal{A}, P)$  se llama espacio de probabilidad.  $P(A)$  se lee como la probabilidad del evento  $A$ .

Algunas consecuencias de la definición (1.2.3) son:

1.  $P(\emptyset) = 0$ .
2. Sean  $A$  y  $B$  eventos. Si  $A \subset B$ , entonces  $P(A) \leq P(B)$ .
3. Si  $A_1, A_2, \dots, A_n$  son eventos disjuntos dos a dos, entonces

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k).$$

4.  $P(A^c) = 1 - P(A)$ , para todo  $A \in \mathcal{A}$ .
5. Si  $\{A_n\}$  es una sucesión contable de eventos, entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n), \quad \text{desigualdad de Boole.}$$

### 1.2.1. Variable aleatoria

**Definición 1.2.4.** Una variable aleatoria  $X$  es una función de valor real cuyo dominio es  $\Omega$  y la cual es  $\mathcal{A}$ -medible, esto es, para cualquier número real  $x$ ,  $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}$ .

El conjunto  $\{\omega \in \Omega : X(\omega) \leq x\}$  se llama conjunto de eventos elementales, se denotará por  $[X \leq x]$ .

Si  $X$  es una variable aleatoria, la función de distribución  $F_X$  se define por

$$F_X(x) = P[X \leq x], \quad \text{para todo } x \in \mathbb{R}.$$



Note que diferentes variables aleatorias pueden tener la misma función de distribución. Por ejemplo, sea  $\Omega = \{C, S\}$ , si  $P(C) = P(S) = 1/2$  y si  $X$  y  $Y$  son variables aleatorias definidas por  $X(C) = 1, X(S) = 0, Y(C) = 0$  y  $Y(S) = 1$ , entonces

$$F_X(x) = F_Y(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1/2, & \text{si } 0 \leq x < 1 \\ 1, & \text{si } x \geq 1. \end{cases}$$

Si  $X$  es una variable aleatoria, entonces la función de distribución  $F_X$  tiene las siguientes propiedades:

1.  $F_X$  es no decreciente, es decir, si  $-\infty < a < b < \infty$ , entonces

$$F_X(a) \leq F_X(b)$$

2.  $\lim_{x \rightarrow \infty} F_X(x) = 1$  y  $\lim_{x \rightarrow -\infty} F_X(x) = 0$
3.  $F_X$  es continua por la derecha, esto es,

$$\lim_{h \downarrow 0} F_X(x+h) = F_X(x), \quad \forall x.$$

Una función de distribución  $F$  se llama absolutamente continua, si existe una función medible Borel  $f$  sobre  $\mathbb{R}$  tal que

$$F(x) = \int_{-\infty}^x f(t) dt \quad \forall x.$$

La función  $f$  se llama la densidad de  $F$ .

Si  $X$  es una variable aleatoria con función de distribución absolutamente continua y densidad  $f$ , entonces el valor esperado de  $X$  es dado por

$$E(X) = \int_{\mathbb{R}} x f(x) dx,$$

siempre que la integral sea finita.

### 1.2.2. Procesos estocásticos

Una variable aleatoria siempre tiene asociada una distribución de probabilidad que mide la probabilidad de ocurrencia de sus distintos resultados. Cuando la variable aleatoria cambia con el tiempo, se le puede asociar una distribución de probabilidad que también varía con el tiempo. En tales ambientes resulta útil definir un proceso estocástico [57].

**Definición 1.2.5.** *Sea  $I \subseteq \mathbb{R}$  un conjunto de índices y  $(\Omega, \mathcal{A}, P)$  un espacio de probabilidad. Una función  $X : I \times \Omega \rightarrow \mathbb{R}^n$  es un proceso estocástico si para cada  $t \in I$  fijo, la función  $X_t : \Omega \rightarrow \mathbb{R}^n$  es una variable aleatoria, que representa el valor del proceso  $X(t, \omega)$ ,  $\omega \in \Omega$ . Si  $\omega \in \Omega$  es fijo, la aplicación  $I \rightarrow \mathbb{R}^n$  tal que  $t \mapsto X_t(\omega)$  se llama la trayectoria o realización del proceso  $X$ .*

Los valores que toma el proceso en  $\mathbb{R}^n$  se llaman estados del proceso. Si el conjunto  $I$  es contable, el proceso estocástico  $X$  se dice que es de tiempo discreto. Por otro lado, si  $I$  es un intervalo de los reales no negativos, el proceso estocástico es de tiempo continuo.

Si  $X$  es un proceso estocástico continuo, entonces

- i)  $X$  es independiente si para todo  $t, s \in I$   $s \neq t$ , las variables aleatorias asociadas  $X_s$  y  $X_t$  son independientes.
- ii)  $X$  es independientemente distribuida, si la distribución de probabilidad  $F_{X_t}$  es la misma para cada  $t \in I$ .
- iii)  $X$  tiene incrementos independientes si para cada  $n \geq 1$  y para cualquier partición del intervalo  $I$ ,  $t_0 < t_1 < \dots < t_n$ , las diferencias

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

son variables aleatorias independientes.

- iv)  $X$  tiene incrementos estacionarios si  $X_t - X_s \stackrel{d}{=} X_{t+h} - X_{t+s}$  para cada  $t, s, t+h, s+h$  en  $I$ ,  $s < t$  y  $h > 0$ . El símbolo  $\stackrel{d}{=}$  significa que los términos en comparación tienen la misma distribución de probabilidad. La estacionariedad de un proceso estocástico alude a que la distribución de probabilidad de la diferencia entre dos variables aleatorias permanece invariante bajo cualquier traslación temporal.

Una serie de tiempo es la realización de un proceso estocástico. En otras palabras, una serie de tiempo se puede considerar como una colección de variables aleatorias  $\{X_t : t \in I\}$  ( p.e., [30], [51] o [71]).

## 1.3. Teoría asintótica

La distribución de un estimador puede cambiar con el tamaño muestral. En ocasiones no es posible obtener cuantitativamente el valor medio de un estimador para saber si es insesgado o no. Lo mismo puede ocurrir con su varianza para un tamaño de muestra dado. En estas situaciones determinar las propiedades analíticas del estimador en muestras finitas es muy complicado y se pasa a estudiar las propiedades asintóticas. El conocimiento del comportamiento en el límite de la distribución de un estimador, puede utilizarse para inferir una distribución aproximada para el estimador obtenido en una muestra finita. Para ello necesitaremos conceptos de teoría asintótica.

Cuando se desea determinar si un estimador es bueno, es decir, consistente, eficiente, suficiente o sesgado, es difícil determinarlo. Sin embargo se puede aproximar al comportamiento a partir de su distribución para tamaños muestrales altos. Esto se puede hacer a partir del límite de la distribución del estimador.

### 1.3.1. Convergencia en probabilidad.

**Definición 1.3.1. Convergencia en probabilidad** La variable aleatoria  $x_n$  converge en probabilidad a una constante  $c$ , si

$$\lim_{n \rightarrow \infty} \text{Prob}(|x_n - c| > \epsilon) = 0$$

para cualquier  $\epsilon > 0$ .

La convergencia en probabilidad implica que los valores cercanos a  $c$  que toma la variable son cada vez más probables, a medida que  $n$  aumenta.

**Definición 1.3.2. Estimador Consistente.** Un estimador  $\hat{\theta}$  de un parámetro  $\theta$  es un estimador consistente de  $\theta$  si y sólo si

$$\lim_{n \rightarrow \infty} \text{Prob}(|\hat{\theta} - \theta| < \epsilon) = 1$$

para cualquier  $\epsilon > 0$ .

### 1.3.2. Convergencia en distribución y distribución límite

Sea  $x_n$  una sucesión de variables aleatorias, cuyo elemento representativo  $x_n$ , es una variable aleatoria obtenida de una muestra de tamaño  $n$ . Si  $x_n$  tiene una función de distribución  $F_n(x)$ , es porque converge a la función de distribución  $F(x)$  de la variable aleatoria  $x$  en todo punto de continuidad de  $F(x)$ .

**Definición 1.3.3. Convergencia en Distribución.** Decimos que la sucesión de variables aleatorias  $\{x_n\}$  converge en distribución a una variable aleatoria  $x$  con función de distribución  $F(x)$  si

$$\lim_{n \rightarrow \infty} |F_n(x) - F(x)| = 0$$

en todos los puntos en los que  $F(x)$  sea continua.

Cabe notar que la convergencia en distribución esta relacionada con la distribución de probabilidad asociada a  $\{x_n\}$  a medida que  $n$  aumenta; no implica la convergencia de los valores que toma cada variable aleatoria  $x_n$ .

**Definición 1.3.4. Distribución límite** Si  $x_n$  converge en distribución a  $x$ , siendo  $F(x)$  la función de distribución de  $x$ , entonces  $F(x)$  es la distribución límite de  $x$ . Se representa  $x_n \xrightarrow{d} x$ .

### 1.3.3. Distribuciones asintóticas

Las distribuciones asintóticas obtenidas a partir del teorema de límite central dependen de parámetros desconocidos, ahora bien, lo que haremos es derivar las distribuciones asintóticas de los estimadores que nos interesan. La forma más común de plantear una distribución asintótica es construirla a partir de la distribución límite conocida de la variable aleatoria. Si

$$\sqrt{n}[(\bar{x}_n - \mu)/\sigma] \xrightarrow{d} N[0, 1],$$

entonces, aproximadamente, o asintóticamente,  $x_n \sim N[\mu, \sigma^2/n]$ , lo que escribiremos como

$$x_n \xrightarrow{a} N[\mu, \sigma^2/n].$$

Mediante la afirmación de que  $\bar{x}_n$  se distribuye asintóticamente como una normal con media  $\mu$  y varianza  $\sigma^2/n$ , quiere decir que esta distribución

normal es una aproximación a la verdadera distribución finita, no que la verdadera distribución es exactamente una normal.

Extendiendo la definición, supongamos que  $\hat{\theta}$  es un estimador del vector de parámetros  $\theta$ . La distribución asintótica del vector  $\hat{\theta}$  se obtiene de la distribución límite,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N[0, \mathbf{V}] \quad (1.3.1)$$

con  $\mathbf{V}$  una matriz de covarianza de  $2 \times 2$ , lo que implica que

$$\hat{\theta} \xrightarrow{d} N[\theta, \frac{1}{n}\mathbf{V}].$$

La matriz de covarianzas de la distribución asintótica es la **matriz de covarianzas asintóticas** y se designa por

$$Asy.Var[\hat{\theta}] = \frac{1}{n}\mathbf{V}.$$

Recordemos que una matriz es semidefinida positiva si  $x'Ax \geq 0$ ,  $\forall x \in \mathbb{R}$ , con  $A$  simétrica y  $f(x) = x'Ax$ , donde  $x'$  es el vector transpuesto ([4]).

**Definición 1.3.5. Normalidad y eficiencia asintótica.**  $\hat{\theta}$  es asintóticamente normal si 1.3.1 se cumple y asintóticamente eficiente si la diferencia entre la matriz de covarianzas de cualquier otro estimador consistente que sea asintóticamente normal y  $1/n\mathbf{V}$  es una matriz semidefinida positiva.



## CAPÍTULO 2

---

### Wavelets

---

#### 2.1. Introducción a los Wavelets

El origen de la descomposición de una señal en wavelets está en la necesidad de conocer las características y particularidades de la señal en diferentes instantes de tiempo. La principal virtud de las wavelets es que permite modelar procesos que dependen fuertemente del tiempo y para los cuales su comportamiento no tiene porqué ser suave [1], [13], [15], [16], [22]. Una de las ventajas de las wavelets frente a los métodos clásicos, como la transformada de Fourier, es que en el segundo caso se maneja una base de funciones bien localizada en frecuencia pero no en tiempo, esto es, el análisis en frecuencia obtenido del análisis de Fourier es insensible a perturbaciones que supongan variaciones instantáneas y puntuales de la señal como picos debidos a conmutaciones o variaciones muy lentas como tendencias. En otras palabras, si  $f$  es una señal ( $f$  es una función definida en todo  $\mathbb{R}$  y tiene energía finita  $\int_{-\infty}^{\infty} |f(t)|^2 dt$ ). La transformada de Fourier  $\hat{f}(\omega)$  proporciona la información global de la señal en el tiempo localizada en frecuencia. Sin embargo,  $\hat{f}(\omega)$  no particulariza la información para intervalos de tiempo específicos, ya que

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$$

y la integración es sobre todo tiempo ([24]). Así, la imagen obtenida no contiene información sobre tiempos específicos, sino que sólo permite calcular el espectro de amplitud total  $|\hat{f}(\omega)|$ , mientras que la mayoría de las wavelets interesantes presentan una buena localización en tiempo y en frecuencia, disponiendo incluso de bases de wavelets con soporte compacto.

En este capítulo se presenta una introducción a las transformadas de Fourier y wavelets.

## 2.2. Transformada de Fourier

En esta sección se recordará la definición y algunas propiedades importantes de la transformada de Fourier. En particular, se hará un resumen de resultados básicos de análisis de Fourier omitiendo sus pruebas, las cuales se pueden encontrar en algunos de los siguientes textos [7], [24], [53], [65], [69].

**Definición 2.2.1.** Sea  $f \in L_1(\mathbb{R})$  y  $\omega \in \mathbb{R}$ . La transformada de Fourier de  $f$  en  $\omega$  se define por

$$\hat{f}(\omega) := \int_{\mathbb{R}} f(t)e^{-i\omega t} dt \quad (2.2.1)$$

donde  $tx = \sum_{j=1}^n t_j x_j$ , es el producto interno usual de  $\mathbb{R}^n$ .

Como

$$\int_{\mathbb{R}} |f(t)| |e^{-it\omega}| dt = \int_{\mathbb{R}} |f(t)| dt = \|f\|_{L_1} < \infty$$

se tiene que la transformada de Fourier está bien definida. La aplicación  $f \mapsto \hat{f}$  se llama transformación de Fourier y se denota por  $\mathcal{F}$  ( $\mathcal{F}(f) = \hat{f}$ ). La función  $\hat{f}$  es continua y tiende a cero cuando  $|\omega| \rightarrow \infty$  (Lema de Riemann-Lebesgue). Es claro que  $\mathcal{F}(af + bg) = a\mathcal{F}(f) + b\mathcal{F}(g)$ , para cada  $a, b \in \mathbb{R}$ .

En general  $\hat{f}$  no es una función integrable, por ejemplo, sea

$$f(t) = \begin{cases} 1, & |t| < 1; \\ 0, & |t| > 1. \end{cases}$$

Entonces

$$\begin{aligned} \hat{f}(\omega) &= \int_{-1}^1 e^{-it\omega} dt = \left[ \frac{e^{-i\omega} - e^{i\omega}}{-i\omega} \right] \\ &= 2 \frac{\text{sen } \omega}{\omega} \notin L_1(\mathbb{R}). \end{aligned}$$



Si  $\hat{f}(\omega)$  es integrable, entonces existe una versión continua de  $f$  y se puede obtener la fórmula de inversión de Fourier

$$f(t) = \mathcal{F}^{-1}(\hat{f}(\omega)) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) e^{i\omega t} d\omega. \quad (2.2.2)$$

La siguiente proposición recoge algunas propiedades fundamentales de la transformada de Fourier.

**Proposición 2.2.2.** Sean  $f, g \in L_1(\mathbb{R})$ , entonces

1.  $\widehat{(T_x f)}(\omega) = e^{-i\omega x} \hat{f}(\omega)$ , donde  $(T_a f)(t) = f(t - a)$ .
2.  $(T_x \hat{f})(\omega) = \widehat{(e^{ix(\cdot)} f)}(\omega)$
3.  $\widehat{f * g} = \hat{f} \hat{g}$
4. Si  $\epsilon > 0$  y  $g_\epsilon(t) = g(\epsilon t)$  entonces  $\hat{g}_\epsilon(\omega) = \epsilon^{-1} \hat{g}(\omega/\epsilon)$ .

Otro resultado útil es el siguiente: Si  $f, g \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ , entonces

$$\|f\|_2^2 = \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega \quad (\text{fórmula de Plancherel}) \quad (2.2.3)$$

$$\langle f, g \rangle_2 = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} d\omega \quad (\text{fórmula de Parseval}). \quad (2.2.4)$$

Por extensión, la transformada de Fourier se puede definir para cualquier  $f \in L_2(\mathbb{R})$ . En virtud a que el espacio  $L_1(\mathbb{R}) \cap L_2(\mathbb{R})$  es denso en  $L_2(\mathbb{R})$ . Luego, por isometría (excepto por el factor  $1/2\pi$ ) se define  $\hat{f}$  para cualquier  $f \in L_2(\mathbb{R})$ , y las fórmulas (2.2.3) y (2.2.4) permanecen válidas para todo  $f, g \in L_2(\mathbb{R})$ .

En teoría de señales, la cantidad  $\|f\|_2$  mide la energía de la señal, mientras que  $\|\hat{f}\|_2$  representa el espectro de potencia de  $f$ .

Si  $f$  es tal que  $\int_{\mathbb{R}} |t|^k |f(t)| dt < \infty$ , para algún entero  $k \geq 1$ , entonces

$$\frac{d^k}{d\omega^k} \hat{f}(\omega) = \int_{\mathbb{R}} (-it)^k e^{-i\omega t} f(t) dt. \quad (2.2.5)$$

Recíprocamente, si  $\int_{\mathbb{R}} |\omega|^k |\hat{f}(\omega)| d\omega < \infty$ , entonces

$$(i\omega)^k \hat{f}(\omega) = \mathcal{F}(f^{(k)})(\omega). \quad (2.2.6)$$

### 2.2.1. Serie de Fourier

Sea  $f$  una función  $2\pi$ -periódica en  $\mathbb{R}$ . Se escribirá  $f \in L_p(0, 2\pi)$  si

$$f(t)\chi_{[0,2\pi]}(t) \in L_p(0, 2\pi), \quad p \geq 1.$$

Cualquier función  $f$ ,  $2\pi$ -periódica en  $\mathbb{R}$ , tal que  $f \in L_2(0, 2\pi)$ , se puede representar por una serie de Fourier convergente en  $L_2(0, 2\pi)$

$$f(t) = \sum_n c_n e^{int},$$

donde los coeficientes de Fourier son dados por

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt.$$

Se puede verificar que si  $f \in L_1(\mathbb{R})$ , entonces la serie

$$S(t) = \sum_k f(t + 2k\pi) \tag{2.2.7}$$

converge casi para todo  $t$  y pertenece a  $L_1(0, 2\pi)$ . Además, los coeficientes de Fourier de  $S(t)$  están dados por

$$c_k = \frac{1}{2\pi} \hat{f}(k) = \mathcal{F}^{-1}(f)(-k).$$

En efecto, para ver la expresión (2.2.7), basta probar que

$$\int_0^{2\pi} \sum_k |f(t + 2k\pi)| dt < \infty.$$

Para la segunda parte se calcula los coeficientes de Fourier

$$\frac{1}{2\pi} \int_0^{2\pi} \left[ \sum_k f(t + 2k\pi) \right] e^{-ikt} dt.$$

Intercambiando la suma con la integral se obtiene

$$\begin{aligned} \sum_k \frac{1}{2\pi} \int_0^{2\pi} f(t + 2k\pi) e^{-ikt} dt &= \sum_k \frac{1}{2\pi} \int_{2\pi k}^{2\pi(k+1)} f(z) e^{-ikz} dz \\ &= \frac{1}{2\pi} \hat{f}(k). \end{aligned}$$

## 2.3. Transformadas wavelets

El análisis wavelets es un método de descomposición de una función o señal usando funciones especiales, las wavelets. La descomposición es similar a la de la transformada de Fourier, donde una señal  $f(t)$  se descompone en una suma infinita de armónicos  $e^{i\omega t}$  de frecuencias  $\omega \in \mathbb{R}$ , cuyas amplitudes son los valores de la transformada de Fourier de  $f$ ,  $\hat{f}(\omega)$ :

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega, \quad \text{donde} \quad \hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.$$

El análisis de Fourier tiene el defecto de la no localidad: el comportamiento de una función en un conjunto abierto, no importa cuán pequeño, influye en el comportamiento global de la transformada de Fourier. No se captan los aspectos locales de la señal tales como cambios bruscos, saltos o picos, que se han de determinar a partir de su reconstrucción.

### 2.3.1. Transformada wavelet continua

La teoría wavelets se basa en la representación de una función en términos de una familia biparamétrica de dilataciones y traslaciones de una función fija  $\psi$ , la wavelet madre que, en general, no es senoidal. Por ejemplo,

$$f(t) = \int_{\mathbb{R}^2} \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \mathcal{W}_\psi f(a, b) da db$$

en donde  $\mathcal{W}_\psi f$  es una transformada de  $f$  definida adecuadamente. También se tiene de modo alterno un desarrollo en serie

$$f(t) = \sum_{j,k} c_{j,k} 2^{j/2} \psi(2^j t - k)$$

en donde se suma sobre las dilataciones en progresión geométrica. Para conservar la norma en  $L_2(\mathbb{R})$  de la wavelet madre  $\psi$ , se insertan los factores  $\frac{1}{\sqrt{|a|}}$  y  $2^{j/2}$ , respectivamente.

**Definición 2.3.1.** Una wavelet  $\psi$  es una función cuadrado integrable tal que la siguiente condición de admisibilidad se tiene

$$C_\psi := \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (2.3.1)$$

donde  $\hat{\psi}(\omega)$  es la transformada de Fourier de  $\psi$ .

**Observación 2.3.1.** Si además  $\psi \in L_1(\mathbb{R})$ , entonces la condición (2.3.1) implica que  $\int_{\mathbb{R}} \psi(t)dt = 0$ . En efecto, por el Lema de Riemann-Lebesgue ([53]),  $\lim_{\omega \rightarrow \infty} \hat{\psi}(\omega) = 0$  y la transformada de Fourier es continua, lo cual implica que  $0 = \hat{\psi}(0) = \int_{\mathbb{R}} \psi(t)dt$ .

Sea  $\psi \in L_2(\mathbb{R})$ . La función dilatada y trasladada se define por

$$\psi_{a,b}(t) := \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad a, b \in \mathbb{R}, \quad a \neq 0.$$

Esta función se obtiene a partir de  $\psi$ , primero por dilatación en el factor  $a$  y, luego, por traslación en  $b$ . Es claro que  $\|\psi_{a,b}\|_2 = \|\psi\|_2$ .

**Definición 2.3.2.** Para  $f, \psi \in L_2(\mathbb{R})$ , la expresión

$$\mathcal{W}_\psi f(a, b) := \int_{\mathbb{R}} f(t) \overline{\psi_{a,b}(t)} dt \tag{2.3.2}$$

se llama la transformada wavelet de  $f$ .

Por la desigualdad de Cauchy, se ve que  $\mathcal{W}_\psi f$  es una función acotada con  $|\mathcal{W}_\psi f(a, b)| \leq \|f\|_2 \|\psi\|_2$ . Note también que

$$\mathcal{W}_\psi f(a, b) = \langle f, \psi_{a,b} \rangle_{L_2(\mathbb{R})} = \langle f, \psi_{a,b} \rangle.$$

La transformada wavelet  $\mathcal{W}_\psi f$  de  $f$  puede ser descrita en términos del producto de convolución. La convolución de dos funciones  $f, g \in L_2(\mathbb{R})$  es dada por

$$(f * g)(t) = \int_{\mathbb{R}} f(t-z)g(z)dz.$$

Observe que esta fórmula está definida para al menos todo  $t \in \mathbb{R}$ , pero  $f * g$  no necesariamente está en  $L_2(\mathbb{R})$ . Usando la notación  $\tilde{\psi}(t) = \overline{\psi(-t)}$ , se tiene  $\mathcal{W}_\psi f(a, b) = (f * \tilde{\psi}_{a,0})(b)$ . Note también que  $\hat{\psi}_{a,b}(\omega) = \sqrt{|a|} \hat{\psi}(a\omega) e^{-i\omega b}$ . Estos hechos se aplicarán en la prueba de la siguiente proposición, la cual establece la fórmula de Plancherel para la transformada wavelet.

**Proposición 2.3.3.** Sea  $\psi \in L_2(\mathbb{R})$  y satisface la condición (2.3.1). Entonces para cualquier  $f \in L_2(\mathbb{R})$ , las siguientes relaciones se tienen

1. *Isometría*

$$\int_{\mathbb{R}} |f(t)|^2 dt = \frac{1}{C_\psi} \int_{\mathbb{R}^2} |\mathcal{W}_\psi f(a, b)|^2 db \frac{da}{a^2}$$

2. Fórmula de inversión

$$f(t) = \frac{1}{C_\psi} \int_{\mathbb{R}^2} \mathcal{W}_\psi f(a, b) \psi_{a,b}(t) db \frac{da}{a^2}$$

*Demostración.* Es fácil verificar que  $(f * \tilde{\psi}_{a,0})(b) = \sqrt{|a|} \mathcal{F}^{-1}\{\hat{f}(\omega) \tilde{\hat{\psi}}(a\omega)\}$ .

En consecuencia,

$$\begin{aligned} \int_{\mathbb{R}^2} |\mathcal{W}_\psi f(a, b)|^2 db \frac{da}{a^2} &= \int_{\mathbb{R}} \int_{\mathbb{R}} |(f * \tilde{\psi}_{a,0})(b)|^2 db \frac{da}{a^2} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |a| |\mathcal{F}^{-1}(\hat{f}(\cdot) \tilde{\hat{\psi}}(a\cdot))(\omega)|^2 d\omega \frac{da}{a^2} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |\hat{f}(\omega)|^2 |\hat{\psi}(a\omega)|^2 d\omega \frac{da}{|a|} \\ &= \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \left[ \int_{\mathbb{R}} |\hat{\psi}(a\omega)|^2 \frac{da}{|a|} \right] d\omega \\ &= C_\psi \int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega = C_\psi \|f\|_2^2. \end{aligned}$$

Observe que se utilizó el teorema de Fubini y la fórmula de Plancherel para la transformada de Fourier.

Para simplificar los cálculos en la fórmula de inversión, suponga que  $f, \hat{f} \in L_1(\mathbb{R})$ .

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{W}_\psi f(a, b) \psi_{a,b}(t) db &= \sqrt{|a|} \int_{\mathbb{R}} \mathcal{F}^{-1}(\hat{f}(\cdot) \tilde{\hat{\psi}}(a\cdot))(\omega) \psi_{a,b}(t) d\omega \\ &= \sqrt{|a|} \int_{\mathbb{R}} \hat{f}(\omega) \tilde{\hat{\psi}}(a\omega) \mathcal{F}^{-1}(g)(\omega) d\omega, \end{aligned}$$

donde  $g(b) := \psi_{a,b}(t)$ . Ahora, la transformada inversa de Fourier de  $g$  es

$$\begin{aligned} \mathcal{F}^{-1}(g)(\omega) &= \frac{1}{2\pi} \int_{\mathbb{R}} g(b) e^{i\omega b} db \\ &= \frac{1}{2\pi} \sqrt{|a|} \int_{\mathbb{R}} \psi(z) e^{-ia\omega z} e^{i\omega t} dz \\ &= \frac{1}{2\pi} \sqrt{|a|} \hat{\psi}(a\omega) e^{i\omega t}. \end{aligned}$$

Sustituyendo e integrando respecto a  $a^{-2}da$  se obtiene

$$\begin{aligned}
 \int_{\mathbb{R}^2} \mathcal{W}_\psi f(a, b) \overline{\mathcal{W}_\psi g(a, b)} db \frac{da}{a^2} &= \frac{1}{2\pi} \int_{\mathbb{R}} |a| \left[ \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{\psi}(a\omega)} e^{i\omega t} d\omega \right] \frac{da}{a^2} \\
 &= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) \left[ \int_{\mathbb{R}} |\hat{\psi}(a\omega)|^2 \frac{da}{|a|} \right] e^{i\omega t} d\omega \\
 &= C_\psi \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) e^{i\omega t} d\omega \\
 &= C_\psi f(t).
 \end{aligned}$$

□

Otro resultado de interés que se presentará en la siguiente proposición, es la fórmula de Parseval para la transformada wavelet.

**Proposición 2.3.4.** *Sea  $\psi \in L_2(\mathbb{R})$  y satisface la condición (2.3.1). Entonces para cualquier  $f, g \in L_2(\mathbb{R})$ , se tienen*

$$\langle f, g \rangle_{L_2(\mathbb{R})} = \frac{1}{C_\psi} \int_{\mathbb{R}^2} \mathcal{W}_\psi f(a, b) \overline{\mathcal{W}_\psi g(a, b)} \frac{dad b}{a^2}$$

*Demostración.* Como  $(f * \tilde{\psi}_{a,0})(b) = \sqrt{|a|} \mathcal{F}^{-1}\{\hat{f}(\omega) \tilde{\psi}(a\omega)\}$  o de manera equivalente,  $\mathcal{F}(f * \tilde{\psi}_{a,0})(\omega) = \sqrt{|a|} \hat{f}(\omega) \tilde{\psi}(a\omega)$ , entonces

$$\int_{\mathbb{R}} \mathcal{W}_\psi f(a, b) \overline{\mathcal{W}_\psi g(a, b)} db = |a| \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} |\hat{\psi}(a\omega)|^2 d\omega,$$

ahora, integrando respecto a  $a^{-2}da$  se sigue

$$\begin{aligned}
 \int_{\mathbb{R}^2} \mathcal{W}_\psi f(a, b) \overline{\mathcal{W}_\psi g(a, b)} db \frac{da}{a^2} &= \int_{\mathbb{R}} |a| \left[ \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} |\hat{\psi}(a\omega)|^2 d\omega \right] \frac{da}{a^2} \\
 &= \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} \left[ \int_{\mathbb{R}} |\hat{\psi}(a\omega)|^2 \frac{da}{|a|} \right] d\omega \\
 &= C_\psi \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} d\omega \\
 &= C_\psi \langle \hat{f}, \hat{g} \rangle_{L_2(\mathbb{R})} = C_\psi \langle f, g \rangle_{L_2(\mathbb{R})}.
 \end{aligned}$$

Note que se aplicó el teorema de Fubini, y en el último renglón de la expresión anterior, la fórmula de Parseval para la transformada de Fourier. □

En la siguiente proposición se listan algunas propiedades.

**Proposición 2.3.5.** *Sean  $\psi$  y  $\varphi$  wavelets y  $f, g \in L_2(\mathbb{R})$ . Entonces*

1.  $\mathcal{W}_\psi(\alpha f + \beta g)(a, b) = \alpha \mathcal{W}_\psi f(a, b) + \beta \mathcal{W}_\psi g(a, b)$ ,  $\alpha, \beta \in \mathbb{R}$ .
2.  $\mathcal{W}_{\alpha\psi + \beta\varphi} f(a, b) = \bar{\alpha} \mathcal{W}_\psi f(a, b) + \bar{\beta} \mathcal{W}_\varphi f(a, b)$ ,  $\alpha, \beta \in \mathbb{R}$ .
3.  $\mathcal{W}_\psi(T_c f)(a, b) = \mathcal{W}_\psi f(a, b - c)$ , donde  $T_c$  es el operador traslación definido por  $T_c f(t) = f(t - c)$ .
4.  $\mathcal{W}_\psi(D_c f)(a, b) = \sqrt{c} \mathcal{W}_\psi f(ca, cb)$ , donde  $D_c$  es el operador dilatación definido por  $D_c f(t) = \sqrt{c} f(ct)$ .

### 2.3.2. Transformada wavelet discreta

La transformada wavelet continua introduce cierta redundancia, pues la señal original se puede reconstruir completamente calculando  $\mathcal{W}_\psi f(a, \cdot)$  para una cantidad numerable de escalas, por ejemplo, potencias enteras de 2. Esto es, si se elige la escala  $a = 2^{-j}$  para cada  $j \in \mathbb{Z}$ , y también se discretiza en el dominio del tiempo en los puntos  $b = 2^{-j}k$ ,  $k \in \mathbb{Z}$ , la familia de wavelets será ahora dada por

$$\psi_{2^{-j}, 2^{-j}k}(t) = \frac{1}{\sqrt{2^{-j}}} \psi\left(\frac{t - 2^{-j}k}{2^{-j}}\right) = 2^{j/2} \psi(2^j t - k), \quad \forall j, k \in \mathbb{Z}.$$

Se utilizará la notación  $\psi_{jk}$  para denotar la wavelet  $\psi$  comprimida  $2^j$  y trasladada el entero  $k$ , es decir,  $\psi_{jk}(t) = 2^{j/2} \psi(2^j t - k)$ .

Con la elección de  $a = 2^{-j}$  y  $b = 2^{-j}k$ , observe que el muestreo en el tiempo se ajusta proporcionalmente a la escala, es decir, a mayor escala se toma puntos más distantes, ya que se busca información global, mientras que a menor escala se buscan detalles de la señal, por tal motivo se muestrea en puntos menos distantes entre si. Para otras elecciones de  $a$  y  $b$  se puede consultar [12].

**Definición 2.3.6.** *Una función  $\psi \in L_2(\mathbb{R})$  es una wavelet si la familia de funciones  $\psi_{jk}$  definidas por*

$$\psi_{jk}(t) = 2^{j/2} \psi(2^j t - k), \quad \forall j, k \in \mathbb{Z}, \tag{2.3.3}$$

*es una base ortonormal en el espacio  $L_2(\mathbb{R})$ .*

Una condición suficiente para la reconstrucción de una señal  $f$  es que la familia de dilatadas y trasladadas  $\psi_{jk}$  forme una base ortonormal en el espacio  $L_2(\mathbb{R})$ , ver [19] y [31] para más detalles. Si esto se tiene, cualquier función  $f \in L_2(\mathbb{R})$  se puede escribir como

$$f(t) = \sum_{j,k} c_{j,k} \psi_{jk}(t) \quad (2.3.4)$$

o teniendo en cuenta (2.3.3) como

$$f(t) = \sum_{j,k} c_{j,k} 2^{j/2} \psi(2^j t - k),$$

donde  $c_{j,k} = \langle f, \psi_{2^{-j}, 2^{-j}k} \rangle = \mathcal{W}_\psi f(2^{-j}, 2^{-j}k)$ .

**Definición 2.3.7.** Para cada  $f \in L_2(\mathbb{R})$  el conjunto bidimensional de coeficientes

$$c_{j,k} = \langle f, \psi_{jk} \rangle = \int_{\mathbb{R}} 2^{j/2} f(t) \overline{\psi(2^j t - k)} dt$$

se llama la transformada wavelet discreta de  $f$ .

En consecuencia, la expresión (2.3.4) se puede escribir en forma alterna como

$$f(t) = \sum_{j,k} \langle f(t), \psi_{jk}(t) \rangle \psi_{jk}(t). \quad (2.3.5)$$

La serie (2.3.5) se llama representación wavelet de  $f$ .

**Observación 2.3.2.**  $\psi_{jk}(t)$  es muy apropiada para representar detalles más finos de la señal como oscilaciones rápidas. Los coeficientes wavelet  $c_{j,k}$  miden la cantidad de fluctuaciones sobre el punto  $t = 2^{-j}k$  con una frecuencia determinada por el índice de dilatación  $j$ .

Es interesante notar que  $c_{j,k} = \mathcal{W}_\psi f(2^{-j}, 2^{-j}k)$  es la transformada wavelet de  $f$  en el punto  $(2^{-j}, 2^{-j}k)$ . Estos coeficientes analizan la señal mediante la wavelet madre  $\psi$ .

## 2.4. Filtros Lineales

El comportamiento de una variable puede verse afectado por la interacción con otras variables, en especial variables exógenas; esto puede generarle



cambios y comportamientos extraños. Un método conocido para extraer o separar la información original y las perturbaciones, es el método de filtrado.

El filtrado es un campo de investigación universal usado por los científicos en astronomía, biología, ingeniería, y física, como también en economía y finanzas. Tradicionalmente, los filtros en economía y finanzas son usados para extraer componentes de una serie de tiempo tales como tendencias, volatilidad, estacionalidad, ciclos de negocios, y ruidos [26].

Popularmente el filtro es un elemento que utilizamos en nuestro cotidiano vivir para separar un componente de otros, que se encuentran mezclados; análogamente en el contexto matemático el filtrado es un procedimiento que se usa en el tratamiento de señales que permite separar o extraer un componente de un conjunto de componentes de una señal por medio de una transformación lineal, por lo general la convolución.

En palabras cortas, un *filtro lineal* es un método que tiene como objetivo la identificación y extracción de ciertas características de las series de tiempo (para nuestro caso) ampliamente usados en el procesamiento de señales.

### 2.4.1. Filtros con dominio en el tiempo

Una serie de tiempo discreta es una sucesión de observaciones ordenadas en el tiempo, donde el tiempo puede tomar valores desde menos infinito hasta infinito,

$$\{x_t\}_{t=-\infty}^{\infty} = \{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\}$$

La definición de serie de tiempo se profundizará en el capítulo 3.

El esquema de un proceso de filtrado es

$$x_t \rightarrow \boxed{\text{filtro}} \rightarrow y_t \quad (2.4.1)$$

Un filtro lineal convierte una serie de tiempo  $x_t$  en otra serie de tiempo  $y_t$  a partir de la transformación lineal (2.4.1). La serie  $y_t$  que se obtiene a la salida del filtro es el resultado de la convolución de la entrada  $x_t$  con un vector de coeficientes  $\omega_t$ . Los elementos del vector  $\omega_t$  son los coeficientes del filtro.

La convolución de la entrada el vector  $x_t$  con los coeficientes del vector  $\omega_t$  se expresan matemáticamente como:

$$y_t = \sum_{i=-\infty}^{\infty} \omega_i x_{t-i} \quad (2.4.2)$$

Un filtro que use únicamente valores pasados y presentes de la serie de tiempo  $x_t$  es denominado Filtro Causal o Filtro Físicamente Realizable (FFR).

Cuando se hace un filtrado a una serie en el tiempo y no en la frecuencia, éstos se clasifican de acuerdo con la señal de respuesta, la cual es particular o especial, además la señal de salida  $y_t$  será invariante en el tiempo si  $F(x_{t+h}) = y_{t+h}$ , para todo  $h \in \mathbb{Z}$ .

El *impulso de respuesta* de una señal de impulso unitario<sup>1</sup> es la sucesión de salida del filtro, si éste es finito el filtro se denomina impulso de respuesta finito FIR (finite impulse response), en caso contrario se denomina filtro de impulso de respuesta infinito IIR (infinity impulse response).

### Filtro de Respuesta del Impulso Finito

Los filtros FIR (que sólo puede ser implementado en tiempo discreto) pueden ser descritos como una suma ponderada de entradas con un determinado retardo. Para estos filtros, si la entrada en un determinado instante es cero, la salida será cero a partir de un instante posterior a los retardos inducidos por el filtro. De este modo, solo existirá respuesta por un tiempo finito.

El filtro FIR básico se caracteriza por

$$y_t = \sum_{i=-N}^M \omega_i x_{t-i}$$

Generalmente, la Formula de un filtro FIR en economía y finanzas es un promedio móvil centrado, cuya estructura es

$$y_t = \frac{1}{M + N + 1} (x_{t-M} + \cdots + x_{t-1} + x_t + x_{t+1} + \cdots + x_{t+N})$$

El impulso de respuesta de este filtro es finito y de la forma:

---

<sup>1</sup>La señal de impulso unitario es de la forma

$$x_t = \begin{cases} 1, & \text{si } t = 0; \\ 0, & \text{si } t \neq 0. \end{cases}$$

$$w_i = \begin{cases} \frac{1}{M+N+1}, & \text{si } i = -N, \dots, -1, 0, 1, \dots, M; \\ 0, & \text{en otros casos.} \end{cases}$$

### Filtro de Respuesta al Impulso Infinito

Los filtros IIR, por el contrario, pueden presentar salida aún cuando la entrada sea cero, si las condiciones iniciales son distintas de cero. La energía del filtro decaerá con el tiempo, pero no llegará a ser nula. Por tanto, la respuesta al impulso se extiende infinitamente.

La forma general de un filtro IIR

$$y_t = \sum_{i=1}^L \alpha_i y_{t-i} + \sum_{i=0}^M \omega_i x_{t-i}$$

Donde  $L$  es el valor de rezago de la salida  $y_t$  y  $M$  es el valor de rezago de la entrada  $x_t$ .

### 2.4.2. Filtros con Dominio en la Frecuencia

Una señal puede tener varios componentes periódicos (seno-coseno) posiblemente con diferentes amplitudes, diferentes fases y diferentes frecuencias, esto puede originar una sucesión finita o infinita. La serie que se genera de esta señal tiene la particularidad de presentar una periodicidad determinada por ejemplo series de tiempo de periodicidad mensual, trimestral, semestral o anual. Esto nos permite realizar el análisis desde el dominio de la frecuencia, la base de este análisis aplicado a las series temporales es mediante la transformada de Fourier.

### Frecuencia de Respuesta

En la sección anterior se dijo que la *función de impulso de respuesta* se utiliza para describir y clasificar filtros lineales con dominio en el tiempo; otra forma de clasificarlos es a partir de la *función de frecuencia de respuesta* o *función de transferencia*, siempre y cuando el filtro lineal sea con dominio en la frecuencia.

$$H(f) = \sum_{k=-\infty}^{\infty} \omega_k e^{-i2\pi f k} \quad (2.4.3)$$

donde  $f$  es la frecuencia,  $\omega_k$  es la función de respuesta al impulso de un filtro.

Un aspecto importante, es notar que la función  $H(f)$  es la versión discreta a la presentada en la expresión 2.2.1, la transformada de Fourier de la función impulso de respuesta. Para el análisis de de una señal a partir de la frecuencia, se considera una señal con frecuencia  $f$  conocida, de la forma

$$x_t = e^{i2\pi ft} \quad (2.4.4)$$

En general, si la señal de entrada tiene la forma de la expresión 2.4.4, el filtro lineal 2.4.2 toma la forma

$$\begin{aligned} y_t &= \sum_{k=-\infty}^{\infty} \omega_k e^{i2\pi f(t-k)} \\ &= e^{i2\pi ft} \left( \sum_{k=-\infty}^{\infty} \omega_k e^{i2\pi fk} \right) \end{aligned} \quad (2.4.5)$$

Nótese que el término entre paréntesis es la frecuencia de respuesta (2.4.3), por lo que reescribiendo  $y_t$  tendríamos que

$$y_t = H(f) e^{i2\pi ft}$$

La función de frecuencia de respuesta  $H(f)$  expresada en coordenadas polares, puede verse como

$$H(f) = G(f) e^{i\theta(f)}$$

Esta representación permite identificar dos aspectos importantes de la función frecuencia de respuesta; la primera es la magnitud de la función frecuencia de respuesta  $|H(f)|$  que determina la ganancia de dicha función; es decir,  $G(f)$ . Así  $G(f)$  recibe el nombre de función *ganancia*. La otra parte es la función de fase  $e^{i\theta(f)}$ , que determina el ángulo de fase  $\theta$  del filtro.

### 2.4.3. Filtros Pasa Baja y Pasa Alta

Un filtro se clasifica como pasa baja o pasa alta dependiendo de la relación entre la función ganancia y la frecuencia; si la ganancia es grande cuando la frecuencia es baja o pequeña cuando la frecuencia es alta, se tiene un filtro

pasa baja (relación inversa); en sentido contrario es pasa alta, es decir, si la ganancia es pequeña cuando la frecuencia es baja o grande cuando la frecuencia es alta, se tiene un filtro pasa alta (relación directa).

| Frecuencia | Ganancia |        |
|------------|----------|--------|
|            | Pequeña  | Grande |
| Baja       | PA       | PB     |
| Alta       | PB       | PA     |

PA: filtro pasa alta

PB: filtro pasa baja

## 2.5. Filtro Haar

Al considerar la transformada wavelet discreta (TWD) Haar de escala unitaria para  $\{y_t\}_{t=1}^T$ , donde  $T$  es par, los coeficientes wavelets y de escala están dados por

$$\begin{aligned}
 W_{t,1} &= \frac{1}{\sqrt{2}}(y_{2t} - y_{2t-1}), & t = 1, 2, 3, 4, \dots, \frac{T}{2} \\
 V_{t,1} &= \frac{1}{\sqrt{2}}(y_{2t} + y_{2t-1}), & t = 1, 2, 3, 4, \dots, \frac{T}{2}
 \end{aligned}
 \tag{2.5.1}$$

Los coeficientes wavelet  $\{W_{t,1}\}$  capturan el comportamiento de  $\{y_t\}$  a través de frecuencias altas de ancho de banda  $[\frac{1}{2}, 1]$ , mientras que los coeficientes de escala  $\{V_{t,1}\}$  capturan el comportamiento en frecuencias bajas con ancho de banda  $[0, \frac{1}{2}]$ . La energía total de  $\{y_t\}$  está dada por la suma de las energías de  $\{W_{t,1}\}$  y  $\{V_{t,1}\}$ . De aquí se puede notar que la energía de los coeficientes de escala dominan a los coeficientes wavelet para procesos de raíz unitaria.

Para este caso, Fan y Gençay proponen el test estadístico:

$$\hat{S}_{T,1} = \frac{\sum_{t=1}^{T/2} V_{t,1}^2}{\sum_{t=1}^{T/2} V_{t,1}^2 + \sum_{t=1}^{T/2} W_{t,1}^2}
 \tag{2.5.2}$$

Las hipótesis para las series de tiempo que se tratarán en la sección siguiente

y con las que se probarán o no la estacionalidad a través de test de raíz unitaria, son:

$$H_0 : \rho = 1 \text{ contra } H_1 : |\rho| < 1 \quad (2.5.3)$$

Dadas dichas hipótesis, bajo  $H_0$   $\hat{S}_{T,1}$  debe acercarse a 1, dado que  $\sum_{t=1}^{T/2} V_{t,1}^2$  domina a  $\sum_{t=1}^{T/2} W_{t,1}^2$ ; mientras que bajo  $H_1$ , debe ser mucho menor que 1.

La distribución asintótica de  $\hat{S}_{T,1}$  bajo  $H_0$ , tiende a

$$T(\hat{S}_{T,1} - 1) \rightarrow -\frac{\gamma_0}{\lambda_v^2 \int_0^1 [W(r)]^2 dr}$$

donde  $\lambda_v^2 = 4\omega^2$ .

En la distribución asintótica nula se encuentran los parámetros  $\gamma_0$  y  $\lambda_v^2$  (o  $\omega^2$ ). Para estimar dichos parámetros se utiliza  $\hat{\gamma}_0 = T^{-1} \sum_{t=1}^T \hat{u}_t^2$  el cual es un estimador consistente para  $\gamma_0$  y  $\hat{\omega}^2 = 4\hat{\gamma}_0 + 2 \sum_{j=1}^q [1 - j/(q+1)] \hat{\gamma}_j$  el cual es un estimador consistente siempre que se estime con un estimador kernel no paramétrico, por ejemplo el kernel Bartlett.

Bajo las condiciones anteriores y si  $\hat{\gamma}_v^2 = 4\hat{\omega}^2$ , se define el test de Fan y Gençay, como:

$$FG_1 = \frac{T\hat{\lambda}_v^2}{\hat{\gamma}_0} [\hat{S}_{T,1} - 1] \quad (2.5.4)$$

El test estadístico  $FG_1$  bajo la hipótesis nula tiene una distribución límite

$$-\frac{1}{\int_0^1 [W(r)]^2 dr} \quad (2.5.5)$$

## 2.6. Filtro Daubechies

Al considerar en forma general un filtro Daubechies de soporte compacto  $\{h_l\}_{l=0}^{L-1}$  de escala unitaria, los coeficientes wavelet y los coeficientes de escala están dados por

$$\begin{aligned}
 W_{t,1} &= \sum_{l=0}^{L-1} h_l y_{2t-l} \\
 V_{t,1} &= \sum_{l=0}^{L-1} g_l y_{2t-l}
 \end{aligned}
 \tag{2.6.1}$$

Donde  $t = L_1, L_1 + 1, \dots, T/2$  con  $L_1 = L/2$ . Al igual que el filtro Haar los coeficientes wavelet extraen la información a partir de frecuencias altas y los coeficientes de escala extraen la información a partir de los coeficientes de escala; sin embargo para cualquier filtro Daubechies los  $W_{t,1}$  son estacionarios bajo las hipótesis 2.5.3 y los  $V_{t,1}$  son no estacionarios bajo  $H_0$  y estacionarios bajo  $H_1$ . Lo anterior implica que bajo el análisis de energía, que los coeficientes de escala dominan a los coeficientes wavelet bajo  $H_0$  y sobre esta característica se formula el test estadístico

$$\hat{S}_{T,1}^L = \frac{\sum_{t=L_1}^{T/2} V_{t,1}^2}{\sum_{t=L_1}^{T/2} V_{t,1}^2 + \sum_{t=L_1}^{T/2} W_{t,1}^2}
 \tag{2.6.2}$$

La distribución asintótica para  $\hat{S}_{T,1}^L$  está dada por:

*i)*  $\hat{S}_{T,1}^L = 1 + o_p(1)$  bajo  $H_0$  y  $\hat{S}_{T,1}^L = c_L + o_p(1)$  bajo  $H_1$

*ii)*  $\left(\frac{T}{2}\right)(\hat{S}_{T,1}^L - 1) \implies -\frac{E(W_{t,1}^2)}{\lambda_v^2 \int_1^0 [W(r)]^2 dr}$  bajo  $H_0$

Con  $c_L = \frac{E(V_{t,1}^2)}{E(V_{t,1}^2) + E(W_{t,1}^2)} < 1$

El test estadístico de Fan y Gençay para un filtro Daubechies, es:

$$FG_1^L = \frac{T}{2} \left( \frac{\hat{\lambda}_v^2}{\hat{v}_{y,1}^2} \right) \left[ \hat{S}_{T,1}^L - 1 \right]
 \tag{2.6.3}$$

El test estadístico  $FG_1^L$  bajo la hipótesis nula tiene la misma distribución límite de  $FG_1$ , ecuación 2.5.5.





## CAPÍTULO 3

---

### Panel de Datos

---

En este capítulo se presentarán los conceptos sobre Panel de datos, sus ventajas y desventajas, los test que determinan el modelo econométrico a utilizar (efectos fijos ó efectos aleatorios) y algunos aspectos metodológicos de la técnica de panel; los cuales dan el fundamento teórico para el análisis del caso de aplicación que se desarrollará en el capítulo 4.

Los modelos usados en el análisis económico se pueden clasificar según los datos utilizados y según las relaciones supuestas entre las variables que intervienen en éstos.

Cuando se realizan estudios económicos, en el análisis de la información pueden existir, entre otras, la dimensión temporal y la dimensión estructural. La primera hace referencia al análisis de series de tiempo, que incorpora información de las variables de estudio en un periodo de tiempo determinado. La segunda representa el análisis de la información para las unidades individuales de estudio restringidas en un momento determinado del tiempo. De las anteriores, interesa obtener conclusiones que se deriven de los modelos estimados y que proporcionen relaciones de causalidad o de comportamiento entre diferentes clases de variables a partir de los datos suministrados.

### 3.1. Panel de datos

**Definición 3.1.1.** *Un panel de datos es un conjunto de datos que combina series temporales con unidades de sección cruzada o de corte transversal (países, ciudades, bancos, regiones, empresas, hogares, etc).*

### 3.2. Modelo general de un panel de datos

La estructura básica para un panel de datos es un modelo de regresión lineal de la forma

$$y_{it} = \mathbf{x}'_{it}\beta + \mathbf{z}'_i\alpha + \varepsilon_{it} \quad (3.2.1)$$

donde  $i = 1, 2, \dots, N$ ;  $t = 1, 2, \dots, T$ . Acá  $i$  es la unidad de estudio (corte transversal),  $t$  se refiere a la dimensión en el tiempo,  $\beta$  es un vector de  $k$  parámetros y  $\mathbf{x}_{it}$  es la  $i$ -ésima observación al momento  $t$  para las  $k$  variables explicativas. La heterogeneidad o efecto individual es  $\mathbf{z}'_i\alpha$ , donde  $\mathbf{z}_i$  está conformada por constantes y un conjunto de individuos o un grupo específico de variables, los cuales pueden ser observables (p.e. género, raza, etc.) o no observables (p.e. características específicas de familias, destrezas, gustos, etc.) todas invariantes en el tiempo  $t$ . Si  $\mathbf{z}_i$  es observable para todos los individuos, entonces el modelo se reduce a un modelo de regresión lineal clásico ([29]). En caso contrario, se tienen:

1. Regresión Total: Si  $\mathbf{z}_i$  contiene solamente términos constantes, el método de mínimos cuadrados ordinarios genera estimadores consistentes y eficientes para  $\alpha$  y el vector de pendientes  $\beta$ .
2. Efectos Fijos: Cuando  $\mathbf{z}_i$  sea no observable y esté correlacionada con  $\mathbf{x}_{it}$ , entonces el estimador de mínimos cuadrados para  $\beta$  será inconsistente. Sin embargo el modelo

$$y_{it} = \mathbf{x}'_{it}\beta + \alpha_i + \epsilon_{it} \quad (3.2.2)$$

donde  $\alpha_i = \mathbf{z}'_i\alpha$ , representa todos los efectos observables. Debe hacerse notar que en el presente se da una pérdida importante de grados de libertad.

3. Efectos Aleatorios: Este modelo considera que los efectos individuales no son independientes entre sí, sino que están distribuidos aleatoriamente alrededor de un valor dado. Una práctica común en el análisis de regresión es asumir que el gran número de factores que afecta el valor de la variable dependiente pero que no han sido excluidas explícitamente como variables independientes del modelo, pueden resumirse apropiadamente en la perturbación aleatoria. El modelo puede ser formulado como:

$$y_{it} = \mathbf{x}'_{it}\beta + E[\mathbf{z}'_i\alpha] + \{\mathbf{z}'_i\alpha - E[\mathbf{z}'_i\alpha]\} + \varepsilon_{it}$$

$$y_{it} = \mathbf{x}'_{it}\beta + \alpha + \mu_i + \varepsilon_{it} \quad (3.2.3)$$

Estos efectos aleatorios se aproxima a especificar que  $\mu_i$  es un elemento aleatorio de un grupo específico, similar a  $\varepsilon_{it}$  excepto que para cada grupo hay una gráfica que representa idénticamente la regresión para cada periodo. El investigador hace inferencia condicional o marginal respecto a una población.

4. Test de Especificación de Hausman: Esta prueba permite determinar qué modelo es el más adecuado para el panel de datos que se está analizando, si es el de efectos fijos o de efectos aleatorios. El test de Hausman se utiliza para analizar la posible correlación entre los  $\alpha_i$  y los regresores. Se basa en la idea que bajo la hipótesis de no correlación, los modelos OLS, LSDV y GLS <sup>1</sup> son consistentes, pero el OLS es ineficiente, mientras que en la hipótesis alternativa, el OLS es consistente, pero el GLS no lo es. Por lo tanto, bajo la hipótesis nula, los dos estimadores difieren sistemáticamente, y el test puede basarse sobre sus diferencias. Otro ingrediente esencial para el test es la matriz de covarianza de el vector diferencia,  $[b - \hat{\beta}]$ :

$$Var[b - \hat{\beta}] = Var[b] + Var[\hat{\beta}] - 2Cov[b, \hat{\beta}] \quad (3.2.4)$$

---

<sup>1</sup>OLS: Mínimos Cuadrados Ordinarios, LSDV: Mínimo Cuadrados de Variable Dummy, y GLS: Minimos Cuadrados Generalizados

## 34 Panel de Datos

---

El resultado esencial de Hausman es que la covarianza de un estimador eficiente y la diferencia del estimador ineficiente, es cero, lo cual implica que

$$\text{Cov}[(b - \hat{\beta}), \hat{\beta}] = \text{Cov}[b, \hat{\beta}] - \text{Var}[\hat{\beta}] = 0$$

o que

$$\text{Cov}[b - \hat{\beta}] = \text{Var}[\hat{\beta}]$$

reemplazando este resultado en (3.2.4) la matriz de covarianza requerida para el test,

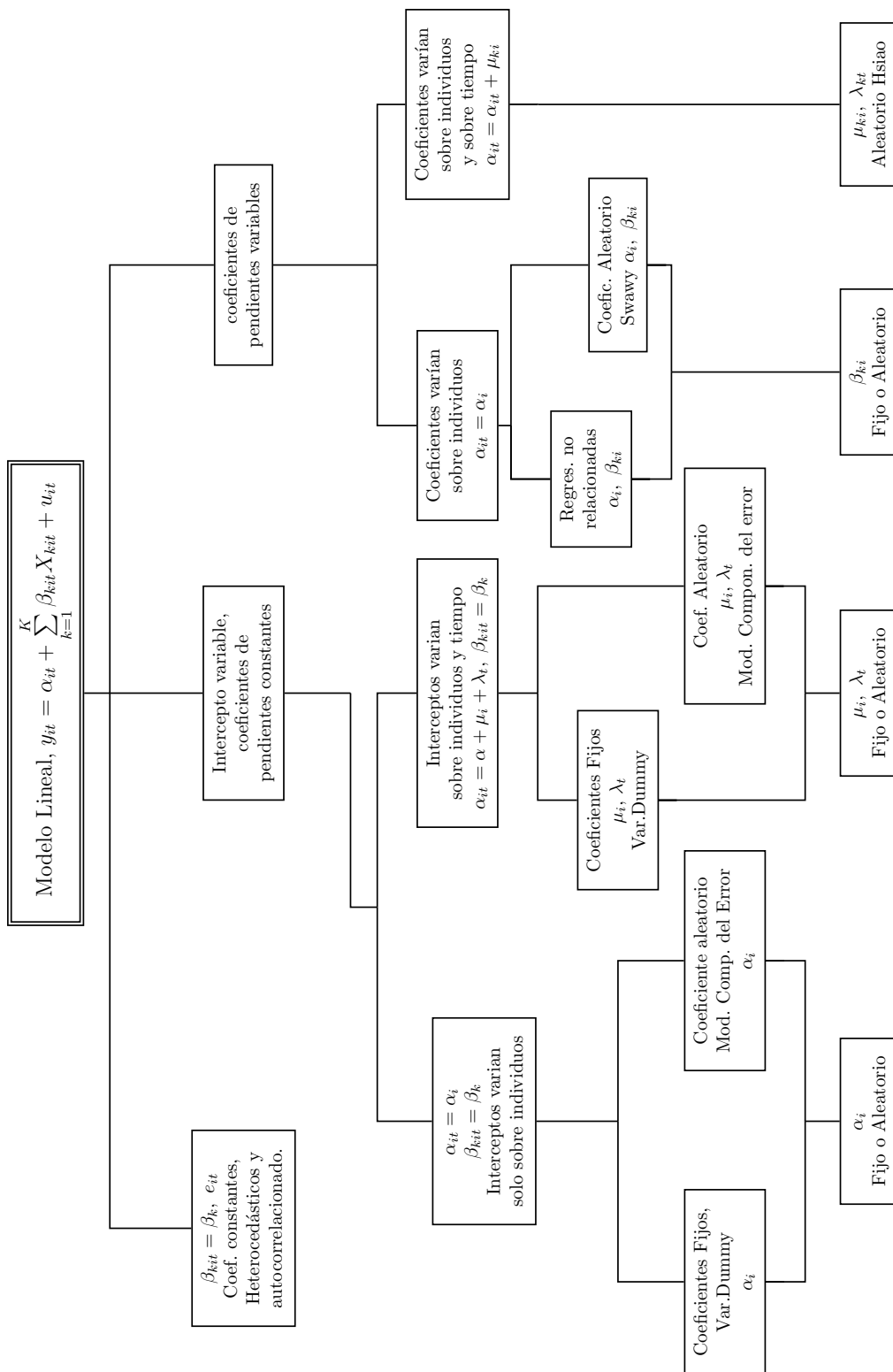
$$\text{Var}[b - \hat{\beta}] = \text{Var}[b] - \text{Var}[\hat{\beta}] = \Psi. \quad (3.2.5)$$

El test  $\chi^2$  se basa en el criterio de Wald:

$$W = \chi^2[K - 1] = [b - \hat{\beta}]' \hat{\Psi}^{-1} [b - \hat{\beta}].$$

Para  $\hat{\Psi}$ , usamos la matriz de covarianza del estimador de pendientes en el modelo LSDV y en el modelo de efecto aleatorio, excluyendo el término constante. Bajo la hipótesis nula,  $W$  tiene una distribución límite  $\chi^2$  con  $K - 1$  grados de libertad.

MODELOS ALTERNATIVOS PARA COMBINAR DATOS DE SERIES DE TIEMPO Y DE CORTE TRANSVERSAL



Tomado y adaptado de: Judge, et. al. (1980), capítulo 13.

### 3.3. Criterios para la selección del modelo

Cuando el investigador quiere hacer inferencia debe decidir si va a trabajar respecto a las características de la población o sobre los efectos que se encuentran en la muestra. Si decide trabajar sobre una muestra aleatoria; es decir, hacer inferencias sobre una población, la estructura apropiada para su análisis es de *tipo aleatorio*. Mientras que si toma una muestra seleccionada a conveniencia, el modelo de *efectos fijos* será el apropiado.

Además, si el objetivo del estudio se centra en los coeficientes de las pendientes de los parámetros y no en las diferencias individuales, se debe elegir un modelo que las elimine y que trabaje la heterogeneidad no observable como aleatoria (incorporándolas en el término de error), lo que modifica la varianza del modelo, mientras que en el modelo de efectos fijos la heterogeneidad no observable se incorpora en la ordenada del modelo.

Otro factor que afecta la selección del modelo radica en el tamaño de las dimensiones, tanto temporal como estructural. Cuando  $t$  es pequeño y  $N$  grande los resultados obtenidos por los dos modelos difieren sustancialmente, además se genera gran cantidad de parámetros de efectos fijos respecto al número de datos disponibles, quienes cuentan con parámetros poco confiables y una estimación ineficiente.

### 3.4. Modelo de efectos fijos

El modelo (3.2.2) se puede escribir, como

$$y_{it} = \mathbf{i}\alpha_i + \mathbf{X}_i\beta + \epsilon_{it},$$

suponiendo que el término  $\alpha_i$  contiene las diferencias entre unidades y debido a ello, dicho parámetro debe ser estimado. En términos matriciales, tenemos:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \mathbf{i} & 0 & \dots & 0 \\ 0 & \mathbf{i} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{i} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} \beta + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

o

$$\mathbf{y} = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \dots \quad \mathbf{d}_n \quad \mathbf{X}] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \epsilon$$

donde  $\mathbf{d}_i$  es una variable dummy<sup>2</sup> que indica la  $i$ -ésima unidad. Reuniendo las  $nT$  filas se obtiene

$$\mathbf{y} = \mathbf{D}\alpha + \mathbf{X}\beta + \epsilon,$$

con  $\mathbf{D} = [d_1 \ d_2 \ \dots \ d_n]_{nT \times n}$ . Este modelo se denomina *mínimos cuadrados de variables ficticias*, MCVF.

Algunos supuestos necesarios para el modelo de efectos fijos, son: Sea

$$\{(y_{i1}, \dots, y_{iT}, x_{i1}, \dots, x_{iT}, \eta_i), i = 1, \dots, N\}$$

una muestra aleatoria y

$$y_{it} = x'_{it}\beta + \eta_i + v_{it}$$

el modelo. Además,

1. *Supuesto Uno:*

$$E(v_t|x_i, \eta_i) = 0 \quad (t = 1, \dots, T),$$

donde  $v_i = (v_{i1}, \dots, v_{iT})'$  y  $x_i = (x_{i1}, \dots, x_{iT})'$ . Tanto  $y_{it}$  como el vector  $k \times 1$  de variables explicativas  $x_{it}$  son observables, mientras  $\eta_i$  es un regresor no observado invariante en el tiempo.

2. *Supuesto Dos:*

$$Var(v_t|x_i, \eta_i) = \sigma^2 I_T.$$

Bajo este supuesto los errores son condicionalmente homocedásticos y no serialmente correlacionados. Bajo el supuesto Uno, tenemos:

$$E(y_i|x_i, \eta_i) = X_i\beta + \eta_i\iota, \tag{3.4.1}$$

donde  $y_i = (y_{i1}, \dots, y_{iT})'$ ,  $\iota$  es un vector  $T \times 1$  de unos y  $X_i = (x_{i1}, \dots, x_{iT})'$  es una matrix  $T \times k$ . La implicación de (3.4.1) para el valor esperado de  $y_i$  dado  $x_i$  es

$$E(y_i|x_i) = X_i\beta + E(\eta_i|x_i)\iota. \tag{3.4.2}$$

---

<sup>2</sup>Una variable dummy, binaria o ficticia es aquella que toma valor de 1 para algunas observaciones indicando la presencia de un efecto sobre miembros de un grupo y 0 para el resto de observaciones.

Sin embargo, bajo el supuesto Dos

$$\text{Var}(y_i|x_i, \eta_i) = \sigma^2 I_T \quad (3.4.3)$$

Lo cual implica que

$$\text{Var}(y_i|x_i) = \sigma^2 I_T + \text{Var}(\eta_i|x_i)\iota\iota' \quad (3.4.4)$$

3. *Supuesto Tres:*

$$E(v_{it}|x_i) = 0 \quad (t = 1, 2, \dots, T).$$

4. *Supuesto Cuatro:*

$$\text{Var}(v_{it}|x_i) = \sigma^2 I_T.$$

Frecuentemente se utiliza  $E(v_{it}|x_i) = 0$  a pesar de ser una suposición débil, sin embargo es conveniente hacerlo pues en las aplicaciones se dificultaría imaginar como  $E(v_{it}|x_i) = 0$  tiende hacia  $E(v_{it}|x_i, \eta_i) = 0$ .

### 3.5. Contraste de significatividad de los efectos de grupo

La razón  $t$  habitual para  $a_i$  implica un contraste de la hipótesis de que  $\alpha_i$  es igual a cero. Pero, normalmente, esta hipótesis no es útil en un contexto de regresión. Si estamos interesados en las diferencias entre grupos, podemos contrastar la hipótesis de que los términos constantes son todos iguales, mediante un contraste  $F$ . Bajo la hipótesis nula, el estimador eficiente coincide con mínimos cuadrados agrupados. La razón  $F$  utilizada para el contraste es

$$F(n-1, nT-n-K) = \frac{(R_u^2 - R_p^2)/(n-1)}{(1 - R_u^2)/(nT-n-K)} \quad (3.5.1)$$

donde  $u$  indica el modelo no restringido y  $p$  indica el modelo agrupado, o restringido, con un único término constante para todos. (Se puede utilizar también la suma de errores al cuadrado, si resulta más conveniente). Si fuese más cómodo, también podría estimarse el modelo con una única constante



y  $n - 1$  variables Dummies. Los demás resultados no cambian, y en vez de estimar  $\alpha_i$ , cada coeficiente de las variables Dummies será una estimación de  $\alpha_i - \alpha_1$ . El contraste  $F$  de que los coeficientes de las  $n - 1$  variables Dummies son cero es idéntico al anterior. Es importante tener presente que, aunque los resultados estadísticos sean los mismos, la interpretación de los coeficientes de las variables Dummies en las dos formulaciones son diferentes.

### 3.6. Los estimadores intra y entre grupos

Podemos formular el modelo de regresión de las siguientes tres formas. Primero, la formulación original es

$$y_{it} = \alpha_i + \beta' x_{it} + \epsilon_{it}. \quad (3.6.1)$$

En términos de desviaciones de las medias del grupo,

$$y_{it} - \bar{y}_{i.} = \beta'(x_{it} - \bar{x}_{i.}) + \epsilon_{it} - \bar{\epsilon}_{i.}, \quad (3.6.2)$$

mientras que en términos de las medias de grupo,

$$\bar{y}_{i.} = \alpha_i + \beta' \bar{x}_{i.} + \bar{\epsilon}_{i.}. \quad (3.6.3)$$

Los tres son modelos de regresión clásica y, en principio, los tres podrían ser estimados, al menos consistentemente, aunque no eficientemente, por mínimos cuadrados ordinarios. Consideremos, entonces, las matrices de sumas de cuadrados y productos cruzados que se utilizarían en cada caso, donde nos centraremos solamente en la estimación de  $\beta$ . En (3.6.1), los momentos serían sobre las medias totales,  $\bar{y}$  y  $\bar{x}$ , y utilizaríamos las sumas totales de cuadrados y productos cruzados,

$$S'_{xx} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})(x_{it} - \bar{x})'$$

y

$$S'_{xy} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})$$

Para (3.6.2), como los datos están ya en desviaciones, las medias de  $(x_{it} - \bar{x}_{i.})$  y  $(y_{it} - \bar{y}_{i.})$  son cero. las matrices de momentos son sumas de

cuadrados y productos cruzados intra-grupos (es decir, desviaciones de las medias de los grupos),

$$S_{xx}^w = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i\cdot})(x_{it} - \bar{x}_{i\cdot})'$$

y

$$S_{xy}^w = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_{i\cdot})(y_{it} - \bar{y}_{i\cdot})'$$

Finalmente, para (3.6.3), las medias de las medias de los grupos es la media total. Las matrices de momentos son las sumas de cuadrados y productos cruzados entre-grupos.

$$S_{xx}^b = \sum_{i=1}^n T(\bar{x}_{i\cdot} - \bar{\bar{x}})(\bar{x}_{i\cdot} - \bar{\bar{x}})'$$

y

$$S_{xy}^b = \sum_{i=1}^n T(\bar{x}_{i\cdot} - \bar{\bar{x}})(\bar{y}_{i\cdot} - \bar{\bar{y}})$$

Es fácil comprobar que

$$S'_{xx} = S_{xx}^w + S_{xx}^b \quad \text{y} \quad S'_{xy} = S_{xy}^w + S_{xy}^b.$$

Hay, por lo tanto, tres posibles estimadores de mínimos cuadrados de  $\beta$ , que corresponden a la descomposición analizada. El estimador de mínimos cuadrados es

$$\mathbf{b}' = [S'_{xx}]^{-1} S'_{xy} = [S_{xx}^w + S_{xx}^b]^{-1} [S_{xy}^w + S_{xy}^b]. \quad (3.6.4)$$

El estimador **intra-grupos** es

$$\mathbf{b}^w = [S_{xx}^w]^{-1} S_{xy}^w.$$

Este es el estimador MCVF. Un estimador alternativo sería el estimador **entre-grupos**,

$$\mathbf{b}^b = [S_{xx}^b]^{-1} S_{xy}^b.$$

Este es el estimador de mínimos cuadrados de (3.6.3) en los  $n$  conjuntos de medias de grupos. De la expresión anterior

$$S_{xy}^w = S_{xx}^w \mathbf{b}^w \quad \text{y} \quad S_{xy}^b = S_{xx}^b \mathbf{b}^b.$$

Insertando estos resultados en (3.6.4), vemos que el estimador de MCO es un promedio ponderado matricialmente, de los estimadores intra y entre grupos:

$$\mathbf{b}^t = \mathbf{F}^w \mathbf{b}^w + \mathbf{F}^b \mathbf{b}^b,$$

donde

$$\mathbf{F}^w = [S_{xx}^w + S_{xx}^b]^{-1} S_{xx}^w = \mathbf{I} - \mathbf{F}^b.$$

### 3.7. Paneles no balanceados y efectos fijos

Los paneles en que los tamaños de grupos difieren son comunes y se conocen como *paneles no balanceados*. Las modificaciones necesarias para permitir tamaños desiguales, son: el tamaño muestral completo es  $\sum_{i=1}^n T_i$  en vez de  $nT$ , y las medias de los grupos deben basarse en  $T_i$ , que varía entre los grupos. Las medias totales para los regresores son

$$\bar{x} = \frac{\sum_{i=1}^n \sum_{t=1}^{T_i} x_{it}}{\sum_{i=1}^n T_i} = \frac{\sum_{i=1}^n T_i \bar{x}_i}{\sum_{i=1}^n T_i} = \sum_{i=1}^n w_i \bar{x}_i,$$

donde  $w_i = T_i / (\sum_{i=1}^n T_i)$ . Si los grupos son de igual tamaño,  $w_i = 1/n$ , la matriz de momentos

$$\mathbf{S}_{xx}^w = \mathbf{X}' \mathbf{M}_d \mathbf{X}$$

es una suma de matrices de sumas de cuadrados y productos cruzados,

$$\sum_{i=1}^n \mathbf{X}'_i \mathbf{M}_i^o \mathbf{X}_i = \sum_{i=1}^n \left( \sum_{t=1}^{T_i} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)$$

sumadas a través de los grupos, denominada la suma de cuadrados **intra-grupos**.

### 3.8. Efectos aleatorios

Dentro algún contexto puede ser más apropiado interpretar los términos constantes específicos de la unidad, como distribuidos aleatoriamente entre las unidades de sección cruzada. Esto es apropiado si creemos que las

## 42 Panel de Datos

---

unidades de sección cruzada de la muestra son extracciones muestrales de una población grande. Retomando (3.2.3), el análisis de familias, se puede interpretar como el conjunto de factores, no incluidos en la regresión, que son específicos en esa familia y además que

$$\begin{aligned}
 E[\epsilon_{it}] &= E[\mu] = 0, \\
 E[\epsilon_{it}^2] &= \sigma_\epsilon^2, \\
 E[u_i^2] &= \sigma_u^2, \\
 E[\epsilon_{it}u_j] &= \sigma_\epsilon^2, \\
 E[\epsilon_{it}\epsilon_{js}] &= 0; \quad \text{si } t \neq s \text{ o } i \neq j, \\
 E[u_iu_j] &= 0; \quad \text{si } i \neq j.
 \end{aligned} \tag{3.8.1}$$

Reescribiendo (3.2.3) en bloques de  $T$  observaciones, tenemos:

$$w_{it} = \epsilon_{it} + u_i \quad \text{y} \quad w_i = [w_{i1}, w_{i2}, \dots, w_{iT}]',$$

el cual se denomina *modelo de componentes del error*. Por consiguiente,

$$\begin{aligned}
 E[w_{it}^2] &= \sigma_\epsilon^2 + \sigma_u^2, \\
 E[w_{it}w_{is}] &= \sigma_u^2, \quad t \neq s.
 \end{aligned}$$

Para las  $T$  observaciones de la unidad  $i$ , sea  $\mathbf{\Omega} = E[w_iw_i']$ . Entonces,

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_\epsilon^2 + \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_\epsilon^2 + \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_\epsilon^2 + \sigma_u^2 \end{bmatrix} = \sigma_\epsilon^2 \mathbf{I} + \sigma_u^2 \mathbf{ii}', \tag{3.8.2}$$

donde  $\mathbf{i}$  es un vector columna  $T \times 1$  de unos. Como las observaciones  $i$  y  $j$  son independientes, la matriz de varianzas y covarianzas de los errores para  $nT$  observaciones, es

$$\mathbf{V} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Omega} \end{bmatrix}$$

## 3.9. Heterocedasticidad

El problema de heterocedasticidad se presenta cuando es violado el supuesto de varianza constante de los errores de la función de regresión. La heterocedasticidad tiene que ver con la relación entre una o más de las variables independientes del modelo y el cuadrado de los errores estimados a partir de la regresión. Este problema se manifiesta en un crecimiento o decrecimiento de la varianza del modelo.

La presencia de heterocedasticidad es muy común en regresiones estimadas a partir de datos de corte transversal. Por ejemplo, cuando se recolectan datos provenientes de estratos, de regiones, por tamaño de la familia o por tipo de empresa. En general, puede presentarse en estudios que incluyen grupos con comportamientos marcados a lo largo de toda la muestra; por ejemplo, la variable ingreso monetario del hogar según el estrato, pues se puede pensar que la varianza del ingreso monetario del grupo de alta riqueza es más alta que la del grupo de escasos recursos.

El problema de heterocedasticidad repercute directamente sobre la estimación de los parámetros de la regresión. Los estimadores seguirán siendo insesgados y consistentes pero no eficientes. La heterocedasticidad causa la subestimación o sobre estimación de la varianza del modelo de regresión, por lo tanto el valor del error estándar de los parámetros, el valor de los estadísticos  $t$  y los intervalos de confianza cambian con respecto a los resultados que deberían obtenerse en ausencia de heterocedasticidad. En este sentido, la presencia de heterocedasticidad en el modelo de regresión hace que las pruebas de hipótesis no tengan validez estadística o que las inferencias sean erróneas.

### 3.9.1. Detección de la heterocedasticidad

A continuación se presentan los métodos para detectar la existencia de heterocedasticidad:

1. **Análisis de residuales:** Este método permite evaluar gráficamente si existe heterocedasticidad causada por una variable independiente en particular o por todo el conjunto de variables independientes. Para el primer caso se elabora un diagrama de dispersión entre  $x_i$

y  $\epsilon_t^2$  (cuadrado del término de error) donde  $x_t$  es el regresor que el investigador supone genera la heterocedasticidad. En el segundo caso, se construye el diagrama de dispersión entre  $y_t$  estimado y  $\epsilon^2$ . Si estas gráficas muestran alguna tendencia específica, puede afirmarse que existe heterocedasticidad en el modelo de regresión. No obstante esta metodología es indicativa y no esta basada en una prueba estadística.

2. **Análisis de regresión:** Es la utilización de una o más regresiones auxiliares. La regresión no se estima entre las variables independientes, sino entre el cuadrado del término de error y el conjunto de regresores del modelo original. Dentro de este método se encuentran las pruebas de Park, White, Glejser, Breusch-Pagan-Godfrey y Golfeld-Quandt.

### 3.10. Autocorrelación

El problema de autocorrelación se presenta en una regresión cuando los errores de las diferentes observaciones están relacionados en el tiempo. Esto indica que el efecto de los errores en el tiempo no es instantáneo sino por el contrario es persistente en el tiempo. La autocorrelación es más común en series ordenadas en el tiempo que en información proveniente de encuestas en un tiempo fijo (sección cruzada). La autocorrelación puede estar relacionada con los ciclos económicos; generalmente ésta se presenta en un modelo con variables macroeconómicas donde en el tiempo ocurre un evidente comportamiento tendencial.

Otra causa de la autocorrelación es la presencia de sesgo de especificación en el modelo; principalmente por omisión de variables importantes, las cuales pasan a formar parte del error de la regresión. La autocorrelación puede ser también generada en casos donde se usa una forma funcional incorrecta del modelo, esto hace que los datos se ajusten a una forma funcional que no es la más adecuada. Se argumenta, que la manipulación de información puede llegar a generar también autocorrelación. Un caso típico se presenta en las cuentas nacionales, donde muchos datos son obtenidos a partir de otros, aplicando técnicas de interpolación o extrapolación. Por ejemplo, cuando se convierten datos diarios a semanales. Finalmente, modelos especiales como los de rezagos distribuidos y los autoregresivos pueden originar autocorrelación. Entre las consecuencias de la autocorrelación se

tiene la sobrestimación o subestimación de los estadísticos  $t$  que juzgan la significancia de las variables independientes en el modelo. Aunque los estimadores siguen siendo insesgados y consistentes son ineficientes. En este sentido se afecta la validez estadística de las pruebas de hipótesis.

### 3.10.1. Detección de la autocorrelación

Los métodos más comunes para detectar autocorrelación son:

1. **Análisis de residuales:** este método plantea la construcción de diagramas de dispersión para los errores en función de tiempo o en función de un período inmediatamente anterior. El primer paso es estimar el modelo original por MCO. Luego los errores estimados de la regresión son graficados en un eje de coordenadas para identificar si existe alguna tendencia de los mismos en el tiempo, o de estos con su primer rezago.
2. **El estadístico de Durbin-Watson:** Esta prueba es válida para aplicar en errores que se modelan como un proceso autoregresivo de orden 1 “AR(1)”, como el mostrado a continuación:

$$\epsilon_t = \rho\epsilon_{t-1} + \nu_t$$

El estadístico  $d$  oscila entre 0 y 4. Si este se aproxima a 0, se dice que existe autocorrelación positiva (relación directa entre los errores), por el contrario si  $d$  se aproxima a 4, existe autocorrelación negativa (relación inversa entre los errores). El Durbin-Watson ( $d$ ) se estima de la siguiente manera:

$$d = \frac{\sum_{t=2}^n (\epsilon_t - \epsilon_{t-1})}{\sum_{t=1}^n \epsilon_t^2} = 2(1 - \hat{\rho}), \quad (3.10.1)$$

donde  $\hat{\rho}$  es el coeficiente de autocorrelación de orden 1, el cual puede despejarse directamente de (3.10.1),

$$\hat{\rho} = 1 - \frac{d}{2}.$$

La hipótesis planteada es entonces

$$H_o : \rho_{\epsilon_t, \epsilon_{t-1}} = 0, \text{ (no existe correlación entre los errores)}$$

$$H_a : \rho_{\epsilon_t, \epsilon_{t-1}} \neq 0, \text{ (existe correlación entre los errores)}.$$

El estadístico Durbin-Watson puede ser comparado con su respectivo tabulado, teniendo en cuenta el número de observaciones contenidas en la muestra y el número de regresores. Se debe tener en cuenta que  $d$  es utilizado para identificar solo autocorrelación de orden 1 siempre y cuando el modelo tenga intercepto. Además no puede usarse en el caso de modelos autoregresivos.

Prueba de Breusch-Godfrey. Esta es una prueba similar a la prueba de White. Se diferencia de ésta en que la variable dependiente de la regresión auxiliar es el término de error  $\epsilon_t$  y los regresores sus respectivos rezagos hasta el orden deseado por el investigador. Adicionalmente son incluidos los regresores usados en el modelo original. La hipótesis nula corresponde a que todos los coeficientes de autocorrelación de orden (los coeficientes que acompañan a los residuos rezagados en la regresión auxiliar) son iguales a cero, mientras la hipótesis alterna es que al menos uno de ellos es distinto de cero.

El estadístico de prueba es  $(n - s)R^2 \sim \chi_s^2$ , donde  $s$  es el número de errores rezagados en la regresión auxiliar. Para probar autocorrelación de orden uno, que es la práctica más común,  $s$  será igual a uno. La hipótesis nula es rechazada cuando  $(n - s)R^2 > \chi_s^2$  a un nivel de significancia  $\alpha$ ; en este caso se concluye que hay autocorrelación ([75], [29]).



## CAPÍTULO 4

---

### Caso de Aplicación

---

Este capítulo se presenta en dos momentos; en el primero, se estudia de manera detallada la demostración del test  $\hat{W}_1$ , planteado por Hong y Kao, publicado en la revista *Econométrica* “Wavelet-based testing for correlation of unknown form in panel models” [32]. Para dicho análisis las demostraciones de las proposiciones, los teoremas del 1 al 6 y el corolario 1, se agrupan de acuerdo a la finalidad de cada una, e.g. convergencia en probabilidad. En el segundo momento se determina la veracidad de la confiabilidad del test de heterocedasticidad consistente  $\hat{W}_1$  a través de la función  $\text{Wavetest}(\text{resid}, N, T, J, W)$  la cual se programó de nuevo en Matlab. Así se contrastan los resultados de [32], [42] y los obtenidos en el presente trabajo.

#### 4.1. Verificación de la Demostración de $\hat{W}_1$

El test propuesto por [32],  $W_1$ , que detecta correlación serial en modelos de panel estáticos o dinámicos, a través de los componentes de error  $\epsilon_{it}$ , es:

$$\hat{W}_1 = \left( \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \hat{\alpha}_{ijk}^2 - \hat{M} \right) / \hat{V}^{\frac{1}{2}} \quad (4.1.1)$$

Donde:

$$\begin{aligned} \hat{\alpha}_{ijk} &\equiv (2\pi)^{-1/2} \sum_{h=1-T_i}^{T_i-1} \hat{R}_i(h) \hat{\Psi}_{jk}^*(h),^1 \\ \hat{R}_i(h) &\equiv T_i^{-1} \sum_{t=|h|+1}^{T_i} \hat{v}_{it} \hat{v}_{it-|h|} \quad (h = 0, \pm 1, \dots, \pm(T_i - 1)).^2 \\ \hat{M} &\equiv \sum_{i=1}^n \hat{R}_i^2(0) M_{i0}, \\ \hat{V} &\equiv \sum_{i=1}^n \hat{R}_i^4(0) V_{i0}, \\ M_{i0} &\equiv \sum_{h=1}^{T_i-1} (1 - h/T_i) b_{J_i}(h, h), \\ V_{i0} &\equiv 4 \sum_{h=1}^{T_i} \sum_{m=1}^{T_i} (1 - h/T_i)(1 - m/T_i) b_{J_i}^2(h, m) \\ a_J(h, m) &\equiv \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \hat{\Psi}_{jk}(h) \hat{\Psi}_{jk}^*(m),^3 \\ b_J(h, m) &\equiv 2Re[a_J(h, m) + a_J(h, -m)],^4 \\ \Psi_{jk}(\omega) &= (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} \hat{\Psi}_{jk}(h) e^{ih\omega},^5 \\ \hat{\Psi}_{ij}(h) &\equiv (2\pi)^{-1/2} \int_{-\pi}^{\pi} \Psi_{jk}(\omega) e^{-ih\omega} d\omega,^6 \\ \hat{\alpha}_{ijk} &\equiv (2\pi)^{-1/2} \sum_{h=1-T_i}^{T_i-1} \hat{\rho}_i(h) \hat{\Psi}_{jk}(h), \text{ donde} \\ \hat{\rho}_i(h) &\equiv \hat{R}_i(h) / \hat{R}_i(0). \end{aligned}$$

El modelo de panel de datos planteado para el trabajo con el test  $W_1$ , [37] es:

$$Y_{it} = \alpha + X'_{it} \beta + \mu_i + \lambda_t + u_{it}, t = 1, 2, \dots, T; i = 1, 2, \dots, n$$

donde  $X_{it}$  puede ser estático o dinámico en la forma de incluir valores de desfase de  $Y_{it}$ ,  $\mu_i$  es un efecto individual y  $\lambda_t$  es el efecto de tiempo común.

---

<sup>1</sup>  $\hat{\alpha}_{ijk}$ , coeficiente empírico wavelet.

<sup>2</sup>  $\hat{R}_i(h)$  función de autocovarianza.

<sup>3</sup>  $a_J$  es un valor real

<sup>4</sup>  $b_J$  es un valores real

<sup>5</sup>  $\Psi_{ij}(\omega)$ , Espectro de un Wavelet.

<sup>6</sup>  $\hat{\Psi}_{ij}(h)$ , transformada de Fourier.

Las hipótesis planteadas son, hipótesis nula

$$\mathbb{H}_0 : cov(u_{it}, u_{it-|h|}) = 0$$

para todo  $h \neq 0$  e  $i$ , e hipótesis alternativa

$$\mathbb{H}_1 : cov(u_{it}, u_{it-|h|}) \neq 0$$

para todo  $h \neq 0$  y para algunos  $i$ . La prueba estadística se construye utilizando la función de densidad espectral en la que el supuesto  $h$  es conocido bajo la hipótesis alternativa relajada.

La prueba de  $\mathbb{H}_0$  se realiza con el residual estimado degradado

$$\hat{v} = \hat{u}_{it} - \bar{u}_i - \bar{u}_{.t} + \bar{u} \quad (t = 1, 2, \dots, T; i = 1, 2, \dots, n)$$

donde

$$\hat{u}_{it} = Y_{it} - X'_{it}\hat{\beta}$$

$$\bar{u}_i = T_i^{-1} \sum_{t=1}^{T_i} \hat{u}_{it}$$

$$\bar{u}_{.t} = n^{-1} \sum_{i=1}^n \hat{u}_{it}$$

$$\bar{u} = (nT_i)^{-1} \sum_{i=1}^n \sum_{t=1}^{T_i} \hat{u}_{it}$$

y  $\hat{\beta}$  estimadores consistentes bajo la hipótesis nula  $\mathbb{H}_0$ .

En lugar de utilizar la función de autocovarianza dada por  $R_i(h) = E(\nu_{it}, \nu_{it-|h|})$ , Hong y Kao utilizan el espectro de potencia  $f_i(\omega) = (2\pi)^{-1} \sum_{h=-\infty}^{+\infty} R_i(h)e^{-ih\omega}$ ,  $\omega \in [-\pi, \pi]$  para construir la prueba estadística, ya que puede contener información sobre la correlación serial en todos los rezagos.

También, en lugar de emplear la representación de Fourier de la densidad espectral, manejan la densidad espectral basado en wavelets  $\Psi_{jk}(\omega)$ , usando la base wavelet mencionada  $\psi \in L^2(\mathbf{R})$ ,  $\Psi_{jk}(\omega)$  definida como:

$$\Psi_{jk}(\omega) = (2\pi)^{-1/2} \sum_{m=-\infty}^{+\infty} \psi_{jk} \left( \frac{\omega}{2\pi} + m \right), \quad \omega \in [-\pi, \pi]$$

donde,  $\Psi_{jk}(\omega)$  capta eficazmente los picos locales y los picos de densidad espectral cambiando el efecto índice de tiempo  $k$ .

Sobre la base de los coeficientes wavelet empíricos  $\hat{\alpha}_{ijk} \equiv (2\pi)^{-1/2} \sum_{h=1-T_i}^{T_i-1} \hat{R}_i(h) \hat{\Psi}_{jk}^*(h)$ , la prueba estadística de heterocedasticidad coherente  $\hat{W}_1$  y la prueba estadística de heterocedasticidad con correlación  $\hat{W}_2$ , así como su distribución de probabilidad, bajo la hipótesis nula  $H_0$ , descrita:

$$\hat{W}_1 = \frac{\sum_{i=1}^n \left[ 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^J} \hat{\alpha}_{ijk}^2 - \hat{R}_i^2(0)(2^{J_i+1}-1) \right]}{2 \left[ \sum_{i=1}^n \hat{R}_i^4(0)(2^{J_i+1}-1) \right]^{1/2}} \xrightarrow{d} N(0, 1)$$

$$\hat{W}_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\left[ 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^J} \hat{\alpha}_{ijk}^2 - (2^{J_i+1}-1) \right]}{2(2^{J_i+1}-1)^{1/2}} \xrightarrow{d} N(0, 1)$$

A continuación se plantean las respectivas demostraciones.

#### 4.1.1. Acotamiento del factor de traslación y los rezagos

**Lema 4.1.1.** *Supongamos que los supuestos 1 y 2 (Pág. 1524-1525, [32]) se cumplen, y sea  $b_{J_i}(h, m)$  como está en  $\hat{W}_2$ . Entonces para cualquier  $J_i, T_i \in \mathbb{Z}^+$  y  $C$  una cota constante que es independiente de  $i, T_i$  y  $J_i$  se tiene:*

- (i)  $b_{J_i}(h, m)$  es un valor real,  $b_{J_i}(0, m) = b_{J_i}(h, 0) = 0$  y  $b_{J_i}(h, m) = b_{J_i}(m, h)$ ;
- (ii)  $\sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} h^v |b_{J_i}(h, m)| \leq C 2^{(1+v)(J_i+1)}$  para  $0 \leq v \leq \frac{1}{2}$ ;
- (iii)  $\sum_{h=1}^{T_i-1} \left[ \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 \leq C 2^{(J_i+1)}$ ;
- (iv)  $\sum_{h=1}^{T_i-1} \sum_{h_2=1}^{T_i-1} \left[ \sum_{m=1}^{T_i-1} |b_{J_i}(h_1, m) b_{J_i}(h_2, m)| \right]^2 \leq C (J_i + 1) 2^{(J_i+1)}$ ;
- (v)  $\left| \sum_{h=1}^{T_i-1} b_{J_i}(h, h) - (2^{J_i+1} - 1) \right| \leq C [(J_i + 1) + 2^{(J_i+1)}(2^{J_i+1}/T_i)^{(2\tau-1)}]$ ; con  $\tau$  como en la Suposición 2 (Pág. 1525, [32]);
- (vi)  $\left| \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}^2(h, m) - 2(2^{J_i+1} - 1) \right| \leq C [(J_i + 1)^2 + 2^{J_i+1}(2^{J_i+1}/T_i)^{(2\tau-1)}]$
- (vii)  $\sup_{1 \leq h, m \leq T_i-1} |b_{J_i}(h, m)| \leq C (J_i + 1)$ ;
- (viii)  $\sup_{1 \leq h \leq T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \leq C (J_i + 1)$

*Demostración del Lema (4.1.1).* Este lema se amplía en Lee y Hong (2001, 4.1.1), quienes consideran el caso tanto para  $J_i \equiv J \rightarrow \infty$  como para  $T_i \equiv T \rightarrow \infty$ . Para detalles de la demostración, ver Hong y Kao (2002).  $\square$

### 4.1.2. Convergencia en Probabilidad

Las siguientes demostraciones muestran que efectivamente los coeficientes wavelets y la varianza convergen en probabilidad 0 y la razón entre las varianzas converge en probabilidad 1.

**Teorema 4.1.2.** *Sea  $\hat{\alpha}_{ijk}$  y  $\bar{\alpha}_{ijk}$  definidas en la ecuación (4.1.1) y en la ecuación (4.1.15), y  $V_{nT} \equiv \sum_{i=1}^n \sigma_i^8 V_{i0}$ , donde  $V_{i0}$  se asume como en el segundo test estadístico  $\hat{W}_2$ , entonces  $V_{nT}^{-1/2} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk}^2 - \bar{\alpha}_{ijk}^2) \xrightarrow{p} 0$ .*

*Demostración del Teorema (4.1.2).* Para demostrar el teorema basta con verificar que  $\hat{\alpha}_{ijk}^2 - \bar{\alpha}_{ijk}^2 = (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})^2 + 2(\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})\bar{\alpha}_{ijk}$ .  $\square$

**Teorema 4.1.3.** *Sean  $\hat{M}$  y  $\hat{V}$  como están definidos en el segundo test estadístico  $\hat{W}_2$ . Entonces  $V_{nT}^{-1/2}(\hat{M} - M_{nT}) \xrightarrow{p} 0$  y  $\hat{V}/V_{nT} \xrightarrow{p} 1$*

*Demostración del Teorema (4.1.3).* Recordemos la definición de  $\hat{M}$  y  $\hat{V}$  en el primer test estadístico  $\hat{W}_1$ . Siguiendo un razonamiento análogo a la demostración del teorema 3 (4.1.3), podemos obtener  $\hat{M} = M_{nT}[1 + o_p(1)]$  y  $\hat{V} = V_{nT}[1 + o_p(1)]$ , de lo cual se obtiene  $(n_A T)^{-1} \hat{V}^{-1/2} \hat{W}_1 = (n_A T)^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \hat{\alpha}_{ijk}^2 + O_p(1)$  dado que  $M_{nT} \leq C \sum_{i=1}^n (2^{J_i+1}) = O(V_{nT})$ , y  $V_{nT}/n_A T \rightarrow 0$  por  $(n_A T)^{-1} \sum_{i=1}^n 2^{J_i+1} \rightarrow 0$ . Recordemos que se debe demostrar:

(a)  $(n_A T)^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk}^2 - \alpha_{ijk}^2) \xrightarrow{p} 0$

(b)  $n_A^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \alpha_{ijk}^2 = (n_A T)^{-1} \sum_{i=1}^n 2\pi c_i Q(f_i, f_{i0}) + o(1)$ ,  
donde  $\alpha_{ijk}^2$  está definida en la suposición 2 (Pág. 1525,[32]) teniendo como  $\alpha_{ijk} \equiv \int_{-\pi}^{\pi} f_i(\omega) \Psi_{jk}(\omega) d\omega$

Primero demostramos (a), puesto que

$$\begin{aligned}
 & (n_AT)^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk}^2 - \alpha_{ijk}^2) = \\
 & = (n_AT)^{-1} \sum_{i=1}^n 2\pi (n_AT)^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} [(\hat{\alpha}_{ijk} - \alpha_{ijk})^2 + 2(\hat{\alpha}_{ijk} - \alpha_{ijk})\alpha_{ijk}]
 \end{aligned} \tag{4.1.2}$$

Es suficiente demostrar que el primer término de (4.1.2) desaparece en probabilidad. El segundo término de (4.1.2) desaparece en probabilidad, entonces por la desigualdad de Cauchy-Schwarz y el hecho que  $(n_AT)^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \alpha_{ijk}^2 \leq C \sup_{i \in \mathbb{N}_A} Q(f_i, f_0) \leq C^2$ . Note que si  $\hat{\alpha}_{ijk} - \alpha_{ijk} = (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk}) + (\bar{\alpha}_{ijk} - \alpha_{ijk})$ , obtenemos:

$$\begin{aligned}
 \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \alpha_{ijk})^2 & \leq 2 \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} [(\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk}) + (\bar{\alpha}_{ijk} - \alpha_{ijk})^2] \\
 & \equiv 2(\hat{M}_{71} + \hat{M}_{72})
 \end{aligned} \tag{4.1.3}$$

Siguiendo la demostración análoga a la proposición (4.1.14), podemos obtener

$$(n_AT)^{-1} \hat{M}_{71} = O_p[(n_AT)^{-1} + (n_AT)^{-1} V_{nT}] \tag{4.1.4}$$

bajo las suposiciones 1 a 6 (Pág. 1524-1531, [32]) y  $\mathbb{H}_A$ . Nótese que hemos obtenido una razón más lenta bajo  $\mathbb{H}_A$  que bajo  $\mathbb{H}_0$ . Para el segundo término en (4.1.3), además podemos descomponer a  $\hat{M}_{72}$  en

$$\hat{M}_{72} \leq 2 \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} [(\bar{\alpha}_{ijk} - E\bar{\alpha}_{ijk})^2 + (E\bar{\alpha}_{ijk} - \alpha_{ijk})^2] \equiv 2(\hat{M}_{721} + \hat{M}_{722}) \tag{4.1.5}$$

Consideramos el primer término en (4.1.5). Tenemos que el  $\sup_{1 \leq h \leq T_i-1} \text{var}[\bar{R}_i(h)] \leq CT_i^{-1}$ , lo cual sigue de la afirmación

$$\text{var}[\bar{R}_i(h)] = T_i^{-1} \sum_{l=1-T_i}^{T_i-1} (1 + |l|T_i)[R_i^2(l) + R_i(l-h)R_i(l+h) + \kappa_i(h, l, l+h)]$$

Cf Hannan (1970, pág. 209). Por lo tanto, nosotros tenemos

$$\hat{M}_{721} \leq \sum_{i=1}^n T_i \sup_{1 \leq h \leq T_i-1} \text{var}[\bar{R}_i(h)] \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| = O(V_{nT})$$

Para el segundo término en (4.1.5), nótese que  $|E\bar{\alpha}_{ijk} - \alpha_{ijk}| \leq (2\pi)^{-1/2} T_i^{-1} \sum_{h=-\infty}^{\infty} |hR_i(h)\hat{\Psi}_{jk}(h)|$  y así tenemos

$$\begin{aligned} \hat{M}_{722} &\leq \sum_{i=1}^n \sum_{j=1}^{J_i} \sum_{k=1}^{2^j} \left[ \sum_{h=1-T_i}^{T_i-1} R_i^2(h) \right] \left[ \sum_{h=-\infty}^{\infty} h^2 |\hat{\Psi}_{jk}^2(h)| \right] \\ &= O \left[ (2^{2\bar{J}}/T) \sum_{i=1}^n (2^{J_i+1}) \right] = o(V_{nT}) \end{aligned}$$

dada la suposición 2 (Pág. 1525,[32]) y  $2^{2\bar{J}}/T \rightarrow 0$ . Siguiendo la desigualdad de Markov  $(n_A T)^{-1} \hat{M}_{72} = O_P[(n_A)^{-1} V_{nT}]$ . Así, (4.1.3), (4.1.4) y  $V_{nT}/(n_A T) \rightarrow 0$  implica (a).

Ahora demostraremos a (b). Continuamos con

$$\begin{aligned} (n_A T)^{-1} \sum_{i=1}^n 2\pi \sum_{j=0}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk} - (n_A T)^{-1} \sum_{i=1}^n 2\pi \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \alpha_{ijk} &\leq \\ &\leq C \sup_{i \in \mathbb{N}_A} \sum_{j=J_i+1}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk}^2 \rightarrow 0 \end{aligned}$$

Como el  $\min_{1 \leq i \leq n} (J_i) \rightarrow \infty$  y  $Q_i(f_i, f_0) = \sum_{j=0}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk}^2 \leq C$ . Así se completa la demostración del test  $\hat{W}_1$ .  $\square$

**Proposición 4.1.4.**

$$V_{nT}^{-1/2} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk}) \bar{\alpha}_{ijk} \xrightarrow{p} 0$$

*Demostración de la proposición 4.1.4.* Recordando

$$\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk} = (2\pi)^{(-1/2)} \sum_{c=1}^9 \sum_{h=1-T_i}^{T_i-1} \hat{\xi}_{ci}(h) \hat{\Psi}_{jk}(h)$$

Podemos escribir

$$\begin{aligned} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk}) \bar{\alpha}_{ijk} &= \sum_{c=1}^9 \left\{ \sum_{i=1}^n T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \left[ \sum_{h=1-T_i}^{T_i-1} \hat{\xi}_{ci}(h) \hat{\Psi}_{jk}(h) \right] \bar{\alpha}_{ijk} \right\} \\ &\equiv \sum_{c=1}^9 \hat{\delta}_c \end{aligned} \quad (4.1.6)$$

Mostraremos  $V_{nT}^{-1/2} \hat{\delta} \xrightarrow{d} 0$  para  $1 \leq c \leq 9$ . Primero, hacemos

$$\begin{aligned} V_{nT}^{-1/2} | \hat{\delta}_1 + \hat{\delta}_8 + \hat{\delta}_9 | &\leq (\hat{A}_1 + \hat{A}_8 + \hat{A}_9)^{1/2} \left( \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk} \right)^{1/2} \\ &= O_p[n^{-3/4} V_{nT}^{1/4} + (V_n T / nT)^{1/2}] \end{aligned} \quad (4.1.7)$$

donde  $V_{nT}^{-1} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk}^2 = O_p(1)$  por Lema (4.1.1)(v) y

$$E \bar{\alpha}_{ijk}^2 \leq C T_i^{-1} \sum_{h=1-T_1}^{T_i-1} | \hat{\Psi}_{jk} 2\pi h |^2.$$

A continuación, consideramos el segundo termino  $\hat{\delta}_2$  en (4.1.6). Escribiremos



$$\begin{aligned}\hat{\delta}_2 &= (\hat{\beta} - \beta)' \sum_{i=1}^n 2\pi \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \left[ \sum_{h=1-T_1}^{T_i-1} \Gamma_{ixv}(h) \hat{\Psi}_{jk}(h) + \sum_{h=1-T_1}^{T_i-1} [\tilde{\Gamma}_{ixv} - \Gamma_{ixv}(h)] \hat{\Psi}_{jk}(h) \right] \tilde{\alpha}_{ijk} \\ &\equiv (\hat{\beta} - \beta)' \hat{M}_3 + (\hat{\beta} - \beta)' \hat{M}_4,\end{aligned}\quad (4.1.8)$$

Para el primer término  $\hat{M}_3$  señalando que  $\{\tilde{\alpha}_{ijk}\}$  es una secuencia independiente a través de  $i$  con media cero, obtenemos

$$\begin{aligned}E\hat{M}_3^2 &= \sum_{i=1}^n T_i^2 E \left\| \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \Gamma_{ixv}(h) \bar{R}_i(m) \right\|^2 \\ &\leq \sum_{i=1}^n \sigma_i^4 T_i \left[ \sum_{h=1}^{T_i-1} \|\Gamma_{ixv}(h)\| \right]^2 \sup_{1 \leq h \leq T_i-1} \left[ \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \right]^2 \\ &= O \left[ T \sum_{i=1}^n (J_i + 1)^2 \right] = O(TV_{nT})\end{aligned}\quad (4.1.9)$$

dado la suposición 5 (Pág. 1528, [32]), Lema (4.1.1, viii), y  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ .

Resulta que  $V_{nT}^{-1/2} (\hat{\beta} - \beta)' \hat{M}_3 = O_p(n^{-1/2})$  por la inecuación de Chebyshev. Para el segundo termino  $\hat{M}_4$  en (4.1.8), tenemos

$$V_{nT}^{-1/2} | (\hat{\beta} - \beta)' \hat{M}_4 | \leq V_{nT}^{-1/2} \| (\hat{\beta} - \beta) \| \hat{M}_2^{1/2} \left( \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \tilde{\alpha}_{ijk}^2 \right)^{1/2} = O_p[(nT)^{-1} V_{nT}^{1/2}]$$

donde  $\hat{M}_2 = O_p[(nT)^{-1} V_{nT}]$  es conocido en (4.1.20). Se observa (4.1.8) que  $V_{nT}^{-1/2} \hat{\delta}_2 = O_p(n^{-1/2} + (nT)^{-1} V_{nT}^{1/2})$ . Del mismo modo tenemos que  $V_{nT}^{-1/2} \hat{\delta}_3 = O_p(n^{-1/2} + (nT)^{-1} V_{nT}^{1/2})$ .

Ahora consideramos  $\hat{\alpha}_4$  en (4.1.6).  $\hat{\alpha}_4 = \sum_{i=1}^n T_i \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \hat{\xi}_{4i}(h) \bar{R}_i(m)$ .

Usando el Lema (4.1.1, ii) y  $V_{nT} \leq C \sum_{i=1}^n 2J_i + 1$ , podemos obtener

$$\begin{aligned}
 E\hat{\delta}_4^2 &= \sum_{i=1}^n \sum_{l=1}^n T_i T_l \sum_{h_1=1}^{T_i-1} \sum_{h_2=1}^{T_l-1} \sum_{m_1=1}^{T_i-1} \sum_{m_2=1}^{T_l-1} b_{J_i}(h_1, m_1) b_{J_l}(h_2, m_2) \\
 &\quad \times E[\hat{\xi}_{4i}(h_1) \bar{R}_i(m_1) \hat{\xi}_{4l}(h_2) \bar{R}_l(m_2)] \\
 &\leq CT^{-1} \sum_{i=1}^n \left[ \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 + CT^{-2} \left[ \sum_{i=1}^n \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 \\
 &\leq C(2^{\bar{J}}/T) V_{nT} + CT^{-2} V_{nT}^2 \tag{4.1.10}
 \end{aligned}$$

donde la primera desigualdad se desprende de:

- a)  $|E[\hat{\xi}_{4i}(h_1) \bar{R}_i(m_1) \hat{\xi}_{4i}(h_2) \bar{R}_i(m_2)]| \leq CT_i^{-3/2} T_l^{-3/2}$ ,
- b) para  $i \neq l$ ,  $|E[\hat{\xi}_{4i}(h_1) \bar{R}_i(m_1) \hat{\xi}_{4l}(h_2) \bar{R}_l(m_2)]| \leq CT_i^{-2} T_l^{-2}$ , puede ser demostrado mediante la expansión de los datos bajo  $\mathbb{H}_0$ , donde los  $\{v_{it}\}$  coinciden con  $\{\varepsilon_{it}\}$ , y como también están i.i.d. en  $E(v_{it}^8) \leq C$  para cada  $i$ , y  $\{v_{it}\}$  y  $\{v_{lt}\}$  son mutuamente independiente para  $i \neq l$ ; por lo tanto, tenemos  $V_{nT}^{-1/2} \hat{\delta}_4 = O_p(2^{\bar{J}/2}/T^{1/2} + V_{nT}^{1/2}/T)$  por la inecuación de Chebyshev. Similarmemente podemos obtener  $V_{nT}^{-1/2} \hat{\delta}_5 = O_p(2^{\bar{J}_2}/T^{1/2} + V_{nT}^{1/2}/T)$ .

Próximamente, consideramos  $\hat{\delta}_6$ . Escribiendo

$$\hat{\delta}_6 = \sum_{i=1}^n \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \hat{\xi}_{6i}(h) \bar{R}_i(m), \text{ donde } \hat{\xi}_{6i}(h) = T_i^{-1} \sum_{t=h+1}^{T_i} v_{it} \bar{v}_{t-h},$$

como en (4.1.16). Usando el Lemma (4.1.1)(ii) y  $v_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ , podemos obtener

$$\begin{aligned}
 E\hat{\delta}_6^2 &= \sum_{i=1}^n \sum_{l=1}^n T_i T_l \sum_{h_1=1}^{T_i-1} \sum_{h_2=1}^{T_l-1} \sum_{m_1=1}^{T_i-1} \sum_{m_2=1}^{T_l-1} b_{J_i}(h_1, m_1) b_{J_l}(h_2, m_2) \\
 &\quad E[\hat{\xi}_{6i}(h_1) \bar{R}_i(m_1) \hat{\xi}_{6l}(h_2) \bar{R}_l(m_2)] \\
 &\leq Cn^{-1} \sum_{i=1}^n \left[ \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 + Cn^{-2} \left[ \sum_{i=1}^n \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 \\
 &\leq C(2^{\bar{J}}/T) V_{nT} + Cn^{-2} V_{nT}^2 \tag{4.1.11}
 \end{aligned}$$

para la primera inecuación usaremos el hecho que

a)  $|E[\hat{\xi}_{6i}(h_1)\bar{R}_i(m_1)\hat{\xi}_{6i}(h_2)\bar{R}_i(m_2)]| \leq CT^{-2}n^{-1}$ ; y

b) para  $i \neq l$ ,  $|E[\hat{\xi}_{6i}(h_1)\bar{R}_i(m_1)\hat{\xi}_{6l}(h_2)\bar{R}_l(m_2)]| \leq CT_i^{-1}T_l^{-2}n^{-2}$ ,

pueden ser demostrado mediante la expansión de los  $\{v_{it}\}$  los cuales tienen la propiedad que están idénticamente distribuidos y la independencia entre  $\{v_{it}\}$  y  $\{v_{lt}\}$  para  $i \neq l$  bajo  $H_0$ . Se sigue por la inecuación de Chebyshev y  $2^J/n \rightarrow 0$  de modo que  $V_{nT}^{-1/2}\hat{\delta}_6 = O_p(2^{J_2}/T^{1/2} + V_{nT}^{1/2}/n) \xrightarrow{p} 0$ . Similarmente hacemos  $V_{nT}^{-1/2}\hat{\delta}_7 = O_p(2^{J_2}/T^{1/2} + V_{nT}^{1/2}/n) \xrightarrow{p} 0$ . Hemos demostrado que  $V_{nT}^{-1/2}\hat{\delta}_c \xrightarrow{p} 0$  para  $1 \leq c \leq 9$  dado  $\max_{1 \leq i \leq n} 2^{2(J_i+1)}/(n^2 + T) \rightarrow 0$ . La Proposición 4.1.4 depende de la ecuación (4.1.6).  $\square$

### 4.1.3. Convergencia en Distribución Normal

**Teorema 4.1.5.** Tomando  $M_{nT} \equiv \sum_{i=1}^n \sigma_i^4 M_{i0}$ ,  $M_{i0}$ , como en el primer test estadístico  $\hat{W}_1$ . Entonces  $V_{nT}^{-1/2}(\sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk}^2 - M_{nT}) \xrightarrow{d} N(0, 1)$ .

*Demostración del Teorema (4.1.5).* Recordando la definición de  $\bar{\alpha}_{ijk}$  en la ecuación(4.1.15) de la suposición 8 (Pág. 1534,[32]), podemos escribir

$$\begin{aligned} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{2^j} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk} &= \sum_{i=1}^n T_i \sum_{h=1}^{T-1} \sum_{m=1}^{T-1} b_{J_i}(h, m) \bar{R}_i(h) \bar{R}_i(m) \\ &\equiv \sum_{i=1}^n (\hat{A}_i + \hat{B}_{1i} - \hat{B}_{2i} - \hat{B}_{3i}), \end{aligned}$$

donde

$$\begin{aligned}\hat{A}_i &\equiv 2T_i^{-1} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h,m) \sum_{t=2}^{T_i} \sum_{s=1}^{t-1} v_{it}v_{it-j}v_{is}v_{is-m} \quad (\text{por simetría de } b_{J_i}(\cdot, \cdot)), \\ \hat{B}_{1i} &\equiv T_i^{-1} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h,m) \sum_{t=1}^{T_i} v_{it}^2 v_{it-h} v_{it-m}, \\ \hat{B}_{2i} &\equiv T_i^{-1} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h,m) \sum_{t=1}^h \sum_{s=m+1}^{T_i} v_{it}v_{it-h}v_{is}v_{is-m}, \\ \hat{B}_{3i} &\equiv T_i^{-1} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h,m) \sum_{t=1}^{T_i} \sum_{s=1}^m v_{it}v_{it-k}v_{is}v_{is-m}\end{aligned}$$

Note que bajo  $\mathbb{H}_0$ ,  $\{v_{it}\}$  coincide con  $\{\varepsilon_{it}\}$ , y son i.i.d. para cada  $i$ ,  $\{v_{it}\}$  y  $\{v_{ls}\}$  son independientes  $i \neq l$  y todo  $t, s$ .  $\square$

**Proposición 4.1.6.** *Bajo las condiciones de la proposición (4.1.10), se tiene que  $V_{nT}^{-1/2} \sum_{i=1}^n \hat{U}_i \xrightarrow{d} N(0, 1)$*

*Demostración de la proposición (4.1.6).* Escribimos

$$\begin{aligned}\hat{U}_i &= T_i^{-1} \sum_{t=q_i+2}^{T_i} U_{it}, \text{ donde} \\ U_{it} &\equiv 2v_{it} \sum_{h=i}^{q_i} v_{it-h} H_{i,t-q_i-1}(h), \\ H_{i,t-q_i-1}(h) &\equiv \sum_{m=i}^{q_i} b_{J_i}(h, m) S_{i,t-q_i-1}(m) \text{ y} \\ S_{i,t-q_i-1}(m) &\equiv \sum_{s=i}^{i,t-q_i-1} v_{is} v_{is-m}.\end{aligned}$$

Entonces  $\hat{U} = \sum_{t=q_0+2}^{\bar{T}} U_t$ , donde  $\bar{T} \equiv \max_{1 \leq i \leq n} (T_i)$ ,  $U_t \equiv \sum_{i=1}^n U_{it} \mathbf{1}(q_i \leq t \leq T_i)$ , y  $\mathbf{1}(\cdot)$  es la función indicador.

Se coloca  $\mathcal{F}_t = \bigotimes_{i=1}^n \mathcal{F}_{it}$ , donde  $\mathcal{F}_{it}$  es el sigma campo generado por  $\{v_{is}, s \leq t\}_{i=1}^n$ . Puesto que  $\{v_{it}v_{it-m}\}$  es independiente de  $H_{i,t-q_i-1}(h)$  para  $0 \leq h \leq q_i$ ,  $\{U_t, \mathcal{F}_{t-1}\}$ , es una secuencia de diferencias de martingalas adaptadas (m.d.s), con

$$E\hat{U}^2 = \sum_{t=q_0+2}^{\bar{T}} EU_t^2 = \sum_{t=q_0+2}^{\bar{T}} \sum_{i=1}^n E(U_{it}^2) \mathbf{1}(q_i \leq t \leq T_i) = V_{nT}[1 + o(1)]$$

dado que  $q_i \rightarrow \infty$ ,  $q_i/2^{J_i} \rightarrow \infty$  y  $\bar{q}^2/T \rightarrow 0$ . Aplicamos el Teorema Central del Límite para martingalas de Brown (1971) y así verificar las siguientes condiciones:

$$(a) \text{ var}^{-2}(\hat{U}) \sum_{t=q_0+2}^{\tilde{T}} E\{U_t^2 \mathbf{1}[|U_t| \geq \epsilon \text{var}^{1/2}(\hat{U})]\} \rightarrow 0 \text{ para todo } \epsilon > 0$$

$$(b) \text{ var}^{-2}(\hat{U}^2) T^{-2} \sum_{t=q_0+2}^{\tilde{T}} E(U_t^2 | \mathcal{F}_{t-1}) \xrightarrow{p} \mathbf{1}.$$

Primero verificaremos (a) demostrando que  $V_{nT}^{-2} \sum_{t=q_0+2}^{\tilde{T}} EU_t^4 \rightarrow 0$ . Dado  $t$ ,  $\{U_{it}\}$  es una sucesión con media cero independiente de  $i$ , así tenemos que  $EU_t^4 \leq C[\sum_{i=1}^n T_i^{-2} (EU_{it}^4)^{1/2} \mathbf{1}(q_i \leq t \leq T_i)]^2$ . Sin embargo, siguiendo a Lee y Hong (2001, demostración del teorema 1 (4.1.2)), podemos obtener que para cada  $i$  y para  $T_i$  suficientemente grande,  $EU_{it}^4 \leq Ct^2 \sum_{h=1}^{q_i} \sum_{m=1}^{q_i} b_{J_i}^2(h, m)$ . Lo que conlleva a  $V_{nT}^{-2} \sum_{t=1}^{\tilde{T}} EU_t^4 = O(T^{-1}) \rightarrow 0$ . Por tanto la condición (a) se cumple.

Ahora verificaremos la condición (b) demostrando que  $V_{nT}^{-2} E(\tilde{U}^2 - E\hat{U}^2) \rightarrow 0$ , donde

$$\tilde{U}^2 \equiv \sum_{t=q_0+2}^{\tilde{T}} E(U_t^2 | \mathcal{F}_{t-1}),$$

$$E(U_t^2 | \mathcal{F}_{t-1}) \equiv E \left\{ \left[ T_i^{-1} \sum_{i=1}^n U_{it} \mathbf{1}(q_i \leq t \leq T_i) \right]^2 \middle| \mathcal{F}_{t-1} \right\} \quad (4.1.12)$$

$$= \sum_{i=1}^n T_i^{-1} E(U_{it}^2 | \mathcal{F}_{it-1}) \mathbf{1}(q_i \leq t \leq T_i) \quad (4.1.13)$$

Donde la segunda igualdad sigue del hecho que para cada  $t$ ,  $\{U_{it}\}$  es una sucesión aleatoria de media cero e independiente de  $i$ ,  $\{U_{it}, \mathcal{F}_{it-1}\}$  es una m.d.s. en Lee y Hong (2001) demostrando que para cada  $i$  y  $T_i$  suficientemente grande,

$$E \left\{ \sum_{t=q_i+2}^{T_i} [E(U_{it}^2 | \mathcal{F}_{i,t-1}) - EU_{it}^2] \right\}^2 \leq C(\bar{q}/T) \left[ \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 + C(J_i+1)2^{J_i}$$

De allí que

$$E(\tilde{U}^2 - E\tilde{U}^2)^2 = \sum_{i=1}^n E \left\{ \sum_{t=q_i+2}^{T_i} [E(U_{it}^2 | \mathcal{F}_{i,t-1}) - EU_{it}^2] \right\}^2 \\ \leq C(\bar{q}/T)V_{nT}^2 + C(\bar{J} + 1)V_{nT}$$

para todo  $T$  suficientemente grande, donde la igualdad sigue del hecho que  $\{\sum_{t=q_i+2}^{T_i} [E(U_{it}^2 | \mathcal{F}_{i,t-1}) - EU_{it}^2]\}$  es una sucesión aleatoria de media cero e independiente de  $i$ , y la inecuación sigue del Lema (4.1.1, ii) y de que  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ . Luego, dado que  $\bar{q}^2/T \rightarrow 0$  y  $2^{2\bar{J}}/T \rightarrow 0$ , tenemos  $V_{nT}^{-2} E(\tilde{U}^2 - E\tilde{U}^2)^2 = O(\bar{q}/T) + O[V_{nT}^{-2} \sum_{i=1}^n (J_i+1)2^{J_i}] \rightarrow 0$  cuando  $n, T \rightarrow \infty$ . Así, la condición (b) se cumple y por tanto  $V_{nT}^{-1/2} \hat{U} \xrightarrow{d} N(0, 1)$  por el teorema de Brown.  $\square$

**Corolario 4.1.7.** *Suponga que se mantienen los supuestos del 1 al 9 (pág. 1524, 1525, 1528, 1531, 1533-1535, [32]) y  $\hat{J}_0$  esta dado como:  $\hat{J}_0 \equiv \max\{\lfloor \frac{1}{2q+1} \lg_2(2q\lambda_q \hat{\xi}_0(q)T) - 1 \rfloor, 0\}$ . Si  $\{\varepsilon_{it}\}$  en (4.1.26) es i.i.d. para cada  $i$ , entonces  $\hat{W}_1(\hat{J}_0) \xrightarrow{d} N(0, 1)$  y  $\hat{W}_2(\hat{J}_0) \xrightarrow{d} N(0, 1)$ .*

*Demostración del Corolario (4.1.7).* El resultado se sigue del Teorema 5 (4.1.12), por el supuesto 9 (Pág. 1535-1536, [32]) que implica,  $2^{\hat{J}}/2^J - 1 = o_p(T^{-1/(2(2q+1))}) = o_p(2^{-J/2})$ , donde el estocástico de escala mas fina  $J$  está dado por:  $2^{J+1} \equiv \max\{\lfloor 2\alpha_q^2 \hat{\zeta}_0(q)T \rfloor^{1/(2q+1)}, 0\}$ . Este último satisface las condiciones del Teorema 5 (4.1.12).  $\square$

#### 4.1.4. Igualdad de varianzas y acotamiento

**Proposición 4.1.8.**

$$V_{nT}^{-1/2} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})^2 = O_P[V_{nT}^{-1/2} + (n^{-1} + T^{-1})V_{nT}^{1/2}]$$

*Demostración de la Proposición 4.1.8.* Por la definición de  $\hat{v}_{it}$  en:

$$\hat{v}_{it} = \hat{u}_{it} - \bar{u}_i - \bar{u}_t + \bar{u}, \quad (t = 1, 2, \dots, T_i; i = 1, 2, \dots, n)$$

$$\begin{aligned} \text{donde } \hat{u}_{it} &\equiv Y_{it} - X'_{it}\hat{\beta}, \\ \bar{u}_i &\equiv T_i^{-1} \sum_{t=1}^{T_i} \hat{u}_{it}, \\ \bar{u}_{\cdot t} &\equiv n^{-1} \sum_{i=1}^n \hat{u}_{it}, \text{ y} \\ \bar{u} &\equiv n^{-1} \sum_{i=1}^n T_i^{-1} \sum_{t=1}^{T_i} \hat{u}_{it}, \end{aligned}$$

teniendo que  $\hat{\beta}$  es un estimador consistente para  $\beta$  dentro de  $\mathbb{H}_0$ , tomamos  $\hat{v}_{it} \equiv \tilde{v}_{it} - \tilde{X}'_{it}(\hat{\beta} - \beta)$ , donde  $\tilde{X}'_{it}$  y  $\tilde{v}_{it}$  son los de la suposición 5 (pág. 1528, [32]). Nótese que bajo  $\mathbb{H}_0$ ,  $v_{it}$  en:  $v_{it} \equiv \varepsilon_{it} + (\beta - \beta^*)'(X_{it} - E\bar{X}_i - E\bar{X}_{\cdot t} + E\bar{X})$ , coincide con los verdaderos errores  $\varepsilon_{it}$  en:

$$\hat{\beta} \equiv \left[ \sum_{i=1}^n \sum_{t=1}^{T_i} \tilde{X}_{it} \tilde{X}'_{it} \right]^{-1} \left[ \sum_{i=1}^n \sum_{s=1}^{T_i} \tilde{X}_{it} \tilde{Y}_{it} \right]$$

$$\begin{aligned} \text{donde } \tilde{X}_{it} &\equiv X_{it} - \bar{X}_i - \bar{X}_{\cdot t} + \bar{X}, \\ \bar{X}_i &\equiv T_i^{-1} \sum_{t=1}^{T_i} X_{it}, \\ \bar{X}_{\cdot t} &\equiv n^{-1} \sum_{i=1}^n X_{it} \text{ y} \\ \bar{X} &\equiv n^{-1} \sum_{i=1}^n T_i^{-1} \sum_{t=1}^{T_i} X_{it}, \end{aligned}$$

los cuales i.i.d. para cada  $i$ ,  $v_{it}$  y  $v_{ls}$  son independientes para todo  $i \neq l$  y todo  $t, s$ .

Tomando  $\tilde{R}_i(h) \equiv T^{-1} \sum_{t=|h|+1}^T \tilde{v}_{it} \tilde{v}_{it-|h|}$ , escribimos

$$\begin{aligned} \hat{R}_i(h) - \tilde{R}_i(h) &= (\hat{\beta} - \beta)' \Gamma_{ixx}(h) (\hat{\beta} - \beta) - (\hat{\beta} - \beta)' \Gamma_{ixv}(h) - (\hat{\beta} - \beta)' \Gamma_{ivx}(h) \\ &\equiv \sum_{c=1}^3 \hat{\xi}_{ci}(h) \end{aligned} \quad (4.1.14)$$

donde  $\tilde{\Gamma}_{ixx}(h) \equiv \bigwedge T_i^{-1} \sum_{t=|h|+1}^{T_i} \tilde{X}_{it} \tilde{X}'_{it-|h|}$ ,  $\tilde{\Gamma}_{ixv}(h)$  y  $\tilde{\Gamma}_{ivx}(h)$  son las de la suposición 5 (pág. 1528, [32]).

Recordando la definición de  $\hat{R}_i(h)$  en:

$$\begin{aligned} \bar{f}_i(\omega) &\equiv (2\pi)^{-1} \bar{R}_i(0) + \sum_{j=1}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk} \Psi_{jk}(\omega), \quad \omega \in [-\pi, \pi], \text{ donde} \\ \bar{R}_i(h) &\equiv T_i^{-1} \sum_{t=h+1}^{T_i} v_{it} v_{it-|h|} \text{ y } \bar{\alpha}_{ijk} \equiv (2\pi)^{-1/2} \sum_{h=1-T_i}^{T_i-1} \bar{R}_i(h) \hat{\Psi}_{jk}^*(h). \end{aligned} \quad (4.1.15)$$

podemos escribir

$$\begin{aligned}\widehat{R}_i(h) - \widetilde{R}_i(h) &= T_i^{-1} \sum_{t=|h|+1}^{T_i} (-\widetilde{v}_i \cdot v_{it} - \widetilde{v}_i \cdot \widetilde{v}_{it-|h|} - v_{it} \widetilde{v}_{t-|h|} - \widetilde{v}_t \widetilde{v}_{it-|h|} + \widetilde{v} v_{it} + \widetilde{v} \widetilde{v}_{it-|h|}) \\ &= \sum_{c=4}^9 \widehat{\xi}_{ci}(h)\end{aligned}\quad (4.1.16)$$

Tomando (4.1.14) y (4.1.16), tenemos  $\widehat{\alpha}_{ijk} - \widetilde{\alpha}_{ijk} = (2\pi)^{-1/2} \sum_{c=1}^9 \sum_{h=1-T_i}^{T_i-1} \widehat{\xi}_{ci}(h) \widehat{\Psi}_{jk}(h)$ , de aquí que

$$\begin{aligned}\sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} (\widehat{\alpha}_{ijk} - \widetilde{\alpha}_{ijk})^2 \widehat{\xi}_{ci}(h) &\leq 2^5 \sum_{c=1}^9 \left\{ \sum_{i=1}^n T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \left[ \sum_{h=1-T_i}^{T_i-1} \widehat{\xi}_{ci}(h) \widehat{\Psi}_{jk}(h) \right]^2 \right\} \\ &\equiv 2^5 \sum_{c=1}^9 \widehat{A}_c.\end{aligned}\quad (4.1.17)$$

Ahora se demostrará que  $V_{nT}^{-1/2} \widehat{A}_c \xrightarrow{p} 0$  para  $1 \leq c \leq 9$ . Primero se considera  $\widehat{A}_1$ . Para (4.1.16), tenemos  $|\widehat{\xi}_{1i}(h)| \leq \|\widehat{\beta} - \beta\|^2 \|\widehat{\Gamma}_{ixx}(h)\| \leq \|\widehat{\beta} - \beta\|^2 \|\widehat{\Gamma}_{ixx}(0)\|$ . Sea  $b_{ji}(h, m)$  como está definido en el Lema (4.1.1). Entonces tenemos

$$\begin{aligned}V_{nT}^{-1/2} |\widehat{A}_1| &= V_{nT}^{-1/2} \left| \sum_{i=1}^n T_i \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \widehat{\xi}_{1i}(h) \widehat{\xi}_{1i}(m) \right| \\ &\leq V_{nT}^{-1/2} \|\widehat{\beta} - \beta\|^4 \left[ \sum_{i=1}^n T_i \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}^2(h, m) \right]^{-1/2} \left[ \sum_{i=1}^n T_i^3 \|\widehat{\Gamma}_{ixx}(0)\|^4 \right]^{-1/2} \\ &= O_p(n^{-3/2})\end{aligned}$$

dato el Lema (4.1.1)(vi),  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ , la suposición de la 3 a la 5 (pág. 1528, [32]))  $T_i \leq CT$  y  $\sigma_i^2 \in [c, C]$ .

Ahora, consideremos el segundo término  $\widehat{A}_2$  en (4.1.17). Recordando que  $\Gamma_{ix\nu}(h) \equiv p \lim_{T_i \rightarrow \infty} \widetilde{\Gamma}_{ix\nu}(h)$ , tenemos  $\widehat{\xi}_{2i}(h) = (\widehat{\beta} - \beta)' \Gamma_{ix\nu}(h) (\widehat{\beta} -$



$\beta)'[\tilde{\Gamma}_{ix\nu}(h) - \Gamma_{ix\nu}(h)]$ . Así,

$$\begin{aligned} \hat{A}_2 &\leq 2\|\hat{\beta} - \beta\|^2 \\ &\sum_{i=1}^n T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \left\{ \left\| \sum_{h=1-T}^{T-1} \hat{\Gamma}_{ix\nu}(h) \hat{\Psi}_{jk}(h) \right\|^2 + \left\| \sum_{h=1-T}^{T-1} [\tilde{\Gamma}_{ix\nu}(h) - \Gamma_{ix\nu}(h)] \hat{\Psi}_{jk}(h) \right\|^2 \right\} \\ &\equiv 2\|\hat{\beta} - \beta\|^2 \hat{M}_1 + 2\|\hat{\beta} - \beta\|^2 \hat{M}_2 \end{aligned} \quad (4.1.18)$$

Ahora consideramos  $\hat{M}_1$  en (4.1.18). Sea  $\Lambda_{ijk}^{x\nu} \equiv \int_{-\pi}^{\pi} f_{ix\nu}(\omega) \Psi_{jk}(\omega) d\omega$ , donde  $f_{ix\nu}(\omega) \equiv (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \Gamma_{ix\nu}(h) e^{-ij\omega}$ , entonces por la identidad de Parseval's  $\Lambda_{ijk}^{x\nu} = (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} \Gamma_{ix\nu}(h) \hat{\Psi}_{jk}(h)$  y  $\sum_{h=1-T_i}^{T_i-1} \Gamma_{ix\nu}(h) \hat{\Psi}_{jk}(h) = (2\pi)^{1/2} \Lambda_{ijk}^{x\nu} + \sum_{|h| \geq T_i} \Gamma_{ix\nu}(h) \hat{\Psi}_{ij}(h)$ . Por la desigualdad de Cauchy-Schwarz, tenemos

$$\begin{aligned} \hat{M}_1 &\leq 2 \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \|\Lambda_{ijk}^{x\nu}\|^2 + 2 \sum_{i=1}^n T_i \sum_{|h| \geq T_i} \|\Gamma_{ix\nu}(h)\|^2 \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \sum_{|h| \geq T_i} |\hat{\Psi}_{jk}(h)|^2 \\ &= O(nT) + o[nT(2^{\bar{J}/T})^{2\tau}] \\ &= O(nT) \end{aligned} \quad (4.1.19)$$

dato  $2^{\bar{J}/T} \rightarrow 0$ , donde  $\bar{J} \equiv \max_{1 \leq i \leq n} (J_i)$ . Aquí, tenemos que usar el hecho de que

a) Por la identidad de Parseval's y la suposición 5 (Pág. 1528, [32]), se tiene

$$\sum_{j=0}^{J_i} \times \sum_{k=1}^{2^j} \|\Lambda_{ijk}^{x\nu}(h)\|^2 = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \|\Gamma_{ix\nu}(h)\|^2 < C$$

b) Por la suposición 5 (Pág. 1528, [32]) y  $T_i = c_i T \geq cT$ , se tiene

$$\sum_{|h| \geq T_i} \|\Gamma_{ix\nu}(h)\|^2 = o(T^{-1})$$

c) Dada la suposición 2 (Pág. 1525, [32]), se tiene

$$\sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \sum_{h \geq |T_i|} |\hat{\Psi}_{jk}(h)|^2 \leq \sum_{j=0}^{J_i} \sum_{h \geq |T_i|} |\hat{\psi}(2\pi h/2^j)|^2 \leq C 2^{2\tau J_i} / T_i^{2\tau-1}$$

Por el segundo término  $\widehat{M}_2$  en (4.1.18), se tiene:

$$\begin{aligned} \|\widehat{\beta} - \beta\|^2 \widehat{M}_2 &\leq \|\widehat{\beta} - \beta\|^2 \sum_{i=1}^n T_i \sum_{k=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \|\widetilde{\Gamma}_{ix\nu}(h) - \Gamma_{ix\nu}(h)\| \|\widetilde{\Gamma}_{ix\nu}(m) - \Gamma_{ix\nu}(m)\| \\ &= O_p \left[ (nT)^{-1} \sum_{i=1}^n \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right] \\ &= O_p[(nT)^{-1} V_{nT}] \end{aligned} \quad (4.1.20)$$

dato el Lema (4.1.1)(ii),  $V_{nT} \leq C \sum_{i=1}^T 2^{J_i+1}$ , y la suposición 5 (Pág. 1528, [32]). Combinando (4.1.18) y (4.1.19) obtenemos  $V_{nT}^{-1} \widehat{A}_2 = O_p(V_{nT}^{-1} + (nT)^{-1} V_{nT}^{1/2})$ . De manera similar, tenemos  $V_{nT}^{-1/2} \widehat{A}_3 = O_p(V_{nT}^{-1/2} + (nT)^{-1} V_{nT}^{1/2})$ .

Ahora consideremos el término  $\widehat{A}_4$  en (4.1.17). Por la desigualdad de Cauchy-Schwarz y el hecho que bajo  $\mathbb{H}_0$ ,  $\{v_{it}\}$  coincide con  $\{\varepsilon_{it}\}$  son i.i.d. con  $E v_{it}^8 \leq C$  para cada  $i$ , tenemos  $E(\bar{v}_i^2 |T_i^{-1} \sum_{t=h+1}^{T_i} v_{it}| |T_i^{-1} \sum_{t=m+1}^{T_i} v_{it}|) \leq C T_i^{-2}$  para  $h, m > 0$ . De la desigualdad de Markov, el Lema (4.1.1)(ii) y  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$  se tiene

$$\begin{aligned} V_{nT}^{-1/2} \widehat{A}_4 &\leq V_{nT}^{-1/2} \sum_{i=1}^n T_i \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \bar{v}_i^2 \left| T_i^{-1} \sum_{t=h+1}^{T_i} v_{it} \right| \left| T_i^{-1} \sum_{t=m+1}^{T_i} v_{it} \right| \\ &= O_p(T^{-1} V_{nT}^{1/2}) \end{aligned} \quad (4.1.21)$$

De manera similar para  $\widehat{A}_5$ , tenemos  $V_{nT}^{-1/2} \widehat{A}_5 = O_p(T^{-1} V_{nT}^{1/2})$ .

Para la  $\widehat{A}_6$  en (4.1.17), note que  $v_{it}$  y  $\bar{v}_{t-h}$  son independientes para  $h > 0$  bajo  $\mathbb{H}_0$ , por la desigualdad de Cauchy-Schwarz y  $E v_{it}^8 \geq C$ , tenemos  $E(|\bar{v}_{t-h} \bar{v}_{t-m}| |T_i^{-1} \sum_{t=h+1}^{T_i} v_{it}| |T_i^{-1} \sum_{t=m+1}^{T_i} v_{it}|) \leq C(nT_i)^{-1}$  para  $h, m > 0$ , llegando a  $V_{nT}^{-1/2} \widehat{A}_6 = O_p(n^{-1} V_{nT}^{1/2})$ . De manera similar, tenemos  $V_{nT}^{-1/2} \widehat{A}_7 = O_p(n^{-1} V_{nT}^{1/2})$ .

Finalmente, dado  $E(\bar{v}^2 |T_i^{-1} \sum_{t=h+1}^{T_i} v_{it}| |T_i^{-1} \sum_{t=m+1}^{T_i} v_{it}|) \leq C n^{-1} T_i^{-2}$  para  $h, m > 0$ , bajo  $\mathbb{H}_0$ , tenemos  $V_{nT}^{-1/2} \widehat{A}_c = O_p[(nT)^{-1} V_{nT}^{1/2}]$  para  $c = 8, 9$ .

Así hemos demostrado que  $V_{nT}^{-1/2} \hat{A}_c \xrightarrow{p} 0$  para todo  $1 \leq c \leq 9$  dado  $\max_{i \leq i \leq n} 2^{2(J_i+1)}/(n^2 + T) \rightarrow 0$ . La proposición (4.1.8) sigue de (4.1.17).  $\square$

**Proposición 4.1.9.**

$$V_{nT}^{-1/2} \left( \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk}^2 - M_{nT} \right) = V_{nT}^{-1/2} \sum_{i=1}^n \hat{A}_i + o_p(1)$$

Primero descomponemos  $\hat{A}_i$  en los términos con  $t - s > q_i$  y  $t - s \leq q_i$ , para algún entero  $q_i \in \mathbb{Z}^+$

$$\begin{aligned} \hat{A}_i &= 2T_i^{-1} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} b_{J_i}(h, m) \left( \sum_{t=q_i+2}^{T_i} \sum_{s=1}^{t-q_i-1} + \sum_{t=2}^{T_i} \sum_{s=\max(t-q_i, 1)}^{t-1} \right) v_{it} v_{it-h} v_{is} v_{is-m} \\ &\equiv \hat{B}_i + \hat{B}_{4i}. \end{aligned} \quad (4.1.22)$$

Además, descomponemos  $\hat{B}_i$ ,

$$\begin{aligned} \hat{B}_i &= 2T_i^{-1} \left( \sum_{h=1}^{q_i} \sum_{m=1}^{q_i} + \sum_{h=1}^{q_i} \sum_{m=q_i+1}^{n-1} + \sum_{h=q_i+1}^{n-1} \sum_{m=1}^{n-1} \right) b_{J_i}(h, m) \sum_{t=q_i+2}^{T_i} \sum_{s=1}^{t-q_i-1} v_{it} v_{it-h} v_{is} v_{is-m} \\ &\equiv \hat{U}_i + \hat{B}_{5i} + \hat{B}_{6i}. \end{aligned} \quad (4.1.23)$$

*Demostración de la proposición (4.1.9).* Recordamos la definición de  $M_{nT}$  en el teorema 4.1.5. Obtenemos

$$\sum_{i=1}^n \left( 2\pi T_i \sum_{j=0}^{J_i} \sum_{k=1}^{2^j} \bar{\alpha}_{ijk}^2 - M_{nT} \right) = \sum_{i=1}^n \hat{A}_i + \sum_{i=1}^n (\hat{B}_{1i} - \alpha_i^4 M_{i0}) - \sum_{i=1}^n \hat{B}_{2i} - \sum_{i=1}^n \hat{B}_{3i}.$$

Demostraremos

- a)  $V_{nT}^{-1/2} (\sum_{i=1}^n \hat{B}_{1i} - M_{nT}) \xrightarrow{p} 0$ ,
- b)  $V_{nT}^{-1/2} \sum_{i=1}^n \hat{B}_{2i} \xrightarrow{p} 0$ ,
- c)  $V_{nT}^{-1/2} \times \sum_{i=1}^n \hat{B}_{3i} \xrightarrow{p} 0$ ,

(a) Observemos que  $\widehat{B}_{1i}$  tiene una estructura similar a  $\widehat{B}_{1n}$  en Lee y Hong (2001). Siguiendo el razonamiento de Lee y Hong (2001) y usando el Lema 4.1.1(ii), podemos obtener para cada  $i$  y  $T_i$  suficientemente grande,

$$E(\widehat{B}_{1i} - E\widehat{B}_{1i})^2 \leq CT_i^{-1} \left[ \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 \leq C^2(2^{\widetilde{J}}/T) \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)|$$

Puesto que  $\{\widehat{B}_{1i}\}$  es una sucesión aleatoria independiente de  $i$  y  $\sum_{i=1}^n E\widehat{B}_{1i} = M_{nT}$ , tenemos  $E(\widehat{B}_{1i} - M_{nT})^2 = \sum_{i=1}^n E(\widehat{B}_{1i} - E\widehat{B}_{1i})^2 = O(V_{nT}2^{\widetilde{J}}/T)$  dado el Lema 4.1.1 (ii) y  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ . Luego, por la desigualdad de Chebyshev y  $2^{2\widetilde{J}}/T \xrightarrow{p} 0$ , tenemos  $V_{nT}^{-1/2}(\sum_{i=1}^n \widehat{B}_{1i} - M_{nT}) = O_p[(2^{\widetilde{J}}/T)^{1/2}] = o_p(1)$ .

(b) Ahora consideremos  $\widehat{B}_{2i}$ . Siguiendo a Lee y Hong (2001), tenemos  $E\widehat{B}_{2i}^2 \leq CT_i^{-1} \times [\sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)|]^3$ . Entonces, por el hecho que  $\widehat{B}_{2i}$  es una sucesión aleatoria con media cero independiente de  $i$ , por el Lema 4.1.1 (ii) y  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ , tenemos  $E(\sum_{i=1}^n \widehat{B}_{2i})^2 = \sum_{i=1}^n E\widehat{B}_{2i}^2 = O_p[(2^{\widetilde{J}}/T)V_{nT}]$ . Luego  $V_{nT}^{-1/2} \sum_{i=1}^n \widehat{B}_{2i} \xrightarrow{p} 0$  por la desigualdad de Chebyshev y  $2^{2\widetilde{J}}/T \rightarrow 0$ .

(c) Por razonamiento similar que (b), se obtiene  $V_{nT}^{-1/2} \times \sum_{i=1}^n \widehat{B}_{3i} \xrightarrow{p} 0$ . □

**Proposición 4.1.10.** *Teniendo los Supuestos 1 y 2 (Pág. 1524-1525, [32])  $2^{2\widetilde{J}}/T \rightarrow 0$ ,  $q_i \equiv q_i(T_i) \rightarrow \infty$ ,  $q_i/2^{J_i} \rightarrow \infty$ , y  $q_i^2/T_i \rightarrow 0$ , donde  $\widetilde{J} \equiv \max_{1 \leq i \leq n}(J_i)$ . Si  $\{v_{it}\}$  son i.i.d. para cada  $i$ , entonces  $V_{nT}^{-1/2} \sum_{i=1}^n \widehat{A}_i = V_{nT}^{-1/2} \sum_{i=1}^n \widehat{U}_i + o_p(1)$ .*

*Demostración de la proposición (4.1.10).* Dado (4.1.22) y (4.1.23), tenemos  $\widehat{A}_i = \widehat{U}_i + \widehat{B}_{4i} + \widehat{B}_{5i} + \widehat{B}_{6i}$ . Es suficiente mostrar que  $V_{nT}^{-1/2} \sum_{i=1}^n \widehat{B}_{ci} \xrightarrow{p} 0$  para  $c = 4, 5, 6$ .

(a) Consideremos primero  $\widehat{B}_{4i}$  en (4.1.22). De Lee y Hong (2001, demostración del Teorema 1), tenemos para cada  $i$  y  $T_i$  suficientemente grande,

$$E\hat{B}_{4i}^2 \leq C(q_i/T_i) \left[ \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)| \right]^2 \leq C^2(\bar{q}2^{\bar{J}}/T) \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |b_{J_i}(h, m)|$$

donde  $\bar{q} \equiv \max_{1 \leq i \leq n} (q_i)$  y la última inecuación sigue del Lema 4.1.1(ii). Luego, usando el hecho que  $\{\hat{B}_{4i}\}$  es una sucesión aleatoria con media cero independiente de  $i$ . Por el Lema 4.1.1(ii) y  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ , tenemos  $E(\sum_{i=1}^n \hat{B}_{4i})^2 = \sum_{i=1}^n E\hat{B}_{4i}^2 = O(V_{nT}\bar{q}2^{\bar{J}}/T)$ . Así, la desigualdad de Chebyshev,  $\bar{q}^2/T \rightarrow 0$  y  $2^{2\bar{J}}/T \rightarrow 0$ , implica  $V_{nT}^{-1/2} \sum_{i=1}^n \hat{B}_{4i} \xrightarrow{p} 0$ .

(b) Ahora consideramos  $\hat{B}_{5i}$  como en (4.1.23). Por la definición de  $b_{J_i}(h, m)$ , la desigualdad de Cauchy-Schwarz y la suposición 2 (Pág. 1525,[32]), tenemos

$$E\hat{B}_{5i}^2 = \sigma_i^4 \sum_{h=1}^{T-1} \sum_{m>q_i}^{T-1} b_{J_i}^2(h, m) \leq C \sum_{j=0}^{J_i} \sum_{m>q_i} |\hat{\psi}(2\pi m/2^j)|^2 \leq C^2 2^{2\tau J_i} / q_i^{2\tau-1}$$

Por tanto,  $E(\sum_{i=1}^n \hat{B}_{5i})^2 = \sum_{i=1}^n E\hat{B}_{5i}^2 \leq C(2^{\bar{J}}/q_0)^{2\tau-1} \sum_{i=1}^n 2^{J_i}$ , donde  $q_0 \equiv \min_{1 \leq i \leq n} (q_i)$ . Seguido por la desigualdad de Chebyshev y  $2^{\bar{J}}/q_0 \rightarrow 0$  para que  $V_{nT}^{-1/2} \sum_{i=1}^n \hat{B}_{5i} = O_p[(2^{\bar{J}}/q_0)^{2\tau-1}] = o_p(1)$ .

(c) Finalmente, consideramos  $\hat{B}_{6i}$  en (4.1.23). Siguiendo a Lee y Hong (2001, demostración del Teorema 1 (4.1.2)), obtenemos

$$E\hat{B}_{6i}^2 \leq C2^{2\tau J_i} / q_i^{2\tau-1} + CT_i^{-1} \left[ \sum_{h=1}^{T-1} \sum_{m=1}^{T-1} |b_{J_i}(h, m)| \right]^2 \quad (4.1.24)$$

$$\leq C2^{J_i} (2^{J_i}/q_0)^{2\tau-1} + C(2^{\bar{J}}/T) \sum_{h=1}^{T-1} \sum_{m=1}^{T-1} |b_{J_i}(h, m)| \quad (4.1.25)$$

Por el Lema (4.1.1, ii) y la suposición 2 (Pág. 1525,[32]). Así  $V_{nT}^{-1/2} \sum_{i=1}^n \hat{B}_{6i} = O_p[(2^{\bar{J}}/q_0)^{(2\tau-1)} + (2^{\bar{J}}/T)^{1/2}] \xrightarrow{p} 0$ . Por la desigualdad de Chebyshev  $V_{nT} \leq C \sum_{i=1}^n 2^{J_i+1}$ ,  $2^{\bar{J}}/q_0 \rightarrow 0$  y  $2^{2\bar{J}}/T \rightarrow 0$ . Así se completa la demostración de la proposición (4.1.18).  $\square$

#### 4.1.5. Monotonía para $a_i$ en $Q(f_i, f_{i0})$

**Teorema 4.1.11.** *Supónganse que las suposiciones de la 1 a la 6 (Pág. 1524-1531, [32]) son verdaderas,  $n = \gamma T^\varsigma$  para  $\gamma \in (0, \infty)$  y  $\varsigma \in (0, \infty)$ , y  $2^{J_i+1} = a_i T_i^\nu$  para  $a_i \in [c, C]$  y  $\nu \in (0, 1)$ . Si  $a_i$  es monótonamente creciente en  $Q(f_i, f_{i0})$  y  $T_i = T$ , para todo  $i$ .*

*Demostración del Teorema (4.1.11).* Dado  $T_i = c_i T$  y  $2^{J_i+1} = a_i T_i^\nu$ , tenemos  $2^{J_i+1} = b_i T^\nu$ , donde  $b_i \equiv a_i c_i^\nu$ . Cuando  $T \rightarrow \infty$ , tenemos:

$$V_{nT}^0 \equiv n^{-1} \sum_{i=1}^n \sigma_i^8 (2^{J_i+1} - 1) = T^\nu \left( n^{-1} \sum_{i=1}^n \sigma_i^8 b_i \right) = \bar{b} T^\nu [1 + o(1)],$$

donde  $\bar{b} = n^{-1} \sum_{i=1}^n \sigma_i^8 b_i$ . Se deduce del Teorema 2 (4.1.5) y  $\hat{V}/V_{nt}^0 \xrightarrow{p} 1$  tal que

$$\begin{aligned} (n_A T)^{-1} (\bar{b} T^\nu)^{1/2} \hat{W}_1 &= n_A^{-1} \sum_{i \in N_A} c_i Q(f_i, f_{i0}) + o_p(1) \text{ y,} \\ (n_A T)^{-1} (\bar{b} T^\nu)^{1/2} \hat{W}_2 &= n_A^{-1} \sum_{i \in N_A} (\bar{b}/\sigma_i^8 b_i)^{1/2} c_i Q_i + o_p(1). \end{aligned}$$

Para  $c = 1, 2$ , ponemos  $S_{nT}^{(c)} \equiv -2 \ln[1 - \Phi(\hat{W}_c)]$ , donde  $\Phi(\cdot)$  es la  $N(0, 1)$  CDF. Porque  $\ln[1 - \Phi(z)] = -\frac{1}{2} z^2 [1 + o(1)]$  como  $z \rightarrow +\infty$ , tenemos

$$\begin{aligned} (n_A T)^{-2} \bar{b} T^\nu S_{nT}^{(1)} &= \left[ n_A^{-1} \sum_{i \in N_A} c_i Q(f_i, f_{i0}) \right]^2 + o_p(1) \\ (n_A T)^{-2} \bar{b} T^\nu S_{nT}^{(1)} &= \left[ n_A^{-1} \sum_{i \in N_A} (\bar{b}/\sigma_i^8 b_i)^{1/2} c_i Q(f_i, f_{i0}) \right]^2 + o_p(1). \end{aligned}$$

Suponga  $\{T_i^{(1)}\}_{i=1}^{n^{(1)}}$  y  $\{T_i^{(2)}\}_{i=1}^{n^{(2)}}$  son dos secuencias de tamaños de muestra utilizados para  $\hat{W}_1$  y  $\hat{W}_2$  respectivamente de modo que  $S_{n^{(1)}T^{(1)}}^{(1)}/S_{n^{(2)}T^{(2)}}^{(2)} \rightarrow 1$  como  $n^{(1)}, n^{(2)}, T^{(1)}$ , y  $T^{(2)} \rightarrow \infty$ . Entonces eficiencia relativa de Bahadur

de  $\hat{W}_1$  a  $\hat{W}_2$

$$\begin{aligned} BE(\hat{W}_1 : \hat{W}_2) &\equiv \lim \frac{\sum_{i=1}^{n^{(2)}} T_i^{(2)}}{\sum_{i=1}^{n^{(1)}} T_i^{(1)}} = \lim \frac{\left( \frac{1}{n^{(2)}} \sum_{i=1}^{n^2} c_i \right) n^{(2)} T^{(2)}}{\left( \frac{1}{n^{(1)}} \sum_{i=1}^{n^{(1)}} c_i \right) n^{(1)} T^{(1)}} \\ &= \left( \frac{\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \sqrt{b/(\sigma_i^8 b_i)} c_i Q(f_i, f_{i0})}{\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n c_i Q(f_i, f_{i0})} \right)^{(1+k)/(3-\nu)} \end{aligned}$$

De la última igualdad se tiene  $n^{(c)} = \gamma [T^{(c)}]^k$  para  $c = 1, 2$ . Por lo tanto,  $BE(\hat{W}_1 : \hat{W}_2) > 1$  si  $a_i$  es una función monótonamente creciente  $Q(f_i, f_{i0})$  y  $c_i = c$  (i.e.,  $T_i = T$ ) para todo  $i$ .  $\square$

#### 4.1.6. Convergencia en probabilidad y en distribución

**Teorema 4.1.12.** *Supónganse que las suposiciones de la 1 a la 5 y la 7 son verdaderas (pág. 1524-1529 y 1533,[32]), y  $\hat{J}$  es un dato de manejo de escala fina con  $\frac{2^{\hat{J}}}{2^J} = 1 + o_p(2^{-J/2})$ , donde  $J$  es una escala fina no estocástica, tal que  $\frac{2^J}{(n^2+T)} \rightarrow 0$ , como  $n \rightarrow \infty$ . Si  $\{\varepsilon_{it}\}$  en:*

$$Y_{it} = \alpha + X'_{it}\beta + \mu_i + \lambda_t + \varepsilon_{it}, \quad (t = 1, \dots, T_i; i = 1, \dots, n; n, T_i \in \mathbb{Z}^+), \quad (4.1.26)$$

es i.i.d para cada  $i$ , entonces  $W_1(\hat{J}) - W_1(J) \xrightarrow{p} 0$ ,  $W_2(\hat{J}) - W_2(J) \xrightarrow{p} 0$ ,  $W_1(\hat{J}) \xrightarrow{p} N(0, 1)$  y  $W_2(\hat{J}) \xrightarrow{p} N(0, 1)$ .

*Demostración del Teorema (4.1.12).* Solamente consideramos  $\hat{W}_1(\hat{J})$ ; la prueba para  $\hat{W}_2(\hat{J})$  es similar. Escribimos

$$\begin{aligned} &\hat{W}_1(\hat{J}) - \hat{W}_2(J) = \\ &= \hat{V}(\hat{J})^{-1/2} \left\{ \sum_{i=1}^n 2\pi T_i \sum_{j=J}^{\hat{J}} \sum_k^{2^j} \hat{\alpha}_{ijk}^2 - [\hat{M}(\hat{J}) - \hat{M}(J)] \right\} - [\hat{V}(\hat{J})^{1/2}/\hat{V}(J)^{1/2} - 1] \hat{W}_1(J) \end{aligned}$$

## 70 Caso de Aplicación

Dado  $\hat{W}_1(J)$  por Teorema 1 (4.1.2) y  $\hat{V}(J)/V_{nT} \xrightarrow{p} 1$  por Teorema 3 (4.1.3), basta para  $\hat{W}_1(\hat{J}) - \hat{W}_1(J) \xrightarrow{p} 0$  y  $\hat{W}_1(\hat{J}) \xrightarrow{d} N(0, 1)$  si (a)  $V_{nT}^{-1/2} \left\{ \sum_{i=1}^n 2\pi T_i \sum_{j=J}^{\hat{J}} \sum_k^{2^j} \hat{\alpha}_{ijk}^2 - [\hat{M}(\hat{J}) - \hat{M}(J)] \right\} \xrightarrow{p} 0$  y (b)  $\hat{V}(\hat{J})/\hat{V}(J) \xrightarrow{p} 1$ .

Mostramos primero a (a). Descomponiendo

$$\begin{aligned} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{\hat{J}} \sum_{k=1}^{2^j} \hat{\alpha}_{ijk}^2 &= \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{\hat{J}} \sum_{k=1}^{2^j} [(\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})^2 + \bar{\alpha}_{ijk}^2 + (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})\bar{\alpha}_{ijk}] \\ &\equiv \hat{G}_1 + \hat{G}_2 + 2\hat{G}_3 \end{aligned} \quad (4.1.27)$$

Para el primer término de (4.1.27), escribimos

$$\hat{G}_1 = \sum_{i=1}^n 2\pi T_i \left( \sum_{j=0}^{\hat{J}} - \sum_{j=0}^J \right) \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})^2 \equiv \hat{G}_{11} - \hat{G}_{12}.$$

Por la proposición 1 (4.1.8), tenemos  $V_{nT}^{-1/2} \hat{G}_{12} \xrightarrow{p} 0$ . A continuación, para cualquiera de las constantes dadas  $M > 0$  y  $\epsilon > 0$ , tenemos

$$P(\hat{G}_{11} > \epsilon) \leq P(\hat{G}_{11} > \epsilon 2^{J/2} M | 2^{\hat{J}}/2^J - 1| \leq \epsilon) + P(2^{J/2} M | 2^{\hat{J}}/2^J - 1| > \epsilon)$$

donde el segundo término se desvanece a 0 como  $n, T \rightarrow \infty$  dado  $2^{J/2} |2^{\hat{J}}/2^J - 1| \xrightarrow{p} 0$ . Para la primera probabilidad, dado  $2^{J/2} M |2^{\hat{J}}/2^J - 1| \leq \epsilon$ , tenemos para todos los  $n$  y  $T_i$  suficientemente grande,  $V_{nT}^{-1/2} \hat{G}_{12} \leq V_{nT}^{-1/2} \sum_{i=1}^n 2\pi T_i \sum_{j=0}^{\hat{J}} \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk} - \bar{\alpha}_{ijk})^2 \xrightarrow{p} 0$  por proposición 1 (4.1.8). Por lo tanto,  $V_{nT}^{-1/2} \hat{G}_1 = o_p(1)$ .

A continuación, consideramos  $\hat{G}_2$ . Por simetría de  $b_J(\cdot, \cdot)$  escribimos

$$\begin{aligned} \hat{G}_2 &= \sum_{i=1}^n \sigma_i^4 \sum_{h=1}^{T_i-1} [b_{\hat{j}}(h, h) - b_J(h, h)] + \sum_{i=1}^n \sum_{h=1}^{T_i-1} [T_i^{-1} \bar{R}_i^2(h) - \sigma_i^4] [b_{\hat{j}}(h, h) - b_J(h, h)] \\ &\quad + 2 \sum_{i=1}^n T_i \sum_{h=1}^{T_i-1} \sum_{m=1}^{h-1} \bar{R}_i(h) \bar{R}_i(m) [b_{\hat{j}}(h, m) - b_J(h, m)] \\ &\equiv \hat{G}_{21} + \hat{G}_{22} + 2\hat{G}_{23}. \end{aligned} \quad (4.1.28)$$



Para el último término  $\hat{G}_{23}$  de  $\hat{G}_2$  (4.1.28), tenemos para cualquier constante  $M > 0$  y  $\epsilon > 0$ ,

$$P(V_{nT}^{-1/2}|\hat{G}_{23}| > \epsilon) \leq P(V_{nT}^{-1/2}|\hat{G}_{23}| > \epsilon, 2^{J/2}M|2^{\hat{J}}/2^J - 1| \leq \epsilon) \\ + P(2^{J/2}M|2^{\hat{J}}/2^J - 1| > \epsilon)$$

donde, de nuevo, el segundo término se convierte en 0 para  $n, t \rightarrow \infty$ . Ahora demostramos que la primera probabilidad desaparece. Tomando  $\bar{T} \equiv \max_{1 \leq i \leq n}(T_i)$ , como se hizo antes. Dado  $2^{J/2}M|2^{\hat{J}}/2^J - 1| \leq \epsilon$  y la definición de  $a_J(h, m)$  como en  $\hat{W}_1$ , tenemos

$$|a_J(h, m) - a_J(h, m)| = 2\pi \sum_{i=\lceil \log_2 2^J(1-\epsilon/M2^{J/2}) \rceil}^{\lceil \log_2 2^J(1+\epsilon/M2^{J/2}) \rceil} |c_j(h, m)\hat{\psi}(2\pi h/2^j)\hat{\psi}^*(2\pi m/2^j)|,$$

donde

$$c_j(h, m) \equiv 2^{-j} \sum_{k=1}^{2^j} e^{i2\pi(m-h)k/2^j} = \begin{cases} 1 & \text{si } m = h + 2^j r \text{ para cualquier entero } r \\ 0 & \text{de lo contrario} \end{cases} \quad (4.1.29)$$

Cf. Priestley (1981, p.392, (6.19)). De ahí que para todo  $n$  y  $T$  suficientemente grande se deduce,

$$E|\hat{G}_{23}| \leq \sum_{h=1-\bar{T}}^{\bar{T}-1} \sum_{m=1-\bar{T}}^{\bar{T}-1} E \left| \sum_{i=1}^n \mathbf{1}(h \leq T_i) \mathbf{1}(m \leq T_i) T_i \bar{R}_i(h) \bar{R}_i(m) \right| \\ \times 2\pi \sum_{i=\lceil \log_2 2^J(1-\epsilon/2^{J/2}M) \rceil}^{\lceil \log_2 2^J(1+\epsilon/2^{J/2}M) \rceil} |c_j(h, m)\hat{\psi}(2\pi h/2^j)\hat{\psi}^*(2\pi m/2^j)|, \\ \leq 2\pi n^{1/2} \sum_{i=\lceil \log_2 2^J(1-\epsilon/2^{J/2}M) \rceil}^{\lceil \log_2 2^J(1+\epsilon/2^{J/2}M) \rceil} 2^j \left[ 2^{-j} \sum_{h=1-\bar{T}}^{\bar{T}-1} |\hat{\psi}(2\pi h/2^j)| \right] \\ \times \left[ \sum_{r=-\infty}^{\infty} |\hat{\psi}(2\pi h/2^j) + 2\pi r| \right] \\ \leq Cn^{1/2}2^{J+1}\epsilon/2^{J/2}M = O(\epsilon V_{nT}^{1/2}/M)$$

## 72 Caso de Aplicación

donde la segunda desigualdad sigue, dada la suposición 2 (pág. 1525,[32]). Por lo tanto,  $P(V_{nT}^{-1/2}|\hat{G}_{23}| > \epsilon, 2^{J/2}M|2^{\hat{J}}/2^J - 1| \leq \epsilon)$  también se desvanece a 0. En consecuencia, tenemos  $V_{nT}^{-1/2}\hat{G}_{23} \xrightarrow{p} 0$ . Del mismo modo, también podemos obtener  $V_{nT}^{-1/2}\hat{G}_{22} \xrightarrow{p} 0$  y

$$V_{nT}^{-1/2}\{\hat{G}_{21} - [\hat{M}(\hat{J}) - \hat{M}(J)]\} = V_{nT}^{-1/2}[\sigma_i^4 - \hat{R}_i^2(0)][b_{\hat{J}}(h, h) - b_J(h, h)] \xrightarrow{p} 0$$

donde  $n^{-1} \sum_{i=1}^n [\hat{R}_i^2(0) - \sigma_i^4] = O_p[(nT)^{-1/2}]$  dado los Supuestos 3-5 (Pág. 1528, [32]), y que bajo  $H_0$ ,  $\{v_{it}\}$  coincide con  $\{\epsilon_{it}\}$  y así es i.i.d. para cada  $i$ . Se deduce de  $\hat{G}_2$  (4.1.28) que  $V_{nT}^{-1/2}\{\hat{G}_2 - [\hat{M}(\hat{J}) - \hat{M}(J)]\} \xrightarrow{p} 0$

A continuación, por la desigualdad de Cauchy-Schwarz, tenemos

$$\begin{aligned} V_{nT}^{-1/2}|\hat{G}_3| &\leq (V_{nT}^{-1/2}\hat{G}_1)^{1/2}(V_{nT}^{-1/2}|\hat{G}_2|)^{1/2} \\ &= O_p[V_{nT}^{-1/4} + (T^{-1/2} + n^{-1/2}V_{nT}^{1/4})]o_p n^{1/4} = o_p(1) \end{aligned}$$

por la Proposición 1 (4.1.8) y el hecho de que

$$n^{-1/2}V_{nT}^{-1/2}\hat{G}_2 = n^{-1/2}V_{nT}^{-1/2}(\hat{G}_{21} + \hat{G}_{22} + \hat{G}_{23}) \xrightarrow{p} 0$$

como se puede demostrar utilizando el razonamiento similar a la de  $\hat{G}_{23}$ . El resultado de (a) sigue de la ecuación 4.1.27.

(b) Para mostrar  $\hat{V}(\hat{J})/\hat{V}(J) = 1 + o_p(1)$  basta con mostrar  $\hat{V}(\hat{J})/V_{nT} \xrightarrow{p} 1$  dado  $\hat{V}(J)/V_{nT} \xrightarrow{p} 1$  por Teorema 3 (4.1.3). Recordando las definiciones de  $\hat{V}(\hat{J})$  y  $V_{nT}$  podemos utilizar un razonamiento análogo al  $\hat{G}_{23}$  para obtener

$$[\hat{V}(\hat{J}) - V_{nT}]/V_{nT} = V_{nT}^{-1} \sum_{i=1}^n [\hat{R}_i^4(0) - \sigma_i^8] \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} [b_{\hat{J}}(h, m) - b_J(h, m)] \xrightarrow{p} 0,$$

donde hemos utilizado el hecho de que

$$n^{-1} \sum_{i=1}^n [\hat{R}_i^4(0) - \sigma_i^8] = O_p[(nT)^{-1/2}]$$

Por lo tanto,  $\hat{V}(\hat{J})/\hat{V}(J) \xrightarrow{p} 1$ . Resulta que  $[\hat{V}(\hat{J})/\hat{V}(J) - 1]\hat{W}_1(J) \xrightarrow{p} 0$  dado que  $\hat{W}_1(J) = O_p(1)$  por Teorema 4.1.2. Por lo tanto, contiene el resultado (b), y tendríamos  $\hat{W}_1(\hat{J}) - \hat{W}_1(J) \xrightarrow{p} 0$ , y  $\hat{W}_1(\hat{J}) \xrightarrow{p} N(0, 1)$ .

Esto completa la demostración.  $\square$

### 4.1.7. Convergencia entre densidades espectrales

**Teorema 4.1.13.** *Suponga los supuestos del 1 al 8 (pág. 1524, 1525, 1528, 1531, 1533, 1534, [32]) son verdaderas,  $\lambda_q \in (0, \infty)$ ,  $J \rightarrow \infty$ ,  $2^{J_i}/T_i \rightarrow 0$  cuando  $T_i \rightarrow \infty$ .*

Entonces: (a) Para cada  $i$ ,  $Q(\hat{f}_i, f_i) = Q(\bar{f}_i, f_i) + o_p(2^{J_i}/T_i + 2^{-2_q J_i})$ , y

$$EQ(\bar{f}_i, f_i) = \frac{2^{J_i+1}}{T_i} \int_{-\pi}^{\pi} f_i^2(\omega) d\omega + (2^{-2_q(J_i+1)}) \lambda_q \int_{-\pi}^{\pi} [f_i^q(\omega)]^2 d\omega + O(2^{J_i}/T_i + 2^{-2_q J_i})$$

(b) Si además  $J_i = J$  para todo  $i$  y  $\frac{T_i}{T} = c_i$ , entonces:

$$n^{-1} \sum_{i=1}^n Q(\hat{f}_i, f_i) = n^{-1} \sum_{i=1}^n Q(\bar{f}_i, f_i) + o_p(2^J/T + 2^{-2_q J}), \text{ y}$$

$$\begin{aligned} n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, f_i) &= \frac{2^{J+1}}{T} n^{-1} \sum_{i=1}^n c_i^{-1} \int_{-\pi}^{\pi} f_i^2(\omega) d\omega \\ &+ 2^{-2_q(J+1)} \lambda_q n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} [f_i^{(q)}(\omega)]^2 d\omega \\ &+ o(2^J/T + 2^{-2_q J}) \end{aligned}$$

*Demostración del Teorema (4.1.13).* Se demostrará el literal (b) solamente. la prueba para el literal (a) es similar y más simple.

Primero mostraremos que

$$n^{-1} \sum_{i=1}^n Q(\hat{f}_i, f_i) = n^{-1} \sum_{i=1}^n Q(\bar{f}_i, f_i) + O_p(2^J/T + 2^{-2_q J})$$

Escribiendo

$$\begin{aligned}
 & n^{-1} \sum_{i=1}^n [Q(\hat{f}_i, f_i) - Q(\bar{f}_i, f_i)] \\
 &= n^{-1} \sum_{i=1}^n Q(\hat{f}_i, \bar{f}_i) + 2n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} [(\hat{f}_i(\omega) - \bar{f}_i(\omega))][(\hat{f}_i(\omega) - f_i(\omega))] d\omega \\
 &\equiv \hat{Q}_1 + 2\hat{Q}_2
 \end{aligned} \tag{4.1.30}$$

Para  $\hat{Q}_1$  en (4.1.30), por identidad de Parseval (4.1.4), y  $V_{nT} \alpha n2^{J+1}$ , tenemos:

$$\hat{Q}_1 = n^{-1} \sum_{i=1}^n \sum_{j=0}^j \sum_{k=1}^{2^j} (\hat{\alpha}_{ijk}, \bar{\alpha}_{ijk})^2 = O_p[(nT)^{-1} + 2^J/nT] = O_p(2^J/T)$$

Como  $n, T \rightarrow \infty$ . Nuevamente, tenemos  $\hat{Q}_2 = O_p(2^J/T + 2^{-2qj})$  por la inecuación de Cauchy Schwarz,  $\hat{Q}_1 = O_p(2^J/T)$ ,  $n^{-1} \sum_{i=1}^n Q(\bar{f}_i, f_i) = O_p(2^J/T + 2^{-2qj})$ , se sigue por la desigualdad de Markov's,  $n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, f_i) = O_p(2^J/T + 2^{-2qj})$ . Este último se muestra a continuación. Calculamos (b)  $n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, f_i)$ , entonces escribimos:

$$n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, f_i) = n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, E\bar{f}_i) + n^{-1} \sum_{i=1}^n Q(E\bar{f}_i, f_i) \tag{4.1.31}$$

Consideramos en primer lugar el segundo término de (4.1.31). Por la ortonormalidad de  $\{\Psi_{jk}(\cdot)\}$ , obtenemos:

$$n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, f_i) = n^{-1} \sum_{i=1}^n \sum_{i=J+1}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk}^2 + n^{-1} \sum_{i=1}^n \sum_{j=0}^J \sum_{k=1}^{2^j} (E\bar{\alpha}_{ijk} - \alpha_{ijk})^2 \tag{4.1.32}$$

Para el primer término de (4.1.31), utilizando (4.1.29) y (4.1.33):

$$\alpha_{ijk} = (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} R_i(h) \hat{\Psi}_{jk}^*(h), \tag{4.1.33}$$

tenemos:

$$\begin{aligned} \sum_{i=J+1}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk}^2 + n^{-1} &= \sum_{j=J+1}^{\infty} \sum_{h=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} R(h)R(m)c_j(h, m)\hat{\psi}(2\pi h/2^j)\hat{\psi}^*(2\pi m/2^j) \\ &= \sum_{j=J+1}^{\infty} \sum_{h=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} R(h)R(h+2^j r)\hat{\psi}(2\pi h/2^j)\hat{\psi}^*(2\pi h/2^j + 2\pi r) \end{aligned}$$

Ahora se evalúan los términos correspondientes a  $r = 0$  y  $r \neq 0$ , respectivamente. Para  $r = 0$  tenemos:  $\sum_{j=J+1}^{\infty} \sum_{h=-\infty}^{\infty} R_i^2(h) |\hat{\psi}(2\pi h/2^j)|^2$

$$\begin{aligned} &= \sum_{j=J+1}^{\infty} (2\pi/2^j)^q \sum_{h=-\infty}^{\infty} \frac{|\hat{\psi}(2\pi h/2^j)|^2}{|(2\pi h/2^j)|^{2q}} h^2 R_i^2(h) \\ &= \lim_{z \rightarrow \infty} \frac{|\psi(z)|^2}{|z|^{2q}} \sum_{j=J+1}^{\infty} (2\pi/2^j)^q \left[ \sum_{h=-\infty}^{\infty} h^2 R_i^2(h) \right] [1 + o(1)] \\ &= \lambda_q 2^{-2q(J+1)} \int_{-\pi}^{\pi} [f_i^{(q)}(w)]^2 dw [1 + o(1)], \end{aligned}$$

Definiendo a nivel generalizada la derivada espectral de  $f_i(\omega)$ :

$$f_i^{(q)}(\omega) \equiv (2\pi)^{-1} \sum_{h=-\infty}^{\infty} |h|^q R_i(\omega) e^{-ih\omega}, \quad \{\omega \in [-\pi, \pi]. \tag{4.1.34}$$

También definimos una medida  $\lambda_q \in (0, \infty)$  de la suavidad de  $\hat{\psi}(\cdot)$  a 0.

$$\lambda_q \equiv -\frac{(2\pi)^{2q+1}}{1 - 2^{-2q}} \lim_{z \rightarrow 0} \frac{|\hat{\psi}(z)|^2}{|z|^{2q}}. \tag{4.1.35}$$

Donde  $f_i^{(q)}(\cdot)$  y  $\lambda_q$  están definidas en (4.1.34) y (4.1.35), y  $o(1)$  es uniforme en  $i$  y  $\omega \in [-\pi, \pi]$ , para los términos correspondientes a  $r \neq 0$ , se puede demostrar que  $2^{-2q(J+1)}$ . De esto se desprende que, para el primer término de (4.1.32),

$$n^{-1} \sum_{i=1}^n \sum_{j=J+1}^{\infty} \sum_{k=1}^{2^j} \alpha_{ijk}^2 = 2^{-2q(J+1)} \lambda_q^2 n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} [f_i^{(q)}(\omega)]^2 d\omega + o(2^{-2q(J+1)}). \quad (4.1.36)$$

Para el segundo término de (4.1.32), por el Lema (4.1.1, vii) y la suposición 8 (Pág. 1534, [32]). Tenemos:

$$\begin{aligned} & n^{-1} \sum_{i=1}^n \sum_{j=0}^J \sum_{k=1}^{2^j} (E\bar{\alpha}_{ijk} - \alpha_{ijk})^2 \\ &= n^{-1} \sum_{i=1}^n \sum_{j=0}^J \sum_{k=1}^{2^j} [T_i^{-1} \sum_{h=1-T_i}^{T_i-1} |h| R_i(h) \psi_{jk}(2\pi h) + \sum_{|h| \geq T_i} R_i(h) \psi_{jk}(2\pi h)] \\ &\leq 4Cn^{-1} \sum_{i=1}^n T_i^{-2} \sum_{h=1}^{T_i-1} \sum_{m=1}^{T_i-1} |hR_i(h)R_i(m)b_i(h, m)| = O((J+1)/T^2) \end{aligned} \quad (4.1.37)$$

Por último, consideramos que el primer término de (4.1.31), es el factor de variación. Escribimos

$$\begin{aligned} n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, E\bar{f}_i) &= n^{-1} \sum_{i=1}^n \sum_{h=1-T_i}^{T_i-1} \sum_{m=1-T_i}^{T_i-1} b_{J_i}(h, m) \text{cov}[\bar{R}_i(h), \bar{R}_i(m)] \\ &= n^{-1} \sum_{i=1}^n \sum_{h=1-T_i}^{T_i-1} \sum_{m=1-T_i}^{T_i-1} b_{J_i}(h, m) T_i^{-1} \sum_l \left[ 1 - \frac{\eta(l) + m}{T_i} \right] \\ &\quad \times [R_i(l) + R_i(l + m - h) + R_i(l + m)R_i(l - h) + k_i(l, h, m - h)] \\ &\equiv \Omega_{1nT} + \Omega_{2nT} + \Omega_{3nT}, \end{aligned}$$

es decir, donde  $\eta(l) = l$  si  $l > 0$ ,  $\eta(l) = 0$  si  $h - m \leq l \leq 0$ , y  $\eta(l) = -l + h - m$  si  $-(T_i - h) + 1 \leq (l) \leq h - m$ . Cf. Priestley (1981, p. 326). Dada la Suposición 6 (ver Pág. 1531, [32]) y el Lema (4.1.1, vii), tenemos  $|\Omega_{2nT}| \leq C(J+1)$  y  $|\Omega_{3nT}| \leq C(J+1)$ .

Para el primer término  $\Omega_{1nT}$  podemos escribir

$$\begin{aligned}\Omega_{1nT} &= n^{-1} \sum_{i=1}^n \sum_{h=1-T_i}^{T_i-1} b_{J_i}(h, h) T_i^{-1} \sum_l \left(1 - \frac{h}{T_i}\right) R_i^2(l) \\ &+ n^{-1} \sum_{i=1}^n \sum_{h=1-T_i}^{T_i-1} \sum_{|r|=1}^{T_i-1} b_{J_i}(h, h+r) T_i^{-1} \sum_l \left[1 - \frac{\eta(l)+h-r}{T_i}\right] R_i(l) R_i(l+r) \\ &= n^{-1} \sum_{i=1}^n T_i^{-1} (2^{J_i+1} - 1) \sum_{h=-\infty}^{\infty} R_i^2(h) + O[(J+1)/T],\end{aligned}$$

Donde hemos utilizado el Lema (4.1.1, v) para el primer término, lo que corresponde a  $h = m$ ; el segundo término corresponde a  $h \neq m$  y esto es  $O[(J+1)/T]$  uniforme en  $i$  dado  $\sum_{h=-\infty}^{\infty} |R_i(h)| \leq C$  y el Lema (4.1.1, v). Se deduce que  $J \rightarrow \infty$ ,

$$n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, E\bar{f}_i) = \frac{2^{J+1}}{T} n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} f_i^2(\omega) d\omega + O(2^J/T) \quad (4.1.38)$$

Retomando de (4.1.31) hasta (4.1.38) y  $J \rightarrow \infty$ , obtenemos:

$$\begin{aligned}n^{-1} \sum_{i=1}^n EQ(\bar{f}_i, E\bar{f}_i) &= \frac{2^{J+1}}{T} n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} f_i^2(\omega) d\omega + (2^{-2q(J+1)} \lambda_q n^{-1} \sum_{i=1}^n \int_{-\pi}^{\pi} [f_i^{(q)}(\omega)]^2 d\omega \\ &+ o(2^J/T + 2^{-2qJ}).\end{aligned}$$

Esto completa la demostración del teorema 6 (4.1.13).  $\square$

## 4.2. Verificación de las hipótesis

En esta sección, se determinó la veracidad de la confiabilidad del test de heterocedasticidad consistente  $\hat{W}_1$ , probado con datos reales, para ello se programó una función denominada *Wavetest(resid, N, T, J, W)* en Matlab.

En la tabla No.1 (a), (b) y (c) se muestran los valores calculados del test  $\hat{W}_1$  con la función *Wavetest* a partir de los residuales obtenidos mediante el modelo estático en la sección 4.2 de la tesis de maestría de [77].

## 78 Caso de Aplicación

---

El modelo estático obtenido, fue

$$\log(m_{it}) = \beta(t) \log(y_{it}) + \gamma_t + \eta_i + v_{it} \quad (4.2.1)$$

con *coeficiente de escala*  $\beta(t) = \beta_1 + \beta_2 * t + \beta_3 * t^2$ .

Donde:

$\gamma_t$ : Captura el efecto temporal generado por cambios en la tasa de interés agregada sobre la demanda de dinero para la firma  $i$  en el periodo  $t$ .

$\eta_i$ : Efecto firma.

$v_{it}$ : Término de error.

$\lg(m_{it})$ : Logaritmo de los saldos monetarios reales.

$\lg(y_{it})$ : Logaritmo de los ingresos operacionales.

Bajo el modelo (4.2.1) y con ayuda del software Matlab, se calcularon los residuales<sup>7</sup> a partir de una muestra aleatoria de 570 datos. Los wavelets utilizados para dicho cálculo, fueron *Spline de primer orden (S1)* y *segundo orden (S2)*.

*Spline de primer orden (S1)*

$$\hat{\psi}(z) = e^{iz/2} (2\pi)^{-1/2} \frac{\sin^4(z/4)}{(z/4)^2} \left[ \frac{P_3(z/4 + \pi/4)}{P_3(z/2)P_3(z/4)} \right]^{1/2} \quad (4.2.2)$$

y *Spline de segundo orden (S2)*

$$\hat{\psi}(z) = -ie^{i\omega/2} (2\pi)^{-1/2} \frac{\sin^6(z/4)}{(z/4)^3} \left[ \frac{P_5(z/4 + \pi/4)}{P_5(z/2)P_5(z/4)} \right]^{1/2} \quad (4.2.3)$$

donde

$P_3(z) \equiv \frac{2}{3} + \frac{1}{3} \cos(2z)$ , y  $P_5(z) \equiv \frac{1}{30} \cos^2(2z) + \frac{13}{30} \cos(2z) + \frac{8}{15}$ , con factores de escala  $j = 4, \dots, 10$ . A partir de éstos y usando niveles de significancia de  $\alpha = 5\%$  y de  $\alpha = 10\%$ , se encontró:

- Para el método de efectos fijos (ver tabla No. 1.(a)), el spline de primer orden o wavelet Franklin, bajo los factores de escala y niveles

---

<sup>7</sup>Al calcular el test de Hausman, a partir de éstos, encontraron que el modelo a seguir era el de efectos fijos; dado que  $p = 0,0075$ , con  $\alpha = 0,05$ .



de significancia mencionados, los valores obtenidos con el test  $\hat{W}_1$  caen en zona de aceptación; generando la aceptación de la hipótesis nula,  $\mathbb{H}_0$ . Mientras que para el Spline de segundo orden bajo los factores de escala  $j = 8, 9, 10$ , los valores obtenidos con el test  $\hat{W}_1$  caen en zona de rechazo para ambos niveles de significancia utilizados; por lo que se rechaza la hipótesis nula,  $\mathbb{H}_0$ .

- Para el método de efectos aleatorios (ver tabla No. 1.(b)) bajo los spline de primer y segundo orden, con los valores obtenidos a partir del test  $\hat{W}_1$ , se acepta la hipótesis nula  $\mathbb{H}_0$ , dado que los valores caen en zona de aceptación.
- En el método de efectos totales (ver tabla No. 1.(c)) con los spline de primer y segundo orden, se observa que ambos permiten obtener valores para  $\hat{W}_1$  por debajo de los valores críticos en el nivel de significancia del 10 %, lo que permite aceptar la hipótesis nula  $\mathbb{H}_0$ ; aunque con el spline de segundo orden para un valor de  $j = 10$  el valor queda sobre el valor crítico superior, se asume que dicha situación no afecta la aceptación de  $\mathbb{H}_0$ .

| Efectos             | (a) Fijos   |             | (b) Aleatorio |             | (c) Total   |             |
|---------------------|-------------|-------------|---------------|-------------|-------------|-------------|
|                     | S1          | S2          | S1            | S2          | S1          | S2          |
| Factor de escala, j | $\hat{W}_1$ | $\hat{W}_1$ | $\hat{W}_1$   | $\hat{W}_1$ | $\hat{W}_1$ | $\hat{W}_1$ |
| 4                   | 0.51        | 1.46        | 0.32          | 1.31        | 1.64        | 0.62        |
| 5                   | 1.33        | 1.34        | 0.97          | 1.09        | 1.53        | 0.91        |
| 6                   | 0.97        | 1.19        | 1.43          | 1.27        | 1.48        | 0.73        |
| 7                   | 1.12        | 1.61        | 1.10          | 1.13        | 1.14        | 1.01        |
| 8                   | 1.51        | 2.52        | 1.24          | 1.12        | 1.53        | 0.52        |
| 9                   | 1.36        | 3.14        | 1.12          | 1.25        | 1.04        | 1.53        |
| 10                  | 0.84        | 2.38        | 1.27          | 1.07        | 1.02        | 1.65        |

Tabla No.1 (a) Test calculado  $\hat{W}_1$  con Matlab

El filtro para la función de escala Daubechies de orden  $p$  tiene la expresión,

$$\hat{h}(\omega) = \sqrt{2} \left( \frac{1+e^{-i\omega}}{2} \right)^p R(e^{-i\omega})$$

con  $p \geq 1$  y  $R(e^{-i\omega})$  un polinomio trigonométrico [40].

## 80 Caso de Aplicación

Utilizando la función de escala madre de los Daubechies de orden 4 (*db4*) y de orden 6 (*db6*), definida por los coeficientes filtros (ver Tabla No. 2), con  $p = 2$  se tiene:

|   | Coeficientes Daubechies 4 ( <i>db4</i> ) | Coeficientes Daubechies 6 ( <i>db6</i> ) |
|---|--|--|
| n | $h_p(n)(p = 3)$                          | $h_p(n)(p = 2)$                          |
| 0 | 0.4829629131445341                       | 0.235233603892                           |
| 1 | 0.8365163037378079                       | 0.570558457915                           |
| 2 | 0.2241438680420134                       | 0.325182500262                           |
| 3 | -0.1294095225512604                      | -0.095467207783                          |
| 4 |  | -0.060416104155                          |
| 5 |  | 0.024908749865                           |

Tabla No. 2 *db4* y *db6*

En la Tabla No.4 (a,b,c) se muestran los valores obtenidos con los Wavelet Daubechies 4 (*db4*) y los Daubechies 6 (*db6*), encontrando:

- Con el método de efecto fijos, efectos aleatorios y efectos totales se acepta la hipótesis Nula  $\mathbb{H}_0$  con  $\hat{W}_1$ , tanto para los Daubechies 4 (*db4*) y los Daubechies 6 (*db6*), para toda  $j = 1, 2, 3, \dots, 10$

| Efectos             | (a) Fijos   |             | (b) Aleatorio |             | (c) Total   |             |
|---------------------|-------------|-------------|---------------|-------------|-------------|-------------|
| Wavelet             | db4         | db6         | db4           | db6         | db4         | db6         |
| Factor de escala, j | $\hat{W}_1$ | $\hat{W}_1$ | $\hat{W}_1$   | $\hat{W}_1$ | $\hat{W}_1$ | $\hat{W}_1$ |
| 0                   | 0.67        | 0.67        | 0.53          | 0.53        | 0.67        | 0.67        |
| 1                   | 1.23        | 1.23        | 0.65          | 0.65        | 0.81        | 0.81        |
| 2                   | 0.42        | 0.42        | 0.71          | 0.71        | 0.88        | 0.88        |
| 3                   | 1.55        | 1.55        | 0.76          | 0.76        | 0.96        | 0.96        |
| 4                   | 1.66        | 1.66        | 0.81          | 0.81        | 1.08        | 1.08        |
| 5                   | 0.72        | 0.72        | 0.87          | 0.87        | 1.14        | 1.14        |
| 6                   | 0.89        | 0.89        | 0.84          | 0.84        | 1.17        | 1.17        |
| 7                   | 0.97        | 0.97        | 0.92          | 0.92        | 1.19        | 1.19        |
| 8                   | 1.01        | 1.01        | 0.97          | 0.97        | 1.23        | 1.23        |
| 9                   | 0.87        | 0.87        | 1.01          | 1.01        | 1.27        | 1.27        |
| 10                  | 0.78        | 0.78        | 1.17          | 1.17        | 1.32        | 1.32        |

Tabla No.4 (a) Test calculado  $\hat{W}_1$  con Matlab

### 4.3. Contraste de Resultados

En esta sección se comparan los resultados obtenidos por [42] y los obtenidos en esta modelación, teniendo en cuenta las hipótesis planteadas:

$$H_0 : cov(\epsilon_{it}, \epsilon_{it-|h|}) = 0, \text{ para todo } h \neq 0 \text{ y todo } i. \quad (4.3.1)$$

$$H_a : cov(\epsilon_{it}, \epsilon_{it-|h|}) \neq 0, \text{ al menos para } h \neq 0 \text{ y algún } i. \quad (4.3.2)$$

Para Spline de primer orden (S1) y segundo orden (S2), se encontró:

| (a) Efectos Fijos para el test $\hat{W}_1$ |      |      |                        |         |        |
|--|------|------|------------------------|---------|--------|
| Datos Tovar y Montilla                     |      |      | Datos Martinez y Villa |         |        |
| Factor de escala, j                        | S1   | S2   | Factor de escala, j    | S1      | S2     |
| 4  | 0.51 | 1.46 | 4                      | 60587.0 | 1371.4 |
| 5  | 1.33 | 1.34 | 5                      | 7594.9  | 429.3  |
| 6  | 0.97 | 1.19 | 6                      | 2478.3  | 253.1  |
| 7  | 1.12 | 1.61 | 7                      | 1367.0  | 237.8  |
| 8  | 1.51 | 2.52 | 8                      | 1065.6  | 266.5  |
| 9  | 1.36 | 3.14 | 9                      | 998.3   | 321.0  |
| 10   | 0.84 | 2.38 | 10                     | 1039    | 400.3  |

Tabla No. 5 (a) Contraste de Resultados calculados para el test  $\hat{W}_1$  con MatLab.

| (b) Efectos Aleatorios para el test $\hat{W}_1$ |      |      |                        |      |      |
|---|------|------|------------------------|------|------|
| Datos Tovar y Montilla                          |      |      | Datos Martinez y Villa |      |      |
| Factor de escala, j                             | S1   | S2   | Factor de escala, j    | S1   | S2   |
| 4   | 0.32 | 1.31 | 4                      | 0.72 | 1.01 |
| 5   | 0.97 | 1.09 | 5                      | 1.17 | 1.39 |
| 6   | 1.43 | 1.27 | 6                      | 1.36 | 1.17 |
| 7   | 1.10 | 1.13 | 7                      | 1.22 | 1.08 |
| 8   | 1.24 | 1.12 | 8                      | 1.14 | 1.04 |
| 9   | 1.12 | 1.25 | 9                      | 1.10 | 1.03 |
| 10  | 1.27 | 1.07 | 10                     | 1.07 | 1.02 |

Tabla No. 5 (b) Contraste de Resultados calculados para el test  $\hat{W}_1$  con MatLab.

| (c) Efectos Totales para el test $\hat{W}_1$ |      |      |                        |       |       |
|--|------|------|------------------------|-------|-------|
| Datos Tovar y Montilla                       |      |      | Datos Martinez y Villa |       |       |
| Factor de escala, j                          | S1   | S2   | Factor de escala, j    | S1    | S2    |
| 4  | 1.64 | 0.62 | 4                      | 0.953 | 1.325 |
| 5  | 1.53 | 0.91 | 5                      | 1.539 | 1.839 |
| 6  | 1.48 | 0.73 | 6                      | 1.785 | 1.549 |
| 7  | 1.14 | 1.01 | 7                      | 1.613 | 1.428 |
| 8  | 1.53 | 0.52 | 8                      | 1.507 | 1.380 |
| 9  | 1.04 | 1.53 | 9                      | 1.446 | 1.357 |
| 10   | 1.02 | 1.65 | 10                     | 1.409 | 1.345 |

Tabla No. 5 (c) Contraste de Resultados calculados para el test  $\hat{W}_1$  con MatLab.

Tomando como referencia los datos obtenidos en las las Tablas No. 5 (a) Efectos Fijos, Tabla No. 5 (b) Efectos Aleatorios y Tabla No.5 (c) Efectos Totales, se determino:

- En nuestra tesis de maestría para los Spline de primer orden ( $S1$ ) y los Spline de segundo orden ( $S2$ ), con los métodos de efectos fijos, aleatorios y totales, bajo los factores de escala y niveles de significancia mencionados, los valores obtenidos con el test  $\hat{W}_1$  caen en zona de aceptación, por tanto se acepta la hipótesis nula  $H_0$ .
- También, los valores calculados con los Wavelets Daubechies 4 ( $db4$ ) y los Daubechies 6 ( $db6$ ), se encontró que con los métodos efectos fijos, efectos aleatorios y efectos totales se acepta la hipótesis nula  $H_0$ , para toda  $j = 1, 2, 3, \dots, 10$ , en el test  $\hat{W}_1$ .
- Mientras que en la tesis de [42], los valores que obtuvieron para los Spline de primer orden ( $S1$ ) y los Spline de segundo orden ( $S2$ ), aplicando el método de efectos fijos, estos caen en zona de rechazo generando el rechazo de la hipótesis nula  $H_0$ . Para efectos aleatorios, se acepta la hipótesis nula  $H_0$ . Para efectos totales, se acepta hipótesis nula  $H_0$  con  $\alpha = 5\%$ , mientras que con  $\alpha = 10\%$  se rechaza para  $j = 6$  para el spline de primer orden  $S1$  y para  $j = 5$  con el Spline de segundo orden, calculados en el test  $\hat{W}_1$ .

## 4.4. Conclusiones

En econometría, el trabajo con panel de datos, consiste en poder determinar si la información (datos de corte transversal y las series de tiempo) está serialmente correlacionada, es un aspecto de suma importancia. El presente trabajo de investigación se fundamentó en dicha premisa. Aprovechando el test  $\hat{W}_1$  elaborado por [32], el modelo planteado por [77], y los resultados de la investigación de [42], se buscó determinar la confiabilidad del test.

El estudio detallado y minucioso de la demostración del test  $\hat{W}_1$  permitió comprobar que su estructura es sólida dado que garantiza que el factor de traslación y los rezagos están acotados, que los coeficientes wavelet y la varianza convergen en probabilidad, adicional la varianza converge en distribución Normal lo que facilita la toma de decisiones frente a la hipótesis nula ( $\mathbb{H}_0$ ), otro aspecto fundamental es que al garantizar la convergencia de los coeficientes ( $a_i$ ) del espectro de potencia ( $f$ ), se garantiza que la forma cuadrática ( $Q(f, f_0)$ ) es positiva, demostrando de esta manera que el test si captura la correlación serial.

De esta forma, se concluye que el test  $\hat{W}_1$  es confiable en estructura.

Otro elemento importante para garantizar la utilidad del test  $\hat{W}_1$  fue ratificar, a partir de datos reales, los resultados obtenidos por [32] mediante la simulación Monte Carlo. Dado que [32] plantean que los test  $\hat{W}_1$  y  $\hat{W}_2$  solo necesitan los residuales y los resultados de [42] no ratificaron dicha situación, se agregaron más parámetros o argumentos a la modelación en Matlab. En este caso y de acuerdo a la estructura, éstos permitieron tener valores mucho más acordes a lo esperado.

Se considera que la incorporación del número de categorías o empresas ( $N$ ), número de datos por categoría ( $T$ ), factor de escala ( $j$ ) permitieron que los valores obtenidos mediante la función *wavetest* analizara la presencia o no de correlación serial y a partir de ella se aceptara o rechazara la hipótesis nula.

Por lo expuesto con anterioridad, se puede concluir que el test  $\hat{W}_1$  es confiable, i.e., el test si captura la existencia o no de correlación serial, siendo el test no sensible al cambio de la función wavelet utilizada al interior del test.

El lector interesado en la función *wavetest* comunicarse con los autores.<sup>8</sup>

---

<sup>8</sup>Contactar a los e-mail: monymontilla@gmail.com y rito700715@hotmail.com



## CAPÍTULO 5

---

### Notación

---

- $a$  parámetro de dilatación
- $\otimes$  producto tensorial
- $a_i$  coeficientes del desplazamiento  $\tilde{u}$
- $b$  parámetro de translación
- $b(\cdot, \cdot)$  forma bilineal
- $b_i$  coeficientes del desplazamiento  $\tilde{\omega}$
- $b_{j,k}$  coeficientes de  $f_j$  en  $V_j$
- $\beta_i$  Coeficiente de regresión
- $\hat{\beta}_i$  Coeficiente parcial (estimado) de la regresora  $X_i$
- $C_\psi$  constante de normalización de la condición de admisibilidad
- $c_{j,k}$  coeficientes wavelets de  $\hat{u}_{j+1}$
- $c_k$  coeficientes de las funciones de escala
- $f$  término independiente del problema diferencial
- $F$  funcional correspondiente a la formulación variacional
- $F_X$  función de distribución
- $h_j$  filtro wavelets de los Daubechies
- $h = 2^j$  espaciado entre puntos diádicos de la escala  $j$
- $h_k$  coeficientes del filtro de la función de escala
- $h_j, k$  funciones de escala sistema de Haar
- $\tilde{h}$  vector filtro wavelet  $\{h_n, n \in Z\}$   $H$  espacio de Hilbert
- $H(f)$  función de frecuencia de respuesta

- $I$  momento de inercia  
 $I_0$  momento de inercia  
 $j$  escala  
 $L$  operador diferencial  
 $L(\omega)$  polinomio trigonométrico  
 $m + 1$  orden de B-spline, splines polinomiales de grado  $m$   
 $N$  orden de la función de escala de Daubechies ( $N + 1$  coeficientes)  
 $O$  y  $o$  Notación de Landau, comparación asintótica de funciones. permite establecer la cota inferior asintótica, la cota superior asintótica y la cota ajustada asintótica.  
 $p(x)$  coeficiente en el problema de valores de contorno  
 $p_k$  coeficientes de la función de escala como combinación lineal de B-splines  
 $\tilde{p}_k$  coeficientes de la función de escala de Daubechies  
 $T$  matriz de la transformación entre el espacio físico y el espacio wavelet  
 $T_c$  operador de translación  
 $\tilde{u}$  desplazamientos en la dirección  $x$   
 $\hat{V}_j^I$  Subespacio de B-Splines interiores  
 $U$  vector de desplazamientos Sección  
 $\hat{v}_j$  incremento en el espacio  $\hat{V}_{j+1}$   
 $\tilde{v}$  desplazamientos en la dirección  $y$   
 $V$  subespacio de  $H$  donde se plantea el problema variacional  
 $V_j$  subespacios cerrados de  $L^2(R)$  que constituyen un AMR  
 $\hat{w}$  grados de libertad en el espacio físico  
 $w_i$  desplazamiento del nodo  $i$  en la dirección transversal  
 $w_k$  coeficientes de la wavelet como combinación lineal de B-splines  
 $\tilde{w}$  desplazamientos en la dirección  $z$   
 $\hat{w}$  desplazamientos transversales en el espacio físico  
 $W_j$  complemento ortonormal de  $V_j$  en  $V_{j+1}$ ,  $W_j \oplus V_j = V_{j+1}$   
 $\hat{W}_j^I$  complemento ortonormal de  $\hat{V}_j^I$  en  $\hat{V}_{j+1}^I$   
 $\alpha$  coeficientes wavelet  
 $\hat{\alpha}$  grados de libertad en el espacio wavelet  
 $\alpha_k$  coeficientes wavelets del desplazamiento  $w$   
 $\alpha_j$  coeficientes wavelets del sistema  
 $\alpha_w$  es el vector de coeficientes wavelets correspondiente a  $\hat{w}$   
 $\alpha_{j,k}(t)$  coeficientes donde  $t > 0$   
 $\tilde{\alpha}_{j,k}$  coeficientes de  $\hat{u}_j$   
 $\hat{\alpha}_j$  coeficientes de  $\hat{u}_j$  solución de Galerkin Modificado



- 
- $[\hat{\alpha}_{j+1}]$  coeficientes de  $\hat{u}_j$  en las funciones de escala de  $\hat{V}_{j+1}$   
 $\hat{\beta}^{\gamma-1}(\omega)$  transformada de Fourier de una B-spline fraccional  
 $\delta$  variación de un operador  
 $\delta_{m,n}$  delta de kroneker  
 $\eta$  parámetro de frecuencia  
 $\varepsilon$  vector de deformaciones de componentes  $\varepsilon_x$  y  $\kappa$   
 $\epsilon$  umbral del criterio de corte en el Algoritmo  
 $\phi(x)$  función de escala madre en un análisis de multiresolución  
 $\phi_N$  función de escala de Daubechies de orden  $N$   
 $\phi_{j,k}(x)$  funciones generadas componiendo traslaciones y dilataciones de  $\phi(x)$   
 $\hat{\phi}(\omega)$  transformada de Fourier de  $\phi$ , definida por  $\int_{-\infty}^{\infty} \phi(t)e^{-2\pi it} dt$   
 $\hat{\phi}_{m+1}(\omega)$  transformada de Fourier de  $\phi_{m+1}$   
 $\hat{\varphi}_0(\omega)$  función de  $\omega$  acotada  
 $\Lambda$  constante de continuidad de  $F$   
 $\Gamma$  frontera del dominio de definición del operador  $L$   
 $\nu$  coeficiente de Poisson  
 $\pi u$  el elemento del espacio  $V_h$  que interpola a  $u$  en los nodos  
 $\pi_r$  polinomios de grado  $r$   
 $\rho$  constante de coercividad de la forma bilineal  $a$   
 $\omega$  variable de frecuencia  
 $\xi$  es la coordenada local,  $0 \leq \xi \leq 1$   
 $\psi(x)$  wavelet madre  
 $\psi_N(x)$  wavelet de orden  $N$  en el AMR Daubechies  
 $\psi_{j,k}(x)$  funciones generadas componiendo traslaciones y dilataciones de  $\psi(x)$   
 $\Omega$  dominio de definición del operador  $L$   
 $\bigoplus$  suma directa de los subespacios  
 $\stackrel{d}{=}$  significa que los términos en comparación tienen la misma distribución de probabilidad



---

## Bibliografía

---

- [1] Aldroubi, A. The wavelet transform: A surfing guide, pp 3-36 in *Wavelets in Medicine and Biology*, A. Audroubi, M. Unser (eds.), CRC Press, New York, 1996.
- [2] Arellano, M. and Bover O., *La Econometría Datos de Panel*, Investigaciones Económicas(Segunda época). Vol XIV, no. 1, 1990, páginas 3-45.
- [3] Arellano, M. *Panel Data Econometrics*. Oxford University Press, New York, 2003
- [4] Asmar, Ch. Abraham. *Tópicos en teoría de matrices*, Primera edición, Medellín, 1995.
- [5] Baltagi, B., *Econometrics Analsis of Panel Data*, 2<sup>nd</sup> edition. Wiley, 2001.
- [6] Billingsley, P. *Probability and Measure*, 3<sup>rd</sup> edition. John Wiley & Sons, Inc. New York, 1995.
- [7] Boggess, A., Narcowich, F. J. *A First Course in Wavelets with Fourier Analysis*. Prentice Hall, New Jersey, 2001.
- [8] Bonani, A., Durand, S. and Weiss, G. *Wavelets obtained by continuous deformations of the Haar wavelet*, *Revista Mat. Iberoamericana*, Vol. 12, No 1, 1996.

- 
- [9] Burdisso, T., *Estimación de una función de costos para los bancos privados Argentinos utilizando datos de panel*. Banco Central de la Republica de Argentina, documento de trabajo N° 3, 1997.
- [10] Burrus, C. S., Gopinath, R. A., Guo, H. *Introduction to Wavelets and Wavelets Transforms A Primer*. Prentice Hall, New Jersey, 1998.
- [11] Cobacho, T. Ma. Belén and Bosch M. Mariano. *Contrastes de Hipótesis de Datos en Panel*, XIII Jornadas de ASEPUMA.
- [12] Chui, C. K. *An Introduction to Wavelets*. Academic Press, Boston, 1992.
- [13] Chui, C. K. *Wavelets: A Mathematical Tool for Signal Analysis*. SIAM Monographs on Mathematical Modeling and Computation, Philadelphia, 1997.
- [14] Chung, K. L. *Elementary Probability Theory with Stochastic Processes*. Springer-Verlag. New York, 1979.
- [15] Clarke, R. J. *Transform Coding of Images*. Academic Press, San Diego, CA, 1985.
- [16] Clarke, R. J. *Digital Compression of Still Images and Video*. Academic Press, San Diego, CA, 1995.
- [17] Cryer, J. D. *Time Series Analysis*, Duxbury Press. Boston, 1986.
- [18] Daubechies, I. Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, 41 1988, 909-996.
- [19] Daubechies, I. *Ten Lectures on Wavelets*, CBMS Series 61, SIAM, Philadelphia 1992.
- [20] Daubechies, I. The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inform. Theory*, vol. 36 1990, 961-1005.
- [21] Delyon, B. and Juditsky, A. Estimating wavelets coefficients, pp 151-168 in *Wavelets and Statistics*, Springer-Verlag, Tokyo, 1995.
- [22] Donoho, D. *Nonlinear Wavelet Methods for Recovery of Signals, Densities and Spectra from Indirect and Noise Data*. "Different Perspectives on Wavelets", Proceedings of Symposia in Pure Math., AMS., I Daubechies, Edt., 47, 173-205, 1993.

- 
- [23] Fan, Y. and Gençay, R. *Unit Root Tests with Wavelets*. Revista Econometric Theory, Vol. 26, pp. 1305-1331, 2010.
- [24] Folland, G. B. *Real Analysis*, 2<sup>nd</sup> edition. John Wiley & Sons, Inc. New York, 1999.
- [25] Gabor, D. *Theory of communications*, J. Inst. Elect. Eng. London. 93(III) 1946, 429-457.
- [26] Gençay, R. y otros. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press. 2002.
- [27] Gersho, A. and Gray R. M. *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992.
- [28] Gonzalez, R. C. and Woods, R. E. *Digital Image Processing*. Addison-Wesley, New York, 1992.
- [29] Greene, W.H. *Análisis Económico*. Quinta Edición, Prentice Hall, 1998.
- [30] Hamilton, J. D. *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, 1994.
- [31] Hernández, E. Weiss, G. *A First Course on Wavelets*. CRC Press, Boca Raton, FL, 1996.
- [32] Hong, Y. and Kao, Ch. *Wavelet-based testing for serial correlation of unknown form in panel models*. Econometrica, Vol. 72, N° 5 (Septiembre, 2004), Pág. 1519-1563.
- [33] Hsiao, Ch. *Analysis of Panel Data*. Econometric Society Monographs. Cambridge University Press, 1986.
- [34] Kikut, O. *Uso de datos de tiempo y sección cruzada*. Monografía, Febrero, 1999.
- [35] Irino, T., Kawahara, H. Signal reconstruction from modified auditory wavelet transform, *IEEE Trans. Signal Process.*, vol 41, 1993, 3549-3553.
- [36] Lee, J. and Hong. *Testing for Serial Correlation of Unknown form using Wavelets Methods*. Econometrics Theory, 2001, Pág. 17, 386-423.

- 
- [37] Li, Yushu. and N.G. Andersson, Fredrik *A simple wavelet-based test for serial correlation in panel data models*. Discussion paper, INSTITUTT FOR FORETAKSØKONOMI, March 2014.
- [38] Mallat, S. Multiresolution approximations and wavelet orthonormal bases for  $L^2(\mathbb{R}^d)$ , *Trans. of Amer. Math. Soc.* 315, 1989, 69-87.
- [39] Mallat, S. A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, 1989, 674-693.
- [40] Mallat, S., *textita wavelet tour of signal processing*, second edition, Academic Press, 1999.
- [41] Martin, G. *Introducción a la econometría*. Prentice Hall, 1997.
- [42] Martínez, J. y Villa, A. *Correlación serial con Wavelet*. Universidad EAFIT, Medellín, Septiembre de 2007.
- [43] MATLAB, *El lenguaje de cálculo técnico*. MathWorks, España. <https://www.es.mathworks.com>. 2014.
- [44] Mulligan, C.B. Scale Economies, the Value of Time, and the Demand for Money: Longitudinal Evidence from Firm. *Journal of Political Economy*, 105, 1061-1079, 1997
- [45] Mulligan, C.B. The Demand for Money by Firms: Some Additional Empirical Results, Discussion Paper 125, Federal Reserve Bank of Minneapolis, 1-52, 1997.
- [46] Meyer, Y. *Ondelettes et opérateurs, I: Ondelettes*. Herman, Paris, 1990.
- [47] Mulligan, C.B. *Scale Economies, the Value of Time, and the Demand for Money: Longitudinal Evidence from Firms*, *Journal of Political Economy*, 105, 1061-1079, 1997.
- [48] Mulligan, C.B. *The Demand for Money by Firms: Some Additional Empirical Results*, Discussion Paper 125, Federal Reserve Bank of Minneapolis, 1-52, 1997.

- 
- [49] Newey, W. y D.McFadden. Large Sample Estimation and Hypothesis Testing, en Engle y McFadden(eds) Handbook of Econometrics, IV, North-Holland, 1994.
- [50] Oppenheim, A., Shafer, R. *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [51] Percival, D. B., Walde, A. T. *Wavelet Methods for Time Series Analysis*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2000.
- [52] Pindyck, R. and Rubinfeld, D. *Econometric models and Economics forecasts*. McGraw-Hill International Edition, Fourth Edition, 1993.
- [53] Pinsky, M. A. *Introduction to Fourier Analysis and wavelet*, Brooks/Cole, NJ, 2001.
- [54] Proakis, J. G., Manolakis, D. G. *Signal Processing: Principles, Algorithms and Applications*, Macmilan, New York, 1992.
- [55] Quak, E. and Weyrich, N. Decomposition and reconstruction algorithms for spline wavelet on a bounded interval, *Appl. and Comp. Harmonic Anal.* (ACHA), 1 1994, 217-231.
- [56] Rioul, O., Duhamel, P. Fast algorithms for discrete and continuous wavelet transform, *IEEE Trans. Inform. Theory*, vol. IT-38 1992, 569-586.
- [57] Ross, S. M. *Introduction to Probability Models*, 8<sup>nd</sup> edition. Academic Press. New York, 2003.
- [58] Shapiro, J. M. Embedded image coding using zerotrees of wavelet coefficients, *IEEE Trans. Signal Proc.* 41 1993, 3445-3462.
- [59] Shen, X. A Galerkin-wavelet method for a singular convolution equation on the real line, *J. Int. Equa. Appl.* 12 2000, 157-176.
- [60] Strang, G. and Nguyen, T. *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Cambridge, MA, 1996.
- [61] Stremmer, F. G. *Introduction to Communication Systems*, 2<sup>nd</sup> edition. Addison-Wesley Publishing Company, Inc. Massachusetts, 1982.

- 
- [62] Sun, W. and Zhou, X. Sampling theorem for wavelet subspaces: error estimate and irregular sampling, *IEEE Trans. Signal Process.*, vol. 48 2000, 223-226.
- [63] Unser, M. and Aldroubi, A. B-spline processing I: Theory and II: Efficient design and application, *IEEE Trans. Signal Process.*, vol 41 1993, 821-848.
- [64] Unser, M. A practical guide to implementation of the wavelet transform, pp 37-76 in *Wavelets in Medicine and Biology*, A. Audroubi, M. Unser (eds.), CRC Press, New York 1996.
- [65] Walnut, D. *An Introduction to Wavelets Analysis*. Birkhäuser, Boston, 2002.
- [66] Walter, G. G. Approximation with impulse trains, *Results Math.* 34 1998, 185-196.
- [67] Walter, G. G. and Shen, X. A substitute for summability in wavelet expansions, *Appl. Numer. Harmon. Anal.*, Birkhäuser Boston, Boston, MA 1999, 51-63.
- [68] Walter, G. G. and Shen, X. Deconvolution using Meyer wavelets, *J. Integral Equations Appl.* 11 1999, 515-534.
- [69] Walter, G. G., Shen, X. *Wavelets and Other Orthogonal Systems*, 2<sup>nd</sup> edition. Chapman & Hall/CRC, Boca Raton, FL, 2001.
- [70] Wayne, A. F. *Introduction to Statistical Time Series*, Second Edition, A Wiley-Interscience Publication, John Wiley and Sons, INC. 1996.
- [71] Wei, W. W. S. *Time Series Analysis*, 2<sup>nd</sup> edition. Addison-Wesley Publishing Company, Inc. Boston, 2006.
- [72] Yamada, K., Ohkitani, K. Orthonormal wavelet expansion and its application to turbulence, *Progr. Theoret. Phys.*, vol. 83-5, 1990, 819-823.
- [73] Yamada, K., Ohkitani, K. An identification of energy cascade in turbulence by orthonormal wavelet analysis, *Progr. Theoret. Phys.*, vol. 84-4, 1991, 799-815.



- 
- [74] Yamada, K., Ohkitani, K. Orthonormal wavelet analysis of turbulence, *Fluid Dynam. Res.*, vol. 8, 1991, 101-115.
- [75] Rosales, R. y Bonilla, J. A. *Introducción a la econometría*. Series apuntes de clase Cede, octubre de 2006.
- [76] Wooldridge, J. *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, 2001.
- [77] Zuluaga, F. I. *Econometría de Datos de Panel: Revisión y una Aplicación*. Universidad Eafit, Medellín, julio de 2005.

---

## Índice alfabético

---

- Asintótica
  - eficiencia y normalidad, 11
- asintótica
  - Distribución, 10
- Autocorrelación, 44
- coeficiente
  - de escala, 78
- Contraste de Significatividad, 38
- Convergencia
  - en Distribución, 10
  - en probabilidad, 9
- convolución, 4
- Efectos Aleatorios, 33, 41
- Efectos Fijos, 32, 36
- Estimador
  - Consistente, 9
  - Intra y entre grupos, 39
- Fourier
  - coeficientes de, 16
  - fórmula de inversión de, 15
  - serie de, 16
  - transformada de, 14
- función
  - característica, 5
  - dilatada, 18
  - trasladada, 18
- funciones
  - cuadrado integrable, 3
  - integrables, 3
  - ortogonales, 3
- Hausman,
  - Test de Especificación de , 33
- Heterocedasticidad, 43
- límite
  - Distribución, 10
- Mínimos cuadrados
  - de variables ficticias, 37
- modelo de componentes
  - del error, 42
- Panel de datos, 32
- Paneles no balanceados, 41
- Parseval
  - fórmula de, 15
- Plancharel
  - fórmula de, 15
- probabilidad , 5

- espacio de, 6
- Regresión Total, 32
- señal, 13
- Teoría
  - asintótica, 9
- variable
  - dummy, ficticia, 37
- wavelet , 17
  - coeficientes, 22
  - fórmula de inversión, 19
  - fórmula de Parseval, 20
  - fórmula de Plancherel, 18
  - serie, 22
  - transformada continua, 18
  - transformada discreta, 22