UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DEPARTAMENTO DE MATEMÁTICAS

TESIS DOCTORAL

ESTIMACIÓN POLINÓMICO LOCAL DEL ESPECTRO POBLACIONAL

JUAN ARTILES ROMERO

Las Palmas de Gran Canaria, diciembre de 1997

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

17/1997-98

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 CUM LAUDE POR UNANIMIDAD

Las Palmas de Gran Canaria a 5 de diciembre de 1997.

El/a Presidente/a: Drª. Dª. Pilar Ibarrola Muñoz,

El/a Secretario/a: Drª. Dª. Carmen Nieves Hernández Flores,

El/a Vocal: Dr. D. Vicente Quesada Paloma,

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

El/a Vocal: Dr. D. José María Limiñana Cañal,

El Doctorando: D. Juan Artiles Romero,

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DEPARTAMENTO DE MATEMÁTICAS

TESIS DOCTORAL

ESTIMACIÓN POLINÓMICO LOCAL DEL ESPECTRO POBLACIONAL

Juan Artiles Romero

1997

BIBLIOTECA UNIVERSITARIA
LAS PALMAS DE G. CANARIA
N.º Documento 537.380
N.º Copia 537.382

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

ESTIMACIÓN POLINÓMICO LOCAL DEL ESPECTRO POBLACIONAL

Juan Artiles Romero

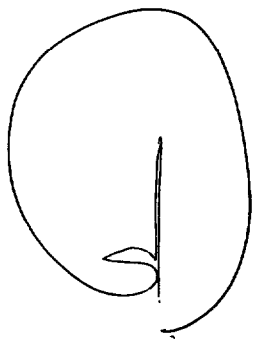
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA, 1997

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 ***Estimación Polinómica Local del Espectro Poblacional*** ha sido realizada bajo su dirección por el Licenciado en Ciencias Matemáticas Don Juan Artiles Romero, y constituye su Tesis para optar al grado de Doctor en Matemáticas.

Y para que así conste, en cumplimiento de la legislación vigente y a efectos a que haya lugar, firma la presente en Las Palmas de Gran Canaria, a catorce de octubre de mil novecientos noventa y siete.



© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Quiero expresar mi gratitud al profesor Dr. D. Pedro Saavedra Santana por sus valiosas orientaciones y eficaz ayuda.

Mi agradecimiento también a los compañeros del Departamento de Matemáticas de la Universidad de Las Palmas de Gran Canaria que de una manera u otra me prestaron ayuda y estímulo.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

ÍNDICE

PRÓLOGO/iv

CAPÍTULO I. ESTIMACIÓN ESPECTRAL.

1.1 Introducción./2

1.2 Procesos estocásticos estacionarios./3

1.3 Relación entre la densidad espectral y la función de autocovarianza./5

1.4 Representación espectral de procesos estacionarios./7

1.5 Procesos lineales generales./10

1.6 Estimación espectral: análisis del periodograma./11

1.7 Estimación de la densidad espectral mediante alisamientos del periodograma./15

1.8 El problema de la elección del ancho de banda./17

1.9 El bootstrap./22

1.10 Estimación de la función de densidad espectral utilizando la metodología bootstrap./24

1.10.1 Método bootstrap para la estimación de la densidad espectral./24

1.10.2 Validez asintótica del método./26

CAPÍTULO II. ANÁLISIS ESPECTRAL CON SERIES REPLICADAS.

2.1 Introducción./31

2.2 El modelo de efectos fijos./32

2.3 El modelo de efectos aleatorios./34

2.4 Procesos lineales de coeficientes aleatorios./35

2.5 El modelo de ruido blanco./36

2.6 Modelo general de efectos aleatorios./37

2.7 Estimación bootstrap del espectro poblacional./40

2.8 Alisamientos mediante núcleo del periodograma medio./42

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

2.9 Aproximación bootstrap para la estimación del espectro poblacional, utilizando como estimador alisamientos del periodograma medio./44

2.10 Ancho de banda óptimo./46

2.12 Estimación del ancho de banda óptimo./48

CAPÍTULO III. ESTIMACIÓN POLINÓMICA LOCAL DEL ESPECTRO POBLACIONAL.

3.1 Introducción./51

3.2 Estimación polinómica local del espectro poblacional./52

3.3 Sesgo y varianza del estimador./56

3.4 Comportamiento cerca de la frontera./66

3.5 Núcleos equivalentes./67

3.6 Distribución asintótica del estimador./68

3.7 Ancho de banda óptimo./70

3.8 Selección del ancho de banda./72

CAPÍTULO IV. ESTIMACIÓN BOOTSTRAP.

4.1 Introducción./75

4.2 Algoritmo bootstrap./75

4.3 Desarrollos de Edgeworth./77

4.4 El bootstrap y los desarrollos de Edgeworth./80

4.5 Aplicación de los desarrollos de Edgeworth a la estimación del espectro poblacional./84

4.6 El bootstrap y los desarrollos de Edgeworth en la estimación del espectro poblacional./85

4.7 Corrección del sesgo./87

4.8 Corrección del sesgo para el pivote del espectro poblacional./88

4.9. Simulaciones./90

CONCLUSIONES/98

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

APÉNDICE./101

BIBLIOGRAFÍA./116

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

PRÓLOGO.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

PRÓLOGO.

Al final de la última centuria, A. Schuster estaba interesado en buscar periodicidades en datos, pero para poder estimar el espectro sobre la totalidad del recorrido de las frecuencias tenemos que llegar al análisis espectral tal como lo conocemos hoy. La función de densidad espectral es la herramienta natural para considerar las propiedades frecuenciales de una serie temporal. El análisis en el dominio de la frecuencia es muy apropiado en campos como la ingeniería eléctrica, geofísica, meteorología, ciencias del mar, etc.

El análisis espectral es una metodología fundamental en el estudio de la señal. Después de la invención de la radio, a comienzos del siglo XX, los ingenieros electrónicos reconocieron la importancia de modelizar la señal a través de procesos estocásticos estacionarios. Shannon (1948) utilizó los modelos de procesos aleatorios para formular una teoría que ha sido la base de la comunicación digital. Wiener (1948) formuló la teoría de las señales aleatorias y la aplicó a la resolución de problemas de filtrajes. La invención del radar durante las 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 (1960), conocido como filtro de Kalman. Dicho algoritmo se utiliza en sistemas de navegación para la exploración de las profundidades del mar.

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.

Sin embargo, la replicación de experimentos en las ciencias biomédicas, es más la regla que la excepción. El registro de un electroencefalograma o la oscilación de concentración de una hormona pueden modelizarse a través de procesos estacionarios. La observación de las diversas trayectorias de un proceso abre nuevas perspectivas para las investigaciones en estas ciencias, en el campo de la estimación espectral.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Diggle y Al-Wasel (1993) estudiaron la concentración de la 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. 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 basado en la teoría asintótica del periodograma, en el cual cada individuo posee una densidad espectral específica $f_l(\omega)= f(\omega)Z_l(\omega)$. 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 ($f_l(\omega)= f(\omega)Z_l(\omega)$). 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.

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 la realización de un proceso estocástico. Diggle y Al-Wasel llaman a la media de este proceso *espectro poblacional.*

Hernández (1997), estudia dos estimadores de la función de densidad espectral poblacional: el periodograma medio cuando el número de individuos r es bastante grande comparado con el número de observaciones N, y a través de alisamientos mediante núcleos del periodograma medio para el caso en que N es bastante mayor que

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

r. Parte del modelo multiplicativo de efectos aleatorios, pero sin suponer que el proceso $\{Z(\omega)\}$ sea un ruido blanco, y sin considerar a priori que sigue una distribución determinada.

La presente monografía está estructurada en cuatro capítulos. En el capítulo primero, sintetizamos algunos resultados sobre la estimación de la densidad espectral, cuando se observa una única serie de un proceso lineal general, el cual, bajo determinadas condiciones, asegura la existencia de distribución espectral absolutamente continua. Entre estos resultados podemos destacar las propiedades asintóticas del periodograma, que pese a ser un estimador inconsistente de la función de densidad espectral, sirve de base para conseguir estimadores consistentes, mediante su alisamiento por medio de ciertas funciones de núcleo. Asimismo, se indica la importancia de la elección del parámetro de alisamiento o ancho de banda y el problema de disponer de un buen selector de ancho de banda. También hacemos una breve introdución del método bootstrap, que permite estimar la distribución de cierta variable aleatoria, en base a los datos observados de ella.

En el capítulo segundo, realizamos una revisión del problema de series replicadas, planteado por Diggle y Al-Wasel (1993), y cómo lo resuelve Hernández (1997) bajo condiciones más plausibles, proponiendo dos tipos de estimadores de la función de densidad espectral poblacional.

El tercero y cuarto capítulos constituyen nuestra aportación al problema de estimación del espectro poblacional. En el tercer capítulo proponemos el estimador polinómico local como estimador de la función de densidad espectral poblacional, cuando exista. Éste consiste en ajustar para cada punto ω_0, un polinomio de grado *p*, a los puntos (ω_j, Y_j), (siendo $Y_j = \bar{I}(\omega_j)$), usando mínimos cuadrados ponderados con el núcleo $K_h(\omega_j - \omega_0)$. Estudiaremos el sesgo y varianza del estimador para el caso general de ajuste polinómico de grado *p*, y obtendremos además, las expresiones del sesgo y de la varianza, tanto cuando se trata de ajuste lineal local, como cuadrático

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

local. Asimismo se estudiará la distribución asintótica del estimador. Por último, se propone un método de validación cruzada, para estimar el ancho de banda óptimo.

Comenzaremos el capítulo cuarto proponiendo un algoritmo bootstrap para obtener una aproximación bootstrap del pivote $\sqrt{r}\dfrac{\hat{f}(\omega;p,h)-E\left[\hat{f}(\omega;p,h)\right]}{\sqrt{Var\left[Z(\omega)\right]}}$. Introduciremos los desarrollos de Edgeworth, y a partir de ellos estudiaremos el orden de aproximación del pivote bootstrap al real.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO I.

ESTIMACIÓN ESPECTRAL.

CAPITULO I. ESTIMACIÓN ESPECTRAL.

1.1 Introducción.

El análisis espectral es una metodología desarrollada originalmente en el campo de la física y particularmente en el análisis de la señal. Su objetivo esencial es el estudio del espectro de un proceso estacionario. Bajo determinadas condiciones, un proceso estacionario puede expresarse como una superposición de movimientos armónicos de diferentes frecuencias y amplitudes. El espectro del proceso es la distribución que a cada rango de frecuencias le asigna su contribución a la varianza del proceso.

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.

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 muy eficaz, liberando al matemático de problemas complejos. Además, el hecho de que los modernos ordenadores cada vez aporten un menor coste computacional, hace que el bootstrap se consolide como una importante herramienta estadística.

Iniciamos este capítulo con los conceptos de proceso estocástico estacionario, función de densidad espectral y representación espectral de un proceso estacionario. Introducimos el concepto de proceso lineal general para a continuación, enunciar la condición necesaria y suficiente para que un proceso estacionario tenga distribución espectral absolutamente continua. Tal condición consiste en que el proceso estacionario sea lineal, y sus coeficientes decrezcan a cero de una forma determinada.

Abordamos también el problema de estimación de la función de densidad espectral sobre la base de la observación de una única trayectoria del proceso o serie

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

temporal. Dado que el periodograma es un estimador inconsistente, hacemos una revisión de los estimadores mediante alisamientos del periodograma con ciertas funciones denominadas de núcleo, las cuales dependen del parámetro ancho de banda, cuya elección adecuada es de importancia crucial. Se indica también cuando un selector de ancho de banda es asintóticamente óptimo.

Por último, introducimos la idea del bootstrap, el concepto de métrica de Mallows, el método bootstrap para estimar la función de densidad espectral y la validez asintótica de este método bajo condiciones precisas que recoge Franke y Härdle (1992) en el teorema 1.16.

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.

Un proceso estocástico $\{X(t)\}_{t\in R}$ se dice estrictamente estacionario si $\forall n\in N$, $\forall t_1, \ldots,t_n \in R$ y $\forall h>0$ las variables $\left(X_{t_1},\ldots,X_{t_n}\right)$ y $\left(X_{t_1+h},\ldots,X_{t_n+h}\right)$ 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 $R(s,t)=\mathrm{Cov}(X_s,X_t)$, si existe, será sólo función de t-s. Estas propiedades

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

autocovarianza $R(s,t)=\mathrm{Cov}(X_s,X_t)$, 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. $E\left[|X(t)|^2\right]<\infty,\ \forall t\in T$
2. $E\left[X(t)\right]=\mu,\forall t\in T$
3. $Cov\left[X(t),X(t+\tau)\right]=R(\tau)$ (independiente de t)

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

1. $R(0)=\sigma^2$.
2. $|R(\tau)|\le R(0)$ para cualquier τ.
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.

La historia temporal $X(t)$ de una trayectoria de un proceso estacionario no hay razón para suponer que es periódica (salvo el caso del proceso armónico cuando los ω_j son conmensurables), y por lo tanto no puede ser representada por una serie de Fourier discreta. Además, $X(t)$ continúa indefinidamente en el tiempo y la condición $\int_{-\infty}^{\infty}|X(t)|dt<\infty$ no se satisface. Por consiguiente, no es posible hallar la transformada de Fourier de $X(t)$ directamente, para obtener información sobre la composición frecuencial del proceso. Esta dificultad puede superarse, analizando por los métodos de Fourier, no la propia trayectoria, sino su función de autocovarianza.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

1.3 Relación entre la densidad espectral y la función de autocovarianza.

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

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

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

■

Corolario1.2 Para que la función de autocovarianza $R(\tau)$ posea transformada de Fourier y exista $f(\omega)$ para cualquier ω, es suficiente que:

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

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

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

■

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

Teorema 1.3 (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 una función de distribución de una medida finita en $(-\infty,\infty)$, tal que, para todo τ, $R(\tau)$ puede expresarse por:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

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 τ. Luego las integrales en las que intervienen se convierten en sumas. Así, la expresión (1.3.7) se transforma en:

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

- 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 2π, 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 1.4 (Wold) Una condición necesaria y suficiente para que la sucesión $\{R(\tau);\ \tau=0,\ \pm1,\ \pm2,\ ...\}$ sea la función de autocovarianza de algún proceso estacionario discreto $\{X_t\}_{t\in Z}$ es que exista una función $F(\omega)$, con las propiedades de función de distribución de una medida finita en $(-\pi,\pi)$, tal que:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

■

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

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

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

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

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

Teorema 1.5 Un proceso de incrementos ortogonales $\{Z(\omega)\}_{\omega\in R}$ es continuo en media cuadrática si y sólo si $F(\omega)$ es continua en ω.

■

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 $\sum_{i=0}^{n-1} g(\omega_i)\big(Z(\omega_{i+1})-Z(\omega_i)\big)$ sobre una partición $\omega_a=\omega_0<\omega_1<...<\omega_n=\omega_b$, cuando el radio de la partición ρ tiende a cero, es decir:

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

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 1.6 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

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

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:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

Esta ecuación es llamada la representación espectral de *X(t)*.

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:

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

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 1.7 (Representación espectral) 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

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

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

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

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:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Teorema 1.8 Sea $\{X_t\}_{t\in Z}$ 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

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

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

■

1.5 Procesos lineales generales.

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

$$X_t = \sum_{u=-\infty}^{\infty} g_u \varepsilon_{t-u} \qquad (1.5.1)$$

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

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

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

Teorema 1.9 Un proceso $\{X_t\}_{t\in Z}$ tiene distribución espectral absolutamente continua si y sólo si $\{X_t\}_{t\in Z}$ es un proceso lineal general dado por (1.5.1).

■

En los siguientes epígrafes del presente capítulo se aborda el problema de la estimación de la función de densidad espectral a partir de la observación de una serie temporal.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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:

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

siendo $\{X_t\}_{t=1}^{N}$ la serie temporal observada, tal que $E[X_t]=0$. Revisamos a continuación las propiedades del periodograma como estimador de la densidad espectral.

Proposición 1.10 Si $f(\omega)$ satisface la condición de Lipschitz de orden uno, el periodograma $I_X(\omega)$ es un estimador asintóticamente centrado de $f(\omega)$.

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

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

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

se tiene que $\forall\theta\neq 0, F_N(\theta)\xrightarrow[N\to\infty]{} 0$ 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:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

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

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

simplificando (1.6.4) se obtiene

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

Expresando el periodograma en función de la transformada finita de Fourier como sigue:

$$I_X(\omega)=\left|\zeta_X(\omega)\right|^2 \tag{1.6.6}$$

y sustituyendo el valor de esta transformada, se obtiene

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

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 1.11 Sea $\{X_t\}$ un proceso lineal general definido de la forma dada por (1.5.1) siendo los $\{\varepsilon_t\}$ variables aleatorias incorreladas, con $E[\varepsilon_t]=0$, $E\left[\varepsilon_t^2\right]=\sigma_\varepsilon^2$ y $\sum_{u=-\infty}^{\infty}|g_u||u|^{1/2}<\infty$. Entonces:

$$\zeta_X(\omega)=\Gamma(\omega)\zeta_\varepsilon(\omega)+r_N(\omega) \tag{1.6.7}$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

■

Teorema 1.12 Sea $\{X_t\}$ un proceso lineal general definido por (1.5.1), verificando que los $\{\varepsilon_t\}$ son independientes con $E[\varepsilon_t]=0$, $E\left[\varepsilon_t^2\right]=\sigma_\varepsilon^2$, $E\left[\varepsilon_t^4\right]<\infty$, y $\sum_{u=-\infty}^{\infty}|g_u||u|^\alpha<\infty$, $\alpha>0$. Entonces

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

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

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

uniformemente en ω.

■

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

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

de donde

$$E\left[I_X(\omega_j)\right]\approx f(\omega_j) \qquad Var\left[I_X(\omega_j)\right]\approx f^2(\omega_j) \qquad (1.6.10)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 ω_1 y ω_2. Dado que no hay problema de confusión de la notación, consideraremos $I_X = I$.

Teorema 1.13 Sea $\{X_t\}$ un proceso lineal general definido por (1.5.1) con los $\{\varepsilon_t\}$ y $\{g_u\}$ verificando las condiciones del teorema 1.12. Entonces:

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

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 ω_1 y ω_2.

■

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

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

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

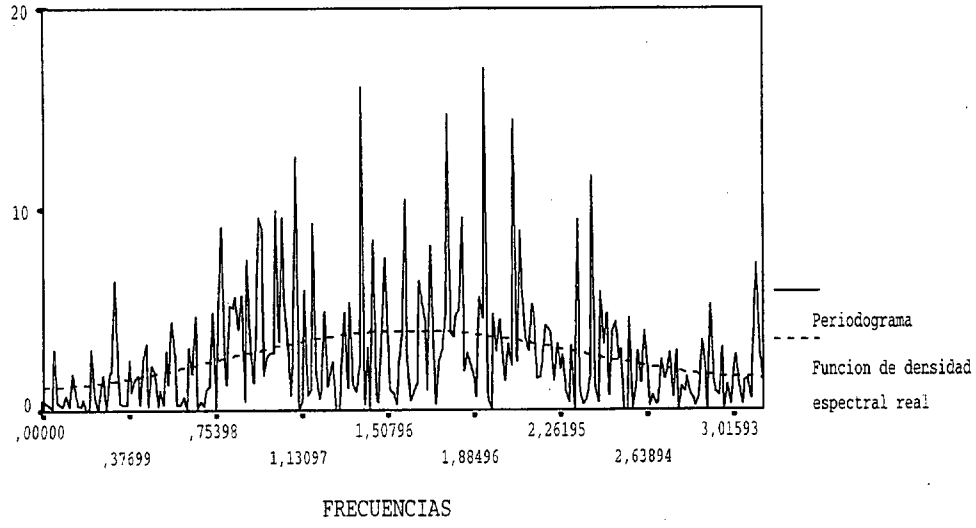
$$Var\{I(\omega)\} \to f^2(\omega) \tag{1.6.13}$$

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 no consistente dado que $Var[I(\omega)] \approx f^2(\omega)$. Además, para las frecuencias ω_1 y ω_2 se tiene $Cov\{I(\omega_1), I(\omega_2)\} \xrightarrow[N\to\infty]{} 0$, 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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

deduce que el periodograma es un mal estimador de la función de densidad. Mencionamos por último que $f(\omega)$ la hemos considerado una función continua 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 ω, y esto es lo que se verá en el próximo epígrafe.

Figura 1.1



1.7 Estimación de la densidad espectral mediante alisamientos del periodograma.

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

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

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 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_k.
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)

Teorema 1.14 Sea $\{X_t, -\infty < t < \infty\}$ un proceso lineal

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

donde ξ_t, son variables aleatorias independientes e idénticamente distribuidas, que satisfacen:

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

y siendo la función característica de ξ_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

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

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

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

■

Teorema 1.15 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 1.12 y $K(\theta)$ la función de núcleo que cumple las condiciones (1.7.2). Si $h \to 0$ y $Nh \to \infty$, entonces $\forall |\omega| < \pi$, se tiene:

i) $NhVar\left(\hat{f}(\omega;h)\right) \to \sigma^2(\omega) = f^2(\omega)\frac{1}{2\pi}\int_{-\infty}^{\infty} K^2(\theta)d\theta$ cuando $N \to \infty$ (1.7.4)

ii) $\sqrt{Nh}\left\{\hat{f}(\omega;h) - E\left[\hat{f}(\omega;h)\right]\right\} \to Z$ en distribución. (1.7.5)

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

■

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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:

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

(1.8.1) puede expresarse alternativamente por:

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

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

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 $\hat{f}(\omega;h)$

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

y

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

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

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

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 $\left|f''(\omega)\right|$ es grande, es decir, existe un máximo local en ω, deberemos tomar un ancho de banda h pequeño para reducir el sesgo. Por el contrario, si $\left|f''(\omega)\right|$ es pequeño, podemos tomar un h algo mayor de

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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, de manera que el $MSPE(\omega;h)$, dado por la expresión asintótica (1.8.5), converja a cero si $h \to 0$ y $Nh \to \infty$.

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

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

El cálculo de un ancho de banda óptimo para cada frecuencia ω, 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:

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

Por tanto, h_0 queda definido como:

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

De aquí se obtiene:

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

Obviamente, $AMSPE(h_0) = O(N^{-4/5})$.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Hemos obtenido expresiones para el ancho de banda óptimo, considerado como el que minimiza el MSE o el MISE (la integral del MSE). Pero desgraciadamente estas fórmulas emplean las funciones desconocidas f y f''. Hasta ahora, solo somos capaces de calcular el ancho de banda óptimo para la estimación de f, cuando ya conocemos f. A continuación vamos a ver algunas técnicas, para obtener una elección razonable de h cuando no conocemos f.

Un método que usa los datos $X_1,...X_N$, para obtener un ancho de banda $\hat{h}$, es llamado un "selector de ancho de banda". Selectores de anchos de banda que logre minimizar el AMISE ($\hat{f}_h$) se dicen ser consistente con respecto al MISE.

Supongamos que h está en algún conjunto $H_N \subseteq R_N^+$ de interés. Sea la distancia $d(\hat{f}_h, f)$, la cual denota una medida de precisión para el estimador $\hat{f}_h$. Siguiendo a Shibata (1981), un selector de ancho de banda $\hat{h}_0$ se dice asintóticamente óptimo con

respecto a d cuando $\lim_{N\to\infty}\left[\frac{d(\hat{f}_{\hat{h}_0}, f)}{\inf_{h\in H_N} d(\hat{f}_h, f)}\right] = 1$ con probabilidad 1.

Estas distancias del estimador de núcleo $\hat{f}_h$ a f puede obtenerse de distintas maneras, entre ellas podemos indicar: el error cuadrático medio MSE ($\hat{f}_h$), el error cuadrático integrado ISE($\hat{f}_h$), el error cuadrático medio relativo o porcentaje de error cuadrático medio MSPE($\hat{f}_h$), el MISE($\hat{f}_h$), etc.

Cualquier ancho de banda h_0 que minimice algunas de estas distancias es un ancho de banda óptimo. Una cuestión importante es, ¿cuál es la elección más apropiada para el ancho de banda óptimo h_0? La razón de convergencia a h_{ISE} de cualquier selector de ancho de banda no puede ser más rápido que $N^{-1/10}$ (Härdle y otros 1988), mientras que la razón de convergencia de selectores de ancho de banda a h_{MISE} puede ser tan rápido como $N^{-1/2}$ (Hall y Marron, 1991).

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Beltrao y Bloomfield usaron selectores de anchos de banda, para el caso de la estimación del espectro. Utilizan el porcentaje de error cuadrático medio MSPE como distancia para obtener el ancho de banda local. Para obtener un criterio que refleje las propiedades de $\hat{f}_h$ sobre la totalidad del intervalo, es natural integrar sobre [0,π]. En el contexto de nuestros estimadores discretos es más conveniente sumar, y también por conveniencia omitimos los puntos extremos. Sin embargo, describe el resultado como el error cuadrático medio integrado

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

Existen varios métodos para la selección del ancho de banda de un estimador de núcleo a partir de los datos. Esto lleva a la cuestión natural, ¿cómo comparar los métodos? Alguna idea puede obtenerse a través del análisis asintótico de cada selector de ancho de banda que lleva típicamente a una razón de convergencia del ancho de banda seleccionado, a algún ancha de banda óptimo.

La manera usual de cuantificar el cumplimiento teórico de un selector de ancho de banda particular $\hat{h}$ es a través de una distribución asintótica, típicamente de la forma

$$N^{\nu}\left(\frac{\hat{h}}{h_0} - 1\right) \xrightarrow{d} N(\mu, \sigma^2) \qquad (1.8.9)$$

donde μ y σ^2 solo dependen de f y K (pero no de N), ν es un número racional y h_0 es algún ancho de banda óptimo.

La cantidad $\frac{\hat{h}}{h_0} - 1$ es llamada el error relativo de $\hat{h}$, teniendo en cuenta que el ancho de banda es un parámetro escala. Si $\hat{h}$ satisface (1.8.9), entonces diremos que $\hat{h}$ tiene una razón relativa de convergencia a h_0 de orden $N^{-\nu}$ con varianza relativa σ^2. Por tanto, valores más grandes de ν corresponde a los selectores de ancho de banda que

[1] La siguiente notación indica covergencia en distribución: $\xrightarrow{d}$.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

converge al óptimo más rápido. Además, valores más pequeños de σ^2 corresponde a selectores más estables de anchos de banda.

Método de validación cruzada.

El principio de la validación cruzada es dejar, por un momento, un punto de los datos fuera y seleccionar el valor de h bajo el cual los puntos de los datos perdidos estan mejor representados por el resto de los datos. El estimador que se obtiene con los puntos de los datos, excepto el punto para el cual se hace la estimación se conoce como "leave one out" y se emplea cuando hay independencia entre dicho punto y los demás, por ejemplo en regresión.

Otro tipo de validación cruzada consiste en dividir la muestra aleatoriamente en dos partes y seleccionar el valor de h bajo el cual los puntos de los datos de un grupo estan mejor representados por el otro. Es tipo se utiliza en el caso que cada punto no es independiente de los demás, como ocurre con las frecuencias en el problema de la estimación de la densidad espectral.

1.9 El 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.

La idea del bootstrap es extraordinariamente simple. Sea $X=(X_1, ..., X_n)$ 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:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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, $X^*=\left(X_1^*,\ldots,X_n^*\right)$.

Paso 3. Se aproxima la distribución de $R(X,F)$ por la distribución bootstrap de $R^*=R\left(X^*,F_n\right)$. 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 pivotales. 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}, ..., X^{*B} y a partir de ellas se obtiene un histograma con los correspondientes valores $R\left(X^{*1},\hat{F}\right)$, $R\left(X^{*2},\hat{F}\right)$, ..., $R\left(X^{*B},\hat{F}\right)$ que nos da una aproximación de la distribución bootstrap de R.

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Métrica de Mallows

Damos ahora la definición de la métrica de Mallows en un conjunto de distribuciones de probabilidad sobre un espacio de Banach, pues este concepto es importante para estudiar la validez asintótica del método bootstrap.

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 γ definidas sobre la correspondiente σ-álgebra de Borel que verifican $\int \|x\|^p \gamma(dx) < \infty$. La métrica de Mallows se define en Γ_p del modo siguiente: Dado $\alpha, \beta \in \Gamma_p$

$$d_p(\alpha,\beta) = \inf_{\substack{X \equiv \alpha \\ Y \equiv \beta}} \left\{ E\left[(X-Y)^p \right] \right\}^{1/p} \qquad (1.9.1)$$

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

1.10 Estimación de la función de densidad espectral utilizando la metodología bootstrap.

Daremos ahora el procedimiento usado por Franke y Härdle (1992) para estimar la función de densidad espectral de un proceso lineal general, así como los teoremas que justifican su validez asintótica en el sentido de Bickel y Freedman.

1.10.1 Método bootstrap para la estimación de la densidad espectral.

Sea $\{X_1,..., X_N\}$ 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:

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

$$I(\omega_j)=f(\omega_j)\varepsilon_j \qquad j=1, ..., [N/2] \tag{1.10.1}$$

siendo las variables residuales ε_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 $h_I>0$, y se estiman los residuos a través de

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

y reescalándolos en la forma:

$$\tilde{\varepsilon}_j = \frac{\hat{\varepsilon}_j}{\hat{\varepsilon}_.} \qquad j=1, ..., [N/2]$$

donde

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

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

$$\forall j=1, ..., [N/2], \forall k=1, ..., [N/2] \qquad P\left\{\varepsilon_j^* = \tilde{\varepsilon}_k\right\} = \frac{1}{[N/2]}$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 ε_j^*, siguen la distribución empírica de la muestra $\{\tilde{\varepsilon}_1, ..., \tilde{\varepsilon}_{[N/2]}\}$, con respecto a la cual tienen media uno, $E^*[\varepsilon_j^*] = 1$.

1.10.2 Validez asintótica del método.

Consideremos la cantidad pivotal

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

y su aproximación bootstrap

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

La validez asintótica se entenderá si las distribuciones de ambas cantidades pivotales 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 1.16 Sea $\{X_t, -\infty < n < \infty\}$ un proceso lineal que toma valores reales

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

donde ξ_t, son variables aleatorias independientes e idénticamente distribuidas, que satisfacen:

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

y siendo la función característica de ξ_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

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

Para $N \to \infty$, 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:

$$h \approx N^{-1/5}, \; h_I \to 0 \text{ tal que } \left(N h_I^4\right)^{-1} = O(1), \; g \to 0 \text{ tal que } h/g \to 0.$$

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 $h/g \to 0$, 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 1.17 Sea $\{X_t, -\infty < t < \infty\}$ un proceso lineal que toma valores reales que verifica las hipótesis del teorema (1.17). Sea K una función núcleo que cumple las condiciones (1.7.2). Sea $\hat{f}''(\omega)$ un estimador consistente de $f''(\omega)$. Sea

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

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

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

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:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

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

El teorema nos indica que si se hace una correción del sesgo en el pivote bootstrap, se pueden tomar los anchos de banda h y g de orden óptimo, consiguiéndose de esta manera mejores estimaciones.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO II.

ANÁLISIS ESPECTRAL CON SERIES REPLICADAS.

CAPÍTULO II. ANÁLISIS ESPECTRAL CON SERIES REPLICADAS.

2.1 Introducción.

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.

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 $f(\omega)$ como el valor medio de dicho proceso. A partir de este modelo estiman el espectro poblacional realizando ciertas parametrizaciones de los procesos que forman el modelo.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Introducimos en este capítulo la idea de proceso lineal de coeficientes aleatorios y a continuación se hace la distinción entre modelo de ruido blanco propuesto por Diggle y Al-Wasel (1993) y el modelo general de efectos aleatorios considerado por Hernández (1997). Además se hace una revisión de las propiedades asintóticas del periodograma medio, observándose que constituye un estimador consistente, cuando r es suficientemente grande. En el caso que se disponga también de un número considerable de observaciones de las series temporales, es más adecuado realizar alisamientos del periodograma medio mediante núcleos. Revisamos las propiedades de este estimador, y el algoritmo bootstrap correspondiente, que permite una estimación bootstrap del espectro poblacional.

2.2 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 $\{1, ..., N\}$, da lugar a una serie temporal $\{x_t / t=1, ..., N\}$, 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 independientes en los mismos instantes de tiempo $\{1, ..., N\}$, dando lugar a r series de la forma $\{x_{lt} / l=1, ..., r; t=1, ..., N\}$. Para ello seleccionamos una muestra aleatoria de r objetos o individuos de una población $\{O_1, ..., O_r\}$ 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 $\{x_{lt} / t=1, ..., N\}$ en la frecuencia de Fourier $\omega_j = 2\pi j/N$, el cual viene dado por la siguiente expresión:

$$I_l(\omega_j) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} x_{lt} e^{it\omega_j}\right|^2 \qquad (2.2.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

$$\bar{I}_{\cdot}(\omega_j)=\frac{1}{r}\sum_{l=1}^{r} I_l(\omega_j) \qquad (2.2.2)$$

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

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

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

3) $Cov\left[\bar{I}_{\cdot}(\omega_j),\bar{I}_{\cdot}(\omega_k)\right]=\frac{C}{r}f(\omega_j)f(\omega_k)+\frac{1}{r}O\left(\frac{1}{N^{\alpha}}\right)$

siendo $C=\left[\frac{e}{N}+\frac{2\pi}{N}\left\{F_N(\omega_j+\omega_k)+F_N(\omega_j-\omega_k)\right\}\right]; e=\left(k_4(\varepsilon)/\sigma_{\varepsilon}^4\right)=\left\{E\left[\varepsilon_t^4\right]-3\right\}$ y $F_N(\theta)$ es el núcleo de Fejer.

Es conocido que el periodograma se puede expresar según el siguiente modelo multiplicativo

$$I_l(\omega_j)=f(\omega_j)U_{lj}+R_{lj} \qquad (2.2.3)$$

donde R_{lj} es asintóticamente despreciable, U_{lj} son variables aleatorias independientes en l y asintóticamente independientes en j, y con distribución exponencial de parámetro 1 para $j=1,\ldots[(N-1)/2]$. Bajo este modelo, el periodograma medio definido por (2.2.2), es un estimador asintóticamente centrado y consistente en r para la función de densidad espectral $f(\omega)$.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

El modelo (2.2.3) 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. Llamaremos a este modelo de *efectos fijos.*

2.3 El modelo de efectos aleatorios.

Diggle y Al-Wasel (1993) consideran 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 $\left\{X_{lj}^{(5)} / l = 1,...,8\ j = 0,5,...,90\right\}$, repitiéndolo posteriormente a intervalos de un minuto durante un periodo de observación de una hora $\left\{X_{lj}^{(1)} / l = 1,...,8\ j = 0,1,...,60\right\}$. Inicialmente, consideran que todas las series proceden de un mismo proceso estacionario lineal general $\left\{X_t / t \in Z^+\right\}$, observado sobre diferentes individuos, y por lo tanto, la densidad espectral es común a todos ellos. De acuerdo con el modelo de efectos fijos, fácilmente se obtiene que para cada frecuencia ω_j, el coeficiente de variación de los periodogramas $I_l(\omega_j)$ $l=1, ..., r$ es aproximadamente:

$$CV\left(I_l(\omega_j)\right) = \frac{\sqrt{Var\left(I_l(\omega_j)\right)}}{E\left[I_l(\omega_j)\right]} \approx \frac{f(\omega_j)}{f(\omega_j)} = 1 \qquad (2.3.1)$$

Al disponerse de series replicadas, este coeficiente de variación puede estimarse alternativamente, estimando $E\left[I_l(\omega_j)\right]$ por $\frac{1}{r}\sum_{l=1}^{r} I_l(\omega_j)$ y $Var\left[I_l(\omega_j)\right]$ por $\frac{1}{r-1}\sum_{i=1}^{r}\left(I_i(\omega_j) - \bar{I}(\omega_j)\right)^2$. De acuerdo con la teoría asintótica del periodograma, si representamos los coeficientes de variación estimados, frente a las frecuencias, la gráfica correspondiente debe oscilar alrededor de uno. Diggle y Al-Wasel (1993)

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

obtienen los coeficientes de variación de los periodogramas de las series observadas a intervalos de cinco y de un minuto, notando que el coeficiente de variación de los primeros toma valores próximos a 1, mientras el segundo es mucho mayor. Este último resultado, hace que el modelo de efectos fijos sea 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 $f_l(\omega)$. Siguiendo esto, el modelo multiplicativo queda de la siguiente forma:

$$I_l(\omega_j) = f_l(\omega_j)U_{lj} + R_{lj} \qquad (2.3.2)$$

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.* Por consiguiente, se puede expresar que:

$$f_l(\omega) = f(\omega)Z_l(\omega) \qquad (2.3.3)$$

donde $\{Z_l(\omega)\}_{l=1}^{r}$ son r realizaciones independientes de un proceso estocástico $\{Z(\omega)\}$. Utilizando (2.3.3) podemos expresar (2.3.2) como

$$I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj} \qquad l=1, ..., r\ ;\ j=1, ..., [(N-1)/2] \qquad (2.3.4)$$

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 dentro de cada individuo.

2.4 Procesos lineales de coeficientes aleatorios.

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$X_n = \sum_{k=-\infty}^{\infty} b_k \xi_{n-k} \qquad (2.4.1)$$

donde $\{\xi_n\}$ son variables aleatorias independientes e idénticamente distribuidas con $E[\xi_n]=0$ y $E[\xi_n^2]=\sigma_\xi^2$ y los coeficientes b_k verifican que $\sum_{k=-\infty}^{\infty} b_k^2 < \infty$.

El modelo (2.3.4) para el periodograma individual, puede corresponder 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

$$X_{l,n} = X_n(O_l) = \sum_{k=-\infty}^{\infty} b_k(O_l)\xi_{n-k}(O_l) \qquad (2.4.2)$$

donde

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

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

2.5 El modelo de ruido blanco.

Diggle y Al-Wasel (1993) propusieron el siguiente modelo de ruido blanco:

$$I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj} \qquad l=1, \ldots, r\ ;\ j=1, \ldots, [(N-1)/2] \qquad (2.5.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 $Var\left[Z\left(\omega_j\right)\right]=\varsigma^2$ para toda frecuencia ω_j.

ii) $Z(\omega_j)$ y $Z(\omega_k)$ independientes para todo $j \neq k$.

Estableciendo que $W_{lj} = Z_l\left(\omega_j\right)U_{lj}$ podemos expresar (2.5.1) como

$$I_l\left(\omega_j\right) = f\left(\omega_j\right)W_{lj} \qquad (2.5.2)$$

A partir de este modelo, Diggle y Al-Wasel estiman el espectro $f(\omega)$ suponiendo que el proceso $Z(\omega)$ sigue la distribución de probabilidad $\beta(\kappa,1-\kappa)$.

2.6 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 (2.3.4).

Consideremos r series temporales $\{x_{lt}; l=1, ..., r; t=1, ..., N\}$, 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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

serie temporal podemos calcular el periodograma $I_l(\omega)$ de acuerdo con (2.2.1). Sea el modelo multiplicativo:

$$I_l(\omega_j) = f_l(\omega_j)U_{lj} + R_{lj}, \qquad \omega_j = \frac{2\pi j}{N}; j=1, ..., [N-1/2] \quad (2.6.1)$$

donde R_{lj} es asintóticamente despreciable, las variables residuales son independientes en l, asintóticamente independientes en j y con distribución común exponencial 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 (2.6.1) y para N suficientemente grande se tiene:

$$I_l(\omega_j) = f(\omega_j)Z_l(\omega_j)U_{lj} \qquad (2.6.2)$$

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

i) $E[Z(\omega)]=1 \quad \forall\omega$

ii) $Var[Z(\omega)] = \varsigma_Z^2(\omega) < \infty$

iii) El proceso $\{Z(\omega)\}$ es independiente de la sucesión $\{U_{lj}\}$.

Si expresamos $W_{lj}=Z_l(\omega_j)U_{lj}$, podemos escribir (2.6.2) como:

$$I_l(\omega_j) = f(\omega_j)W_{lj} \qquad l=1, ..., r; j=1, ..., [(N-1)/2] \qquad (2.6.3)$$

donde para cada frecuencia ω_j fija, $\{W_{lj}\}$ son variables aleatorias independientes e idénticamente distribuidas, con distribución F_j de varianza $Var[W_{lj}] = \sigma^2(\omega_j)$ finita. Es inmediato probar que $1 \le \sigma^2(\omega_j) < \infty \ \forall j$. Según se deduce de este modelo, las variables

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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 primer capítulo, 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. Hernández (1997) formula las siguientes proposiciones, que dan las propiedades asintóticas de $\bar{I}(\omega_j)$ como estimador del espectro poblacional $f(\omega_j)$ en el modelo de efectos aleatorios descrito por (2.6.3).

Proposición 2.1 Si cada densidad espectral $f_l(\omega_j)$ verifica una condición de Lipschitz de orden 1, de la forma $|f_l(\omega_2)-f_l(\omega_1)|<K|\omega_2-\omega_1|$ para $|\omega_2-\omega_1|\to 0$, entonces

$$E\left[\bar{I}_r(\omega_j)\right]=f(\omega_j)+O\left(\frac{LnN}{N}\right)$$

■

Proposición 2.2 Bajo las consideraciones del modelo (2.6.3), se verifica:

$$i)\ Var\left[\bar{I}_r(\omega_j)\right]=\frac{f^2(\omega_j)\sigma_j^2}{r}$$

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

■

En el contexto de una única serie temporal $\{X_t\}$ con densidad espectral $f(\omega)$ y siendo $I(\omega)$ el correspondiente periodograma, Zhao-Guo y Hannan (1980) probaron, que bajo determinadas condiciones, las variables $U_j=I(\omega_j)/f(\omega_j)$, $(\omega_j=2\pi j/N;\ j\neq 0,\pm[N/2])$, tienen función de distribución asintótica *1-exp(-x)*. Si en el

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

contexto de las series replicadas, admitimos que $E[U_{lj}]=1$, y teniendo en cuenta además, que $Z_l(\omega_j)$ y U_{lj} son independientes y $E[Z_l(\omega_j)]=1$, podemos razonablemente suponer, al menos para grandes valores de *N*, que $E[W_{lj}]=1$.

Corolario 2.3 Bajo las mismas condiciones anteriores, si $E[W_{lj}]=1;\forall i,j$ entonces $\bar{I}(\omega_j)$ es un estimador centrado y consistente de $f(\omega_j)$.

■

Para construir intervalos de confianza del espectro poblacional $f(\omega)$, usando los últimos resultados obtenidos, es necesario estimar la varianza σ_j^2. Esto puede ser logrado aproximando las variables no observables W_{lj}, por medio de $\hat{W}_{lj}=I_l(\omega_j)/\bar{I}(\omega_j)$, y luego obtener el estimador $\hat{\sigma}_j^2=\sum_{l=1}^{r}\{\hat{W}_{lj}-1\}^2/(r-1)$.

2.7 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 $\{x_{lt}; l=1, ..., r; t=1, ..., N\}$. 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 paramétricos. 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:

$$I_l(\omega_j)=f(\omega_j)W_{lj} \qquad (2.7.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

El algoritmo bootstrap que proponemos 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 (2.7.1)

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

Denotamos por $\mathbf{W}_l = \left(\hat{W}_{l1}, \ldots, \hat{W}_{l[N/2]}\right)$.

Paso 2 Se toman B muestras bootstrap $\left\{\mathbf{W}_1^*, \ldots, \mathbf{W}_r^*\right\}$ de la distribución empírica de $\left\{\mathbf{W}_1, \ldots, \mathbf{W}_r\right\}$, es decir, $\forall l = 1, \ldots, r$,

$$P\left(\mathbf{W}_l^* = \mathbf{W}_l^*\right) = \frac{1}{r}, \ k = 1, \ldots, r.$$

y se define el periodograma bootstrap de cada frecuencia mediante:

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

siendo $\hat{W}_{lj}^*$ la j-ésima componente de $\mathbf{W}_l$.

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

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

Para probar la validez asintótica de este algoritmo bootstrap, Hernández (1997) demuestra el siguiente teorema:

Teorema 2.4 Consideremos el modelo general de efectos aleatorios dado por la expresión (2.6.1). Si $E\left[W_{lj}\right] = 1; \forall i, j$ entonces:

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

siendo d_2 la métrica de Mallows, definida en el epígrafe 1.9 del capítulo I.

■

2.8 Alisamientos mediante núcleo del periodograma medio.

Consideramos r series temporales $\{x_{lt} / l=1, ...,r ; t=1, ..., N\}$ observadas 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 (2.6.2) 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:

$$\hat{f}(\omega;h)=\frac{1}{Nh}\sum_{j=-[N/2]}^{[N/2]} K\left(\frac{\omega-\omega_j}{h}\right)\cdot \bar{I}(\omega_j) \qquad (2.8.1)$$

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 (2.8.1). Denotamos por $\hat{f}_l(\omega;h)$ el estimador de núcleo de la densidad espectral correspondiente al objeto l-ésimo,

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

y utilizándolo en la expresión (2.8.1), tenemos:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\hat{f}(\omega;h)=\frac{1}{r}\sum_{l=1}^{r}\hat{f}_l(\omega;h) \qquad (2.8.2)$$

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 que habría que emplear para cada estimación individual, lo que supondrá que el estimador global tenga menor sesgo.

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

Proposición 2.5 Consideremos el modelo general de efectos aleatorios descrito en el epígrafe 2.6 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 (2.8.1). Supongamos además que la función de núcleo K verifica las condiciones (1.7.2). Si $N\to\infty$, $h\to 0$, $\left(Nh^4\right)^{-1}=O(1)$ entonces, para todo ω tal que $|\omega|<\pi$, el sesgo de $\hat{f}(\omega;h)$ satisface:

$$E\left[\hat{f}(\omega;h)\right]-f(\omega)=\frac{h^2}{2}f''(\omega)+o\left(h^2\right) \qquad (2.8.3)$$

■

Teorema 2.6 Sea un proceso lineal de coeficientes aleatorios, de tal forma que dada una sucesión $\{b_{lk}\}$, el proceso lineal

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

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_{lt}, el proceso $Z(\omega)$ es determinístico puro,

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

lo que supone que $Cov(Z(\omega_j),Z(\omega_k))=\varphi(\omega_j)\varphi(\omega_k)$, donde $\varphi(\omega)=\sqrt{Var[Z(\omega)]}$. Si además $h\to 0$, $Nh\to\infty$ y $\varphi(\omega)>0$ Hernández (1997) obtiene la expresión asintótica de la varianza para $|\omega|<\pi$

$$Var\left[\hat{f}(\omega;h)\right]=\frac{f^2(\omega)}{Nhr}E\left[Z^2(\omega)\right]\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(\theta)d\theta+\frac{f^2(\omega)}{r}Var\left[Z(\omega)\right] \qquad (2.8.4)$$

■

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 ω. 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 $(Nhr)^{-1}$.

Teorema 2.7 Consideremos las condiciones del modelo general de efectos aleatorios descrito en el epígrafe 2.6, y supongamos que el espectro poblacional $f(\omega)$, $Var[Z(\omega)]$ y $Cov(Z(\omega),Z(\omega'))$ son funciones acotadas, siendo además $\forall\omega$, $Var[Z(\omega)]>0$. Entonces, para $h\to 0$, $Nh\to\infty$, $r\to\infty$, Hernández (1997) demuestra que

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

■

2.9 Aproximación bootstrap para la estimación del espectro poblacional utilizando como estimador alisamientos del periodograma medio.

Consideremos el modelo general de efectos aleatorios $I_l(\omega_j)=f(\omega_j)W_{lj}$, descrito

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

en 3.6, con la $Var[Z(\omega)]>0$ para todo ω. Esto garantiza, bajo las condiciones señaladas, que la varianza de $\hat{f}(\omega;h)$ es de orden r^{-1} en todas las frecuencias ω. Proponemos el siguiente algoritmo bootstrap para aproximar la distribución del pivote

$$\sqrt{r}\left\{\frac{\hat{f}(\omega;h)-f(\omega)}{f(\omega)}\right\} \qquad (2.9.1)$$

Paso 1 Las variables no observables W_{lj}, de acuerdo con (2.6.3), se aproximan por:

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

y se obtienen los vectores $\mathbf{W}_l=\left(\hat{W}_{l1},...,\hat{W}_{l[N/2]}\right)$.

Paso 2 Se selecciona un ancho de banda de remuestreo g y se toman B muestras bootstrap $\left\{\mathbf{W}_1^*,....,\mathbf{W}_r^*\right\}$ de la distribución empírica de $\left\{\mathbf{W}_1,....,\mathbf{W}_r\right\}$. 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:

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

Utilizando este algoritmo aproximamos el pivote dado por (2.10.1) a través del siguiente pivote bootstrap

$$\sqrt{r}\left\{\frac{\hat{f}^*(\omega;h,g)-\hat{f}(\omega;g)}{\hat{f}(\omega;g)}\right\} \qquad (2.9.2)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

2.10 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:

$$MSPE(\omega;h)=\frac{E\left\{\hat{f}(\omega;h)-f(\omega)\right\}^2}{f^2(\omega)} \qquad (2.10.1)$$

(2.10.1) puede expresarse alternativamente por:

$$MSPE(\omega;h)=\frac{v^2(\omega)+b^2(\omega)}{f^2(\omega)} \qquad (2.10.2)$$

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

De acuerdo a los resultados asintóticos obtenidos en el epígrafe 2.8

$$E\left[\hat{f}(\omega,h)\right]-f(\omega)=\frac{h^2}{2}f''(\omega)+o(h^2) \qquad (2.10.3)$$

$$Var\left[\hat{f}(\omega,h)\right]=\frac{f^2(\omega)}{Nhr}E\left[Z^2(\omega)\right]\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(\theta)d\theta+\frac{f^2(\omega)}{r}Var\left[Z(\omega)\right]+o(r^{-1}) \qquad (2.10.4)$$

Al sustituir (2.10.3) y (2.10.4) en (2.10.2), se tiene el error asintótico:

$$MSPE(\omega,h)=\left\{\frac{1}{2}h^2\frac{f''(\omega)}{f(\omega)}\frac{1}{2\pi}\int_{-\infty}^{\infty}\theta^2K(\theta)d\theta\right\}^2+$$

$$+\frac{E\left[Z^2(\omega)\right]}{Nhr}\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(\theta)d\theta+\frac{Var\left[Z(\omega)\right]}{r} \qquad (2.10.5)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

$$h(\omega)=\left(\frac{E\left[Z^2(\omega)\right]}{2\pi Nr}\int_{-\infty}^{\infty}K^2(\theta)d\theta\right)^{1/5}\left(\frac{f(\omega)}{f''(\omega)}\right)^{2/5} \qquad (2.10.6)$$

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)

$$AMSPE(h)=\frac{1}{[(N-1)/2]}\sum_{j=1}^{[(N-1)/2]}MSPE\left(\omega_j,h\right) \qquad (2.10.7)$$

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

$$h=\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(\theta)d\theta\Big/Nr\right)^{1/5}\left\{\sum_{j=1}^{[(N-1)/2]}E\left[Z^2\left(\omega_j\right)\right]\Big/\sum_{j=1}^{[(N-1)/2]}\left\{f''(\omega_j)/f(\omega_j)\right\}^2\right\}^{1/5} \qquad (2.10.8)$$

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. Se puede seleccionar ambos parámetros g y h del orden $(Nr)^{-1/5}$. Pero se debe hacer una corrección del sesgo.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:

$$\hat{f}''(\omega,l)=\frac{1}{Nl^3}\sum_{j=-[N/2]}^{[N/2]}K_{(2)}\left(\frac{\omega-\omega_j}{h_2}\right)\bar{I}\left(\omega_j\right) \qquad (2.10.9)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

$$\int_{-\infty}^{\infty} K_{(2)}(u)du = \int_{-\infty}^{\infty} uK_{(2)}(u)du = 0 \quad \int_{-\infty}^{\infty} u^2 K_{(2)}(u)du = 2 \qquad (2.10.10)$$

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

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

$$\sqrt{r}\left\{\frac{\hat{f}^*(\omega,h,g) - \hat{f}_c(\omega,h,g) + \frac{h^2}{2}\hat{f}''(\omega,l)}{\hat{f}(\omega,g)}\right\} \qquad (2.10.11)$$

donde

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

y $\hat{f}''(\omega;l)$ viene dado por la expresión (2.10.9).

2.11 Estimación del ancho de banda óptimo.

De la expresión (2.10.8) se deduce que el ancho de banda óptimo depende de varios parámetros como son $f''(\omega)$, $f(\omega)$, y $E\left[Z^2(\omega)\right]$ 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 (2.10.8) los parámetros desconocidos por estimaciones realizadas con los mismos datos.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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.

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 el valor de h que minimice la expresión

$$\min_h \sum_{j=1}^{[N/2]}\left\{\bar{I}^{(r_2)}(\omega_j)-\hat{f}^{(r_1)}(\omega_j;h)\right\}^2 \qquad (2.11.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO III.

ESTIMACIÓN POLINÓMICA LOCAL DEL.ESPECTRO POBLACIONAL.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO III. ESTIMACIÓN POLINÓMICA LOCAL DEL ESPECTRO POBLACIONAL.

3.1 Introducción.

En este capítulo estudiaremos una clase de estimadores tipo núcleo, llamada "estimadores de núcleo polinómicos locales". Estos estiman la función de densidad espectral para un punto particular, por ajuste local a un polinomio de grado *p*, a los datos, por medio de mínimos cuadrados ponderados. Esta clase incluye un caso especial, el estimador de Nadaraya-Watson, ya que se puede mostrar que corresponde a ajustar un polinomio de grado cero, que es constante local.

Las favorables propiedades de los estimadores de núcleo polinómicos locales en regresión no paramétrica, han sido establecidas por Fan (1992).

Los estimadores de núcleo polinómicos locales tienen la ventaja de que se adaptan tanto al caso de diseño fijo, como al diseño aleatorio, aunque en el contexto de la estimación de la densidad espectral, se trata de diseño fijo equiespaciado.

Obtenemos una estimación polinómica local, de la función de densidad espectral poblacional a partir del modelo multiplicativo

$$\bar{I}(\omega_j) = f(\omega_j)\overline{W}_j \quad j = 1,\ldots,[(N-1)/2]$$

Partiendo de este modelo y sabiendo que $E[\bar{I}(\omega_j)] = f(\omega_j) + O\left(\frac{LnN}{N}\right)$, se conseguirá mejores ajustes cuanto mayor sea *N*, ya que no hay un sesgo añadido, al de la estimación polinómica local.

Obtenemos una expresión general del sesgo y la varianza del estimador, para un ajuste polinómico local de grado *p*, bajo ciertas condiciones. Y aunque nos centramos en

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

la estimación de la función $f(\omega)$, se indica también con que facilidad se puede estimar sus derivadas hasta el orden p.

Además hallamos las expresiones asintóticas del sesgo y varianza de los estimadores lineal local y cuadrático local, así como la distribución asintótica del estimador.

Es importante destacar las ventajas de los estimadores polinómicos locales con respecto a los de núcleo, pues carecen de efectos frontera, es decir el orden del sesgo en la frontera es el mismo que en los puntos interiores. Además en los picos y valles el sesgo de estos es menor que en los estimadores de núcleo.

Damos el concepto de "núcleo equivalente", que permite expresar los estimadores polinómicos locales, en la forma sencilla que se obtienen los estimadores de núcleo, con la diferencia que el "núcleo equivalente", depende de los puntos del diseño y del grado del polinomio. De tal modo que puede adaptarse automáticamente a la estimación en la frontera.

Por último, obtenemos el ancho de banda óptimo, que es de orden $(Nr)^{-1/5}$ en ajuste lineal, y de orden $(Nr)^{-1/9}$ en ajuste cuadrático. Hallando la estimación del ancho de banda óptimo, por un método de validación cruzada.

3.2 Estimación polinómica local del espectro poblacional.

Consideremos el modelo multiplicativo correspondiente a r series temporales $\{x_{lt};\ l=1, \ldots, r;\ t=1, \ldots, N\}$, obtenidas de procesos lineales observados en los mismos instantes de tiempo. Dicho modelo, según vimos en el capítulo anterior, lo podemos expresar como sigue

$$I_l(\omega_j) = f(\omega_j)W_{lj} \qquad l=1, \ldots, r;\ j=1, \ldots, [(N-1)/2] \qquad (3.2.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde para cada frecuencia ω_j fija, $\{W_{lj}\}$ son variables aleatorias independientes en l, asintóticamente independientes en j e idénticamente distribuidas, con distribución F_j de varianza $Var[W_{lj}] = \sigma^2(\omega_j)$ finita. Podemos obtener el siguiente modelo

$$Y_j = \bar{I}(\omega_j) = f(\omega_j)\overline{W}_j \qquad j=1, \ldots, [(N-1)/2] \qquad (3.2.2)$$

en el cual

$$E[Y_j] = E[\bar{I}(\omega_j)] = E[I_l(\omega_j)] = f(\omega_j) + O\left(\frac{LnN}{N}\right)$$

$$Var[Y_j] = Var[\bar{I}(\omega_j)] = \frac{f^2(\omega_j)}{r}\sigma^2(\omega_j)$$

Sea p el grado del polinomio a ser ajustado. Para un punto ω_0 el estimador es obtenido ajustando el polinomio

$$\beta_0 + \beta_1(\omega-\omega_0) + \ldots + \beta_p(\omega-\omega_0)^p$$

a los (ω_j, Y_j), usando mínimos cuadrados ponderados con el núcleo $K_h(\omega_j-\omega_0)$.

Denotamos por $\hat{\beta}_j$ $j=0,\ldots,p$ la solución que minimiza

$$\sum_{j=1}^{N}\left\{Y_j - \sum_{j=0}^{p}\beta_j(\omega_j-\omega_0)^j\right\}^2 K_h(\omega_j-\omega_0) \qquad (3.2.3)$$

donde $K_h(\omega_j-\omega_0) = \frac{1}{h}K\left(\frac{\omega_j-\omega_0}{h}\right)$.

Es obvio del desarrollo de Taylor de $f(\omega)$

$$f(\omega) \approx f(\omega_0) + f'(\omega_0)(\omega-\omega_0) + \frac{f''(\omega_0)}{2}(\omega-\omega_0)^2 + \ldots + \frac{f^{(p)}(\omega_0)}{p!}(\omega-\omega_0)^p$$

que $\hat{f}_v(\omega_0) = v!\hat{\beta}_v$ es un estimador de $f^{(v)}(\omega_0), v = 0,1,\ldots,p$.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Es conveniente trabajar con notación matricial. Denotamos por X la matriz de diseño del problema (3.2.3):

$$X=\begin{pmatrix} 1 & \omega_1-\omega_0 & \dots & (\omega_1-\omega_0)^p \\ 1 & \omega_2-\omega_0 & \dots & (\omega_2-\omega_0)^p \\ .. & \dots & \dots & \dots \\ 1 & \omega_N-\omega_0 & \dots & (\omega_N-\omega_0)^p \end{pmatrix} \qquad Y=\begin{pmatrix} Y_1 \\ : \\ Y_N \end{pmatrix} \qquad \hat{\beta}=\begin{pmatrix} \hat{\beta}_0 \\ : \\ \hat{\beta}_p \end{pmatrix}$$

Además, sea W la matriz diagonal NxN de pesos $W=diag\left\{K_h(\omega_j-\omega_0)\right\}$. Entonces el problema de mínimos cuadrados ponderados (3.2.3) se puede escribir de la siguiente manera

$$\min_{\beta}(Y-X\beta)^t W(Y-X\beta) \quad \text{con } \beta=(\beta_0,\dots,\beta_p)^t$$

El vector solución es proporcionado por la teoría de los mínimos cuadrados ponderados y es dado por

$$\hat{\beta}=(X^tWX)^{-1}X^tWY$$

suponiendo la invertibilidad de $X^t WX$.

Ya que el estimador de $f(\omega_0)$ es $\hat{\beta}_0$ obtenemos

$$\hat{f}(\omega_0;p,h)=e_1^t(X^tWX)^{-1}X^tWY \qquad (3.2.4)$$

donde e_1 es el vector (p+1)x1, que tiene 1 en la primera componente y cero en el resto y $\hat{f}(\omega_0;p,h)$ denota el estimador en el punto ω_0, con un polinomio local de grado p y un ancho de banda h.

Para estimar cualquiera de las derivadas de orden p de $f(\omega_0)$, basta hallar

$$\hat{f}^{(v)}(\omega_0;p,h)=v!e_v^t(X^tWX)^{-1}X^tWY$$

donde e_v es el vector (p+1)x1, que tiene 1 en la componente que ocupa el lugar v+1 y cero en el resto.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Una fórmula simple existe para el estimador de Nadaraya-Watson (p=0)

$$\hat{f}(\omega_0;0,h)=\frac{\sum_{j=1}^{N}K_h(\omega_j-\omega_0)Y_j}{\sum_{j=1}^{N}K_h(\omega_j-\omega_0)}$$

para el estimador lineal local (p=1)

$$\hat{f}(\omega_0;1,h)=N^{-1}\sum_{j=1}^{N}\frac{\left\{\hat{S}_2(\omega_0;h)-\hat{S}_1(\omega_0;h)(\omega_j-\omega_0)\right\}K_h(\omega_j-\omega_0)Y_j}{\hat{S}_2(\omega_0;h)\hat{S}_0(\omega_0;h)-\hat{S}_1(\omega_0;h)^2}$$

donde $\hat{S}_l(\omega_0;h)=N^{-1}\sum_{j=1}^{N}(\omega_j-\omega_0)^l K_h(\omega_j-\omega_0)$ (3.2.5)

y para el estimador cuadrático local (p=2)

$$\hat{f}(\omega_0;2,h)=\frac{T_0\hat{S}_2\hat{S}_4+T_1\hat{S}_2\hat{S}_3+T_2\hat{S}_1\hat{S}_3-T_0\hat{S}_3^2-T_1\hat{S}_1\hat{S}_4-T_2\hat{S}_2^2}{\hat{S}_0\hat{S}_2\hat{S}_4+2\hat{S}_1\hat{S}_2\hat{S}_3-\hat{S}_0\hat{S}_3^2-\hat{S}_1^2\hat{S}_4-\hat{S}_2^3}$$

donde $\hat{S}_l=\hat{S}_l(\omega_0;h)$ y $T_l=T_l(\omega_0;h)=N^{-1}\sum_{j=1}^{N}(\omega_j-\omega_0)^l Y_j K_h(\omega_j-\omega_0)$.

El estimador polinómico local de grado p se puede expresar en general como sigue

$$\hat{f}(\omega_0;p,h)=e_1^t(X^tWX)^{-1}X^tWY=e_1^t\begin{pmatrix}\hat{S}_0 & \hat{S}_1 & .. & \hat{S}_p\\ \hat{S}_1 & \hat{S}_2 & .. & \hat{S}_{p+1}\\ .. & .. & .. & ..\\ \hat{S}_p & \hat{S}_{p+1} & .. & \hat{S}_{2p}\end{pmatrix}^{-1}\begin{pmatrix}T_0\\ T_1\\ ..\\ T_p\end{pmatrix}.$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

3.3 Sesgo y varianza del estimador.

Teorema 3.1 Bajo las condiciones siguientes:

1) El proceso $\{X_{lt}\}$ es un proceso lineal gaussiano de coeficientes aleatorios, de tal forma que dada una sucesión $\{b_{lk}\}$,

$$X_{lt} = \sum_{k=-\infty}^{\infty} b_{lk}\xi_{l,t-k} \quad \text{con} \quad \sum_{k=-\infty}^{\infty} kb_{lk} < \infty \quad a.s.,$$

y donde $\xi_{l,t}$ son independientes, con distribución $N(0,\sigma^2)$.

2) La función de densidad espectral poblacional $f(\omega)$ no se anula en $[-\pi,\pi]$, es $p+1$ veces diferenciable con $f^{(p+1)}(\omega)$ continua en la vecindad de ω_0 si p es impar, $p+2$ veces diferenciable con $f^{(p+2)}(\omega)$ continua en la vecindad de ω_0 si p es par.

3) La función de núcleo $K(\theta)$ es simétrica y no negativa en el intervalo $(-\infty,\infty)$, tiene soporte compacto $[-\kappa,\kappa]$ y es uniformemente Lipschitziana.

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

5) El ancho de banda $h=h_N$ es una sucesión que satisface $h\to 0$ y $Nh\to\infty$ cuando $N\to\infty$.

6) El punto ω_0 para el cual tomamos la estimación satisface $h<\omega_0<\pi-h \quad \forall N\geq N_0$ donde N_0 es fijo. Entonces el sesgo del estimador polinómico local de grado p es

$$b(\omega_0) = E\left[\hat{f}(\omega_0;p,h)\right] - f(\omega_0) = e_1'(X'WX)^{-1}X'WM - f(\omega_0)$$

siendo $M=E[Y]=E\begin{pmatrix} Y_1 \\ Y_2 \\ .. \\ Y_N \end{pmatrix} = \begin{pmatrix} f(\omega_1) \\ f(\omega_2) \\ .. \\ f(\omega_N) \end{pmatrix}$ asintóticamente en N. ■

Corolario 3.2 En el caso de ajuste lineal local el sesgo es como sigue

$$b(\omega_0) = \frac{1}{2}f''(\omega_0)h^2\mu_2(K) + o(h^2)$$

siendo $\mu_2(K) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u^2K(u)du$ ■

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

En efecto,

$$b(\omega_0)= E\left[\hat{f}(\omega_0;1,h)\right]- f(\omega_0) = e_1^t (X^t WX)^{-1} X^t WM - f(\omega_0) \cong$$

$$\frac{1}{2} f''(\omega_0) e_1^t (X^t WX)^{-1} X^t W \begin{pmatrix} (\omega_1-\omega_0)^2 \\ (\omega_2-\omega_0)^2 \\ .. \\ (\omega_N-\omega_0)^2 \end{pmatrix}$$

y como la matriz de diseño es ahora

$$X= \begin{pmatrix} 1 & \omega_1-\omega_0 \\ .. & .. \\ 1 & \omega_N-\omega_0 \end{pmatrix}$$

de acuerdo con el teorema de Taylor

$$f(\omega_j) = f(\omega_0) + f'(\omega_0)(\omega_j-\omega_0) + \frac{f''(\omega_0)}{2}(\omega_j-\omega_0)^2 + \ldots$$

que implica que $M= X\begin{pmatrix} f(\omega_0) \\ f'(\omega_0) \end{pmatrix} + \frac{1}{2} f''(\omega_0) \begin{pmatrix} (\omega_1-\omega_0)^2 \\ (\omega_2-\omega_0)^2 \\ .. \\ (\omega_N-\omega_0)^2 \end{pmatrix} + \ldots$

luego $E\left[\hat{f}(\omega_0;1,h)\right] = e_1^t (X^t WX)^{-1} X^t WM = e_1^t (X^t WX)^{-1} X^t W \; X \begin{pmatrix} f(\omega_0) \\ f'(\omega_0) \end{pmatrix} +$

$$e_1^t (X^t WX)^{-1} X^t W \frac{1}{2} f''(\omega_0) \begin{pmatrix} (\omega_1-\omega_0)^2 \\ (\omega_2-\omega_0)^2 \\ .. \\ (\omega_N-\omega_0)^2 \end{pmatrix} + \ldots = e_1^t \begin{pmatrix} f(\omega_0) \\ f'(\omega_0) \end{pmatrix} +$$

$$\frac{1}{2} f''(\omega_0)\, e_1^t (X^t WX)^{-1} X^t W \begin{pmatrix} (\omega_1-\omega_0)^2 \\ (\omega_2-\omega_0)^2 \\ .. \\ (\omega_N-\omega_0)^2 \end{pmatrix} + \ldots =$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$f(\omega_0)+\frac{1}{2}f''(\omega_0)\, e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^2\\(\omega_2-\omega_0)^2\\..\\(\omega_N-\omega_0)^2\end{pmatrix}+...$, es decir el sesgo de $\hat{f}(\omega_0;1,h)$ es

$$b(\omega_0)=E\left[\hat{f}(\omega_0;1,h)\right]-f(\omega_0)\cong\frac{1}{2}f''(\omega_0)e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^2\\(\omega_2-\omega_0)^2\\..\\(\omega_N-\omega_0)^2\end{pmatrix}$$

Para calcular el término sesgo obtenemos

$$N^{-1}X^tWX=\begin{pmatrix}\hat{S}_0(\omega_0;h) & \hat{S}_1(\omega_0;h)\\ \hat{S}_1(\omega_0;h) & \hat{S}_2(\omega_0;h)\end{pmatrix} \text{ y } N^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^2\\(\omega_2-\omega_0)^2\\..\\(\omega_N-\omega_0)^2\end{pmatrix}=\begin{pmatrix}\hat{S}_2(\omega_0;h)\\ \hat{S}_3(\omega_0;h)\end{pmatrix}.$$

Aproximando $\hat{S}_l(\omega_0;h)=N^{-1}\sum_{j=1}^{N}(\omega_j-\omega_0)^l K_h(\omega_j-\omega_0)\cong\frac{h^l}{2\pi}\int_{-\infty}^{\infty}u^lK(u)du+O(N^{-1})=$

$h^l\mu_l(K)+O(N^{-1})$ (3.3.1),

teniendo en cuenta las hipótesis establecidas. Por último sustituyendo las expresiones obtenidas

$$b(\omega_0)\cong\frac{1}{2}f''(\omega_0)(1,0)(N^{-1}X^tWX)^{-1}N^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^2\\(\omega_2-\omega_0)^2\\..\\(\omega_N-\omega_0)^2\end{pmatrix}=$$

$$\frac{1}{2}f''(\omega_0)(1,0)\begin{pmatrix}1 & 0\\ 0 & \frac{1}{h^2\mu_2(K)}\end{pmatrix}\begin{pmatrix}h^2\mu_2(K)\\0\end{pmatrix}\Rightarrow b(\omega_0)=\frac{1}{2}f''(\omega_0)h^2\mu_2(K)+o(h^2).$$

Corolario 3.3 En el caso de ajuste cuadrático local el sesgo es

$$b(\omega_0)=\frac{1}{24}f^{(IV)}(\omega_0)h^4\frac{\mu_4^2-\mu_6\mu_2}{\mu_4-\mu_2^2}+o(h^4)$$

siendo $\mu_l(K)=\frac{1}{2\pi}\int_{-\infty}^{\infty}u^lK(u)du$.

■

Para obtener este resultado partimos de la matriz de diseño

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$X=\begin{pmatrix} 1 & \omega_1-\omega_0 & (\omega_1-\omega_0)^2 \\ .. & .. & .. \\ 1 & \omega_N-\omega_0 & (\omega_N-\omega_0)^2 \end{pmatrix}$$

$$M= X\begin{pmatrix} f(\omega_0) \\ f'(\omega_0) \\ \frac{1}{2}f''(\omega_0) \end{pmatrix}+\frac{1}{6}f'''(\omega_0)\begin{pmatrix} (\omega_1-\omega_0)^3 \\ (\omega_2-\omega_0)^3 \\ .. \\ (\omega_N-\omega_0)^3 \end{pmatrix}+\frac{1}{4^!}f^{(IV)}(\omega_0)\begin{pmatrix} (\omega_1-\omega_0)^4 \\ (\omega_2-\omega_0)^4 \\ .. \\ (\omega_N-\omega_0)^4 \end{pmatrix}$$

$$\text{luego } E\left[\hat{f}(\omega_0;2,h)\right]=e_1^t(X^tWX)^{-1}X^tWM=e_1^t(X^tWX)^{-1}X^tW\ X\begin{pmatrix} f(\omega_0) \\ f'(\omega_0) \\ \frac{1}{2}f''(\omega_0) \end{pmatrix}+$$

$$e_1^t(X^tWX)^{-1}X^tW\frac{1}{6}f'''(\omega_0)\begin{pmatrix} (\omega_1-\omega_0)^3 \\ (\omega_2-\omega_0)^3 \\ .. \\ (\omega_N-\omega_0)^3 \end{pmatrix}+ e_1^t(X^tWX)^{-1}X^tW\frac{1}{4^!}f^{(IV)}(\omega_0)\begin{pmatrix} (\omega_1-\omega_0)^4 \\ (\omega_2-\omega_0)^4 \\ .. \\ (\omega_N-\omega_0)^4 \end{pmatrix}$$

$$=f(\omega_0)+\frac{1}{6}f'''(\omega_0)\ e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix} (\omega_1-\omega_0)^3 \\ (\omega_2-\omega_0)^3 \\ .. \\ (\omega_N-\omega_0)^3 \end{pmatrix}+$$

$$\frac{1}{4^!}f^{(IV)}(\omega_0)\ e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix} (\omega_1-\omega_0)^4 \\ (\omega_2-\omega_0)^4 \\ .. \\ (\omega_N-\omega_0)^4 \end{pmatrix}. \text{ Por lo tanto}$$

$$b(\omega_0)=E\left[\hat{f}(\omega_0;2,h)\right]-f(\omega_0)\cong\frac{1}{6}f'''(\omega_0)e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix} (\omega_1-\omega_0)^3 \\ (\omega_2-\omega_0)^3 \\ .. \\ (\omega_N-\omega_0)^3 \end{pmatrix}+$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\frac{1}{4!}f^{(IV)}(\omega_0)e_1^t(X^tWX)^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^4\\(\omega_2-\omega_0)^4\\..\\(\omega_N-\omega_0)^4\end{pmatrix}.$$

Para calcular el término sesgo expresamos

$$N^{-1}X^tWX=\begin{pmatrix}\hat{S}_0(\omega_0;h) & \hat{S}_1(\omega_0;h) & \hat{S}_2(\omega_0;h)\\ \hat{S}_1(\omega_0;h) & \hat{S}_2(\omega_0;h) & \hat{S}_3(\omega_0;h)\\ \hat{S}_2(\omega_0;h) & \hat{S}_3(\omega_0;h) & \hat{S}_4(\omega_0;h)\end{pmatrix}, \text{ hallamos su inversa}$$

$$N(X^tWX)^{-1}=\begin{pmatrix}\hat{S}_0(\omega_0;h) & \hat{S}_1(\omega_0;h) & \hat{S}_2(\omega_0;h)\\ \hat{S}_1(\omega_0;h) & \hat{S}_2(\omega_0;h) & \hat{S}_3(\omega_0;h)\\ \hat{S}_2(\omega_0;h) & \hat{S}_3(\omega_0;h) & \hat{S}_4(\omega_0;h)\end{pmatrix}^{-1}=$$

$$\begin{pmatrix}\frac{\mu_4}{\mu_4-\mu_2^2} & 0 & \frac{-\mu_2}{h^2(\mu_4-\mu_2^2)}\\ 0 & \frac{1}{h^2\mu_2} & 0\\ \frac{\mu_2}{h^2(\mu_4-\mu_2^2)} & 0 & \frac{1}{h^4(\mu_4-\mu_2^2)}\end{pmatrix}, \text{ haciendo la aproximación}$$

$\hat{S}_l(\omega_0;h)=N^{-1}\sum_{j=1}^{N}(\omega_j-\omega_0)^l K_h(\omega_j-\omega_0)\cong h^l\mu_l(K)+O(N^{-1})$. De la misma manera

$$N^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^3\\(\omega_2-\omega_0)^3\\..\\(\omega_N-\omega_0)^3\end{pmatrix}=\begin{pmatrix}\hat{S}_3(\omega_0;h)\\ \hat{S}_4(\omega_0;h)\\ \hat{S}_5(\omega_0;h)\end{pmatrix}=\begin{pmatrix}0\\ h^4\mu_4\\ 0\end{pmatrix} \quad \text{y}$$

$$N^{-1}X^tW\begin{pmatrix}(\omega_1-\omega_0)^4\\(\omega_2-\omega_0)^4\\..\\(\omega_N-\omega_0)^4\end{pmatrix}=\begin{pmatrix}\hat{S}_4(\omega_0;h)\\ \hat{S}_5(\omega_0;h)\\ \hat{S}_6(\omega_0;h)\end{pmatrix}=\begin{pmatrix}h^4\mu_4\\ 0\\ h^6\mu_6\end{pmatrix}$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

luego

$$b(\omega_0)=\frac{1}{4!}f^{(IV)}(\omega_0)(1,0,0)\begin{pmatrix} \frac{\mu_4}{\mu_4-\mu_2^2} & 0 & \frac{-\mu_2}{h^2(\mu_4-\mu_2^2)} \\ 0 & \frac{1}{h^2\mu_2} & 0 \\ \frac{\mu_2}{h^2(\mu_4-\mu_2^2)} & 0 & \frac{1}{h^4(\mu_4-\mu_2^2)} \end{pmatrix}\begin{pmatrix} h^4\mu_4 \\ 0 \\ h^6\mu_6 \end{pmatrix}\Rightarrow$$

$$b(\omega_0)=\frac{1}{24}f^{(IV)}(\omega_0)h^4\frac{\mu_4^2-\mu_6\mu_2}{\mu_4-\mu_2^2}+o(h^4).$$

Teorema 3.4 Bajo las condiciones 1), 2), 3) y 4) del teorema 3.1 y considerando además la hipótesis siguiente:

1) $\sigma^2(\omega)$ tiene derivada continua en la vecindad de ω_0,

entonces la varianza del estimador polinómico local de grado p es

$$Var\left[\hat{f}(\omega_0;p,h)\right]=e_1^t(X^tWX)^{-1}X^t\Sigma X(X^tWX)^{-1}e_1$$

siendo $\Sigma=WVW$ y $V=E\left\{(Y-E[Y])(Y-E[Y])^t\right\}=\begin{pmatrix} v_1 & v_{12} & .. & v_{1N} \\ v_{21} & v_2 & .. & v_{2N} \\ .. & .. & .. & .. \\ v_{N1} & v_{N2} & .. & v_N \end{pmatrix}$ cuyos

términos asintóticamente toman los valores siguientes:

$$v_j=v(\omega_j)=Var[Y_j]=Var\left[\bar{I}(\omega_j)\right]=\frac{f^2(\omega_j)\sigma^2(\omega_j)}{r} \quad \text{y}$$

$$v_{jj'}=Cov[Y_j,Y_{j'}]=f(\omega_j)f(\omega_{j'})Cov\left(\overline{W}_j,\overline{W}_{j'}\right)=\frac{f(\omega_j)f(\omega_{j'})}{r}Cov(W_{lj},W_{lj'})\cong$$

$$\frac{f(\omega_j)f(\omega_{j'})}{r}Cov(Z_l(\omega_j),Z_l(\omega_{j'}))=\frac{f(\omega_j)f(\omega_{j'})}{r}\Psi(\omega_j,\omega_{j'}) \qquad (3.2.2)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde $\Psi(\omega_j,\omega_{j'})=Cov(Z_l(\omega_j),Z_l(\omega_{j'}))\quad j\neq j'$.

■

Se puede obtener además una aproximación de la varianza, descomponiendo la

$$\text{matriz } \Sigma=\begin{pmatrix} v_1K_1^2 & v_{12}K_1K_2 & .. & v_{1N}K_1K_N \\ v_{21}K_2K_1 & v_2K_2^2 & .. & v_{2N}K_2K_N \\ .. & .. & .. & .. \\ v_{N1}K_NK_1 & v_{N2}K_NK_2 & .. & v_NK_N^2 \end{pmatrix}=\Sigma_1+\Sigma_2$$

$$\text{siendo } \Sigma_1=\begin{pmatrix} v_1K_1^2 & 0 & .. & 0 \\ 0 & v_2K_2^2 & .. & 0 \\ .. & .. & .. & .. \\ 0 & 0 & .. & v_NK_N^2 \end{pmatrix} \text{ y } \Sigma_2=\begin{pmatrix} 0 & v_{12}K_1K_2 & .. & v_{1N}K_1K_N \\ v_{21}K_2K_1 & 0 & .. & v_{2N}K_2K_N \\ .. & .. & .. & .. \\ v_{N1}K_NK_1 & v_{N2}K_NK_2 & .. & 0 \end{pmatrix}$$

donde se ha utilizado la notación más abreviada $K_h(\omega_j-\omega_0)=K_j$.

Y calculamos

$$N^{-1}(X^t\Sigma_1X)=\begin{pmatrix} \hat{S}_0' & \hat{S}_1' & .. & \hat{S}_p' \\ \hat{S}_1' & \hat{S}_2' & .. & \hat{S}_{p+1}' \\ .. & .. & .. & .. \\ \hat{S}_p' & \hat{S}_{p+1}' & .. & \hat{S}_{2p}' \end{pmatrix} \quad \text{siendo}$$

$$\hat{S}_l'=\frac{N^{-1}}{h^2}\sum_{j=1}^{N}(\omega_j-\omega_0)^lK^2\left(\frac{\omega_j-\omega_0}{h}\right)v(\omega_j)\cong$$

$$\frac{h^l}{2\pi h}\left[v(\omega_0)\int_{-\infty}^{\infty}u^lK^2(u)du+v'(\omega_0)h\int_{-\infty}^{\infty}u^{l+1}K^2(u)du\right] \text{ y}$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$N^{-1}(X'WX)=\begin{pmatrix}\hat{S}_0 & \hat{S}_1 & .. & \hat{S}_p\\ \hat{S}_1 & \hat{S}_2 & .. & \hat{S}_{p+1}\\ .. & .. & .. & ..\\ \hat{S}_p & \hat{S}_{p+1} & .. & \hat{S}_{2p}\end{pmatrix} \text{ siendo}$$

$$\hat{S}_l=\frac{N^{-1}}{h}\sum_{j=1}^{N}(\omega_j-\omega_0)^l K\left(\frac{\omega_j-\omega_0}{h}\right)\cong\frac{1}{2\pi}h^l\int_{-\infty}^{\infty}u^l K(u)du$$

Los términos de estas expresiones para el grado p considerado, aproximados según se ha indicado, se sustituyen en la expresión general de la varianza, dando como resultado el primer término de la varianza del estimador polinómico local.

Para obtener el segundo término de la varianza, hallamos

$$N^{-2}(X'\Sigma_2 X)=\begin{pmatrix}a_{00} & a_{01} & .. & a_{0p}\\ a_{10} & a_{11} & .. & a_{1p}\\ .. & .. & .. & ..\\ a_{p0} & a_{p1} & .. & a_{pp}\end{pmatrix} \text{ siendo}$$

$$a_{ll'}=\frac{1}{N^2}\sum_j\sum_{j'}(\omega_j-\omega_0)^l(\omega_{j'}-\omega_0)^{l'}v_{jj'}K_jK_{j'}\qquad j\neq j'$$

Corolario 3.5 Para el caso de estimación lineal local, la varianza es

$$Var\left[\hat{f}(\omega_0;1,h)\right]=\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du+\frac{1}{r}f^2(\omega_0)Var\left[Z(\omega_0)\right]+o\left(r^{-1}\right)$$

si $Var\left[Z(\omega_0)\right]>0$.

■

Para obtener el primer término de la varianza, basta calcular

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$N(X'WX)^{-1}=\begin{pmatrix}1 & 0\\ 0 & \dfrac{1}{h^2\mu_2(K)}\end{pmatrix}$$

$N^{-1}(X'\Sigma_1 X)=\begin{pmatrix}\hat{S}_0' & \hat{S}_1'\\ \hat{S}_1' & \hat{S}_2'\end{pmatrix}$ y sustituir en

$$e_1' N(X'WX)^{-1}N^{-2}(X'\Sigma_1 X)N(X'WX)^{-1}e_1=\frac{1}{N}\hat{S}_0'=\frac{1}{Nh}v(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du=$$

$$\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du$$

y para obtener el segundo término, sustituimos $N^{-2}(X'\Sigma_2 X)=\begin{pmatrix}a_{00} & a_{01}\\ a_{10} & a_{11}\end{pmatrix}$ en

$$e_1' N(X'WX)^{-1}N^{-2}(X'\Sigma_2 X)N(X'WX)^{-1}e_1=a_{00}=N^{-2}\sum_j\sum_{\substack{j'\\ j\neq j'}}v_{jj'}K_jK_{j'}\cong$$

$$\frac{1}{N^2h^2}\sum_j\sum_{j'}K\left(\frac{\omega_j-\omega_0}{h}\right)K\left(\frac{\omega_{j'}-\omega_0}{h}\right)f(\omega_j)f(\omega_{j'})\frac{\Psi(\omega_j,\omega_{j'})}{r}\ \text{(aplicando (3.3.2))}$$

$$\cong\frac{1}{4\pi^2}\frac{f^2(\omega_0)}{r}\iint K(u)K(v)\Psi(\omega_0+hu,\omega_0+hv)\cong\frac{f^2(\omega_0)Var[Z(\omega_0)]}{r}.$$

Corolario 3.6 Para la estimación cuadrático local, la varianza es

$$Var\left[\hat{f}(\omega_0;2,h)\right]=\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{\mu_4^2\mu_0(K^2)-2\mu_2\mu_4\mu_2(K^2)+\mu_2^2\mu_4(K^2)}{(\mu_4-\mu_2^2)^2}+$$

$$\frac{1}{r}f^2(\omega_0)Var[Z(\omega_0)]\frac{\mu_4^2+\mu_2^2}{(\mu_4-\mu_2^2)^2}+o\left(r^{-1}\right).$$

■

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Vamos a obtener el primer término sustituyendo, en la expresión general de la varianza, $e_1^t N(X^tWX)^{-1}N^{-2}(X^t\Sigma_1 X)N(X^tWX)^{-1}e_1$ los términos siguientes,

$$N^{-1}(X^t\Sigma_1 X)=\begin{pmatrix}\hat{S}_0' & \hat{S}_1' & \hat{S}_2'\\ \hat{S}_1' & \hat{S}_2' & \hat{S}_3'\\ \hat{S}_2' & \hat{S}_3' & \hat{S}_4'\end{pmatrix} \quad \text{y} \quad N(X^tWX)^{-1}=\begin{pmatrix}\frac{\mu_4}{\mu_4-\mu_2^2} & 0 & \frac{-\mu_2}{h^2(\mu_4-\mu_2^2)}\\ 0 & \frac{1}{h^2\mu_2} & 0\\ \frac{-\mu_2}{h^2(\mu_4-\mu_2^2)} & 0 & \frac{1}{h^4(\mu_4-\mu_2^2)}\end{pmatrix}$$

de aquí que, $e_1^t N(X^tWX)^{-1}N^{-2}(X^t\Sigma_1 X)N(X^tWX)^{-1}e_1$

$$\cong N^{-1}\frac{1}{h^4(\mu_4-\mu_2^2)^2}\left(h^4\mu_4^2\hat{S}_0'-2h^2\mu_2\mu_4\hat{S}_2'+\mu_2^2\hat{S}_4'\right)$$

$$\cong N^{-1}\frac{1}{h^4(\mu_4-\mu_2^2)^2}\left(h^4\mu_4^2\frac{1}{h}\mu_0(K^2)v(\omega_0)-2h^2\mu_2\mu_4 h\mu_2(K^2)v(\omega_0)+\mu_2^2h^3\mu_4(K^2)v(\omega_0)\right)$$

$$=\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{\mu_4^2\mu_0(K^2)-2\mu_2\mu_4\mu_2(K^2)+\mu_2^2\mu_4(K^2)}{(\mu_4-\mu_2^2)^2}.$$

Para hallar el segundo término, obtenemos $N^{-2}(X^t\Sigma_2 X)=\begin{pmatrix}a_{00} & a_{01} & a_{02}\\ a_{10} & a_{11} & a_{12}\\ a_{20} & a_{21} & a_{22}\end{pmatrix}$

y sustituimos en la expresión

$$e_1^t N(X^tWX)^{-1}N^{-2}(X^t\Sigma_2 X)N(X^tWX)^{-1}e_1$$

$$\cong\frac{1}{h^4(\mu_4-\mu_2^2)^2}\left(h^4\mu_4^2a_{00}-2h^2\mu_2\mu_4a_{02}+\mu_2^2a_{22}\right)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\cong \frac{f^2(\omega_0)Var[Z(\omega_0)]}{r}\frac{\mu_4^2+\mu_2^2}{(\mu_4-\mu_2^2)^2},$$

ya que $a_{00} \cong \frac{f^2(\omega_0)Var[Z(\omega_0)]}{r}$ $a_{02} \cong 0$ y $a_{22} \cong \frac{f^2(\omega_0)Var[Z(\omega_0)]}{r} h^4\mu_2^2$.

3.4 Comportamiento cerca de la frontera.

Vamos a investigar ahora el cumplimiento del estimador polinómico local cerca de la frontera.

Hacemos el estudio en el intervalo $[0,\pi]$. Anteriormente evitamos los puntos frontera tomando $h<\omega<\pi-h$. Supongamos ahora que $\omega=\omega_n=\alpha h$ donde $0\leq\alpha<\pi$, es decir es una sucesión de puntos situados en la frontera izquierda.

Los momentos incompletos de K son denotados por $\mu_{l,\alpha} = \int_{-\alpha}^{\pi} z^l K(z)dz$.

En el caso, por ejemplo, de estimación lineal local,

$$n^{-1}X^tWX = \begin{pmatrix} \mu_{0,\alpha}(K)+o(1) & h\mu_{1,\alpha}(K)+o(h) \\ h\mu_{1,\alpha}(K)+o(h) & h^2\mu_{2,\alpha}(K)+o(h^2) \end{pmatrix} \quad \text{y}$$

$$n^{-1}X^tW\begin{pmatrix} (\omega_1-\omega_0)^2 \\ \vdots \\ (\omega_n-\omega_0)^2 \end{pmatrix} = \begin{pmatrix} h^2\mu_{2,\alpha}(K)+O(h^2) \\ h^3\mu_{3,\alpha}(K)+O(h^3) \end{pmatrix}.$$ Esto lleva a una expresión del sesgo de

la forma

$$E\left[\hat{f}(\omega_0;1,h)\right]-f(\omega_0) = \frac{1}{2}f''(\omega_0)h^2Q_\alpha(K)+o(h^2)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde $Q_\alpha(K)=\dfrac{\mu'^{2}_{2\alpha}(K)-\mu_{1,\alpha}(K)\mu_{3,\alpha}(K)}{\mu_{2,\alpha}(K)\mu_{0,\alpha}(K)-\mu'^{2}_{1\alpha}(K)}$. Por tanto, el sesgo en la frontera de $\hat{f}(\omega_0;1,h)$ es de la misma forma que en el interior, excepto que la constante dependiente del núcleo es diferente. Así que, como se puede observar, el sesgo en la frontera es del mismo orden que en el interior. Esta es una destacable diferencia del método de estimación de núcleo, el cual presenta un sesgo considerable en la frontera.

3.5 Núcleos equivalentes.

El estimador polinómico local se puede representar en términos de un núcleo equivalente.

Podemos escribir el estimador

$$\hat{f}(\omega_0;p,h)=e_1^t(X^tWX)^{-1}X^tWY=\sum_{j=1}^{N}W_N\left(\frac{\omega_j-\omega_0}{h}\right)Y_j \qquad \text{siendo}$$

$$W_N(u)=\frac{1}{h}e_1^tS_N^{-1}\left(1,uh,\ldots(uh)^p\right)^tK(u) \qquad \text{donde}$$

$$S_N=X^tWX=N\begin{pmatrix}\hat{S}_0 & \hat{S}_1 & .. & \hat{S}_p\\ \hat{S}_1 & \hat{S}_2 & .. & \hat{S}_{p+1}\\ .. & .. & .. & ..\\ \hat{S}_{p+1} & \hat{S}_{p+2} & .. & \hat{S}_{2p}\end{pmatrix} \quad \text{y } \hat{S}_l \text{ viene dado por (3.2.5).}$$

La expresión anterior revela que el estimador $\hat{f}(\omega_0;p,h)$ es bastante semejante a un estimador de núcleo convencional, excepto que el núcleo W_N depende de los puntos del diseño y de su localización. Esto explica por qué el ajuste polinómico local puede adaptarse automáticamente a varios diseños y a la estimación en la frontera.

Podemos expresar $S_N=NHSH\{1+o(1)\}$, siendo $H=diag\ (1,h,\ldots,h^p)$ y

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$S = \begin{pmatrix} \mu_0 & \mu_1 & .. & \mu_p \\ \mu_1 & \mu_2 & .. & \mu_{p+1} \\ .. & .. & .. & .. \\ \mu_p & \mu_{p+1} & .. & \mu_{2p} \end{pmatrix}$, para ello basta aplicar la expresión (3.3.1). De esta manera podemos escribir $\hat{f}(\omega_0; p, h) = \frac{1}{Nh} \sum_{j=1}^{N} K_p^* \left(\frac{\omega_j - \omega_0}{h} \right) Y_j \{1 + o(1)\}$ donde

$K_p^*(u) = e_1^t S^{-1} (1, u, ..., u^p)^t K(u) = \sum_{k=0}^{p} (S^k u^k) K(u)$, siendo S^k los elementos de la primera fila de S^{-1}. A $K_p^*(u)$ se le llama núcleo equivalente y se ha notado con subíndice p para enfatizar su dependencia del grado del polinomio.

3.6 Distribución asintótica del estimador.

Sea la cantidad pivotal

$$\sqrt{r} \frac{\hat{f}(\omega; p, h_N) - E\left[\hat{f}(\omega; p, h_N)\right]}{f(\omega)\varphi(\omega)} = \frac{1}{\sqrt{r}} \sum_{l=1}^{r} \frac{\hat{f}_l(\omega; p, h_N) - E\left[\hat{f}(\omega; p, h_N)\right]}{f(\omega)\varphi(\omega)} = \frac{1}{\sqrt{r}} \sum_{l=1}^{r} \xi_{lN}(\omega, h) \tag{3.6.1}$$

siendo $\varphi(\omega) = \sqrt{VarZ(\omega)}$.

Estableceremos condiciones para las variables $\{\xi_{lN}(\omega)\}$, en orden a que la cantidad pivotal (3.6.1), tenga límite en distribución para $h \to 0, N \to \infty$ y $r = r(N) \to \infty$.

Lema 3.7 Si $E\left[|\xi_{lN}(\omega, h)|\right] \leq C; \forall \omega, h, l, N$ entonces

$\sqrt{r} \frac{\hat{f}(\omega; p, h_N) - E\left[\hat{f}(\omega; p, h_N)\right]}{f(\omega)\varphi(\omega)}$ tiene distribución infinitamente divisible.

■

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Demostración

Para ω, h y N fijos $\xi_{1,N}(\omega),\ldots,\xi_{r,N}(\omega)$ son variables independientes e idénticamente distribuidas. Siendo $r = r(N) \to \infty$ para $N \to \infty$.

$$\text{Además, } \forall \varepsilon > 0,\ \lim_{N\to\infty} \max_{1\le l\le r} P\left(\frac{|\xi_{lN}(\omega,h)|}{\sqrt{r}} > \varepsilon\right) = \lim_{N\to\infty} P\left(|\xi_{lN}(\omega,h)| > \varepsilon\sqrt{r}\right)$$

$$\le \lim_{N\to\infty} \frac{E\left[|\xi_{lN}(\omega,h)|\right]}{\varepsilon\sqrt{r}} \to 0.$$

Teorema 3.8 Si *i)* $E\left[|\xi_{lN}(\omega,h)|\right] \le C;\ \forall \omega, h, l, N$

y

$$ii)\ P\left(|\xi_{lN}(\omega,h)| > \varepsilon\sqrt{r}\right) = o\left(r^{-1}\right) \text{ entonces } \forall \omega, h$$

$$\sqrt{r}\,\frac{\hat{f}(\omega;p,h_N) - E\left[\hat{f}(\omega;p,h_N)\right]}{f(\omega)\varphi(\omega)} \xrightarrow{d} N(0,1)$$

■

Demostración

Por el lema 3.7, $\sqrt{r}\,\dfrac{\hat{f}(\omega;p,h_N) - E\left[\hat{f}(\omega;p,h_N)\right]}{f(\omega)\varphi(\omega)}$ converge en distribución, a una distribución infinitamente divisible. Veamos que este límite es la distribución $N(0,1)$.

En efecto, $\forall \varepsilon > 0$ sea

$$P\left(\max_{1\le l\le r(N)} \frac{|\xi_{lN}(\omega,h)|}{\sqrt{r}} > \varepsilon\right) = 1 - P\left(\max_{1\le l\le r(N)} \frac{|\xi_{lN}(\omega,h)|}{\sqrt{r}} \le \varepsilon\right) = 1 - \prod_{l=1}^{r} P\left(|\xi_{lN}(\omega,h)| \le \varepsilon\sqrt{r}\right) =$$

$$1 - \left\{1 - P\left(|\xi_{lN}(\omega,h)| > \varepsilon\sqrt{r}\right)\right\}^r = 1 - \left\{1 - o\left(r^{-1}\right)\right\}^r.$$

Ahora dado que $r = r(N) \to \infty$ para $N \to \infty \Rightarrow \lim_{N\to\infty} P\left(\max_{1\le l\le r(N)} \dfrac{|\xi_{lN}(\omega,h)|}{\sqrt{r}} > \varepsilon\right) = 0.$

Por tanto,

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\frac{1}{\sqrt{r(N)}}\sum_{l=1}^{r(N)}\xi_{lN}(\omega,h)\xrightarrow[N\to\infty]{d}N(0,1).$$

3.7 Ancho de banda óptimo.

La elección del ancho de banda juega un papel crucial en cualquier técnica de alisamiento. Consideramos como ancho de banda óptimo aquel que minimiza el porcentaje de error cuadrático medio (Mean Square Percentage Error MSPE), que se puede expresar por:

$$MSPE\left(\hat{f}(\omega_0;p,h)\right)=\frac{b^2(\omega_0)+v(\omega_0)}{f^2(\omega_0).} \qquad (3.7.1)$$

De acuerdo con las expresiones asintóticas del sesgo y la varianza obtenidas anteriormente para el estimador lineal local

$$b(\omega_0)=\frac{1}{2}f''(\omega_0)h^2\mu_2(K)$$

$$Var\left[\hat{f}(\omega_0;1,h)\right]=\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du+\frac{1}{r}f^2(\omega_0)Var\left[Z(\omega_0)\right]$$

y sustituyendo estas expresiones en (3.7.1), tenemos el porcentaje de error cuadrático medio asintótico,

$$AMSPE\left(\hat{f}(\omega_0;1,h)\right)=\frac{1}{4}h^4\mu_2^2(K)\left[\frac{f''(\omega_0)}{f(\omega_0)}\right]^2+$$

$$\frac{1}{Nhr}\sigma^2(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du+\frac{Var\left[Z(\omega_0)\right]}{r}$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Al minimizar esta expresión, obtenemos el ancho de banda óptimo asintótico del estimador lineal local

$$h=\left[\frac{\sigma^2(\omega_0)\frac{1}{2\pi}\int_{-\infty}^{\infty}K^2(u)du}{\mu_2^2(K)}\right]^{1/5}\left[\frac{f(\omega_0)}{f''(\omega_0)}\right]^{2/5}(Nr)^{-1/5}$$

De la misma manera podemos obtener el ancho de banda óptimo asintótico del estimador cuadrático local sustituyendo en (3.7.1) las expresiones asintóticas del sesgo y la varianza siguientes:

$$b(\omega_0)=\frac{1}{24}f^{(IV)}(\omega_0)h^4\frac{\mu_4^2-\mu_6\mu_2}{\mu_4-\mu_2^2}$$

$$Var\left[\hat{f}(\omega_0;2,h)\right]=\frac{1}{Nhr}f^2(\omega_0)\sigma^2(\omega_0)\frac{\mu_4^2\mu_0(K^2)-2\mu_2\mu_4\mu_2(K^2)+\mu_2^2\mu_4(K^2)}{(\mu_4-\mu_2^2)^2}+$$

$$\frac{f^2(\omega_0)Var[Z(\omega_0)]}{r}\frac{\mu_4^2+\mu_2^2}{(\mu_4-\mu_2^2)^2}.$$

obtenemos

$$AMSPE\left(\hat{f}(\omega_0;2,h)\right)=\frac{1}{576}h^8\left[\frac{f^{(IV)}(\omega_0)}{f(\omega_0)}\right]^2\left[\frac{\mu_4^2-\mu_6\mu_2}{\mu_4-\mu_2^2}\right]^2+$$

$$\frac{1}{Nhr}\sigma^2(\omega_0)\frac{\mu_4^2\mu_0(K^2)-2\mu_2\mu_4\mu_2(K^2)+\mu_2^2\mu_4(K^2)}{(\mu_4-\mu_2^2)^2}+\frac{1}{r}Var[Z(\omega_0)]\frac{\mu_4^2+\mu_2^2}{(\mu_4-\mu_2^2)^2}.$$

Por último minimizando la expresión anterior obtenemos

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$h=\left[\frac{\sigma^2(\omega_0)\dfrac{\mu_4^2\mu_0(K^2)-2\mu_2\mu_4\mu_2(K^2)+\mu_2^2\mu_4(K^2)}{(\mu_4-\mu_2^2)^2}}{\dfrac{8}{576}\left[\dfrac{f^{(IV)}(\omega_0)}{f(\omega_0)}\right]^2\left[\dfrac{\mu_4^2-\mu_6\mu_2}{\mu_4-\mu_2^2}\right]^2}\right]^{1/9}(Nr)^{-1/9}$$

3.8 Selección del ancho de banda.

En la práctica, el h óptimo no puede calcularse exactamente, ya que depende de la función desconocida $f(\omega)$. En este punto se plantea un método de estimación del h óptimo.

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

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

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

$$\min_h \sum_{j=1}^{[N/2]}\left\{\bar{I}^{(r_2)}(\omega_j)-\hat{f}^{(r_1)}(\omega_j;h)\right\}^2 \qquad (3.8.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Seguidamente veremos la justificación teórica de este procedimiento como selector de ancho de banda. 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

$$\frac{1}{[N/2]}\sum_{j=1}^{[N/2]}E\left[\left(\hat{f}(\omega_j;p,h)-f(\omega_j)\right)^2\right] \qquad (3.8.2)$$

Veamos que minimizar (3.8.2) es aproximadamente equivalente a minimizar

$$\frac{1}{[N/2]}\sum_{j=1}^{[N/2]}E\left[\left(\hat{f}^{(r_1)}(\omega_j;p,h)-\bar{I}^{(r_2)}(\omega_j)\right)^2\right] \qquad (3.8.3)$$

En efecto, como $f(\omega_j)$ no es conocida, podemos considerar una estimación, basada en el periodograma medio, el cual se obtiene, con datos diferentes a los utilizados para estimar $\hat{f}^{(r_1)}(\omega_j;p,h)$. Sea $\bar{I}^{(r_2)}(\omega_j)$ el periodograma medio referido. Se tiene entonces:

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

$$+E\left[\sum_{j=1}^{[N/2]}\left\{f(\omega_j)-\hat{f}^{(r_1)}(\omega_j;p,h)\right\}^2\right]$$

Dado que $E\left[\sum_{j=1}^{[N/2]}\left\{\bar{I}^{(r_2)}(\omega_j)-f(\omega_j)\right\}^2\right]$ es independiente del ancho de banda h, queda establecido que la minimización de (3.8.2) es aproximadamente equivalente a minimizar (3.8.3), lo cual justifica el método de validación cruzada propuesto para la selección del parámetro de alisamiento.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO IV.

ESTIMACIÓN BOOTSTRAP.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CAPÍTULO IV. ESTIMACIÓN BOOTSTRAP.

4.1 Introducción.

En el capítulo anterior, hemos obtenido una aproximación normal al estimador de núcleo polinómico local. Daremos ahora una aproximación bootstrap y compararemos ésta con la aproximación normal, utilizando los desarrollos de Edgeworth.

4.2 Algoritmo bootstrap.

Consideremos las r series temporales observadas en los mismos instantes $\{X_{lt} / l = 1,\ldots,r; t = 1,\ldots,N\}$.

El estimador polinómico lineal podemos expresarlo en los términos

$$\hat{f}_{r,N}(\omega;1,h) = \frac{1}{Nh} \sum_{j=-[N/2]}^{[N/2]} K_{(r,N)}^{(p)} \left(\frac{\omega_j - \omega}{h} \right) \bar{I}_{\cdot}(\omega_j) \qquad (4.2.1)$$

siendo $\bar{I}_{\cdot}(\omega_j) = \frac{1}{r} \sum_{l=1}^{r} I_l(\omega_j)$ y $K_{(r,N)}^{(p)}$ es el núcleo equivalente, correspondiente a un ajuste polinómico local de grado p (p=1, en este caso). El algoritmo bootstrap que proponemos es similar al dado por Härdle y Bowman (1988), en regresión no paramétrica, Franke y Härdle (1992) en estimación espectral, en el contexto de una única serie temporal y Hernández (1997) para la estimación del espectro poblacional, mediante series replicadas.

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\sqrt{r}\left\{\frac{\hat{f}(\omega;p,h)-E\left[\hat{f}(\omega;p,h)\right]}{f(\omega)}\right\} \qquad (4.2.2)$$

Paso 1 Las variables no observables W_{lj}, se aproximan por:

$$\tilde{W}_{lj}=\frac{I_l(\omega_j)}{\hat{f}(\omega_j;p,h^I)} \qquad l=1,\ldots,r;\ j=1,\ldots,[N/2]$$

siendo h^I un ancho de banda inicial. De acuerdo con la hipótesis $E\left[W_{lj}\right]=1$,consideramos $\hat{W}_{lj}=\frac{\tilde{W}_{lj}}{\overline{W}_{.j}}$, siendo $\overline{W}_{.j}=(1/r)\sum_{l=1}^{r}\tilde{W}_{lj}$ y se obtienen los vectores $\mathbf{W}_l=\left(\hat{W}_{l1},\ldots,\hat{W}_{l[N/2]}\right)$.

Paso 2 Se toman B muestras bootstrap de tamaño r $\mathbf{W}_l^*=\left(\hat{W}_{l1}^*,\ldots,\hat{W}_{l[N/2]}^*\right)$ de la distribución empírica en $R^{[N/2]}$ de $\{\mathbf{W}_1,\ldots,\mathbf{W}_r\}$. Se define el periodograma bootstrap de cada frecuencia mediante:

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

siendo g el ancho de banda de remuestreo.

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

$$\hat{f}^*\left(\omega_j;p,h,\ g\right)=\frac{1}{Nh}\sum_{j=-[N/2]}^{[N/2]}K^*\left(\frac{\omega-\omega_j}{h}\right)\bar{I}^*\left(\omega_j\right) \qquad (4.2.3)$$

Utilizando este algoritmo aproximamos el pivote dado por (4.2.2) a través del siguiente pivote bootstrap

$$\sqrt{r}\left\{\frac{\hat{f}^*(\omega;p,h,g)-\hat{f}_c(\omega;p,h,g)}{\hat{f}(\omega;p,g)}\right\} \qquad (4.2.4)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

siendo $\hat{f}_c(\omega; p,h,g) = E^*\left[\hat{f}^*(\omega; p,h,g)\right]$.

4.3 Desarrollos de Edgeworth.

En este epígrafe haremos una revisión de los fundamentos de los desarrollos de Edgeworth. Estos desarrollos permiten evaluar, bajo determinadas condiciones, el orden del error que se comete cuando se aproxima la distribución de ciertas variables, por la distribución normal o por otra distribución.

En el siguiente epígrafe, utilizamos estos desarrollos para evaluar, el orden de error cometido, cuando aproximamos la distribución del pivote por el correspondiente bootstrap, a través del cual determinamos intervalos de confianza para el espectro poblacional.

En líneas generales, la idea de los desarrollos de Edgeworth es como sigue: si $\hat{\theta}$ es el estimador de θ_0, basado en una muestra aleatoria de tamaño n , y si $n^{1/2}\left(\hat{\theta}-\theta_0\right)$ tiene distribución asintótica normal de media cero y varianza σ^2, en muchas situaciones la función de distribución de $n^{1/2}\left(\hat{\theta}-\theta_0\right)$ puede desarrollarse en series de potencias de $n^{1/2}$ de la forma:

$$P\left(n^{-1/2}\left(\hat{\theta}-\theta_0\right)/\sigma \le x\right) = \Phi(x) + n^{-1/2}p_1(x)\phi(x) + \ldots + n^{-j/2}p_j(x)\phi(x) + \ldots \qquad (4.3.1)$$

donde $\Phi(x)$ es la función de distribución de la N(0,1) y $\phi(x)$ su función de densidad. Las funciones p_j son polinomios cuyos coeficientes dependen directamente de los cumulantes de la distribución de $\hat{\theta}-\theta_0$. La expresión (4.3.1) recibe el nombre de desarrollo de Edgeworth. Tal expresión tiene un evidente significado. Muestra como la distribución de la cantidad pivotal puede aproximarse a la distribución normal, y que el error cometido en tal aproximación es del orden $n^{-1/2}$.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Desarrollos paralelos a los Edgeworth, son los desarrollos de Cornish-Fisher, los cuales pueden considerarse una inversión de (4.3.1). Sea

$$P(n^{1/2}(\hat{\theta}-\theta_0)/\sigma \le u_\alpha) = \alpha \qquad (4.3.2)$$

El cuantil u_α puede, bajo determinadas condiciones, desarrollarse en la forma:

$$u_\alpha = z_\alpha + n^{-1/2} p_{11}(z_\alpha) + n^{-2/2} p_{21}(z_\alpha) + \ldots + n^{-j/2} p_{j1}(z_\alpha) + \ldots \qquad (4.3.3)$$

siendo z_α el α-ésimo cuantil de la distribución N(0,1) ($\Phi(z_\alpha)=\alpha$) y las funciones $p_{j1}(x)$ polinomios que pueden obtenerse, a partir de los polinomios $p_j(x)$ correspondientes a los desarrollos de Edgeworth. Los desarrollos del tipo (4.3.3), llamados de Cornish-Fisher, permiten evaluar el orden de los errores cometidos en la construcción de los intervalos de confianza, cuando la distribución del pivote utilizado se aproxima por la distribución normal. Asimismo permiten evaluar el error correspondiente cuando la distribución del pivote se aproxima por la distribución bootstrap.

Un caso particularmente simple, es el correspondiente a la estimación de la media μ de una población, por la media muestral $\bar{x}$. Si la varianza poblacional es σ^2, el pivote
$S_n = n^{1/2}(\bar{x}-\mu)/\sigma$ tendrá una función de distribución que puede expresarse por

$$P(S_n \le x) = \Phi(x) + n^{-1/2} p_1(x)\phi(x) + \ldots + n^{-j/2} p_j(x)\phi(x) + o\left(n^{-j/2}\right) \qquad (4.3.4)$$

uniformemente en x.

En este caso, condiciones suficientes para la validez de (4.3.4) son:

$$E\left[|X|^{j+2}\right] < \infty \quad \text{y} \quad \limsup_{|t|\to\infty} |\chi(t)| < 1 \qquad (4.3.5)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde $\chi(t)$ es la función característica de la variable $Y=(X-\mu)/\sigma$. La última condición llamada condición de Cramer, es válida si la distribución de la variable X es no singular.

Si una variable X tiene distribución absolutamente continua y momento de orden tres finito ($E[X^3]<\infty$), (4.3.4) puede expresarse como

$$P(S_n \le x) = \Phi(x) + n^{-1/2} p_1(x)\phi(x) + o(n^{-1}) \qquad (4.3.6)$$

uniformemente en x, siendo $p_1(x) = -\frac{\kappa_3}{6}(x^2-1)$ y κ_3 el cumulante de orden 3 de Y ($\kappa_3 = E[(X-\mu)^3/\sigma^3]$). Dado que en casi todas las situaciones prácticas, σ^2 es desconocida, podemos estimarla por $\hat{\sigma}^2$ y considerar el pivote estudentizado

$$T_n = n^{1/2}(\bar{x}-\mu)/\hat{\sigma} \qquad (4.3.7)$$

para el cual pueden utilizarse análogos desarrollos.

Estableceremos a continuación situaciones más generales, en las que son válidos los desarrollos de Edgeworth.

Consideremos vectores aleatorios $X_1,\dots X_n$ en R^d, independientes e identicamente distribuidos y con media $\mu = E[X_1]$. Sea $A: R^d \to R$, suficientemente suave, satisfaciendo $A(\mu)=0$. Sea $\chi(t)$ la función característica d-dimensional de X_i. El siguiente teorema debido a Bhattacharya y Ghosh (1978), da condiciones generales para la validez de los desarrollos de Edgeworth.

Teorema 4.1 Supongamos que la función A tiene $j+2$ derivadas continuas en un entorno de $\mu = E[X]$, A(μ)=0, $E[\|X\|^{j+2}]<\infty$ y que la función característica satisface $\limsup_{|t|\to\infty}|\chi(t)|<1$. Sea además $\sigma^2 = Var(A(X_i))>0$. Entonces

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$P(n^{1/2} A(\overline{X})/\sigma \le x) = \Phi(x) + n^{-1/2} p_1(x)\phi(x) + \ldots + n^{-j/2} p_j(x)\phi(x) + o\left(n^{-j/2}\right) \qquad (4.3.8)$$

uniformemente en x, siendo $p_j(x)$ un polinomio de grado, a lo sumo $3j-1$, impar si j es par, y par si j es impar.

Para los cuantiles de las cantidades pivotales, el siguiente teorema da condiciones para la existencia de los desarrollos de Cornish-Fisher.

Teorema 4.2 Supongamos las mismas condiciones del teorema 4.1 para la distribución del vector aleatorio X y la función A. Sea el α-ésimo cuantil

$$w_\alpha = \inf\left\{x; P(n^{1/2} A(\overline{X}) \le x) \ge \alpha\right\}.$$

Entonces

$$w_\alpha = z_\alpha + n^{-1/2} p_{11}(z_\alpha) + n^{-2/2} p_{21}(z_\alpha) + \ldots + n^{-j/2} p_{j1}(z_\alpha) + o\left(n^{-j/2}\right)$$

uniformemente en $\varepsilon<\alpha<1-\varepsilon \ \forall \varepsilon>0$ y siendo p_{j1} polinomios de grado a lo sumo $j+1$.

4.4 El bootstrap y los desarrollos de Edgeworth.

En el epígrafe anterior se han dado condiciones para que la distribución de una cantidad pivotal T_n admita un desarrollo de Edgeworth. En aquellos casos en los que T_n tiene distribución normal asintótica, el error que se comete al hacer la aproximación normal es de orden $n^{-1/2}$. En el presente epígrafe, damos condiciones para aproximar la distribución bootstrap por la normal y a través de los desarrollos de Edgeworth, veremos que el error de tal aproximación es $O_P\left(n^{-1/2}\right)$. De ambos desarrollos de Edgeworth (el del pivote T y su análogo bootstrap T^*), deduciremos bajo adecuadas condiciones, que la diferencia entre la distribución de T y la distribución bootstrap es de orden $O_P\left(n^{-1}\right)$. Ello significa que la aproximación bootstrap a la distribución de T es asintóticamente mejor aproximación que la normal. Ello nos permitirá, en el caso de la estimación del

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

espectro poblacional, afirmar que el bootstrap es mejor procedimiento para la construcción de intervalos de confianza, que el basado en la distribución normal.

Consideremos entonces una muestra aleatoria $X=\{X_1,...,X_n\}$, obtenida de una población d-dimensional y sea $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$. Supongamos que deseamos estimar un parámetro $\theta_0 = g(\mu)$, siendo $g: R^d \rightarrow R$ una función suave de $\mu = E[X]$. El parámetro θ_0 se estima mediante $\hat{\theta} = g(\overline{X})$, el cual tiene varianza asintótica $n^{-1}\sigma^2 = n^{-1}h(\mu)^2$, donde h es una función suave conocida. Sea $\hat{\sigma}^2$ el estimador de σ^2.

Se desea entonces estimar la distribución del pivote

$$A(x) = \frac{g(x) - g(\mu)}{h(\mu)} \quad \text{o} \quad A(x) = \frac{g(x) - g(\mu)}{h(x)}$$

Sea la muestra bootstrap $X^* = \{X_1^*,..., X_n^*\}$ $\overline{X}^* = n^{-1}\sum_{i=1}^{n} X_i^*$ y sea $\hat{\theta}^* = g(\overline{X}^*)$ y $\hat{\sigma}^* = h(\overline{X}^*)$.

El pivote análogo a $A(x)$ será $\hat{A}(x) = \frac{g(x) - g(\overline{X})}{h(\overline{X})}$ o $\hat{A}(x) = \frac{g(x) - g(\overline{X})}{h(x)}$. El estimador bootstrap de $A(\overline{X})$ estará entonces dado por $\hat{A}(\overline{X}^*)$.

El siguiente teorema da condiciones bajo las cuales la distribución de $\hat{A}(\overline{X}^*)$, condicionado por la muestra, admite un desarrollo de Edgeworth, análogo al expresado en el teorema 4.1.

Teorema 4.3 Sea $\lambda>0$ y sea $l=l(\lambda)$, un entero positivo suficientemente grande (cuyo valor no se especifica). Supongamos que las funciones g y h tienen $\nu+3$ derivadas

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

acotadas en un entorno de μ, $E\left[\|X\|^1\right] < \infty$ y que la función característica de X satisface la condición de Cramer $\limsup_{|t|\to\infty} |\chi(t)| < 1$. Existe entonces una constante C, tal que

$$P\left(\sup_{-\infty<x<\infty} \left| P^*\left(n^{1/2}\hat{A}(\overline{X}^*) \le x\right) - \Phi(x) - \sum_{j=1}^{\nu} n^{-j/2}\hat{\pi}_j(x)\phi(x) \right| > Cn^{-(\nu+1)/2} \right) = O\left(n^{-\lambda}\right)$$

y

$$P\left(\max_{1\le j\le \nu} \sup_x \left(1+|x|\right)^{-(3j-1)} \left|\hat{\pi}_j(x)\right| > C \right) = O\left(n^{-\lambda}\right).$$

Los teoremas anteriores, nos permiten deducir, que bajo ciertas condiciones, la aproximación bootstrap a la distribución de un pivote, es mejor que la aproximación normal.

Así por ejemplo, supongamos que un pivote T admite un desarrollo de Edgeworth de la forma

$$P(T \le x) = \Phi(x) + n^{-1/2} p(F,x) + O\left(n^{-1}\right) \qquad (4.3.1)$$

uniformemente en x. Con la notación $p(F,x)$, queremos destacar que los coeficientes del polinomio, dependen de la distribución original de los datos F.

El estimador bootstrap de T admite entonces, el desarrollo análogo

$$P^*\left(T^* \le x\right) = \Phi(x) + n^{-1/2} p(F_n,x) + O_P\left(n^{-1}\right) \qquad (4.3.2)$$

siendo F_n la distribución empírica correspondiente a la muestra. Se tiene en general, que

$$p(F_n,x) - p(F,x) = O_P\left(n^{-1/2}\right)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Si a la expresión (4.3.2.) le restamos la (4.3.1), obtenemos

$$P^*\left(T^* \le x\right) - P\left(T \le x\right) = O_P\left(n^{-1}\right) \qquad (4.3.3)$$

Se tiene entonces que al aproximar la distribución de T, por la distribución bootstrap, el error cometido es de orden n^{-1}, según (4.3.3), mientras que el error cometido al aproximarla por la distribución normal, es de orden $n^{-1/2}$, según (4.3.1).

En los siguientes epígrafes, analizaremos la posibilidad de utilizar los desarrollos de Edgeworth, sobre la cantidad pivotal correspondiente a la estimación del espectro poblacional.

Si sobre la muestra de objetos seleccionados, pudiésemos determinar sus densidades espectrales exactas $f_1(\omega),\ldots,f_r(\omega)$, los resultados anteriores serían de aplicación directa, sobre esta muestra aleatoria, para cada frecuencia fija ω. Dado que estas variables no pueden observarse directamente, dispondremos al menos de estimaciones, $\hat{f}_1(\omega),\ldots,\hat{f}_r(\omega)$. Ello hará más complicado establecer una teoría rigurosa, sobre desarrollos de Edgeworth alternativos para el sistema de variables $\left\{\xi_{l,N}(\omega;h)\right\}_{l=1,\ldots,r(N)}$ definidas en (3.6.1), que nos permiten establecer condiciones precisas, bajo las cuales las aproximaciones bootstrap, sean mejores que la aproximación normal.

Abramovich y Singh (1985), proponen a través de los desarrollos de Edgeworth, correcciones de la asimetría de un pivote T, de tal forma que el pivote corregido T_1, aproxima mejor a la distribución normal que T. Propondremos en este sentido una extensión de estos resultados, a la estimación del espectro poblacional.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

4.5 Aplicación de los desarrollos de Edgeworth a la estimación del espectro poblacional.

Consideremos nuevamente la cantidad pivotal

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

supuesto que $\varphi(\omega)=Var(Z(\omega))>0;\forall\omega\in[-\pi,\pi]$.

Esta cantidad pivotal puede expresarse alternativamente por:

$$\frac{1}{\sqrt{r}}\sum_{l=1}^{r}\xi_{l,N}(\omega) \text{ donde } \xi_{l,N}(\omega)=\sqrt{r}\frac{\hat{f}(\omega;p,h)-E\left[\hat{f}(\omega;p,h)\right]}{f(\omega)\varphi(\omega)}.$$

En el capítulo anterior dimos condiciones bajo

$$\lim_{N\to\infty}P\left(\max_{1\le l\le r}\frac{\left|\xi_{l,N}(\omega)\right|}{\sqrt{r}}>\varepsilon\right)=0,\ \forall\varepsilon,N$$

Esto supone que $S_{N,r}(\omega)=\frac{1}{\sqrt{r}}\sum_{l=1}^{r}\xi_{l,N}(\omega)\xrightarrow{d}N(0,1)$.

De acuerdo con lo expuesto en el epígrafe (4.3.5), para cada ω, N, r, h, podemos expresar:

si $E\left[\xi_{l,N}^{3}(\omega)\right]<\infty$, $\limsup_{|t|\to\infty}\left|\chi_{\xi_{l,N}}(t)\right|<1$

$$P\left(S_{N,r}(\omega)\le x\right)=\Phi(x)+r^{-1/2}p_{N,1}(x)\phi(x)+O\left(r^{-1},N\right) \qquad (4.5.1)$$

donde

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$p_{N,1}(x) = -\frac{\kappa_{3,N}(\omega)}{6}(x^2 - 1) \qquad (4.5.2)$$

siendo $\kappa_{3,N}(\omega) = E\left[\xi_{l,N}^3(\omega)\right]$.

El desarrollo (4.5.1) aparentemente evalúa el error que se comete cuando se aproxima la distribución de S_N por $\Phi(x)$. En este caso, (4.5.2) da la forma exacta del polinomio $p_{N,1}(x)$, el cual obviamente depende de N, a través del cumulante $\kappa_{3,N}(\omega)$ de las variables $\xi_{l,N}(\omega)$, para cada frecuencia fija $\omega \in [-\pi, \pi]$. Ello significa que al crecer r, y por tanto variar N, el término $r^{-1/2} p_{N,1}(x)\phi(x)$ del desarrollo de Edgeworth podría no tender a cero, o hacerlo a una tasa más lenta que $r^{-1/2}$.

Si $\kappa_{3,N}(\omega) \leq C \ \forall N$, entonces puede garantizarse que el referido término tiende a cero al menos con la tasa $r^{-1/2}$.

La expresión $O_N(r^{-1})$ no puede interpretarse en el sentido habitual de $\lim_{r\to\infty} O_N(r^{-1}) < \infty$ salvo que $\left|O_N(r^{-1})\right| < D, \forall N$.

De lo anterior se concluye que si:

$\left|\kappa_{3,N}(\omega)\right| < C$ y $\left|O_N(r^{-1})\right| < D, \forall N, \omega$ entonces,

$$\sup_x r^{1/2}\left\{P(S_N \leq x) - \Phi(x)\right\} \xrightarrow[r\to\infty]{} 0.$$

4.6 El bootstrap y los desarrollos de Edgeworth en la estimación del espectro poblacional.

En epígrafes anteriores hemos propuesto aproximar la distribución de

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

$$\sqrt{r}\frac{\hat{f}(\omega;p,h)-E\left[\hat{f}(\omega;p,h)\right]}{f(\omega)\varphi(\omega)}$$

por

$$\sqrt{r}\frac{\hat{f}^*(\omega;p,h)-f_c(\omega;p,h,g)}{\hat{f}(\omega;p,h)\hat{\varphi}(\omega)}$$

donde $f_c(\omega;p,h,g)=E^*\left[\hat{f}^*(\omega;p,h,g)\right]$.

Consideremos entonces la cantidad pivotal bootstrap

$$S^*_{N,r}(\omega)=\sqrt{r}\frac{\hat{f}^*(\omega;p,h)-f_c(\omega;p,h,g)}{\hat{f}(\omega;p,h)\hat{\varphi}(\omega)}$$

que puede expresarse por

$$S^*_{N,r}(\omega)=\frac{1}{\sqrt{r}}\sum_{l=1}^{r}\frac{\hat{f}^*_l(\omega;p,h)-f_c(\omega;p,h,g)}{\hat{f}(\omega;p,h)\hat{\varphi}(\omega)}.$$

Parece razonable, que podamos encontrar condiciones bajo las cuales

$$P\left(S^*_{N,r}(\omega)\le x/\left\{I_l(\omega_j)\right\}_{lj}\right)=\Phi(x)+r^{-1/2}p^*_{N,1}(x)\phi(x)+O_P\left(r^{-1}\right)$$

donde $p^*_{N,1}(x)=-\frac{\kappa^*_{3,N}(\omega)}{6}(x^2-1)$ y $\kappa^*_{3,N}=E^*\left[\left\{\frac{\hat{f}^*_l(\omega;p,h)-f_c(\omega;p,h,g)}{\hat{f}(\omega;p,h)\hat{\varphi}(\omega)}\right\}^3\right]$.

Si además, $p^*_{N,1}(x)-p_{N,1}(x)=O_P\left(r^{-1/2}\right)$. Entendiendo que

$$\lim_{\lambda\to\infty}\limsup_{r\to\infty}P\left(r^{1/2}\left|p^*_{N,1}(x)-p_{N,1}(x)\right|>\lambda\right)=0$$

En tal caso, $P\left(S^*_{N,r}(\omega)\le x/\left\{I_l(\omega_j)\right\}_{lj}\right)-P\left(S_{N,r}(\omega)\le x\right)=O_P\left(r^{-1}\right)$. Lo cual significa que la aproximación bootstrap, mejora la aproximación normal.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

4.7 Corrección del sesgo.

Abramovitch y Singh (1985), introducen un procedimiento de modificación de pivotes estadísticos en sucesivas etapas, para mejorar la aproximación normal. El siguiente teorema prueba, bajo determinadas condiciones, que una cierta modificación del pivote estadístico, tiene distribución más próxima a la distribución estándar.

Teorema 4.4 Sea un estadístico T, basado en una muestra de variables aleatorias, independientes e identicamente distribuidas, con distribución f de tal forma que admite el desarrollo de Edgeworth

$$P(T \leq x) = \Phi(x) + n^{-1/2} p(x)\phi(x) + o(n^{-1/2})$$

uniformemente en x. Sea $\hat{p}_n$ un estimador de $p(T)$ que satisface la condición

$$P\left(\left|\hat{p}_n - p(T)\right| > \varepsilon\right) = o\left(n^{-1/2}\right); \quad \forall \varepsilon$$

Entonces, T_1 definido por $T_1 = T + n^{-1/2}\hat{p}_n$ tiene el siguiente desarrollo de Edgeworth

$$P(\mathrm{T}_1 \leq \mathrm{x}) = \Phi(x) + o(n^{-1/2})$$

uniformemente en x.

En realidad, si T_1 admite un desarrollo de Edgeworth, este tiene la forma:

$$P(\mathrm{T}_1 \leq \mathrm{x}) = \Phi(x) + n^{-1} q(x)\phi(x) + o(n^{-1})$$

uniformemente en x. En este caso, el teorema anterior sugiere realizar la corrección en segunda fase

$$T_2 = T_1 + n^{-1}\hat{q}_n$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde $\hat{q}_n$ es un estimador $q(T_1)$ que satisface

$$P\left(\left|\hat{q}_n - q(T_1)\right| > \varepsilon\right) = o\left(n^{-1}\right); \forall \varepsilon > 0$$

De esta forma, T_2 no tendría un término de orden n^{-1}, en el supuesto que a su vez admitiera un desarrollo de Edgeworth.

Teóricamente, el proceso descrito puede realizarse en k etapas. No obstante, con muestras pequeñas, probablemente no merezca la pena realizar una corrección de segunda etapa, pues la correcciones de alto orden suponen la estimación de momentos de alto orden, los cuales se estiman con poca precisión.

Cabe preguntarse ahora si tiene alguna ventaja, realizar el bootstrap sobre T_1, en lugar de hacer uso de la aproximación a la distribución normal.

4.8 Corrección del sesgo para el pivote del espectro poblacional.

A lo largo de este epígrafe haremos una serie de consideraciones relacionadas con la posible extensión del teorema de Abramovitch y Singh (1985), para la corrección del pivote correspondiente al espectro poblacional. Sea pues nuevamente

$$T_{r,N}(\omega) = \sqrt{r}\frac{\hat{f}(\omega;p,h) - E\left[\hat{f}(\omega;p,h)\right]}{f(\omega)\varphi(\omega)} = \frac{1}{\sqrt{r}}\sum_{l=1}^{r}\xi_{l,N}(\omega)$$

Supongamos que se dan las condiciones del teorema 4.4 y que $\frac{1}{\sqrt{r}}\sum_{l=1}^{r}\xi_{l,N_r}(\omega) \xrightarrow[r\to\infty]{} N(0,1)$; donde $N_r \to \infty$. Supongamos que además la distribución de $T_{r,N}(\omega)$ admite el desarrollo

$$P\left(T_{r,N}(\omega) \le x\right) = \Phi(x) + r^{-1/2}p_N(x)\phi(x) + o(r^{-1/2};N) \qquad (4.8.1)$$

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

siendo

$$p_N(x) = -\frac{\kappa_{3,N}}{6}(x^2 - 1)$$

y

$$\kappa_{3,N}(\omega) = E\left[\xi_{l,N}^3(\omega)\right].$$

Supongamos que se dan las condiciones

$$\kappa_{3,N}(\omega) = E\left[\xi_{l,N}^3(\omega)\right] \leq C; \forall N$$

y

$$\limsup_{r\to\infty} {}_x\, r^{1/2} o\left(r^{1/2}; N\right) = 0.$$

En tal caso, el desarrollo anterior informa acerca de la velocidad de convergencia, de la distribución de la cantida pivotal $T_{r,N}(\omega)$, a la distribución normal estándar.

Bajo las condiciones mencionadas, propondremos una corrección del pivote $T_{r,N}(\omega)$, en el sentido propuesto por Abramovitch y Singh (1985), para pivotes basados en muestras aleatorias de una distribución de probabilidad.

El fundamento de nuestra aproximación consiste, en estimar una densidad espectral individual $f_l(\omega)$ mediante alisamientos del correspondiente periodograma individual $I_l(\omega)$. El ancho de banda h se selecciona utilizando el método del capítulo anterior, por tanto este es común para todas las densidades espectrales individuales.

La elección de un ancho de banda común, no solo simplifica notablemente el problema computacional, sino que además garantiza, que para un valor fijo de N y ω, $\xi_{1,N}(\omega),\ldots,\xi_{r,N}(\omega)$ sean variables independientes e identicamente distribuidas.

Dado que las variables $f_1(\omega),\ldots f_r(\omega)$ son desconocidas, consideramos sus estimaciones $\hat{f}_1(\omega;h),\ldots,\hat{f}_r(\omega;h)$, bien por estimaciones de núcleos, o por estimaciones polinómicas locales.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Consideraremos entonces los siguientes estadísticos

$$\hat{\mu}_1 = \frac{1}{r}\sum_{l=1}^{r} \hat{f}_l(\omega;h) = \hat{f}(\omega;h)$$

$$\hat{\mu}_2 = \frac{1}{r-1}\sum_{l=1}^{r} \left\{ \hat{f}_l(\omega;h) - \hat{\mu}_1 \right\}^2$$

$$\hat{\mu}_3 = \frac{1}{r}\sum_{l=1}^{r} \left\{ \hat{f}_l(\omega;h) - \hat{\mu}_2 \right\}^3$$

De esta forma, consideramos como estimador de $P_N\left(\mathrm{T}_{r,N}(\omega)\right)$ en (4.7.1)

$$\hat{p}_{N,r}(\omega) = -\frac{\hat{\mu}_3}{6\hat{\mu}_2^{1/2}}\left(T_{r,N}^2(\omega) - 1\right)$$

Y proponemos como estimador modificado

$$T_{r,N}^{(1)}(\omega) = T_{r,N}(\omega) + r^{-1/2}\hat{p}_{N,r}(\omega).$$

El problema que se plantea ahora es determinar condiciones que nos permitan afirmar que:

$$P\left(T_{r,N}^{(1)}(\omega) \le x\right) = \Phi(x) + o\left(r^{-1/2}\right)$$

uniformemente en x. Esto supondría que la distribución de probabilidad del pivote modificado $T_{r,N}^{(1)}(\omega)$, aproxima mejor la distribución normal, que el pivote $T_{r,N}(\omega)$.

4.9 Simulaciones.

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

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

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

donde el vector de coeficientes (ϕ_0, ϕ_1) es aleatorio y sigue una distribución normal bivariante con vector de media [3,3] y con matriz de varianza-covarianza $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Las $\{\varepsilon_t\}$ son variables aleatorias independientes, normalmente distribuidas con media cero y varianza 1, y además independientes de (ϕ_0, ϕ_1).

En todas las gráficas representamos para cada frecuencia el espectro poblacional real. en color rojo. La figura 1 representa las dos bandas de confianza al 95% obtenidas unas en base a las propiedades asintóticas del periodograma medio en color azul y otras al utilizar el algoritmo bootstrap descrito en el epígrafe 2.7 en color verde. Las estimaciones se obtienen en base a 792 series de 60 observaciones por objeto. La gráfica presenta un aspecto adecuadamente suave, lo cual es consecuencia de las propiedades teóricas del periodograma medio.

El periodograma medio proporciona adecuadas estimaciones 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*, en los alisamientos mediante núcleo. 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.

Para hacer la estimación del espectro poblacional, por alisamientos mediante núcleos del periodograma medio y mediante el estimador polinómico local hemos utilizado el núcleo de Bartlett-Priestley, que viene dado por:

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

La figura 2 representa el intervalo de confianza bootstrap al 95% del espectro poblacional, en color verde, obtenidas por estimación mediante núcleo, utilizando el algoritmo bootstrap descrito en el epígrafe 2.10.

Las figuras 3 y 4 representan las bandas de confianza al 95%, obtenidas por estimación polinómica local (grado 1), mediante el algoritmo bootstrap descrito en el epígrafe 4.2, en color verde, y a través de las propiedades asintóticas del estimador para $r=100$ y $N=100$ en la figura 3 y para $r=100$ y $N=200$ en la figura 4, en color azul.

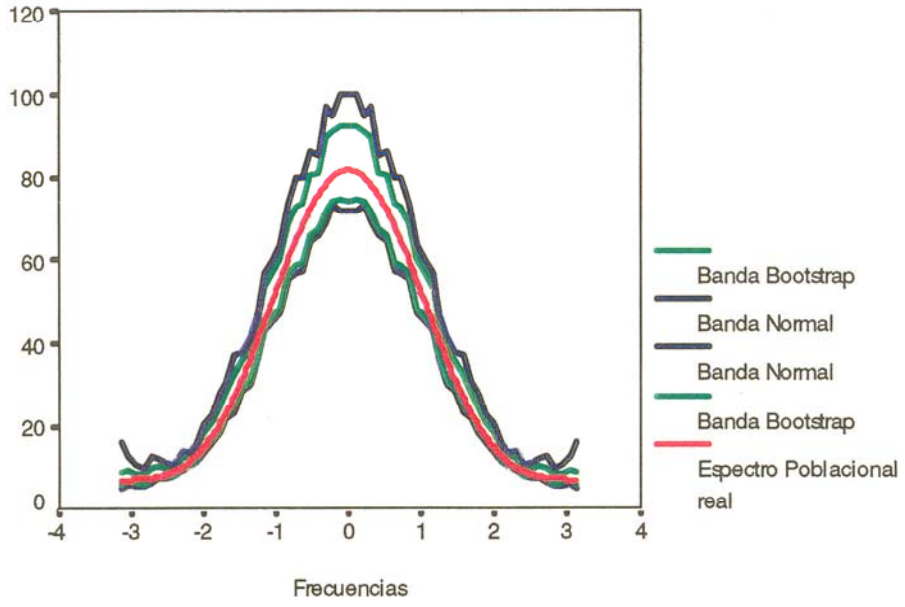
Por último, la figura 5 representa para $r=100$ y $N=100$ la comparación de las bandas bootstrap obtenidas por el estimador de núcleo en color verde, y por el estimador polinómico local en color azul, observándose la ventaja del estimador polinómico local frente a los de núcleo, principalmente en el pico y en la frontera.

Puede observarse también en las figuras 1, 3 y 4, como las bandas de confianza bootstrap aproximan mejor a la función de densidad espectral poblacional verdadera que las obtenidas por la distribución normal. Lo cual corrobora los resultados obtenidos en el capítulo IV con los desarrollos de Edgeworth.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Figura 1

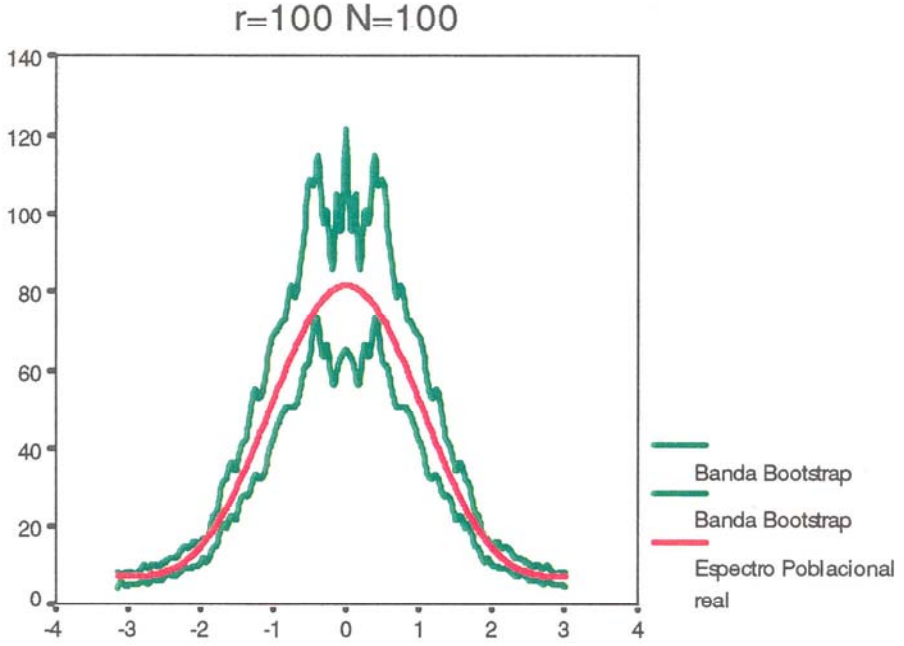
r=792 N=60



Estimación de la función de densidad espectral poblacional mediante el periodograma medio

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Figura 2

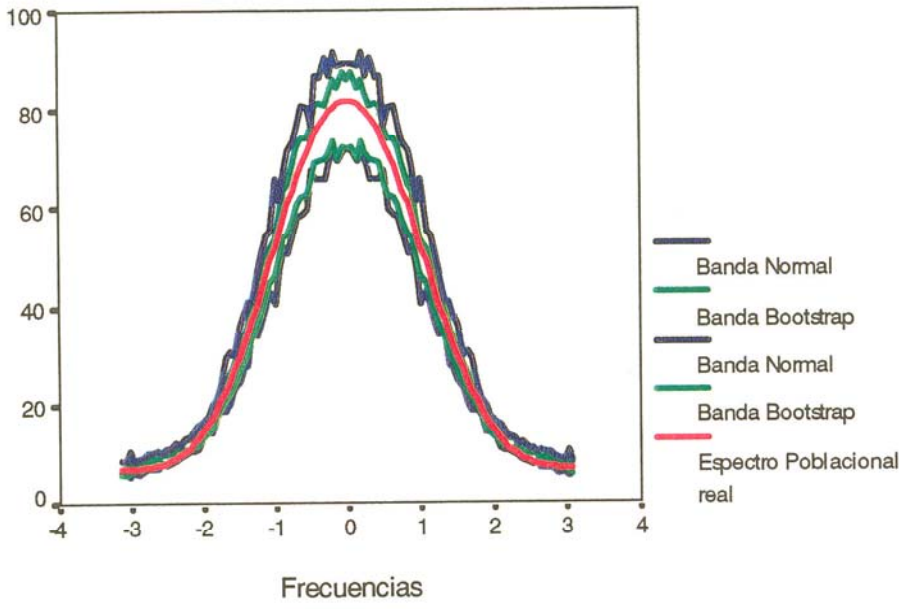


Estimación de la función de densidad espectral poblacional a través de alisamientos mediante núcleo

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Figura 3

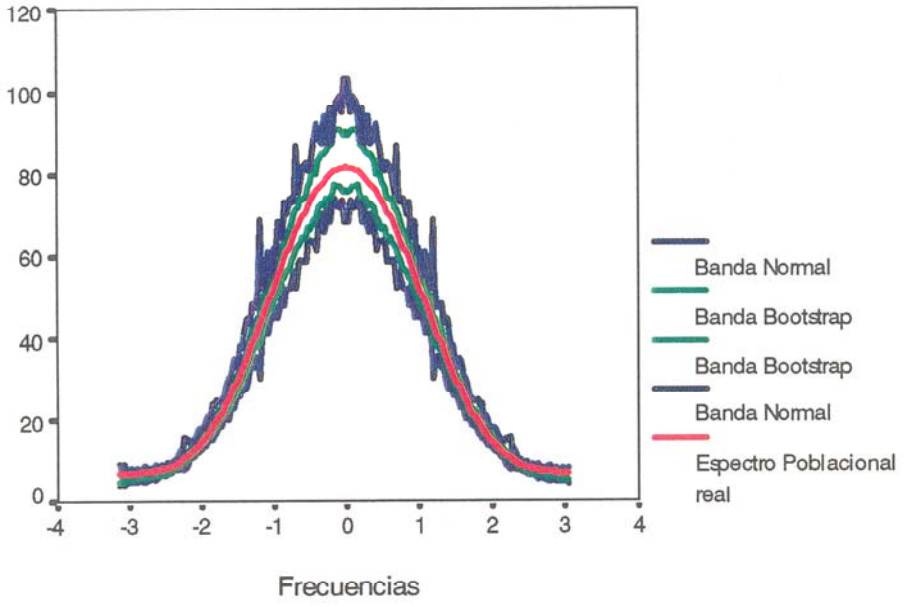
r=100 N=100



© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

FIGURA 4

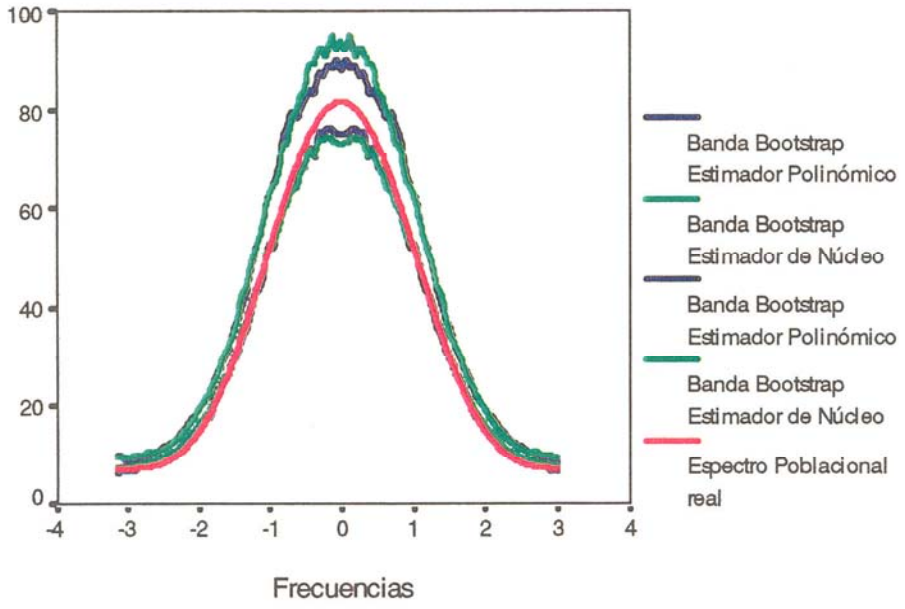
r=100 N=200



© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

FIGURA 5

r=100 N=100



Estimación de la función de densidad espectral poblacional mediante núcleo y a través del estimador polinómico local

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CONCLUSIONES.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

CONCLUSIONES.

♠ Hemos obtenido el sesgo, la varianza y la distribución asintótica del estimador polinómico local de la función de densidad espectral poblacional, partiendo del modelo multiplicativo de efectos aleatorios.

♠ Destacamos que la varianza se descompone en una componente *intraindividual* y otra componente que corresponde a los efectos aleatorios del modelo. De este resultado obtenemos, bajo condiciones precisas, que el orden del parámetro de alisamiento óptimo es $(Nr)^{-1/5}$, para el caso de ajuste lineal local.

♠ Resaltamos las ventajas de los estimadores polinómicos locales frente a los de núcleos: por la carencia de efectos frontera, ya que el sesgo en la frontera en los estimadores polinómicos locales, automáticamente es del mismo orden que en el interior, y en los puntos interiores, pues en regiones suaves un ajuste constante o lineal es recomendable, mientras que en regiones con picos o valles, un ajuste cuadrático o cúbico es preferible. Estas ventajas se hacen patentes si se observan los gráficos que hemos obtenido mediante la simulación de un MA(2).

♠ Señalamos como desventaja de estos estimadores polinómicos locales que al tenerse que contemplar p parámetros más en los cálculos, que son los coeficientes del polinomio de grado p, lleva consigo un mayor coste computacional.

♠ Hemos obtenido expresiones para los estimadores polinómicos locales, similares a las de los estimadores de núcleo, mediante el uso de los "núcleos equivalentes". Esto hace que consideremos los estimadores de núcleo, como un caso particular de los estimadores polinómicos locales.

♠ Hemos estimado el ancho de banda óptimo por un método de validación cruzada, dividiendo el número de individuos de la muestra aleatoriamente en dos partes.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

♠ Por último aplicamos los desarrollos de Edgeworth, y observamos que la aproximación bootstrap del pivote $\sqrt{r}\dfrac{\hat{f}(\omega;p,h)-E\left[\hat{f}(\omega;p,h)\right]}{\sqrt{Var\left[Z(\omega)\right]}}$ es del orden r^{-1}, en lugar del orden $r^{-1/2}$ que se consigue haciendo la aproximación mediante la normal.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

APÉNDICE.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

APÉNDICE.

Se presentan a continuación los listados de tres programas realizados en Pascal y dos programas en el paquete Statistical Analysis System (SAS 6.06), ejecutados en un ordenador Pentium 133 Mhz, 8 Mb, bajo Windows 95.

Listado 1. Programa en Pascal que genera *r* realizaciones de un proceso MA(2) con coeficientes aleatorios.

```
uses crt;

var i,t,n,r,j,nm:integer;
   mb1,mb2,v1,v2,c12,u,v,z1,z2,a,b,c,
   e0,e1,e2,std,b1,b2,k1,k2,k3,k4,freq,hw,hw2:real;
   y:array[1..100,1..100] of real;
   f:text;
   ka:char;

function normal(s:real):real;
var u,v,z:real;
 begin
  u:=random;
  v:=random;
  z:=sqrt(-2*ln(u))*cos(2*pi*v);
  normal:=s*z;
 end;

 {Programa Principal}
 begin
  TextBackground(lightblue);
  clrscr;
  gotoxy(3,3);write('E[B1] = ');read(mb1);
  gotoxy(3,4);write('E[B2] = ');read(mb2);
  repeat
   gotoxy(3,5);write('Varianza B1 = ');read(v1);
   gotoxy(3,6);write('Varianza B2 = ');read(v2);
   gotoxy(3,7);write('COV(B1,B2) = ');read(c12);
  until sqr(c12)<=v1*v2;
  gotoxy(3,8);write('Varianza del ruido blanco = ');read(std);
  repeat
   gotoxy(3,9);write('N£mero de sujetos = ');read(r);
  until r<=100;
  repeat
   gotoxy(3,10);write('N£mero de observaciones = ');read(n);
  until n<=100;
  c:=sqrt(v2);
  if c=0 then b:=0 else b:=c12/c;
  a:=sqrt(v1-b*b);
  randomize;

  assign(f,'\sas\coef.dat');rewrite(f);
  {Simulaci¢n de las r series}
  for i:=1 to r do
   begin
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
 gotoxy(75,2);
 write(i);
 u:=random;
 v:=random;
 z1:=sqrt(-2*ln(u))*cos(2*pi*v);
 z2:=sqrt(-2*ln(u))*sin(2*pi*v);
 b1:=mb1+a*z1+b*z2;
 b2:=mb2+c*z2;
 writeln(f,b1:10:5,b2:10:5);
 e1:=normal(std);
 e2:=normal(std);
 for t:=1 to n do
  begin
   e0:=normal(std);
   y[i,t]:=e0+b1*e1+b2*e2;
   e2:=e1;
   e1:=e0;
  end;
 end;
close(f);

{Escritura de las r series}
assign(f,'\sas\mm2.dat');
rewrite(f);
for t:=1 to n do
 begin
  for i:=1 to r do
   begin
    write(f,y[i,t]:9:3);
    if 9*i mod 72=0 then writeln(f);
    if i=r then writeln(f);
   end;
 end;
close(f);

{C lculo del espectro poblacional}
nm:=round(n/2);
k1:=sqr(std)/(2*Pi);
k2:=1+v1+sqr(mb1)+v2+sqr(mb2);
k3:=2*(mb1+c12+mb1*mb2);
k4:=2*mb2;
assign(f,'\sas\espobla.dat');
rewrite(f);
for j:=0 to nm do
 begin
  freq:=(2*pi*j)/n;
  hw:=k1*(k2+k3*cos(freq)+k4*cos(2*freq));
  writeln(f,freq:10:5,hw:10:4);
 end;
close(f);
assign(f,'\sas\derivada.dat');
rewrite(f);
for j:=0 to nm do
 begin
  freq:=(2*pi*j)/n;
  hw2:=-k1*k3*cos(freq)-4*k1*k4*cos(2*freq);
  writeln(f,hw2:10:4);
 end;
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
 close(f);

 gotoxy(40,23);
 write('Pulse cualquier tecla para continuar');
 ka:=readkey;
 clrscr;
end.
```

Listado 2. Programa en SAS que obtiene los periodogramas individuales y medio de las series generadas por el programa del listado 1.

```
goptions device=egal;
data datos;
infile 'mm2.dat';
input x1-x100;

/* Calculo de los periodogramas de los r individuos */
proc spectra s out=espect;
  var x1-x99;
run;
data b;
set espect;
file 'periodo.dat';
put (s_01-s_99)(12.5);



/* Calculo de los Z(w) */
data c;set espect;
array in{99} s_01-s_99;
r=99;
Pi=3.14159265;
se=datetime();
se=int(se);
file 'pm.dat';
/* Calculamos el periodograma medio para cada una de las frecuencias */
pm=0;
do i=1 to r;
  pm=pm+in{i};
 end;

pm=pm/r;
put pm 12.5;
```

Listado 3. Programa en Pascal que estima las variables W_{ij} del modelo multiplicativo dado por (3.2.1)

```
uses crt;
const dimax=200;
type vector=array[-dimax..dimax] of real;

var i,j,k,n,t,inf,sup,grado,r:integer;
    paso,h1,h2,h3,h4,h5,h6,m1,m2,m3,m4,m5,m6,mise,c123,c124,a,bb,c,hopt,dis:real;
    ka:char;
    f,g:text;
    x,pm,fwg,w,m:vector;
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
function bp(x:real):real;
 begin
  if abs(x)<=1 then bp:=3*(1-x*x)/4 else bp:=0;
 end;

procedure mise1(t:integer;h:real;x,y:vector;var mse:real;var b:vector);
var i,j:integer;
   a,c,d,e,g,k:real;
 begin
  mse:=0;
  for i:=-t to t do
   begin
    a:=0;c:=0;d:=0;e:=0;g:=0;
    for j:=-t to t do
     begin
      if j<>i then
       begin
        k:=bp((x[i]-x[j])/h);
        a:=a+k*y[j];
        g:=g+k*(x[i]-x[j]);
        c:=c+k*(x[i]-x[j])*y[j];
        d:=d+k*(x[i]-x[j])*(x[i]-x[j]);
        e:=e+k;
       end;
     end;
    if (e*d-g*g)<>0 then b[i]:=(a*d-c*g)/(e*d-g*g);
    mse:=mse+sqr(y[i]-b[i]);
   end;
 end;

procedure mise2(t:integer;h:real;x,y:vector;var mse:real;var b:vector);
var i,j:integer;
   a,c,k,s0,s1,s2,s3,s4,t0,t1,t2:real;
 begin
  mse:=0;
  for i:=-t to t do
   begin
    s0:=0;s1:=0;s2:=0;s3:=0;s4:=0;t0:=0;t1:=0;t2:=0;
    for j:=-t to t do
     begin
      if j<>i then
       begin
        if h<>0 then k:=bp((x[i]-x[j])/h);
        s0:=s0+k;
        s1:=(x[i]-x[j])*k+s1;
        s2:=sqr(x[i]-x[j])*k+s2;
        s3:=sqr(x[i]-x[j])*(x[i]-x[j])*k+s3;
        s4:=sqr(x[i]-x[j])*sqr(x[i]-x[j])*k+s4;
        t0:=k*y[j]+t0;
        t1:=(x[i]-x[j])*k*y[j]+t1;
        t2:=sqr(x[i]-x[j])*k*y[j]+t2;
       end;
     end;
    c:=s0*s2*s4+2*s1*s2*s3-sqr(s2)*s2-sqr(s1)*s4-s0*sqr(s3);
    a:=t0*s2*s4+t1*s2*s3+t2*s1*s3-t2*sqr(s2)-t0*sqr(s3)-t1*s1*s4;
   if c<>0 then b[i]:=a/c;
   mse:=mse+sqr(y[i]-b[i]);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
  end;
 end;


 {PROGRAMA PRINCIPAL}
begin
 TextBackground(lightblue);
 clrscr;
 gotoxy(15,12);
 write('*** GENERA W(I,J) UTILIZANDO POLINOMIOS LOCALES ***');
 gotoxy(40,23);
 write('Pulse cualquier tecla para continuar');
 ka:=readkey;
 clrscr;
 gotoxy(3,9);write('N£mero de individuos= ');read(r);
 readln;
 assign(f,'\sas\pm.dat');
 reset(f);
 t:=-1;
 while not eoln(f) do
  begin
   t:=t+1;
   gotoxy(75,2);
   write(t);
   readln(f,pm[t]);
   pm[-t]:=pm[t];
  end;
 close(f);
 n:=2*t;
 for j:=-t to t do
 x[j]:=2*PI*j/n;
 gotoxy(3,11);
 write('Grado de los polinomios locales ==> ');
 read(grado);
 h1:=2*PI/n;
 h1:=2*h1;
 h4:=(x[t]-x[0])/2;
 h2:=h1+(h4-h1)/3;
 h3:=h1+2*(h4-h1)/3;

  {CALCULO DEL BANDWIDTH OPTIMO}
if grado=2 then
  begin
   mise2(t,h1,x,pm,m1,m);
   mise2(t,h2,x,pm,m2,m);
   mise2(t,h3,x,pm,m3,m);
   mise2(t,h4,x,pm,m4,m);
  end
 else
  begin
   mise1(t,h1,x,pm,m1,m);
   mise1(t,h2,x,pm,m2,m);
   mise1(t,h3,x,pm,m3,m);
   mise1(t,h4,x,pm,m4,m);
  end;
 c123:=((m1-m2)/(h1-h2)-(m1-m3)/(h1-h3))/(h2-h3);
 c124:=((m1-m2)/(h1-h2)-(m1-m4)/(h1-h4))/(h2-h4);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
a:=(c123-c124)/(h3-h4);
bb:=c123-(h1+h2+h3)*a;
c:=((m1-m2)/(h1-h2))-(h1*h1+h1*h2+h2*h2)*a-(h1+h2)*bb;
ka:=readkey;
if m1<m2 then hopt:=h1 else
 begin
  if a=0 then hopt:=-c/(2*bb) else
   begin
    dis:=bb*bb-3*a*c;
    if dis<0 then
     begin
      if a<0 then hopt:=h4;
      if a>0 then hopt:=h1;
     end
    else
     begin
      h5:=(-bb+sqrt(dis))/(3*a);
      h6:=(-bb-sqrt(dis))/(3*a);
      if grado=2 then
       begin
        mise2(t,h5,x,pm,m5,m);
        mise2(t,h6,x,pm,m6,m);
       end
      else
       begin
        mise1(t,h5,x,pm,m5,m);
        mise1(t,h6,x,pm,m6,m);
       end;
      if m5<=m6 then hopt:=h5 else hopt:=h6;
       end;
    end;
 end;
clrscr;
gotoxy(3,9);
write('EL ANCHO DE BANDA OPTIMO ES=',hopt);
gotoxy(40,23);
write('Pulse cualquier tecla para continuar');
ka:=readkey;
clrscr;
if grado=2 then mise2(t,hopt,x,pm,mise,m)
 else mise1(t,hopt,x,pm,mise,m);
for j:=-t to t do
 fwg[j]:=m[j];
assign(f,'\sas\periodo.dat');
assign(g,'\sas\w.dat');
rewrite(g);
reset(f);
for j:=0 to t do w[j]:=0;
for j:=0 to t do
 begin
  for i:=1 to r do
   begin
    gotoxy(75,2);
    write(j);
    read(f,paso);
    paso:=paso/fwg[j];
    w[j]:=w[j]+paso;
    write(g,paso:10:5);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
   end;
  writeln(g);
 end;
close(f);
close(g);
for j:=0 to t do w[j]:=w[j]/r;

assign(f,'\sas\w.dat');
assign(g,'\sas\ww.dat');
rewrite(g);
reset(f);
for j:=0 to t do
 begin
  for i:=1 to r do
   begin
    read(f,paso);
    paso:=paso/w[j];
    writeln(g,paso:10:5);
   end;
 end;
close(f);
close(g);
end.
```

Listado 4. Programa en SAS que genera las *b* muestras bootstrap.

```
goptions device=egal;

data datos;
infile 'c:\sas\ww.dat';
input w1-w99;

proc transpose data=datos out=salida prefix=fq;
 run;

proc univariate data=salida noprint;
 var fq1-fq51;
  output out=varf var=varf1-varf51;
 run;

data varf;
set varf;
file 'varf.dat';
put varf1-varf51;

/* Calculo de los w bootstrap*/

data c;
set datos;
array w{99} w1-w99;
array wbm{100} wbm1-wbm100;
r=99;b=100;
Pi=3.14159265;
se=datetime();
se=int(se);
file 'c:\sas\boot.dat';
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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

   do k=1 to b;
    wbm{k}=0;
    do i=1 to r;
    l=ranuni(se);
    l=int((r-1)*l)+1;
    wbm{k}=wbm{k}+w{l};
    end;
    wbm{k}=wbm{k}/r;
    put wbm{k} 10.5;
   end;
```

Listado 5. Programa en Pascal donde se implementa el algoritmo bootstrap dado en el epígrafe 4.2.

```
uses crt;
const dimax=120;
type vector=array[-dimax..dimax] of real;

var i,j,k,n,b,t,inf,sup,grado,r:longint;
   paso1,paso2,paso,h1,h2,h3,h4,h5,h6,m1,m2,m3,m4,m5,m6,
   mise,c123,c124,a,bb,c,hopt,dis:real;
   ka:char;
   f,g:text;
   x,yb,pm,fwg,fwh,wb,m,hw,f2,alfa1,alfa2,varf,normal1,normal2:vector;

function bp(x:real):real;
 begin
  if abs(x)<=1 then bp:=3*(1-x*x)/4 else bp:=0;
 end;

procedure mise1(t:integer;h:real;x,y:vector;var mse:real;var b:vector);
var i,j:integer;
   a,c,d,e,g,k:real;
 begin
  mse:=0;
  for i:=-t to t do
   begin
    a:=0;c:=0;d:=0;e:=0;g:=0;
    for j:=-t to t do
     begin
      if j<>i then
       begin
        k:=bp((x[i]-x[j])/h);
        a:=a+k*y[j];
        g:=g+k*(x[i]-x[j]);
        c:=c+k*(x[i]-x[j])*y[j];
        d:=d+k*(x[i]-x[j])*(x[i]-x[j]);
        e:=e+k;
       end;
     end;
    if (e*d-g*g)<>0 then b[i]:=(a*d-c*g)/(e*d-g*g);
    mse:=mse+sqr(y[i]-b[i]);
   end;
 end;
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
procedure mise2(t:integer;h:real;x,y:vector;var mse:real;var b:vector);
var i,j:integer;
   a,c,k,s0,s1,s2,s3,s4,t0,t1,t2:real;
 begin
  mse:=0;
  for i:=-t to t do
   begin
    s0:=0;s1:=0;s2:=0;s3:=0;s4:=0;t0:=0;t1:=0;t2:=0;
    for j:=-t to t do
     begin
      if j<>i then
       begin
        if h<>0 then k:=bp((x[i]-x[j])/h);
        s0:=s0+k;
        s1:=(x[i]-x[j])*k+s1;
        s2:=sqr(x[i]-x[j])*k+s2;
        s3:=sqr(x[i]-x[j])*(x[i]-x[j])*k+s3;
        s4:=sqr(x[i]-x[j])*sqr(x[i]-x[j])*k+s4;
        t0:=k*y[j]+t0;
        t1:=(x[i]-x[j])*k*y[j]+t1;
        t2:=sqr(x[i]-x[j])*k*y[j]+t2;
        end;
     end;
     c:=s0*s2*s4+2*s1*s2*s3-sqr(s2)*s2-sqr(s1)*s4-s0*sqr(s3);
     a:=t0*s2*s4+t1*s2*s3+t2*s1*s3-t2*sqr(s2)-t0*sqr(s3)-t1*s1*s4;
    if c<>0 then b[i]:=a/c;
    mse:=mse+sqr(y[i]-b[i]);
   end;
end;



procedure trans(t,b:integer);
var n,i,j,k:integer;
   x:real;
   y:array[1..1000] of real;
   f,g:text;

begin
 assign(f,'\sas\mboot.dat');
 assign(g,'\sas\tmboot.dat');
 rewrite(g);
 for i:=1 to t+1 do
  begin
   reset(f);
   for j:=1 to b do
    begin
     for k:=1 to i do
      read(f,x);
     y[j]:=x;
     for k:=i+1 to t+1 do
      read(f,x);
    end;
   for j:=1 to b do
    write(g,y[j]:10:4);
   writeln(g);
   close(f);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
  end;
 close(g);
 end;

procedure ordena(t,b:integer);
 var i,j,k:integer;
    a,c:real;
    x:array[0..300] of real;
    f,g:text;
 begin
  assign(f,'\sas\tmboot.dat');
  assign(g,'\sas\ordena.dat');
  reset(f);
  rewrite(g);
  for j:=1 to t+1 do
   begin
    for k:=1 to b do
    read(f,x[k]);
    for i:=1 to b-1 do
     begin
      for k:=i+1 to b do
       begin
        a:=x[i];c:=x[k];
        if a>c then
         begin
          x[i]:=c;
          x[k]:=a;
         end;
       end;
     end;
    for k:=1 to b do write(g,x[k]:10:4);
    writeln(g);
   end;
  close(f);
  close(g);
 end;

procedure deri2(t:integer;h:real;x,y:vector;var b:vector);
var i,j:integer;
   a,c,k,s0,s1,s2,s3,s4,t0,t1,t2:real;
 begin
  for i:=-t to t do
   begin
    s0:=0;s1:=0;s2:=0;s3:=0;s4:=0;t0:=0;t1:=0;t2:=0;
    for j:=-t to t do
     begin
      if j<>i then
       begin
        if h<>0 then k:=bp((x[i]-x[j])/h);
        s0:=s0+k;
        s1:=(x[i]-x[j])*k+s1;
        s2:=sqr(x[i]-x[j])*k+s2;
        s3:=sqr(x[i]-x[j])*(x[i]-x[j])*k+s3;
        s4:=sqr(x[i]-x[j])*sqr(x[i]-x[j])*k+s4;
        t0:=k*y[j]+t0;
        t1:=(x[i]-x[j])*k*y[j]+t1;
        t2:=sqr(x[i]-x[j])*k*y[j]+t2;
       end;
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
   end;
   c:=s0*s2*s4+2*s1*s2*s3-sqr(s2)*s2-sqr(s1)*s4-s0*sqr(s3);
   a:=s1*s3*t0+s1*s2*t1+s0*s2*t2-sqr(s2)*t0-s0*s3*t1-sqr(s1)*t2;
  if c<>0 then b[i]:=a/c;
 end;
end;


{PROGRAMA PRINCIPAL}
begin
 TextBackground(lightblue);
 clrscr;
 gotoxy(16,12);
 write('*** ESTIMACION BOOTSTRAP POLINOMICO LOCAL ***');
 gotoxy(40,23);
 write('Pulse cualquier tecla para continuar');
 ka:=readkey;
 clrscr;
 gotoxy(3,7);write('Tama¤o de la muestra bootstrap  b= ');read(b);
 readln;
 gotoxy(3,9);write('N£mero de individuos= ');read(r);
 readln;
 assign(f,'\sas\pm.dat');
 reset(f);
 t:=-1;
 while not eoln(f) do
  begin
   t:=t+1;
   gotoxy(75,2);
   write(t);
   readln(f,pm[t]);
   pm[-t]:=pm[t];
  end;
 close(f);
 n:=2*t;
 for j:=-t to t do
 x[j]:=2*PI*j/n;
 gotoxy(3,11);
 write('Grado de los polinomios locales ==> ');
 read(grado);
 h1:=2*PI/n;
 h1:=2*h1;
 h4:=(x[t]-x[0])/2;
 h2:=h1+(h4-h1)/3;
 h3:=h1+2*(h4-h1)/3;

 {CALCULO DEL BANDWIDTH OPTIMO}
 if grado=2 then
  begin
   mise2(t,h1,x,pm,m1,m);
   mise2(t,h2,x,pm,m2,m);
   mise2(t,h3,x,pm,m3,m);
   mise2(t,h4,x,pm,m4,m);
  end
 else
  begin
   misc1(t,h1,x,pm,m1,m);
   mise1(t,h2,x,pm,m2,m);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
 mise1(t,h3,x,pm,m3,m);
 mise1(t,h4,x,pm,m4,m);
end;
c123:=((m1-m2)/(h1-h2)-(m1-m3)/(h1-h3))/(h2-h3);
c124:=((m1-m2)/(h1-h2)-(m1-m4)/(h1-h4))/(h2-h4);
a:=(c123-c124)/(h3-h4);
bb:=c123-(h1+h2+h3)*a;
c:=((m1-m2)/(h1-h2))-(h1*h1+h1*h2+h2*h2)*a-(h1+h2)*bb;

if m1<m2 then hopt:=h1 else
 begin
  if a=0 then hopt:=-c/(2*bb) else
   begin
    dis:=bb*bb-3*a*c;
    if dis<0 then
     begin
      if a<0 then hopt:=h4;
      if a>0 then hopt:=h1;
     end
    else
     begin
      h5:=(-bb+sqrt(dis))/(3*a);
      h6:=(-bb-sqrt(dis))/(3*a);
     if grado=2 then
      begin
       mise2(t,h5,x,pm,m5,m);
       mise2(t,h6,x,pm,m6,m);
      end
      else
       begin
        mise1(t,h5,x,pm,m5,m);
        mise1(t,h6,x,pm,m6,m);
       end;
      if m5<=m6 then hopt:=h5 else hopt:=h6;
     end;
    end;
 end;
if grado=2 then mise2(t,1.5*hopt,x,pm,mise,m)
 else mise1(t,1.5*hopt,x,pm,mise,m);
for j:=-t to t do
 fwg[j]:=m[j];
if grado=2 then mise2(t,hopt,x,pm,mise,m)
 else mise1(t,hopt,x,pm,mise,m);
for j:=-t to t do
  fwh[j]:=m[j];
deri2(t,hopt,x,pm,f2);
assign(f,'\sas\varf.dat');
reset(f);
for j:=0 to t do
 begin
  read(f,paso);
  if paso<1 then paso:=1.1;
  varf[j]:=sqrt((paso-1)/2);
  varf[-j]:=varf[j];
 end;
close(f);
assign(f,'\sas\boot.dat');
assign(g,'\sas\mboot.dat');
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
rewrite(g);
reset(f);
clrscr;
for k:=1 to b do
begin
 for j:=0 to t do
  begin
   read(f,wb[j]);
   wb[-j]:=wb[j];
  end;
 for j:=-t to t do
   yb[j]:=fwg[j]*(wb[j]-1);
 gotoxy(58,2);
 write('muestra bootstrap=',k);
 if grado=2 then mise2(t,hopt,x,yb,mise,m)
 else mise1(t,hopt,x,yb,mise,m);
 for j:=0 to t do
  begin
   paso:=(sqrt(r)*m[j])/(fwg[j]*varf[j]);
   write(g,paso:9:3);
  end;
 writeln(g);
end;
close(f);
close(g);
trans(t,b);
ordena(t,b);
inf:=trunc(2.5*b/100)+1;
sup:=trunc(97.5*b/100);

assign(f,'\sas\ordena.dat');
reset(f);
clrscr;
for j:=0 to t do
  begin
   for k:=1 to inf do read(f,paso1);
   for k:=inf+1 to sup do read(f,paso2);
   for k:=sup+1 to b do read(f,paso);
   alfa1[j]:=(fwh[j]-sqr(hopt)*f2[j]/10)/(1+(paso2*varf[j])/sqrt(r));
   alfa1[-j]:=alfa1[j];
   alfa2[j]:=(fwh[j]-sqr(hopt)*f2[j]/10)/(1+(paso1*varf[j])/sqrt(r));
   alfa2[-j]:=alfa2[j];
   normal1[j]:=(fwh[j]-sqr(hopt)*f2[j]/10)/(1+((-1.96)*varf[j])/sqrt(r));
   normal1[-j]:=normal1[j];
   normal2[j]:=(fwh[j]-sqr(hopt)*f2[j]/10)/(1+(1.96*varf[j])/sqrt(r));
   normal2[-j]:=normal2[j];
  end;
 close(f);
 assign(g,'\sas\inter.dat');
 rewrite(g);
 for j:=-t to t do
 begin
  writeln(g,alfa1[j]:10:4,alfa2[j]:10:4,normal1[j]:10:4,normal2[j]:10:4);
 end;

 assign(f,'\sas\espobla.dat');
 assign(g,'\sas\hw.dat');
 reset(f);
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

```
 for j:=0 to t do
  begin
   readln(f,x[j],hw[j]);
   x[-j]:=-x[j];
   hw[-j]:=hw[j];
  end;
 close(f);
 rewrite(g);
 for j:=-t to t do writeln(g,x[j]:12:4,hw[j]:12:4);
 close(g);

end.
```

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

BIBLIOGRAFÍA.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

BIBLIOGRAFÍA.

Abramovitch, K and Singh, K (1985). Edgeworth Corrected Pivotal Statistics an the Bootstrap. *The Annals of Statistics.* **13**, *116-132.*

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.*

Artiles, J. (1986). *Cuestiones Notables de Análisis Espectral.* Tesina. Universidad de La Laguna.

Artiles, J., Hernández, C. y Saavedra, P. (1997). Corrección del sesgo en estimación bootstrap del espectro poblacional. XXIII Congreso Nacional de Estadística e Investigación Operativa. Valencia. *Actas del Congreso.* 55.7-55.8.

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.*

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the Vality of the Formal Edgeworth Expansion. *The Annals of Statistics.* **6**, *434-451.*

Bickel, P. and Doksum, K. (1977). *Mathematical Statistics.* Holden-Day.

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.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Box, G. and Jenkins, G. (1976). *Time Series Analysis, Forecasting and Control.* Holden-Day.

Brillinger, D. R. and Krishnaiah, P. R. (1983). *Time Series in the Frequency Domain.* North-Holland.

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 González, 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 dependen 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.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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. (1992). Design-Adaptive Nonparametric Regression. *Journal of American Statistical Association.* **87**, *998-1004.*

Fan, J. and Gijbels, I. (1992). Variable Bandwidth and Local Linear Regression Smoothers. *The Annals of Statistics.* **20**, *2008-2036.*

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* Chapman and Hall.

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 Appoach.* Academic Press.

Franke, J. and Härdle, W. (1992). On-Boorstrapping 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. and Kneip, A. (1995). Searching for Structure in Curves Samples. *Journal of American Statistical Association.* **90** ,*1179-1188.*

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.*

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.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag.

Härdle, W. (1991). *Smoothing Techniques*. Springer-Verlag.

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., Hall, P. and Marron, J. S. (1988). How Far Are Automatically Chosen Regression Smoothing Parameters From Their Optimum? *Journal of the American Statistical Association.* **83,** *86-95.*

Härdle, W. and Marron, J. (1985). Optimal Bandwidth Selection in Nonparametric Regression Function Estimation. *The Annals of Statistics.* **13**, *1465-1481.*

Härdle, W. and Schimek, M. G. (1996). *Statistical Theory and Computational Aspects of Smoothing* . Springer-Verlag.

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.*

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall.

Hernández, C., Artiles, J. and Saavedra, P., (1996). Determining the Bandwidth of a Kernel Spectrum Estimate with Replicated Time Series. XII Symposium on Computational Statistics. COMPSTAT´96. Barcelona. *Proceedins in Computational Statistics 1996.* Short Comunications and Posters. 193-194.

Hernández, C. N., Artiles, J. and Saavedra, P. (1996). Bootstrap Estimation of Population Spectrum with Replicated Time Series. *Academia Canaria de Ciencias.* 7. 19-29.

Hernández, C. N., Artiles, J. and Saavedra, P. (1996). Estimation of the Population Spectrum with Replicated Time Series. *Computational Statistics and Data Analysis.* Enviada la 2ª revisión.

Hernández, C. N. (1997). *Análisis Espectral con Series Replicadas: Estimación del Espectro Poblacional.* Tesis Doctoral. Universidad de Las Palmas de Gran Canaria.

Kalman, R. (1960). A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **82**, *35-45.*

Koopmans, L. H. (1974). *The Spectral Analysis of Time Series.* Academic Press.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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.*

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 Aplications.* **34**, *171-180.*

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* Wiley

Ramlau-Hansen, H. (1983). Smoothing Counting Process Intensities by Means of Kernel Functions. *The Annals of Statistics.* **11**, *453-466.*

Rice, J. (1984). Bandwidth Choise for Nonparametric Regression. *The Annals of Statistics.* **12**, *1215-1230.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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.*

Saavedra, P., Hernández, C. y Artiles, J. (1995). Estimación Espectral con Series Replicadas. XXII Congreso Nacional de Estadística e Investigación Operativa. Sevilla. *Actas del Congreso.* 79-80.

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.

Shibata, R. (1981). An optimal autorregresive spectral estimate. *The Annals of Statistics.* **9**, *300-306.*

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Aplications to Statistics.* Wiley.

Shumway, R. H. (1988). *Applied Statistical Time Series Analysis.* Prentice-Hall International.

Simonoff, J. (1996). *Smoothing Methods in Statistics.* Springer.

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.

Stone, C. J. (1982). Optimal Global Rates of Convergence for Nonparametric Regression. *The Annals of Statistics.* ***10****, 1040-1053.*

Sun, J. and Loader, C. (1994). Simultaneous Confidence Bands for Linear Regression and Smoothing. *The Annals of Statistics. 22, 1328-1345.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003

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.*

Youndjé, E., Sardá, P. and Vieu, P. (1996). Optimal Smooth Hazard Estimates. *Test.* **5**, *379-394.*

Zhao-Guo, C. and Hannan, E. (1980). The Distribution of Periodogram Ordinates. *Journal of Time Series Analysis.* **1**, *73-83.*

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2003