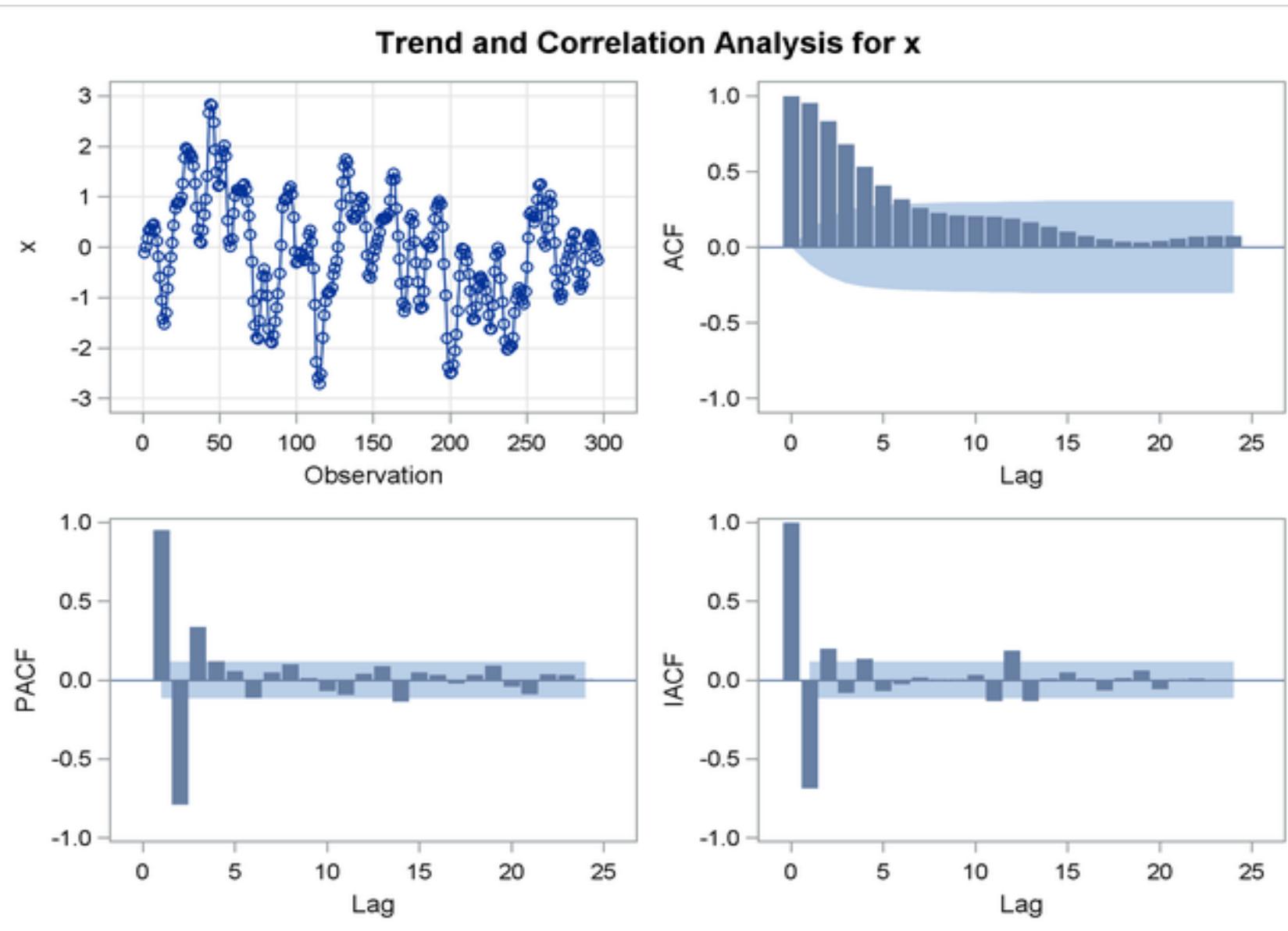


# La metodología Box-Jenkins: los modelos ARIMA



Oscar Centeno Mora

# Preámbulo

- Los modelos ARIMA proporcionan otro enfoque para el pronóstico en las series de tiempo.
- Los modelos de suavizamiento exponencial y los modelos ARIMA son los enfoques clásicos más utilizados para la predicción de series temporales, y proporcionan enfoques complementarios a un problema: mientras que los modelos de suavizamiento exponencial se basan en una descripción de los componentes de la serie, los modelos ARIMA tienen como objetivo explicar las relaciones pasadas en la serie.
- Iniciamos el estudio de los modelos de la segunda generación, los modelos  $ARIMA(p, d, q)(P, D, Q)$  donde el estudio de la serie se lleva a cabo mediante el conocimiento de la asociación de los datos mediante las correlaciones totales y las correlaciones parciales.



# Preámbulo

- En el estudio de los modelos  $ARIMA(p, d, q)(P, D, Q)$ , debemos pasar tanto por los fundamentos de estos para así poder comprender qué es lo que nos impulsa a realizar el proceso analítico de dichos procesos.
- Los presentes modelos estocásticos rompen con cualquier método analítico estudiado hasta ahora. Tomo la siguiente cita sobre la metodología de Box-Jenkins:

*“Esta metodología es diferente a la mayoría de los métodos para generar pronósticos porque no supone ningún patrón en particular en los datos históricos de la series que se van a pronosticar. Se fundamenta en un enfoque iterativo para identificar primeramente un modelo posible a partir de una clase general de modelos. Luego, el modelo seleccionado es puesto a prueba de acuerdo a ciertos estadísticos de bondad y de ajuste. Finalmente, adecuado el modelo, se procede a realizar el pronóstico de la serie en estudio”.*



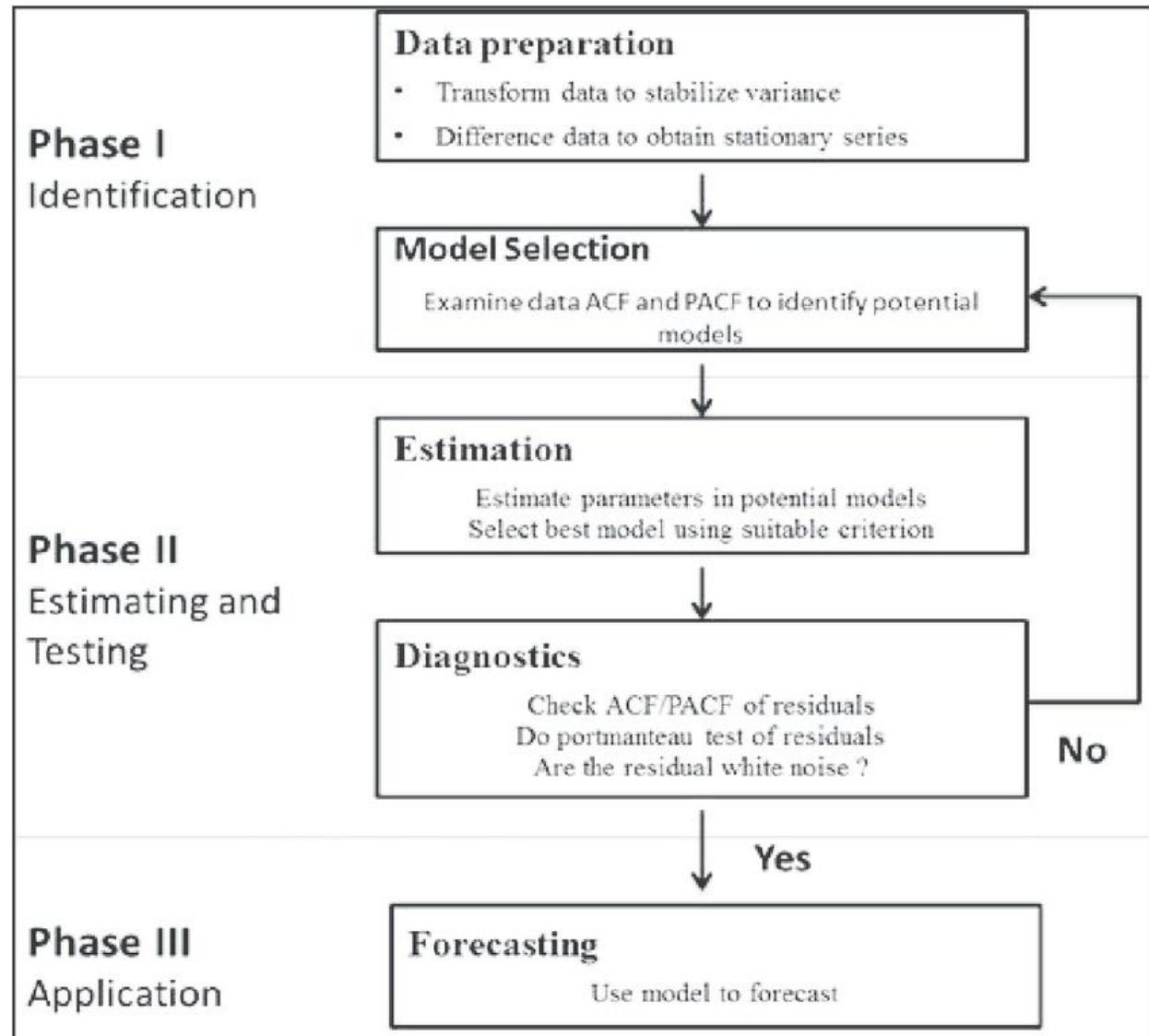
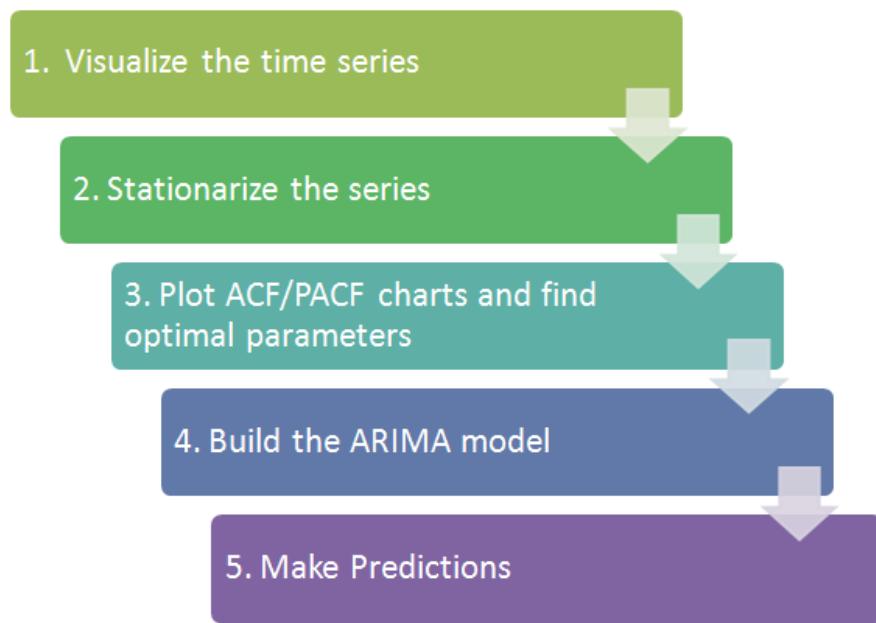
# Preámbulo

- Al estudiar una serie de tiempo mediante la metodología de Box-Jenkins, un modelo ARIMA posee las siguientes etapas:
  1. Descripción de la serie: corroboración de la estacionaridad
  2. Estacionarización la serie.
  3. Identificación del proceso  $ARIMA(p, d, q)(P, D, Q)$
  4. Estimación del proceso  $ARIMA(p, d, q)(P, D, Q)$
  5. Pruebas de bondad y ajuste del modelo (diagnósticos)
  6. Pronóstico.
- Entre los punto 3-4-5, es donde se suele dar la interacción del proceso. De forma muy honesta, no suelo brindar mucha importancia a la corroboración de los supuestos, me fundamento más en aquel modelo que posea los mejores valores de rendimiento... aunque de igual forma los estudiaremos.



# Preámbulo

- La siguiente figura ilustra las 3 principales fases que posee el método de Box-Jenkins en la determinación del modelo  $ARIMA(p, d, q)(P, D, Q)$ .



# Índice

1

Fundamentos del método  
Box-Jenkins

4

Diagnósticos / bondad y  
ajuste del modelo ARIMA

2

La estacionaridad y las  
pruebas de raíz unitaria

5

Pronósticos del  
 $ARIMA(p, d, q)(P, D, Q)$

3

Identificación y  
estimación de un proceso  
ARIMA

6

¿ `auto.arima()` ?

# Índice

7

ARIMA vs ETS

8

Resumen del Box-Jenkins

# Índice

1

Fundamentos del método  
Box-Jenkins

# Índice

1

Fundamentos del método  
Box-Jenkins

Momentos de una  
serie temporal

Los procesos  
estacionarios

El ruido blanco

El teorema de Wold

Los proceso ARMA

Autocorrelaciones:  
total y parcial

El correlograma

# Índice

1

Fundamentos del método  
Box-Jenkins

Momentos de una  
serie temporal

# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

- El objetivo de esta sección es introducir y formalizar la noción de proceso temporal, y más particularmente la clase de procesos ARMA que son particularmente útiles para describir el comportamiento de las series temporales univariadas, al identificar el proceso que les corresponde.
- Para lo anterior, se deben introducir una serie de nociones esenciales para el análisis de las series cronológicas, y, ante todo, las nociones de estacionariedad, ruido blanco, la ecuación de Wold, y las autocorrelaciones totales y parciales.
- El estudio de las series cronológicas mediante la metodología de Box-Jenkins se fundamenta en un cierto número de propiedades probabilísticas, momentos teóricos, convergencias, distribuciones conjuntas y marginales, etc., que no se desarrollaran en su totalidad. Iniciemos volviendo a formalizar lo que entendemos por una serie temporal y la importancia de sus momento teóricos.



# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

Recordemos que en el primer tema habíamos definido qué esta una serie temporal. ¿Qué es lo que habíamos planteado?

Considérese una muestra de realizaciones de  $T$  variables aleatorias indexadas (continuas) en el tiempo, en donde notamos el procese temporal general como  $y_t$ , que va desde  $t = 1, \dots, T$ . El proceso  $y_t$  se compone por:

$$\{y_1, y_2, y_3, \dots, y_T\}$$

Para simplificar, supongamos que la colección de realizaciones de  $T$  variables aleatorias son consideradas *independientes e idénticamente distribuidas* (como sigue lo denominaremos como un proceso *i.i.d*) con una distribución normal. Este proceso *i.i.d* lo podemos denotar como:

$$Y_t \sim N(\mu, \rho^2), \quad t = 1, \dots, T$$

# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

- La muestra observada constituye una realización particular sobre las  $T$  unidades de tiempo de un mismo proceso aleatorio que generan los datos de la variable  $y_t$ . El proceso generado (o DGP por *Data Generating Process*) se considera de esta forma como una *serie estocástica temporal*.
- Entonces, ¿entendemos bien que queremos decir con que  $y_t$  es proceso generado o un DGP? Esto es que  $y_t$  está gobernado o caracterizado por un determina proceso DGP el cuál explica su comportamiento en el tiempo.
- De forma aún más general, el proceso  $y_t$  en el absoluto de todas las posibles realizaciones, se explica por un proceso aún más general que posee un número infinito de períodos. A esta colección de observaciones las podemos escribir de la siguiente forma:

$$\{y_t\}_{t=-\infty}^{\infty} = \{\dots, y_{-1}, y_0, y_1, y_2, \dots, y_T, y_{T+1}, \dots\}$$

# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

- Estas variables aleatorias poseen una función de densidad, notadas como  $f_{Y_t}(y_t)$ , que corresponden a la densidad incondicional de  $y_t$ .
- En el presente caso, esta densidad incondicional está dada por la relación:

$$f_{y_t}(y_t) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-y_t^2}{2\sigma^2}}$$



- Partiendo de la función de densidad de  $y_t$ , es posible determinar el conjunto de momentos de esta variable aleatoria. En el análisis de las series de tiempo, a menudo nos limitamos a la definición de los momentos de orden uno (esperanza) y de orden dos (la función de autocovariancia). Para el presente curso, nos enfocaremos sobre todo al primero momentos, la relación del segundo a nivel de las correlaciones temporales.



# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

- La esperanza teórica del proceso  $Y_t$  está determinada por la fórmula:

$$E(Y_t) = \int_{-\infty}^{\infty} y_t f(y_t) dy_t$$

- Esta integral puede ser aproximada por el límite de probabilidad, denotado  $\text{plim}$ , de la suma infinita siguiente:

$$E(Y_t) = \text{plim}_{I \rightarrow \infty} \sum_{i=1}^I y_t^i$$

- O de forma equivalente:

$$\frac{1}{I} \sum_{i=1}^I y_t^i \underset{I \rightarrow \infty}{\text{p}} E(Y_t)$$

## Fundamentos de la metodología de Box-Jenkins: la serie cronológica

Utilizamos aquí la noción de convergencia en probabilidad (o límite en probabilidad) ya que, recordemos, los términos  $\{y_t^i\}$  son realizaciones de una variable aleatoria, por lo que la suma de estos términos es en sí una variable aleatoria.

Para nuestro interés, cuando las variables  $y_t$ ,  $\forall t \in \mathbb{Z}$ , son independientemente e idénticamente distribuidas (*i.i.d*), entonces podemos determinar la esperanza  $E(Y_t)$  no sólo de las realizaciones de esta variable denominada  $\{y_t^i\}$ , sino también simplemente a partir de las realizaciones de las variables  $y_1, y_2, \dots$ , denotadas  $\{y_t\}_{t=-\infty}^{\infty}$ .

De hecho, bajo la hipótesis i.i.d. estas diferentes variables tienen la misma distribución y en particular el mismo momento de orden uno. Por lo tanto, obtenemos la relación usual:

$$\frac{1}{T} \sum_{t=1}^T y_t \xrightarrow[T \rightarrow \infty]{p} E(Y_t) \quad \forall t \in \mathbb{Z}$$

# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

- En cuanto a los momentos de 2 orden de la variable  $y_t$ , en este curso no nos enfocaremos a la volatilidad del proceso. Estudiaremos el 2nd orden en relación a las posibles correlaciones (o covariancia) entre la variable  $y_t$  y las variables  $y_{t-1}, y_{t-2}, y_{t-3}, \dots$ , entre los valores actuales y pasados de la serie temporal estudiada.
- Esta información (covarianza con valores pasados y futuros) se resume en la función de autocovariancia, denotado generalmente  $\gamma_t(h)$ . La covariancia se escribe como:

$$\gamma_t(h) = E\{[Y_t - E(Y_t)][Y_{t-h} - E(Y_t)]\}$$

- Esto se puede reescribir de la siguiente forma:

$$\begin{aligned}\gamma_t(h) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [y_t - E(y_t)] [y_{t-h} - E(y_{t-h})] \\ &\quad \times f_{Y_t, Y_{t-1}, \dots, Y_{t-h}}(y_t, y_{t-1}, \dots, y_{t-h}) dy_t dy_{t-1} \dots dy_{t-h}\end{aligned}$$

# Fundamentos de la metodología de Box-Jenkins: la serie cronológica

A partir de la función  $\gamma_t(h)$  podemos obtener la variancia. Esta se puede definir de la siguiente forma:

$$\gamma_t(0) = E \left\{ [Y_t - E(Y_t)]^2 \right\} = \text{var}(Y_t)$$

Como para el momento de orden 1, se puede brindar una función en términos del límite en probabilidad:

$$\frac{1}{I} \sum_{i=1}^I [y_t^i - E(Y_t)] [y_{t-h}^i - E(Y_{t-h})] \xrightarrow[I \rightarrow \infty]{p} \gamma_t(h)$$

Cuando las variables  $y_t, t \in Z$  son *i.i.d.*, podemos suponer que el conjunto de los momentos de orden dos son idénticos para todas las variables  $y_t$ , la función de auto covariancia puede definirse por:

$$\frac{1}{T} \sum_{t=1}^T [y_t - E(Y_t)] [y_{t-h} - E(Y_{t-h})] \xrightarrow[T \rightarrow \infty]{p} \gamma_t(h) = \gamma(h) \quad \forall t \in \mathbb{Z}$$

De forma general, la función de autocorrelación puede depender del índice temporal  $t$ , esto es la variable  $y_t$ . Es por eso que se conserva la notación  $\gamma_t(h)$ .

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

# Fundamentos de la metodología de Box-Jenkins: procesos estacionarios

Comenzamos por plantear la definición de un proceso estacionario en sentido estricto (o estacionariedad fuerte) para estudiar las propiedades de la estacionariedad de segundo orden o estacionariedad débil. A partir de esto estudiaremos los procesos estacionarios particulares que son los ruidos blancos.

## La estacionariedad fuerte

Se dice que el proceso  $x_t$  es estrictamente o fuertemente estacionario si para las  $n$ -duplas de tiempo  $t_1 < t_2 < \dots < t_n$ , la serie  $(x_{t_1+h}, \dots, x_{t_n+h})$  posee la misma ley de probabilidad que la serie  $(x_{t_1}, \dots, x_{t_n})$ .

O también se puede definir como

Un proceso es dicho estacionario en el sentido estricto si, para todos los valores  $j_1, j_2, \dots, j_n$  la distribución conjunta de la serie  $(x_t, x_{t+j_1}, \dots, x_{t+j_n})$  depende únicamente de los intervalos de tiempo  $j_1, j_2, \dots, j_n$  y es independiente del periodo  $t$ .

# Fundamentos de la metodología de Box-Jenkins: procesos estacionarios

## La estacionaridad de orden dos o suave

En la práctica, nos limitamos generalmente a utilizar la estacionaridad de segundo orden o estacionaridad suave del proceso en estudio.

**Definición:** un proceso  $(x_t)$  es dicho estacionario de segundo orden o suave si las tres condiciones siguientes son satisfechas:

- $E(x_t^2) < \infty$
- $E(x_t) = m, \text{ independiente de } t$
- $\text{cov}(x_t, x_{t+h}) = E[(x_{t+h} - m)(x_t - m)] = \gamma(h), \text{ independiente de } t$

La primera condición  $E(x_t^2) < \infty$  garantiza simplemente la existencia (o convergencia) de los momentos de orden dos.

La segunda condición  $E(x_t) = m$  testifica sobre los momentos de orden uno y significa que las variables aleatorias  $x_t$  deben poseer la misma esperanza sin importar el tiempo  $t$ . En otras palabras, la esperanza matemática del proceso  $x_t$  debe ser constante e independiente del tiempo  $t$ .

# Fundamentos de la metodología de Box-Jenkins: procesos estacionarios

Finalmente, la tercera condición  $\gamma(h)$ , independiente del tiempo  $t$ , argumenta sobre los momentos de orden 2 resumidos por la función de auto-covariancia. Esto quiere decir que la función de auto covariancia del proceso  $x_t$  debe ser constante e independiente del tiempo  $t$ .



Bajo estas tres condiciones, hay entonces una o más representaciones posibles de la serie temporal  $x_t$ , estas representaciones o procesos tienen la misma distribución en el tiempo (mismos momentos).



Esta suposición de invariancia de la distribución a través del tiempo permite por lo tanto acotarse o apropiarse a una cierta clase de procesos DGP.

En resumen, un proceso es estacionario en el segundo orden si todos sus momentos son independiente del tiempo. Posteriormente, el término de estacionaridad lo utilizaremos para la estacionaridad débil.

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

El ruido blanco

# Fundamentos de la metodología de Box-Jenkins: ruido blanco

- Entre la clase de procesos estacionarios, hay procesos particulares que son los procesos de ruido blanco (*o white noise*). Estos procesos se usan muy a menudo en el análisis de series temporales porque constituyen, en cierto modo, la base de todo el proceso tiempo. De hecho, veremos más adelante que cualquier proceso estacionario puede escribirse como una suma ponderada de ruidos blancos (teorema de Wold).
- Un ruido blanco es un proceso estacionario con incrementos independientes. También hablamos de i.i.d. (variables independientes e idénticamente distribuidas). Así que vamos a empezar definiendo lo que es un proceso con incrementos independientes.
- Un proceso  $x_t$  es un proceso de incrementos independientes si para cualquier n-tupla de tiempo  $t_1 < t_2 < \dots < t_n$ , las variables aleatorias reales  $x_{t_2} - x_{t_1}, \dots, x_{t_n} - x_{t_{n-1}}$  son independientes.
- El proceso  $x_t$  es un proceso estacionario con incrementos independientes si la ley de probabilidad de incrementos ( $x_{t+h} \rightarrow x_t$ ) es independiente de  $t$ . Por lo tanto, es una clase particular de procesos estacionarios.
- Sin embargo, en la práctica la siguiente definición de ruido blanco es generalmente más adoptada que la recién vista.

# Fundamentos de la metodología de Box-Jenkins: ruido blanco

**Definición:** un proceso ( $x_t$ ) es un ruido blanco si satisface las siguientes dos condiciones:

$$E(x_t) = 0$$

$$\gamma(h) = E(x_t x_{t-h}) = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \forall h \neq 0 \end{cases}$$

La primera condición significa que la expectativa del proceso es independiente del tiempo, y además que es nulo. La segunda condición implica independencia de la función de auto-covariancia con respecto al tiempo (estacionariedad). Pero también implica que los términos de las auto-covariancias (para  $h \neq 0$ ) son nulos. Sólo la varianza es distinta de cero.

En otras palabras, esto significa que los ruidos blancos son estacionarios. El nivel de la serie que se considera hoy no tiene ningún impacto en su nivel de mañana, al igual que el nivel de ayer no tiene ningún impacto en el nivel de hoy. Esta es la razón por la que el término "ruido blanco" deriva de la analogía en el dominio de la frecuencia entre la densidad espectral de una variable i.i.d. (constante) y el espectro de la luz blanca en el espectro de color (el término de ruido blanco viene del análisis espectral derivado de la física).

Además, hablamos de ruido blanco gaussiano cuando la ley de probabilidad del proceso es Gaussiana, y lo podemos describir como:

$$\varepsilon_t \quad i.i.d. \sim N(0, \sigma^2)$$

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

El ruido blanco

El teorema de Wold

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

El teorema de Wold (1938) es el teorema fundamental del análisis de las series temporales estacionarias. El teorema de Wold se presenta como sigue:

Todo proceso estacionario  $x_t$  puede ser representado según la forma:

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \kappa_t$$

$\psi_j$  satisface que  $\psi_0 = 1$

$\psi_j \in \mathbb{R}$ ,

$\sum_{j=0}^{\infty} \psi_j^2 < \infty$

$\varepsilon_t$  es un ruido blanco i.i.d.  $(0, \sigma^2)$

Se dice que la suma de los choques pasados corresponden al componente estocástico lineal de  $x_t$ . El término  $\kappa_t$  denota el componente lineal determinístico tal que  $\text{cov}(\kappa_t, \varepsilon_{t-j}) = 0$ .

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

La prueba de este teorema se da en Wold (1938), véase Sargent, Teoría macroeconómica, Boston Academic Press, pp 286-290 para la intuición de la demostración (<http://faculty.wcas.northwestern.edu/~lchrist/finc520/wold.pdf>).

Así, según el teorema de Wold, si omitimos el componente determinista  $\kappa_t$ , cualquier proceso estacionario puede ser escrito como una suma ponderada infinita de choques pasados, representados por un ruido blanco de varianza finita. La fuerte implicación de este teorema es que si se conocen los pesos  $\psi_j$ , y si la varianza  $\sigma_\varepsilon^2$  del ruido blanco es conocida, es posible proponer una representación de cualquier proceso estacionaria. Esta representación también se denomina representación de *media móvil infinita*.

Queda por comprender lo que puede ser este componente lineal determinista  $\kappa_t$ . La condición  $cov(\kappa_t, \varepsilon_{t-j}) = 0$ , implica que este término es, por definición (determinista) independiente de los choques.

Entonces el caso más simple es el de un proceso estacionario  $x_t$  de la expectativa no igual a cero, tal que  $E(x_t) = m \neq 0$ . Dado que el ruido blanco es por definición un proceso centrado, una suma ponderada de estos choques está centrada en sí misma. En consecuencia, la representación de Wold del proceso  $(x_t)$  supone que sumamos a esta suma ponderada de shocks pasados, un componente determinista que no es otro que la expectativa del proceso:  $\kappa_t = m$  (una constante cualquiera).

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

Por lo tanto, se tiene que:

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + m$$

Y podemos verificar que:

$$E(x_t) = E\left[\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right] + m = \sum_{j=0}^{\infty} \psi_j E(\varepsilon_{t-j}) + m = m$$

## Condiciones sobre las ponderaciones

El enunciado del teorema de Wold presupone tres condiciones sobre las ponderaciones de  $\psi_j$ . La primera condición  $\psi_0 = 1$  impone en primer lugar, que el peso del choque presente en  $\varepsilon_t$  en la definición del proceso es unitario. Esto es simplemente una condición de normalización que trata la determinación de la varianza del ruido blanco. Considere el siguiente proceso:

$$x_t = \sum_{j=0}^{\infty} \psi_j v_{t-j} = \left(\frac{1}{2}\right) v_t + \left(\frac{1}{2}\right)^2 v_{t-1} + \left(\frac{1}{2}\right)^3 v_{t-2} + \dots$$

Con  $v_t \sim N(0, \sigma_v^2)$  y  $\sigma_v^2 = 1$

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

Es posible normalizar la varianza del ruido blanco para que el peso del choque presente sea unitario. Para ello hacemos:

$$\varepsilon_t = \frac{1}{2} v_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

El proceso  $x_t$  puede entonces ser expresado en función del ruido blanco  $\varepsilon_t$  de la siguiente forma:

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \varepsilon_t + \left(\frac{1}{2}\right) \varepsilon_{t-1} + \left(\frac{1}{2}\right)^2 \varepsilon_{t-2} + \left(\frac{1}{2}\right)^3 \varepsilon_{t-3} + \dots$$

De esta forma se obtiene la representación de Wold de su proceso  $x_t$ , donde el peso del choque contemporáneo  $\varepsilon_t$  es unitario.

La segunda condición,  $\psi_j$ , aunque totalmente trivial, indica que las ponderaciones de los choques pasados pueden ser eventualmente nulos para ciertos retrasos. En contraposición, la tercera condición,  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ , es muy importante.

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

## Definición: convergencia del proceso

La condición  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ , dicha como sumatoria de cuadrados (o integrabilidad de los cuadrados) asegura la existencia de los momentos de orden dos de proceso ( $x_t, t \in \mathbb{Z}$ ). Bajo esta condición, decimos que  $x_t$  converge en promedio cuadrático.

Para comprender cómo esta condición que la sumabilidad asegura la existencia de momentos teóricos de orden dos, debemos volver a la definición de la función generadora de autocovariancia del proceso  $x_t$  y usar la representación de Wold. Si asumimos por simplicidad que  $E(x_t) = 0$ , obtenemos:

$$\gamma(h) = E(x_{t+h}x_t) = E\left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t+h-j} \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right)$$

Según la definición de un ruido blanco sabemos que  $E(\varepsilon_t \varepsilon_{t-j}) = 0, j \neq 0$ ,  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

Por lo tanto, puesto que la esperanza es un operador lineal, la función  $\gamma(h)$  puede ser reescrita como sigue:

$$\gamma(h) = \begin{cases} \sum_{j=0}^{\infty} \psi_j^2 E(\varepsilon_{t-j}^2) & \text{si } h = 0 \\ \sum_{j=0}^{\infty} \psi_{j+h} \psi_j E(\varepsilon_{t-j}^2) & \forall h \neq 0 \end{cases}$$

De donde se puede obtener mediante simplificación:

$$\gamma(h) = \begin{cases} \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2 & \text{si } h = 0 \\ \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_{j+h} \psi_j & \forall h \neq 0 \end{cases}$$

Se puede demostrar que la condición  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$  implica necesariamente  $\sum_{j=0}^{\infty} \psi_{j+h} \psi_j < \infty, h \neq 0$ . Esta condición **garantiza la convergencia**, et por lo tanto la existencia del total de momentos de orden  $k$  del proceso  $x_t$ .

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

## Previsiones a partir de la descomposición de Wold

Una de las consecuencias del teorema de Wold es que si conocemos los ponderadores  $\psi_j$ , y si además conocemos la variancia  $\sigma_\varepsilon^2$  del ruido blanco, estamos en posesión de representar no importa cual tipo de proceso estacionario ¿Qué queremos decir con esto? Conocemos entonces el proceso que lo gobierna.

A partir de esto podemos a su vez realizar cualquier tipo de previsión, denominada como  $\hat{x}_{T+h}$  para el proceso  $x_t$  a la fecha  $T + h$ , a partir de una muestra de  $T$  observaciones de  $x_t$ . Para comenzar, consideraremos una previsión al periodo 1 ( $h = 1$ ). La mejor previsión posible de  $x_{t+1}$ , denotada por  $\hat{x}_{T+1}$ , sabiendo que se conocen los valores,  $x_T, x_{T-1}, x_{T-2}, \dots$ , está dada por la esperanza condicional  $E(x_{t+1}|x_T, x_{T-1}, x_{T-2}, \dots)$ .

Según el teorema de Wold, el conocimiento de los valores pasados de  $x_t$  es equivalente al conocimiento de los valores pasados de los choques  $\varepsilon_T, \varepsilon_{T-1}, \varepsilon_{T-2}, \dots$ . En efecto sabemos que:

$$x_T = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T-j} = \varepsilon_T + \psi_1 \varepsilon_{T-1} + \psi_2 \varepsilon_{T-2} + \psi_3 \varepsilon_{T-3} + \dots$$

$$x_{T-1} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T-j-1} = \varepsilon_{T-1} + \psi_1 \varepsilon_{T-2} + \psi_2 \varepsilon_{T-3} + \psi_3 \varepsilon_{T-4} + \dots$$

$$x_{T-p} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T-j-p} = \varepsilon_{T-p} + \psi_1 \varepsilon_{T-p-1} + \psi_2 \varepsilon_{T-p-2} + \psi_3 \varepsilon_{T-p-3} + \dots$$

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

Por consecuencia, el espacio de condicionamiento de nuestra previsión puede ser modificado de la siguiente forma:

$$\hat{x}_{T+1} = E(x_{T+1}|x_T, x_{T-1}, x_{T-2}, \dots) = E(x_{T+1}|\varepsilon_T, \varepsilon_{T-1}, \varepsilon_{T-2}, \dots)$$

Ahora lo debemos enmarcar en la representación de Wold del proceso  $x_{T+1}$ :

$$x_{T+1} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T-j+1} = \varepsilon_{T+1} + \psi_1 \varepsilon_T + \psi_2 \varepsilon_{T-1} + \psi_3 \varepsilon_{T-2} + \dots$$

A partir de lo anterior, se puede determinar la esperanza condicional, y entonces el pronóstico para  $h = 1$ . Vamos a simplificar bastante para llegar finalmente al resultado esperado:

$$\hat{x}_{T+1} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T-j+1} = \psi_1 \varepsilon_T + \psi_2 \varepsilon_{T-1} + \psi_3 \varepsilon_{T-2} + \dots$$

Y por consiguiente el error de previsión:

$$x_{T+1} - \hat{x}_{T+1} = \varepsilon_{T+1}$$

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

Por definición, este error de predicción es de cero y de varianza igual  $\sigma_\varepsilon^2$ . Por lo tanto, para un proceso estacionario, el error óptimo de predicción en un horizonte de período corresponde a la innovación fundamental (choque) asociada a este período. Además, este error de predicción tiene "buenas" propiedades, ya que es estrictamente independiente del error de predicción que podría haberse cometido en las fechas anteriores, ya que  $E(\varepsilon_T, \varepsilon_{T-j}) = 0$ .

Estos resultados se pueden generalizar para cualquier orden de previsión. Los resultados siguientes se muestran de la misma manera. Sea  $\hat{x}_{T+h}$  la mejor previsión posible para el proceso  $x_t$  para la fecha  $T + h$ . Su predicción está dada por la ecuación:

$$\hat{x}_{T+h} = \sum_{j=1}^{\infty} \psi_j \varepsilon_{T-j+1}$$

El error de previsión asociado sería entonces:

$$x_{T+h} - \hat{x}_{T+h} = \sum_{j=1}^{\infty} \psi_j \varepsilon_{T+h-j}$$

Para saber más sobre el teorema de Wold, ver el siguiente enlace:  
<http://faculty.wcas.northwestern.edu/~lchrist/finc520/wold.pdf>

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

## El operador de retraso

El teorema de Wold se da a menudo introduciendo un polinomio definido por el operador del retraso. Más generalmente, los modelos de series temporales se expresan a menudo como un polinomio de retraso. Por lo tanto, comenzaremos definiendo lo que es el operador de retraso, entonces re-expresaremos el teorema de Wold usando un polinomio de retraso.

El operador de retraso (denotado  $L$  por retraso o  $B$  según las obras) se define como sigue: Consideramos un proceso estocástico ( $x_t$ ), el operador de retraso, denotado  $L$ , se define por la relación:

$$Lx_t = x_{t-1} \quad \forall t \in \mathbb{Z}$$

Este operador permite así definir una aplicación que a cualquier variable  $x_t$  asocia la variable retrasada (de ahí su nombre)  $x_{t-1}$ .

Partiendo de la definición de este operador de retraso, entonces es posible construir polinomios con coeficientes reales en este operador. Por ejemplo, considere un polinomio  $\Psi(L)$  de orden  $q$ :

$$\Psi(L) = \psi_0 L^0 + \psi_1 L^1 + \psi_2 L^2 + \cdots + \psi_q L^q$$

# Fundamentos de la metodología de Box-Jenkins: teorema de Wold

La aplicación del polinomio a un proceso aleatorio ( $\varepsilon_t$ ) nos brinda la ecuación:

$$\begin{aligned}\Psi(L)\varepsilon_t &= \psi_0 L^0 \varepsilon_t + \psi_1 L^1 \varepsilon_t + \psi_2 L^2 \varepsilon_t + \cdots + \psi_q L^q \varepsilon_t \\ \Psi(L)\varepsilon_t &= \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots + \psi_q \varepsilon_{t-q}\end{aligned}$$

- Si  $\psi_0 = 1$  y  $q \rightarrow \infty$  encontramos aquí una forma de media móvil infinita idéntica a la del teorema de Wold. La nueva afirmación de este teorema (idéntica, por supuesto, a la precedente) es la siguiente: todo proceso estacionario puede ser representado como:

$$xt = \Psi(L)\varepsilon t + kt$$

donde el polinomio de grado infinito está definido por  $\Psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$ ,  
los parámetros  $\psi_j$  satisfacen que  $\psi_0 = 1, \psi_j \in R, \forall j \in N^*, \sum_{j=0}^{\infty} \psi_j^2 < \infty$ , y donde  
 $\varepsilon t$  es un ruido blanco i.i.d.  $(0, \sigma_{\varepsilon}^2)$ ,

El componente lineal determinístico  $kt$  verifica que  $cov(k_t, \varepsilon_{t-j}) = 0, \forall j \in Z$ .

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

El ruido blanco

El teorema de Wold

Los proceso ARMA

# Fundamentos de la metodología de Box-Jenkins: MA, AR y ARMA

- Hasta ahora hemos visto que cualquier proceso estacionario podría escribirse en la forma de una suma ponderada infinita de choques pasados (teorema de Wold). La descomposición de Wold es una primera representación posible de una serie temporal utilizando la información pasada ya sea en sus valores o errores pasados. Sin embargo, esta representación nunca es la representación óptima de todas las representaciones del mismo proceso (¡¿CÓMO?!).
- Si se aplicara la descomposición de Wold, esto implicaría que estimamos una infinidad de parámetros (el  $\psi_j$  y el  $\sigma_\varepsilon^2$ ). Por lo tanto, en la práctica, deben buscar otras representaciones para los procesos temporales más eficientes sin que caigamos en la sobre parametrización del proceso que se desea determinar.
- Entre las representaciones más usadas se encuentran las representaciones ARMA que son explicaciones **Auto Regresivas de Medias Móviles**. Esta representación consiste en la adición de un componente Autorregresivo de orden finito (AR) y un componente de media móvil de orden finito (MA). Los procesos AR, MA y ARMA se trabajan bajo procesos  $x_t$  que son estacionarios. Veamos cómo es que los representamos

# Fundamentos de la metodología de Box-Jenkins: MA, AR y ARMA

## Los procesos MA

Un proceso estacionario  $x_t$  satisface una representación *MA* de orden  $q$ , denotado como  $MA(q)$ , si y solo si:

$$x_t = c + \Theta(L) \varepsilon_t$$

Con  $E(x_t) = c$ .

El polinomio  $\Theta(L)$  se define como  $\Theta(L) = \sum_{j=0}^q \theta_j L^j$ , donde

1.  $\forall j < q$ ,
2.  $\theta_j \in \mathbb{R}$ ,
3.  $\theta_0 = 1$  y
4.  $\theta_q \in \mathbb{R}^*$ ,
5.  $\varepsilon_t$  que está i.i.d.  $(0, \sigma_\varepsilon^2)$

# Fundamentos de la metodología de Box-Jenkins: MA, AR y ARMA

## Los procesos AR

El proceso estacionario  $x_t$  satisface una representación *AR* de orden  $p$ , denotado como  $AR(p)$ , si y solo si:

$$\Phi(L) xt = \varepsilon t$$

- El polinomio  $\Phi(L)$  se define como  $\Phi(L) = \sum_{j=0}^q \phi_j L^j$ , donde

1.  $\forall j < p$ ,
2.  $\phi_j \in \mathbb{R}$ ,
3.  $\phi_0 = 1$
4.  $\phi_q \in \mathbb{R}^*$ ,
5.  $\varepsilon t$  que está i.i.d.  $(0, \sigma_\varepsilon^2)$

# Fundamentos de la metodología de Box-Jenkins: MA, AR y ARMA

## Los procesos ARMA

El proceso estacionario  $x_t$  satisface una representación *ARMA* de orden  $p$  y  $q$ , denotado como  $ARMA(p, q)$ , si y solo si

$$\Phi(L)x_t = c + \Theta(L)\varepsilon_t$$

Con  $c \in \mathbb{R}$ .

$$\Phi(L) = \sum_{j=0}^q \phi_j L^j$$

$$\Theta(L) = \sum_{j=0}^q \theta_j L^j \text{ en donde:}$$

1.  $\forall j < p$ ,
2.  $(\phi_j, \theta_p) \in \mathbb{R}^2$ ,
3.  $(\phi_0, \theta_0) = 1$
4.  $(\phi_q, \theta_p) \in \mathbb{R}^2$ ,
5.  $\varepsilon_t$  que está i.i.d.  $(0, \sigma_\varepsilon^2)$

# Fundamentos de la metodología de Box-Jenkins: MA, AR y ARMA

- Por lo tanto, para un proceso estacionario  $x_t$ , vamos a aproximar el proceso que gobierna la serie  $y_t$  mediante un modelo Autoregresivo (AR), un modelo de Medias Móviles (MA), o un modelo Autoregresivo de Medias Móviles (ARMA).



- Surgen tres preguntas:

1. ¿Cómo es que procedemos con la identificación de estos modelos?
2. Pero antes, ¿cómo pasamos de un proceso no estacionario  $y_t$  a un proceso estacionario?
3. Y finalmente, ¿estos modelos contemplan la estacionalidad de los datos?



- Todo a su debido tiempo... contestemos antes como identificaciones estos procesos, luego como obtenemos una serie estacionaria, y finalmente resolvemos el problema de la estacionalidad en los procesos ARMA.

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

El ruido blanco

El teorema de Wold

Los proceso ARMA

Autocorrelaciones:  
total y parcial

# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

La identificación de un proceso AR( $p$ ), MA( $p$ ) o ARMA( $p,q$ ), se llevan a cabo mediante el conocimiento del proceso que lo genera. Para poder saber esto, tenemos que, a partir de una serie estacionaria, saber como está su relación total e intermediaria respecto a los valores que la generan. Necesitamos para esto conocer para esto las autocorrelaciones parciales y totales del proceso como tal.

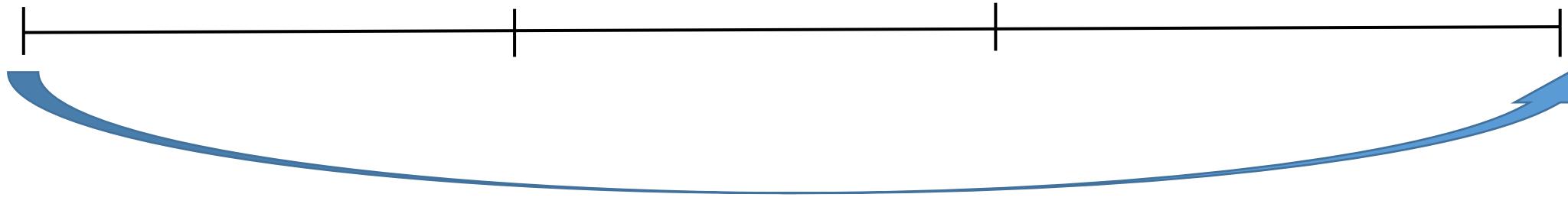
¿Qué son entonces las autocorrelaciones parciales y totales? De inmediato lo explicamos.



# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

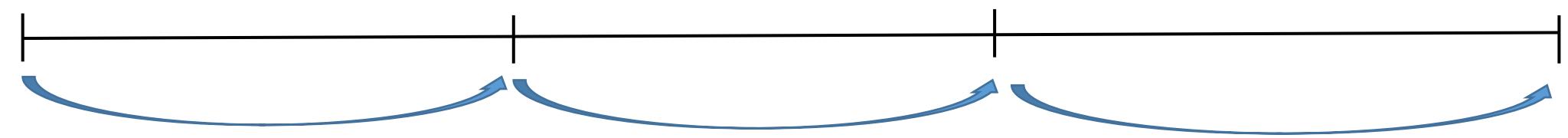
Autocorrelación  
total

$x_t$                      $x_{t+1}$                      $x_{t+2}$                      $x_{t+3}$



Autocorrelación  
parcial

$x_t$                      $x_{t+1}$                      $x_{t+2}$                      $x_{t+3}$



# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

La correlación entre  $x_t$  y  $x_{t+h}$  proviene de dos fuentes: primeramente sobre la correlación que puede existir entre  $x_t$  y  $x_{t+h}$ , *luego de las correlaciones intermedias de  $x_{t+1}, x_{t+2}, \dots, x_{t+h}$*  por lo tanto, el valor de la función de autocorrelación toma tanto la forma total como parcial en el análisis. Estudiemos primeramente la correlación total (ACF).

La ACF trata de apreciar la dependencia existente entre las observaciones que constituyen una trayectoria del proceso. El uso de un coeficiente de correlación aparece entonces naturalmente indicado, de ahí el valor de la función de autocorrelación entre dos momentos.

- La función de autocorrelación total está en función de  $k$ , el valor de la correlación entre  $x_t$  y  $x_{t+k}$ , se rige por la siguiente fórmula:

$$\rho_k = \frac{Cov(x_t, x_{t+k})}{[V(x_t)]^{1/2}[V(x_{t+k})]^{1/2}} = \frac{\gamma_k}{\gamma_0}$$

- Observamos que la simetría de la función de autocovariancia conduce a la simetría de la función de la autocorrelación:  $\rho_k = \rho_{-k}$

# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

- Los coeficientes  $\rho_k$  son generalmente no observables. Las estimaciones utilizadas son los coeficientes de correlación empíricos usuales,  $r_k$ , que poseen por ecuación:

$$r_k = \frac{\sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2}$$

- Para procesos gaussianos, Bartlett [1946] derivó los valores asintóticos de las variancias y covariancias de los estimadores. Estos dependen de forma compleja de los términos  $\rho_k$ ,  $k = -\infty, \dots, \infty$ :

$$V(r_k) \approx \frac{1}{T} \sum_{j=-\infty}^{+\infty} (\rho_j^2 + \rho_{k+j}\rho_{-k+j} - 4\rho_k\rho_j\rho_{-k+j} + 2\rho_j^2\rho_k^2), \text{ y}$$

$$\text{Cov}(r_k, r_{k+s}) \approx \frac{1}{T} \sum_{j=-\infty}^{+\infty} \rho_j \rho_{j+s}$$

# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

- La autocorrelación parcial mide la relación intermediaria entre las observaciones de  $k$  periodos una vez tomadas en cuenta las consideraciones de las relaciones entre las observaciones de 1,2,...,  $k-1$  periodos. Siendo la correlación una medida de asociación linear, es natural evaluar esta correlación parcial por el coeficiente de  $x_{t-k}$  en la regresión de  $x_t$  para  $x_{t-1}, x_{t-2}, \dots, x_{t-k}$ .
- Esta función asocia entonces a los  $k$  valores, los coeficientes  $\phi_{kk}$  que tienen por escritura:

$$x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \cdots + \phi_{kk}x_{t-k} + u_t \quad \text{para } k = 1, 2, \dots, K.$$

Los valores  $\phi_{kk}$  se pueden obtener mediante el método de estimación por MCO. Podemos, suponiendo que el proceso es estacionario, dar una representación a la relación entre autocorrelaciones parciales y autocorrelaciones totales, una representación conocida como las ecuaciones de Yule-Walker: de hecho, a partir de la escritura anterior deducimos el término general de orden  $i$  ( $i > 0$ ) que la función de autocorrelación de un modelo autorregresivo de orden  $k$  está dada por la ecuación:

$$\begin{aligned}\gamma_i &= E[\phi_{k1}x_{t-1}x_{t-i} + \phi_{k2}x_{t-2}x_{t-i} + \cdots + \phi_{kn}x_{t-n}x_{t-i} + \cdots + u_tx_{t-i}] \\ &= \phi_{k1}\gamma_{i-1} + \phi_{k2}\gamma_{i-2} + \cdots + \phi_{kn}\gamma_{i-n} + \cdots\end{aligned}$$

# Fundamentos de la metodología de Box-Jenkins: autocorrelación total

Sea entonces, la ecuación puesta en términos de la función de auto correlación:

$$\rho_i = \phi_{k1}\rho_{i-1} + \phi_{k2}\rho_{i-2} + \cdots + \phi_{kn}\rho_{i-n} + \cdots$$

Obtenemos , haciendo variar  $i=1,2,\dots,k$ , el conjunto de ecuaciones de Yule-Walker:

$$\rho_1 = \phi_{k1} + \phi_{k2}\rho_1 + \cdots + \phi_{kn}\rho_{n-1} + \cdots$$

$$\rho_2 = \phi_{k1}\rho_1 + \phi_{k2} + \cdots + \phi_{kn}\rho_{n-2} + \cdots$$

$$\rho_3 = \phi_{k1}\rho_2 + \phi_{k2}\rho_1 + \cdots + \phi_{kn}\rho_{n-3} + \cdots$$

⋮

$$\rho_k = \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \cdots + \phi_{kn}\rho_{k-n} + \cdots$$

Las autocorrelaciones parciales provienen de la explicación intermediaria entre las diferentes correlaciones. Estas se suelen representar en el sistema matricial de Yule-Walker como:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}$$



Las autocorrelaciones parciales se obtienen a partir del despeje del vector  $\phi$

# Índice

1

Fundamentos del método  
Box-Jenkins

¿Qué es una serie  
cronológica?

Los procesos  
estacionarios

El ruido blanco

El teorema de Wold

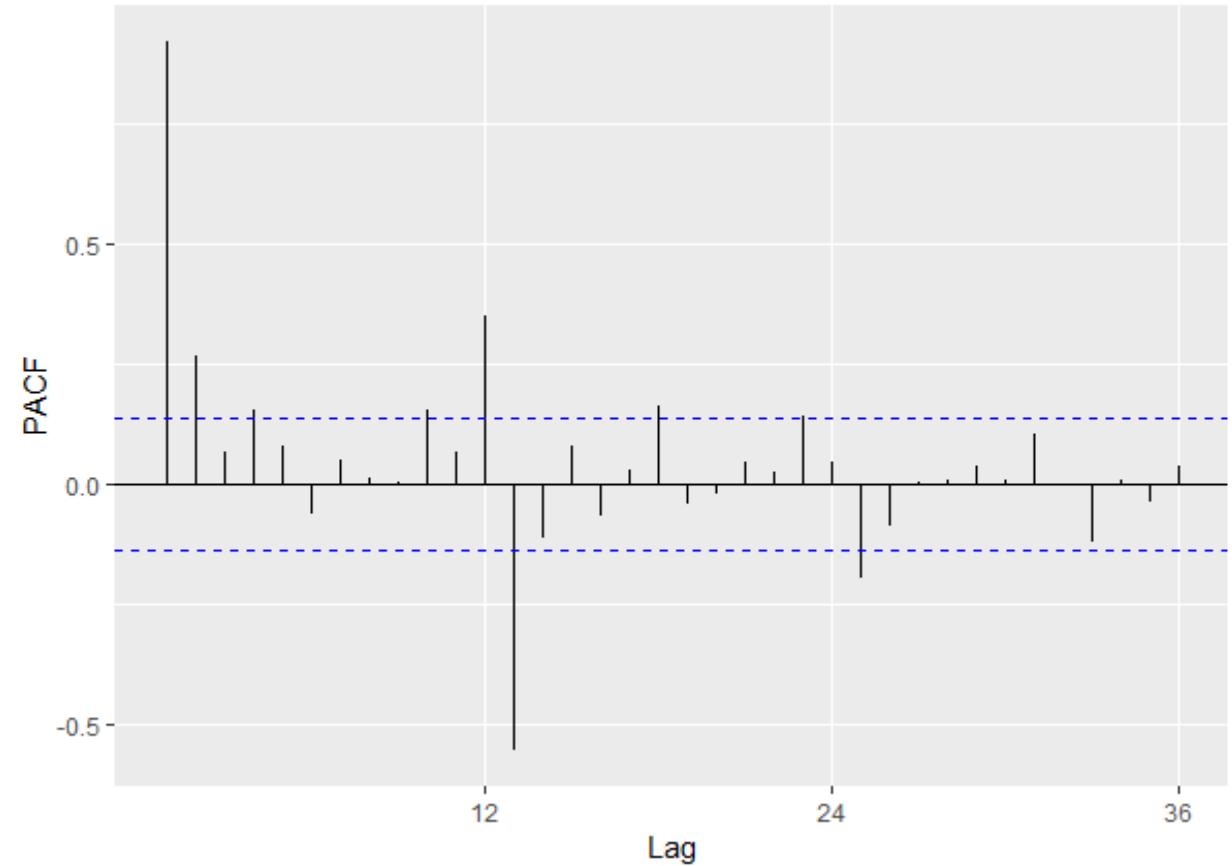
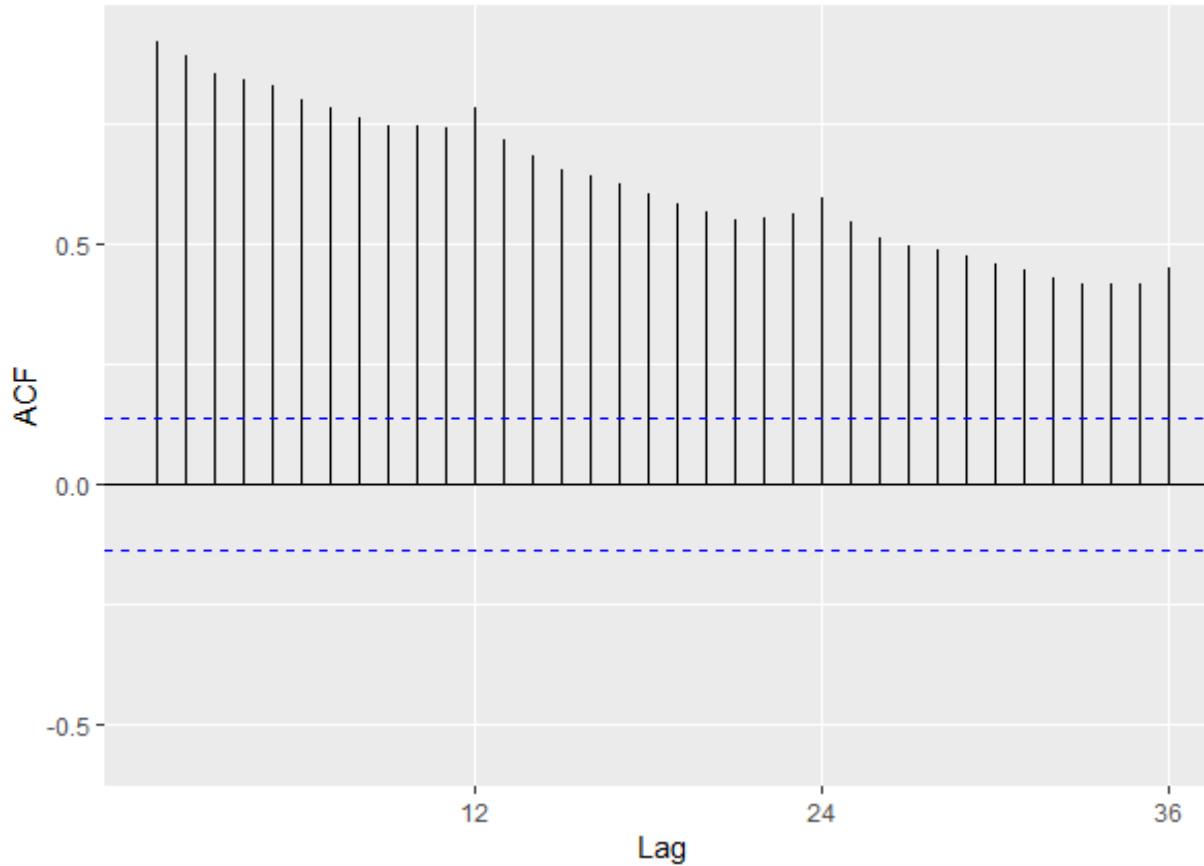
Los proceso ARMA

Autocorrelaciones:  
total y parcial

El correlograma

# Fundamentos de la metodología de Box-Jenkins: el correlograma

Finalmente, para identificar cuál es el proceso que es gobernado por la serie, utilizamos el correlograma, el cual muestra tanto las autocorrelaciones totales y parciales para así determinar cuál es el proceso de la serie. Sobre este, hablaremos más adelante en la sección de identificación y estimación del  $ARIMA(p, d, q)(P, D, Q)$ .



# Índice

1

Fundamentos del método  
Box-Jenkins

2

La estacionaridad y las  
pruebas de raíz unitaria

# Índice

1

Fundamentos del método  
Box-Jenkins

2

La estacionaridad y las  
pruebas de raíz unitaria

La estacionaridad

Pruebas de raíz  
unitaria

# Índice

1

Fundamentos del método  
Box-Jenkins

2

La estacionaridad y las  
pruebas de raíz unitaria

La estacionaridad

# La estacionaridad

En secciones anteriores habíamos definido la estacionaridad de la siguiente forma:

**Definición:** un proceso  $(x_t)$  es dicho estacionario o suave si las tres condiciones siguientes son satisfechas:

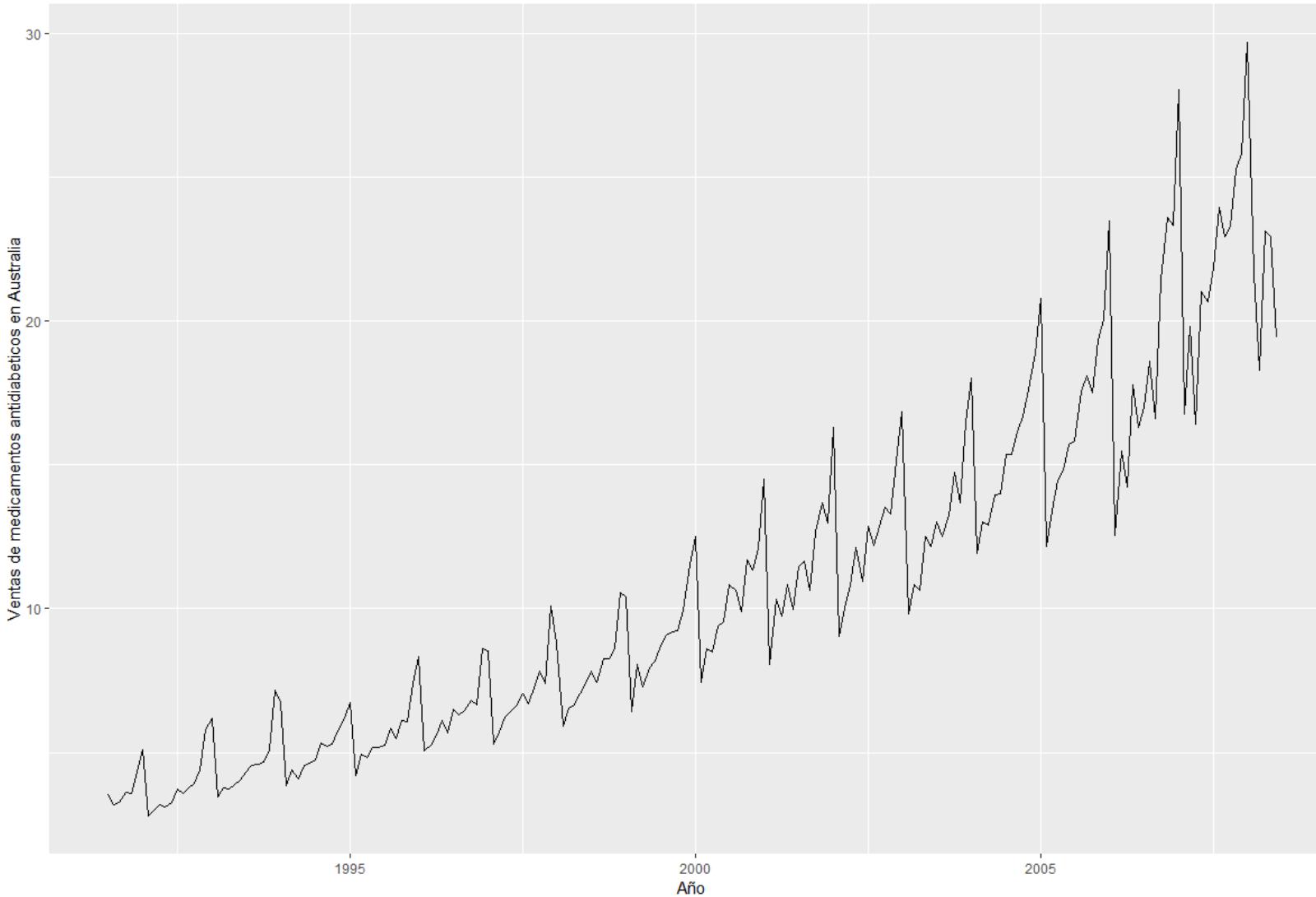
- $E(x_t^2) < \infty$
- $E(x_t) = m,$  *independiente de t*
- $\text{cov}(x_t, x_{t+h}) = E[(x_{t+h} - m)(x_t - m)] = \gamma(h),$  *independiente de t*

La primera condición  $E(x_t^2) < \infty$  garantiza simplemente la existencia (o convergencia) de los momentos de orden dos. La segunda condición  $E(x_t) = m$  testifica sobre los momentos de orden uno y significa que las variables aleatorias  $x_t$  deben poseer esperanza constante sin importar el tiempo  $t$ . En otra palabras, la esperanza matemática del proceso  $x_t$  debe ser independiente del tiempo  $t$ . Finalmente, la tercera condición  $\gamma(h)$ , independiente del tiempo  $t$ , argumenta sobre los momentos de orden 2 resumidos por la función de autocovariancia. Esto quiere decir que la función de auto covariancia del proceso  $x_t$  debe ser independiente / constante del tiempo  $t$ .

A grosso modo, en la práctica, para obtener una serie estacionaria, debemos asegurarnos que la serie posea tanto una media constante y una variancia constante en el tiempo.

# La estacionaridad

Analicemos el archivo a10 de la librería ffp2. Estas son ventas de antibióticos. ¿Es la siguiente serie estacionaria?

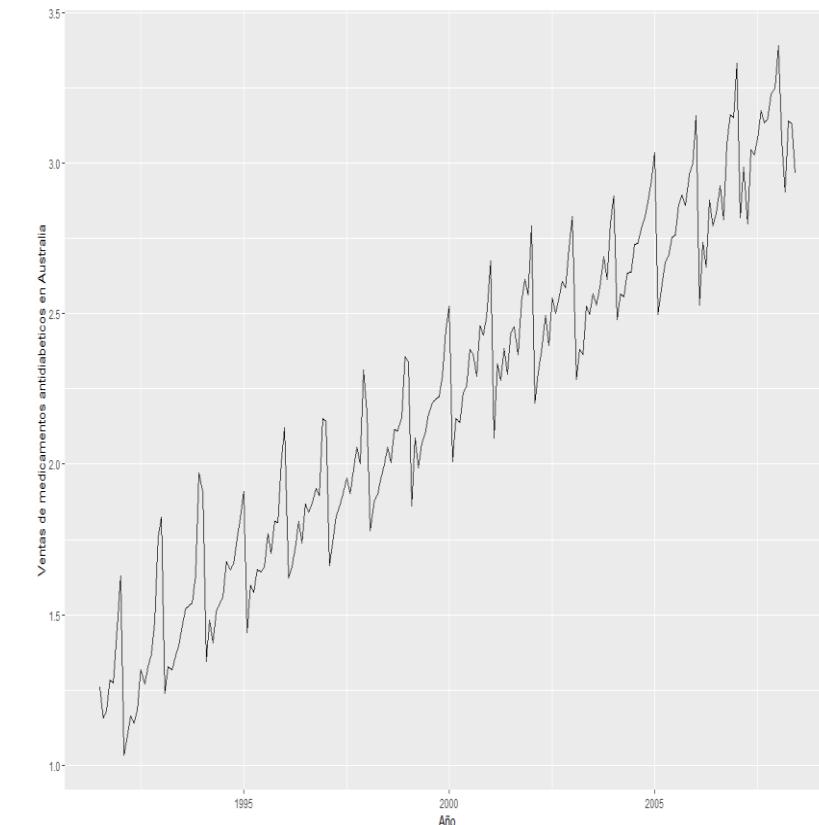


Explique las tres formas en que **NO** se cumple la estacionaridad

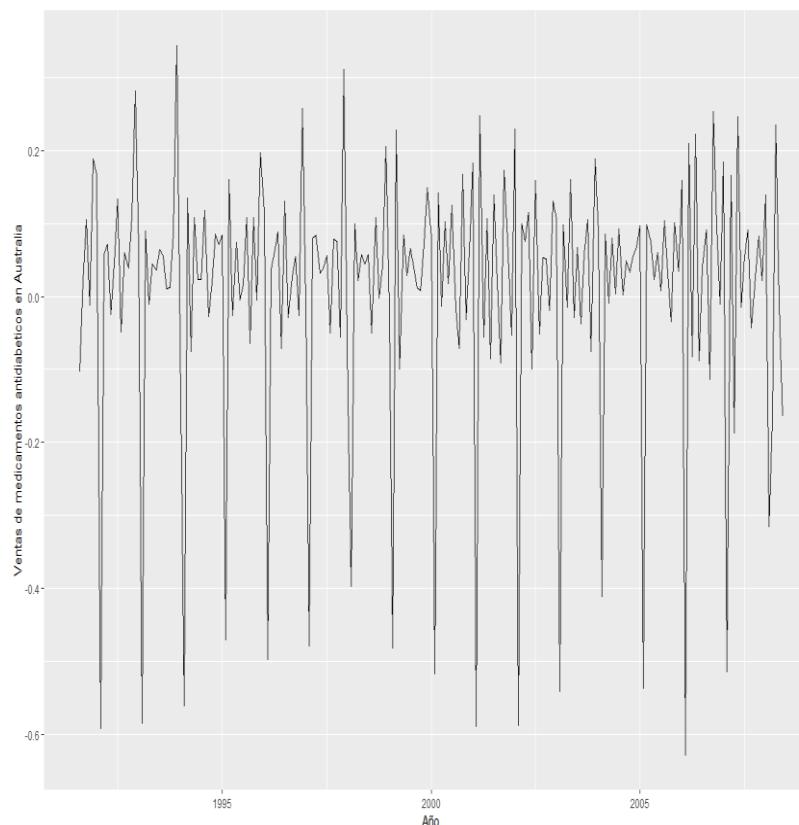
# La estacionaridad

Para obtener estacionaridad, debemos tener primero una serie con variancia constante, luego quitarle el efecto de la tendencia (diferenciación) y finalmente no seguir viendo correlaciones de las observaciones en el tiempo (diferenciar en la periodicidad  $s$ ). Veamos entonces las 3 series.

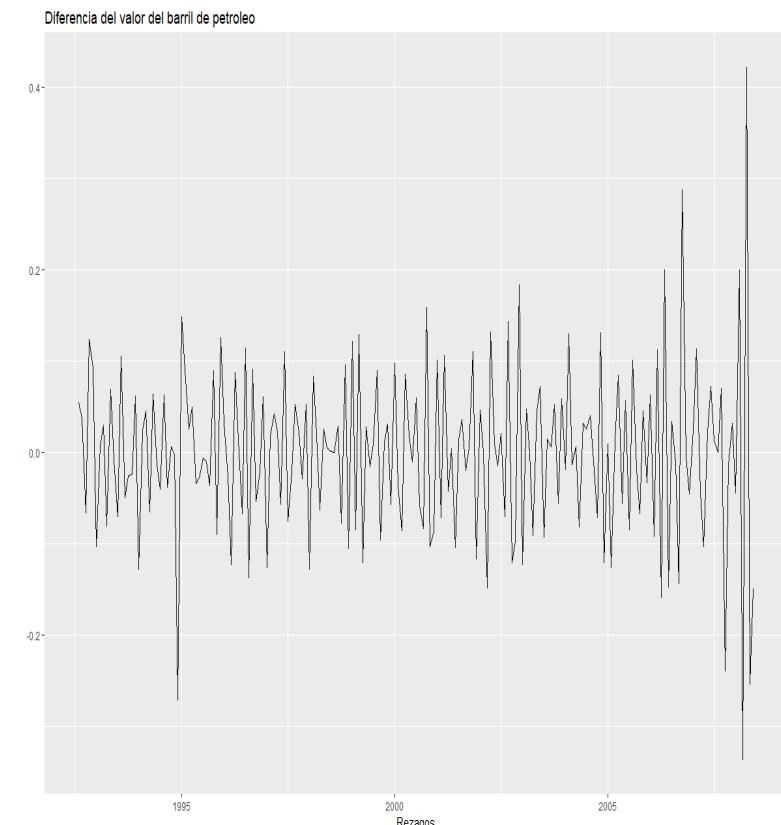
Serie con log



Serie con log y diferencia 1



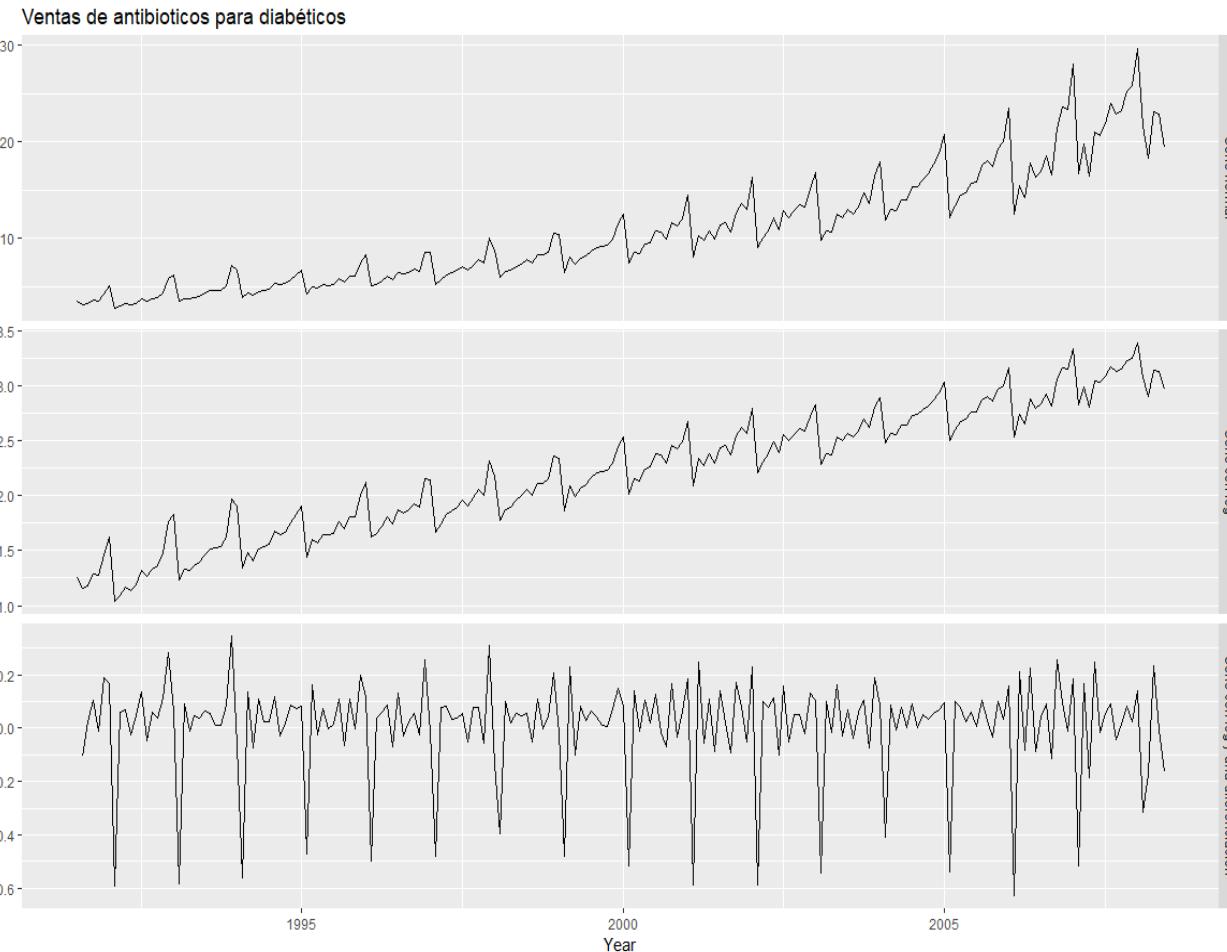
Serie con log y diferencia 1 y 12



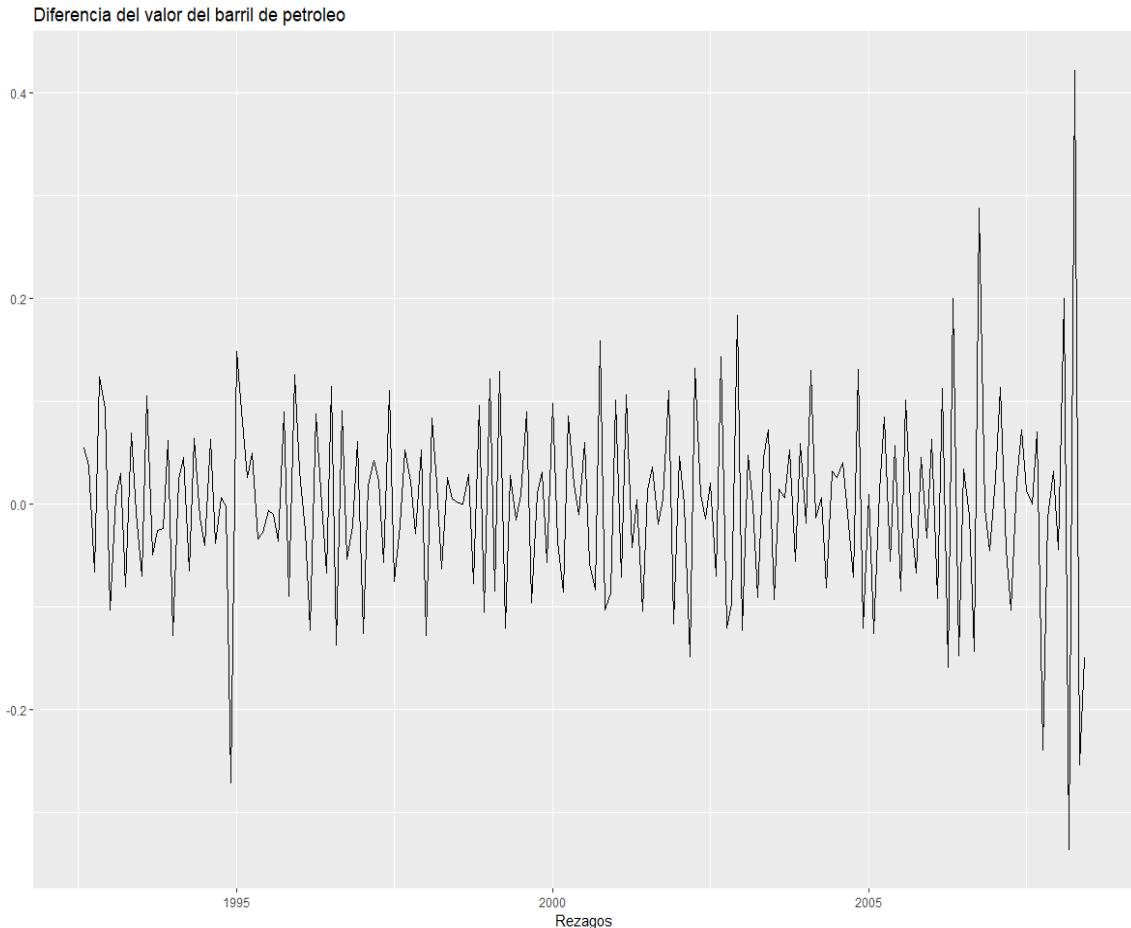
# La estacionaridad

Finalmente, veamos el efecto de pasar de una serie  $y_t$  sin estacionarizar hasta llegar a la serie  $x_t$  estacionarizada.

Serie no estacionaria  $y_t$

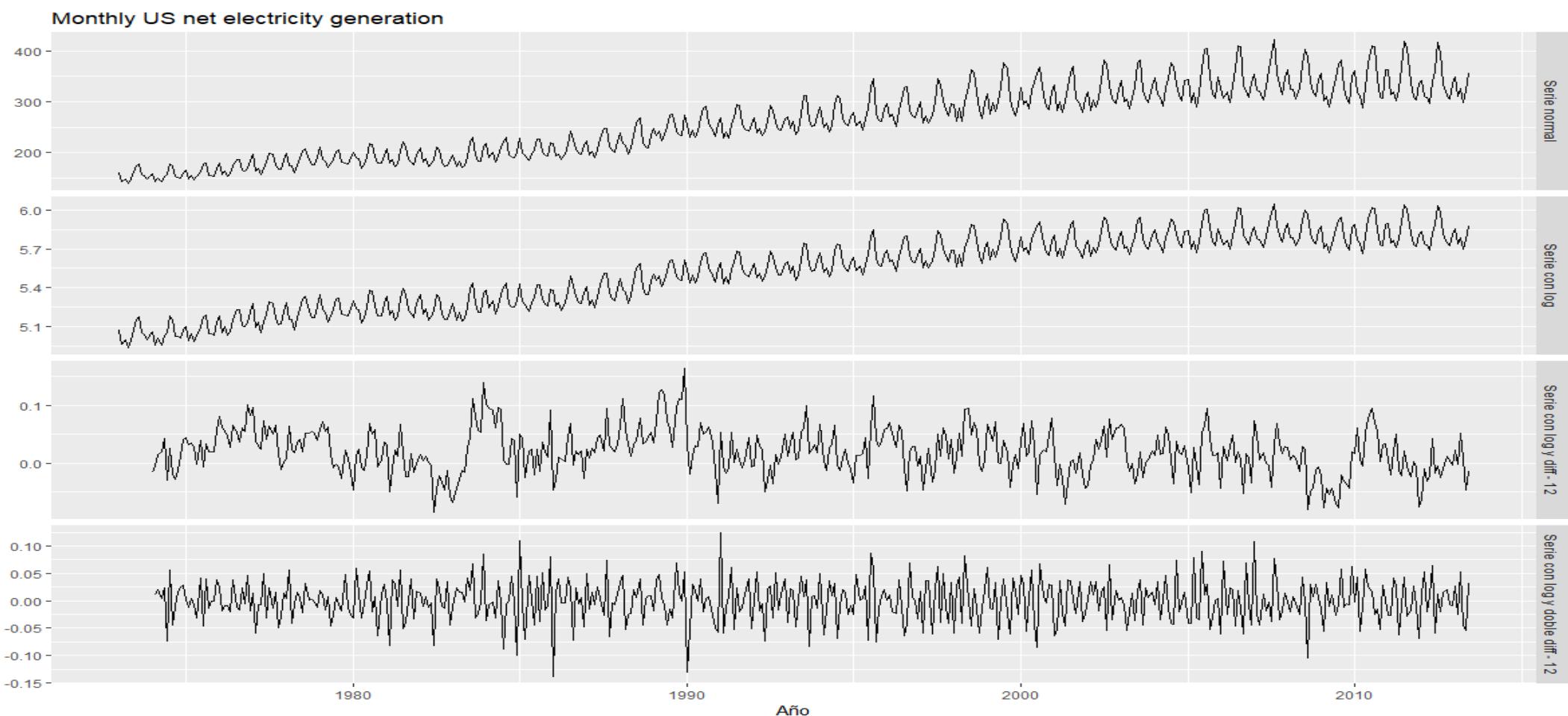


Serie estacionaria  $x_t$



# La estacionaridad

Tenemos que asegurarnos que efectivamente todas las posesiones de la estacionaridad se cumplan. Muchas veces el aplicar una única diferenciación de orden 1, no produce realmente una serie estacionaria, y se podría optar por volver a re-diferenciar la serie para evitar los diferentes niveles en la serie. Otras veces simplemente pensamos que por hacer una de las dos diferenciaciones, estamos cumpliendo el supuesto de la media constante. Véase el siguiente ejemplo:



# La estacionaridad

En estos momentos, seguro se preguntan : ¿no será posible corroborar la estacionaridad también mediante estadísticos con pruebas de hipótesis asociadas? Pues claro, para eso debemos hacer pruebas a las raíces unitarias. Pasemos es esto de inmediato.



# Índice

1

Fundamentos del método  
Box-Jenkins

2

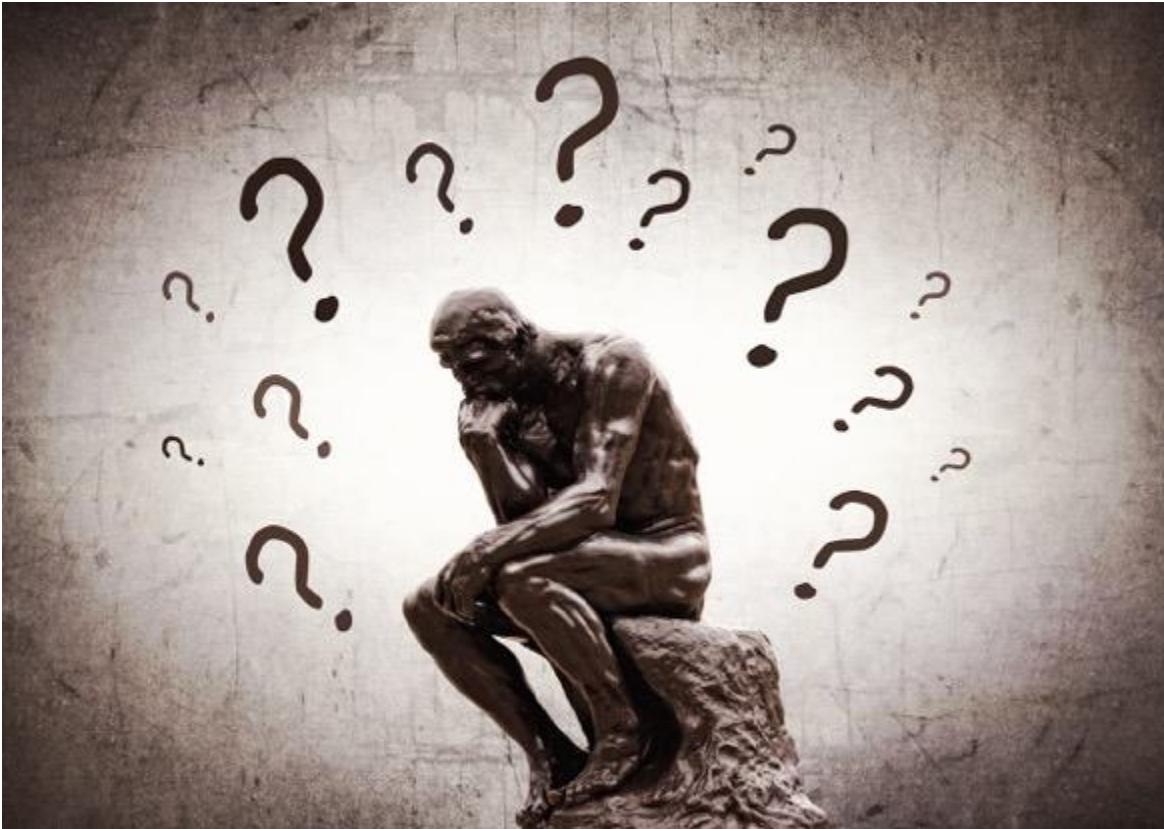
La estacionaridad y las  
pruebas de raíz unitaria

La estacionaridad

Pruebas de raíz  
unitaria

# Pruebas de raíz unitaria

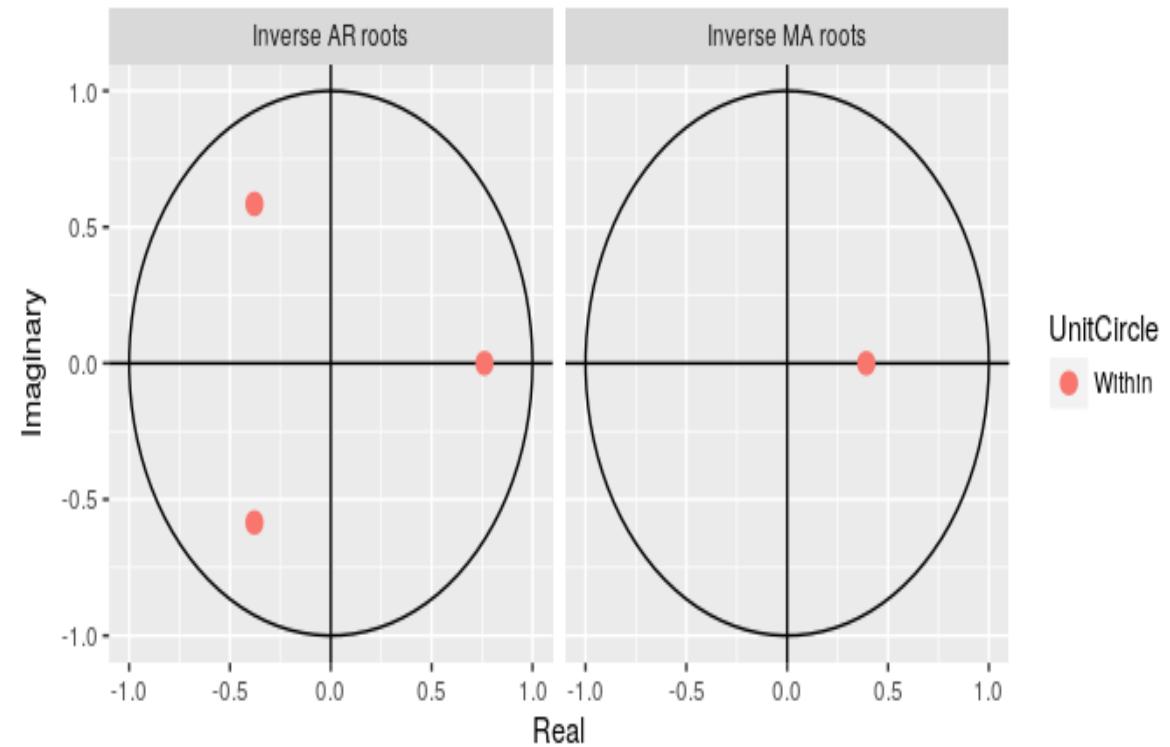
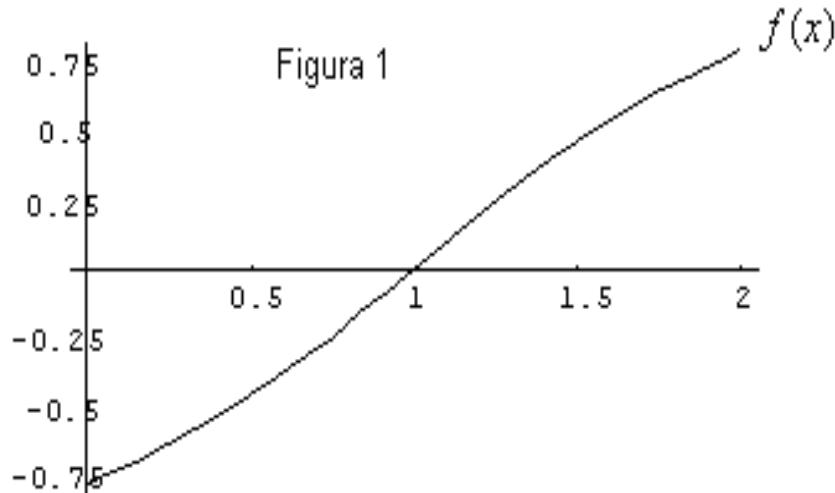
¿Qué es una raíz unitaria?



# Pruebas de raíz unitaria

Cuando se tiene una ecuación, las soluciones encontradas son llamadas raíces unitarias y corresponden a aquellos valores de  $x$  que hacen igual a cero la ecuación. La raíz de la ecuación es el punto donde la función cruza el eje x. Una raíz que es igual a 1 en valor absoluto es llamada raíz unitaria. La raíz unitaria también puede ser presentada o explicada dentro del círculo trigonométrico. Dentro del presente contexto, la presencia de raíces unitarias en el proceso  $y_t$ , son las responsables de que el proceso no sea estacionario, y producen que realmente nuestro proceso sea un Random Walk como estimación (<https://www.ssc.wisc.edu/~sdurlauf/includes/pdf/Trends%20versus%20Random%20Walks%20in%20Time%20Series%20Analysis.pdf>).

$$y_t = \delta y_{t-1} + \varepsilon_t$$



# Pruebas de raíz unitaria

Para saber si una serie es o estacionaria, podemos hacer :

1. Verificación gráfica
2. Analizar los correlogramas, y
3. Hacer pruebas de raíz unitaria.

Una raíz unitaria es una característica de los procesos que evolucionan a través del tiempo y que puede causar problemas en inferencia estadística en modelos de series de tiempo. Un proceso estocástico lineal tiene una raíz unitaria si el valor de la raíz de la ecuación característica del proceso es igual a 1 (decimos que un proceso I(1)). Por lo tanto, tal proceso no es estacionario. Si las demás raíces de la ecuación característica se encuentran dentro del círculo unitario - es decir, tienen un valor absoluto menor a uno - entonces la primera diferencia del proceso es estacionaria.

Existen muchas pruebas formales para verificar la existencia de raíces unitarias en una serie temporal. Las que veremos son: 1. Dickey-Fuller (DF) 2. Dickey-Fuller aumentada (ADF) y la Phillip-Perrón (PP).

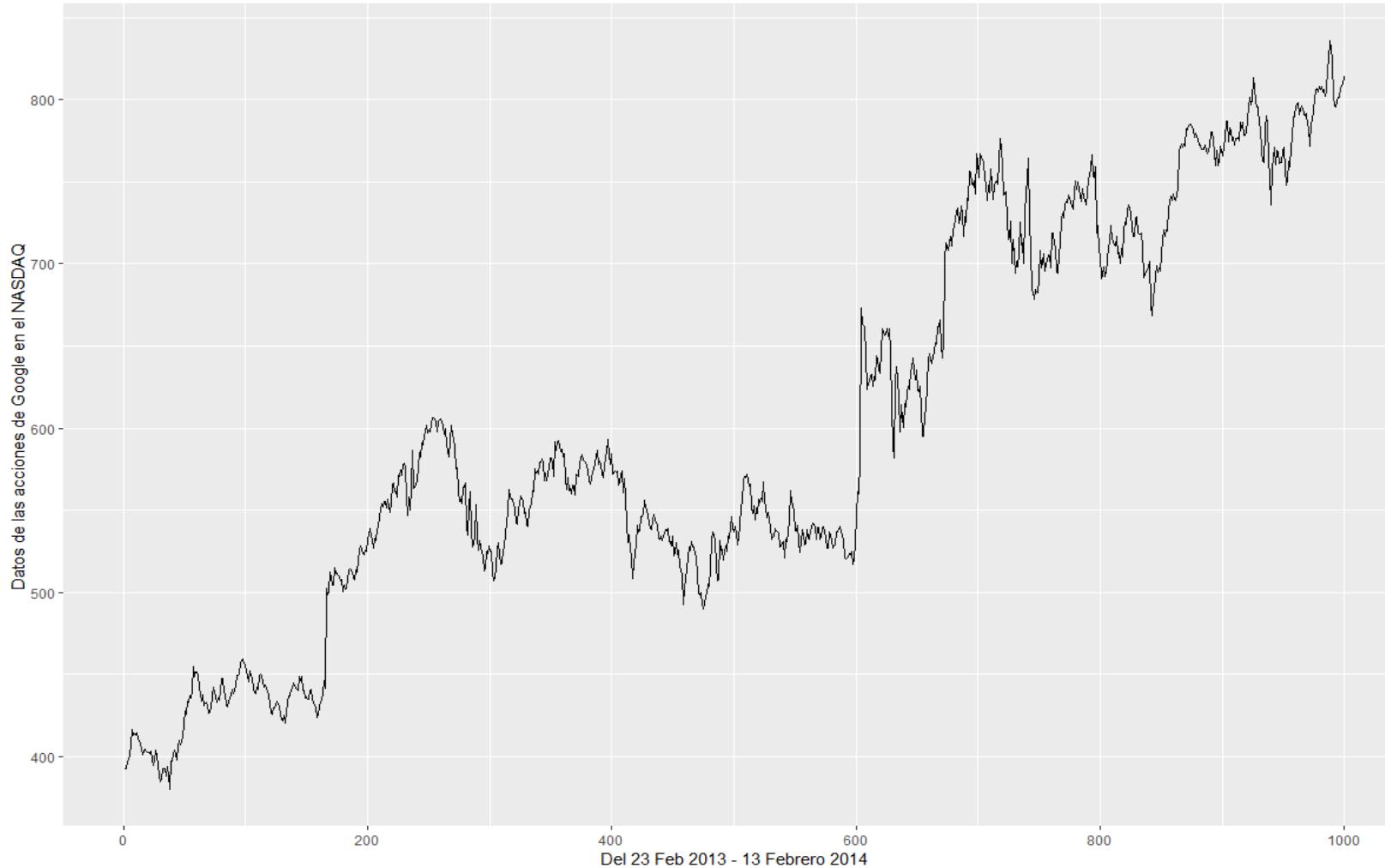
El presente curso no profundiza sobre la teoría de las raíces unitarias, ni sus distintas modalidades, ni sus estadísticos, valores críticos, función de probabilidades con momentos browsianos, ni las tabla de Davidson y Mackinnon. Nos conformaremos con saber en qué punto podremos o no hacer diferenciación.

Para más referencias, ver los siguientes documentos:

<https://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf>

# Pruebas de raíz unitaria

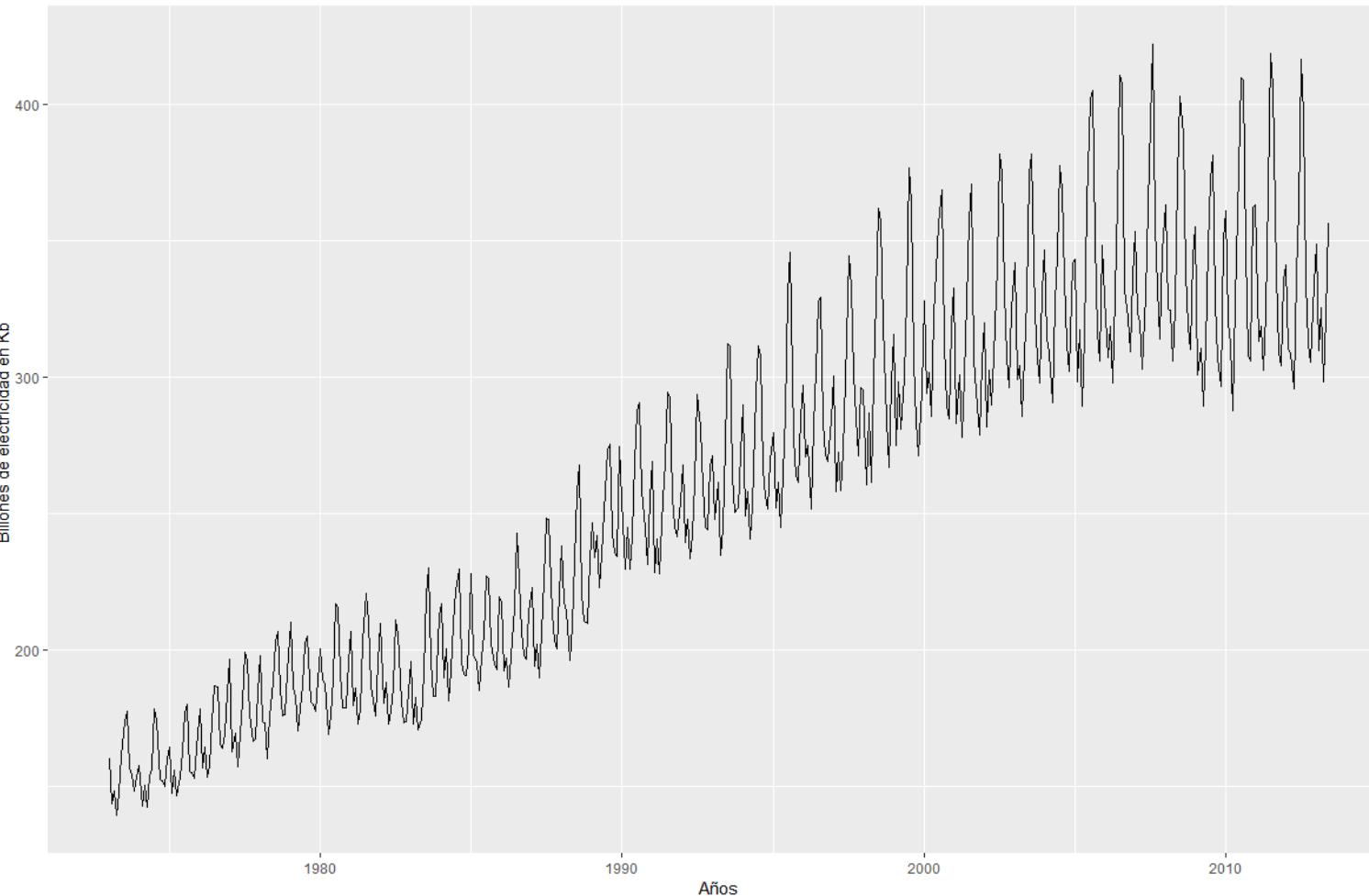
Para el ejemplo de las acciones de google, queremos saber si hay una raíz unitaria dada la presencia de tendencia. Para esto utilizamos la función ndiffs(). Nótese que el resultado fue de “1”, indicando la presencia de una raíz unitaria



```
> ndiffs(goog)
[1] 1
>
```

# Pruebas de raíz unitaria

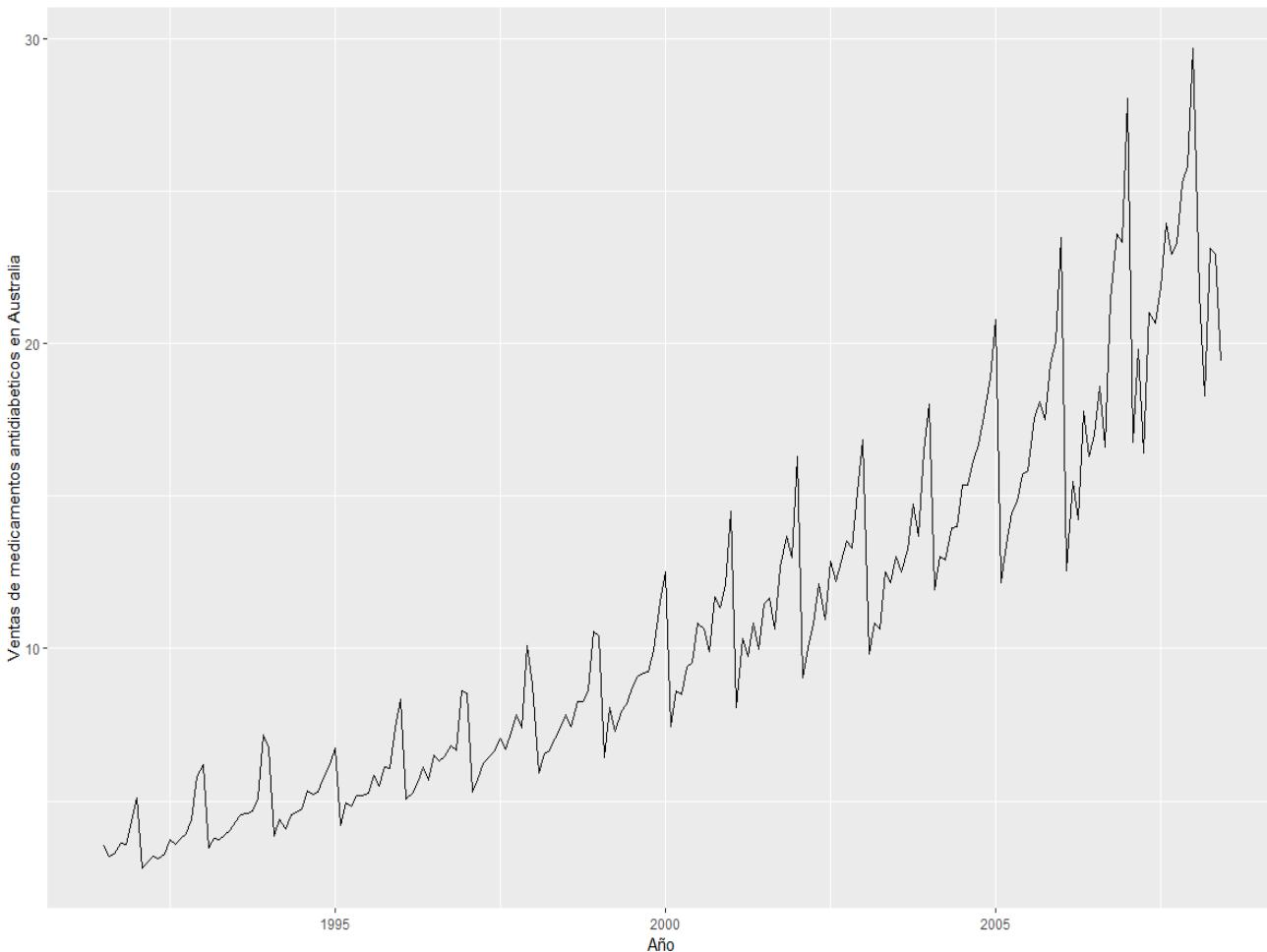
Para el ejemplo de demanda eléctrica, queremos saber si hay una raíz unitaria tanto a nivel de la tendencia como de la estacionalidad. Para esto utilizamos la función *ndiffs()* y *diff(lag=)*. Nótese que el resultado fue de “1”, en ambos casos, indicando la presencia de una raíz unitaria.



```
> usmelec %>% log() %>% nsdiffs()
[1] 1
>
> usmelec %>% log() %>% diff(lag=12) %>% ndiffs()
[1] 1
```

# Pruebas de raíz unitaria

Para el ejemplo de la venta de medicamentos, queremos saber si hay una raíz unitaria tanto a nivel de la tendencia como de la estacionalidad. Para esto utilizamos la función *ndiffs()* y *diff(lag=)*. Nótese que el resultado fue de “1”, aunque tuve que hacer diversas combinaciones...



```
#Datos sobre ventas de medicamentos para diabéticos
a<- a10
b<- log(a10)
c<- diff(log(a10),1)
d<-diff(log(a10),1,12)

a10 %>% log() %>% nsdiffs()
```

```
a %>% log() %>% nsdiffs()
b %>% log() %>% nsdiffs()
c %>% log() %>% nsdiffs()
d %>% log() %>% nsdiffs()
```



```
> a %>% log() %>% nsdiffs()
[1] 1
> b %>% log() %>% nsdiffs()
[1] 1
> c %>% log() %>% nsdiffs()
[1] 0
Warning message:
In log(.) : NaNs produced
> d %>% log() %>% nsdiffs()
[1] 1
Warning message:
In log(.) : NaNs produced
```

# Índice

1

Fundamentos del método  
Box-Jenkins

2

La estacionaridad y las  
pruebas de raíz unitaria

3

Identificación y  
estimación de un proceso  
ARIMA

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

ARMA(p,q)

Identificación de  
procesos

Estimación y adición  
de parámetros

Modelos ARIMA  
estacionales

Estimación ARIMA  
estacional

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

# Identificación : los procesos MA(q)

El proceso MA(q) admite dos condiciones:

1. El componente determinístico del proceso  $x_t$  es una constante  $\mu_x$ .
2. La realización en el tiempo t del proceso  $x_t$  es la suma ponderada de  $q$  valores presentes y pasados de un ruido blanco  $u_t$ .

De esta forma se tiene que un proceso  $x_t$  está gobernado por un MA(q) y este se expresa como

$$x_t = X_t - \mu_x = u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q}$$

Con memoria  $E[u_t] = 0$

$$E[u_t u_{t+s}] = \begin{cases} \sigma_u^2 & \text{si } s = 0 \\ 0 & \text{others cases} \end{cases}$$

Está claro que estos procesos son estacionarios y esto sin ningún tipo de restricción que se le haya impuesto al a los parámetros teta cuando son constantes y con número finito de estos. Por otro lado, puede ser necesario hacer referencia a una reescritura de un proceso  $MA(q)$  dado en términos de procesos autorregresivos, para especificar su función de autocorrelación parcial. Con el fin de garantizar la singularidad de esta reescritura, se debe asegurarse de que uno está en presencia de un proceso invertible MA. Antes de especificar estas condiciones de inversión, se debe introducir el operador  $B$  de desplazamiento, lo que simplificará en gran medida las escrituras.

# Identificación : los procesos MA(q)

## El operador de desplazamiento $B$

Este se define como:  $B^k x_t = x_{t-k}$ .

**Recordar:** Si se aplica a una constante tenemos:  $B^k c = c$

Para un proceso  $MA(q)$ , se tiene por ecuación:

$$x_t = u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q}$$

Lo cual se puede reescribir como

$$x_t = (1 - \vartheta_1 B - \vartheta_2 B^2 - \cdots - \vartheta_q B^q) u_t$$

Donde  $\vartheta(B)$  es un polinomio en  $B$  de grado  $q$ .

Veamos ahora cómo sería la presentación para un  $MA(1)$ .

# Identificación : los procesos MA(q)

## El proceso MA(1)

Su escritura está dada por la siguiente ecuación:  $x_t = u_t - \vartheta_1 u_{t-1} = (1 - \vartheta_1 B)u_t$

El cálculo de los dos primeros momentos sería:

$$E[x_t] = 0 \text{ o también } E[X_t] = \mu_X$$

$$Var[x_t] = \gamma_0 = (1 + \vartheta_1^2)\sigma_u^2$$

$$\begin{aligned} Cov(x_t, x_{t-1}) &= \gamma_1 = E[x_t, x_{t-1}] = E[(u_t - \vartheta_1 u_{t-1})(u_{t-1} - \vartheta_1 u_{t-2})] = -\vartheta_1 \sigma_u^2 \\ Cov(x_t, x_{t-k}) &= \gamma_k = E[x_t, x_{t-k}] = E[(u_t - \vartheta_1 u_{t-1})(u_{t-k} - \vartheta_1 u_{t-k-1})] = 0 \quad k > 1 \end{aligned}$$

## Función de autocorrelación de un MA(1)

Está dada por las siguientes ecuaciones:

$$Cor(x_t, x_{t-1}) = \rho_1 \equiv \frac{\gamma_1}{\gamma_0} = -\frac{\vartheta_1}{(1 + \vartheta_1^2)}$$

$$Cor(x_t, x_{t-k}) = \rho_k \equiv \frac{\gamma_k}{\gamma_0} = 0 \quad k > 1$$

- La única autocorrelación no nula es por lo tanto  $\rho_1$ . La memoria del proceso es de un único período en la autocorrelación total. Veremos así que  $\vartheta_1$  y  $\rho_1$  son de signo opuesto.

# Identificación : los procesos MA(q)

## La función de autocorrelación parcial de un proceso MA(1)

Sabemos que solo  $\rho_0$  y  $\rho_1$  no son nulas (con  $\rho_0 = 1$ ). Y entonces, si elegimos las ecuaciones de Yule-Walker tenemos que:

$$\begin{bmatrix} \rho_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}, k = 1, 2, \dots, K$$

De donde obtenemos que:

$$\begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, k = 1, 2, \dots, K$$

# Identificación : los procesos MA(q)

La solución del sistema para los valores sucesivos de k-expresiones de los coeficientes de las auto correlaciones parciales son  $\phi_{kk}$ :

$$\phi_{11} = \rho_1$$

$$\phi_{11} = -\frac{\vartheta_1}{1 + \vartheta_1^2}$$

$$\phi_{22} = \frac{\rho_1^2}{\rho_1^2 - 1}$$

$$\phi_{22} = -\frac{\vartheta_1^2}{1 + \vartheta_1^2 + \vartheta_1^4}$$

$$\phi_{33} = \frac{\rho_1^3}{1 - 2\rho_1^2}$$

$$\phi_{33} = -\frac{\vartheta_1^3}{1 + \vartheta_1^2 + \vartheta_1^4 + \vartheta_1^6}$$

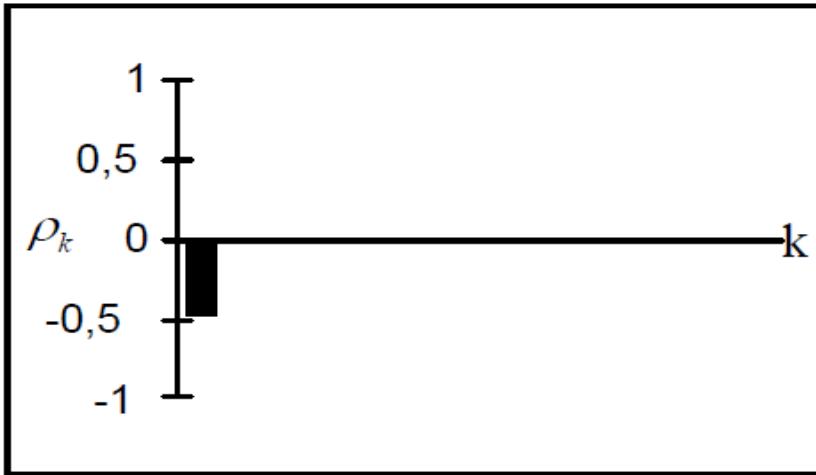
Y de forma general:

$$\phi_{kk} = -\frac{\vartheta_1^k}{\sum_{i=0}^k \vartheta_1^{2i}}$$

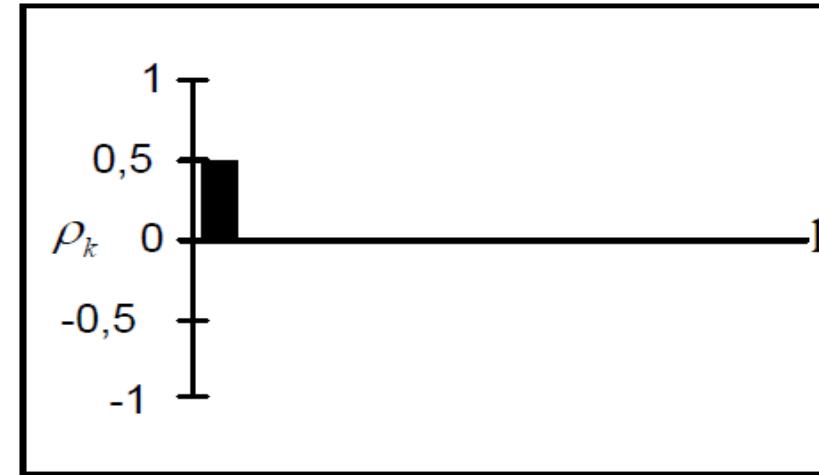
Los coeficientes de autocorrelación parcial irán en decrecimiento en función de  $k$ , y serán todos negativos si  $\vartheta_1 > 0$ , o se alternarán si  $\vartheta_1 < 0$ . Los dos ejemplos siguientes ilustran los rasgos característicos de las funciones de auto correlación total y parcial para un MA(1).

# Identificación : los procesos MA(q)

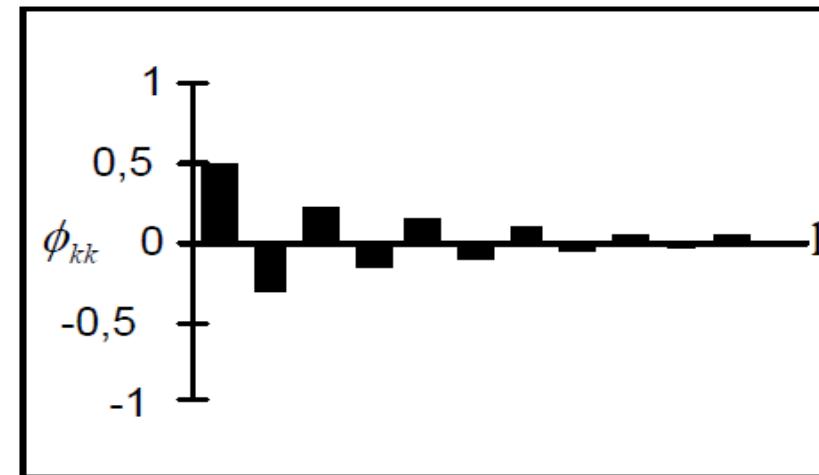
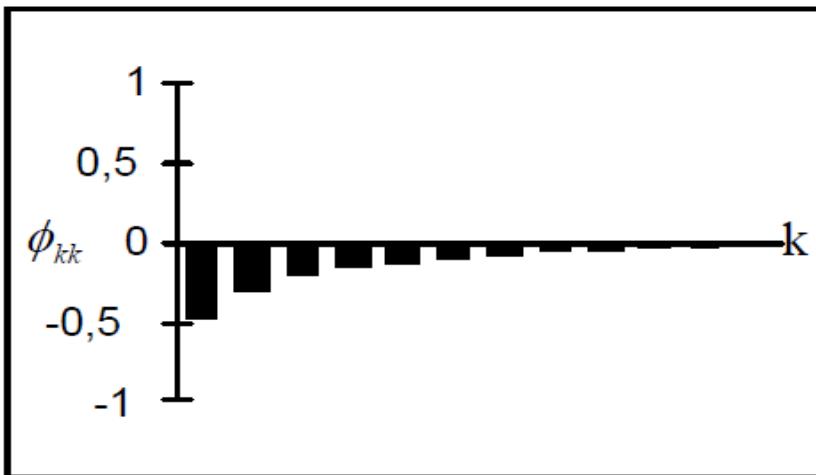
¿ Cuál es  $\vartheta_1 > 0$ , y cuál  $\vartheta_1 < 0$ ?



$$x_t = u_t - 0,8u_{t-1}$$



$$x_t = u_t + 0,8u_{t-1}$$



# Identificación : los procesos MA(q)

## El proceso MA(1)

Su escritura está dada por la siguiente ecuación:  $x_t = u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} = (1 - \vartheta_1 B - \vartheta_2 B^2)u_t$

El cálculo de los dos primeros momentos sería:

$$E[x_t] = 0 \text{ o también } E[X_t] = \mu_X$$

$$Var[x_t] = \gamma_0 = (1 + \vartheta_1^2 + \vartheta_2^2)\sigma_u^2$$

$$\begin{aligned} Cov(x_t, x_{t-1}) &= \gamma_1 = E[x_t x_{t-1}] = E[(u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2})(u_{t-1} - \vartheta_1 u_{t-2} - \vartheta_2 u_{t-3})] = -\vartheta_1(1 - \vartheta_2)\sigma_u^2 \\ Cov(x_t, x_{t-2}) &= \gamma_2 = E[x_t x_{t-2}] = E[(u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2})(u_{t-2} - \vartheta_1 u_{t-3} - \vartheta_2 u_{t-4})] = -\vartheta_1\sigma_u^2 \\ Cov(x_t, x_{t-k}) &= \gamma_k = E[x_t x_{t-k}] = 0 \quad k > 2 \end{aligned}$$

## Función de autocorrelación de un MA(2)

Está dada por las siguientes ecuaciones:

$$Cor(x_t, x_{t-1}) = \rho_1 \equiv \frac{\gamma_1}{\gamma_0} = -\frac{\vartheta_1(1 - \vartheta_2)}{(1 + \vartheta_1^2 + \vartheta_2^2)}$$

$$Cor(x_t, x_{t-2}) = \rho_2 \equiv \frac{\gamma_2}{\gamma_0} = -\frac{\vartheta_2}{(1 + \vartheta_1^2 + \vartheta_2^2)}$$

$$Cor(x_t, x_{t-k}) = \rho_k \equiv \frac{\gamma_k}{\gamma_0} = 0 \quad k > 2$$

La única autocorrelación no nula son para  $\rho_0, \rho_1, \rho_2$ . La memoria del proceso es de dos períodos.

# Identificación : los procesos MA(q)

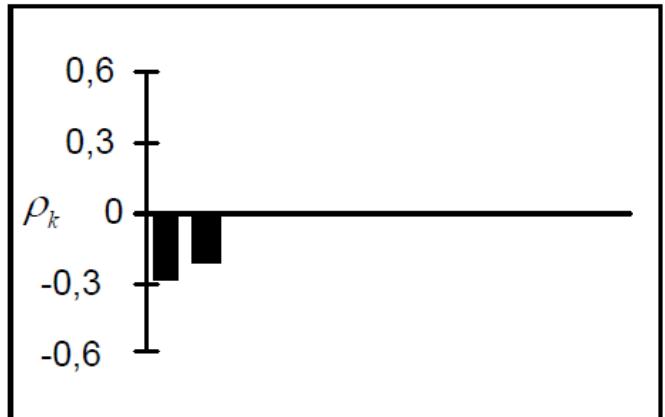
- Si consideramos las ecuaciones de Yule-Walker, los coeficientes de la función de autocorrelación parcial se obtienen resolviendo el siguiente sistema:

$$\begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \phi_{k4} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \cdots & 0 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \cdots & 0 \\ 0 & \rho_2 & \rho_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad k = 1, 2, \dots, K$$

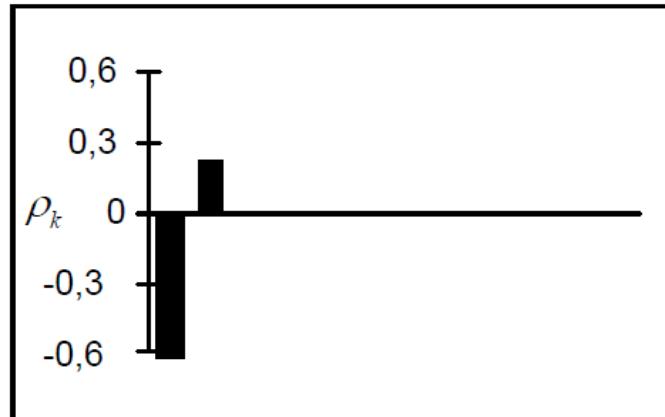
Las expresiones son complejas. Nos damos cuenta que como para el  $MA(1)$ , la función de autocorrelación parcial no se anulan más allá de un orden  $k$ .

# Identificación : los procesos MA(q)

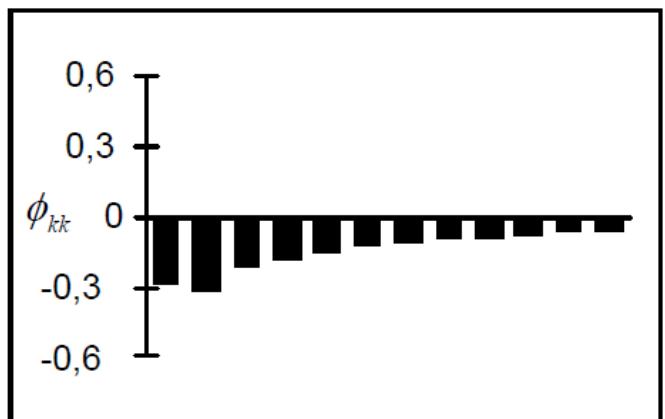
Los correogramas de abajo presentan diferentes procesos para un MA(2)



$$x_t = u_t - .6u_{t-1} - .3u_{t-2}$$

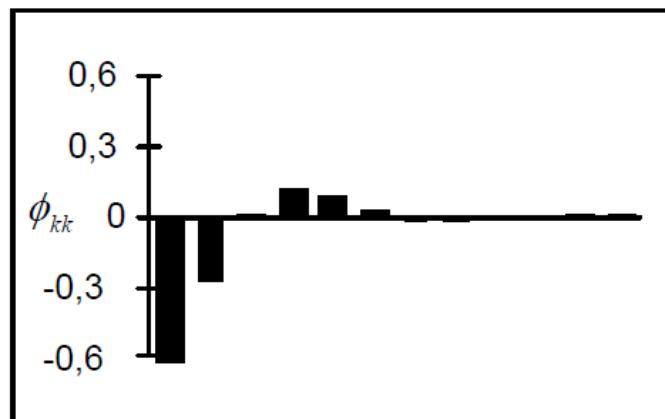


$$x_t = u_t - .8u_{t-1} + .4u_{t-2}$$



$$w_1 = 1.08$$

$$w_2 = -3.08$$



$$w_1 = 1 - 1.22i$$

$$w_2 = 1 + 1.22i$$

# Identificación : los procesos MA(q)

## El proceso $MA(q)$

Partiendo de la expresión

$$x_t = u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q}$$

¿Qué tendríamos como valores de primer y segundo orden?

La función de autocovariancia estaría dado por lo siguiente:

$$\gamma_k = \begin{cases} (-\vartheta_k + \vartheta_1 \vartheta_{k+1} + \vartheta_2 \vartheta_{k+2} + \cdots + \vartheta_q \vartheta_{k+q}) \sigma_u^2, & k = 0, 1, 2, \dots, q \\ 0, & k > q \end{cases}$$

La función de autocorrelación de un  $MA(q)$  se deduce de forma inmediata de la función de autocovariancia, y posee una memoria de  $q$  periodos:

$$\rho_k = \begin{cases} \frac{(-\vartheta_k + \vartheta_1 \vartheta_{k+1} + \vartheta_2 \vartheta_{k+2} + \cdots + \vartheta_q \vartheta_{k+q})}{1 + \vartheta_1^2 + \vartheta_2^2 + \cdots + \vartheta_q^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases}$$

La función de autocorrelación parcial de un  $MA(q)$ , para las expresiones de los elementos  $\phi_{kk}$  pueden obtener a partir de las ecuaciones de Yule-Walker, pero son bastante complejas.

# Identificación : los procesos MA( $q$ )

- Hay 3 cosas que son importantes notar:
  1. Las autocorrelaciones parciales son infinitas. La razón es dado que el proceso  $MA(q)$  bajo la hipótesis de inversibilidad así lo genera.
  2. La serie de coeficientes  $\phi_{kk}$  converge a “0”. Esto es una consecuencia de la hipótesis de inversibilidad.
  3. La convergencia es de tipo senoidal.
  4. ¿Algo más?

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

# Identificación : los procesos AR(p)

El proceso  $AR(p)$  admite dos condiciones:

1. El componente determinístico del proceso  $x_t$  es una constante  $\mu_x$ .
2. La realización en el tiempo t del proceso  $x_t$  es la suma ponderada de  $p$  valores presentes y pasados de un ruido blanco  $u_t$ .

De esta forma se tiene que un proceso  $x_t$  está gobernado por un  $AR(p)$  y este se expresa como

$$x_t = X_t - \mu_x = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + u_t$$

Con memoria  $E[u_t] = 0$

$$E[u_t u_{t+s}] = \begin{cases} \sigma_u^2 & \text{si } s = 0 \\ 0 & \text{otros casos} \end{cases}$$

También a la expresión anterior:

$$x_t = X_t - \mu_x = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + u_t$$

Se puede re escribir como sigue:

$$\phi(B)x_t = u_t,$$

$$\text{con } \phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$$

# Identificación : los procesos AR(p)

- Estos procesos autoregresivos son por naturaleza invertibles, sin embargo no son necesariamente estacionarios: esto requiere volver a escribir una forma infinita de orden  $MA$  que impone una restricción en las raíces del polinomio característico  $\emptyset(B)$ .
- Esta referencia a una reescritura en la forma de un orden infinito de  $MA$  permite caracterizar la memoria del proceso  $AR$  medida en la función de autocorrelación: este último será infinito en el sentido de que las autocorrelaciones entre  $x_t$  y  $x_{t-k}$  no se cancelen para cualquiera que sea el valor de  $k$ . De hecho, los términos  $u_{t-k}, u_{t-k-1}, u_{t-k-2}, \dots$ , están presentes simultáneamente en cada una de las realizaciones  $x_t$  de los tiempos  $t$  y  $t - k$ . Por otra parte, y aquí de nuevo encontramos la simetría con las condiciones de invertibilidad discutidas en la  $MA$ , las condiciones de estacionaridad impondrán la convergencia hacia cero de la serie de autocorrelaciones.
- Veamos que la expresión anterior supone que el proceso es centrado. Claro que se puede trabajar con el proceso original, pero trabajamos con procesos centrados por comodidad:

$$\begin{aligned} x_t &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + u_t \\ &\Updownarrow \\ (X_t - \mu_X) &= \phi_1(X_{t-1} - \mu_X) + \phi_2(X_{t-2} - \mu_X) + \cdots + \phi_p(X_{t-p} - \mu_X) + u_t \\ &\Updownarrow \\ X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + (1 - \phi_1 - \phi_2 - \cdots - \phi_p)\mu_X + u_t \\ &\Updownarrow \\ X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + c + u_t \end{aligned}$$

- De donde se obtiene:  $E[X_t] = \mu_X = \frac{c}{(1 - \phi_1 - \phi_2 - \cdots - \phi_p)}$

- Para simplificar la expresión, mantendremos la expresión de las realizaciones centradas.

# Identificación : los procesos AR(p)

## El proceso AR(1)

Su escritura está dada por la siguiente ecuación:  $x_t = \phi_1 x_{t-1} + u_t$ , o también:  $(1 - \phi_1 B)x_t = u_t$

Mediante la condición de estacionaridad se puede obtener el <sup>pasaje a una escritura MA:</sup>

$$x_t = (1 - \phi_1 B)^{-1}u_t, \quad \text{o también de la forma: } x_t = \sum_{i=0}^{\infty} \phi_1^i u_{t-i}$$

Con condición de que  $|\phi_1| < 1$

Para los momentos matemáticos, recordamos que trabajamos sobre una serie centrada. La función de autocovariancia se calcula como sigue:

$$x_t = \phi_1 x_{t-1} + u_t \Rightarrow x_t x_{t-k} = \phi_1 x_{t-1} x_{t-k} + u_t x_{t-k}$$

Recordamos que:  $E[u_t x_{t-k}] = 0$  si  $k > 0$

$$E[u_t x_t] = E[u_t (\phi_1 x_{t-1} + u_t)] = \sigma_u^2$$

Y especificamos que:  $\gamma_0 = \phi_1 \gamma_1 + \sigma_u^2$

$$\gamma_k = E[x_t x_{t-k}] = \phi_1 E[x_{t-1} x_{t-k}] = \phi_1 \gamma_{k-1} \quad \text{si } k > 0$$

# Identificación : los procesos AR(p)

## Función de autocorrelación total de un método AR(1)

Si partimos de las expresiones anteriores, y lo dividimos por  $\gamma_0$ , obtenemos por función de autocorrelación:

$$\rho_k = \phi_1 \rho_{k-1} \quad \text{si } k > 0 \quad \text{o su similar} \quad \rho_k = \phi_1^k$$

Por lo que el valor de  $\gamma_0$  se representa como:

$$1 = \phi_1 \rho_1 + \frac{\sigma_u^2}{\gamma_0} \Rightarrow \gamma_0 = \frac{\sigma_u^2}{1 - \phi_1 \rho_1} = \frac{\sigma_u^2}{1 - \phi_1^2}$$

Esta función se caracteriza por una disminución exponencial de los términos todos positivos si  $\phi_1 > 0$ , alternando en signo si  $\phi_1 < 0$ . Por último, cabe señalar que la función de autocorrelación y la función de autocorrelación obedecen, al ruido blanco a la misma ecuación que el proceso.

En la práctica, esto significa, por ejemplo, que si  $\phi_1$  es positivo, entonces las autocovariancias y autocorrelaciones también serán positivas y una realización  $x_t$  mayor que cero (o  $x_t$  mayor que  $\mu_X$ ) será seguida de realizaciones más positivas que negativas y viceversa. Por otro lado, si  $\phi_1$  es negativo, entonces las autocovariancias y autocorrelaciones alternan en signo y en este caso, la trayectoria observada a menudo debe "perforar" su media.

# Identificación : los procesos AR(p)

La función de autocorrelación parcial del  $AR(1)$

Estos se obtienen fácilmente si nos acordamos que el coeficiente de autocorrelación parcial de orden  $k$  es el coeficiente de  $x_{t-k}$  para  $x_{t-1}, x_{t-2}, \dots, x_{t-k}$ . Consideramos lo siguiente como una regresión lineal:

$$x_t = \phi_{11}x_{t-1} + u_t$$

$$x_t = \phi_{21}x_{t-1} + \phi_{22}x_{t-2} + u_t$$

⋮

$$x_t = \phi_{K1}x_{t-1} + \phi_{K2}x_{t-2} + \dots + \phi_{KK}x_{t-K} + u_t$$

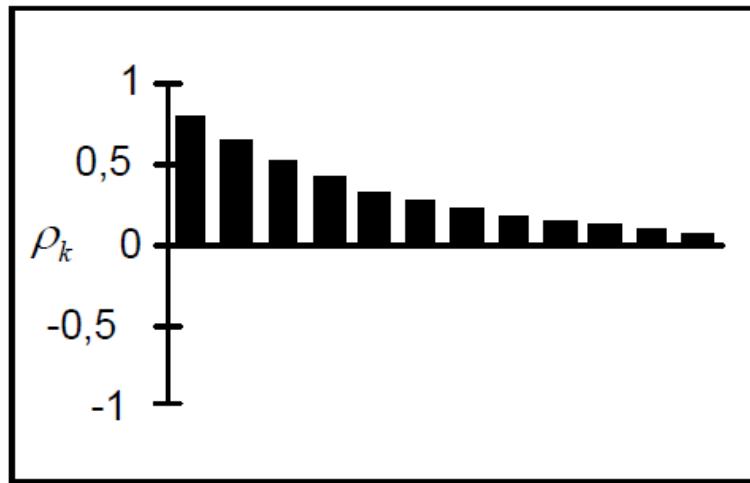
Si el verdadero proceso es un modelo  $AR(1)$ , entonces:

$$x_t = \phi_1 x_{t-1} + u_t,$$

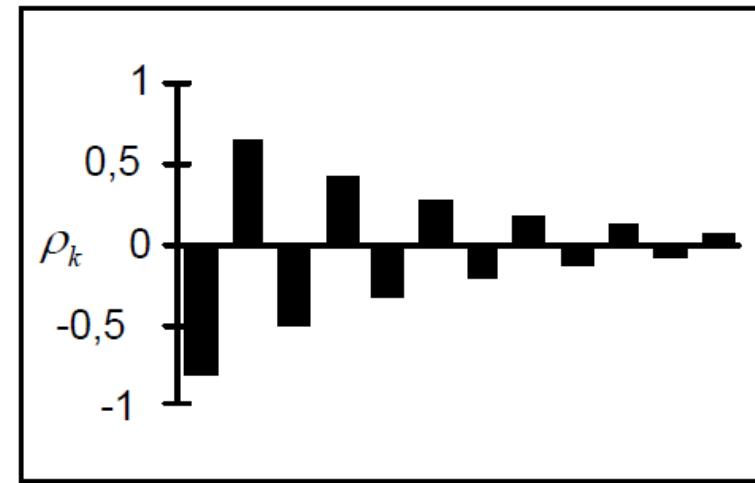
y entonces tenemos que:  $\phi_{11} = \phi_1, \phi_{22} = \phi_{33} = \dots = 0$

# Identificación : los procesos AR(p)

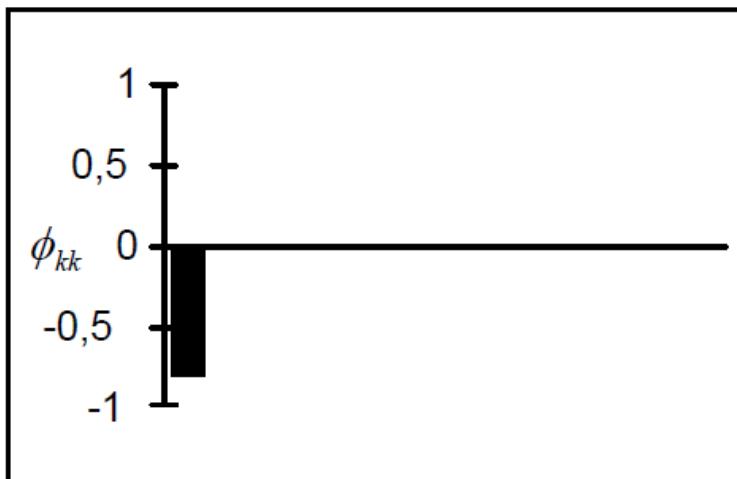
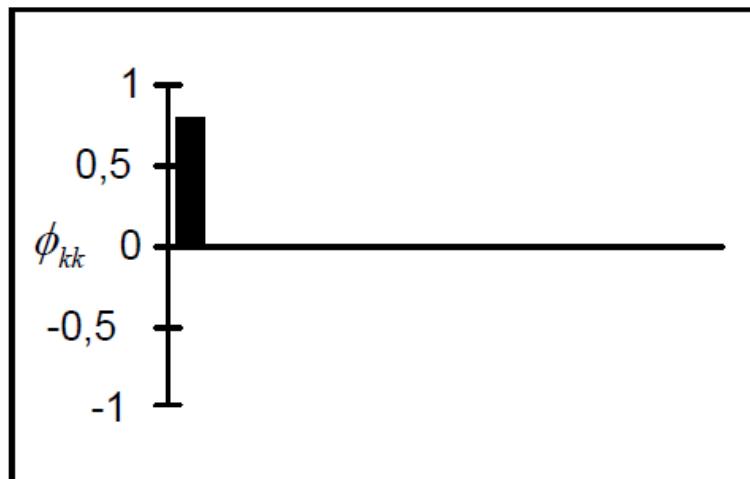
Se representan las funciones de autocorrelación total y parcial para un proceso AR(1).



$$x_t = .8x_{t-1} + u_t$$



$$x_t = -.8x_{t-1} + u_t$$



# Identificación : los procesos AR(p)

## El proceso AR(2)

Su escritura está por :  $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t$ , *o también:*  $(1 - \phi_1 B - \phi_2 B^2)x_t = u_t$

Mediante la condición de estacionaridad se puede obtener el pasaje a una escritura MA:

$$x_t = (1 - \phi_1 B - \phi_2 B^2)^{-1}u_t, \quad \text{Como condición de restricciones se tiene que } \begin{cases} \phi_1 + \phi_2 < 1 \\ -\phi_1 + \phi_2 < 1 \\ -1 < \phi_2 < 1 \end{cases}$$

La función de auto-covariancia se calcula fácilmente. De hecho

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t \Rightarrow x_t x_{t-k} = \phi_1 x_{t-1} x_{t-k} + \phi_2 x_{t-2} x_{t-k} + u_t x_{t-k}$$

Y entonces:

$$E[u_t x_{t-k}] = 0 \text{ si } k > 0 \text{ et } E[u_t x_t] = E[u_t (\phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t)] = \sigma_u^2,$$

Tenemos finalmente:  $\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_u^2$

$$\gamma_k = E[x_t x_{t-k}] = \phi_1 E[x_{t-1} x_{t-k}] + \phi_2 E[x_{t-2} x_{t-k}] = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \text{ si } k > 0$$

# Identificación : los procesos AR(p)

## La función de auto correlación de un $AR(2)$

Dividiendo las autocovariancias por  $\gamma_0$ , obtener como funciones:  $1 = \phi_1\rho_1 + \phi_2\rho_2 + \frac{\sigma_u^2}{\gamma_0} \Rightarrow \gamma_0 = \frac{\sigma_u^2}{1 - \phi_1\rho_1 - \phi_2\rho_2}$

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}, k > 0$$

Como para el caso de AR(1), la función de autocovariacia y la función de autocorrelación obedecen, según el ruido blanco, a la misma ecuación del proceso. En estas condiciones, la hipótesis de estacionaridad implica la convergencia hacia cero de la serie de correlaciones, esta convergencia son exponenciales si las raíces del polinomio  $(1 - \phi_1Z - \phi_2Z^2)$  son reales, y de tipo sinoidales si las raíces son complejas.

Finalmente, vemos que la ecuación anterior permite calcular todos los valores de  $\rho_k$ , si  $k \geq 3$ , en función de  $\rho_1$  y  $\rho_2$ , y entonces los parámetros autoregresivos, sabiendo que:

$$\text{Si } k = 1: \quad \rho_1 = \phi_1 + \phi_2\rho_1 \Rightarrow \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

$$\text{Si } k = 2: \quad \rho_2 = \phi_1\rho_1 + \phi_2 \Rightarrow \rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$$

# Identificación : los procesos AR(p)

## La función de auto correlación parcial de un AR(2)

Estos se obtienen fácilmente si nos acordamos que el coeficiente de autocorrelación parcial de orden  $k$  es el coeficiente de  $x_{t-k}$  para  $x_{t-1}, x_{t-2}, \dots, x_{t-k}$ . Consideramos lo siguiente como una regresión lineal:

$$\begin{aligned}x_t &= \phi_{11}x_{t-1} + u_t \\x_t &= \phi_{21}x_{t-1} + \phi_{22}x_{t-2} + u_t \\&\vdots \\x_t &= \phi_{K1}x_{t-1} + \phi_{K2}x_{t-2} + \cdots + \phi_{KK}x_{t-K} + u_t\end{aligned}$$

Si el verdadero proceso es un modelo AR(2), entonces  $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + u_t$ ,  
y entonces tenemos que:  $\phi_{11} = \phi_1, \phi_{22} = \phi_2, \dots = 0$

La expresión de los coeficientes parciales se pueden obtener a partir de la ecuación de Yule-Walker:

$$k=1 : \phi_{11} = \rho_1$$

$$k=2 :$$

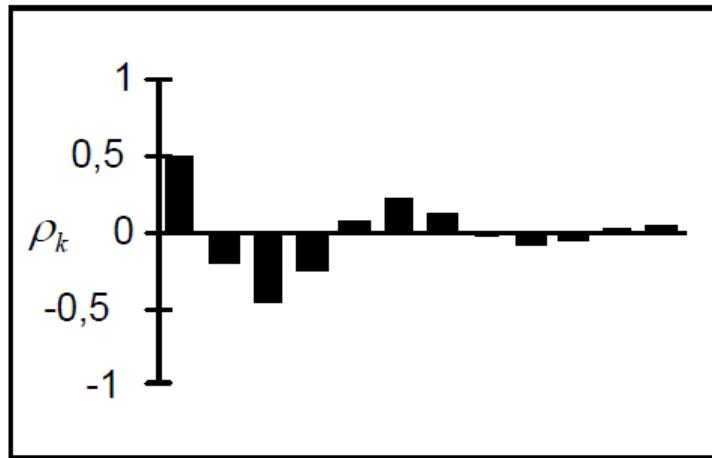
$$\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{pmatrix} \phi_{21} \\ \phi_{22} \end{pmatrix} \Rightarrow \begin{pmatrix} \phi_{21} \\ \phi_{22} \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}$$

$$\Rightarrow \phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \phi_2$$

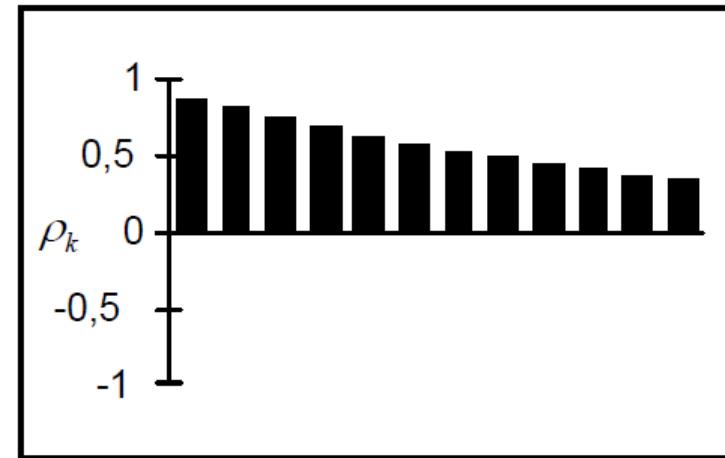
$$k \geq 3 : \phi_{kk} = 0$$

# Identificación : los procesos AR(p)

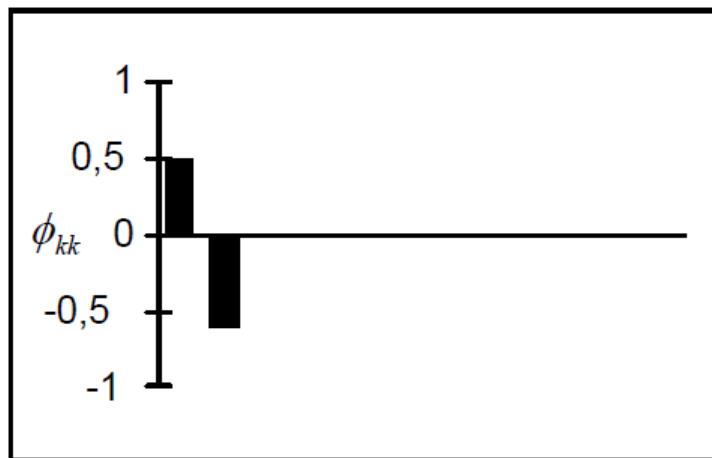
Se representan las funciones de autocorrelación total y parcial para un proceso AR(2).



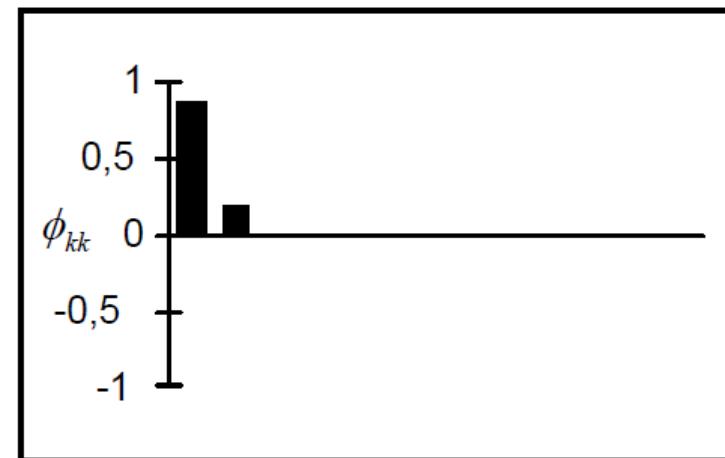
$$x_t = .8x_{t-1} - .6x_{t-2} + u_t$$



$$x_t = .7x_{t-1} + .2x_{t-2} + u_t$$



$$w_1 = .67 - 1.11i, w_2 = .67 + 1.11i$$



$$w_1 = 1.09, w_2 = -4.59$$

# Identificación : los procesos AR(p)

## El proceso $AR(p)$

Su escritura es :

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + u_t, \quad o \quad (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) x_t = u_t$$

Mediante la condición de estacionaridad se puede obtener el pasaje a una escritura MA:

$$x_t = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)^{-1} u_t,$$

Tenemos que la obtención de una autocorrelación total se rige por la ecuación:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$$

# Identificación : los procesos AR(p)

## La función de auto correlación parcial de un $AR(p)$

Estos se obtienen fácilmente si nos acordamos que el coeficiente de autocorrelación parcial de orden  $k$  es el coeficiente de  $x_{t-k}$  para  $x_{t-1}, x_{t-2}, \dots, x_{t-k}$ . Consideramos lo siguiente como una regresión lineal:

$$\begin{aligned}x_t &= \phi_{11}x_{t-1} + u_t \\x_t &= \phi_{21}x_{t-1} + \phi_{22}x_{t-2} + u_t \\&\vdots \\x_t &= \phi_{p-1,1}x_{t-1} + \phi_{p-1,2}x_{t-2} + \cdots + \phi_{p-1,p-1}x_{t-p+1} + u_t \\x_t &= \phi_{p1}x_{t-1} + \phi_{p2}x_{t-2} + \cdots + \phi_{p,p}x_{t-p} + u_t \\x_t &= \phi_{p+1,1}x_{t-1} + \phi_{p+1,2}x_{t-2} + \cdots + \phi_{p+1,p}x_{t-p} + \phi_{p+1,p+1}x_{t-p+1} + u_t \\&\vdots \\x_t &= \phi_{K1}x_{t-1} + \phi_{K2}x_{t-2} + \cdots + \phi_{KK}x_{t-K} + u_t, K>p+1\end{aligned}$$

Cuando el orden  $p$  se supera, sucede que:  $\phi_{K,p+1} = \phi_{K,p+2} = \cdots = \phi_{K,K} = 0$

Cuando  $k < p$ , uno está en el caso de la omisión clásica de variables relevantes con correlación no nula entre variables y variables omitidas. Los coeficientes estarán sesgados, pero, lo que es más importante, generalmente no serán cero. En otras palabras, La función no es cero hasta llegar después del cero. También es posible expresar los diferentes coeficientes  $\phi_{kk}$  en función de las autocorrelaciones  $\rho_1, \rho_2, \rho_p$ , resolviendo el sistema de ecuaciones de Yule-Walker para valores sucesivos de  $k$ .

# Identificación : los procesos AR(p)

De lo anterior tenemos que:

$$\text{pour } k=1 : \quad \phi_{11} = \rho_1$$

pour  $k=2 :$

$$\begin{pmatrix} \phi_{21} \\ \phi_{22} \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}$$

pour  $k=3 :$

$$\begin{pmatrix} \phi_{31} \\ \phi_{32} \\ \phi_{33} \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix}$$

pour  $k=p :$

$$\begin{pmatrix} \phi_{p1} \\ \phi_{p2} \\ \vdots \\ \phi_{pp} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \ddots & \rho_{p-2} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix}$$

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

ARMA(p,q)

# Identificación : los procesos ARMA(p,q)

El proceso  $ARMA(p, q)$  admite dos condiciones:

1. El componente determinístico del proceso  $x_t$  es una constante  $\mu_x$ .
2. La realización en el tiempo  $t$  del proceso  $x_t$  es la suma ponderada de  $p$  y  $q$  valores presentes y pasados de un ruido blanco  $u_t$ .

De esta forma se tiene que un proceso  $x_t$  está gobernado por un  $ARMA(p, q)$  y este se expresa como

$$x_t = X_t - \mu_x = \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \cdots + v_t \quad y \quad v_t = u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q}$$

O también de la forma:

$$\emptyset(B)x_t = \vartheta(B)u_t, \text{ con} \quad \emptyset(B) = (1 - \emptyset_1 B - \emptyset_1 B^2 - \cdots - \emptyset_p B^p) \quad y \\ \vartheta(B) = (1 - \vartheta_1 B - \vartheta_2 B^2 - \cdots - \vartheta_q B^q)$$

Con memoria  $E[u_t] = 0$

$$E[u_t u_{t+s}] = \begin{cases} \sigma_u^2 & \text{si } s = 0 \\ 0 & \text{otros casos} \end{cases}$$

# Identificación : los procesos ARMA(p,q)

También podemos trabajar con datos no centrados. En este caso, un cálculo simple muestra que es necesario introducir una constante  $C$  en la ecuación del filtro de manera que:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + c + u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q}$$
$$E[X] = \mu_X = \frac{c}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

Acá deberemos de ser atentos a las condiciones de estacionaridad y de invisibilidad. El proceso  $ARMA(p, q)$  debe poder considerar:

1. Sea como una media móvil infinita:  $x_t = \emptyset(B)^{-1}\vartheta(B)u_t$
2. Tener la autorización de ser escrito de forma autoregresivo infinita a partir de:  $u_t = \vartheta(B)^{-1}\emptyset(B)u_t$

Las raíces de los polinomios  $\vartheta(B)$  y  $\emptyset(B)$  deben ser en modulo superior a la unidad.

Debemos precisar que la escritura  $ARMA(p, q)$  es la mínima del proceso considerado. Eso quiere decir que no existe raíces comunes para los dos polinomios  $\vartheta(B)$  y  $\emptyset(B)$ . Admitamos, para ilustrar este punto, que los polinomios constituyentes de un filtro ARMA,  $\Phi(B)$  y  $\Theta(B)$ , tengan una raíz común  $\lambda^{-1}$ . En estas condiciones, se tendrá:

$$\begin{aligned} \Phi(B)x_t &= \Theta(B)u_t \\ \Updownarrow \\ (1 - \lambda B)\emptyset(B)x_t &= (1 - \lambda B)\vartheta(B)u_t \\ \Updownarrow \\ \emptyset(B)x_t &= \vartheta(B)u_t \end{aligned}$$

De esta forma, todo proceso  $ARMA(p, q)$  posee una infinidad de representaciones  $ARMA(p+m, q+m)$  equivalentes si multiplicáramos la representación mínima por un mismo polinomio de grado  $m$  a la derecha y la izquierda.

# Identificación : los procesos ARMA(p,q)

## El ARMA(1,1)

Su escritura es:

$$\begin{aligned}x_t &= \phi_1 x_{t-1} + u_t - \vartheta_1 u_{t-1}, \quad o \\(1 - \phi_1 B)x_t &= (1 - \vartheta_1 B)u_t\end{aligned}$$

Las condiciones de estacionaridad permiten el pasar a una escritura *MA*:

$$x_t = (1 - \phi_1 B)^{-1}v_t, o : x_t = \sum_{i=0}^{\infty} \phi_1^i v_{t-i} = (1 - \vartheta_1 B) \sum_{i=0}^{\infty} \phi_1^i u_{t-i}$$

lo cual exige que  $|\phi_1| < 1$

La condiciones de inversibilidad permiten el pasar a una escritura *AR*:

$$u_t = (1 - \vartheta_1 B)^{-1}(1 - \phi_1 B)x_t, \quad o \text{ también} :$$

$$u_t = (1 - \phi_1 B) \sum_{i=0}^{\infty} \vartheta_1^i x_{t-i}$$

Lo cual exige que  $|\vartheta_1| < 1$

# Identificación : los procesos ARMA(p,q)

La función de auto-covariancia del proceso se calcula como sigue:

$$\begin{aligned}\gamma_0 &= E[x_t^2] = E[(\phi_1 x_{t-1} + u_t - \vartheta_1 u_{t-1})^2] \\&= E[\phi_1^2 x_{t-1}^2 + u_t^2 + \vartheta_1^2 u_{t-1}^2 + 2\phi_1 x_{t-1} u_t - 2\phi_1 \vartheta_1 x_{t-1} u_{t-1} - 2\vartheta_1 u_t u_{t-1}] \\&= \phi_1^2 E[x_{t-1}^2] + E[u_t^2] + \vartheta_1^2 E[u_{t-1}^2] - 2\phi_1 \vartheta_1 E[x_{t-1} u_{t-1}] \\&= \phi_1^2 \gamma_0 + \sigma_u^2 + \vartheta_1^2 \sigma_u^2 - 2\phi_1 \vartheta_1 E[(\phi_1 x_{t-2} + u_{t-1} - \vartheta_1 u_{t-2}) u_{t-1}] \\&= \phi_1^2 \gamma_0 + \sigma_u^2 + \vartheta_1^2 \sigma_u^2 - 2\phi_1 \vartheta_1 \sigma_u^2 \\&\Rightarrow \gamma_0 = \frac{(1 + \vartheta_1^2 - 2\phi_1 \vartheta_1)}{1 - \phi_1^2} \sigma_u^2\end{aligned}$$

Para un periodo detenemos:

$$\begin{aligned}\gamma_1 &= E[x_t x_{t-1}] = E[(\phi_1 x_{t-1} + u_t - \vartheta_1 u_{t-1}) x_{t-1}] \\&= E[\phi_1 x_{t-1}^2 + u_t x_{t-1} - \vartheta_1 u_{t-1} x_{t-1}] \\&= \phi_1 \gamma_0 - \vartheta_1 E[u_{t-1} (\phi_1 x_{t-2} + u_{t-1} - \vartheta_1 u_{t-2})] \\&= \phi_1 \gamma_0 - \vartheta_1 E[u_{t-1}^2] \\&= \phi_1 \gamma_0 - \vartheta_1 \sigma_u^2\end{aligned}$$

# Identificación : los procesos ARMA(p,q)

Finalmente, para  $k \geq 2$ :

$$\begin{aligned}\gamma_k &= E[x_t x_{t-k}] = E[(\phi_1 x_{t-1} + u_t - \vartheta_1 u_{t-1}) x_{t-k}] \\ &= E[\phi_1 x_{t-1} x_{t-k} + u_t x_{t-k} - \vartheta_1 u_{t-1} x_{t-k}] \\ &= \phi_1 \gamma_{k-1} \\ &= \phi_1^{k-1} \gamma_1\end{aligned}$$

- La función de autocorrelación de un ARMA (1,1), recordando que  $\rho_k = \frac{\gamma_k}{\gamma_0}$ , se establece como:

$$\rho_1 = \phi_1 - \frac{\vartheta_1(1 - \phi_1^2)}{1 + \vartheta_1^2 - 2\phi_1\vartheta_1}$$

- Y se generaliza como:  $\rho_k = \phi_1 \rho_{k-1} = \phi_1^{k-1} \rho_1$

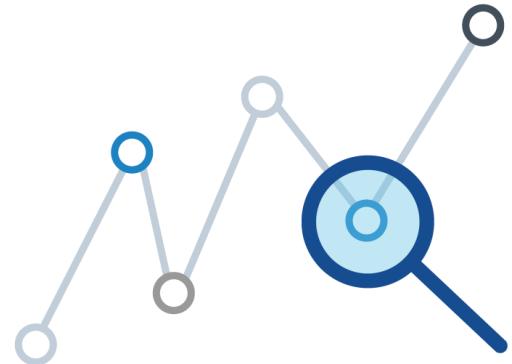
- Se observa un decrecimiento geométrico de los coeficientes de auto-correlación a partir del primero, característica vista ya en el AR(1): el valor de  $\rho_1$  es el mismo que el de  $(\phi_1 - \vartheta_1)$

# Identificación : los procesos ARMA(p,q)

Dado que el proceso tiene una representación equivalente AR de orden infinito, sabemos ahora que la función de autocorrelación parcial también será infinita.

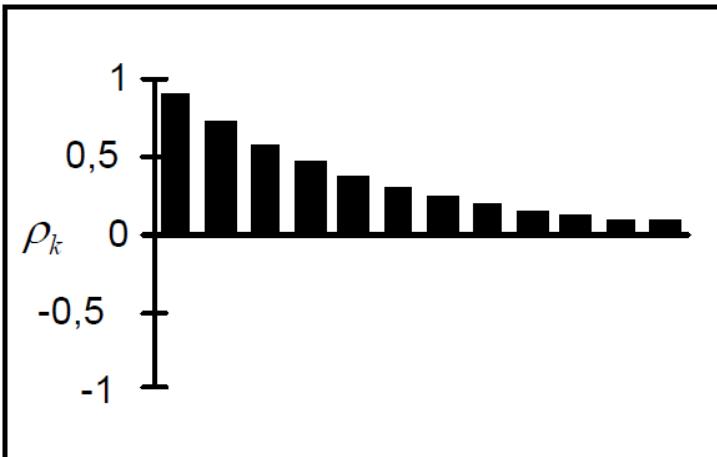
Dada la condición de inversibilidad, la secuencia de coeficientes de la autocorrelación parcial tiende a cero, siguiendo una tendencia cercana a la de un MA(1). Las expresiones están en función de  $\rho_1, \rho_2, \dots$  y entonces, según los resultados anteriores, de  $\varnothing_1$  y de  $\vartheta_1$  se obtiene de la forma habitual mediante la resolución de las  $k$  ecuaciones de Yule-Walker.

En los gráficos siguientes presentamos diversos dendogramas para un ARMA(1,1). Los dos primeros son típicos, un ARMA, el tercero es un ARMA, pero uno podría inclinarse a un AR o MA.



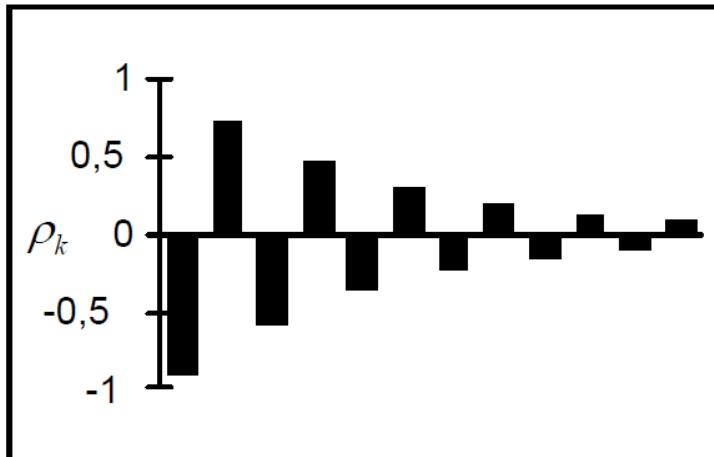
# Identificación : los procesos ARMA(p,q)

Proceso ARMA 1



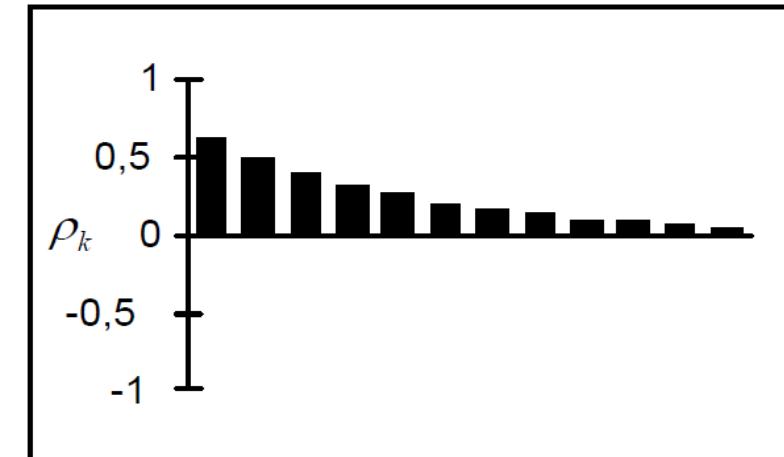
$$x_t = .8x_{t-1} + u_t + .8u_{t-1}$$

Proceso ARMA 2

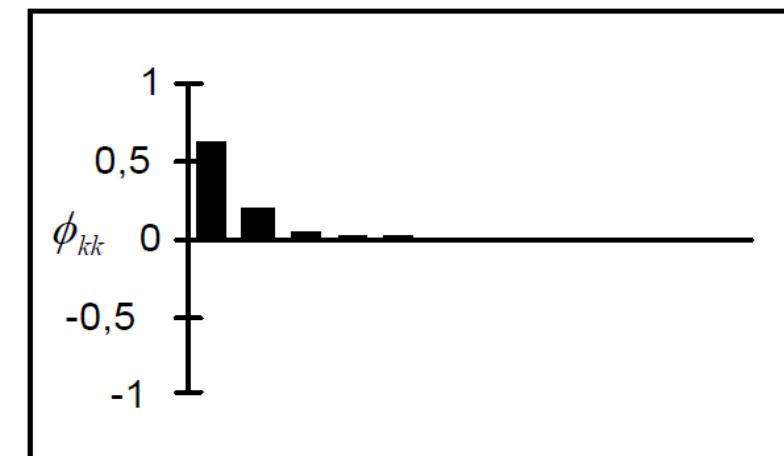
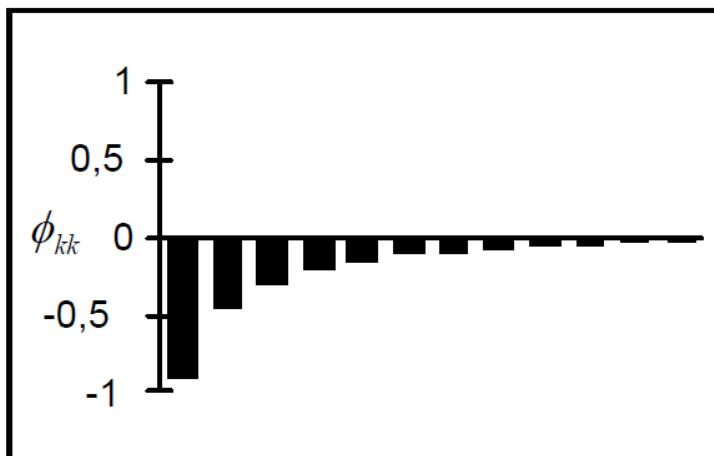
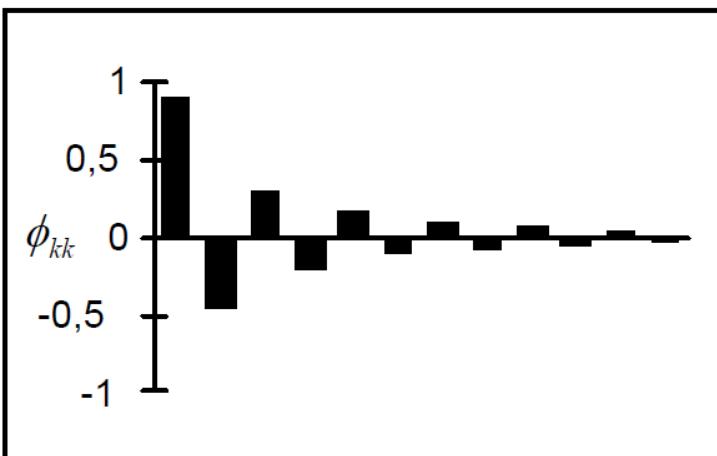


$$x_t = -.8x_{t-1} + u_t - .8u_{t-1}$$

Proceso ARMA 3



$$x_t = .8x_{t-1} + u_t - .3u_{t-1}$$



# Identificación : los procesos ARMA(p,q)

## El ARMA(p,q)

Su escritura está dada por la siguiente ecuación:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + u_t - \vartheta_1 u_{t-1} - \vartheta_2 u_{t-2} - \cdots - \vartheta_q u_{t-q},$$

O también  $\phi(B)x_t = \vartheta(B) u_t$

$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$$

$$\vartheta(B) = (1 - \vartheta_1 B - \vartheta_2 B^2 - \cdots - \vartheta_q B^q)$$

Se imponen las condiciones habituales:

- La estacionaridad: las raíces de  $\phi(B)$  son exteriores al circulo unitario, de manera que  $u_t$  es una innovación.
- La inversibilidad: las raíces de  $\vartheta(B)$  son exteriores al circulo unitario.

La función de auto-covariancia se obtiene de la siguiente forma:

$$\begin{aligned}\gamma_0 &= E[x_t^2] = E[x_t(\phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + u_t - \vartheta_1 u_{t-1} - \cdots - \vartheta_q u_{t-q})] \\ &= \phi_1 \gamma_1 + \cdots + \phi_p \gamma_p + E[x_t u_t] - \vartheta_1 E[x_t u_{t-1}] - \cdots - \vartheta_q E[x_t u_{t-q}] \\ &= \phi_1 \gamma_1 + \cdots + \phi_p \gamma_p + (1 + \vartheta_1^2 + \cdots + \vartheta_q^2) \sigma_u^2\end{aligned}$$

# Identificación : los procesos ARMA(p,q)

$$\begin{aligned}\gamma_1 &= E[x_t x_{t-1}] = E[(\phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + u_t - \vartheta_1 u_{t-1} - \cdots - \vartheta_q u_{t-q}) x_{t-1}] \\ &= \phi_1 \gamma_0 + \cdots + \phi_p \gamma_{p-1} - \vartheta_1 E[u_{t-1} x_{t-1}] - \cdots - \vartheta_q E[u_{t-q} x_{t-1}] \\ &= \phi_1 \gamma_0 + \cdots + \phi_p \gamma_{p-1} - (\vartheta_1 + \vartheta_1 \vartheta_2 + \cdots + \vartheta_{q-1} \vartheta_q) \sigma_u^2\end{aligned}$$

⋮

$$\begin{aligned}\gamma_q &= E[x_t x_{t-q}] = E[(\phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + u_t - \vartheta_1 u_{t-1} - \cdots - \vartheta_q u_{t-q}) x_{t-q}] \\ &= \phi_1 \gamma_{q-1} + \cdots + \phi_p \gamma_{q-p} - \vartheta_q E[u_{t-q} x_{t-q}] \\ &= \phi_1 \gamma_{q-1} + \cdots + \phi_p \gamma_{q-p} - \vartheta_q \sigma_u^2\end{aligned}$$

Dado que,  $k > q$ , las auto-covariancias siguen la siguiente ecuación recurrente:

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p}$$

# Identificación : los procesos ARMA(p,q)

## La función de auto-correlación total de un ARMA(p,q)

Según lo anterior, es fácil ver que los primeros  $q$  coeficientes de la auto-correlación dependerá complejamente de los coeficientes phi y theta. Posteriormente, los elementos de esta función obedecen a la ecuación de recurrencia típica de un proceso  $AR(p)$ :

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \quad k > q$$

Así pues, encontramos resultados ya vistos: una convergencia hacia cero, ligada a la condición de estacionariedad, dominada por exponenciales o sinusoides dependiendo de si las raíces del polinomio característico del componente AR son complejas.

## La función de auto-correlación parcial de un ARMA(p,q)

Debido a la infinita representación autorregresiva cuya existencia está asegurada por las condiciones de inversibilidad, esta función también converge a cero. Ella se aproxima las evoluciones que caracterizan la función de autocorrelación parcial de un proceso MA puro. Sin embargo, sus elementos serán funciones complejas de los dos conjuntos de coeficientes phi y theta.

En la práctica, es fácil ver que la discriminación entre las funciones de autocorrelación parcial de una MA(q) y una ARMA(p,q) no es algo simple como hemos comprobado previamente en el caso del ARMA(1,1).

# Identificación : resumen final de los procesos ARMA(p,q)

En resumen, es posible presentar las principales características de las dos funciones autocorrelación, total y parcial, de los diversos procesos estacionarios hasta ahora considerado. Véase el siguiente cuadro resumen:

Proceso	Auto-correlación total	Auto-correlación parcial
MA(q)	Anulación después de los q primeros coeficientes	Decrecimiento
AR(p)	Decrecimiento	Anulación después de los p primeros coeficientes
ARMA(p,q)	Decrecimiento	Decrecimiento

Se resalta en particular la dualidad perfecta de los procesos MA y AR, dejando por una parte, la posibilidad de elegir entre una u otra representaciones y, por otro lado, seleccionar un orden  $p$  o  $q$  según sea el caso. Como ya se ha señalado, la selección de un proceso ARMA y, además, su  $p$  y  $q$  el examen visual de las dos funciones parece mucho más problemático.

Es necesario continuar con la descripción de este último, que sin embargo sigue siendo útil, cuando se trata de identificar con medios de autocorrelaciones teóricas totales y parciales, pero de sus respectivas estimaciones, para continuar con los pasos de estimar los parámetros de por una parte, y la validación por otro lado.

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

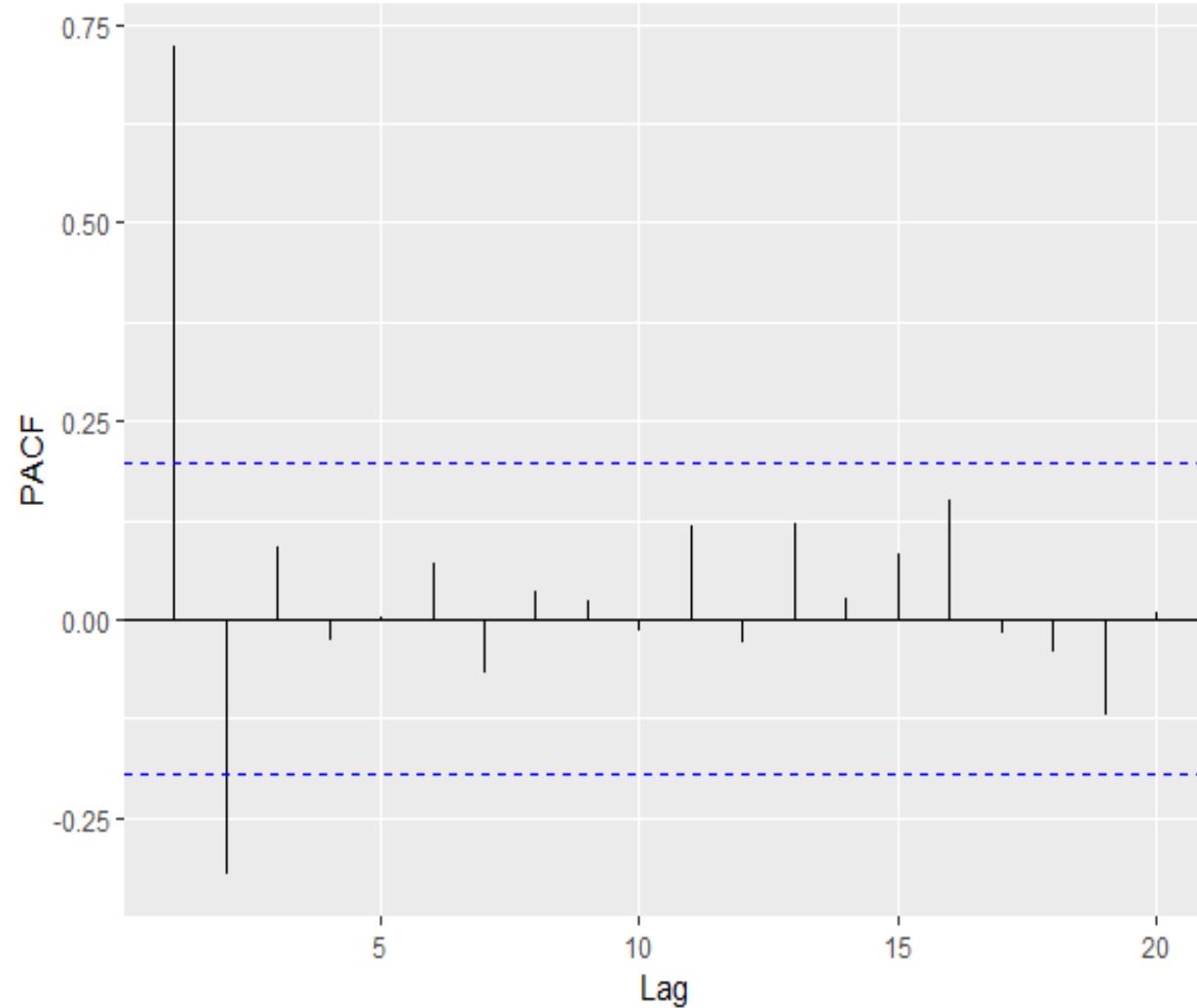
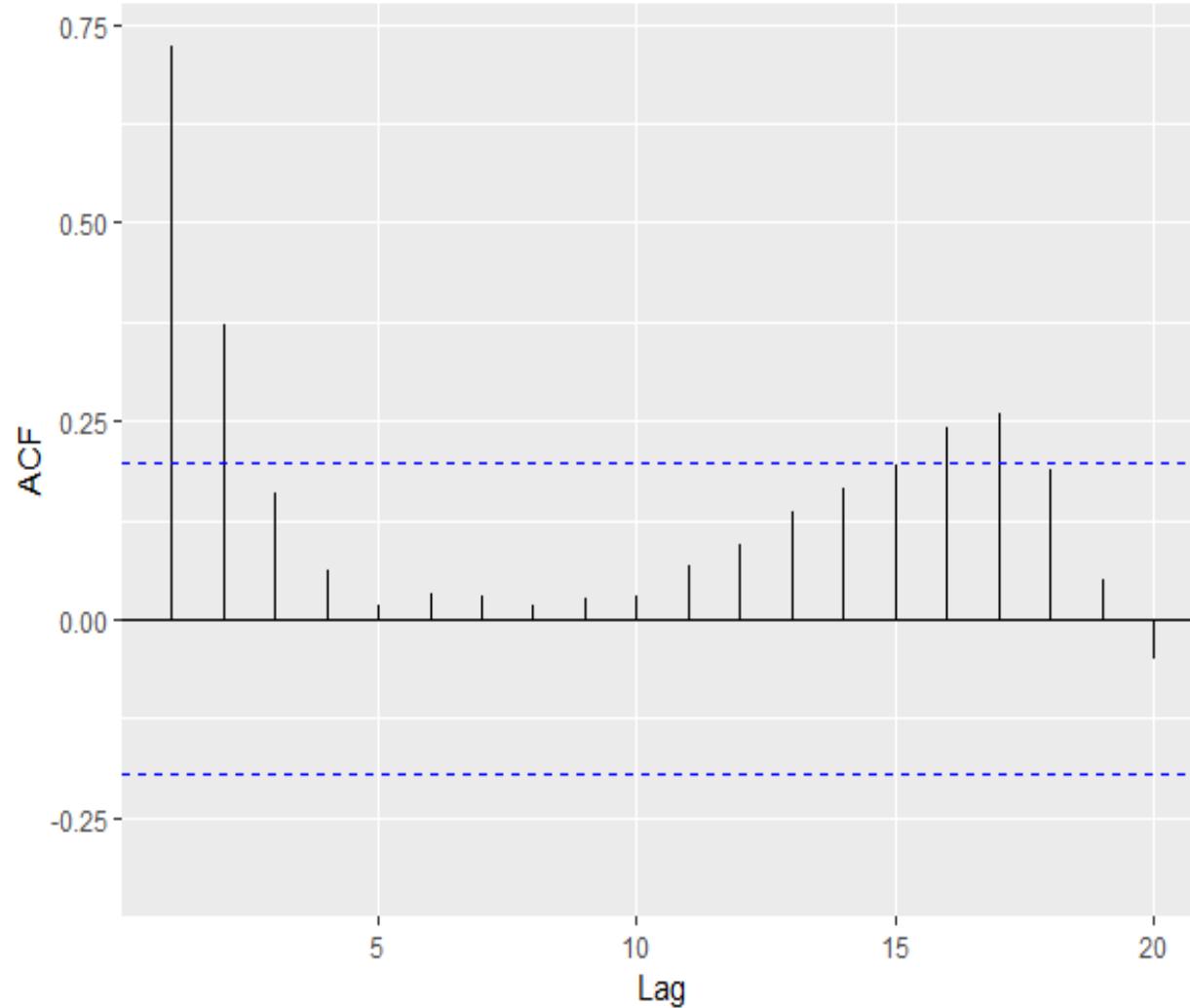
AR(p)

ARMA(p,q)

Identificación de  
procesos

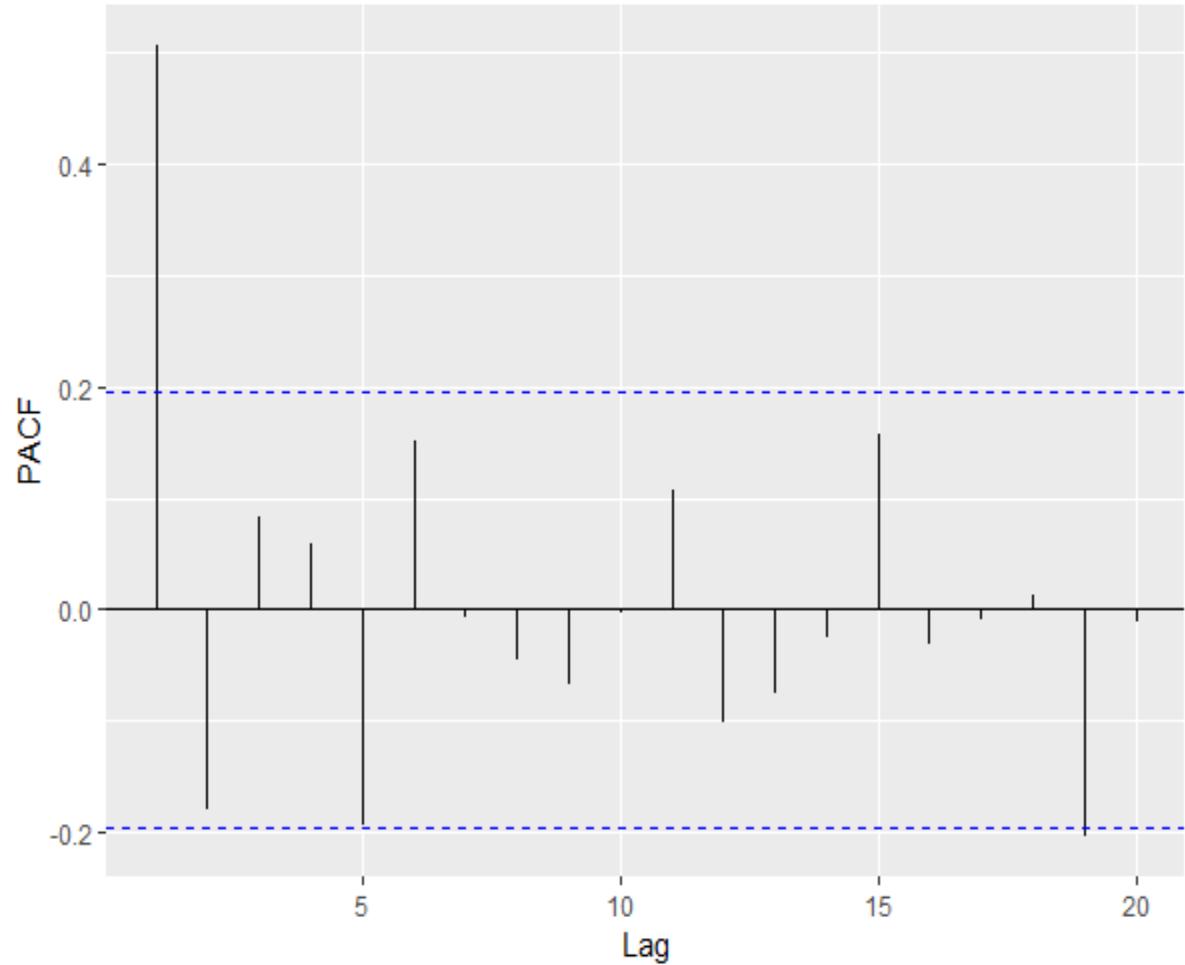
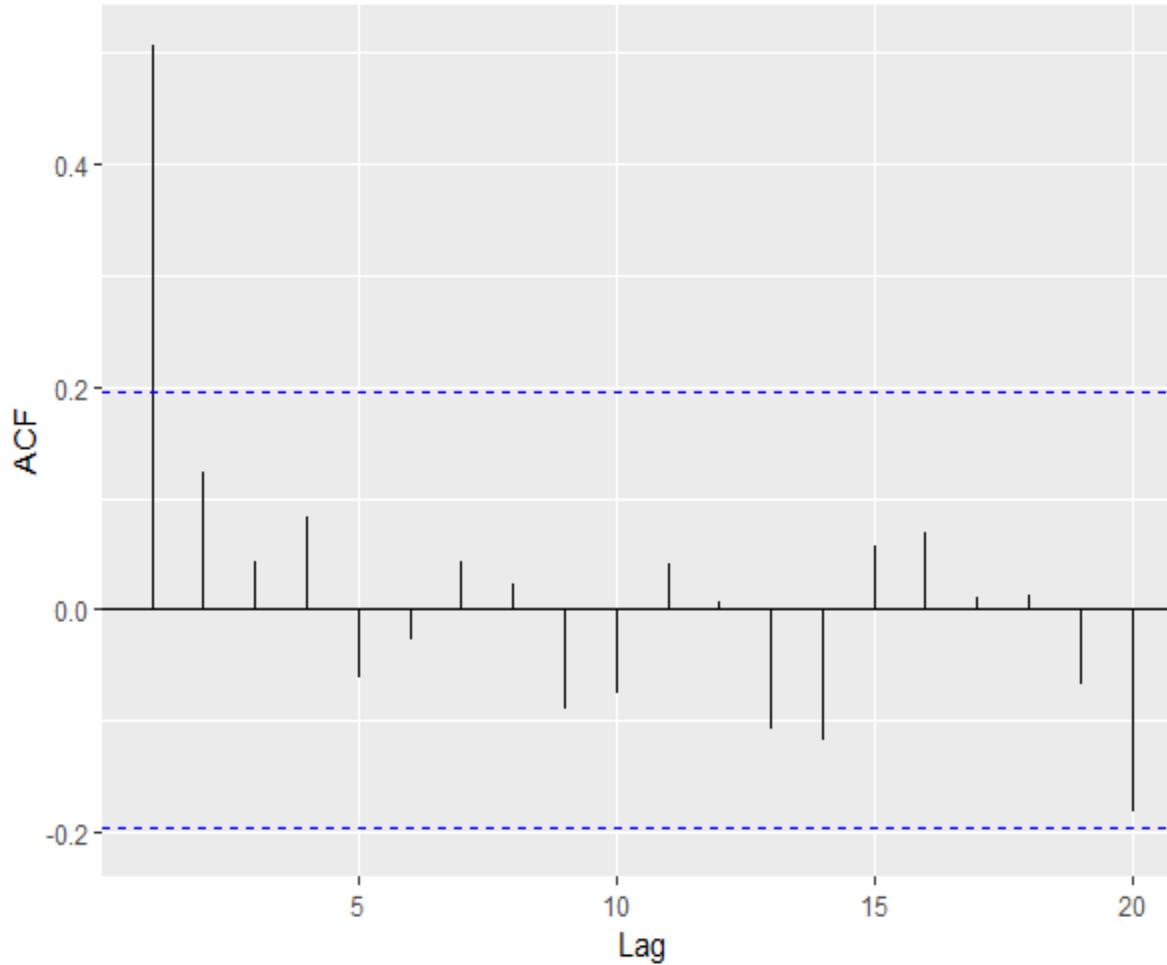
# Identificación: ¿cuál es el proceso?

Este es un proceso generado mediante una simulación. ¿Qué creen que es este proceso?



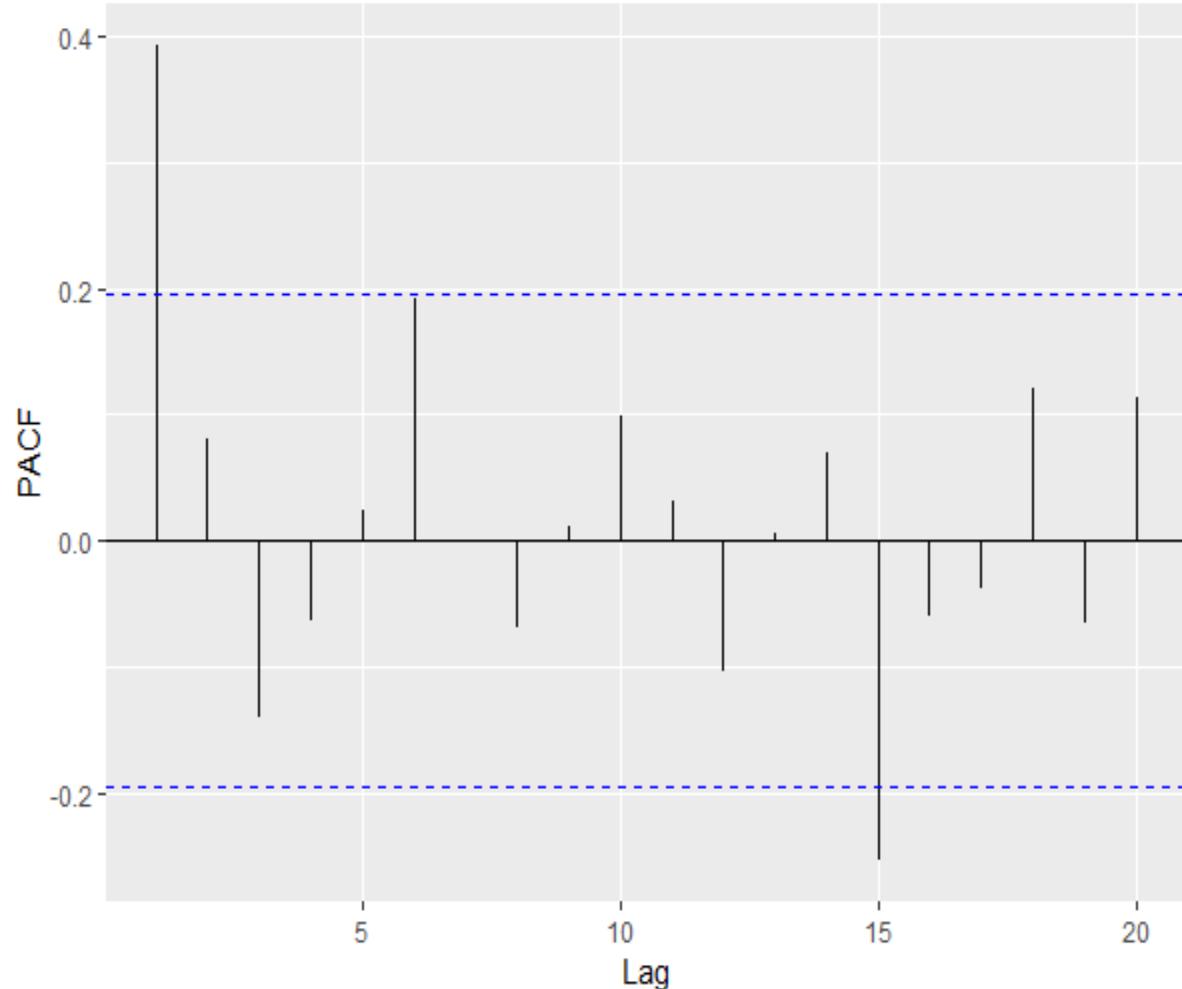
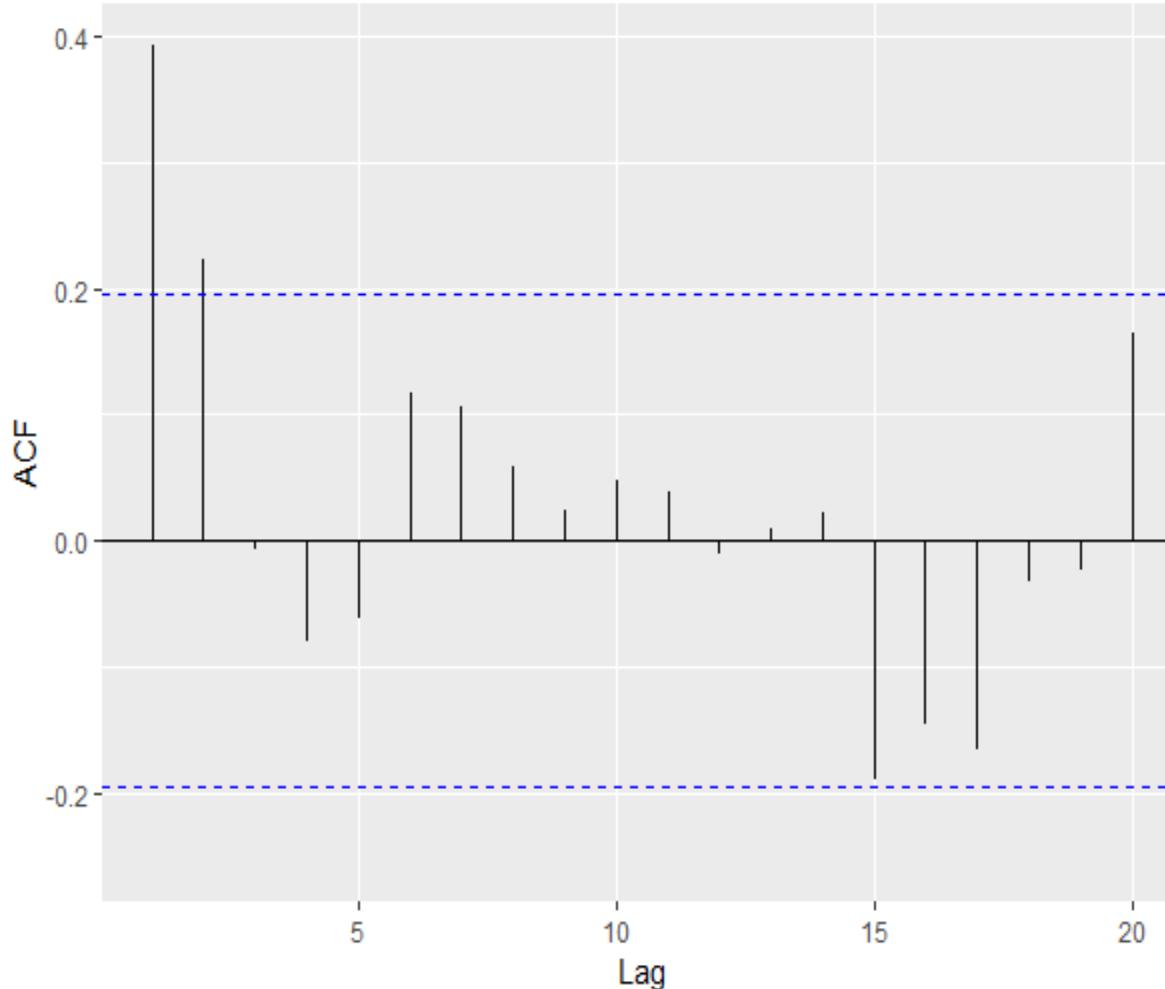
# Identificación: ¿cuál es el proceso?

Otro ejemplo. Este es un proceso generado mediante una simulación. ¿Qué creen que es este proceso?



# Identificación: ¿cuál es el proceso?

Por último. Este es un proceso generado mediante una simulación. ¿Qué creen que es este proceso?



# Identificación: ¿cuál es el proceso?

En la identificación del modelo, partimos de una serie  $y_t$  que no es estacionaria, para luego de haber aplicado las diferenciaciones y el log, llegar a una serie  $x_t$  que si lo es. En este proceso, debemos hacer un correcto análisis del correlograma para así determinar o identificar un posible modelo  $ARIMA(p,d,q)(P,D,Q)_s$ .

Los procesos anteriores resultaron ser:

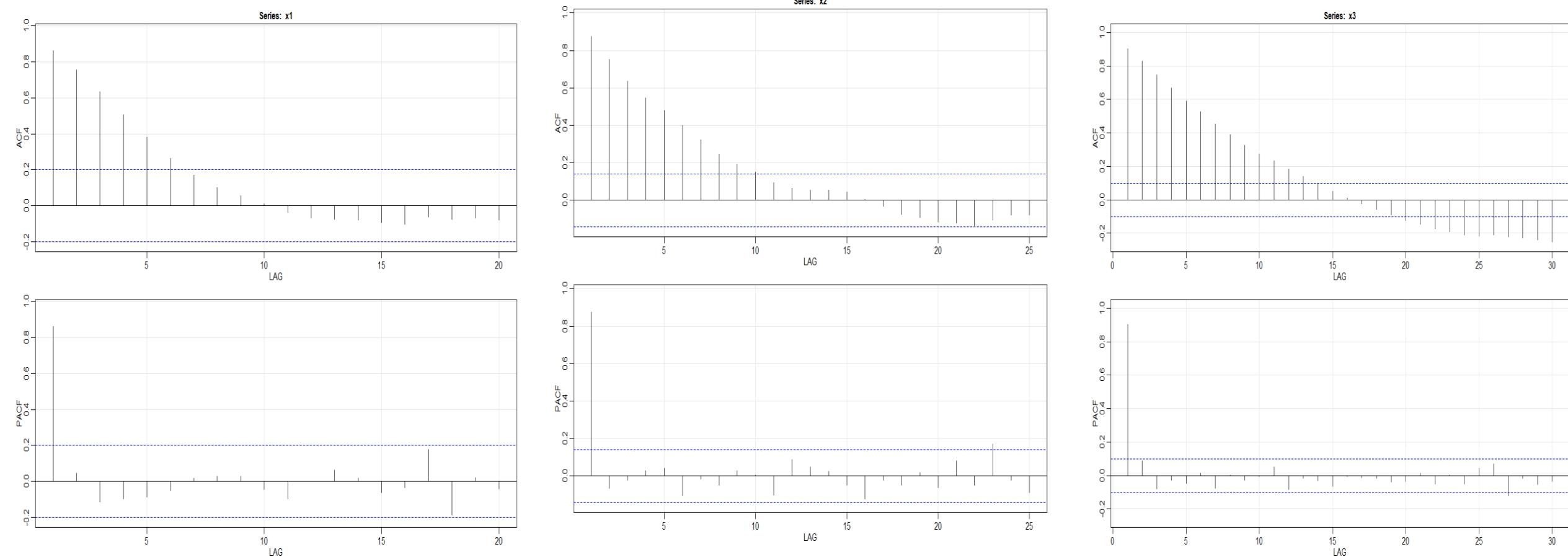
1.  $x_t = 0.9x_{t-1} - 0.2x_{t-2} + u_t$   $ARIMA(2,0,0)$
2.  $x_t = u_t + 0.8u_{t-1} + 0.4u_{t-2}$   $ARIMA(0,0,2)$
3.  $x_t = 0.9x_{t-1} - 0.2x_{t-2} + u_t + 0.8u_{t-1} + 0.4u_{t-2}$   $ARIMA(2,0,2)$

Las identificaciones anteriores fueron tramposas. De buenas a primeras resulta difícil identificar los  $AR(2)$ ,  $MA(2)$  y sobre todo los  $ARMA(2,2)$ . Resulta aconsejable primero estimar los modelos con órdenes más bajos, y luego ir aumentando el orden del ARIMA. Eso lo llamamos como la sobre parametrización o adición de términos, y lo veremos en la siguiente sub-sección.

Se debe de tener especial cuidado a la hora de identificar el modelo, y no caer en la trampa de fijarse si las distintas autocorrelaciones totales y parciales sobrepasan o no los límites de confianza. Si se hiciera caso omiso a esto, entonces no podríamos verdaderamente identificar un modelo ARIMA, y se podría llegar a tener un modelo  $ARIMA(0,0,0)$ , el cual es un ruido blanco puro, y eso no es lo que deseamos... Expliquemos esto.

# Identificación: ¿cuál es el proceso?

En todos los casos estamos estimando un AR(1):  $x_t = 0.9x_{t-1} + u_t$ , pero con tamaños de muestras distintos ( $n = 100, n = 200, n = 300$ ). Nótese como para los distintos tamaños de  $n$ , los límites o intervalos de confianza van siendo diferentes. Se debe de tener mucho cuidado: lo que importa es identificar el proceso que gobierna el proceso  $x_t$ , y aunque son una guía, los límites de confianza pueden confundir el proceso y despistar el correcto modelo ARIMA a la hora tanto de identificar como de estimar el proceso de  $x_t$ .



# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

ARMA(p,q)

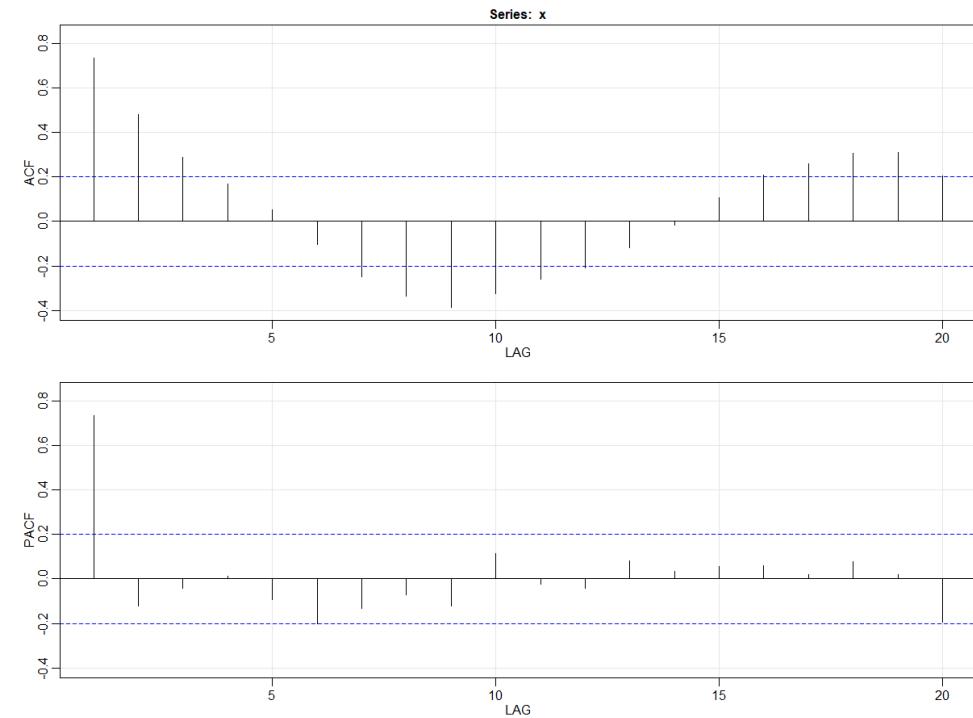
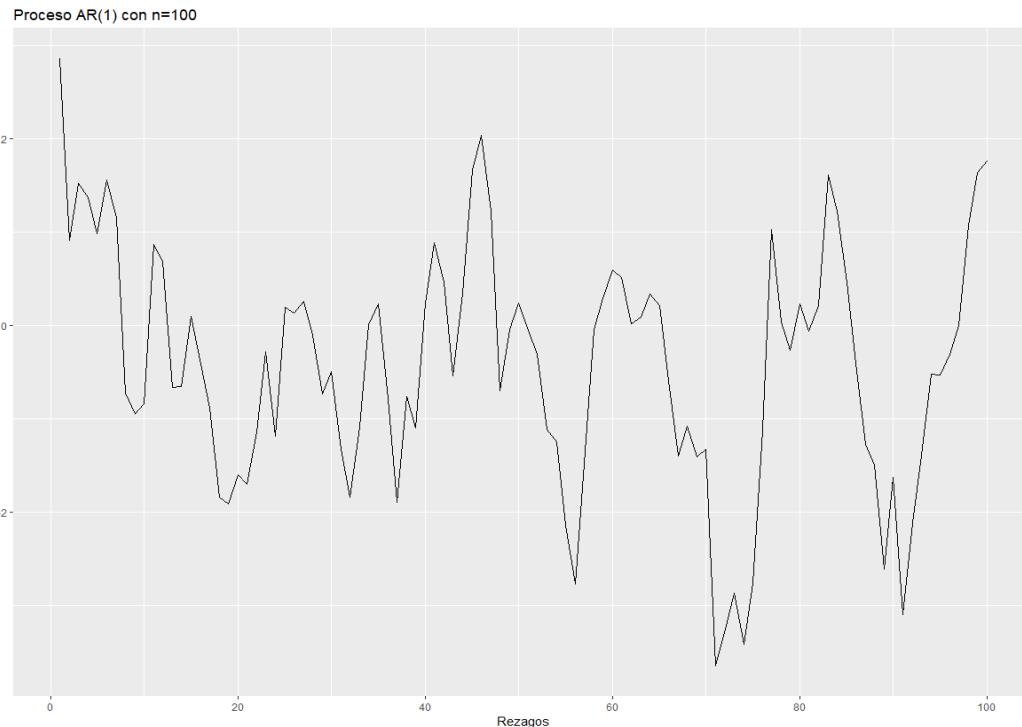
Identificación de  
procesos

Estimación y adición  
de parámetros

# Estimación de los procesos ARMA

La estimación de los modelos  $ARIMA(p, d, q)(P, D, Q)_s$  se llevan a cabo mediante la optimización de la verosimilitud, y para eso recurre a la optimización numérica en la búsqueda de los mejores coeficientes ARIMA que maximizan dicho criterio.

En R se pueden utilizar diversas funciones para llevar a cabo la estimación del proceso: `arima()`, `Arima()`, `auto.arima()`, etc. Nosotros trabajaremos con la función de `sarima()`, lo cual no posee tantas restricciones como las otras, y permite estimar modelos tanto no estacionales como estacionales. Para eso, generaremos un proceso AR(1), con un coeficiente de 0.9, a partir de 100 observaciones, y veamos su autocorrelograma.



# Estimación de los procesos ARMA

Aunque sabemos el proceso generador, procedemos a estimar el modelo como un *ARIMA* (1,0,0) mediante la función sarima: *sarima(x, p = 1, d = 0, q = 0)*. Los valores o argumentos p,d,q no son necesarios. Obtenemos como resultados:

Estimación

```
Call:  
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  
Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,  
REPORT = 1, reltol = tol))  
  
Coefficients:  
      ar1    xmean  
     0.8473  0.8621  
  s.e.  0.0514  0.6065  
  
sigma^2 estimated as 0.9452:  log likelihood = -139.71,  aic = 285.41
```

Significancia

```
$ttable  
  Estimate      SE t.value p.value  
ar1    0.8473 0.0514 16.4834 0.0000  
xmean  0.8621 0.6065  1.4215 0.1583
```

Criterios de información



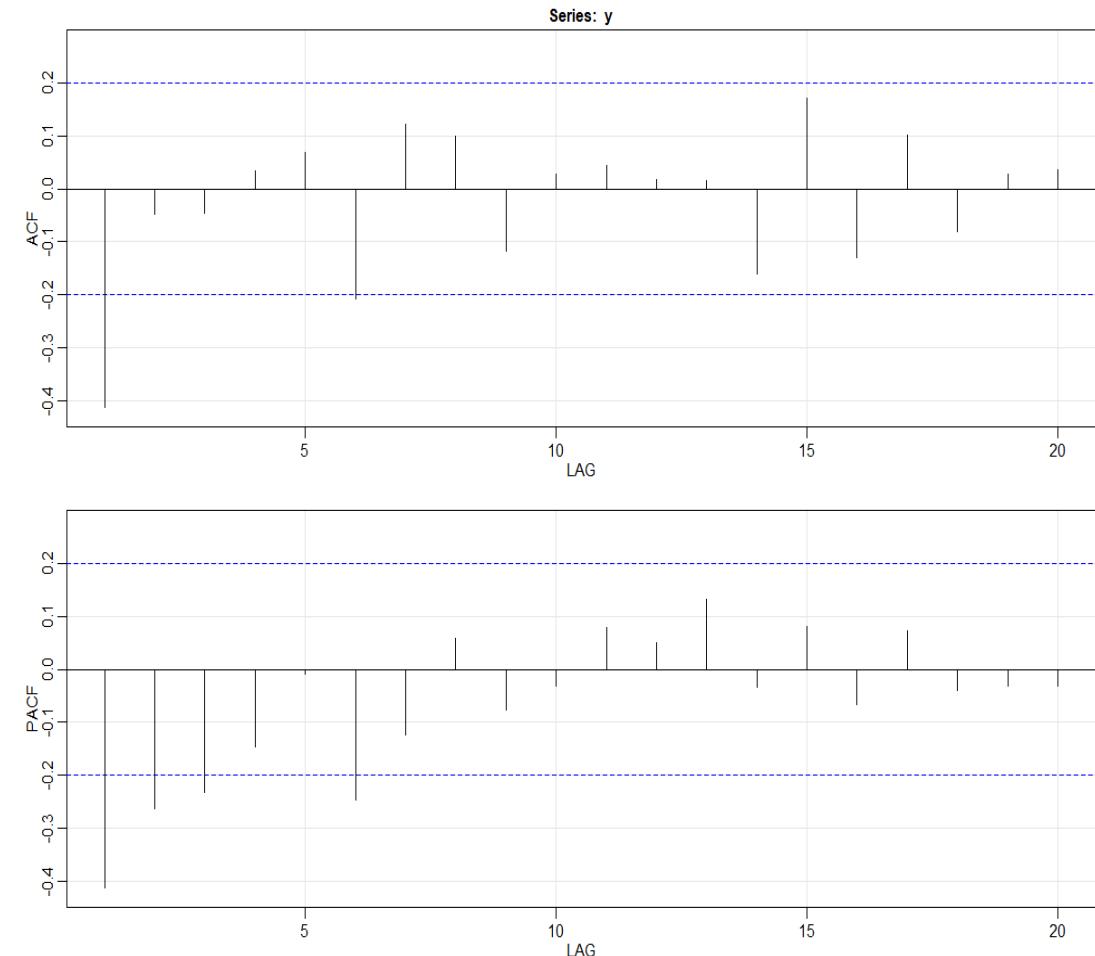
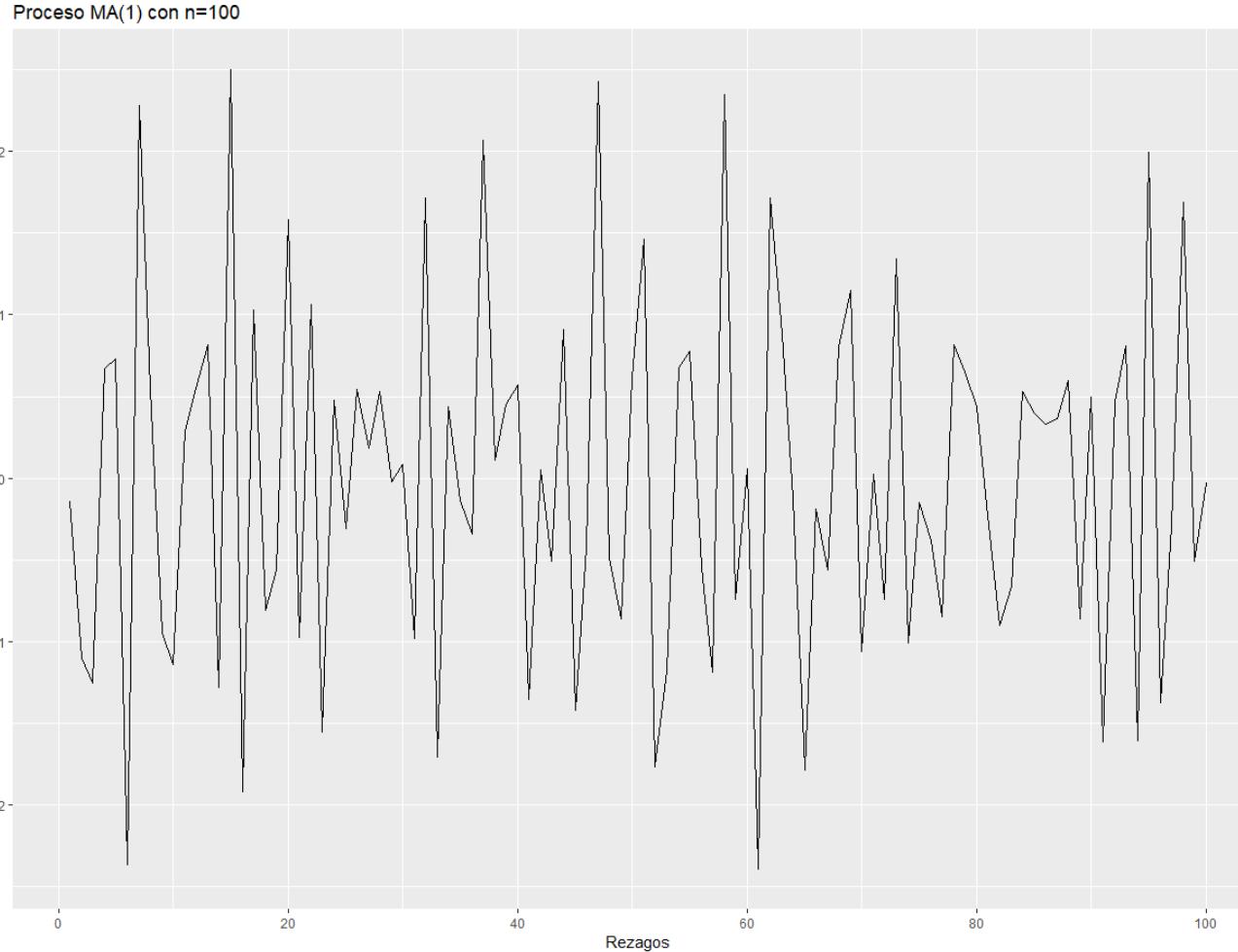
```
$AIC  
[1] 0.9836098  
  
$AICC  
[1] 1.00611  
  
$BIC  
[1] 0.03571325
```

Nótese que:

1. El valor estimado del *AR*(1) es de 0,8473
2. El coeficiente *AR*(1) es significativo

# Estimación de los procesos ARMA

Realicemos el mismo ejercicio pero ahora estimado un proceso MA(1), con un coeficiente de -0.8, a partir de 100 observaciones, y veamos su representación y su autocorrelograma.



# Estimación de los procesos ARMA

Aunque sabemos el proceso generador, procedemos a estimar el modelo como un *ARIMA* (0,0,1) mediante la función *sarima*: *sarima(x, p = 0, d = 0, q = 1)*. Obtenemos como resultados:

## Estimación

```
Call:  
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  
Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,  
REPORT = 1, reltol = tol))  
  
Coefficients:  
      ma1     xmean  
    -0.7594 -0.0047  
  s.e.   0.0660  0.0226  
  
sigma^2 estimated as 0.821:  log likelihood = -132.46,  aic = 270.92
```

## Significancia

```
$ttable  
      Estimate      SE   t.value p.value  
ma1    -0.7594 0.0660 -11.5131  0.000  
xmean   -0.0047 0.0226  -0.2076  0.836
```

## Criterios de información



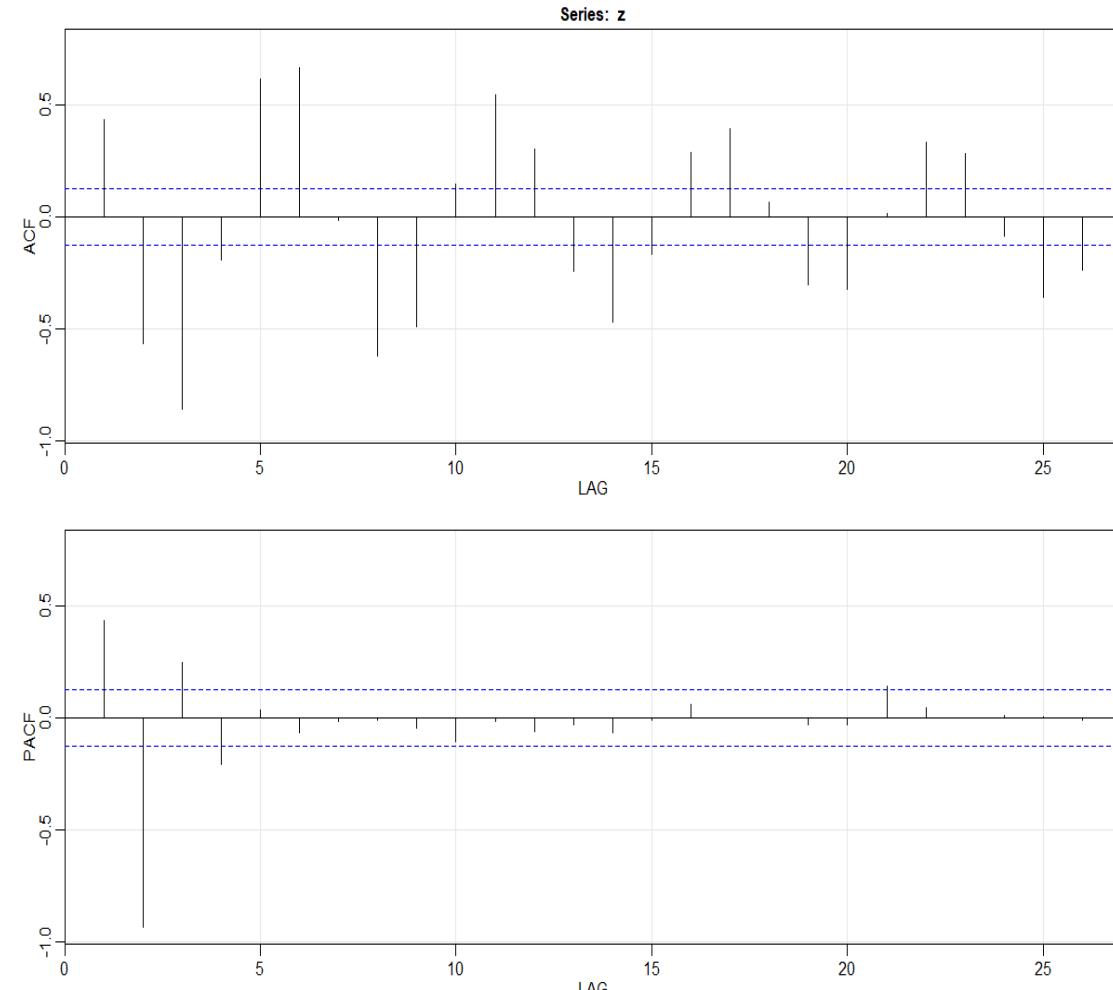
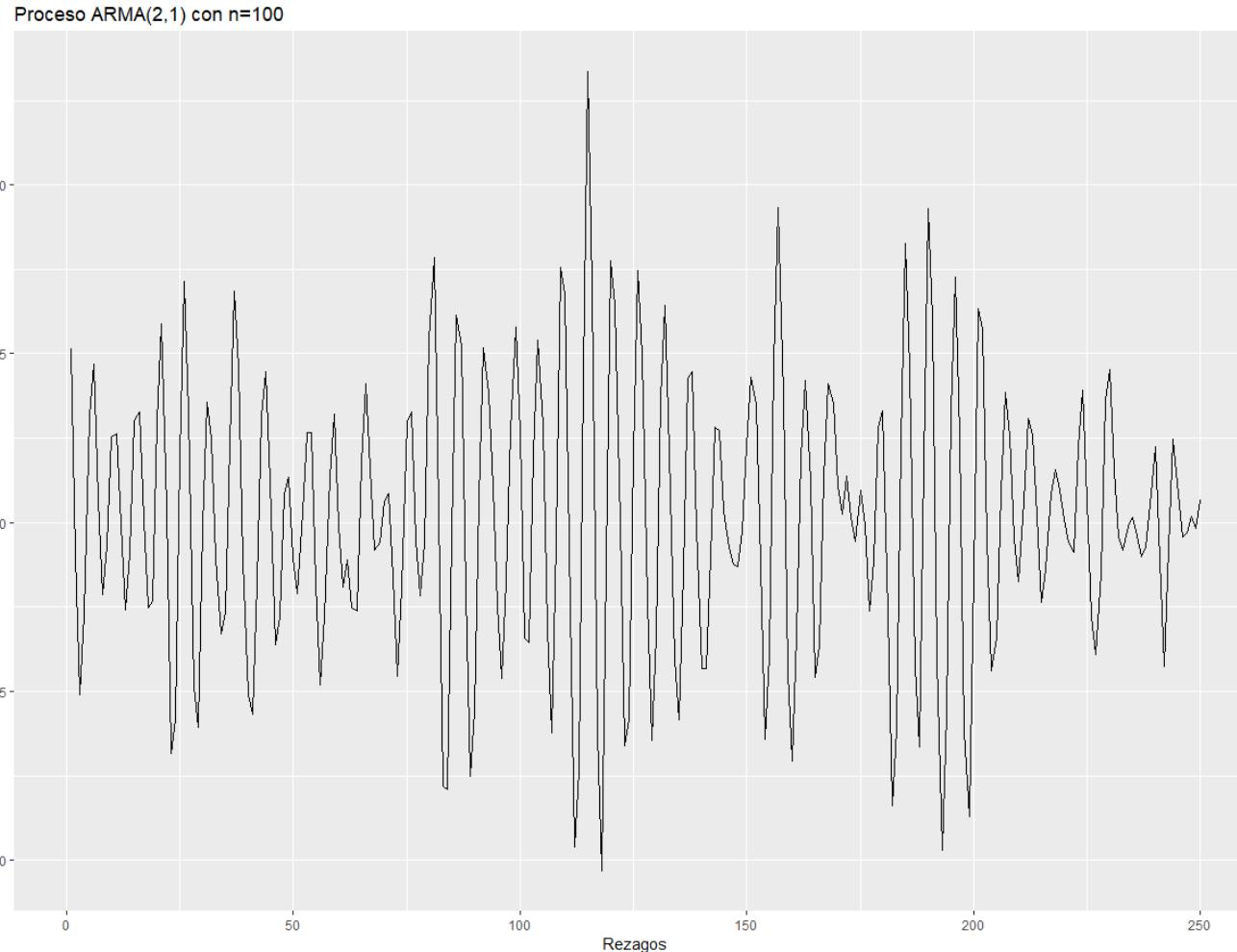
```
$AIC  
[1] 0.8427677  
  
$AICC  
[1] 0.8652677  
  
$BIC  
[1] -0.1051289
```

Nótese que:

1. El valor estimado del *MA*(1) es de -0.759
2. El coeficiente *MA* (1) es significativo

# Estimación de los procesos ARMA

Realicemos el mismo ejercicio pero ahora estimado un proceso ARMA(2,1), con coeficientes AR de 0.8 y -0.9, y MA de 0.8, a partir de 100 observaciones, y veamos su representación y su autocorrelograma.



# Estimación de los procesos ARMA

Aunque sabemos el proceso generador, procedemos a estimar el modelo como un *ARIMA* (2,0,1) mediante la función *sarima*: *sarima(x, p = 2, d = 0, q = 1)*. Obtenemos como resultados:

Estimación

```
Call:  
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  
Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,  
REPORT = 1, reltol = tol))  
  
Coefficients:  
      ar1     ar2     ma1   xmean  
    0.7753 -0.8811  0.7912  0.0385  
  s.e.  0.0313  0.0308  0.0423  0.0988  
  
sigma^2 estimated as 0.926:  log likelihood = -347.98,  aic = 705.95
```

Significancia

```
$ttable  
  Estimate      SE   t.value p.value  
ar1    0.7753 0.0313  24.7374  0.0000  
ar2   -0.8811 0.0308 -28.6287  0.0000  
ma1    0.7912 0.0423  18.7195  0.0000  
xmean   0.0385 0.0988   0.3896  0.6972
```

Criterios de información



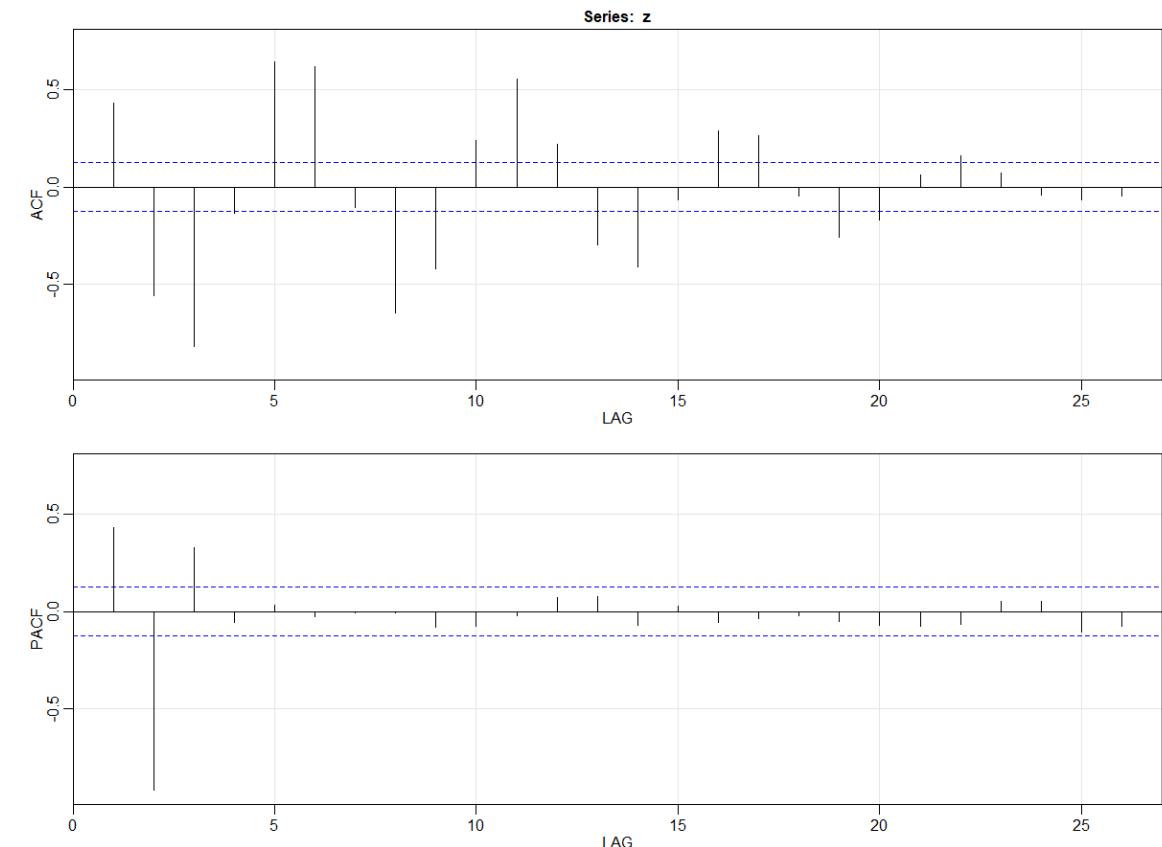
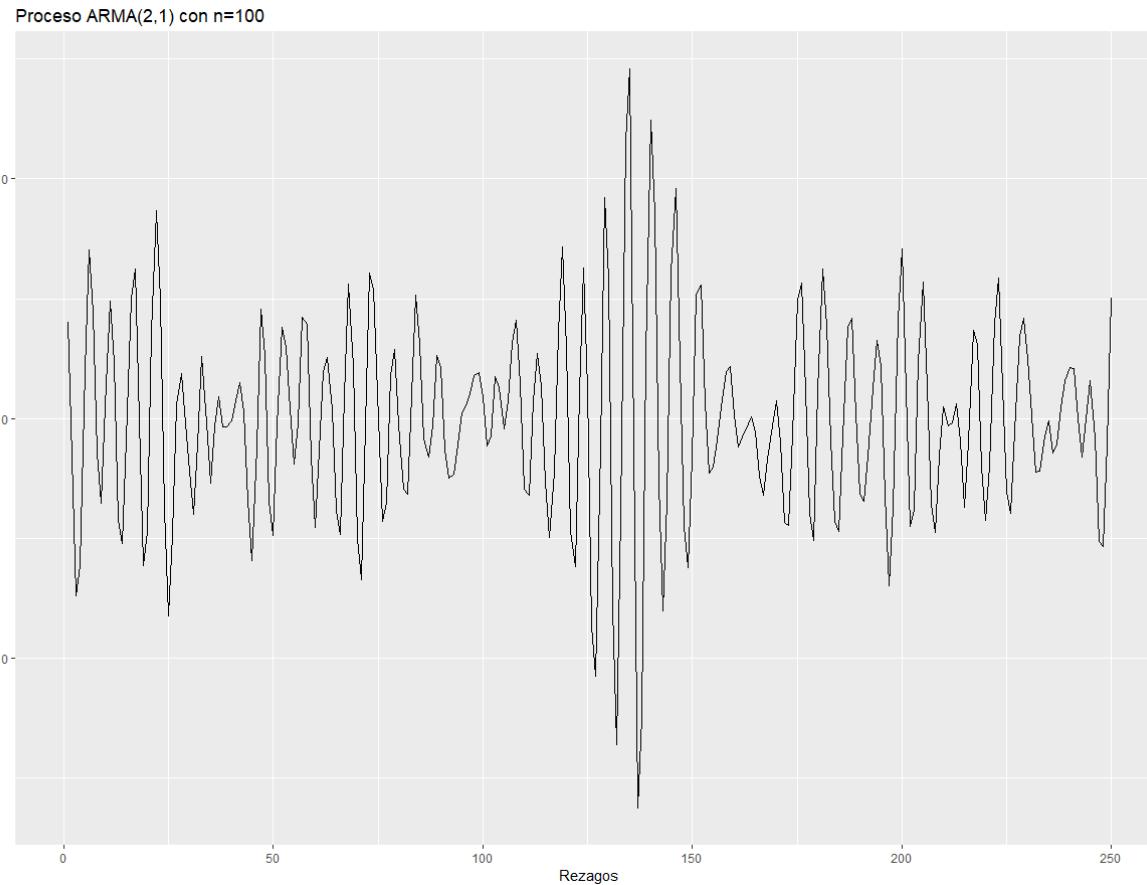
```
$AIC  
[1] 0.955093  
  
$AICC  
[1] 0.9640766  
  
$BIC  
[1] 0.01143635
```

Nótese que:

1. El valor estimado del ar1 es de 0.77, del ar2 de -0.88, y del ma1 de 0.8.
2. Los coeficientes son significativos.

# Estimación de los procesos ARMA

El último modelo ARMA(2,1) es complicado de identificar, y lo recomendable es utilizar el criterio de sobre y sub parametrizar o adicionar parámetros ARMA en la búsqueda del proceso que gobierna  $x_t$ . Lo que hacemos es ver hasta que punto los coeficientes son o no son significativos a la hora de agregar parámetros, en donde respetamos siempre el principio de ***parsimonía***. Volvamos a ver el modelo ARMA(2,1).



# Estimación de los procesos ARMA

ARMA(1,0)

```
> mod.1 $ttable
    Estimate      SE  t.value p.value
ar1     0.4371 0.057  7.6678  0.0000
xmean   -0.0803 0.418 -0.1922  0.8478
> |
```

ARMA(0,1)

```
> mod.2 $ttable
    Estimate      SE  t.value p.value
ma1     0.9817 0.0098 100.0004 0.0000
xmean   -0.0818 0.3152 -0.2594  0.7956
> |
```

ARMA(1,1)

```
> mod.3 $ttable
    Estimate      SE  t.value p.value
ar1     0.3968 0.0586  6.7684  0.0000
ma1     0.9729 0.0106  92.1925 0.0000
xmean   -0.0459 0.4771 -0.0961  0.9235
> |
```

ARMA(2,1)

```
> mod.4 $ttable
    Estimate      SE  t.value p.value
ar1     0.7825 0.0279  28.0299  0.0000
ar2    -0.9016 0.0271 -33.3298  0.0000
ma1     0.8288 0.0479  17.3097  0.0000
xmean   -0.1009 0.1106 -0.9123  0.3625
> |
```

ARMA(2,2)

```
> mod.5 $ttable
    Estimate      SE  t.value p.value
ar1     0.7919 0.0299  26.5118  0.0000
ar2    -0.9038 0.0267 -33.8004  0.0000
ma1     0.7909 0.0654  12.0918  0.0000
ma2    -0.0555 0.0693  -0.8013  0.4237
xmean   -0.1012 0.1055 -0.9595  0.3383
> |
```

ARMA(3,2)

```
> mod.6 $ttable
    Estimate      SE  t.value p.value
ar1     0.0295 0.1913  0.1540  0.8777
ar2    -0.3065 0.1483 -2.0664  0.0398
ar3    -0.6902 0.1683 -4.1004  0.0001
ma1     1.5441 0.2102  7.3452  0.0000
ma2     0.5722 0.1882  3.0404  0.0026
xmean   -0.1020 0.1066 -0.9576  0.3392
> |
```

¿Por qué el ARMA(2,1) es el mejor modelo? Es el modelo que posee la mayor cantidad de coeficientes donde todos son significativos. Los AR(1) y MA(1) sobre estiman o carecen de cierta información, y los ARMA(2,2) y ARMA(3,2) sub estiman el verdadero proceso dado que agrega términos que no son significativos.

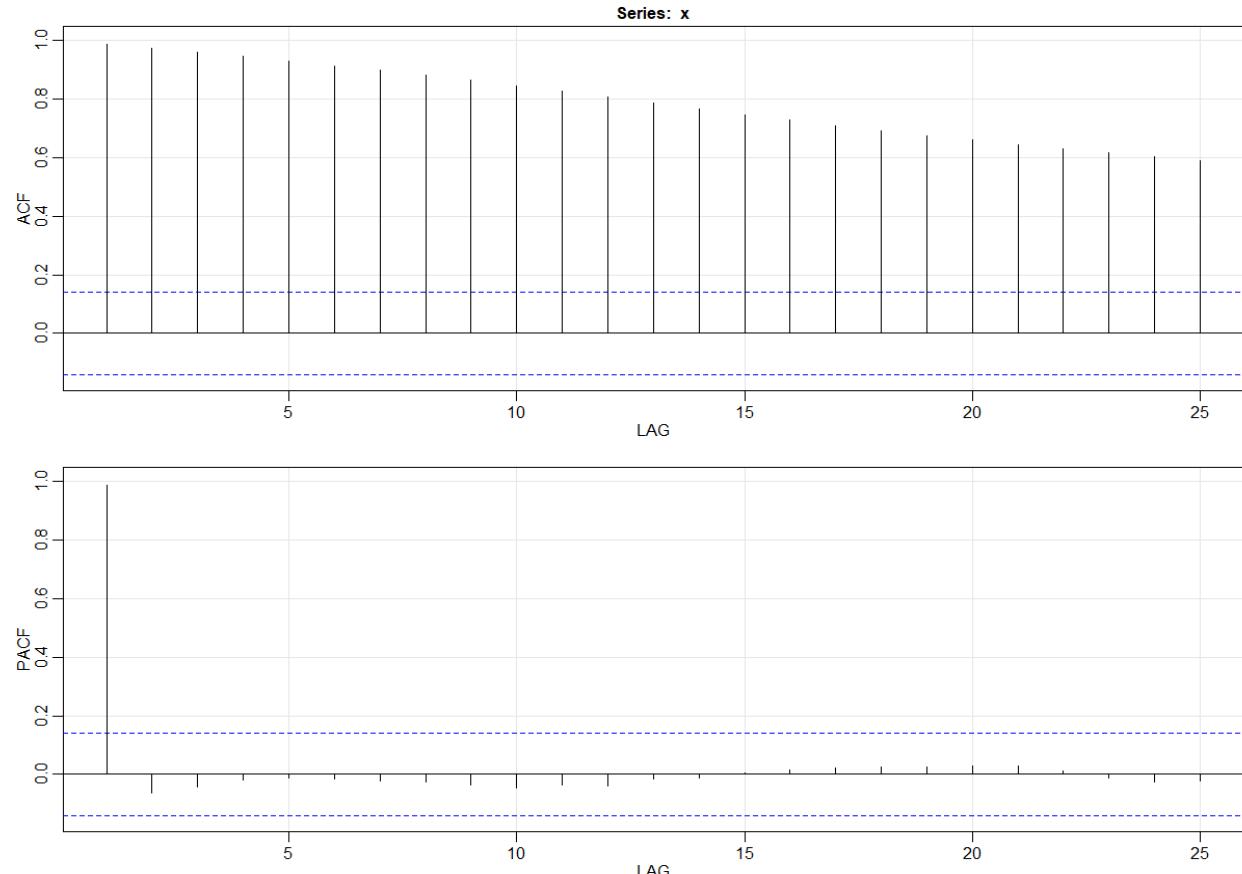
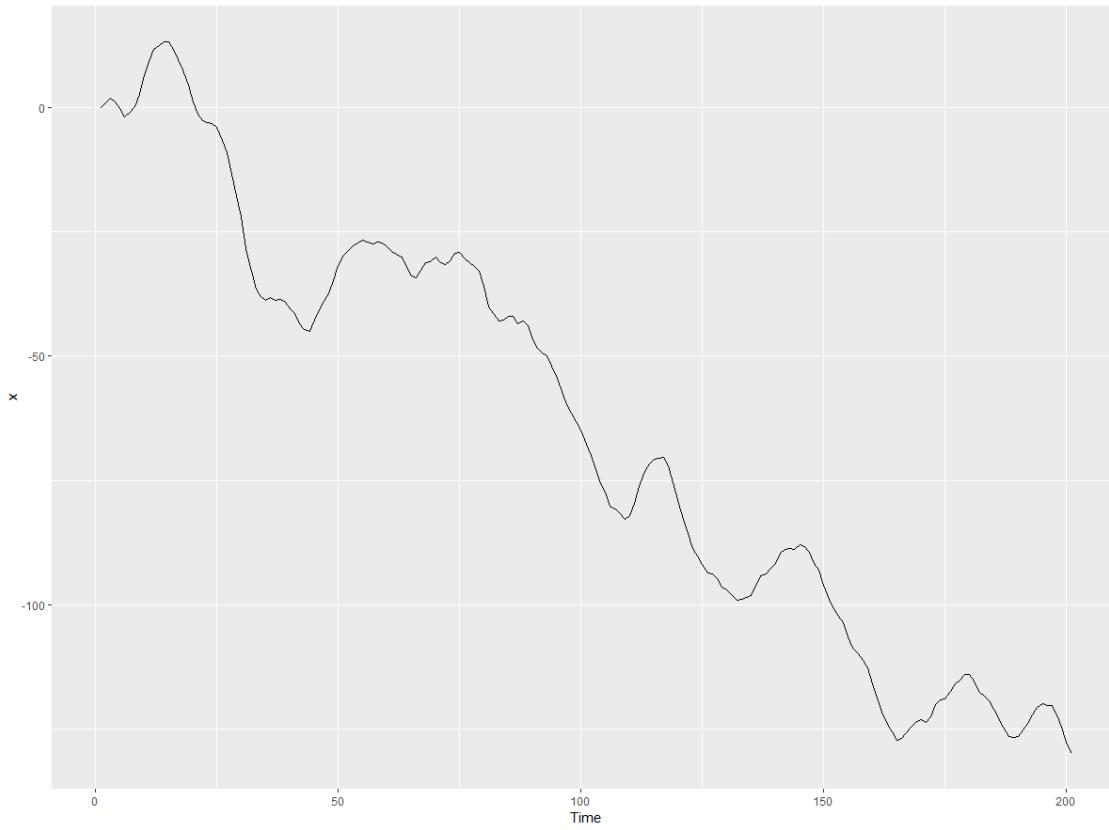
# Estimación de los procesos ARMA

En los modelos  $ARIMA(p, d, q)$ , ¿qué ocurrió con la  $d$ , lo cual hace referencia al término *Integrated*?



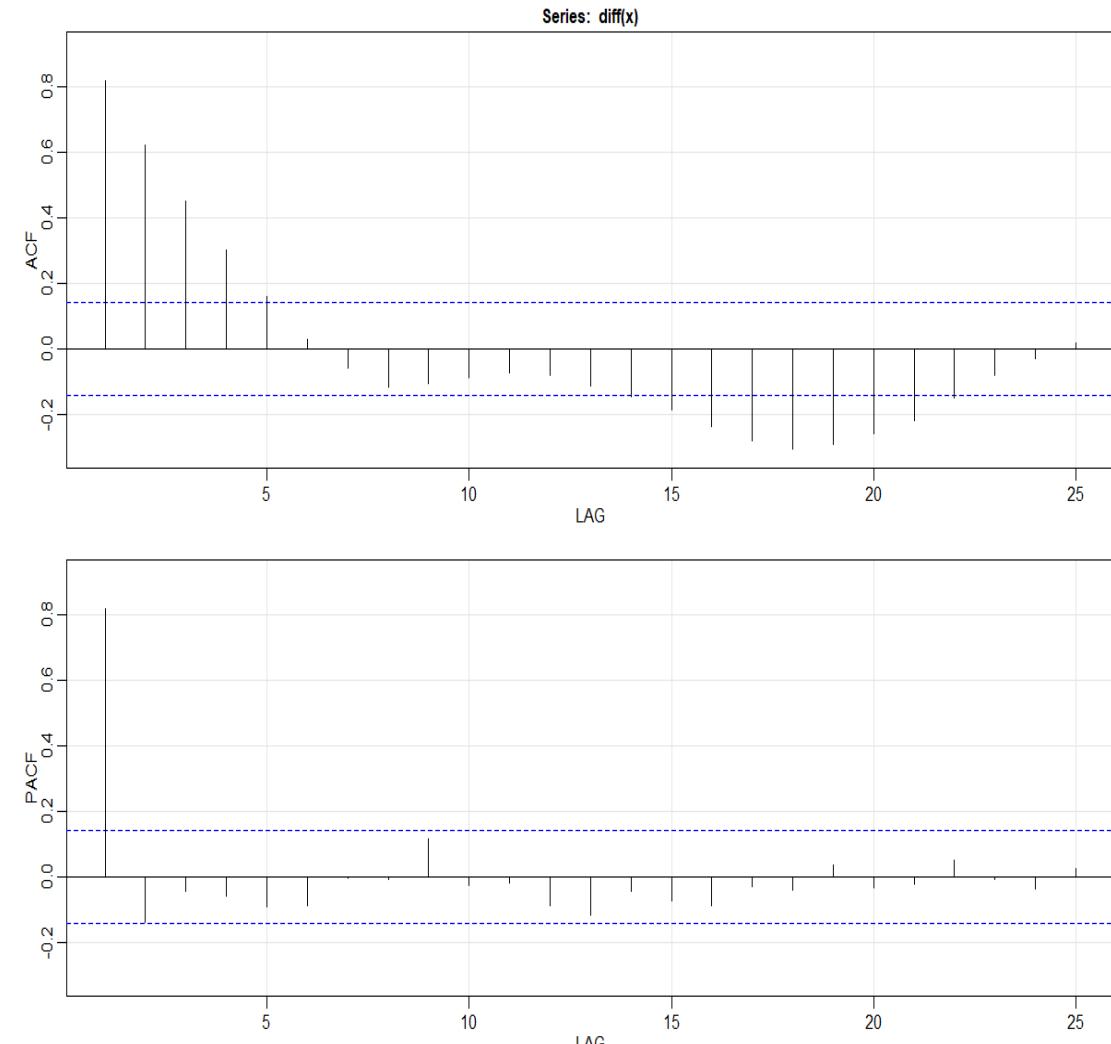
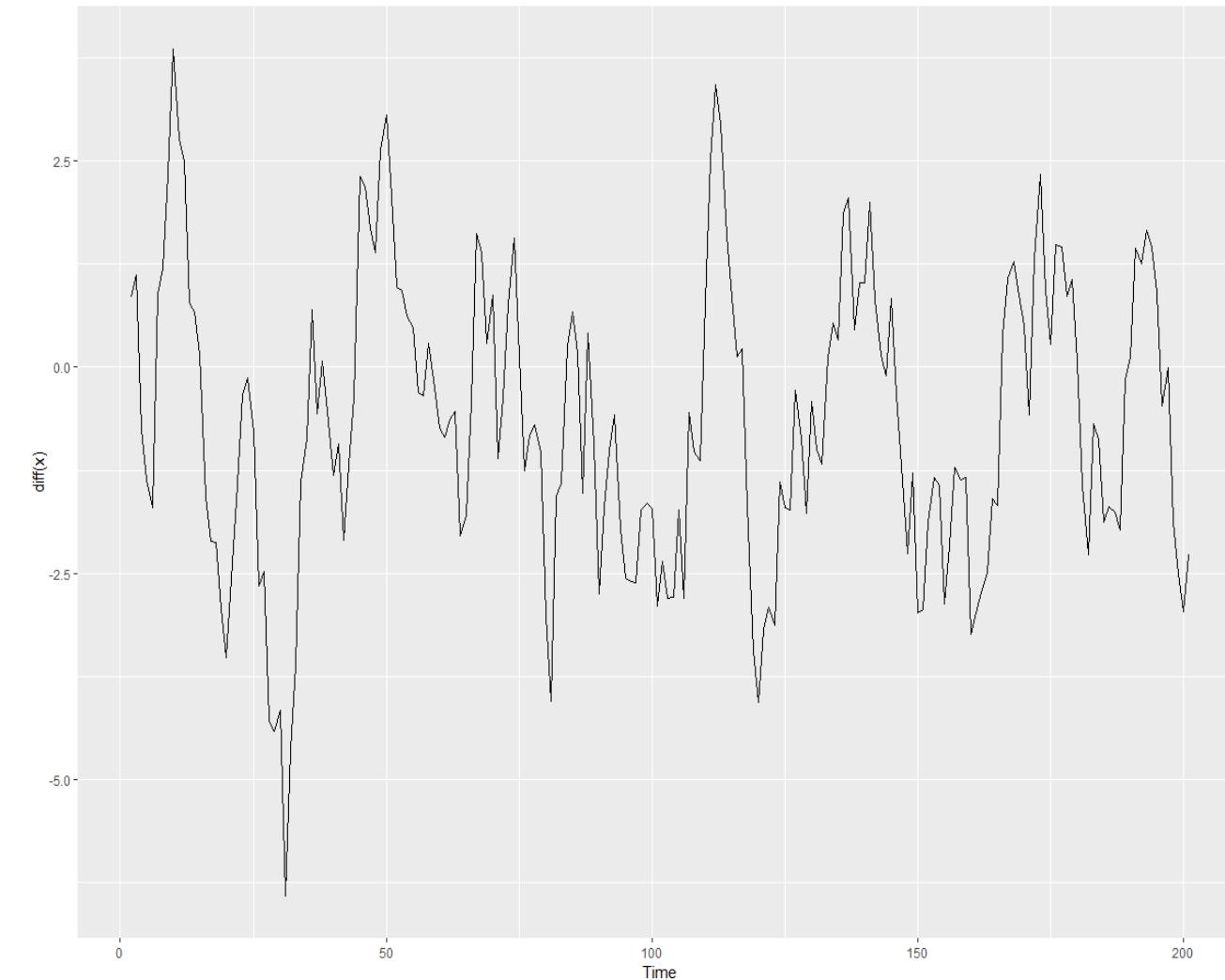
# Estimación de los procesos ARMA

El termino  $d$  en los modelos  $ARIMA(p, d, q)$  se utilizan para estimar el modelo original el cual no suele ser estacionario. Así por ejemplo, todos los modelos generados hasta ahora eran procesos estacionarios, por lo que no nos vimos en la necesidad de estimar algún término  $d$ . Estimemos el siguiente modelo  $Y_t = 0.9Y_{t-1} + u_t$ , para  $Y_t = \nabla X_t = X_t - X_{t-1}$  el cual NO es proceso estacionario.



# Estimación de los procesos ARMA

Veamos el proceso modelo  $Y_t = 0.9Y_{t-1} + u_t$ , para  $Y_t = \nabla X_t = X_t - X_{t-1}$ , mediante la aplicación de la diferenciación de primer orden.



# Estimación de los procesos ARMA

Finalmente, se pasó de un proceso  $Y_t$  no estacionario a  $X_t$  que si lo es. Estimaremos entonces el proceso ARIMA pero para el modelo en cuestión, el  $Y_t$ , a partir de la identificación que vimos en el modelo estacionario  $X_t$ .

Estimación

```
Call:  
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  
Q), period = s), xreg = constant, optim.control = list(trace = trc, REPORT = 1,  
reltol = tol))  
  
Coefficients:  
      ar1  constant  
      0.9356   0.1627  
  s.e.  0.0239   1.0245  
  
sigma^2 estimated as 0.9924:  log likelihood = -284.07,  aic = 574.13
```

Significancia

	Estimate	SE	t.value	p.value
ar1	0.9356	0.0239	39.2141	0.000
constant	0.1627	1.0245	0.1588	0.874

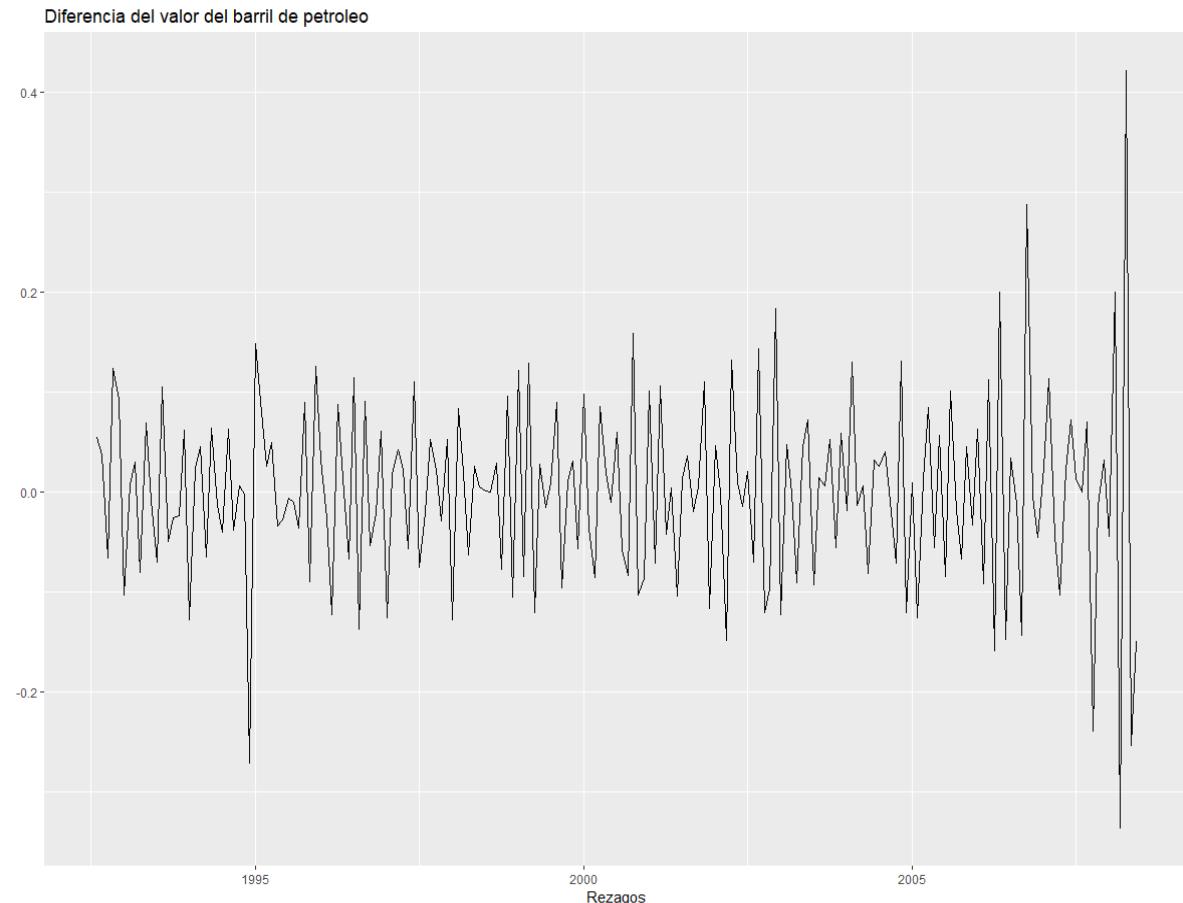
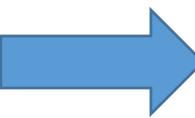
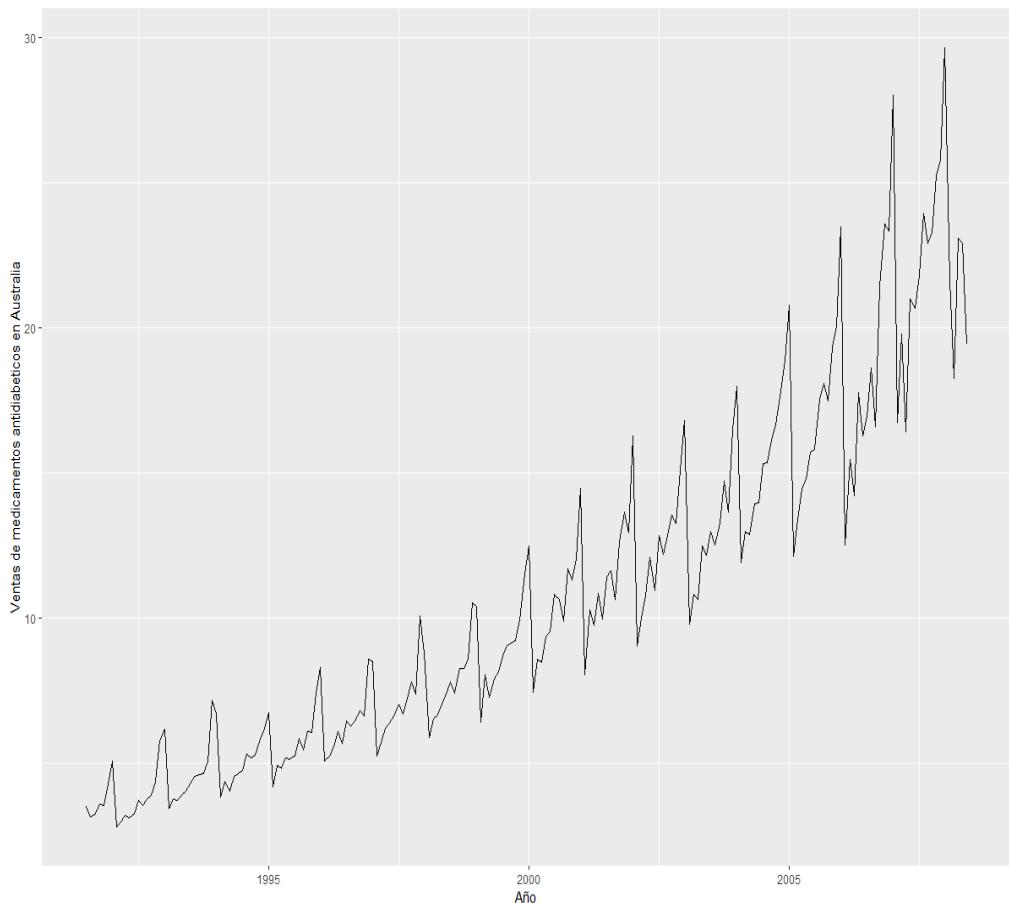
Criterios de información



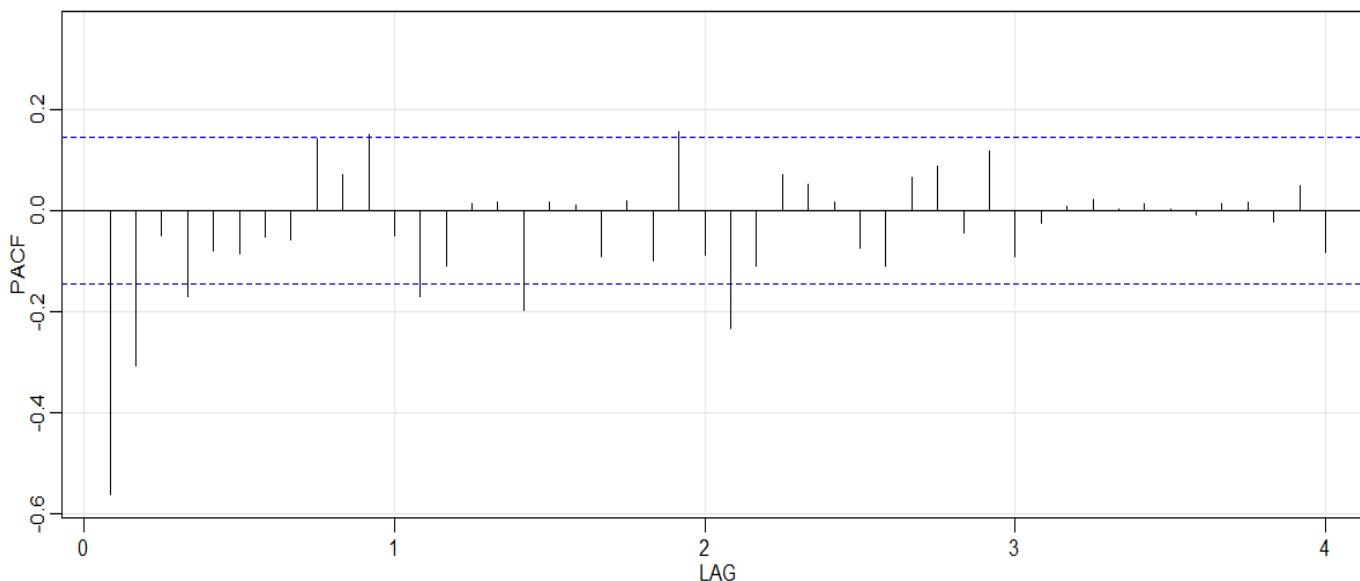
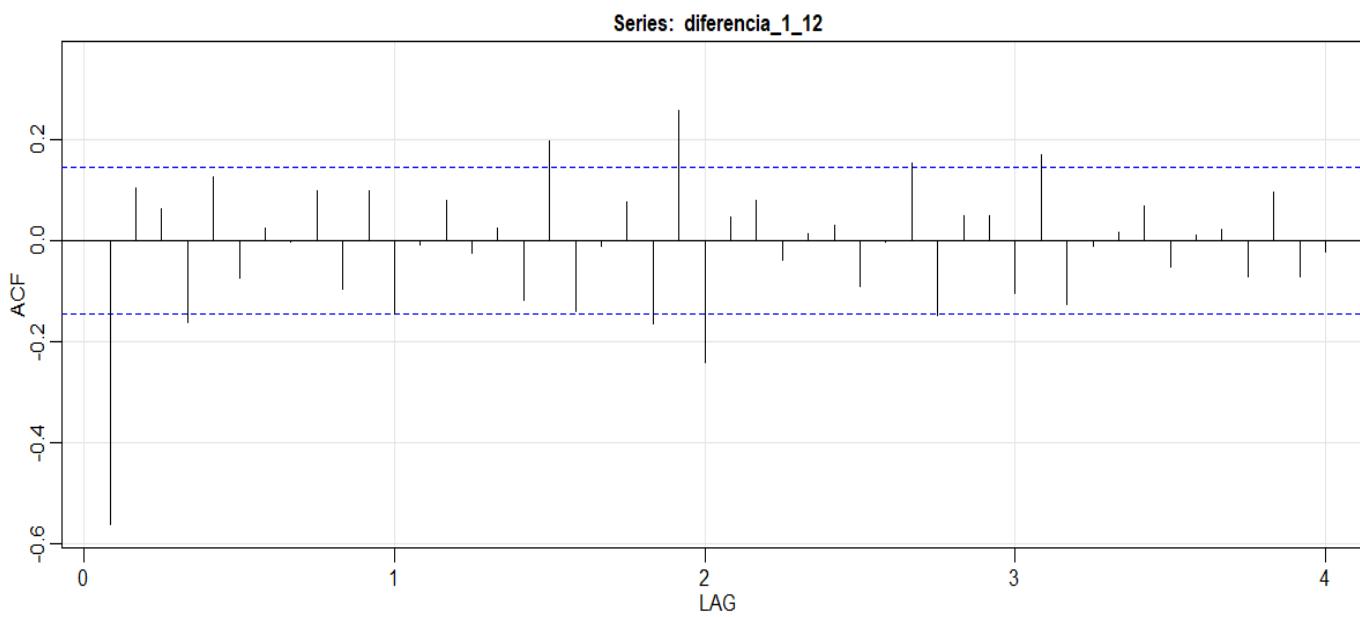
```
$AIC  
[1] 1.01228  
  
$AICC  
[1] 1.022836  
  
$BIC  
[1] 0.04514885
```

# Estimación de los procesos ARMA

Analicemos el archivo de datos a10, sobre las ventas de medicamentos en Australia. Nótese que la serie es no estacionaria, y pues tendremos que hacerla estacionaria como se hizo anteriormente. Apliquemos entonces el log, y las diferenciaciones de orden 1 y 12.



# Estimación de los procesos ARMA



A partir de la serie estacionaria realizamos el proceso de identificación. Dentro de los procesos no estacionarios, ¿qué logramos ver? Parece ser un modelo MA(1), dado que las ACF y la caída pausada de las ACFP.

Nótese que, en los períodos o rezagos 12, 24, 36, etc., se nota la influencia de ambas correlaciones. Por ahora las dejaremos de lado, pero es imprescindible hacer el análisis tanto de la parte no estacional como estacional.

# Estimación de los procesos ARMA

Estimemos un modelo *ARIMA*(0,1,1) para las ventas de medicamentos en Australia.

## Estimación

```
Call:  
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  
Q), period = S), xreg = constant, optim.control = list(trace = trc, REPORT = 1,  
reltol = tol))  
  
Coefficients:  
      ma1   constant  
     -0.8577    0.0971  
  s.e.  0.0442    0.0203
```

## Significancia

	Estimate	SE	t.value	p.value
ma1	-0.8577	0.0442	-19.4012	0
constant	0.0971	0.0203	4.7770	0

## Criterios de información



```
$AIC  
[1] 2.376587  
  
$AICC  
[1] 2.386979  
  
$BIC  
[1] 1.409117
```

Aunque que la estimación se hizo sobre la serie original a10. visualmente parece ser que el *ARIMA*(0,1,1) es correcto, sin embargo este carece aún de mucha información por no contemplar la parte estacional.

# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

ARMA(p,q)

Identificación de  
procesos

Estimación y adición  
de parámetros

Modelos ARIMA  
estacionales

# Los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Sin entrar en la desagregación anterior de los modelos  $AR(p)$ ,  $MA(q)$ , y  $ARMA(p, q)$ , los modelos ARIMA estacionales poseen características similares a los no estacionales, con la gran diferencia que los coeficientes se posicionan sobre  $S$ , o el periodo estacional. La representación está dada por la siguiente ecuación:

$$\Phi_P(B^s)(1 - B_s)^D x_t = c + \Theta_Q(B^s)u_t$$

P: # parámetros **AR** estacionales

D: # de diferenciaciones estacionales

Q: # parámetros **MA** estacionales

donde:  $u_t$  ruido blanco y además

$$\Phi_P(B^s) = (1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \dots - \Phi_{Ps} B^{Ps}) \quad \text{Polinomio en } B^s \text{ de grado P}$$

$$\Theta_Q(B^s) = (1 - \Theta_s B^s - \Theta_{2s} B^{2s} - \dots - \Theta_{Qs} B^{Qs}) \quad \text{Polinomio en } B^s \text{ de grado Q}$$

Es raro ver un ARIMA únicamente estacional... De forma general para el  $ARIMA(p, d, q)(P, D, Q)_s$  tenemos:

$$\underbrace{\Phi_p(B)}_{\text{regular}} \underbrace{\Phi_{Ps}(B^s)}_{\text{estacional}} \underbrace{(1 - B)^d}_{\text{regular}} \underbrace{(1 - B_s)^D}_{\text{estacional}} \tilde{Z}_t = \underbrace{\Theta_q(B)}_{\text{regular}} \underbrace{\Theta_{Qs}(B^s)}_{\text{estacional}} a_t$$

regular

estacional

regular

estacional

regular

estacional

$$\Phi(B) = (1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p)$$

p: orden del polinomio  $\Phi(B)$

$$\Phi_s(B) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps})$$

P: orden del polinomio  $\Phi_s(B)$

$$\Theta(B) = (1 - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q)$$

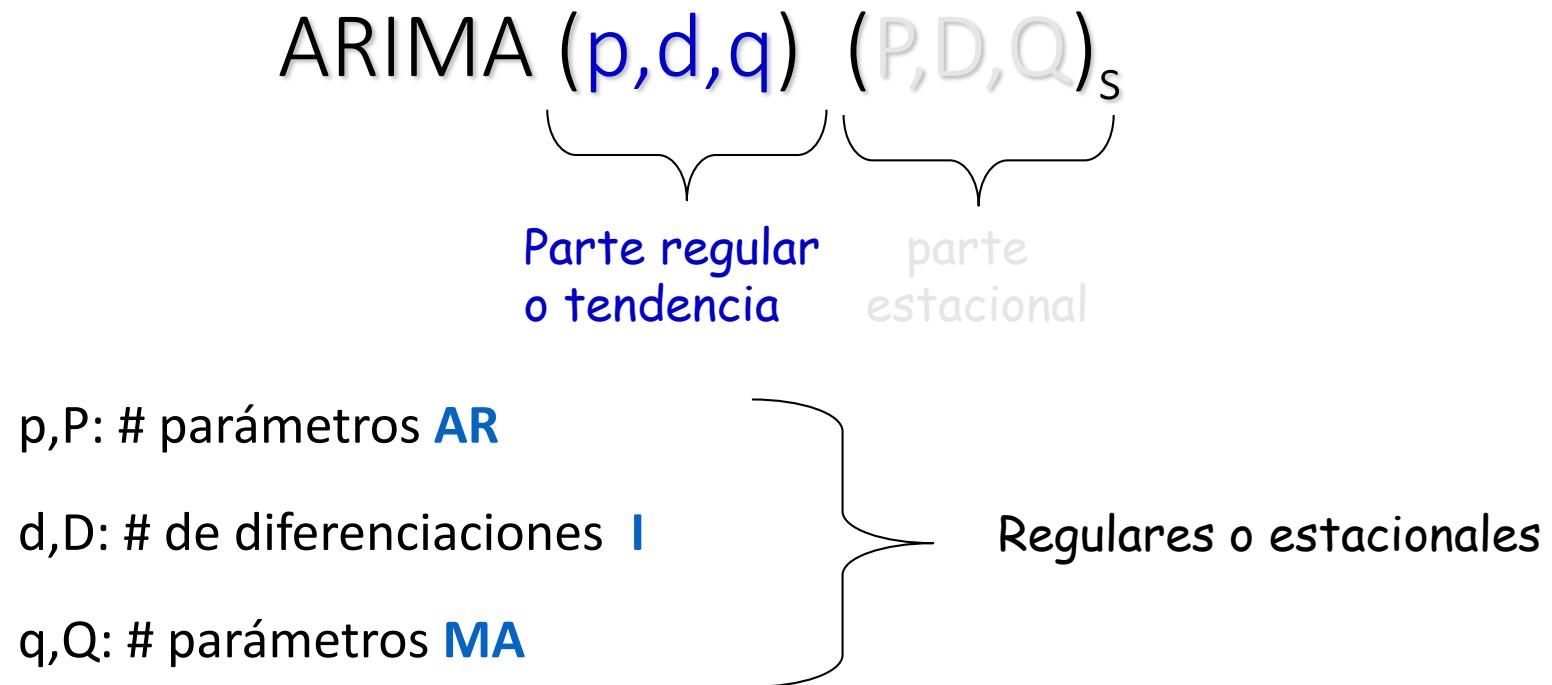
q: orden del polinomio  $\Theta(B)$

$$\Theta_s(B) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs})$$

Q: orden del polinomio  $\Theta_s(B)$

# Los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Por lo tanto, podemos resumir el modelo ARIMA no estacional como estacional de la siguiente forma:



Nótese que en el fondo la diferencia está en cómo en el proceso de estimación se van ajustar coeficientes *AR* o *MA* para la periodicidad *S*. Con respecto al proceso de identificación, utilizamos el autocorrelograma en una serie estacionaria  $x_t$ , pero la identificación de la parte estacional se hace observando los rezagos de orden *S*,  $S_2$ ,  $S_3$ ,..., etc.

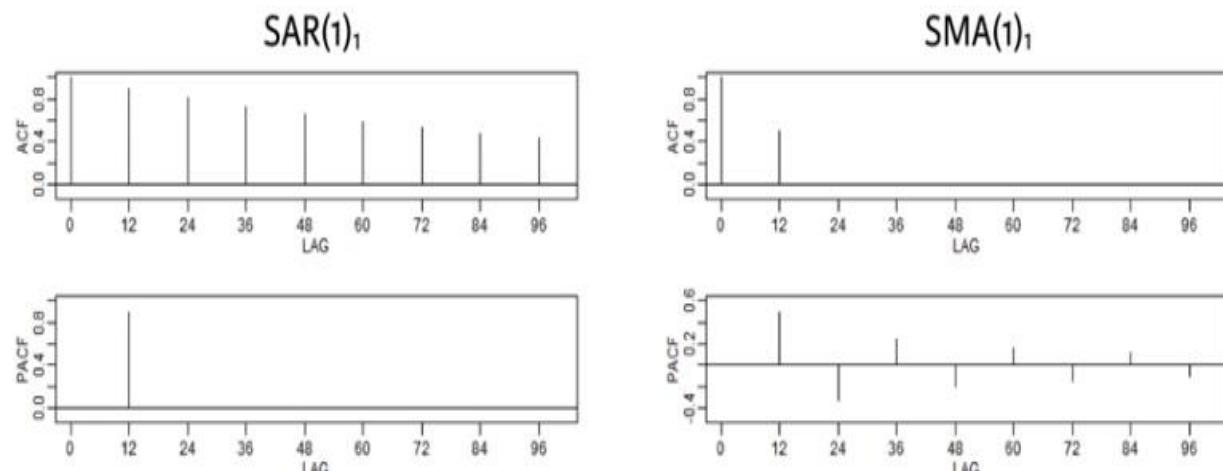
# Los modelos $ARIMA(p, d, q)(P, D, Q)_s$

De forma análoga, podemos tener la siguiente tabla en la identificación de la parte estacional del proceso ARIMA (la otra tabla la obtuve del curso de ARIMA del datacamp.com):

Proceso	Auto-correlación total	Auto-correlación parcial
MA(Q)	Anulación después de los Q primeros coeficientes	Decrecimiento
AR(P)	Decrecimiento	Anulación después de los P primeros coeficientes
ARMA(P,Q)	Decrecimiento	Decrecimiento

	SAR( $P_s$ )	SMA( $Q_s$ )	SARMA( $P, Q_s$ )
ACF*	Tails off	Cuts off lag QS	Tails off
PACF*	Cuts off lag PS	Tails off	Tails off

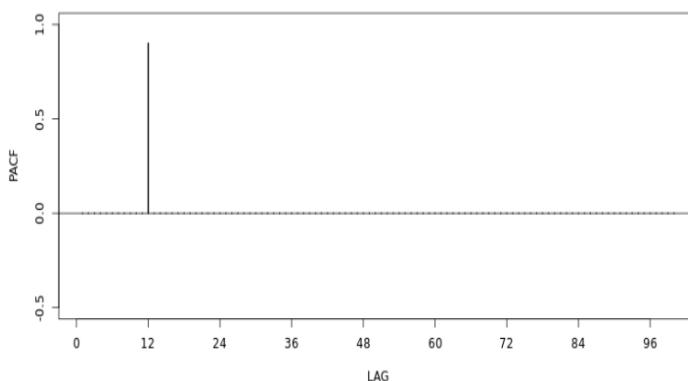
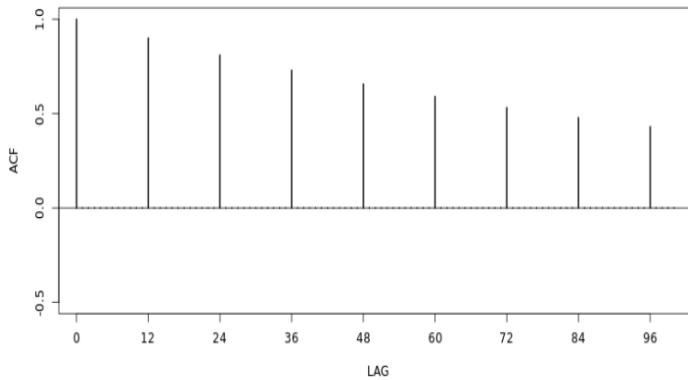
\* The values at the nonseasonal lags are zero



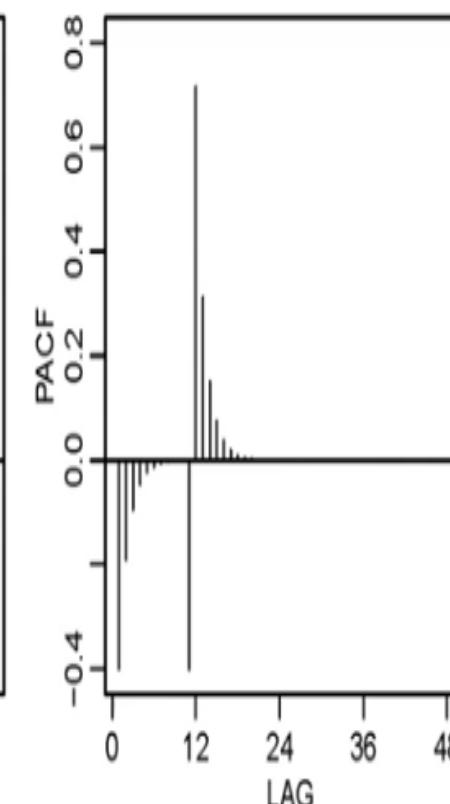
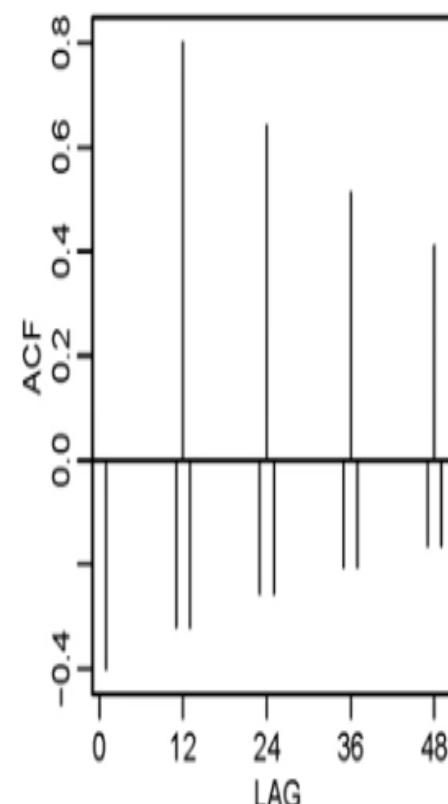
# Los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Veamos los siguientes ejemplos:

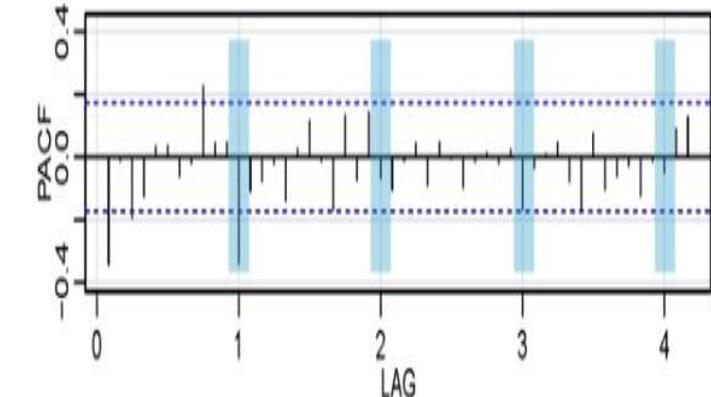
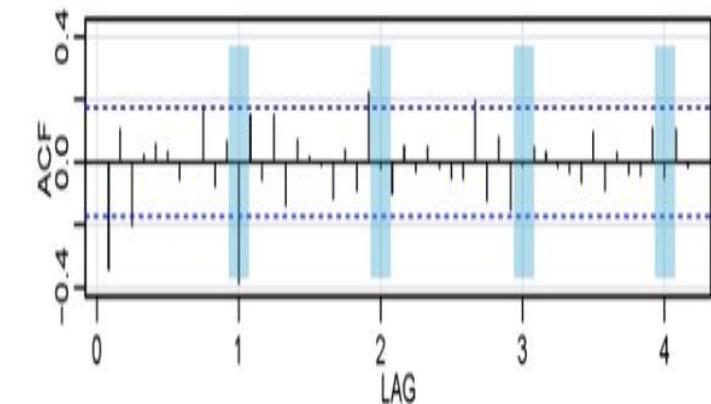
¿Cuál es el proceso estacional?



¿Cuál es el proceso total?



¿Cuál es el proceso total?



# Índice

3

Identificación y  
estimación de un proceso  
ARIMA

MA(q)

AR(p)

ARMA(p,q)

Identificación de  
procesos

Estimación y adición  
de parámetros

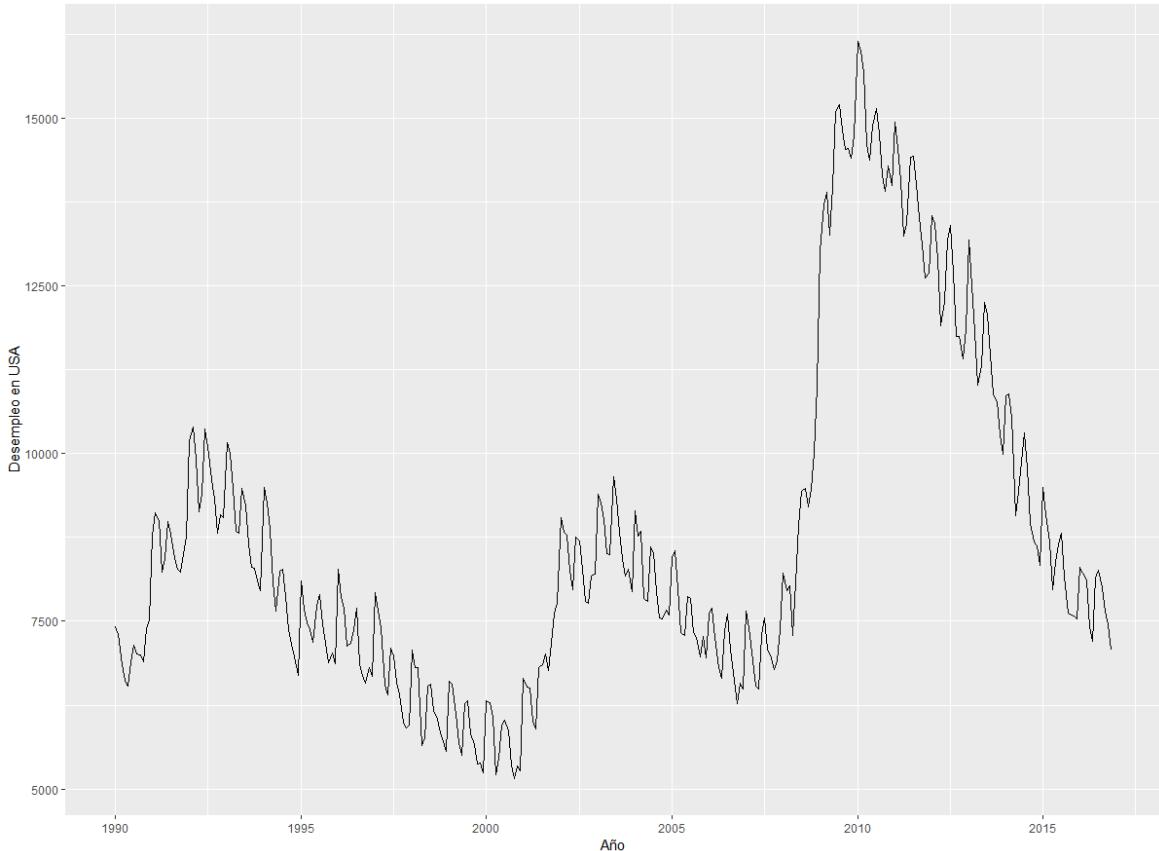
Modelos ARIMA  
estacionales

Estimación ARIMA  
estacional

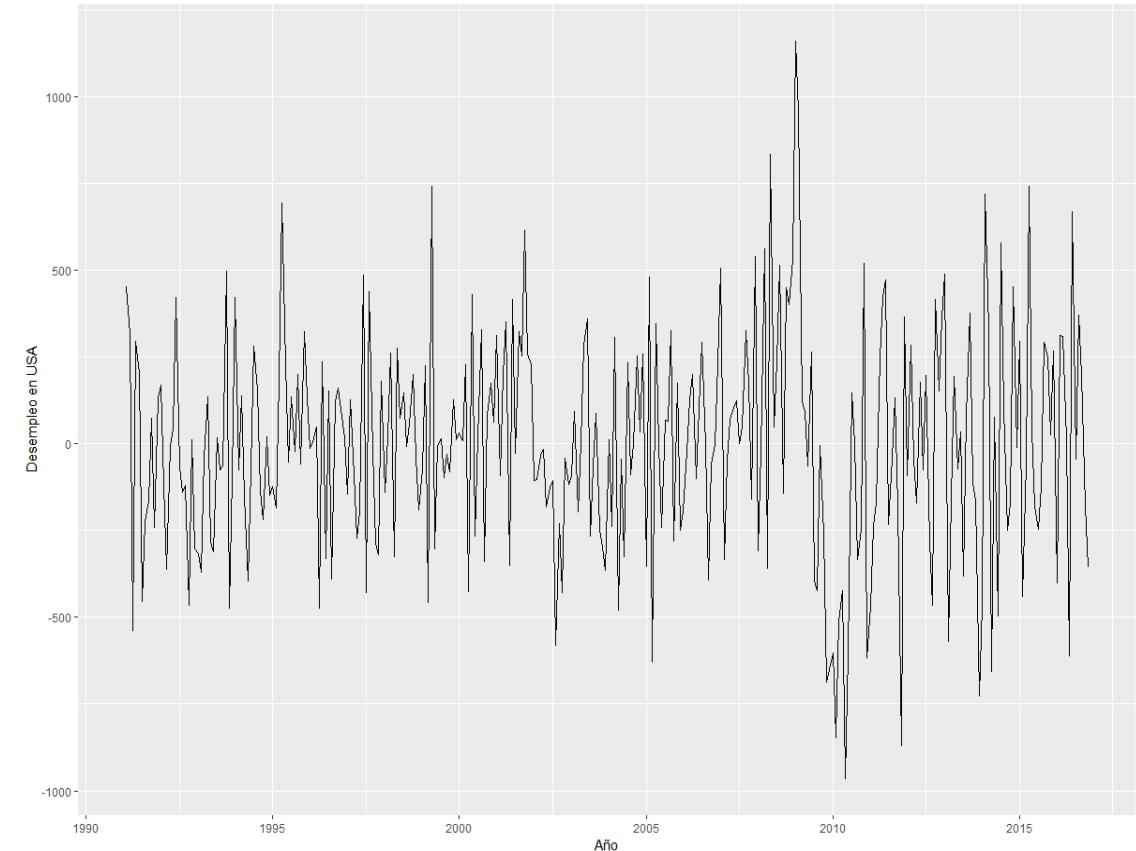
# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Al obtener la serie estacionaria  $x_t$  de  $y_t$ , y habiendo identificado el proceso tanto en la parte no estacional como estacional, se procede a estimar, para la serie original,  $y_t$ , el proceso que gobierna la serie. Veamos tres ejemplos, empezando por la serie unemp, referente al desempleo en USA.

Serie original  $y_t$

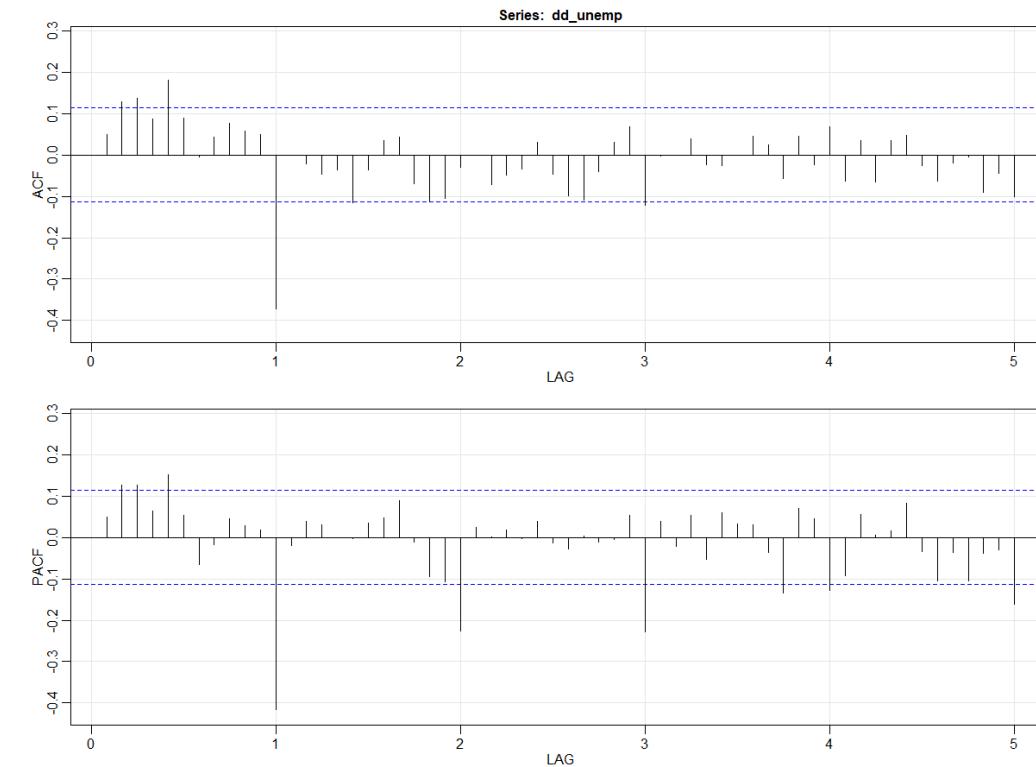


Serie estacionaria  $x_t$

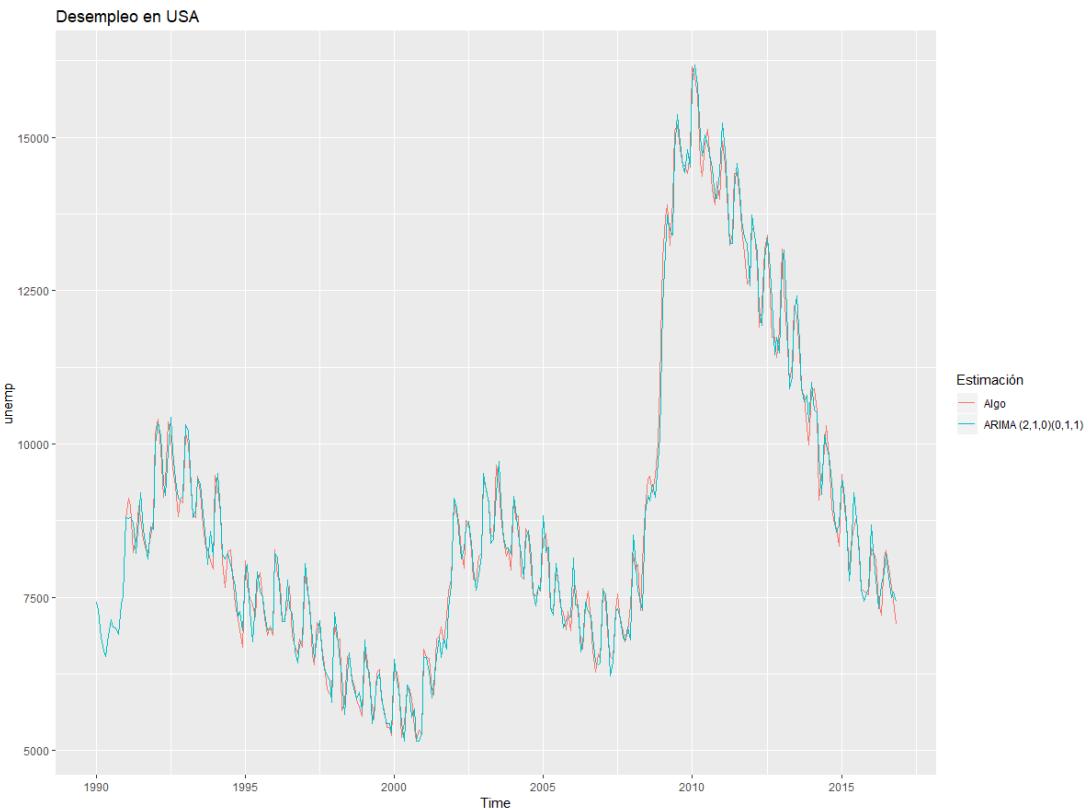


# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Correlograma



$ARIMA(2,1,0)(0,1,1)_{12}$

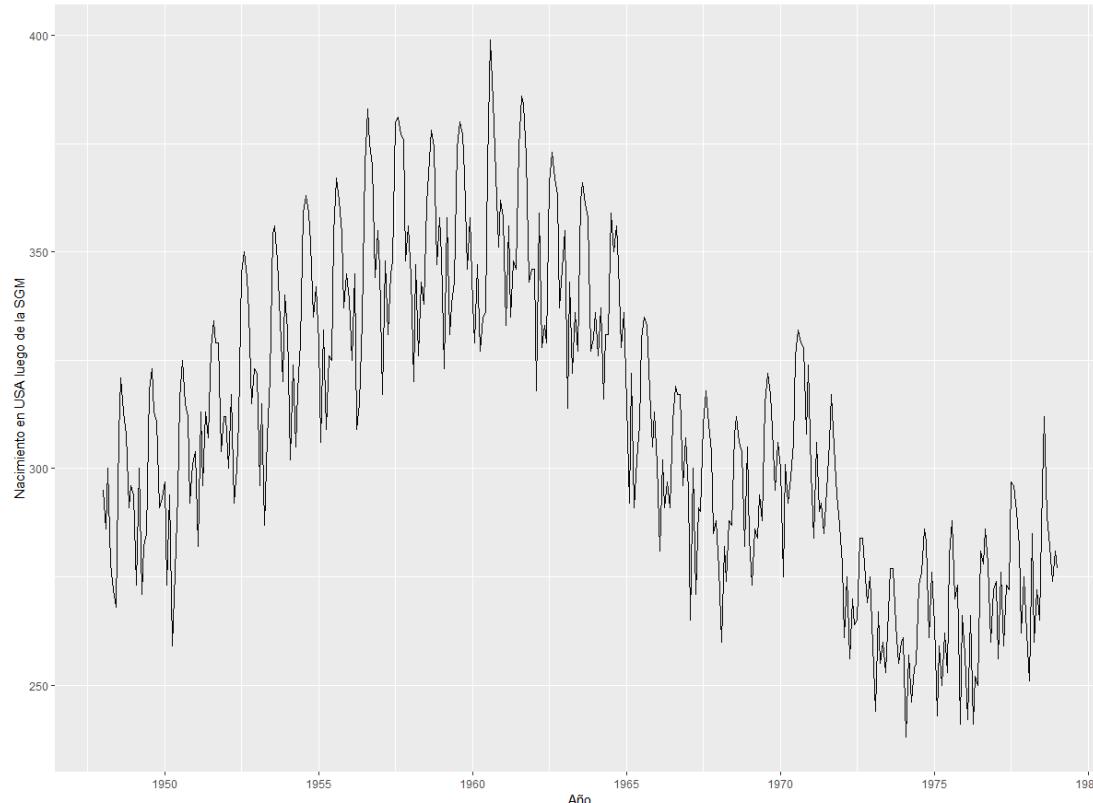


Para la presente serie se estimó un  $ARIMA(2,1,0)(0,1,1)_{12}$  y los resultados parecen ser satisfactorios. Es importante adicionar parámetros, y comparar los otros modelos con medidas de rendimiento, además de corroborar los diagnósticos.

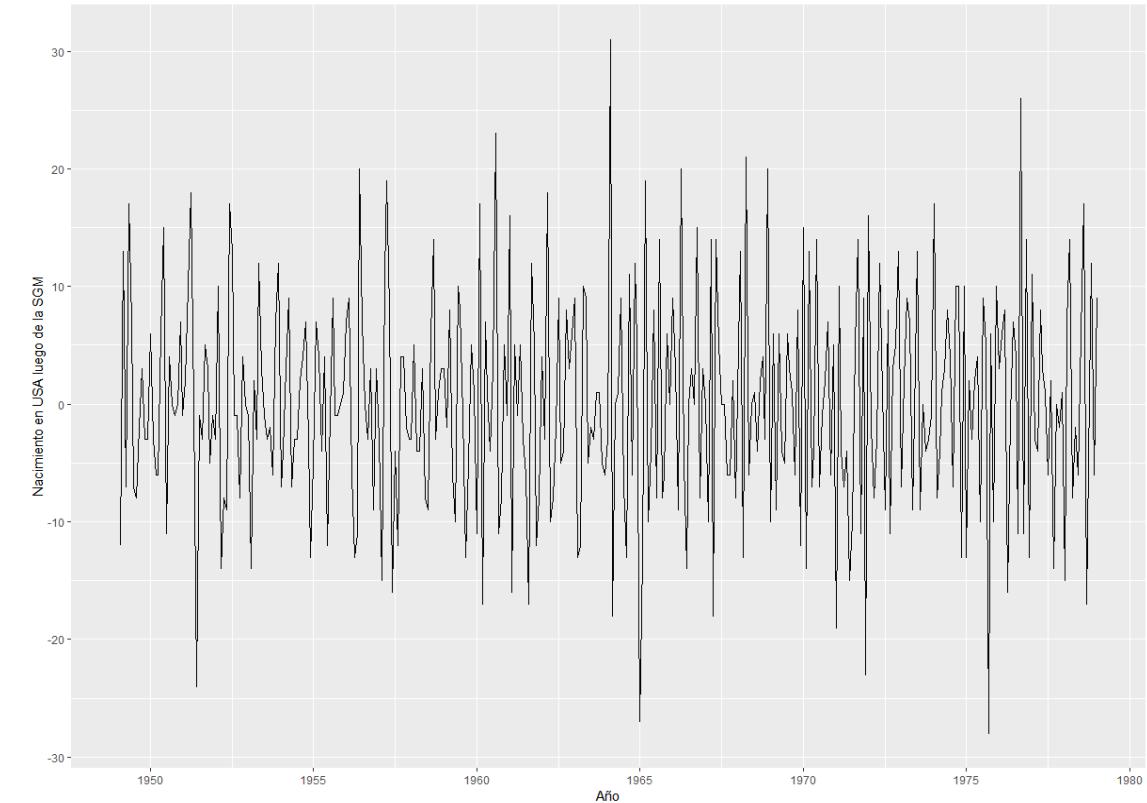
# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Veamos el siguiente ejemplo sobre los nacimientos en USA después de la segunda Guerra Mundial.

Serie original  $y_t$

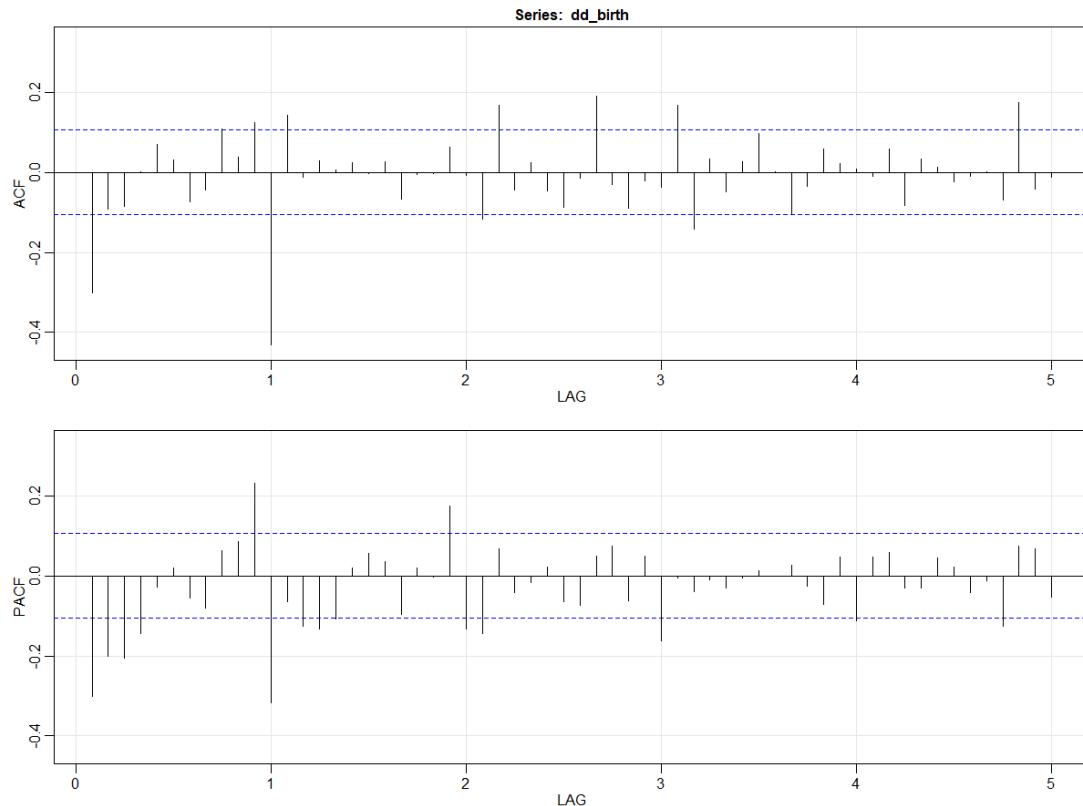


Serie estacionaria  $x_t$

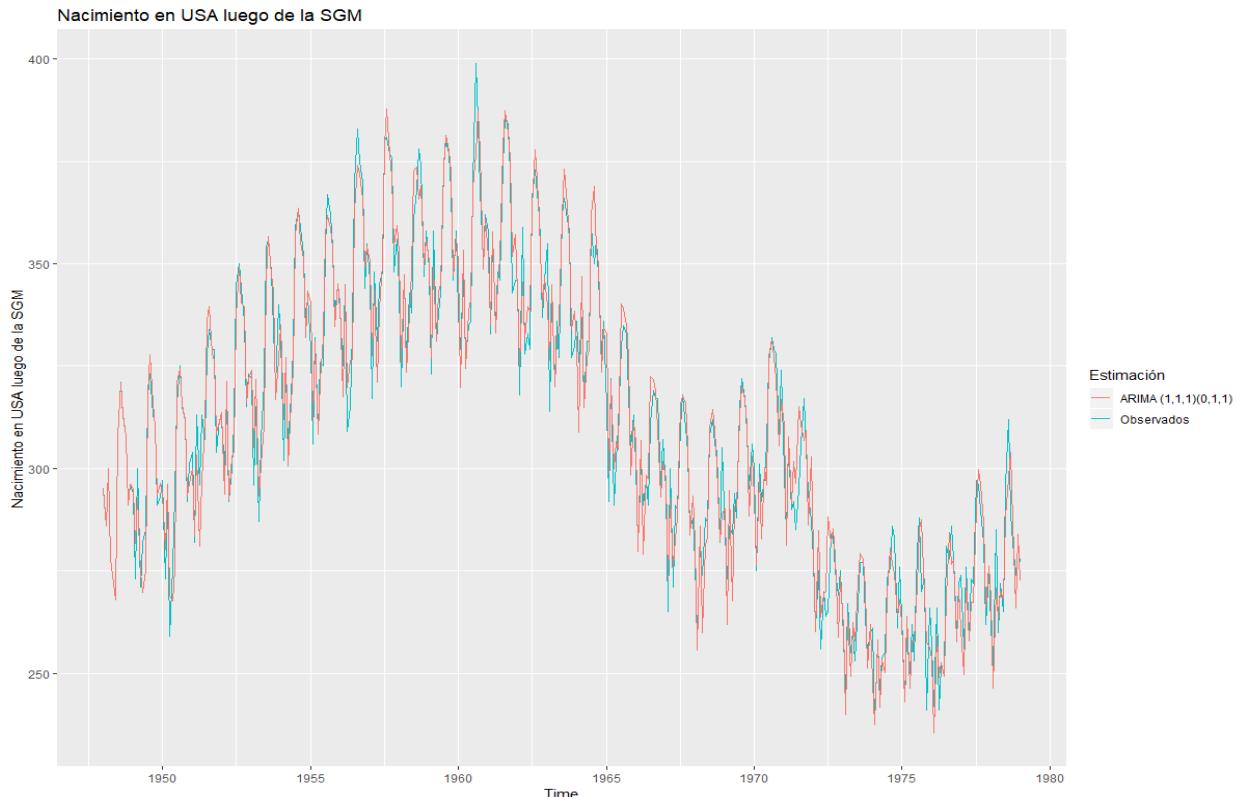


# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Correlograma



$ARIMA(1,1,1)(0,1,1)_{12}$

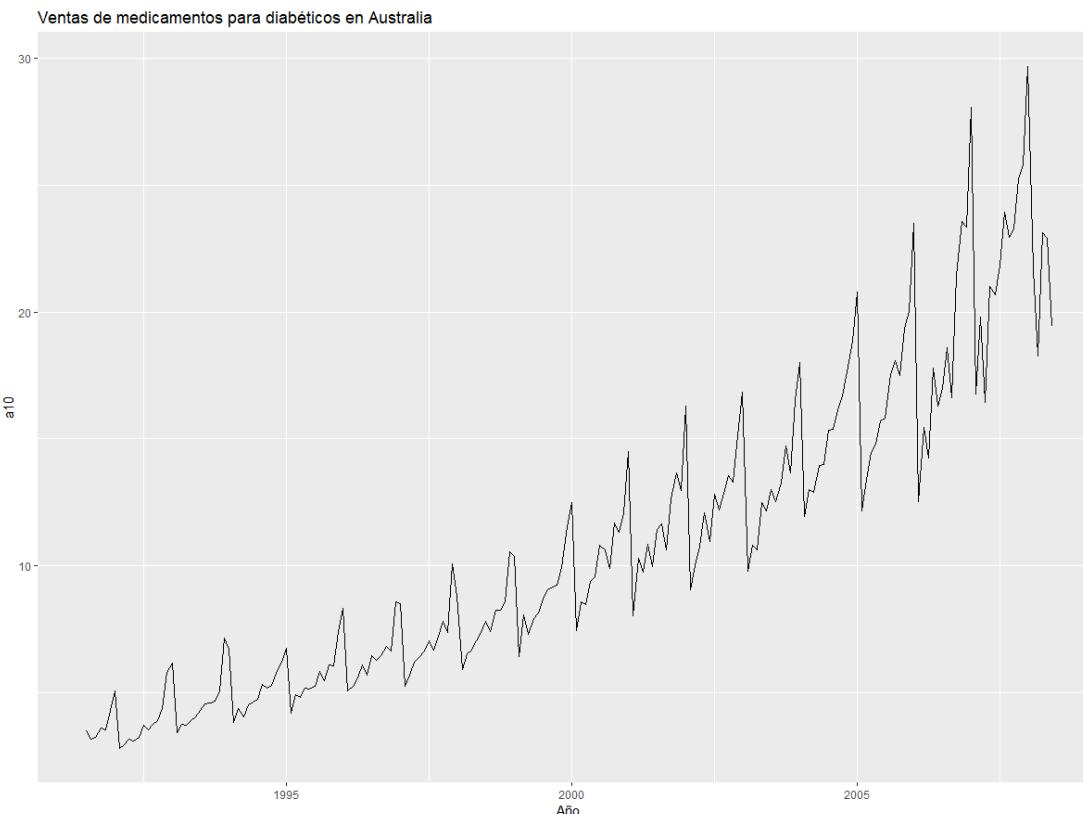


Para la presente serie se estimó un  $ARIMA(1,1,1)(0,1,1)_{12}$  y los resultados parecen ser satisfactorios. Es importante adicionar parámetros, y comparar los otros modelos con medidas de rendimiento, además de corroborar los diagnósticos.

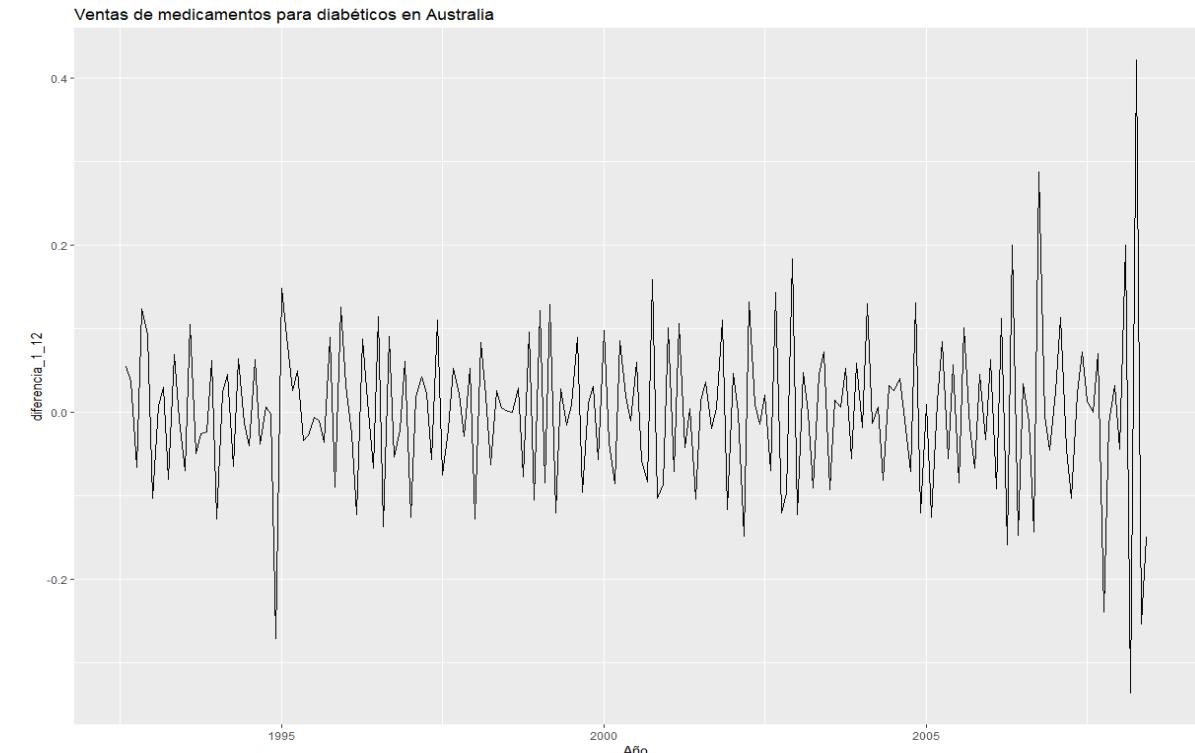
# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Finalmente, volvamos al ejemplo de las ventas de medicamentos en Australia. Este es el único caso que no es muy convincente...

Serie original  $y_t$

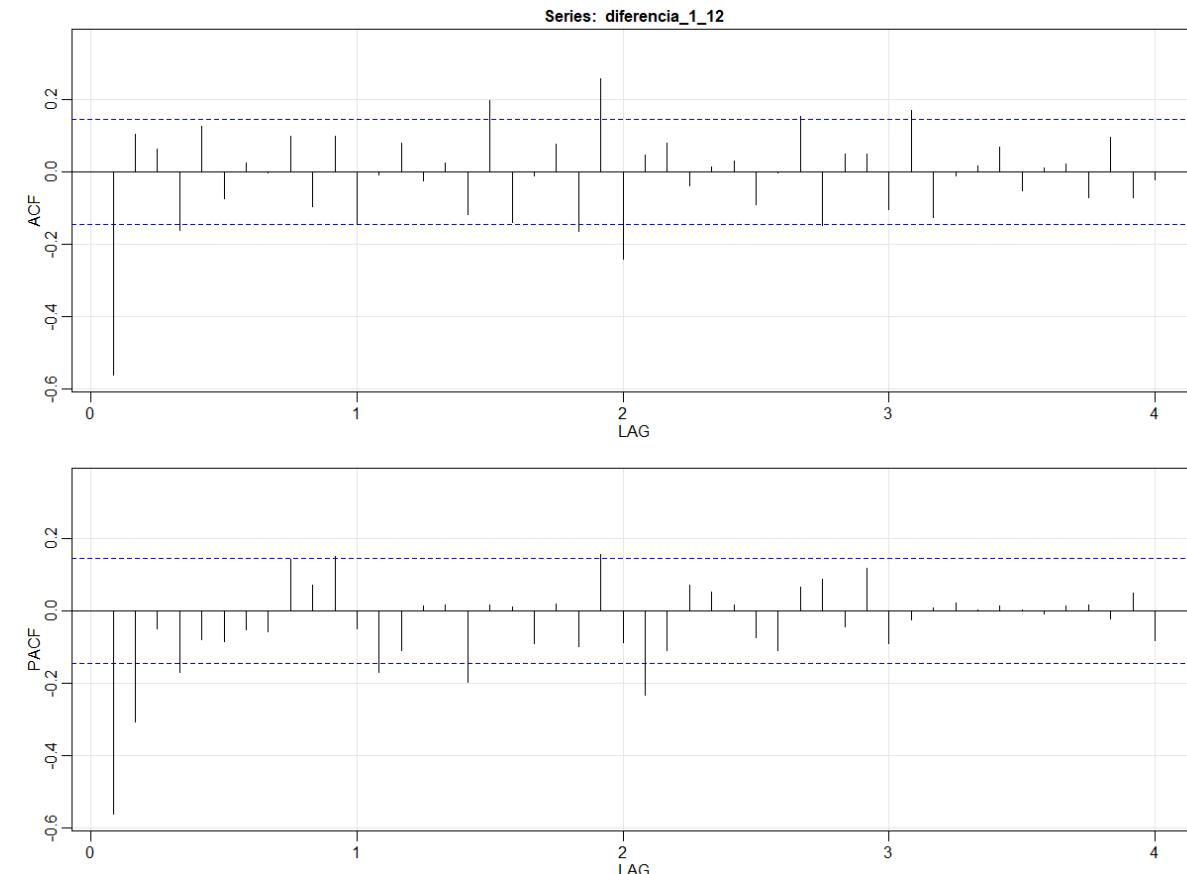


Serie estacionaria  $x_t$

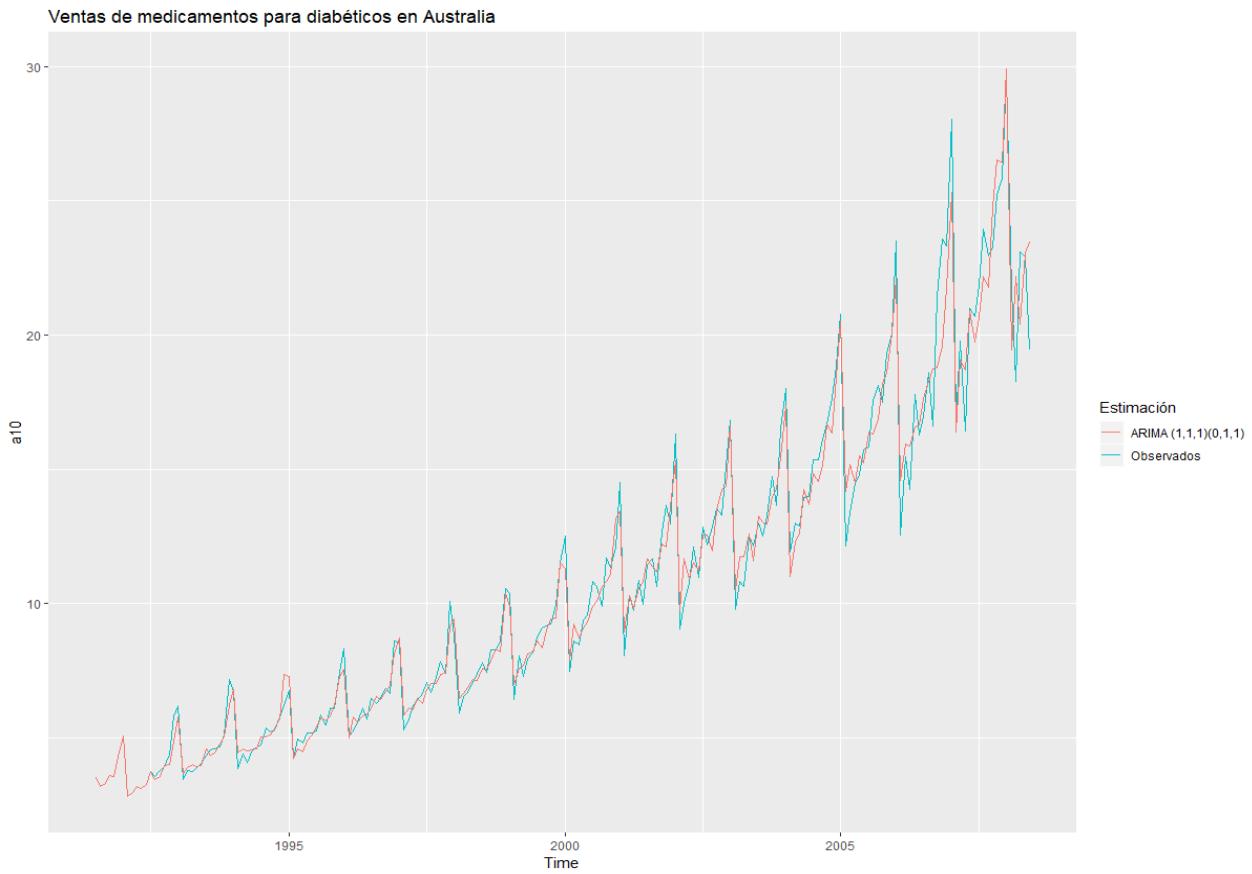


# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_s$

Correlograma



$ARIMA(1,1,1)(0,1,1)_{12}$



Para la presente serie se estimó un  $ARIMA(1,1,1)(0,1,1)_{12}$  y los resultados NO son muy satisfactorios. Es importante adicionar parámetros, y comparar los otros modelos con medidas de rendimiento, además de corroborar los diagnósticos... en este caso no se cumplen diversos supuestos...

# Estimación de los modelos $ARIMA(p, d, q)(P, D, Q)_S$

No hace falta algo antes de pasar al pronóstico, y es verificar o hacer un análisis más profundo de los residuos para saber si cumplen con la propiedad de ser un ruido blanco. A esto lo solemos denominar como la prueba de los diagnósticos.



# Índice

1

Fundamentos del método  
Box-Jenkins

4

Diagnósticos / bondad y  
ajuste del modelo ARIMA

2

La estacionaridad y las  
pruebas de raíz unitaria

3

Identificación y  
estimación de un proceso  
ARIMA

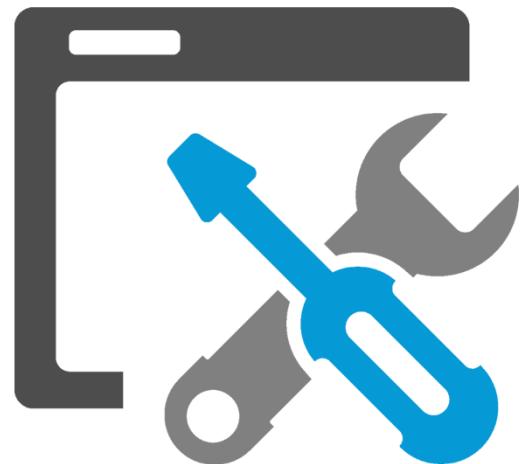
# Diagnósticos: ¿son los residuos adecuados?

El análisis de los diagnósticos nos permiten testificar 4 cosas:

1. Patrones en los residuos estandarizados.
2. Residuos de las ACF.
3. Normalidad en los residuos
4. El correlación o no de los residuos, lo cual indica que se está omitiendo un término ARIMA al modelo, además del cumplimiento de la no correlación entre los residuos (ruido blanco).



Si bien es cierto que los  $ARIMA(p, d, q)(P, D, Q)_{12}$  son modelos paramétricos que se ajustan tanto a la propiedad de la estacionariedad como a la propiedad de ruido blanco en los residuos, el no cumplimiento de algún supuesto no se traduce por el no tener que llevar a cabo el pronóstico bajo un determinado modelo. Es posible que en la búsqueda de algún  $ARIMA(p, d, q)(P, D, Q)_{12}$  no se llegue a encontrar un proceso que cumpla con los supuesto.



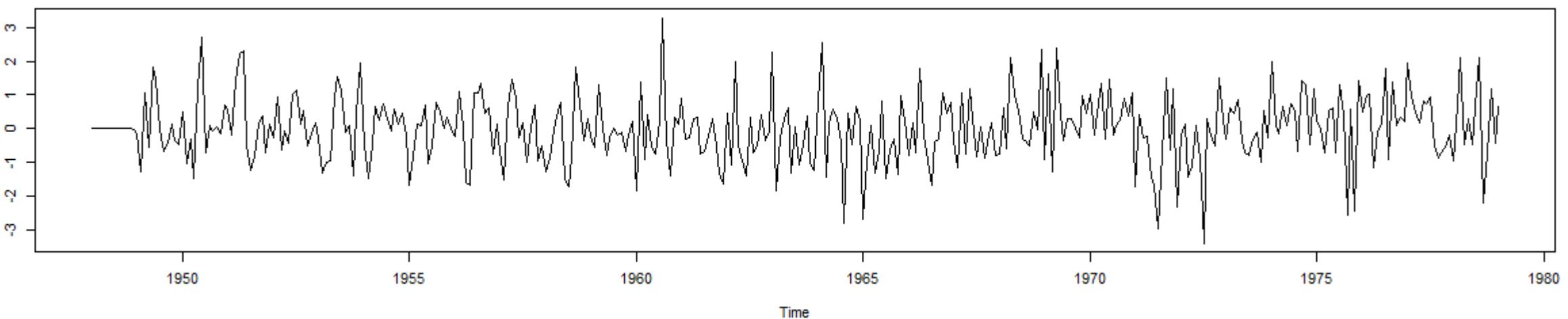
A título personal, considero más importante el análisis o la validez interna del modelo, y escoger aquel modelo que posea los criterios o estadísticas de rendimiento más pequeñas.

# Diagnósticos: ¿son los residuos adecuados?

1. Patrones en los residuos: los residuos estandarizados es una forma de analizar si la serie estimada posee errores que se comportan como ruidos blancos. De hecho, se busca que los residuos sigan una secuencia de ruido blanco con media cero y varianza uno. La fórmula esta dada por:

$$\frac{r_i}{\sqrt{\frac{1}{T-1} \sum_{i=1}^T r_i^2}}$$

Un ejemplo podría ser:



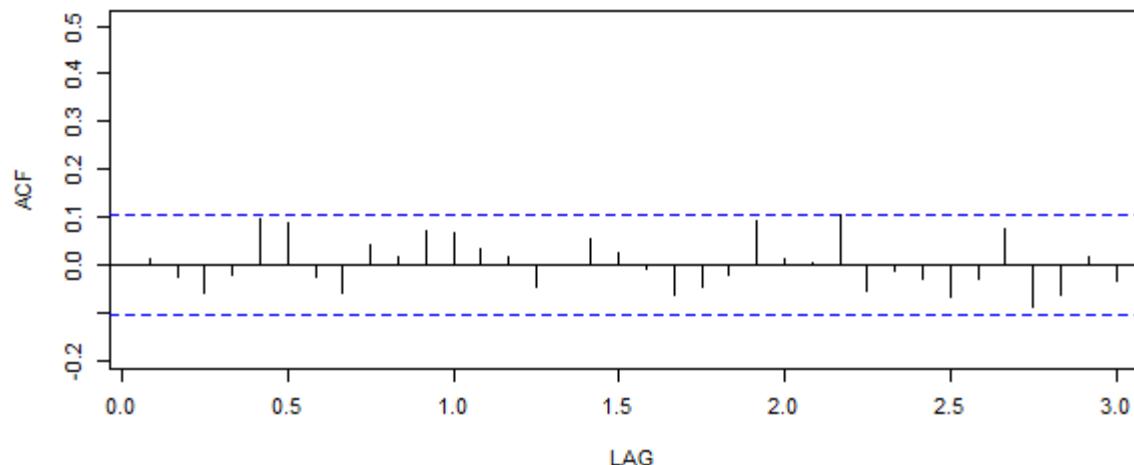
# Diagnósticos: ¿son los residuos adecuados?

2. Residuos de las ACF: otra forma de verificar el cumplimiento es al inspeccionar si los ACF residuales se comportan como un ruido blanco. Este gráfico, como el anterior, también nos puede testificar si hay algún cambio estructural o valor de influencia particular. Los residuales poseen como estimación puntal :

$$\pm \frac{ACF_i - \frac{\alpha}{2}}{\sqrt{n}} \quad \text{e intervalos dado por}$$

$$\pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{n} \left( 1 + 2 \sum_{i=1}^k \rho^{i^2} \right)}$$

Un ejemplo podría ser:

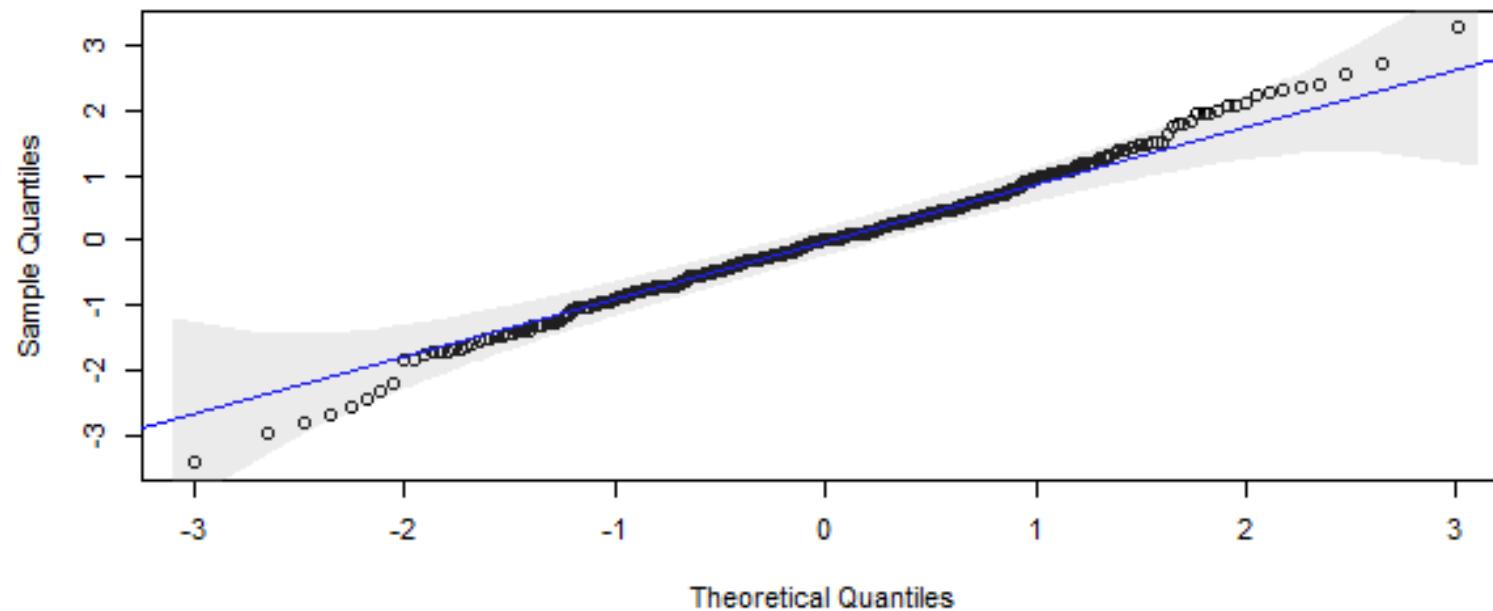


# Diagnósticos: ¿son los residuos adecuados?

3. Gráfico QQ: al supuesto de que el ruido blanco se distribuye como una distribución normal, con media y variancia constante, verificamos lo anterior con un gráfico QQ. Su ecuación resulta ser:

$$\frac{k - a}{(T + 1 - 2a)} \quad \text{para } a \text{ entre } 0 \text{ y } \frac{1}{2}$$

Un ejemplo podría ser:



# Diagnósticos: ¿son los residuos adecuados?

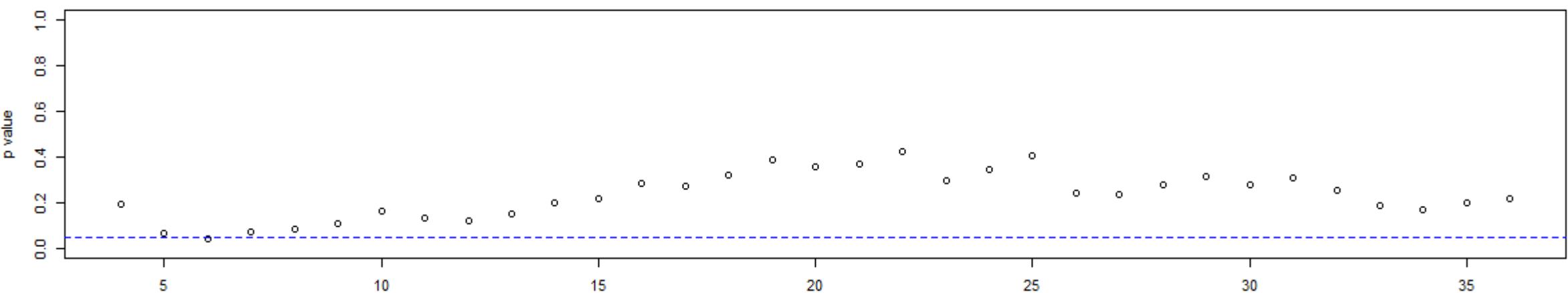
4. Ljung-Box-Pierce: la prueba verifica si los residuos poseen o no correlaciones. Si no hay correlaciones, el proceso es un ruido blanco, de lo contrario, la no estacionaridad podría ser que se esté omitiendo algún componente o ponderador (ARIMA), y se debe replantear el proceso como tal. La estadística es la siguiente:

$$Q = T(T + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k}$$

$T$  es el tamaño de la serie

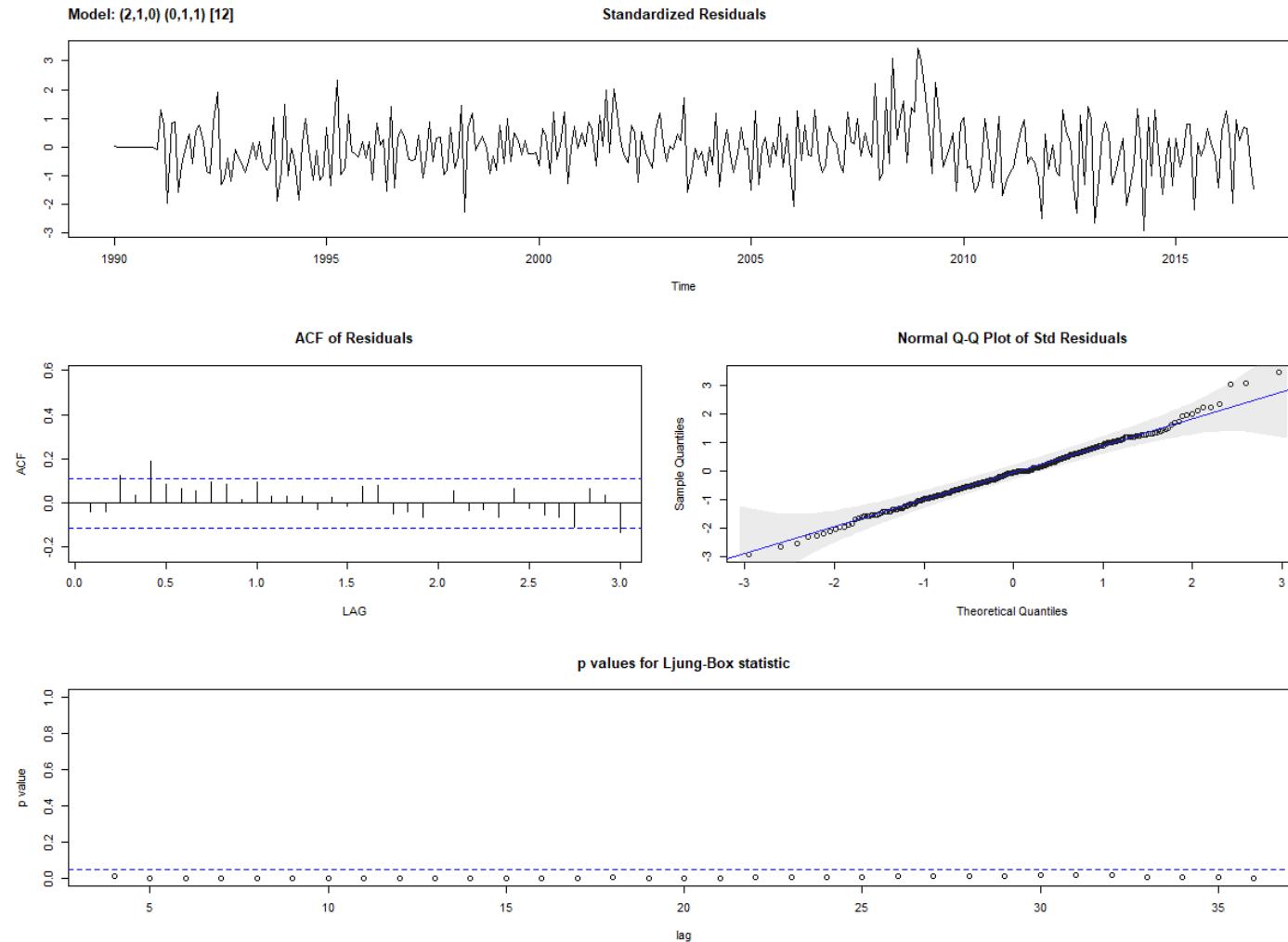
$\hat{\rho}_k$  es la autocorrelación de la muestra en el retraso  $k$

$h$  es el número de retrasos que está probando. Un ejemplo podría ser:



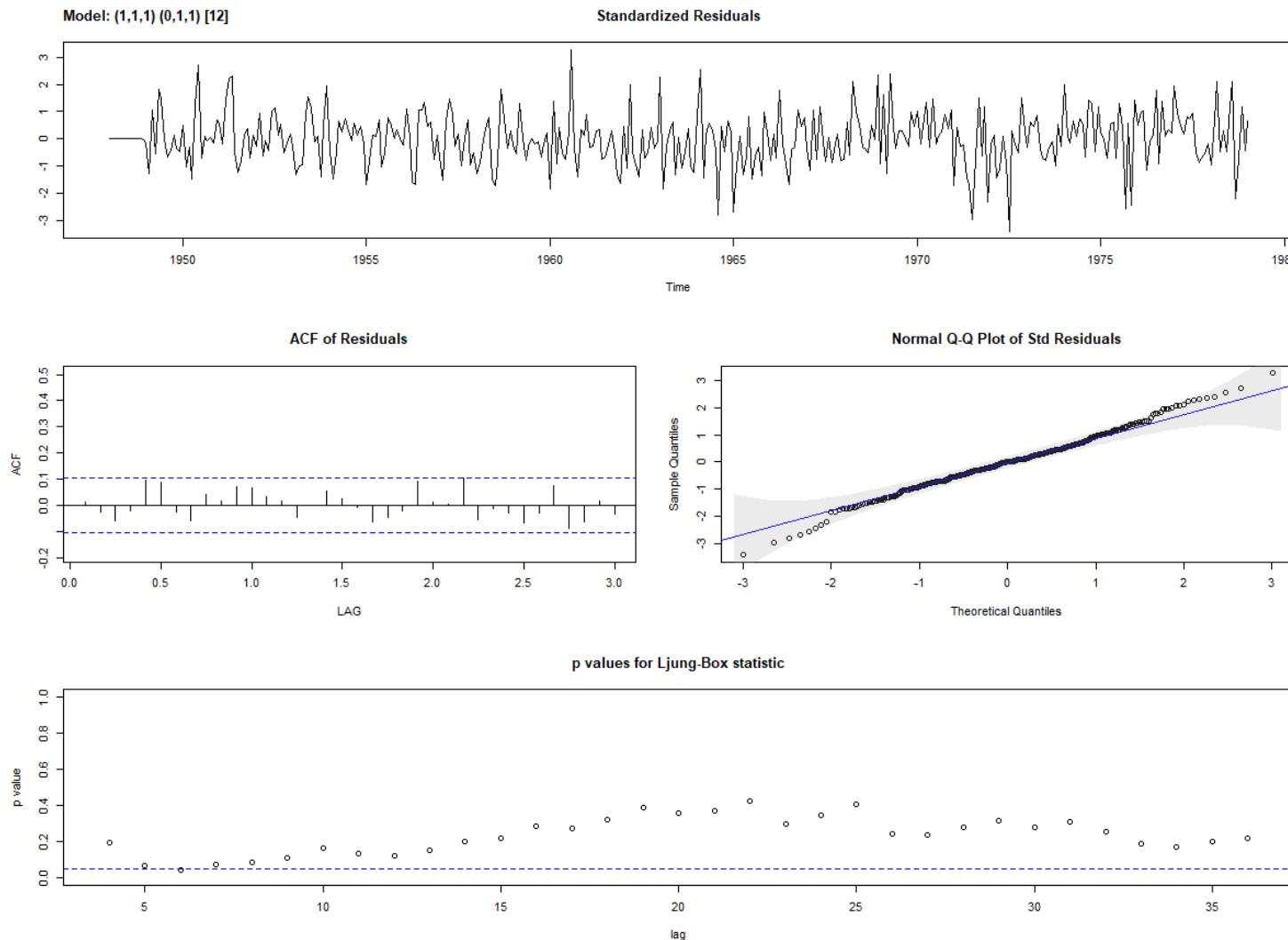
# Diagnósticos: ¿son los residuos adecuados?

Veamos los diagnósticos para el ejemplo del modelo estimado de *unem*. ¿Qué podemos decir?



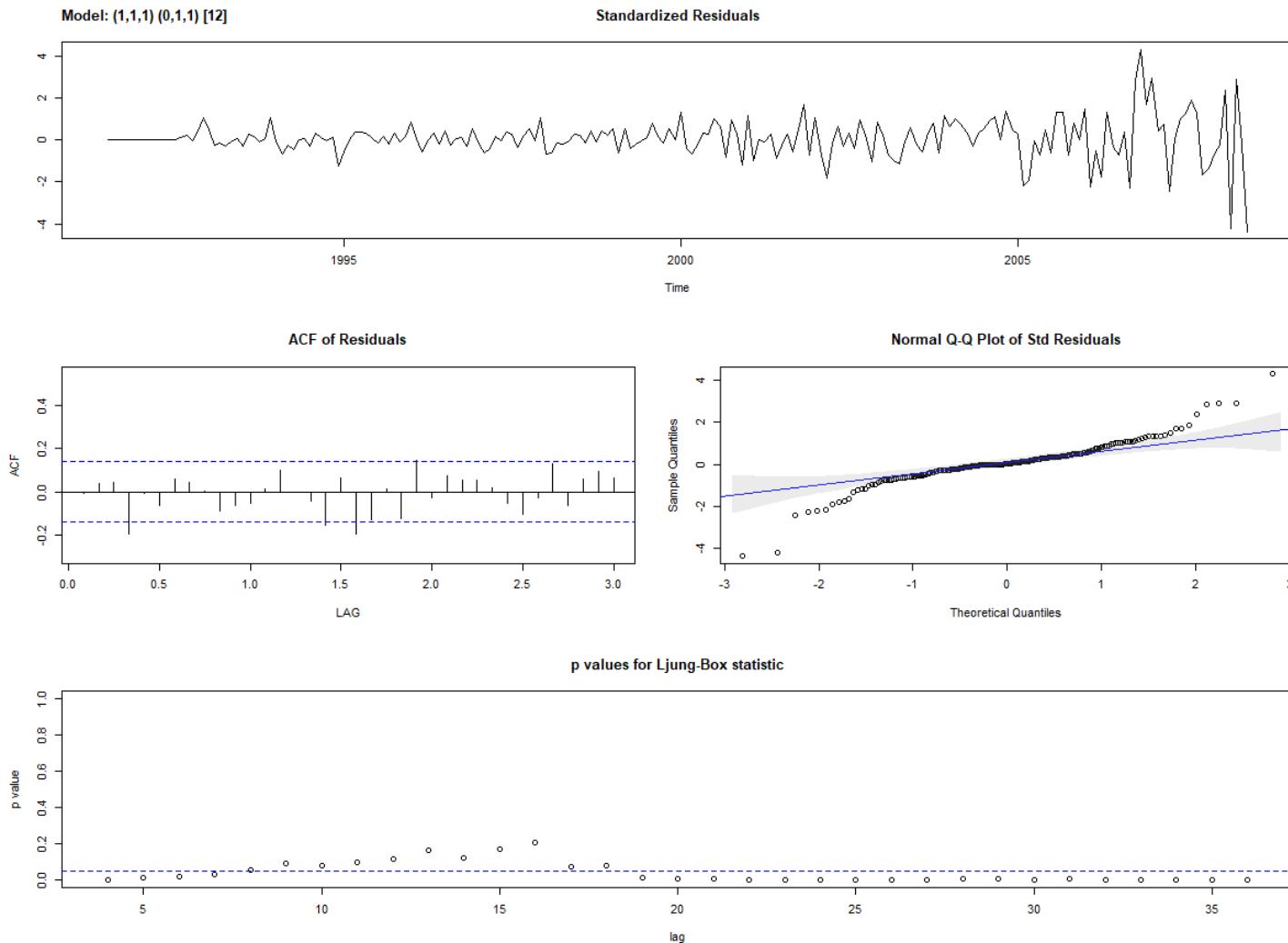
# Diagnósticos: ¿son los residuos adecuados?

Veamos los diagnósticos para el ejemplo del modelo estimado de *birth*. ¿Qué podemos decir?



# Diagnósticos: ¿son los residuos adecuados?

Veamos los diagnósticos para el ejemplo del modelo estimado de *a10*. ¿Qué podemos decir?



# Índice

1

Fundamentos del método  
Box-Jenkins

4

Diagnósticos / bondad y  
ajuste del modelo ARIMA

2

La estacionaridad y las  
pruebas de raíz unitaria

5

Pronósticos del  
 $ARIMA(p, d, q)(P, D, Q)$

3

Identificación y  
estimación de un proceso  
ARIMA

# Pronósticos del $ARIMA(p, d, q)(P, D, Q)_s$

## Estimación puntual

Los pronósticos puntuales se pueden calcular usando el concepto del valor esperado condicional a las observaciones pasadas como para los modelos exponenciales:  $\hat{y}_{T+h|T}$ . Por ejemplo en la obtención de  $y_{T+1}$ , esta se puede obtener, para un proceso ARMA(3,1):

$$y_t = (1 + \hat{\phi}_1)y_{t-1} - (\hat{\phi}_1 - \hat{\phi}_2)y_{t-2} - (\hat{\phi}_2 - \hat{\phi}_3)y_{t-3} - \hat{\phi}_3y_{t-4} + u_t + \hat{\theta}_1u_{t-1}$$

$$\hat{y}_{T+1|T} = (1 + \hat{\phi}_1)y_T - (\hat{\phi}_1 - \hat{\phi}_2)y_{T-1} - (\hat{\phi}_2 - \hat{\phi}_3)y_{T-2} - \hat{\phi}_3y_{T-3} + \hat{\theta}_1u_t$$

$$\hat{y}_{T+2|T} = (1 + \hat{\phi}_1)y_{T+1|T} - (\hat{\phi}_1 - \hat{\phi}_2)y_T - (\hat{\phi}_2 - \hat{\phi}_3)y_{T-1} - \hat{\phi}_3y_{T-2} + \hat{\theta}_1u_t \dots$$

¿Por qué en última ecuación no poseemos términos de perturbación?

El proceso continúa de esta manera para todos los períodos de tiempo futuros. De esta manera, se puede obtener cualquier número de pronósticos de puntos para encontrar residuos.

# Pronósticos del $ARIMA(p, d, q)(P, D, Q)_s$

## Estimación por intervalos

El cálculo de los intervalos de predicción de ARIMA es algo más complejo. Sabemos que por la ecuación de Wold:

$y_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$ , a partir de esto, la estimación de variabilidad en el tiempo está dada por:

$$\hat{\sigma}_h = \hat{\sigma}^2 \left[ 1 + \sum_{i=1}^{h-1} \hat{\theta}_i^2 \right], \quad \text{para } h = 2, 3, \dots$$

El intervalo de predicción está dado por  $\hat{y}_{T+h|T} \pm 1.96\sqrt{\hat{\sigma}_h}$

Los intervalos de predicción para los modelos ARIMA se basan en suposiciones de que los residuos no están correlacionados y están normalmente se distribuyen. Si alguna de estas suposiciones no se cumple, entonces los intervalos de predicción pueden ser incorrectos. Por este motivo, siempre hay que analizar el ACF y el histograma de los residuos para verificar los supuestos antes de producir intervalos de predicción. En general, los intervalos de predicción de los modelos ARIMA aumentan a medida que aumenta el horizonte de pronóstico. Para los modelos estacionarios, convergerán, de modo que los intervalos de predicción para horizontes largos son esencialmente los mismos. Para el caso  $d > 1$ , los intervalos de predicción continuarán creciendo en el futuro. Al igual que con la mayoría de los cálculos de intervalo de predicción, los intervalos basados en ARIMA tienden a ser demasiado estrechos. Esto ocurre porque solo se ha tenido en cuenta la variación en los errores. También existe una variación en las estimaciones de los parámetros y en el orden del modelo, que no se ha incluido en el cálculo. Finalmente, el cálculo supone que los patrones históricos que se han modelado continuarán en el período de pronóstico.

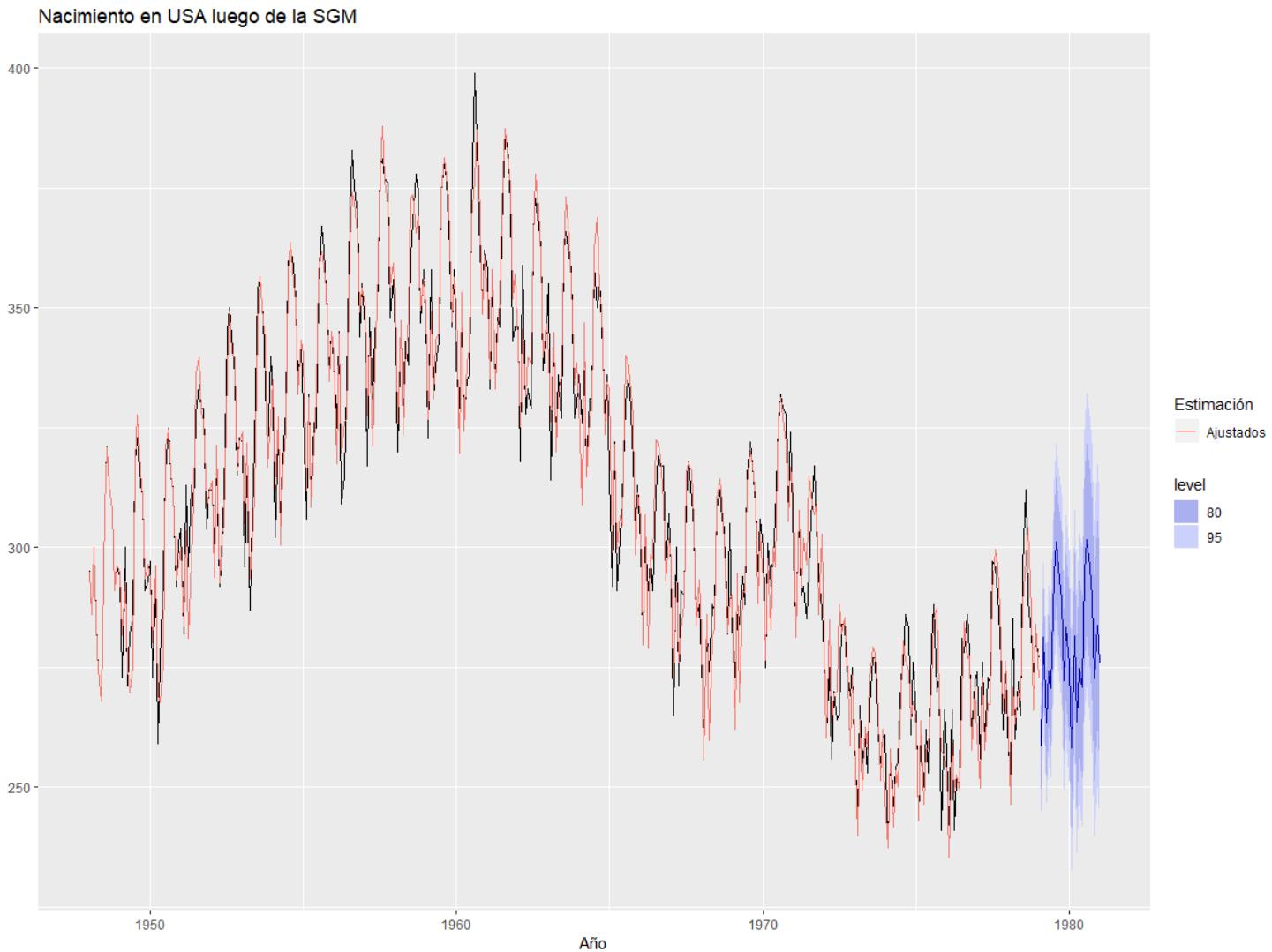
# Pronósticos del $ARIMA(p, d, q)(P, D, Q)_s$

Veamos los pronósticos para el ejemplo del modelo estimado de ***unem***. ¿Qué podemos decir?



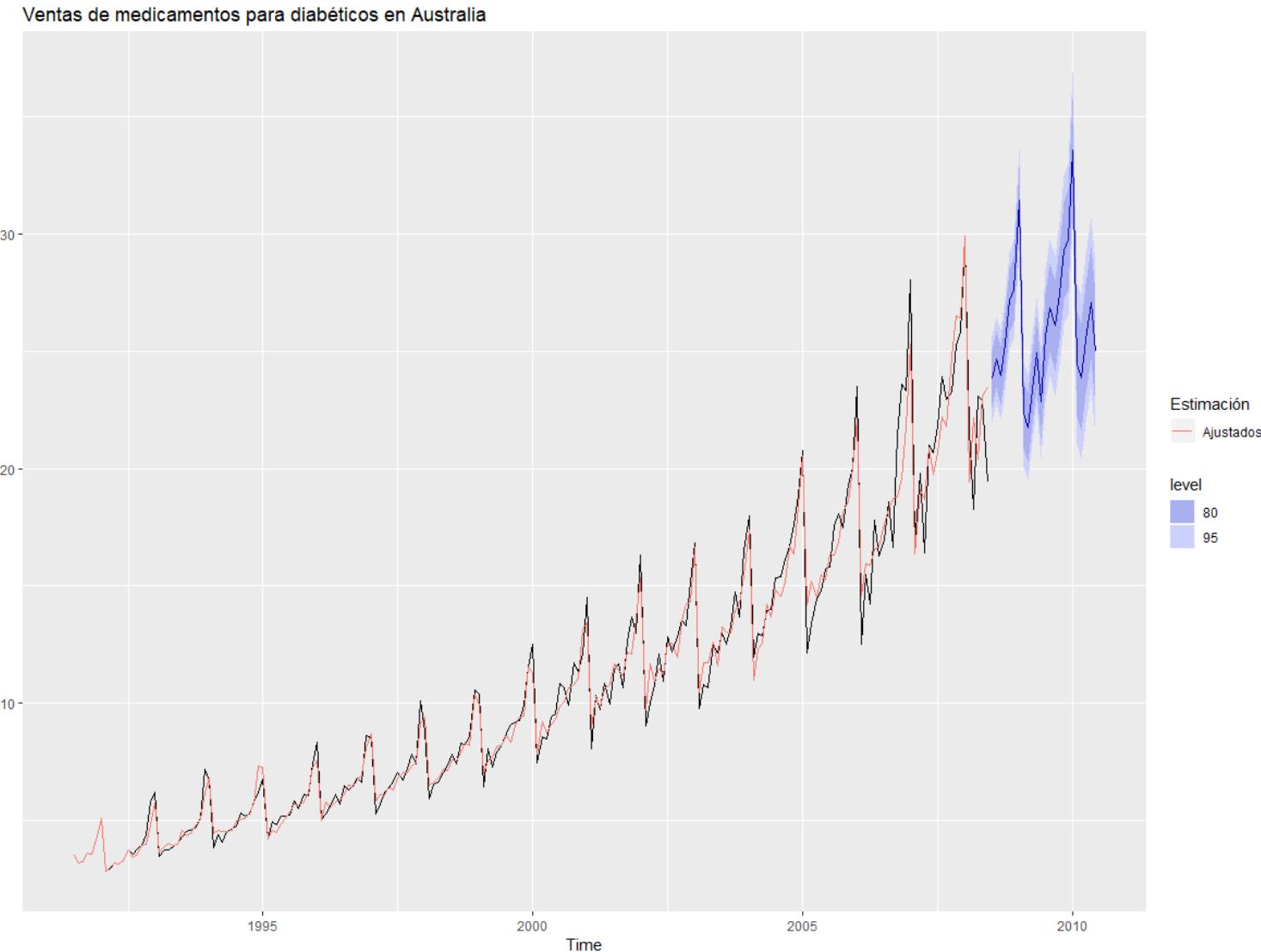
# Pronósticos del $ARIMA(p, d, q)(P, D, Q)_s$

Veamos los pronósticos para el ejemplo del modelo estimado de ***birth***. ¿Qué podemos decir?



# Pronósticos del $ARIMA(p, d, q)(P, D, Q)_s$

Veamos los pronósticos para el ejemplo del modelo estimado de **a10**. ¿Qué podemos decir?



# Índice

1

Fundamentos del método  
Box-Jenkins

4

Diagnósticos / bondad y  
ajuste del modelo ARIMA

2

La estacionaridad y las  
pruebas de raíz unitaria

5

Pronósticos del  
 $ARIMA(p, d, q)(P, D, Q)$

3

Identificación y  
estimación de un proceso  
ARIMA

6

¿ `auto.arima()` ?

# La función auto.arima()

La función `auto.arima()` en R utiliza una variación del algoritmo Hyndman-Khandakar (Hyndman & Khandakar, 2008), que combina pruebas de raíz unitaria, minimización del AICc y de la MLE para obtener un modelo ARIMA. Los argumentos del `auto.arima()` proporcionan muchas variaciones en el algoritmo: lo que se describe aquí es el comportamiento predeterminado bajo ciertas restricciones en la función.

El procedimiento predeterminado utiliza algunas aproximaciones para acelerar la búsqueda... Estas aproximaciones se pueden evitar con el argumento `approximation=FALSE`. Es posible que el modelo AICc mínimo no se encuentre debido a estas aproximaciones, o debido al uso de un procedimiento stepwise.

Se buscará un conjunto mucho más grande de modelos si se usa el argumento `stepwise = FALSE`.

En lo particular, considero que la función `auto.arima()` no siempre brinda los mejores resultados a causa de la forma de encontrar el proceso, pero si podría llegar a ser una buena guía.

## Hyndman-Khandakar algorithm for automatic ARIMA modelling

1. The number of differences  $0 \leq d \leq 2$  is determined using repeated KPSS tests.
2. The values of  $p$  and  $q$  are then chosen by minimising the AICc after differencing the data  $d$  times.  
Rather than considering every possible combination of  $p$  and  $q$ , the algorithm uses a stepwise search to traverse the model space.
  - a. Four initial models are fitted:
    - ARIMA(0,  $d$ , 0),
    - ARIMA(2,  $d$ , 2),
    - ARIMA(1,  $d$ , 0),
    - ARIMA(0,  $d$ , 1).A constant is included unless  $d = 2$ . If  $d \leq 1$ , an additional model is also fitted:
    - ARIMA(0,  $d$ , 0) without a constant.
  - b. The best model (with the smallest AICc value) fitted in step (a) is set to be the “current model”.
  - c. Variations on the current model are considered:
    - vary  $p$  and/or  $q$  from the current model by  $\pm 1$ ;
    - include/exclude  $c$  from the current model.The best model considered so far (either the current model or one of these variations) becomes the new current model.
  - d. Repeat Step 2(c) until no lower AICc can be found.

# La función auto.arima()

Para la serie **unem**, comparemos los resultados del  $ARIMA(2,1,0)(0,1,1)_{12}$  con el auto.arima. ¿Qué podemos decir?

```
Series: unemp
ARIMA(2,1,0)(0,1,1)[12]

Coefficients:
      ar1     ar2     sma1
    0.0946  0.2081 -0.8512
  s.e.  0.0556  0.0556  0.0423

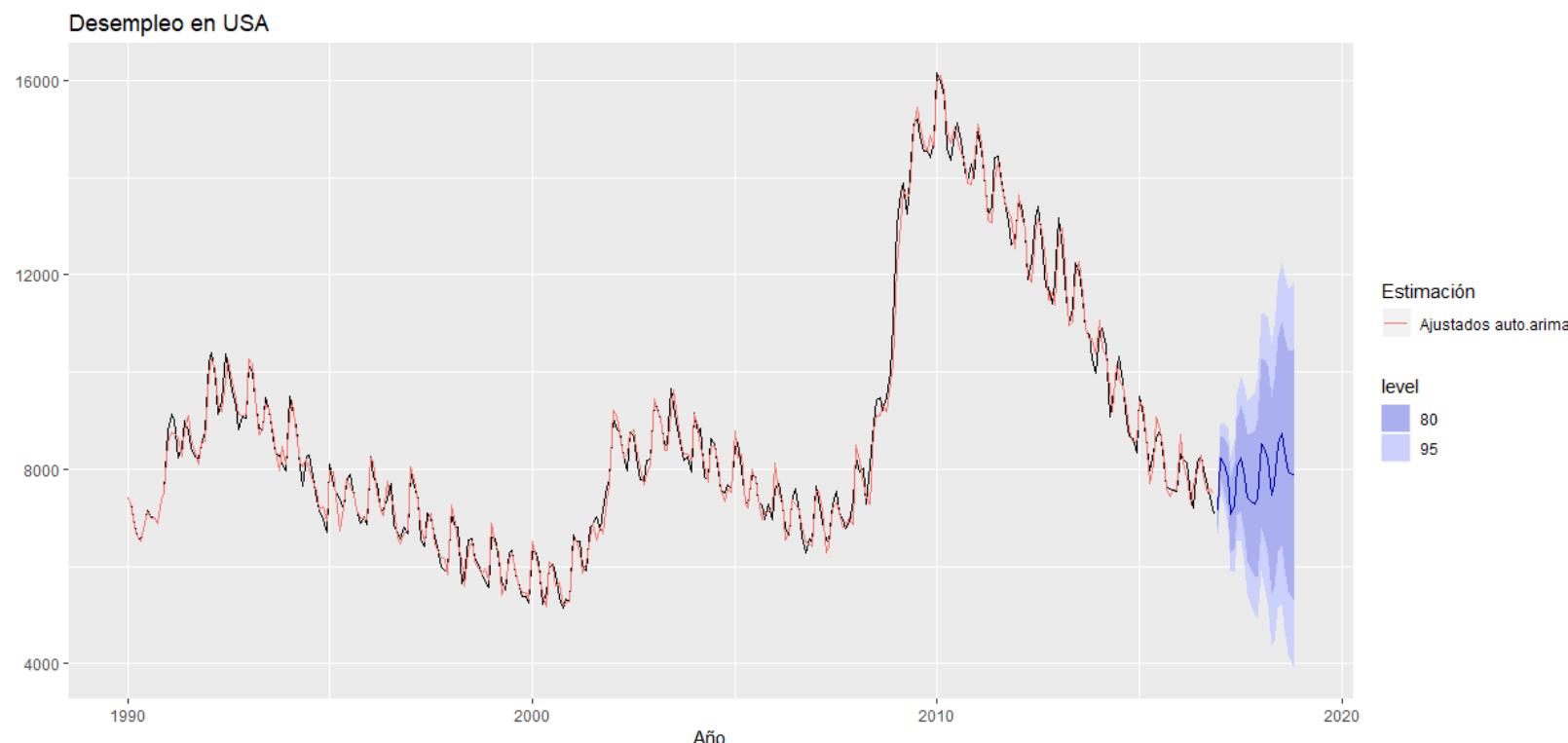
sigma^2 estimated as 66176:  log likelihood=-2166.66
AIC=4341.31   AICc=4341.44   BIC=4356.26
```

```
Series: unemp
ARIMA(2,0,2)(1,1,2)[12]

Coefficients:
      ar1     ar2     ma1     ma2     sar1     sma1     sma2
    1.9434 -0.9465 -0.9730  0.1391  0.0602 -0.8595 -0.0001
  s.e.  0.0290  0.0288  0.0647  0.0597  0.5469  0.5447  0.4561

sigma^2 estimated as 60422:  log likelihood=-2158.67
AIC=4333.34   AICc=4333.82   BIC=4363.26
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA	-15.12	250.79	194.75	-0.17	2.27	0.19	-0.04
auto.arima	0.06	0.90	0.58	0.03	4.90	0.44	0.00



# La función auto.arima()

Para la serie ***birth***, comparemos los resultados del ARIMA  $ARIMA(1,1,1)(0,1,1)_{12}$ , con el auto.arima. ¿Qué podemos decir?

```
Series: birth
ARIMA(1,1,1)(0,1,1)[12]

Coefficients:
      ar1      ma1      sma1
      0.3038 -0.7006 -0.8000
  s.e.  0.0865  0.0604  0.0441

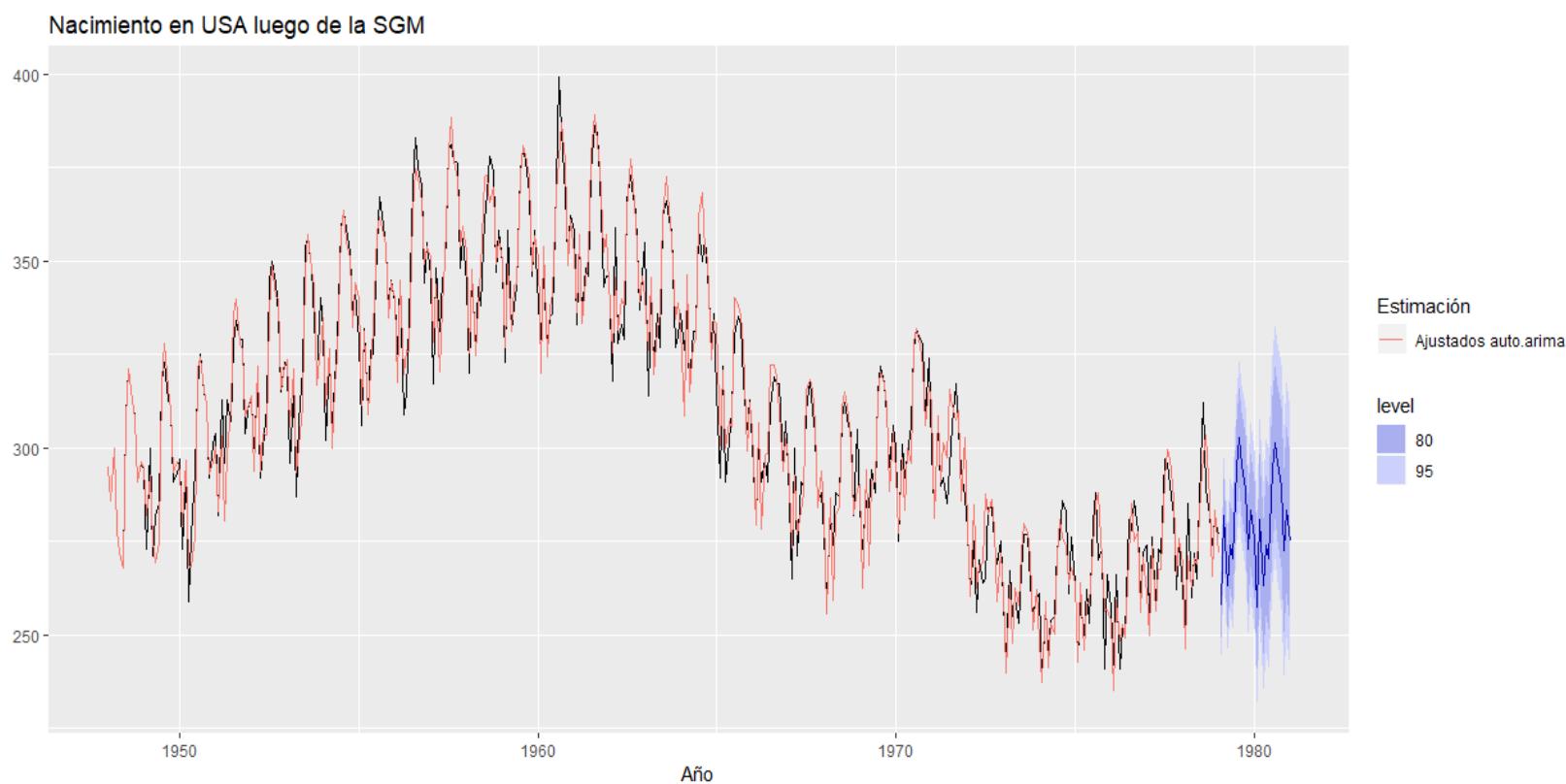
sigma^2 estimated as 46.3:  log likelihood=-1205.93
AIC=2419.85  AICc=2419.97  BIC=2435.4
```

```
Series: birth
ARIMA(1,1,1)(0,1,1)[12]

Coefficients:
      ar1      ma1      sma1
      0.3038 -0.7006 -0.8000
  s.e.  0.0865  0.0604  0.0441

sigma^2 estimated as 46.3:  log likelihood=-1205.93
AIC=2419.85  AICc=2419.97  BIC=2435.4
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA	-0.05	6.66	5.08	-0.02	1.66	0.52	0.01
auto.arima	-0.10	6.63	5.06	-0.03	1.66	0.52	0.01



# La función auto.arima()

Para la serie **a10**, comparemos los resultados del  $ARIMA(1,1,1)(0,1,1)_{12}$ , con el `auto.arima`. ¿Qué podemos decir?

```
Series: a10
ARIMA(1,1,1)(0,1,1)[12]

Coefficients:
      ar1      ma1      smal
     -0.2504  -0.6674  -0.4725
  s.e.  0.1007   0.0870   0.0641

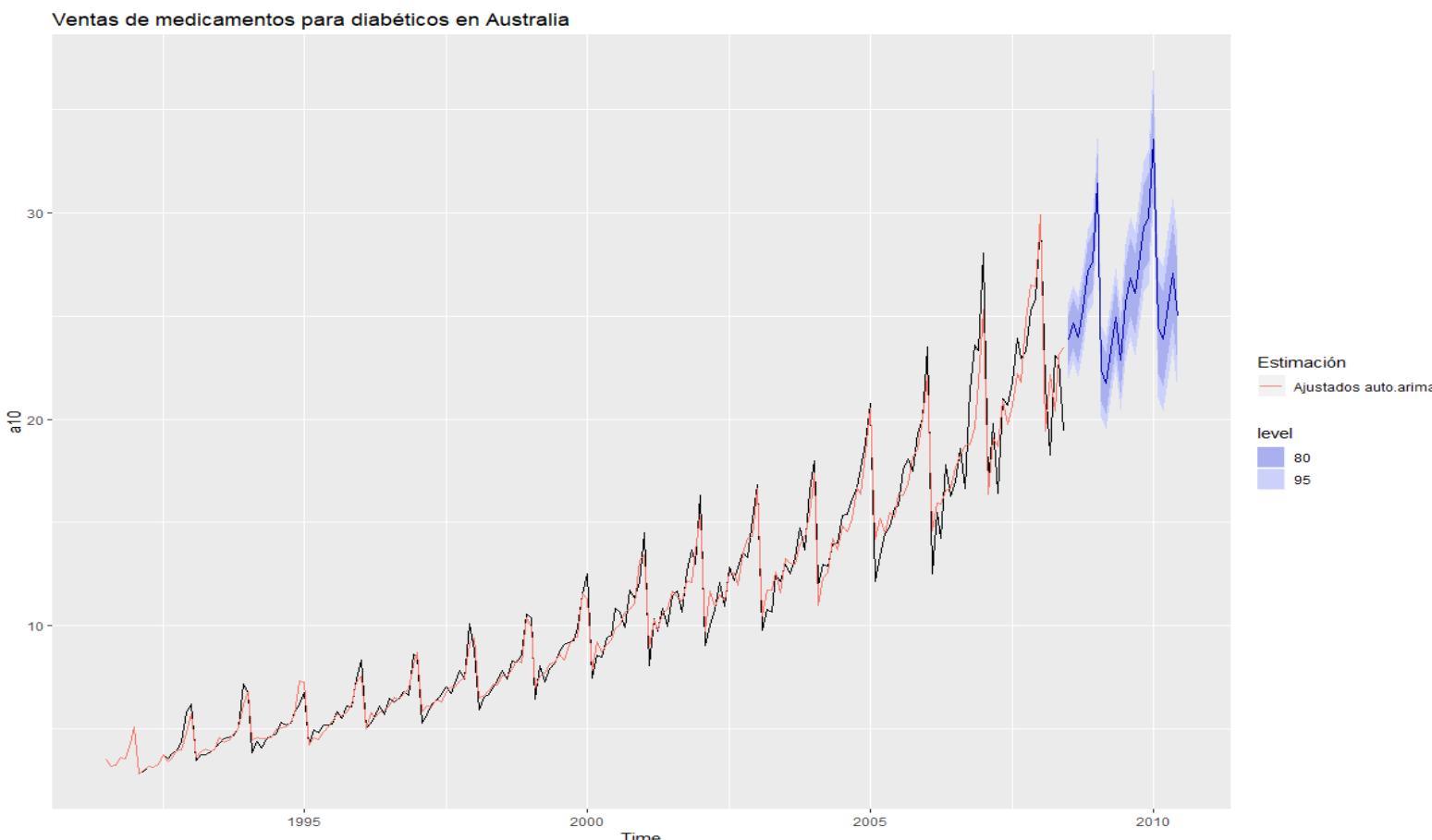
sigma^2 estimated as 0.8756:  log likelihood=-258.82
AIC=525.63  AICc=525.85  BIC=538.64
```

```
Series: a10
ARIMA(1,1,1)(0,1,1)[12]

Coefficients:
      ar1      ma1      smal
     -0.2504  -0.6674  -0.4725
  s.e.  0.1007   0.0870   0.0641

sigma^2 estimated as 0.8756:  log likelihood=-258.82
AIC=525.63  AICc=525.85  BIC=538.64
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA	0.06	0.9	0.58	0.03	4.9	0.44	0
auto.arima	0.06	0.9	0.58	0.03	4.9	0.44	0



# Índice

7

ARIMA vs ETS

# La relación de los modelos ETS y los ARIMA

Es un mito común de las personas que suelen saber de series temporales cuando dicen que los modelos ARIMA son más generales que el suavizado exponencial, y que entonces los ARIMA pueden llegar a ser peores que el de suavizamiento exponencial.

En TODOS los casos los modelos de suavizado exponencial lineal son casos especiales de los modelos ARIMA (un subconjunto de los modelos ARIMA); los modelos de suavizado exponencial *no lineal* no tienen equivalentes al ARIMA (estos no fueron cubiertos en el presente curso). Por otro lado, también hay muchos modelos ARIMA que no tienen contrapartes de suavizado exponencial (por ejemplo, no estudiamos modelos ARIMA no lineales, bajo ciertas restricciones en los estimadores y errores, etc). En particular, todos los modelos ETS no son estacionarios, mientras que algunos modelos ARIMA son estacionarios. Aunque teóricamente podemos llegar a encontrar relaciones, en las variantes de ambos casos no se posee relación alguna.

Recordar que ambos son sumas ponderadas, y de ahí que se puede llegar a obtener cierta igualdad. Sin embargo, para largos periodos o memorias, los  $ARIMA(p, d, q)(P, D, Q)_s$  suelen ser más poderosos.

Una de las mayores diferencias es que los modelos ETS con tendencia estacional o sin amortiguamiento, ambos poseen raíces unitarias (es decir, necesitan dos niveles de diferenciación para hacerlos estacionarios...). Todos los demás modelos de ETS tienen una raíz unitaria (necesitan un nivel de diferenciación para hacerlos estacionarios).

La tabla de la siguiente diapositiva muestra ciertas relaciones teóricas entre los modelos ETS y los ARIMA.

# La relación de los modelos ETS y los ARIMA

Table 8.3: Equivalence relationships between ETS and ARIMA models.

ETS model	ARIMA model	Parameters
ETS(A,N,N)	ARIMA(0,1,1)	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)	$\theta_1 = \alpha + \beta - 2$ $\theta_2 = 1 - \alpha$
ETS(A,A <sub>d</sub> ,N)	ARIMA(1,1,2)	$\phi_1 = \phi$ $\theta_1 = \alpha + \phi\beta - 1 - \phi$ $\theta_2 = (1 - \alpha)\phi$
ETS(A,N,A)	ARIMA(0,1,m)(0,1,0) <sub>m</sub>	
ETS(A,A,A)	ARIMA(0,1,m + 1)(0,1,0) <sub>m</sub>	
ETS(A,A <sub>d</sub> ,A)	ARIMA(0,1,m + 1)(0,1,0) <sub>m</sub>	

Para más información sobre la comparación entre los modelos exponenciales y los ARIMA:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.3756&rep=rep1&type=pdf>

<https://www.scss.tcd.ie/~dahyotr/ST3010/RzDTimeSeriesForecasting.pdf>

[http://web.vu.lt/mif/a.buteikis/wp-content/uploads/2018/02/Lecture\\_03.pdf](http://web.vu.lt/mif/a.buteikis/wp-content/uploads/2018/02/Lecture_03.pdf)

<https://robjhyndman.com/uwafiles/fpp-notes.pdf>

# Índice

7

ARIMA vs ETS

8

Resumen del Box-Jenkins

# Resumen del método Box-Jenkins

Finalmente, el método de Box-Jenkins se puede resumir de la siguiente forma (no olvidar particionar la serie en un conjunto de entrenamiento y otro de validación.... por favor... ):

1. Analice la serie original  $y_t$ . Determine cuales podrían los factores no estacionarias.
2. Pase de una serie no estacionaria  $y_t$  a una serie estacionaria  $x_t$
3. Realice la o las posibles identificaciones del proceso
4. Estime las o las series.
5. Adicione términos o parámetros a la serie
6. Conjuntamente con la estimación y adición de parámetros, realice los diagnósticos. Recuerde que el concepto de parsimonia es importante.
7. Haga un cuadro resumen con las medidas de rendimiento de los modelos. Escoja el mejor modelo ARIMA.
8. Realice el pronostico con los  $h$  periodos que quiere predecir a futuro.
9. No olvidar analizar si los pronósticos siguen el mismo patrón de la serie, y verificar si el modelo por ende puede ser correcto.



# Conclusión

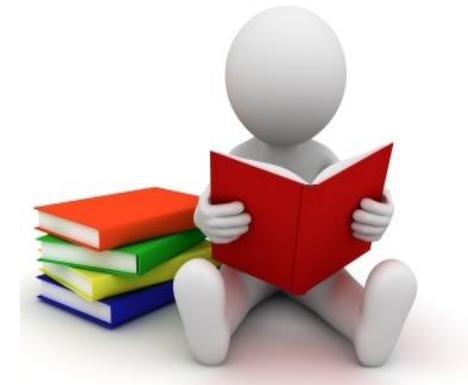
El capítulo presentó la metodología de Box-Jenkins en la estimación de los modelos  $ARIMA(p, d, q)(P, D, Q)_s$ .

Se estudiaron los fundamentos de los métodos ARIMA: procesos generadores, principio de estacionaridad, ruido blanco, la ecuación de Wold, las autocorrelaciones, y el dendograma.

Los modelos  $ARIMA(p, d, q)(P, D, Q)_s$  provienen tanto de la explicación de componentes autoregresivos en las observaciones, como factores de medias móviles para los residuos según la memoria del proceso. Se debe de tener en cuenta tanto la identificación no estacional como estacional.

El análisis de una serie temporal pasa por las etapas de estacionar la seria, llevar a cabo la identificación, estimación, corroboración de los supuestos en los residuos, elección del mejor modelo, y finalmente el pronóstico en el horizonte  $h$ .

En la próxima sesión se estudiaran los modelos ARIMA que poseen cambios en su estructura (intervención), y ciertas soluciones en la presencia de valores extremos.



The  
End