ESTADÍSTICA ESPAÑOLA
Vol. 44, Núm. 150, 2002, págs. 133 a 159

Una revisión de los métodos de remuestreo en series temporales

por

ANDRÉS M. ALONSO

Departamento de Matemáticas

Universidad Autónoma de Madrid

DANIEL PEÑA

Departamento de Estadística y Econometría

Universidad Carlos III de Madrid

JUAN ROMO

Departamento de Estadística y Econometría

Universidad Carlos III de Madrid

RESUMEN

Desde Efron y Tibshirani (1986), se han propuesto varios métodos de remuestreo para datos temporales. En este artículo, presentamos los principales métodos de remuestreo desarrollados para series temporales, centrándonos en el jackknife por bloques móviles, el bootstrap por bloques móviles, y en el bootstrap para modelos autorregresivos, y proponemos nuevas alternativas para los métodos de remuestreo basados en bloques de observaciones.

Palabras claves: series temporales, bootstrap, jackknife.

Clasificación AMS: 62G09.

1. INTRODUCCIÓN

En muchos procedimientos estadísticos es necesario conocer determinadas características de la distribución muestral de los estadísticos o los estimadores empleados. Así, por ejemplo, en el contraste de hipótesis o en la construcción de intervalos de confianza se necesitan los percentiles de la distribución muestral del estadístico, mientras que en problemas de estimación es esencial tener alguna medida de la exactitud (v.g. el sesgo, la varianza o el error cuadrático medio) del estimador obtenido.

Un enfoque clásico para obtener medidas de exactitud de un estimador es calcularlas mediante análogos empíricos de las fórmulas explícitas obtenidas bajo un modelo determinado. Sin embargo, en la mayoría de los estadísticos es muy difícil o imposible obtener fórmulas exactas y explícitas de las medidas de exactitud. Una solución es considerar una aproximación o una expansión asintótica de la medida de exactitud, y estimarla mediante el análogo muestral de la aproximación obtenida. Este enfoque tiene tres desventajas fundamentales: (i) la fórmula exacta o su aproximación están basadas en el modelo postulado, y cuando este es “ligeramente” incorrecto, el estimador obtenido de la exactitud puede resultar no válido, (ii) las derivaciones de las expresiones son difíciles y tediosas en la mayoría de los casos, y (iii) en ocasiones, requieren tamaños muestrales grandes cuando se utilizan expansiones asintóticas (ver sección 1.2 de Shao y Tu (1995)).

Los métodos de remuestreo o submuestreo reemplazan las derivaciones teóricas del enfoque anterior por la evaluación de los estadísticos en remuestras o submuestras obtenidas a partir de los datos originales, y mediante estos valores se obtienen estimadores de las medidas de exactitud o de la distribución muestral del estadístico. Los métodos de remuestreo más populares en la literatura estadística son: *jackknife* de Quenouille (1949) y Tukey (1958), y *bootstrap* de Efron (1979). Más recientemente, Politis y Romano (1994a) introducen el submuestreo, que puede considerarse como una importante generalización de los resultados de Shao y Wu (1989) y Wu (1990) sobre la utilización del *d-jackknife* para obtener estimadores de la varianza y de la distribución muestral.

Un gran número de artículos han probado la validez de estos métodos, o han proporcionado variantes válidas, en situaciones como regresión (Efron (1979), Wu (1986), Politis et al. (1997)), datos censurados (Efron y Tibshirani (1986)), y series temporales (Efron y Tibshirani (1986), Carlstein (1986), Künsch (1989), Bose (1990), Liu y Singh (1992), Kreiss y Franke (1992), Bühlmann (1997), entre otros). Los libros de Efron y Tibshirani (1993), Shao y Tu (1995) y Davison y Hinkley (1997) constituyen valiosas referencias sobre los métodos de remuestreo y cubren una amplia gama de aplicaciones.

En la elaboración de esta revisión hemos puesto énfasis en los métodos de remuestreo desarrollados para series temporales. Como señalan las revisiones de Li y Maddala (1996), Cao (1999), y Berkowitz y Kilian (2000), estos métodos pueden subdividirse teniendo en cuenta si utilizan o no un modelo para la estructura de dependencia de la serie temporal: el enfoque **b**asado en **m**odelo (*BM*) aplica el *bootstrap* de Efron (1979) a los residuos obtenidos a partir de ajustar a la serie temporal un modelo "adecuado" (v.g. modelos autorregresivos en Efron y Tibshirani (1986); modelos de medias móviles en Bose (1990); modelos ARMA en Kreiss y Franke (1992), y el enfoque **n**o **b**asado en **m**odelo (*NBM*) en el que se toman remuestras de bloques de observaciones consecutivas de la serie temporal (v.g. Carlstein (1986) propone estimadores de la varianza basado en bloques no solapados; Künsch (1989) y Liu y Singh (1992) introducen el *jackknife* y el *bootstrap* con bloques solapados para procesos α-*mixing* y *m*-dependientes, respectivamente; Politis y Romano (1992b) estudian un *bootstrap* por bloques de estadísticos; Politis y Romano (1994b) introducen el *bootstrap* estacionario basado en bloques de longitud aleatoria).

Más recientemente, aparecen los trabajos de Paparoditis (1996) sobre *bootstrap* en procesos VAR de orden infinito, y el AR(∞)-*bootstrap* de Kreiss (1992) o el *sieve bootstrap* en Bühlmann (1997) y Bickel y Bühlmann (1999), que podrían incluirse dentro del enfoque *BM*, en los que se asume que la serie temporal sigue un modelo que pertenece a una determinada clase de modelos (v.g. modelos lineales), se aproxima este por una sucesión de modelos de orden finito (el orden o complejidad del modelo crece con el tamaño muestral) pertenecientes a la clase postulada, y se remuestrean los residuos de los modelos de orden finito. En Alonso et al. (2002) se propone un procedimiento basado en el *sieve bootstrap* para construir intervalos de predicción en modelos lineales.

Sherman (1998) evalúa los dos enfoques, *BM* y *NBM*, como estimadores de la varianza, y concluye que dentro de la clase de modelos ARMA y para tamaños muestrales moderados (n = 200) se obtiene poca ganancia con el enfoque *BM* bajo el modelo correcto, aunque esta aumenta para tamaños mayores (n = 1000). Por otra parte, bajo especificaciones "ligeramente" incorrectas, los estimadores *BM* resultan inconsistentes, y en muestras finitas tienen ECM mayor que los *NBM*; su sesgo es similar al de los estimadores *NBM*, pero su varianza es mayor.

Davison y Hall (1993) y Li y Maddala (1996) señalan que los estimadores de la varianza obtenidos mediante *bootstrap* por bloques son sesgados, debido a que este método no reproduce, y de hecho modifica, la estructura de dependencia de la serie temporal. Notemos que, si bien en la serie temporal original los bloques de observaciones son dependientes, en la versión *bootstrap* los bloques se toman como independientes. Los métodos que se proponen en Alonso et al. (2000), se dirigen a la consecución de métodos de remuestreo no basados en modelos que corrijan tales defectos.

El resto del artículo se divide en las siguientes secciones: en la Sección 2 introducimos brevemente los métodos de remuestreo en el contexto de observaciones independientes e igualmente distribuidas como base para la presentación en la Sección 3 de los resultados fundamentales de los enfoques *BM* y *NBM* mencionados anteriormente.

2. MÉTODOS DE REMUESTREO PARA DATOS I.I.D.

2.1. Jackknife

Quenouille (1949) introduce un método para estimar el sesgo de un estimador (posteriormente recibiría el nombre de *jackknife*) mediante su cálculo en conjuntos que se forman a partir de la muestra original omitiendo una observación. Formalmente, Efron y Tibshirani (1993) lo definen como sigue:

Sea $\mathbf{X} = (X_1, X_2 \ldots, X_N)$ una muestra de tamaño N y sea $T_N = T_N(\mathbf{X})$ un estimador de θ. Las *muestras jackknife* son muestras en que se excluye una observación cada vez, i.e., $\mathbf{X}_{(i)} = (X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_N)$, para i = 1,2,...,N, y el estadístico $T_N^{(i)} = T_{N-1}(\mathbf{X}_{(i)})$ es la i-ésima *réplica jackknife* del estimador.

Una vez definidas las réplicas *jackknife*, podemos definir los estimadores *jackknife* del sesgo (b_{Jack}) y de la varianza (v_{Jack}):

$$b_{Jack} = (N-1)(\overline{T}_N - T_N)\,, \tag{1}$$

$$v_{Jack} = \frac{N-1}{N} \sum_{i=1}^{N} \left(T_N^{(i)} - \overline{T}_N\right)^2, \tag{2}$$

donde $\overline{T}_N = N^{-1} \sum_{i=1}^{N} T_N^{(i)}$.

Con el estimador b_{Jack} podemos obtener un estimador *jackknife* de θ dado por: $T_{Jack} = T_N - b_{Jack}$ que bajo determinadas condiciones de regularidad tiene menor sesgo que T_N , aunque en ocasiones resulta de mayor error cuadrático medio (ver ejemplo 1.3 de Shao y Tu (1995)).

El estimador de la varianza, v_{Jack}, es consistente para datos independientes e igualmente distribuidos para un gran número de estadísticos que incluyen funciones de la media muestral, estadísticos U, V, y L, de M-estimadores, y en estadísticos basados en distribuciones empíricas generalizadas. Los resultados de consistencia se obtienen bajo condiciones de suavidad de los estadísticos, que usualmente son más restrictivas que las condiciones utilizadas para obtener la normalidad asintótica de los estadísticos. Un ejemplo clásico de inconsistencia de v_{Jack} es la estimación de la varianza de la mediana muestral. En el capítulo 2 de Shao y Tu (1995) se presentan distintas variantes de diferenciabilidad bajo las cuales es posible obtener consistencia de v_{Jack}.

En Shao y Wu (1989) se propone una solución para estadísticos menos suaves, como los cuantiles muestrales, en la que las réplicas *jackknife* se obtienen de muestras *jackknife* en las que se han excluido d observaciones:

Sea $\mathbf{X} = (X_1,X_2,...,X_N)$ una muestra de tamaño N, sea d un número entero menor que N y sea $r = N - d$. Sea $S_{N,r}$ la colección de todos los subconjuntos de $\{1,2,\ldots,N\}$ de tamaño r. Para cualquier $\mathbf{s} = \{i_1,i_2,...,i_r\} \in S_{N,r}$ obtenemos la *réplica d-jackknife*, $T_N^{\mathbf{s}} = T_r(X_{i1},X_{i2},...,X_{ir})$. El estimador *d-jackknife* de la varianza es:

$$v_{Jack-d} = \frac{r}{dC} \sum_{s \in S_{N,r}} \left(T_N^{\mathbf{s}} - T_N^{(\bullet)}\right)^2, \tag{3}$$

donde $T_N^{(\bullet)} = C^{-1} \sum_{s \in S_{N,r}} T_N^{\mathbf{s}}$ y $C = \#\left(S_{N,r}\right) = \binom{N}{d}$

La consistencia e insesgadez asintótica del estimador $v_{Jack\text{-}d}$ se establece para estadísticos que admitan la expansión siguiente:

$$T_N = \theta + N^{-1} \sum_{i=1}^{N} \phi_F(X_i) + R_N, \tag{4}$$

donde $\phi_F(X_i)$ tiene media 0 y varianza $\sigma_F^2 > 0$, y $R_N = o_P(N^{-1/2})$. Esta condición es menos restrictiva que la diferenciabilidad de T_N. En Shao y Wu (1989) se establecen condiciones para $d = d(N)$ bajo las cuales se tiene la consistencia de $v_{Jack\text{-}d}$ basadas en medidas de suavidad de T_N.

Hasta este punto, solo se ha utilizado *jackknife* para estimar medidas de exactitud, pero en Wu (1990) se presenta el primer resultado sobre la estimación de la distribución muestral de T_N mediante un estimador que se denomina histograma *jackknife*:

Sea T_N un estimador de θ, y $H_N(x) = \Pr\{\sqrt{N}(T_N - \theta) \leq x\}$ la distribución que deseamos estimar. Se define el *histograma jackknife* como:

$$H_{Jack}(x) = Pr_*\left\{\sqrt{\tfrac{Nr}{d}}(T_N^s - T_N) \le x\right\} = C^{-1}\sum_{s\in S_{N,r}} I\left(\sqrt{\tfrac{Nr}{d}}(T_N^s - T_N) \le x\right) \qquad (5)$$

donde Pr_* denota la probabilidad bajo el esquema de muestreo de *d-jackknife* e $I(.)$ denota la función indicatriz.

Wu (1990) establece que $\|H_{Jack} - H_N\|_\infty \xrightarrow{P} 0$, para estimadores que admitan la expansión (4) bajo condiciones sobre d similares a las utilizadas en Shao y Wu (1989). En Politis y Romano (1994a) se generalizan estos resultados bajo hipótesis menos restrictivas:

Sea T_N un estimador de θ, y $H_N(x) = Pr\{\tau_N(T_N - \theta) \le x\}$ la distribución que deseamos estimar. Se define un estimador por submuestras de tamaño d de $H_N(x)$ como:

$$H_{Sub}(x) = C^{-1}\sum_{s\in S_{N,d}} I\left(\tau_d(T_d^s - T_N) \le x\right) \qquad (6)$$

donde $S_{N,d}$ es la colección de subconjuntos de $\{1, 2, \ldots, N\}$ de tamaño d.

Politis y Romano (1994a) obtienen que $\|H_{Sub} - H_N\|_\infty \xrightarrow{P} 0$ bajo la hipótesis de que H_N converge débilmente a una distribución límite, H, que se asume continua, y las siguientes condiciones sobre las constantes normalizadoras y el tamaño de la submuestra: $\tau_d/\tau_N \to 0$ y $d/N \to 0$ cuando $N \to \infty$.

2.2. Bootstrap

El método *bootstrap* propuesto en Efron (1979) es conceptualmente más sencillo que los métodos de remuestreo precedentes. Su pieza clave es la utilización "extrema" del principio de analogía (*plug-in)* que constituye uno de los métodos más simples utilizados para obtener un estimador de un parámetro $\theta = \theta(P)$ donde P es el modelo estadístico postulado. Un estimador *plug-in* o análogo es $\hat{\theta} = \theta(\hat{P})$, donde $\hat{P}$ es un estimador de P. En lo que sigue combinamos el modo en que Efron y Tibshirani (1986), y Shao y Tu (1995) presentan el método bootstrap en una situación general:

Sea $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ un conjunto de datos (no necesariamente I.I.D.) generados por el modelo estadístico P, y sea $T(\mathbf{X})$ el estadístico cuya distribución $\mathbf{L}(T;P)$ deseamos estimar. El método *bootstrap* propone como estimador de $\mathbf{L}(T;P)$ la distribución $\mathbf{L}^*(T^*;\hat{P})$ del estadístico $T^* = T(\mathbf{X}^*)$, donde $\mathbf{X}^*$ es un conjunto de datos generado por el modelo estimado $\hat{P}$.

Notemos que si $\hat{P} = P$, entonces las distribuciones $\mathbf{L}(T;P)$ y $\mathbf{L}^*(T^*;\hat{P})$ coinciden. En algunos casos (ver ejemplo 1.5 de Shao y Tu (1995)) coinciden aún cuando

$\hat{P} \neq P$. De manera que si tenemos un buen estimador de P, es lógico suponer que $\mathbf{L}^*(T^*; \hat{P})$ se aproximará a $\mathbf{L}(T; P)$. Cientos de artículos han comprobado tal suposición, básicamente demostrando resultados del tipo $\rho(\mathbf{L}(T;P), \mathbf{L}^*(T^*; \hat{P})) \rightarrow 0$ en probabilidad o casi seguramente, donde ρ denota una métrica entre medidas de probabilidad, por ejemplo: la distancia del supremo, ρ_∞, como en Singh (1981), o la distancia de Mallows, ρ_r, como en Bickel y Freedman (1981), dos trabajos pioneros en los aspectos teóricos de *bootstrap.*

Otros artículos, que han contribuido a popularizar al *bootstrap* como método de estimación de $\mathbf{L}(T; P)$ son los dedicados a compararlo con otros métodos de estimación, como la aproximación normal o las expansiones de Edgeworth. En esos trabajos se ha mostrado para algunos estadísticos (v.g. $\sqrt{n}(\overline{X}_n - \mu)$ o $\sqrt{n}(\overline{X}_n - \mu)/\hat{\sigma}$) que: (i) el *bootstrap* tiene la misma o más rápida tasa de convergencia que la aproximación normal, y es equivalente a la expansión de Edgeworth de primer orden, y (ii) el *bootstrap* es superior en error cuadrático medio asintótico a la aproximación normal, y equivalente o superior a la expansión de Edgeworth de primer orden (ver en el capítulo 3 de Shao y Tu (1995) resultados para el caso I.I.D., y las referencias fundamentales en este tema).

Concentrándonos en el caso I.I.D. donde el modelo P es la función de distribución, F, que sigue X_i , la forma en que hemos presentado el *bootstrap* incluye el *bootstrap* clásico, que corresponde a estimar F por la función de distribución empírica, $\hat{F}_N = F_N$, el *bootstrap* suavizado para F continua que utiliza $\hat{F}_N = F_{n,h}$ donde $F_{n,h}$ es la función de distribución asociada a un estimador de la función de densidad f (v.g. en Cuevas y Romo (1997) se utilizan estimadores tipo *kernel*), y el *bootstrap* paramétrico cuando se supone que la función de distribución pertenece a una familia paramétrica F_θ y se utiliza $\hat{F}_N = F_{\hat{\theta}}$, donde $\hat{\theta}$ es un estimador de θ.

Otro aspecto a considerar es el tamaño, m, del conjunto de datos, $\mathbf{X}^*$**,** generado por $\hat{F}_N$, típicamente se utiliza m = n, aunque en Bickel y Freedman (1981) para el caso de la media muestral se considera m general y los resultados asintóticos se obtienen bajo $m\rightarrow\infty$ y $n\rightarrow\infty$. Posteriormente, Bickel et al. (1997) desarrollan en detalle el caso $m = o(n)$ y comprueban que esa versión de *bootstrap* es válida en casos donde el *bootstrap* es inconsistente (v.g. intervalos de confianza para los estadísticos $\min_i\{Xi\}$ y $\max_i\{X_i\}$). En ese artículo interpretan el método de submuestreo de Politis y Romano (1994a) como un *bootstrap* sin reemplazamiento de m observaciones obtenidas de F_N , y muestran que este difiere del *bootstrap* con reemplazamiento en $o_P(1)$ si $m = o(n^{1/2})$.

Hasta este punto hemos mostrado el *bootstrap* como un método para estimar la distribución muestral de un estadístico, pero también se utiliza para obtener estimadores de la varianza. Sin embargo, notemos que los resultados de consistencia de $H_{Boot} = \mathbf{L}(T^*, \hat{P})$ son del tipo convergencia en distribución, y la consistencia de $v_{Boot} = Var^*(H_{Boot})$ es la convergencia del momento de segundo orden de H_{Boot}. Cuando los resultados se obtienen bajo la distancia de Mallows, ρ_r, implican la consistencia de v_{Boot}, pero cuando se utiliza la distancia ρ_∞ es sabido que la convergencia en distribución de una sucesión de variables aleatorias no implica la convergencia de momentos, a menos que la sucesión sea uniformemente integrable. Por otra parte, para la consistencia de V_{Boot} es necesario imponer alguna restricción de momentos sobre T_N^*-T_N para evitar que tome valores excepcionalmente grandes (ver el ejemplo 3.5 de Shao y Tu (1995), un caso de inconsistencia de v_{Boot} para un estadístico regular); típicamente se exige:

$$\max_{i_1,i_2,\ldots,i_N} \left| T_N(X_{i_1}, X_{i_2}, \ldots, X_{i_N}) - T_N \right| / \tau_N \xrightarrow{\text{c.s.}} 0, \tag{7}$$

donde el máximo se toma sobre los números enteros i_k tales que $1 \leq i_1 \leq \cdots \leq i_N \leq N$, y $\{\tau_N\}_{N\in \mathbf{N}}$ es una sucesión de números positivos que satisface: $\liminf_N \tau_N > 0$, y $\tau_N = O(e^{nq})$ con $q \in (0, 1/2)$.

Otro punto importante a considerar es el cálculo de los estimadores *bootstrap*, ya que tenemos $\binom{n+m-1}{m-1}$ muestras distintas en el *bootstrap* estándar, e infinitas en el *bootstrap* suavizado o en el paramétrico. En general no se tienen fórmulas explícitas sobre los estimadores *bootstrap*, y el problema se aborda mediante (i) soluciones analíticas o aproximaciones como el método delta, aproximaciones *jackknife*, o el método de punto de silla (ver resultados y referencias en el capítulo 9 de Davison y Hinkley (1997)), o (ii) aproximaciones de Monte Carlo, cuya versión más simple mostramos a continuación:

Sea $\hat{P}$ el modelo estimado utilizando los datos $\mathbf{X} = (X_1, X_2, ..., X_N)$, y queremos obtener aproximaciones de los estimadores *bootstrap* del sesgo, la varianza y la distribución muestral del estadístico $T(\mathbf{X})$ dados por:

$$b_{Boot} = E^*\left[T(\mathbf{X}^*)\right] - T(\mathbf{X}), \tag{8}$$

$$v_{Boot} = Var^*\left(T(\mathbf{X}^*)\right), \tag{9}$$

$$H_{Boot}(x) = Pr^*\left\{T(\mathbf{X}^*) \leq x\right\}, \tag{10}$$

donde $\mathbf{X}^*$ es una muestra obtenida de $\hat{P}$, y E^*, Var^* y Pr^* denotan la esperanza, la varianza y la probabilidad *bootstrap* dada la muestra **X.** Un procedimiento de Monte Carlo para estimar (8) - (10) es:

1. Generar B muestras independientes $\mathbf{X}^{*(b)}$ a partir de $\hat{P}$, con b = 1,2,…,B.

2. Calcular $T^{*(b)} = T(\mathbf{X}^{*(b)})$ para b = 1, 2, …, B, y aproximar b_{Boot}, v_{Boot}, *y* $H_{Boot}(x)$ por:

$$b_{Boot}^{(B)} = \frac{1}{B}\sum_{b=1}^{B} T^{*(b)} - T(\mathbf{X}), \tag{11}$$

$$v_{Boot}^{(B)} = \frac{1}{B}\sum_{b=1}^{B}\left(T^{*(b)} - \frac{1}{B}\sum_{b=1}^{B} T^{*(b)}\right)^2, \tag{12}$$

$$H_{Boot}^{(B)}(x) = \frac{1}{B}\sum_{b=1}^{B} I\left(T^{*(b)} \le x\right). \tag{13}$$

La selección de B no es un problema sencillo, Efron y Tibshirani (1986) sugieren utilizar B entre 50 y 200 para estimadores de momentos, y al menos 1000 remuestras para estimadores de distribuciones o cuantiles. Más recientemente, Andrews y Buchinsky (1998) y Davidson y MacKinnon (2000) proponen procedimientos para seleccionar el número de remuestras en la estimación de errores estándar, intervalos y regiones de confianza, y contrastes de hipótesis.

3. MÉTODOS DE REMUESTREO PARA SERIES TEMPORALES

3.1. Métodos basados en modelos

En este apartado se presentan los resultados fundamentales de los métodos *bootstrap* para series temporales que utilizan modelos de la estructura de dependencia presente en la serie. Dado que las primeras aplicaciones de este enfoque BM se realizaron bajo formulaciones paramétricas de los modelos, v.g. AR(p) en Freedman (1984) y Efron y Tibshirani (1986), y ARMA(p, q) en Kreiss y Franke (1992), se le suele denominar *enfoque paramétrico*; sin embargo, su característica esencial es la utilización de los residuos del modelo estimado, sea de forma paramétrica como en las referencias anteriores o de forma no paramétrica como en Franke et al. (1997).

Básicamente, la idea de estos métodos está desarrollada en el apartado 2.2 y únicamente tendremos que especificar qué entendemos por el modelo estadístico P en el caso de series temporales. Una formulación bastante general de P que engloba los modelos utilizados en las referencias fundamentales donde se ha utilizado este enfoque es:

$$X_t = g(\mathbf{F}_{-\infty}^{t-1}, \varepsilon_t), \tag{14}$$

donde $\{\varepsilon_t\}_{t\in \mathbf{Z}}$ es una sucesión de variables aleatorias I.I.D. que siguen una distribución F_ε, $\mathbf{F}_{-\infty}^{t-1}$ es el conjunto de información hasta el momento t-1 (podemos suponer que es el conjunto formado por $(X_1,X_2,...,X_{t-1})$ y $(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_{t-1})$) y g es una función de enlace. En esta formulación el modelo P es el par (g, F_ε).

De manera que, para aplicar *bootstrap* a un estadístico T(**X**) calculado en series temporales **X** generadas por P, necesitaremos un estimador $\hat{P} = (\hat{g}, \hat{F}_{\hat{\varepsilon}})$ y estimaremos las características de T(**X**) mediante sus análogas de $T^*(\mathbf{X}^*)$, i.e., el estadístico calculado en la remuestra $\mathbf{X}^*$ generada por $\hat{P}$. Si bien esta formulación es “general”, los resultados de consistencia se han obtenido para estadísticos concretos (típicamente, estimadores de los parámetros del modelo paramétrico postulado, o estimadores no paramétricos de g) en determinados modelos (AR(p), MA(q), ARMA(p, q), autorregresiones no paramétricas, entre otros) sobre los que suele imponerse la hipótesis de estacionariedad junto con restricciones de momentos sobre F_ε.

Por supuesto, también en este caso necesitaremos procedimientos para calcular los estimadores *bootstrap*, y el método de Monte Carlo es la solución más empleada. Ilustramos este enfoque en el caso de un proceso autorregresivo de orden p:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t, \tag{15}$$

donde $\{\varepsilon_t\}_{t\in \mathbf{Z}}$ es una sucesión de variables I.I.D. con $E[\varepsilon_t] = 0$ y $E[\varepsilon_t^2] = \sigma_\varepsilon^2$. Uno de los procedimientos planteados es como sigue:

1. Obtener estimaciones a partir de $\mathbf{X}=(X_1,X_2,...,X_N)$ de los parámetros autorregresivos, $\phi = (\phi_1, \phi_2, \dots, \phi_p)$.

2. Calcular los residuos, $\hat{\varepsilon}_t = X_t - \sum_{i=1}^{p} \hat{\phi}_i X_{t-i}$, para t = p + 1,p + 2,...,N.

3. Calcular la función de distribución empírica de los residuos centrados, $\tilde{\varepsilon}_t = \hat{\varepsilon}_t - (N-p)^{-1} \sum_{t=p+1}^{N} \hat{\varepsilon}_t$, mediante:

$$F_N^{\tilde{\varepsilon}}(x) = \frac{1}{N-p} \sum_{t=p+1}^{N} I(\tilde{\varepsilon}_t \le x). \tag{16}$$

Y repetir los siguientes pasos para b = 1, 2 ...,B:

4. Obtener N - p observaciones independientes e igualmente distribuidas según $F_N^{\tilde{\varepsilon}}$ que denotamos ($\varepsilon_{p+1}^{*(b)}, \varepsilon_{p+2}^{*(b)}, \dots, \varepsilon_N^{*(b)}$).

5. Fijar los primeros p valores de la serie ($X_1^{*(b)}$, $X_2^{*(b)}$,..., $X_p^{*(b)}$) y calcular las restantes observaciones de la remuestra, ($X_{p+1}^{*(b)}, X_{p+2}^{*(b)}, \ldots, X_N^{*(b)}$), mediante:

$$X_t^{*(b)} = \sum_{i=1}^{p} \hat{\phi}_i X_{t-i}^{*(b)} + \varepsilon_t^{*(b)} . \tag{17}$$

6. Utilizando la remuestra, $\mathbf{X}^{*(b)} = (X_1^{*(b)}, X_2^{*(b)}, \ldots, X_n^{*(b)})$, calcular los análogos *bootstrap* de los estimadores.

El procedimiento anterior merece algunos comentarios:

• Los estimadores $\hat{\phi}$ en el paso (1), son usualmente los estimadores de mínimos cuadrados (v.g. Efron y Tibshirani (1986), Stine (1987) y Thombs y Schucany (1990)). También se tienen resultados de convergencia para otros métodos de estimación, v.g. Kreiss y Franke (1992) prueban la consistencia para M-estimadores de los parámetros ϕ y θ de un modelo ARMA(p,q) estacionario, y Stoffer y Wall (1991) para estimadores máximo verosímiles gaussianos de los parámetros de un modelo en el espacio de los estados.

• Stine (1987) sugiere reescalar los residuos centrados mediante:

$$\tilde{\varepsilon}_t = \sqrt{\frac{n-p}{n-2p}} \left(\hat{\varepsilon}_t - (N-p)^{-1} \sum_{t=p+1}^{N} \hat{\varepsilon}_t \right). \tag{18}$$

De esta manera tenemos que $E^*[\varepsilon_t^{*2}] = \hat{\sigma}^2 = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\varepsilon}_t^{*2}$.

• Como estimador de F_ε suele utilizarse la función de distribución empírica de los residuos (centrados y reescalados), pero pueden emplearse estimadores asociados a estimadores kernel de la función de densidad como en Cao et al. (1997), donde se ilustra el buen comportamiento de la versión suavizada del procedimiento anterior.

• Para los valores iniciales ($X_1^{*(b)}, X_2^{*(b)}, \ldots, X_p^{*(b)}$), Efron y Tibshirani (1986) proponen utilizar ($X_1, X_2, \ldots, X_p$), o pueden obtenerse de una selección aleatoria de los n - p + 1 bloques de *p* observaciones consecutivas como en Stine (1987).

• Este método de remuestreo como señalan Cao et al. (1997) es válido si lo que se desea es estimar características de los parámetros del modelo. Sin embargo, para el caso de intervalos de predicción debe ser modificado, pues no replica la distribución condicional de las observaciones futuras X_{t+k} con k > 0 dada la serie observada **X**, en particular para un modelo AR(p) dadas las últimas p observaciones. Notemos que en este caso nos interesa la probabilidad $\Pr\{X_{t+k} < x | X_{t-p+1}, \ldots, X_t\}$, y su análogo *bootstrap* es $Pr^*\{X_{t+k}{}^* \leq x|X^*_{t-p+1}, ..., X^*_t\}$, de manera que se impone $X_{t-p+1}{}^*=X_{t-p+1}$,..., $X_t{}^*=X_t$; ver Thombs y Schucany (1990) y la extensión a modelos ARI(p,d) de García-Jurado et al. (1995). En Alonso et al. (2002) se generaliza el enfoque de Cao et al. (1997) a una clase general de modelos lineales estacionarios que incluye como caso particular a los modelos ARMA(p, q*).*

Otros trabajos, dentro del enfoque BM se han dirigido a considerar estadísticos asociados a contrastes de hipótesis, por ejemplo: contrastes de raíz unidad en Ferretti y Romo (1996) y en Heimann y Kreiss (1996), contraste de especificación de modelos ARFIMA en Delgado e Hidalgo (1998), y contrastes de cointegración en Li y Maddala (1997).

En Kreiss (1988), Kreiss (1992) y Bühlmann (1997) se supone que el proceso es estacionario y que admite una representación autorregresiva infinita:

$$X_t = \sum_{i=1}^{\infty} \phi_i X_{t-i} + \varepsilon_t \,, \tag{19}$$

donde $\{\varepsilon_t\}$ es una sucesión de variables I.I.D. con $E[\varepsilon_t] = 0$ y $E[\varepsilon_t^2] = \sigma^2$, y proponen como estimador de la distribución muestral del estadístico de interés (estimadores Yule-Walker de ϕ en Kreiss (1992), y $T_N = T(F^{(m)})$ donde $F^{(m)}$ es la distribución empírica m-dimensional en Bühlmann (1997)) la distribución *bootstrap* del estadístico calculado en una remuestra que se construye a partir de la estimación de un modelo AR(p(n)). Kreiss (1992) y Bühlmann (1997) establecen la consistencia del procedimiento (*sieve bootstrap*) bajo condiciones sobre (i) el orden, $p = p(n)$, del modelo autorregresivo, (ii) existencia de momentos de ε_t, (iii) restricciones sobre la representación autorregresiva. En Bühlmann (1997) se imponen condiciones sobre la diferenciabilidad de los funcionales T. Bickel y Bühlmann (1999) extienden los resultados de *sieve bootstrap* para estadísticos no regulares.

3.2. Métodos no basados en modelos

En ocasiones la estructura de dependencia de la serie temporal no permite reconocer el modelo “correcto” y, como mencionan Künsch (1989), Li y Maddala (1996) e ilustra Sherman (1998), la utilización de un modelo mal especificado en los métodos BM conduce a estimadores inconsistentes. Nótese que esto también puede ocurrir en el caso I.I.D. si empleamos el *bootstrap* paramétrico con $F_{\hat{\theta}}$ (i.e. si suponemos un modelo para la distribución) y la distribución real, F, está muy alejada de la familia paramétrica postulada. En el caso I.I.D. una solución es utilizar en el *bootstrap* la función de distribución empírica que es una estimación (no basada en modelo) de F, y otra solución es el submuestreo.

Singh (1981) y Liu y Singh (1992) muestran que el *bootstrap* estándar y el *jackknife* son inconsistentes (como estimadores de la varianza) si las observaciones son m-dependientes. Los métodos que presentamos en esta sección son modificaciones de estos dos procedimientos que corrigen este "defecto".

Carlstein (1986) propone un estimador para la varianza de un estadístico, $\sqrt{N}\,T_N$, basado en los valores que toma en submuestras no solapadas de **X**:

Sea **X** = $(X_1,X_2,...,X_N)$ un vector de observaciones de una serie temporal estrictamente estacionaria. Sean $l = l(n)$ *y* $k = k(n)$ dos enteros positivos tales que $k = \lceil n/l \rceil$. Se definen las siguientes submuestras no solapadas de **X** : $\mathbf{X}_{jl}^{l}$ = $(X_{jl+1},X_{jl+2},...,X_{jl+l})$ para j = 0,1,...,k-1. Sea $T_{jl}^{l} = T_l(\mathbf{X}_{jl}^{l})$ el estadístico T evaluado en la submuestra $\mathbf{X}_{jl}^{l}$. Se propone como estimador de la varianza asintótica de $\sqrt{N}T_N$ a:

$$\hat{\sigma}_N^2 = \frac{l}{k}\sum_{j=0}^{k-1}\left(T_{jl}^{l} - \overline{T}\right)^2, \tag{20}$$

donde $\overline{T} = k^{-1}\sum_{j=0}^{k-1} T_{jl}^{l}$.

Carlstein (1986) demuestra que $\hat{\sigma}_N^2 \xrightarrow{L_2} \sigma^2 = \lim_N \mathrm{Var}(\sqrt{N}T_N)$, suponiendo que $\{X_t\}_{t\in \mathbf{Z}}$ *es* α-mixing, bajo condiciones de integrabilidad uniforme de $n^2(T_N - E[T_N])^4$, $l\to\infty$ y $l = o(n)$. Este método recuerda al *jackknife* por grupos en el caso I.I.D. (ver en Shao y Wu (1989)), y puede considerarse como un antecesor con bloques no solapados del submuestreo de Politis y Romano (1994a).

Künsch (1989) y Liu y Singh (1992) de manera independiente proponen el *jackknife* y el *bootstrap* por bloques móviles (MBJ de "*moving blocks jackknife*" y MBB de "*moving blocks bootstrap*") para series α-*mixing* y m-dependientes, respectivamente. A continuación presentamos el MBJ y MBB según Liu y Singh (1992):

MBJ: Sea **X** = $(X_1,X_2,...,X_N)$ un vector de observaciones de una serie temporal estrictamente estacionaria. Sea l un número entero menor que N. Se definen las siguientes submuestras o bloques de observaciones: $\mathbf{X}j^{l}$ = $(X_{j+1},X_{j+2},...,X_{j+l})$ para j=0,1,...,N-l. Sea $T_N^{(j)}$ el estadístico evaluado en $\mathbf{X}\backslash\mathbf{X}j^{l}$ = $(X_1,X_2,..., X_j,X_{j+l+1},...,X_N)$. Se propone como estimador de la varianza de $\sqrt{N}T_N$ a:

$$v_{MBJ} = \frac{(N-l)^2}{l(N-l+1)}\sum_{j=0}^{N-l}\left(T_N^{(j)} - T_N\right)^2. \tag{21}$$

MBB: Supongamos que se seleccionan con reemplazamiento k bloques de los N-l+1 posibles, que denotamos por $\mathbf{X}_{ji}^{*}$, con i = 1, 2, …, k. Sea h = kl el tamaño de la serie $\mathbf{X}^{*}$ que resulta de unir los bloques $\mathbf{X}_{ji}^{*}$, i.e. $\mathbf{X}^{*} = (\mathbf{X}_{j1}^{*}, \mathbf{X}_{j2}^{*}, \ldots, \mathbf{X}_{jk}^{*})$. Sea T_h^{*} el estadístico calculado en la serie $\mathbf{X}^{*}$. Se propone como estimador de la distribución muestral H_N de $\sqrt{N}(T_N - T(F))$ a:

$$H_{MBB}(x) = Pr^{*}\left\{\sqrt{h}(T_h^{*} - T_N) \le x\right\}. \qquad (22)$$

Liu y Singh (1992) demuestran que $v_{MBJ} \xrightarrow{c.s.} \sigma^2$ y que $\rho_\infty(H_{MBB}, H_N) \to 0$ en probabilidad, suponiendo que $\{X_t\}_{t\in \mathbf{Z}}$ es m-dependiente, bajo diferenciabilidad del funcional T(·) que define al estadístico $T_N = T(F_N)$, condiciones de momentos de la diferencial de T(·), *y* $l\to\infty$, $l = o(n)$ *o* $l = o(n^{1/3})$ en MBJ dependiendo del orden de diferenciabilidad impuesto, y $l = o(n^{1/2}/\ln(n))$ en MBB.

Künsch (1989) propone un MBJ y un MBB similares a Liu y Singh (1992) orientado a estadísticos definidos mediante funcionales que dependen de la distribución m-dimensional de $(X_1, X_2, \ldots, X_m)$:

Sea $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ un vector de observaciones de una serie temporal estrictamente estacionaria. Sean $\mathbf{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ con t = 1,2,…,n=N-m+1, bloques de m observaciones consecutivas. Sea T un funcional que define al estadístico $T_N = T(\rho_N^m)$, donde $\rho_N^m = (N-m+1)^{-1}\sum_{t=1}^{N-m+1}\delta_{(X_t,\ldots,X_{t+m-1})} = n^{-1}\sum_{t=1}^{n}\delta_{\mathbf{X}_t}$ es la distribución empírica m-dimensional.

MBJ: Se infravaloran o subponderan l "observaciones" consecutivas $(\mathbf{X}_{j+1}, \mathbf{X}_{j+2}, \ldots, \mathbf{X}_{j+l})$, i.e., l bloques consecutivos de m observaciones. Sea $T_N^{(j)} = T(\rho_N^{m,(j)})$ el estadístico calculado sobre la distribución empírica en la que se han infravalorado l bloques: $\rho_N^{m(j)} = (n - \|w_n\|_1)^{-1}\sum_{t=1}^{n}(1 - w_n(t-j))\delta_{\mathbf{X}_t}$ donde los pesos w_n cumplen que $0 \le w_n(i) \le 1$ para $i\in\mathbf{Z}$, $w_n(i) > 0$ si y sólo si $1 \le i \le l$, y $\|w_n\|_1 = \sum_{i=1}^{l} w_n(i)$. Se propone como estimador de la varianza de $\sqrt{N}T_N$ a:

$$v_{MBJ} = \frac{(n - \|w_n\|_1)^2}{\|w_n\|_2^2 (n-l+1)} \sum_{j=0}^{n-l}\left(T_N^{(j)} - T_N^{(\bullet)}\right)^2, \qquad (23)$$

donde $T_N^{(\bullet)} = (n-l+1)^{-1}\sum_{j=0}^{n-l} T_N^{(j)}$ y $\|w_n\|_2^2 = \sum_{i=1}^{l} w_n(i)^2$.

MBB: Se asume n = kl, y se seleccionan k bloques de l "observaciones" consecutivas de los n - l + 1 bloques posibles que denotamos por $\mathbf{Y}_{ji}^{*} = (\mathbf{X}_{ji+1}, \mathbf{X}_{ji+2}, \ldots, \mathbf{X}_{ji+l})$ con

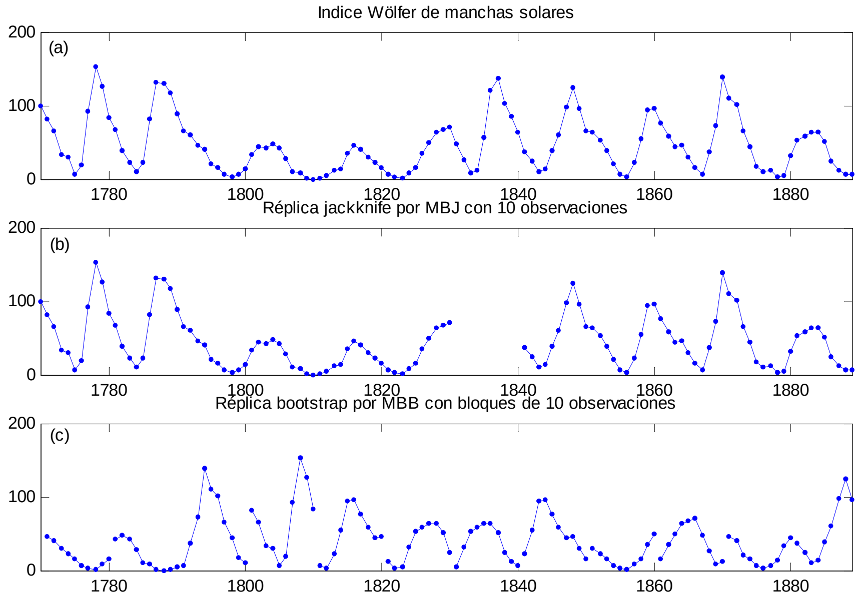
$i = 1,2,...,k$. Sea $\mathbf{X}^*$ la serie que resulta de unir los k bloques, $\mathbf{X}^* = (\mathbf{Y}_{j1}^*, \mathbf{Y}_{j2}^*, ..., \mathbf{Y}_{jk}^*) = (\mathbf{X}_1^*, \mathbf{X}_2^*, ..., \mathbf{X}_n^*)$, sea ρ_N^{m*} la distribución empírica m-dimensional *bootstrap*, $\rho_N^{m*} = n^{-1}\sum_{i=1}^{k}\sum_{t=1}^{l}\delta_{\mathbf{X}_{ji+t}^*} = n^{-1}\sum_{t=1}^{n}\delta_{\mathbf{X}_t^*}$, y sea $T_N^* = T(\rho_N^{m*})$. Se propone como estimador de la distribución muestral H_N de $\sqrt{n}(T_N - T(\rho^m))$ a:

$$H_{MBB}(x) = \Pr{}^*\left\{\sqrt{n}\left(T_N^* - E^*\left[T_N^*\right]\right) \leq x\right\}. \tag{24}$$

Nótese que los estimadores (21)-(22), para el caso m = 1 y cuando $w_n(i) = 1$ para $1 \leq i \leq l$ (i.e., en el caso en que se eliminan los bloques) coinciden con (23)-(24) salvo en que utilizan $T_N^{(\bullet)}$ y $E^*[T_N^*]$ en lugar de T_N, respectivamente. En la figura 1 ilustramos la formación de una muestra *jackknife* por MBJ con eliminación del bloque y por MBB. Los datos corresponden a la serie del índice Wölfer de manchas solares en el período 1770-1889, estudiada por Efron y Tibshirani (1986) y por Künsch (1989).

Figura 1

(a) SERIE ORIGINAL (b) MUESTRA POR MBJ, (c) MUESTRA POR MBB



Künsch (1989) demuestra que $v_{MBJ} \xrightarrow{P} \sigma^2$ para los siguientes tipos de estadísticos:

1. $T_N = g\left(n^{-1}\sum_{t=1}^{n} h(X_t, X_{t+1}, \ldots, X_{t+m-1})\right) = g\left(n^{-1}\sum_{t=1}^{n} h(\mathbf{X}_t)\right)$, donde $h: R^m \rightarrow R^q$, y $g: R^q \rightarrow R$.

2. $T_N = N^{-q}\sum_{t_1=1}^{N}\sum_{t_2=1}^{N}\cdots\sum_{t_q=1}^{N}\phi\left(X_{t_1}, X_{t_2}, \ldots, X_{t_q}\right)$, donde $\phi: R^q \rightarrow R$ es una función simétrica en sus argumentos.

3. Estadísticos definidos implícitamente, como las soluciones de:

$$\sum_{t=1}^{n}\psi(X_t, X_{t+1}, \ldots, X_{t+m-1}; T_N) = \sum_{t=1}^{n}\psi(\mathbf{X}_t; T_N) = 0 .$$

Los estadísticos anteriores admiten la representación

$$T_N = T\left(\rho_N^m\right) + n^{-1}\sum_{t=1}^{n} IF_\rho(\mathbf{X}_t) + R_N, \qquad (25)$$

donde $R_N = o_P(n^{-1/2})$, e IF_ρ es la curva de influencia de T. La consistencia de (23) se obtiene bajo condiciones sobre g, h, ϕ y ψ, (típicamente, diferenciabilidad y acotación de la función o sus derivadas), restricciones en la dependencia temporal y condiciones sobre el tamaño del bloque del tipo: $l \to \infty$ y $l = o(n^{1/2})$.

En MBB, Künsch (1989) establece para la media muestral que $\rho_\infty(H_{MBB}, H_N) \to 0$ c.s. para procesos α-mixing bajo restricciones de momentos $E[|X_t|^\delta]$ con $\delta > 2$, y $l = o(n^{1/2-1/\delta})$. En Politis y Romano (1992a) se obtienen resultados análogos bajo condiciones menos restrictivas. Para extender los resultados a otros estadísticos, en particular cuando m > 1, es necesario establecer la convergencia débil del proceso empírico *bootstrap*, W_N^*, al proceso empírico, W_N. Bühlmann (1994) obtiene este resultado para procesos α*-mixing* con restricciones en la dependencia, continuidad de la distribución de X, y $l = o(n^{1/2-\varepsilon})$ con $\varepsilon > 0$, y lo utiliza para demostrar la consistencia de los GM-estimadores de los parámetros de un proceso AR(p).

Politis y Romano (1992b) proponen un esquema de remuestreo de bloques de estadísticos calculados en bloques de observaciones, que permite considerar estadísticos que dependan de la distribución infinito-dimensional:

Sea $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ un vector de observaciones de una serie temporal estrictamente estacionaria. Sean $\mathbf{X}_{iL}^m = (X_{(i-1)L+1}, X_{(i-1)L+2}, \ldots, X_{(i-1)L+m})$ con $i = 1, 2, \ldots$, $Q = \lceil (N - m)/L \rceil + 1$ y sea $T_{iL}^m = T_m(\mathbf{X}_{iL}^m)$ el estadístico T evaluado en el bloque $\mathbf{X}_{iL}^m$. Sea $\overline{T}_N = Q^{-1}\sum_{i=1}^{Q} T_{iL}^m$ el estadístico cuya varianza y distribución muestral deseamos estimar. Se definen los siguientes bloques de estadísticos $T_j = (T^m_{((j-1)h+1)L}, T^m_{((j-1)h+2)L}, \ldots, T^m_{((j-1)h+b)L})$ para $j = 1, 2, \ldots, q = \lceil (Q - b)/h \rceil + 1$.

MBJ: Sea $\overline{T}_{N,j}$ la media de los T_{iL}^m restantes después de eliminar el j-ésimo bloque. Se propone como estimador de la varianza de $\sqrt{Q}\,\overline{T}_N$ a:

$$v_{MBJ} = \frac{(Q-b)^2}{bq}\sum_{j=1}^{q}\left(\overline{T}_{N,j} - \overline{T}_N\right)^2. \qquad (26)$$

MBB: Se seleccionan k bloques con reemplazamiento de $(\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_q)$. Sea $s = bk$ el tamaño del bloque de estadísticos resultante, $(T_1^*, T_2^*, \ldots, T_s^*)$. Se propone como estimador de la distribución muestral H_N de $\sqrt{Q}\left(\overline{T}_N - \theta\right)$ a:

$$H_{MBB}(x) = \Pr{}^{*}\left\{\sqrt{s}(\overline{T}_s^{*} - \overline{T}_N) \leq x\right\} \tag{27}$$

donde $\overline{T}_s^{*} = s^{-1}\sum_{i=1}^{s} T_i^{*}$.

Politis y Romano (1992b) demuestran que $v_{MBJ} \xrightarrow{P} \sigma^2$ y que $\rho_\infty(H_{MBB}, H_N) \to 0$ en probabilidad para procesos α-*mixing* con restricciones de dependencia y sobre el tamaño de los bloques, $m = o(N)$, $L \sim am$ con $a \in (0, 1]$, $b \to \infty$, $b = o(Q)$ y $h = O(b)$. Un inconveniente de este procedimiento es que sólo puede aplicarse a estadísticos de la forma $\overline{T}_N = Q^{-1}\sum_{i=1}^{Q} T_{iL}^{m}$.

En Bühlmann y Künsch (1995) se extienden los resultados de Künsch (1989) a estadísticos que dependen de la distribución m-dimensional siendo m una función creciente de n, bajo la condición de que el tamaño del bloque, l, crezca más rápido que m. Por otra parte, muestran que la condición $L \sim am$ y $h = O(b)$ de los parámetros que controlan el solapamiento entre bloques en Politis y Romano (1992b) conduce a pérdida de eficiencia en la estimación de la varianza.

Un aspecto importante en los métodos MBJ y MBB anteriores es la determinación de l (en Politis y Romano (1992b) de L, h, b, k). Bühlmann y Künsch (1994) lo estudian para estimadores de la varianza por MBB en estadísticos lineales. En Hall et al. (1995) proponen un método iterativo para determinar l que se basa en que el tamaño de bloque “óptimo” es de orden $n^{1/3}$, $n^{1/4}$, y $n^{1/5}$ para estimadores de la varianza (o del sesgo), estimadores unilaterales de la distribución, y estimadores bilaterales de la distribución, respectivamente.

Un problema que presenta MBB es que la serie generada no es estacionaria, pues por ejemplo, $E^{*}[X_t^{*} X_{t+1}^{*}]$ depende de t, en particular de si el par (X_t^{*}, X_{t+1}^{*}) está en el medio de un bloque o forma parte de los extremos final e inicial de dos bloques consecutivos. En Politis y Romano (1994b) se propone el *bootstrap estacionario* que corrige este problema mediante bloques de longitud aleatoria.

Sea $\mathbf{X} = (X_1, X_2, ..., X_N)$ un vector de observaciones de una serie estrictamente estacionaria. Se definen los bloques de observaciones $\mathbf{X}^{iL} = (X_i, X_{i+1}, ..., X_{i+L-1})$, y en el caso en que X_j sea tal que $j > N$ se define $X_j = X_{j\,mod(N)}$. Sea $p \in (0, 1)$ y sea $\{L_i\}_{i \in \mathbf{N}}$ una sucesión de variables aleatorias I.I.D. con distribución geométrica de parámetro p, i.e. $\Pr\{L = l\} = (1-p)^{l-1}p$. La pseudo-serie $\mathbf{X}^{*} = (X_1^{*}, X_2^{*}, ..., X_N^{*})$ está formada por las N primeras observaciones de la unión de los bloques: $(\mathbf{X}_{i_1}^{L_1}, \mathbf{X}_{i_2}^{L_2}, \ldots, \mathbf{X}_{i_{N1}}^{L_N})$. Se proponen como estimadores de la varianza de $\sqrt{N}\,T_N$ y de la distribución muestral de $\sqrt{N}\,(T_N - \theta)$ a sus análogos *bootstrap*:

$$v_{BE} = \mathrm{Var}^{*}\left(\sqrt{N}T_N^{*}\right), \tag{28}$$

$$H_{BE}(x) = Pr^*\left\{\sqrt{N}\left(T_N^* - E^*\left[T_N^*\right]\right) \le x\right\}. \qquad (29)$$

Politis y Romano (1994b) demuestran que las series generadas, $\mathbf{X}^*$, son estacionarias, que $v_{BE} \xrightarrow{P} \sigma^2$ para la media muestral y que $\rho_\infty(H_{BE},H_N) \to 0$ en probabilidad para la media muestral y funciones suaves de la media, suponiendo que $\{X_t\}_{t\in}$ es α-*mixing* con restricciones en la dependencia, restricciones de momentos, y $p=p(N)\to 0$ y $Np(N)\to\infty$.

Carlstein et al. (1998) reinterpretan las remuestras de *bootstrap* por bloques como realizaciones de un proceso cuyos estados son los bloques definidos y las probabilidades de transición son constantes, i.e., no dependen de los estados. Generalizan este enfoque al considerar probabilidades de transición variables y dependientes de las observaciones finales e iniciales de los bloques que quedarán unidos. A pesar de lo atractivo que resulta este método de *bootstrap* por bloques pareados, tiene el inconveniente de los aspectos técnicos que involucra. Nótese que Carlstein et al. (1998) estudian solo el caso de un estimador de la varianza del estadístico media muestral para una serie temporal que se obtiene de un proceso estocástico continuo en el que se supone que la frecuencia de muestreo aumenta con el tamaño muestral. No obstante, los resultados de sus simulaciones apuntan a que los procedimientos que corrijan los efectos de la unión completamente aleatoria de los bloques, obtendrán estimadores con menor sesgo y con poco efecto en la varianza.

Motivados por esto último, en Alonso et al. (2000) se propone una modificación del *jackknife* por bloques móviles (MBJ) y del *bootstrap* por bloques móviles (MBB). Las observaciones eliminadas en MBJ se consideran como datos faltantes y se sustituyen por una estimación que tiene en cuenta la estructura de dependencia presente en la serie temporal. En ese caso el estimador de la varianza es una varianza muestral ponderada del estadístico calculado sobre muestras de series “completas”. Podemos interpretar este procedimiento como una manera de realizar una transición suave entre los bloques con peso completo del *jackknife* por bloques. En la figura 2 ilustramos la formación de una muestra *jackknife* por MBJ con eliminación del bloque y por M^2BJ (de *missing moving blocks jackknife*) en el que se estima el bloque omitido. En ambas muestras hemos eliminado las observaciones correspondientes al período 1830-1839. Para la estimación de los datos omitidos hemos supuesto que el modelo generador es un AR(2), como se sugiere en Box y Jenkins (1976).

Extendemos este procedimiento al *bootstrap* por bloques. Para ello definimos un bloque de observaciones faltantes entre los bloques que forman la serie o remuestra *bootstrap*. En este caso, podemos interpretar el procedimiento como un mecanismo de unión de los bloques, y en ese sentido se asemeja al *bootstrap* por bloques pareados de Carlstein et al. (1998). Una manera de interpretar una remuestra por M^2BB (de *missing moving blocks bootstrap*) es tomar l + k como el tamaño de los bloques en MBB y entonces considerar las últimas k observaciones en cada bloque como datos omitidos. En la figura 3 ilustramos la formación de una muestra *bootstrap* por MBB con bloques de tamaño l = 10, y por M^2BB con bloques de tamaño l + k = 9 + 1 en 3(b) y de tamaño l + k = 8 + 2 en 3(c). Para la estimación de los datos omitidos hemos utilizado el AR(2) estimado en la muestra $(X_1, X_2, ..., X_N)$.

Figura 2

(a) MUESTRA JACKKNIFE POR MBJ, (b) MUESTRA JACKKNIFE POR M^2BJ

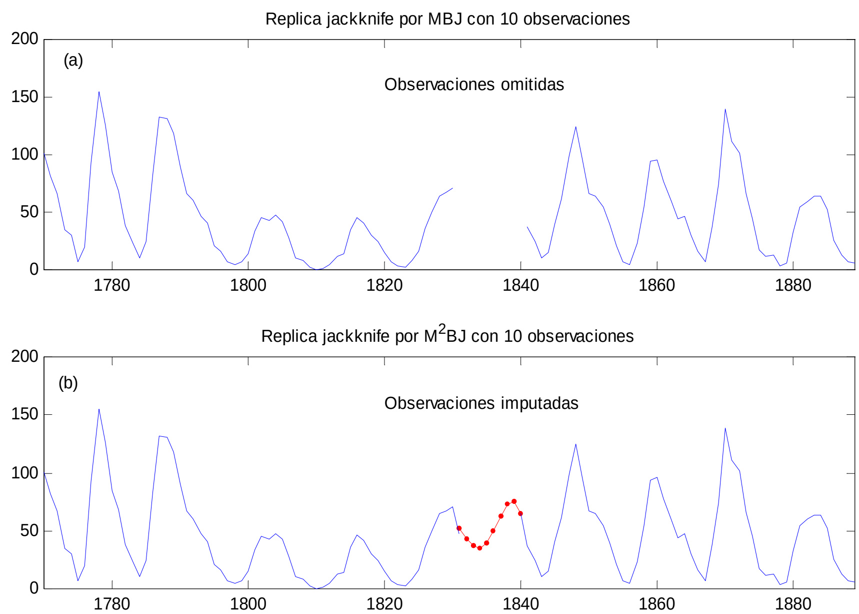
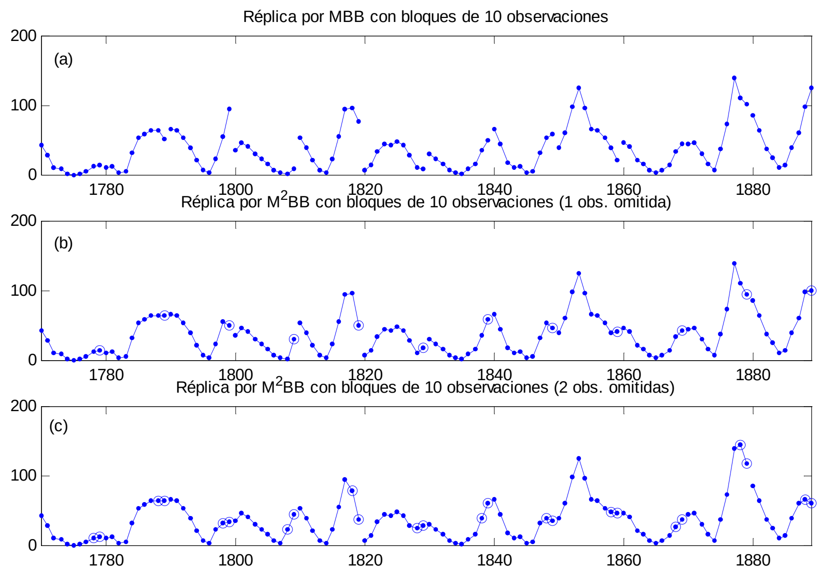


Figura 3

(a) MUESTRA BOOTSTRAP POR MBB, (b) MUESTRA BOOTSTRAP POR M^2BB CON UNA OBSERVACIÓN OMITIDA EN CADA BLOQUE, (c) MUESTRA BOOTSTRAP POR M^2BB CON DOS OBSERVACIONES OMITIDAS EN CADA BLOQUE



Politis y Romano (1994a) introducen el *submuestreo* para campos estocásticos homogéneos, que incluyen como caso particular las series estrictamente estacionarias. Por simplicidad, describimos el método en este caso:

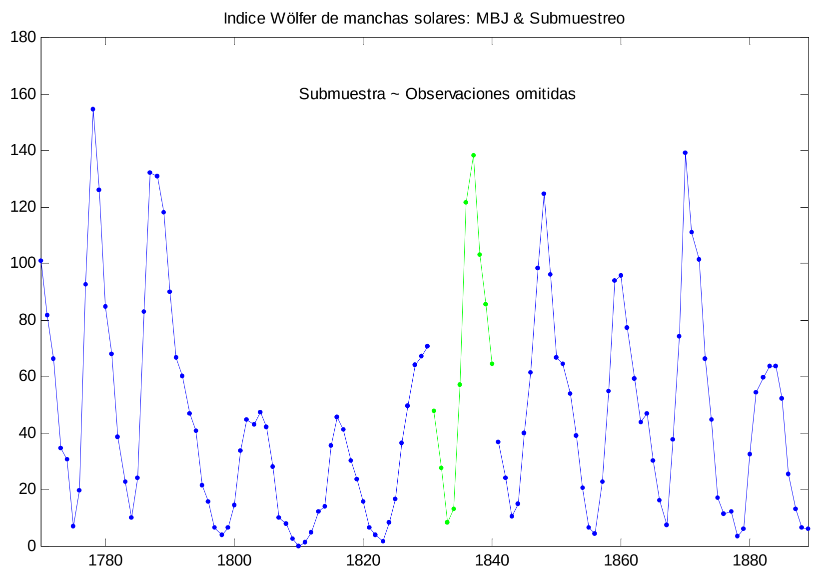
Sea $\mathbf{X} = (X_1, X_2, ..., X_N)$ un vector de observaciones de una serie estacionaria. Sean $\mathbf{X}_j = \mathbf{X}_j^l = (X_j, X_{j+1}, ..., X_{j+l-1})$ con j = 1,2,...,n-l+1 bloques o submuestras de l observaciones consecutivas. Sea $T_l^{(j)} = T_l(\mathbf{X}_j)$ el estadístico, T, evaluado en la submuestra $\mathbf{X}_j$. Se propone como estimador de la distribución muestral, H_N, de $\tau_N (T_N - T(F))$ a:

$$H_{Sub}(x) = \frac{1}{n-l+1} \sum_{j=1}^{n-l+1} I\left(\tau_l \left(T_l^{(j)} - T_N\right) \leq x\right). \tag{30}$$

La hipótesis básica en el submuestreo es que H_N converja débilmente a una distribución límite, H. Politis y Romano (1994a) demuestran que $\rho_\infty(H_{Sub}, H_N) \to 0$ en probabilidad bajo continuidad de H, suponiendo que $\{X_t\}_{t\in \mathbf{Z}}$ es α-*mixing*, $\tau_l / \tau_N \rightarrow 0$, $l\rightarrow\infty$ y $l/n\rightarrow 0$. Politis et al. (1997) extienden, bajo hipótesis similares, los resultados a series temporales heterocedásticas, y ejemplifican su uso para la media muestral k-dimensional, funciones suaves de la media, y los estimadores de mínimos cuadrados de una regresión lineal múltiple. Por último, en la figura 4 ilustramos la formación de una submuestra con tamaño l = 10 y su estrecha relación con el MBJ, hecho que se utiliza en Alonso et al. (2000) para establecer la consistencia de los estimadores de la distribución de estadísticos lineales.

Figura 4

MUESTRA JACKKNIFE POR MBJ Y SUBMUESTRA



REFERENCIAS

ALONSO, A. M., PEÑA, D., Y ROMO, J. (2000). «Resampling time series by missing values techniques». *Working Paper* 00-42, Universidad Carlos III de Madrid, Madrid.

ALONSO, A. M., PEÑA, D., Y ROMO, J. (2002). «Forecasting time series with sieve bootstrap». *Journal of Statistical Planning and Inference***,** 100, 1-11.

ANDREWS, D. W. K. Y BUCHINSKY, M. (1998). «On the number of bootstrap repetitions for bootstrap standard errors, confidence intervals, confidence regions, and tests». *Cowles Foundation Discussion Paper* 1141R, Yale University, Yale.

BERKOWITZ, J. Y KILIAN, L. (2000). «Recent developments in bootstrapping time series». *Econometric Reviews,* 19**,** 1-48.

BICKEL, P. J. Y BÜHLMANN, P. (1999). «A new mixing notion and functional central limit theorems for a sieve bootstrap in time series». *Bernoulli, 5***,** 413-446.

BICKEL, P. J. Y FREEDMAN, D. A. (1981). «Some asymptotic theory for the bootstrap». *The Annals of Statistics,* 9, 1196-1217.

BICKEL, P. J., G¨OTZE, F., Y VAN ZWET, W. R. (1997). «Resampling fewer than n observations: Gains, losses, and remedies for losses». *Statistica Sinica***,** 7, 1-31.

BOSE, A. (1990). «Bootstrap in moving average models». *Annals of the Institute of Statistical Mathematics,* 42, 753-768.

BOX, G. E. P. Y JENKINS, G. M. (1976). «Time Series Analysis: Forecasting and Control». San Francisco: Holden-Day.

BÜHLMANN, P. (1994). «Blockwise bootstrapped empirical process for stationary sequences». *The Annals of Statistics,* 22, 995-1012.

BÜHLMANN, P. (1997). «Sieve bootstrap for time series». *Bernoulli,* 3**,** 123-148.

BÜHLMANN, P. Y KÜNSCH, H. R. (1994). «Block length selection in the bootstrap for time series». *Research Report* 72, Eidgenössische Technische Hochschule, Zürich.

BÜHLMANN, P. Y KÜNSCH, H. R. (1995). «The blockwise bootstrap for general parameters of stationary time series». *Scandinavian Journal of Statistics. Theory and Applications,* 22**,** 35-54.

CAO, R. (1999). «An overview of bootstrap methods for estimating and predicting time series». *TEST,* 8, 95-116.

CAO, R., FEBRERO-BANDE, M., GONZÁLEZ-MANTEIGA, W., PRADA-SÁNCHEZ, J.M., Y GARCÍA-JURADO, I. (1997). «Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes». *Communications in Statistics. Simulation and Computation***,** 26, 961-978.

CARLSTEIN, E. (1986). «The use of subseries values for estimating the variance of a general statistics from a stationary sequence». *The Annals of Statistics, 14***,** 1171-1194.

CARLSTEIN, E., DO, K., HALL, P., HESTERBERG, T., Y KÜNSCH, H. R. (1998). «Matched-block bootstrap for dependent data». *Bernoulli,* 4*,* 305-328.

CUEVAS, A. Y ROMO, J. (1997). «Differentiable functionals and smoothed bootstrap». *Annals of the Institute of Statistical Mathematics,* 49**,** 355-370.

DAVIDSON, R. Y MACKINNON, J. G. (2000). «Bootstrap tests: How many bootstraps?» *Econometric Reviews,* 19, 55-68.

DAVISON, A. C. Y HALL, P. (1993). «On studentizing and blocking methods for implementing the bootstrap with dependent data». *Australian Journal of Statistics,* 35, 215-224.

DAVISON, A. C. Y HINKLEY, D. V. (1997). «Bootstrap Methods and their Applications». Cambridge: Cambridge University Press.

DELGADO, M. A. E HIDALGO, J. (1998). «Consistent specification testing of stationary processes with long-range dependence: Asymptotic and bootstrap tests». *Working Paper* 98-50, Universidad Carlos III de Madrid, Madrid.

EFRON, B. (1979). «Bootstrap methods: Another look at the jackknife». *The Annals of Statistics*, 7, 1-26.

EFRON, B. Y TIBSHIRANI, R. J. (1986). «Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy». *Statistical Science***,** 1*,* 54-77.

EFRON, B. Y TIBSHIRANI, R. J. (1993). «An introduction to the bootstrap». New York: Chapman & Hall.

FERRETTI, N. Y ROMO, J. (1996). «Unit root bootstrap tests for AR(1) models». *Biometrika*, 83, 849-860.

FRANKE, J., KREISS, J.-P., Y MAMMEN, E. (1997). «Bootstrap of kernel smoothing in nonlinear time series». *Technical Report*, Technische Universität Braunschweig, Braunschweig.

FREEDMAN, D. A. (1984). «On bootstrapping two-stage least-square estimates in stationary linear models». *The Annals of Statistics,* 12**,** 827-842.

GARCÍA-JURADO, I., GONZÁLEZ-MANTEIGA, W., PRADA-SÁNCHEZ, J.M., FEBRERO-BANDE, M., Y CAO, R. (1995). «Predicting using Box-Jenkins, nonparametric, and bootstrap techniques». *Technometrics,* 37, 303-310.

HALL, P., HOROWITZ, J. L., Y JING, B. Y. (1995). «On blocking rules for the bootstrap with dependent data». *Biometrika***,** 82, 561-574.

HEIMANN, G. Y KREISS, J. P. (1996). «Bootstrapping general first order autoregression». *Statistics & Probability Letters,* 30, 87-98.

KREISS, J.-P. (1988). «Asymptotical Inference of Stochastic Processes». *Ph. D. thesis*, Universität Hamburg.

KREISS, J.-P. (1992). «Bootstrap procedures for AR(∞)-processes». In: *Bootstrapping and Related Technique*s*,* K. H. Jöckel, G. Rothe y W. Sendler eds., 107-113. Heidelberg: Springer-Verlag.

KREISS, J.-P. Y FRANKE, J. (1992). «Bootstrapping stationary autoregressive moving average models». *Journal of Time Series Analysis,* 13, 297-317.

KÜNSCH, H. R. (1989). «The jackknife and the bootstrap for general stationary observations». *The Annals of Statistics,* 17, 1217-1241.

LI, H. Y MADDALA, G. (1996). «Bootstrapping time series models». *Econometric Theory,* 15**,** 115-195.

LI, H. Y MADDALA, G. (1997). «Bootstrapping cointegrating regressions». *Journal of Econometrics,* 80*,* 297-318.

LIU, R. Y. Y SINGH, K. (1992). «Moving blocks jackknife and bootstrap capture weak dependence». In: *Exploring the Limits of Bootstrap,* R. Lepage y L. Billard eds., 225-248. New York: Wiley.

PAPARODITIS, E. (1996). «Bootstrapping autoregressive and moving average parameter estimates of infinite order vector processes». *Journal of Multivariate Analysis*, 57*,* 277-296.

POLITIS, D. N. Y ROMANO, J. F. (1992a). «A circular block-resampling procedure for stationary data». In: *Exploring the Limits of Bootstrap,* R. Lepage y L. Billard eds., 263-270. New York: Wiley.

POLITIS, D. N. Y ROMANO, J. F. (1992b). «A general resampling scheme for triangular arrays of α-mixing variables with application to the problem of spectral density estimation». *The Annals of Statistics,* 20, 1985-2007.

POLITIS, D. N. Y ROMANO, J. F. (1994a). «Large sample confidence regions based on subsamples under minimal assumptions». *The Annals of Statistics,* 22**,** 2031-2050.

POLITIS, D. N. Y ROMANO, J. F. (1994b). «The stationary bootstrap». *Journal of the American Statistical Association*, 89*,* 1303-1313.

POLITIS, D. N., ROMANO, J. F., Y WOLF, M. (1997). «Subsampling for heteroskedatic time series». *Journal of Econometrics,* 81**,** 281-317.

QUENOUILLE, M. (1949). «Approximation test of correlation in time series». *Journal of the Royal Statistical Society. Series B,* 11, 18-84.

SHAO, J. Y TU, D. (1995). «The Jackknife and Bootstrap». New York: Springer-Verlag.

SHAO, J. Y WU, C. F. J. (1989). «A general theory for jackknife variance estimation». *The Annals of Statistics,* 17*,* 1176-1197.

SHERMAN, M. (1998). «Efficiency and robustness in subsampling for dependent data». *Journal of Statistical Planning and Inference,* 75**,** 133-146.

SINGH, K. (1981). «On the asymptotic accuracy of Efron's bootstrap». *The Annals of Statistics,* 9*,* 1187-1195.

STINE, R. A. (1987). «Estimating properties of autoregressive forecasts». *Journal of the American Statistical Association,* 82**,** 1072-1078.

STOFFER, D. S. Y WALL, K. D. (1991). «Bootstrapping space-state models: Gaussian maximum likelihood estimation and Kalman filter». *Journal of the American Statistical Association,* 86**,** 1024-1033.

THOMBS, L. A. Y SCHUCANY, W. R. (1990). «Bootstrap prediction intervals for autoregression». *Journal of the American Statistical Association,* 85**,** 486-492.

TUKEY, J. (1958). «Bias and confidence in not quite large samples». *The Annals of Mathematical Statistics,* 29**,** 614.

WU, C. F. J. (1986). «Jackknife, bootstrap and others resampling methods in regression analysis». *The Annals of Statistics,* 14, 1261-1350.

WU, C. F. J. (1990). «On the asymptotic properties of the jackknife histogram». *The Annals of Statistics,* 1***8***, 1438-1452.

A REVIEW OF RESAMPLING METHODS IN TIME SERIES

SUMMARY

Since Efron and Tibshirani (1986), several resampling methods have been proposed for time series data. In this paper, we present the main resampling methods developed for time series, we focus the presentation in the moving blocks jackknife, the moving blocks bootstrap, and the bootstrap in autoregressive models. Also, we propose some new alternatives for resampling methods based on blocks of observations.

Key words: Time series, bootstrap, jackknife.

AMS Classification: 62G09.