Bussi, Javier
Marí, Gonzalo
Méndez Fernanda
Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística

ESTIMACIÓN DE LA FUNCIÓN DE AUTOCORRELACIÓN EN MODELOS AR(1) CON MÉTODOS DE REPLICACIONES

Resumen

La función de autocorrelación (FAC) es una herramienta fundamental en el análisis de series de tiempo lineales, entre otras muchas cuestiones, para la identificación del modelo. La estimación muestral de la FAC es altamente sensible a la presencia de observaciones extremas. El objetivo del presente trabajo es comparar distintos estimadores de la FAC propuestos en la literatura con cinco estimadores basados en métodos de replicaciones a través del sesgo y el Error Cuadrático Medio (ECM). Cuatro de estos estimadores son variaciones basadas en la técnica *Jackknife* para series de tiempo con bloques móviles. El restante es una adaptación del estimador a través del método Bootstrap Rápido y Robusto (FRB) con el fin de estimar la FAC. Se compararon las estimaciones para el rezago de orden 1 de la FAC de un modelo AR(1). El estimador basado en el método FRB parece ser un serio competidor del estimador altamente robusto MG ya que es superior en cuanto a sesgo y ECM en los casos de ausencia de outliers y resulta tener un comportamiento similar en los casos simulados con un outlier con Φ=±0,9; ±0,6. Lo mismo ocurre si se lo compara con el estimador Trun2, el cual tiene un buen comportamiento cuando existen outliers. Para los casos en donde el valor de Φ en valor absoluto es 0,3, el desempeño del método FRB decae notablemente. Los estimadores basados en el método *Jackknife* se comportan razonablemente bien en presencia de observaciones extremas pero no logran en ningún caso superar el desempeño logrado por el estimador MG.

Abstract

The autocorrelation function (ACF) is a fundamental tool in the analysis of linear time series, among other things, for model identification. The sample estimate of the ACF is highly sensitive to the presense of outliers. The aim of this study is to compare differents estimator of the ACF proposed in the literature with five estimators based on resampling methods through the bias and the Mean Square Error (MSE). Four of these estimatos are variations based on Jackknife for time series with moving blocks. The other is an adaptation of the estimator through the method Fast and Robust Bootstratp (FRB) in order to estimate the ACF. The estimates of the lag of order 1 of the ACF of an AR(1) model were compared. The estimator based on the FRB method seems to be a serious competitor of the highly robust estimator MG as it is better in terms of bias and MSE in the cases without outliers, and found to have a similar behaviour in simulated cases with one outlier if Φ=±0.9; ±0.6. The same is true when compared with the Trun2 estimator, which has a good behavior when there are outliers. For cases with small absolute values of Φ, such as 0.3, the FRB method performance declines significantly. Estimates based on the Jackknife method behave reasonably well in the presence of outliers but canot, in any case, overcome the performance achieved by the MG estimator.

Palabras claves: FUNCIÓN DE AUTOCORRELACIÓN, ESTIMADORES ROBUSTOS, JACKKNIFE, BOOTSTRAP

INTRODUCCIÓN

La función de autocorrelación (FAC) es una herramienta fundamental en el análisis de series de tiempo lineales. Además de los fines descriptivos, es utilizada, entre otras muchas cuestiones, para la identificación del modelo. La estimación muestral de la FAC es altamente sensible a la presencia de observaciones extremas, generándose sesgos de distinto tipo. Varias alternativas robustas han sido propuestas para solucionar o atenuar el problema. El objetivo del presente trabajo es comparar distintos estimadores de la FAC propuestos en la literatura con cinco estimadores basados en métodos de replicaciones a través del sesgo de los estimadores y el Error Cuadrático Medio (ECM). Cuatro de estos estimadores son variaciones basadas en la técnica *Jackknife* para series de tiempo con bloques móviles. El restante es una adaptación del estimador a través del método Boootstrap Rápido y Robusto (FRB: Fast and Robust Bootstrap) con el fin de estimar la FAC. Se compararon las estimaciones para el rezago de orden 1 de la FAC de un modelo autoregresivo del mismo orden AR(1).

METODOLOGÍA

Sea un proceso estacionario z_t, se define la función de autocorrelación (FAC) de dos valores cualquiera de la variable z_t en los tiempos t y $t+k$ como

$$\rho(t,t+k)=\rho(k)=\frac{Cov(z_t;\ z_{t+k})}{\sqrt{Var(z_t)\,Var(z_{t+k})}}=\frac{\gamma(k)}{\gamma(0)},k=0,\pm1,\pm2,\ldots$$

A partir de una serie de tiempo observada $z_1,z_2,\ldots,z_n$, se proponen diversos estimadores de la FAC. En primer lugar, se consideró la estimación muestral clásica por el método de los momentos (Est) la cual resulta sensible a la presencia de observaciones extremas. La misma se obtiene de la siguiente forma

$$\hat{\rho}_s(k)=\frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}=\frac{\sum_{t=1}^{n-k}(z_t-\bar{z})(z_{t+k}-\bar{z})}{\sum_{t=1}^{n}(z_t-\bar{z})^2},\qquad k=0,1,2,\ldots,(n-1)$$

donde $\bar{z}=\frac{1}{n}\sum_{t=1}^{n}z_t$

Como ya se mencionó, el estimador Est resulta sensible a la aparición de valores considerados outliers debido a que se basa en estimadores no robustos. Por tal motivo, se presentan una serie de estimadores robustos presentes en la bibliografía. Se consideró un estimador robusto de la FAC que trunca un porcentaje de las observaciones más extremas. Se postularon dos valores para el truncamiento: 2% y 5% que dieron origen a los estimadores Trun2 y Trun5, respectivamente (Chan & Wei, 1992). Por otro lado, se presenta un estimador basado en un estimador de escala altamente robusto (MG) (Ma & Genton, 2000). Luego se consideraron estimadores basados en los métodos de replicaciones

Jackknife y *Bootstrap*. Para el primero de ellos, cuatro estimadores basados en la técnica *Jackknife* para series de tiempo con bloques móviles en donde el primero de ellos utiliza la estimación muestral de la FAC (JackMB), los dos siguientes utilizan los estimadores robustos que truncan un porcentaje (2%, 5%) de las observaciones extremas (JackMB2, JackMB5 respectivamente) (Chambers, 2013), y el cuarto combina la estimación robusta MG y la técnica *Jackknife*, que inserta este estimador computándolo en cada réplica seleccionada (JackMB-MG). Por último se postuló el estimador a través del método *Bootstrap Rápido y Robusto* (FRB). Se partió de la estimación propuesta para los estimadores de regresión y se lo adaptó al caso de un modelo autorregresivo de primer orden AR(1). Para cada serie generada se obtuvo un estimador FRB replicando 1000 muestras bootstrap. Una vez estimado el parámetro correspondiente, el mismo fue utilizado en la estimación de la FAC.

A continuación se presenta brevemente, los estimadores propuestos.

Estimador truncado de la función de autocorrelación

Dada una serie de tiempo observada ordenada $z_{(1)} \leq z_{(2)} \leq \ \dots \ \leq z_{(n)}$, la función de autocorrelación muestral α-recortada o truncada (Trunα) está definida por:

$$\hat{\rho}_{T,\mathrm{Trun}\alpha}(k) = \frac{\hat{\gamma}_{T,\alpha}(k)}{\hat{\gamma}_{T,\alpha}(0)}, \qquad k = 0,1,2,\dots,(n-1)$$

donde

$$\hat{\gamma}_{T,\alpha}(k) = \frac{1}{\sum_{t=k+1}^{n} L_{t-k}^{(\alpha)} L_t^{(\alpha)}} \left\{ \sum_{t=k+1}^{n} \left(z_{t-k} - \bar{z}^{(\alpha)}\right)\left(z_t - \bar{z}^{(\alpha)}\right) L_{t-k}^{(\alpha)} L_t^{(\alpha)} \right\}$$

con

$$\bar{z}^{(\alpha)} = \frac{\sum_{t=1}^{n} z_t \, L_t^{(\alpha)}}{\sum_{t=1}^{n} L_t^{(\alpha)}}$$

y

$$L_t^{(\alpha)} = \begin{cases} 0 & si \ \ z_t \leq \ z_{(g)} \text{ ó } \ z_t \geq z_{(n-g+1)} \\ 1 & en\ otro\ caso \end{cases},$$

donde g es la parte entera de $[\alpha n]$ y $0 \leq \alpha < 0.5$.

Para el presente trabajo se consideraron dos valores de α, 2 y 5 por ciento.

Estimador de escala altamente robusto de la función de autocorrelación

Propuesto por Ma y Genton (2000), el estimador propuesto, a diferencia del estimador empírico, se basa en una medida de escala en lugar de una medida de posición. El estimador robusto viene dado por

$$\hat{\gamma}_{MG}(k) = {}^{1}\!/_{4}\left(Q_{n-k}^{2}(\boldsymbol{u}+\boldsymbol{v}) - Q_{n-k}^{2}(\boldsymbol{u}-\boldsymbol{v})\right)$$

donde $\boldsymbol{u} = (z_1, z_2, \dots, z_{n-k})$ y $\boldsymbol{v} = (z_k, z_{k+1}, \dots, z_n)$ y $Q_n(\boldsymbol{x})$ es un estimador de escala robusto propuesto por Rousseeuw y Croux (1992, 1993) que se calcula de la siguiente forma

$$Q_n(\boldsymbol{x}) = c\left\{|z_i - z_j|;\ i < j\right\}_{(h)}$$

donde *c* es una constante elegida de forma tal que el estimador sea consistente (*c=2.21191* para la distribución normal), $h = \left[\left(\binom{n}{2}+2\right)/4\right]+1$, $\left\{|z_i - z_j|;\ i < j\right\}_{(h)}$ representa la estadística de orden h-ésima de las diferencias absolutas entre pares de la serie observada.

Estimador basados en el método Jackknife

El método *Jackknife* fue desarrollado por Quenouille (1949) dentro del ámbito de las series de tiempo. El mismo consiste en la eliminación de una o varias observaciones de un conjunto de datos denominándose al conjunto resultante réplica, en la estimación del parámetro con la muestra reducida, y en repetir este procedimiento un número determinado de veces, obteniendo a partir del promedio de las estimaciones, un estimador del parámetro de interés.

Uno de los métodos de selección de réplicas es el denominado bloques móviles, donde, dado un bloque de longitud fija, el mismo se mueve a través de la serie de tiempo de a una observación por vez, calculando en cada movimiento el estimador correspondiente a la réplica luego de eliminar de la serie original el bloque móvil.

El método puede resumirse en los siguientes pasos

1. Definir un bloque de longitud n_M, lo que determina M bloques. Eliminar de la serie original las primeras n_M observaciones.
2. Con la muestra reducida estimar la FAC a partir de un método propuesto (*met*). Llamemos $\hat{\rho}_{met(-1)}(k)$ a este estimador, donde (-1) indica que fue eliminado el primer bloque.
3. Mover el bloque una observación y repetir el paso 2.
4. Repetir el paso 3 hasta llegar al final de la serie
5. El estimador de la FAC a partir de este método es

$$\hat{\gamma}_{Jack}(k) = \frac{N}{M-1}\hat{\rho}_{met}(k) - \frac{(N-M+1)}{(M-1)}\frac{1}{M}\sum_{i=1}^{M}\hat{\rho}_{met(-i)}(k)$$

Como puede observarse, el método Jackknife necesita definir con qué método se realizarán las estimaciones en el punto 2. Para la presente investigación, se consideraron el método basado en los momentos (JackMB), los estimadores truncados con 2 y 5% de fracción de truncamiento (JackMB2 y JackMB5 respectivamente), y el método debido a Ma y Genton (JackMB-MG)

Estimador basados en el método Bootstrap Rápido y Robusto (FRB)

Un modelo AR(1) puede ser escrito de la siguiente forma $z_t = \emptyset z_{t-1} + a_t$, donde $\emptyset$ es la función de autocorrelación para el rezago 1. Luego, la estimación de la misma, puede ser visto un problema de estimación del coeficiente de regresión en un modelo de regresión simple. Se considerará el método FRB para la estimación del mismo.

Se describe a continuación el método FRB en el caso de regresión lineal en general. Se considera el escenario donde se cuenta con variables explicativas aleatorias y donde se aplica el método a estimadores de regresión MM. Las observaciones $\mathbf{x}_i$ no están prefijadas como en diseño de experimentos sino que son variables aleatorias que se observan conjuntamente con la variable dependiente y_i. Se describe brevemente a continuación el modelo lineal. Las observaciones $\mathbf{x}_i$ tienen dimensión *px1* donde se considerará el caso de *p=2*, es decir, una regresión lineal con una variable explicativa y una ordenada al origen. Se cuenta con *n* observaciones independientes e idénticamente distribuidas (i.i.d) de vectores aleatorios $(yi, \mathbf{z}_i{}')'$ con distribución común *H* donde $\boldsymbol{x}i \ = (1, \mathbf{z}_i{}')'$.

Se considera el modelo lineal:

$$y_i = \mathbf{x}_i^{'} \boldsymbol{\beta}_0 + \sigma_0 \varepsilon_i \qquad\qquad i = 1, \dots, n$$

La situación ideal corresponde a suponer que y_i y $\mathbf{z}_i$ son independientes, donde $y_i \sim F_0$ siendo ésta alguna distribución simétrica (en general la normal estándar), $\mathbf{z}_i \sim G_0$, $(y_i, \mathbf{z}_i{}')' \sim H_0$. Con el fin de considerar la presencia de observaciones atípicas (outliers) y otros desvíos con respecto al modelo clásico, se supone que la distribución *H* pertenece al entorno de contaminación:

$$H_\epsilon = \{H = (1-\epsilon)H_0 + \epsilon H^*\}$$

donde $0 \le \epsilon \le 1/2$ y H^* es una distribución cualquiera no especificada.

Los estimadores MM de regresión están basados en dos funciones de pérdida ρ_0 y ρ_1, las cuales determinan la eficiencia y el punto de quiebre de la estimación, respectivamente**.**

El estimador MM, $\widehat{\boldsymbol{\beta}}_n$, satisface la ecuación:

$$\frac{1}{n}\sum_{i=1}^{n} \rho_1^{'}\left(\frac{y_i - \boldsymbol{x}_i^{'}\widehat{\boldsymbol{\beta}}_n}{\hat{\sigma}_n}\right) = 0$$

donde $\hat{\sigma}_n$ es un estimador S de escala que minimiza el M estimador de escala $\hat{\sigma}_n(\boldsymbol{\beta})$ definido en la ecuación:

$$\frac{1}{n}\sum_{i=1}^{n}\rho_0\left(\frac{y_i-\mathbf{x}_\mathrm{i}^{\prime}\widehat{\boldsymbol{\beta}}_\mathrm{n}}{\hat{\sigma}_n}\right)=0.$$

La distribución asintótica de los estimadores MM ya ha sido estudiada el caso del modelo central paramétrico donde $H=H_0$. Sin embargo, esta situación no se presenta en situaciones donde se pretende utilizar estimadores MM altamente robustos. El método FRB descripto a continuación conduce a estimadores consistentes de la covariancia de $\widehat{\boldsymbol{\beta}}_\mathrm{n}$ bajo condiciones generales que incluye el caso en que $H\in H_{\epsilon}$.

Sea $\widetilde{\boldsymbol{\beta}}_\mathrm{n}$ el estimador de regresión *S* asociado a la ecuación

$$\frac{1}{n}\sum_{i=1}^{n}\rho_0\left(\frac{y_i-\mathbf{x}_\mathrm{i}^{\prime}\widehat{\boldsymbol{\beta}}_\mathrm{n}}{\hat{\sigma}_n}\right)=0$$

Se desea realizar inferencias acerca del parámetro de regresión $\boldsymbol{\beta}_0$. Utilizando el mismo criterio de "plug-in" del método BC propuesto por Efron (1979), se propone el siguiente método que produce un gran número de estimadores de regresión $\widehat{\boldsymbol{\beta}}_\mathrm{n}^{*}$ recalculados a partir de las muestras generadas. A partir de la función de distribución empírica de esta estadística se estima la distribución muestral del estimador de regresión $\widehat{\boldsymbol{\beta}}_\mathrm{n}$. Para cada vector $(yi,\boldsymbol{z_i}^{\prime})^{\prime}$ perteneciente a la muestra se definen los residuos asociados con $\widehat{\boldsymbol{\beta}}_\mathrm{n}$ y $\widetilde{\boldsymbol{\beta}}_\mathrm{n}$:

$$r_i=y_i-\mathbf{x}_\mathrm{i}^{\prime}\widehat{\boldsymbol{\beta}}_\mathrm{n}$$
$$\tilde{r}_i=y_i-\mathbf{x}_\mathrm{i}^{\prime}\widetilde{\boldsymbol{\beta}}_\mathrm{n}$$

Se debe notar que tanto $\widehat{\boldsymbol{\beta}}_\mathrm{n}$ como $\hat{\sigma}_n$ pueden ser expresados como el resultado de un ajuste por mínimos cuadrados ponderados (MCP). Para $i=1,\dots,n$ se definen los pesos ω_i y v_i como:

$$\omega_i=\rho_1^{\prime}(r_i/\hat{\sigma}_n)/r_i$$
$$v_i=\frac{\hat{\sigma}_n}{nb}\frac{\rho_0(\tilde{r}_i/\hat{\sigma}_n)}{\tilde{r}_i}$$

pudiendo entonces representarse los estimadores de la siguiente manera:

$$\widehat{\boldsymbol{\beta}}_n=\left[\sum_{i=1}^{n}\omega_i\mathbf{x}_\mathrm{i}\mathbf{x}_\mathrm{i}^{\prime}\right]^{-1}\sum_{i=1}^{n}\omega_i\mathbf{x}_\mathrm{i}y_i$$

$$\hat{\sigma}_n = \sum_{i=1}^{n} v_i \left(y_i - \mathbf{x}_{\mathrm{i}}^{'} \widetilde{\boldsymbol{\beta}}_{\mathrm{n}} \right)$$

Si se considera la muestra bootstrap de las observaciones $\left\{ (y_i^*, \mathbf{x}_i^*)^{'}, i = 1, \dots, n \right\}$ y se definen las variables aleatorias:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^* = \left[\sum_{i=1}^{n} \omega_i^* \mathbf{x}_{\mathrm{i}}^* \mathbf{x}_{\mathrm{i}}^{*'} \right]^{-1} \sum_{i=1}^{n} \omega_i^* \mathbf{x}_i^* y_i^*$$

$$\hat{\sigma}_n = \sum_{i=1}^{n} v_i^* \left(y_i^* - \mathbf{x}_{\mathrm{i}}^{*'} \widetilde{\boldsymbol{\beta}}_{\mathrm{n}} \right)$$

donde:

$$\omega_i^* = \rho_1^{'}(r_i^*/\hat{\sigma}_n)/r_i^*$$

$$v_i^* = \frac{\hat{\sigma}_n}{nb} \frac{\rho_0(\tilde{r}_i^*/\hat{\sigma}_n)}{\tilde{r}_i^*}$$

$$r_i^* = y_i^* - \mathbf{x}_{\mathrm{i}}^{*'} \widehat{\boldsymbol{\beta}}_{\mathrm{n}}$$

$$\tilde{r}_i^* = y_i^* - \mathbf{x}_{\mathrm{i}}^{*'} \widetilde{\boldsymbol{\beta}}_{\mathrm{n}},$$

para $i = 1, \dots, n$.

Es importante notar que los estimadores $\widehat{\boldsymbol{\beta}}_{\mathrm{n}}$, $\widetilde{\boldsymbol{\beta}}_{\mathrm{n}}$ y $\hat{\sigma}_n$ no se recalculan para cada muestra bootstrap. Por lo tanto los estimadores recalculados pueden no reflejar la variabilidad del vector aleatorio $\left(\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^{'}, \hat{\sigma}_n \right)^{'}$ debido a que los pesos ω^* y v^* son calculados utilizando estimaciones fijas. Con el fin de ajustar los resultados a causa de esta situación, se aplica una corrección lineal a los estimadores $\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^*$ y $\hat{\sigma}_n^*$ recalculados y se los combina.

Sea entonces:

$$\mathbf{M}_{\mathrm{n}} = \hat{\sigma}_n \left[\sum_{i=1}^{n} \rho_1^{''}(r_i/\hat{\sigma}_n, \mathbf{x}_{\mathrm{i}}) \mathbf{x}_{\mathrm{i}} \mathbf{x}_{\mathrm{i}}^{'} \right]^{-1} \sum_{i=1}^{n} \omega_i \mathbf{x}_{\mathrm{i}} \mathbf{x}_{\mathrm{i}}^{'}$$

$$\mathbf{d}_{\mathrm{n}} = a_n^{-1} \left[\sum_{i=1}^{n} \rho_1^{''}(r_i/\hat{\sigma}_n, \mathbf{x}_{\mathrm{i}}) \mathbf{x}_{\mathrm{i}} \mathbf{x}_{\mathrm{i}}^{'} \right]^{-1} \sum_{i=1}^{n} \rho_1^{''}(r_i/\hat{\sigma}_n, \mathbf{x}_{\mathrm{i}}) r_i \, \mathbf{x}_{\mathrm{i}}$$

$$a_n = \frac{1}{nb} \left[\sum_{i=1}^{n} \rho_0^{'}(\tilde{r}_i/\hat{\sigma}_n) \, \tilde{r}_i/\hat{\sigma}_n \right]$$

El valor recalculado de $\widehat{\boldsymbol{\beta}}_{\mathrm{n}} - \boldsymbol{\beta}$ a través del bootstrap rápido está dado por:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^{\mathrm{R}*} - \widehat{\boldsymbol{\beta}}_{\mathrm{n}} = \mathbf{M}_{\mathrm{n}}\left(\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^{*} - \widehat{\boldsymbol{\beta}}_{\mathrm{n}}\right) + \mathbf{d}_{\mathrm{n}}(\hat{\sigma}_n^{*} - \hat{\sigma}_n)$$

Es importante notar que al recalcular $\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^{\mathrm{R}*} - \widehat{\boldsymbol{\beta}}_{\mathrm{n}}$ se resuelve el sistema de ecuaciones

$$\widehat{\boldsymbol{\beta}}_{\mathrm{n}}^{*} = \left[\sum_{i=1}^{n} \omega_i^{*} \mathbf{x}_{\mathrm{i}}^{*} \mathbf{x}_{\mathrm{i}}^{*\prime}\right]^{-1} \sum_{i=1}^{n} \omega_i^{*} \mathbf{x}_i^{*} y_i^{*}$$

y se calcula el promedio ponderado

$$\hat{\sigma}_n = \sum_{i=1}^{n} v_i^{*} \left(y_i^{*} - \mathbf{x}_{\mathrm{i}}^{*\prime} \widetilde{\boldsymbol{\beta}}_{\mathrm{n}}\right)$$

Los factores de corrección $\mathbf{M}_n$, $\mathbf{d}_n$ y a_n surgen de resolver dos sistemas lineales y un promedio ponderado respectivamente, y se calculan solo una vez con la muestra completa. Para estimadores de regresión MM con ρ_0' redecendiente, los pesos ω_i brindan al método estabilidad ante la presencia de outliers. Los puntos que estén alejados estarán asociados a pesos bajos en

$$\widehat{\boldsymbol{\beta}}_n = \left[\sum_{i=1}^{n} \omega_i \mathbf{x}_{\mathrm{i}} \mathbf{x}_{\mathrm{i}}'\right]^{-1} \sum_{i=1}^{n} \omega_i \mathbf{x}_{\mathrm{i}} y_i$$

y en

$$\hat{\sigma}_n = \sum_{i=1}^{n} v_i \left(y_i - \mathbf{x}_{\mathrm{i}}' \widetilde{\boldsymbol{\beta}}_{\mathrm{n}}\right)$$

Los puntos atípicos extremos, es decir aquellos con residuos asociados muy grandes, recibirán pesos nulos y por lo tanto no tendrán efecto alguno sobre los coeficientes recalculados. Los pesos v_i utilizados en el recálculo de la escala son también decrecientes a medida que aumenta el valor absoluto de los residuos y en consecuencia las observaciones atípicas son menos influyentes también en el recalculo de $\hat{\sigma}_n^{*}$.

RESULTADOS

Con el objetivo de comparar los estimadores presentados, se realizó un estudio de simulaciones donde se generaron 1000 series de longitud (N) igual a 60, 120 y 180, de procesos AR(1) (Φ=±0,9; ±0,6; ±0,3), sin outliers y contaminadas con un outlier del tipo aditivo o innovativo de tamaño $w = 7$ en el momento $t = N/2 + 2$. Para detallar la comparación se cotejó la FAC teórica de los modelos simulados con cada una de las estimaciones propuestas. Se calcularon el sesgo y el error cuadrático medio de cada estimador sobre las 1000 replicaciones para el rezago k= 1.

En las situaciones en las que no se presentan observaciones atípicas, el estimador FRB es superior a los restantes en todos los casos, tanto en el sesgo como en el ECM, salvo en alguna situación en donde su desempeño es similar al estimador Est y MG. Estos dos últimos son los que presentan el segundo mejor rendimiento en ambos aspectos (Gráficos 1 al 4).

Gráfico Nº 1 : AR(1) $\phi = 0.9$ – Sin Outliers – Rezago=1

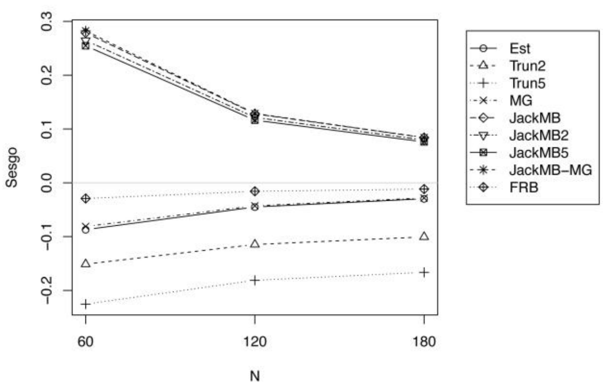


Gráfico Nº 2 : AR(1) $\phi = 0.9$ – Sin Outliers – Rezago=1

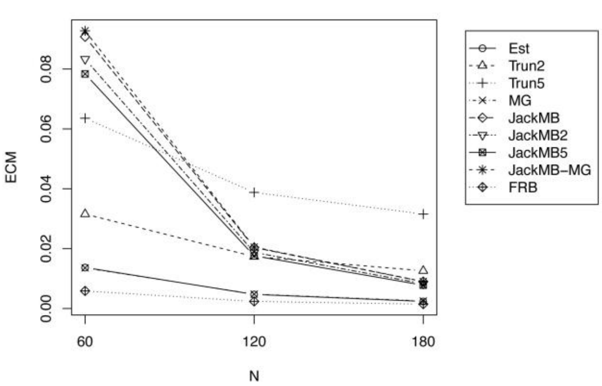


Gráfico Nº 3 : AR(1) $\phi = -0.3$ – Sin Outliers – Rezago=1

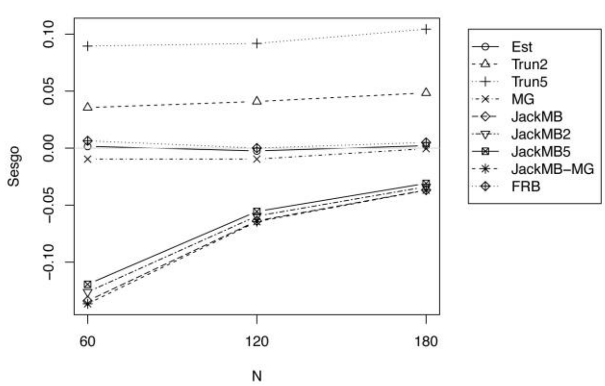
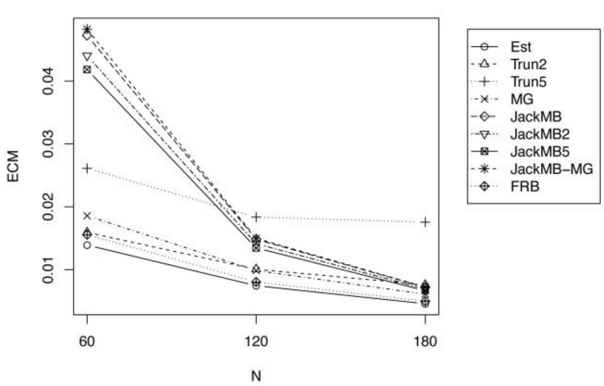
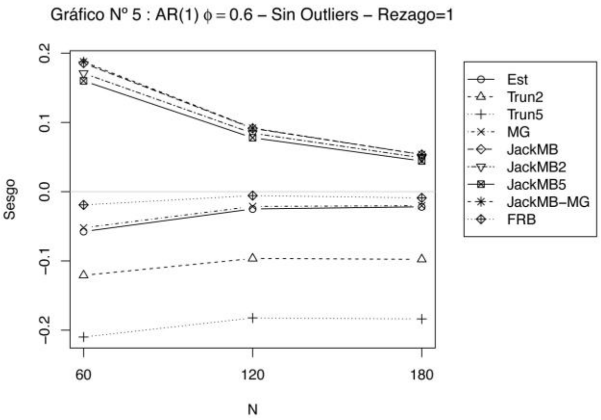
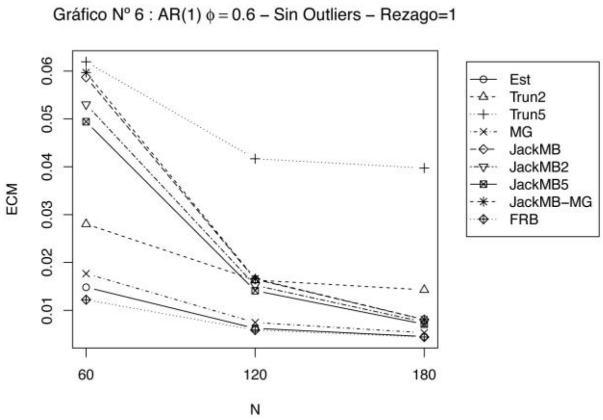


Gráfico Nº 4 : AR(1) $\phi = -0.3$ – Sin Outliers – Rezago=1

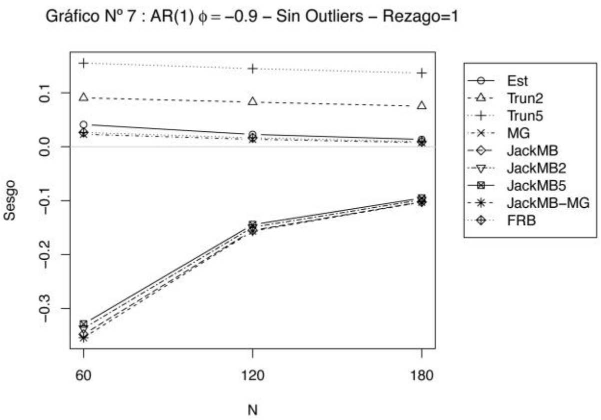


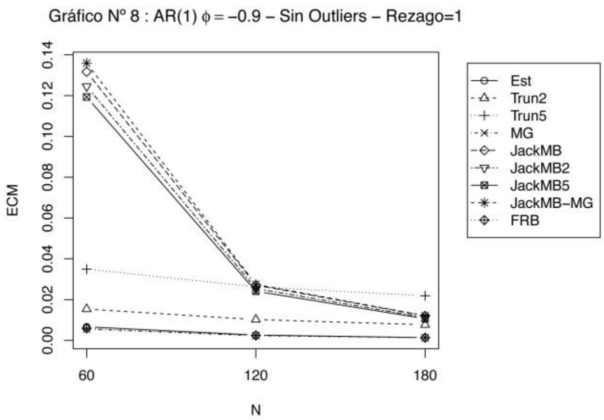
El estimador que presenta el siguiente mejor rendimiento es el Trun2. Los métodos *Jackknife* presentan comportamientos similares entre ellos y levemente inferiores al Trun2 en series cortas (n=60) pero similares para series más largas (Gráficos 5 y 6).





El Trun5 parece tener el desempeño más débil, salvo para series cortas (n=60) y con respecto al ECM, en donde en algunos casos es superior a los métodos *Jackknife* (Gráficos 7 y 8).





En las situaciones en las que se presentan observaciones atípicas (tanto aditivas como innovativas), los estimadores MG, FRB y Trun2 tienen los mejores rendimientos en comparación con el resto de los estimadores. El estimador MG y FRB tienen comportamientos similares y claramente superiores al Trun2 en cuanto a sesgo y ECM para los casos en donde $\Phi=\pm0{,}9$; $\pm0{,}6$ (Gráficos 9 y 10).

Gráfico Nº 9 : AR(1) $\phi = 0.9$ – Outlier Aditivo (w=7) – Rezago=1

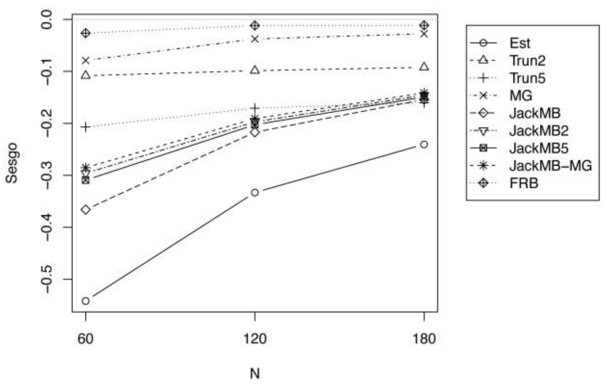
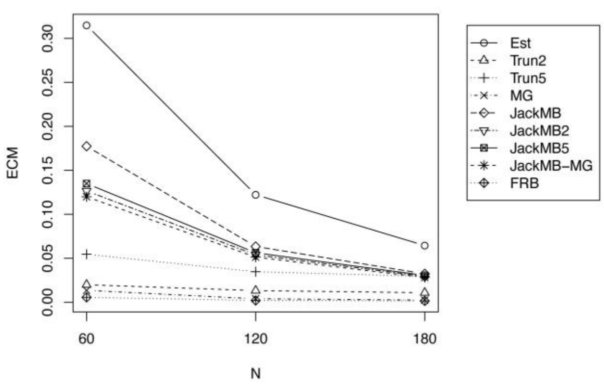


Gráfico Nº 10 : AR(1) $\phi = 0.9$ – Outlier Aditivo (w=7) – Rezago=1



En el caso donde Φ=±0,3, el rendimiento del FRB decae siendo superado por el Trun2, en especial en series cortas (n=60) o intermedias (n=120) (Gráficos 11 y 12).

Gráfico Nº 11 : AR(1) $\phi = 0.3$ – Outlier Innovativo (w=7) – Rezago=1

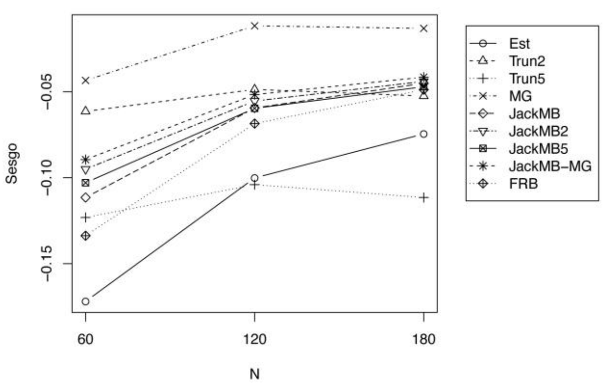
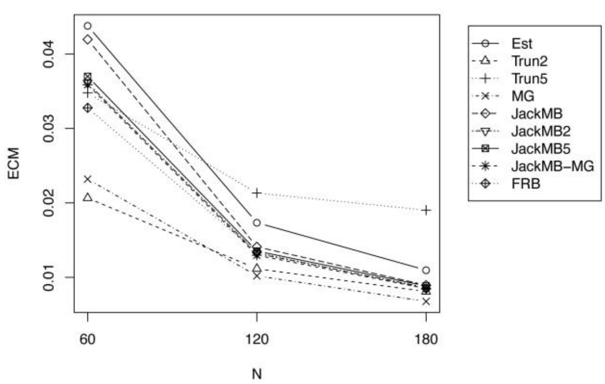


Gráfico Nº 12 : AR(1) $\phi = 0.3$ – Outlier Innovativo (w=7) – Rezago=1



Los cuatro estimadores basados en el método *Jackknife* presentan desempeños similares, pero siendo siempre mejor al compararlos entre ellos, la versión combinada con el estimador robusto MG. Son superiores al estimador Est pero de peor rendimiento comparado con el Trun2 salvo en series largas (n=180) (Gráficos 13 y 14).

Gráfico Nº 13 : AR(1) $\phi = -0.6$ – Outlier Aditivo (w=7) – Rezago=1

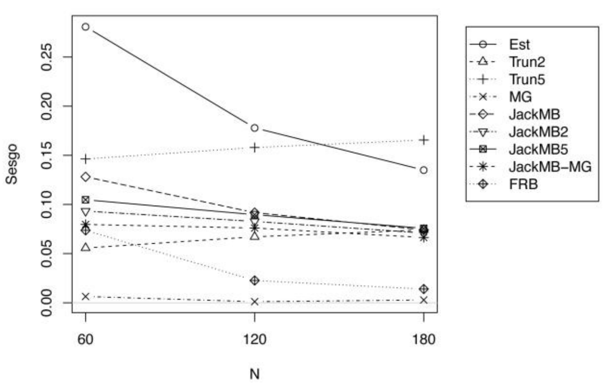
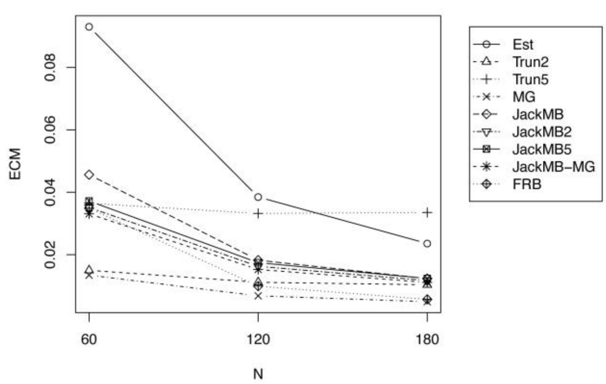


Gráfico Nº 14 : AR(1) $\phi = -0.6$ – Outlier Aditivo (w=7) – Rezago=1



El estimador Trun5 parece tener el comportamiento más pobre, e inclusive peor que el del estimador EST en series largas (n=180) (Gráficos 15 y 16).

Gráfico Nº 15 : AR(1) $\phi = 0.6$ – Outlier Innovativo (w=7) – Rezago=1

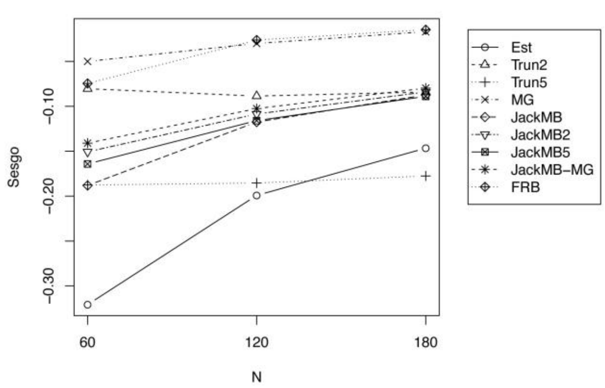
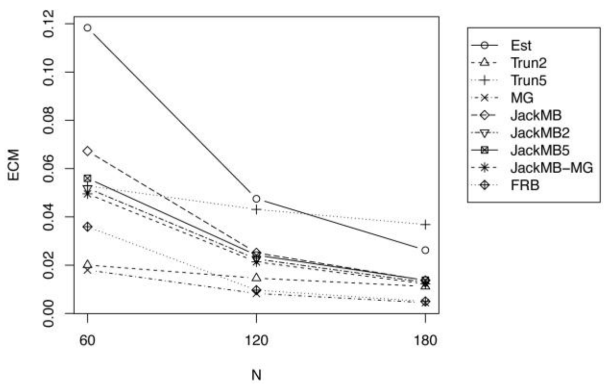


Gráfico Nº 16 : AR(1) $\phi = 0.6$ – Outlier Innovativo (w=7) – Rezago=1



CONCLUSIONES

Los métodos de replicaciones para la estimación de la FAC de un modelo de serie de tiempo AR(1) presentaron resultados diversos. El estimador basado en el método FRB parece ser un serio competidor del estimador altamente robusto MG ya que es superior en cuanto a sesgo y ECM en los casos de ausencia de outliers y resulta tener un comportamiento similar en los casos simulados con un outlier (aditivo o innovativo) siempre y cuando Φ=±0,9; ±0,6. Lo mismo ocurre si se lo compara con el estimador Trun2, el cual tiene un buen comportamiento cuando existen outliers. Para los casos en donde el valor de Φ en valor absoluto es 0,3, el desempeño del método FRB decae notablemente. Los estimadores basados en el método *Jackknife* se comportan razonablemente bien en presencia de observaciones extremas pero no logran en ningún caso superar el desempeño logrado por el estimador MG.

DISCUSIÓN

En primer resulta interesante probar que ocurriría con las estimaciones de la FAC para rezagos más altos. La estimación de la misma a través del método FRB no debería significar una adaptación demasiado complicada del método, pudiendo además obtenerse la estimación de los otros métodos de manera relativamente sencilla.

En segundo lugar el método FRB parecería perder sus buenas propiedades cuando la FAC corresponde a un modelo de serie de tiempo AR(1) con valores de Φ bajos, con lo cual se podría explorar este comportamiento para un mayor número de posibles valores del parámetro que abarquen un rango considerable.

En tercer lugar, si bien el método FRB resultaría ser un posible estimador altamente robusto alternativo a MG y superior al método *Jackknife* en cualquiera de las versiones aplicadas en este trabajo, es importante destacar que a pesar de su característica “Rápida” presente en su propia definición, los tiempos computacionales de cálculo son notoriamente superiores al de los otros métodos, que no requieren de tiempo computacional de cómputo. Este punto también es importante de notar en casos donde sea necesario computar la FAC para un gran número de series.

Es de interés además explorar y comparar el rendimiento de los métodos de replicaciones para estimar la FAC en el caso de modelos MA puros y procesos ARMA.

Por último, futuros estudios deberían estar enfocados a la evaluación de los métodos de estimación presentados ante la existencia de un mayor número de outliers.

REFERENCIAS BIBLIOGRÁFICAS

Bonifazi, F. (2015). *Estimación robusta de la función de autocorrelación*. Tesina de grado de la Licenciatura en Estadística, Facultad de Ciencias Económicas y Estadística, UNR.

Bonifazi, F.; Méndez, F. (2014). Estimación robusta de la función de autocorrelación. *Decimoctavas Jornadas "Investigaciones en la Facultad" de Ciencias Económicas y Estadística.*

Box, G. E. P.; Jenkins, D. A. (1976). *Time Series Analysis and Control*, 2nd edition. Holden-Day.

Box, G. E. P.; Pierce, D. A. (1970). Distribution of residual autocorrelations in autorregresive-integrated moving average time series models. *J. American Statistica Association*, 65, 1509-1526.

Buishand, T.A.; Beersma, J.J. (1993). Jackknife Tests for Differences in Autocorrelation between Climate Time Series. *Journal of Climate,*1, 6, 2490-2495.

Chambers, M. J. (2013), Jackknife Estimation of Stationary Autoregressive Models. *Journal of Econometrics*, 171, Issue 1, January, 142–157

Chan, W.; Wei, W. (1992), A comparison of some estimators of time series autocorrelations. *Computational Statistics & Data Analysis*, 14, 146-163.

Dürre, A.; Fried, R.; Liboschik, T. (2014). *Robust estimation of (partial) autocorrelation*. Discussion Paper.

Efron, B., Tibshirani, R. (1993)**.** *An Introduction to the Bootstrap***.** New York: Chapman & Hall.

Efron, B. (1979). Bootstrap Methods: another look at the Jackknife. *Annals of Statistics*, 7, 1-26.

Ma, Y.; Genton, M. (2000), Highly robust estimation of the autocovariance function. *Journal of Time Series Analysis*, 21, Nº6, 663-684.

Maronna, R.A.; Martin, R.D.; Yohai, V.J. (2006), *Robust Statistics: Theory and Methods*. John Wiley and Sons.

Quenouille, M.H. (1949). Approximate tests of correlation in time series. Mathematical Proceedings of the Cambridge Philosophical Society, 45, 483-484

Rousseeuw, P. J.; Croux, C. (1992), Explicit scale estimators with high breakdown point. *L1- Statistical Analysis and Related Methods,* 77-92.

Rousseeuw, P. J.; Croux, C. (1993), Alternatives to the median absolute deviation. *Journal of the American Statistical Association,* 88, 1273-83.

Salibian-Barrera, M. (2000). *Contributions to the theory of robust inference*. Ph.D. thesis, Dept. Statist., Univ. British Columbia,Vancouver.

Salibian-Barrera, M., Van Aelst S., Willems G. (2006). PCA Based on Multivariate MM-Estimators with Fast and Robust Bootstrap. *Journal of the American Statistical Association,* 101, 1198-1211.

Salibian-Barrera, M., Zamar, R. H. (2002). Bootstrapping robust estimates of regression.*The Annals of Statistics*, 30, 556-582.

Van Aelst, S., Willems, G. (2005).Multivariate regression S-estimators for robust estimation and inference.*StatisticaSinica*, 15, 981-1001.

Van Aelst, S., Willems, G. (2013). Fast and robust bootstrap for multivariate inference: The R package FRB. *Journal of Statistical Software*, 53 (3),1–32.URL: http://www.jstatsoft.org/v53/i03/.