

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA**

**DEPARTAMENTO DE MATEMÁTICAS**



**TESIS DOCTORAL**

**ANÁLISIS ESPECTRAL CON SERIES REPLICADAS:  
ESTIMACIÓN DEL ESPECTRO POBLACIONAL**

**CARMEN NIEVES HERNÁNDEZ FLORES**

Las Palmas de Gran Canaria, 1997

24/1996-97

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA  
UNIDAD DE TERCER CICLO Y POSTGRADO**

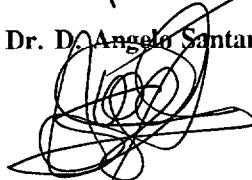
Reunido el día de la fecha, el Tribunal nombrado por el Excmo. Sr. Rector Magfco. de esta Universidad, el/a aspirante expuso esta TESIS DOCTORAL.

Terminada la lectura y contestadas por el/a Doctorando/a las objeciones formuladas por los señores miembros del Tribunal, éste calificó dicho trabajo con la nota de APTO COM LAUDE POR UNANIMIDAD Las Palmas de Gran Canaria a 6 de Marzo de 1997.

El/a Presidente/a: Dr. D. Manuel del Río Bueno,



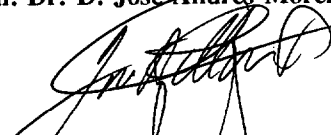
El/a Secretario/a: Dr. D. Angelo Santana del Pino,



El/a Vocal: Dr. D. José Miguel Pacheco Castela,



El/a Vocal: Dr. D. José Andrés Moreno Pérez,



El/a Vocal: Dr. D. Francisco José Vázquez Polo,



El/a Doctorando/a: D<sup>a</sup>. Carmen Nieves Hernández Flores,



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DEPARTAMENTO DE MATEMÁTICAS



TESIS DOCTORAL

ANÁLISIS ESPECTRAL CON SERIES REPLICADAS:  
Estimación del espectro poblacional.

Carmen Nieves Hernández Flores

1997

ANÁLISIS ESPECTRAL CON SERIES REPLICADAS:  
Estimación del espectro poblacional.

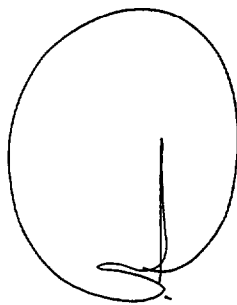
Carmen Nieves Hernández Flores

*UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA, 1997*

PEDRO SAAVEDRA SANTANA, Catedrático de Universidad del Área de Conocimiento de Estadística e Investigación Operativa del Departamento de Matemáticas de la Universidad de Las Palmas de Gran Canaria,

**CERTIFICA:** Que la presente memoria titulada *Análisis espectral con series replicadas: estimación del espectro poblacional* ha sido realizada bajo mi dirección por la Licenciada en Matemáticas (Especialidad de Estadística e Investigación Operativa) Doña Carmen Nieves Hernández Flores, y constituye su Tesis para optar al grado de Doctora en Matemáticas.

Y para que así conste, en cumplimiento de la legislación vigente y a los efectos a que haya lugar, firmo la presente en Las Palmas de Gran Canaria, a dieciseis de enero de mil novecientos noventa y siete.

A handwritten signature in black ink, enclosed within a hand-drawn oval. The signature is stylized and appears to be the name 'Pedro Saavedra Santana'.

Es mi deseo expresar al profesor Dr. D. Pedro Saavedra Santana mi más sincera gratitud por la entrega, entusiasmo y tiempo que ha invertido en la dirección de esta memoria. Sus valiosos consejos, aportaciones y críticas así como sus palabras de ánimo han hecho posible la consecución de este trabajo.

Este agradecimiento es extensivo a mis compañeros del departamento de Matemáticas, en especial al profesor J. Artiles por sus valiosos comentarios y buen humor y a los profesores J. Rocha, M.P. Quintana y J.J. González por su apoyo y ayuda tan entusiasta; y a todos aquellos que han deseado que este proyecto alcanzara su objetivo. También quiero dar las gracias a mi familia y en especial a Moreno por su comprensión, apoyo y empeño.

# ÍNDICE

## PRÓLOGO/iii

### CAPÍTULO I. ANÁLISIS ESPECTRAL.

- 1.1 Introducción./1
- 1.2 Procesos estocásticos estacionarios./2
- 1.3 Concepto de espectro: distribución espectral./4
- 1.4 Representación Espectral de un Procesos Estacionarios./9
- 1.5 Procesos lineales generales./12
- 1.6 Estimación espectral: análisis del periodograma./13
- 1.7 Estimación de la densidad espectral mediante alisamientos del periodograma./18
- 1.8 El problema de la elección del ancho de banda óptimo./21

### CAPITULO II. MÉTODO BOOTSTRAP EN EL ANÁLISIS ESPECTRAL.

- 2.1 Introducción./26
- 2.2 Validez asintótica del bootstrap./28
- 2.3 Estimación de la función de densidad espectral utilizando la metodología bootstrap./41
  - 2.3.1 Método Bootstrap para la estimación de la densidad espectral./41
  - 2.3.2 Validez asintótica del método./43

### CAPITULO III. SERIES TEMPORALES REPLICADAS: ANÁLISIS DEL PERIODOGRAMA MEDIO.

- 3.1 Introducción./47
- 3.2 El modelo de efectos fijos./48
- 3.3 El modelo de efectos aleatorios./50
- 3.4 Procesos lineales de coeficientes aleatorios./52
- 3.5 El modelo de ruido blanco./53

- 3.6 Modelo general de efectos aleatorios./56
- 3.7 Estimación Bootstrap del espectro poblacional./59
- 3.8 Simulaciones./64

## **CAPITULO IV. ESTIMACIÓN DEL ESPECTRO POBLACIONAL MEDIANTE ALISAMIENTO DEL PERIODOGRAMA MEDIO.**

- 4.1 Introducción./67
- 4.2 Alisamientos mediante núcleo del periodograma medio./68
- 4.3 Aproximación Bootstrap para la estimación del espectro poblacional utilizando como estimador alisamientos del periodograma medio./76
- 4.4 Ancho de banda óptimo./77
- 4.5 Estimación del ancho de banda óptimo./80
- 4.6 Simulaciones./82

**APÉNDICE./88**

**BIBLIOGRAFÍA./104**



## PRÓLOGO.

El análisis espectral es una metodología ampliamente desarrollada debido fundamentalmente a sus aplicaciones en el campo de la física y particularmente en el estudio de la señal. Después de la invención de la radio, a comienzos del siglo XX, los ingenieros electrónicos reconocieron el notable interés de modelizar la señal a través de procesos estocásticos estacionarios. Wiener (1948) formuló la teoría de las señales aleatorias y la aplicó a la resolución de problemas de filtrajes. Shannon (1948) utilizó los modelos de procesos aleatorios para formular una teoría que ha sido la base de la comunicación digital. La invención del radar durante la segunda guerra mundial condujo al desarrollo de nuevos algoritmos para la detección de señales débiles. El algoritmo más significativo para la localización de una posición fue desarrollado por Kalman en 1960. Dicho algoritmo se utiliza en sistemas de navegación para la exploración del espacio profundo.

El problema inferencial de estimación de una señal modelizada mediante un proceso estocástico se basa generalmente en la observación de una única trayectoria del proceso. Sus propiedades ergódicas permiten estimar características tales como su media o función de autocovarianza, en base a la única trayectoria observada. La excelente monografía de Priestley (1981) recoge los resultados más importantes acerca de la estimación de la función de densidad espectral de un proceso lineal. Dicha estimación se basa en alisamientos del periodograma mediante estimadores de núcleo. Beltrao y Bloomfield (1987), estudian de forma más precisa el problema de la selección del ancho de banda para los referidos estimadores. Franke y Härdle (1992) utilizan el bootstrap para la estimación de la densidad espectral y estudian la validez del método de acuerdo con el criterio de Bickel y Freedman (1981).

Las aplicaciones del análisis espectral en el campo de la economía se realizan también mediante la observación de una única serie temporal. Sin embargo, el escaso número de observaciones de la serie en este campo, hace poco viable la estimación de la densidad espectral mediante suavizamientos del periodograma correspondiente.

En las ciencias biomédicas, sin embargo, la replicación de experimentos es más la regla que la excepción. El registro de un electroencefalograma o la oscilación de concentración de una hormona, pueden también modelizarse a través de procesos estacionarios. La observación de diversas trayectorias del proceso abre nuevas perspectivas en el campo de la estimación espectral. Diggle y Al-Wasel (1993) estudiaron la concentración de hormona LH en una muestra de ocho sujetos. Supusieron inicialmente que las series de mediciones correspondientes a cada uno de los sujetos eran trayectorias de un mismo proceso estacionario con distribución espectral absolutamente continua. En este contexto, analizaron las propiedades del periodograma medio como estimador de la función de densidad espectral. Ellos observaron no obstante, que la hipótesis de una misma densidad espectral para todos los individuos, era inconsistente con los datos. Por este motivo, propusieron un modelo alternativo, en el cual cada individuo posee una densidad espectral específica. Dado que la existencia de densidad espectral supone que el correspondiente proceso sea lineal, la selección de un sujeto es equivalente a la selección de la sucesión de coeficientes del proceso. Por esta razón, podemos considerar que el modelo de Diggle y Al-Wasel es un *proceso estacionario lineal de coeficientes aleatorios*. Por otro lado, la selección de un individuo supone también la selección aleatoria de una función de densidad espectral. De esta forma, dicha densidad espectral puede considerarse a su vez un proceso estocástico. Diggle y Al-Wasel llaman a la media de este proceso *espectro poblacional* y proponen una metodología de estimación basada en las series temporales replicadas.

A nuestro juicio, esta metodología se fundamenta en hipótesis poco verosímiles. El modelo que proponen está basado en la teoría asintótica del periodograma dada por Hartigan (1980), y tiene la forma  $I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj}$ , donde  $I_l(\omega_j)$  representa el  $l$ -ésimo periodograma,  $f(\omega_j)$  el espectro poblacional,  $Z_l(\omega_j)$  un proceso tal que  $f(\omega_j)Z_l(\omega_j)$  es la  $l$ -ésima función de densidad espectral y  $U_{lj}$  son las variables residuales. A este modelo lo denominaron modelo de ruido blanco, posiblemente por la hipótesis de que las variables  $Z(\omega)$  son independientes y con la misma distribución de probabilidad. Esta hipótesis no es muy plausible, pues si el espectro poblacional es una función *suave*, no resulta difícil imaginar el aspecto de cada densidad espectral específica ( $f_l(\omega) = f(\omega)Z_l(\omega)$ ). Ellos admiten la poca consistencia de esta suposición y sugieren formas de distribución alternativas para el proceso  $Z(\omega)$ . Por otra parte, parametrizan el modelo con el objetivo de obtener una expresión computacionalmente manejable del estimador de máxima verosimilitud del espectro poblacional. El único fundamento de estas hipótesis es precisamente la simplicidad computacional.

Es bien sabido que el periodograma correspondiente a una serie temporal es un estimador insatisfactorio de la función de densidad espectral, dada su inconsistencia. Los alisamientos del periodograma que revisamos en el primer capítulo, y la aproximación bootstrap de su distribución en el segundo capítulo, parecen estimaciones adecuadas de la densidad espectral. Sin embargo, en el modelo multiplicativo de efectos aleatorios, el periodograma medio resulta ser consistente. En esta memoria hemos desarrollado una teoría asintótica para el periodograma medio, utilizando el modelo multiplicativo de efectos aleatorios, pero sin suponer que el proceso  $\{Z(\omega)\}$  sea un ruido blanco, y por supuesto, sin hacer parametrizaciones injustificadas del mismo. De esta teoría podemos determinar las correspondientes regiones de confianza para el espectro poblacional. Para aquellos casos en los que no consideremos utilizar la teoría asintótica, hemos propuesto un algoritmo bootstrap para la obtención de los intervalos de confianza. Dicho algoritmo puede considerarse como un análogo del

utilizado por Freedman (1981) para los modelos de regresión lineal. Un aspecto importante en el uso del bootstrap, es la validez asintótica del mismo en el sentido de Bickel y Freedman (1981). En el capítulo segundo hemos intentado realizar una revisión de este problema. La idea fundamental consiste en considerar una métrica adecuada entre distribuciones de probabilidad, de tal forma, que la cantidad pivotal en la que se basa la construcción del intervalo de confianza, esté próxima en el sentido de la métrica considerada de su aproximación bootstrap. En el capítulo tercero, hemos considerado un pivote natural basado en el periodograma medio para la determinación de los intervalos de confianza del espectro poblacional y su aproximación bootstrap, y hemos probado la validez del método en el referido sentido de Bickel y Freedman.

Sin duda, el periodograma medio puede ser un adecuado estimador del espectro poblacional para cada frecuencia fija, pero su aspecto global puede llevar a valoraciones erróneas acerca de la importancia de determinadas bandas de frecuencias. Ello se debe a la naturaleza de cada uno de los periodogramas individuales, pues para cada objeto fijo, el aumento del número de observaciones de la serie supone que las variables residuales  $U_{ij}$  tiendan a ser incorreladas. Priestley utiliza para el periodograma la siguiente expresión: *the periodogram has an erratic and wildly fluctuating form*. El periodograma medio suaviza este comportamiento, pero para muestras pequeñas este suavizamiento es claramente insuficiente. Esto nos conduce en el cuarto capítulo a la consideración de estimadores alternativos del espectro poblacional. Un estimador natural es sin duda el obtenido mediante suavizamientos del periodograma medio.

Uno de los problemas clave de los alisamientos, tanto en el contexto de la regresión como en el de la estimación de la función de densidad de probabilidad o la función de densidad espectral, es la elección de un adecuado parámetro de alisamiento o ancho de banda  $h$  (bandwidth). En el contexto de la estimación de la función de densidad espectral en base a la observación de una única serie temporal del proceso correspondiente, utilizando suavizamientos del periodograma, la

magnitud del parámetro de alisamiento  $h$  debe depender del número  $N$  de observaciones de la serie temporal. Franke y Härdle (1992) obtienen las propiedades asintóticas del estimador para el caso en el que  $h \rightarrow 0$  y  $Nh^2 \rightarrow \infty$ . En realidad, el orden del  $h$  que minimiza el error cuadrático medio es  $N^{-1/5}$ . La replicación de series introduce obviamente el parámetro  $r$  (número de objetos). Ello plantea un nuevo problema; a saber; ¿cuál debe ser en este caso el orden del parámetro de alisamiento?; o más precisamente: ¿cómo debe influir  $r$  en el ancho de banda?. Téngase en cuenta que el modelo que hemos considerado no es un proceso lineal (con densidad espectral) del que se observan  $r$  series. Es un proceso lineal de coeficientes aleatorios, de tal forma que la selección de cada objeto supone la selección de la sucesión de coeficientes, lo cual supone una nueva fuente de variabilidad.

Hemos obtenido propiedades asintóticas para el estimador del espectro poblacional definido mediante alisamientos del periodograma medio y bajo determinadas condiciones de los parámetros  $h$ ,  $N$  y  $r$ . Destacamos la descomposición de la varianza en una componente *intraindividual* y otra componente corresponde a los efectos aleatorios del modelo. De este resultado obtenemos, bajo condiciones precisas, que el orden del parámetro de alisamiento óptimo es  $(Nr)^{-1/5}$ .

Hemos introducido también el bootstrap para aproximar la distribución del estimador definido mediante suavizamientos del periodograma medio, y hemos comparado las correspondientes bandas de confianza con las obtenidas a través de los resultados asintóticos.

## CAPITULO I. ANÁLISIS ESPECTRAL.

### 1.1 Introducción.

El análisis espectral es una metodología desarrollada originalmente en el campo de la física y particularmente del análisis de la señal. Su objetivo esencial es el estudio del espectro de un proceso estacionario, es decir, de un proceso que tiene determinadas propiedades constantes a lo largo del tiempo. Uno de los resultados esenciales de los procesos estocásticos estacionarios es el llamado teorema de la representación espectral. Este resultado afirma que bajo determinadas condiciones, un proceso estacionario puede expresarse como una cierta superposición de movimientos armónicos de diferentes frecuencias. Cada uno de los armónicos aporta una determinada contribución a la varianza del proceso. Esta varianza en problemas físicos puede identificarse con conceptos de energía. De forma imprecisa, el espectro es la distribución que a las bandas de frecuencias le asigna su contribución a la varianza total del proceso.

El presente capítulo se inicia con los conceptos de proceso estocástico y proceso estocástico estacionario, y a continuación se hace una aproximación heurística al concepto de función de densidad espectral. Ésta es insuficiente para determinar, en un sentido amplio, el espectro de un proceso, pues, al igual que una distribución de probabilidad, un proceso estacionario no tiene necesariamente densidad espectral. Este concepto de distribución espectral sólo puede darse de forma general a partir del teorema de representación espectral. Un resultado fundamental para los objetivos de esta memoria, consiste en dar una condición necesaria y suficiente para que un proceso estacionario, tanto discreto como continuo, tenga función de densidad espectral. Tal condición consiste en que el proceso estacionario sea un proceso lineal, cuyos coeficientes decrezcan a cero de una forma determinada.

Abordamos también el problema de la estimación de la función de densidad espectral sobre la base de la observación de una única trayectoria del proceso o serie temporal. El periodograma es la herramienta natural para la estimación de la densidad espectral. Sin embargo, como señala Priestley (1981), el periodograma es un estimador *pobre* de la densidad espectral, pues aunque es asintóticamente insesgado, es inconsistente. Las técnicas de alisamiento del periodograma introducidas por Grenander y Rosenblatt (1953), conducen a los llamados estimadores de núcleo. Estos dan soluciones satisfactorias, pues aunque tienen un sesgo mayor que el periodograma, bajo determinadas condiciones son consistentes. Revisamos el problema de la elección del ancho de banda óptimo. Desde nuestro punto de vista, Priestley, no aborda de forma muy precisa este problema. Beltrao y Bloomfield (1987) aportan métodos más precisos para la selección del parámetro de alisamiento óptimo. Finalmente se dan las propiedades asintóticas del estimador de núcleo de la densidad espectral, en función del número de observaciones de la serie temporal y del ancho de banda considerado.

## 1.2 Procesos estocásticos estacionarios.

Un proceso estocástico es una colección de variables aleatorias  $\{X(t)\}_{t \in T}$ , definidas sobre un mismo espacio de probabilidades  $(\Omega, \mathcal{A}, P)$ . El conjunto  $T$ , llamado espacio de parámetros, generalmente representa el tiempo y por tanto sería un subconjunto de  $R$ , pero puede representar otros espacios, tal como ocurre con los procesos espaciales. Por su parte, el espacio probabilístico  $(\Omega, \mathcal{A}, P)$  representa normalmente un cierto fenómeno aleatorio que evoluciona con el tiempo y, por tanto,  $X(t)$  mide alguna característica del fenómeno en el instante  $t$ .

De acuerdo con esta definición, a cada  $\omega \in \Omega$  fijo, el proceso le asocia una función real definida sobre  $T$ , denominada trayectoria del proceso. Así podemos

considerar un proceso estocástico como una función de dos variables,  $X: T \times \Omega \longrightarrow R$ , tal que:

- Para cada  $t$  fijo,  $X(t, \cdot): \Omega \longrightarrow R$  es una variable aleatoria.
- Para cada  $\omega$  fijo,  $X(\cdot, \omega): T \longrightarrow R$  es una trayectoria del proceso.

Si llamamos  $R^T$  al espacio de todas las funciones de  $T$  en  $R$ , también podemos interpretar un proceso estocástico como una aplicación  $X: \Omega \longrightarrow R^T$ , tal que a cada  $\omega \in \Omega$  le hace corresponder la función  $X(\cdot, \omega): T \longrightarrow R$ .

Entre las condiciones de regularidad que se le pueden imponer a un proceso estocástico para hacerlo a la vez más útil y manejable, ocupa un lugar importante la condición de estacionariedad. De hecho, en muchas situaciones aparecen sistemas que evolucionan bajo el efecto de influencias, que aún siendo aleatorias, se conservan más o menos invariantes a lo largo del tiempo. A esta idea intuitiva responde el concepto de proceso estacionario y para expresarlo en términos concretos habrá que imponer que las leyes de probabilidad que rigen la evolución sean constantes en el tiempo.

Un proceso estocástico  $\{X(t)\}_{t \in R}$  se dice estrictamente estacionario si  $\forall n \in N, \forall t_1, \dots, t_n \in R$  y  $\forall h > 0$  las variables  $(X_{t_1}, \dots, X_{t_n})$  y  $(X_{t_1+h}, \dots, X_{t_n+h})$  tienen la misma distribución.

En particular todas las variables  $X_t$  tendrán la misma distribución, y por tanto la misma media  $E[X_t] = \mu$ , si es que ésta existe. También, las distribuciones conjuntas  $(X_s, X_t)$  han de depender sólo de la diferencia  $t-s$ , con lo cual la función de autocovarianza  $R(s, t) = \text{Cov}(X_s, X_t)$ , si existe, será sólo función de  $t-s$ . Estas propiedades de los procesos estacionarios conducen a definir los procesos estacionarios en sentido débil.



Un proceso  $\{X(t)\}_{t \in T}$  es estacionario en sentido débil si:

1.  $E[|X(t)|^2] < \infty \quad \forall t \in T$
2.  $E[X(t)] = \mu, \quad \forall t \in T$
3.  $Cov[X(t), X(t+\tau)] = R(\tau)$  (independiente de  $t$ )

La función de autocovarianza  $R(\tau)$  tiene las siguientes propiedades:

1.  $R(0) = \sigma^2$ .
2.  $|R(\tau)| \leq R(0)$  para cualquier  $\tau$ .
3. Cuando  $\{X(t)\}$  es un proceso que toma valores reales,  $\forall \tau$ ,  $R(-\tau) = R(\tau)$ , es decir  $R(\tau)$  es una función par.

### 1.3 Concepto de espectro: distribución espectral.

El concepto de espectro está basado en el hecho de que un proceso estacionario es una superposición de movimientos armónicos de diferentes frecuencias y amplitudes. Esta idea la concretaremos posteriormente en el teorema de representación espectral. Cada uno de los referidos movimientos armónicos aporta una *contribución* a la varianza del proceso, varianza que es constante a lo largo del tiempo. La idea de espectro puede asociarse a la distribución que evalúa la aportación de un determinado conjunto de frecuencias a la varianza del proceso. Precisaremos estas ideas mediante el siguiente desarrollo.

Consideremos una trayectoria  $X(t)$  de un proceso estacionario en tiempo continuo, la cual suponemos que es una función continua. En general no podremos suponer que

$$\int_{-\infty}^{\infty} |X(t)| dt < \infty$$

Sin embargo, a partir de  $X(t)$  podemos definir:

$$X_{\tau}(t) = \begin{cases} X(t) & |t| \leq \tau \\ 0 & |t| > \tau \end{cases} \quad (1.3.1)$$

Esta función es absolutamente integrable, por lo que se podría expresar como la transformada de Fourier de una cierta función  $G_{\tau}$ , como

$$X_{\tau}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\tau}(\omega) e^{i\omega t} d\omega \quad (1.3.2)$$

donde

$$G_{\tau}(\omega) = \int_{-\infty}^{\infty} X_{\tau}(t) e^{-i\omega t} dt = \int_{-\tau}^{\tau} X(t) e^{-i\omega t} dt \quad (1.3.3)$$

Considerando la relación de Parseval

$$\int_{-\infty}^{\infty} X_{\tau}^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\tau}^2(\omega) d\omega$$

el primer término representa la energía total de  $X_{\tau}(t)$ , utilizando la analogía de considerar  $X_{\tau}(t)$  como la intensidad de corriente eléctrica que pasa a través de una resistencia unitaria durante el intervalo de tiempo  $(-\infty, \infty)$ . De aquí que se hable del concepto de *energía* y se interprete  $|G_{\tau}(\omega)|^2 d\omega$  como la contribución a la energía total correspondiente a las frecuencias que están entre  $\omega$  y  $\omega+d\omega$ .

Puede observarse que si  $\tau$  tiende a infinito,  $X(t)$  y  $X_{\tau}(t)$  toman igual valor para todo  $t$ , y de esta forma  $\left\{ \lim_{\tau \rightarrow \infty} |G_{\tau}(\omega)|^2 \right\}$  describe las propiedades de  $X(t)$ .

En general este límite no tiene por qué existir, ya que  $X(t)$  en general no es absolutamente integrable. Además, la energía disipada por un proceso estocástico en  $(-\infty, \infty)$  es infinita, análogamente a lo que le ocurre a las funciones periódicas. Para resolver este problema introducimos el concepto de potencia como *energía por unidad de tiempo* y esto hace que el límite

$$\lim_{\tau \rightarrow \infty} \frac{|G_{\tau}(\omega)|^2}{2\tau} \quad (1.3.4)$$

pueda ser finito para cualquier  $\omega$ . Se puede dar una interpretación física a

$\lim_{\tau \rightarrow \infty} \frac{|G_{\tau}(\omega)|^2}{2\tau} d\omega$  como la contribución a la potencia total de  $X(t)$  de las

componentes con frecuencias entre  $\omega$  y  $\omega+d\omega$ .

Mencionamos en este punto que todo el desarrollo anterior corresponde a una única realización del proceso. Obviamente, para definir la potencia espectral del proceso debemos hacerlo utilizando el valor medio de  $|G_{\tau}(\omega)|^2/2\tau$  sobre todas las realizaciones del proceso.

$$f(\omega) = \lim_{\tau \rightarrow \infty} \left[ E \left\{ |G_{\tau}(\omega)|^2 / 2\tau \right\} \right] \quad (1.3.5)$$

en los puntos en que existe  $f(\omega)$ , y tiene la siguiente interpretación física:

$f(\omega)d\omega$  = la contribución media, de todas las realizaciones, a la potencia total de las componentes de  $X(t)$  con frecuencias entre  $\omega$  y  $\omega+d\omega$

La función  $f(\omega)$  se denomina función de potencia espectral o densidad espectral no normalizada del proceso  $X(t)$ .

Se define el espectro integrado  $F(\omega)$ , como la contribución media a la potencia total de todas aquellas frecuencias menores o iguales a  $\omega$ ,

$$F(\omega) = \int_{-\infty}^{\omega} f(\theta) d\theta \quad (1.3.6)$$

Veamos seguidamente un teorema que da la relación existente entre la función de potencia espectral  $f(\omega)$  y la función de autocovarianza  $R(\tau)$  de un proceso estacionario de parámetro continuo.

*Teorema 1.1* Sea  $\{X(t)\}$  un proceso estacionario continuo, de media cero, con función de densidad de potencia espectral  $f(\omega)$ , la cual existe para todo  $\omega$ , y con función de autocovarianza,  $R(\tau)$ . Entonces,  $f(\omega)$  es la transformada de Fourier de  $R(\tau)$ . Es decir

$$f(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau \quad (1.3.7)$$

Para que la función de autocovarianza  $R(\tau)$  posea transformada de Fourier y exista  $f(\omega)$  para cualquier  $\omega$ , es suficiente que:

$$\int_{-\infty}^{\infty} |R(\tau)| d\tau < \infty \quad (1.3.8)$$

entonces, si  $R(\tau)$  es continua, podemos expresarla como la transformada inversa de  $f(\omega)$ :

$$R(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} f(\omega) d\omega \quad (1.3.9)$$

El siguiente teorema establece la relación entre la función de autocovarianza de un proceso estacionario y su distribución espectral.

*Teorema 1.2 (Wiener-Khintchine)* La condición necesaria y suficiente para que  $R(\tau)$  sea la función de autocovarianza de un proceso estacionario continuo  $\{X(t)\}$ , es que exista una función  $F(\omega)$ , que tenga las propiedades de función de

distribución de una medida finita en  $(-\infty, \infty)$ , tal que, para todo  $\tau$ ,  $R(\tau)$  puede expresarse por:

$$R(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} dF(\omega) \quad (1.3.10)$$

la expresión (1.3.10) se denomina *representación espectral de la función de autocovarianza*.

Para los procesos estacionarios de parámetro discreto  $\{X_t\}$ , existe una representación espectral análoga. Hay que hacer no obstante las siguientes observaciones:

- $R(\tau)$  está definida para valores enteros de  $\tau$ . Luego las integrales en las que intervienen se convierten en sumas. Así, la expresión (1.3.7) se transforma en:

$$f(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} e^{-i\omega\tau} R(\tau) \quad (1.3.11)$$

- Todas las funciones espectrales van a estar definidas solamente para frecuencias en el rango  $-\pi < \omega < \pi$ , ya que el proceso sólo se observa en instantes enteros. Los armónicos cuyas frecuencias se diferencien en múltiplos de  $2\pi$ , tienen efectos indistinguibles si la observación se hace sólo en los momentos enteros, ya que  $e^{i\omega\tau} = e^{i(\omega+2\pi)\tau}$ . Esto es lo que se denomina *efecto aliasing*. Podemos afirmar que toda frecuencia fuera de  $(-\pi, \pi)$  tiene un *alias* dentro de este intervalo.

El análogo al teorema de Wiener- Khintchine para procesos discretos es el teorema de Wold cuyo enunciado es el siguiente:

*Teorema 1.3 (Wold)* Una condición necesaria y suficiente para que la sucesión  $\{R(\tau); \tau=0, \pm 1, \pm 2, \dots\}$  sea la función de autocovarianza de algún proceso estacionario discreto  $\{X_t\}_{t \in \mathbb{Z}}$  es que exista una función  $F(\omega)$ , con las propiedades de función de distribución de una medida finita en  $(-\pi, \pi)$ , tal que:

$$R(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega) \quad (1.3.12)$$

## 1.4. Representación Espectral de Procesos Estacionarios.

El teorema de representación espectral formaliza la idea de que un proceso estacionario es una superposición de movimientos armónicos de diferentes frecuencias.

Definimos en primer lugar los procesos de incrementos ortogonales, a través de los cuales se modelizará el espectro de un proceso estacionario.

Un proceso  $\{Z(\omega)\}_{\omega \in \mathbb{R}}$  se dice de incrementos ortogonales si verifica:

1.  $\forall \omega_s, \omega_t \in T, E[Z(\omega_t) - Z(\omega_s)] = 0$
2.  $\forall \omega_s, \omega_t \in T, E\left[|Z(\omega_t) - Z(\omega_s)|^2\right] < \infty$
3.  $\forall \omega_1 < \omega_2 < \omega_3 < \omega_4 \in T, E\left[(Z(\omega_4) - Z(\omega_3))(Z(\omega_2) - Z(\omega_1))\right] = 0$

Un proceso de incrementos ortogonales define una función  $F(\omega)$ , monótona, no decreciente y determinada salvo una constante, a través de la relación:

$$E\left[|Z_{\omega_t} - Z_{\omega_s}|^2\right] = F(\omega_t) - F(\omega_s) \quad (1.4.1)$$

La función  $F$  recibe el nombre de función de varianza incremental. Como consecuencia directa de la expresión anterior se tiene el siguiente teorema:

*Teorema 1.4* Un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in R}$  es continuo en media cuadrática si y sólo si  $F(\omega)$  es continua en  $\omega$ .

Pasamos a definir la integral de una función no aleatoria respecto a un proceso de incrementos ortogonales.

Definimos la integral, sobre un intervalo  $(\omega_a, \omega_b) \subset R$ , de una función no aleatoria  $g(\omega)$  con respecto a un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in T}$ , como el límite, en media cuadrática, de las sumas  $\sum_{i=0}^{n-1} g(\omega_i)(Z(\omega_{i+1}) - Z(\omega_i))$  sobre una partición  $\omega_a = \omega_0 < \omega_1 < \dots < \omega_n = \omega_b$ , cuando el radio de la partición  $\rho$  tiende a cero, es decir:

$$\int_{\omega_a}^{\omega_b} g(\omega) dZ(\omega) = L.i.m. \sum_{i=0}^{n-1} g(\omega_i)(Z(\omega_{i+1}) - Z(\omega_i)) \quad (1.4.2)$$

donde la expresión *L.i.m.* se interpreta como límite en media cuadrática.

El siguiente teorema da la condición necesaria y suficiente para la existencia de esta integral.

*Teorema 1.5* La integral de una función  $g(\omega)$  sobre  $(\omega_a, \omega_b) \subset R$  con respecto a un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in R}$  existe si y sólo si

$$\int_{\omega_a}^{\omega_b} |g(\omega)|^2 dF(\omega) < \infty \quad (1.4.3)$$

Y entonces:

$$E\left[\int_{\omega_a}^{\omega_b} g(\omega)dZ(\omega)\right] = 0$$

y

$$E\left[\int_{\omega_a}^{\omega_b} g(\omega)dZ(\omega)\right]^2 = \int_{\omega_a}^{\omega_b} |g(\omega)|^2 dF(\omega)$$

A partir de un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in R}$  cuya función de varianza incremental  $F(\omega)$  esté acotada, puede definirse un proceso estacionario  $\{X(t)\}_{t \in R}$ , mediante la relación:

$$X(t) = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega) \tag{1.4.4}$$

La función de autocovarianza del proceso  $\{X(t)\}_{t \in R}$  es precisamente la transformada de Fourier de la función de varianza incremental  $F$  del proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in R}$ , es decir:

$$R(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} dF(\omega) \tag{1.4.5}$$

El teorema que enunciamos a continuación da condiciones bajo las cuales  $X(t)$  puede expresarse como la transformada de Fourier de un proceso de incrementos ortogonales:

*Teorema 1.6* Sea  $\{X(t)\}_{t \in R}$  un proceso estacionario y continuo en media cuadrática, con función de autocovarianza  $R(\tau)$ . Entonces existe un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in R}$  tal que

$$X(t) = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega) \tag{1.4.6}$$

estando definida la integral en media cuadrática. Además, si  $\omega_1 \leq \omega_2$



$$E\left[|Z(\omega_2) - Z(\omega_1)|^2\right] = F(\omega_2) - F(\omega_1)$$

siendo  $F(\omega)$  la función correspondiente a  $R(\tau)$  en el teorema de Wiener-Khintchine.

De forma análoga para los procesos estacionarios de parámetro discreto se tiene el siguiente teorema:

*Teorema 1.7* Sea  $\{X_t\}_{t \in \mathbb{Z}}$  un proceso estacionario de parámetro discreto, con función de autocovarianza  $R(\tau)$ . Entonces existe un proceso de incrementos ortogonales  $\{Z(\omega)\}_{\omega \in (-\pi, \pi)}$  tal que

$$X_t = \int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega) \tag{1.4.7}$$

donde  $E\left[|Z(\omega_2) - Z(\omega_1)|^2\right] = F(\omega_2) - F(\omega_1) \quad \forall \omega_1, \omega_2 \in (-\pi, \pi)$

## 1.5 Procesos lineales generales.

Diremos que un proceso estocástico  $\{X_t\}_{t \in \mathbb{Z}}$ , es un proceso lineal general, si puede expresarse de la forma:

$$X_t = \sum_{u=-\infty}^{\infty} g_u \varepsilon_{t-u} \tag{1.5.1}$$

siendo  $\{\varepsilon_t\}_{t \in \mathbb{Z}}$  un proceso de variables incorreladas, que satisface:

$$E[\varepsilon_t] = 0 \quad \text{y} \quad E[\varepsilon_t^2] = \sigma_\varepsilon^2$$

y la sucesión de coeficientes  $\{g_u\}$  verifican  $\sum_{u=-\infty}^{\infty} g_u^2 < \infty$ .

*Teorema 1.8* Un proceso  $\{X_t\}_{t \in \mathbb{Z}}$  tiene distribución espectral absolutamente continua si y sólo si  $\{X_t\}_{t \in \mathbb{Z}}$  es un proceso lineal general dado por (1.5.1).

En el resto del capítulo se plantea el problema de estimar las propiedades espectrales de los procesos lineales generales a partir de la observación de una serie temporal.

## 1.6 Estimación espectral: análisis del periodograma.

Se aborda en el presente epígrafe el problema de la estimación de la función de densidad espectral de procesos estacionarios lineales, en base a la observación de una trayectoria del proceso en un conjunto finito de instantes. Tal trayectoria recibe el nombre de serie temporal. La base de la estimación de la función de densidad espectral es el periodograma, el cual se define por:

$$I_X(\omega) = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{-i\omega t} \right|^2 \quad (1.6.1)$$

siendo  $\{X_t\}_{t=1}^N$  la serie temporal observada y  $\omega_j = 2\pi j/N$ . Revisamos a continuación las propiedades del periodograma como estimador de la densidad espectral.

*Proposición 1.9* El periodograma  $I_X(\omega)$  es un estimador asintóticamente centrado de  $f(\omega)$ .

$$E[I_X(\omega)] = \int_{-\pi}^{\pi} f(\theta) F_N(\theta - \omega) d\theta = f(\omega) + O\left(\frac{\ln N}{N}\right)$$

siendo  $F_N(\theta)$  el llamado núcleo de Fejer definido por:

$$F_N(\theta) = \frac{1}{2\pi} \frac{\text{sen}^2(N\theta/2)}{N \text{sen}^2(\theta/2)} \quad (1.6.2)$$

se tiene que  $\forall \theta \neq 0, F_N(\theta) \xrightarrow{N \rightarrow \infty} 0$  y  $\int_{-\pi}^{\pi} F_N(\theta) d\theta = 1$

Sea  $\zeta_X(\omega)$  la transformada finita de Fourier de la serie temporal  $\{X_t\}_{t=1}^N$  definida por:

$$\zeta_X(\omega) = \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N X_t e^{-i\omega t} \quad -\pi \leq \omega \leq \pi \quad (1.6.3)$$

Utilizando la representación espectral de  $X_t$  dada (1.4.7) y siendo  $Z(\theta)$  un proceso de incrementos ortogonales, la expresión (1.6.3) nos queda

$$\zeta_X(\omega) = \int_{-\pi}^{\pi} \left( \frac{1}{\sqrt{2\pi N}} \sum_{t=1}^N e^{it(\theta - \omega)} \right) dZ(\theta) \quad (1.6.4)$$

simplificando (1.6.4) se obtiene

$$\zeta_X(\omega) = \int_{-\pi}^{\pi} F_N^{1/2}(\theta - \omega) e^{i \frac{(N+1)(\theta - \omega)}{2}} dZ(\theta) \quad (1.6.5)$$

De esta forma podemos expresar el periodograma en función de la transformada finita de Fourier como sigue:

$$I_X(\omega) = |\zeta_X(\omega)|^2 \quad (1.6.6)$$

Se exponen a continuación una serie de teoremas para los procesos lineales generales, que relacionan las propiedades del proceso  $\{X_t\}$  con las del proceso  $\{\varepsilon_t\}$ .

*Teorema 1.10* Sea  $\{X_t\}$  un proceso lineal general definido de la forma dada por (1.5.1) siendo los  $\{\varepsilon_t\}$  variables aleatorias incorreladas, con  $E[\varepsilon_t]=0$ ,

$E[\varepsilon_t^2]=\sigma_\varepsilon^2$  y  $\sum_{u=-\infty}^{\infty} |g_u||u|^{1/2} < \infty$ . Entonces:

$$\zeta_X(\omega) = \Gamma(\omega)\zeta_\varepsilon(\omega) + r_N(\omega) \quad (1.6.7)$$

siendo  $\Gamma(\omega) = \sum_{u=-\infty}^{\infty} g_u e^{-i\omega u}$  y  $E[|r_N(\omega)|^2] = O(1/N)$ , uniformemente en  $\omega$ . Si además  $\{\varepsilon_t\}$  son independientes con  $E[\varepsilon_t^4] < \infty$ , entonces  $E[|r_N(\omega)|^4] = O(1/N^2)$  uniformemente en  $\omega$ .

*Teorema 1.11* Sea  $\{X_t\}$  un proceso lineal general definido por (1.5.1), verificando que los  $\{\varepsilon_t\}$  son independientes con  $E[\varepsilon_t]=0$ ,  $E[\varepsilon_t^2]=\sigma_\varepsilon^2$ ,  $E[\varepsilon_t^4] < \infty$ ,

y  $\sum_{u=-\infty}^{\infty} |g_u||u|^\alpha < \infty$ ,  $\alpha > 0$ . Entonces

$$I_X(\omega) = 2\pi f(\omega) \frac{1}{\sigma_\varepsilon^2} I_\varepsilon(\omega) + R_N(\omega) \quad (1.6.8)$$

donde  $I_X$  e  $I_\varepsilon$  son los periodogramas de  $\{X_t\}_{t=1}^N$  y  $\{\varepsilon_t\}_{t=1}^N$  respectivamente, y

$$E[|R_N(\omega)|^2] = O(1/N^{2\alpha})$$

uniformemente en  $\omega$ .

En el caso particular en que las  $\{\varepsilon_t\}$  sean normales se puede concluir que las  $\{I_X(\omega_p)/f(\omega_p)\}$ ,  $p=0, 1, \dots, [N/2]$ ,  $(\omega_p=2\pi p/N)$  son independientes para  $p \neq 0, N/2$  y asintóticamente distribuida según una  $\chi_2^2$ , es decir:

$$I_X(\omega_p) \approx \begin{cases} f(\omega_p) \chi_2^2/2 & p \neq 0, N/2 \\ f(\omega_p) \chi_1^2 & p = 0, N/2 \end{cases} \quad (1.6.9)$$

de donde

$$E[I_X(\omega_p)] \approx f(\omega_p) \quad \text{Var}[I_X(\omega_p)] \approx f^2(\omega_p) \quad (1.6.10)$$

Como puede deducirse de (1.6.10), el periodograma no es un estimador consistente de  $f(\omega)$ . El siguiente teorema da la expresión de la covarianza del periodograma para dos frecuencias próximas  $\omega_1$  y  $\omega_2$ . Dado que no hay problema de confusión de la notación, consideraremos  $I_X = I$ .

*Teorema 1.12* Sea  $\{X_t\}$  un proceso lineal general definido por (1.5.1) con los  $\{\varepsilon_t\}$  y  $\{g_u\}$  verificando las condiciones del teorema 1.11. Entonces:

$$\text{Cov}\{I(\omega_1), I(\omega_2)\} = \left[ \frac{e}{N} + \frac{2\pi}{N} \{F_N(\omega_1 + \omega_2) + F_N(\omega_1 - \omega_2)\} \right] f(\omega_1) f(\omega_2) + O(1/N^\alpha) \quad (1.6.11)$$

donde  $e = \{E[\varepsilon_t^4] - 3\}$ ,  $F_N(\theta)$  es el núcleo de Fejer definido por (1.6.2), y el resto es del orden  $O(1/N^\alpha)$  uniforme en  $\omega_1$  y  $\omega_2$ .

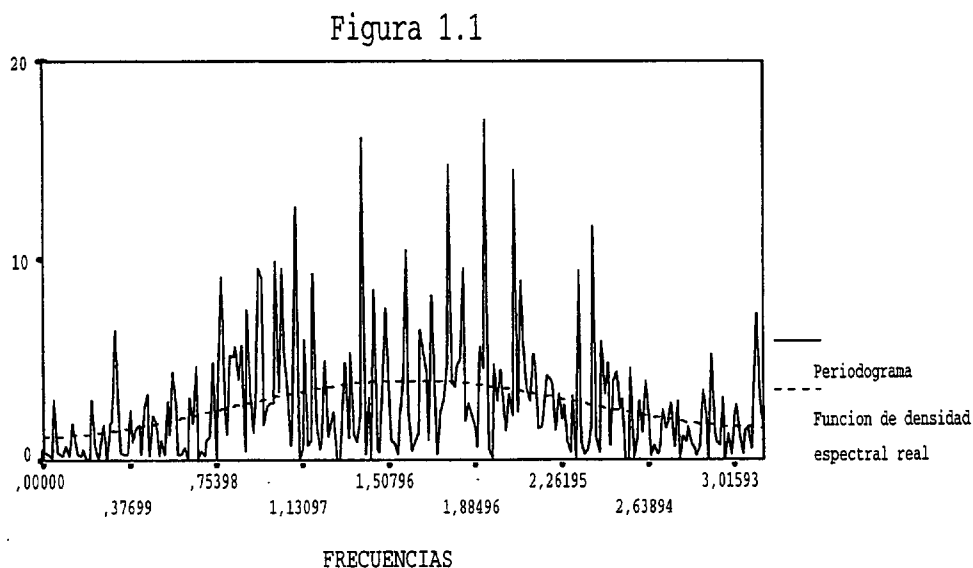
Así pues, en el caso  $\omega_1 = \omega_2 = \omega$  obtenemos la expresión de la varianza

$$\text{Var}\{I(\omega)\} = f^2(\omega)(1 + e/N) + O(1/N^\alpha) \quad (1.6.12)$$

cuando  $N \rightarrow \infty$ , la expresión asintótica de la varianza es:

$$\text{Var}\{I(\omega)\} \rightarrow f^2(\omega) \tag{1.6.13}$$

Indicamos a continuación algunas de las propiedades que se deducen directamente de los teoremas enunciados anteriormente. En primer lugar, el periodograma es un estimador asintóticamente centrado de  $f(\omega)$ , aunque la varianza no disminuye al aumentar la muestra ya que  $\text{Var}[I(\omega)] \approx f^2(\omega)$ . Además, para las frecuencias  $\omega_1$  y  $\omega_2$  se tiene  $\text{Cov}\{I(\omega_1), I(\omega_2)\} \xrightarrow{N \rightarrow \infty} 0$ , de donde se deduce que para frecuencias próximas,  $I(\omega_1)$  no da información sobre  $I(\omega_2)$ . Estas propiedades las vemos gráficamente en la figura 1.1 donde representamos la función de densidad espectral  $f(\omega)$  de un proceso MA(2) en trazo discontinuo, mientras que en trazo continuo aparece el periodograma obtenido al observar el proceso en 500 instantes de tiempo. Simplemente observando la gráfica se deduce que el periodograma es un mal estimador de la función de densidad. Mencionamos por último que  $f(\omega)$  lo hemos considerado continuo y suave, mientras que el periodograma  $I(\omega)$  presenta grandes irregularidades. Obviamente un procedimiento natural para mejorar las propiedades de este estimador consiste en suavizar localmente los valores de  $I(\omega)$  en un entorno de  $\omega$ , y esto es lo que se verá en el próximo epígrafe.



## 1.7 Estimación de la densidad espectral mediante alisamientos del periodograma.

Sea la serie  $\{X_t\}_{t=1}^N$  correspondiente al proceso estacionario  $\{X_t\}_{t \in \mathbb{Z}}$  con función de densidad espectral  $f(\omega)$ .

Consideramos ahora la estimación de  $f(\omega)$  por un estimador de núcleo dado por:

$$\hat{f}(\omega; h) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) I(\omega_j) \quad (1.7.1)$$

siendo  $I(\omega)$  el periodograma correspondiente a la serie temporal tal como se definió en (1.6.1) y siendo la función núcleo  $K(\theta)$  una función simétrica, no negativa en  $\mathbb{R}$ . Destacamos la dependencia del estimador  $\hat{f}(\omega; h)$  del parámetro de alisamiento o ancho de banda  $h$ , parámetro que condiciona fuertemente al referido estimador.

Analizamos a continuación las propiedades asintóticas del estimador  $\hat{f}(\omega; h)$ . Para ello, se le imponen las siguientes restricciones a la función núcleo  $K(\theta)$ , que pasamos a enunciar:

1.  $K(\theta)$  es una función simétrica y no negativa definida en el intervalo  $(-\infty, \infty)$ .
2.  $K$  tiene soporte compacto  $[-\kappa, \kappa]$  y es uniformemente Lipschitziana con constante  $L_\kappa$ .
3.  $\frac{1}{2\pi} \int_{-\infty}^{\infty} K(\theta) d\theta = 1$  y  $\frac{1}{2\pi} \int_{-\infty}^{\infty} \theta^2 K(\theta) d\theta = 1$ .

$$(1.7.2)$$

*Lema 1.13* (Franke y Härdle (1992)). Sea  $K$  una función núcleo que verifica las condiciones (1.7.2), y para todo  $|\omega| < \pi - \kappa h$ , sea  $p(\omega)$  una función continua y dos veces diferenciable en  $[\omega - \kappa h, \omega + \kappa h]$ . Entonces para  $h \rightarrow 0$  y  $Nh \rightarrow \infty$ , se tiene:

$$\left| \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) p(\omega_j) - p(\omega) - \frac{h^2}{2} p''(\omega) \right| \leq$$

$$\leq \frac{c}{Nh} \left\{ \sup_{\theta} |p(\theta)| + h \sup_{\theta} |p'(\theta)| \right\} + \frac{h^2}{2} \sup_{\theta} |p''(\theta) - p''(\omega)|$$

siendo  $c$  una constante y donde el supremo se calcula en el intervalo  $[\omega - \kappa h, \omega + \kappa h]$ .

*Teorema 1.14* (Franke y Härdle (1992)) Sea  $\{X_t, -\infty < t < \infty\}$  un proceso lineal

$$X_t = \sum_{k=-\infty}^{\infty} b_k \xi_{t-k}$$

donde  $\xi_t$ , son variables aleatorias independientes e idénticamente distribuidas, que satisfacen:

$$\forall l, t \quad E[\xi_{l,t}] = 0 \text{ y } E[\xi_{l,t}^2] = 1, \quad E|\xi_{l,t}^5| < \infty$$

y siendo la función característica de  $\xi_t$ ,  $q(u)$  tal que;  $\forall \delta > 0 \sup\{|q(u)|; |u| \geq \delta\} < 1$ .

Además suponemos que la función de densidad espectral  $f(\omega)$  del proceso lineal  $\{X_t\}$  es dos veces diferenciable en  $[-\pi, \pi]$ , y

$$\sum |kb_k| < \infty$$



siendo  $K$  la función núcleo que verifica las condiciones (1.7.2). Entonces para  $N \rightarrow \infty$ ,  $h \rightarrow 0$ , de forma que  $(Nh^4)^{-1} = O(1)$ , el sesgo del estimador  $\hat{f}(\omega; h)$  verifica:

$$E[\hat{f}(\omega; h)] - f(\omega) = \frac{h^2}{2} f''(\omega) + o(h^2) + O\left(\frac{\ln N}{N}\right) \quad (1.7.3)$$

uniformemente para  $|\omega| < \pi - \kappa h$ .

*Teorema 1.15* (Franke y Härdle (1992)). Sea  $\{X_t, -\infty < t < \infty\}$  un proceso lineal

$$X_t = \sum_{k=-\infty}^{\infty} b_k \xi_{t-k}$$

que verifica las condiciones del teorema 1.14 y  $K(\theta)$  la función de núcleo que cumple las condiciones (1.7.2). Si  $h \rightarrow 0$  y  $Nh^2 \rightarrow \infty$ , entonces  $\forall |\omega| < \pi$ , se tiene:

$$i) Nh \text{Var}(\hat{f}(\omega; h)) \rightarrow \sigma^2 = f^2(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta \text{ cuando } N \rightarrow \infty \quad (1.7.4)$$

$$ii) \sqrt{Nh} \{ \hat{f}(\omega; h) - E[\hat{f}(\omega; h)] \} \rightarrow Z \text{ en distribución.} \quad (1.7.5)$$

siendo  $Z$  una variable aleatoria Gaussiana de media cero y varianza  $\sigma^2$ .

Un aspecto clave en la estimación de la función de densidad espectral utilizando alisamientos mediante núcleo del periodograma es la elección del parámetro ancho de banda  $h$ .

### 1.8. El problema de la elección del ancho de banda óptimo.

Como hemos indicado anteriormente, la elección de la función núcleo no es especialmente relevante en los estimadores de núcleo. No así la elección del ancho de banda, el cual influye notablemente en el estimador. Siguiendo a Parzen (1957), consideramos como ancho de banda óptimo aquél que minimiza el porcentaje de error cuadrático medio (Mean Square Percentage Error MSPE), definido por:

$$MSPE(\omega; h) = \frac{E\{\hat{f}(\omega; h) - f(\omega)\}^2}{f^2(\omega)} \quad (1.8.1)$$

(1.8.1) puede expresarse alternativamente por:

$$MSPE(\omega; h) = \frac{v^2(\omega) + b^2(\omega)}{f^2(\omega)} \quad (1.8.2)$$

donde  $b(\omega) = (E[\hat{f}(\omega; h)] - f(\omega))$  y  $v^2(\omega) = Var[\hat{f}(\omega; h)]$ .

Si se verifican las condiciones de los teoremas 1.14 y 1.15 se tienen las siguientes expresiones asintóticas del sesgo y la varianza del estimador  $\hat{f}(\omega; h)$

$$b(\omega) = \frac{h^2}{2} f''(\omega) + o(h^2) + O\left(\frac{LnN}{N}\right) \quad (1.8.3)$$

y

$$Var\{\hat{f}(\omega; h)\} = \frac{f^2(\omega)}{Nh} \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta + o\left(\frac{1}{Nh}\right) \quad (1.8.4)$$

sustituyendo estas expresiones en (1.8.2), el MSPE puede expresarse también por:

$$MSPE(\omega; h) = \frac{h^4}{4} \left\{ \frac{f''(\omega)}{f(\omega)} \right\}^2 + \frac{1}{Nh} \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta + o\left(\frac{1}{Nh}\right) \quad (1.8.5)$$

Como puede deducirse de la expresión (1.8.3), el sesgo depende, por un lado, del grado de curvatura de la función de densidad espectral dada por  $f''(\omega)$ , y por otro, del cuadrado del ancho de banda  $h$ . De esta forma si  $|f''(\omega)|$  es grande, es decir, existe un máximo local en  $\omega$ , deberemos tomar un ancho de banda  $h$  pequeño para reducir el sesgo. Por el contrario, si  $|f''(\omega)|$  es pequeño, podemos tomar un  $h$  algo mayor de manera que no contribuya a aumentar el sesgo y a cambio conseguimos una reducción de la varianza, según se desprende de la expresión (1.8.5). Por ello, el ancho de banda óptimo debe controlar a la vez el sesgo y la varianza del estimador, y para ello, el  $MSPE(\omega;h)$ , dado por su expresión asintótica (1.8.5), converge a cero si  $h \rightarrow 0$  y  $Nh \rightarrow \infty$ .

Así pues, la expresión que se obtiene para  $h$ , es:

$$h_0(\omega) = N^{-1/5} \left\{ \frac{f(\omega)}{f''(\omega)} \right\}^{2/5} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta \right\}^{1/5} \quad (1.8.6)$$

El cálculo de un ancho de banda óptimo para cada frecuencia  $\omega$ , puede resultar excesivamente costoso desde el punto de vista computacional. Beltrao y Bloomfield (1987) proponen considerar un ancho de banda global óptimo como aquel que minimiza el promedio de la proporción de error cuadrático medio (Average Mean Square Percentage Error (AMSPE)), definido por:

$$AMSPE(h) = \frac{1}{[N/2]} \sum_{j=1}^{[N/2]} MSPE(\omega_j; h) \quad (1.8.7)$$

Por tanto,  $h_0$  queda definido como:

$$AMSPE(h_0) = \inf_h AMSPE(h)$$

De aquí se obtiene:

$$h_0 = N^{-1/5} \left\{ \sum_{j=1}^{[N/2]} \left\{ \frac{f''(\omega)}{f(\omega)} \right\}^2 \right\}^{-1/5} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta \right\}^{1/5} \quad (1.8.8)$$

Obviamente,  $AMSPE(h_0) = O(N^{-4/5})$ . En la práctica,  $h_0$  no puede calcularse exactamente dado que depende de la función  $f(\omega)$ . En este punto, se plantean los dos métodos alternativos para la estimación  $h_0$ .

*Método Plug-in.*

Este método consiste en sustituir en (1.8.8)  $f(\omega)$  por  $\hat{f}(\omega; h_l)$  donde  $h_l$  es un ancho de banda piloto y  $f''(\omega)$  por

$$\hat{f}''(\omega; l) = \frac{1}{Nl^3} \sum_{j=-[N/2]}^{[N/2]} K_{(2)} \left[ \frac{\omega - \omega_j}{l} \right] I(\omega_j) \quad (1.8.9)$$

donde el núcleo  $K_{(2)}(\theta)$  satisface las siguientes condiciones:

$$\int_{-\infty}^{\infty} K_{(2)}(u) du = \int_{-\infty}^{\infty} u K_{(2)}(u) du = 0 \quad \int_{-\infty}^{\infty} u^2 K_{(2)}(u) du = 2$$

siendo el parámetro de alisamiento  $l$  tal que  $Nl^5 \rightarrow \infty$  cuando  $l \rightarrow 0$ .

*Método de validación cruzada.*

Los métodos de validación cruzada pueden clasificarse a grandes rasgos en validación cruzada de la pseudo-verosimilitud y validación cruzada de mínimos cuadrados. Ambas formas de validación cruzada consisten en construir estimadores de la densidad espectral que no utilicen todos los datos disponibles en relación a minimizar una expresión aproximada de una medida de discrepancia entre el estimador y la verdadera función de densidad espectral.

Ilustraremos ahora el uso de la validación cruzada para la estimación de la densidad espectral en base a la observación de una serie temporal del proceso correspondiente. El método que exponemos se debe a Beltrao y Bloomfield (1987). La forma de verosimilitud de validación cruzada aproximada que obtienen es:

$$CVLL(h) = \frac{1}{[N/2]} \sum_{j=1}^{[N/2]} \left\{ \ln[\hat{f}_j(\omega_j; h)] + \frac{I(\omega_j)}{\hat{f}_j(\omega_j; h)} \right\} \quad (1.8.10)$$

donde  $\hat{f}_j(\omega_j; h)$  se definen por:

$$\hat{f}_j(\omega_j; h) = \frac{1}{(N-1)h} \sum_{k \neq j} K\left(\frac{\omega_j - \omega_k}{h}\right) I(\omega_k) \quad (1.8.11)$$

El siguiente teorema da las condiciones bajo las cuales el ancho de banda que maximiza la pseudo-verosimilitud (1.8.10) es asintóticamente equivalente,  $N \rightarrow \infty$ , a tomar el que minimiza AMSPE.

*Teorema 1.16* (Beltrao y Bloomfield (1987)). Se considera un proceso estacionario Gaussiano  $\{X_t\}_{t \in \mathbb{Z}}$  que verifica las siguientes propiedades:

i)  $E[X_0] = 0$ .

ii) La función de autocovarianza  $R(\tau) = E[X_0 X_\tau]$ , tal que:

$$\sum_{\tau=1}^{\infty} \tau |R(\tau)| < \infty$$

iii) La función de densidad espectral es definida positiva

iv) La función de núcleo  $K(\theta)$ , simétrica, no negativa y acotada. Además, verifica

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} K(\theta) d\theta = 1 \quad \text{y} \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta^2 K(\theta) d\theta < \infty$$

v)  $K(\theta) = \int_{-\infty}^{\infty} k(u) e^{-iu\theta} du$  donde  $k(u)$  tiene soporte compacto.

Además, suponemos que  $h_N$  verifica:

vi)  $h_N^{-1} = O(N^\alpha)$ , para  $\alpha < 2/5$

vii)  $h_N = o(1)$ .

Entonces para valores de  $N \rightarrow \infty$ , se verifica:

$$\begin{aligned} \frac{1}{[N/2]} CVLL(h) &= \frac{1}{[N/2]} \sum_j \{Ln[f(\omega_j)] + I(\omega_j)/f(\omega_j)\} + \\ &+ \frac{1}{2} AMSPE(h) + o_p(AMSPE(h)) \end{aligned}$$

## CAPITULO II. MÉTODO BOOTSTRAP EN EL ANÁLISIS ESPECTRAL.

### 2.1 Introducción.

El problema de la inferencia estadística consiste en analizar datos generados por una distribución de probabilidad desconocida con la finalidad, precisamente, de estimar aspectos de dicha distribución. Los métodos estadísticos son actualmente un soporte esencial del método científico. Staudte y Sheather (1990) señalan cinco etapas esenciales en la estadística del presente siglo. La primera está impulsada por los trabajos de Fisher, Neyman y Pearson durante los años veinte. La segunda está influenciada por la teoría de probabilidades de Kolmogorov (1933). La tercera está ligada al descubrimiento de la alta eficiencia de los procedimientos no paramétricos. Y la última relacionada con la robustificación de muchos métodos clásicos estadísticos. Desde nuestro punto de vista, habría que añadir una nueva etapa que comienza en 1979 con el trabajo de Efron titulado: *Bootstrap Methods: Another look at the Jackknife*.

La complejidad de muchos modelos hace frecuentemente muy difícil su estimación. Las hipótesis paramétricas pueden también resultar poco verosímiles y los métodos de estimación ser pocos robustos. En este contexto, el bootstrap de Efron puede resultar una herramienta de muy alta eficiencia, liberando al matemático de problemas complejos. Además, el hecho de que los modernos ordenadores cada vez aportan un menor coste computacional, hace que el bootstrap se consolide como una importante herramienta estadística.

La idea del bootstrap es extraordinariamente simple. Sea  $X=(X_1, \dots, X_n)$  una muestra aleatoria de una distribución  $F$  sobre  $R$ . El problema que se desea resolver es el siguiente. Dada una variable aleatoria  $R(X,F)$ , posiblemente dependiente de  $X$  y de la distribución desconocida  $F$ , estimar la distribución de  $R$  en base a los

datos observados. El método bootstrap puede esquematizarse de la siguiente forma:

*Paso 1.* Se considera la distribución empírica de la muestra

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{[X_i, \infty)}(t)$$

*Paso 2.* De la distribución empírica  $F_n$  se extrae una muestra aleatoria (con reemplazamiento) de tamaño  $n$ ,  $X^* = (X_1^*, \dots, X_n^*)$ .

*Paso 3.* Se aproxima la distribución de  $R(X, F)$  por la distribución bootstrap de  $R^* = R(X^*, F_n)$ . Esta distribución, se obtiene inducida por el mecanismo del paso 2 para extraer muestras aleatorias.

La mejor aproximación de la distribución de  $R^*$  a la de  $R$  depende en cierta medida de la forma de  $R$ , esto es una cuestión importante relativa al problema de las cantidades pivotaes. La parte más compleja del método es la obtención de la distribución bootstrap. Para ello se consideran los siguientes métodos:

A.- Cálculo teórico directo, obteniéndose la distribución bootstrap exacta.

B.- Método de Monte Carlo para obtener una aproximación de la distribución bootstrap. Se generan mediante simulación  $B$  muestras bootstrap de tamaño  $n$ ,  $X^{*1}, X^{*2}, \dots, X^{*B}$  y a partir de ellas se obtiene un histograma con los correspondientes valores  $R(X^{*1}, \hat{F}), R(X^{*2}, \hat{F}), \dots, R(X^{*B}, \hat{F})$  que nos da una aproximación de la distribución bootstrap de  $R$ .



C.- Método de desarrollos en series de Taylor para aproxima la distribución bootstrap de  $R(X^*, F_n)$  se hacen desarrollos de Taylor en el punto  $(1/n, \dots, 1/n)$

## 2.2 Validez asintótica del bootstrap.

Introduciremos en este epígrafe los elementos sobre los que se basa la validez asintótica de los procedimientos bootstrap en el sentido desarrollado por el excelente trabajo de Bickel y Freedman (1981). Haremos una revisión de los conceptos de puente Browniano, proceso empírico y convergencia débil de procesos. Introduciremos la métrica de Mallows y estableceremos resultados acerca de su relación con la convergencia débil. Como indicamos en el epígrafe anterior, el bootstrap pretende aproximar la distribución de determinadas variables  $R(X, F)$  por su *equivalente bootstrap*  $R^*(X^*, \hat{F}_n)$ . Tal aproximación será válida en algún sentido asintótico, cuando sea posible definir una distancia entre tales distribuciones y dicha distancia converja en algún sentido a cero.

### Movimiento Browniano

Consideremos un espacio probabilístico  $(\Omega, \mathcal{A}, P)$ . Un proceso  $\{X(t)\}_{t \geq 0}$  se dice que es un movimiento browniano si verifica las siguientes propiedades.

(i)  $X(0) = 0$

(ii)  $\forall t_0, \dots, t_n; 0 \leq t_0 < t_1 < \dots < t_n$ , las variables  $X(t_k) - X(t_{k-1})$   $1 \leq k \leq n$  son independientes (Proceso de incrementos independientes).

(iii)  $0 \leq s < t$ ,  $X(t) - X(s) \equiv N(\mu(t-s); \sigma^2(t-s))$ .  $\mu$  recibe el nombre de *deriva* y  $\sigma^2$  *varianza*.

Es fácil probar que  $Cov(X(t_i), X(t_j)) = \sigma^2 \min\{t_i, t_j\}$ . En el caso particular en que  $\mu=0$  y  $\sigma^2=1$ , el movimiento Browniano recibe el nombre de movimiento Browniano normalizado.

De la definición anterior resulta obvio que,  $\forall t_1, \dots, t_n; (X(t_1), \dots, X(t_n))$  tiene distribución normal multivariante con vector de medias 0 y matriz de covarianzas

$$\begin{bmatrix} t_1 & t_1 \wedge t_2 & \cdots & t_1 \wedge t_n \\ t_2 \wedge t_1 & t_2 & \cdots & t_2 \wedge t_n \\ \vdots & \vdots & \vdots & \vdots \\ t_n \wedge t_1 & t_n \wedge t_2 & \cdots & t_n \end{bmatrix}$$

donde por  $t_i \wedge t_j$  representamos  $\min\{t_i, t_j\}$ . Este conjunto de distribuciones finito dimensionales satisface las condiciones de consistencia de Kolmogorov, por lo que definen una única medida de probabilidad  $P$  sobre el conjunto  $R^{(0,\infty)}$ .

Establecemos a continuación la continuidad de las trayectorias del movimiento Browniano. Para ello tendremos en cuenta el siguiente resultado debido a Kolmogorov.

*Teorema 2.1* Sea  $\{X(t)\}_{t \in I}$  ( $I$  intervalo de  $R$ ) un proceso estocástico tal que  $E|X(t) - X(s)|^\beta \leq C|t - s|^{1+\alpha}$  para constantes positivas  $C, \alpha, \beta$ . Entonces, existe una versión de  $X(t)$  con trayectorias continuas con probabilidad 1. Si  $\{X(t)\}_{t \geq 0}$  es un movimiento Browniano (normalizado), los incrementos  $X(t) - X(s)$  tienen distribución  $N(0, t-s)$ , por lo que

$$E|X(t) - X(s)|^{2n} = \frac{(2n)!}{2^n n!} |t - s|^{2n}$$

De la identidad anterior se sigue por tanto, que existe una versión de  $\{X(t)\}_{t \geq 0}$  con trayectorias continuas con probabilidad uno.

### Puente Browniano.

Consideremos un proceso  $\{U(t)\}_{t \in [0,1]}$ , con trayectorias continuas y tal que:

$$E[U(t)] = 0 \text{ y } Cov(U(s), U(t)) = s \wedge t - st ; \quad 0 \leq s, t \leq 1 \quad (2.2.1)$$

La existencia de un tal proceso puede establecerse a través de la existencia del movimiento Browniano probada anteriormente. En efecto, consideremos un movimiento Browniano  $\{X(t)\}_{t \in [0,1]}$ , y sea el proceso

$$U(t) = X(t) - tX(1)$$

Si  $X(t)$  es la versión del movimiento Browniano con trayectorias continuas, resulta trivial probar que se verifica (2.2.1), de donde se sigue la existencia de una versión del puente browniano con trayectorias continuas.

### Convergencia débil de medidas de probabilidad

Sea  $S$  un espacio métrico y  $\mathfrak{S}$  la correspondiente  $\sigma$ -álgebra de Borel. Sea además  $\{P_n\}_{n \in \mathbb{N}}$  y  $P_0$  medidas de probabilidad definidas sobre  $(S, \mathfrak{S})$ . Decimos que la sucesión de probabilidades  $P_n$  converge débilmente a  $P_0$  ( $P_n \Rightarrow P_0$ ) si y sólo si  $P_n(A) \rightarrow P_0(A)$ ,  $\forall A \in \mathfrak{S}$ , tal que  $P_0(\partial A) = 0$  ( $\partial A$  representa la frontera de  $A$ ).

El siguiente teorema da condiciones equivalentes a la convergencia débil.

*Teorema 2.2* Sean  $\{P_n\}_{n \in \mathbb{N}}$ ,  $P_0$  medidas de probabilidad sobre  $(S, \mathfrak{S})$ . Son equivalentes las siguientes condiciones

- (i)  $P_n \Rightarrow P_0$
- (ii)  $\int f dP_n \rightarrow \int f dP_0 ; \forall f: S \rightarrow \mathfrak{R}$ , acotada y uniformemente continua.

$$(iii) \limsup_n P_n(F) \leq P_0(F), \forall F \text{ cerrado.}$$

$$(iv) \liminf_n P_n(G) \geq P_0(G), \forall G \text{ abierto.}$$

La convergencia débil de procesos estocásticos lógicamente se entenderá en el sentido de la convergencia débil de sus correspondientes distribuciones de probabilidad.

### Métrica de Mallows

Damos ahora la definición de la métrica de Mallows en un conjunto de distribuciones de probabilidad sobre un espacio de Banach. Veremos también un resultado debido a Bickel y Freedman (1981), el cual establece la equivalencia de la convergencia según la métrica de Mallows, con la convergencia débil.

Consideremos un espacio de Banach  $B$  cuya norma denotamos por  $\|\cdot\|$ , y sea  $\Gamma_p = \Gamma_p(B)$  el conjunto de las medidas de probabilidad  $\gamma$  definidas sobre la correspondiente  $\sigma$ -álgebra de Borel que verifican  $\int \|x\|^p \gamma(dx) < \infty$ . La métrica de Mallows se define en  $\Gamma_p$  del modo siguiente: Dado  $\alpha, \beta \in \Gamma_p$

$$d_p(\alpha, \beta) = \inf_{\substack{X \equiv \alpha \\ Y \equiv \beta}} \left\{ E[(X - Y)^p] \right\}^{1/p} \quad (2.2.2)$$

( $X \equiv \alpha$  significa que  $X$  tiene distribución de probabilidad  $\alpha$ )

*Lema 2.3* (Bickel y Freedman (1981))

i) El ínfimo al que hace referencia la definición de la métrica de Mallows es accesible.

ii)  $d_p$  es una métrica sobre  $\Gamma_p$ .

*Lema 2.4* (Bickel y Freedman (1981)) Sean  $\alpha_n, \alpha \in \Gamma_p$ . Entonces,  $d_p(\alpha_n, \alpha) \xrightarrow{n \uparrow \infty} 0$  es equivalente a cada una de las siguientes condiciones:

$$i) \alpha_n \Rightarrow \alpha \text{ y } \int \|x\|^p \alpha_n(dx) \rightarrow \int \|x\|^p \alpha(dx).$$

$$ii) \alpha_n \Rightarrow \alpha \text{ y } \|x\|^p \text{ es uniformemente } \alpha_n \text{ integrable.}$$

$$iii) \int \phi d\alpha_n \rightarrow \int \phi d\alpha \text{ para cualquier función continua } \phi \text{ tal que } \phi(x) = O(\|x\|^p).$$

Sea  $D=D[0,1]$  el espacio de todas las funciones reales  $x$  definidas sobre  $[0,1]$ , que son continuas a la derecha y con límites a la izquierda. Sobre este espacio podemos considerar la métrica del supremo definida por:

$$\|x - y\| = \sup_{t \in [0,1]} |x(t) - y(t)| \tag{2.2.3}$$

El espacio  $(D, \|\cdot\|)$  es un espacio métrico completo, pero no separable. Es posible, no obstante, definir sobre el conjunto  $D$  métricas alternativas  $d$ , bajo las cuales el espacio  $(D, d)$  es métrico, completo y separable. Consideramos la métrica de Skorohod definida por:

$$d(x, y) = \inf_{\lambda \in \Lambda} \left\{ \max \{ \|x - y \circ \lambda\|, \|\lambda - I\| \} \right\} \tag{2.2.4}$$

donde  $\Lambda$  es el conjunto de todas las funciones continuas y crecientes de  $[0,1]$  en  $[0,1]$ . De acuerdo con Billingsley (1968) se tiene que  $(D, d)$  es un espacio métrico completo y separable.

**Convergencia finito dimensional y convergencia débil.**

Consideramos una sucesión de procesos  $\{X_n\}_{n \in \mathbb{N}}$ , definidos sobre un mismo espacio de probabilidad y  $X_0$ , otro proceso estocástico, tal que  $X_n \Rightarrow X_0$ . Ello significa que la correspondiente sucesión de distribuciones de probabilidad

$P^{X_n}$  converge débilmente a  $P^{X_0}$ . Esto supone que  $\forall t_1, \dots, t_k \in T$  (espacio de parámetros del proceso)  $P^{X_n} \circ \pi_{t_1, \dots, t_k}^{-1} \Rightarrow P^{X_0} \circ \pi_{t_1, \dots, t_k}^{-1}$  ( $\pi_{t_1, \dots, t_k}(f) = (f(t_1), \dots, f(t_k))$ ). Obviamente,  $P^{X_0} \circ \pi_{t_1, \dots, t_k}^{-1}$  es la distribución finito-dimensional ( $k$ -dimensional) del proceso  $X_n$  para los instantes  $t_1, \dots, t_k$ . Todo esto puede expresarse diciendo que la convergencia débil de una sucesión de procesos implica la convergencia de las correspondientes distribuciones finito dimensionales. Sin embargo, el recíproco de esta propiedad no es cierto en general.

Daremos condiciones a continuación bajo las cuales la convergencia finito dimensional implica la convergencia débil. Para ello definiremos en primer lugar el concepto de *tight*.

Una sucesión de medidas de probabilidad  $\{P_n\}$  sobre un espacio métrico  $(M, \delta)$  es un *tight*, si  $\forall \varepsilon > 0; \exists K \subset M, K$  compacto, tal que  $P_n(K) \geq 1 - \varepsilon$ . Una sucesión de elementos aleatorios con valores en un espacio métrico  $(M, \delta)$  es un *tight*, si la correspondiente sucesión de medidas inducidas es un *tight*.

El siguiente resultado da las condiciones bajo las cuales la convergencia finito dimensional implica la convergencia débil.

*Teorema 2.5* Sea  $(M, \delta)$  un espacio separable y  $\{X_n\}_{n \in \mathbb{N}}$  una sucesión de procesos con valores en  $(M, \delta)$  tal que  $X_n \xrightarrow{f.d.} X_0$  (convergencia finito dimensional) y  $\{X_n\}_{n \in \mathbb{N}}$  es un *tight*, entonces  $X_n \Rightarrow X_0$  sobre  $(M, \mathcal{M}_\delta, \delta)$  donde  $\mathcal{M}_\delta$  es la  $\sigma$ -álgebra de Borel en  $M$  inducida por la métrica  $\delta$ .

Si consideramos una sucesión de procesos  $\{X_n\}$ , cuyas trayectorias con probabilidad uno están en el espacio  $D = D[0, 1]$ , la convergencia finito dimensional a un proceso  $X_0$  implicará la convergencia débil, si además de ser

$\{X_n\}$  un *tight*, se considera sobre  $D$  una métrica  $\delta$ , bajo la cual el espacio  $(D, \delta)$  es separable. Esto, según vimos anteriormente, es posible si  $\delta$  es la métrica de Skorohod.

### Convergencia de procesos empíricos

Sea  $X_1, \dots, X_n$  una muestra aleatoria de una distribución  $F$  sobre  $R$  y sea la distribución empírica  $F_n$  definida por:

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n I_{[X_i, \infty)}(t) \quad (2.2.5)$$

Nuestro objetivo ahora consiste en estudiar la convergencia débil del proceso  $\sqrt{n}(\hat{F}_n(t) - F(t))$ . Para ello consideramos en primer lugar  $n$  variables aleatorias  $\xi_1, \dots, \xi_n$ , independientes y con distribución uniforme en  $[0,1]$ . Definimos el proceso empírico uniforme  $G_n$  por:

$$G_n(t) = \frac{1}{n} \sum_{i=1}^n I_{[\xi_i, \infty)}(t)$$

Resulta obvio que  $nG_n(t) \cong \text{Binomial}(n, t)$ , por lo que:

$$E[G_n(t)] = t \text{ y } n\text{Cov}(G_n(s), G_n(t)) = s \wedge t - st$$

Consideramos también el proceso

$$U_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{I_{[\xi_i, \infty)}(t) - t\} = \sqrt{n}[G_n(t) - t]$$

se tiene que:

$$E[U_n(t)] = 0 \text{ y } \text{Cov}(U_n(s), U_n(t)) = s \wedge t - st$$

De acuerdo con el teorema central del límite, podemos establecer la existencia de un puente Browniano  $U(t)$ , tal que

$$U_n(t) \xrightarrow[f.d.]{n \uparrow \infty} U$$

Obsérvese que las trayectorias del proceso  $U_n(t)$ , para todo  $n$ , son elementos del espacio  $D[0,1]$ .

Establecemos a continuación la convergencia débil de la sucesión  $U_n(t)$  a un puente Browniano. Tal resultado, como veremos, puede obtenerse de resultados más generales que daremos ahora

Consideremos una doble sucesión  $\{C_{ni}; i=1, \dots, n\}$  que verifica la llamada condición u.a.n. (Uniformly asymptotically negligible), es decir:

$$\max_{1 \leq i \leq n} \frac{c_{ni}^2}{\sum_{j=1}^n c_{nj}^2} \xrightarrow{n \rightarrow \infty} 0 \quad (2.2.6)$$

Sea además una doble sucesión  $\{G_{ni}(t); i=1, \dots, n\}$  de funciones de distribución de probabilidad arbitrarias en  $[0,1]$  y consideremos el proceso definido por

$$Z_n(t) = \frac{1}{\sum_{j=1}^n c_{nj}^2} \sum_{i=1}^n c_{ni} \left[ I_{[\xi_i, \infty)}(t) - G_{ni}(t) \right] \quad t \in [0,1]$$

donde  $\{\xi_i; i=1, \dots, n\}$  son variables aleatorias independientes e idénticamente distribuidas, con distribución  $U[0,1]$ . Consideramos también la función

$$\gamma_n(t) = \frac{\sum_{i=1}^n c_{ni}^2 G_{ni}(t)}{\sum_{j=1}^n c_{nj}^2}$$



donde  $\{\xi_i; i=1, \dots, n\}$  son variables aleatorias independientes e idénticamente distribuidas, con distribución  $U[0,1]$ . Consideramos también la función

$$\gamma_n(t) = \frac{\sum_{i=1}^n c_{ni}^2 G_{ni}(t)}{\sum_{j=1}^n c_{nj}^2}$$

y la función incremento

$$\gamma_n(s,t) = \gamma_n(t) - \gamma_n(s) \quad 0 \leq s \leq t \leq 1$$

Tenemos entonces el siguiente resultado

*Teorema 2.6* (Shorack, 1986) Supongamos  $\{G_{ni}; i=1, \dots, n\}$  verificando la condición u.a.n. si

$$a_m = \overline{\lim}_{n \rightarrow \infty} \max_{1 \leq k \leq m} \gamma_n\left(\frac{k-1}{m}, \frac{k}{m}\right) \longrightarrow 0 \quad m \uparrow \infty$$

Entonces  $\exists \{Z(t)\}$ , proceso con trayectorias casi seguro en  $(D, \Delta)$  tal que  $Z_n \Rightarrow Z$ ,  $n \uparrow \infty$  si y sólo si

$$K_n(s,t) = \text{Cov}(Z_n(s), Z_n(t)) \longrightarrow K(s,t), \quad n \uparrow \infty; \quad 0 \leq s \leq t \leq 1$$

En este caso,  $Z$  es un proceso normal, de media cero y función de covarianza  $K$  y con trayectorias casi seguro continuas. ( $\Delta$  es la  $\sigma$ -álgebra de Borel generada por la métrica de Skorohod)

De este resultado general, podemos deducir que el proceso empírico anteriormente considerado  $U_n(t)$  converge débilmente a un puente Browniano.

En efecto; consideramos  $C_{ni} = \frac{1}{n}$ ; para todo  $i=1, \dots, n$  y  $G_{ni}(t)=t$ , para todos los  $t, n, i$ . El proceso  $Z_n(t)$  adoptará entonces la expresión:

$$Z_n(t) = \sqrt{n}[G_n(t) - t] = U_n(t)$$

también

$$\gamma_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{1}{n^2} t = t \quad \text{y} \quad \gamma_n(s, t) = t - s, \quad 0 \leq s \leq t \leq 1$$

Por tanto

$$a_m = \overline{\text{Lim}}_{n \rightarrow \infty} \max_{1 \leq k \leq m} \gamma_n\left(\frac{k-1}{m}, \frac{k}{m}\right) = \overline{\text{Lim}}_{n \rightarrow \infty} \max_{1 \leq k \leq m} \frac{1}{n} = 0$$

Además

$$K_n(s, t) = \text{Cov}(G_n(s), G_n(t)) = s \wedge t - st$$

Estamos entonces en condiciones de enunciar el siguiente corolario

*Corolario 2.7*  $U_n \Rightarrow U$ , siendo  $U$  un puente Browniano.

Podemos establecer el resultado buscado acerca de la convergencia del proceso empírico

$$\sqrt{n}(F_n(t) - F(t))$$

donde  $F_n(t)$  es la función de distribución empírica de una muestra aleatoria de la distribución de probabilidad  $F$ . Para ello, basta considerar

$$U_n(F) = \sqrt{n}[G_n(F) - F]$$

Obviamente

$$G_n(F(t)) = \frac{1}{n} \sum_{i=1}^n I_{[\xi_i, \infty)}(F(t))$$

siendo  $\xi_i = F(X_i)$  y  $X_1, \dots, X_n$  la referida muestra aleatoria de la distribución  $F$ . Si esta distribución es continua,  $\xi_1, \dots, \xi_n$  construirán una muestra aleatoria de una distribución  $U[0,1]$ .

Ahora bien,

$$G_n(F(t)) = \frac{1}{n} \sum_{i=1}^n I_{[\xi_i, \infty)}(F(t)) = \frac{1}{n} \sum_{i=1}^n I_{[\xi_i, \infty)}(t) = F_n(t)$$

tenemos entonces que

$$\sqrt{n}(F_n - F) \Rightarrow U(F)$$

siendo  $U$  un puente Browniano.

No es necesaria la continuidad de la función de distribución  $F$  para establecer el resultado anterior. Sea

$$F^{-1}(y) = \inf\{x; F(x) \geq y\}$$

Podemos entonces considerar la muestra aleatoria  $X_1, \dots, X_n$  generada por  $F^{-1}(\xi_1), \dots, F^{-1}(\xi_n)$ , donde  $\xi_1, \dots, \xi_n$  es una muestra aleatoria de la distribución  $U[0,1]$ . Bajo estas consideraciones se establece la referida convergencia débil del proceso  $\sqrt{n}(F_n - F)$ . Utilizamos ahora la siguiente notación de Bickel y Freedman (1981). Sea  $\Gamma$  el conjunto de todas las funciones de distribución en la norma del supremo. Sobre el espacio  $D^{[0,1]}$ , seguiremos considerando la  $\sigma$ -álgebra de Borel  $\Delta$  generada por la métrica de Skorohod  $\delta$ , por lo que el espacio  $(D, \delta)$  es métrico, completo y separable. Para cualquier  $F \in \Gamma$ , sea  $X_1, \dots, X_n$  una muestra aleatoria de  $F$ , y sea el proceso empírico

$$W_{F_n}(t) = \sqrt{n}[F_n(t) - F(t)]$$

Si llamamos  $\psi_m(F)$  a la distribución del proceso  $W_{F_n}$ , obviamente  $\psi_m(F)$  es una distribución de probabilidad sobre  $(D, \Delta)$ , la cual converge débilmente a  $U(F)$ , siendo  $U$  un puente Browniano.

### El bootstrap en procesos empíricos.

Sea  $X_1, \dots, X_n$  una muestra aleatoria de  $F$  y sea  $F_n$  la correspondiente distribución empírica. Consideramos a su vez una muestra aleatoria  $X_1^*, \dots, X_m^*$  (bootstrap) de  $F_n$  y sea la correspondiente distribución empírica

$$F_{nm}(t) = \frac{1}{m} \sum_{i=1}^m I_{[X_i^*, \infty)}(t)$$

Tenemos entonces que el proceso empírico

$$W_{nm}(t) = \sqrt{m} [F_{nm}(t) - F_n(t)]$$

convergerá débilmente a un proceso  $U(F)$ , siendo  $U$  un puente Browniano. La demostración de este resultado se debe a Bickel y Freedman (1981). Utilizando un argumento debido a Shorack (1982) se simplifica notablemente la demostración.

En primer lugar, si tenemos una sucesión de funciones de distribución  $F_n$  que converge débilmente hacia una función de distribución  $F_0$ , podemos construir fácilmente un espacio de probabilidad, y una sucesión de variables aleatorias  $\{X_n\}$  sobre dicho espacio de probabilidad, de tal manera que cada variable  $X_n$  tenga distribución  $F_n$ , y otra variable aleatoria  $X_0$  sobre el mismo espacio de probabilidad y con distribución  $F_0$ , de tal manera que la sucesión  $\{X_n\}$  converja casi seguramente a  $X_0$ . La idea de tal construcción consiste en observar una variable aleatoria  $\xi$  con distribución  $U[0,1]$  y definir  $X_n = F_n^{-1}(\xi)$  y  $X_0 = F_0^{-1}(\xi)$ .

Consideramos ahora la muestra aleatoria  $X_1, \dots, X_n$  de la distribución  $F$ , y la correspondiente distribución empírica  $F_n$ . Sea asimismo  $X_{n1}^*, \dots, X_{nm}^*$  una muestra aleatoria de tamaño  $m$  de la distribución empírica  $F_n$ . Por un argumento similar al párrafo anterior, podemos considerar que cada  $X_{ni}^* = F_n^{-1}(\xi_{mi}^*)$ , donde  $\xi_{m1}^*, \dots, \xi_{mm}^*$  son variables aleatorias independientes e idénticamente distribuidas, con distribución  $U[0,1]$  y cuyo proceso empírico  $U_m^*$  satisface  $\|U_m^* - U\| \xrightarrow[m \uparrow \infty]{a.s.} 0$ , (*a.s.* convergencia casi seguro) siendo  $U$  un puente Browniano. Esta construcción se basa en la convergencia débil de la sucesión de procesos  $U_m^*$  a un puente Browniano  $U$ . Por tanto, llamando  $F_{nm}$  a la distribución empírica de  $X_{n1}^*, \dots, X_{nm}^*$ , tenemos:

$$\begin{aligned} \|\sqrt{m}(F_{nm} - F_n) - U(F)\| &= \|U_m^*(F_n) - U(F)\| \leq \\ &\leq \|U_m^*(F_n) - U(F_n)\| + \|U(F_n) - U(F)\| \leq \|U_m^* - U\| + \|U(F_n) - U(F)\| \end{aligned}$$

La desigualdad  $\|U_m^*(F_n) - U(F_n)\| \leq \|U_m^* - U\|$  es trivial dada la forma de la norma del supremo. Por construcción se tiene que  $\|U_m^* - U\| \xrightarrow[n \uparrow \infty]{a.s.} 0$ . Por el teorema de Glivenko-Cantelli,  $\|F_n - F\| \xrightarrow[n \uparrow \infty]{a.s.} 0$ , y dada la continuidad casi segura de las trayectorias del puente browniano, podemos afirmar también que  $\|U(F_n) - U(F)\| \xrightarrow[n \uparrow \infty]{a.s.} 0$ . Por tanto

$$\|\sqrt{m}(F_{nm} - F_n) - U(F)\| \xrightarrow{a.s.} 0, \quad n \wedge m \rightarrow \infty$$

Este resultado nos permite enunciar el siguiente teorema

**Teorema 2.8** Dada la sucesión de variables aleatorias  $\{X_n\}_{n \in \mathbb{N}}$

$$\sqrt{m}(F_{nm} - F_n) \Rightarrow U(F) \quad n \wedge m \rightarrow \infty$$

La importancia de este resultado es notable desde el punto de vista de la validez asintótica del método bootstrap. Según hemos visto, la distribución asintótica del pivote  $\sqrt{n}(F_n - F)$  es la distribución del proceso  $U(F)$ , siendo  $U$  un puente Browniano. Este último teorema a su vez garantiza que la distribución asintótica del pivote bootstrap  $\sqrt{m}(F_{nm} - F_n)$  es también la distribución de un proceso  $U(F)$  con  $U$  puente Browniano. Dado que el método de Monte Carlo nos permite obtener mediante remuestreos aleatorios una aproximación a la distribución de  $\sqrt{m}(F_{nm} - F_n)$ , ello nos permite disponer de una aproximación para la distribución del pivote  $\sqrt{n}(F_n - F)$ .

### 2.3. Estimación de la función de densidad espectral utilizando la metodología bootstrap.

A lo largo de los últimos años el bootstrap se ha utilizado para la estimación de muy variados modelos, entre otros, los modelos de procesos estacionarios. Freedman (1984), Efron y Tibshirani (1986), Swanepoel y van Wyk (1986) y Kreiss y Franke (1989) utilizan el bootstrap para la estimación de modelos paramétricos de series temporales. Franke y Härdle (1992) usan esta metodología para estimar la función de densidad espectral de un proceso lineal general. Daremos ahora este último procedimiento, así como los teoremas que justifican su validez asintótica en el sentido de Bickel y Freedman.

#### 2.3.1. Método bootstrap para la estimación de la densidad espectral.

Sea  $\{X_1, \dots, X_N\}$  una serie temporal correspondiente a un proceso estacionario con densidad espectral  $f(\omega)$ . Según se vio en 1.7, un estimador de  $f(\omega)$  es:

$$\hat{f}(\omega; h) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) I(\omega_j)$$

siendo  $K(\theta)$  la función núcleo, que es simétrica y no negativa; y  $I(\omega_j)$  el periodograma correspondiente a la serie observada.

Franke y Härdle (1992) consideran el modelo multiplicativo

$$I(\omega_j) = f(\omega_j) \varepsilon_j \quad j=1, \dots, [N/2] \quad (2.3.1)$$

siendo las variables residuales  $\varepsilon_j$  asintóticamente independientes e idénticamente distribuidas. El algoritmo bootstrap que proponen es similar al de Härdle y Bowman (1988) para modelos de regresión no paramétrica. El algoritmo opera como si las variables residuales fuesen independientes. Esta situación hará más compleja la prueba de la validez asintótica de la aproximación bootstrap.

*Paso 1.* Se considera un ancho de banda global piloto  $h_1 > 0$ , y se estiman los residuos a través de

$$\hat{\varepsilon}_j = \frac{I(\omega_j)}{\hat{f}(\omega_j; h_1)}$$

y reescalándolos en la forma:

$$\tilde{\varepsilon}_j = \frac{\hat{\varepsilon}_j}{\hat{\varepsilon}} \quad j=1, \dots, [N/2]$$

donde

$$\hat{\varepsilon} = \frac{1}{[N/2]} \sum_{j=1}^{[N/2]} \hat{\varepsilon}_j$$

*Paso 2.* Se obtienen  $B$  muestras bootstrap  $\{\varepsilon_1^*, \dots, \varepsilon_{[N/2]}^*\}$  de la distribución empírica de  $\{\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_{[N/2]}\}$ , de tal forma que:

$$\forall j=1, \dots, [N/2], \forall k=1, \dots, [N/2] \quad P\{\varepsilon_j^* = \tilde{\varepsilon}_k\} = \frac{1}{[N/2]}$$

Utilizando (2.3.1) se obtienen los periodogramas bootstrap

$$I^*(\omega_k) = I^*(-\omega_k) = \hat{f}(\omega_k; g) \varepsilon_k^*$$

Por convenio se toma  $I^*(0) = 0$ . Finalmente se considera como estimador bootstrap de  $f(\omega)$

$$\hat{f}^*(\omega; h, g) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) I^*(\omega_j) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \hat{f}(\omega_j; g) \varepsilon_j^*$$

Tal y como se obtienen en el paso 2 los residuos bootstrap  $\varepsilon_j^*$ , siguen la distribución empírica de la muestra  $\{\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_{[N/2]}\}$ , con respecto a la cual tienen media uno,  $E^*[\varepsilon_j^*] = 1$ .

### 2.3.2. Validez asintótica del método.

Consideremos la cantidad pivotal

$$\sqrt{Nh} \{ \hat{f}(\omega; h) - f(\omega) \} / f(\omega)$$

y su aproximación bootstrap

$$\sqrt{Nh} \{ \hat{f}^*(\omega; h, g) - \hat{f}(\omega; g) \} / \hat{f}(\omega; g)$$



La validez asintótica se entenderá si las distribuciones de ambas cantidades pivotaes están próximas de acuerdo con la métrica de Mallows. El siguiente teorema de Franke y Härdle da las condiciones precisas bajo las cuales la métrica de Mallows converge a cero en probabilidad.

*Teorema 2.9* Sea  $\{X_t, -\infty < n < \infty\}$  un proceso lineal que toma valores reales

$$X_t = \sum_{k=-\infty}^{\infty} b_k \xi_{t-k}$$

donde  $\xi_t$ , son variables aleatorias independientes e idénticamente distribuidas, que satisfacen:

$$\forall l, t, \quad E[\xi_{l,t}] = 0 \quad \text{y} \quad E[\xi_{l,t}^2] = 1, \quad E|\xi_{l,t}^5| < \infty$$

y siendo la función característica de  $\xi_n$ ,  $q(u)$  tal que;  $\forall \delta > 0, \sup\{|q(u)|; |u| \geq \delta\} < 1$ .

Además, suponemos que la función de densidad espectral  $f(\omega)$  del proceso lineal  $\{X_n\}$  es no nula y dos veces diferenciable en  $[-\pi, \pi]$ , y que se satisface la acotación  $\sum |kb_k| < \infty$ . Sea además  $K$  una función núcleo verificando las condiciones (1.7.2). Sea  $\kappa(u)$  la transformada de Fourier de  $K(\theta)$ , y supongamos que el siguiente límite existe y es finito y distinto de cero

$$\lim_{u \rightarrow 0} \frac{\kappa(0) - \kappa(u)}{u^2}$$

Para  $N \rightarrow \infty$ , se considera  $h$  el ancho de banda del estimador de interés,  $h_I$  el ancho de banda inicial, y  $g$  el ancho de banda del remuestreo, que verifican:

$$h \approx N^{-1/5}, \quad h_I \rightarrow 0 \quad \text{tal que} \quad (Nh_I^4)^{-1} = O(1), \quad g \rightarrow 0 \quad \text{tal que} \quad h/g \rightarrow 0.$$

Bajo estas condiciones, el algoritmo bootstrap descrito anteriormente, satisface:

$$d_2 \left[ \sqrt{Nh} \frac{\hat{f}(\omega; h) - f(\omega)}{f(\omega)}; \sqrt{Nh} \frac{\hat{f}^*(\omega; h, g) - \hat{f}(\omega; g)}{\hat{f}(\omega; g)} \right] \longrightarrow 0 \text{ en probabilidad}$$

En el problema de estimación del modelo de regresión por estimadores de núcleo, Härdle y Bowman (1988) utilizan un algoritmo bootstrap similar, pero tomando el ancho de banda de remuestreo del orden  $N^{-1/5}$ . Aunque inicialmente el ancho de banda óptimo es de este orden, tal elección presenta el problema de que el sesgo del pivote bootstrap, no aproxima adecuadamente al sesgo del pivote real. Esta es la clave de que en las condiciones del teorema anterior, se exija que  $h/g \rightarrow 0$ , lo que supone que si  $h$  es de orden  $N^{-1/5}$ ,  $g$  debe ser de orden  $N^{-\alpha}$  con  $\alpha < 1/5$ . Härdle y Bowman, para poder elegir el parámetro  $g$  del orden  $N^{-1/5}$ , estiman separadamente el sesgo del pivote bootstrap a través de la estimación de  $f''(\omega)$ . Franke y Härdle consideran también una elección del parámetro  $g$  del orden  $N^{-1/5}$ , estimando el sesgo bootstrap separadamente. El siguiente teorema, también de Franke y Härdle, recoge el procedimiento alternativo de estimación separada del sesgo.

*Teorema 2.10* Sea  $\{X_t, -\infty < t < \infty\}$  un proceso lineal que toma valores reales que verifica las hipótesis del teorema (2.9). Sea  $K$  una función núcleo que cumple las condiciones (1.7.2). Sea  $\hat{f}''(\omega)$  un estimador consistente de  $f''(\omega)$ . Sea

$$\hat{f}_c(\omega; h, g) = E^*[\hat{f}^*(\omega; h, g)] = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \hat{f}(\omega_j; g)$$

$\hat{f}_c$  corresponde a la esperanza condicionada de  $\hat{f}^*(\omega; h, g)$  respecto a los datos originales. Entonces, para  $N \rightarrow \infty$ ,  $h \approx N^{-1/5}$ ,  $g \approx N^{-1/5}$ ,  $h_I \rightarrow 0$  tal que  $(Nh_I^4)^{-1} = O(1)$ , se tiene

$$d_2 \left[ \sqrt{Nh} \frac{\hat{f}(\omega; h) - f(\omega)}{f(\omega)}; \sqrt{Nh} \frac{\hat{f}^*(\omega; h, g) - \hat{f}_c(\omega; h, g) + (h^2/2)\hat{f}''(\omega)}{\hat{f}(\omega; g)} \right] \rightarrow 0$$
 en probabilidad

Como estimador consistente de  $f''(\omega)$  puede tomarse, tal y como se vió en el epígrafe 1.8, un estimador de núcleo dado por la expresión (1.8.9)

$$\hat{f}''(\omega, l) = \frac{1}{Nl^3} \sum_{j=-[N/2]}^{[N/2]} K_{(2)} \left[ \frac{\omega - \omega_j}{l} \right] I(\omega_j)$$

verificando el núcleo  $K_{(2)}(\theta)$  las siguientes condiciones:

$$\int_{-\infty}^{\infty} K_{(2)}(u) du = \int_{-\infty}^{\infty} u K_{(2)}(u) du = 0 \quad \int_{-\infty}^{\infty} u^2 K_{(2)}(u) du = 2$$

y el ancho de banda  $l$  de forma que  $Nl^5 \rightarrow \infty$  cuando  $l \rightarrow 0$ .

## CAPITULO III. SERIES TEMPORALES REPLICADAS: ANÁLISIS DEL PERIODOGRAMA MEDIO.

### 3.1 Introducción.

Como se ha indicado ya en esta memoria, el análisis espectral es una metodología altamente desarrollada para el caso de observaciones de una única serie temporal. Esto, como se ha dicho anteriormente, es debido a que originalmente sus aplicaciones se hicieron en el campo de la teoría de la señal, donde sólo se dispone de una única serie para la estimación del espectro. Sin embargo, en las ciencias biomédicas, la replicación de experimentos es más la regla que la excepción. Así por ejemplo, podemos disponer de una muestra de individuos de una cierta población y sobre cada uno de ellos observar un mismo proceso estocástico estacionario, por lo general, en los mismos instantes de tiempo. Señales analógicas tales como un electroencefalograma pueden digitalizarse en los mismos instantes de tiempo. También, en problemas astronómicos puede ser conveniente hacer replicas de la señal luminosa de una estrella.

Diggle y Al-Wasel (1993) consideraron series temporales de mediciones de concentración de la hormona LH en muestras de sangre en 8 individuos. Inicialmente supusieron que todas las series procedían de la observación de un mismo proceso estocástico estacionario con función de densidad espectral común  $f(\omega)$  y propusieron el periodograma medio como estimador de la densidad espectral. Tales mediciones las realizaron primeramente cada 5 minutos y durante una hora y media. Posteriormente aumentaron el número de observaciones a una cada minuto y durante una hora. En este último caso observaron que la teoría asintótica del periodograma era incompatible con la hipótesis de la existencia de una densidad espectral común para todos los procesos de los cuales se obtuvieron las referidas series. Ello les llevó a formular el llamado modelo de efectos

aleatorios de ruido blanco, cuyo fundamento consiste en que cada sujeto tiene su propia función de densidad espectral  $f_i(\omega)$ . Consideraron además, que cada función  $f_i(\omega)$  es una realización de un proceso  $R(\omega)$  y definen el espectro poblacional  $f(\omega)$  la media de dicho proceso. A partir de este modelo estiman el espectro poblacional realizando ciertas parametrizaciones de los procesos que forman el modelo.

Desde nuestro punto de vista, los argumentos utilizados por Diggle y Al-Wasel (1993) para justificar su modelo de ruido blanco, merecen ciertas consideraciones. Por esta razón, introducimos un modelo general de efectos aleatorios (epígrafe 3.6) basado en hipótesis más flexibles y verosímiles. Además, desarrollamos una teoría asintótica del periodograma medio como estimador del espectro poblacional. Proponemos asimismo una aproximación bootstrap para la distribución del periodograma medio y estudiaremos su validez en el sentido dado por Bickel y Freedman (1981) revisado en el capítulo anterior.

### 3.2 El modelo de efectos fijos.

Consideramos inicialmente un proceso estocástico estacionario  $\{X_t, t \in Z^+\}$  con distribución espectral absolutamente continua, siendo  $f(\omega)$  su función de densidad espectral, la cual se desea estimar. La observación del proceso en un conjunto de instantes  $\{1, \dots, N\}$ , da lugar a una serie temporal  $\{x_t, t=1, \dots, N\}$ , cuyo periodograma, tal y como se vio en el capítulo primero, constituye la base de la estimación de la función de densidad espectral  $f(\omega)$ .

Supongamos ahora que del mismo proceso  $\{X_t, t \in Z^+\}$  obtenemos  $r$  realizaciones en los mismos instantes de tiempo  $\{1, \dots, N\}$ , dando lugar a  $r$  series de la forma  $\{x_{lt}, l=1, \dots, r; t=1, \dots, N\}$ . Para ello seleccionamos una muestra aleatoria de  $r$  objetos o individuos de una población  $\{O_1, \dots, O_r\}$  y sobre cada uno

de los cuales obtenemos una serie que corresponde a la observación del proceso  $\{X_t, t \in Z^+\}$ . Sea  $I_l(\omega_j)$  el periodograma correspondiente de la  $l$ -ésima serie temporal  $\{x_{lt} / t=1, \dots, N\}$  en la frecuencia de Fourier  $\omega_j = 2\pi j/N$ , el cual viene dado por la siguiente expresión:

$$I_l(\omega_j) = \frac{1}{2\pi N} \left| \sum_{t=1}^N x_{lt} e^{i\omega_j t} \right|^2 \quad (3.2.1)$$

En base a los datos correspondientes a las  $r$  series temporales, es razonable considerar como estimador de la función de densidad espectral  $f(\omega)$ , el periodograma medio definido por

$$\bar{I}(\omega_j) = \frac{1}{r} \sum_{l=1}^r I_l(\omega_j) \quad (3.2.2)$$

Siguiendo los resultados acerca del periodograma dados en el epígrafe 1.6 (capítulo I), podemos escribir:

$$1) E[\bar{I}(\omega)] = f(\omega) + O\left(\frac{\ln N}{N}\right)$$

$$2) \text{Var}[\bar{I}(\omega)] = \frac{f^2(\omega)}{r} \left(1 + \frac{e}{N}\right) + \frac{1}{r} O\left(\frac{1}{N^\alpha}\right)$$

$$3) \text{Cov}[\bar{I}(\omega_1), \bar{I}(\omega_2)] = \frac{C}{r} f(\omega_1) f(\omega_2) + \frac{1}{r} O\left(\frac{1}{N^\alpha}\right)$$

siendo  $C = \left[ \frac{e}{N} + \frac{2\pi}{N} \{F_N(\omega_1 + \omega_2) + F_N(\omega_1 - \omega_2)\} \right]$ ;  $e = \{E[\varepsilon_t^4] - 3\}$  y  $F_N(\theta)$  es el núcleo de Fejer definido por (1.6.2).

De acuerdo con la teoría asintótica del periodograma dada por Hartigan (1980), podemos considerar el modelo alternativo dado por:

$$I_l(\omega_j) = f(\omega_j)U_{lj} \quad (3.2.3)$$

donde  $U_{lj}$  son variables aleatorias independientes en  $l$  y asintóticamente independientes en  $j$ , y con distribución exponencial de parámetro 1. Bajo este modelo, el periodograma medio definido por (3.2.2), es un estimador centrado y consistente en  $r$  para la función de densidad espectral  $f(\omega)$ . Además, para cada  $N$  fijo

$$\sqrt{r} \frac{\bar{I}(\omega_j) - f(\omega_j)}{f(\omega_j)} \xrightarrow{r \rightarrow \infty} N(0,1) \text{ en distribución} \quad (3.2.4)$$

El modelo (3.2.3) está basado en el hecho de que todas las series temporales proceden de la observación de un mismo proceso estacionario lineal con función de densidad espectral  $f(\omega)$ . El uso de este modelo, lógicamente, es aceptable cuando el número de observaciones por objeto es razonablemente grande. A este modelo lo denominamos *modelo de efectos fijos*.

### 3.3 El modelo de efectos aleatorios.

Diggle y Al-Wasel (1993) consideraron la determinación del patrón de variación de la concentración de hormona LH en una cierta población de individuos. Para ello utilizaron medidas de concentración en sangre de esta hormona en ocho mujeres sanas en los mismos instantes de tiempo a intervalos de 5 minutos durante un periodo de una hora y media  $\{X_{lj}^{(5)} / l=1, \dots, 8; j=0, 5, \dots, 90\}$ , repitiéndolo posteriormente a intervalos de un minuto durante un periodo de observación de una hora  $\{X_{lj}^{(1)} / l=1, \dots, 8; j=0, 1, \dots, 60\}$ . Inicialmente, consideraron que todas las series procedían de un mismo proceso estacionario lineal general  $\{X_t / t \in Z^+\}$ ,

observado sobre diferentes individuos, y por lo tanto, la densidad espectral es común a todos ellos. De acuerdo con el modelo de efectos fijos, se obtiene fácilmente que para cada frecuencia  $\omega_j$ , el coeficiente de variación de los periodogramas  $I_l(\omega_j)$   $l=1, \dots, r$  es aproximadamente:

$$CV(I_l(\omega_j)) = \frac{\sqrt{\text{Var}(I_l(\omega_j))}}{E[I_l(\omega_j)]} \approx \frac{f(\omega_j)}{f(\omega_j)} = 1 \quad (3.3.1)$$

Al disponer de series replicadas, este coeficiente de variación puede estimarse alternativamente, estimando  $E[I_l(\omega_j)]$  por  $\frac{1}{r} \sum_{i=1}^r I_i(\omega_j)$  y  $\text{Var}[I_l(\omega_j)]$  por  $\frac{1}{r-1} \sum_{i=1}^r (I_i(\omega_j) - \bar{I}(\omega_j))^2$ . De acuerdo con la teoría asintótica del periodograma, si representamos los coeficientes de variación estimados frente a las distintas frecuencias, la gráfica correspondiente debe oscilar alrededor de uno. Diggle y Al-Wasel (1993) obtienen los coeficientes de variación de los periodogramas de las series observadas a intervalos de cinco y de un minuto, observando que los coeficientes de variación de los primeros toman valores alrededor de 1, mientras que para las segundas series, las oscilaciones se producen claramente por encima de uno. Este último resultado, prueba que el modelo de efectos fijos es inconsistente con los datos, advirtiéndose que hay una fuente de variabilidad superpuesta a la variabilidad del periodograma. Las representaciones gráficas de las series de cada uno de los individuos sugieren a su vez que éstas pueden proceder de diferentes procesos, por lo que proponen un modelo alternativo donde cada individuo posee su propia función de densidad espectral  $f_l(\omega)$ . De acuerdo con estos razonamientos, proponen el modelo:

$$I_l(\omega_j) = f_l(\omega_j) U_{lj} \quad (3.3.2)$$

Según los referidos resultados asintóticos acerca del periodograma debido a Hartigan (1980), y de manera análoga al modelo de efectos fijos, las variables aleatorias  $U_{lj}$  son independientes en  $l$ , asintóticamente independientes en  $j$  y con



distribución común exponencial de parámetro 1. Diggle y Al-Wasel consideran que cada  $f_l(\omega)$  es una realización de un proceso  $R(\omega)$  y definen  $f(\omega)=E[R(\omega)]$  denominándolo *espectro poblacional*. Se tiene entonces:

$$f_l(\omega) = f(\omega)Z_l(\omega) \quad (3.3.3)$$

donde  $\{Z_l(\omega)\}_{l=1}^r$  son  $r$  realizaciones independientes de un proceso estocástico  $\{Z(\omega)\}$ . Utilizando (3.3.3) podemos expresar (3.3.2) como

$$I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj} \quad l=1, \dots, r; j=1, \dots, [(N-1)/2] \quad (3.3.4)$$

Este modelo puede considerarse como un modelo de efectos aleatorios, donde el proceso  $Z(\omega)$  expresa la variabilidad entre los objetos y las variables  $U_{lj}$  la variabilidad residual.

### 3.4 Procesos lineales de coeficientes aleatorios.

De acuerdo con el teorema 1.8 del epígrafe 1.5, un proceso  $\{X_t / t \in Z^+\}$  que posee un espectro absolutamente continuo, es un proceso lineal general, y por esta razón, se puede expresar por

$$X_t = \sum_{k=-\infty}^{\infty} b_k \xi_{t-k} \quad (3.4.1)$$

donde  $\{\xi_t\}$  son variables aleatorias independientes e idénticamente distribuidas con  $E[\xi_t]=0$  y  $E[\xi_t^2]=\sigma_\xi^2$ , verificando los coeficientes  $b_k$  la condición

$$\sum_{k=-\infty}^{\infty} b_k^2 < \infty.$$

El modelo (3.3.4) para el periodograma individual, corresponde de hecho a un proceso doblemente estocástico. Dado que cada serie individual procede de la observación de un proceso estacionario con distribución espectral absolutamente continua, tal proceso individual es un proceso lineal, por lo que el proceso global puede considerarse como un proceso lineal de coeficientes aleatorios. En definitiva, la selección de cada individuo es equivalente a seleccionar los coeficientes del proceso de acuerdo con una determinada distribución de probabilidad. Desde el punto de vista espectral, esta selección es equivalente a la observación del proceso  $Z(\omega)$ .

Se puede expresar el proceso en los siguientes términos. Para un objeto  $O_l$

$$X_{l,t} = X_t(O_l) = \sum_{k=-\infty}^{\infty} b_k(O_l) \xi_{t-k}(O_l) \quad (3.4.2)$$

donde

$$\sum_{k=-\infty}^{\infty} b_k^2 < \infty \quad (\text{a.s})^1$$

y donde cada sucesión de variables aleatorias  $\{\xi_{l,t}\}$  verifican  $E[\xi_{l,t}] = 0$  y  $E[\xi_{l,t}^2] = \sigma_\xi^2$ , son independientes en  $l$  y  $t$ , además de independientes del proceso  $\{b_k\}_{k \in \mathbb{Z}}$ .

### 3.5 El modelo de ruido blanco.

A lo largo de este epígrafe desarrollamos el siguiente modelo de ruido blanco propuesto por Diggle y Al-Wasel (1993):

$$I_l(\omega_j) = f(\omega_j) Z_l(\omega_j) U_{lj} \quad l=1, \dots, r; j=1, \dots, [(N-1)/2] \quad (3.5.1)$$

<sup>1</sup> del inglés *Almost Sure*

donde  $\{Z_l(\omega)\}_{l=1}^r$  son  $r$  copias independientes de un proceso estocástico  $\{Z(\omega)\}$ , que verifica:

- i)  $E[Z(\omega_j)] = 1$  y  $Var[Z(\omega_j)] = \zeta^2$  para toda frecuencia  $\omega_j$ .
- ii)  $Z(\omega_j)$  y  $Z(\omega_k)$  independientes para todo  $j \neq k$ .

Estableciendo que  $W_{lj} = Z_l(\omega_j)U_{lj}$  podemos expresar (3.5.1) como

$$I_l(\omega_j) = f(\omega_j)W_{lj} \quad (3.5.2)$$

A partir de este modelo, Diggle y Al-Wasel estiman el espectro  $f(\omega)$  realizando ciertas parametrizaciones de los procesos que forman el modelo. En primer lugar suponen una cierta distribución para el proceso  $Z(\omega)$ , de tal forma que el producto  $W_{lj} = Z_l(\omega_j)U_{lj}$  tenga una determinada distribución de probabilidad. La falta de evidencia para establecer el tipo de distribución que siguen las  $W_{lj}$  les lleva a considerar aquélla que simplifica la obtención de un estimador de máxima verosimilitud. Damos a continuación la proposición, en que se basan dichos autores para considerar la distribución de las  $W_{lj}$ .

*Proposición 3.1* Sean  $Z$  y  $U$  dos variables aleatorias con funciones de densidad de probabilidad respectivas

$$f(z) = \{\Gamma(\kappa)\Gamma(1-\kappa)\}^{-1} \kappa(\kappa z)^{\kappa-1}(1-\kappa z)^{-\kappa} \quad \text{para } 0 \leq z \leq \kappa^{-1}$$

para algún  $0 < \kappa < 1$ , y

$$g(u) = e^{-u} \quad \text{para } u \geq 0$$

entonces la variable aleatoria  $W=ZU$  tiene la siguiente función de densidad

$$h(\omega) = \{\Gamma(\kappa)\}^{-1} \kappa^\kappa \omega^{\kappa-1} e^{-\kappa\omega} \quad \text{para } \omega \geq 0$$

Utilizando la proposición 3.1 en el modelo de ruido blanco expresado por (3.5.2) se deduce que las variables aleatorias  $W_{lj}$  son independientes y siguen una distribución gamma con parámetro de escala  $\kappa$  y parámetro de forma  $\kappa$ , para algún valor  $\kappa \leq 1$ . De esta forma, si con  $f_j$  denotamos  $f(\omega_j)$  podemos expresar la función de densidad del periodograma del individuo  $l$ -ésimo  $I_l(\omega_j)$  de la siguiente forma:

$$h_{lj}(y) = \{\Gamma(\kappa)\}^{-1} (\kappa/f_j)^\kappa y^{\kappa-1} e^{-\kappa y/f_j} \quad \text{para } y \geq 0 \quad (3.5.3)$$

El valor observado del periodograma  $l$ -ésimo en la frecuencia  $j$ -ésima  $I_l(\omega_j)$  lo denotamos por  $y_{lj}$ , de esta forma el logaritmo de la verosimilitud de los datos puede entonces expresarse por:

$$l = [(N-1)/2]r \{ \kappa \log \kappa - \text{Log} \Gamma(\kappa) \} - \kappa r \sum_{j=1}^{[(N-1)/2]} \text{Log} f_j + \\ + (\kappa - 1) \sum_{l=1}^r \sum_{j=1}^{[(N-1)/2]} \text{Log} y_{lj} - \kappa \sum_{j=1}^{[(N-1)/2]} f_j^{-1} \sum_{l=1}^r y_{lj} \quad (3.5.4)$$

De aquí se deduce como estimador de máxima verosimilitud del espectro poblacional el periodograma medio

$$\hat{f}(\omega_j) = \bar{y}_j$$

Sustituyendo este estimador en la expresión (3.5.4) y usando el método de Newton-Raphson para obtener el valor de  $\kappa$ , llegamos a:

$$\hat{\kappa} = \frac{1}{[(N-1)/2]} \sum_{j=1}^{[(N-1)/2]} \bar{y}_j^2 / s_j^2$$

De acuerdo con las consideraciones hechas en el modelo de ruido blanco descrito por (3.5.2),  $\hat{f}(\omega_j) = \bar{I}(\omega_j)$  sigue una distribución gamma, de media  $f(\omega_j)$  y parámetro de forma  $r\kappa$ . Utilizando este resultado, Diggle y Al-Wasel construyen

intervalos de confianza de  $f(\omega_j)$  haciendo uso del estimador  $\hat{\kappa}$  de máxima verosimilitud del parámetro  $\kappa$ .

### 3.6 Modelo general de efectos aleatorios.

Los argumentos utilizados por Diggle y Al-Wasel para justificar el modelo de ruido blanco merecen ciertas consideraciones. En primer lugar suponen que el proceso  $Z(\omega)$  está formado por variables independientes. Si el espectro poblacional es una función *suave*, es fácil averiguar el aspecto de las densidades espectrales individuales  $f_l(\omega)=f(\omega)Z_l(\omega)$  bajo la mencionada hipótesis de independencia. Ellos mismos consideran poco verosímil esta hipótesis y plantean formas alternativas para el proceso  $\{Z(\omega)\}$ , como puede ser incorporar una estructura de autocovarianza positiva para el proceso  $\{Z(\omega)\}$ . Tampoco parece muy adecuada la hipótesis de que la distribución de  $Z(\omega)$  no dependa de la frecuencia. Por todo ello consideramos necesario hacer hipótesis más flexibles para el modelo de efectos aleatorios propuesto por Diggle y Al-Wasel (1993). Debido a esto, en el presente epígrafe proponemos una generalización del modelo descrito por (3.3.4).

Consideremos  $r$  series temporales  $\{x_{lt}; l=1, \dots, r; t=1, \dots, N\}$ , correspondientes a procesos lineales observados en los mismos instantes de tiempo. Para cada uno de los procesos supondremos que existe la función de densidad espectral  $f_l(\omega)$ . Para la  $l$ -ésima serie temporal podemos calcular el periodograma  $I_l(\omega)$  de acuerdo con (3.2.1). Teniendo en cuenta las propiedades asintóticas del periodograma dadas por Hartigan (1980), podemos expresar  $I_l(\omega)$  de la siguiente forma:

$$I_l(\omega_j) = f_l(\omega_j)U_{lj}, \quad \omega_j = \frac{2\pi j}{N}; j=1, \dots, N \quad (3.6.1)$$

donde las variables residuales  $U_{lj}$  son independientes en  $l$ , asintóticamente independientes en  $j$  y con distribución común exponenciales de parámetro 1.

Siguiendo a Diggle y Al-Wasel, supondremos que  $f_l(\omega)$  es una realización de un proceso estocástico  $R(\omega)$ , y consideramos el espectro poblacional definido por  $f(\omega)=E[R(\omega)]$ . Podemos expresar la función de densidad espectral del  $l$ -ésimo individuo por  $f_l(\omega)=f(\omega)Z_l(\omega)$ , donde  $Z_l(\omega)$  son copias independientes de un proceso estocástico  $Z(\omega) = R(\omega)/f(\omega)$ . Obviamente, utilizando esto en (3.6.1) se tiene:

$$I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj} \quad (3.6.2)$$

A continuación asumimos una serie de hipótesis sobre el proceso  $\{Z(\omega)\}$

- i)  $\forall \omega, E[Z(\omega)]=1$
- ii)  $Var[Z(\omega)] = \zeta_z^2(\omega) < \infty$
- iii) El proceso  $\{Z(\omega)\}$  es independiente de la sucesión  $\{U_{lj}\}$ .

Si establecemos que  $W_{lj}=Z_l(\omega_j)U_{lj}$ , podemos expresar (3.6.2) como:

$$I_l(\omega_j) = f(\omega_j)W_{lj} \quad l=1, \dots, r; j=1, \dots, [(N-1)/2] \quad (3.6.3)$$

donde para cada frecuencia  $\omega_j$  fija,  $\{W_{lj}\}$  son variables aleatorias independientes e idénticamente distribuidas, con distribución  $F_j$  de media uno y varianza  $Var[W_{lj}] = \sigma_j^2$  finita. Es fácil probar que  $1 \leq \sigma_j^2 < \infty$ , para todo  $j$ . Según se deduce de este modelo, las variables  $W_{lj}$  tienen dos fuentes de variabilidad: una debida a las diferencias entre individuos, que viene dada por  $Z_l(\omega_j)$ ; y otra, procedente de la variable residual  $U_{lj}$ .

Tal y como hemos visto en el capítulo primero, el periodograma no es un buen estimador de la función de densidad espectral de un proceso estacionario, debido a su inconsistencia. Sin embargo, en el caso de series replicadas, el periodograma medio puede considerarse un estimador satisfactorio cuando el número de objetos  $r$  es grande. Las siguientes proposiciones dan las propiedades asintóticas de  $\bar{I}(\omega_j)$  como estimador del espectro poblacional  $f(\omega)$  en el modelo de efectos aleatorios descrito por (3.6.3).

*Proposición 3.2* Bajo las condiciones del modelo general de efectos aleatorios dado por la expresión (3.6.3), el periodograma medio verifica:

$$E[\bar{I}(\omega_j)] = f(\omega_j)$$

$$\text{Var}[\bar{I}(\omega_j)] = \frac{f^2(\omega_j)\sigma_j^2}{r}.$$

*Proposición 3.3* Bajo las consideraciones del modelo (3.6.3) y si además  $\forall \varepsilon > 0, \forall i$ , se cumple que  $P(|W_{ij} - 1| > \varepsilon\sqrt{r}) = o(1/r)$  entonces

$$\sqrt{r} \left\{ \frac{\bar{I}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\} \xrightarrow[r \rightarrow \infty]{d} N(0, \sigma_j).$$

*Demostración*

$\forall j = 1, \dots, [(N-1)/2]$ ,  $\sqrt{r} \left\{ \frac{\bar{I}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\} = \sum_{l=1}^r \frac{W_{lj} - 1}{\sqrt{r}}$ , es por tanto suficiente

probar que,  $\forall \varepsilon > 0$  y  $j = 1, \dots, [(N-1)/2]$

$$P\left(\max_{1 \leq l \leq r} \left| \frac{W_{lj} - 1}{\sqrt{r}} \right| > \varepsilon\right) \xrightarrow[r \rightarrow \infty]{} 0$$

En efecto

$$\begin{aligned}
 P\left(\max_{1 \leq l \leq r} \left| \frac{W_{lj} - 1}{\sqrt{r}} \right| > \varepsilon\right) &= 1 - P\left(\max_{1 \leq l \leq r} |W_{lj} - 1| \leq \varepsilon\sqrt{r}\right) = \\
 &= 1 - \left\{1 - P(|W_{lj} - 1| > \varepsilon\sqrt{r})\right\}^r \xrightarrow{r \rightarrow \infty} 0
 \end{aligned}$$

dato que por hipótesis es  $P(|W_{lj} - 1| > \varepsilon\sqrt{r}) = o(1/r)$ . Se verifica por tanto la proposición. ■

Utilizando esta última proposición podemos construir intervalos de confianza del espectro poblacional  $f(\omega)$  basados en el pivote  $\sqrt{r} \{ \bar{I}(\omega_j) - f(\omega_j) \} / f(\omega_j)$ . Sin embargo, desconocemos la varianza  $\sigma_j^2$  de las variables  $W_{lj}$ , siendo necesario utilizar un estimador de las mismas a partir de las variables observadas  $\hat{W}_{lj} = I_l(\omega_j) / \bar{I}(\omega_j)$ , mediante la expresión  $\hat{\sigma}_j^2 = \sum_{l=1}^r \{ \hat{W}_{lj} - 1 \}^2 / (r-1)$ .

### 3.7 Estimación Bootstrap del espectro poblacional.

Proponemos ahora un método bootstrap para la estimación del espectro poblacional en base a los datos aportados por  $r$  series temporales  $\{x_{lt}; l=1, \dots, r; t=1, \dots, N\}$ . El procedimiento es similar al utilizado por Freedman (1981) para modelos de regresión lineal y por Härdle y Bowman (1988) para modelos de regresión no paramétrica. Utilizamos como estimador del espectro poblacional el periodograma medio definido por:

$$\hat{f}(\omega_j) = \frac{1}{r} \sum_{l=1}^r I_l(\omega_j) = \bar{I}(\omega_j)$$

El modelo general de efectos aleatorios propuesto en el epígrafe anterior viene dado por:



$$I_l(\omega_j) = f(\omega_j)W_{lj} \quad (3.7.1)$$

Para cada frecuencia  $\omega_j$  fija, las variables aleatorias  $W_{lj}$  son independientes e idénticamente distribuidas, con función de distribución  $F_j$  desconocida. Utilizando la metodología bootstrap reemplazamos la distribución  $F_j$  por la distribución empírica  $F_{rj}$  de  $\{W_{lj}\}_{l=1}^r$ . Dado que las variables  $\{W_{lj}\}$  no son observables, utilizando (3.7.1) podemos aproximarlas por:

$$\hat{W}_{lj} = \frac{I_l(\omega_j)}{\hat{f}(\omega_j)} \quad (3.7.2)$$

Obviamente,  $\frac{1}{r} \sum_{l=1}^r \hat{W}_{lj} = 1$ . Llamamos  $\hat{F}_{rj}$  a la distribución empírica correspondiente a  $\{\hat{W}_{lj}\}$ . El paso siguiente es la obtención de las variables bootstrap  $\{\hat{W}_{lj}^*\}$  tomando muestras aleatorias de tamaño  $r$  de la distribución empírica  $\hat{F}_{rj}$ , para cada una de las frecuencias  $\omega_j$ :

$$\hat{I}_l^*(\omega_j) = \hat{I}_l^*(-\omega_j) = \hat{f}(\omega_j)\hat{W}_{lj}^* \quad (3.7.3)$$

y el estimador bootstrap del espectro poblacional,  $\hat{f}^*(\omega)$ , utilizando la expresión

$$\hat{f}^*(\omega_j) = \frac{1}{r} \sum_{l=1}^r \hat{I}_l^*(\omega_j) \quad (3.7.4)$$

que es el periodograma bootstrap medio.

El algoritmo bootstrap para aproximar la distribución de  $\hat{f}(\omega)$ , puede esquematizarse de la siguiente forma.

Paso 1 Las variables  $W_{lj}$  se estiman siguiendo la expresión (3.7.2)

$$\hat{W}_{lj} = \frac{I_l(\omega_j)}{\hat{f}(\omega_j)} \quad l=1, \dots, r ; j=1, \dots, [(N-1)/2]$$

Paso 2 Para cada una de las frecuencias  $\omega_j$ , se toman  $B$  muestras bootstrap  $\{\hat{W}_{1j}^*, \dots, \hat{W}_{rj}^*\}$  de  $\{\hat{W}_{1j}, \dots, \hat{W}_{rj}\}$  y se define el periodograma bootstrap de cada frecuencia mediante:

$$\hat{I}_l^*(\omega_j) = \hat{I}_l^*(-\omega_j) = \hat{f}(\omega_j) \hat{W}_{lj}^*$$

Paso 3 Por último, obtenemos el estimador bootstrap del espectro poblacional

$$\hat{f}^*(\omega_j) = \frac{1}{r} \sum_{l=1}^r \hat{I}_l^*(\omega_j)$$

Obviamente se verifica:

$$E^*[\hat{f}^*(\omega_j)] = \frac{1}{r} \sum_{l=1}^r E^*[\hat{I}_l^*(\omega_j)] = \frac{1}{r} \sum_{l=1}^r E^*[\hat{f}(\omega_j) \hat{W}_{lj}^*] = \hat{f}(\omega_j)$$

Para probar la validez asintótica de este algoritmo bootstrap, como hemos visto en el capítulo segundo, consideramos la distancia de Mallows dada por la expresión (2.2.2). Ahora bien, teniendo en cuenta las propiedades asintóticas que hemos visto del periodograma medio, tomamos como cantidad pivotal real

$$\sqrt{r} \left\{ \frac{\hat{f}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\}$$

y como aproximación bootstrap

$$\sqrt{r} \left\{ \frac{\hat{f}^*(\omega_j) - \hat{f}(\omega_j)}{\hat{f}(\omega_j)} \right\}.$$

En el siguiente teorema se comprueba cómo la distancia de Mallows de estas dos cantidades pivotaes tiende a cero en probabilidad cuando  $r \rightarrow \infty$ .

*Teorema 3.4* Consideremos el modelo general de efectos aleatorios dado por la expresión (3.6.1). Se verifica:

$$d_2 \left( \sqrt{r} \left\{ \frac{\hat{f}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\}, \sqrt{r} \left\{ \frac{\hat{f}^*(\omega_j) - \hat{f}(\omega_j)}{\hat{f}(\omega_j)} \right\} \right) \xrightarrow{r \rightarrow \infty} 0 \text{ en probabilidad}$$

siendo  $d_2$  la métrica de Mallows.

### *Demostración*

Para cada frecuencia  $\omega_j$ , sea  $F_j$  la función de distribución de las variables  $W_{lj}$  y  $F_{jr}$  la función de distribución empírica correspondiente a  $\{W_{1j}, \dots, W_{rj}\}$ . De acuerdo con lo visto en los apartados anteriores, podemos expresar la cantidad pivotal real como

$$\sqrt{r} \left\{ \frac{\hat{f}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\} = \frac{1}{\sqrt{r}} \sum_{l=1}^r (W_{lj} - 1)$$

y análogamente su aproximación bootstrap

$$\sqrt{r} \left\{ \frac{\hat{f}^*(\omega_j) - \hat{f}(\omega_j)}{\hat{f}(\omega_j)} \right\} = \frac{1}{\sqrt{r}} \sum_{l=1}^r (\hat{W}_{lj}^* - 1)$$

Sustituyendo estas aproximaciones en la expresión de la distancia, se tiene:

$$d_2^2 \left( \sqrt{r} \left\{ \frac{\hat{f}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\}, \sqrt{r} \left\{ \frac{\hat{f}^*(\omega_j) - \hat{f}(\omega_j)}{\hat{f}(\omega_j)} \right\} \right) = d_2^2 \left( \frac{1}{\sqrt{r}} \sum_{l=1}^r (W_{lj} - 1), \frac{1}{\sqrt{r}} \sum_{l=1}^r (\hat{W}_{lj}^* - 1) \right)$$

Ahora bien, de acuerdo con el lema 8.6 de Bickel y Freedman (1981), se tiene la siguiente desigualdad

$$d_2^2\left(\frac{1}{\sqrt{r}}\sum_{l=1}^r(W_{lj}-1), \frac{1}{\sqrt{r}}\sum_{l=1}^r(\hat{W}_{lj}^*-1)\right) \leq \frac{1}{r}\sum_{l=1}^r d_2^2(W_{lj}-1, \hat{W}_{lj}^*-1) = d_2^2(W_{lj}, \hat{W}_{lj}^*)$$

Consideremos ahora para cada  $j$ , una sucesión de variables aleatorias  $W'_{lj}$ , independientes y cuya distribución es la distribución empírica  $F_{rj}$ . Es fácil comprobar la siguiente desigualdad:

$$d_2^2(W_{lj}, \hat{W}_{lj}^*) \leq 2d_2^2(W_{lj}, W'_{lj}) + 2d_2^2(W'_{lj}, \hat{W}_{lj}^*)$$

La distancia  $d_2^2(W_{lj}, \hat{W}_{lj}^*)$  tiende a cero en probabilidad, si ambos términos del segundo miembro de la desigualdad tienden a la vez a cero en probabilidad. Utilizando el Lema 8.4 de Bickel y Freedman (1981) la demostración de  $d_2^2(W_{lj}, W'_{lj})$  es inmediata ya que

$$d_2^2(W_{lj}, W'_{lj}) = d_2^2(F_j, F_{rj}) \xrightarrow{r \rightarrow \infty} 0 \quad (\text{c.s.})$$

Para demostrar que  $d_2^2(W'_{lj}, \hat{W}_{lj}^*)$  tiende a cero en probabilidad, acotamos superiormente esta distancia considerando la distribución conjunta de los  $\{W_{lj}\}$  y  $\{\hat{W}_{lj}\}$  que asigna una probabilidad  $1/r$  a cada  $(W_{lj}, \hat{W}_{lj})$

$$d_2^2(W'_{lj}, \hat{W}_{lj}^*) \leq \frac{1}{r}\sum_{l=1}^r (W_{lj} - \hat{W}_{lj})^2$$

Remitiéndonos a las expresiones (3.7.1) y (3.7.2) se tiene que  $W_{lj} = I_l(\omega_j)/f(\omega_j)$  y que  $\hat{W}_{lj} = I_l(\omega_j)/\hat{f}(\omega_j)$ , sustituyendo en la desigualdad anterior, nos queda

$$\frac{1}{r}\sum_{l=1}^r (W_{lj} - \hat{W}_{lj})^2 = \frac{1}{r}\sum_{l=1}^r \left(\frac{I_l(\omega_j)}{f(\omega_j)} - \frac{I_l(\omega_j)}{\hat{f}(\omega_j)}\right)^2 = \left(\frac{\hat{f}(\omega_j) - f(\omega_j)}{\hat{f}(\omega_j)f(\omega_j)}\right)^2 \frac{1}{r}\sum_{l=1}^r I_l^2(\omega_j)$$

Ahora bien, como  $\hat{f}(\omega_j)$  converge en media cuadrática a  $f(\omega_j)$  y

$\frac{1}{r} \sum_{l=1}^r I_l^2(\omega_j)$  converge en probabilidad

$$\frac{1}{r} \sum_{l=1}^r I_l^2(\omega_j) = \frac{f^2(\omega_j)}{r} \sum_{l=1}^r W_{lj}^2 \xrightarrow{p} f^2(\omega_j)(\sigma_j^2 + 1)$$

se tiene entonces:

$$\left( \frac{\hat{f}(\omega_j) - f(\omega_j)}{\hat{f}(\omega_j)f(\omega_j)} \right)^2 \frac{1}{r} \sum_{l=1}^r I_l^2(\omega_j) \xrightarrow{r \rightarrow \infty} 0 \quad \text{en probabilidad.}$$

Con todo lo visto anteriormente podemos concluir que el algoritmo propuesto anteriormente cumple el principio de validación bootstrap, puesto que

$$d_2 \left( \sqrt{r} \left\{ \frac{\hat{f}(\omega_j) - f(\omega_j)}{f(\omega_j)} \right\}, \sqrt{r} \left\{ \frac{\hat{f}^*(\omega_j) - \hat{f}(\omega_j)}{\hat{f}(\omega_j)} \right\} \right) \xrightarrow{r \rightarrow \infty} 0 \quad \text{en probabilidad}$$

■

### 3.8 Simulaciones.

Para ilustrar los resultados anteriores, hemos simulado un proceso de medias móviles de orden dos con coeficientes aleatorios. El proceso tiene la forma:

$$X_t = \phi_0 \varepsilon_t + \phi_1 \varepsilon_{t-1}$$

donde el vector de coeficientes  $(\phi_0, \phi_1)$  es aleatorio y sigue una distribución normal bivalente con vector de media  $[0.2, -4]$  y con matriz de varianza-covarianza  $\begin{bmatrix} 0.01 & -0.007 \\ -0.007 & 0.01 \end{bmatrix}$ . Las  $\{\varepsilon_t\}$  son variables aleatorias independientes,

normalmente distribuidas con media cero y varianza 1, y además, independientes de  $(\phi_0, \phi_1)$ .

Las gráficas 3.1 y 3.2 representan simultáneamente para cada frecuencia el espectro poblacional real, las dos bandas de confianza al 95% obtenidas unas en base a las propiedades asintóticas del periodograma medio y otras al utilizar el algoritmo bootstrap descrito en el epígrafe 3.7. Ambas estimaciones se obtienen en base a series de 60 observaciones por objeto. La gráfica 3.1 está basada en 100 series y las bandas obtenidas contienen para casi todas las frecuencias al espectro poblacional. Esto da idea de que la estimación del espectro poblacional para cada frecuencia es aceptable. Sin embargo, la impresión global de la gráfica es claramente deficiente, pues presenta numerosos picos y valles, lo cual puede conducir a ideas erróneas acerca de la importancia de determinadas bandas de frecuencias. La gráfica 3.2, utiliza un número muy superior de objetos ( $r=495$ ) y presenta un aspecto adecuadamente suave, lo cual es consecuencia de las propiedades teóricas del periodograma medio. Por último, señalar cómo en ambas gráficas las bandas de confianzas bootstrap y las obtenidas por la distribución normal coinciden prácticamente.

El periodograma medio proporciona estimaciones adecuadas del espectro poblacional para cada frecuencia. Sin embargo, los picos y valles del estimador pueden conducir a ideas erróneas acerca de la importancia de determinadas frecuencias. Las técnicas de alisamiento pueden dar una estimación global más adecuada, sobre todo, cuando se dispone de un suficiente número de observaciones por objeto. En definitiva, nos permiten obtener estimadores con menor varianza, aunque introducen un problema de sesgo, que puede ser importante en los extremos de la función  $f$ . Por tal motivo, no son adecuados en los casos en que el valor de  $r$  es mucho mayor que  $N$ , situaciones que son habituales en el campo de la medicina, donde el número de observaciones por individuo suele ser pequeño. Bajo estas circunstancias es más adecuado utilizar el periodograma medio como estimador del espectro poblacional. En el próximo

capítulo trataremos los casos en que es más adecuado la estimación del espectro poblacional utilizando alisamientos mediante núcleos del periodograma medio.

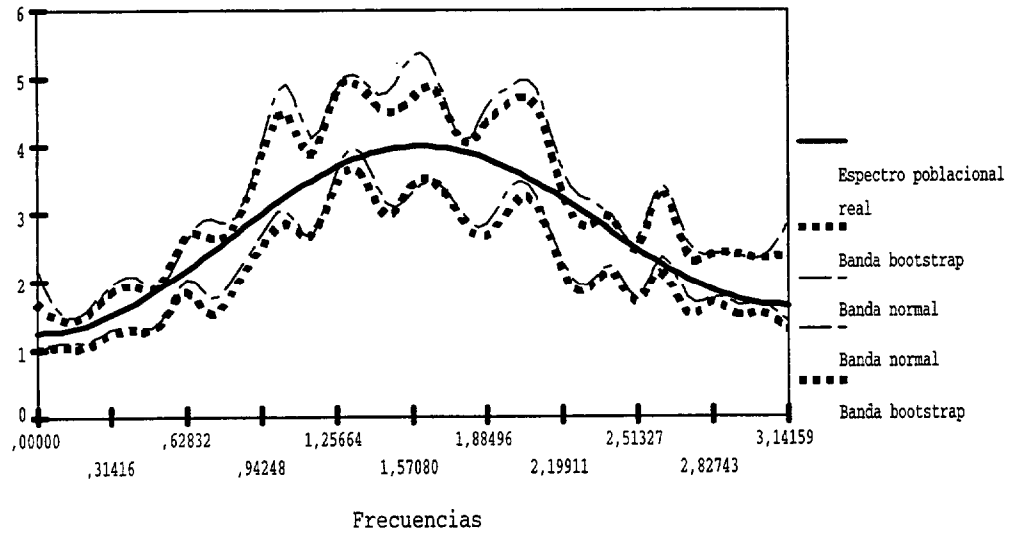


Figura 3.1. Intervalos de confianza al 95% obtenidos al utilizar el algoritmo dado en 3.7. y la proposición 3.3. ( $r=100$ ,  $N=60$ )

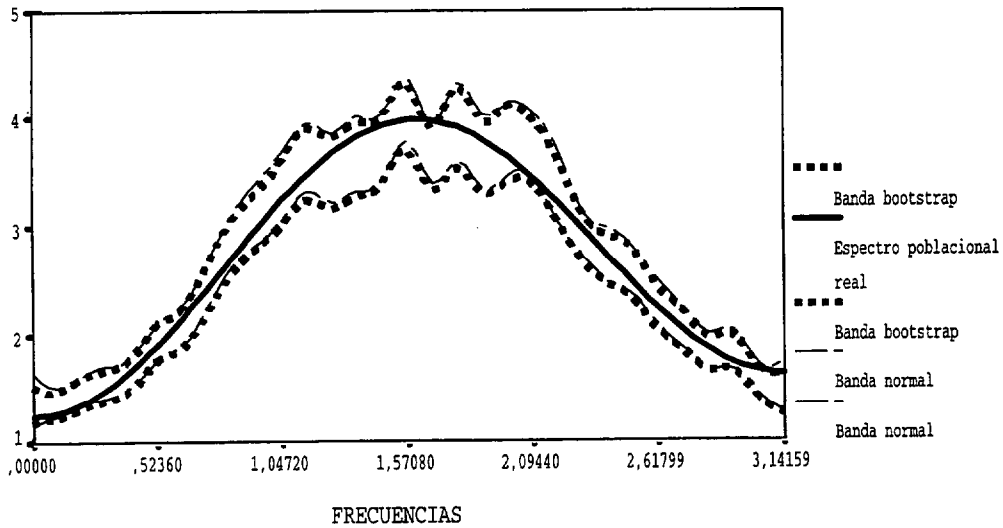


Figura 3.2. Intervalos de confianza al 95% obtenidos al utilizar el algoritmo dado en 3.7. y la proposición 3.4. ( $r=500$ ,  $N=60$ )

## CAPITULO IV. ESTIMACIÓN DEL ESPECTRO POBLACIONAL MEDIANTE ALISAMIENTO DEL PERIODOGRAMA MEDIO.

### 4.1 Introducción

La metodología de estimación del espectro poblacional basada en el periodograma medio conduce a resultados satisfactorios para cada frecuencia individual, al menos cuando se dispone de un amplio número de objetos. Sin embargo, los picos y valles que presenta el estimador, pueden conducir a ideas erróneas acerca de la importancia de determinadas bandas de frecuencia. Las técnicas de alisamiento del periodograma medio pueden ofrecer mejores estimaciones del periodograma en el sentido de eliminar los referidos picos y valles que procedan de la propia naturaleza de los periodogramas. El estimador obtenido a través de los alisamientos tendrá previsiblemente menor varianza que el periodograma medio, aunque introduzca un sesgo, cuyo orden dependerá esencialmente del ancho de banda utilizado.

En el presente capítulo consideramos nuevamente el modelo general de efectos aleatorios y proponemos estimadores del espectro poblacional basados en alisamientos del periodograma medio mediante estimadores de núcleo. Tres parámetros del modelo son esenciales en orden a analizar las propiedades del estimador propuesto, a saber; el número de objetos o series  $r$ , el número de observaciones de cada serie  $N$  y el ancho de banda  $h$ . Bajo determinadas condiciones, daremos resultados acerca del sesgo, varianza y distribución asintótica del estimador. Aproximaremos alternativamente la distribución del estimador mediante el bootstrap. Contemplaremos el problema de la elección del ancho de banda óptimo proponiendo un método de validación cruzada para su



estimación. Finalmente aplicaremos los métodos propuestos a las series obtenidas en el capítulo anterior mediante la simulación de un proceso estacionario lineal de coeficientes aleatorios y compararemos las bandas de confianza obtenidas a través de los resultados asintóticos con aquellas obtenidas por la aproximación bootstrap.

## 4.2 Alisamientos mediante núcleo del periodograma medio.

Consideramos  $r$  series temporales  $\{x_{lt} / l=1, \dots, r ; t=1, \dots, N\}$  observada en los mismos instantes de tiempo. Adoptaremos nuevamente el modelo general de efectos aleatorios, para el  $l$ -ésimo periodograma  $I_l(\omega_j) = f(\omega_j) Z_l(\omega_j) U_{lj}$ , definido en (3.6.2) y sea  $\bar{I}(\omega_j) = (1/r) \sum_{l=1}^r I_l(\omega_j)$  el periodograma medio. El estimador de núcleo del espectro poblacional tendrá entonces la forma:

$$\hat{f}(\omega; h) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \cdot \bar{I}(\omega_j) \quad (4.2.1)$$

donde la función de núcleo  $K(\theta)$  es simétrica, no negativa y verifica:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} K(\theta) d\theta = 1$$

Como es bien sabido en las estimaciones de núcleo, la elección de la función  $K(\theta)$  no es relevante. No así la elección del parámetro de alisamiento  $h$  del que depende fuertemente el estimador  $\hat{f}(\omega; h)$ . Veamos a continuación otra forma de expresar el estimador de núcleo dado por (4.2.1). Denotamos por  $\hat{f}_l(\omega; h)$  el estimador de núcleo de la densidad espectral correspondiente al objeto  $l$ -ésimo,

$$\hat{f}_l(\omega; h) = \frac{1}{Nh} \sum_j K\left(\frac{\omega - \omega_j}{h}\right) I_l(\omega_j); \quad l=1, \dots, r$$

y utilizándolo en la expresión (4.2.1), tenemos:

$$\hat{f}(\omega; h) = \frac{1}{r} \sum_{l=1}^r \hat{f}_l(\omega; h) \quad (4.2.2)$$

Podría pensarse que la estimación del espectro poblacional no es más que el promedio de las estimaciones de cada una de las densidades espectrales individuales. Veremos sin embargo que el parámetro de alisamiento está afectado por el número de series. Tales parámetros podrán ser de menor orden que los habría que emplear para cada estimación individual, lo que supondrá que el estimador global tenga menor sesgo. Por otra parte, la aproximación bootstrap la planteamos *a lo largo de los individuos* y no *a lo largo de los instantes de observación* como hacen Franke y Härdle. Esto como veremos, evita el problema de la dependencia de las variables residuales del periodograma.

Los siguientes resultados dan las propiedades asintóticas ( $N \rightarrow \infty$ ) del estimador  $\hat{f}(\omega; h)$ .

*Proposición 4.1* Consideremos el modelo general de efectos aleatorios descrito en el epígrafe 3.6 y supongamos que el espectro poblacional  $f(\omega)$  es una función no nula y dos veces diferenciable en  $[-\pi, \pi]$  y que  $\hat{f}(\omega; h)$  es un estimador de núcleo definido en (4.2.1). Supongamos además que la función de núcleo  $K$  verifica las condiciones (1.7.2). Si  $N \rightarrow \infty$ ,  $h \rightarrow 0$ ,  $(Nh^4)^{-1} = O(1)$  entonces, para todo  $\omega$  tal que  $|\omega| < \pi$ , el sesgo de  $\hat{f}(\omega; h)$  satisface:

$$E[\hat{f}(\omega; h)] - f(\omega) = \frac{h^2}{2} f''(\omega) + o(h^2) \quad (4.2.4)$$

*Demostración*

De acuerdo con la proposición 3.3, podemos escribir

$$E[\hat{f}(\omega; h)] = \frac{1}{Nh} \sum_j K\left(\frac{\omega - \omega_j}{h}\right) E[\bar{I}(\omega_j)] = \frac{1}{Nh} \sum_j K\left(\frac{\omega - \omega_j}{h}\right) f(\omega_j)$$

utilizando el Lema 1.13 llegamos a (4.2.4). ■

*Lema 4.2* Si  $K$  es una función núcleo verificando las condiciones (1.7.2) y  $p$  una función diferenciable, entonces, uniformemente en  $|\omega| < \pi - \kappa h$  se tiene:

$$\left| \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) p(\omega_j) - \frac{1}{2\pi h} \int_{-\infty}^{\infty} K\left(\frac{\omega - \theta}{h}\right) p(\theta) d\theta \right|$$

$$\leq \frac{c}{Nh} \left\{ \sup_{\theta} |p(\theta)| + h \sup_{\theta} |p'(\theta)| \right\}$$

siendo  $c$  una constante y donde el supremo se calcula en el intervalo  $[\omega - \kappa h, \omega + \kappa h]$ .

Consideremos ahora un proceso lineal de coeficientes aleatorios, de tal forma que dada una sucesión  $\{b_{lk}\}$ , el proceso lineal

$$X_{lt} = \sum_{k=-\infty}^{\infty} b_{lk} \xi_{l,t-k}$$

que satisface las condiciones del teorema 1.14. Asumimos además, que condicionando por  $b_{lk}$ , las funciones de densidad espectral individual  $f_l(\omega)$  son funciones no nulas y dos veces diferenciables en el intervalo  $[-\pi, \pi]$ . Supondremos lo mismo para el espectro poblacional  $f(\omega)$ . Dada la forma del

proceso  $X_{it}$ , el proceso  $Z(\omega)$  es determinístico puro, lo que supone que  $Cov(Z(\omega_j), Z(\omega_k)) = \varphi(\omega_j)\varphi(\omega_k)$ , donde  $\varphi(\omega) = \sqrt{Var[Z(\omega)]}$ .

Calcularemos ahora una expresión aproximada para la varianza en el caso de ser  $h \rightarrow 0$  y  $Nh^2 \rightarrow \infty$ . Tal expresión como veremos tiene la forma aproximada para  $|\omega| < \pi$

$$Var[\hat{f}(\omega; h)] \approx \frac{f^2(\omega)}{Nhr} E[Z^2(\omega)] \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta + \frac{f^2(\omega)}{r} Var[Z(\omega)] \quad (4.2.5)$$

Cada una de las componentes de la varianza tiene un significado preciso. La primera componente corresponde a una varianza *intraclase*. Obsérvese que esta componente es de orden  $(Nhr)^{-1}$ . La segunda componente corresponde a la variabilidad entre los diferentes sujetos o poblaciones. La homogeneidad espectral de estos, supondría en el modelo que la varianza de  $Z(\omega)$  es nula en todas las frecuencias  $\omega$ . Obsérvese también que esta componente, caso de ser  $Var[Z(\omega)] > 0$ , es de orden  $r^{-1}$  y además la varianza global del estimador es también de orden  $r^{-1}$ . Si por el contrario,  $Var[Z(\omega)] = 0$ , dicha varianza global es de orden  $(Nhr)^{-1}$ .

Veamos como podemos obtener la aproximación (4.2.5). Utilizando la expresión (4.2.2), se tiene

$$Var(\hat{f}(\omega; h)) = Var\left(\frac{1}{r} \sum_{l=1}^r \hat{f}_l(\omega; h)\right) = \frac{1}{r} E\left[\left(\hat{f}_l(\omega; h) - E[\hat{f}_l(\omega; h)]\right)^2\right] =$$

Denotamos por  $J_r$  la  $\sigma$ -álgebra generada por las  $r$  sucesiones de variables aleatorias  $\{b_{lk}; l=1, \dots, r; k \in \mathbb{Z}^+\}$ . Podemos entonces escribir:

$$Var[\hat{f}(\omega; h)] = \frac{1}{r} E\left[\left(\hat{f}_l(\omega; h) - E[\hat{f}_l(\omega; h)/J_r] + E[\hat{f}_l(\omega; h)/J_r] - E[\hat{f}(\omega; h)]\right)^2\right] =$$

$$= \frac{1}{r} E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 \right] + \frac{1}{r} E \left[ \left( E \left[ \hat{f}_l(\omega; h) / J_r \right] - E \left[ \hat{f}(\omega; h) \right] \right)^2 \right] +$$

$$+ \frac{2}{r} E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right) \left( E \left[ \hat{f}_l(\omega; h) / J_r \right] - E \left[ \hat{f}(\omega; h) \right] \right) \right] \quad (4.2.6)$$

Pasamos a estudiar cada uno de los términos del segundo miembro de la expresión (4.2.6)

i) El primer término se puede expresar por:

$$\frac{1}{r} E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 \right] = \frac{1}{r} E \left[ E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 / J_r \right] \right]$$

Como se verifican las condiciones del teorema 1.15 podemos concluir cuando  $h \rightarrow 0$  y  $Nh^2 \rightarrow \infty$ , que  $\forall |\omega| < \pi$

$$E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 / J_r \right] \approx \frac{1}{Nh} f_l^2(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta$$

utilizando este resultado en la expresión anterior se tiene

$$\frac{1}{r} E \left[ E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 / J_r \right] \right] \approx \frac{1}{r} E \left[ \frac{1}{Nh} f_l^2(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta \right]$$

y de esta forma podemos concluir que

$$Nh E \left[ \left( \hat{f}_l(\omega; h) - E \left[ \hat{f}_l(\omega; h) / J_r \right] \right)^2 \right] \xrightarrow[Nh^2 \rightarrow \infty]{\substack{(a.s.) \\ h \rightarrow 0}} f^2(\omega) E \left[ Z^2(\omega) \right] \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta$$

ii) Desarrollando la expresión  $\frac{1}{r} E \left[ \left( E \left[ \hat{f}_l(\omega; h) / J_r \right] - E \left[ \hat{f}(\omega; h) \right] \right)^2 \right]$  llegamos a

$$= \frac{1}{rN^2h^2} \sum_j \sum_k K\left(\frac{\omega - \omega_j}{h}\right) K\left(\frac{\omega - \omega_k}{h}\right) f(\omega_j) f(\omega_k) E\left[\left(Z_l(\omega_j) - 1\right)\left(Z_l(\omega_k) - 1\right)\right]$$

Ahora bien, dado que  $Cov(Z(\omega_j), Z(\omega_k)) = \varphi(\omega_j)\varphi(\omega_k)$  y teniendo en cuenta el lema 4.2, podemos hacer la aproximación:

$$\frac{1}{Nh} \sum_j K\left(\frac{\omega - \omega_j}{h}\right) f(\omega_j)\varphi(\omega_j) \approx \frac{1}{2\pi h} \int_{-\infty}^{\infty} K\left(\frac{\omega - \xi}{h}\right) f(\xi)\varphi(\xi) d\xi \quad (4.2.7)$$

utilizando (4.2.7) y la hipótesis de que el espectro poblacional  $f(\omega)$  es dos veces diferenciable; y por otro lado la hipótesis de que  $\psi(\omega)$  es una función diferenciable, se tiene:

$$\frac{1}{r} E\left[\left(E\left[\hat{f}_l(\omega; h)/J_r\right] - E\left[\hat{f}(\omega; h)\right]\right)^2\right] \approx \frac{f^2(\omega)Var[Z(\omega)] + O(h^2)}{r}$$

iii) Por último, desarrollando la expresión

$$\begin{aligned} & \frac{2}{r} E\left[\left(\hat{f}_l(\omega; h) - E\left[\hat{f}_l(\omega; h)/J_r\right]\right)\left(E\left[\hat{f}_l(\omega; h)/J_r\right] - E\left[\hat{f}(\omega; h)\right]\right)\right] = \\ & = \frac{1}{r} E\left[\left(\frac{1}{Nh} \sum_j K\left(\frac{\omega - \omega_j}{h}\right) f(\omega_j) Z_l(\omega_j) (U_{lj} - 1)\right)\left(\frac{1}{Nh} \sum_k K\left(\frac{\omega - \omega_k}{h}\right) f(\omega_k) (Z_l(\omega_k) - 1)\right)\right] = \\ & = \frac{1}{rN^2h^2} \sum_j \sum_k K\left(\frac{\omega - \omega_j}{h}\right) K\left(\frac{\omega - \omega_k}{h}\right) f(\omega_j) f(\omega_k) E\left[Z_l(\omega_j) (U_{lj} - 1) (Z_l(\omega_k) - 1)\right] = 0 \end{aligned}$$

Dado que al ser  $Z_l(\omega)$  independiente de  $U_{lj}$ , se verifica:

$$E\left[Z_l(\omega_j) (Z_l(\omega_k) - 1) (U_{lj} - 1)\right] = E\left[Z_l(\omega_j) (Z_l(\omega_k) - 1)\right] E\left[U_{lj} - 1\right] = 0$$

Sustituyendo todo esto en la expresión (4.2.6) se llega a (4.2.5)

*Teorema 4.3* Consideremos un proceso lineal de coeficientes aleatorios, de tal forma que dada una sucesión  $\{b_{lk}\}$ , el proceso lineal

$$X_{ln} = \sum_{k=-\infty}^{\infty} b_{lk} \xi_{l,n-k}$$

verifica las condiciones del teorema 1.14. Asumimos además, que condicionando por  $\{b_{lk}\}$ , las funciones de densidad espectral individual  $f_l(\omega)$  son funciones no nulas y dos veces diferenciables en el intervalo  $[-\pi, \pi]$ . Supondremos lo mismo para el espectro poblacional  $f(\omega)$ . Consideremos también que  $\{Z(\omega)\}$  es un proceso determinístico puro y  $\varphi(\omega) = \sqrt{\text{Var}[Z(\omega)]}$  una función diferenciable. Entonces,

para  $h \rightarrow 0$  y  $Nh^2 \rightarrow \infty$  y para todo  $\omega$  tal que  $|\omega| < \pi$ , se tiene:

$$r \text{Var}[\hat{f}(\omega; h)] \rightarrow f^2(\omega) \text{Var}[Z(\omega)] \quad (4.2.8)$$

y

$$Nhr E \left[ E \left[ \left\{ \hat{f}_l(\omega; h) - E[\hat{f}_l(\omega; h)/J_r] \right\}^2 / J_r \right] \right] \xrightarrow{(a.s.)} f^2(\omega) E[Z^2(\omega)] \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta \quad (4.2.9)$$

*Teorema 4.4* Consideremos las condiciones del modelo general de efectos aleatorios descrito en el epígrafe 3.6. y supongamos que el espectro poblacional  $f(\omega)$  es una función acotada,  $\forall \omega$ ,  $\text{Var}[Z(\omega)] > 0$ , y  $\forall l, j$   $E[(W_{lj} - 1)^4] = C < \infty$ .

Entonces, para  $h \rightarrow 0$ ,  $Nh^2 \rightarrow \infty$ ,  $r \rightarrow \infty$ ,

$$\sqrt{r} \left\{ \hat{f}(\omega; h) - E[\hat{f}(\omega; h)] \right\} \xrightarrow{d} N(0, f^2(\omega) \text{Var}[Z(\omega)])$$

*Demostración*

$$\sqrt{r}\{\hat{f}(\omega;h) - E[\hat{f}(\omega;h)]\} = \sum_{l=1}^r \left\{ \frac{\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]}{\sqrt{r}} \right\}$$

es obvio que  $\forall \varepsilon > 0$ ,  $\lim_{N \rightarrow \infty} \max_{1 \leq l \leq r} P \left( \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} > \varepsilon \right) = 0$ . Ello supone que

el pivote  $\sqrt{r}\{\hat{f}(\omega;h) - E[\hat{f}(\omega;h)]\} \xrightarrow{d} Z$ , siendo  $Z$  una distribución infinitamente divisible. Esto se deduce al utilizar el teorema 14.9.b., pág. 373, de Burrill (1972), donde  $Z$  sigue una distribución normal si y sólo si se verifica:

$$\lim_{N \rightarrow \infty} P \left( \max_{1 \leq l \leq r} \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} > \varepsilon \right) = 0 \quad (4.2.10)$$

En efecto,

$$\begin{aligned} P \left( \max_{1 \leq l \leq r} \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} > \varepsilon \right) &= 1 - \prod_{l=1}^r P \left( \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} \leq \varepsilon \right) \\ &= 1 - \prod_{l=1}^r \left\{ 1 - P \left( \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} > \varepsilon \right) \right\} \end{aligned}$$

Es obvio que la expresión (4.2.10) es válida si y sólo si

$$\lim_{N \rightarrow \infty} \prod_{l=1}^r \left\{ 1 - P \left( \frac{|\hat{f}_l(\omega;h) - E[\hat{f}_l(\omega;h)]|}{\sqrt{r}} > \varepsilon \right) \right\} = 1 \quad (4.2.11)$$

Utilizando la siguiente desigualdad [Sección 13-6, Burrill (1972)]:



$$1 - \sum_{v=1}^n b_v \leq \prod_{v=1}^n (1 - b_v) \leq \exp\left(-\sum_{v=1}^n b_v\right)$$

es obvio que (4.2.11) ocurre si y sólo si

$$\lim_{N \rightarrow \infty} \sum_{i=1}^r P\left\{\left|\hat{f}_i(\omega; h) - E\left[\hat{f}_i(\omega; h)\right]\right| \geq \varepsilon\sqrt{r}\right\} = 0 \quad (4.2.12)$$

En efecto, ya que:

$$\begin{aligned} P\left\{\left|\hat{f}_i(\omega; h) - E\left[\hat{f}_i(\omega; h)\right]\right| \geq \varepsilon\sqrt{r}\right\} &\leq \frac{E\left[\left(\hat{f}_i(\omega; h) - E\left[\hat{f}_i(\omega; h)\right]\right)^4\right]}{\varepsilon^4 r^2} \leq \\ &\leq \frac{1}{\varepsilon^4 r^2} E\left[(W_{i_0 j_0} - 1)^4\right] \leq \frac{C}{\varepsilon^4 r^2} \end{aligned}$$

### 4.3 Aproximación Bootstrap para la estimación del espectro poblacional utilizando como estimador alisamientos del periodograma medio.

Consideremos el modelo general de efectos aleatorios  $I_i(\omega_j) = f(\omega_j)W_{ij}$ , descrito en 3.6, con  $\text{Var}[Z(\omega)] > 0$  para todo  $\omega$ . Ello garantiza, bajo las condiciones señaladas, la varianza de  $\hat{f}(\omega; h)$  es de orden  $r^{-1}$  en todas las frecuencias  $\omega$ . Proponemos el siguiente algoritmo bootstrap para aproximar la distribución del pivote

$$\sqrt{r} \left\{ \frac{\hat{f}(\omega; h) - f(\omega)}{f(\omega)} \right\} \quad (4.3.1)$$

Paso 1 Las variables no observables  $W_{lj}$ , de acuerdo con (3.6.3), se aproximan por:

$$\hat{W}_{lj} = \frac{I_l(\omega_j)}{\bar{I}(\omega_j)} \quad l=1, \dots, r; j=1, \dots, [N/2]$$

Paso 2 Se selecciona un ancho de banda de remuestreo  $g$  y para cada una de las frecuencias  $\omega_j$ , se toman  $B$  muestras bootstrap  $\{\hat{W}_{1j}^*, \dots, \hat{W}_{rj}^*\}$  de la distribución empírica correspondiente a la muestra  $\{\hat{W}_{1j}, \dots, \hat{W}_{rj}\}$ . Para cada una de ellas, se obtiene el periodograma bootstrap del  $l$ -ésimo individuo de la siguiente forma:

$$I_l^*(\omega_j) = I_l^*(-\omega_j) = \hat{f}(\omega_j; g) \hat{W}_{lj}^*$$

Paso 3 Se elige finalmente un ancho de banda  $h$  y se obtiene la estimación bootstrap del espectro poblacional por:

$$\hat{f}^*(\omega_j; h, g) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \bar{I}^*(\omega_j) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \hat{f}(\omega_j; g) \hat{W}_{lj}^*$$

Utilizando este algoritmo aproximamos el pivote dado por (4.3.1) a través del siguiente pivote bootstrap

$$\sqrt{r} \left\{ \frac{\hat{f}^*(\omega; h, g) - \hat{f}(\omega; g)}{\hat{f}(\omega; g)} \right\} \quad (4.3.2)$$

#### 4.4 Ancho de banda óptimo.

Sin duda, el parámetro de alisamiento es un elemento clave en cualquier técnica de alisamiento. Propondremos en este epígrafe la forma de elección de un

ancho de banda global óptimo basado en el porcentaje de error cuadrático medio (MSPE, Mean Square Percentage Error). De acuerdo con Parzen (1957), y como vimos en el capítulo primero consideramos:

$$MSPE(\omega; h) = \frac{E\{\hat{f}(\omega; h) - f(\omega)\}^2}{f^2(\omega)} \quad (4.4.1)$$

(4.4.1) puede expresarse alternativamente por:

$$MSPE(\omega; h) = \frac{v^2(\omega) + b^2(\omega)}{f^2(\omega)} \quad (4.4.2)$$

donde  $b(\omega) = (E[\hat{f}(\omega; h)] - f(\omega))$  y  $v^2(\omega) = Var[\hat{f}(\omega; h)]$ .

De acuerdo a los resultados asintóticos obtenidos en 4.2.

$$E[\hat{f}(\omega, h)] - f(\omega) \approx \frac{h^2}{2} f''(\omega) \quad (4.4.3)$$

$$Var[\hat{f}(\omega, h)] \approx \frac{f^2(\omega)}{Nhr} E[Z^2(\omega)] \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta + \frac{f^2(\omega)}{r} Var[Z(\omega)] \quad (4.4.4)$$

Al sustituir (4.4.3) y (4.4.4) en (4.4.2), se tiene:

$$MSPE(\omega, h) \approx \left\{ \frac{1}{2} h^2 \frac{f''(\omega)}{f(\omega)} \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta^2 K(\theta) d\theta \right\}^2 + \frac{E[Z^2(\omega)]}{Nhr} \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta + \frac{Var[Z(\omega)]}{r} \quad (4.4.5)$$

Al minimizar esta expresión, obtenemos el ancho de banda  $h$  óptimo para cada frecuencia  $\omega$ , que es de la siguiente forma

$$h(\omega) = \left( \frac{E[Z^2(\omega)]}{2\pi Nr} \int_{-\infty}^{\infty} K^2(\theta) d\theta \right)^{1/5} \left( \frac{f(\omega)}{f''(\omega)} \right)^{2/5} \quad (4.4.6)$$

Como consecuencia de esta expresión obtenemos un ancho de banda  $h$  de orden  $(Nr)^{-1/5}$ , que es óptimo para cada frecuencia.

Podemos alternativamente determinar un ancho de banda global en un sentido análogo al propuesto por Beltrao y Bloomfield (1987) basado en la media del porcentaje del error cuadrático (AMSPE Average Mean Square Percentage Error)

$$AMSPE(h) = \frac{1}{[(N-1)/2]} \sum_{j=1}^{[(N-1)/2]} MSPE(\omega_j, h) \quad (4.4.7)$$

de acuerdo con esto, el ancho de banda global óptimo es de la forma:

$$h = \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} K^2(\theta) d\theta / Nr \right)^{1/5} \left\{ \frac{\sum_{j=1}^{[(N-1)/2]} E[Z^2(\omega_j)]}{\sum_{j=1}^{[(N-1)/2]} \{f''(\omega_j)/f(\omega_j)\}^2} \right\}^{1/5} \quad (4.4.8)$$

El algoritmo descrito en el epígrafe anterior utiliza dos anchos de banda; a saber: el ancho de banda de remuestreo  $g$ , y el ancho de banda para el bootstrap  $h$ . Inicialmente podríamos considerar seleccionar ambos parámetros  $g$  y  $h$  del orden  $(Nr)^{-1/5}$ . En previsión de que tal elección para  $g$  pudiera no aproximar bien el sesgo real, como ocurre en la estimación de modelos de regresión no paramétrica (Härdle y Bowman (1988)) y en la estimación de la densidad espectral (Franke y Härdle (1992)), podemos estimar separadamente el sesgo, mediante la estimación de la segunda derivada del espectro poblacional. Tal estimación tiene la forma:

$$\hat{f}''(\omega, l) = \frac{1}{Nl^3} \sum_{j=-[N/2]}^{[N/2]} K_{(2)}\left(\frac{\omega - \omega_j}{h_2}\right) \bar{I}(\omega_j) \quad (4.4.9)$$

donde el núcleo  $K_{(2)}(\theta)$  satisface las siguientes condiciones:

$$\int_{-\infty}^{\infty} K_{(2)}(u)du = \int_{-\infty}^{\infty} uK_{(2)}(u)du = 0 \quad \int_{-\infty}^{\infty} u^2 K_{(2)}(u)du = 2 \quad (4.4.10)$$

y el ancho de banda  $l$  de forma que  $Nl^5 \rightarrow \infty$  cuando  $l \rightarrow 0$ .

De acuerdo con Härdle y Bowman, consideramos el pivote bootstrap alternativo

$$\sqrt{r} \left\{ \frac{\hat{f}^*(\omega, h, g) - \hat{f}_c(\omega, h, g) + \frac{h^2}{2} \hat{f}''(\omega, l)}{\hat{f}(\omega, g)} \right\} \quad (4.4.11)$$

donde

$$\hat{f}_c(\omega, h, g) = E^*[\hat{f}^*(\omega, h, g)] = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \hat{f}(\omega_j; g)$$

y  $\hat{f}''(\omega; l)$  viene dado por la expresión (4.4.9).

## 4.5 Estimación del ancho de banda óptimo.

De la expresión (4.4.8) se deduce que el ancho de banda óptimo depende de varios parámetros como son  $f''(\omega)$ ,  $f(\omega)$ , y  $E[Z^2(\omega)]$  que son desconocidos.

Por ello pasamos a considerar dos métodos alternativos para obtener una estimación del ancho de banda óptimo.

1.- Método Plug-in.

Este método estima el ancho de banda óptimo  $h$  al sustituir en la expresión (4.4.8) los parámetros desconocidos por estimaciones realizadas con los mismos datos. Evidentemente, la presencia de varios parámetros y su necesaria estimación para obtener mediante el método Plug-in el valor de  $h$ , hacen pensar que el método pueda ser inadecuado. Pensamos que en este caso, los métodos de validación cruzada puedan ser de mayor efectividad como veremos a continuación.

2.- Método de validación cruzada.

Damos a continuación un método de validación cruzada para la estimación del ancho de banda óptimo y una justificación del mismo.

*Paso 1.* Se divide la muestra de tamaño  $r$  en dos grupos de tamaños  $r_1$  y  $r_2$ .

*Paso 2.* Estimamos  $f(\omega)$  de la siguiente forma, para el grupo 1 mediante estimadores de núcleo

$$\hat{f}^{(r_1)}(\omega; h) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega - \omega_j}{h}\right) \bar{I}^{(r_1)}(\omega_j)$$

Para el segundo grupo, estimamos el espectro poblacional utilizando el periodograma medio  $\bar{I}^{(r_2)}(\omega)$ .

*Paso 3.* Se obtiene  $h$  que minimiza la expresión

$$\min_h \sum_{j=1}^{[N/2]} \left\{ \bar{I}^{(r_2)}(\omega_j) - \hat{f}^{(r_1)}(\omega_j; h) \right\}^2 \quad (4.5.1)$$

Seguidamente veremos la justificación teórica de este procedimiento para la elección de un ancho de banda óptimo. Un criterio natural de selección de

ancho de banda óptimo es elegir aquel que minimice la media de los errores cuadráticos medios de cada una de las frecuencias

$$\frac{1}{[N/2]} \sum_{j=1}^{[N/2]} E \left[ \left( \hat{f}(\omega_j; h) - f(\omega_j) \right)^2 \right] \quad (4.5.2)$$

Veamos que minimizar esta (4.5.2) es aproximadamente equivalente a minimizar

$$\frac{1}{[N/2]} \sum_{j=1}^{[N/2]} E \left[ \left( \hat{f}^{(\tau)}(\omega_j; h) - \bar{I}^{(\tau)}(\omega_j) \right)^2 \right] \quad (4.5.3)$$

En efecto, como  $f(\omega_j)$  no es conocida, podemos considerar una estimación basada en el periodograma medio, obtenido con datos diferentes a los utilizados para estimar  $\hat{f}^{(\tau)}(\omega_j; h)$ . Sea  $\bar{I}^{(\tau)}(\omega_j)$  el periodograma medio referido. Se tiene entonces:

$$\begin{aligned} E \left[ \sum_{j=1}^{[N/2]} \left\{ \bar{I}^{(\tau)}(\omega_j) - \hat{f}^{(\tau)}(\omega_j; h) \right\}^2 \right] &= E \left[ \sum_{j=1}^{[N/2]} \left\{ \bar{I}^{(\tau)}(\omega_j) - f(\omega_j) \right\}^2 \right] + \\ &+ E \left[ \sum_{j=1}^{[N/2]} \left\{ f(\omega_j) - \hat{f}^{(\tau)}(\omega_j; h) \right\}^2 \right] \end{aligned}$$

Dado que  $E \left[ \sum_{j=1}^{[N/2]} \left\{ \bar{I}^{(\tau)}(\omega_j) - f(\omega_j) \right\}^2 \right]$  es independiente del ancho de banda  $h$ ,

queda establecido que la minimización (4.5.2) es aproximadamente equivalente a minimizar (4.5.3), lo cual justifica el método de validación cruzada propuesto para la selección del parámetro de alisamiento.

## 4.6 Simulaciones.

Los métodos propuestos en los epígrafes anteriores se aplican a las series obtenidas en el epígrafe 3.8 mediante la simulación de un proceso de medias móviles de orden dos con coeficientes aleatorios. El núcleo utilizado para la estimación del espectro poblacional mediante alisamientos del periodograma medio es el núcleo de Bartlett-Priestley, que viene dado por:

$$K(\theta) = \begin{cases} \frac{3}{2}(1 - (\theta/\pi)^2) & |\theta| \leq \pi \\ 0 & |\theta| > \pi \end{cases}$$

En la figura 4.1 utilizamos el algoritmo bootstrap descrito en el epígrafe 4.3 para obtener el intervalo de confianza bootstrap al 95% del espectro poblacional. Las series utilizadas son exactamente las mismas que las de la gráfica 3.1 del epígrafe 3.8. Se ve como el problema de sobre valoración de frecuencias que aparecía en la gráfica 3.5 se ha resuelto al hacer alisamientos del periodograma medio, de forma que obtenemos una mejor estimación global del espectro poblacional.

La gráfica 4.2 se obtiene en base a 60 series de 200 observaciones por objeto. Para cada frecuencia se han representado simultáneamente el verdadero espectro poblacional  $f(\omega)$ , el periodograma medio  $\bar{I}(\omega)$  y un estimador de núcleo obtenido mediante alisamientos del periodograma medio  $\hat{f}(\omega; h)$  dado por la expresión (4.2.1) con  $h=0.112$ . Las gráficas 4.3 y 4.4 se obtienen a partir de las mismas series que la gráfica 4.2. El intervalo bootstrap del 95% de la gráfica 4.3 se obtiene al aproximar el pivote real  $\sqrt{r} \{ \hat{f}(\omega, h) - f(\omega) \} / f(\omega)$  por el pivote bootstrap  $\sqrt{r} \{ \hat{f}^*(\omega; h, g) - \hat{f}(\omega; g) \} / \hat{f}(\omega; g)$ , la banda obtenida es suficientemente suave aunque se observa un cierto sesgo. No obstante, dicho problema de sesgo se soluciona en la figura 4.4 al utilizar el pivote bootstrap corregido



$\sqrt{r} \left\{ \hat{f}^*(\omega; h, g) - \hat{f}_c(\omega; h, g) + \frac{h^2}{2} \hat{f}''(\omega; l) \right\} / \hat{f}(\omega; g)$ . Las gráficas 4.5 y 4.6

utilizan el mismo número de series que las anteriores pero un número muy superior de observaciones por objeto ( $N=500$ ). De forma análoga a 4.3 y 4.4 se comparan los intervalos bootstrap construidos con los pivotes descritos anteriormente y además se comparan con el intervalo de confianza obtenido al utilizar los resultados asintóticos del teorema 4.4. Destacar como ambas bandas coinciden prácticamente.

Por último en la tabla 1 se comparan los anchos de banda obtenidos por los métodos de validación cruzada y Plug-in antes descritos, así como los errores AMSPE, deduciéndose la eficacia del algoritmo de validación cruzada.

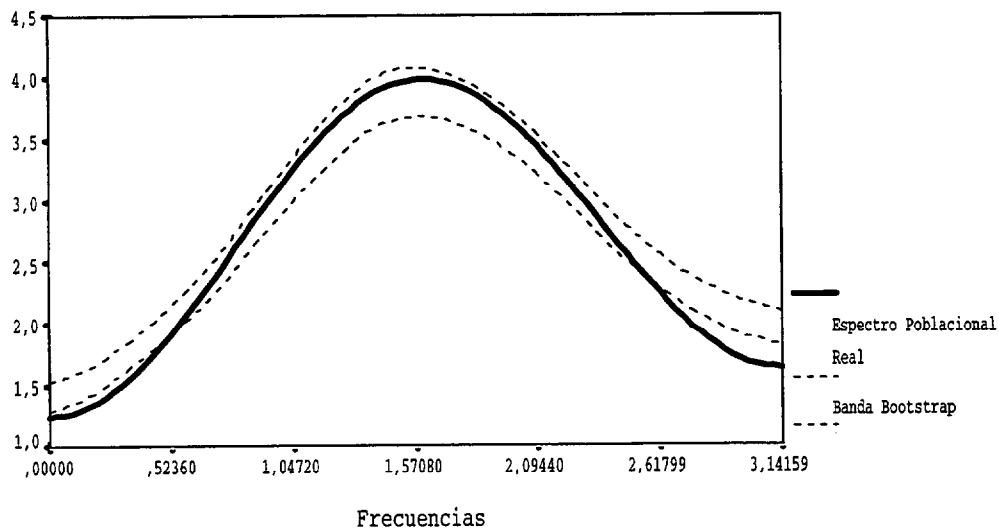


Figura 4.1. Intervalo bootstrap al 95% obtenido con el algoritmo descrito en 4.3 ( $r=100$ ,  $N=60$ ,  $h=0.3$  y  $g=0.25$ )

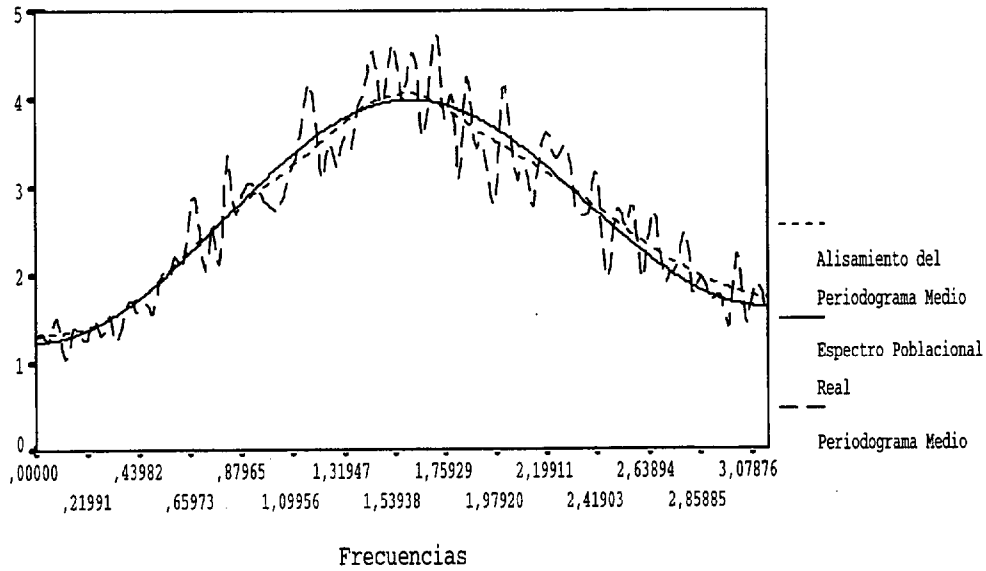


Figura 4.2

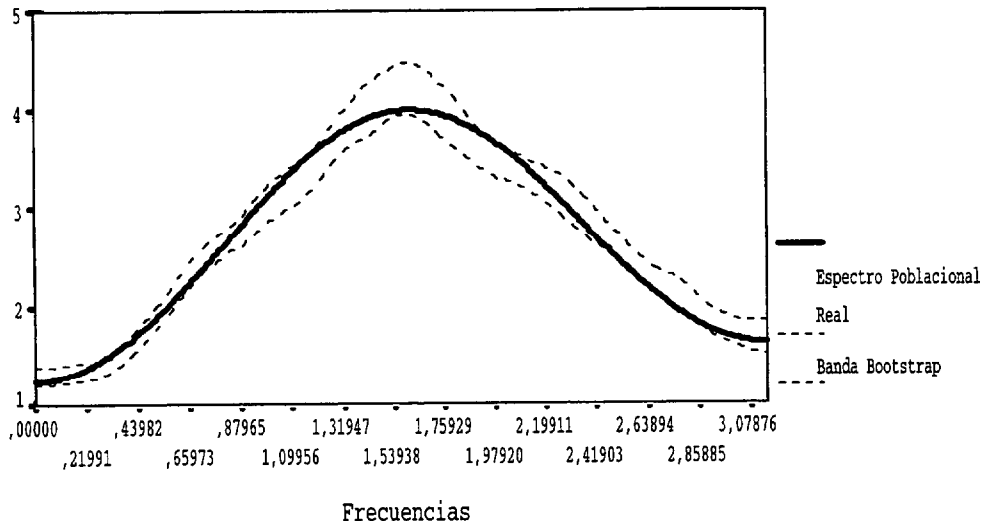


Figura 4.3 Intervalo Bootstrap al 95% obtenido con el algoritmo descrito en 4.3 tomando anchos de banda óptimos ( $r=60$ ,  $N=200$ ,  $h=0.112$  y  $g=0.09$ )

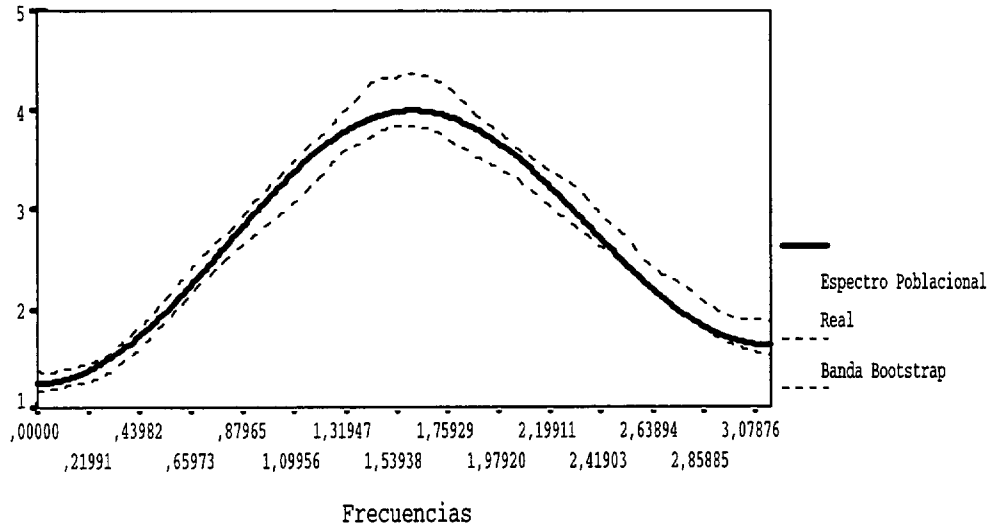


Figura 4.4 Intervalo al 95% obtenido al corregir el algoritmo bootstrap dado por 4.3 según Härdle y Bowman ( $r=60$ ,  $N=200$ ,  $h=0.112$  y  $g=0.112$ )

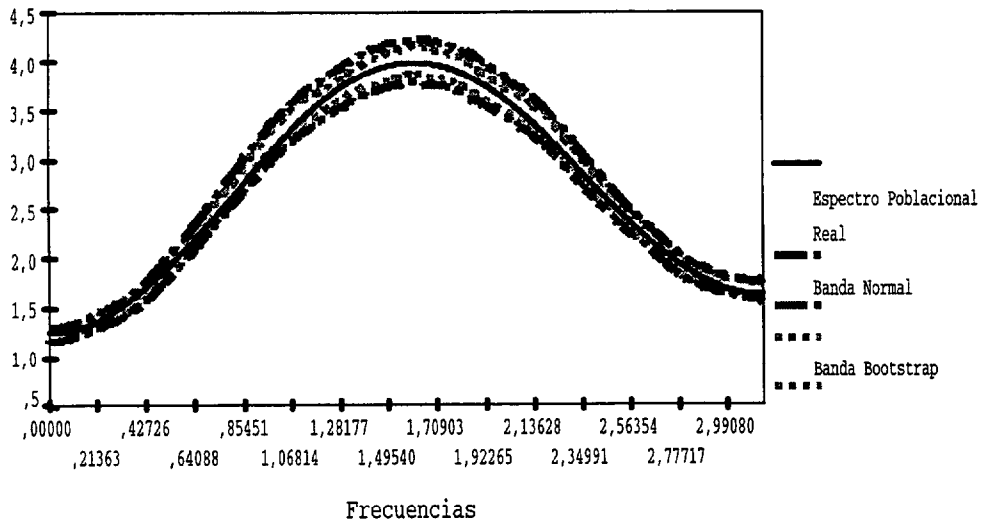


Figura 4.5 Intervalo Bootstrap al 95% obtenido con el algoritmo descrito en 4.3 ( $r=100$ ,  $N=500$ ,  $h=0.076$  y  $g=0.05$ )

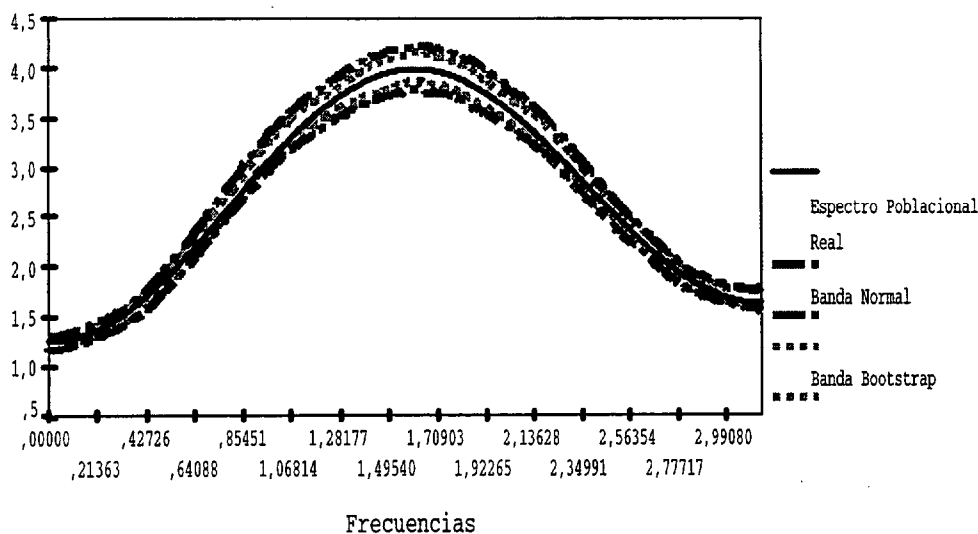


Figura 4.6 Intervalo al 95% obtenido al corregir el sesgo del algoritmo propuesto en 4.3 ( $r=100$ ,  $N=500$ ,  $h=0.076$  y  $g=0.076$ )

Tabla 1

$N$	60	100	200	500
30	$hcv=0.16700$ $AMSPEcv=1.32217$ $hplug=0.42532$ $AMSPEplug=9.81784$	$hcv=0.084$ $AMSPEcv=2.10769$ $hplug=0.3642$ $AMSPEplug=10.95634$	$hcv=0.15000$ $AMSPEcv=1.94542$ $hplug=0.30061$ $AMSPEplug=11.17195$	$hcv=0.09200$ $AMSPEcv=1.48378$ $hplug=0.25239$ $AMSPEplug=12.96201$
45	$hcv=0.14300$ $AMSPEcv=0.70175$ $hplug=0.36624$ $AMSPEplug=5.71249$	$hcv=0.12500$ $AMSPEcv=1.79561$ $hplug=0.31765$ $AMSPEplug=5.66489$	$hcv=0.13400$ $AMSPEcv=1.64223$ $hplug=0.27383$ $AMSPEplug=6.71927$	$hcv=0.08300$ $AMSPEcv=2.09823$ $hplug=0.23340$ $AMSPEplug=11.30690$
60	$hcv=0.13900$ $AMSPEcv=1.18628$ $hplug=0.33583$ $AMSPEplug=5.53363$	$hcv=0.02700$ $AMSPEcv=2.88659$ $hplug=0.30754$ $AMSPEplug=4.65369$	$hcv=0.11200$ $AMSPEcv=0.71902$ $hplug=0.25956$ $AMSPEplug=5.95729$	$hcv=0.11200$ $AMSPEcv=2.90357$ $hplug=0.22145$ $AMSPEplug=10.41049$
99		$hcv=0.07400$ $AMSPEcv=1.05487$ $hplug=0.27374$ $AMSPEplug=4.23467$	$hcv=0.10000$ $AMSPEcv=0.36013$ $hplug=0.23361$ $AMSPEplug=3.83591$	$hcv=0.07600$ $AMSPEcv=0.29136$ $hplug=0.19693$ $AMSPEplug=4.98237$

## APÉNDICE.

Se presentan a continuación los listados de tres programas realizados en el paquete Statistical Analysis System (SAS 6.06), y ejecutandos en una estación de trabajo HP900. En estos programas se han implementado los algoritmos propuestos en los capítulos III y IV.

### Listado 1. Implementación del algoritmo bootstrap propuesto en el epígrafe 3.7. (Capítulo III)

```

filename grafout pipe 'lp -d necs62p';
goptions device=xcolor;
goptions gaccess=sasgaedt;
goptions gprolog='25210d0a'x;
goptions gsfmode=replace;
goptions gsfname=grafout;
options linesize=70;

data datos;
se=1;
n=60;
r=99;
b=r;
s=datetime();
s=int(s);
Pi=3.14159265;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;
array f1i(99) f1i1-f1i99;
array f2i(99) f2i1-f2i99;
array sx(99) sx1-sx99;
array e(99) e1-e99;
array d(99) d1-d99;
array f(99) f1-f99;
array x(99) x1-x99;

/* Generamos aleatoriamente los parametros f1i f2i del modelo ma(2) */

do i=1 to r;
  y1=rannor(s);
  y2=rannor(s);
  f1i(i)=0.092581*y1-0.03781*y2+nu1;
  f2i(i)=-0.03781*y1+0.09258*y2+nu2;
  sx(i)=sqrt(se**2*(1+f1i(i)**2+f2i(i)**2));
  d(i)=se*rannor(s);
  f(i)=se*rannor(s);

```

```

end;

/* Generacion de r muestras de ma(2) */

do j=1 to n;
  do i=1 to r;
    e(i)=se*rannor(s);
    x(i)=e(i)+f1i(i)*d(i)+f2i(i)*f(i);
    f(i)=d(i);
    d(i)=e(i);
  end;
  output;
end;

/* Calculo de los periodogramas de los r individuos */

proc spectra p out=espect;
  var x1-x99;
run;

data c;
set espect;
array in{99} p_01-p_99;
array w{99} w1-w99;
array wbm{99} wbm1-wbm99;
array flb{99} flb1-flb99;
array qlb{99} qlb1-qlb99;
n=60;
r=99;
b=r;
Pi=3.14159265;
s=datetime();
s=int(s);
file 'pivote.dat';

/* Calculamos el periodograma medio para cada una de las frecuencias */

pm=0;
do i=1 to r;
  in(i)=in(i)/(4*Pi);
  pm=in(i)+pm;
end;
pm=pm/r;

/* Paso 1 estimamos W(i,j) */

do i=1 to r;
  w{i}=in{i}/pm;
end;

/* Paso 2, para cada frecuencia wj obtenemos b muestras bootstrap de los errores medios de los individuos */

do j=1 to N;
  if _n_=j then do;
    do k=1 to b;

```

```

    wbm{k}=0;
    do i=1 to r;
      l=ranuni(s);
      l=int(r*l)+1;
      wbm{k}=wbm{k}+w{l};
    end;
    wbm{k}=wbm{k}/r;
  end;
end;
end;

/* Paso 3, calculo del estimador bootstrap de la funcion de densidad espectral */
/* y del pivote bootstrap */

do k=1 to b;
  flb{k}=pm*wbm{k};
  qlb{k}=sqrt(r)*(flb{k}-pm);
  put( qlb{k})(10.5);
end;

/* Tenemos en un conjunto solo los pivotes bootstrap */

data pivbot;
array qb{99} qb1-qb99;
b=99;
infile 'pivote.dat';
do k=1 to b;
  input qb{k};
end;

proc transpose data=pivbot out=tranpiv1 prefix=v;
run;

data tranpiv;
set tranpiv1;
if _n_<=99 ;

/* Buscamos los percentiles 2.5 */

proc univariate data=tranpiv noprint;
var v1-v31;
output out=inferior pctlpts=2.5 pctlpre=v1-v31 pctlname=inf;
run;

proc transpose data=inferior out=qinf prefix=inf;
run;

/* Buscamos los percentiles 97.5 */

proc univariate data=tranpiv noprint;
var v1-v31;
output out=superior pctlpts=97.5 pctlpre=v1-v31 pctlname=sup;
run;

proc transpose data=superior out=qsup prefix=sup;
run;

```

```

/* Cálculo del intervalo de confianza */

data inter;
set qsup;
set qinf;
set c;
n=60;
r=99;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;
se=1;
file 'dr99n60.dat';
hw=(se**2/(2*Pi))*(1+v1+nu1**2+v2+nu2**2+(2*nu1+2*(cov12+nu1*nu2))*cos(freq)+2*nu2*c
os(2*freq));
pinfer=pm-(sup1/sqrt(r));
psuper=pm-(inf1/sqrt(r));
put( freq pinfer hw psuper) ( 10.5);
title 'MA(2) tma2 r=99 n=60';
symbol1 interpol=join value=dot height=0.1 color=blue;
symbol2 interpol=join value=dot height=0.1 color=red;
symbol3 interpol=join value=square height=0.1 color=green;
symbol4 interpol=join value=square height=0.1 color=yellow;
proc gplot data=inter;
plot (pinfer psuper hw)*freq /overlay;
run;

```

## Listado 2. Implementación del algoritmo bootstrap propuesto en el epígrafe 4.3. corregido el problema del sesgo siguiendo a Härdle y Bowman (Capítulo IV)

```

filename grafout pipe 'lp -d necs62p';
goptions device=xcolor;
goptions gaccess=sasgaedt;
goptions gprolog='25210d0a'x;
goptions gsfname=replace;
goptions gsfname=grafout;
options linesize=70;

data datos;
se=1;
n=200;
r=99;
s=datetime();
s=int(s);
Pi=3.14159265;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;
array fli(99) fli1-fli99;

```



```

array f2i(99) f2i1-f2i99;
array sx(99) sx1-sx99;
array e(99) e1-e99;
array d(99) d1-d99;
array f(99) f1-f99;
array x(99) x1-x99;

/* Generamos los parametros f1i f2i del modelo ma(2) */

do i=1 to r;
  y1=rannor(s);
  y2=rannor(s);
  f1i(i)=0.092581*y1-0.03781*y2+nu1;
  f2i(i)=-0.03781*y1+0.09258*y2+nu2;
  sx(i)=sqrt(se**2*(1+f1i(i)**2+f2i(i)**2));
  d(i)=se*rannor(s);
  f(i)=se*rannor(s);
end;

/* Generacion de r muestras de ma(2) */

do j=1 to n;
  do i=1 to r;
    e(i)=se*rannor(s);
    x(i)=e(i)+f1i(i)*d(i)+f2i(i)*f(i);
    f(i)=d(i);
    d(i)=e(i);
  end;
  output;
end;

/* Calculo de los periodogramas de los r individuos */

proc spectra p out=espect;
  var x1-x99;
run;

/* Conjunto de frecuencias*/

data freq;
set espect;
drop p_01-p_99 period;

/* Calculo de los Z(w) */

data z;
set datos;
set freq;
se=1;
n=200;
r=99;
Pi=3.14159265;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;

```

```

k1=se**2/(2*Pi);
k2=1+v1+nu1**2+v2+nu2**2;
k3=2*nu1+2*(cov12+nu1*nu2);
k4=2*nu2;
hw=k1*(k2+k3*cos(freq)+k4*cos(2*freq));
array f1i(99) f1i1-f1i99;
array f2i(99) f2i1-f2i99;
array sx(99) sx1-sx99;
array z(99) z1-z99;
varz=0;
do l=1 to r;
  z(l)=se/(2*Pi*sx(l))*(f1i(l)**2+f2i(l)**2+2*f2i(l)*cos(freq));
  z(l)=z(l)/hw;
  varz=(z(l)-1)**2+varz;
end;
varz=varz/(r-1);

data z;
set z;
drop f1i1-f1i99 f2i1-f2i99 sx1-sx99 se n Pi r
  nu1 nu2 v1 v2 cov12 k1 k2 k3 k4 hw freq i l
  y1 y2 s z1-z99 f1-f99 d1-d99 x1-x99 e1-e99 j;

data c;
set espec;
array in{99} p_01-p_99;
array w{99} w1-w99;
array wbm{99} wbm1-wbm99;
n=200;
r=99;
b=r;
Pi=3.14159265;
s=datetime();
s=int(s);

/* Calculamos el periodograma medio para cada una de las frecuencias */

pm=0;
do i=1 to r;
  in{i}=in{i}/(4*Pi);
  pm=pm+in{i};
end;
pm=pm/r;

/* Generamos los W(i,j) */

do i=1 to r;
  w{i}=in{i}/pm;
end;

/* estimacion de la varianza de los Wij */

vari=0;
do i=1 to r;
  vari=(w{i}-1)**2+vari;
end;
vari=vari/(r-1);

```

```

/* Para cada frecuencia wj obtenemos b muestras bootstrap de los errores medios de los individuos
*/

```

```

do j=1 to N;
  if _n_=j then do;
    do k=1 to b;
      wbm{k}=0;
      do i=1 to r;
        l=ranuni(s);
        l=int(r*l)+1;
        wbm{k}=wbm{k}+w{l};
      end;
      wbm{k}=wbm{k}/r;
    end;
  end;
end;

data d;
set c;
drop freq period p_01-p_99 w1-w99 wbm1-wbm99 n r b Pi s i j k l vari;
run;

data varian;
set c;
drop freq period p_01-p_99 w1-w99 wbm1-wbm99 n r b Pi s i j k l pm;
run;

proc transpose data=d out=e1 prefix=ww;
run;

data f1;
set e1;
if _n_<=1;

data _null_;
set f1;
array pr{101} ww1-ww101;
Pi=3.14159265;
r=99;
g=0.08;
h=g;
n=200;
nh=int(h*n/2);
m=int(n/2)+1;
file 'periali.dat';
do k=1 to m;
  gg=0;
  ggm=0;
  do j=1 to m;
    if abs(k-j)<=nh then do;
      gg=(1-(2*(k-j)/(n*h))**2)*pr(j)+gg;
      ggm=(1-(2*(k-j)/(n*h))**2);
    end;
  end;
do j=2 to m;
  if abs(k+j)<=nh then do;

```

```

        gg=(1-(2*(k+j)/(n*h))**2)*pr(j)+gg;
        ggm=(1-(2*(k+j)/(n*h))**2)
    end;
end;
ggm=ggm*3/(2*h*n);
gg=gg/ggm;
put (gg)(10.5);
end;

data g;
infile 'periali.dat';
array g(101) g1-g101;
do k=1 to 101;
    input( g(k))( 10.5);
end;

data _null_;
set g;
array pr{ 101} g1-g101;
Pi=3.14159265;
r=99;
h=0.1;
n=200;
nh=int(h*n/2);
m=int(n/2)+1;
file 'periali.dat';
do k=1 to m;
    gg=0;
    ggm=0;
    do j=1 to m;
        if abs(k-j)<=nh then do;
            gg=(1-(2*(k-j)/(n*h))**2)*pr(j)+gg;
            ggm=(1-(2*(k-j)/(n*h))**2);
        end;
    end;
    do j=2 to m;
        if abs(k+j)<=nh then do;
            gg=(1-(2*(k+j)/(n*h))**2)*pr(j)+gg;
            ggm=(1-(2*(k+j)/(n*h))**2);
        end;
    end;
    ggm=ggm*3/(2*h*n);
    gg=gg/ggm;
    put ( gg)(10.5);
end;

data gg;
infile 'periali.dat';
array g(101) g1-g101;
do k=1 to 101;
    input( g(k))( 10.5);
end;

proc transpose data=g out=g1 prefix=pali;
run;

data g2;

```

```

set g1;
if _n_ <= 101;

proc transpose data=gg out=gg1 prefix=palig;
run;

data gg2;
set gg1;
if _n_ <= 101;

data fwg;
set g2;
set gg2;
r=99;
n=200;
Pi=3.14159265;
g=0.08;
h=0.1;
fg=pali1*3/(2*g*n);
fc=palig1*(3/(2*g*n))*(3/(2*h*n));

data _null_;
set f1;
array pr{101} ww1-ww101;
file 'periali.dat';
Pi=3.14159265;
r=99;
h=0.1; /* Hay que variar el valor del bandwidth */
n=200;
nh=int(h*n/2);
m=int(n/2)+1;
do k=1 to m;
  gg=0;
  ggm=0;
  do j=1 to m;
    if abs(k-j) <= nh then do;
      gg=(1-(2*(k-j)/(n*h))**2)*pr(j)+gg;
      ggm=(1-(2*(k-j)/(n*h))**2);
    end;
  end;
  do j=2 to m;
    if abs(k+j) <= nh then do;
      gg=(1-(2*(k+j)/(n*h))**2)*pr(j)+gg;
      ggm=(1-(2*(k+j)/(n*h))**2);
    end;
  end;
  ggm=ggm*3/(2*n*h);
  gg=gg/ggm;
  put( gg)(10.5);
end;

data fffg;
infile 'periali.dat';
array g(101) g1-g101;
do k=1 to 101;
  input( g(k))( 10.5);
end;

```

```

proc transpose data=fffg out=fffg1 prefix=pali;
run;

data ffg2;
set ffg1;
if _n_ <= 101;

data fwh;
set ffg2;
r=99;
n=200;
Pi=3.14159265;
h=0.1;
fh=pali1*3/(2*h*n);

/* Obtenemos los periodogramas medios bootstrap de cada una de las b muestras bootstrap */

data _null_;
set c;
set fwg;
array pb{99} pb1-pb99;
array wbm{99} wbm1-wbm99;
file 'pboots.dat';
title 'h=0.1';
b=99;
do k=1 to b;
  pb{k}=fg*wbm{k};
  put( pb{k})( 10.5);
end;

/* Calculo del estimador bootstrap fg */

data db;
array pb(99) pb1-pb99;
infile 'pboots.dat';
do i=1 to 99;
input pb(i);
end;

proc transpose data=db out=e1b prefix=wwb;
run;

data f1b;
set e1b;
if _n_ <= 99;

/* Estimacion de la funcion de densidad poblacional bootstrap */

data _null_;
set f1b;
array pr{101} ww1-ww101;
file 'pbali.dat';
Pi=3.14159265;
r=99;
b=r;
h=0.1;

```

```

n=200;
nh=int(h*n/2);
m=int(n/2)+1;
do k=1 to m;
  gg=0;
  ggm=0;
  do j=1 to m;
    if abs(k-j)<=nh then do;
      gg=(1-(2*(k-j)/(n*h))**2)*pr(j)+gg;
      ggm=(1-(2*(k-j)/(n*h))**2);
    end;
  end;
  do j=2 to m;
    if abs(k+j)<=nh then do;
      gg=(1-(2*(k+j)/(n*h))**2)*pr(j)+gg;
      ggm=(1-(2*(k+j)/(n*h))**2);
    end;
  end;
  ggm=ggm*3/(2*n*h);
  gg=gg/ggm;
  do i=1 to b;
    put( gg)(10.5);
    output;
  end;
end;

data gb;
infile 'pbali.dat';
array g(101) g1-g101;
do k=1 to 101;
  do i=1 to 99;
    input( g(k))( 10.5);
  end;
end;

proc transpose data=gb out=g1b prefix=pbali;
run;

data g2b;
set g1b;
if _n_<=101;

/* Obtenemos los pivotes bootstrap */

data _null_;
set g2b;
set fwg;
set espect;
array fbali(99) pbali1-pbali99;
array qb{99} qb1-qb99;
r=99;
b=r;
n=200;
h=0.1;
Pi=3.14159265;
nu1=0.2;
nu2=-4;

```

```

v1=0.01;
v2=0.01;
cov12=-0.007;
se=1;
k1=se**2/(2*Pi);
k2=1+v1+nu1**2+v2+nu2**2;
k3=2*nu1+2*(cov12+nu1*nu2);
k4=2*nu2;
hw2=-k1*k3*cos(freq)-4*k1*k4*cos(2*freq);
bw=((h**2)/2)*hw2*((Pi**2)/5);
file 'pivote.dat';
do k=1 to b;
  fbali{k}=fbali{k}*3/(2*h*n); /* div r*/
  qb{k}=(fbali{k}-fc+bw)/fg;
  put( qb{k})( 10.5);
end;

/* Tenemos en un conjunto solo los pivotes bootstrap */

data pivbot;
array qb{99} qb1-qb99;
b=99;
infile 'pivote.dat';
do k=1 to b;
  input qb{k};
end;

/* Transponemos los dataos por frecuencias */

proc transpose data=pivbot out=tranpiv1 prefix=v;
run;

data tranpiv;
set tranpiv1;
if _n_<=99 ;

/* Buscamos los percentiles 2.5 */

proc univariate data=tranpiv noprint;
var v1-v101;
output out=inferior pctlpts=2.5 pctlpre=v1-v101 pctlname=inf;
run;

proc transpose data=inferior out=qinf prefix=inf;
run;

/* Buscamos los percentiles 97.5 */

proc univariate data=tranpiv noprint;
var v1-v101;
output out=superior pctlpts=97.5 pctlpre=v1-v101 pctlname=sup;
run;

proc transpose data=superior out=qsup prefix=sup;
run;

/* calculo del intervalo de confianza */

```



```

data inter;
set varian;
set qinf;
set qsup;
set fwh;
set freq;
set z;
set d;
n=200;
h=0.1;
r=99;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;
se=1;
Pi=3.14159265;
t=tinv(0.975,r-1);
k1=se**2/(2*Pi);
k2=1+v1+nu1**2+v2+nu2**2;
k3=2*nu1+2*(cov12+nu1*nu2);
k4=2*nu2;
hw=k1*(k2+k3*cos(freq)+k4*cos(2*freq));
hw2=-k1*k3*cos(freq)-4*k1*k4*cos(2*freq);
bw=((h**2)/2)*hw2*(Pi**2)/5;
fi=sqrt((vari+1)*6/(5*n*h*r)+varz/(4*(Pi**2)*r));
pinfer=fh/(1+sup1);
psuper=fh/(1+inf1);
nsuper=(fh-bw)/(1-1.96*fi);
ninfer=(fh-bw)/(1+1.96*fi);
title 'MA(2) tma2 y normal sesgo r=99 n=200';
symbol1 interpol=join value=dot height=0.1 color=blue;
symbol2 interpol=join value=dot height=0.1 color=blue;
symbol3 interpol=join value=square height=0.1 color=green;
symbol4 interpol=join value=square height=0.1 color=red;
symbol5 interpol=join value=square height=0.1 color=red;
proc gplot data=inter;
plot (pinfer psuper hw ninfer nsuper)*freq /overlay;
run;

```

### Listado 3. Implementación del algoritmo de validación cruzada propuesto en el epígrafe 4.4. (Capítulo IV).

```

filename grafout pipe 'lp -d necs62p';
goptions device=xcolor;
goptions gaccess=sasgaedt;
goptions gprolog='25210d0a'x;
goptions gsfmode=replace;
goptions gsfname=grafout;
options linesize=70;

data datos;
se=1;

```

```

n=100;
r=90;
s=datetime();
s=int(s);
Pi=3.14159265;
nu1=0.2;
nu2=-4;
v1=0.01;
v2=0.01;
cov12=-0.007;
array f1i(90) f1i1-f1i90;
array f2i(90) f2i1-f2i90;
array sx(90) sx1-sx90;
array e(90) e1-e90;
array d(90) d1-d90;
array f(90) f1-f90;
array x(90) x1-x90;

/* Generamos los parametros f1i f2i del modelo ma(2) */

do i=1 to r;
  y1=rannor(s);
  y2=rannor(s);
  f1i(i)=0.092581*y1-0.03781*y2+nu1;
  f2i(i)=-0.03781*y1+0.09258*y2+nu2;
  sx(i)=sqrt(se**2*(1+f1i(i)**2+f2i(i)**2));
  d(i)=se*rannor(s);
  f(i)=se*rannor(s);
end;

/* Generacion de r muestras de ma(2) */

do j=1 to n;
  do i=1 to r;
    e(i)=se*rannor(s);
    x(i)=e(i)+f1i(i)*d(i)+f2i(i)*f(i);
    f(i)=d(i);
    d(i)=e(i);
  end;
  output;
end;

/* Calculo de los periodogramas de los r individuos */

proc spectra p out=espect;
  var x1-x90;
run;

data pm;
set espect;
array in{90} p_01-p_90;
n=100;
r=90;
Pi=3.14159265;
r1=45;
pm1=0;
do i=1 to r1;

```

```

    in{i}=in{i}/(4*Pi);
    pm1=pm1+in{i};
end;
pm1=pm1/r1;
pm2=0;
do i=r1+1 to r;
    in{i}=in{i}/(4*Pi);
    pm2=pm2+in{i};
end;
pm2=pm2/(r-r1);

data pm2;
set pm;
drop freq period p_01-p_90 n r pi r1 pm1 i;

data freq;
set pm;
drop period p_01-p_90 n r pi r1 pm1 pm2 i;

data pm;
set pm;
drop freq period p_01-p_90 n r pi r1 pm2 i;

proc transpose data=pm out=e1 prefix=ww;
run;

data f1;
set e1;
if _n_<=1;

data f1;
set f1;
drop _name_;

proc transpose data=pm2 out=e1 prefix=www;
run;

data f2;
set e1;
if _n_<=1;

data f2;
set f2;
drop _name_;

data vc;
set f1;
set f2;
array pr{51} ww1-ww51;
array p{51} www1-www51;
array fh{51} fh1-fh51;
n=100;
r1=45;
r=r1;
Pi=3.14159265;
m=int(n/2)+1;
h0=0;

```

```

min=0;
i=0;
do h=0 to 1 by 0.01;
nh=int(h*n/2);
do k=1 to m;
  if h=0 then fh{k}=pr{k};
  else do
    gg=0;
    ggm=0;
    do j=1 to m;
      if abs(k-j)<=nh then do;
        gg=(1-(2*(k-j)/(n*h))**2)*pr(j)+gg;
        ggm=(1-(2*(k-j)/(n*h))**2)+ggm;
      end;
    end;
    do j=2 to m;
      if abs(k+j)<=nh then do;
        gg=(1-(2*(k+j)/(n*h))**2)*pr(j)+gg;
        ggm=(1-(2*(k+j)/(n*h))**2)+ggm;
      end;
    end;
    fh{k}=gg*(3/(2*n*h));
  end;
end;

s=0;
do j=1 to m;
  s=s+(p{j}-fh{j})**2;
end;
t=s;
g=h;
output;
if h=0 then min=s;
if min>s then do;
  h0=h;
  min=s;
end;
end;

symbol1 interpol=join value=dot height=0.1 color=blue;

proc gplot data=vc;
plot t*g;
run;

data vc2;
set vc;
file 'h2.dat';
put( t g)(10.5);

proc print data=vc;
var h0;
run;

```

**BIBLIOGRAFÍA.**

- Aebi, M. and Embrechts, P. and Mikosch, T. (1994). Stochastic Discounting, Aggregate Claims, and the Bootstrap. *Adv. Appl. Prob.* **26** 183-206.
- Altman, N. S. (1990). Kernel Smoothing of Data With Correlated Errors. *Journal of American Statistical Association.* **85**, 749-759.
- Basawa, I., Mallik, A., McCormick, W., Reeves, J. and Taylor, R. (1991). Bootstrapping Unstable First-Order Autorregressive Processes. *The Annals of Statistics*, **19**, 1098-1101.
- Beltrao, K and Bloomfield, P. (1987). Determining the Bandwidth of a Kernel Spectrum Estimate. *Journal of Time Series Analysis.* **8**, 1. 21-38.
- Bendat, J. and Piersol, A. (1986). *Random Data: Analysis and Measurement Procedures*. Wiley.
- Beran, R. (1984). Jackknife Approximations to Bootstrap Estimates. *The Annals of Statistics*. **12**, 101-118.
- Bickel, P. and Freedman, D. (1981). Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics.* **9**, 1196-1217.
- Bickel, P, Vijayan, N and Wang P. (1992). Nonparametric Inference under Biase Sampling from a Finite Population. *The Annals of Statistics.* **20**, 853-878.
- Billingsley, P. (1968). *Convergence of Probability Measure*. Wiley.
- Blackman, R. B. and Tukey, J. W. (1958). The Measurement of Power Spectrum from the point of view of Communications Engineering. *Bell Systems Tech. Journal.* **37**, 185-282 and 485-569.
- Brockmann, M., Gasser, T. and Herrmann, E. (1993). Locally Adaptive Bandwidth Choise for Kernel Regression Estimators. *Journal of American Statistical Association.* **88**, 1302-1309.
- Burrill, C. W. (1972). *Measure, integration and Probability*. Mc-GrawHill.
- Cao, R., Cuevas, A. and Gonzalez, W. (1994). A comparative study of several smoothing methods in density estimation. *Computational Statistical and Data Analysis*. **17**, 153-170.
- Chen, C. H. (1989). *Applied Time Series Analysis*. World Scientific.

- Chen, Z. and Hannan, E. (1980). The Distribution of periodogram ordinates. *Journal of Time Series Analysis*. **1**, 73-83.
- Chiu, S.T. (1990). On the Asymptotic Distributions of Bandwidth Estimates. *The Annals of Statistics*. **18**, 1696-1711.
- Chiu, S.T. (1991). Bandwidth Selection for Kernel Density Estimation. *The Annals of Statistics*. **19**, 1883-1905.
- Chu, C.K. and Marron, J. S. (1991). Comparison of two bandwidth selector with dependent errors. *The Annals of Statistics*. **19**, 1906-1918.
- Diggle, P. J. (1990). *Time Series: A Biostatistical Introduction*. Oxford Science Publications.
- Diggle, P.J. and Al-Wasel, I. (1993). On Periodogram-Based Spectral Estimation for Replicated Time Series. *Developments in Time Series Analysis*. Chapman and Hall.
- Do, K. and, Hall, P. (1991). On importance resampling for the bootstrap. *Biometrika*. **78**, 161-167.
- Dunn, O. J. and Clark, V.A. (1987). *Applied Statistics: Analysis of Variance and Regression*. Wiley.
- Efron, B. (1979). Bootstrap methods: Another Look at the Jackknife. *The Annals of Statistics*. **7**, 1-26
- Efron, B. (1994). Missing Data, Imputation, and the Bootstrap. *Journal of the American Statistical Association*. **89**, 463-479.
- Efron, B. and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*. **1**, 54-77.
- Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.
- Fan, J. and Gijbels, I. (1992). Variable Bandwidth and Local Linear Regression Smoothers. *The Annals of Statistics*. **20**, 2008-2036.
- Fan, J. and Marron J. S. (1992). Best Possible Constant for Bandwidth Selection. *The The Annals of Statistics*. **20**, 2057-2070.
- Faraway, J. and Jhun, M. (1990). Bootstrap Choice of Bandwidth for Density Estimation. *Journal of American Statistical Association*. **85**, 1119-1122

- Ferguson, T. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press.
- Franke, J and Härdle, W. (1992). On-Boostrapping Kernel Spectral Estimates. *The Annals of Statistics*. **20**, 121-145.
- Freedman, D. (1981). Bootstrapping Regression Models. *The Annals of Statistics*. **19**, 1218-1228.
- Freedman, D. (1984). On Bootstrapping two-stage least-squares estimates in stationary linear models. *The Annals of Statistics*. **12**, 827-842.
- Gasser, T., Müller, H., Köhler, W., Molinari, L. and Prader, A. (1984) Nonparametric regression analysis of growth curves. *The Annals of Statistics*. **12**, 210-229.
- Gasser, T. and Kneip, A. (1995). Searching for Structure in curves Samples. *Journal of American Statistical Association*. **90** ,1179-1188.
- Grenander, U. and Rosenblatt, M. (1953). Statistical spectral analysis arising from stationary stochastic processes. *Ann. Math. Statist*. **24**, 537-558.
- Gu, C. and Qiu, C. (1993). Smoothing Spline Density Estimation:Theory. *The Annals of Statistics*. **21**, 217-234..
- Hall, P. (1992). On Global Properties of Variables Bandwidth Density Estimators. *The Annals of Statistics*. **20**, 762-778.
- Hall, P (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag.
- Härdle, W. (1991). *Smoothing Techniques*. Springer-Verlag.
- Härdle, W. and Marron, J. (1985). Optimal Bandwidth Selection in Nonparametric Regression Function Estimation. *The Annals of Statistics*. **13**, 1465-1481.
- Härdle, W. and Bowman, A. (1988). Bootstrapping in Nonparametric Regression: Local Adaptive Smoothing and Confidence Bands. *Journal of the American Statistical Association*. **83**, 102-110.
- Härdle, W. and Vieu, P. (1992). Kernel Regression Smoothing of Time Series. *Journal of Time Series Analysis*. **13**, 209-232.
- Hartigan, J. (1990). Perturbed Periodogram Estimates of Variance. *International Statistical Revice*. **58**, 1-7.
- Hästie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. Chapman and Hall.

- Kalman, R. (1960). A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **82**, 35-45.
- Kreiss, J. and Franke, J. (1992). Bootstrapping Stationary Autoregressive Moving-Average Models. *Journal of Time Series Analysis.* **13**, 297-317.
- Künsch, H. (1989). The Jackknife and the Bootstrap for general Stationary Observations. *The Annals of Statistics.* **17**, 1217-1241.
- Lo, A. (1986). A Bayesian Bootstrap for Finite Population. *The Annals of Statistics.* **16**, 1684-1695.
- Major, P. (1978). On the variance principle for sums of independent, identically distributed random variables. *J. Multivariate Anal.* **8**, 487-501.
- Marron, J. S. (1985). An asymptotically efficient solution to the bandwidth problem of kernel density estimation. *The Annals of Statistics* **13**, 1011-1023.
- Messer, K. and Goldstein, L. (1993). A new class of kernels for nonparametric curve estimation. *The Annals of Statistics* **21**, 179-195.
- Milton, S. (1994). *Estadística para Biología y Ciencias de la Salud*. McGraw-Hill.
- Nadaraya, E. A. (1989). *Nonparametric Estimation of Probability Densities and Regression Curves*. Kluwer Academic Publishers.
- Paparoditis, E. and Streitberg, B. (1991). Order Identification Statistics in Stationary Autoregressive Moving-Average Models: Vector Autocorrelations and Bootstrap. *Journal of Time Series Analysis.* **13**, 415-434.
- Park, B. and Marron, J. S. (1990). Comparison of Data-Driven Bandwidth Selectors. *Journal of American Statistical Association.* **85**, 66-72.
- Parzen, E. (1957). On choosing an estimate of the spectral density function of stationary time series. *Ann. Math. Statist.* **28**, 921-932.
- Priestley, M.B. (1981). *Spectral Analysis and Time Series*. Wiley
- Politis, D. Romano, J. and Lai T. (1992). Bootstrap Confidence Bands for Spectra and Cross-Spectra. *IEEE Transactions on signal processing.* **40**, 1206-1215.
- Pötzelberger, K. (1990). A characterization of Random-coefficient  $AR(1)$ . Models. *Stochastic Processes and their Applications.* **34**, 171-180.



- Ramlau-Hansen, H. (1983). Smoothing Counting Process Intensities by Means of Kernel Functions. *The Annals of Statistics*. **11**, 453-466.
- Rice, J. (1984). Bandwidth choice for nonparametric Regression. *The Annals of Statistics*. **12**, 1215-1230.
- Romano, J. (1988). Bootstrapping the Mode. *Inst. Statist. Math.* **40**, 565-586.
- Romano, J. (1989). Bootstrap and randomization test of some nonparametric hypotheses. *The Annals of Statistics* **17**, 141-159.
- Rubin, D. (1981). The Bayesian Bootstrap. *The Annals of Statistics*. **9**, 130-134.
- Ruppert, D. and Wand, M.P. (1994). Multivariate Locally weighted least squares regression. *The Annals of Statistics*. **22**, 1346-1370
- Simonoff, J. (1996). *Smoothing Methods in Statistics*. Springer.
- Shibata, R. (1981). An optimal autorregressive spectral estimate. *The Annals of Statistics*. **9**, 300-306.
- Shanmugan, K. and Breipohl, A. (1988). *Random Signals: Detection, Estimation and Data Analysis*. Wiley.
- Shannon, C. (1948). A Mathematical Theory of Communication. *Bell Systems Tech. J.* **27**, 379-423 and 623-656.
- Shao, J. and Wu, C. F. J. (1989). A general theory for Jackknife variance estimation. *The Annals of Statistics*. **17**, 1176-1197.
- Shao, J and Tu, D. (1995). *The Jackknife and Bootstrap*. Springer-Verlag.
- Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics*. Wiley.
- Staniswalis, J. (1989). Local Bandwidth Selection for Kernel Estimates. *Journal of the American Statistical Association*. **84**, 284-288.
- Staudte, R. and Sheather, S. (1990). *Robust Estimation and Testing*. Wiley.
- Sun, J. and Loader, C. (1994). Simultaneous Confidence Bands for Linear Regression and Smoothing. *The Annals of Statistics*. **22**, 1328-1345.
- Swanepoel, J. and Wyk, J. (1986). The bootstrap applied to power spectral density function estimation. *Biometrika*. **73**, 135-141.

- Terrell, G. and Scott, D. (1992). Variable Kernel Density Estimation. *The Annals of Statistics*. **20**, 1236-1265.
- Thombs, L. and Schucany, W. (1990). Bootstrap Prediction Intervals for Autoregression. *Journal of the American Statistical Association*. **85**, 286-492.
- Tong, H. (1993). *Non-linear Time Series*. Clarendon Press Oxford.
- Wand, M.P. and Jones M.C. (1995). *Kernel Smoothing*. Chapman-Hall.
- Wang, S. (1995). Optimizing the Smoothed bootstrap. *Ann. Inst. Statist. Math.* **47**, 65-80.
- Wiener, N. (1948). *Cybernetics*. MIT Press.
- Wu,C. (1986). Jackknife, Bootstrap and other Resampling Methods in Regression Analysis. *The Annals of Statistics*. **14**, 1261-1295.
- Wu,C. (1990). On the asymptotic properties of the jackknife histogram. *The Annals of Statistics*. **18**, 1438-1452.
- Zhao-Guo, C. and Hannan, E. (1980). The Distribution of Periodogram Ordinates. *Journal of Time Series Analysis*. **1**, 73-83.