# Técnicas de Pronóstico

Trabajo 2

## **Análisis de Serie de Tiempo:**

**Modelos ARMA-SARMA para los Residuos Estructurales**

Maria Isabel Arango Acevedo y David Fernando Rivera Olarte

*Octubre de 2017*

##### Resumen

Este documento presenta los análisis realizados a la serie de tiempo relacionada con el turismo en Alemania respecto a las estadías de una sola noche, así como un comparativo entre modelo escogido para pronosticar el último año de la serie y los datos observados de la serie para el mismo año. En esta entrega ser profundiza en el análisis de los residuos estructurales y su ajuste con análisis de modelos ARMA y S-ARMA.

**Palabras claves**: Serie de tiempo, AIC, BIC, AR, MA, ARMA, S-ARMA.

##### Abstract

This paper presents the analyzes of the time series related to tourism in Germany with regard to single-night stays, as well as a comparison between the model chosen to forecast the last year of the series and the observed data from the series for the same year. In this delivery, we will deepen the analysis of the structural residues and their adjustment with analysis of ARMA and S-ARMA models.

**Key words**: Time series, AIC, BIC, AR, MA, ARMA, S-ARMA.

# Introducción

La serie de datos trabajada corresponde a “permanencia en una noche” (overnight stay), que corresponde a turismo en Alemania. Según la lectura de datos, se está analizando el período comprendido entre 01/01/2005 y 01/06/2017, y se leen en miles de estadías en hoteles, moteles, casas de huéspedes, hospitales/clínicas y spas.

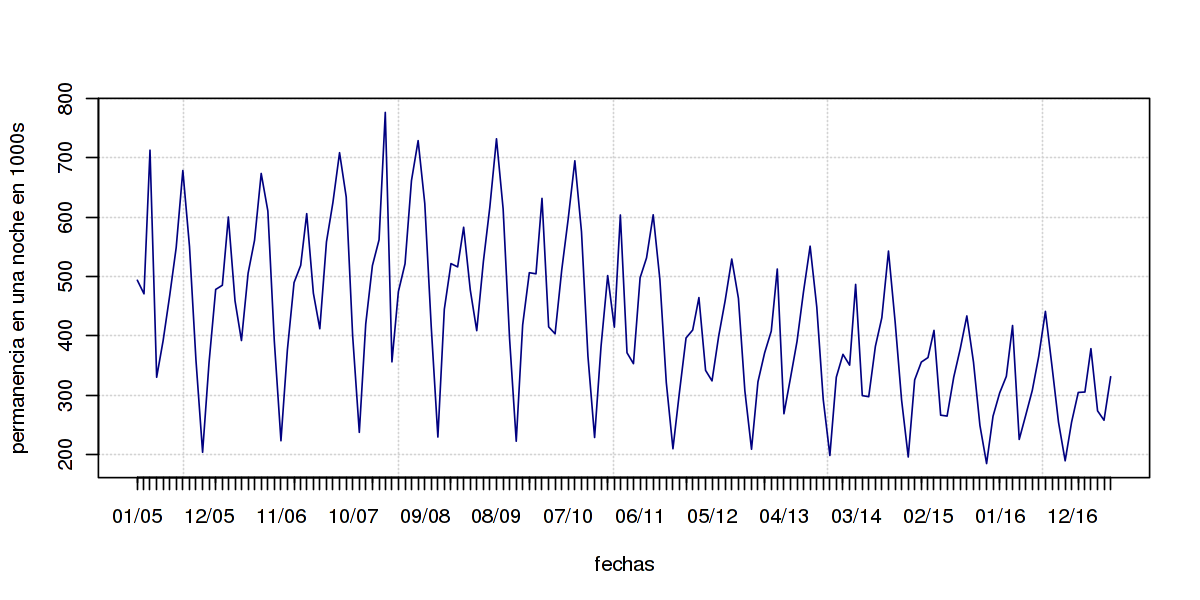


Figura 1: Estadías de una noche - Turismo Suiza. [[1]](#footnote-1)

Dentro de los modelos analizados en el entregable anterior, se tiene que el modelo que mejor se ajustó respecto a los datos observados y en pronósticos era una ecuación del tipo exponencial cubica con estacionalidad marcada con indicadoras:

(1)

Con este modelo, se obtuvieron los siguientes resultados de datos observados sobre los ajustados:

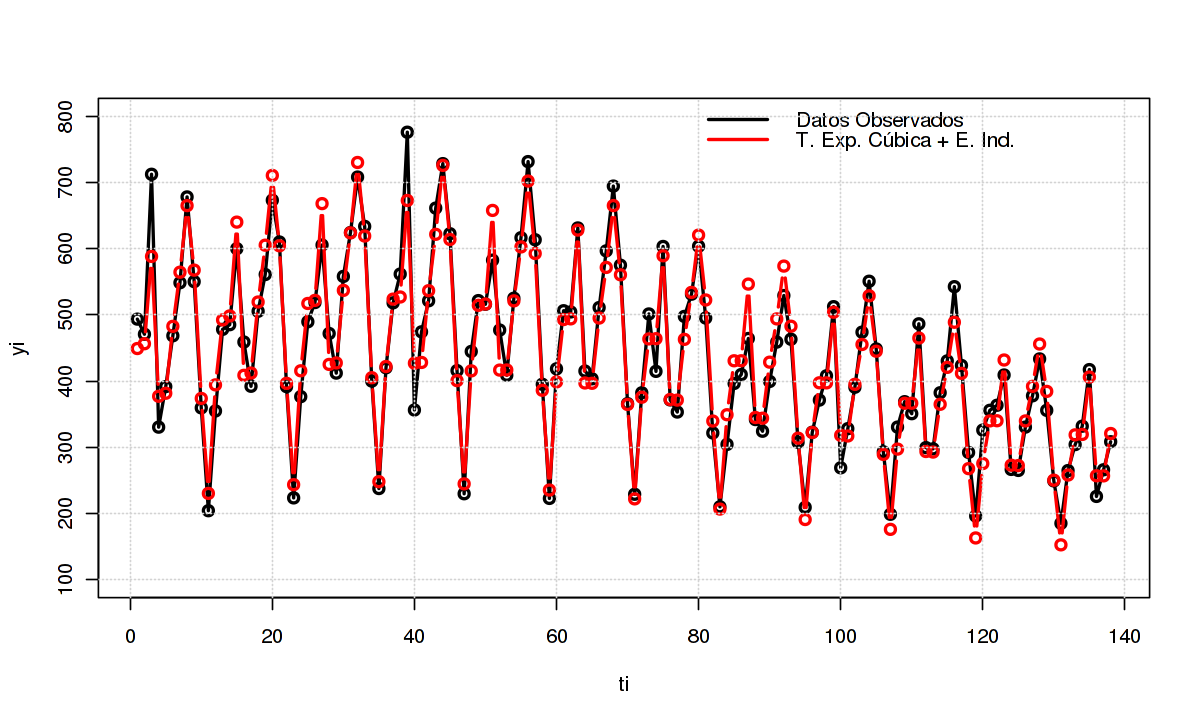


Figura 2- Datos Observados vs Datos Ajustados.

Y de acuerdo con el modelo, pronóstico de los últimos 12 datos de la serie de tiempo:

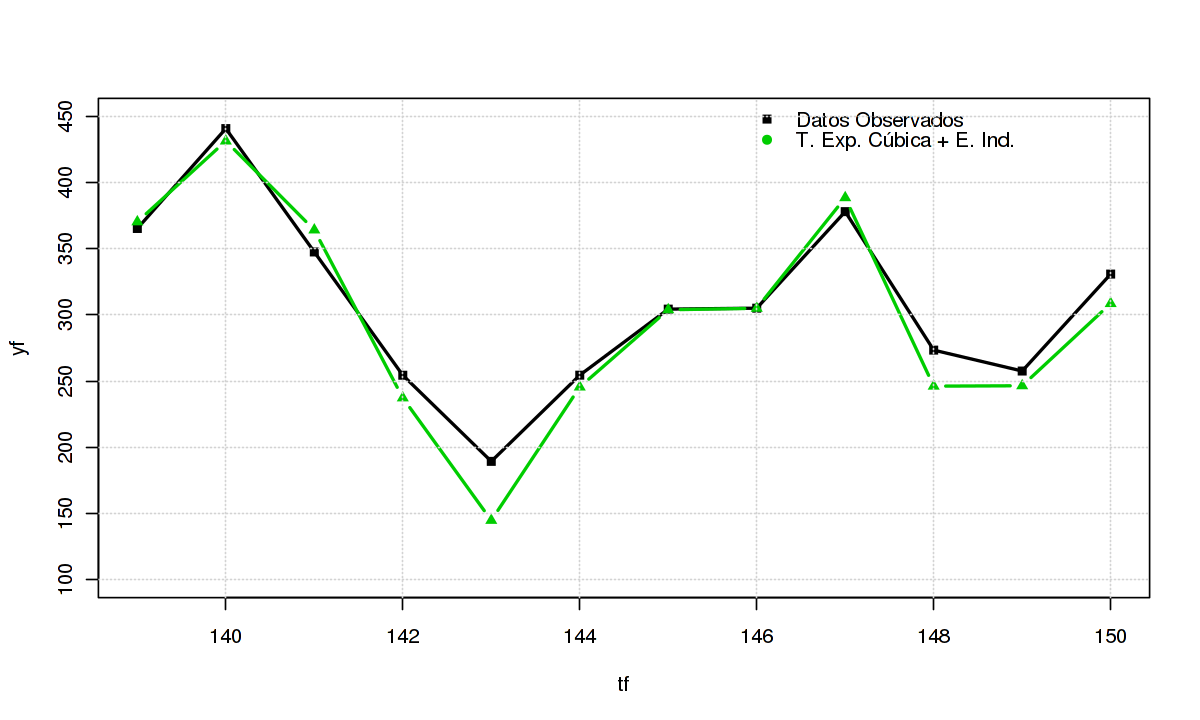


Figura 3- Pronóstico de la Serie

Con estos resultados, es el objetivo del presente entregable hacer el modelamiento de los residuos estructurales a través de modelos ARMA o S-ARMA, de acuerdo con criterios de ajuste mediante estadísticos como AIC, BIC y MAPE para los resultados en los pronósticos.

# Pruebas de Incorrelación

Como primera medida, se determina si los residuos estructurales son ruido blanco. A continuación, se expone el comportamiento de esta componente en el tiempo, la gráfica de función de autocorrelación y la gráfica de función de autocorrelación parcial:

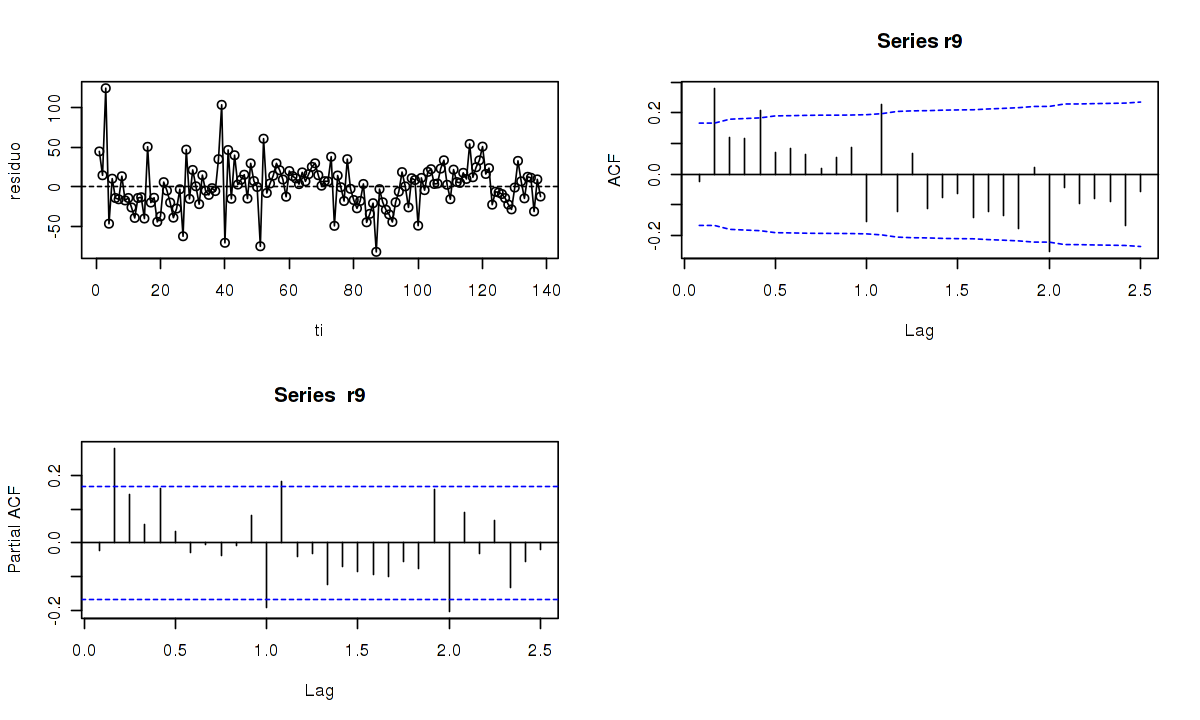


Figura 4- ACF, PACF de los residuos estructurales.

Para los datos en cuestión se aplican las pruebas Ljung-Box, obteniendo:

Tabla 1- Pruebas Ljung-Box residuos estructurales

|  |  |  |
| --- | --- | --- |
| **Rezagos** |  | **p-value** |
| 5 | 22 | 0,0005948 |
| 15 | 40 | 0,0003984 |
| 25 | 69 | 0,005491 |

Y, por último, se aplica la prueba Durbin-Watson generalizada, obteniendo:

Tabla 2- Pruebas Durbin Watson Generalizadas

|  |  |  |  |
| --- | --- | --- | --- |
| **lag** | **Autocorrelation** | **D-W Statistic** | **p-value** |
| 1 | -0.02645121 | 2.033.631 | 0.954 |
| 2 | 0.27874281 | 1.420.686 | 0.000 |
| 3 | 0.11643249 | 1.605.692 | 0.020 |
| 4 | 0.11450047 | 1.592.067 | 0.026 |
| 5 | 0.20484650 | 1.409.328 | 0.000 |
| 6 | 0.06643548 | 1.682.556 | 0.134 |
| 7 | 0.07913294 | 1.655.257 | 0.102 |
| 8 | 0.06130401 | 1.681.354 | 0.176 |
| 9 | 0.01538833 | 1.771.043 | 0.472 |
| 10 | 0.05284372 | 1.687.167 | 0.256 |
| 11 | 0.08517716 | 1.612.608 | 0.150 |
| 12 | -0.15715372 | 2.083.252 | 0.158 |
| 13 | 0.22455291 | 1.317.485 | 0.000 |
| 14 | -0.12341638 | 2.011.580 | 0.256 |
| 15 | 0.06463627 | 1.622.502 | 0.256 |

Concluyendo de lo anterior:

* Determinación de componentes AR o MA a partir de las gráficas AFC y PACF.
  + Los residuos no se encuentran contenidos dentro de las bandas den Bartlett en la gráfica ACF, sugiriendo que no se trata de un ruido blanco y posibilitando los análisis de presencia de modelos AR o MA.
  + MA(q): Dado que la gráfica de ACF no termina abruptamente[[2]](#footnote-2), no es posible determinar que la se trate de una serie tipo MA(q)
  + De la PACF se pueden observar los rezagos 2, 5, 12 y 13 superando de las bandas de Bartlett, sugiriendo un posible modelo AR, ya que, desde ese punto en adelante, los rezagos se ven contenidos dentro de las bandas.
* Criterios para aceptación de la prueba Ljung-Box
  + Si p-value <0.05: Se puede rechazar la hipótesis nula suponiendo un 5% de posibilidades de cometer un error[[3]](#footnote-3). Entonces se puede asumir que sus rezagos muestran dependencia el uno del otro.
* Criterios para aceptación de la prueba Durbin-Watson
  + Se aplica DWG ya se tiene un modelo diferente a un AR(1), de acuerdo lo encontrado en las evaluaciones de las gráficas ACF, y de la tabla se tiene que con el estadístico d < 2 indicando posible autocorrelación positiva en la serie[[4]](#footnote-4). Para la mayoría de los valores de los lag evaluados, este valor es menor a 2, indicando que se presenta autocorrelación en esos rezagos de la serie.

De acuerdo con los puntos expuestos se determina que el residuo estructural trabajado no corresponde a ruido blanco y se debe proceder a modelarlos, de acuerdo al objetivo propuesto, como ARMA o S-ARMA.

# Identificación

## **3.1** **Armasubsets**

Por medio del uso de la función armasubsets de la librería TSA, se determinarán los posibles modelos ARMA o S-ARMA presente en la serie de datos analizados.

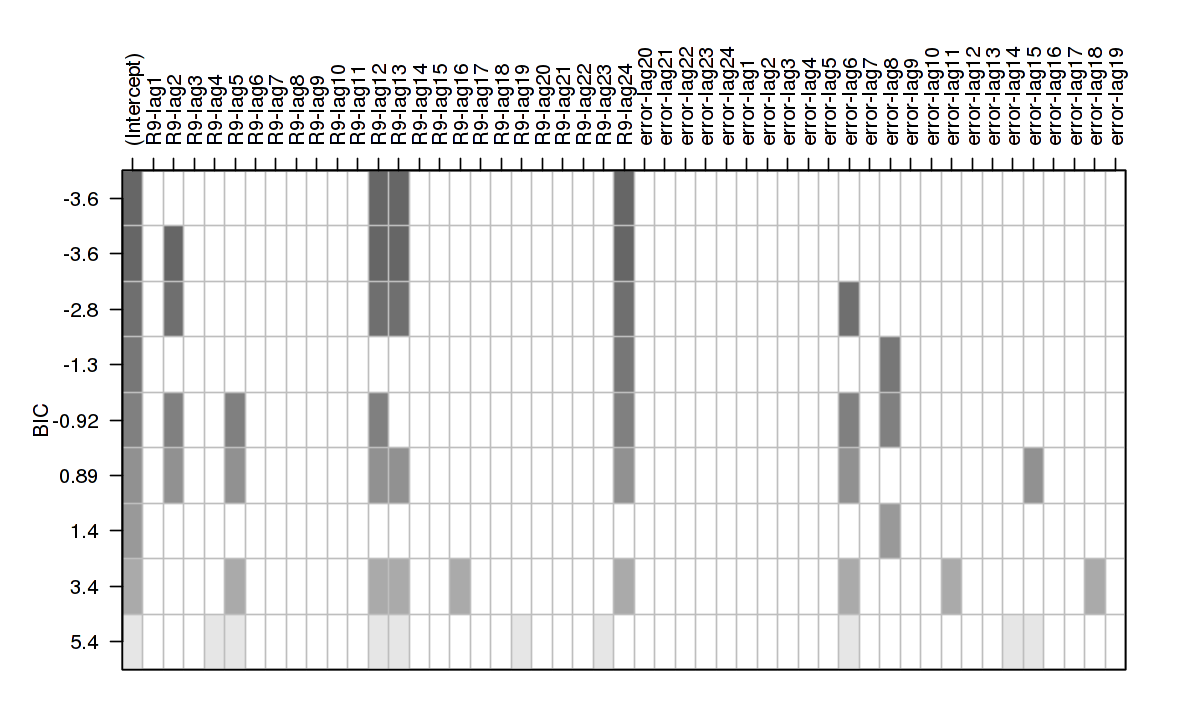


Figura 5- Resultados armasubsets.

Por tener un valor mayor de BIC las componentes MA respecto a las AR estarían descartadas del modelo[[5]](#footnote-5), sugiriéndose entonces un modelo AR como sigue: **p = 2;12;13**.

## **3.2 SelectModel-FitAr**

Con la aplicación de estas funciones, se trata de encontrar modelos AR(p) que se ajusten al comportamiento de los residuos, bajo el criterio del modelo que arroje mejor AIC, obteniéndose el modelo **AR(5)**, siendo este uno de los rezagos presentes en las gráficas del *armasubset*, pero sin ser la más predominante en la lectura del BIC.

## **3.3 Auto.arima**

Para esta función se debe colocar la parametrización de detección de componentes estacionales y estacionarias, con rezagos máximos de 24 periodos con el fin de detectar componentes altas.

Con esta configuración se obtiene un modelo **S-ARMA ﻿(3,1)(3,1)[12]**.

## **3.4 Autosmarfit**

A través de esta función, se trata de buscar un posible ARMA – S-ARMA, reportando el siguiente modelo **ARMA(2,1)**.

# Estimación

De acuerdo con los modelos resaltados en los anteriores puntos, se realizan los cálculos de AIC para cada uno de ellos:

Tabla 3- Comparación AIC para los modelos ARMA - S-ARMA

|  |  |
| --- | --- |
| **modelo** | **AIC** |
| ARIMA(2,0,0) | 1.324,32 |
| ARIMA(5,0,0) | 1.322,28 |
| ARIMA(12,0,0) | 1.327,90 |
| ARIMA(13,0,0) | 1.324,88 |
| **ARIMA(3,0,1)(3,0,1)[12]** | **1.300,84** |
| **ARIMA(2,0,1)** | **1.320,61** |

De aquí se observa que los modelos **ARIMA(3,0,1)(3,0,1)[12]** y **ARIMA(2,0,1)** representa y con ellos se procede a realizar las validaciones respectivas de curvas AFC, PACF, qqplot, densidad, periodograma acumulado y la prueba Ljung-Box.

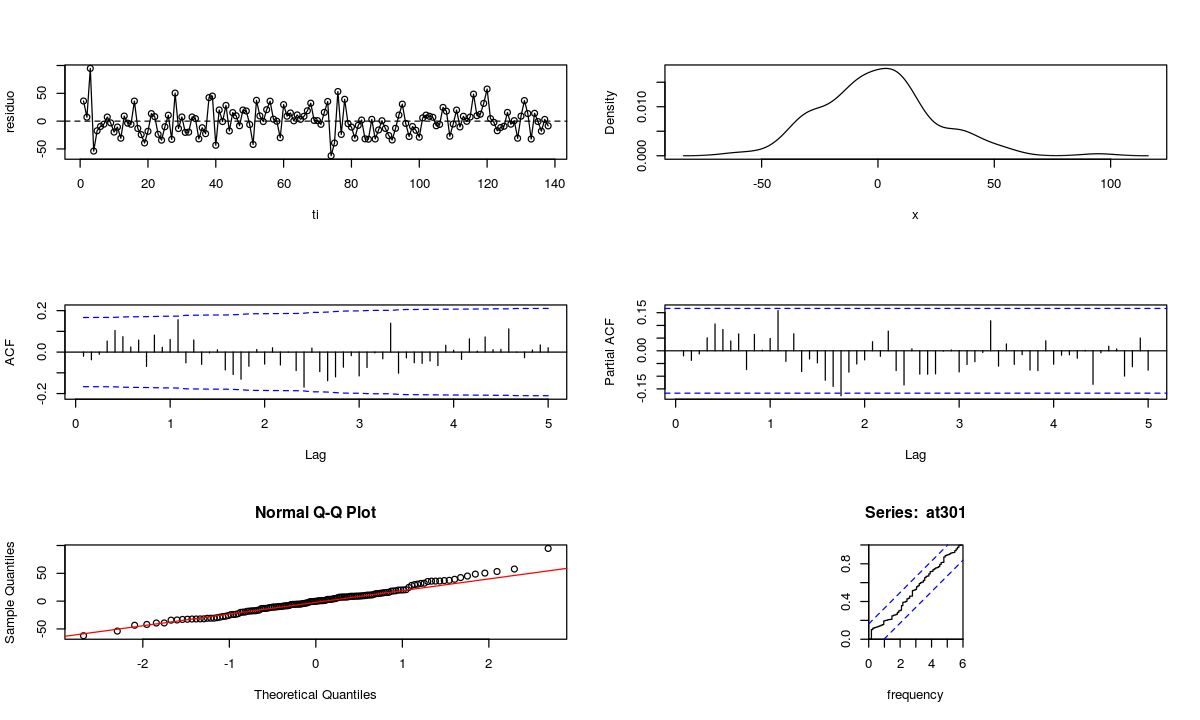


Figura 6 - Validaciones de RB - ARIMA(3,0,1)(3,0,1)[12]

Para este modelo de residuales, se observa una distribución de densidad con cierta simetría y abultamiento al lado derecho, asi como puntos extendidos en la parte derecha de la gráfica, comprobándose en el QQplot. Para la gráfica AFC se observa que todos los rezagos están contenidos dentro de la bandas de Bartlett, siendo el comportamiento esperado para una serie de Ruido Blanco.

1. Christoph Sax, “Switzerland’s data series in one place”, 2017 <http://www.dataseries.org/> [consultado 17 septiembre 2017]. [↑](#footnote-ref-1)
2. Norman Giraldo, *Notas de Clase Series de Tiempo con R*, 2006 p112. [↑](#footnote-ref-2)
3. Norman Giraldo, *Notas de Clase Series de Tiempo con R*, 2006 p.94. [↑](#footnote-ref-3)
4. Norman Giraldo, *Notas de Clase Series de Tiempo con R*, 2006 p.99. [↑](#footnote-ref-4)
5. Norman Giraldo, *Notas de Clase Series de Tiempo con R*, 2006 p.141. [↑](#footnote-ref-5)