

**TESIS DOCTORAL**

**AÑO 2015**

**Algunos Desarrollos y Aplicaciones del  
Análisis Simbólico al Estudio de Series  
Temporales**

**Isidro Morales González**

**Ingeniero Superior de Telecomunicaciones**

**DEPARTAMENTO DE ECONOMÍA CUANTITATIVA I**

**FACULTAD DE CIENCIAS ECONOMICAS Y  
EMPRESARIALES**

**UNIVERSIDAD NACIONAL DE EDUCACIÓN A  
DISTANCIA**

**Director : Dr. Mariano Matilla-García**

**Codirector: Dr. Manuel Ruiz Marín**

DEPARTAMENTO DE ECONOMÍA CUANTITATIVA I  
FACULTAD DE CIENCIAS ECONOMICAS Y EMPRESARIALES  
UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

**Algunos Desarrollos y Aplicaciones del Análisis Simbólico al  
Estudio de Series Temporales**

Tesis presentada por

Isidro Morales González

Ingeniero Superior de Telecomunicaciones

Director : Dr. Mariano Matilla-García

Codirector: Dr. Manuel Ruiz Marín

## Dedicatorias

*Quiero dedicar esta Tesis a la memoria de D. Jose Vicente García Sestafe sin cuyo apadrinamiento, sabiduría y estímulo intelectual nunca habría hecho posible esta tesis. También quiero dedicar esta tesis a su mujer Carmen.*

*A Don Julián Rodríguez, por su ayuda y apoyo durante el período de redacción de esta tesis.*

*A mi director de Tesis Mariano Matilla-García por su ayuda incondicional en la elaboración de esta tesis.*

*A mis padres que se han preocupado de mí desde el momento en que llegué a este mundo, me han formado para saber cómo luchar y salir victorioso ante las diversas adversidades de la vida. Muchos años después, sus enseñanzas continúan, y nunca estaré lo suficientemente agradecido a su magisterio de la vida.*

*A mi hermano Alberto.*

*A la memoria de mis abuelos y mi tío Ramón.*

*A mi amigo Benjamín, deseando su pronta recuperación.*

*A todos los que me apoyaron para escribir y concluir esta tesis.*

*Para ellos es esta dedicatoria de tesis, pues es a ellos a quienes se las debo por su apoyo incondicional.*

## **Agradecimientos**

Mi más sincero agradecimiento a mis directores de tesis; al Dr. Mariano Matilla García, por su dirección y comprensión durante el desarrollo de esta Tesis y al Dr. Manuel Ruiz Marín por la revisión, comentarios y sugerencias a este trabajo.

Quiero mostrar también mi agradecimiento a María Jesus Alonso Gil, jefa de mi departamento de Trabajo (Area de Control de Acciones y Núcleo de Telefónica Móviles España), por su concesión del tiempo necesario para realizar las estancias en la UNED, con el fin de llevar a cabo esta Tesis.

A mis amigos Oscar, Mario, Jose Luis y Nuria.

A mis compañeros de trabajo.

Gracias a todos ellos por su apoyo en la realización de esta Tesis.

# Índice general

<b>Dedicatorias</b>	<b>II</b>
<b>Agradecimientos</b>	<b>III</b>
<b>Lista de Figuras</b>	<b>IX</b>
<b>Lista de Tablas</b>	<b>XIII</b>
<b>Lista de Apéndices</b>	<b>XIV</b>
<b>Lista de Términos</b>	<b>XVII</b>
<b>Capítulo 1 Introducción</b>	<b>1</b>
<b>Capítulo 2 Contraste de Heterocedasticidad Condicional de Forma Desconocida en Series Temporales Mediante Dinámica Simbólica</b>	<b>8</b>
2.1 Introducción	8
2.2 Introducción a la Entropía Simbólica	11
2.3 De la Entropía Simbólica a la Entropía de Permutación	12
2.4 De la Entropía Simbólica a la Construcción de Tests No Paramétricos. El Test de Independencia $G(m)$	13
2.4.1 Propiedades del Test $G(m)$	18
2.4.2 El Conjunto de los Indices	19
2.4.3 Estudio Empírico de la Convergencia Asintótica del Test $G(m)$	20

2.5	Marco Estadístico de Definición y Diseño de los Tests	21
2.6	Test Semi-Paramétrico de Heterocedasticidad basado en Permutaciones $G_{h_P}(m)$	23
2.7	Test Semi-Paramétrico de Heterocedasticidad basado en la Mediana $G_{h_{Me}}(m)$	26
2.8	Test Semi-Paramétrico de Heterocedasticidad basado en los N-Cuantiles $G_{h_{NQ}}(m)$	29
2.8.1	Selección práctica del parámetro $m$ y el tamaño de muestra $T$	33
2.9	Pruebas Empíricas	34
2.9.1	Simulaciones	35
2.9.1.1	Tamaño	35
2.9.1.2	Potencia	40
2.9.2	Comparación con Otros Tests de Heterocedasticidad	44
2.10	Relación de los resultados de Tamaño y el solapamiento de las $m - historias$	56
2.11	Conclusiones	63

### **Capítulo 3 Aplicación de la Entropía de Tsallis a las Técnicas de Análisis Simbólico para el Contraste de Heterocedasticidad Condicional de Forma Desconocida en Series Temporales**

3.1	Introducción	65
3.2	Introducción al Formalismo de Tsallis	72
3.2.1	Introducción a la Entropía de Tsallis	72
3.2.2	Introducción a las Matemáticas de Tsallis	73
3.2.3	Introducción a la Estadística de Tsallis	75
3.3	Estadísticas, Entropía de Tsallis y Heterocedasticidad	77
3.3.1	Cuantificación del grado de Dependencia de la Varianza	80
3.4	De la Entropía de Tsallis a la Entropía de Permutación de Tsallis	81
3.5	Marco Estadístico de Definición y Diseño de los Tests	82

3.6	Aplicación del Formalismo de Tsallis al Análisis Simbólico. Los Tests $G_h^q(m)$	83
3.6.1	El Test de Independencia $G^q(m)$	83
3.6.2	Los Tests $G_{h_P}^q(m)$ , $G_{h_{M_e}}^q(m)$ y $G_{h_{N_Q}}^q(m)$	86
3.6.3	Estimación Teórica del Índice Entrópico $q$	88
3.6.4	Estimación Empírica del Índice Entrópico $q$	92
3.7	El Test Bootstrap $T_{h_b}^q(m)$	93
3.7.1	Procedimiento del Test $T_h^q(\tau, m)$ basado en Bootstrap	94
3.7.2	Selección de los parámetros $\tau$ y $q$	97
3.8	Pruebas Empíricas	97
3.8.1	Simulaciones	99
3.8.1.1	Tamaño de los Tests $G_{h_P}^q(m)$ , $G_{h_{M_e}}^q(m)$ y $G_{h_{N_Q}}^q(m)$	99
3.8.1.2	Tamaño del Test $T_h^q(m)$	99
3.8.1.3	Potencia de los Tests $G_{h_P}^q(m)$ , $G_{h_{M_e}}^q(m)$ y $G_{h_{N_Q}}^q(m)$	102
3.8.1.4	Potencia del Test $T_h^q(\tau, m)$	104
3.8.2	Comparación con Otros Tests de Heterocedasticidad	106
3.9	Conclusiones	111

## **Capítulo 4 Selección y Contraste Estadístico de los Parámetros de Inmersión para la Reconstrucción del Espacio de Estados en Series Caóticas Temporales.**

		<b>113</b>
4.1	Introducción	113
4.2	El Método C-C	117
4.2.1	La Integral de Correlación y el Test BDS	117
4.2.2	El Método C-C	119
4.2.3	Desventajas del Método C-C	122
4.2.4	Mejoras del Método C-C	123
4.3	Aplicación de Simbolización basada en las distancias absolutas $f_{r_e}$	125

4.4	Aplicación de Simbolización basada en las distancias relativas	
$f_{r_e}^1$		131
4.5	Comportamiento de la Entropía Simbólica $h_\tau(m)$ con $p$ y $r_e$ para tamaños de muestra finito	136
4.6	Selección del parámetro $r_e$	141
4.7	Análisis de Simulación	143
4.8	Efectos del Ruido en la Precisión (Error Estándar $\sigma$ ) y Exactitud (Sesgo) de los Métodos de Selección de los Parámetros de Inmersión (Retardo y Ventana Temporales) para la Reconstrucción del Espacio de Estados de un Sistema Caótico	149
4.8.1	Efectos del Ruido en la Exactitud (Sesgo)	159
4.8.2	Efectos del Ruido en la Precisión (Error Estándar $\sigma$ ) : Contraste Estadístico de los Parámetros de Inmersión (Retardo y Ventana Temporales) para la Reconstrucción del Espacio de Estados de un Sistema Caótico	160
4.8.2.1	Marco Estadístico de Definición y Diseño del Test	161
4.8.2.2	Procedimiento del Test Bootstrap	167
4.8.3	Pruebas Empíricas	168
4.8.3.1	Pruebas de Tamaño	170
4.8.3.2	Pruebas de Potencia	171
4.9	Conclusiones	175
	<b>Capítulo 5 Conclusiones y Líneas Futuras de Investigación</b>	<b>177</b>
5.1	Conclusiones	177
5.2	Futuras líneas de Investigación	180
	<b>Bibliografía</b>	<b>184</b>

## Índice de figuras

Figura N°	Título	Página
4.1	$h_\tau(m)/m$ ( $m = 6$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ , $c = 45.92$ y $\tau_s = 0.01$ para la Aplicación de Simbolización $f_{re}$ .	138
4.2	$h_\tau(m)/m$ ( $m = 5$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ , $c = 45.92$ y $\tau_s = 0.01$ para la aplicación Simbólica $f_{re}$ .	139
4.3	$h_\tau(m)/m$ ( $m = 4$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ , $c = 45.92$ y $\tau_s = 0.01$ para la aplicación Simbólica $f_{re}$ .	139
4.4	$h_\tau(m)$ ( $m = 6$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ , $c = 45.92$ y $\tau_s = 0.01$ para la aplicación Simbólica $f_{re}^1$ .	140
4.5	$h_\tau(m)/m$ ( $m = 5$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ y $c = 45.92$ , $\tau_s = 0.01$ para la aplicación Simbólica $f_{re}^1$ .	140
4.6	$h_\tau(m)/m$ ( $m = 4$ ) para la variable $x$ del sistema de Lorenz de la ecuación 4.31 con $a = 16.0$ , $b = 4.0$ , $c = 45.92$ , $\tau_s = 0.01$ para la aplicación Simbólica $f_{re}^1$ .	141
4.7	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}$ .	149

4.8	$h_\tau(m)/m$ con $m = 6$ y $R = 0,4$ para la serie $\{X_t\}$ del Sistema de Rossler para la Aplicación $f_{re}$ .	150
4.9	$h_\tau(m)/m$ con $m = 6$ y $R = 0,275$ para la serie $\{X_t\}$ del Sistema de Duffing para la Aplicación $f_{re}$ .	150
4.10	$h_\tau(m)/m$ con $m = 6$ y $R = 0,25$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}^1$ .	151
4.11	$h_\tau(m)/m$ con $m = 6$ y $R = 0,1$ para la serie $\{X_t\}$ del Sistema de Rossler para la Aplicación $f_{re}^1$ .	151
4.12	$h_\tau(m)/m$ y $m = 6$ con $R = 0,25$ para la serie $\{X_t\}$ del Sistema de Duffing para la Aplicación $f_{re}^1$ .	153
4.13	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}$ . Factor de ruido $\eta = 10\%$ .	155
4.14	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}$ . Factor de ruido $\eta = 30\%$ .	156
4.15	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}$ . Factor de ruido $\eta = 50\%$ .	156
4.16	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}$ . Factor de ruido $\eta = 100\%$ .	157
4.17	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}^1$ . Factor de ruido $\eta = 10\%$ .	157
4.18	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}^1$ . Factor de ruido $\eta = 30\%$ .	158
4.19	Entropía Simbólica por símbolo $h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}^1$ . Factor de ruido $\eta = 50\%$ .	158
4.20	$h_\tau(m)/m$ con $m = 6$ y $R = 1,2$ para la serie $\{X_t\}$ del Sistema de Lorenz para la Aplicación $f_{re}^1$ . Factor de ruido $\eta = 100\%$ .	159

## Índice de tablas

Tabla N <sup>o</sup>	Título	Página
2.1	Valores de algunos Estadísticos para $\chi^2(5)$ y $\hat{G}(3)$	21
2.2	<i>Ejemplo de m-historias y Símbolos</i>	31
2.3	<i>Símbolos y Probabilidades de Símbolo 2.40.</i>	32
2.4	Modelos bajo la Hipótesis Nula	36
2.5	Tamaño Empírico de $G_{h_P}(m)$	37
2.6	Tamaño Empírico de $G_{h_{M_e}}(m)$	38
2.7	Tamaño Empírico de $G_{h_{N_Q}}(m)$	39
2.8	Modelos bajo la Hipótesis Alternativa	43
2.9	Potencia Empírica de $G_{h_P}(m)$	45
2.10	Potencia Empírica de $G_{h_{M_e}}(m)$	46
2.11	Potencia Empírica de $G_{h_{N_Q}}(m)$	47
2.12	Tamaño Empírico de $G_{h_{N_Q}}(4)$ y varios Tests	52
2.13	Potencia Empírica de $G_{h_{N_Q}}(4)$ y Varios Tests con Modelos X-ARCH	53
2.14	Potencia Empírica de $G_{h_{N_Q}}(4)$ y varios Tests con Modelos SV	55
2.15	Potencia Empírica de $G_{h_{N_Q}}(4)$ y varios Tests con otras estructuras ARCH no lineales	56
2.16	Número de Observaciones con solapamiento ( $k = 1$ ) y sin solapamiento ( $k \geq m$ ) de <i>m-historias</i> para los tres tests	58
2.17	Número de Observaciones con solapamiento ( $k = 1$ ) y sin solapamiento ( $k \geq m$ ) de <i>m-historias</i> para $G_{h_P}(m)$ y $G_{h_{M_e}}(m)$ (1).	58

2.18	Tamaño Empírico de $G_{h_P}(m)$ con y sin solapamiento de símbolo $m$	60
2.19	Tamaño Empírico de $G_{h_{Me}}(m)$ con y sin solapamiento de símbolo $m$	61
2.20	Tamaño Empírico de $G_{h_{NQ}}(m)$ con y sin solapamiento de símbolo $m$	62
3.1	Estimación del Índice Entrópico $q$ bajo $\mathbb{H}_0$	90
3.2	Estimación del Índice Entrópico $q$ bajo $\mathbb{H}_1$	91
3.3	Tamaño Empírico de $T_h^q(\tau, 4)$ , $G_{h_P}(4)$ , $G_{h_{Me}}(4)$ y $G_{h_{NQ}}(4)$ para $T=500$	100
3.4	Tamaño Empírico de $T_h^q(\tau, 4)$ para $T=1000$	101
3.5	Potencia Empírica de los tests $G_{h_i}(4)$ y los tests $G_{h_i}^q(4)$	103
3.6	Potencia Empírica de $T_h^q(\tau, 4)$	105
3.7	Tamaño Empírico de $T_h^{1,5}(2, 4)$ y varios Tests	107
3.8	Potencia empírica de $T_h^{1,5}(2, 4)$ y Varios Tests con Modelos $X-ARCH$	108
3.9	Potencia empírica de $T_{h_b}(4)$ y varios Tests con Modelos SV	109
3.10	Potencia Empírica de $T_h^{1,5}(2, 4)$ y varios Tests con otras estructuras ARCH no lineales	110
4.1	<i>m-historias y Símbolos para <math>\tau = 1</math> de 4.29</i>	128
4.2	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 1</math></i>	129
4.3	<i>m-historias y Símbolos para <math>\tau = 2</math> de 4.29</i>	129
4.4	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 2</math></i>	129
4.5	<i>m-historias y Símbolos para <math>\tau = 3</math> de 4.29</i>	129
4.6	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 3</math></i>	129
4.7	<i>m-historias y Símbolos para <math>\tau = 4</math> de 4.29</i>	130
4.8	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 4</math></i>	130
4.9	<i>m-historias y Símbolos para <math>\tau = 5</math> de 4.29</i>	130
4.10	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 5</math></i>	130
4.11	<i>m-historias y Símbolos para <math>\tau = 6</math> de 4.29</i>	130
4.12	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 6</math></i>	130
4.13	<i>m-historias y Símbolos para <math>\tau = 1</math> de 4.29</i>	133

4.14	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 1</math></i>	134
4.15	<i>m-historias y Símbolos para <math>\tau = 2</math> de 4.29</i>	134
4.16	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 2</math></i>	134
4.17	<i>m-historias y Símbolos para <math>\tau = 3</math> de 4.29</i>	134
4.18	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 3</math></i>	134
4.19	<i>m-historias y Símbolos para <math>\tau = 4</math> de 4.29</i>	135
4.20	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 4</math></i>	135
4.21	<i>m-historias y Símbolos para <math>\tau = 5</math> de 4.29</i>	135
4.22	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 5</math></i>	135
4.23	<i>m-historias y Símbolos para <math>\tau = 6</math> de 4.29</i>	135
4.24	<i>Símbolos y Probabilidades de Símbolo de 4.29 para <math>\tau = 6</math></i>	135
4.25	Resumen Criterios de Selección de los Parámetros de Reconstrucción $\tau$ y $\tau_w$ de los Métodos MI, C-C y C-C-Sc	148
4.26	Resumen Criterios de Selección de los Parámetros de Reconstrucción $\tau$ y $\tau_w$ de los Métodos $f_{re}$ y $f_{re}^1$	148
4.27	Valor óptimo de $R^{opt}$ para los Sistemas Caóticos de Lorenz, Rossler y Duffing.	149
4.29	LLE y parámetros de reconstrucción $\tau(t)$ y $\tau_w(t)$ para los métodos de selección MI, C-C y C-C-Sc.	152
4.30	LLE y parámetros de reconstrucción $\tau_d(t)$ y $\tau_w(t)$ para los métodos de selección $f_{re}$ y $f_{re}^1$ .	152
4.31	Dimensión de Correlación $D_2$ para los métodos de selección C-C y C-C-Sc, $f_{re}$ y $f_{re}^1$ .	152
4.32	Valores medios de selección de los parámetros de inmersión de retardo y ventana temporales en función de los niveles de ruido.	160
4.33	Modelos bajo la Hipótesis Nula.	170
4.34	Valor de $\overline{R^{opt}}$ para los DGP's Normal $N(0, 1)$ y Uniforme $U(0, 1)$ .	171
4.35	Tamaño Empírico de varios Modelos de Selección.	171

4.36	Tamaño Empírico para los métodos $f_{re}$ y $f_{re}^1$ con los valores de $\overline{R_e^{opt}}$ de la Tabla 4.34.	171
4.37	Modelos bajo la Hipótesis Alternativa.	172
4.38	Valor del radio normalizado $\overline{R_e^{opt}}$ para el Sistema Caóticos de Lorenz con varios niveles de ruido.	174
4.39	Potencia Empírica de varios modelos de reconstrucción.	174
4.40	Potencia Empírica para los métodos $f_{re}$ y $f_{re}^1$ con los valores de $\overline{R_e^{opt}}$ de la Tabla 4.34.	174

## Lista de Apéndices

Apéndice	Título	Página
<b>A</b>	<b>Prueba de la Convergencia de <math>G^{q \rightarrow 1}(m) \sim \chi^2(k)</math></b>	<b>194</b>

## Lista de Términos

- $\tau_w$ : Ventana de retardo máxima llamada ventana de inmersión que relaciona la dimensión de inmersión  $p$  y el retardo temporal  $\tau$
- $h_\tau(m)$ : Entropía simbólica que depende del retardo temporal.
- $P_w$ : Periodo medio orbital de los sistemas caóticos de baja dimensión que muestran pseudo-periodicidad.
- $r_e$ : Parámetro de las Aplicaciones de Simbolización  $f_{r_e}$  y  $f_{r_e}^1$  que llamamos Radio y se define como la distancia estadística con el que se comparan los valores de la serie caótica temporal observada  $\{X_t\}_{t \in T}$ .
- AR-GARCH: Proceso Estocástico Generalizado Autorregresivo Condicionalmente Heterocedástico cuya media sigue un Proceso Estocástico Autorregresivo.
- AR: Proceso Estocástico Autorregresivo.
- ARCH-M: Proceso Estocástico Autorregresivo Condicionalmente Heterocedástico cuya media sigue los movimientos de la varianza.
- ARCH: Proceso Estocástico Autorregresivo Condicionalmente Heterocedástico.
- ARFIMA: Proceso Estocástico Autorregresivo Integrado Fraccional de Media Móvil.
- ARSV: Proceso de Volatilidad Estocástica Autorregresivos o de memoria corta.
- B: Número de realizaciones Bootstrap
- BDS: Prueba estadística no paramétrica basada en la estimación de la integral de correlación en varias dimensiones, para probar la independencia de una serie de tiempo. Prueba desarrollada por Brock, Dechert y Scheinkman en 1986.
- BP: Test de Heterocedasticidad de Breuch-Pagan.

- C-C- $S_c$ : Tipo de técnica de ventana de retardo temporal de series temporales univariantes basada en el método de reconstrucción C-C así como en la Integral de Correlación Média.
- C-C: Tipo de técnica de ventana de retardo temporal de series temporales univariantes donde se asume que el retardo temporal y la dimensión de inmersión son interdependientes. Fue propuesta por Kim, Eykholt y Salas en 1999.
- DGP: Data Generation Process
- EGARCH: Proceso Estocástico Generalizado Exponencial Condicionalmente Heterocedástico.
- G(m): Test Simbólico de Independencia temporal desarrollado por Matilla-García en 2007.
- G-P: Grassberger y Procaccia, autores del estadístico insesgado para estimar la integral de correlación.
- GARCH: Proceso Estocástico Autorregresivo Generalizado Condicionalmente Heterocedástico.
- GQ: Test de Heterocedasticidad de Goldfeld y Quandt.
- h(m): Entropía Simbólica
- i.i.d: Proceso Estocástico Independiente y Uniformemente Distribuido.
- LLE: Cálculo del Exponente de Lyapunov más grande una serie temporal para medir el grado de separación de las órbitas de un sistema caótico que inicialmente estaban muy cercanas.
- LM: Método de los Multiplicadores de Lagrange.
- LMSV: Proceso de Volatilidad Estocástica de memoria larga propuestos por Breidt, Crato y de Lima en 1998.
- m: Tamaño de Símbolo. Dimensión de la historia en la que se inserta la serie  $X(t)$  correspondiente a una Aplicación de Simbolización f.
- OLS: Regresión de Mínimos Cuadrados Ordinarios

- p: Dimensión de reconstrucción o Dimensión de inmersión en la que se inserta la serie temporal observada  $X(t)$  para reconstruir el espacio de fases de un atractor caótico.
- q: Parámetro real llamado Índice Entrópico relacionado con el grado de No-Extensividad del Sistema.
- SG(m): Test Simbólico de Independencia Espacial desarrollado por Matilla-García y Ruiz en 2008.
- SV: Proceso de Volatilidad Estocástica, propuestos por Taylor en 1986.
- T: Número total de observaciones de la serie de tiempo.

# Capítulo 1

## Introducción

La dinámica simbólica es una técnica utilizada de forma generalizada para la investigación de los sistemas dinámicos no lineales (para un análisis exhaustivo ver Hao y Zheng, 1998 [45]). El proceso de simbolización de una serie temporal está basada en el método de los retardos temporales, introducido por Packard et al. [63], para llevar a cabo la reconstrucción del espacio de estados. Esta reconstrucción, demostrada en el teorema de Takens [75], se realiza partiendo de una serie temporal escalar extrayendo toda la información relevante (relativa a la dinámica subyacente), tales como dependencias, periodicidad o cambios en la complejidad.

Basándose en los detalles del procedimiento de reconstrucción del atractor de Takens, surgió el método de Bandt y Pompe [8], que consiste en medir la complejidad de los sistemas unidimensionales sobre la base de la distribución de los patrones ordinales en el sistema. En relación a este concepto ha surgido una gran cantidad de literatura relevante en varias áreas de la Ciencia [6]. Este método consiste ni más ni menos que analizar las relaciones de orden entre los valores del sistema. Es evidente que cuando se analizan los patrones ordinales se pierde algo de información, pero el punto crucial es que en muchos casos no se pierde la información estructural intrínseca y por consiguiente se puede extraer de estos patrones ordinales la información relevante para su análisis.

La idea del análisis ordinal de series es transformar una serie unidimensional, por

ejemplo una serie temporal  $X_{t-4}, X_{t-3}, X_{t-2}, X_{t-1}, X_t, X_{t+1}$ , en una serie de patrones ordinales que describen la estructura de orden de vectores retardados dependientes temporalmente  $(X_t, X_{t+\tau}, \dots, X_{t+(m-1)\tau})$  para un retardo  $\tau$  y dimensión de inmersión  $m$ . Una vez obtenidos los patrones ordinales y su distribución, uno puede usar una gran diversidad de características de la estadística teórica y de la teoría de la información para cuantificar y visualizar cambios o eventos interesantes en las series temporales. Por ejemplo, la entropía de permutación introducida por Bandt y Pompe [8] está basada en la entropía de Shannon. Algunas de las ventajas de estos métodos es que son conceptualmente simples, robustos, flexibles y computacionalmente rápidos.

En esta tesis usamos como herramienta fundamental de inferencia estadística el análisis de patrones, que es un campo más amplio que el análisis ordinal y posee una de las ventajas atractivas del análisis de patrones ordinales: ésta consiste en que no requiere casi supuestos sobre el proceso generador para llevar a cabo la inferencia estadística. Esto es atractivo para muchas herramientas potenciales basadas en este método, que se puede considerar libre de modelo, ya que evita realizar supuestos innecesarios y proporciona resultados más generales.

En este trabajo de investigación consideraremos aquellos aspectos de la inferencia estadística y aplicaciones del análisis de patrones en aquellas áreas científicas que usan los tests estadísticos bajo un nivel mínimo de hipótesis en el contexto del análisis de series temporales. Algunos de los artículos más importantes que tratan sobre las características estadísticas de los patrones ordinales son los de Keller et al. [49], Bandt y Shina [9] y Matilla-García y Ruiz [60], [61]. Las áreas científicas potencialmente beneficiarias son aquellas que usan fuentes de datos no experimentales. También es importante resaltar que la utilización del análisis de patrones (ordinales o no ordinales) no es nuevo en muchas de estas áreas.

En economía y finanzas, los indicadores de recesión utilizados para seguir los ciclos económicos, y los indicadores de mercado alzistas y bajistas usados para caracterizar el mercado de valores son dos claros ejemplos en los que los datos se transforman y analizan en términos de patrones particulares o “símbolos”, véase por ejemplo [58]. En

geografía, el análisis estadístico y la descripción de mapas están hechos en referencia a variables cualitativas y es por eso por lo que el análisis simbólico puede ser también especialmente clarificador (véase [70]). En la estadística espacial, la dependencia económica espacial ha sido estudiada recientemente mediante la transformación de los datos en patrones [55]. En ensayos clínicos, la repetición de la medida de la condición de un individuo dado es también un ejemplo de análisis simbólico [24]. En genética, la aproximación a través del análisis simbólico [71] permite construir tests potentes de asociación genética, y especialmente para la identificación de efectos raros, que son fundamentales para comprender la etiología de las enfermedades complejas, básicamente por que se deben a la combinación de factores genéticos y ambientales. Finalmente en meteorología, gran parte del interés se centra en la exactitud de la predicción de las variables climáticas simbolizadas [73].

Sin embargo, aunque la aplicación de la Dinámica Simbólica es generalizada en muchas áreas de la Ciencia, aún hay necesidad de comprobar su eficacia en el campo de las series financieras, como es por ejemplo la detección del fenómeno de la heterocedasticidad. Por ello, el primer objetivo de esta tesis es comprobar que el Análisis Simbólico es una herramienta válida y potente para la construcción de tests estadísticos competitivos en el campo del contraste de la heterocedasticidad, el cual es un hecho frecuente y relevante en la mayoría de las series temporales financieras. La heterocedasticidad implica que la varianza del proceso no es constante a lo largo del tiempo y se puede ganar una información (económicamente relevante) primero detectándola y después modelándola. Mientras que hay un acuerdo general en la necesidad de la detección y modelado del fenómeno heterocedástico, no hay un único test que haya conseguido una amplia aceptación comparable al test de autocorrelación de Durbin-Watson. Además el menú de tests disponibles continúa creciendo año a año y a pesar de ello una gran cantidad presenta carencias teóricas en cuanto a los supuestos adoptados o prácticas en cuanto al rendimiento de los mismos. A diferencia de las técnicas en las que se basan mayoría de tests utilizables en la actualidad, nosotros adoptamos un camino diferente ya que los tests que proponemos no sólo son novedosos por que estudian la presencia de

heterocedasticidad mediante la *Dinámica Simbólica* sino que se basan en la *Teoría de la Información*. El fundamento de estos tests es fácil de explicar, dada una serie temporal  $\{X_t\}$ , estudiamos la dependencia presente en la varianza de las series trasladando el problema al terreno de la dinámica simbólica y entonces usamos la medida de Entropía de Shannon asociada a estos símbolos para contrastar la heterocedasticidad presente en las series temporales.

Como se mostrará a lo largo de los capítulos dos y tres de esta tesis, estas herramientas nos permiten construir tests omnibus de heterocedasticidad de naturaleza semiparamétrica simples, consistentes, fáciles de computar y competitivos en términos de tamaño y potencia con los tests de heterocedasticidad más conocidos; como el test de White [84], el test de Breuch-Pagan-Godfrey ([19], [36]), el test de Goldfeld-Quandt [38] o el test de Engle [29]. Otro aspecto de importante relevancia de nuestros tests es que los resultados de tamaño de muestra finita no difieren del tamaño asintótico, y de esta forma se asegura la aplicabilidad general y reproducibilidad de los tests.

En la mayoría de los casos, cuando utilizamos la Entropía de Shannon como medida de la Información extraída mediante el análisis de patrones, estamos pasando por alto la dinámica microscópica subyacente en la serie temporal bajo estudio y asumimos de antemano que estamos tratando con sistemas que interaccionan con fuerzas de corto alcance. Es sobradamente conocido que las Entropías de Shannon y Boltzmann Gibbs (la Entropía equivalente a la Entropía de Shannon en el campo de la Termodinámica) funcionan perfectamente, pero dentro de ciertos límites, con sistemas compuestos de subsistemas independientes o de corto alcance cuyos subsistemas pueden acceder a todo el espacio de estados disponible.

En realidad, una cantidad significativa de sistemas complejos dentro de los campos de la Física [7, 11, 10, 12], Medicina [22, 74], Economía [5, 17], Telecomunicaciones [2] y otros conocidos hoy en día se han adaptado de una manera más o menos forzada a las Entropías de Boltzmann Gibbs o de Shannon. Estos sistemas, caracterizados por estar fuera del equilibrio, exhiben correlaciones de largo alcance, memoria o propiedades fractales. Para este tipo de sistemas la Entropía de Tsallis que emerge de forma natu-

ral con la mecánica estadística no extensiva, se convierte en la función entrópica más apropiada [79, 80]. Concretamente, el uso de los conceptos y métodos de la mecánica estadística y termodinámica en economía y finanzas ha resultado ser muy productivo desde hace tiempo. Una de estas aplicaciones recientes, fruto de las ideas asociadas con la mecánica extensiva estadística, consiste en la capacidad de reproducir estadística y dinámicamente los procesos de naturaleza heterocedástica, por ser sistemas complejos que muestran correlaciones de largo alcance [26]. Estos procesos cubren desde los modelos ARCH, GARCH [26] hasta los modelos de Volatilidad Estocástica SV [28]. Por esta razón se puede utilizar la Entropía de Tsallis para discriminar mejor este tipo de procesos que la Entropía de Shannon, para un rango apropiado de valores del índice entrópico  $q$ .

Así, el segundo objetivo de esta tesis es mostrar que la Entropía de Tsallis, cuando se usa en el marco del Análisis Simbólico, posibilita la construcción de tests estadísticos más competitivos que los basados en la Entropía de Shannon. Concretamente, en el Capítulo 3 se propone y desarrolla un nuevo test bootstrap semiparamétrico de tipo omnibus que utiliza la Entropía de Tsallis como medida entrópica de la distribución de probabilidades generada por la Dinámica Simbólica. Como se mostrará a lo largo del capítulo, el nuevo test es más potente frente a la mayoría de modelos heterocedásticos que los tests simbólicos basados en la Entropía de Shannon propuestos en el capítulo dos; manteniendo al mismo tiempo prestaciones no diferentes de los teóricos en los resultados de tamaño. Como principal novedad, el test demuestra ser significativamente más potente que sus contrapartidas de Shannon aplicadas a los procesos de Volatilidad Estocástica SV; procesos que han demostrado ser incapaces de capturar. Además, otra característica importante es que el procedimiento consigue evitar el problema de encontrar el valor óptimo del índice entrópico, ya que integra dentro del resultado del propio test la búsqueda del valor más adecuado.

En el contexto de los sistemas caóticos, la inmersión de los datos experimentales es el primer paso común a muchas técnicas de análisis dinámico. Siguiendo el procedimiento de Takens, una serie temporal escalar  $\{X_t\}_{t \in I}$ , se usa para construir vectores en  $\mathbb{R}^p$  de la

forma  $X_p(t) = (X, X_{t+\tau}, X_{t+2\tau}, \dots, X_{t+(p-1)\tau})$ , donde  $\tau$  es un retardo temporal aplicado a  $\{X_t\}$  y  $p$  es la dimensión del Pseudo espacio de fases o dimensión reconstruida<sup>1</sup>. Para valores apropiados de  $p$  y  $\tau$  se define una aplicación suave  $P : X_t \rightarrow X_{t+1}$  que reconstruye la dinámica subyacente. La elección de los parámetros apropiados  $p$  y  $\tau$  es crucial para el análisis subsiguiente. La precisión de  $\tau$  y  $p$  está directamente relacionado con la exactitud de los invariantes dinámicos del atractor extraño en la reconstrucción del espacio de fases. Para realizar una reconstrucción adecuada hay dos puntos de vista. Uno es que  $p$  y  $\tau$  no están correlacionados, es decir,  $p$  y  $\tau$  se pueden seleccionar de forma independiente (Takens demostró que  $p$  y  $\tau$  son independientes en una serie caótica temporal de longitud infinita y libre de ruido). Bajo esta regla de oro, Albano et al. propuso el algoritmo G-P [4] para calcular la dimensión de inmersión  $p$ . Entre los criterios para seleccionar el retardo  $\tau$  uno de los más utilizados es el método de la Información Mútua [33]. El segundo punto de vista es que  $p$  y  $\tau$  están fuertemente relacionados ya que las series temporales en el mundo real no son infinitamente largas ni están libres del ruido. Un gran número de experimentos indican que  $\tau$  y  $p$  están ligados con la ventana temporal  $\tau_w = (p-1)\tau$  para la reconstrucción del espacio de estados ya que  $\tau_w$  es relativamente inalterable. Uno de los métodos más importantes para estimar de forma combinada  $p$  y  $\tau$  es el método C-C basado en la integral de correlación [50]. Como en gran cantidad de los métodos sugeridos hasta la fecha para la reconstrucción del espacio de estados, el método de la Información Mútua [33] y el método C-C seleccionan los parámetros de reconstrucción  $\tau$  y  $\tau_w$  en función de la existencia de extremos locales en los funcionales propuestos por cada método para caracterizar la información de la serie temporal. La inspección y selección de estos extremos se realiza hasta la fecha de forma “visual”, es decir, se “dibuja” el funcional de cada método en función del retardo y se identifican visualmente los extremos locales donde se localizan el retardo y ventana temporales  $\tau$  y  $\tau_w$ . Sin embargo, las series temporales reales están

---

<sup>1</sup>Aunque como hemos señalado al principio del capítulo, el Análisis Simbólico se deriva del método de reconstrucción del espacio de estados de un sistema caótico, con el fin de no dar lugar a equívocos, se designa con  $m$  la dimensión de inmersión de una serie temporal  $\{X_t\}_{t \in I}$  cuando se somete a un proceso de simbolización; en cambio se denomina con  $p$  a la dimensión de reconstrucción (inmersión) de un Sistema Caótico.

contaminadas por ruido por lo que la precisión (error estándar) y exactitud (sesgo) de los métodos de selección en la identificación de los parámetros de reconstrucción disminuye conforme aumenta su nivel. De esta manera no podemos estar seguros de si los parámetros elegidos son significativos desde un punto de vista Informacional o son en cambio, valores espúreos causados por el ruido. Esto es debido a que en la actualidad no existe un método estadístico riguroso y definitivo para determinar la significatividad del extremo local seleccionado frente a sus valores adyacentes, es decir, un método que nos permita acotar la precisión o error estándar de los parámetros seleccionados para cada método de reconstrucción.

A la vista de las limitaciones mencionadas, el tercer objetivo de esta tesis es proponer y desarrollar un test estadístico no paramétrico que contraste la significatividad de los extremos locales seleccionados en los retardos  $\tau_w$  y  $\tau$ . Como se mostrará de forma pormenorizada en el capítulo cuarto de esta tesis, el test nos permite discriminar el retardo seleccionado con los retardos adyacentes y acotar la precisión de los parámetros seleccionados para cada método de reconstrucción utilizado. De esta forma el test nos ofrece como ventaja principal un criterio para seleccionar un método de reconstrucción en función del nivel de precisión que es capaz de alcanzar en la identificación de los parámetros  $\tau_w$  y  $\tau$ . Otra de las ventajas importantes del test propuesto es que los resultados de tamaño del test no difieren del nominal y lo que es más importante, no realiza supuestos extra sobre el funcional a especificar, además de su simplicidad computacional. Aparte del criterio estadístico de selección, también proponemos como principal novedad, dos nuevos métodos de selección de los parámetros  $\tau$  y  $\tau_w$  basados en el análisis de patrones.

## Capítulo 2

# Contraste de Heterocedasticidad Condicional de Forma Desconocida en Series Temporales Mediante Dinámica Simbólica

### 2.1. Introducción

Desde el artículo fundamental de Engle [29] en el que introdujo los modelos ARCH, uno de los problemas fundamentales de la metodología econométrica es contrastar la presencia de heterocedasticidad en las perturbaciones de una serie temporal. La literatura estadística y econométrica que trata sobre el problema del contraste de la heterocedasticidad es bastante extensa. En el contexto de la regresión lineal, los procedimientos más populares incluyen los tests de Goldfeld-Quandt [38], el test de Breuch-Pagan-Godfrey ([36], [19]) y el test general de White's [84]. Otros métodos propuestos incluyen los tests de Verosimilitud (LR) contra alternativas específicas (véase por ejemplo Harvey [42]) y los “procedimientos robustos”, tales como el test máximo de Goldfeld y Quandt [38]. Estos métodos toman la varianza como parámetro de ruido que tiene que tenerse en cuenta cuando se realizan inferencias sobre otros parámetros del modelo (tales como los coeficientes de regresión).

Como resultado de la popularidad actual de los modelos ARCH, GARCH [15] y otros similares (véase [32] para estudio excelente del modelado GARCH), la detección de la

presencia de heterocedasticidad condicional de naturaleza estocástica se ha convertido en una cuestión muy importante, y se han propuesto un gran número de tests para detectar la presencia de tales efectos; véase por ejemplo los tests de Engle [29], Lee and King [52], Bera y Ra [13] y Hong and Shehadeh [46].

A pesar del gran espectro de tests disponibles, la gran mayoría manifiesta uno o más de los problemas que vamos a mencionar a continuación y que son referidos por la literatura econométrica. La mayoría tienen que ver con el número de supuestos necesarios para obtener las distribuciones de los tests estadísticos. Varios estudios comparan varios tests de heterocedasticidad desde los puntos de vista de fiabilidad y potencia; véase por ejemplo, McKinnon y White [59], Godfrey [37], y en conexión con los tests GARCH, Engle, Hendry y Trumble [30], Lee y King [52] o Bera y Ra [13]. La mayoría de las referencias citadas incluyen pruebas empíricas sobre el rendimiento relativo de varios de los tests mencionados. Los principales hallazgos (quizá influenciados por considerar algunos supuestos especiales y de diseño de las simulaciones) que emergen de estos estudios son los siguientes: (i) ningún test analizado tiene más potencia contra todas las alternativas; (ii) los tests basados en residuos OLS son los que poseen mejor rendimiento (iii) el nivel actual de tests justificados por su naturaleza asintótica está bastante lejos de los niveles nominales: algunos tests están sobredimensionados, mientras que otros están fuertemente infradimensionados (iv) los tests exactos son los que mejor parados salen de su comparación con los tests asintóticos pero son los más difíciles de implementar en la práctica.

Para superar los problemas planteados por estos tests proponemos utilizar el análisis simbólico. El análisis simbólico es una herramienta potente para el estudio de los sistemas dinámicos. El propósito de la dinámica simbólica es proporcionar un dibujo simplificado y entendible de los aspectos más relevantes de la dinámica compleja. Para este fin, el espacio de estados se particiona en un número finito de trozos, y las trayectorias exactas se resumen con la trayectoria relativa a esa partición. Estas así llamadas trayectorias de grano grueso terminan por ser realizaciones de un proceso aleatorio estacionario con un alfabeto finito. En particular, la entropía de un sistema dinámico se

puede aproximar por la entropía de Shannon de cualquiera de sus dinámicas simbólicas (cuanto más fina la partición, mejor la aproximación) y que denominamos *Entropía Simbólica*. Hoy, la dinámica simbólica es un campo independiente de la física teórica y de las matemáticas aplicadas con aplicaciones tan importantes como la criptología, análisis de series temporales, y compresión de datos.

En este capítulo, pretendemos proponer algunos tests que utilizan el análisis simbólico para contrastar la presencia de heterocedasticidad en el contexto de una especificación correcta de la media condicional. Esta es la razón por la que los tests se pueden considerar de naturaleza semiparamétrica. Además nuestros tests no requieren conocimiento a priori de las posibles alternativas (incluyendo tanto formas funcionales como estructuras del retardo de la varianza condicional). Los predecesores naturales de estos tests son los tests de dependencia en series temporales desarrollados por Matilla y Ruiz en [60] y en [61] los cuales están basados en la entropía de permutación.

Este capítulo se organiza como sigue. Las Secciones 2.2 y 2.3 muestran una metodología flexible para obtener patrones (ordinales o no) y llegar a un test general que depende de la entropía “simbólica” asociada con un proceso. Concretamente en la sección 2.3 se describe como ejemplo el proceso de obtención de los tests de independencia  $G(m)$  y  $SG(m)$  propuestos por Matilla [60] y Matilla y Ruiz [61] mediante el uso de patrones ordinales y la entropía de permutaciones. En la Sección 2.4 se presenta el marco estadístico de definición y diseño de los tres tests de heterocedasticidad que se proponen en este capítulo. En las secciones 2.5, 2.6 y 2.7 se describen los tres tests propuestos. El objetivo de la sección 2.8 es doble: primeramente mostrar el comportamiento de los tests propuestos en términos de tamaño y potencia para tamaños de muestra finita, y segundo comparar el comportamiento del test más competitivo de los tres con otros disponibles. En la sección 2.9 se analiza el impacto en el tamaño nominal de uno de los parámetros de diseño de los tests, denominado solapamiento de símbolo. Finalmente en la sección 2.10 se exponen las conclusiones.

## 2.2. Introducción a la Entropía Simbólica

En esta sección daremos algunas definiciones y notaciones referentes al campo del análisis simbólico.

Supongamos que queremos construir un test estadístico para contrastar una hipótesis nula  $H_0$  sobre la serie  $\{X_t\}_{t \in I}$ . Para un entero positivo  $m \geq 2$  denotamos como  $\Gamma_m$  el conjunto de elementos finito de orden  $n$ . Ahora definimos la aplicación  $f$  que llamaremos *Aplicación de Simbolización* :

$$f : \mathbb{R}^m \rightarrow \Gamma_m$$

donde  $\pi_i = f(X_m(t)) \in \Gamma_m$ .

Llamaremos *símbolo* a un elemento  $\pi_i$  of  $\Gamma_m$ . El entero positivo  $m$  es normalmente conocido como *dimensión de inmersión*. En lo que sigue consideraremos que la serie temporal  $\{X_t\}_{t \in I}$  está inmersa en un espacio  $m - dimensional$  como sigue:

$$X_m(t) = (X_t, X_{t+1}, \dots, X_{t+(m-1)}) \text{ para } t \in T \quad (2.1)$$

donde  $I$  es un conjunto de índices temporales. Además diremos que  $t \in T$  es de tipo  $\pi_i$  si y solo si  $\pi_i = f(X_m(t))$ . Así pues la  $m$  no es ni más ni menos que la longitud de la ventana de  $X_m(t)$ . La dimensión de inmersión determina el tamaño de los patrones investigados para medir la complejidad de las series.

Por lo tanto  $f(X_m(t)) = \pi_i$ , es decir, cada  $m - historia$   $X_m(t)$  es convertida en un único símbolo  $\pi_i$  de  $\Gamma_m$ . Además, si la aplicación de simbolización es tal que, bajo la hipótesis nula de nuestro test, todos los símbolos tienen la misma probabilidad de ocurrencia, entonces diremos que  $f$  es una *Aplicación de Simbolización Estándar*.

Llamemos a  $p_\pi$  la probabilidad de ocurrencia de un símbolo  $\pi$  y a  $n_{\pi_i}$  la cardinalidad del subconjunto  $I$  formado por todos los elementos de tipo  $\pi_i$

$$n_{\pi_i} = \{ t \in I \mid f(X_m(t)) = \pi_i \}$$

Por lo que la frecuencia relativa de un símbolo  $\pi \in \Gamma_m$  viene dada por

$$p(\pi) = p_\pi = \frac{\{t \in I \mid t \text{ es de tipo } \pi\}}{T - (m - 1)}$$

donde  $T$  es la cardinalidad del conjunto  $I$ .

Ahora, bajo este escenario, podemos definir la *entropía simbólica* de un proceso temporal  $\{X_t\}_{t \in I}$  para una dimensión  $m \geq 2$ . Esta entropía se define como la Entropía de Shannon de  $n$  símbolos distintos como sigue:

$$h(m) = - \sum_{\pi \in \Gamma} p_\pi \ln(p_\pi) \quad (2.2)$$

La entropía simbólica,  $h(m)$ , es la información contenida en comparar las  $m$ -*historias* generadas por el proceso temporal. Observar que si la aplicación de simbolización es estándar, entonces  $0 \leq h(m) \leq \ln(n)$  donde el límite inferior se alcanza cuando sólo ocurre un símbolo, y el límite superior para un sistema totalmente aleatorio (secuencia i.i.d temporal) donde todos los posibles  $n$  símbolos aparecen con la misma probabilidad.

### 2.3. De la Entropía Simbólica a la Entropía de Permutación

Partiendo de la idea de entropía simbólica (ecuación 2.2), Matilla-García y Ruiz en [61] utilizaron patrones de orden para definir una Aplicación de Simbolización Estándar y construir un test de independencia serial. A continuación describiremos a modo ilustrativo cómo se define dicha aplicación.

Sea  $\{X_t\}_{t \in I}$  una serie temporal de valores reales. Definimos como  $S_m$  el grupo simétrico de orden  $m!$  (para un entero positivo  $m \geq 2$ ), esto es el grupo formado por todas las permutaciones de longitud  $m$ . Sea  $\pi_i = (i_1, i_2, \dots, i_m) \in S_m$ . Ahora vamos a definir el patrón ordinal  $\pi_i = (i_1, i_2, \dots, i_m) \in S_m$  en un tiempo  $t \in I$  de una  $m$ -*historia*. Para este fin consideraremos que la serie temporal  $\{X_t\}_{t \in I}$  está inmersa en un espacio  $m$ -dimensional como sigue:

$$X_m(t) = (X_{t+1}, X_{t+2}, \dots, X_{t+m}) \text{ para } t \in I \quad (2.3)$$

Entonces diremos que  $t$  es de tipo- $\pi_i$  si y solo si  $\pi_i = (i_1, i_2, \dots, i_m)$  es el único símbolo en el grupo  $S_m$  que satisface las siguientes dos condiciones:

$$(a) \quad X_{t+i_1} \leq X_{t+i_2} \leq \dots \leq X_{t+i_m}, \text{ y}$$

$$(b) \quad i_{s-1} < i_s \text{ if } X_{t+i_{s-1}} = X_{t+i_s}$$

La condición (b) garantiza la unicidad del símbolo  $\pi_i$ . Esto se justifica si los valores de  $X_t$  tienen una distribución continua de manera que los valores iguales son poco comunes, con una probabilidad teórica de ocurrencia igual a 0.

Ahora definimos la aplicación de simbolización como  $f : \mathbb{R}^m \rightarrow S_m$  dada por

$$f(X_m(t)) = (i_1, i_2, \dots, i_m) \quad (2.4)$$

donde  $(i_1, i_2, \dots, i_m) \in S_m$  de tal manera que  $t$  es de tipo- $(i_1, i_2, \dots, i_m)$ . Para esta aplicación de simbolización, la entropía simbólica de la ecuación 2.2 es la bien conocida entropía de permutación de Bandt y Pompe [8].

## 2.4. De la Entropía Simbólica a la Construcción de Tests No Paramétricos. El Test de Independencia $G(m)$

Supongamos que queremos construir un test estadístico para una serie temporal  $\{X_t\}_{t \in I}$ . Como hemos visto en las secciones anteriores (2.2 y 2.3) el proceso de construcción de un test estadístico que utiliza como herramienta el análisis simbólico pasa en primer lugar por definir una aplicación de simbolización  $f$ . A continuación tenemos que aplicar  $f$  a  $\{X_t\}_{t \in I}$ , de la que resulta un conjunto de símbolos a los que podemos asignar la entropía simbólica como medida de su distribución estadística. También hemos visto en la sección 2.1 que  $f$  es una *Aplicación de Simbolización Estándar* si, bajo la hipótesis nula del test que queremos construir, todos los símbolos tienen la misma probabilidad de ocurrir. Cuando esto sucede, se puede demostrar como veremos a lo largo de este

apartado que una transformación afín de la entropía simbólica se distribuye asintóticamente como una distribución chi-cuadrado:

$$2R[\ln(n) - h(\Gamma_m)] \sim \chi_k^2, \quad (2.5)$$

bajo la aplicación de simbolización  $f : \mathbb{R}^m \rightarrow \Gamma_m$ , siendo  $\Gamma_m$  el grupo de elementos de orden  $n$  y dimensión  $m$ . Donde  $n$  es el número de símbolos de  $\Gamma_m$ ,  $T$  es la cardinalidad del conjunto  $I$ ,  $R = T - (m - 1)$  es el número de  $m - \text{historias}$  resultantes del proceso de simbolización  $f$  de  $\{X_t\}_{t \in I}$  y  $k$  expresa los grados de libertad que dependerá del conjunto de símbolos.

Hasta ahora la literatura se ha concentrado en estudiar las *Aplicación de Simbolización Estándar* por ser, a priori, de tratamiento estadístico más formal y sencillo en comparación con las *Aplicaciones de Simbolización no Estándar*, es decir, aquellas en las que la probabilidad de ocurrencia de los símbolos es diferente bajo la hipótesis nula.

Como ejemplo ilustrativo del uso de las Aplicaciones de Simbolización Estándar, Matilla-García y Ruiz en [61] y López et al. en [55], utilizaron la *Entropía de Permutación* y la *Entropía Simbólica* para obtener unos tests de independencia temporal y espacial llamados  $G(m)$  y  $SG(m)$  respectivamente. A continuación describiremos el proceso de construcción en el que se basan ambos tests en el contexto de las series temporales (test  $G(m)$ ).

Sea  $\{X_t\}_{t \in I}$  un proceso temporal y  $m$  su dimensión de inmersión. Para construir el test de independencia temporal se considera la hipótesis nula:

$$\mathbb{H}_0 : \{X_t\}_{t \in I} \text{ i.i.d} \quad (2.6)$$

contra cualquier otra alternativa.

Ahora, para un símbolo  $\pi_i = (i_1, i_2, \dots, i_m) \in \Gamma_m$  definimos una variable aleatoria  $Z_{\pi_i t}$  como sigue:

$$Z_{\pi_i t} = \begin{cases} 1 & \text{si } f(X_m(t)) = \pi_i \\ 0 & \text{en los demás casos} \end{cases} \quad (2.7)$$

esto es, tenemos que  $Z_{\pi_i t} = 1$  si y solo si  $t$  es de tipo- $\pi_i$ ,  $Z_{\pi_i t} = 0$  en otro caso.

Así,  $Z_{\pi_i t}$  es una variable de Bernoulli con probabilidad de "éxito"  $p_{\pi_i}$ , donde éxito significa que  $t$  es del tipo- $\pi_i$ . Es inmediato ver que se cumple que:

$$\sum_{i=1}^n p_{\pi_i} = 1 \quad (2.8)$$

Asumamos que  $I$  es finito y de orden  $T$ . Entonces estamos interesados en saber cuantos  $t$ 's son de tipo- $\pi_i$  para todo símbolo  $\pi_i$  que pertenece a  $\Gamma_m$ . Para contestar a esta pregunta construimos la siguiente variable:

$$Y_{\pi_i} = \sum_{t=1}^R Z_{\pi_i t} \quad (2.9)$$

La variable  $Y_{\pi_i}$  puede tomar los valores  $\{0, 1, 2, \dots, R\}$ . Entonces se deduce que  $Y_{\pi_i}$  es quasi una variable aleatoria binomial

$$Y_{\pi_i} \approx B(R, p_{\pi_i}) \quad (2.10)$$

Bajo la hipótesis  $\mathbb{H}_1$  la función densidad de probabilidad conjunta de las  $n$  variables  $(Y_{\pi_1}, Y_{\pi_2}, \dots, Y_{\pi_n})$  es:

$$P(Y_{\pi_1} = a_1, Y_{\pi_2} = a_2, \dots, Y_{\pi_n} = a_n) = \frac{(a_1 + a_2 + \dots + a_n)!}{(a_1! a_2! \dots a_n!)} p_{\pi_1}^{a_1} p_{\pi_2}^{a_2} \dots p_{\pi_n}^{a_n} \quad (2.11)$$

donde  $a_1 + a_2 + \dots + a_n = R$ . Consecuentemente, la distribución conjunta de las  $n$  variables  $(Y_{\pi_1}, Y_{\pi_2}, \dots, Y_{\pi_n})$  es una distribución multinomial.

La función de verosimilitud de la distribución 2.11 es:

$$L(p_{\pi_1}, p_{\pi_2}, \dots, p_{\pi_n}) = \frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!)} p_{\pi_1}^{n_{\pi_1}} p_{\pi_2}^{n_{\pi_2}} \dots p_{\pi_n}^{n_{\pi_n}} \quad (2.12)$$

y puesto que  $\sum_{i=1}^n p_{\pi_i} = 1$ , se deduce que

$$L(p_{\pi_1}, p_{\pi_2}, \dots, p_{\pi_n}) = \frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_n}!)} p_{\pi_1}^{n_{\pi_1}} p_{\pi_2}^{n_{\pi_2}} \dots (1 - p_{\pi_1} - p_{\pi_2} - \dots - p_{\pi_{n-1}})^{n_{\pi_n}} \quad (2.13)$$

Por lo que el logaritmo de la función verosimilitud queda como

$$\begin{aligned} Ln(L(p_{\pi_1}, p_{\pi_2}, \dots, p_{\pi_n})) &= Ln\left(\frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_n}!)}\right) + \dots \\ &\dots + \sum_{i=1}^{n-1} n_{\pi_i} Ln(p_{\pi_i}) + n_{\pi_n} Ln(1 - p_{\pi_1} - p_{\pi_2} - \dots - p_{\pi_{n-1}}) \end{aligned} \quad (2.14)$$

Para obtener los estimadores de máxima verosimilitud  $\hat{p}_{\pi_i}$  de  $p_{\pi_i}$  para todo  $i = 1, 2, \dots, n$ , resolvemos la siguiente ecuación

$$\frac{\partial Ln(L(p_{\pi_1}, p_{\pi_2}, \dots, p_{\pi_n}))}{\partial p_{\pi_i}} = 0 \quad (2.15)$$

obteniendo

$$\hat{p}_{\pi_i} = \frac{n_{\pi_i}}{R} \quad (2.16)$$

El estadístico del cociente de verosimilitudes es (véase por ejemplo, Lehmann [53]):

$$\begin{aligned} \lambda(Y) &= \frac{\frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_n}!)} p_{\pi_1}^{n_{\pi_1}} p_{\pi_2}^{n_{\pi_2}} \dots p_{\pi_n}^{n_{\pi_n}}}{\frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_n}!)} \hat{p}_{\pi_1}^{n_{\pi_1}} \hat{p}_{\pi_2}^{n_{\pi_2}} \dots \hat{p}_{\pi_n}^{n_{\pi_n}}} = \frac{\prod_{i=1}^n p_{\pi_i}^{n_{\pi_i}}}{\prod_{i=1}^n \left(\frac{n_{\pi_i}}{R}\right)^{n_{\pi_i}}} = \\ &= R^{\sum_{i=1}^n n_{\pi_i}} \prod_{i=1}^n \left(\frac{p_{\pi_i}}{n_{\pi_i}}\right)^{n_{\pi_i}} = R^R \prod_{i=1}^n \left(\frac{p_{\pi_i}}{n_{\pi_i}}\right)^{n_{\pi_i}} \end{aligned} \quad (2.17)$$

Además  $G(m) = -2Ln(\lambda(Y))$  tiende asintóticamente a una distribución Chi-cuadrado con  $k$  grados de libertad, donde  $k$  es igual al número de parámetros desconocidos bajo  $\mathbb{H}_0$  menos el número de parámetros desconocidos bajo  $\mathbb{H}_1$  (Lehmann [53]).

Así

$$G(m) = -2Ln(\lambda(Y)) = -2 \left[ RLn(R) + \sum_{i=1}^n n_{\pi_i} Ln \left( \frac{p_{\pi_i}}{n_{\pi_i}} \right) \right] \sim \chi_k^2 \quad (2.18)$$

Si la aplicación  $f$  de simbolización es estándar, esto es, bajo la hipótesis nula todos los símbolos tienen la misma probabilidad de ocurrencia,  $p_{\pi_i} = \frac{1}{n}$  para todo  $i = 1, 2, \dots, n$  entonces se deduce que

$$\begin{aligned} G(m) &= -2R \left[ Ln(R) + \sum_{i=1}^n \frac{n_{\pi_i}}{R} Ln \left( \frac{p_{\pi_i}}{n_{\pi_i}} \right) \right] \\ &= -2R \left[ Ln(R) + \sum_{i=1}^n \frac{n_{\pi_i}}{R} Ln \left( \frac{p_{\pi_i}}{n_{\pi_i}} \right) \right] \\ &= -2R \left[ Ln(R) + \sum_{i=1}^n \frac{n_{\pi_i}}{R} (Ln \left( \frac{1}{n} \right) - Ln(n_{\pi_i})) \right] \\ &= -2R \left[ Ln(R) + \sum_{i=1}^n \frac{n_{\pi_i}}{R} (Ln \left( \frac{1}{n} \right) - Ln \left( \frac{n_{\pi_i}}{R} \right) - Ln(R)) \right] \end{aligned} \quad (2.19)$$

Ahora, si tenemos en cuenta que  $h(m) = - \sum_{\pi \in \Gamma} p_{\pi_i} \ln(p_{\pi_i}) = - \sum_{i=1}^n \frac{n_{\pi_i}}{R} Ln \left( \frac{n_{\pi_i}}{R} \right)$ , tenemos que

$$G(m) = -2R \left[ Ln \left( \frac{1}{n} \right) + h(m) \right] = 2R [Ln(n) - h(m)] \quad (2.20)$$

Y se obtiene así el teorema enunciado por Matilla y Ruiz [61] :

**Teorema 1.** *Sea  $\{X_t\}_{t \in I}$  un proceso temporal de valores reales con  $|T| = R$ . Asumamos que existe una Aplicación Simbólica estándar  $f$  para  $\{X_t\}_{t \in I}$ . Denotemos por  $h(m)$  la entropía simbólica definida en 2.2 para una dimensión de inmersión  $m \geq 2$ , con  $m \in \mathbb{N}$ . Si el proceso temporal  $\{X_t\}_{t \in I}$  es i.i.d, entonces la transformación afín de la Entropía Simbólica*

$$G(m) = 2R [Ln(n) - h(m)] \quad (2.21)$$

*está distribuida asintóticamente como una  $\chi_k^2$ .*

*Sea  $a$  un número real con  $0 \leq a \leq 1$ . Sea  $\chi_k^2$  de tal manera que*

$$P(\chi_k^2 > \chi_\alpha^2) = \alpha$$

Entonces para contrastar

$$\mathbb{H}_0 : \{X_t\}_{t \in I} \text{ i.i.d}$$

la regla de decisión en la aplicación del test  $G(m)$  con un nivel de confianza del  $100(1 - \alpha)\%$  es:

$$\begin{aligned} \text{Si } 0 \leq G(m) \leq \chi_\alpha^2 & \quad \text{Aceptar } \mathbb{H}_0 \\ \text{En otro Caso} & \quad \text{Rechazar } \mathbb{H}_0 \end{aligned}$$

#### 2.4.1. Propiedades del Test $G(m)$

Como se probó en Matilla y Ruiz [61], si un proceso  $\{X_t\}_{t \in I}$  admite una aplicación de simbolización estándar, entonces el test  $G(m)$  es consistente para una amplia variedad de procesos seriales dependientes. Esta propiedad implica que el test rechazará asintóticamente la hipótesis de independencia serial cuando hay dependencia dentro de las  $m$ -vecindades. Por dependencia serial de orden  $\leq m$  se entiende que cualquiera que sea la estructura del proceso, existe dependencia entre la variable aleatoria en el punto  $t$  y sus  $m$ -vecindades (o parte de ellas). De esta propiedad se deduce el siguiente teorema (Matilla y Ruiz [61]):

**Teorema 2.** *Sea  $\{X_t\}_{t \in I}$  un proceso estacionario estricto, y  $m \geq 2$  con  $m \in \mathbb{N}$ , si se asume que existe un proceso de simbolización estándar  $f$  para  $\{X_t\}_{t \in I}$ . Entonces*

$$\lim_{R \rightarrow \infty} \Pr(\hat{G}(m) > C) = 1 \tag{2.22}$$

bajo dependencia serial de orden  $\leq m$  para todo  $0 < C < \infty$ ,  $C \in \mathbb{R}$ . Siendo  $\hat{G}(m)$  el estimador de  $G(m)$ .

### 2.4.2. El Conjunto de los Índices

Hemos visto en la sección anterior (2.4) que el test  $G(m)$  se distribuye asintóticamente bajo la hipótesis nula como una distribución *Chi-Cuadrado* con  $k$  grados de libertad. Además en [60] se mostró cómo los grados de libertad de esta distribución depende de la estrategia de elección de los índices. Si los índices no se solapan, bajo la hipótesis nula, todos los símbolos tendrán una probabilidad específica de aparecer. Por ejemplo, todos los símbolos tendrán la misma probabilidad de aparecer bajo la hipótesis nula de independencia serial si el conjunto de índices que se escoge no se solapa. Este conjunto de índices escogidos de la siguiente manera

$$T' = \{0, m - 1, 2(m - 1), \dots, t(m - 1), \dots\}$$

no se solapará, y entonces la distribución asintótica será  $\chi_{m!-1}^2$ .

Algunas veces por el propio carácter del problema bajo estudio, el investigador no tendrá más remedio que construir aplicaciones de simbolización para los cuales el conjunto  $T'$  se solapa. Esta situación puede aparecer cuando el número de observaciones (tamaño muestral) es relativamente escaso comparado con el número de símbolos. Por ejemplo, en economía, aunque no necesariamente en finanzas, la escasez de datos podría ser una restricción habitual.

Bajo estas circunstancias hay varias estrategias o posibilidades, algunas de las cuales han sido estudiadas en varios artículos ([70] y [55]):

- (i) Controlar el grado de solapamiento (como se propuso en [55]), donde en el contexto de procesos espaciales, los autores muestran como disminuir el impacto de las dependencias entre los indicadores teniendo control del solapamiento de las *m-historias* para asegurar una buena aproximación binomial de los indicadores dependientes. Esta técnica del grado de solapamiento es directamente utilizable en el contexto de series temporales.
- (ii) Utilizar resultados asintóticos para los procesos estacionarios *m-dependientes*. Particularmente, hay algunos teoremas centrales del límite para secuencias *m-dependientes* que pueden ser fácilmente usados para derivar una distribución asintótica normal están-

dar. La normal estándar se consigue siempre y cuando las variables aleatorias indicador utilizadas para obtener la frecuencia de un símbolo dado (permutación) están uniformemente acotadas. Aunque esta posibilidad es teóricamente atractiva, tiene la desventaja de la necesidad de obtener la varianza de la suma de *m-dependientes* variables aleatorias indicador. Esta varianza depende de las probabilidades de transición que definen la probabilidad de un estado al otro dentro del rango de *m-dependencia*. Además dependerá de *m*, y por consiguiente, generalmente para  $m > 3$ , se necesitarán métodos de simulación numérica para definir una varianza correcta aproximada.

Independientemente de si usamos la opción (i) o (ii), para resolver el problema del solapamiento de índices, el estadístico obtenido tiene que garantizar que el test, dado un tamaño de muestra, no rechace demasiado más allá del nivel de significación. Si se da esta situación práctica entonces uno puede estar seguro de usar este estadístico.

Una tercera forma es construir un test diferente usando técnicas estadísticas de bootstrap. En este caso, es importante asegurarse que los procedimientos de remuestreo están de acuerdo con la hipótesis nula.

### 2.4.3. Estudio Empírico de la Convergencia Asintótica del Test $G(m)$

Una vez que hemos definido teóricamente el test  $G(m)$  así como sus propiedades, nuestro objetivo en esta sección es comprobar si el test converge empíricamente bajo la hipótesis nula  $\mathbb{H}_0$  de independencia serial a la distribución asintótica  $\chi^2(k)$  de  $k$  grados de libertad con  $k = m! - 1$  en un escenario ideal de no solapamiento de las *m-historias*. Con este fin realizamos un experimento de tamaño generando una distribución normal  $N(0, 1)$  de  $T=5000$  observaciones para un tamaño de símbolo  $m=3$ . Este procedimiento se repitió 1000 veces.

Para comprobar la convergencia de la distribución finita del estadístico  $\hat{G}(3)$  a la distribución  $\chi^2(5)$ , hemos comparado en la Tabla 2.1 los valores de algunos estadísticos de tendencia central tales como la Media, la Varianza y la Mediana junto con algunos de forma como el Coeficiente de Simetría y la Curtosis. Tal y como esperábamos, podemos observar que los valores de los estadísticos de la distribución teórica  $\chi^2(5)$  son

Tabla 2.1: Valores de algunos Estadísticos para  $\chi^2(5)$  y  $\widehat{G}(3)$ 

Estadísticos	$\chi^2(5)$	$\widehat{G}(3)$
Media	5	4.958
Varianza	10	9.619
Mediana	4.333	4.329
Coefficiente de simetría	1.264	1.12
Curtosis	5.4	5.34

muy parecidos a los obtenidos en el test  $\widehat{G}(3)$ .

## 2.5. Marco Estadístico de Definición y Diseño de los Tests

Supongamos  $\{Y_t\}_{t \in I}$  un proceso estrictamente estacionario. Ahora consideremos un proceso general estocástico de naturaleza temporal

$$Y_t = F(I_{t-1}, \theta) + \varepsilon_t, \quad t = 1, 2, \dots \quad (2.23)$$

donde  $I_{t-1}$  es la información disponible en el tiempo  $t - 1$ , la cual puede contener variables dependientes retardadas, shocks retardados y variables exógenas actuales y retardadas;  $F(I_{t-1}, \theta)$  es un modelo paramétrico de  $E(Y_t | I_{t-1})$ , la media condicional de  $Y_t$  dado  $I_{t-1}$  y  $\theta \in \Theta$  es un parámetro finito dimensional.

Asumamos para la media condicional  $E(Y_t | I_{t-1})$  que

$$F(I_{t-1}, \theta) = X_t' \psi(\theta), \quad (2.24)$$

donde  $X_t$  es un vector  $d \times 1$  en  $I_{t-1}$ ,  $\Psi$  es un parámetro  $d \times 1$  que depende de  $\theta$ , y el entero  $d$  puede ser finito o infinito. Esto es  $F(I_{t-1}, \theta)$  es un modelo paramétrico lineal de  $E(Y_t | I_{t-1})$  y el modelo  $\{Y_t\}_{t \in T}$  es lineal y correctamente especificado en media.<sup>1</sup>

Ya que  $F(I_{t-1}, \theta)$  y  $I_{t-1}$  son conocidos respecto a la media condicional, finalmente estimaremos el siguiente modelo general

<sup>1</sup>La terminología y definición de media condicional está extraída del artículo de Hong y Lee [47].

$$Y_t = X_t' \psi(\theta) + \varepsilon_t, \quad t = 1, 2, \dots \quad (2.25)$$

Y los errores del modelo  $\varepsilon_t = Y_t - F(I_{t-1}, \theta)_t$  tienen media 0, y varianza incondicional  $E(\varepsilon_t^2) = \sigma^2$ .

En la práctica consideraremos que  $X_t$  sólo posee variables endógenas retardadas, es decir,  $X_t = \{Y_{t-1}\}_{t \in T}$ ,  $I_{t-1} = I(1)$  con lo que la ecuación del proceso toma la forma

$$Y_t = \beta Y_{t-1} + \varepsilon_t \quad t = 1, \dots, T \quad (2.26)$$

Por lo tanto, para construir un test en presencia de heterocedasticidad en  $\{Y_t\}_{t \in I}$ , asumiremos la hipótesis nula de homocedasticidad condicional ( $\mathbb{H}_0$ ) la cual implica determinar si la varianza condicional  $E(\varepsilon_t^2 | I_{t-1}) = h_t = g(\varepsilon_t^2)$  es (una variante temporal) constante

$$\mathbb{H}_0 : E(\varepsilon_t^2 | I_{t-1}) = h_t = h, \quad (2.27)$$

contra la hipótesis alternativa de heterocedasticidad condicional ( $\mathbb{H}_1$ ), que implica determinar si varianza condicional  $h_t = g(\varepsilon_t^2)$  es una función de  $t$ , donde  $g(\cdot)$  es una función general desconocida de  $\varepsilon_t^2$ .

Ahora expresamos los residuos estimados por

$$\hat{\varepsilon}_t = Y_t - \hat{\beta} Y_{t-1}, \quad (2.28)$$

donde  $\hat{\beta} Y_{t-1}$ , como hemos dicho antes, es un estimador de la media condicional. Ejemplos de  $\hat{\beta}$  son los *Estimadores de Mínimos Cuadrados Ordinarios* o de *Máxima Verosimilitud*. Debido a la asunción de este supuesto, una de las características fundamentales de los tests de detección de heterocedasticidad que vamos a proponer en las siguientes secciones es que son *Tests de Naturaleza Semiparamétrica*. En la práctica, primeramente calcularemos un estimador  $T^{1/2}$ -consistente para  $\beta$ , el cual, normalmente, puede ser un estimador de tipo-OLS como clase particular de la clase más general de estimadores no lineales de mínimos cuadrados

$$\hat{\beta} = \arg \min_{\beta} T^{-1} \sum_{t=1}^T [Y_t - F(I_{t-1}, \theta)_t]^2. \quad (2.29)$$

Dicho estimador no toma en cuenta la posible estructura de la varianza, pero estima

consistentemente  $\beta$ .

Una vez proporcionado que la media condicional está correctamente especificada, se sigue que  $\hat{\varepsilon}_t \xrightarrow{p} \varepsilon_t$ , y por lo tanto es obvio que contrastar la hipótesis nula sobre  $E(\hat{\varepsilon}_t^2|I_{t-1})$  es asintóticamente equivalente a contrastar  $E(\varepsilon_t^2|I_{t-1})$ . En este caso las series temporales de los residuos estimados  $\hat{\varepsilon}_t^2$  son apropiadas para contrastar la hipótesis nula de homocedasticidad y consecuentemente ser analizadas desde el punto de vista de la Dinámica Simbólica. Como resultado de la estrategia presentada, construiremos nuestros tests focalizándonos en la serie de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$ .

En las siguientes secciones, con el fin de construir varios tests de heterocedasticidad, propondremos varias aplicaciones de simbolización estándar  $f$  para la nueva serie temporal de valores reales  $\{\hat{\varepsilon}_t^2\}_{t \in I}$ . En estas condiciones, nuestro objetivo es que cuando estamos bajo la hipótesis nula de homocedasticidad, al aplicar  $f$ , los símbolos resultantes son equiprobables. En cambio cuando aplicamos  $f$  bajo  $\mathbb{H}_1$ , la estructura de la varianza es dependiente y no todos los símbolos tienen la misma probabilidad de ocurrencia. Por ejemplo, es un hecho bien conocido que muchas series temporales económicas y financieras exhiben períodos de varianza inusualmente grande, seguida de períodos de relativa tranquilidad. Para estos procesos, el fenómeno conocido como “*clustering de la volatilidad*” se reflejará en probabilidades de ocurrencia distintas para cada símbolo, por lo que la hipótesis nula será rechazada.

## 2.6. Test Semi-Paramétrico de Heterocedasticidad basado en

### Permutaciones $G_{h_P}(m)$

A continuación vamos a construir nuestro primer test de heterocedasticidad mediante el uso de los patrones ordinales (permutaciones) y de la Entropía de Permutación; sabiendo de antemano que los patrones ordinales no parecen ser óptimos para el estudio de patrones relacionados con la estructura de la varianza ya que básicamente la varianza no está relacionada con los patrones ordinales. Como ya anticipamos anteriormente, Matilla y Ruiz [61] en la propuesta de su test  $G(m)$  utilizaron los patrones ordinales para estudiar la presencia de dependencia serial mediante la definición de una aplicación

de simbolización estándar  $f$  desde el espacio de los valores reales al de la Dinámica Simbólica; y después la entropía de permutación como medida asociada a la distribución de los símbolos resultantes de  $f$ .

Así, la característica fundamental de este test yace en encontrar algún tipo de estructura o dependencia serial de naturaleza lineal o no lineal en la serie de errores al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  del modelo 2.25 basándose en el análisis de los patrones ordinales generador por  $f$ , la aplicación de simbolización que describimos en la sección 2.3, y que volvemos a describir con más detalle aquí.

Sea  $m \geq 2$  un entero positivo, indicamos como  $S_m$  el grupo simétrico de orden  $m!$ , esto es, el grupo formado por todas las permutaciones de longitud  $m$ . Sea  $\pi_i = (i_1, i_2, \dots, i_m)$  un símbolo del grupo simétrico  $S_m$ . Ahora definimos un patrón ordinal para el símbolo  $\pi_i = (i_1, i_2, \dots, i_m) \in S_m$  en un tiempo dado  $t$ ,  $t \in I$ , es decir, una aplicación  $f$  que convierte  $\hat{\varepsilon}_m^2(t)$  en un símbolo  $\pi_i$  siendo :

$$\hat{\varepsilon}_m^2(t) = (\hat{\varepsilon}_{t+1}^2, \hat{\varepsilon}_{t+2}^2, \dots, \hat{\varepsilon}_{t+m}^2) \text{ para } t \in I$$

Entonces diremos que  $t \in I$  es de tipo- $\pi_i$  si y sólo si  $\pi_i = (i_1, i_2, \dots, i_m)$  es el único símbolo en el grupo  $S_m$  resultante de aplicar  $f$  bajo las 2 siguientes condiciones:

$$(a) \quad \hat{\varepsilon}_{t+1}^2 \leq \hat{\varepsilon}_{t+2}^2 \leq \dots \leq \hat{\varepsilon}_{t+m}^2, \text{ y}$$

$$(b) \quad i_{s-1} < i_s \text{ si } \hat{\varepsilon}_{t+i_{s-1}}^2 = \hat{\varepsilon}_{t+i_s}^2,$$

donde  $T$  es un conjunto de índices con  $i_s \in \{1, 2, \dots, n\}$ . Notar que las condiciones arriba mencionadas aseguran la univocidad de  $\hat{\varepsilon}_m^2(t)$  para todos los  $t \in I$ , esto significa que  $\hat{\varepsilon}_m^2(t)$  es convertida bajo  $f$  en un único símbolo  $\pi_i$ . Este símbolo describe como el ordenamiento de los índices temporales  $t + 0 < t + 1 < \dots < t + (m - 1)$  es convertido en el ordenamiento de los valores de la serie temporal bajo escrutinio. Para ver esto, lo ilustraremos con un ejemplo práctico.

Tomemos como dimensión de inmersión  $m = 3$ . Entonces el grupo simétrico es

$$S_3 = \{(0, 1, 2), (0, 2, 1), (1, 0, 2), (1, 2, 0), (2, 0, 1), (2, 1, 0)\}$$

Consideremos ahora que tenemos la siguiente serie temporal de siete valores de residuos al cuadrado

$$\{\hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6, \hat{\varepsilon}_4^2 = 4, \hat{\varepsilon}_5^2 = 3, \hat{\varepsilon}_6^2 = 11, \hat{\varepsilon}_7^2 = 2\} \quad (2.30)$$

Entonces para  $t = 2$  tenemos que  $\hat{\varepsilon}_{t+2}^2 = 4 < \hat{\varepsilon}_{t+1}^2 = 6 < \hat{\varepsilon}_{t+0}^2 = 9$  y por lo tanto para el período  $t = 2$ , el símbolo es de tipo-(2,1,0).

La 3-historia  $\hat{\varepsilon}_3^2(1) = (\hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6)$  está representada por el símbolo (0, 2, 1);  $\hat{\varepsilon}_3^2(2) = (9, 6, 4)$  y  $\hat{\varepsilon}_3^2(3) = (6, 4, 3)$  están representados por el símbolo (2, 1, 0);  $\hat{\varepsilon}_3^2(4) = (4, 3, 11)$  está representada por el símbolo (1, 0, 2); y finalmente  $\hat{\varepsilon}_3^2(5) = (3, 11, 2)$  está representada por el símbolo (2, 0, 1). Así obtenemos que  $p((0, 1, 2)) = 0 = p((1, 2, 0))$ ,  $p((0, 2, 1)) = \frac{1}{5}$ ,  $p((1, 0, 2)) = \frac{1}{5}$ ,  $p((2, 0, 1)) = \frac{1}{5}$  y  $p((2, 1, 0)) = \frac{2}{5}$ . Y la entropía de permutación asociada es  $h(3) = -3\frac{1}{5}\ln(\frac{1}{5}) - \frac{2}{5}\ln(\frac{2}{5}) \approx 1.33$ .

Por lo tanto la aplicación de simbolización  $f_{h_P}:\mathbb{R}^m \rightarrow S$  queda definida como:

$$f_{h_P}(\hat{\varepsilon}_m^2(t)) = f_h(\hat{\varepsilon}_{t+i_1}^2, \hat{\varepsilon}_{t+i_2}^2, \dots, \hat{\varepsilon}_{t+i_m}^2) = \pi_m(t) = (i_1, i_2, \dots, i_m) \quad (2.31)$$

$$\text{siendo } \hat{\varepsilon}_{t+i_1}^2 \leq \hat{\varepsilon}_{t+i_2}^2 \leq \dots \leq \hat{\varepsilon}_{t+i_m}^2 \text{ y } i_{s-1} < i_s \text{ si } \hat{\varepsilon}_{t+i_{s-1}}^2 = \hat{\varepsilon}_{t+i_s}^2$$

donde  $\hat{\varepsilon}_t$  representa los residuos generados por un estimador consistente de los parámetros del modelo.

Para cualquier  $t$ , fijemos  $\hat{\varepsilon}_m^2(t) = (\hat{\varepsilon}_{t+i_1}^2, \hat{\varepsilon}_{t+i_2}^2, \dots, \hat{\varepsilon}_{t+i_m}^2)$ , entonces este procedimiento de simbolización establece una correspondencia, para cada  $t$ , entre cada  $\hat{\varepsilon}_{t+i_j}^2$  para  $i_j = i_1, \dots, i_m$  y un patrón de ordenación (símbolo) o permutación  $\pi_i = f_{h_P}(\hat{\varepsilon}_m^2(t)) = f_{h_P}(\hat{\varepsilon}_{t+1}^2, \hat{\varepsilon}_{t+2}^2, \dots, \hat{\varepsilon}_{t+m}^2) = (i_1, i_2, \dots, i_m)$ .

Con la aplicación de simbolización propuesta, bajo la hipótesis  $\mathbb{H}_0$  de errores homocedásticos, la varianza no puede predecirse y por lo tanto los  $\hat{\varepsilon}_{t+i_j}^2$  para  $i_j = i_1, \dots, i_m$  serán independientes. Como consecuencia de ello, todas las permutaciones  $\hat{\varepsilon}_m^2(t)$  resultantes del proceso de ordenación de los  $\hat{\varepsilon}_{t+i}^2$ , los  $\pi_i = f_{h_P}(\hat{\varepsilon}_m^2(t)) = (i_1, i_2, \dots, i_m)$  serán equiprobables y la aplicación de simbolización  $f_{h_P}$  será una *Aplicación de Simbolización*

*Estándar.* Por el contrario, bajo la hipótesis alternativa  $\mathbb{H}_1$  existirá dependencia serial entre los  $\widehat{\varepsilon}_{t+i}^2$ , y los  $\pi_i = f_{h_P}(\widehat{\varepsilon}_m^2(t)) = (i_1, i_2, \dots, i_m)$  ya no serán equiprobables.

Además bajo la aplicación  $f_{h_P}$ , el conjunto de símbolos resultantes tiene una cardinalidad de  $m!$  y un número de grados de libertad  $k = m! - 1$ <sup>2</sup>.

Por lo tanto, procediendo como en 2.4 obtenemos el estadístico  $G_{h_P}(m)$  basado en la aplicación de simbolización definida por  $f_{h_P} : \mathbb{R}^m \rightarrow S$ , bajo la hipótesis nula  $\mathbb{H}_0$  2.27:

$$\mathbb{H}_0 : \left\{ \widehat{\varepsilon}_t^2 \right\}_{t \in I} \text{ i.i.d} \quad (2.32)$$

contra cualquier otra alternativa heterocedástica:

$$\mathbb{H}_1 : \left\{ \widehat{\varepsilon}_t^2 \right\}_{t \in I} \text{ no i.i.d}$$

y que se expresa como

$$G_{h_P}(m) = 2(T - m + 1)[\ln(m!) - h(S = S_m)] \quad (2.33)$$

el cual se distribuye asintóticamente como una  $\chi_{m!-1}^2$  cuando 2.32 es verdadera. Donde  $T - m + 1 = R$  es la cardinalidad de los índices.

## 2.7. Test Semi-Paramétrico de Heterocedasticidad basado en la Mediana $G_{h_{Me}}(m)$

En esta sección construiremos un nuevo test de heterocedasticidad con la aplicación de simbolización estándar  $f$  definida por López et al. [55] que utilizaron para estudiar la presencia de dependencia en series de tipo espacial. Análogamente al artículo de López et al. [55] también usaremos la entropía simbólica 2.2 para medir la distribución asociada a los símbolos de la aplicación  $f$  obteniendo el estadístico asociado siguiendo el procedimiento descrito en 2.4.

La principal característica de este test consiste en encontrar algún tipo de estructura

---

<sup>2</sup>El número de grados de libertad es el resultado del número total de símbolos  $m!$  sometida a la condición  $\sum_{i=1}^{m!} p_i = 1$ .

en la serie de errores al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  del modelo 2.25 basándose en el análisis comparativo de los residuos al cuadrado adyacentes en el tiempo con respecto al estadístico de la Mediana. Para ello, proponemos en esta sección la siguiente aplicación de simbolización  $f$  de  $\mathbb{R}^m$  en  $\Gamma$ , siendo  $\Gamma = \{\pi_1, \pi_2, \dots, \pi_n\}$  un conjunto de  $n$  símbolos: Expresemos como  $M_e$  la mediana de la serie temporal  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  y sea

$$J_t = \begin{cases} 0 & \text{si } \hat{\varepsilon}_t^2 \leq M_e \\ 1 & \text{en otro caso} \end{cases}$$

Ahora definimos la función indicador

$$\delta_{tt+k} = \begin{cases} 0 & \text{si } J_t \neq J_{t+k} \\ 1 & \text{en otro caso} \end{cases}$$

Por lo que la aplicación de simbolización así descrita  $f_{h_{M_e}} : \mathbb{R}^m \rightarrow \Gamma$  se define como:

$$\pi_i(t) = f_{h_{M_e}}(\hat{\varepsilon}_m^2(t)) = f_{h_{M_e}}(\hat{\varepsilon}_t^2, \hat{\varepsilon}_{t+1}^2, \dots, \hat{\varepsilon}_{t+m-1}^2) = (\delta_{tt+1}, \delta_{tt+2}, \dots, \delta_{tt+m-1}) \quad (2.34)$$

donde  $\hat{\varepsilon}_t$  representa los residuos generados por un estimador consistente de los parámetros del modelo.

Si para cualquier  $t$ , expresamos  $\hat{\varepsilon}_m^2(t) = (\hat{\varepsilon}_t^2, \hat{\varepsilon}_{t+1}^2, \dots, \hat{\varepsilon}_{t+m-1}^2)$ , e indicamos como  $N_t = \{t+1, t+2, \dots, t+m-1\}$  los  $m-1$  residuos al cuadrado adyacentes de  $t+1$ , entonces este procedimiento de simbolización consiste en comparar para cada tiempo  $t$ , el valor de  $\hat{\varepsilon}_t^2$  con  $\hat{\varepsilon}_{t+k}^2$  para todos los  $t \in N_t$ . Entonces si  $J_t = J_{t+k}$  significa que  $\hat{\varepsilon}_t^2$  y  $\hat{\varepsilon}_{t+k}^2$  son menores que, o mayores que,  $M_e$  y así  $f_{h_{M_e}}(\hat{\varepsilon}_m^2(t)) = \pi_i$  tendrá el valor 1 en la entrada  $k$ . Ahora veamos como opera la aplicación de simbolización con el siguiente ejemplo: Tomemos como dimensión de inmersión  $m = 3$ . Entonces el conjunto de símbolos  $\Gamma_m$  es

$$\Gamma_3 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

Supongamos ahora que tenemos la serie temporal 2.30 de los siete valores de residuos al cuadrado del apartado 2.6.

$$\{\hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6, \hat{\varepsilon}_4^2 = 4, \hat{\varepsilon}_5^2 = 3, \hat{\varepsilon}_6^2 = 11, \hat{\varepsilon}_7^2 = 2\} \quad (2.35)$$

La mediana de la serie es  $M_e = 4$ . Entonces para  $t = 2$  tenemos que  $\hat{\varepsilon}_t^2 = 9 > 4$ ;  $\hat{\varepsilon}_{t+1}^2 = 6 > 4$ ;  $\hat{\varepsilon}_{t+2}^2 = 4 \leq 4$  y  $J_t = 1$ ;  $J_{t+1} = 1$ ;  $J_{t+2} = 0$  con lo que  $\delta_{(t)(t+1)} = 1$ ;  $\delta_{(t)(t+2)} = 0$  de manera que para el período  $t = 2$  es un símbolo de tipo-(1,0).

Por lo tanto tenemos que la 3-historia  $\hat{\varepsilon}_3^2(1) = (\hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6)$  está representada por el símbolo (0,0);  $\hat{\varepsilon}_3^2(2) = (9, 6, 4)$  está representada por el símbolo (1,0);  $\hat{\varepsilon}_3^2(3) = (6, 4, 3)$  está representada por el símbolo (0,0);  $\hat{\varepsilon}_3^2(4) = (4, 3, 11)$  está representada por el símbolo (1,0); y finalmente  $\hat{\varepsilon}_3^2(5) = (3, 11, 2)$  está representada por el símbolo (0,1). Así obtenemos que  $p((0,0)) = \frac{2}{5}$ ,  $p((0,1)) = \frac{1}{5}$ ,  $p((1,0)) = \frac{2}{5}$ ,  $p((1,1)) = 0$ . Y la entropía de símbolo asociada es  $h(3) = -\frac{1}{5}\ln(\frac{1}{5}) - \frac{2}{5}\ln(\frac{2}{5}) - \frac{2}{5}\ln(\frac{2}{5}) \approx -1.0995$ .

Así pues bajo la hipótesis nula de homocedasticidad de los errores  $\{\hat{\varepsilon}_t^2\}_{t \in I} \mathbb{H}_0$  2.27:

$$\mathbb{H}_0 : \{\hat{\varepsilon}_t^2\}_{t \in I} \text{ i.i.d} \quad (2.36)$$

y con la aplicación de simbolización sugerida  $f_{h_{M_e}}$ , la varianza no es predecible, por lo que no hay relación entre los  $\delta_{t_{i_1}t_{i_j}}$  y en consecuencia todos los símbolos  $\pi_i = f_{h_{M_e}}(\hat{\varepsilon}_m^2(t)) = (\delta_{t+i_1t+i_2}, \delta_{t+i_1t+i_3}, \dots, \delta_{t+i_1t+i_m})$  tienen la misma probabilidad de ocurrencia  $\frac{1}{2^{m-1}}$ . Sin embargo bajo la hipótesis alternativa  $\mathbb{H}_1$  de heterocedasticidad, la varianza es predecible y existirá dependencia serial entre los  $\delta_{t_{i_1}t_{i_j}}$ ,  $i_j = i_2, i_3, \dots, i_{m-1}$ . De esta forma la aplicación de simbolización  $f_{h_{M_e}}$  es una *Aplicación de Simbolización Estándar*.

Con esta aplicación  $f_{h_{M_e}}$ , el conjunto de símbolos tiene una cardinalidad de  $n = 2^{m-1}$  y un número de grados de libertad  $k = n - 1$ .

Siguiendo el proceso descrito en el apartado 2.4 con la aplicación de simbolización definida por  $f_{h_{M_e}} : \mathbb{R}^m \rightarrow \Gamma$ , obtenemos el estadístico  $G_{h_{M_e}}(m)$  cuya expresión es

$$G_{h_{M_e}}(m) = 2(T - m + 1)[\ln(2^{m-1}) - h(\Gamma)] \quad (2.37)$$

que se distribuye asintóticamente como una  $\chi_{2^{m-1}-1}^2$  cuando 2.36 es verdadera.

## 2.8. Test Semi-Paramétrico de Heterocedasticidad basado en los N-Cuantiles $G_{h_{NQ}}(m)$

Esta propuesta de test de heterocedasticidad se basa, a diferencia de las dos anteriores, en una aplicación  $f$  de simbolización estándar novedosa. De modo similar a los dos tests anteriores, a los símbolos resultantes de la aplicación  $f$ , les aplicamos la entropía simbólica 2.2 para medir su distribución y obtener un estadístico siguiendo el procedimiento descrito en el apartado 2.4.

La característica primordial de este test radica en encontrar algún tipo de estructura lineal o no lineal en la serie de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  del modelo 2.25 basándose en la clasificación de los residuos al cuadrado en función de la partición de su distribución estadística a través de sus cuantiles. Para ello, proponemos la siguiente aplicación de simbolización  $f_{h_{NQ}}$  de  $\mathbb{R}^m$  en  $\Gamma$ , siendo  $\Gamma = \{\pi_1, \pi_2, \dots, \pi_n\}$  un conjunto de  $n$  símbolos:

Denotemos por  $C_1, \dots, C_{N-1}$  los  $N-1$  cuantiles que dividen los valores de la serie temporal de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  en  $N_C$  intervalos regulares.

Ahora definimos la función

$$Q(\hat{\varepsilon}_t^2) = \begin{cases} 1 & \text{si } 0 \leq \hat{\varepsilon}_t^2 \leq C_1 \\ z & \text{si } C_{z-1} \leq \hat{\varepsilon}_t^2 \leq C_z \quad z = 2, \dots, N-1 \\ N & \text{si } C_{N-1} \leq \hat{\varepsilon}_t^2 \leq +\infty \end{cases} \quad (2.38)$$

ya que el dominio de definición de  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  es  $[0, +\infty)$ .

Entonces, la aplicación de simbolización  $f_{h_{NQ}} : \mathbb{R}^m \rightarrow \Gamma$  se define como:

$$\pi_i(t) = f_{h_{NQ}}(\hat{\varepsilon}_m^2(t)) = f_h(\hat{\varepsilon}_{t+i_1}^2, \hat{\varepsilon}_{t+i_2}^2, \dots, \hat{\varepsilon}_{t+i_m}^2) = (\hat{Q}(\hat{\varepsilon}_{t+i_1}^2), \hat{Q}(\hat{\varepsilon}_{t+i_2}^2), \dots, \hat{Q}(\hat{\varepsilon}_{t+i_m}^2)), \quad (2.39)$$

donde  $\hat{\varepsilon}_t$  representa los residuos generados por un estimador consistente de los parámetros del modelo.

Si para cualquier  $t$  expresamos  $\hat{\varepsilon}_m^2(t) = (\hat{\varepsilon}_{t+i_1}^2, \hat{\varepsilon}_{t+i_2}^2, \dots, \hat{\varepsilon}_{t+i_m}^2)$ , entonces este procedimiento de simbolización establece una correspondencia, para cada  $t$ , entre cada  $\hat{\varepsilon}_{t+i_j}^2$  para  $i_j = i_1, \dots, i_m$  y un intervalo (symbol  $\pi_i$ ) de los  $N_C$  definido por los  $\hat{C}_1, \dots, \hat{C}_{N-1}$ ,

los cuales son cuantiles muestrales de  $\{\hat{\varepsilon}_t^2\}_{t \in I}$ .

Vamos a ver cómo opera la aplicación de simbolización con el siguiente ejemplo:

Tomemos como dimensión de inmersión  $m = 3$  y  $N_C = 3$ . Entonces el conjunto de símbolos  $\Gamma_m$  es

$$\Gamma_3 = \left\{ \begin{array}{l} (1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 2, 1), (1, 2, 2), (1, 2, 3), \\ (1, 3, 1), (1, 3, 2), (1, 3, 3), (2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 2, 1), \\ (2, 2, 2), (2, 2, 3), (2, 3, 1), (2, 3, 2), (2, 3, 3), (3, 1, 1), (3, 1, 2), \\ (3, 1, 3), (3, 2, 1), (3, 2, 2), (3, 2, 3), (3, 3, 1), (3, 3, 2), (3, 3, 3) \end{array} \right\}$$

Ahora consideremos que tenemos la siguiente serie temporal de veinte y nueve valores de residuos al cuadrado obtenidos de la regresión OLS de una serie temporal objeto de análisis  $Y_t$ .

$$\left\{ \begin{array}{l} \hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6, \hat{\varepsilon}_4^2 = 4, \hat{\varepsilon}_5^2 = 3, \hat{\varepsilon}_6^2 = 11, \hat{\varepsilon}_7^2 = 2, \\ \hat{\varepsilon}_8^2 = 4, \hat{\varepsilon}_9^2 = 2, \hat{\varepsilon}_{10}^2 = 10, \hat{\varepsilon}_{11}^2 = 14, \hat{\varepsilon}_{12}^2 = 13, \hat{\varepsilon}_{13}^2 = 25, \hat{\varepsilon}_{14}^2 = 22, \hat{\varepsilon}_{15}^2 = 18, \\ \hat{\varepsilon}_{16}^2 = 14, \hat{\varepsilon}_{17}^2 = 6, \hat{\varepsilon}_{18}^2 = 3, \hat{\varepsilon}_{19}^2 = 2, \hat{\varepsilon}_{20}^2 = 5, \hat{\varepsilon}_{21}^2 = 2, \hat{\varepsilon}_{22}^2 = 4, \\ \hat{\varepsilon}_{23}^2 = 8, \hat{\varepsilon}_{24}^2 = 10, \hat{\varepsilon}_{25}^2 = 15, \hat{\varepsilon}_{26}^2 = 23, \hat{\varepsilon}_{27}^2 = 25, \hat{\varepsilon}_{28}^2 = 22, \hat{\varepsilon}_{29}^2 = 16 \end{array} \right\} \quad (2.40)$$

Los dos cuantiles muestrales  $\hat{C}_1, \hat{C}_2$  que dividen la distribución empírica de la serie en tres intervalos  $N_3$  son  $[\hat{C}_1 = 4, \hat{C}_2 = 13.6]$  y los intervalos correspondientes a dichos cuantiles son  $[0, 4), [4, 13.6)$  y  $[13.6, 25]$ . Entonces para  $t = 2$  tenemos que  $4 < \hat{\varepsilon}_t^2 = 9 < 13.6$ ;  $4 < \hat{\varepsilon}_{t+1}^2 = 6 < 13.6$ ;  $4 \leq \hat{\varepsilon}_{t+2}^2 = 4 < 13.6$  y  $Q_t = 2$ ;  $Q_{t+1} = 2$ ;  $Q_{t+2} = 2$  así que tenemos que para el período  $t = 2$  es de tipo-(2,2,2).

En la Tabla 2.2 mostramos las 3-historias de la serie y los símbolos que las representan obteniendo de esta manera las probabilidades de símbolos que se representan en la Tabla 2.3. Así pues, la entropía de símbolo asociada a las probabilidades calculadas es

$$h(3) = 9 \frac{1}{27} \ln\left(\frac{1}{27}\right) - 3 \frac{2}{27} \ln\left(\frac{2}{27}\right) - \frac{3}{27} \ln\left(\frac{3}{27}\right) - \frac{4}{27} \ln\left(\frac{4}{27}\right) - \frac{5}{27} \ln\left(\frac{5}{27}\right) \approx 2.51.$$

En consecuencia, bajo la hipótesis nula de homocedasticidad de los errores homocedásticos  $\{\hat{\varepsilon}_t^2\}_{t \in I}$   $\mathbb{H}_0$  (Ecuación 2.27)

$$\mathbb{H}_0 : \{\hat{\varepsilon}_t^2\}_{t \in I} \text{ i.i.d} \quad (2.41)$$

Tabla 2.2: *Ejemplo de m-historias y Símbolos*

<i>nombre m-Historia</i>	<i>m – Historia</i>	<i>Símbolo</i>
$\hat{\varepsilon}_3^2(1)$	$(\hat{\varepsilon}_1^2 = 1, \hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6)$	$(1, 2, 2)$
$\hat{\varepsilon}_3^2(2)$	$(\hat{\varepsilon}_2^2 = 9, \hat{\varepsilon}_3^2 = 6, \hat{\varepsilon}_4^2 = 4)$	$(2, 2, 2)$
$\hat{\varepsilon}_3^2(3)$	$(\hat{\varepsilon}_3^2 = 6, \hat{\varepsilon}_4^2 = 4, \hat{\varepsilon}_5^2 = 3)$	$(2, 2, 1)$
$\hat{\varepsilon}_3^2(4)$	$(\hat{\varepsilon}_4^2 = 4, \hat{\varepsilon}_5^2 = 3, \hat{\varepsilon}_6^2 = 11)$	$(2, 1, 2)$
$\hat{\varepsilon}_3^2(5)$	$(\hat{\varepsilon}_5^2 = 3, \hat{\varepsilon}_6^2 = 11, \hat{\varepsilon}_7^2 = 2)$	$(1, 2, 1)$
$\hat{\varepsilon}_3^2(6)$	$(\hat{\varepsilon}_6^2 = 11, \hat{\varepsilon}_7^2 = 2, \hat{\varepsilon}_8^2 = 4)$	$(2, 1, 2)$
$\hat{\varepsilon}_3^2(7)$	$(\hat{\varepsilon}_7^2 = 2, \hat{\varepsilon}_8^2 = 4, \hat{\varepsilon}_9^2 = 2)$	$(1, 2, 1)$
$\hat{\varepsilon}_3^2(8)$	$(\hat{\varepsilon}_8^2 = 4, \hat{\varepsilon}_9^2 = 2, \hat{\varepsilon}_{10}^2 = 10)$	$(2, 1, 2)$
$\hat{\varepsilon}_3^2(9)$	$(\hat{\varepsilon}_9^2 = 2, \hat{\varepsilon}_{10}^2 = 10, \hat{\varepsilon}_{11}^2 = 14)$	$(1, 2, 3)$
$\hat{\varepsilon}_3^2(10)$	$(\hat{\varepsilon}_{10}^2 = 10, \hat{\varepsilon}_{11}^2 = 14, \hat{\varepsilon}_{12}^2 = 13)$	$(2, 3, 2)$
$\hat{\varepsilon}_3^2(11)$	$(\hat{\varepsilon}_{11}^2 = 14, \hat{\varepsilon}_{12}^2 = 13, \hat{\varepsilon}_{13}^2 = 25)$	$(3, 2, 3)$
$\hat{\varepsilon}_3^2(12)$	$(\hat{\varepsilon}_{12}^2 = 13, \hat{\varepsilon}_{13}^2 = 25, \hat{\varepsilon}_{14}^2 = 22)$	$(2, 3, 3)$
$\hat{\varepsilon}_3^2(13)$	$(\hat{\varepsilon}_{13}^2 = 25, \hat{\varepsilon}_{14}^2 = 22, \hat{\varepsilon}_{15}^2 = 18)$	$(3, 3, 3)$
$\hat{\varepsilon}_3^2(14)$	$(\hat{\varepsilon}_{14}^2 = 22, \hat{\varepsilon}_{15}^2 = 18, \hat{\varepsilon}_{16}^2 = 14)$	$(3, 3, 3)$
$\hat{\varepsilon}_3^2(15)$	$(\hat{\varepsilon}_{15}^2 = 18, \hat{\varepsilon}_{16}^2 = 14, \hat{\varepsilon}_{17}^2 = 6)$	$(3, 3, 2)$
$\hat{\varepsilon}_3^2(16)$	$(\hat{\varepsilon}_{16}^2 = 14, \hat{\varepsilon}_{17}^2 = 6, \hat{\varepsilon}_{18}^2 = 3)$	$(3, 2, 1)$
$\hat{\varepsilon}_3^2(17)$	$(\hat{\varepsilon}_{17}^2 = 6, \hat{\varepsilon}_{18}^2 = 3, \hat{\varepsilon}_{19}^2 = 2)$	$(2, 1, 1)$
$\hat{\varepsilon}_3^2(18)$	$(\hat{\varepsilon}_{18}^2 = 3, \hat{\varepsilon}_{19}^2 = 2, \hat{\varepsilon}_{20}^2 = 5)$	$(1, 1, 2)$
$\hat{\varepsilon}_3^2(19)$	$(\hat{\varepsilon}_{19}^2 = 2, \hat{\varepsilon}_{20}^2 = 5, \hat{\varepsilon}_{21}^2 = 2)$	$(1, 2, 1)$
$\hat{\varepsilon}_3^2(20)$	$(\hat{\varepsilon}_{20}^2 = 5, \hat{\varepsilon}_{21}^2 = 2, \hat{\varepsilon}_{22}^2 = 4)$	$(2, 1, 2)$
$\hat{\varepsilon}_3^2(21)$	$(\hat{\varepsilon}_{21}^2 = 2, \hat{\varepsilon}_{22}^2 = 4, \hat{\varepsilon}_{23}^2 = 8)$	$(1, 2, 2)$
$\hat{\varepsilon}_3^2(22)$	$(\hat{\varepsilon}_{22}^2 = 4, \hat{\varepsilon}_{23}^2 = 8, \hat{\varepsilon}_{24}^2 = 10)$	$(2, 2, 2)$
$\hat{\varepsilon}_3^2(23)$	$(\hat{\varepsilon}_{23}^2 = 8, \hat{\varepsilon}_{24}^2 = 10, \hat{\varepsilon}_{25}^2 = 15)$	$(2, 2, 3)$
$\hat{\varepsilon}_3^2(24)$	$(\hat{\varepsilon}_{24}^2 = 10, \hat{\varepsilon}_{25}^2 = 15, \hat{\varepsilon}_{26}^2 = 23)$	$(2, 3, 3)$
$\hat{\varepsilon}_3^2(25)$	$(\hat{\varepsilon}_{25}^2 = 15, \hat{\varepsilon}_{26}^2 = 23, \hat{\varepsilon}_{27}^2 = 25)$	$(3, 3, 3)$
$\hat{\varepsilon}_3^2(26)$	$(\hat{\varepsilon}_{26}^2 = 23, \hat{\varepsilon}_{27}^2 = 25, \hat{\varepsilon}_{28}^2 = 22)$	$(3, 3, 3)$
$\hat{\varepsilon}_3^2(27)$	$(\hat{\varepsilon}_{27}^2 = 25, \hat{\varepsilon}_{28}^2 = 22, \hat{\varepsilon}_{29}^2 = 16)$	$(3, 3, 3)$

Tabla 2.3: *Símbolos y Probabilidades de Símbolo 2.40.*

<i>Símbolo</i>	<i>Probabilidad de Símbolo</i>
(1,1,1)	0
(1,1,2)	$\frac{1}{27}$
(1,1,3)	0
(1,2,1)	$\frac{3}{27}$
(1,2,2)	$\frac{2}{27}$
(1,2,3)	$\frac{1}{27}$
(1,3,1)	0
(1,3,2)	0
(1,3,3)	0
(2,1,1)	$\frac{1}{27}$
(2,1,2)	$\frac{4}{27}$
(2,1,3)	0
(2,2,1)	$\frac{1}{27}$
(2,2,2)	$\frac{2}{27}$
(2,2,3)	$\frac{1}{27}$
(2,3,1)	0
(2,3,2)	$\frac{1}{27}$
(2,3,3)	$\frac{2}{27}$
(3,1,1)	0
(3,1,2)	0
(3,1,3)	0
(3,2,1)	$\frac{1}{27}$
(3,2,2)	0
(3,2,3)	$\frac{1}{27}$
(3,3,1)	0
(3,3,2)	$\frac{1}{27}$
(3,3,3)	$\frac{5}{27}$

cada  $\hat{\varepsilon}_{t+i_j}^2$  caerá en cada cuantil con la misma probabilidad para  $i_j = i_1, \dots, i_m$ , y entonces cada símbolo tendrá la misma probabilidad de ocurrencia. En este sentido la aplicación de simbolización  $f_{h_{NQ}}$ , al igual que las aplicaciones  $f_{h_{Me}}$  y  $f_{h_P}$  definidas en 2.6 y 2.7 respectivamente, es una *Aplicación de Simbolización Estándar*. En cambio, cuando estamos bajo la hipótesis alternativa, esto es  $\hat{h}_t = g(\hat{\varepsilon}_t^2)$ , o explícitamente,

$$\mathbb{H}_1 : \left\{ \hat{\varepsilon}_t^2 \right\}_{t \in I} \text{ no } i.i.d$$

no todos los símbolos tendrán la misma probabilidad de ocurrencia ya que hay una dependencia de la estructura lineal o no lineal en la varianza..

Proponiendo un valor del número de intervalos  $N_C = 2$ , los valores de  $\left\{ \hat{\varepsilon}_t^2 \right\}_{t \in I}$  que estarán por encima de  $C_1 = M_e(\left\{ \hat{\varepsilon}_t^2 \right\})$  (siendo  $M_e(x)$  la mediana de  $x$ ) y los que estarán por debajo del Cuantil  $C_1$  que divide los dos intervalos, ocurrirán más o menos con la misma frecuencia. Este número de intervalos no nos permite diferenciar los diferentes tipos de “actividades” propias de las series heterocedásticas. Sin embargo, si seleccionamos un número de intervalos  $N_C = 3$  podemos discriminar, a nivel simbólico, entre valores de  $\left\{ \hat{\varepsilon}_t^2 \right\}_{t \in I}$  que indican actividades de extrema tranquilidad, valores propios de una actividad “normal” de la volatilidad y valores extremos (cola) de la distribución de  $\left\{ \hat{\varepsilon}_t^2 \right\}_{t \in I}$ . De este modo tres intervalos deberían ser suficientes para capturar las desviaciones de la hipótesis nula.

De igual forma que en las secciones anteriores, el test obtenido de esta aplicación de simbolización definida por  $f_{h_{NQ}} : \mathbb{R}^m \rightarrow \Gamma$  es:

$$G_{h_{NQ}}(m) = 2(T - m + 1)[\ln(3^m) - h(\Gamma)] \quad (2.42)$$

que se distribuye asintóticamente como una  $\chi_{3^m-1}^2$  cuando 2.41 es verdadera.

Notar que con esta aplicación de simbolización el conjunto de símbolos tiene una cardinalidad de  $n = 3^m$  y  $k = n - 1$  (número de símbolos menos uno) grados de libertad.

### 2.8.1. Selección práctica del parámetro $m$ y el tamaño de muestra $T$

El hecho de cuantizar el espacio de estados mediante una aplicación de simbolización de  $m$  símbolos, conlleva una pérdida irreversible de información; aunque nuestro objetivo

es que esta pérdida sea la menor posible teniendo en cuenta la característica estocástica concreta que se quiere capturar en cada caso. Con el fin de mantener este coste de pérdida de información a niveles bajos y usar toda la información disponible proponemos considerar que  $t$  cubre todos los enteros del conjunto  $1 \leq t \leq T - m + 1$ . Esto implica que, a pesar del hecho que  $Y_t$  e  $Y_{t+k}$  son independientes si  $k \neq 1, 2, \dots, (m-1)$ , ahora la variable  $Y_t$  está correlada con  $Y_{t+1}, Y_{t+2}, \dots, Y_{t+(m-1)}$  para  $m = 3, 4, 5, \dots$ . Si  $m$  es pequeño comparado con  $T - m + 1$ , las dependencias entre  $Y_t$  y  $Y_{t+k}$  son débiles y pueden ser despreciadas. Parcialmente por esta razón recomendamos seleccionar un valor de  $m$  pequeño ( $m \leq 6$ ) y siempre en relación con el tamaño de muestra  $T$  (véase López et al. [55] y el apartado 2.4.2).

Además de lo mencionado anteriormente, es preciso percatarse que el tamaño de muestra  $T$  debe ser más grande que el número de símbolos necesario para la construcción de los tests, es decir, se debe tener al menos el mismo número de  $m - \text{historias}$  que símbolos posibles. Ahora bien, cuando el test  $\chi^2_\alpha$  es aplicado en práctica, y todas las frecuencias esperadas son  $\geq 5$ , la distribución tabulada  $\chi^2$ , da como regla, el valor de  $\chi^2_\alpha$  con una aproximación suficiente para los propósitos ordinarios (véase Rothagi [69]). Por esta razón, requerimos trabajar con conjunto de datos que contienen al menos 5 veces el número posible de símbolos. Por ejemplo, un conjunto de datos de 200 observaciones es el conjunto de datos más pequeño que se puede considerar para una dimensión de inmersión de  $m = 5$  con el test  $G_{h_P}(m)$ , 160 con el test  $G_{h_{M_e}}(m)$  y 1215 con el test  $G_{h_{N_Q}}(m)$ . Más allá de  $m = 6$  los requerimientos de datos no son realistas para series temporales económicas, así que no usaremos tales dimensiones. En general para un conjunto de  $T$  observaciones, la dimensión de inmersión será el  $m$  más grande que satisface  $5m! \leq T$ ,  $5 \cdot 2^{m-1} \leq T$ ,  $5 \cdot 3^m \leq T$ , respectivamente para los tres tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$ .

## 2.9. Pruebas Empíricas

En esta sección investigaremos el rendimiento finito muestral de los tests simbólicos de heterocedasticidad  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$ . Nuestro primer objetivo es in-

investigar si la distribución asintótica es una aproximación adecuada a la distribución muestral finita bajo la hipótesis nula. Para analizar el rendimiento finito muestral de los estadísticos hemos diseñado algunos experimentos de Monte Carlo: se han creado procesos generadores de datos (DGP's)<sup>3</sup> de tamaños muestrales ( $T$ ) de 100, 500 y 1000 observaciones. Estos tamaños muestrales son compatibles con las longitudes habituales encontradas en series temporales económicas. Hemos usado la dimensión de inmersión de  $m = 4$ , excepto para el tamaño más pequeño de muestra ( $T = 100$ ), para el que hemos utilizado  $m = 3$ . Para cada DGP este procedimiento se repitió 1000 veces y la proporción de rechazos de la hipótesis nula de homocedasticidad fue calculada usando un nivel de significación del 5%. El segundo objetivo es comparar nuestro test más competitivo desde el punto de vista de tamaño y potencia con algunos de los tests de heterocedasticidad de espíritu similar más populares, concretamente los tests de White [84], Breuch-Pagan ([19]), Goldfeld-Quandt [38] y Engle[29].

### 2.9.1. Simulaciones

#### 2.9.1.1. Tamaño

Para examinar el tamaño estadístico de los tests bajo  $\mathbb{H}_0$ , consideramos varios DGPs. Dada la naturaleza estocástica de los datos financieros y económicos, tomamos en cuenta el siguiente modelo general:

$$M(X_t, \varepsilon_t) : Y_t = a + bX_t + \varepsilon_t$$

con  $a = b = 1$  para el que construimos 16 DGP's (S.1-S.15), donde cada uno es un modelo diferente compatible con la hipótesis nula. Es un hecho estudiado y comprobado que los shocks externos del mercado (llamados  $\varepsilon_t$ ) generalmente no siguen una distribución normal. Por esta razón hemos utilizado varios tipos de shocks estocásticos como por ejemplo los que obedecen a las distribuciones *Normal*, *t-Student*, *Chi-Cuadrado*, *Uniforme* y también la *Log-Normal*. También hemos usado para caracterizar la variable independiente  $X_t$ , las siguientes distribuciones: una distribución *Chi-Cuadrado*, una

---

<sup>3</sup>Nosotros proponemos usar solapamiento de las  $m - historias$  dada la relativa escasez del número de observaciones.

Tabla 2.4: Modelos bajo la Hipótesis Nula

DGP: $Y_t = a + bX_t + \varepsilon_t$	Variable Independiente $X_t$	Innovaciones, $\varepsilon_t$
S.1	$\chi_1^2$	$N(0, 1)$
S.2	$\chi_1^2$	$\chi_4^2$
S.3	$\chi_1^2$	$Log - N(0, 1)$
S.4	$\chi_1^2$	$t_4$
S.5	$\chi_1^2$	$U(-1, 1)$
S.6	$U(1, 31)$	$N(0, 1)$
S.7	$U(1, 31)$	$\chi_4^2$
S.8	$U(1, 31)$	$Log - N(0, 1)$
S.9	$U(1, 31)$	$t_4$
S.10	$U(1, 31)$	$U(-1, 1)$
S.11	$Log - N(3, 1)$	$N(0, 1)$
S.12	$Log - N(3, 1)$	$\chi_4^2$
S.13	$Log - N(3, 1)$	$Log - N(0, 1)$
S.14	$Log - N(3, 1)$	$t_4$
S.15	$Log - N(3, 1)$	$U(-1, 1)$
S.16	$0,5y_{t-1}$	$N(0, 1)$

distribución *Uniforme* en el intervalo (1,31), y una distribución *Log-Normal* con media tres y varianza unidad. Finalmente hemos añadido una estructura autorregresiva, esto es  $X_t = f(Y_{t-1})$ , en la variable independiente del modelo estudiado (DGP S.16). La varianza de este proceso autorregresivo es homocedástica y por eso está también bajo el supuesto de la hipótesis nula. Todos estos DGP's se describen en la Tabla 2.4.

La Tablas 2.5, 2.6 y 2.7 reportan los tamaños empíricos de los tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  bajo dieciséis configuraciones de la  $\mathbb{H}_0$  para un nivel del 5%. Hay varios hechos que podemos observar de este experimento. Primeramente, el tamaño del test  $G_{h_{N_Q}}(m)$  está bien controlado independientemente de la característica estocástica de los shocks y de la variable explicativa,  $X_t$  para  $m = 3$ ,  $T = 100$  y  $m = 4$ ,  $T = 500$ ,  $T = 1000$ . Como podemos ver, los valores de probabilidad están casi siempre en el intervalo [0.036, 0.064], que se corresponden con una intervalo de confianza del 95% para un tamaño del 5% y 1000 iteraciones. En cambio, para esta configuración de los parámetros de simulación, los tests  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$ , tienden a rechazar fuertemente por encima del tamaño nominal; siendo esto más evidente en el test  $G_{h_{M_e}}(m)$ . Una explicación para estas elevadas tasas de rechazo puede ser la alta correlación inducida

Tabla 2.5: Tamaño Empírico de  $G_{h_P}(m)$ 

	T=100		T=500	T=1000
	$m = 3$	$m = 4$	$m = 4$	$m = 4$
S.1	0.041	0.071	0.041	0.043
S.2	0.03	0.067	0.048	0.048
S.3	0.03	0.071	0.049	0.04
S.4	0.036	0.068	0.053	0.059
S.5	0.043	0.075	0.051	0.045
S.6	0.042	0.064	0.041	0.04
S.7	0.032	0.065	0.035	0.044
S.8	0.031	0.07	0.043	0.032
S.9	0.028	0.052	0.051	0.039
S.10	0.038	0.075	0.086	0.065
S.11	0.045	0.085	0.084	0.092
S.12	0.047	0.084	0.077	0.081
S.13	0.05	0.088	0.08	0.092
S.14	0.05	0.084	0.075	0.086
S.15	0.049	0.083	0.084	0.072
S.16	0.032	0.06	0.049	0.05

entre  $Y_t$  e  $Y_{t+k}$  por las aplicaciones de simbolización  $f_{h_P}$  y  $f_{h_{M_e}}$  incluso cumpliendo las condiciones de selección de tamaño de símbolo  $m$  en relación con el tamaño de muestra  $T$  (ver apartado 2.8.1). Estudiaremos el efecto comparativo del grado de solapamiento de los símbolos para los tres tests en la sección 2.10.

Por último señalar que los resultados de tamaño para el test  $G_{h_{N_Q}}(m)$  siempre están por encima del 10% con  $m = 4$  y  $T = 100$ . Esto es debido a que el tamaño de muestra es pequeño en comparación con el número de símbolos, ya que no se cumple la regla general de selección del tamaño de muestra  $5 \cdot (3^m - 1) \leq T \rightarrow 5 \cdot (3^4 - 1) > 100$  descrita en el apartado 2.8.1.

Tabla 2.6: Tamaño Empírico de  $G_{h_{M_e}}(m)$ 

	T=100		T=500	T=1000
	$m = 3$	$m = 4$	$m = 4$	$m = 4$
S.1	0.093	0.116	0.1	0.097
S.2	0.08	0.099	0.094	0.114
S.3	0.081	0.113	0.092	0.107
S.4	0.086	0.117	0.112	0.095
S.5	0.086	0.112	0.114	0.097
S.6	0.109	0.114	0.092	0.106
S.7	0.091	0.112	0.094	0.116
S.8	0.089	0.105	0.087	0.119
S.9	0.1	0.1020	0.104	0.08
S.10	0.097	0.112	0.113	0.097
S.11	0.06	0.145	0.09	0.094
S.12	0.074	0.116	0.09	0.122
S.13	0.109	0.099	0.128	0.108
S.14	0.087	0.066	0.105	0.12
S.15	0.076	0.142	0.119	0.107
S.16	0.054	0.099	0.085	0.089

Tabla 2.7: Tamaño Empírico de  $G_{h_{NQ}}(m)$ 

	T=100		T=500	T=1000
	$m = 3$	$m = 4$	$m = 4$	$m = 4$
S.1	0.026	0.11	0.053	0.041
S.2	0.033	0.129	0.039	0.042
S.3	0.03	0.123	0.044	0.037
S.4	0.029	0.148	0.054	0.05
S.5	0.029	0.139	0.042	0.049
S.6	0.027	0.12	0.048	0.042
S.7	0.02	0.131	0.054	0.033
S.8	0.032	0.131	0.043	0.043
S.9	0.028	0.135	0.038	0.048
S.10	0.027	0.131	0.038	0.049
S.11	0.022	0.153	0.047	0.047
S.12	0.023	0.143	0.063	0.053
S.13	0.032	0.133	0.060	0.051
S.14	0.028	0.155	0.056	0.050
S.15	0.037	0.141	0.050	0.051
S.16	0.03	0.128	0.034	0.044

### 2.9.1.2. Potencia

A continuación examinamos la potencia de los tres tests, es decir, la probabilidad de rechazar  $\mathbb{H}_0$  cuando es falsa (i.e, cuando hay heterocedasticidad). Intentamos evaluar los tests en diferentes contextos que son relativamente frecuentes en series temporales típicas financieras. Para este fin hemos considerado 25 procesos (P1-P25) los cuales se recogen en la Tabla 2.8. Los primeros 15 procesos comparten las mismas estructuras y difieren en el tipo de shocks  $\varepsilon_t$ , bajo cada estructura. La base de la construcción de estas estructuras son los modelos autorregresivos de heterocedasticidad condicional (ARCH) introducidos por Engle, y su extensión GARCH (ARCH generalizado) debida a [14]. En estos modelos, el concepto clave es la varianza condicional, es decir, la varianza condicional del pasado. Esta especificación particular es capaz de capturar los principales hechos estudiados que caracterizan a las series financieras. En los modelos clásicos GARCH, la varianza condicional está expresada como una función lineal de los valores cuadráticos del pasado de la serie. A este respecto, consideramos los modelos bien conocidos ARCH(1), GARCH(1,1) y también estudiamos el modelo ARGARCH(1,1,1) donde la media del proceso es un modelo autorregresivo y la varianza es heterocedástica del tipo GARCH(1,1). La parte lineal AR(1) está correctamente especificada para  $E(Y_t | I_{t-1}) = \alpha_1 Y_{t-1}$  y el parámetro autorregresivo puede ser estimado consistentemente a través de su estimador OLS.

Para evaluar el impacto de los shocks estocásticos en el comportamiento de los tests bajo muestra finita, para cada uno de estos tres modelos consideramos 5 tipos de shocks estocásticos:  $N(0, 1)$ ,  $\chi_1^2$ ,  $t_5$ ,  $\text{Log} - N(0, 0,75)$  y  $U(-1, 1)$ .

También es frecuente que la media de la serie temporal bajo estudio dependa de su propia media condicional. Por ejemplo, los agentes con aversión al riesgo requerirán que se les compense por poseer un valor de riesgo. Dado que el riesgo de un valor se puede medir mediante la varianza de los rendimientos, la prima de riesgo será una función creciente de la varianza condicional de los rendimientos. Esta idea se puede capturar mediante los modelos ARCH-M como se refleja en el modelo P.16.

Una característica fundamental de los precios de los valores es que las “malas” noticias

parecen tener un efecto más pronunciado en la volatilidad que la que tienen las “buenas” noticias, así hay algún grado de asimetría respecto de los shocks o innovaciones. Muchos estudios empíricos sobre las series de stocks muestran una correlación negativa entre las innovaciones actuales al cuadrado y las innovaciones pasadas: si la distribución condicional fuera simétrica en las variables pasadas, tal correlación sería igual a cero. Sin embargo la asimetría condicional es un hecho comprobado: la volatilidad aumenta debido a un descenso del precio tiene un efecto generalmente más fuerte que un aumento del precios de la misma magnitud. También estamos interesados en estudiar el comportamiento del test en presencia de asimetrías. Para este fin consideramos P17 y P18.

Los procesos del P.19 al P.25 pertenecen al tipo de modelos de volatilidad estocástica (modelos SV, propuestos por Taylor [78]). Al igual que los modelos GARCH, son utilizados para modelizar series de volatilidad no constante, pero se diferencian de ellos por que consideran que la varianza condicionada es en sí misma un proceso aleatorio. Uno de las principales ventajas de los modelos de volatilidad estocástica respecto a los modelos GARCH es que consiguen una especificación más sencilla (pues contienen menos parámetros) pero en cambio, el proceso de estimación es más dificultoso pues requiere métodos de estimación de máxima verosimilitud.

Los modelos P.19 a P.22 son los llamados procesos de Volatilidad Estocástica con Memoria Corta (ARSV). La ecuación estructural de estos modelos es igual a la de los modelos GARCH  $y_t = u_t$ ,  $u_t = \sigma_t \varepsilon_t$ , pero la diferencia es que la varianza ya no depende de los factores observables sino que se supone que el logaritmo de la varianza condicionada sigue un proceso lineal AR(p).

Los resultados de los trabajos de algunos autores, sugieren, en primer lugar, que la varianza condicional de los rendimientos de activos financieros pueden tener mayor memoria que la típicamente capturada por los modelos tipo ARCH o ARSV. En segundo lugar, estos trabajos confirman las limitaciones de los modelos con raíces unitarias para reproducir la persistencia observada en los datos financieros, en favor de los procesos de memoria larga. Para tener en cuenta esta idea, se han generalizado los modelos de

volatilidad estocástica (SV), dando lugar a los Modelos de Volatilidad Estocástica con Memoria-Larga (LMSV) propuestos por Breidt, Crato y de Lima [18] y Harvey [43]. En este caso, el logaritmo de la varianza condicionada sigue un proceso ARFIMA  $(p,d,q)$ . Estos modelos implican que los rendimientos son una secuencia de martingalas en diferencias, es decir, incorreladas, mientras que la serie de las potencias de los valores absolutos de los rendimientos poseen correlaciones que decaen lentamente en el tiempo, lo cual contrasta con el decrecimiento exponencial impuesto por los modelos ARCH o SV de memoria corta (ARSV). Para este caso consideramos los modelos LMSV(1,d,0) de P.23 y P.24 <sup>4</sup>.

Las Tablas 2.9, 2.10 y 2.11 reportan las tasas de rechazo empírico de los tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  bajo los DGP's P.1-P.25, para  $T=100$  ( $m=3$ ),  $T=100$  ( $m=4$ ), 500 ( $m=4$ ) y 1000 ( $m=4$ ).

Como podemos ver, la potencia mejora rápidamente para los tres tests  $G_h(m)$  bajo los modelos heterocedásticos X-ARCH (DGP's P.1-P.18) cuando el tamaño de muestra y de símbolo aumentan, independientemente del proceso estocástico subyacente, siendo significativamente alta en todos los modelos para el test  $G_{h_{N_Q}}(m)$  y  $T=1000$ . Este test sigue teniendo una potencia significativamente alta si reducimos el tamaño muestral a la mitad,  $T=500$ .

La potencia del test  $G_{h_{M_e}}(m)$  también se mantiene cercana al 100 % para casi todos los modelos X-ARCH (DGP's P.1-P.18)  $T=1000$ , excepto para los DGP's P.4 y P.10, es decir, para el modelo ARCH(1) bajo las distribuciones subyacentes Chi-cuadrado (1) y Log-Normal (0,0.75) (5). Sin embargo, la razón por la que el test  $G_{h_{M_e}}(m)$  tiene tan buenos resultados en potencia se debe a que rechaza fuertemente en las pruebas de tamaño, como pudimos comprobar en el apartado 2.9.1.1.

Las tasas de rechazo obtenidas para el test  $G_{h_P}(m)$  no son buenas en general, aunque se acercan al 100 % en los DGP's P.4, P.5, P.6 y P.7. Este hecho viene a corroborar lo que ya anticipábamos en el punto 2.6 y es que la varianza condicional no está muy relacionada con los patrones ordinales.

---

<sup>4</sup>Los valores de los parámetros de los modelos SV P.19 a P.24 están tomados de [66]

Tabla 2.8: Modelos bajo la Hipótesis Alternativa

	Modelos	Innovaciones
P.1	ARCH(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2$	$\varepsilon_t : N(0, 1).$
P.2	GARCH(1,1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2$	$\varepsilon_t : N(0, 1).$
P.3	AR-GARCH(1,1,1): $y_t = 0,1y_{t-1} + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2.$	$\varepsilon_t : N(0, 1).$
P.4	ARCH(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2.$	$\varepsilon_t : \chi_1^2$
P.5	GARCH(1,1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2$	$\varepsilon_t : \chi_1^2$
P.6	AR-GARCH(1,1,1): $y_t = 0,1y_{t-1} + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2.$	$\varepsilon_t : \chi_1^2$
P.7	ARCH(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2.$	$\varepsilon_t : t_5$
P.8	GARCH(1,1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2$	$\varepsilon_t : t_5$
P.9	AR-GARCH(1,1,1): $y_t = 0,1y_{t-1} + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2.$	$\varepsilon_t : t_5$
P.10	ARCH(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2.$	$\varepsilon_t : \text{Log} - N(0, 0,75)$
P.11	GARCH(1,1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2$	$\varepsilon_t : \text{Log} - N(0, 0,75)$
P.12	AR-GARCH(1,1,1): $y_t = 0,1y_{t-1} + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2.$	$\varepsilon_t : \text{Log} - N(0, 0,75)$
P.13	ARCH(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2.$	$\varepsilon_t : U(-1, 1)$
P.14	GARCH(1,1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2$	$\varepsilon_t : U(-1, 1)$
P.15	AR-GARCH(1,1,1): $y_t = 0,1y_{t-1} + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,4u_{t-1}^2 + 0,5\sigma_{t-1}^2.$	$\varepsilon_t : U(-1, 1)$
P.16	ARCH-M: $y_t = -1,37 + 0,6\sigma_t + u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,7u_{t-1}^2$	$\varepsilon_t : N(0, 1).$
P.17	QARCH: $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 + 0,6u_{t-1}^2 + 0,3\sigma_{t-1}^2 - 0,15u_{t-1}\sigma_{t-1}$	$\varepsilon_t : N(0, 1).$
P.18	EGARCH: $y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = ((\varepsilon_{t-1}^2)^{0,6}(\varepsilon_{t-1}^2)^{0,3})/((\sigma_{t-1}^2)^{0,3}).$	$\varepsilon_t : N(0, 1).$
P.19	ARSV(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \log(\sigma_t^2) = 0,95\log(\sigma_{t-1}^2) + \eta_t, 0,22 = e^{\sigma_h^2} - 1, \sigma_\eta^2 = (1 - 0,95^2)\sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, \sigma_\eta^2)$
P.20	ARSV(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \log(\sigma_t^2) = 0,95\log(\sigma_{t-1}^2) + \eta_t, 1,72 = e^{\sigma_h^2} - 1, \sigma_\eta^2 = (1 - 0,95^2)\sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, \sigma_\eta^2)$
P.21	ARSV(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \log(\sigma_t^2) = 0,98\log(\sigma_{t-1}^2) + \eta_t, 0,22 = e^{\sigma_h^2} - 1, \sigma_\eta^2 = (1 - 0,98^2)\sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, \sigma_\eta^2)$
P.22	ARSV(1): $y_t = u_t, u_t = \sigma_t \varepsilon_t, \log(\sigma_t^2) = 0,98\log(\sigma_{t-1}^2) + \eta_t, 1,72 = e^{\sigma_h^2} - 1, \sigma_\eta^2 = (1 - 0,98^2)\sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, \sigma_\eta^2)$
P.23	LMSV(1,d,0): $y_t = u_t, u_t = \sigma_t \varepsilon_t, (1 - L)^{0,2}(1 - 0,9L)\log(\sigma_t^2) = \eta_t, 1,72 = e^{\sigma_h^2} - 1, 0,01 = (1 - 0,9^2)\sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, 1)$
P.24	LMSV(1,d,0): $y_t = u_t, u_t = \sigma_t \varepsilon_t, (1 - L)^{0,45}\log(\sigma_t^2) = \eta_t, 1,72 = e^{\sigma_h^2} - 1, 0,01 = \sigma_h^2$	$\varepsilon_t : N(0, 1)$ $\eta_t : N(0, 1)$

En lo que respecta a los modelos ARSV(1), la potencia de los tres tests aumenta conforme lo hace el coeficiente de persistencia alcanzando niveles cercanos al 100 % cuando éste es igual a 0.98 (P.19-P.22). Sin embargo, ninguno de los tres tests consigue capturar adecuadamente los DGP's pertenecientes a los procesos LMSV(1,d,0) (P.23 y P.24). Una de las razones podría residir en que los tres tests  $G_h(m)$  sólo son capaces de detectar dependencia de orden menor o igual que  $m$  (tamaño de símbolo) en la varianza y los procesos LMSV P.23 y P.24 al tener el parámetro de memoria a largo plazo mayor que cero, producen dependencias de larga distancia en la volatilidad que son difíciles de detectar por nuestros tests. Esta limitación, es una de las propiedades fundamentales de los tests  $G_h(m)$  (ver apartado 2.4.1).

Por último también podemos observar que la potencia de los tres tests se deteriora para el tamaño de muestra  $T=100$ , por lo que se puede decir que su uso, para este número de observaciones, no está recomendado.

### 2.9.2. Comparación con Otros Tests de Heterocedasticidad

Hasta ahora se ha visto que los tres estadísticos propuestos  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$ , poseen asintóticamente una distribución límite bajo la hipótesis nula, son independientes de la distribución subyacente y consistentes en mayor o menor medida, dependiendo del test considerado, contra una amplia clase de alternativas. También se ha comprobado en los apartados 2.9.1.1 y 2.9.1.2 que de los tres tests propuestos, el test  $G_{h_{NQ}}(m)$  es el que tiene mejor comportamiento en términos de tamaño y potencia y es por esto por lo que es el mejor candidato para el contraste de heterocedasticidad en series temporales.

El siguiente paso es comparar el estadístico  $G_{h_{NQ}}(m)$  para  $m=4$  con otros tests de heterocedasticidad conocidos, como el test de White [84], el test de Breuch-Pagan-Godfrey ([19], [36]), el test de Goldfeld-Quandt [38] y el test de Engle [29]. El contraste se

Tabla 2.9: Potencia Empírica de  $G_{h_P}(m)$ 

	$T=100$		$T=500$	$T=1000$
	$m=3$	$m=4$	$m=4$	$m=4$
P.1	0.065	0.154	0.704	0.966
P.2	0.033	0.057	0.053	0.066
P.3	0.02	0.078	0.144	0.295
P.4	0.114	0.206	0.594	0.868
P.5	0.11	0.166	0.911	0.998
P.6	0.119	0.247	0.954	1
P.7	0.078	0.211	0.852	0.997
P.8	0.032	0.071	0.057	0.068
P.9	0.024	0.095	0.151	0.305
P.10	0.051	0.11	0.423	0.713
P.11	0.068	0.137	0.905	0.999
P.12	0.083	0.17	0.975	1
P.13	0.33	0.523	0.999	1
P.14	0.059	0.122	0.441	0.808
P.15	0.121	0.211	0.857	0.972
P.16	0.05	0.168	0.701	0.972
P.17	0.064	0.102	0.43	0.858
P.18	0.141	0.435	0.999	1
P.19	0.023	0.067	0.05	0.04
P.20	0.043	0.07	0.057	0.058
P.21	0.032	0.083	0.05	0.04
P.22	0.024	0.064	0.045	0.038
P.23	0.038	0.072	0.1	0.203
P.24	0.032	0.066	0.04	0.064

Tabla 2.10: Potencia Empírica de  $G_{h_{M_e}}(m)$ 

	$T=100$		$T=500$	$T=1000$
	$m=3$	$m=4$	$m=4$	$m=4$
P.1	0.35	0.355	0.969	1
P.2	0.263	0.355	0.964	1
P.3	0.354	0.419	0.968	1
P.4	0.079	0.092	0.249	0.416
P.5	0.114	0.179	0.774	0.946
P.6	0.2	0.219	0.794	0.955
P.7	0.478	0.585	0.998	1
P.8	0.647	0.709	0.999	1
P.9	0.676	0.704	1	1
P.10	0.195	0.215	0.598	0.697
P.11	0.36	0.529	0.997	1
P.12	0.436	0.573	1	1
P.13	0.678	0.703	1	1
P.14	0.779	0.723	1	1
P.15	0.712	0.74	1	1
P.16	0.345	0.367	0.988	1
P.17	0.415	0.577	0.999	1
P.18	0.471	0.601	1	1
P.19	0.024	0.134	0.142	0.22
P.20	0.147	0.155	0.781	0.982
P.21	0.04	0.064	0.164	0.261
P.22	0.097	0.105	0.791	0.974
P.23	0.07	0.138	0.0202	0.319
P.24	0.039	0.06	0.085	0.081

Tabla 2.11: Potencia Empírica de  $G_{h_{NQ}}(m)$ 

	$T=100$		$T=500$	$T=1000$
	$m=3$	$m=4$	$m=4$	$m=4$
P.1	0.283	0.428	0.977	0.998
P.2	0.217	0.404	0.971	1
P.3	0.259	0.438	0.96	1
P.4	0.357	0.464	0.986	1
P.5	0.596	0.714	1	1
P.6	0.733	0.797	1	1
P.7	0.436	0.571	0.999	1
P.8	0.614	0.715	1	1
P.9	0.597	0.697	1	1
P.10	0.409	0.593	0.998	1
P.11	0.94	0.929	1	1
P.12	0.983	0.992	1	1
P.13	0.747	0.819	1	1
P.14	0.743	0.834	1	1
P.15	0.703	0.831	1	1
P.16	0.29	0.42	0.977	1
P.17	0.47	0.595	0.998	1
P.18	0.339	0.475	0.999	1
P.19	0.049	0.131	0.094	0.13
P.20	0.107	0.266	0.784	0.986
P.21	0.038	0.122	0.114	0.156
P.22	0.066	0.193	0.736	0.971
P.23	0.038	0.16	0.118	0.193
P.24	0.036	0.144	0.062	0.062

realiza comparando el tamaño y la potencia para tamaños de muestra medio (comportamiento con muestra finita) contra algunos de los DGP's descritos en 2.8 y otros nuevos que vamos a sugerir para  $T=500$ .

Describiremos los tests utilizados para la comparación considerando el modelo lineal

$$y_t = x_t' \beta + \varepsilon_t, t = 1, \dots, T \quad (2.43)$$

$$\varepsilon_t = \sigma_t u_t, t = 1, \dots, T \quad (2.44)$$

donde  $x_t = (x_{t1}, x_{t2}, \dots, x_{tk})'$ ,  $X \equiv [x_1, \dots, x_T]'$  es una matriz de dimensión  $T \times k$ ,  $\beta = (\beta_1, \dots, \beta_k)'$  es un vector de  $k \times 1$  coeficientes desconocidos,  $\sigma_1, \dots, \sigma_T$  son (posiblemente) parámetros de escala aleatorios, y  $u_t = (u_1, \dots, u_T)'$  es un vector aleatorio con una distribución continua condicional sobre  $X$  completamente especificada.

Los tests de White [84] y Breusch-Pagan-Godfrey ([19], [36]) son, seguramente, dos de los tests de heterocedasticidad basados en regresión auxiliar mejor conocidos de la literatura econométrica.

Para introducir estos tests en su forma más simple, consideremos la siguiente regresión auxiliar:

$$\hat{\varepsilon}_t^2 = z_t' \alpha + \omega_t, t = 1, \dots, T \quad (2.45)$$

donde  $z_t = (1, z_{t2}, \dots, z_{tm})'$  es un vector de  $m$  regresores fijos sobre los que puede depender  $\sigma_t$ ,  $\alpha = (1, \alpha_1, \dots, \alpha_m)'$  y  $\omega_t, t=1, \dots, T$  son términos de error.<sup>5</sup>

El criterio LM de Breusch-Pagan-Godfrey (BPG) se puede obtener como la suma de los cuadrados de la regresión (SCR) asociada con 2.45 dividida por  $2\hat{\sigma}^4$ . El test estadístico de White (W) es  $T$  veces el  $R^2$  no centrado de la regresión 2.45 usando para  $z_t$  las  $r \times 1$  observaciones de las variables no redundantes del vector  $x_t \otimes x_t$ . Estos tests se pueden derivar como tests de tipo-LM contra alternativas de la forma

$$\mathbb{H}_A : \sigma_t^2 = g(z_t' \alpha) \quad (2.46)$$

donde  $g(\cdot)$  es una función dos veces diferenciable. Bajo  $\mathbb{H}_0$  y varias condiciones estándar de regularidad asintóticas, tenemos:

---

<sup>5</sup>Véase Pagan y Hall [64] para una interpretación formal de la regresión del test.

$$BPG \sim \chi^2(m - 1),$$

$$W \sim \chi^2(r - 1),$$

donde el símbolo  $\sim$  indica que el test estadístico está asintóticamente distribuido como se indica (bajo  $\mathbb{H}_0$  cuando  $T \rightarrow \infty$ ).

El test más familiar de la clase que contrasta un punto de ruptura desconocido de la varianza es el test de Goldfeld y Quandt [38], el cual implica separar la muestra ordenada en tres subconjuntos y computar regresiones OLS separadas sobre el primer y último subconjuntos. Sea  $T_i$ , por  $i=1,2,3$  designamos el número de observaciones en cada uno de estos subconjuntos  $T = (T_1 + T_2 + T_3)$ . Teniendo en cuenta 2.43-2.44 y

$$u_1, \dots, u_T,$$

son i.i.d de acuerdo a alguna distribución  $F_0$ , en particular es bastante común asumir que

$$u_1, \dots, u_T \sim N(0, 1),$$

lo cual supone que  $\varepsilon_1, \dots, \varepsilon_T$  son i.i.d  $N(0, \sigma^2)$  bajo  $H_0$ . El test estadístico es

$$GQ(T_1, T_3, k) = \frac{S_3/(T_3 - k)}{S_1/(T_1 - k)}$$

donde  $S_1$  y  $S_3$  son la suma de los cuadrados de los residuos de las observaciones del primer subconjunto  $T_1$  y del último subconjunto  $T_3$  (donde  $k < T_1$  y  $k < T_3$ ). Bajo la nula  $GQ(T_1, T_3, k) \sim F(T_3 - k, T_1 - k)$ . Los parámetros empleados en este caso son  $T_2 = T/5$ ,  $T_1 = T_3 = (T - T_2)/2$ .

En el contexto de heterocedasticidad condicional, las regresiones artificiales proporcionan una manera fácil de contrastar los efectos GARCH. Engle [29] propuso un test LM basado en el siguiente escenario:

$$y_t = x_t' \beta + \varepsilon_t, \quad t = 1, \dots, T$$

$$\varepsilon_t |_{t-1} \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

donde  $|_{t-1}$  significa condicionar la información hasta  $t - 1$  inclusive. La hipótesis de homocedasticidad se puede formular como sigue:

$$\mathbb{H}_0 : \alpha_1 = \dots = \alpha_q = 0.$$

La comparación se completa considerando el test de Engle para retardo correcto en la varianza condicional, y se presenta su resultado bajo la etiqueta “Engle”. El test de Engle se calcula a través del producto  $TR^2$ , donde  $T$  es el tamaño de muestra,  $R^2$  es el coeficiente de determinación en la regresión OLS de los residuos al cuadrado  $\hat{\varepsilon}_t^2$  sobre una constante y  $\hat{\varepsilon}_{t-i}^2$  ( $i=1, \dots, q$ ). Bajo condiciones estándar de regularidad, la distribución asintótica en la hipótesis nula de este estadístico es  $\chi^2(q)$ .

Por último, en el contexto de los Modelos de Volatilidad Estocástica introduciremos alguno de los tests analizados en Rodríguez y Ruiz [66] junto con el nuevo test  $Q_i^*(M)$  propuesto por estos dos autores con el fin de tener una visión más realista de la competitividad del test  $G_{h_{NQ}}(m)$  con aquellos diseñados específicamente para detectar procesos SV.

En primer lugar consideraremos el test propuesto por Harvey y Streibel [44] que se enfoca específicamente en el modelo ARSV (1)

$$\begin{aligned} y_t &= \sigma_* \varepsilon_t \sigma_t, & t &= 1, \dots, T \\ \log(\sigma_t^2) &= \phi \log(\sigma_{t-1}^2) + \eta_t \end{aligned} \quad (2.47)$$

donde  $\sigma_*$  es un parámetro de escala,  $\sigma_t$  es la volatilidad, y  $\varepsilon_t$  y  $\eta_t$  son procesos gaussianos de ruido blanco mutuamente independientes de media cero y varianzas uno y  $\sigma_\eta^2$  respectivamente. El modelo es estacionario si  $|\phi| < 1$ . La varianza del proceso del logaritmo de la volatilidad está dado por  $\sigma_h^2 = (1 - \phi^2)\sigma_\eta^2$ , y se asume que es finita y fija. Si la persistencia del parámetro  $\phi$  es próxima a la unidad, entonces el valor de  $\sigma_\eta^2$  debe estar cerca de cero para un valor dado de la varianza de  $\log(\sigma_t^2)$  y la volatilidad evoluciona muy suavemente en el tiempo. En el límite, si  $\phi = 1$  entonces  $\sigma_\eta^2 = 0$  y  $y_t$  es condicionalmente homocedástico. De esta forma Harvey y Streibel proponen [44] contrastar  $\sigma_\eta^2 = 0$  usando

$$NM = -T^{-1} \sum_{k=1}^{T-1} r(k)k \quad (2.48)$$

donde  $r(k) = \sum_{t=k+1}^T (y_t - \bar{y}_t)(y_{t-k} - \bar{y}_t) / \sum_{t=1}^T (y_t - \bar{y}_t)^2$  es la autocorrelación de orden  $k$ ,  $\bar{y}_t = \sum_{k=1}^{T-1} y_t / T$  y  $T$  es el tamaño muestral. Los dos autores muestran que si el momento de segundo orden de  $y_t$  es finito, el estadístico NM posee asintóticamente la distribución de Crámer-von Mises, para la cual el valor crítico del 5% corresponde a 0.461.

En segundo lugar, Rodríguez y Ruiz [66] proponen contrastar la homocedasticidad condicional usando la información contenida en las autocorrelaciones muestrales de las transformaciones no-lineales  $f$  del proceso subyacente  $y_t$ , es decir, contrastar si las autocorrelaciones muestrales son diferentes de cero. Además el test tiene en cuenta que bajo la hipótesis nula, si existe el momento de cuarto orden de  $f(y_t)$ , las autocorrelaciones muestrales de  $f(y_t)$  son asintóticamente variables aleatorias normales independientes e idénticamente distribuidas con media cero y varianza  $1/T$  con lo que el test también incorpora información sobre los posibles patrones entre los sucesivos coeficientes de autocorrelación  $r(k)$ . Así, el test bajo estas condiciones toma la forma

$$Q_i^*(M) = T \sum_{k=1}^{M-i} \left[ \sum_{k=1}^{T-1} \tilde{r}(k+1) \right]^2, \quad i = 0, \dots, M-1 \quad (2.49)$$

donde  $\tilde{r}(k+1)$  es la autocorrelación muestral estandarizada de orden  $k+1$ ,  $\tilde{r}(k+1) = \sqrt{(T+2)(T-k-1)} r(k+1)$  y  $M$  es el número de autocorrelaciones consideradas. La distribución de  $Q_i^*(M)$  se puede aproximar por una distribución gamma,  $\zeta(\theta, \tau)$  con parámetros  $\theta = a^2/b$  y  $\tau = a/2b$ , con  $a = (i+1)(M-i)$  y  $b = (M-i)(i+1)^2 + 2 \sum_{j=1}^{i+1} (M-i-j)(i+1-j)^2$ .

Para comparar el comportamiento de muestra finito de estos tests, presentamos en la Tabla 2.12 el tamaño de los tests para un nivel del 5% basado en 1000 iteraciones de cada uno de los DGP's descritos en la Tabla 2.4 para  $T=500$  observaciones.

Como podemos observar, los niveles empíricos de los tests  $G_{h_{NQ}}(m)$ , White y Engle

Tabla 2.12: Tamaño Empírico de  $G_{h_{NQ}}(4)$  y varios Tests

	$G_{h_{NQ}}(4)$	White	BP	GQ	Engle
S.1	0.053	0.045	0.17	0.175	0.033
S.2	0.059	0.035	0.14	0.117	0.036
S.3	0.065	0.045	0.17	0.159	0.045
S.4	0.064	0.044	0.25	0.297	0.029
S.5	0.058	0.048	0.02	0.038	0.033
S.6	0.063	0.061	0.001	0.004	0.065
S.7	0.051	0.053	0.006	0.004	0.049
S.8	0.058	0.049	0.003	0.01	0.054
S.9	0.054	0.05	0.007	0.008	0.051
S.10	0.06	0.049	0.008	0.012	0.056
S.11	0.061	0.035	0.25	0.344	0.016
S.12	0.056	0.043	0.25	0.337	0.024
S.13	0.067	0.036	0.28	0.356	0.022
S.14	0.064	0.048	0.24	0.325	0.017
S.15	0.061	0.037	0.26	0.345	0.02
S.16	0.063	0.044	0.041	0.036	0.036

Tabla 2.13: Potencia Empírica de  $G_{h_{NQ}}(4)$  y Varios Tests con Modelos X-ARCH

DGPs	$G_{h_{NQ}}(4)$	White	Engle
P.2	0.97	1	1
P.5	1	0.941	0.8
P.8	1	0.982	0.994
P.11	1	1	0.925
P.14	1	1	1
P.18	1	0.66	0.72

son correctos, ya que éstos se mantienen casi siempre en el intervalo de probabilidad  $[0.036, 0.064]$  correspondiente a un 95 % de confianza para un tamaño del 5 % y 1000 iteraciones. Siendo los tests de White y Engle un poco más conservativos que el test  $G_{h_{NQ}}(m)$ . Por el contrario, los tests BP (Breuch-Pagan) y GQ (Goldfeld y Quandt) rechazan fuertemente para algunos DGP's (S.1, S.2, S.3, S.4, S.11, S.12, S.13, S.14 y S.15). Este es un problema serio por que estos tests podrían llevar al usuario a una conclusión errónea una vez que el investigador fija los niveles de confianza de estos tests.

El siguiente paso de la comparación es considerar el comportamiento de los tests bajo la hipótesis alternativa de heterocedasticidad. Para este fin usamos algunos de los modelos reunidos en la Tabla 2.8, particularmente y en primer lugar todos aquellos con estructura GARCH(1,1), como en el modelo estándar clásico, también incluimos uno más realista como el modelo EGARCH. La comparación se realiza para los tres estadísticos siguientes:  $G_{h_{NQ}}(4)$ , White, y Engle, ya que estos tests presentan tamaños correctos <sup>6</sup> en nuestro experimento previo. Notar que si se consideraran los tests BP y GQ, los resultados de potencia estarían contaminados por su tendencia fuerte a rechazar. Los resultados se presentan en la Tabla 2.13.

<sup>6</sup>Los resultados para todos los tests y todos los modelos están disponibles en los autores.

El resultado más notable es que el test  $G_{h_{NQ}}(4)$  supera a los otros dos excepto en el caso P.2 que es uno de los casos para el que los errores de la estructura GARCH(1,1) siguen una distribución normal estándar. Sin embargo, para este caso las diferencias en términos de potencia no son realmente relevantes.

Como podemos ver también, el test  $G_{h_{NQ}}(4)$  claramente domina a los otros tests cuando se permiten estructuras asimétricas no lineales deterministas en la varianza condicional. Como se comentó previamente, es un hecho constatado empíricamente que las sorpresas negativas del mercado parecen incrementar más la volatilidad que las sorpresas positivas. Debido a que un precio más bajo de la acción reduce el valor del activo neto de la deuda privada, un descenso agudo en los precios de las acciones aumenta el apalancamiento privado y podría de esta forma aumentar el riesgo de poseer acciones. Por esta razón, es conveniente para la forma estructural de la varianza condicional poder soportar este comportamiento asimétrico. El modelo P.18 posee una especificación ARCH no lineal que tiene en cuenta shocks asimétricos. Para este modelo, los tests de White y Engle no son competitivos en términos de potencia, a pesar del hecho de considerar innovaciones con distribución estándar normal. Sin embargo, el test  $G_{h_{NQ}}(4)$  obtiene niveles de rechazo espectaculares<sup>7</sup>, lo cual sugiere que es razonable considerar a este test como superior a los otros dos tests cuando tenemos en cuenta estructuras X-ARCH clásicas y realistas en la varianza condicional. Esta conclusión está reforzada por el hecho de que para usar el test  $G_{h_{NQ}}(4)$  los supuestos requeridos son mínimos.

En la Tabla 2.15, también incluimos en el análisis comparativo los modelos de Volatilidad Estocástica (Modelos SV), de memoria a corto plazo ARSV (P.20 y P.22) y largo plazo LMSV (P.23 y P.24) de la Tabla 2.8, los cuales han atraído recientemente la atención de algunos especialistas dedicados al modelado de datos de alta frecuencia como es el caso de los datos financieros recabados dentro del mismo día. Además como hemos comentado en párrafos anteriores, ampliamos la comparación con los tests SV: NM y  $Q_7^*(24)$ , extraídos del artículo de Julio Rodríguez y Esther Ruiz [66].

---

<sup>7</sup>Aunque no se reporta aquí, las diferencias en términos de potencia aumentan cuando se consideran innovaciones no estándar.

Tabla 2.14: Potencia Empírica de  $G_{h_{NQ}}(4)$  y varios Tests con Modelos SV

DGPs	$G_{h_{NQ}}(4)$	White	Engle	NM	$Q_7^*(24)$
P.19	0.094	0.459	0.323	0.58	0.75
P.20	0.784	0.819	0.771	1	1
P.21	0.114	0.315	0.252	0.6	0.73
P.22	0.736	0.782	0.806	1	1
P.23	0.118	0.127	0.249	0.376	0.573
P.24	0.062	0.54	0.587	0.382	0.38

Como podemos observar el test  $G_{h_{NQ}}(4)$  no es claramente competitivo con los otros cuatro tests en los modelos ARSV (P.19 y P.21) cuando el coeficiente de persistencia es bajo. Esta desventaja es aún más notoria con respecto al test  $Q_7^*(24)$  donde los niveles de rechazo siempre se mantienen iguales o superiores al 75 %. La potencia del test  $G_{h_{NQ}}(4)$  en los modelos P.20 y P22 aunque es inferior a los demás tests (con excepción del test de Engle para el modelo P.20), se mantiene en niveles alrededor del 75 %. El test  $G_{h_{NQ}}(4)$  también resulta perdedor con los cuatro tests en los procesos LMSV, si bien las tasas de rechazo mostradas en todos los casos se mantienen por debajo del 50 % .

Finalmente consideramos otras especificaciones ARCH, no lineales y más complejas para mostrar que las ganancias relevantes en potencia del test  $G_{h_{NQ}}(m)$  aplican a otros modelos. El modelo P.25 es una variante del P.18:

$$y_t = u_t, u_t = \sigma_t \varepsilon_t, \log \sigma_t^2 = -2,4 + 0,75 \log \sigma_{t-1}^2 + 0,5 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,15 \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

El modelo P.26 fue introducido por Brock [20]. En este caso  $\sigma_t^2$  está determinísticamente determinada por la aplicación “tienda de campaña” indicada abajo, con  $\sigma_0^2 \in (0, 1)$ . Este modelo es claramente heterocedástico, pero no exhibe clustering de la volatilidad, ya que la correlación empírica serial de  $\{\sigma_t^2\}$  se aproxima a cero en muestras grandes para casi todos los valores de  $\sigma_0^2$ .

Tabla 2.15: Potencia Empírica de  $G_{h_{NQ}}(4)$  y varios Tests con otras estructuras ARCH no lineales

DGPs	$G_{h_{NQ}}(4)$	White	Engle
P.25	0.99	0.012	0.012
P.26	1	0.282	0.784

$$y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 - 2|\sigma_{t-1}^2 - 0,5|$$

en ambos modelos los errores son estándar normales.

Los resultados se muestran en la Tabla 2.15. En este caso se puede ver claramente que los resultados de potencia del test  $G_{h_{NQ}}(4)$  son superiores a los tests de White y Engle cuando la complejidad de la estructura no lineal aumenta.

## 2.10. Relación de los resultados de Tamaño y el solapamiento de las $m - historias$

Como se vio en los apartados 2.4.2 y 2.8.1, el modo en que seleccionamos las  $m - historias$  de la serie  $Y_t$  a contrastar puede influir en los resultados de tamaño de los tests ya que  $Y_t$  e  $Y_{t+k}$  pueden ser independientes o no, dependiendo de si  $k$  es mayor que  $m - 1$ , aunque si  $m$  es pequeño comparado con el número de observaciones  $T - m + 1$ , en general estas correlaciones se pueden despreciar.

También pudimos ver en el apartado 2.9.1.1 que las tasas de rechazo mostradas por los tests  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$  eran en general bastante superiores al nivel de significación establecido del 5%, más evidente este fenómeno en el test  $G_{h_{M_e}}(m)$ , y se sugirió que una de las posibles razones para este comportamiento podría ser el solapamiento entre las  $m - historias$  de la serie temporal a contrastar.

Para comprobar cómo afecta el grado de correlación producido por el solapamiento de las  $m - historias$   $Y_m(t)$  e  $Y_m(t+k)$  en los resultados de tamaño de los tests reproducimos

los experimentos que describimos en la sección 2.9.1.1 para los tamaños de símbolo  $m = 3$  y  $m = 4$  con solapamiento  $(Y_t, Y_{t+1}, Y_{t+2}, \dots, Y_{t+T-m+1})$  y sin solapamiento de las  $m - \text{historias}$   $(Y_t, Y_{t+m}, Y_{t+2m}, \dots, Y_{t+T-m+1})$  cumpliendo en todo momento la regla de selección del tamaño de muestra particularizado para cada test (ver sección 2.8.1), es decir :

$$5 \cdot m! \leq T_1$$

$$5 \cdot 2^{m-1} \leq T_2$$

$$5 \cdot 3^m \leq T_3$$

siendo  $T_1$ ,  $T_2$  y  $T_3$  los tamaños de muestra finitos para los tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  respectivamente.

Además en los experimentos sin solapamiento de las  $m - \text{historias}$  exigimos que el número de las mismas sea similar a las del escenario con solapamiento, por ello la regla de selección del tamaño de muestra  $T_i$ ,  $i = 1, 2, 3$  para cada test es:

$$5 \cdot m \cdot m! \leq T_1$$

$$5 \cdot m \cdot 2^{m-1} \leq T_2$$

$$5 \cdot m \cdot 3^m \leq T_2$$

es decir, que el número de observaciones cuando no tenemos solapamiento de las  $m - \text{historias}$  es  $m$  veces más grande que cuando imponemos la condición de no solapamiento.

Los tamaños de muestra seleccionados para los experimentos toman los valores que reflejamos en la Tabla 2.16 ( $k = 1$ ).

Tabla 2.16: Número de Observaciones con solapamiento ( $k = 1$ ) y sin solapamiento ( $k \geq m$ ) de  $m$ -historias para los tres tests

$k = 1$	$G_{h_P}(m)$	$G_{h_{M_e}}(m)$	$G_{h_{N_Q}}(m)$
$m = 3$	30	20	135
$m = 4$	120	40	500
$k \geq m$	$G_{h_P}(m)$	$G_{h_{M_e}}(m)$	$G_{h_{N_Q}}(m)$
$m = 3$	90	60	405
$m = 4$	480	160	2000

Tabla 2.17: Número de Observaciones con solapamiento ( $k = 1$ ) y sin solapamiento ( $k \geq m$ ) de  $m$ -historias para  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$  (1).

$k = 1$	$G_{h_P}(m)$	$G_{h_{M_e}}(m)$
$m = 3$	135	135
$m = 4$	500	500
$k \geq m$	$G_{h_P}(m)$	$G_{h_{M_e}}(m)$
$m = 3$	405	405
$m = 4$	2000	2000

Como podemos observar en los test  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$ , el número de observaciones es muy pequeño, por lo que para evitar efectos debidos a el tamaño de la muestra en los resultados repetimos los experimentos para estos dos tests con los tamaños de muestra seleccionados para el test  $G_{h_{N_Q}}(m)$  de la Tabla 2.17.

Para comparar el comportamiento de muestra finito con y sin solapamiento, presentamos en las Tablas 2.18, 2.19 y 2.20 el tamaño de los tres tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  respectivamente para un nivel de significatividad del 5% basado en 1000 iteraciones de los procesos descritos en la Tabla 2.4 con los distintos tamaños de muestra  $T$  de las Tablas 2.16 y 2.17.

Los resultados muestran claramente cómo los niveles empíricos de  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$  mejoran ostensiblemente cuando no hay solapamiento de símbolo, estando en la mayoría de los procesos en niveles cercanos al tamaño nominal del 5%. Esto no ocurre con el test  $G_{h_{N_Q}}(m)$ , el cual parece más insensible al grado de solapamiento de las  $m - historias$ . Como conclusión, el comportamiento en tamaño de los tests  $G_{h_P}(m)$  y  $G_{h_{M_e}}(m)$  se corrige y es correcto cuando las  $m - historias$  de  $Y_m(t)$  e  $Y_m(t+k)$  son independientes, es decir, cuando no hay solapamiento de símbolo y el número de  $m - historias$  crece. Esto resulta lógicamente a costa de una degradación en los resultados de potencia.

Tabla 2.18: Tamaño Empírico de  $G_{h_P}(m)$  con y sin solapamiento de símbolo  $m$

	$m=3$				$m=4$			
	$k=1$ $T=30$	$k=m$ $T=90$	$k=1$ $T=135$	$k=m$ $T=405$	$k=1$ $T=120$	$k=m$ $T=480$	$k=1$ $T=500$	$k=m$ $T=2000$
S.1	0.026	0.068	0.023	0.044	0.051	0.081	0.041	0.05
S.2	0.03	0.064	0.028	0.048	0.055	0.092	0.048	0.055
S.3	0.018	0.064	0.029	0.055	0.063	0.08	0.049	0.058
S.4	0.035	0.076	0.033	0.057	0.088	0.078	0.053	0.057
S.5	0.033	0.06	0.031	0.062	0.073	0.085	0.051	0.048
S.6	0.026	0.056	0.016	0.05	0.077	0.084	0.041	0.042
S.7	0.029	0.063	0.029	0.048	0.054	0.073	0.035	0.052
S.8	0.036	0.066	0.026	0.048	0.045	0.087	0.043	0.069
S.9	0.035	0.066	0.019	0.053	0.07	0.057	0.051	0.057
S.10	0.032	0.066	0.029	0.044	0.05	0.065	0.086	0.043
S.11	0.042	0.093	0.048	0.07	0.084	0.076	0.084	0.069
S.12	0.041	0.057	0.04	0.06	0.072	0.071	0.077	0.077
S.13	0.029	0.079	0.037	0.055	0.093	0.094	0.08	0.069
S.14	0.038	0.082	0.048	0.059	0.084	0.084	0.075	0.067
S.15	0.03	0.068	0.048	0.051	0.08	0.08	0.084	0.072
S.16	0.034	0.063	0.021	0.061	0.058	0.058	0.049	0.055

Tabla 2.19: Tamaño Empírico de  $G_{h_{Me}}(m)$  con y sin solapamiento de símbolo  $m$

	$m=3$						$m=4$					
	$k=1$	$k=m$	$k=1$	$k=m$	$k=1$	$k=m$	$k=1$	$k=m$	$k=1$	$k=m$	$k=1$	$k=m$
	$T=20$	$T=60$	$T=135$	$T=405$	$T=40$	$T=160$	$T=500$	$T=2000$				
S.1	0.134	0.062	0.074	0.062	0.135	0.064	0.1	0.064				
S.2	0.117	0.066	0.077	0.054	0.115	0.062	0.094	0.043				
S.3	0.096	0.052	0.07	0.057	0.132	0.053	0.092	0.045				
S.4	0.129	0.071	0.085	0.056	0.151	0.064	0.112	0.056				
S.5	0.124	0.066	0.077	0.046	0.135	0.08	0.114	0.037				
S.6	0.122	0.074	0.083	0.042	0.142	0.073	0.092	0.057				
S.7	0.128	0.054	0.095	0.057	0.122	0.069	0.094	0.052				
S.8	0.123	0.064	0.077	0.04	0.136	0.071	0.087	0.045				
S.9	0.127	0.061	0.089	0.057	0.127	0.057	0.104	0.046				
S.10	0.14	0.056	0.088	0.045	0.115	0.067	0.113	0.062				
S.11	0.124	0.041	0.111	0.04	0.141	0.086	0.09	0.052				
S.12	0.191	0.047	0.079	0.058	0.098	0.068	0.09	0.048				
S.13	0.098	0.016	0.105	0.047	0.156	0.09	0.128	0.042				
S.14	0.12	0.068	0.103	0.029	0.154	0.042	0.105	0.053				
S.15	0.123	0.053	0.091	0.035	0.077	0.07	0.119	0.055				
S.16	0.101	0.08	0.062	0.039	0.116	0.066	0.085	0.057				

Tabla 2.20: Tamaño Empírico de  $G_{h_{NQ}}(m)$  con y sin solapamiento de símbolo  $m$ 

	$m=3$		$m=4$	
	$k=1$ $T=135$	$k=m$ $T=405$	$k=1$ $T=500$	$k=m$ $T=2000$
S.1	0.045	0.055	0.053	0.05
S.2	0.029	0.05	0.039	0.052
S.3	0.029	0.05	0.044	0.061
S.4	0.026	0.04	0.054	0.063
S.5	0.032	0.05	0.042	0.052
S.6	0.028	0.053	0.048	0.057
S.7	0.023	0.055	0.054	0.056
S.8	0.031	0.045	0.043	0.058
S.9	0.035	0.039	0.038	0.051
S.10	0.025	0.04	0.038	0.057
S.11	0.034	0.051	0.047	0.06
S.12	0.038	0.059	0.063	0.061
S.13	0.054	0.054	0.060	0.054
S.14	0.038	0.066	0.056	0.063
S.15	0.038	0.052	0.050	0.06
S.16	0.037	0.042	0.034	0.05

## 2.11. Conclusiones

En este capítulo hemos propuesto tres tests,  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$ , semi-paramétricos de heterocedasticidad basados en el análisis de la distribución de patrones denominado *Análisis Simbólico* que tiene su origen en el análisis de patrones ordinales sugerido por Bandt y Pompe. Estas técnicas estadísticas han sido propuestas y desarrolladas entre otros por M. Matilla-García y Manuel Ruiz.

Los tres tests se basan en el hecho que la hipótesis nula se puede expresar en probabilidades asociadas con un conjunto dado de patrones. Luego, bajo la hipótesis nula de homocedasticidad y de una forma sencilla hemos extendido a otros patrones diferentes de los ordinales la relación entre la entropía de simbólica y las distribuciones estadísticas chi-cuadrado.

Además, los tres tests analizan las series temporales usando hipótesis mínimas, mirando específicamente a la premisa de homocedasticidad con la media condicional correctamente especificada. Estos nuevos tests están basados en la extracción de características heterocedásticas encontradas en la distribución de los patrones de la serie de los errores estimados al cuadrado.

De los tres tests analizados en términos de tamaño y potencia, el test  $G_{h_{NQ}}(m)$  es el que mejor comportamiento ha presentado. Este nuevo test ha demostrado ser consistente contra una amplia clase de procesos heterocedásticos. El tamaño empírico no difiere del tamaño teórico, lo cual es una propiedad interesante que garantiza la aplicabilidad general y reproducibilidad del test. También hemos comprobado que el test, a parte de ser menos exigente en términos de supuestos, es muy competitivo contra los tests más populares de heterocedasticidad, salvo en el caso de procesos de volatilidad estocástica (Modelos SV). Además se ha visto como el nuevo test supera a los otros tests disponibles cuando hay una estructura no lineal compleja en la varianza del proceso. Por último, dos ventajas adicionales son su simplicidad computacional y sus bajos tiempos de computación.

Demostrado el buen rendimiento del test  $G_{h_{NQ}}(m)$  para detectar desviaciones respecto de la hipótesis nula de homocedasticidad, el siguiente paso es, y dado el pobre rendi-

miento mostrado frente a los modelos de Volatilidad Estocástica, mejorar las pobres tasas de rechazo presentadas por este test frente a este tipo de modelos.

## Capítulo 3

# Aplicación de la Entropía de Tsallis a las Técnicas de Análisis Simbólico para el Contraste de Heterocedasticidad Condicional de Forma Desconocida en Series Temporales

### 3.1. Introducción

Como se ha visto en el capítulo anterior, hemos propuesto tres tests semiparamétricos de heterocedasticidad mediante las técnicas de análisis simbólico. De los tres tests analizados, uno de ellos se comprobó que era claramente competitivo en términos de tamaño y potencia con los tests de heterocedasticidad más populares. Este tipo de tests posee varias características reseñables. La primera es que están basados en el concepto general de Entropía Simbólica, una generalización de la Entropía de Permutación. La segunda es que son de naturaleza semiparamétrica. La tercera es que a diferencia de otros tests existentes, éstos parecen comportarse muy bien ante una amplia gama de modelos heterocedásticos con forma funcional desconocida. Por último, su comportamiento en términos de potencia no es bueno cuando el funcional que define las perturbaciones heterocedásticas es de naturaleza estocástica compleja con correlaciones a largo plazo, como en el caso de los Modelos de Volatilidad Estocástica de memoria larga (LMSV). Para intentar superar este inconveniente podemos acometer varias estrategias. En pri-

mer lugar aumentar el número de  $m$  – *historias* de la serie hasta detectar el proceso. En segundo lugar podemos aumentar la dimensión de inmersión  $m$ , lo cual nos permite, según pudimos comprobar, aumentar la potencia. Estas dos primeras estrategias funcionan para procesos heterocedásticos de complejidad relativa, pero no para los Modelos S.V de memoria larga. Ambas estrategias han resultado ineficientes frente a este tipo de modelos. La tercera estrategia, objeto de estudio del presente capítulo, consiste en cambiar o más formalmente generalizar la forma de medir la Información Simbólica del proceso heterocedástico para que tenga en cuenta las dependencias de largo alcance. Para llevar a cabo este objetivo, tenemos que estudiar qué forma o formas funcionales son las más adecuadas para medir la información contenida en los procesos heterocedásticos de memoria larga. Para ello vamos a comenzar recordando la definición formal de Entropía Simbólica, que se define como la Entropía de Shannon de  $n$  símbolos para cada dimensión de inmersión  $m$ :

$$h(m) = - \sum_{\pi \in \Gamma} p_{\pi} \ln(p_{\pi}) \quad (3.1)$$

Sin embargo, la Entropía de Shannon es una reinención en versión binaria de la forma funcional de Boltzmann-Gibbs interpretada en términos de información y teoría de la Comunicación. Así, dada una probabilidad de distribución  $P = \{p_i\}$  ( $i = 1, \dots, N$ ), siendo  $p_i$  la probabilidad del sistema de estar en el microestado  $p_i$  la Entropía de BG tiene la forma:

$$S_{BG}(P) = -k \sum_{i=1}^N p_i \ln(p_i) \quad (3.2)$$

donde  $k$  es la constante de Boltzmann y  $N$  es el número total de microestados Si todos los estados son equiprobables conduce al famoso principio de Boltzmann ( $N = W$ ):

$$S = k \log W \quad (3.3)$$

Si establecemos  $k = 1$  y usamos la base  $b$  para la función logarítmica ambos funcionales

entrópicos resultan ser equivalentes

$$S_S(P) = - \sum_{i=1}^N p_i \ln_b(p_i) \quad (3.4)$$

Es común usar el logaritmo natural para la entropía de BG mientras que la base 2 tiene la ventaja de entregar las cantidades de información en bits. Por lo tanto lo que caracteriza las medidas de BG y Shannon es la aditividad de la información. Dados dos sistemas descritos por dos distribuciones de probabilidad independientes A y B (i.e  $P(A \cap B) = P(A)P(B)$ ), usando la aditividad de la información significa que:

$$S_S(A \cap B) = S_S(A) + S_S(B/A) \quad (3.5)$$

y

$$S_S(B/A) = - \sum_{i=1}^N p_i(A) S_S(B/A = A_i) \quad (3.6)$$

es la entropía condicional.

Este es el caso de los sistemas extensivos, i.e. sistemas donde la entropía está dada por la suma de todas las entropías de sus partes como se ha hecho tradicionalmente en la mecánica estadística estándar y en la teoría de la Información. Para asegurar la aditividad de estos sistemas es necesario utilizar el logaritmo. Esta es la razón real que subyace detrás de la presencia ubicua de la función logaritmo en la teoría de la información, y podemos decir que cualquier modificación refleja una desviación de la ley de aditividad. La aditividad sin embargo no se preserva siempre, especialmente en los sistemas complejos no lineales, fenómenos débilmente ergódicos que también son abundantes en la naturaleza (fenómenos biológicos, socio-económicos, fenómenos cognitivos...etc ), esencialmente son conducidos por interacciones microscópicas que son de largo alcance tanto en el espacio como en el tiempo (fuerzas de largo alcance, memoria no-markoviana, condiciones de frontera fractales..). Estamos tratando en este caso con sistemas no-extensivos, sistemas que han recibido mucha atención últimamente.

Hasta hace relativamente poco tiempo la entropía y sus formas entrópicas han ido dis-

curriendo por senderos retorcidos, ya que desde hace años la gente que trabajaba en Teoría de la Información, Complejidad, Cibernética y Sistemas Dinámicos No Lineales y Físicos entre otros han introducido cerca de 40 funcionales entrópicos. En 1960 Renyi propuso una nueva forma que recupera la Entropía de Shannon como caso especial. Su forma es siempre extensiva pero no siempre cóncava (o convexa) con respecto a la distribución de probabilidad. En 1987, presuntamente con fines cibernéticos Harvda y Charvat introdujeron, aunque con un prefactor diferente (ajustado para variables binarias) la forma entrópica denotada por  $S_q$ , y estudiaron alguna de sus propiedades matemáticas. Un año más tarde Vajda, además estudió esta forma, citando a Harvda y Charvat. En 1970, en el contexto de la información y control, Darockey redescubrió esta forma. Más tarde en 1975, Sharma y Mittal introdujeron y estudiaron algunas propiedades de una forma de dos parámetros que recupera ambas formas de Renyi y  $S_q$  como instancias particulares. En 1988, en la comunidad física, se produjo un nuevo redescubrimiento realizado por Constantino Tsallis, el cual inspirado en los Multifractales, postuló la forma entrópica  $S_q$  como posible nuevo camino para la generalización de la mecánica estadística de BG, a través de la optimización de  $S_q$  y obtuvo una distribución de equilibrio la cual ,genéricamente, seguía la ley de potencias, y reproducía el factor celebrado de Boltzmann como límite particular cuando  $q \rightarrow 1$ . Esta constituye una generalización de la entropía de BG a sistemas no-extensivos. Tsallis notó que si la no-extensividad entra en juego, las cosas se describen mejor por distribuciones que siguen una ley potencial,  $p_i^q$ , las así llamadas *q-probabilidades*, es decir, probabilidades de escala, donde  $q$  es un parámetro real, es el índice entrópico que debería ser determinado a priori de la dinámica microscópica. Sin embargo la elección apropiada de este parámetro es un problema que permanece todavía abierto [48]. Esto introduce la posibilidad formal de no hacer equiparables los eventos comunes y los raros, como en las estadísticas de BG o Shannon, sino de resaltarlos o atenuarlos de acuerdo a los parámetros elegidos (en sistemas complejos los eventos raros pueden tener efectos dramáticos sobre la evolución global). Como consecuencia de ello, la función densidad de probabilidad derivada del principio variacional relacionado con la entropía de Tsallis

resulta ser la función densidad de probabilidad *q-gaussiana* [80].

$$p(x) = A_q e_q^{-B_q(x-\mu_q)^2}$$

siendo  $A_q$  y  $B_q$  parámetros dependientes de  $q$  y  $e_q$  la función *q-exponencial*. En el límite  $q \rightarrow 1$  se recupera la distribución Gaussiana asociada a la Entropía de BG en estado de Equilibrio Térmico.

Existen gran cantidad de mecanismos dinámicos que conducen a las funciones densidad de probabilidad gaussianas, como por ejemplo las que fueron derivadas por L. Bachelier en 1900 en su trabajo sobre la dinámica de los cambios en el precio del Mercado de París y en 1905 por Albert Einstein en su trabajo sobre el movimiento Browniano. También partiendo de la dinámica de Langevin uno es capaz de escribir la correspondiente ecuación de Fokker-Plank y de ella obtener como solución la distribución gaussiana. Análogamente, también es posible obtener la familia de distribuciones *q-gaussianas*, partiendo de ciertas clases de ecuaciones diferenciales estocásticas y sus ecuaciones de Fokker Plank asociadas. Como consecuencia de ello, los dos teoremas básicos del límite central (1) Una distribución con momento finito de segundo orden convolucionada (distribución de la suma de  $N$  variables independientes), se aproxima en el límite cuando las convoluciones  $N \rightarrow \infty$  a un atractor Gaussiano (2) Una distribución convolucionada con un momento de segundo orden divergente, se aproxima en el límite a una distribución de Lévy  $L_\gamma(x)$  ( $0 < \gamma < 2$ ); se generalizan con diferentes tipos de dinámicas asociadas a la convolución. Así, a parte de las dinámicas asociadas a las distribuciones Gaussiana o de Lévy, existen en la naturaleza dinámicas mucho más complejas, por ejemplo aquellas asociadas con la variedad de ecuaciones no lineales de Fokker-Plank que implican correlaciones no lineales, ruido multiplicativo u otros efectos. Es decir, las convoluciones simples permiten sólo un comportamiento asintótico para  $q = 1$  (Gaussiana) ó  $q > 5/3$  (distribuciones de Lévy) pero no permiten distribuciones con colas pesadas, asociadas con el intervalo  $1 < q \leq 5/3$ . Pero muchos sistemas complejos tanto en la naturaleza como en las ciencias sociales exhiben expo-

nentes que pertenecen precisamente a ese intervalo. Ejemplos de este tipo de sistemas los encontramos en el ámbito económico y financiero y que poseen las características típicas de los modelos ARCH y GARCH, en los que el proceso de la variable heterocedástica (e.g los rendimientos) se puede demostrar (ver [26]) que posee conexión con aquellos que ocurren en sistemas de no-equilibrio en estado estacionario de larga duración. Concretamente, se ha demostrado que existe una relación funcional entre la distribución de los modelos ARCH-GARCH y la distribución  $q$ -Gaussiana. Otro ejemplo de la conexión de los modelos heterocedásticos con la distribución  $q$ -Gaussiana son los modelos S.V (Volatilidad Estocástica), en los cuales se parte de ecuaciones diferenciales estocásticas no lineales, como las ecuaciones generalizadas de Fokker-Plank, y se obtienen diferentes tipos de soluciones [80]. En su versión lineal la ecuación de Fokker Plank tiene como soluciones la distribución Gaussiana (v.a i.i.d de varianza finita) y de Lévy (v.a i.i.d de varianza infinita) y en su versión no lineal las distribuciones  $(q, \alpha)$ -estables ( $q$ -varianza infinita); y la distribución  $q$ -Gaussiana ( $q$ -varianza finita) <sup>1</sup>. Estas dos últimas distribuciones violan de la hipótesis de independencia, pero no una débil violación con correlaciones que desaparecen gradualmente cuando  $t \rightarrow \infty$ , sino que está producida por correlaciones globales que persisten en el infinito.

Las diferentes consideraciones mencionadas anteriormente nos permiten conjeturar que la Entropía Simbólica de Tsallis, es decir la medida entrópica de Tsallis evaluada después usar el Análisis Simbólico para extraer la distribución de probabilidad asociada, debería ser un cuantificador más potente que la entropía Simbólica de Shannon para caracterizar los procesos heterocedásticos. Así, la Entropía Simbólica de Tsallis constituye una generalización de la Entropía de Permutación de Tsallis del mismo modo como la Entropía Simbólica constituye una generalización de la Entropía de Permutación. Algún precedente en este sentido podemos encontrarlo en el trabajo de L. Zunino et. al [89], los cuales introdujeron el concepto de *Entropía de Permutación de Tsallis*

---

<sup>1</sup>Según sea el tipo de distribución a que da lugar, el Teorema del Límite Central toma distintas formas. En las distribuciones Gaussianas se cumple el teorema del Límite Central-CLT. En las distribuciones de Lévy se cumple el Teorema del Límite Central de Lévy-Gnedenko y en las distribuciones tipo  $q$ -Gaussiana se cumple el teorema del límite Central Generalizado  $q$ -CLT.

como herramienta de descripción de los procesos estocásticos de naturaleza fractal.

El objetivo del presente capítulo es mejorar, a la luz de las  $q$  – estadísticas de Tsallis, los pobres resultados de potencia mostrados por los tests del capítulo anterior  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$  en procesos heterocedásticos de largo alcance como los modelos LMSV. Para este fin encontraremos en primer lugar la relación formal entre la Entropía de Tsallis y la distribución de los símbolos en los tests  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$ . Esto dará lugar a unas nuevas expresiones de los tests, llamados  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  que permiten incorporar la información no extensiva presente en las series heterocedásticas de memoria larga. En segundo lugar evaluaremos el rendimiento de tamaño y potencia de los nuevos tests estimando teóricamente el índice entrópico  $q$  de las series temporales utilizadas en el contraste. El objetivo de este ejercicio, meramente teórico dado que no se dispone hasta el momento de un método formal aceptado para la estimación no paramétrica del índice  $q$ , es mostrar que al reemplazar la Entropía de Shannon por la Entropía de Tsallis en la forma de medir la distribución de los símbolos se mejora significativamente los resultados de potencia de los tests propuestos. Por último propondremos un nuevo test de heterocedasticidad, de aplicación práctica como principal novedad, cuyas características principales son que 1) utiliza el análisis de patrones, 2) se basa en la técnica bootstrap, por lo que ya no se fundamenta como los tests  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  en resultados asintóticos, 3) es semiparamétrico, 4) utiliza la Entropía de Tsallis como medida de la distribución de los símbolos y 5) no necesita estimar extrínsecamente el índice entrópico  $q$ . Los resultados de potencia de este nuevo test muestran una mejora significativa en la detección de los procesos heterocedásticos con autocorrelaciones temporales (memoria larga), mientras se mantienen los resultados de tamaño en el nivel nominal. Esto nos permiten corroborar empíricamente que la Entropía de Tsallis es una buena medida de la distribución de los símbolos en las series heterocedásticas de largo alcance.

Este capítulo se organiza como sigue. En la sección 3.2 se realiza una breve introducción al formalismo matemático y estadístico inducido por la Entropía de Tsallis. En la sección 3.4 se explica la conexión funcional entre heterocedasticidad y Entropía de

Tsallis. En la sección 3.5 se describe el marco de diseño de los nuevos tests que se van a proponer en este capítulo. En la sección 3.6 se describe el procedimiento de generalización de los tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  mediante el formalismo de Tsallis. En la sección 3.7 se propone un nuevo test bootstrap de heterocedasticidad, también en el contexto de la entropía de Tsallis. En la sección 3.8 se presentan los resultados de tamaño y potencia de los tests propuestos. En la sección 3.9 se compara el nuevo test bootstrap propuesto con los tests de heterocedasticidad más populares. Finalmente en la sección 3.10 se aportan las conclusiones.

## 3.2. Introducción al Formalismo de Tsallis

### 3.2.1. Introducción a la Entropía de Tsallis

En esta sección daremos algunos conceptos y notaciones referentes la Entropía de Tsallis que son necesarios para una posterior comprensión del proceso de generalización de los tests propuestos en el capítulo anterior<sup>2</sup>.

La generalización de la Entropía de Shannon que postuló Tsallis en el marco de la no-extensividad en sistemas discretos se define de la siguiente manera:

$$S_q^{(d)} = \frac{1 - \sum_{i=1}^N p_i^q}{q - 1} = \frac{1}{q - 1} \sum_{i=1}^N p_i (1 - p_i^{q-1}) \quad (3.7)$$

donde  $(p_i)_{i=1}^n$  es la distribución de probabilidad. La entropía de Tsallis recupera la entropía de Shannon cuando  $q \rightarrow 1$ <sup>3</sup>

$$\lim_{q \rightarrow 1} S_q^{(d)} = - \sum_{i=1}^n p_i \ln p_i \quad (3.8)$$

La Entropía de Tsallis se extiende a la ley pseudo-aditiva

$$S_q(A + B) = S(A) + S(B) - (q - 1)S(A)S(B) \quad (3.9)$$

y decimos que un sistema es no extensivo a causa de la cuasi-aditividad de las Entro-

<sup>2</sup>Para una completa descripción de la teoría formal de la estadística no-extensiva de Tsallis ver [80].

<sup>3</sup>Cuando  $q \rightarrow 1$ ,  $p_i^{q-1} = e^{(q-1)\ln(p_i)} \sim 1 + (q-1)\ln(p_i)$ .

pías. En relación con la propiedad de pseudo-aditividad 3.9 de  $S_q$ , se han introducido generalizaciones de funciones, distribuciones de probabilidad y operaciones matemáticas elementales que dan lugar a un nuevo formalismo matemático-estadístico conocido como las Matemáticas de Tsallis y las Estadísticas de Tsallis de las que a continuación haremos un breve resumen.

### 3.2.2. Introducción a las Matemáticas de Tsallis

La base matemática de la estadística de Tsallis proviene de las expresiones deformadas de las funciones exponencial y logarítmicas que son la función *q-logarítmica*:

$$\ln_q x := \frac{x^{1-q} - 1}{1 - q} \quad (x \geq 0, q \text{ para } t \in \mathbb{R}) \quad (3.10)$$

la cual para  $q = 1$  se convierte en el logaritmo natural y su función inversa la función *q-exponencial*:

$$\exp_q(x) := \begin{cases} [1 + (1 - q)x]^{1/(1-q)} & \text{si } 1 + (1 - q)x > 0 \\ 0 & \text{resto} \end{cases} \quad (x \in \mathbb{R}) \quad (3.11)$$

la cual se convierte en la función exponencial para  $q = 1$ . La importancia del *q-logaritmo* es que satisface la ley pseudo-aditiva:

$$\ln_q xy := \ln_q x + \ln_q y + (1 - q)\ln_q x \ln_q y \quad (3.12)$$

Usando la función *q-logarítmica*, la entropía de Tsallis puede escribirse como la versión *q-deformada* de la Entropía de Shannon:

$$S_q^{(d)} = - \sum_{i=1}^N p_i^q \ln_q p_i = \sum_{i=1}^N p_i^q \ln_q((1/(p_i))) = \langle \ln_q((1/(p_i))) \rangle_{lin} \quad (3.13)$$

Detrás de las Estadísticas de Tsallis se han encontrado nuevas operaciones determinadas por las funciones *q-logarítmicas* y *q-exponenciales* las cuales emergen naturalmente de la Entropía de Tsallis.

Nueva operación de Multiplicación  $\otimes_q$  que satisface las siguientes operaciones:

$$\begin{aligned} \ln_q(x \otimes_q y) &= \ln_q x + \ln_q y \\ \exp_q(x) \otimes_q \exp_q(y) &= \exp_q(x + y) \end{aligned} \quad (3.14)$$

Esto nos conduce a la definición de  $\otimes_q$  entre 2 números positivos

$$x \otimes_q y = \begin{cases} [x^{1-q} + y^{1-q} - 1]^{\frac{1}{1-q}} & \text{si } x > 0, y > 0 \\ & x^{1-q} + y^{1-q} - 1 > 0 \\ 0 & \text{resto} \end{cases} \quad (3.15)$$

esta expresión recibe el nombre de *q-producto*. El *q-producto* recupera el producto usual de tal manera que  $\lim_{q \rightarrow 1}(x \otimes_q y)$ . Las propiedades del *q-producto* son casi todas las del producto usual salvo que la propiedad distributiva no se conserva en general. La operación inversa del *q-producto*  $\otimes_q$  se simboliza como  $\oslash_q$  y se define como :

$$x \oslash_q y = \begin{cases} [x^{1-q} - y^{1-q} + 1]^{\frac{1}{1-q}} & \text{si } x > 0, y > 0 \\ & x^{1-q} - y^{1-q} + 1 > 0 \\ 0 & \text{resto} \end{cases} \quad (3.16)$$

de manera análoga que para el *q-producto*  $\otimes_q$  se cumplen las siguientes propiedades por parte de  $\oslash_q$ :

$$\begin{aligned} \ln_q(x \oslash_q y) &= \ln_q(x) - \ln_q(y) \\ \exp_q(x) \oslash_q \exp_q(y) &= \exp_q(x - y) \end{aligned} \quad (3.17)$$

No solo el *q-producto*, sino también la *q-suma*  $\oplus_q$  se puede deducir de manera similar al *q-producto*  $\otimes_q$ . La *q-suma*  $\oplus_q$  se define como:

$$x \oplus_q y = x + y + (1 - q)xy \quad (x, y \in \mathbb{R}, q \in \mathbb{R}^+) \quad (3.18)$$

A su vez de la definición de la *q-suma*  $\oplus_q$  se derivan las siguientes propiedades:

$$\begin{aligned} \ln_q(xy) &= \ln_q(x) \oplus_q \ln_q(y) \\ \exp_q(x)\exp_q(y) &= \exp_q(x \oplus_q y) \end{aligned} \quad (3.19)$$

También se pueden definir el  $q$ -coeficiente binomial y multinomial respectivamente como

$$\begin{aligned} \left[ \frac{n}{k} \right]_q &= (n!_q) \oslash_q [(k!_q \otimes_q ((n-k)!_q)] \quad (n, k (\leq n) \in \mathbb{N}) \\ \left[ \frac{n}{n_1 \cdots n_k} \right]_q &= \left[ \sum_{i=1}^n l^{1-q} - \sum_{i_1=1}^{n_1} i_1^{1-q} \dots - \sum_{i_k=1}^{n_k} i_k^{1-q} + 1 \right] \end{aligned} \quad (3.20)$$

### 3.2.3. Introducción a la Estadística de Tsallis

Recordando brevemente el principio de máxima entropía para la entropía de Boltzmann-Gibbs bajo las restricciones :

$$\int f(x)dx = 1, \quad \frac{\int x^2 f(x)dx}{\int f(y)dy} = \sigma^2, \quad \langle x \rangle = 0$$

ésta conduce a la función densidad de probabilidad Gaussiana:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \quad (3.21)$$

Así pues, la distribución Gaussiana es la distribución que maximiza la Entropía de Gibbs-Boltzmann bajo restricciones apropiadas, actuando como atractor de los sistemas independientes con varianza finita. Análogamente podemos aplicar el principio de máxima entropía para la Entropía de Tsallis en su versión continua :

$$S_q(f(x)) = \frac{1 - \int f(x)^q}{q - 1} \quad (3.22)$$

bajo las restricciones naturales:

$$\int f(x)dx = 1 \quad (3.23)$$

correspondiente a la normalización y

$$\frac{\int x f(x)^q dx}{\int f(x)^q dy} \equiv \langle x \rangle_q = \mu_q \quad (3.24)$$

$$\frac{\int (x - \mu_q)^2 f(x)^q dx}{\int f(x)^q dy} \equiv \langle (x - \mu_q)^2 \rangle_q = \sigma_q^2$$

correspondientes a la  $q$ -media y  $q$ -varianza generalizadas de una cantidad relevante  $x$ ,

respectivamente, donde

$$P(x) = \frac{f(x)^q}{\int f(x)^q dy} \quad (3.25)$$

es la expresión de la Distribución Escort [80].

Todo ello conduce a la denominada función densidad de probabilidad *q-Gaussiana* [80], [28]

$$f_q(x) = \frac{\exp_q(-\beta_q x^2)}{\int \exp_q(-\beta_q y^2) dy} = A_q [1 + (q-1)B_q(x - \mu_q)^2]^{(1/(1-q))}, \quad q < 3 \quad (3.26)$$

donde

$$A_q = \begin{cases} \frac{\Gamma(\frac{5-3q}{2-2q})}{\Gamma(\frac{1-q}{2-2q})} \sqrt{\frac{1-q}{\pi}} B_q & \Leftarrow q < 1 \\ \frac{\Gamma(\frac{1}{q-1})}{\Gamma(\frac{3-q}{2q-2})} \sqrt{\frac{1-q}{\pi}} B_q & \Leftarrow q > 1 \end{cases} \quad (3.27)$$

y

$$B_q = [(3-q)\sigma_q^2]^{-1} \quad (3.28)$$

Como se puede observar, el límite superior  $q = 3$  garantiza la normalizabilidad. Partiendo de la definición de la *q-exponencial*, podemos reescribir la función densidad de probabilidad como:

$$f_q(x) = A_q e_q^{-B_q(x-\mu_q)^2} \quad (3.29)$$

y los momentos de segundo orden estándar y generalizados se relacionan, en el intervalo  $1 \leq q < 5/3$  por la expresión:

$$\bar{\sigma}^2 = \bar{\sigma}_q^2 \frac{3-q}{5-3q} \quad (3.30)$$

Teniendo en cuenta lo expuesto anteriormente, según los diferentes rangos de valores que puede tomar  $q$  se deducen las siguientes propiedades de  $f(x)$ :

- Para  $q \rightarrow 1$ , la distribución *q-gaussiana* recupera la distribución *Gaussiana* usual.

- Para  $q > 1$  el comportamiento asintótico de las distribuciones sigue una ley de potencias.
- El rango  $1 < q \leq 5/3$  pertenece a las distribuciones *q-gaussianas* con características heavy tail (y varianza finita).
- El rango  $q > 5/3$  es característico de las Distribuciones de Lévy.
- Para  $q = 2$ ,  $f(x)$  da lugar a la Distribución Lorentziana.
- La varianza de estas distribuciones es finita para  $q < 5/3$  y diverge para  $5/3 < q < 3$ .
- Para  $q \geq 3$  la función no es normalizable e inaceptable para una distribución de probabilidades.

### 3.3. Estadísticas, Entropía de Tsallis y Heterocedasticidad

Para mostrar que la Entropía de Tsallis es un funcional adecuado para describir las series temporales con dependencia temporal en la varianza, vamos a reseñar de forma breve en este apartado cómo las Estadísticas de Tsallis pueden reproducir las propiedades dinámicas y estadísticas principales que caracterizan algunos de los procesos heterocedásticos, y cuyo desarrollo detallado se describe en varios de los artículos publicados por Constantino Tsallis y Silvio Duarte Queirós (veáse por ejemplo [26, 27, 28]).

En primer lugar describimos la relación de las Estadísticas de Tsallis con los procesos heterocedásticos de forma funcional determinista, *ARCH* y *GARCH*. De [28], se parte de la expresión introducida por Engle para los procesos *ARCH*( $n$ ) para una variable heterocedástica  $u_t$

$$u_t = \sigma_t v_t \tag{3.31}$$

donde  $\sigma_t$  es la volatilidad y  $v_t$  es una variable i.i.d asociada con una función densidad de probabilidad  $P_n(\nu)$ . La dinámica de las volatilidades, en primera aproximación sigue

una dependencia lineal de los  $n$  valores previos de  $[u_t^2]$

$$\sigma_t^2 = \alpha + \sum_{i=1}^n \beta_i u_{t-i}^2 \quad (\alpha, \beta_i > 0) \quad (3.32)$$

Después, con el objetivo de dar más flexibilidad a la forma funcional de  $\sigma_t^2$  Bollerslev generalizó la ecuación 3.32, definiendo el proceso  $GARCH(n,m)$

$$\sigma_t^2 = \alpha + \sum_{i=1}^n \beta_i u_{t-i}^2 + \sum_{j=1}^m \gamma_j \sigma_{t-j}^2 \quad (\alpha, \beta_i > 0) \quad (3.33)$$

Se puede verificar que la distribución  $P(u)$  de la variable estocástica  $u$ , tiene para cada instante de tiempo  $t$ , la misma forma funcional que la distribución del ruido  $P(\nu)$  pero con varianza  $\sigma_t$ . Esta propiedad permite mirar al proceso  $\{u_t\}$  como un proceso similar a aquellos que ocurren en algunos de los *sistemas de no-equilibrio* con un estado estacionario duradero. Así, de este principio se han podido derivar relaciones, primero para el proceso  $ARCH(1)$  y después para el  $GARCH(1,1)$  [28] entre los parámetros  $\beta$  y  $\gamma$ , y las distribuciones  $P(u)$  y  $P_n(\nu)$ , donde en ésta última se asume que tiene la siguiente forma de distribución  $q_n - Gaussiana$

$$P_n(\nu) = \frac{A_{q_n}}{\left[1 + \frac{q_n-1}{5-3q_n} \nu_t^2\right]^{\frac{1}{q_n-1}}} = A_{q_n} e^{-\nu^2/(5-3q_n)}, \quad (q_n < 5/3) \quad (3.34)$$

Haciendo el Ansatz  $P(u) \simeq p(u)$ , donde  $p(u)$  es la función densidad de probabilidad  $q$ -*Gaussiana* que maximiza  $S_q$  (expresión 3.7) y haciendo coincidir los momentos de orden 2 y 4, es posible establecer para el proceso  $GARCH(1,1)$ , una relación conteniendo los parámetros dinámicos  $\beta$  y  $\gamma$  y los índices entrópicos  $q$  y  $q_n$  :

$$\beta(5-3q_n)(2-q) = \sqrt{(q-q_n) - [(5-3q_n)(2-q) - \gamma^2(5-3q)(2-q_n)]} - \gamma(q-q_n) \quad (3.35)$$

Para  $\gamma = 0$  se reduce al correspondiente  $ARCH(1)$

$$q = \frac{q_n + 2b^2(5-3q_n)}{1 + b^2(5-3q_n)} \quad (3.36)$$

y para  $\beta = \gamma = 0$ ,  $q = q_n$ . En [28] se ha comprobado la validez empírica de las expresiones 3.35 y 3.36 .

En segundo lugar, los modelos de volatilidad estocástica y las distribuciones *q-Gaussianas* se relacionan en el contexto de los sistemas dinámicos no lineales denominados mesoscópicos [80], en el sentido que a la evolución temporal de los ingredientes determinísticos se añade un ruido estocástico. Los modelos de volatilidad estocástica son soluciones de la ecuación de difusión de Fokker-Plank, la cual describe básicamente la evolución temporal de la distribución de probabilidad en el espacio de fases de los sistemas mesoscópicos:

$$\frac{\delta p(x, t)}{\delta t} = D \frac{\delta^2 p(x, t)}{\delta x^2} \quad (D > 0) \quad (3.37)$$

donde  $D$  es el coeficiente de difusión y cuya solución básica correspondiente a la condición inicial  $p(x, 0) = \delta(x)$  es

$$p(x, t) = \frac{1}{\sqrt{2\pi Dt}} e^{-x^2/2Dt} \quad (3.38)$$

La generalización en el contexto de la Entropía de Tsallis da lugar a

$$\frac{\delta^\beta p(x, t)}{\delta t} = D_{\gamma, q} \frac{\delta^\gamma [p(x, t)]^{2-q}}{\delta |x|^\gamma} \quad (0 < \gamma \leq 2; q < 3) \quad (3.39)$$

que para  $\beta = 1$  da lugar a cuatro tipos de soluciones estables, en el sentido que si partimos con una solución arbitraria  $p(x, 0)$ , asintóticamente se aproximan a cualquiera de las cuatro tipos de distribuciones que se describen:

1. *La Clase de las Distribuciones Gaussianas*, correspondiente a los valores de los parámetros  $q = 1$  y  $\gamma = 2$ . Sus soluciones básicas son la distribuciones gaussianas. Esto se corresponde con el *Teorema del Límite Central G-CLT*. Este teorema dice que si tenemos la suma o media aritmética de  $N$  variables probabilísticamente independientes y de varianzas finitas, entonces su distribución se aproxima a una gaussiana para  $N \rightarrow \infty$ .
2. *La Clase de las Distribuciones de Levy o clase  $\alpha$*  con  $\alpha = \gamma$  correspondiente con

los valores  $q = 1$  y  $0 < \gamma < 2$ . Sus soluciones básicas son las distribuciones de Lévy (también llamadas distribuciones *a-estables*). Se corresponde con el *Teorema central de Lévy-Gnedenko (L-CLT)*. Este teorema dice que si tenemos la suma o media arimética de  $N$  variables aleatorias que son probabilísticamente independientes y tienen varianza infinita entonces la distribución se aproxima a una distribución de Lévy cuando  $N \rightarrow \infty$ .

3. *La Clase de las Distribuciones  $q$ -Gaussianas* correspondiente con  $\gamma = 2$  y  $q \neq 1$ . Sus soluciones básicas son las distribuciones  *$q$ -Gaussianas*. En estas circunstancias se establece el *Teorema Central Generalizado ( $q$ -G-CLT)* y se corresponde con la violación de la hipótesis de independencia. Una dependencia que es de largo alcance y que no se diluye con  $N \rightarrow \infty$ .
4. *La Clase de las Distribuciones  $(q, \alpha)$  estables* se corresponde con  $\gamma \equiv 2\alpha < 2$  y  $q \neq 1$ . Sus soluciones básicas son las distribuciones  $(q, \alpha)$  estables.

### 3.3.1. Cuantificación del grado de Dependencia de la Varianza

Una forma interesante de evaluar numéricamente el grado de dependencia no lineal en la varianza de una serie temporal heterocedástica se describe en el artículo de Duarte Queirós [27], donde como principal novedad se propone el uso de la generalización no-extensiva de la medida de Información de Kullback-Leibler.

$$I_q = -\sum_t p(\varepsilon_t^2, \varepsilon_{t+\tau}^2) \frac{\left[ \frac{p'(\varepsilon_t^2, \varepsilon_{t+\tau}^2)}{p(\varepsilon_t^2, \varepsilon_{t+\tau}^2)} \right]^{1-q} - 1}{1 - q} \quad (3.40)$$

donde  $p'(\varepsilon_t^2, \varepsilon_{t+\tau}^2) = p(\varepsilon_t^2)p(\varepsilon_{t+\tau}^2) = |p(\varepsilon_t^2)|^2$  (asumiendo estacionariedad), y  $\varepsilon_t^2$  son los residuos al cuadrado del modelo 2.25. Esta medida basada en el concepto de información ha sido capaz de proporcionar resultados interesantes en los problemas de dependencia (ver [27]).

El ratio 3.41 junto con la ecuación 3.42 permiten definir un criterio para medir el grado de dependencia entre variables aleatorias. La cuantificación del grado de dependencia se realiza a través del valor de  $q^{op}$ , el cual se corresponde con el punto de inflexión de la versión normalizada de  $I_q$ ,

$$R_q \equiv \frac{I_q}{I_q^{max}} \in [0, 1] \quad (3.41)$$

donde  $I_q^{max}$  es el valor de  $I_q$  cuando las variables  $\varepsilon_t$  y  $\varepsilon_{t+\tau}$  presentan una relación biunívoca y viene definida por la expresión (para una descripción completa ver [16]).

$$I_q^{max}(\varepsilon_t^2, \varepsilon_{t+\tau}^2) = -E_{p(x,y)}^q(x, y) [\ln_q(p(\varepsilon_t^2) + (1-q)\ln_q p(\varepsilon_t^2)\ln_q p(\varepsilon_{t+\tau}^2))] \quad (3.42)$$

donde  $E_{p(x,y)}^q$  denota el valor de la  $q$ -esperanza con respecto a la distribución  $p(x, y)$ .

$R_q$  varía entre 0 (correspondiente a correlación cero, i.e, independencia entre  $\varepsilon_t^2$  y  $\varepsilon_{t+\tau}^2$ ) a 1 (correspondiente a la correlación máxima, i.e., aplicación uno a uno entre  $\varepsilon_t^2$  y  $\varepsilon_{t+\tau}^2$ ). Notar que hay un  $q^{op}$  para un determinado grado de correlación, esto quiere decir que el gradiente del índice  $R$  es más sensible y por lo tanto más capaz de detectar pequeñas variaciones de la correlación entre variables. El criterio  $R$  es (i) *consistente*, en el sentido que varía de forma monótona con el grado de dependencia de la variable aleatoria (ii) *versátil*, por que incluye un parámetro, el índice entrópico  $q$ , que se puede tunear a un valor óptimo  $q^{opt}$  para cada problema específico y para el cual la sensibilidad del criterio es máxima (iii) *óptimo*, ya que sus dos valores extremos se corresponden con la independencia completa y la dependencia completa.

Para series infinitas se puede demostrar que, cuando las variables son completamente dependientes  $q^{op} = 0$  mientras que  $q^{op} = \infty$  cuando las variables son independientes. Además, para sistemas finitos hay un nivel de ruido,  $q_n^{op}$ , que se consigue después de un retardo temporal finito  $\tau$  (para más detalles ver [27]).

### 3.4. De la Entropía de Tsallis a la Entropía de Permutación de Tsallis

En esta sección describimos el concepto de Entropía de Permutación de Tsallis Normalizada, la cual ya se anticipó en [48]. Esto nos servirá para introducir *la Entropía Simbólica de Tsallis* en la sección 3.6. Para ello retomamos los conceptos ya presentados en las secciones 2.2 y 2.3.

Dada una serie temporal  $\{X_t\}_{t \in I}$  una serie temporal de valores reales, una dimensión de inmersión  $m > 1$  e  $I$  un conjunto de índices temporales, consideramos los patrones ordinales de orden  $m$  generados por

$$X_m(t) = (X_{t+1}, X_{t+2}, \dots, X_{t+m}) \text{ para } t \in I$$

que da lugar a  $m!$  permutaciones posibles  $\pi_i$  de orden  $m$ , cuya distribución de probabilidad  $f = \{p(\pi_i), i = 1, \dots, m!\}$  está dada por las frecuencias relativas

$$p(\pi) = p_\pi = \frac{\{t \in T \mid t \text{ tiene patrón ordinal de tipo } \pi\}}{T - (m - 1)}$$

siendo  $T$  la cardinalidad del conjunto  $I$ .

Entonces la *Entropía de Permutación de Tsallis Normalizada* es solamente la Entropía de tsallis normalizada asociada con la distribución de probabilidad  $f$ ,

$$H_{S_q} = S/S_{q, \max} = \frac{1}{1 - m^{1-q}} \sum_{i=1}^{m!} p(\pi_i) - (p(\pi_i))^q \quad (3.43)$$

donde el valor particular  $S_{q, \max} = (1 - m^{1-q})/(q - 1)$  se alcanza para la distribución equiprobable  $f_e = \{1/m!, \dots, 1/m!\}$ .

### 3.5. Marco Estadístico de Definición y Diseño de los Tests

El marco teórico de diseño de los nuevos tests propuestos de este capítulo es el definido en el apartado 2.5 del capítulo dos, por lo que estos tests también son de naturaleza semiparamétrica. Esto es, suponemos que tenemos un proceso genérico temporal estacionario  $\{Y_t\}_{t \in I}$  lineal y correctamente especificado en media de la forma

$$Y_t = \beta Y_{t-1} + \varepsilon_t \quad t = 1, \dots, T.$$

y la hipótesis nula

$$\mathbb{H}_0 : E(\varepsilon_t^2 | I_{t-1}) = h_t = h,$$

contra la hipótesis alternativa de heterocedasticidad condicional ( $\mathbb{H}_1$ ), en la que la varianza condicional  $h_t = g(\varepsilon_t^2)$  es una función de  $t$ , y  $g(\cdot)$  es una función general desconocida de  $\varepsilon_t^2$ .

Los residuos estimados resultan de la expresión

$$\hat{\varepsilon}_t = Y_t - \hat{\beta}Y_{t-1}, \quad (3.44)$$

donde  $\hat{\beta}Y_{t-1}$  es un estimador OLS o de Máxima Verosimilitud de la media condicional y  $\hat{\beta}$  un estimador consistente de  $\beta$ .

De esta forma  $\hat{\varepsilon}_t \xrightarrow{p} \varepsilon_t$ , por lo que contrastar la hipótesis nula sobre  $E(\hat{\varepsilon}_t^2|I_{t-1})$  es asintóticamente equivalente a contrastar  $E(\varepsilon_t^2|I_{t-1})$ . Por este motivo la esencia de los nuevos tests que se van a proponer reside en el análisis simbólico de la serie de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in I}$ .

### 3.6. Aplicación del Formalismo de Tsallis al Análisis Simbólico.

#### Los Tests $G_h^q(m)$

El siguiente paso consiste en obtener, en el contexto de la Entropía de Tsallis, unos nuevos tests de heterocedasticidad. En primer lugar obtendremos una nueva expresión del test no paramétrico  $G(m)$  de independencia que llamaremos  $G^q(m)$ . Una vez obtenida la expresión analítica del test  $q$ -generalizado  $G^q(m)$ , hallaremos las expresiones correspondientes a los tests  $G_{h_p}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_Q}(m)$  que se describieron en el capítulo anterior (ver secciones 2.6, 2.7 y 2.8) y que denominaremos  $G_{h_p}^q(m)$ ,  $G_{h_{M_e}}^q(m)$  y  $G_{h_Q}^q(m)$  respectivamente.

#### 3.6.1. El Test de Independencia $G^q(m)$

Sea  $\{X_t\}_{t \in I}$  un proceso temporal y  $m$  una dimensión fija integrada en el mismo. Como hemos visto en la sección 3.5 para construir el test de independencia temporal en el contexto de la entropía de Tsallis se considera la hipótesis nula:

$$H_0 : \{X_t\}_{t \in I} \text{ i.i.d} \quad (3.45)$$

contra cualquier alternativa de dependencia  $\{X_t\}_{t \in I}$  no *i.i.d.*

Procediendo de forma análoga a la sección 2.4, la expresión de la función de Verosimilitud bajo la hipótesis  $\mathbb{H}_0$  era:

$$L(p_{\pi_1}, p_{\pi_1}, \dots, p_{\pi_n}) = \frac{R!}{(n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!)} p_{\pi_1}^{n_{\pi_1}} p_{\pi_2}^{n_{\pi_2}} \dots p_{\pi_n}^{n_{\pi_n}} \quad (3.46)$$

Para contemplar la posibilidad de *q-dependencia* en el test  $G_h(m)$  bajo la hipótesis de dependencia en la varianza (hipótesis  $\mathbb{H}_1$ ), aplicamos el *q-producto de Tsallis* (ver apartado 3.2.2) a la función de verosimilitud<sup>4</sup> y resulta:

$$L^q(p_{\pi_1}, p_{\pi_1}, \dots, p_{\pi_n}) = \frac{R!}{n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!} \otimes_q \overbrace{p_{\pi_1}^q \otimes_q \dots \otimes_q p_{\pi_1}^q}^{n_{\pi_1}} \cdot \overbrace{p_{\pi_2}^q \otimes_q \dots \otimes_q p_{\pi_2}^q}^{n_{\pi_2}} \dots \overbrace{p_{\pi_n}^q \otimes_q \dots \otimes_q p_{\pi_n}^q}^{n_{\pi_n}} \quad (3.47)$$

Aplicando el *q-Logaritmo y sus propiedades* a la función *q-Verosimilitud* resultante obtenemos la expresión:

$$\begin{aligned} Ln_q(L_q(p_{\pi_1}, \dots, p_{\pi_n})) &= Ln_q\left(\frac{R!}{n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!}\right) + Ln_q\left(\overbrace{p_{\pi_1}^q \otimes_q \dots \otimes_q p_{\pi_1}^q}^{n_{\pi_1}}\right) + \dots \\ &+ Ln_q\left(\overbrace{p_{\pi_2}^q \otimes_q \dots \otimes_q p_{\pi_2}^q}^{n_{\pi_2}}\right) + \dots + Ln_q\left(\overbrace{p_{\pi_n}^q \otimes_q \dots \otimes_q p_{\pi_n}^q}^{n_{\pi_n}}\right) = \\ &= Ln_q\left(\frac{R!}{n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!}\right) + \sum_{i=1}^{n_{\pi_1}} Ln_q(p_{\pi_1}^q) + \dots + \sum_{i=1}^{n_{\pi_n}} Ln_q(p_{\pi_n}^q) \end{aligned} \quad (3.48)$$

Para obtener el estimador  $\hat{p}_{\pi_i}^q$  de  $p_{\pi_i}^q$  para todo  $i = 1, 2, \dots, n$ , resolvemos las ecuaciones

$$\frac{\partial Ln_q(L_q(p_{\pi_1}^q, p_{\pi_2}^q, \dots, p_{\pi_n}^q))}{\partial p_{\pi_i}^q} = 0 \quad (3.49)$$

dando como resultado

$$\hat{p}_{\pi_i}^q = \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \quad (3.50)$$

Como se puede observar el estimador generalizado  $\hat{p}_{\pi_i}^q$  de  $p_{\pi_i}^q$  recupera  $\hat{p}_{\pi_i}$  cuando  $q \rightarrow 1$

<sup>4</sup>Para indicar que estamos trabajando en el contexto de las matemáticas Tsallis, designamos las probabilidades de símbolo  $p_{\pi_i}$  con el superíndice  $q$ , es decir, como  $p_{\pi_i}^q$ .

$$\lim_{q \rightarrow 1} \widehat{p}_{\pi_i}^q = \widehat{p}_{\pi_i} = \frac{n_{\pi_i}}{R} \quad (3.51)$$

ya que  $\sum_{i=1}^n n_{\pi_i} = R$  por definición. La versión  $q$ -generalizada del estadístico del ratio de las verosimilitudes resulta de aplicar la operación inversa al  $q$ -producto  $\otimes_q$  a las  $q$ -verosimilitudes :

$$\lambda_q(Y) = \frac{\left( \frac{R!}{n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!} \otimes_q \overbrace{p_{\pi_1}^q \otimes_q \dots \otimes_q p_{\pi_1}^q}^{n_{\pi_1}} \otimes_q \dots \otimes_q \overbrace{p_{\pi_n}^q \otimes_q \dots \otimes_q p_{\pi_n}^q}^{n_{\pi_n}} \right)}{\left( \frac{R!}{n_{\pi_1}! n_{\pi_2}! \dots n_{\pi_m}!} \otimes_q \overbrace{\widehat{p}_{\pi_1}^q \otimes_q \dots \otimes_q \widehat{p}_{\pi_1}^q}^{n_{\pi_1}} \otimes_q \dots \otimes_q \overbrace{\widehat{p}_{\pi_i}^q \otimes_q \dots \otimes_q \widehat{p}_{\pi_i}^q}^{n_{\pi_n}} \right)} \otimes_q$$

Podemos observar fácilmente que si  $q \rightarrow 1$  entonces la expresión de  $\lambda_q$  recupera  $\lambda$  ya que los los  $q$ -productos se convierten en productos y las  $q$ -divisiones en divisiones. Así mismo, en el apartado 2.4 se vió que  $G(m) = -2Ln(\lambda(Y))$  sigue asintóticamente una distribución *Chi-Cuadrado* con  $k$ -grados de libertad bajo la hipótesis nula de independencia:

$$G(m) = -2Ln(\lambda(Y)) \sim \chi_k^2 \quad (3.52)$$

Análogamente se puede demostrar como

$$G^q(m) = -2Ln_q(\lambda_q(Y)) \quad (3.53)$$

sigue asintóticamente una distribución *Chi-cuadrado* con  $k$ -grados de libertad bajo la hipótesis nula de independencia en la varianza, es decir, cuando  $q \rightarrow 1$  (ver apéndice A).

Ahora aplicamos el  $q$ -Logaritmo a la versión  $q$ -generalizada del estadístico del ratio de verosimilitudes  $\lambda_q(Y)$  de manera análoga al proceso realizado para derivar la expresión

analítica de  $G(m)$

$$\begin{aligned}
G^q(m) &= -2Ln_q(\lambda_q(Y)) = -2 \left( \sum_{i=1}^{n_{\pi_1}} Ln_q(p_{\pi_1}) + \dots + \sum_{i=1}^{n_{\pi_n}} Ln_q(p_{\pi_n}) - \dots = \right. \\
&- \left. \left( \sum_{i=1}^{n_{\pi_1}} Ln_q(\hat{p}_{\pi_1}) + \dots + \sum_{i=1}^{n_{\pi_n}} Ln_q(\hat{p}_{\pi_n}) \right) \right) = -2 \left( \sum_{i=1}^n n_{\pi_i} Ln_q(p_{\pi_i}) - \sum_{i=1}^n n_{\pi_i} Ln_q(\hat{p}_{\pi_i}) \right) \\
&= \left( \frac{-2}{1-q} \right) \left( \sum_{i=1}^n n_{\pi_i} (p_{\pi_i})^{1-q} - \sum_{i=1}^n n_{\pi_i} \left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} \right) \quad (3.54)
\end{aligned}$$

aplicando la condición de simbolización estándar ( $p_{\pi_i} = \frac{1}{n}$ ) bajo la hipótesis nula

$$G^q(m) = \left( \frac{-2}{1-q} \right) \left( \sum_{i=1}^n n_{\pi_i} \left( \frac{1}{n} \right)^{1-q} - \sum_{i=1}^n n_{\pi_i} \left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} \right) \quad (3.55)$$

y teniendo en cuenta que

$$h_q(m) = \frac{1 - \sum_{i=1}^n (p_{\pi_i})^q}{q-1} = \frac{1 - \sum_{i=1}^n \left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^q}{q-1} \quad (3.56)$$

siendo  $h_q(m)$  la expresión de la Entropía de Tsallis en su versión discreta 3.7 tenemos

$$G^q(m) = \frac{-2R}{q-1} \left[ - \left( \frac{1}{n} \right)^{1-q} + \frac{1}{1 - (q-1)h_q(m)} \right] \quad (3.57)$$

Por último, se puede comprobar que cuando  $q \rightarrow 1$ ,  $G_q(m)$  recupera  $G(m)$  teniendo en cuenta que en el límite, cuando  $q \rightarrow 1$ ,  $p^{q-1} = e^{(q-1)\ln p} \sim 1 + (q-1)\ln p$ .

$$G_{q \rightarrow 1}(m) = G(m) = -2R \left[ Ln\left(\frac{1}{n}\right) + h(m) \right] = 2R [Ln(n) - h(m)] \quad (3.58)$$

### 3.6.2. Los Tests $G_{h_P}^q(m)$ , $G_{h_{Me}}^q(m)$ y $G_{h_{NQ}}^q(m)$

Como vimos en la sección 2.27 la formalización de la hipótesis nula  $\mathbb{H}_0$  de homocedasticidad es:

$$\mathbb{H}_0 : \left\{ \hat{\varepsilon}_t^2 \right\}_{t \in I} \text{ i.i.d}$$

contra cualquier otra alternativa heterocedástica:

$$\mathbb{H}_1 : \{\widehat{\varepsilon}_t^2\}_{t \in I} \text{ no i.i.d}$$

También se vió en 2.6, 2.7 y 2.8 que las aplicaciones de simbolización  $f_{h_P}$ ,  $f_{h_{M_e}}$  y  $f_{h_{N_Q}}$  de los tests  $G_{h_P}(m)$ ,  $G_{h_{M_e}}(m)$  y  $G_{h_{N_Q}}(m)$  daban lugar a las expresiones 2.33, 2.37 y 2.42 bajo la hipótesis nula de homocedasticidad:

$$G_{h_P}(m) = 2(T - m + 1)[\ln(m!) - h(S = S_m)] \sim \chi_{m!}^2$$

$$G_{h_{M_e}}(m) = 2(T - m + 1)[\ln(2^{m-1}) - h(\Gamma)] \sim \chi_{2^{m-1}}^2$$

$$G_{h_{N_Q}}(m) = 2(T - m + 1)[\ln(3^m) - h(\Gamma)] \sim \chi_{3^m-1}^2$$

que no son más que distintas expresiones del test  $G(m)$  de independencia particularizado para cada una de las aplicaciones de simbolización  $f_{h_P}$ ,  $f_{h_{M_e}}$  y  $f_{h_{N_Q}}$  en la serie temporal de residuos al cuadrado  $\{\widehat{\varepsilon}_t^2\}_{t \in I}$ .

De la misma forma partiendo de la expresión 3.57 del test de independencia  $G^q(m)$  3.57

$$G_q(m) = \frac{-2R}{q-1} \left[ - \left( \frac{1}{n} \right)^{1-q} + \frac{1}{1 - (q-1)h_q(m)} \right]$$

los tests  $q$ -generalizados  $G_{h_P}^q(m)$ ,  $G_{h_{M_e}}^q(m)$  y  $G_{h_{N_Q}}^q(m)$  a los que da lugar se expresan formalmente como  $G^q(m)$ , pero particularizando  $n$  con el número de símbolos generado por cada aplicación de simbolización

$$\begin{aligned} G_{h_P}^q(m) &= \frac{-2(T-m+1)}{q-1} \left[ - \left( \frac{1}{m!} \right)^{1-q} + \frac{1}{1 - (q-1)h_q(m)} \right] \sim \chi^2(m! - 1) \\ G_{h_{M_e}}^q(m) &= \frac{-2(T-m+1)}{q-1} \left[ - \left( \frac{1}{2^{m-1}} \right)^{1-q} + \frac{1}{1 - (q-1)h_q(m)} \right] \sim \chi^2(2^{m-1} - 1) \\ G_{h_{N_Q}}^q(m) &= \frac{-2(T-m+1)}{q-1} \left[ - \left( \frac{1}{3^m} \right)^{1-q} + \frac{1}{1 - (q-1)h_q(m)} \right] \sim \chi^2(3^m - 1) \end{aligned} \quad (3.59)$$

la convergencia asintótica de las expresiones 3.59 a la distribución  $\chi_k^2$  bajo la hipótesis nula está garantizada, ya que bajo la condición de homocedasticidad el índice entrópico  $q \rightarrow 1$  y  $\lim_{q \rightarrow 1} G^q(m) = G(m)$  (ver apéndice A).

Lógicamente, una cuestión práctica que surge de forma natural a la hora de implementar estos tests es elegir correctamente el índice entrópico  $q$  característico del proceso temporal a contrastar. Como ya hemos anticipado en la introducción, aunque el pro-

blema de la elección del valor correcto del índice entrópico  $q$  está aún abierto, en las siguientes secciones discutiremos algunas estrategias para la estimación de dicho parámetro.

### 3.6.3. Estimación Teórica del Índice Entrópico $q$

El índice entrópico  $q$  se puede considerar como el grado de memoria a largo plazo de una serie temporal. Por ejemplo, Gell-mann y Tsallis ([34]) estimaron  $q \approx 1,4$  para datos financieros de alta frecuencia (rendimientos y volúmenes). Como se puede observar, esto ocurre para valores grandes de  $q$ , en los cuales se enfatizan grandes actividades de la volatilidad. Concretamente en el caso de los procesos de varianza finita el rango de recorrido de  $q$  se encuentra, para los procesos heterocedásticos, en el intervalo  $1 < q \leq 5/3$ , mientras que en los procesos homocedásticos el valor de  $q$  se restringe a la unidad ( $q = 1$ ); puesto que como hemos visto en el apartado 3.2.3, es en ese intervalo donde las distribuciones  $q$ -gaussianas poseen varianza finita.

A continuación queremos estimar paramétricamente los valores de  $q$  característicos de cada uno de los DGP's descritos en la sección 2.9.1 del capítulo uno y que se presentaron en las Tablas 2.4 y 2.8 para el tamaño de muestra  $T=500$ . Este paso previo es necesario ya que en las secciones posteriores queremos realizar pruebas de tamaño y potencia a los tests  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_Q}^q(m)$  para un tamaño de muestra  $T=500$  y un tamaño de símbolo  $m = 4$ . El objetivo no es otro que evaluar las prestaciones de los tres tests con respecto a sus versiones “*extensivas*”, esto es, cuando incorporan la Entropía de Shannon como medida de la distribución de los símbolos generados.

El método que vamos a emplear para estimar el índice entrópico  $q$  es similar al empleado en [35] y consiste en encontrar el valor de  $\hat{q}$  que minimiza la suma de los errores cuadráticos entre la función de distribución generada por la pd.f  $q$ -gaussiana en el intervalo  $(1 \leq q \leq 5/3)$  y la función de distribución empírica generada por los datos de la serie temporal  $Y(t)$ . El proceso de estimación consiste en los siguientes pasos:

1. Sea  $Y_{ri}(t) = \{Y_{ri_1}, Y_{ri_2}, \dots, Y_{ri_T}\}_{i=1, \dots, 500}$  la réplica  $i$  de 500 réplicas de Montecarlo del DGP  $Y(t)$ ,  $Y_i(t) = \{Y_{i_1}, Y_{i_2}, \dots, Y_{i_T}\}_{i=1, \dots, 500}$  de tamaño de muestra  $T$  sobre el que queremos estimar su  $q$  entrópico característico.
2. Se escoge un valor de  $q$  del intervalo  $(1 \leq q < 5/3)$ .
3. Estimamos la función de distribución empírica de la media de las réplicas  $Y_{ri}(t)$  de  $Y(t)$   $F(E(Y_{ri}(t))) = P(E(Y_{ri}(t)) \leq x)$ <sup>5</sup>. Siendo el estimador de la media de cada réplica  $Y_{ri}(t)$ ,  $E(Y_{ri}(t))$ ,  $E(\widehat{Y_{ri}(t)}) = \overline{Y_{ri}(t)} = \frac{T}{\sum_{i=1}^T Y_{rij}/T$ .
4. Se calcula la función de distribución  $q$  – *gaussiana* correspondiente a los valores de  $\{\overline{Y_{ri}(t)}\}_{i=1}^{500}$  con el valor de  $q$  seleccionado,  $\widehat{F}_q = P_q(\overline{Y_{ri}(t)} \leq x)$ . La función de distribución  $\widehat{F}_q$  se estima a partir de la expresión formal de la función de densidad  $q$ -*gaussiana* 3.26 y de las expresiones de 3.27, 3.28 y 3.24.
5. Se calcula la suma de los errores cuadráticos entre la función de distribución empírica y la función de distribución  $q$ -*gaussiana*  $d = (\widehat{F}_q(\overline{Y_{ri}(t)}) - \widehat{F}(\overline{Y_{ri}(t)}))$ .
6. Repetimos los pasos 3, 4 y 5  $B$  veces para obtener  $B = 100$  realizaciones de  $\{d_b\}_{b=1}^B$ .
7. Calculamos el valor de la distancia entre la función de distribución empírica y la distribución  $q$  – *Gaussiana* como  $\widehat{d}_q = P_{95}(\{d_b\}_{b=1}^B)$ .
8. Repetimos los pasos 3, 4, 5, 6 y 7 para los  $q$ 's del intervalo  $(1 \leq q < 5/3)$  a incrementos  $\Delta q = 0.01$ . Como consecuencia del proceso de iteración en el proceso de incremento de  $q$ , obtenemos distintos tipos de funciones monótonas  $f(\widehat{q}) = \widehat{d}_q$  en el intervalo  $[1, 5/3)$  que nos hace confiar razonablemente en el algoritmo de estimación empleado.
9. El valor de  $\widehat{q}_i$  óptimo es el correspondiente al  $\widehat{d}_{q_i}$  que minimiza la serie  $\{\widehat{d}_q\}_{q=1}^{q=5/3}$ .

La Tabla 3.1 reporta el valor de Índice Entrópico estimado  $\widehat{q}$ , con una imprecisión

---

<sup>5</sup>La Función de distribución muestral se calcula de forma sencilla como  $\widehat{F}(E(Y(t))) = P(E(Y(t)) \leq x) = \frac{\text{Número de Observaciones de } E(Y(t)) \leq x}{\text{Número Total de Observaciones}}$

Tabla 3.1: Estimación del Índice Entrópico  $q$  bajo  $\mathbb{H}_0$ 

$\hat{q} \pm 0.019$	
DGP	T=500
S.1	1
S.2	1
S.3	1
S.4	1
S.5	1
S.6	1
S.7	1
S.8	1
S.9	1
S.10	1
S.11	1
S.12	1
S.13	1
S.14	1
S.15	1
S.16	1

de  $\pm 0.019$  para un 95 % de confianza en el intervalo de estimación con un tamaño de réplicas de Monte Carlo de 500, para cada uno de los 16 procesos homocedásticos S.1-S.16 de la Tabla 2.4. Lo fundamental que podemos observar es que el valor estimado  $\hat{q}$  es igual, tal como esperábamos, a la unidad para todas las configuraciones de la  $\mathbb{H}_0$  contempladas, por lo que como esperábamos, bajo la hipótesis nula los tests  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_Q}^q(m)$  toman su forma estándar  $G_{h_p}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_Q}(m)$  ya que como vimos en el apartado 3.6.1  $\lim_{q \rightarrow 1} G^q(m) = G(m)$ .

La Tabla 3.2 reporta el valor estimado  $\hat{q}$  característico (con un valor de imprecisión de  $\pm 0.019$  correspondiente al 95 % de confianza) correspondiente a 24 configuraciones bajo la  $\mathbb{H}_1$  correspondientes a la Tabla 2.8.

Como podemos ver el valor estimado  $\hat{q}$  en la mayoría de los casos es mayor que la unidad, confirmando que existe relación, tal y como se describió en el apartado 3.3, entre heterocedasticidad e índice entrópico  $q$ . Es importante resaltar que el índice  $\hat{q}$  de los procesos SV (P.19-P.24) ARSV y LMSV es mayor que la unidad, por lo que podemos esperar que la incorporación del índice entrópico estimado  $\hat{q}$  en los tests  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_Q}^q(m)$  pueda mejorar las prestaciones de potencia de los tests.

Tabla 3.2: Estimación del Índice Entrópico  $q$  bajo  $\mathbb{H}_1$ 

$\hat{q} \pm 0.019$	
DGP	T=500
P.1	1.3
P.2	1.39
P.3	1.34
P.4	1
P.5	1.4
P.6	1.64
P.7	1.43
P.8	1.64
P.9	1.5
P.10	1.37
P.11	1
P.12	1
P.13	1.31
P.14	1.23
P.15	1.21
P.16	1
P.17	1
P.18	1
P.19	1.33
P.20	1.34
P.21	1.22
P.22	1.29
P.23	1.3
P.24	1.29

También observamos que hay algunos procesos heterocedásticos para los que el valor de  $\hat{q}$  es igual a la unidad (P.16-P.18, P.11,P.12). Esto puede ser debido a que las aproximaciones “gruesas” que hemos adoptado en el proceso de estimación del valor de  $\hat{q}$  no logran ser suficientemente precisas en algunos DGPs. Sin embargo veremos en el apartado 3.8 que en algunos procesos la introducción del índice entrópico  $q$  no produce una mejora sustancial de la potencia con los nuevos tests propuestos debido a que ésta ya es en los resultados presentados en 2.9.1.2, muy cercana al 100% .

#### 3.6.4. Estimación Empírica del Índice Entrópico $q$

El cálculo empírico del parámetro  $q$  es un problema abierto hasta la fecha, es decir, no hay una metodología formal no paramétrica que permita estimar el índice entrópico  $q$  cuando se desconoce la distribución de las observaciones y éstas además no son de gran tamaño. No nos sirve por tanto, la metodología expuesta en el apartado 3.6.3, ya que ésta estima mediante simulaciones de MonteCarlo el parámetro  $q$  conociendo de antemano el modelo generador de los valores de la serie.

Como ya se comentó en el apartado 3.3.1, en Duarte y Queirós [27], se propone estimar el índice entrópico  $q$  basándose en la versión normalizada  $R_q$  de  $I_q$ ; siendo  $I_q$  la medida generalizada de Información No-Extensiva de la serie de los residuos al cuadrado y de su versión desplazada  $\tau$  unidades, tal y como ya se describió en el apartado 3.3.1.

$$R_q \equiv \frac{I_q}{I_q^{max}} \quad (3.60)$$

Por lo que el valor óptimo del índice entrópico  $q^{op}$  se corresponde con el punto de inflexión de  $R_q$ . Siendo

$$I_q = -\sum_t p(\varepsilon_t^2, \varepsilon_{t+\tau}^2) \frac{\left[ \frac{p'(\varepsilon_t^2, \varepsilon_{t+\tau}^2)}{p(\varepsilon_t^2, \varepsilon_{t+\tau}^2)} \right]^{1-q} - 1}{1 - q} \quad (3.61)$$

con  $p'(\varepsilon_t^2, \varepsilon_{t+\tau}^2) = p(\varepsilon_t^2)p(\varepsilon_{t+\tau}^2) = |p(\varepsilon_t^2)|^2$ .

La cuestión que permanece abierta a raíz de lo expuesto anteriormente es cómo estimar de forma no paramétrica la densidad de probabilidad bidimensional de  $\varepsilon_t^2$  y  $\varepsilon_{t+\tau}^2$ ,  $p(\varepsilon_t^2, \varepsilon_{t+\tau}^2)$ . En Abramov [3] se propone por ejemplo un algoritmo para estimar funcio-

nes densidad de probabilidad multidimensional. La posible aplicación de estos algoritmos para estimar de forma no paramétrica la densidad de probabilidad bidimensional  $p(\varepsilon_t^2, \varepsilon_{t+\tau}^2)$ , y por lo tanto índice  $q$ , queda fuera del marco de investigación de esta tesis, aunque puede abrir futuras líneas de trabajo en el desarrollo del procedimiento de estimación del índice entrópico  $q$ .

### 3.7. El Test Bootstrap $T_{h_b}^q(m)$

En las secciones anteriores hemos mostrado que los tres tests  $G_{h_p}^q(m)$ ,  $G_{h_{M_e}}^q(m)$  y  $G_{h_Q}^q(m)$  aunque parecen prometedores en la teoría en sus prestaciones de tamaño y potencia, no pueden ser utilizados de forma práctica hasta la fecha, dado que, como hemos comentado en el apartado 3.6.4, no disponemos de un procedimiento empírico de estimación no paramétrico de muestra finita del índice  $q$ . Por este motivo, en esta sección proponemos un nuevo test bootstrap de heterocedasticidad de carácter semiparamétrico y basado en la *Entropía de Tsallis* que denominaremos  $T_{h_b}^q(m)$ . Como sabemos tanto los modelos GARCH (Heterocedasticidad Condicional Autorregresiva Generalizada) como los modelos ARSV (Volatilidad Estocástica Autorregresiva) generan series cuyos cuadrados están autocorrelados, es decir muestran dependencia entre ellos. Además, en el apartado 3.3.1 hemos visto que el grado de dependencia no lineal de una serie temporal se puede medir a través de la información conjunta (ecuación 3.40) en su versión no-extensiva, de la varianza temporal estimada (serie de los residuos al cuadrado  $\hat{\varepsilon}_t^2$ ) y de la misma retardada “ $d$ ” unidades  $\hat{\varepsilon}_{t+d}^2$ . Esto implica comparar la probabilidad de densidad conjunta de ambas series con las probabilidades marginales bajo la hipótesis nula  $\mathbb{H}_0$  de homocedasticidad.

Siguiendo un camino similar, comparamos la Entropía Simbólica de la serie del producto de orden  $d$  de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+d}^2\}$  con la Entropía Simbólica de su versión remuestreada bajo la hipótesis nula de homocedasticidad. Sólo bajo la hipótesis alternativa de heterocedasticidad la diferencia entre ambas entropías deberá ser estadísticamente significativa. La Entropía Simbólica está asociada a la aplicación de simbolización  $f_{h_{NQ}}$  basada en los *N-Cuantiles* que se ha descrito en la sección 2.8 del

Capítulo 2. Finalmente para tener en cuenta la propiedad de no-extensividad asociada a los fenómenos heterocedasticos (ver sección 3.3), sustituiremos, para cuantificar la Entropía Simbólica, la Entropía de Shannon por la Entropía de Tsallis.

### 3.7.1. Procedimiento del Test $T_h^q(\tau, m)$ basado en Bootstrap

Sea la serie de los residuos al cuadrado de la serie temporal estacionaria  $\{Y_t\}_{t \in I}$  correspondiente al modelo 2.25,  $\{\hat{\varepsilon}_t^2\}_{t \in I}$  y  $\{\hat{\varepsilon}_t^2\}_{t \in I}^b$  la serie de los residuos al cuadrado remuestreados de forma independiente  $\{\hat{\varepsilon}_t^2\}_{t \in I}$ .

**Definición 1.** Sea  $\tau$  un entero positivo. Diremos que estamos bajo la hipótesis nula de homocedasticidad si  $\forall q \in [1, 5/3)$ ,  $q \in \mathbb{R}$  y  $\tau \in \mathbb{N}$

$$h_m^q(f_{h_{NQ}}(\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2)) \geq h_m^q(f_{h_{NQ}}(\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2))^b \quad (3.62)$$

donde  $\{h_m^q(f_{h_{NQ}}(\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2))\}$  y  $\{h_m^q(f_{h_{NQ}}(\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2))^b\}$  son las *entropías simbólicas de tsallis* de la series de las autocorrelaciones de orden  $\tau$  de los residuos al cuadrado de la serie original y remuestreada respectivamente. La Entropía simbólica de Tsallis se calcula según la expresión 3.7,  $q$  es el índice entrópico de Tsallis y  $f_{h_{NQ}}$  es la aplicación de simbolización  $f_{h_{NQ}} : \mathbb{R}^m \rightarrow \Gamma_m$  de la sección del capítulo uno 2.8 que se aplica a la nueva serie  $\{\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2\}_{t \in I}$  definida como:

$$\begin{aligned} \pi_i(t) &= f_{h_{NQ}}(\hat{\varepsilon}_m^2(t) \cdot \hat{\varepsilon}_m^2(t + \tau)) = f_{h_{NQ}}(\hat{\varepsilon}_{t+i_1}^2 \cdot \hat{\varepsilon}_{t+i_1+\tau}^2, \hat{\varepsilon}_{t+i_2}^2 \cdot \hat{\varepsilon}_{t+i_2+\tau}^2, \dots, \hat{\varepsilon}_{t+i_m}^2 \cdot \hat{\varepsilon}_{t+i_m+\tau}^2) = \\ &= (\hat{Q}(\hat{\varepsilon}_{t+i_1}^2 \cdot \hat{\varepsilon}_{t+i_1+\tau}^2), \hat{Q}(\hat{\varepsilon}_{t+i_2}^2 \cdot \hat{\varepsilon}_{t+i_2+\tau}^2), \dots, \hat{Q}(\hat{\varepsilon}_{t+i_m}^2 \cdot \hat{\varepsilon}_{t+i_m+\tau}^2)) \end{aligned} \quad (3.63)$$

Siendo la función  $\hat{Q}()$

$$Q(\hat{\varepsilon}_m^2(t) \cdot \hat{\varepsilon}_m^2(t + \tau)) = \begin{cases} 1 & \text{if } 0 \leq \hat{\varepsilon}_m^2(t) \cdot \hat{\varepsilon}_m^2(t + \tau) \leq C_1 \\ z & \text{if } C_{z-1} \leq \hat{\varepsilon}_m^2(t) \cdot \hat{\varepsilon}_m^2(t + \tau) \leq C_z \quad z = 2, \dots, N - 1 \\ N & \text{if } C_{N-1} \leq \hat{\varepsilon}_m^2(t) \cdot \hat{\varepsilon}_m^2(t + \tau) \leq +\infty \end{cases} \quad (3.64)$$

y  $C_1, \dots, C_{N-1}$  los N-1 cuantiles que dividen los valores de la serie temporal de la autocorrelación de orden  $\tau$  de los residuos al cuadrado  $\{\hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau}^2\}_{t \in T}$  en  $N_C$  intervalos

regulares. Al igual que en el apartado 2.8,  $N_C = 3$ .

Con estas consideraciones proponemos llevar a cabo un test no paramétrico de una cola con la siguiente hipótesis nula

$$\mathbb{H}_0 : \left\{ \widehat{\varepsilon}_t^2 \right\}_{t \in I} \text{ i.i.d,} \quad (3.65)$$

con el siguiente estadístico:

$$\hat{T}_h^q(\tau, m) = h_m^q(f_{h_{NQ}}(\widehat{\varepsilon}_t^2 \cdot \widehat{\varepsilon}_{t+\tau}^2)) \forall q \in [1, 5/3) \forall \tau \in \mathbb{N} \quad (3.66)$$

Si estamos bajo la hipótesis nula de homocedasticidad entonces  $\hat{T}_h^q(\tau, m)$  no es significativamente diferente de la entropía simbólica bajo la hipótesis nula de homocedasticidad  $\forall q \in [1, 5/3)$  y  $\forall \tau \in \mathbb{N}$ . En el caso contrario, bajo la hipótesis alternativa  $\mathbb{H}_1$  entonces  $\exists q \in [1, 5/3)$  y  $\exists \tau \in \mathbb{N}$  que maximiza estadísticamente significativa el valor de la diferencia  $\hat{T}_{h_0}^q(\tau, m) - \hat{T}_{h_1}^q(\tau, m)$ .

A diferencia de los tests asintóticos  $G_h$  descritos en los apartados anteriores, en este caso no se conoce la distribución del estadístico  $\hat{T}_h^q(\tau, m)$  bajo la hipótesis nula de homocedasticidad, por lo que necesitamos estimar su distribución a partir de la distribución empírica, esto constituye el aspecto fundamental del *Bootstrap no paramétrico*. Para que el test basado en bootstrap sea asintóticamente independiente del DGP bootstrapeado, tenemos que asegurar que el DGP remuestreado respeta la hipótesis nula de homocedasticidad en condiciones de estacionariedad. Para este fin, utilizaremos el *remuestreo independiente*, ya que con la aplicación de simbolización  $f_{h_{NQ}}$  utilizada de la sección 2.8, la entropía simbólica es equivalente al caso homocedástico, esto es, la Entropía Simbólica de  $f_{h_{NQ}}$  aplicada a la serie del producto los residuos al cuadrado y su versión retardada  $\tau$  unidades (correlación de los residuos al cuadrado) de una serie una i.i.d es equivalente a la Entropía Simbólica de  $f_{h_{NQ}}$  aplicada a la serie de los residuos al cuadrado y su versión retardada  $\tau$  unidades de una serie homocedástica. En esta ocasión la Aplicación de Simbolización  $f_{h_{NQ}}$  bajo la hipótesis  $\mathbb{H}_0$  de independencia en la varianza es una Aplicación de Simbolización *no estándar*, esto es, los símbolos resultantes de la aplicación no son equiprobables, a diferencia del caso que se describió en la sección 2.8, en la cual, como recordamos, la Aplicación de Simbolización era

Estándar.

Por lo tanto el procedimiento del test bootstrap, con  $B$  réplicas bootstrap se compone de los siguientes pasos:

1. Elegir  $q_s \in [1, 5/3)$  y  $\tau_0 \in \mathbb{N}$ .
2. Mediante el remuestreo de la serie de residuos al cuadrado  $\{\hat{\varepsilon}_t^2\}_{t \in T}$  obtenida de  $\{Y_t\}_{t \in T}$ , obtener la serie de residuos remuestreados  $\{\hat{\varepsilon}_t^2\}_{t \in I}^b$  siendo  $b$  el número de observaciones remuestreadas.
3. Para las observaciones remuestreadas  $\{\hat{\varepsilon}_t^2\}_{t \in I}^b$  obtener la Entropía Simbólica de Tsallis expresada como :

$$\hat{h}_m^b(\tau_0, q_s) = \hat{h}_m^{q_s}(f_{h_{NQ}} \{ \hat{\varepsilon}_t^2 \cdot \hat{\varepsilon}_{t+\tau_0}^2 \}^b) \quad (3.67)$$

4. Repetir  $B-1$  veces los pasos 2 y 3 para obtener las  $B$  realizaciones remuestreadas de la entropía simbólica de Tsallis  $\{\hat{h}_m^b(\tau_0, q_s)\}_{b=1}^B$
5. Calcular el  $p_{boots}$ -valor:

$$p_{boots} - valor(\hat{h}_m^{q_s}(\tau_0)) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}(\hat{h}_m^b(\tau_0, q_s) \leq \hat{h}_m^{q_s}(\tau_0)) \quad (3.68)$$

donde  $\mathbf{1}(\cdot)$  es la función indicador que asigna 1 si la sentencia es verdadera y 0 en otro caso.

6. Elegir el valor mínimo de los  $p_{boots} - valor(\hat{h}_m^{q_s}(\tau_0))$  con  $q_s \in [1, 5/3)$
7. Rechazar la hipótesis nula de homocedasticidad si

$$\min_{\substack{q_s \in [1, 5/3) \\ \tau_0 \in \mathbb{N}}} \left\{ p_{boots} - valor(\hat{h}_m^{q_s}(\tau_0)) \right\} < \alpha \quad (3.69)$$

para un tamaño nominal  $\alpha$ .

### 3.7.2. Selección de los parámetros $\tau$ y $q$

Una cuestión práctica en la implementación del test  $T_h^q(\tau, m)$  es la elección del orden del retardo  $\tau$  y el índice entrópico  $q$ . Como ventaja, nuestro test bootstrap proporciona en el propio estadístico  $\hat{T}_h^q(\tau, m)$  el criterio para elegir  $\tau$  y  $q$ , ya que, hasta cierto punto, deja que sean los datos por sí mismos a través de la dinámica simbólica, los que pregonen los valores de  $\tau$  y  $q$  más adecuados. Así los valores de  $\tau$  y  $q$  óptimos serán aquellos que minimicen el valor de  $p_{boots} - valor(\hat{h}_m^q(\tau_0))$ , con la condición que éste se encuentre lo más cerca posible del tamaño nominal  $\alpha$ . Como podemos ver, a diferencia del problema que adolecen los tests  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$ , en este caso hemos conseguido estimar dentro del propio procedimiento del test  $T_h^q(\tau, m)$  el parámetro entrópico  $q$  del proceso que minimiza  $\{p_{boots} - valor(\hat{h}_m^{q_s}(\tau_0))\}$ . Es preciso decir que por razones eminentemente prácticas no hemos dedicado tiempo para encontrar la relación entre el parámetro  $q_s$  que minimiza el  $\{p_{boots} - valor(\hat{h}_m^{q_s}(\tau_0))\}$  y el índice entrópico  $q$  característico del proceso  $\{Y_t\}_{t \in I}$ . Además, como se puede observar, el proceso de estimación de los valores óptimos de  $\tau$  y  $q$  se realiza mediante un proceso heurístico ya que no hay procedimiento formal de búsqueda de  $\tau$  y  $q$ , aunque la búsqueda será normalmente en orden creciente; lo cual puede aumentar los tiempos de computación del test<sup>6</sup>.

## 3.8. Pruebas Empíricas

En esta sección investigaremos el rendimiento muestral finito de los nuevos tests simbólicos de heterocedasticidad  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  y  $T_h^q(\tau, m)$ . El primer objetivo es verificar si la introducción de los valores del índice entrópico teórico  $q$  en los tests  $G_{h_p}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  conlleva un aumento significativo de los resultados de potencia con respecto a los tests  $G_{h_p}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$ , mientras se mantienen los niveles de tamaño cerca del nominal  $\alpha$ . Para comprobarlo vamos a comparar los resultados de tamaño y potencia en los tests  $G_{h_p}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$  con los

<sup>6</sup>Sin embargo, como podremos ver en la sección siguiente 3.8 con pequeños valores del retardo  $\tau$  es suficiente para que el estadístico  $\hat{T}_h^q(\tau, m)$  rechace la hipótesis nula de homocedasticidad.

obtenidos en  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  introduciendo en estos últimos los valores de  $\hat{q}$  estimados de la sección 2.9. En segundo lugar estudiaremos las prestaciones de tamaño y potencia del Test  $T_{hb}^q(m)$  y lo compararemos con los de los tests  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$ .

Es importante resaltar que al utilizar los valores  $\hat{q}$  estimados del índice entrópico  $q$  de la sección 3.6.3, la validez e interpretación de los resultados de tamaño y potencia obtenidos en los tests  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  se restringe únicamente a un ámbito teórico y constituyen más bien un límite del rendimiento de estos tests. Como ya hemos indicado en las anteriores secciones, el motivo se debe a la carencia en la literatura científica actual de un método formal que permita estimar no-paramétricamente el índice entrópico  $q$ .

Para analizar el rendimiento finito muestral de todos los tests hemos utilizado parte de la configuración de los experimentos de Monte Carlo descritos en la sección 2.9, creando procesos generadores de datos (DGP's) <sup>7</sup> de tamaño muestral  $T=500$  observaciones, por ser un tamaño compatible con las longitudes habituales encontradas en las series temporales económicas. Asimismo hemos empleado como tamaño de símbolo  $m = 4$ . Para cada DGP este procedimiento se repitió 1000 veces y la proporción de rechazos de la hipótesis nula de homocedasticidad fue calculada usando un nivel de significación del 5%. Además, para el caso del test  $T_h^q(\tau, m)$ , las simulaciones se realizaron con los valores de los parámetros de retardo  $\tau = 1$ ,  $\tau = 2$  y el índice entrópico  $q = 1$ ,  $q = 1.2$ ,  $q = 1.5$  y  $q = 1.6$ <sup>8</sup>. Por último realizaremos, de forma análoga a como se hizo en la sección 2.9.2, la comparativa del test  $T_h^q(\tau, m)$  con algunos de los tests de heterocedasticidad más conocidos como son el de White [84], Breuch-Pagan [19], Goldfeld-Quandt [38] y Engle[29] .

---

<sup>7</sup>Nosotros proponemos usar solapamiento de las  $m$ -historias dada la relativa escasez del número de observaciones.

<sup>8</sup>Los resultados de rendimiento obtenidos para valores del retardo superiores se comprobaron que eran inferiores a los realizados para los valores de retardo escogidos en las pruebas  $\tau = 1$  y  $\tau = 2$ .

### 3.8.1. Simulaciones

#### 3.8.1.1. Tamaño de los Tests $G_{h_P}^q(m)$ , $G_{h_{M_e}}^q(m)$ y $G_{h_{N_Q}}^q(m)$

Para examinar el tamaño estadístico de los tests bajo  $\mathbb{H}_0$  consideramos los mismos DGPs de la sección 2.9.1.1 que se recogen en la Tabla 2.4 junto con los valores estimados del índice entrópico  $\hat{q}$  de la Tabla 3.1. Como podemos observar, todos los valores estimados del índice característico  $\hat{q}$  son iguales a la unidad bajo las 16 configuraciones de la  $\mathbb{H}_0$  por lo que bajo este supuesto, los tests  $G_{h_P}^q(4)$ ,  $G_{h_{M_e}}^q(4)$  y  $G_{h_{N_Q}}^q(4)$  se convierten en los tests  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{N_Q}}(4)$ . Como consecuencia se comprueba que los resultados de tamaño de los tests  $G_{h_P}^q(4)$ ,  $G_{h_{M_e}}^q(4)$  y  $G_{h_{N_Q}}^q(4)$  se reducen a los de los tests  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{N_Q}}(4)$ , valores que, como hemos visto anteriormente, se recogen en las Tablas 2.5, 2.6 y 2.7.

#### 3.8.1.2. Tamaño del Test $T_h^q(m)$

A continuación investigaremos el tamaño del test  $T_h^q(\tau, 4)$  bajo  $\mathbb{H}_0$  para  $T=500$  observaciones con los valores de los parámetros  $\tau = 1, 2$  y  $q = 1, 1.2, 1.5, 1.6$ ; considerando los 16 modelos del apartado 2.9.1.1. También estudiaremos el comportamiento asintótico del test  $T_h^q(\tau, 4)$  con  $T=1000$  pero sólo para algunos de los modelos considerados con el fin de mostrar la convergencia del test  $T_h^q(\tau, 4)$  al tamaño nominal  $\alpha$ .

La Tabla 3.3 reporta el tamaño empírico del Test  $T_h^q(\tau, 4)$  comparado con los tamaños de los tests  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{N_Q}}(4)$  bajo dieciséis configuraciones de la  $\mathbb{H}_0$  para un nivel del 5%. Primero podemos observar que el tamaño del test  $T_h^q(\tau, 4)$  está bien controlado para cualquier configuración de los parámetros  $\tau$  y  $q$  contemplados y todos los modelos considerados. Es tamaño se encuentra siempre en el intervalo de probabilidad  $[0.036, 0.064]$  correspondiente a un 95% de confianza para un tamaño del 5% y 1000 iteraciones de cada DGP. También se puede advertir que los niveles de rechazo, para esta configuración de los parámetros de simulación, son claramente inferiores a las de los tests  $G_{h_P}(4)$  y  $G_{h_{M_e}}(4)$  y se mantienen en niveles similares a los del test  $G_{h_{N_Q}}(4)$ .

Tabla 3.3: Tamaño Empírico de  $T_h^q(\tau, 4)$ ,  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{N_Q}}(4)$  para  $T=500$ 

	$T_h^q(\tau, 4)$										$G_{h_P}(4)$	$G_{h_{M_e}}(4)$	$G_{h_{N_Q}}(4)$	
	$\tau=1$													
	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$	$q=1$	$q=1.2$				$q=1.5$
S.1	0.044	0.039	0.037	0.054	0.054	0.054	0.054	0.042	0.042	0.054	0.053	0.041	0.1	0.053
S.2	0.057	0.057	0.044	0.051	0.057	0.055	0.043	0.055	0.043	0.043	0.034	0.048	0.094	0.039
S.3	0.039	0.049	0.054	0.049	0.042	0.042	0.052	0.042	0.042	0.052	0.049	0.049	0.092	0.044
S.4	0.041	0.051	0.053	0.048	0.06	0.046	0.06	0.046	0.046	0.06	0.042	0.053	0.112	0.054
S.5	0.056	0.041	0.064	0.039	0.052	0.046	0.052	0.046	0.046	0.052	0.049	0.051	0.114	0.042
S.6	0.061	0.045	0.055	0.05	0.057	0.039	0.054	0.039	0.039	0.054	0.06	0.041	0.092	0.048
S.7	0.048	0.047	0.043	0.045	0.042	0.052	0.039	0.052	0.039	0.049	0.049	0.035	0.094	0.054
S.8	0.052	0.044	0.05	0.043	0.049	0.032	0.042	0.032	0.042	0.042	0.059	0.043	0.087	0.043
S.9	0.062	0.053	0.056	0.052	0.053	0.059	0.045	0.059	0.045	0.045	0.043	0.051	0.104	0.038
S.10	0.047	0.041	0.042	0.044	0.039	0.044	0.045	0.044	0.044	0.045	0.046	0.086	0.113	0.038
S.11	0.051	0.052	0.048	0.056	0.052	0.049	0.052	0.049	0.049	0.052	0.045	0.084	0.09	0.047
S.12	0.059	0.056	0.049	0.055	0.06	0.046	0.063	0.046	0.046	0.063	0.049	0.077	0.09	0.063
S.13	0.049	0.057	0.042	0.043	0.054	0.053	0.056	0.053	0.053	0.056	0.057	0.08	0.128	0.060
S.14	0.06	0.055	0.058	0.051	0.055	0.048	0.05	0.048	0.048	0.05	0.053	0.075	0.105	0.056
S.15	0.044	0.057	0.047	0.045	0.054	0.057	0.057	0.057	0.057	0.057	0.058	0.084	0.119	0.050
S.16	0.051	0.043	0.041	0.045	0.048	0.045	0.062	0.045	0.045	0.062	0.052	0.049	0.085	0.034

Tabla 3.4: Tamaño Empírico de  $T_h^q(\tau, 4)$  para  $T=1000$ 

	$T_h^q(\tau, 4)$							
	$\tau=1$				$\tau=2$			
	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$
S.1	0.045	0.054	0.032	0.041	0.051	0.057	0.036	0.05
S.2	0.057	0.052	0.051	0.042	0.042	0.052	0.044	0.047
S.3	0.045	0.041	0.047	0.043	0.053	0.047	0.035	0.054

La Tabla 3.4 pone de relieve el tamaño empírico del test  $T_h^q(\tau, 4)$  para  $T=1000$  bajo los modelos homocedásticos S.1, S.2 y S.3. Como se puede ver, para tamaños de muestra grande ( $T=1000$ ), el test sigue estando bien dimensionado, y las frecuencias de rechazo ocurren aproximadamente a las tasas nominales de diseño. El tamaño de muestra finita no difiere significativamente en este caso tampoco del tamaño teórico, con tasas similares a las del test  $G_{h_{NQ}}(4)$ .

Es preciso aclarar que los resultados de tamaño del test  $T_h^q(\tau, 4)$  no siguen, a diferencia de los tests  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{NQ}}(4)$  una distribución asintótica conocida. Por este motivo hemos tenido que utilizar la técnica Bootstrap<sup>9</sup>. Sin embargo podemos observar que los niveles empíricos del test  $T_h^q(\tau, 4)$  son similares a los obtenidos en los tests  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{NQ}}(4)$  para  $q = 1$  con  $T = 500$  y  $T = 1000$ . Estos resultados nos permiten conjeturar que la distribución del estadístico  $\hat{T}_h^q(\tau, 4)$  para  $q = 1$  bajo la hipótesis nula puede estar relacionada con la de los estadísticos  $G_{h_P}(4)$ ,  $G_{h_{M_e}}(4)$  y  $G_{h_{NQ}}(4)$ .

<sup>9</sup>Esta es una de las aplicaciones de las técnicas Bootstrap. Es preciso matizar que entre otras, las técnicas Bootstrap también se utilizan cuando se quieren refinar los resultados de tamaño con muestras pequeñas, conociéndose la distribución del estadístico bajo la hipótesis nula.

### 3.8.1.3. Potencia de los Tests $G_{h_P}^q(m)$ , $G_{h_{Me}}^q(m)$ y $G_{h_{NQ}}^q(m)$

A continuación examinamos la potencia de los tests  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  para  $m = 4$ . Evaluamos los tests considerando los procesos (P1-P24) de la Tabla 2.8 y comparamos los resultados con los obtenidos en los tests  $G_{h_P}(4)$ ,  $G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  para  $T = 500$  recogidos en las Tablas 2.9, 2.10 y 2.11.

La Tabla 3.5 reporta las tasas de rechazo empírico de los tests  $G_{h_P}(4)$ ,  $G_{h_{Me}}(4)$ ,  $G_{h_{NQ}}(4)$ ,  $G_{h_P}^q(4)$ ,  $G_{h_{Me}}^q(4)$  y  $G_{h_{NQ}}^q(4)$  bajo los DGP's P.1-P.24 de la Tabla 2.8 para  $T=500$  y  $m=4$ .

Como podemos observar la potencia mejora en los tres tests  $G_{h_P}^q(4)$ ,  $G_{h_{Me}}^q(4)$  y  $G_{h_{NQ}}^q(4)$  con respecto a los tests  $G_{h_P}(4)$ ,  $G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  bajo todos los modelos heterocedásticos, excepto para aquellos en los que el índice entrópico estimado  $\hat{q}$  es igual a la unidad (P.4, P.11, P.12, P.16, P.17, P.18). Este incremento de la potencia es más apreciable en los modelos de volatilidad estocástica SV, sobre todo en los modelos LMSV (P.23 y P.24).

Lo más destacable es que la potencia del test  $G_{h_{NQ}}^q(4)$  es la mayor de todos los tests y es igual al 100 % para casi todos los modelos, incluso los SV.

Tabla 3.5: Potencia Empírica de los tests  $G_{h_i}(4)$  y los tests  $G_{h_i}^q(4)$ 

DGP's	$G_{h_P}(4)$	$G_{h_{Me}}(4)$	$G_{h_{NQ}}(4)$	$G_{h_{Me}}^q(4)$	$G_{h_{Me}}^q(4)$	$G_{h_{NQ}}^q(4)$
P.1	0.704	0.969	0.977	0.995	0.977	1
P.2	0.053	0.964	0.971	0.959	0.971	1
P.3	0.144	0.968	0.96	0.997	0.96	1
P.4	0.594	0.249	0.986	0.916	0.986	0.986
P.5	0.911	0.774	1	1	1	1
P.6	0.954	0.794	1	1	1	1
P.7	0.852	0.998	0.999	1	0.999	1
P.8	0.057	0.999	1	1	1	1
P.9	0.151	1	1	1	1	1
P.10	0.423	0.598	0.998	1	0.998	1
P.11	0.905	0.997	1	1	1	1
P.12	0.975	1	1	1	1	1
P.13	0.999	1	1	1	1	1
P.14	0.441	1	1	0.996	1	1
P.15	0.857	1	1	1	1	1
P.16	0.701	0.988	0.977	0.988	0.977	0.977
P.17	0.43	0.999	0.998	0.999	0.998	0.998
P.18	0.999	1	0.999	1	0.999	0.999
P.19	0.05	0.142	0.094	0.607	0.094	1
P.20	0.057	0.781	0.784	0.982	0.784	1
P.21	0.05	0.164	0.114	0.455	0.114	1
P.22	0.045	0.791	0.736	0.953	0.736	1
P.23	0.1	0.0202	0.118	0.596	0.118	1
P.24	0.04	0.085	0.062	0.887	0.062	1

### 3.8.1.4. Potencia del Test $T_h^q(\tau, m)$

En la Tabla 3.6 presentamos las tasas de rechazo del test  $T_h^q(\tau, 4)$  para los valores de los parámetros de retardo  $\tau = 1, \tau = 2$  e índice entrópico  $q = 1, q = 1.2, q = 1.5$  y  $q = 1.6$  considerando los procesos (P1-P24) de la Tabla 2.8. También incluimos los resultados obtenidos en los tests  $G_{h_P}(4), G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  para  $T = 500$  recabados de las Tablas 2.9, 2.10 y 2.11 respectivamente.

En primer lugar podemos observar que hay cierta tendencia a aumentar los niveles de rechazo para valores del parámetro  $q > 1$  y del retardo  $\tau = 2$ . En general, de los valores de  $\tau$  y  $q$  contemplados, la potencia empírica del test  $T_h^q(\tau, 4)$  es más alta para los valores del retardo  $\tau = 2$  e índice entrópico  $q = 1.5$ .

En segundo lugar, las tasas de rechazo son significativamente altas y son similares a los del test  $G_{h_{NQ}}(4)$  para los modelos heterocedásticos X-ARCH (DGP's P.1-P.18) independientemente de los valores de  $\tau$  y  $q$  excepto en el caso de P.4, bajo el cual el test  $T_h^q(\tau, 4)$  sólo alcanza niveles de potencia entre el 65 % y 74 % para  $\tau = 1$ , claramente inferiores al 98 % alcanzado en el test  $G_{h_{NQ}}(4)$ .

En tercer lugar el test  $T_h^q(\tau, 4)$  consigue mejorar de forma significativa los niveles de potencia mostrados por los tests  $G_{h_P}(4), G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  cuando los DGP's son los modelos de volatilidad estocástica SV (P.19-P.24):

- En los modelos ARSV(1) (P.19-P.22), la potencia del test  $T_h^q(\tau, 4)$  aumenta con el coeficiente de persistencia alcanzando niveles cercanos al 100 % cuando éste es igual a 0.98 (P.19-P.22) excepto en P.22 y el valor del retardo es igual  $\tau = 1$ . En este caso la tasa de rechazo es bastante baja. Por el contrario, para el valor del retardo  $\tau = 2$  la potencia del test  $T_h^q(\tau, 4)$  es en todos los casos superior a la mostrada por los tests  $G_{h_P}(4), G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  para todos los valores de  $q$ .
- En los modelos LMSV(1,d,0) (P.23 y P.24) el test  $T_h^q(4)$  domina claramente a los tests  $G_{h_P}(4), G_{h_{Me}}(4)$  y  $G_{h_{NQ}}(4)$  para los valores  $\tau = 2$  y  $q > 1$ . Esta superioridad se hace más notoria en el modelo P.24, precisamente cuando el parámetro de memoria  $d$  a largo plazo se hace más significativo ( $d=0.45$ ). Es importante notar

Tabla 3.6: Potencia Empírica de  $T_h^q(\tau, 4)$ 

		$T_h^q(\tau, 4)$												$G_{h_P}(4)$	$G_{h_{M_e}}(4)$	$G_{h_{NQ}}(4)$
		$\tau=1$						$\tau=2$								
	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$	$q=1$	$q=1.2$	$q=1.5$	$q=1.6$				
P.1	0.959	0.965	0.959	0.958	0.829	0.86	0.89	0.9	0.704	0.969	0.977					
P.2	0.98	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.053	0.964	0.971					
P.3	0.988	0.976	0.984	0.992	0.96	0.97	0.99	0.98	0.144	0.968	0.96					
P.4	0.65	0.68	0.74	0.71	0.795	0.816	0.845	0.833	0.594	0.249	0.986					
P.5	0.88	0.946	0.943	0.95	0.99	0.99	1	1	0.911	0.774	1					
P.6	0.98	0.97	0.98	1	1	1	1	1	0.954	0.794	1					
P.7	1	1	0.99	1	0.94	0.93	0.95	0.99	0.852	0.998	0.999					
P.8	1	0.99	0.99	1	1	1	1	1	0.057	0.999	1					
P.9	1	1	1	1	1	1	1	1	0.151	1	1					
P.10	0.913	0.91	0.944	0.948	0.946	0.959	0.978	0.979	0.423	0.598	0.998					
P.11	1	1	1	1	1	1	1	1	0.905	0.997	1					
P.12	1	1	1	1	1	1	1	1	0.975	1	1					
P.13	1	1	1	1	1	1	1	1	0.999	1	1					
P.14	1	1	1	1	1	1	1	1	0.441	1	1					
P.15	1	1	1	1	1	1	1	1	0.857	1	1					
P.16	1	1	1	1	0.91	0.92	0.96	0.94	0.701	0.988	0.977					
P.17	1	1	1	1	0.99	1	1	1	0.43	0.999	0.998					
P.18	0.78	0.782	0.734	0.736	0.876	0.893	0.95	0.933	0.999	1	0.999					
P.19	0.19	0.214	0.222	0.261	0.195	0.196	0.227	0.231	0.05	0.142	0.094					
P.20	0.869	0.895	0.902	0.916	0.9	0.9	0.92	0.92	0.057	0.781	0.784					
P.21	0.19	0.19	0.227	0.222	0.2	0.19	0.23	0.198	0.05	0.164	0.114					
P.22	0.069	0.062	0.046	0.056	0.88	0.9	0.9	0.918	0.045	0.791	0.736					
P.23	0.039	0.04	0.039	0.053	0.343	0.374	0.35	0.378	0.1	0.02	0.118					
P.24	0.697	0.713	0.743	0.766	0.728	0.701	0.776	0.764	0.04	0.085	0.062					

que en este caso se hace más patente la relación empírica entre valores grandes del parámetro de memoria  $q$  y los procesos de memoria a largo plazo LMSV.

### 3.8.2. Comparación con Otros Tests de Heterocedasticidad

Como hemos visto en las secciones anteriores, al no disponer hasta la fecha de un método que estime el índice entrópico  $q$  de forma no paramétrica en series con tamaños de muestra finita, no podemos comparar de forma páctica los tres tests  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  con otros tests de heterocedasticidad. Así pues, en este caso vamos a comparar el nuevo estadístico  $T_h^q(\tau, m)$  para  $m=4$ , con el parámetro de retardo  $\tau = 2$  e índice entrópico  $q = 1,5$  con los tests de heterocedasticidad de White [84] y Engle [29], para un tamaño muestral  $T = 500$ , como ya hicimos para el test  $G_{h_{NQ}}(4)$  en el apartado 2.9.2. La comparativa se ha realizando considerando los DGP's descritos en la Tabla 2.4 para  $T=500$  observaciones. Hemos elegido el valor del retardo  $\tau = 2$  y el del índice entrópico  $q = 1.5$ , ya que el test  $T_h^q(\tau, 4)$  posee para estos valores los mejores resultados de potencia frente a estos DGP's.

Para comparar el rendimiento de muestra finita de estos tests, presentamos en la Tabla 3.7 el tamaño de los tests para un nivel del 5% sobre 1000 iteraciones de cada uno para los DGP's considerados. Los resultados, para los tests de White y Engle, ya se mostraron en la Tabla 2.12 y volvemos a reproducirlos aquí junto con los resultados de tamaño del Test  $T_h^{1,5}(2, 4)$ . Como podemos ver, los niveles empíricos de  $T_h^{1,5}(2, 4)$ , White y Engle son correctos, tendiendo estos dos últimos a ser un poco más conservativos.

El siguiente paso es estudiar el comportamiento de los Tests bajo la hipótesis alternativa de heterocedasticidad. Al igual que hicimos en la sección 2.9.2, usaremos algunos de los DGP's de la Tabla 2.8. Específicamente y en primer lugar todos aquellos con estructura GARCH(1,1), así como el modelo EGARCH por ser un modelo más realista. Los resultados de potencia para los modelos con estructura GARCH(1,1) y EGARCH de la Tabla 2.13 de los estadísticos de White y Engle se presentan de nuevo en la Tabla 3.8 junto con los del estadístico  $T_h^{1,5}(2, 4)$ .

Tabla 3.7: Tamaño Empírico de  $T_h^{1,5}(2, 4)$  y varios Tests

	$T_h^{1,5}(2, 4)$	White	Engle
S.1	0.054	0.045	0.033
S.2	0.043	0.035	0.036
S.3	0.052	0.045	0.045
S.4	0.06	0.044	0.029
S.5	0.052	0.048	0.033
S.6	0.054	0.061	0.065
S.7	0.039	0.053	0.049
S.8	0.042	0.049	0.054
S.9	0.045	0.05	0.051
S.10	0.045	0.049	0.056
S.11	0.067	0.035	0.016
S.12	0.063	0.043	0.024
S.13	0.056	0.036	0.022
S.14	0.05	0.048	0.017
S.15	0.057	0.037	0.02
S.16	0.062	0.044	0.051

Tabla 3.8: Potencia empírica de  $T_h^{1,5}(2, 4)$  y Varios Tests con Modelos  $X-ARCH$ 

DGPs	$T_h^{1,5}(2, 4)$	White	Engle
P.2	0.98	1	1
P.5	1	0.941	0.8
P.8	1	0.982	0.994
P.11	1	1	0.925
P.14	1	1	1
P.18	0.94	0.66	0.72

En primer lugar observamos que el test  $T_h^{1,5}(2, 4)$  supera a los otros dos excepto en el caso P.2 que es uno de los casos para el que los errores de la estructura GARCH(1,1) siguen una distribución normal estándar. Sin embargo, para este caso las diferencias en términos de potencia no son realmente relevantes.

En segundo lugar el test  $T_h^{1,5}(2, 4)$ , al igual que ocurría con el test  $G_{h_{NQ}}(4)$  del capítulo uno, domina claramente a los otros tests cuando la varianza condicional posee estructuras asimétricas no lineales deterministas. Para el modelo P.18 (especificación ARCH no lineal que tiene en cuenta shocks asimétricos), los tests de White y Engle no son competitivos en términos de potencia. Sin embargo, el test  $T_{h_b}(4)$  obtiene niveles de rechazo espectaculares<sup>10</sup>, lo cual sugiere que es razonable considerar a este test superior a los tests de White y Engle cuando tenemos en cuenta estructuras X-ARCH clásicas y realistas en la varianza condicional.

Por último, en la Tabla 3.9 se muestra la comparativa del rendimiento de potencia de los estadísticos  $T_h^{1,5}(2, 4)$ , White y Engle respecto de los modelos de Volatilidad Estocástica de memoria a corto plazo ARSV (P.20 y P.22) y largo plazo LMSV (P.23 y P.24) descritos en la Tabla 2.8. También introducimos los resultados de potencia de los estadísticos específicos para los modelos SV, NM y  $Q_7^*(24)$  de la Tabla 2.15, descritos en el apartado 2.9.2.

<sup>10</sup>Aunque no se reporta aquí, las diferencias en términos de potencia aumentan cuando se consideran innovaciones no estándar.

Tabla 3.9: Potencia empírica de  $T_{h_b}(4)$  y varios Tests con Modelos SV

DGPs	$T_{h_b}^{1,5}(2, 4)$	White	Engle	NM	$Q_7^*(24)$
P.19	0.22	0.459	0.323	0.58	0.75
P.20	0.92	0.819	0.771	1	1
P.21	0.23	0.315	0.252	0.6	0.73
P.22	0.9	0.782	0.806	1	1
P.23	0.35	0.127	0.249	0.376	0.573
P.24	0.77	0.54	0.587	0.382	0.38

Podemos observar que el test  $T_h^{1,5}(2, 4)$  es claramente superior a los tests de White y Engle en los modelos ARSV P.20 y P.22, cuando el coeficiente de persistencia es alto. También vemos que el test  $T_h^{1,5}(2, 4)$  es superado claramente por los tests NM y  $Q_7^*(24)$  en todos los procesos ARSV (P.19 a P.22), aunque en los modelos P.20 y P.21 las tasas de rechazo de los tres tests son prácticamente equivalentes.

En el caso de los procesos LMSV (P.23 y P.24), el test  $T_h^{1,5}(2, 4)$  destaca notablemente (nivel por encima del 75 %) sobre los otros tests cuando el DGP considerado es el modelo P.24, precisamente en este modelo es cuando el parámetro de memoria a largo plazo es más significativo. En el DGP P.23, aunque las prestaciones de potencia del estadístico  $T_h^{1,5}(2, 4)$  son superiores respecto a los estadísticos de White y Engle, la potencia del estadístico  $T_h^{1,5}(2, 4)$ , está por debajo del 40 %, debido sobre todo al bajo nivel de los parámetros de persistencia y memoria a largo plazo. Las tasas de rechazo del test  $T_h^{1,5}(2, 4)$  en este caso se mantienen equivalentes al test NM, pero inferiores al test  $Q_7^*(24)$ .

Finalmente, como ya hicimos en la sección 2.9.2, también consideramos el DGP P.25, que es una variante del modelo EGARCH P.18 de la Tabla 2.8 :

$$y_t = u_t, u_t = \sigma_t \varepsilon_t, \log \sigma_t^2 = -2,4 + 0,75 \log \sigma_{t-1}^2 + 0,5 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,15 \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

Tabla 3.10: Potencia Empírica de  $T_h^{1,5}(2, 4)$  y varios Tests con otras estructuras ARCH no lineales

DGPs	$T_h^{1,5}(2, 4)$	White	Engle
P.25	1	0.012	0.012
P.26	1	0.282	0.784

y el modelo P.26 introducido por Brock [20].

$$y_t = u_t, u_t = \sigma_t \varepsilon_t, \sigma_t^2 = 1 - 2|\sigma_{t-1}^2 - 0,5|$$

siguiendo los errores en ambos modelos, una distribución estándar normal.

Los resultados se presentan en la Tabla 3.10. Como podemos ver, cuando se aumenta la complejidad de la estructura no lineal de la varianza, el comportamiento  $T_h^{1,5}(2, 4)$  es superior a los tests de White y Engle.

### 3.9. Conclusiones

El objetivo de este capítulo ha sido investigar la posibilidad de incorporar, en el contexto del Análisis Simbólico, la Entropía de Tsallis como medida “no-extensiva” que permita cuantificar estadísticamente de forma más eficiente que la Entropía de Shannon, los procesos heterocedásticos que exhiben memoria y clustering de la volatilidad (Modelos X-ARCH, SV). Al igual que la Entropía de Shannon, la Entropía de Tsallis se evalúa después de usar las Aplicaciones de Simbolización para extraer la distribución de probabilidades de símbolo asociadas, esto da lugar a una nueva forma de Entropía que llamaremos Entropía Simbólica de Tsallis y que es un concepto más amplio que la Entropía de Permutación de Tsallis descrita en trabajos anteriores [89]. Como resultado, a partir de la medida de la Entropía Simbólica de Tsallis se han obtenido unos nuevos tests semiparamétricos (de media correctamente especificada) denominados  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  y  $T_h^q(\tau, m)$ , cuyos resultados de tamaño y potencia han revelado que pueden detectar de forma efectiva los fenómenos heterocedásticos. La demostración de la relación entre Entropía Simbólica de Tsallis y heterocedasticidad se ha puesto de manifiesto:

1. A nivel teórico, mediante la generalización no-extensiva de los estadísticos  $G_{h_P}(m)$ ,  $G_{h_{Me}}(m)$  y  $G_{h_{NQ}}(m)$  a los estadísticos  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$ . Un obstáculo importante para la aplicabilidad empírica de los tests  $G_{h_P}^q(m)$ ,  $G_{h_{Me}}^q(m)$  y  $G_{h_{NQ}}^q(m)$  es la carencia en la actualidad de un procedimiento formal de estimación no-paramétrica del índice entrópico  $q$ . Es por ello que las pruebas de potencia realizadas con los valores estimados de  $q$  de cada DGP suponen una cota superior en el rendimiento de los tests. Aun así es importante destacar que los resultados de potencia muestran una mejora significativa, cercana al 100 % en casi todos los modelos, especialmente en los modelos SV.
2. A nivel empírico, mediante la propuesta de un nuevo test semiparamétrico de heterocedasticidad basado en bootstrap denominado  $T_h^q(\tau, m)$ . El test puede detectar eficientemente una amplia gama de modelos heterocedásticos, incluidos los

modelos de Volatilidad Estocástica y demuestra tener una gran potencia omnibus en relación con algunos de los tests de heterocedasticidad más importantes. En este caso, como novedad, la estimación del índice entrópico  $q$  y del parámetro de retardo  $\tau$  se realiza en el propio proceso de cálculo del estadístico  $T_h^q(\tau, m)$ . Una propiedad interesante es que los resultados de tamaño de muestra finita no difieren del tamaño nominal, lo cual nos permite garantizar la aplicabilidad general y reproducibilidad del test.

## Capítulo 4

# Selección y Contraste Estadístico de los Parámetros de Inmersión para la Reconstrucción del Espacio de Estados en Series Caóticas Temporales.

### 4.1. Introducción

La teoría de la reconstrucción del espacio de fases sugerido por Packard, Takens et al. [63],[75] es la base del análisis y predicción de los sistemas caóticos temporales. Sugiere que los sistemas caóticos temporales pueden ser reconstruidos a partir de una sola coordenada de un sistema dinámico. El teorema fundamental de reconstrucción, no da ninguna restricción sobre el retardo temporal constante  $\tau$ , mientras que para  $p$ , dimensión del espacio de fases, establece una condición suficiente (pero no necesaria) dada por  $p \geq 2d + 1$ , donde  $d$  es la dimensión fractal del atractor subyacente. El teorema de reconstrucción del espacio de fases o método de los retardos consiste en insertar una serie temporal escalar  $\{X_t\}_{t \in I}$  en un espacio  $p$ -dimensional  $X_p(t) = (X_t, X_{t+\tau}, \dots, X_{t+(p-1)\tau})$  para  $t \in I$ , donde  $\tau$  es el retardo temporal para la reconstrucción,  $p$  es la dimensión de inmersión,  $I$  es el conjunto de índices temporales de cardinalidad  $T$  y  $M$  es el número de puntos insertados en el espacio  $p$ -dimensional con  $M = T - (p - 1)\tau$ . Hay resultados teóricos que demuestran que el espacio de estados con los apropiados  $p$  y  $\tau$  tienen la misma calidad de difeomorfismo que el sistema

dinámico original.

La calidad de la reconstrucción afecta a la simulación y a modelización directamente. La selección de los parámetros  $p$  y  $\tau$  en dicha reconstrucción, no sólo es importante, sino que constituye un reto. Una elección inapropiada puede resultar en la indicación espúrea de una estructura no lineal donde no hay ninguna presente. Y a la inversa, una elección inapropiada puede resultar en un fallo al resolver estructuras que realmente están presentes en los datos. Hay una gran cantidad y creciente literatura, aunque en parte contradictoria, en la que se describen métodos candidatos de selección de los parámetros de reconstrucción, por lo que se puede decir que no hay único método óptimo de reconstrucción. Es decir, el criterio de inmersión óptima de una serie temporal no se puede determinar sólo basándose en la serie temporal. En general, la estrategia óptima de inmersión depende la serie temporal y de la medida aplicada (exponentes de Lyapunov, Dimensión de Correlación) [23]. A pesar de la confusión imperante en la literatura actual, podemos decir que hay dos puntos importantes a que tener en cuenta a la hora de seleccionar los parámetros  $p$  y  $\tau$ , ver por ejemplo [83],[57, 81, 86, 88]:

1. La selección de  $p$  es independiente de la selección de  $\tau$ . Las selecciones de  $p$  y  $\tau$  están basadas en tres métodos. El primero es el método de la correlación serial, tales como la función de autocorrelación, la información mútua, correlación de orden superior ..etc; el segundo es el método basado en la expansión del espacio de estados, tales como el factor de llenado, el desplazamiento medio, SVF (fracción de valores singulares) y otros; el último incluye el método de la múltiple función de autocorrelación y el método de la función de autocorrelación múltiple no-sesgada.
2. Los parámetros  $p$  y  $\tau$  están relacionados debido a que los conjuntos de datos reales son finitos y tienen ruido. Los tests muestran que los modelos óptimos de reconstrucción de la dinámica del atractor no dependen de la elección del retardo de inmersión elegido  $\tau$ , sino también del retardo máximo  $(p-1)\tau$ . A esta ventana de retardo máxima se la denomina ventana de inmersión  $\tau_w = (p-1)\tau$ , y es de particular importancia por que determina, en cierto sentido, la cantidad de información transferida de la serie temporal a los vectores  $\{X_p(t)\}$ . En este caso,

el retardo temporal  $\tau$  varía con la dimensión de inmersión  $p$ . Así, se originan algunos métodos de cálculo combinados, tales como el método C-C y derivados del mismo, la longitud de la ventana temporal, la inmersión automática y otros. Además en [72] se demuestra que para series temporales largas y libres de ruido, la ventana de inmersión óptima  $\tau_w$  minimiza el error de predicción del modelo de una sola fase. Este método de estimar  $\tau_w$  ha probado ser consistente y robusto. En este caso los autores afirman que en general no hay un único retardo temporal característico  $\tau$  y que la reconstrucción óptima se puede obtener considerando el vector de retardos  $(\tau_1, \tau_2, \dots, \tau_k)$  donde  $0 < \tau_i < \tau_{i+1} \leq \tau_w$ .

Muchos investigadores están de acuerdo con el punto 2 expuesto. Estos consideran que el método de la Información Mutua es engorroso computacionalmente, mientras que la función de autocorrelación al sólo tratar la dependencia lineal de las series temporales, puede arrojar un valor incorrecto del retardo temporal  $\tau_d$ . Algunos de los métodos más populares para tratar de hallar  $p$  y  $\tau$ , son aquellos que se basan en el método C-C, ya que dan simultáneamente el retardo temporal  $\tau$  y el retardo de la ventana temporal  $\tau_w$  aplicando la integral de correlación. Basados en resultados estadísticos, aunque los métodos C-C carecen de soporte teórico, funcionan bien en la práctica y poseen algunas ventajas como facilidad de operación, baja complejidad algorítmica, fiabilidad y robustez cuando se trata con pocos datos.

Como hemos visto en los capítulos anteriores, las técnicas basadas en el Análisis Simbólico son herramientas útiles en la detección de dependencias en las series temporales. Puesto que los métodos de reconstrucción del espacio de fases se basan en mayor medida en la detección de los retardos para los cuales se da algún tipo de dependencia lineal o no lineal, podemos estudiar si el análisis de patrones es una herramienta adecuada para la selección del retardo temporal  $\tau$  y la ventana de inmersión  $\tau_w$  de un sistema caótico. Así, el primer objetivo de este capítulo es proponer dos nuevos métodos de selección basados en sendas aplicaciones de simbolización con el fin de elegir de forma óptima los parámetros de inmersión  $\tau, \tau_w$ , y comparar su rendimiento con el de algunos métodos de selección utilizados actualmente.

Las series caóticas temporales reales están contaminadas por ruido. El ruido influye en la selección de los parámetros  $\tau$  y  $\tau_w$  de dos formas: (i) En primer lugar contribuye a un aumento del sesgo o desviación de los valores de  $\tau$  y  $\tau_w$  respecto de los hallados en condiciones libres de ruido; es lo que se denomina en ingeniería y física "exactitud" o error sistemático del método de selección; (ii) En segundo lugar el ruido aleatorio da lugar a que los valores de  $\tau$  y  $\tau_w$  seleccionados varíen "dispersándose" cuando repetimos el proceso de selección en las mismas condiciones, es decir, se origina una disminución de la precisión o lo que es lo mismo, un aumento del error estándar. Para controlar la variabilidad causada por el ruido aleatorio es necesario que el proceso de selección de  $\tau$  y  $\tau_w$ , independientemente del método de reconstrucción utilizado, esté basado en algún procedimiento estadístico, ya que la mera selección "visual" de los mismos puede llevar a una mala elección de los retardos que aumente el error en la reconstrucción del espacio de estados del sistema. La escasez de procedimientos estadísticos hasta la fecha para facilitar la selección de los parámetros de inmersión  $(\tau, \tau_w)$  contrasta con su importancia práctica. Así pues, el segundo objetivo de este capítulo es proponer un test estadístico no paramétrico de selección del retardo  $\tau$  y ventana temporales  $\tau_w$  adaptado a cualquier método de reconstrucción elegido.

Este capítulo se organiza como sigue. La sección 4.2 describe el método C-C y algunos métodos derivados basados en la Integral de Correlación. En las secciones 4.3 y 4.4 se describen dos aplicaciones de simbolización para detectar los retardos  $\tau$  y  $\tau_w$ . Las secciones 4.5 y 4.6 describen el comportamiento y optimización paramétrica respectivamente de las aplicaciones de simbolización. En la sección 4.7 se evalúa el comportamiento de las aplicaciones de simbolización junto con algunos métodos de reconstrucción para tres sistemas caóticos en condiciones ideales libres de ruido. La sección 4.8 describe los efectos del ruido en los distintos métodos de selección del retardo y ventana de inmersión y además se propone un test de selección de estos parámetros, el cual se describe en los apartados 4.8.2.1 y 4.8.2.2. En el apartado 4.8.3 se muestra el comportamiento estadístico en términos de tamaño y potencia del test propuesto en la selección del retardo y ventana temporales cuando se utilizan varios métodos de

selección. Finalmente en la sección 4.9 se exponen las conclusiones.

## 4.2. El Método C-C

El método C-C es una técnica de ventana  $\tau_w$  basada en la integral de correlación para la selección de los parámetros de reconstrucción  $\tau$  (retardo temporal) y  $p$  (dimensión de inmersión) en series temporales univariantes, la cual asume que ambos parámetros son interdependientes (ver [50]).

Sea  $p$  suficientemente grande, y el intervalo temporal  $[\tau_0, \tau_k]$  está dado de antemano, entonces uno puede encontrar un retardo apropiado  $\tau$  del intervalo  $[\tau_0, \tau_k]$  y una ventana de retardo temporal  $\tau_w = (p - 1)\tau$  a través del funcional  $S(p, T, r, \tau)$  y además poder estimar estima la dimensión de inmersión  $p$ .

El Método C-C está basado en las propiedades del funcional

$$S(p, T, r, \tau) = C(p, T, r, \tau) - C^p(1, T, r, 1) \quad (4.1)$$

del test de independencia  $BDS(p, T, r)$ . Entonces  $S(p, T, r, \tau)$  se puede interpretar como una medida adimensional de la dependencia serial de una serie temporal no lineal.

### 4.2.1. La Integral de Correlación y el Test BDS

Sea la serie temporal caótica  $\{X_t\}_{t \in I}$ ,  $X_p(t) = (X_t, X_{t+\tau}, \dots, X_{t+(p-1)\tau})$ ,  $t \in I$  donde  $I$  es el conjunto de índices temporales de cardinalidad  $T$  (número de observaciones de la serie),  $p$  es la dimensión de inmersión y  $\tau$  es el retardo temporal.

Así la integral de correlación para la serie embebida  $\{X_p(t)\}$  es:

$$C(p, T, r, \tau) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \theta(r - d_{ij}), \quad r > 0 \quad (4.2)$$

donde  $M$  es el conjunto de puntos insertados en el espacio  $p$ -dimensional,  $M = T - (p - 1)\tau$ ,  $\theta(x)$  es la función de Heaviside:  $\theta(x) = 0$ , si  $x < 0$ ;  $\theta(x) = 1$ , si  $x \geq 0$ , y  $d_{ij} = \|X_p^i - X_p^j\|$  denota la norma del supremo de los  $X_p$  vectores  $i, j$ .

La Integral de Correlación es una función de distribución acumulada, la cual indica la probabilidad de que dos puntos en el espacio de estados no sea más grande que  $r$ . La

distancia entre un par de puntos se denota como la norma suprema de la diferencia entre dos vectores. Si el límite de  $C(m, T, r, \tau)$  cuando  $N \rightarrow \infty$  existe para cada  $r$ , entonces la fracción de todos los puntos del espacio de estados que están dentro de un radio  $r$  dentro de cada punto se indica por  $C(p, r, \tau) = \lim_{N \rightarrow \infty} C(p, T, r, \tau)$  y la dimensión de correlación se define como  $D_2(p, \tau) = \lim_{r \rightarrow 0} [\log C(p, r, \tau) / \log r]$ . Ya que  $T$  es finito para conjuntos de datos reales, no se puede hacer tender  $r$  a 0; en lugar de eso se busca una región lineal para la pendiente  $D_2(p, \tau)$  en la gráfica de  $\log C(p, T, r, \tau)$  vs.  $\log r$ .

Brock et al. [20, 21] estudiaron el estadístico BDS, el cual está basado en la integral de correlación, con el fin de contrastar la hipótesis nula de la independencia e i.i.d de un conjunto de datos dado. El test ha sido particularmente útil para sistemas caóticos y sistemas estocásticos no lineales. Por completitud, revisamos brevemente el estadístico BDS.

Sea  $F$  la distribución invariante de una variable  $X$  en el espacio de estados y definamos la integral de correlación como sigue:

$$C(p, r) = \iint \theta(r - \|X - Y\|) dF(X) dF(Y), \quad r > 0 \quad (4.3)$$

Si  $X$  es iid, entonces usando  $\theta(r - \|X_p - Y_p\|) = \prod_{k=1}^p \theta(r - \|X_k - Y_k\|)$  lleva a  $C(p, r) = C^p(1, r)$ , donde

$$C(1, r) = \int [F(x+r) - F(x-r)] dF(x) \equiv C \quad (4.4)$$

Denker y Keller mostraron que  $C(p, T, r)$  es un estimador U-estadísticos. Brock et al [20, 21] usando la teoría estadística-U para un proceso absolutamente regular probaron que cuando  $T \rightarrow \infty$ ,  $\sqrt{N}[C(p, T, r) - C^p(1, r)]$  se aproxima a una distribución normal con media 0 y varianza

$$\sigma^2(p, r) = 4 \left[ K^p - C^{2p} + 2 \sum_{i=1}^{p-1} (K^{p-i} C^{2i} - C^{2p}) \right], \quad (4.5)$$

donde

$$K \equiv \int [F(x+r) - F(x-r)]^2 dF(x) \quad (4.6)$$

(asumiendo que  $K > C^2$ ), el estadístico BDS definido como

$$BDS(p, T, r) = \frac{\sqrt{T}}{\sigma(p, r)} [C(p, T, r) - C^p(1, r)] \quad (4.7)$$

se aproxima a una distribución normal. Sin embargo, ya que la distribución  $F$  es generalmente desconocida, no podemos obtener los valores de  $C$ ,  $K$  y la varianza  $\sigma^2(m, r)$  para las definiciones de arriba. En lugar de eso, la integral de correlación  $C(1, r)$  y la varianza  $\sigma^2(p, r)$  deben ser estimados a partir de los datos de las observaciones. Así  $C(1, r)$  es estimado a partir de  $C(1, T, r, \tau)$ , y  $\sigma^2(p, r)$  es estimado por

$$\hat{\sigma}^2 = 4 \left\{ p(p-1)\hat{C}^{2(p-1)}(\hat{K} - \hat{C}^2) + (\hat{K}^p - \hat{C}^{2p}) + 2 \sum_{i=1}^{p-1} [\hat{C}^{2i}(\hat{K}^{p-i} - \hat{C}^{2(p-i)}) - p\hat{C}^{2(p-i)}(\hat{K} - \hat{C}^2)] \right\} \quad (4.8)$$

donde  $\hat{C} = C(p, T, r, \tau)$  y

$$\hat{K} = \frac{6}{M(M-1)(M-2)} \sum_{1 \leq i < j < k \leq M} \theta(r - d_{ij})\theta(r - d_{jk}), \quad r > 0 \quad (4.9)$$

con  $M = T - (p-1)\tau$ . Entonces bajo la hipótesis de iid, si  $\hat{K} - \hat{C}^2$  y  $p > 1$ , el estadístico BDS toma la forma

$$BDS(p, T, r) = \frac{\sqrt{T}}{\hat{\sigma}} [C(p, T, r, \tau) - C^p(1, T, r, \tau)] \quad (4.10)$$

#### 4.2.2. El Método C-C

El algoritmo consiste en subdividir la serie  $\{X_t\}_{t \in I}$ , en  $\tau$  series temporales disjuntas, donde  $\tau$  es el retardo temporal,  $I$  es el conjunto de índices temporales y  $T$  es su cardinalidad ( $t=1, 2, \dots, T$ ). Así, para estimar  $S(p, T, r, \tau)$  se define el promedio del estadístico dado por la ecuación 4.1 como sigue:

$$S(p, T, r, \tau) = \frac{1}{\tau} \sum_{s=1}^{\tau} [C_s(p, T/\tau, r, \tau) - C_s^p(1, T/\tau, r, \tau)] \quad (4.11)$$

Mediante esta manipulación se consigue al mismo tiempo estudiar la dependencia no lineal y eliminar correlaciones temporales espúreas. En la ecuación 4.11,  $s$  es un índice que se refiere a la serie temporal disjunta sobre la que se aplica la integral de correlación  $C(p, T/\tau, r, \tau)$ . Es decir, que si tenemos  $k$  series temporales disjuntas en las que

dividimos la serie original de longitud  $T$  y  $s=1, 2, \dots, k$ ;  $C_s$  se refiere a la integral de correlación aplicada sobre la serie temporal  $s$  disjunta de longitud  $T/k$ . Esto lo podemos ver en la manera que se calcula  $S(p, T, r, t)$ :

Para  $\tau = 1$  se tiene una única serie temporal  $\{X_1, X_2, X_3, \dots, X_T\}$  y

$$S(p, T, r, 1) = C(p, T, r, 1) - C^p(1, T, r, 1) \quad (4.12)$$

Para  $\tau = 2$  se tienen dos series temporales disjuntas  $\{X_1, X_3, \dots, X_{T-1}\}$  y  $\{X_2, X_4, \dots, X_T\}$ , cada una de longitud  $T/2$ , así que se promedian los valores de  $S(p, T/2, r, 2)$  para las dos series:

$$S(p, T/2, r, 2) = \frac{1}{2} [C_1(p, T/2, r, 2) - C_1^p(1, T/2, r, 2)] + \frac{1}{2} [C_2(p, T/2, r, 2) - C_2^p(1, T/2, r, 2)] \quad (4.13)$$

Para un  $t$  general, la expresión se convierte en

$$S(p, T/2, r, \tau) = \frac{1}{\tau} \sum_{s=1}^{\tau} [C_s(p, N/\tau, r, \tau) - C_s^p(1, T/\tau, r, \tau)] \quad (4.14)$$

Y finalmente cuando  $N \rightarrow \infty$  se puede escribir

$$S(p, r, \tau) = \frac{1}{\tau} \sum_{s=1}^{\tau} [C_\tau(p, r, \tau) - C_s^p(1, r, \tau)], \quad p = 2, 3, \dots \quad (4.15)$$

Para  $p$  y  $\tau$  fijos  $S(p, r, \tau)$  será idénticamente nulo para todos los  $r$  si los datos son iid y  $T \rightarrow \infty$ . Sin embargo, los datos son reales y pueden ser serialmente dependientes, por lo que en general  $S(p, r, \tau) \neq 0$ . Así, los tiempos óptimos locales serán o los pasos por cero de  $S(p, r, \tau)$  o los puntos en los que  $S(p, r, \tau)$  sufre la menor variación con  $r$  ya que indica una distribución de puntos cercana a la uniforme. Es decir, la cantidad  $\frac{\partial S(p, r, \tau)}{\partial r}$  es mínima y  $S(p, r, \tau)$  posee un mínimo local en  $\tau$ . Por ello, se seleccionan ciertos valores representativos  $r_j$ , y se define la cantidad

$$\Delta S(p, \tau) = \max(S(p, r_j, \tau)) - \min(S(p, r_j, \tau)) \quad (4.16)$$

el cual es una medida de la variación de  $S_\tau(p, r_j, \tau)$  con  $r$ . Los tiempos locales óptimos  $\tau$  son los pasos por cero de  $S(p, r_j, \tau) \sim \tau$  y los mínimos de  $\Delta S(p, r_j, \tau) \sim \tau$ . Los pasos

por cero de  $S(p, r_j, \tau) \sim \tau$  deberían ser aproximadamente los mismos para todos los  $p$  y  $r$ , y los mínimos de  $S(p, r_j, \tau) \sim \tau$  deberían ser casi los mismos para todos los  $p$  (de otra forma, el retardo no sería localmente óptimo). El retardo temporal  $\tau$  corresponderá al primero de estos tiempos localmente óptimos.

Elecciones apropiadas de  $p$ ,  $T$  y  $r$  pueden encontrarse examinando el estadístico BDS. Brock et al[20] realizaron esta investigación usando series temporales generadas por varias distribuciones asintóticas. Los resultados concluyeron que  $p$  debería estar entre 2 y 5 y  $r$  debería estar entre  $\sigma/2$  y  $2\sigma$  donde  $\sigma = std(x)$  es la desviación estándar de la serie temporal. Además, las distribuciones asintóticas estaban bien aproximadas por series temporales cuando  $T \geq 500$ .

Después se parte de las ecuaciones 4.15 y 4.16 para definir los siguientes promedios (ecuaciones 4.17 y 4.18), escogiendo para ello cuatro  $r_i$  representativos en el rango  $\sigma/2 \leq r \leq 2\sigma$ ,  $r_1 = 0,5$ ,  $r_1 = 1$ ,  $r_1 = 1,5$  y  $r_1 = 2$

$$\bar{S}(\tau) = \frac{1}{16} \sum_{p=1}^4 \sum_{j=1}^4 S(p, r_j, \tau) \quad (4.17)$$

$$\Delta\bar{S}(\tau) = \frac{1}{4} \sum_{p=1}^4 \Delta S(p, \tau) \quad (4.18)$$

y se observa el primer paso por cero de  $\bar{S}(\tau)$  o el primer mínimo local de  $\Delta\bar{S}(\tau)$  para encontrar el primer tiempo óptimo local de independencia de los datos, el cual da el retardo temporal óptimo  $\tau$ . El tiempo óptimo es el retardo temporal  $\tau$  para el cual ambas magnitudes  $\bar{S}(\tau)$  y  $\Delta\bar{S}(\tau)$  son cercanas a 0. Si se asigna igual importancia a ambas cantidades, se puede mirar simplemente al mínimo de la cantidad

$$S_{cor}(\tau) = \Delta\bar{S}(\tau) + |\bar{S}(\tau)| \quad (4.19)$$

y el tiempo local de esta cantidad es el retardo de la ventana temporal  $\tau_w$ .

### 4.2.3. Desventajas del Método C-C

El método C-C posee algunas desventajas, algunas de las cuales se describen en [82] y que merece la pena mencionar:

1. Idealmente el retardo que hace mínimo  $S_{cor}(\tau)$  es el retardo óptimo de la ventana temporal, pero si por ejemplo, se evalúan diferentes muestras de 3000 puntos de diferentes intervalos, hay algunos puntos mínimos locales cuyos valores están mucho más cercanos al mínimo  $S_{cor}(\tau)$  y que perturban la estimación del mínimo de  $S_{cor}(\tau)$  y consecuentemente de la ventana temporal  $\tau_w$ .
2. En la práctica, el primer cruce por cero de  $\bar{S}(\tau)$  no es igual al primer mínimo local de  $\Delta\bar{S}(\tau)$ . Pero para las series temporales con período  $P_w$  para  $\tau = kP_w$ ,  $k = 1, 2, \dots$ , uno de los puntos es en su mayoría no solamente el primer paso por cero de  $\bar{S}(\tau)$ , sino también el mínimo de  $\Delta\bar{S}(\tau)$ , lo cual puede llevar a conclusiones paradójicas. Por ello, es mejor sugerir tomar solamente el primer mínimo local de  $\Delta\bar{S}(\tau)$  como el primer retardo local óptimo  $\tau$ .
3. Estrictamente hablando un sistema caótico no tiene período. Para sistemas caóticos de baja dimensión con período  $T$ , el período medio orbital  $P_w$  es el período medio generado por las oscilaciones del atractor caótico en las órbitas del espacio de estados. El modo de cálculo de 4.15 lleva al siguiente resultado: si  $\tau = kP_w$ ,  $k = 1, 2, \dots$ , entonces  $\Delta\bar{S}(\tau) = 0$ , y  $\Delta\bar{S}(\tau)$  muestra oscilaciones de alta frecuencia que aumentan con el aumento de  $\tau$ . Cuando el valor del retardo óptimo  $\tau$  es suficientemente grande, las oscilaciones de alta frecuencia pueden afectar incluso a la estimación del primer mínimo local de  $\Delta\bar{S}(\tau)$ .
4. La evolución de cualquier variable en un sistema dinámico está relacionada con otras variables, las cuales interaccionan entre sí. La integral de correlación definida por 4.2 calcula la norma del supremo; y no es apropiada para medir la relatividad de las distancias mediante el valor absoluto más grande de un vector *p-dimensional*.

#### 4.2.4. Mejoras del Método C-C

Ultimamente se han sugerido nuevos métodos que mejoran algunas de las desventajas del método C-C original mencionadas arriba:

1. En [83] se propone un nuevo método llamado C-C-1 que introduce una nueva cantidad  $S(\tau)$ , basada en la teoría de período  $P_w$ , para estimar el retardo de la ventana temporal  $\tau_w$  de una serie caótica temporal mediante el cálculo directo de la magnitud. Por lo tanto, el retardo óptimo de la ventana temporal  $\tau_w$  se toma como el primer período de la serie caótica temporal con un mínimo local en  $S(\tau)$ . Además, el retardo óptimo temporal  $\tau$  se obtiene a partir del primer mínimo local del promedio de la cantidad  $\Delta\bar{S}_2(\tau)$ . El promedio  $\Delta\bar{S}_2(\tau)$  se obtiene de la cantidad estadística  $S_2(p, r, \tau)$  de la misma manera que se llega a  $\Delta\bar{S}(\tau)$  a partir de  $S(p, r, \tau)$  y se define como

$$S_2(p, r, \tau) = \frac{1}{h} \sum_{s=1}^h \left[ C_s(p, r, \tau) - \left[ \frac{1}{p} C_s(1, r, \tau) \right]^p \right], \quad p = 2, 3, \dots \quad (4.20)$$

Donde  $h$  es un parámetro ajustable, entero positivo, para compensar la precisión y la velocidad de cálculo, independiente del retardo temporal  $\tau$ , que subdivide el espacio de fases reconstruido  $X = \{X(\tau)\}$ , siendo los  $X(\tau)$

$$\begin{cases} X(1) = \{X_1, X_{h+1}, \dots, X_{[T/h]-h+1}\} \\ X(2) = \{X_2, X_{h+1}, \dots, X_{[T/h]-h+2}\} \\ \dots \\ X(h) = \{X_h, X_{p+h}, \dots, X_{[T/h]}\} \end{cases} \quad (4.21)$$

$$\begin{cases} x(1) = \{X_1(1), X_1(2), \dots, X_1(p), X_{h+1}(1), X_{h+1}(2), \dots, X_{h+1}(p), \dots\} \\ x(2) = \{X_2(1), X_2(2), \dots, X_2(p), X_{h+2}(1), X_{h+2}(2), \dots, X_{h+2}(p), \dots\} \\ \dots \\ x(h) = \{X_h(1), X_h(2), \dots, X_h(p), X_{h+h}(1), X_{h+h}(2), \dots, X_{h+h}(p), \dots\} \end{cases} \quad (4.22)$$

A diferencia del método C-C, en este caso la dimensión de inmersión  $p$  se ajusta racionalmente. La robustez de este método alcanza el 40 % mientras que con el método C-C original, ésta sólo llega al 30 %.

2. En [82] se propone un nuevo algoritmo (método C-C-3) que propone modificar la métrica de la integral de correlación

$$C_3(p, T, r(p), \tau) = \frac{2}{M(M-1)} \sum_{1 \leq i \leq j < M} \theta(r(p) - d_{ij}), \quad r > 0 \quad (4.23)$$

donde  $r(p)$  es el radio de búsqueda,  $r(p) > 0$ ,  $r(p) = 1$  si  $p = 1$ ,  $r(p) = r^* \sqrt{(p+1)/2}$ , si  $p > 1$  y  $d_{ij} = \left\| X_p^i - X_p^j \right\|_{(2)}$  denota la norma cuadrática en lugar de la norma del máximo en 4.2 . Además se extiende el radio de búsqueda alargando el rango de variación de la serie temporal  $\sigma = std(x) * (1 + cv/3)$ , donde  $cv$  es el coeficiente de variación de la serie temporal caótica,  $cv = std(x)/mean(x)$ . Luego a partir de  $C_3(p, r(p), \tau)$  se define los estadístico  $S_3(p, r(p), \tau)$ ,  $\Delta \bar{S}_3(\tau)$  y  $S_{3cor}(p, T, r, \tau)$  definidas de la misma forma que en las ecuaciones 4.15, 4.18 y 4.19 respectivamente. La ventana temporal óptima  $\tau_w$  se estima buscando el primer mínimo local del estadístico  $S_{3cor}(p, T, r, \tau)$  basándose en la teoría de los sistemas caóticos de período T. El retardo óptimo temporal  $\tau$  se estima a partir del primer mínimo local del promedio  $\Delta \bar{S}_3(\tau)$ . Los resultados obtenidos son más estables (70 %) y apropiados que los obtenidos para del método C-C(30 %).

3. En [54] se propone un nuevo algoritmo para seleccionar la ventana de inmersión. Para ello, en el nuevo algoritmo se define la Integral de Correlación media:

$$C_l(p, r, \tau) = \frac{1}{\tau} \sum_{s=1}^{\tau} C_s(p, r, \tau) \quad (4.24)$$

La dispersión entre el estadístico de la Integral de Correlación Media  $\bar{C}(\tau)$  y la variable estadística promedio  $\bar{S}(\tau)$  se usa en un nuevo indicador llamado  $S_c$  para

eliminar las fluctuaciones y los picos locales.

$$\begin{aligned}\overline{C}(\tau) &= \frac{1}{16} \sum_{p=2}^5 \sum_{j=1}^4 C_l(p, r_j, \tau) \\ \overline{S}_c(\tau) &= \overline{C}(\tau) - \overline{S}(\tau)\end{aligned}\tag{4.25}$$

La ventana de inmersión  $\tau_w$  se estima basándose en la pseudo-periodicidad  $P_w$  de la Integral de Correlación Media y que es igual a la pseudo-periodicidad  $P_w$  del sistema caótico. Esta se localiza en el primero de los máximos locales periódicos de la cantidad  $\overline{S}_c(\tau)$ . Los resultados de las simulaciones verifican que el nuevo algoritmo (que nosotros llamaremos C-C- $S_c$ ) es superior a otras alternativas para el cálculo de la ventana de inmersión y usado junto con el algoritmo C-C para la localizar el retardo temporal  $\tau$ , determinan conjuntamente de forma óptima los parámetros de reconstrucción del espacio de estados.

### 4.3. Aplicación de Simbolización basada en las distancias

#### absolutas $f_{r_e}$

Como alternativa a los métodos de reconstrucción basados en el C-C, se propone utilizar la *Dinámica Simbólica*. Sea de la serie temporal de valores reales  $\{X_t\}_{t \in I}$ , siendo  $I$  el conjunto de índices temporales y  $T$  su cardinalidad o número de observaciones de la serie temporal. Ahora definimos la siguiente aplicación de simbolización  $f_{r_e}$  de  $\mathbb{R}^m$  en  $\Gamma$ , siendo  $\Gamma = \{\pi_1, \pi_2, \dots, \pi_n\}$  el conjunto de  $n$  símbolos que genera la aplicación  $f_{r_e}$  <sup>1</sup>:

Expresemos como  $r_e$  un radio de la serie temporal  $\{X_t\}_{t \in I}$  y sea

$$J_t = \begin{cases} 0 & X_t \leq r_e \\ 1 & \text{en otro caso} \end{cases}$$

<sup>1</sup>En este capítulo el parámetro  $p$ , dimensión de inmersión para la reconstrucción del atractor de un sistema caótico, tiene una utilidad diferente de la dimensión de inmersión  $m$  en el contexto de la *Dinámica Simbólica*. La dimensión de inmersión  $m$ , como ya se describió en el capítulo uno, es el espacio  $m$  – *dimensional* en que se embebe la serie temporal  $\{X_t\}_{t \in I}$ . Una vez obtenida la serie de  $m$ -*historias* resultantes  $\{X_m(t)\}$  se le aplica la simbolización  $f_{r_e} : \mathbb{R}^m \rightarrow \Gamma_m$ , siendo  $\Gamma_m$  el conjunto de elementos de orden  $n$  y dimensión  $m$ .

Ahora definimos la función indicador

$$\delta_{tt+k} = \begin{cases} 0 & \text{si } J_t \neq J_{t+k} \\ 1 & \text{en otro caso} \end{cases}$$

Por lo que la aplicación de simbolización así descrita  $f_{r_e} : \mathbb{R}^m \rightarrow \Gamma$  se define como:

$$\pi_i(t) = f_{r_e}(X_m(t)) = f_{r_e}(X_t, X_{t+\tau}, \dots, X_{t+(m-1)\tau}) = (\delta_{tt+\tau}, \delta_{tt+2\tau}, \dots, \delta_{tt+(m-1)\tau}), \quad (4.26)$$

Si para cualquier  $t$ , expresamos  $X_m(t) = (X_t, X_{t+\tau}, \dots, X_{t+(m-1)\tau})$ , e indicamos como  $N_t = \{t + \tau, t + 2\tau, \dots, t + (m-1)\tau\}$  las  $m-1$  observaciones adyacentes de la  $m$ -historia de  $t$ , entonces este procedimiento de simbolización consiste en comparar para cada tiempo  $t$ , el valor de  $X_t$  con  $X_{t+k}$  (con  $k = t + \tau, t + 2\tau, \dots, t + (m-1)\tau$ , para todos los  $t \in N_t$ . Por lo tanto si  $J_t = J_{t+k}$  significa que  $X_t$  y  $X_{t+k}$  son menores que, o mayores que la magnitud  $r_e$  y así  $f_{r_e}(X_m(t)) = \pi_i$  tendrá el valor 1 en la entrada  $k$ .

Como podemos ver, esta Aplicación de Simbolización no es sino una extensión de la descrita en el apartado 2.7, donde  $r_e = \text{Mediana}(x)$ . Además, al igual que en 2.7, el conjunto de símbolos generado por la Aplicación de Simbolización tiene una cardinalidad de  $n = 2^{m-1}$ .

La *Entropía  $\tau$ -Simbólica* asociada, presentada en [62] se define como

$$h_\tau(\tau) = - \sum_{\pi \in \Gamma} p_{\pi_\tau} \ln(p_{\pi_\tau}) \quad (4.27)$$

donde la frecuencia relativa de un símbolo  $\pi_\tau \in \Gamma_m$  viene dada por

$$p(\pi, \tau) = p_{\pi_\tau} = \frac{\{t \in T \mid t \text{ es de tipo } \pi\}}{T - (m-1)\tau} \quad (4.28)$$

Si  $\tau$  es un retardo óptimo de reconstrucción, entonces la dependencia entre  $X_t$  y  $X_{t+\tau}, X_{t+2\tau}, \dots, X_{t+k\tau}$  es mínima, con  $k$  un número entero mayor que 0. Por lo tanto, con la aplicación de simbolización propuesta, para una dimensión  $m$  dada, una primera tentativa de selección del retardo  $\tau$  óptimo es escoger cualquiera de los retardos  $\tau$  que se corresponden con máximos locales de la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$ , pues para estos

retardos la dependencia entre  $X_t$  y  $X_{t+\tau}, X_{t+2\tau}, \dots, X_{t+k\tau}$  es mínima (máxima entropía  $\rightarrow$  mínima información). En nuestro caso, con el fin de evitar la amplificación del ruido debido a la elección de un retardo  $\tau$  demasiado bajo (i.e efecto *redundancia*) pero al mismo tiempo evitar el efecto *irrelevancia* (cuando se elige un  $\tau$  es demasiado grande), esto es, cuando se corre el riesgo de perder información relevante; seleccionamos el *primer máximo local de la Entropía Simbólica*  $h_\tau(m)$ .

Con respecto al retardo de la ventana temporal  $\tau_w = (p - 1)\tau$ , ésta se puede asociar con el período medio orbital  $P_w$  de sistemas caóticos de baja dimensión que muestran pseudo-periodicidad. Todos los puntos  $X_{t+kP_w}$  correspondientes al período medio orbital  $\tau = P_w$ , donde  $k$  es un entero más grande que cero, están en la misma sección de Poincaré del espacio de estados para un  $p$  y  $\tau$  fijos,  $T \rightarrow \infty$ . Así, los  $X_{t+kP_w}$  están en la misma sección de Poincaré y se encuentran muy próximos entre sí, por lo que el retardo  $\tau = P_w$  se pueden considerar como el tiempo de dependencia de los datos. Esta característica se corresponde con un valor mínimo local de la Entropía Simbólica  $h_\tau(m)$  en los retardos  $\tau = kP_w$ . Con esta aplicación de simbolización, la mayoría de los puntos  $X_{t+kP_w}$ , que se encuentran muy próximos entre sí en la misma sección de Poincaré, son casi todos mayores o menores que  $r_e$  y los  $\delta_{tt+kP_w}$  tienen casi todos el mismo valor. Para estos retardos  $\tau = kP_w$ , existen unos pocos símbolos  $\pi_i$  cuyas probabilidades  $p(\pi_i, kP_w) = p_{kP_w}(f_{r_e}(X_m(t))) = \pi_i$  son mucho mayores que el resto, dando lugar a que la Entropía Simbólica  $h_\tau(m)$  sea mínima en esos retardos. Así estimaremos el retardo de óptimo de la ventana temporal  $\tau_w$  con el primer mínimo relativo local en  $\tau = P_w$  de los mínimos relativos locales periódicos en  $\tau = kP_w$  de la *Entropía  $\tau$ -Simbólica*  $h_\tau(m)$ .

Ahora veamos como opera la aplicación de simbolización con el siguiente ejemplo:

Tomemos como dimensión de inmersión  $m = 3$ . Entonces el conjunto de símbolos  $\Gamma_m$  es

$$\Gamma_3 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

Tabla 4.1:  $m$ -historias y Símbolos para  $\tau = 1$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
1	$x_3^1(1)$	$(x_1 = 3, x_2 = 1, x_3 = 11)$	$(1, 0)$
1	$x_3^1(2)$	$(x_2 = 1, x_3 = 11, x_4 = 15)$	$(0, 0)$
1	$x_3^1(3)$	$(x_3 = 11, x_4 = 15, x_5 = 3)$	$(1, 0)$
1	$x_3^1(4)$	$(x_4 = 15, x_5 = 13, x_6 = 6)$	$(0, 0)$
1	$x_3^1(5)$	$(x_5 = 3, x_6 = 6, x_7 = 18)$	$(1, 0)$
1	$x_3^1(6)$	$(x_6 = 6, x_7 = 18, x_8 = 12)$	$(0, 0)$
1	$x_3^1(7)$	$(x_7 = 18, x_8 = 12, x_9 = 11)$	$(1, 0)$
1	$x_3^1(8)$	$(x_8 = 12, x_9 = 5, x_{10} = 4)$	$(0, 0)$
1	$x_3^1(9)$	$(x_9 = 5, x_{10} = 4, x_{11} = 13)$	$(1, 0)$
1	$x_3^1(10)$	$(x_{10} = 4, x_{11} = 13, x_{12} = 12)$	$(0, 0)$
1	$x_3^1(11)$	$(x_{11} = 13, x_{12} = 12, x_{13} = 8)$	$(1, 0)$
1	$x_3^1(12)$	$(x_{12} = 12, x_{13} = 8, x_{14} = 2)$	$(0, 0)$

Supongamos ahora que tenemos una serie temporal :

$$x_t = \{x_1 = 3, x_2 = 1, x_3 = 11, x_4 = 15, x_5 = 3, x_6 = 6, x_7 = 18, x_8 = 12, \quad (4.29)$$

$$x_9 = 5, x_{10} = 4, x_{11} = 13, x_{12} = 12, x_{13} = 8, x_{14} = 2\}$$

Ahora elijamos un valor de  $r_e = 10$ . Entonces para  $t = 2$  y  $\tau = 1$  tenemos que  $x_t = 1 < 10$ ;  $x_{t+1} = 11 > 10$ ;  $x_{t+2} = 15 \geq 10$  y  $i_t = 0$ ;  $i_{t+1} = 1$ ;  $i_{t+2} = 1$  con lo que  $\delta_{(t)(t+1)} = 0$ ;  $\delta_{(t)(t+2)} = 0$  y por lo tanto tenemos que para el período  $t = 2$  es de tipo-(0,0).

En las Tablas 4.1, 4.3, 4.5, 4.7, 4.9 y 4.11 mostramos las 3-historias de la serie y los símbolos que las representan para los retardos  $\tau = 1, 2, 3, 4, 5$  y 6 respectivamente. De esta manera obtenemos las probabilidades de símbolos asociadas a cada retardo que se representan respectivamente en las Tablas 4.2, 4.4, 4.6, 4.8, 4.10 y 4.12. Así pues, las entropías de símbolo asociadas a las probabilidades calculadas para cada retardo son:  $h_{\tau=1}(3) = 9\frac{1}{27}\ln(\frac{1}{27}) - 3\frac{2}{27}\ln(\frac{2}{27}) - \frac{3}{27}\ln(\frac{3}{27}) - \frac{4}{27}\ln(\frac{4}{27}) - \frac{5}{27}\ln(\frac{5}{27}) \approx 2.51$ ,  $h_{\tau=2}(3) = -\frac{10}{10}\ln(\frac{10}{10}) = 0$ ,  $h_{\tau=3}(3) = -\frac{4}{8}\ln(\frac{4}{8}) - \frac{1}{8}\ln(\frac{1}{8}) - \frac{3}{8}\ln(\frac{3}{8}) \approx 0.97$ ,  $h_{\tau=4}(3) = -\frac{5}{5}\ln(\frac{5}{5}) = 0$ ,  $h_{\tau=5}(3) = -\frac{2}{4}\ln(\frac{2}{4}) - \frac{2}{4}\ln(\frac{2}{4}) \approx 0.69$ ,  $h_{\tau=6}(3) = 0$ . Esto es, la función Entropía  $\tau$ -simbólica toma los valores  $h_\tau(3) = \{0.69, 0, 0.97, 0, 0.69, 0\}$ .

Tabla 4.2: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 1$ 

$\tau$	Símbolo	Probabilidad de Símbolo
1	(1, 0)	$\frac{6}{12}$
1	(0, 0)	$\frac{6}{12}$
1	(0, 1)	0
1	(1, 1)	0

Tabla 4.3:  $m$ -historias y Símbolos para  $\tau = 2$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
2	$x_3^1(1)$	$(x_1 = 3, x_2 = 11, x_3 = 13)$	(0, 1)
2	$x_3^1(2)$	$(x_2 = 1, x_3 = 15, x_4 = 6)$	(0, 1)
2	$x_3^1(3)$	$(x_3 = 11, x_4 = 3, x_5 = 8)$	(0, 1)
2	$x_3^1(4)$	$(x_4 = 15, x_5 = 6, x_6 = 12)$	(0, 1)
2	$x_3^1(5)$	$(x_5 = 3, x_6 = 18, x_7 = 15)$	(0, 1)
2	$x_3^1(6)$	$(x_6 = 6, x_7 = 12, x_8 = 4)$	(0, 1)
2	$x_3^1(7)$	$(x_7 = 18, x_8 = 5, x_9 = 13)$	(0, 1)
2	$x_3^1(8)$	$(x_8 = 12, x_9 = 4, x_{10} = 12)$	(0, 0)
2	$x_3^1(9)$	$(x_9 = 5, x_{10} = 13, x_{11} = 8)$	(1, 0)
2	$x_3^1(10)$	$(x_{10} = 4, x_{11} = 12, x_{12} = 2)$	(0, 0)

Tabla 4.4: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 2$ 

$\tau$	Símbolo	Probabilidad de Símbolo
2	(1, 0)	0
2	(0, 0)	0
2	(0, 1)	1
2	(1, 1)	0

Tabla 4.5:  $m$ -historias y Símbolos para  $\tau = 3$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
3	$x_3^1(1)$	$(x_1 = 3, x_2 = 15, x_3 = 18)$	(0, 0)
3	$x_3^1(2)$	$(x_2 = 1, x_3 = 3, x_4 = 12)$	(0, 1)
3	$x_3^1(3)$	$(x_3 = 11, x_4 = 6, x_5 = 5)$	(0, 0)
3	$x_3^1(4)$	$(x_4 = 15, x_5 = 18, x_6 = 4)$	(1, 0)
3	$x_3^1(5)$	$(x_5 = 3, x_6 = 12, x_7 = 13)$	(0, 0)
3	$x_3^1(6)$	$(x_6 = 6, x_7 = 5, x_8 = 12)$	(1, 0)
3	$x_3^1(7)$	$(x_7 = 18, x_8 = 4, x_9 = 8)$	(0, 0)
3	$x_3^1(8)$	$(x_8 = 12, x_9 = 13, x_{10} = 2)$	(1, 0)

Tabla 4.6: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 3$ 

$\tau$	Símbolo	Probabilidad de Símbolo
3	(1, 0)	$\frac{3}{8}$
3	(0, 0)	$\frac{4}{8}$
3	(0, 1)	$\frac{1}{8}$
3	(1, 1)	0

Tabla 4.7: *m*-historias y Símbolos para  $\tau = 4$  de 4.29

$\tau$	nombre <i>m</i> -Historia	<i>m</i> – Historia	Símbolo
4	$x_3^1(1)$	$(x_1 = 3, x_2 = 3, x_3 = 5)$	(1, 1)
4	$x_3^1(2)$	$(x_2 = 1, x_3 = 6, x_4 = 4)$	(1, 1)
4	$x_3^1(3)$	$(x_3 = 11, x_4 = 18, x_5 = 13)$	(1, 1)
4	$x_3^1(4)$	$(x_4 = 15, x_5 = 12, x_6 = 12)$	(1, 1)
4	$x_3^1(5)$	$(x_5 = 3, x_6 = 5, x_7 = 8)$	(1, 1)

Tabla 4.8: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 4$ 

$\tau$	Símbolo	Probabilidad de Símbolo
4	(1, 0)	0
4	(0, 0)	0
4	(0, 1)	0
4	(1, 1)	1

Tabla 4.9: *m*-historias y Símbolos para  $\tau = 5$  de 4.29

$\tau$	nombre <i>m</i> -Historia	<i>m</i> – Historia	Símbolo
5	$x_3^1(1)$	$(x_1 = 3, x_2 = 6, x_3 = 13)$	(1, 0)
5	$x_3^1(2)$	$(x_2 = 1, x_3 = 18, x_4 = 12)$	(0, 0)
5	$x_3^1(3)$	$(x_3 = 11, x_4 = 12, x_5 = 8)$	(1, 0)
5	$x_3^1(4)$	$(x_4 = 15, x_5 = 5, x_6 = 2)$	(0, 0)

Tabla 4.10: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 5$ 

$\tau$	Símbolo	Probabilidad de Símbolo
5	(1, 0)	$\frac{2}{4}$
5	(0, 0)	$\frac{4}{4}$
5	(0, 1)	0
5	(1, 1)	0

Tabla 4.11: *m*-historias y Símbolos para  $\tau = 6$  de 4.29

$\tau$	nombre <i>m</i> -Historia	<i>m</i> – Historia	Símbolo
6	$x_3^1(1)$	$(x_1 = 3, x_2 = 18, x_3 = 8)$	(0, 1)
6	$x_3^1(2)$	$(x_2 = 1, x_3 = 12, x_4 = 2)$	(0, 1)

Tabla 4.12: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 6$ 

$\tau$	Símbolo	Probabilidad de Símbolo
6	(1, 0)	0
6	(0, 0)	0
6	(0, 1)	1
6	(1, 1)	0

#### 4.4. Aplicación de Simbolización basada en las distancias

relativas  $f_{r_e}^1$

La segunda Aplicación de Simbolización  $f_{r_e}^1$  que proponemos en esta sección es una variante de la descrita en el apartado anterior  $f_{r_e}$ . La aplicación de simbolización  $f_{r_e}^1$  se define de  $\mathbb{R}^m$  en  $\Gamma$ , siendo  $\Gamma = \{\pi_1, \pi_2, \dots, \pi_n\}$  un conjunto de  $n$  símbolos :

Sea la serie temporal  $\{X_t\}_{t \in I}$  y  $r_e$  el parámetro para medir la distancia entre los elementos de la serie entonces

$$\delta_{tt+k} = \begin{cases} 0 & \text{si } |X_t - X_{t+k}| \leq r_e \\ 1 & \text{en otro caso} \end{cases}$$

Por lo que la aplicación de simbolización así descrita  $f_{r_e}^1 : \mathbb{R}^m \rightarrow \Gamma$  se define como:

$$\pi_i(t) = f_{r_e}^1(X_m(t)) = f_{r_e}^1(X_t, X_{t+\tau}, \dots, X_{t+(m-1)\tau}) = (\delta_{tt+\tau}, \delta_{tt+2\tau}, \dots, \delta_{tt+(m-1)\tau}) \quad (4.30)$$

Si para cualquier  $t$ , expresamos  $X_m(t) = (X_t, X_{t+\tau}, \dots, X_{t+(m-1)\tau})$ , e indicamos como  $N_t = \{t + \tau, t + 2\tau, \dots, t + (m-1)\tau\}$  las  $m-1$  observaciones adyacentes de la  $m$ -historia de  $t$ , entonces este procedimiento de simbolización consiste en comparar para cada tiempo  $t$ , el valor de  $X_t$  con  $X_{t+k}$  (con  $k = t + \tau, t + 2\tau, \dots, t + (m-1)\tau$ ), para todos los  $t \in N_t$ . Entonces si  $|X_t - X_{t+k}| \leq r_e$  significa que  $X_t$  y  $X_{t+k}$  son menores que la magnitud  $r_e$  y así  $f_{r_e}^1(X_m(t)) = \pi_i$  tendrá el valor 0 en la entrada  $k$ . Si por el contrario,  $|X_t - X_{t+k}| > r_e$ ,  $f_{r_e}^1(X_m(t)) = \pi_i$  tendrá el valor 1 en la entrada  $k$ . Como podemos ver, el conjunto de símbolos generado por la aplicación de simbolización  $f_{r_e}^1$  es, al igual que  $f_{r_e}^1$  de  $n = 2^{m-1}$ .

Como hemos dicho anteriormente, cuando  $\tau$  es un retardo óptimo de reconstrucción, existe una dependencia mínima entre  $X_t$  y  $X_{t+\tau}, X_{t+2\tau}, \dots, X_{t+k\tau}$ . Con la aplicación de simbolización propuesta, para una dimensión de inmersión  $m$ , al igual que para la aplicación  $f_{r_e}$  una primera tentativa de selección del retardo  $\tau$  óptimo es escoger cualquiera de los retardos  $\tau$  que se corresponden con máximos locales de la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$ , pues para estos retardos la dependencia entre  $X_t$  y  $X_{t+\tau}, X_{t+2\tau}, \dots, X_{t+k\tau}$  es mínima. Con el fin de evitar la amplificación de ruido debido a los efectos de *redundancia e*

*irrelevancia* seleccionamos el *primer máximo local de la Entropía Simbólica*  $h_\tau(m)$ . Así, los tiempos óptimos locales serán aquellos que correspondan a máximos relativos de la *Entropía  $\tau$ -simbólica* y el retardo temporal óptimo  $\tau_d$  se corresponderá a cualquiera de estos máximos locales, sujetos a la condición  $\tau_d \leq T_p$  (siendo  $T_p$  el pseudo-período del sistema caótico). En este caso, igualmente que en el apartado 4.3, proponemos elegir el primero de los máximos locales.

Con respecto al retardo de la ventana temporal  $\tau_w = (p - 1)\tau$ , como hemos dicho anteriormente, ésta se asocia con el período medio orbital  $P_w$  de sistemas caóticos de baja dimensión con pseudo-periodicidad. También hemos comentado que los puntos  $X_{t+kP_w}$ , con  $k > 0$ , correspondientes al período medio orbital  $\tau = P_w$  están muy próximos entre sí en la misma sección de Poincaré del espacio de estados para un  $p$  y  $\tau$  fijos,  $T \rightarrow \infty$ . En la sección de Poincaré donde se encuentran los  $X_{t+kP_w}$ , éstos están muy próximos entre sí y distribuidos más o menos uniformemente a diferencia de los  $X_{t+k\tau \neq t+kP_w}$ . Esta característica de la distribución de los  $X_{t+kP_w}$ , se corresponde con un valor máximo local de la Entropía Simbólica  $h_\tau(m)$  en los retardos  $\tau = kP_w$ , ya que los puntos  $X_{t+k\tau}$ , con  $\tau \neq kP_w$  se encuentran fuera de la sección de Poincaré y están la mayoría fuera del radio  $r_e$ , por lo que los  $\delta_{tt+kP_w}$  tienen el valor 1 en esos puntos. En los retardos  $\tau \neq kP_w$ , el símbolo del tipo  $\pi_i = (1 \underbrace{.,1^{m-1}}_., 1)$  posee una frecuencia mayor que para el resto de símbolos. La Entropía Simbólica  $h_{\tau \neq kP_w}(m)$  es menor por consiguiente que la Entropía Simbólica  $h_{\tau = kP_w}(m)$ . Así pues estimaremos el retardo de óptimo de la ventana temporal  $\tau_w$  con el primer máximo relativo local en  $\tau = P_w$  de los máximos relativos locales periódicos en  $\tau = kP_w$  de la *Entropía  $\tau$ -Simbólica*  $h_\tau(m)$ .

Veamos como opera la aplicación de simbolización con el siguiente ejemplo:

Tomemos como dimensión de inmersión  $m = 3$ . Entonces el conjunto de símbolos  $\Gamma_m$  es

$$\Gamma_3 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

Tabla 4.13:  $m$ -historias y Símbolos para  $\tau = 1$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
1	$x_3^1(1)$	$(x_1 = 3, x_2 = 1, x_3 = 11)$	$(0, 0)$
1	$x_3^1(2)$	$(x_2 = 1, x_3 = 11, x_4 = 15)$	$(1, 1)$
1	$x_3^1(3)$	$(x_3 = 11, x_4 = 15, x_5 = 3)$	$(0, 0)$
1	$x_3^1(4)$	$(x_4 = 15, x_5 = 13, x_6 = 6)$	$(1, 0)$
1	$x_3^1(5)$	$(x_5 = 3, x_6 = 6, x_7 = 18)$	$(0, 1)$
1	$x_3^1(6)$	$(x_6 = 6, x_7 = 18, x_8 = 12)$	$(1, 0)$
1	$x_3^1(7)$	$(x_7 = 18, x_8 = 12, x_9 = 11)$	$(0, 1)$
1	$x_3^1(8)$	$(x_8 = 12, x_9 = 5, x_{10} = 4)$	$(0, 0)$
1	$x_3^1(9)$	$(x_9 = 5, x_{10} = 4, x_{11} = 13)$	$(0, 0)$
1	$x_3^1(10)$	$(x_{10} = 4, x_{11} = 13, x_{12} = 12)$	$(0, 0)$
1	$x_3^1(11)$	$(x_{11} = 13, x_{12} = 12, x_{13} = 8)$	$(0, 0)$
1	$x_3^1(12)$	$(x_{12} = 12, x_{13} = 8, x_{14} = 2)$	$(0, 0)$

Supongamos ahora que tenemos la serie temporal del ejemplo 4.29 :

$$x_t = \{x_1 = 3, x_2 = 1, x_3 = 11, x_4 = 15, x_5 = 3, x_6 = 6, x_7 = 18, x_8 = 12, \\ x_9 = 5, x_{10} = 4, x_{11} = 13, x_{12} = 12, x_{13} = 8, x_{14} = 2\}$$

Ahora consideremos un  $r_e = 9$ . Entonces para  $t = 2$  y  $\tau = 1$  tenemos que  $|x_t - x_{t+1}| = |1 - 11| > 9$ ;  $|x_t - x_{t+2}| = |3 - 15| > 9$  con lo que  $\delta_{(t)(t+1)} = 1$ ;  $\delta_{(t)(t+2)} = 1$  y por lo tanto tenemos que para el período  $t = 2$  es de tipo-(1,1).

En las Tablas 4.13, 4.15, 4.17, 4.19, 4.21 y 4.23 mostramos las 3-historias de la serie y los símbolos que las representan para los retardos  $\tau = 1, 2, 3, 4, 5$  y 6 respectivamente. De esta manera obtenemos las probabilidades de símbolos asociadas a cada retardo que se representan respectivamente en las Tablas 4.14, 4.16, 4.18, 4.20, 4.22 y 4.24. Así pues, las entropías de símbolo asociadas a las probabilidades calculadas para cada retardo son:  $h_{\tau=1}(3) = -\frac{6}{12}\ln(\frac{6}{12}) - \frac{6}{12}\ln(\frac{6}{12}) \approx 1,11$ ,  $h_{\tau=2}(3) = -\frac{7}{10}\ln(\frac{7}{10}) - \frac{3}{10}\ln(\frac{3}{10}) \approx 0,36$ ,  $h_{\tau=3}(3) = -\frac{2}{8}\ln(\frac{2}{8}) - \frac{4}{8}\ln(\frac{4}{8}) - \frac{2}{8}\ln(\frac{2}{8}) \approx 1,04$ ,  $h_{\tau=4}(3) = 0$ ,  $h_{\tau=5}(3) = -\frac{1}{4}\ln(\frac{1}{4}) - \frac{1}{4}\ln(\frac{1}{4}) - \frac{2}{4}\ln(\frac{2}{4}) \approx 1,03$ ,  $h_{\tau=6}(3) = 0$ . Esto es, la función *Entropía  $\tau$ -simbólica* toma los valores  $h_{\tau}(3) = \{1,11, 0,36, 1,04, 0, 1,03, 0\}$ .

Tabla 4.14: *Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 1$* 

$\tau$	Símbolo	Probabilidad de Símbolo
1	(1, 0)	$\frac{2}{12}$
1	(0, 0)	$\frac{7}{12}$
1	(0, 1)	$\frac{2}{12}$
1	(1, 1)	$\frac{1}{12}$

Tabla 4.15: *m-historias y Símbolos para  $\tau = 2$  de 4.29*

$\tau$	nombre <i>m</i> -Historia	<i>m</i> - Historia	Símbolo
2	$x_3^1(1)$	$(x_1 = 3, x_2 = 11, x_3 = 13)$	(0, 0)
2	$x_3^1(2)$	$(x_2 = 1, x_3 = 15, x_4 = 6)$	(1, 0)
2	$x_3^1(3)$	$(x_3 = 11, x_4 = 3, x_5 = 8)$	(0, 0)
2	$x_3^1(4)$	$(x_4 = 15, x_5 = 6, x_6 = 12)$	(0, 0)
2	$x_3^1(5)$	$(x_5 = 3, x_6 = 18, x_7 = 15)$	(1, 0)
2	$x_3^1(6)$	$(x_6 = 6, x_7 = 12, x_8 = 4)$	(0, 0)
2	$x_3^1(7)$	$(x_7 = 18, x_8 = 5, x_9 = 13)$	(1, 0)
2	$x_3^1(8)$	$(x_8 = 12, x_9 = 4, x_{10} = 12)$	(0, 0)
2	$x_3^1(9)$	$(x_9 = 5, x_{10} = 13, x_{11} = 8)$	(0, 0)
2	$x_3^1(10)$	$(x_{10} = 4, x_{11} = 12, x_{12} = 2)$	(0, 0)

Tabla 4.16: *Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 2$* 

$\tau$	Símbolo	Probabilidad de Símbolo
2	(1, 0)	0
2	(0, 0)	$\frac{7}{10}$
2	(0, 1)	$\frac{3}{10}$
2	(1, 1)	0

Tabla 4.17: *m-historias y Símbolos para  $\tau = 3$  de 4.29*

$\tau$	nombre <i>m</i> -Historia	<i>m</i> - Historia	Símbolo
3	$x_3^1(1)$	$(x_1 = 3, x_2 = 15, x_3 = 18)$	(1, 1)
3	$x_3^1(2)$	$(x_2 = 1, x_3 = 3, x_4 = 12)$	(0, 1)
3	$x_3^1(3)$	$(x_3 = 11, x_4 = 6, x_5 = 5)$	(0, 0)
3	$x_3^1(4)$	$(x_4 = 15, x_5 = 18, x_6 = 4)$	(0, 1)
3	$x_3^1(5)$	$(x_5 = 3, x_6 = 12, x_7 = 13)$	(0, 1)
3	$x_3^1(6)$	$(x_6 = 6, x_7 = 5, x_8 = 12)$	(0, 0)
3	$x_3^1(7)$	$(x_7 = 18, x_8 = 4, x_9 = 8)$	(1, 1)
3	$x_3^1(8)$	$(x_8 = 12, x_9 = 13, x_{10} = 2)$	(0, 1)

Tabla 4.18: *Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 3$* 

$\tau$	Símbolo	Probabilidad de Símbolo
3	(1, 0)	0
3	(0, 0)	$\frac{2}{8}$
3	(0, 1)	$\frac{4}{8}$
3	(1, 1)	$\frac{2}{8}$

Tabla 4.19:  $m$ -historias y Símbolos para  $\tau = 4$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
4	$x_3^1(1)$	$(x_1 = 3, x_2 = 3, x_3 = 5)$	$(0, 0)$
4	$x_3^1(2)$	$(x_2 = 1, x_3 = 6, x_4 = 4)$	$(0, 0)$
4	$x_3^1(3)$	$(x_3 = 11, x_4 = 18, x_5 = 13)$	$(0, 0)$
4	$x_3^1(4)$	$(x_4 = 15, x_5 = 12, x_6 = 12)$	$(0, 0)$
4	$x_3^1(5)$	$(x_5 = 3, x_6 = 5, x_7 = 8)$	$(0, 0)$

Tabla 4.20: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 4$ 

$\tau$	Símbolo	Probabilidad de Símbolo
4	$(1, 0)$	0
4	$(0, 0)$	0
4	$(0, 1)$	0
4	$(1, 1)$	1

Tabla 4.21:  $m$ -historias y Símbolos para  $\tau = 5$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
5	$x_3^1(1)$	$(x_1 = 3, x_2 = 6, x_3 = 13)$	$(0, 1)$
5	$x_3^1(2)$	$(x_2 = 1, x_3 = 18, x_4 = 12)$	$(1, 1)$
5	$x_3^1(3)$	$(x_3 = 11, x_4 = 12, x_5 = 8)$	$(0, 0)$
5	$x_3^1(4)$	$(x_4 = 15, x_5 = 5, x_6 = 2)$	$(1, 1)$

Tabla 4.22: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 5$ 

$\tau$	Símbolo	Probabilidad de Símbolo
5	$(1, 0)$	0
5	$(0, 0)$	$\frac{1}{4}$
5	$(0, 1)$	$\frac{1}{4}$
5	$(1, 1)$	$\frac{2}{4}$

Tabla 4.23:  $m$ -historias y Símbolos para  $\tau = 6$  de 4.29

$\tau$	nombre $m$ -Historia	$m$ - Historia	Símbolo
6	$x_3^1(1)$	$(x_1 = 3, x_2 = 18, x_3 = 8)$	$(1, 0)$
6	$x_3^1(2)$	$(x_2 = 1, x_3 = 12, x_4 = 2)$	$(1, 0)$

Tabla 4.24: Símbolos y Probabilidades de Símbolo de 4.29 para  $\tau = 6$ 

$\tau$	Símbolo	Probabilidad de Símbolo
6	$(1, 0)$	1
6	$(0, 0)$	0
6	$(0, 1)$	0
6	$(1, 1)$	0

#### 4.5. Comportamiento de la Entropía Simbólica $h_\tau(m)$ con $p$ y $r_e$ para tamaños de muestra finito

En la sección 2.8.1 describimos una serie de criterios generales para la selección del tamaño de símbolo  $m$  de las aplicaciones de simbolización y vimos que en general el tamaño de símbolo  $m$  venía limitado por a) el número de observaciones  $T$  mínimo para la caracterización de la distribución bajo la hipótesis nula y b) La correlación de  $X_m(t)$  con  $X_m(t + \tau), X_m(t + 2\tau), \dots, X_m(t + (m - 1)\tau)$  exige un tamaño  $m$  es pequeño comparado con  $T - (m - 1)\tau$ , para que las dependencias entre  $X_m(t)$  y  $X_m(t + k)$  se consideren débiles y puedan ser despreciadas. Bajo estas dos premisas el tamaño de símbolo recomendado es un valor de  $m$  pequeño ( $m \leq 6$ ) y siempre en relación con el tamaño de muestra  $T$ , ya que obviamente el número de símbolos de las aplicaciones de simbolización nunca debe ser mayor que el tamaño de muestra  $T$ . En el caso de las series temporales caóticas, el número de observaciones disponibles no será mucho mayor que el tamaño de las series de las variables macroeconómicas por lo que podemos operar con el mismo rango de tamaño de símbolo  $m$  ( $m \leq 6$ ) recomendado en la sección 2.8.1. Con respecto al parámetro distancia  $r_e$ , si éste es muy grande y excede el tamaño del atractor, la entropía  $h_\tau(m)$  alcanzará su valor de saturación mínimo ya que casi todos los puntos estarán dentro del radio  $r_e$ , independientemente del valor de  $\tau$ . En cambio si  $r_e$  es muy pequeño, el valor de la entropía  $h_\tau(m)$  tendrá su valor mínimo cuando  $\tau$  es pequeño, es decir, cuando las correlaciones entre  $X(t)$  y  $X(t + k)$  son grandes; ya que para estos valores, casi todos los puntos están fuera de  $r_e$ . Además, a medida que las correlaciones entre  $X(t)$  y  $X(t + k)$  desaparecen con el tiempo, la entropía alcanza su valor máximo, puesto que los puntos se distribuyen uniformemente alrededor de  $r_e$  independientemente del valor de  $\tau$ .

Como ejemplo representativo ilustramos el comportamiento de la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$  aplicada a las simbolizaciones  $f_{r_e}$  (sección 4.3), y  $f_{r_e}^1$  (sección 4.4) con los parámetros  $m$  (tamaño de símbolo) y  $r_e$  para el sistema de Lorenz de tres ecuaciones

diferenciales [56]:

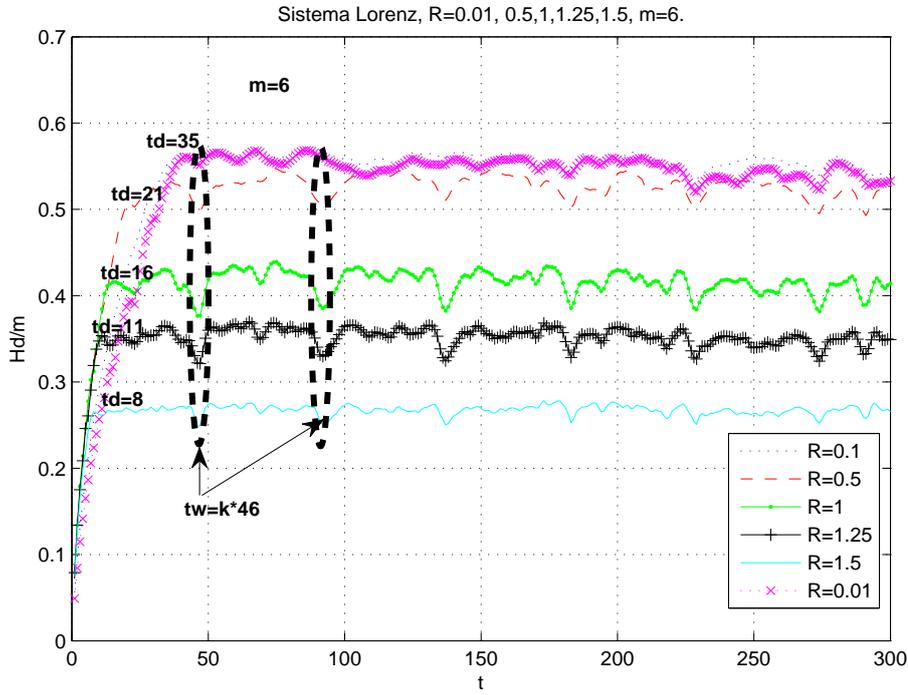
$$\dot{x} = -a(x - y), \dot{y} = -xz + cx - y, \dot{z} = xy - bz \quad (4.31)$$

donde  $a$ ,  $b$  y  $c$  son constantes. Resolvemos este sistema de ecuaciones con  $a = 16.0$ ,  $b = 4.0$  y  $c = 45.92$  y  $\tau_s = 0.01$  (paso de integración) y valores iniciales  $x_0 = -1$ ,  $y_0 = 0$  y  $z_0 = 1$  para 3000 puntos de la variable  $x$ . Para manejar resultados comparativos, representamos los valores de la entropía por símbolo  $h_\tau(m)/m$  para varios valores de  $m$  ( $m=4,5$  y  $6$ ) y  $r_e$ , este último lo representamos con el *parámetro normalizado*  $R = r_e/\sigma$  ( $\sigma = std(X(t))$ ), con  $R=0.01, 0.1, 0.5, 1, 1.25, 1.5$  para las aplicaciones de simbolización  $f_{r_e}$  y  $f_{r_e}^1$ . Los resultados se muestran en las figuras 4.1, 4.2, 4.3 para la aplicación  $f_{r_e}$  y 4.4, 4.5, 4.6 para el caso de la aplicación  $f_{r_e}^1$ . Lo primero que podemos observar es que en las dos aplicaciones el retardo temporal óptimo correspondiente al primer máximo de  $h_\tau(m)/m$  depende del radio  $r_e$ . Así, nuestras aplicaciones de simbolización  $f_{r_e}$  y  $f_{r_e}^1$  proponen varios candidatos a retardo temporal  $\tau^2$  en función del radio  $r_e$  que utilicemos. Además éstos van apareciendo con más precisión conforme aumentamos el tamaño de símbolo  $m$ . En cada una de las figuras aparecen señalados los primeros máximos locales de  $h_\tau(m)/m$  en función de  $r_e$ . Los retardos temporales, con el criterio establecido, se eligen en estos puntos, los cuales como se puede observar se encuentran en  $\tau = 8, 11, 16, 21$  y  $35$  para  $f_{r_e}$  y  $\tau = 10, 11, 13$  y  $14$  para  $f_{r_e}^1$ . Como podemos ver uno de los retardos encontrados, el retardo  $\tau = 10$ , está de acuerdo con el retardos temporal encontrado por Abarbanel et al. [1] usando el primer mínimo de la Información Mútua; mientras que el retardo  $\tau = 11$  coincide con el hallado por Kim et. al. [50] con el método C-C y Wei-Dong Cai et. al. en [82] con el método C-C-3. También se puede ver cómo para la gran mayoría valores del radio  $r_e$ , la *Entropía  $\tau$ -simbólica*  $h_\tau(m)/m$  tiene extremos locales en retardos  $\tau$  con periodicidad  $P_w = 46$  independientemente de la Aplicación de Simbolización empleada. En este grupo de retardos la *Entropía  $\tau$ -simbólica*  $h_\tau(m)/m$  tiene valores muy superiores o muy inferiores al resto de retardos. Según los criterios de

---

<sup>2</sup>En todas las figuras representadas este capítulo, el retardo temporal  $\tau$  óptimo seleccionado se designa con la letra  $t_d$  y la ventana temporal  $\tau_w$  con  $t_w$ .

Figura 4.1:  $h_\tau(m)/m$  ( $m = 6$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$ ,  $c = 45.92$  y  $\tau_s = 0.01$  para la Aplicación de Simbolización  $f_{r_e}$ .



selección de la ventana de inmersión  $\tau_w$  expuestos en los apartados 4.3 y 4.4 la ventana temporal es el primero de estos extremos locales periódicos, en este caso  $\tau_w = 46$ . Este valor está de acuerdo con el valor del retardo de ventana temporal  $\tau_w = 46$  encontrado con los métodos C-C-3 y C-C en Li [54] (para más detalles de ambos métodos ver sección 4.2.4).

Figura 4.2:  $h_\tau(m)/m$  ( $m = 5$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$ ,  $c = 45.92$  y  $\tau_s = 0.01$  para la aplicación Simbólica  $f_{re}$ .

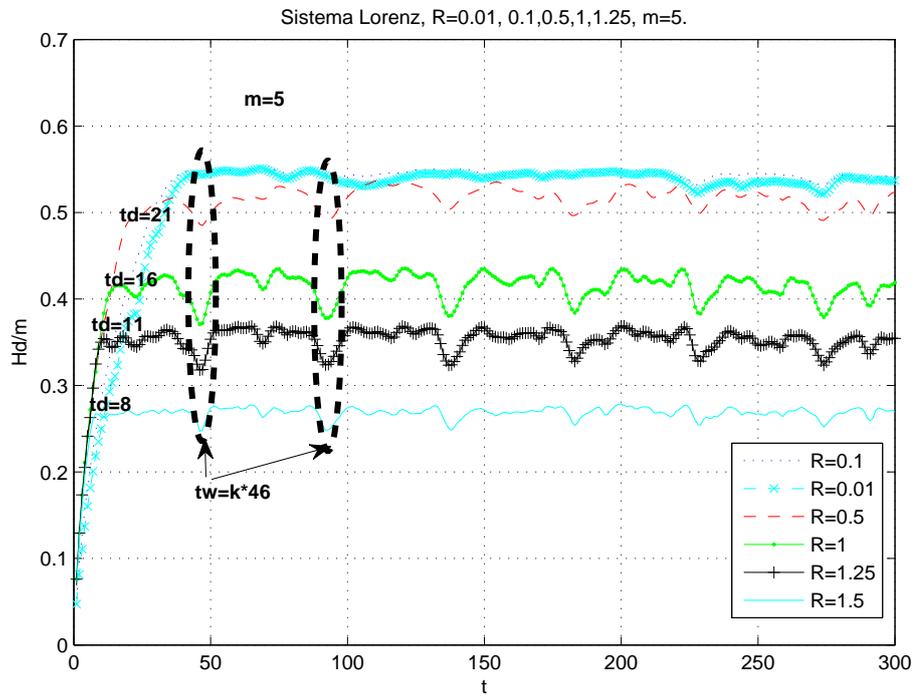


Figura 4.3:  $h_\tau(m)/m$  ( $m = 4$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$ ,  $c = 45.92$  y  $\tau_s = 0.01$  para la aplicación Simbólica  $f_{re}$ .

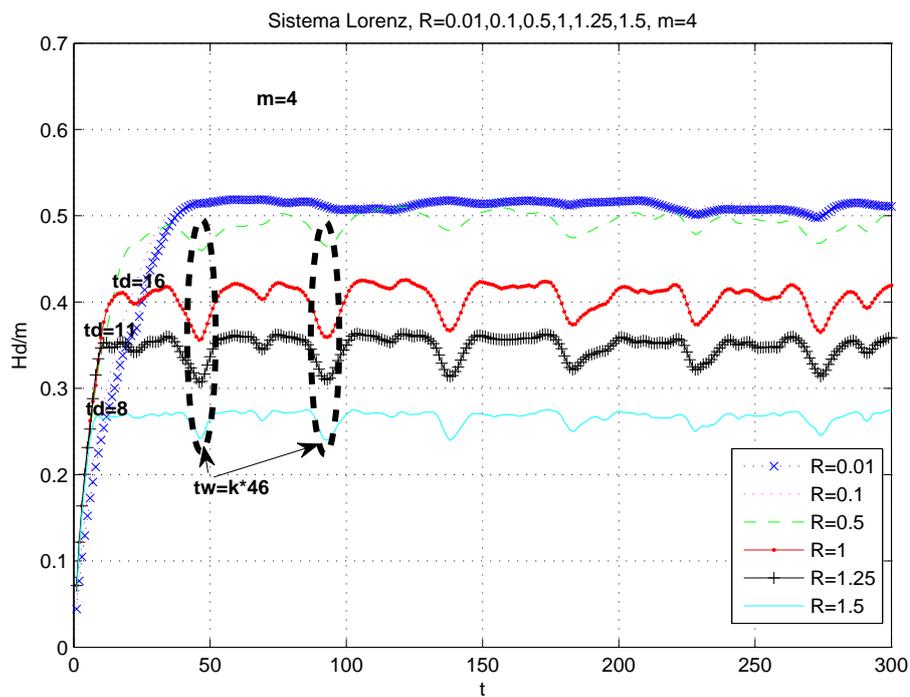


Figura 4.4:  $h_\tau(m)$  ( $m = 6$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$ ,  $c = 45.92$  y  $\tau_s = 0.01$  para la aplicación Simbólica  $f_{r_e}^1$ .

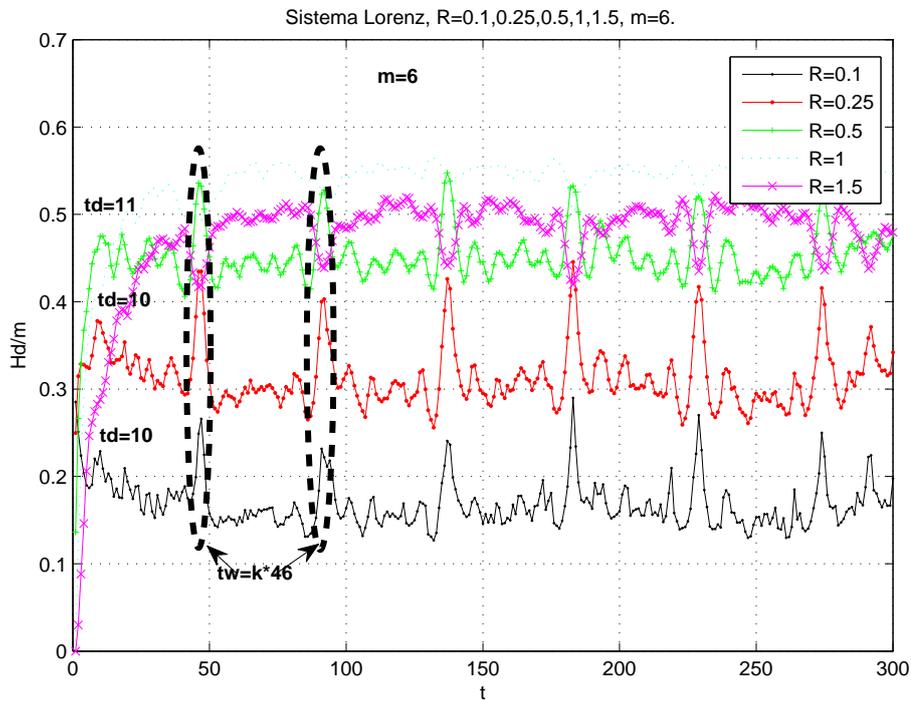


Figura 4.5:  $h_\tau(m)/m$  ( $m = 5$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$  y  $c = 45.92$ ,  $\tau_s = 0.01$  para la aplicación Simbólica  $f_{r_e}^1$ .

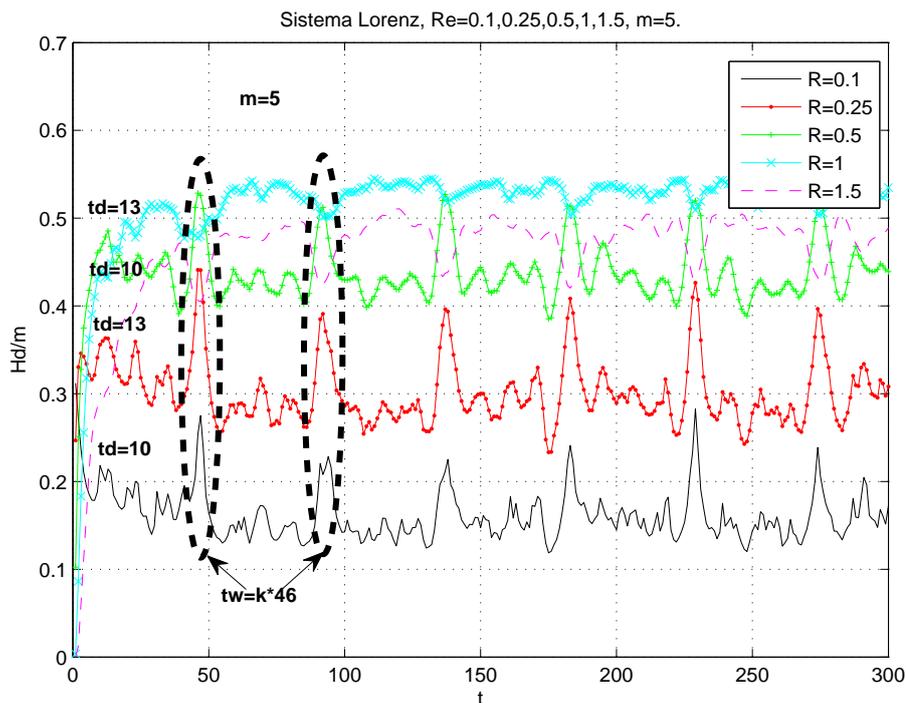
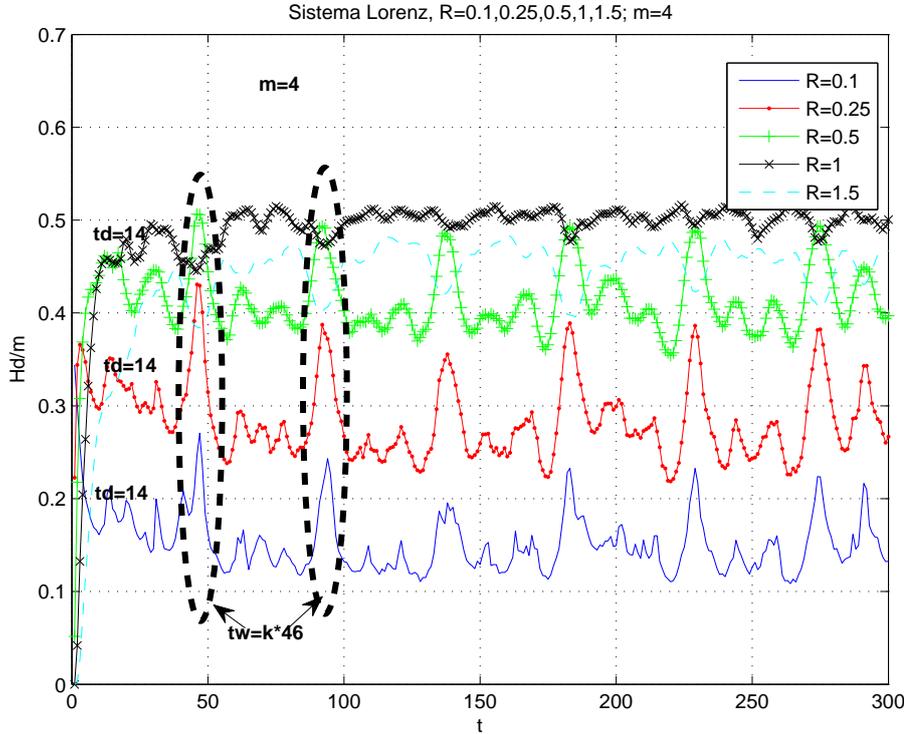


Figura 4.6:  $h_\tau(m)/m$  ( $m = 4$ ) para la variable  $x$  del sistema de Lorenz de la ecuación 4.31 con  $a = 16.0$ ,  $b = 4.0$ ,  $c = 45.92$ ,  $\tau_s = 0.01$  para la aplicación Simbólica  $f_{r_e}^1$ .



#### 4.6. Selección del parámetro $r_e$

Es obvio que las aplicaciones de simbolización  $f_{r_e}$  y  $f_{r_e}^1$  dependen del valor del parámetro de distancia  $r_e$  y consecuentemente la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$ , por lo que una cuestión práctica para su cálculo es la elección de dicho parámetro  $r_e$ . Como ventaja, la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$  proporciona el criterio para elegir el parámetro  $r_e$ , ya que, con ciertos límites, deja que sean los datos del sistema caótico en estudio los que, a través de la *Entropía  $\tau$ -simbólica*  $h_\tau(m)$ , muestren el valor óptimo de  $r_e$ .

Para la detección del parámetro correcto  $r_e$ , consideramos el  $r_e^{opt}$  relevante como aquel (entre todos los posibles valores  $r_e$ ) que maximiza el valor absoluto del Indicador  $FP$  [65], siendo  $FP$  una *función pico* que asocia una puntuación  $FP(i, x_i)$ , para un punto  $x_i$  de posición  $i$ -ésima en la entropía simbólica  $h_\tau(m)$  asociada a la serie  $\{X_t\}_{t \in I}$  con la cual pretendemos detectar los picos periódicos locales de  $h_\tau(m)$ . Si el retardo de la ventana óptima se produce en  $t = P_w$ , siendo  $P_w$  el período medio orbital cuya entropía  $\tau$ -simbólica  $h_\tau(m)$  se encuentra en el primero (por orden temporal) de los  $k$  máximos

o  $k$  mínimos locales en  $t = k \cdot P_w$ , para nosotros el parámetro distancia  $r_e^{opt}$  será aquel que maximice/minimice “los picos de la Entropía  $\tau$ -Simbólica  $h_\tau(m)$ ”, es decir, el valor de  $h_\tau(m)$  en los  $t = k \cdot P_w$  respecto de los retardos temporales adyacentes. Estos retardos  $t = k \cdot P_w$  múltiplos del período medio orbital, se caracterizan por el hecho que su configuración espacial (representada por el valor de su entropía simbólica) varía brúscamente respecto de sus retardos adyacentes en un entorno local. y esa variación es máxima cuando  $r_e = r_e^{opt}$ .

La cuestión importante es cómo calcular la función  $FP$ . A continuación describiremos el método, descrito ampliamente en [65], utilizado para encontrar el valor óptimo de  $r_e = r_e^{opt}$ .

Sea  $k$  un entero mayor que cero y  $N^+(k, i) = \langle h_{i+1}, h_{i+2}, \dots, h_{i+k} \rangle$  la secuencia de  $k$  vecinos temporales por la derecha de  $h_i$ , i.e, los  $k$  puntos que siguen inmediatamente al punto  $h_i$  de la serie  $\{h_\tau(m)\}$ .  $N^-(k, i) = \langle h_{i-1}, h_{i-2}, \dots, h_{i-k} \rangle$  se define similarmente como el conjunto de los  $k$  vecinos temporales por la izquierda (previos) de  $x_i$ . Sea  $N'(k, i) = N^+(k, i) \bullet \{h_i\} N^-(k, i)$ . Entonces sea  $k < i < N - k$ , para un punto dado  $h_i$ , la siguiente función  $FP_1$  calcula la media de (i) el máximo entre las distancias con signo de  $h_i$  con sus  $k$  vecinos a la izquierda y (ii) el máximo entre las distancias con signo de  $h_i$  con sus  $k$  vecinos a la derecha. Normalmente son adecuados valores bajos de  $k$  (i.e 3 a 5) si la mayoría de los picos son “delgados”. Los valores de  $FP_1$  indican la significancia de la altura del pico en el instante  $i$ .

$$FP^+(k, i, h_i) = \frac{\max \{h_i - h_{i-1}, h_i - h_{i-2}, \dots, h_i - h_{i-k}\} + \max \{h_i - h_{i+1}, h_i - h_{i+2}, \dots, h_i - h_{i+k}\}}{2} \quad (4.32)$$

que es la versión de  $FP_1$  para detectar los picos cuando son máximos locales.

Si quisiéramos detectar los picos de la serie  $\{h_\tau(m)\}$  cuando éstos son mínimos locales entonces  $S$  toma la forma:

$$FP^-(k, i, h_i) = \frac{\min \{h_i - h_{i-1}, h_i - h_{i-2}, \dots, h_i - h_{i-k}\} + \min \{h_i - h_{i+1}, h_i - h_{i+2}, \dots, h_i - h_{i+k}\}}{2} \quad (4.33)$$

Una vez que tenemos la serie  $\{FP(k, i, h_i)\}$  se escoge el máximo valor del indicador  $\max\{FP^+(k, i, h_i)\}$  cuando  $\{h_\tau(m)\}$  posee picos máximos locales o el mínimo valor del indicador  $\min\{FP^-(k, i, h_i)\}$  cuando  $\{h_\tau(m)\}$  posee picos mínimos locales.

Por lo tanto el procedimiento de selección del parámetro óptimo  $r_e$  de las aplicaciones  $f_{r_e}$  y  $f_{r_e}^1$  se resume en los siguientes pasos:

1. Sea el conjunto  $O = \emptyset$
2. Elegir  $R = \frac{r_e}{std(\{X_t\})} \in (0, \frac{\max(abs(\{X_t\}))}{std(\{X_t\})})^3$ , esto es el parámetro distancia  $r_e$  normalizado por la desviación típica de  $\{X_t\}$ ,  $R = \frac{r_e}{std(X_t)}$ .
3. Calcular la serie *Entropía  $\tau$ -simbólica*  $h_\tau(m)$  correspondiente a  $f_{r_e}$  y  $f_{r_e}^1$  para  $R$  y  $\tau \leq (m - 1)T$ . Siendo  $T$  la longitud de la serie  $\{X_t\}$ .
4. Calcular  $FP_1 = \max\{FP^+(k, i, h_i)\}$  y  $FP_2 = \min\{FP^-(k, i, h_i)\}$ . Hallar  $FP_m = \max\{FP_1, |FP_2|\}$  y  $O = O \cup \{FP_m\}$ .
5. Volver al punto 1 hasta que  $R < \max(abs(\{X_t\}))$ .
6. Seleccionar el  $R$  que hace  $\max_{R \in (0, \max(abs(\{X_t\})))} \{FP_m\}$ .

## 4.7. Análisis de Simulación

Para ilustrar el comportamiento de los métodos de selección de los parámetros de reconstrucción basados en las aplicaciones de simbolización  $f_{r_e}$  y  $f_{r_e}^1$ , los utilizamos en algunos sistemas caóticos conocidos. Para cada uno de los sistemas se generó una serie temporal  $\{X_t\}_{t \in I}$  de variable  $x$  de 3000 puntos (donde los 3000 puntos se seleccionan del rango 3001-6000 de 100000 puntos simulados) y después calculamos las *Entropías  $\tau$ -Simbólicas*  $h_\tau(m)$  correspondientes a las dos aplicaciones de simbolización  $f_{r_e}$  y  $f_{r_e}^1$  con el tamaño de símbolo  $m = 6$  con los valores de  $R^{opt}$  ( $r_e^{opt}$  normalizados por la desviación típica de la serie caótica  $\{X_t\}_{t \in I}$ ) estimados mediante el método de selección descrito en

---

<sup>3</sup>  $\frac{\max(abs(\{X_t\}))}{std(\{X_t\})}$  es el máximo valor que puede tomar el parámetro  $R$ , ya que el recorrido de  $r_e$  es el intervalo  $(0, \max(abs(\{X_t\}))$ ). En la práctica, un intervalo razonable para encontrar  $R$  estaría en el intervalo  $(0, 2]$ , tal y como propone Brock en [20] para la elección del parámetro  $r$  de la Integral de Correlación  $C(p, T, r, \tau)$ ,

el apartado 4.5. Los  $R^{opt}$  calculados para cada uno de los sistemas caóticos se adjuntan en la Tabla 4.27. Además, con el fin de refrescar algunos de los criterios de selección de los principales métodos de reconstrucción que se analizan en esta tesis; en la tabla 4.25 resumimos los criterios de selección de los parámetros  $\tau$  y  $\tau_w$  para el método de Información Mútua de Fraser y Swinney [33], el método C-C y el método C-C- $S_c$ . Con el mismo fin, también resumimos los criterios de selección de  $\tau_d$  y  $\tau_w$  para los métodos propuestos  $f_{re}$  y  $f_{re}^1$  en la tabla 4.26. Los sistemas analizados son los siguientes:

1. El sistema caótico de Lorenz de tres ecuaciones diferenciales acopladas [56]

$$\begin{aligned}\dot{x} &= -a(x - y) \\ \dot{y} &= -xz + cx - y \\ \dot{z} &= xy - bz\end{aligned}\tag{4.34}$$

con  $a$ ,  $b$  y  $c$  constantes. Como en el apartado 4.5, se resuelve el sistema de ecuaciones con  $a = 16,0$ ,  $b = 4,0$  y  $c = 45,92$  con  $\tau_s = 0.01$  (paso de integración) y valores iniciales  $x_0 = -1$ ,  $y_0 = 0$  y  $z_0 = 1$ . Los resultados de la entropías simbólica  $h_\tau(6)$  se muestran en la figura 4.7 para la aplicación  $f_{re}$  y en la figura 4.10 para la aplicación  $f_{re}^1$ . Es fácil encontrar en ambas figuras el primer máximo local. En el caso de  $f_{re}$  el primer máximo se corresponde con el retardo temporal  $\tau = 12$  mientras que para  $f_{re}^1$  éste se localiza en el retardo temporal  $\tau = 9$ . El retardo temporal  $\tau = 12$  está de acuerdo con el retardo encontrado por Wei-Dong Cai et. al. en [83], mientras que el retardo  $\tau_d = 9$  está muy cerca de los retardos temporales hallados ( $\tau = 10$ ,  $\tau = 11$  y  $\tau = 12$ ) en [50], [83] y [82]. De los resultados mostrados de la figura 4.7 podemos observar que la entropía simbólica  $h_\tau(6)$  exhibe fuerte pseudoperiodicidad y tiene mínimos locales cuando  $t = kP_w$ , donde  $P_w$  es el periodo medio orbital, el tiempo de dependencia de los datos. Concretamente, posee mínimos locales en 46, 92, 138, 184, 230 y 276. De esta forma la ventana de inmersión óptima  $\tau_w$  es igual a 46, el primer máximo local pseudoperiódico. Así estimamos el valor de la dimensión de inmersión  $p$  a partir de

la relación  $\tau_w = (p - 1)\tau$ , por lo que  $p$  es aproximadamente igual a 5. Por último, en la figura 4.10 se puede observar que la entropía simbólica  $h_\tau(6)$  de  $f_{re}^1$  exhibe máximos locales en los puntos 46, 92, 138, 184, 230 y 276, por lo que al igual que ocurre con la aplicación  $f_{re}$ , seleccionamos la ventana de inmersión en el retardo  $\tau_w = 46$ . En ambos casos, la ventana de inmersión  $\tau_w = 46$  está de acuerdo con el valor de la ventana encontrado en los trabajos anteriores de Wei-Dong Cai et. al. [82, 83]. En este caso la dimensión de inmersión es aproximadamente igual a  $p=6$ .

2. El sistema caótico de Rossler de tres ecuaciones diferenciales acopladas [68]. Las ecuaciones de Rossler están dadas por

$$\begin{aligned} \dot{x} &= -y - z & (4.35) \\ \dot{y} &= x + dy \\ \dot{z} &= z(x - f) + e \end{aligned}$$

donde  $d = 0,15$ ,  $e = 0,2$  y  $f = 10$ , paso de integración  $\tau_s = 0,05$  y valores iniciales  $x_0 = -1$ ,  $y_0 = 0$  y  $z_0 = 1$ . La entropía simbólica  $h_\tau(6)$  de la aplicación  $f_{re}$  se muestra en la figura 4.8, donde se observan los mínimos locales de  $h_\tau(6)$  en  $t = 121$ , 242 y por consiguiente el periódico caótico se encuentra en  $P_w = 121$ . Por lo tanto la ventana de inmersión es igual al retardo  $\tau_w = 121$ . Además el retardo óptimo temporal se localiza en el primer máximo local  $\tau = 18$ . Este valor del retardo está muy cerca del retardo  $\tau = 17$  encontrado en [50] y [82], mientras que el retardo de la ventana temporal está de acuerdo con el retardo  $\tau_w = 121$  de [82]. Consecuentemente la dimensión de inmersión es  $p = 8$ . En la figura 4.11 podemos observar que bajo la aplicación  $f_{re}^1$ , los máximos locales de  $h_\tau(6)$  se localizan también en los puntos  $t = 121$ , 242; mientras que el retardo óptimo temporal se da en el primer máximo local  $\tau = 24$ . En este caso, al igual que en la aplicación  $f_{re}$ , la ventana de inmersión se encuentra en el retardo  $\tau_w = 121$  y la dimensión de inmersión, con los valores de los parámetros de reconstrucción

$\tau = 24$  y  $\tau_w = 121$ , es igual a  $p = 6$ .

### 3. El sistema caótico de Duffing

$$\begin{aligned}\dot{x} &= y & (4.36) \\ \dot{y} &= -gy - kx(1 + x^2) + l\cos z \\ \dot{z} &= m\end{aligned}$$

donde  $g$ ,  $k$ ,  $l$  y  $m$  son constantes. Resolvemos este sistema de ecuaciones para  $g = 0,05$ ,  $k = 0,25$ ,  $l = 7,5$  y  $m = 1$ , con condiciones iniciales  $x_0 = -1$ ,  $y_0 = 0$  y  $z_0 = 1$  con un paso de integración  $\tau_s = 0.05$ . De los resultados en la figura 4.9 podemos observar que bajo la aplicación  $f_{re}$ , la entropía simbólica  $h_\tau(6)$  tiene mínimos locales pseudo-periódicos cuando  $t$  es igual a 126 y 252. Por lo tanto la ventana de dimensión óptima  $\tau_w$  es igual a 126 aproximadamente y el retardo temporal óptimo se localiza en el primer máximo local  $\tau = 14$ . Los valores de ambos parámetros están de acuerdo con los encontrados en [83] y [82]. Como resultado, la dimensión de inmersión  $p$  es igual a 10. En la figura 4.12 la entropía simbólica  $h_\tau(6)$  de  $f_{re}^1$  tiene máximos locales en los puntos  $t = 126, 252$ , por lo que la ventana de inmersión coincide con el retardo  $\tau_w = 126$ . También podemos ver que el retardo temporal óptimo se encuentra en el primer máximo local  $\tau = 11$  y la dimensión de inmersión es igual a  $p = 12$ .

A partir de los resultados obtenidos, la siguiente cuestión que debemos abordar es si los valores de los parámetros  $p$  y  $\tau_d$  son adecuados para la reconstrucción del espacio de estados:

1. En primer lugar podemos hacer una validación de los parámetros de reconstrucción calculando la caracterización geométrica del sistema caótico y ello se realiza con el cálculo del exponente de Lyapunov mayor de una serie temporal (LLE). El LLE es una cantidad significativa de la medida de un sistema caótico y cuantifica la evolución de la separación de dos órbitas a lo largo del tiempo que inicialmente estaban muy cercanas. Un sistema es caótico si el LLE es mayor que cero, pero el

espacio de estados debe reconstruirse antes de calcular el LLE de una serie temporal. Así, mediante la caracterización LLE podemos decidir si los parámetros elegidos son buenos o no. Primeramente, los parámetros seleccionados se aplican para reconstruir el espacio de estados de una variable de una serie temporal y entonces calculamos los LLE de las series temporales correspondientes mediante un método de cálculo para conjuntos de datos pequeños <sup>4</sup>. Los resultados de los LLE y los parámetros de reconstrucción  $\tau$  y  $\tau_w$  se muestran en las Tablas 4.30 y 4.29 para los métodos de reconstrucción C-C, C-C- $S_c$  y los basados en las aplicaciones  $f_{re}$  y  $f_{re}^1$  bajo los sistemas caóticos de Lorenz, Rossler y Duffing. El método C-C, como hemos comentado, es uno de los métodos más populares utilizados para el proceso de selección de los parámetros de reconstrucción  $\tau$  y  $\tau_w$  [50]. Así mismo, para superar varias de las desventajas del método C-C (descritas en el apartado 4.2.4), hemos creído conveniente realizar la comparación también con su variante C-C- $S_c$ . También incluimos en la Tabla 4.29 los retardos temporales encontrados mediante el célebre método de la Información Mutua  $MI$  de Fraser y Swinney [33]. Por último, en la última columna de las tablas se incluyen los valores de los LLE de referencia estimados en [76, 87, 67]. Podemos observar que los valores de LLE son muy cercanos a los calculados en [76, 87, 67] para casi todos los métodos en los tres sistemas caóticos, excepto en el sistema de Lorenz para el método C-C, en este caso el valor estimado del LLE es un poco peor que los valores de referencia. También podemos destacar que el parámetro de reconstrucción  $\tau_w$  está de acuerdo en los tres métodos C-C- $S_c$ ,  $f_{re}$  y  $f_{re}^1$  bajo los tres sistemas caóticos. Sin embargo, aunque los valores estimados del retardo temporal  $\tau_d$  no coinciden en casi ninguno de los métodos, éstos no difieren demasiado en la mayoría de los casos.

2. En segundo lugar podemos comprobar si los parámetros de reconstrucción están de acuerdo con la dimensión de correlación estimada. La dimensión de correlación es una medida importante de la complejidad del atractor de un sistema caótico. La

---

<sup>4</sup>Para una descripción detallada del método utilizado para el cálculo del LLE ver [67].

Tabla 4.25: Resumen Criterios de Selección de los Parámetros de Reconstrucción  $\tau$  y  $\tau_w$  de los Métodos MI, C-C y C-C- $S_c$

Parámetro de Selección	Método MI	Método C-C	Método C-C- $S_c$
$\tau$	1º Mín $I(x(t)x(t + \tau))$	1º Mín $\Delta\bar{S}_1(t)$	
$\tau_w$		Mín $S_{corr}(t)$	1º Máx $\bar{S}_c(t)$ en $t = kT_p$

Tabla 4.26: Resumen Criterios de Selección de los Parámetros de Reconstrucción  $\tau$  y  $\tau_w$  de los Métodos  $f_{re}$  y  $f_{re}^1$

Parámetro de Selección	Método $f_{re}$	Método $f_{re}^1$
$\tau$	1º Máx $h_\tau(m)/m$	1º Máx $h_\tau(m)/m$
$\tau_w$	1º Mín $h_\tau(m)/m$	1º Mín $h_\tau(m)/m$ en $t = kT_p$

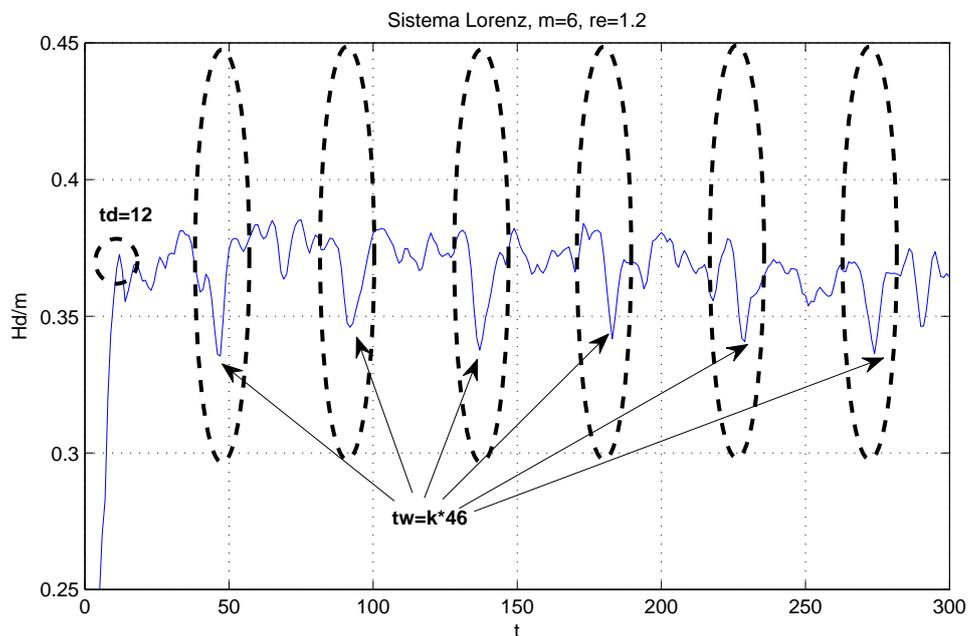
dimensión de correlación se calcula en primer lugar estimando la integral de correlación  $C(r)$  mediante el método de Grassberger y Proccacia [39, 40]. Luego se calcula la dimensión de correlación mediante la expresión  $D_2 = d\log(C(r))/d\log(r)$  en el límite cuando  $\Delta r \rightarrow 0$ ,  $r \rightarrow 0$  y  $N \rightarrow \infty$ , donde  $r$  es la distancia de separación entre dos puntos arbitrarios en la órbita del espacio de estados. En la Tabla 4.31 se muestran los resultados de la dimensión de correlación para las aplicaciones  $f_{re}$ ,  $f_{re}^1$ , C-C y C-C- $S_c$  en los sistemas caóticos de Lorenz, Rossler y Duffing junto con los valores teóricos estimados. Se puede observar claramente que en los tres sistemas caóticos, los valores estimados de la dimensión de correlación para los métodos C-C- $S_c$ ,  $f_{re}$  y  $f_{re}^1$  están muy cerca de los valores teóricos.

En resumen, en vista de los resultados obtenidos, los métodos basados en las aplicaciones de simbolización  $f_{re}$  y  $f_{re}^1$  han demostrado ser competitivos y efectivos frente a los métodos C-C y C-C- $S_c$  por lo que podemos decir que ambos métodos se pueden usar de forma satisfactoria para la selección de unos buenos parámetros de reconstrucción  $\tau$  y  $\tau_w$ .

Tabla 4.27: Valor óptimo de  $R^{opt}$  para los Sistemas Caóticos de Lorenz, Rossler y Duffing.

Sistema	Parámetros	$f_{re}$	$f_{re}^1$
Lorenz	$a=16, b=4.0, c=45.92$	1.2	0.25
Rosler	$d=0.2, e=0.4, f=5.7$	0.4	0.1
Duffng	$g=0.05, k=0.25, l=7.5, m=1$	0.275	0.25

Figura 4.7:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}$ .



#### 4.8. Efectos del Ruido en la Precisión (Error Estándar $\sigma$ ) y Exactitud (Sesgo) de los Métodos de Selección de los Parámetros de Inmersión (Retardo y Ventana Temporales) para la Reconstrucción del Espacio de Estados de un Sistema Caótico

En los apartados anteriores hemos comparado varios métodos de selección de los parámetros de reconstrucción del espacio de estados  $\tau$  y  $\tau_w$  de varios sistemas caóticos

Figura 4.8:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 0,4$  para la serie  $\{X_t\}$  del Sistema de Rossler para la Aplicación  $f_{re}$ .

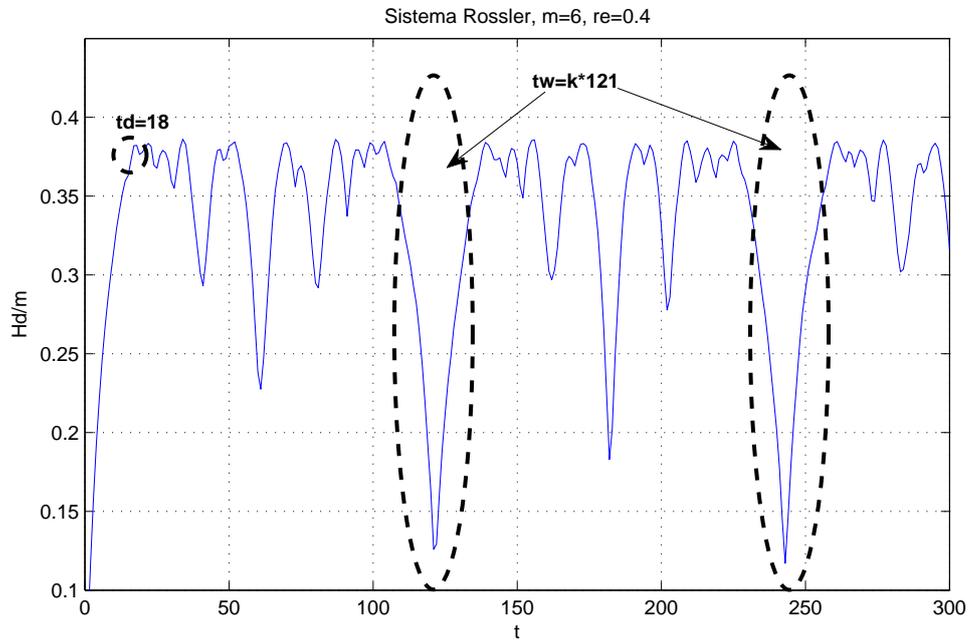


Figura 4.9:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 0,275$  para la serie  $\{X_t\}$  del Sistema de Duffing para la Aplicación  $f_{re}$ .

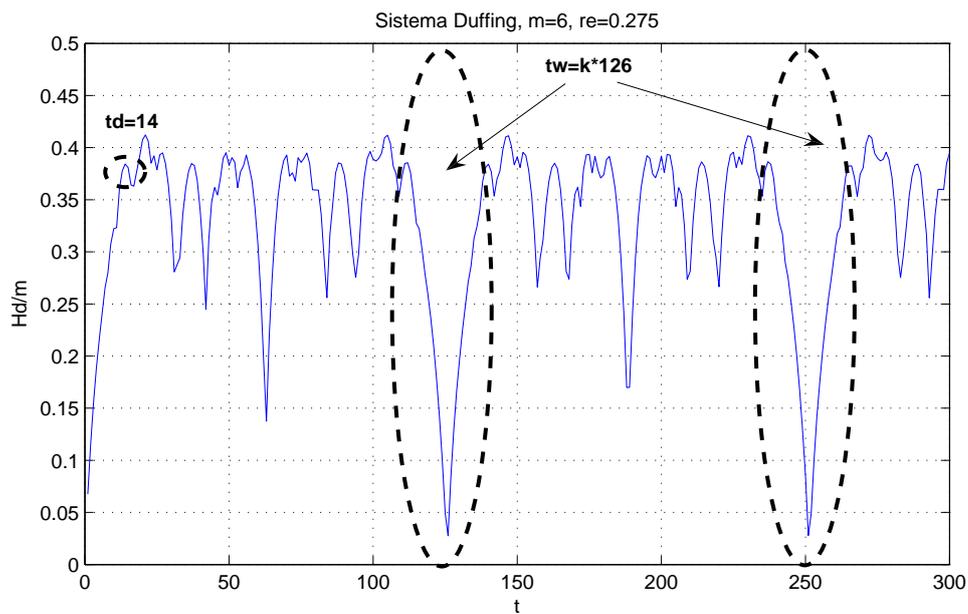


Figura 4.10:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 0,25$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}^1$ .

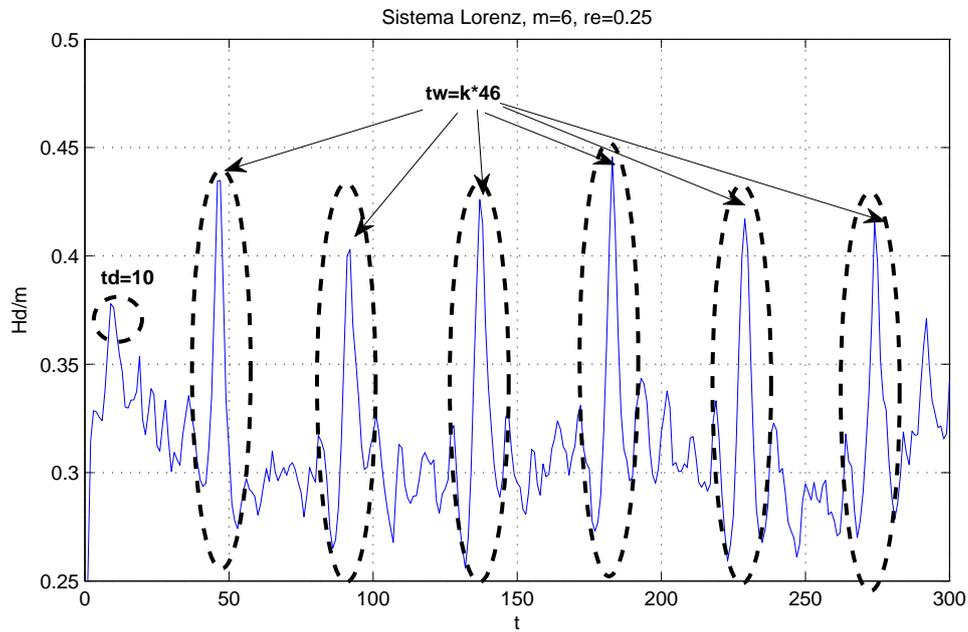


Figura 4.11:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 0,1$  para la serie  $\{X_t\}$  del Sistema de Rossler para la Aplicación  $f_{re}^1$ .

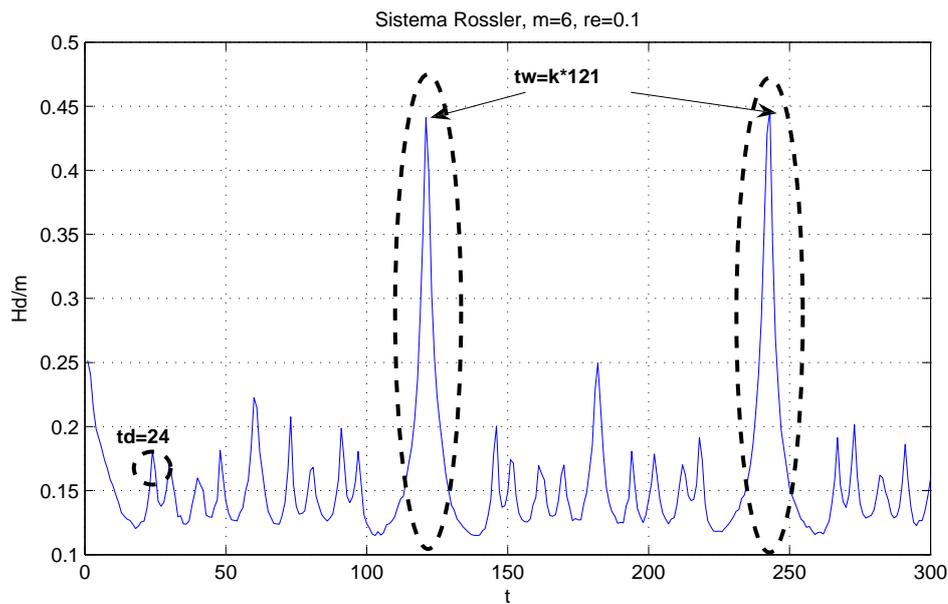


Tabla 4.29: LLE y parámetros de reconstrucción  $\tau(t)$  y  $\tau_w(t)$  para los métodos de selección  $MI$ , C-C y C-C-Sc.

Sistema	Parámetros	$\tau_s$	$MI$	Método C-C			Método C-C-Sc				
			$\tau(t)$	$\tau(t)$	$\tau_w(t)$	$p$	LLE	$\tau_w(t)$	$p$	LLE	Ref
Lorenz	$a=16, b=4.0, c=45.92$	0.01	0.11(11)	0.1(10)	1(100)	11	0.94	0.46(46)	5	1.5	1.5
Rosler	$d=0.15, e=0.2, f=10$	0.05	1(20)	0.85(17)	9.55(191)	12	0.095	6.05(121)	8	0.099	0.09
Duffng	$g=0.05, k=0.25, l=7.5, m=1$	0.05	0.6(12)	0.7(14)	8.05(161)	12	0.168	6.3(126)	10	0.184	0.183

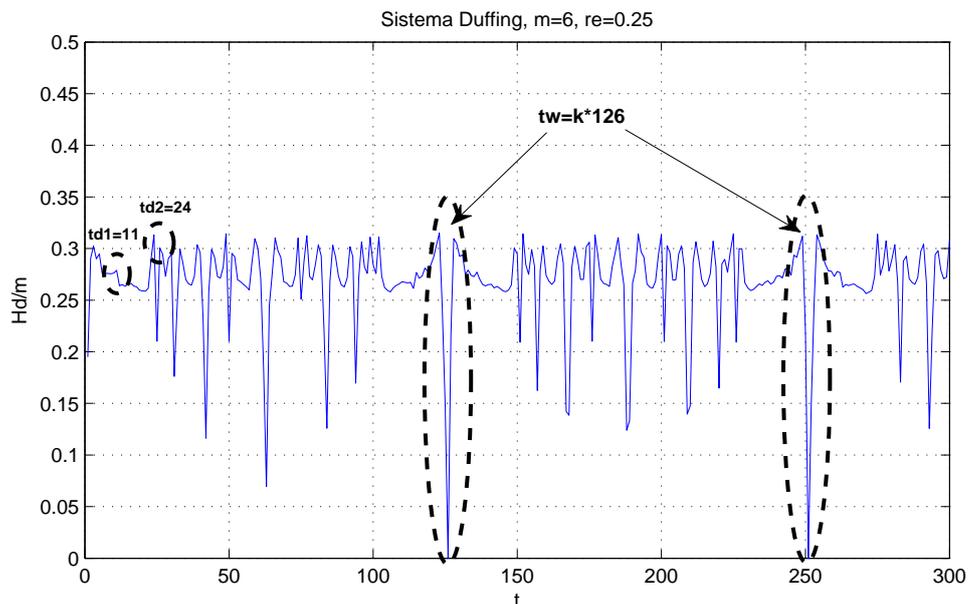
Tabla 4.30: LLE y parámetros de reconstrucción  $\tau_d(t)$  y  $\tau_w(t)$  para los métodos de selección  $f_{re}$  y  $f_{re}^1$ .

Sistema	Parámetros	$\tau_s$	$f_{re}$			$f_{re}^1$					
			$\tau(t)$	$\tau_w(t)$	$p$	LLE	$\tau(t)$	$\tau_w(t)$	$p$	LLE	Ref
Lorenz	$a=16, b=4.0, c=45.92$	0.01	0.12(12)	0.46(46)	5	1.518	0.09(9)	0.46(46)	6	1.56	1.5
Rosler	$d=0.15, e=0.2, f=10$	0.05	0.9(18)	6.05(121)	8	0.099	1.2(24)	6.05(121)	6	0.099	0.09
Duffng	$g=0.05, k=0.25, l=7.5, m=1$	0.05	0.7(14)	6.3(126)	10	0.184	0.55(11)	6.3(126)	12	0.189	0.183

Tabla 4.31: Dimensión de Correlación  $D_2$  para los métodos de selección C-C y C-C-Sc,  $f_{re}$  y  $f_{re}^1$ .

Sistema	Parámetros	Método C-C			Método C-C-Sc			$f_{re}^1$	Ref
Lorenz	$a=16, b=4.0, c=45.92$	1.93		2.004	2.008	2.01	2.06	2.06	
Rosler	$d=0.15, e=0.2, f=10$	2.13		2.08	2.09	2.08	2.01	2.01	
Duffng	$g=0.05, k=0.25, l=7.5, m=1$	2.1		2.21	2.21	2.3	2.23	2.23	

Figura 4.12:  $h_\tau(m)/m$  y  $m = 6$  con  $R = 0,25$  para la serie  $\{X_t\}$  del Sistema de Duffing para la Aplicación  $f_{re}^1$ .

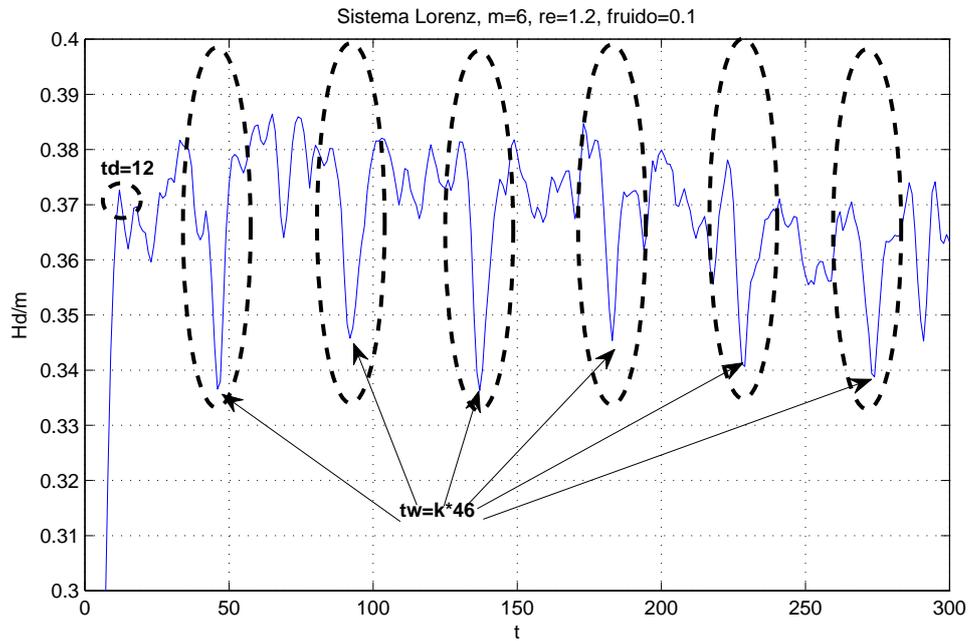


de longitud limitada en un entorno ideal libre de ruido. Sin embargo, éste no será el caso que nos encontremos en la realidad, lo normal es que las señales caóticas estén contaminadas por ruido aleatorio. El ruido aleatorio influye en la exactitud y precisión de los métodos de selección de los parámetros  $\tau$  y  $\tau_w$  de forma que la calidad de la reconstrucción del atractor del sistema caótico se ve deteriorada en mayor o menor medida. Podemos diferenciar el proceso de selección del retardo y la ventana temporales en condiciones de ruido en dos fases: a) en primer lugar es necesario encontrar los máximos/mínimos locales de Información (según los criterios establecidos por el algoritmo utilizado) candidatos a los retardos  $\tau$  y  $\tau_w$  y b) en segundo lugar hay que contrastar si estos máximos/mínimos son retardos de Información significativos respecto a sus retardos vecinos. La primera parte está relacionada con la exactitud o sesgo del algoritmo de selección, es decir, cuánto nos desviamos en media de los valores de los retardos óptimos  $\tau$  y  $\tau_w$  en condiciones libres de ruido. La segunda está relacionada con la precisión o error estándar en la selección de los parámetros del algoritmo, es decir cómo de dispersos están los retardos  $\tau$  y  $\tau_w$  seleccionados bajo condiciones de selección reiterada. La influencia del ruido tanto en la exactitud como en la precisión de la selección de los parámetros de reconstrucción se traduce en que por ejemplo, si

el retardo temporal es muy diferente al óptimo, se pueden presentar casos de *redundancia* cuando  $X_t$  y  $X_{t+k\tau}$  están muy próximos ( $k = 1, \dots, (p - 1)$ ). En este caso las medidas consecutivas de los vectores reconstruidos darán casi los mismos resultados y los vectores topológicos reconstruidos se estrechan a lo largo de la línea de identidad, en el espacio de inmersión  $m$ -dimensional, conduciendo a una dimensión del atractor reconstruido cercano a uno. También nos podemos encontrar con el fenómeno de *irrelevancia* cuando  $x_t$  y  $X_{t+k\tau}$  están muy alejados, en este caso concreto los  $X_t$  y  $X_{t+k\tau}$  poseen escasa relación y el vector reconstruido está formado por componentes irrelevantes. Con respecto a la ventana temporal  $\tau_w$ , si ésta, para un retardo temporal fijo  $\tau$ , es muy diferente a la óptima, la selección de ventanas de inmersión más grandes conlleva dimensiones más grandes que las requeridas por los datos y por lo tanto a un exceso de computación o incluso una amplificación de la “contaminación” mediante la introducción de ruido en las dimensiones adicionales del espacio de inmersión donde la dinámica del atractor real no está operando. Al contrario, si la ventana de inmersión  $\tau_w$  es más pequeña que la óptima entonces da lugar a la selección de dimensiones más pequeñas que la del atractor por lo que éste no se desdobra totalmente y algunos de los puntos encontrados como vecinos en esa dimensión pueden estar realmente lejos en la dimensión correcta y de esta manera el proceso de reconstrucción resulta fallido.

Para estudiar el efecto del ruido en la selección de los retardos  $\tau$  y  $\tau_w$  añadimos ruido gaussiano a la serie temporal de Lorenz. Específicamente, examinamos las series temporales  $X_t = X_t^L + \eta\sigma\varepsilon_i$  donde  $\{X_t^L\}$  es la serie temporal de Lorenz libre de ruido,  $\sigma$  es su desviación típica,  $\varepsilon_i$  es una variable aleatoria i.i.d gaussiana con media cero y desviación típica de 1, y  $\eta$  es el “nivel de fuerza” del ruido. Hemos añadido niveles del ruido del 10 %, 30 %, 50 % y 100 % ( $\eta = 0.1, 0.3, 0.5$  y  $1.0$ ) a la serie temporal de Lorenz. Además las Aplicaciones de Simbolización  $f_{r_e}$  y  $f_{r_e}^1$  se calcularon para cada uno de estos niveles de ruido con los valores de  $R^{opt}$  de la Tabla 4.27. Junto con estos dos métodos, también se estudió el efecto del ruido para los algoritmos de selección MI (Información Mutua), C-C y C-C( $S_c$ ). Además en las figuras 4.13 a 4.16 para  $f_{r_e}$  y 4.17 a 4.20 para  $f_{r_e}^1$  podemos ver gráficamente el efecto de cada uno de los niveles de ruido

Figura 4.13:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}$ . Factor de ruido  $\eta = 10\%$ .



en la selección del retardo  $\tau$  y ventana  $\tau_w$  temporales. Como se puede observar, en las dos aplicaciones  $f_{re}$  y  $f_{re}^1$  el estimador de  $\tau_w$  permanece prácticamente inalterado para todos los valores de  $\eta$ , pero no para el estimador de  $\tau$  cuando  $\eta=0.5$  y  $\eta=1$ .

Figura 4.14:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}$ . Factor de ruido  $\eta = 30\%$ .

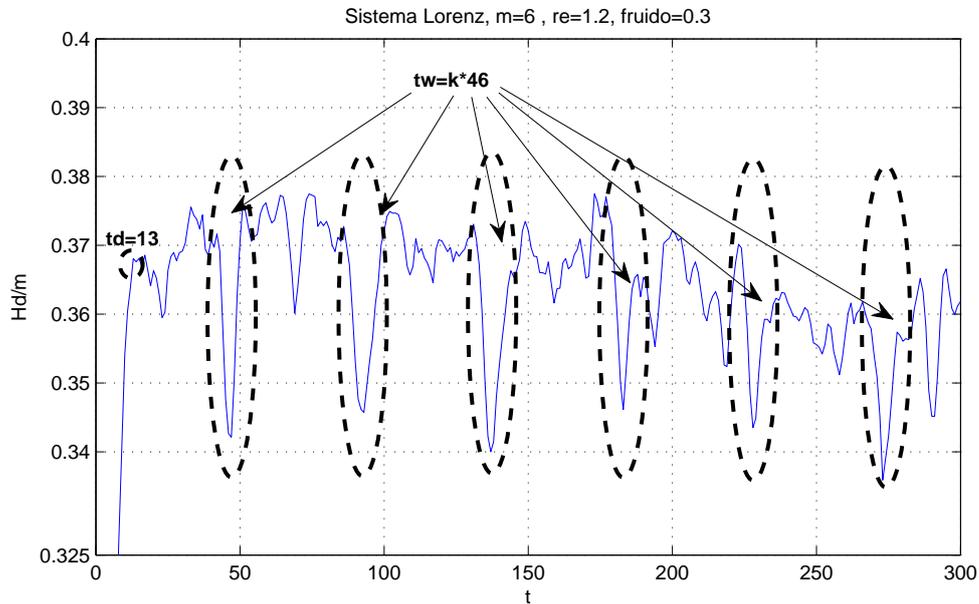


Figura 4.15:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}$ . Factor de ruido  $\eta = 50\%$ .

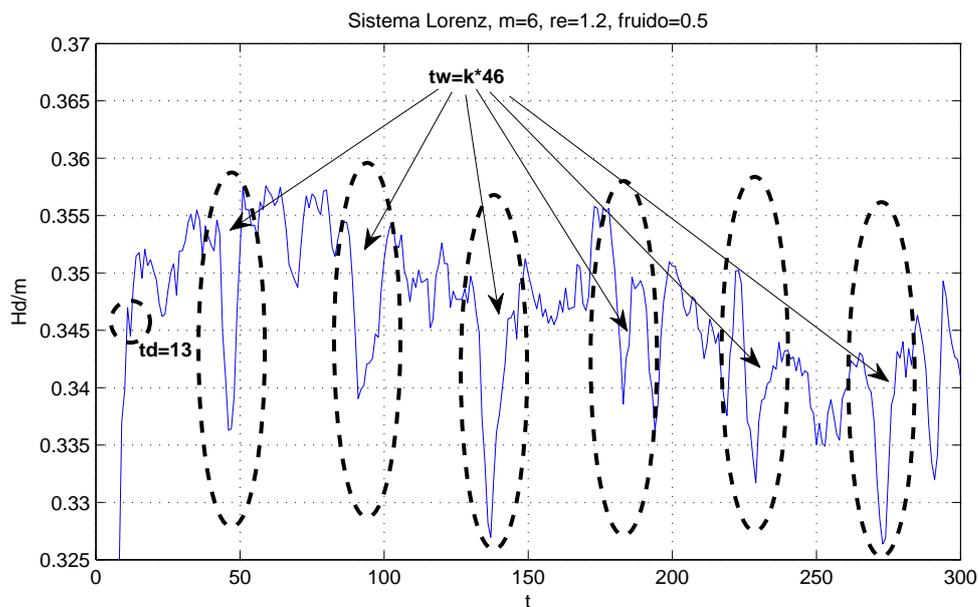


Figura 4.16:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}$ . Factor de ruido  $\eta = 100\%$ .

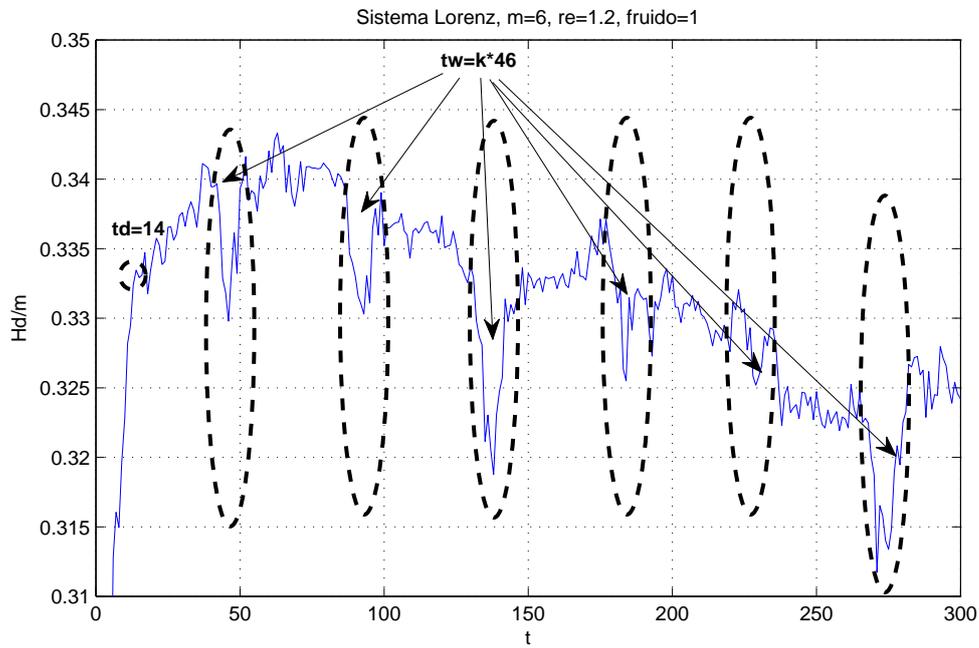


Figura 4.17:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}^1$ . Factor de ruido  $\eta = 10\%$

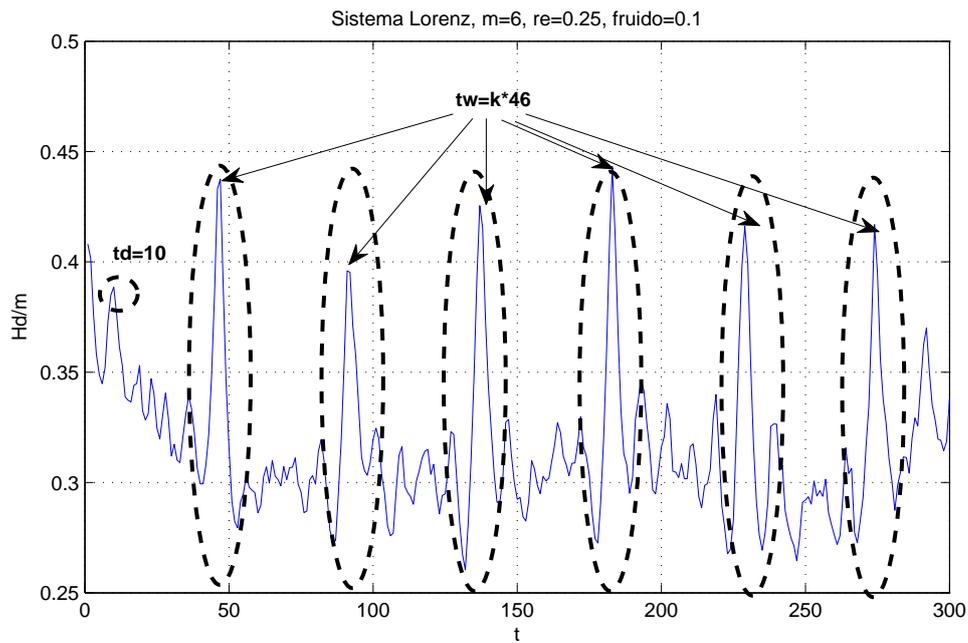


Figura 4.18:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}^1$ . Factor de ruido  $\eta = 30\%$ .

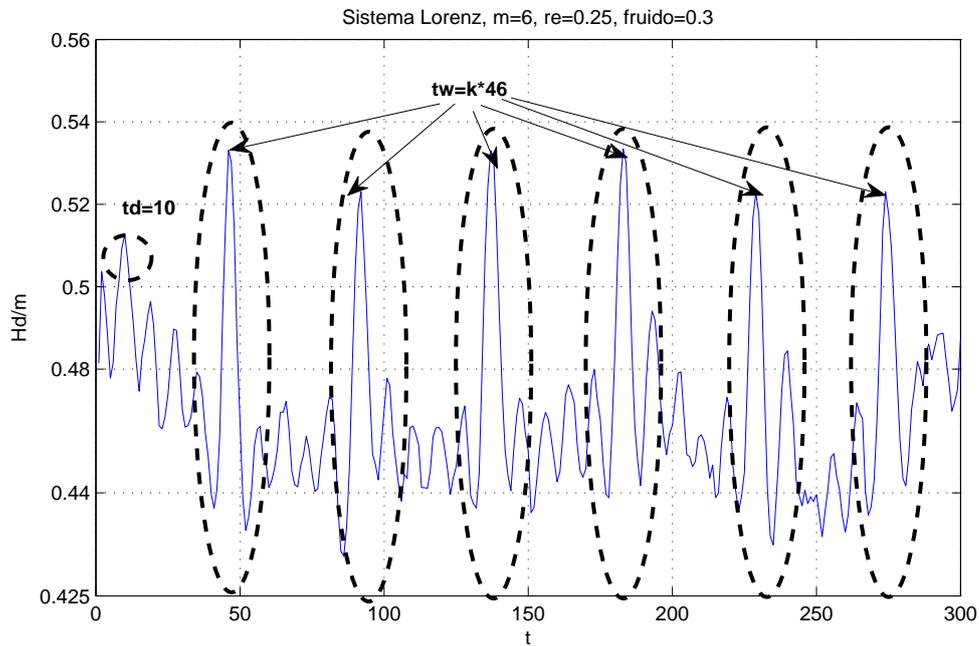


Figura 4.19: Entropía Simbólica por símbolo  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}^1$ . Factor de ruido  $\eta = 50\%$ .

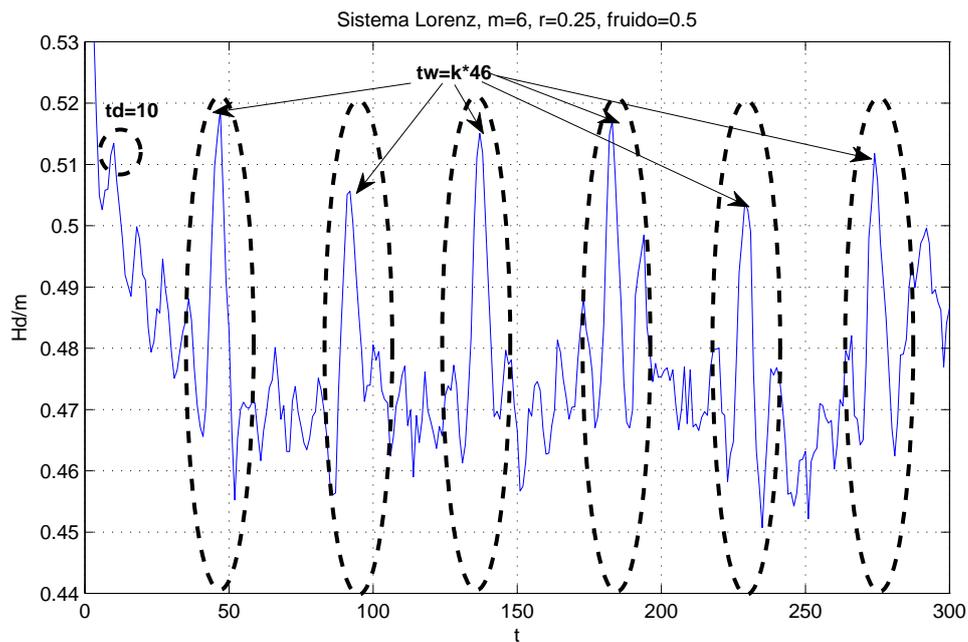
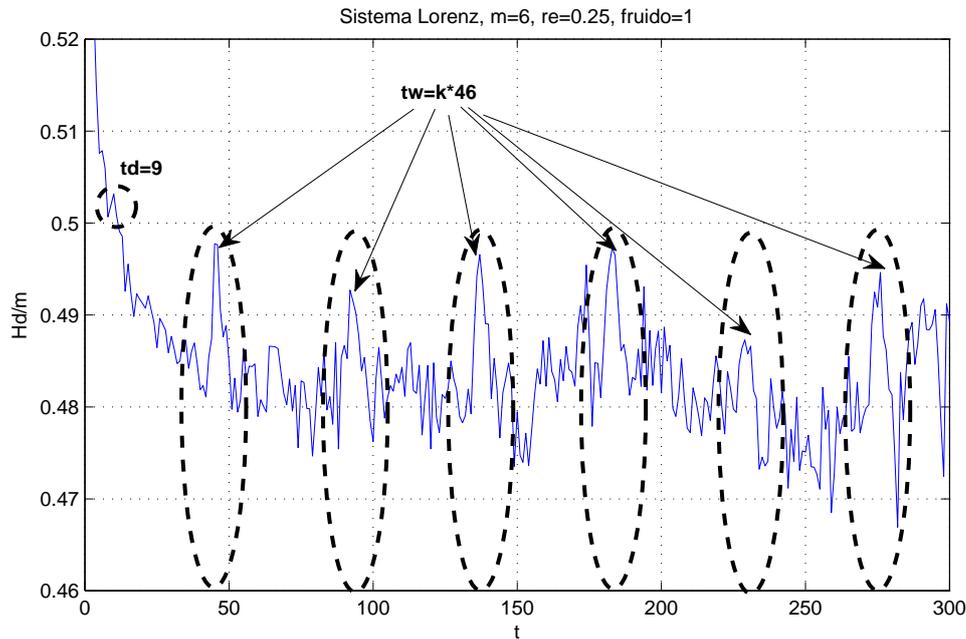


Figura 4.20:  $h_\tau(m)/m$  con  $m = 6$  y  $R = 1,2$  para la serie  $\{X_t\}$  del Sistema de Lorenz para la Aplicación  $f_{re}^1$ . Factor de ruido  $\eta = 100\%$ .



#### 4.8.1. Efectos del Ruido en la Exactitud (Sesgo)

Para mostrar el efecto del ruido en la exactitud o sesgo de los métodos MI (Información Mútua), C-C, C-C( $S_c$ ),  $f_{re}$  y  $f_{re}^1$  en la selección de  $\tau$  y  $\tau_w$  computamos 1000 réplicas del modelo de Lorenz del apartado 4.7 para cada uno de los métodos de selección y cada nivel de ruido (10%, 30%, 50% y 100%). Después calculamos el valor medio de los funcionales de Información correspondientes a cada uno de los métodos de selección y por último seleccionamos los extremos relativos de dichos funcionales correspondientes (máximo o mínimo local) a cada uno de los parámetros de inmersión  $\tau$  y  $\tau_w$ . Los resultados del retardo y ventana temporales medios para cada uno de los métodos de selección y cada nivel de ruido se muestran en las Tabla 4.32. Al lado de cada valor seleccionado de  $\tau$  y  $\tau_w$  se adjuntan entre paréntesis sus valores en condiciones libres de ruido (ver Tablas 4.29 y 4.30). Como podemos observar en los métodos  $f_{re}$ ,  $f_{re}^1$  y C-C( $S_c$ ) los valores estimados de la ventana temporal  $\tau_w$  apenas cambian conforme aumenta el nivel de ruido, si bien en los métodos C-C y C-C( $S_c$ ) estos valores se mantienen más estables. También vemos que los valores medios estimados del retardo óptimo temporal  $\tau$  fluctúan en todos los métodos cuando aumenta el nivel de ruido,

Tabla 4.32: Valores medios de selección de los parámetros de inmersión de retardo y ventana temporales en función de los niveles de ruido.

	MI	C-C		C-C( $S_c$ )	$f_{re}$		$f_{re}^1$	
	$\tau$	$\tau$	$\tau_w$	$\tau_w$	$\tau$	$\tau_w$	$\tau$	$\tau_w$
P.1	12(11)	13(10)	100(100)	46(46)	13(12)	46(46)	9(9)	46(46)
P.2	14(11)	15(10)	100(100)	46(46)	14(12)	47(46)	9(9)	46(46)
P.3	11(11)	15(10)	100(100)	46(46)	16(12)	47(46)	8(9)	46(46)
P.4	8(11)	15(10)	97(100)	46(46)	8(12)	46(46)	9(9)	46(46)

aunque la el nivel de variación es mayor en los métodos MI,  $f_{re}$  y  $f_{re}^1$ . A pesar de ello, los valores estimados de  $\tau$  permanecen siempre en un factor de 4 dentro de los valores correctos. En general se puede observar una mayor robustez al ruido de los métodos C-C y C-C( $S_c$ ), esta robustez es debida al hecho que las ecuaciones 4.17 y 4.18 suponen promediar el ruido por lo que éste al compensarse tiende a 0.

#### 4.8.2. Efectos del Ruido en la Precisión (Error Estándar $\sigma$ ) : Contraste Estadístico de los Parámetros de Inmersión (Retardo y Ventana Temporales) para la Reconstrucción del Espacio de Estados de un Sistema Caótico

Para comprobar los efectos del ruido en la precisión vamos a tomar un camino diferente y vamos a proponer un test de contraste de significatividad de los retardos de inmersión  $\tau$  y  $\tau_w$  seleccionados basado en la teoría de la información. Más precisamente la selección del retardo y ventana temporales  $\tau$  y  $\tau_w$  como extremos del funcional de Información se realizará contrastando si esos parámetros contienen más información que sus vecinos. Con el fin de mostrar cómo contrastar la validez estadística de la elección de cualquier parámetro de inmersión ( $\tau$  y  $\tau_w$ ) se ha construido un test no paramétrico de una cola basado en bootstrap<sup>5</sup>.

<sup>5</sup>A partir de ahora, dentro de los apartados 4.8.2 y 4.8.3, al retardo temporal  $\tau$  candidato a retardo óptimo lo denominaremos  $\tau^*$  para distinguirlo de la variable retardo genérica, que la designaremos con  $\tau$ , cuando necesitemos referirnos a ambos parámetros dentro del mismo contexto.

#### 4.8.2.1. Marco Estadístico de Definición y Diseño del Test

Supongamos que tenemos una serie temporal  $\{X_t\}_{t \in I}$ , siendo  $I$  un conjunto de índices temporales de cardinalidad  $T$ , según el proceso generador de datos caótico de la forma

$$x_t = G(X_{t-1}) + \varepsilon_t \quad (4.37)$$

donde  $X_{t-1} = \{x_{t-1}, x_{t-2}, \dots\}$ ,  $G$  es una función medible no lineal de  $X_{t-1}$  y  $\varepsilon_t$  es el residual i.i.d independiente de  $X_{t-1} \forall t$  con  $t = 2, \dots, T$ , siendo  $T$  el número de elementos de la serie. Éste es un caso particular del modelo estocástico de serie temporal

$$x_t = G(X_{t-1}, \theta) + \varepsilon_t$$

donde  $G$  es una función medible lineal o no lineal de  $X_{t-1} = \{x_{t-1}, x_{t-2}, \dots\}$  y  $\theta$  es un vector de parámetros. En estas condiciones, según se demostró en [62], la importancia de la identificación del retardo temporal óptimo de una serie, radica en el hecho de que tanto en una serie de tiempo con datos observados, como en una generada por un proceso de la forma 4.37, existe un retardo temporal significativo que aporta una mayor información para reproducir la estructura dinámica correcta del proceso subyacente. Basándonos en esta proposición, vamos a derivar a continuación condiciones de significatividad para los retardos de reconstrucción  $\tau$  y  $\tau_w$ .

Supongamos que hemos empleado un método de selección “ $X$ ” para seleccionar como candidatos a parámetros de reconstrucción un retardo  $\tau = \tau^*$  y una ventana  $\tau_w$  temporales de la serie caótica  $x_t$ . Entonces decimos que  $\tau^*$  y  $\tau_w$  son significativos si en ambos retardos existen extremos locales en sentido estricto del funcional de Información  $F(I_t)$  generado por el proceso  $\{X_t\}_{t \in I}$  y asociado al método “ $X$ ”. Ahora bien, los retardos  $\tau^*$  y  $\tau_w$  se han seleccionado de forma “visual” en la gráfica del funcional  $F(I_t)$  en función del retardo  $\tau$ ,  $F(\tau)_{\tau=1,2,\dots,n}$ . Debido a que  $\{X_t\}_{t \in I}$  está contaminada con “ruido” entonces necesitamos saber si los valores de  $F(I_t)$  seleccionados en  $\tau^*$  y  $\tau_w$  son extremos locales significativos o son efectos espúreos del ruido. Es por esto que es necesario contrastar estadísticamente la existencia de extremo local en  $\tau^*$  y  $\tau_w$ , es

decir, si los valores de  $F(\tau^*)$  y  $F(\tau_w)$  son máximos o mínimos locales respecto de los retardos  $\tau$  adyacentes de sus entornos  $Ent(\tau^*)$  y  $Ent(\tau_w)$ . La existencia de extremo local en los parámetros elegidos  $\tau^*$  y  $\tau_w$  se establece con la condición de que además de cumplirse la relación de orden necesaria derivada de la condición de extremo local, los diferenciales de Información con signo  $F(\tau^*)-F(\tau)$  y  $F(\tau_w)-F(\tau)$  sean estadísticamente significativos respecto a los retardos de sus entornos locales  $Ent(\tau^*)$  y  $Ent(\tau_w)$ . Si la relevancia del diferencial de Información de  $\tau^*$  y  $\tau_w$  no es estadísticamente significativa respecto a cualquier otro posible retardo de sus entornos  $\tau \in Ent(\tau^*)$ ,  $\tau \in Ent(\tau_w)$  nada nos impide seleccionar cualquier otro  $\tau$  adyacente como candidato a parámetro de reconstrucción. Esto nos lleva a ampliar el número de candidatos a retardo y ventana óptimos, lo que disminuye lógicamente la precisión del método de selección de estos retardos. En lo que sigue procederemos a definir las relaciones de orden necesarias para la condición de extremo local de un funcional de Información  $F(I_t)$ . Esto es necesario para posteriormente formular las hipótesis nula y alternativa del test que se va a proponer.

Sean  $\tau^*$  y  $\tau_w$  el retardo y la ventana temporales donde existen extremos locales estrictos (máximos o mínimos) de la función de Información  $F(I_t)$ . Entonces existe un entorno reducido con centro en  $\tau^*$  y  $\tau_w$ ,  $Ent(\tau^*)$  y  $Ent(\tau_w)$  para los cuales se cumple que

$$\begin{aligned} F(\tau) < F(\tau^*), \quad F(\tau) < F(\tau_w) \text{ máximo local estricto} \\ F(\tau) > F(\tau^*), \quad F(\tau) > F(\tau_w) \text{ mínimo local estricto} \end{aligned} \tag{4.38}$$

$\forall \tau \in Ent(\tau^*), \forall \tau \in Ent(\tau_w)$  según sean máximos o mínimos locales respectivamente. Si las relaciones son además de igualdad estamos entonces ante la condición de máximo y mínimo local no estricto respectivamente

$$\begin{aligned} F(\tau) \leq F(\tau^*), \quad F(\tau) \leq F(\tau_w) \text{ máximo local no estricto} \\ F(\tau) \geq F(\tau^*), \quad F(\tau) \geq F(\tau_w) \text{ mínimo local no estricto} \end{aligned} \tag{4.39}$$

$\forall \tau \in Ent(\tau^*), \forall \tau \in Ent(\tau_w)$ .

Como hemos dicho, las relaciones de orden 4,38 son estadísticamente significativas cuando los diferenciales de Información con signo  $F(\tau^*)-F(\tau)$  y  $F(\tau_w)-F(\tau)$  son relevantes para todos los retardos de sus entornos  $\tau \in Ent(\tau^*)$  y  $\tau \in Ent(\tau_w)$  respectivamente. En estas condiciones la precisión del método de selección es máxima. Al contrario, si estas relaciones de orden no son significativas, entonces el diferencial informacional de los parámetros  $\tau^*$  y  $\tau_w$  en correspondencia con los retardos  $\tau \in Ent(\tau^*)$  y  $\tau \in Ent(\tau_w)$  será irrelevante.

Bajo estas premisas se propone llevar a cabo un test no paramétrico de una cola cuya hipótesis nula consiste en contrastar las relaciones de orden 4.38 necesarias para cumplir la condición de extremo local estricto:

$\mathbb{H}_0$ : *El Funcional de Información  $F(I_t)$  no tiene un extremo local estricto en  $\tau=\tau_f$ ,  $\tau_f=\tau^*$ ,  $\tau_w$ . Esto equivale a decir que *El Funcional de Información  $F(I_t)$  en  $\tau_f$ ,  $F(\tau_f)$ , no es significativamente mayor/menor que  $F(\tau)$  para  $\tau \neq \tau_f$ ,  $\tau \in Ent(\tau_f)$ .**

Donde  $\tau_f$  es un retardo candidato a retardo o ventana temporales y  $\tau$  es un retardo que pertenece al entorno de  $\tau_f$ ,  $Ent(\tau_f)$ . Bajo la hipótesis nula el funcional  $F(\tau_p)$  no es significativamente mayor/menor que  $F(\tau)$ , para cualquier retardo  $\tau \neq \tau_f$  que pertenece al entorno reducido de  $\tau_f$ ,  $Ent(\tau_f)$ . Por lo tanto aceptar la hipótesis  $\mathbb{H}_0$  para algún  $\tau \in Ent(\tau_f)$  significa que  $F(\tau_f)$  no es extremo local estricto en  $\tau = \tau_f$  en  $Ent(\tau_f)$  y  $\tau_f$  no es un retardo de inmersión significativo.

Teniendo en cuenta las propiedades de información de los funcionales estudiados en este capítulo: Integral de Correlación (a través de los funcionales  $S_{corr}$  y  $\Delta\bar{S}_1$ ), Información Mutua (MI) y Entropía  $\tau$ -Simbólica ( $h_\tau(m)$ ), la hipótesis nula de no extremo local toma la forma

$$\delta(\tau) = F(\tau_f) - F(\tau) \leq 0 \quad (4.40)$$

si queremos contrastar  $F(\tau_f)$  como candidato a máximo local. Si lo que queremos es contrastar  $F(\tau_f)$  como candidato a mínimo local, la versión correspondiente de la hipótesis nula sería

$$\delta(\tau) = F(\tau_f) - F(\tau) \geq 0 \quad (4.41)$$

Entonces la hipótesis alternativa toma la forma

$\mathbb{H}_1$ : *El Funcional de Información  $F(I_t)$  en  $\tau_f$ ,  $F(\tau_f)$ , es significativamente mayor/menor que  $F(\tau)$   $\tau \neq \tau_f$ ,  $\tau \in Ent(\tau_f)$ .*

Entonces bajo la hipótesis alternativa el retardo  $\tau_f$  es un retardo significativo respecto de  $\tau$  de  $F$  en  $Ent(\tau_f)$ . Cuando  $F(\tau_f)$  es significativamente mayor que  $F(\tau)$  la hipótesis alternativa toma la forma

$$\delta(\tau) = F(\tau_f) - F(\tau) > 0 \quad (4.42)$$

y si  $F(\tau_f)$  es significativamente menor que  $F(\tau)$  entonces la hipótesis alternativa es

$$\delta(\tau) = F(\tau_f) - F(\tau) < 0 \quad (4.43)$$

Rechazar la hipótesis nula  $\forall \tau \in Ent(\tau_f)$  supone identificar el retardo  $\tau_f$  como extremo (máximo/mínimo) local en sentido estricto de forma estadísticamente significativa. Si hay algún retardo  $\tau$  que pertenece al entorno de  $\tau_f$ ,  $Ent(\tau_f)$ , para el que se acepta la hipótesis nula, entonces  $\tau_f$  sólo se puede considerar a lo sumo extremo (máximo/mínimo) local en sentido no estricto. Si trasladamos las condiciones establecidas por las hipótesis nula y alternativa a los funcionales de Información de Información Mutua, Integral de Correlación y Entropía Simbólica, éstos quedan de la siguiente forma:

1. Información Mútua: En este caso el parámetro que tenemos que contrastar es el retardo temporal óptimo  $\tau^*$ . La condición de retardo temporal óptimo corresponde a el primer mínimo de la función de Información Mútua [33]  $I(x_t, x_{t-\tau})$  por lo que la hipótesis nula y la alternativa toman la forma respectivamente

$$\mathbb{H}_0 : I(x_t, x_{t-\tau^*}) - I(x_t, x_{t-\tau}) \geq 0 \quad (4.44)$$

$$\mathbb{H}_1 : I(x_t, x_{t-\tau^*}) - I(x_t, x_{t-\tau}) < 0 \quad (4.45)$$

$$\forall \tau \in Ent(\tau^*).$$

2. Integral de Correlación. Método C-C. Funcionales  $\Delta\overline{S}_1$  y  $S_{corr}$ . Como ya se describió en el apartado 4.2.2, el retardo temporal óptimo en el método C-C corresponde al mínimo de la cantidad  $\Delta\overline{S}_1$ , mientras que la ventana temporal se haya en el mínimo absoluto de la cantidad  $S_{corr}$ . De esta forma las expresiones de hipótesis nula y alternativa son respectivamente para la el retardo  $\tau^*$  y ventana temporal  $\tau_w$

$$\mathbb{H}_0 : \Delta\overline{S}_1(\tau^*) - \Delta\overline{S}_1(\tau) \geq 0 \quad (4.46)$$

$$\mathbb{H}_1 : \Delta\overline{S}_1(\tau^*) - \Delta\overline{S}_1(\tau) < 0 \quad (4.47)$$

$\forall \tau \in Ent(\tau^*)$  y para la condición de ventana temporal

$$\mathbb{H}_0 : S_{corr}(\tau_w) - S_{corr}(\tau) \geq 0 \quad (4.48)$$

$$\mathbb{H}_1 : S_{corr}(\tau_w) - S_{corr}(\tau) < 0 \quad (4.49)$$

$\forall \tau \in Ent(\tau_w)$ .

3. Integral de Correlación Media. Método C-C- $S_c$ . En este caso, como ya se vió en el apartado 4.2.4, la ventana de inmersión óptima corresponde al primer máximo local periódico de la cantidad  $\overline{S}_c(t)$ , por lo que la hipótesis nula y alternativa toman la forma

$$\mathbb{H}_0 : \overline{S}_c(\tau_w) - \overline{S}_c(\tau) \leq 0 \quad (4.50)$$

$$\mathbb{H}_1 : \overline{S}_c(\tau_w) - \overline{S}_c(\tau) > 0 \quad (4.51)$$

$\forall \tau \in Ent(\tau_w)$ .

4. Entropía simbólica. Aplicación  $f_{re}$ . El retardo temporal óptimo se estima como cualquiera de los primeros máximos locales de la Entropía Simbólica de la aplicación  $f_{re}$  y la ventana temporal con el primer mínimo local periódico de la entropía simbólica  $h_\tau(m)$  (ver apartado 4.3) . Con estos criterios las hipótesis

nula y alternativas para el retardo temporal son

$$\mathbb{H}_0 :: h_{\tau^*}(m) - h_{\tau}(m) \leq 0 \quad (4.52)$$

$$\mathbb{H}_1 : h_{\tau^*}(m) - h_{\tau}(m) > 0 \quad (4.53)$$

$\forall \tau \in Ent(\tau^*)$  y para la ventana temporal

$$\mathbb{H}_0 : h_{\tau_w}(m) - h_{\tau}(m) \geq 0 \quad (4.54)$$

$$\mathbb{H}_1 : h_{\tau_w}(m) - h_{\tau}(m) < 0 \quad (4.55)$$

$\forall \tau \in Ent(\tau_w)$ .

5. Entropía simbólica. Aplicación  $f_{re}^1$ . En el apartado 4.4 se estableció como criterio para estimar el retardo temporal óptimo la selección de cualquiera de los primeros máximos locales de la Entropía Simbólica de la aplicación  $f_{re}^1$  y la ventana temporal con el primer máximo local periódico de la entropía simbólica  $h_{\tau}(m)$  (ver apartado 4.3). Así las hipótesis nula y alternativas para el retardo temporal toman la forma

$$\mathbb{H}_0 : h_{\tau^*}(m) - h_{\tau}(m) \leq 0 \quad (4.56)$$

$$\mathbb{H}_1 : h_{\tau^*}(m) - h_{\tau}(m) > 0 \quad (4.57)$$

$\forall \tau \in Ent(\tau^*)$  y para la ventana temporal

$$\mathbb{H}_0 : h_{\tau_w}(m) - h_{\tau}(m) \leq 0 \quad (4.58)$$

$$\mathbb{H}_1 : h_{\tau_w}(m) - h_{\tau}(m) > 0 \quad (4.59)$$

$\forall \tau \in Ent(\tau_w)$ .

Para que el test basado en bootstrap asegure que el DGP remuestreado respeta la hipótesis nula utilizaremos el remuestreo independiente. Con ello, el procedimiento para aplicar el test estadístico tipo bootstrap es el siguiente.

#### 4.8.2.2. Procedimiento del Test Bootstrap

El procedimiento del test bootstrap, con  $B$  réplicas bootstrap se compone de los siguientes pasos:

1. Considerar la serie temporal procedente de un proceso generador de datos caóticos  $\{X_t\}_{t \in I}$ .
2. Calcular el funcional de Información  $F$  sobre la serie  $\{X_t\}_{t \in I}$ .
3. Encontrar el candidato a retardo de inmersión  $\tau_f = \tau^*$ ,  $\tau_w$  que es extremo del funcional de Información  $F$ . Donde el funcional  $F$  está dado por las cantidades de Información Mútua,  $S_{corr}$  y  $\Delta \overline{S}_1$  (método  $C - C$ ),  $\overline{S}_c(t)$  (Método C-C- $S_c$ ) y  $h_\tau(m)$  (Entropía simbólica de las aplicaciones  $f_{re}$  y  $f_{re}^1$ ).
4. Considerar un entorno reducido de retardos con centro en el retardo  $\tau_f$ ,  $Ent(\tau_f)$ .
5. Considerar la estimación  $F(\tau)$  para un conjunto de retardos  $\tau \neq \tau_f$  que pertenecen a  $Ent(\tau_f)$ .
6. Estimar  $\hat{\delta}(\tau) = F(\tau_f) - F(\tau)$ .
7. Obtener una serie independiente bootstrap  $\{X_t\}_{t \in T}^b$  donde  $(b)$  indica que se refiere a la muestra bootstrap.
8. Para cada serie bootstrap  $\{X_t\}_{t \in I}^b$  repetir los pasos 2, 3, 4, 5 y 6 para obtener:
  - a)  $F^b(\tau_f)$
  - b)  $F^b(\tau)$  con  $\tau \neq \tau_p$  que pertenece a  $Ent(\tau_f)$
  - c)  $\hat{\delta}^b(\tau)$ , 4.40, 4.41 de la serie bootstrap.

Se repiten los pasos 7 y 8  $B - 1$ , veces para obtener  $B$  realizaciones bootstrap del estadístico  $\{\hat{\delta}^b(\tau)\}_{b=1}^B$ .

9. Calcular el  $p$ -valor como

$$p_{boots} - valor(\hat{\delta}(\tau)) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}(\hat{\delta}^b(\tau) \leq \hat{\delta}(\tau)) \quad (4.60)$$

cuando queramos contrastar un máximo y

$$p_{boots} - valor(\hat{\delta}(\tau)) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}(\hat{\delta}^b(\tau) \geq \hat{\delta}(\tau)) \quad (4.61)$$

cuando el retardo a contrastar sea un mínimo.  $\mathbf{1}(\bullet)$  es la función indicadora que asigna 1 si el argumento es verdadero y 0 cuando es falso.

La regla de decisión del test es rechazar la hipótesis nula si el  $p_{boots} - valor$  satisface  $p_{boots} - valor < \alpha$  para un nivel de significancia  $\alpha$  dado, y no se rechaza en caso contrario.

### 4.8.3. Pruebas Empíricas

En esta sección investigaremos el rendimiento finito muestral del test de contraste de precisión estadístico del retardo y ventana temporales  $\hat{\delta}(\tau)$  con distintos métodos de reconstrucción del espacio de estados en series temporales caóticas. Para ello hemos diseñado algunos experimentos de Monte Carlo: se han creado procesos generadores de datos (DGP's) de tamaño muestral ( $T$ ) de 3000 observaciones. Estos tamaños muestrales son compatibles con las longitudes habituales encontradas en series temporales caóticas. Para comprobar el rendimiento del test  $\hat{\delta}(\tau)$  hemos utilizado los métodos de selección de Fraser y Swinney [33], el método C-C (ver apartado 4.2.2), el método C-C- $S_c$  (ver apartado 4.2.4) así como los métodos basados en la entropía simbólica  $h_\tau(m)$ ,  $f_{re}$  y  $f_{re}^1$  (ver apartados 4.3 y 4.4). Además en los métodos  $f_{re}$  y  $f_{re}^1$  se ha considerado un tamaño de símbolo  $m = 6$  y los valores de  $R^{opt}$  se seleccionaron según:

1. El modelo de Lorenz en condiciones libre de ruido que se muestran en la Tabla 4.27.
2. Los valores medios de  $R^{opt}$ ,  $\overline{R^{opt}}$ , calculados para cada uno de los DGP's propuestos en los experimentos de tamaño y potencia de las Tablas 4.33 y 4.37 respectivamente. Estos valores son los valores medios del radio  $\overline{R^{opt}}$  correspondientes a la simulación de Monte Carlo de 1000 réplicas de cada uno de los DGP's.

Los radios medios  $\overline{R^{opt}}$  para los DGP's de tamaño y potencia se recogen en las Tablas 4.34 y 4.38 respectivamente.

Para la realización de las pruebas de tamaño y potencia se han considerado los valores de los retardos óptimos  $\tau = \tau^*$  y  $\tau_w$  estimados para cada método de selección en el caso de una serie temporal  $\{X_t\}_{t \in I}$  de 3000 puntos (donde los 3000 puntos se seleccionan en el rango 3001-6000 de 100000 puntos simulados) generada a partir del sistema caótico de Lorenz libre de ruido tal y como se ha descrito en el apartado 4.7. Esto supone considerar que la exactitud de los métodos es perfecta a la hora de evaluar su precisión. Los valores de  $\tau^*$  y  $\tau_w$  estimados para cada uno de los métodos de selección se recogen en las Tablas 4.29 y 4.30. A efectos prácticos, para comprobar el rendimiento de tamaño y potencia del test, las pruebas se realizaron contrastando los retardos seleccionados  $\tau^*$  y  $\tau_w$  con los retardos  $\tau^* - 1$  y  $\tau_w - 1$ . Como hemos explicado en la sección anterior, el contraste de  $\tau^*$  y  $\tau_w$  como retardos de inmersión significativos requiere tener en cuenta el conjunto de retardos pertenecientes a los entornos de  $\tau^*$  y  $\tau_w$ . Por lo tanto en este caso, rechazar la hipótesis nula para  $\tau^*$  y  $\tau_w$ , sólo asegura estadísticamente que ambos son retardos significativos respecto de  $\tau^* - 1$  y  $\tau_w - 1$ . Es decir, cumplen la condición de *extremos locales no estrictos*. No podemos decir nada de la significatividad estadística de  $\tau^*$  y  $\tau_w$  respecto de los retardos  $\tau^* + 1$  y  $\tau_w + 1$ . Cuando consecuencia de aceptar la hipótesis nula para  $\tau^*$  y  $\tau_w$  con el conjunto de intervalos del entorno  $[\tau^* - 1, \tau^* + 1]$  o  $[\tau_w - 1, \tau_w + 1]$  la precisión del método considerado disminuye en un factor de 2 por lo que debe aumentarse el entorno de contraste de ambos parámetros a los intervalos  $[\tau^* - k, \tau^* + k]$ ,  $[\tau_w - k, \tau_w + k]$  con  $k = 2, 3, \dots$  de forma sucesiva hasta conseguir acotar el intervalo de precisión.

Por último, para cada DGP el procedimiento se repitió 1000 veces y la proporción de rechazos de la hipótesis nula de homocedasticidad fue calculada usando un nivel de significación del 5%.

Tabla 4.33: Modelos bajo la Hipótesis Nula.

DGP: $X_t = \varepsilon_t$	Innovaciones $\varepsilon_t$
S.1	$N(0, 1)$
S.2	$U(0, 1)$

#### 4.8.3.1. Pruebas de Tamaño

Para examinar el tamaño estadístico del test bajo  $\mathbb{H}_0$  con cada uno de los métodos de selección hemos considerado dos DGP's. Teniendo en cuenta que bajo  $\mathbb{H}_0$  no hay dependencia estadística entre los parámetros seleccionados como retardos relevantes  $\tau^*$  y  $\tau_w$  y sus retardos adyacentes  $\tau^* - 1$  y  $\tau_w - 1$ , se han construido los 2 siguientes DGP's compatibles con la hipótesis nula y que ambos obedecen al modelo general

$$M(\varepsilon_t) : X_t = \varepsilon_t, \varepsilon_t \text{ i.i.d} \quad (4.62)$$

donde hemos usado para las innovaciones  $\varepsilon_t$  dos tipos de shocks estocásticos i.i.d que obedecen a las distribuciones *Normal* (media 0 y desviación típica la unidad) y *Uniforme* (distribución uniforme en el intervalo (0,1)). Estos dos modelos se describen en la Tabla 4.33.

La Tabla 4.35 reporta los tamaños empíricos de los modelos de reconstrucción para los retardos de inmersión  $\tau_d$  y  $\tau_w$  cuando se contrastan con los retardos  $\tau^* - 1$  y  $\tau_w - 1$  respectivamente bajo dos configuraciones de la  $\mathbb{H}_0$  para un nivel del 5%. Para este escenario, los métodos  $f_{r_e}$  y  $f_{r_e}^1$  se calcularon con los radios normalizados  $R^{opt}$  con el modelo de Lorenz de la Tabla 4.27. Podemos observar que el tamaño del test bajo todos los modelos de reconstrucción está bien controlado independientemente de la distribución y del tipo de retardo (retardo y ventana temporales).

En la Tabla 4.36 se recogen los resultados de tamaño de los modelos  $f_{r_e}$  y  $f_{r_e}^1$  calculados con los valores medios del radio normalizado  $\overline{R^{opt}}$  de la Tabla 4.34 mediante simulación de Monte Carlo para cada uno de los modelos de la Tabla 4.33. Como se puede ver en este caso, en ambos métodos y para el retardo  $\tau^*$  y ventana  $\tau_w$  temporales, los niveles de tamaño también permanecen cercanos al tamaño nominal del 5% independientemente

Tabla 4.34: Valor de  $\overline{R^{opt}}$  para los DGP's Normal  $N(0, 1)$  y Uniforme  $U(0, 1)$ .

Modelo	$f_{re}$	$f_{re}^1$
S.1	0.7	0.2
S.2	0.85	0.15

Tabla 4.35: Tamaño Empírico de varios Modelos de Selección.

	MI	C-C		C-C( $S_c$ )		$f_{re}$		$f_{re}^1$	
	$\tau^*$	$\tau^*$	$\tau_w$	$\tau_w$	$\tau^*$	$\tau_w$	$\tau^*$	$\tau_w$	
S.1	0.049	0.041	0.025	0.048	0.058	0.046	0.035	0.049	
S.2	0.044	0.054	0.035	0.035	0.044	0.04	0.045	0.051	

del tipo de distribución estadística.

#### 4.8.3.2. Pruebas de Potencia

A continuación examinamos la potencia del test con cada uno de los métodos de reconstrucción considerados (Información Mútua, Método C-C, Método C-C- $S_c$ ,  $f_{re}$  y  $f_{re}^1$ ), es decir, la probabilidad de rechazar  $\mathbb{H}_0$  cuando es falsa (i.e, cuando los retardos de inmersión son significativos como extremos respecto de sus vecinos  $\tau^* - 1$  y  $\tau_w - 1$ ). Para este fin añadimos ruido gaussiano a la serie temporal de Lorenz (4.34):

$$\dot{x} = -a(x - y) \quad (4.63)$$

$$\dot{y} = -xz + cx - y$$

$$\dot{z} = xy - bz$$

Tabla 4.36: Tamaño Empírico para los métodos  $f_{re}$  y  $f_{re}^1$  con los valores de  $\overline{R^{opt}}$  de la Tabla 4.34.

	$f_{re}$		$f_{re}^1$	
	$\tau^*$	$\tau_w$	$\tau^*$	$\tau_w$
S.1	0.048	0.06	0.041	0.051
S.2	0.034	0.055	0.055	0.049

Tabla 4.37: Modelos bajo la Hipótesis Alternativa.

DGP: $X_t = X_t^L + \eta\sigma\varepsilon_i$	Modelo	Nivel de ruido $\eta$	Innovaciones $\varepsilon_t$
P.1	$X_t = X_t^L + 0,1\sigma\varepsilon_i$	$\eta = 0,1$	$N(0, 1)$
P.2	$X_t = X_t^L + 0,3\sigma\varepsilon_i$	$\eta = 0,3$	$N(0, 1)$
P.3	$X_t = X_t^L + 0,5\sigma\varepsilon_i$	$\eta = 0,5$	$N(0, 1)$
P.4	$X_t = X_t^L + \sigma\varepsilon_i$	$\eta = 1$	$N(0, 1)$

con la configuración de parámetros descrita en el apartado 4.7,  $a = 16.0$ ,  $b = 16.0$  y  $c = 45.92$ . Así, construimos las series temporales de la forma que corresponden al modelo general

$$M(X_t, \varepsilon_t) : X_t = X_t^L + \eta\sigma\varepsilon_i \quad (4.64)$$

donde  $X_t^L$  es la serie temporal de Lorenz libre de ruido,  $\sigma$  es su desviación típica,  $\varepsilon_i$  es una variable aleatoria i.i.d gaussiana con media cero y desviación típica de 1, y  $\eta$  es el “nivel de fuerza” del ruido. Consideramos 4 procesos, P.1 a P.4, correspondientes a los niveles del ruido del 10 %, 30 %, 50 % y 100 % ( $\eta = 0,1, 0,3, 0,5$  y  $1,0$ ) que se describen en la Tabla 4.37.

Las Tabla 4.39 reporta las tasas de rechazo empírico de los métodos de selección de retardo y ventana temporales bajo los DGP's P.1-P.4 del modelo de Lorenz de 3000 puntos. Los métodos de selección  $f_{r_e}$  y  $f_{r_e}^1$  se computaron con los valores del radio normalizado  $R^{opt}$  calculados para la serie de Lorenz en condiciones libre de ruido e indicados en la Tabla 4.27.

Como podemos ver la potencia o precisión se degrada, como es de esperar, conforme aumenta el nivel de ruido en todos los métodos de selección; siendo más evidente este efecto en los dos retardos  $\tau^*$  y  $\tau_w$  para los métodos  $f_{r_e}$  y  $f_{r_e}^1$  y C-C. También se ve claramente que el método C-C( $S_c$ ) es el más robusto (respecto del retardo  $\tau_w$ ) contra el ruido de todos los métodos de selección ya que la potencia empírica sólo se degrada al 62 % cuando el nivel de ruido es del 100 % (P.4), permaneciendo en niveles del 100 % en el resto de los DGP's. El método MI (Información Mútua) como se puede advertir, es el más robusto frente al ruido en la selección del retardo  $\tau^*$ . Sin embargo observamos que en el método C-C, los valores de  $\tau^*$  y  $\tau_w$  no son significativos con respecto a  $\tau^* - 1$

y  $\tau_w - 1$  para ninguno de los DGP's; ni siquiera cuando el nivel de ruido es muy bajo (DGP P.1  $\eta = 0,1$ ). Los métodos  $f_{r_e}$  y  $f_{r_e}^1$  poseen tasas de rechazo aceptables con niveles de ruido por debajo del 50% aunque por encima del 50% la potencia se degrada considerablemente. Además el método  $f_{r_e}$  posee mejores resultados de potencia que el método  $f_{r_e}^1$  tanto para el retardo como la ventana temporales en todos los DGPs. Finalmente, a la vista de los resultados (a falta de probar por supuesto con  $\tau^* + 1$  y  $\tau_w + 1$ ) se puede concluir que la precisión obtenida es máxima cuando utilizamos el método MI para seleccionar el retardo  $\tau^*$  y el método C-C( $S_c$ ) cuando seleccionamos la ventana temporal  $\tau_w$ .

Por último, en la Tabla 4.40 se presentan las tasas de potencia de los métodos  $f_{r_e}$  y  $f_{r_e}^1$  calculados con los valores medios del radio  $\overline{R^{opt}}$  de la Tabla 4.38 mediante simulación de Monte Carlo para cada uno de los modelos de la Tabla 4.37. Se puede observar como las tasas de rechazo empírico para ambos tipos de retardo en general son similares a los mostrados en la Tabla 4.39 por debajo de niveles del ruido del 50%. Cuando el factor de ruido aumenta por encima del 50%, los resultados de potencia del método  $f_{r_e}$  se degradan de forma significativa para los dos tipos de retardo  $\tau^*$  y  $\tau_w$ .

Tabla 4.38: Valor del radio normalizado  $\overline{R_e^{opt}}$  para el Sistema Caóticos de Lorenz con varios niveles de ruido.

Modelo	$f_{re}$	$f_{re}^1$
P.1	1	0.25
P.2	0.95	0.3
P.3	0.9	0.3
P.4	0.85	0.2

Tabla 4.39: Potencia Empírica de varios modelos de reconstrucción.

	MI		C-C		C-C( $S_c$ )		$f_{re}$		$f_{re}^1$	
	$\tau^*$	$\tau^*$	$\tau_w$	$\tau_w$	$\tau^*$	$\tau_w$	$\tau^*$	$\tau_w$	$\tau^*$	$\tau_w$
P.1	0.941	0.03	0.058	1	1	1	0.89	1		
P.2	0.989	0.025	0.034	1	0.95	1	0.28	0.756		
P.3	0.921	0.03	0.055	1	0.735	0.89	0.095	0.28		
P.4	0.54	0.04	0.035	0.62	0.455	0.22	0.025	0.043		

Tabla 4.40: Potencia Empírica para los métodos  $f_{re}$  y  $f_{re}^1$  con los valores de  $\overline{R_e^{opt}}$  de la Tabla 4.34.

	$f_{re}$		$f_{re}^1$	
	$\tau_d$	$\tau_w$	$\tau_d$	$\tau_w$
P.1	1	1	0.86	1
P.2	1	1	0.32	0.854
P.3	0.96	0.72	0.077	0.31
P.4	0.53	0.05	0.055	0.097

## 4.9. Conclusiones

En este capítulo hemos propuesto dos aplicaciones de simbolización denominadas  $f_{r_e}$  y  $f_{r_e}^1$  para la selección del retardo  $\tau$  y ventana  $\tau_w$  temporales. Ambos tipos de retardo se utilizan en la reconstrucción del espacio de estados de un sistema caótico. En primer lugar se ha analizado su comportamiento con tres sistemas caóticos en condiciones libres de ruido y después se han comparado con los métodos de selección de retardos basados en la Integral de Correlación (Métodos C-C y C-C- $S_c$ ) así como el método de la Información Mútua de Fraser y Swinney [33]. Para verificar la idoneidad de la selección de ambos parámetros  $\tau$  y  $\tau_w$  se han calculado las cantidades LLE (Largest Lyapunov Exponent) y Dimensión de Correlación para cada uno de los métodos analizados. En la mayoría de los casos se ha comprobado que ambas simbolizaciones son competitivas con los métodos de reconstrucción inspeccionados.

En segundo lugar se ha analizado la precisión de los métodos de selección de los parámetros de reconstrucción  $\tau$  y  $\tau_w$  cuando se enfrentan a series caóticas temporales contaminadas con ruido. La precisión de cada método se evalúa a través de un test que contrasta la significatividad de los parámetros de reconstrucción seleccionados  $\tau$  y  $\tau_w$ . El test contrasta la significatividad del retardo  $\tau$  y la ventana  $\tau_w$  temporales en función de si cumplen las relaciones de orden del Funcional de Información específico (máximo o mínimo) de cada método de selección con el resto de retardos de sus entornos. Para un método de reconstrucción dado y un Funcional de Información que lo caracteriza, cuando los parámetros  $\tau$  y  $\tau_w$  son significativos con respecto a todos los retardos que pertenecen a sus entornos locales, el Funcional de Información adquiere la condición de extremo local significativo en  $\tau$  y  $\tau_w$  y por lo tanto se pueden elegir como los únicos retardos de reconstrucción significativos de ese entorno local. Cuando esto no se cumple, el error estándar aumenta, y por lo tanto el intervalo de definición de los retardos y consecuentemente se degrada la calidad de la reconstrucción.

Concretamente el contraste que se propone es un test bootstrap no paramétrico de una cola. Una característica importante del test es que no realiza suposiciones sobre el tipo

de funcional a contrastar, lo que garantiza su aplicabilidad a cualquier tipo de método de reconstrucción. Otra ventaja es su simplicidad computacional, pero delega el tiempo de computación en el tipo de funcional específico de cada método de reconstrucción. Hemos presentado los principales resultados del test en una simulación de Monte Carlo en la cual hemos incluido los métodos de selección de Información Mútua, C-C, C-C- $S_c$  y dos nuevos métodos propuestos basados en Dinámica Simbólica denominados  $f_{r_e}$  y  $f_{r_e}^1$ . Estos resultados nos permiten establecer que el test, tiene en general el tamaño empírico exacto frente a la hipótesis nula de no significatividad de los parámetros  $\tau$  y  $\tau_w$  frente a sus retardos vecinos. El test también posee buenos resultados de potencia/precisión en todos los métodos analizados para niveles bajos de ruido. Sin embargo las prestaciones de potencia se degradan cuando aumenta el nivel de ruido, si bien el método C-C- $S_c$  es el que menos se degrada proporcionalmente. Además se ha comprobado que los dos métodos basados en las simbolizaciones  $f_{r_e}$  y  $f_{r_e}^1$  poseen en general una precisión similar a los otros métodos de selección, con excepción del método C-C- $S_c$ , que es el que mejores resultados de precisión exhibe de todos los analizados.

## Capítulo 5

### Conclusiones y Líneas Futuras de Investigación

#### 5.1. Conclusiones

En esta Tesis hemos extendido las aplicaciones del análisis simbólico a algunas áreas científicas. Algunas de ellas, tales como las que hacen uso de datos no experimentales, necesitan realizar inferencia estadística bajo un mínimo nivel de supuestos. Concretamente, en el área de las finanzas, hemos utilizado el análisis de patrones para construir algunos tests de heterocedasticidad competitivos con los más populares. Además, en el campo del análisis de los sistemas caóticos, hemos propuesto el análisis simbólico como método de reconstrucción de los atractores caóticos a partir de las series temporales escalares.

En primer lugar se han obtenido resultados que contribuyen, de manera relevante al contraste de procesos heterocedásticos mediante el uso del análisis de patrones (simbólicos) como herramientas estadísticas. Para llevar a cabo tal fin, la primera contribución de esta tesis es la construcción de un nuevo test Omnibus Semiparamétrico de Heterocedasticidad llamado  $G_{h_{NQ}}(m)$ . Este test extrae las características heterocedásticas mediante la evaluación de la distribución de probabilidades asociadas a la serie de los residuos al cuadrado, asumiendo previamente que la media está correctamente especificada. Además el tamaño empírico no difiere del tamaño teórico, lo cual es una propiedad interesante para la aplicabilidad y reproducibilidad del test. Una de las ventajas

importantes del test es que a parte de ser menos exigente en términos de supuestos, es muy competitivo contra la mayoría de tests de heterocedasticidad. Este nuevo test supera a los otros tests disponibles cuando hay una estructura compleja no lineal en la varianza condicional desconocida del proceso. Otras ventajas competitivas del test es su simplicidad computacional y por lo tanto sus bajos tiempos de procesamiento computacional.

También se ha estudiado en esta tesis el comportamiento del test  $G_{h_{NQ}}(m)$  con respecto a los modelos de Volatilidad Estocástica, los cuales han atraído recientemente la atención de algunos investigadores dedicados a la modelización de datos de alta frecuencia como los encontrados en datos financieros intradía. Se ha comprobado empíricamente que el test  $G_{h_{NQ}}(m)$  no captura adecuadamente este tipo de modelos. Los modelos de Volatilidad Estocástica, por pertenecer al conjunto de sistemas caracterizados por interacciones de largo alcance, memoria larga o multifractalidad se describen de forma más adecuada mediante la Entropía no Extensiva de Tsallis, conocida como  $q$ -entropía que se caracteriza por el índice entrópico  $q$ . Por esta razón, la segunda contribución de esta Tesis consiste en utilizar la Entropía Simbólica de Tsallis en un nuevo test Omnibus Semiparamétrico Bootstrap denominado  $T_h^q(\tau, m)$  con el fin de cuantificar de forma más eficiente los procesos de Volatilidad Estocástica. El nuevo test  $T_h^q(\tau, m)$  utiliza la Entropía de Tsallis evaluada después de usar el análisis de patrones para extraer las características heterocedásticas de la distribución de probabilidades de la serie de residuos al cuadrado asumiendo que la media está correctamente especificada. Como principal ventaja, el test consigue capturar la gran mayoría de modelos heterocedásticos que el test  $G_{h_{NQ}}(m)$  no era capaz de detectar. Además el test  $T_h^q(\tau, m)$  sigue siendo consistente contra una amplia clase de procesos heterocedásticos mientras mantiene el tamaño empírico en niveles parecidos a los nominales. En este aspecto, el test es claramente más competitivo que su antecesor  $G_{h_{NQ}}(m)$  en relación a la mayoría de tests de heterocedasticidad disponibles.

La última aportación relevante de esta Tesis consiste en el desarrollo y propuesta de algunos métodos de reconstrucción basado en el análisis de patrones; así como de un

nuevo test estadístico para seleccionar los retardos de reconstrucción del atractor de un sistema caótico. En un entorno ideal libre de ruido, el retardo temporal  $\tau$  y la dimensión de inmersión  $p$  se obtienen a través de la inmersión, según el método de los retardos propuesto por Takens, de la serie temporal no lineal observada del sistema. Para estimar  $\tau$  se usan diversos métodos, entre ellos la función de autocorrelación, el método de la información mútua y el método C-C basado en la integral de correlación, así como sus derivados. Algunos investigadores han sugerido que uno no debería elegir un retardo temporal  $\tau$ , independiente de la dimensión de inmersión  $m$ , sino que uno debería elegir un valor apropiado de la ventana temporal  $\tau_w = (m - 1)\tau$ , el cual es el tiempo total abarcado por las componentes de cada punto inmerso. Además, con el fin de evaluar la utilidad del análisis de patrones en el desarrollo de métodos de selección de los parámetros de reconstrucción, en esta Tesis se proponen dos nuevos métodos llamados  $f_{re}$  y  $f_{re}^1$ . Estos dos métodos se han comprobado que son competitivos en entornos ideales libres de ruido, con algunos de los principales métodos de reconstrucción disponibles hasta la fecha. Por otra parte, no existe un método estándar para estimar  $\tau_w$  pero se puede usar el método C-C (y derivados de este método) para estimarlo. Todos estos métodos seleccionan los retardos  $\tau$  y  $\tau_w$  en base a localizar visualmente los extremos relativos de los funcionales que miden la Información contenida en la serie temporal caótica. Sin embargo las series caóticas reales están contaminadas con ruido, por lo que se necesita un criterio estadístico, no disponible en la actualidad, que permita fundamentar la selección de los retardos de reconstrucción  $\tau$  y  $\tau_w$  de cada método utilizado para estimarlos. Por esta razón en esta Tesis se propone un test bootstrap no paramétrico de una cola. Este nuevo test comprueba si el retardo elegido, en el cual se localiza un extremo local del funcional de información característico de cada método de selección, es estadísticamente mayor o menor (en función de los requerimientos impuestos por el extremo local) que cualquiera de los retardos de su entorno. Se ha comprobado que para todos los métodos de selección el tamaño empírico no difiere del teórico. En presencia de series caóticas con ruido, el test contrasta la significatividad de los retardos  $\tau$  y  $\tau_w$  de algunos métodos de reconstrucción respecto de sus vecinos

inmediatos.

## 5.2. Futuras líneas de Investigación

En esta sección se presentan algunas líneas de investigación que pueden ser objeto de interés teniendo en cuenta el trabajo expuesto en este tesis.

Como ya hemos señalado anteriormente, el contraste de heterocedasticidad es necesario para modelar la varianza de las series temporales, especialmente las financieras. En relación con los tests de heterocedasticidad que hemos propuesto, para poder realizar dicho contraste, uno de los supuestos necesarios asume que la media condicionada del modelo está correctamente especificada. En este sentido, desde el punto de vista del modelado de series temporales, es interesante también poder contrastar la naturaleza de la media condicional. Muchas teorías económicas en el contexto de series temporales tienen implicaciones solamente en la dinámica de la media condicional. Por ejemplo, la hipótesis de mercado eficiente o la de expectativas racionales, normalmente tienen implicaciones si y solo si en la dinámica de la media condicional de las variables económicas subyacentes. Así, la hipótesis de mercado eficiente establece que el rendimiento esperado de los activos dada la información disponible, es cero o al menos constante en el tiempo. Aunque la teoría económica puede sugerir una relación no lineal, no da una forma concreta funcional de la media condicional de las variables económicas. Por esto es importante contrastar la especificación en la media condicional, ya que una mala especificación en la media puede llevar a conclusiones erróneas sobre teorías económicas, hipótesis o pronósticos subóptimos.

En el contexto del modelado de series temporales es importante determinar primero si una serie temporal es lineal en media; es decir, si la media condicional de un proceso es una combinación lineal de las variables de un conjunto de Información. Así, podemos encontrarnos con un gran número de tests de linealidad en media. Éstos se pueden dividir en dos grupos, aquellos que se derivan sin tener en cuenta ninguna alternativa no lineal específica como por ejemplo el test de Tsay (1986) [77] o el test de White (1989) [85]. El otro grupo comprende aquellos tests que se construyen contra una alternativa

no lineal alternativa, como por ejemplo el test de Hamilton (2001) [41]. Sin embargo, en la práctica, la teoría económica con frecuencia no señala una única alternativa no lineal. Por ello, es deseable tener en estos escenarios un test Omnibus contra cualquier desviación de la linealidad.

Por otra parte, la teoría económica, al tiempo que tiene implicaciones sobre la dinámica de la media condicional, normalmente no dice nada sobre la dinámica de los momentos condicionales de orden superior como por ejemplo la varianza. Así, es importante desarrollar tests de modelos de media condicional que sean robustos a la heterocedasticidad. Cuando los tests no tienen la capacidad de ajustarse a la heterocedasticidad conducen a niveles inapropiados y dan lugar a conclusiones erróneas, por ejemplo, la gran mayoría de tests de no linealidad en media no son robustos a la heterocedasticidad, incluidos los mencionados anteriormente. Como ejemplo de test Omnibus de no linealidad en media robusto a la heterocedasticidad tenemos al test de Hong (2005) [47]. Sin embargo este tipo de test cuenta entre algunas de sus desventajas con asumir bastantes supuestos en lo que se refiere a las condiciones de convergencia asintótica.

Por este motivo, otra posible línea futura de trabajo podría consistir en desarrollar un nuevo test Omnibus que, basado en el análisis de patrones, permita contrastar la linealidad o no linealidad en la media condicional, siendo robusto al mismo tiempo a la forma de la varianza condicionada. Esta aportación sería especialmente interesante ya que los tests que se construyen en el dominio del *Análisis Simbólico*, como hemos podido comprobar en esta tesis, requieren un mínimo de supuestos, son simples, flexibles y de computación eficiente.

Otra posible e interesante línea de trabajo consiste en desarrollar un procedimiento para estimar de manera no paramétrica en series de longitud moderada, el índice entrópico  $q$ . Como hemos visto en el Capítulo tres, la estimación del parámetro  $q$ , al seguir permaneciendo como un problema abierto hasta la fecha, impide el cálculo del estadístico  $G_{h_{NQ}}^q(m)$ . También pudimos comprobar que cuando se estima correctamente el índice  $q$ , la potencia del test  $G_{h_{NQ}}^q(m)$  es cercana al 100 % para la práctica totalidad de modelos heterocedásticos contrastados. Es por esto que la propuesta de un método de

estimación del parámetro  $q$  podría significar una mejora significativa del rendimiento en potencia respecto del test  $G_{h_{NQ}}(m)$  original propuesto en el capítulo tres de este Tesis.

En el Capítulo tres también hemos visto que el cálculo del estadístico  $\hat{T}_h^q(\tau, m)$  consistía en encontrar los valores de los parámetros  $\tau$  y  $q$  que maximizaran dicho valor pero cercanos del tamaño nominal. Dicho procedimiento de maximización se lleva a cabo mediante un procedimiento de búsqueda de fuerza bruta. En este sentido sería conveniente proponer un algoritmo de optimización que encontrase los valores óptimos de  $\tau$  y  $q$  de una manera más eficiente que la propuesta en esta Tesis.

En el Capítulo Cuatro, vimos que los métodos de selección de los parámetros de reconstrucción de un atractor caótico,  $p$  (dimensión de inmersión) y  $\tau$  (retardo), se basan en aplicaciones de Simbolización que dependen de un parámetro, el radio "estadístico"  $r_e$ . Lógicamente la dependencia de un parámetro agravada por la presencia de ruido, degrada la precisión de los métodos de selección propuestos; haciéndolos más sensibles al ruido. Así, otra posible vía de investigación podría consistir en proponer un método de selección basado en una Aplicación de Simbolización libre de parámetros que ayude a escoger el retardo y la dimensión de inmersión de una manera más robusta frente al ruido. Por ejemplo, se podría proponer un nuevo funcional que promediara la Entropía de Símbolo para varios  $m$ .

Otra posible línea de investigación tiene que ver con el contraste estadístico de selección de los parámetros de reconstrucción (ventana y retardo de inmersión) de los atractores caóticos y que fue propuesto en el Capítulo cuatro. Como vimos, para que un retardo sea candidato a retardo de reconstrucción una de las condiciones que debe cumplir es la de extremo relativo local en sentido estricto del funcional de información del método de selección. El test contrasta la significatividad (respecto a la Información contenida) de los retardos de reconstrucción candidatos a extremo relativo con uno de los retardos pertenecientes al intervalo reducido con centro en el retardo candidato. Rechazar la hipótesis nula supone que el retardo  $\tau_f$  candidato a parámetro de reconstrucción es significativo (i.e cumple una relación de orden estricta) respecto a uno de los retardos

pertenecientes a su intervalo  $(\tau_f - 1, \tau_f + 1, \tau_f - 2, \tau_f + 2, \dots)$ . Pero, como hemos dicho, para poder seleccionar dicho retardo como candidato a parámetro de reconstrucción, éste debe cumplir la condición de extremo relativo local estricto. Para cumplir esta condición, el test de selección debe rechazar la hipótesis nula para todos los retardos pertenecientes al intervalo reducido con centro en el retardo candidato,  $(\tau_f - 1$  y  $\tau_f + 1)$  o  $(\tau_f - 2$  y  $\tau_f + 2, \dots)$ . Por esta razón, una propuesta de mejora sería desarrollar un test de selección cuyo rechazo de la hipótesis nula signifique cumplir la condición de extremo relativo local en sentido estricto. Esto implicaría contrastar simultáneamente respecto a todos los retardos del intervalo reducido con centro en el retardo seleccionado.

## Bibliografía

- [1] Abarbanel, H.D.I. & Brown, R.& Sidorowich, J.J. & Tsimring L.S (1993), *Rev. Mod. Phys.* 65, 1331.
- [2] Abe, S.& Suzuki, N. (2003): Itineration of the Internet over nonequilibrium stationary states in Tsallis statistics, *Phys. Rev. E* 67, 016106.
- [3] Abramov, Rafail V. (2009): The multidimensional moment-constrained maximum entropy problem: A BFGS algorithm with constraint scaling, *Journal of Computational Physics* 228:1, 96-108.
- [4] Albano, A. M. et al (1988): SVD and Grassberger–Procaccia algorithm, *Phys. Rev. A*, 38(6): 3017–3026
- [5] Anteneodo, C.& Tsallis, C. & Martínez, A.S. (2002), *Europhys. Lett.* 59, 635.
- [6] Amigó, J.M. (2010): *Permutation Complexity in Dynamical Systems*. Springer Series Synergetics, Berlin.
- [7] Arimitsu, T. & Arimitsu, N. (2002), *Physica A* 305, 218.
- [8] Bandt, C. & Pompe, B. (2002): Permutation entropy: a natural complexity measure for time series, *Phys. Rev. Lett.* 88: 174102
- [9] Bandt, C. & Shiha, F. (2007): Order patterns in time series, *Journal of Time Series Analysis*, 28, 646-665.
- [10] Beck, C. & Lewis, G. S. & Swinney, H. L. (2001), *Phys. Rev. E* 63, 035303.

- [11] Beck, C. (2001), *Phys. Rev. Lett.* 87, 180601.
- [12] Bediaga, I. & Curado, E. M. F. & Miranda, J. (2000), *Physica A* 286, 156.
- [13] Bera, A. K. & Ra. S. (1995): A Test for the Presence of Conditional Heteroscedasticity Within ARCH-M Framework. *Econometric Reviews*, 14, pp 473-485.
- [14] Bollerslev, T. (1986): Generalized Autorregressive conditional Heteroskedasticity, *Journal of Econometrics*, 31 pp 307-327.
- [15] Bollerslev, T. & Engle R. F. & Nelson D. B. (1994): "ARCH Models," in *Handbook of Econometrics*, Volume IV, ed. by R. F. Engle, and D. L. McFadden, chap. 49, pp. 2959–3038. North-Holland, Amsterdam.
- [16] Borland, L. & Plastino, A. R & Tsallis, C. (1998), *J. Math. Phys.* 39, 6490 [Errata: *J. Math. Phys.* 40, 2196 (1999)].
- [17] Borland, L. (2002), *Phys. Rev. Lett.* 89, 098701; *Quantitative Finance* 2, 415.
- [18] Breidt, F.J. & Crato, N. & De Lima, P. (1998): On the detection and estimation of long memory in stochastic volatility, *J. Econometrics* 83, pp 325-348.
- [19] Breusch, T. S. & Pagan, A. R. (1979): A Simple Test for Heteroscedasticity and Random Coefficient Variation, *Econometrica*, 47, pp 1287–1294.
- [20] Brock, W. A. & Hsieh & B. LeBaron (1991): *Nonlinear Dynamics, Chaos and Inestability : Statistical Theory and Economic Evidence*. MIT Press: Cambridge, M.A.
- [21] Brock, W. A. & Dechert W.D & Scheinkman J.A. & LeBaron B.,(1996). *Econ. Rev.* 15, 197.
- [22] Celedon, J.H.Q. (1997): *Maxima entropia aplicada a tomografia de tempos de transito*, Master Thesis (Universidade Federal da Bahia - Brazil, 1997).
- [23] Celluci, C.J & Albano A.M. & Rapp P.E (2003): Comparative Study of Embedding Methods, *Phys. Rev. E* 67, 66210.

- [24] Conaway, M.R. (1989): Analysis of repeated categorical measurements with conditional likelihood methods, *Journal of the American Statistical Association*, 84, pp 53-62.
- [25] Demos, A.& Sentana, A. (1998): Testing for GARCH Effects: A One-Sided Approach. *Journal of econometrics*, 86, pp 97-127.
- [26] Duarte Queirós, S. M. & Anteneodo Celia & Tsallis Constantino (2005): Power-law distributions in economics: a nonextensive statistical approach. *Noise and Fluctuations in Econophysics and Finance*, eds. D. Abbot, J.-P. Bouchaud, X. Gabaix and J.L. McCauley, *Proc. of SPIE 5848*, 151 (SPIE, Bellingham, WA, 2005)
- [27] Duarte Queirós, S.M (2008): On discrete stochastic processes with long-lasting time dependence in the variance, *The European Physical Journal B* 66, pp 137-148 (2008).
- [28] Duarte Queirós, S.M & Tsallis, C. (2008): On the connection between financial processes with stochastic volatility and non extensive statistical mechanics. *Eur. Phy. J. B* 66, pp 137-148 (2008).
- [29] Engle, R. F. (1982): Autorregresive Conditional Heteroskedasticity with estimates of the variance of U.K inflation, *Econometrica*, 50, pp 987-1008.
- [30] Engle, R. F. (1985): ARCH: Selected Readings, *Advanced Texts in Econometrics*, Oxford University Press, Oxford, U.K.
- [31] Engle & Hendry & Trumble (1985): Small-Sample Properties of ARCH Estimators and Test, *Canadian Journal of Economics*. 18, pp 66-93.
- [32] Francq, C. & Zakoidan, J.M. (2010): *GARCH Models*, Jonh Wiley and Sons, United Kingdom.
- [33] Fraser, A.M.& Swinney, H.L.(1986), *Phys. Rev.*, A 33, 1134.

- [34] Gell-Mann, M. & Tsallis, C. (2004): *Nonextensive Entropy: Interdisciplinary Applications*. Oxford University Press, Oxford.
- [35] Gencay, R. & Nicola G. (2006): *Crash of '87-was it expected? Aggregated market fears and long range dependence*.
- [36] Godfrey, L. G. (1978): *Testing for Multiplicative Heteroscedasticity*, *Journal of Econometrics*, 8, 227–236.
- [37] Godfrey, L. G. (1996): *Some Results on the Glejser and Koenker Tests of Heteroscedasticity*, *Journal of Econometrics*, 72, pp 275–299.
- [38] Goldfeld, S. M. & Quandt, R. (1965): *Some Tests for Heteroscedasticity*, *Journal of the American Statistical Association*, 60, pp 539–547.
- [39] Grassberger P. & Procaccia I. (1983), *Phys. Rev. Lett.* 50, 345.
- [40] Grassberger P. & Procaccia I. (1983), *Physica D* 7, 153.
- [41] Hamilton, J. D. (2001): *A parametric Approach to Flexible Nonlinear Inference*, *Econometrica*, 69, pp 537-573.
- [42] Harvey, A. C. (1976): *Estimating Regression Models with Multiplicative Heteroscedasticity*. *Econometrica*, 44, pp 461–465.
- [43] Harvey, A. C. (1998): *Long memory in stochastic volatility*. In: Knight, J., Satchell, S. (Eds.), *Forecasting volatility in financial markets*. Butterworth-Heinemann, London.
- [44] Harvey, A. Cy Streibel, M. (1998): *Testing for slowly changing level with special reference to stochastic volatility*. *J. Econometrics* 87, 167-189. In *Recent Developments in Time Series*, 2003 (Reproducido por P. Newbold y S.J. Leybourne). Edward Elgar Publishing.
- [45] Harvey, A. C. & Streibel, M. (1998): *Testing for a slowly changing level with special reference to stochastic volatility*, *J. Econometrics* 87, pp 167-189, In *Recent*

- Developments in Time Series, 2003 (Reproduced by P. Newbold and S. J. Leybourne). Edward Elgar Publishing. Hao, B.& Zheng, W: Applied Symbolic Dynamics and Chaos World Scientific, Singapore.
- [46] Hong, Y. & Shehadeh, R. D. (1999): A New Test for ARCH Effects and its Finite-Sample Performance, *Journal of Business and Economic Statistics*, 17, pp 91–108.
- [47] Hong, Y & Lee, Y (2005): Generalized Spectral Tests for Conditional Mean Models in Time Series with Conditional Heteroscedasticity of Unknown Form, *The Review of Economic Studies*, Vol. 72, No. 2 (Apr., 2005), pp. 499-541.
- [48] Kalimeri, M. & Papadimitrou, C. & Balasi, G.& Eftaxias, K. (2008): Dynamical Complexity detection in pre-seismic emissions using nonadditive Tsallis Entropy, *Physica A* 387 (2008) 1161-1172.
- [49] Keller, K. & Sinn, M. & Emonds, J. (2007): Time series from the ordinal viewpoint, *Stochastics and Dynamics*, 07:02, pp 247-272 .
- [50] Kim, H.S. & Eykholt R. & Salas, J.D. (1999) : Nonlinear dynamics, delay times, and embedding windows, *Physica D* 127.
- [51] Lee, J.H.H (1991): A Lagrange Multiplier Test for Garch Models. *Economic letters*, 37, pp 265-271.
- [52] Lee, J.H & King, M.L.(1993): A Locally Most Mean Powerful Based Score Test for ARCH and GARCH Regression Disturbances, *Journal of Business and Economic Statistics*, 11, 17-27, Correction 12 (1994),139.
- [53] Lehmann, E.L (1986): *Testing Statistical Hypothesis*. John Wiley & Sons, New York.
- [54] Li, Y. & Song, Y. & Li, C. (2010): Selection of Parameters for Phase Space Reconstruction of Chaotic Time Series, *Bio-Inspired Computing: Theories and Applications (BIC-TA)*, 2010 IEEE Fifth International Conference on, pp 30-33.

- [55] Lopez, F. & Matilla-Garcia, M. & Mur, J. M. & Ruiz, M (2010): Non-Parametric Spatial Independence Test Using Symbolic Entropy, *Regional Science and Urban Economics*, 40, pp 106-115.
- [56] Lorenz, E.N. & Atoms J. (1963). *Sci.* 20, 130.
- [57] Lu, X.Q. & Cao B. & Zeng, M. et al (2006) : An algorithm of selecting delay time in the mutual information method, *Chinese J. Comput. Physics* 23, pp 184–188.
- [58] Lunde, A. & A. G. Timmermann (2004): Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets, *Journal of Business & Economic Statistics*, 22 (3), pp 253-273.
- [59] Mackinnon, J.G. & White (1985): Some Heteroscedasticity-consistent Covariance Matrix Estimators with Improved Finite Sample Properties, *Journal of Econometrics*, 29, pp 305-325.
- [60] Matilla-García, M. (2007): A non-parametric test for independence based on symbolic dynamics, *Journal of Economic Dynamics and Control*, 31(12), pp 3889-3903.
- [61] Matilla-García, M. & Ruiz, M. (2008): A non-parametric independence test, *Journal of Econometrics* 144, pp 139-155.
- [62] Matilla-Garcia, M. & Ruiz, M. (2009): Detection of nonlinear structure in time series, *Economics Letters*, Volume 105, Issue 1, October 2009, pp 1–6.
- [63] Packard, N.H. & Crutchfield, J.P. & Farmer, J.D. & Shaw, R.S. (1980), *Phys. Rev. Lett.* 45 712.
- [64] Pagan, A. R. & Hall, A. D. (1983): Diagnostic Tests As Residual Analysis, *Econometric Reviews*, 2, 159–218, Comments and Reply, pp 219-254.
- [65] Palshikar, G. (2009): Simple algorithms for peak detection in time-series, *Proc. 1st Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence*.

- [66] Rodríguez, J. & Ruiz, E. (2005): A powerful Test for conditional heterocedasticity for financial time series with highly persistent volatilities, *Statistica Sinica* 15(2005), pp 505-525.
- [67] Rosenstein, M. T. & Collins, J. J. & De Luca, C. J. (1992): A practical method for calculating largest Lyapunov exponents from small data sets, *Physica D: Nonlinear Phenomena*. Volume 65, 1–2, 15 Mayo 1993, pp 117–134.
- [68] Rossler, O.E, (1976), *Phys. Lett. A* 57, 397.
- [69] Rothagi, V.K. (1976): *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, New York.
- [70] Ruiz, M. & Lopez, F. & Paez, A. (2010): Testing for spatial association of qualitative data using symbolic dynamics, *Journal of Geographical Systems*, 12(3), pp 281-309.
- [71] Ruiz, M. & Matilla-García, M. & García, J.A. & Susillo, J.L. & Romo A. & Gonzalez, A. & Ruiz, A. & Gayan, J. (2010): An Entropy Test for Single-Locus Genetic Association Analysis. *BMC Genetics*.
- [72] Small, M. (2005): *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*, World Scientific Series on Nonlinear Science Series A, Volume 52.
- [73] Stephenson, D.B. (2000): Use of the “Odds Ratio” for Diagnosing Forecast Skill, *Weather and Forecasting*, 15, pp 221-232.
- [74] Sun, Y. & Chan, K.L. & Krishnan, S.M. & Dutt, D.N. (1999). Tsallis multiscale entropy for the analysis of nonlinear dynamical behavior of ECG signals, in *Medical Diagnostic Techniques and Procedures*, eds. Megha Singh et al (Narosa Publishing House, London, 1999), p. 49.
- [75] Takens, F. (1981): *Detecting Strange Attractors in Turbulence*. *Dynamical Systems and Turbulence*. Springer, Berlin.

- [76] Tang, L. & Liang, J. (2011): C-C Method to Phase Space Reconstruction Based on Multivariate Time Series, Intelligent Control and Information Processing (ICICIP), 2011 2nd International Conference on.
- [77] Tsay, R. S. (1986): Nonlinearity Tests for Time Series, *Biometrika* 73, pp 461-466.
- [78] Taylor, S.J. (1986): *Modelling Financial Time Series*. Wiley, New York.
- [79] Tsallis, C. (1995): Nonextensive thermostatics and fractals, *Fractals* 3, 541.
- [80] Tsallis, C. (2009): *Introduction To Nonextensive Statistical Mechanics - Approaching A Complex World*. Brasil. Springer Science + Business Media.
- [81] Wang, Y. & Xu, W. (2006): The methods and performance of phase space reconstruction for the time series in Lorenz system, *J. Vibration Engrg.* 19, pp 277-282.
- [82] Wei-Dong Cai & Qin Yi-Qin & Yang Bing-Ru (2007) : Selection of Delay Time Window and Delay Time in Phase Space Reconstruction, IEEE Computer Society, International Conference on Computational Intelligence and Security. pp 526-530.
- [83] Wei-Dong Cai & Qin Yi-Qin & Yang Bing-Ru (2008) : Determination of Phase-Space Reconstruction Parameters of chaotic time series, *Kybernetika-Volumen* 44, numero 4, pp 557-570.
- [84] White, H. (1980): A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity, *Econometrica*, 48, 817-838.
- [85] White, H. (1989): An additional Hidden Unit Test for neglected Nonlinearity in Multilayer FeedForward Networks, *Proceedings of the International Joint Conference on Neural Networks* (New York, IEEE Press), pp 451-455.
- [86] Xiu, C. B. & Liu, X. D. & Zhang, Y. H. (2003): Selection of embedding dimension and delay time in the phase space reconstruction, *Trans. Beijing Institute of Technology* 23, pp 219-224.

- [87] Zeni, R.A. & Gallas J. A. C. (1995): Lyapunov Exponents for a Duffing Oscillator, *Physica D* 89, p. 79-85.
  
- [88] Zhang, Y. & Ren, C. L. (2005): The methods to confirm the dimension of reconstructed phase space, *J. National University of Defense Technology* 27 (2005), pp 101–105.
  
- [89] Zunino, L.& Pérez, D.G. & Kowalsky, M.T & Caravaglia, M. & Plastino, M. & Rosso O.A (2008): Fractional Brownian motion, fractional gaussian noise and Tsallis permutation Entropy, *Physical A* 387 (2008), pp 6057-6068.

# APÉNDICES

## Apéndice A

### Prueba de la Convergencia de $G^{q \rightarrow 1}(m) \sim \chi^2(k)$

En este apéndice se va a probar la convergencia del estadístico  $G^q(m)$ ,  $G_{q \rightarrow 1}(m) \sim \chi^2(k)$  bajo la hipótesis nula  $H_0$  de independencia (apartado 3.6.1), esto es, cuando el índice entrópico  $q \rightarrow 1$ . Esta relación es muy importante pues nos dice que los resultados de tamaño del test  $G(m)$  son asintóticamente equivalentes al del test  $G^q(m)$  cuando el índice entrópico  $q \rightarrow 1$ .

Partiendo de la expresión 3.53 :

$$\begin{aligned}
 G^q(m) &= -2Ln_q(\lambda_q(Y)) = -2 \left( \sum_{i=1}^{n_{\sigma_1}} Ln_q(p_{\pi_1}) + \dots + \sum_{i=1}^{n_{\sigma_n}} Ln_q(p_{\pi_n}) - \dots \right. \\
 &\quad \left. \dots - \left( \sum_{i=1}^{n_{\sigma_1}} Ln_q(\hat{p}_{\pi_1}) + \dots + \sum_{i=1}^{n_{\sigma_n}} Ln_q(\hat{p}_{\pi_n}) \right) \right) \quad (A.1)
 \end{aligned}$$

con

$$p_{\pi_i} = \frac{n_{\pi_i} + \delta_i}{R} \text{ y } \hat{p}_{\pi_i} = \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \text{ de 3.50} \quad (A.2)$$

suponiendo además que  $\sum_{i=1}^n \delta_i = 0$  y desarrollando la expresión del  $\ln_q$  que se vio en 3.2.2

$$\ln_q x := \frac{x^{1-q} - 1}{1 - q} (x \geq 0, q \in \mathbb{R}) \quad (A.3)$$

$$\begin{aligned}
G_q(m) &= -2 \left( \sum_{i=1}^n n_{\pi_i} \left( \frac{p_{\pi_i}^{1-q} - 1}{1-q} \right) - \sum_{i=1}^{n_{\sigma_1}} n_{\sigma_i} \left( \frac{\hat{p}_{\pi_i}^{1-q} - 1}{1-q} \right) \right) \\
&= -2 \left( \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( \frac{n_{\pi_i} + \delta_i}{R} \right)^{1-q} - 1}{1-q} \right) - \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) \right) = \\
&= -2 \left( \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( 1 + \frac{R\delta_i}{n_{\pi_i}} \right)^{1-q} - 1}{1-q} \right) - \sum_{i=1}^n n_{\sigma_i} \left( \frac{\left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) \right) \quad (\text{A.4})
\end{aligned}$$

Si ahora hacemos un desarrollo de Taylor alrededor de  $\frac{\delta_i}{n_{\pi_i}} = 0$  (el punto en que  $p_{\pi_i}$  y  $\hat{p}_{\pi_i}$  coinciden) y usando que  $(1+x)^n \approx nx + \frac{n(n-1)}{2}x^2 + O(x^3)$  tenemos:

$$\begin{aligned}
G_q(m) &= -2 \left[ \sum_{i=1}^n n_{\pi_i} \left( \frac{(1-q) \left( \frac{\delta}{n_{\pi_i}} \right) - \frac{q(1-q)}{2} \left( \frac{\delta}{n_{\pi_i}} \right)^2 - 1}{1-q} \right) - \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( \frac{n_{\pi_i}^{\frac{1}{q}}}{\sum_{i=1}^n n_{\pi_i}^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) \right] \Rightarrow \\
&\Rightarrow G_q(m) = -2 \left[ \sum_{i=1}^n n_{\pi_i} \left( \left( \frac{\delta_i}{n_{\pi_i}} \right) - \frac{q}{2} \left( \frac{\delta_i}{n_{\pi_i}} \right)^2 - 1 \right) - \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) \right] \Rightarrow \\
&\Rightarrow G_q(m) = -2 \left[ \sum_{i=1}^n \left( R(\delta_i) - \frac{q n_{\pi_i}}{2} \left( \frac{\delta_i}{n_{\pi_i}} \right)^2 - n_{\pi_i} \right) - \sum_{i=1}^n n_{\pi_i} \left( \frac{\left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) \right] \Rightarrow \\
&\Rightarrow G_q(m) = -2 \sum_{i=1}^n R\delta_i + \sum_{i=1}^n \frac{qR^2}{n_{\pi_i}} (\delta_i)^2 + 2R + 2 \sum_{i=1}^n n_{\sigma_i} \left( \frac{\left( \frac{(n_{\pi_i})^{\frac{1}{q}}}{\sum_{i=1}^n (n_{\pi_i})^{\frac{1}{q}}} \right)^{1-q} - 1}{1-q} \right) + O(\delta_i^3) \simeq
\end{aligned}$$

$$G_q(m) \approx \sum_{i=1}^n \frac{qR^2}{n_{\sigma_i}} (\delta_i)^2 \Rightarrow \lim_{q \rightarrow 1} G^q(m) \simeq \sum_{i=1}^n \frac{R^2}{n_{\pi_i}} (\delta_i)^2 \approx \chi^2(k) \quad (\text{A.5})$$

como queríamos probar.