



**Bonifazi, Fernanda**  
**Méndez Fernanda**

*Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística*

## **ESTIMACIÓN ROBUSTA DE LA PRUEBA DE *PORTMANTEAU***

### **Resumen**

En este trabajo se considera el problema de la comprobación de la idoneidad de un modelo de series de tiempo en presencia de valores atípicos. La prueba de significación conjunta de *portmanteau* se generaliza a un estadístico robusto basado en el recorte o truncamiento de observaciones extremas (Chan, 1994). Se compara el desempeño de este nuevo estimador y del estimador clásico mediante un estudio de simulaciones y una aplicación a datos reales. Los resultados indican que el nuevo estimador es preferible a la alternativa clásica cuando las observaciones están contaminadas por valores atípicos. También se dan comentarios sobre cada estimador individual.

### **Abstract**

In this paper the problem of testing the adequacy of a time series model in the presence of outliers is considered. The classical portmanteau statistic is generalized to a robust estimator based on the trimming of extreme observations (Chan, 1994). The performance of this new estimator and the classical one is compared in a simulation study and by an application to real data. The results indicate that the new estimator is preferable to the other alternatives when the observations are contaminated by outliers. Comments on each individual estimator are also given.

**Palabras claves:** valores atípicos; modelo ARMA; diagnosis del modelo; prueba de *portmanteau*; estimadores robustos.



## INTRODUCCIÓN

La estrategia de construcción de modelos para series de tiempo reales desarrollada por Box & Jenkins (1976) consta de tres etapas: identificación, estimación y diagnóstico.

La primera consiste en identificar el posible modelo que sigue la serie, lo que requiere decidir qué transformaciones aplicar para convertir la serie observada en una serie estacionaria y, además, determinar un modelo ARMA para la serie estacionaria.

Una vez seleccionado provisionalmente un modelo para la serie estacionaria, se pasa a la segunda etapa de estimación, donde los parámetros AR y MA del modelo se estiman por máxima verosimilitud y se obtienen sus errores estándar y los residuos del modelo.

La tercera etapa es la de diagnóstico, donde se verifica que el modelo proporciona un ajuste adecuado y que los supuestos básicos, implícitos en el modelo, se satisfacen; de no cumplirse los supuestos, se determinan las modificaciones necesarias y, de hecho, se repiten las etapas anteriores hasta obtener un modelo adecuado.

La etapa de diagnóstico o verificación de la metodología de Box y Jenkins tiene su origen en la idea de que todo modelo es erróneo, puesto que los modelos son representaciones simplificadas de la realidad. Lógicamente, si hay que elegir entre varios modelos, se debe elegir aquel que presente menos fallas, o poco importantes; por este motivo se pone en tela de juicio todos los posibles modelos para detectar sus fallas (que se miden como violaciones a los supuestos que fundamentan el modelo).

Uno de los estadísticos más conocidos para probar la adecuación de un modelo de series de tiempo es el estadístico de Ljung-Box o también denominada la prueba de significación conjunta de *portmanteau* (Box & Pierce, 1970; Ljung & Box, 1978). Se pueden considerar dos casos básicos de datos atípicos en series de tiempo definidos por Fox (1972): los aditivos, que son análogos a los que ocurren en datos estáticos, como por ejemplo un error de medición o de transcripción de un dato, y los innovativos, que son específicos de situaciones dinámicas. En general, los habituales procedimientos de estimación por máxima verosimilitud o mínimos cuadrados funcionan adecuadamente cuando el proceso no está contaminado por *outliers*. Sin embargo, valores atípicos aditivos dan lugar a serios problemas en la identificación y estimación del modelo. Desafortunadamente, cuando ellos están presentes, los métodos usuales de identificación, tales como la función de autocorrelación muestral (FACM) y la función de autocorrelación parcial muestral (FACPM) no son robustos, dado que la estructura de correlación es seriamente perturbada (Chan & Wei, 1992; Bonifazi & Méndez, 2014; Duerre et al, 2014; Bonifazi, 2015). Este problema podría aliviarse mediante estadísticas robustas de bondad de ajuste.

En este trabajo se presenta un estadístico robusto para la prueba de *portmanteau*, propuesto por Chan (1994), basado en el recorte o eliminación de las observaciones extremas. Su *performance* se compara, mediante un estudio de simulación bajo diversos escenarios y la aplicación con datos reales, con el desempeño del estadístico clásico de Ljung-Box.

En la sección siguiente se describen brevemente los estadísticos de la prueba de *portmanteau* utilizados; seguidamente se presenta el estudio de simulación, la aplicación con datos reales y los resultados hallados.



## METODOLOGÍA

### Diagnosís del modelo

Una de las formas más claras y simples para detectar violaciones a los supuestos de los modelos es a través del análisis de residuos. Se considera residuo a aquella parte de las observaciones que no la explica el modelo, es decir los residuos miden las discrepancias entre los valores observados y los valores estimados por el modelo. Además, cuando el tamaño de la muestra es grande, los errores aleatorios y los residuos son esencialmente iguales; por esta razón, al analizar los residuos observados,  $\{\hat{a}_t\}$ , se analiza básicamente lo que debería ser una realización del proceso de ruido blanco,  $\{a_t\}$ .

La diagnosis del modelo requiere comprobar que las hipótesis básicas realizadas respecto a los residuos son ciertas. Estos deben tener: a) media marginal igual a cero, b) variancia marginal constante, c) falta de correlación para cualquier retardo y d) distribución normal.

De las cuatros condiciones establecidas en la etapa de diagnosis en relación a las distribuciones marginales se puede observar que la a) es poco restrictiva. Es posible que el modelo sea muy incorrecto, y sin embargo se verifique a). Por otro lado esta condición puede no verificarse si existen valores atípicos. La condición b) de variancia marginal constante es más fuerte. La condición c), falta de incorrelación para cualquier retardo, es central para asegurar que el modelo es adecuado. Finalmente, la condición de normalidad d) es conveniente, porque nos garantiza que la incorrelación implica la independencia, y que no estamos dejando información por modelizar.

La hipótesis de falta de correlación o que los residuos provienen de un modelo ruido blanco puede verificarse, primero, calculando la FACM de los residuos mediante:

$$\hat{\rho}_k(\hat{a}) = \frac{\sum_{t=1}^{n-k} (\hat{a}_t - \bar{a})(\hat{a}_{t+k} - \bar{a})}{\sum_{t=1}^n (\hat{a}_t - \bar{a})^2}, k = 1, 2, \dots, (n - 1) \quad (1)$$

donde  $\hat{a}_t$  son los residuos estimados y  $\bar{a}$  es la media de los  $n$  residuos. Si los residuos son independientes, los coeficientes  $\hat{\rho}_k$ , para  $k$  no muy pequeño, se distribuyen aproximadamente normal con media cero y variancia asintótica constante,  $1/n$ . La variancia asintótica es válida para  $k$  grande, pero no para los primeros retardos. Para los primeros rezagos la variancia puede ser mucho menor que  $1/n$ , por tanto, este valor debe considerarse como un límite máximo de las variancias de las autocorrelaciones residuales.

El procedimiento habitual de verificar incorrelación de los residuos es dibujar dos líneas paralelas a distancia  $2/\sqrt{n}$  del origen, y comprobar si todos los coeficientes  $\hat{\rho}_k$  están dentro de estos límites de confianza.

No obstante esta prueba de significación no es del todo válida para autocorrelaciones correspondientes a retrasos pequeños ( $k \leq 3$ ) y conviene por ello realizar una prueba conjunta simultánea de la significación de las primeras  $K$  autocorrelaciones, como se indica a continuación.



### **Contraste de Ljung-Box clásico**

Un contraste global de que los primeros  $K$  coeficientes son cero es el contraste de Ljung-Box (1978), también conocido como prueba de significación de *portmanteau*. Se postula la hipótesis nula conjunta:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0, \quad (2)$$

con el estadístico

$$Q_{LB}(K) = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2(\hat{a})}{n-k} \quad (3)$$

donde  $\hat{\rho}_k(\hat{a})$  es la autocorrelación muestral de los residuos de orden  $k$ .

Este estadístico de prueba es una modificación del originalmente propuesto por Box & Pierce (1970)<sup>1</sup>, que resulta ser más adecuado para realizar el contraste.

Bajo la hipótesis nula que el modelo ARMA es adecuado,  $Q_{LB}$  se distribuye asintóticamente como una  $\chi^2$  con grados de libertad igual al número de coeficientes en la suma ( $K$ ) menos el número de parámetros estimados,  $m$ . Para modelos no estacionales ARMA( $p, d, q$ )( $P, D, Q$ )<sub>s</sub>,  $m = p + q + 1$  o  $m = p + q$ , según el modelo tenga o no constante; y para los estacionales sin constante  $m = P + p + Q + q$ . Se concluye que el modelo es inadecuado si el valor obtenido de  $Q_{LB}(K)$  con (3) es mayor que el percentil  $(1 - \alpha)$  de la distribución  $\chi^2$  con  $(K - m)$  grados de libertad,  $\chi^2_{(1-\alpha);(K-m)}$ . En general, se rechaza la hipótesis de incorrelación de los residuos cuando la probabilidad  $P(\chi^2_{(K-m)} > Q_{LB}(K))$  sea pequeña (menor que  $\alpha$ ).

### **Estadística de portmanteau truncada o recortada**

Una primera idea intuitiva para limitar la influencia de *outliers* es descartar o al menos asignar baja ponderación a aquellos valores extremos grandes o pequeños de la serie temporal, donde los valores atípicos se determinan a partir de la distribución marginal de la variable, haciendo caso omiso de la dependencia de serie.

Chan (1994) propuso un nuevo método para la prueba de *portmanteau* basado en el truncamiento, debido a que el recorte o truncamiento es un método muy simple y eficiente para incrementar la resistencia de una estadística frente a la presencia de valores atípicos o extremos.

El test de *portmanteau* truncado se obtiene de forma similar al estadístico introducido por Box & Pierce (1970). En la estadística truncada las autocorrelaciones residuales son obtenidas como una extensión de la función de autocorrelación muestral  $\alpha$ -truncada propuesta por Chan & Wei (1992).

---

<sup>1</sup> Estadístico  $Q$  de Box & Pierce (1970):  $Q_{BP}(K) = n \sum_{k=1}^K \hat{\rho}_k^2(\hat{a})$ . Conviene hacer notar que en ocasiones, la aproximación lograda con  $Q_{BP}$  no es del todo apropiada. Por esta razón, Ljung & Box (1978) modificaron esta expresión para obtener el estadístico que resulta ser más adecuado para realizar la prueba.



Sea  $\hat{a}_{(p+1)} \leq \hat{a}_{(p+2)} \leq \dots \leq \hat{a}_{(n)}$  los residuos ordenados de un modelo ARMA( $p, q$ ) estimado, se define la función de autocorrelación residual truncada como:

$$\hat{\rho}_{T,\alpha}(k) = \frac{\hat{\gamma}_{T,\alpha}(k)}{\hat{\gamma}_{T,\alpha}(0)} \quad (4)$$

donde

$$\hat{\gamma}_{T,\alpha}(k) = \frac{1}{\sum_{t=p+k+1}^n L_{t-k}^{(\alpha)} L_t^{(\alpha)}} \left\{ \sum_{t=p+k+1}^n \hat{a}_{t-k} \hat{a}_t L_{t-k}^{(\alpha)} L_t^{(\alpha)} \right\} \quad (5)$$

y

$$L_t^{(\alpha)} = \begin{cases} 0 & \text{si } \hat{a}_t \leq \hat{a}_{(g)} \text{ o } \hat{a}_t \geq \hat{a}_{(n-g+1)}, \\ 1 & \text{en otro caso,} \end{cases}$$

para  $p+1 \leq t \leq n$  y donde  $g$  es la parte entera de  $[\alpha n]$  y  $0 \leq \alpha < 0,5$ .

Se define

$$C_L(k) = \frac{1}{n} \sum_{t=p+k+1}^n L_{t-k}^{(\alpha)} L_t^{(\alpha)} \quad (6)$$

y se asume que su límite es  $v_k$ ,  $\lim_{n \rightarrow \infty} C_L(k) = v_k$ , y existe para  $k$  finito.

Sea

$$Q_{T,\alpha}(K) = \sum_{k=1}^K (n v_k) \left( \hat{\rho}_{T,\alpha}(k) \right)^2. \quad (7)$$

$v_k$  es desconocido y puede ser reemplazado por:

$$\hat{v}_k = \frac{1}{n} \sum_{t=p+k+1}^n L_{t-k}^{(\alpha)} L_t^{(\alpha)}. \quad (8)$$

Luego,  $Q_{T,\alpha}(K)$  se distribuye asintóticamente como una chi-cuadrado con  $(K - m)$  grados de libertad.

$$Q_{T,\alpha}(K) \xrightarrow{d} \chi_{(K-m)}^2 \quad (9)$$



## APLICACIÓN

### Simulaciones

Se estudia la robustez y el desempeño de las estadísticas propuestas mediante un estudio de simulación. En una primera instancia, se presentan los resultados de las simulaciones de procesos ARMA no contaminados con valores atípicos. Luego se evalúa el desempeño de las estadísticas en diferentes procesos ARMA contaminados por *outliers* aditivos o innovativos. Se utiliza el programa estadístico R.

#### Escenario 1: serie no contaminada con datos atípicos

Se generan 1000 muestras de procesos ARMA(1,1) de tamaño  $n = 20, 50$  y  $100$  con parámetros  $(\phi, \theta) = (0.5, 0), (-0.5, 0), (0.8, 0), (0, 0.8)$  y  $(0.8, 0.2)$ . Se calculan el promedio y la variancia para cada estimador sobre las 1000 replicaciones, como también los percentiles del 5% y 10% superiores de la distribución muestral, considerando el contraste  $H_0: \rho_1 = \dots = \rho_K = 0$ , para  $K = 6, 12, 18$  y  $24$ . Se considera  $\alpha = 2\%$  y  $5\%$  en el cálculo de la estimación de la FACMT. Los resultados se presentan en la Tabla 1. En la misma sólo se muestran algunas de las combinaciones de los valores de los parámetros puesto que los resultados son muy similares; los restantes resultados se muestran en la Tabla A.1 en el Anexo A.

Cuando no existen valores atípicos, la distribución muestral de las 1000 replicaciones de la estadística de Ljung-Box se aproxima a una distribución  $\chi^2$  con  $(K - m)$  grados de libertad (Figura 2). Por tanto, la esperanza estimada o el valor promedio de la misma se aproxima a  $(K - m)$  en todos los procesos simulados y su variancia toma valores cercanos a  $2 * (K - m)$ . También, se puede observar que los percentiles superiores del 5 y 10% de la distribución muestral son similares a los obtenidos a partir de la distribución teórica. Al analizar la distribución muestral de la estadística de bondad de ajuste mediante el método truncado, se observan comportamientos similares a los obtenidos a partir del método clásico, sin notar grandes discrepancias entre sus estimaciones.

En términos generales, se puede apreciar que ambas estadísticas tienen una buena *performance* en situaciones no afectadas por valores atípicos.



Tabla 1: Medidas descriptivas de la estadística de bondad de ajuste basada en 1000 repeticiones de tamaño 20, 50 y 100 para procesos ARMA(1,1)

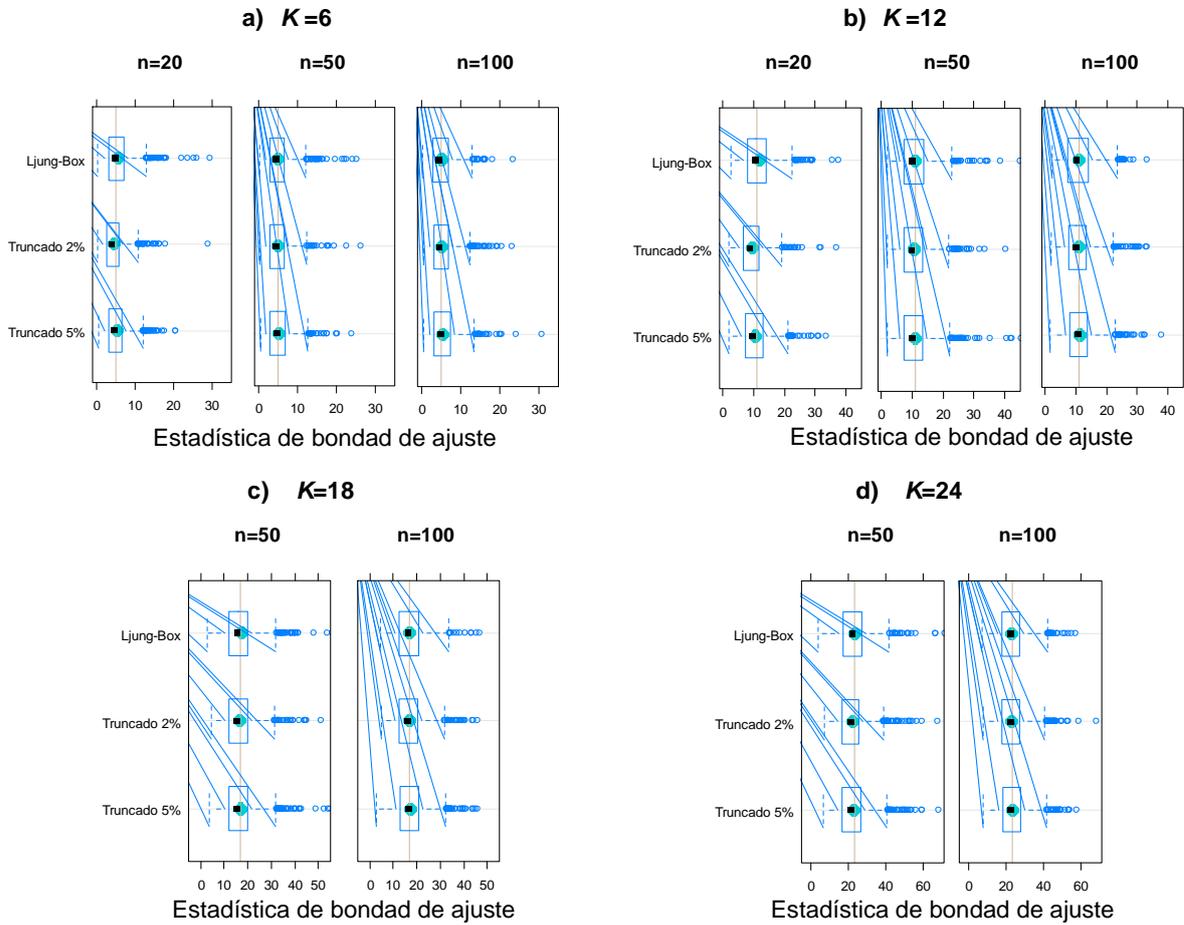
Estadística de bondad de ajuste		K= 6			12			18			24		
		n= 20	50	100	20	50	100	20	50	100	20	50	100
<b>a) <math>(\phi, \theta) = (0.5, 0)</math></b>													
Teórico	Promedio	5			11			17			23		
	Variación	10			22			34			46		
	10%	9,24			17,28			24,77			32,01		
	5%	11,07			19,68			27,59			35,17		
Estimador Ljung-Box	Promedio	5,5	5,1	5,1	11,6	11,1	11,0	-	17,1	17,0	-	23,1	22,9
	Variación	13,4	10,9	9,8	28,6	27,4	24,1	-	50,1	40,0	-	70,5	57,9
	10%	10,20	9,17	9,56	18,17	18,27	17,52	-	26,94	25,37	-	33,74	33,00
	5%	12,03	11,45	11,05	20,89	20,98	20,55	-	30,17	28,27	-	38,78	36,92
Estimador truncado $\alpha=2\%$	Promedio	4,4	5,1	5,3	9,5	10,7	11,0	-	16,5	16,9	-	22,1	22,8
	Variación	8,1	9,2	10,2	17,4	22,4	24,8	-	41,7	38,6	-	61,4	54,8
	10%	8,06	9,21	9,61	14,85	16,78	17,72	-	24,54	25,03	-	31,52	31,58
	5%	9,79	10,76	11,67	16,97	19,45	20,48	-	27,64	27,23	-	34,96	35,66
Estimador truncado $\alpha=5\%$	Promedio	5,2	5,3	5,6	10,6	11,1	11,4	-	16,7	17,4	-	22,4	23,4
	Variación	8,9	9,9	11,0	20,2	27,8	24,8	-	48,2	41,8	-	68,9	59,9
	10%	9,30	9,39	9,59	16,78	17,55	18,08	-	25,26	25,98	-	32,25	33,83
	5%	11,07	11,30	11,49	19,34	20,82	20,93	-	29,24	29,17	-	37,45	37,38
<b>b) <math>(\phi, \theta) = (0, 0,8)</math></b>													
Teórico	Promedio	5			11			17			23		
	Variación	10			22			34			46		
	10%	9,24			17,28			24,77			32,01		
	5%	11,07			19,68			27,59			35,17		
Estimador Ljung-Box	Promedio	5,5	5,1	5,2	11,6	10,9	11,2	-	16,9	17,2	-	22,6	23,1
	Variación	11,5	9,3	11,0	24,9	25,4	27,8	-	46,0	44,4	-	63,5	66,6
	10%	9,60	9,10	9,53	18,10	17,57	17,93	-	25,32	26,00	-	32,76	34,02
	5%	11,91	10,47	11,78	20,64	20,47	21,28	-	29,73	30,19	-	38,41	38,80
Estimador truncado $\alpha=2\%$	Promedio	4,8	5,2	5,5	10,1	10,8	11,4	-	16,5	17,3	-	22,0	23,3
	Variación	8,7	9,8	10,4	19,0	25,4	26,6	-	44,5	44,9	-	61,5	68,0
	10%	8,38	9,42	9,87	15,70	17,69	17,68	-	25,18	25,23	-	32,47	33,08
	5%	10,33	10,79	11,48	18,13	20,42	20,29	-	29,46	29,68	-	36,34	37,25
Estimador truncado $\alpha=5\%$	Promedio	5,5	5,6	5,8	11,0	11,5	11,8	-	17,3	17,8	-	23,0	24,0
	Variación	11,8	11,7	12,1	24,7	27,9	30,4	-	48,7	55,4	-	71,9	78,8
	10%	10,00	10,02	10,16	17,53	18,15	18,36	-	26,02	26,29	-	33,48	34,60
	5%	12,57	12,15	11,95	20,72	21,07	21,21	-	29,61	30,47	-	38,49	39,57
<b>c) <math>(\phi, \theta) = (0.8, 0,2)</math></b>													
Teórico	Promedio	4			10			16			22		
	Variación	8			20			32			44		
	10%	7,78			15,99			23,54			30,81		
	5%	9,49			18,31			26,30			33,92		
Estimador Ljung-Box	Promedio	4,5	4,2	4,2	10,4	9,9	10,3	-	15,5	16,4	-	21,4	22,3
	Variación	7,8	6,9	7,4	19,6	18,7	21,9	-	33,4	39,1	-	52,1	59,2
	10%	8,08	7,61	7,65	15,74	15,83	16,37	-	23,05	24,36	-	30,83	31,81
	5%	9,69	9,19	9,27	18,65	18,19	18,12	-	26,46	27,68	-	34,84	36,46
Estimador truncado $\alpha=2\%$	Promedio	3,6	4,5	4,8	8,5	9,9	10,8	-	15,4	16,7	-	21,0	22,6
	Variación	4,9	7,9	8,0	11,8	20,9	22,3	-	35,3	38,3	-	52,5	59,1
	10%	6,51	8,18	8,74	12,75	15,55	16,65	-	23,21	24,68	-	30,18	31,98
	5%	7,58	9,25	10,08	14,49	18,55	19,36	-	26,43	28,02	-	34,61	36,86
Estimador truncado $\alpha=5\%$	Promedio	4,8	5,0	5,4	10,0	10,6	11,3	-	16,2	17,3	-	21,8	23,3
	Variación	7,5	9,5	10,5	18,0	24,3	25,1	-	41,9	42,2	-	64,0	67,6
	10%	8,39	9,01	9,61	15,17	16,63	17,71	-	24,07	25,87	-	31,95	33,82
	5%	10,26	10,63	11,41	17,77	19,57	20,95	-	28,43	28,47	-	36,92	38,11

\*Para muestras chicas no es posible realizar el cálculo de autocorrelaciones para rezagos grandes, por tanto no es posible el cálculo de la estadística de bondad de ajuste para K grande (K=18 y 24).

Fuente: Elaboración propia a partir de datos simulados.



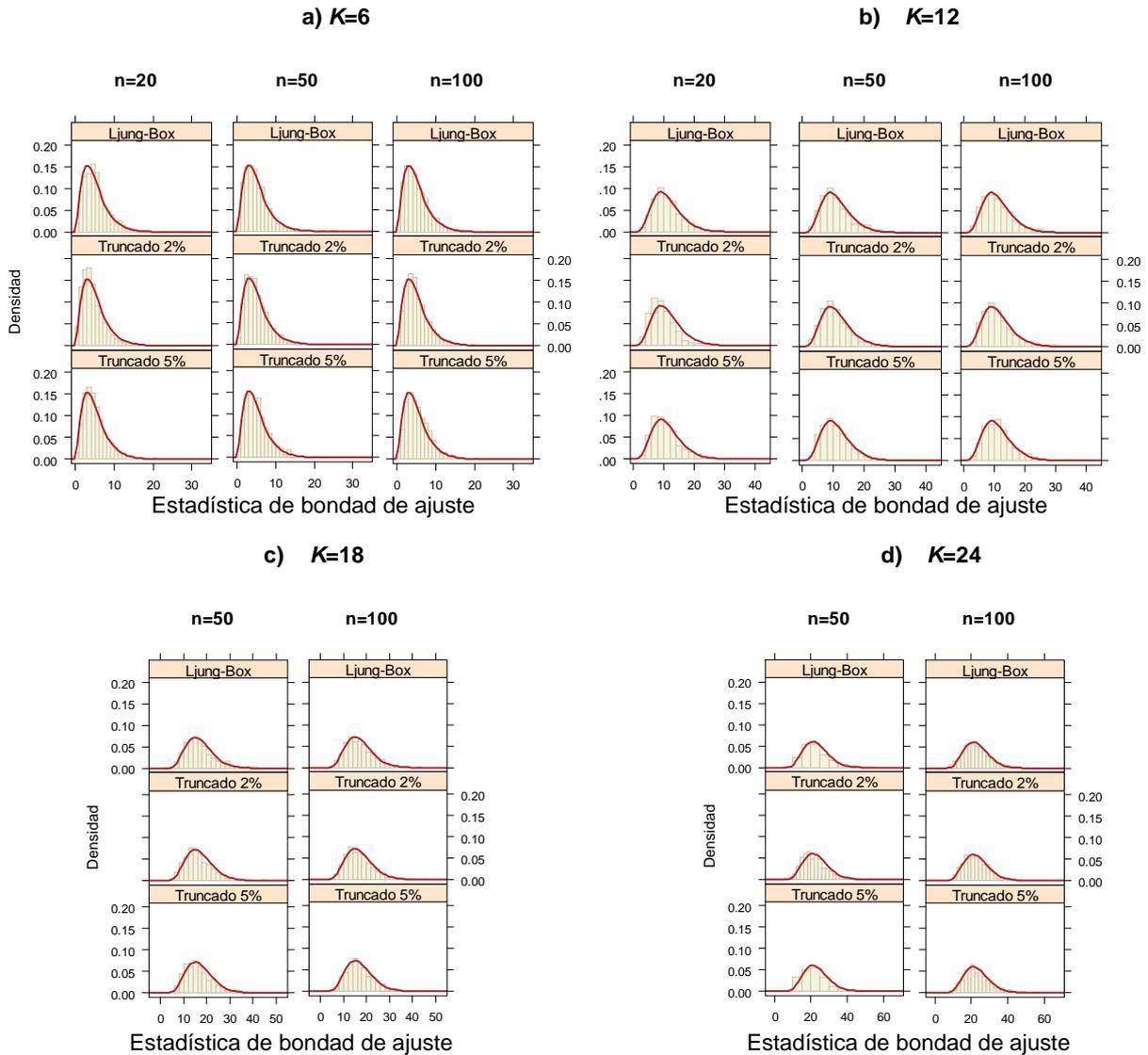
Figura 1: Diagramas de caja de la estadística de bondad de ajuste para las 1000 muestras de tamaño 20, 50 y 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$



Fuente: Elaboración propia a partir de datos simulados.



Figura 2: Histogramas de frecuencias de la estadística de bondad de ajuste para las 1000 muestras de tamaño 20, 50 y 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$



Fuente: Elaboración propia a partir de datos simulados.



## Escenario 2: serie contaminada por un dato atípico

En esta sección se consideran los efectos de la presencia de datos atípicos aditivos<sup>2</sup> (OA) o innovativos<sup>3</sup> (OI) en la estimación de la función de autocorrelación del proceso subyacente libre de *outliers*. Las series temporales se generan a partir de procesos ARMA(1,1) como en el escenario 1. Se usan las mismas seis combinaciones de valores de parámetros con tamaño de muestra  $n = 100$ . El valor extremo se encuentra en el instante  $t = 64$  con magnitudes  $\omega_A = \omega_I = sd^4$ ,  $2 \times sd$ ,  $7 \times sd$  y  $10 \times sd$ . El experimento se repite 1000 veces y el promedio y variancia, así como los percentiles del 5% y 10% superiores de la distribución muestral de cada estimador, se presentan en las Tablas 2 y 3.

Cuando la serie está afectada por un *outlier* innovativo, existen otros tipos de inconvenientes que deben ser abordados con mayor detalle: la estimación de los parámetros mediante el método clásico de máxima verosimilitud para los procesos MA y ARMA no converge ante la presencia de un OI de gran magnitud. Dado este inconveniente, en la Tabla 3 sólo se presentan los resultados de la estadística de bondad de ajuste para los procesos autorregresivos,  $(\phi, \theta) = (0.5, 0)$ ,  $(-0.5, 0)$  y  $(0.8, 0)$ .

Los resultados de las restantes combinaciones de valores de parámetros frente a la presencia de un OA se muestran en la Tabla A.2 del anexo A, ya que son similares a los mostrados.

Se puede apreciar que la estadística  $Q_{LB}(K)$  parece ser sustancialmente distorsionada en presencia de datos atípicos del tipo aditivo o innovativo, no siendo el caso de la estadística robusta  $Q_{T,\alpha}(K)$ . Cuando el efecto del *outlier* es pequeño, en general las distribuciones muestrales de ambas estadísticas estudiadas se aproximan a sus distribuciones teóricas (Figura 4 y Figura 6). Sin embargo, cuando el efecto del atípico es grande, la distribución de la estadística clásica se ve seriamente afectada pudiendo conducir a conclusiones erróneas en la prueba de bondad de ajuste del modelo. En relación al estimador truncado, en general, se observa que la variancia de la estadística de bondad de ajuste es sobrestimada para valores extremos grandes, disminuyendo a medida que el efecto del atípico disminuye. Además, no se aprecian grandes discrepancias para los diferentes valores de  $\alpha$  seleccionados.

En conclusión,  $Q_{T,\alpha}(K)$  es preferible a la estadística de bondad de ajuste clásica cuando la serie está contaminada por un valor atípico del tipo aditivo o innovativo.

---

<sup>2</sup>Se dice que ha ocurrido un atípico aditivo (OA) sobre una serie temporal en el instante  $T$  si el valor de la serie se genera en ese instante de manera distinta al resto, con una magnitud igual a  $\omega_A$ .

<sup>3</sup> Se dice que ha ocurrido un atípico innovativo (OI) sobre una serie temporal en el instante  $T$  si el valor de la serie a partir de ese instante es afectado a través de la memoria del sistema descrita por  $\frac{\theta(B)}{\phi(B)}$ , con una magnitud  $\omega_I$ .

<sup>4</sup> $sd$  es el desvío estándar teórico del proceso ARMA (1,1).



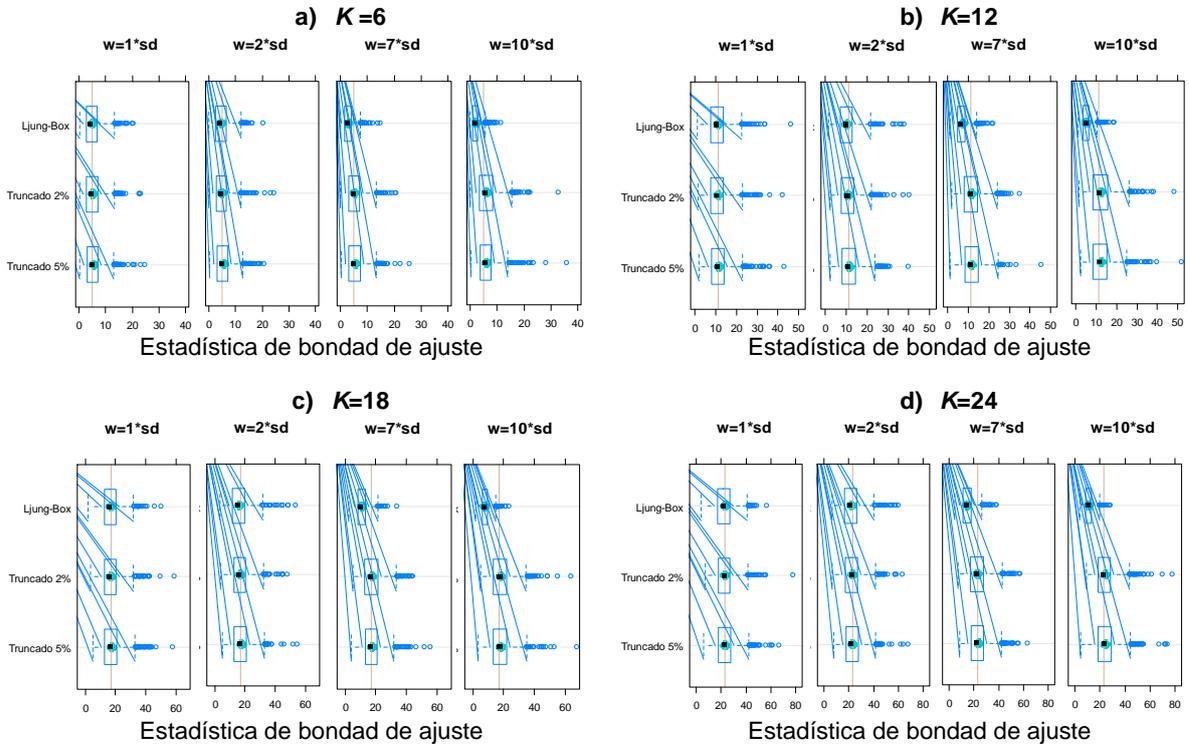
Tabla 2: Medidas descriptivas de la estadística de bondad de ajuste basada en 1000 repeticiones de procesos ARMA(1,1) de tamaño 100 con un OA en  $t=64$

Estadística de bondad de ajuste	K=6				12				18				24				
	$\omega=sd$	2*sd	7*sd	10*sd	$\omega=sd$	2*sd	7*sd	10*sd	$\omega=sd$	2*sd	7*sd	10*sd	w=sd	2*sd	7*sd	10*sd	
<b>a) <math>(\phi, \theta) = (0.5, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variancia	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,2	4,9	3,1	2,3	11,1	10,5	6,9	5,1	16,9	16,3	10,9	8,1	22,9	22,2	14,9	11,1
	Variancia	10,6	8,6	3,7	2,2	25,3	23,9	9,6	5,4	42,0	42,5	16,2	8,9	62,4	61,5	23,5	13,2
	10%	9,48	8,87	5,56	4,07	17,66	16,78	11,02	8,00	25,06	24,60	16,12	11,86	32,40	31,96	21,07	15,64
	5%	11,51	10,25	6,67	4,93	20,26	19,56	12,50	9,09	28,13	28,10	18,37	13,58	36,12	35,85	23,64	18,03
Estimador truncado $\alpha=2\%$	Promedio	5,4	5,2	5,5	6,3	11,1	11,0	11,6	12,3	17,1	16,9	17,5	18,3	23,1	23,0	23,2	24,3
	Variancia	11,0	9,5	10,3	14,7	25,6	23,1	25,6	31,8	42,8	39,9	44,5	50,4	60,0	60,8	64,5	73,6
	10%	9,70	9,13	9,67	11,39	17,46	17,27	18,39	19,59	24,64	25,07	26,65	27,50	32,18	33,17	34,15	35,41
	5%	11,73	10,96	11,79	14,10	20,73	19,86	21,21	22,04	29,49	28,08	29,76	31,68	36,35	36,57	38,22	39,74
Estimador truncado $\alpha=5\%$	Promedio	5,7	5,6	5,7	6,3	11,6	11,5	11,7	12,3	17,6	17,4	17,6	18,3	23,4	23,4	23,6	24,5
	Variancia	11,2	10,0	10,7	14,4	26,2	25,3	24,2	31,2	44,3	42,1	42,3	48,5	61,6	63,2	63,1	74,3
	10%	9,86	9,83	10,09	10,99	18,19	17,96	18,25	19,38	25,97	26,04	26,29	27,17	33,80	34,33	33,36	34,77
	5%	12,36	11,70	11,59	13,20	20,96	21,00	20,41	22,34	29,68	29,63	29,14	30,77	37,25	38,15	38,15	39,50
<b>b) <math>(\phi, \theta) = (0, 0,8)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variancia	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,2	4,8	2,9	2,1	11,0	10,3	6,5	4,6	16,8	16,0	10,0	7,3	22,7	21,7	13,7	10,0
	Variancia	10,3	9,0	3,3	1,6	22,9	20,7	8,2	4,2	37,5	36,2	13,7	7,1	55,2	52,9	20,5	10,1
	10%	9,53	8,89	5,23	3,71	16,90	16,40	10,24	7,28	24,33	23,89	14,77	10,60	32,69	31,16	19,52	14,03
	5%	11,21	10,69	6,13	4,57	19,73	19,19	11,50	8,46	28,51	26,98	16,92	12,10	36,22	35,18	21,89	15,55
Estimador truncado $\alpha=2\%$	Promedio	5,6	5,4	5,9	6,5	11,4	11,2	11,7	12,5	17,3	17,3	17,4	18,4	23,2	23,2	23,2	24,5
	Variancia	11,9	11,1	12,8	19,2	28,2	25,8	25,9	35,4	46,0	44,9	42,7	50,3	64,6	63,1	61,8	72,3
	10%	10,17	9,83	10,52	11,48	17,84	18,03	18,46	19,68	25,76	25,84	25,99	27,46	33,89	33,15	32,62	34,80
	5%	12,16	11,96	12,41	14,66	21,19	20,79	21,46	22,97	29,58	29,32	29,50	31,11	37,38	39,32	36,26	41,25
Estimador truncado $\alpha=5\%$	Promedio	5,7	5,8	6,0	6,3	11,6	11,7	11,9	12,4	17,6	17,8	17,8	18,6	23,5	23,8	23,9	24,6
	Variancia	11,6	13,3	13,9	17,2	26,0	28,5	28,6	36,2	43,3	47,0	48,8	57,1	64,2	68,7	75,4	76,6
	10%	10,36	10,34	10,28	11,06	18,58	18,33	18,43	20,06	26,18	25,87	26,35	27,83	33,36	34,07	33,76	36,12
	5%	12,43	12,54	12,50	13,82	21,72	21,28	21,06	23,48	29,72	29,10	29,49	33,45	38,16	37,96	38,10	41,63
<b>c) <math>(\phi, \theta) = (0,8, 0,2)</math></b>																	
Teórico	Promedio	4				10				16				22			
	Variancia	8				20				32				44			
	10%	7,78				15,99				23,54				30,81			
	5%	9,49				18,31				26,30				33,92			
Estimador Ljung-Box	Promedio	4,1	3,9	2,3	1,5	10,1	9,6	5,5	3,7	16,0	15,1	8,8	6,1	21,6	20,9	12,2	8,6
	Variancia	6,7	6,8	2,5	0,9	18,8	19,6	6,6	2,6	31,9	36,0	10,7	4,8	49,4	54,6	16,0	7,4
	10%	7,31	7,54	4,32	2,76	15,97	15,30	8,94	5,75	23,39	22,30	12,96	8,96	31,22	30,20	17,07	12,19
	5%	9,04	9,10	5,53	3,41	17,97	17,58	10,21	6,53	26,09	26,35	14,94	10,07	33,75	34,55	19,77	13,66
Estimador truncado $\alpha=2\%$	Promedio	4,7	4,7	6,0	7,2	10,4	10,6	11,7	13,3	16,1	16,4	17,4	19,4	21,8	22,3	23,1	25,4
	Variancia	8,4	8,6	13,8	22,9	20,4	24,5	28,2	42,4	34,2	41,7	44,2	65,0	52,2	58,7	63,2	88,4
	10%	8,43	8,25	10,42	12,80	16,45	16,74	18,43	21,19	23,33	24,36	26,45	29,26	30,58	32,26	33,85	36,80
	5%	10,23	10,13	12,87	15,75	18,40	19,49	21,62	25,02	26,05	27,59	29,85	33,06	34,85	36,10	38,27	42,27
Estimador truncado $\alpha=5\%$	Promedio	5,3	5,2	6,0	6,7	11,1	11,0	11,8	12,7	16,9	16,8	17,8	18,8	22,7	22,6	23,7	24,7
	Variancia	9,6	9,3	13,0	19,3	22,4	22,9	25,8	37,1	37,8	40,2	46,7	60,9	56,4	56,6	65,6	86,3
	10%	9,41	9,14	10,71	11,91	17,55	17,21	18,17	20,45	25,24	25,33	26,44	28,27	32,52	32,45	33,47	36,38
	5%	11,36	11,19	13,47	14,80	19,70	19,64	21,40	23,23	27,47	28,75	29,83	32,85	36,10	36,63	37,39	41,34

Fuente: Elaboración propia a partir de datos simulados.



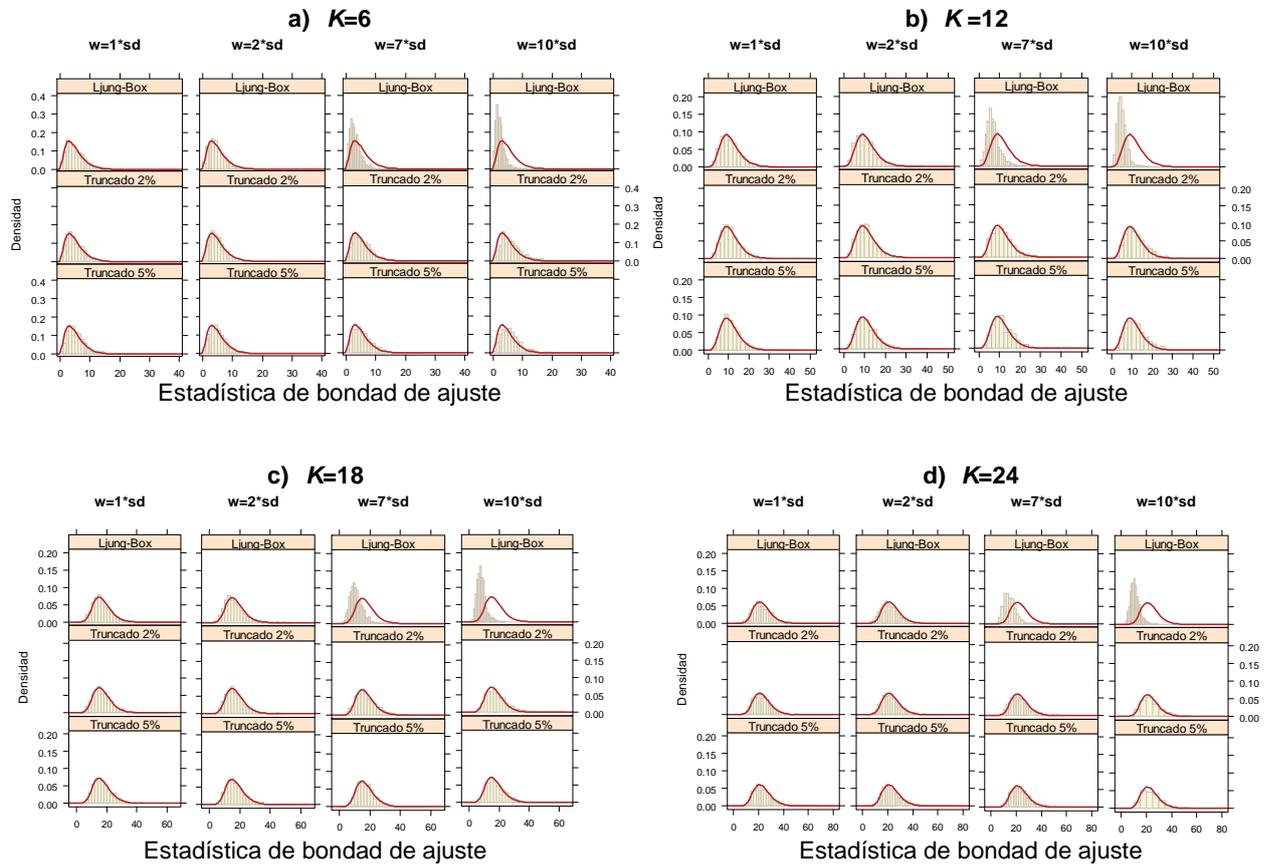
Figura 3: Diagramas de caja de la estadística de bondad de ajuste para las 1000 muestras de tamaño 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$  y un OA en  $t=64$



Fuente: Elaboración propia a partir de datos simulados.



Figura 4: Histogramas de frecuencias de la estadística de bondad de ajuste para las 1000 muestras de tamaño 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$  y un OA en  $t=64$



Fuente: Elaboración propia a partir de datos simulados.



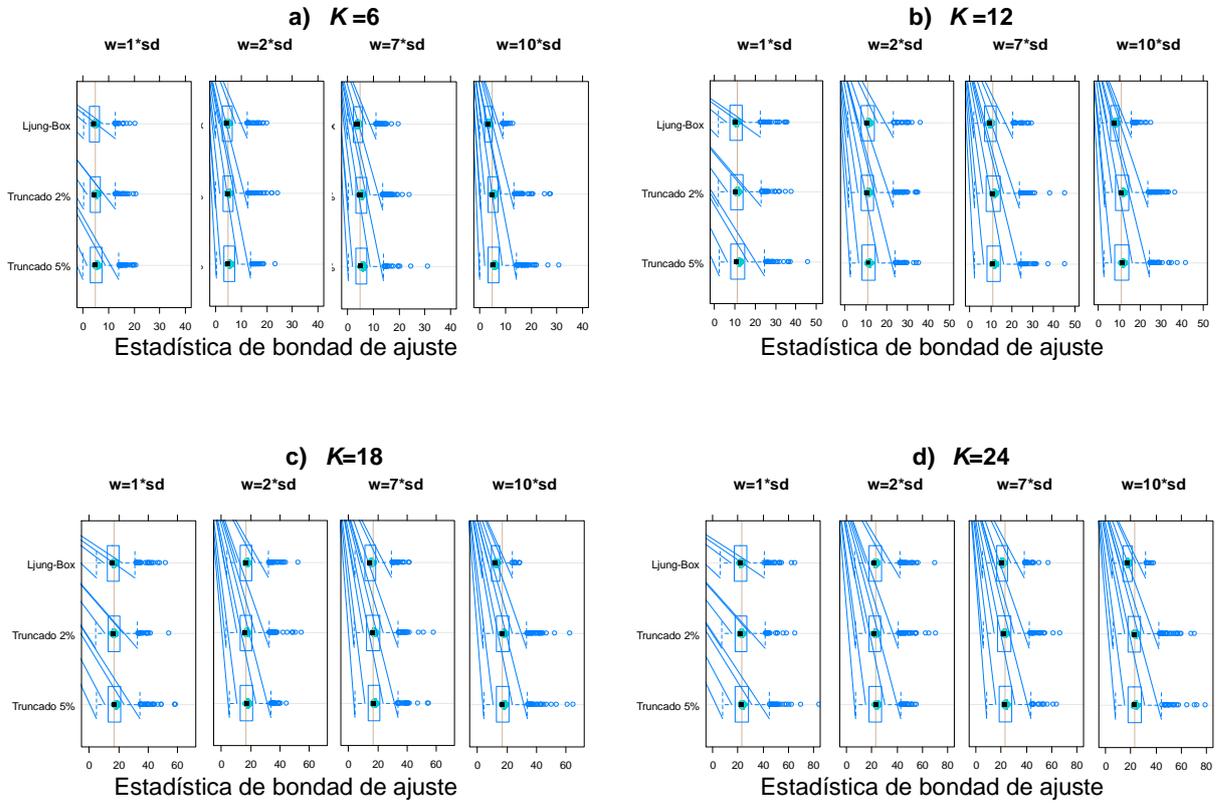
Tabla 3: Medidas descriptivas de la estadística de bondad de ajuste basada en 1000 repeticiones de procesos ARMA(1,1) de tamaño 100 con un OI en  $t=64$

Estadística de bondad de ajuste		K=6				12				18				24			
		$\omega=sd$	2°sd	7°sd	10°sd	$\omega=sd$	2°sd	7°sd	10°sd	$\omega=sd$	2°sd	7°sd	10°sd	w=sd	2°sd	7°sd	10°sd
<b>a) <math>(\phi, \theta) = (0.5, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variación	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,0	5,1	4,5	3,6	10,9	11,2	9,9	8,0	16,8	17,2	15,4	12,6	22,8	23,3	21,3	17,7
	Variación	9,4	9,8	8,1	4,6	24,9	24,2	18,4	11,5	42,2	40,7	32,8	18,6	58,7	61,3	48,7	29,6
	10%	9,36	9,43	8,19	6,48	17,03	17,36	15,62	12,50	25,23	25,39	22,81	18,79	32,43	33,16	30,44	24,98
	5%	11,16	11,06	9,95	7,79	20,08	20,01	17,37	14,16	28,37	29,00	25,71	20,65	36,60	37,98	33,69	27,95
Estimador truncado $\alpha=2\%$	Promedio	5,3	5,4	5,8	5,8	11,3	11,3	11,6	11,8	17,1	17,3	17,5	17,8	23,0	23,3	23,3	23,8
	Variación	10,0	11,2	12,2	12,2	23,9	26,2	27,5	27,5	39,6	44,5	47,9	47,8	58,4	68,6	69,0	76,2
	10%	9,58	9,50	10,31	10,16	17,63	17,76	18,34	18,46	25,33	25,54	26,36	26,56	33,39	33,49	34,24	34,41
	5%	11,71	12,23	13,04	12,07	20,23	20,14	21,26	21,20	28,20	29,65	30,38	30,41	36,92	38,54	39,30	39,16
Estimador truncado $\alpha=5\%$	Promedio	5,7	5,7	5,8	5,9	11,8	11,8	11,8	11,9	17,9	17,8	17,8	17,9	24,0	23,8	23,8	24,0
	Variación	11,9	11,3	11,5	13,4	30,2	25,1	27,1	29,8	52,3	42,1	44,5	49,9	78,3	63,9	62,6	75,5
	10%	9,99	10,33	10,16	10,26	18,73	18,29	18,36	18,84	26,99	26,54	26,47	26,71	35,39	34,32	35,00	34,90
	5%	12,48	12,25	11,89	12,28	21,75	20,85	21,53	22,25	31,33	30,50	30,01	30,25	39,85	39,23	38,10	39,35
<b>b) <math>(\phi, \theta) = (0.8, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variación	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,4	5,3	3,7	2,6	11,3	11,4	8,1	5,8	17,3	17,6	12,7	9,0	23,5	23,8	17,7	12,5
	Variación	11,8	10,3	5,1	2,4	27,8	24,8	11,5	5,8	44,8	41,5	20,0	9,2	67,7	59,5	30,5	14,4
	10%	9,89	9,63	6,87	4,88	17,71	18,05	12,68	9,15	25,61	25,79	18,20	13,22	33,27	34,02	24,81	17,65
	5%	12,07	11,25	8,18	5,76	20,90	20,42	14,41	10,14	29,21	29,33	20,73	14,84	39,18	37,82	27,86	19,53
Estimador truncado $\alpha=2\%$	Promedio	5,4	5,4	5,8	5,8	11,3	11,1	11,7	11,8	17,2	17,1	17,5	17,8	23,1	23,0	23,4	23,7
	Variación	12,1	10,1	12,4	12,4	27,8	22,7	26,1	29,1	44,0	37,3	42,9	46,5	64,1	55,7	60,0	67,9
	10%	9,82	9,75	10,58	10,72	18,32	17,41	18,40	18,98	25,50	24,99	26,68	26,83	32,91	32,93	33,55	34,23
	5%	11,93	11,39	12,70	12,65	20,69	19,23	21,63	21,73	28,12	27,69	29,80	29,44	36,70	37,42	38,28	39,27
Estimador truncado $\alpha=5\%$	Promedio	5,7	5,5	6,0	5,8	11,7	11,4	11,9	11,7	17,7	17,4	17,8	17,6	23,7	23,4	23,7	23,6
	Variación	11,7	10,3	14,0	12,4	25,7	27,6	30,7	29,4	45,3	45,5	51,8	47,7	65,2	67,3	69,3	66,0
	10%	10,05	9,58	10,62	10,28	18,34	18,14	18,83	18,58	26,56	26,13	26,17	26,61	34,20	33,80	33,86	34,39
	5%	11,69	11,33	12,51	12,38	20,69	21,02	22,24	22,02	30,20	28,93	29,95	31,41	38,30	37,30	38,16	38,75
<b>c) <math>(\phi, \theta) = (-0.5, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variación	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,1	5,1	4,5	3,5	11,1	11,3	9,9	8,0	17,0	17,2	15,5	12,7	22,9	23,2	21,3	17,6
	Variación	10,2	10,3	7,9	4,9	23,1	25,4	17,7	11,7	38,5	42,4	30,5	20,2	58,3	65,5	46,0	28,3
	10%	9,30	9,22	8,10	6,48	17,30	17,98	15,17	12,61	25,01	26,07	23,14	18,68	33,01	33,61	30,69	24,59
	5%	11,18	10,99	9,68	7,92	20,01	20,60	17,52	14,31	28,79	29,14	25,15	21,22	36,34	38,00	33,80	27,29
Estimador truncado $\alpha=2\%$	Promedio	5,4	5,4	5,7	6,0	11,3	11,3	11,8	12,0	17,2	17,2	17,6	18,1	23,1	23,2	23,7	24,1
	Variación	11,1	11,1	11,6	14,2	24,5	23,4	25,2	33,8	41,8	40,1	40,9	56,9	60,4	59,3	61,3	77,7
	10%	9,55	9,60	10,12	10,58	17,34	17,27	18,14	19,85	25,59	25,29	25,83	27,14	33,17	33,68	33,60	34,72
	5%	11,29	11,63	12,11	13,15	20,16	20,73	20,89	23,43	28,49	29,49	28,92	31,62	37,95	37,45	37,50	38,88
Estimador truncado $\alpha=5\%$	Promedio	5,8	5,7	5,9	6,2	11,8	11,7	11,8	12,4	17,8	17,6	17,6	18,5	23,8	23,7	23,6	24,7
	Variación	11,6	12,6	12,5	15,0	27,1	31,2	25,7	32,6	48,0	50,5	41,7	53,1	70,6	71,0	61,8	72,8
	10%	10,35	10,22	10,59	11,34	18,79	18,60	18,37	20,15	26,38	26,55	25,99	27,77	34,62	35,01	34,76	35,58
	5%	12,25	12,60	12,45	13,32	21,34	21,86	20,84	22,83	31,31	31,22	28,88	32,25	39,48	40,33	38,10	40,85

Fuente: Elaboración propia a partir de datos simulados.



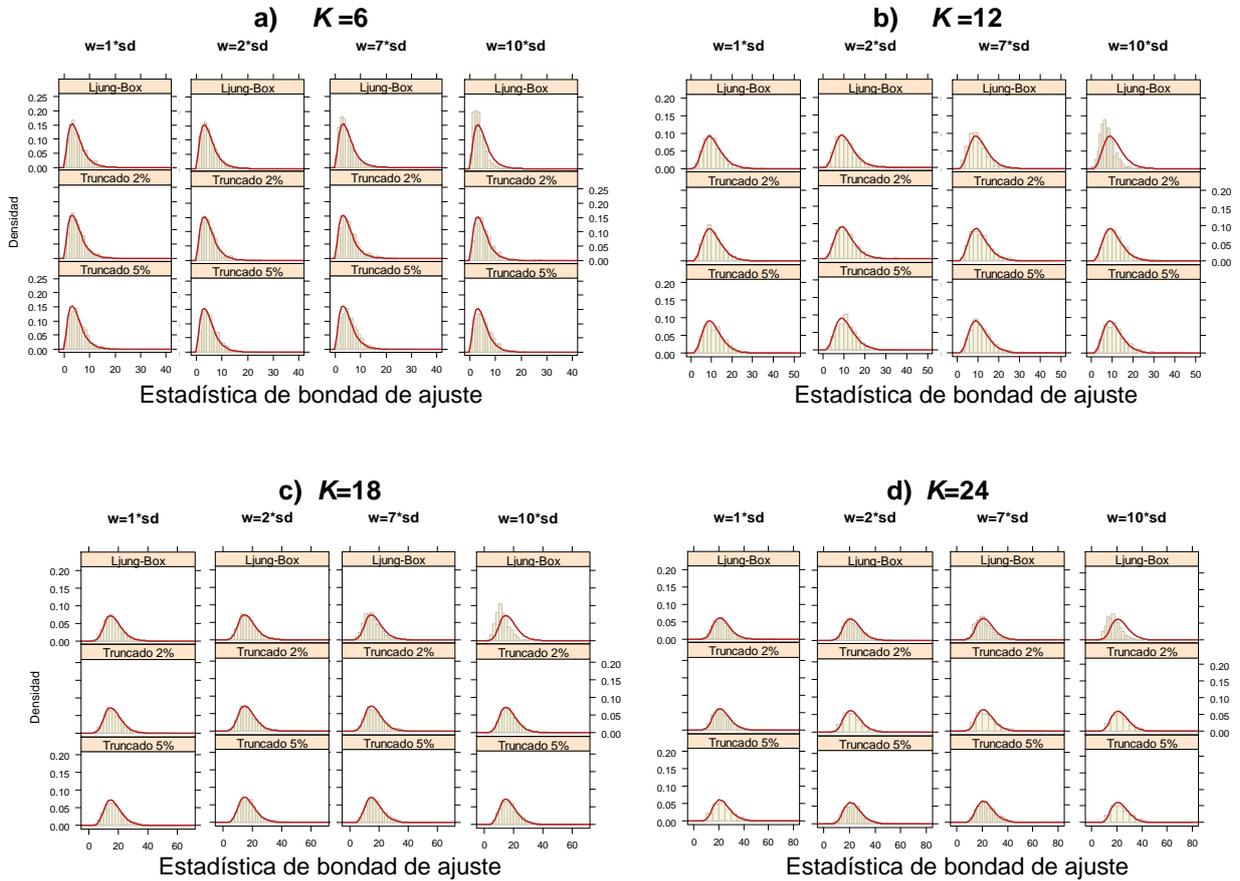
Figura 5: Diagramas de caja de la estadística de bondad de ajuste para las 1000 muestras de tamaño 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$  y un OI en  $t=64$



Fuente: Elaboración propia a partir de datos simulados.



Figura 6: Histogramas de frecuencias de la estadística de bondad de ajuste para las 1000 muestras de tamaño 100 de un proceso ARMA(1,1) con parámetros  $(\phi, \theta) = (0.5, 0)$  y un OI en  $t=64$



Fuente: Elaboración propia a partir de datos simulados.

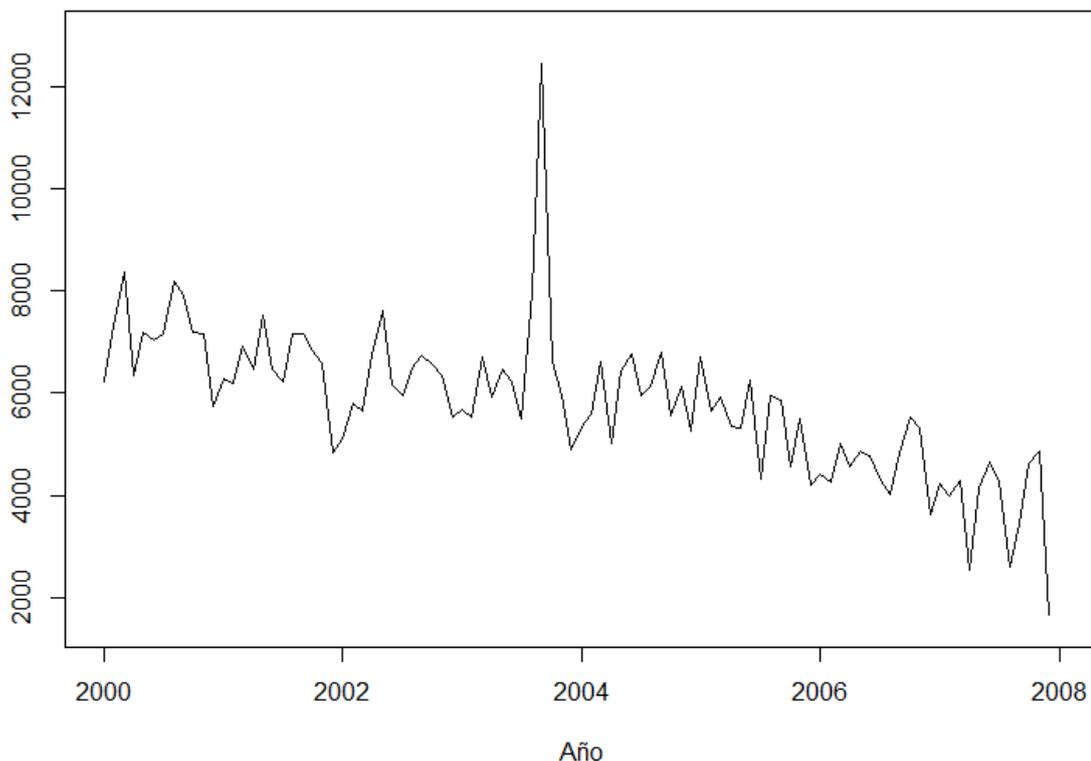


## Aplicación con datos reales

Con el fin de evaluar y comparar el comportamiento de la estadística clásica de Ljung-Box y la estadística robusta de Chan se analizan las mismas en un conjunto de datos donde se detecta la presencia de un posible valor atípico. Con el mismo se pretende mostrar, a través de la aplicación de las estadísticas estudiadas, los inconvenientes que conlleva la presencia de datos extremos en la prueba de bondad de ajuste y, además, evidenciar las ventajas de los estimadores robustos.

La serie de tiempo a analizar corresponde a la cantidad de consultas que se atendieron en consultorio externo del Hospital Roque Sáenz Peña (HRSP) realizadas por pacientes residentes en la ciudad de Rosario durante el período enero 2000 a diciembre 2007 (Figura 7). Los datos de esta serie fueron facilitados por la Secretaria de Salud Pública de la Municipalidad de Rosario (Tabla B.1–Anexo B).

Figura 7: Consultas en consultorio externo del HRSP de pacientes residentes en la ciudad de Rosario. Período: enero 2000 - diciembre 2007.



Fuente: Secretaria de Salud Pública de la Municipalidad de Rosario, Santa Fe, Argentina.

El gráfico de la serie (Figura 7) muestra que la misma no es estable ya que las consultas realizadas en consultorio externo en el HRSP no oscilan alrededor de un valor fijo sino que disminuyen con los años. Se observa que la serie presenta una tendencia globalmente decreciente.

Además, en septiembre del año 2003 se detecta la presencia de un valor inconsistente comparado con el resto de la serie, se observa un elevado número de consultas para este mes. Por tanto, se puede conjeturar que esta observación corresponde a un valor atípico.



Bonifazi (2015) aplicó métodos robustos para estimar la función de autocorrelación e identificar un modelo adecuado para dicha serie. Mediante la función de autocorrelación muestral robusta (Ma & Genton, 2000) y la aplicación de criterios de selección de modelos se identificó que un modelo razonable para representar la serie del número de consultas en consultorio externo del HRSP de pacientes residentes en la ciudad de Rosario es un proceso  $SARIMA(0,1,2)(0,1,1)_{12}$ :

$$(1 - B)(1 - B^{12})z_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta B^{12})a_t$$

El modelo estimado por el método máxima verosimilitud es:

$$(1 - B)(1 - B^{12})z_t = (1 - 0.537 B - 0.325 B^2)(1 - 0.803 B^{12})a_t,$$

a partir del cual se calculan los residuos del modelo, mediante la diferencia entre los valores observados y los valores estimados por el modelo.

Luego de estimar el modelo, se debe realizar la siguiente etapa correspondiente a la diagnosis. En esta etapa se verifica mediante la prueba de bondad de ajuste que el modelo proporciona un ajuste adecuado. Entonces, el contraste de bondad de ajuste que prueba la hipótesis de falta de correlación, es decir que los residuos provienen de un modelo ruido blanco se puede expresar:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0, \quad \text{para } K = 6, 12, 18 \text{ y } 24$$

Se utilizan las estadísticas  $Q_{LB}(K)$  y  $Q_{T,\alpha}(K)$  para su verificación. Las mismas son presentadas en la Tabla 4.

Tabla 4: Estadísticas de la prueba de bondad de ajuste

K	gl	$Q_{LB}(K)$		$Q_{T,\alpha=2\%}(K)$		$Q_{T,\alpha=5\%}(K)$	
		Estadística	p-value	Estadística	p-value	Estadística	p-value
6	3	2.54	0.4679	3.39	0.3355	4.32	0.2285
12	9	3.98	0.9124	7.24	0.6124	7.48	0.5875
18	15	12.10	0.6711	12.55	0.6369	8.88	0.8840
24	21	20.86	0.4672	20.04	0.5190	10.62	0.9696

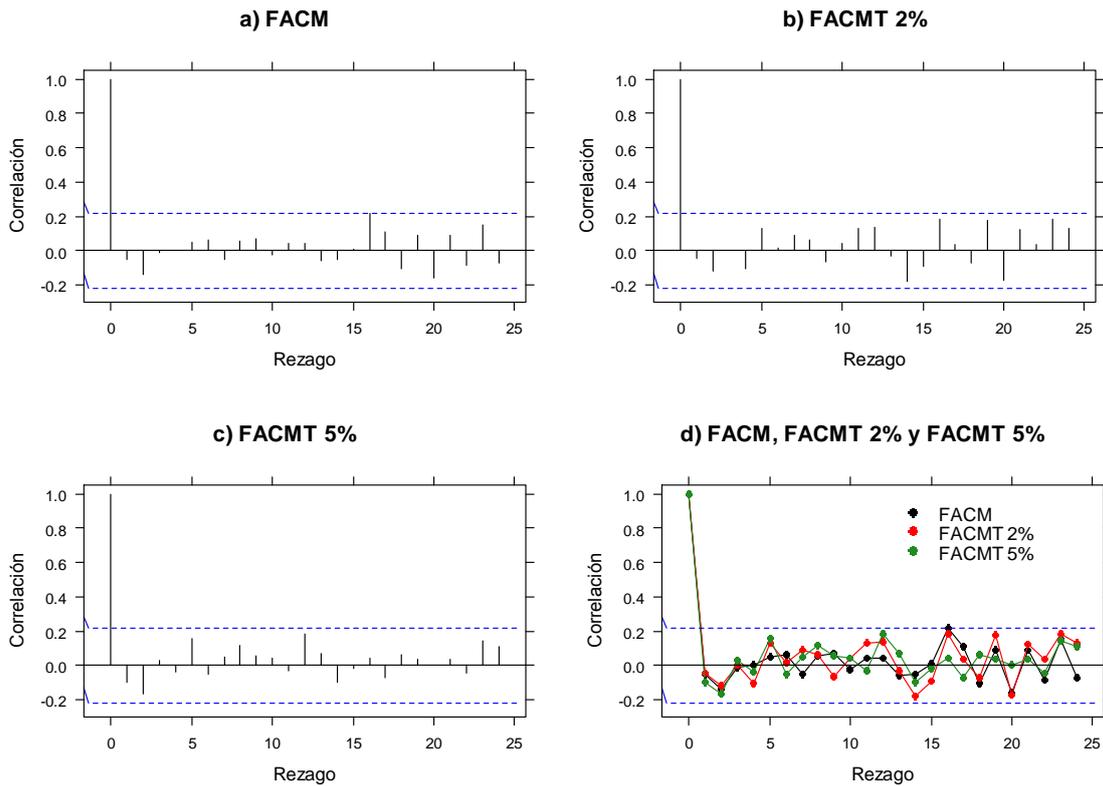
Fuente: Elaboración propia en base a datos de la Secretaria de Salud Pública de la Municipalidad de Rosario.

Al evaluar el comportamiento de la estadística de bondad de ajuste para el modelo que representa al número de consultas en consultorio externo del HRSP mediante ambos métodos estudiados, se observa que para autocorrelaciones correspondientes a los rezagos chicos ( $K = 6$  y  $12$ ), la estadística clásica  $Q_{LB}(K)$  nos provee valores más pequeños que para el estimador truncado. Esto se debe a que las autocorrelaciones estimadas mediante el método de los momentos son más pequeñas para los primeros rezagos (Figura 8). Sin embargo, para rezagos más grandes la FACM y FACMT con  $\alpha = 2\%$  presentan valores similares, y superiores a la FACMT con  $\alpha = 5\%$ , que conducen a valores superiores en la estadística de *portmanteau* en relación con el estimador truncado con  $\alpha = 5\%$ .



En efecto, al calcular conjuntamente la estadística de bondad de ajuste del modelo  $SARIMA(0,1,2)(0,1,1)_{12}$  por el método de Ljung-Box y el método truncado se obtuvieron algunas discrepancias. Sin embargo, las mismas no afectan a las conclusiones arribadas en la prueba de hipótesis, ya que en todos los casos se concluiría que los residuos están incorrelacionados.

Figura 8: FAC estimada para los residuos del modelo



Fuente: Elaboración propia en base a datos de la Secretaria de Salud Pública de la Municipalidad de Rosario.



## CONCLUSIONES

El estadístico de significación conjunta de *portmanteau* es usado comúnmente para evaluar la bondad de ajuste de los modelos de series de tiempo. Sin embargo, el desempeño de esta prueba puede depender de la presencia de una o varias observaciones atípicas en la serie. En este trabajo se presenta un estadístico robusto para la prueba de *portmanteau*, propuesto por Chan (1994), basado en el recorte o eliminación de las observaciones extremas. Se investiga la sensibilidad en presencia de valores atípicos de este estimador robusto y del estadístico clásico de Ljung-Box, mediante un estudio de simulación bajo diversos escenarios y la aplicación con datos reales. Se observa que, cuando la serie no presenta *outliers*, el comportamiento del estadístico de *portmanteau* robusto es similar al de la prueba clásica en todos los aspectos considerados. Sin embargo, si hay valores atípicos, la distribución de la estadística clásica difiere de la de una variable aleatoria chi-cuadrada, pudiendo conducir a conclusiones erróneas en la prueba de bondad de ajuste del modelo. Por otra parte, la distribución del estimador truncado pareciera aproximarse razonablemente a la teoría asintótica. En conclusión, se puede pensar que la prueba conjunta propuesta por Chan (1994) es útil en la modelización robusta de series temporales. Queda pendiente el análisis del desempeño de estos estimadores en presencia de múltiples datos atípicos en la serie y en el caso de procesos con estacionalidad. Además, en una futura línea de investigación, se puede abordar los inconvenientes producidos por la presencia de un valor atípico innovativo en las estimaciones máximo verosímiles de los parámetros de procesos MA y ARMA.



## REFERENCIAS BIBLIOGRÁFICAS

- Arranz, M. A.** (2005). Portmanteau Test Statistic in Time Series. *Tol-Project*
- Box, G. E. P.; Jenkins, D. A.** (1976). *Time Series Analysis and Control*, 2<sup>nd</sup> 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 Statistical Association*, 65, 1509-1526.
- 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*.
- Brockwell, P. J.; Davis, R. A.** (1991). *Time Series: Theory and Methods*. Springer Verlag.
- Chan, W.; Wei, W.** (1992). A comparison of some estimators of time series autocorrelations. *Computational Statistics & Data Analysis*, 14, 146-163.
- Chan, W.S.** (1994). On portmanteau goodness-of-fit test in robust time series modelling. *Computational Statistics*, 9, 301-310.
- Duerre, A.; Fried, R.; Liboschik, T.** (2014). Robust estimation of (partial) autocorrelation. *SFB 823 Discussion Paper 12/14*.
- Fox, A. J.** (1972). Outliers in Time Series. *Journal Royal Statistical Society*, B 34, 350-363.
- Guerrero Guzmán, V.** (2003). *Análisis estadístico de series de tiempo económicas. Segunda edición*. Thomson.
- Ljung, G. M.; Box, G. E. P.** (1978). On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303.
- Li, W. K.** (2004). *Diagnostic checks in time series*. Chapman & Hall/CRC.
- Li, W. K.** (1988). A Goodness of Fit Test in Robust Time Series Modelling. *Biometrika*, 75, 2, 355-361.
- Lee, A. H.; Yick J. S.; Hui, Y. V.** (2001). Sensitivity of the portmanteau statistic in time series modeling. *Journal of Applied Statistics*, 28,6, 691-702.
- Maronna, R.A.; Martin, R.D.; Yohai, V.J.** (2006). *Robust Statistics: Theory and Methods*. John Wiley and Sons.
- Peña, D.** (2005). *Análisis de series temporales*. Alianza Editorial.
- Peña, D.; Tiao, G.; Tsay, R.** (2001). *A course in time series analysis*. John Wiley & Sons, Inc.
- Wei, W.** (2006). *Time Series Analyses. Univariate and Multivariate Methods. Second edition*. Pearson Education, Inc.



**ANEXO**

**ANEXO A - Simulaciones**

Tabla A.1: Medidas descriptivas de la estadística de bondad de ajuste basada en 1000 repeticiones de tamaño 20, 50 y 100 para procesos ARMA(1,1)

Estadística de bondad de ajuste		K= 6			12			18			24		
		n= 20	50	100	20	50	100	20	50	100	20	50	100
<b>a) <math>(\phi, \theta) = (0.8, 0)</math></b>													
<b>Teórico</b>	Promedio	5			11			17			23		
	Variancia	10			22			34			46		
	10%	9,24			17,28			24,77			32,01		
	5%	11,07			19,68			27,59			35,17		
<b>Estimador Ljung-Box</b>	Promedio	5,7	5,5	5,3	11,9	11,8	11,4	-	17,9	17,4	-	24,2	23,5
	Variancia	12,5	11,5	11,1	28,3	28,1	25,2	-	51,7	41,1	-	74,5	62,8
	10%	9,90	9,92	9,66	18,24	19,06	18,10	-	27,20	25,60	-	34,66	33,45
	5%	11,69	11,33	12,05	21,48	21,38	21,25	-	30,90	29,31	-	39,82	38,00
<b>Estimador truncado <math>\alpha=2\%</math></b>	Promedio	4,7	5,1	5,2	9,9	10,8	11,1	-	16,7	17,0	-	22,4	22,9
	Variancia	8,4	9,3	9,2	19,1	24,5	21,9	-	44,3	37,5	-	67,7	57,6
	10%	8,28	9,10	9,25	15,65	17,02	17,20	-	24,85	24,45	-	32,62	32,25
	5%	10,08	10,83	11,01	18,11	20,04	19,69	-	29,57	27,68	-	37,36	37,13
<b>Estimador truncado <math>\alpha=5\%</math></b>	Promedio	5,1	5,4	5,5	10,6	11,2	11,4	-	17,0	17,4	-	22,7	23,1
	Variancia	10,2	11,3	11,4	21,7	27,5	24,9	-	47,0	42,5	-	71,6	62,9
	10%	9,12	9,67	9,98	16,65	17,97	17,82	-	25,46	25,57	-	34,27	33,54
	5%	11,51	12,01	11,67	19,07	20,98	20,57	-	29,25	28,89	-	38,36	37,90
<b>b) <math>(\phi, \theta) = (-0.5, 0)</math></b>													
<b>Teórico</b>	Promedio	5			11			17			23		
	Variancia	10			22			34			46		
	10%	9,24			17,28			24,77			32,01		
	5%	11,07			19,68			27,59			35,17		
<b>Estimador Ljung-Box</b>	Promedio	5,1	5,1	5,2	11,0	11,0	11,0	-	16,9	17,0	-	22,9	23,1
	Variancia	10,3	9,7	10,9	24,1	23,5	26,3	-	42,1	42,5	-	65,2	62,9
	10%	8,98	9,40	9,50	17,33	17,49	17,14	-	25,65	24,77	-	33,13	33,52
	5%	10,83	11,10	11,19	19,83	20,37	20,82	-	28,95	29,10	-	37,49	37,88
<b>Estimador truncado <math>\alpha=2\%</math></b>	Promedio	4,4	5,3	5,5	9,5	10,9	11,3	-	16,5	17,2	-	22,2	23,3
	Variancia	8,2	10,1	11,7	18,8	24,0	28,0	-	42,9	45,2	-	63,0	67,9
	10%	7,99	9,40	9,80	15,04	17,33	18,29	-	24,91	25,80	-	32,32	33,65
	5%	9,75	11,40	11,76	17,56	20,58	21,69	-	29,02	30,16	-	37,00	38,12
<b>Estimador truncado <math>\alpha=5\%</math></b>	Promedio	5,2	5,4	5,7	10,7	11,1	11,6	-	16,7	17,5	-	22,3	23,6
	Variancia	10,4	10,0	11,8	22,3	24,7	26,7	-	44,6	47,4	-	63,6	71,6
	10%	9,13	9,39	10,23	16,22	18,07	18,55	-	25,15	26,18	-	32,46	34,27
	5%	11,33	11,40	12,57	19,31	20,63	21,54	-	29,82	30,59	-	37,61	40,41

Fuente: Elaboración propia a partir de datos simulados.



Tabla A.2: Medidas descriptivas de la estadística de bondad de ajuste basada en 1000 repeticiones de procesos ARMA(1,1) de tamaño 100 con un OA en  $t=64$

Estadística de bondad de ajuste		K=6				12				18				24			
		$\omega=sd$	2°sd	7°sd	10°sd	$\omega=sd$	2°sd	7°sd	10°sd	$\omega=sd$	2°sd	7°sd	10°sd	w=sd	2°sd	7°sd	10°sd
<b>a) <math>(\phi, \theta) = (0.8, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variancia	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,1	4,6	2,4	1,5	10,8	10,0	5,1	3,3	16,6	15,5	8,0	5,1	22,3	21,1	11,0	7,1
	Variancia	10,3	8,8	2,0	0,9	23,7	20,9	4,6	2,1	40,0	36,8	7,9	3,7	57,4	52,3	12,0	5,3
	10%	9,11	8,40	4,13	2,73	17,04	16,19	7,93	5,20	24,85	23,34	11,92	7,73	31,61	30,93	15,67	10,22
	5%	10,73	10,10	4,89	3,26	19,35	18,40	9,21	5,89	28,33	26,40	13,29	8,72	36,21	34,13	17,16	11,31
Estimador truncado $\alpha=2\%$	Promedio	5,3	5,2	6,1	6,8	11,1	10,7	12,1	12,8	17,0	16,6	18,0	18,9	22,8	22,6	24,1	24,9
	Variancia	11,5	9,7	13,4	20,6	27,9	20,8	29,9	35,9	46,4	35,7	50,2	56,7	67,1	51,7	72,8	77,0
	10%	9,60	9,30	10,74	12,36	17,89	16,89	19,38	20,56	25,77	24,56	26,99	28,42	33,10	31,89	35,26	35,68
	5%	11,33	11,31	12,76	15,88	21,25	18,73	22,62	24,56	29,95	27,31	31,33	32,65	37,95	35,07	39,99	40,16
Estimador truncado $\alpha=5\%$	Promedio	5,6	5,5	6,2	6,5	11,4	11,3	12,2	12,7	17,4	17,3	18,1	18,8	23,2	23,3	24,3	24,8
	Variancia	11,6	10,1	14,7	18,0	27,9	22,3	31,2	39,2	41,7	40,1	50,5	68,6	60,6	59,2	76,3	97,0
	10%	9,67	9,42	10,85	12,12	17,92	17,65	19,68	20,98	25,31	25,76	27,24	28,99	33,61	33,06	35,43	36,82
	5%	11,81	11,53	12,80	15,26	21,23	20,35	22,48	24,45	29,24	29,01	30,97	32,86	37,39	37,46	40,68	42,21
<b>b) <math>(\phi, \theta) = (-0.5, 0)</math></b>																	
Teórico	Promedio	5				11				17				23			
	Variancia	10				22				34				46			
	10%	9,24				17,28				24,77				32,01			
	5%	11,07				19,68				27,59				35,17			
Estimador Ljung-Box	Promedio	5,0	4,9	3,2	2,3	10,9	10,6	6,9	5,1	17,0	16,5	10,8	8,0	22,8	22,3	14,7	11,0
	Variancia	8,8	9,1	4,0	2,2	23,1	23,3	10,0	4,5	43,2	40,3	16,7	7,8	68,6	58,5	25,4	11,6
	10%	8,63	8,70	5,78	4,29	17,24	17,10	10,92	7,98	25,36	24,84	16,62	11,62	33,95	31,65	21,54	15,34
	5%	10,26	10,46	7,03	5,16	19,71	19,58	12,74	8,90	28,74	28,60	18,55	12,94	37,83	36,23	24,70	17,17
Estimador truncado $\alpha=2\%$	Promedio	5,5	5,2	5,7	6,2	11,3	11,0	11,6	12,1	17,1	16,9	17,5	18,1	22,9	22,7	23,4	23,9
	Variancia	10,4	10,4	11,6	13,6	23,8	23,9	25,3	28,5	42,1	41,3	40,8	45,7	64,9	59,3	58,7	68,4
	10%	9,76	9,27	10,43	10,99	17,98	17,01	18,16	18,84	24,92	24,91	26,00	26,17	32,32	32,65	33,87	33,85
	5%	11,46	10,99	12,41	12,61	20,18	19,85	20,59	21,01	28,39	28,70	29,16	29,01	37,23	37,58	37,25	37,73
Estimador truncado $\alpha=5\%$	Promedio	5,7	5,5	6,1	6,1	11,7	11,4	12,0	12,1	17,5	17,4	18,0	18,0	23,3	23,3	23,9	23,7
	Variancia	10,8	12,0	14,2	13,2	25,5	25,0	30,6	29,8	42,2	44,8	48,8	50,7	64,9	65,5	70,4	70,9
	10%	10,18	9,67	10,58	10,39	18,11	17,73	19,06	19,10	26,33	25,21	26,83	27,07	33,18	32,79	34,03	34,23
	5%	12,10	11,90	13,06	12,58	20,63	20,39	21,74	22,07	30,09	30,10	30,15	30,88	38,12	38,58	38,41	39,53

Fuente: Elaboración propia a partir de datos simulados.



## ANEXO B - Aplicación con datos reales

**Tabla B.1:** Consultas de pacientes residentes en la ciudad de Rosario realizadas en los consultorios externos del HRSP. Período: enero 2000 a diciembre 2007

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Promedio
<b>2000</b>	6218	7301	8372	6347	7197	7031	7165	8186	7912	7183	7152	5723	<b>7149</b>
<b>2001</b>	6279	6197	6911	6458	7514	6447	6214	7150	7143	6836	6584	4835	<b>6547</b>
<b>2002</b>	5118	5812	5636	6734	7610	6157	5952	6482	6737	6548	6335	5527	<b>6221</b>
<b>2003</b>	5675	5539	6712	5905	6466	6219	5502	7980	12450	6626	5944	4896	<b>6660</b>
<b>2004</b>	5354	5586	6608	5002	6419	6749	5952	6118	6804	5560	6136	5268	<b>5963</b>
<b>2005</b>	6711	5637	5931	5346	5309	6265	4305	5950	5854	4558	5502	4200	<b>5464</b>
<b>2006</b>	4404	4253	5027	4547	4847	4778	4294	4026	4824	5533	5319	3629	<b>4623</b>
<b>2007</b>	4235	3980	4289	2550	4126	4645	4260	2611	3428	4624	4870	1665	<b>3774</b>
<b>Promedio</b>	<b>5499</b>	<b>5538</b>	<b>6186</b>	<b>5361</b>	<b>6186</b>	<b>6036</b>	<b>5456</b>	<b>6063</b>	<b>6894</b>	<b>5934</b>	<b>5980</b>	<b>4468</b>	<b>5800</b>

Fuente: Secretaria de Salud Pública de la Municipalidad de Rosario, Santa Fe, Argentina.